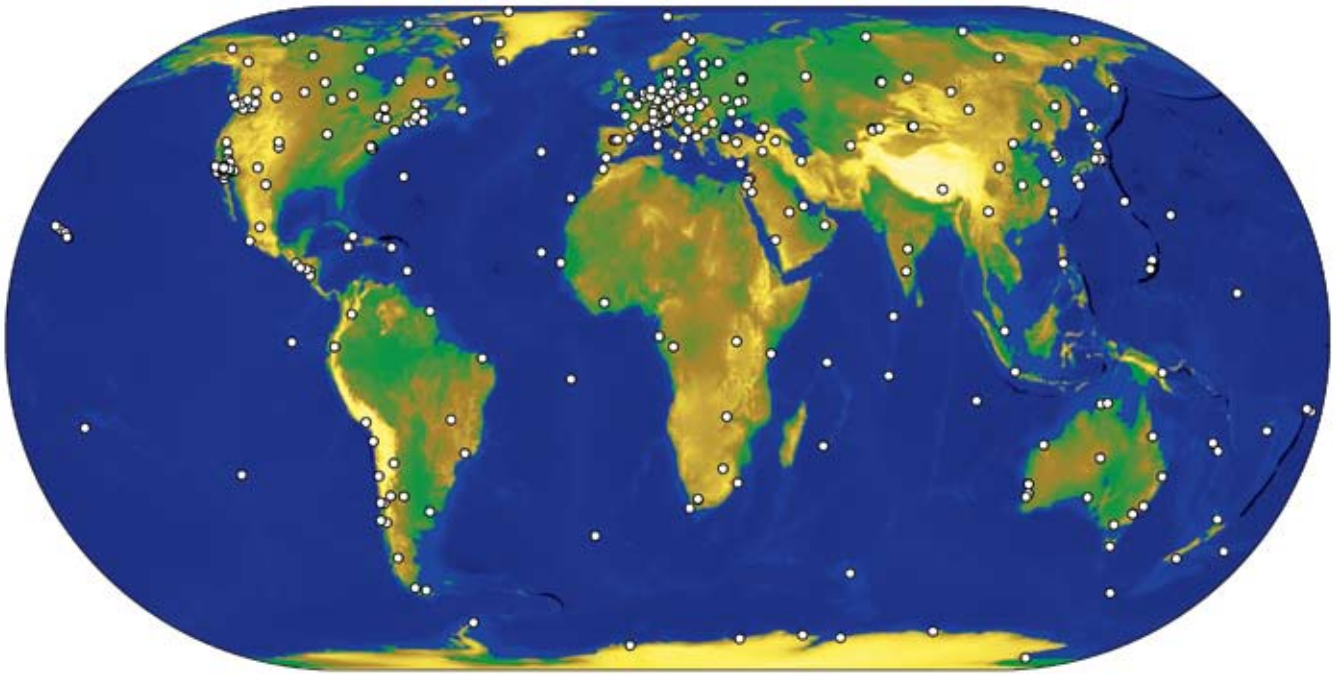
Interpolating Reference Data

Kinematic Positioning Using Public GNSS Networks



International GNSS Service Network

High-accuracy users of public GNSS reference archives don't always have access to data that matches the high sampling frequencies needed for real-time kinematic techniques. This column proposes a method for interpolating GNSS observation residuals so as to produce data that can be used in kinematic applications.

TORBEN SCHÜLER
INSTITUTE OF GEODESY AND NAVIGATION,
UNIVERSITY FAF MUNICH, GERMANY

Kinematic GNSS data processing algorithms are constantly being improved to work over long inter-station distances. Within the GNSS community some users have a need for publicly available data for high-frequency (equal to or greater than one hertz) kinematic applications. These applications may be separated into two classes: those using post-processed data and those requiring real-time or near real-time data.

A variety of commercial and public GNSS service providers offer reference data derived from GNSS satellite broadcasts and archived in near real-time, which can be used for long-baseline kinematic applications. Several commercial vendors provide reference data through satellite-based channels at relatively high frequency rates (up to one Hertz).

A number of public service providers, however, offer data free of charge via the Internet, notably the International GNSS Service (IGS) Tracking Network with approximately 350 stations distributed globally. The IGS network has a relatively high density over North-America and Europe. Further densifications exist regionally, such as the National Continuously Operating Reference Stations (CORS) Network supported by the U.S. National Geodetic Survey, the EUREF Permanent Network in Europe, and the Shenzehn CORS network in southeastern China.

