Ensemble Based Subsidence application to the Ameland gas field – long term subsidence study part two (LTS-II) continued study

NAM

Doc nr. EP201710200509
Version 0. Status: FINAL
DD 20-10-2017

Contents

1 Summary

This report describes the work done to explain and reduce the difference between the NAM-predicted long-term subsidence and the field data in the Ameland field, The Netherlands: The subsidence appears to continue even though the depletion rates have been decreasing significantly over the past years. As a result, the maximum subsidence and the subsidence patterns predicted by previous subsidence models did not fully agree with the field data, making predictions of subsidence uncertain. The long-term subsidence study addresses this issue with part two (LTS-II) being a follow-up of part one that was finished in 2015 and outlined in NAM (2015a). The study work in part one (LTS-I) concluded on the relevant physical mechanisms that could explain the anomalous subsidence observed above the Ameland field. The main objective of the LTS-II project is the implementation of the results obtained in LTS-I into a new confrontation workflow, applied on a real field case: the Ameland field. A unique feature of this developed confrontation workflow is that it considers the uncertainty of geodetic data, the reservoir model and the geomechanical model in the forecasting of subsidence resulting from gas production.

A key element of the new workflow is that model results are compared with subsidence measurements in an objective way. The workflow incorporates the findings of the LTS-I study and as demonstrated here is able to identify the most likely model factors, like reservoir and aquifer depletion scenarios, and parameter values for the compaction model and influence function (i.e. a model that translates the compaction source to surface subsidence). Moreover, posterior (after confrontation with the geodetic data) probability distributions for the model data show redefined distributions for the input values of the model data. The workflow consists of 5 main components as visualised in Figure 1, that are summarised here.

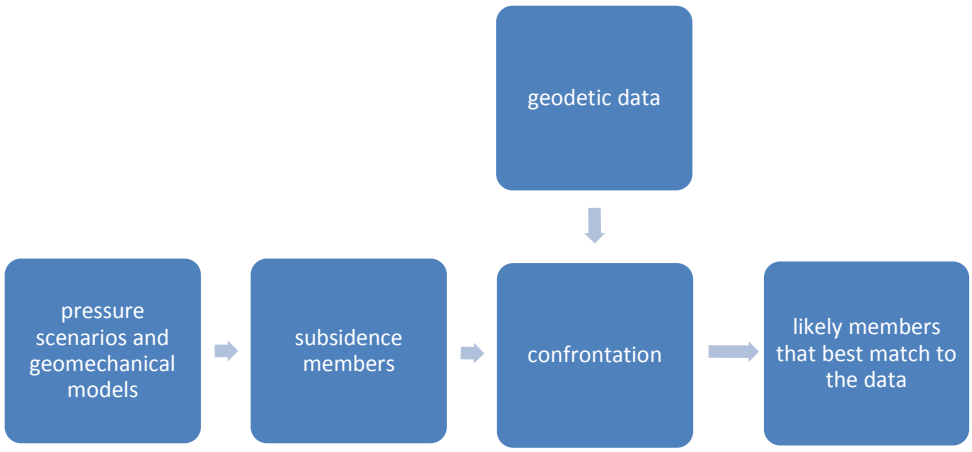


Figure 1. Main components of the confrontation workflow as adopted in this study.

Results of the LTS-II study demonstrate that a successful workflow was created that confronts subsidence model results with measurements in an objective way. The workflow incorporates the findings of the LTS-I study and as demonstrated here can identify the most likely model factors, like reservoir and aquifer depletion scenarios and parameter values for the subsidence model. Moreover,

posterior (after confrontation with the geodetic data) probability distributions for the model parameter values show redefined distributions when compared to the prior input values of the model parameters.

Pressure scenarios
To describe the possible variation in the pressure scenarios for Ameland, 58 reservoir scenarios were created that were all history matched to the available pressure and production data. The scenarios are characterized by different levels of depletion in time of both the western and eastern aquifers. Because of the long production history of the field (since 1986) combined with the regular measurements of the pressures in the producing wells, the development of the pressure in the gas bearing part of the structure is narrowly constrained and therefore shows little variation between the 58 models.

Subsidence members
For every pressure scenario, the parameter values of the compaction model and influence function were varied in a Monte Carlo simulation. In the current study, one generic compaction model is used, the RTCiM model (Pruiksma et al, 2015), with the addition of a linear elastic branch. For the influence function, a modified Geertsma and van Opstal (1973) model is used with the addition of a time dependent shape factor that is calibrated to the viscous behaviour of a salt layer embedded in an elastic overburden, modelled by a finite element model. The two time-dependent processes (one in the compaction and one in the influence function) have a distinctly different effect on the subsidence rate. Subsidence as a result of reservoir compaction always leads to a change of the subsidence bowl *volume*. In contrast to this, the viscous behaviour of the salt only impacts the *shape* of the subsidence bowl and not the volume.
Each parameter picked by the Monte Carlo procedure results in a subsidence model member, with a group of members defining the ensemble.

Geodetic data
Several significant innovations were applied in the use of geodetic data. Observations from levelling and GPS techniques have been used in the format of spatio-temporal (i.e. in space and time) double-differences. Uncertainties are described by a fully populated covariance matrix, that also takes shallow movements into account (idealisation noise). Outlier removal has been implemented in a formal way.

Confrontation
Each member is confronted to the geodetic data. The resulting test statistic defines the probability and weight of the specific member. The used theory to calculate the test statistic is based on Nepveu et al. (2010) with modifications, explained in this report.

Output and results
The modelled results versus the geodetic data are presented for two E-W profiles over the island and for many individual double-differences. The overall fit to the data is good, matching better to the location of the deepest point of the bowl and matching better to the benchmarks on the eastern part of the island as compared to results from earlier models. The weighted average was also used to calculate the subsidence rates in the Pinkegat area along with their confidence intervals. It is observed that the confidence interval is narrow and is, at present, far below the defined boundary for the rates (called the “meegroeivermogen”).

The same workflow has been applied to two “emergency stop scenarios”. The results show a clear reduction in the subsidence rates after a hypothetical stop in 1996, but also smaller but clear reduction following a hypothetical stop in 2016. The latter being smaller can be explained by the reduced effect of a later stop on the pressures in the gas field, knowing that most of the gas has been produced by 2016. For a hypothetical stop in 1996, the production stop is ten years after the start of production, which is comparable to the current field life of the Waddenzee fields like Nes and Moddergat. Based on this analogue it can therefore be concluded that an emergency stop scenario in the Waddenzee for fields like Nes and Moddergat would result in a significant decrease of the subsidence rate in the Waddenzee.
A verification of the effect of more geodetic data on the accepted model ensemble members when bringing in longer time series has also been performed as part of this study. The results show a narrowing band specifically in the early monitoring period, i.e. early measurement campaigns quickly narrow the uncertainty.

In summary, the application of further detailed (statistical) analysis to pressure scenarios, field data and geomechanical models improved our ability to describe the Ameland subsidence, giving confidence in the ability of our models to predict subsidence in the Ameland field.

2 Introduction

Thanks to extensive subsidence monitoring, NAM observed that subsidence at the Ameland field continues after production-induced depletion has slowed down significantly from the rates in the first decades of the field (80s and 90s). More precisely, the rate of subsidence per unit depletion (at a particular location above the Ameland field) is increasing over time, suggesting a time-dependent (combination of) mechanism(s) at work in or around the reservoir.

This report describes the results of the Long-Term Subsidence Study part 2 (LTS-II), which is a follow-up study of part 1 (LTS-I, NAM, 2015a). Both studies are aimed to better understand the observed spatial and temporal behaviour of subsidence measurements above the Rotliegend gas reservoirs in the northern part of the Netherlands, with LTS-I focussing on the possible physical causes and LTS-II applying this knowledge for an actual field case: subsidence above the Ameland field.

The LTS research is framed around geodetic observations above the Ameland field that show a continuing subsidence even after the rate of reservoir pressure depletion has slowed down significantly. This behaviour was not well understood in 2011 but was mathematically resolved by the introduction of a function describing a time-dependence of subsidence on depletion (the Time decay model, Mossop, 2012). This new function resolves largely the mismatch between model predictions and subsequent survey measurements, but the possible underlying physical mechanisms were not properly understood. This situation was unsatisfactory for NAM, SodM and other stakeholders like “de Waddenvereniging” and led to a condition in the approval of the licence (“winningsplan”) that allows gas production in the Waddenzee area.

As mentioned before, LTS-I (NAM, 2015a) focused on the possible mechanisms that could explain such observations. The main conclusions from this study are listed below.

2.1 Summary of part 1 (NAM, 2015a)

The main findings of the LTS-I study are listed below:

- The time dependent subsidence is “real”, i.e. it results from physical mechanisms operating in and around the Ameland reservoir and not an artefact of noise and uncertainty in the geodetic data.
- Time dependent creep behaviour is observed and predicted to be associated with compaction of the sandstone in the gas reservoir, pressure diffusion and partial depletion of the aquifers as well as flow of the overlying salt. Salt flow in isolation appears not to be a plausible explanation for time dependent subsidence, while the compaction and pressure depletion models remain viable hypotheses within the possible uncertainty ranges.
- Deformation experiments of Rotliegend reservoir core material under in-situ conditions show that reservoir compaction involves a porosity-dependent element of inelastic deformation through grain cracking and an elastic (reversible) element. The contribution of non-reversible inelastic strain increases with increasing porosity.
- The subsidence modelling precision can significantly be improved by taking correlation structures in the surveillance data into account. By appropriately differencing the survey data, biases as well as complexities in covariance structures can be reduced. In addition, methods have been developed to identify and handle outlier measurements, data reduction

techniques for large geodetic data sets, as well as improvements to processing, including GPS data.

- An improved and more formal statistical method is proposed to validate and test the quality of subsidence predictions against the survey data. It is based on a Bayesian framework that can provide a coherent structure for the creation of initial models built on prior information, the objective updating of these models using collected geodetic data and the quantitative testing of future predictions.

2.2 Previous LTS-II reports

The LTS-II research started in March 2016 by a consortium of NAM, TNO and TU Delft. A workplan (NAM, 2016d) was written with the aim to deploy the knowledge on geomechanics and ensemble based confrontation techniques by TNO (TNO, 2017) and geodetic processing by TU Delft (van Leijen et al., 2017) on Ameland. The application to the Ameland field and the results of the first confrontation of the subsidence model members to the geodetic data are described in two reports (NAM, 2017, 2017b).
Even though significant technical progress was recognised by the reviewers, the Dutch regulator SodM (State Supervision of Mines) concluded that the work should be improved by a number of modifications (SodM letter, 2017), to be addressed in a follow-up study (this report):

1. Improve the fit of the subsurface models to the measured data (see Section 7.4)
2. Include the effect of the large observed overpressure in the subsurface model (see Section 4.2.2)
3. Use a single generic compaction model instead of 4 specific compaction models (see Section 4.2.1)
4. Use a realistic range of salt parameter values (see Section 4.3.1)
5. Consequently, consider some neglected co-variances (see Chapter 6)
6. Resolve geomechanical modelling issues that were introduced by the TNO program AEsubs (see Section 4.3)

The current report is written as a stand-alone report, i.e. it can be read without first digesting the previous LTS reports. It describes in detail the latest insights in Ameland subsidence modelling, and presents a complete overview of the confrontation workflow, addressing the six suggestions for improvement listed in the SodM letter.

3 Static and dynamic reservoir modelling of the Ameland area

3.1 Summary of the static model

The static model provides the properties to the dynamic reservoir model. The model is described in detail in NAM (2010). A summary of that work is presented in below.

The objective of the study in 2010 was to prepare dynamic reservoir model scenarios that incorporate the key uncertainties that pertain to subsidence modelling of the Ameland gas fields and the aquifers. Before 2010, a Petrel static reservoir model was built in 2005 for late life field development planning using the analysis of the comprehensive Ameland Field Review (subsurface study during 2005-2007). Subsequently, a Petrel static reservoir model rebuild was undertaken in 2009 to incorporate new insights and analysis based on field performance and well results. In 2010, these models were deemed unsuitable to underpin subsidence modelling for an update of the Ameland Winningsplan, hence, a full subsurface model rebuild was undertaken, described in the NAM (2010) report. The main characteristics of this model are:

- Rebuild from scratch delivering a new base case static model for a new large dynamic model that delivered pressure information for subsidence modelling in the whole Waddenzee area, including Ameland fields and Waddenzee fields (Ameland-Oost, Ameland-Noord, Ameland-Westgat; Nes-Noord area; Ternaard; Nes, Moddergat, including all southern aquifers (Figure 2).
- Vertical layer thickness of only 1 m.
- 2 scenarios based on porosity distribution. The scenario based on a continuous porosity deemed to be more likely and is chosen as the base case static model that serves the dynamic model in the LTS-II study.
- Application of new (sequential Gaussian simulation) techniques for property modelling honouring geological trends.
- Based on the 2009 PreSDM (Prestack depth migration) seismic cube.

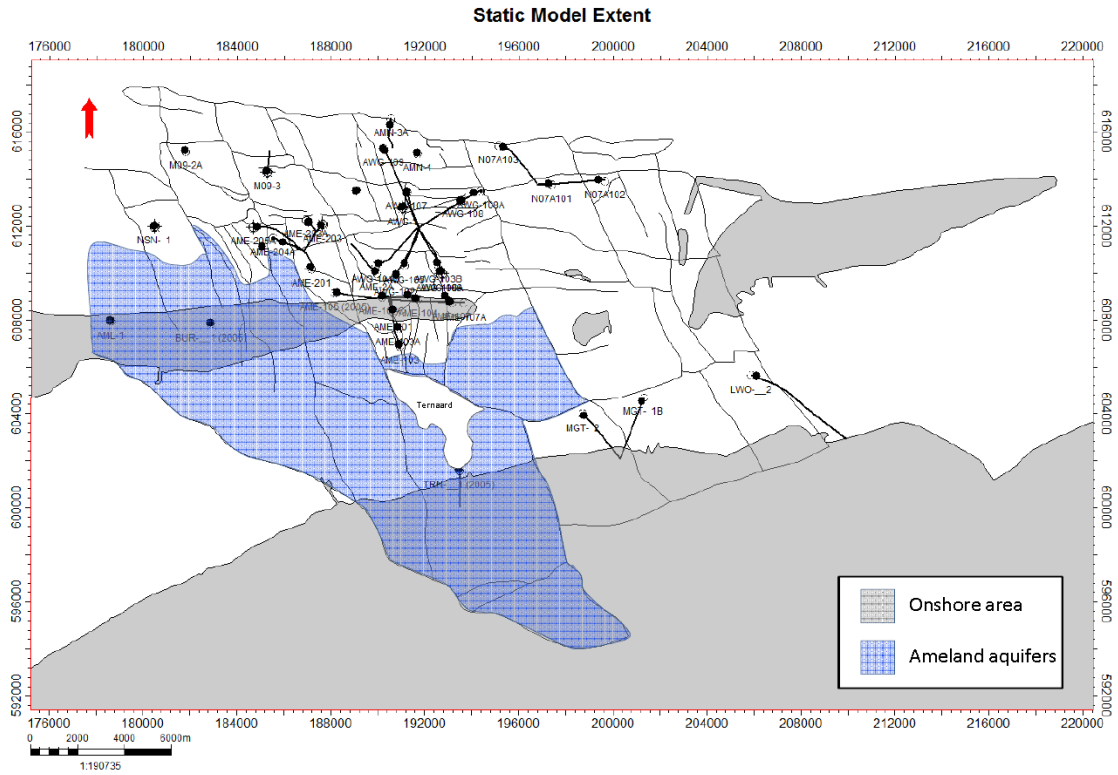


Figure 2. Outline of Petrel2009 model showing model coverage of aquifers and Ternaard gas field. The model extends into the Waddenzee fields (Modddergat, Nes, and Lauwerszee-Oost) and onshore NL to cover areas of surface subsidence monitoring. Also, the well paths of the wells used in the NAM (2010) report are visualised.

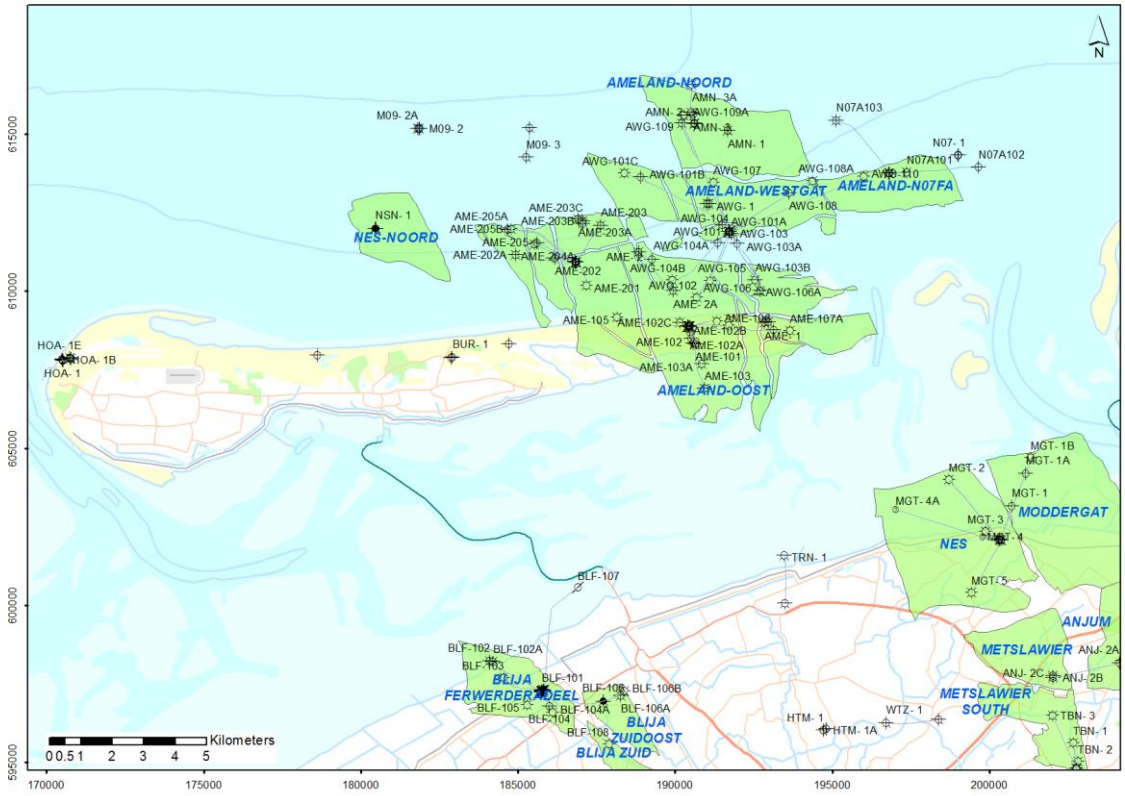


Figure 3. Overview of the fields and drilled wells in the Waddenzee area (status 2017).

3.2 Dynamic modelling

As described in the previous section, a base case static model was constructed and used to build the dynamic model. Even though the full area of interest is present in the dynamic model, it was decided to only investigate the effect on the subsidence because of the production from the Ameland-Oost, Ameland-Westgat and Ameland-N07FA fields (Figure 3). The production of the Ameland-Oost field will have a dominant effect on the subsidence data because:

- The geodetic measurements are above or close to this field
- It is the biggest field (area- and Gas Initially In Place (GIIP)-wise)
- It has the highest cumulative gas production
- It has, at present, the lowest reservoir pressure and therefore the highest depletion.

Therefore, in the sections below, the focus will be on the Ameland-Oost field, and the approach, applied in the dynamic modelling domain, will be explained.

3.2.1 Ameland-Oost field overview

The Ameland-Oost field is in the Dutch part of the Southern North Sea area. The field was discovered in 1964 by well AME-1 drilled from the Ameland Island. The field contains several intra field faults, further sub-dividing the field into several producing blocks with different degree of connectivity across the faults. The fault blocks in this field are E11-15, E12-13, E14, E21, E22, E23, E24, E25 and E27 (Figure 4).

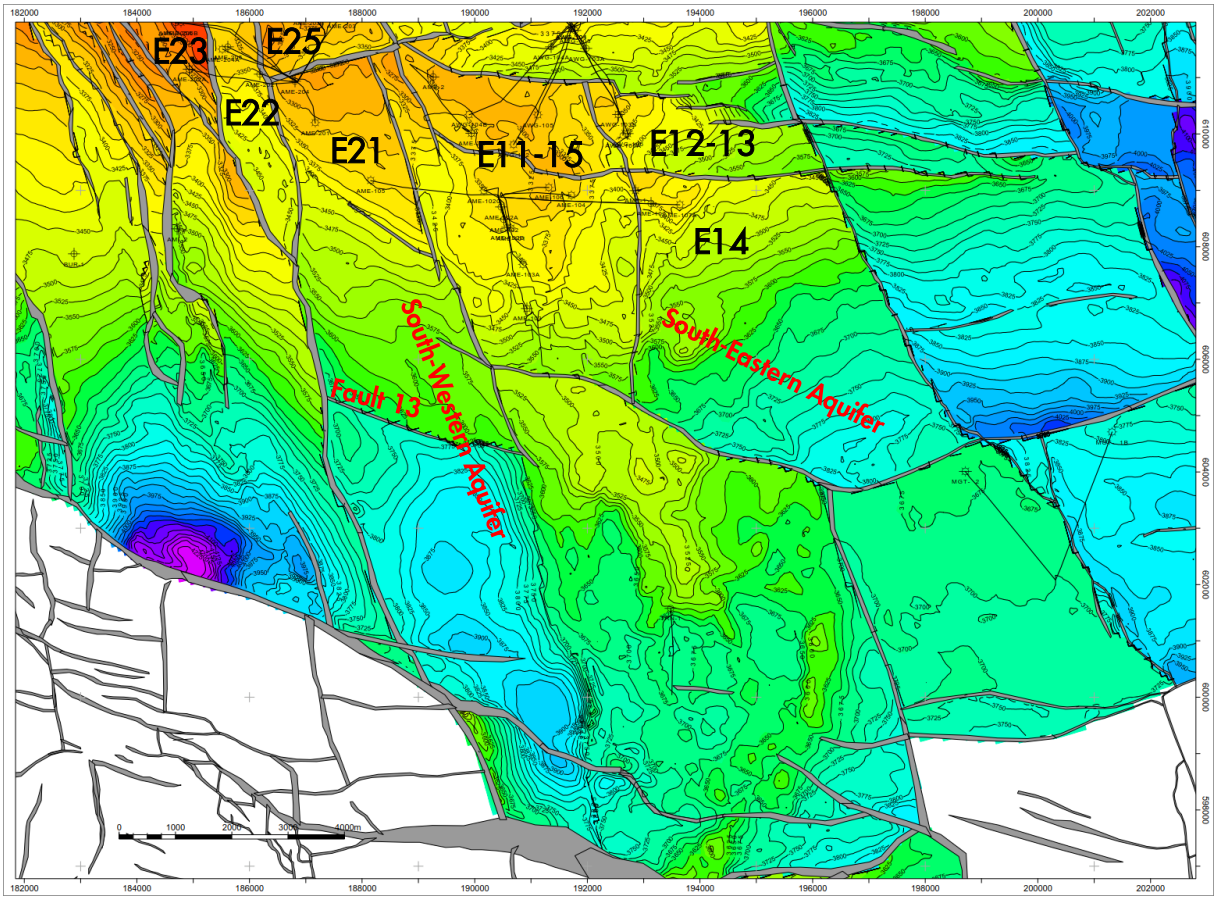


Figure 4. Ameland-Oost field overview. The contour lines are the depth of the top of the reservoir in meters.

3.2.2 Ameland-Oost field development history

Gas production from the field commenced in 1986 with wells AME-102, AME-104 and AWG-102, which were drilled into the central part of block E11-15, the biggest fault block of the Ameland-Oost field (Figure 4). Drilling continued, and in 2009 the last Ameland-Oost well was drilled, completed and brought to production.

The initial reservoir pressure of the field (as measured by several Repeat Formation Tests, giving direct bottom hole pressure measurements) was 555 bar. Today, the pressure varies significantly across the field, ranging from 45 bar in the central part (block E11-15) up to 400 bar on the edge of the gas reservoir (block E23). The main reasons for such a difference in reservoir pressures are:

1. the geological structure of the field: Ameland-Oost consists of several fault blocks with a different degree of inter- and intra-block communication
2. production timing: production from E11-15 block started in 1986, while block E23 was brought to production only in 2009.

3.2.3 Dynamic modelling process

The current dynamic model is based on a static model built in 2010. It is the same model that is used for the measurement and control cycle (NAM, 2016a), i.e. an annual close the loop cycle that allows for the calibration of subsurface models to the acquired data. The long production history of Ameland, together with the firm geological concept mean that the remaining uncertainty in the static model has only limited effect on the pressure uncertainty over the gas field itself as the input is constrained by the actual pressure measurements.

The main idea behind the reservoir simulation exercise during this study was to achieve variability in reservoir pressure while honouring the existing field pressure and production data, i.e. while maintaining a history match. In this way, possible depletion (hence, subsidence) variation could be investigated.

Since the gas part of the field is well covered by producing wells and there is 30+ years of production and pressure measurement history, significant changes in reservoir pressure in the gas part would distort the history match. On the other hand, there is no well (i.e. no pressures measurements) control in the aquifers in the southeast and southwest of the field (Figure 4). Hence, the applied approach was:

1. to investigate the possible variation for several reservoir parameters by changing each within a defined range,
2. while maintaining history match in the gas part of the field, and
3. see what effect it will have on the depletion in the south eastern and southwestern aquifers.

Reservoir parameters, which were selected to be altered during this study, were (in italic, the parameter name is defined):

- Multiplier to the residual gas saturation in aquifers – *Sgr*
- Fault 13 transmissibility in the southwestern aquifer (E21 block) – *Fault_13*
- Permeability multiplier in the southwestern aquifer (E21 block) – *E21_Kaqf*
- Permeability multiplier in the southeastern aquifer (E11-15 + E14 block) – *E11_15_Kaqf*
- GIIP multiplier in the central part of the Ameland-Oost field (E11-15 block) – *E11_15_GIIP*

The first four parameters would directly affect the behaviour and strength of the aquifers while the fifth parameter was used to balance out energy increase coming from aquifers to the gas leg to maintain the history match of the gas bearing part to the production data.
In more detail, the *Fault_13* parameter was used to investigate the possible degree of communication in the southwestern aquifer that could affect the subsidence in that area; the *GIIP* multiplier was mainly used as a matching parameter to account for effects caused by *Sgr* and *Kaqf* changes, which, in turn, were the main parameters defining aquifer depletion.

Values for the parameters were picked up from the ranges described below.

Table 1. Range of values for Sgr.

Parameter	Value range		
Sgr	0	0.5 (5%)	1 (10%)

Sgr value of 0 would mean no residual gas saturation in the aquifer at all, while the values 0.5 and 1 would result in 5% and 10% residual gas saturation respectively.

Table 2. Range of values for Fault_13.

Parameter	Value range	
Fault_13	0	1

The values 0 and 1 were used in the model as seal factor for the Fault 13, which separates the aquifer of the E21 block in 2 parts. A seal factor of 0 represents a fully sealing fault while a seal factor 1 makes this fault fully transmissible.

Table 3. Range of values for E21_Kaqf.

Parameter	Value range			
E21_Kaqf	0.01	0.1	1	5

These numbers were used not as absolute permeability values in the model, but instead as multipliers to the current permeability/transmissibility areal distribution in the aquifer part of the E21 block.

Table 4. Range of values for E11_15_Kaqf.

Parameter	Value range								
E11_15_Kaqf	0.001	0.005	0.01	0.025	0.05	0.1	0.25	0.5	1

Just like with the E21_Kaqf parameter, these values are permeability/transmissibility multipliers for the south-eastern aquifer of the Ameland-Oost field.

Table 5. Range of values for E11_15_GIIP.

Parameter	Value range		
E11_15_GIIP	0.88	0.9	0.91

These values were used as multipliers to the volume of Gas Initially In Place (GIIP) in the E11-15 block of the Ameland-Oost field.

Several (dynamic) subsurface realizations were generated in the work process. Parameter ranges were tested to find minimum and maximum values, beyond which models' behaviours deviate significantly from the observed historical performance of the field. The effect of those parameters and their mutual interdependencies were thoroughly analysed, i.e. for permeability multipliers, high

values of the parameter will make the aquifer “very active”, resulting in high reservoir pressure in the gas part and very high water production from the wells (what is not observed). On the other hand, low values of permeability multiplier parameters will make the aquifer “weak” (or, in other terms, immobile). In this case, there is virtually no depletion, hence, no subsidence, in the aquifer, while some level of subsidence is observed in that area. Increase of residual gas saturation (Sgr) brings more energy to the aquifer, making it more “active”. To counter this effect in the gas part of the field (to preserve history match there), permeability multipliers and GIIP multipliers had to be reduced accordingly. This explains how the boundaries of the abovementioned ranges were defined.

The careful selection of exact values for different parameters allowed generating a set of history matched scenarios, from which the final set of 58 unique scenarios was selected. All the scenarios from the final set are acceptably history matched in the gas part of the field, while allowing significant variation in aquifers depletion. The parameters for these 58 scenarios are listed in Table 6.

Table 6. Final set of selected 58 models with their parameters' values.

Model Name	Sgr	Fault_13	E21_Kaqf	E11_15_Kaqf	E11_15_GIIP
90	1	0	1	1	0.88
91	1	0	0.1	1	0.88
92	1	0	0.01	1	0.88
93	1	1	1	1	0.88
94	1	1	0.1	1	0.88
95	0.5	0	1	1	0.88
96	0.5	0	0.1	1	0.88
97	0.5	0	0.01	1	0.88
98	0.5	1	1	1	0.88
100	0	0	5	1	0.88
101	0	0	1	1	0.88
102	0	0	0.1	1	0.88
103	0	0	0.01	1	0.88
104	0	1	1	1	0.88
902	1	0	1	0.05	0.9
903	1	0	1	0.025	0.9
981	0.5	1	1	0.1	0.88
982	0.5	1	1	0.05	0.91
983	0.5	1	1	0.025	0.9
1001	0	0	5	0.5	0.88
1002	0	0	5	0.25	0.88
1003	0	0	5	0.1	0.88
1011	0	0	1	0.5	0.88
1012	0	0	1	0.25	0.88
1013	0	0	1	0.1	0.88
9001	1	0	1	0.001	0.9
9002	1	0	1	0.005	0.9
9003	1	0	1	0.01	0.9
9101	1	0	0.1	0.001	0.9
9102	1	0	0.1	0.005	0.9
9103	1	0	0.1	0.01	0.9
9201	1	0	0.01	0.001	0.9
9202	1	0	0.01	0.005	0.9
9203	1	0	0.01	0.01	0.9
9301	1	1	1	0.001	0.9
9302	1	1	1	0.005	0.9
9303	1	1	1	0.01	0.9
9501	0.5	0	1	0.001	0.9
9502	0.5	0	1	0.005	0.9
9503	0.5	0	1	0.01	0.9
9601	0.5	0	0.1	0.001	0.9
9602	0.5	0	0.1	0.005	0.9
9603	0.5	0	0.1	0.01	0.9
9701	0.5	0	0.01	0.001	0.9
9702	0.5	0	0.01	0.005	0.9
9703	0.5	0	0.01	0.01	0.9
9801	0.5	1	1	0.001	0.9
9803	0.5	1	1	0.01	0.9
10001	0	0	5	0.01	0.9
10002	0	0	5	0.05	0.9
10101	0	0	1	0.01	0.9
10102	0	0	1	0.05	0.9
10201	0	0	0.1	0.01	0.9
10202	0	0	0.1	0.05	0.9
10301	0	0	0.01	0.01	0.9
10302	0	0	0.01	0.05	0.9
10401	0	1	1	0.01	0.9
10402	0	1	1	0.05	0.9

3.2.4 Simulation results

To monitor pressure changes in the aquifer and to check that the history match in the gas part is still preserved, several locations were selected in the model. This also allowed more efficient and visual comparison of the models against each other. Figure 5 below depicts the position of these locations.

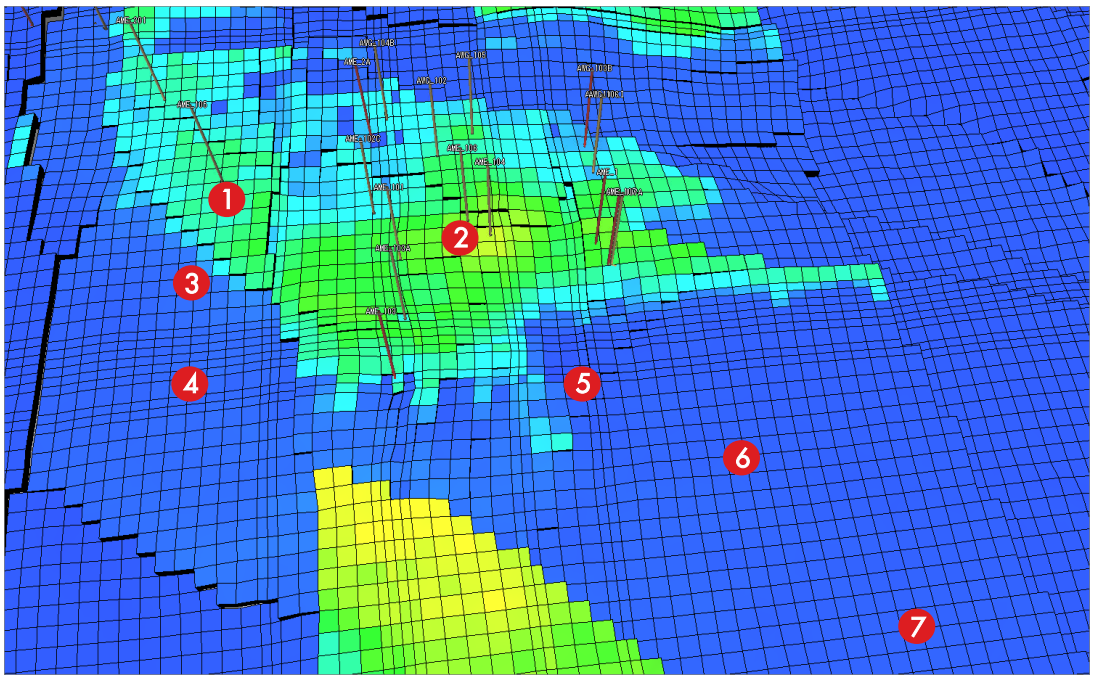


Figure 5. Selected locations for pressure/water production control in Ameland-Oost.

For history matching purposes, it was important to select locations close to existing wells in the main producing blocks, so a direct (with all associated uncertainties involved) pressure measurement history would be available. Additionally, modelled water production in those locations was calculated and compared to measured data. Ameland does not produce much water, so all subsurface scenarios with excessive water production in control points were excluded during the simulation workflow.

Location 1 in block E21 is the AME-105 well, the well closest to the southwestern aquifer. This is an appropriate location to check the water production rates and reservoir pressures, since the possible increase in the southwestern aquifer strength will be first felt in this well.

Location 2 in block E11-15 is the AME-106 well, one of the main production wells in the Ameland-Oost field, which is still producing and has a reliable pressure data history.

By checking these 2 locations, it could be very quickly confirmed whether the history match was still preserved after varying certain parameters.

There are no wells drilled into southern aquifers of the Ameland-Oost field. To get an understanding of how parameters influence the aquifer pressures and to what extent, several control locations were selected based on the geometry of the aquifers. Locations 3 and 4 were selected in the northern and central parts for the southwestern aquifer (E21 aquifer).

Locations 5, 6 and 7 were chosen in the northern, central and southern parts of the south-eastern aquifer (E11-15 + E14 aquifer).

Figure 6 and Figure 7 below show the quality of the history match in Locations 1 and 2 respectively. The lines on those plots are simulation results (reservoir pressure on the left, water production on the right) for all 58 selected models. Red diamonds on the left plots are measured reservoir pressures observed in wells AME-105 (Location 1) and AME-106 (Location 2). There are no direct water measurements available for the individual wells, so the water production check here is more qualitative, rather than quantitative. To put in numbers – it should not exceed the value of approximately 50 m^3/day. Note that the present measured Ameland water production so far has always been very small (50-60 m^3/day).

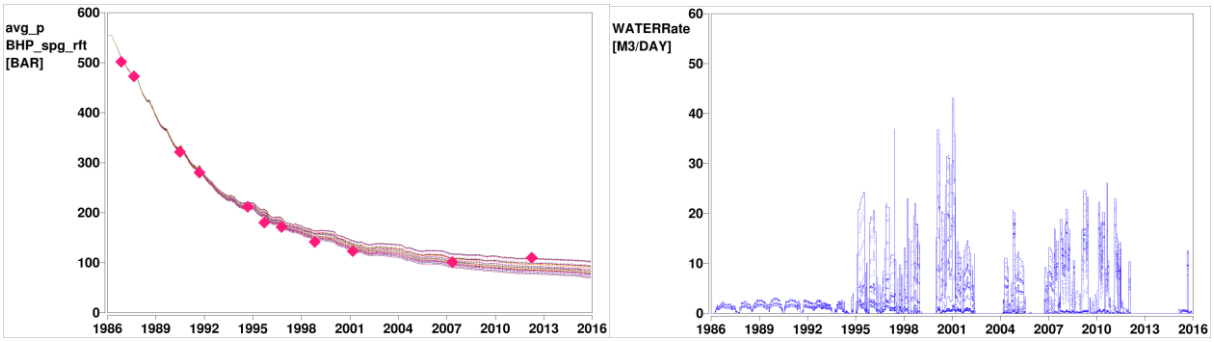


Figure 6. Left: History match quality of 58 models in Location 1. Red dots in the left-hand graph represent the data while the lines show the modelled results. Right: modelled water production.

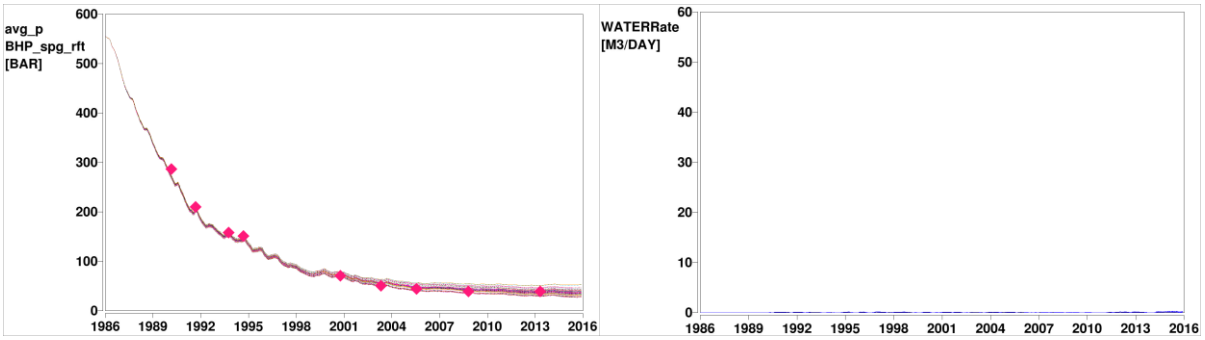


Figure 7. Left: History match quality of 58 scenarios in Location 2. Red dots in the left-hand graph represent the data while the lines show the modelled results. Right: modelled water production.

In Location 1, the modelled reservoir pressures follow the measured data and, together with relatively low water production, it is concluded that Location 1 is history matched.
In Location 2 an almost perfect history match is achieved with very minor deviation of simulated pressures from observed ones. Virtually no water production is observed in Location 2, which is also in line with the actual observations.
Altogether this means, that though some variations are there, all 58 selected models are history matched. Figure 8 and Figure 9 below show the variation in reservoir pressure in Locations 3 and 4, respectively, for the same set of 58 history matched models. Figure 10, Figure 11 and Figure 12 show the reservoir pressure is in Locations 5, 6 and 7 respectively for the same set of models. These figures do all show a large variation of possible pressures in the aquifer, up to 500 bar.

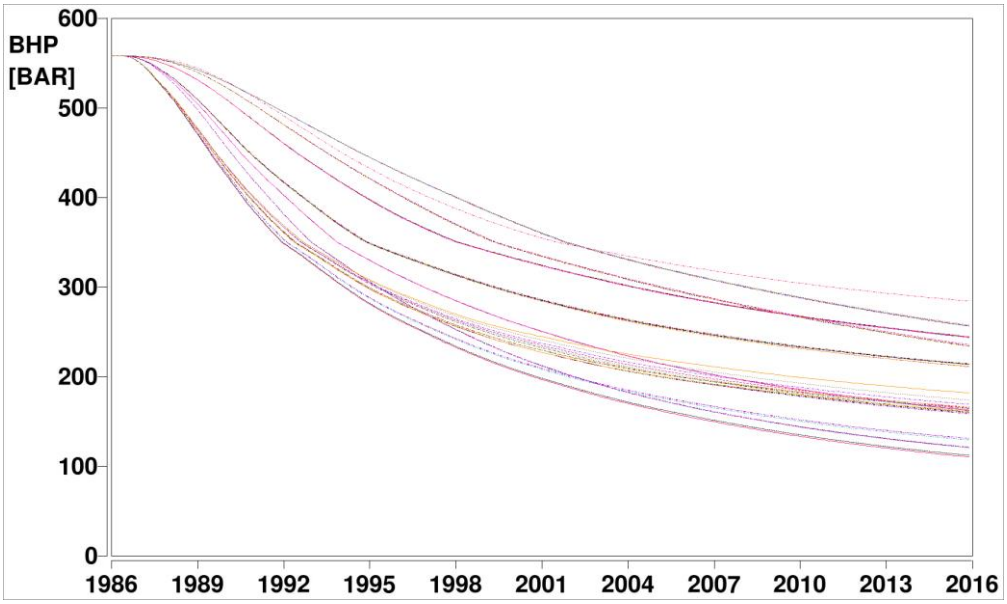


Figure 8. Pressure variation with time for the 58 scenarios in the northern part of the southwestern aquifer (Location 3).

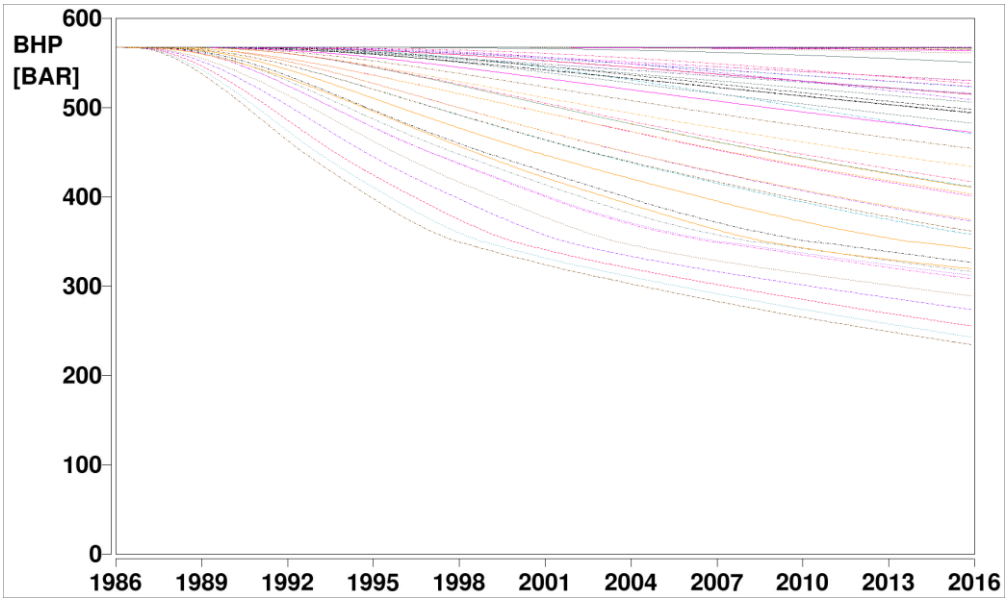


Figure 9. Pressure variation with time for the 58 scenarios in the central part of the southwestern aquifer (Location 4).

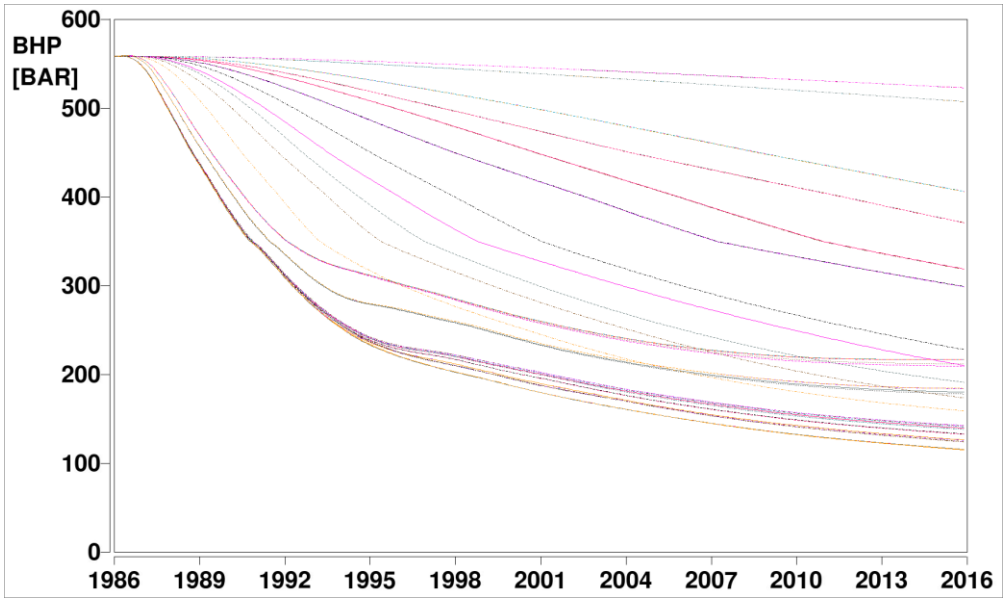


Figure 10. Pressure variation with time for the 58 scenarios in the northern part of the south-eastern aquifer (Location 5).

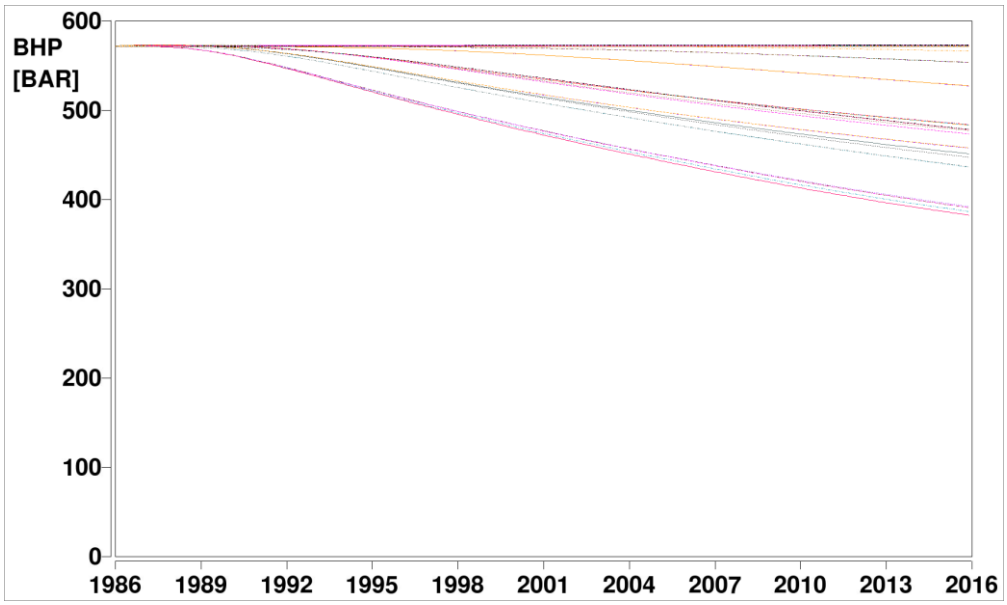


Figure 11. Pressure variation with time for the 58 scenarios in the central part of the south-eastern aquifer (Location 6).

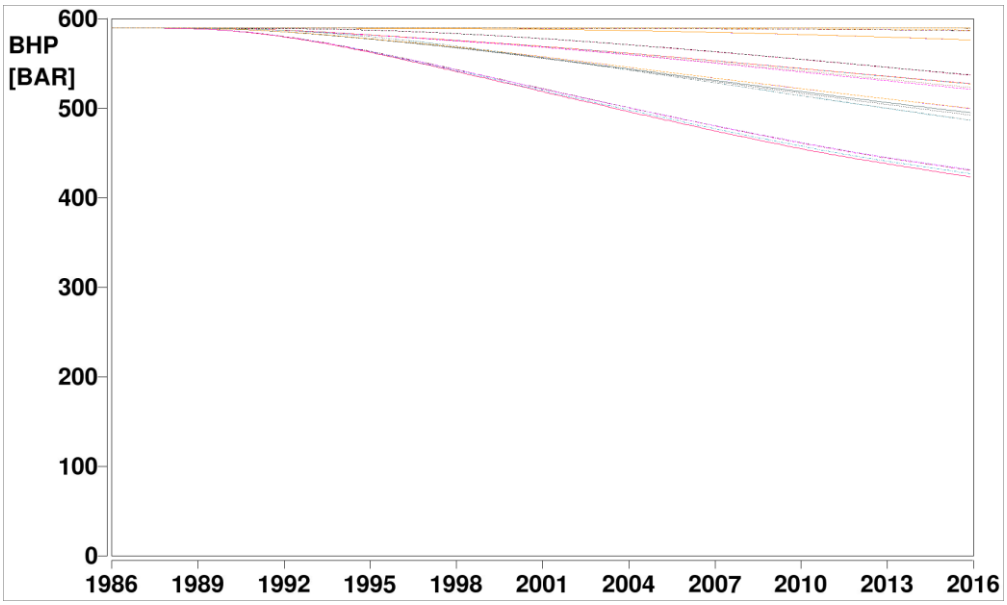


Figure 12. Pressure variation with time for the 58 scenarios in the southern part of the south-eastern aquifer (Location 7).

Overall, this result of the simulation study shows (Figure 13), that even though a history match in the gas part of the field is achieved, aquifer depletion can vary significantly – up to a few hundred bar for the same location. This adds significantly to the possible variation in subsidence in the area, since subsidence (to large extent) is driven by pressure depletion. The full set of 58 history matched scenarios was handed over to the geomechanics team for further study and investigation.

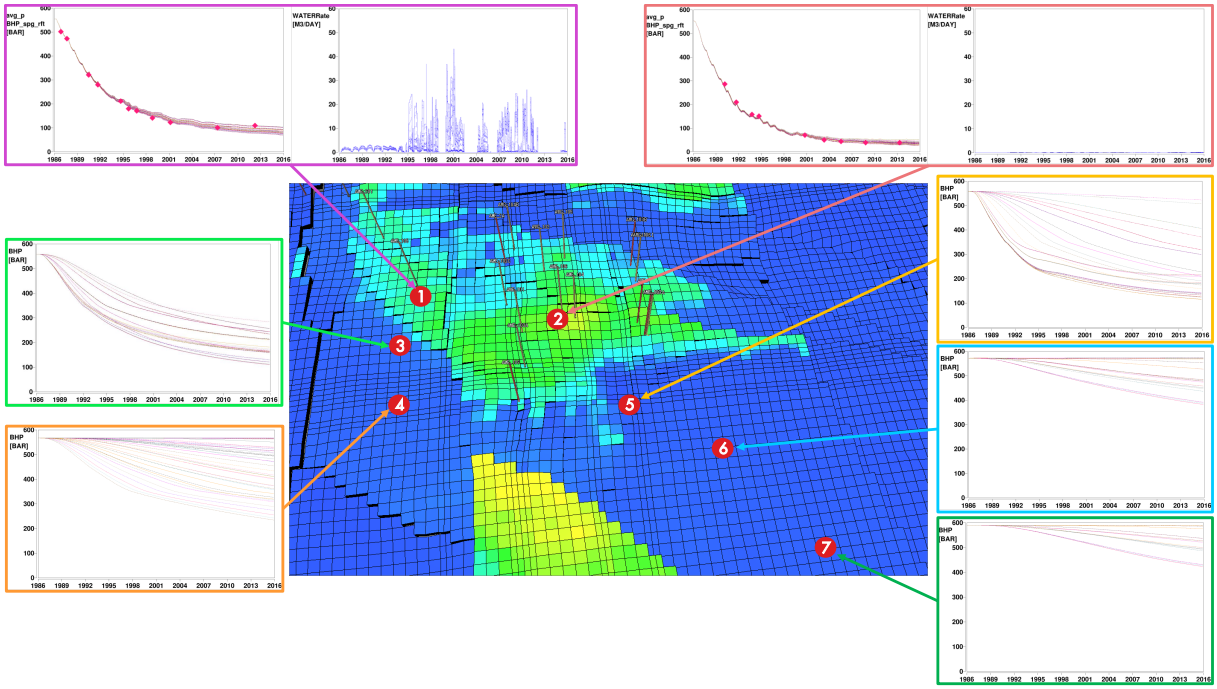


Figure 13. Summary of the simulation study. The plots in this figure are presented and explained in the previous sections.

4 Geomechanical model

4.1 Introduction

To calculate reservoir compaction and surface subsidence for the 58 history matched models (describing reservoir pressure depletion and depleting thickness) provided by the reservoir engineer, a geomechanical model is required. The geomechanical model has two components: a compaction model to calculate reservoir strain resulting from pressure depletion and an influence function to translate the compaction strain into a subsidence signal at the surface.

4.2 The compaction model

The number of publications on subsidence and compaction in the Netherlands has increased significantly over recent years (e.g. van Thienen-Visser et al. 2015, Fokker and van Thienen-Visser 2016, van der Wal and van Eijs, 2016). This is probably due to the observed link between reservoir compaction and induced seismicity in the Groningen field. The compaction models presented in these publications and used in NAM are summarized below:

Linear compaction model

The simplest geomechanical compaction model assumes compaction of a depleting reservoir layer to have a linear elastic relationship with depletion, depleting thickness and the uniaxial compressibility (c_m) of the reservoir rock. This is a good first order approximation and a suitable compaction model for short term (up to 5 years) predictions but fails to predict the temporal behaviour of the observed subsidence above the Ameland gas field.

Bi-linear

Geodetic observations above the gas fields in The Netherlands show an increase of the subsidence rate after the first years of production, which is a phenomenon observed globally in many oil- and gas fields (Hettema et al., 2002). The first phenomenological model used in NAM (till 2011) to match this observation was a bi-linear compaction model consisting of two branches with different C_m values, where a stiffer branch changes at a certain transition pressure to a softer branch. From geodetic observations above the Ameland field (NAM, 2011, 2015a) it became clear that the bi-linear model could not describe the ongoing subsidence observed above this field at the end of field life when the depletion rate had decreased significantly.

Time-decay

Both the delayed response of compaction at the start of production and at the end of production can be reasonably fitted by the NAM Time-decay model (Mossop, 2012). Although the precise cause of a volumetric time decay process is unknown, the time-decay model is the simplest model that reasonably matches the spatial-temporal developments of the subsidence above the Ameland field. The strain of the constrained volume, e_{ii}, at a point, $\mathbf{x}$, in the reservoir is then the instantaneous product of pressure change, Δp, and constrained uniaxial compressibility, c_m, but now convolved with a time decay function.

$$e_{ii}(\mathbf{x},t) = \Delta p(\mathbf{x},t)\; c_m(\mathbf{x}) \;*_t \frac{1}{\tau}\exp\left[\frac{-t}{\tau}\right]$$

In this formula, t, is time, $*_t$, is the convolution operator with respect to time and, τ, is a time-decay constant. The best fitting time decay constants for the Groningen field were found by inversion using

a semi-analytic geomechanical model and typically have values of 3 to 8 years. For the Ameland field, the numbers that were used to calibrate to the subsidence varied between 5 and 7 years (NAM, 2016b).
It is quite possible that the observed time-decay is not a material property of the reservoir rock, but could be due to particularities of the reservoir geometry, pore fluids or some other factor. Therefore, the time-decay constant for one reservoir is not necessarily appropriate for another reservoir simply based on rock type. Even though the time-decay model leads to reasonable fit to the observed subsidence above the Ameland field, it provides a worse fit to stress-strain curves for core plug tests where the loading has been varied during the test (Pruiksma et al., 2015).

Rate type compaction model
De Waal (1986) proposed a rate type compaction model where the compaction (strain) of a sandstone is dependent on the loading rate of the rock. This model originates from soil mechanics principles (e.g. Bjerrum, 1967) but is applied as well to describe the stress strain response of a sandstone plug in the laboratory. In the Netherlands, a form of this model is the most accepted model for settlement calculations in soft soil. The model is also known as the a,b,c isotachen model (den Haan, 2003). Pruiksma et al. (2015) described the application of this model to laboratory experiments on Rotliegend sandstone core material. Improvements to the original work of the De Waal led to the definition of the isotach (i) formulation of the rate type compaction model (RTCiM) which was also implemented by NAM (e.g. van der Wal and van Eijs, 2016).

The ideal compaction (+ influence model) should be able to describe both the lab data and the complex temporal subsidence behaviour observed above the Ameland field. The RTCiM model is used in this study, as it gives the best match to core experiments when compared to any of the other compaction models (see sources in Table 7). Another advantage of the RTCiM model is that with its versatility to span a wide range of temporal behaviours, it prevents the workflow from allocating all the observed temporal complexity to the influence function. The RTCiM can behave more linear with depletion or exhibit time decay and temporal characteristics. None of the other models is as versatile. This choice concurs as well with the findings of the LTS-I research.

Pruiksma et al. (2015) demonstrated that the linear, bi-linear and Time Decay compaction models can be mathematically derived from the more complex RTCiM model but this requires a zero value for one or more parameters. This observation conflicts with another objective of this study, which is the application of realistic parameter value ranges for the RTCiM in the confrontation workflow. Section 4.2.3 describes the realistic value ranges for the RTCiM model that don't incorporate zero values.
This study considers the realistic ranges of the RTCiM to be more important excluding therefore the possibility of testing the simpler compaction models.

The current study implements the RTCiM framework proposed by TNO (2017), with the addition of a linear branch to describe the effects of the significant overpressure observed in the Ameland field. This addition of the overpressure component is described in Section 4.2.2.

4.2.1 RTCiM implementation

At a change of the loading rate, a first direct strain response, ε_d, is recorded followed by a more gradual response referred to as the secular strain, ε_s. The total strain is defined as the sum of a direct part and a time dependent secular part:

$$\varepsilon = \varepsilon_d + \varepsilon_s$$

The rate type isotach compaction model implemented in this workflow is derived from an explicit Euler finite-difference scheme keeping a constant time step Δt. To calculate the compaction of one grid block (x, y), the applied numerical scheme can be divided into 5 steps (TNO, 2017):

1) From the current effective vertical stress $\sigma'(t)$ and strain $\varepsilon(t)$, calculate the secular strain rate as:

$$\dot{\varepsilon}_s(t) = \left(\frac{\varepsilon(t)-\varepsilon_0}{\sigma'(t)} - C_{m_d}\right) \dot{\sigma}'_{ref} \left(\frac{\varepsilon(t)-\varepsilon_0}{\sigma'(t)\,.C_{m_{ref}}}\right)^{-1/b}$$

The vertical effective stress is derived from the reservoir depth, z_r, and the mean density, ρ_{mean}, of the subsurface up to the reservoir top and the reservoir gas or fluid pressure, $P(t)$, as:

$$\sigma'(t) = (\rho_{mean}.g.z_r) - P(t)$$

At t_0 (the onset of pressure depletion/production), the direct elastic strain $\varepsilon_d(t_0)$ and secular or creep strain $\varepsilon_s(t_0)$ are both considered to be equal to zero, and thus total strain $\varepsilon(t_0)$ is set to zero. The reference total strain is expressed as:

$$\varepsilon_0 = -C_{m_{ref}}\,.\sigma'_{ref}$$

with the reference vertical effective stress $\sigma'_{ref} = (\rho_{mean}.g.z_r) - (\rho_{water}.g.z_r)$, i.e. the vertical effective stress for a hydrostatically pressured rock.
Three material parameters ($C_{m_{ref}}$, C_{m_d}, b) and one state parameter ($\dot{\sigma}'_{ref}$) are needed to compute the rate type compaction. $C_{m_{ref}}$ is the reference compaction coefficient corresponding to the pre-depletion loading rate, and thus by definition quite high. Parameter C_{m_d} is the direct compaction coefficient and dedicated to map out the direct effect at the change of loading rate, closely linked to the elastic response of the rock. In the scenario of the change of loading rate due to the onset of pressure depletion, C_{m_d} is expected to be low, to mimic the stiff (elastic) response of the reservoir rocks. The ranges and source for the values of these parameters are described in Section 4.2.3.

2) The second step of the Euler scheme calculates the increase in creep strain as:

$$\Delta\varepsilon_s = \dot{\varepsilon}_s(t)\,.\Delta t$$

and update the creep strain as:

$$\varepsilon_s(t+\Delta t) \rightarrow \varepsilon_s(t) + \Delta\varepsilon_s$$

3) Following a linear stress-strain relationship, the direct elastic strain is:

$$\varepsilon_d(t+\Delta t) = C_{m_d}.(\sigma'(t+\Delta t) - \sigma'_{ref})$$

4) Finally, the total cumulative strain is:

$$\varepsilon(t+\Delta t) = \varepsilon_s(t+\Delta t) + \varepsilon_d(t+\Delta t)$$

The total cumulative compaction is:

$$V_{comp}(t + \Delta t) = -\varepsilon(t + \Delta t) \cdot V(t_0)$$

with $V(t_0)$ the grid block net volume, assumed constant over time. Accounting for changes in grid block net volume would not significantly change the compaction. After this final calculation, the workflow returns to the first step for the next time step. It is important to note that the present rate type isotach compaction model is attempting to describe the delay and persistence in subsidence rates at the onset and arrest of production, by only considering the reservoir compaction and assuming a purely elastic linear response of the rocks surrounding the reservoir rocks. The creep of a possible visco-elastic salt layer on top of the reservoir might also contribute to the non-linearity in the subsidence, which will be addressed in Section 0.

4.2.2 Adding a linear branch to the RTCiM model

The Ameland gas field initial pore pressure was around 555 bar, implying an overpressure of around 165 bar relative to a hydrostatic pore pressure gradient of 0.11 bar/m. The overpressure is a result of a large, low density, gas column that developed after the deepest burial of the rock. This means that the Ameland field saw in geological history a phase of pore pressure increase and effective stress decrease. During burial, the reservoir rock compacted partly via an inelastic (visco-plastic) process and partly via an elastic process. During the gas fill phase the vertical effective stress was reduced with a partial rebound of the elastic strain only. A similar situation would occur as well in case of previous deeper burial and subsequent uplift, also called “over consolidation” (de Waal and Smits, 1985).
This elastic deformation path is also followed during the first stage of the depletion (as result of the gas production) until the depletion is equal to the initial overpressure. Once the overpressure has been relieved, inelastic deformation of the rock matrix starts to play a role in the compaction behaviour, which is captured by the RTCiM model. The rate type behaviour then takes over from the linear elastic compaction behaviour at the point when the pore pressure is below normal hydrostatic pressure.

