Statistics Netherlands

2015 Scientific Paper

Trend changes in ground subsidence in Groningen

update November 2015

Frank Pijpers
D. Jan van der Laan

Contents

Nederlands

Deze rapportage is een verslag van onderzoek dat is uitgevoerd in het kader van fase 1 van een onderzoeksproject door het CBS in opdracht van Staatstoezicht op de Mijnen (SodM). Dit onderzoek is ten behoeve van een statistische onderbouwing van het meet- en regelprotocol voor gasexploitatie in de provincie Groningen. Het onderwerp van dit rapport is een heranalyse van trends in de bodemdaling in de provincie Groningen, gerapporteerd in fase 0, in december 2014 en in fase 1 in mei 2015. Voor de voorliggende analyse is de tijdreeks voor de GPS gegevens aangevuld tot aan 3 oktober 2015.

Zoals ook bleek uit de eerdere analyses is er een statistisch significante trendbreuk in de bodemdaling ongeveer 2 maanden nadat de productie sterk was gereduceerd. De trendwijziging kan zich geleidelijk gemanifesteerd hebben over een periode van enkele weken, en er is daarom een onzekerheid van ruwweg een week of twee over de centrale datum van deze overgang.

English

This is a report on the research that has been carried out within phase 1 of a project being carried out by Statistics Netherlands and commissioned by State Supervision of Mines (SodM). This research is part of the underpinning of the statistical methods employed to support the protocol for measurement and regulation of the production of natural gas in the province of Groningen. The subject of this report is to re-analyse the trends in the ground subsidence in and around that region, first reported in phase 0, in December 2014 and later in phase 1 in May 2015. For this re-analysis the time series of the GPS data has been updated with more recent measurements up to October 3, 2015.

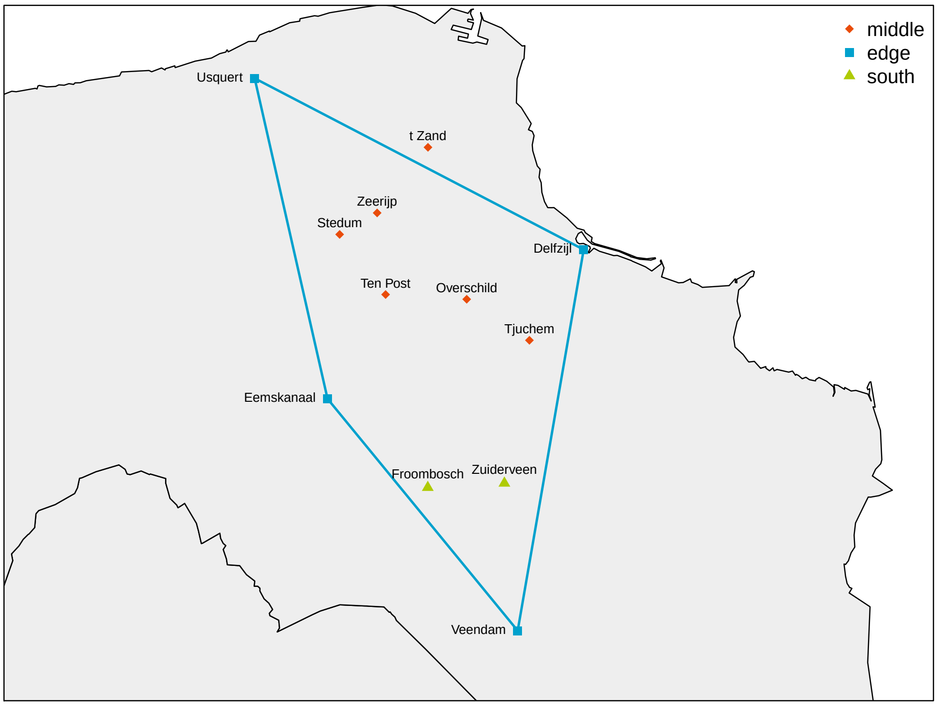
In accordance with the earlier report, it is found that there has been a significant change in the ground subsidence. The changeover in the trend of ground subsidence can have become manifest gradually over a period of several weeks, which implies that the central date of the transition is also uncertain by roughly a week or two.

1 Introduction

This is an update to the previous report studying ground subsidence in Groningen. The calculations performed in previous reports (Pijpers, 2014; Pijpers and van der Laan, 2015) are redone using more recent GPS data. The time series being analysed in this report contain the entire period that was analysed in the previous reports, but is now extended forward to cover the epoch up to October 2 2015. The same methodology is used as in previous reports.

In addition a cross correlation technique is employed to explore the relationship between production rate variations and GPS signals at ten Post (see section 3.3), through which the effect of more modest increases and decreases in gas production rates might become traceable in GPS ground subsidence measurements, rather than focus just on the very substantial decrease in gas production rate of January 2014.

Figure 2.1 The locations of the GPS stations from which data are available.



2 Data

Data are available for 13 stations for the period 2013-09-13 to 2015-10-03. Not all stations have data available for the complete period. From March 2014 GPS data is available for most stations. The stations can be divided into three groups as indicated in figure 2.1. A group 'middle'; these are in the production field in which the production has been reduced in January of 2014. A group 'edge': these are outside the main field, although at the station Eemskanaal (eems) there are wells in production. Finally, a group 'south' where production has not been reduced.

