Integrating Leveling and InSAR Networks for the Monitoring of Subsidence Due to Gas Extraction

Dennis Odijk

1 Introduction

Due to its situation below sealevel, the monitoring of land subsidence is an important issue in the Netherlands. In the northeast, the exploitation of the large Groningen gas field and its smaller adjacent fields has resulted in smooth subsidence bowls up to several kilometers in diameter with a velocity of typically one centimeter per year. Since the beginning of gas production in the 1960s spirit leveling surveys have been carried out on a regular basis. Although this technique allows for the detection of very small surface deformations, it is rather time-consuming and thus expensive. The introduction of space-geodetic techniques like GPS and the interferometric use of Synthetic Aperture Radar (InSAR) has offered new opportunities for precise deformation monitoring. In particular, using the InSAR technique at relatively low costs (when compared to leveling) relatively large subsidence areas can be monitored. InSAR provides vertical displacements between coherent points in two SAR acquisitions (images). Additional benefits of InSAR are that it is not necessary to physically access the deformation area and the high spatial and temporal density of the data. Besides, sub-centimeter accuracy has been reported for InSAR-derived surface deformations, see e.g. Massonnet and Feigl (1998).

In this article it is investigated to what extent an integration of leveling observations and vertical displacement data derived from InSAR results in a more efficient deformation monitoring system than based on leveling data only. GPS-derived vertical displacements are not considered here because of their lower spatial density compared to InSAR. Simulations have been carried out to study the performance of an integration of leveling and InSAR vertical displacement data for subsidence estimation. Aspects that play a role in these design computations are the spatial and temporal density of the leveling and InSAR networks, in combination with the assumptions concerning the stochastic properties of the observations. Before discussing these simulations, in the next sections an extensive overview is provided of the procedure we applied for the mathematical modeling of land subsidence due to gas extraction.

2 A spatio-temporal model for land subsidence

The developed procedure for the monitoring in gas field areas is based on a spatio-temporal modeling of the subsidence. This section describes this modeling in detail.

2.1 Vertical deformations in gas field areas

In this paper we consider slow-term land subsidence due to gas extraction, which ranges from a few millimeters up to a few centimeters a year. However, a point in such a subsidence area may also be subject to movements due to other causes (not due to gas extraction). In presence of all kinds of deformations, the height of a point \( i \) at time \( t \), denoted as \( H_{i,t} \), can be written as the following function (Kenselaar and Quadvlieg 2001):

\[
H_{i,t} = H_{i,t_0} + \zeta_{i,t-t_0} + \eta_{i,t},
\]

where \( t_0 \) denotes the start time of the deformation, \( H_{i,t_0} \) the height of the point before or at the beginning of deformation (initial height), \( \zeta_{i,t-t_0} \) the motion of the point due to gas extraction, and \( \eta_{i,t} \) the motions of the point that are not caused by gas extraction. Motions not caused by gas extraction are usually due to local effects (for example ground water variations).
The underscore in Eq. (1) emphasizes the stochastic nature of a variable. The deformation due to gas extraction, i.e. $g_{i,t-t_0}$, is assumed as a stochastic process here, since in practice it is usually impossible to completely model it by a deterministic function. In Sect. 2.2 this so-called trend-signal modeling is discussed in detail. The deformation of the points that is not caused by gas extraction, $\eta_{i,t}$ in Eq. (1), is also assumed as a stochastic process. Suppose that we have two epochs, denoted as $t_1$ and $t_m$, then the following covariance function is used to describe the temporal behavior of local point instabilities:

$$\sigma_{\eta_{i},\eta_{i,t-m}} = \sigma^2_B (\min[t_1,t_m] - t_1),$$  \hspace{1cm} (2)

where $\min[\cdot,\cdot]$ denotes the minimum operator, $t_1$ the first epoch at which the measurements are obtained and $\sigma_B$, a point-specific standard deviation (also referred to as point noise). From Eq. (2) note that the point instabilities are allowed to become larger with time. In fact, they are modeled as a random-walk process. These point instabilities are uncorrelated in space assumed.

\subsection{2.2 Modeling deformations due to gas extraction}

In our procedure for subsidence estimation the vertical deformation due to gas extraction is decomposed into a deterministic trend, denoted as $z$, plus a stochastic signal, denoted as $\xi$.

$$g_{i,t-t_0} = z_{i,t-t_0} + \xi_{i,t-t_0}. \hspace{1cm} (3)$$

The purpose of the estimation procedure is to solve for this sum of trend and signal. For the application of subsidence in gas field areas that is slow-term, spatio-temporal models have proven to be valid for both trend and signal.

\subsubsection{2.2.1 Modeling the deformation trend}

It is known that subsidence of the surface above deep gas reservoirs shows strong coherence, both in space and time. In among others (Houtenbos 2000) it was shown that for the deformation trend the adoption of smooth 7-parameter spatiotemporal subsidence bowls can be applied. In this paper we adopt this method. The subsidence trend is modeled as a superposition of $n_B$ bowls, in which the center of each bowl subsides with an unknown but constant velocity. In addition, it is assumed that this subsidence velocity exponentially decreases with increasing distance from the center of the bowl. Mathematically, we may then write the trend $z_{i,t-t_0}$ in Eq. (3) as:

$$z_{i,t-t_0} = \sum_{B=1}^{n_B} z_{i,t-t_0,B}, \text{ where } z_{i,t-t_0,B} = \begin{cases} 0 & \text{for } t \leq t_{0,B} \\ v_B (t-t_{0,B}) \exp \{-1/2 r_i^2/\sigma^2_B\} & \text{for } t \geq t_{0,B} \end{cases},$$  \hspace{1cm} (4)

