J. BEZRUCHKA

AN ADAPTIVE DISPLACEMENT DETECTION ALGORITHM

Juraj BEZRUCHKA

Email: juraj.bezrucka@stuba.sk

Research field: near real time monitoring of GNSS networks, the detection of possible displacements

Address: Department of Theoretical Geodesy, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia

ABSTRACT

During permanent GNSS observations a displacement of the network's station may occur. Not all changes are significant enough to raise an alarm. However, the critical value of the station's position change may vary significantly for various reasons, such as the antenna model, receiver or observation interval. The coordinates gained by GNSS methods are also usually characterized by a higher degree of precision of the horizontal components than those of the vertical ones. It is virtually impossible to set a single critical value for each component of each station of the network; therefore, an automatic algorithm evaluating the estimates should be used. The paper suggests an adaptive algorithm using a combination of the Kalman filter and a modified moving average technique to reveal sudden changes in a GNSS station's position, while allowing it to adapt to long-term changes.

KEY WORDS

- Displacement detection,
- Kalman filter,
- adaptive algorithm.

1 INTRODUCTION

When analyzing GNSS network solutions, especially when short observations are used, the process is characterized by a rather noisy output and overestimated root mean square error (henceforth called "RMSE"). The noise may result in an outlying position estimate. However, a critical value that helps to distinguish between an outlier and a real displacement may vary considerably for different stations due to various effects and can even change for a single station during long-term observations. Therefore, it is impossible to set a universal critical value valid for the whole network.

If a displacement station occurs at the station during the observation, it should be reported. The goal of this paper is to introduce a very simple adaptive one-pass algorithm that evaluates the results of the processing and reveals possible displacements. A specific feature of the algorithm is the use of two adjustable meta parameters binding the past observations with the current one. The parameters allow for the use of the same setting for all stations of the network, regardless of their performance.

The paper is divided into three parts. The first describes the use of a Kalman filter when processing GNSS network solutions. The main part first describes the a priori test of the network shift, then focuses on a new adaptive algorithm for displacement detection, and finally introduces several optional extensions of the algorithm. The last part briefly summarizes the features of the introduced algorithm.

2 KALMAN FILTER IN PERMANENT GNSS NETWORK PROCESSING

The analysis of the network performance is based on the Bernese GPS Software (Dach et al., 2007) position estimates. These estimates are further processed and analyzed by the Kalman filter-based (Kalman, 1960) algorithm **kane**s (Bezrucka, 2010). Although the input data are not direct observations, for the purposes of the Kalman filtering, the estimated positions are regarded as observed. The program output represents an estimate of the current position and current velocity for each component of each station based on previ-

ous estimates and new observations of the coordinate components. The model can be expressed as

$$\mathbf{z}_t = \mathbf{H}_t \cdot \mathbf{x}_t, \quad (1)$$

where $\mathbf{z}$ represents an observation, in this case the station's coordinates from SINEX transformed to a local coordinate system (the first occurrence is set as a reference position); $\mathbf{H}$ is an observation matrix; and $\mathbf{x}$ represents the unknown parameters – the estimated position and velocity for epoch t . It is recommended to use local coordinates when estimating the unknown parameters, because the RMSE of the vertical component is significantly worse than that of the horizontal coordinates. This could not be taken into account if geocentric coordinates were used.

The principles of the Kalman filter have been thoroughly explained in many works (e.g., Welch-Bishop 2001, Haykin 2001). It consists of two main parts – a *prediction* and an *update*. In the prediction part the current values of the unknown parameters $\mathbf{x}_t$ and their covariance matrix $\mathbf{P}_t$ are predicted using a transition matrix $\mathbf{A}$, the last estimate $\mathbf{x}_{t-1}$ and covariance matrix $\mathbf{P}_{t-1}$, respectively.

$$\mathbf{x}_t^- = \mathbf{A}_t \mathbf{x}_{t-1} \quad (2)$$

$$\mathbf{P}_t^- = \mathbf{A}_t \mathbf{P}_{t-1} \mathbf{A}_t^T + \mathbf{Q} \quad (3)$$

