A TESTING PROCEDURE FOR SUBSIDENCE ANALYSIS

Frank Kenselaar

Delft University of Technology, Department of Mathematical Geodesy and Positioning, Thijsseweg 11, 2629 JA Delft, The Netherlands (f.kenselaar@geo.tudelft.nl)

Abstract

Land subsidence is an important topic in The Netherlands, due to its situation below sea level and the resulting large-scale waterworks this requires. Except for natural causes, like post-glacial rebound and compaction of the soft peat soil, man-made causes for land subsidence, like the extraction of ground water, oil, gas and salt, have to be monitored with great care. However, the determination of land subsidence with magnitudes of less than one centimeter per year requires high-accuracy measurement and modelling methods, and scrupulous data processing.
Therefore, the monitoring of land subsidence is based on an extensive adjustment and testing procedure of the levelling data. The optimal deformation model, e.g. a kinematic model per benchmark or a spatial-temporal model for a subsidence bowl, is determined by testing the actual solution (null hypothesis) against several alternative hypotheses, resulting in adaptation of the data and/or the functional or stochastic model in case of rejection. The paper discusses this testing procedure and various alternative hypotheses like datasnooping, identification tests, tests for deviant behavior of individual benchmarks and epoch tests. The most likely alternative hypothesis indicates an adaptation of the model or data, resulting in an improved solution acting as new null hypothesis. The testing procedure is demonstrated on a multi-epoch data set.

1. Introduction: land subsidence modelling

A typical example of land subsidence in the Netherlands consists of a smooth subsidence bowl of up to several kilometers in diameter, with a subsidence velocity of millimeters up to a few centimeters per year. In the western part of the country compaction and oxidation of the peat soil are natural causes of land subsidence, often induced or intensified by a lowering of the (very high) ground water level. Direct manmade causes for land subsidence are the extraction of water, salt, oil and gas. Especially in the northeast, the exploitation of the large Groningen gas field and its adjacent smaller fields result in smooth subsidence bowls with a typical velocity of one centimeter per year. Although maximum subsidence is only a few decimeters, subsiding areas are usually very well monitored. One of the main reasons is that the subsidence should be accompanied by careful water management to avoid the ground water level becoming too high for agricultural activities.
To monitor this low magnitude land subsidence, precise spirit levelling of benchmarks is still the primary measurement method. Other methods have been investigated, like GPS height determination (see e.g. DE HEUS ET AL. 1999) and interferometric SAR. Their accuracy under practical circumstances is however not yet high enough to replace spirit levelling. Therefore, throughout this paper precise spirit levelling measurements will be considered the main data source for monitoring and modelling of land subsidence. However, many principles of the presented testing procedure can also be applied to other observation types.

The general model for subsidence modelling based on benchmark heights reads

$$H_{i,t} = H_{i,t_0} + z_{i,t-t_0}\,, \tag{1}$$

with $H_{i,t}$ the height of benchmark i at time t, H_{i,t_0} its height at reference time t_0, and $z_{i,t-t_0}$ the subsidence since t_0, according to some specified subsidence model. If data is available from before subsidence occurred, the reference time t_0 can be adopted or estimated as the beginning of subsidence and H_{i,t_0} as the initial benchmark height before subsidence. Otherwise a reference time needs to be chosen, without any specific interpretation.

The specific subsidence model at hand is not relevant here since the presented testing procedure can be applied generally. The subsidence model can e.g. be specified per benchmark, as a kinematic function of the time lag since t_0, $z_{i,t-t_0} = z_i(t-t_0, par)$. In VERHOEF AND DE HEUS 1995 a polynomial is estimated per benchmark, if necessary including a discontinuity or breakpoint. When the behavior of the benchmarks shows a regional coherence, then it might be possible to estimate a spatial-temporal model for all benchmarks within a local area, $z_{i,t-t_0} = z(x_i, y_i, t-t_0, par)$. For subsidence modelling above deep gas reservoirs, smooth 7-parameter spatial-temporal subsidence bowls were successfully applied (HOUTENBOS 2000).