where $t_{0,B}$ denotes the start time of subsidence of bowl $B$, $v_B$ the subsidence velocity in the center of this bowl, and $r_i$ the standardized radius from the center of this bowl to point $i$. The subsidence velocity at point $i$ can be computed as $v_B \exp \{-1/2 r_i^2/\sigma^2_B\}$. It is assumed that each bowl has an ellipsoidal shape in the horizontal plane, such that the standardized radius can be written as follows:

$$r_i^2 = \left( \frac{(x_i-x_{c,B})\sin \phi_B + (y_i-y_{c,B})\cos \phi_B}{a_B} \right)^2 + \left( \frac{(x_i-x_{c,B})\cos \phi_B - (y_i-y_{c,B})\sin \phi_B}{b_B} \right)^2,$$  \hspace{1cm} (5)

where $(x_i, y_i)$ denote the (known) horizontal coordinates of point $i$, $(x_{c,B}, y_{c,B})$ the coordinates of the center of bowl $B$, $(a_B, b_B)$ the half long respectively half short axes of this bowl, and $\phi_B$ the argument of the long axis of this bowl. To recover the deformation trend we therefore need to determine seven parameters for each bowl: the two temporal parameters $v_B$ and $t_{0,B}$, plus the five spatial parameters $(x_{c,B}, y_{c,B}, a_B, b_B$ and $\phi_B$.

\subsubsection{2.2.2 Modeling the deformation signal}

Unfortunately, for many gas fields it is not possible to describe the land subsidence due to gas extraction using only the trend model. For example, for the large and complex Groningen gas field in the Netherlands a deterministic trend only turns out to be insufficient (Houtenbos 2000). Hence, the signal term $\xi_{i,t-t_0}$ in Eq. (3) accounts for the imperfection of the adopted trend model with respect to the actual land subsidence. This term is also referred to as model noise, see e.g. Kenselaar and Quadflieg (2001). In our procedure these discrepancies are modeled as a stochastic process with both temporal and spatial components. Conform Houtenbos (2000) it is assumed that in the temporal domain the signal can be described by a random-walk process, while in the spatial domain it can be modeled by a Gaussian covariance function, this latter function is an exponential depending on the squared distance between points, see e.g. Wackernagel (1998). Suppose we have two points, denoted as $i$ and $j$, and two epochs, denoted as $t_1$ and $t_m$, then the covariance between the signal at point $i$ at epoch $t_1$ and the signal at point $j$ at epoch $t_m$ is modeled as:

$$\sigma_{\xi_{i,t_1}\xi_{j,t_m}} = \sigma^2_S (\min[t_1,t_m] - t_1) \exp \{ - (l_{ij}/L_S)^2 \}. \hspace{1cm} (6)$$

In this covariance function $\sigma^2_S$ denotes the standard deviation of the signal, $t_1$ the first measurement epoch, $l_{ij}$ the horizontal distance between points $i$ and $j$ and $L_S$ the correlation distance of the signal.

\section{3 Observation equations}

To estimate subsidence by means of the trend-signal model, in this section it will be connected to two types of observations: spirit leveling measurements and vertical displacements as derived from modern space-geodetic...
techniques like InSAR or GPS. This section deals with their observation equations.

3.1 Leveling data

In case of conventional leveling, an observed height difference between two points \(i\) and \(j\) on epoch \(t\), denoted as \(w_{ij,t}\), can be written as:

\[
w_{ij,t} = -H_{ij,t} + H_{j,i,t} + \xi_{ij,t},
\]

where \(H_{ij,t}\) denotes the height of point \(i\) at time \(t\) and \(H_{j,i,t}\) the height of point \(j\) at time \(t\). Random measurement errors are contained in the term \(\xi_{ij,t}\). By inserting Eq. (1) into Eq. (7), we may rewrite the observation equation as function of initial heights, vertical deformations and a term due to noise:

\[
w_{ij,t} = -H_{ij,t} + H_{j,i,t} - z_{ij,t-l_0} + g_{ij,t-l_0} - n_{ij,t} + n_{ji,t} + \xi_{ij,t}.
\]

 Decomposing the vertical deformation due to gas extraction into a trend and signal term, see Eq. (3), we obtain:

\[
w_{ij,t} = -H_{ij,t} + H_{j,i,t} - z_{ij,t-l_0} + g_{ij,t-l_0} - n_{ij,t} + n_{ji,t} + \xi_{ij,t} + \eta_{ij,t} - \xi_{ij,t}.
\]

where all stochastic terms, i.e. the deformation signal \(\xi\), point noise \(\eta\) and measurement noise \(\xi\) are collected in one random error term, denoted as \(\xi_{ij,t}\). Note that since the mathematical expectation of all three individual components is zero, also the combined noise term has zero expectation. A common assumption for the measurement noise of leveling data is to assume them normally (Gaussian) distributed and uncorrelated, both in space and time. In addition, the measurement noise is proportional to the length of the measured trajectory, denoted as \(s_{ij}\):

\[
\sigma_{w_{ij}}^2 = \sigma_v^2 s_{ij}.
\]

In case of precise leveling data, the standard deviation is usually within the interval \(0.5 < \sigma_v < 1 \text{ mm}/\sqrt{\text{km}}\).

3.2 Vertical displacement data

3.2.1 The general case

Vertical displacements derived from space-geodetic techniques like InSAR or GPS are nothing else than temporal differences of two sets of heights of a collection of points. In case of InSAR we may have vertical deformations for the time interval between two SAR acquisitions. In case of GPS we may have the differences between the heights of the stations of a permanent network at two epochs.