Consequently, publicly available GNSS reference station data are becoming more and more attractive for use in kinematic applications. Unfortunately, these archived data normally exhibit sampling intervals as large as 30 seconds, whereas kinematic positioning requires sampling frequencies in the range of 1 to 20 Hz or even higher. For instance, only sites belonging to the IGS Low Earth Orbit (LEO) network — a subset of the IGS tracking network — provide 1 Hz data. There are many scientific applications, such as hydrographic surveying or airborne photogrammetry, and gravimetry, that can benefit from using publicly available, "freeware data" to generate 1 Hz data.

The decision of the U.S. government to turn off S/A (Selective Availability) in May 2000 opened the door for accurate densification of reference station data, but this opportunity has not received proper attention yet. In the future, the availability of the future Galileo system, combined with existing GPS and GLONASS satellite constellations, will further improve the positioning performance.

For the time being, however, in using the low-frequency public reference network data the need to interpolate reference station measurements in the time domain arises, because both rover and reference station data must be available with the same temporal resolution (and synchronously). This article will discuss techniques by which low-frequency reference station data can be used to generate higher sampling rates for use in kinematic applications, both post-processed and in real-time.

The Algorithm

In this article, GNSS data interpolation is not carried out on the measurements directly. Instead, the known orbit information is used to compute observation residuals that will be interpolated afterwards. This procedure is slightly more complicated than a direct interpolation of the measurements — for example, using high-order polynomials — but it is still easy to implement.

The observation equation for a pseudorange measurement *PR* at reference station *A* to satellite *i* at epoch *k* is

$$PR^{i}_{A,L1,k} = \rho^{i}_{A,k} + \delta S^{i}_{A,Trop,k} + \delta S^{i}_{A,L1,Ion,k} + \left(\delta t_{Ak} - \delta t^{i}_{k}\right) \cdot c \quad (1)$$

ρ	*distance*
δS_{Trop}	*tropospheric propagation delay*
δS_{Ion}	*ionospheric delay*
δS_{A}	*receiver clock error*
δt^{i}	*satellite clock error*

(All equations are only presented for the primary L1 carrier frequency for the sake of simplicity.) Most of the dynamics are contained in the geometrical distance ρ because the satellites are cruising at a speed of around 3 km/s.

Fortunately, both satellite and reference station positions are precisely known, that is, the term ρ can be easily evaluated. The atmospheric errors, on the contrary, are considered as slowly varying error terms. We employ default models (subscript 'MOD') to predict these terms. The satellite clock error δt^{i} is treated as a quantity that does not show any unexpected behavior within the interpolation interval, now that S/A is — and likely to remain — switched off. Finally, the receiver clock error must be estimated:

$$\delta t_{Ak} = \frac{PR^{i}_{A,L1,k} - \rho^{i}_{A,k} - \delta S^{i}_{A,L1,Ion,MOD,k} - \delta S^{i}_{A,Trop,MOD,k}}{c} + \delta t^{i}_{k} \quad (2)$$

Of course, any sudden jumps of the receiver clock term, which might occur when a new satellite rises or an observed one drops below the elevation mask angle, negatively affects the interpolation process. In order to avoid such jumps, the receiver clock error is always deduced from the master satellite — the one with highest elevation. Alternatively, one could estimate the clock error from the entire ensemble of measurements employing a strong elevation-dependent weighting scheme that gives very little weight to low-elevation satellites and considerably larger weight to those of high-elevation satellite measurements.

For all required epochs, the pseudorange residual

$$\Delta PR^{i}_{A,L1,k} = PR^{i}_{A,L1,k} - \rho^{i}_{A,k} - c \cdot (\delta t_{A_k} - \delta t^{i}_{k}) \quad (3)$$

One of the in-house reference stations of University FAF Munich, Germany (above) Ciutadella Reference Station, ENVISAT Project, Spain (inset)

can be computed and basically represents the slowly varying terms of atmospheric propagation delays. We omitted the predicted errors of the tropospheric and ionospheric delays in this formula because these modeled quantities remain almost constant during two subsequent epochs and, thus, do not require any further treatment.