The update part updates the original prediction using the new observation $\mathbf{z}_t$ and the Kalman gain $\mathbf{K}_t$ according to Eqs. (4) – (6)

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \quad (4)$$

$$\mathbf{x}_t = \mathbf{x}_t^- + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \mathbf{x}_t^-) \quad (5)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_t^- \quad (6)$$

where $\mathbf{Q}$ and $\mathbf{R}$ represent the white noise covariance matrix of the process and the observation covariance matrix respectively. $\mathbf{Q}$ uses empirical values: 10^{-8} for position estimates $Q_{1,1} - Q_{3,3}$ and 10^{-18} for velocity estimates $Q_{4,4} - Q_{6,6}$. As we expect the coordinate components and their velocities to be independent, the non-diagonal elements $Q_{i,j}$ are set to zero. For better results and to eliminate the use of arbitrary values, a modification can be made according to (Bos et al. 2005). The observation covariance matrix $\mathbf{R}$ is taken from SINEX and transformed to the local coordinate system. These values are used, although they are rather noisy and characterized by an overestimated RMSE as a result of ignoring various effects (Krynski-Zanimonskiy, 2003). Both the RMSE and covariance matrix are, however, characterized by stable values with only minor changes; therefore, they can be exploited to detect a station's unexpected behavior. This approach takes into account no covariance between different stations; the stations are processed independently. We expect that the coordinates between two subsequent epochs may be considered as independent.

3 THE DISPLACEMENT DETECTION TESTS

3.1 A priori test of the network shift

The first test performed on the a priori residuals is one testing a possible network shift. The a priori residual used in Eq. (5)

$$\mathbf{e}_t = \mathbf{z}_t - \mathbf{H}_t \mathbf{x}_t^- \quad (7)$$

represents the difference between an expectation based on the prediction part of the Kalman filter and an observation.

The analysis center of the Slovak University of Technology uses a single reference station when processing a GNSS network solution. This approach has been chosen due to its low impact on the network design as the most appropriate method for discovering possible earth surface deformations. If more reference points were used, the risk of the network shift would be minimized, and the internal precision would probably be higher too. On the other hand, the forced positions of several stations may result in apparent deformations at stations not used as reference points, which could be misinterpreted as a real change in position and thus could raise a false alarm or vice versa – although less probably – that some deformations may be overlooked.

As mentioned above, referencing to a single station results in a major disadvantage. The data from a reference station are treated as almost perfect coordinates (the RMSE is set very close to zero). However, a displacement at this very station or the lower quality data observed may cause an apparent shift in the whole network.

This was the reason to introduce a network shift test in the prediction part. The test evaluates the a priori differences of all the geocentric coordinates components $[X, Y, Z]$ of all the stations (with the exception of the reference station) using a simple quartile test [Mendenhall-Sincich 1988]. The upper and lower limits are determined according to (Tukey 1977) as

$$UL = UQ + c * IQR \quad (8)$$

$$LL = LQ - c * IQR \quad (9)$$

where UL , LL are the upper and lower limits respectively; UQ , LQ are the upper and lower quartiles respectively; $IQR = UQ - LQ$ is interquartile; and c is a coefficient (for 95% probability, $c \approx 1.5$). Using the test we estimate the approximate shift and the limits for outliers (LL , UL). Therefore, if zero does not fall within the limits, the network is assumed to be significantly shifted. The program allows for the correction of the shift by moving the whole network according to the previously determined value and the use of the corrected coordinates for further analysis. Multiple tests performed on various networks and parameters (observation interval, observation duration, etc.) show this technique provides better results and is therefore recommended.

3.2 The Adaptive Displacement Detection

3.2.1 A posteriori test of the residuals

The new parameter estimates are available after the new observations are processed by the Kalman filter. Through the use of a simple transformation, the current position estimate of the observed station can be calculated.