Subsidence modelling from benchmark heights is actually the last part of the data analysis. Since the benchmark heights are no measurable quantities, they have been derived from e.g. levelling networks at several epochs. The complete analysis procedure then consists of three parts (DE HEUS ET AL. 1995): 1) Single epoch analysis of the levelling networks; 2) Stability analysis of the levelling benchmarks and connection of the levelling networks to stable reference points (DE HEUS ET AL. 1994); and 3) Estimation of a kinematic subsidence model per benchmark, or a spatial-temporal model for a group of benchmarks. In each part, optimal analysis requires systematic testing for outliers and of model deficiencies.

Recently, this stepwise analysis procedure was successfully replaced by an integrated analysis, based on straightforward estimation of the subsidence model from the original levelling data (KENSELAAR AND QUADVLIEG 2001). With (1) the observation equation for a levelled spatial height difference $h_{ij,t}$, between benchmarks i and j at time t, can then be written as

$$\begin{aligned} \underline{h}_{ij,t} &= -H_{i,t} + H_{j,t} + \underline{e}_{ij,t} \\ &= -H_{i,t_0} + H_{j,t_0} - z_{i,t-t_0} + z_{j,t-t_0} + \underline{e}_{ij,t} \ , \end{aligned} \tag{2}$$

where the stochastic noise term $\underline{e}_{ij,t}$ accounts for the measurement noise and possibly model inaccuracies.

Usually, subsidence monitoring is based on levelling networks measured at discrete epochs. Within an epoch all observations share the same time label. Observation equation (2) however also allows for a strict kinematic analysis approach, where each levelling observation has its own individual time label. The data is then not organized in epochs, but considered to compose one multi-temporal kinematic network. An example of this approach is described in DE BRUIJNE ET AL. 2001.

In this paper we will assume the subsidence modelling to be based on a straightforward analysis of the original levelling observations, like with (2). We will also assume the levelling data is organized in epoch networks, where the m_k levelling observations within each epoch $k = 1 \dots K$ share the same time label t_k and are gathered in a vector h_k. Unknowns are the heights of nb benchmarks under study and np parameters of the subsidence model. According to practice we will allow that in an epoch network not all, but only nb_k benchmarks have been measured. Based on observation equations (2) the (linearized) model of observation equations can then be written as

$$\begin{pmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_k \\ \vdots \\ \underline{h}_K \end{pmatrix} = \begin{pmatrix} W_1 & & & & \\ & \ddots & & & \\ & & W_k & & \\ & & & \ddots & \\ & & & & W_K \end{pmatrix} \begin{pmatrix} P_1 & Z_1 \\ \vdots & \vdots \\ P_k & Z_k \\ \vdots & \vdots \\ P_K & Z_K \end{pmatrix} \begin{pmatrix} H_0 \\ par \end{pmatrix} + \begin{pmatrix} \underline{e}_1 \\ \vdots \\ \underline{e}_k \\ \vdots \\ \underline{e}_K \end{pmatrix}, \tag{3}$$

with: h_k vector of m_k (linearized) levelled height differences in the levelling network at epoch k;
W_k $m_k \times nb_k$ matrix relating the levelled height differences and benchmark heights at epoch k (the levelling network matrix);
P_k $nb_k \times nb$ permutation matrix, selecting the nb_k benchmarks occupied at epoch k from the complete set of nb benchmarks;
Z_k $nb_k \times np$ coefficient matrix of the (linearized) subsidence model;
H_0 vector of nb (linearized) unknown benchmark heights at reference time t_0;
par vector of np (linearized) unknown parameters of the subsidence model;
$\underline{e}_k$ vector of m_k unknown stochastic noise terms;

In (3) the number of observations is $\sum_{k=1}^{K} m_k$ and the number of unknowns is $nb + np$.

2. Hypothesis testing

The subsidence analysis is based on determination of the optimal subsidence model fitting through the levelling data, meanwhile identifying outliers in the data. The precise levelling measurements and low magnitude subsidence require scrupulous data processing in order to discriminate between significant subsidence, observation noise, outliers and model imperfections. Therefore a combination of least-squares estimation and statistical hypothesis testing is advised.