Analogously, the carrier-phase residual reads

$$\Delta\varphi^{i}_{A,L1,k} = \varphi^{i}_{A,L1,k} - \frac{\rho^{i}_{A,k} + c \cdot (\delta t_{A_k} - \delta t^{i}_{k})}{\lambda_{L1}} \quad (4)$$

and is given in cycles here. It contains the atmospheric effects plus the (non-varying) ambiguity unknown. Although the receiver clock error is computed with the help of pseudorange measurements, this fact has very little effect because the positioning software will employ double differences that largely reduce clock error inadequacies.

The interpolation of the observation residuals can now be performed with the following linear function

$$\Delta\varphi^{i}_{A,L1,t} = a_0 + a_1 \cdot t \quad (5)$$

where the coefficients a_0 and a_1 are derived using the observation residuals of epochs t_k and t_{k+1} with $t_k < t < t_{k+1}$. Note that quadratic interpo-

lation showed no improvement during the tests described in this article. The same statement is true regarding the use of high-precision IGS orbits versus broadcast orbit information. The latter is less accurate, but not problematic either, because orbit errors are slowly varying and can also be compensated by the linear interpolation procedure for short time intervals.

The interpolated carrier phase then reads

$$\varphi^{i}_{A,L1,t} = \frac{\rho^{i}_{A,t} + c \bullet (\delta t_{A_t} - \delta t^{i}_{t})}{\lambda_{L1}} + \Delta\varphi^{i}_{A,L1,t} \quad (6)$$

with ρ_t and δt_t^i being computed with knowledge of the orbit information and δt_{At}, δt^i_t being interpolated in the same manner as the observation residual $\Delta\phi_t$. An algorithm as depicted before was implemented in the in-house software "PrePos GNSS Suite - Module RInter" <*http://www.unibw.de/ifen/software/prepos*> and used in the tests described later.

The main drawback of interpolation compared with the use of reference data recorded at the nominal frequency of the roving receiver is that interpolation cannot be carried out when (non-fixable) cycle slips occur. However, because reference station data collection is a static scenario, cycle slips tend to be less frequent in comparison to kinematic scenarios, and we can often fix them when they do occur.

Positioning Tests

Positioning experiments were carried out on the campus area of the University FAF Munich, Germany, in order to demonstrate the feasibility of this interpolation approach. A short baseline (about 40 meters) was established with one receiver acting as rover and the other as reference station. Data were collected for 37 minutes at a nominal sampling frequency of 1 Hz. Afterwards, the rover data were decimated to 5, 10, 15, and 30 seconds and interpolated back to the nominal frequency. We decimated the data by using the "worst case," for example, for the 15-second interval all those records belonging to epochs different from 0s, 15s, 30s, 45s, etc. were simply deleted. Afterwards, the decimated data were interpolated using the algorithm described earlier, and compared to the original data.

As this baseline is very short, atmospheric errors are effectively eliminated when double difference measurements are formed for GNSS positioning. This statement is true when using the original data. However, for the decimated data we can expect those atmospheric

Sampling interval (in seconds)	Double Difference	North	East	Radial
1	-	4.3 mm	2.4 mm	7.5 mm
5	2.4 mm	4.5 mm	2.4 mm	8.1 mm
10	3.3 mm	5.0 mm	2.7 mm	8.9 mm
15	4.0 mm	5.4 mm	2.7 mm	10.2 mm
30	6.0 mm	7.4 mm	3.6 mm	14.4 mm

TABLE 1. Standard deviation of interpolated carrier phases and positioning accuracy with respect to reference coordinates (1-second results: original data set, no interpolation).