Note that a set of heights within a SAR image or corresponding to GPS points at a certain epoch are never absolute heights, but always relative to a (non-stochastic assumed) reference point. Since it is not always known which point is selected as reference and whether this point is also reference for another epoch, in our estimation procedure the heights of the reference points are additional unknown parameters. Suppose that at epoch \(t_1\) we have a set of heights which are relative to point A, denoted as \(h_{A,i,t_1}\), while at epoch \(t_m\) we have a set of heights which are relative to point B, denoted as \(h_{B,i,t_m}\) (see Fig. 1), then for a point \(i\) the difference between the two heights, denoted as \(v_{i,t_1,t_m}\), can be written as:

\[
v_{i,t_1,t_m} = -h_{A,i,t_1} + h_{B,i,t_m} + \xi_{i,t_1,t_m}.
\]

where the measurement noise of this vertical displacement observation is contained in the term \(\xi_{i,t_1,t_m}\). Inserting \(h_{A,i,t_1} = h_{A,t_1} - h_{A,t_0}\) and \(h_{B,i,t_m} = h_{B,t_m} - h_{B,t_0}\) yields:

\[
v_{i,t_1,t_m} = -h_{A,t_1} + h_{B,t_m} + h_{A,t_1} - h_{B,t_0} + \xi_{i,t_1,t_m},
\]

where \(h_{A,t_1}\) denotes the height of reference A at time \(t_1\) and \(h_{B,t_m}\) the height of reference B at time \(t_m\). However, this observation equation cannot be used yet to estimate deformations, since one should realize that GPS or InSAR-derived heights are geometric heights, i.e. they are heights above a certain reference ellipsoid, e.g. WGS84. When InSAR- or GPS-derived vertical displacement data are integrated with leveling data, the observation equation should – like the leveling observation equation (7) – be expressed as orthometric heights, i.e. heights with respect to the geoid. Both geometric heights and orthometric heights are connected as follows, for example the heights of point \(i\) at epoch \(t\):

\[
h_{i,t} = h_{i,t} + N_{i,t}.
\]

Here \(N_{i,t}\) denotes the height of the geoid above the reference ellipsoid (geoid undulation), see Fig. 2. The observation equation as function of the orthometric heights of point \(i\) now becomes:

\[
v_{i,t_1,t_m} = -H_{i,t_1} + H_{i,t_m} - N_{i,t_1} + N_{i,t_m} + h_{A,t_1} - h_{B,t_m} + \xi_{i,t_1,t_m}.
\]
constant in time, i.e. \( N_{i,t} = N_{i,t_0} \). The observation equation then reduces to:

\[
\begin{align*}
\varepsilon_{i,t_{0}} &= -H_{i,t} + H_{i,t_0} + h_{A,t_0} - h_{B,t_0} + \xi_{i,t_{0}}. \\
\end{align*}
\]

This observation equation can in a next step be expressed as function of the vertical deformations by inserting expressions for \( H_{i,t} \) and \( H_{i,t_0} \), conform Eq. (1):

\[
\begin{align*}
\varepsilon_{i,t_{0}} &= -z_{i,t_0} + z_{i,t_{0}} + h_{A,t_0} - h_{B,t_0} - t_{i,t_0} + t_{i,t_0} + \xi_{i,t_{0}}. \\
\end{align*}
\]

Note that the initial height of point \( i \) is eliminated in this equation. Decomposing the vertical deformation due to gas extraction into a trend and signal term, conform Eq. (3), we obtain the following observation equation:

\[
\begin{align*}
\varepsilon_{i,t_{0}} &= -z_{i,t_0} + z_{i,t_{0}} + h_{A,t_0} - h_{B,t_0} - t_{i,t_0} + t_{i,t_0} + \xi_{i,t_{0}}. \\
\end{align*}
\]

Similar to the leveling observation equation, the random error term \( \xi_{i,t_{0}} \) in Eq. (17) is a combination of deformation signal, point noise and measurement noise. As mentioned, there are no initial heights in this observation equation, in contrast to the leveling observation equation (9). On the other hand, in observation equation (17) we have to account for a difference in reference heights \( A \) and \( B \). It should be noted that also the same reference height can be used at both epochs. Suppose that this is point \( A \), then we have to account for the height difference of point \( A \) in time: \( h_{A,t} - h_{A,t_0} \). It can also be true that we a priori know that this reference point is stable. In that case the difference in reference heights completely disappears from observation equation (17).

3.2.2 InSAR-derived vertical displacement data

In this paper we focus on the integration of leveling data with InSAR-derived vertical displacement data, rather than GPS-derived vertical displacements. This choice is not only motivated by the cost aspect, but also by the spatial density of InSAR-derived deformation data (up to 1600 points/km²), which is much higher than of GPS data. The temporal density of InSAR-derived data is lower than that of (permanent) GPS data, because of the revisit interval of the SAR sensors onboard radar satellites, which can be more than one month. However, this is not really an issue for slow-term deformation problems as due to gas extraction. The collection of InSAR-derived deformations for one time interval is referred to as a vertical displacement map. In order to make assumptions on the measurement noise of InSAR-derived vertical displacements, it is necessary to briefly explain the way these data are generated from raw InSAR observations.

Generation of an InSAR displacement map

Synthetic Aperture Radar sensors measure both magnitude and phase of the transmitted electromagnetic signal that is backscattered from the earth’s surface. The phase measurement is used to derive information on heights and deformations of the terrain. This phase represents a combination of the distance between the SAR antenna and the surface, and the surface scattering effect. If a second SAR data set is collected, then from comparing the phase of the second image with the phase of the first, an interferogram can be formed. The basic principle of interferometric SAR is that if the surface characteristics are identical for both images, the phase difference between corresponding pixels in both images only accounts for the distance-dependent component. The surface scattering effects will be the same when the distance between the orbital positions of the radar (called the baseline) is small. This is established by making benefit of the repeating constellation of SAR sensors onboard radar satellites. For deformation applications the baseline should ideally be zero in order to eliminate topographic information from the interferogram (since this may bias true surface deformations). However, due to satellites drifting from their nominal orbits, baselines will be non-zero. Fortunately, the interferogram can be corrected for topographic information using an external digital elevation model (DEM). As a result, a differential interferogram is obtained that only contains the topographic changes that occurred between the two SAR acquisitions.