In (3) the levelling data of all epochs available is written as a linear(ized) model of observation equations,

$$\underline{y} = Ax + \underline{e}\,; \quad Q_y\,, \tag{4}$$

with y a vector of levelling observations, Ax the model of observation equations, with the unknown vector x containing the benchmark heights as well as the parameters of the subsidence model, $\underline{e}$ the vector of stochastic noise terms and Q_y the variance-covariance matrix describing their dispersion. Except for the quite well known levelling measurement precision, the variance-covariance matrix could also account for the limited accuracy of the subsidence model to describe the actual land subsidence (KENSELAAR AND QUADVLIEG 2001).
From (4), minimum variance unbiased estimates for both x and e can be obtained by least-squares adjustment. In case the subsidence model consists of non-linear functions in the unknown parameters, the model needs to be linearized and an iterative least-squares solution is required. Statistical testing is based on the least-squares residuals and their variance-covariance matrix, which can be computed as (see e.g. TEUNISSEN 2000A).

$$\underline{\hat{e}} = P_A^{\perp}\,\underline{y}\,; \quad Q_{\hat{e}} = P_A^{\perp} Q_y\,, \quad \text{with} \quad P_A^{\perp} = I - A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1}\,. \tag{5}$$

In hypothesis testing, the actual model and data – called the null hypothesis (H_0) – is tested against one or more alternative hypotheses (H_A), specifying a model adaptation or possible error(s). The alternative hypothesis is specified as a linear extension of the null hypothesis with q additional error components, q being the dimension of the test,

$$\mathrm{H}_0\!: \;\; \underline{y} = Ax + \underline{e}\,, \quad \text{versus} \quad \mathrm{H}_A\!: \;\; \underline{y} = Ax + C\nabla + \underline{e}\,, \tag{6}$$

where $C\nabla$ is a linear(ized) model extension with vector ∇ containing q unknown (suggested) errors and C describing their influence on the observables.
One is likely to reject the null hypothesis, in favor of the alternative hypothesis, if the estimated errors are significant. The general test-statistic for testing the null hypothesis against a specific alternative hypothesis can be computed from the least-squares residuals (5) under the null hypothesis (TEUNISSEN 2000B):

$$\begin{aligned}\underline{T}_q &= \underline{\hat{\nabla}}^T Q_{\hat{\nabla}}^{-1} \underline{\hat{\nabla}} \\ &= \underline{\hat{e}} Q_y^{-1} C (C^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C)^{-1} C^T Q_y^{-1} \underline{\hat{e}}.\end{aligned} \tag{7}$$

Assuming the observables obey a Normal distribution, the test-statistic is Chi-squared distributed with q degrees of freedom. Thus, the null hypothesis is rejected, in favor of the alternative hypothesis, if $T_q > \chi^2_{\alpha_q}(q)$, with α the level of significance.

For alternative hypotheses assuming only one error (q = 1), ∇ reduces to a scalar, C to a vector c, and (7) can be simplified to

$$\underline{T}_1 = \frac{(c^T Q_y^{-1} \underline{\hat{e}})^2}{c^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c}. \tag{8}$$

Before making any further inferences, a so called overall model test of the null hypothesis is computed, as a general check whether the data fits the subsidence model within the tolerances as specified in the stochastic model. This test quantity can easily be computed as the length of the vector of least-squares residuals. The dimension of the test equals the redundancy (r) of the model:

$$\underline{T}_r = \underline{\hat{e}}^T Q_y^{-1} \underline{\hat{e}} \sim \chi^2(r) \quad \text{with} \quad r = \textstyle\sum_{k=1}^{K} m_k - (nb + np). \tag{9}$$

As long as the overall model test is rejected, identification of the most likely error is supported by testing the null hypothesis against various alternative hypotheses.