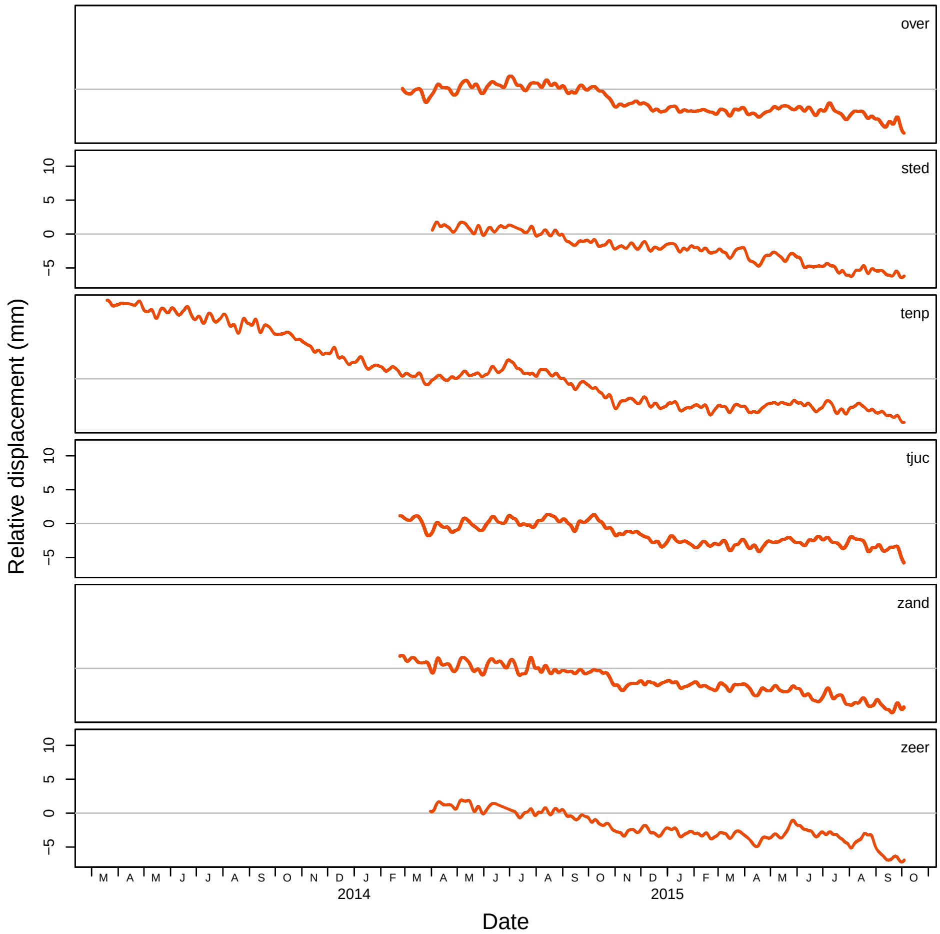
Figure 2.2 shows the filtered time series for each of the stations in 'middle' and 'edge'. The time series were filtered using a moving average filter as described in Pijpers (2014) to remove the high frequency noise, which is not of interest for current research. Since the absolute height of the stations is not of interest, the height has been normalised by subtracting the average height in the period 2014-02-19--2014-11-15.

The quantity of interest is the 'sagging' of the stations in group 'Middle' and in group 'South' with respect to those in group 'Edge'. It is this differential displacement over a spatial scale of the order of, or smaller than, the field that traces the 'subsidence bowl' which is due to compaction of the layer from which gas is extracted. Therefore the displacements, averaged for each group, are subtracted in the sense 'Middle' minus 'Edge' and 'South' minus 'Edge'. The result is shown in figure 2.3.

To assess the extent to which correlations might still be present in the differential displacement, two additional differential time series are shown in figure 2.4. One measure concerns the differential displacement between stations that are all within group 'Middle', in the sense (Zeerijp+ten Post) - (Stedum+Overschild) which is shown as an orange line. The other concerns

Figure 2.2 The time series of the GPS height after filtering out intermediate time scale variations. The average of its time series over the period from 2014-02-19 (day 50) to 2014-11-15 (day 214) is subtracted.

(a) GPS stations in group `middle'



(b) GPS stations in group `edge'

