Austrian Contributions to the

XXIV General Assembly of the International Union of Geodesy and Geophysics

July 2 – 13, 2007, Perugia, Italy

TECHNISCHE UNIVERSITÄT WIEN

VIENNA UNIVERSITY OF TECHNOLOGY

An der **Fakultät für Mathematik und Geo-information** der Technischen Universität Wien ist am **Institut für Geodäsie und Geo-physik** eine unbefristete Stelle für eine/n

Universitätsprofessor/in für Ingenieurgeodäsie

ab 1.10.2008 zu besetzen.

Nähere Informationen finden Sie unter **http://info.tuwien.ac.at/dekzent/ Ausschreibungen.html**

Österreichische Zeitschrift für Vermessung & Geoinformation

Special Issue
XXIV General Assembly of the International Union of Geodesy and Geophysics (IUGG), Perugia, Italy

ed. by Fritz K. Brunner, Heribert Kahmen and Harald Schuh
(All papers of this issue were peer-reviewed.)

ÖSTERREICHISCHE GEODÄTISCHE KOMMISSION

ÖGK

Editorial of Special Issue

The Austrian Geodetic Commission (ÖGK) is the national organisation in Austria related to the International Association of Geodesy (IAG) which in turn, is one of the eight associations of the International Union of Geodesy and Geophysics (IUGG).

Every four years IUGG and its associations hold a General Assembly. The IAG uses these general assemblies to present the work performed, to review its own structure, to appoint new officers and to organise a suite of symposia. The XXIV General Assembly of IUGG will take place in Perugia, Italy, July 2 to July 13, 2007. It has been a tradition to expect that member countries of IAG prepare a National Report about the geodetic activities during the past four years for presentation at the general assemblies. However, the ÖGK has decided to prepare a collection of geodetic research papers by Austrian institutions and scientists on the occasion of the forthcoming General Assembly instead of providing a National Report.

The main task of this special issue of the *Österreichische Zeitschrift für Vermessung & Geoinformation (VGI)* is twofold: (i) to give the international geodetic community an idea about the various research topics to which Austrian geodesists are contributing currently, and (ii) to show to those Austrian colleagues, who are usually involved in practical work, the new developments Austrian scientists are involved in. The range of research topics indicates how closely geodesy is related to its neighbouring disciplines such as geophysics, metrology, meteorology, space science, and computer science. Joint efforts in different fields are needed to make contributions in solving problems currently faced by modern society, such as precise navigation on Earth or understanding of the – 'System Earth' – global changes.

The authors of this Special Issue of the VGI represent the two Austrian universities, which offer the full programme in geodesy, i.e., the Vienna University of Technology and the Graz University of Technology, the Federal Office of Metrology and Surveying (BEV), the Department of Meteorology und Geophysics of the Vienna University, and the Space Research Institute (IWF) of the Austrian Academy of Sciences. The following 19 papers highlight major parts of the current geodetic research in Austria. The sequence of the papers in this Special Issue reflects the structure of the commissions of IAG.

On behalf of the ÖGK and all authors, the three guest-editors would like to thank the Austrian Society of Geodesy and Geoinformation (OVG) for providing this platform. In particular we are very grateful to the editor-in-chief of VGI, Stefan Klotz, and his two deputies, Ernst Zahn and Andreas Pammer, for taking care of the layout of the papers and solving all technical problems, often at short notice.

Finally, it should be mentioned that all papers of this issue were peer-reviewed to ensure a high quality of the published articles. The efforts of the reviewers are gratefully acknowledged.

Fritz K. Brunner, Heribert Kahmen, Harald Schuh
(Guest-editors of this Special Issue)

Automated quasi-realtime prediction of GNSS clock corrections

Veronika Broederbauer, Michael Opitz and Robert Weber

Abstract

For positioning and time transfer applications with GNSS knowledge about satellite specific orbits and clock-corrections is required. The ACs (Analysis Centers) of the IGS (International GNSS Service) provide satellite ephemeris as well as clock corrections to GPST (GPS-Time) in sp3-format for free via internet with a latency of about one day. These orbit and clock information result from a least squares adjustment of the parameters based on observations from the IGS reference station network.

The accurate and reliable prediction of satellite clocks and orbits is an indispensable condition of all GNSS based positioning applications in real-time. While the orbits are output to an integration of the well-known force field the clock corrections to GPST (GPS-Time) have to be extrapolated by means of an experienced prediction model. The model used for predicting GPS and GLONASS satellite clocks within program GNSS-VC (GNSS-Vienna Clocks) contains basically the coefficients of a quadratic polynomial as well as an amplitude and a phase shift of an once per revolution periodic term. These parameters were initially determined in a least squares adjustment based on the observed part of the IGS Ultra-Rapid clock solutions. Since October 2006 the program GNSS-VC is operated in a fully automated mode.

To get rid of the 3 hours delay of the IGS Ultra-Rapid-solution we developed a KF (Kalman-Filter) approach which allows to issue clock predictions in near real-time. This is important, because the accuracy especially of predicted clock-corrections decreases rapidly with time. Parameters in the KF are again the 3 coefficients of a quadratic polynomial. After an initial pre-determination of the parameters the KF continuously updates the model using real-time clock corrections calculated from a one-minute data stream based on observations of the RT-IGS network (Real-Time IGS; more than 50 almost globally distributed stations). These once-per-minute clock correction data are output of the program RTR-Control [5]. Clock predictions are calculated every 15 minutes for the upcoming 6 hours period.

We present comparisons of our clock predictions with the Ultra-Rapid and the Rapid solutions of the IGS and with solutions of individual ACs of the IGS. The results of GNSS-VC can be obtained from the institutes webpage [10].

Kurzfassung

Für Positionierungs- und Zeitübertragungsaufgaben mittels GNSS benötigt der Nutzer Informationen über die Satellitenbahndaten und -uhren. Die Analysis Centers (ACs) des IGS (International GNSS Service) stellen die Bahnkoordinaten sowie die Abweichungen der GPS und GLONASS Satellitenuhren zu GPST (GPS-Time) im sp3-Format zur Verfügung. Diese Dateien sind jeweils am folgenden Tag über einen freien ftp-Server erhältlich. Die Bahn- und Uhrinformationen sind das Ergebnis einer Parameterschätzung (vermittelnder Ausgleich nach der Methode der kleinsten Quadrate) auf Basis der Beobachtungsdaten des IGS-Stationsnetzes.

Für Echtzeit- oder beinahe Echtzeit-Anwendungen ist es notwendig, die Satellitenbahnen und -uhren für einen begrenzten Zeitraum vorauszurechnen. An der TU-Wien wurde das Programm GNSS-VC (GNSS-Vienna Clocks) entwickelt, welches seit Oktober 2006 mittels eines eigenen Prädiktionsmodells Satellitenuhrkorrekturdaten über einen Zeitraum von 12 Stunden prädiziert. Modellparameter sind die Koeffizienten eines quadratischen Polynoms sowie die Amplitude und die Phasenverschiebung einer zusätzlichen Sinusschwingung mit der Periodendauer eines Satellitenumlaufs. Als Eingangsgrößen dienen die Ultra-Rapid Produkte des IGS, welche dem Nutzer mit einer Verzögerung von ca. 3 Stunden auf der Homepage des IGS [11] zur Verfügung stehen. Die mittels GNSS-VC prädizierten Uhrkorrekturen unterliegen deshalb ebenfalls einer entsprechenden Verspätung.

Um die Uhrkorrekturdaten auch in quasi-Echtzeit, also ohne die oben beschriebene Verzögerung, vorhersagen zu können, wurde im Anschluss ein Prädiktionsalgorithmus auf Basis eines KF (Kalman-Filters) entwickelt. Dies ist insofern wichtig, als prädizierte Uhrkorrekturen mit fortschreitender Zeit deutlich an Genauigkeit verlieren. Nach der Bestimmung von Startwerten für die Modellparameter (Koeffizienten eines quadratischen Polynoms) werden diese mit Hilfe des KFs in regelmäßigen Intervallen aktualisiert. Als Eingangsgrößen werden dafür Echtzeit-Uhrkorrekturen des Programms RTR-Control herangezogen [5]. Dieses errechnet Uhrkorrekturdaten im Minutentakt, basierend auf einer Lösung des globalen RT-IGS Stationsnetzes (Real-Time IGS Network), welches zur Zeit mehr als 50 Stationen umfasst. Die über den KF berechneten Uhrdaten werden jeweils alle 15 Minuten für die folgenden 6 Stunden ermittelt. Ihre Genauigkeit liegt dabei im 2-Nanosekunden-Bereich, was einem radialen Distanzfehler von ca. 60 cm entspricht.

1. Introduction

Precise point determination by means of GPS relies to a considerable extent on the quality of available satellite orbits and clock offsets with respect to GPST. GPST itself differs from the International Coordinated Time (UTC) by an amount of up to 40 ns (nanoseconds), ignoring leap-second differences. For many geodetic applications using differencing schemes the broadcast ephemeris, currently issued with an accuracy of about ±2 m and a clock rms (root mean square) of ±3-5 ns, are sufficient. To achieve highest precision, especially over baseline lengths larger than 20 km, the user is well-advised to take advantage of precise ephemeris provided by the IGS [11]. Moreover, when analysing pseudorange and phase data, these products allow users to determine consistent coordinates and clock values even for an isolated GPS receiver with an internal accuracy of a few centimetres (PPP, precise point positioning) [8].

The IGS provides so-called "IGS Final Orbits" of all GPS satellites on a weekly basis since the end of 1993. Compared to broadcast orbit information these ephemeris are more accurate by a factor of about 100, i.e. a few centimetres with a satellite clock rms of less than ±0.1 ns. They are available for post processing applications with a delay of 13 days (counted from the last day of the week which is contained in the orbit-file). IGS combined clocks are based on a linear alignment to GPST separately for each day. So while the internal stability of ±0.1 ns is quite good, the day-to-day stability of this reference is poor. Besides, the IGS also provides so-called "Rapid" solutions (IGR) with a slightly lower quality which are available at 17.00 UTC the following day. IGS Final and Rapid solutions are available from the IGS website at [11] free of charge.

It is worth mentioning that a joint project of the IGS and the BIPM (Bureau International des Poids et Mesures) aims to develop and demonstrate the operational capabilities of satellite navigation systems for time transfer [1, 9]. In this context concepts to steer the IGS time scale to UTC were investigated [7]. This will allow for a general dissemination of accurate and easily accessible UTC in the near future.

Moreover, since November 2000, the IGS distributes "Ultra-Rapid" products (IGU) comprising precise GPS satellite orbits and satellite clocks for real-time or near real-time applications. This solution, issued 4 times daily, contains both an observed and a predicted part. Both cover a period of 24 hours. While the orbits of the predicted part are output to an integration of the well-known force field the clocks have to be extrapolated by means of a sophisticated prediction model.

Nevertheless, for real-time applications which require high accuracy the situation should be further improved. Broadcast orbits and clocks are, e. g. for PPP, to imprecise, while the problem with the Ultra-Rapid products is their latency of 3 hours. A real-time prediction of satellite orbits and clocks without any latency and with a high accuracy is at the moment still not available. For this reason we developed the program GNSS-VC (see chapter 2) which has been enhanced recently by a KF-approach (see chapter 3).

2. GNSS-VC – basic principle

In the past years we developed at the IGG (Institute of Geodesy and Geophysics, Research Unit Advanced Geodesy, TU-Vienna) a program to predict GNSS satellite clock corrections. The basic version of GNSS-VC determines the parameters of a prediction model which reads:

$$p(t) = a \cdot t^2 + b \cdot t + c + A_0 \cdot \sin(\omega \cdot t + \phi) \quad (1)$$

with: a, b, c ... polynomial coefficients
$\quad\quad A_0$ ... amplitude
$\quad\quad \omega$ ... frequency
$\quad\quad \phi$ ... phase shift

The nature of the model results from the assumption that the satellites oscillator operates on a slightly shifted and drifting frequency with respect to (w.r.t.) the nominal frequency. Clock observations (phase-observations of this oscillator) are obtained by integration of this deviating satellite frequency and their description therefore yields in a quadratic polynomial. Most satellite clocks satisfy this model very well (see figure 1a).
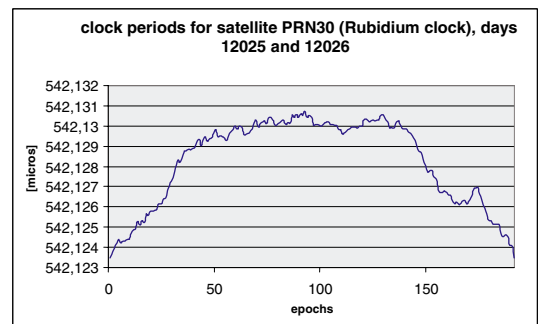
Figure 1a: IGU clock corrections of PRN30 (Rb), GPS week 1202, days 5, 6

Moreover we add a cyclic term to our primary model because of a periodic effect (w.r.t. the satellites revolution time) which basically concerns satellites working on Cesium (Cs) clocks. We assume that the periodic effect of the satellite clock data relate to temperature variations. The Cesium clocks also show a slightly worse stability over time than Rubidium (Rb) clocks, displayed in figure 1b by means of the Allan-variance [1].
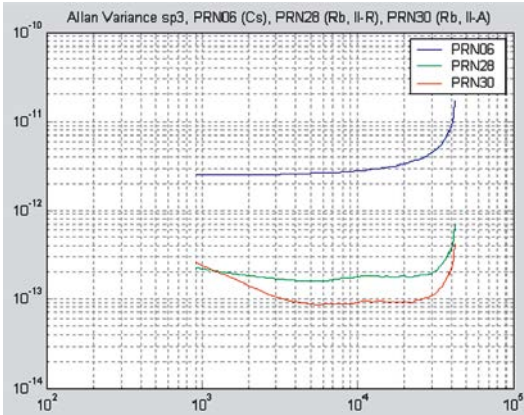
Figure 1b: Allan-variance for different clock types, x-axis in seconds

If the Allan Variance is displayed with respect to time in a logarithmic scale on both axis, one can identify the noise characteristic of the signal. The value of parameter α in table 1 allows to distinguish between different kinds of noise. α is parameter of a model Sy(f) describing the power spectral density of time or frequency fluctuations. For frequency analysis this model reads as

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha \qquad (2)$$

 is a constant depending on the source data and usually ranges between -2 and $+2$. Figure 1c shows the different model slopes of the Allan Variance representing various noise characteristics of a signal. p denotes the slope of the straight line.

α	Noise
2	White phase noise
1	Flicker phase noise
0	White frequency noise
-1	Flicker frequency noise
-2	Random walk frequency noise

Table 1: Allan-variance – noise characteristics
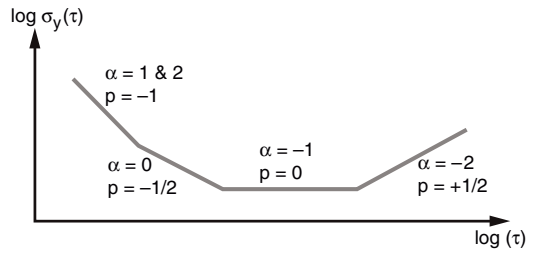
Figure 1c: Allan-variance – noise characteristics

Because we used in figure 1b 24 hours of clock observations with a temporal resolution of 900 s the variance graphs of PRN06 and PRN28 mainly cover the flicker noise part while the more stable clock of PRN30 shows white frequency noise up to 4000 s. The extreme steep gradient of the graphs at the longest interval around 40000 s are artefacts of the computation model used.
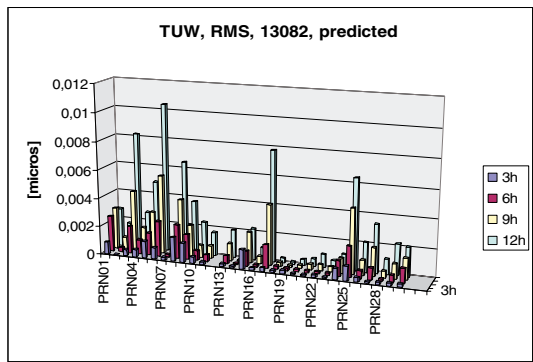
Figure 1d: rms of TUW clock predictions w.r.t. IGR, GPS week1308, day 2

In order to explore reliable model parameters, valid over a span of about 12 hours, we use the available observed part of the IGS Ultra-Rapid solutions of the past 48 hours. A least squares adjustment determines the coefficients of the prediction model presented by equation (1).

To evaluate of the quality of our results we calculated the rms w.r.t. to the IGS Rapid clock corrections (see figure 1d) for different prediction time spans. Every interval starts at 00:00:00, lasting for 3, 6, 9, and 12 hours.

For more detailed information about the algorithm used in GNSS-VC as well as comparisons of the results (w.r.t. the IGS Rapid solution and the ACs Ultra-Rapid solutions) we refer to [2] and to various posters of the authors at the institutes homepage [10].

3. Kalman-filtering

Because of the latency of the Ultra-Rapid solutions (3 hours) as well as the necessary computation time the delay of our predicted clock corrections is about 3 hours and 15 minutes. To get rid of this time delay we developed an algorithm based on a KF [3], using input-data with a time resolution of one minute. These data results from the program RTR-Control, which uses the observations of the RT-IGS station network to calculate real-time clock corrections every 15 seconds. Afterwards the clock corrections are combined in one smoothed solution per minute. The resulting data can be retrieved by registered users from [12], where the clock correction table is actualized every minute. These data also serve as input for the KF-algorithm in GNSS-VC.

To set up reliable initial values of our parameters we use 10 consecutive epochs of the one-minute-data of RTR-Control. We determine an offset and the linear coefficient of our parameter vector. The offset might be the last input-value of the initial values (representing the first prediction epoch), the linear coefficient corresponds to the slope of a straight line through the first to the last data-point. The initial value of the quadratic coefficient of the polynomial is zero. Currently we do not evaluate the coefficients of the periodic term given in equation (1). Moreover with the small amount of input data of 10 clock corrections in the first test version it does not make sense to evaluate an amplitude and a phase shift. Thus the KF approach does not distinguish any more between different clock types (Cesium or Rubidium) because the parameter vector is used for predictions over not more than 6 hours while the cyclic term has a period of about 12 hours (revolution period).

After the initial process we update the parameter vector every minute. If there appears a gap in the data the new parameters are predicted as long as the gap does not exceed a limited time window, which at the moment is set to 10 minutes. These results will be published via internet (see chapter 5). Every 15 minutes the new parameters are used for predicting clock corrections over an interval of 6 hours. Within this time span the rms of the predictions w.r.t to the IGS Rapid solutions remains at the two-nanoseconds-level (or below) for most of the satellites.

The results are also stored in a format similar to the sp3-format with a time interval between consecutive epochs of 15 minutes. The reason is just convenience for comparisons with the IGS

products. Of course in the future the time resolution of the predicted clock corrections can vary depending on the application.

4. Comparisons

To get information about the quality of our predictions we computed the rms w.r.t. the IGS Rapid solution by means of:

$$rms_i = \sqrt{\left(\sum (\Delta IGR_i)^2\right)/n} \qquad (3)$$

with: ΔIGR_i difference between the IGS Rapid solutions and TUW
i satellite
n number of epochs

The rms contains the whole prediction time span over 6 hours. As in former studies, e.g. [2], we take into account an offset between the two data sets as well as a linear drift in time. The following figures 2a-c show the improvement of the KF predictions at the begin of a test-day in GPS-week 1401.
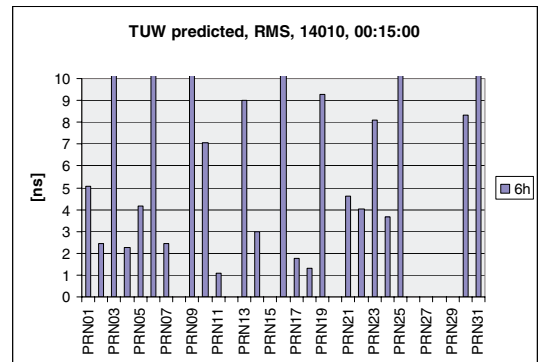
Figure 2a: rms w.r.t. IGR for week 1401, day 1, prediction time 00:15:00
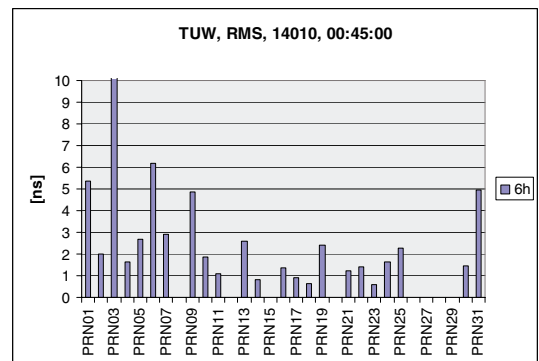
Figure 2b: rms w.r.t. IGR for week 1401, day 1, prediction time 00:45:00
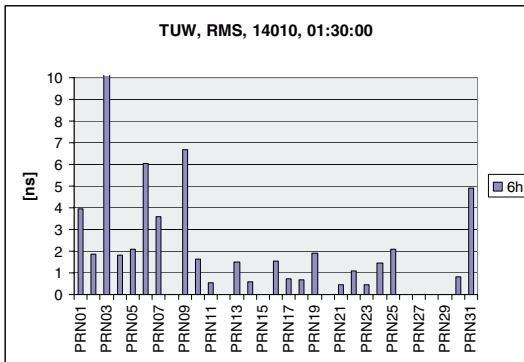
Figure 2c: rms w.r.t. IGR for week 1401, day 1, prediction time 01:30:00

It is obvious that the quality of the predicted clock corrections increases over the time (for an easy comparison the y-axis of all diagrams is scaled to 10 ns). For the first prediction epoch (figure 2a) the model-parameters are, w.r.t. the input data of this day, calculated from the ten initial values as described above and corrected for just one epoch with the KF. Afterwards the predicted clock corrections improve very fast and already after 1.5 hours (figure 4.b) the rms-values for more than half of the given satellites come close to the expected 2 ns-level. After this time span the KF delivers stable output data. Of course it is also possible, that the rms of a satellite gets worse for shorter time spans. This happens if the input data worsens or clock corrections lack for a couple of epochs in the input file. For small gaps the prediction parameters are not corrected for the new epoch but just calculated with the prediction-rules defined in the transition-matrix of the KF. Therefore the predicted clock corrections get continuously worse until at a certain time the predictions are stopped (currently after 10 minutes). Whether satellites are missing or not solely depends on the input data of RTR-Control.

5. Further work

In the near future we will improve the current test-version of GNSS-VC to provide reliable clock predictions via the "IGU – Real Time Monitoring"-homepage [12] of the IGS-RT working group. Because the IGS-RT network (figure 3) is currently still to sparse to track all satellites by enough stations at any time there remain gaps in the observations. Therefore it is necessary to consider reasonable intervals to start and stop the clock correction predictions of single satellites.
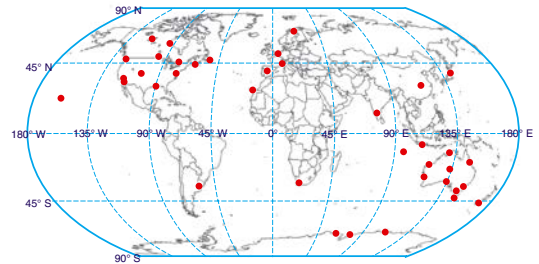
Figure 3: IGS-RT network, spring 2007

According to figures 2a-c it does not make sense to start with the predictions just a couple of minutes after the first parameter vector has been calculated in the KF. A good starting interval might be about 1.5 hours following the fact that the rms gets stable after about this time span. Secondly, there is no reasonable breakpoint for the case of gaps in the input data at the moment. Currently this parameter is set to 10 minutes, but it has to be checked whether it is possible to expand this limit. One might consider a threshold for the rms (e.g. ± 10 ns, comparable with a radial distance error of ca. ± 3 m), so that satellites beyond this value will not be available in the output files. Therefore the predicted clock data will have to be compared to the IGS Ultra-Rapid clock corrections right after their predetermination.

6. Summary

As shown above it is possible to compute GNSS clock correction data in quasi real-time for a limited time span (about 6 hours) with a simple prediction model in form of a quadratic polynomial based on a KF. The rms remains in most cases below ± 2 ns for the 6 hours prediction interval. For the future (July 2007) we plan to provide our clock products via the institutes webpage. The advantage of this new clock solutions is that it is computed in almost real-time whereas the former predicted clock corrections as well as the IGS- and the ACs Ultra-Rapid solutions are released with a delay of about 3 hours.

References

[1] *Audoin C., Guinot B.*: Measurement of Time and Frequency. Cambridge University Press, Cambrigde, 2001

[2] *Broederbauer V., Weber R.*: Results of Modelling GPS Satellite Clocks. VGI, Heft 1/2003

[3] *Hofmann-Wellenhof B., Legat K., Wieser A.*: Navigation. Springer-Verlag Wien, 2003

[4] *Opitz M., Weber R., M. Caissy*: Real Time Integrity Monitoring of the IGU Satellite Orbits by Means of the RTIGS Network. (Poster, EGU 2006)

[5] *Opitz M., Weber R.*: Real Time Monitoring of IGS Products within the RTIGS Network. Proceedings of the IGS Workshop, Darmstadt, 8 – 12 May 2006 (in print)

[6] *Ray J.*: IGS/BIMP Time Transfer Pilot Project. IGS 1999 Technical Reports, November 2000

[7] *Senior K., Koppang P., Matsakis D., Ray J.*: Developing an IGS Time Scale. Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on, Volume: 50, pp. 585-593

[8] *Zumberge J.F., Helfin M.B., Jefferson D.C., Watkins M.M., Webb F.H.*: Precise point positioning for the efficient and robust analysis of GPS data from large networks. Journal of Geophysical Research (JGR), Vol. 102, No. B3, pp. 5005-5017, 1997

[9] BIPM: http://www.bipm.fr/

[10] IGG TU-Vienna: http://www.hg.tuwien.ac.at/

[11] IGS: http://igscb.jpl.nasa.gov/

[12] IGU Real Time Monitoring: http://rtclocks.hg.tuwien.ac.at/

Contact

Dipl.-Ing. Veronika Broderbauer: Department of Advanced Geodesy, TU-Vienna, Gusshausstr. 27-29/E1281, A-1040 Wien, Austria. E-mail: veronika@mars.hg.tuwien.ac.at

Dipl.-Ing. Michael Opitz: Department of Advanced Geodesy, TU-Vienna, Gusshausstr. 27-29/E1281, A-1040 Wien, Austria. E-mail: mopitz@mars.hg.tuwien.ac.at

Dr. Robert Weber: Department of Advanced Geodesy, TU-Vienna, Gusshausstr. 27-29/E1281, A-1040 Wien, Austria. E-mail: rweber@mars.hg.tuwien.ac.at

Changes in the OLG GPS time series due to new adjustment models

Cornelia Haslinger, Sandro Krauss and Günter Stangl

Abstract

IGS [1] (IGS = International GNSS Service, GNSS = Global Navigation Satellite System) and EPN (European Permanent Network) Local Analysis Centres changed their adjustment strategy, starting with GPS week 1400. The most important were the change from relative to absolute antenna phase calibrations, the transition to the reference frame IGS05/ITRF2005 (International Terrestrial Reference Frame 2005) and the lowering of the minimum elevation. The jumps in the time series of the coordinates in two networks of the analysis centre OLG (Observatory Lustbuehel Graz), AMON (Austrian Monitoring Network) und MON (Monitoring Oriental Network), are shown in examples.

Kurzfassung

Die Auswertungsstrategie von IGS [1] (IGS = International GNSS Service, GNSS = Global Navigation Satellite System) und EPN (European Permanent Network) wurde mit der GPS-Woche 1400 massiv geändert. Die wichtigsten Änderungen sind die Verwendung von absoluten Kalibrierwerten für die Antennen, die Senkung der Mindestelevation und ein neues Referenzsystem IGS05/ITRF2005 (International Terrestrial Reference Frame 2005). Die Auswirkungen auf die Zeitreihen zweier nach den internationalen Standards ausgewerteter Netze des Analysezentrums OLG (Observatory Lustbuehel Graz), AMON (Austrian Monitoring Network) und MON (Monitoring Oriental Network), werden exemplarisch beschrieben.

1. Introduction

Starting with GPS week 1400 (November 2006) the analysis centres of IGS altered officially several items in their computation strategy of the GPS orbits. The main differences are the use of absolute antenna phase calibrations for space vehicles and ground receivers and the adoption of a new reference frame IGS05. The EPN analysis centres also changed their strategies at the same time, adding some other features which should improve the repeatability of the coordinates, especially of the vertical component. The major changes are given in Table 1. Not all required basic data are available at this time (March 2007). Concerning the calibration values a sub-set has no absolute calibrations; especially certain combinations of antenna plus dome are not yet measured. The fallback rules are to use the calibration values of the corresponding antenna without dome which might lead to improper results. IGS05 is derived from a set of coordinates computed with absolute calibration values for several months. However, it is very well aligned to ITRF2005 which is based on relative calibrations for the GPS contribution [2]. Presently there are no coordinates and velocities for several EPN stations in ITRF2005. Based on a transformation

formula [3] [4] the ETRF (European Terrestrial Reference Frame), the realization of ETRS89 (European Terrestrial Reference System 1989) is connected to ITRF. Because several countries adopted ETRS89 as the basic reference system for geodesy it is necessary to know how the coordinates of the realizations (= permanent and epoch GPS stations) will change. Based on the coordinate series derived from processing data from two networks of permanent GPS stations this study highlights the different processing strategy. The transformation of AMON into ETRS89 is needed as a basis of the geodetic reference frame of Austria. The effect of the change on the coordinates is compared to another network, MON, which covers the Eastern Mediterranean and the Western Indian Ocean. It is too early to investigate the effect on the station velocities but it is expected that already some conclusions for the future estimations can be drawn.

2. Networks

Two networks, which are computed at a weekly basis since at least 1999 at the analysis centre OLG are the basis for this study.

The first, AMON, serves as the Austrian realization of the reference systems ITRS2000

Item	Old Strategy	New Strategy	Remarks
Antenna Phase Calibrations	Relative from IGS and NGS	Absolute from IGS plus individual calibrations, if available	Robot measurements and transformed values relative → absolute
Reference Frame	ITRF2000	IGS05	IGS05 based on a transformation of ITRF2005
Minimum Elevation	10 degrees	3 degrees	Optional (recommended)
Troposphere Model	Dry Niell	Wet Niell	Estimation of zenith delays
Horizontal Gradients of the Troposphere	No	Yes, 24 h tilting	Optional (recommended)

Table 1: Major changes in processing before and after GPS week 1400, EPN analysis centres.
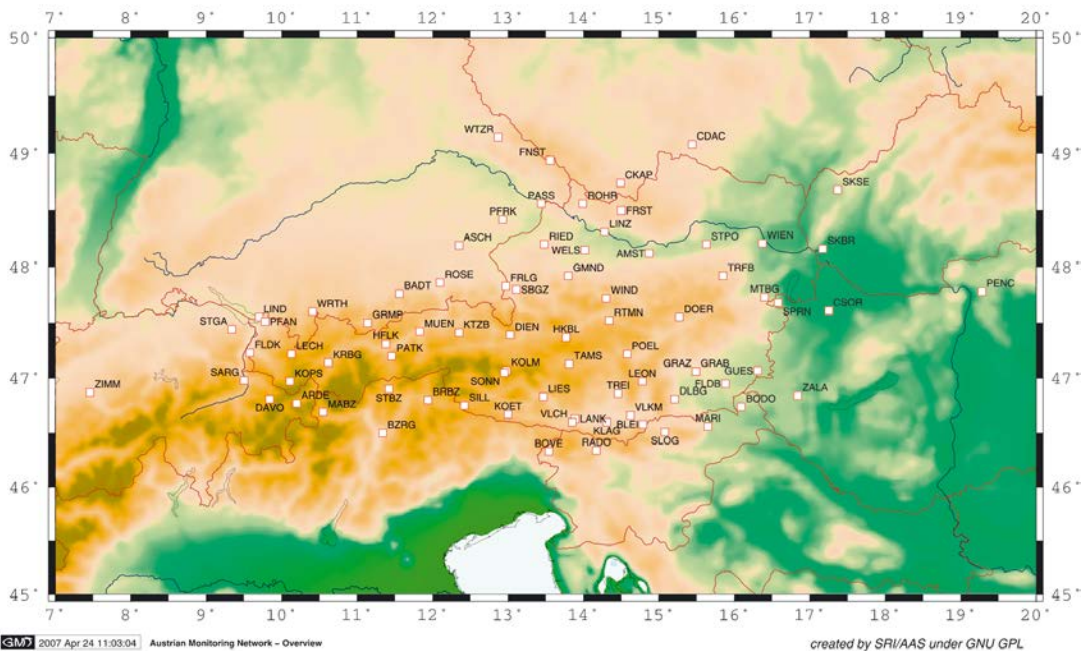
Figure 1: Austrian Monitoring Network (overview)

and ETRS89. It is a very compact and precise network consisting of about 80 stations. As can be seen from Figure 1 the network do not only cover Austria, also parts from Germany, Italy, Switzerland, Slovenia, Hungary, Slovakia and the Czech Republic are included.

The second network, MON, covers the Eastern Mediterranean, Arabia and the Western half of the Indian Ocean. Compared to the AMON network, where the mean baseline length is approximately 50 km, this one covers a larger area with baselines up to 3500 km. Figure 2 shows the distribution of more than 50 stations, which presently are part of the network.

In Table 2 the stations of AMON and MON are listed, arranged according to their status (IGS, EPN or both). GPS sites which are neither IGS nor EPN stations are here not mentioned.
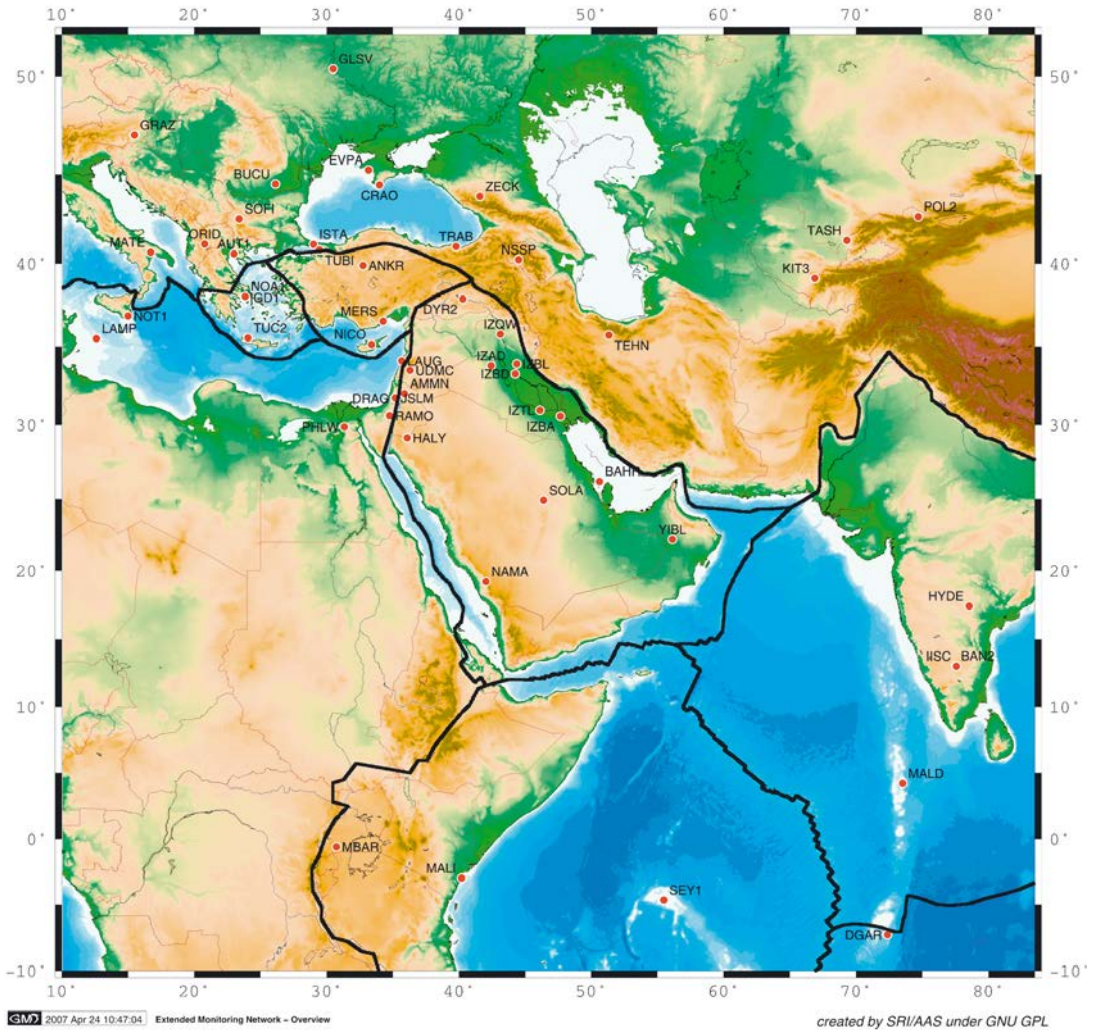
Figure 2: Monitoring Oriental Network (overview)

IGS	IGS+EPN	EPN
AMMN, BAHR, BAN2, CRAO, DGAR, DYR2, HALY, HYDE, IISC, KIT3, MBAR, MTBG, NAMA, POL2, SOLA, TEHN, YIBL	ANKR, BUCU, BZRG, DRAG, GLSV, GRAZ, HFLK, ISTA, MATE, NICO, NOT1, NSSP, ORID, PENC, RAMO, SOFI, TRAB, TUBI, WTZR, ZECK, ZIMM	AUT1, EVPA, LINZ, NOA1, PFAN, SBGZ, TRFB, TUC2

Table 2: AMON and MON stations according to their status

The receivers and antennas used at the GPS stations are from TRIMBLE, LEICA, ASHTECH, JAVAD and TOPCON. Due to the consideration of individual antenna calibrations since November 2006 the radome types of the stations became very important. The radome types which are used

in the two networks are listed in Table 3 including the number of stations where the radomes are applied.

Radom	Number of stations (AMON)	Number of stations (MON)
NONE	19	24
TZGD	10	—
LEIC	9	—
LEIS	9	2
BEVA	11	—
GRAZ	16	—
UNAV	1	2
SNOW	1	7
CONE	2	—
SCIS	—	2
SCIT	—	1

Table 3: Radome types

MON includes no stations with calibrated antennas, but AMON contains nine GPS sites with calibrated antenna-radome combinations (Table 4).

3. Generation of Time Series

For both networks the weekly solutions generated from daily ones are used. The processing was done to the standards of EPN [5] of the valid epoch. MON was reprocessed for the years 1999-2001 because the original network was small and needed to be extended. Despite different processing algorithms in the past the changes seemed to be negligible. Only the last change can be seen in the time series producing substantial offsets. In a first step the normal equations were stacked together with equal weight. Because this was already done in the past, the offsets and outliers found have been applied already. The time series were revised and searched for new offsets or old offsets to be corrected or new outliers. The reference was chosen from stations with consistent coordinates and velocities in ITRF2000 plus some EPN and AMON stations to keep the balance in the Austrian network. MON was attached to stations which are all supposed to be on the Eurasian Plate. Each weekly solution was transformed to the system of selected stations using minimum constraints.

The different quality in precision between AMON and MON is shown in Figure 3. The distribution of the r.m.s values from the weekly repeatability show that the lateral residuals of AMON are concentrated at the 1 mm level, the vertical ones around 3 mm. The MON stations have not a good distribution, the good ones are at the 3 mm (lateral) and 5 mm (vertical) level. There are some stations in the Indian Ocean which have r.m.s worse than 10 mm, the largest (Up component) has 34 mm r.m.s. All stations summed up in Figure 3 are active since more than one year. The recent ones normally show larger r.m.s. because the seasonal variations have a larger influence.

GPS station	Receiver	Antenna	Radome
DOER	TRIMBLE NETRS	TRM41249.00	TZGD
FLDB	TRIMBLE NETRS	TRM41249.00	BEV A
HKBL	TRIMBLE NETRS	TRM29659.00+50	BEV A
KRBG	TRIMBLE NETRS	TRM41249.00	BEV A
MUEN	LEICA SR520	LEIAT504	BEV A
POEL	TRIMBLE NETRS	TRM41249.00	TZGD
TAMS	TRIMBLE NETRS	TRM41249.00	TZGD
WELS	TRIMBLE NETRS	TRM41249.00	BEV A
WTZR	TPS E_GGD	AOAD/M_T	NONE

Table 4: Calibrated antenna-radome combinations of AMON
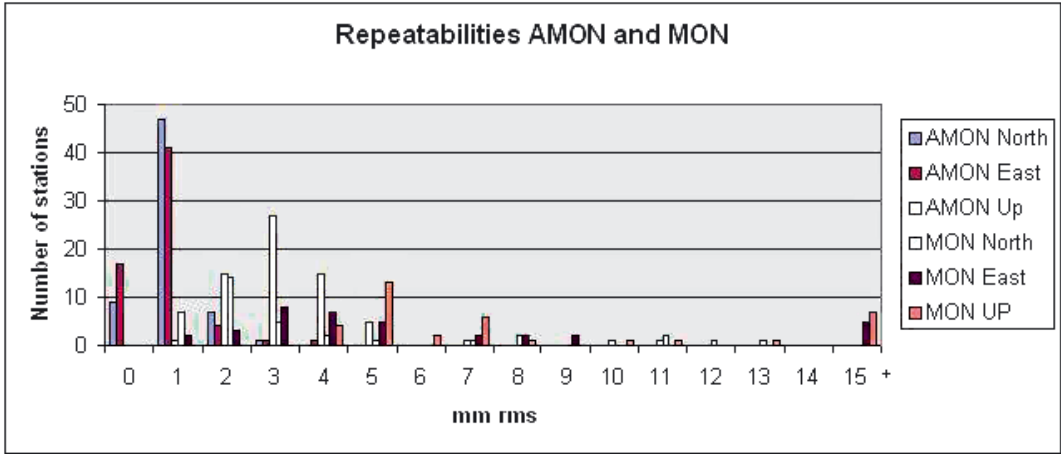
Figure 3: Histogram of r.m.s values for AMON and MON in North/East/Up

4. Offsets and Outliers

Offsets are jumps in the time series, which remain during plenty of weeks. Outliers on the other hand show up for fewer weeks, frequently only one week and the values are larger than those of offsets. Both effects can be caused by equipment modifications and processing changes. Outliers can also be affected by bad weather conditions like hard snowfall. The detection of offsets and outliers occurs graphically. Outliers larger than ± 10 mm in North and East and larger than ± 20 mm in Up are eliminated. Offsets are estimated with an accuracy of 0.1 mm and applied to the time series.

Figure 4 shows the time series of Krahberg / Landeck in Austria. The blue circle marks an outlier in the North component which was not eliminated due to its small value (about 7 mm).

The green lines in Figure 4 present the equipment changes (GPS weeks 1334, 1339 and 1353), whereas the change of antenna and receiver in GPS week 1334 causes an offset of -2.1 mm in North, $+2.4$ mm in East and -9.1 mm in the Up component. The processing changes since GPS week 1400 associate also offsets (yellow line) which are not yet estimated due to the low amount of available data since that point in time.
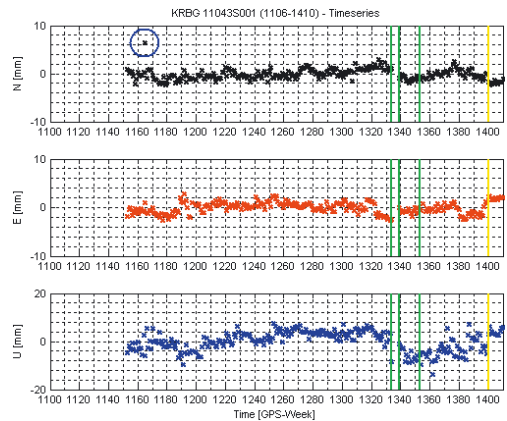
Figure 4: Example of typical offsets in the time series (KRBG)
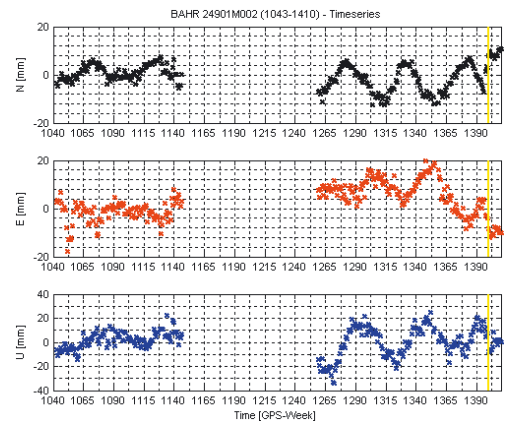
Figure 5: Time series with large periodic features (BAHR)

The GPS station in Manama/Bahrain features periodic effects in each direction (Figure 5). The amplitudes are nearly 10 mm in North and East and 20 mm in the Up component. The periodicity shows an annual term and is influenced by the seasons. The first part of the time series, which results from the reprocessing, shows a smaller amplitude compared with the end of the series. This effect is caused by the strategy changes, where the former strategy, which was adopted at the reprocessing and is now a part of the new strategy (low elevation angle, horizontal gradients, Niell wet part), seems to smooth the time series in this part.

For closer investigation of amplitudes and phases it is planned to generate the phase spectrum of the time series to get more information about seasonal effects.

5. Comparison and Results

The changes due to the new adjustment strategies can be seen in the time series. Because there were many changes the contribution of each change sums up to a common change which seems to have a general component and an individual one. To get an insight into the different contributions GPS week 1400 was investigated by varying the major changes:

- Relative calibrations + minimum elevation 10 degrees, no horizontal gradients (old solution),
- Relative calibrations + minimum elevation 10 degrees + horizontal gradients,
- Relative calibrations + minimum elevation 3 degrees, no horizontal gradients,
- Relative calibrations + minimum elevation 3 degrees + horizontal gradients,
- Absolute calibrations + minimum elevation 10 degrees, no horizontal gradients,
- Absolute calibrations + minimum elevation 10 degrees + horizontal gradients,
- Absolute calibrations + minimum elevation 3 degrees, no horizontal gradients.

These seven variants were tested against the new solution (absolute calibrations + minimum elevation 3 degrees + horizontal gradients). It turned out that generally AMON is not too sensitive to the changes whereas in MON at some stations changes of several centimetres appear. In Table 5 examples of stations are given with slight changes (ORID) and with the maxima in the vertical component. The two most relevant cases (old

strategy and old strategy with absolute calibration values) are shown. Due to the fact that the Bernese Software (BSW) in version 4.2 [6] ignored the calibration values the differences at SBGZ should be considered as the change to the correct antenna type. Worth of mention is that the coordinates of SBGZ in the EUREF (European Reference Frame) [7] sub-network were always correct. The stations GUES and SKSE demonstrate that stations with low elevation observations are only affected if the minimum elevation is lowered.

The very poor quality at DGAR and SEY in MON is the result of large data gaps which occur nearly every day. DGAR shows a consistent difference whether absolute or relative calibrations are applied whereas SEY1 might suffer from a combined effect (calibration values, data gaps). Considering the estimation at low elevations we can see that both stations are very sensitive. In the horizontal component the comparison between old and new strategy yields a difference up to 20 mm. Introducing absolute calibration values reduces the difference to 5 mm. This means that a combination of different azimuth/elevation corrections of the antenna and a change in the model of the troposphere may be responsible for those changes. The influence of the horizontal gradients seems to be in the range of several millimetres. The baselines configuration can reduce the effect of observations at low elevations whereby large changes are almost randomly distributed. The Bernese Software offers a baseline adjustment strategy (OBSMAX), which uses the number of observations as optimization criterion. Also the calibration of different antenna types (new calibration, relative transformed to absolute values only) may play a certain role. Altogether the common effect for both networks due to the change is rather small (< 2 mm for AMON, 5 mm for MON).

6. Conclusions

The effect of the new adjustment strategy on the time series is well seen. In the future it would not be possible to derive station velocities without further actions. The best way would be to reprocess the old solutions with the new strategy. The reprocessing of MON in 2006 for the years 1999-2001 showed that approximately one year reprocessing would request one month of computation time. A quicker way would be to estimate the offsets caused by the changed strategy. Probably this is not too incorrect because very few stations observed below five degrees in the past and one

Station	Calibration absolute Minimum elevation 10 degrees, no horizontal gradients [mm]	Calibration relative Minimum elevation 10 degrees, no horizontal gradients (old solution) [mm]
SBGZ	9.2	10.1
GUES	−1.0	13.7
SKSE	−0.1	−10.9
ORID	1.2	−0.6
SEY1	−12.7	8.1
DGAR	−23.0	−24.9

Table 5: Change of Up component for GPS week 1400 against new adjustment strategy

important contribution of the offsets comes from the lowering of the minimum elevation at some stations. The second important contribution, the new calibration values, cannot be investigated very easily because most of the antenna-radome combinations have been neglected in the past. The use and the meaning of the horizontal gradients should be investigated further more. Despite the positive effect there might be a geophysical influence in the seasonal variations which is now filtered away. Concerning the Austrian national reference system estimating the offsets can be estimated and applied. It is a political question which coordinates should be "official", the old ones or the new ones. As shown above the maximum change in Austria will be about 10 mm in the vertical component. It has to be decided if this change should be applied for a handful of stations. A large distortion in the Austrian national reference network will not take place because the average changes are much smaller. Concerning MON the new adjustment would bring a big benefit in the future because the time series become considerably smoother. For this reason it seems really necessary and worthwhile to consider a reprocessing in the next years.

References

[1] *Dow, J.M., Neilan, R. E., Gendt, G., The International GPS Service (IGS)*: Celebrating the 10th Anniversary and Looking to the Next Decade, Adv. Space Res. 36 vol. 36, no. 3, pp. 320-326, 2005. doi:10.1016/j.asr.2005.05.125.

[2] *Ferland, R.*: IGS05 Fine Tuning, IGSMail 5456.

[3] *Boucher, C., Altamimi, Z. (2001)*: Specifications for reference frame fixing in the analysis of a EUREF GPS-campaign, available from http://lareg.ensg.ign.fr/EUREF/memo.pdf.

[4] *Altamimi, Z., Boucher, C. (2002)*: The ITRS and ETRS89 Relationship: New Results from ITRF2000, EUREF Publication No. 10, Mitteilungen des Bundesamtes für Kartographie und Geodäsie Band 23, Frankfurt/Main, pp. 49-52.

[5] guidelines for EPN Analysis Centres (2002), http://www.oma.be/guidelines

[6] *Hugentobler, U., Dach, R., Fridez, P., Meindl, M. (2006)*: Bernese GPS Software, Version 5.0, Draft

[7] *Adam J., W. Augath, C. Boucher, C. Bruyninx, A. Caporali, E. Gubler, W. Gurtner, H. Habrich, B. Harsson, H. Hornik, J. Ihde, A. Kenyeres, H. van der Marel, H. Seeger, J. Simek, G. Stangl, J. Torres, G. Weber, 2002*: Status of the European Reference Frame – EUREF, International Association of Geodesy Symposia, IAG Scientific Assembly, Springer, ed. J. Adam and K.-P. Schwarz, Vol. 125, pp 42-46

Contact

Dipl.-Ing. Cornelia Haslinger: Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: cornelia.haslinger@oeaw.ac.at
Dipl.-Ing. Sandro Krauss: Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: sandro.krauss@oeaw.ac.at
Dipl.-Ing. Mag. Dr. phil. Günter Stangl: Federal Office of Metrology and Surveying, Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: guenter.stangl@oeaw.ac.at

The Intra-Plate Velocities of GPS Permanent Stations of the Eastern Alps

Cornelia Haslinger, Sandro Krauss and Günter Stangl

Abstract

Since 2000 the GPS (Global Positioning System) permanent stations in Austria are monitored at a weekly basis, applying the international guidelines of analysis centres of IGS [1] (IGS = International GNSS Service, GNSS = Global Navigation Satellite System) and EPN (European Permanent Network). The number of stations was about 80 at the beginning of 2007. Station velocities have been derived from coordinate time series by taking into account offsets and detecting outliers. The estimated velocities have a precision of 1 mm/year laterally and 1–3 mm/year vertically. These velocities are reduced by the rotational velocity of the Eurasian Plate, derived from ITRF2000 (International Terrestrial Reference Frame 2000) [3], in order to investigate intra-plate movements. Apart from some local movements the velocities are in the range of 0–3 mm/year, but can already be grouped into different clusters. The movement of the Alpine Forelands is identical to the rotation of the Eurasian Plate whereas the region between the Alps and the Dinarides seems to undergo an eastward movement. Within the Eastern Alps the situation is still unclear due to some reasons, e.g., young stations, poor coverage and local movements.

Kurzfassung

Seit 2000 werden GPS (Global Positioning System) Permanent Stationen nach den internationalen Richtlinien der Analysezentren IGS [1] (IGS = International GNSS Service, GNSS = Global Navigation Satellite System) und EPN (European Permanent Network) wöchentlich ausgewertet. Die Anzahl der Stationen betrug 80 zu Beginn des Jahres 2007. Von den Zeitreihen der Koordinaten wurden Geschwindigkeiten unter Berücksichtigung von Sprüngen und Erfassung von Ausreißern abgeleitet. Die geschätzten Geschwindigkeiten weisen eine Präzision von 1.0 mm/Jahr lateral und 1–3 mm vertikal auf. Um die Bewegungen innerhalb einer Platte untersuchen zu können, wird die Rotationsgeschwindigkeit der Eurasischen Platte, welche vom ITRF2000 (International Terrestrial Reference Frame 2000) [3] hergeleitet wurde, von den geschätzten Geschwindigkeiten abgezogen. Abgesehen von einigen lokalen Bewegungen bewegen sich die Geschwindigkeiten in einem Schwankungsbereich von 0 bis 3 mm/Jahr, können aber bereits in verschiedene Gruppen unterteilt werden. Die Bewegung des Alpinen Vorlandes ist identisch der Rotation der Eurasischen Platte, wohingegen die Region zwischen den östlichen Alpen und den Dinariden einer Bewegung nach Osten zu folgen scheint. Innerhalb der Ostalpen ist die Situation aus verschiedenen Gründen, wie z.B., dem geringen Alter mancher Stationen, der mangelhaften Flächendeckung und lokaler Bewegungen, nach wie vor unklar.

1. Introduction

The European Terrestrial Reference System ETRS89 is derived from the International Terrestrial Reference System (ITRS) by removing the rotation of the Eurasian Plate. It was designed to keep the changes of the coordinates within an acceptable range for national mapping agencies. Austria adopted the system in 2003. The realization of ETRS89 is derived from the ITRS realizations (e.g. ITRF2000). Austria covers a part of the Eurasian Plate, which cannot be considered as totally stable. Significant deviations from the rotation of this plate show up in some parts. The residual motion is less than 10 mm/year, but can already be determined. The following article describes the present state of velocity estimation from about 80 GPS permanent stations in and around the Eastern Alps. The GPS network is still being augmented. Many stations are not older than two years. Hence, former results [4] need to be refined. However, there is already an overview of the situation, providing a range of possible intra-plate velocities in Austria and its neighbourhood. Two questions should be answered. First, how will the Austrian reference network be deformed by station velocities for the next decade, and second, is it possible to detect the causes of these residual velocities, related to the geokinematic situation of the Eastern Alps.

2. Coverage

The AMON (Austrian Monitoring Network) consists of about 80 stations in Austria, Germany, Switzerland, Italy, Slovenia, Hungary, Slovakia

and the Czech Republic. Table 1 shows an overview of the station owners and the number of stations under their responsibility.   AMON includes several different types of antennas and radomes. Further information can be found at ([10] → AMON network).

The GPS stations of the AMON are divided in two groups in Figure 1. Those marked with white squares were used for velocity estimations, as they cover many years of observations and show very stable velocities.

Agency, Country	Link	#
BEV, AT	http://www.bev.gv.at	29
ÖAW, AT	http://www.oeaw.ac.at	4
BLVA, DE	http://www.geodaten.bayern.de/bvv_web/blva	10
KELAG, AT	http://www.kelag.at	9
Swisstopo, CH	http://www.swisstopo.ch	5
BEWAG, AT	http://www.bewag.at	2
Geodetska Uprava, SI	http://www.gu.gov.si	5
BKG, DE	http://www.ifag.de	1
LINZ AG, AT	http://www.linzag.at	1
FÖMI, HU	http://www.fomi.hu	4
Ufficio Geodetico, IT	http://www.provincia.bz.it/catasto-librofondiario/4103/geodetico.htm	4
GKU Bratislava, SK	http://www.gku.sk	2
Zememericky Urad, CZ	http://geoportal.cuzk.cz	2

Table 1: Overview of the station owners

Figure 1: GPS Stations used for velocity estimation (white squares)

< 0.5 year	0.5 – 1 year	1 – 2 year	> 2 years
BADT (2), CDAC, CKAP, CSOR, SILL, SKBR, SKSE, SPRN, ZALA	BODO, BRBZ, BZRG, DOER, KOLM, MABZ, POEL, STBZ	AMST, BOVE, DIEN, FLDB, FRST, KOPS (1), RADO, SLOG, TAMS	ARDE, ASCH, BADT (1), BLEI, DAVO, DLBG, FLDK, FNST, FRLG, GMND, GRAB, GRAZ, GRMP, GUES, HFLK, HKBL, KLAG, KOET, KOPS (2), KRBG, KTZB, LANK, LECH, LEON, LIES, LIND, LINZ, MARI, MTBG, MUEN, PASS, PATK, PENC, PFAN, PFRK, RIED, ROHR, ROSE, RTMN, SARG, SBGZ, SONN, STGA, STOPP, TREI, TRFB, VLCH, VLKM, WELS, WIEN, WIND, WRTH, WTZR, ZIMM

Table 2: Age of the stations

Stations, which are in operation for less than one year show velocities strongly varying in size and direction. For this reason these stations marked by black filled circles will only be considered for future velocity estimations. Table 2 shows the age of the stations.

3. Velocity Estimation

Estimation of station velocities follow a similar procedure used in the Time Series Project of EPN [11]. Weekly solutions, derived from daily solutions, were stacked together to form the "RAW Time Series". Then outliers and offsets were defined. Outliers, which are characterized by a single value, which differs from the neighbouring weeks by more than 10 mm, were removed. Offsets are estimated if more than 10 consecutive weeks differ from the previous ones. According to the average r.m.s. of the weekly repeatability, which is about one millimetre for the lateral and about 2.5 mm for the vertical component, it was possible to detect offsets down to two millimetres. All normal equations were used since the start of AMON in 2001. This implies that the normal equations processed with the Bernese Software (BSW) version 4.2 [5] had to be first converted into the format of BSW 5.0. For the a priori changes (outliers, offsets, dome numbers, split of time series) a file with extension .STA was created and updated for each velocity computation. The software needs this file to be able to consider the changes in the processing. Outliers and offsets of the EPN stations were taken and checked to see if an improvement occurred in the time series. Most of them are applied, however, for the stations GRAZ, HFLK and LINZ they had to be modified. Additionally the changes and deletions already known by the OLG (Observatory Lustbuehel Graz) were added. As a reference, the ITRF2000 coordinates/velocities were chosen for the stations GRAZ, HFLK, WTZR and ZIMM. To strengthen the network, the EPN coordinates of PENC and PFAN were added in the same system, because jumps at those stations are already

applied. To balance the reference for the area of Austria a nine years old station with a very smooth time series, STPO (St. Pölten), was added in the East. Their respective coordinates and velocities were estimated, while the a priori values were only used for the datum definition.

The first combination of normal equations leads to time series, which are comparable to the type of "CLEANED Time Series" of EPN. Afterwards the series were checked for new outliers and offsets, and for revisions of former changes. If necessary the changes were introduced and the combination was repeated. From a total of 78 stations, those including less than 53 weeks of observations were removed from the investigations. Most of the stations experience seasonal changes in their coordinates, especially in the height component. Because the amplitudes can reach 20 mm, it is unadvisable to consider their velocities, yet. Therefore, 14 stations plus the test station WIEB were removed from investigation, thus leaving 63 stations providing observations for velocity calculation. The still growing network makes it necessary to repeat the combination. Presently a cycle of 2–3 months seems to be sufficient to monitor the stations for outliers and offsets. A yearly velocity estimation should be sufficient to integrate additional stations into a reliable velocity field. The lower time limit where a station velocity can be seen as well determined amounts to one year. Compared to earlier studies [7] this shows an improvement of the estimation in the last years due to this investigation of a regional area, the station densification and the modernization of analysis strategies.

4. Results

For the 63 stations mentioned above the rotation of the Eurasian Plate defined by ITRF2000 was removed thus clearly revealing the differences of velocities over the network area. For comfort the residual velocities are called "intra-plate velocities" without preconditioning the tectonic inter-

pretation. For some stations it is possible to check the derived velocities from other references, like ITRF2000, ITRF2005 and EPN. For that reason the full values are compared in Table 3. It should be noted that the system velocities (−0.2/0.1/−1.8 mm/year) of ITRF2005 [9], relative to the other systems are removed.

Apart from the SBGZ velocity in ITRF2000, which is ostensibly obsolete, the values in Table 3 are very similar. The differences at the 1mm/year level are mostly caused by the application of offsets to the coordinate time series. As the seven stations are at least observing since 1999, the length of the time series has almost no influence. The offsets may affect the time series considerably, e.g., when the equipment is changed triennially.

STATION	VELOCITIES [MM/YEAR]						SYSTEM
	VEL$_X$	VEL$_Y$	VEL$_Z$	VEL$_N$	VEL$_E$	VEL$_U$	
GRAZ 1001M002	−16.4	18.9	9.8	14.5	22.61	−0.2	AMON
	−16.9	18.2	8.9	14.4	22.1	−1.3	ITRF2005
	−17.6	18.1	8.2	14.5	22.1	−2.3	ITRF2000
	−14.9	19.2	10.8	14.1	22.5	−2.3	EPN
WTZR 14201M010	−15.8	17.4	8.5	14.3	20.5	−1.1	AMON
	−15.3	17.4	9.6	14.6	20.4	0.0	ITRF2005
	−15.7	17.2	8.7	14.4	20.3	−0.9	ITRF2000
	−15.7	17.2	8.6	14.3	20.3	−1.0	EPN
ZIMM 14001M004	−13.0	18.8	10.9	15.1	20.3	0.8	AMON
	−12.8	18.2	10.9	15.0	19.7	0.9	ITRF2005
	−13.8	18.5	10.0	15.1	20.1	−0.4	ITRF2000
	−13.1	18.8	10.6	14.9	20.3	0.5	EPN
HFLK 11006S003	−13.7	18.6	10.3	14.2	20.9	1.0	AMON
	−15.0	18.2	10.6	15.4	20.8	0.3	ITRF2005
	−13.4	18.7	11.1	14.5	21.0	1.8	ITRF2000
	−15.5	18.5	9.4	14.9	21.2	−0.9	EPN
PENC 11206M006	−18.4	17.3	7.3	13.5	22.4	−2.4	AMON
	−18.1	17.3	7.6	13.5	22.3	−2.0	ITRF2005
	−16.6	18.1	8.2	12.7	22.6	−0.4	ITRF2000
	−18.5	17.6	6.9	13.3	22.7	−2.7	EPN
SBGZ 11031S001	−14.9	18.8	9.4	13.9	21.7	0.1	AMON
	9.7	25.8	41.7	16.7	22.9	41.2	ITRF2000
	−15.4	18.6	9.1	14.1	21.6	−0.5	EPN
PFAN 11005S002	−13.3	19.0	10.5	14.4	21.0	1.1	AMON
	−13.7	18.7	11.2	15.2	20.8	1.3	ITRF2000
	−13.1	19.1	10.8	14.4	21.0	1.4	EPN

Table 3: Comparison of station velocities of ITRF2005, ITRF2000 and EPN with AMON

AMON … Results from the OLG network AMON
EPN … Results from the European Reference Frame (EUREF) [6] Permanent Network, Special Project Time Series
ITRF2000 … Results from the International Earth rotation and Reference systems Service (IERS) solution of ITRF2000
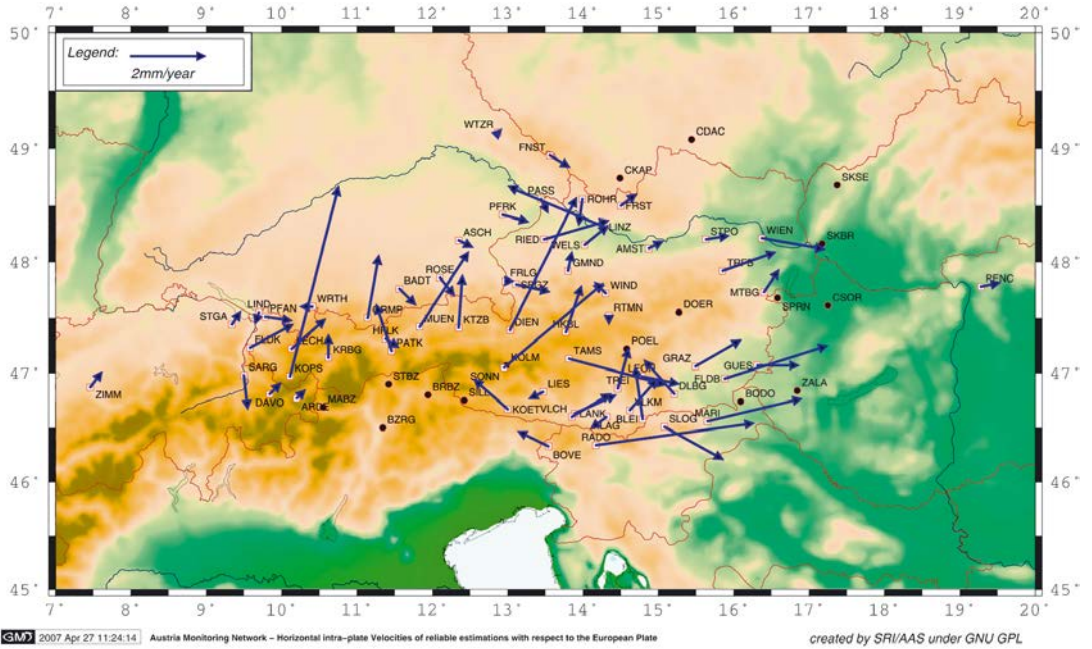ITRF2005 … Results from IERS solution of ITRF2005

Figure 2: Horizontal velocities of reliable estimations (AMON), with respect to the Eurasian Plate
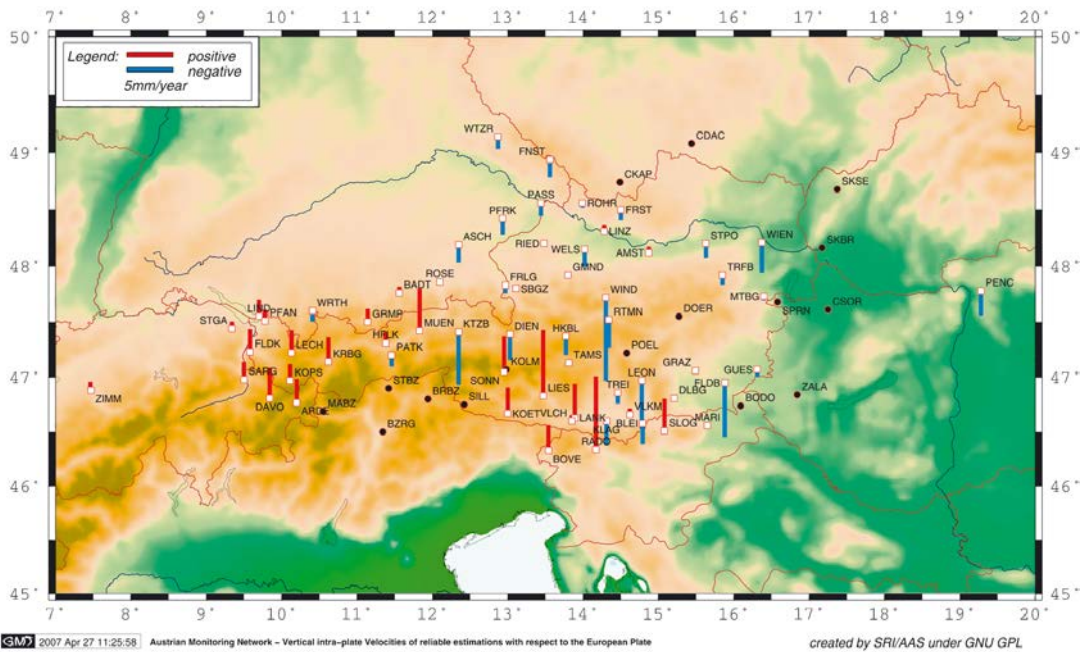

Figure 3: Vertical velocities of reliable estimations (AMON), with respect to the Eurasian Plate

The horizontal velocities in Figure 2 show the smallness of the "intra-plate velocities". The average velocity is about 1mm/year. The large deviations in velocity (>2 mm/year) at the stations DIEN, GRMP, SONN, KOPS, KTZB, LINZ, MUEN, TAMS and WIEN can be considered as local movements, or to due the short time series with offsets, which can not be estimated yet. The movements in the Southeast seem to be a real feature, which might be a structure moving into the Pannonian Basin. The stations in the Northern Alpine Forelands have a similar rotation to the one of the Eurasian plate, but the feature changes in the Alps. Presently, the station coverage is still too poor to discern between different groups. Figure 3 shows the vertical movements where only the small part of the Eurasian movement in ITRF2000 was subtracted. Again some stations show a local behaviour, mainly arising due to antenna changes. Two features can be recognized, a slight rising of the Alpine stations and a constant height or slightly falling in the forelands. Those quite old stations in the West, quasi unaffected by antenna changes, might give the best impression of the rising of the Alps.

5. Conclusions

Presently a large part of the Eastern Alps is already covered by GPS permanent stations. The station velocities can be derived from the time series produced by network adjustments similar to the international standards of EPN. The application of offsets and outliers lead to "cleaned time series" where velocities can be estimated with a residual error of about 1mm/year or better in the lateral components, and 1-3mm/year in the vertical component. The residual velocities with respect to the Eurasian Plate show significant deviations from a unique rotation. Beside individual station problems, which can be solved in the next years, local movements may be discerned from velocity groups. For a national network the general distortion reaches about 10–20 mm per decade over the Austrian territory. Because the local movements can be much higher it is advisable to include at least three stations for transformation in every densification. Nevertheless, even in ETRS89 station coordinates and those derived from them cannot be considered as constant for a generation. There should be a procedure either to apply velocities from a velocity field to every precise network and refer the coordinates to a certain epoch, or at least to apply an official overall change of all coordinates of the Austrian reference network,

for every decade. The geokinematic interpretation is just starting. The next step should be the construction of a velocity field. Subsequently, a stress field could be derived and used for investigations in geodynamics. It should be noted that up to now the geological background of the individual stations has not been investigated and therefore common movements cannot be easily assigned to a geological structure. The seismological experiments CELEBRATION2000 and ALP2002 [8] have provided a huge amount of information related to geological structure. Last but not least, this paper is just a starting point for more up-to-date intra-plate studies over Austria.

References

[1] *Dow, J. M., Neilan, R. E., Gendt, G.*: The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade, Adv. Space Res. 36 vol. 36, no. 3, pp. 320-326, 2005. doi:10.1016/j.asr.2005.05.125.

[2] *Ihde, J., Baker, T., Bruyninx, C., Francis, O., Amalvict, M., Kenyeres, A., Mäkinen, J., Shipman, S., Simek, J. and Wilmes, H. (2005)*: Development of a European Combined Geodetic Network (ECGN). Journal of Geodynamics, n. 40, pp. 450-460

[3] *Altamimi, Z., Sillard, P., Boucher, C. (2002)*: ITRF2000: A New Release of the International Terrestrial Reference Frame, J. Geophys. Res, 2002

[4] *Haslinger, C., Stangl, G. (2006)*: The First Austrian Velocity Field derived from GPS, EUREF Publication No. 16, Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Frankfurt/Main, (in print).