Limitations of InSAR

Using the described procedure a vertical displacement map with a very high resolution (about 25 x 25 m²) can be obtained. Since the accuracy of the phase measurement is a fraction of (half) the radar wavelength (2–3 mm for distance measurements of typical radars), this in principle allows for mm-accuracy of vertical deformations.
The phase measurements are ambiguous: only the fractional phase can be measured while the number of whole cycles is unknown.

- Inaccuracies in the satellite orbits cause baseline errors which result in increased phase noise.

- Differences in atmospheric conditions between the two SAR acquisitions cause spatially correlated errors in the phase measurements.

- The surface backscattering characteristics within one resolution cell (pixel) change in time, which causes weak phase correlation (interferometric coherence) between both images. This loss of coherence is known as temporal decorrelation and makes InSAR unsuitable for, for example, vegetated areas.

For the design studies described in Sect. 6 it is however assumed that the first two limitations are overcome. It is assumed that for every point in the displacement map the phase can be successfully unwrapped and that absolute phases are obtained. In addition, it is assumed that errors due to the satellite orbits are absent by using precise orbits.

Conventional vs. permanent-scatterer InSAR

The two other limitations of the InSAR technique, i.e. the atmospheric errors and the temporal decorrelation, are currently important topics of research. Recently the so-called permanent-scatterers InSAR technique has been developed to isolate coherent pixels and to solve the atmospheric errors (Ferretti et al. 2001). These coherent pixels (the permanent scatterers) are usually stable natural reflectors, like roofs of buildings or rocks. Although the locations of the permanent scatterers cannot be predicted, their density will vary between 50 and 400 points/km². Despite this lower density compared to the conventional InSAR technique in which just two acquisitions are used to compute a displacement map (about 1600 points/km²), it is still much higher than obtained with traditional geodetic techniques like leveling and GPS. Drawback is that for this technique many, e.g. more than 30, SAR images are needed, from which the atmospheric errors are estimated (Hanssen and Ferretti 2002). For the simulations described in Sect. 6 it is assumed that we have InSAR-derived displacement maps based on the conventional approach which can be biased by atmospheric errors, as well as maps based on the permanent-scatterer approach free of atmospheric errors.

Assumptions InSAR measurement noise

In this article the atmospheric errors – when present – are modeled as random errors, which means that they are part of the InSAR measurement noise term $\sigma_{\text{v,\text{p}}}^{2}$ in Eq. (17). As a consequence, the measurement noise becomes spatially correlated. It is assumed that the covariance between points $i$ and $j$ in each map can be described by

$$
\sigma_{\text{v,\text{p}}}^{2} = \sigma_{\text{a}}^{2} \delta_{ij} + \sigma_{\text{e}}^{2} \exp\left\{ - \left( \frac{t_{ij} - t_{g}}{a_{g}} \right)^{2} \right\}.
$$

(18)

In this covariance function $\sigma_{\text{a}}^{2}$ denotes the variance of the phase measurement and the Delta function is defined as $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. The atmospheric variance is denoted as $\sigma_{\text{e}}^{2}$ and the correlation distance of the atmospheric errors as $a_{g}$. The variance of each point in a displacement map due to measurement noise follows thus as $\sigma_{\text{v,\text{p}}}^{2} + \sigma_{\text{a}}^{2}$. Note that in absence of atmospheric errors, we set $\sigma_{\text{e}}^{2} = 0$. Whereas the phase standard deviation is at mm-level, the atmospheric standard deviation is typical of the order of 1 cm. The correlation distance of the atmospheric errors can be a few km.

When different maps have a common start date or end date, the InSAR measurement noise should also address temporal correlation. This correlation is positive when two maps have either the same start or end date (in case of so-called single-master interferograms; see Hanssen et al. 2003), whereas the correlation is negative when the end date of the first map corresponds to the start date of the second map (in case of so-called cascaded interferograms; see ibid). Taking these considerations into account, the covariance between point $i$ of a displacement map over time interval $(t_{l}, t_{m})$ and point $j$ of a displacement map over time interval $(t_{f}, t_{g})$ can be written as:

$$
\sigma_{\text{v,\text{p}}}^{2}(t_{l}, t_{m}) \text{ and } (t_{f}, t_{g}) = \begin{cases} 
\sigma_{\text{v,\text{p}}}^{2}, & \text{if } t_{l} = t_{f} \land t_{m} = t_{g}, \\
\frac{1}{2} \sigma_{\text{v,\text{p}}}^{2}, & \text{if } t_{l} = t_{f} \land t_{m} \neq t_{g}, \\
0, & \text{otherwise}.
\end{cases}
$$

(19)

4 Mathematical model

4.1 Functional model

For deformation estimation, the observation equations derived in Sect. 3 are connected to the seven trend parameters per subsidence bowl through Eqs. (4) and (5). These trend parameters are computed by means of an integrated least-squares adjustment of all leveling and/or vertical displacement data simultaneously. Since both Eqs. (4) and (5) are nonlinear, we first need to linearize the observation equations. For the linearization we need approximate values of all unknown parameters. After linearization, we may reformulate the linearized observations equations as a Gauss-Markov model:

$$
E\left\{ \begin{bmatrix} \Delta W \\ \Delta V \end{bmatrix} \right\} = \begin{bmatrix} \begin{bmatrix} W_{\text{sub}} \end{bmatrix}_{\text{tgt}} & \begin{bmatrix} W_{\text{sub}} \end{bmatrix}_{\text{sub}} & 0 \\ 0 & V_{\text{sub}} & A_{\text{sh}} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \Delta h_{\text{lt}} \\ \Delta h_{\text{AB}} \end{bmatrix}_{\text{tgt}} \end{bmatrix} \right\} = \begin{bmatrix} \Delta p \end{bmatrix},
$$