3. Alternative hypotheses for subsidence analysis

In this section we will specify some alternative hypotheses that can be useful in subsidence analysis. All alternative hypotheses will be written as linear extensions of the null hypothesis (3),

$$\begin{pmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_k \\ \vdots \\ \underline{h}_K \end{pmatrix} = \begin{pmatrix} W_1 & & & & \\ & \ddots & & & \\ & & W_k & & \\ & & & \ddots & \\ & & & & W_K \end{pmatrix} \begin{pmatrix} P_1 & Z_1 \\ \vdots & \vdots \\ P_k & Z_k \\ \vdots & \vdots \\ P_K & Z_K \end{pmatrix} \begin{pmatrix} H_0 \\ par \end{pmatrix} + \begin{pmatrix} C_1 \\ \vdots \\ C_k \\ \vdots \\ C_K \end{pmatrix} \nabla + \begin{pmatrix} \underline{e}_1 \\ \vdots \\ \underline{e}_k \\ \vdots \\ \underline{e}_K \end{pmatrix}, \tag{10}$$

with ∇ a vector containing q suggested unknown errors or model adaptations and the $m_k \times q$ matrices C_k specifying their influence on the levelling observations of epoch k. The following types of hypothesis tests correspond with specific choices of q and C_k.

Observation test
With this type of hypothesis we test for an individual error in the levelling data. If we suspect levelling observation $h_{ij,k}$, all matrices $C_1 \ldots C_K$ in (10) reduce to zero vectors, except for C_k that reduces to a vector

$$c_{ij,k} = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 1 \end{pmatrix}^T, \tag{11}$$

with the 1 in the m_k-vector $c_{ij,k}$ at the position corresponding with the suspected levelling observation in epoch k. Systematic testing of all $\sum_{k=1}^{K} m_k$ observations is often referred to as datasnooping.

Identification test

In this type of alternative hypothesis a deviating behavior of a specific benchmark in only one epoch is suggested. In case of benchmark i in epoch k, all matrices $C_1 \dots C_K$ in (10) reduce to zero-vectors, except for C_k that reduces to a vector

$$c_{i,k} = W_k P_k c_i \ , \quad \text{with} \quad c_i = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 1 \end{pmatrix}^T , \tag{12}$$

with the 1 in the nb-vector c_i at the position corresponding with benchmark i. Systematic testing of all benchmarks in all epochs implies K times nb tests. However, if a benchmark is not measured in epoch k, $c_{i,k}$ reduces to a zero-vector and no test-statistic can be computed. The number of relevant tests reads $\sum_{k=1}^{K} nb_k$.

Point test

This type of alternative hypothesis investigates whether a benchmark has a significant deviating behavior throughout all epochs. For benchmark i, the matrices C_k, k = 1…K in (10) have dimension $m_k \times q_i$ and can be written as

$$C_{i,k} = \begin{pmatrix} 0 & \cdots & 0 & W_k P_k c_i & 0 & \cdots & 0 \end{pmatrix}, \tag{13}$$

with c_i as in (12) and the vector $W_k P_k c_i$ as column number $k = 1 \dots q_i$ of $C_{i,k}$. The dimension of the test (q_i) equals the redundancy of the benchmark, i.e. the number of epochs K minus the minimal number of epochs needed to determine its height and any benchmark specific subsidence parameters. Since $W_k P_k c_i$ becomes a zero-vector for epochs where the benchmark is not measured, the matrix could contain zero-columns, making computation of the test-statistic impossible. Elimination of these zero-columns and their corresponding errors will reduce q_i to its true value and solve the problem.

The point test can be considered an overall test per benchmark since the dimension of the test equals the benchmark redundancy, while no specific causes of the deviation are defined. In total nb point test-statistics can be computed.

As an example, consider a subsidence model assuming an individual, constant velocity per benchmark. Per benchmark then two unknowns need to be estimated (benchmark height and velocity). If the benchmark is measured in all epochs, then q_i = K−2 and the model extension of the alternative hypothesis (10) reads

$$\begin{pmatrix} W_1 P_1 c_i & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & W_k P_k c_i & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & W_{q_i} P_{q_i} c_i \\ 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \nabla_{i,1} \\ \vdots \\ \nabla_{i,k} \\ \vdots \\ \nabla_{i,q_i} \end{pmatrix} .$$

Epoch test

This type of alternative hypothesis acts as an overall test per epoch. It indicates that a complete epoch deviates from the model, without further specification. In case of epoch k, all matrices $C_1 \dots C_K$ in (10) are zero-matrices, except for C_k that equals the $m_k \times nb_k$ matrix

$$C_k = W_k \ . \tag{14}$$

Considering each epoch, K epoch test-statistics can be computed. The large number of additional errors estimated ($q = m_k$) may cause rank deficiency problems. An epoch test can only be computed if the subsidence model could also be determined from the data without the epoch considered.