[5] *Hugentobler, U., Dach, R., Fridez, P., Meindl, M. (2006)*: Bernese GPS Software, Version 5.0, Draft

[6] *Adam J., W. Augath, C. Boucher, C. Bruyninx, A. Caporali, E. Gubler, W. Gurtner, H. Habrich, B. Harsson, H. Hornik, J. Ihde, A. Kenyeres, H. van der Marel, H. Seeger, J. Simek, G. Stangl, J. Torres, G. Weber, 2002*: Status of the European Reference Frame – EUREF, International Association of Geodesy Symposia, IAG Scientific Assembly, Springer, ed. J. Adam and K.-P. Schwarz, Vol. 125, pp 42-46

[7] *Blewitt, G., Lavallée, D., (2002)*: Effect of annual signals on geodetic velocity, Journal of Geophysical Research, Vol. 107, No. B7, 10.1029/2001JB000570,2002

[8] *Brückl, E., Mitterbauer, U., Behm, M. and Working Groups CELEBRATION 2000&ALP2002, 2006*: Studies on Crustal Structure and Gravity in the Eastern Alps. In Geodetic Deformation Monitoring: From Geophysical to Engineering Roles. IAG Symposium 131. F. Sanso and A.J. Gil (Eds). Springer, pp.181-192.

[9] http://itrf.ensg.igns.fr/ITRF_solutions/2005 – Webpage of ITRF 2005

[10] http://gps.iwf.oeaw.ac.at/ – Webpage of the Observatory Lustbuehel Graz

[11] http://www.epncb.om.be/_dataproducts/timeseries – Webpage of the European Permanent Network / Special Project / Time Series

[12] http://gmt.soest.hawaii.edu/ – Webpage of Generic Mapping Tools

Contact

Dipl.-Ing. Cornelia Haslinger: Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: cornelia.haslinger@oeaw.ac.at

Dipl.-Ing. Sandro Krauss: Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: sandro.krauss@oeaw.ac.at

Dipl.-Ing. Mag. Dr. phil. Günter Stangl: Federal Office of Metrology and Surveying, Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, E-mail: guenter.stangl@oeaw.ac.at

vgi

Determination of water vapor trends from VLBI observations

Robert Heinkelmann, Michael Schmidt, Johannes Böhm and Harald Schuh

Abstract

Very Long Baseline Interferometry (VLBI) observations can be analyzed to derive precipitable water, which can be an important contribution for meteorological research. Since some of the VLBI stations observe for more than twenty years, it is possible to determine long-term trends. In this work we introduce two methods for the climate trend determination: the robust estimation of a linear trend using the bounded influence by standardized residuals (BIBER) – estimator, and a multi-resolution quadratic normalized B-spline wavelet model for the representation of linear and non-linear trend characteristics. If the trend is modeled by the wavelets instead of a solely linear term, the rms of the residuals becomes significantly smaller. Robust estimated linear trends of water vapor at twelve globally distributed VLBI sites are presented.

Kurzfassung

Beobachtungen der Radiointerferometrie auf langen Basislinien (VLBI) können zur Bestimmung des ausfällbaren Wassers in der Atmosphäre über den Messstationen herangezogen werden. Die Zeitreihen können einen wichtigen Beitrag für die meteorologische Forschung leisten. Da von manchen VLBI Stationen Beobachtungsreihen über mehr als zwanzig Jahre vorliegen, können langfristige Trends berechnet werden. In dieser Studie werden zwei Methoden zur Trendbestimmung untersucht: Zum einen werden lineare Trendanteile mit dem BIBER-Schätzer (bounded influence by standardized residuals) bestimmt, zum anderen werden lineare und nicht-lineare Eigenschaften der Trendkomponente mit normierten, quadratischen B-spline Wavelets dargestellt. Wird der Trendanteil im Gegensatz zu einem linearen Term durch die Wavelets beschrieben, so sind die Standardabweichungen der Residuen signifikant kleiner. Robust geschätzte lineare Trends an zwölf global verteilten VLBI Stationen werden präsentiert.

1. Introduction

Water vapor in the atmosphere is an important storage and energy transfer medium of the global water household, influencing Earth's climate in many ways. Among the greenhouse gases water vapor holds an important and difficult status: On the one hand it enforces global warming acting as a filter of Earth's long wave radiation (albedo). On the other hand it can reduce the absorption of solar energy in form of clouds. The spatio-temporal distribution of water vapor in the troposphere (the neutral atmosphere) can be subject to rapid variations, and therefore, it is difficult to be measured and modeled. Vertical profiles of the partial pressure of water vapor in the troposphere can sparsely be measured by radiosondes. Observational techniques such as water vapor radiometer or solar spectrometer detect the characteristic absorption of water vapor, but are bound to good weather conditions, and the orbit of the sun, respectively. Therefore, we agree with Hagemann et al. [2003], that present observing systems are inadequate to monitor water vapor and its spatio-temporal distribution properly. Since it is questionable

whether climate trends can be calculated from reanalysis data [Bengtsson et al., 2004], such as the 40-years re-analysis product (ERA-40) of the European Centre of Medium-Range Weather Forecasts (ECMWF), the space-geodetic techniques, in particular Very Long Baseline Interferometry (VLBI), can give an important contribution to meteorological research.

In Section 2 we give a general introduction on the determination of water vapor.

Then we discuss our approach for the determination of trends and other signal components in Section 3. Due to VLBI's organizational and observational characteristics its time-series are usually irregularly sampled and clumped. Possible causes of a trend in time-series of precipitable water (PW) are considered and investigated in Section 4. Zenith wet delays (ZWD) for this study are taken from long time-series determined by several Analysis Centers (AC) of the International VLBI Service for Geodesy and Astrometry (IVS) [Schlüter et al., 2002] combined on the level of parameter estimates [Heinkelmann et al., 2007].

2. Determination of water vapor by space-geodetic techniques

The determination of water vapor by space-geodetic techniques operating on radio wave-lengths, e.g. VLBI and Global Navigation Satellite Systems (GNSS), is based on the excess delay of radio signals traveling through the wet tropospheric medium and through the corresponding hydrostatic medium

$$\Delta L(E, A) = mf_h(E) \cdot ZHD + mf_w(E) \cdot ZWD+$$
$$+mf_g cot(E)[G_N \cos(A) + G_E \sin(A)] \tag{1}$$

depending on elevation E (deg) and azimuth A (deg) angles of an observation. The total tropospheric delay in the line of sight ΔL(m) completely describes the effects of the tropospheric medium on VLBI observables. In Equation (1), the hydrostatic and wet mapping functions mf_h, mf_w are assumed to be known and the gradient mapping function mf_g is either the wet or the hydrostatic mapping function [MacMillan, 1995]. While ZWD(m) and the gradient components in north G_N(m) and east direction G_E(m) are estimated [Koch, 1997], zenith hydrostatic delays ZHD (m) can be computed by Davis et al. [1985]

$$ZHD = \frac{0,0022768\,p}{1\,-\,0,00266\cos(2\phi)\,-\,0,00028h} \tag{2}$$

and are typically considered as corrections to the vector of observations. In Equation (2) the atmospheric surface pressure p (hPa), the latitude ϕ and height h(km) refer to the antenna reference point (ARP), i.e. the intersection of antenna axes of the VLBI telescope, or the antenna's phase center in the case of GNSS.

Following Bevis et al. [1993] the vertically integrated water vapor above the ARP can be expressed in terms of precipitable water PW(m) depending on the ZWD

$$PW = \Pi \cdot ZWD \tag{3}$$

where the factor Π is given by

$$\Pi = \frac{10^6}{\rho_l \,\cdot\, R_v[(k_3/T_m)\,+\,k_2']} \tag{4}$$

In Equation (4) $\rho_l = 998.2$ (kg m^{-3}) denotes the temperature-dependent density of liquid water at 20 (°C) and $R_v = 461.525$ (J kg^{-1} K^{-1})

denotes the specific gas constant of water vapor. The variables k_2' and k_3 are refractivity constants and T_m is a weighted mean temperature of the atmosphere above the ARP. The correlation between T_m and the measured surface temperature T (K) was studied by Bevis et al. [1993], who analyzed two years of radiosonde observations over the United States territory. They found a linear relation: $T_m = 70.2 + 0.72T$.

3. Assessment of climate signals

In general, a climate signal y_t can be considered as linear combination of a constant mean, a trend component, one or more cycles, extreme events or outliers, and a noise component

$$y_t = mean + trend + cycle + extrema + noise \tag{5}$$

While the cyclic component consists of those parts of the signal, which are reproduced during certain periods, the trend component describes the temporal change of the signal in a non-cyclic sense. In general the trend component is not linear. Outliers can be identified either manually or automatically by applying robust estimation. The noise component is characterized by a zero expectation and a constant standard deviation. In our analysis we compare approximations of climate signals by two approaches: First we calculate and remove the mean value of the signal. Then we identify the periods of cycles by spectral analysis and determine the corresponding amplitudes and phases of sinusoids. In the sequel, we approximate the trend component by a simple linear term determined by robust regression. Alternatively we describe linear and non-linear characteristics of the trend by wavelet modeling.

3.1. Fourier expansion of cyclic components in $\mathbf{ZWD}$ time-series

Fig. 1 shows the spectra of ZWD at Gilmore Creek determined by three methods: the discrete Fourier transform, the Lomb-Scargle periodogram [Lomb, 1976], and the 'CLEANed-spectrum' obtained by the CLEAN-algorithm [Roberts et al., 1987]. The discrete Fourier spectrum shows peaks at periods >1 year, which are not significant in the spectra derived with the two other methods. Since the Lomb-Scargle periodogram and in particular the CLEAN-algorithm are designed to operate on unequally spaced data, we consider those periods to be artifacts due to the irregular sampling of the time-series. The discrete Fourier spectrum in addition shows some power along the shorter Fourier-periods (1/3 year, 1/4 year, etc.).
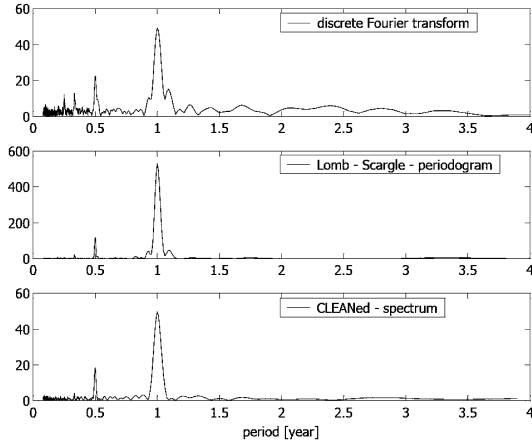
Figure 1: Three spectra of zenith wet delays (ZWD) at GILCREEK (Fairbanks, Alaska, USA) obtained by discrete Fourier transformation (top), Lomb-Scargle periodogram (mid), and CLEAN-algorithm (bottom). The abscissae give the amplitude (mm), or the power (mm²) in case of the Lomb-Scargle periodogram, and the ordinates the periods (years).

These shorter periods have no physical meaning, but are no artifacts as well. They emerge from the non-sinusoidal shape of the cyclic component. As a consequence annual, semi-annual and the next six periods are considered for the cyclic component.

$$y_t - mean = cycle + w_t$$

$$cycle = \sum_{i=1}^{8} [A_i \sin(2\pi f_i t) + B_i \cos(2\pi f_i t)] \qquad (6)$$

In Equation (6) A_i and B_i are the $i = 8$ pairs of coefficients characterizing amplitudes and phases of sinusoids with frequencies f_i corresponding to annual, semi-annual and 6 successive periods. The variable wt denotes a residual signal, composed of $trend$, $extrema$, and $noise$ components. After the f_i are identified from the spectra, the coefficients A_i and B_i of Equation (6) are determined by least-squares method; see e.g. Koch [1997].

3.2. Determination of linear trend by robust regression

After the cyclic component (Section 3.1) has been identified and removed, a linear trend can be estimated from the residual signal w_t, e.g. by robust estimation using the BIBER-estimator [Wicki, 2001]. However, in general, a climate trend is not fully described by a linear term and thus, it should not be modeled by a linear term only.

3.3. Trend approximation using normalized quadratic B-splines

Linear and non-linear characteristics of a climate trend can be assessed and represented by wavelet modeling. Therefore, we apply the normalized quadratic B-spline function as one dimensional scaling function. The normalized B-spline $N_{j,k}^d(x)$ of order d is defined recursively at equally spaced knots $t_0^j, t_1^j, \ldots, t_{m_j+d}^j$ and represented in levels $j = 0, \ldots, J$ of different resolutions [Schmidt, 2007] with $k = 0, \ldots, m_j - 1$ and $m = 1, \ldots, d$ as

$$N_{j,k}^m(x) = \frac{x - t_k^j}{t_{k+m}^j - t_k^j} N_{j,k}^{m-1}(x) + \frac{t_{k+m+1}^j - x}{t_{k+m+1}^j - t_{k+1}^j} N_{j,k+1}^{m-1}(x) \qquad (7)$$

and with initial values

$$N_{j,k}^0(x) = \begin{cases} 1 & if \quad t_k^j \le x < t_{k+1}^j \\ 0 & else \end{cases}; \qquad (8)$$

see e.g. Stollnitz et al. [1995]. In general, the B-spline of order d is compactly supported, i.e. its values are different from zero only in a finite range on the real axis. Since we want to use this approach for regional modeling we introduce the so-called endpoint-interpolated B-splines of order d defined on the unit interval $[0, 1[$.

For our investigations we choose with $d = 2$ the normalized quadratic B-spline functions $N_{jk}^2(x)$. In this case the knot sequence is given as

$$0 = t_0^j = t_1^j = t_2^j < t_3^j < t_4^j < \ldots < t_{m_j}^j = t_{m_{j+1}}^j = t_{m_{j+2}}^j = 1 \qquad (9)$$

with $m_j = 2^j + 2$. Fig. 2 shows the normalized quadratic B-spline, for resolution level $j = 2$.
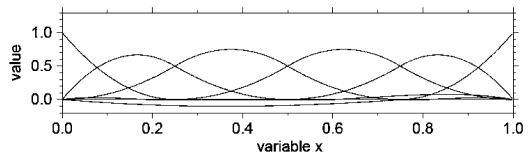
Figure 2: Normalized quadratic B-splines of resolution level .

Coefficients at the highest resolution level J are estimated by least-squares method, e.g. Koch [1997]. Then the coefficients related to the lower resolution levels $0 \le j < J$ can successively be expressed by linear combinations starting from the highest level J. This so-called pyramid algorithm is the basic tool of the decomposition of the input signal into frequency-dependent detail signals, which is known as the multi-

resolution representation; see Schmidt [2007]. The level-j approximation means the sum of the detail signals until resolution level j. The highest resolution level J is implicitly given such that the relation $m_j = 2^J + 2 < n$ holds, where n denotes the number of observations.

When neglecting non-significant coefficients, the number of coefficients used for the representation of signals is reduced significantly. This can be interpreted as statistical data compression. The significance of the wavelet-coefficients can be assessed e.g. by hypothesis testing.

3.4. Comparison of the two approaches for the trend approximation
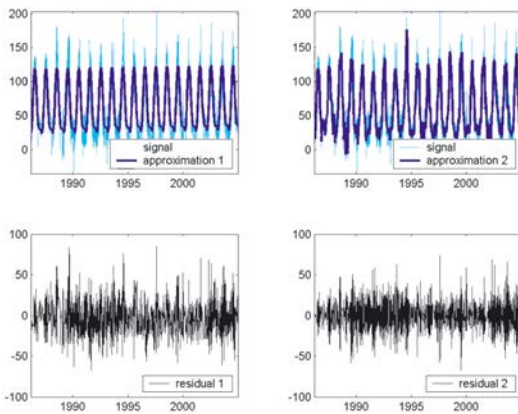
Figure 3: Zenith wet delays (ZWD) from International VLBI Service for Geodesy and Astrometry (IVS) combined series at GILCREEK (Fairbanks, Alaska, USA) between 1986.0 and 2005.0 (cyan). In blue approximation (1) by sinusoids, a linear and a constant term on the left ($\sigma = 20.7$ mm), as well as approximation (2) by sinusoids, B-spline wavelet modeling and a constant term on the right ($\sigma = 17.0$ mm). The corresponding residual signals are displayed in black (bottom). The abscissae give the ZWD (mm) and the ordinates the time (years).

Figure (3) exemplarily illustrates the ZWD for station Gilmore Creek in cyan (top). The blue curves display the two different approximations: On the left the trend component is modeled by a linear term only, while at the right side the trend component is represented by normalized quadratic B-splines until level $J = 7$. Cycles are considered by a Fourier expansion for both approaches. The rms of the noise component by modeling a linear trend only is 20.7 mm, whereas wavelet modeling yields a significantly smaller rms of 17.0 mm. Choosing a higher value for J the approximation can be improved.

4. Analysis of the trend component

Reconsidering Equations (1–7) a trend in time-series of PW will mainly be due to a trend of inherent ZWD. In addition, variations of wet and hydrostatic mapping functions and of ZHD, which mainly depend on the atmospheric pressure p, can affect the trend of PW. The temporal change of the proportionality factor Π due to the mean atmosphere temperature T_m and the density of liquid water ρ_1 is of negligible size, i.e., the proportionality factor Π shows no significant variations itself. Trends in north and east gradients are typically of negligible size and do not have to be considered for the trend analysis of ZWD, or PW, respectively.

4.1. Atmospheric pressure

The atmospheric pressure is the mayor quantity for the determination of ZHD. Since ZHD are subtracted from the observations before the estimation of ZWD, inherent trends will propagate to the ZWD via the mapping functions. For most of the VLBI sites atmospheric pressure is automatically recorded in-situ, with a specific temporal resolution, and provided at the epoch of each observation. In-situ measurements are the source with the highest resolution. If outliers and missing values are appropriately replaced and significant shifts of the running mean are eliminated, in-situ atmospheric pressures are the most reliable pressure data. If in-situ atmospheric pressure data are not available, values should be taken from a numerical weather model (NWM). E.g. ERA-40, or operational data-sets of the ECMWF. If NWM are unavailable, we suggest the empirical global temperature and pressure (GPT) model [Böhm et al., 2007]. For the homogenization of in-situ atmospheric pressures as well as for the replacement of outliers and data gaps, the use of pressure time-series derived from NWM, such as the ERA-40, is very valuable and suggested by the authors. Significant shifts of the running mean of atmospheric pressure time-series recorded at VLBI stations [Heinkelmann et al., 2005] can be found by the application of a standard normal homogeneity test (SNHT) [e.g. Tuomenvirta and Alexandersson, 1996]. Shifts of the running mean of pressure time-series can be e.g. due to calibration, replacement, or relocation, in particular in height, of a pressure sensor. Figure (4) shows homogenized pressure time-series for station GILCREEK (Fairbanks, Alaska, USA) on the bottom, as well as ZHD on the top in cyan. An approximation by wavelet modeling is displayed in blue.
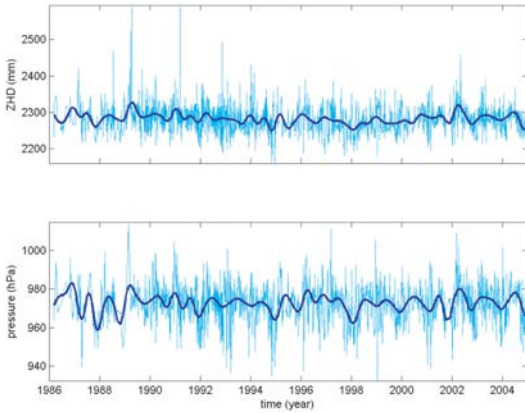
Figure 4: ZHD (top) from International VLBI Service for Geodesy and Astrometry (IVS) combined series at Gilmore Creek (Fairbanks, Alaska, USA) between 1986.0 and 2005.0 in mm and atmosphere pressure (bottom) in hPa are plotted in cyan. Approximations are displayed in blue.
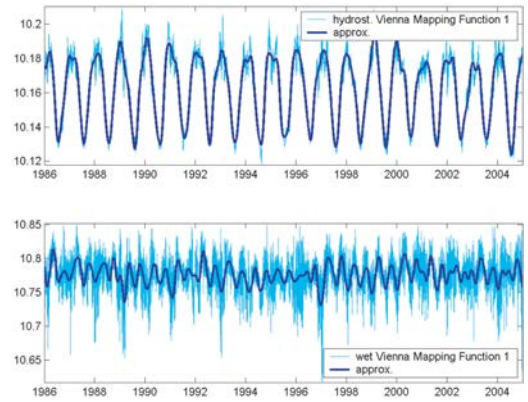
Figure 5: Hydrostatic (top) and wet (bottom) Vienna Mapping Functions 1 (VMF1) at Gilmore Creek (Fairbanks, Alaska, USA) for elevation angle E (deg) in cyan as well as approximations in blue. The abscissae give the unitless values of the mapping functions and the ordinates the time (years).

4.2. Mapping functions

Trends in mapping functions directly and indirectly affect the trend of ZWD. The wet mapping function is the partial derivative of the observation with regards to ZWD. In addition the interdependency between ZWD, ZHD and the two gradient components are governed by the wet and hydrostatic mapping functions.

Various authors provide mapping functions by time-series of coefficients for particular stations or in form of global grids. Currently, the most accurate mapping functions globally available are the Vienna Mapping Functions 1 (VMF1) [Böhm et al., 2006]

$$mf(E) = \frac{1 + \dfrac{a}{1 + \dfrac{b}{1+c}}}{\sin(E) + \dfrac{a}{\sin(E) + \dfrac{b}{\sin(E)+c}}} \qquad (10)$$

Here, wet and hydrostatic mapping functions mf are given by three coefficients a, b, c. The coefficients a are determined from raytracing through ECMWF data and are provided with a temporal resolution of $t = 6$ (hours). The coefficients b and c are calculated from empirical equations. The mapping functions depend on the elevation angle E of on observation and the latitude of a station. Figure (5) exemplarily displays the VMF1 at Gilmore Creek for an elevation angle of $E = 5$ (deg).

5. Results and discussion

VLBI is able to provide time-series of ZWD, which can be transformed to PW, i.e., high-quality information about the amount of water vapor above the VLBI stations.

If time-series of ZWD, or PW, are interpreted in terms of climate signals, the atmospheric pressure and mapping function time-series used for the determination of ZWD have to be considered. In particular the time-series of atmospheric pressure need to be homogenized, i.e. shifts of the running mean value due to sensor relocation, replacement, etc. need to be identified and removed. Trends in atmospheric pressure and mapping function time-series can cause apparent trends in ZWD (Equations 1 and 2).

Table 1 gives the linear trends determined by robust estimation (Section 3.2) for the twelve most frequently observing VLBI sites [Behrend and Baver, 2005]. At some of those VLBI sites ZWD are available for more than 20 years. Thus, due to the large number of inherent ZWD estimates the formal errors of the linear trends are usually very small (see Table 1). However, the noise component dominates the climate signal. E.g. for the observed trend of 0.29 (mm/year) and the standard deviation of the noise component of 20.7 (mm) at station GILCREEK, the trend to noise ratio of ~ 0.014 indicates a very low significance of the trend. Instead of the formal errors of the linear trends, we suggest to consider the 1σ-level of the noise, to assess the significance of the trend.

station	time span	lin. trend (mm/year)	formal error (mm/year)	1-σ of the noise (mm)
ALGOPARK	1995–2005	−0.24	0.09	39.5
FORTLEZA	1995–2005	0.99	0.06	35.7
GILCREEK	1986–2006	0.29	0.01	20.7
HARTRAO	1993–2005	0.63	0.06	34.4
HOBART26	1993–2005	−0.31	0.06	31.1
KOKEE	1993–2005	0.78	0.04	33.7
MATERA	1993–2005	0.48	0.04	26.2
NYALES20	1995–2005	1.61	0.02	17.4
SESHAN25	1990–2005	−1.95	0.31	57.7
TSUKUB32	1997–2005	−2.33	0.22	51.2
WESTFORD	1986–2006	−0.15	0.07	47.4
WETTZELL	1986–2006	−0.07	0.02	20.2

Table 1: Linear trend estimates (mm/year) from International VLBI Service for Geodesy and Astrometry (IVS) combined series [Heinkelmann et al., 2007]

Comparison of linear trend estimates [Heinkelmann et al., 2007] shows better agreement of linear trends in case of synchronized data. This effect is probably also due to the large noise component.

In comparison to the linear trend model, the normalized quadratic B-spline wavelets appeared to be more capable of representing linear and nonlinear characteristics of the trend component. E.g. at Gilmore Creek the sigma of the residuals drops from 20.7 (mm) to 17.0 (mm), if the trend component is modeled by a quadratic normalized B-spline of level $J = 7$, instead of a linear term only. In particular inter-annual variations of the amplitude of the seasonal signal are much better described by the wavelet model. In our example the variations of the annual amplitude of the seasonal signal at Gilmore Creek exceed the size of the linear trend by 1 to 2 orders of magnitude.

Acknowledgements

References

[1] *Behrend D. and Baver K. (editors) [2005]*: International VLBI Service for Geodesy and Astrometry. IVS Annual Report 2004, NASA/TP-2005-212772, NASA, Greenbelt

[2] *Bengtsson L., Hagemann S., Hodges K. I [2004]*: Can climate trends be calculated from reanalysis data? Journal of Geophysical Research. Vol. 109(D11111), doi:10.1029/2004JD004536

[3] *Bevis M., Businger S., Chiswell S., Herring T. A., Anthes R. A., Rocken C., Ware R. H. [1993]*: GPS meteorology: mapping zenith wet delays onto precipitable water. Journal of Applied Meteorology. Vol. 33, 379-386

[4] *Böhm J., Heinkelmann R., Schuh H. [2007]*: Short Note: A global model of pressure and temperature for geodetic application. Journal of Geodesy. doi:10.1007/s00190-007-0135-3

[5] *Böhm J., Werl B., Schuh H. [2006]*: Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. Journal of Geophysical Research. Vol. 111, doi:10.1029/2005JB003629

[6] *Davis J. L., Herring T. A., Shapiro I. I., Rogers A. E. E., Elgered G. [1985]*: Geodesy by radio interferometry: Effects of atmospheric modelling errors on estimates of baseline length. Radio Science. Vol. 20(6), 1593-1607

[7] *Hagemann S., Bengtsson L., Gendt G. [2003]*: On the determination of atmospheric water vapor from GPS measurements. Journal of Geophysical Research. Vol. 108(D21), doi:10.1029/2002JD003235

[8] *Heinkelmann R., Böhm J., Schuh H. [2005]*: Homogenization of surface pressure recordings and its impact on long-term series of VLBI tropospheric parameters. In: Proceedings of the 17th Working Meeting on European VLBI for Geodesy and Astrometry, Vennebusch M., Nothnagel A. (editors), INAF–Instituto di Radioastronomia–Sezione di NOTO, Italy, 74-78

[9] *Heinkelmann R., Böhm J., Schuh H., Bolotin S., Engelhardt G., MacMillan D. S., Negusini M., Skurikhina E., Tesmer V., Titov O. [2007]*: Combination of long time-series of troposphere zenith delays observed by VLBI. Journal of Geodesy. doi:10.1007/s00190-007-0147-z

[10] *Koch K. R. [1997]*: Parameterschätzung und Hypothesentests. 3rd edition. Dümmler, Bonn, 368

[11] *Lomb N. R. [1976]*: Least-squares frequency analysis of unequally spaced data, Astrophysics and Space Science. Vol. 39, 447-462

[12] *MacMillan D. S. [1995]*: Atmospheric gradients from very long baseline interferometry observations. Geophysical Research Letters. Vol. 22(9), 1041-1044

[13] *Roberts D. H., Lehar J., Dreher J. W. [1987]*: Time series analysis with CLEAN. I. Derivation of a spectrum. Astronomical Journal. Vol. 93(4), 968-989

[14] *Schlüter W., Himwich E., Nothnagel A., Vandenberg N., Whitney A. [2002]*: IVS and its important role in the maintenance of the global reference systems. Advances in Space Research. Vol. 30(2), 145-150

[15] *Schmidt, M. [2007]*: Wavelet modelling in support of IRI. Advances in Space Research. doi:10.1016/j.asr.2006.09.030

[16] *Stollnitz E. J., DeRose T. D., Salesin D. H. [1995]*: Wavelets for Computer Graphics: A Primer. Institute of Electrical and Electronics Engineering - Computer Graphics and Applications. Vol. 15(3), 76-84 (part 1), and Vol. 15(4), 75-85 (part 2)

[17] *Tuomenvirta H., Alexandersson H. [1996]*: Review on the methodology of the standard normal homogeneity test (SNHT). In: Proceedings of the 1st seminar for homogenization of surface climatological data, Hungarian Meteorological Service (editors), 35-45

[18] *Wicki F. [2001]*: Robust estimator for the adjustment of geodetic networks. In: Proceedings of the 1st international symposium on robust statistics and fuzzy techniques in geodesy and GIS. Carosio A. and Kutterer H. (editors), International Association of Geodesy – Special Study Group 4.190. Non-probabilistic assessment in geodetic data analysis, Swiss Federal Institute of Technology Zurich (ETH). Institute of Geodesy and Photogrammetry. IGP – Bericht Nr. 295, 53-60

Contact

Dipl.-Ing. Robert Heinkelmann, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, A – 1040 Vienna, Austria.
E-mail: rob@mars.hg.tuwien.ac.at

PD Dr.-Ing. habil. Michael Schmidt, Deutsches Geodätisches Forschungsinstitut, Alfons-Goppel-Str. 11, D – 80539 Munich, Germany. E-mail: schmidt@dgfi.badw.de

Dr. Johannes Böhm, Institute of Geodesy and Geophysics, Vienna University of Technology. 27-29, A – 1040 Vienna, Austria.
E-mail: johannes.boehm@tuwien.ac.at

Prof. Dr. Harald Schuh, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, A – 1040 Vienna, Austria.
E-mail: harald.schuh@tuwien.ac.at

vgi

Global models of the ionosphere obtained by integration of GNSS and satellite altimetry data

Sonya Todorova, Harald Schuh, Thomas Hobiger, Manuel Hernández-Pajares

Abstract

The high free-electron and ion density in the ionosphere disturbs both the group and phase velocity of the signals of all space geodetic techniques, operating in the microwave band. In first approximation this delay is proportional to the so-called Slant Total Electron Content (STEC) along the ray path and can be corrected only if the measurements are carried out at two distinct frequencies. On the other hand, this effect allows information to be gained about the parameters of the ionosphere in terms of Total Electron Content (TEC) values. The classical input data for the development of Global Ionosphere Maps (GIM) of the total electron content is obtained from dual-frequency Global Navigation Satellite System (GNSS) observations. However, the GNSS stations are inhomogeneously distributed, with large gaps particularly over the sea surface, which lowers the precision of the GIM over these areas. On their part, dual-frequency satellite altimetry missions such as Jason-1 provide information about the ionosphere precisely above the sea surface. Due to the limited spread of the measurements and some open questions related to their systematic errors, the ionospheric data from satellite altimetry is used only for cross-validation of the GNSS GIM so far. It can be anticipated however, that some specifics of the ionosphere parameters derived by satellite altimetry will partly balance the inhomogeneity of the GNSS data. In this study we create two-hourly GIM from GNSS data and additionally introduce satellite altimetry observations, which help to compensate the insufficient GNSS coverage of the oceans. Furthermore, this method allows the independent estimation of systematic instrumental errors, affecting the two types of measurements. Thus, besides the daily values of the Differential Code Biases (DCB) for all GNSS satellites and receivers, also a constant daily bias for the Jason-1 satellite is estimated and investigated.

Kurzfassung

Durch die hohe Dichte von freien Ionen und Elektronen in der Ionosphäre werden die Beobachtungen aller geodätischen Weltraumverfahren, die im Mikrowellenbereich operieren, verzögert. Die Laufzeitverzögerung der Beobachtungsstrahlen ist in erster Näherung proportional zum so genannten Gesamtelektronengehalt entlang des Strahlenwegs (Slant Total Electron Content, STEC). Dieser Effekt kann nur dann korrigiert werden, wenn die Messungen auf zwei verschiedenen Frequenzen erfolgen. Auf diese Weise lässt sich aber auch Information über die Ionosphärenparameter in Form von TEC-Werten gewinnen. Die klassischen Eingabedaten für die Entwicklung globaler Karten der Ionosphäre (Global Ionosphere Maps, GIM) sind Zweifrequenzbeobachtungen des Globalen Satellitennavigationssystems (Global Navigation Satellite System, GNSS). Die GNSS-Stationen sind jedoch nicht homogen auf der Erde verteilt, wobei vor allem die Meeresoberfläche schlecht abgedeckt ist. Andererseits liefern die Zweifrequenz-Messungen von Satellitenaltimetrie Missionen wie Jason-1 Information für die Ionosphärenparameter genau über den Ozeanen. Aufgrund der begrenzten Verteilung dieser Messungen, sowie einiger offenen Fragen bezüglich der systematischen Fehler, werden die Altimetrie Daten derzeit nur zur Validierung der GNSS GIM genutzt. Man kann jedoch annehmen, dass gewisse Besonderheiten der Ionosphärenparameter, die von Satellitenaltimetrie-Messungen erhalten werden, die Inkonsistenzen der GNSS Beobachtungen ausgleichen können. In dieser Studie werden für die Erzeugung globaler Ionosphärenkarten in zweistündigen Intervallen neben GNSS auch Messungen aus Satellitenaltimetrie herangezogen, deren Verteilung die mangelhafte GNSS-Abdeckung der Meeresoberfläche auszugleichen hilft. Außerdem erlaubt diese Methode die unabhängige Schätzung von systematischen, technik-spezifischen Fehlern. Deshalb wird neben den täglichen Werten der instrumentellen Einflüsse (Differential Code Biases, DCB) aller GNSS Satelliten und Empfänger auch ein konstanter täglicher Jason-1 Messfehler geschätzt und untersucht.

1. The ionosphere and its impact on space geodetic techniques

The Earth's ionosphere is defined as that part of the upper atmosphere where the density of free electrons and ions is high enough to influence the propagation of electromagnetic radio frequency waves [9]. The boundaries of this area are roughly set between 50 and 1000 km above the Earth's surface. The ionisation process is primarily driven by the Sun's activity and varies strongly with time, as well as with geographical location. The diurnal

maximum of free electrons and ions occurs around local noon; as for the spatial variations, the low latitude and equatorial regions are characterized by stronger ionization than the high latitudes.

When electromagnetic waves travel through the ionosphere, the integration between the electromagnetic field and the free electrons influences both the speed and the propagation direction of the signals. This effect is known as ionospheric refraction [10] and has to be considered in the determination of the propagation velocity of the signals of all space geodetic techniques operating with electromagnetic waves (Fig. 1).
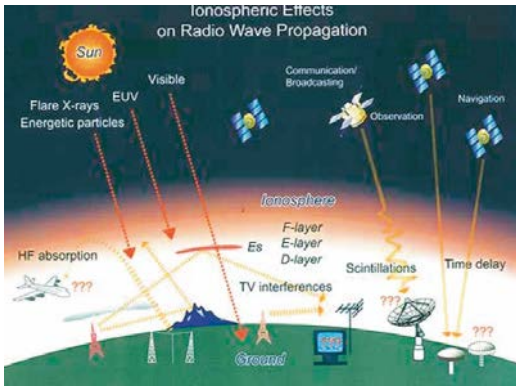
Figure 1: Ionospheric effects on radio wave propagation (http://www2.nict.go.jp/)

The ionospheric refraction disturbs the group and phase velocity of the signals by the same amount but with different sign. The effect of the refraction can be determined in terms of STEC (Slant Total Electron Content), which is the integral of the electron density along the signal path S - see equations (1) and (2). This quantity represents the total amount of free electrons in a cylinder with a cross section of 1m^2 and the slant signal path as axis. STEC is measured in Total Electron Content Units (TECU), with 1 TECU equivalent to 10^{16} electrons/m^2. Equations (1) and (2) express the effect in meters of the ionised medium on phase and group signal propagation:

$$I_{phase} = \int_S \left[\left(1 - \frac{40.28 N_e}{f^2} \right) - 1 \right] dS = -\frac{40.28}{f^2} \int_S N_e \, dS =$$
$$= -\frac{40.28 \cdot 10^{16}}{f^2} STEC, \qquad (1)$$

$$I_{group} = \int_S \left[\left(1 + \frac{40.28 N_e}{f^2} \right) - 1 \right] dS = \frac{40.28}{f^2} \int_S N_e \, dS =$$
$$= \frac{40.28 \cdot 10^{16}}{f^2} STEC, \qquad (2)$$

f ... carrier frequency in Hz,
N_e ... free electron density in the medium,
S ... slant signal path.

The measurements of nearly all space geodetic techniques operating with electromagnetic waves are carried out at two different radio frequencies, which allows the ionospheric influence to be eliminated by linear combinations of the observations. On the other hand, in this way information about the ionosphere parameters can be obtained. If the behaviour of the ionosphere is known, the ionospheric refraction can be computed via equations (1) and (2) and used to correct single-frequency measurements.

1.1. Probing the ionosphere through GNSS

The Global Navigation Satellite System (GNSS), presently consisting on GPS (Global Positioning System) and GLONASS (GLObal NAvigation Satellites System), provides information about the ionospheric refraction, enabling high resolution ionosphere imaging in longitude, latitude, and time (e.g. [2], [17], and references therein). Both observables of the system – carrier phase and code measurements – are affected by the ionosphere. According to equations (1) and (2), this effect depends on the signal frequency f and on the STEC between the satellite and the receiver. Thus, forming the so-called geometry-free linear combination by subtracting simultaneous observations at the two different frequencies L1 and L2, and in this way removing all frequency-independent effects (such as clock errors, troposphere delay etc.), leads to an observable, which contains only the ionospheric refraction and the inter-frequency hardware biases Δbk and Δb_i (usually in ns), associated with the satellite k and the receiver i. In this work carrier phase smoothed code observations are used and the ionospheric observable reads as:

$$\Phi_{i,4}^k = \Phi_{i,1}^k - \Phi_{1,2}^k = -\xi_4 \alpha I_i^k + c(\Delta b^k - \Delta b_i), \qquad (3)$$

$\Phi_{i,1}^k, \ \Phi_{1,2}^k$ ... carrier phase observations at the two frequencies, corrected by the carrier phase ambiguities,

I_i^k ... the ionospheric refraction between the satellite and the receiver related to L1 (in meters),

$\xi_4 = 1 - f_1^2/f_2^2$ ... factor for relating the ionospheric refraction on L4 to L1,

α ... constant used to convert meters into TECU.

The ionospheric parameters derived from the geometry-free linear combination are affected by inter-frequency hardware biases (e.g. [15]), also called Differential Code Biases (DCB); when modelling the ionosphere it is necessary to estimate them as additional unknowns.

In 1998 a special Ionosphere Working Group of the International GNSS Service (IGS) was initiated for developing global ionospheric TEC maps ([8] and [12]). Up to now, four Analysis Centres (AC) - Centre for Orbit Determination in Europe (CODE) [13] and [17], European Space Agency (ESA) [7], Jet Propulsion Laboratory (JPL) [15], and Universidad Politécnica de Cataluña (UPC) [11], deliver daily global maps of vertical TEC and DCB values in the IONospheric EXchange (IONEX) format [16] by using different estimation methods. Since the end of 2005 a combined IGS solution is also available.

1.2. Ionosphere parameters provided by satellite altimetry

Satellite altimetry missions with double-frequency radar altimeter on-board, such as TOPEX/ Poseidon (T/P) and Jason-1 provide information about the ionosphere in the form of Vertical Total Electron Content (VTEC). The T/P mission was launched in August 1992 for observing the ocean circulation and was operational till October 2005. Jason-1, launched in December 2001, is the follow-on to T/P and has inherited its main features – orbit, instruments, measurement accuracy, etc. The orbit altitude of the two missions is 1336 km. The primary sensor of both T/P and Jason-1 is the NASA Radar Altimeter operating at 13.6 GHz (Ku-band) and 5.3 GHz (C-band), simultaneously. The two widely separated frequencies allow TEC to be detected directly from the nadir altimetry sampling data along the satellite track [14].

1.3. TEC derived from GNSS and satellite altimetry - key issues

The TEC estimates from GNSS and from satellite altimetry measurements have often been compared in order to asses the precision of the two techniques (e.g. [3], and references therein). Generally, the agreement between GNSS and altimetry derived TEC is good, but there are still some contradictions, which need further investigation. One important topic is the better understanding of the frequency-dependent systematic errors in the altimetry measurements, which would bias both the sea-level height and the TEC estimates [4]. The TEC values obtained by satellite altimetry are expected to be lower than the ones coming from GNSS because opposite to GNSS, the altimetry satellites do not sample the topside ionosphere due to their lower orbit altitude. However, several studies have shown that T/P and Jason-1 systematically overestimate the vertical TEC by about 3–4 TECU compared to the values delivered by GNSS (i.e. [1] and [3]). On the other hand, most of the ionosphere models from GNSS data are based on the Single Layer Model (SLM, described in section 2.), which does not account well for the ionospheric contribution above the altitude of the altimetry missions [3]. Furthermore, it has to be pointed out that when using SLM the STEC values derived from GNSS measurements have to be converted into vertical TEC (VTEC), while the altimetry missions deliver directly the vertical values. The mapping function (see equation (4)) used for this convertion is a potential error source for the GNSS TEC estimates. Finally, for comparing with altimetry TEC, the values derived from GNSS have to be interpolated for regions far from the observing stations – that is above the oceans, i.e. such comparisons are performed in the worst scenario for GNSS.

The differences between GNSS and altimetry derived TEC as well as the systematic errors of the Jason-1 satellite are treated in more detail in subsection 3.2.

2. Development of 3D global ionosphere maps

The global maps created in this study represent the ionosphere in longitude, latitude and time and are based on the Single Layer Model (SLM). SLM assumes that all free electrons are concentrated in an infinitesimally thin layer above the Earth's surface. The height H of this thin shell is usually set at the height, where the highest electron density is expected, which is between 300 and 450 km. A signal transmitted from the satellite to the receiver crosses the ionospheric shell in the so-called ionospheric pierce point. The zenith angle at that point is z' and the signal arrives at the ground station with zenith distance z. The relation between the measured slant TEC along the ray path and the vertical value at the pierce point is given by a mapping function (4). In this study the Modified Single Layer Model (MSLM) was adopted (as in [5], and [6]) and the mapping function for the transformation between STEC and VTEC reads as:

$$F(z) = \frac{STEC}{VTEC} = \frac{1}{\cos z'} = \frac{1}{\sqrt{1 - \left(\frac{R_e}{R_e + H}\sin(\alpha z)\right)^2}}, \qquad (4)$$

with $\alpha = 0.9782$, $H = 506.7$ km and R_e – the Earth radius.

The GNSS-derived STEC values are extracted from the geometry-free linear combination applied on dual-frequency carrier-phase smoothed code observations, as shown in equation (3). Data from around 190 stations of the International GNSS Service (IGS) is used with sampling rate of 30 seconds. In the case of satellite altimetry, the ionospheric refraction extracted from the double-frequency measurements of Jason-1 is adopted and converted into VTEC by a factor depending on the operational frequency of the altimeter. A spherical harmonic extension up to degree and order 15 is chosen for the global representation of VTEC, as a function of geomagnetic latitude and sun-fixed longitude [17]:

$$E_V(\beta,s)=$$
$$= \sum_{n=0}^{n_{max}} \sum_{m=0}^{n} \tilde{P}_{nm}(\sin\beta)(a_{nm}\cos(ms)+b_{nm}\sin(ms)), \tag{5}$$

   E_V ... Vertical Total Electron Content,

   β ... geomagnetic latitude of the ionospheric pierce point,

   $s = \lambda_G + UT - \pi$ ... sun-fixed longitude of the ionospheric pierce point,

   λ_G ... geographical longitude,

   $\tilde{P}_{nm} = N_{nm}P_{nm}$ ... normalized Legendre function from degree n and order m,

   a_{nm}, b_{nm} ... unknown coefficients of the spherical extension.

A software based on Matlab was developed for computation of 12 two-hourly global VTEC maps per day, the corresponding RMS (Root Mean Square) maps, and daily values of the DCB for all GNSS satellites and receivers. The VTEC and RMS values are estimated for grid points in an interval of $\pm 5°$ in longitude and $\pm 2.5°$ in latitude. The final outputs are in the IONEX [16] format.

For the combination of GNSS and altimetry data a least-squares adjustment (Gauss-Markov model) is applied on each set of observations and then the normal equations are combined by adding the relevant matrices. At this stage of our work, we adopt equal weights ($p_{GNSS} = 1$) for all GNSS observations in both the GNSS-only and the combined solution. As for the relative weighting of the altimetry data, different strategies are possible. On the one hand, due to the much higher number of GNSS measurements compared to satellite altimetry, the Jason-1 data should be over weighted, in order to increase its impact on the combined GIM. For the combined GIM presented in section 3.2 we adopt the a priori standard deviation $\sigma_0 = 0.25$ TECU ($p_{ALT} = 4$) for the altimetry measurements, which was determined experimentally. In the case of overweighting the Jason-1 data, however, it becomes crucial to assess the bias between GNSS and altimetry TEC, discussed in section 1.3. On the other hand, if we take into account the higher noise of the altimetry measurements compared to the carrier-phase smoothed code observations from GNSS, a lower weight should be applied on the Jason-1 derived observations than on the ones from GNSS. It has to be pointed out, that the relative weighting acts like a scaling factor for the contribution of the altimetry data in the combined GIM. It is a very complex issue, depending on the different spatial and temporal distribution of the observations and on their specific systematic errors. Therefore, the relative weighting of the two types of measurements needs to be optimised and is a matter of further investigation. Nevertheless, the spatial distribution of the altimetry observations (see Fig. 2) partly balances the gaps over the oceans between the GNSS stations, which is the main motivation behind adding altimetry data to the GNSS ionosphere model.
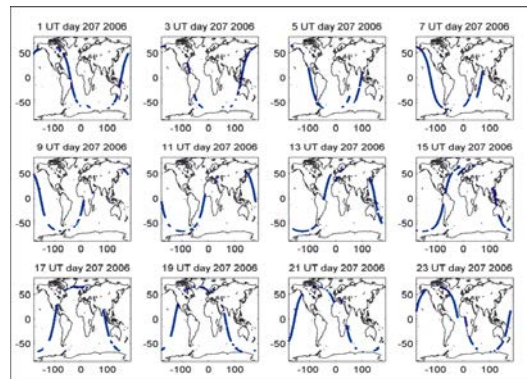
Figure 2: Jason-1 footprints in two-hourly snapshots, day 207 2006

3. Results

The ionosphere models computed within this work are referred to as IGG (Institute of Geodesy and Geophysics) GNSS-only or IGG COMB (GNSS

combined with altimetry data) Global Ionosphere Maps. The presented GIM refer to the 26th of July in 2006 (DOY 207). For validation of the obtained results, all IGG VTEC maps are routinely compared with the GIM provided by the IGS Analysis Centre CODE (Center for Orbit Determination in Europe, [5]). The bias and standard deviation of the difference CODE minus IGG are shown in the figure capture of the IGG VTEC maps (Fig.3a). It has to be mentioned, that the reference epochs of the IGG GIM (01, 03, 05, …, 23 UT) were set up at a very early stage of our work without taking into account the IGS conventions, and they do not overlap with the ones adopted in the IGS GIM (00, 02, 04, …, 24 UT).

Therefore, an interpolation in time is performed when computing the differences between the CODE and IGG maps, which can worsen the results of the comparison. The alignment of the IGG GIM reference epochs to the IGS convention is in progress.

3.1. IGG GNSS-only solution

The two-hourly IGG GNSS-only VTEC and RMS maps in two hours intervals for day 207 in 2006 are shown in figures 3a and 3b. The ionosphere maximum, which appears around local noon as travelling along with the Sun, is clearly visible in the VTEC maps (Fig. 3a). As anticipated, the precision of the GIM (Fig. 3b) is lower in areas where no GNSS sites are located, which is mainly above the sea surface.

In parallel to VTEC, the differential code biases for all GNSS satellites and ground stations are computed daily as constant values, with a zero-mean condition imposed on the satellite DCB. The results for day 207 in 2006 are shown in figures 4a and 4b. The estimated (IGG) values agree very well with the monthly DCB provided by the IGS Analysis Center CODE [5], especially for the GPS satellites and receivers. The lager differences in some of the DCB of the GLONASS system are subject of further investigation.
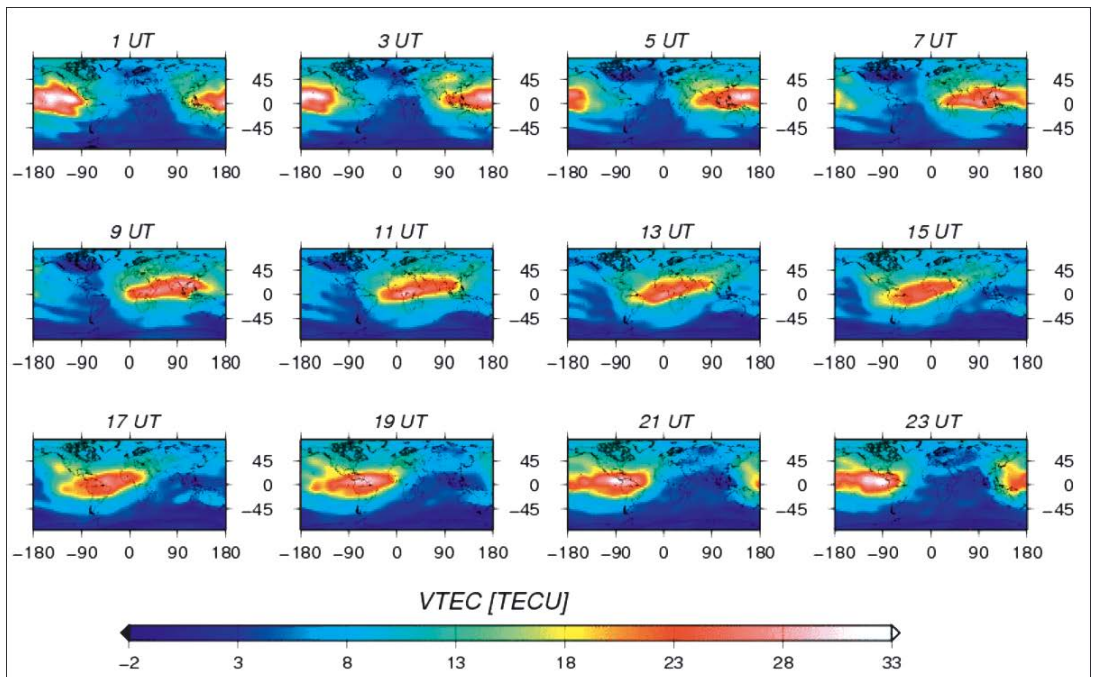
Figure 3a: VTEC, IGG GNSS-only, DOY 207 2006; bias$^{code-igg}$ = −0.10 TECIU, std$^{code-igg}$ = 0.97 TECU
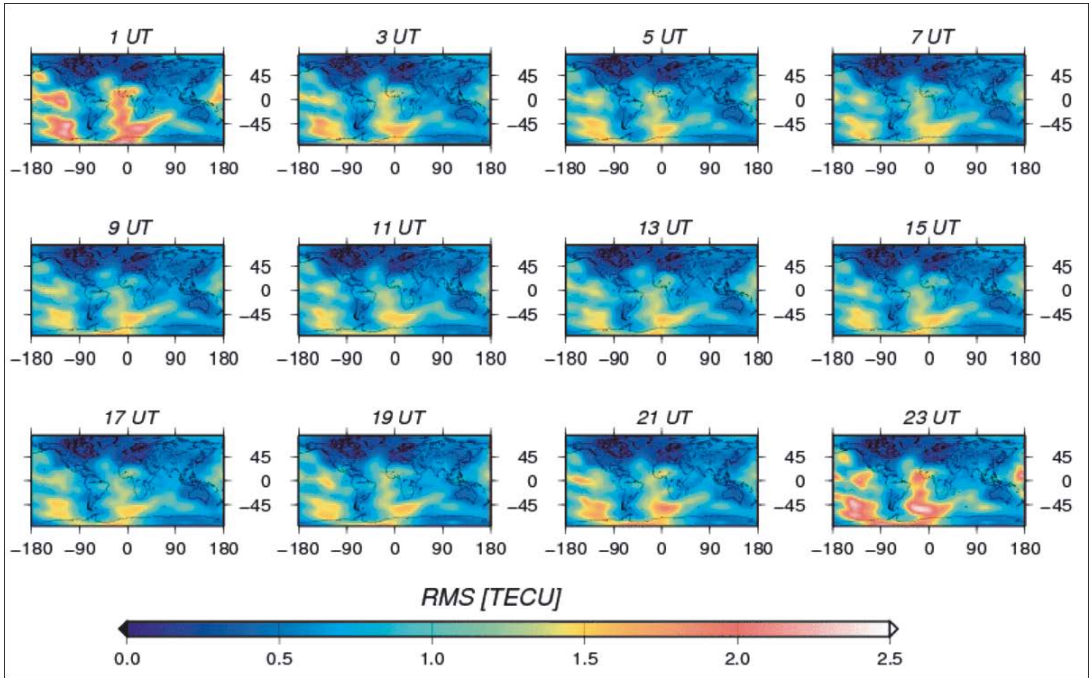
Figure 3b: RMS, IGG GNSS-only, DOY 207 2006
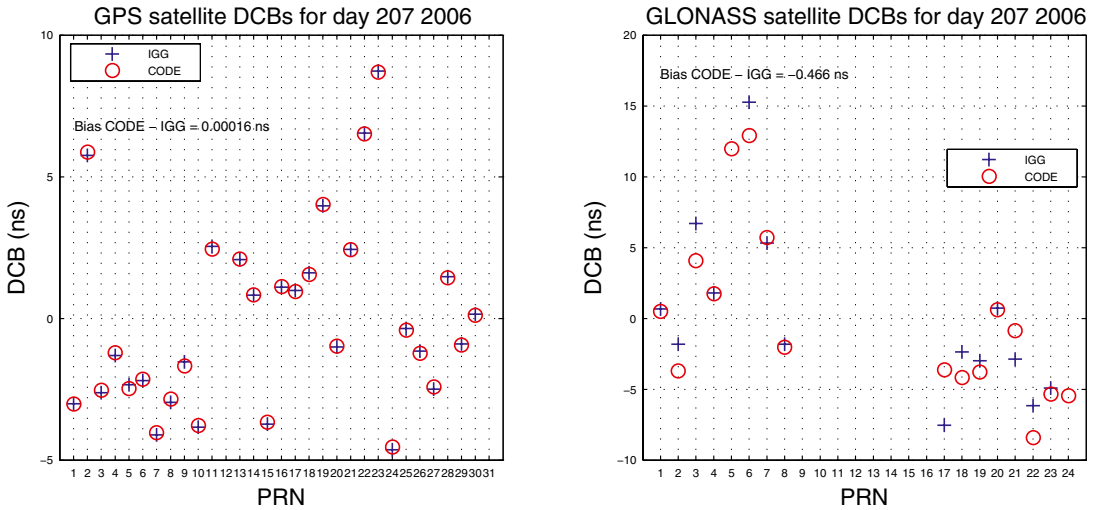
Figure 4a: Estimated GPS and GLONASS satellite DCB, IGG vs. CODE, day 207 2006

GPS stations DCBs for day 207 2006

GLONASS stations DCBs for day 207 2006
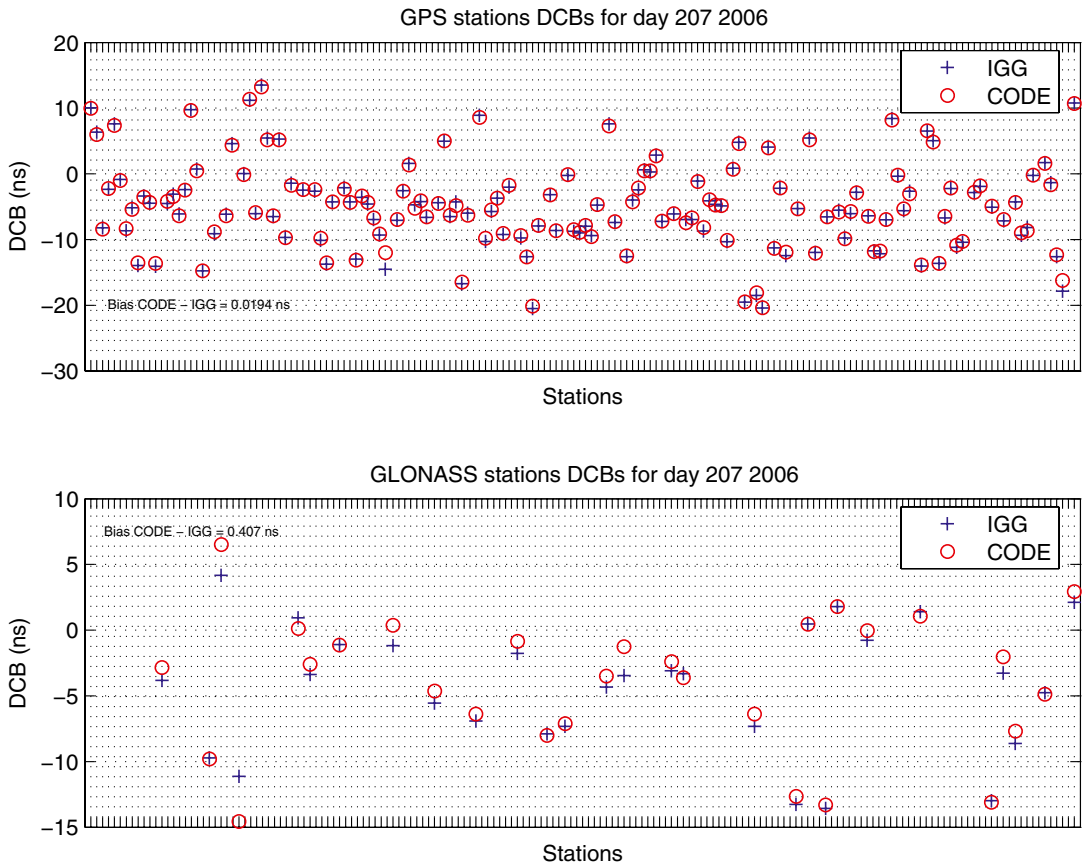
Figure 4b: Estimated GPS and GLONASS receiver DCB, IGG vs. CODE, day 207 2006

3.2. IGG COMB solution

The impact of altimetry data integration on the estimated GIM is evident over the areas coinciding with the footprints of Jason-1, shown in Fig.2. In the case of overweighting the altimetry data, the combination with Jason-1 measurements causes a decrease of the RMS up to 2 TECU and a general trend for increase of the VTEC values along the Jason-1 track (Fig.5). However, the bias between GNSS and altimetry TEC is a very important issue especially if a higher weight is applied to the altimetry data than to the GNSS measurements.

As mentioned in 1.3, several studies show that despite of the lower orbit altitude of the altimetry satellites, the vertical TEC delivered by these missions is higher than the values obtained from GNSS. Due to this contradiction it can be assumed, that the altimetry measurements are biased by an instrumental offset, similar to the GNSS DCB. The combination of ionosphere data from GNSS and altimetry, realised by stacking of the normal equations, allows the independent estimation of technique-specific time delays additionally to the combined ionospheric parameters. Thus, we developed combined ionosphere models for several subsequent days in 2006 with additional estimation of one constant Jason-1 bias per day, referred to as JB (Table 1). Since the Jason-1 bias is computed as a single unknown, it includes the effect of the plasmaspheric component (VTEC above ~1300 km height above the Earth's surface), additionally to the actual JASON instrumental offset.

DOY	201	202	203	204	205	206	207	208	209	210	211	212	213	214
JB	3.52	3.60	3.30	3.28	3.42	3.68	3.30	3.19	3.73	3.76	3.71	3.09	3.74	3.55

Table 1: Estimated daily Jason-1 bias in TECU

IGG GIM, VTEC along the Jason-1 track, day 207 2006

IGG GIM, RMS along the Jason-1 track, day 207 2006
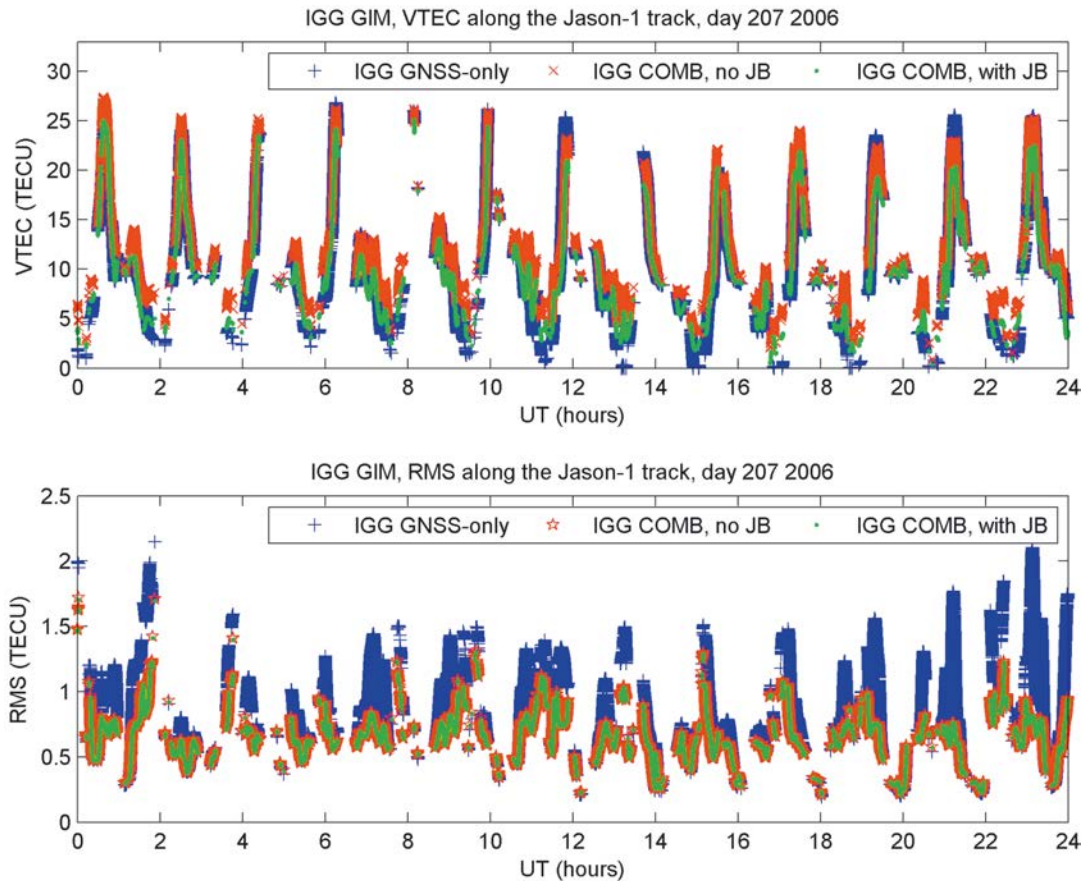
Figure 5: VTEC and RMS interpolated along the Jason-1 track from the different IGG GIM, DOY 207 2006

In order to demonstrate the differences between the GNSS-only and the combined IGG GIM, the VTEC and RMS values along the Jason-1 track were interpolated from the different IGG global maps for day 207 in 2006 – IGG GNSS-only GIM, IGG COMB GIM without estimation of Jason-1 bias, and IGG COMB GIM with Jason-1 bias estimated as a constant – and plotted as a function of time (Fig.5). Regarding the TEC values (Fig.5, upper plot), the IGG COMB without estimation of Jason-1 bias (IGG COMB, no JB) seems to generally overestimate the VTEC compared to the GNSS-only solution. The introduction of JB compensates this effect and acts as a negative offset as to the IGG COMB solution without JB, because the amount of TECU determined as Jason-1 bias (3.3 TECU for that day) does not contribute to the VTEC obtained from the altimetry data anymore. Nevertheless, there are also differences between the IGG GNSS-only and COMB with JB GIM. In the domain of low

ionosphere activity (TEC below 10 TECU), which is mostly at mid and high latitudes, a general trend for increase of the VTEC values along the Jason-1 track is visible. This effect can be interpreted as the positive contribution of the altimetry data in areas, where nearly no GNSS observations are available. However, there is also a decrease of VTEC in the combined model, coinciding with the ionospheric maximum as it travels with the Sun along the geomagnetic equator. This decrease can be related to the insufficient performance of the altimetry measurements in low latitudes, caused by the contribution of the topside ionosphere. As already mentioned, the altimetry measurements do not account for the topside ionosphere and therefore, despite of the discussed TEC overestimation, the integration of altimetry data in the GNSS GIM leads to a decrease of the obtained TEC over the area where the plasmaspheric contribution reaches its maximum (see also Fig.6). As for the precision of the

global maps, the lower plot in Fig.5 clearly shows the decrease of the RMS along the Jason-1 track caused by the combination. As it could be expected, the introduction of the Jason-1 bias has nearly no impact on the RMS of the combined GIM.

In order to investigate the obtained results and examine the self-consistency of our approach, we applied the same procedure, which is used for routine validation of the rapid and final global ionosphere maps produced by the IGS Analysis Centres (AC) [12]. In this validation procedure raw VTEC delivered by Jason-1 along its track is compared with the corresponding values interpolated from the global maps from GNSS data. The comparison was performed for 15 consecutive days in 2006 including the IGG GNSS-only and the IGG COMB GIM with estimated Jason-1 bias, as well as the official combined IGS GIM. Fig. 6 shows the mean bias for the regarded days of the final IGS solution and the two different IGG GIM compared to Jason-1. The comparison is performed in time (upper plot) and in latitude (lower plot). The difference between the raw Jason-1 VTEC and both the IGG and IGS GNSS-only GIM has a mean of about 3 TECU, with the IGS GIM performing slightly better than the IGG GNSS-only GIM. The magnitude of this difference generally corresponds to the estimated Jason-1 biases, shown in Table 1. As expected, with mean differences of -0.19 TECU in time and -0.09 TECU in latitude, the IGG COMB GIM with estimated JB coincides better with the Jason-1 raw data (after removing the computed offset), than the IGG and IGS GNSS-only solutions.
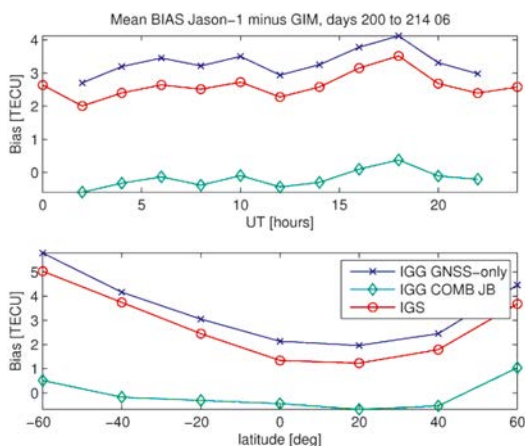
Figure 6: ΔVTEC, Jason-1 minus IGG and IGS GIM, DOY 207 2006

The plasmaspheric component is evident in the lower plot of Fig.6 as a decrease of the difference between Jason-1 and the GNSS-only solutions along the JASON orbital track at lower latitudes, where the topside ionosphere – not sampled from Jason-1 – reaches its maximum. As a next step of our study, the estimated constant Jason-1 bias will be replaced by a proper function of the latitude, which should account for the topside ionosphere. Such approach will improve the combined solution and, on the other hand, it could be used as a tool for rough estimation of the plasmaspheric component.

4. Conclusions and outlook

The combined GIM from GNSS and altimetry data have the potential to contribute to the accuracy of the global ionosphere maps especially over the seas, where none or only a few GNSS stations are located (worst case for GNSS). Still, the combined GIM must be optimised with main focus on the technique-specific error sources and the relative weighting of the individual results from the different techniques. Furthermore, the implementation of an appropriate function for modelling the Jason-1 offset could provide information about the topside ionosphere and improve the performance of the combined GIM. As a next step, a modified coordinate system can be adopted for improving the representation of the ionosphere and in this way enhancing the agreement between GNSS and altimetry derived TEC [1]. In order to achieve a global coverage and higher accuracy and reliability of the ionosphere models, the combination method can be adopted also for ionospheric data from other space geodetic techniques, such as VLBI and DORIS. For an objective validation of the results, they will be applied on single-frequency measurements, and then compared to the corresponding results from dual-frequency, in which the ionospheric effect is corrected.

References

[1] *Azpilicueta, F., Brunini, C., Radicella, S. M.*: Global ionospheric maps from GPS observations using modip latitude. Advances in Space Research, Volume 38, Issue 11, pp. 2324-2331, 2006

[2] *Brunini, C.*: Global ionospheric models from GPS measurements. PhD thesis, Universidad Nacional de La Plata, La Plata, Argentina, 1997

[3] *Brunini, C., Meza, A., Bosch, W.*: Temporal and spatial variability of the bias between TOPEX- and GPS-derived total electron content. J. Geod. 79, 4-5, pp. 175-188, 2005

[4] *Chelton, D.B., Ries, J.C., Haines, B.J., Fu, L.L., Callahan, P.S.*: Satellite altimetry. In: Satellite altimetry and Earth sciences. Academic, London, pp. 57-64, 2001

[5] Center for Orbit Determination in Europe, Global Ionosphere Maps Produced by CODE, http://www.aiub.unibe.ch/ionosphere/

[6] *Dach, R., Hugentobler, U., Fridez, P., and Meindl, M.* (Eds.), Bernese GPS Software, Version 5.0. Astronomical Institute, University of Bern, 2007

[7] *Feltens, J.*, Chapman profile approach for 3-D global TEC representation. In: Proceeding of the IGS AC Workshop, pp. 285-297, Dow, J.M., Kouba, J., and Springer, T. (Eds.), Darmstadt, Germany, February 9-11, 1998

[8] *Feltens, J., Schaer, S.*: IGS Products for the Ionosphere. In: Proceeding of the IGS AC Workshop, pp. 225-232, Dow, J.M., Kouba, J., and Springer, T. (Eds.), Darmstadt, Germany, February 9-11, 1998

[9] *Hargreaves, J.K.*: The solar-terrestrial environment. In: Cambridge atmospheric and space science series, Cambridge University Press, 1992

[10] *Hartmann, G.K., Leitinger, R.*: Range errors due to ionospheric and tropospheric effects for signals above 100 MHz. Bulletin Géodesique 58, pp. 109-136, 1984

[11] *Hernández-Pajares, M., Juan J.M., Sanz, J.*: New approaches in global ionospheric determination using ground GPS data. J. Atmos. Solar Terrestrial Phys. 61, pp. 1237-1247, 1999

[12] *Hernández-Pajares M.*: IGS Ionosphere WG Status Report: Performance of IGS Ionosphere TEC Maps, IGS Workshop, Bern, March 2004

[13] *Hugentobler, U., M. Meindl, G. Beutler, H. Bock, R. Dach, A. Jäggi, C. Urschl, L. Mervart, M. Rothacher, S. Schaer, E. Brockmann, D. Ineichen, A. Wiget, U. Wild, G. Weber, H. Habrich, and C. Boucher*: CODE IGS Analysis Center Technical Report 2003/2004, in IGS 2004 Technical Reports, Ken Gowey et al. (Eds.), IGS Central Bureau, Jet Propulsion Laboratory, Pasadena, California, USA, 2006

[14] *Imel, D.A.*: Evaluation of the TOPEX/POSEIDON dual-frequency ionosphere correction. J. Geophys. Res. 99, C12, pp. 24895-24906, 1994

[15] *Mannucci, A. J., Wilson, B., Yuan, D., Linqwister, U., and Runge, T.*: A global mapping technique for GPS-derived ionospheric total electron content measurements. Radio Sci. 33, pp. 565–582, 1998

[16] *Schaer, S., Gurtner, W., Feltens, J.*: IONEX: The IONosphere map eXchange format version 1. In: Proceeding of the IGS AC Workshop, pp. 233-247, Dow, J.M., Kouba, J., and Springer, T. (Eds.), Darmstadt, Germany, February 9-11, 1998

[17] *Schaer, S.*: Mapping and predicting the Earth's ionosphere using the Global Positioning System. PhD thesis, Bern University, Switzerland, 1999

[18] *Todorova, S., Hobiger, T., Schuh, H.*: Using the Global Navigation Satellite System and Satellite Altimetry for combined Global Ionosphere Maps, submitted to Advances in Space Research, 2007

Contact

Dipl.-Ing. Sonya Todorova: Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: stodo@mars.hg.tuwien.ac.at

Prof. Harald Schuh: Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: harald.schuh@tuwien.ac.at

Dr. Thomas Hobiger: Space-Time Standards Group, Kashima Space Research Center, NICT, 893-1 Hirai, Kashima, 314-0012 Ibaraki, Japan. E-mail: hobiger@nict.go.jp

Dr. Manuel Hernández-Pajares: Group of Astronomy and Geomatics, Univ. Politécnica de Catalunya, Jordi Girona 1, C3, 08034, Barcelona, Spain. E-mail: manuel@ma4.upc.edu

vgi

Compilation of a new Bouguer gravity data base in Austria

Bruno Meurers and Diethard Ruess

Abstract

All gravity data acquired in Austria by several institutions during the past 50 years have been re-processed in order to get a high accurate data base for interpretation and geoid determination. The paper addresses the problem of data homogenization and quality assessment. Detection of gross coordinate errors has been performed by comparing with high resolution digital terrain models. Erroneous horizontal coordinates have been corrected by digitizing of modern topographic maps and by utilizing the digital cadastre. Another focus is set on the problem of correction errors (mass correction, free air correction) in order to estimate the accuracy of the Bouguer gravity.