(20)
where:

\[ E\{\cdot\} \quad : \quad \text{mathematical expectation operator} \]
\[ \Delta \nu \quad : \quad \text{observed-minus-computed vector of leveling data} \]
\[ \Delta \nu \quad : \quad \text{observed-minus-computed vector of vertical displacement data} \]
\[ \Delta H_{t_0} \quad : \quad \text{incremental vector of leveling \textit{null} heights} \]
\[ \Delta p \quad : \quad \text{incremental vector of subsidence parameters} \]
\[ \Delta h_{AB} \quad : \quad \text{incremental vector of reference height differences} \]
\[ W \quad : \quad \text{matrix to transform heights to leveling height differences} \]
\[ V \quad : \quad \text{matrix to transform heights to vertical displacement height differences} \]
\[ A_{hgt} \quad : \quad \text{matrix selecting null heights for leveling observations} \]
\[ A_{sub} \quad : \quad \text{matrix with coefficients due to linearization of bowl equations} \]
\[ A_{shft} \quad : \quad \text{matrix needed for the parameterization of shift parameters} \]

Model (20), referred to as the \textit{functional} model, distinguishes between three groups in parameter vector \( x \). The first group, \( \Delta H_{t_0} \), contains the \textit{initial or null heights} corresponding to the leveling points. It is not possible to estimate all null heights in the adjustment: we have to constrain one of these heights to its a priori value (or zero). The choice of this so-called \textit{S-basis} (see Baarda 1973) is arbitrary since it does not influence the estimated subsidence. It is also not necessary that this height remains stable. The second group of parameters are the \textit{subsidence parameters}. For \( n_B \) bowls we have
\[
\Delta p = (\Delta p_1^T, \ldots, \Delta p_n_B^T)^T, \quad \text{where} \quad \Delta p_B = (\Delta v_B, \Delta t_{0, B}, \Delta x_{c, B}, \Delta y_{c, B}, \Delta \beta_B, \Delta \phi_B)^T
\]
is the vector of seven parameters of the \( B \)th bowl. The subsidence parameters are of course the parameters we are primarily interested in. In this context it should be noted that for the start time of subsidence, \( \Delta t_{0, B} \), being estimable, we need data of \textit{before} the gas production started. Otherwise it should be constrained to an a priori value. The third parameter group, \( \Delta h_{AB} \), contains the differences in reference heights \( \text{(shifts)} \) of the vertical displacement data between different epochs. Not the individual reference heights are estimable in the adjustment, but only combined terms, for example
\[
h_{A,t_0} - h_{B,t_0} \text{ (see Eq. (17))}
\]
So for each set \( \text{or map} \) of vertical displacement observations \text{one} shift is parameterized.

Note from the design matrix \( A \) of model (20) that the partial design matrices corresponding to the leveling null heights \( A_{hgt} \) and the subsidence parameters \( A_{sub} \) are pre-multiplied by matrices \( W \) and \( V \). Both matrices, which only contain ones and zeros, are introduced to perform \textit{differencing} operations. The reason for introducing these matrices is that it is easier to set up the partial design matrices at the absolute point and epoch level, than relative in space (in case of leveling), or relative in time (in case of vertical displacement observations). Both matrices \( W \) and \( V \) only contain \{minus\} ones and zeros. This pre-multiplication is not necessary for the partial design matrix of the shift parameters, since the estimable shift parameters can be regarded as parameters that are already differenced. Hence, matrix \( A_{shft} \) is a simple vector having ones at all entries.

To set up the partial design matrix of the subsidence parameters, \( A_{sub} \), we also need the \textit{horizontal coordinates} of all involved leveling and vertical displacement points. In order to perform an integral adjustment of both types of observations, these horizontal positions must be defined in one and the same coordinate system. If this is not the case, coordinate transformations should be applied.

### 4.2 Stochastic model

The stochastic properties of the leveling and vertical displacement data are modeled in the \textit{stochastic model}, i.e. the variance-covariance \((\text{vc-})\) matrix. This reads:
\[
D \begin{bmatrix} \Delta \nu \\ \Delta \nu \end{bmatrix} = W \begin{bmatrix} Q^w \quad 0 \\ 0 \quad Q^p \end{bmatrix} W^T \quad \text{(21)}
\]

where:
\[ D \{\cdot\} \quad : \quad \text{mathematical dispersion operator} \]
\[ Q^w \quad : \quad \text{vc-matrix of the leveling data} \]
\[ Q^p \quad : \quad \text{vc-matrix of the vertical displacement data} \]
\[ Q_\eta \quad : \quad \text{vc-matrix of the point instabilities} \]
\[ Q_z \quad : \quad \text{vc-matrix of the deformation signal} \]

As shown, the vc-matrix of the observations consists of two parts: a block-diagonal part (left of the plus sign in Eq. (21)), and a part that consists of a multiplication of matrices (right of the plus sign). The first part models the measurement noise. Both measurement noise vc-matrices, \( Q^w \) and \( Q^p \), are constructed using the covariance functions in Eqs. (10) and (19), respectively. The second part models both point noise and deformation signal. The point noise vc-matrix, \( Q_\eta \), is created using the covariance function in Eq. (2), while the vc-matrix accounting for the deformation signal, \( Q_z \), is filled using the covariance function in Eq. (6). Note that since the point noise is uncorrelated in space, there is no correlation between the leveling and vertical displacement data due to point noise (since it can be proved that \( WQ_\eta V^T = 0 \)). Both types of observations become correlated however, because the spatially correlated deformation signal implies that \( WQ_z V^T \neq 0 \).