While the pore pressure is above the hydrostatic pressure (i.e. while the grid block is overpressured), the strain response is expected to be elastic. Therefore, while $\sigma'(t) < \sigma'_{ref}$, the secular creep component ε_s is assumed zero such that the total strain developed in each grid block is given by $C_{m_{lin}}.(\sigma'(t + \Delta t) - \sigma'_{ref})$. The C_{m_d} in the RTCiM model also represents the elastic behaviour of the rock, therefore $C_{m_{lin}} = C_{m_d}$.

4.2.3 Parameter values for the RTCiM model

The a priori parameter ranges for the RTCiM model are based on the calibration of the RTCiM model to laboratory experiments. Table 7 presents a list of values deduced from various literature sources.

Table 7. Parameter values for the RTCiM model from various sources.

$C_{m_{ref}}$ [10^{-5} bar^{-1}]	C_{m_d} [10^{-5} bar^{-1}]	b [-]	*Source*
0.57 x f(por)	0.26 *$C_{m_{ref}}$- 0.44*$C_{m_{ref}}$	0.017-0.021	TNO (2013)
0.7 - 1.29	0.42 - 0.69	0.01 – 0.03	NAM (2016)
1-5		0.008-0.022	De Waal (1986)
0.46-2.6	0.26 *$C_{m_{ref}}$- 0.56 *$C_{m_{ref}}$	0.01 – 0.024	De Waal et al. (2015)

Rather than using a unique value for the $C_{m_{ref}}$ independent of any other parameter, the influence of the porosity on the $C_{m_{ref}}$ was investigated. The objective is to find possible relations between C_m and porosity to match to the core plug experiments performed on Dutch Rotliegendes sandstone (Figure 14). The C_m value in this figure shows a large scatter with the porosity where both linear and non-linear functions might fit the data. The posterior ensemble should not be impacted by an a priori subjective choice of a possible relation between the two parameters. A general exponential function has been defined that could describe different levels of complexity (linear to exponential):

$$C_{m_{ref}}(Por) = C_4 + C_2 10^{C_3(Por - C_1)}$$

The parameters of this function $C_1 - C_4$ are chosen such that the $C_{m_{ref}}$ is always increasing with an increasing value for the porosity. They are also chosen such that the resulting C_m-Porosity relation fall within the search (grey) area shown in Figure 14. All four parameters in this function interact on the C_m-Porosity relation. Because of this interaction, only parameters C_1 and C_2 have a fixed range with parameter values selected by the Monte Carlo process. With the selected C_1 and C_2 the parameter range for C_3 and C_4 are determined such that the C_m value at 7 % porosity is between 0.25 and 2.5 x 10^{-5} bar^{-1} and C_m value at 20% porosity is between 0.25 and 7 x 10^{-5} bar^{-1}. The parameters C_3 and $C_4$are drawn from these ranges. The resulting C_m space is shown as the grey area in Figure 14 where the space is defined by lines with slopes having a starting value for zero porosity and a description of the curvature.

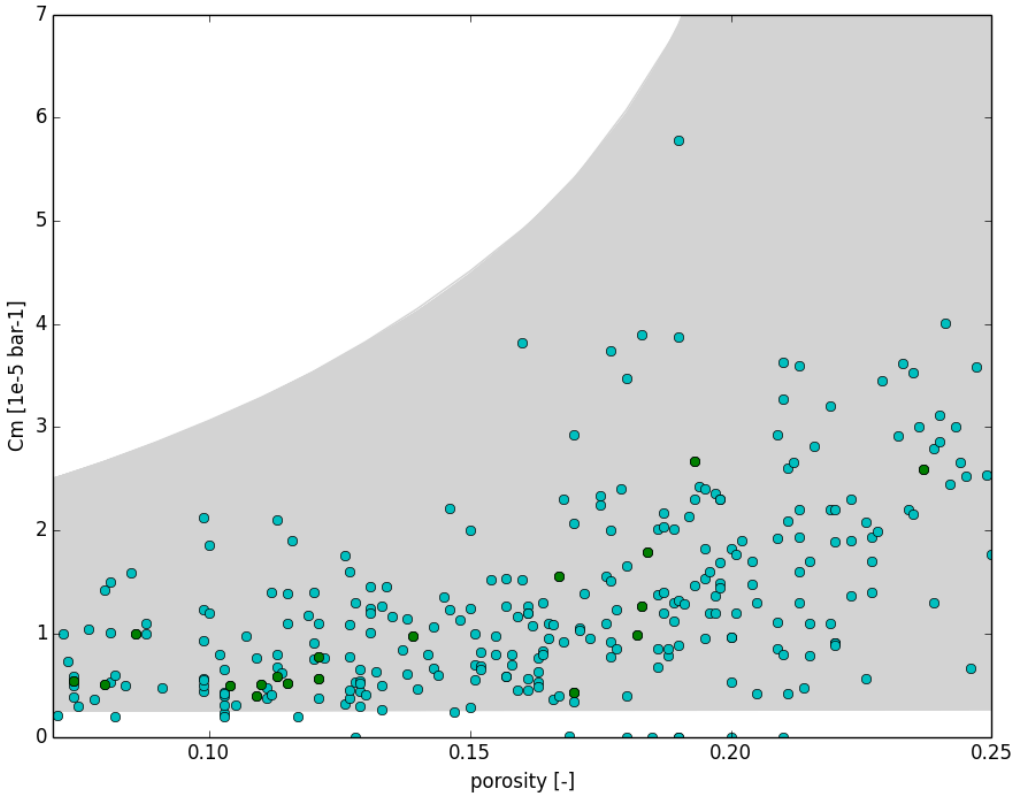


Figure 14. C_m-Porosity values from core plug experiments. The cyan dots represent all North-Netherlands core plug experiments, the green dots the core plug experiments of Ameland. The grey area is the *a priori* range for the C_m-Porosity relations

Figure 14 shows that the grey area spans a larger area, towards higher values, than given by the results for the core plug experiments. The reason for it is two-fold, as the average and effective C_m will be different after upscaling (will be explained in Section 4.5), and the compressibility in the field can be different from that observed in the lab due to different loading rates.

When comparing properties used for the RTCiM model, especially the $C_{m_{ref}}$ directly with those obtained from core measurements, the applied upscaling process should be kept in mind. For the compaction calculation, the compressibility is determined based on the upscaled porosity, and the compressibility is not upscaled itself since a prior compressibility grid does not exist. If the relation between porosity and compressibility is non-linear, as seems to be the case here, then the compressibility determined from the upscaled porosity (= model compressibility) is different from the average compressibility of the area (= effective compressibility).

The effect of the porosity upscaling on the resulting C_m values can be imagined as follows. From core data, it appears there is a possible exponential relation between porosity and compressibility, $Cm = f(\phi)$. Consider the area that is to be upscaled has some porosity distribution, with average (= upscaled) porosity ϕ_{av}, and some compressibility distribution [linked to the porosity via $Cm =$

$f(\phi)$] with average or effective compressibility Cm_{eff}. Since the compressibility itself is not upscaled, the value of the compressibility as used by the model will be based on the upscaled (or average) porosity as $Cm_{model} = f(\phi_{av})$. For an exponentially increasing function, $f(\phi)$, the effective compressibility, Cm_{eff}, will be higher than the determined model compressibility, Cm_{model}, derived from the upscaled porosity. For example, assume the area to have a Gaussian porosity distribution, then the compressibility for this area will follow a log-normal distribution. The average, or effective, compressibility, Cm_{eff}, of this area will be the average of the log-normal distribution, which will in general be *higher* than the Cm_{model} derived from the average porosity: the average of a log-normal distribution (Cm_{eff}) is generally higher than its mode (Cm_{model}). Therefore, for an upscaled model to obtain the same effective compaction as the non-upscaled model, it should apply a slightly higher porosity – compressibility relation. Hence, the range should be chosen slightly wider than the ranges from core measurements for the models to use the same effective compressibility. The discrepancy between the two becomes larger for wider porosity distributions. Note that it also holds for other porosity distributions, like uniform distribution.

In addition to the effect resulting from the upscaling of the grids in size, there can also be a difference in the compressibility derived from lab data and observed in the field, that originates from differences in applied loading rates (de Waal, 1986). C_m-Porosity relations on the field time scales are different from those on lab time scales.

Therefore, the C_m values that appear in the accepted members, after the confrontation with the data, cannot be compared directly to the core plug experiments.

The *a priori* parameter value ranges used in this study for both the C_m-Porosity relation and the RTCiM compaction model are provided in Table 8.

Table 8. Selected parameter ranges for the RTCiM compaction model.

Parameter	range
C_1	0.02 - 0.17
C_2	0.01 - 2.5
C_3/C_4	*
$C_{m_{ref}}$	= $C_4 + C_2 10^{C_3(Por-C_1)}$
$C_{m_{lin}}$	= C_{m_d}
$C_{m_d}/C_{m_{ref}}$	0.2- 0.7
b	0.01 - 0.03

* Parameters C_3 and C_4 are derived based on the values of C_1 and C_2 such that the used Cm-Porosity relations cover the full range of observations from core plug experiments (grey area of Figure 14).

4.3 The influence model

The influence model describes the translation of the reservoir compaction into surface deformation. Next to the correct physical conditions of the model, a second requirement of the confrontation workflow is that the expressions in the influence model can be calculated fast. The well-known and widely applied theory published by Geertsma (1973) fulfils these requirements, building on the physical principles of elasticity theory.
However, the bowl shape resulting from this model is in many applications too wide when compared with geodetic data in the Netherlands (Geertsma & van Opstal, 1973). In their paper, they proposed to add an infinitely stiff reflective boundary or rigid basement (RGB) to the solution, which sits below the reservoir. This additional RGB is partly justified by stiffness data. In general, it is known that the stiffness increases with depth but not with a sudden step to an infinite stiff layer. Still the solution provides a reasonable fit to the measured subsidence bowls in the Netherlands (NAM, 2015).
Note that the effect of a *stiff* layer *below* the reservoir on the shape of the subsidence bowl is similar to the effect of a *soft* layer *above* the reservoir and therefore a "soft" response of the salt layer above the reservoir could act as an additional (temporal) component to the steepening of the subsidence bowl, postulated in the LTS-I study (Marketos et al. ,2015, 2016).

The account for the viscous and therefore temporal effect, the influence function for the Ameland case requires an additional component. TNO (2011) introduced the idea to move the RGB with time to mimic the behaviour of a viscous salt layer that is present in the overburden. Their solution not only approaches results from far more complex finite element (FE) models, it also allows for a very fast routine that is very suitable for the Monte Carlo approach used here. The formula that is implemented in this study is the same as the TNO equation:

$$c/k(t) = c/k(0) + d(c/k) * \left(1 - e^{\frac{-(t-t_0)}{\tau_{salt}}}\right)$$

Where c is the depth of the reservoir, $k(0)$ the depth of the RGB, τ_{salt} , the relaxation time of the field specific salt/overburden combination, $c/k(0)$ the starting position of the rigid basement and $d(c/k)$ the translation of the basement.

Unlike the salt viscosity, the relaxation time of the salt in this equation is not a physical property of the salt itself. It is a measure of the full response of an overburden where a salt layer is embedded between other elastic layers. Therefore, the value of this relaxation time can only be derived from another model, e.g. a finite element model. In this study, the Ameland FE model is used as a reference. This model is described in NAM (2011) and can be regarded as the state-of-the-art geomechanical model for the Ameland field to provide the subsidence volume contribution caused by the Ameland gas production in the annual measurement and control process as part of the "Hand aan de Kraan" procedure. The model contains the structural grids that represent the 3D geology of the Ameland area. The requirement of an additional numerical model complicates the application of the confrontation workflow on other fields. In a possible application for a new field situation, a simple FEM should be built to provide the input to new calibrations. This FE model requires, in that case, the input of viscosity values for the salt. The basis for the value range that was used is explained in the next Section.

4.3.1 Value range for the salt viscosity

Finding the right experimental data to calibrate the salt parameters in the models used in this study is difficult. The level of applied shear stress (S1-S3) in laboratory experiments on salt should ideally be close to the expected shear stresses caused by reservoir compaction in the field. From the Geomec model it can be observed that the induced shear stresses in the salt show values up to 1 MPa. In most of the laboratory experiments performed on natural salts, the applied shear stresses are much larger, thereby involving other creep mechanisms than only linear pressure solution creep, which is the dominant mechanism in this pressure-temperature domain (Marketos et al., 2015). Experimental data at conditions close to the stress conditions that would apply to our case are presented by IFG (2006) and Berest et al. (2012). IFG (2006) provides experimental data at ambient temperatures on the Zechstein salt that is mined close to the Ameland field (Harlingen, Frisia Zout B.V.). Although, the applied level of shear stress in the IFG (2006) experiments is very low, it is still higher than required for direct use in the Ameland models. The results from Berest et al. (2012) were performed at a differential stress of 1 bar (0.1 MPa), which is at the low side of the possible induced shear stress range of the model, keeping in mind that the experiments were conducted at a temperature of only 14.4 °C. Via the equations shown in Section 4.3.2, the viscosity can be recalculated to a value at a temperature of 105 °C, similar to the temperature of the other experiments but more importantly, similar to the ambient average temperature of the salt above the Ameland field.
Figure 15 shows the strong non-linear dependency of the viscosity with the differential stress. The extrapolation of the IFG data to lower stress values gives a good match to our temperature corrected data of Berest et al. (2012). This result would imply that the possible range for the salt viscosity value is rather well constrained to a range of around $5 \cdot 10^{15}$ to $5 \cdot 10^{16}$ Pa·s at the stress levels valid for the Ameland case. Section 4.3.2 demonstrates that this viscosity range results in very fast relaxation times (half times) for the salt, i.e. between 1 month and 1 year.

However, indirect evidence for viscosity ranges at low shear stress might come as well from the process of salt diapirism (e.g. Keken et al., 1993). The speed of this process can be inferred from geological observation and subsequently be used for the calibration of the salt viscosity value in geomechanical models. This was exactly the objective of the study by Keken et al. (1993), which concluded a range of viscosity values between 10^{17} and 10^{20} Pa·s at strain rates between 10^{-12} and 10^{-15} s^{-1}, i.e. significantly higher values for the viscosity than obtained by the above described experiments. Combining all these observations leads to the conclusion that the uncertainty of the viscosity value is large, up to almost 5 orders of magnitude, and that a possible *a priori* range should be influential at all time scales during and after the gas production process.

The next Sections describe how a 'pulse model' translates viscosity values into a relaxation time for the RGB expression, τ_{salt}. This work demonstrates that the range between $5 \cdot 10^{15}$ and $5 \cdot 10^{18}$ Pa·s addresses the earlier conclusion on the *a priori* range.

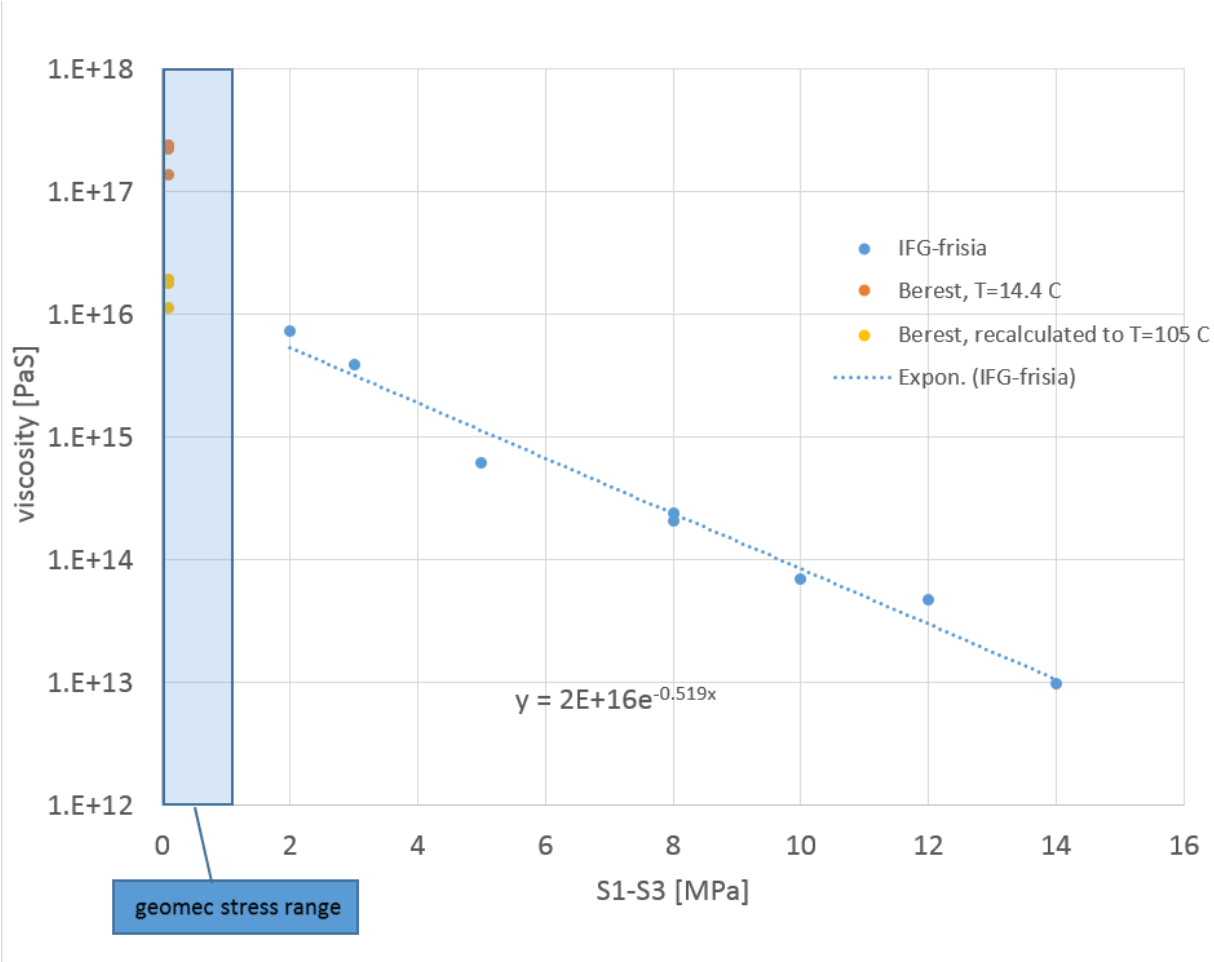


Figure 15. Viscosity versus differential stress for experiments; result found in literature (Berest) or provided by salt mining operators (IFG-Frisia).

4.3.2 Results of the 'pulse-model' for different viscosities

The obtained range for the viscosity led to the definition of 5 FE model scenarios where the viscosity value is varied in the salt layer above the reservoir. In these scenarios, the total depletion of the field was applied as a load case in just one month (for the pulse model, see also Section 4.4.2). The objective of using a "pulse-model", is to compare the response from a certain salt viscosity in the FE model to a clear event in time, and match its behaviour with a relaxation time for the moving-rigid-basement model, without obscuring the signal by adding other temporal components like pressure changes.

The main geomechanical layers and values for the elastic parameters of the FE model are listed in Table 9 (NAM, 2011). The stiffness or Young's modulus (E) of the reservoir layer is lower when compared to the overburden layer. This is mainly caused by the assumption that these low porous rocks outside the reservoir would respond more to the undrained modulus, i.e. reflecting a stiffer response of the rock because of the effect that fluids in the rock cannot escape during small stress changes. The relative higher values of the dynamic moduli, measured by a sonic tool, are representative for this undrained response and used in the model. Obviously, the porous reservoir response or compaction is better reflected by the drained moduli.

Table 9. Average values of the finite element model.

Layer name	Average depth of top of layer [m]	Mean E [GPa]	Mean ν	Viscosity Pa·s	Cm [10^{-5}bar^{-1}]
North Sea	0	2	0.3	-	3.7
Chalk	1000	10	0.25	-	0.83
Cretaceous/Jurassic	1500	16	0.25	-	0.52
Zechstein	2000	30	0.35	variable	0.21
Ten Boer Claystone	3200	40	0.2	-	0.23
Rotliegendes Sandstone (reservoir)	3300	8	0.2	-	1.13
Limburg	3410	40	0.2	-	0.23

Creep parameter values in Geomec (the FE software used) were translated into viscosity values as defined by a Maxwell model (pers. Comm. P. Fokker, TNO and P. Fokker, Shell) following the scheme of equations below.

In Geomec vertical strain rate is defined by:

$$\dot{\varepsilon}_{33} = D(\sigma_{33} - \sigma_{11})^n$$

with $D = A \cdot exp\left(-\frac{Q}{RT}\right)$

Here A the "creep factor" [MPa^{-1}s^{-1}], Q the activation energy [J mol^{-1}], R the gas constant [J mol^{-1}K^{-1}] and T the temperature [K].

With n being 1 for linear viscosity and for the condition that radial strain (ε_{11} and ε_{22}) is zero and $\sigma_{11} = \sigma_{22}$.
Maxwell viscosity is defined by:

$2\mu\dot{\varepsilon}_{ij} = \frac{\mu}{\eta}\left(\sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij}\right)$ and for $ij = 33$

$\dot{\varepsilon}_{33} = \frac{1}{2\eta}\left(\sigma_{33} - \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33})\right)$ and $\sigma_{11} = \sigma_{22}$,

Therefore $D = \frac{1}{3\eta}$

The response of a "pulse model" scenario is visualised via a cross section through the deformation bowl in Figure 16. The red line (least subsidence) in this graph shows the linear elastic response of the model, while the curves below this red line are a result of the viscous flow of the salt. This figure shows the definition of a model halftime and a maximum extra subsidence caused by the salt flow, indicated by the blue arrows in the figure. The maximum extra subsidence is equal for all scenarios, whereas the model halftime depends on the chosen viscosity.

Table 10 shows the scenario halftime values for 5 different viscosities. These results show that all viscosity values in the range would impact the spatial-temporal shape of the subsidence bowl during the lifetime or production time of the field.

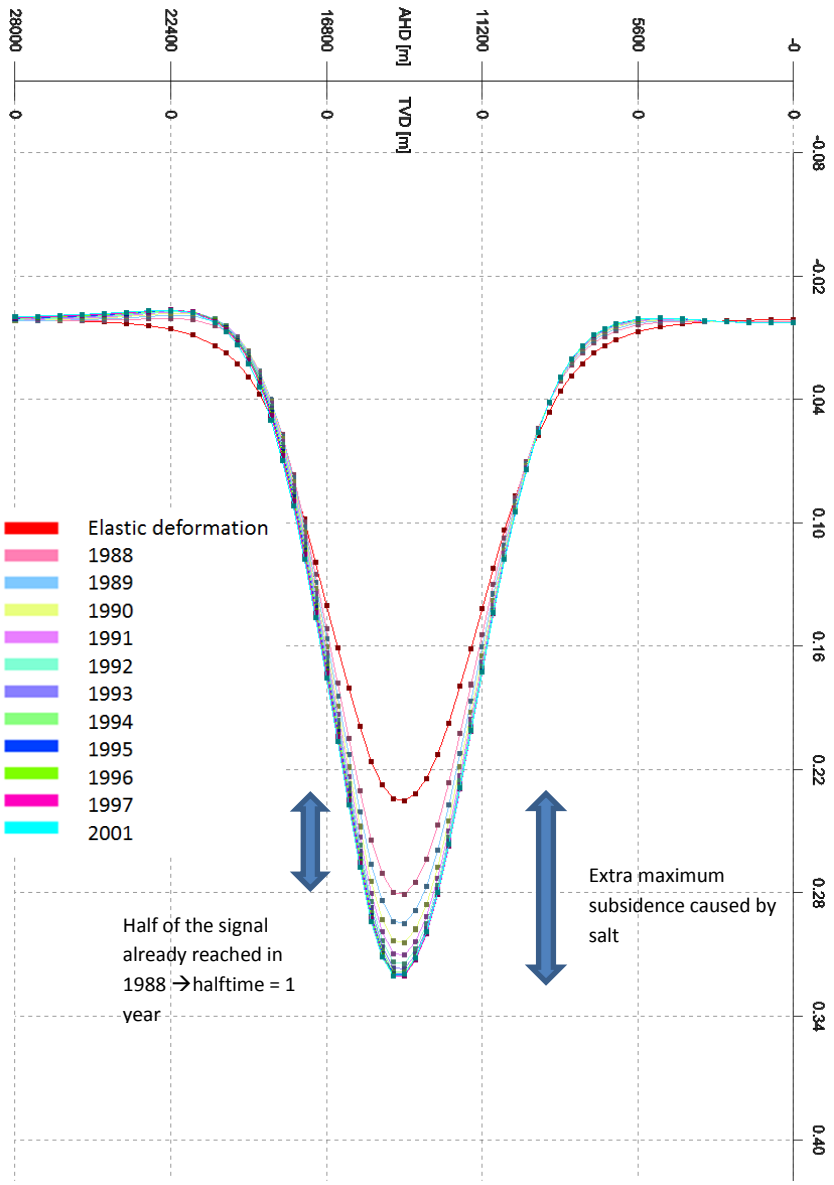


Figure 16. Result of the Geomec pulse model for a viscosity of $1 \cdot 10^{17}$ Pa·s. The use of the term 'halftime' in Table 10 is also explained in this figure. Within this model the full field depletion has been taken place in the first month. The elastic deformation is the direct response to this pressure pulse. The additional deformation shown for the later years is caused by the effect of salt only.

Table 10. Results for the signal half time because of viscosity variation.

Model name	A factor in Geomec ($MPa^{-1}S^{-1}$)	Maxwell Viscosity (Pa·s)	Signal half time estimated (year)
Ame_Pulse_ZE-A2.26E-07_TD7_C1_I2	$2.26\ 10^{-7}$	$5.0\ 10^{15}$	0.1
Ame_Pulse_ZE-A1.13E-08_TD7_C1_I2	$1.13\ 10^{-8}$	$1.0\ 10^{17}$	1.5
Ame_Pulse_ZE-A1.69E-09_TD7_C1_I2	$1.69\ 10^{-9}$	$6.7\ 10^{17}$	10
Ame_Pulse_ZE-A1.13E-09_ TD7_C1_I2	$1.13\ 10^{-9}$	$1.0\ 10^{18}$	15
Ame_Pulse_ZE-A2.26E-10_TD7_C1_I2	$2.26\ 10^{-10}$	$5.0\ 10^{18}$	70

The values for the "moving RGB" were derived from matching the analytical model with the output of the Geomec scenarios. The compaction model in both Geomec and Geertsma & van Opstal models is a linear model and therefore the viscosity of the salt is the only time dependent process in these scenarios. Several scenarios, uniquely defined by different values for the salt viscosity, were matched with the analytical model and results are provided in Table 11. The parameters were derived from different parts of the maximum subsidence vs time curve (Figure 17). $c/k(0)$ was determined by the initial subsidence, $d(c/k)$ by the total effect of the viscous part, and τ_{salt} from the duration in which the creep took place.

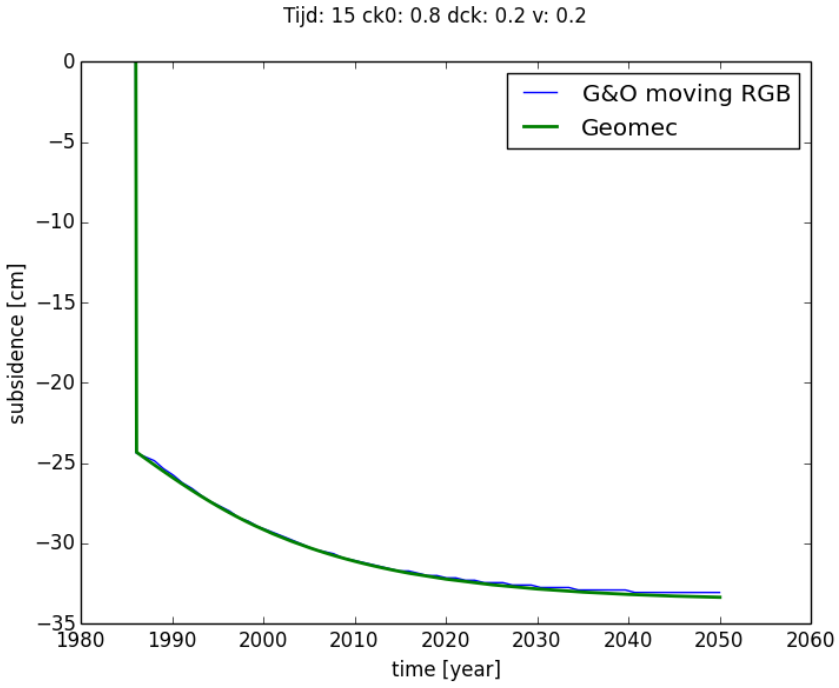


Figure 17. Subsidence versus time curve for a salt viscosity of 6.7 10^{17} Pa·s. The elastic deformation is the direct response to this pressure pulse. The additional deformation shown for the later years is caused by the effect of salt only.

The matching exercise resulted in fixed parameters for the $c/k(0)$ and $d(c/k)$. These fixed values are simply a result of this comparison using a single location in both models. These values may

change with salt thickness and thickness of the ROCLT, the claystone layer between the salt and the reservoir. To investigate this possible effect, the same matching exercise was performed at different locations of the models but yielded small differences for each parameter as demonstrated by Figure 18. For a fixed set of calibrated values, it is shown that the results of the Geomec model versus the Geertsma and van Opstal solution with the moving rigid basement are not impacted by thickness variation of the salt layer and/or ROCLT. More evidence for the small impact of a heterogenous subsurface on the shape of the bowl is provided in the next Section.

Table 11 provides an overview of values that were obtained during this calibration process and shows that the τ_{salt} values are roughly 50% higher than the half time values that are listed in Table 10. From the literature, a large *a priori* range for the τ_{salt} was deduced, varying between 0.3 and 100 years. The other two parameters in the equation, $c/k(0)$ and $d(c/k)$ show less variation in the obtained values. Narrow ranges are defined around these values (respectively 0.75 - 0.85 and 0.15 - 0.25).

Table 11. Match of the analytical model to the Geomec output for different salt viscosities.

Geomec creep factor (s^{-1})	viscosity [Pa·s]	ν	τ_{salt} [year]	$c/k(0)$	$d(c/k)$
$2.26\ 10^{-7}$	$5\ 10^{15}$	0.2	0.3	0.87	0.15
1.1310^{-8}	$1\ 10^{17}$	0.2	2	0.8	0.2
$1.69\ 10^{-9}$	$6.7\ 10^{17}$	0.2	15	0.8	0.2
$2.26\ 10^{-10}$	$5\ 10^{18}$	0.2	100	0.8	0.2

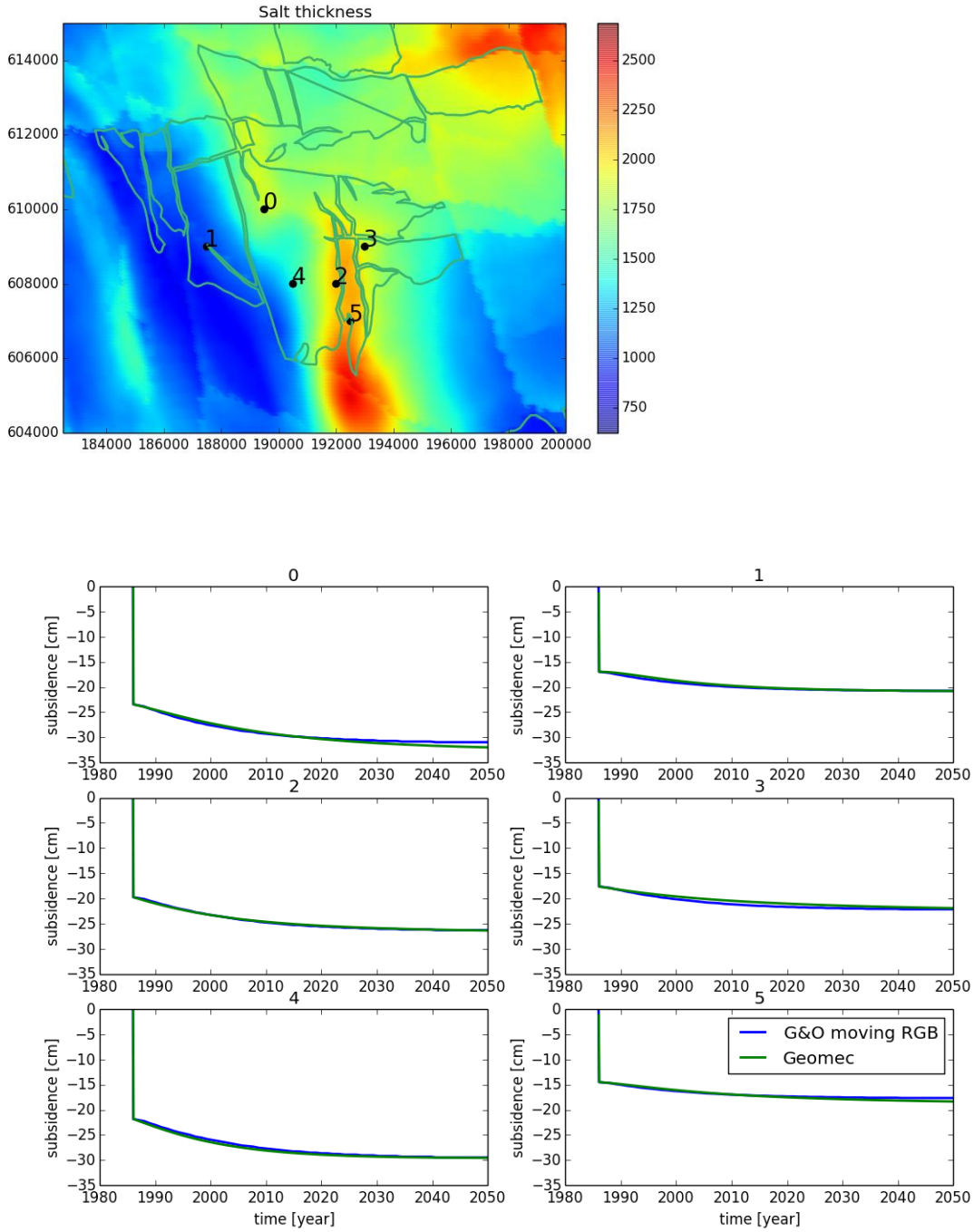


Figure 18. Comparison of Geomec and Geertsma and van Opstal moving RGB solution. The subsidence with time is presented at different locations having a different thickness (metres) of the salt package (see top figure).

4.4 The possible role of the overburden on the position of the subsidence bowl

The previous Sections demonstrate that a complex 3D model can be represented by a simplified analytical solution. However, this solution does not account for the possible influence of a 3D varying salt layer. A model with a realistic 3D overburden with different values for the geomechanical behaviour of this overburden might position the subsidence bowl differently than a simple "layer cake" model of the subsurface. To verify the existence of such a 3D effect, several tests were performed. First an elastic (time independent) overburden is examined in Section 4.4.1 and then possible time dependent effects are examined in Section 4.4.2.

4.4.1 Possible time-independent effects

In the 3D model, the largest difference in thickness exists in the salt layer. This layer is therefore used to test the hypothesis that a different value for the Young's modulus would lead to a different (time-independent) position of the bowl. To test the hypothesis, two scenarios with different elastic properties for the salt were created.

Scenario A uses a Young's modulus value of 3.5 GPa while Scenario B uses a value of 35 GPa. Note that these values span a range larger than what is realistic for the Zechstein salt elastic values. Both scenarios have been tested in Geomec by a "pulse model" as described before, i.e. the total depletion applied in one "time" step. lists the position and value of the maximum modelled subsidence for both scenarios. Although the difference in the value of the Young's modulus of the salt causes a difference in the amount of the subsidence, the position of maximum is the same. The accuracy of this assessment is however defined by the grid size of the Geomec model, which is in our case 500 m.

Table 12 lists the position and value of the maximum modelled subsidence for both scenarios. Although the difference in the value of the Young's modulus of the salt causes a difference in the amount of the subsidence, the position of maximum is the same. The accuracy of this assessment is however defined by the grid size of the Geomec model, which is in our case 500 m.

Table 12. Position and value of the maximum subsidence for variation of Young's modulus.

	Esalt 3.5 GPa			Esalt 35 GPa		
year	x	y	Value [cm]	x	y	Value [cm]
1987	190563	609163	26.5	190563	609163	24.7

A possible influence of Poisson's ratio has been investigated for a model with a fixed value of the Young's modulus of 35 GPa. Table 13 shows that the position of the deepest point is not influenced by Poisson's ratio.

Table 13. Position and value of the maximum subsidence for variation of Poisson's ratio.

	Esalt 35 GPa, ν 0.35			Esalt 35 GPa, ν 0.1		
year	x	y	Value [cm]	x	y	Value [cm]
1987	190563	609163	24.7	190563	609163	25.0

The results of these tests show that the combination of a strong heterogeneous and spatially variable thickness of the overburden doesn't affect the position of the deepest point of the subsidence bowl. This is further illustrated in Figure 19. The shape of the contour lines is very similar for both scenarios.

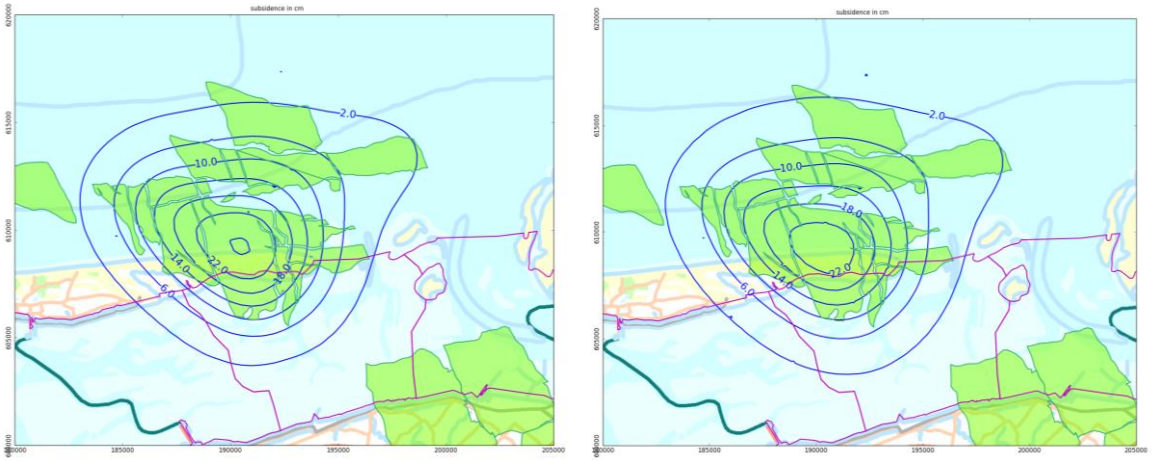


Figure 19. Left: total subsidence after full depletion for a salt layer having a Young's modulus of 3.5 GPa. Right: for a salt layer having a Young's modulus of 35 GPa.

4.4.2 Possible time dependent effects of viscous salt on the position of the deepest point

The pulse model is used again to rule out any *time dependent* effects of a combination of a complex 3D structure and the viscous behaviour of salt on the position of the subsidence bowl. Using various values for the viscosity, the impact of this depletion "pulse" on the salt above the reservoir is analysed. The hypothesis to test is that the position of the subsidence bowl is not influenced by the 3D structure of the salt. This would provide confidence that incorporation of a simple overburden in the analytical model would not compromise the results.

Table 14. Position of the bowl for two salt viscosity values.

	Viscosity 5 10^{18} Pa·s			Viscosity 5 10^{15} Pa·s		
year	x	y	Value [cm]	x	y	Value [cm]
Feb 1986	190563	609163	24.7	190563	609163	27.3
1990	190563	609163	25.0	190563	609163	33.6
2001	190563	609163	25.7	190563	609163	33.6
2010	190563	609163	26.2	190563	609163	33.6
2020	190563	609163	26.7	190563	609163	33.6
2030	190563	609163	27.3	190563	609163	33.6
2040	190563	609163	28.2	190563	609163	33.6

Table 14 demonstrates that the location of the deepest point of the subsidence bowl is not affected by the 3D thickness variation of the salt layer, both for a high viscous and a low viscous salt. Figure 20 shows the subsidence contours after 10 years (left) and after 30 years (right) for a pulse model with a salt viscosity of 6.7 10^{17} Pa·s. The shape of the contour lines is very similar leading to the conclusion that there is no significant time dependent spatial influence of the salt layer on the shape of the subsidence bowl.

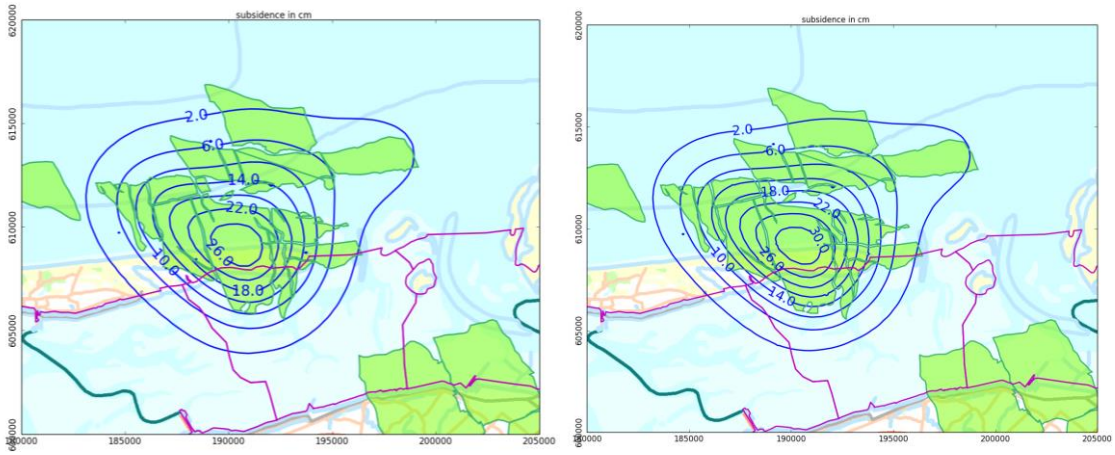


Figure 20. Subsidence contours for a "pulse" test model with a salt viscosity of 6.7E+17 Pa·s. Left, subsidence after 10 years, right: subsidence after 30 years.

4.4.3 Earlier published LTS-I results by Marketos et al. (2015)

Figure 36 of the Marketos et al. (2015) report (Figure 21, this report) shows the results of the deformation for different time-steps via a contour map and 3 cross sections through the subsidence bowl. Figure 21 c seems to suggest a shift in the location of the bowl, which possibly contradicts the observations described in the previous Section.

The FEM code, used by Marketos et al. (2015) is developed by the University of Utrecht (UU) and they were asked (pers. com. University of Utrecht) to replicate a simulation as described in Section 4.4.2 to verify the results from NAM with the "pulse model". Again, like NAM, the UU simulated the total Ameland depletion as one load case at t=0, where after the salt relaxes the imposed shear stresses over time. Contour plots of the normalised subsidence over time are presented in Figure 22. These plots indicate that there is no significant change in shape of the subsidence contours. Also, the location of the centre of the subsidence bowl does not shift over time. These results confirm the obtained results by NAM.

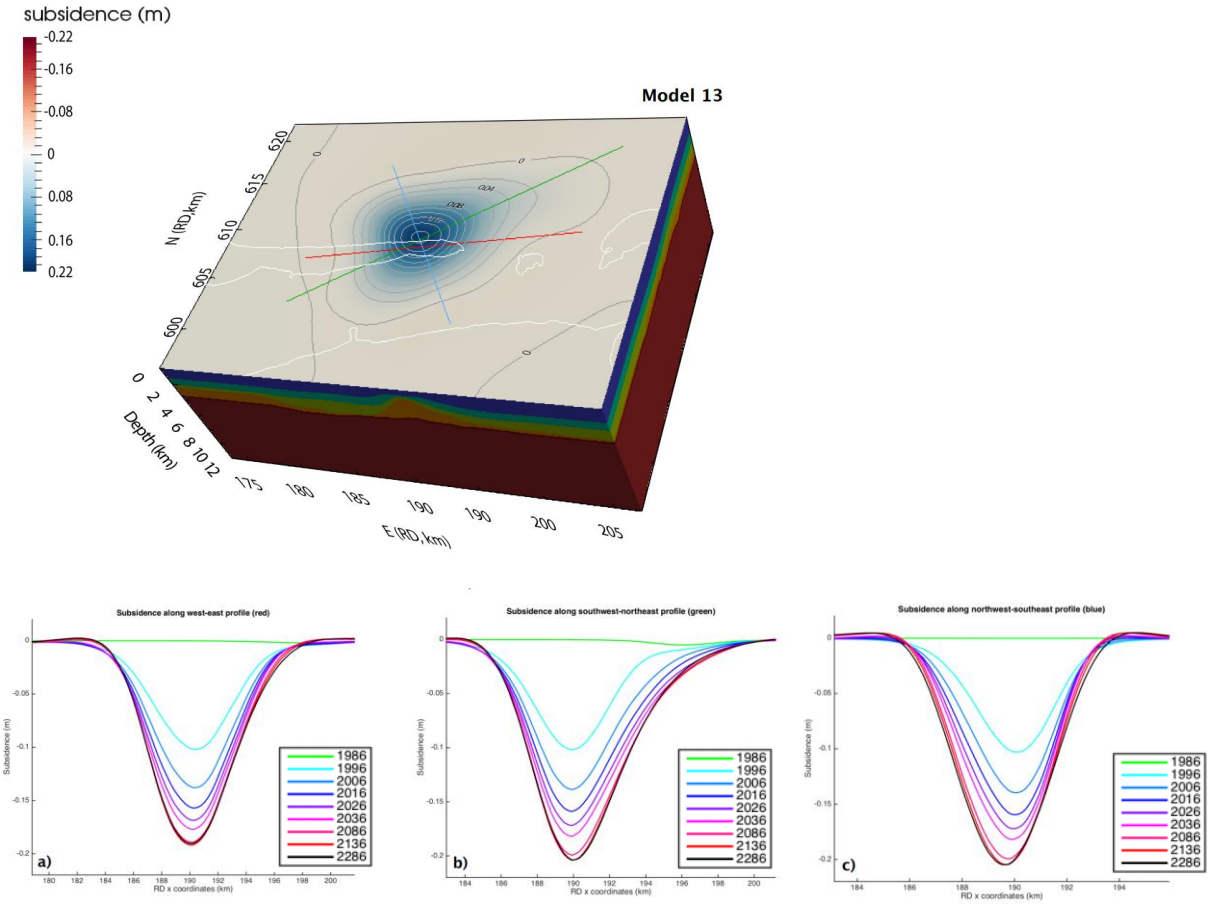


Figure 21. Deformation in time after 150 years of evolution (figure above). The subplots a, b, and c show Cross sections through the model showing vertical deformations (figure 36 from Marketos et al. 2015).

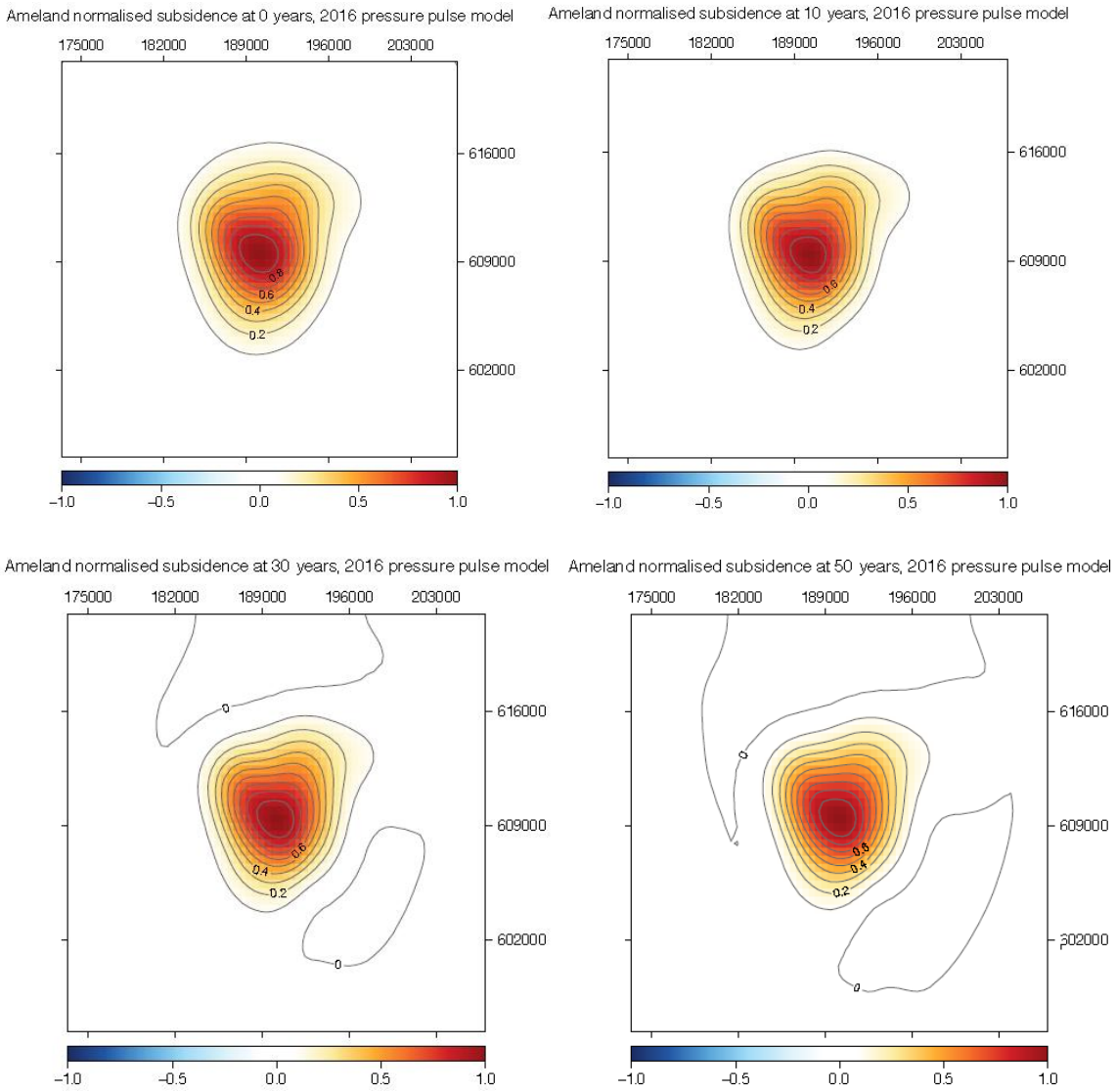


Figure 22. Normalised subsidence at resp. 0, 10, 30 and 50 years from UU FEM model for the "pulse model".

4.4.4 Conclusions on possible time-dependent and time-independent effects from a spatial varying overburden.

The tests described here and the further analysis by the University of Utrecht suggest that the (time) dependent influence of the spatial thickness variation in the overburden layers above the Ameland field on the location of the subsidence bowl is small, over time periods up to 50 years. This strengthens the conclusion that the influence function as used in the analytical model is a reasonable approximation of the complex model to translate the reservoir strain to the surface. To support the above statement, Figure 23 shows the results of a comparison between the analytical and numerical model for a pulse model. The shape of the contours of both models are very similar. The small differences in the magnitude of the subsidence that can be observed in these figures are most likely caused by the stiff elastic layers surrounding the reservoir in the much more detailed 3D finite element Geomec model.

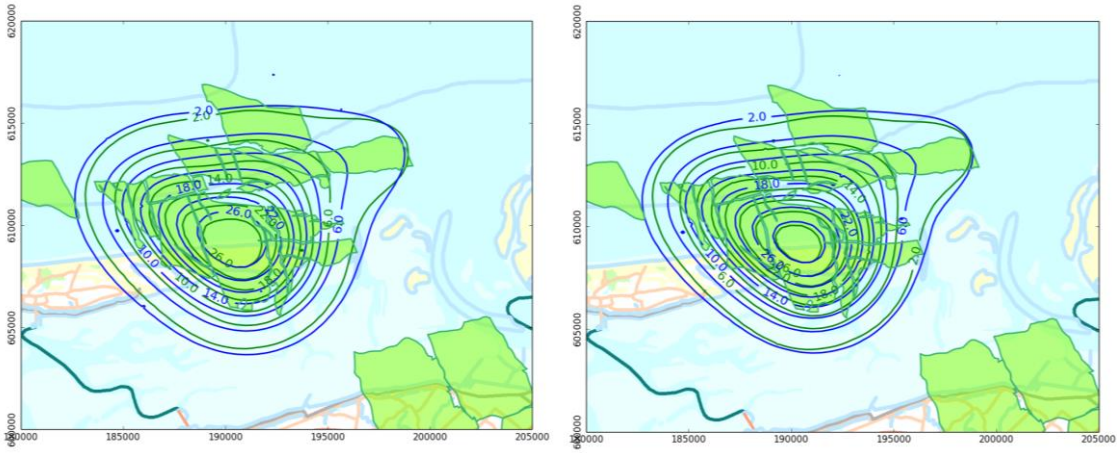


Figure 23. Comparing the analytical model (blue contour lines) with the complex FE model (green contour lines). Left figure: results of the pulse model for 1999, right figure: 2016.

4.5 Upscaling

Ideally the compaction calculations are run on the same grid size as used in the dynamic reservoir simulator (MoReS). However, the grid used in the MoReS dynamic reservoir simulator is very fine and contains too many grid cells to be used for computationally efficient compaction calculations. An upscaling step is introduced to generate a regular, one layer grid with horizontal dimensions of 500m x 500m as input for the compaction models. For each grid block, the location, grid block volume, porosity and the pressure over time is specified.

Upscaling is common practice in the building of subsurface models. When building a geological model, a geologist "upscales" very detailed measurements and observations from cores, into a fine scale geological model. When building the dynamic reservoir simulator in MoReS, this very fine scale geological model with millions of grid cells is "upscaled" into a coarser grid of a few hundred thousand grid cells to ensure that the computation time does not exceed 12-24 hours. The resulting input for the geomechanical modelling work flow is a grid with pressures that change over time. To enable the computation of many Monte Carlo simulations (in the case of this study 1.2 mln models were simulated) a further "upscaling" had to take place to convert the MoReS grid into a single layer of grid cells of 500m x 500m with changing pressure over time.

The upscaling process from MoReS to the geomechanical model is kept simple by using the (arithmetic) upscaling equations below. This set of equations preserves bulk volume [$V_{tot} = \sum_{k=1}^{n} V_k$], pore volume [$V_{pore} = \emptyset_{av} \cdot V_{tot} = \sum_{k=1}^{n} \emptyset_k . V_k$] and [$P_{av} \cdot V_{pore} = \sum_{k=1}^{n} P_k . V_{pore,k}$] with n the total number of fine scaled grid blocks.

The vertical upscaling is performed first, followed by a horizontal up-scaling over a coarser grid.

The *vertically* averaged pressures P_{av}, volumes V_{tot}, porosities $\emptyset_{av}$ and grid block centre positions X_{av} and Y_{av} for one vertical column of the 3-D grid can be expressed as:

$$P_{av} = \frac{\sum_{k=1}^{n} \emptyset_k . P_k . V_k}{\emptyset_{av} . V_{tot}}$$

$$V_{tot} = \sum_{k=1}^{n} V_k$$

$$\emptyset_{av} = \sum_{k=1}^{n} \emptyset_k . \frac{V_k}{V_{tot}}$$

$$X_{av} = \sum_{k=1}^{n} X_k . \frac{V_k}{V_{tot}}$$

$$Y_{av} = \sum_{k=1}^{n} Y_k . \frac{V_k}{V_{tot}}$$

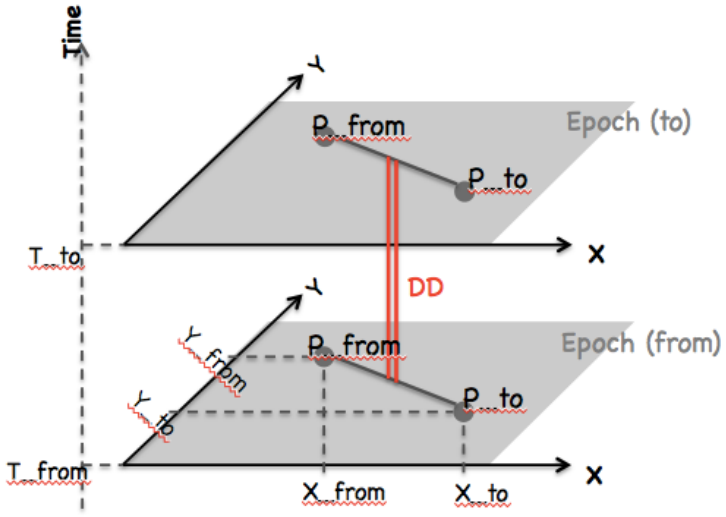
The first step of vertical upscaling yields a 2D irregular fine grid for P_{av}, $\emptyset_{av}$, V_{tot} at upscaled positions (X_{av}, Y_{av}).
The next step involves the horizontal upscaling of the vertically up-scaled grid. Basically, the horizontal upscaling follows the same equations as used for the vertical up-scaling. The end product is a 2D coarser regular grid.

5 Geodetic data

Levelling and GPS (Global Positioning System) data have been prepared for subsidence modelling following the new approach proposed in LTS-I (see next section). Data processing has been carried out by NAM with substantial support by Delft University of Technology (Van Leijen et al., 2017).

5.1 Innovations as recommended in LTS-I (2013-2015)

1. The observations from different survey techniques are not combined prior to subsidence modelling in a single dataset. Handling separate datasets for levelling and GPS avoids assumption-based interpolation that would be required for their alignment.
2. Spatio-temporal double-differences are used instead of the previously used temporal referenced single-differences. The confrontation of measurements and subsidence model predictions takes place at the level of changes of surface positions between two points in space and two epochs in time (see Figure 24). Thus, no assumptions on a stable reference point are necessary. The original recommendation from LTS-I was to use multiple reference points and multiple reference epochs to form the set of double-differences. This recommendation has been revised in the meantime (Samiei-Esfahany and Bähr, 2017), because the assumption that using a single reference point and epoch would introduce a bias turned out to be invalid. Therefore, many double-differences in LTS-II have identical reference points and epochs.



$$\text{DD\_OBS} = \underbrace{(H_{P_{\text{to}}}^{T_{\text{to}}} - H_{P_{\text{from}}}^{T_{\text{to}}})}_{\Delta H^{T_{\text{to}}}} - \underbrace{(H_{P_{\text{to}}}^{T^{\text{from}}} - H_{P_{\text{from}}}^{T_{\text{from}}})}_{\Delta H^{T_{\text{from}}}}$$

H_p^T is the height of benchmark P at epoch T

Figure 24. Concept of double-differences (figure provided by TU Delft).

3. The stochastic properties of the geodetic observations are described by their full covariance matrix. Compared to previous approaches this yields a more realistic assessment of uncertainties.
4. Individual outliers are identified by a more formal and objective approach. The recommendation from LTS-I was to identify outliers based on an a priori subsidence model

(based on a reservoir scenario and geomechanical parameters) and to account for residual model uncertainties by relaxing the threshold in a way that only the most obvious outliers are rejected. This approach turned out to be too conservative, which is why the a priori subsidence model was replaced by the assumption that the observed deformation is smooth in space and time. This allowed for a stricter outlier identification while "tuning" the measurements towards a previous model prediction could be avoided.

5. The discrepancy between measurements and subsidence model predictions originates not only from uncertainty of geodetic observations (measurement noise) and model uncertainty but also from shallow movements. These movements, like building settlements, cannot be attributed to measurement imperfection and are not considered in the geomechanical model either. Thus, they are subsumed by the so-called idealisation noise, which is now taken into account.

5.2 Survey data and measurement noise

Figure 25 provides an overview of the datasets used in the study.

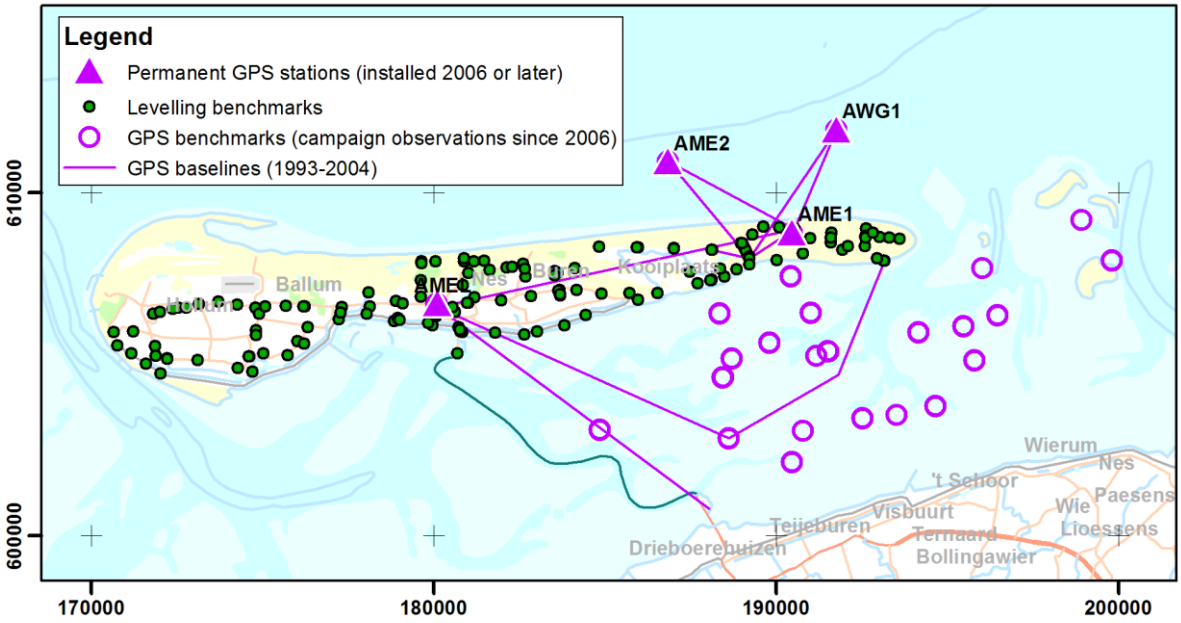


Figure 25. Spatial distribution of geodetic observations used in this study (final dataset, after outlier handling).

5.2.1 Levelling

Levelling data from 26 campaigns between 1986 and 2017 have been used in this study. An attempt was made to complement optical levelling data by hydrostatic levelling observations. This technique of measuring height differences by means of a water-filled tube has been applied by Rijkswaterstaat in the study area until 2002. Thus, also benchmarks in the Waddenzee could be connected to the levelling network on the island. However, the levelling observations in the Waddenzee are too sparse for effective quality assurance (see section 5.4.1) and have thus been excluded from the final dataset. To quantify the measurement uncertainty, standard models have been used.

5.2.2 GPS network (since 2006)

The more recent GPS observations originate from a network of permanent monitoring stations established in 2006. Four of them have been used in LTS-II (see Table 15). The network is extended

campaign-wise by the benchmarks in the Waddenzee. The first campaign was in 2006, and the last campaign considered for this project was in 2016. Benchmarks on the platforms AWG1 and AME2 have been observed in several campaigns before installing continuously operating stations in 2014.

Table 15. Continuous GPS stations used in this study.