Kurzfassung

Alle verfügbaren Schweredaten Österreichs, die in den letzten 50 Jahren an verschiedenen Institutionen entstanden sind, werden neu aufbereitet um einen hochgenauen Datenbestand zu erhalten, der für geophysikalische Interpretationen und für die Geoidbestimmung verwendet werden kann. Der Artikel spricht das Problem der Homogenisierung und der Beurteilung der Daten an. Grobe Koordinatenfehler wurden aus dem Höhenvergleich mit einem hochauflösenden digitalen Geländemodell aufgedeckt. Fehlerhaften Lagekoordinaten wurden durch Digitalisierung moderner topographischer Karten bzw. durch Verwendung der digitalen Katastralmappe verbessert. Weiters wird das Problem der Korrekturfehler angesprochen (Massenkorrektur, Freiluftkorrektur) um die Genauigkeit des Bougerschwerefeldes abzuschätzen.

1. Introduction

Gravity data available in Austria have been acquired during the past 50 years by

- Federal Office of Metrology and Surveying (BEV), Austria
- Institute of Meteorology and Geophysics, University of Vienna, Austria
- Institute of Geophysics, Mining University of Leoben, Austria
- Institute of Geophysics, Technical University of Clausthal, Germany
- OMV AG, Austria,
- Institute of Geophysics, Technical University of Vienna, Austria

In the early beginning, the gravity network design was motivated by the establishment of an orthometric height system in Austria. The first gravity map of Austria was an important by-product (Senftl, 1965). Accordingly, the gravity stations are mainly arranged along levelling lines and consequently often situated within regions of very local anomalies due to the gravity effect of sedimentary valley fillings, for example. While their vertical coordinates were well determined, their horizontal coordinates had to be obtained by topographic map digitization based on maps

available in that time. Oil exploration companies built up the first gravity networks of areal character. However, as the target was hydrocarbon exploration, surveying was confined to the Alpine Foreland, the Vienna basin and to parts of the Flysch and Calcareous zone of the Eastern Alps (Zych 1988). For better understanding the crustal structure of the Alps additional gravity profiles were established across the Eastern Alps (Ehrismann et al. 1969, 1973, 1976, Götze et al. 1978) during the late 1960ies and 1970ies. In all these cases, vertical coordinates were derived from precise levelling methods while the horizontal ones were read from topographic maps. Due to limited accessibility, stations were widely missing in rugged mountainous terrain, and gravity had to be interpolated there.

However, Steinhauser et al. (1990) showed that interpolation errors of up to 100 μms^{-2} can occur in the Bouguer anomaly pattern when stations are arranged along profiles exclusively. Therefore, the first areal investigation with a lot of stations even high up mountain flanks and tops has been done during the late 1970ies on the so called Gravimetric Alpine Traverse (Meurers et al. 1987) and by the gravimetric research group of the Technical University of Clausthal (Germany)
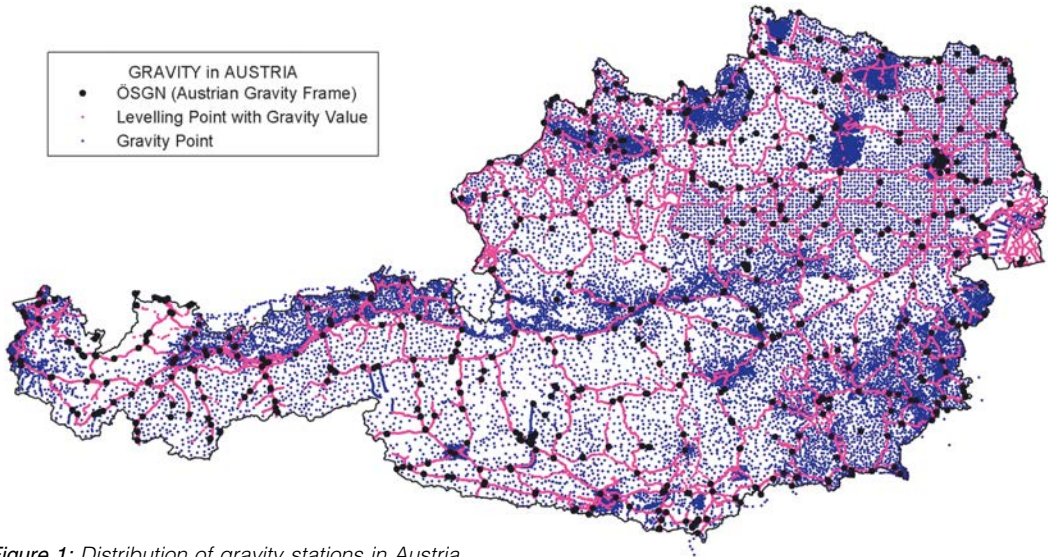
Figure 1: Distribution of gravity stations in Austria

(e.g. Götze et al. 1979, Schmidt 1985) in the adjoining central part of the Eastern Alps. The westernmost and the south-east part of Austria were surveyed by the Mining University of Leoben (Posch and Walach 1990, Walach 1990, Walach and Winter 1994). The first Bouguer gravity map of entire Austria since Senftl's (1965) work has been compiled about 15 years ago (Kraiger and Kühtreiber 1992, Kraiger 1993). In that time wide areas, especially along the Alpine crest, were not or poorly covered by gravity stations, and no high resolution digital terrain models could be utilized for precise mass corrections. The gaps remaining especially along the crest of the Eastern Alps have been filled since 1990 in cooperation of the Institute of Meteorology and Geophysics (University of Vienna), the Central Institute for Meteorology and Geodynamics (Vienna) and the Department of Physical Geodesy of TU Graz by applying GPS techniques and helicopter transportation in otherwise un-accessible mountainous regions. Presently the gravity map is supported by 54000 stations (Fig. 1). The average station interval is less than 3 km even in the high mountains resulting to an average station density of 1 station/9 km^2 or higher.

2. Data processing

During the past 50 years both data acquisition and processing methods have been continually improved. Depending on the origin, the available gravity data refers to different datum and exhibits different quality and accuracy. Additionally, digital terrain models of high spatial resolution (<50 m)

have been established and could be used for more accurate mass corrections (Graf 1996). Therefore, data homogenization is required.

3. Gravity datum and calibration

The industrial data (OMV) was tied to an own gravity base net. Compared to the Austrian gravity base net (AGBN/OeSGN), which is supported by absolute gravity observations (Ruess and Gold 1995, Ruess 2001, 2002), the OMV base net has not only an offset but also a slightly different scale due to limited calibration accuracy. Therefore, both scale factor and offset had to be determined by least squares adjustment of numerous ties between the OMV base net and AGBN (Meurers 1992a). The gravimeter calibration was controlled by observations at the Hochkar Calibration Line in Austria which is constraint by absolute gravity measurements (Meurers and Ruess 2001). Finally the gravity of all stations was transformed to the absolute gravity datum established by Ruess (2002).

4. Correcting erroneous coordinates

Horizontal coordinates have been obtained in many cases by topographic map digitization based on those maps which were available when the stations have been established. Mostly maps with a scale of 1:50000 have been used for that purpose. Therefore, generally the coordinate accuracy is estimated as ±25 m in x- and y-direction, but even higher errors can occur occasionally. However, apart from these random
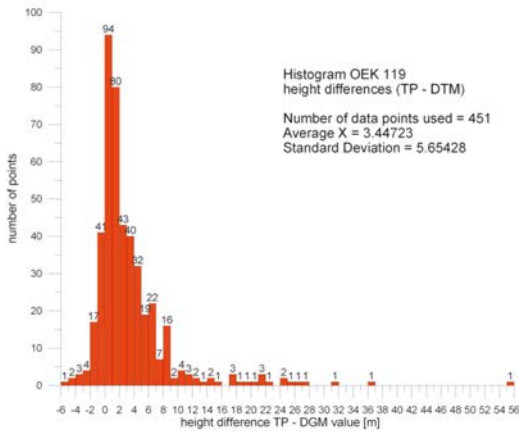
Figure 2: Histogram of the differences between the elevations of geodetic benchmarks (TP) and their corresponding interpolated DTM heights within a mountainous area (Austrian map sheet 119).

Figure 3: Histogram of the differences between the elevations of all gravity stations and their corresponding interpolated DTM heights within the same area as investigated in Fig. 2 (Austrian map sheet 119).

errors, the low quality of some older maps causes even systematic errors. In order to check gross coordinate errors we compared the station heights with those obtained by interpolating a high resolution digital terrain model (DTM) with 50 m spacing. This method of course works well only where interpolation does not seriously smooth out the true topography. Fig. 2 demonstrates the capability of this method by applying the procedure on geodetic benchmarks in Austria which typically exhibit coordinate errors of less than 10 cm. Fig. 2 shows the histogram of the differences between the elevations of the

geodetic benchmarks (H_{TP}) and those of the DTM (H_{DTM}) on the Austrian map sheet 119 (Inn valley, Tyrol). Large differences belong to benchmarks situated on extreme terrain (steep rugged rock or peaks), where gravity stations normally are not established. In Fig. 3 the actual situation for all gravity stations within the same area is displayed. Large discrepancies potentially point to errors of horizontal or vertical coordinates, which have to be checked and re-evaluated if necessary.

Erroneous coordinates have been corrected by making use of modern topographic maps, orthophoto maps and by utilizing the digital cadastre, based on the original gravity station descriptions or sketches. An example of a digital cadastre map (DKM) used for fixing the horizontal position of the gravity stations is shown in Fig 4.

Figure 4: Example of a digital cadastre map (DKM) used for fixing the horizontal position of the gravity stations.

5. Bouguer gravity determination

Normal gravity at the gravity stations has been calculated by applying a Taylor series expansion to the 2nd order both in elevation and geometrical flattening (Wenzel 1985) based on the Geodetic Reference System 1980 (Moritz 1984). Atmospheric corrections according to Wenzel (1985) were performed. The height system used corresponds to orthometric rather than to ellipsoidal heights. The expected geophysical indirect effect varies between 8.5 and 10 mGal[1] in Austria (Vajda et al. 2006).

---

1) 1 mGal $= 10^{-5}$ m/s^2
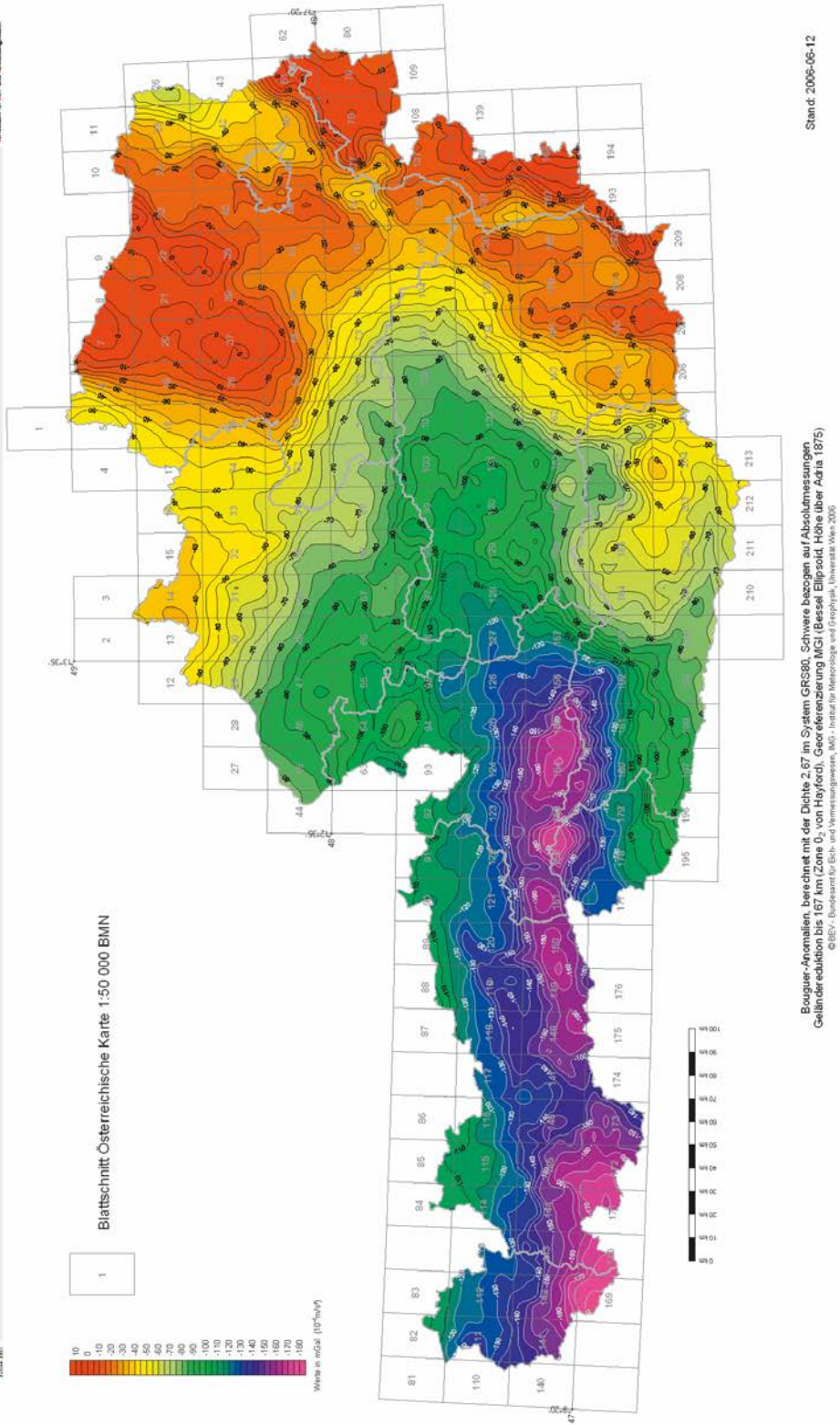
Figure 5: Bouguer gravity map of Austria. Absolute gravity datum, GRS 1980, mass corrections: spherical, Hayford zone O2, density 2670 kg/m³.

The complete mass correction was calculated by using modern and high accurate mass correction methods (Meurers et al. 2001). We applied a spherical mass correction (167 km radius, Hayford zone O2) assuming constant density of 2670 kgm^{-3}. This value is close to the mean surface rock density in the investigated area. The truncation of the mass correction area beyond 167 km is justified as it generates only a very small height dependant effect (Meurers 1992b). Based on a digital terrain model (Graf 1996) with high spatial resolution (50 m spacing), the topography close to the gravity stations (< 1000 m) was approximated by a polyhedral surface, which is able to cope with rugged terrain and allows for exact determination of the corresponding gravity effect (Götze and Lahmeyer 1988). Additionally the gravity effect of lakes has been corrected, which is an important issue in high mountainous areas (Steinhauser et al. 1990).

Fig. 5 shows the new Bouguer gravity map of Austria as coloured map in 10 mGal intervals. The geo-reference (coordinates) is based on the national Austrian system (MGI).

Acknowledgement

References

[1] *Ehrismann, W., Rosenbach, O., Steinhauser, P., 1969*: Untersuchungen zur Korrelation zwischen Freiluftanomalie und Stationshöhe im Hochgebirge. ÖZ f. Verm. u. Photogr., Jgg.57, 183-191, Wien.

[2] *Ehrismann, W., Leppich, W., Lettau, O., Rosenbach, O. and Steinhauser, P., 1973*: Gravimetrische Detail-Untersuchungen in den westlichen Hohen Tauern. Z.f.Geoph., 39, 115-130.

[3] *Ehrismann, W., Götze, H.J., Leppich, W., Lettau, O., Rosenbach, O., Schöler, W., Steinhauser, P., 1976*: Gravimetrische Feldmessungen und Modellrechnungen des Krimmler Achentales und Obersulzbachtales. Geologische Rundschau 65, 767-778.

[4] *Götze, H.J., Lahmeyer, B., 1988*: Application of three-dimensional interactive modeling in gravity and magnetics. Geophysics, 53, 1096-1108.

[5] *Götze, H.J., Rosenbach, O., Schöler, W., 1978*: Gravimetric measurements on three N-S profiles through the Eastern Alps – Observational results and preliminary modeling. In: Alps, Apennines, Hellenides. Inter-Union Comm. on Geodynamics, Scient. Rep., 38, 44-49, Stuttgart.

[6] *Götze, H.J., Rosenbach, O., Schöler, W., 1979*: Gravimetrische Untersuchungen in den östlichen Zentralalpen. Geol. Rdsch., 68, 1, 61-82.

[7] *Graf J., 1996*: Das digitale Geländemodell für Geoidberechnungen und Schwerereduktionen in Österreich. Proceed. 7. intern. meeting on Alpine Gravimetry. Österr. Beitr. Met. Geoph., 14, 121-136, ZAMG, Wien.

[8] *Kraiger G., 1993*: Die neue Schwerekarte von Österreich (Bouguer-Isanomalien). 6. Int. Alpengrav. Koll., Leoben 1993, Österr. Beitr. Met. Geoph., 8, 39-53.

[9] *Kraiger G. Kühtreiber, N., 1992*: Preliminary Results of a New Bouguer Map of Austria. Geodesy and Physics of the Earth: Geodetic Contributions to Geodynamics. 7th International Symposium "Geodesy and Physics of the Earth, Symposium No. 112, Berlin, Springer, 1992, p.133.

[10] *Meurers, B., 1992a*: Bearbeitung der Schweredaten der OMV-AG. Unveröff. Firmenbericht. OMV-AG Wien.

[11] *Meurers, B., 1992b*: Untersuchungen zur Bestimmung und Analyse des Schwerefeldes im Hochgebirge am Beispiel der Ostalpen. Österr. Beitr. zu Met. u. Geoph., 6, 146. ZAMG Publ.Nr. 343, Wien.

[12] *Meurers, B., Ruess, D., 2001*: Gravity Measurements at the Hochkar Calibration Line (HCL). Proceed. 8. intern. meeting on Alpine Gravimetry. Österr. Beitr. Met. Geoph., 26, 209-216, ZAMG, Wien.

[13] *Meurers, B., Ruess, D., Graf, J., 2001*: A Program System for High Precise Bouguer Gravity Determination. Proceed. 8th intern. meeting on Alpine Gravimetry, Leoben 2000, Österr. Beitr. Met. Geoph., 26, 217-226,ZAMG, Wien.

[14] *Meurers, B., Ruess, D., Steinhauser, P., 1987*: The Gravimetric Alpine Traverse, in Geodynamics of the Eastern Alps, Flügel H. and Faupl F. (Eds), Deuticke, 334-344, Vienna 1987, ISBN 3-7005-4528-2.

[15] *Moritz, H., 1984*: Geodetic reference system 1980. Bull. Geod., 54, 3, 395-405.

[16] *Posch, E., Walach, G., 1989*: Das Bouguerschwerefeld in Vorarlberg und im Bereich der Übergangszone zwischen West- und Ostalpen. 5. Int. Alpengrav. Koll., Graz 1989, Österr. Beitr. Met. Geoph., 2, 147-151.

[17] *Ruess, D., 200*: Absolute Schweremessungen in Österreich seit 1987. In: Hammerl, Ch., Lenhardt, W., Steinacker, R., Steinhauser, P., (eds.), 2001: Die Zentralanstalt für Meteorologie und Geodynamik 1851 - 2001, 150 Jahre Meteorologie und Geophysik in Österreich, 590 – 607, Leykam, Graz 2001, ISBN 3-7011-7437-7.

[18] *Ruess, D., 2002*: Der Beitrag Österreichs an UNIGRACE – Unification of Gravity Systems of Central and Eastern European Countries. VGI, 3+4/2002, 129-139, Wien.

[19] *Ruess, D., Gold, W., 1996*: The Austrian Gravity Base Net 1995 stabilized by Absolute Gravity Measurements and Connected to the European Network Adjustment. VGI, 3/96, 275-283.

[20] *Senftl, E., 1965*: Bouguer – Isoanomalienkarte von Österreich, 1:1.000.000, BEV-Wien.

[21] *Schmidt, S., 1989*: Untersuchungen zum regionalen Verlauf des Vertikalgradienten der Schwere im Hochgebirge. Ph.D. thesis, TU Clausthal.

[22] *Steinhauser, P., Meurers, B., Ruess, D., 1990*: Gravity investigations in mountainous areas. Exploration Geophysics, 21, 161-168.

[23] *Walach, G., 1990*: The Gravity Field in Southern Austria. Mitt. Geodät. Inst. TU Graz, 67, 135–140.

[24] *Walach, G., Winter, P., 1994*: Bouguer Anomaly Map of Styria. „Gravity and Geoid", Proc. Joint Symposium IUGG.

[25] *Wenzel, F., 1985*: Hochauflösende Kugelfunktionsmodelle für das Gravitationspotential der Erde. Wiss. Arb. Fachr. Vermessungswesen Univ. Hannover, Nr. 137.

[26] *Vajda, P., Vanicek, P., Meurers, B., 2006*: A new physical foundation for anomalous gravity. Stud. Geophys. Geod., 50, 189-216.

[27] *Zych, D., 1988*: 30 Jahre Gravimetermessungen der ÖMV Aktiengesellschaft in Österreich und ihre geologisch-geophysikalische Interpretation. Arch. f. Lagerst.forsch. Geol. B.-A., 9, 155-175.

Contact

Ao.Univ.Prof. Dr. Bruno Meurers: Institute of Meteorology and Geophysics, University of Vienna, Althanstraße 14, UZA II, A - 1090 Wien. E-mail: bruno.meurers@univie.ac.at

Dr. Diethard Ruess: Federal Office of Metrology and Surveying (BEV), Vienna, Austria. E-mail: Diethard.Ruess@bev.gv.at

Operational data processing of ESA's GOCE gravity field mission

Roland Pail

Abstract

The operational scientific processing of GOCE data will be performed by the European GOCE Gravity Consortium (EGG-C) in the framework of the ESA-funded project "GOCE High-level Processing Facility" (HPF). One key component of the HPF hardware and software system is the processing of a spherical harmonic Earth's gravity field model and the corresponding full variance-covariance matrix from the precise GOCE orbit and satellite gravity gradiometry data. In parallel to two other HPF teams, this task is performed by the "Sub-processing Facility (SPF) 6000". The second main task of SPF6000 is the production of quick-look gravity field products in parallel to the GOCE mission for system diagnosis purposes. The paper gives an overview of the operational SPF6000 software system that has been implemented and integrated at the facilities of TU Graz. On the basis of a numerical case study, which is based on the data of an ESA GOCE end-to-end simulation, the processing architecture is presented, and several aspects of the involved functional and stochastic models are addressed.

Kurzfassung

Die operationelle Prozessierung von GOCE-Daten wird im Rahmen des ESA-Projektes "GOCE High-level Processing Facility" (HPF) von einem Konsortium, gebildet aus 10 europäischen Forschungsinstituten (European GOCE Gravity Consortium, EGG-C) erfolgen. Eine Hauptkomponente dieses dezentralen Hardware- und Software-Systems ist die Berechnung eines globalen GOCE-Erdschwerefeldmodells, parametrisiert mittels sphärischen harmonischen Koeffizienten und der zugehörigen Varianz-Kovarianzmatrix, aus GOCE-Orbit- und Gradiometriedaten. Neben zwei weiteren HPF-Teams wird diese Aufgabe von der sogenannten "Sub-processing Facility (SPF) 6000" durchgeführt. Die zweite Hauptaufgabe der SPF6000 besteht in der fortlaufenden Produktion von schnellen Schwerefeldlösungen während der Mission als Beitrag zur Missionskontrolle. In dieser Arbeit wird das operationelle Softwaresystem der SPF6000 vorgestellt, das am Standort TU Graz implementiert und integriert wurde. An Hand einer numerischen Simulationsstudie, die auf Daten einer GOCE-Simulation der ESA beruht, werden die Prozessierungs-Architektur und ausgewählte Aspekte hinsichtlich der zu Grunde liegenden funktionalen und stochastischen Modelle präsentiert und diskutiert.

1. Introduction

The dedicated satellite gravity mission GOCE (Gravity field and steady-state Ocean Circulation Explorer; [6]), the first Earth Explorer Core Mission in the context of ESA's Living Planet programme, strives for a high-accuracy, high-resolution global model of the Earth's static gravity field. GOCE is based on a sensor fusion concept: satellite-to-satellite tracking in the high-low mode (hl-SST) using GPS, and satellite gravity gradiometry (SGG). During the (at least) two GOCE measurement phases of 6 months each, GOCE will provide a huge data set consisting of several 100 million orbit and gravity gradiometry data, which contains abundant information about the gravity field of the Earth on a near-global scale, from very low (derived mostly from hl-SST) to high (derived mostly from SGG) frequencies.

The mathematical model for the parameterization of the global Earth's gravity field is usually based on an expansion into spherical harmonics.

The gravitational potential V can be expressed in a spherical coordinate system (r, ϑ, λ) by:

$$V(r,\vartheta,\lambda) = \frac{GM}{R} \sum_{l=0}^{lmax} \left(\frac{R}{r}\right)^{l+1} \sum_{m=0}^{l} \overline{P}_{lm}(\cos\vartheta) \cdot \\ \cdot \left[\overline{C}_{lm}\cos(m\lambda) + \overline{S}_{lm}\sin(m\lambda) \right] \quad (1)$$

where G is the gravitational constant, M and R are the Earth's mass and reference radius, $\overline{P}_{lm}$ are the fully normalized Legendre polynomials of degree l and order m, and $\{\overline{C}_{lm}, \overline{S}_{lm}\}$ are the corresponding spherical harmonic coefficients.

In the case of a gravity field model resolution complete to degree and order $l_{max} = 250$, this yields approximately 63000 unknown spherical harmonic coefficients $\{\overline{C}_{lm}, \overline{S}_{lm}\}$. The estimation of these coefficients from the complementary hl-SST and SGG data sets is a demanding numerical and computational task, and therefore efficient solution strategies are required to solve the corresponding large normal equation systems. During the last decade, several approaches have been developed to perform this task (e.g. [22] [24]

[8] [16] [14]). In [16] [17] the rigorous solution of the large normal equation matrix by means of a parallel processing strategy implemented on a Linux-PC cluster was proposed.

The scientific data processing (Level 1b to Level 2[1)]) is performed by the "European GOCE Gravity Consortium" (EGG-C), a consortium of 10 European university and research institutes, in the framework of the ESA-funded project "GOCE High-Level Processing Facility" (HPF; [23]). Table 1 gives an overview of the project partners of GOCE HPF.

The HPF project is jointly managed by IAPG and SRON, principal investigator is Prof. Rainer Rummel (IAPG). Table 2 lists the main work packages of GOCE HPF. Usually, several partners contribute to one work package.

The SPF6000 is a co-operation of TU Graz, Austrian Academy of Sciences, University of Bonn, and TU Munich, under the lead of TU Graz. The two main tasks of SPF6000 are:

1. the computation of a global Earth's gravity field model from GOCE data, parameterized by spherical harmonic coefficients, and the corresponding error estimates in terms of a full variance-covariance matrix;

2. the continuous production of quick-look gravity field solutions in parallel to the mission as a tool for GOCE system diagnosis and mission control.

Acronym	Institution
AIUB	Astronomical Institute of Bern, University of Bern, Switzerland
CNES	Centre National d'Etudes Spatiales, Groupe de Recherche de Géodésie Spatiale, Toulouse, France
FAE/A&S	aculty of Aerospace Engineering, Astrodynamics & Satellite systems, TU Delft, The Netherlands
GFZ	GeoForschungsZentrum Potsdam, Germany
IAPG	Institute of Astronomical and Physical Geodesy, TU Munich, Germany
ITG	Institute of Theoretical Geodesy, University of Bonn, Germany
POLIMI	DIIAR – Sezione Rilevamento, Politecnico di Milano, Italy
SRON	National Institute for Space Research, Utrecht, The Netherlands
TUG	Institute of Navigation and Satellite Geodesy, TU Graz, Austria
UCPH	Department of Geophysics, University of Copenhagen, Denmark

Table 1: EGG-C members

SPF	Task	SPF lead
2000	Central Processing Facility (CPF)	SRON
3000	Scientific Pre-processing and External Calibration	SRON
4000	Orbit Determination	FAE/A&S
5000	Gravity Field Determination – Direct Approach	CNES
6000	Gravity Field Determination – Time-wise Approach	TUG
7000	Gravity Field Determination – Space-wise Approach	POLIMI
8000	Level 2 Products Validation	IAPG

Table 2: Work packages / SPFs of GOCE HPF

[1)] Level 1 b data are (internally) calibrated instrument time series, while Level 2 data are products which are generated in the framework of HPF, such as precise GOCE orbits, externally calibrated gravity gradients, and finally the gravity field model coefficients and corresponding covariance information.
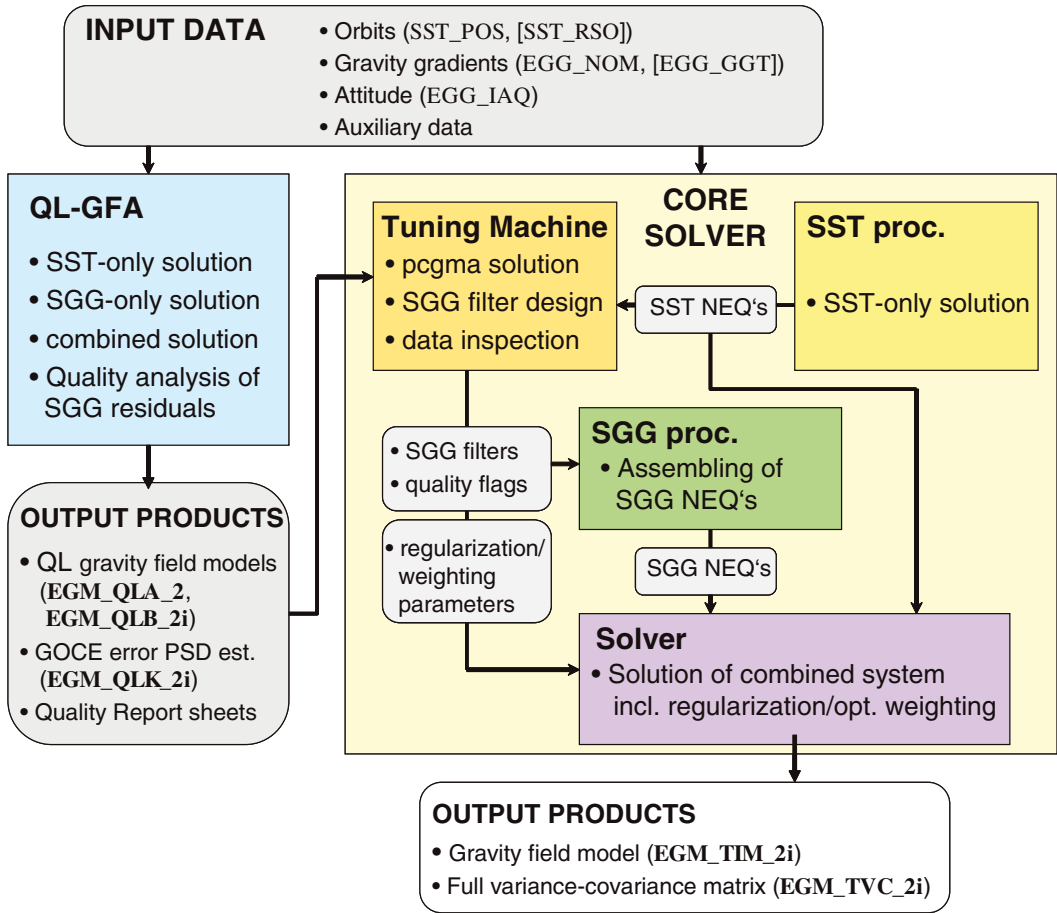
Figure 1: SPF6000 software architecture and product flow. The official acronyms of the input and output products are given in brackets.

2. Software architecture

Fig. 1 shows the architectural design, the main components and the product flow through the SPF6000 software system. It is conceived in a highly modular manner that allows the investigation of specific aspects of gravity modelling such as filtering, numerical stability and optimum regularization, complementary relations of SST and SGG and their optimum weighting.

The data transfer between SPF6000 and the central HPF data repository CPF (Central Processing Facility) is managed via automated interfaces. At SPF6000, the data are stored on a central access local data server.

The software system is composed of two main components: the Quick-Look Gravity Field Analysis (QL-GFA), and the Core Solver (CS), which will be briefly described in the following.

2.1. Quick-Look Gravity Field Analysis (QL-GFA)

This stand-alone software system performs the computation of fast approximate gravity field solutions based on SGG and hl-SST data, for the purpose to derive a fast diagnosis of the GOCE system performance and of the input data in parallel to the mission with short latencies. These gravity field products are input to ESA's calibration/validation activities in the frame of the GOCE mission control.

Key tasks of QL-GFA are:

- Check of SGG and hl-SST input data and analysis of partial / incomplete SGG and hl-SST data sets.

- Computation of quick-look gravity field models (SGG-only, SST-only, combined SST+SGG) aiming at a fast analysis of the information content of the input data on the level of the gravity field solution. Additionally, quick-look gravity solutions are statistically tested against reference gravity models.

- Estimation of the gradiometer error PSD (power spectral density) from the residuals of a SGG-only gravity field analysis, and application of previously defined statistical hypothesis test strategies in time and frequency domain [10].

- Production of Diagnosis Report Sheets: All these system diagnosis products are reported by means of a standardized Diagnosis Report Sheet.

QL-GFA solutions complete to degree/order 250 can be processed within the order of one to two hours on a standard PC. The efficiency and speed of QL-GFA is founded mainly on the application of FFT techniques (semi-analytical approach), the assumption of block-diagonality of the normal equation matrix, and also on a simplified filter strategy in the spectral domain to cope with the coloured noise characteristics of the gradiometer. Deviations from this assumption are incorporated by means of an iterative procedure [21] [18]. A detailed description of the functionality of QL-GFA can be found in [20].

QL-GFA will be applied at two stages: Quick-Look-A (QL-A) is applied to Level 1b preliminary orbits (accuracy $\sim$ 10 m) and the Level 1b gravity gradients. The main purpose at this stage is a rough check of the SGG time series, with special concern on the testing of the SGG error PSD. For QL-A, consecutive gravity field solutions will be available in a daily interval. They will be generated fully automated with a latency of maximum 4 hours after arrival of all required input data. The achievable accuracy is mainly dependent on the correct (internal) calibration of the Level 1b gradients.

Quick-Look-B (QL-B) is applied to the Level 2 rapid science orbit solution (accuracy in the decimetre range) and the externally calibrated gravity gradients. Consecutive gravity field solutions will be available in a weekly interval, with a latency of 2 days after the availability of all input data. The maximum degree and order for the

QL-GFA gravity field models will be optimized with respect to the global coverage of the input data.

2.2. Core Solver (CS)

The objective of the CS is to compute a high-accuracy, high-resolution spherical harmonic model including a quality description of the static Earth's gravity field from GOCE SGG and SST observations. The parameterization of the model will be complete at least up to degree and order 200, and a resolution up to degree and order 250 is envisaged, depending on the actual accuracy of the SGG observations.

The *Tuning Machine*, whose development, implementation and integration is completely in the responsibility of the HPF work package partner University of Bonn, consists of two main modules:

- *pcgma* (pre-conditioned conjugate gradient multiple adjustment; [4]): This module acts as a stand-alone gravity field solution strategy, using the sparse structure of the normal equations [3], and is used to verify and tune the involved software components of the CS in many respects, e.g., to derive optimum regularization and weighting parameters.

- *Data analysis tool*: The data inspection and filter design tool is used to verify external and internal products, and to define the filter coefficients [27] which will be used in the Final Solver.

The *Final Solver* consists of the following main modules:

- *SST processor*: The information content of the SST data is exploited by making use of the precise GOCE orbit solutions (Precise Science Orbits; PSOs) expressed in terms of position and velocity information. The principle of energy conservation in a closed system is applied [7] [1] [2]. The software can process both kinematic (purely geometric) and reduced-dynamic (using a-priori models of the external forces and the gravity field) orbit solutions. However, kinematic orbit solutions are preferred, because they do not contain any a-priori information about the gravity field.

- *SGG processor*: The SGG processing is based on the position information given by the PSOs, and the externally calibrated gravity gradients defined in the Gradiometer Reference Frame (GRF). The complications arising from the coloured noise of the gradiometer are managed by a recursive filter procedure in time domain [24] [25] [26] [28] [16]. The SGG processor

assembles the full normal equations applying parallel processing on a Linux-PC-Cluster.

■ *Solver*: The mathematical models for SGG and SST data are combined to the overall mathematical model by means of superposition of the normal equations, applying variance component estimation [9] for an optimum weighting of the individual data types. The solution is processed applying a parallelized Cholesky reduction. The ill-posedness of the normal equations due to the polar gaps is managed by Spherical Cap Regularization [12] [13], a regularization technique which is tailored to the problem of the non-polar GOCE orbit configuration. Together with the GOCE gravity field model coefficients, a statistical error description in terms of the full mean square error matrix (variance-covariance matrix plus regularization bias) is processed.

2.3. Product Flow

Chronologically, the first processing steps will be performed by QL-GFA (cf. Fig. 1). The output products are not only transferred to the GOCE mission control via CPF, but are also used as prior information for the CS.

In the Core Solver processing, the SST and SGG normal equations are assembled separately. The SST normal equations (and other internal products) are transferred to the Tuning Machine and the Final Solver. In the Tuning Machine, the SGG normal equations are set-up using a sparse matrix scheme, and, after combination with the full SST normal equations, gravity field solutions are computed applying the pcgma algorithm. The residuals of the adjustment are analyzed by the data inspection tool, and filter coefficients, regularization and weighting parameters are derived, which are provided to the Final Solver. Here, the full SGG normal equations are assembled, and optimally combined with the SST normal equations. Finally, the gravity field coefficients and the mean square error matrix are computed rigorously.

3. Numerical case study

The operability of the software system shall be demonstrated by a numerical case study. In the present case, the main objective was the computation of optimum GOCE gravity field models complete to D/O 200.

3.1. Test data sets

The numerical case study is based on the data of an ESA GOCE end-to-end simulation [5]. This test configuration was also used during the official ESA Acceptance Review 2 for the testing of the final operational software (at the end of the development phase) in the framework of the HPF, which was performed in spring/summer 2006. The test data sets consist of:

■ *Gravity gradients*: 60 days of 1 Hz rate simulated gravity gradients defined in the GRF, based on the gravity model EGM96 [11] complete to degree/order 360, superimposed by colored noise (cf. Fig. 2, black curve).

■ *Orbit*: The gradients are defined along an orbit with GOCE characteristics (inclination $i = 96,5°$, eccentricity $e < 2 \cdot 10^{-3}$, mean altitude ~ 240 km). The orbit positions (and velocities) were generated by orbit integration, based on the gravity model EGM96, complete to degree/order 200, and including a full external force model and drag free and attitude control (DFAC) simulation. The DFAC performs the attitude control of the satellite and the compensation of non-conservative forces by means of ion thrusters [6].

■ *Attitude*: The orientation of the satellite body axes (and hence the GRF) with respect to the inertial frame is given in terms of quaternions, which are computed from a combination of star tracker and gradiometer information. Correspondingly, they include attitude biases and noise [15], related to the star tracker and gradiometer inaccuracies modelled in the end-to-end simulation.
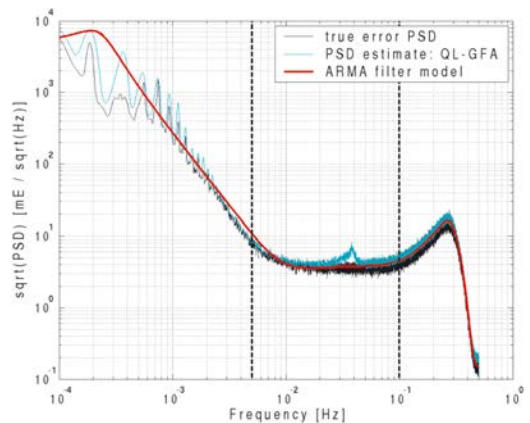
Figure 2: Gradiometer error PSDs for the gradiometer component V_{XX}: true PSD (black), QL PSD estimate (light blue), ARMA digital filter model (red); measurement bandwidth: 5 – 100 mHz.

3.2. Results: QL-GFA

Due to the limited space, this paper concentrates on the results of the Core Solver processing. The results of the QL-GFA, based on the test configuration described in section 3.1, are presented in a separate paper [20].

3.3. Results: Core Solver

In the following, the results of the main CS components will be presented. Selected issues of this processing, such as the regularization, the optimum filtering of the SGG observations and normal equations, the actual parameterization, etc. are addressed in [19] in more detail.

3.3.1. SST processing

The SST processing, which is based on the energy integral method, was applied to kinematic orbits. A numerical differentiation procedure using the Newton-Gregory method [1] [2] was applied to the kinematic orbit positions, which results in orbit velocities representing, after application of accelerometry to cope with the non-conservative forces, the basic pseudo-observations. The SST normal equations are set-up complete to degree/order 90, which turned out to be sufficient to finally obtain a smooth combined SST+SGG solution. The light blue curve in Fig. 3 shows the resulting SST-only solution in terms of the degree error median

$$\sigma_1 = median_m\left\{\left|\overline{R}_{lm}^{(est)} - \overline{R}_{lm}^{(EGM)}\right|\right\} \qquad (2)$$

where $\overline{R}_{lm} = \{\overline{C}_{lm}; \overline{S}_{lm}\}$ are the fully normalized spherical harmonic coefficients, (est) denotes the estimated quantities, and (EGM) refers to the reference model EGM96.
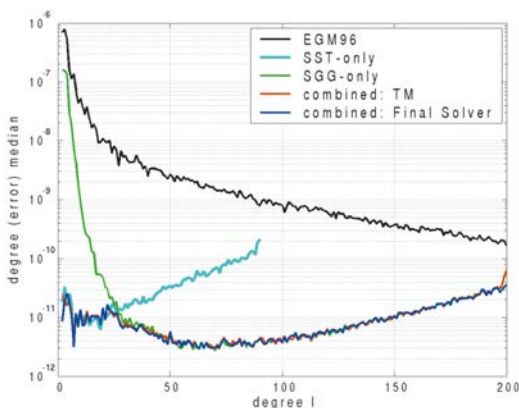
Figure 3: Degree (error) medians per degree of selected Core Solver gravity field solutions.

The corresponding SST normal equations are transferred to the Tuning Machine and the Final Solver, where they are used as input product.

3.3.2. Tuning Machine (TM)

One main task of the TM is the approximation of an appropriate SGG digital filter model to introduce the correct metrics to the SGG normal equation system [24] [25] [16] [28]. Fig. 2 shows the error PSD of the gravity gradient tensor component V_{XX} (black curve), and the corresponding filter model using a cascaded ARMA filter with an effective filter order of 52 (red curve). The other main diagonal tensor components V_{YY} and V_{ZZ} (not shown) have similar error characteristics. The corresponding cascaded filter models have an effective filter order of 42 (V_{YY}) and 32 (V_{ZZ}). (Additionally, the PSD derived by QL-GFA is displayed as light blue curve. For more details confer [20].)

A combined gravity field solution, based on the SST normal equation complete to degree/order 90 described above, and SGG normal equations complete to degree/order 200, was computed by pcgma. The red curve in Fig. 3 shows the results in terms of the degree error median.

3.3.3. SGG processing

The full SGG normal equations were assembled on a Linux-PC-Cluster, which was installed under the umbrella of the initiative "Scientific Supercomputing" at TU Graz. The key parameters of this Beowulf cluster are: 54 Dual-Xeon 2.6GHz PCs with 1–2 GB RAM, GigaBit-Ethernet connection, performance 210 GFlops.

The final goal of this simulation was an optimum gravity field solution complete to degree/order 200. Since the signal content of the SGG inputdata is degree/order 360, a spectral leakage effect due to the non-parameterized signals fromdegree 201 to 360 has to be expected. In [25], it was shown that the spectral leakage effect mainly affects the coefficients in the spectral region close to the upper limit of resolution (in the present case 200). Therefore, in order to reduce the effect, SGG normal equations complete to degree/order 204 are assembled, and the final solution is truncated at degree 200, thus eliminating the coefficients of degree 201 to 204, which absorb most of the unresolved high-frequency signals. During the assembling of the SGG normal equations, the digital filter model derived by the Tuning Machine was applied to the observations and the columns of the design matrix [24] [25] [16] [28].
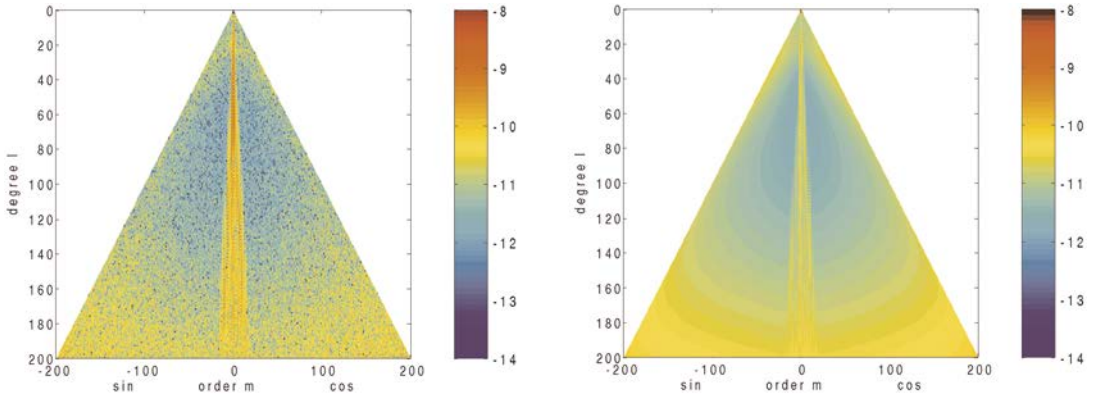
Figure 4: Combined gravity field solution: coefficient deviations from EGM96 (left) and MSE estimates (right). Scaled in $\log_{10}(|\ldots|)$.
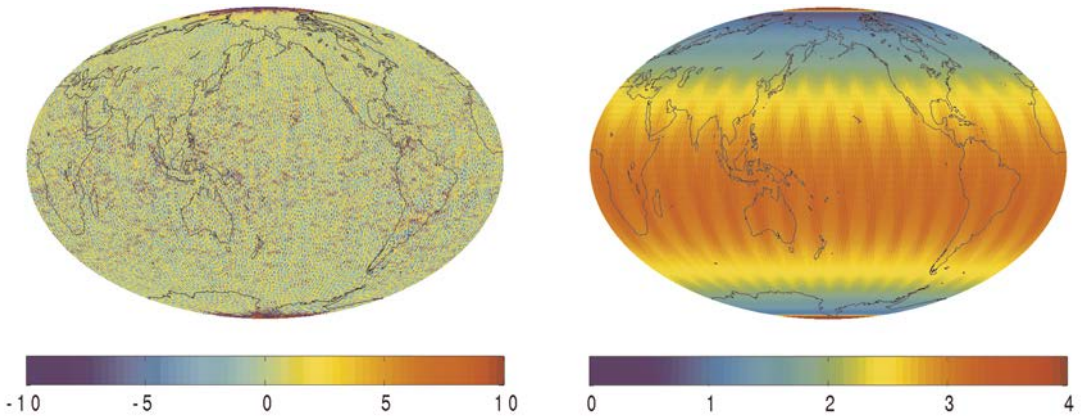
Figure 5: Combined gravity field solution: height anomaly deviations [cm] from EGM96 (left) and corresponding standard deviations (right), degree/order 200.

3.3.4. Final Solver

After the assembling of the SST (D/O 90) and the SGG (D/O 204) normal equations, they are superposed and solved by a rigorous parallel solver. The memory size of the upper triangle of the normal equations (double precision arithmetics) is about 6.5 GBytes for the degree/order 204 system. An optimum weighting based on variance-component estimation [9] among the individual normal equation systems was applied. The optimum weighting factor was computed by the Tuning Machine.

Finally, the large combined SST+SGG normal equation system, complete to degree/order 204, was solved rigorously, and afterwards truncated at degree/order 200 in order to reduce spectral

leakage. Fig. 4 shows the coefficient deviations from the reference gravity field model EGM96 (left), as well as the corresponding error estimates (right). Evidently, the absolute errors and the statistical error estimates are quite consistent, except of the (near-)zonals, whose accuracy is slightly overestimated.

The corresponding degree error median of this solution is displayed as dark blue curve in Fig. 3. Evidently, it is stabilized in the low-degree range mainly by the SST component (light blue curve), and dominated by SGG (green curve) from degree 25 onwards.

The combined solutions processed by the Tuning Machine (red curve) and the Final Solver (dark blue curve) show a very good agreement.

The fact that two independent methods and implementations obtain practically identical results supports the conclusion that the remaining coefficient errors are due to the noise of the input data, but are not produced by insufficiencies of the processing algorithms.

Based on the coefficient estimates of the combined solution of the Final Solver (Fig. 4, left), cumulative height anomaly errors at degree/order 200 have been processed, and are displayed in Fig. 5 (left).

The standard deviations of the height anomaly σ_ζ and gravity anomaly $\sigma_{\Delta g}$ difference fields in the latitudinal region $-83.5° < \varphi < 83.5°$, which is covered by GOCE observations, are $\sigma_\zeta = 2.93$ cm and $\sigma_{\Delta g} = 0.81$ mGal, respectively. Considering the fact that only 2 months of input data have been used in this simulation, it can be concluded, that under the presently made noise assumptions of the input products the GOCE mission specifications of $\sigma_\zeta^{(spec)} = 1 - 2$ cm height anomaly and $\sigma_{\Delta g}^{(spec)} = 1$ mGal gravity anomaly accuracy at degree/order 200 can be achieved. Further, it shall be emphasized that the solution presented in this paper is a GOCE-only solution in a strict sense, i.e., it does not contain any external gravity field models as prior information.

Together with the coefficient solution, also a full variance-covariance matrix (approx. 20 GBytes in ASCII format), complete to degree/order 200, was output of this processing. In order to prove the plausibility of this matrix, a rigorous covariance propagation was performed to propagate the coefficient errors to height anomaly errors on a global grid. Fig. 5 (right) shows the specific error structure of this field. Compared with the amplitude of absolute errors (Fig. 5, left), their statistical error estimates match very well, which proves consistency of this numerical closed-loop simulation study.

4. Summary and conclusions

In this paper the architectural design and the main modules of the SPF6000 are described. The software is now fully implemented, and the hardware and software system is integrated and has been tested in the frame of the "Acceptance Review 2" of the HPF project. The software is accepted by ESA for the operational GOCE gravity field processing. The data flow through the SPF6000 and the interplay of the system modules is described based on a numerical case study applying the official ESA test data, and the main output products are presented as an example for a multitude of test scenarios which have been processed to validate the software system extensively during the development phase. Finally, the SPF6000 is ready to process GOCE gravity field solutions, and it will become operable after the availability of the first real GOCE data.

References

[1] _Badura, T._: Gravity Field Analysis from Satellite Orbit Information applying the Energy Integral Approach. Dissertation, 109pp., TU Graz, 2006.

[2] _Badura, T., Sakulin, C., Gruber, C., and Klostius, R._: Derivation of the CHAMP-only global gravity field model TUG-CHAMP04 applying the energy integral approach. Stud. geophys. geod., 50, 59–74, 2006.

[3] _Boxhammer, Ch._: Effiziente numerische Verfahren zur sphärischen harmonischen Analyse von Satellitendaten. Dissertation, 95pp., University of Bonn, 2006.

[4] _Boxhammer, Ch., and Schuh, W.-D._: GOCE gravity field modeling: computational aspects - free kite numbering scheme. In: Rummel et al. (eds.): Observation of the Earth System from Space, 209–224, Springer, Berlin-Heidelberg, 2006.

[5] _De Sanctis, S, Guiotto, A., and Parisch, M._: Simulator Design Specification. Technical Note, GO-SP-AI-0001, Alenia Spazio, Turin, Italy, 2002.

[6] European Space Agency. Gravity Field and Steady-State Ocean Circulation Mission. Reports for mission selection, The four candidate Earth explorer core missions, SP-1233(1), European Space Agency, Noordwijk, 1999.

[7] _Földvary, L., Svehla, D., Gerlach, Ch., Wermuth, M., Gruber, T., Rummel, R., Rothacher, M., Frommknecht, B., Peters, T., and Steigenberger, P._: Gravity Model TUM-2Sp Based on the Energy Balance Approach and Kinematic CHAMP Orbits. In: Reigber et al. (eds.), Earth Observation with CHAMP-Results from Three Years in Orbit, 13–18, Springer Verlag, 2004.

[8] _Klees, R., Koop, R, Visser, P.N.A.M., and van den Ijssel, J._: Efficient gravity field recovery from GOCE gravity gradient observations. J. Geod., 74, 561–571, 2000.

[9] _Koch, K.-R., and Kusche, J._ :Regularization of geopotential determination from satellite data by variance components. J. Geod., 76, 259–268, 2002.

[10] _Lackner, B._: Datainspection and Hypothesis Tests of very long Time Series applied to GOCE Satellite Gravity Gradiometry Data. Dissertation, 187pp., Graz University of Technology, Graz, 2006.

[11] *Lemoine, F.G., Kenyon, S.C., Factor, J.K., Trimmer, R.G., Pavlis, N.K., Chinn, D.S., Cox, C.M., Klosko, S.M., Luthcke, S.B., Torrence, M.H., Wang, Y.M., Williamson, R.G., Pavlis, E.C., Rapp, R.H., and Olson, T.R.*: The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, 1998.

[12] *Metzler, B., and Pail R.*: GOCE data processing: the Spherical Cap Regularization Approach. Stud. Geophys. Geod., 49, 441–462, 2005.

[13] *Metzler, B., and Pail, R.*: Spherical Cap Regularization of GOCE normal equation systems. Proc. 3rd GOCE User Workshop, Frascati, November 2006, European Space Agency, Noordwijk, 2006.

[14] *Migliaccio, R., Reguzzoni, M., and Sansó, F*: Spacewise approach to satellite gravity field determinations in the presence of coloured noise. J. Geod, 78, 304–313, 2003.

[15] *Pail, R.*: A parametric study on the impact of satellite attitude errors on GOCE gravity field recovery. J. Geod., 79, 231–241, 2005.

[16] *Pail, R., and Plank, G.* :Assessment of three numerical solution strategies for gravity field recovery from GOCE satellite gravity gradiometry implemented on a parallel platform. J. Geod., 76, 462–474, 2002.

[17] *Pail, R., and Plank, G.*: GOCE Gravity Field Processing Strategy. Stud. Geophys. Geod., 48, 289–308, 2004.

[18] *Pail, R., and Wermuth, M.* :GOCE SGG and SST quick-look gravity field analysis. Advances in Geosciences, 1, 5–9, 2003.

[19] *Pail, R., Metzler, B., Lackner, B., Preimesberger, T., Höck, E., Schuh, W.-D., Alkathib, H., Boxhammer, Ch., Siemes, Ch., and Wermuth, M.*: GOCE Gravity Field Analysis in the Framework of HPF: Operational Software System and Simulation Results. Proc. 3rd GOCE User Workshop, ESRIN, Frascati, November 2006, SP-627, 249–256, European Space Agency, Noordwijk, 2007.

[20] *Pail, R., Metzler, B., Preimesberger, T., Lackner, B., and Wermuth, M.*: GOCE Quick-Look Gravity Field Analysis in the Framework of HPF. Proc. 3rd GOCE User Workshop, ESRIN, Frascati, November 2006, SP-627, 325–332, European Space Agency, Noordwijk, 2007.

[21] *Preimesberger, T., and Pail, R.*: GOCE quick-look gravity solution: application of the semianalytic approach in the case of data gaps and non-repeat orbits. Studia geoph. geod., 47, 435–453, 2003.

[22] *Rummel, R., van Gelderen, M., Koop, R., Schrama, E., Sansó, F., Brovelli, M., Miggliaccio, F., and Sacerdote, F*: Spherical harmonic analysis of satellite gradiometry. Neth. Geod. Comm., Publications on Geodesy, 39, Delft, The Netherlands, 1993.

[23] *Rummel R, Gruber, T., and Koop, R.*: High Level Processing Facility for GOCE: Products and Processing Strategy. Proc 2nd International GOCE User Workshop, Frascati, 2004.

[24] *Schuh, W.-D.*: Tailored Numerical Solution Strategies for the Global Determination of the Earth's Gravity Field, Mitteilungen geod. Inst. TU Graz, 81, Graz Univ. of Technology, Graz, 1996.

[25] *Schuh, W.-D.*: Improved modelling of SGG-data sets by advanced filter strategies. ESA-Project ''From Eötvös to mGal+'', Final Report, ESA/ESTEC Contract 14287/00/NL/DC, WP 2, 113–181, ESA, Noordwijk, 2002.

[26] *Schuh, W.-D.*: The processing of band-limited measurements; filtering techniques in the least squares context and in the presence of data gaps. In: Beutler et al. (eds.): Earth Gravity Field from Space – From Sensors to Earth Sciences, Space Science Reviews, 108, 67–78, 2003.

[27] *Schuh, W.-D., and Kargoll, B.*: The numerical treatment of the downward continuation problem for the gravity potential. In: Sanso (ed.): Hotine-Marussi-Symposium, IAG Symposia, Lecture Notes in Earth Sciences, 127, 22–31, Springer, Berlin-Heidelberg, 2004

[28] *Schuh, W.-D., Pail, R., and Plank, G.*: Assessment of different numerical solution strategies for gravity field recovery. Proc. 1st Internat. GOCE User Workshop, April 2001, WP-188, 87–95, ESTEC Noordwijk, European Space Agency, 2001.

Contact

a.o. Univ.Prof. Dr.techn. Roland Pail: Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria.
E-mail: pail@geomatics.tu-graz.ac.at   vgi

Determination of Earth rotation variations by means of VLBI and GPS and comparison to conventional models

Sigrid Englich, Paulo Jorge Mendes Cerveira, Robert Weber, Harald Schuh

Abstract

Oceanic and solid Earth tides induce periodic signals in the Earth rotation parameters (ERP), i.e., the pole coordinates (x_p, y_p) and universal time (UT1) or length of day (LOD), respectively. The oceanic tides cause variations with diurnal and semidiurnal periods in all parameters, whereas the zonal Earth tides mainly influence the rotational speed of the Earth, and thus UT1 and LOD. These signals show periods from $\sim$ 5 days to 18.6 years. Today, the ERP and their variations can be observed by modern space geodetic techniques with an unprecedented accuracy and temporal resolution. For the investigation of short and long period tidal effects ERP series of one year were computed from GPS observational data with hourly and 6-hours intervals. One ERP series with daily resolution was generated from six years of VLBI observations. The high-resolution GPS-based ERP series was analysed w.r.t. daily and sub-daily tidal variations. The variations in the VLBI-based dUT1 series and the GPS-based LOD series of lower temporal resolution were examined for periods from 5 to 32 days. The observed periods and corresponding amplitudes were compared to the theory of tidal excitation in order to reveal and, if possible, interpret remaining fluctuations in the ERP series. We retrieved significant signal peaks at the periods of the major tidal constituents, as well as at non-tidal periods from 5 to 32 days. The occurrence of the unexpected variations is supposed to be primarily a consequence of insufficient modelling of the atmospheric excitation.

Kurzfassung

Die Gezeiten der festen Erde und der Ozeane verursachen periodische Schwankungen der Erdrotationsparameter (ERP), d.h. der Polkoordinaten (x_p, y_p) und der Weltzeit (UT1), beziehungsweise deren zeitlicher Ableitung, der Tageslänge (LOD). Die Meeresgezeiten bewirken eintägige und halbtägige Variationen in allen Parametern, während die zonalen Erdgezeiten vorwiegend die Rotationsgeschwindigkeit der Erde und damit UT1 und LOD beeinflussen. Diese Effekte haben Perioden von $\sim$ 5 Tagen bis zu 18,6 Jahren. Die ERP und ihre Änderungen können heute in hoher zeitlicher Auflösung mit modernen geodätischen Weltraumverfahren beobachtet werden. Zur Betrachtung der kurz- und langperiodischen Gezeiteneffekte wurden ERP-Zeitreihen mit einstündiger und 6-stündiger Auflösung aus GPS-Beobachtungen eines Jahres generiert. Aus VLBI-Beobachtungsdaten von sechs Jahren wurden ERP, mit zeitlicher Auflösung von 24 Stunden berechnet. Die hochauflösende GPS-basierte ERP-Serie wurde bezüglich täglicher und subtäglicher Variationen analysiert. Schwankungen in der VLBI-basierten dUT1 und in der GPS-basierten LOD-Serie wurden auf Perioden von 5 bis 32 Tagen untersucht. Die beobachteten Perioden und die dazugehörigen Amplituden wurden mit theoretischen Modellen für die Anregung durch Gezeiten verglichen, um verbleibende Fluktuationen in den ERP-Serien aufzudecken und gegebenenfalls zu interpretieren. Signifikante Signale wurden sowohl mit Perioden der Hauptgezeitenterme, als auch mit nicht gezeitenbezogenen Perioden, im langperiodischen Bereich von 5 bis 32 Tagen, gefunden. Das Auftreten der unerwarteten Variationen ist vermutlich eine Folge unzureichender Modellierung der atmosphärischen Anregung.

1. Introduction

The rotational behaviour of the Earth is highly connected to various geophysical processes. We distinguish between two sets of parameters for describing Earth rotation. The direction of the rotation axis in an Earth-fixed system is expressed by two parameters x_p and y_p, their variation is called polar motion (PM). The parameter dUT1, which is the difference between universal time UT1 and UTC (universal time coordinated), gives access to variations of the rotational speed of the Earth. The three parameters x_p, y_p and dUT1 are

called Earth rotation parameters (ERP). Depending on the purpose, also the time derivative of dUT1, i.e., the length of day (LOD) is assigned to this parameter set. Earth rotation variations of geodynamic origin are typically studied by observing the Earth rotation parameters. The ERP together with precession-nutation form the Earth orientation parameters (EOP). The EOP are the linking elements between celestial and terrestrial reference frames, i.e., they are transformation arguments. This transformation could in principle be reduced to an arbitrary spatial

rotation, expressed by three time-dependent angles (so-called Euler angles) [1]. Since these angles would change very fast with time and because of the scientific purpose of studying geophysical phenomena, the transformation is favourably specified by a sequence of rotation matrices. The conventional systems used are the International Celestial Reference System (ICRS) and the International Terrestrial Reference System (ITRS) and their corresponding reference frames. These systems are established and maintained by the IERS (International Earth Rotation and Reference Systems Service). The two transformations in use are either based on the ecliptic and the equator, or the new transformation paradigm making use of the so-called NRO (Non-Rotating-Origin) according to the IAU-resolutions 2000 [2]. Before the adoption of the IAU resolutions 2000, polar motion referred to the so-called Celestial Ephemeris Pole (CEP). The CEP was defined such that there are no nearly diurnal motions of this celestial pole with respect to either the space-fixed or the body-fixed system. Along with the new transformation paradigm the CEP was replaced by the Celestial Intermediate Pole (CIP). Its definition is an extension of that of the CEP in the high frequency domain and coincides with that of the CEP in the low frequency domain (for the exact definition we refer to [3]). The concept of the CIP is realized within the IAU 2000A Precession-Nutation Model. The transformation based on ecliptic and equator is shortly described in the following because it is still used in most of the software packages related to the analysis of observation data from space geodetic techniques. In this representation the transformation is performed by the following sequence of rotation matrices [4]:

$$x_i(t) = P(t)N(t)R_3(-\theta(t))R_1(y_p(t))R_2(x_p(t))x_e(t) \tag{1a}$$

with $N(t) = R_1(-\varepsilon_A)R_3(\Delta\psi)R_1(\varepsilon_A + \Delta\varepsilon)$ (1b)

$x_i(t)$ and $x_e(t)$ are vectors in the celestial and in the terrestrial reference system, respectively. Both precession matrix $P(t)$ and nutation matrix $N(t)$ describe the change of the direction of the Earth axis in space; the matrix of sidereal time $R_3(-\theta)$ represents the actual rotation angle of the Earth, and the polar motion matrix $W = R_1(y_p(t))R_2(x_p(t))$ indicates the position of the Earth axis with respect to the Earth-fixed reference system, where $x_p(t)$ and $y_p(t)$ are the coordinates of the CIP in the Earth-fixed reference system (ITRS) at time t. The obliquity of the ecliptic is denoted by ε_A, and $\Delta\varepsilon$ and $\Delta\psi$ designate nutation in obliquity and longitude. The actual

rotation of the Earth around the CIP is given via the true sidereal time in Greenwich θ (Greenwich Apparent Sidereal Time: GAST). The true sidereal time is deduced from the mean sidereal time (GMST) considering nutation corrections. UTC is converted to UT1 by applying dUT1, which can be observed by space geodetic techniques. GMST can be derived from UT1 by a definite formula [5]. The five quantities x_p, y_p, UT1 − UTC = dUT1, $\Delta\varepsilon$ and $\Delta\psi$ represent the Earth orientation parameters mentioned above. It is necessary to evaluate this transformation between the terrestrial and the celestial reference system in all space-based geodetic analyses, because the observations are made between objects of which the coordinates vary in one system and change slowly, or not at all, in the other system. A stable realisation of the systems along with high-precision observation data from space geodetic techniques gives thus access to high-quality Earth rotation parameters and their variations.

The major effects, detectable in the ERP time series, result from deformations of the solid Earth, mass displacements and fluctuations in oceans and atmosphere and from the interactions between these systems. Mass displacements cause changes of the inertia tensor of the Earth (mass- or pressure-terms), whereas mass motions relative to the reference system alter the relative angular momentum (motion- or wind-terms). The Earth principal moments of inertia are moreover changed periodically due to the reactions of the oceans and the solid Earth to the tidal potential exerted by the Moon and Sun. These variations of the inertia tensor translate into the position of the rotation axis (PM) as well as in the Earth spin rate (LOD) and its total rotational phase angle (UT1).

In the following we describe the derivation of ERP time series from observations of the Global Positioning System (GPS) and of Very Long Baseline Interferometry (VLBI). Subsequently, the variations in the ERP series are examined for periods from a few hours up to 32 days, focusing on signals excited by solid Earth and ocean tides.

2. Determination of Earth rotation parameters

The determination of ERP is typically solved by means of least squares adjustment or Kalman-filtering. The Earth orientation parameters are estimated from the observations of space geodetic techniques, mostly together with numerous other parameters, such as satellite orbits, station coordinates or troposphere parameters.

Geodetic Parameters	VLBI	GPS, GLONASS	DORIS	SLR	LLR
Position of the Quasars	X				
Satellite orbit parameters		X	X	X	
Lunar orbit parameters					X
Nutation	X	X[*]			X[*]
Polar motion	X	X	X	X	X
UT1	X				X
Length of day	X	X	X	X	
Station coordinates, Station velocities	X	X	X	X	X

[*] ... only nutation rates can be determined by these techniques

Table 1: Parameters relating to Earth rotation from different space geodetic techniques (taken from [4])

GLONASS – Global'naya Navigatsionnaya Sputnikowaya Sistema
DORIS – Doppler Orbitography Radiopositioning Integrated by Satellite
SLR – Satellite Laser Ranging, LLR – Lunar Laser Ranging

Nutation parameters are thereby, in any case, estimated as corrections to an a priori model. The most current and accurate model is the IAU2000A precession-nutation model. There are essential distinctions between the space geodetic techniques to be considered, w.r.t. the determination of EOP, depending on whether the technique observes satellites or quasars. Because of the necessity to determine orbital elements simultaneously to the EOP, satellite techniques cannot determine dUT1 and nutation offsets directly, but only their temporal changes (rates) [6]. VLBI observations enable to directly estimate dUT1 and nutation offsets, but because of the lack of continuous observations it is more difficult to determine rates. Unfortunately, due to economic and logistic factors, a continuous monitoring of Earth rotation by VLBI is at present not feasible. Polar motion can be determined likewise from both of the mentioned techniques. Table 1 gives an outline of all space geodetic techniques and their ability to determine geodetic parameters related to Earth rotation.

2.1. GPS ERP series

A globally distributed network of stations is a prerequisite to derive ERP from GPS observations. Hence it is indispensable to use automated processing, in order to cope with the huge amount of data. All estimations were therefore performed with the Bernese Processing Engine (BPE) of Bernese GPS Software Version 5.0. The required observation data is kindly provided by the International GNSS Service (IGS) [7]. Our processing network consists of 113 carefully chosen and fairly stable stations, which belong to the IGS05 reference frame. The network, including the used standard baseline configuration is shown in figure 1. The red lines designate baselines ≤2000 km, which are accented here, because the ambiguity resolution is working best for baselines up to this length. By processing data of the year 2005 we derived a year-long ERP series. Since the processing of large GPS networks is very time consuming, and because of limited computing capacities, it was not yet possible to compute longer ERP time series. ERP were set up as continuous piecewise-linear functions at hourly and 6-hours intervals, and were solved for weekly solutions. In the Bernese GPS Software the ERP series are computed by default with respect to a sub-daily ERP model for ocean tide corrections. In our study the Eanes model was used, which is recommended in the IERS Conventions 2003 [8]. The geodetic datum definition uses a minimum constraint solution, imposing a no-net-rotation condition on the station coordinates from IGS05.snx [9]. As orbit information, final orbits of the CODE (Center for Orbit Determination in Europe [10]) were taken and held fixed in the least squares adjustment. We chose to take a priori ERP information from CODE for consistency reasons, since the CODE products were also processed with the Bernese GPS Software. All calculations were performed using absolute antenna phase center corrections. Site displacements due to
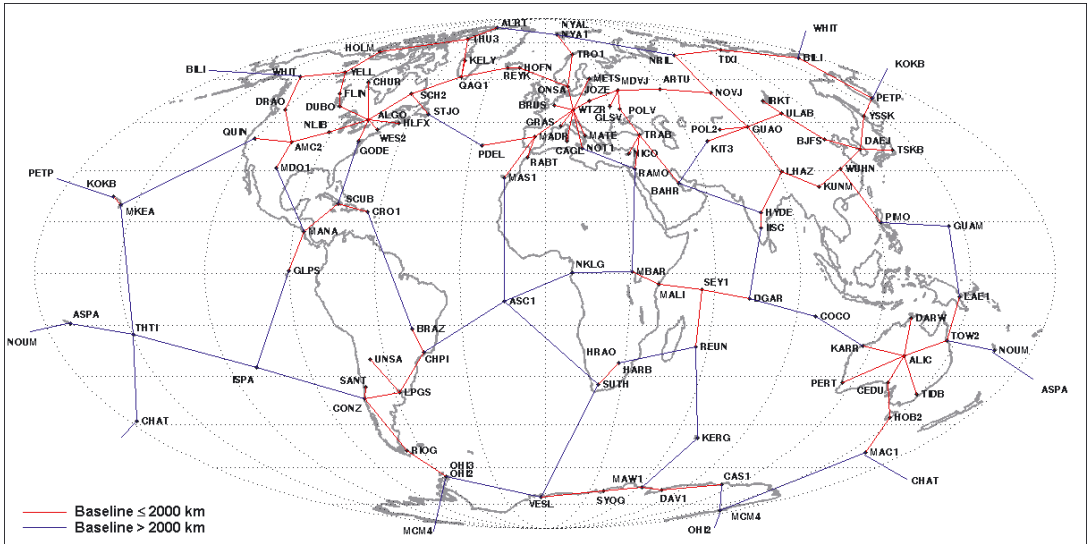
Figure 1: GPS processing network.

ocean loading were taken into account with the FES2004 model. Nutation was modelled using IAU2000A according to the IERS Conventions 2003.