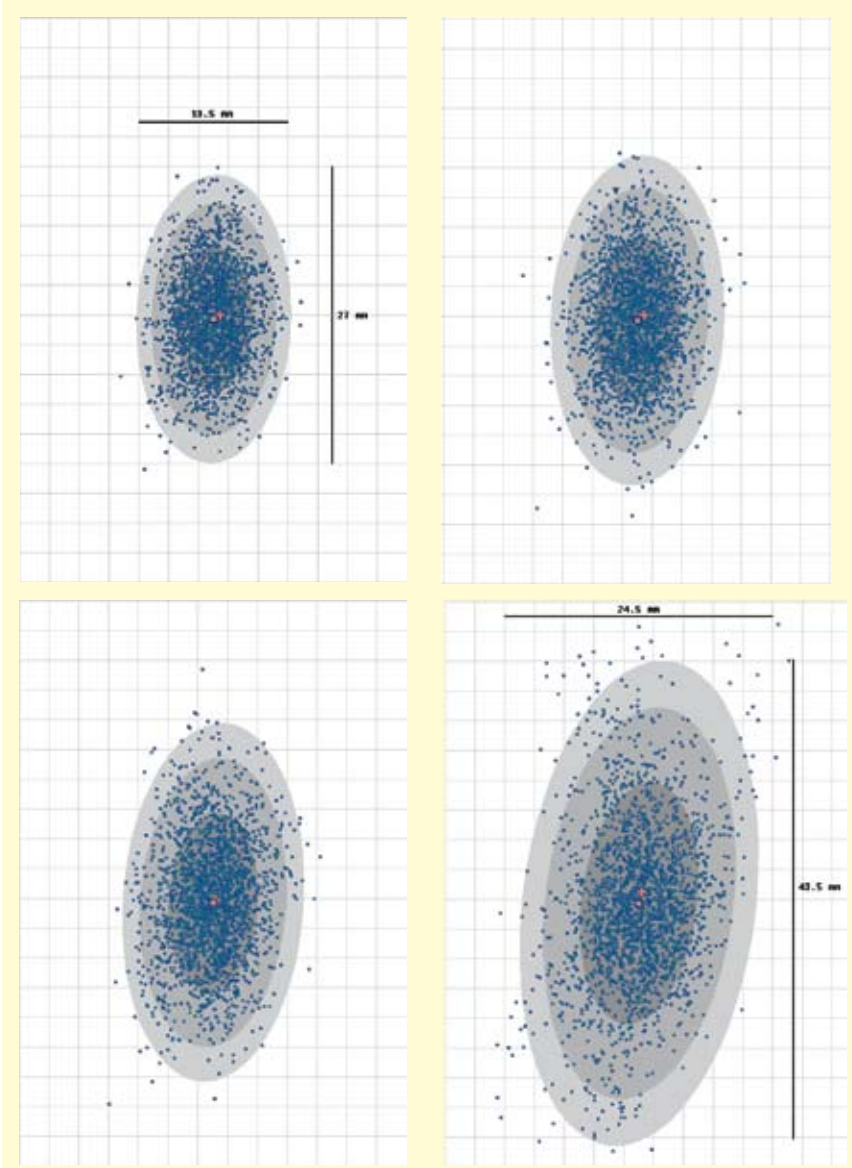


FIGURE 1 Horizontal position scatterplots using the original data (1s) as well as selected interpolated datasets (10s, 15s and 30s); same scaling for all plots. The red dot represents the reference position; the gray shaded areas represent confidence ellipses of probabilities (from center) of 68, 95 and 99 percent.

errors (and other remaining contributors to the error budget) will still show variations that depart from linear nature (the approach used for interpolation). So, we must anticipate a larger interpolation error for longer data sampling intervals, which maps into the double difference measurements.