Name	Location	Installation
AME1	Production location on Ameland	2006
AME2	Production platform	2014
AMEL	Nes village	2014
AWG1	Production platform	2014

The data from the four continuously operating stations have been sampled at the average dates of the campaigns as well as an additional sampling epoch in January 2017. They have been corrected for temperature effects, atmospheric loading as well as annual and semi-annual harmonics (Van Leijen et al., 2017). The latter correction compensates for systematic errors that are inherent to the GPS system. The measurement uncertainty is modelled by the stochastic model proposed in LTS-I (Williams 2015), which comprises the uncertainty of determining the antenna position and a setup error. This setup error of 1.5 mm is applied to the campaign data only and accounts for the yearly mounting of the equipment.

The previously mentioned concept of spatio-temporal differencing has been applied to the GPS data in a modified way. NAM obtains the processed data in form of displacements with respect to an assumedly stable reference network of currently more than 10 stations. Their stability is closely monitored in yearly intervals. The spatial differences are formed with respect to a virtual reference station representing the whole network of reference stations.

For reliability purposes, the monitored benchmarks in the Waddenzee consist of clusters of three. Observations from all three benchmarks of a cluster have been used to maximise the information content and avoid a subjective and unnecessary preselection.

5.2.3 GPS baselines (1993-2004)

Prior to 2006, only few GPS observations have been carried out by NAM in the study area, mainly to establish a connection from the island to the platforms AWG1 and AME2. Benchmarks have been occupied and processed in pairs (baselines). Since very little is documented about the processing approach from more than a decade ago, the opportunities to a tailor stochastic model were limited. Hence, a conservative standard deviation of 11 mm has been assumed for these height differences, based on experience with comparable datasets.

5.2.4 InSAR

Despite the availability of InSAR (Interferometric Synthetic Aperture Radar) data from Ameland, these have not been used in the study for various reasons. The spatial coverage in the eastern part of the island, where the gas field is located, is very sparse. In addition, the quality of the available measurements is poor due to lack of stable Radar reflectors like well-founded buildings in that area. This together with pending imperfections in stochastic modelling would require a very conservative approach for quantifying the uncertainty. After all, there would have been little (if any) added value to this project while the effort of data preparation would have been disproportionately large.

5.3 Idealisation noise

This noise component describes displacements of the measurement points due to shallow movements like building settlement or soil compaction. Together with the measurement noise, it models the uncertainty of geodetic observations when used to quantify subsidence caused by (deep) reservoir compaction. The model differentiates between two components:

- *Temporal component:* This subsumes all shallow effects that are correlated in time and uncorrelated in space. It can be considered as an autonomous movement of an individual benchmark that has nothing in common with the behaviour of neighbouring benchmarks. An example would be settlement of the individual building a benchmark is attached to.
- *Spatio-temporal component:* This subsumes all shallow effects that are correlated in both space and time. It can be considered as a coherence in movement with neighbouring benchmarks, whereas the level of coherence decreases with distance. Examples would be the compaction of a shallow peat layer or the response to ground water level changes.

Both components can be described by a five-parametric model (Samiei-Esfahany and Bähr, 2015 and 2017), for which different representations were discussed during LTS-I. For the spatial dependency of the spatio-temporal component both an exponential ($e^{-\left(\frac{\Delta d}{L}\right)}$) and a Gaussian ($e^{-\left(\frac{\Delta d}{L}\right)^2}$) model were considered. After an exponential model had been chosen in the final LTS-I report due to a slightly better fit, it was pointed out in the LTS-I review that this model is suboptimal, because it creates unrealistically shaped functions in least squares collocation. Against this background, it seemed reasonable to use a Gaussian representation in LTS-II. The full model reads for the variance of a double-difference observation:

$$\sigma_{DD}^2 = \underbrace{2\boldsymbol{\sigma_t^2}\Delta t^{\boldsymbol{p_t}}}_{\text{temp. comp.}} + \underbrace{2\boldsymbol{\sigma_s^2}\left(1 - e^{-\left(\frac{\Delta d}{\boldsymbol{L}}\right)^2}\right)\Delta t^{\boldsymbol{p_s}}}_{\text{spatio-temp. component}}$$

During the two phases of LTS, two datasets have been used to estimate idealisation noise parameters (see Figure 26):

1. Calibration of both the temporal and the spatio-temporal component to an **onshore** levelling dataset from the LTS-I study area in Northern Friesland. In that area, both components can be estimated due to absence of gas production (Samiei-Esfahany and Bähr, 2015 and 2017). Both components turned out to be significant.
2. Calibration of the temporal component to the relative intra-cluster movements of the **offshore** benchmarks (van Leijen et al., 2017). Levelling between these benchmarks, that are placed in clusters of three with mutual distances ≤ 15 m, provide the opportunity to estimate the temporal component in isolation. This is possible, because all spatially correlated signal components cancel out for short distances.

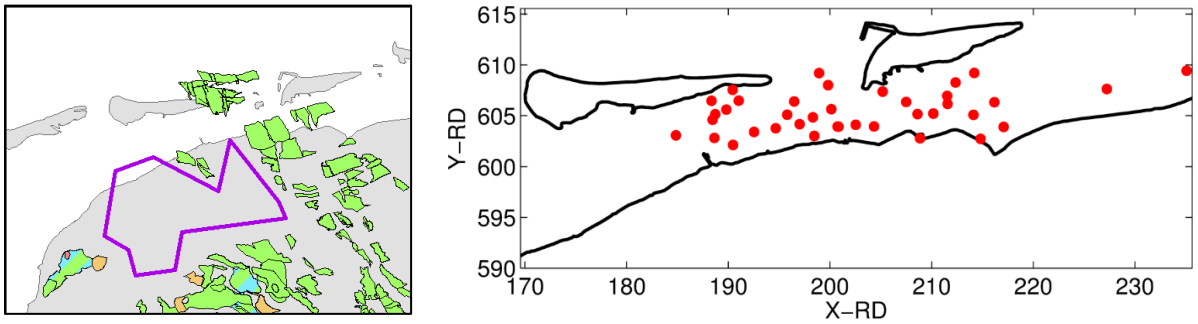


Figure 26. Study areas for the determination of idealisation noise parameters. Left: LTS-I onshore dataset (Samiei-Esfahany and Bähr, 2015 and 2017). Right: LTS-II offshore dataset (van Leijen et al. 2017).

The non-uniqueness of the model parameter estimation triggered an examination of its robustness in more detail. For that, the offshore dataset was a logical starting point, because a two-parametric model for the temporal component is less complex than a five-parametric model for both components. From Figure 27, it can be seen that even though the estimation is stable, it is not well-constrained. Whereas point A marks the global minimum, the parameter set marked by B gives a slightly less optimal but still reasonable fit. In the figure on the right-hand side can be seen that models A and B yield similar standard deviations for time spans up to 9 years, which is the maximum time span covered by the underlying dataset. Model A, however, performs poor for longer time spans. It suggests quite high standard deviations, increasing almost linearly with time. This is not an expected behaviour for benchmark settlement, which is suspected to be the primary driving mechanism for spatially uncorrelated autonomous movements. Settlement would be expected to decelerate over time.

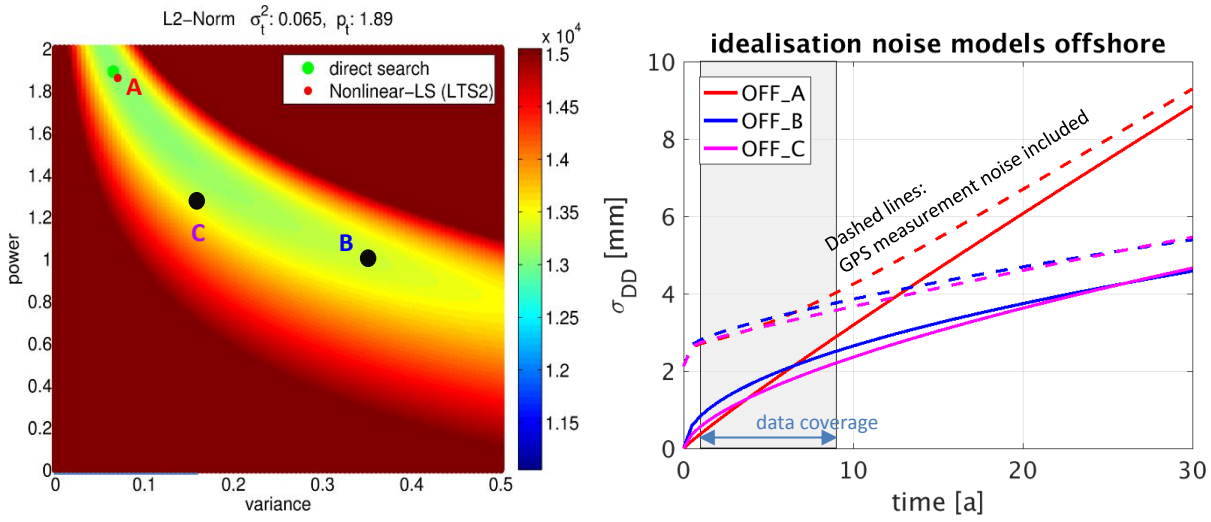


Figure 27. Estimation of the temporal idealisation noise component from intra-cluster levellings between the Waddenzee benchmarks. Left: Solution space: A greenish colour (corresponding to a low L2 norm) indicates a good agreement between data and model. Point A marks the global minimum. Point B is arbitrarily chosen to represent an extreme case with a still reasonable fit. Point C represents the parameters estimated by Houtenbos and Kenselaar (2001). Right: Contribution of the temporal idealisation noise component to the standard deviations of double-difference measurements for the models indicated on the left-hand side. See Table 17 for the model parameters.

For comparison, also a third model is considered in Figure 27: Model C (Houtenbos and Kenselaar, 2001)[1] is close to model B, still reasonably close to the minimum area in the solution space (see Figure 27) but based on a considerably larger dataset (see Table 16). Model C is slightly less optimal than model A for describing the idealisation noise in the LTS-II offshore dataset, because it does not minimise the objective function and is not tailored to the benchmark type in the Waddenzee. However, it is more appropriate for longer time spans. Due to the size of the underlying dataset, which includes observations from the whole of the Netherlands, it is also considered a more precise estimate. Therefore, model C is considered most appropriate to be used for the temporal idealisation noise component of offshore benchmarks in LTS-II. Note that this doesn't mean that a spatio-temporal component is absent for these benchmarks; but it is not possible to determine its parameters based on the offshore dataset.

Table 16. Comparison of studies to estimate idealisation noise parameters: Houtenbos and Kenselaar (2001), LTS-I onshore dataset (Samiei-Esfahany and Bähr, 2015 and 2017), LTS-II offshore dataset (Van Leijen et al., 2017).

	Houtenbos and Kenselaar (2001)	**LTS-I onshore**	**LTS-II offshore**
Data	unadjusted sections	adjusted networks	benchmark clusters
Area	whole Netherlands	Northern Friesland	part of Waddenzee
Temporal sampling	0.5...40 a (Δt = 1 a)	5...25 a ($\Delta t \approx$ 5 a)	1...9 a ($\Delta t \approx$ 1 a)
Spatial sampling	0.05...1.6 km (Δs = 0.1 km)	0.25...12 km (Δs = 0.5 km)	0 km
temporal component	most stable estimation (large dataset)	poorly constrained for t < 5 a	poorly constrained for t > 9 a
spatio-temporal component	not significant (poorly constrained)	estimation possible	not estimable

The dataset of Houtenbos and Kenselaar (2001) also suffered from too short spatial distances to reliably estimate the parameters of a spatio-temporal idealisation noise component. Therefore, the calibration of a full five-parametric model (including both temporal and spatio-temporal component) is tackled with the LTS-I onshore dataset. It had already been concluded that this estimation is very unstable. It fails for most of the simulations of Samiei-Esfahany and Bähr (2017). Therefore, it was deemed necessary to constrain at least one parameter to a given value. For providing an appropriate constraint, the model of Houtenbos and Kenselaar (2001; "OFF_C") is a suitable candidate, because it is based on a large set of onshore observations from the whole of the Netherlands. Tests have shown that the result doesn't change significantly if not only p_t but also σ_t^2 is constrained to this model. Constraining both parameters instead of one is also convenient, because this yields a homogeneous parameter set for both offshore and onshore observations.

The parameters of the resulting idealisation noise model "ON_C" that comprises both a temporal and a spatio-temporal component are specified in Table 17, and the corresponding standard deviations for double-differences are evaluated in Figure 28. In conclusion, ON_C is considered the most appropriate idealisation noise model and has been applied in this study to both the onshore and the offshore benchmarks. Nevertheless, it should be noted that even though two of the five parameters have been constrained, the three estimated parameters are still uncertain to some degree. (The standard deviation of σ_s^2 is about half the parameter value itself.) With a correlation coefficient between σ_s^2 and p_s of -0.95, the uniqueness of the parameters is also arguable. This should be kept in mind when using the parameters.

[1] 3-component model for "other" (i. e., non-underground) benchmarks with outlier removal up to 2.5 σ from Houtenbos and Kenselaar (2001), Table 5.5. Mind the different definition of p_t in that publication.

Table 17. Parameters of different idealisation noise models with corresponding standard deviations.

	σ_s^2 [mm²/km/a^ps]	L [m]	p_s [–]	σ_t^2 [mm^2/a^pt]	p_t [–]	model description
OFF_A	–	–	–	**0.07** ±0.003	**1.86** ±0.021	LTS-II offshore dataset, global minimum
OFF_B	–	–	–	**0.350** –	**1.000** –	LTS-II offshore dataset, extreme case
OFF_C	–	–	–	**0.16** ±0.002	**1.24** ±0.002	Houtenbos and Kenselaar (2001)
ON_C	**0.465** ±0.229	**3956** ±390	**1.590** ±0.154	**0.160** –	**1.240** –	LTS-I onshore dataset, σ_t^2 and p_t constrained to Houtenbos and Kenselaar (2001)

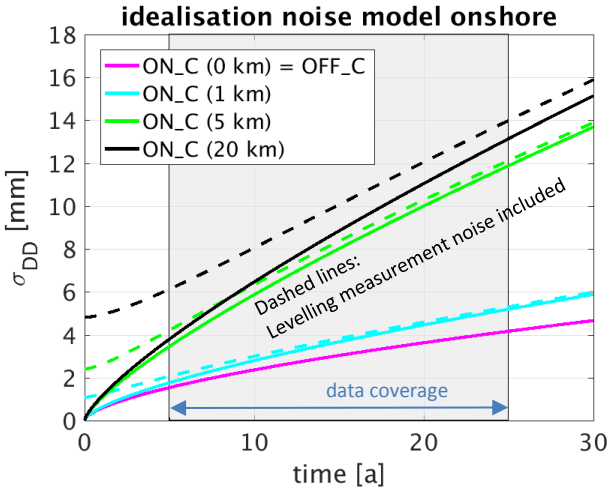


Figure 28. Idealisation noise model estimated from the LTS-I onshore dataset, evaluated for standard deviations of double-differences covering spatial distances of 0, 1, 5 and 20 km.

5.4 Outlier handling

In the following, the different outlier handling procedures for levelling and GPS are explained in detail.

5.4.1 Levelling

For levelling data, a dedicated outlier handling approach has been proposed in LTS-I (Bähr and Samiei-Esfahany, 2015) and implemented by Van Leijen et al. (2017) with the intention to be applied in LTS-II. It will be referred to as LTS-I approach in the following. It is based on time series analysis per benchmark of double-differences with a common reference point and a common reference epoch. The expected deformation of an a priori subsidence model is subtracted from these time series. Subsequently, hypotheses on disturbances, identification errors and abnormal behaviour are tested on these time series of model residuals (see Figure 29).

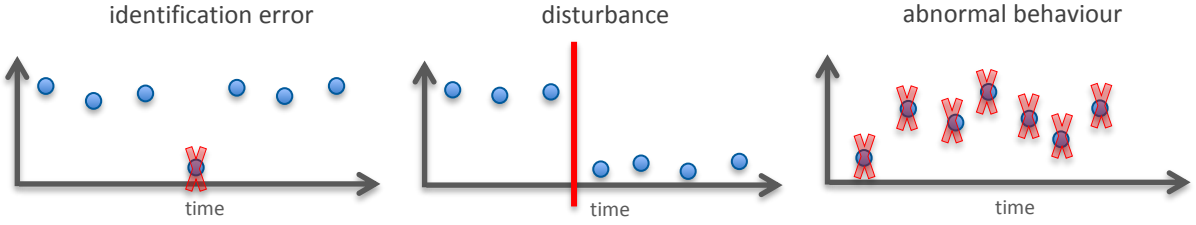


Figure 29. Types of alternative hypotheses for residual time series and corresponding actions: Identification errors are removed from the dataset. Where disturbances are detected, time series are split in two. In case of abnormal behaviour, the entire time series is rejected.

Disturbances are jumps in a time series and can be addressed by splitting the time series in two parts. Identification errors are outliers in a single epoch, caused for instance by misidentification of a benchmark. The appropriate remedy is here to reject the single observation from the affected epoch. Abnormal behaviour is stated if the time series of residuals does not match the presumed stochastic model. The stochastic model for the testing, comprises of both measurement noise and idealisation noise. To cover residual imperfections in this stochastic model, and since subsidence modelling imperfections and uncertainties are not covered at all, the testing threshold is largely increased.

Drawbacks of this approach are its poor sensitivity, the assumption of a stable reference point, incomplete exploitation of spatio-temporal relationships and limited fitness for short time series. The consequences of its application were that many outliers remained undetected, because the conservative framework allows only for detection of clearly obvious outliers. This resulted in far too high test statistic values for the confrontation of geodetic observations and subsidence model predictions.

To mitigate the large test statistics, some additional outlier handling was applied by visualising time series that were suspected of containing outliers based on some auxiliary analysis. Some more potential outliers have been assessed manually, and the dataset has only been altered in case of obvious outliers. Altering the dataset or “handling” outliers does not necessarily imply complete rejection of benchmarks. In most cases, depending on the type of error, time series are split in two independent parts (disturbances), or single observations are excluded from the dataset (see Figure 29).

Even after these remedies, the test statistic value was significantly above the expectation value. This may be an indication for a substantial (subsidence) model mismatch, but it became clear that remaining outliers in the geodetic data made a substantial contribution. Since all opportunities of conservative and (almost) assumption-free outlier handling had been exploited, other routes had to be explored. During progress meetings with experts there was general agreement on spatio-temporal smoothness of subsidence signals that are caused by hydrocarbon production. This smoothness assumption has been exploited to detect some more still significant but less obvious outliers. To achieve this, the workflow has been complemented by a spatio-temporal analysis of the geodetic observations, and there was no longer a need to use an a priori subsidence model based on geomechanical parameters and reservoir scenarios. An overview flowchart is shown in Figure 30.

The spatio-temporal analysis (Methode Houtenbos; TCBB, 2009) on the right-hand side works as follows. The levelled height differences are subdivided into two contributions: the height differences at a reference epoch (here: 1986) as estimable parameters and the deformation that has taken place since then. The latter is considered a stochastic signal that can be predicted for any point in space and time within the observed area and time interval. This is achieved by means of least squares collocation or Best Linear Unbiased Prediction (BLUP), respectively.

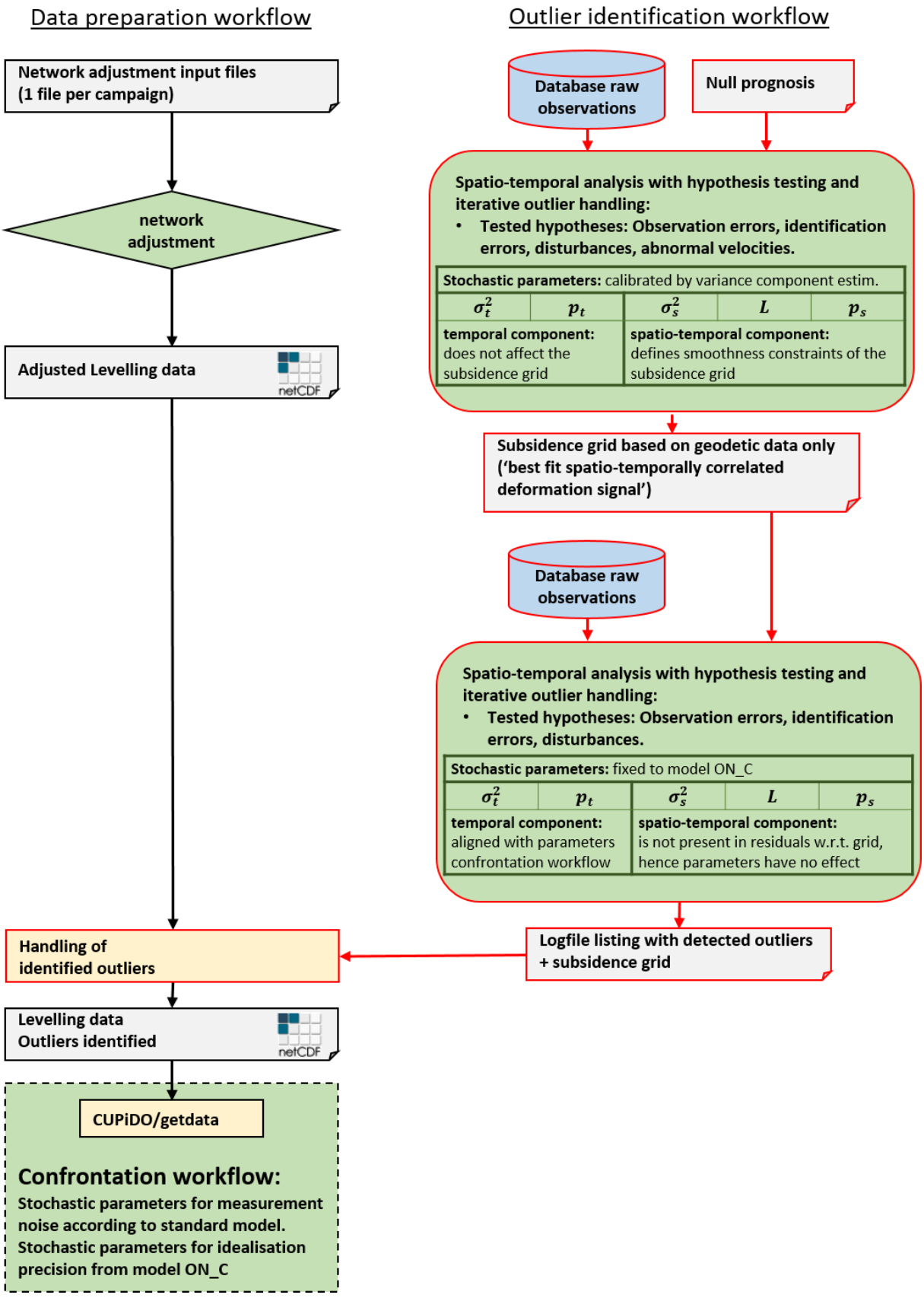


Figure 30. Outlier handling workflow for levelling data.

To minimise the prediction bias, a first order approximation of the deformation signal needs to be subtracted from the observations beforehand. Usually, an existing subsidence model is used for this purpose. But this would also contaminate the result with subsidence modelling assumptions. To avoid this, a two-step approach is pursued: As an initial approximation, zero deformation (a “null prognosis”) is used to estimate a spatio-temporally correlated subsidence signal from the geodetic data. From this signal, a gridded output is generated, which is used as approximation of the subsidence signal in a second run.

During the spatio-temporal analysis, outlier hypotheses are tested, and the data model is adapted accordingly in subsequent iterations. The hypotheses cover observation errors (i. e., inconsistencies in single observations that would cause significant loop misclosures in an epoch-wise adjustment), identification errors (see Figure 29) disturbances (see also Figure 29) and abnormal velocities. The latter are related to autonomously behaving benchmarks with a significantly deviating linear subsidence rate. Abnormal velocities are disregarded as a hypothesis in the second run, because they cannot easily be integrated into the confrontation workflow. They would be detected in only a limited number of cases, which are covered by other hypothesis types.

The list of detected outliers is exploited automatedly to handle all detected identification errors and disturbances in the LTS-II dataset according to Figure 29. (The absence of observation errors is guaranteed by the prior campaign-wise network adjustment). A second result from the spatio-temporal analysis is a subsidence grid representing the predicted smooth surface based on geodetic data only. This is used for an additional crosscheck following the LTS-I outlier detection approach. In addition, only a few observations are removed manually for unquestionable and documented reasons (see section 5.4.4).

A drawback of using the spatio-temporal analysis is that it is only valid in the area and time interval that is covered by observations with sufficient density. Since this is only the case for the island of Ameland, all levelling observations outside the island are excluded from the confrontation workflow, because their reliability cannot be guaranteed at the same level of confidence. Note that GPS data outside the island are still used. Whereas the GPS campaign data (since 2006) have some redundancy within the benchmark clusters, the GPS baseline data (until 2004) is considered of limited influence due to its larger uncertainty.

5.4.2 Consistency of stochastic modelling

Outlier assessment is always based on a data quality assumption. In geodesy, this assumption is usually parameterised in a stochastic model. If an observation does not satisfy the stochastic model with a high likelihood, it is considered an outlier. Since (levelling) outliers are identified in a separate workflow prior to the confrontation with subsidence modelling, it is essential to ensure that the stochastic models assumed in both steps are aligned with each other.

The idealisation noise model used in the confrontation workflow consists of five parameters: two parameters (σ_t^2 and p_t) for the temporal component and three parameters (σ_s^2, L and p_s) for the spatio-temporal component. The parameters for the temporal component have been determined by Houtenbos and Kenselaar (2001) from a dataset with outliers removed. The parameterisation of the spatio-temporal component is based on the variogram approach from Samiei-Esfahany and Bähr (2015, 2017). This approach uses a robust algorithm, which is insensitive to outliers in the underlying dataset.

All five stochastic parameters have been calibrated under the assumption that the underlying dataset is free of outliers. Consequently, it is consistent to also remove outliers from the data before applying these stochastic parameters in the confrontation workflow. Still, the principal consistency criterion is that the stochastic model parameters themselves are identical in both outlier removal and confrontation with subsidence modelling. This is the case for the second iteration of the spatio-temporal analysis in the outlier identification procedure. In the first iteration, however, the five stochastic parameters have been calibrated by variance component estimation to obtain an optimal fitting smooth subsidence grid in space and time. This is necessary, because the signal does not comprise only idealisation noise but also the full subsidence signal due to reservoir compaction. Both contributions need to be modelled combinedly by the same a priori unknown stochastic parameters. After estimating the subsidence grid and subtracting it from the observations in the second iteration, no spatio-temporally correlated signal is present in the residuals. Hence, the three corresponding parameters have no effect on the resulting set of outliers. All that matters are the two parameters of the temporal component, which are aligned with the confrontation workflow.

5.4.3 GPS

For the GPS part, the outlier handling procedure from Van Leijen et al. (2017) has been applied. Outlier testing is based on redundant information that is available from the clusters of three benchmarks. Observations have been excluded in case of significant changes in the relative heights within a cluster.

Since significant settlement effects can be observed on benchmarks placed in 2006 (M-points: M***N, M***M, M***Z), all observations on these benchmarks from 2006 had been excluded from the dataset beforehand. As it is unclear if the settlement is still an issue in the subsequent years, observations from one M-cluster in 2007 and two M-clusters in 2008 have additionally been excluded.

In contrast to the GPS campaigns, the continuous GPS stations have no inherent means of validation. The reliability of the vertical component is assured by comparison with levelling (for AME1 and AMEL) or InSAR (for AWG1 and AME2). For the horizontal GPS observations, which are available only for AME1 (11 years) and AMEL (2 years) there are no redundant measurement points to be compared with. When it comes to uncertainties, there is absolutely no insight into eventual horizontal movement of the monument, nor are there any models for horizontal idealisation noise. While especially the horizontal movement of AME1 (close to the centre of the subsidence bowl) would be a valuable contribution to subsidence modelling, which has been stressed by many subsurface experts in the past, there is a substantial risk that unmodeled shallow effects leak into the model of the deep subsurface. To mitigate this risk, the horizontal GPS observations have not been used in LTS-II.

5.4.4 Detailed documentation

This section transparently provides detailed information on data selection and outlier removal.

Initially, all available data inside the predefined area of interest (see Figure 45) have been considered except for two levelling campaigns (279H08 and 279H09). For these, contradictory indications existed regarding the year of observation (1987 or 1988), and it was not possible to resolve this inconsistency. In addition, horizontal GPS observations have been disregarded as explained in

section 5.4.3). Figure 31 gives an overview of the usage of all observations that remained after these previously mentioned exclusions.

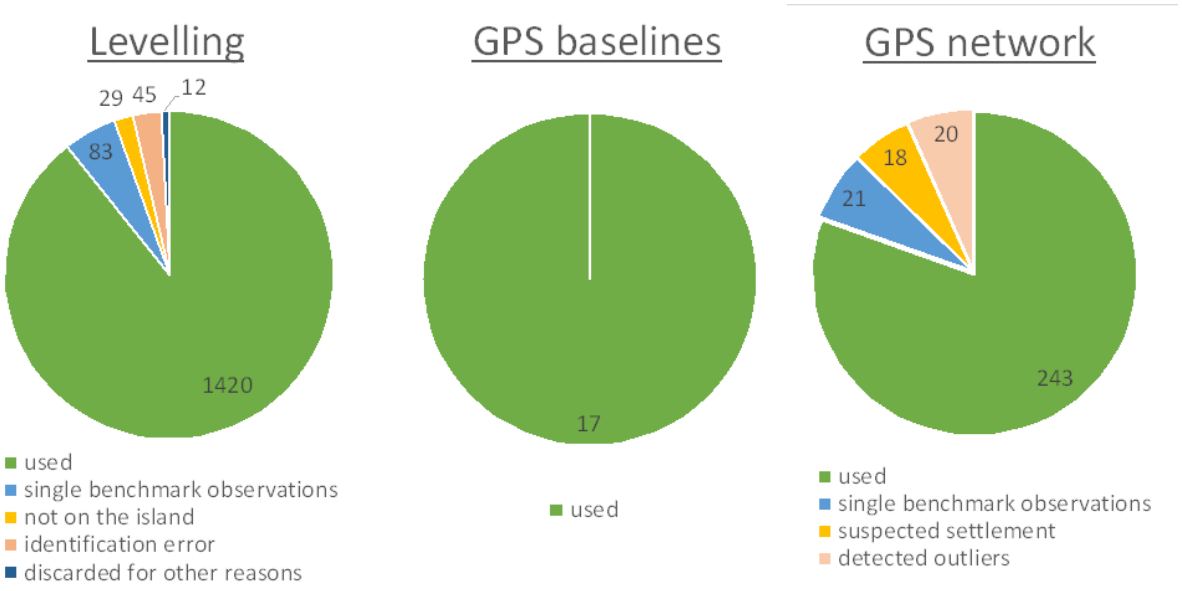


Figure 31. Usage of available (spatial) single-difference observations for the three observation techniques (1589 levelling, 17 GPS baselines until 2004, 303 GPS network since 2006).

While the confrontation between geodetic observations and subsidence modelling takes places on the level of double-differences, outlier identification and removal is applied to single-differences. There are 104 benchmarks in the dataset that are observed only once (83 by levelling, 21 by GPS). Their observations cannot be differenced with a second epoch. Hence, they cannot contribute to any double-difference and remain unused.

Furthermore, it has been explained in section 5.2.1 that levelling observations outside the island of Ameland are disregarded, because their reliability cannot be assured by a spatio-temporal analysis. This concerns another 29 observations that are also excluded.

Among the remaining observations, 45 identification errors have been detected, and 12 other observations have been excluded for individual reasons elucidated in section 5.4.5.

Among the GPS observations of benchmarks that have been observed more than once, 18 observations have been excluded due to suspected settlement (M-points placed in the year 2006). Another 20 observations have been excluded, because they have been identified as an outlier.

The spatial distribution of excluded observations is visualised in Figure 32. Except for benchmark 002C0117, where 10 out of 10 observations have been excluded (see section 5.4.5), the spread of outliers over the whole area of interest does not suggest any systematic impact on the modelling. For the GPS clusters in the Waddenzee it should be noted that there is no epoch, for which all three observations within a cluster have been removed due to outliers. This has only happened for the M-benchmarks due to suspected settlement effects in the years 2006-2008.

Figure 33 gives an overview of the detected disturbances. There are in total 9 benchmarks affected by one disturbance, and 1 benchmark (002C0112) is affected by two disturbances. A disturbance does not imply that (spatial single-difference) observations are removed. However, when combining

spatial single-differences to spatio-temporal double-differences, one double-difference less can be formed per disturbance. Thus, even though no observation is removed, one piece of information is lost per disturbance. Note that many disturbances have limited impact, because they occur close to the beginning or the end of a time series, or the time series is relatively short anyway.

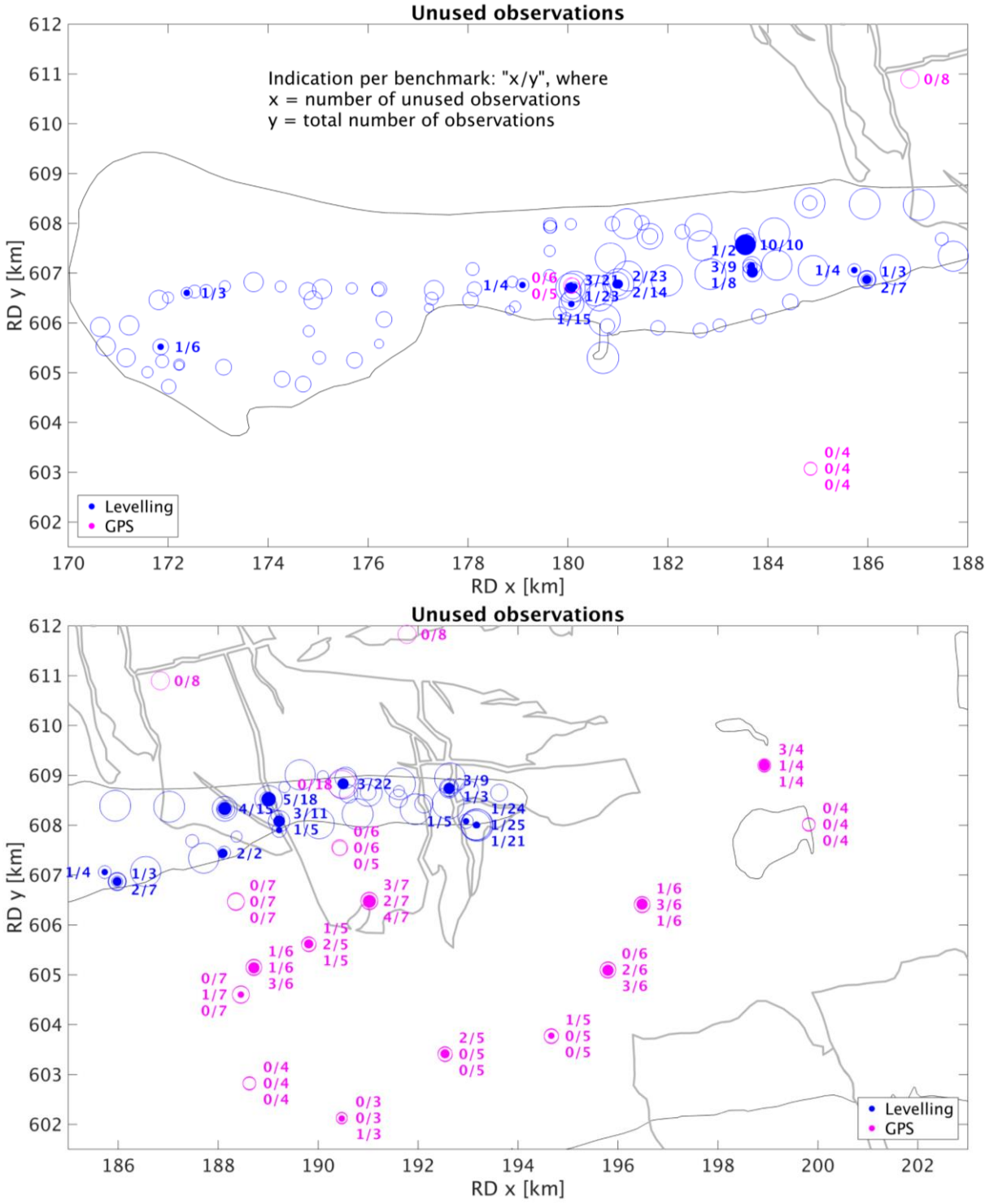


Figure 32. Visual overview of unused observations. Every benchmark is depicted by a circle, the area of which is proportional to the number of (spatial single-difference) observations to this benchmark. If one or more observations are not used, this is indicated by a coloured circle within, its area being proportional to the number of unused observations. The number of unused observations and the total number of observations are indicated next to the benchmark location.

For levelling observations, this indication is only given at benchmarks with minimally one unused observation to retain the clarity of the plot.

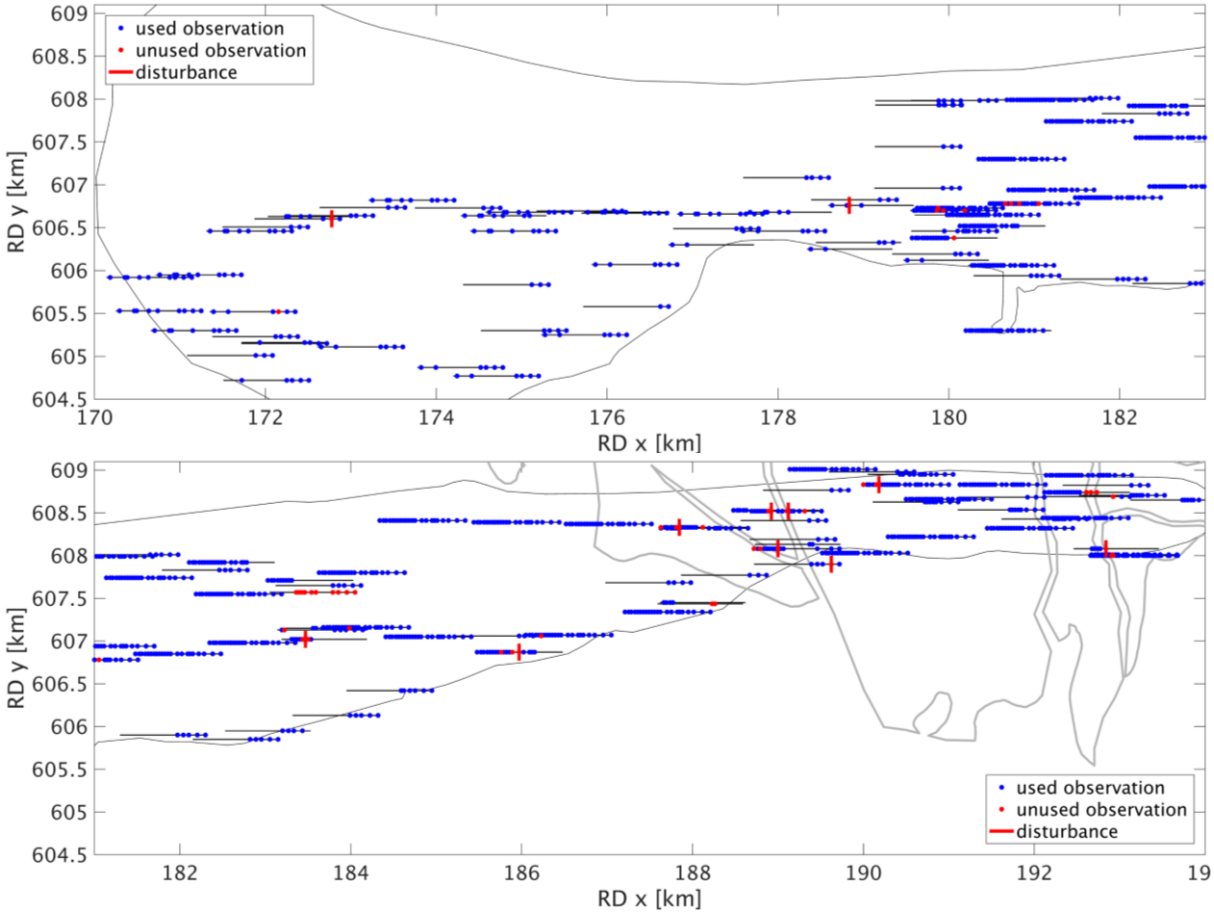


Figure 33. Visual overview over the localisation of disturbances in space and time. Every benchmark is depicted by a horizontal line representing a time bar from 1986 until 2017. For every (spatial single-difference) observation, a blue dot is placed on the time bar at the epoch when it was taken. If the dot is red, the observation has been excluded due to a detected identification error or for another documented reason (see section 5.4.5). Finally, the 11 red vertical bars represent disturbances, depicting the epochs at which the time series have been split in two. Note that the time bars of benchmarks on the same spot may overlap each other. This can hamper the interpretation but does not distort the general picture.

5.4.5 Individual cases

In two cases, all observations of a benchmark have been removed from the dataset ("discarded for other reasons" in Figure 31).

0009994: This benchmark has been observed twice (in 2006 and in 2007), observing an uplift of 3.5 cm. Whereas this is an obvious outlier, it cannot be stated that either of the two observations is an identification error. The entire time series of two observations behaves abnormally and has therefore been removed completely.

002C0117: From Figure 34 can be seen that the double-difference observations of this benchmark jump up and down in the order of centimetres and do not match with observations from benchmarks nearby. Benchmark 002C0117 is a bolt on the top of a mushroom-shaped guidepost ("paddenstoel") that is standing askew. This is sufficient indication to conclude that the benchmark monumentation itself is highly unstable. To avoid speculations on when exactly the instability began, all 10

observations have been removed. This is a tolerable loss, seeing that there are sufficient observations on other benchmarks nearby.

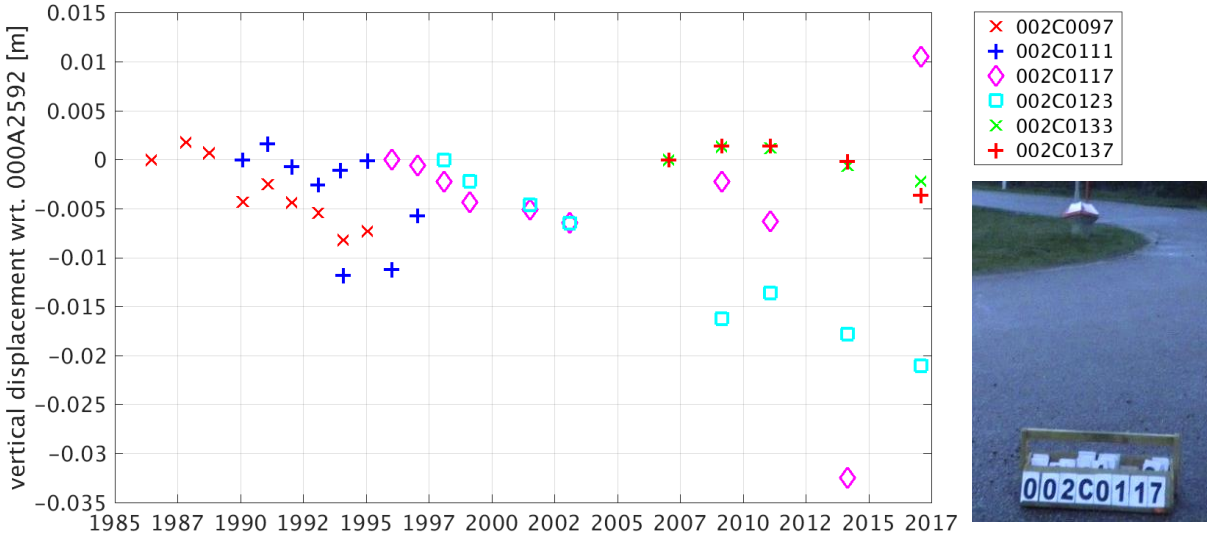


Figure 34. Double-difference time series of benchmark 002C0117 together with time series of other benchmarks within 600 m distance. The temporal reference is the epoch of first observation for every individual benchmark.

6 Workflow confrontation of model results with measured data

6.1 General, description

A probabilistic workflow was developed with the objective to confront model results with the measured data. This confrontation workflow is founded on the workflow described by TNO (2017) with modifications to both the functional form and stochastic form as described in chapter 4 and in this chapter. In our study, the workflow automates the choice of compaction and influence model input variables and confronts the modelled double differences with the measured data for each model member at the location of the benchmarks. A measure of the fit to the data is obtained by calculating the match of the model with the data expressed by a χ^2 value. Next to delivering a probabilistic confrontation workflow, the objective of this study is also to provide an objective statistical description of the outcome, i.e. an expectation case based on the weighted average and a 95% confidence interval of the posterior model ensemble.

Nomenclature

The confrontation workflow uses several words that deserve more explanation and definition:

- Reservoir scenario: a scenario based on a set of reservoir simulator (MoReS) parameter values that describe the spatial evolution of pressures with time.
- Compaction model: a model that translates the change of effective stress due to pressure depletion into reservoir strain.
- Influence function: function which translates the reservoir compaction to surface subsidence.
- Geomechanical model: combination of compaction model and influence function.
- Subsidence model: combination of geomechanical model and a reservoir scenario.
- Member: a model member represents a single run of the subsidence model with specific values for each of the parameters in the influence function and the compaction model and a specific reservoir scenario.
- Ensemble: a collection of members that result from a Monte Carlo analysis.

The confrontation workflow requires the following input:

1. geodetic data;
2. an ensemble of reservoir flow model simulation results (pressure *scenarios*);
3. a prior ensemble of geomechanical realisations or members being built on top of the pressure scenarios (i.e. subsidence members), varying the values of the parameters for both the compaction model and influence function.

The workflow selects randomly values in the geomechanical parameter space and calculates compaction and subsidence using the modified Geertsma and van Opstal (moving) rigid basement approach. The geodetic data are processed into double-difference data and a corresponding covariance matrix. The modelled ensemble is confronted with the geodetic data. A schematic representation of the workflow is provided in Figure 35.

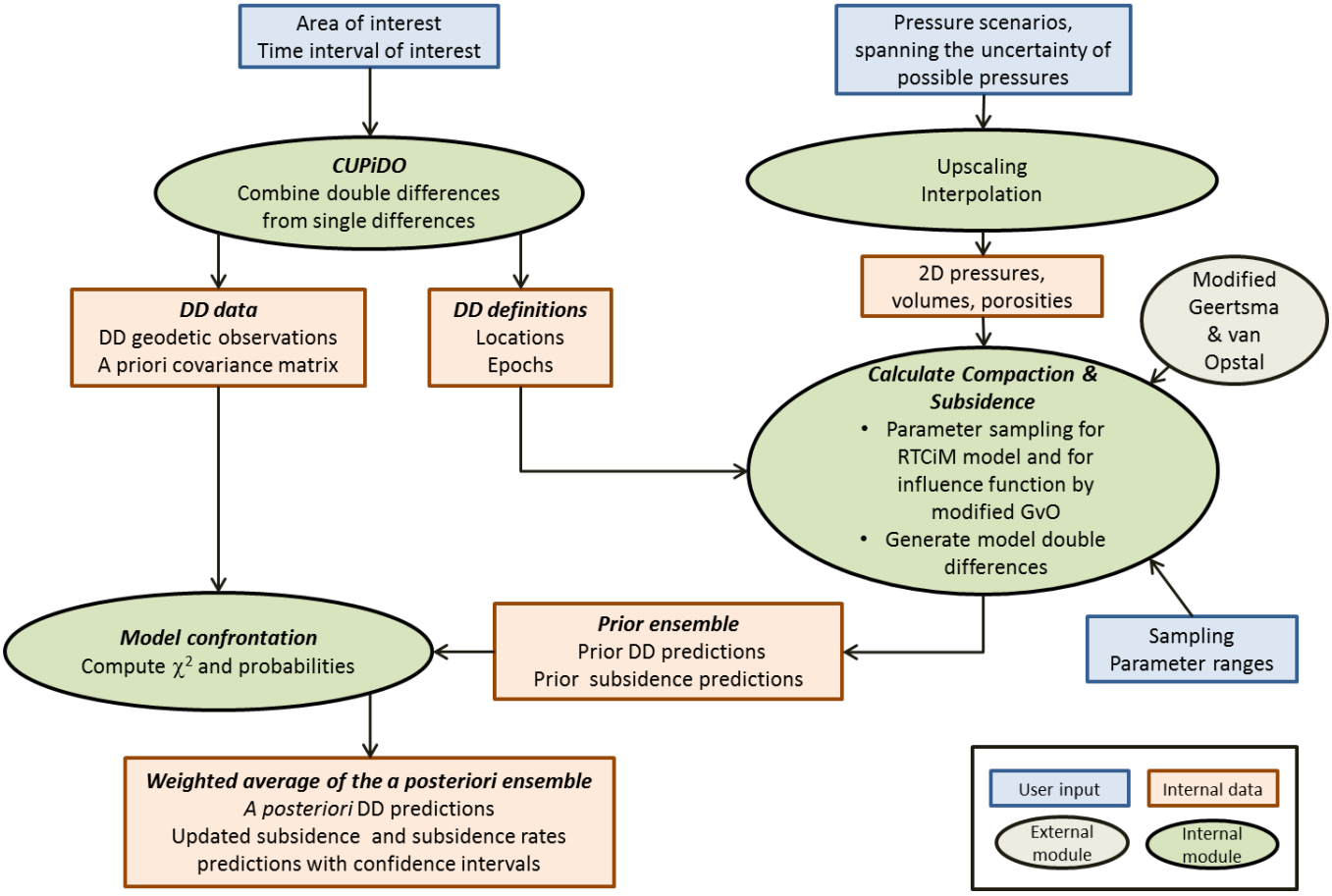


Figure 35. Schematic representation of the NAM confrontation workflow (modified from TNO, 2017).

The confrontation workflow contains 4 internal modules that form the core of the program.

6.2 CUPiDO

The module CUPiDO (Connecting Undifferenced Points in Deformation Observations; working title "getdata") has been created by TU Delft as an interface between geodetic data and the confrontation workflow (Van Leijen et al., 2017). It selects geodetic observations for a specified area and time period of interest. The output is a non-redundant[2] set of double-differenced displacements (see Figure 24) with an appropriate stochastic description in form of a fully populated covariance matrix (see Figure 36). Figure 37 shows the standard deviations of all double-differences as a function of distance and time.

[2] The non-redundancy is a feature of CUPiDO version 1.1.4, which has been released after the publication of (Van Leijen et al., 2017).

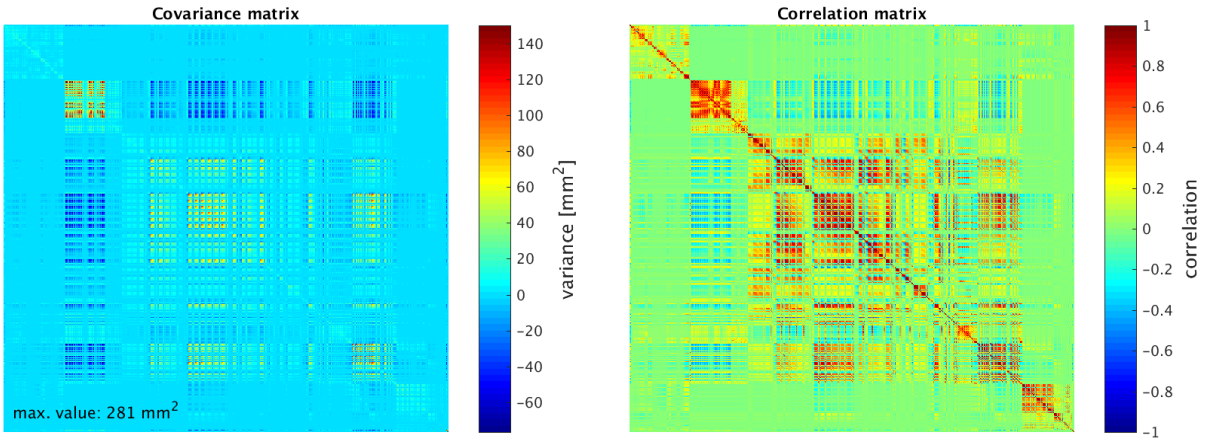


Figure 36. Covariance matrix (left) of the double-difference observations used in LTS-II, including measurement noise and idealisation noise with both temporal and spatio-temporal component according to model ON_C (see section 5.3). The off-diagonal elements describe the stochastic interdependence. The correlation matrix (right) is a normalized form of the covariance matrix, which better visualizes the correlation structure. Note that the colour scale on the left-hand side has been truncated; values between 150 mm² and 281 mm² are all in the same colour.

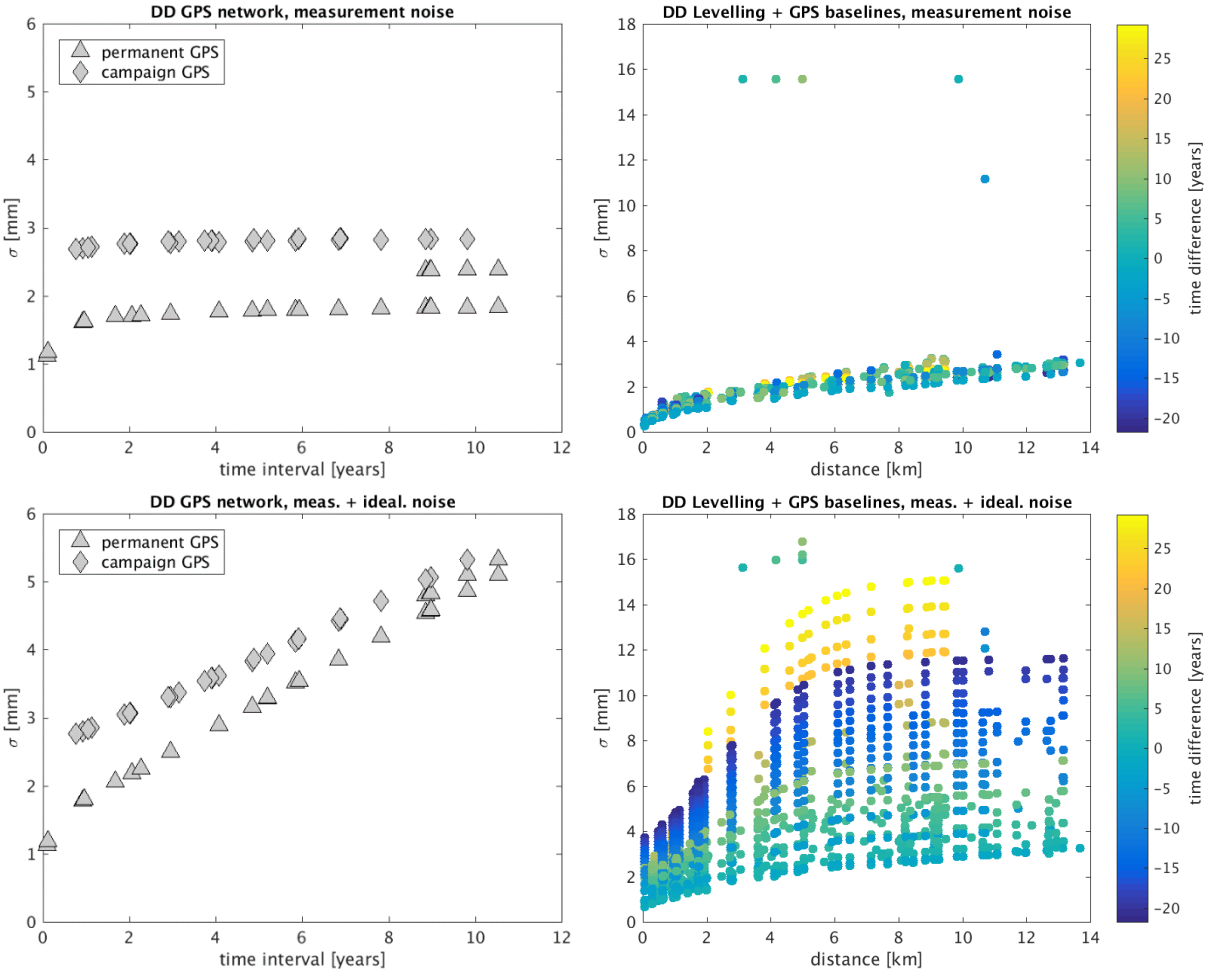


Figure 37. Standard deviations of the individual double-difference observations. Whereas this visualization does not convey the full stochastic information, it provides a good impression of the precision of the individual double-differences. The upper row shows the contribution of the measurement noise, which is related to the uncertainty of the measurement itself. The lower row displays the full uncertainty budget. This also comprises both temporal and spatio-temporal component of the idealisation noise according to model ON_C (see section 5.3). The figures on the left-hand cover GPS network observations (permanent stations 2006-2017 and campaigns 2006-2016), and the figures on the right-hand side cover levelling (1986-2017) and GPS baseline observations (1993-2004).

6.3 The particle filter in the confrontation workflow

The confrontation workflow (TNO, 2017) confronts the model members with the geodetic data including the covariance matrix. To do so, double-differences at the locations of the benchmarks and the acquisition times are computed for the model members.

TNO (2017) describes two methods:

1. The Red-Flag confrontational method (Nepveu et al., 2010).
2. The Ensemble Smoother approach (Fokker et al. 2016).

The LTS-II study results are based on the application of a particle filter, such as the Red-Flag approach, but with modifications as explained in this Section. The goodness of fit between the model results and the geodetic data in the particle filter is expressed by the value of the χ^2 test statistic that is also used for the likelihood calculation of a specific model member.

The ensemble smoother is a methodology that mixes and conditions the output of individual model results to create new model results, finally iterating to a narrower distribution of possible realisations. The mixing of the member solutions makes a mapping to the original reservoir scenarios and geomechanical models less transparent. The aim of LTS-II is to demonstrate the tool and provide more insight in plausible scenarios and input values to the subsidence models. This is the reason why the particle filter method has been chosen.

The particle filter that is used in the confrontation between geodetic data and subsidence models is a modified version of the Red-Flag methodology as described by TNO (2017). The modified particle filter that is used in this study does not exhibit the ensemble 'collapse' feature of particle filters (e.g. Snyder et al. 2008), and hence can produce a subsidence model confidence interval.

The following modifications have been applied to the Red-Flag methodology from TNO (2017):

1. Revision of the covariance matrix used in the χ^2 test statistic.
2. The application of the particle filter to the test statistic value (sum of weighted differences between geodetic data and subsidence models), instead of to the vector of differences. This implies using the χ^2 probability density function instead of the multivariate normal distribution.

In this Section, these modifications are explained.

6.3.1 The covariance matrix in the χ^2 test statistic

The test statistic applied in the confrontation between geodetic data and subsidence models in TNO (2017) is defined for each model ensemble member as:

$$\chi^2 = \left(dd_r^{prior} - dd\right)^T C_{dd}^{-1} \left(dd_r^{prior} - dd\right),$$

where dd_r^{prior} represents the vector of model double-differences of an ensemble member, dd is the vector of geodetic data double-differences, and C_{dd}^{-1} the inverse of the covariance matrix of the geodetic data. The subsidence model members are considered deterministic.

In the revision of the covariance matrix used in the χ^2 test statistic in the particle filter, the concept of potential residual geomechanical model imperfections is introduced.

The revision is explained from the quadrant model in Figure 38. This model discriminates functional models (parametric relations) and stochastic models (statistical relations) in the geodetic data and subsidence modelling space.

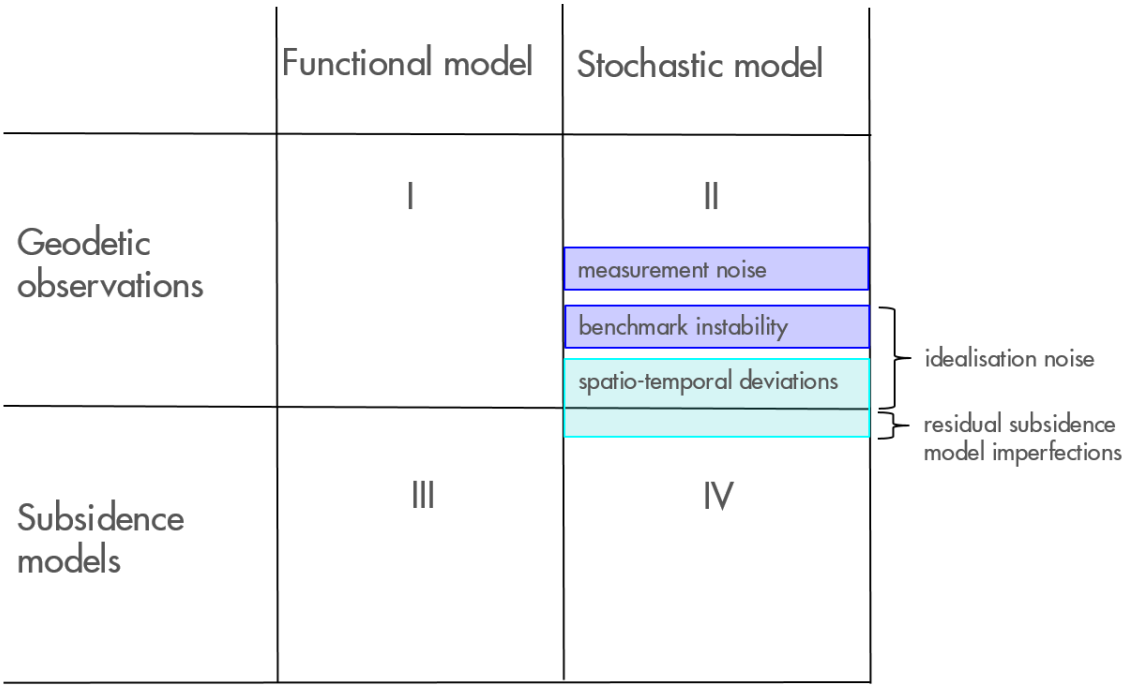


Figure 38. Quadrant model. The functional model describes the functional relationship between observations and parameters. The stochastic model describes the expectable misfit of the functional model due to residual uncertainties. Geodetic observations quantify relative displacements of surface benchmarks. Subsidence models describe the expected surface deformation caused by hydrocarbon extraction. (Residual) spatio-temporally correlated deviations between geodetic observations and subsidence models can be attributed to both quadrant II (idealisation noise) and quadrant IV (subsidence model imperfections).

As explained in Section 5.3, the stochastic model of geodetic observations may contain besides measurement noise, also idealisation noise. Idealisation noise needs to be incorporated if the benchmark displacements are not entirely representative for the subsidence signal of interest (subsidence due to hydrocarbon extraction), for example due to water extraction or heterogeneities in the shallow subsurface. It comprises of a temporal component and/or a spatio-temporal component. The stochastic model of geodetic data double-differences including both idealisation noise components can be defined as (van Leijen et al., 2017):

$$C_{dd} = C_{e_{dd}} + C_{t_{dd}} + C_{st_{dd}},$$

where C denotes the covariance matrix (please be aware that the covariance matrix is denoted as Q in the notation by van Leijen et al. (2017)). The covariance matrices of the measurement noise, the temporal component of idealisation noise, and the spatio-temporal component of idealisation noise are respectively $C_{e_{dd}}$, $C_{t_{dd}}$, and $C_{st_{dd}}$.