An “unreliable” value can be either a temporarily changed position or an observation with a low degree of precision. It is important to avoid reporting a permanently changed position as a bad observation. This is provided by updating the test parameter, which could be the difference between the expected and observed positions (“observation error”), or an observation error weighted by the RMSE of the observation, etc.

The duration of the observation, or observation window, is a very important and influential parameter of processing output tests. It is obvious that the precision of a 3-hour-long observation must be significantly lower than that of a week-long campaign. The advantage of a short observation window is a more realistic and timely estimate of a changing position compared to a weekly position with averaged tidal and other important effects. The observations with a higher RMSE are spread less densely; therefore, the differences between the expected and observed or estimated values are higher. How can we determine the critical value that points out the problem in the network and distinguish the real displacements from erroneous or less precise observations? Defining the empirical critical values for the coordinate residuals seems to be the simplest way. However, when using GPS technology, the vertical coordinate component has a several times larger standard error compared to the horizontal components. Therefore, we should define the various critical values for each component. Permanent GNSS networks are seldom homogenous (different antennae, receivers), and the local environment is never perfect. This would lead to defining specific critical values for each site.

Such a system would be too complicated, although we have not yet mentioned the various lengths of observations. There is a motivation to create a self-determining model, which defines the critical values automatically according to the performance of the permanent station. The removal of rigid critical values and their replacement with autonomously updating values allow the system to adapt in cases of changed conditions (receiver replaced, changes of field of view, etc.)

Sometimes a minor change in the GNSS station’s position may occur. Whether it is a natural process (an earthquake, a landslide) or the result of human action (antenna experiments, pillar changes), the system should be able to adapt to changed conditions and still be able to handle outliers. Ting et al. (2007) suggests an effective outlier-robust Kalman filter. This approach required substantial changes in the original algorithm; therefore, another approach has been chosen. The aforementioned filter, however, may be tested in an alternate version of the algorithm.

The test value U consists of two components. The first parameter is proportional to the observation residual, i.e., the difference between the observed and expected (predicted) positions. The second parameter is proportional to the standard error of either observation or estimate; in our case, the geometric average of both has been cho-

sen. To emphasize the importance of the residual, it is suggested to use the square of its value and compute the test value as

$$U = (A - B)^2 \cdot (m_a m_b), \quad (10)$$

where A , m_a are the observed position and its RMSE, respectively, B , m_b are the predicted position and its RMSE, respectively.

Eq. (10) can be modified arbitrarily to fit the user’s needs. To emphasize the effect of the displacement and to lower the importance of the RMSE, the square root of the product $m_a m_b$ may be used or can even be omitted. Although the RMSE rarely changes significantly, the tests show that this information is useful for a better performance of the algorithm.

The test parameter has been initialized. To allow for the adaptation of the parameter, it has to be updated on a regular basis. The Kalman filter is a recursive algorithm; it only uses the last estimate and a new observation to update the estimate parameters. The test parameter should be updated in the same fashion, i.e., using the last estimate of the test parameter V_{n-1} as the expected value and a gain U based on the new observation. The coefficient $\omega = 0.8$ ensures a performance similar to a moving average of the last 5 observations. The test parameter is updated as

$$V_n = \omega \cdot V_{n-1} + nw \cdot U, \quad (11)$$

where $nw = (1 - \omega)$ is the weight of the new estimate.

3.2.2 Detection of outlying and problematic observations and estimates

An unreliable value, particularly a significantly shifted one with a small standard error, may dramatically degrade the estimates in the next several epochs (depending on the size of the shift). To avoid such a situation two additional parameters were introduced to manipulate the weight of the new observation/estimate if there is an indication that the value is an outlier. The parameters P1 and P2 divide the results into 3 categories according to their “reliability”. The test value is now the ratio of the current value U to the usual (averaged) value V_{n-1} , i.e.

$$T = \frac{U}{V_{n-1}}. \quad (12)$$

The parameters may be set to any values, depending on how strictly the test criteria should be defined. For daily or weekly data, reasonable values are P1=1 and P2=16, while hourly data are needed to loosen the criteria to P1=1.5 and P2=30. These are the only parameters that enable us to perform the tuning of the model, although these should also be preferably set up automatically in future updates of our algorithm.