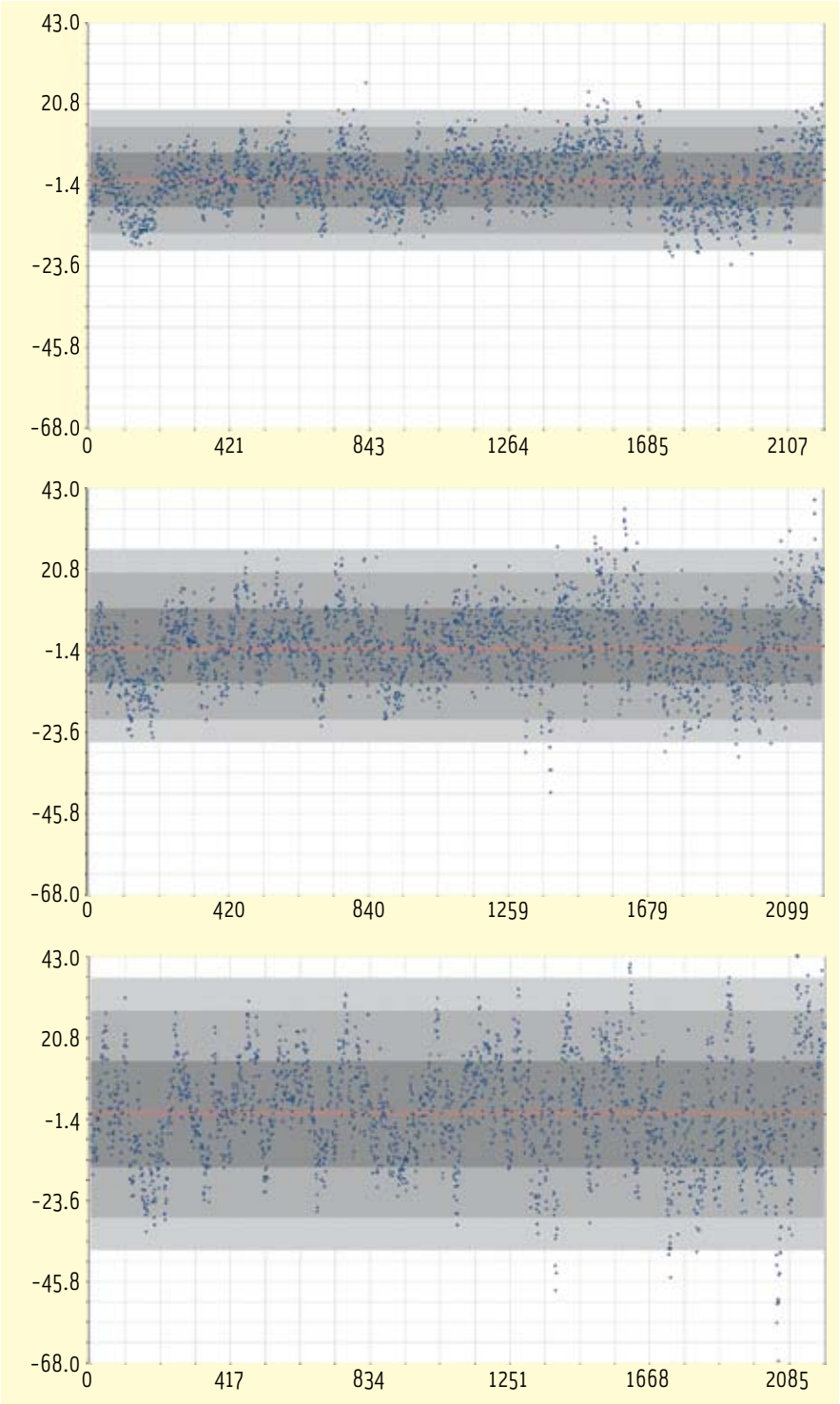


FIGURE 2 Vertical scatter diagrams using the original data (1 s) as well as selected interpolated data sets (15 s and 30 s); y-axis in units of millimeters, x-axis in seconds. The red line represents the reference position; the gray shaded areas represent probabilities of 68, 95 and 99 percent (darker to lighter shades, respectively).

Please note that this contribution concentrates on L1 measurements. L2 data were analyzed in the 2006 article by this author cited in the Additional Resources section at the end of this article. That analysis showed an agreement similar to the L1 observations.

Carrier Phase Measurements

Table 1 shows results obtained from interpolated carrier-phase measurements on L1. The first column gives the sampling interval of the input data for the interpolator. The one-second data are not interpolated. The second column depicts the agreement of the original versus those double differences containing interpolated data. As expected, we can see that the longer the sampling interval becomes, the lower the agreement with the original data will be.

Columns 3 and 4 portray the horizontal positioning accuracy (standard deviation from difference to the known reference coordinates, rms), and the last column provides information about the height accuracy. All values represent the standard deviations of a single-shot result (epoch-to-epoch solution), that is, *not* the precision of the averaged/adjusted mean coordinate result.

Figure 1 graphically illustrates these results (horizontal scatter plots for 1 s, 10 s, 15 s and 30 s), and Figure 2 (vertical scatter diagrams for 1 s, 15 s and 30 s). The empirical confidence ellipses for probabilities of 68, 95, and 99 percent were drawn in these scatter plots in shades of gray. The corresponding confidence intervals of the radial coordinate channel are indicated as well. The red cross/red line indicates the reference position.

The positioning accuracy using interpolated data does not suffer much for intervals of 5 to 10 seconds: The agreement (1 σ) of the horizontal positions is almost identical regarding the 5-second data and less than 1 millimeter worse for the 10-second data.