If the spatial and temporal evolution of pressures and geomechanical processes in quadrant III (see Figure 38) are known and functionally modelled, and if the model solution space is completely sampled, there would be no need to incorporate residual model imperfections in the stochastic model (quadrant IV). However, the subsidence models are representations of reality to a certain accuracy, and therefore unmodeled subsurface processes will cause residual spatio-temporally correlated deviations between geodetic data and subsurface models. The uncertainty caused by these deviations is also known as epistemic uncertainty or structural uncertainty, i.e. a systematic uncertainty due to processes that are not (exactly) known thus not incorporated in the functional model.

The subsidence models aim to capture the processes related to subsidence induced by hydrocarbon production. Many thousands of simulations are calculated for each ensemble to analyse the long-term effects and the uncertainties. However, the complexity of the models is bounded to manage the calculation time of simulations. If residual model imperfections exist and are not taken into account, these lead to χ^2 test statistic values in the particle filter that systematically exceed the expectation value of the test statistic. To account for this, the potential residual geomechanical model imperfections can be accounted for in the stochastic model by a covariance matrix that represents the residuals that are correlated both in time and space:

$$C_{st_{mm}}$$

Incorporation of model uncertainties in the stochastic model is effectively the same as including a covariance matrix to represent uncertainties in the forward modelling, such as the covariance matrix C_T in Tarantola (2005).

Note that both residual subsidence model imperfections (quadrant IV) and the spatio-temporal component of idealisation precision (quadrant II) are correlated in space and time. If the residual spatio-temporally correlated deviations between geodetic observations and subsidence models have their origin in both benchmark idealisation noise and geomechanical model imperfections, their covariance matrix can be denoted as:

$$C_{st} = C_{st_{dd}} + C_{st_{mm}}$$

with elements that can be attributed to both quadrant II ($C_{st_{dd}}$) and IV ($C_{st_{mm}}$).

Incorporation of residual model imperfections leads to a reformulation of the test statistic from:

$$\chi^2 = \left(dd_r^{prior} - dd\right)^T C_{dd}^{-1} \left(dd_r^{prior} - dd\right)$$

to

$$\chi^2 = \left(dd_r^{prior} - dd\right)^T \left(C_{dd} + C_{st_{mm}}\right)^{-1} \left(dd_r^{prior} - dd\right)$$

which equals

$$\chi^2 = \left(dd_r^{prior} - dd\right)^T \left(C_{e_{dd}} + C_{t_{dd}} + C_{st_{dd}} + C_{st_{mm}}\right)^{-1} \left(dd_r^{prior} - dd\right)$$

where $C_{e_{dd}}$represents measurement noise, $C_{t_{dd}}$ represents the temporal component of idealisation noise ('benchmark instability'), $C_{st_{dd}}$ represents the spatio-temporal component of idealisation noise, and $C_{st_{mm}}$ residual subsidence model imperfections. Note that benchmark idealisation noise and residual subsidence model imperfections do not necessarily always have to be present; their presence can vary per area as well as its magnitude.

Idealisation noise describes the representativeness of geodetic benchmarks for the subsidence signal of interest (subsidence due to hydrocarbon production). The stochastic model parameters of idealisation noise can be estimated from spatio-temporally correlated benchmark movements before the onset of gas production or in a 'stable' area nearby which is representative regarding benchmark type and shallow subsurface (see Section 5.3).

Subsidence model imperfections are minimized by exploring the full solution space of reservoir scenarios and geomechanical models, and by densely sampling the parameter space. However, as stated above, the subsidence models are representations of reality to a certain accuracy, and therefore unmodeled subsurface processes can cause residual spatio-temporally correlated deviations between geodetic data and subsidence models. If the residual model imperfections cannot be obtained by forward modelling of subsurface uncertainties, they could be estimated from the geodetic data via geodetic spatio-temporal analysis as soon as geodetic observations from multiple acquisition campaigns are available.

6.3.2 Application of the χ^2 distribution in the probability computation

The original Red-Flag methodology uses the multivariate normal distribution to assign a probability to the vector of differences between a model ensemble member and the geodetic data (TNO, 2017):

$$P\left(dd_r^{prior}\middle|dd\right) = \frac{P\left(dd_r^{prior}\right).P\left(dd\middle|dd_r^{prior}\right)}{\sum_{i=1}^{Ne} P(dd_i^{prior}).P\left(dd\middle|dd_i^{prior}\right)}$$

$$P\left(dd\middle|dd_r^{prior}\right) = exp[-\frac{\chi^2}{2}]$$

Please note that a factor 1/N is incorrectly present in Eq. 30 in (TNO, 2017), and has been removed here. If the probability of all prior ensemble members is equal (in other words, the reservoir scenarios and geomechanical models have equal a-priori probability), the probability of an ensemble member is the ratio between the probability of this specific ensemble member and the sum of the probabilities of the entire model ensemble.

Application of this methodology results in a 'collapse' of the ensemble into a few members having a high probability and all other members having a probability of 0. Even in the case of χ^2 values being close to each other, the weighing as determined by this probability definition causes an extreme discrimination between them, which is a counter intuitive result. This feature of particle filtering methods like Red-Flag is observed as well in many other disciplines in cases where a large set of independent variables is used (e.g. Snyder et al. 2008).

A solution to this problem is found in using the probability of the χ^2 test statistic *value,* which has a χ^2 distribution, to estimate the probability for a specific ensemble member. The χ^2 probability density function reads:

$$f_\nu(u) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} u^{\frac{\nu}{2}-1} e^{-u/2}$$

where u is the squared sum of weighted double-difference residuals (the test statistic χ^2), and ν is the number of residuals. For a large ν, the χ^2 distribution can be approximated by a normal distribution with expectation ν and variance 2ν:

$$P\left(dd \middle| dd_r^{prior}\right) \propto \exp\left\{-\frac{(\chi^2-\nu)^2}{4\nu}\right\}$$

Section 6.3.4 shows via simulations that using the χ^2 probability density function results in a band of accepted model members that comprises of all models that will be accepted when applying hypothesis testing of the individual models against the geodetic data, using a pre-defined level of significance α, which determines the critical value for the test statistic.

Before describing the simulation results, the analogy with geodetic adjustment and testing theory is demonstrated, and it will be shown that the χ^2 test statistic used in the particle filter has a χ^2 probability density function with degrees of freedom that is equal to the number of double-differences.

6.3.3 Analogy with geodetic adjustment and testing theory

The application of geodetic adjustment and testing theory [Teunissen (2000a) and Teunissen (2000b)] starts with the formulation of the functional model and the stochastic model. The functional model describes the relation between the observations and the unknown parameters, whereas the stochastic model describes the precision of the observations.

Starting from the vector of observations and its corresponding covariance matrix:

$$\underline{y} = \begin{bmatrix} \underline{dd}_{dd} \\ \underline{dd}_{mm} \end{bmatrix} \quad Q_y = \begin{bmatrix} C_{dd} & 0 \\ 0 & C_{st_{mm}} \end{bmatrix}$$

with the vector $\underline{y}$ comprising of the 'observed' double-differences from the geodetic data ($\underline{dd}_{dd}$) and the subsidence model ($\underline{dd}_{mm}$). The precision of the double-differences is described by the covariance matrices C_{dd} and $C_{st_{mm}}$ for the geodetic data and the subsidence model, respectively. The covariance matrix of the geodetic data contains the measurement precision of the geodetic technique (and additionally the idealisation precision), whereas the covariance matrix of the subsidence model represents the residual model imperfections. The stochastic models of the geodetic data and the subsidence model are here considered to be independent and hence uncorrelated.

The unknown parameters are defined as the double-difference surface displacements. The functional model can be formulated in terms of condition equations, stating that the expectation value of the misclosures (t) between the vectors $\underline{dd}_{dd}$ and $\underline{dd}_{mm}$ is 0:

$$B^T = \begin{bmatrix} -I & I \end{bmatrix} \quad \underline{t} = B^T \underline{y} \quad B^T E\{\underline{y}\} = 0$$

$$Q_t = B^T Q_y B = C_{dd} + C_{st_{mm}}$$

where B is the matrix containing the conditions on the observations, I is the identity matrix, and Q_t the covariance matrix of the differences ($\underline{t}$). Q_t can be obtained by the variance propagation law. Since the geodetic data and subsidence model double-differences are considered uncorrelated, Q_t is the sum of the covariance matrices of the geodetic data and the subsidence model.

In geodetic testing theory (hypothesis testing on the functional and stochastic model), the test statistic $\underline{T}_{q=\nu}$ is defined as:

$$\underline{T}_{q=\nu} = \underline{t}^T Q_t^{-1} \underline{t} = (\underline{dd}_{mm} - \underline{dd}_{dd})^T (C_{dd} + C_{st_{mm}})^{-1} (\underline{dd}_{mm} - \underline{dd}_{dd})$$

$$\underline{T}_{q=\nu} \quad \sim \quad \chi^2(\nu, 0)$$

where ν equals the redundancy in the estimation (assuming the matrix B has full rank), equal to the number of conditions (misclosures). The test statistic $\underline{T}_{q=\nu}$ has a central chi-squared distribution with ν degrees of freedom. Please note that $\chi^2(\nu, 0)$ denotes the χ^2 **distribution**, which should not be confused with the χ^2 **test statistic variable** as defined in Section 6.3.1:

$$\chi^2 = \left(dd_r^{prior} - dd\right)^T \left(C_{dd} + C_{st_{mm}}\right)^{-1} \left(dd_r^{prior} - dd\right)$$

Comparing the particle filter χ^2 test statistic to the $\underline{T}_{q=\nu}$ test statistic definition, it can be seen that they are similar, since dd_r^{prior} is a specific realisation (ensemble member) of subsidence model double-differences $\underline{dd}_{mm}$, and dd is the vector of geodetic data double-differences $\underline{dd}_{dd}$.

Instead of the test statistic $\underline{T}_{q=\nu}$, the Overall Model Test is generally used in geodesy to assess both the functional and stochastic model considering the observations. The Overall Model Test statistic is defined as:

$$\underline{\hat{\sigma}}^2 = \frac{\underline{T}_{q=\nu}}{\nu} = \frac{\underline{t}^T Q_t^{-1} \underline{t}}{\nu}$$

which has a Fisher distribution,

$$\underline{\hat{\sigma}}^2 \quad \sim \quad F(\nu, \infty, 0)$$

and an expectation value of 1 (if both functional and stochastic model are correct).

Note that the Overall Model Test is equal to the particle filter test statistic χ^2/N, where N is equal to the number of misclosures between the geodetic data and the subsidence model double-differences.

6.3.4 Effect of the probability computation modification

This Section shows via simulations the effect of the applied modification in the probability computation in Section 6.3.2 on the band of accepted model ensemble members (see Section 7.3.1 for a definition of 'accepted members') . The model ensemble in these simulations comprises of hypothetical length expansion models as a function of temperature. Unknowns are the initial length of an object and its expansion coefficient (this implies 2 model parameters). The measurements are uncorrelated in this example, and have a normal distribution.

The level of significance α has been chosen 0.05. With the level of significance, the critical value is determined for the χ^2 test statistic value, to either accept (test statistic is lower than critical value) or reject (test statistic is higher than the critical value) the model ensemble member.

With the simulations, the aim is to find answers to the following main questions:

1. Based on the χ^2 test statistic value, which models are accepted with a certain level of significance?
2. Is the band of these accepted models equal to the band of accepted model members from the (modified) particle filter?
3. What is the difference with respect to the collapse behaviour of the original particle filter (applying the multivariate normal distribution)?
4. What is the effect of a different number of observations?

Based on the simulations and visualisations, these questions can be answered as follows:

- Application of the χ^2 probability density distribution in the particle filter leads to a band of accepted model members that is similar to the band of accepted members based on comparison of the χ^2 test statistic values to the critical value (Figure 39).
- Application of the multivariate normal distribution leads to a narrower model confidence interval compared to that from the χ^2 probability density distribution. The discrepancy becomes larger when the number of observations increases (Figure 40, Figure 41). This effect is present for different numbers of model parameters.

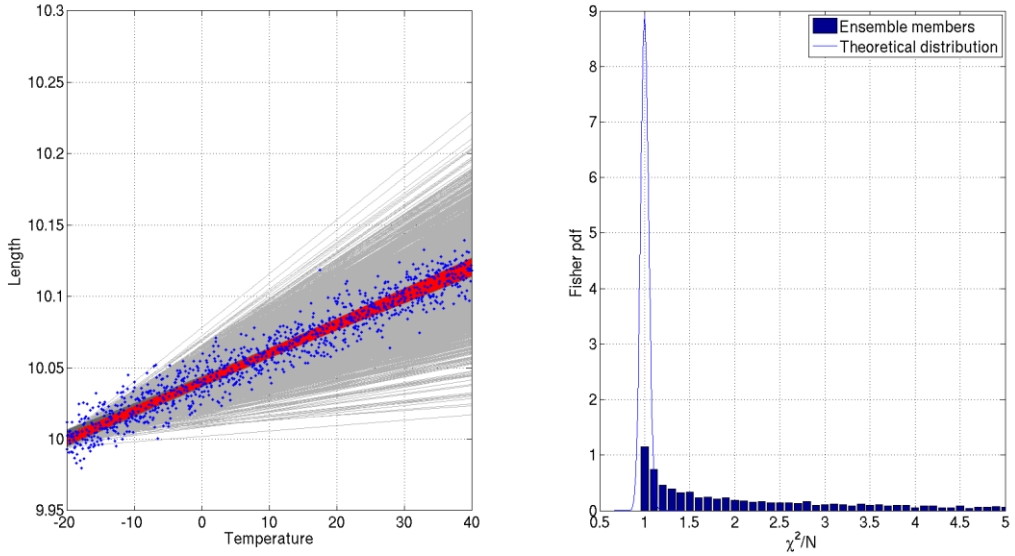


Figure 39. Results for the χ^2 probability density function. Left: in light grey the model ensemble, in dark grey (overlapping with red) the accepted members based on the χ^2 test statistic value (below the critical value), in red the accepted members from the particle filter, in blue the measurements. **The band of the accepted members from the particle filter is similar to the band of accepted members based on the χ^2 test statistic value compared to the critical value** (the red and the dark grey band fully overlap). Right: in blue bars the histogram of the Overall Model Test statistics (χ^2 test statistic value divided by number of differences between model and measurements (N)); the blue line is the theoretical Fisher distribution.

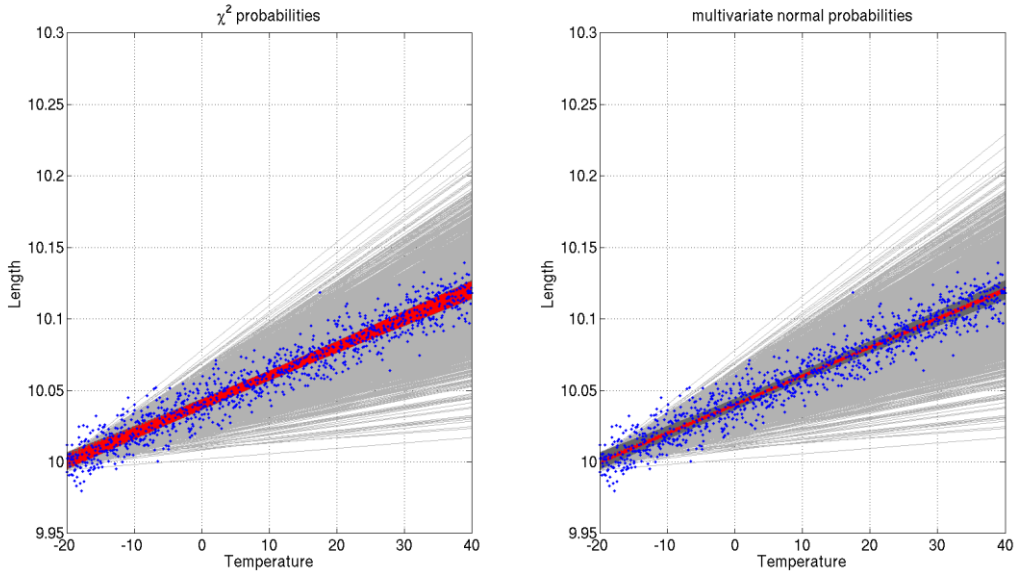


Figure 40. In light grey the model ensemble, in dark grey the accepted members based on the χ^2 test statistic value (below the critical value), in red the accepted members from the particle filter, in blue the measurements. Left: particle filter based on χ^2 probability density function (evaluating χ^2 test statistic value). Right: particle filter based on multivariate normal distribution (evaluating the vector of differences between model ensemble member and measurements). **It can be seen that the particle filter model confidence interval is similar to the interval of accepted members when applying the χ^2 probability density function, while it shows the collapse behaviour when applying the multivariate normal distribution** (red band is much narrower than the dark grey band).

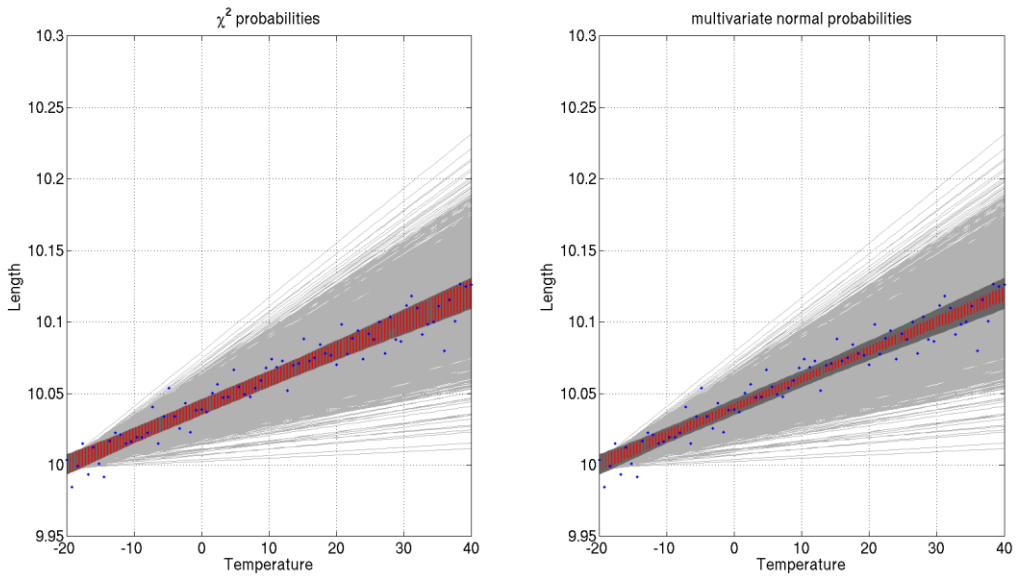


Figure 41. Same simulation as for Figure 40, but with **75** instead of **1000** observations. In light grey the model ensemble, in dark grey the accepted members based on the χ^2 test statistic value (below the critical value), in red the accepted members from the particle filter, in blue the measurements. Left: particle filter based on χ^2 probability density function (evaluating χ^2 test statistic value). Right: particle filter based on multivariate normal distribution (evaluating the vector of differences between model ensemble member and measurements). **The collapse behaviour when applying the multivariate normal distribution is less for lower number of measurements, but still is present. A different number of measurements has no effect when applying the χ^2 probability density function: The Red Flag model confidence interval is still similar to the interval of accepted members.**

7 The confrontation between geodetic data and subsidence models

This chapter describes the results of the confrontation of the selected reservoir scenarios and geomechanical members with the geodetic data for the Ameland case. Each model member will give a subsidence forecast, and from the confrontation with the geodetic data one can infer the likelihood of that specific member depending on the match with the data, as discussed in Chapter 6. Knowing the likelihood, a weighted average and a 95% confidence interval can be determined. These values provide a statistical description of the quality of model predictions. First, Section 7.1 describes the generation of the model ensemble, consisting of the selection of reservoir scenarios, geomechanical model parameters, and the resulting subsidence members, and a selection of the benchmarks used to generate double-differences. The construction of the covariance matrix used in the confrontation is described in Section 7.2. The direct results of the confrontation in terms of member probabilities are presented in Section 7.3, followed by the results in terms of subsidence at the benchmark locations in Section 7.4 and subsidence rates in the Pinkegat area in Section 7.5. The chapter is concluded by a presentation of additional model insights obtained from the confrontation in Section 7.6, and a discussion of the results of the application of the workflow over different time intervals in Section 7.7.

7.1 Generation of model ensemble

The confrontation requires an input ensemble of subsidence model members that comprises all properties and behaviours deemed possible with prior knowledge. The range of pressure depletion or reservoir scenarios that all honour the production and pressure data, yet display a wide range of pressure evolution in areas with no subsurface data, is described in Chapter 3. The prior knowledge and parameter selection for the compaction model is described in Section 4.2, and those for the translation of the compaction to subsidence through the overburden, i.e. the influence model are described in Chapter 0. The resulting pressure scenarios that are used in the workflow are summarised here, followed by a summary of the geomechanical parameters for the compaction and the influence function.

7.1.1 Reservoir scenarios

A wide variety of reservoir properties lead, as described in Chapter 3, to a set of 58 different reservoir scenarios that all honour the available gas and water production and pressure data. The set of scenarios covers significant diversity in terms of pressure depletion in the areas where little or no pressure data is available, most noticeable in the aquifers away from the well locations. Figure 42 Figure 43, and in more detail the figures in Appendix 1, shows the pressure depletion of all the reservoir scenarios contained in the ensemble at several locations across the area of interest.

Figure 42. Depletion of the reservoir and adjacent aquifers for the year 2016 for all 58 reservoir scenarios.

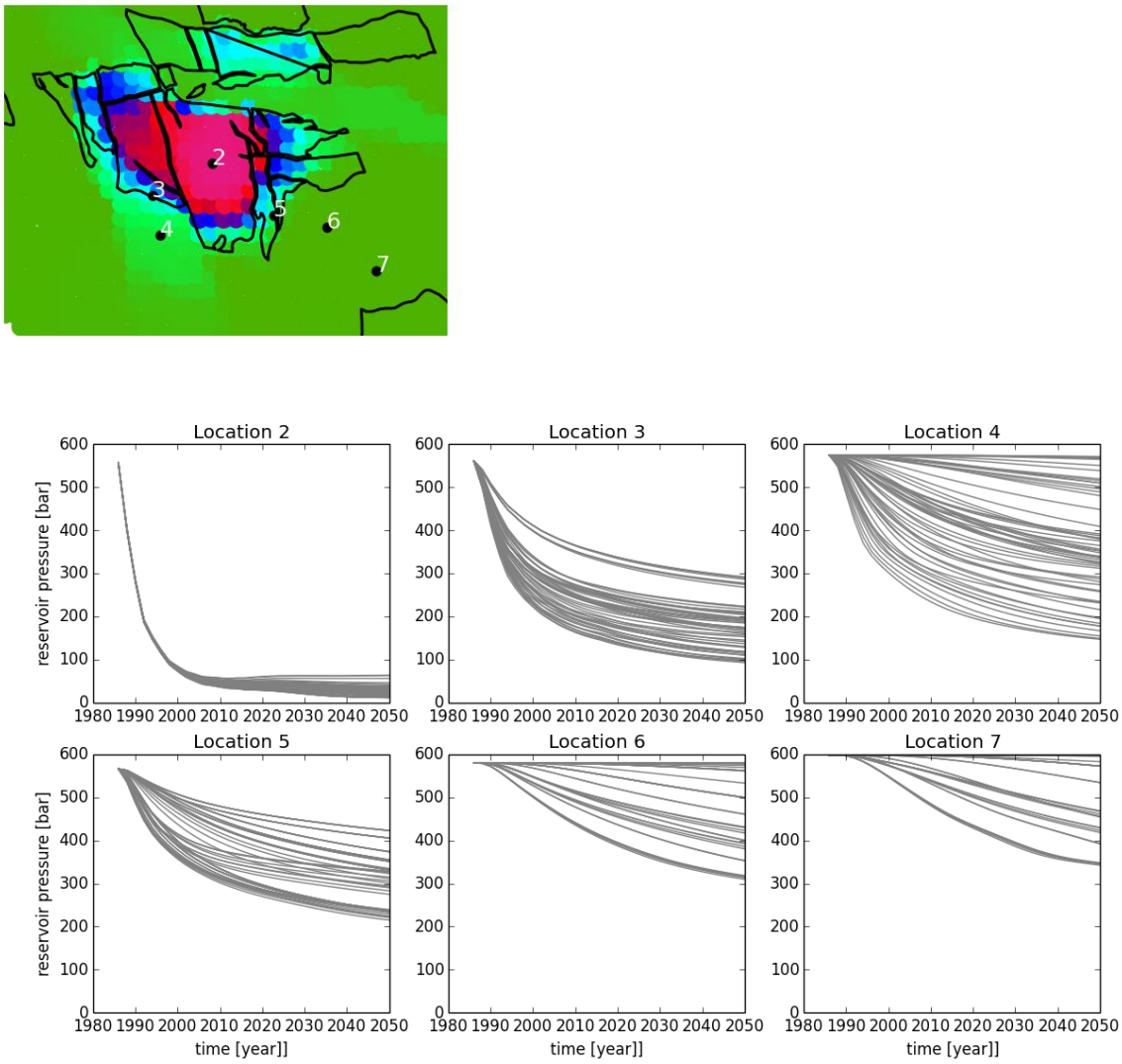


Figure 43. The bottom figure shows the pressure vs time for all selected reservoir scenarios for the six points shown in the top figure.

7.1.2 Selection of parameter values for the geomechanical model

The parameter set for the compaction model is based on coreplug experiment results and literature studies (Section 4.2.3). The salt layer above the reservoir is implemented in the Geertsma and van Opstal influence function using a moving rigid basement that iscalibrated to Geomec simulations (Section 4.3.2). The full parameter set used for both compaction parameters and the influence function parameters in the Monte Carlo simulation is presented in Table 18.

Table 18. Parameter values and ranges used in the Monte Carlo simulation.

Type	Parameter	Range / value
C_m - Porosity relation: $C_{m_{ref}}(Por) = C_4 + C_2 10^{C_3(Por - C_1)}$	C_1	0.02 - 0.17
	C_2	0.01 - 2.5
	C_3/C_4	*
Salt (moving rigid basement)	$c/k(0)$	0.75 - 0.85
	$d(c/k)$	0.15 - 0.25
	τ_{salt} [year]	0.3 - 100.0
RTCiM parameters	$C_{m_d}/C_{m_{ref}}$	0.2 - 0.7
	b	0.01 - 0.03
Overburden	Poisson's ratio	0.2

* Parameters C_3 and C_4 are derived based on the values of C_1 and C_2 such that the used C_m-Porosity relations cover the full range of observations from core plug experiments (grey area of Figure 14).

7.1.3 Subsidence members

The calculation of one subsidence member requires one reservoir scenario (out of 58 possible scenarios), together with a set of parameters for both the compaction model and influence function, provided by the Monte Carlo process. To arrive at the full prior ensemble of subsidence members, combinations are generated with reservoir pressure scenarios and geomechanical parameters. All the available pressure scenarios have the same (prior) likelihood. Also, the probability distribution for the geomechanical parameters is unknown and therefore assumed to be uniformly distributed.

Since the solution space is high-dimensional, care must be taken to adequately sample the full space. One way to ensure all corners of the multidimensional space are probed is by using a Latin Hyper Cube (LHC) method for generating sets of parameters (McCay et al., 2000). The LHC method divides the ranges for all parameters into equal subdivisions, which are then populated randomly.

It is not practical in the workflow to randomly select a pressure scenario for each case, therefore a set of geomechanical model parameters, selected via the LHC method, is run in combination with one pressure scenario. Subsequently the same is done for the next pressure scenario. The number of geomechanical model parameter sets to be run for each pressure scenario is difficult to define up front. It depends not only on the number of parameters but also on their correlation and their impact on the outcome. One wants to maximize the number of parameter sets such that the parameter space is adequately sampled, thereby increasing the likelihood that the best combination of parameters is obtained, whilst also generating a dense range of subsidence predictions, which aids the statistics in terms of subsidence from the output. On the other hand, one cannot produce an endless amount of subsidence members from a practical point of view, hence an optimum should be found.

The current study examines the adequate number of members, by creating a large set of randomly generated subsidence members based on one reservoir scenario, and looking at the resulting χ^2 distribution. Figure 44 shows the lowest χ^2 values for subsidence ensembles with a different number of members. After about 10,000-15,000 members, the lowest five χ^2 seem to have converged, indicating that adding many more members would not result in finding significantly better members in the solution space. Therefore, it was decided that 20,000 sets of geomechanical models for each

reservoir scenario would adequately sample the solution space, giving a total of 1,160,000 members in the final ensemble.

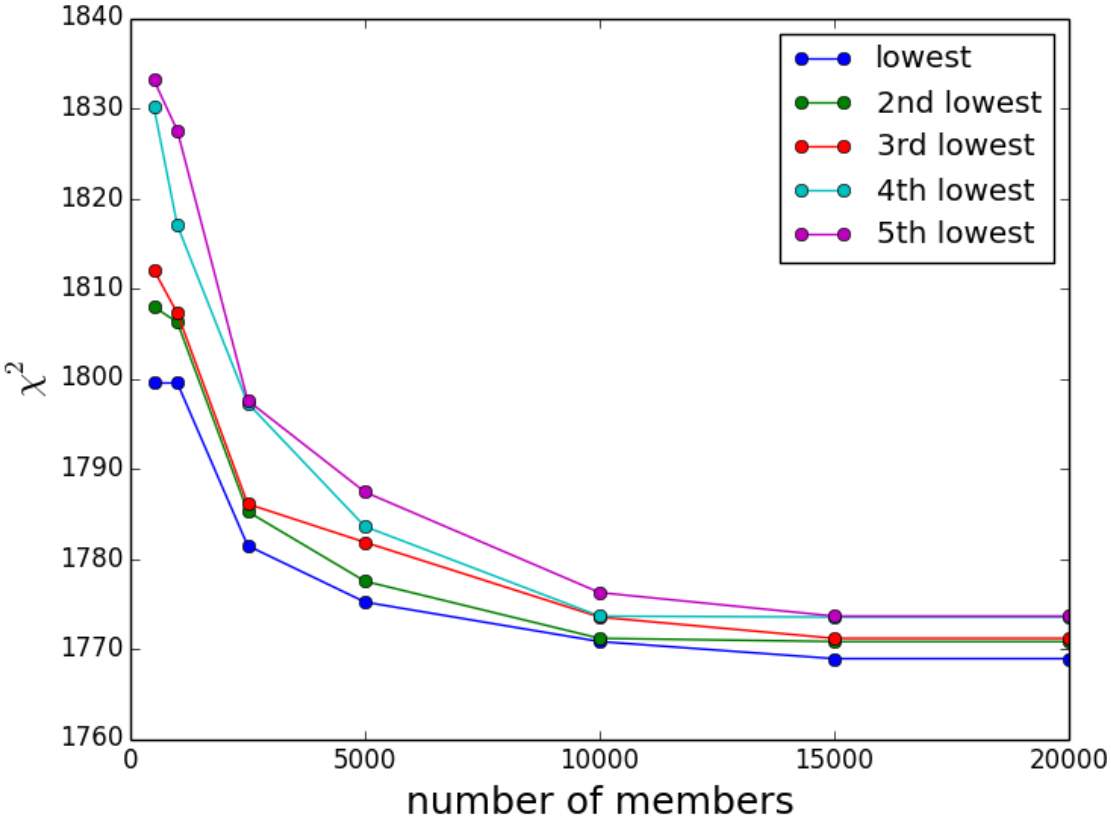


Figure 44. Lowest χ^2 values as a function of the number of subsidence members, all based on the same reservoir scenario. A total of 20,000 members was generated, from which smaller subsets were randomly sampled. For all these subsets, the lowest five χ^2 values are plotted, showing that the distribution does not change much after approximately 10,000-15,000 members.

7.1.4 Selection of benchmarks

The benchmarks selected for the confrontation between modelled and measured double-differences is based on a polygon which covers the influence area of the Ameland subsidence bowl. In total, 1417 double differences have been formed by the CUPiDO tool from the selected geodetic data. The same double differences are evaluated for all subsidence model members and used in the confrontation workflow. The blue dots in Figure 45 mark the benchmark locations at which these double differences are taken.

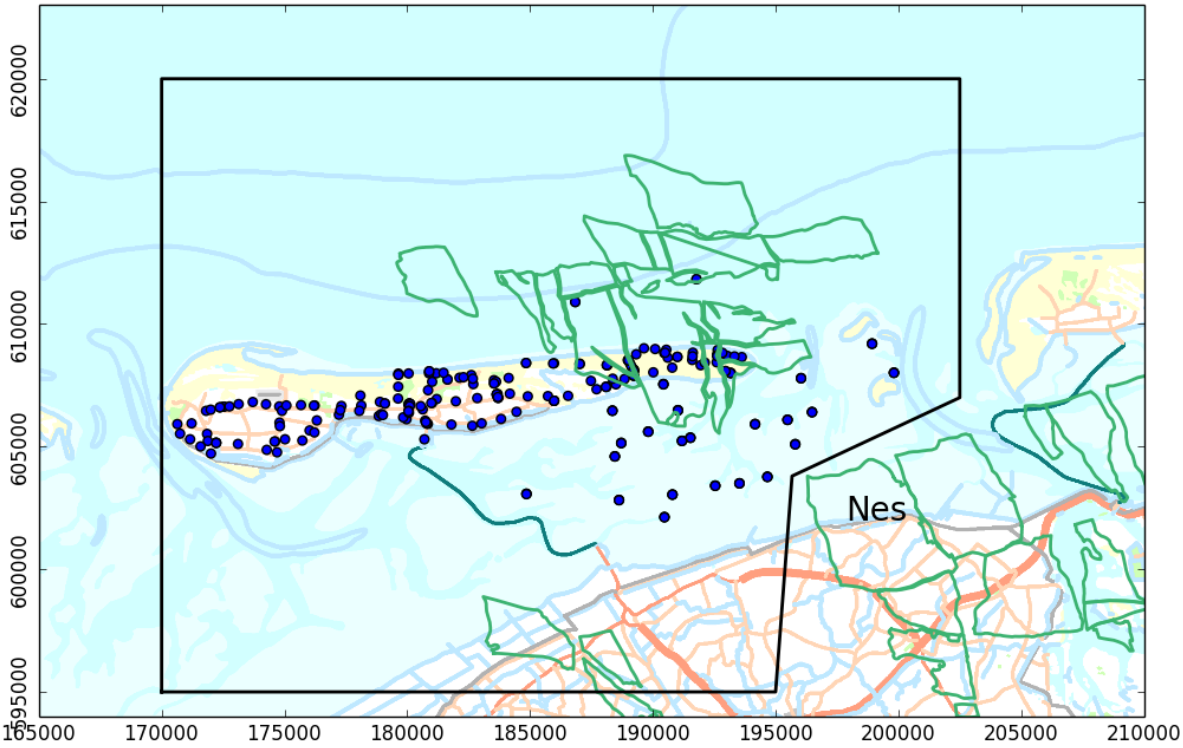


Figure 45. Selected benchmarks (blue dots) within the areal selection polygon (black outline).

7.2 Covariance matrix in the particle filter

In the particle filter test statistic for the confrontation between geodetic observations and subsidence model members, the spatio-temporal component of benchmark idealisation noise $C_{st_{dd}}$ and residual subsidence model imperfections $C_{st_{mm}}$ are stochastically modelled in the covariance matrix C_{st}:

$$C_{st} = C_{st_{dd}} + C_{st_{mm}}$$

The covariance matrix C_{st} has a direct effect on the subsidence model confidence interval resulting from the confrontation using the particle filter, and therefore should be well-considered. As is displayed in Figure 38, spatio-temporally correlated deviations between geodetic observations and subsidence models can have their origin both in quadrant II (spatio-temporal component of idealisation noise) and quadrant IV (residual subsidence model imperfections).

For the benchmark idealisation noise, the spatio-temporal component of the idealisation noise can be estimated either from geodetic observations before the onset of gas production, or from the geodetic observations in a neighbouring area that is representative. In the gas production area on the Ameland island there are not enough measurements before 1986 to estimate the spatio-temporal component of the idealisation noise from. Attempts have been carried out to estimate spatio-temporally correlated noise on the western part of the Ameland island via geodetic spatio-temporal analysis (Methode Houtenbos; TCBB, 2009), but no consistent estimates could be obtained. This can be due to the limited spatial extent of the area or the limited number of observations. Hence, it has been chosen to model idealisation precision with parameters obtained in the LTS-I

study area in Northern Friesland. This is parameter set ON_C from Section 5.3, which includes both the temporal component ($C_{t_{dd}}$) and the spatio-temporal component ($C_{st_{dd}}$) of benchmark idealisation noise.

To evaluate the characteristics of the idealisation noise parameters from parameter set ON_C, the 1417 geodetic data double-differences in the LTS-II application for Ameland are grouped with respect to distance and time, and the residuals with respect to the best fit subsidence model member are compared to the standard deviation from ON_C (see Figure 47 and Figure 48). The best fit subsidence model member in this context is the ensemble member with the lowest χ^2 test statistic value, when considering measurement noise ($C_{e_{dd}}$)and benchmark instability noise ($C_{t_{dd}}$) only (see Figure 46). From Figure 46, Figure 47, and Figure 48, the following can be deduced:

- Time correlated noise is present, but for some benchmark groups model ON_C overestimates time correlated noise, while for other groups it may be slightly underestimated.
- Distance correlated noise seems not present for all benchmark groups. Especially for the benchmarks groups in subplots with time between campaigns larger than 20 years, model ON_C overestimates distance correlated noise. The majority of double differences with long time spans involve benchmarks on the western part of Ameland (west from Nes).
- Besides benchmark idealisation noise, it is likely that residual model imperfections are present as well in the gas production areas (see for example the difference of ~3 cm between subsidence model and geodetic data double-differences in Figure 46 at the eastern part of the island).

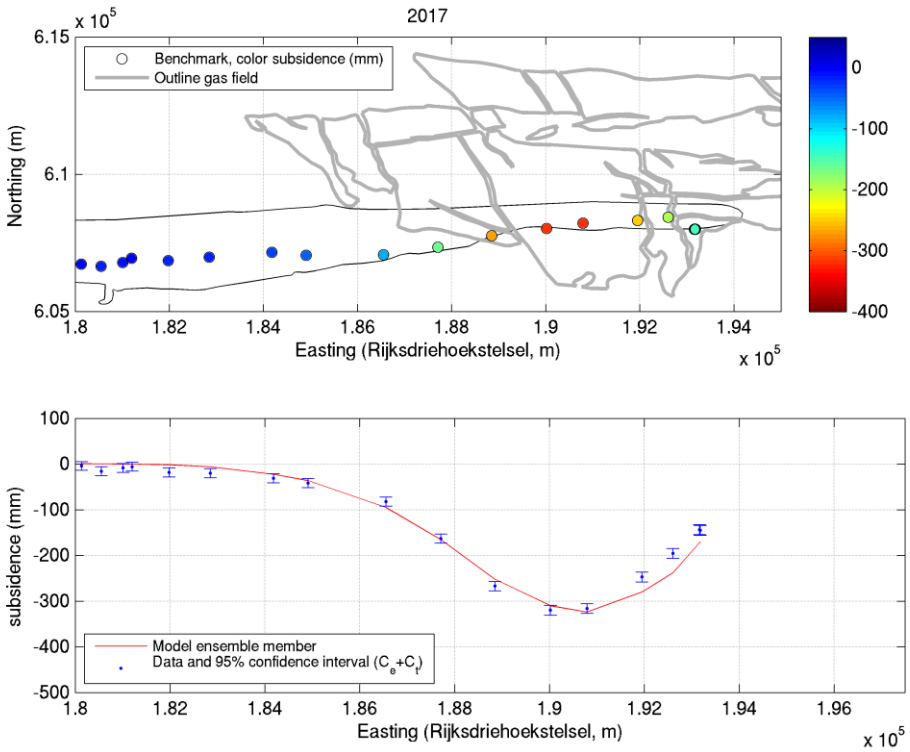


Figure 46. Southern profile showing benchmark displacements between 1986 and 2017 with respect to benchmark 000A2592 (Nes). The error bars on the benchmarks consist of measurement and benchmark instability noise ($C_{e_{dd}}$+ $C_{t_{dd}}$). The red line is the model ensemble member with the best fit to the geodetic data, where the goodness of fit has been evaluated based on the noise components $C_{e_{dd}}$+ $C_{t_{dd}}$.

Summarizing, while the model ON_C seems to overestimate idealisation noise for multiple benchmarks on Ameland, it does account for (part of) the residual model imperfections (it covers both elements of the covariance matrix C_{st}). Hence model ON_C is an approximate stochastic model for the LTS-II application. Refining the spatio-temporal component by means of a geodetic spatio-temporal analysis may lead to a stochastic model that is even closer to reality. However, this would have the disadvantage that the geodetic data is then used twice in the process and that it can only be done 'a-posteriori'.

However, the disadvantage has to be acknowledged that the stochastic modelling of spatio-temporally correlated deviations according to ON_C does not accommodate individual benchmark properties and does not account for local variations in subsidence model uncertainty. In spite of these residual imperfections, the chosen stochastic model allows for a reasonable amount of acceptable subsidence model members, and hence is able to generate a weighted average from subsidence model members that are physically possible, including a confidence interval (see Figure 49).

Since the test statistic values in the confrontation workflow are close to the expectation value (χ^2/N values 1.23 and higher, see Section 7.3), model ON_C has been used in its original form to reflect spatio-temporally correlated deviations in the covariance matrix. For the application of the particle filter in other areas, the stochastic modelling of spatio-temporal deviations should always be carefully evaluated.

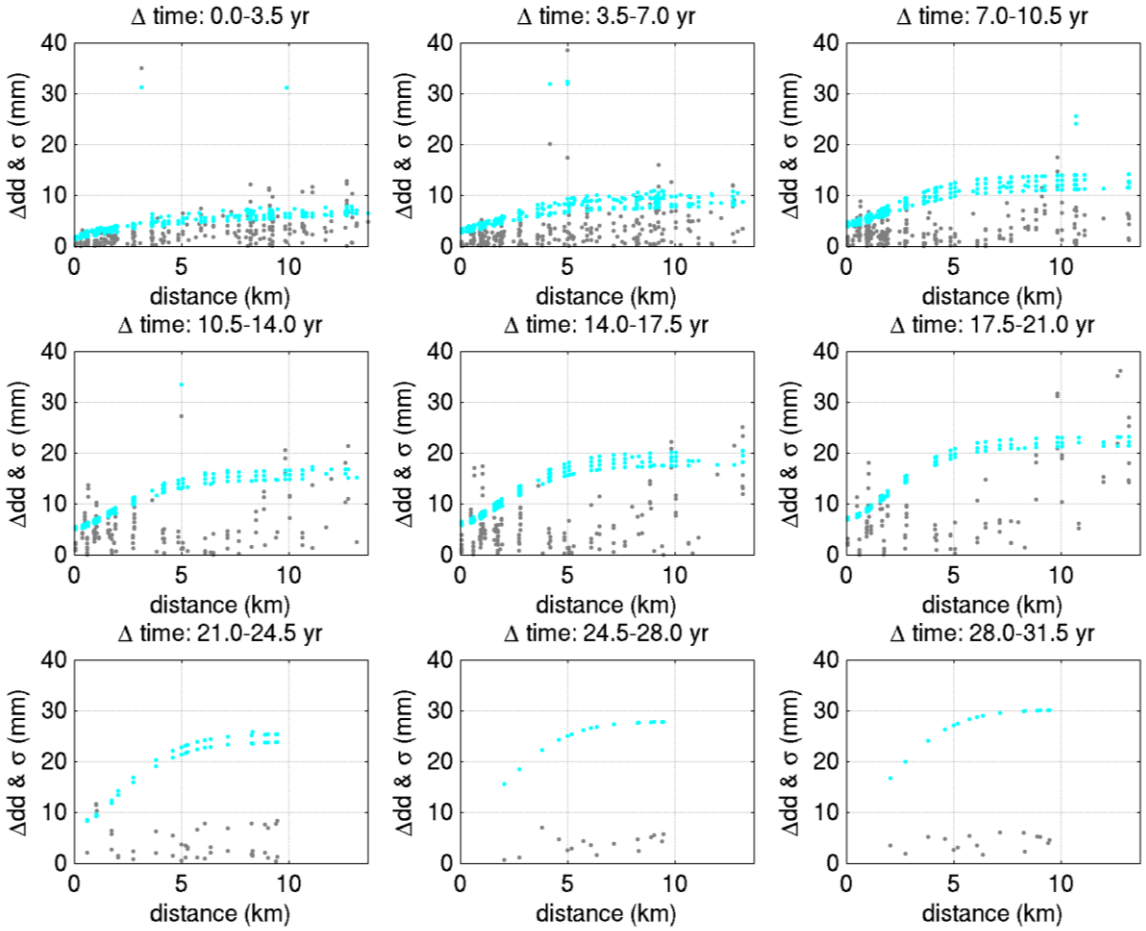


Figure 47. In grey the double-difference residuals between the geodetic data and best fit subsidence model based on measurement and benchmark instability noise only ($C_{e_{dd}}$+ $C_{t_{dd}}$). In light blue the two time the standard deviation (2-sigma) from the full covariance matrix based on model ON_C ($C_{e_{dd}}$+ $C_{t_{dd}}$+ $C_{st_{dd}}$). The double-differences are grouped based on distance between benchmarks, and the double-difference residuals and standard deviations are shown as a function of time.

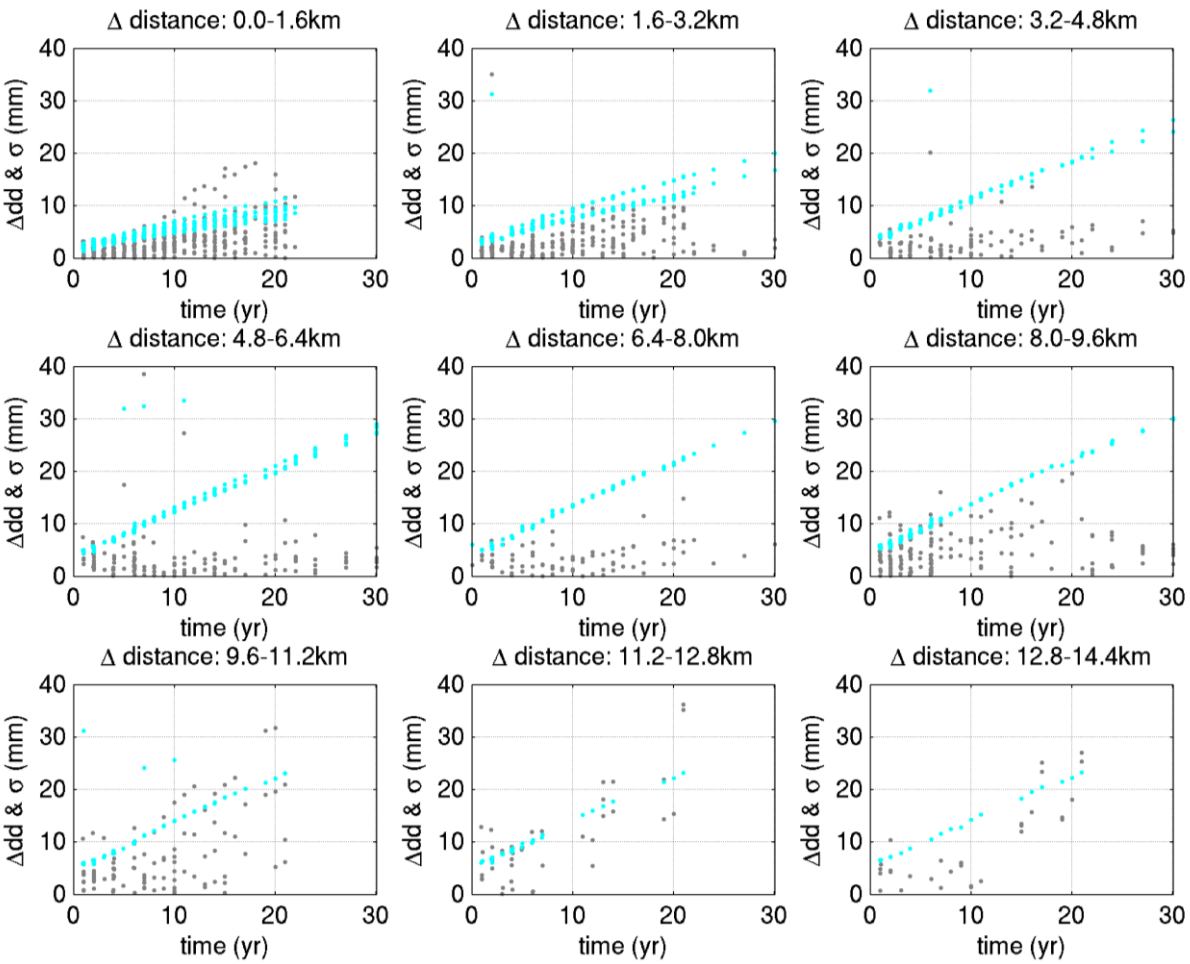


Figure 48. In grey the double-difference residuals between the geodetic data and best fit subsidence model based on measurement and benchmark instability noise only ($C_{e_{dd}}$+ $C_{t_{dd}}$). In light blue two times the standard deviation (2-sigma) from the full covariance matrix based on model ON_C ($C_{e_{dd}}$+ $C_{t_{dd}}$+ $C_{st_{dd}}$). The double-differences are grouped based on time between measurement campaigns, and the double-difference residuals and standard deviations are shown as a function of distance.

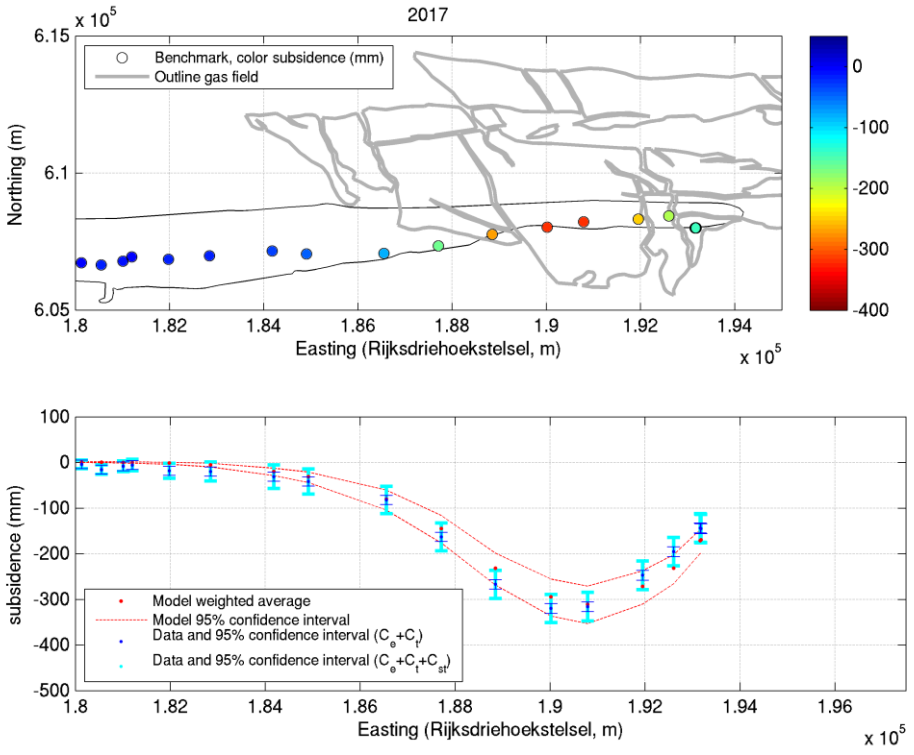


Figure 49. Southern profile showing benchmark displacements between 1986 and 2017 with respect to benchmark 000A2592 (Nes). In red: the model weighted average from the particle filter with 95% confidence interval. In blue the geodetic data double-differences with 95% confidence interval, with (light blue) and without (dark blue) stochastically modelled spatio-temporally correlated deviations (C_{st}) according to ON_C.

To perform a crosscheck on the results of the particle filter, it is validated if the fit between geodetic data and subsidence model weighted averages matches the confidence intervals. While the validation takes into account the covariance structure of both the geodetic data and the subsidence models, the correlations between them are neglected. This is a simplification, because these correlations are not zero. They are dependent, since the weights for computing the weighted averages of the subsidence models are determined based on the fit of the subsidence model members to the geodetic data.

The one-dimensional w-test statistics (Teunissen, 2000b) are used for the evaluation of the misclosures between the geodetic data double-differences $\underline{dd}_{dd}$ and the weighted average subsidence model double-differences $\underline{dd}_{mm}$:

$$\underline{t} = \underline{dd}_{dd} - \underline{dd}_{mm},$$

with covariance matrix of the misclosures Q_t:

$$Q_t = C_{e_{dd}} + C_{t_{dd}} + C_{mm}$$

where the covariance matrix of geodetic data double-differences contains measurement ($C_{e_{dd}}$) and benchmark instability noise ($C_{t_{dd}}$), and C_{mm} is the a-posteriori subsidence model covariance matrix from the particle filter as defined in Eq. 35 in TNO, 2017. Figure 50 shows the structure of the geodetic data and model covariance matrices that are used in the crosscheck. The covariance matrix C_{st}, which models spatio-temporally correlated deviations between geodetic data and subsidence model members has deliberately not been included, to avoid that residual model imperfections are addressed twice in the crosscheck.

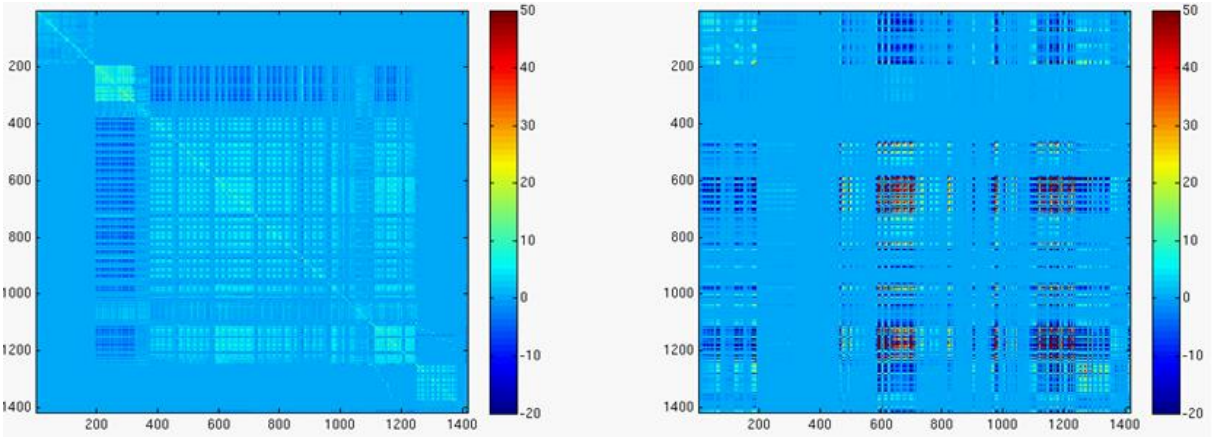


Figure 50. Structure of the covariance matrices $C_{e_{dd}}$+ $C_{t_{dd}}$ (left figure: geodetic data, measurement and benchmark instability noise), and C_{mm} (right figure: a-posteriori model covariance matrix) for the 1417 double-differences used in the confrontation. (Co)variances are in mm².

The one-dimensional w-test statistics provide a test statistic per misclosure between geodetic data and model weighted average double-differences, taking the full covariance structure into account:

$$\underline{w} = \frac{c_t^T Q_t^{-1} \underline{t}}{\sqrt{c_t^T Q_t^{-1} c_t}},$$

where c_t is a vector with 1 at the double-difference to be tested and zeros elsewhere.

Figure 51 shows the histogram of the w-test statistics for the 1417 double-differences used in the particle filter for Ameland. With 6.3% of the double-differences exceeding the critical values for the 95% confidence interval, it can be concluded that the particle filter has been able to generate a model weighted average with a confidence interval that captures the geodetic data double-differences. A complementary (simplified) validation using w-test statistics has been carried out by replacing the a-posteriori subsidence model covariance matrix (C_{mm}) with the a-priori covariance matrix for spatio-temporal deviations (C_{st}). In this situation, 5.2% of the double-difference misclosures where rejected based on the w-test statistics.

It can be noted that for some double-differences between benchmarks (see Appendix 2), the time series lies on the upper or lower bound of the model confidence interval. Note that the subsequent double-differences are correlated, and that the correlation structure is taken into account in the w-test statistics. Subsequent double-differences are therefore not 'independently' accepted or rejected. For more background on the effect of correlation on confidence intervals, see Appendix 3.

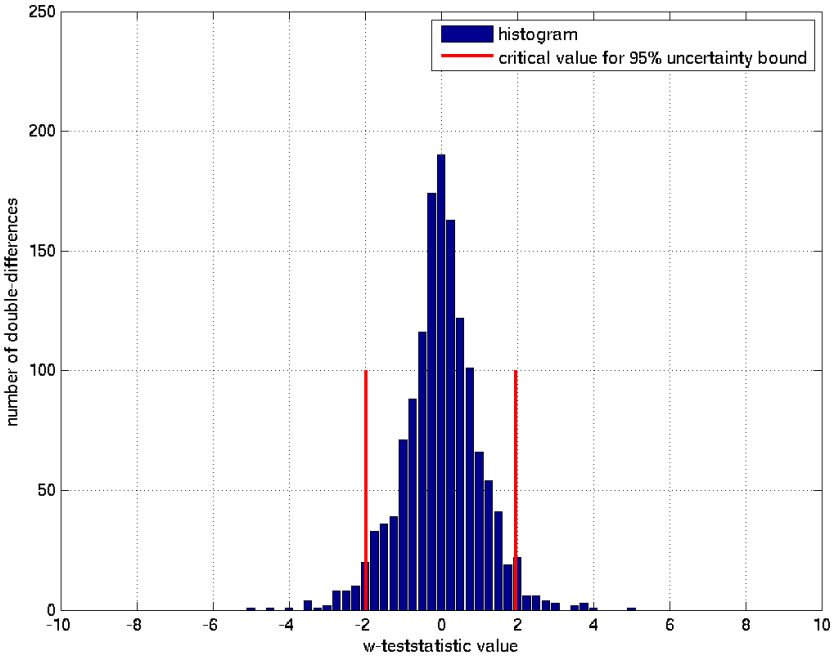


Figure 51. W-test statistics for the 1417 double-differences between geodetic data and weighted averages from the particle filter. The red lines show the critical values for the 95% error bound. 6.3% of the double-difference misclosures exceed the critical values.

7.3 Confrontation between geodetic data and model ensemble members

The results of the confrontation between the complete set of geodetic data from the selection of benchmarks with the full model ensemble is presented in this section. First, the definition of the terms 'confidence interval' and 'accepted members' as used in the remainder of this chapter is given, followed by a presentation and discussion of the results from the confrontation.

7.3.1 Use of terms 'confidence interval' and 'accepted members'

As the use of terms like 'confidence interval' and 'number of accepted members' can be rather ambiguous, the definition of the terms as used in the current study is provided here.

A confidence interval is mentioned or depicted for parameters such as the subsidence, subsidence rate, and posterior model parameter values. Only two-sided, 95% confidence intervals are used, which are determined from the cumulative member probability as illustrated in Figure 52. After the confrontation, each member has been assigned a probability. First, the values of the particular *variable*, for which the confidence interval is to be determined, from all the members are sorted from low to high values. These variable values can be, for example, the resulting subsidence at a specific location, as determined by all the members, or the compressibility values used by all the members. The member probability is then summed to obtain a cumulative probability curve as a function of the particular variable. The bounds of the two-sided 95% confidence interval are determined by finding the variable value at a cumulative probability of 2.5% (lower bound) and at 97.5% (upper bound). Using this definition means that there is a 2.5% probability to find a variable value below the lower bound, and a 2.5% probability to find it above the upper bound.

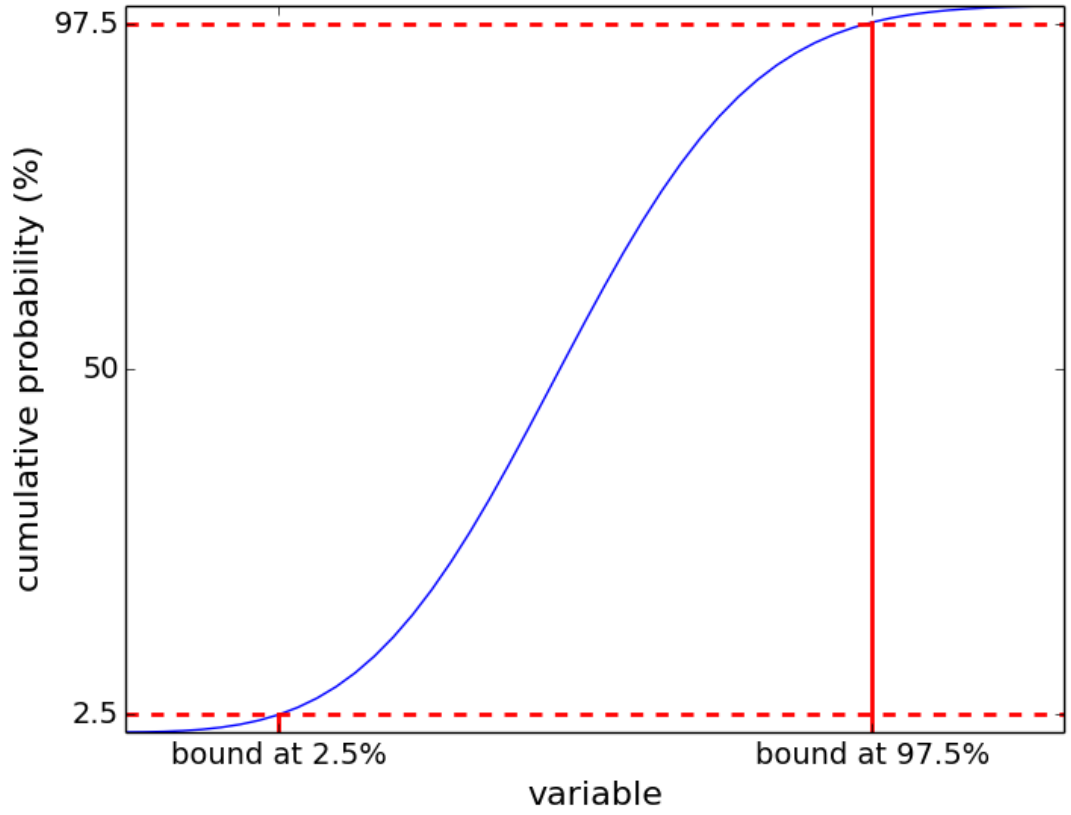


Figure 52. Illustration of the definition of a confidence interval as used in this report.