Depending on the subsidence model at hand various, types of alternative hypotheses can be designed to test whether specific model extensions would yield a significant improvement. Two examples:

Breakpoint test
In case the subsidence model consists of an individual temporal polynomial function per benchmark, one could think of extension with one or more higher order terms, or inclusion of a discontinuity in the benchmark behavior, called breakpoint in VERHOEF AND DE HEUS 1995. For example, starting with a subsidence model assuming an individual, constant velocity per benchmark, we can test the null hypothesis against an alternative hypothesis suggesting a breakpoint at epoch time t_b. For benchmark i, all matrices C_k in (10) for $t_k \leq t_b$ reduce to zero-vectors, while for $t_k > t_b$ matrices C_k reduce to vectors

$$c_{i,bk} = (t_k - t_b) W_k P_k c_i \,, \tag{15}$$

with c_i as in (12). This alternative hypothesis still assumes a linear subsidence for this benchmark, but with a different velocity before and after the breakpoint. Systematically, each of the nb benchmarks can be tested for a breakpoint in each (but the first and last) epoch.

ALB test
Using a spatial-temporal subsidence model as null hypothesis, one could for instance test for a specific deviation of a benchmark from the regional coherent model. For example, a subsidence model assuming per benchmark a constant, but spatially varying velocity, can be tested for benchmarks with velocities that differ significantly from the spatial model. In KENSELAAR AND MARTENS 2000, this is introduced as so-called 'autonomous linear behavior' (ALB). Taking benchmark i, all matrices C_k in (10) reduce to vectors

$$c_{i,1k} = (t_k - t_1) W_k P_k c_i \,, \tag{16}$$

with c_i as in (12) and t_1 the time label of the first epoch. The ALB is assumed to be independent of an eventual beginning of the spatial-temporal subsidence model at t_0. Systematically, each of the nb benchmarks can be tested for ALB.

4. Estimation and testing procedure

In geodetic positioning problems the model is usually known very precisely. The data processing therefore focuses on blunder detection in the observations. The testing procedure can often be limited to datasnooping (observation tests). For subsidence analysis however, the model has a physical character and is only known up to an approximation of the true local behavior of the earth's surface. In data processing, blunder detection and the determination of the best fitting model can not be fully discriminated. The testing procedure should therefore consist of several types of alternative hypothesis, testing the data and both the functional and stochastic model. Since most types of tests are sequentially computed for all observations, benchmarks or epochs, the null hypothesis is opposed against a large number of alternatives.
Identification of the most likely alternative hypothesis is trivial as long as all tests have the same number of degrees of freedom. The largest (rejected) test-statistic then points to the first candidate for adaptation of the data or the model. For example, from the previous section all test-statistics of the observation test, identification test, breakpoint test and ALB test follow a Chi-square

distribution with one degree of freedom and share the same critical value $\chi^2_{\alpha_1}(1)$. However, the point test, epoch test and overall model test have different degrees of freedom and different distributions. Their test-statistics can not be compared straightforwardly. In DE HEUS ET AL. 1994 it was therefore proposed to compare the quotients of test-statistic and critical value instead, $T_q \big/ \chi^2_{\alpha_q}(q)$, while fixing the power of all tests to 50%. A test-quotient larger than one is then rejected and the largest test-quotient indicates the most likely alternative hypothesis amongst all tests performed.

The testing procedure has to be completed stepwise, where in each step the data, the functional model or the stochastic model is adapted, as indicated by the largest test-quotient. The improved model is adjusted and tested again, until all tests are accepted. The estimation and testing procedure can be summarized as follows:

1. ***(Iterative) least-squares adjustment*** of the null hypothesis subsidence model (3);
2. ***Computation of test-statistics*** for the overall model test (9) and numerous alternative hypothesis using (7) and (8), e.g. a selection from the types of alternative hypotheses (11) – (16);
3. ***Identification*** of the most likely alternative hypothesis by selection of the largest rejected test-quotient (as long as the null hypothesis model is rejected);
4. ***Adaptation*** of the subsidence model following the suggestion of the most likely alternative hypothesis, the improved model enters as new null hypothesis in step 1.