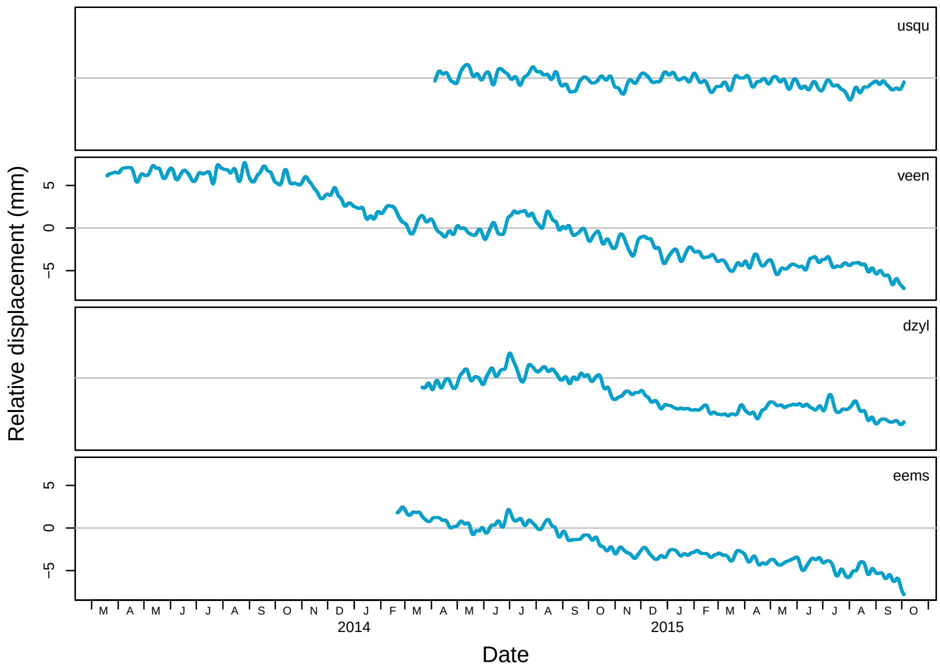
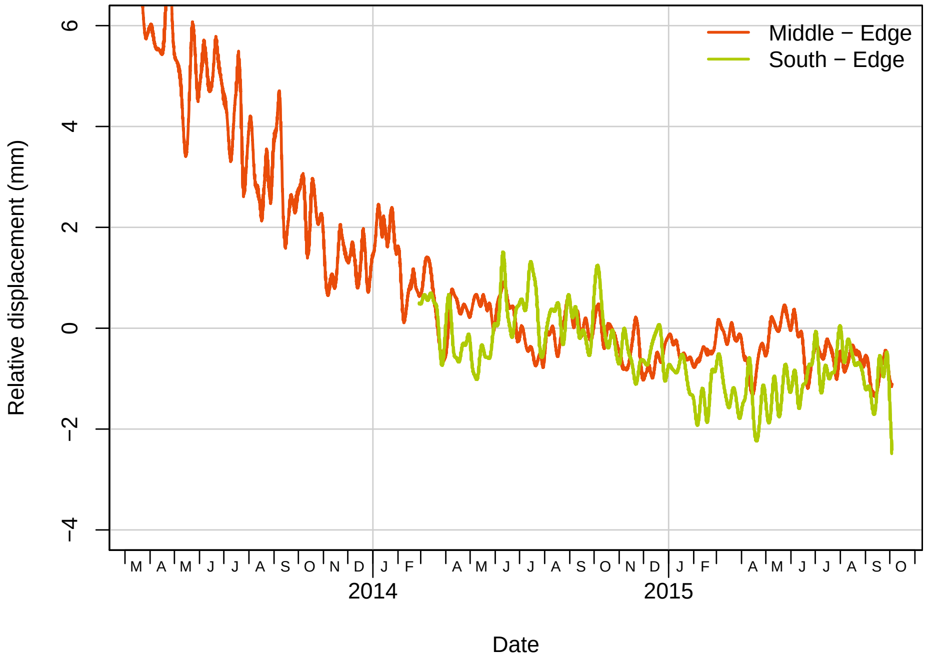
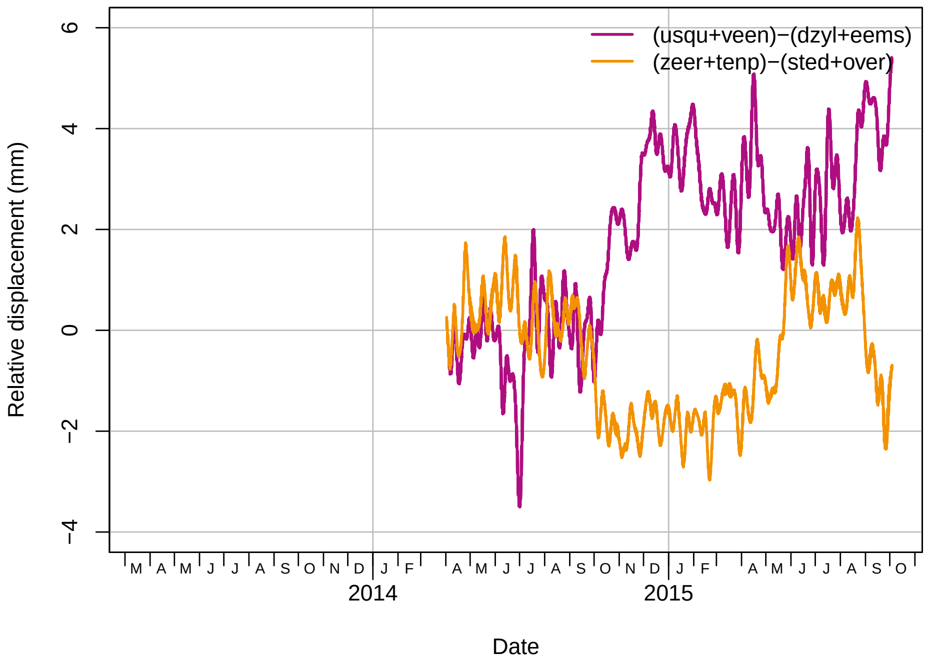


Figure 2.3 The difference between the average heights for the groups `middle' and `south' and the group `edge'.



the differential displacement between stations that are all within group ‘Edge’, in the sense (Usquert+Veendam) - (Delfzijl+Eemskanaal) shown as a purple line. In the absence of remaining biases these differential measurements should not show any trends.

Figure 2.4 The difference between the average heights for stations all within the group `middle' and all within the group `edge'.



While it is clear that for most of 2014 no trend is present in these sets of differential displacements, in October there is a sudden change in particular in the group ‘Edge’, but also to some extent in the group ‘Middle’. In the report of May 2015 (Pijpers and van der Laan, 2015) it is seen that a different calibration procedure and processing as performed by the TU Delft, produced clearly different results that are also closer to 0 for the period October 2014 to January 2015. Evidently there are still some issues that require resolving in the standard processing. From figure 2.2(b) it appears that of the GPS stations in the group 'edge', Usquert appears not to follow the trend of the other stations in that group which could produce the result seen in figure 2.4.

3 Analysis of trend changes

3.1 Linear fit through average trends

Since the reduction of the gas production at some locations near stations in the group 'Middle' has been rapid, it is of interest to explore whether a break can be found in the linear component of the downward drift for the time series for 'Middle' minus 'Edge'. To this end one can fit not only a single straight line, using standard least-squares fitting, but also introduce a break-point with a different straight line fit before and after that point. Two types of breaks are modelled: a continuous break which we will call a 'kink' and a discontinuous break which we will call a 'jump'.

Figure 3.1 shows the Akaike Information Criterion as a function of the position of the break. The minimum for the model with a 'kink' is at 2014-02-22 and that of the model with the 'jump' is a 2014-04-04. Figure 3.2 shows the time series with the three models (the two with break and the one without break). It is clear that the linear model without break doesn't describe the time series. The rate of relative descent has decreased after the first quarter of 2014.

It should be noted that while this type of fitting with a breakpoint is useful in order to asses whether subsidence trend is different between the start and the current end of the measured time series, it does not necessarily imply that in reality the changeover is as abrupt as this. A more gradual change in trend is allowed by the data.