Although not part of the confrontation workflow, a member acceptance test is performed for illustrative purposes of the results in some figures and discussion. In an acceptance test, one determines *how many members* would be accepted based on a certain cumulative probability cut-off. The number of accepted members itself is not used, but can be viewed as a metric or indication of how many members give a considerable contribution to a weighted average. It gives insight especially when comparing these numbers, to get a sense of the contribution of different reservoir scenarios, or when comparing the results from different data sets (for example, the results from confrontations performed over different time intervals). Figure 53 illustrates how the number of accepted members is determined. Instead of first sorting a variable value from low to high, the members are sorted from highest to lowest assigned probability. The member probability is then summed to obtain a cumulative probability curve as a function of the number of members, starting with the most likely member. The number of accepted members is determined by finding the number of members required to get a cumulative probability of 95% (only a 95% acceptance interval is used in the report). Using this definition means that the best *n* members fall within the acceptance level, and their combined probability adds up to 95%. The members outside the accepted interval are the members with the lowest probabilities assigned to them, which add up to 5%.

Note that acceptance tests are only performed in this study for illustrative purposes. Actual member selection and rejection is not performed at any stage; all members in the ensemble contribute to the weighted average and are taken into account when determining confidence intervals.

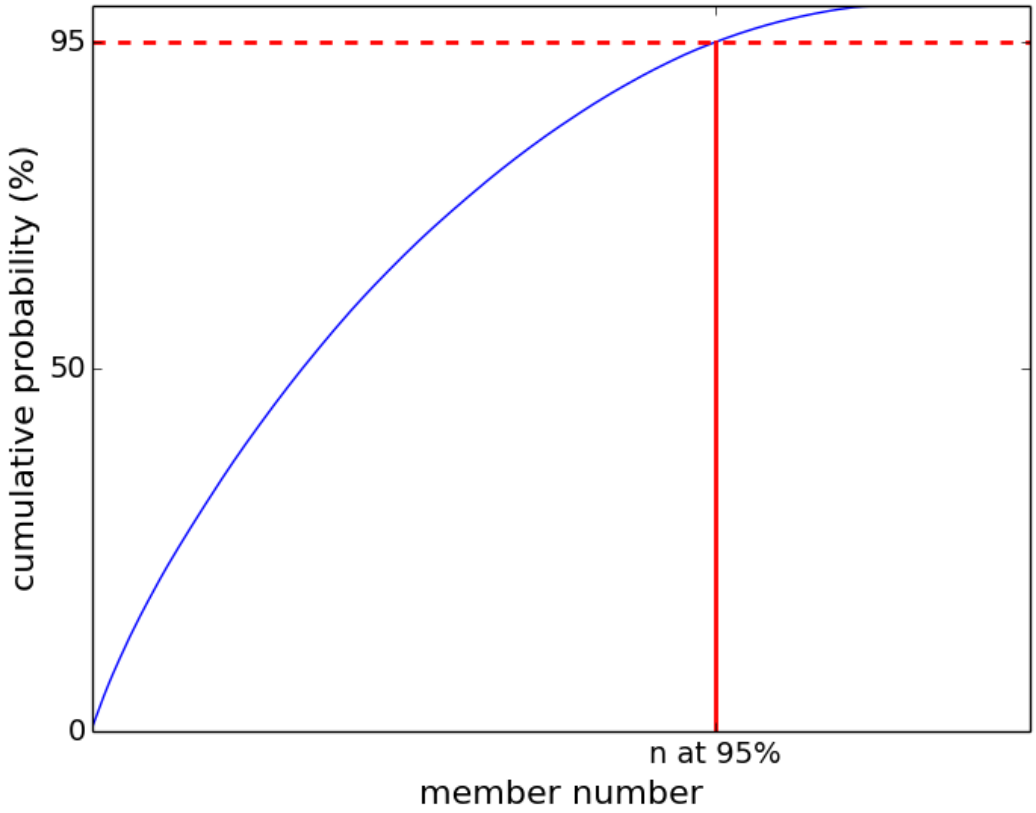


Figure 53. Illustration of the determination of the number of accepted members as used in this report.

7.3.2 Results of the confrontation

Each member of the full ensemble is confronted with the geodetic data, which, in combination with the covariance matrix to incorporate the uncertainties and correlations, assigns a χ^2 value (or χ^2/N, where N is the number of double-differences) to each member. The lower the value of χ^2/N, the better the fit of the particular member to the data. Figure 54 below shows the resulting χ^2/N values (where in this case N = 1417 for the confrontation with the geodetic data in the time interval 1986 - 2017) for all the 1,160,000 members that make up the full ensemble. The ensemble is generated based on 58 different reservoir scenarios. For each reservoir scenario, 20,000 subsidence members were generated. The first 20,000 members in the left plot of Figure 54 therefore all used the same reservoir scenario while varying the geomechanical parameters; the second set of 20,000 uses the next reservoir scenario, etc.

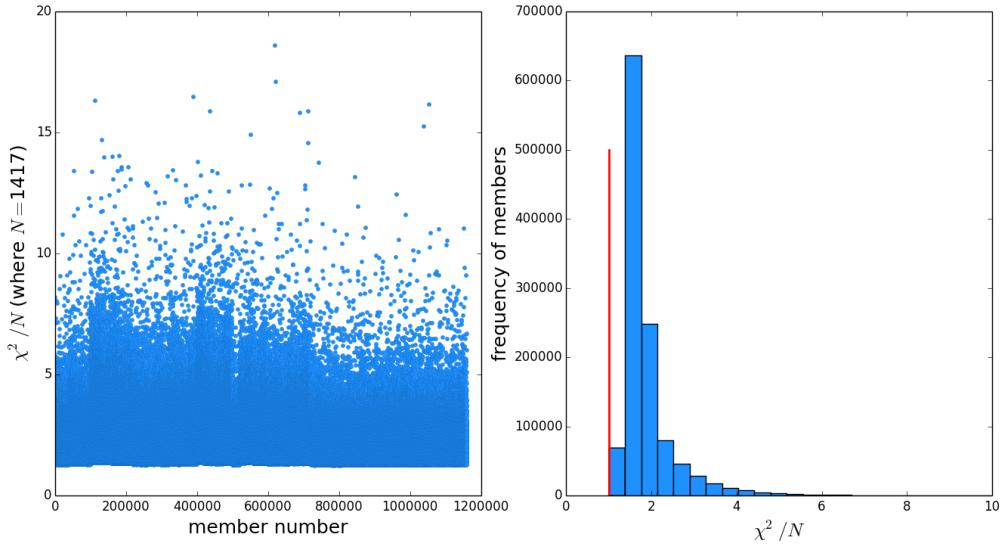


Figure 54. Left: Assigned χ^2/N values for all members, from the confrontation of the members with the geodetic data available for the time period 1986 - 2017. For the case here N is 1417 (= number of double differences). Right: histogram of the χ^2/N values. The red line marks $\chi^2/N=1$.

Figure 55 shows the result of converting the χ^2 values assigned to each member to a weight or normalized probability, using the method described in Section 6.3 . The members with the highest χ^2 values get the lowest probability, and only those with a very good match to the geodetic data obtain a probability that is significant and visible on these scales. The figure demonstrates the significant reduction in members with an appreciable probability after confronting the members with the data, reducing the ensemble of almost 1.2M members to only a few thousand with a significant probability.

The plot on the right-hand side of Figure 55 shows the cumulative probability, after sorting the members from highest to lowest probability. Note that the x-axis is limited here to only show the contribution of the 10,000 most likely members to the cumulative probability. If one applies a 95% acceptance level to the set of members, a total 1,870 out of 1,160,000 would pass, as indicated by the red dashed lines.

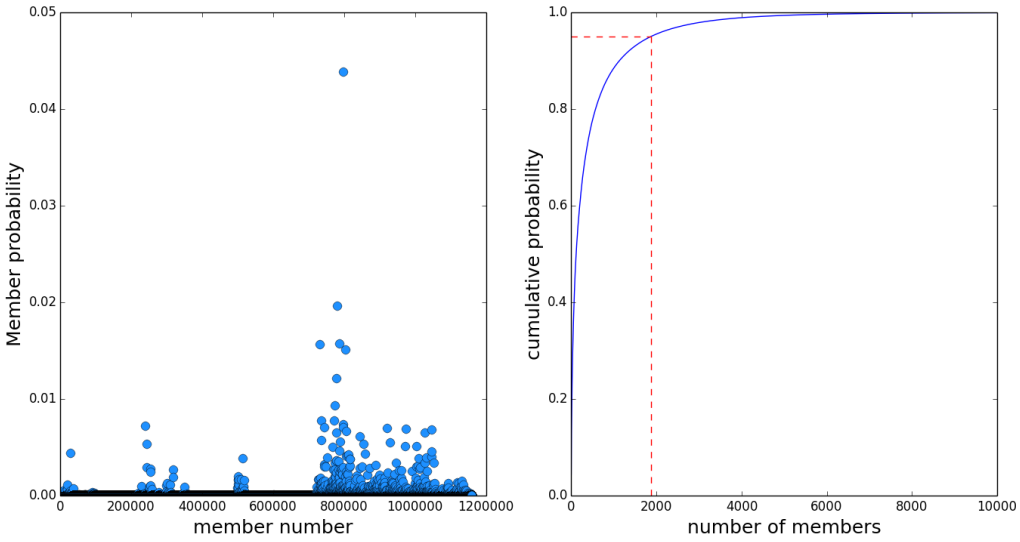


Figure 55. Left: Weight or normalised probability for each member in the ensemble, using the assigned χ^2 values from the confrontation with the geodetic data. Right: cumulative probability, after sorting the members from highest to lowest probability. When applying a 95% acceptance level, 1870 members will be accepted (indicated by the red dashed lines).

7.4 Subsidence

After the confrontation with the geodetic data, a weight or relative probability has been assigned to each of the members, discussed in the previous sections, which is used to determine an expectation case of the subsidence (defined as the weighted average of the subsidence models from all the ensemble members), together with 95% confidence intervals. Profiles and time series visualisations of this weighted average and its 95% confidence interval including the measured subsidence can be found in Appendix 2. The model weighted average and the confidence intervals are shown up to 2040 for the levelling double-differences.

The confrontation between geodetic data and subsidence models has been performed at double-difference level, with various references in space and time. In the LTS-II study, the application of the CUPiDO tool (Section 6.2) has resulted in 1417 non-redundant geodetic data double-differences with 6 difference references in space, and multiple references in time. For visualisation purposes (time series and profiles), they have been transformed to a common reference in time and space.

The common reference in space is:

- The underground benchmark 000A2592 in Nes for levelling campaigns.
- The GPS reference network for the GPS campaigns.

The common reference in time is:

- For the levelling time series and profile visualisations: 1986.
- For the GPS campaigns (time series only): earliest campaign per GPS benchmark.

To obtain the confidence interval of the transformed double-differences, the covariance matrix is transformed as well by application of the propagation law:

$$C_{dd}{}^{t1,p1} = S\, C_{dd}{}^{t2,p2} S^T ,$$

where S is the transformation matrix (and S^T its transposed), which specifies the linear combinations to perform the transformation to another reference in time and space (here from double-differences with respect to time t2 and benchmark p2 to double-differences with respect to time t1 and benchmark p1).

Appendix 2 contains time series and profile visualisations that provide insight in the match between geodetic data and the subsidence models, including the confidence intervals:

1. Levelling profile plots along a northern and a southern trajectory on the island.
2. The full time series for all levelling benchmarks that have been measured in both 1986 and 2017.
3. Time series for the GPS benchmarks with respect to the earliest measurement campaign of the benchmark.

All time series and profile visualisations contain the following information – see the example in Figure 56:

- **Blue**: geodetic data double-differences with 95% confidence interval derived from the covariance matrix considering measurement noise and the temporal component of idealisation noise ('benchmark instability'): $C_{e_{dd}} + C_{t_{dd}}$.
- **Red dots**: model weighted average computed using the probabilities determined from the particle filter (see Section 6.3).
- **Dashed red lines**: model 95% confidence interval based on the probability profiles per double-difference in space and time. The confidence interval is evaluated for benchmark locations or observation epochs, respectively, and linearly interpolated in between.

The covariance matrix C_{st} , which represents spatio-temporally correlated deviations between geodetic data and subsidence model double-differences, has been used in the confrontation workflow. The workflow results in the subsidence model weighted average including the confidence interval, and is described in Section 7.2. In the LTS-II study, C_{st} is modelled with parameters obtained in an area in Friesland that is assumed not to be affected by subsidence due to gas extraction. It is shown in Section 7.2 that the application of C_{st} on Ameland overestimates the spatio-temporal component of idealisation noise for multiple benchmarks, and that it likely also accounts for localized residual model imperfections. Since the contribution of C_{st} cannot be differentiated, the confidence intervals in the visualisations contain the only the components that can certainly be attributed to the geodetic data (i. e., measurement and benchmark instability noise; $C_{e_{dd}} + C_{t_{dd}}$).

In the visualisations, detected outliers (Section 5.4) have been excluded. An exception has been made for benchmark 002D0079, where the full, consecutive time series is displayed, even though a disturbance was detected in 1992. In the confrontation, however, no temporal differences are formed across the disturbance.

The geodetic data double-differences set in the confrontation with subsidence model members contains various references in space and time. To ease the visualisation, the levelling time series and profiles have been converted to the spatial reference benchmark 000A2592 and the temporal reference 1986. This conversion involves a linear recombination of double-differences along with a rigorous propagation of the full covariance matrix. Figure 57 shows a time series example for benchmark 002D0069 with spatial reference 002D0070. The geodetic data time series shows relative uplift, which is in agreement with the subsidence model confidence interval.

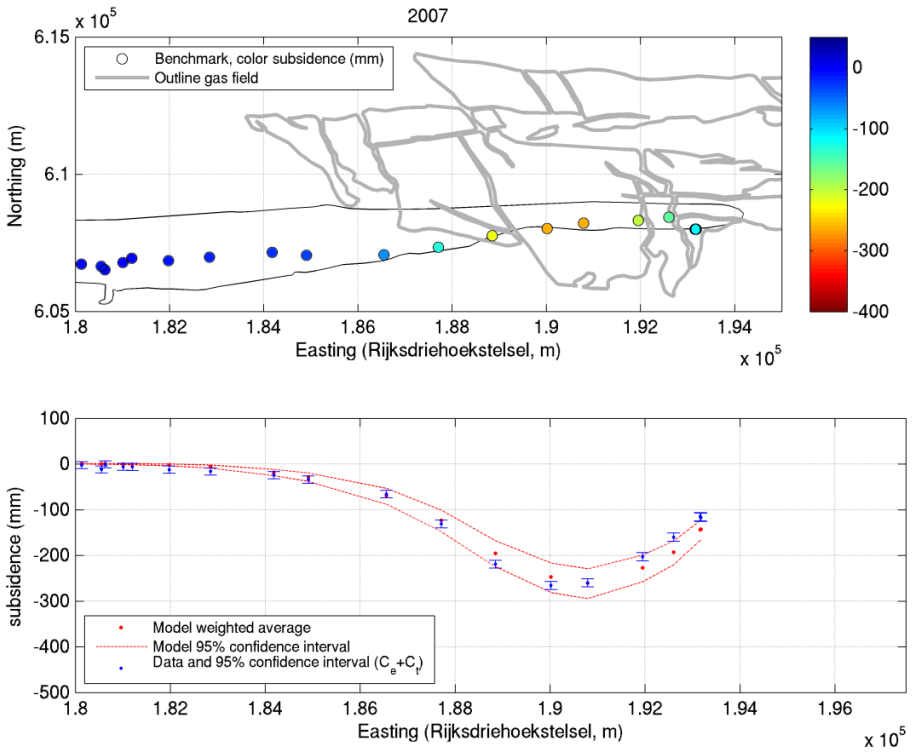


Figure 56. Example of a profile visualisation: geodetic levelling double-differences along the southern profile on Ameland for the time interval 1986-2007, with respect to benchmark 000A2592 (Nes). In red the model weighted average from the particle filter, including the 95% confidence interval. In blue the geodetic data double-differences including 95% confidence interval, representing measurement noise (C_e) and benchmark instability noise (the temporal component of idealisation noise: C_t).

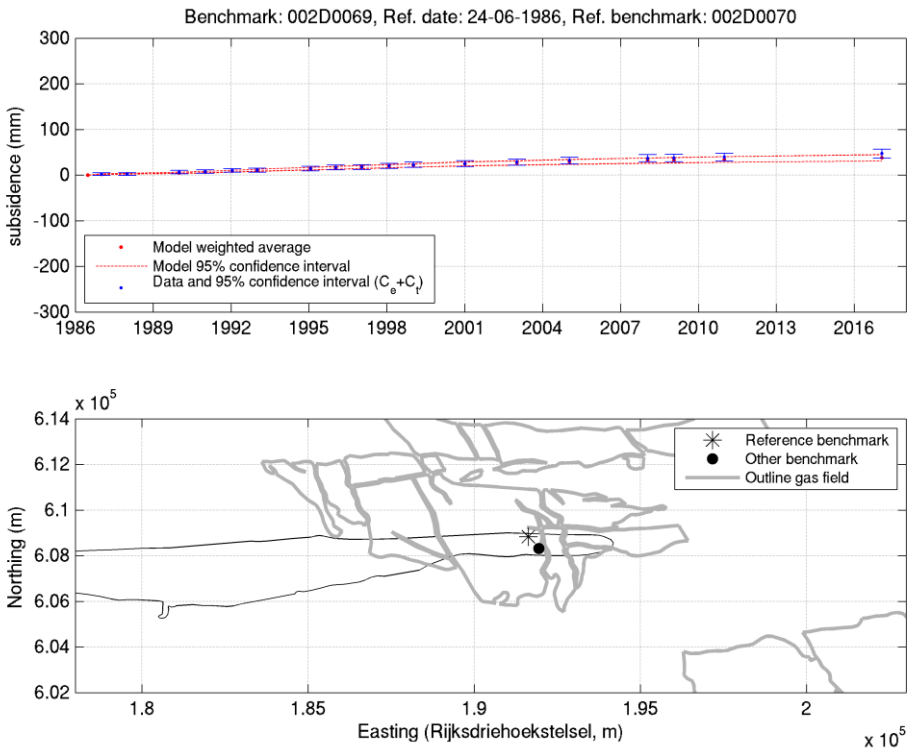


Figure 57. Double-differences of benchmark 002D0069 with respect to benchmark 002D0070 and 1986. In red the model weighted average from the particle filter, including the 95% confidence interval. In blue the geodetic data double-differences including 95% confidence interval, representing measurement noise (C_e) and benchmark instability noise (the temporal component of idealisation noise: C_t).

The 95% subsidence model confidence interval is determined according to the procedure explained in Section 7.3:

- the model ensemble member displacements for a specific double-difference in space and time are sorted from low to high,
- the cumulative probability is computed for the sorted displacements, using the probabilities that have been computed from the χ^2 test statistic values in the confrontation workflow,
- the lower and upper bound are determined two sided, at respectively 2.5% and 97.5% cumulative probability.

This procedure is carried out individually for every double-difference. For each double-difference, the probabilities of the model ensemble members stay the same, but they may appear in a different order in the cumulative probability profile. Figure 58 shows an example of the determination of the confidence interval for one double-difference. Similar confidence intervals can be obtained using the a-posteriori model covariance matrix (Eq. 35 in TNO, 2017), but the procedure from Section 7.3 has the advantage that it can account for potential non-symmetry in the probability distribution.

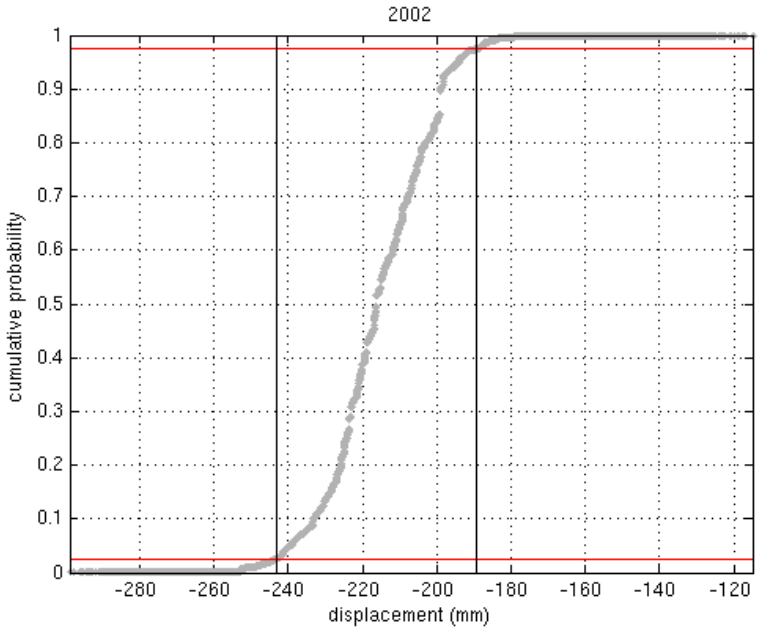


Figure 58. Cumulative probability of all model ensemble members for the double-difference displacement between 1986 and 2002, between benchmark 000A2592 (Nes) and 002D0074. The grey dots represent the cumulative probability profile. The red lines are the 2.5% and 97.5% lower and upper bound. The black lines show the interval of displacements that is in the 95% confidence interval.

7.5 Subsidence rates

For each member, the subsidence volume in the Pinkegat area (see Figure 59 for the Pinkegat boundaries) has been calculated over time, which is converted to an average subsidence rate, which in turn is averaged over 6 years, to obtain subsidence rates for each member in the ensemble in line with those reported in the M&R framework (NAM, 2012). The 6-year-average of the rate in year Y, rate(Y), is calculated from the subsidence rates, S, three years before and two years after that specific year:

rate (Y) = [S(Y-3) + S(Y-2) + S(Y-1) + S(Y) + S(Y+1) + S(Y+2)] / 6.

After the confrontation with the geodetic data, a weight or relative probability has been assigned to each of the members, as discussed in the previous sections, which is used to determine an expectation case of the subsidence rate (defined as the weighted average of the subsidence rates from all the ensemble members), together with 95% confidence intervals. Note that the members in the ensemble only use reservoir scenarios from the Ameland field, hence all reported subsidence rates do not include the effect from other fields in the area.

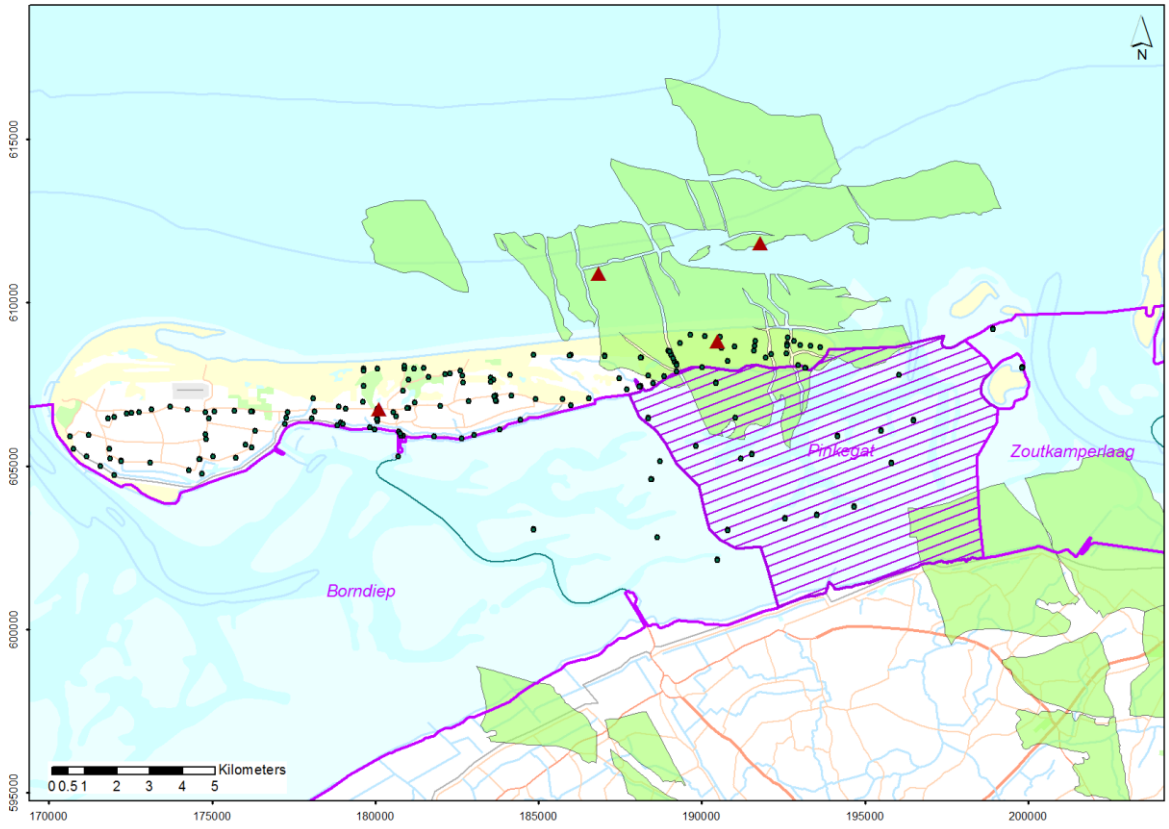


Figure 59. Location and boundaries of the different sand sharing areas in the Waddenzee in pink, including the Pinkegat area. Also shown are the gas fields in the North of the Netherlands (green), the locations of the benchmarks (dots) and permanent GPS stations (red triangles).

Figure 60 shows the resulting cumulative probability as a function of increasing subsidence rates in two different years (1991 and 2015). Here, the subsidence rate is first sorted in each year from lowest subsidence rate to highest, and the probability cumulated accordingly. In both cases, a well-defined distribution of the subsidence rate probability can be observed, being localised with boundaries that can be described in terms of confidence levels. The localised nature of the obtained probability distribution becomes especially pronounced when considering that some members in the ensemble produce subsidence rates are far beyond the limits of the plots here, but the weights assigned through the confrontation method to these members are so insignificant that their contribution to the cumulative probability is negligible.

Since these probability curves are smooth and densely sampled, their boundaries can be parameterized in terms of confidence levels, to facilitate the visualization of its development over time. A two-sided 95% confidence interval is applied, i.e. there is a 2.5% probability that the subsidence rates are below the lower bound and a 2.5% probability that the subsidence rates are higher than the upper bound in the figures of subsidence rate versus time. The location for these two bounds are indicated for the years 1991 and 2015 in Figure 60. Figure 61 shows the development of the confidence bounds over time, together with the weighted average or expectation rate, for the Pinkegat area. At the bottom of the figure, probability distributions of the subsidence rates are shown at three different times, all showing a localised distribution.

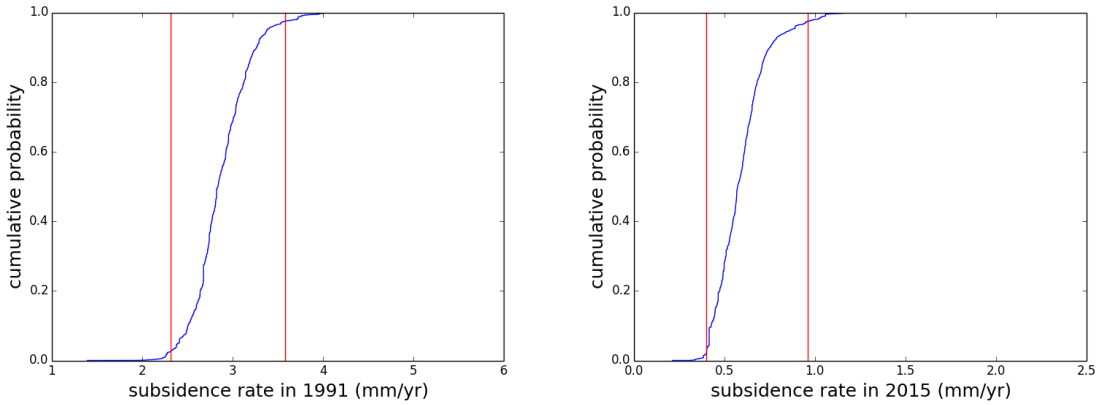


Figure 60. Cumulative probability as a function of subsidence rate in the year 1991 (left) and 2015 (right). Red lines indicate lower bound (2.5%) and upper bound (97.5%) of the 95% confidence interval.

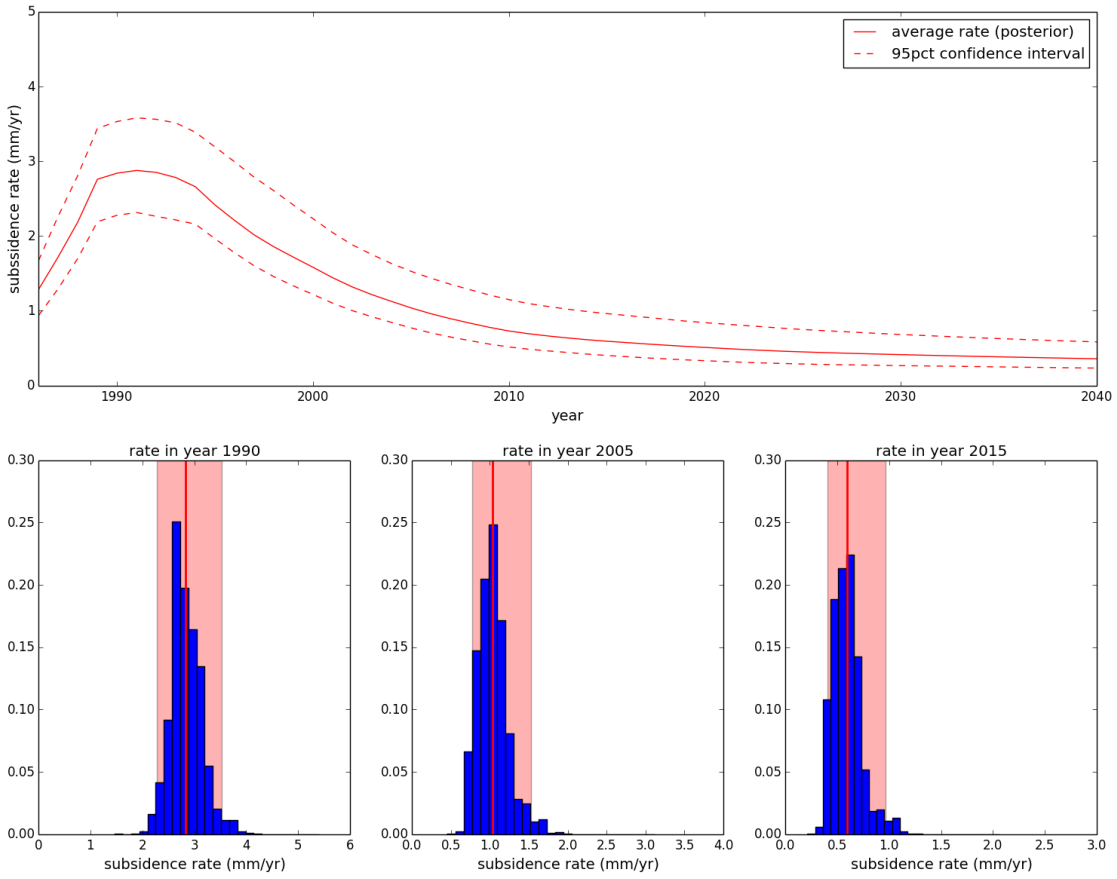


Figure 61. Contribution of the Ameland field to the subsidence rates for the Pinkegat area, starting in 1986 (start of production) and forecast up to 2040. Bottom plots show the probability distributions of the subsidence rate at three different times (1990, 2005 and 2015). The thick red line in the histograms marks the weighted average subsidence rate in that year, and the pink area marks the 95% confidence interval.

Figure 62 shows the same expectation subsidence rate together with the 95% confidence interval as in Figure 61, but now on top of the expected sea level rise rate as given by Ministerie van Economische Zaken (2016). To the right the same figure, showing only the Ameland contribution to the M&R 2016 (NAM, 2016a) results.

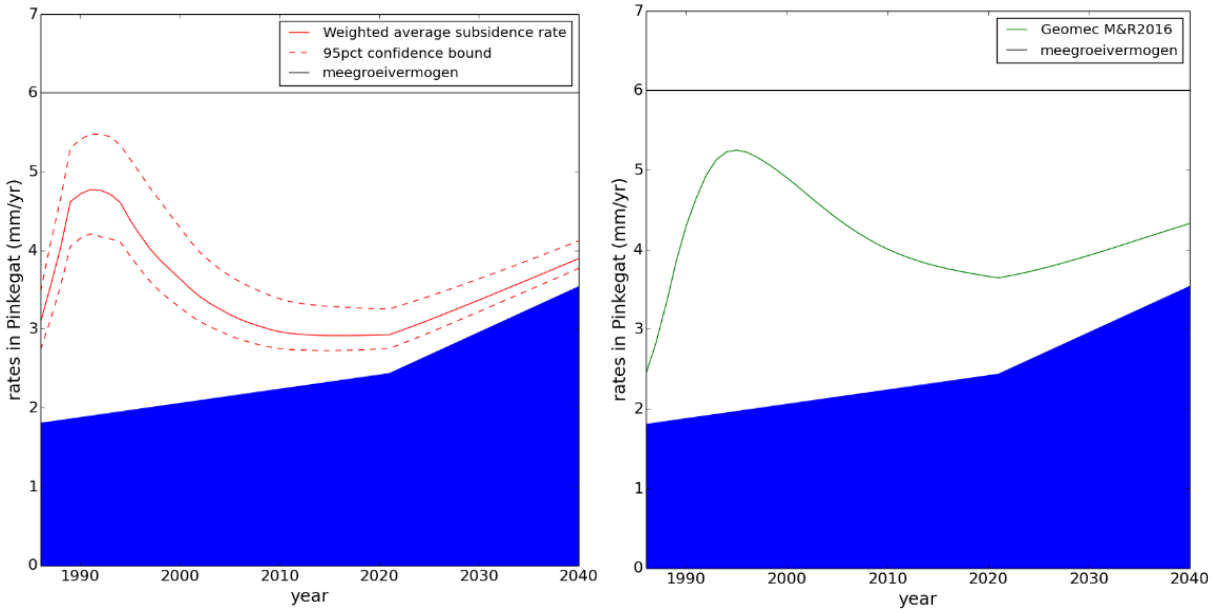


Figure 62. Left: Subsidence rate (weighted average) with 95% confidence interval on top of the predicted sea-level rise rates for the Pinkegat area. Right: subsidence rates in Pinkegat for the M&R 2016 result.

The M&R 2016 result shows higher values for the subsidence rates when compared to the LTS-II expectation case but a direct comparison with the LTS-II results is difficult because the M&R model is based on a different reservoir scenario and on a different geomechanical model.

In more detail, the two observations listed below largely explain why the LTS-II expectation case shows a lower rate than the M&R 2016 result.

1. The M&R 2016 base case clearly overpredicts the subsidence above the southwestern aquifer, indicated by three benchmarks that are located within the yellow circle in Figure 63. The likely cause is a too high depletion of this aquifer. The choice in the M&R approach to align all parameter values in the reservoir model for both the southwestern and southeastern aquifers led to an overprediction of the subsidence in the southwestern aquifer, while the modelled subsidence above the southeastern aquifer is matched to the data.
2. The LTS-II expectation case shows an underprediction of the subsidence above the southeastern aquifer (Figure 64). This is a result of the applied LTS-II method that selects the best member based on all measurements. In trying to match the modelled results to the steeper flank of the subsidence bowl, the workflow selects a reservoir scenario with less depletion in the eastern part of both the reservoir and aquifer. Combined with the stiff response of the linear branch this leads to an underprediction of the subsidence above the southeastern aquifer. Also, in the LTS-II study, the parameter values of the two aquifers were varied independently.
   Note that the M005 benchmark could be influenced by a few mm by the subsidence caused by the gas production from the Nes field, situated south of this benchmark.

Based on these observations it is concluded that the M&R 2016 base case represents a conservative scenario, overpredicting the subsidence above the southwestern aquifer. This is caused by parameter choices in the M&R 2016 reservoir scenario. The LTS-II expectation case

slightly underpredicts the subsidence above the southeastern aquifer. This is mainly caused by the LTS-II confrontation workflow that obtains a best fit to all the data. This best fit may show small local variations as a result of this general assumption. Please note that there are only very few measurements above the southeastern aquifer to compare with the subsidence model members.

It is recommended in the next M&R cycle to use the best of both worlds and, more precisely, alter the southwestern aquifer property values independent from the southeastern aquifer property values. This will likely lead to a reduction of the subsidence rate in Pinkegat in the next M&R cylcle.

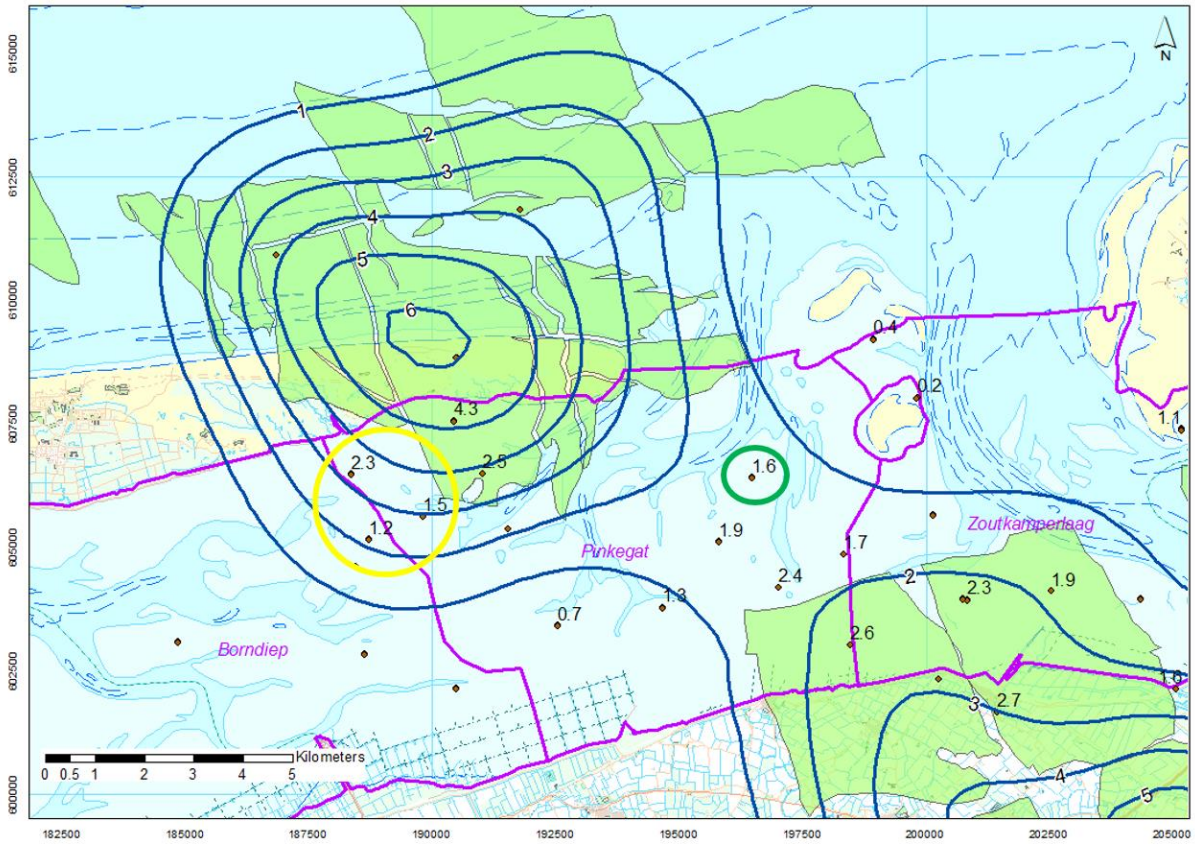


Figure 63. Results for the M&R cycle (NAM 2016a) with in the yellow circle the three benchmarks indicating an overprediction of the subsidence by the model. The green circle indicates the position of benchmark M005. The subsidence is shown in cm for the period 2006 – 2015.

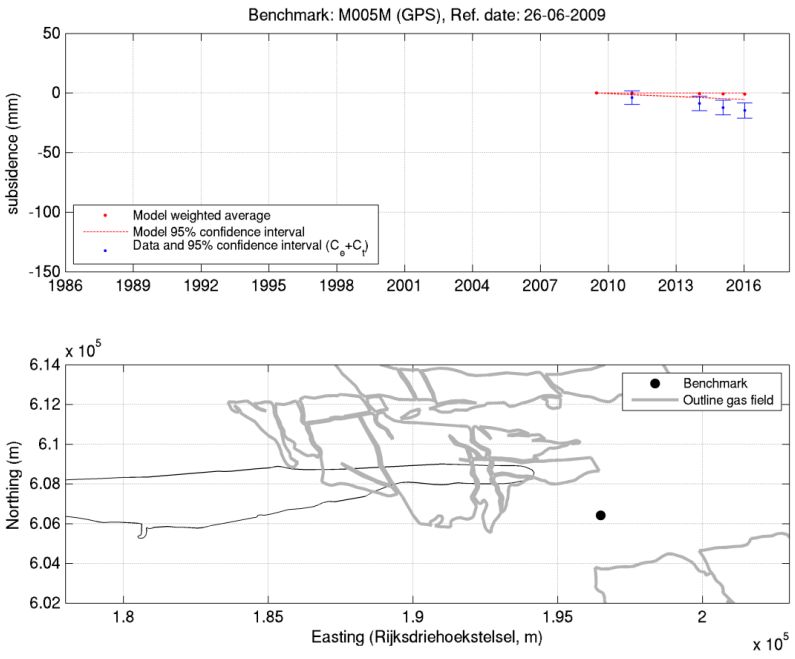


Figure 64. Data versus model for benchmark M005M (within the green circle Figure 63) above the eastern aquifer indicating that the LTS-II expectation case under predicts the measured subsidence. The model confidence interval is computed from the model ensemble that reflects subsidence due to the Ameland gas fields only.

7.5.1 Effect of emergency stop on subsidence rates

The "hand-on-the-tap" procedure that it is in place for the Waddenzee fields (the gas fields Nes, Moddergat, Lauwersoog Oost, Lauwersoog West and Vierhuizen), defines control measures in the form of production reduction if subsidence rates might reach the maximum allowed levels. These control measures are highlighted in a so-called "remweg scenario" or emergency stop scenario, i.e. a scenario that demonstrates the effect of a production stop in these fields on the subsidence rates in the sand sharing areas like Pinkegat.

The member probabilities obtained from the confrontation of the full data set with geodetic data can also be used to determine the expected subsidence in case the production would be stopped at a given moment. Figure 61 shows the result of the potential implementation of a production stop and clearly demonstrates the effectiveness of implementing the production reduction control measure.

Figure 65 shows the effects of a production stop of the Ameland field in 1996 and one in 2016 on the expected subsidence rate in Pinkegat. The subsidence ensembles were generated by altering the pressure scenarios of the full production forecast to now reflect the pressure development as a result of a production stop. The exact same geomechanical *a posteriori* model parameters were subsequently applied to arrive at the same set of members with the same member probability, only now having a different pressure depletion after the stop dates. The expected subsidence rate is again the weighted average of all the members in the ensemble. The resulting expected subsidence rates for the emergency stop scenarios are plotted in Figure 65, together with that from the full production. The bottom graph in Figure 65 plots the cumulative production for Ameland over time, where the solid lines are actual production and dashed are forecast for the full production and hypothetical production for the two stop scenarios.

Especially the emergency stop in 1996 has a significant impact on the subsidence rate, since a large part of the production is still to come after 1996. Even though the majority of the field's production has already occurred by 2016, an emergency stop at such a late time in the field's development is still expected to have a significant effect based on these results, as the stop in 2016 reduces the expected subsidence rate by about a factor of two.

The current production lifetime of the Waddenzee fields is around 11 years and the relative effect of a hypothetical production stop in those fields would be more alike the 1996 production stop in Ameland that is described in the section above. By making this comparison, it is concluded that a production stop in the Waddenzee fields is an effective measure to reduce the subsidence rate in the sand sharing areas.

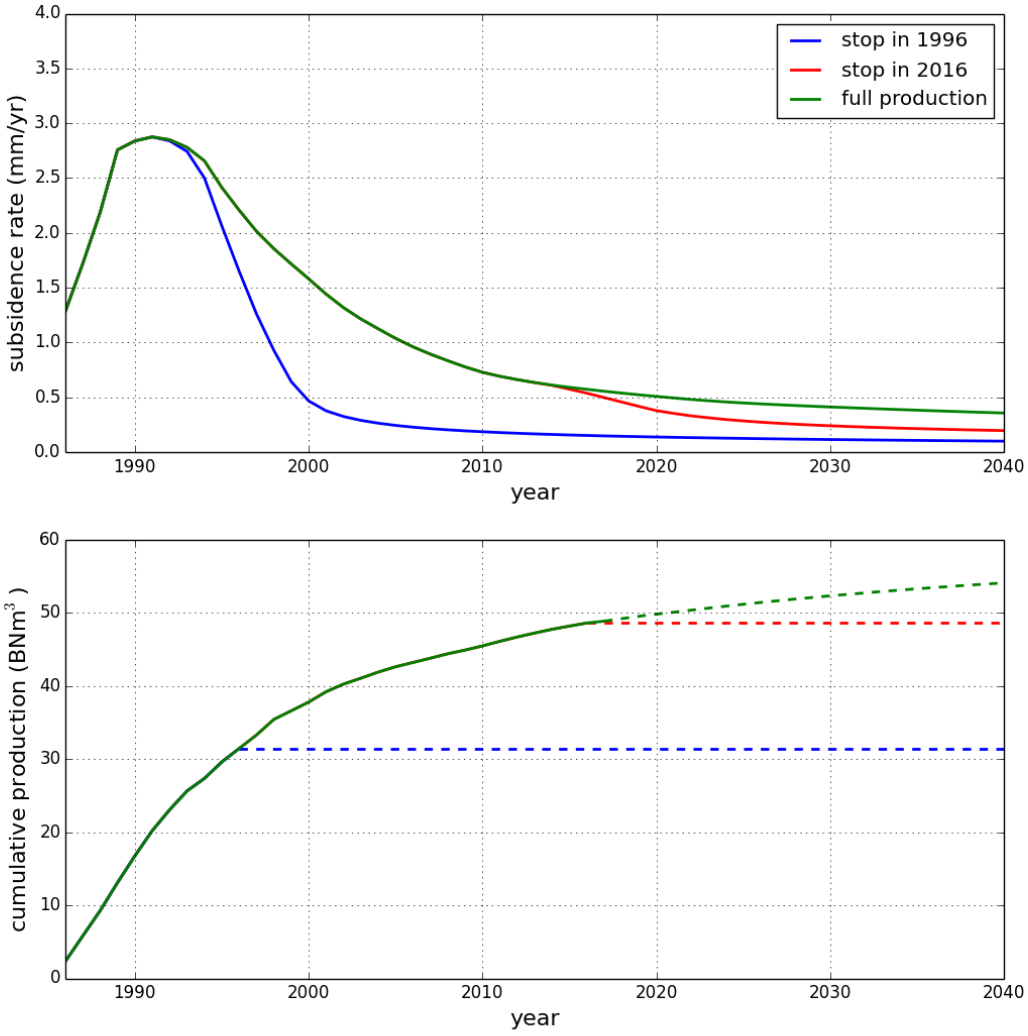


Figure 65. Top graph shows the subsidence rate for full production scenarios (green), emergency stop in 1996 (blue) and emergency stop in 2016 (red). Subsidence rate plotted here is the weighted average, i.e. expectation case, after confrontation with the full set of geodetic data. Bottom graph shows the cumulative production. Actual production is plotted by the solid line (up to October 2017). The green dashed line shows the production forecast for the full production scenario. The red and blue dashed lines mark the total cumulative production for a hypothetical stop in 2016 and 1996, respectively.

7.6 Additional insights from the confrontation

The outcomes of the confrontation of the full subsidence ensemble with the geodetic data are presented in this section. First, the resulting probability distributions as a function of the pressure scenario are shown in Section 7.6.1, to give an impression of the more likely depletion scenarios and reservoir parameter sets. Section 7.6.2 and 7.6.3 present additional insights into the parameter sensitivity of the geomechanical model.

7.6.1 Probability distributions of reservoir scenarios and reservoir parameters

If all input parameters were truly randomly distributed amongst the subsidence members, one should expect to obtain a randomly distributed set of χ^2 values and corresponding member probabilities as a function of member number in Figure 54 and Figure 55, respectively. Although the geomechanical parameters are randomly distributed among the members here, the pressure scenarios are not, as each group of 20,000 members is based on the same reservoir scenario (i.e. first 20,000 are based on the first scenario, second 20,000 on the second scenario, etc). The apparent structure that is especially pronounced in the member probability plot in Figure 55 suggests some reservoir scenarios can produce more probable subsidence members in combination with the geomechanical parameters, whereas other scenarios seem to be incapable of producing a likely subsidence member with any parameter set.

The summed probability of all 20,000 members per reservoir member is plotted in Figure 66, to better assess the likelihood of each reservoir scenario the results are tabulated in Table 19. These summed member probability values represent the total contribution of each reservoir scenario to the weighted averages, and can thus be viewed as a likelihood of the reservoir scenarios, and is referred to as 'scenario probability' in this study. Remarkably, some reservoir scenarios have a contribution of virtually zero while at the same time there are reservoir scenarios with a scenario probability of 10% and higher.

Besides the scenario probability, Figure 76 also shows the number of members from each reservoir scenario that would be accepted when a 95% acceptance level is applied (see Section 7.3.1). In general, it can be observed that more individual members would be accepted from scenarios with a higher probability. A total of 1,870 members out of 1,160,000 fall within the 95% acceptance level. The scenario probability and the number of accepted members are listed in Table 19 for all reservoir scenarios. Like the scenario probability, the number of accepted members gives an indication of the relative contribution of each scenario to the final results. Some interesting observations can be made, for example models M91 and M9203 have almost the same scenario probability, however almost three times as many members from M9203 fall inside the 95% acceptance level. This means that individual members from M91 are able to produce, on average, a better match to the data.

The probability per reservoir scenario can help in investigating which aspects of the pressure depletion improve the match of modelled to measured subsidence. In Figure 67, the lateral depletion in 2016 for all reservoir scenarios are plotted in order of descending scenario probability. Figure 68 shows the depletion over time for all the models, where the thickness of the line indicates the scenario probability. The figures show that the reservoir models with the highest probability have moderate depletion in the western aquifer and a moderate depletion in the northern part of the eastern aquifer, and no or limited depletion in the central and southern part of the eastern aquifer.

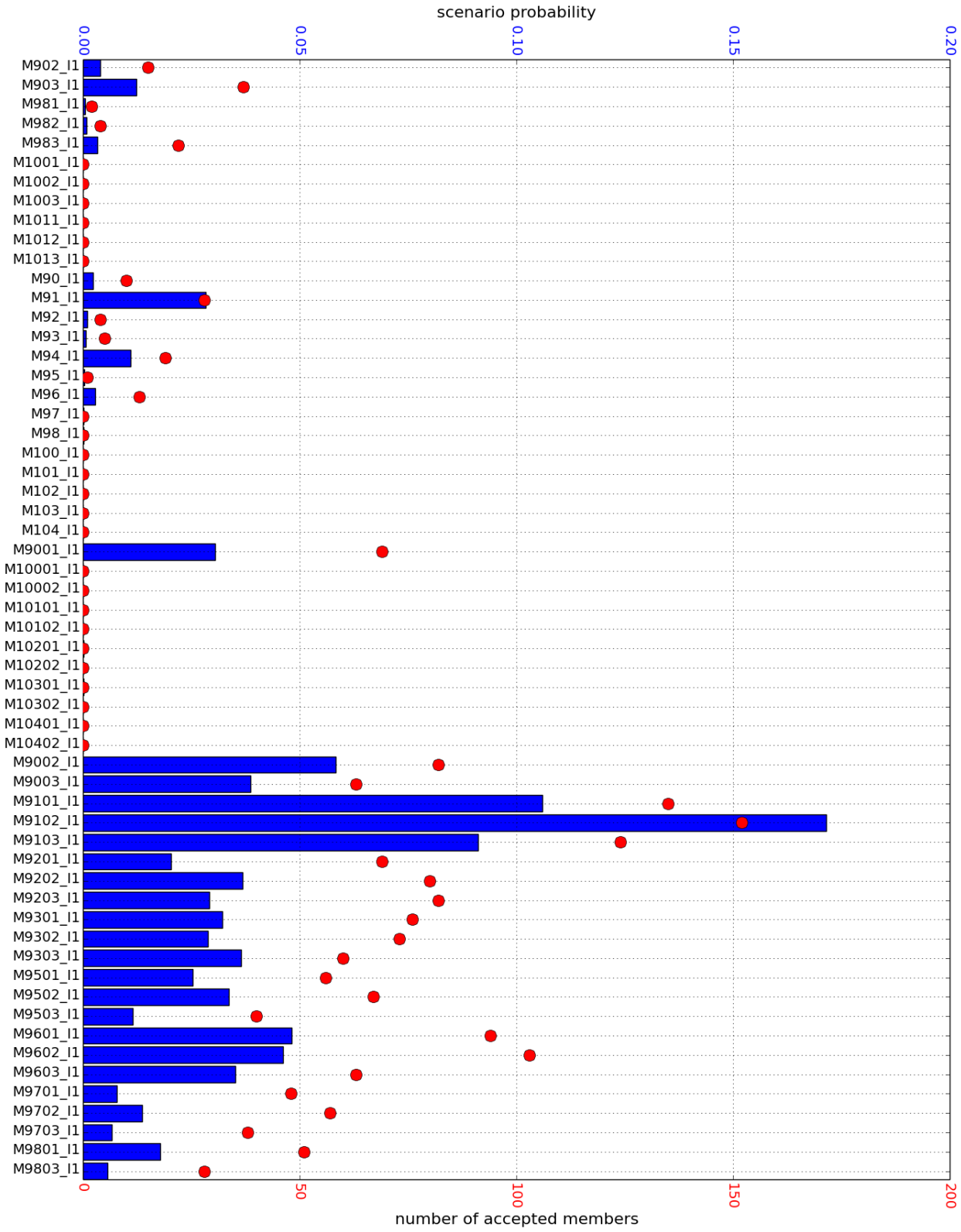


Figure 66. Assessment of the probability per reservoir scenario. The blue bars show the sum of the member probability per reservoir pressure scenario (58 scenarios in total) The distribution shows that clearly some scenarios result in a better match to the geodetic data than others. The number of members from each reservoir scenario that would be accepted in a 95% acceptance level is indicated by the red dots.

Table 19. Copy of Table 6, now including the scenario probability and the number of accepted members in a 95% acceptance level.

Model Name	Sgr	Fault_13	E21_Kaqf	E11_15_Kaqf	E11_15_GIIP multiplier	Scenario probability %	Number of accepted members
90	1	0	1	1	0.88	0.23	10
91	1	0	0.1	1	0.88	2.83	28
92	1	0	0.01	1	0.88	0.09	4
93	1	1	1	1	0.88	0.06	5
94	1	1	0.1	1	0.88	1.1	19
95	0.5	0	1	1	0.88	0.03	1
96	0.5	0	0.1	1	0.88	0.28	13
97	0.5	0	0.01	1	0.88	0.02	0
98	0.5	1	1	1	0.88	0.02	0
100	0	0	5	1	0.88	0	0
101	0	0	1	1	0.88	0	0
102	0	0	0.1	1	0.88	0	0
103	0	0	0.01	1	0.88	0	0
104	0	1	1	1	0.88	0	0
902	1	0	1	0.05	0.9	0.4	15
903	1	0	1	0.025	0.9	1.24	37
981	0.5	1	1	0.1	0.88	0.05	2
982	0.5	1	1	0.05	0.91	0.08	4
983	0.5	1	1	0.025	0.9	0.33	22
1001	0	0	5	0.5	0.88	0	0
1002	0	0	5	0.25	0.88	0	0
1003	0	0	5	0.1	0.88	0	0
1011	0	0	1	0.5	0.88	0	0
1012	0	0	1	0.25	0.88	0	0
1013	0	0	1	0.1	0.88	0	0
9001	1	0	1	0.001	0.9	3.06	69
9002	1	0	1	0.005	0.9	5.83	82
9003	1	0	1	0.01	0.9	3.87	63
9101	1	0	0.1	0.001	0.9	10.6	135
9102	1	0	0.1	0.005	0.9	17.14	152
9103	1	0	0.1	0.01	0.9	9.11	124
9201	1	0	0.01	0.001	0.9	2.04	69
9202	1	0	0.01	0.005	0.9	3.69	80
9203	1	0	0.01	0.01	0.9	2.91	82
9301	1	1	1	0.001	0.9	3.22	76
9302	1	1	1	0.005	0.9	2.89	73
9303	1	1	1	0.01	0.9	3.66	60
9501	0.5	0	1	0.001	0.9	2.53	56
9502	0.5	0	1	0.005	0.9	3.36	67
9503	0.5	0	1	0.01	0.9	1.16	40
9601	0.5	0	0.1	0.001	0.9	4.82	94
9602	0.5	0	0.1	0.005	0.9	4.62	103
9603	0.5	0	0.1	0.01	0.9	3.52	63
9701	0.5	0	0.01	0.001	0.9	0.78	48
9702	0.5	0	0.01	0.005	0.9	1.37	57
9703	0.5	0	0.01	0.01	0.9	0.67	38
9801	0.5	1	1	0.001	0.9	1.78	51
9803	0.5	1	1	0.01	0.9	0.57	28
10001	0	0	5	0.01	0.9	0	0
10002	0	0	5	0.05	0.9	0	0
10101	0	0	1	0.01	0.9	0	0
10102	0	0	1	0.05	0.9	0	0
10201	0	0	0.1	0.01	0.9	0.02	0
10202	0	0	0.1	0.05	0.9	0	0
10301	0	0	0.01	0.01	0.9	0.01	0
10302	0	0	0.01	0.05	0.9	0	0
10401	0	1	1	0.01	0.9	0	0
10402	0	1	1	0.05	0.9	0	0

Figure 67. Lateral depletion for the year 2016 for all pressure scenarios, sorted on probability (between brackets) For comparison, the last scenario is the pressure scenario used for the Ameland field in the M&R 2016 (not used explicitly in the LTS-II study).

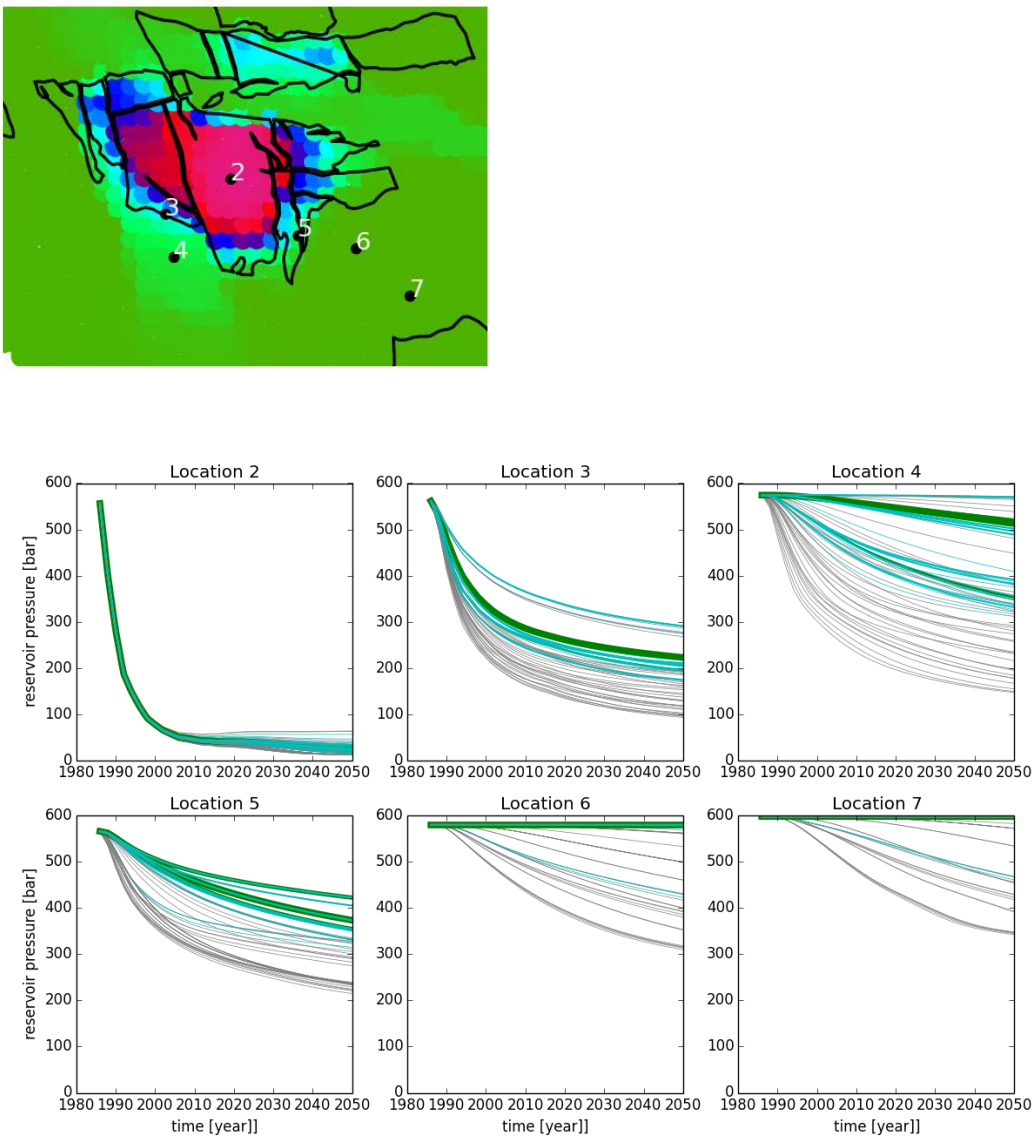


Figure 68. The bottom figure shows the pressure vs time for the four points shown in the top figure. The thicker lines represent more probable reservoir scenario's. The colours represent also the probability: green lines higher than 5%, cyan higher than 0.5%, grey lines the lower probabilities.

The probability per pressure scenario can also be related to some key reservoir parameters used in the pressure depletion modelling. Examples are displayed in Figure 69, Figure 70 Figure 71 and Figure 72 (please refer to Section 3.2.3, Table 1-Table 5 for detailed parameters' description).

Figure 69 shows the scenario probability together with the residual gas saturation multipliers used in each of the scenarios. Most often, a good match with the subsidence data is reached with an Sgr multiplier of 1 (resulting in 10% of residual gas in the aquifer) or 0.5 (equal to 5% Sgr). At the same time, all the scenarios with Sgr multiplier of 0 (i.e. no residual gas in the aquifer) are at the bottom of the list with very low probabilities, meaning that it is very hard (up to impossible) to achieve a good subsidence match with no residual gas in the model, even though these scenarios give a good match to production and pressure data.