The first three steps can be automated completely. Automation of the fourth step is only possible for simple adjustment problems, e.g. when only datasnooping is involved. For a complicated problem like the determination of a subsidence model, possible improvements concern the data, the functional model, as well as the stochastic model. This requires both experience and specialist knowledge of the measurement process and the subsidence problem at hand. Therefore, the analyst has to make an expert decision, although the identified most likely alternative hypothesis gives valuable suggestions:

- The ***observation test*** suggests an error in a levelled height difference. When identified as most likely alternative hypothesis, the observation can be excluded from the data set or reweighted to diminish its influence, assuming the error source can not be reconstructed.
- The ***point test*** acts as an unspecified identifier for a benchmark deviating from the subsidence model. Other benchmark related tests, like the identification, breakpoint or ALB test, investigate specific types of deviations. When marked as most likely alternative hypothesis, the benchmark can be excluded from the data set in all epochs. Graphical visualization of the benchmark behavior throughout all epochs could suggest a more specific alternative hypothesis.
- The ***identification test*** marks a benchmark to deviate from the model in one of the epochs. When pointed as most likely alternative hypothesis, the benchmark can be omitted from the data set in that epoch, by combining its related observations to new ones. In this way the benchmark is excluded, still using all the levelling data. Further data screening could e.g. reveal that in one of the epoch networks a wrong benchmark (number) was used.
- The ***breakpoint*** and ***ALB tests*** suggest that for a specific benchmark the subsidence model should be extended with one or more parameters. When identified as most likely alternative hypothesis, one should actually adapt the functional model and estimate a breakpoint or ALB velocity for the benchmark concerned. Again, graphical visualization of the benchmark behavior can support this decision.
- The ***epoch test*** acts as an unspecified identifier for an epoch that is significantly deviating from the subsidence model. When the largest rejected test-quotient concerns an epoch test, further investigation is necessary since usually one is not willing to exclude a complete epoch

from the data set. Rejection of the epoch test could e.g. be caused by remaining systematic errors in the levelling network, or the subsidence model assumed fails, specifically concerning the temporal behavior. It could also be a reason for adaptation of the stochastic model, e.g. increasing the measurement standard deviation for that epoch.

- In case the ***overall model test*** has the largest rejected test-quotient, the data and model just do not match, without a specific alternative hypothesis indicating the problem. Further investigation is necessary. For example, too optimistic assumptions have been made for the stochastic model. A subsidence model that is only a rough approximation of the physical reality can not be expected to fit the data within the narrow margins of levelling measurement precision only.

5. Real data example

In this section the presented testing procedure will be demonstrated on a relatively small and convenient, multi-epoch data set. The data is gathered to monitor land subsidence above the so-called Roswinkel gas field. This is a small, isolated gas field south of the large Groningen field. Five epochs of levelling networks were measured in 1980, 1985, 1990, 1994 and 1997, the first one before gas production in 1983 (see figure 1). Note that the epoch levelling networks have very low redundancy. Thus, single epoch analysis of the network will yield little information about the data quality. However, the effects of deep gas extraction on the benchmark heights can be well approximated by a continuous spatial-temporal subsidence model as described in HOUTENBOS 2000, KENSELAAR AND QUADVLIEG 2001. This model assumes a smooth subsidence bowl with ellipsoidal contour lines, where each point has a constant subsidence velocity that is exponentially decreasing with the distance to the center of the subsidence bowl. Except for the initial benchmark heights, seven parameters can be estimated, determining the subsidence bowl:

t_0 the time of beginning of subsidence;
s the linear subsidence velocity in the center of the subsidence bowl;
x_0 and y_0 the position of the center of the subsidence bowl;
a, b and φ the half long and short axes and the orientation of the subsidence ellipse.

The unknown benchmark heights and subsidence parameters are estimated straightforwardly from the levelling data of all epochs.