The height component — as always with GNSS positioning — suffers a bit more, as seen in Figure 2. For the 30-second data in Figure 1, however, this

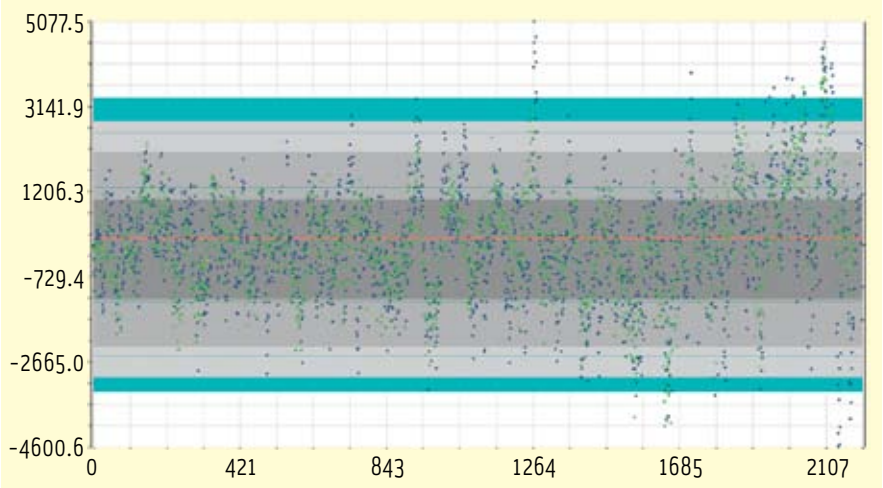


FIGURE 3 Vertical scatter diagram using the original pseudo-range data (1s, blue dots) and the interpolated 30-s-dataset (yellow-green dots). Empirical 68%, 95% and 99% confidence intervals are indicated in shades of gray for the interpolated and in cyan for the original dataset (y-axis in units of millimeters, x-axis in seconds).

Sampling interval	Double Difference	North	East	Radial
1 s	-	0.74 m	0.47 m	1.30 m
5 s	0.14 m	0.72 m	0.46 m	1.27 m
10 s	0.25 m	0.69 m	0.44 m	1.21 m
15 s	0.33 m	0.65 m	0.43 m	1.17 m
30 s	0.47 m	0.64 m	0.43 m	1.12 m

TABLE 2. Standard deviation of interpolated pseudoranges and positioning accuracy with respect to reference coordinates.

component is already 7 millimeters less accurate (almost twice the error) than the results obtained from the original data set, and the horizontal scatter is more than 3 millimeters larger (about 75 percent more)— a tendency that is also evident when having a look at Figure 2.

Pseudorange Measurements

Although only carrier phase measurements are needed for high-precision applications, it is still worth having a closer look at the interpolated pseudorange measurements. Table 2 depicts the results for this type of observation analogously to Table 1. Data processing with the in-house software was conducted without any smoothing of the pseudorange data.

The results are partly astonishing: The agreement between the original double differences and the interpolated ones is as expected. Again, the larger the sampling interval becomes the more the standard deviation increases. However, the positioning accuracy increases (i.e., the rms error decreases) with increasing sampling interval! The increase in accuracy is only marginal (about 10 centimeters for the horizontal position and less than 20 centimeters for the vertical channel) compared to the ordinary noise of pseudoranges, but we still must find a suitable explanation for this apparent contradiction.

Figure 3 helps to understand this phenomenon. The reason is related to the fact that multipath plays a dominant role for the (un-smoothed) pseudoranges. This error contributor is much more emphasized for this type of observations than for carrier phases, whereas other contributors, such as changes in ionospheric delay, play an important role for the carrier-phase interpolation, but only a marginal one when looking at the ranges.

The standard deviation of the double differences (see second column of Tables 1 and 2) is obtained from the comparison of the interpolated with the original observations. These values do not represent any kind of "accuracy," but simply indicate the agreement of interpolated and the native data. In other words, this standard deviation reveals how precisely the original data can be reconstructed from the interpolated ones. The better the errors inherent in the original measurements can be reconstructed, the smaller this standard deviation will become — but this has nothing to do with the accuracy of the observations.

On the contrary, the standard deviation of the coordinate components is computed taking into account the known reference position. Assuming that this reference matches the true position, the coordinate rms values shown in the tables will indicate the accuracy of the positions.

Looking at Figure 3 we can easily see typical multipath signatures in the time series. Furthermore, we can find several events where the peak errors of short-term multipath were cut off during the interpolation (compare yellow-green dots versus blue dots indicating heights from original data). As a matter of fact, the multipath errors inherent in the interpolated data can never be larger than that present in this discrete, original time series as long as only a linear interpolator is employed — but it can be smaller! As a result, the positioning accuracy becomes better in this test trial, and, generally speaking, it will be better as long as multipath errors are dominant in comparison to the other contributors to the error budget.