2.2. VLBI ERP series

All estimations of VLBI-based ERP were computed with the VLBI software OCCAM Version 6.1. OCCAM is a research and development package for the analysis of geodetic and geodynamic VLBI experiments. The ERP series generated for this study uses the observation data of all geodetic 24-hour experiments carried out from the beginning of 2000 to the end of 2005 and the so-called Intensive sessions with duration of one hour. Sessions which are not suitable for reliable EOP determination due to their limited extension in one or more components were excluded, as suggested in [11] by the Analysis Coordinator of the IVS (International VLBI Service for Geodesy and Astrometry). The observation data of these measurements is provided by the IVS and available via ftp-download.

The applied reference frame for the VLBI-derived series is the global VLBI solution IGG05R01 [12] of the Institute of Geodesy and Geophysics in Vienna which is a Special Analysis Center of the IVS. Coordinates of stations and sources are fixed to the IGG05R01 frame. Hence the estimates are Earth orientation, troposphere, and clock parameters. Within the processing of the 24-hour experiments nutation offsets, PM and

dUT1 were set up once per session. When processing Intensive sessions, nutation and PM have to be fixed to a priori values and solely one dUT1 value is estimated per session. The same a priori nutation model used for the GPS solutions was also applied in VLBI analysis. Site displacements due to ocean loading were modelled based on the ocean tide model GOT00.2. An ERP-series with sub-daily resolution from VLBI observation data is not presented here because this approach demands for another processing strategy, which is still under development.

3. Investigation of Earth rotation variations

The objective of this study was to examine the variations of Earth rotation parameters induced by tidal fluctuations.

3.1. Diurnal and sub-diurnal tidal variations in PM and LOD

The driving forces for the Earth rotation variations of short periods are the changes in ocean heights and currents, which are driven by the tidal acceleration due to the Sun and Moon. Additionally, also atmospheric excitation of PM and dUT1 has to be expected at diurnal and semidiurnal frequencies. Since the excitation of Earth rotation by diurnal and semidiurnal ocean tides is up to two orders of magnitude larger than the corresponding atmospheric influence [13], this aspect will not be treated in the present work. As previously mentioned, sub-daily GPS-derived ERP were estimated w.r.t. a sub-daily model for ocean tide

corrections. Thus the model values were re-applied to the GPS-based ERP series for the analysis. In order to investigate daily and sub-daily variations, the low-frequency variations were removed by subtracting a smooth ERP series, based on the C04 series of the IERS. The Figures 2a-2c show amplitude spectra of the GPS-derived x_p and y_p series and of the GPS-derived LOD series. The spectra for the GPS ERP variations were generated using the CLEAN-algorithm according to [14].

The most distinct peaks in the frequency content of the GPS-based ERP variations are the eight major ocean tidal terms: Q1, O1, P1, and K1 in the diurnal and N2, M2, S2, and K2 in the semidiurnal band. The P1 and K2 peaks have been magnified in the pictures in order to visualise that these terms are separated from K1 and S2, respectively. In a first approach we estimated amplitudes of the eight major ocean tides from the set of GPS-derived ERP. The amplitudes were determined in a least squares adjustment using the following representation for the tidally driven variations:

$$\Delta X(t) = \sum_{j=1}^{n} [sx_j \sin \varphi_j(t) + cx_j \cos \varphi_j(t)] \quad (2a)$$

$$\Delta Y(t) = \sum_{j=1}^{n} [sy_j \sin \varphi_j(t) + cy_j \cos \varphi_j(t)]$$

$$\Delta LOD(t) = \sum_{j=1}^{n} [sl_j \sin \varphi_j(t) + cl_j \cos \varphi_j(t)]$$

$$\varphi_j(t) = \sum_{j=1}^{6} N_{ij} F_i(t) \quad (2b)$$

$\Delta X(t)$, $\Delta Y(t)$ and $\Delta LOD(t)$ denote the observed tidal variations in the corresponding quantities; the sx_j, cx_j, sy_j, cy_j, sl_j and cl_j are the corresponding sine and cosine amplitudes, while n specifies the number of tides considered. The angle argument $\varphi_j(t)$ is built as a linear combination of the five fundamental arguments and GMST $+ \pi$; as the sixth argument (see [8]). Each tide j is thereby characterized by a special sequence of integer multipliers N_{ij}.
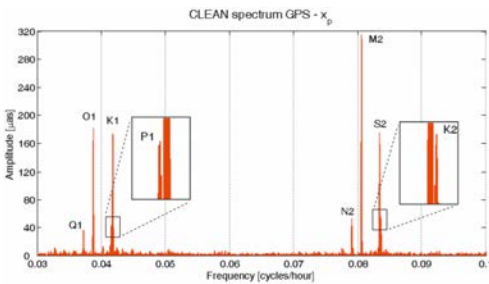
Figure 2a: CLEAN spectrum of the diurnal and sub-diurnal tidal frequency bands from GPS-derived x_p time series.
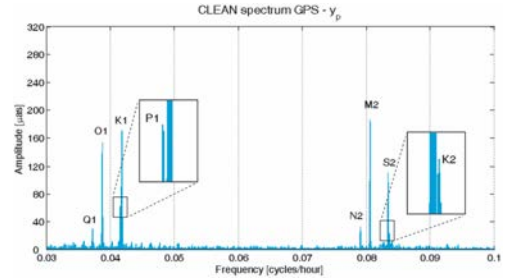
Figure 2b: CLEAN spectrum of the diurnal and sub-diurnal tidal frequency bands from GPS-derived y_p time series.
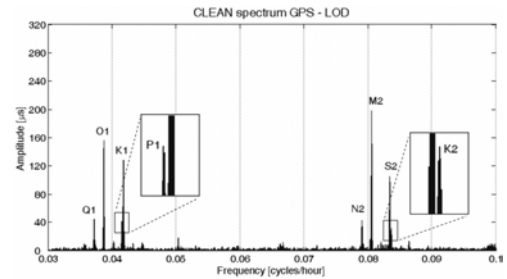
Figure 2c: CLEAN spectrum of the diurnal and sub-diurnal tidal frequency bands from GPS-derived LOD time series.

In the tables 2a–c the sine and cosine amplitudes of the tidal variations from our estimate are compared to the sum of the corresponding IERS amplitudes and the amplitudes of those neighbouring periods, which differ by one cycle only after a full revolution of the lunar node. The current IERS model for ERP variations caused by ocean tides is an extension of the model calculated by Ray [15] from Satellite Altimetry data. The modulus of the amplitudes estimated in the least squares adjustment agrees to the magnitude of the constituents found by the CLEAN-algorithm within three times the formal error of the adjustment. When we compare the estimated signal magnitudes to the values of the IERS model the overall agreement is about 50% (the last row of each table gives the difference of the absolute values of the observed amplitudes minus the combined model amplitudes). With three times the formal error as a criterion the least accordance between observations and model is found in the y_p variations, where accordance is given only for the Q1, N2 and M2 terms. For the x_p variations the values agree for the Q1, N2, M2 and K2 tides. The closest agreement arises for the LOD variations, where all terms except the O1 term meet the criterion. The most peculiar

Period [hours]	Tide	ΔX (observed) Amplitude of				ΔX (IERS) Comb. amplitude of		Diff. |Amp| obs-IERS
		sin [μas]		cos [μas]		sin [μas]	cos [μas]	[μas]
26.87	Q1	8.1	±2.4	36.1	±2.4	7.4	31.3	4.9
25.82	O1	53.8	±2.4	173.9	±2.4	58.0	158.0	13.7
24.07	P1	26.8	±2.4	38.8	±2.4	25.8	50.6	−10.2
23.93	K1	−97.9	±2.4	−152.8	±2.4	−86.5	−169.3	−8.6
12.66	N2	−54.6	±2.4	−3.8	±2.4	−54.8	−12.4	−1.4
12.42	M2	−322.1	±2.4	−21.7	±2.4	−317.9	−26.0	3.9
12.00	S2	−155.8	±2.3	79.7	±2.4	−144.1	63.6	17.5
11.97	K2	−60.7	±2.4	4.5	±2.4	−49.9	24.9	5.1

Table 2a: *Major ocean tidal terms for ΔX (observed) vs. major ocean tidal terms + sideband terms for ΔX according to IERS.*

Period [hours]	Tide	ΔY (observed) Amplitude of				ΔY (IERS) Comb. amplitude of		Diff. |Amp| obs-IERS
		sin [μas]		cos [μas]		sin [μas]	cos [μas]	[μas]
26.87	Q1	−29.7	±2.5	4.8	±2.5	−31.3	7.4	−2.1
25.82	O1	−148.1	±2.5	42.5	±2.5	−158.0	58.0	−14.3
24.07	P1	−60.7	±2.5	23.5	±2.5	−50.6	25.8	8.3
23.93	K1	153.2	±2.4	−97.3	±2.5	169.3	−86.5	−8.6
12.66	N2	10.8	±2.5	32.0	±2.5	10.7	30.8	1.2
12.42	M2	50.8	±2.5	183.5	±2.5	36.2	188.6	−1.7
12.00	S2	68.1	±2.4	90.8	±2.5	59.2	86.6	8.6
11.97	K2	8.4	±2.5	22.8	±2.5	23.0	30.0	−13.5

Table 2b: *Major ocean tidal terms for ΔY (observed) vs. major ocean tidal terms + sideband terms for ΔY according to IERS.*

Period [hours]	Tide	ΔLOD (observed) Amplitude of				ΔLOD (IERS) Comb. amplitude of		Diff. |Amp| obs-IERS
		sin [μas]		cos [μas]		sin [μas]	cos [μas]	[μas]
26.87	Q1	−13.5	±3.6	−42.9	±3.6	−16.7	−34.1	7.0
25.82	O1	−86.8	±3.6	129.9	±3.6	−83.8	−111.3	17.0
24.07	P1	−22.3	±3.6	−36.4	±3.6	−19.2	−34.2	3.5
23.93	K1	63.2	± 3.6	116.6	± 3.6	60.1	123.9	−5.1
12.66	N2	−18.2	± 3.6	40.6	± 3.6	−17.9	43.5	−2.5
12.42	M2	−90.6	± 3.6	179.8	± 3.6	−83.6	189.2	−5.6
12.00	S2	3.0	± 3.6	105.6	± 3.6	−2.0	94.8	10.8
11.97	K2	−5.5	± 3.6	32.0	± 3.7	0.7	34.4	−2.0

Table 2c: *Major ocean tidal terms for ΔLOD (observed) vs. major ocean tidal terms + sideband terms for ΔLOD according to IERS.*

difference results in the variations of all three parameters for the O1 term, which is with the exception of x_p, larger than for the other terms. It is not probable that these differences are technique specific artefacts since the regarded tide is not close to any resonance frequency of the satellite revolution period. To what extend the observed differences are deficiencies in the IERS model has to be investigated on the basis of a longer time series and compared to results of a second time series, gained from another technique, such as VLBI. For the future, it is intended to scrutinize the ERP variations (derived from both techniques) for non-tidal signals as well, and to solve for an extended set of ocean tide amplitudes, including the sidebands. Therefore it will be necessary to constrain the ratio of the sideband amplitude and the amplitude of the major tide (see [16] for further information).

3.2. Long-period tidal variations in dUT1/LOD

The second point of interest of our work is the impact of zonal Earth tides on the rotation of the Earth. The deformations of the Earth, caused by the zonal part of the tidal potential, periodically change the principal moments of inertia. This leads in accordance to the conservation of angular momentum to periodic changes in the rotation speed and thus in dUT1 and LOD. The variations show periods from ∼5 days to 18.6 years. For the investigation of these effects we used a continuous one-year GPS LOD series with a 6-hour resolution and a six year VLBI dUT1 series, given in daily intervals, except for time periods where no measurements were available. We estimated the amplitudes of 10 terms from the LOD series and of 12 terms from the dUT1 series at fixed periods. Periods are given in units of days: 6.86, 7.10, 9.13, 9.56, 12.81, 13.66, (13.78), 14.77, 23.94, (27.09), 27.56, 31.81. The amplitudes for the periods in braces were estimated from the dUT1 series only, because these terms would not be separable from the one-year LOD time series. One term (at 12.81 days) was added with the amplitude expected to be around zero, for testing the correctness of the approach. The observation equations were set up quite similar as for the estimation of the ocean tide amplitudes, with the exception that only the five fundamental arguments are used in this approach. The representation reads as follows:

$$\delta UT1(t) = \sum_{j=1}^{n} [B_j \sin \varphi_j(t) + C_j \cos \varphi_j(t)] \qquad (3a)$$

$$\delta LOD(t) = \sum_{j=1}^{n} [B'_j \cos \varphi_j(t) + C'_j \sin \varphi_j(t)]$$

$$\varphi_j(t) = \sum_{j=1}^{5} N_{ij} F_i(t) \qquad (3b)$$

Actually we solved only for the B amplitudes and neglected the so-called out-of-phase terms, which are denoted as C amplitudes here, because their magnitude is close to zero, according to the model currently recommended by the IERS [17].

For the derivation of the amplitudes corresponding to periods <32 days it was necessary to filter tidal signals of longer periods and strong atmospheric signals from the original data. The influence of the atmosphere was reduced using atmospheric angular momentum data of the NCEP (U.S. National Centers for Environmental Prediction) [18]. Additionally we subtracted the term at 13.63 days, since it is of reasonable size and strongly correlates to the term at 13.66 days. At least 18.6 years of data would be required to separate these two periods. The remaining low-frequency signal was removed by applying a highpass filter. Figures 3a and 3b display the frequency content of the LOD and dUT1 signals after the filtering process.
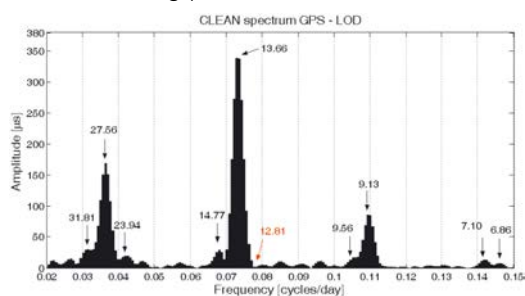
Figure 3a: CLEAN spectrum of filtered GPS-derived LOD time series with periods given in days.
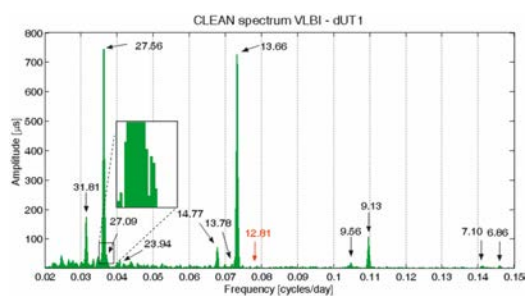
Figure 3b: CLEAN spectrum of filtered VLBI-derived dUT1 time series with periods given in days.

It can be seen that the 9 resp. 11 terms sought for are existent in the spectra and that there is almost no signal at the period 12.81 days with the expected zero amplitude. There is small signal content remaining also at non-tidal periods, such as between 10–12 days, 15–22 days and over 32 days. These peaks are most probably the results of insufficient modelling of the atmospheric influence.

Both, the GPS-derived LOD variations and the VLBI-derived dUT1 variations were additionally subjected to a second spectral analysis based on a heuristic method [19]. The purpose of the heuristic analysis is not to estimate the amplitudes at predefined periods but to determine the periods with the largest amplitudes. Calibrated for the search of 9 periods (in the range of 5 to 32 days), the algorithm found amplitude maxima in the LOD variations at the following periods: 31.81, 31.63, 27.37, 23.66, 14.74, 14.17, 13.65, 9.47 and 9.12 days. When restricting the search range to below 8 days also peaks at the periods 7.06 and 6.44 days can be found. But they are obviously of less energy than the unexpected peaks at 23.66 and 14.17 days. A small tidal constituent would be expected at 23.94 days, but it is already visible in the amplitude spectrum (Figure 3a) that there is significantly more signal around this period than predicted from the IERS model. The origin of this oversized peak is not quite clear, but we assume that it is also due to insufficient modelling of atmospheric influence. The detected peak at 14.17 days is possibly the result of an aliasing of the residual diurnal tidal signature of the O1 tide to this period. This effect has been observed in GPS height time series so far [20] and has to be investigated for LOD series in further studies.

With the constraint to search for 11 peaks (between 5 and 32 days), the following periods resulted from the heuristic analysis of the dUT1 variations: 31.79, 29.89, 28.84, 27.91, 27.54, 26.93, 24.56, 14.76, 13.78, 13.66 and 9.13 days. Again, when we limit the search range to below 10 days also peaks at 9.54, 7.09 and 6.86 days are found. As the algorithm detects periods corresponding to amplitude maxima we have to state that these results are congruent to the CLEAN spectral analysis (Figure 3b), where we clearly see the signal content between 30 and 27.6 days. The occurrence of these peaks and also of the peak at 24.56 days is also supposed to be the consequence of the subtraction of mismodeled atmospheric excitation.

Table 3 shows our results of the least squares estimation on the left hand side. On the right side the zonal tidal terms for the estimated periods plus neighbouring periods are given according to the IERS Conventions 2003 [8]. The estimated terms result as a combined amplitude of the principal and the sideband terms. With three times the estimation error as a criterion and taking into account also the contribution of the sideband amplitudes the majority of the estimated terms agrees with the values of the IERS model. In both cases this does not hold for the terms 9.13 and 13.66 days, where the estimated amplitudes are smaller than the combined amplitudes predicted from the model. For the dUT1 variations also the estimated terms at 27.09, 27.56 and 31.81 do not match the combined terms of the IERS model. This mismatch, especially in the case of the dUT1 variations where we investigated a time series of six years, gives rise to the question if it is adequate to compare the estimates simply with the sum of the amplitudes of the considered periods and their sidebands. It might be more appropriate for the analysis of a time series of this length to constrain the sideband amplitudes and estimate them as well. The smaller size of the fortnightly terms could be a consequence of not reducing the long-period ocean tide (treated for example in [21]), which is existent at the same period but with a different phase.

VLBI δUT1			IERS δUT1		GPS δLOD			IERS δLOD	
Period [d]	Sin [μs]	RMS [μs]	Period [d]	Sin [μs]	Period [d]	Cos [μs]	RMS [μs]	Period [d]	Cos [μs]
			6.85	-4.0				6.85	3.8
6.86	-13.9	3.9	6.86	-10.0	6.86	9.5	2.5	6.86	9.1
			7.09	-5.0				7.09	4.5
7.10	-10.3	3.9	7.10	-12.0	7.10	11.9	2.5	7.10	10.9
			9.12	-41.0				9.12	28.4
9.13	-102.9	3.9	9.13	-100.0	9.13	87.4	2.5	9.13	68.5
			9.54	-8.0				9.54	5.4
9.56	-19.3	3.9	9.56	-20.0	9.56	15.8	2.5	9.56	13.0
12.81	8.5	3.9	12.81	2.0	12.81	-2.7	2.5	12.81	-1.1
13.66	-740.3	3.9	13.66	-779.0	13.66	349.1	2.5	13.66	358.4
13.78	-26.5	3.9	13.78	-34.0					
			14.73	5.0				14.73	-2.0
14.77	-75.1	3.9	14.77	-74.0	14.77	28.9	2.5	14.77	31.4
			14.80	-5.0				14.80	2.2
			23.86	5.0				23.86	-1.3
23.94	6.0	3.9	23.94	10.0	23.94	-3.8	2.5	23.94	-2.6
			26.98	18.0					
27.09	37.5	3.9	27.09	44.0					
			27.44	54.0				27.44	-12.3
27.56	-774.7	3.9	27.56	-833.0	27.56	168.6	2.5	27.56	189.9
			27.67	55.0				27.67	-12.5
			31.66	12.0				31.66	-2.4
31.81	-171.6	3.9	31.81	-184.0	31.81	31.3	2.5	31.81	36.3
			31.96	13.0				31.96	-2.6

Table 3: Zonal tidal terms.

4. Conclusions and Outlook

We investigated Earth rotation variations at tidal frequencies on the basis of GPS- and VLBI-derived Earth rotation parameters. The estimated ocean tide amplitudes are regarded as preliminary results, since fairly longer time series would be necessary to provide robust and reliable solutions. A comparable parallel computation of the ocean tidal terms from VLBI-derived ERP would essentially improve the reliability. Before this intention can be put into practice though, we will still have to focus on the derivation of high-quality sub-daily ERP from VLBI observations. Concerning the obtained zonal tidal terms for LOD and dUT1 we conclude that the magnitudes agree quite well with the values predicted by the IERS model. But as mentioned above, the way of comparison has to be reconsidered and the adjustment model needs to be modified accordingly. Additionally, it is intended to investigate the effect of the long-period ocean tides. In general it is definitely desirable to refine the method of removing non-tidal signals, such as the atmospheric influence and to generate ERP time series which allow for the separation of all adjacent periods and the estimation of periods ›32 days. Thus, it would be possible to avoid affecting the observed signal by a priori model information and thereby to resolve the source of the hitherto inscrutable peaks at periods of around 30 and 28 days in the dUT1 variations from VLBI and around 24 days in the LOD variations from GPS. Interesting conclusions will be feasible also from the analysis of combined GPS-VLBI ERP series, an important objective of our future work.

Acknowledgements

References

[1] *Moritz H. and Mueller I.I. (1987)*: Earth Rotation, Theory and Observation. The Ungar Publishing Company, New York.

[2] IAU-Resolutions 2000, http://syrte.obspm.fr/IAU_resolutions/Resol-UAI.htm

[3] *Capitaine N. (2000)*: Definition of the Celestial Ephemeris Pole and the Celestial Ephemeris Origin. Towards models and constants for sub-microarcsecond astrometry, Proceedings of IAU Colloquium 180 held at the U.S. Naval Observatory, Washington, DC, USA, 27-30 March 2000, Washington, DC: U.S. Naval Observatory, 2000 xix, 427 p. Edited by Kenneth J. Johnston, Dennis D. McCarthy, Brian J. Luzum, and George H. Kaplan., p. 153

[4] *Schuh H., Dill R., Greiner-Mai H., Kutterer H., Müller J., Nothnagel A., Richter B., Rothacher M., Schreiber U. and Soffel M. (2003)*: Erdrotation und globale dynamische Prozesse. Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Band 32. Verlag des Bundesamtes für Kartographie und Geodäsie, Frankfurt am Main.

[5] *Schödlbauer A. (2000)*: Geodätische Astronomie. Walter de Gruyter, Berlin.

[6] *Rothacher M., Beutler G., Herring T.A. and Weber R. (1999)*: Estimation of Nutation using the Global Positioning System. Journal of Geophysical Research, Vol. 104, NO. B3, p. 4835–4859.

[7] IGS: http://igscb.jpl.nasa.gov/

[8] IERS Conventions (2003). Dennis D. McCarthy and Gérard Petit. (IERS Technical Note; 32). Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.

[9] ftp://igscb.jpl.nasa.gov/igscb/station/coord/IGS05.snx

[10] ftp://ftp.unibe.ch/aiub/CODE/

[11] http://vlbi.geod.uni-bonn.de/IVS-AC/data/exclude.txt

[12] *Heinkelmann R., Boehm J, Schuh H. and Tesmer V. (2006)*: Global VLBI solution IGG05R01, in IVS 2006 General Meeting Proceedings, edited by D. Behrend and K.D. Baver, NASA/CP-2006-214140, 2006.

[13] *Brzezinski A., Bizouard Ch. and Petrov S.D (2002)*: Influence of the Atmosphere on Earth Rotation: what new can be learned from recent atmospheric angular momentum estimates? Surveys in Geophysics, 23, p. 33–69.

[14] *Baisch S. and Bokelmann G.H.R. (1999)*: Spectral analysis with incomplete time series: an example from seismology. Computers & Geosciences, 25, p. 739–750.

[15] *Ray R.D., Steinberg D.J., Chao B.F., Carthwright D.E. (1994)*: Diurnal and Semidiurnal Variations in the Earth's Rotation Rate Induced by Oceanic Tides. Science, 264, p. 830–832.

[16] *Gipson J.M (1996)*: Very long baseline interferometry determination of neglected tidal terms in high-frequency Earth rotation variation. Journal of Geophysical Research, Vol. 101, NO. B12, p. 28,051–28,064.

[17] *Defraigne P. and Smits I. (1999)*: Length of day variations due to zonal tides for an inelastic earth in non-hydrostatic equilibrium. Geophysical Journal International, 139, p. 563–572.

[18] *Salstein D.A. and Rosen R.D. (1997)*: Global momentum and energy signals from reanalysis systems. 7th Conf. on Climate Variations, American Meteorological Society, Boston, MA, p. 344–348.

[19] *Mautz R. (2001)*: Zur Lösung nichtlinearer Ausgleichungsprobleme bei der Bestimmung von Frequenzen in Zeitreihen. Dissertation, Fachbereich 9 – Bauingenieurwesen und Angewandte Geowissenschaften der Technischen Universität Berlin.

[20] *Penna N.T. and Stewart M.P. (2003)*: Aliased tidal signatures in continuous GPS height time series. Geophysical Research Letters, Vol. 30, NO. 23, 2184, doi:10.1029/2003GL018828.

[21] *Kantha L.H. and Stewart J.S. (1998)*: Long-period lunar fortnightly and monthly ocean tides. Journal of Geophysical Research, Vol. 103, NO. C6, p. 12639–12647.

Contact

Dipl.-Ing. Sigrid Englich, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: senglich@mars.hg.tuwien.ac.at

Dr. Paulo Jorge Mendes Cerveira, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: mendes@mars.hg.tuwien.ac.at

Dr. Robert Weber, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: rweber@mars.hg.tuwien.ac.at

Prof. Harald Schuh, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: hschuh@mars.hg.tuwien.ac.at

The instantaneous Earth rotation – still inaccessible?

Paulo Jorge Mendes Cerveira, Robert Weber and Harald Schuh

Abstract

Nowadays, space geodesy, such as Very Long Baseline Interferometry (VLBI) and Global Navigation Satellite System (GNSS) and Satellite Laser Ranging (SLR), allows determining Earth orientation to a fraction of 1 milliarcsecond with daily to hourly resolution. This should not prevent us from studying other innovative and powerful technologies. A new emerging technology, called ring laser gyroscope, is a high-precision tool that provides us with extra information in the daily and sub-daily time domain. The experimental determination of the amplitudes of the forced diurnal polar motion is for example exclusively allocated to the ring laser technique. Another aspect, in which ring lasers could emphasize their supremacy, is the determination of the motion of the Earth-fixed frame w.r.t. the instantaneous Earth rotation axis. However, present ring lasers are huge constructions extremely sensitive to external effects, e.g., temperature variations.

This paper illuminates the relationship between various Earth rotation axes, in the large sense, and discusses the separation between polar motion and nutation of these axes in space and w.r.t. the Earth body. The second part covers the expected benefit of ring laser observables.

Kurzfassung

Die geodätischen Weltraumtechniken, wie VLBI, GNSS und SLR, erlauben heute die Bestimmung der Orientierung einer mittleren Polachse der Erde im Raum und relativ zum Erdkörper mit einer Genauigkeit einiger Zehntel Millibogensekunden und einer zeitlichen Auflösung im Stundenbereich. Dieser offensichtliche Erfolg sollte uns aber nicht daran hindern, auch neue innovative Techniken zur Bestimmung der Erdrotation, wie z.B. Ringlaser, näher zu untersuchen. Ringlaser sind hochpräzise Instrumente, die uns speziell bei subtäglicher Auflösung erlauben, die Bewegung der festen Erdkruste relativ zur wahren Rotationsachse zu verfolgen. Es muss allerdings bemerkt werden, dass die Ringlaser große bauliche Anlagen erfordern und äußerst sensibel auf Einflüsse wie Temperaturänderung etc. reagieren.

Der vorliegende Artikel beleuchtet den Zusammenhang der unterschiedlichen Erddrehachsen und diskutiert demgemäss die in Polbewegung und Nutation aufgespaltenen Bewegungen dieser Achsen im Raum und relativ zum Erdkörper. Speziell werden dann im zweiten Teil die neuen von den Ringlasern zu erwartenden Messgrößen behandelt.

1. Fundamentals

Generally, two fundamental frames of reference are used for studies of Earth rotation: the quasi-inertial celestial coordinate system tightened to the directions of extragalactic sources, and the terrestrial coordinate system attached to several observatories, perpetually in motion, located on the Earth surface [20]. The first realization holds the acronym ICRF (International Celestial Reference Frame), whilst the second is called ITRF (International Terrestrial Reference Frame) [1]. The study of Earth rotation, in its basics, is already rather complicated. But, in the case of the Earth composed of an atmosphere, oceans, crust, mantle, liquid outer and solid inner core [11, 13, 15], things turn out to get extremely intriguing. In that case, the concept of Earth rotation, as a whole, is deprived of its initial sense. Which axis of rotation is meant, when the terminology "Earth rotation" appears in literature, and what are the underlying reference frames? In theoretical derivations, the perturbed instantaneous rotation vector is usually described in the "Tisserand mean-mantle" frame. In this frame, the relative angular momentum of the crust and mantle vanishes, while the one of the atmosphere, oceans and core does not. Another frame, introduced by Darwin G.H. [10], is chosen such that the diagonal components of the Earth inertia tensor keep constant over time [16]. Such frames considerably simplify the theoretical study of geophysical influences on Earth rotation, but how are such frames realized? Do such frames allow for a convenient comparison of excitation functions to astronomical and geodetic observations?

For some decades, space geodetic techniques were the primary means to monitor Earth rotation variations [2, 23, 26, 32]. Let us first recapitulate some basic notions relevant for discussing Earth rotation displayed in Figure 1.

If we set external torques (due to the presence of external bodies of our solar system) acting on the Earth equal to zero, the angular momentum vector $\vec{L}$ remains constant over time w.r.t. an inertial frame. Hence we assume that the Earth system, i.e., without the Moon, is a closed and conservative system. Under this hypothesis, the conservation of angular momentum admits an exchange between the Earth inertia tensor Θ, the rotation vector $\vec{\Omega}$ and the relative angular momentum $\vec{H}$. Figure 1 shows the relationship between these three vectors in the figure axis system ($\vec{x}$, $\vec{z}$) [19]. The $\vec{z}$-component is directed towards the Earth greatest moment of inertia, while the $\vec{x}$-component is pointed towards the Earth smallest moment of inertia. The direction of the pole of the Earth-fixed frame is denoted by $\vec{P}_{ITRF}$, i.e., the pole of the most recent ITRF. If the length of day (LOD) increases, hypothesizing a constant inertia tensor, the relative angular momentum increases too. Fortunately, the inertia tensor absorbs most of the surplus angular momentum, leaving relative angular momentum quasi unaffected. Space geodetic techniques usually observe an intermediate axis, different from the Earth rotation axis with a time resolution of one day and even shorter.
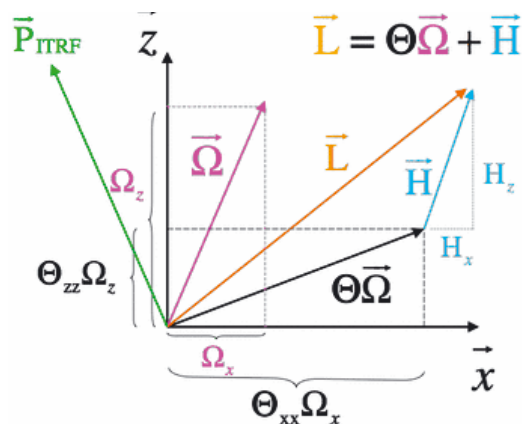
Figure 1: Relationship between the Earth angular momentum $\vec{L}$, the inertia tensor Θ, the rotation vector $\vec{\Omega}$ and the relative angular momentum $\vec{H}$ w.r.t. the principal unit axes of inertia $\vec{x}$ and $\vec{z}$.

Recently, the emerging ring laser technology, making use of the Sagnac effect, developed rapidly and provides the instantaneous Earth rotation vector without looking outside the Earth [8, 25]. The Sagnac effect is obtained by the frequency difference between counter-rotating beams, when the system or Earth rotates.

2. The motion of the Earth rotation axis

The motion of the rotation axis of an elastic Earth in space has been first modeled by W. Schweydar in 1917 [27]. Since then, Earth models evolved, and this motion has been modeled semi-analytically (IAU 2000) to an accuracy of less than 200 μas [17, 31, 6]. Recently, Krasinsky developed this motion to near observational perfection groping the 100 μas accuracy level (ca. 3 mm on Earth surface), for daily values [14]. Figure 2 displays the most relevant poles of reference to Earth rotation and will be essential in the understanding of subsection 2.1. These poles are: the pole of the ICRF, the initial Celestial Intermediate Pole CIP (CIP_0), the final CIP, and the pole of the ITRF. An unsolved task is to predict the remaining offsets between the celestial pole CIP_0 and the CIP. The Earth rotation pole is located close to the CIP: its formulation w.r.t. the CIP_0 will be presented in subsection 2.1. On the other hand, the motion of the non-rigid Earth w.r.t. the Earth rotation axis, is still unpredictable at the 200 milliarcsecond (mas) level [12, 34], and is therefore perpetually monitored. Annual meteorological effects drive the yearly signal of this motion, while the prediction of the free component of polar motion, called Chandler Wobble (CW), remains a topic under investigation since more than one century. In recent literature, the description "free nutation" is reserved for other resonant modes such as the Free Core Nutation (FCN). This terminology is also consistent with the definitions provided by the IERS Conventions 2003 (see Figure 4). Discussions and papers on the multiple-peaks spectrum of the CW have existed in the past, and recent publications remind us of this possible feature [21]. This splitting of frequencies would at least explain the great Chandler Wobble change during the years 1930–1940. However, it should be noted that the multiple-peaks model is only one possible way of describing the time-variable CW signal. Moreover, this model is rather controversial because it is based solely on data analysis. So far, nobody could give any physical reasoning for it. The great change of the CW in the 1930's can be interpreted also within other models as shown in [5, 29]. A great recent progress in the excitation studies of this free wobble has been attained using the model in which the CW signal is described as the free oscillator with damping excited by the random processes taking place in the dynamically coupled system atmosphere-oceans [33, 13].
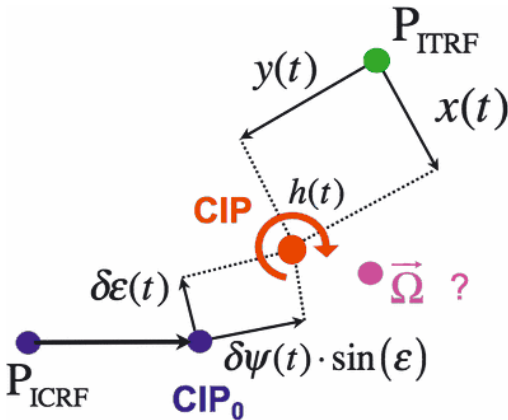
Figure 2: Representation of relevant poles of reference to Earth rotation.

2.1. Polar motion and universal time vs. instantaneous rotation vector

Reminding Figure 2, the rotation matrix $B(t)$, which in our restricted case accounts only for polar motion, $x(t)$ and $y(t)$, and the hour angle of the true equinox of date $h(t)$, transforms vectors from a terrestrial (TRS) to an intermediate reference system [12], i.e., the CIP:

$$X(t) = \begin{pmatrix} \cos[x(t)] & 0 & -\sin[x(t)] \\ 0 & 1 & 0 \\ \sin[x(t)] & 0 & \cos[x(t)] \end{pmatrix} \quad (1)$$

$$Y(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos[y(t)] & \sin[y(t)] \\ 0 & -\sin[y(t)] & \cos[y(t)] \end{pmatrix} \quad (2)$$

$$h(t) = \int \Omega_0 \cdot (1 + z(t))\, dt \quad (3)$$

$$S(t) = \begin{pmatrix} \cos[h(t)] & -\sin[h(t)] & 0 \\ \sin[h(t)] & \cos[h(t)] & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

$$B(t) = S(t) \cdot X(t) \cdot Y(t) \quad (5)$$

The skew-symmetric tensor $\Omega(t)$ allows recovering the instantaneous Earth rotation vector w.r.t. the TRS:

$$\Omega(t) = B^T(t) \cdot \delta_t B(t) = \begin{pmatrix} 0 & \Omega_z(t) & -\Omega_y(t) \\ -\Omega_z(t) & 0 & \Omega_x(t) \\ \Omega_y(t) & -\Omega_x(t) & 0 \end{pmatrix} \quad (6)$$

In deriving the instantaneous Earth rotation vector, it is assumed that the precession/nutation model is perfectly known [6]. The instantaneous Earth rotation vector, which is considered in the

dynamical theories and observed by the ring laser, concerns Earth rotation relative to the inertial space. In equation (7), the spatial reference is an intermediate reference system, which moves in inertial space (transition from P_{ICRF} to CIP_0, then to CIP, as shown in Figure 2) causing additional contributions to $\vec{\Omega}(t)$. A detailed discussion of this problem is given in [4].

The traditional way for obtaining the Earth rotation vector keeps only terms to first order in small quantities, i.e., polar motion $x(t)$ and $y(t)$ and departure $z(t)$ from uniform spin at the mean sidereal rotation speed Ω_0 of the Earth [4, 34,12]. Let us consider the non-linearized skew-symmetric tensor $\Omega(t)$ and give a complete geometrical interpretation of the motion of the Earth rotation vector w.r.t. the normalized vector $\vec{r}_{CIP_0}$ of the CIP_0:

$$\vec{\Omega}(t) = \begin{pmatrix} \Omega_x(t) \\ \Omega_y(t) \\ \Omega_z(t) \end{pmatrix} = \quad (7)$$
$$= \begin{pmatrix} \Omega_0 \sin[x(t)] + z(t) \cdot \Omega_0 \sin[x(t)] - \partial_t y(t) \\ -\Omega_0 \cos[x(t)] \cdot \sin[y(t)] - \Omega_0 \cdot z(t) \cdot \cos[x(t)] \cdot \sin[y(t)] - \cos[y(t)] \cdot \partial_t x(t) \\ \Omega_0 \cos[x(t)] \cdot \cos[y(t)] + \Omega_0 \cdot z(t) \cdot \cos[x(t)] \cdot \cos[y(t)] - \sin[y(t)] \cdot \partial_t x(t) \end{pmatrix}$$

which can be synthesized into:

$$\begin{pmatrix} \Omega_z(t) \\ -\Omega_y(t) \\ \Omega_x(t) \end{pmatrix} = \quad (8)$$
$$= \partial_t h(t) \cdot \vec{r}_{CIP_0} + \partial_{y(t)} \vec{r}_{CIP_0} \cdot \partial_t x(t) + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \cdot \partial_t y(t)$$

Formula (8) relates the Earth rotation vector to the initial CIP, i.e. the CIP_0 (see equation 10). Its magnitude reads:

$$\left| \vec{\Omega}(t) \right| =$$
$$= \left\{ \begin{matrix} \Omega_0^2 + \Omega_0^2 \cdot z(t)^2 + [\partial_t x(t)]^2 - 2\Omega_0 \sin[x(t)] \cdot \partial_t y(t) + \dots \\ [\partial_t y(t)]^2 + 2\Omega_0 \cdot z(t) \cdot [\Omega_0 - \sin[x(t)] \cdot \partial_t y(t)] \end{matrix} \right\}^{\frac{1}{2}}$$
$$(9)$$

To first approximation, the initial unit CIP vector, and its partial derivative w.r.t. polar motion in the y-component, follows as:

$$\vec{r}_{CIP_0} = \begin{pmatrix} \cos[x(t)] \cdot \cos[y(t)] \\ \cos[x(t)] \cdot \sin[y(t)] \\ \sin[x(t)] \end{pmatrix} \approx \begin{pmatrix} 1 \\ y(t) \\ x(t \end{pmatrix} \quad (10)$$

$$\partial_{y(t)} \vec{r}_{CIP_0} = \begin{pmatrix} -\sin[y(t)] \\ \cos[y(t)] \\ 0 \end{pmatrix} \approx \begin{pmatrix} -y(t) \\ 1 \\ 0 \end{pmatrix} \quad (11)$$

Keeping only terms to first order in small quantities reduces the Earth rotation vector and its magnitude to the familiar form [4, 12]:

$$\begin{pmatrix} \Omega_z(t) \\ -\Omega_y(t) \\ \Omega_x(t) \end{pmatrix} \approx \begin{pmatrix} \Omega_0 \cdot [1 + z(t)] \\ \Omega_0 \cdot y(t) + \partial_t x(t) \\ \Omega_o \cdot x(t) - \partial_t y(t) \end{pmatrix} \qquad (12)$$

$$\left| \vec{\Omega}(t) \right| \approx \Omega_0 \cdot (1 + z(t)) \qquad (13)$$

We would like to make clear that present polar motion and universal time estimates, in fact, do contain much more information than only the no-net-rotation (NNR) of the station coordinate corrections. The reason is that more than 70% of Earth surface is covered by the oceans where no geodetic sites are established. All signals in the oceans, which do not significantly affect station coordinates, do have their signatures in the Earth rotation parameters.
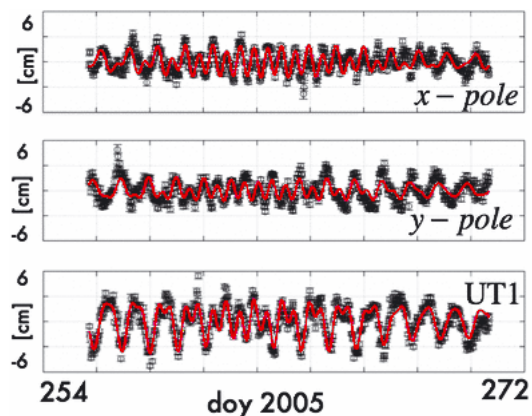
Figure 3: De-trended hourly polar motion and UT1 variations from VLBI (black circles) and results by using a source code of the ocean tidal model worked out by Ray et al. (1994), extended by Chao et al. (1996) and interpolated by Eanes for smaller tidal constituents (red line) during the VLBI CONT05 campaign. Variations are projected to the Earth surface and given in units of [cm].

Figure 3 shows de-trended hourly polar motion and universal time offsets (a smooth low-frequency function has been removed) during the VLBI CONT05 campaign. The VLBI software package OCCAM61E has been used [28]. The difference between the black circles and the red line, which is predicted from the most recent ocean tidal model adapted by Eanes [3], is in median less than 6 mm for each component. In fact, the tables provided by the IERS Conventions 2003 are based on the ocean tidal model worked out by [22] and extended by [9] by computing additionally influence on polar motion. Eanes

interpolated this model for smaller tidal constituents in a standard way, using the tidal potential and the admittance function [3].

2.2. Polar motion versus nutation

The basic and clear distinction between polar motion and precession/nutation started getting more complicated since space geodetic measurements allowed for subdaily temporal resolution of Earth rotation parameters, i.e., polar motion and universal time or length of day (LOD).

For interested readers in historical aspects, discussing the rather confusing terminology in the literature from different points of views, we refer to more extensive studies, e.g., [7, 18, 35].

Till recently, space geodetic techniques remained the only practical means of monitoring Earth rotation variations. However, space geodesy needs to attach their observations to external objects, e.g., quasars, stars, planets, satellites, and are therefore principally unable to record clean geophysical polar motion without aliasing from precession/nutation [8].

Space geodetic techniques observe only the time dependent transformation matrix $B(t)$ (including precession and nutation) between the ITRF and the ICRF. The components of the instantaneous rotation vector are not needed to define this matrix. Moreover, the dynamical theories of Earth rotation can also be expressed in terms of the components of the matrix $B(t)$. This matrix has three degrees of freedom, and hence can be described by three mutually independent parameters. For various reasons, e.g., tradition, continuity of software, etc., the scientific community decided to keep the convention of using five parameters (two for polar motion, two for precession-nutation and one for spin). Therefore it became necessary to introduce an additional constraint (see Figure 4) to remove the redundancy of the parameters. This artificial problem and its solution correspond to the title of section.

The solution of this artificial problem was conventionally adopted by the scientific community and is anchored in the International Earth rotation and Reference systems Service (IERS) Conventions 2003, which define polar motion and nutation as depending on the frequency of the motion [24]. Here, every motion in the terrestrial reference frame with a retrograde period from 16 to 48 sidereal hours is attributed to nutation; all other pro- and retrograde sidereal periods are considered to pertain to polar motion (see

Figure 4). Retrograde means the opposite direction to Earth rotation. The retrograde nearly diurnal ocean tidal terms as well as terms excited by the atmosphere are considered to be part of nutation. From the astronomical point of view, this is perfectly correct, and represents one additional part of forced nutation (not externally, but internal to the Earth system).
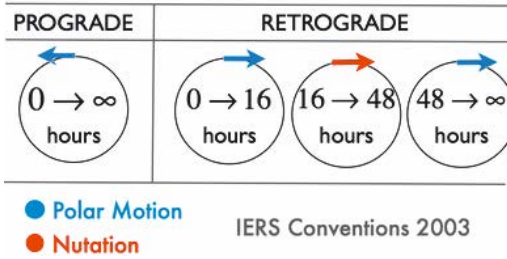
Figure 4: Definition of polar motion and nutation following the IERS Conventions 2003.

Polar motion

Phenomenon	Period [days]	Amplitude [cm]	VLBI	GNSS	Ring laser
Chandler	~432	600	✔	✔	?
Annual	~365	300	✔	✔	?
von Oppolzer	~1	60	✔*	?	✔
Ocean tides	~0.5, ~1	3	✔	✔	?

Figure 5: Sensitivity of space geodetic techniques and ring lasers to polar motion signals. Question marks denote that those techniques are not yet able to sense these signals but will probably do in future. (is included in nutation).*

Length of Day

Phenomenon	Period [days]	Amplitude [cm]	VLBI	GNSS	Ring laser
Solid Earth tides	13.7, 182, ...	50	✔	✔	(✔)
Atmosphere (winds, pressure)	10-60	50	✔	✔	(✔)
Ocean tides	~0.5, ~1	3	✔	✔	?

Figure 6: Sensitivity of space geodetic techniques and ring lasers to length of day signals. Question marks denote that those techniques are not yet able to sense these signals but will probably do in future.

The ring laser gyroscope observes a completely different quantity, i.e., the instantaneous rotation vector, which according to equation (6) is obtained from the product of the transposed matrix $B(t)$ and its time derivative. Each terrestrial motion of this vector, polar motion, is associated by the celestial motion, nutation. Another difference is that the observations of space geodesy are more or less limited in time resolution while the ring laser gyroscope can (at least potentially) perform continuous recordings.

Nevertheless, space geodetic techniques do detect many signals in polar motion and length of day down to the cm-level. The current performance of ring laser gyroscopes seems also being able to detect other effects beyond the Oppolzer terms (see section 3, Figure 5). This statement is also applicable to length of day variations (Figure 6). The potential of ring laser data remains still to be exploited. Recent data sets look very promising.

3. The ring laser gyroscope

In the last decade, the forced diurnal polar motion felt nearly into oblivion, since space geodetic techniques take this signal into account as part of precession/nutation models, e.g., IAU 2000. But recently a revival in modern Sagnac instruments, especially in ring laser gyroscopes emerged. These sophisticated instruments are sensitive to absolute rotation. Recent papers report of ring laser gyroscopes being the first tools to detect forced diurnal polar motion (arising due to lunisolar torques), first predicted by Ritter T. von Oppolzer (see Figure 7), already in the nineteenth century [30]. In reality, present single oriented circuits record the angular variation formed by the instantaneous Earth rotation vector with the area vector normal to the circuit, and are therefore sensitive to the frequency signature of the forced diurnal polar motion.

The gyrometric phase shift of a ring laser measurement is proportional to $Q(t)$:

$$Q(t) = A(t) \cdot \vec{n}_0(t) \cdot \vec{\Omega}(t) \qquad (14)$$

where $A(t)$ is the area formed by the beam circuit, $\vec{n}_0(t)$ is its unit normal vector, and $\vec{\Omega}(t)$ is the Earth rotation vector. Any variation in the angular speed or the magnitude of the Earth rotation vector translates immediately into the phase shift or $Q(t)$. The components of $\vec{n}_0(t)$ must be given in the terrestrial reference frame, however, need not being highly accurate. But, changes in $A(t)$ and $\vec{n}_0(t)$, due to pressure or

temperature changes are critical, given the high sensitivity requirements of the ring laser. For that reason the orientation of the ring laser is monitored by tiltmeters. However, the latter are not sensitive to locally induced rotational motions of the ground arising from shear stresses. Tidal deformations are such motions, but these harmonic tidal frequencies can be easily accounted for. Deformations arising from atmospheric and hydrological loading are other possible sources, however produce mainly surface normal stresses on the soil. In any case, local rotational motions, e.g., due to tremendous storms such as the European windstorm Kyrill in January 2007, should not be neglected when interpreting variations of $Q(t)$ and thus indirectly of $\vec{\Omega}(t)$.
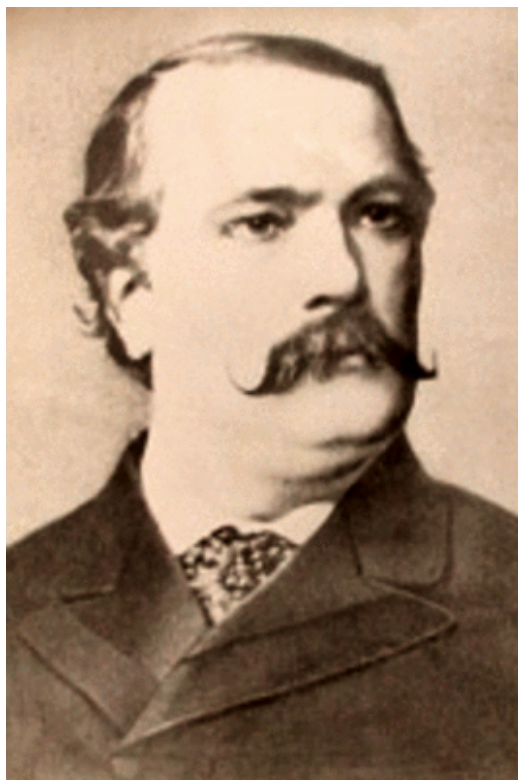

Figure 7: Prof. Theodor Ritter von Oppolzer.

In principle, three independently oriented circuits would be sufficient at one single site to determine the Earth rotation vector $\vec{\Omega}(t)$ unambiguously if the a priori knowledge of $A(t)$ and $\vec{n}_0(t)$ is sufficiently accurate. In the case that $A(t)$ is exactly known, five unknowns remain for one single site. Three sites equipped with three independently oriented circuits would be suffi-

cient to determine their respective tilts $\vec{n}_0(t)$, and the Earth rotation vector $\vec{\Omega}(t)$ unambiguously.
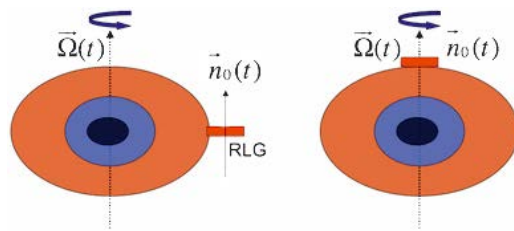

Figure 8: Equatorial (left) and polar mounting (right) of a single circuit ring laser (RLG).

At first, Figure 8 explores one further special case regarding only one single-circuit ring laser, located at the equator and having its areal unit normal vector nearly parallel to the Earth rotation vector, i.e., the angular departure of $\vec{n}_0(t)$ from $\vec{\Omega}(t)$. With this mounting, the ring laser will be mainly sensitive to the change in magnitude Ω of $\vec{\Omega}(t)$, i.e., to LOD variations. It will remain quasi unaffected, from both polar motion and local tilt. And secondly, in the same logic, a polar mounting with $\vec{n}_0(t)$ parallel to $\vec{\Omega}(t)$ senses primarily LOD variations.

4. Conclusions

Very Long Baseline Interferometry (VLBI) is a geodetic technique, which enables to estimate parameters related to the shape of the Earth and orientation in inertial space. Such parameters are the difference between universal time corrected for polar motion and Universal Coordinated Time (UT1-UTC), polar motion, and nutation/precession [16]. Besides, VLBI is the most accurate technique having access to the Celestial Reference Frame (CRF) through observations of quasars. Today, polar motion and universal time observations with a temporal resolution as short as semi-hourly are obtained from VLBI, with the potential to reveal short period and episodic events with signatures below the 100µas (7.5µas) level. The Global Navigation Satellite System (GNSS) contributes effectively to nutation rates, polar motion, and length of day parameters and Satellite Laser Ranging (SLR) is also observing polar motion, LOD and geocentre variations.

Nonetheless, the ring laser provides additional information, which is not yet, and presumably in principle never, accessible through space geodetic measurements. In the future, ring lasers could become standard high-precision tools, which could measure the finest rotational phenomena of the Earth. The operation of high-

sensitivity gyroscopes (three independently oriented circuits) at several (at least three) highly stable sites would be a real breakthrough in Earth rotation science. The combination of VLBI and ring laser data at the observation level would presumably relieve one shortcoming, i.e., provide a clean physical separation between nutation offsets and polar motion for high-frequency estimates. And finally, some systematic errors of the individual monitoring techniques could potentially be identified by comparison of both independent types of measurements.

References

[1] Altamimi, Z., P. Sillard, and C. Boucher, (2002). ITRF2000: A New Release of the International Terrestrial Reference Frame for Earth Science Applications, J. Geophys. Res., 107 (B10), 2214, doi: 10.1029/2001JB000561.

[2] Beutler, G., (2005). Methods of Celestial Mechanics, Volume I&II, in cooperation with Prof. Leos Mervart and Dr. Andreas Verdun, Astronomy and Astrophysics Library, Springer.

[3] Bizouard, C., R. Eanes, R. Ray, P. Brosche, P. Defraigne, S. Dickman, D. Gambis, and R. Gross, (2004). Tidal Variations in the Earth's Rotation, pp. 92-98, IERS Conventions 2003, eds. McCarthy, D.D., and G. Petit. http://maia.usno.navy.mil/conv2003.html

[4] Brzezinski, A., and N. Capitaine (1993). The use of precise observations of the Celestial Ephemeris Pole in the analysis of geophysical excitation of Earth rotation, J. Geophys. Res., Vol. 98 (B4), pp. 6667-6676, doi: 10.1029/92JB02874.

[5] Brzezinski, A., (2005). Review of the Chandler Wobble and its excitation, Cahiers du Centre Européen de Gédynamique et de Séismologie, Vol. 24, pp. 109-120.

[6] Capitaine, N., P.M. Mathews, P. Wallace, P. Bretagnon, R. Gross, T. Herring, G. Kaplan, D. McCarthy, B. Richter, and P. Simon, (2004). Transformation Between the Celestial and Terrestrial Systems, pp. 33-56, IERS Conventions 2003, eds. McCarthy, D.D., and G. Petit. http://maia.usno.navy.mil/conv2003.html

[7] Chao, B.F., (1985). As the World Turns, EOS, Trans. Amer. Geophys. Union, Vol. 46, pp. 766-770.

[8] Chao, B.F., (1991). As the World Turns, II, EOS, Trans. Amer. Geophys. Union, Vol. 72, pp. 550-551.

[9] Chao, B.F., R. D. Ray, J. M. Gipson, G. D. Egbert, and C. Ma, (1996). Diurnal/semidiurnal polar motion excited by oceanic tidal angular momentum, J. Geophys. Res., Vol. 101 (B9), pp. 20151-20164, doi: 10.1029/96JB01649.

[10] Darwin, G.H., (1876). On the Influence of Geological Changes on the Earth's Axis of Rotation, Proceedings of the Royal Society of London, Vol. 25, (1876-1877), pp. 328-332.

[11] Dziewonski, A.M., and D.L. Anderson (1981). Preliminary Reference Earth Model (PREM), Phys. Earth Planet. Int., Vol. 25, pp. 297-356.

[12] Gross, R.S., (2007). Earth Rotation Variations – Long Period, in "Treatise on Geophysics", Vol. 3 "Physical Geodesy", ed. T. Herring, Elsevier, in press.

[13] Jeffreys, H., (1962). The Earth, Cambridge: Cambridge University Press.

[14] Krasinsky, G.A., (2006). Numerical theory of rotation of the deformable Earth with the two-layer fluid core. Part 1: Mathematical model. Celestial Mech. Dyn. Astr., Vol. 96, pp. 169-217, doi: 10.1007/s10569-006-9038-5.

[15] Lamb, H., (1932). Hydrodynamics, 6th Ed. (November 1993), Dover Publications.

[16] Lambeck, K., (1980). The Earth Variable Rotation: Geophysical Causes and Consequences, Cambridge: Cambridge University Press.

[17] Mathews, P.M., T.A. Herring, and B.A. Buffett, (2002). Modelling of Nutation and Precession: New Nutation Series for Non-Rigid Earth and Insights into the Earth Interior, J. Geophys. Res., Vol. 107 (B4), doi: 10.1029/2001JB000390.

[18] Melchior, P., (1980). For a Clear Terminology in the Polar Motion Investigations. Nutation and the Earth's Rotation: Proceedings from IAU Symposium no. 78 held in Kiev, USSR 23-28 May, 1977. Edited by Raynor L. Duncombe. International Astronomical Union. Symposium no. 78, Dordrecht, Holland; Boston: D. Reidel Pub. Co., pp. 17-21.

[19] Meurers, B., (2002). Erdgezeiten und Eigenschwingungen, lecture notes, University of Vienna.

[20] Moritz, H., and I.I. Mueller, (1987). Earth Rotation. Theory and Observation, New York: Ungar.

[21] Pan, C., (2007). Observed multiple frequencies of the Chandler wobble, J. Geodyn., doi: 10.1016/j.jog.2006.112.004.

[22] Ray, R.D., D.J. Steinberg, B.F. Chao, and D.E. Cartwright, (1994). Diurnal and semidiurnal variations in the Earth's rotation rate induced by oceanic tides, Science, Vol. 264, pp. 830-832.

[23] Rothacher, M., G. Beutler, T. A. Herring, and R. Weber, (1999). Estimation of Nutation using the Global Positioning System, J. Geophys. Res., Vol. 104 (B3), pp. 4835-4860, doi: 10.1029/1998JB900078.

[24] Schlüter, W., E. Himwich, A. Nothnagel, N. Vandenberg, and A. Whitney, (2002). IVS and Its Important Role in the Maintenance of the Global Reference Systems, Advances in Space Research, Vol. 30, No. 2, pp. 145-150.

[25] Schreiber, K.U., A. Velikoseltsev, M. Rothacher, T. Klügel, G.E. Stedman, and D.L. Wiltshire, (2004). Direct measurement of diurnal polar motion by ring laser gyroscopes, J. Geophys. Res., Vol. 109 (B06405), doi: 10.1029/2003JB002803.

[26] Schuh, H., R. Dill, H. Greiner-Mai, H. Kutterer, J. Mueller, A. Nothnagel, B. Richter, M. Rothacher, U. Schreiber, and M. Soffel, (2003). Erdrotation und Globale Dynamische Prozesse, Mitteilungen des Bundesamtes für Kartographie und Geodäsie; 32, Verlag des Bundesamtes für Kartographie und Geodäsie.

[27] *Schweydar, W.V., (1917). Die Bewegung der Drehachse der Elastischen Erde im Erdkörper und im Raum, Astron. Nachr., Vol 203, pp. 103-114.*

[28] *Titov, O.A, V. Tesmer, and J. Boehm, (2001). Version 5.0 of the OCCAM VLBI Software. User Guide. AUSLIG Technical Reports 7, AUSLIG, Canberra.*

[29] *Vondrak, J., and C. Ron (2005). The Great Chandler Wobble Change in 1923-1940 Re-Visited, Cahiers du Centre Européen de Gédynamique et de Séismologie, Vol. 24, pp. 39-47.*

[30] *von Oppolzer, T. Ritter, (1882). Bahnbestimmung der Kometen und Planeten, Vol. 1, (2nd Ed.), Leipzig, pp. 152-155.*

[31] *Wahr, J.M., (1981). Body Tides on an Elliptical, Rotating, Elastic and Oceanless Earth, Geophys. J. R. Astr. Soc., Vol. 64, pp. 677-703.*

[32] *Weber, R., (1999). The Ability of the GPS to Monitor Earth Rotation Variation, in Acta Geodaetica et Geophysica Hungarica, Vol. 34, Number 4, pp. 457-473, Akademiai Kiado, Budapest. Res. Lett., Vol. 30(14),1742, doi: 10.1029/2003GL017546.*

[33] *Yule, G.U., (1927). Phil. Trans. Royal Soc. A, Vol. 226, pp. 267-298.*

[34] *Zharkov, V.N., S.M. Molodensky, A. Brzezinski, E. Groten, and P. Varga, (1996). The Earth and its Rotation: Low Frequency Geodynamics, Herbert Wichman Verlag.*

[35] *Zürn, W., (1997). The Nearly-Diurnal Free Wobble-Resonance, pp. 96-109, in Lectures Notes in Earth Sciences, Tidal Phenomena, eds. H. Wilhelm, W. Zürn and H.-G. Wenzel, Springer-Verlag.*

Contact

DI Dr. techn. Paulo Jorge Mendes Cerveira, Vienna University of Technology, Advanced Geodesy, Institute of Geodesy and Geophysics, Gusshaustrasse 27-29, A-1040 Vienna, email: mendes@mars.hg.tuwien.ac.at

Ao. Univ. Prof. Dr. techn. Robert Weber, Vienna University of Technology, Advanced Geodesy, Institute of Geodesy and Geophysics, Gusshaustrasse 27-29, A-1040 Vienna, email: rweber@mars.hg.tuwien.ac.at

O. Univ. Prof. Dr. Ing. Harald Schuh, Vienna University of Technology, Advanced Geodesy, Institute of Geodesy and Geophysics, Gusshaustrasse 27-29, A-1040 Vienna, email: hschuh@mars.hg.tuwien.ac.at

vgi

A New Seismic Model of the Eastern Alps and its Relevance for Geodesy and Geodynamics

*Michael Behm, Ewald Brückl, Ulrike Mitterbauer,
CELEBRATION 2000 and ALP 2002 Working Groups*

Abstract

Between 1997 and 2003, Central Europe was the target of large international seismic programs to investigate the upper lithosphere. So far, new results concerning the Eastern Alpine region include a 3D model of the P-wave velocity in the crust and a map of the Mohorovicic discontinuity (Moho) and interpretations of two representative profiles.. Taking gravity data into account, these models provide also insights into the density distribution of the lithosphere and isostatic compensation. The lateral distribution of the crustal velocities and the new Moho map enable to determine a significant fragmentation of the lithosphere. The inferred kinematics of the crustal blocks is closely related to Neogene tectonics and displacements observed by GPS.

Kurzfassung

In den Jahren 1997–2003 wurden in Zentraleuropa unter Beteiligung von 17 Nationen ausgedehnte seismische Experimente durchgeführt, welche der Untersuchung der oberen Lithosphäre dienen. Die bisherigen Ergebnisse für den Bereich der Ostalpen und umliegender Gebiete umfassen die 3-dimensionale Verteilung der P-Wellengeschwindigkeit in der Erdkruste und eine neue Tiefen- und Strukturkarte der Mohorovicic-Diskontinuität. In Verbindung mit Schweredaten ermöglicht dieses Modell Rückschlüsse auf die Dichteverteilung in der Lithosphäre und die Isostasie. Die Struktur der Mohorovicic-Diskontinuität und die laterale Verteilung der P-Wellengeschwindigkeiten innerhalb der Erdkruste ergibt eine Fragmentierung der Lithosphäre. Die daraus ableitbare Kinematik der tektonischen Blöcke steht mit neogenen geologischen Vorgängen und aktuellen, geodätisch beobachteten Verschiebungen in engem Zusammenhang.

Geology and tectonic setting of the Eastern Alps and their surrounding provinces

The major geologic units of the Eastern Alps and their surrounding tectonic provinces are shown in **Figure 1** [Schmid et al. 2004, Oberhauser 1980, Franke and Żelaźniewicz 2000]. The Bohemian massif in the north represents the European platform. To the south, European crust dips below the Molasse basin, the foreland of the Alpine orogen. The Molasse basin is overthrusted to the north by the accretionary wedge of the Eastern Alps, which comprises of the Flysch belt and the Austro-Alpine nappes. European crust has been exhumed in the Tauern Window. The Periadriatic lineament (PAL) separates the Eastern Alps from the Southern Alps. The Southern Alps share similar lithologies with the Eastern Alps, but exhibit a southward directed vergency. They are bounded to the south by the External Dinarides and the Adriatic foreland (Po plain and peninsula Istria). To the north-east, the Eastern Alps continue into the Carpathians, while the Pannonian domain, which comprises parts of the Internal Dinarides and the Tisza unit, marks the south-eastern border of the Eastern Alps. The Pannonian domain is interrupted by the Mid Hungarian Line (MHL), an important SW-NE trending fault zone.

The tectonic structure of the Alps results from a long and ongoing evolution, initiated contemporary with the opening of the Atlantic Ocean in the early Jurassic, approximately 180 Million years ago. Major geodynamic processes of the Eastern Alps include a first orogenic cycle in the Cretaceous (Eoalpine phase), resulting from the subduction of the Triassic Meliata Ocean, the subduction of the Alpine Tethys in the Tertiary, and the subsequent continent-continent collision between the European and Adriatic-Apulian plates. Crustal shortening of the Eastern Alps in north-south direction followed. The maximum extent of shortening is assumed to be 100 km, which corresponds to 50% of the original width. Since the Late Oligocene and Early Miocene, the ongoing north-south oriented compression of the Eastern Alps has been accompanied by vertical and lateral extrusion and tectonic escape of large crustal wedges to the unconstrained margin represented by the Pannonian basin in the east.

Major normal (e.g., Brenner, Tauern East) and strike-slip fault systems (Periadriatic line, Salzach-Enns-Mariazell-Puchberg line, Mur-Mürz line) were formed or reactivated by this tectonic process [Ratschbacher et al. 1991]. Within this tectonic regime, the Vienna basin and several intra-alpine basins were generated by the pull-apart mechanism.
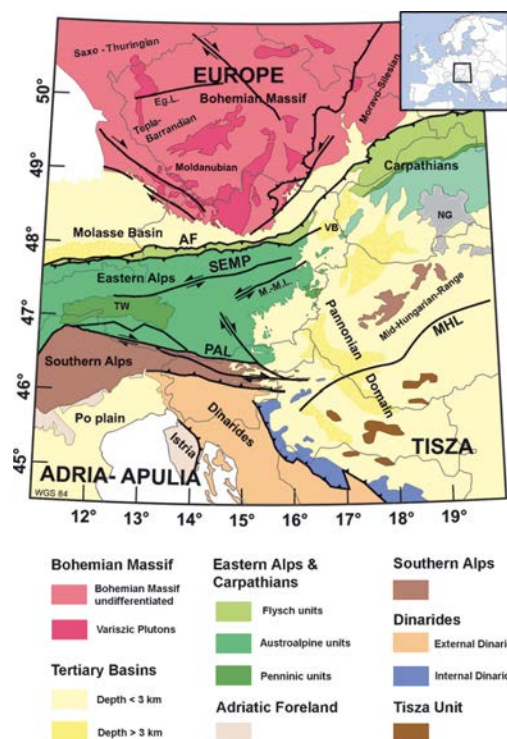

Figure 1: Tectonic setting of the investigated area [Schmid et al. 2004, Oberhauser 1980, Franke and Želaźniewicz 2000]. SEMP: Salzach-Enns-Mariazell-Puchberg line; M.-M.L: Mur-Mürz line; Eg.L: Eger line; AF: Alpine Front; PAL: Perdiadriatic lineament; MHL: Mid-Hungarian Line; TW: Tauern Window; VB: Vienna Basin; NG: Neogene Volcanics

Key questions concerning the geodynamic evolution of the Eastern Alpine lithosphere are the direction of subduction, the significance of major fault systems and their imprints in the deep crust, the continuation of large tectonic units into depth and the composition of the lower crust. Seismic models are of great importance for addressing these subjects.

Seismic experiments

Seismic investigations in the Eastern Alps started in the early 1960's with refraction lines around the quarry Eschenlohe [Giese and Prodehl 1976]. Among further important refraction lines are ALP75, ALP77, and ALP78, which cover large parts of Austria and north-eastern Italy. Several 2D interpretations derive from these experiments [Miller et al. 1977; Aric et al. 1987; Yan and Mechie 1989; Scarascia and Cassinis 1997]. TRANSALP was a large interdisciplinary project targeting the crustal structure of the Eastern Alps along transect from Munich to Venice [Lüschen et al. 2004; Bleibinhaus and Gebrande 2006]. Its core formed a 300 km long NS oriented reflection line. Three short reflections lines (altogether ~100 km length) were shot between 1992 and 2001 at the eastern edge of the Eastern Alps [Weber et al. 1996; Grassl et al. 2004], focussing on Penninic units and their relation to deeper structures.
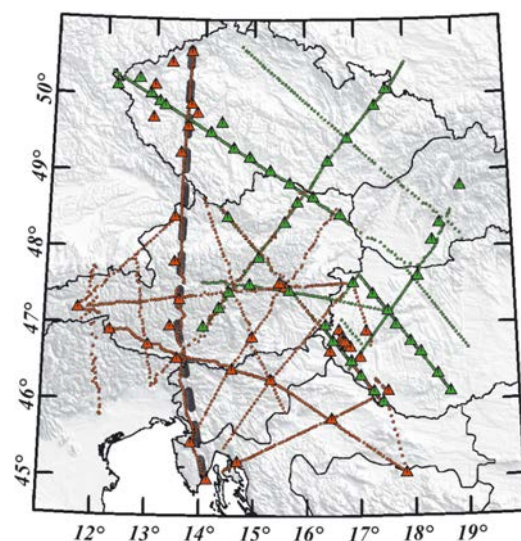
Figure 2: Layout of the experiments CELEBRATION 2000 (green; 3rd deployment) and ALP 2002 (red). Triangles indicate shot locations, and small circles represent recorders. The dotted grey line shows the orientation of the profile Alp01.

CELEBRATION 2000 and ALP 2002 were among an unprecedented series of large international 3D wide-angle refraction and reflection (WAR/R) experiments to investigate the lithosphere of Central Europe [Guterch et al. 2003]. Both projects comprised the combined efforts of 15 nations. ALP 2002 was initiated and led by the Institute of Geodesy and Geophysics at the Vienna University of Technology [Brückl et al. 2003]. The layout of the projects is shown in **Figure 2**. In this study we use data from the 3rd deployment of CELEBRATION 2000, which included 55 shots and 844 receivers deployed

along 7 profiles (total length of approximately 2800 km). Furthermore all 39 shots and 947 receivers (4300 km profile length) along 13 profiles from the ALP 2002 experiment were integrated. Shot charges were 300 kg on average and were distributed to 5 boreholes of 50 m depth. The receivers were mainly single channel recorders with pre-programmed recording time windows. The average receiver distance was 2.9 km on high density profiles and 5.8 km on low density profiles. Approximately 900 record sections and a total of about 79,000 traces were obtained.