As mentioned before, P1 and P2 split all the tested values into three categories:

- 1) $T \leq P1$: the estimate is supposedly correct; it is within the boundaries of a regular occurrence or in a close vicinity (depending on P1).
- 2) $P1 < T \leq P2$: this value is outside a regular occurrence. The higher the T , the lower the weight. The weight is lowered gradually.

- 3) $P2 < T$: The value is too far from the usual occurrence; therefore, its weight $nw = 0$. This value does not influence the new estimate unless the “ignore parameter” is defined (explained in section 3.2.4 in detail), and the prediction may be used instead (depending on the options).

To ensure a seamless transition between the full and partial weights, we suggest the following piecewise function for the updated weight:

$$f(T) = \begin{cases} nw & T \leq P1 \\ g\left(1 - \frac{T-P1}{P2-P1}\right)nw & P1 < T \leq P2 \\ 0 & P2 < T \end{cases}, \quad (13)$$

where $nw = 1 - \omega$ is the default weight of the tested value and g is an arbitrary increasing bijection $f: [0,1] \rightarrow [0,1]$, e.g. $g(x) = x^2$, $g(x) = \sqrt{x}$ or simply $g(x) = x$.

The weight of the test value of the previous epoch is then updated to $\omega = f(T)$.

This method not only detects the problematic observations, but can also be used to test the final estimates. If a significant change occurs in the station's behavior, the system will adapt to it. However, a single outlier will be ignored.

The parameters $P1$ and $P2$ can be set arbitrarily according to the requirements: the higher the value, the less strict is the test.

To avoid false alarms, it is suggested to use an additional parameter to define the number of unreliable values allowed. This option is described further.

3.2.3 Testing observations and estimates

The algorithm can be used to detect outlying observations as well as estimates. The observations are usually scattered in a wider area and have a higher RMSE than the estimates. Therefore, they should not use the same test parameters $P1$ and $P2$. It is rec-

Fig. 1 Comparison of less and more strict testing of observations and estimate residuals.

Fig. 2 Adaptation to changed conditions.

ommended to set the P1 higher and the P2 lower for testing the estimates.

Figure 1 shows the difference between a less (upper) and a more (lower) strict testing of the observations (left) and estimates (right). To avoid false alarms it is suggested to use less strict values, especially for parameter P2.

Figure 2 shows the extremely changed performance of the station and the algorithm's ability to adapt to such a change. When the position estimates started to be scattered more widely, the zone of acceptable positions approximately kept its original width. However, after several days the acceptable zone gets wider, and the observations influence the estimate updates again. The figure 2 parameter settings were stricter.

3.2.4 Ignore bad observations

If an observation is supposed to be either unreliable or an outlier, it is ignored according to Eq. (13). However, a simple optional extension to the algorithm, which requires only a resettable bad observation count and a simple condition, allows us to ignore only a certain number of supposedly bad observations. After the number of consecutive bad observations exceeds a limit (henceforth called BOL), the bad or problematic observations are taken into account again to allow for the adjustment to the changed conditions.

Figure 3 shows two plots – on the left all the observations are used for updating the estimates; on the right plot, if one or two bad observations occur, they are ignored and the predicted values are output as the estimate. Only the third (and next) subsequent bad observation is taken into account for updating the estimates. This may bring some latency to adapting the estimate; on the other hand, the estimates are smoother unless a radical change happens. Figure 3 shows a station with problems during the winter months. Ignoring the bad observations made the estimates smoother. Thanks to this approach, alerts were raised when problems emerged, which was unlike the original approach that exploited all the observations regardless of their quality. It is suggested to use only lower values for the BOL parameter to keep the adaptation ability of the system.

3.3 Backward check

The system described in section 3.2 is suitable for detecting outliers or sudden position changes. If the station's motion changes its direction or velocity, the algorithm recognizes it as a long-term change and does not evaluate it as an "event".