3.2 Alternative model

One possible issue with the previous fit of the break in the trend, is that the number of measurements contributing to each of the average lines changes with time. Before Februari 2014 there is only one measurement station contributing to each of the lines. The previous model does not take this into account: one would expect the model to give more weight to differences at later dates as more measurement stations are contributing to the averages.

One possibility to avoid this problem is to not calculate the difference, but to model all time series combined. The difference in slopes between the two groups can be modelled explicitly and it is possible to test if this difference is significant.

A piecewise linear continuous fit can be written as

$$y_i = \beta_0 + \beta_1 t_i + \beta_2 s(t - t_0) t_i + e_i, \qquad (1)$$

Figure 3.1 AIC as a function of the position of the break in the linear model fit of the time series of the difference between the average of the group `middle' and `edge'. The minimum of the `kink' is at 2014-02-22 and that of the `jump' is at 2014-04-04.

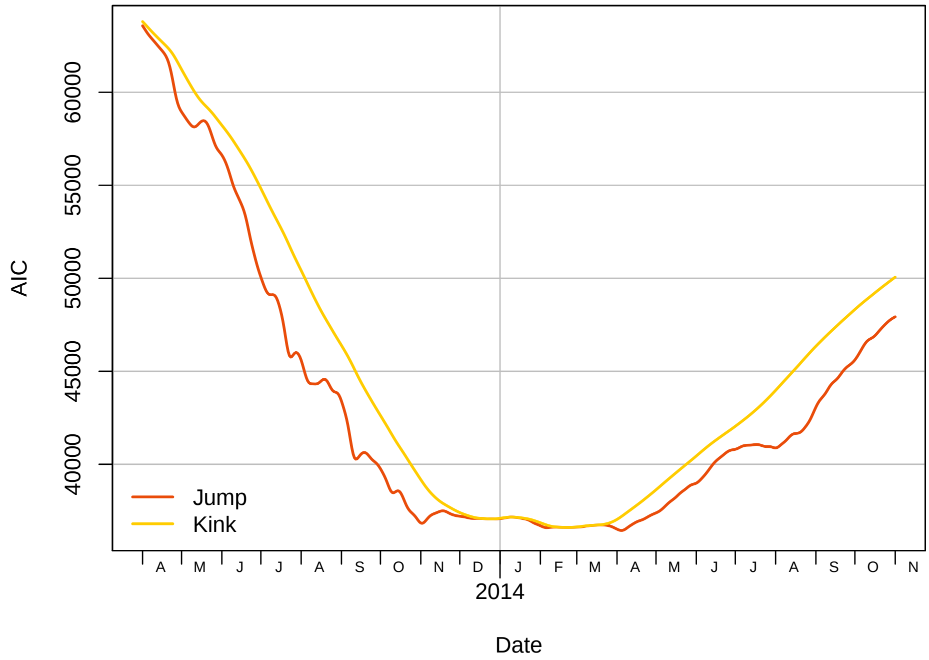


Figure 3.2 The trends fitted to the difference between the group averages of `middle' and `edge'.

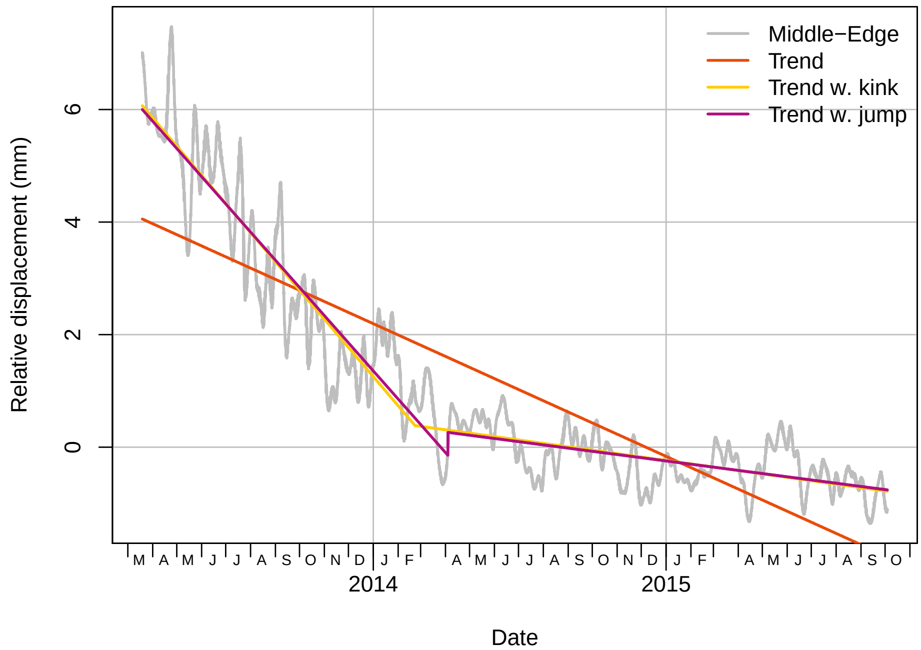
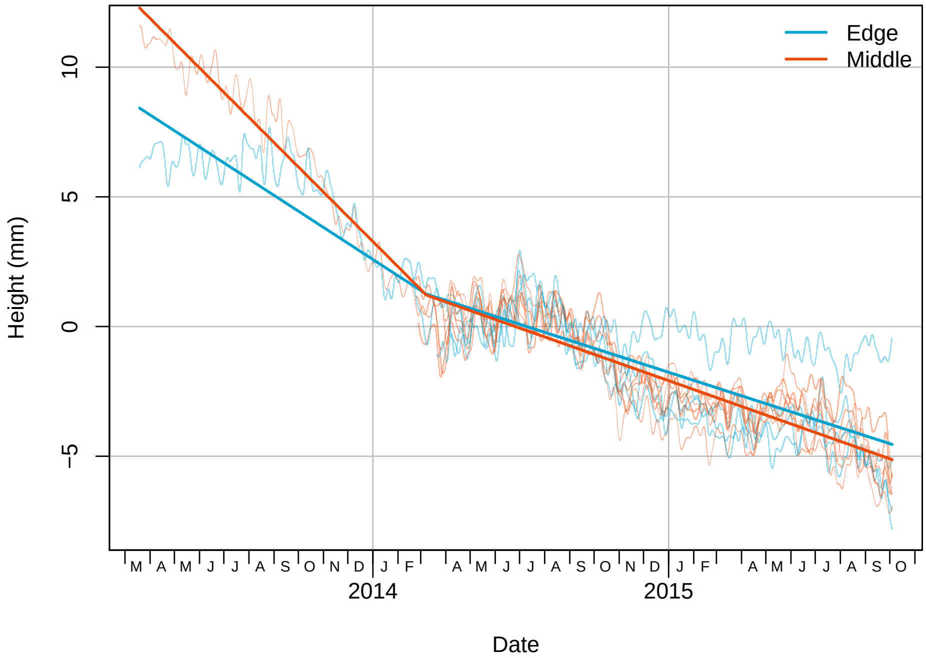