Interpretation of seismic data

The wave field reflects the complex tectonic setting (**Figure 3**). Our interpretation aims at modelling the P-wave velocity distribution in the crust and the uppermost mantle. Thus we utilise compressional waves that dive through the crust (Pg), which are reflected from the Moho (PmP), and which travel through the uppermost mantle (Pn). The velocity distribution is described in

terms of smooth variations and/or first-order discontinuities (e.g., Moho).

The traditional interpretation method of WAR/R data is 2D interactive modelling based on ray tracing. This method provides detailed information on the velocity distribution within the crust and the uppermost mantle, in particular when intracrustal reflections (Pc) are additionally considered. However, this kind of interpretation is restricted to inline data where shots and receivers are aligned along one profile. Inline data amount to only ∼20% of the whole data set. The remaining ∼80% are crossline data, which represent recordings of shots with a lateral offset to receivers arranged along a profile.

Besides interactive modelling, other common methods are 2D and 3D travel time tomography that require high signal-to-noise (S/N) ratio of the data [e.g., Hole 1992]. In particular young orogens, such as the Eastern Alps, are characterized by complicated structures that lead to scattering of seismic energy and therefore a low
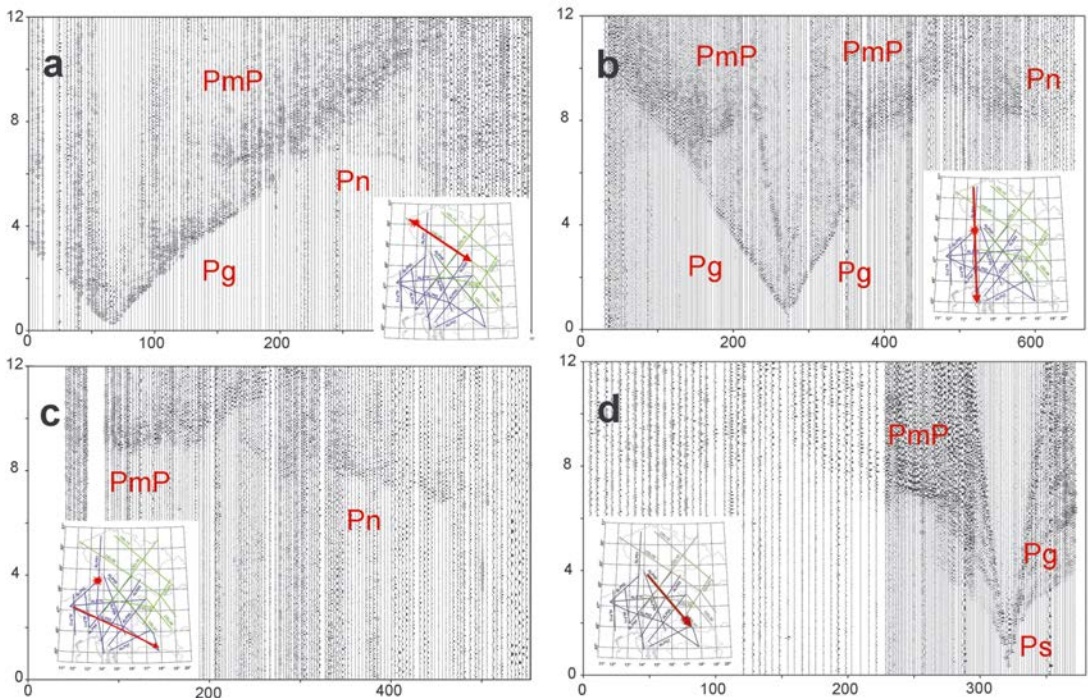
Figure 3: Examples of record sections, reduction velocity (linear move-out correction) is 8 km/s. Horizontal axis: profile coordinate [km]; Vertical axis: Reduced arrival time (**a**) *inline recording along 400 km of profile CEL09;* (**b**) *inline recording along 650 km of profile Alp01;* (**c**) *crossline recording along 550 km of profile Alp02;* (**d**) *inline recording along 370 km of profile CEL07. Note the varying quality of diving waves through sediments (Ps) and crust (Pg), of waves reflected from the Moho (PmP), and of waves travelling through the uppermost mantle (Pn).*
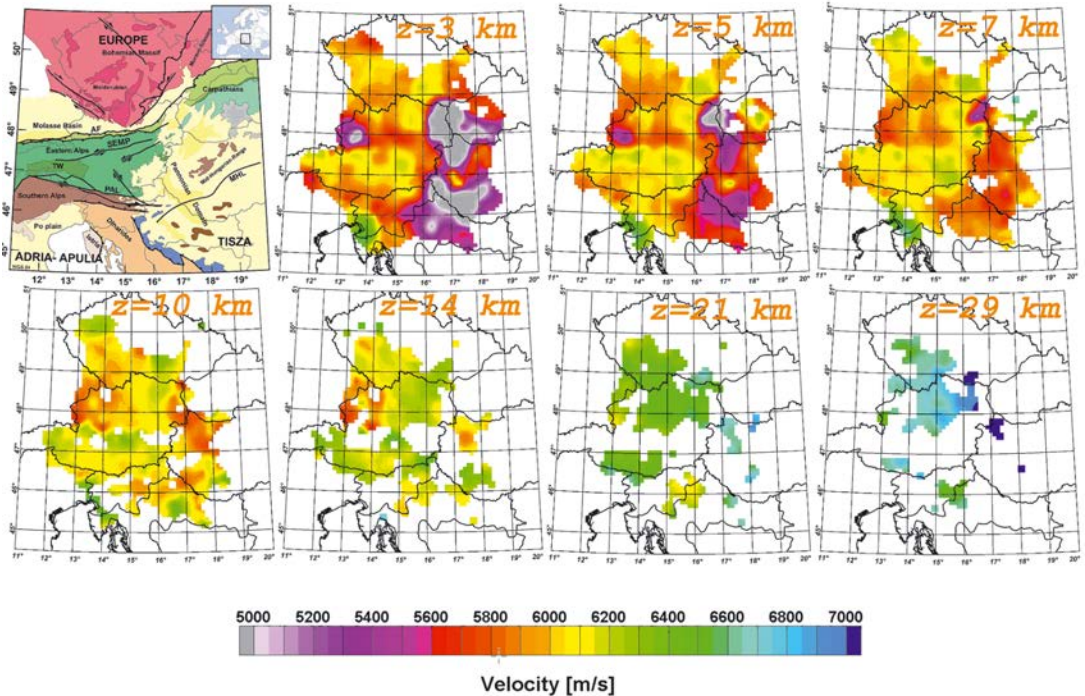
Figure 4: 7 selected horizontal slices at different depths through the 3D P-wave velocity model of the crust.

S/N ratio. In case of 3D tomography, it has to be taken into account that travel times from crossline recordings are much more difficult to interpret than from inline recordings.

For these reasons we further developed and applied 3D methods, which increase the S/N ratio and simplify the wave field. This ensures the use of the full data, regardless wether they represent crossline or inline recordings. The approach is based on a signal detection algorithm [Astiz et al. 1996] and stacking and inversion techniques especially designed for 3D WAR/R data [Behm et al. 2007]. The outcome of the methods are robust 3D models, which are not restricted to vertical sections along the individual recording profiles, but provide real 3D coverage.

In case of the crust, we stack Pg waves. Travel times are picked from stacked data and are inverted for P-wave velocities. Stacked data are supplemented by travel times from the original single-fold record sections (**Figure 3**) wherever the S/N ratio is high. Finally, a smooth 3D P-wave velocity model of the crust is obtained. Horizontal slices through this model are shown in **Figure 4**. In the uppermost crust (0–3 km), the velocities

correlate well with the tectonic setting. Low velocities represent Neogene basin fillings. Very high velocities are found in the unfolded Adriatic foreland. The most remarkable features down to depths of 10 km are pronounced velocity contrasts between segments that are related to different terrains of the Bohemian massif (Saxothuringian, Moldanubian, Moravo-Silesian) and Austro-Alpine units. These contrasts are also partially identified in the middle crust (10–20 km). Furthermore, in this depth range the Mid-Hungarian line correlates with a separation of low from high velocities. In the lower crust (> 20 km), the coverage of the model allows to interpret "normal" velocities in the Alps and the north-western part of the Bohemian Massif, compared to "anomalous" high velocities at the transition from the Bohemian Massif to the Alps and Pannonian Domain.

The Moho topography has been evaluated by a delay time approach based on stacked Pn waves. Delay times are proxy of Moho depths and are converted to depths with the velocity model of the crust. Again, the model provided by stacked data is supplemented by travel times of single fold
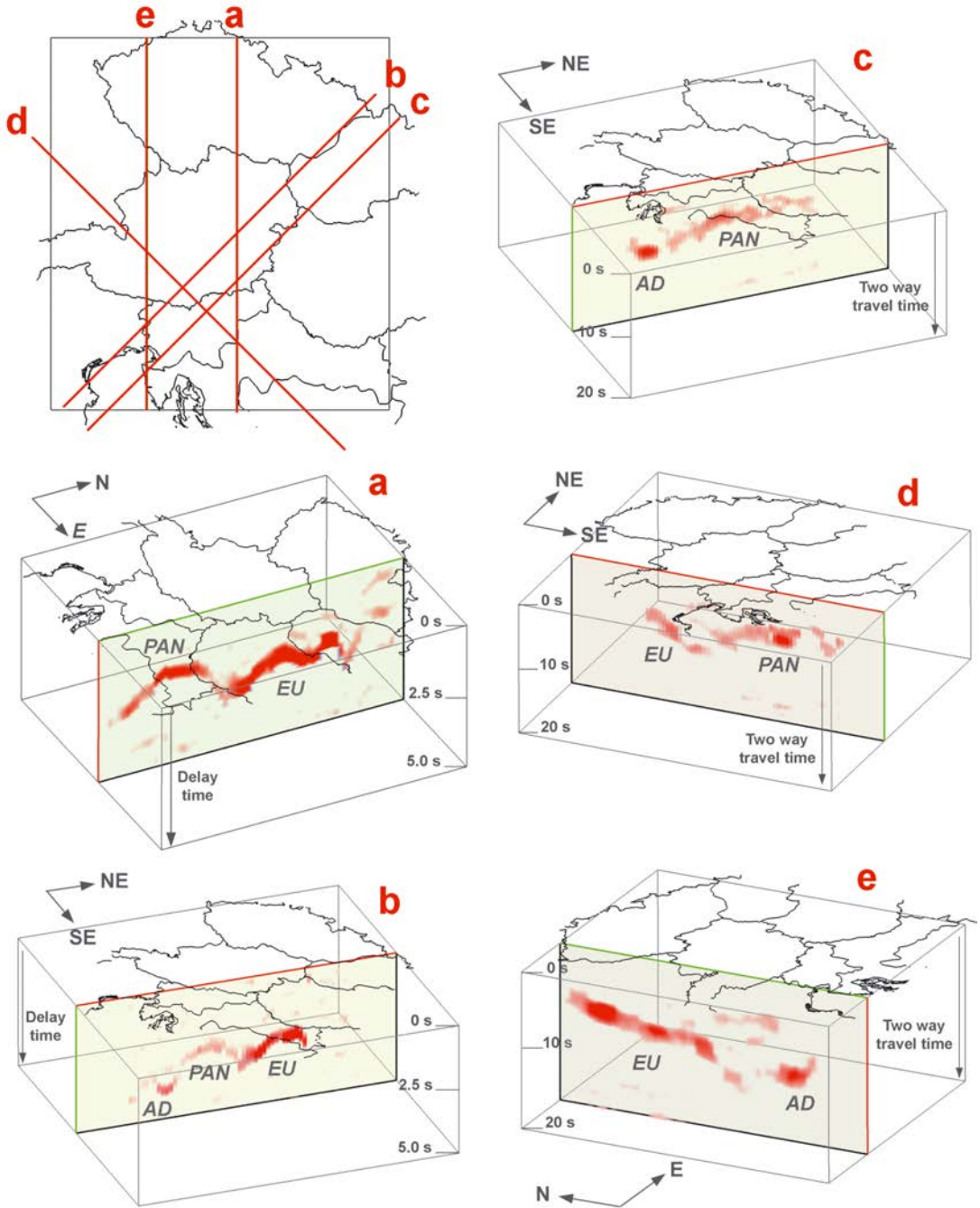
Figure 5: Slices through the 3D volume of Pn (a, b) and PmP stacks (c–e), highlighting the observed Moho fragmentation. EU: European Moho; AD: Adriatic Moho; PAN: Pannonian Moho. The "red bands" are stacked waves associated with the Moho, and the intensity of the red colour is equivalent to the amplitude of the stacked waves. The stacks are shown in the time domain and the reference level (zero time) is the depth of 10 km. For PmP stacks, the range of the time axis (two way zero offset travel time) is 20 s. For Pn stacks, the range of the time axis (delay time) is 5 s.

traces. Further, reflections from the Moho (PmP) are stacked to gain additional information on the Moho topography. The procedures yield 3D volumes of stacked Pn and PmP waves such that the coverage is not restricted to the profiles.

The stacked Pn and PmP data further enable to identify vertical offsets at the Moho. We interpret a pronounced fragmentation of the Moho (**Figures 5a–5e**) and identify the new plate fragment "Pannonia" located between the European platform in the North and the Adriatic microplate in the South-West. The Pannonian Moho is underthrusted by both the European and Adriatic Moho. The regions of underthrusting correlate well with the Alpidic and Dinaric orogens. The Moho depths vary between 24 km in the Pannonian Domain and 51 km at the central part of the Eastern Alps where the three plates collide (**Figure 6**). Previous 2D investigations [Yan and Mechie 1989, Scarascia and Cassinis 1993] are in agreement with the introduction of the Pannonian fragment.
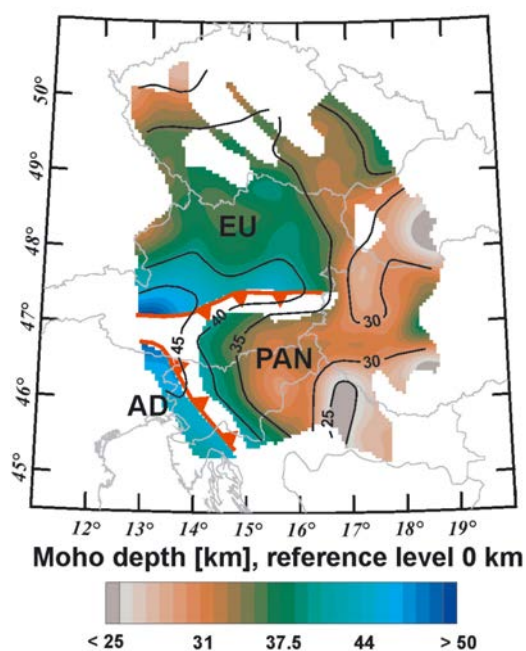
Figure 6: Moho depth map, colour coded with super-imposed continuous isolines. The fragmentation into European plate (EU), Adriatic microplate (AD) and the Pannonian fragment (PAN), as derived from the 3D volume of stacked Pn and PmP waves, is visualized by thrust symbols. Gaps result from lack of data.

The analysis of Pn waves leads also to a map of the velocity distribution in the uppermost mantle (**Figure 7**). Although the uppermost mantle velocities are not as well constrained as the crustal velocities, we can separate relatively low velocities in the Pannonian domain from higher velocities in the Adriatic part. Lower upper mantle velocities are also found at the transition from the Bohemian Massif to the Alpine area.
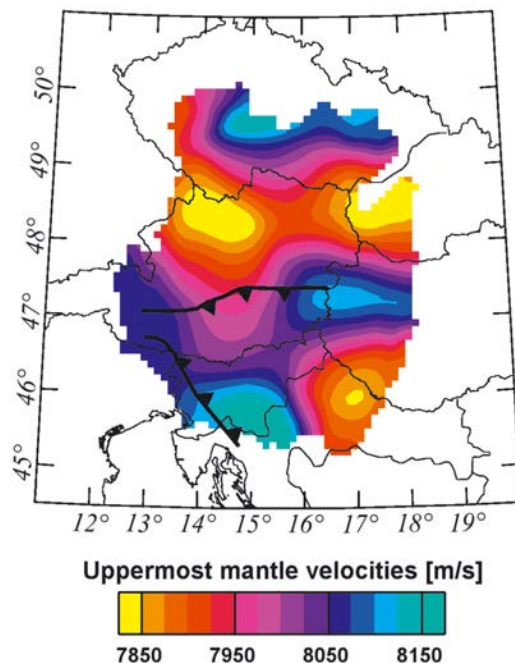
Figure 7: Map of the uppermost mantle velocity (Pn velocity). The interpreted fragmentation of the Moho is indicated by thrust symbols.

Additionally to 3D modelling, interactive 2D modelling based on ray tracing has been used to interpret profiles with high density receiver layout and a high number of shots. By interactive modelling, the velocity model along the profile is optimized by trial and error until a satisfying fit between observed and modelled seismic phases is achieved. Travel times, rays and synthetic seismograms were calculated with the ray tracing package SEIS83 [Červený and Pšenčík 1983]. The initial model is generated by tomographic inversions of first arrival travel time data [Hole 1992] and a priori information on near surface structures like Neogene basins [Brix and Schultz 1993; Saftić et al. 2003].

The main profiles of the ALP 2002 experiment are Alp01 (650 km, N-S, 212 receivers, 13 shots) and Alp02 (550 km, WNW-ESE, 151 receivers, 8 shots). In **Figure 8** we show a preliminary interpretation of Alp01 [Behm et al. 2006], the final interpretation will be presented by Brückl et

al. [in print]. The most important findings from these profiles are a down-dipping European Moho below the Adriatic mantle (Alp01) and a sudden Moho uprise at the transition from the Alpine area to Pannonian domain (Alp02). In general, the results confirm the main findings of the 3D model as the fragmentation into three plate fragments and certain high and low velocity regions in the crust. By careful evaluation of crustal and Moho reflections (Pc and PmP phases) the 2D modelling based on ray tracing yields also velocity information in parts of the lower crust where the 3D-velocity model lacks coverage.
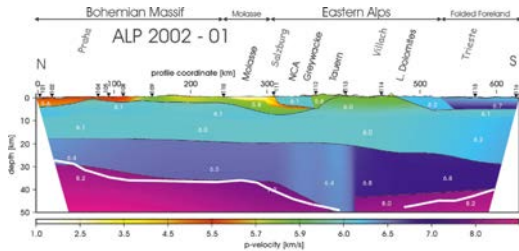
Figure 8: 2D model along the profile Alp01 derived by interactive ray tracing. Numbers show the velocity in km/ s. The Moho is located at the transition from the blue to the violet coloured region. The white line shows the Moho depth obtained from 3D modelling. Inverted triangles and numbers at the surface indicate shot points. Vertical exaggeration is 5:1.

Analysis of sensitivity and accuracy [Behm 2006] provide estimations of errors of the velocity models and the Moho depth map. On average, the accuracy of velocities ranges from ±60 m/s in the upper crust to ±100 m/s in the middle and lower crust. Depending on the coverage, the accuracy of Moho depths varies between ±1 km and ±3 km.

Implementation of gravity data

Senftl (1965) derived a Bouguer gravity map of Austria from data acquired for hydrocarbon exploration and geodetic purpose along the benchmarks of levelling lines based on approxi-mately 2000 stations. The gravimetic net has been densified since the Gravimetric Alpine Traverse [Meurers et al. 1987, Meurers 1993]. Use of this data was made in the eastern part of our investigation area by Lillie et al. [1994], who focused on regional studies extending from the Eastern Alps to the Carpahtians and Pannonian Basin. In the western part of our investigation area the Eschenlohe seismic profiles [Braitenberg et al. 1997, Dal Moro et al. 1998] and the TRANSALP transect served as constraints for gravimetric inversions and investigations of isostasy [Ebbing

et al. 2001, Ebbing 2002, Braitenberg et al. 2002, Ebbing 2004]. A gravimetric study, using the seismic 3D-model of the Eastern Alps and their surroundings derived from CELELEBRATION 2000 and ALP 2002 data, as described above, was made by Brückl et al. [2006a]. The main purpose of this study was to supply additional constraints for P-wave velocities in the lower crust derived from gravity, since the seismic 3D model does not cover these depths sufficiently. Methods and results of this study will be described briefly in the following paragraphs.

The used gravity data (**Figure 9**) are compiled from "New Austrian Bouguer Map" [Kraiger and Kühtreiber 1992] and the Bouguer gravity map of the West East European Gravity Project (http:// www.getech.com). The density for reduction of masses is d = 2670 kg/m^3. The authors use the orthometric height system in such a way that rather masses above the geoid are reduced than those above the ellipsoid. This leads to the geophysical indirect effect that in Austria is widely negligible at the accuracy level of the Bouguer gravity.
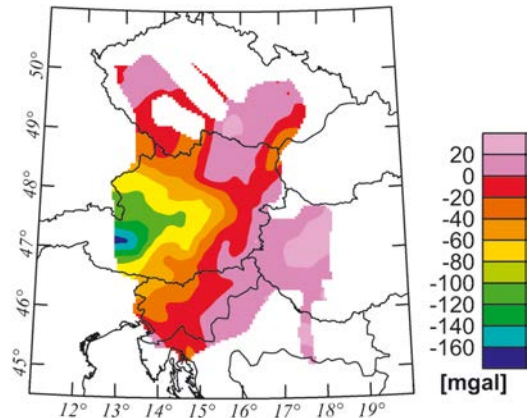
Figure 9: Bouguer gravity map of the investigation area compiled from data of the West East European Gravity Project and the Ldquo;New Austrian Bouguer Map" (for references see text).

Densities can be estimated from correlations with seismic P-wave velocities. Sobolev and Babeyko [1994] derived a relation between P-wave velocity and density that considers the thermal gradient and the mineral transformations within the crust with changing PT conditions. As only few observations on the geothermal gradient are available within the Eastern Alps [Sachsen-hofer 2001], we chose the Christensen and Mooney [1995] velocity-density relationship. This relationship requires only P-wave velocity data to

estimate crustal densities. The Moho depth map (Figure 6) is a further constraint to the gravimetric modelling. However, as pointed out before, the seismic velocities of the lower crust are the most uncertain parameters, and these errors are also propagated to the Moho depths. For this reason we convert the Moho depth map to the time domain. The Moho depth in the time domain (Tm) is the travel time of a P-wave from surface to Moho along a vertical straight ray. This quantity is much less sensitive to errors of crustal velocities than the Moho depth itself. Therefore, Tm has been used as constraint for our gravimetric modelling.

The 3D seismic model covers the upper crust down to 10 km depth almost entirely. Thus we decide to generate a density model from this data by the Christensen and Mooney velocity-density relationship and subtract its gravity effect from the Bouguer anomaly. The next step is to calculate the gravity effect of the middle and lower crust between $Z = -10$ km and the Moho. For this purpose we select a series of reasonable linear velocity-depth functions, convert them to density-depth functions by the Christensen-Mooney relationship, and calculate Moho depths from Tm for each velocity-depth function. We assume that an optimum velocity-depth function and the corresponding density-depth function removes the correlation between Tm and the residual of the Bouguer gravity after subtraction of the effect of the whole crust. The velocity function with 6102 m/s at 10 km depth and 7174 m/s at 50 km depth fulfills this condition. Corresponding densities are 2758 kg/m^3 (10 km) and 3210 kg/m^3 (50 km). The mantle density is assumed to be 3270 kg/m^3. Our approach to determine an optimum average density-depth function for the middle and lower crust is similar to Nettleton's method to derive the density of topographic masses. In Nettleton's method the optimum topographic density corresponds to a zero correlation between Bouguer anomaly and terrain elevation.

The residual Bouguer gravity after subtraction of the gravity effect of the whole crust is shown in Figure 10. The reference density model of the upper lithosphere is 2670 kg/m^3 for the uppermost 10km, 2900 kg/m^3 from 10 km to Moho and 3270 kg/m^3 for the upper mantle. The residual Bouguer gravity shows significant correlations with tectonic structures (compare Figures 10 and 1). Gravity lows correlate with the South Bohemian Pluton, which has significantly lower densities than the metamorphic part of the Bohemian Massif [Meurers 1993], and with the Tauern Window [see also Ebbing 2002]. Evidence for an Adriatic-

Apulian plate dipping below the assumed tectonic block "Pannonia" and the Dinarides is indicated by positive anomalies. Strong positive anomalies are also found in the transition from the Northern Pannonian domain (Vienna Basin) to the European platform, where the seismic models show high velocities in the lower crust.
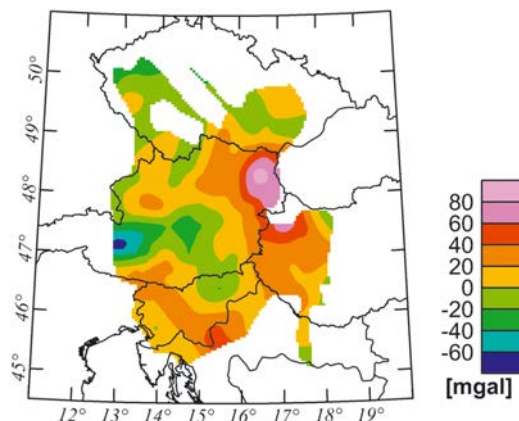

Figure 10: Residual Bouguer gravity, the effect of the whole crust subtracted.

The sources of the residual Bouguer gravity are most probably located within the crust, as could be demonstrated by tests with sources at different depth levels. Finally, we model the residual Bouguer gravity by superimposing density variations on the average density below 10 km depth. Back transformation of the densities to velocities yields average P-wave velocities for the crust below 10 km depth, which are mainly gravimetrically derived (Figure 11a). For comparison the average P-wave velocities derived only from seismic data (3D and 2D models) are shown in Figure 11b. The differences between the seismically and gravimetrically determined velocities vary between −381 m/s and 362 m/s with a mean and standard deviation of −7±147 m/s. Corresponding values for the Moho depths are −3.8 km, +2.7 km, and +0.1±1.2 km. These figures are in accordance with the Moho depth errors we estimated for the seismic model. However, only few structural features in the two velocity maps (Figures 11a, b) show similarity, which should be the topic of further investigations

Finally, density models from seismic and gravimetric modelling are combined. The vertical loading stress at $Z = 50$ km (a level below the orogen root), derived from the loads of this model is shown in Figure 12. The base area of the vertical colums is 20 km times 20 km and only local compensation is considered. The standard
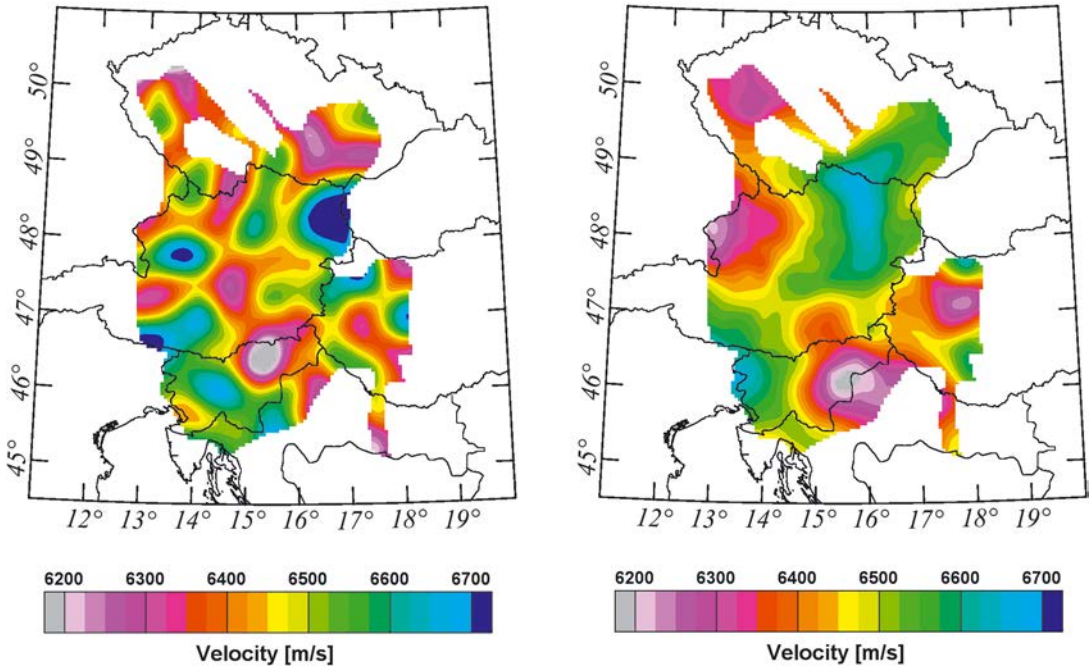
Figure 11: Average P-wave velocity of the middle and lower crust from 10 km to Moho depth; (a) *from gravimetric data;* (b) *from seismic data.*

deviation from Airy isostatic equilibrium at $Z = 50$ km is ± 6 MPa. The corresponding value at $Z = 0$ km is ± 11 MPa. These figures and the corresponding variances demonstrate that the major part of the surface loads is compensated by Airy isostasy. Ebbing et al. [2006] derived deviations from isostatic equilibrium in the same order of magnitude in the central part of the Eastern Alps (TRANSALP) at depth levels between 60 and 300 km.

Figure 12: Vertical loading stress at $Z = 50$ km, derived from the loads given by our density model. The values are relative to the reference lithosphere.

Geodynamic interpretation

The outcome of the CELEBRATION 2000 and ALP 2002 projects concerning the Eastern Alpine region can be related to geodynamic processes. Some structures in the geophysical models, like the seismic velocity distribution in the Bohemian massif, or the low velocity European basement below the Molasse basin, find their explanation in processes that took place during pre-Alpidic orogenetic cycles, or the development of the Penninic Ocean. The high upper crust velocities at the peninsula Istria suit the concept of an Adriatic indenter. The high lower crust densities below the Vienna basin and the north-western Pannonian basin are relevant for isostatic compensation and may be explained by magmatic underplating during the generation of these basins by the pull-apart mechanism.

The most prominent tectonic structure revealed by the CELEBRATION 2000 and ALP 2002 projects is the Pannonian fragment. According to our interpretation [Brückl et al. in print] the Adriatic microplate and the Pannonian fragment formed one single unit during the collision process (Late Cretaceous until Early Oligocene). Since the onset of extrusion and tectonic escape in Late Oligocene / Early Miocene, significant upper

crustal tension took place. The corresponding flattening of surface topography in the eastern part of the Eastern and Southern Alps has been compensated by a pronounced thinning of the whole crust and an upward jump of the Moho discontinuity of ~10 km. This interpretation is further supported by low seismic velocities in the upper and middle crust in this area. Furthermore the Moho jump could also facilitate continuing convergence between Europe and the Adriatic microplate by underthrusting of Adriatic mantle below Pannonian mantle. The uppermost mantle velocities (**Figure 7**) may support this interpretation, if we assume a rigid Adriatic indenter with high velocities. The lower velocities in the eastern part of the Pannonian fragment (former Adriatic plate) are possibly related to the increased heat flow below the Pannonian basin.

A sketch of the geodynamic situation in key area of tectonic processes shows the main Neogene tectonic movements, which are related to the geophysical structures of our models (**Figure 13**). This scheme of approximately N-S directed convergence between European plate and Adriatic microplate, uplift of the Tauern Window due to vertical extrusion, and lateral escape of the Pannonian fragment approximately to the SSE corresponds well with geologic estimates of continuing convergence and tectonic escape [Linzer et al. 2002], actual seismicity [Reinecker and Lenhardt 1999] and crustal

movements in lateral [Grenerczy and Kenyeres 2006] and vertical [Höggerl 1989] directions.

Conclusion

The application of stacking, travel time tomography, and delay time methods to 3D seismic data resulted in a detailed image of the P-wave velocity distribution in the crust and a new map of the Moho discontinuity. Complementary interpretations of selected profiles with dense receiver spacing and sufficient number of inline shots revealed crustal reflectors, supplied additional information on the velocity of the middle and lower crust, and refined the image of the Moho discontinuity. Furthermore, the seismic model of the upper crust (down to 10 km depth) and the Moho depth information in time domain (travel time of the normal incidence PmP reflection) were used to constrain gravimetric modelling of Bouguer gravity. P-wave velocities were converted to densities by the use of a well-established relation [Christensen and Mooney 1995] and the gravimetrically determined densities of the middle and lower crust were back transformed to P-wave velocities using the inverse relation. On average these velocities agree well with the seismic model. However, only few features on the maps of both velocities (**Figure 11 a, b**) correlate well. Therefore, the density model should be refined by using the latest Bouguer anomaly map [Bielik et al. 2006] and implementing detailed information on upper

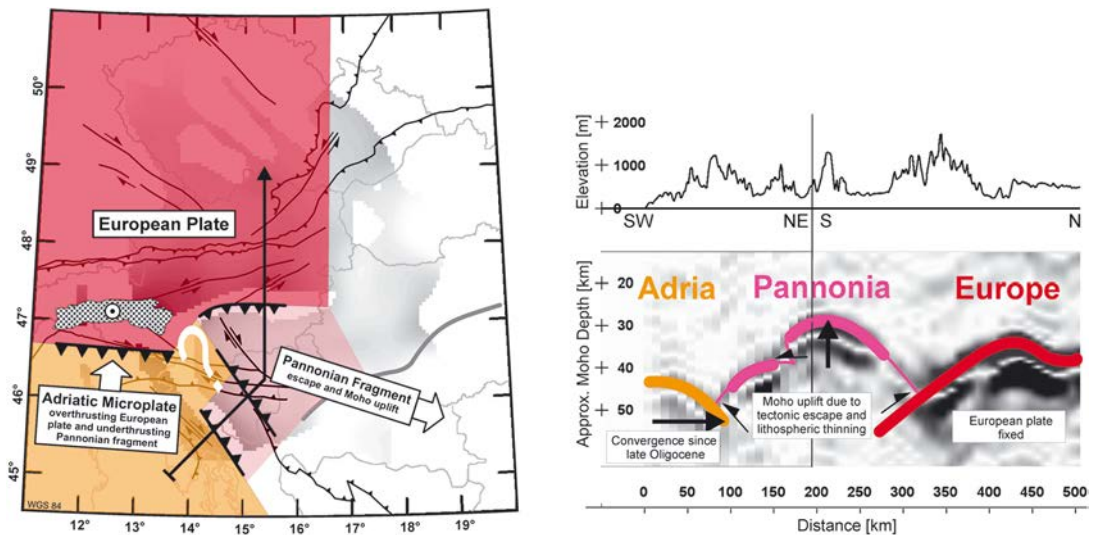

Figure 13: Geodynamic Model: (a) kinematics of European plate, Adriatic microplate and Pannonian fragment; arrows and other symbols indicate direction of underthrusting of upper mantle, vertical extrusion of Tauern Window, and relative horizontal displacement due to collision and tectonic escape; (b) cross section through PmP stack and its tectonic interpretation along profile shown in (a).

crustal densities [e.g., Steinhauser et al. 1984]. The seismic data contain also S-wave phases (Sg). This additional information has currently been exploited, and may help to refine the relations between density and seismic velocities. Finally, the density model should reach an accuracy that makes it useful for geoid determinations.

The seismic model, in particular the topography of the Moho discontinuity can be closely related to inter-plate collision and to tectonic extrusion and escape processes since the Tertiary. Seismicity and geodetically observed deformations indicate that these processes are still active. The development of a geodynamic model combining geological data, geophysical determined structures, and observed actual processes is a challenge for future work. So far, our geophysical investigations comprise the crust and uppermost mantle. The ongoing ALPASS project [Brückl et al. 2006b] is a passive seismic monitoring program based on a temporary network. It targets the upper mantle including the asthenophere by teleseismic tomography and the activity of large scale fault systems (e.g., Mur-Mürztal). The additional information on deep structures will increase our knowledge of the large scale plate tectonic processes, and will also improve gravimetric modelling (influence of asthenosphere topography on Bouguer anomaly). The high resolution image and quantification (magnitude, moment tensor, slip vector) of seismic activity at fault systems together with long term GPS observations will provide a sound basis for an improved understanding of present tectonic processes.

References

[1] *Aric, K., and. Gutdeutsch, R., 1987*: Geophysical aspects of the crustal structures of the Eastern Alps, in Geodynamics of the Eastern Alps, edited by H. W. Fluegel and P. Faupl, Franz Deuticke Verlag, 309-360.

[2] *Astiz, L., Earle, P., and Shearer, P., 1996*: Global stacking of broadband seismograms. Seismological Research Letters, 67, 4, 8-18.

[3] *Behm, M. 2006*: Accuracy and resolution of a 3D seismic model of the Eastern Alps. Ph.D. thesis, Vienna University of Technology.

[4] *Behm, M., Brückl, E., Bleibinhaus, F., Grad, M., Mertl, S., and CELEBRATION 2000 & ALP2002 Working Groups, 2006*: A new seismic model of the Eastern Alpine crust. Proceedings at First European Conference on Earthquake Engineering and Seismology (a joint event of the 13th ECEE & 30th General Assembly of the ESC) Geneva, Switzerland, 3-8 September 2006.

[5] *Behm, M., Brückl, E., Chwatal, W., Thybo, H., 2007*: Application of stacking techniques to 3D wide-angle reflection and refraction data from the Eastern Alps. Geophys. J. Int., in print, doi: 10.1111/j.1365-246X.2007.03393.x.

[6] *Bielek, M., and the CELEBRATION 2000 Potential Field Team, 2006*: Gravity Map of Central Europe (CELEBRATION 2000area). A040; EGU-A-01556.

[7] *Bleibinhaus, F., Brückl, E., and ALP 2002 Working Group, 2006*: Wide-angle observations of ALP 2002 shots on the TRANSALP profile: Linking the two DSS projects, Tectonophysics, 414 (1-4), 71-78, doi:10.1016/j.tecto.2005.10.027

[8] *Bleibinhaus, F., and Gebrande, H., 2006*: Crustal structure of the Eastern Alps along the TRANSALP profile from wide-angle seismic tomography, Tectonophysics, 414 (1-4), 51-69, doi:10.1016/j.tecto.2005.10.028.

[9] *Braitenberg, C., Ebbing, J., Götze, H.-J., 2002*: Inverse modelling of elastic thickness by convolution method - the Eastern Alps as a case example. Earth and Planetary Science Letters, 2002, 387-404.

[10] *Braitenberg, C., Pettenati, F., Zadro, M., 1997*: Spectral and classical methods in the evaluation of Moho undulations from gravity data: the NE-Italian Alps and isostasy. Journal of Geodynamics, 23, 5-22.

[11] *Brix, F., and Schultz, O., 1993*: Erdöl und Erdgas in Österreich, Verlag: Naturhistorisches Museum Wien und F. Berger, Horn, Wien, 688.

[12] *Brückl, E. Bleibinhaus, F., Gosar, A, Grad, M., Guterch, A., Hrubcová, P., Keller, G.R, Majdański, M., Šumanovac, F., Tiira, T., Yliniemi, J., Hegedüs, E., Thybo, H., 200X*: Crustal Structure Due to Collisional and Escape Tectonics in the Eastern Alps Region Based on Profiles Alp01 and Alp02 from the ALP 2002 Seismic Experiment. J. Geophys. Res., in print, doi:10.1029/2006JB004687.

[13] *Brückl, E., Mitterbauer, U., and Behm, M. 2006a*: Studies on crustal structure and gravity in the Eastern Alps, in Geodetic Deformation Monitoring: From Geophysical to Engineering Roles. IAG Symposium 131, editors F. Sanso and A. J. Gil, Springer, 181-192.

[14] *Brückl, E. & THE ALPASS TEAM, 2006b*: ALPASS – Passive seismic monitoring in the Eastern Alps. Poster Presentation at EGU, General Assembly, Vienna, Austria, 02-07 April 2006. Geophysical Research Abstracts, Vol. 8, 06385, 2006. SRef-ID: 1607-7962/gra/EGU06-A-06385.

[15] *Brückl, E., Bodoky, T., Gosar, A., Grad, M., Guterch, A., Hajnal, Z., Hegedüs, E., Hrubcová, P., Keller, G.R., Špičák, A., Šumanovac, F., Thybo, H., Weber, F., and ALP 2002 Working Group, 2003a*: ALP 2002 seismic experiments. Stud. Geoph. Geod., 47, 671-679.

[16] *Červeny, V., and Pšenčik, I., 1983*: SEIS83 – Numerical modelling of seismic wave fields in 2-D laterally varying layered structures by the ray method, in Documentation of Earthquake Algorithms, edited by E. R. Engdal, Rep. SE-35, World Data Center for Solid Earth Geophysics, Boulder, Colo., 36-40.

[17] *Christensen, N.I. and Mooney, W.D., 1995*: Seismic velocity structure and composition of the continental crust: A global view, J. Geophys. Res., 100, 97619788.

[18] *Dal Moro G., Braitenberg, C., and Zadro, M., 1998*: Geometry and mechanical and crustal properties in NE Italy based on seismic and gravity data. B olletino di Geofisica Teorica ed Applicata,Vol. 39, N.1, 37-46.

[19] *Ebbing J., Braitenberg, C., Götze, H.-J., 2001*: Forward and inverse modelling of gravity revealing insight into crustal structures of the Eastern Alps. Tectonophysics, 337/3-4, 191-208.

[20] *Ebbing, J., 2002*: 3-D Dichteverteilung und isostatisches Verhalten der Lithosphäre in den Ostalpen. Thesis, Freie Universität Berlin, 143.

[21] *Ebbing, J., 2004*: The crustal structure of the Eastern Alps from a combination of 3D gravity modelling and isostatic investigations. Tectonophysics, 380/1-2, 80-104.

[22] *Ebbing J., Braitenberg, C., Götze, H.-J., 2006*: The lithospheric density structure of the Eastern Alps, Tectonophysics, 414, 145-155.

[23] *Franke, W., and Żelaźniewicz, A., 2000*: The eastern termination of the Variscides: terrane correlation and kinematic evolution, in: Orogenic processes: quantification and modelling in the Variscan Belt, edited by W. Franke, W. Haak, O. Oncken, and D. Tanner, Geological Society, London, Special Publications, 179, 63-85.

[24] *Giese, P., and Prodehl, C., 1976*: Main features of crustal structure in the Alps, in Explosion seismology in Central Europe, edited by P. Giese and C. Prodehl, Springer, 347-375.

[25] *Grassl, H., Neubauer, F., Millahn, K., Weber, F., 2004*: Seismic image of the deep crust at the eastern margin of the Alps (Austria): indications for crustal extension in a convergent origin. Tectonophysics 380, 105-122.

[26] *Grenerczy, G., and Kenyeres, A., 2006*: Crustal deformation between Adria and the European platform from space geodesy, in: The Adria microplate: GPS geodesy, tectonics and hazards, edited by N. Pinter, G. Grenerczy, J. Weber, S. Stein, and D. Medak, NATO Science Series IV: Earth and Environmental Sciences, 61, 321-334.

[27] *Guterch, A., Grad, M., Keller, G.R., Posgay, K., Vozár, J., Špičák, A., Brückl, E., Hajnal, Z., Thybo, H., Selvi, O., and CELEBRATION 2000 Experiment Team 2003*: CELEBRATION 2000 Seismic Experiment, Stud. Geoph. Geod., 47, 659-669.

[28] *Höggerl, N., 1989*: Rezente Höhenänderungen in Österreich abgeleitet aus Präzisionsnivellement-Messungen, Österreichische Beiträge zu Meteorologie und Geophysik, 2, 161-173.

[29] *Hole, J. A., 1992*: Nonlinear high-resolution three-dimensional seismic travel time tomography, J. Geophys. Res., 97, 6553-6562.

[30] *Hrubcová, P., Środa, P., Špičák, A., Guterch, A., Grad, M., Keller, G. R., Brückl, E., Thybo, H.,2005*: Crustal and uppermost mantle structure of the Bohemian massif based on CELEBRATION 2000 data, J. Geophys. Res., 110, B11305, doi:10.1029/2004JB003080.

[31] *Kraiger, G., and Kühtreiber, N., 1992*: Preliminary results of a new Bouguer Map of Austria. Geodesy and Physics of the Earth: Geodetic Contributions to Geodynamics, 7 th Symposium Nr. 112, 5-10 October, Potsdam. Eds. H. Montag and C. Reigber: Springer-Verlag, p.133.

[32] *Lillie, R.J., Bielik, M., Babuska, V., Plomerova, J., 1994*: Gravity modelling of the lithosphere in the Eastern Alpine - Western Carpathian - Pannonian Basin region. Tectonophysics, 231, 215-235.

[33] *Linzer, H.-G., Decker, K., Pereson, H., Dell'Mour, R,. Frisch, W., 2002*: Balancing lateral orogenic float of the Eastern Alps, Tectonophysics, 354, 211-237.

[34] *Lüschen, E., Lammer, B., Gebrande, H., Millahn, K., Nicolich, R., and TRANSALP Working Group 2004*: Orogenic structure of the Eastern Alps, Europe, from TRANSALP deep seismic reflection profiling, Tectonophysics, 388, 85-102.

[35] *Meurers, B.,1993*: Die Böhmische Masse Österreichs im Schwerefeld. 6. Int. Alpengrav. Koll., Leoben 1993, Österr. Beitr. Met. Geoph., 8, 69-81.

[36] *Meurers, B., Ruess, D., Steinhauser, P., 1987*: The Gravimetric Alpine Traverse. In: Flügel, H., and Faupl, P. (Editors): Geodynamics of the Eastern Alps. Deuticke, Vienna, 334-344.

[37] *Miller, H., Gebrande, H., and Schmedes, E., 1977*: Ein verbessertes Strukturmodell für die Ostalpen, abgeleitet aus refraktionsseismischen Daten unter Berücksichtigung des Alpenlängsprofils, Geol. Rundschau 66, 289-308.

[38] *Oberhauser, R.,1980*: Der Geologische Aufbau Österreichs, Springer, Wien, New York, 699.

[39] *Reinecker, J., and Lenhardt, W. AQ.,1999*: Present-day stress field and deformation in eastern Austria, Int. Journ. Earth Sciences, 88, 532-550.

[40] *Ratschbacher, L., Frisch, W., Linzer, H. G., and Merle, O., 1991*: Lateral extrusion in the Eastern Alps, Part 2.: Structural analysis, Tectonics, 10, 2, 257-271.

[41] *Sachsenhofer, R.,2001*: Syn- and post-collisional heat flow in the Cenozoic Eastern Alps. Int. J. Earth Sciences (Geol. Rundsch.), 90, 579-592.

[42] *Saftić, B., Velić, J., Sztano, O., Juhasz, G., Ivković, Ž., 2003*: Tertiary subsurface facies, source rocks and hydrocarbon reservoirs in the SW part of the Pannonian basin (Northern Croatia and Southwestern Hungary), Geologia Croatica, 56, 101-122.

[43] *Scarascia, S., and Cassinis, R., 1997*: Crustal structures in the central-eastern Alpine sector: a revision of available DSS data, Tectonophysics, 271, 157-188.

[44] *Schmid, S., Fügenschuh, B., Kissling, E., and Schuster R., 2004*: Tectonic map and overall architecture of the Alpine orogen, Swiss Journal of Geosciences, 07,1, 93-117.

[45] *Sobolev, S.V and Babeyko, A.Y., 1994*: Modelling of mineralogical composition, density and elastic wave velocities in anhydrous magamatic rocks. Surveys in Geophysics 15, 515-544.

[46] *Środa, P., Czuba, W., Guterch, A., Tokarski, A.K., Janik, T., Rauch, M., Keller, G. R., Hegedues, E., Vozar J., and CELEBRATION 2000 Working Group, 2006*: Crustal and upper mantle structure of the Western Carpathians from CELEBRATION 2000 profiles CEL01 and CEL04: seismisc models and geological implications. Geophys. J. Int., 167, 737-760.

[47] *Steinhauser, P., Ruess, D., Zych, D., Haitzmann, H., Walach, G., 1984*: The geoid in Austria: Digital models of mean topographic heights and rock densities. Proc. 18th Gen. Ass. IUGG, IAG, Vol. 1, 322-338.

[48] *Weber. F., Schmöller, R., and Fruhwirth, R.K., 1996*: Results of a deep reflection seismic measurement south of Rechnitz/Burgenland/Austria. Geophys. Trans., 40, 79-93.

[49] *Yan, Q.Z., and J. Mechie, 1989*: A fine section through the crust and lower lithosphere along the axial region of the Alps. Geophysical Journal 98, 465-488.

Contact

Univ.Ass.Dr. Dipl.Ing. Michael Behm, Institut für Geodäsie und GeophysikTechnische Universität Wien, Gusshausstrasse 27-29/1282, A-1040 Wien, Austria.
E-mail: mbehm@mail.tuwien.ac.at

O.Univ.Prof. Dr. Dipl.Ing. Ewald Brückl, Institut für Geodäsie und GeophysikTechnische Universität Wien, Gusshausstrasse 27-29/1282, A-1040 Wien, Austria.
E-mail: ebrueckl@mail.tuwien.ac.at

Dipl.Ing. Ulrike Mitterbauer, Institut für Geodäsie und GeophysikTechnische Universität Wien, Gusshausstrasse 27-29/1282, A-1040 Wien, Austria.
E-mail: umitt@mail.tuwien.ac.at

vgi

Static and kinematic testing of tiltmeters: facilities and results

Helmut Woschitz and Klaus Macheiner

Abstract

Today tiltmeters are widely used, often as part of measurement systems. In civil engineering applications, an accuracy of about 0.01° for the inclination is sufficient for many purposes. However, before using a specific type of sensor, it is most important to know about its performance, in static as well as in kinematic situations. But often, the required information is not provided by the manufacturer, and thus tests by the user are essential. We have developed testing facilities and a simple testing sequence for the determination of basic static and dynamic parameters of tiltmeters. The capability of the facilities is described in this article, and the results of the testing sequence are shown for one sensor exemplarily.

Kurzfassung

Neigungssensoren werden heutzutage vielfältig eingesetzt und sind oft Teil komplexer Messsysteme. Bei deren Einsatz im Bauwesen ist für die meisten Anwendungen eine Genauigkeit von ca. 0.01° ausreichend. Trotz dieser auf den ersten Blick nicht besonders herausfordernd erscheinenden Genauigkeit ist es notwendig, das Verhalten des Neigungssensors sowohl im statischen als auch im kinematischen Einsatz zu kennen, um diese Genauigkeit unter jeglichen Bedingungen einhalten zu können. Hersteller von Neigungssensoren stellen allerdings nicht immer oder nur teilweise Qualitätsinformation über ihr Produkt zur Verfügung. Daher sind Tests dringend notwendig. Wir stellen in diesem Beitrag Testeinrichtungen und –abläufe zur Untersuchung von Neigungssensoren vor, mit deren Hilfe einige grundlegende statische und dynamische Kenngrößen abgeleitet werden können. Die Möglichkeiten der Testeinrichtungen und die Ergebnisse für einen ausgewählten Sensor werden gezeigt.

1. Introduction

Today, tiltmeters are often part of measurement systems that are used in static and kinematic applications, see [1] to [4] for examples. In civil engineering applications, an accuracy of 0.001° is often sufficient for the inclination measurements. With this accuracy, a pre-fabricated component (that may be part of a larger structure) of 10 m in length, for example, may be aligned vertically within a tolerance of 10 mm. However, this tolerance must be preserved in static and kinematic conditions.

Usually, a tiltmeter is attached to the component and cannot be removed during its guiding process. Typical time spans for such a guiding process of a component are several hours up to a few weeks. Therefore, the long term performance, the zero point stability, the temperature dependence and the dynamic properties of a tiltmeter are of special interest. Not all manufacturers provide information about these parameters, and especially the specifications of low-cost sensors are often incomplete, see chapter 2 for example. Therefore, additional information about the sensor is necessary in order to guarantee the specified tolerance.

We have established a simple testing sequence (chapter 3) to acquire basic information about individual inclination sensors. Using this sequence, static and dynamic properties can be determined. In chapters 4 and 5, the static and kinematic testing facilities needed for testing are described. The testing sequence was applied to a sample of tiltmeters, and the results will be shown in chapter 6 exemplarily.

2. Tiltmeters

2.1. Tiltmeter samples

There is a variety of different tiltmeters available on the market. Within the last years, we have purchased several inclination sensors from different manufacturers with different operating principles, different precisions and various areas of application. Table 1 gives an overview about these sensors. In addition, the sensor's basic operating principle, as well as its output (analogue, digital), its working range (R), resolution (q), accuracy (a), linearity (l) and stability of zero offset ($\Delta\beta^0$) is listed. The listed information was extracted from manuals or other sources provided by the manufacturers.

Manufacturer	Model	Operating principle	A/D	$R[°]$	$q[°]$	$a[°]$	$l[\%]$	$\Delta\beta^0[°]$
AOSI	EZ-Tilt 5000-15 VIB	fluid / electrodes	A	15	0.004	—	< 0.3	—
Applied Geomechanics	Little Dipper Mod. 906	fluid / electrodes	A	10	0.005	—	0.8	—
Crossbow	CTXA02-T	seismic / capacitive	A	75	0.050	< 0.500	< 0.5	0.2
GeoKon	Model 6350	seismic / vibrating wire	D	10	0.003	0.02	0.3	—
HL-Planartechnik	NS-5/P2	fluid / electrodes	D	5	0.001	0.005	< 0.08	—
Interfels	EL 211.7115	fluid / electrodes	A	10	< 0.001	—	< 0.03	—
KELAG	SCA124T-D04FA	seismic / capacitive	A	90	0.003	—	—	< 0.03
MOBA	Dual Axis Slope Sensor	fluid / electrodes	D	60	0.010	—	0.03	0.1
Schaevitz	LSRP – 14.5	pendulum / servo	A	15	< 0.001	0.006	—	—
SEIKA	N3	fluid / capacitive	A	30	0.005	—	< 0.2	—
Wyler	Zerotronic Type 3-10	pendulum / capacitive	D	10	< 0.001	—	< 0.001	—

Table 1: Sample of tiltmeters and basic specifications

The listed specifications differ widely, and sometimes there is even no information available for a specific property (indicated by "—").

The choice of tiltmeters was based on recommendations given in the literature or personally by other users. Actually, the sample does not include sensors with automatic reversal measurement, mainly because of their high costs.

2.2. Preparation of the sensors

The base plate of some sensors is not adequate for the precise determination of zero offsets (e.g., bad quality in flatness). Therefore, each sensor was attached to a separate base plate that provides three-point bearing and therefore allows the reproducible set up of the sensor.

For the digital sensors, specific data acquisition software had to be programmed to get at least data rates of 10 Hz with a standard PC. Data conversion from raw (e.g., [V], [mA]) to angular units [°] is also performed by these routines. New calibration parameters had to be determined for the sensors, where no or inaccurate parameters were provided by the manufacturer.

3. Testing sequence

For modern instruments, often little information about the operating principle, the setup or the calibration functions applied by the manufacturer is available. Then, the sensor may be considered as a general measurement system that transforms an input signal $x(t)$ into an output signal $y(t)$, e.g. [5].

In the case of a tiltmeter, $x(t)$ will be the inclination applied to the object carrying the sensor. The output $y(t)$ is dependent on the characteristics of the individual components of the sensor. Additionally, the output may be affected by external disturbances. Such disturbances may be caused by vibrations or changing temperatures for example.

There are two simple possibilities to obtain information about the measurement system. The output $y(t)$ is investigated, (a) while external disturbances are applied to the system and the input $x(t)$ is kept constant, or (b) while the input $x(t)$ is changed and no disturbances are present. Both are used in the test sequence described in the following.

A commonly known error source in inclination measurements is the zero offset β^0, which can be determined easily by reversal measurement, [6]. Changing conditions may cause changes of β^0. This is why the repeated determination of β^0 is a central part of the testing sequence. Additionally, tests on the sensor's self-heating and temperature dependence as well as kinematic tests are performed. An outline overview of the testing procedure is given in table 2.

Tests A and B should show the performance of a sensor when self-heating. The fully acclimatised sensors were switched on immediately before starting the measurement. Tests C followed in order to determine β^0 for the sensors at working temperature. For the investigation of the acclimatisation performance (tests D and F), the sensor was heated up to $+40°C$ and cooled down to $-10°C$ in a climatic chamber, respectively

(storage period in the chamber at least 7 hours). Afterwards, they were put to a granite plate (fig. 1), and data acquisition started. In order to avoid effects due to self-heating, all sensors were switched on in the climatic chamber and data output was active. Again, the determination of β^0 followed tests D and F. In order to investigate the stability of β^0 whilst the kinematic tests (test I), β^0 was determined before and after this test, too.

#	Name of the test	duration
A	determination of β^0	1 h
B	self-heating performance	> 36 h
C	determination of β^0	1 h
D	acclimatisation(temperature change: $+40°C \rightarrow +20°C$)	> 36 h
E	determination of β^0	1 h
F	acclimatisation(temperature change: $-10°C \rightarrow +20°C$)	> 36 h
G	determination of β^0	1 h
H	determination of β^0	1 h
I	kinematic test (3 runs)	0.5 h
J	determination of β^0	1 h

Table 2: Overview of the test cycle

Reversal measurements were used to determine β^0, where data were sampled for 60 seconds in each position. Prior to data acquisition, a waiting period of 60 seconds was used in order to avoid settling errors. The determination of β^0 was repeated 10 times in every test to check the repeatability.

For all static experiments (A-H, J), the sampling period was set to 10 Hz. This is considered to be a good compromise, when considering the need of many kinematic applications. However, some digital sensors provide lower sampling rates.

The duration of tests B, D and F was longer than 36 hours in order to be able to detect drifts after the acclimatisation period in the beginning.

For practical reasons, the static tests were carried out simultaneously for several sensors (in our case four). Contrary, the kinematic tests were done individually, as the kinematic test facility (see chapter 4) can carry only one sensor. Moreover, the sampling frequency was higher in the case of kinematic testing (9600 Hz for analogue sensors, maximum sampling frequency possible for digital sensors), based on the idea to determine natural frequencies.

If a project gets to a well defined stage and more information about the site conditions

become available, a more intensive investigation of appropriate sensors should be carried out. At least, the whole test cycle should be repeated for the sensors under consideration.

4. Static test facility

The performance of tiltmeters might depend on various factors (e.g., temperature, mechanical stress). When testing an instrument, only one of these factors should be varied to get the system's response for one single dependency. Testing is done best in a laboratory with constant temperature and a stable foundation. The geodetic laboratory of the institute (approx. 33×6.5 m^2 in size) is climatically controlled (temperature: $22.0°C \pm 0.5°C$, humidity: $50\% \pm 10\%$) and set up on a stable base.

For the tests described above, a stable base plate is needed in order to allow the reproducible set up of the sensors. For that purpose we use a plane granite plate which is set up in the laboratory. Figure 1 shows this plate with the required mounting accessories and four sensors under test.

Figure 1: Granite plate with mounting accessories and sensors under test

The stability of the granite plate is continuously monitored by a Leica Nivel20 inclination sensor, which is known to be highly precise under stable thermal conditions. Figure 2 shows the result of this continuous monitoring during a time period of two years.

A heat radiation shield (pink Styrofoam plates in fig. 1) was used for not disturbing the Nivel20 during the acclimatisation experiments. However, it was still slightly disturbed (outliers indicated by red dots in fig. 2).
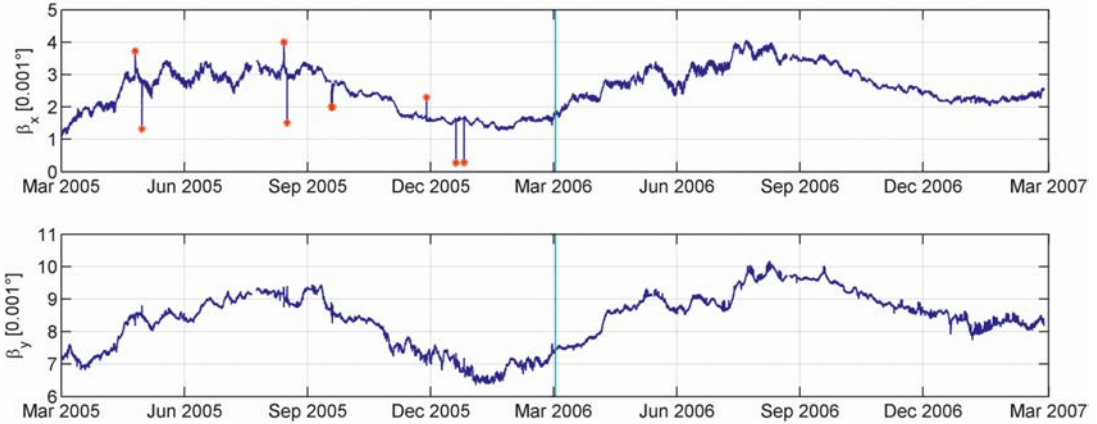
Figure 2: *Biaxial Stability of the granite plate monitored by a Leica Nivel20.*

An annual cycle is apparent for both axes, where the inclinations change for about 0.003°. Thus, the stability of the granite plate is sufficient for the static tests, which do not last longer than two days.

The laboratory also houses a climatic chamber (volume 1 m^3, temperature range: -10°C to $+40$°C), which was used to expose the sensors to different temperatures prior to the acclimatisation experiments.

5. Kinematic test facility
5.1. Testing functions

The performance of measurement systems may be investigated using testing functions, see [5] for example. The response $y(t)$ of a system is investigated by changing the input $x(t)$ according to a known testing function. The kinematic test facility described later was constructed in order to generate a testing function $f_P(t)$, consisting of a ramp, $f_R(t)$, and a step function, $f_S(t)$, as shown in fig. 3.
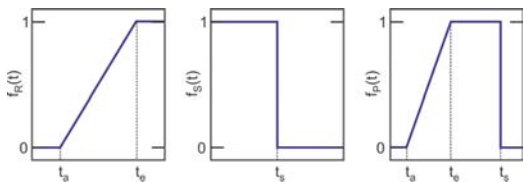

Figure 3: *Ramp (left), step (middle) and composed testing function (right).*

A tiltmeter represents a measurement system of second order, for example see [6] or [7]. Applying the testing function $f_P(t)$ to the system

and investigating the system response, typical parameters like the natural frequency f and the damping coefficient δ may be determined (for details see [8]).

5.2. Design and experimental set up

A specific test facility is needed in order to apply the testing function $f_P(t)$ (fig. 3, right) to the sensor under test. Its principle is illustrated in fig. 4.
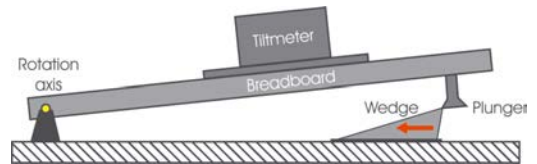
Figure 4: *Principle of the kinematic test facility.*

The tiltmeter is attached to a breadboard, which may be rotated about a horizontal axis. The rotation is induced by a horizontally moving wedge (1 mm/s) that is attached to a translation stage and lifts up a plunger, which is fixed to the bottom of the breadboard. The amount of the inclination change (ramp function) is about 0.45°. After the plunger has reached the upper end of the wedge, the breadboard suddenly returns to its initial horizontal position and thus realizes the step function. The height of the fall does not exceed 2 mm, in order to avoid damaging the sensor by the shock induced by the impact. The time of the impact and its load after the fall are determined by using accelerometers mounted additionally at the breadboard (fig. 5). During the experiments, impact loads of up to 40 g were measured.

Figure 5: Kinematic test facility with attached tiltmeter and accelerometers.

The kinematic test facility is set up in the laboratory of the institute on the same concrete table as the granite plate (see chapter 4). Using this facility, characteristic parameters like the natural frequency, the damping coefficient or the decay time may be determined for each individual sensor.
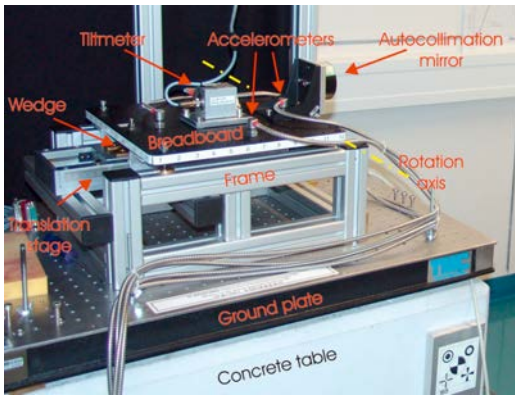
Three PCs and a data logger are used to operate the test facility. Time synchronization between the PCs is done using a synchronisation signal recorded simultaneously by all PCs (for details see [8]). In order to determine the time and load of the impact precisely, a sampling frequency of 9600 Hz was used.

For the analysis of the inclination data, reference values for the testing function realized by the testing facility are necessary (comparison of sensor output with reference values). These values were determined using a highly precise sensor (Wyler Zerotronic). In order to keep dynamic influences as low as possible, the velocity of the translation stage was rather slow (0.1 mm/s) in this case. These values were controlled at 45 discrete points using autocollimation with a Wild T2000 theodolite. Differences between the two methods did not exceed 0.001°, which is sufficient for our investigations ($\sigma < 0.01°$).

6. Analysis and results

All sensors listed in table 1 were tested in the same way, using the test sequence described in chapter 3. Here, the results of one sensor (Kelag SCA124T – D04FA, referred as KE in the following) are shown exemplarily. More details and further results are provided in [8] and [9].

6.1. Temperature experiments

Data were acquired with 10 Hz in order to get similar heating-up effects of the sensor as they might occur in field use. Before further processing, data of consecutive 10 second intervals were averaged in order to reduce data. This is applicable for the static tests, where inclinations change slowly.

Figure 6 shows the inclinations β measured with sensor KE for test B and F (see table 2) exemplarily. The values shown are already reduced for β^0 (determined by test C and G, respectively). Thus, the data at the end of the plots represent the true inclination of the granite plate, which was independently determined with the Nivel20 as $-0.009°$.

During both tests, the stability of the granite plate was better than 0.001°. Thus, values differing from $-0.009°$ are originated by the self-heating and the acclimatisation of the sensor. In the beginning of the time series conditions change rapidly and strong variations can be seen. These variations are much larger than the resolution of the sensor (see table 1). Non-linear drifts as well as sudden inclination changes or even peaks are visible. These effects may be caused by the sudden expansion of individual parts of the sensor. As such effects are hardly reproducible and thus cannot be modelled, the range (rg) covered by the changes is numerically given in fig. 6 as one indicator. The end of this initial phase ($EoIF$) is indicated by a dashed vertical (red) line. After $EoIF$, the conditions are stable and the output values should remain constant. But even then, small linear drifts are present.

The position $EoIF$ is determined by means of linear regression analysis. This was favoured over e.g., modelling a measurement system of first order, because some sensors (so does the KE) correct the measured values by means of internal temperature measurements, so the remaining effects may differ significantly from a modelled system. Computation is done using outlier detection in order to find the section in the beginning of the time series. The a-priori variances σ^2 needed for the model test were determined for each sensor empirically, using a linear sample of the data set. In the case the empirical variance gets zero (i.e., for sensors with low noise and a large resolution), the variance of quantisation noise ($\sigma^2 = \frac{q^2}{12}$, see [10], pp.193) is used instead.
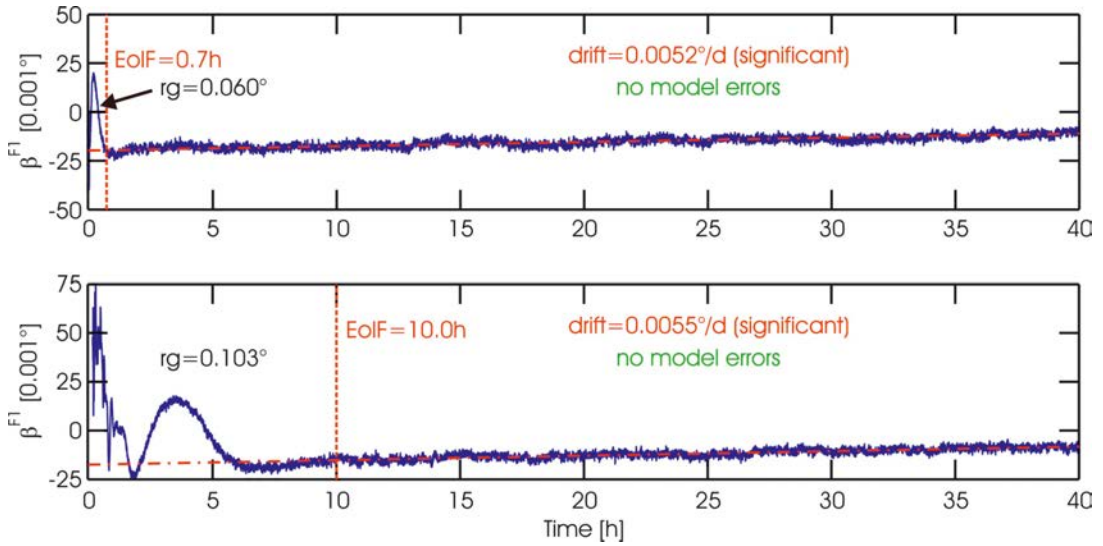
Figure 6: Inclination changes of sensor KE, caused by self-heating (test B, upper) and acclimatisation from −10° C to +20° C (test F, lower).

The slope of the regression line is an indicator for the drift of the sensor. It is determined by the data after $EoIF$ only. A numerical value is given for the drift, if the slope significantly deviates from zero (confidence level 99.7%) and its magnitude is larger than a threshold. There, the threshold is set by practical considerations, e.g., 0.002° per day if the sensor is used five days continuously in the guiding process of an object and the effect of the drift should remain smaller than 0.01° within that time.

6.2. Kinematic results

The difference ($\Delta\beta$) between the reference values of the kinematic test facility and the output of the sensor is shown in the upper part of fig. 7. The blue line represents the originally sampled data (9600 Hz, see section 5.2), and the yellow line shows a moving average (10 Hz, which may be a proper sampling frequency for various applications). Caused by the testing function $f_P(t)$, fig. 7 exhibits four parts: a region before the ramp (a), a section during the slope of the ramp (b), a part after the ramp (c) and an area after the step (d), see lower part of fig. 7.

Regarding the sections (a) and (d), where no inclination changes occur, robust means (m) and standard deviations (s) were calculated. The median absolute deviation (MAD, [11]) was used for that purpose. The means m and the confidence intervals (95 %) are depicted by red

and green lines, respectively. For the case of this sensor, there is no significant shift (Δm) in the robust means before (a) and after the impact (d).

The following sensor properties may be obtained by investigating the sensor's behaviour in sections (a) to (d) of fig. 7:

- In the case of an ideal sensor, $\Delta\beta$ equals zero all the time, i.e. reference values and sensor output are identical.
- A shift in the robust means of section (a) and (d) would indicate a change of the zero point β^0 due to the impact. In both sections, the sensor's output should hold a constant level, since no inclination changes take place.
- A delayed response of the sensor causes an offset in the ramp section (b).
- If a slope appears in the ramp section (b), data of the sensor may be affected by a "scale factor". However, this is only representative for the small range of the test facility (0 to 0.45°) and may rather be caused by the calibration function used by the sensor. If needed, the scale factor over the whole working range of the sensor must be determined separately.
- After the ramp, section (c), the values for $\Delta\beta$ should return to the zero line. A remaining offset would be the consequence of the afore-mentioned "scale factor".

The section of the step response is shown in fig. 8 in detail. This illustration corresponds to the end of

section (c) and the beginning of section (d) in fig. 7. The time stamp of the impact (dashed vertical line) is determined by the signal of the accelerometers.

The step response of sensor KE exhibits the behaviour of a measurement system of second order (section 5.1), with an additional exponential term. The eigenfrequency f was determined as $f = 75\,Hz$. After the impact, the signal oscillates with an initial range (rg) of 9.7°. To characterize the decay behaviour, three thresholds were defined: (i) The value t_1 corresponds to the time after which a band of ±5 % of the maximum amplitude is not exceeded any more [5]. (ii) Time t_2 indicates when the signal stays within a confidence level of 95 %.