### 5 Estimation and testing procedure

A powerful consequence of formulating an integrated adjustment model (20) is that it enables \textit{statistical hypothesis testing} of both leveling and vertical displacement data.
Note that the spatio-temporal subsidence model allows for testing on (gross) errors in the data even when there is no redundancy in the individual data sets. For example, when a simple leveling line is measured at two epochs, the individual networks do not have redundancy since no loops can be formed. Redundancy is however introduced when the two epochs are connected by means of the spatio-temporal model (provided that the start time of subsidence $\Delta t_{0,B}$ is constrained; otherwise at least three leveling epochs should be used). Statistical testing of the data of individual vertical displacement maps is impossible when no spatio-temporal model is used. In contrast to leveling the individual maps are not redundant and hence it is not possible to form loops or whatsoever. However, it becomes possible to test the data of an individual map when the spatio-temporal model is introduced (provided that the start time of subsidence $\Delta t_{0,B}$ is constrained; otherwise at least two maps should be used).

With the functional and stochastic models as specified in the previous section, the best fitting subsidence model is determined in a stepwise procedure of least-squares adjustment, statistical hypothesis testing and adaptation of both the data and the model. In each step the actual model and data (the so-called null-hypothesis model) are tested against a number of alternative hypotheses, each suggesting a specific model adaptation or possible error(s) in the data. For an overview of the alternative hypotheses relevant for subsidence monitoring we refer to Kenselaar (2001).

5.1 Solving the trend model

When the null-hypothesis is not rejected eventually, the least-squares solution for the unknown parameters of the trend model, denoted as $\hat{p}$, plus its variance-covariance matrix, denoted as $Q_\mathbf{\hat{z}}$, is obtained as follows:

$$\hat{z} = \left(A^T Q_y^{-1} A\right)^{-1} A^T Q_y^{-1} \mathbf{y}_z = \left(A^T Q_y^{-1} A\right)^{-1} \mathbf{y}_z.$$

The least-squares solution of the subsidence parameters is part of vector $\hat{z}$. It reads $\hat{p} = p^0 + \Delta \hat{p}$, where $p^0$ is the vector of approximate values. Using the solution for the subsidence parameters, the least-squares solution for the trend at any arbitrary point and at any epoch (not necessarily at the locations of the measurement data or at the measurement epochs) is obtained by inserting the entries of $\hat{p}$ into Eq. (4). For a point $G$ at epoch $t_n$ this solution is denoted as $\hat{z}_{G,t_n-0}$.

5.2 Estimation of signal and total subsidence

In order to resolve the total subsidence at any point at any epoch, besides the trend also the signal needs to be solved, see Eq. (3). The signal is computed using the vector of least-squares residuals of the trend model. This vector is computed as:

$$\hat{\xi} = P_A^\perp \mathbf{y}_z,$$

where the orthogonal projector is computed as $P_A^\perp = I - A \left(A^T Q_y^{-1} A\right)^{-1} A^T Q_y^{-1}$. The least-squares solution of the signal at a point $G$ at epoch $t_n$ is now obtained as:

$$\hat{z}_{G,t_n-0} = q^T G_{t_n-0} M^T Q_y^{-1} \hat{\xi}.$$

where $M^T = [V^T, V^T]$, and $q_{G,t_n-0}$ a vector accounting for the covariances between the signal at point $G$ at epoch $t_n$ and the signals at the observed points and epochs. In fact, this computation of the signal for a specific point at a specific epoch corresponds to least-squares collocation. With both trend and signal solved, the least-squares solution of the total subsidence easily follows as:

$$\hat{z}_{G,t_n-0} = \hat{z}_{G,t_n-0} + \hat{z}_{G,t_n-0}^*.$$

When the subsidence is estimated in this way for a grid over the subsidence area, it is possible to create contour maps of the estimated/collocated subsidence.

6 Simulations

In order to gather experience into the estimability of a spatio-temporal subsidence bowl from integrated leveling and InSAR-derived vertical displacement data, simulations were carried out. Advantage of conducting simulations is that the performance of the different processing strategies can then easily be assessed by comparing the estimated subsidence with a ground truth. In practice the true subsidence is of course unknown and then it can be rather difficult to separate the subsidence due to gas extraction from the subsidence caused by other driving mechanisms (for example, ground water variations).

6.1 Simulated subsidence

The design studies described in this article are a continuation of the simulations described in Kenselaar and Martens (2000), in which only leveling measurements were used. In our design studies we simulated the same single subsidence bowl as in (ibid), but we also added a signal to the subsidence trend. This deformation signal was generated at a grid of $49 \times 49$ points over the subsidence area of $25 \times 25$ km, using the random generator of Matlab. For this stochastic process we assumed a standard deviation of $2 \text{mm} / \sqrt{\text{yr}}$ and a correlation distance of 3 km, see Eq. (6). These values can be expected in practice (Houtenbos 2000). Tab. 1 summarizes the characteristics of the simulated subsidence. Fig. 3 shows contour maps of this subsidence for 1 January 1991 (5.5 years after
the start), while Fig. 4 shows contour maps for 1 January 1999 (13.5 years after the start). Note that due to the simulated deformation signal, there is some small uplift in the northwestern part of the area for 1 January 1999. From the figures it seems that the contribution of the signal to the total subsidence can range up to 20% at some locations.

6.2 Simulated observations

Precise leveling data are simulated for the time span 1983–1999 at five epochs, at a regular interval of four years. The leveling network is a regular grid of lines that completely covers the area of $25 \times 25$ km, with point distances of 3125 m. In the center of the subsidence area even points are added which are halfway the leveling trajectories, see Fig. 5 (left). Note that this dense and regular leveling network consisting of 121 points connected by 184 observations will probably yield very good results, but is very expensive to measure and maintain it on a regular basis (450(!) km of leveling per epoch). In the simulations it will be referred to as network $L$.