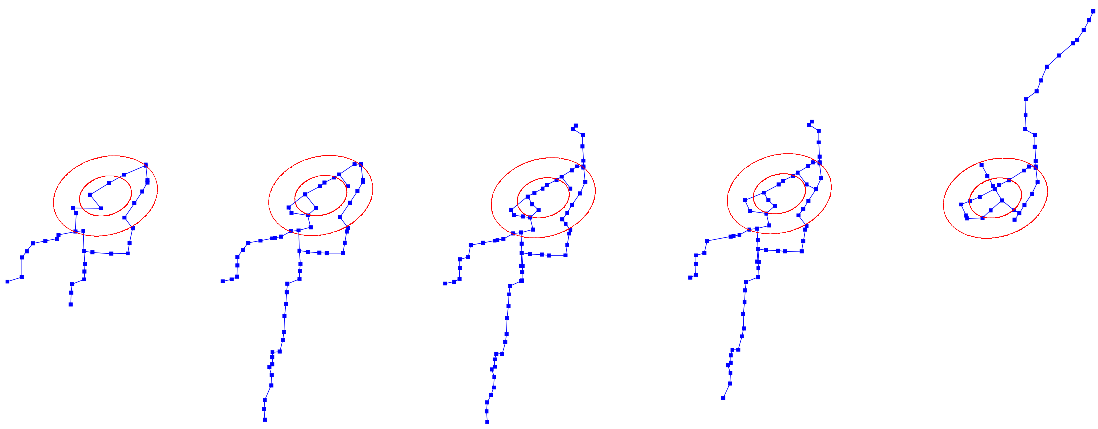

Figure 1: The Roswinkel data set, from left to right the epoch networks of 1980, 1985, 1990, 1994 and 1997, and the contours of the subsidence bowl.

Initially the stochastic model only accounted for the levelling measurement precision. For the standard deviation of the levelling observations an a-priori value of 0.7 mm/√km was assumed for all epoch networks. In the stepwise estimation and testing procedure the actual null hypothesis was tested with an overall model test and the following alternative hypotheses: observation tests (for each levelling observation); identification tests (for each benchmark in each epoch); point tests (for each benchmark); ALB tests (for each benchmark); and epoch tests (for each epoch).

Successive adaptations in the data or the model were undertaken, based on evaluating the alternative hypothesis with the largest rejected test-quotient. In table 1 a summary is given of the successive test results and model adaptations.

In the first four steps, the largest test-quotients concerned benchmarks that clearly deviated from the subsidence model. In three cases the benchmark showed so-called autonomous linear behavior (ALB), i.e. the benchmark velocity is still constant, but significantly deviating from its velocity according to the subsidence model. The model was extended with an additional velocity parameter for such a benchmark. As an example, the estimated velocity for benchmark 18A0090 is 1.3 mm/year lower than according to the subsidence model. One benchmark deviates from the model in a more complex way and was eliminated from the data set. Isolated misbehaving benchmarks are often ascribed to other causes than gas extraction, like local groundwater lowering.

Step	Alternative hypothesis with largest test-quotient (value between brackets)	Model adaptation
1	ALB test benchmark 18A0090 (57.14)	additional velocity parameter
2	ALB test benchmark 18A0006 (27.43)	additional velocity parameter
3	Point test benchmark hp0650 (26.57)	benchmark excluded
4	ALB test benchmark 18A0089 (10.97)	additional velocity parameter
5	Epoch test network 1980 (5.78)	point noise, $\sigma = 0.6$ mm/√year
6	Observ. test 18C0095 – 18C0152 in 1994 (1.67)	observation excluded
7	Epoch test network 1980 (1.08)	for epoch 1980 $\sigma = 0.8$ mm/√km
8	Largest test-quotient is 0.98 (overall model test)	accept model

Table 1: Summary of successive test results and model adaptations.

In step 5 however, the epoch test for 1980 was pointed as most significant. Since the test-statistic still exceeded the critical value with a factor 5.78, it was concluded that the subsidence model would not fit the data within relaxation for levelling measurement precision only. The stochastic model was extended by accounting for so-called point noise, allowing a stochastic variation of the benchmark heights, with a standard deviation of 0.6 mm/√year (see KENSELAAR AND QUADVLIEG 2001).
Adding the additional point noise terms to the stochastic model significantly decreased the test-quotients. Their maximum now pointed to an observation, that was excluded from the data. In step 7, epoch 1980 was again marked as problematic, with a test-quotient just over one. Slightly increasing the measurement standard deviation to 0.8 mm/√km, for just this first epoch network, resulted in an accepted model.