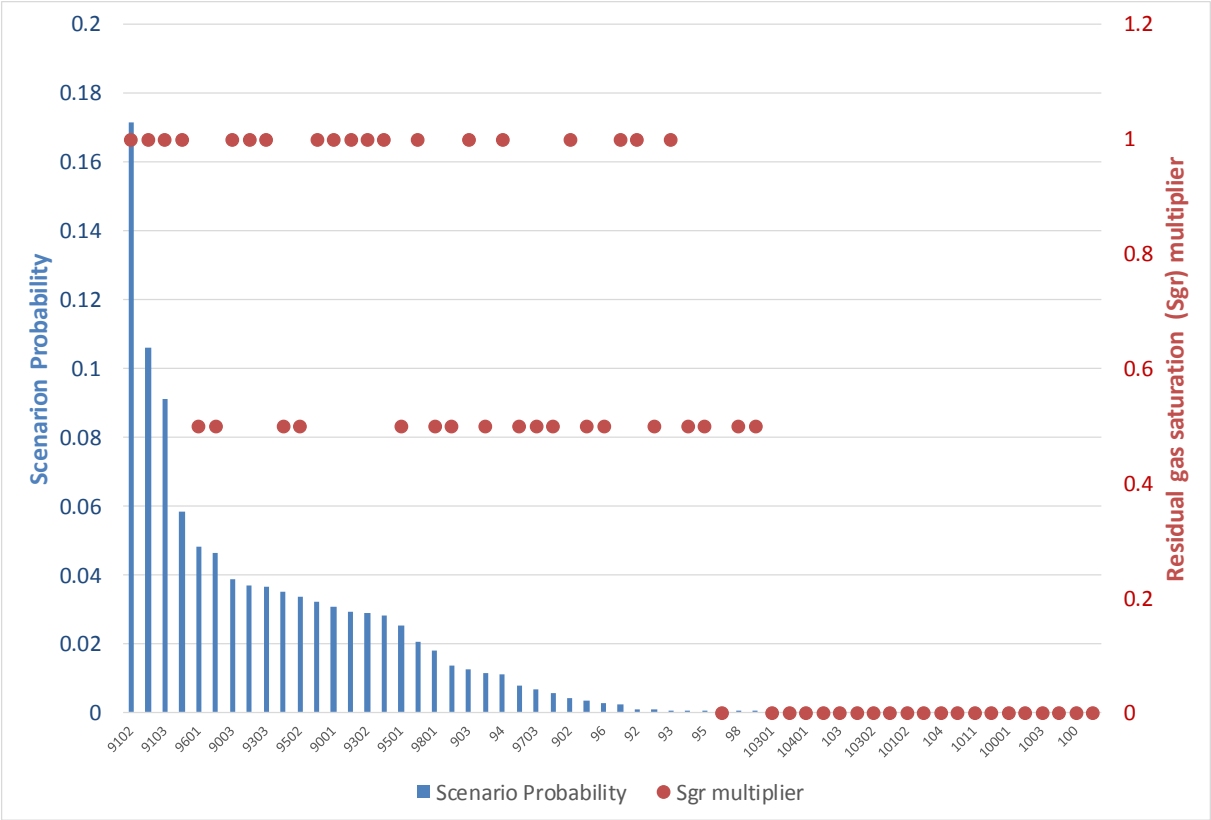


Figure 69. Scenario probability and residual gas saturation multiplier in the aquifers. For Sgr = 1, there is a 10% residual gas saturation in the aquifer; Sgr = 0.5 means 5% residual gas saturation; Sgr = 0 means no residual gas saturation.

Figure 70 shows the scenario probability together with the aquifer permeability multiplier, applied in the southeastern aquifer of the Ameland-Oost field. This southeastern aquifer shows a good correlation between the permeability multiplier and the probability of the reservoir scenario. Lower values of this multiplier seem to be more likely than higher values, implying that it is more likely from the confrontation result that a southeastern with a low permeability is more likely. Figure 71 compares the scenario probabilities against permeability multipliers, but now applied in the southwestern aquifer of the Ameland-Oost field (block E21). Conclusions which may be drawn from analysis of both plots are as follows: the three reservoir models with the highest average probability all use an aquifer permeability multiplier of 0.1, while a good subsidence match is also achievable for aquifer permeability multipliers of 0.01 and 1, but almost no scenarios were accepted with aquifer permeability multiplier of 5.

Additional to the aquifer permeability in the southwestern aquifer, the transmissibility of Fault 13, which separates the aquifer in a northern and southern compartment (see Figure 4), appears to have a strong impact on the scenario probability, as depicted by Figure 72. The scenarios with the highest probability all have a transmissibility of Fault 13 of zero, suggesting that the depletion of the southwestern aquifer is most likely constrained to its northern section.

This loop between reservoir engineering and geomechanic domains was not foreseen as a direct objective on beforehand, but now is considered to be essential in improving of understanding of the dynamic response of the Ameland field. Purely from a reservoir engineering point of view, all 58 selected scenarios may be treated as equally likely, but this finding clearly indicates what subsurface

scenarios are more likely from a geomechanics point of view, significantly reducing the uncertainty of a number of subsurface parameters.

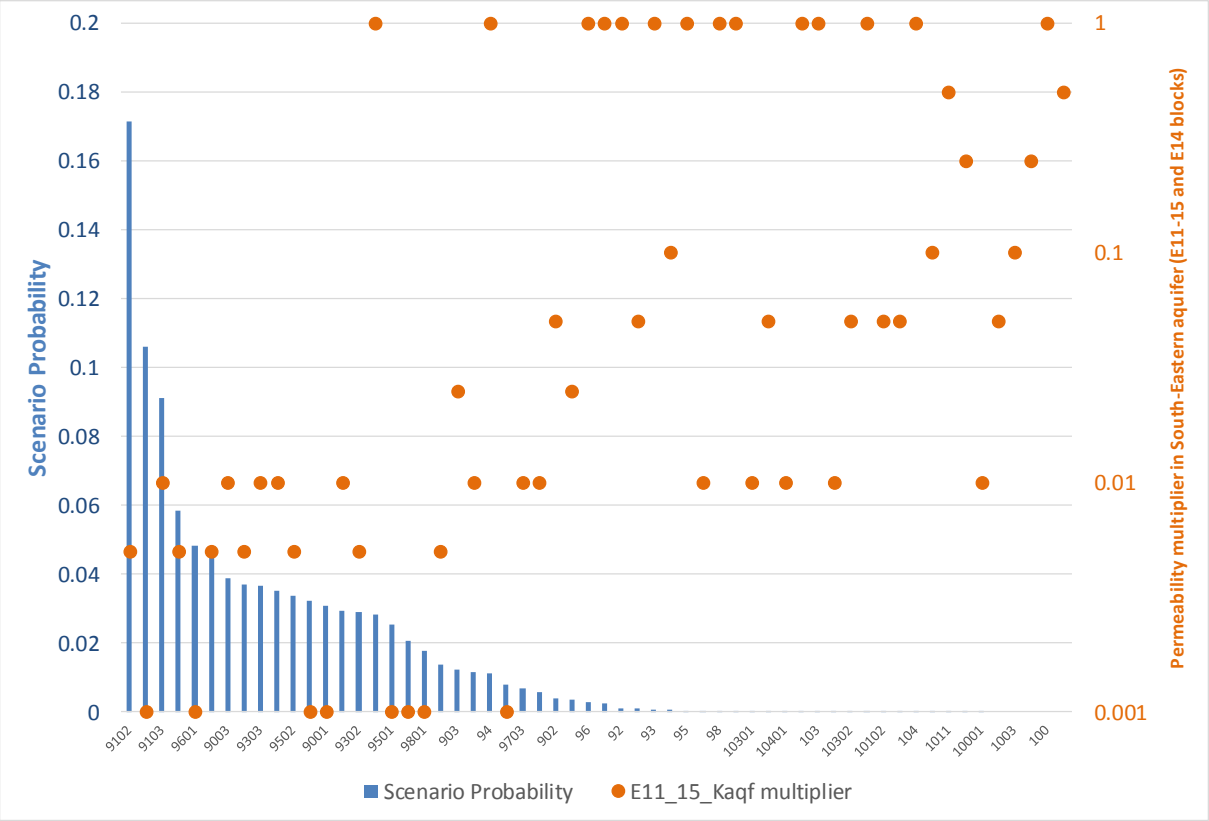


Figure 70. Scenario probability and permeability multiplier in the south-eastern south-eastern aquifer.

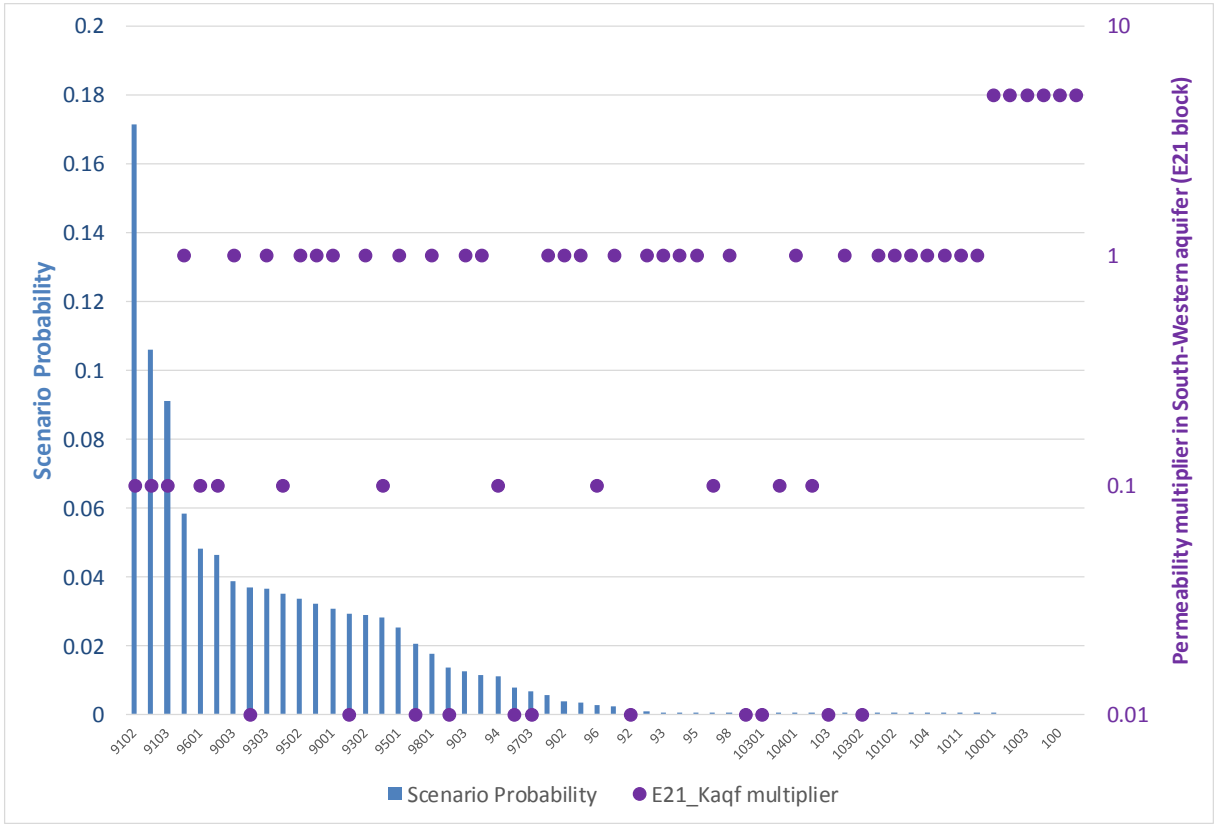


Figure 71. Scenario probability and permeability multiplier in the southwestern aquifer.

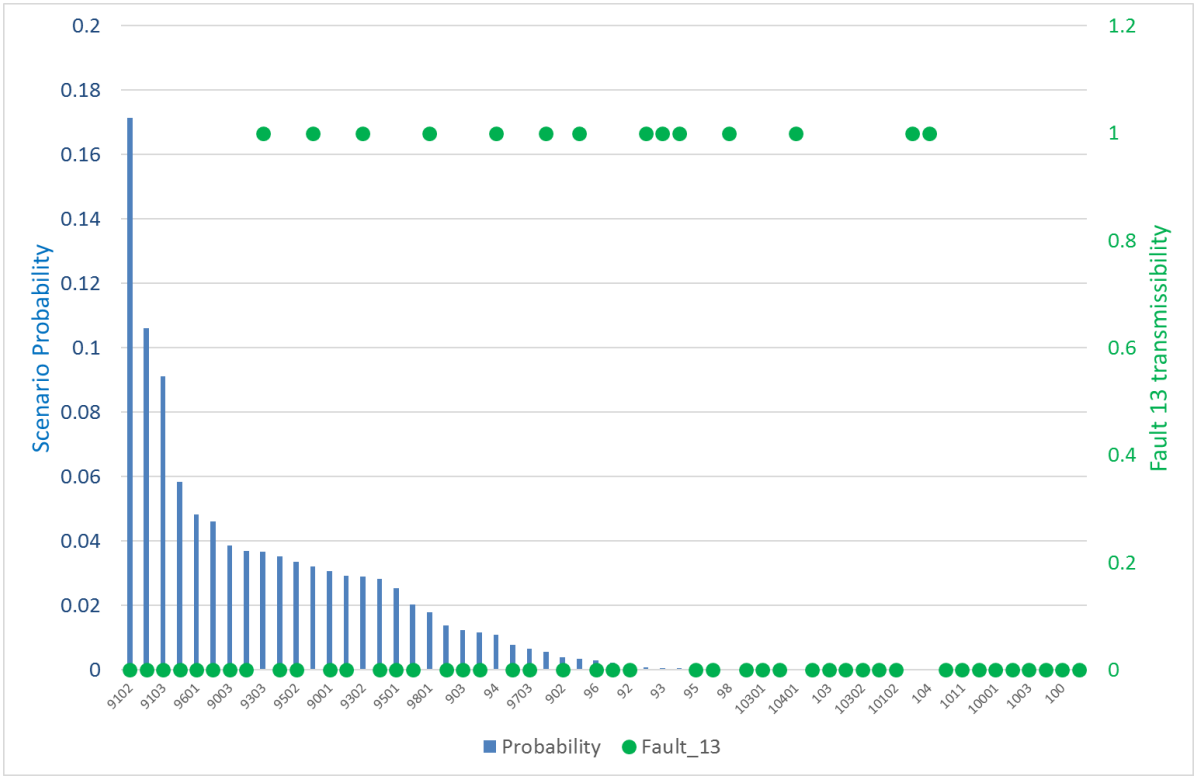


Figure 72. Scenario probability and transmissibility of Fault 13, which divides the Southwestern aquifer (see Figure 4 for location of the fault).

7.6.2 Probability distributions of geomechanical model parameters

Next to an assessment of the probability of the reservoir scenarios, the probability of the resulting geomechanical parameters used in the Monte Carlo simulation is studied. First, the C_m-Porosity relations are shown in Figure 73, where the darker green lines have a higher probability than the lighter green lines. First of all, it can be observed that all lines are positioned in the upper half of the possible range. This means that the C_m values are high and therefore that the compaction volume per bar depletion is also high when compared to a field like Groningen where inversion demonstrates that the inverted C_m values are situated in the bottom part of the same plot (van der Wal en van Eijs, 2016). Secondly, the relationship with the porosity is for quite some members weak. It means that even with a homogeneous subsurface distribution of the porosity, a good fit can be achieved.

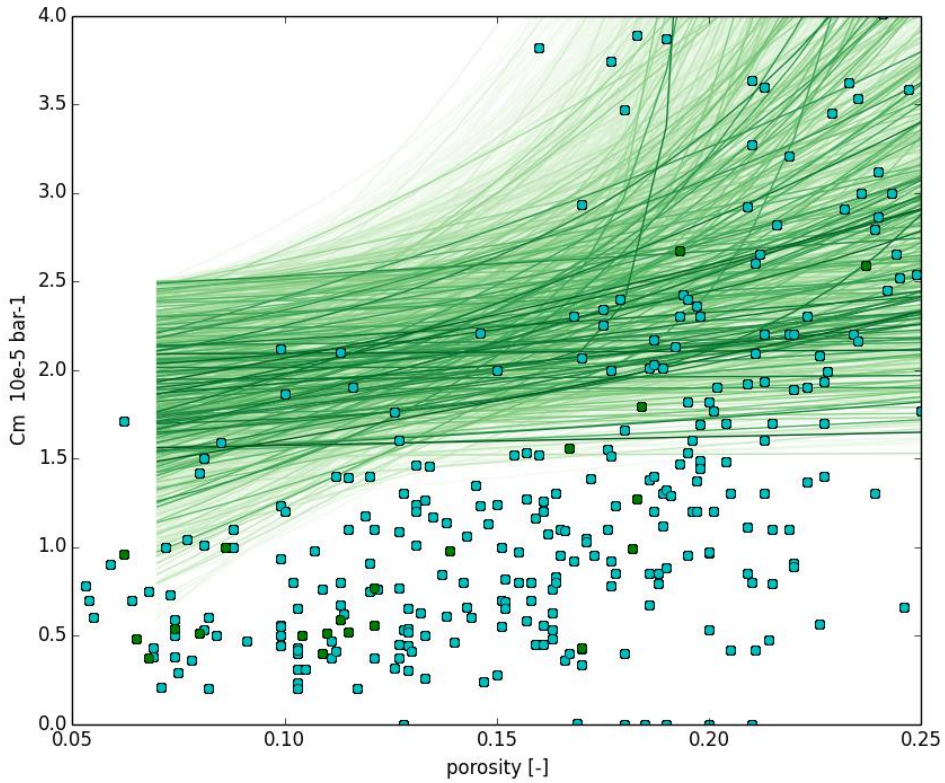


Figure 73. C_m-Porosity relation for the accepted members. Darker green lines indicate a higher probability. The dots represent the core plug experiments results for Ameland cores (green) and other Rotliegendes cores (cyan).

The five graphs shown in Figure 74 present the cumulative probability distribution of the 5 most relevant geomechanical parameters. A steep curve for a particular parameter implies that the workflow converged towards a narrow possible range of values, like e.g. the values for the C_{md}/C_{ref}. Vice versa, a shallow slope indicates a wide range of possible values. Salt behaviour and time-dependent compaction are two competing mechanisms, as described earlier, with the influence of salt on the bowl shape being volume neutral. The results show that it is likely that the salt flow, at present, still influences the bowl shape above the gas field. However, the dominant mechanism at present for the subsidence is reservoir compaction.

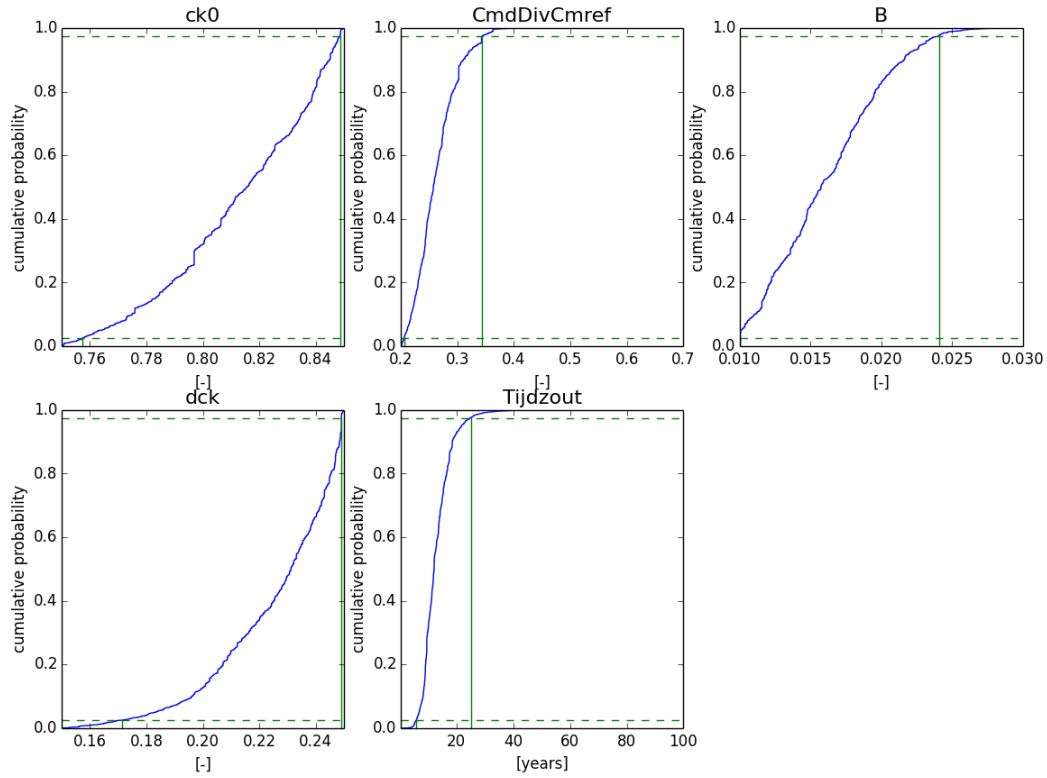


Figure 74. Probability distribution of the geomechanical parameters, the bottom part shows the different parameters with its range (x-axis) and cumulative probability (y-axis). The green lines indicate the 95% confidence interval

7.6.3 Parameter sensitivity

A possible other way to investigate the influence of the individual parameters on the results is the calculation and visualisation of tornado charts. These plots show the difference in subsidence, for 2040 with respect to the subsidence as forecasted by the member with having the highest probability value, at four different locations. The geomechanical parameters were varied individually within the value range that results from the *a posteriori* 95% confidence interval shown in Figure 74, while keeping the other parameters fixed at the values used for the most likely member. The C_m-Porosity range is defined by the green lines in Figure 74. The impact of the reservoir scenarios in the tornado charts is governed by selecting the member with the highest probability per reservoir scenario.
The difference between the, for the year 2040, forecasted subsidence, resulting from the parameter variation, and the forecast that results from the most probable member is indicated on the x-axis of the charts.
The bar right to the parameter name indicates the impact of the parameter value variation on the subsidence forecast at the specific location, with a wider bar indicating a larger influence on the result.
Note that for the tornado chart all parameters were fixed to those of the most probable member and only one parameter at the time is varied. The tornado chart therefore only indicates the sensitivity per parameter on the subsidence and is not a measure of a confidence interval. This latter remark is related to the fact that some of the parameters show a dependency with other parameters, which is not honoured by the workflow underlying the tornado chart. In other words, this workflow creates members that would be not accepted by the confrontation workflow. Figure 75 shows that the

influence of the reservoir scenario is more prominent in the aquifer area, while the uncertainty of the C_m-Porosity relation dominates the results on the island. This can be explained by the fact that the variation in possible pressure depletion is small in the gas part while it is large in the aquifer part. Also, it is observed that the influence of the b-value of the RTCiM is smaller in the aquifer areas, which can be explained by the fact that most of the reservoir scenarios show reservoir pressures that are still above the hydrostatic pressure. At those locations, only the added linear branch is valid, excluding a possible effect of the remaining RTCiM parameters.

Finally, it is observed that the salt viscosity has a small influence on the forecasted result. Within the range, governed by the *a-posteriori* results, it is forecasted that the salt viscosity does not influence the subsidence anymore in 2040.

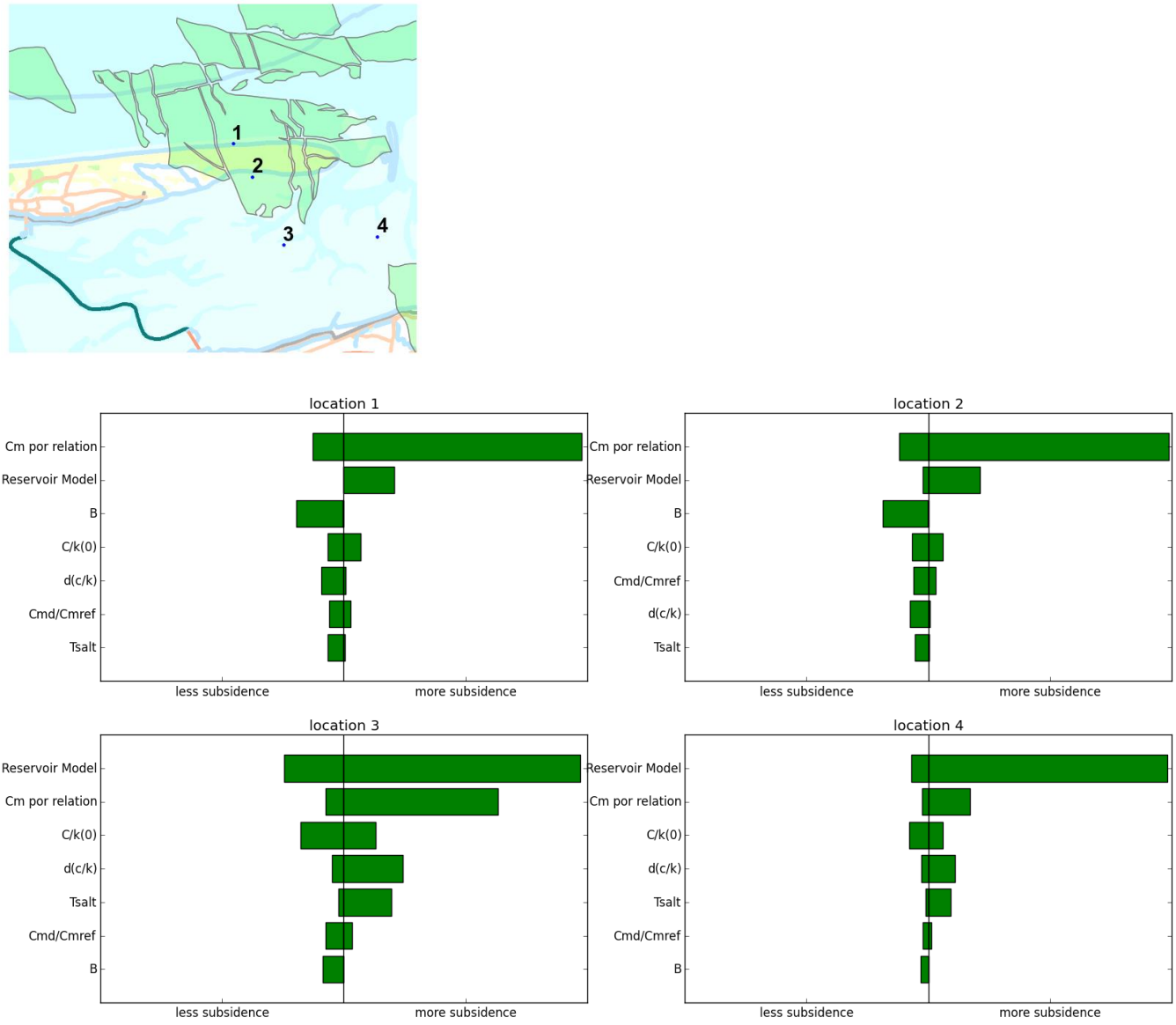


Figure 75. Tornado charts that result from the variation of geomechanical parameters and reservoir scenarios, where value ranges are based on the *a posteriori* results. The x-axis shows the impact of this variation on the subsidence forecast in 2040 compared to the subsidence forecast by the most probable member.

7.7 Results of the confrontation for different time intervals

One of the main aspects of the workflow is to make predictions of the subsidence rates in the future. It is important to demonstrate that the workflow can reliably forecast future subsidence. This is investigated by observing the effect of including more geodetic data (i.e. longer time series) on the posterior result. The confrontation is first performed between models and a limited time series of geodetic data (the first 6 years of the geodetic data). Next, the confrontation is performed on a more complete data set and finally it is performed on the full dataset. It is expected that the use of more data will improve the confidence of the predictions (i.e. more narrow intervals), but the prediction resulting from these two confrontations should not wildly differ, the confidence intervals should just get narrower. To compare the outcomes, the resulting χ^2/N and probability distributions, and the resulting effect on the expected subsidence as well as the changes in the band of accepted membersis presented.

As a first data set, all geodetic data between 1986 and up to the end of 1993 is used, which, using CUPiDO, gives a set of 318 double differences. For a second partial set, the confrontation is performed with all geodetic data between 1986 and the end of 2001, which consists of a set of 654 double differences. The exact same ensemble of members as in Figure 54 was used. The two partial confrontations result in the χ^2/N distributions plotted in Figure 76, together with those from the confrontation with the full set of geodetic data between 1986 and 2017. Note that the value of N differs between the sets, since the number of double differences changes as data is added.

The lowest χ^2/N in the 1986-1993, 1986-2001 and 1986-2017 sets are 1.51, 1.38 and 1.23, respectively. Looking at the scale of the vertical axes in the left plots, it is observed that the χ^2/N values decrease when more data is added. After converting these χ^2/N values to probabilities and member weights, one can plot the subsidence predictions for the three different confrontations, which is shown in Figure 77 for one of the benchmarks (002D0075). The subsidence time-series plotted here demonstrate that the band of accepted ensemble members based on the geodetic data in the period 1986-1993 only, is wider compared to the confrontation based on the longer periods 1986-2001 and 1986-2017. The band of accepted members determined from the data in the period 1986-2017 is slightly larger than in the 1986-2001 results. A possible reason could be the idealisation noise (see paragraph 5.3), which increases in time.

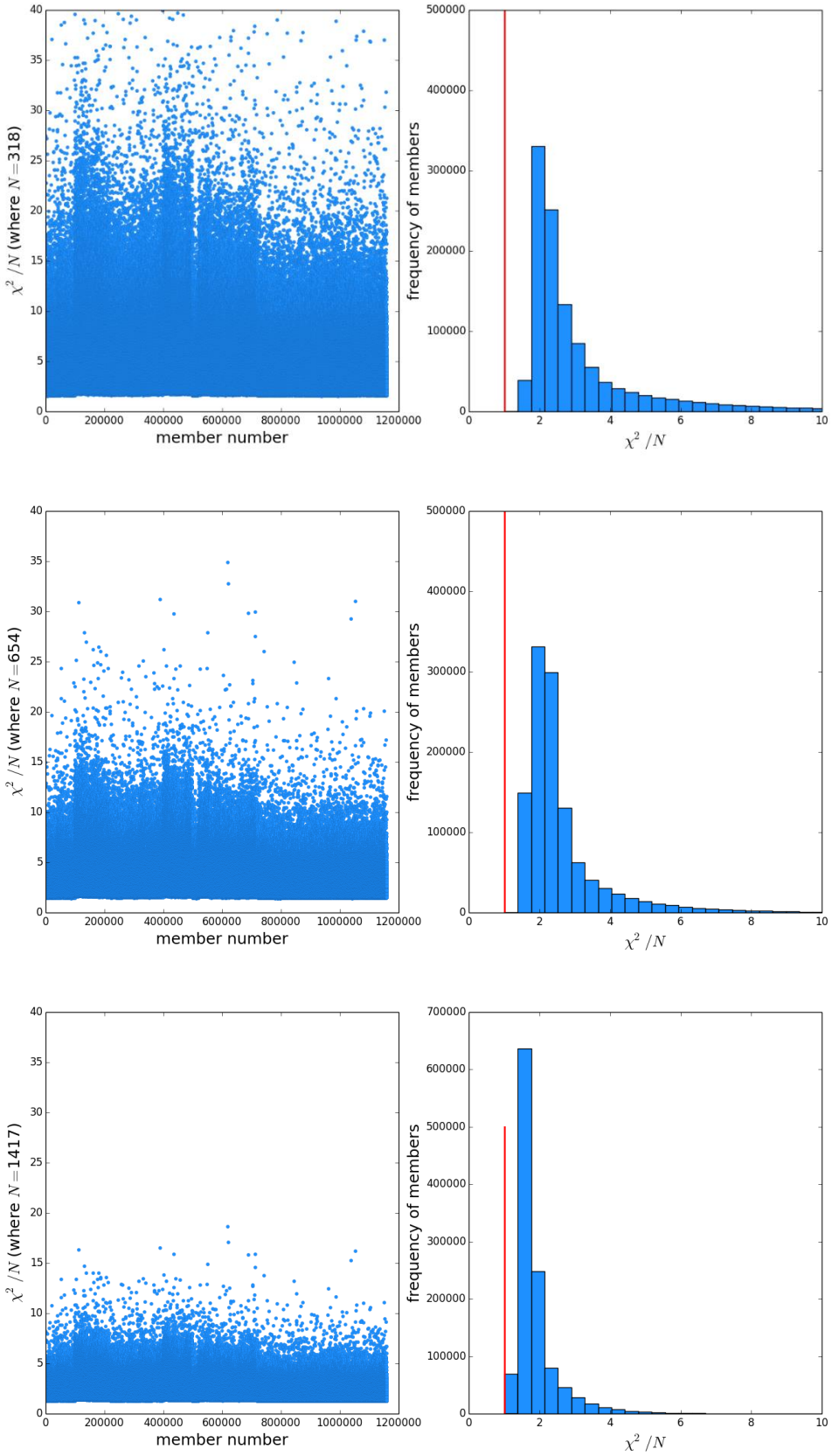


Figure 76. Left: Resulting χ^2/n values. Right: histogram of the χ^2/N values. The red line marks χ^2/N=1. Top: data up to end 1993; middle: data up to end 2001; bottom: full dataset.

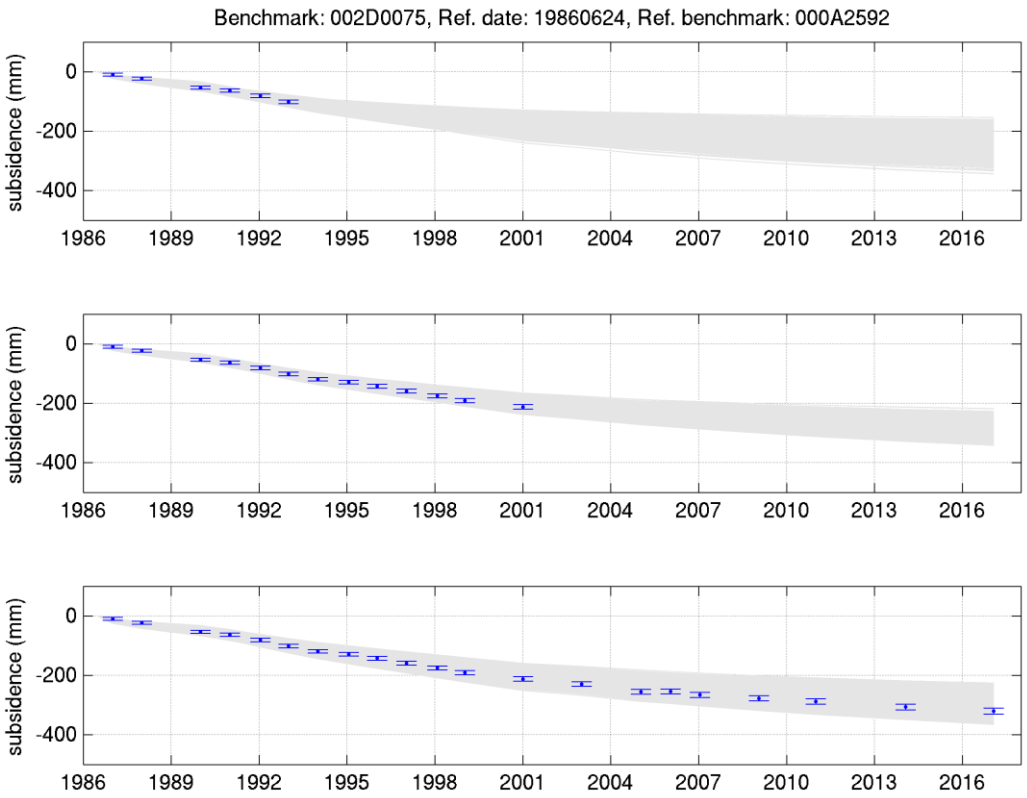


Figure 77. Band of accepted model ensemble members based on confrontation with the geodetic data in the periods 1986-1993, 1986-2001 and 1986-2017 respectively, for the double-differences between benchmarks 000A2592 and 002D0075.

Figure 78 compares the probability of each reservoir scenario; similar to Figure 66, but now for the confrontation in the three different time periods. Although the relative probabilities between reservoir scenarios do change after adding more data to the confrontation, the group of scenarios receiving the highest probabilities remains more or less the same. The general behaviour of the subsidence predictions is, judging from the similarities in probability distributions, not expected to show significant changes over time, as was also observed in Figure 77. The scenario probabilities seem more evenly spread after the confrontation with the least amount of data (1986 – 1993) compared to the confrontation with most data (1986 – 2017), where in the latter case certain scenarios become relatively more probable than others (most notably M9101, M9102 and M9103). In other words, the probability is distributed among fewer, more distinct members as more data is added over time. This quantitatively describes the improvement of the confidence in the predictions with time, as observed in Figure 77.

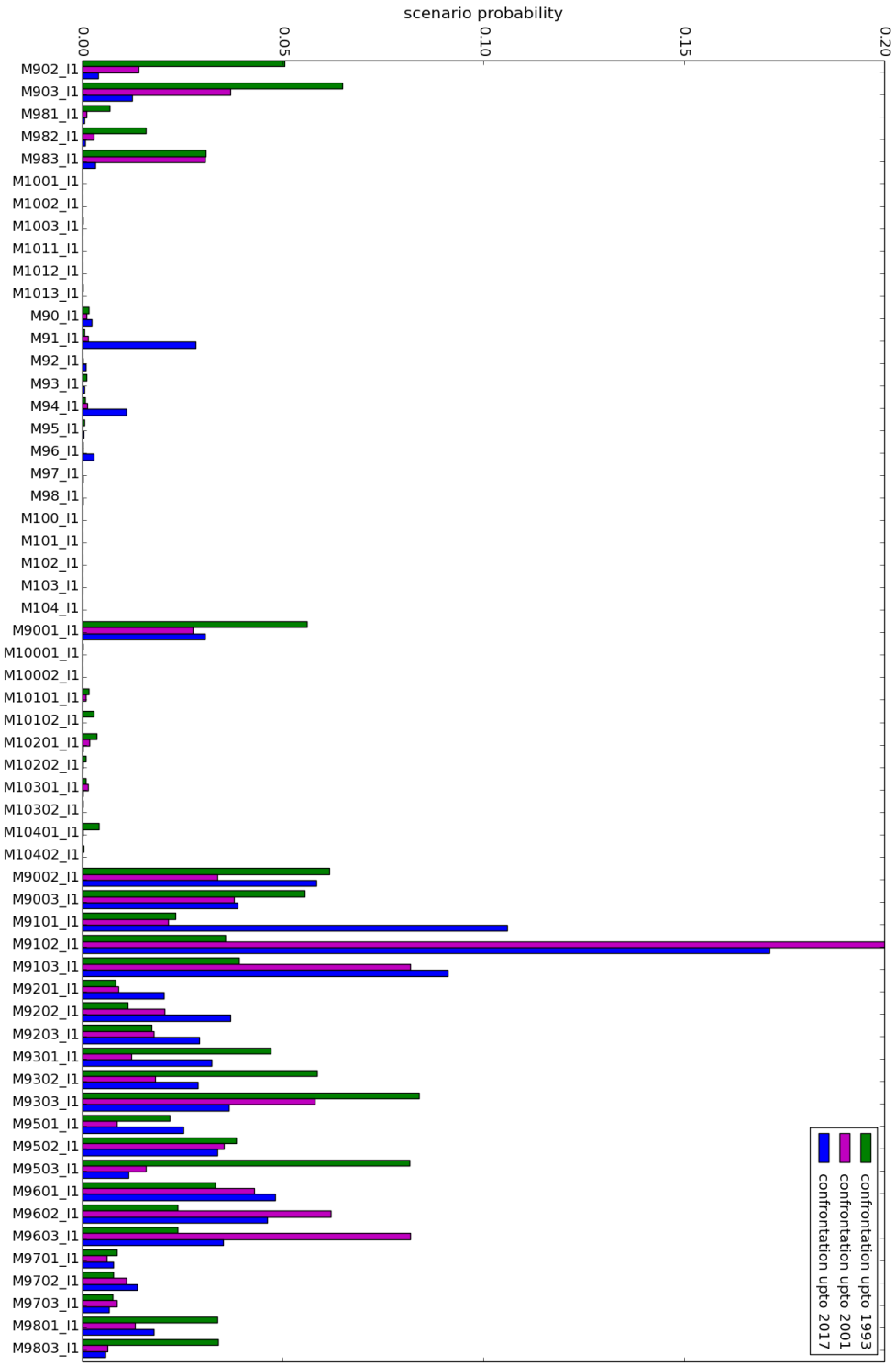


Figure 78: Average reservoir scenario probability after confronting the same ensemble of members with geodetic data in three different time periods (1986 to 1993 in green; 1986 to 2001 in purple; 1986 to 2017 in blue).

From the χ^2/N distributions themselves, one can also observe that the group of most likely ensemble members remains the same for the time periods tested here. A comparison of the χ^2/N values for each member in the full ensemble resulting from the different confrontations is made in Figure 79. In the left figure, a comparison is made between the dataset up to end 1993 to the full dataset and

right a comparison between the dataset up to end of 2001 with the full dataset. Both plots suggest a strong correlation, where poorly-fitting members in one confrontation are also the worst fitting in the other confrontation, and the group of members with lowest χ^2/N values remains the same in all three confrontations. In other words, members that gave a relatively good match to the data in the past, still give a relatively good match after new data has been added.

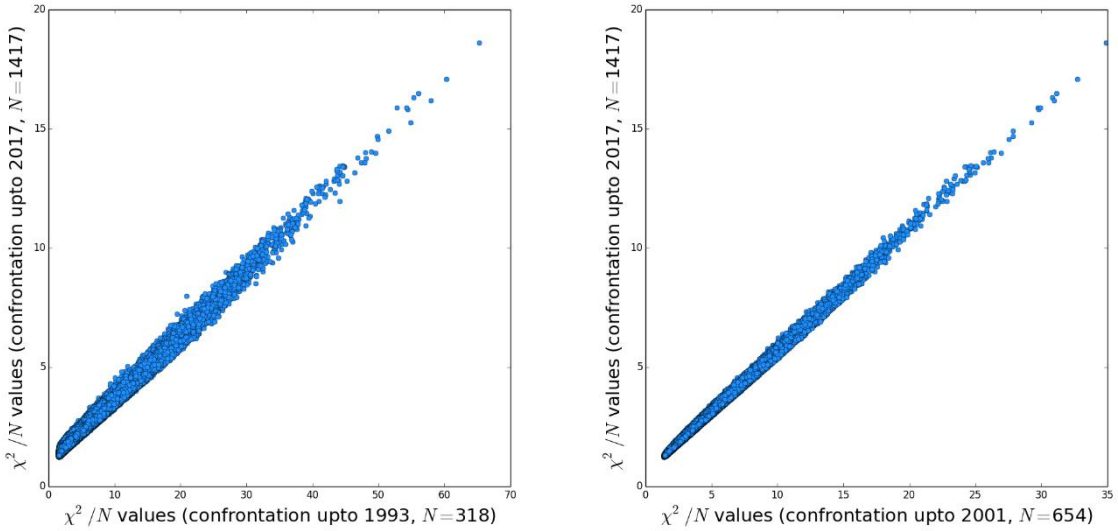


Figure 79. Left: comparison of member χ^2/N values from the confrontation up to the end of 1993 (x-axis) and from the confrontation including the most recent geodetic data from 2017 (y-axis). Right, same result but now for a confrontation up to end 2001.

7.7.1 Application to new fields

A verification of the effect of more geodetic data on the accepted model ensemble members when bringing in longer time series has been described in the previous section. Figure 77 shows a narrowing band specifically in the early monitoring period, i.e. early measurement campaigns quickly narrow the uncertainty.

For a new field, initially, the uncertainty in the longer-term future subsidence can be large, because of the unavailability of constraining subsurface data. This workflow shows that with early measurement campaigns the uncertainty is quickly constrained and therefore the confidence in the prediction of future subsidence improves. Regular geodetic measurement campaigns ensure that subsidence can be maintained within regulatory limits.

8 Conclusions

Results of the LTS-II study demonstrate that a successful workflow was created that confronts subsidence model results with measurements in an objective way. The workflow incorporates the findings of the LTS-I study and as demonstrated here is able to identify the most likely model factors, like reservoir and aquifer depletion scenarios and parameter values for the subsidence model. Moreover, posterior (after confrontation with the geodetic data) probability distributions for the input model data show redefined distributions for the prior input values of the model data. The workflow consists of 5 main components as visualised in Figure 80 and summarised below.

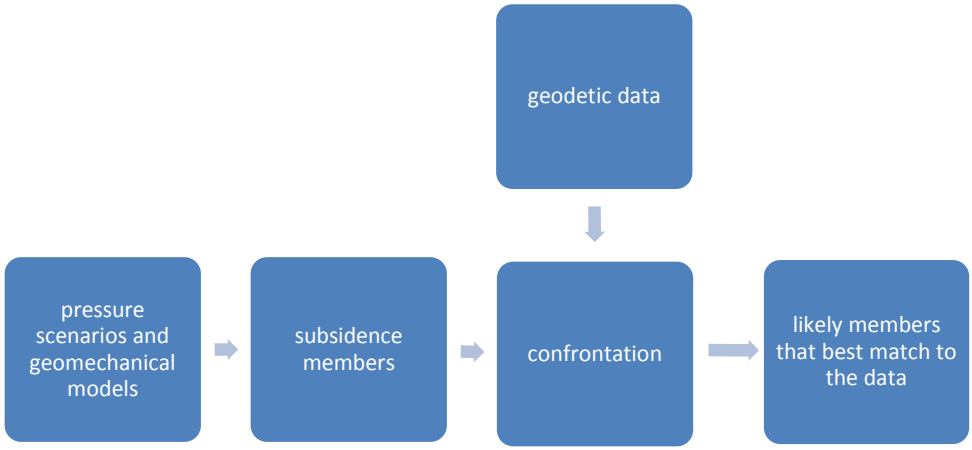


Figure 80. Main components of the confrontation workflow as adopted in this study.

Pressure scenarios

To describe the possible variation in the pressure scenarios for Ameland, 58 reservoir scenarios were created that were all history matched to the available pressure and production data. The scenarios are characterized by different levels of depletion in time of both the western and eastern aquifers. Because of the long production history of the field (since 1986) combined with the regular measurements of the pressures in the producing wells, the development of the pressure in the gas bearing part of the structure is narrowly constrained and therefore shows little variation between the 58 models.

Subsidence members

For every pressure scenario, the parameter values of the compaction model and influence function were varied in a Monte Carlo simulation. In the current study, one generic compaction model is used, the RTCiM model (Pruiksma et al, 2015), with the addition of a linear elastic branch. For the influence function, a modified Geertsma and van Opstal (1973) model is used with the addition of a time dependent shape factor that is calibrated to the viscous behaviour of a salt layer embedded in an elastic overburden, modelled by a finite element model. The two time-dependent processes (one in the compaction and one in the influence function) have a distinctly different effect on the subsidence rate. Subsidence as a result of reservoir compaction always leads to a change of the subsidence bowl *volume*. In contrast to this, the viscous behaviour of the salt only impacts the *shape* of the

subsidence bowl and not the volume.
Each parameter picked by the Monte Carlo procedure results in a subsidence model member, with a group of members defining the ensemble.

Geodetic data
Several significant innovations were applied in the use of geodetic data. Observations from levelling and GPS techniques have been used in the format of spatio-temporal (i.e. in space and time) double-differences. Uncertainties are described by a fully populated covariance matrix, that also takes shallow movements into account (idealisation noise). Outlier removal has been implemented in a formal way.

Confrontation
Each member is confronted to the geodetic data. The resulting test statistic defines the probability and weight of the specific member. The used theory to calculate the test statistic is based on Nepveu et al. (2010) with modifications, explained in this report.

Output and results
The modelled results versus the geodetic data are presented for two E-W profiles over the island and for many individual double-differences. The overall fit to the data is good, matching better to the location of the deepest point of the bowl and matching better to the benchmarks on the eastern part of the island as compared to results from earlier models. The weighted average was also used to calculate the subsidence rates in the Pinkegat area along with their confidence intervals. It is observed that the confidence interval is narrow and is, at present, far below the defined boundary for the rates (called the “meegroeivermogen”).

The same workflow has been applied to two “emergency stop scenarios”. The results show a clear reduction in the subsidence rates after a hypothetical stop in 1996, but also smaller but clear reduction following a hypothetical stop in 2016. The latter being smaller can be explained by the reduced effect of a later stop on the pressures in the gas field, knowing that most of the gas has been produced by 2016. For a hypothetical stop in 1996, the production stop is ten years after the start of production, which is comparable to the current field life of the Waddenzee fields like Nes and Moddergat. Based on this analogue it can therefore be concluded that an emergency stop scenario in the Waddenzee for fields like Nes and Moddergat would result in a significant decrease of the subsidence rate in the Waddenzee.
A verification of the effect of more geodetic data on the accepted model ensemble members when bringing in longer time series has also been performed as part of this study. The results show a narrowing band specifically in the early monitoring period, i.e. early measurement campaigns quickly narrow the uncertainty.

More specific to the 6 points that are addressed in the letter by SodM (2017), the actions taken in this LTS-II study and the results related to these are listed in Table 20.

Table 20 List of requested improvements by SodM (2017) and actions taken including pointers to the relevant sections

	Issue addressed by SodM	Actions and results	More information
1	Improve the fit of the subsidence models to the measured data	• 58 new reservoir scenarios. All history matched against the pressure data. • Improved geomechanical model resulting in better fit. • Improved procedure for geodetic data outlier removal.	Section 3.2, Chapter 4, Section 5.4
2	Include the effect of the large observed overpressure in the subsurface model	• A linear elastic branch has been added to the RTCiM. Elastic compaction will take place as long the reservoir or aquifers are overpressured.	Section 4.2.2
3	Use a single generic compaction model instead of 4 specific compaction models	• Only the versatile RTCiM is used in this study, modified by the addition of a linear elastic branch.	Section 4.2.1
4	Use a realistic range of salt parameter values	• A range for salt viscosity values is based on laboratory measurements and model results described in the literature.	Section 4.3.1
5	Take into account the covariances and its effect on the uncertainty band of forecasted subsidence	• The covariance matrix has been extended with a component that describes the spatio-temporal correlations. • The role of the spatio-temporal component in the covariance matrix has been investigated. • The effects are visualised in multiple figures comparing subsidence model results with geodetic data.	Chapter 6, Section 7.2, Section 7.4, Appendix 3
6	Resolve geomechanical modelling issues that were introduced by the TNO program AEsubs	• The program AEsubs has been replaced by a modified version of the Geertsma & van Opstal solution.	Section 4.3

9 References

Berest, P., Jean Francois Beraud, J.F., Bourcier, M., Dimanov, A., Gharbi, H. (2012) Very slow creep tests on rock samples. Pierre Berest, Mehdi Ghoreychi, Faouzi Hadj Hassen, Michel Tijani. Mechanical Behavior of Salt VII, Apr 2012, France. Taylor & Francis Group, pp.81-88, 2012. https://hal-mines-paristech.archives-ouvertes.fr/hal-00786923/document

Bjerrum, L., 1967, Engineering Geology of Norwegian Normally Consolidated Marine Clays as Related to the Settlements of Buildings, Geotechnique, 17, pp 81-118.

Breunese, J.N., Eijs, R.M.H.E. van, Meer, S. de & Kroon, I.C. (2003) Observation and prediction of the relation between salt creep and land subsidence in solution mining - The Barradeel case. Solution Mining Research Institute Fall meeting 2003, October 5-8, Chester, United Kingdom, abstract.

Cai, Y., Verdel, T., Deck, O. (2014) On the topography influence on subsidence due to horizontal underground mining using the influence function method. Computers and Geotechnics 61. Pp. 328–340

De Waal, J.A. (1986) On the rate type compaction behaviour of sandstone reservoir rock. doctoral thesis.
http://repository.tudelft.nl/islandora/object/uuid:b805782b-2eb4-4f72-98f4-f727c4ea9df0?collection=research

De Waal, J.A. and Smits, R.M.M. (1985) Prediction of reservoir compaction and surface subsidence-field application of a new model. Shell report 7.21.180

De Waal, J.A., van Thienen-Visser, K., Pruiksma, J.P. (2015) Rate type isotach compaction of consolidated sandstone 49th US Rock mechanics/Geomechanics symposium (ARMA), San Franscisco, 28 June - 1 July, Volume: ARMA 15-436

den Haan, E.J., 2003 Het a,b,c-isotachenmodel: hoeksteen van een nieuwe aanpak voor zettings - berekeningen, Geotechniek (4) 28-35
http://www.delftgeosystems.nl/files/files_org/MSettle_abc_isotachenmodel.pdf

Fokker, P.A. (1995) The behaviour of salt and salt caverns. Doctoral Thesis.
http://repository.tudelft.nl/islandora/object/uuid:6847f8e4-3b09-4787-be02-bcce9f0eed06?collection=research

Fokker, P.A., Orlic, B. (2006) Semi-Analytic Modelling of Subsidence. Mathematical Geology, Vol. 38, No. 5, July 2006. DOI: 10.1007/s11004-006-9034-z

Fokker, P.A., Wassing, B.B.T., van Leijen, F.J., Hanssen, R.F., Nieuwland, D.A. (2016) Application of an ensemble smoother with multiple data assimilation to the Bergermeer gas field, using PS-InSAR. Geomechanics for Energy and the Environment, Vol. 5, No. March, 2016, p. 16-28.

Fokker, P.A. and Van Thienen-Visser, K. (2016) Inversion of double-difference measurements from optical levelling for the Groningen gas field, International Journal of Applied Earth Observation and Geoinformation, 49: 1-9.

Geertsma, J. (1973) , Land Subsidence Above Compacting Oil and Gas Reservoirs, J. Petr. Tech., pp.734-744.

Geertsma, J. and van Opstal, G. (1973). A Numerical Technique for Predicting Subsidence Above Compacting Reservoirs, Based on the Nucleus of Strain Concept. Verh. Kon. Ned. Geol. Mijnbouwk. Gen., 28, pp. 63-78.

Hettema, M., Papamichos, E., Schutjens, P. (2002). "Subsidence delay: Field observations and analysis". Oil & Gas Sci. Technol. Rev. IFP 57, 443–458.

Houtenbos, A.P.E.M., F. Kenselaar (2001): Peilmerk hoogte variaties – stochastische analyse van peilmerkbeweging in Nederland. TUDelft, 2001.

IFG (2006) Rock Mechanical Investigations on Rock Salt from the Cavern Well BAS4

Marketos, G., C. J. Spiers, and R. Govers (2016) Impact of rock salt creep law choice on subsidence calculations for hydrocarbon reservoirs overlain by evaporite caprocks, J. Geophys. Res. Solid Earth, 121, 4249–4267, doi:10.1002/2016JB012892.

Marketos, G., Broerse, D.B.T., Spiers, C.J. and Govers, R. (2015) Long-term subsidence study of the Ameland gas field: time-dependence induced by rocksalt flow

McKay, M.D.; Beckman, R.J.; Conover, W.J. (2000). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics, Vol. 42, No. 1,

Ministerie van Economische Zaken(2016), Wijziging instemming gewijzigd winningsplan gaswinning Moddergat, Lauwersoog en Vierhuizen, DGETM-EO / 16110969

Mossop, A., 2012. An explanation for anomalous time dependent subsidence, 46th US Rock Mechanics/Geomechanics Symposium, Chicago, Illinois, USA, 24–27 June 2012, 12-518.

NAM (2010) Subsurface Technical Report: Static Reservoir Modelling of Ameland Fields for Subsidence Prediction.

NAM (2011) Subsurface Technical Report: Subsidence Modeling of Ameland Fields. UIE Report No: EP201105208617

NAM (2012) Gaswinning Moddergat, Lauwersoog, Vierhuizen: Actualisering meet- en regelprotocol n.a.v. wijziging winningsplannen 2011, EP200612202112

NAM (2015) Bodemdaling door aardgaswinning NAM-velden in Groningen, Friesland en het noorden van Drenthe. Status rapport 2015 en Prognose tot het jaar 2080. EP201511213444

NAM (2015a) Wadden Sea Long term Subsidence Studies –Overview report. EP201506209625 http://feitenencijfers.namplatform.nl/download/rapport/2ca6c8d8-c0d4-4c10-8460-672f93b4cdaa?open=true

NAM (2016) RTCiM calibration to Moddergat-3 coreplug experiments. NAM presentation to TNO-AGE and SodM

NAM (2016a) Gaswinning vanaf de locaties Moddergat, Lauwersoog en Vierhuizen. Resultaten uitvoering Meet- en regelcyclus 2015. http://www.commissiemer.nl/projectdocumenten/00000929.pdf?documenttitle=MR%20rapportage%20%202015.pdf

NAM (2016b) Gaswinning vanaf de locaties Moddergat, Lauwersoog en Vierhuizen: Integrale beoordeling en samenvatting van de monitoringresultaten over 2015. http://feitenencijfers.namplatform.nl/download/rapport/dcf5fe13-a1aa-4dea-84e5-dff484326f68?open=true

NAM (2016c) Technical Addendum to the Winningsplan Groningen 2016 Production, Subsidence, induced Earthquakes and Seismic Hazard and Risk Assessment in the Groningen Field.Part II-Subsidence https://www.sodm.nl/binaries/staatstoezicht-op-de-mijnen/documenten/publicaties/2016/06/21/4--technical-addendum-to-the-winningsplan-groningen-2016-chapter-6/4-technical-addendum-to-the-winningsplan-groningen-2016-chapter-6-15042016.pdf

NAM (2016d) Werkplan Long Term Subsidence Fase II (LTS II). Bijlage 4 in Gaswinning vanaf de locaties Moddergat, Lauwersoog en Vierhuizen Resultaten uitvoering Meet- en regelcyclus 2015. EP201602210605

NAM (2017) Ensemble Based Subsidence application to the Ameland gas field – long term subsidence study part two (LTS II). NAM Doc nr. EP201701217189

NAM (2017a) Addendum to report “Ensemble Based Subsidence application to the Ameland gas field – long term subsidence study part two (LTS II)”. NAM Doc nr. EP201703202079

NAM (2017c) LTS-II double-difference time series visualisations. NAM Doc nr. EP201710206381

Nepveu M., Kroon, I. C., Fokker, P. A. (2010) Hoisting a Red Flag: An Early Warning System for Exceeding Subsidence Limits. Mathematical Geosciences. February 2010, Volume 42, Issue 2, pp 187–198

Pruiksma, J.P. & Breunese, J.N. & Thienen-Visser, Karin & De Waal, Hans. (2015). Isotach formulation of the rate type compaction model for sandstone. International Journal of Rock Mechanics and Mining Sciences. 78. 127-132. 10.1016/j.ijrmms.2015.06.002.

Samiei-Esfahany S., H. Bähr (2015): Research and Development Project for Geodetic Deformation Monitoring – Contribution to the research project: “Long-term study on anomalous time-dependent subsidence in the Wadden Sea Region”. NAM report EP201505216980. http://feitenencijfers.namplatform.nl/download/rapport/6af45eca-b02f-4844-8adf-7835b5e0ee55?open=true

Samiei-Esfahany S., H. Bähr (2017): Erratum to the report: ”Research and Development project for Geodetic Deformation Monitoring” – Revision of recommendations from the project: “Long-term study on anomalous time-dependent subsidence in the Wadden sea region”. NAM report EP201701215912.
https://nam-feitenencijfers.data-app.nl/download/rapport/4f112b38-d303-4329-943e-d3f5952e5555?open=true

Snyder C, Bengtsson T, Bickel P, Anderson J. (2008) Obstacles to high-dimensional particle filtering. Monthly Weather Review 136: 4629-4640.

SodM (2017) Last onder dwangsom LTS-studie. Kenmerk 17106531

Spiers, C.J., Urai, J.L., Lister, G.S., Boland, J.N., Zwart, H.J.,(1986) The influence of fluid–rock interaction on the rheologyof salt rock. EUR 10399 EN, Office for Official Publications of the European Communities, Luxembourg, 131 pp.

Spiers, C.J. & Carter, N.L.(1998) Microphysics of rocksalt flow in nature. In: M. Aubertin & H.R.Hardy (Eds.) The Mechanical Behaviour of Salt: Proceedings of the Fourth Conference Series on Rock

Tarantola A. (2005): Inverse Problem theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics.

TCBB (2009): Methode Houtenbos. In: Van meting naar daling: bodemdaling door delfstofwinning. Technische Commissie Bodembeweging, 2009, section 4.2.

Teunissen P. J. G. (2000a): Adjustment Theory. VSSD, Delft.

Teunissen P. J. G. (2000b): Testing Theory. VSSD, Delft.