Figure 3.3 Filtered time series of each of the measurement stations in `edge' and `middle' with the fitted model. The break is located at 2014-03-08.



with $s(t)$ the step function ($s(t) = 1$ for $t \geq 0$ and 0 otherwise). We can then fit each of the time series with a piecewise linear continuous fit:

$$y_i = \begin{cases} \beta_0 + \beta_1 t_i + \beta_2 s(t - t_0) t_i + e_i & \text{if 'Edge',} \\ (\beta_0 + \beta_3) + (\beta_1 + \beta_4) t_i + (\beta_2 + \beta_5) s(t - t_0) t_i + e_i & \text{if 'Middle'.} \end{cases} \quad (2)$$

We can then test if the coefficient β_5 is zero.

The position of the break (t_0) is estimated using the same method as above: estimate the model for each break point and select the model with the smallest value of the AIC.

The resulting fit is shown with the time series of each of stations in figure 3.3. The coefficient β_5 differs significantly from zero. Furthermore, the in 'Edge' the discontinuity is significant ($\beta_2 \neq 0$). Therefore, both 'Edge' and 'Middle' decrease in time and for both this decrease is reduced from approximately 2014-03-14. However, for 'Middle' this decrease is much stronger than for 'Edge', suggesting that this difference is related to the gas production.

3.3 A smooth higher-order fit

As the time series is extended forward, attempting to capture the behaviour of the subsidence in terms of one or two linear trends with a single break or kink is likely to be increasingly inappropriate. Subsequent to January 2014 the gas production in the clusters at or near the more central GPS stations has varied substantially, and was increased, sometimes temporarily, even at those locations where a substantial reduction was implemented in Januari 2014. If the subsidence does respond to the gas extraction, its time behaviour must therefore similarly show accelerations in subsidence as well as slowdowns. An alternative approach to the previous sections is to attempt a smooth fit and examine its slope.

Figure 3.4 The unfiltered differenced time series of `edge' - `middle' with the fitted model : a third-degree polynomial without a break.

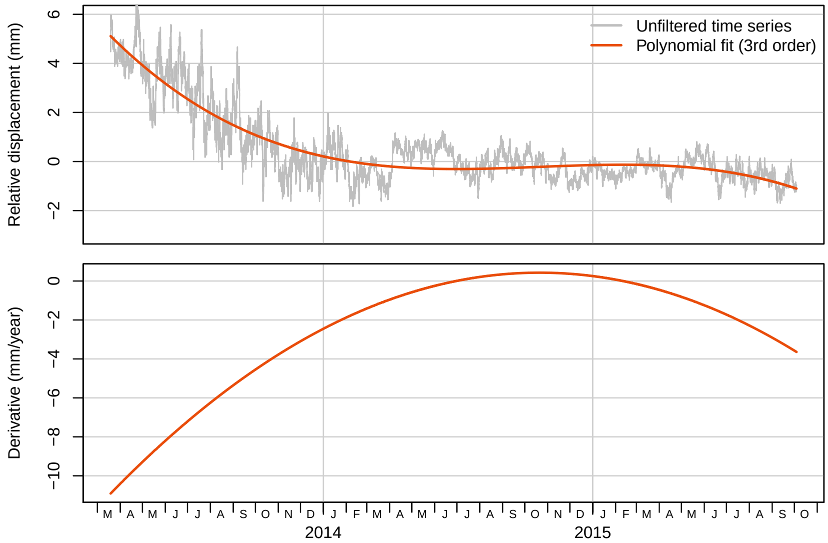


Fig. 3.4 shows a fit of a third-degree polynomial to the unfiltered differenced time seris, with in the lower panel the time derivative of the fit. This shows that the rate of subsidence in absolute value first decreases, and in the most recent months appears to be increasing again. The latter might be an end-point effect and should probably await further confirmation when a further half year of monitoring is available.

4 Re-processing of GPS data by TU Delft

The GPS data used has been processed from the original satellite measurements using one standard algorithm and associated method. In order to assess the effect of some of the assumptions regarding sources of error and their effect, a different algorithm has been applied, developed at the TU Delft. Their dataset of reduced GPS signal per GPS ground location has been made available, so that the same difference technique can be applied, as has been done for the previous reduced GPS datasets.

Fig. 4.1 shows the 'sagging' of the stations in group 'Middle' and in group 'South' with respect to those in group 'Edge' , ie. the equivalent of what is shown in fig.2.3.