In this paper we concentrate on the testing procedure and not on the subsidence modelling itself. It suffices to remark that the redundancy of the accepted model was 134 (225 observed height differences, 81 unknown initial benchmark heights, 7 unknown parameters of the subsidence bowl and 3 individual benchmark velocities). The estimated subsidence velocity in the center of the subsidence bowl was 10.1 mm/year, with a standard deviation of 0.2 mm/year. The standard deviation of the time of beginning of subsidence was 80 days and the standard deviations of the position and size of the subsidence bowl were in the order of 50 meter. In this particular (quite convenient) case the remaining discrepancies between the data and the smooth subsidence trend model were within half a centimeter.

6. Concluding remarks

The described estimation and testing procedure has been implemented in several software packages. On various data sets the procedure has been proven a successful tool in processing levelling data for subsidence analysis. The largest test-quotients supply valuable suggestions for adaptation of the data or extension of the model. The actual adaptation however requires expert knowledge about the levelling data, the subsidence phenomenon, benchmark foundation and local circumstances. Often, the adopted subsidence model is a too approximate assumption to fit the very precise levelling observations. Extension of the stochastic model will then be necessary. Since a stochastic description of these model inaccuracies is often hardly known, various choices could lead to different results. Practically, one could complete the stepwise testing procedure following different adaptation strategies, and investigate whether their resulting subsidence models differ significantly. A more theoretical approach would be to combine the testing procedure with estimation of the stochastic model, e.g. with variance component estimation techniques.

This research was performed in close cooperation with the Geomatics Department of the Nederlandse Aardolie Maatschappij (NAM), who also placed the Roswinkel data at our disposal. Further research is sponsored by Delft University DIOC-funding.

References

Bruijne, A.J.T., F. Kenselaar and F. Kleijer (2001). Kinematic deformation analysis of the first order benchmarks in The Netherlands. In: *Proceedings of the 10th FIG International Symposium on Deformation Measurements*, Orange, California, March 19-22.

Heus, H.M. de, M.H.F. Martens, H.M.E. Verhoef (1994). Stability-analysis as part of the strategy for the analysis of the Groningen gas field levellings. In: *Proceedings of the Perlmutter Workshop on Dynamic Deformation Models*, Haifa, pp 259-272.

Heus, H.M. de, P. Joosten, M.H.F. Martens, H.M.E. Verhoef (1995). Strategy for the analysis of the Groningen gas field levellings - an overview. In: *Land Subsidence, proceedings of the IAHS Fifth International Symposium On Land Subsidence (FISOLS)*, the Hague (ISBN 0-947571-74-4), pp.301-311.

Heus, H.M. de, M.H.F. Martens, H. van der Marel (1999). Re-weighting of GPS baselines for vertical deformation analysis. In: *Proceedings of the 22nd General Assembly of the IUGG*, Birmingham.

Houtenbos, A.P.E.M. (2000). The quantification of subsidence due to gas-extraction in the Netherlands. In: *Land Subsidence, proceedings of the IAHS Sixth International Symposium On Land Subsidence (SISOLS)*, Italy (ISBN 88-87222-06-1), pp. 177-189.

Kenselaar, F., M.H.F. Martens (2000). Spatial-temporal modelling of land subsidence due to gas extraction. In: *Land Subsidence, proceedings of the IAHS Sixth International Symposium On Land Subsidence (SISOLS)*, Italy (ISBN 88-87222-06-1), pp. 383-396.

Kenselaar, F. and R.C.H. Quadvlieg (2001). Trend-signal modelling of land subsidence. In: *Proceedings of the 10th FIG International Symposium on Deformation Measurements*, Orange, California, March 19-22.

Verhoef, H.M.E. and H.M. de Heus (1995). On the estimation of polynomial breakpoints in the subsidence of the Groningen gas field. *Survey Review*, Vol.33, No.2, pp.17-31.

Teunissen, P.J.G. (2000). *Adjustment theory, an introduction*, Series on Mathematical Geodesy and Positioning, Delft University Press.

Teunissen, P.J.G. (2000). *Testing theory, an introduction*, Series on Mathematical Geodesy and Positioning, Delft University Press.