In the carrier phase results (Figure 2), multipath signatures can be seen as well. However, they are not large enough to overcome the expected tendency of a poorer interpolation performance with increased sampling interval due to a less-accurate interpolation of the remaining error contributors.

Real-Time Aspects

Although this article focuses on data interpolation for post-processing (or near real-time) applications, many readers might also be interested in true real-time applications. In this case, the algorithm depicted at the outset can be used in a straightforward manner as an extrapolator. Some results are shown in Tables 3 and 4 as well as in Figure 4.

Linear extrapolation according to Equation (5) was carried out for all quantities needed. The interval of the

reference data was chosen in a "symmetrical manner" with respect to the prediction (extrapolation) interval: For example, the data predicted for 5 seconds in advance (t_{+5}) are based on reference data for epochs t_0 and t_{-5}, because it would not be wise to base an extrapolation on the original data sampling interval of just 1 second. This would quickly magnify extrapolation errors for long-term predictions, according to the laws of error propagation.

Table 3 gives results for the single-shot positioning accuracy with carrier phase measurements. We can see that data prediction is actually possible, but we cannot expect the same accuracy as for data interpolation. Data prediction for +10 seconds showed a horizontal position error of 14 mm, and the vertical coordinate error is 26 mm. Data interpolation from 10-second data provided a horizontal accuracy of six millimeters and less than one centimeter in height (see Table 1).

Figure 4 illustrates that prediction of data one second in advance only shows a minor decrease in accuracy. This type of short-term extrapolation is needed for applications with very strict real-time requirements because in a differential network, data latencies are inevitable. Consequently, small-scale predictions are needed to provide true real-time capabilities.

Nevertheless, the other results also indicate that communication outages of less than 30 seconds are not necessarily fatal for RTK positioning because extrapolation of the reference station data can still be accomplished with reasonable — although not the highest — accuracy. For example, for a 15-second outage in a GSM network, the errors in the original data will amount to a horizontal RMS error of 4.5 millimeters, while the 15-second prediction will lead to a "single-shot" RMS of 19.2 mm.

It seems that an accuracy of about two centimeters will still be sufficient for many kinematic applications. Some improvements to the algorithm could still be made with respect to extrapolation. For instance, better satellite clock error prediction models would enhance the accuracy a bit.

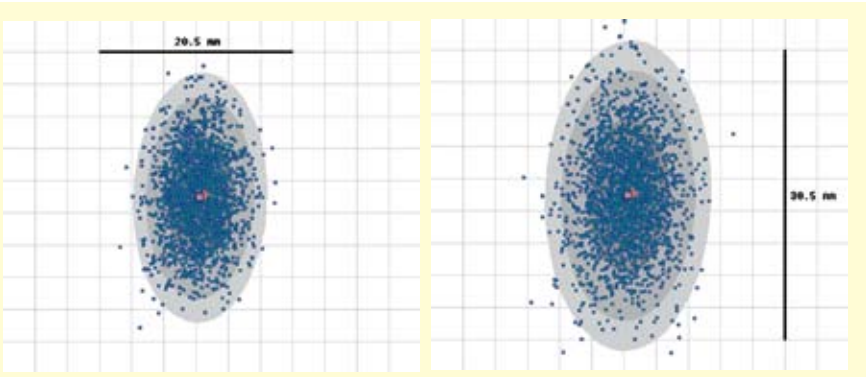


FIGURE 4 Comparison of horizontal position scatter from original data (left) and short-term predictions (+1 s, right).

Extrapolation Interval	North	East	Radial
0 s	4 mm	2 mm	8 mm
1 s	5 mm	3 mm	10 mm
5 s	8 mm	4 mm	17 mm
10 s	14 mm	6 mm	26 mm
15 s	17 mm	9 mm	31 mm
30 s	30 mm	17 mm	44 mm

TABLE 3. Single-shot positioning accuracy (rms) using extrapolated carrier-phase measurements (0 s: original data set).

Extrapolation Interval	North	East	Radial
0 s	0.7 m	0.5 m	1.3 m
1 s	0.8 m	0.5 m	1.4 m
5 s	1.0 m	0.7 m	1.7 m
10 s	1.3 m	0.8 m	2.2 m
15 s	1.3 m	0.8 m	2.3 m
30 s	1.3 m	1.0 m	2.7 m

TABLE 4. Single-shot positioning accuracy (rms) using extrapolated pseudo-range measurements (0 s: original data set).