The differential displacement between stations that are all within group 'Middle', in the sense (Zeerijp+ten Post) – (Stedum+Overschild) is shown as an orange line in fig. 4.2. The other measure concerns the differential displacement between stations that are all within group 'Edge', in the sense (Usquert+Veendam) – (Delfzijl+Eemskanaal) shown as a purple line. In the absence of remaining biases these differential measurements should not show any trends. Comparing fig. 4.2 with fig. 2.4 shows that for the differential time series between the stations that are all in the group 'Edge' the behaviour after October 2014 is very different. This would appear to

Figure 4.1 The time series of the TU Delft processed GPS height data, and after filtering, for the differences between the group averages `Middle'−`Edge' (in red), and `South'−`Edge' (in green). Note that the vertical scale is in mm.

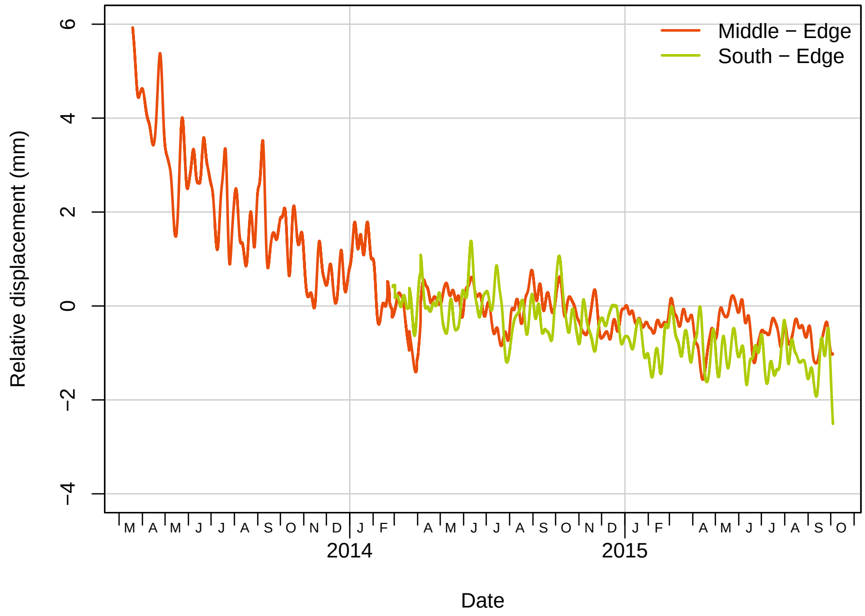
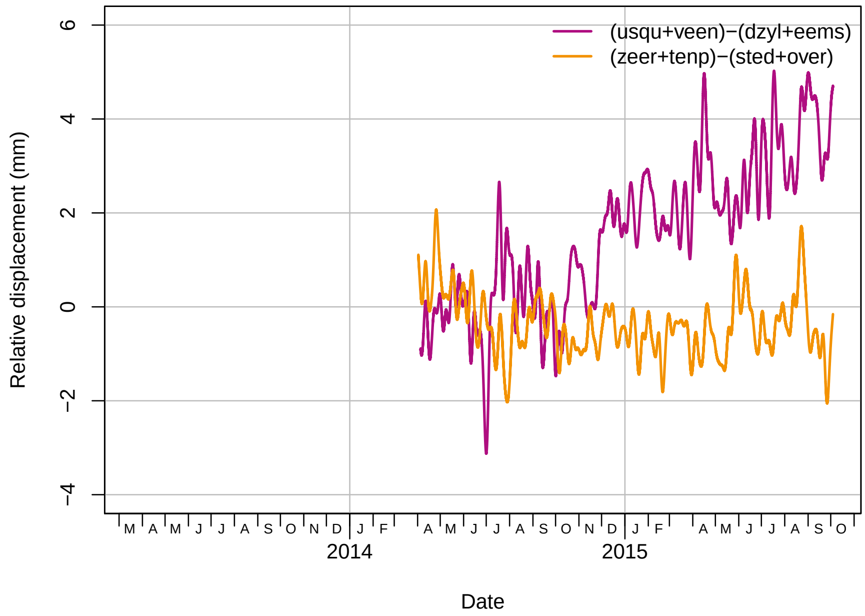


Figure 4.2 The differential time series of the TU Delft processed GPS height data, and after filtering, within groups. Purple line: differential measurements between stations all within group `Edge'. Orange line: differential measurements between stations all within group `Middle'.



indicate that in that epoch there are indeed some calibration issues in the earlier dataset which the reprocessing by the TU Delft have resolved. The same epoch does still show some anomalies for the time series involving stations within the group 'Edge'. The most likely cause is that the station of Usquert shows rather different behaviour from all other stations in that there appears to be very little subsidence at all.

4.1 Fitting to the TU Delft data without additional filtering

While the fitting procedures used are robust to the precise properties of the errors, it is worthwhile to explore to what extent the point-to-point correlation introduced by the filtering process affects the conclusions drawn from the trend fitting. At this stage it is worth remembering that the choice of fits does not have an underlying physical model for the subsidence. The intention is merely to assess whether one can reject the hypothesis that there is no change in average subsidence rate over the entire period of observation.

Figure 4.4 is the equivalent of fig. 3.1 but now the fitting is done on GPS data as processed by the TU Delft without any additional filtering by Statistics Netherlands. While there are differences in

Figure 4.3 AIC score as a function of the position of a break in the linear model fit of the difference between the group averages of `Middle' and `Edge'. The minimum for `jump' (red curve) is at 2014-04-05. For the `kink' (yellow curve) there are two choices which are nearly equally valid, which can be seen from the three broad minima. The final minimum on 2014-03-01 coincides within the uncertainties with what is obtained for the filtered data (fig. 3.1). The slightly better earlier minimum is on 29-11-2013.