In addition to leveling data, InSAR-derived vertical displacement data are simulated for varying time intervals. It is assumed that these displacement maps cover the same time span as the leveling data, so from 1983 to 1999. We assume to have four InSAR maps available that each cover the 4-year time interval between two leveling epochs, i.e. 1983–1987, 1987–1991, 1991–1995 and 1995–1999. Note that these four maps can – without loss of information – be transferred to four maps that have for example a common start date, i.e. 1983–1987, 1983–1991, 1983–1995 and 1983–1999. So in the simulations we could also have used this latter set of maps. The solution using either the first or the latter set will however be exactly the same, since the correlation between the maps is properly accounted for in the adjustment. Each displacement map covers the subsidence area of $25 \times 25$ km.
in the form of a regular grid. In the simulations we varied
the spacing (resolution) of this grid:
- **Network L**<sub>625</sub> (see Fig. 5; middle), a grid with a spac-
ing of 1/3 × 3125 m (the distance between points in
the leveling network) = 1041.67 m. This corresponds
to 25 × 25 = 625 points for the complete subsidence
area. Note that this spacing is much larger than the
size of a resolution cell of real InSAR displacement
data, which is between 25 m (conventional InSAR)
and 160 m (permanent-scatterer InSAR). Due to com-
puter memory overflow it is not possible to use very
dense grids covering an area of 25 × 25 km. It is
however expected that relatively sparse grids are suf-
ficient for this type of (smooth) deformation.
- **Network L**<sub>289</sub> (see Fig. 5; right), a grid with a spac-
ing of 1.5 times the spacing of network L<sub>625</sub>, i.e.
1562.5 m. (Note that this distance corresponds to the
distances between the leveling points of network L
in the center of the subsidence area.) This spacing
corresponds to 17 × 17 = 289 points for the complete
subdivision area, and this number is only 46% of the
number of points of network L<sub>625</sub>.

With the horizontal positions of the leveling and InSAR
points extracted from the network configurations, the
systematic component of the observations was generated
using the simulated vertical deformations. Note that
the initial heights of the leveling benchmarks and the shifts
of the InSAR displacement maps were simulated as zero
values. In addition to this systematic component, for all
leveling and InSAR observations a random error was gen-
erated in order to simulate measurement noise and noise
due to point instabilities. For this purpose again the ran-
dom generator of Matlab was used. The leveling mea-
surement noise was generated with a standard deviation
of 0.7 mm/√km. For the InSAR data we generated two
data sets with different measurement noise in order to
investigate the effect of the presence of atmospheric art-
facts. For both sets a phase standard deviation of 3 mm
was chosen for all points. For the atmosphere-biased map
the observations were generated with an additional ran-
dom error for atmospheric artifacts, having a standard
deviation of 1 cm and a correlation length of 5 km. The
standard deviation of leveling point noise was chosen as
0.6 mm/√yr, being a realistic assumption conform Wy-
at (1989) or Kenselaar and Martens (2000). Since InSAR
point noise is usually larger than leveling point noise
(Houtenbos 2003), a value of 1.2 mm/√yr was assumed
for the InSAR point noise.

The characteristics of the simulated observations are
summarized in Tab. 2.

<table>
<thead>
<tr>
<th>Simulated processing strategies</th>
</tr>
</thead>
</table>

6.3 Simulated processing strategies

In a search for more efficient monitoring strategies we
formulated some strategies for the processing of the sim-
ulated leveling and InSAR data. The central question
hereby is: how accurate can a certain strategy recover the
true subsidence? Since we have simulated the true subsi-
dence, the accuracy of a certain strategy can be assessed
by comparing the estimated subsidence at a certain loca-
tion/epoch with its simulated counterpart. This compa-
rison is conducted on a grid of 49 by 49 (= 2401) points
spanning the subsidence area of 25 × 25 km. The follow-
ing leveling/InSAR strategies are considered:
1. leveling only based on 5 networks \(L\) (L);
2. InSAR only based on 4 maps \(I_{625}\) (I1);
3. InSAR only based on 4 maps \(I_{289}\) (I2);
4. InSAR only based on 4 maps with atmospheric artifacts \(I_{625}^a\) (I3);
5. leveling based on 3 networks integrated with InSAR based on 4 maps with atmospheric artifacts \(I_{625}^a\) (L/I3).

The choice of these strategies is motivated as follows. The leveling-only strategy in which network \(L\) is surveyed every four years is the most expensive strategy but has been included since a leveling-only based strategy is the usual approach for the monitoring of subsidence due to gas extraction. The second and third strategies are InSAR-only based on atmosphere-free displacement maps. From these strategies it immediately follows how well InSAR is able to recover the subsidence, as compared to leveling only. The fourth strategy is also an InSAR-only strategy, but then based on atmosphere-biased displacement maps. In the fifth strategy the four atmosphere-biased InSAR maps are combined with three leveling epochs.

### 6.4 Results and discussion

The relevant results of the processing of each strategy are summarized in Tab. 3.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>I1</th>
<th>I2</th>
<th>I3</th>
<th>L/I3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-test quotient</td>
<td>0.94</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>% of discrepancies &lt; 5 mm, at 1 January 1999</td>
<td>83.9</td>
<td>89.3</td>
<td>68.1</td>
<td>67.8</td>
<td>92.8</td>
</tr>
<tr>
<td>precision subs. velocity center (mm/yr)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td>redundancy</td>
<td>793</td>
<td>2489</td>
<td>1145</td>
<td>2489</td>
<td>2921</td>
</tr>
<tr>
<td>number of epochs</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>number of maps</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

In the table first the F-test quotient is given, which is defined as the ratio of the F-test (overall model test) statistic with the critical value of the test. The expectation of this F-test quotient is 1. Tab. 3 shows that for all strategies this quotient is indeed very close to 1, and this is understandable because we simulated the observations, without adding (gross) errors (aside from the random errors). This is also a check that we simulated the data in a correct way.