TNO (2011) Toetsing van de belasting op de gebruiksruimte in de kombergingsgebieden in Pinkegat en Zoutkamperlaag door bodemdaling ten gevolge van gaswinning onder de Waddenzee. TNO-060-ur-2011-020351c
http://nlog.nl/bodemdaling

TNO (2013) A general framework for rate dependent compaction models for reservoir rock. TNO 2013 R11405

TNO (2013a), Toetsing van de bodemdalingsprognoses en seismische hazard ten gevolge van gaswinning van het Groningen veld, 23 dec 2013, Thienen-Visser et al. TNO report R11953

TNO (2017) Ensemble-based Subsidence Interpretation and Prediction ESIP: Technical Reference Manual. TNO report-R11278

Van der Wal, O. & van Eijs, R., 2016. Subsidence inversion on Groningen using levelling data only. NAM report EP201612206045.
http://feitenencijfers.namplatform.nl/download/rapport/30346ae1-8443-452f-96a6-314ee5c91d41?open=true

van Keken, P.E., Spiers, C.J., van den Berg, A.P., Muyzert, E.J. (1993) The effective viscosity of rocksalt: implementation of steady-state creep laws in numerical models of salt diapirism, In Tectonophysics, Volume 225, Issue 4, 1993, Pages 457-476, ISSN 0040-1951,

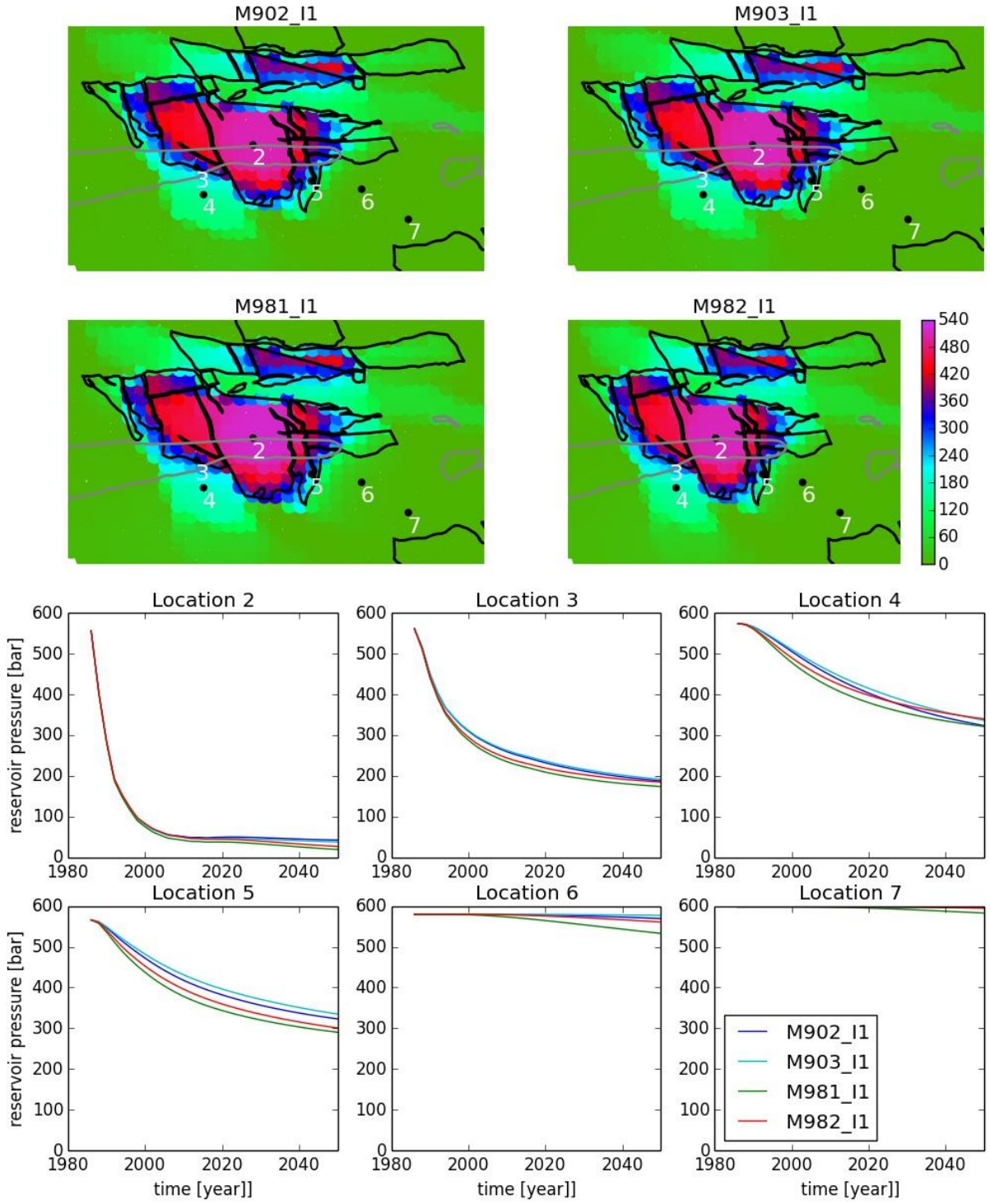
Van Leijen F., Samiei-Esfahany, S., van der Marel, H., Hanssen, R. F. (2017): Uniformization of Geodetic data for deformation analysis – Contribution to the research project: Second phase of the long-term subsidence study in the Wadden Sea Region (LTS2). Delft University of Technology. https://nam-feitenencijfers.data-app.nl/download/rapport/c8030e3d-2873-4003-a5ae-b99c576775aa?open=true

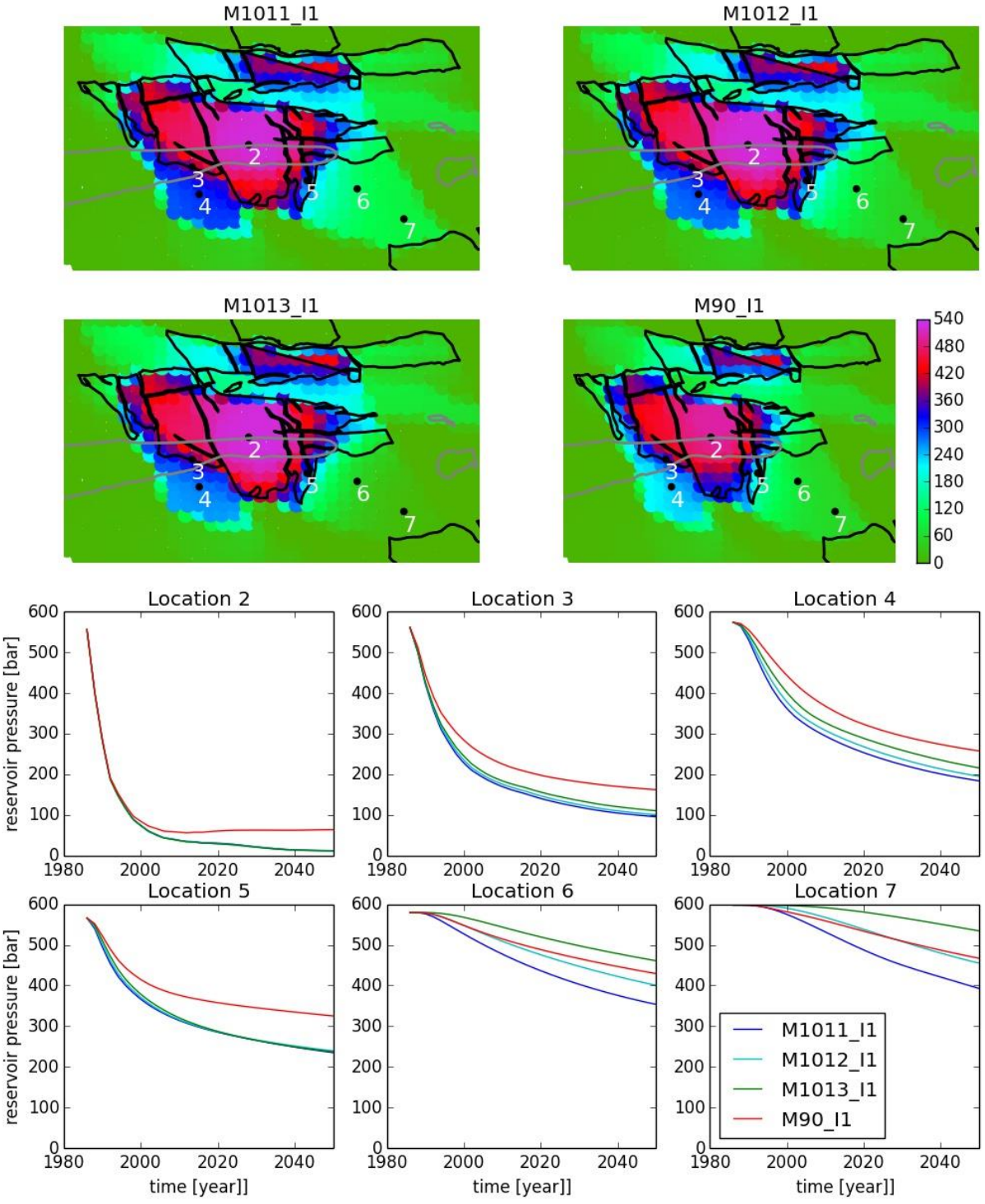
Van Thienen-Visser, K., Pruiksma, J.P. and Breunese, J.N., Compaction and subsidence of the Groningen gas field in the Netherlands, doi: 10.5194/PHIAS 372-367-2015

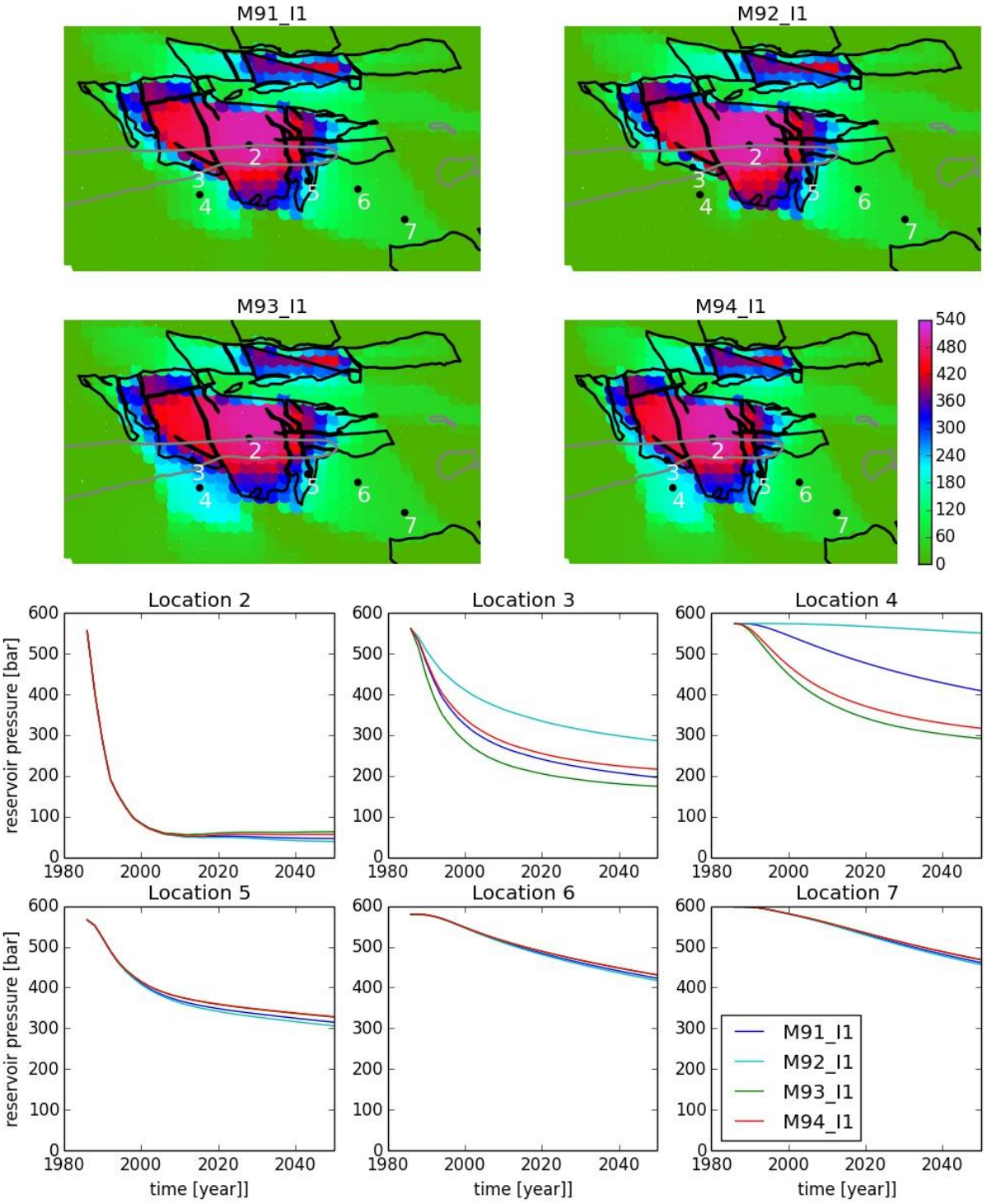
Williams S. (2015): Description of GPS uncertainties within the Long Term Study on Anomalous Time-Dependent Subsidence. National Oceanographic Centre, Liverpool, UK. https://nam-feitenencijfers.data-app.nl/download/rapport/0405dcce-70fa-4dab-86e8-f61933aac6e2?open=true

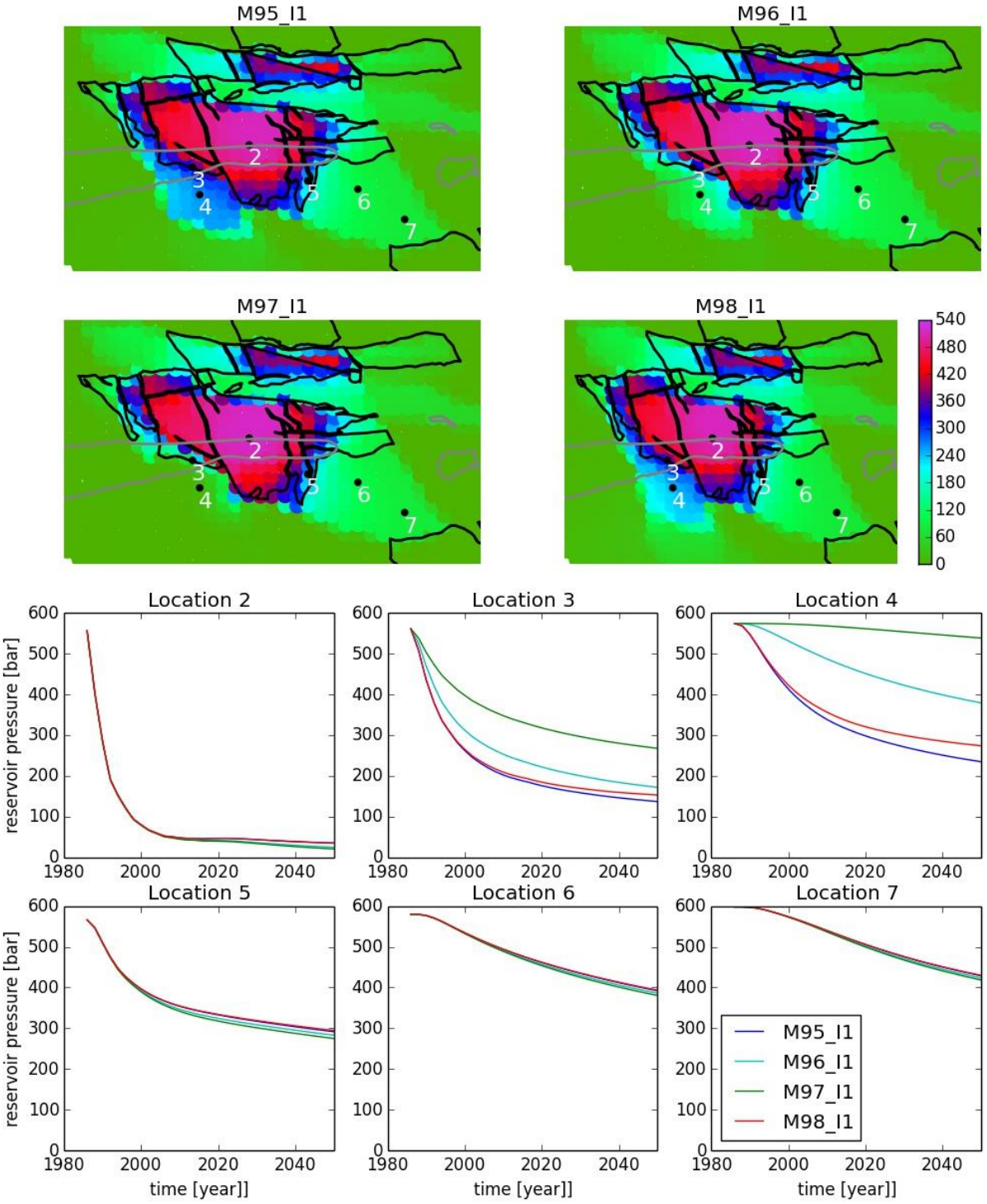
Appendix 1 Overview of the pressure scenarios

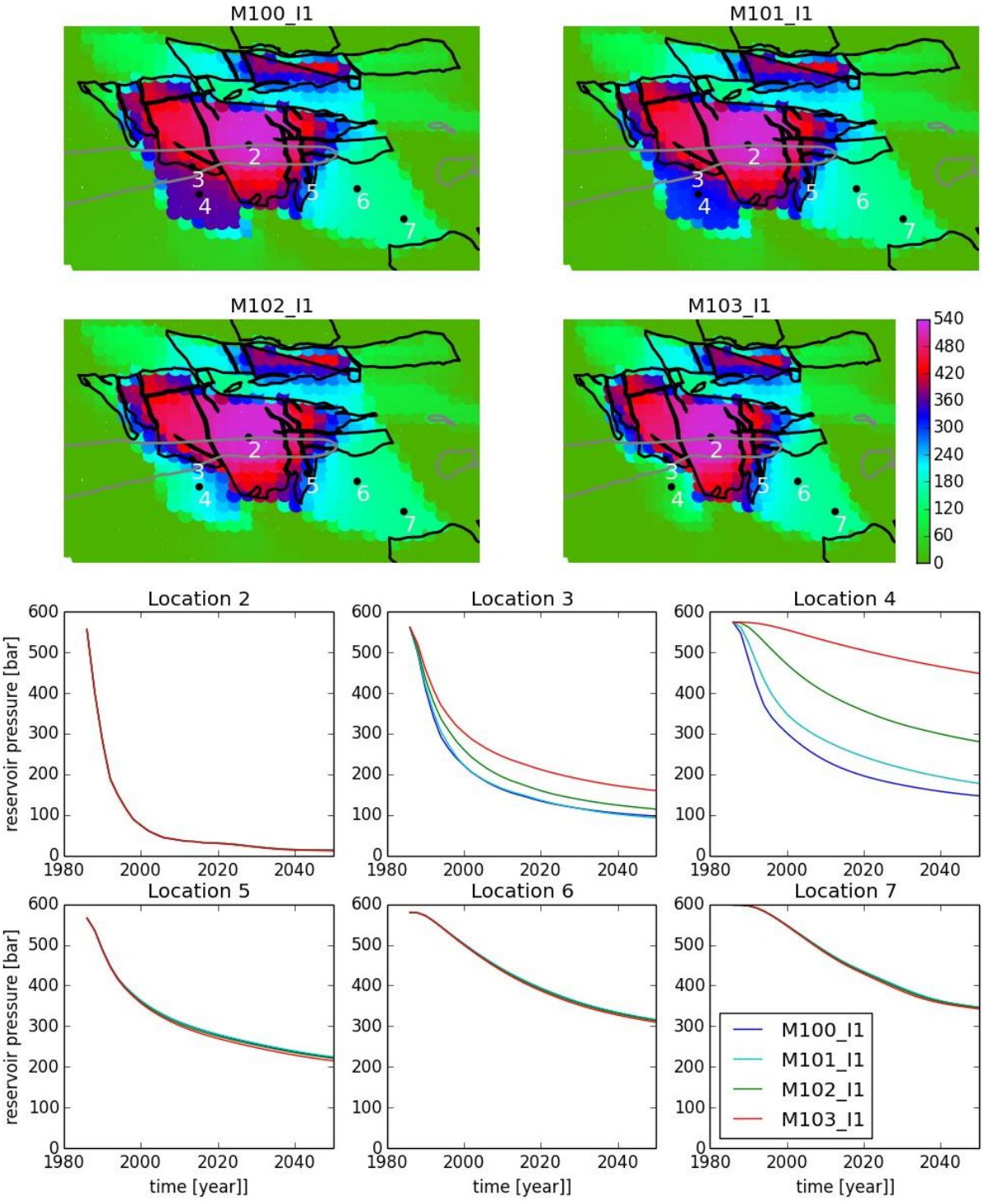
Following plots show the depletion of all reservoir models. The top part shows the areal depletion in 2016, the bottom part show the depletion in time for these reservoir models at different locations.

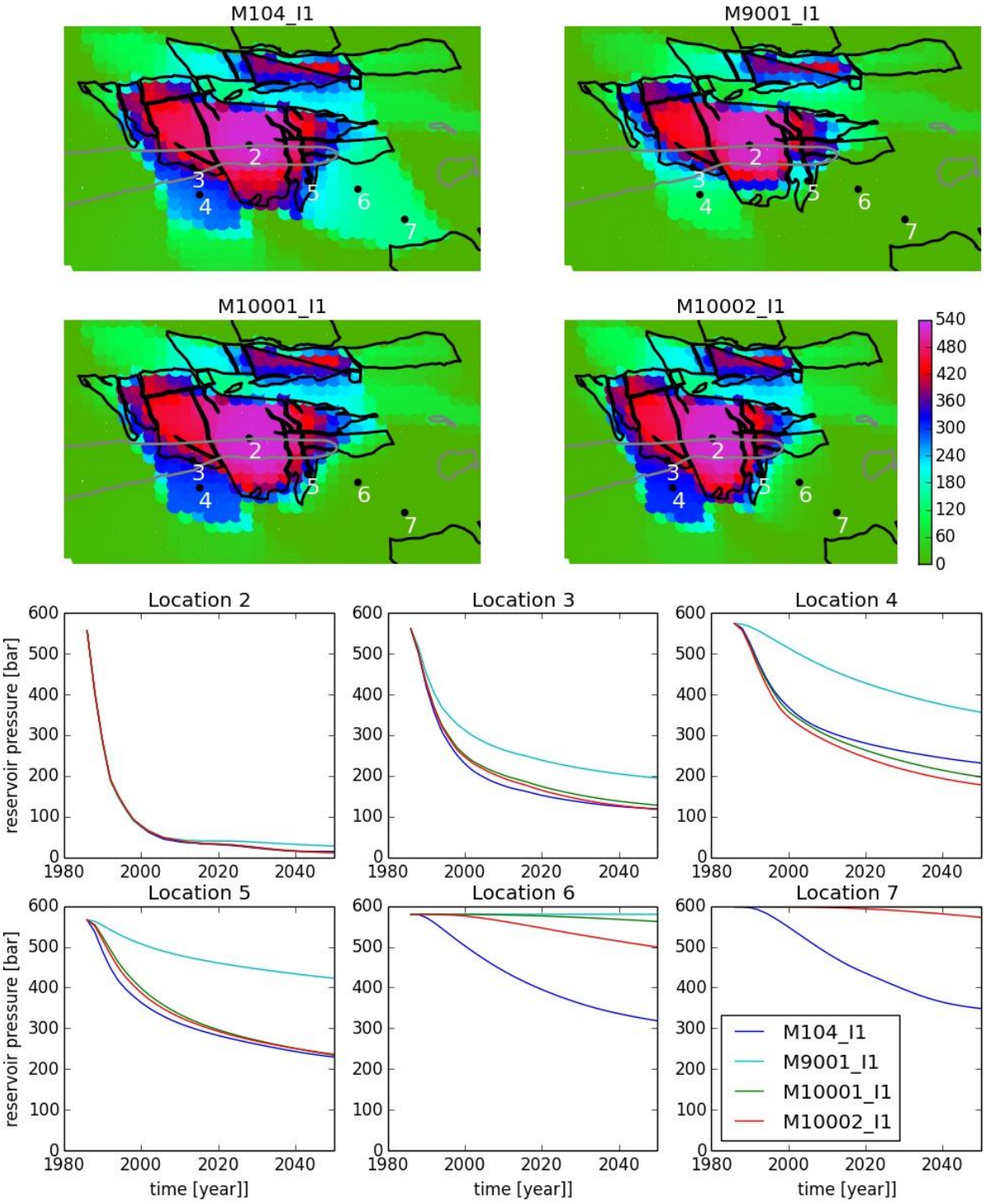


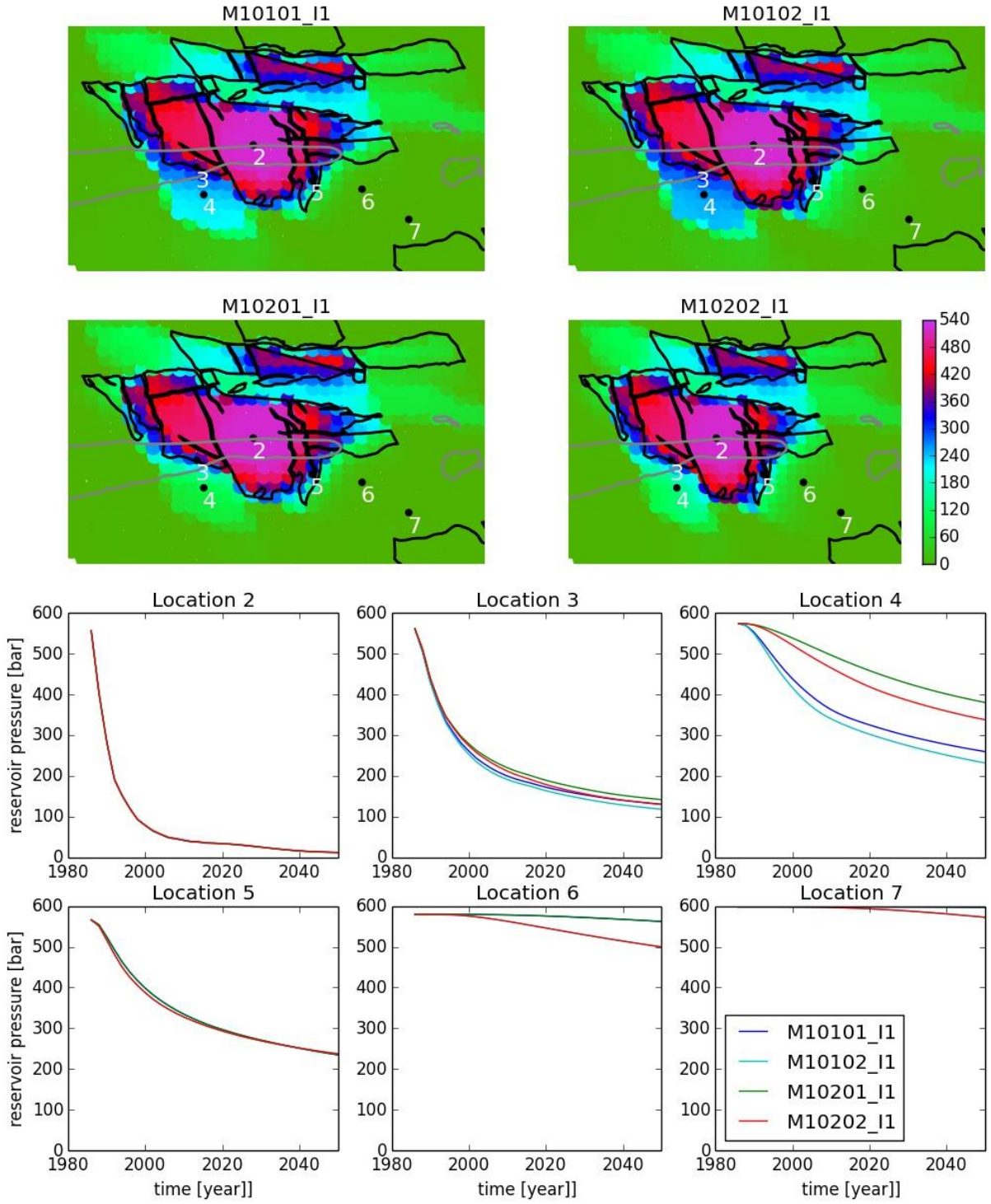

M1011_I1
M1012_I1
M1013_I1
M90_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M1011_I1
M1012_I1
M1013_I1
M90_I1

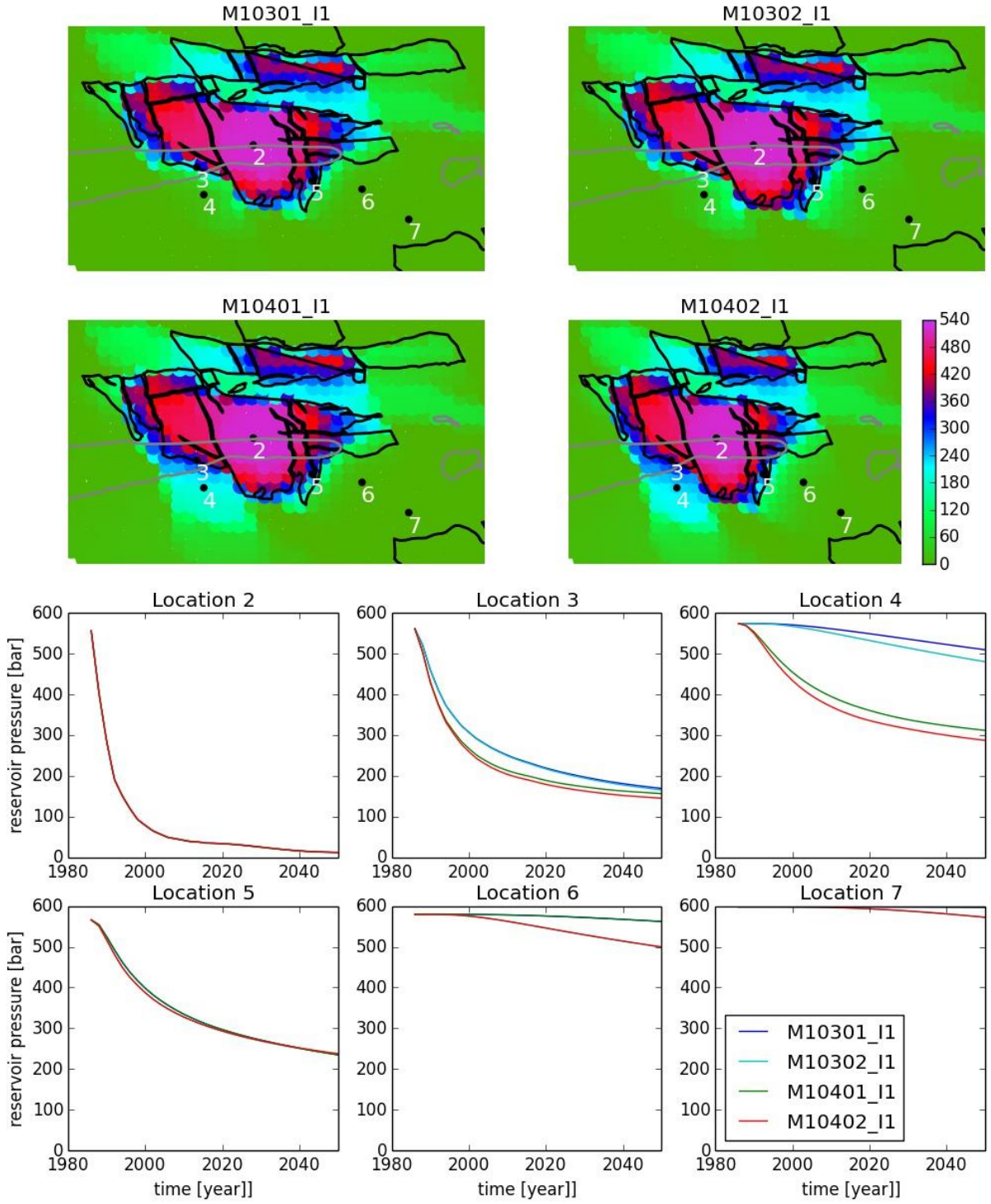

M91_I1
M92_I1
M93_I1
M94_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M91_I1
M92_I1
M93_I1
M94_I1

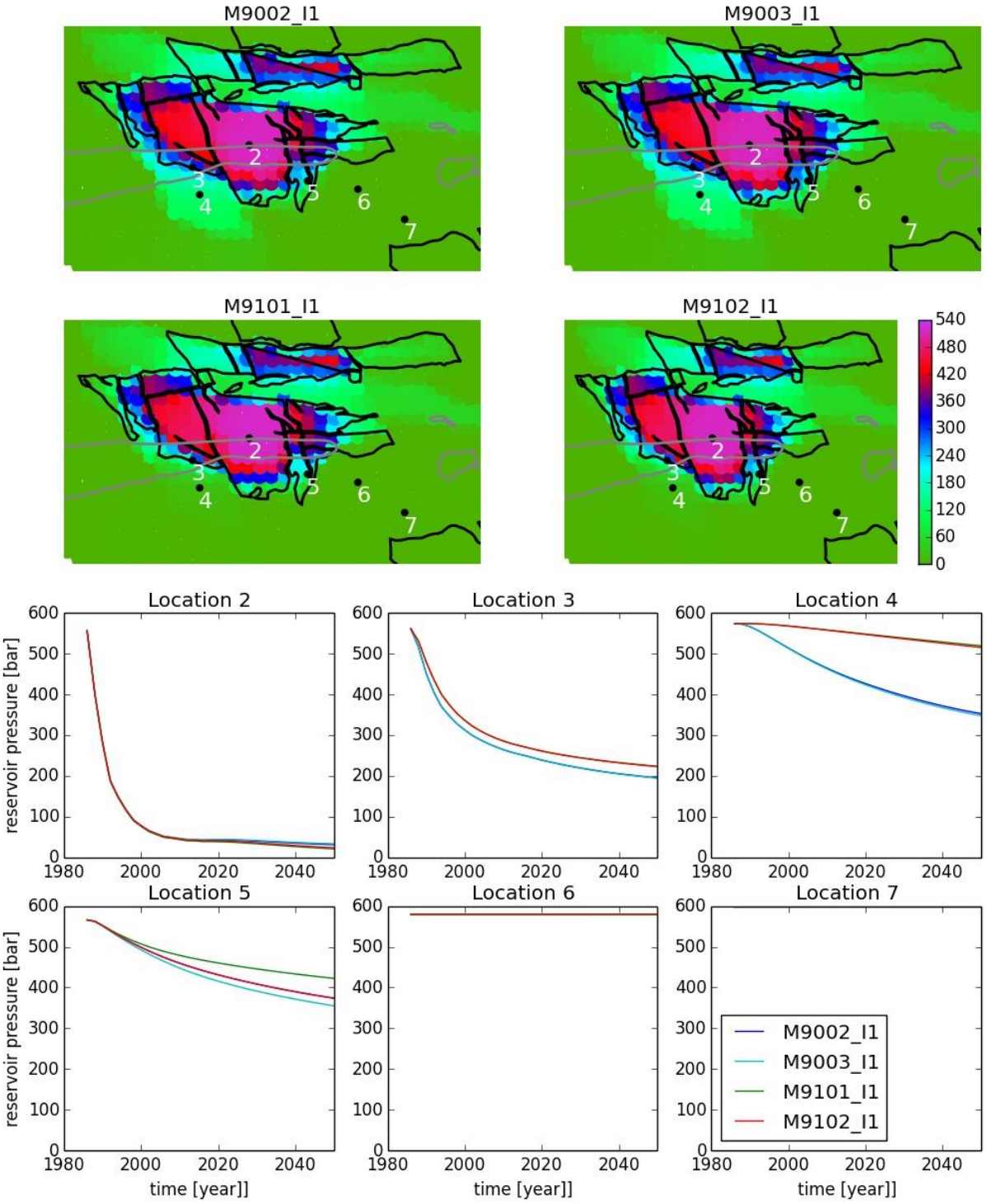

M95_I1
M96_I1
M97_I1
M98_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
0
100
200
300
400
500
600
1980
2000
2020
2040
M95_I1
M96_I1
M97_I1
M98_I1

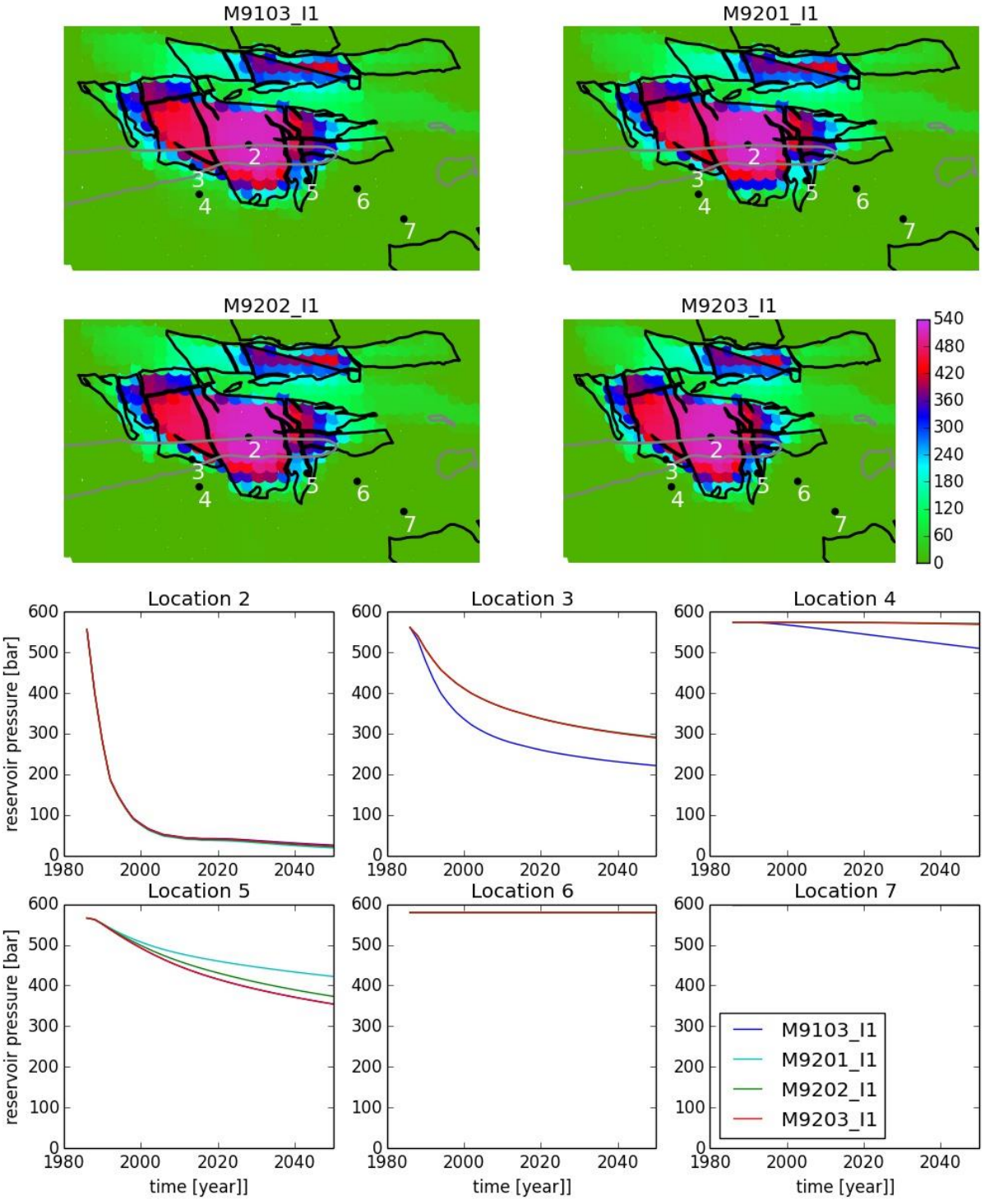

M100_I1
M101_I1
M102_I1
M103_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
0
100
200
300
400
500
600
1980
2000
2020
2040
M100_I1
M101_I1
M102_I1
M103_I1

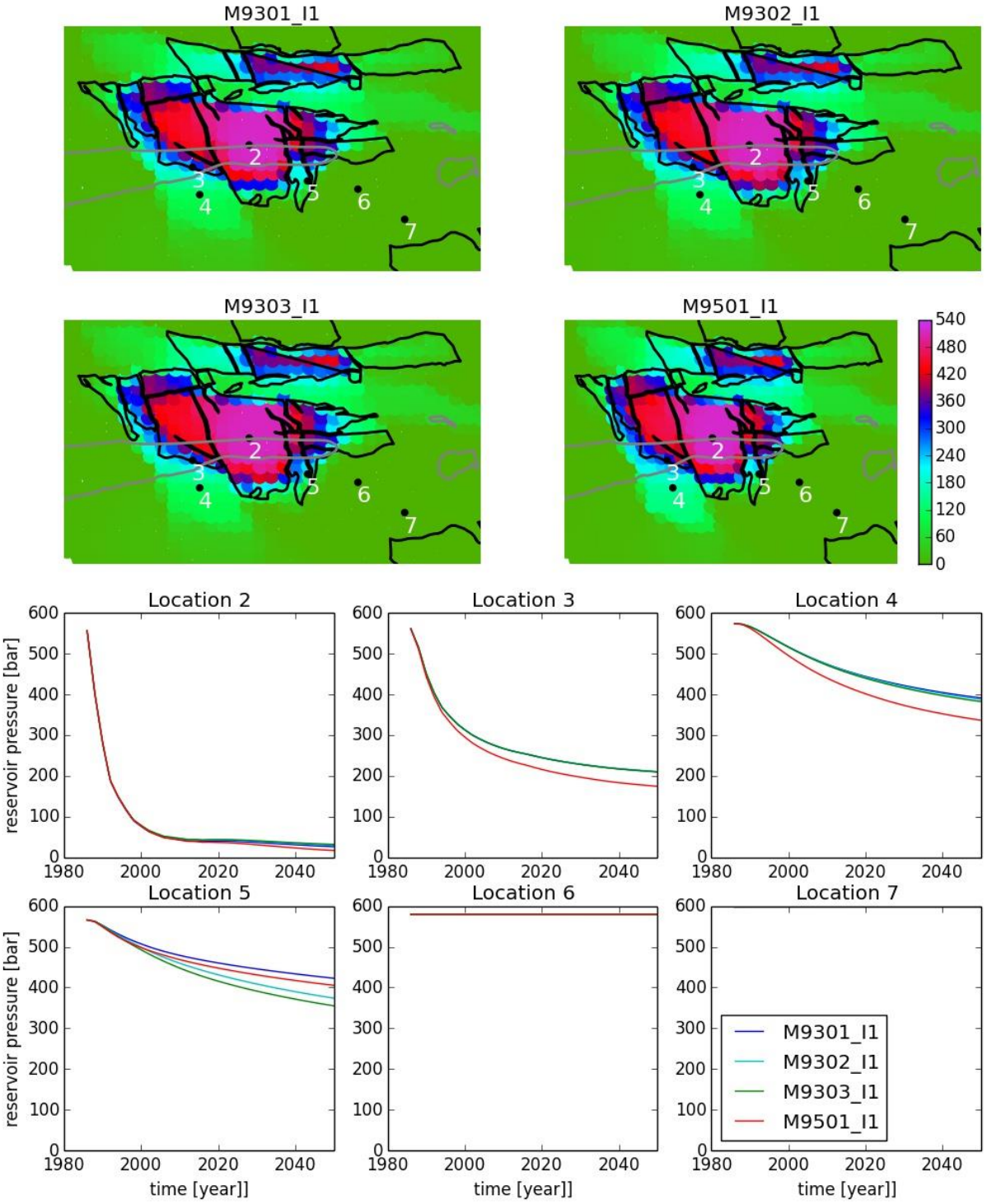

M104_I1
M9001_I1
M10001_I1
M10002_I1
1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
0
100
200
300
400
500
600
1980
2000
2020
2040
M104_I1
M9001_I1
M10001_I1
M10002_I1

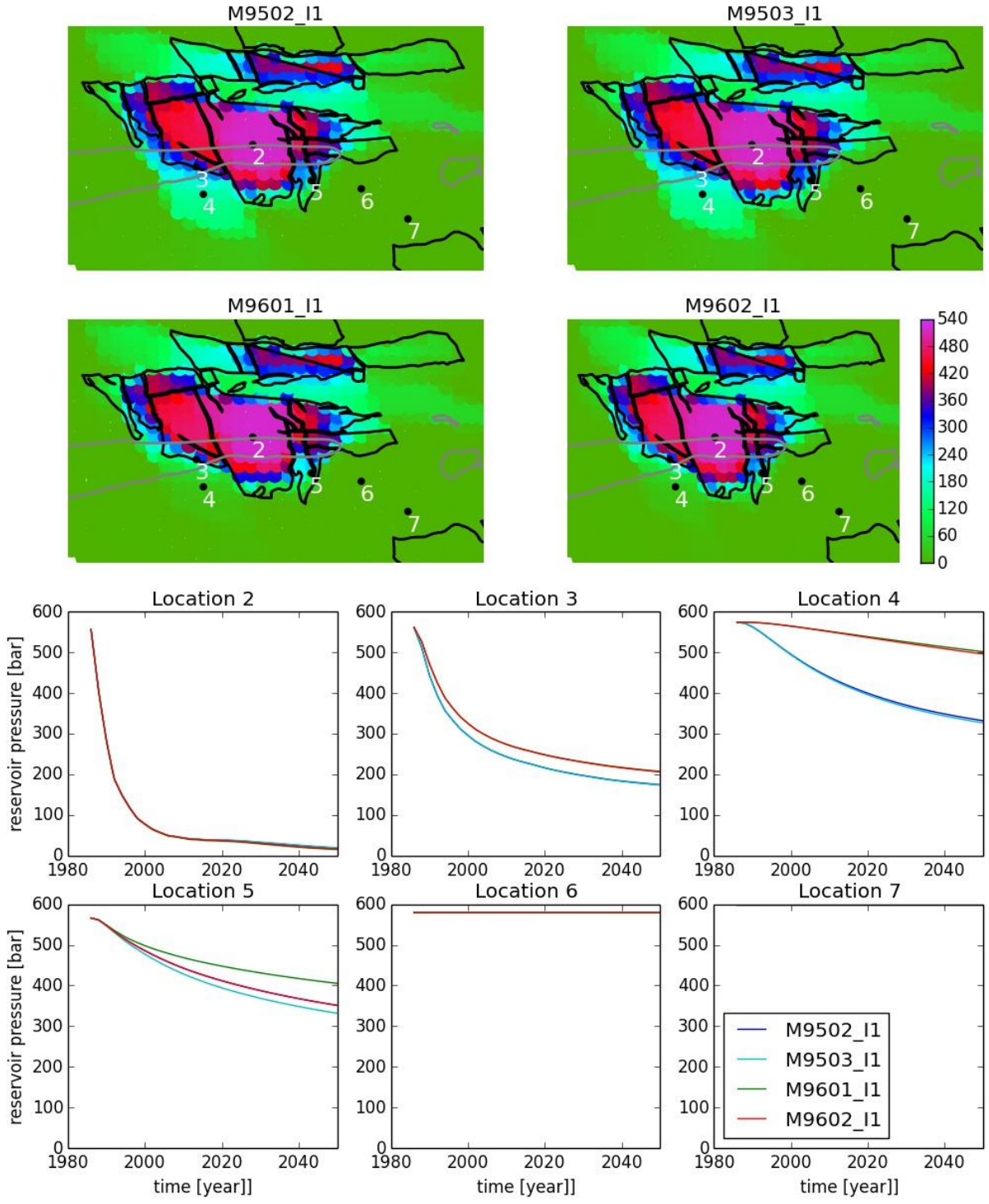

M10101_I1
M10102_I1
M10201_I1
M10202_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M10101_I1
M10102_I1
M10201_I1
M10202_I1

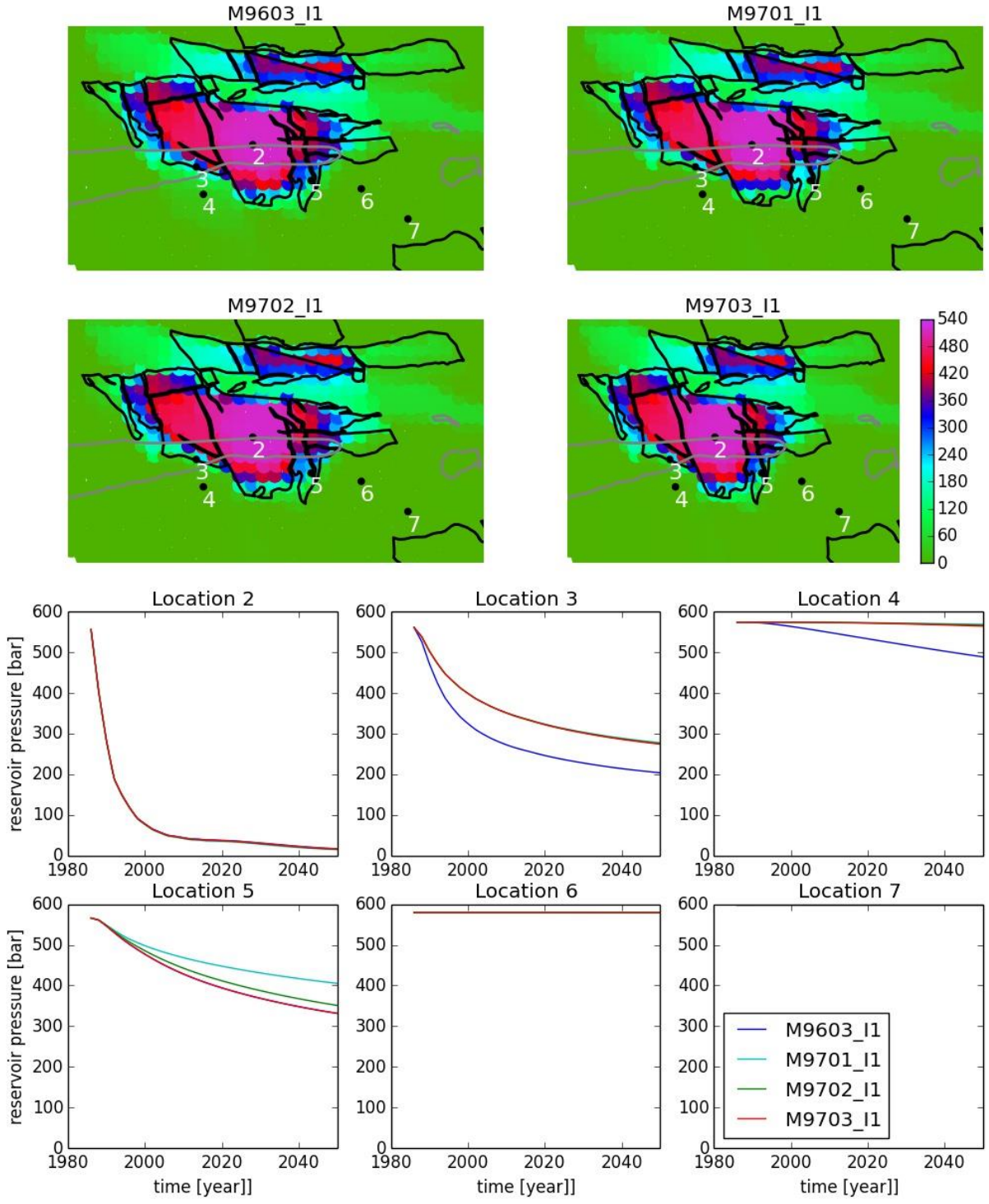

M10301_I1
M10302_I1
M10401_I1
M10402_I1
1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
M10301_I1
M10302_I1
M10401_I1
M10402_I1

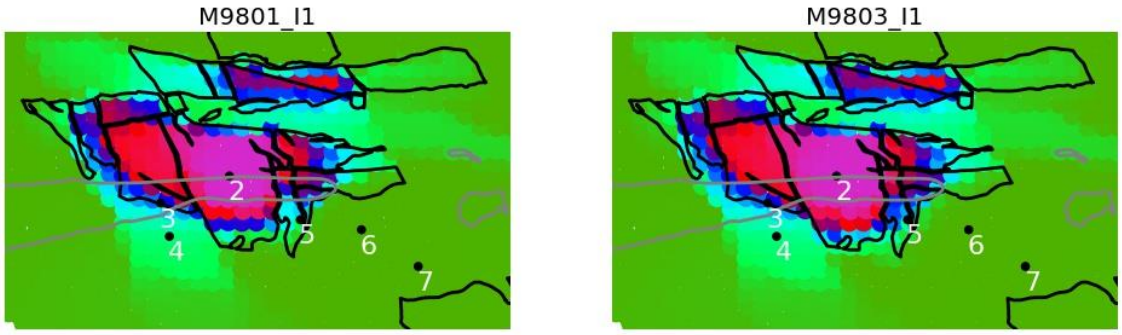

M9002_I1
M9003_I1
M9101_I1
M9102_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M9002_I1
M9003_I1
M9101_I1
M9102_I1

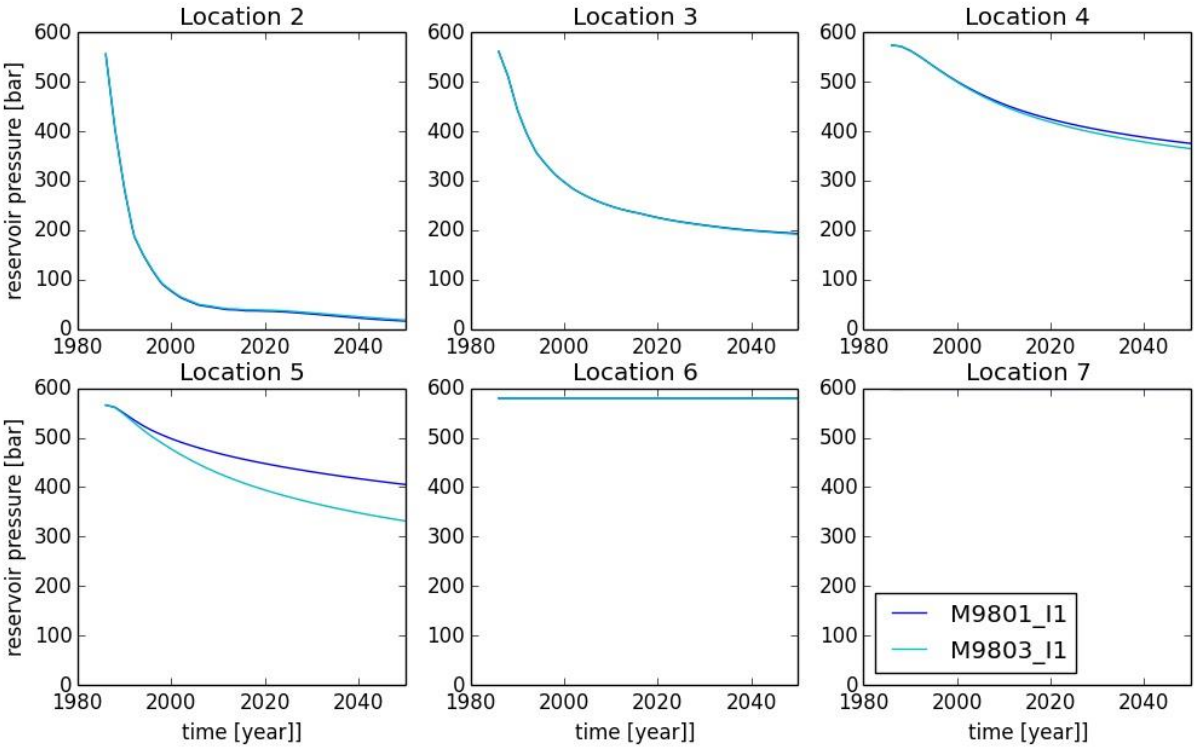

M9103_I1
M9201_I1
M9202_I1
M9203_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
0
100
200
300
400
500
600
1980
2000
2020
2040
M9103_I1
M9201_I1
M9202_I1
M9203_I1


M9301_I1
M9302_I1
M9303_I1
M9501_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M9301_I1
M9302_I1
M9303_I1
M9501_I1


M9502_I1
M9503_I1
M9601_I1
M9602_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M9502_I1
M9503_I1
M9601_I1
M9602_I1


M9603_I1
M9701_I1
M9702_I1
M9703_I1
2
3
4
5
6
7
540
480
420
360
300
240
180
120
60
0
Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
1980
2000
2020
2040
0
100
200
300
400
500
600
M9603_I1
M9701_I1
M9702_I1
M9703_I1


M9801_I1
M9803_I1
2
3
4
5
6
7


Location 2
Location 3
Location 4
Location 5
Location 6
Location 7
reservoir pressure [bar]
time [year]]
0
100
200
300
400
500
600
1980
2000
2020
2040
M9801_I1
M9803_I1

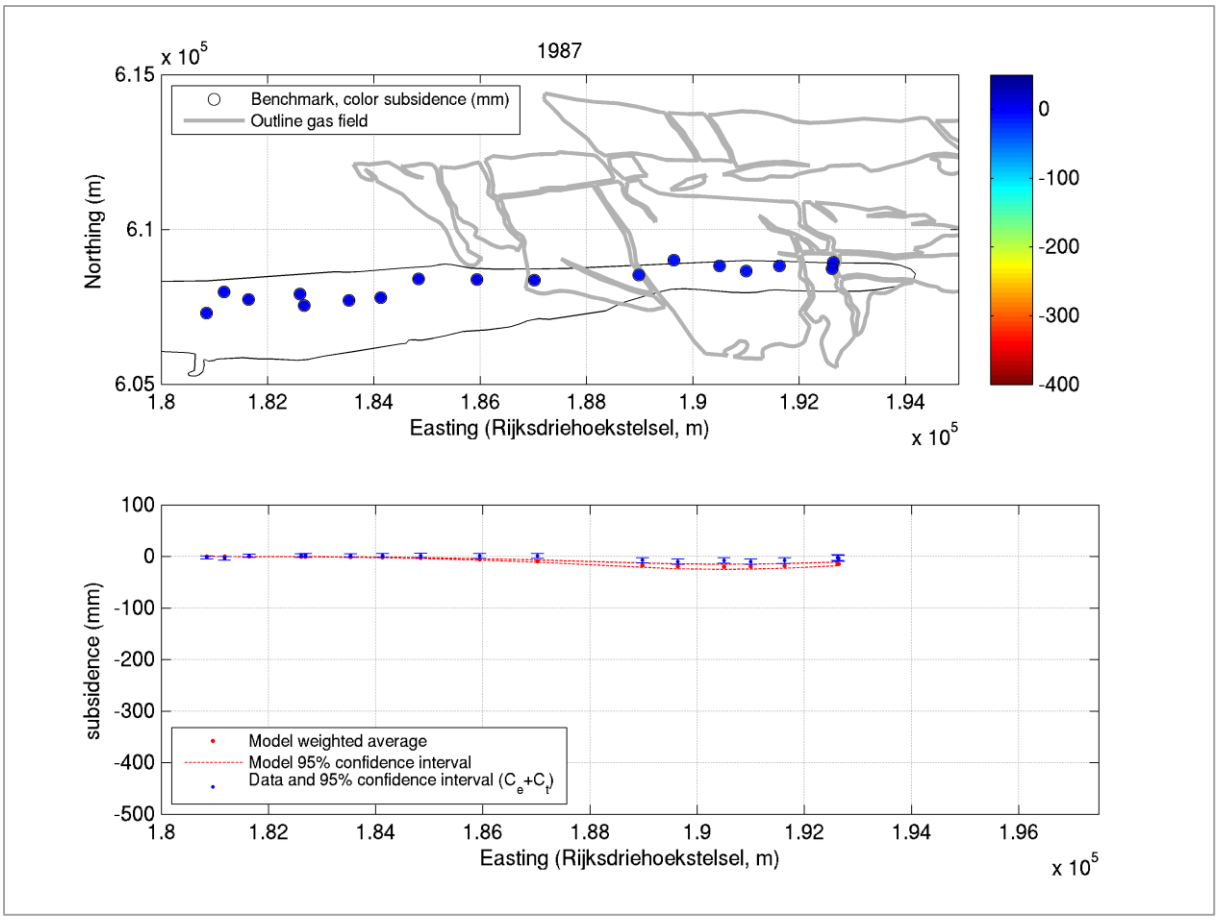
Appendix 2. Subsidence time series and profile visualisations

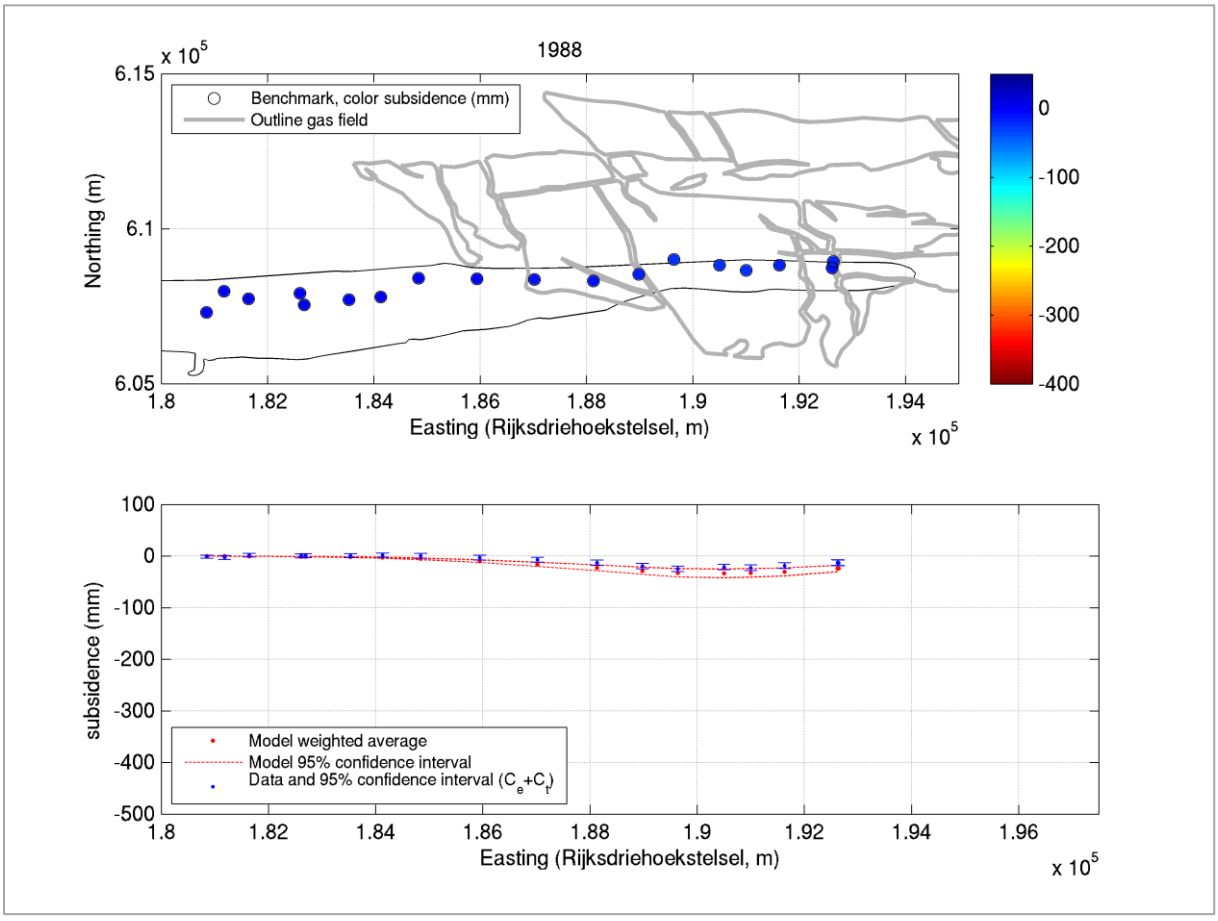
This appendix contains time series and profile visualisations that provide detailed insight in the match between geodetic data and subsidence models, including the confidence intervals:

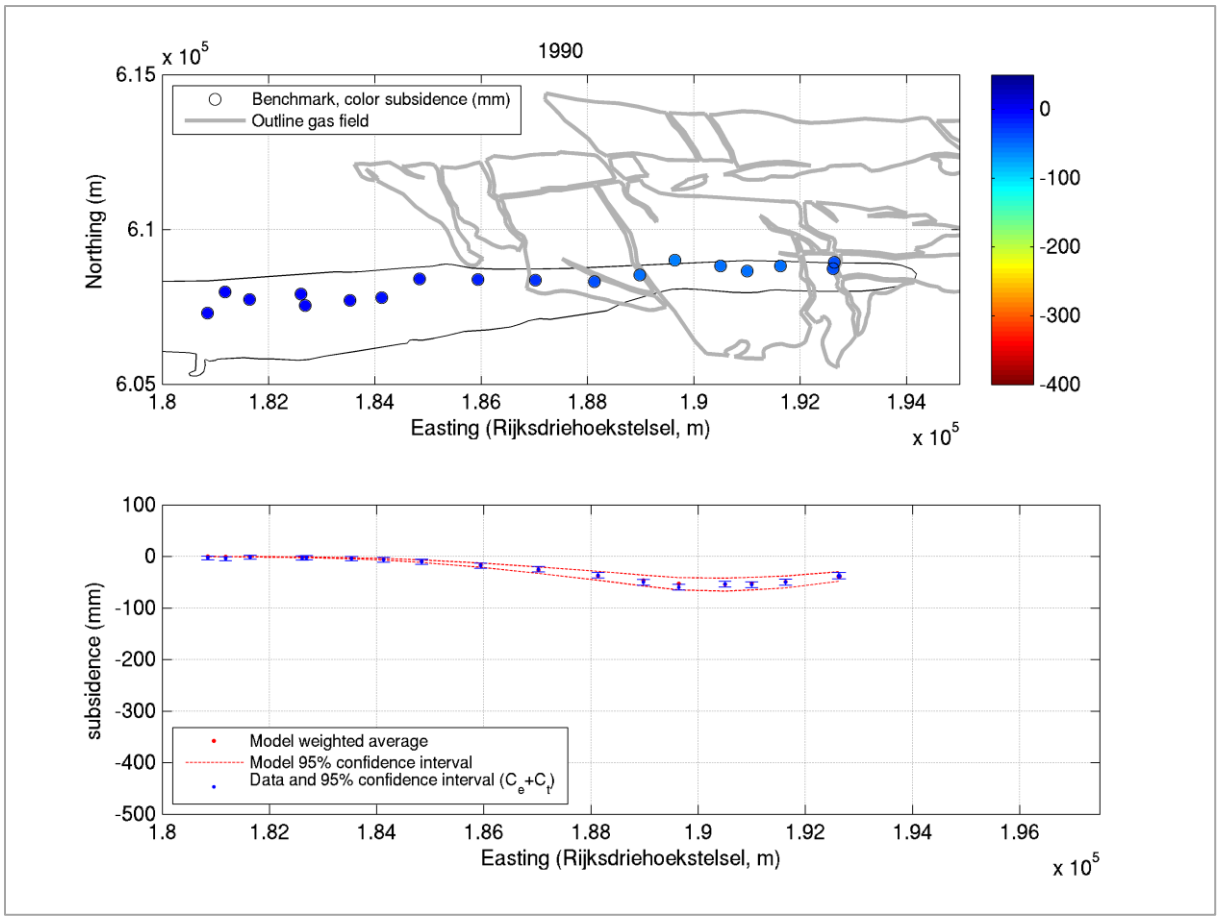
1. Levelling profile plots along a northern and a southern trajectory on the island.
    - For the visualisations, the geodetic data and subsidence model double-differences have been transformed to a common reference in time (1986) and space (benchmark 000A2592 in Nes). See Section 7.4 for more information and an explanation of the transformation of double-differences and the corresponding covariance matrix.
2. The full time series for all levelling benchmarks that have been measured in both 1986 and 2017.
    - As for the profile plots, double-differences have been transformed to a common reference in time (1986) and space (benchmark 000A2592 in Nes).
    - The time series visualisations for all 1417 double-differences used in the confrontation workflow have been complied in a separate document (NAM, 2017c), which is available on request.
3. Time series for the GPS benchmarks with respect to the earliest measurement campaign of the benchmark.
    - In the interpretation of the time series of the Waddenzee benchmarks, it has to be noted that the weighted average and the confidence interval of the subsidence models contain subsidence due the **Ameland gas fields only**. Hence, the Waddenzee benchmarks that are close to the Nes and Moddergat gas fields can be affected by an additional subsidence component that is visible in the geodetic data time series, but not in the subsidence model time series.