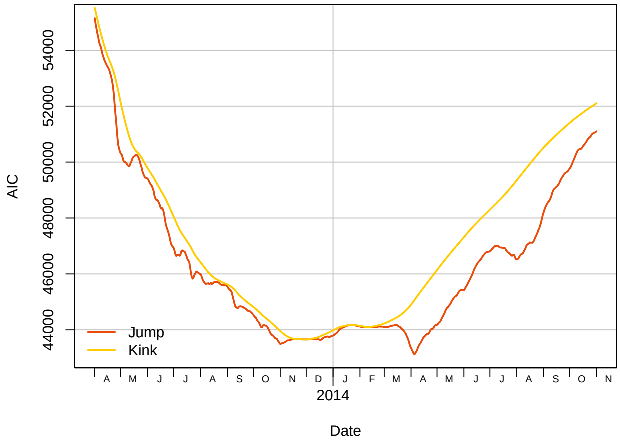
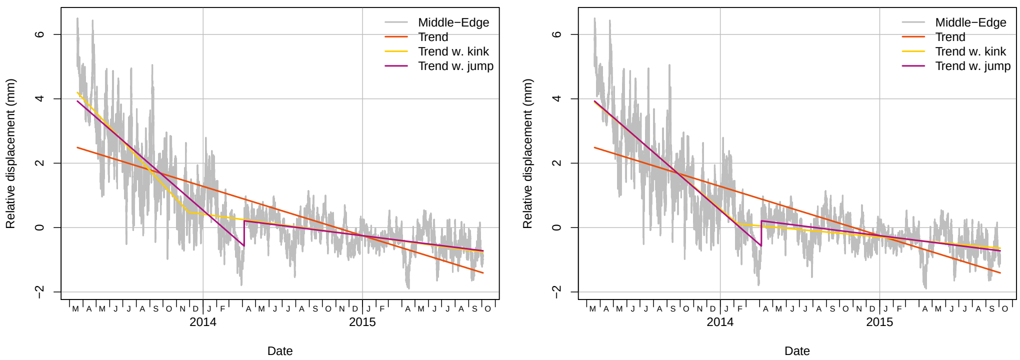


Figure 4.4 The fitted models, with either a jump or a kink, to the data as processed by the TU Delft without any additional filtering by Statistics Netherlands for the best-fitting position of that break and the two possible best-fitting positions for the kink in the linear model fit of the difference between the group averages of `Middle' and `Edge'.



the AIC for the fits as determined for the filtered and unfiltered data, as one would expect, in both cases one concludes that fits for which the average subsidence rate changes over the period of observation are superior and statistically significantly so. The data appear to point to a trend change in late March early April 2014, with an uncertainty of a few weeks either way, possibly because the change is not quite as abrupt as the fits are. With the extension of the time series in 2015 it would perhaps appear odd that the best-fit position for the kink appears to be shifted to an earlier date. However, if indeed in more recent times there is some acceleration in subsidence this would push the average slope of the section of the time series after the break down, which automatically brings about this shift. With such more complex behaviour of trends the longer the measured time series are, it is becomes necessary investigate alternative techniques for determining delay times between gas production variations and subsidence variations.

5 Response function

It is of interest to explore whether an, indirect, relationship between the time series for subsidence and for the gas production rates can be quantified, in terms of a typical delay time. One possible way is to assume that there is a linear relationship between the time series, or a relationship that can be linearised for limited lengths of time, such that the following holds:

$$h'(t) = \int r(t' - t)p(t')dt' \tag{3}$$

in which $h(t)$ is the GPS height at a given station with time derivative $h'(t)$, and p is the gas production rate, for instance as extracted at the cluster nearest to the GPS station. In this formalism $r(t)$ is a transfer function, which is in principle determined by the complex processes governing the adjustment of ground layers overlying the gas reservoir to the extraction of gas from that reservoir. This equation (3) allows for the possibility that the time variation of the production rate becomes 'smeared out' when it gets transferred to the GPS signal, so that r has a finite width as a function of delay time, around a particular peak value. Although modelling of the geophysical dynamics of the ground layers might provide such a function from first principles, even when this r is not known it can be estimated using the measured time series h and p.

The mathematical problem of determining r in such a setting is an example of a well-known class of mathematical problems commonly referred to as inverse problems. There are a number of approaches to this problem (see eg. Tarantola (2004), Pijpers (1997)). Intrinsic to this type of problem is the sensitivity to noise: inverse problems have the property that they are ill-posed and require particular care in solving because otherwise any results are dominated by the noise. While most of the GPS time series are still not long enough to attempt to apply the techniques from this field, the time series for Ten Post, with its nearby production cluster is a good candidate to explore such a technique. Figure 5.1 shows the unfiltered time series h for the GPS measurements (in mm) as well as the gas production, in 10^6 normalised m^3/d.

Given that the data are regularly sampled one can make use of Fourier transforms to facilitate the inversion. If the Fourier transform of $h(t)$ is denoted by $H(\omega)$ then the Fourier transform of its time derivative $h'(t)$ is $-i\omega H(\omega)$. The Fourier transform of the production data $p(t)$ is $P(\omega)$. It is well-known that a convolution integral such as (3) in the time domain, can be written as the ordinary product in the Fourier domain:

$$-i\omega H(\omega) = R(\omega)P(\omega) \tag{4}$$

in which $R(\omega)$ is the Fourier transform of the transfer function r. While it appears straightforward that dividing through by $P(\omega)$ and performing an inverse Fourier transform

Figure 5.1 The time series for the GPS station of ten Post (in mm), and the daily gas production at the nearest production cluster in units of 10^6 normalised m^3.