The second row in Tab. 3 shows the results of the comparison of the estimated subsidence according to each strategy, with our simulated subsidence, in fact the ground truth. In the table the (relative) number of the 2401 grid points is given for which the estimated subsidence agrees within 5 mm with the simulated subsidence at those points. These discrepancies have also been interpolated for the subsidence area and are visualized as contour maps in Figs. 6, 7 and 8. Note that these graphs and also the percentages in Tab. 3 are given for 1 January 1999, the last observed epoch. An exception to this is the contour map left in Fig. 6, which visualizes the discrepancies on 1 January 1991, halfway the time span 1983–1999. This map has been included to demonstrate that the discrepancies generally grow in time. By evaluating the discrepancies only for the last epoch, 1 January 1999, the percentages can be considered as safe lower bounds. In case of the leveling-only strategy \(L\) this lower bound is 84%. In case of InSAR only (strategies I1 and I2) it follows that the four \(I_{625}\)-maps with a spacing of about 1042 m are able to recover the subsidence very precisely: for almost 90% of the area the subsidence is recovered within 5 mm. However, when the spacing of the maps is increased (about 1563 m; in case of the \(I_{289}\) maps) this percentage is reduced to about 68%. Fig. 7 visualizes the discrepancies for both strategies. Note that with an even smaller spacing than of the \(I_{625}\)-maps we probably would have obtained a percentage close to 100%, but this strategy has not been taken into account as due to computer memory overflow problems. The displacement maps in strategies I1 and I2 are simulated without atmospheric errors. When we use the atmosphere-biased \(I_{625}^a\)-maps (strategy I3) the result becomes much worse: for only 68% of the area the recovered subsidence is within 5 mm of the ground truth, versus 90% for strategy I1, in which the maps have the same spacing as in strategy I3. The left contour map in Fig. 8 visualizes these large discrepancies. The map shows that the discrepancies may be up to 20 mm.

Based on the good performance of the InSAR-only strategy I1 versus the performance of the leveling-only strategy \(L\), it is expected that an integration of these maps with leveling data would barely improve the InSAR-only results. So this strategy was not tried. Another situation applies to the atmosphere-biased InSAR data; for these maps an integration with leveling data might be worthwhile. Hence in strategy L/I3 the atmosphere-biased \(I_{625}^a\)-maps are combined with leveling data. However, for this integration we only used three leveling networks (at epochs 1983, 1991 and 1999), since taking all five networks would still lead to a very expensive solution. This integrated L/I3-strategy indeed resulted in a significant improvement: for 93% of the area at 1 January 1999 the subsidence could be recovered with an accuracy better than 5 mm, thus an improvement of about 25% compared to the InSAR-only strategy I3. The right contour map in Fig. 8 visualizes the discrepancies of the L/I3-strategy.

These simulations demonstrate that an integration of leveling and InSAR data only leads to better results when the InSAR data are of minor quality due to for example atmospheric artifacts. Otherwise one may rely on InSAR only for this type of subsidence monitoring. As shown in Tab. 3, the precision of the estimated subsidence velocity is in the integrated case comparable to the precision of the velocity when either good InSAR data or frequently surveyed leveling data are used: about 0.5 mm/yr.
7 Conclusions

There are several advantages of the approach to continuous spatio-temporal modeling of land subsidence due to gas extraction as described in this article:

- It allows for a relatively easy integration of measurement techniques, at observation level. Hereby it is not necessary that the different measurement techniques occupy the same points or are observed at the same epochs (although this was the case in the simulations).
- It enables rigorous statistical testing of all observations, even those observations that are not redundant when considered as individual data sets.
- In contrast to many other deformation modeling schemes, the approach does not require stable reference points.

In this article it has furthermore been demonstrated that despite its benefits (high spatial resolution, relatively low costs), InSAR is currently not yet a fully reliable technique for the precise monitoring of deformations due to gas extraction. Leveling data remain necessary, especially when the InSAR displacement data are contaminated with atmospheric errors or at places where no coherent InSAR data are available, e.g. due to vegetation. In addition, leveling data are indispensable for the detection of deformations that started before 1992, the year the InSAR data

Fig. 6: Discrepancies between estimated and simulated subsidence for 1 January 1991 (left) and 1 January 1999 (right) using the L-strategy. Contours are plotted for each 5 mm difference.

Fig. 7: Discrepancies between estimated and simulated subsidence for 1 January 1999 using the I1-strategy (left) and the I2-strategy (right). Contours are plotted for each 5 mm difference.

Fig. 8: Discrepancies between estimated and simulated subsidence for 1 January 1999 using the I3-strategy (left) and the L/I3-strategy (right). Contours are plotted for each 5 mm difference.
were archived for the first time. The development of the permanent-scatterers InSAR technique is promising, since the problems of low coherence and atmospheric errors are being tackled. However, because the number and location of the permanent scatterers cannot be predicted with this technique, leveling may still be necessary in areas with too few permanent scatterers.

Acknowledgements
This work is performed within project 2.4 "Integrated Deformation Analysis" of the DIOC (Delft Interfaculty Research Center) research program Determination and Prediction of the 3D movements of the Earth’s Surface at the Delft University of Technology in the Netherlands. Raoul Quadvlieg and Adriaan Houtenbos are acknowledged for commenting on an early draft of this article.

References

Author’s address
Dr. Dennis Odijk
Department of Earth Observation and Space systems
Delft University of Technology
P. O. Box 5058, 2600 GB Delft, The Netherlands
d.odijk@lr.tudelft.nl