(iii) The threshold t_3 marks the time stamp when the oscillation is below 0.01°, which was the aim of the investigations. Figure 8 shows that the inclination signal does not exceed the 5 % bounds after $t_1 = 0.05\,s$ and the confidence bounds after $t_2 = 0.10\,s$. Decay time t_3 could not be computed in this case (precision of the signal approx. 0.035°). The short decay times indicate the strong damping of the sensor ($\delta > 100$). At a practically acceptable sampling rate of 10 Hz, for example, the magnitude of the impact's reaction decreases dramatically down to 0.12° due to averaging effects. More details and results of other sensors are given in [8].
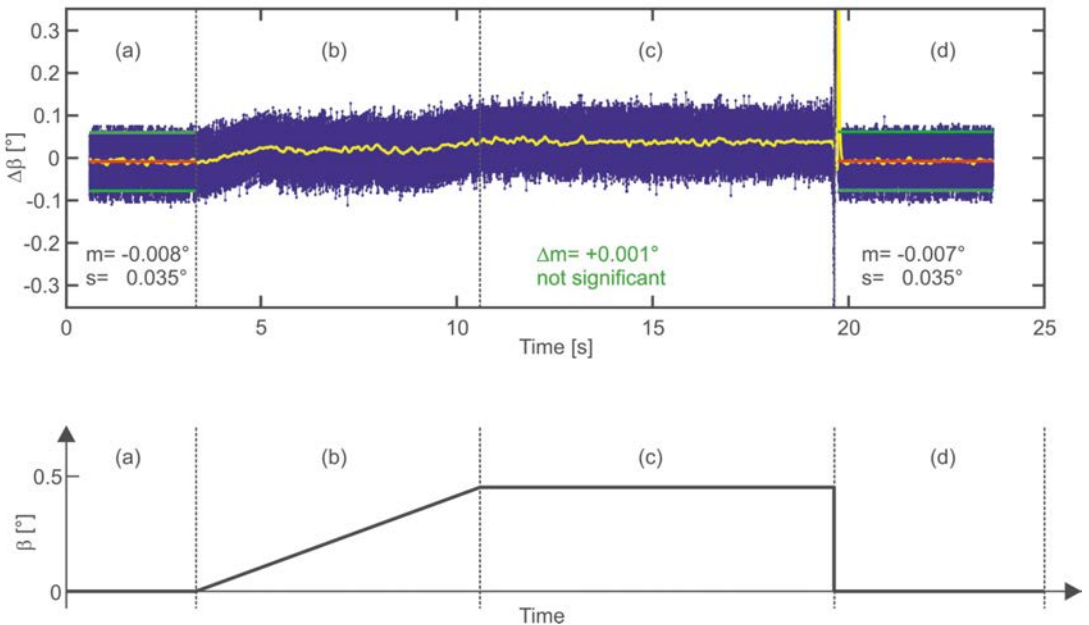

Figure 7: Difference between the reference values of the kinematic test facility and the sensor's output (upper) and the related sections of the testing function (lower).
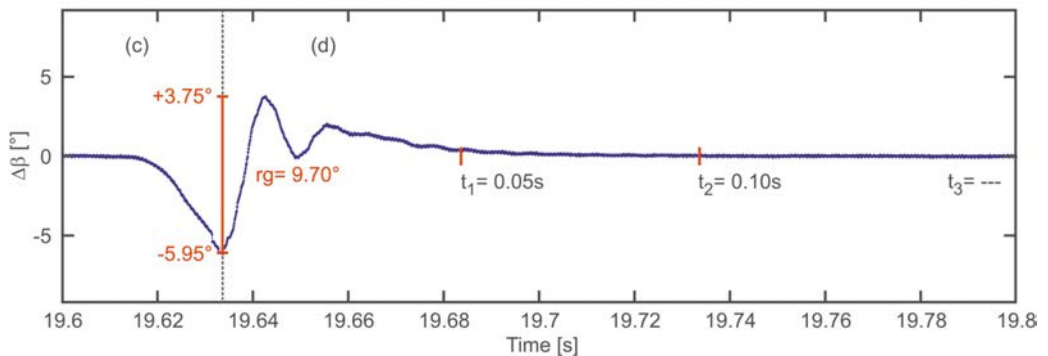

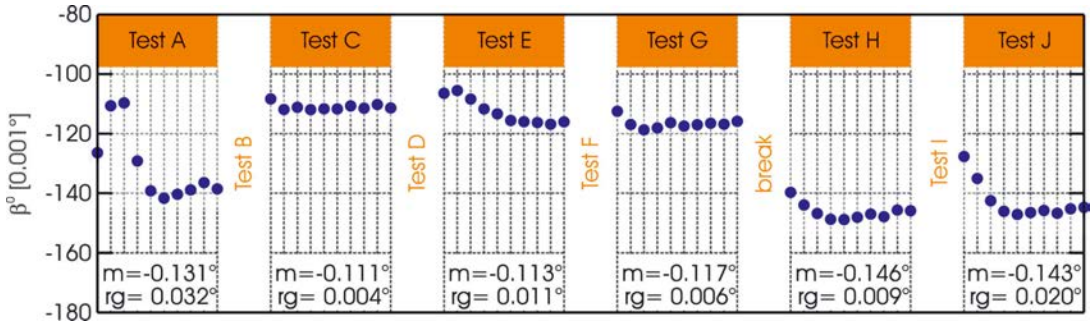Figure 8: Step response of sensor KE.

Figure 9: Zero points of the sensor KE.

6.3. Stability of zero points

In the case of the reversal measurements for the zero points (60 sec data with 10 Hz), data evaluation is quite simple. After robust outlier detection, the remaining values are averaged to give one inclination value per set-up (β^{F1}, β^{F2}). Using these values, the zero point β^0 can be easily computed as well as the inclination β of the surface (granite plate):

$$\beta^0 = \frac{\beta^{F1} + \beta^{F2}}{2}, \quad \beta = \frac{\beta^{F1} - \beta^{F2}}{2} \qquad (1)$$

The resulting zero points for the different experiments are shown in fig. 9 for the sensor KE.

For the assignment of the different tests, table 2 may be used. During each test, β^0 was determined 10 times and each individual β^0 is shown as a blue point. The mean value (m) and the range (rg) of the zero points are given numerically.

Sensor KE shows some variations of β^0 in the beginning of experiment A, which are caused by the self-heating of the sensor. The magnitude is comparable with the changes determined by test B in the initial phase (see fig. 6, upper plot). For the sensor at operating temperature (test C), only little variation is present. The values for β^0 determined immediately after the acclimatisation experiments (tests E and G) show a drift, which may be caused by a release of tensions during reversing the sensor. Tests H and J were carried out a few months later (after the establishment of the kinematic test facility), which may be responsible for the offset of β^0 determined within these tests. However, the observed variations of β^0 are within the range of the stability given by the manufacturer (0.03°). More details are given in [9].

7. Resume

There are large differences in the specifications of commercially available tiltmeters. The lack of some specifications, crucial for certain applications, can only be overcome by individual testing. We have presented some test facilities and a procedure which allow the determination of some basic static and dynamic properties of tiltmeters. The derived static parameters include zero point stability, self-heating effects, temperature dependencies and long term stability. The dynamic parameters comprise natural frequencies, damping characteristics and decay times.

The derived information is the same for all sensors under test, thus allowing a comparison of different sensors, which cannot be done easily using the varying information provided by the manufacturers. Of course additional investigations (e.g., linearity over the whole working range) have to follow, if the sensor was selected for a specific application. However, the presented facilities and procedures may be used to exclude sensors not fulfilling the defined criteria.

The shown sensor KE shows a desirable performance within the static and kinematic tests. The results do not disagree with the manufacturer's specifications. Although the specified stability of the sensor (0.03°) is little worse than our aspired 0.01°, the performance of KE is sufficient for many civil engineering applications.

R. Ruland, Stanford Linear Accelerator Center, Stanford University, and A. Bühlmann, Leica Geosystems AG, provided several tiltmeters. Their cooperation and interest in our investigations are gratefully appreciated.

References

[1] *R. Glaus, N. Lauener, U. Müller and M. Baumeler (2004):* Der Gleismesswagen swiss trolley: Leistungsmerkmale und Anwendungen. In: Ingenieurvermessung 2004 pp. 27-37. 14th International Course on Engineering Surveying. Zürich, March 15-19, 2004

[2] *W. Stempfhuber (2006):* 1D and 3D Systems in Machine Automation. In: Proceedings of the 3rd IAG / 12th FIG Symposium, Baden. May 22-24, 2006

[3] *H. Ingensand (1985):* Ein Beitrag zur Entwicklung und Untersuchung hochgenauer elektronischer Neigungs-messsysteme für kontinuierliche Messungen. Series C, Book No. 308, Bavarian Academy of Sciences, Munich

[4] *E. Grillmayer, A. Wieser and F. K. Brunner (2000):* Einrichtung der Stahlpylone der Donaubrücke bei Pöchlarn. In: Ingenieurvermessung 2000 pp. 63-74. 13th International Course on Engineering Surveying. Munich, March 13-17, 2000

[5] *P. Profos and T. Pfeifer (2001):* Grundlagen der Messtechnik. Fifth Edition. Oldenbourg, Munich

[6] *W. Schwarz (1995):* Vermessungsverfahren im Maschi-nen- und Anlagenbau. Wittwer, Stuttgart

[7] *D. Gross, W. Hauger, W. Schnell and J. Schröder (2004):* Technische Mechanik 3. Kinetik. Eighth Edition. Springer, Berlin

[8] *K. Macheiner, H. Woschitz and F. K. Brunner (2007):* Test von dynamischen Eigenschaften ausgewählter Neigungssensoren. (Submitted for publication)

[9] *H. Woschitz (2007):* Statische Eigenschaften ausge-wählter Neigungssensoren. (Submitted for publication)

[10] *A. V. Oppenheim and R. W. Schafer (1999):* Discrete-Time Signal Processing. Second Edition. Prentice-Hall, Upper Saddle River (NJ)

[11] *L. Sachs (1997):* Angewandte Statistik. Anwendung statistischer Methoden. Eighth Edition. Springer, Berlin

Contact

Dr. Helmut Woschitz: Engineering Geodesy and Measure-ment Systems, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: helmut.woschitz@tugraz.at
Dipl.-Ing. Klaus Macheiner: Engineering Geodesy and Measurement Systems, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: klaus.macheiner@tugraz.at   vgi

A critical assessment of the current EGNOS performance

Michael Opitz, Robert Weber and Wolfgang Winkler

Abstract

The main purpose of this paper is to evaluate the current performance of the European Geostationary Navigation Overlay Service (EGNOS) in comparison to commercial, local DGPS services. In full operational capability (FOC) EGNOS provides orbit and clock-corrections of all GPS satellites, ionospheric delays and integrity information of the GPS system.

The analysis is mainly based on the comparison of the trajectories of a slowly moving vehicle obtained simultaneously by two real-time correction techniques – EGNOS and WEP (Wienstrom Positioning Service Provider). The tests are carried out in urban environments with frequently varying obstructions and on a highway. Therefore the visibility of the EGNOS satellites varies during the test ride. During the trial session also raw data of the rover receivers as well as the reference station was logged. This allows to verify a posteriori the calculated real time position with respect to a reference of sub dm accuracy. Additionally an evaluation of the EGNOS ionospheric model is presented.

Kurzfassung

Die Hauptmotivation für die Arbeit ist die Untersuchung der derzeitigen Verfügbarkeit, Genauigkeit und Stabilität des European Geostationary Overly Service (EGNOS) im Vergleich zu kommerziellen, lokalen DGPS- Anbietern. Im Vollausbau (full operational capability – FOC) wird EGNOS Satellitenbahn- und Uhrenkorrekturen für alle GPS Satelliten, ionosphärische Laufzeitverzögerungen und Integritätsinformationen für das GPS System aussenden.

Die präsentierte Analyse basiert auf dem Vergleich zweier Trajektorien eines sich langsam bewegenden Fahrzeuges. Die Koordinatenlösungen werden gleichzeitig mit zwei unterschiedlichen Echtzeit Korrekturtechniken bestimmt – EGNOS und WEP (Wienstrom Positioning Service Provider). Die Testfahrt wurde in unterschiedlich bebautem Stadtgebiet und auf einer Autobahn ausgeführt, wodurch die Sichtbarkeit der geostationären EGNOS Satelliten während des Tests stark variierte. Um a posteriori die in Echtzeit bestimmten Positionen kontrollieren zu können, wurden die Rohdaten der Rover Stationen und einer Referenzstation gespeichert. Ergänzend wird auch eine Evaluierung des EGNOS- Ionosphärenmodells präsentiert.

1. Introduction – Differential GPS (DGPS) vs. Satellite Based Augmentation Systems (SBAS)

The positioning accuracy of a single frequency GPS- receiver used under optimal conditions (no obstructions and multipath effects, good satellite geometry) is in the range of about ten meters. However, for a wide range of applications an improved accuracy is needed. While random errors as the code measurement noise are receiver dependent and the hard to detect and to reduce multipath effects relate to the permanently changing environment a real improvement of positioning can be achieved by minimizing the systematic error influences on the raw measurements.

Systematic errors are mainly composed by inaccuracies in the satellite orbit- and clock navigation information and by atmospheric effects. The orbit- and clock errors can influence the raw range measurements each by about two meters. The atmospheric error can be separated in a ionospheric part and a tropospheric part. The tropospheric delay adds up to about 2,3 meters in zenith direction and is slightly affected by local weather conditions. The ionospheric delay on the contrary mainly depends on the geographic latitude, the local time and last but not least on the current activity of the sun. Dependent on the elevation of the satellite this delay can add up to several tenths of meters. Both atmospheric delays can be significantly reduced when using correction models (for a detailed description of GPS error sources see [8]).

An option to account for systematic errors in real-time is provided by different augmentation systems like WAAS or EGNOS (see weblink [6] and [1], [2]). All these systems make use of observations gathered in coordinative known base station networks. The mean distance between these stations is usually several hundreds of kilometers. In the systems central

computer facility the network error models for the service area are calculated and then transformed into correction data that is transmitted to the user community (see [2] and weblink [3]). Another option is to make use of correction data offered by 'Local Service Providers' which might be private companies or national mapping authorities. Mean station distances within their GNSS networks are usually 50-80 kilometers (see weblink [7]).

The correction data is broadcasted by means of two different standard formats. Local services make use of the RTCM format (see [7] available via web link [5]) for the transmission of corrections. For example the RTCM v.2.3 message types 1 and 21 contain range corrections for raw pseudorange measurements. These corrections implicitly cover most of the systematic errors. The rover receiver adds the corrections to the raw measurements and does not need to apply any additional corrections (e.g.: error models for ionospheric or tropospheric effects).

Contrary, regional SBAS services use the RTCA format (see [6], available via web link [4]). This data format issues corrections (ionospheric delay, satellite position- and clock correction) valid for the whole service region. Therefore the receiver needs to calculate the range corrections for his position in a primary step before they can be added to the raw measurements.

One of the advantages of local networks is that their correction data covers the tropospheric delay whereas regional networks do not provide this information due to the fact that the tropospheric delay varies locally. A SBAS receiver therefore has to apply a tropospheric model. The update rates of RTCM corrections are generally one second while those for RTCA corrections are lower. For this reasons the positioning results obtained in local networks are generally more accurate. The expected DGPS coordinate accuracy is at the sub- meter level. SBAS positions are expected at the two meter accuracy level. It has to be noted that a major focus of SBAS Systems is the Integrity Monitoring of the GPS satellites (usually not provided by local DGPS services).

A further difference between local DGPS- and SBAS- systems is the transmission of the correction data. DGPS services mainly use GSM connections (in former days also radio links) whereas SBAS Systems broadcast their data via geostationary satellites. Due to the low elevation of geostationary satellites SBAS signals cannot be received everywhere. Therefore the EGNOS signals are also distributed via Internet (SISNeT (Signal in space through the Internet), see [9] and [10]). Via SISNeT the corrections can be received also in heavily obstructed areas.

A potential user has to take into account that the EGNOS signal can be used for free and no costs emerge for the data connection – except when using SISNeT. Local DGPS services charge for their data. Additionally the GSM data link has to be paid.

2. Purpose and description of the experiment

In this paper we want to carry out a comparison of vehicle trajectories applying on the one hand correction data from a local DGPS service provider and from the European SBAS System EGNOS on the other. These trajectories are compared to an aposteriori calculated precise trajectory. The main purpose is to test the current EGNOS performance. A rating between DGPS services and EGNOS has to be based on several components like accuracy, availability, price and intended application and has therefore to be left to the user.

The two different systems tested in this paper are:

■ The local DGPS network WEP

■ The regional SBAS network EGNOS

WEP (see web link [7]) is operated by the Viennese power supplier Wienstrom in cooperation with the OEBB (Austrian Railways) and the BEWAG (power supplier of Burgenland). The network currently covers 12 GPS/GLONASS reference stations in the east of Austria. The central processing unit which provides RTCM data to the users via a GSM router is located in Vienna.

EGNOS is the European SBAS System set up by the European Space Agency (ESA). About 30 RIMS (Ranging and Integrity Monitoring Stations) permanently observe the GPS satellites and transfer the data to the MCCs (Mission Control Centers). The MCCs are monitoring the integrity of the GPS system and calculate the error models and corrections. These data is transmitted by Up-Link stations (NLES – Navigation Land Earth Station) to three geostationary satellites which broadcast the RTCA corrections to the users. The system is highly redundant (e.g. just one MCC is calculating the corrections at a certain point in time – the others are purely backups). The system covers the European area and in future parts of Africa.

Figure 1: Test configuration.

In our experiment we wanted to test the performance of both systems under the same conditions at the same time. To fulfill these specifications we used one antenna and an antenna splitter to divide the received signal into two streams. Therefore both receivers got exactly the same incoming signals (see fig 1).

For the reception of the EGNOS- signal a Septentrio PolaRx2 Receiver was used. A Leica 1200 Receiver equipped with a GSM modem was deployed for operating in DGPS mode. Both receivers are originally two frequency receivers. For this experiment both receivers were set to L1 mode.

3. Static Experiment

To investigate the stability of both systems a static test on a coordinative known point was carried out. The experiment took place on March 9, 2007 between 8:10am and 16:36pm (CET) under perfect conditions (no obstructions and no multipath). The positions were logged in five seconds intervals and then transformed to local coordinates (east, north, height). Figure 2 shows the differences between the real time coordinates and the known coordinates over the whole observation span.
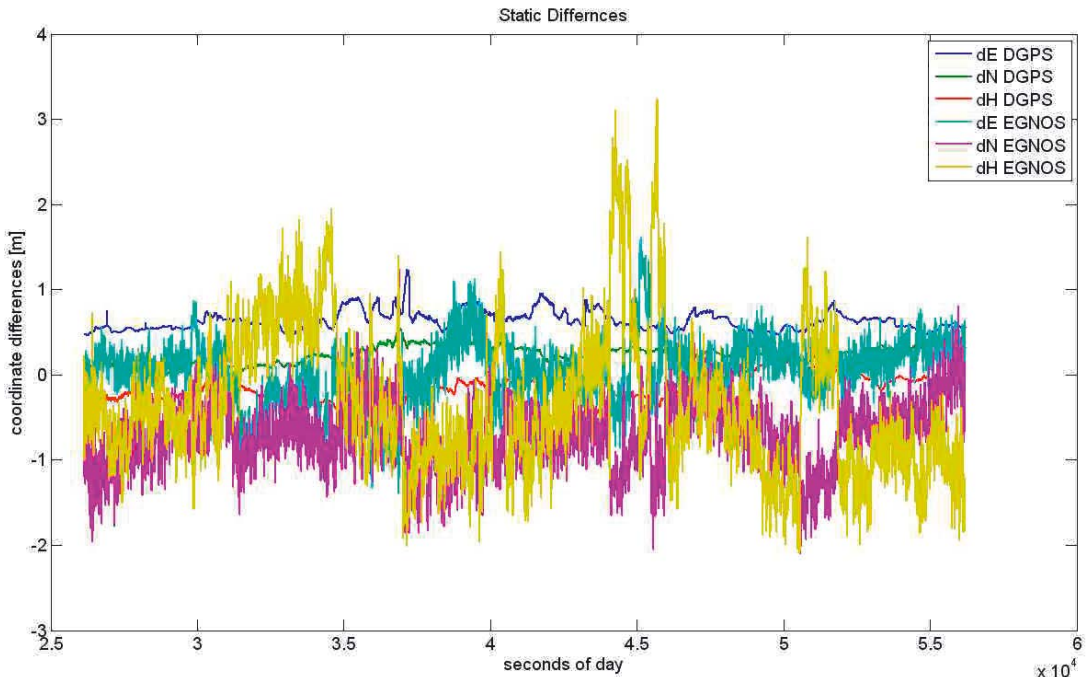

Figure 2: Differences between real time coordinates and a priori known coordinates.

The results are well in the expected range. During the whole observation time the differences between the DGPS coordinates and the reference coordinates are below one meter. No problems with the GSM connection could be observed.

The results obtained by the EGNOS receiver are noisier and the absolute coordinates differ up to more than 2 meters to the a priori known ones. Also it is clearly visible that the EGNOS solution is less stable and that outliers in the height occur occasionally.

Table 1 shows the maximum and mean values of the differences to the known coordinates over the whole observation time (6016 epochs). The superior accuracy of the DGPS coordinates can clearly be seen. The mean values of the EGNOS differences are also below one meter, but single measurements can differ up to 2,10 m in the plane and up to 3,24 m in height.

4. Kinematic Test

For the kinematic test the antenna was mounted on a car. The positions were logged every second and again transformed to the local coordinate system (east, north, height). In order to be able to calculate a posteriori a precise reference trajectory also the raw observation data was logged.

The test was carried out in Vienna on March 8, 2007 between 10:30am and 12:00am (CET). The route was around 9 km long and passed through different obstructed areas. The environment comprised the heavy obstructed inner city of Vienna, wide boulevards as well as a highway. The chosen track was therefore an ideal test area.

Figure 3 shows the 3 trajecories (green lines) on 3 maps of Vienna. Figure 3a represents the DGPS solution. The GSM connection to the DGPS provider did not fail through the entire test period. This line can be seen as the area where GPS positioning in general was possible. A part from some small pieces downtown everywhere enough GPS satellites could be observed.

The EGNOS results are represented by the graphic in 3b. As expected, the geostationary satellites were not visible in the inner city most of the time and therefore no comparison between DGPS and EGNOS coordinates could be performed for this area. It is planned to redo the experiment in this areas with a SISNeT capable receiver.

In the map 3c the a posteriori calculated reference trajectory is shown. Just the areas where the ambiguities could be fixed are plotted. A correct solution of the ambiguities was possible only for small parts of the trajectory. This can be caused on the one hand by the bad satellite geometry (and low number of satellites visible downtown) and on the other hand by the high velocity of the test vehicle on the highway.

(6016 observations)		dE[m]	dN[m]	d2D[m]	dH[m]
EGNOS	Absolute Maximum	1,62	2,10	2,10	3,24
	Mean of Absolute Differences	0,28	0,71	0,81	0,75
DGPS	Absolute Maximum	0,55	0,75	0,78	1,09
	Mean of Absolute Differences	0,23	0,18	0,32	0,22

Table 1: Statistics of the static measurements.

Figure 3a-c: Test Trajectories (a: DGPS, b: EGNOS, c: Post Processing).

(460 observations)		dE[m]	dN[m]	d2D[m]	dH[m]
EGNOS	Absolute Maximum	3,95	5,73	6,95	7,13
	Mean of Absolute Differences	0,74	0,75	1,15	0,97
DGPS	Absolute Maximum	0,51	0,70	0,72	1,16
	Mean of Absolute Differences	0,19	0,16	0,28	0,20

Table 2: Statistics of the kinematic experiment.

Table 2 shows the maximum and mean values of the differences between the EGNOS and DGPS points with respect to (w.r.t.) the a posteriori estimated positions. Solely those points were taken into consideration where DGPS and EGNOS positioning was possible and where in the post processing analysis the ambiguities could be fixed. Therefore only a total of 460 epochs was used for the statistics.

While the results for the DGPS solutions are more or less the same as in the static experiment the EGNOS results are worse. The mean values increased by 20-30cm. But the maximum values are three times worse than in the static experiment. If this effect can be led back to the movement of the car or more likely to a general poor performance of the EGNOS corrections (the system is still under development) could not be clarified and needs further investigations.

Figure 4 and Figure 6 show the differences between the real time and the a posteriori calculated reference coordinates for two sections of the trajectory where an estimation of the reference trajectory was possible.

The first part of figure 4 shows some unexplained variations in the EGNOS results with differences up to 7m. In figure 5 the three trajectories of this part of the experiment are plotted. The drift of the EGNOS results can clearly be seen. The second part of figure 4 displays the typical variations w.r.t. ground truth.

The results of section 2, displayed in figure 6, demonstrate an example for pretty good performance of both systems. Nevertheless, the benefits of the DGPS service are visible. These results are more accurate and more stable.
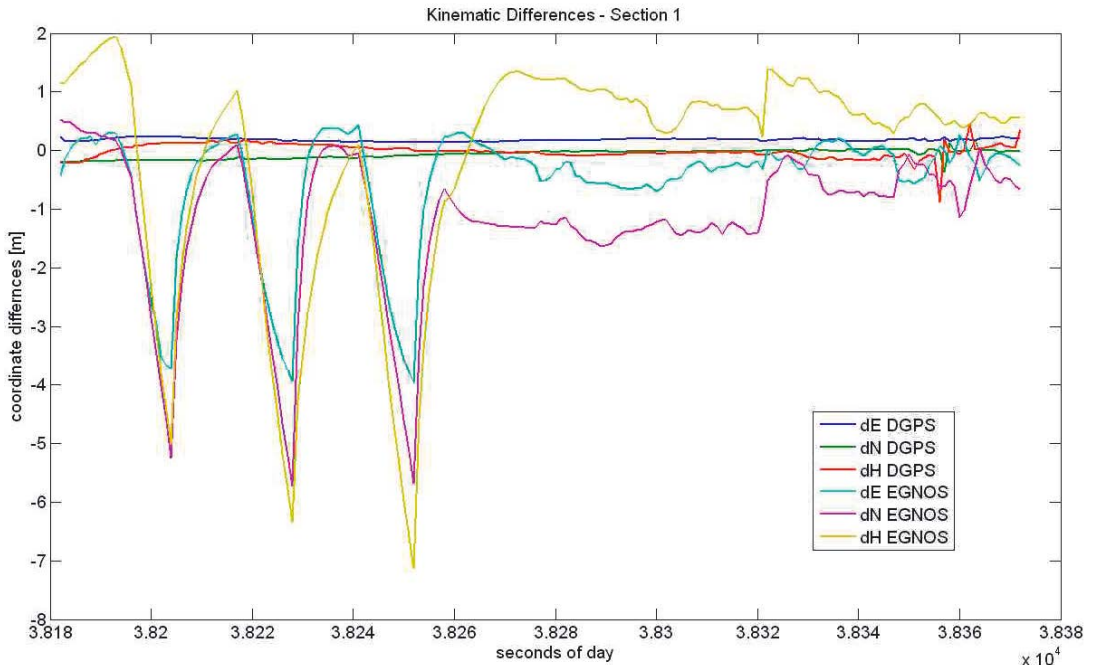

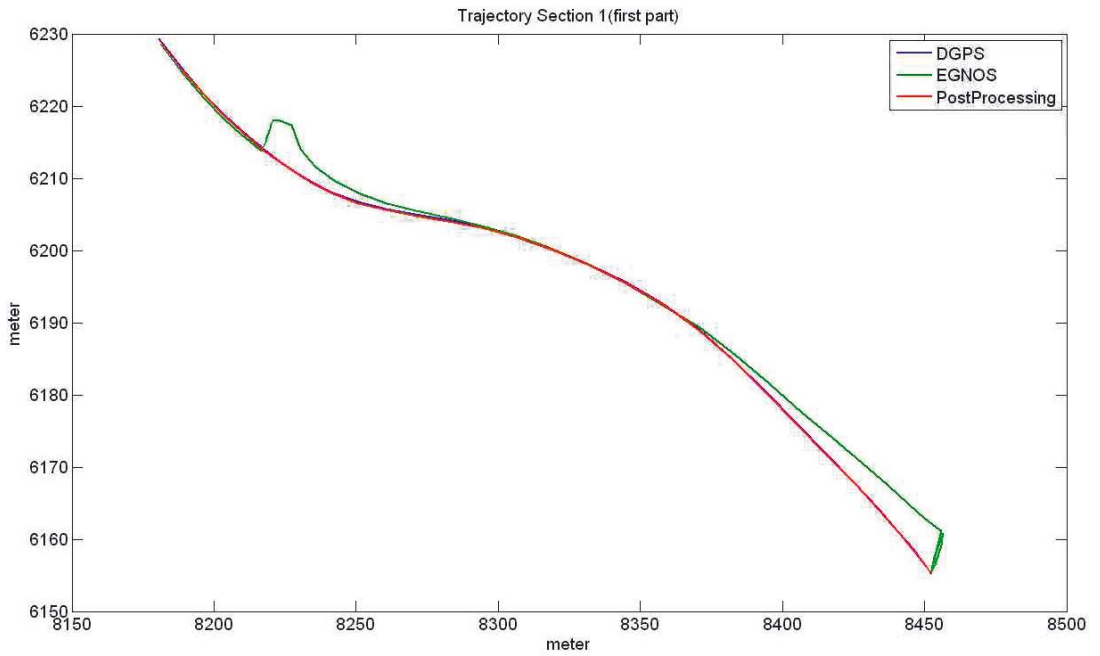*Figure 4: Differences of DGPS and EGNOS coordinates w.r.t. a posteriori estimated coordinates for section 1.*
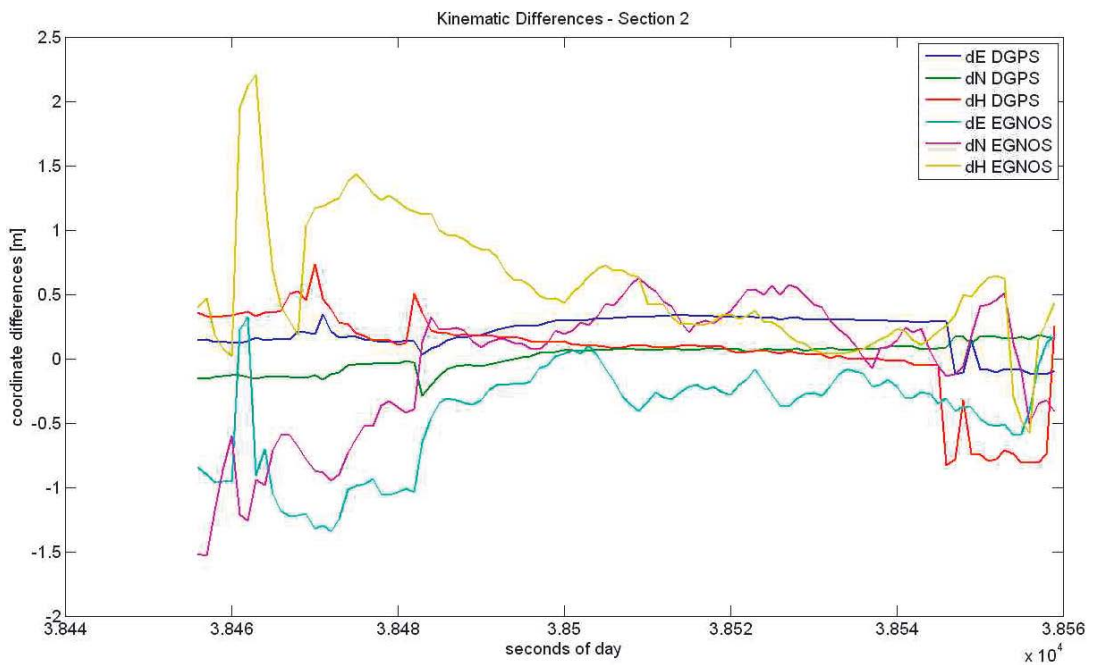
Figure 5: Trajectory plot of section 1.

Figure 6: Differences of DGPS and EGNOS coordinates w.r.t. a posteriori estimated coordinates for section 2.

5. Ionospheric Model Comparison

A very important error source when using single frequency receivers is the ionospheric delay. There are different ionospheric models to diminish the error of raw pseudorange observations. One is the Klobuchar Model (see [3]). It corrects at least 50% of the ionospheric delay. The model is described by eight parameters. A priori values for these parameters are transmitted in the navigation message of the GPS broadcast ephemerides and are valid for some days. With the knowledge of approximate coordinates, the time, the elevation and the azimuth of an observed satellite the ionospheric correction to be applied can be calculated. This model is by default implemented in single frequency receivers.

EGNOS uses the NeQuick model (see [1]). In contrary to GPS, EGNOS transmits the ionospheric delay for well defined grid points (IGPs, ionospheric grid points). The rover receiver has to interpolate the ionospheric delay for its position. The values of the ionospheric delay are updated every five minutes and the accuracy should be better than 50 centimeters. For these reasons the EGNOS model should be superior to the GPS broadcast model (when working with DGPS services the ionospheric delay is implicitly included in the correction data).

Figure 7 shows the differences of the ionospheric delay between the EGNOS- and GPS-model at the European IGPs for one day. The data interval is 2 hours. The first image (most left, upper line) shows the differences for 02:00 am (UTC); the last one (right, bottom line) shows the differences for 12:00 pm (UTC) the same day.

It is visible that the two models do not match very well. The differences range from +2,5m to -2,5m. Generally the EGNOS corrections are lower than the broadcast values. In addition, the delays at various grid points are very often missing in the EGNOS data (blue colored IGPs in fig. 7). An analysis of the ionospheric delay over several days confirmed this effect for most of the considered time (see [11]). This might be a system data processing artifact because EGNOS is still in an introduction phase and therefore the ionospheric corrections currently do not reach the expected accuracy. A detailed analysis of the EGNOS ionospheric model can be found in [11].
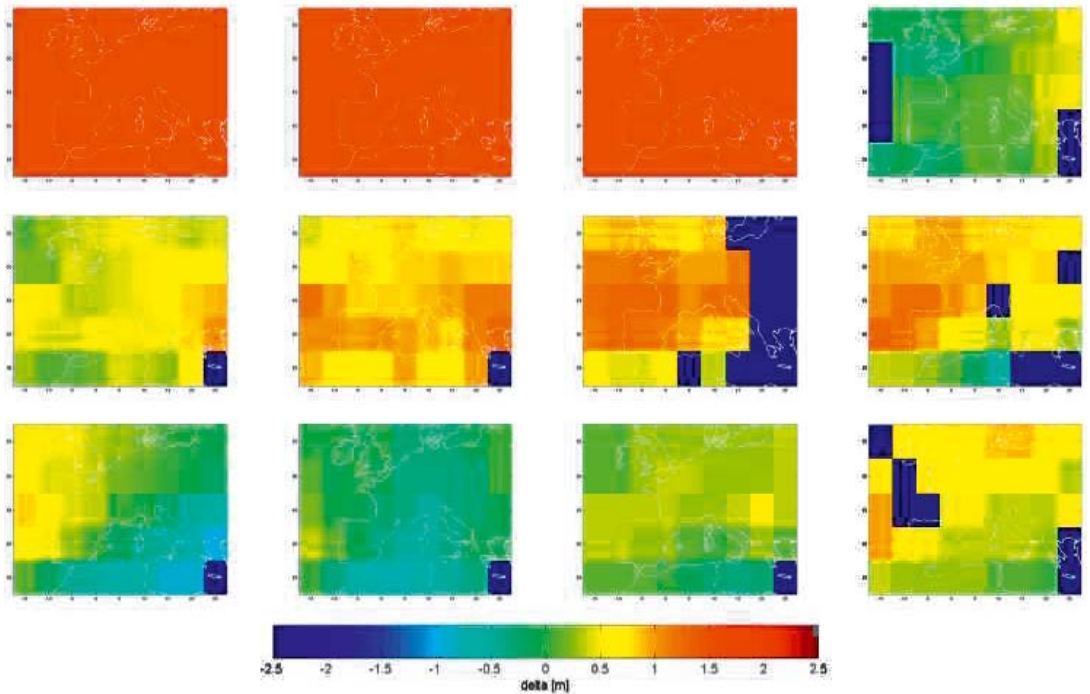
Figure 7: Difference GPS- broadcast-minus SBAS- model for day 189 (year 2006) (see [9]).

6. Summary and Outlook

In summary, the achieved results were as expected. The local DGPS service is superior to the EGNOS service concerning accuracy and availability of the service. On the other hand the user of a DGPS service has to pay for the service and for the GSM connection.

In the heavy obstructed inner city of Vienna positioning was frequently harmed or impossible. Especially the EGNOS signal could not be received there. Thus it is planned to repeat the experiment with a SISNeT receiver. Astonishing was the degraded accuracy when using EGNOS in kinematic mode. A similar behavior was not observed in the DGPS solution.

The raw data is recalculated with the software SISSIM (SISNeT Simulation) for carrying out tests with the corrections provided by EGNOS (see [2] and [3]). The comparison of different ionospheric models showed that the performance and stability of the EGNOS ionospheric corrections is not optimal so far. However, the situation might change after shifting EGNOS from the initial to regular operation.

Finally, we want to thank the company Wienstrom for offering their network services to carry out the presented DGPS tests and the Institute of Engineering Geodesy and Measuring Systems of TU-Graz for loaning the antenna splitter.

References

[1] *Aragón-Ángel A., Orús R., Amarillo M., Hernández-Pajares M., Juan J. M., Sanz J.:* Preliminary NeQuick assessment for future single frequency users of GALILEO, In Proceedings of thr "6th Geomatic Week", Barcelona, Spain February 8-11, 2005.

[2] *ESA (European Space Agency):* EGNOS – The European geostationary navigation overlay system: a cornerstone of Galileo. ESA Publ. Div. (SP-1303), Noordwijk, 2006.

[3] *Klobuchar, J., Ionospheric time- delay algorithm for single- frequency GPS- users, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-23, No.3, 1987.*

[4] *Opitz M. and Weber R.:* Development of a SISNeT-Simulation Software – "SISSIM", In Proceedings of the "Workshop on EGNOS performance and applications", p.173 – p.178, Gdynia, Poland October 27-28, 2005.

[5] *Opitz, M.; Weber, R.:* Development od a SISNeT-Simulation Software – "SISSIM". European Journal of Navigation, Volume 4, Number 2, May 2006.Page 44 – 48.

[6] *Radio Technical Comission for Aeronautics – RTCA (1999):* Minimal Operational Performance Standards for GPS/WAAS Airborne Equipment. Doc. No. Do 229 B.

[7] *Radio Technical Commission for Maritime Services – RTCM (2001):* RTCM recommended standards for differential GNSS (Global Navigation Satellite Systems). Service Version 2.3.

[8] *Seeber, G. (2003):* Satellite Geodesy, 2nd Edition. De Gruyter Verlag, Berlin.

[9] *Toran-Marti F. and Ventura-Traveset J.:* SISNeT User Interface Document, ESA Technical Document, Issue 2, Revision 1, Ref. E-RD-SYS-E31-010. Available through the ESA ESTB Web Site (http://www.esa.int/estb).

[10] *Toran-Marti F. and Ventura-Traveset J.:* The ESA SISNeT Project: Current Status and Future Plans, The European Navigation Conference GNSS, 2004.

[11] *Winkler, W. (2006):* Validierung des EGNOS- Bahn- und Ionosphärenmodells. Diplomarbeit, Institut für Geodäsie und Geophysik, FG Höhere Geodäsie, Technische Universität Wien.

Web-links (March, 2007)

[1] *EGNOS for professionals:*
http://www.egnos-pro.esa.int/index.html

[2] *EGNOS, ESSP:* http://www.essp.be/

[3] *EGNOS information:*
http://www.environmental-studies.de/Teilflachenbewirtschaftung/EGNOS_WAAS/3.html

[4] *RTCA:* http://www.rtca.org/

[5] *RTCM:* http://www.rtcm.org/

[6] *WAAS:* http://gps.faa.gov/Programs/WAAS/waas.htm

[7] *WEP:* http://wep.wienstrom.at/

Contact

Dipl.-Ing. Michael Opitz, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: mopitz@mars.hg.tuwien.ac.at

Dr. Robert Weber, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: rweber@mars.hg.tuwien.ac.at

Dipl.-Ing. Wolfgang Winkler, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.

AIS transponder with integrated EGNOS/Galileo receiver and related maritime standardization and certification aspects

Bernhard Hofmann-Wellenhof, Andreas Lesch, Philipp Berglez, Klaus Aichhorn, Jürgen Seybold and Franz Zangerl

Abstract

Within the Austrian national research project EMAG (Feasibility Study for an Experimental Platform for Multi-modal Applications of Galileo), two innovative developments are carried out: (1) the architectural design of a software based Galileo receiver tailored to the requirements of maritime navigation applications, and (2) the development of an AIS (Automatic Identification System) transponder with integrated GPS/EGNOS receiver being capable of providing visual integrity information to the user.

This paper provides an overview of the EMAG project results by describing the benefits of EGNOS and Galileo for maritime navigation applications in detail. First test results regarding functionality and especially accuracy, availability, and integrity of the GPS/EGNOS receiver integrated into the AIS transponder are presented. Available outputs of the standardization process are also mentioned.

Kurzfassung

Innerhalb des nationalen österreichischen Forschungsprojekts EMAG (Feasibility Study for an Experimental Platform for Multi-modal Applications of Galileo) wurden zwei innovative Entwicklungen durchgeführt: (1) das Design eines Software basierten Galileo Empfängers, der auf die Anforderungen maritimer Navigationsanwendungen maßgeschneidert wurde und (2) die Entwicklung eines AIS (Automatic Identification System) Transponders mit integriertem GPS/EGNOS Empfänger, der visuelle Integritätsinformationen für den Nutzer ausgibt.

Dieses Dokument bietet einen Einblick in die EMAG Projektergebnisse, indem die Vorteile von EGNOS und Galileo für maritime Navigationsanwendungen detailliert beschrieben werden. Erste Testergebnisse bezüglich der Funktionalität und speziell der Genauigkeit, Verfügbarkeit und Integrität des in den AIS Transponder integrierten GPS/EGNOS Empfängers werden vorgestellt. Verfügbare Ergebnisse des Standardisierungsprozesses werden ebenfalls behandelt.

1. Introduction

1.1. General

Worldwide open sea and inland waterways are the most widely used method for the transport of goods. A wide variety of vessels move around the world each day. The traffic of goods in, from, and to European ports reaches 40 million containers per year.

The efficiency, safety, and optimization of marine transportation are key issues. Global navigation satellite systems (GNSS) like GPS or the future Galileo [5] [8] [11] are becoming a fundamental tool yielding innovation and progress to this sector. Many other marine activities such as fishing, oceanography, or oil and gas exploitation will also benefit from the availability of Galileo services [1] [2] [3].

Increased accuracy and integrity, certified services, and a higher signal availability introduced by Galileo will be used by leisure boats, commercial vessels and all ships falling under the safety-of-life-at-sea (SOLAS) convention in every phase of marine navigation, i.e., ocean, coastal, port approach, and harbour manoeuvres, and under all weather conditions.

For marine navigation, regulated by the International Maritime Organization (IMO), Galileo will be an additional means of implementing the regulations on automatic identification systems (AIS) and vessel traffic management systems to increase the navigation safety, the collision prevention, and the economic benefit [12] [13].

Many maritime commercial activities are going to use satellite navigation. When fishing, it helps locating traps and nets. Fleet management, cargo monitoring, delivery and loading schedules are optimized. Even the locating of shipping containers can be facilitated, and satellite navigation could be used for automatic piloting or tracing of barges. Within ports, a system for information

services tailored to each ship's location is being considered.

Handling of containers is crucial for efficient commercial harbour operations.

Sciro in Italy is involved in developing a system using GNSS that automatically tracks containers inside a terminal. This will reduce container handling time and costs, and increase service levels and terminal productivity.

For inland waterways, accuracy and integrity of navigation data are essential to automate precise manoeuvres in narrow rivers and canals. Better navigation can be a major aid to the increasing capacity in inland waterway networks, which in turn contributes to the modal shift targets EC transport policy [4] [9] [12].

1.2. EMAG project

The content of the EMAG project included technical and economical basic research and analysis regarding the feasibility to develop an experimental platform based on a software receiver for multi-modal maritime applications of Galileo. Thus, the aim was to develop a detailed concept for a future realization [6] [7]. In a next step, the proper development of the experimental platform shall be carried out.

Additionally, a functional demonstrator based on the integration of AIS, GPS, and EGNOS was developed. Therefore, a combined GPS/EGNOS receiver was used and integrated into an AIS transponder of ACR Electronics Europe GmbH (formerly Nauticast Navigationssysteme GmbH). The combined position solutions (GPS+EGNOS) and integrity messages are fed into the transponder in real time. Thus, the transponder is capable to broadcast the position of the ship including a quality indicator (integrity flag) for each position solution to other ships and to the control facilities on shore.

2. Maritime Galileo Receiver Development

The implementation of the so-called experimental receiver platform for maritime applications is driven by the need for a receiver, which provides flexibility and openness to novel algorithms and receiver concepts. A software receiver has been found to be the most adequate solution for the above requirements.

2.1. Architectural Hardware Design

A high-level block diagram of the proposed receiver platform hardware is shown in Figure 1.

Three major parts can be identified. The first part comprises all the analog hardware required to preprocess and down-convert the signal received from a GNSS antenna to baseband or to an intermediate frequency. This hardware unit is referred to as analog front end, or radio frequency (RF) front end, respectively.

The second hardware unit comprises the analogue-to-digital converter, reconfigurable digital preprocessing hardware and intermediate memory, which allows for the storage of data at high rates being too high to be transferred to a single processor via a standard bus.

The third module establishes the processing platform for the receiver software. An essential part is the software running on the processing platform.

The chosen architectural concept allows a high degree of flexibility. This is mainly provided by the reconfigurable preprocessing hardware and the receiver software. The proposed platform can be configured to perform more or less the complete digital signal processing by the software. However, the wideband Galileo signals cannot be processed in real time with this approach. For real-time processing, the digital preprocessing hardware can be utilised. On the other hand, for the development of novel concepts and algorithms, a software-based approach is preferred and seems to be the logic consequence. Therefore, this approach has been chosen as driving requirement for the architectural design of the maritime Galileo receiver platform.

2.2. Architectural Firmware Design

Figure 2 shows a high-level context diagram of general GNSS receiver software. The receiver software comprises the following elements:

- The low-level software handling the external interfaces and converting the received data into a defined format, which allows a standardized access to this data.

- The control software establishes the interface to the man-machine interface and allows for the configuration and control of the algorithm software.

- The algorithm software accomplishes the signal processing tasks required to provide the GNSS receiver functionality.

- The algorithm software can be based on a set of predeveloped functions collected in the function library or on dedicated functions. The function library is open to be augmented by
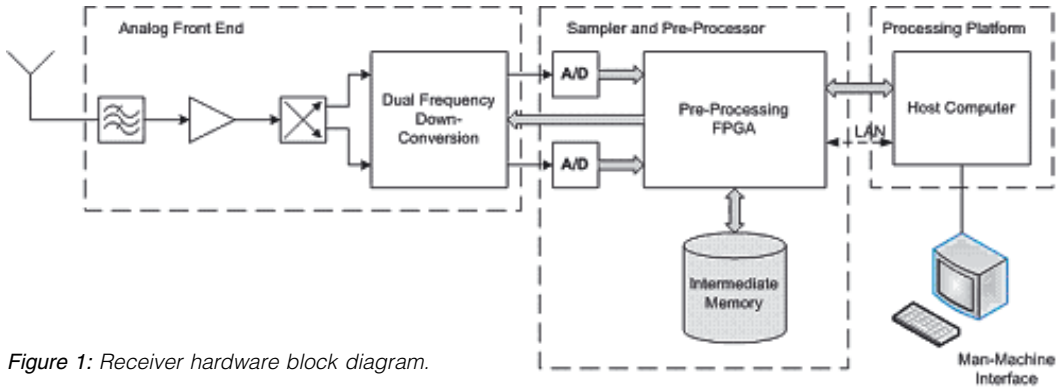
Figure 1: Receiver hardware block diagram.

custom functions. The default functions of the library shall be configurable in order to allow a high degree of flexibility. One specific part of the receiver algorithms is established by the position, velocity, and time (PVT) determination functions.

■ For applications without stringent deadline constraints, a conventional operating system is sufficient. For such applications the interactions of the algorithm software with the operating system are more or less transparent for the software developer. In case a real-time operating system is required, the algorithm software has to use services provided by this operating system. A real-time operating system is recommended in case stringent deadlines have to be met.
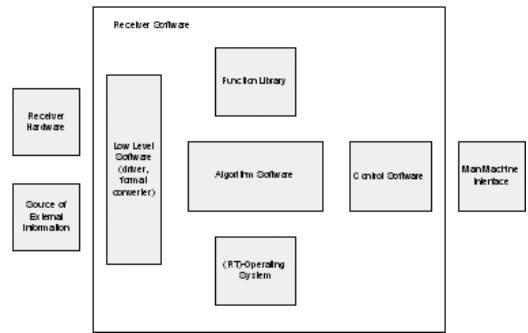
The receiver firmware covers the low-level software, the control software, and the function library. As the Galileo receiver platform shall establish a platform for the development of maritime Galileo receiver concepts and algorithms, the receiver firmware does not implement a specific functionality, but has to provide the flexibility to define the functionality for the desired application. Thus, the receiver platform has to provide an application programming interface (API). An API defines the access to a set of functions, procedures, variables, and data structures using a library that has been written for that purpose. A programmer implementing a specific application puts the blocks together and can write applications consistent with the operating environment in a straightforward way.

Novel concepts or algorithms require additional functionality, which is not available in the default library. Therefore, the Galileo receiver

platform has to allow for augmentation and modification of the library.

Figure 2: Receiver context software diagram.

2.3. Architectural PVT Software Design

The software design for the processing of Galileo measurements and the integration of GPS/EGNOS and Galileo is based on currently available modules used in TeleConsult Austria's IntNav software. The PVT software is designed for the processing of Galileo raw data and the integration of the raw data gathered from Galileo, GPS, and EGNOS applying different algorithms for PVT computation, including quality data as well. Thus, the following main functionalities of the PVT software are implemented:

■ Processing of navigation measurements in real time and in post-processing mode (only for developmental purposes).

■ Processing of (currently simulated) Galileo raw measurements.

■ Processing the combination of Galileo/GPS/EGNOS measurements.

Figure 3: EMAG functional demonstrator overview (left: u-blox GPS/EGNOS receiver module; right: NAUTICAST AIS transponder).

- Output of quality information for the processed position solution.
- Provision of position, velocity, time, and the quality data (integrity) for the application software.

3. EMAG Functional Demonstrator

The functional demonstrator developed and tested within EMAG consists of the AIS transponder NAUTICAST from ACR Electronics Europe GmbH using an integrated low-cost GPS/EGNOS receiver board from u-blox AG as position fixing device (Figure 3).

The transponder communicates with the receiver module via an internal standard serial interface (RS232). Thus, on the one hand the receiver can be configured by the transponder and on the other hand the receiver provides position information and integrity data to the transponder. The transponder is now capable of providing visual warnings to the user if so-called out-of-tolerance conditions appear (the calculated protection level exceeds the defined alarm limit as specified for inland waterway navigation or maritime navigation). Further, the transponder transmits the position of the own ship including the quality information to other ships and to the control facilities on shore.

The test region was selected alongside the river Danube from the lock "Freudenau" upstream to the bridge "Reichsbrücke" (green trajectory). The test center was located in the office of ACR Electronics Europe GmbH at Handelskai 388 (red circle in Figure 4).

Figure 4: Test region along the river Danube in Vienna.

3.1. Operational Tests

The reference installation of the standard inland AIS transponder "NAUTICAST Inland AIS" is used to demonstrate the behavior and interoperability between the developed demonstrator and the live AIS acting along the river Danube. Irrespectively of the type of AIS, currently present traffic should be visible beside the developed demonstrator.

One installation is a standard inland AIS transponder "NAUTICAST Inland AIS" together with a connected inland ECDIS solution in the test center (ACR office). On the screens of this ECDIS solution (Figure 6), the EMAG demonstrator should be visible for tracking and tracing. It is evaluated, whether the pick report could be used for reading out the data transmitted by the demonstrator. Nevertheless, the continuous visibility on the inland ECDIS is the target of this part.

The Attingimus AIS test device is a worldwide approved device for measuring the date flow in and out of an AIS transponder. Most of the developed features are seen in particular on that device because of it's capability to decode the binary data stream of the VDL (VHF Data Link).
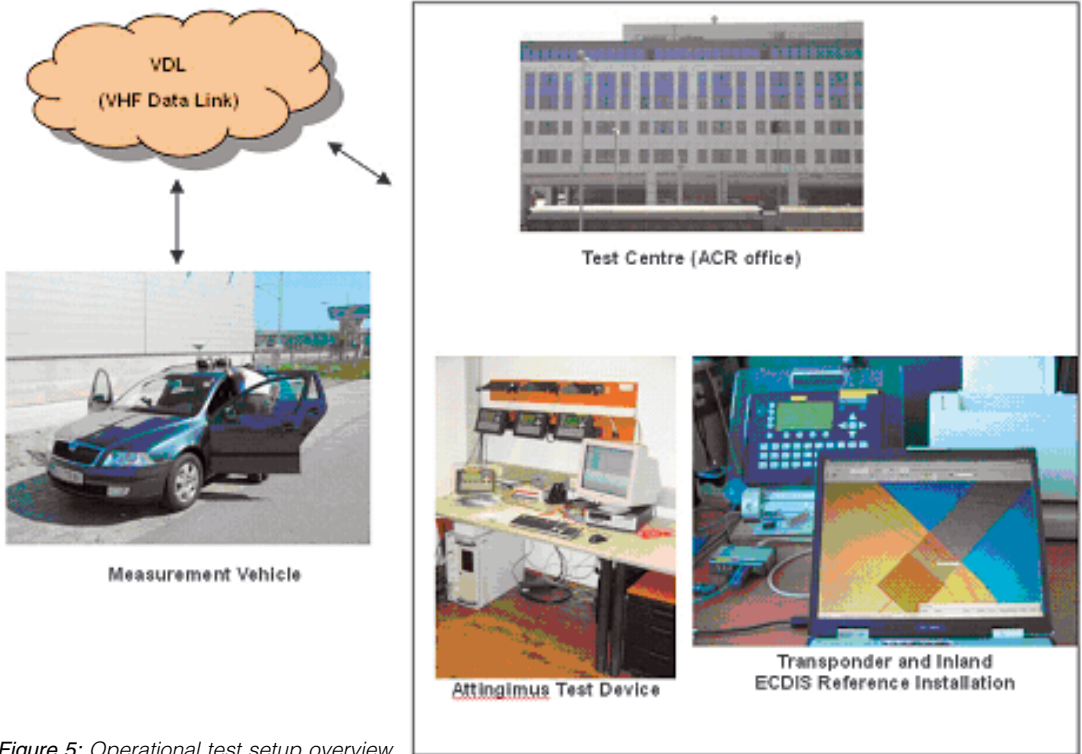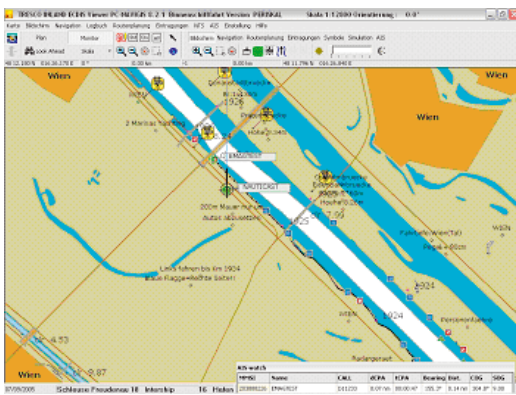
Figure 5: Operational test setup overview.

Figure 6: Visualisation on the presentation interface during the test measurements.

The demonstration with the above setup (Figure 5) showed an operational sample of a future implementation and its requirements. Based on these results (and the essential functional requirements of future EGNOS/Galileo input and output receiver performance characteristics) the inland AIS transponder families (as well as the high seas) electronic positioning fixing device performance could be substantially improved.

3.2. Navigation Performance Tests

The overall test system consists of a GPS reference station (L1/L2 GPS receiver) and the mobile equipment used in a demonstration vehicle. The mobile equipment consists of the AIS transponder with integrated u-blox GPS/EGNOS receiver module – the EMAG functional demonstrator, another GPS receiver (L1/L2 GPS receiver) including antenna (for comparison purposes), an antenna splitter, a notebook for data recording, the power supply, and the necessary cablings. The demo tests were carried out with a car, simulating the trajectory of a vessel, on the Treppelweg.

The analyses and interpretation of the navigation performance are split into three parts. The first part deals with the positioning performance of the GPS/EGNOS receiver compared to a reference trajectory. In the second part, an availability analysis of the receiver with special focus on the EGNOS signal is carried out. In the last part, the computed protection level (integrity information) is verified.

For a significant analysis, the GPS/EGNOS trajectory output by the transponder-internal u-

blox receiver is compared with the computed trajectory of the second receiver. The reference trajectory is measured with a dual-frequency geodetic GPS receiver (Ashtech Z-Xtreme) and is computed by kinematic baseline determination in post-processing mode [10].

In Figure 7, the differences (Y = longitude, X = latitude in meters) between the computed reference trajectory and the trajectory of the transponder-internal GPS/EGNOS receiver are visualised.
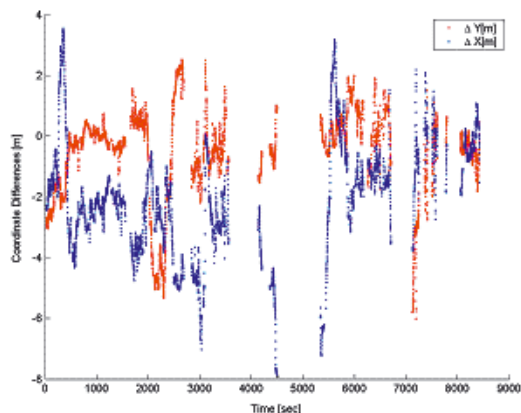

Figure 7: Absolute positioning accuracy of the functional demonstrator.

In a next step, the availability of the EGNOS signal was validated. The percentage of the EGNOS signal availability for the transponder internal receiver amounts 93.1%. Due to various high buildings along the south-west shore of the river Danube, this is a very positive result.

The protection level calculation is presented in Table 1. It indicates the amount of generated visible user warnings in situations, where the calculated protection level exceeds the defined alarm limit as specified by IMO regulations for inland waterway navigation.

The demo tests showed clearly that the low-cost u-blox receiver provides sufficient absolute accuracy (usually better than ±5 m) for the use in inland waterway navigation during normal conditions (no extensive satellite signal shadings). During conditions with various satellite signal shadings occurring (e.g., under bridges,

etc.), the absolute accuracy is worse than ±5 m, e.g., due to bad satellite geometry. Furthermore, the EGNOS signal is sometimes shaded because of high buildings, vegetation, and bridges. It can be assumed that the EGNOS signal availability is much better directly on the river Danube. Nevertheless, due to the low elevation of the EGNOS satellites, signal shadings will occur sometimes.

Suggestions to improve the situation significantly are alternatives to the direct EGNOS signal reception:

- a terrestrial network transmitting the EGNOS corrections (e.g., as described in the ESA project GALEWAT – Galileo and EGNOS for Waterway Transport);
- a direct internet communication link to the EGNOS data messages provided by ESA (SiSNet or, in future, EDS – EGNOS Data Server).

It is worth to mention that the suggestions only include infrastructure improvements to provide better EGNOS signal availability. The tests showed clearly that the u-blox receiver provides sufficient absolute accuracy and EGNOS signal availability during normal conditions.

4. General EGNOS and Galileo Differentiators and Benefits

Satellite navigation applications are currently based on GPS performances, and great technological effort is spent to integrate satellite-derived information with a number of other techniques in order to achieve better positioning accuracy with improved reliability. This scenario will significantly change in the very near future. Still in 2007, EGNOS, the European regional augmentation of GPS, starts to fully provide its services. Later, the global satellite navigation infrastructure will dramatically increase with the advent of Galileo, with full operational capability (FOC) probably around 2010/11. The availability of two or more GNSS (also a renaissance of the Russian GLONASS is to be expected soon), which will significantly increase the total number of available satellites, will enhance the quality of the services - and the number of potential users and applications will rise.

	Transponder internal GPS/EGNOS receiver
Nr. of trajectory points recorded	8774
Nr. of alerts generated	1355
Percentage of exceeded alarm limit (visible user warning generated on the display of the AIS transponder)	15,4%

Table 1: Protection level summary.

Galileo's specific characteristics will bring significant enhancements. Primarily for urban areas, indoor applications, and inland waterway navigation, the design of Galileo signals will improve the availability of the service (broadcast of data-less ranging channels, in addition to the classical pseudorandom ranging codes). Additionally, the high-end professional market will also benefit from Galileo's signal characteristics [9]. Three-carrier phase measurements will be essential for the development of specific three-carrier ambiguity resolution (TCAR) algorithms, leading to centimeter accuracy over large regions.

Simulations demonstrate that availability of positioning services in urban areas or valleys (like the Danube valley in Austria) with "canyons" (satellite visibility obstruction by topography, vegetations, high buildings and skyscrapers) is increased significantly using a combination of Galileo and GPS.

Galileo is and will remain under civil control and has been designed under civil requirements. It is operated in a transparent way, allowing for full service certification.

The need for service guarantee for safety-of-life and commercial applications has been taken into account in the design of the systems. Legal implications of service level commitment are driving the Galileo system implementation. Accountability on requirements regarding service provision has led to clear traceability requirements on detailed system performance history (already implemented in EGNOS). This approach significantly improves the safety of the navigation system, facilitates the detection and investigation of any malfunctioning and allows recording of service-level performance in case of claims. In this context, the integrity information plays an important role.

Law enforcement in the road and maritime domains, road charging and tolling, safety-of-life navigation in all modes of transport will soon rely on dedicated infrastructures with reliability and guarantee characteristics that are simply not available with current systems.

5. Conclusion

The conclusion is split into the development of the experimental maritime Galileo receiver platform and the EMAG functional demonstrator (AIS transponder with integrated GPS/EGNOS receiver board).

5.1. Experimental Galileo Receiver Platform

Within the EMAG project, the first step for the development of an experimental maritime Galileo receiver platform has been carried out. The structures of the hardware, firmware, and software for the receiver are defined and developed. In a planned follow-on project, further development steps (prototyping of hardware components and programming of the software) shall be carried out.

The implementation of an experimental receiver platform is driven by the need for such a receiver, which provides flexibility and openness to novel algorithms and receiver concepts. A software receiver has been found as most adequate solution for these requirements.

5.2. Demonstration System

The operational and navigation performance tests pointed out clearly that the navigation performance of the AIS transponder could be improved significantly by the integration of EGNOS into the system concept. By applying the EGNOS corrections to the GPS position solution, an absolute positioning accuracy of better than ±5 m is achieved during normal operation conditions. This fact is very essential for the compliance of the requirement regarding an absolute positioning accuracy of better than ±5 m as stated in the inland ECDIS standard for operations in the so-called "navigation mode". Further, the user gets visual information via the AIS transponder display if the system provides "safe navigation conditions" (the continuously computed protection level is better than the defined horizontal alarm limit of 25 m as specified for inland waterway navigation.

The main outcome of the demonstration tests points out clearly that the AIS transponder with integrated GPS/EGNOS receiver provides improved positioning and navigation performance and can be extended by the integration of Galileo to further improve the system performance significantly.

Acronyms and Abbreviations

AIS	Automatic Identification System
API	Application Programming Interface
DLR	Deutsches Zentrum für Luft- und Raumfahrt
EC	European Commission
ECDIS	Electronic Chart Display and Information System
EDS	EGNOS data server
EGNOS	European Geostationary Navigation Overlay Service
EMAG	Feasibility Study for an Experimental Platform for Multi-modal Applications of Galileo

ESA   European Space Agency
FOC   Full Operational Capability
GALEWAT   Galileo and EGNOS for Waterway Transport
GATE   Galileo Test Environment
GNSS   Global Navigation Satellite System
GPS   Global Positioning System (U.S.)
IMO   International Maritime Organisation
PVT   Position, Velocity, Time
RF   Radio Frequency
SOLAS   Safety Of Life At Sea
TCAR   Three-Carrier Ambiguity Resolution
VDL   VHF Data Link
VHF   Very High Frequency

References

[1] *European Commission 2000:* Cost benefit analysis results for Galileo. Commission staff working paper.

[2] *European Commission 2000:* Galileo – a new GNSS designed with and for the benefit of all kind of civil users. European Commission DG Energy and Transport. ION GPS 2000.

[3] *European Commission (2003a):* The Galilei Project – Galileo Design Consolidation. Results of the Galilei project funded under the European Union's 5th framework programme.

[4] *European Commission (2003b):* The European Dependence on US GPS and the Galileo Initiative.

[5] *European Commission and European Space Agency (2003):* Galileo Definition Studies Summary. Global Satellite Navigation Services for Europe.

[6] *EMAG – Executive Summary, Version 1.0, dated 21.12.2006; prepared by TeleConsult Austria GmbH, supported by ACR Electronics Europe GmbH, Austrian Aerospace GmbH, Graz University of Technology – Institute of Navigation and Satellite Geodesy.*

[7] *EMAG – Final Report, Version 2.0, dated 21.12.2006; prepared by TeleConsult Austria GmbH, supported by ACR Electronics Europe GmbH, Austrian Aerospace GmbH, Graz University of Technology – Institute of Navigation and Satellite Geodesy.*

[8] *European Space Agency (2003):* Galileo – The European Programme for Global Navigation Services. BR-186 (2nd impression), March 2003.

[9] *Galileo Joint Undertaking (2003):* Business in satellite navigation – An overview of market developments and emerging applications.

[10] *Hofmann-Wellenhof B., Lichtenegger H., Collins J. (2001):* GPS – Theory and practice; Springer, Wien New York, 5th edition;

[11] *European Commission (EC) and European Space Agency (ESA):* Galileo mission high level definition, September 23, 2002, 3rd issue.

[12] *IEC61108-1 Edition 2.0:* Maritime navigation and radiocommunication equipment and system – Global navigation satellite systems (GNSS) - Part 1: Global positioning system (GPS) - Receiver equipment - Performance requirements, methods of testing and required test results; Copyright * 2003 International Electrotechnical Commission.

[13] *Revised Maritime Policy and Requirements for a Future Global Navigation Satellite System (GNSS), International Maritime Organization (IMO), 22/Res. 915, November 29, 2001.*

Contact

Bernhard Hofmann-Wellenhof: Graz University of Technology, Institute of Navigation and Satellite Geodesy, Steyrergasse 30, A-8010 Graz, Austria.
E-Mail: hofmann-wellenhof@tugraz.at,
Tel: +43-316-873-6830, Fax: +43-316-873-8888.

Andreas Lesch: ACR Electronics Europe GmbH, Handelskai 388 / Top 632, A-1020 Vienna, Austria.
E-Mail: andreas.lesch@acr-europe.com.

Philipp Berglez: Graz University of Technology, Institute of Navigation and Satellite Geodesy, Steyrergasse 30, A-8010 Graz, Austria. E-Mail: pberglez@tugraz.at.

Klaus Aichhorn: TeleConsult Austria GmbH, Schwarzbauerweg 3, A-8043 Graz, Austria.
E-Mail: kaichhorn@teleconsult-austria.at.
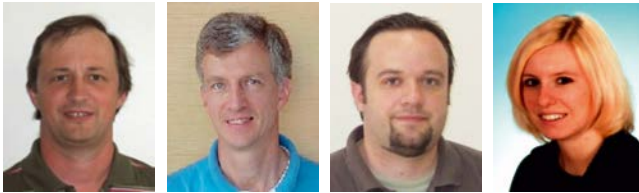
Jürgen Seybold: TeleConsult Austria GmbH, Schwarzbauerweg 3, A-8043 Graz, Austria.
E-Mail: jseybold@teleconsult-austria.at.

Franz Zangerl: Austrian Aerospace GmbH, Stachegasse 16, A-1120 Vienna, Austria. E-Mail: franz.zangerl@space.at.

A Navigation Concept for Visually Impaired Pedestrians in an Urban Environment

Manfred Wieser, Bernhard Hofmann-Wellenhof,
Bernhard Mayerhofer and Bettina Pressl

Abstract

An overall concept of navigation is adapted to the special needs of blind and visually impaired pedestrians and is being developed in national R&D-projects. The development aims at prototypes of navigation systems which cover the total spectrum of navigational components. This comprises an appropriate modeling of the navigational environment, fast routing algorithms generating lists of maneuvers, suitable positioning tools based on sensor fusion and Kalman filtering, reliable map matching algorithms for route checking, and, finally, efficient guidance instructions communicated via an adequate man-machine interface.

Kurzfassung

In zwei F&E-Projekten des Instituts für Navigation und Satellitengeodäsie wird ein allumfassendes Konzept eines Navigationssystems entwickelt, welches den speziellen Bedürfnissen blinder und sehbehinderter Fußgänger angepasst wird. Das Vorhaben zielt jeweils auf einen Prototyp ab, welcher das gesamte Spektrum von Navigationskomponenten abdeckt. Dazu gehören eine entsprechende Modellierung des Umfeldes, in dem navigiert wird, schnelle Routingalgorithmen zur Generierung von Manöverlisten, geeignete Positionierungsmethoden basierend auf Sensorfusion und Kalmanfilterung, zuverlässige Map-Matching-Algorithmen zur Routenverfolgung und schließlich eine effiziente Zielführung mittels eines dem Zweck entsprechenden Man-Machine-Interfaces.

1. Introduction

1.1. The projects

The research projects at the Institute of Navigation and Satellite Geodesy (INAS) of the Graz University of Technology dealing with navigation systems for blind people are called PONTES ("positioning and navigation of visually impaired pedestrians in an urban environment") and ODILIA ("mobility of blind people through user-oriented information and satellite-based navigation"). Both the term PONTES (Latin word for bridges) and the term ODILIA (patron of the blind) underline the benefit of technological development for physically disabled people. Financial support is granted by the Austrian Federal Ministry of Transportation, Innovation, and Technology. The ministry is represented by the Austrian Research Association (FFG) and by its aerospace agency, respectively. The R&D-program in which PONTES and ODILIA are embedded is the Austrian Radionavigation Technology and Integrated Satnav Services and Products Testbed (ARTIST). PONTES has started in May 2005 and ODILIA in September 2006 with a duration of sixteen months per project.

From the very beginning, potential users of a future navigation system for the blind people are involved in the projects. There is a close cooperation between INAS and the Styrian Association of the Blind and Visually Impaired People.

Further cooperations are initiated with Vectronix AG (Heerbrugg) which distributes pedestrian navigation modules and the company AFN (Vienna) which develops user interfaces for the blind and visually impaired.

1.2. The intention

The first essential aspect is to find an application for a pedestrian navigation system with social relevance. The second issue is to develop an overall outdoor-navigation concept, i.e., covering the whole spectrum of navigation from positioning to guidance.

The specific user group requires a tailored system architecture and a user-oriented development. Therefore, blind and visually disabled people are defining their needs and requirements with respect to an appropriate navigation tool and its components, are contributing during the

conceptual phase, and are then testing the prototype system.

The underlying issues of the projects mainly are:

- a two-phase development of demonstrator prototypes where the first phase is covered by PONTES and represents the system's main functionality and where the second phase is performed by ODILIA and includes innovative aspects and sophisticated add-ons;

- the use of existing hardware components, especially for positioning and the man-machine interface;

- a proprietary software development for routing, map matching, and guidance;

- a technical feasibility and acceptance study rather than commercially oriented exploitation and dissemination;

- supplementing common aids for the blind people, e.g., the blindman's stick, and medium-term replacing the guide dog or a coach person (visually enabled).

2. The Concept

The system is self-contained, apart from received positioning signals (e.g., from GPS), and is therefore autonomous, i.e., independent of any control center. The system's architecture is similar to the concept of an in-vehicle navigation system but is tailored to the special needs of the visually impaired pedestrian [1]. The components shown in Fig. 1 appear in three (colored) pairs:

- The navigable map describes the geographic space in terms of the path network and the navigational environment and is the primary source for route planning which computes an optimal path from a starting point to a given destination.