From Table 4 we can conclude that the pseudorange positioning accuracy does not decrease as quickly as the carrier-phase measurements. The height error is about a factor of 2 larger for the 30-second predictions compared to the original data.

Summary and Conclusions

Interpolation of GNSS measurements is a task often undertaken by users who want to exploit publicly available data in high-frequency kinematic scenarios. We presented an algorithm for precise data interpolation that reduces the original measurements with the help of orbit information and applies linear interpolation to the slowly varying observation residuals.

The loss of accuracy is rather marginal when interpolated carrier-phases from data sampled each 5 seconds are used compared to the positions obtained from original data. Even for 10-second data the agreement between original and interpolated data is still within 1 millimeter for the horizontal position and 1.5 millimeter for the vertical coordinate.

These results indicate that reference station data do not necessarily need to be archived with the nominal sampling frequency (often 1 Hz). Instead, an optimal trade-off between storage capacity and sampling interval would be between 5 and 10 seconds. This statement is in agreement with 5-second sampling interval recommended by L. Wanninger in the work cited in the Additional Resources section as well as J. Wickert et al. who recommended 10-second sampling intervals for ground station data used for CHAMP radio occultation measurements.

Data sampled every 30 seconds tends to be a bit too sparse for interpolation when the highest precision is required in kinematic applications. Unfortunately, this is the commonly used output interval of publicly available GPS reference networks. It is therefore recommended to enhance the output interval to at least 15 seconds (or better 10 seconds) in order

to offer useful data to more users, thus expanding the idea of real multipurpose networks.

Author

Torben Schüler received his diploma in geodesy and cartography from the University of Hannover. Afterwards, he joined the Institute of Geodesy and Navigation (University FAF Munich) as a research associate, earned a doctorate, and is now head of the GNSS/INS Laboratory at the same institute. His major research work is focused on precise GNSS positioning, including atmospheric and, in particular, tropospheric delay modeling. Schüler was heavily involved in the development of the standard tropospheric correction model for the Galileo satellite navigation system and is currently investigating the impact of the innovations in satellite navigation on active networking positioning.

"Working Papers" explore the technical and scientific themes that underpin GNSS programs and applications. This regular column is coordinated by Prof. Dr.-Ing. Günter Hein, a member of the European Commission's Galileo Signal Task Force and organizer of the annual Munich Satellite Navigation Summit. He has been a full professor and director of the Institute of Geodesy and Navigation at the University FAF Munich since 1983. In 2002, he received the United States Institute of Navigation Johannes Kepler Award for sustained and significant contributions to the development of satellite navigation. He received his Dipl.-Ing and Dr.-Ing. degrees in geodesy from the University of Darmstadt, Germany. Contact Professor Hein at <Guenter.Hein@unibw-muenchen.de>.

Manufacturers

Equipment used for test trials included a Z-Sensor base station, from Thales Navigation (now **Magellan Navigation, Inc.**) San Dimas, California USA, and, as the rover unit, the RTK GPS 5800 from **Trimble,** Sunnyvale, California USA.

Additional Resources

Articles

[1] Mader, G. L., and M. L. Morrison, "Using Interpolation and Extrapolation Techniques to Yield High Data Rates and Ionosphere Delay Estimates from Continuously Operating GPS Networks," Proceedings of ION GPS 2002, September 24–27, 2002, Portland, Oregon, USA, pp. 2342-2348

[2] Schüler, T., "10 Hz or 10 s?" GPS Solutions, DOI 10.1007/s10291-006-0034-8, in press, Springer Verlag, 2006

[3] Wanninger, L., "Interpolation von GPS-Beobachtungen," Allgemeine Vermessungs-Nachrichten (AVN), Vol. 10/2000, 107:360-363

[4] Wickert, J., and R. Galas, G. Beyerle, R. König, and Ch. Reigber, Ch. "GPS Ground Station Data for CHAMP Radio Occultation Measurements," Phys. Chem. Earth (A), 26 (6-8), 503-511, 2001

WWW Links

IGS Tracking Network: http://igscb.jpl.nasa.gov/network/netindex.html

CORS Network: http://www.ngs.noaa.gov/CORS/

EUREF Network: http://www.epncb.oma.be/