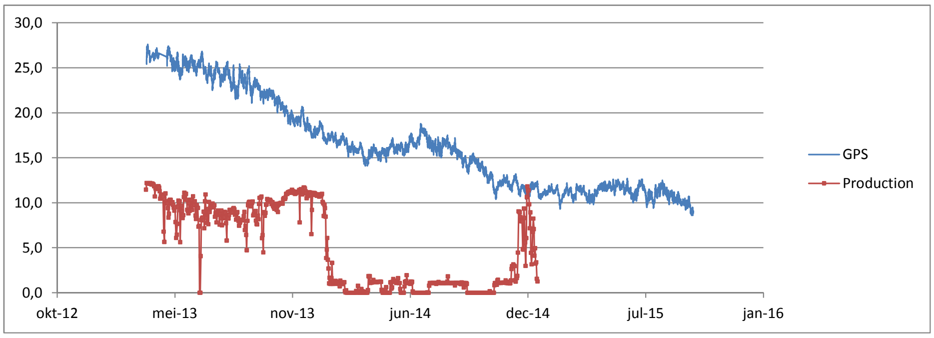
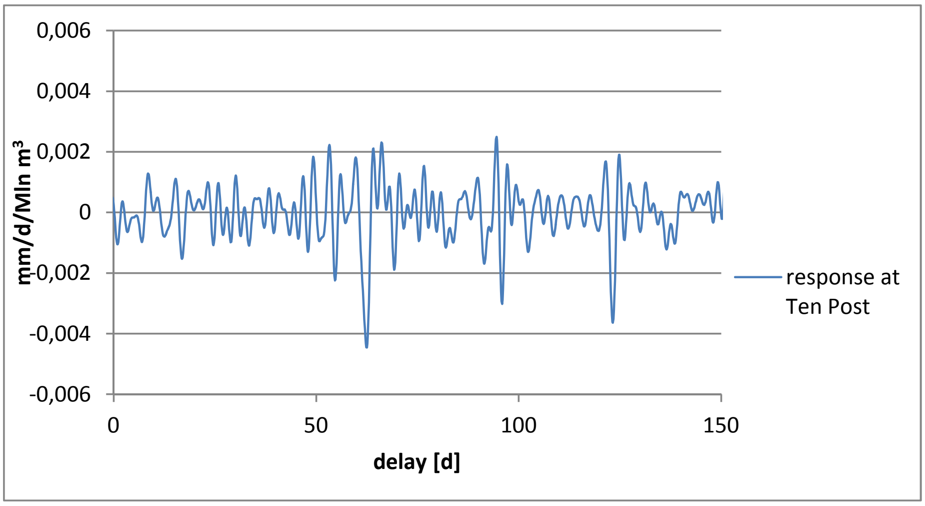


Figure 5.2 The transfer function r for ten Post as a function of time delay in days



would produce $r(t)$, the result is in practice dominated by noise, which is a consequence of the ill-posed nature of the inversion. The standard solution is to regularise the equation which in this case is achieved by calculating instead:

$$R(\omega) = \frac{-i\omega H(\omega)P^{\dagger}(\omega)}{|P(\omega)|^2 + \epsilon P_{ref}^2} \quad (5)$$

In which $P^{\dagger}(\omega)$ is the complex conjugate of $P(\omega)$.

While generally the optimal choice of the regularization parameter ϵP_{ref}^2 would be obtained by for instance cross-validation, here trial and error is used. For P_{ref} the maximum value of $|P(\omega)|$ is chosen and the parameter $\epsilon = 0.0002$. Once $R(\omega)$ is obtained an inverse Fourier transform leads to the function r as a function of time delay which is shown in fig. 5.2. The transfer function can also be determined for negative values of the time delay. This shows no features other than noise, and indeed no features would be expected other than occurring through the 'wrapping around' which occurs in discrete Fourier transforms of time series with a finite total length cf. (Bracewell, 1965). The values of the transfer function at negative time delays can be used however to establish a noise level, which results in an estimate of the standard deviation of $0.68\ \mu m/s/10^6 m^3$.

While the resulting transfer function r in fig. 5.2 is evidently fairly noisy there do appear to be a few statistically significant downward spikes, the first of which occurs at around 62 days delay. The negative value indicates that the slope of h becomes more negative as a response to increased gas extraction and vice versa, which is what one would expect. The delay time of 62 days is very similar to the 9 weeks determined from the analyses in the previous section. The subsequent spikes could perhaps be noise, or they could perhaps also be due to the influence of slightly more distant production clusters to ten Post, with a similar production rate pattern as a function of time.

6 Conclusion

In this paper we presented an update of the results obtained in previous reports (Pijpers, 2014; Pijpers and van der Laan, 2015) using new, more recent data. The results obtained here are in line with the results obtained in the previous reports.

From the GPS data it can be concluded that there is continued subsidence of the ground in the area of the wells where production was reduced in the month of January of 2014. However, the rate of subsidence is lower some time after this reduction. The location of the break is around mid-March or very early in April, ie. approximately 9 weeks after the reduction in production, but there is a fair margin of uncertainty (some 2 weeks) around the exact value of the time gap. In addition a completely independent data reduction of the GPS data, developed at the TU Delft, leads to the same conclusions in our analysis.

While there is clear statistical evidence that a break has occurred in the subsidence rates, and the reduction in subsidence speeds is measurable with a high degree of precision, research is now in progress to perform a more systematic analysis of the correlation between the production time series and the GPS height time series. First results of such an analysis, shown in section 3.3, provide additional support for the typical response time scales of subsidence to production rates of around 9 weeks.

It is important to carry out more of such extensive correlation analyses in view of the fact that at all clusters the production rate continues to vary with time, even at clusters where production was substantially reduced in january 2014. The cluster of ten Post is an example where production was partially resumed towards december of that year. However, for an analysis such as shown in section 3.3 to work well, long time series with low noise levels are required.

References

Bracewell, R. (1965). *The Fourier transform and its applications*. McGraw-Hill. (paperback edition: 1999).

Pijpers, F. (1997). The determination of time delays as an inverse problem - the case of the double quasar 0957+561. *Mon. Not. Roy. Ast. Soc. 289*, 933--944.

Pijpers, F. (2014). Phase 0 report 1 : significance of trend changes in ground subsidence in Groningen. Technical report, Statistics Netherlands.

Pijpers, F. and D. van der Laan (2015). Phase 1 update may 2015: trend changes in ground subsidence in Groningen. Technical report, Statistics Netherlands.

Tarantola, A. (2004). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress: Statistics Netherlands, Grafimedia
Design: Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
verkoop@cbs.nl
Fax +31 45 570 62 68
ISSN 1572-0314