- The positioning module has to forward the position of the pedestrian to the map-matching module which transforms the absolute position to a location relative to a close-by object or a decision point.

- Route guidance directly profits from route planning as well as from map matching. The guidance instructions are mainly transmitted acoustically via the man-machine interface at the correct time.
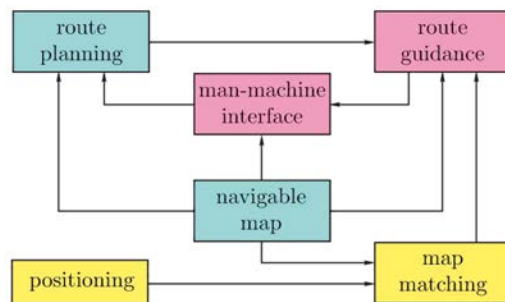

Figure 1: Components of a pedestrian navigation system.

In the following, the components are described in more detail except of the man-machine interface. Although it is of great importance in the case of a blind user and puts big requirements on the information exchange by voice between the navigation device and the person using it, it is not a typical navigation component and is skipped for that reason.

3. The System's Components
3.1. Navigable Map

A navigable map tailored to the needs of a visually impaired person is the digital answer to tactile maps which are normally used by blind people to get acquainted with the surrounding in which they move. In order to become navigable, a digital map must be able to support routing in the sense of route planning, positioning by map-based "artificial" observations (map aiding), map matching as the projection of a position onto the map, and finally guidance on the basis of a maneuver list.

A careful modeling of the urban environment plays a key role for the navigation system's reliability. A visually impaired person must count on the guidance instructions generated by the system, since the person is unable to compensate any misinformation by visual faculty. As a consequence, the navigable map must be geometrically accurate, topologically consistent, and thematically correct, up-to-date, and complete on a topmost level. Furthermore, the map and the corresponding database have to fulfill efficient data storage and modeling rules of national standards or, preferably, those of the Comité Européen de Normalisation (CEN) and the International Organization for Standardization (ISO), respectively.

The modeling of the navigational environment and the generation of the navigable map are achieved by a geographical information system

(GIS) and are based, e.g., on a digital map for an urban area (Fig. 2). The first step is to create an application layer for the special navigation purpose in case of a visually impaired person. This is done by eliminating unnecessary features and attributes (e.g., road markings) and adding those of specific importance (e.g., "acoustic" traffic lights). The second step is to establish a path network for the blind people and the respective database containing all the geometric, topological and thematic information [2].

Figure 2: Extract of the digital city map of Graz.

The data structure of the vector-type path network consists of nodes (e.g., intersection of pavements) and edges (e.g., promenade, zebra crossing) relevant for route planning and of polygon points relevant for guidance. The latter represent obstacles, pre-defined guidance instructions, and points of interest (POIs). Fig. 3 shows an example of a path network (area extracted from Fig. 2) where nodes and edges are colored in orange and polygon points are colored in blue.

Figure 3: Example of a path network for the blind.

3.2. Routing

Planning the optimal route from a start site to a destination site equals finding the shortest path between the two corresponding nodes in the path network. Note that the terms "optimal" and "shortest" represent the safest route in case of the visually impaired pedestrian. In addition, the result of the routing procedure defines a line-based (i.e., one-dimensional) motion. This strongly facilitates subsequent computations, e.g., map matching. In comparison, the trajectory of a sighted pedestrian in general has a degree of freedom which equals two. Mathematically speaking, the node-edge structure of a path network is represented by a valuated graph. A cost function $c(e(i,j))$ assigns a real valuation number $c(i,j)$ to an edge $e(i,j)$ from node $v(i)$ to node $v(j)$. The cost number $c(i,j)$ is influenced by the geometric length of the edge and by safety aspects along the edge, e.g., occurrence of obstacles or dangerousness of road crossings. The shortest-path algorithm then tries to find the path with the minimal sum of cost numbers along the path. Because of the characteristic behavior of the valuation function it may of course happen that a path from $v(i)$ over $v(j)$ to $v(k)$ is shorter in terms of the cost numbers than the direct path from $v(i)$ to $v(k)$.

One of the basic ideas of shortest-path algorithms is to generate a search tree which branches out radially with the start node as a root and grows iteratively till the destination node is reached. The result provides all shortest paths from the start node to all nodes within the tree (Fig. 4).
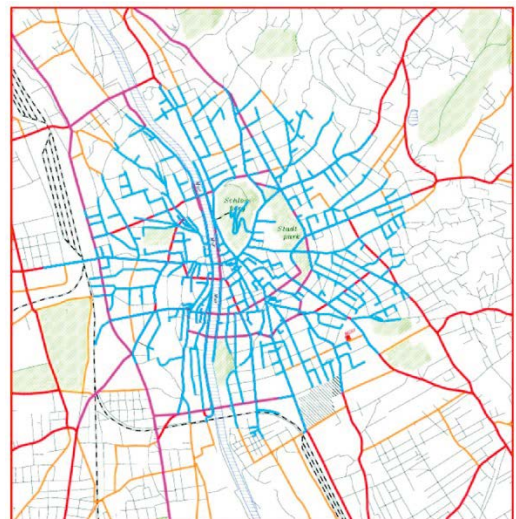

Figure 4: Search tree.

One of the most frequently tree algorithms used is the one by Dijkstra. In any case, the path network has to be efficiently stored by an indexed adjacency list, so that a direct access to all neighbors of a node can be guaranteed. This is needed to keep the algorithm's computation time as low as possible. Another strategy to achieve this is to apply heuristic concepts. Contrary to Dijkstra's algorithm, the A*-algorithm (pronounced as "A star") extends the search tree predominantly towards the destination and much closer to the shortest-path result. A second heuristic approach is the task of spatial reasoning. The cognitive map as the mental representation of the environment and the way how visually impaired human beings reason about space helps to understand whether they use the same criteria as the network algorithms [3, Sect. 14.2].

3.3. Positioning

Compared with conventional pedestrian navigation, the quality requirements on position determination are much higher in the case of blind pedestrians. Accuracy should be at the one-meter level and below. The position must be accurate enough so that objects and locations relevant for guidance instructions are reliably situated within the tactile range of the blindman's stick. Of course, availability must be ensured anytime and anywhere with a top-level integrity.

To achieve the above quality requirements, an integrated concept of position determination is indispensable. On the hardware side this is realized by a multisensor system where the determination of the pedestrian's state vector (position and velocity, and to a certain extent even attitude) has to follow the principles of sensor fusion. Hence, appropriate methods of signal processing, i.e., methods of computing and updating (or filtering) the actual state vector are required. In kinematic mode, optimal filtering is achieved via Kalman filtering based on measurements from the multisensor equipment and on dynamic modeling of the motion.

Typically, the integrated concept for pedestrian navigation is the combination of satellite-based positioning, e.g., GPS using code pseudoranges, and dead reckoning (DR). In the sense of relative positioning, the DR vector is calculated from an oriented direction (course angle) and a range (covered distance). In case of a pedestrian, the range is derived from step detection.

Integrated positioning techniques profit from redundant information. The combination of GPS and DR is a typical example for a complementary redundancy. The sensors are based on different physical principles; and, to some extent, they complement each other: GPS, e.g., is affected by outages due to shadowing effects in urban areas, which can be overcome by DR. In contrast, DR may keep sufficient accuracy only over a short distance traveled.

Although INAS and its cooperative partners have been already involved in the research and development phase for a commercial prototype of pedestrian positioning devices [4], the decision in PONTES and ODILIA was "not to reinvent the wheel" and to use therefore existing pedestrian navigation modules, e.g., the PNM kindly offered by the Swiss Vectronix AG for the duration of the projects. The conceptual design and the technical implementation of the device should ensure that the PNM will fulfill its purpose [5].

The PNM in use has a size of $85 \times 135 \times 35$ mm, is mounted on a belt and should be carried at the back (Fig. 5). The GPS antenna can be fixed on a cap and the RS232 cable for bidirectional data exchange is connected to a PDA or subnotebook. According to the above multisensor concept, the PNM includes a GPS receiver for absolute single point positioning, a magnetometer triad and a gyrocompass for course determination, an accelerometer triad for step detection (derived from frequency analysis mainly in the vertical component), and a barometric altimeter for height determination. The associated firmware calculates the DR position and integrates it with the GPS position using filtering techniques.

Figure 5: PNM, Vectronix AG.

Future improvements of the positioning performance may be achieved for instance by:

- image-based positioning using a head set of two micro-cameras: in principle, the position is derived after feature extraction from the images and object matching in a city model;

- the integration of differential GPS: especially, if the accuracy of the map-matched GPS position is not sufficient (cf. Sect. 3.4);

- applying map aiding: suitable contents (e.g., direction or height information) of the navigable map is used as a kind of pseudoobservation;

- a centralized filter architecture: the sensors are integrated on the measurement level and the data are not preprocessed within the individual sensor.

3.4. Map matching

The process of map matching (MM) "projects" the pedestrian's trajectory onto an edge sequence in the digital map (Fig. 6). The current position is matched to a map point which is directly situated on an edge of the model graph. In other terms, the real position, absolutely defined in a coordinate reference frame, is transformed to a location relative to the nodes of an edge in the graph. This makes sense, since the guidance instructions which have to be communicated according to a current position are unexceptionally stuck to topological relations in the map: e.g., an obstacle warning is announced a few meters prior to the occurrence of the hindrance, certain maneuver instructions (e.g., "turn left") are announced immediately before the direction of the path changes, etc.
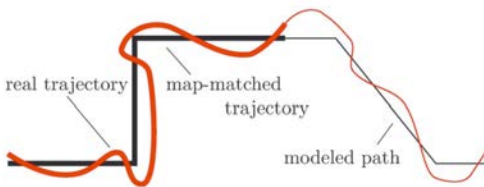

Figure 6: Principle of map matching.

MM directly follows Kalman filtering where the observed trajectory is updated. The resultant filtered trajectory is then blended with the navigable map in the subsequent MM step what finally leads to a map-matched trajectory (Fig. 7). Thus, MM also serves as a kind of filter-based updating. In this respect it must be mentioned that it is mathematically controversial to combine Kalman filtering and MM in a common central filter.
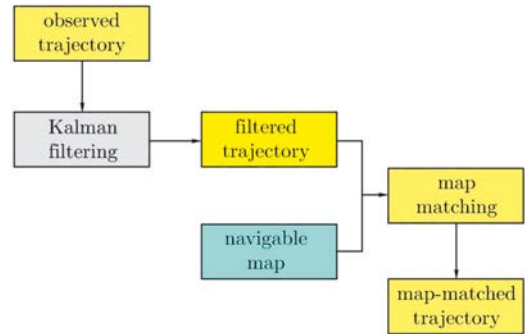

Figure 7: Process of trajectory computation.

MM calculations are complex and time-consuming, since sophisticated MM algorithms perform an edge-to-edge and not a point-to-point matching. Not only the current position is taken into account but also a (partial) "history" of the trajectory. Especially, in the neighborhood of nodes representing a junction with at least three roads, edge-to-edge matching helps to decide which of the branching roads has to be taken when leaving the node [6]. The two elementary edge-based MM strategies use either correlation tests or affine transformations. Both methods measure the similarity between the trajectory and its potential map representatives, and the final trajectory matching corresponds to the nature of a least-squares alignment. MM algorithms must also be able to warn of off-route or even off-road situations. In the first case, the pedestrian is still moving on a path within the network but not on the precomputed one, in the second case, he has totally left the path network [3, Sect. 14.3.1].

3.5. Guidance

As in vehicle navigation, the guidance concept is the one for en route guidance along a line-based trajectory following a precomputed route. As shown in Fig. 8, the main components of this concept are the generation of a maneuver list and the permanent route checking. Including warnings of obstacles and information on POIs, the maneuver list contains all turn-by-turn instructions which strictly rely on the prior map-based route computation. The route checking task is permanently monitoring the pedestrian's map-matched position relative to the points where guidance instructions are to be reported. As the maneuver list tells how to guide, route checking decides when to guide. The appropriate guidance instructions are then communicated acoustically via a head set to the visually impaired pedestrian [3, Sect. 14.3.2].
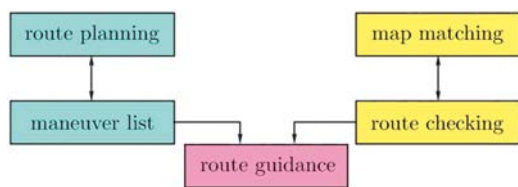
Figure 8: Concept of guidance.

The "vision" for the future is that the above guidance concept supplements usual tactile aids like blister paving and acoustic traffic lights (Fig. 9, upper left image). A further important issue is the capability of warnings with respect to permanent and temporary obstacles. In the permanent case, the information should be available in the database even if the hindrance can be detected by the blindman's stick (Fig. 9, lower right image). Warnings of a temporary obstacle, like the scaffold in Fig. 9 (upper right image), can be performed, if the database is updated continuously ("dynamic" database) or if the blind carries a camera head set. The digital data outcome of the camera is analyzed by digital image processing. Within the scope of ODILIA, further developments of the system may consider indoor guidance and the use of public means of transport like busses or trams. A desired system miniaturization including an appropriate design of the device should be enabled according to the principles of wearable computing and should be tailored to guidance purposes. Finally, effective on-trip guidance can be prepared by pre-trip training on a desktop computer under virtual reality conditions.

Figure 9: Tactile aids and obstacles.

4. Conclusions

The system concept introduced in PONTES and ODILIA should be understood as a navigation system tailored to the needs of visually impaired pedestrians in an urban environment. As the information distribution plays a dominant role in any location-based service [7], an efficient guidance to guarantee save mobility is aimed at in both projects. In the process, the blind pedestrian's location is gained by integrated positioning followed by map matching onto a highly accurate, digital map.

Nevertheless, some questions remain: Is the above approach the right one? Is the undertaking commercially feasible? Are there alternatives? E.g., especially in ODILIA, data acquisition and modeling of the path network are considered: the modeling of the navigational environment is an annoying and time-consuming task which can be accelerated by semi-automatic procedures. Or, instead of data acquisition done in advance, the potential user, guided by a sighted coach, moves along the path network and records his own data.

Whether data acquisition or one of the other components is considered, future developments will probably cast some light on the shadow.

References

[1] *Pressl B., Wieser M.:* A Computer-Based Navigation System Tailored to the Needs of Blind People. In Miesenberger K., Klaus J., Zagler W., Karshmer A. (Eds.): Computers Helping People with Special Needs, 10th International Conference ICCHP, Linz. Springer, Berlin Heidelberg, 2006.

[2] *Pressl B.:* Digitale Karte zur Zielführung in einem Navigationssystem für blinde Personen. Magisterarbeit, Technische Universität Graz, 2005.

[3] *Hofmann-Wellenhof B., Legat K., Wieser M.:* Navigation – principles of positioning and guidance. Springer, Wien, 2003.

[4] *Legat K.:* Pedestrian navigation. Dissertation, Graz University of Technology, (2002).

[5] *Ladetto Q., Merminod B.:* In step with INS – navigation for the blind, tracking emergency crews. GPS World, 13(10), 2002.

[6] *Mayerhofer B.:* Map-Matching in der Fußgängernavigation für blinde Personen. Magisterarbeit, Technische Universität Graz, 2005.

[7] *Wieser M., Mayerhofer B., Pressl B., Hofmann-Wellenhof B., Legat K.:* GIS-gestützte Navigation blinder und sehbehinderter Personen. In Strobl J., Blaschke T., Griesebner G. (Hrsg.): Angewandte Geoinformatik 2006, Beiträge zum 18. AGIT-Symposium Salzburg. Wichmann, Heidelberg, 2006.

Contact

Dr. Manfred Wieser, Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, A-8010 Graz. E-mail:wieser@geomatics.tu-graz.ac.at

Dr. Bernhard Hofmann-Wellenhof, Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, A-8010 Graz. E-mail: hofmann-wellenhof@tugraz.at

Dipl.-Ing. Bernhard Mayerhofer, Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, A-8010 Graz. E-mail: mayerhofer@tugraz.at

Dipl.-Ing. Bettina Pressl, Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, A-8010 Graz. E-mail: pressl@tugraz.at

vgi

Investigations of selected systems for Indoor and Pedestrian Navigation

Günther Retscher

Abstract

Within the last decade navigation systems have become popular. Vehicle drivers have started to trust in the information provided by car navigation systems and even pedestrians are gaining interest in reliable guiding instructions. Most of the available systems on the market, however, are limited to outdoor areas, whereas way-finding within buildings has mostly been neglected so far as location systems are rarely available indoors. In a conducted performance analysis of indoor location systems could be also seen that the user has to expect high costs in using some of these systems (e.g. using infrared and ultrasonic signals) as an installation of a large number of sensors is required in the building. One approach to reduce costs is the use of already available wireless infrastructure such as Wireless Local Area Networks (WLAN). Such a positioning system has been installed in an office building of the Vienna University of Technology and can be employed for location determination of user's which are equipped with a WLAN enabled mobile device. In addition, the use of Radio Frequency Identification (RFID) tags at selected known points, so-called active landmarks, is currently investigated. Then the user can be located using cell-based positioning if he is in the read range of such a tag. Currently the deployment of such a concept in our office building and in the surrounding outdoor environment is investigated. Furthermore a combination with WLAN and dead reckoning shall be performed in the near future to provide continuous position determination of a user.

Kurzfassung

Der Einsatz von Navigationssystemen hat vor allem im letzten Jahrzehnt sehr stark an Popularität gewonnen. Autofahrer vertrauen immer mehr auf die Führungsanweisungen von Autonavigationssystemen und auch Fußgänger nutzen moderne, mobile Systeme. Die meisten am Markt befindlichen Systeme können jedoch nur im Freien genutzt werden, da sie im wesentlichen auf satellitengestützter Positionierung beruhen und die Positionsbestimmung in komplexen Gebäuden vernachlässigt wird. Zur Positionierung in Gebäuden wurden in den letzten Jahren eigene Systeme entwickelt, die z.B. Infrarot, Ultraschall oder Radiosignale sowie digitale Kameras nutzen. Sie stehen jedoch zur Lokalisierung von Personen und Objekten in komplexen, öffentlichen Bürogebäuden und anderen wichtigen Einrichtungen (wie Einkaufszentren, Bahnhöfen, Flughäfen, usw.) kaum zur Verfügung und einige dieser Systeme erfüllen im Hinblick auf die hohen Kosten für die notwendige Installation von Sensoren im Gebäude kaum die Erwartungen der Nutzer. Eine Strategie zur Reduzierung der notwendigen Installationskosten besteht darin, bereits im Gebäude vorhandene Infrastruktur, wie z.B. Wireless Local Area Networks (WLAN), zur Positionierung zu nutzen. Ein solches Positionierungssystem steht nun in unserem Bürogebäude der Technischen Universität Wien zur Verfügung und kann zur Lokalisierung von Nutzern mit einem mobilen, WLAN-fähigen Notebook oder Pocket-PC verwendet werden. Mit diesem System kann man einen Nutzer mit einer durchschnittlichen Genauigkeit von ± 1 bis 4 m positionieren, womit man in der Regel in der Lage ist, den Büroraum anzugeben, in dem sich der Nutzer gerade befindet, bzw. eine kontinuierliche Positionsbestimmung und Zielführung bei einem sich bewegenden Nutzer durchzuführen. Zusätzlich wird nun der Einsatz der Radio Frequency Identification (RFID) Technologie zur Positionierung in Gebäuden und städtischen Umfeld untersucht, wobei die RFID-Tags an ausgewählten, bekannten Punkten, den sog. Active Landmarks, angebracht werden. Der Nutzer kann dann mit Hilfe einer zellbasierten Positionierungsmethode lokalisiert werden, wenn er sich im Empfangsbereich von einem RFID-Tag befindet und ein RFID Lesegerät (z.B. in PC-Kartenformat) verwendet. Zur Zeit bereiten wir die Installation eines solchen Positionierungssystems in unserem Bürogebäude und der umliegenden, urbanen Umgebung vor. Weiters soll in Zukunft eine Kombination mit WLAN und der Koppelnavigation untersucht und verwirklicht werden.

1. Introduction and Overview about Previously Conducted and Current Research at our University

Personal mobility and location-based services have been identified as main research areas in recent years. Most developments have started with emerging navigation and tracking systems in the area of vehicle guidance in the early 1980's.

The avaiability of GPS receivers has significantly changed the navigation market and their integration with other existing techniques has become a common integration scenario.

Nowadays mainly multi-sensor solutions are employed as satellite positioning does not work in every environment, e.g. in urban canyons where satellite signals are frequently blocked or positioning accuracies are low due to a weak

reveiver-satellite geometry (i.e., high DOP (Dilution of Precision)). Based on these developments the author started to work on the improvement of location techniques for personal navigation about a decade ago. In a first study emerging vehicle navigation systems on the European market have been analyzed and the integration of GPS and dead reckoning in future intelligent vehicle navigation systems (IVNS) has been discussed. In the following, the use of alternative location techniques, e.g. the integration of cellular phone positioning into vehicle navigation systems was investigated. Part of the work was carried out in a joint research project[1] with the Department of Land Surveying and Geo-Informatics of the Hong Kong Polytechnic University (see [10]).

Apart from navigation systems for vehicles research efforts have been concentrated on the challenging task of location determination for pedestrians in recent years (see e.g. [5], [12]). In the work package 'Integrated Positioning' of the research project NAVIO[2] at the Vienna University of Techology different aspects of location determination of a pedestrian in a combined indoor and outdoor urban environment have been investigated. Before the start of the project a study was conducted to investigate the possible sensors that are suitable to be integrated in a pedestrian navigation system. It was found that a pedestrian navigation system should include at least a GPS receiver for absolute position determination and dead reckoning (DR) sensors for relative positioning such as a heading sensor and accelerometers for measurement of the direction of motion and the distance travelled as well as a barometric pressure sensor for altitude determination [18]. A multi-sensor system has been developed in NAVIO and first tests of the integrated location sensors were presented in [16]. For the integration of all available sensor observations a new multi-sensor fusion model that makes use of knowledge-based systems was developed (see section 5).

The provision of navigation and guidance services in challenging indoor environments (e.g. in large office buildings, shopping centres, railway stations, airports, etc.) is very important; especially in the case of emergency situations (e.g. to locate firefighters in a rescue situation inside a building). New developments in indoor location techniques have been investigated and their principle of operation, application and performance was analyzed (see [8]). Systems available on the market use signals such as infrared, ultrasonic and radio signals. Most of these systems, however, require expensive installations of a larger number of receivers or transmitters in the indoor environment. To reduce installation costs an approach was choosen which makes use of already available infrastructure, i.e., the use of Wireless LAN (WLAN or WiFi) [1].

At the Vienna University of Technology a location system based on WLAN has just been installed recently in our office building. For that purpose a cooperation with the German research institute IMST GmbH[3] has been established. The WLAN location system 'ipos' is jointly analyzed and further improvements will be made to the system design. A performance test of the system ipos is presented in section 2. Apart from WLAN also other techniques, such as Ultra Wide Band (UWB) [9], [20] and RFID [2], [17] are currently investigated in more detail and their integration into the system design of a pedestrian navigation system is analyzed. The main advantage of the new wireless radio signals in the UWB band will be the improved performance and accuracy compared to the current WLAN standard. Using these signals positioning accuracies on the dm-level can be achieved in the future. If positioning accuracy requirements are not so high then RFID will be a cheap alternative to perform cell-based positioning. If a user is in the read range of an RFID tag which is placed at a known location (so-called active landmark) then he can retrieve the tags information with its current location. The mobile user has to carry only a reader which can be in the form of a PC card and plugged into a mobile device (e.g. a pocket PC). This form of positioning can be applied also to augement GNSS positioning in areas with no satellite visibilty (e.g. in tunnels, under bridges, etc.) in vehicle navigation. For indoor location determination a concept has already been developed (see section 3). System testing will be performed in a localization testbed at the Vienna University of Technology at our department. In the following, first the performance of the WLAN system ipos and then the concept of RFID positioning will be discussed in more detail.

2. Performance of the WLAN Indoor Positioning System ipos

A common approach for the localization of a mobile device by means of WLAN (or WiFi) is based on measurements of received signal strengths of the WLAN signals from surrounding access points at the terminal [1]. An estimate of the location of the terminal is then obtained on the basis of these measurements and a signal propagation model inside the building. The propagation model can be obtained from modelling of the surrounding environment or with prior calibration measurements at certain locations. In the first case, a building model is used to estimate the signal propagation of the radio signals from all available access points; whereas in the second case, the signal strengths are measured at certain known locations in the building. Using the second approach higher positioning accuracies can be obtained as it is based on actual measurements of signal strength values. The calculation of the location of a user takes place in two phases: an offline and an online phase. During the offline phase (i.e., the calibration), which has to be executed only once for each building, a so-called radiomap will be composed. This radiomap can be considered to be a collection of calibration points at different locations in the building, each with a list of radio signal strength indicator (RSSI) values for visible access points at that particular location. This process is also known as fingerprinting. During the online phase, the calibration points are being used to calculate the most probable location of the user, whose actual location is unknown. Further information about the principle of operation of the WLAN positioning system 'ipos' can be found in [15].

In a first study the performance of the WLAN positioning system ipos was analyzed in an office building of IMST GmbH in Germany in a diploma thesis (see [11]). In the tests different scenarios were investigated; e.g. if a standing or moving user can be located in an office room or a predefined area and when he is moving from one office room to another along the corridor. The expected positioning accuracy in the range of ± 1 to 4 m could be confirmed in the tests in most cases. The system is now available in our office building of the Vienna University of Technolgy and makes use of the already installed access points for wireless communication. As the ipos positioning system is a software solution which runs on the mobile device or a server in the network, standard WLAN hardware without any modifications can be used. Recent performance tests have been conducted and their main results are summarized in the following.

For the performance tests the office rooms and corridor between two staircases on the 3rd floor of our building have been selected (see Figure 1). In every office room (with a size of around 19.5 m^2) at least 2 calibration points are located, in the two general teaching rooms which have a size of around 110 m^2 around 10 calibration points and along the cooridor 15 calibration points were positioned. Along the corridor there are three WLAN access points evenly distributed. Apart from those, also the signals of two additional access points located on other levels in the building have been used in the test. Figure 1 shows the location of the test points for a standing user and a classification of the deviations of the position fixes from the selected 'true' location. A few test points along the corridor, in office rooms and the general teaching rooms have been selected. On every test point position fixes were determined over several minutes with a frequency of 1 Hz. The deviations of the position fixes from the truth were in the range of 0 to 3 m for most of the points. As an example, Figure 2 shows the location of around 250 position fixes on one point which is located in the hallway in front of the elevator (i.e., point 1 in Figure 1) in a local coordinate system (where the Y-axis is parallel to the main axis of the building tract; note that many position fixes in Figure 2 have the same location). The standard deviation of the position fixes in the X-coordinates resulted in ± 0.3 m and in Y-coordinates in ± 1.0 m. Remarkable is that the position fixes show a significant systematic offset from the true location (X=111.39 m, Y=109.63 m) where the deviation of the mean value of all position fixes from the true location is 2.1 m. In respective to the location of the mean value, 67.2 % of the individual position fixes have a smaller Y-coordinate than the mean value (i.e., Y=110.26 m) and only 32.8 % have a larger value for the Y-coordinate. In the X-coordinates 46 % of all position fixes have a larger value and 54 % a smaller value than the X-coordinate of the mean value (i.e., X=109.35 m). The mean positional error of all position fixes on point 1 in relation to the truth is ± 2.9 m.

Figure 3 shows the track of a moving user from the start point in the hallway in front of the elevator to the general teaching room. The reference track is shown in green and the track of the moving user in light blue. The deviations from the reference track are in the range of 1 to 4 m. As can be seen from Figure 3, the reference track is followed quite

nicely along the corridor where the deviations are less than 1.3 m. The largest deviations occurred at the end of the track as the user entered the general teaching room. The higher positioning accuracies along the corridor is due to the larger number of calibration points. The number of calibration points in the general teaching room seems not to be sufficient and should be increased. In the example, the track crosses also the wall between the corridor and the room and reaches a maximum deviation inside the room of around 4 m. In this case, an improvement of the solution can be achieved if the building map is used in a post-processing step which does not allow the track of the user to cross walls when there is no door between them. Then the crossing of the wall can be matched to the nearest door between the corridor and the room.

Deviation of position fixes: ○ 0 - 3 m ◉ 3 - 5 m |— 10 m —|
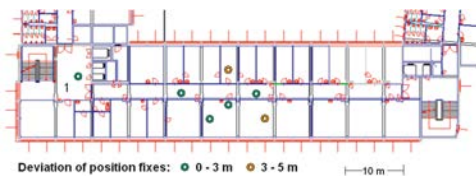
Figure 1: Location of the test points for a standing user in the localization testbed (i.e., 3rd floor of our office building of the Vienna University of Technology) showing a classification of the deviations of the position fixes from the selected 'true' location.
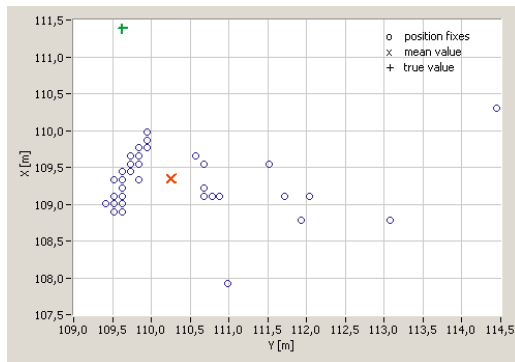
Figure 2: Location of the position fixes for point 1 in the hallway in front of the elevator in respective to its true location.

—— Reference track
—— Track of the moving user |— 10 m —|
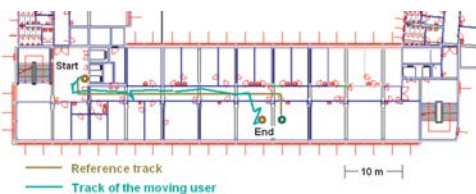
Figure 3: Track of a moving user from the elevator in the hallway to the general teaching room.

3. RFID Positioning and Active Landmarks

Apart from WiFi fingerprinting for indoor location determination the use of RFID for positioning (see [2], [17]) using active landmarks is planned to augment the indoor location determination in araes which are not covered with WLAN. RFID relies on storing and remotely retrieving data using devices called RFID transponders. Typical RFID systems consist of transponders (called tags) with antennas and transceivers (called readers). RFID tags contain silicon chips and antennas to enable them to receive and respond to radio-frequency queries from a RFID reader. RFID tags can be either active, semi-active or passive. Passive tags require no internal power supply, whereas active tags require a power source. The passive tags do have, however, a smaller read range than active tags, i.e., in the range of about 2 mm up to several metres depending on the chosen radio frequency. On the other hand, long range active tags with their own power supply could have a read range of up to 100 m. Another advantage is that active tags have larger memories than passive tags and the ability to store additional information (apart from the tag ID). At present, the smallest active tags are about the size of a coin. Further information about the underlying technology can be found in [3].

For indoor and outdoor pedestrian location determination we propose that RFID tags are installed at known locations in the surrounding environment (so-called active landmarks, at e.g. street crossings, entrances of buildings and offices, at regular distances inside of buildings, etc.). The system user would be equipped with a portable RFID reader module. If the tag's information can be retrieved the user is located in a cell of circular shape with the location of the tag in the centre and a radius equal to the possible read range of the tag. The used location method is referred to as Cell of Origin (CoO) [7]. Several tags located in the smart environment can overlap and define certain cells that intersect. The position of the user can therefore be determined using the network of the tags which can be made available in a database. Figure 4 shows an example for the proposed location of active landmarks equipped with RFID tags on the 3rd floor of our office building at the Vienna University of Technology. In the chosen indoor environment the active RFID tags will be placed at lift entrances and doors to offices, at the staircases at different levels, inside office rooms (e.g. the secretary's office) and the general teaching rooms as well as at regular distances along the corridor. We propose to use wireless

long-range RFID systems from Identec Solutions for the positioning of a pedestrian in the localization testbed. Using the Intelligent Long Range® (ILR®) technique, the user can be located at a distance of up to 100 m [6]. Higher positioning accuracies can be achieved by reducing the sensitivity in the reader. It is therefore possible to limit the read range down to a few meters. Then in the presence of an active landmark the RFID positioning can be used if no GNSS positioning is available due to satellite signal obstructions in urban areas or no WLAN positioning is possible in the indoor environment. The research in this field is conducted in a new research project [4]. In the following section a combination of WLAN and RFID positioning in the localization testbed of our University will be discussed.
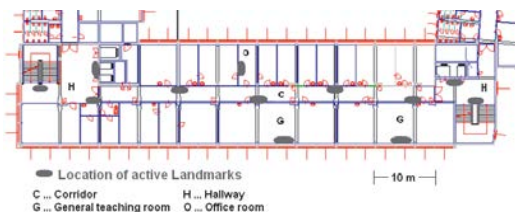

Figure 4: Concept for location placement of active landmarks equipped with RFID tags in the localization testbed.

4. Augmentation of WLAN Positioning with RFID

The performance tests of the WLAN positioning ipos presented in section 2 have also shown that in some areas in the localization testbed the achievable positioning accuracies are lower than usual or that a positioning is not possible at all (e.g. at the boundaries of the covered area). In this areas a calibration was not possible due to the fact that either this area was not accessible or the signal strenghts of the WLAN signals from the chosen five access points were not sufficient. To analyse the WLAN coverage and the achievable positioning accuracies in the testbed a simulation using image processing was performed. For that purpose a simulation software called 'Kingston' was developed that uses the ipos database with the signal strength values of the surrounding access points at certain locations inside the building obtained during the calibration in the offline phase. Figure 5 shows a visualization of the WLAN coverage and achievable positioning accuracy in the localization testbed. The pixel

colour intensity in the image reflects the signal strength and the positioning accuracy on that particular location. A high level of signal strengths to all visible access points results directly in a good positioning accuracy. Using this simulation tool it is now possible to determine if either additional WLAN access points or active landmarks equipped with RFID tags are needed. As an example, Figure 5 shows two locations for proposed RFID tags where one is located in the main staircase on the left and the second in an office room on the right. In simulating the cell-based positioning an isotropic distribution of the range of the RFID system was assumed. The degradiation of the signal strength due to walls, however, was not modelled and has been neglected in this simulation. Based on the simulation tool it is possible to decide how the WLAN positioning system can be augmented meaningful by using RFID. Using a combined indoor location system integrating WLAN and RFID, the areas providing good positioning accuracy in the testbed can therefore be increased significantly.
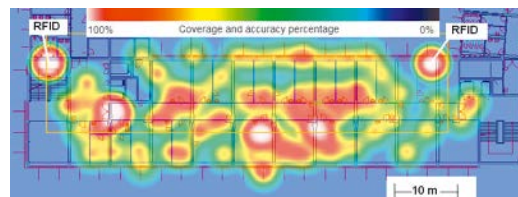

Figure 5: Visualization of the WLAN coverage and achievable positioning accuracy in the localization testbed and augmentation with two active landmarks equipped with RFID tags.

5. Integration with Dead Reckoning (DR) Sensors

In the work package 'Integrated Positioning' of the research project NAVIO[2] a pedestrian navigation system has been developed in recent years (see e.g. [13]). In the system dead reckoning sensors are employed for relative position determination from a given start position and GPS is employed for absolute positioning. Using such a multi-sensor approach the current user's position and the distance travelled, the direction of motion and the changes in altitude can be determined. For the integration of all sensor observations a new fusion model based on knowledge-based systems has been developed ([14], [19]). The approach makes use of a knowledge-based component for a

preprocessing of all available sensor observations. In this preprocessing step outliers and large errors are detected and these observations discarded as well as the quality of the new observations analyzed before their integration in a centralized Kalman filter (see [4]). Then the knowledge about the quality of the current observations from the preprocessing filter can be used to adapt the stochastic Kalman filter model; for example, in the case of large observation errors their weighting can be considerably reduced in the filter. The optimal estimate of the current user's position is then obtained in the filter process. The performance of the NAVIO system has been tested in different site conditions (e.g. sites providing free satellite visibility and in urban areas where satellite signals are frequently blocked). Major test results have been reported in [13]. Using the new multi-sensor fusion approach a high reliabilty and location accuracy for continuous position determination of a pedestrian in urban environments was achieved. In general, it could be shown that a user can be located continuously in outdoor urban environments with a positioning accuracy of a few metres. A next development step deals with the positioning of a user in indoor environment. For that purpose the dead reckoning observations will be augmented with WLAN and RFID positioning if this systems are available. Using the absolute position from WLAN or RFID it is then possible to correct for the drift rates of the DR sensors.

6. Concluding Remarks and Outlook

One might think that a lot has been achieved already in terms of positioning accuracy and system performance in the field of personal navigation and location-based services. The acceptance by the public, however, does not depend only on the achievable location accuracy but also on the reliability of continuous position determination and the integrity of the systems. The author believes that still a lot has to be done in this field of research and he will continue his work in this challenging area. Especially the upcoming introduction of the GALILIEO system will play an important role in navigation applications and if it will be augmented by alternative techniques and location sensors as shown by the author in the presented work the reliability of the continuous position determination can be significantly increased. This will help to improve the acceptance by the user who tries to find his way in an unfamiliar environment no matter if it will be in an urban or a combined indoor/outdoor environment.

References

[1] Bahl P., V. Padmanabhan (2000): Radar: An In-Building RF-based User Location and Tracking System. Proceedings of IEEE Infocom, Tel-Aviv, March, pp. 775-784.

[2] Chon H. D., S. Jun, H. Jung, S. W. An (2004): Using RFID for Accurate Positioning. in: Papers presented at the 2004 International Symposium on GNSS, Sydney, Australia, 6-8 December,2004, 10 pgs.

[3] Finkenzeller K. (2003): RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, 2nd edition, Wiley & Sons LTD, see also http://www.rfid-handbook.de/ (Last date accessed: May 2007).

[4] Gelb A. (Ed.) (1986): Applied Optimal Estimation. The MIT Press, Cambridge, Massachusetts and London, England, 9th print, 374 pgs.

[5] Grejner-Brzezinska D., C. Toth, S. Moafipoor, Y. Jwa (2006): Multi-Sensor Personal Navigator Supported by Human Motion Dynamics Model. in: Papers presented at the 3rd IAG Symposium on Geodesy for Geotechnical and Structural Engineering and 12th FIG Deformation Measurement Symposium, May 22-24, 2006, Baden, Austria, see http://www.fig.net/commission6/baden_2006/PDF/NMM/Grejner-Brzezinska.pdf (Last date accessed: May 2007).

[6] Identec Solutions (2006): Intelligent Long Range® (ILR®) RFID system, Product Information, http://identecsolutions.com/ (Last date accessed: May 2007).

[7] Ingensand H., P. Bitzi (2001): Technologien der GSM-Positionierungsverfahren. Allgemeine Vermessungsnachrichten (AVN), No. 8-9, pp. 286-294.

[8] Kistenich M. (2005): Indoor Positioning: Vergleich von Systemen zur Positionsbestimmung und Navigation in Gebäuden. Diploma Thesis, Institute of Geodesy and Geophysics, Research Group Engineering Geodesy, Vienna University of Technology.

[9] Kong H., Y. Kwon, T. Sung (2004): Comparisons of TDOA Triangulation Solutions for Indoor Positioning. in: Papers presented at the 2004 International Symposium on GNSS, Sydney, Australia, 6-8 December, 2004, 11 pgs.

[10] Mok E., G. Retscher, L. Lau (2002): Development of an Event-Reporting Intelligent Vehicle Navigation System for Areas with Urban Canyons. Survey Review, April 2002, Vol. 36 No. 284, pp. 398-409.

[11] *Moser E. (2006):* Genauigkeits- und Leistungstest eines WLAN Indoor Positionierungssystems. Diploma Thesis, Institute of Geodesy and Geophysics, Research Group Engineering Geodesy, Vienna University of Technology.

[12] *Renaudin V., O. Yalak, P. Tome (2007):* Hybridization of MEMS and Assisted GPS for Pedestrian Navigation. Inside GNSS, January-February 2007, pp. 34-42.

[13] *Retscher G. (2007 a):* Test and Integration of Location Sensors for a Multi-sensor Personal Navigator. The Journal of Navigation, Vol. 60, No. 1, January 2007, pp 107-117.

[14] *Retscher G. (2007 b):* A Knowledge-based Kalman Filter for an Intelligent Pedestrian Navigation System. Paper accepted for publication in Survey Review.

[15] *Retscher G., E. Moser, D. Vredeveld, D. Heberling, J. Pamp (2007):* Performance and Accuracy Test of a WiFi Indoor Positioning System. Journal of Applied Geodesy, Vol. 1, No. 2 (in press).

[16] *Retscher G., M. Thienelt (2004):* NAVIO – A Navigation and Guidance Service for Pedestrians. Journal of Global Positioning Systems, CPGPS, Vol. 3, No. 1-2, pp. 208-217.

[17] *Roth J. (2004):* Data Collection. in: Schiller, J. and Voisard A. (Eds.): Location-Based Services, Morgan Kaufmann Publishers, USA, pp. 175-205.

[18] *Skolaut G. (2002):* Untersuchung von Messsensoren zum Einsatz in Navigationssystemen für Fußgänger. Diploma Thesis, Institute of Geodesy and Geophysics, Research Group Engineering Geodesy, Vienna University of Technology.

[19] *Thienelt M., A. Eichhorn, A. Reiterer (2005):* Konzept eines wissensbasierten Kalman-Filters für die Fußgängerortung (WiKaF). Österreichiche Zeitschrift für Vermessung und Geoinformation (VGI), No. 2, pp. 96-104.

[20] *Win M., R. Scholtz (1998):* On the Performance of Ultra-Wide Bandwith Signals in Dense Multipath Environment, IEEE Commun. Letters, Vol. 2, No. 2, Feb. 1998, pp. 51-53.

Contact

Ass.-Prof. Dr. Günther Retscher, Institute of Geodesy and Geophysics, Research Group Engineering Geodesy, Vienna University of Technology, Gusshausstr. 27-29, A1040 Vienna, Austria, E-mail: gretsch@pop.tuwien.ac.at

vgi

A Knowledge-Based Optical 3D Measurement and Analysis System for Quality Control

Alexander Reiterer, Martin Lehmann, Johannes Fabiankowitsch and Heribert Kahmen

Abstract

In the FWF project P18286 *"Multi-Sensor Deformation Measurement System Supported by Knowledge Based and Cognitive Vision Techniques"* a new kind of image-based measurement system is under development. This system is able to detect un-signalised object points by means of appropriate algorithms – the procedure is divided into three main steps: image pre-processing, automated point detection and interactive point filtering. The system is based on new techniques (originally developed in the area of Artificial Intelligence) which shall be used for the task of deformation measurement, analysis and interpretation. Examples for such techniques are knowledge-based systems, cognitive vision and image understanding methods and case-based reasoning.

Kurzfassung

Derzeit wird im Zuge des FWF Projektes P18286 *"Multi-Sensor Deformation Measurement System Supported by Knowledge Based and Cognitive Vision Techniques"* ein neues bild-basiertes Messsystem entwickelt. Dieses System ist fähig nicht-signalisierte Punkte mit Hilfe geeigneter Algorithmen zu detektieren – die Prozedur ist in drei Teilschritte unterteilt: Bildaufbereitung, automatische Punkterfassung und Punktfilterung. Das System basiert auf neuen Techniken (ursprünglich im Fachgebiet der Künstlichen Intelligenz entwickelt), welche für die Teilschritte der Deformationsmessung, -analyse und -interpretation genutzt werden sollen. Beispiele solcher Techniken sind wissensbasierte Systeme, Cognitive Vision, Bildverstehen und Fallbasiertes Schließen.

1. Introduction

In science and industry (e.g. in architecture, medicine, or construction), highly accurate 3D representations and/or monitoring of objects are required. A great variety of optical 3D measurement techniques like laser scanners, photogrammetric systems, or image-based measurement systems is available to achieve this need.

In comparison with laser scanners, image-based systems measure objects with higher accuracy; compared with photogrammetric systems, they can be used more easily for on-line measurement processes. This will especially be the case, if the measurements can be performed with a high degree of automation.

During the last years, research in the area of image assisted measurement systems has gained an increasing interest. Most systems are working on the basis of user-interaction. Notable are the systems developed by *Leica Geosystems* [18, 19], by *Technische Universität München* [20], by *Ruhr University Bochum* [16] and by *Topcon* [17] (see Figure 1).

The central topic of all image-based measurement systems is the calculation of 3D object coordinates from 2D image coordinates for subsequent processing steps. As mentioned above the concept of these systems is based on a permanent interaction between user and system.

One of the main goals of the research work done at the *Institute of Geodesy and Geophysics* of *Vienna University of Technology* is to automate such image-based measurement systems. *Fabiankowitsch* [6] was the first researcher who experimented with a new measurement system on the basis of the *Leica TM3000* videotheodolite. His work resulted in special measurement methods (filtering techniques) for active targets. In 1994, the research project *"Stereovideometry and Spatial Object Recognition"* commenced. One of the milestones of this project was the work done by *Roic* [14]. The aim of his work was to prepare images of unsignalized targets by using image processing methods, in order to make interactive and automatic spatial surface measurement possible. A few years later *Mischke* [11] developed a powerful measurement system based on two videotheodolites. The system was able to measure active or passive targets and non-signalized points, like intersections of edges or

lines. *Mischke* has implemented the *"Förstner Interest Operator"* to select all remarkable but non-signalized points. *Interest operators* (IOPs) were well-known for offline applications in photogrammetry, but had not been used for videotheodolites before.

The disadvantage of all these measurement systems is the requirement for a well-trained "measurement expert" who has to have certain skills and experience to properly handle the complex system. From sensor orientation, data capturing to the detection and tracking of points of interest, a series of actions and decisions have to be performed. In two research projects at the *Institute of Geodesy and Geophysics* of *Vienna University of Technology (1:2001-2005 - "Theodolite-based and Knowledge-based Multi-Sensor-System"; 2: 2006-2008 - "Multi-Sensor Deformation Measurement System Supported by Knowledge-Based and Cognitive Vision Techniques"*) a new kind of measurement system is under development [13]. This system is based on new techniques (originally developed in the area of Artificial Intelligence) which shall be used for the task of *decision making* and *quality control* (deformation measurement, analysis and interpretation). We report on the state-of-the-art of such a new measurement system, its functionality and development stage.

2. Concept for a new kind of measurement system

An image-based measurement system is a combination of different components:

- image sensors (our system is based on the Leica IATS – see Figure 1c),
- a computer system,
- software (e.g. control system, decision-making system, image processing, etc.),
- accessories.

IATS and videotheodolites have a CCD camera in its optical path. The images of the telescope's visual field are projected onto the camera's CCD chip. The camera is capable of capturing mosaic panoramic images through camera rotation, if the axes of the theodolite are driven by computer controlled motors. With appropriate calibration these images are accurately georeferenced and oriented as the horizontal and vertical angles of rotation are continuously measured and fed into the computer.

An *optical system* for such a system (see Figure 2) was developed by *Leica Geosystems* [19]. It is reduced to a two-lens system consisting of the front and the focus lens. Instead of an eyepiece a CCD sensor is placed in the intermediate focus plane of the objective lens. The image data from the CCD sensor are fed into a computer using a synchronized frame grabber. For the transformation of the measured image points into the object space the camera constant must be known. In an optical system with a focus lens the camera constant, however, changes with the distance of the object.
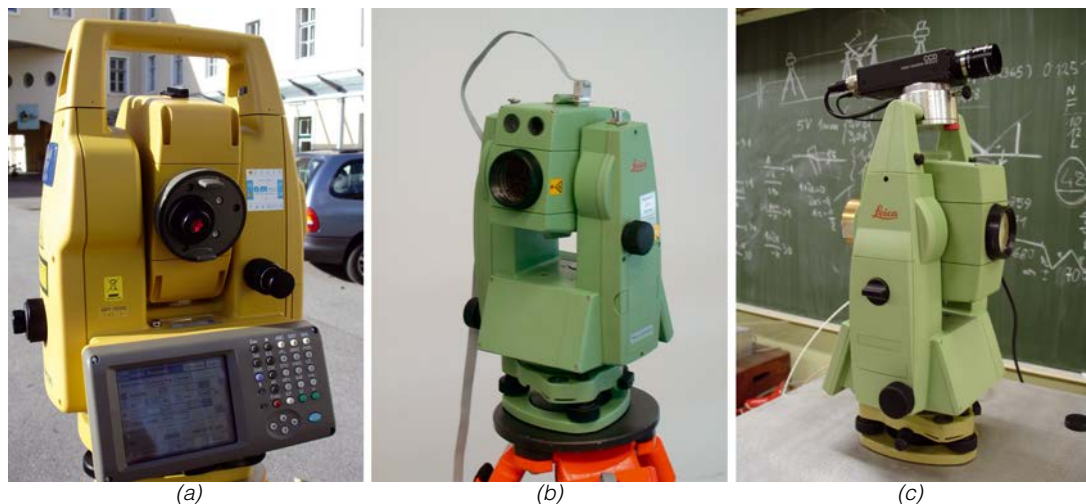

Figure 1: Examples of image-based measurement systems developed by (a) Topcon, by (b) the University of Bochum, and (c) by Leica Geosystems..
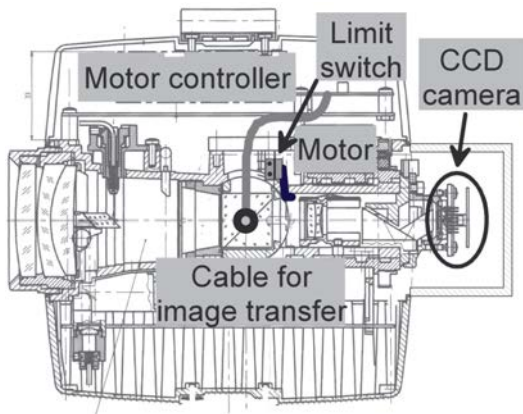
Figure 2: Cross section of the telescope developed by Leica Geosystems [19].

The camera constant can be derived from the focal length. This can be performed automatically if an encoder measures the focus lens position relative to an origin, which is chosen when focusing to infinity. Now the optical mapping model includes not only the theodolites axis errors and the vertical index error, but also errors resulting from a displacement of the projection centre from the intersection of the theodolites axes and from the optical distortions for field points. Consequently calibration of an image-based theodolite has to comprise all these errors.

As mentioned, our new system is based on a combination of the IATS prototype developed by *Leica Geosystems* and a terrestrial Laser Scanner (TLS). The data of the different sensors have to be merged by a special data fusion process, which is designed as a knowledge-based approach (the implementation will be done in the near future). Such a system provides an immense number of 3D data, both from the IATS system and from the laser scanner. This point cloud may be reduced by filtering, even if not very effective. Our approach builds on *cognitive vision techniques*. These methods can be used as well for finding regions of interest as for point detection. A combined system of this kind will be a great challenge in the future (details about data fusion and interpretation are currently in publication/review).

The basic concept of our new measurement system / procedure can be formulated as follows:

■ IATS and TLS are used in a common way for data capturing. Roughly speaking the point-oriented method of the image assisted theodolite is capable to capturing structured regions (e.g. edges and corners) of an object with high

accuracy while the area-oriented method of the laser scanner is able to survey the unstructured regions. The data of the different sensors have to be merged by a suitable data fusion process (as mentioned above). In the long term the combination of image assisted theodolites and laser scanners should yield optimal results for most fields of application.

■ The captured data (images and laser scanner data) are used by an *data interpretation tool* (which is working on the basis of *image understanding* and/or *cognitive vision techniques*) to produce a description of the scene/object. This can be done by recognizing different objects and assigning them to proper categories together with information about the object and relevant parameters. This process results in a special kind of "*object information system*".

■ The information yielded is used as input for a decision system, to produce a list of actions (e.g. ranking of suitable point detection algorithms or the context between different object parts and suitable sensors).

■ Furthermore, on the basis of the generated scene description, regions of interest (ROI) are selected. One ROI is the smallest measurement area in the current framework. For provision of a suitable data base the order of measurements can be predefined either as individual measurements, as repetitive measurements (each hour at same minute, each day at same hour, each month at same day) or can be created automatically.

■ On the basis of the captured measurements a new kind of deformation analysis will be processed – we will call this process in the following as *deformation assessment*. This step results in an appropriate report.

The abstract, simplified measurement and analysis procedure is shown in Figure 3.

The process of *measurement* consists of several steps, including image capturing, image pre-processing, point detection, calculation of 3D point coordinates, etc.

The part of point detection by means of the image-based measurement system can be divided into two main steps: detection of points in the image(s) and measurement of these points in the object space. To process such a procedure, image points have to be transformed into the coordinate system of the measurement system. This is done by a complex mapping function

(using the pinhole camera model) developed by *Walser* [18]. The underlying object measurement is realized on the basis of conventional tacheometric measurement elements – each detected image point is measured in the object space by horizontal and vertical angles and distance.

The integration of a laser scanner device into the measurement procedure is currently only realized by prototypical implementation (as shown in Figure 3 data are used as an overview – data processing and orientation is done in a manual way). A fully integrated system is planned for the near future and could be based on techniques developed by *Jäger* [8].

Image pre-processing and image-point detection algorithms are selected on the basis of extracted image features (e.g. *histogram features*, *Haralick moments*, etc.). The selection and combination of suitable algorithms is done in an automatic way, by a *knowledge-based decision system*. The knowledge which was required to be included was obtained in different ways: from technical literature, e.g. [2, 4], previous projects [6, 11, 13, 14] and from experiments. This knowledge was converted into "*If-Then-Statements*" (rules) followed by coding them for the used development tool. The knowledge-based system has been carried out in *Clips*, a productive tool which provides a complete environment for the construction of rule- and object-based systems [3]. More details about this part can be found in [13].

After having detected points of interest in more than one time epoch, *deformation assessment* can be done (see Figure 3).

3. Deformation assessment for quality control

The main goal of the developed *deformation assessment* is to classify relevant deformations on an unstructured point cloud. This process can be divided indo several steps. As a first step a *classical deformation analysis* (as the absolute deformation network type according to [12]) on the basis of the measured point coordinates can be done (deformation analysis by means of laser scanner data is one of the main goal of the actual development done by [7, 8]).

Coupled to this procedure can be a process, which sets up a description of the movements and distortions of the object. An assessment of the deformation can follow. This "deformation classification" must be outlined in a framework of *local-to-global information integration*, by grouping locally measured deformation into a more informative *deformation pattern*. The identification of *prototypical* discriminative deformation patterns will enable to derive a codebook of deformation characteristics to achieve a vocabulary of prototypical patterns for future reference. Early identification of critical prototypical deformation patterns may initiate focus of attention thereon, requiring a more precise and possibly time consuming analytical process that might even involve human intervention.

We have tested several methods for the task of "*deformation clustering*" (e.g. strain analysis, k-means, hierarchical clustering, etc); for an on-line characterisation of deformations it needs a simple and runtime optimised solution – the development of a compact algorithm is necessary.

The new developed procedure can be divided into several steps. First of all the surface of the object is subdivided into regions which might have deformed in the same way (this step will be processed by means of a new developed knowledge-based image understanding tool or in close collaboration with experts from civil engineering domain). This sub-division into
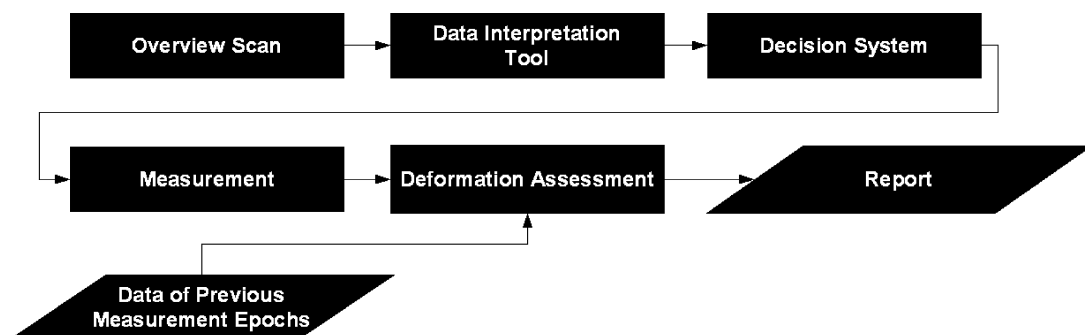

Figure 3: Abstract, simplified measurement and analysis procedure.

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\beta\cos\gamma & \cos\beta\sin\gamma & \sin\beta & t_x \\ \sin\alpha\sin\beta\cos\gamma \ + \ \cos\alpha\sin\gamma & -\sin\alpha\sin\beta\sin\gamma \ + \ \cos\alpha\cos\gamma & -\sin\alpha\cos\beta & t_y \\ -\cos\alpha\sin\beta\cos\gamma \ + \ \sin\alpha\sin\gamma & \cos\alpha\sin\beta\sin\gamma \ + \ \sin\alpha\cos\gamma & \cos\alpha\cos\beta & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Formula 1: Motion of a point described by similarity transformation

regions of interest is founded in the local-to-global information integration as mentioned above. The following calculations and methods are only applied to the single ROIs (due to each ROI can only be modeled as rigid-body motion). In a later step the results of the investigation of these bounded areas are combined to formulate a deformation model for the whole object (rigid-body motion and inner geometry changes).

Inside the ROIs points are detected by means of image processing techniques (so-called *interest operators*). After having measured the points in different time epochs, a classical geodetic deformation analysis is used to find significant motions (this part is based on GOCA) [7]. The determined deformation has to be split into basic motions. Currently, the developed method analyses only *translations* along the coordinate axes and *rotations* around these axes – an extension to other movements/distortions is planned for the near future. The motion of a point $(x, y, z \rightarrow x', y', z')$ is described by an *similarity transformation* [4] (in terms of homogeneous coordinates, modeled with 6 parameters) see Formula 1.

To calculate the parameters (for the rotations around the axes: α, β, γ and for the translation along the axes: t_x, t_y, t_z) inclusive their standard deviation a *Gauss-Helmert* equalization is used. To use this method a minimal number of three points is necessary (the distribution of these points in the considered region is (nearly) irrelevant).

With an artificial testobject, shown in Figure 4, we have done some tests, to proof the accuracy of the algorithm.

As mentioned, the calculated parameters only describe the deformation of the individual region – to make a conclusion about the deformation of the whole object, it is necessary to group these regions by their specific parameters (this combination can be done by statistical clustering methods).

On the basis of well known prototypical "*deformation cases*" a special kind of codebook of deformation characteristics will be implemented. The matching between cases in the database (codebook) and a new (unknown) one will be processed by means of *case-based reasoning* (CBR) (see Figure 5) [1].

4. Conclusion

In this paper, first steps of the development of an optical 3D measurement and analysis system for quality control have been described. The main task of this development has been the automation of different decision makings in the course of the measurement and analysis process. The analysis process is based on a combination between a conventional deformation analysis and c*ase-based reasoning* techniques. For an on-line system a fast execution of all algorithms and processes is necessary.
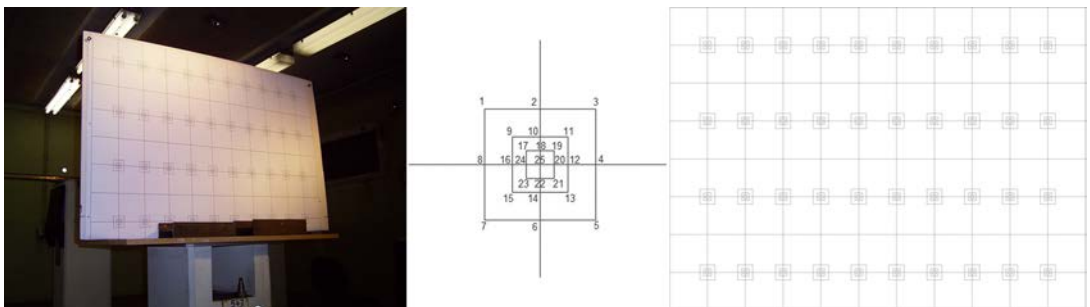

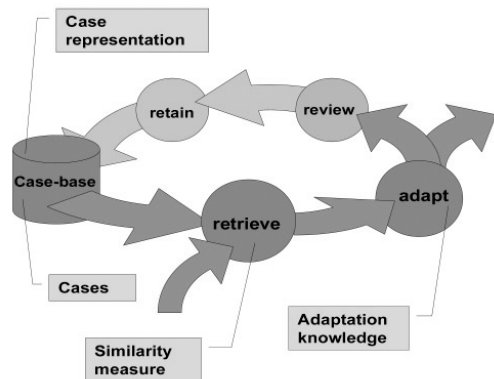*Figure 4: (a) Testobject, (b) gridsector with pointnumbering, (c) grid.*

Figure 5: Case-based reasoning cycle [1].

The vision for the next years is the development of a fully integrated and automated measurement system, supported by image-based measurement techniques and laser scanning techniques. Also the fusion with other sensors (GNSS, PMD, etc.) will be a challenging task.

Such a integrated system represents an approach for an automated on-line working system. The degree of automation can be very high, whereas by decision-making, human interaction remains an important part of the workflow even though the amount of decisions done by the user can be reduced considerably to a minimum.

References

[1] *Aamodt, A. / Plaza, E.:* Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, AICOM 7, 1994.

[2] *Buchmann, P.:* Entwicklung eines Messsystems zur automatischen polaren Objekterfassung am Beispiel der Bauaufnahme. DGK, Volume C, No. 456, München, 1996.

[3] *CLIPS-Project:* http://www.ghg.net/clips/CLIPS.html (as 02/2007).

[4] *Drobniewski, P:* Integration geodätischer und geotechnischer Beobachtungen und Strukturinformationen für eine 3D-Strainanalyse'', PhD, Technische Universität Bergakademie Freiberg, 2005.

[5] *Dunn, F. / Parberry, I.:* 3D Math Primer for Graphics and Game Development, Wordware Publishing Inc., 2002.

[6] *Fabiankowitsch, J.:* Automatische Richtungsmessung mit digitalen Differenzbildern. PhD Thesis, Vienna University of Technology, 1990.

[7] *Goca:* http://www.goca.info/ (as 02/2007).

[8] *Jäger, R.:* GOCA Presentation – Internal Seminar, 2006.

[9] *Jekeli, C.:* Inertial Navigation Systems with Geodic Applications. Gruyter, 2001.

[10] *Juretzko, M. / Scherer, M.:* Hochgenaue polare Fassadenvermessung. Ingenieur-vermessung 2000, XIII. International Course on Engineering Surveying – TU München, pp.400-405, Wittwer Verlag Stuttgart, 2000.

[11] *Mischke, A.:* Entwicklung eines Videotheodolite-Mess-systems zur automatischen Richtungs-messung von nicht signalisierten Objektpunkten. PhD Thesis, Vienna University of Technology, 1998.

[12] *Pelzer, H:* Geodätische Netze in Landes- und Ingenieurvermessung II. 1985.

[13] *Reiterer, A.:* A Knowledge-Based Decision System for an On-Line Videotheodolite-Based Multisensor System. PhD Thesis, Vienna University of Technology, 2004.

[14] *Roic, M.:* Erfassung von nicht signalisierten 3D-Strukturen mit Videotheodoliten. PhD Thesis, Vienna University of Technology, 1996.

[15] *Scherer, M.:* Architectural Surveying by combined Tacheometric and Photo-grammetric tools – About the Realisation of a Synthesis. Workshop "Archeology and Computer", Wien, 2003.

[16] *Scherer, M:* Intelligent Scanning with Robot-Tacheometer and Image Processing – A Low Cost Alternative to 3D Laser Scanning? In: FIG Working Week, Athens, 2004.

[17] *Topcon:* http://www.topcon.com/ (as 02/2007).

[18] *Walser, B.:* Development and calibration of an image assisted total station. PhD thesis, ETH-Zurich, 2003.

[19] *Walser, B. / Braunecker, B.:* Automation of Surveying Systems through Integration of Image Analysis Methods. In: Optical 3-D Measurement Techniques VI, Grün / Kahmen (eds.), Volume I, pp. 191-198, Herbert Wichmann, Karlsruhe, 2003.

[20] *Wasmeier, P.:* The Potential of Object Recognition Using a Servo-tacheometer TCA2003. In: Optical 3-D Measurement Techniques VI, Grün / Kahmen (eds.), Volume II, pp. 48-54, ETH Zurich, 2003.

Contact

Dr. Alexander Reiterer, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-Mail: alexander.reiterer@tuwien.ac.at

Dipl.-Ing. Martin Lehmann, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-Mail: martin.lehmann@tuwien.ac.at

Dr. Johannes Fabiankowitsch, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-Mail: jfabian@pop.tuwien.ac.at

Prof. Dr.-Ing. Heribert Kahmen, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-Mail: Heribert.Kahmen@tuwien.ac.at

Concept of a multi-scale monitoring and evaluation system for landslide disaster prediction

Michaela Haberler-Weber, Andreas Eichhorn and Heribert Kahmen

Abstract

In 2006, OASYS, an EU funded project on a multi-scale monitoring concept for landslides as a basis for an alert system, was completed. 12 institutes from 6 countries tried to merge their multidisciplinary knowledge in the field of landslides and disaster management. The main goal of the research was to develop a cost saving concept for landslide disaster prediction in areas with a higher density of landslides. The present paper reports about the innovative steps and about some highlights of the research, emphasising mainly three tasks:

■ GIS integrated geological evaluations of remote-sensing data to delineate the high-risk areas in regions with a larger number of landslides

■ geometrical analysis of the monitoring data by fuzzy techniques as a basis for the design of the sensor network and

■ geomechanical modelling of the landslide by FD-methods as a basic information for an alarm system.

Kurzfassung

Im Jahr 2006 wurde ein von der EU gefördertes Projekt abgeschlossen, in dem ein mehrstufiges Mess- und Auswertekonzept entwickelt wurde, welches eine Basis für Frühwarnsysteme sein soll. Wissenschaftler von zwölf Instituten aus sechs Ländern kooperierten in interdisziplinären Arbeitsgruppen. Eine besondere Herausforderung war, ein kostensparendes Konzept für Regionen mit einer Vielzahl von Rutschungsgebieten entstehen zu lassen. Dieser Beitrag beschränkt sich auf folgende drei Schwerpunkte:

■ GIS integrierte geologische Evaluierung von Fernerkundungsdaten, um die Grenzen der Gefahrenzonen zu kennzeichnen

■ geometrische Analyse von Beobachtungsdaten mit einem mehrstufigen Fuzzy-System als Basis für ein kostensparendes Design der Sensor-Netzwerke

■ geomechanische Modellierung der Rutschungshänge durch FD-Methoden als Basisinformation für ein Frühwarnsystem

1. Introduction

One main challenge of the research was to develop cost saving solutions which can especially be used in areas with a larger number of landslides. A first proposal of a concept was described in [1].

During the last years, an advanced model was investigated in the project, based on large scale monitoring and evaluation as a first step, regional monitoring as a second step, culminating in a multi-component knowledge-based alert system:

1st step: 'Detection of potential landslides (large scale monitoring and evaluation)'. To get the borders of the moving areas and information about the long-term geodynamical processes a large scale evaluation has to be performed, see Sec. 2, e.g. [2]. This includes e.g. the historical data and remote sensing data, such as aerial photographs, optical and radar images from satellites.

Remote sensing techniques (e.g. In-SAR), differential GPS and tacheometric measurements can be used to obtain additional information about the deformation process, see Fig. 1.

Figure 1: Delineation of risk areas using GPS and total stations.

The measurements are usually performed only three or four times a year, and the results are vector fields describing the displacements and velocities. Based on all displacement information available, an advanced analysis algorithm (see Sec. 3) performs the detection of the so-called taking-off-domains, which are areas where the landslides have their origins, on or below the

surface. In these zones, deformation can be detected at an early stage. Additionally, the taking-off-domains can give a first insight into the possible progress of the landslides.

2nd step: 'High precision permanent measurements in the taking-off-domains'. High precision relative measurement systems (borehole tilt-meters, extensometers, hydrostatic levelling systems, etc.) can be installed in a cost saving way in the area of the taking-off-domains to obtain online information about the geodynamical process (see Sec. 4). This multi-sensor system can measure continuously and can therefore support the real time alert system in this application.

3rd step: 'Impact and risk assessment; development of strategies for knowledge-based alert systems'. The integration of hazard and vulnerability analysis leads to an estimation of the actual risk situation of the affected population. The risk management measures also depend heavily on the specific conditions and include landuse planning, technical measures (e.g. building drainage systems), biological measures (e.g. afforestation) and temporary measures in case of danger.

Three tasks of the project are described below in more detail.

2. GIS integrated geological evaluations of remote-sensing data

The goal of this step is the delineation of high-risk areas in regions with a greater number of landslides. Remote sensing technology embedded in a GIS database can be used as a complementary tool for existing landslide hazard studies, [3]. In the OASYS project LANDSAT ETM, ENVISAT, and ERS imageries are used to produce maps within a 'Landslide Hazard Information System' as layers in a GIS data base with the aim to create user-defined computations of landslide hazard maps. By using earth observation data it is possible to detect traces of past or even recent tectonic movements and mass movements which could be sources for future landslides.

Especially 'Lineament Analysis' based on LANDSAT ETM and radar images can help to delineate those local fracture systems and faults that might influence dynamics and shape of landslides. In this case digital image enhancement is an important step followed by visual interpretation in an interactive manner.

Aim of these techniques is to detect linear or curve linear features in order to find traces of possible slope failure. Combining lineament maps for example with slope degree maps, derived from digital terrain models, helps to delineate areas with higher risk. The term lineament is used for all linear, rectilinear or slightly bended image elements being extracted by image enhancement. Lineaments are symbols for e. g. linear valleys, linear zones of abundant watering, drainage network, peculiar vegetation, landscape and geological anomalies. The definition of the borders of landslide areas can be based on the assumption that lineament systems in a satellite image are closely connected to deformation which is caused by a change in the Earth's crust stress field.

3. Geometrical analysis of monitoring data with fuzzy systems

One aim within the integration of different measurement methods was the precise, continuously monitoring of the taking-off-domains with geotechnical sensors. This information can be used in the alert system to analyse the actual situation.

So one task was the detection of block boundaries between stable and unstable areas or between unstable areas (moving with different velocities in different directions) out of the displacement vectors given by the geodetic deformation measurements. These block boundaries are the optimal places for the installation of geotechnical sensors, which give a very precise relative information on the movement in this area.

Classical deformation measurements are usually done in several epochs; for processing the geodetic deformation analysis is used. This analysis results in displacement vectors for distinct points being observed in each epoch. So an advanced analysis method is necessary, using the displacement vectors for a grouping of observed points into blocks of consistent movement pattern. Within the classical quasi-static geodetic deformation analysis only single point movements can be assessed. A block detection and assessment is only possible manually, i.e. the user can define models describing different blocks, which are assessed by statistical properties (e.g. [4]). This strategy of 'trial and error' is not useful for a large number of object points. So an automated block detection method was necessary.

A pure mathematical approach cannot fully achieve this goal due to the inherent fuzziness of this problem. A human expert solves this task by looking at the graphical representation of the displacement vectors. Keeping in mind other non-mathematical relevant information like geological facts or some properties of the measurement process, he or she can separate the blocks by a combination of mathematical facts and human experience. This kind of processing can be achieved with help of fuzzy systems, where the human way of thinking can be imitated by a rule based expert knowledge.

The task is to find groups of points with a similar pattern of movement, so that the bounderies between these blocks can be identified. Within the automated block detection process, the smallest starting block of 4 points is identified due to some indicators. Then the block is expanded by an iterative algorithm, where the best fitting point is added to the block in each iteration step until no neighbouring points with a similar pattern of movement exist. Two different types of parameters are used as indicators for this block separation (for more details see e.g. [5]):

1. Geodetic parameters: an overdetermined affine coordinate transformation and the derived strain parameters are used to assess the movement pattern of the observed points. If the points of the block under investigation are lying on one block (showing a similar pattern of movement), e.g. the strain parameters and the standard deviation of unit weight of the coordinate transformation are rather small. If in the next iteration step a neighbouring point with another movement pattern is added to this block, these indicators usually are increasing significantly. But for a fully automated analysis, these geodetic indicators are not sufficient. So a second type of parameters is needed, using human expert knowledge in a fuzzy system.