The Kalman filter estimates the current position and velocity to pre-

Fig. 3 Ignoring bad observations can smooth the estimates. The left picture depicts the Kalman estimates of the untreated observations; the right one ignores up to two consecutive supposedly bad observations and outputs the predicted values instead.

Fig. 4 The compliance of the three tests – observation and estimate residuals and a backward projection.

dict a future position. However, the velocity parameter can be used for a “backward prediction”. Using the current estimate, the position can be projected backwards and thus be compared with the station’s position estimate (or observation if required) in the past. The number of periods and intervals for a backward check is optional. This can be particularly suitable when testing a periodically changing position. For instance, if the station oscillates around its mean position within a period of approximately 12 hours, checking 6 hours backwards may cause unwanted alerts, as the station needs 6 hours to reach a maximum change.

REFERENCES

- Bezručka, J. (2010)** *Analýza permanentných sietí GNSS v kvázi-reálnom čase a jej aplikácie. Dizertačná práca. (Near real time analysis of the permanent GNSS networks and its applications. PhD Thesis.)*. Slovak Technical University, Bratislava (in Slovak).
- Bos, R. – Bombois, X. – van den Hof, P.M.J. (2005)** *Designing a Kalman Filter When No Noise Covariance Information Is Available*. 16th IFAC World Congress, Part 1. Prague.
- Dach, R. – Hugentobler, U. – Fridez, P. – Meindl, M. (eds.) (2007)** *Bernese GPS Software Version 5*. Astronomical Institute, University of Bern, 2007.
- Haykin, S. (ed.) (2001)** *Kalman filter and Neural Networks*. Communications Research Laboratory. McMaster University, Hamilton, Ontario, Canada.
- Kalman, R.E. (1960)** *A New Approach to Linear Filtering and Prediction Problems*. Transactions of the ASME – Journal of Basic Engineering, 82 (Series D): pp. 35-45.
- Krynski, J. – Zanimonskiy, J.M. (2003)** *Towards More Reliable Estimation of GPS Positioning Accuracy*. In: *A Window on the Future of Geodesy*. IAG General Assembly Sapporo, Japan, June 30 – July 11, 2003.
- Mendenhall, W. – Sincich, T. (1988)** *Statistics For the Engineering and Computer Sciences - 2nd Edition*. Dellen Publishing Company, San Francisco; London, pp. 25-35.
- Ting, J.-A. – Theodorou, E. – Schaal, S. (2007)** *Learning an Outlier-Robust Kalman Filter*. Computational Learning & Motor Control Laboratory, University of Southern California.
- Tukey, J. W. (1977)** *Box-and-Whisker Plots*. In: *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, pp. 39-43.
- Welch, G. - Bishop, G. (2001)** *An Introduction to the Kalman Filter*. Siggraph 2001, University of North Carolina.

Figure 4 shows the compliance of the three tests – as the observations became more scattered, the estimates also appeared unexpectedly far from the predicted positions, and finally, the backward check shows that the 12-hour and 24-hour backward projections are also far from the expected values.

4 CONCLUSIONS

The paper proposes a very simple one-pass algorithm, which is easy to implement to detect sudden changes in a station’s position. It uses only one set of parameters for all the coordinate components of all the stations of a network, which makes the estimates more homogeneous, regardless of whether the original network was homogeneous or not. As the test parameters adapt to a station’s performance, the same values can be used for testing both horizontal and vertical coordinate components.

The algorithm is proposed for detecting sudden changes in a station’s position or behavior, but it can adapt to changes over the long term. The algorithm not only checks the position changes, but also evaluates the station’s behavior. Several tuning extensions to this algorithm were also introduced for a better performance and usability.

Acknowledgment

This publication is the result of project implementation: National Centre of Earth’s Surface Deformation Diagnostic in the area of Slovakia, ITMS 26220220108 supported by the Research and Development Operational Programme funded by the ERDF.