2. Visual parameters: Human experts do the block detection by looking at the pattern of displacement vectors, selecting all points and the corresponding vectors which show a similar length and a similar direction. This is a typical example for the application of fuzzy systems because no sharp definition of 'similar length' or 'similar direction' can be found. One example will be given here for the indicator 'similarity of direction': If the vectors under investigation are within a range of approximately 20 gon, they can be assessed as 'similar'. The greater the difference in the azimuth, the smaller the

indicator 'similarity' will be within the processing, according to the human way of thinking.

The analysis algorithm was implemented in Matlab®. The fuzzy toolbox provides the basic methods like standard membership functions, rule-based inference algorithms and calculation of the output parameters.

Some examples confirmed the applicability of the block separation algorithm in the first part of the OASYS project. Based on this information of the block boundaries, in the next step of the project the high precision geotechnical sensors can be installed in these areas.

4. FD-modelling of the test slope for a knowledge-based alarm system

4.1. Basic concept of the knowledge-based alarm system

The basic concept for a knowledge-based alarm system includes two complementary strategies for the analysis of the current state of the slope and the alarm level decision: the data- and the knowledge-based system analysis. Its architecture is described in detail in [6].

Besides the collection and investigation of geotechnical and geodetic monitoring data one central task for the data-based system analysis will be the provision of calculation results from calibrated numerical slope models which quantify the inner structure of the slope and the mechanisms of possible failure events. Numerical models can principally be represented by DE (distinct element, e.g. [7]), FD (finite difference, e.g. [8]) or FE (finite element, e.g. [9]) codes. Standard software packages are available e.g. from [10]. The slope models will aim at the calculation of the behaviour of the slope under time variable loads (e.g. mass extraction, changes in the ground-water table). In the alarm system the calculation results shall be mainly used for

- the comparison of model predictions and measured displacements to detect deviations from normal behaviour,
- the realistic simulation of failure events as reaction to critical trigger influences (possibility of prevention) and for
- the calculation of the local and global inner stress distribution of the slope which shall be used for the quantification of numerical safety factors (e.g. *FS* = factor of safety, [11]) which enable the evaluation of the risk potential.

One goal of OASYS was to accomplish a feasability study for numerical slope modelling.

4.2. Selected test slope

The monitoring system of TU Vienna (together with Geodata company) was installed at a test slope in a large open cast lignite mine in Germany. This study site provides nearly lab conditions for the development and evaluation of investigation methods for slope mechanisms, this means well known main trigger events (mass excavation process) and an also well known geological structure.

A cross section of the test slope is shown in Fig. 2. The mass extraction is realised by huge bucket wheel excavators which dig out discrete stages, so called 'berms'. The installed monitoring system for regional scale consists of two parts: a geodetic monitoring system (Fig. 2b) with GPS (GOCA = GNSS/LPS/LS-based online Control and Alarm System, see [12]) and robot tacheometers (GEOROBOT, see [13]). In addition a geotechnical monitoring system was installed in the vicinity of the current digging area (top of berm 12, see [14]). The system consists of 1 piezometer, 1 inclinometer (length $\approx$ 20 m), 6 accelerometers (2 systems with 3 axes each), 1 magnetostrictive extensometer and 2 tiltmeters (see Fig. 2a and [15]).

4.3. Creation of a FD model for data-based analysis

In the feasability study the slope model was restricted to static loading and the 2D-space (to avoid too extensive computing time). The idea of static loading was motivated by mass excavation as dominant trigger event and the observation of new balanced states of the slope as reaction. The numerical model was realised with the FD software FLAC3D 2.10 from HCItasca (see [16]). Its basic geometrical and physical features are

- a horizontal extension of approximately 1500 m and a vertical extension of approximately 500 m,
- a discretisation with more than 100.000 finite meshes (mean extension < 10 m),
- a density due to the geological situation and special areas of interest,
- seven physical parameters (e.g. density, Poisson ratio and bulk modulus) per mesh and
- mass extraction as main trigger event.

The numerical model was calibrated using 'trial and error methods', [16]. The material parameters were derived from geological plans and literature and adapted within a realistic range (a priori assumed uncertainty of 30 %). For this purpose calculated displacements in selected slope points (see Fig. 3) were compared with independent control measurements (e.g. precise levelling or GPS).
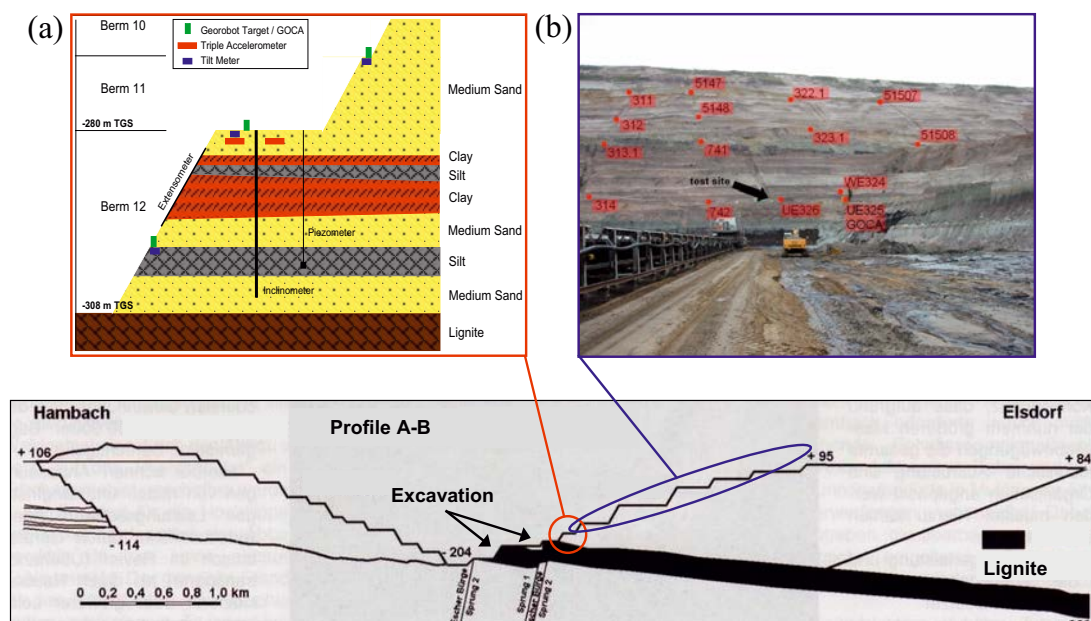

Figure 2: Test-slope with (a) geotechnical and (b) geodetic monitoring system ([14]).
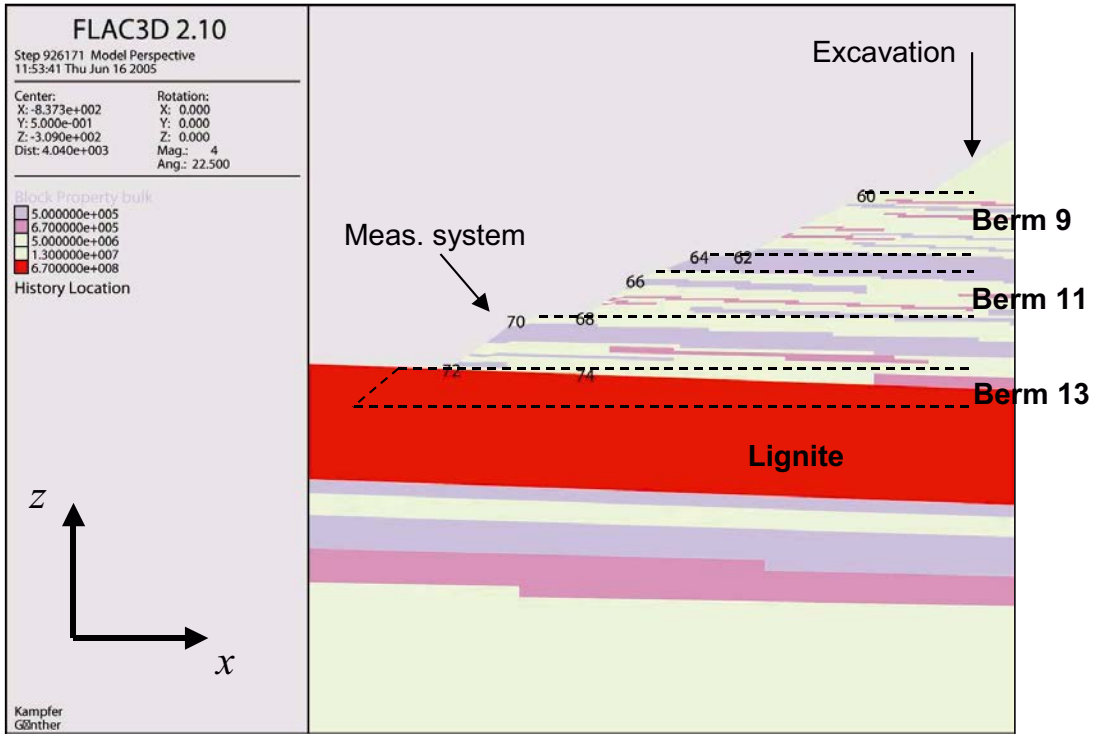
Figure 3: FD modelling of the test-slope with FLAC3D: slope points 60 to 74 (ref. to [16]).

In Fig. 3 it is also shown how the $[x, z]$ slope model coordinate system is defined. The x-axis defines the horizontal and the z-axis the vertical direction of the 2D slope. Positive x shows into the slope and positive z up to the zenith.

4.4. Calculation results from the calibrated FD model

In the years 2004 to 2005 trigger events were performed at the test slope by 13 excavation steps from top ground surface (tgs) down to the lignite layer (Fig. 2 and 3). The resulting berms have a mean height of approximately 30 m. In the FD model this mass extraction could be quantified by the reduction of the associated meshes and the calculation of the resulting states of equilibrium.

In Fig. 4 the comparison between predicted and monitored displacements is exemplarily shown for the sum of excavation steps 12 and 13 as trigger input. The steps were performed in mid and end of January 2005.

Control point 325 is situated on the top of berm 12 next to the geotechnical monitoring system and was not used for the calibration process. Its

horizontal (Δx) and vertical (Δz) displacements are measured by a GPS receiver from the GOCA monitoring system (see Sec. 4.2). The available time series (original coordinates from GOCA output, not smoothed) start in January 2005 and end in February with a scanning rate of $\Delta t = 10$ min. They are reduced to a (nearly) static state of the slope at the beginning of January just before excavation step 12.

As a reaction to steps 12 and 13 the point rises in total with $\Delta z = 5 - 6$ cm and performs a horizontal move 'out of the slope' with $\Delta x \approx -11$ cm. In both cases the movements are not finally stabilised but show an asymptotic behaviour.

The related static FD prediction of the total displacements at the end of step 13 shows a contradiction of only $dx \approx 1$ cm in the horizontal and $dz \approx 2 - 3$ cm in the vertical component. Predicted and real behaviour of point 325 are fitting together within a range of some cm. These first results can be stated as promising for further attempts to predict the at least normal behaviour of the slope.
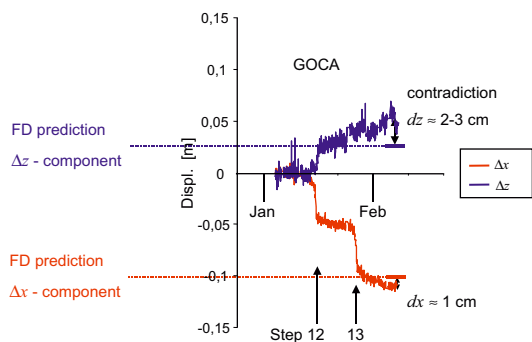
Figure 4: Comparison of FD prediction and measured displacements in control point 325.

Using the calibrated FD model it is also possible to simulate realistic failure events in instable slope areas ([16] and [17]). One simulation result is presented in Fig. 5. In the simulation the real gradient angle of the excavated berms was increased with 15 %.

Fig. 5 shows the static calculation of the total horizontal (x) and vertical (z) movement of control point 70 (also on top of berm 12) as reaction to the 13 excavation steps beginning at tgs with 0 m. The failure event starts when the lignite layer is reached by excavation and causes a horizontal move out of the slope and a vertical sagging with $\approx 1,5\,m$ in each direction. Finally the movement stabilises itself again. These results could be verified by interviews with local experts which classified this combination of slope gradient, excavation quantity and mass movement as realistic.
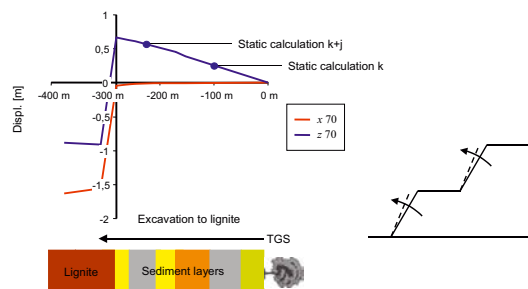

Figure 5: Simulation of a failure event: local slope sagging in control point 70 (ref. to [16]).

This example shows the potential of the calibrated numerical slope model for the simulation of failures in case of critical triggers and how to perform prevention measures. In this case one possible solution could be a reduction of the gradient angle.

5. Conclusions

A multi-scale monitoring and analysis system for detecting landslides was developed and combined with an efficient alarm system. GIS integrated geological evaluations of remote-sensing data and geometrical analysis of monitoring data with fuzzy systems enables to define the more active zones of the landslides and to install there wireless sensor networks. This concept supports cost-saving solutions especially in regions with a larger number of landslides. A concept for a knowledge-based alarm system was presented which aims to produce alarm levels by the interaction of a data- and a knowledge-based system analysis. The data-based system analysis is based on FD-modelling of the slopes which was successfully applied within the OASYS project. It can be stated that the cooperation with partners from different disciplines was the right way to investigate this complex topic. Of course further investigations are necessary to develop a functional alarm system; according research proposals are currently in evaluation.

References

[1] *Kahmen, H., W. Niemeier:* OASYS – Integrated Optimization of Landslide Alert Systems. In: Sansò, F. (Ed.): A Window on the Future of Geodesy, Sapporo, Japan, June 30 – July 11, 2003. International Association of Geodesy Symposia, Volume 128, Springer, Berlin-Heidelberg, 2005, pp. 101-104.

[2] *Theilen-Willige B.:* Seismic hazard localization based on lineament analysis of ERS- and SIR-C- radar-data of the Lake Constance Area and on field check. Proceedings of the European Conference on Synthetic Aperture Radar. 25-27 May 1998 in Friedrichshafen, VDE-Verlag, Berlin, 1998, pp. 447-550.

[3] *Theilen-Willige B.:* Detection of Landslides and High-Risk Areas. Contribution to the Final Report of OASYS, 2006, pp. 11-24.

[4] *Welsch W. (Hrsg.):* Deformationsanalysen '83 – Geometrische Analyse und Interpretation von Deformationen Geodätischer Netze. Beiträge zum Geodätischen Seminar. Schriftenreihe Wissenschaftlicher Studiengang Vermessungswesen Hochschule der Bundeswehr München, Heft 9, 1983.

[5] *Haberler M.:* Einsatz von Fuzzy Methoden zur Detektion konsistenter Punktbewegungen. Geowissenschaftliche Mitteilungen, Schriftenreihe der Studienrichtung Vermessungswesen und Geoinformation der Technischen Universität Wien, Nr. 73, 2005.

[6] *Kahmen, H., A. Eichhorn, M. Haberler-Weber:* A Multi-Scale Monitoring Concept for Landslide Disaster Mitigation. In: Proceedings of the IAG, IAPSO and IABO joint Conference Dynamic Planet 2005, Cairns, International Association of Geodesy Symposia, Volume 130, Springer, Berlin-Heidelberg, 2007, pp. 769 – 775.

[7] *Bicanic, N.:* Discrete Element Methods. In: Stein, de Borst, Hughes Encyclopedia of Computational Mechanics, Vol. 1. Wiley, 2004.

[8] *Ames, W.F.:* Numerical Method for Partial Differential Equations. Academic Press, New York, 1977.

[9] *Gallagher, R.H.:* Finite-Element-Analysis. Springer, Berlin-Heidelberg-New York, 1976.

[10] *HCItasca:* http://www.itascacg.com/, last access 03/07.

[11] *Wittke W., C. Erichsen:* Stability of Rock Slopes. Geotechnical Engineering Handbook, Ernst & Sohn, 2002.

[12] *Jäger, R., S. Kälber, M. Oswald, M. Bertges:* GNSS/GPS/LPS based Online Control and Alarm System (GOCA) - Mathematical Models and Technical Realisation of a System for Natural and Geotechnical Deformation Monitoring and Analysis -. In: Kahmen, H., Chrzanowski, A. (Eds.): 3rd IAG Symp. on Geodesy for Geotechnical and Structural Engineering / 12th FIG Symp. on Deformation Measurements, Baden, 2006.

[13] *Kahmen, H., H. Suhre:* Ein lernfähiges tachymetrisches Vermessungssystem zur Überwachung kinematischer Vorgänge ohne Beobachter. ZfV 108, 1983, pp. 345–351.

[14] *Prader A.:* Measurement Methods and Sensors for Geodetic and Geotechnical Information. OASYS Deliverable 1.3, Vienna, 2003.

[15] *TU Braunschweig, TU Vienna, Geodata GmbH:* Instrumentation, Data Analysis and Decision Support for Landslide Alarm Systems. Report of the project OASYS, 2004.

[16] *Kampfer, G.:* Zusammenfassung der Ergebnisse der geotechnischen Modellierung der RWE Testböschung. Report of the project OASYS, 2005.

[17] *Eichhorn, A.:* Geomechanical Modelling as Central Component of a Landslide Alert System Prototype: Case Study 'Opencast Mine'. Proceedings XXIII FIG Congress, Munich, TS29.4, 2006.

Contact

Dr. Michaela Haberler-Weber, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-mail: Michaela.Haberler-Weber@tuwien.ac.at
Dr. Andreas Eichhorn, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria. E-mail: Andreas.Eichhorn@tuwien.ac.at
Prof. Heribert Kahmen, Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstr. 27-29, 1040 Vienna, Austria.
E-mail: Heribert.Kahmen@tuwien.ac.at

vgi

Value-added services for river information systems

N. Kühtreiber, M. Jandrisits, R. Pfliegl and B. Hofmann-Wellenhof

Abstract

A River Information Service (RIS) for the European east-west transportation axis (Rhine-Main-Danube) will be installed in the near future. The system will rely on a conventional transponder-based network using radio links embedded in a wire-based communication network on shore. By the control of central management facilities tactical traffic information is provided. On-board a ship the positioning data of the vessel itself and the positioning data of the surrounding ships are compiled to generate a tactical traffic image which is displayed by using an Electronic Chart Display and Information System (ECDIS). The positioning technology used for traffic management systems on inland waterways has to fulfill high requirements regarding accuracy (better than 5 m) and availability. For positioning, DGPS with local reference stations along the rivers will be used.

The advantages of the tactical traffic image for the user are, e.g., the increased transport safety by monitoring and guiding the flow of traffic, the support of the steersman by an extensive image of the traffic situation, and the possibility to monitor the transport of dangerous goods. Commercial inland navigation benefits too, by minimizing the costs of the non-productive time at locks and in harbors. The fuel consumption can be reduced through the planning of trips with continuous speed and by simplifying the logistic flow of the inland waterway traffic.

Around 2008-2010, an upgrade of the existing technology will be possible. In order to identify the best-practice method covering the needs of commercial (transport management) as well as nautical and administrative (traffic management) requests in inland waterway navigation, the forthcoming positioning technologies are evaluated.

The project EPRIS (Evaluation of Positioning technologies for the generation of value-added services in the environment of River Information Systems) focused on future systems (EGNOS, Galileo, etc.) to identify migration paths to new positioning and navigation technologies in the system concept of today's RIS.

Kurzfassung

Entlang der europäischen Ost-West Transportachse (Rhein-Main-Donau) wird in naher Zukunft ein River Information Service (RIS) eingerichtet. Dieses baut auf einem konventionellen transponderbasierten Netzwerk auf und nutzt Funkverbindungen zur Datenübertragung neben der landseitigen leitungsbasierten Kommunikationsinfrastruktur. Sogenannte taktische Verkehrsinformationen werden unter der Kontrolle einer zentralen Leitstelle bereitgestellt. An Bord des Schiffs werden Positionsdaten und Schiffsinformationen des eigenen und der umgebenden Schiffe zu einem taktischen Verkehrsbild zusammengefügt und mittels Electronic Chart Display and Information System (ECDIS) auf einem Bildschirm ausgegeben. Die Positionierungstechnologie, die für Managementsysteme auf Binnenwasserstraßen eingesetzt wird, muss hohe Anforderungen an die Genauigkeit (besser als 5 m) und Verfügbarkeit erfüllen. Derzeit wird DGPS mit lokalen Referenzstationen entlang der Wasserstraßen verwendet.

Durch die Nutzung des taktischen Verkehrsbilds ergeben sich eine gesteigerte Verkehrssicherheit aufgrund der möglichen Verkehrsüberwachung, die Unterstützung des Steuermanns durch eine ausführliche Darstellung der aktuellen Verkehrssituation und die Möglichkeit, Gefahrguttransporte überwachen zu können. Wartezeiten an Schleusen und in Häfen werden minimiert, wodurch sich enorme Vorteile für die kommerzielle Binnenschifffahrt ergeben. Zudem kann der Treibstoffverbrauch, durch eine effektive Fahrzeitplanung mit kontinuierlicher Geschwindigkeit und eine Vereinfachung des logistischen Verkehrsablaufs auf der Binnenwasserstraße, gesenkt werden.

Um 2008-2010 wird eine Verbesserung der existierenden Technologie möglich, indem eine Best-Practice Methode identifiziert wird, die die Anforderungen des kommerziellen, nautischen und administrativen Transportmanagements an die Binnenschifffahrt abdeckt. Hierfür werden u.a. die verfügbaren und künftigen Positionierungstechnologien evaluiert.

Das Projekt EPRIS (Evaluierung von Positionierungstechnologien zur Generierung von Mehrwertdiensten im Umfeld von River Information Services) zielte auf zukünftige Systeme (EGNOS, Galileo, etc.) ab um diese neuen Technologien in das derzeitige Systemkonzept von RIS zu integrieren.

1. Introduction

Recent traffic forecasts [1] for goods transport on the Trans European Network (TEN) show an average annual increase of 3-5%. This means doubling the overall goods transport volume within the time frame until 2015. In [2] a more detailed analysis is given showing that the average increase on the highly sensitive east-west transport axis will be enormous. One reason is the economic growth of the Central Eastern European Countries which is higher-than-average. Further analysis show that the major part of this traffic increase will occur on the Trans European Road Network. The road transport segment has been growing continuously during the last 20 years and is cumulating at about 80% of the overall transport volume today. A critical saturation of the road network has been reached causing accidents, delays, and a loss of 300 billion € every year on the European gross domestic product. Significant effort is required to keep the traffic flowing.

The transportation capacity of many inland waterways especially on the river Danube is used at a very low extent. The traffic on the European rivers can be increased significantly over the next years (e.g. see [1]). As the road network reaches its economical and ecological limits a shift of transportation capacities to inland waterways will make the situation easier. Another positive effect of this possible shift is that CO_2 emissions will be cut significantly (cf. Figure 1).
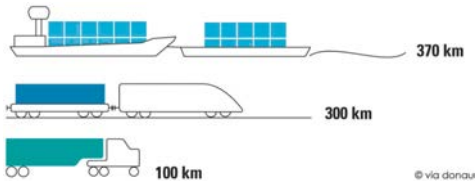
Figure 1: Comparison of the energy consumption of different common carriers.

One problem connected with the river Danube is the occasionally possible big change of the water-level in the upper part of the river. Other problems for the traffic on inland waterways are the low reliability for travel time and ETA (estimated time of arrival) as well as the poor 'status information' about the inland navigation overall. Furthermore, increasing transport operations on the Danube will increase the probability of accidents, waiting periods at locks, and endanger the ecologic balance of the river system.

Most recent technological concepts within RIS can provide online traffic information and position-based nautical support for inland navigation and will transform the inland waterways to reliable and competitive transport ways.

2. Donau River Information Services (DoRIS)

RIS is defined as a concept for harmonized information services to support traffic and transport management in inland waterway navigation including interfaces to other transport modes.

In autumn 1999 the European research project INDRIS (Inland Navigation Demonstrator for River Information Services) showed the technical feasibility of a concept for RIS, on a demonstration site on part of the Austrian Danube (see Final Report of INDRIS).

Due to the INDRIS project results, the Austrian ministry for transport, innovation, and technology decided the stepwise implementation of RIS in Austria. Together with a subsidiary of the ministry, the Danube Transport Development agency (via donau), an INDRIS-based system concept has been developed and a European tendering procedure was initiated. According to the tendering documents a consortium of private companies started to implement a RIS test-center on a limited section of the Danube in Austria which was finished in the third quarter of 2002.

This test-center is used to verify and improve the chosen technical approach and should further serve as an international reference system where all parties involved in the European RIS process get the possibility to perform studies in a working RIS environment.

2.1. Concept

Figure 2 shows the concept of the DoRIS which is based on a fast short-distance radio communication between the involved ships and to land-side control stations exchanging ship identification, position, time, and other transport-related information. A RIS enables the generation of a Tactical Traffic Image (TTI) supporting nautical operation on-board the ship and traffic monitoring by the responsible federal administration at a national traffic control center.

In its concept the river information system DoRIS has two main objectives:

1. to provide an overview of the actual traffic situation and the traffic control by a TTI, which represents the nautical component.
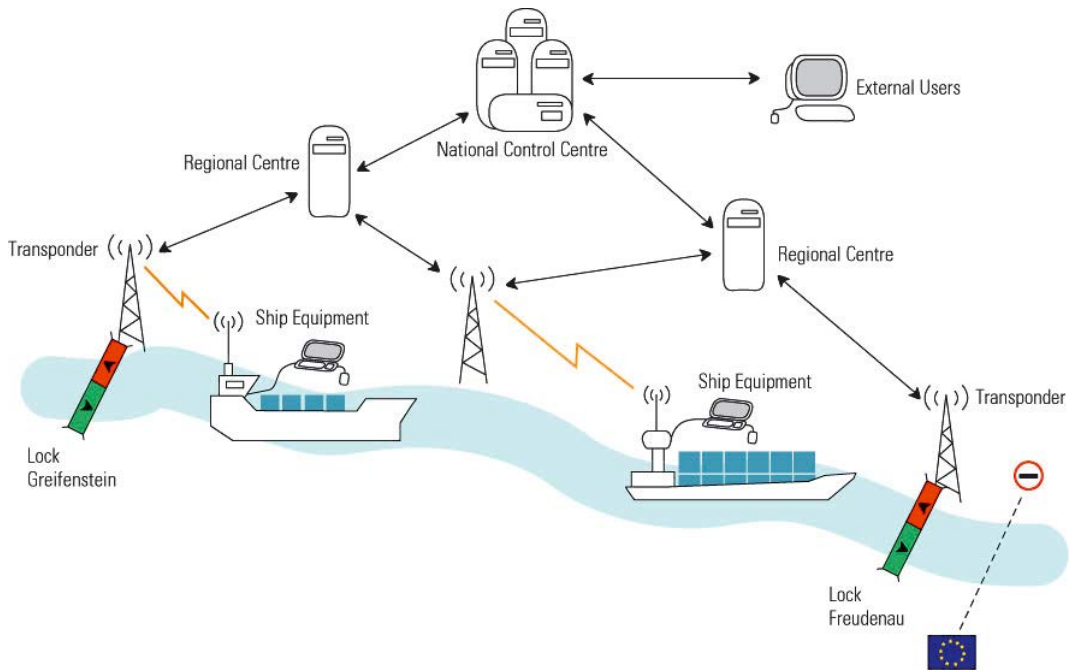
Figure 2: Concept of the Donau River Information Service (DoRIS).

© via donau

2. to provide online traffic information by an information system in order to fulfill the needs of management tasks in inland waterway traffic operation. This can be seen as the navigational component.

2.2. Transponder

One main element of the DoRIS system are AIS (Automatic Identification System) transponders. Within DoRIS, the transponders exchange positioning information and other relevant information via VHF. The transponder is connected to a notebook where a TTI is produced (see below). Each vessel needs to be equipped either with a fixed or a removable transponder. Additional transponders are installed ashore to provide local and governmental authorities with an actual traffic image. Traffic data are processed and stored in a central data base and allow the reconstruction of accidents and the workout of statistics if needed.

2.3. Positioning component

Positioning technologies used for traffic management systems on inland waterways have to fulfill different requirements regarding accuracy, availability, integrity and actuality. The evolution of the needs for navigational purposes as well as the needs for producing value- added services for the traffic and transport management on inland

waterways will be discussed in detail below. At the moment, DGPS with local reference stations along the river will be used in the DoRIS system.

In the project EPRIS emphasis were led on the forthcoming GNSS systems like EGNOS and GALILEO and their contribution to the generation of value-added services.

2.3.1. EGNOS

The regional satellite-based augmentation system (SBAS) EGNOS emits augmentation signals via two geostationary Inmarsat-III satellites and the geostationary Artemis satellite. The augmentation signal is designed in a GPS-like manner and carries differential corrections, integrity data, and ionospheric corrections.

One advantage of EGNOS are lower horizontal and vertical positioning errors compared to GPS. GPS standalone provides 13 m and 22 m for the horizontal and vertical positioning accuracy (95%), respectively (see [6]). With GPS augmented by EGNOS this accuracy can be improved to 1 - 3 m for the horizontal and 3 – 5 m for the vertical component. Even more important is the possibility to get information on the integrity of GPS and EGNOS itself and implied alarm limits for positioning with a time to alarm of better than 6 seconds.
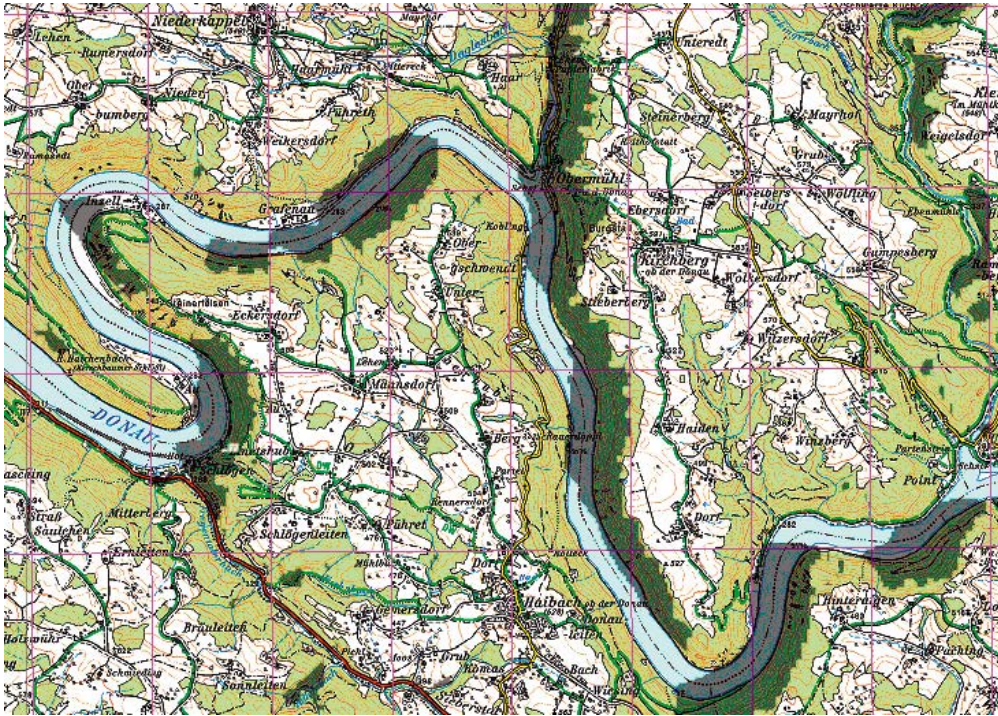
Figure 3: Visibility of IOR along the Danube near the German-Austrian border.
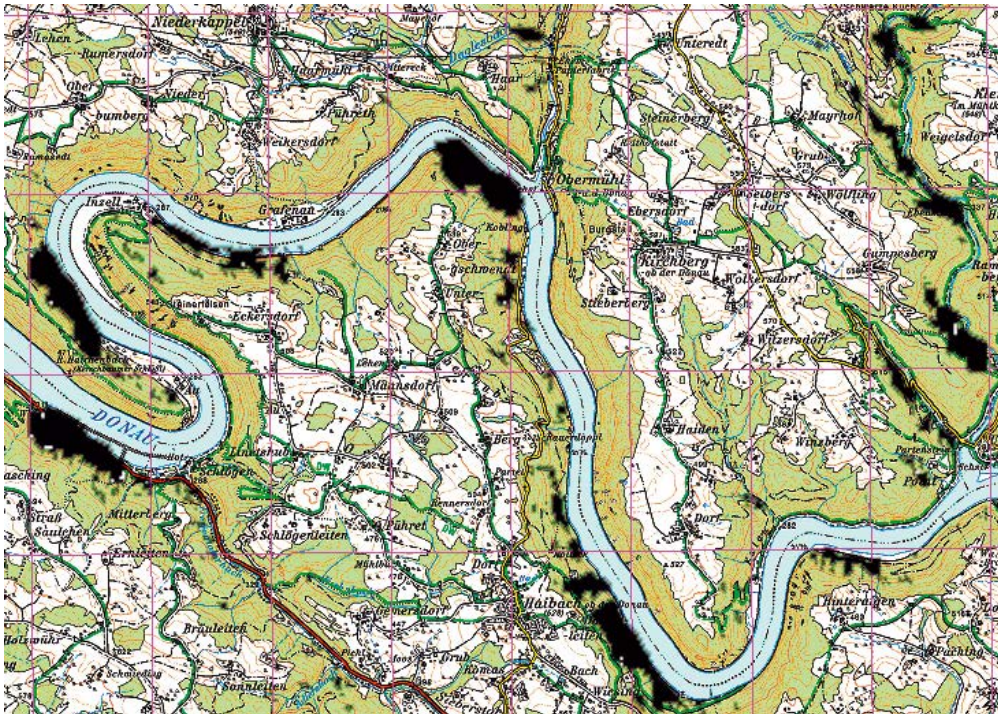*The dark gray pattern marks obstructions of the EGNOS signal.*

Figure 4: Visibility of AOR-E along the Danube near the German-Austrian border.
The dark gray pattern marks obstructions of the EGNOS signal.

The receiving of EGNOS signals needs a line of sight to one or more of the geostationary satellites (AOR-E, IOR, and Artemis). Obstacles (dense vegetation and high rocks ashore, brides etc.) may cut off the connection. As the satellites are placed over the equator they are "visible" in the southern part of the sky regarding locations in the northern hemisphere. This is the reason why in deep valleys running from east to west problems may occur. Figure 3 and Figure 4 show a simulation of the visibility conditions of AOR-E and IOR. All together the approach using EGNOS for positioning is promising. Nevertheless applications requiring redundant EGNOS signals for safety reasons will need additional augmentation measurements.

During the GALEWAT (Galileo and EGNOS for Waterway Transport) project [3], a concept for the provision of the EGNOS correction signals by the use of the terrestrial AIS infrastructure was shown. This concept [3] and [5] pointed out clearly the possibility to provide the EGNOS corrections even in regions, where the EGNOS signal can not be received directly from the geostationary satellites. Nevertheless, this system concept is not realized until now.

2.3.2. GALILEO

The forthcoming European satellite navigation system GALILEO is a civilian system which will provide accurate positioning information on a worldwide basis. GALILEO's navigation and positioning services will be interoperable with GPS and GLONASS. The accuracy of the real-time positioning information will be in the meter range. The availability of selected services will be guaranteed. Also the integrity is high. The GALILEO system will consist of 30 satellites in medium earth orbits (MEO) and will be operable around 2012 presumably.

The foreseen GALILEO-services (open service, commercial service, safety-of-life service and public regulated service) will be complemented by local GALILEO components if higher user requirements concerning accuracy, availability, integrity, and continuity are needed. Further details can be found in [8].

2.4. TTI

By the control of central management facilities, tactical traffic information on a GIS-based application is provided. On-board a ship the position data of the vessel and the position data of the surrounding ships are compiled to generate a TTI which is displayed by an ECDIS (see Figure 5).

The TTI is a function creating a graphical impression of the current traffic situation in a certain area. The TTI can be used in control centers on shore and onboard of an inland waterway vessel. On shore the TTI will support a new method of traffic guidance through monitoring traffic using AIS, GPS and inland-ECDIS technology. Onboard of inland vessels the TTI will become the navigational display using the same technologies. This image is further enhanced by reports from vessels to the information center where an overview of vessels active in the respective area is generated. The AIS technology will improve the TTI in such a traffic center since ships can be identified and additional information is available.

3. Economical and ecological aspects

The deployment of RIS will meet both ecological and economical aspects as well as governmental/administrative requirements.

According to [2] and [7], the main advantages are:
- Increased safety of transport operation.
- Improved catastrophe management and calamity abatement.
- Minimization of the use of the waterway in sections declared as protected areas, e.g., the National Park "Donau-Auen" by commercial shipping.

Economic advantages:
- Improved logistics operation by more efficient use of transport means.
- Improved transshipment process between road/rail and vessels.
- New infrastructure within the information technology.
- Reduced danger of accidents and therefore reduction of follow-up costs.
- Fuel and time savings.
- Integration of an information service for commercial users like ship owners, brokers, ports and tourist businesses, etc.

Administrative advantages (governmental level):
- Common traffic guidance system with integration of the corresponding public authorities, e.g., navigation, customs, defense, and hydraulic engineering authorities.
- Services for the water supply and distribution as well as for the regional planning.
- Fast and efficient information service for skippers ('notices to skippers').
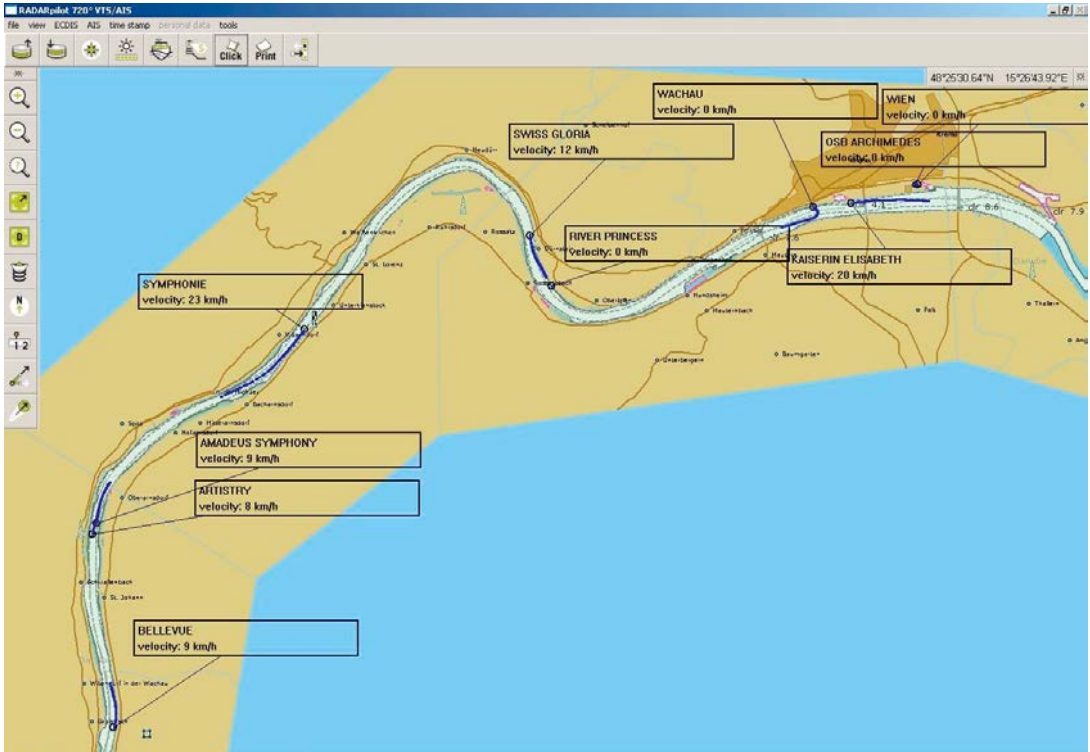
Figure 5: Example of a TTI.

With RIS, the responsible Austrian federal government expects both, improved safety for transport operation on the Danube and improved competitiveness of inland waterway transport as part of modern supply chain management. This will encourage the modal shift of cargo transport from road to inland waterway on the highly saturated east-west road transit transport axis in Europe (Amsterdam, Rotterdam and Antwerp to Black Sea).

4. Value-added services

[1] focuses on future systems (EGNOS, Galileo, etc.) to identify migration paths to new positioning and navigation technologies in the system concept of RIS.

Value-added services indicate services improving basic GNSS data by increasing the quality (accuracy, integrity, availability and continuity) and combining it with additional information (navigation, communication etc.).

The signal in space (SIS) of positioning and navigation systems is free to everyone (neglecting restrictions for the moment). In general the knowledge of an accurate and actual position represents no added value. Normally an added value is achieved if the position information is combined with other information sources (geographical information systems, traffic information etc.) and communication tools.

4.1. Value added chain

Participants of a value-added chain provide additional services for the improvement of the SIS from pure data to the end user. The steps from the generated SIS to an end user application can be described by the example of a traffic service. The basic SIS data is supplemented by GIS-data provider information. Afterwards object data provider and traffic data provider information are added. Before the improved data is delivered a data integration is done. One group to be concerned with this topic is the Research Group Geoinformation at the Technical University Vienna (see e.g. [4]).

4.2. Components of value-added services

Value-added services are generated by several processing steps where the basic information is

supplemented by other data sources. At the beginning the SIS and a number of other data basis (knowledge data basis, GIS etc.) are needed. The data collected from the different sources are combined and reprocessed thus improving the contents of the data.
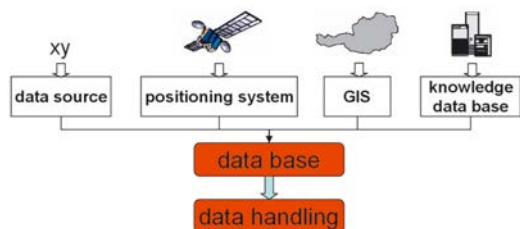

Figure 6: Components of value-added information services step 1.

For the estimation of the added value it is of fundamental importance to know the requirements of the end user for the application environment.

Figure 7: Components of value-added information services step 2.

As a next step a user-specific processing, visualization and storage is done. Finally the value-added service is provided to the user. All components of added values (shown by the red color) contribute to the refinement of the data and thereby to value-added services. The acceptance and perception of value-added services is directly influenced by the time needed for processing, delivering and storing the data thus influencing the willingness to pay for value-added services.
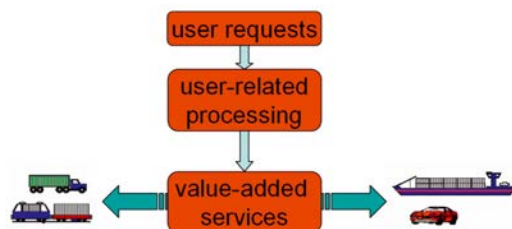

Figure 8: Components of value-added information services step 3.

5. Evaluation of positioning technologies

Within [1] an evaluation of positioning technologies for applications and value-added services in the transport and traffic management of inland

waterways was carried out. The following parameters and assessment criteria were used.

Absolute accuracy: stands for the needed accuracy of the positioning data given in meters.

Spatial and temporal availability: describes the availability of the particular positioning technology needed for the chosen application. Assessment criteria are: low, medium, and high.

Integrity: stands for the need of integrity of the particular positioning technology. Assessment criteria are: none, medium, and high.

5.1. ETA

The ETA is the basis of improved logistic operations for the optimization of the transporting process as a whole. Knowing the ETA will simplify and accelerate the inland waterway navigation process through the improved access to resources of ports and terminals. The time for loading and unloading can be kept as short as possible. ETA has its advantages also in managing operations at locks, supporting customs authorities in planning their resources (personal). Concerning positioning aspects an accuracy of better than 100 m is needed. Integrity and availability need to be high.

5.2. Travel plan

In order to establish a travel plan for ships additional information besides the ETA is needed. Fairway information of the actual and predicted water-levels of the waterway, port cycles, stream velocities etc. are provided. The ecological effect of this value-added service is evident through fuel and time savings. The information will be supplied on a regional basis. The integrity and availability has to be high. The positioning accuracy needs to be 1–50 m.

5.3. Flow of goods

This service incorporates all the information needed for improving the planning and guiding of logistic operations. Mainly information about the cargo and the vessel are incorporated. While for this service no integrity is needed, the availability needed is middle and the accuracy of the position needed is about 100 m.

5.4. Statistics

The information collected in a RIS may be analyzed and described by statistical parameters. Statistics supports numerous management decisions of operational and logistic centers,

public authorities, federal administrations, locks, ports, terminals, ambulance, disaster management etc.

5.5. Traffic information

Traffic information stands for all data needed for guidance and navigation. The information is provided on an inland-ECDIS together with the electronic inland navigation. Two operational modes of the ECDIS exist: in the navigation mode the inland-ECDIS is used for navigating the ship (additional radar information is provided). The information mode of the inland ECDIS is used for information purposes only and no additional radar information is provided. As the traffic information is used for navigation the accuracy needed ranges from 0.5 to 5 m [2] with a high availability and integrity.

5.6. Infrastructure

The service provides information on the actual infrastructure on a regional basis. Fairway information, water-levels, stream velocities etc. are contained. The accuracy needed is 100 m with a high availability and middle integrity.

5.7. Traffic control

The traffic control supports the traffic and transport management on inland waterways. The basis for the traffic guidance are accurate positioning data of all vessels. The traffic guidance uses the TTI and STI. The accuracy needed for the positioning data is high and ranges between 0.5 and 5 m. The availability and integrity need to be high.

5.8. Locks

The positioning data of vessels is used to optimize the time needed for passing through a lock. In other words the service allows to allocate time slots for the ships at a lock in advance. It supports the lock-keeper during the planning stage. The accuracy needed for this service is about 30 m.

5.9. TTI

The TTI is a service creating a display of the current traffic situation (around a ship, in a limited area or the whole river). The advantages of monitoring the traffic are the increased safety of the transport operations and the improved traffic management. In a TTI the different information provided by a RIS are combined: radar data, AIS information, fairway information, water level (actual and predicted), stream velocities. The visualization of the TTI is provided via ECDIS. As the TTI supports the skipper in his responsibility for navigation decisions the positioning accuracy needed is high (between 0.5 and 5 m). Additionally availability and integrity requirements are high. The TTI can be used on shore or onboard of a vessel.

5.10. Automatic charging

With the help of position information automated port and canal prizing can be done. The needed accuracy is 100 m, the availability is high and with middle integrity.

5.11. Safety applications

Applications related to the safety of transport operations on waterways are divided into active and passive safety services. Active safety services are used for the prevention of accidents and improved catastrophe management. Passive safety services support the reconstruction and examination of safety-critical events. The availability as well as the integrity of the position information needed for this service are high.

6. Conclusion

The benefits of value-added services of a RIS in transport and traffic management will improve transport efficiency, will help to increase safety and security on inland waterways and will support the skipper in his actual nautical decisions. Cost-effectiveness, the acceleration and simplification of logistical processes will make the waterways a competitive alternative to rail and road transport systems. By the use of value-added services inland navigation will become reliable, flexible and more planned in future thus increasing its attractiveness.

Acronyms and Abbreviations

AIS	Automatic Identification System
DGPS	Differential Global Positioning System
DoRIS	Danube River Information Services
ECDIS	Electronic Chart Display and Information System
EGNOS	European Geostationary Navigation Overlay Service
EPRIS	Evaluation of Positioning technologies for the generation of value-added services in the environment of River Information Systems
ETA	Estimated Time of Arrival
GIS	Geographic Information System
GLONASS	Global Navigation Satellite System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System

INDRIS	Inland Navigation Demonstrator for River Information Services
RIS	River Information Services
SBAS	Satellite-Based Augmentation System
STI	Strategical Traffic Image
TEN	Trans European Network
TTI	Tactical Traffic Image
VHF	Very High Frequency

References

[1] *EUDET – Evaluation of the Danube Waterway as a Key European Transport Resource, 1996-1999 (EU 4th Framework Programme, via donau)*

[2] *EPRIS – Evaluation of Positioning Technologies for the Generation of Value Added Services in the context of River Information Services, Final Report dated 15.02.2005, via donau, TeleConsult Austria GmbH, Graz University of Technology – Institute of Navigation and Satellite Geodesy, Nauticast Navigationssysteme GmbH (Austrian Space Programme 1st Call)*

[3] *GALEWAT – Galileo and EGNOS for Waterway Transport, Final Report (2005), via donau, INOV, TeleConsult Austria, Kongsberg Seatex,*

[4] *Krek, A., and Frank, A.U.(2000):* 'The Economic Value of Geo Information'. In: Geo-Informations-Systeme – Journal for Spatial Information and Decision Making, 13 (3), pp: 10-12.

[5] *Pfliegl R, Jandrisits M, De Mateo Garcia JC, Hofmann-Wellenhof B, Aichhorn K, Abwerzger G (2004):* Introduction of EGNOS in River Information Services, Proceedings of the 46th International Symposium ELMAR-2004, June 2004.

[6] *Department of Defense (2001):* Global Positioning System standard positioning service performance standard. Available at www.navcen.uscg.gov.

[7] *Final Report of INDRIS, Assessment on the basis of INDRIS of the Rhine, Seine and Danube demonstrator Guidelines and recommendations for River Information Services.*

[8] *Hofmann-Wellenhof B, Legat K, Wieser M (2003):* Navigation: principles of positioning and guidance. Springer, Wien.

Contact

Ao. Univ. Prof. Dr.techn. Norbert Kühtreiber, Institute of Navigation and Satellite Geodesy, Graz University of Technology, Steyrergasse 30, A-8010 Graz, Austria.
E-mail: norbert.kuehtreiber@tugraz.at.

Dipl.- Ing. (FH) M. Jandrisits, previously: via donau, Donau-City-Strasse 1, A-1220 Vienna, Austria, now: ASFINAG Verkehrstelematik GmbH, Klingerstraße 12, A-1230 Vienna, Austria. E-mail: Marko.Jandrisits@asfinag.at.

Dipl.- Ing. Dr.techn. R. Pfliegl, previously: via donau, Donau-City-Strasse 1, A-1220 Vienna, Austria, now: AustriaTech, Donau-City-Strasse 1, A-1220 Vienna, Austria.
E-mail: reinhard.pfliegl@austriatech.org.

Univ. Prof. Dr.techn. B. Hofmann-Wellenhof, TeleConsult Austria GmbH, Schwarzbauerweg 3, A-8043 Graz, Austria. E-mail: howe@teleconsult-austria.at.

Wayfinding for Public Transportation Users as Navigation in a Product of Graphs

Andrew U. Frank

Abstract

Navigation systems help car drivers and pedestrians to find their way in unknown environments; probably the most widely used GIS application. GIScience investigates the theoretical foundations for geoinformation. This article describes a series of recent investigations focusing on finding the shortest path in a network represented as graph. To aid pedestrians in wayfinding and using public transportation systems effectively, they need information not only for their spatial decisions (e.g., also to buy a ticket). These business aspects can be represented as a second graph they navigate. The article shows a novel solution to merge two state-transition graphs using category theory. The resulting formula can be used to program simulation systems or wayfinding programs.

Kurzfassung

Navigationssysteme helfen Autofahrern und Fußgängern ihren Weg zu finden. Sie sind wohl die am weitesten verbreiteten GIS Anwendungen; die Geoinformationswissenschaft legt die theoretischen Grundlagen dazu. Es werden hier drei neuere Untersuchungen, die alle die Suche nach dem kürzesten Weg in einem Graph als zentrale Operation enthalten, vorgestellt.

Fußgänger, die öffentliche Verkehrsmittel verwenden wollen, brauchen nicht nur Anweisungen für ihre Bewegungen im Raum, sondern auch Hinweise auf Benützungsregeln der Verkehrsmittel (Ticket kaufen, abstempeln etc.). Diese Regeln können ebenfalls als Zustands-Übergangsgraph dargestellt werden. Es wird hier eine neuartige mathematische Formulierung für die Verbindung der zwei Zustands-Übergangsgraphen angegeben, die auf Kategorientheorie beruht und zur Programmierung geeignet ist.

1. Introduction

Geodesy makes important contributions to the construction of geoinformation systems (GIS). It is therefore not surprising that some of the pioneering geoinformation researchers were geodesists. Today, universities where a geodesy curriculum is offered have often a chair in geoinformation. The research group for geoinformation in the Institute for Geoinformation and Cartography of the TU Vienna stands in this tradition, with a background in geodesy and surveying and a strong research orientation in geoinformation.

From the technological beginnings of GIS in Canada by Tomlinson in the early 1960s and LIS in Europe (Eichhorn 1979) emerged a need to understand and further develop the foundational theory. Geoinformation Science (GISc) developed, with now several scientific journals and regular conferences, supplementing the large number of applications, national and vendor oriented conferences and magazines. GISc is an interdisciplinary enterprise, bringing together researchers from geography, geodesy, cartography, computer science, mathematics, experimental psychology, linguistics, cognitive science, etc. My contributions over the past 25 years were mostly in putting mathematical theory to use in

GISc. Egenhofer investigated point-set topology to characterize the relations between spatial regions (Egenhofer 1989; Egenhofer et al. 1990; Egenhofer et al. 1991), which opened a field that has now many research publications and his results eventually were integrated in an ISO standard (SQL Multi-Media). Bittner generalized raster based indexing methods, as used in location indices for city maps (indications like "B7", i.e., second row, 7th column), to irregular subdivisions of space (Bittner 1999). This was one of the precursors for the application of rough set theory (Zdzisław 1991) to GIS and is now an active field of research (Bittner et al. 2003).

In this article I address another core theme of GIS, namely, navigation in a discrete network – for example a street network. This can be seen as an application of the geodesic, though not as usual in continuous space, but in discrete space.

The next section introduces navigation as a major GIS application and section 3 abstracts the street-network to a graph. The following section discusses why operations on a graph are seldom expressible as closed, analytical formula. Section 5 reports two investigations using agent simulation to assess the quality of street-network data and signage in a building. The following section

generalizes navigation as graph traversal further, to apply it to connected actions in general, preparing for the application to navigation aids for pedestrians using public transportation systems discussed in the following section. These systems must combine navigation in a street-network with the fulfillment of business rules. Section 8 then gives a general, category theory based, solution for this and similar problems. The concluding section suggests future work.

2. Navigation Is a Major Application of GIS

Computerized car navigation systems are probably the most popular application of geographic information systems. They go back to one of the first thorough analyses of an algorithm by computer pioneer Edsger Dijkstra (1959): the task to find the shortest path in a (street) network. This one or the faster A* algorithm (Hart et al. 1968) are at the core of the programs running in millions of car navigation systems and produce navigation advice for drivers all over the world.

Car navigation needs, besides the algorithm and the hardware to run it on, a schematic representation of the street-network, usually provided by surveyors and cartographers. Car navigation systems became feasible with the construction of computerized street-network graphs in the 1970s and 80s by the U.S. Bureau of the Census (Corbett 1979; Witiuk 1988) and later by many European jurisdictions (see SORSA conferences). The viability of this industry depends on the availability of the data collected and maintained and the quality of the data, in particular, the encoding. This led to a research strand investigating the representation of space (Mark et al. 1991) and how people cognize space (Twaroch 2007), but also to research in the application of ontology to geographic space (Smith et al. 1998; Kuhn 2005).

3. An Abstract View of a Street-Network

The famous problem to find a path leading once over all bridges in Königsberg (Kalinagrad) led Euler to abstract graph theory. For present purposes, a graph is a structure, consisting of nodes and edges (many other terms are in use), related by incidences and adjacency.

Node A in Figure 1 is incident with edge 1. The edges 1 and 2 are adjacent because they have node A in common; similarly, the nodes A and B are adjacent, because they have edge 1 in common.
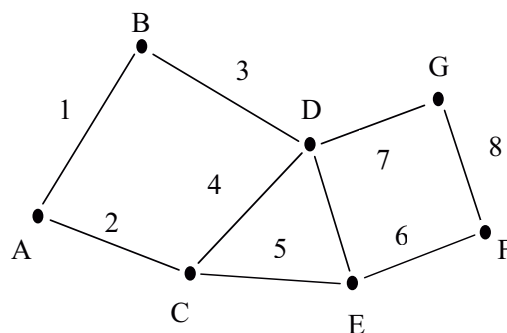

Figure 1: Example street-network graph.

4. Operations on Graphs

In continuous space the shortest path from A to B is called the geodesic (line). In a discrete space like a graph, this problem has in general no analytical solution. Most operations on a graph require some sort of traversal of the graph and at each node or edge traversed an operation is performed. These operations are usually simple: count, find minimum or maximum. Efficient algorithms to determine the shortest path avoid traversing the complete graph and are the more effective the less blind alleys the algorithm chases and the more directly it is targeting the goal; this is the reason why A* is more efficient than Dijkstra's method – if applicable.

For some regular graphs (e.g., complete graphs) analytical solution for some questions are known. Recently random graphs (Barabási 2003) attracted attention, because they allow a statistical analysis. In general, however, solutions for applied graph problems cannot be given by analytical methods and require some sort of traversal, for which we use the conceptual framework of multi-agent simulation. In a computational agent simulation a programmed agent traverses a simulated network (graph) and performs at each node and edge some functions that determine the next action. The behavior of a person navigating in a street-network maps naturally to such a computational agent.

5. Agent Simulation for Navigation Problems

Initially, in our research on navigation problems we approached two quite different questions:
- to assess the affects of data quality of street-network data on navigation (Krek 2002).
- to identify places in a network, where road signs are missing (Raubal 2001).

The first reflects a geodesists concern for the quality of data, whereas the second is exploring novel applications for the geoinformation industry.

5.1. Assess effects of quality of street-networks on navigation

The quality of the data describing the street-network influences the quality of the navigation instructions produced. Before we can address the question to find a relation between data quality and navigation information quality, we have to define data and navigation quality. Quality of street-network data can be – in a first approximation – characterized by

1. the precision of the node positions (coordinate values),

2. the completeness of the connections between the nodes (omissions and commissions).

The quality of the navigation instructions is measured by the length of the path resulting when an agent follows the instructions till the goal is reached. Optimal instructions lead to the shortest path. Incorrect instructions lead to situations where the planned path cannot be executed, for example, because a one-way street cannot be traversed in the intended direction, this results in a longer than optimal path.

Krek constructed a simulation of an agent following instructions to navigate a street-network and she varied the amount of one-way streets not included (omissions) as the dependent variable representing data quality. The result for a part of the street-network of Vienna was:

■ even quite large (meter) position inaccuracies have very little effect on the quality of navigation instructions

■ omissions and commissions of one-way streets up to 25% have nearly no effect on the navigation quality.

5.2. Completeness of signage

Street signs and signs in buildings are crucial in all situations where the visitors do not know and do not intend to learn the environment. Raubal selected an airport situation, where a sufficient number of signs is crucial for passengers to find their departure gate in time. The simulated agents navigate in the graph representing the halls and corridors of the airport and follow only the signs they find at each bifurcation where they must make a decision. Three kinds of errors encountered are:

1. there are no relevant signs at a bifurcation point,

2. the signs do not lead to the desired goal, and

3. following the signs results in an infinite loop.

The simulation permits to test new buildings and the planned signage to avoid confusion and frustration of visitors.

6. Generalization: State Transition Diagram

Navigation in a street-network consists of a series of actions that lead to states; when a bifurcation point is reached, a decision about the next action is taken, following some rule incorporated in the agent. This schema of *action – state – decision – action* is applicable to other problems, where a sequence of steps is necessary to achieve a goal. The correspondence is well known and is widely applied to determine, e.g., the critical path in large projects (CPM). In using a public transportation system, a passenger follows the graph in Figure 2. Realistic business rules may contain bifurcations, when users have options, e.g., buying different types of tickets.
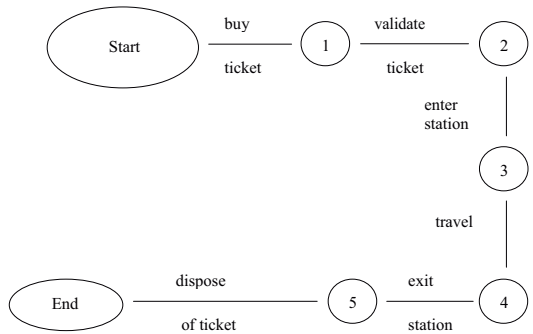

Figure 2: Graph, which describes actions necessary for a legal ride on Vienna's metro.

7. Navigations Systems for Using Public Transportation

The future of large cities is threatened by the overflowing individual vehicle traffic; methods to make public transportation more attractive are urgently needed. We investigated how portable navigation aids (PDA, mobile phone or similar) could be used to assist users. Using a public transportation system requires surprisingly large amounts of information: one needs to know where lines are running and where stops are, but one needs also information about ticketing and where to buy tickets – tobacconist's shop, for example, are not an obvious place! – and what other actions

are required to legally ride a bus or tram (e.g., validation of ticket). Our scientific goal was to understand how to combine spatial navigation with the "navigation" of the state-transition diagram of the public transportation rules (Figure 2 is an extremely simple example); this latter case will be called "business navigation" (Pontikakis 2006). An agent simulation was constructed and demonstrated clearly the typical problems; e.g., how do agents recognize the stop where they have to exit a vehicle?

8. Generalization

The result obtained was only partially satisfactory, because it was based on a single state transition diagram, which integrated spatial and "business navigation". In this section I will show how category theory gives a general solution that allows the integration of two (or more) state transition diagrams formally. Given two navigation problems with state spaces W (wayfinding) and B (business), and actions V and T. The state-transition diagrams are then described by two (partial) functions

w: $W \times V \to W$
b: $B \times T \to B$

A combined state-transition diagram with states S and actions P is described as function

f: $S \times P \to S$.

How to construct f from w and b? The function f, representing the combined state-transition diagram describes the total behavior of the passenger; this is given separately as a set of navigation rules w – which are essentially a graph encoding of the city map with streets and public transportation lines – and the "business rules" b about tickets, how they are obtained, validated, etc.

A fully general solution can be achieved using category theory (Mac Lane 1998), a generalization of universal algebra (Whitehead 1898; Asperti et al. 1991). In this framework, we can write

$S = W \times B$
$P = V + T$

and then

f: $S \times P$

becomes

f: $(W \times B) \times (V + T) \to (W \times B)$

The set of states of the combined wayfinding problem is the product (x) of the states W and B from each problem; each state of the combined problem consists of a pair of states of the individual problem. The set of actions is the union $(+)$ of the actions V and T of the separated problem; a passenger can perform at each decision point a wayfinding or a business action. We construct two trivial operations (isomorphisms) h, k such that

h: $(W \times B) \times V \to (W \times V) \times B$.
k: $(W \times B) \times T \to W \times (B \times T)$

From the given functions w and b we construct functions w' and b'

w': $(W \times V) \times B \to W \times B$
$w' = w \times id \cdot h$
b': $(W \times B) \times T \to W \times B$
$b' = id \times b \cdot k$

id: $A \to A$ denotes in category theory the function that does nothing. The function l is an isomorphism

l: $(W \times B) \times (V + T) \to (W \times B) \times V) + (W \times B) \times T$

because we operate here in the distributive category of sets (Cockett et al. 1992). Combined this gives for f

$f = [w', b'] \cdot l = [(w \times id) \cdot h, \ (id \times b) \cdot k] \cdot l$

This solution translates directly to an executable computer program in a modern programming language that includes second order operations (like [,], $\cdot$, id) and categorical concepts (for example Haskell in Peyton Jones et al. 1999), which produced the expected result and demonstrates the correction of the solution.

Two or more navigation problems in real space and "business logic" space can be directly combined with this formula and we produce a state transition function that can be used for shortest path and similar algorithms. The formula needs to be extended to include possible interactions between space navigation and business rules, e.g., passengers must not enter the station unless they are in possession of a validated ticket. The compact formula together with a description of the interaction may make it possible to analyze the interaction between the two navigation tasks – a probably fruitful field for future research.

9. Conclusions

The development of geographic information systems from a niche technology benefiting from geodesy, surveying, and cartographic theories to an important industry with its own body of theories has led to geographic information science. Assisting humans with navigation and wayfinding tasks is one of the most important applications of GIS – and most likely the one used by more people than any other. The article showed how graph theory and agent simulation are used to analyze the experiences expected for agents with a certain behavior. We have used it to assess the influence of street-network data quality on the quality of navigation instructions and to identify points in a building where existing signage is not sufficient. Such simulations are crucial to identify short comings in signage, which can lead to disasters and loss of life (Hajibabai et al. 2006).

The goal to construct portable navigation aid devices for pedestrians using public transportation systems requires the integration of navigation in space with "navigation in business logic". The novel contribution of this paper is a categorical formula to construct a combined state-transition function from individual ones. The compact representation of the combined formula suggests future research in the interaction between the two wayfinding tasks. Is it possible to identify situations where the individual problems can be solved, but the combined problem – which is the one relevant for the user – forces passengers to wasteful behavior, e.g., forth- and backward movements, or leads even to impossible catch-22 situations (e.g., you need a ticket to enter the station, but tickets are only sold inside)?

The discussion of three practical problems related to the "geodesic" in a discrete graph space and the solution found in a corresponding part of modern algebra could remind geodesists of the advancement achieved by C.F. Gauss in geodesy by extending the mathematics of continuous (real) spaces.

Acknowledgements

References

[1] *Asperti, A. and G. Longo (1991):* Categories, Types and Structures - An Introduction to Category Theory for the Working Computer Scientist. Cambridge, Mass., The MIT Press.

[2] *Barabási, A.-L. (2003):* Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. Cambridge, Massachusetts, Plume.

[3] *Bittner, T. (1999):* Rough Location. Vienna, Austria, Technical University. PhD Thesis.

[4] *Bittner, T. and J. Stell (2003):* Stratified Rough Sets and Vagueness. COSIT 2003, Ittingen, Switzerland, Springer.

[5] *Cockett, J. R. B. and R. A. G. Seely (1992):* Weakly distributive categories. Applications of Categories in Computer Science: Proc. LMS Symp. on the Applications of Category Theory in Computer Science, Durham, Cambridge University Press.

[6] *Corbett, J. P. (1979):* Topological Principles in Cartography, Bureau of the Census, US Department of Commerce.

[7] *Dijkstra, E. W. (1959):* "A Note on Two Problems in Connection with Graphs." Numerische Mathematik (1): 269-271.

[8] *Egenhofer, M. and J. R. Herring (1990):* A Mathematical Framework for the Definition of Topological Relationships. 4th International Symposium on Spatial Data Handling, Zurich, Switzerland, IGU.

[9] *Egenhofer, M. J. (1989):* Spatial Query Languages. University of Maine. PhD thesis.

[10] *Egenhofer, M. J. and R. D. Franzosa (1991):* "Point-Set Topological Spatial Relations." International Journal of Geographical Information Systems 5(2): 161-174.

[11] *Eichhorn, G., Ed. (1979):* Landinformationssysteme. Schriftenreihe Wissenschaft und Technik. Darmstadt, Germany, TH Darmstadt.

[12] *Hajibabai, L., M. Delavar, M. R. Malek and A. Frank (2006):* Agent-Based Simulation for Building Fire Emergency Evacuation. 1st ICA Workshop on Geospatial Analysis and Modeling, Vienna, Austria.

[13] *Hart, P. E., N. J. Nilsson and B. Raphael (1968):* "A Formal Basis for the Heuristic Determination of Minimum Cost Paths." Transactions on Systems Science and Cybernetics SSC 4(2): 100–107.

[14] *Krek, A. (2002):* An Agent-Based Model for Quantifying the Economic Value of Geographic Information. Vienna, Technical University Vienna. PhD Thesis.

[15] *Kuhn, W. (2005):* Geospatial Sematics: Why, of What, and How?

[16] *Mac Lane, S. (1998):* Categories for the Working Mathematician. New York, Berlin, Springer.

[17] *Mark, D. M. and A. U. Frank, Eds. (1991):* Cognitive and Linguistic Aspects of Geographic Space. NATO ASI Series D. Dordrecht, The Netherlands, Kluwer Academic Publishers.

[18] *Peyton Jones, S., J. Hughes and L. Augustsson. (1999):* "Haskell 98: A Non-Strict, Purely Functional Language." from http://www.haskell.org/onlinereport/.

[19] *Pontikakis, E. (2006):* Wayfinding in GIS: Formalization of Basic Needs of a Passenger When Using Public Transportation. Vienna, Technical University Vienna. Doctor of Technical Science.

[20] *Raubal, M. (2001):* Agent-Based Simulation of Human Wayfinding: A Perceptual Model for Unfamiliar Buildings. Vienna, Vienna University of Technology.

[21] *Smith, B. and D. M. Mark (1998):* Ontology and Geographic Kinds. International Symposium on Spatial Data Handling, Vancouver, Canada.

[22] *Twaroch, F. (2007):* Sadbox Geography: How to Structure Space in Formal Models. Vienna, Technical University Vienna. Doctor of Technical Science.

[23] *Whitehead, A. (1898):* A Treatise on Universal Algebra. Cambridge, Cambridge University Press.

[24] *Witiuk, S. W. (1988):* Promoting Spatially-Oriented Referencing Systems Development at the International Level. Information Systems for Government and Business: Trends, Issues, Challenges. Nagoya, Japan, United Nations Centre for Regional Development: 163-169.

[25] *Zdzisław, P. (1991):* Rough Sets: Theoretical Aspects of Reasoning About Data. Dordrecht, Kluwer Academic Publishing.

Contact

Univ. Prof. Dr. Andrew U. Frank, Vienna University of Technology (TU Wien), Department of Geoinformation and Cartography, Gusshausstrasse 27-29/E127, A-1040 Vienna, Austria. E-mail: frank@geoinfo.tuwien.ac.at
vgi

Jetzt mit Mobile Internet

NTRIP*

APOS

Austrian Positioning Service

Um einen Schritt schneller und wirtschaftlicher

APOS - Ansprechpartner
BEV - Bundesamt für Eich- und Vermessungswesen
Kundenservice
Tel.: +43-(0)1 - 21110 - 2160
Fax: +43-(0)1 - 21110 - 2161

E-mail: kundenservice@bev.gv.at
See you: www.bev.gv.at

NTRIP*
Ein vom Bundesamt
für Kartographie und Geodäsie
entwickeltes Verfahren
zur Bereitstellung von
GNSS/GPS - Datenströmen
über das Internet.

BEV - Bundesamt für Eich- und Vermessungswesen

BEV
Bundesamt für Eich- und Vermessungswesen

Wie bringen Sie
Ihre Firma auf
die Überholspur?

Schnallen Sie sich an, die Fahrt geht los. Trimble macht Sie fit für die Zukunft. Die Revolution in der Vermessung ist in vollem Gange. Wir verkaufen Ihnen nicht nur Vermessungstechnologie, sondern stehen an vorderster Front, wenn es darum geht, Ihnen zu geschäftlichem Erfolg zu verhelfen. Das beste Werkzeug dazu, ist das Trimble Connected Site Konzept.

Unser neues Konzept sorgt für eine nahtlose Verbindung zwischen Ihrem Büro und dem Einsatzort, und beinhaltet innovative und hochproduktive Lösungen. Sei es durch die geeigneten Werkzeuge, Techniken und Dienstleistungen oder durch passende Geschäftspartnerschaften - Ihre Effizienz und Produktivität steigt in jedem Fall. Als bekannter führender Anbieter vielseitig einsetzbarer Vermessungslösungen höchster Qualität sind wir für Sie da, um Sie für alle Eventualitäten zu rüsten.

Trimble ist Ihr Partner auf dem Weg in die Zukunft.

Nähere Information erhalten Sie im Internet unter **www.trimble.com/overdrive**

⊛Trimble.

www.trimble.com