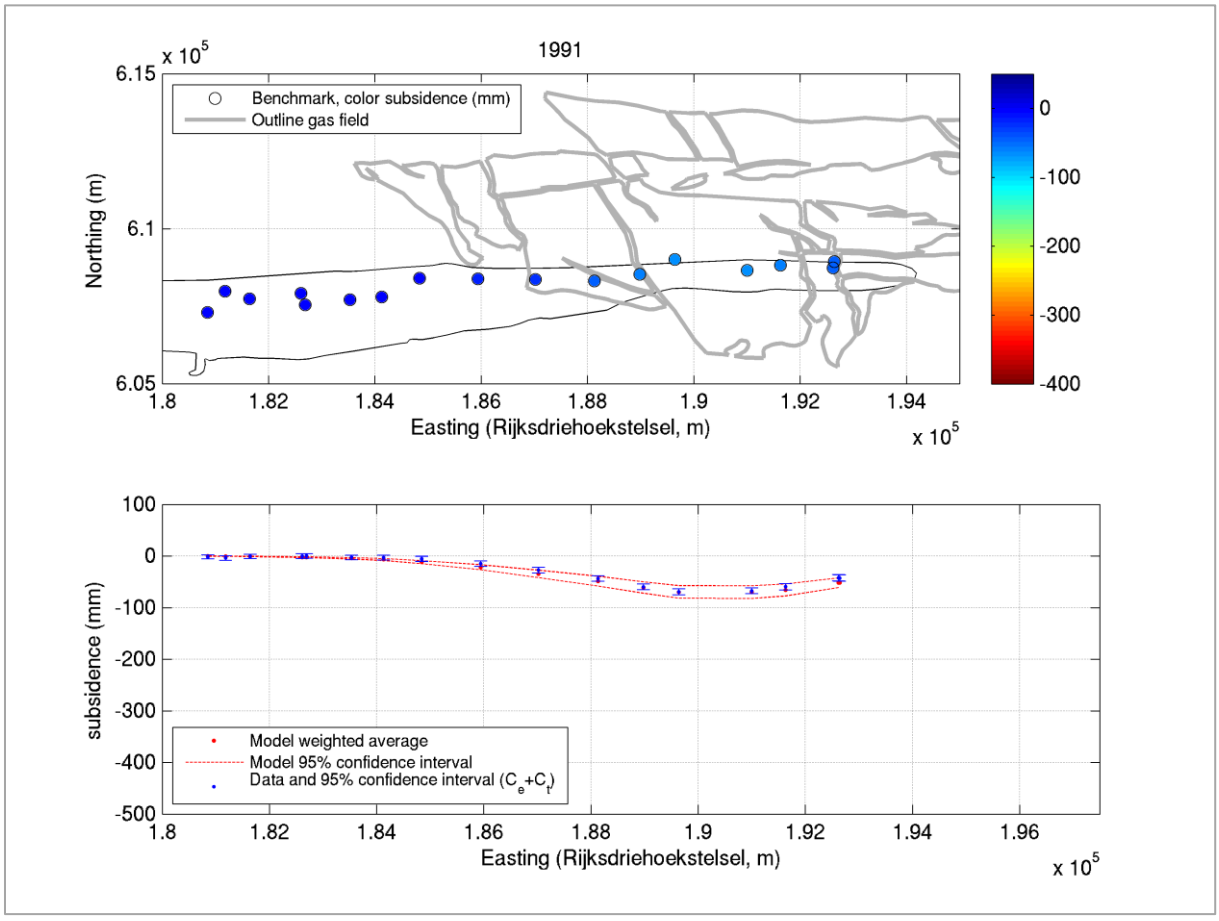
All time series and profile visualisations contain the following information (see for a further explanation Section 7.4) :

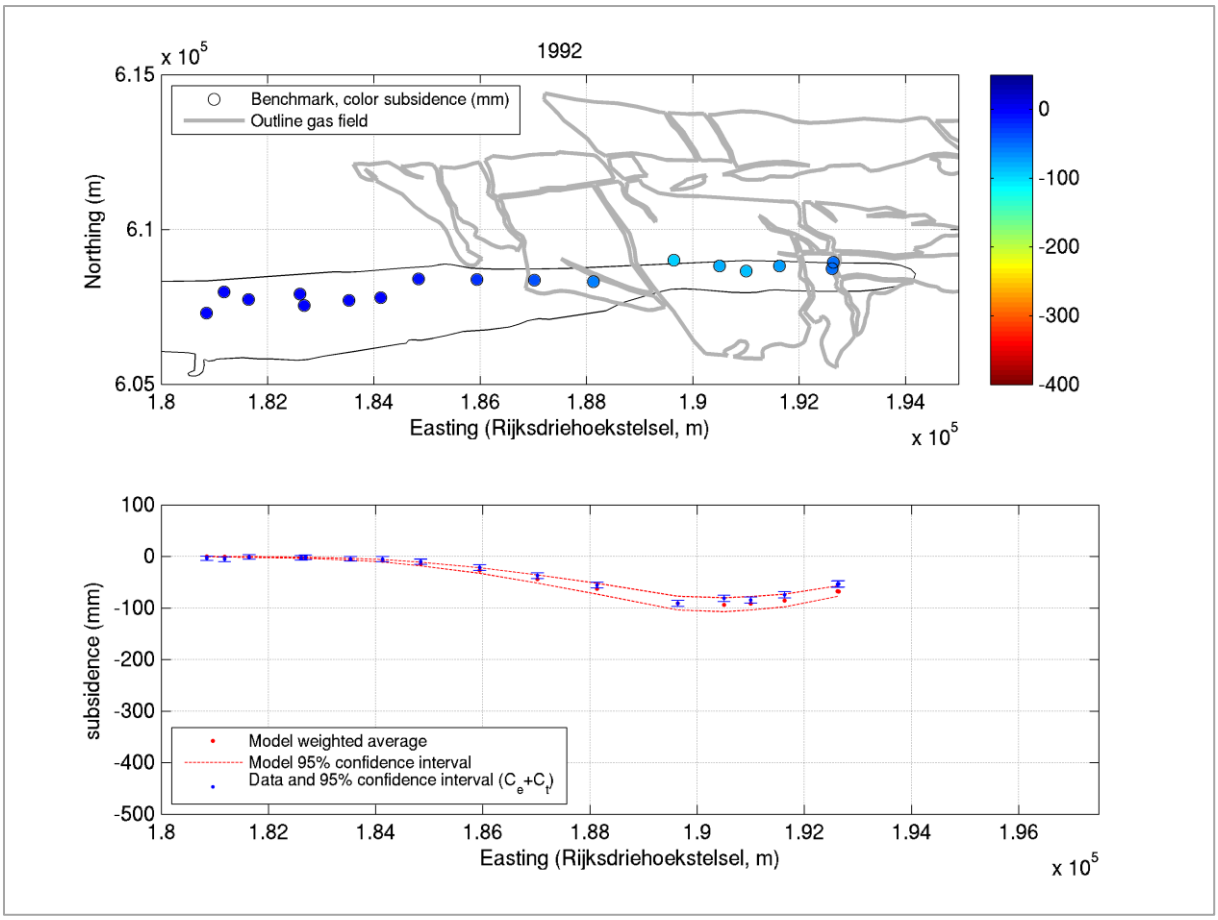
- **Blue**: geodetic data double-differences with 95% confidence interval derived from the covariance matrix considering measurement noise and the temporal component of idealisation noise ('benchmark instability'): $C_{e_{dd}} + C_{t_{dd}}$.
- **Red dots**: model weighted average computed using the probabilities determined from the particle filter (see Section 6.3).
- **Dashed red lines**: model 95% confidence interval based on the probability profiles per double-difference in space and time. The confidence interval is evaluated for benchmark locations or observation epochs, respectively, and linearly interpolated in between.

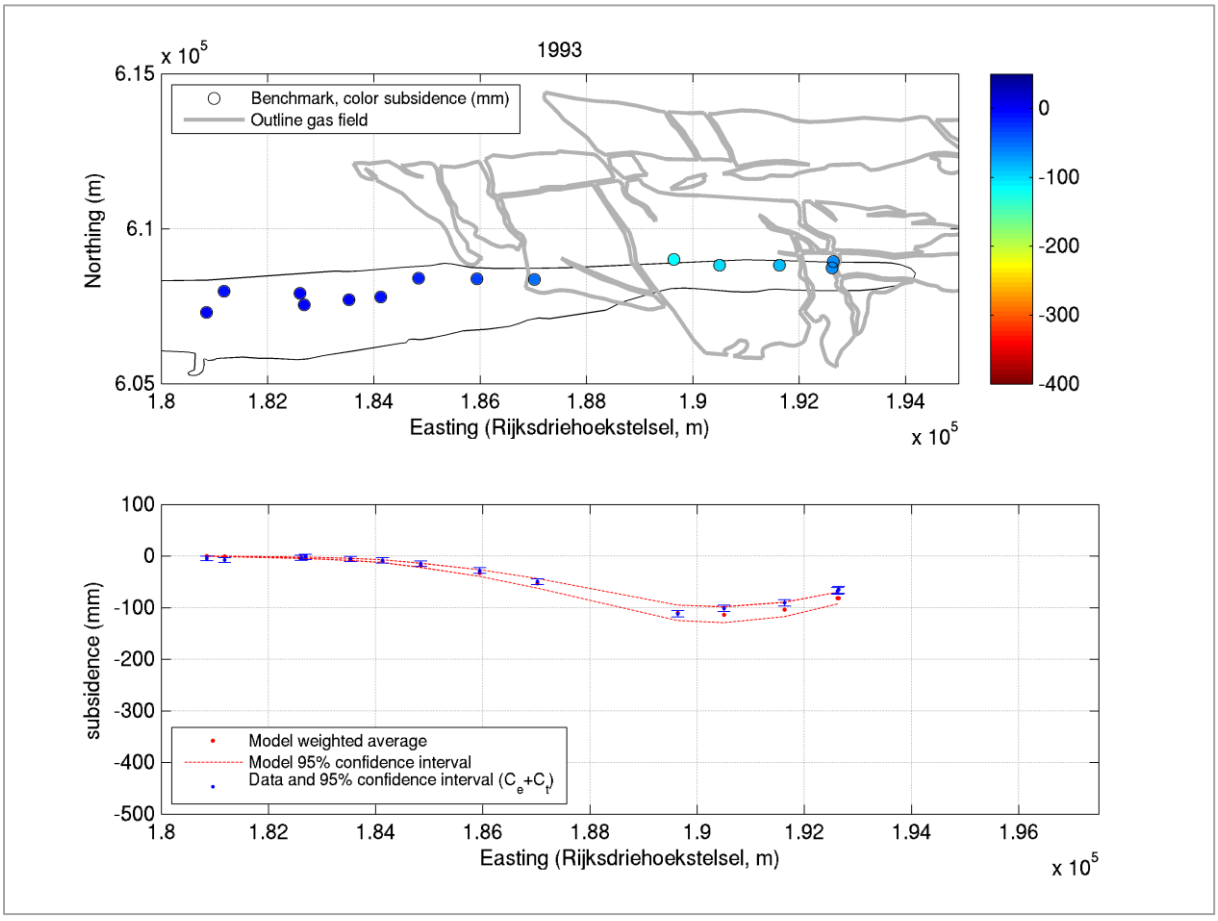

1987
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

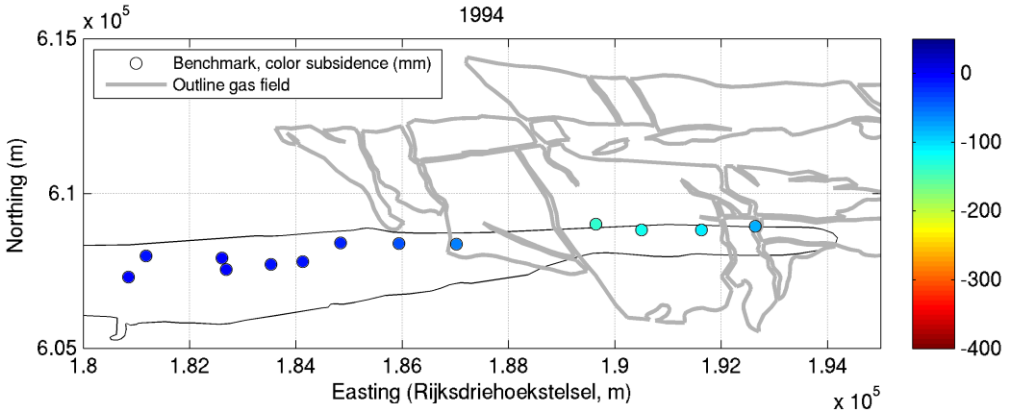

1988
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

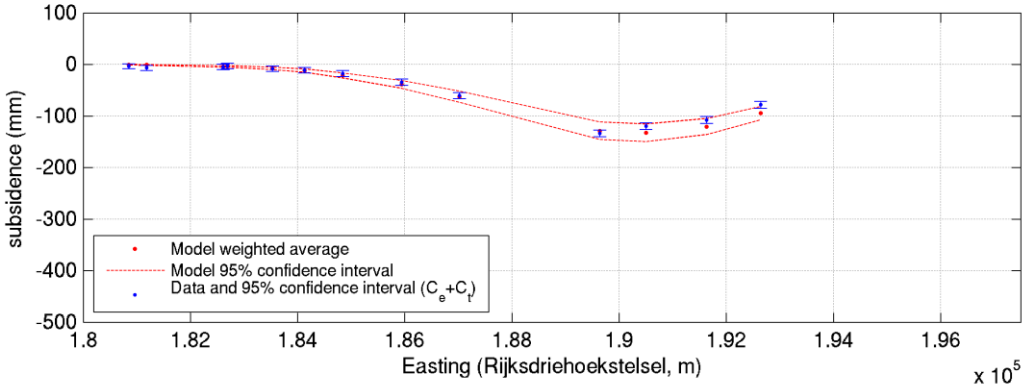

1990
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

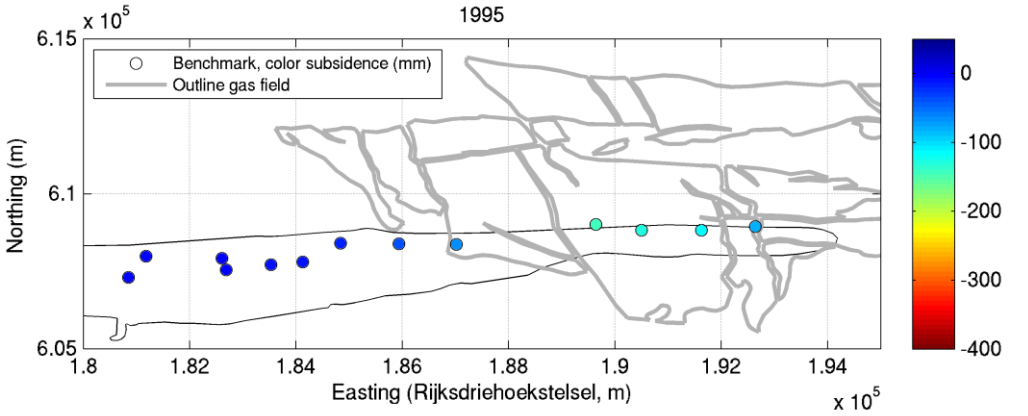

1991
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

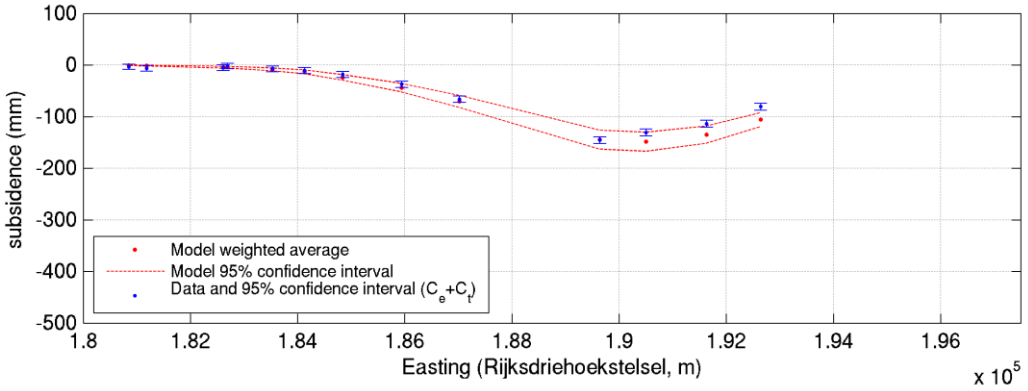

1992
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

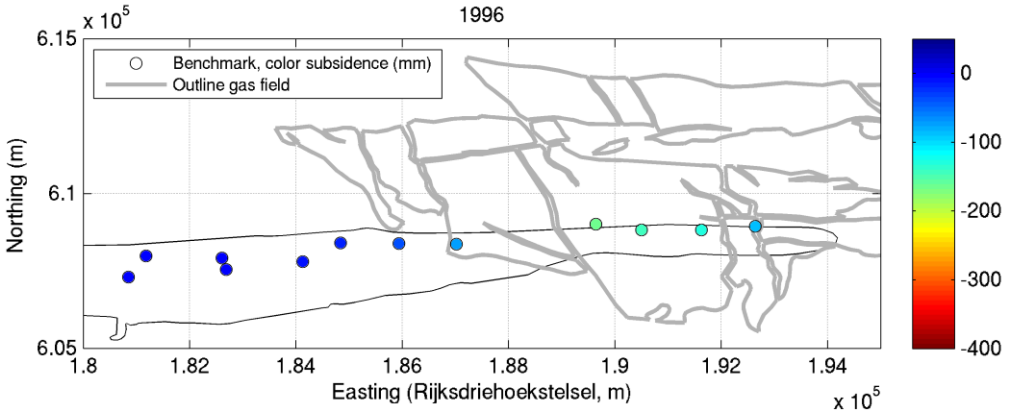

1993
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

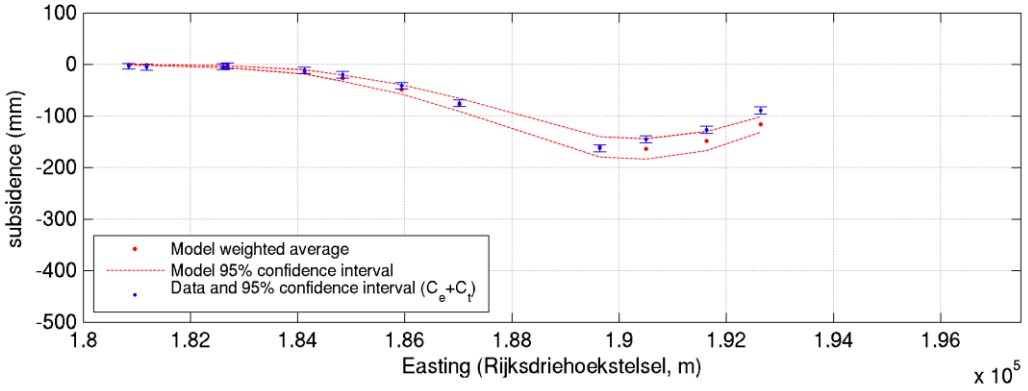

1994
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

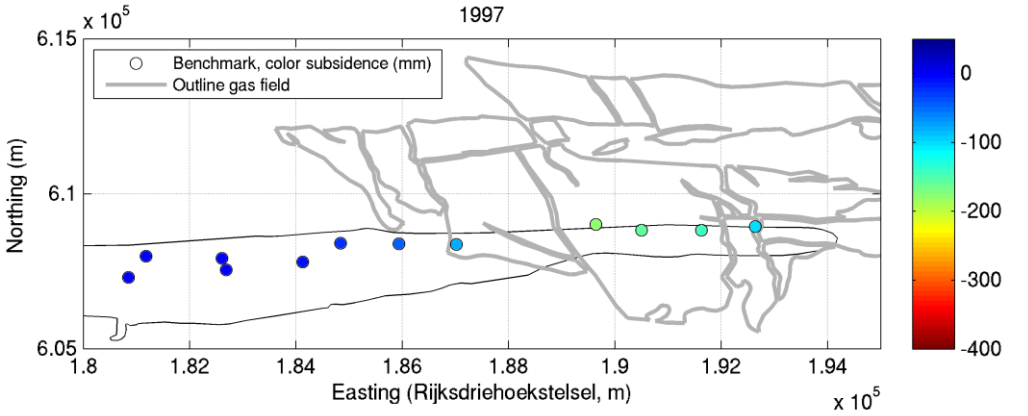

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

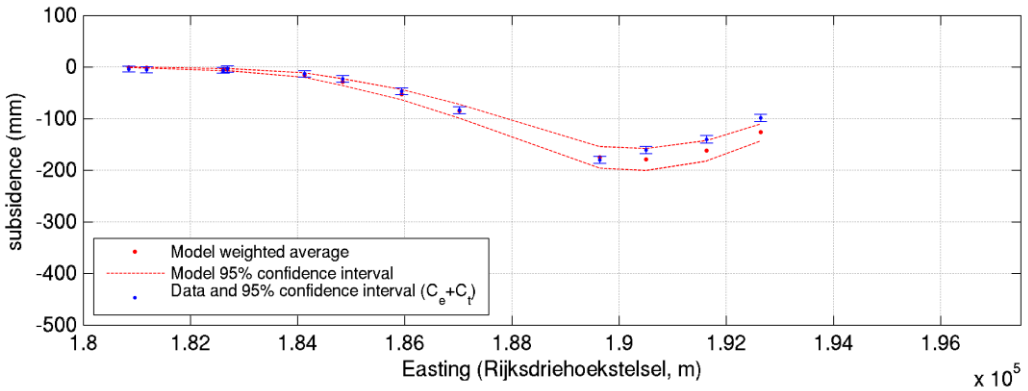

1995
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

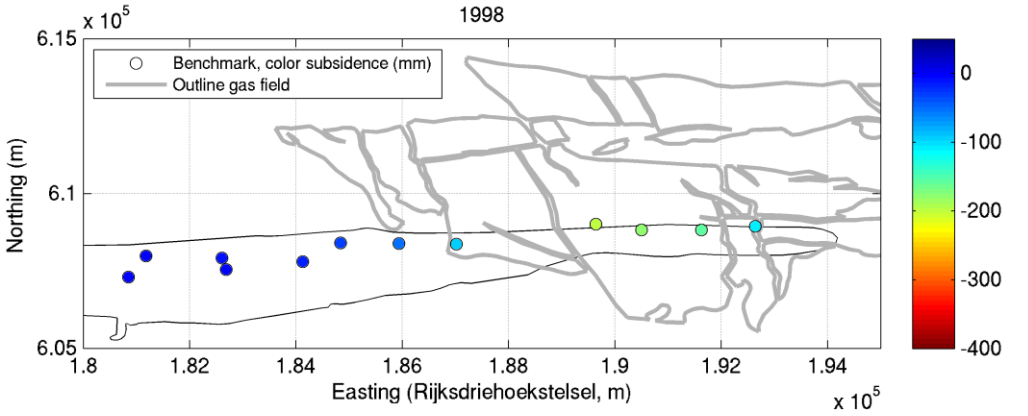

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

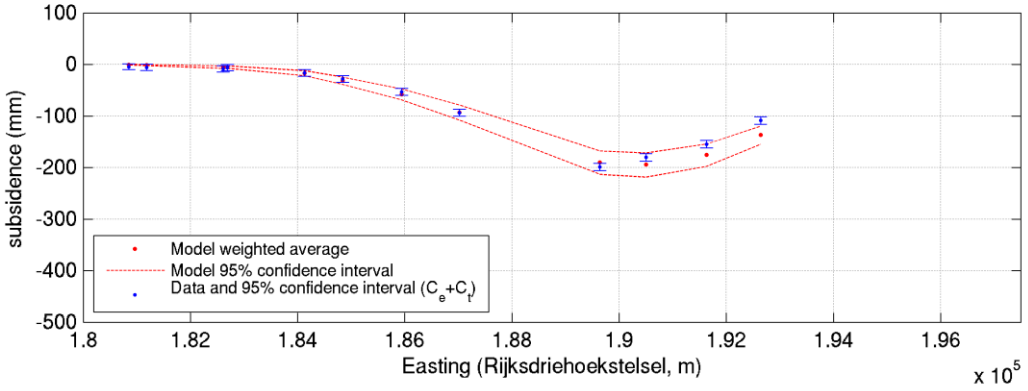

1996
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
6.15
6.1
6.05
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

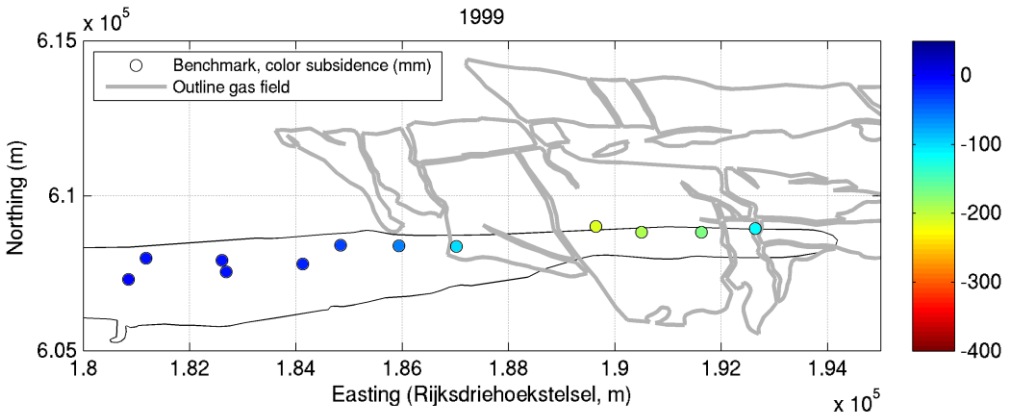

subsidence (mm)
100
0
-100
-200
-300
-400
-500
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

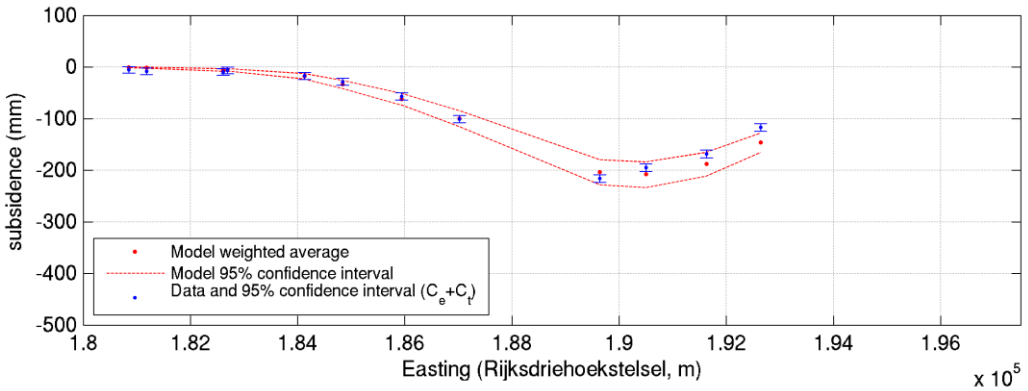

1997
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
6.15
6.1
6.05
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

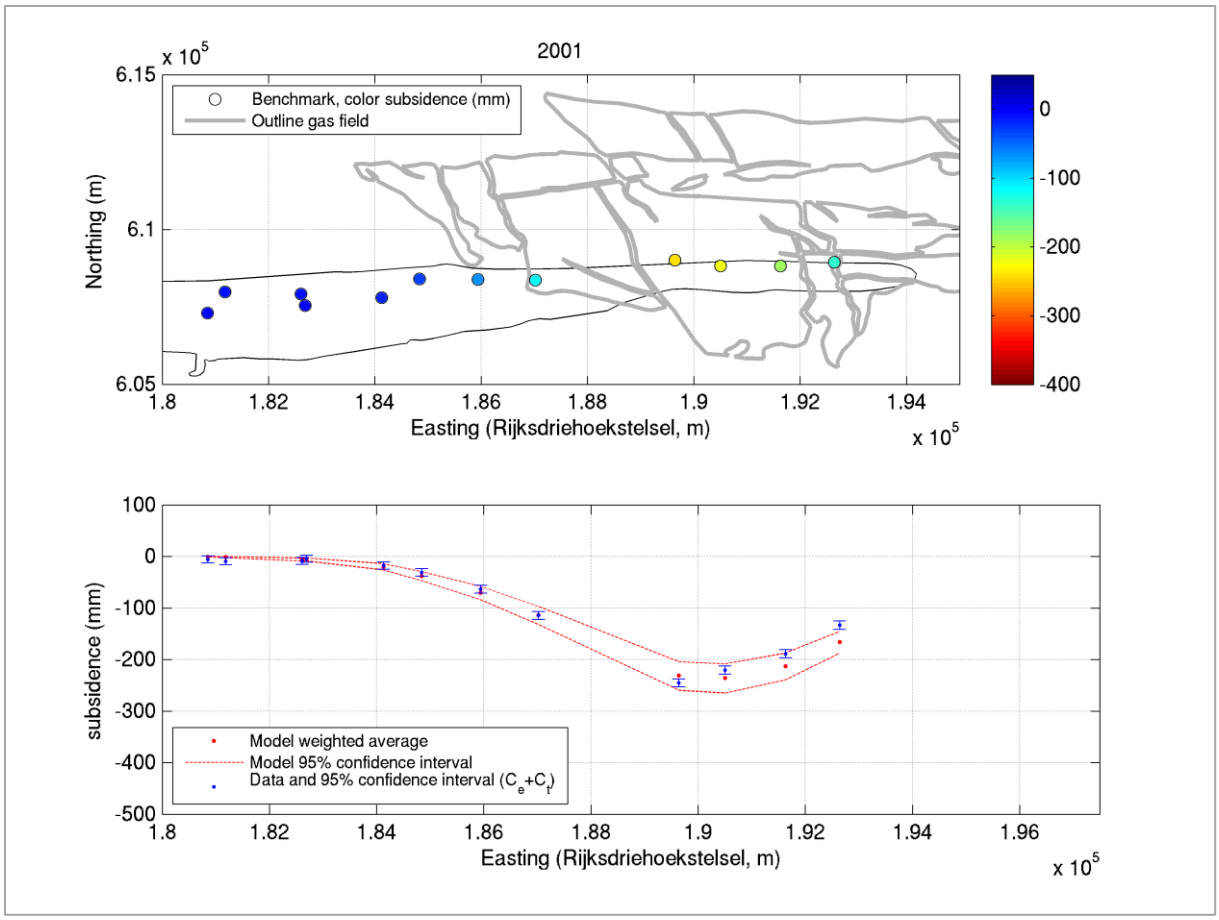

subsidence (mm)
100
0
-100
-200
-300
-400
-500
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

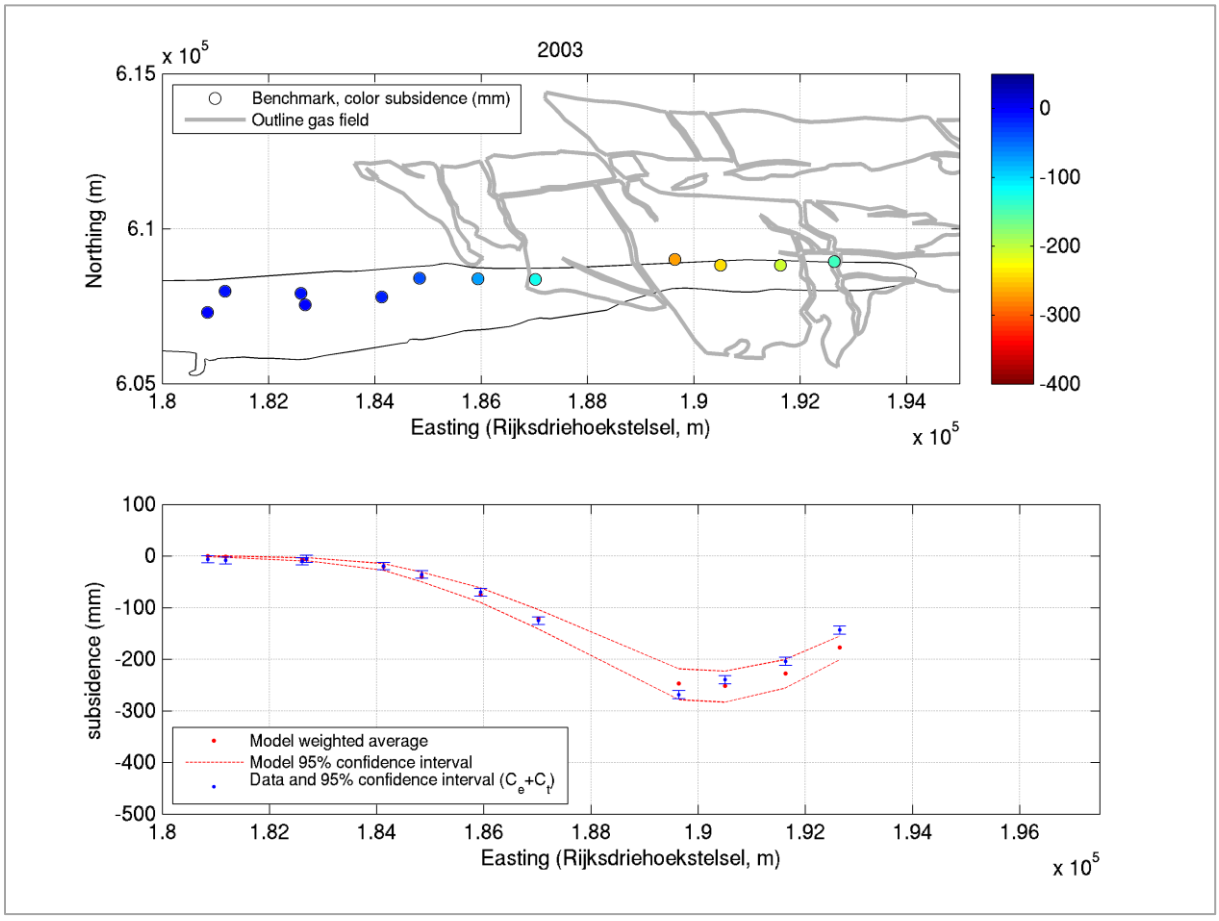

1998
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

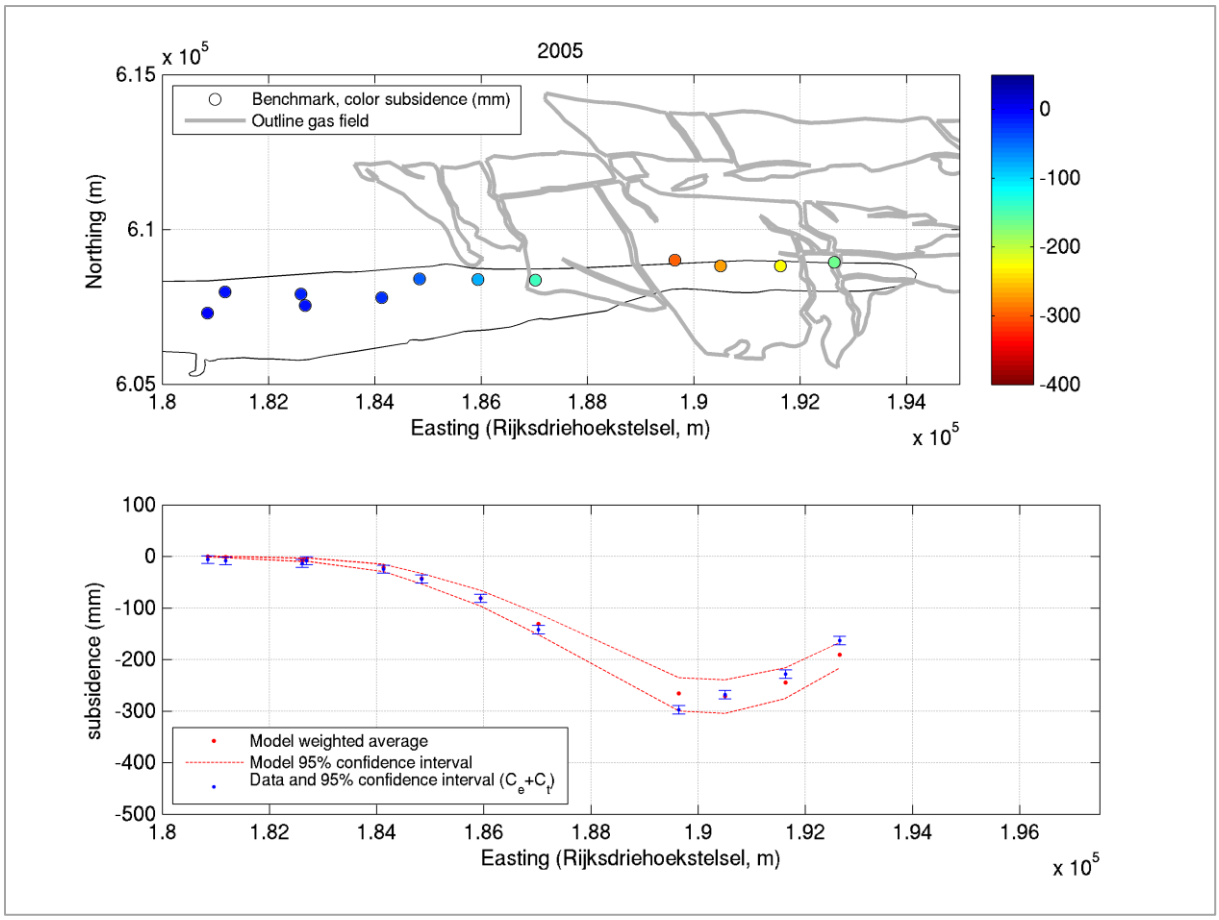

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

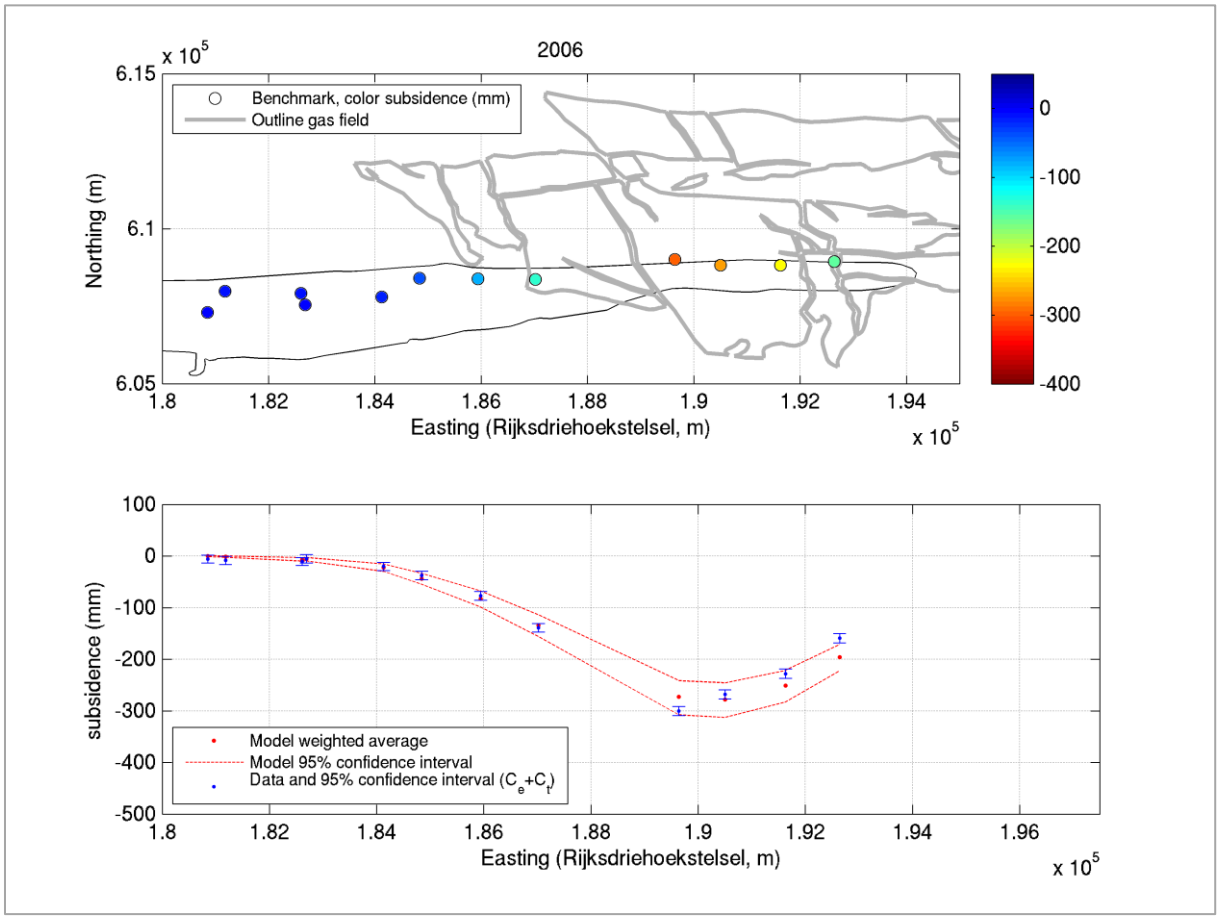

1999
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

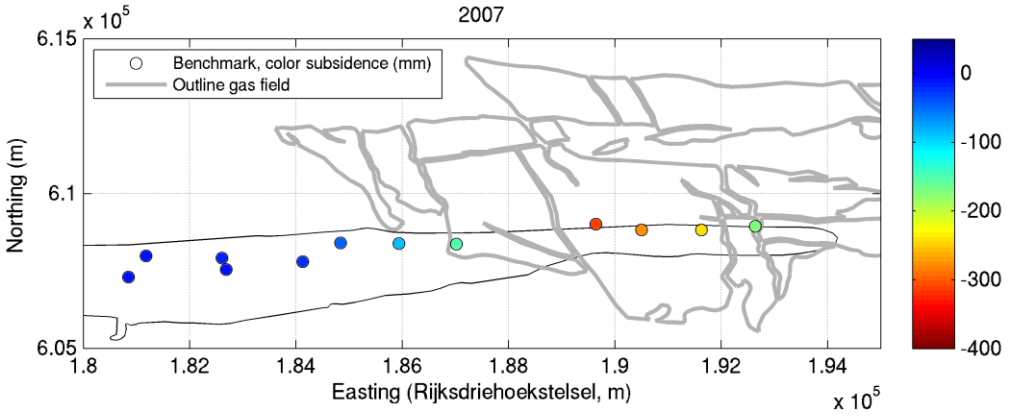

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

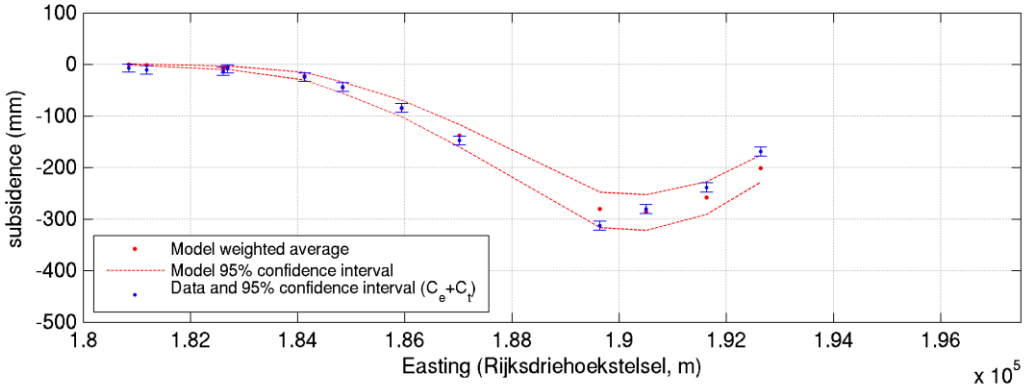

2001
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

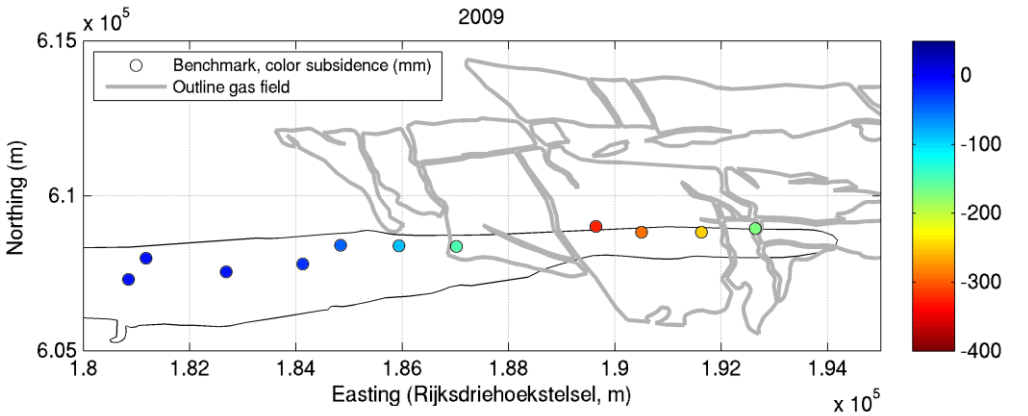

2003
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

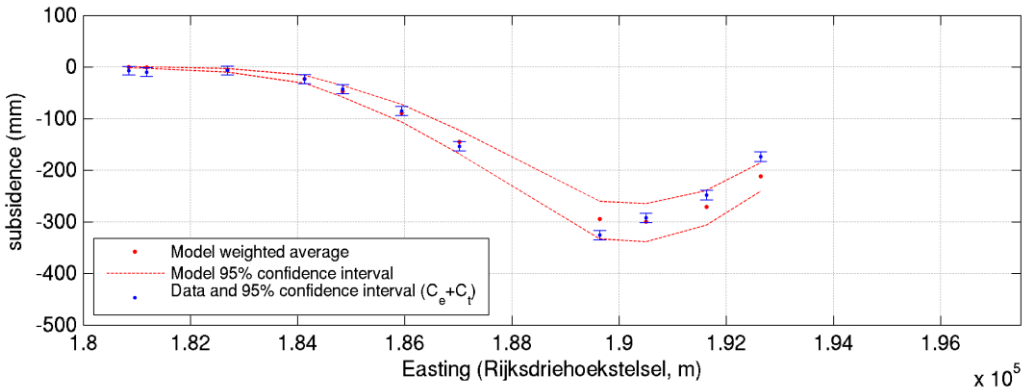

2005
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

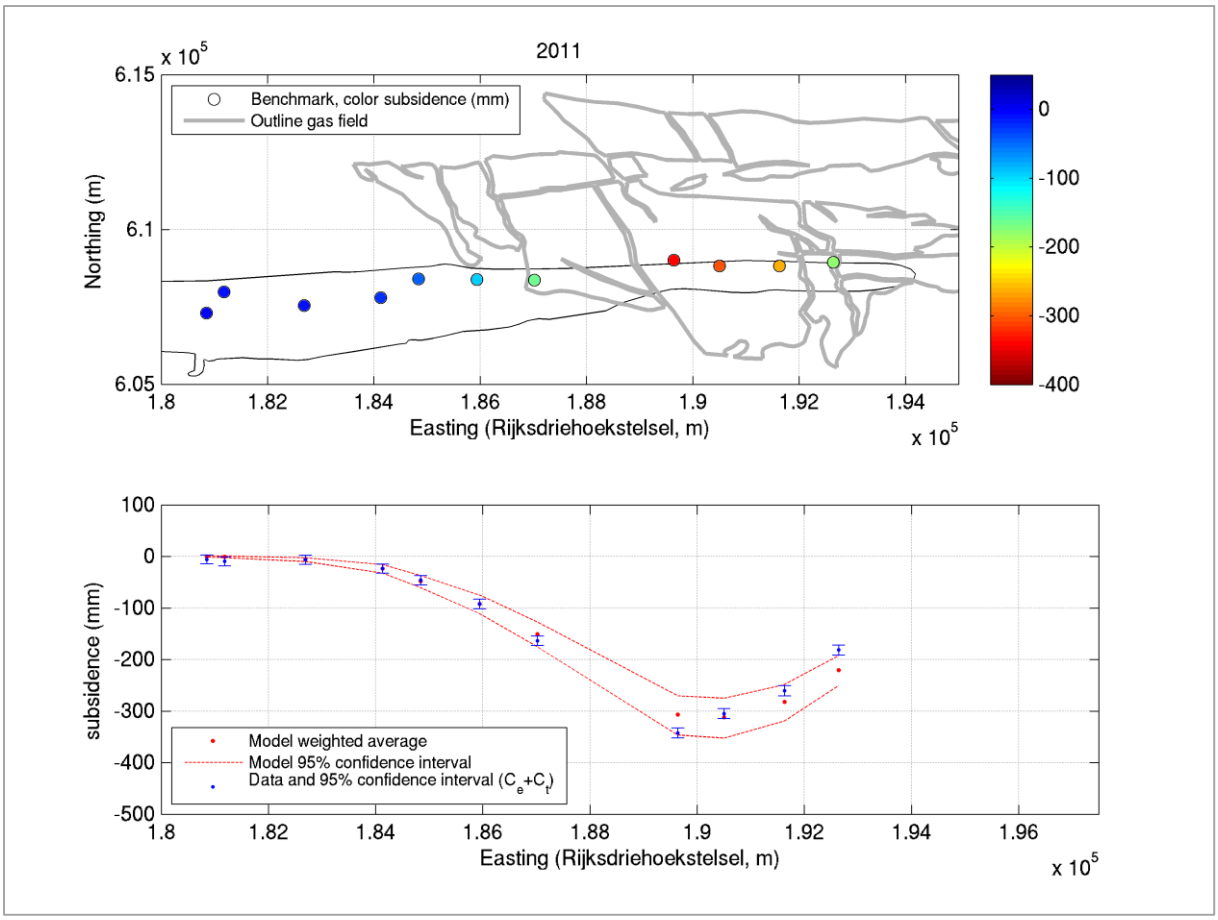

2006
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

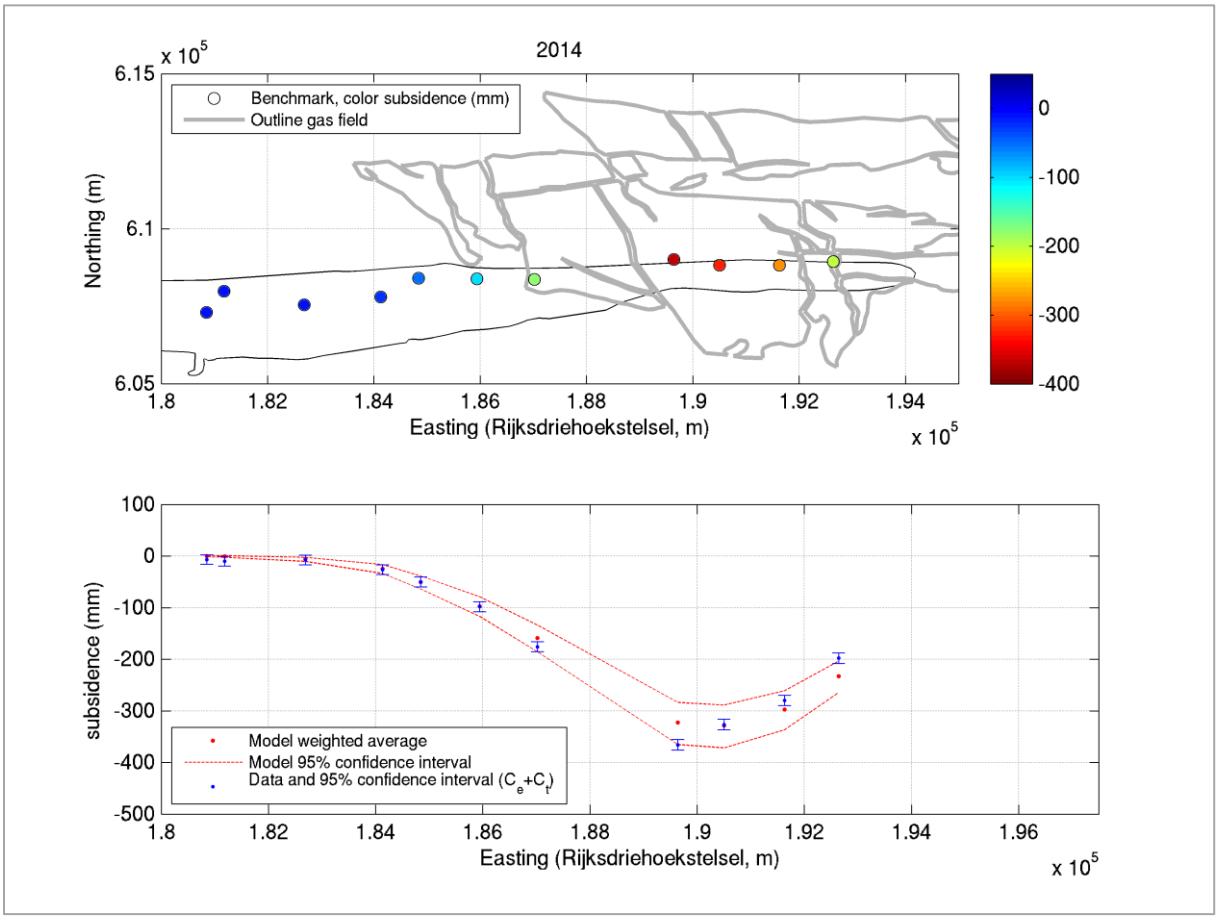

2007
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
6.15
6.1
6.05
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

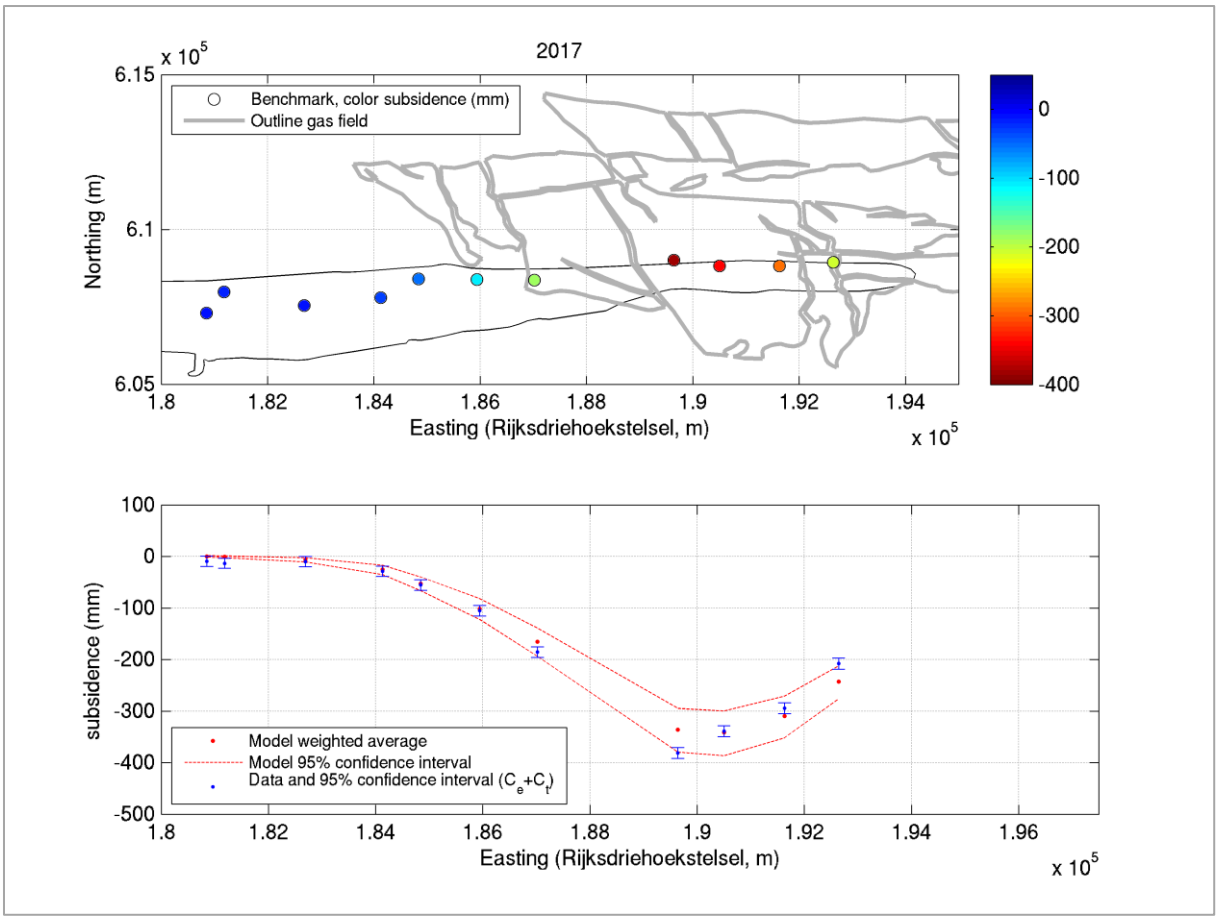

subsidence (mm)
100
0
-100
-200
-300
-400
-500
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

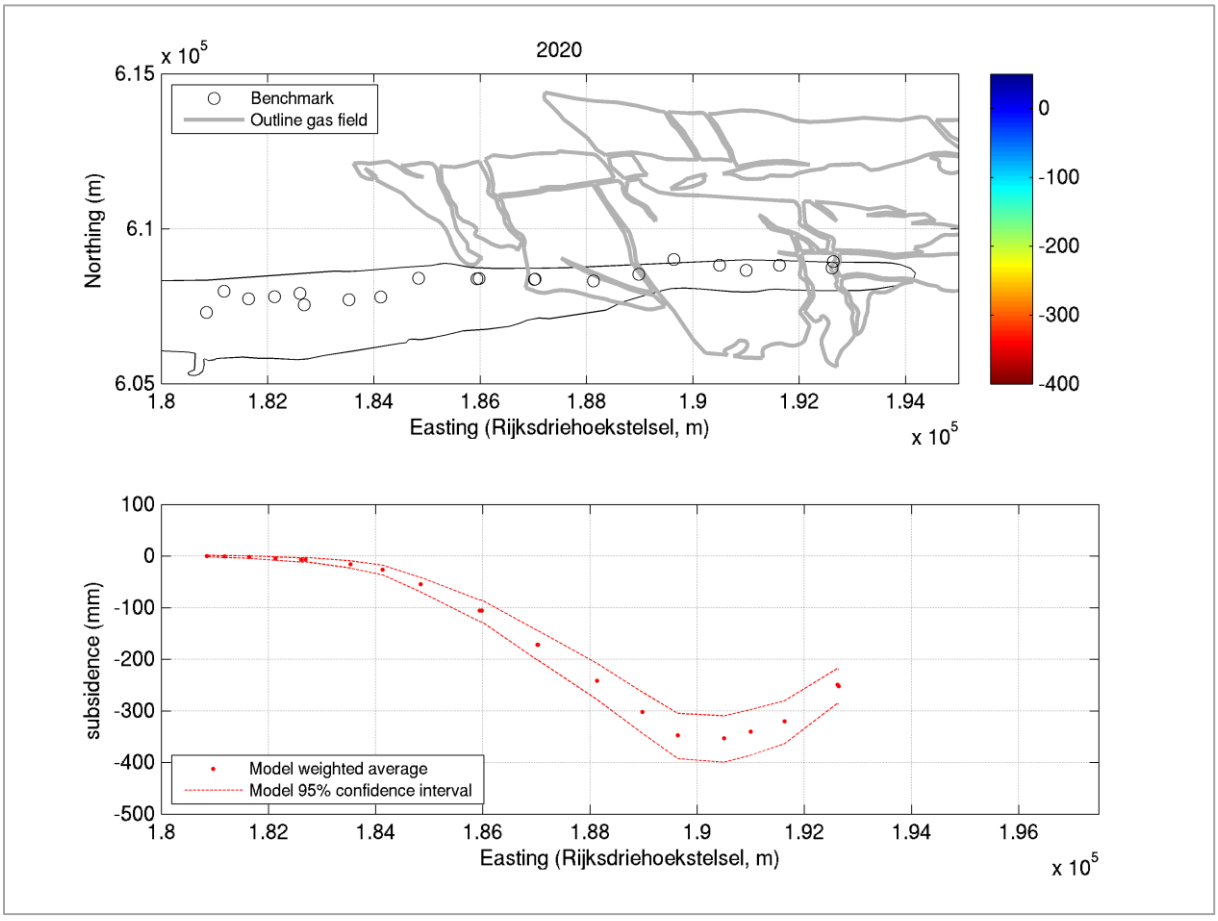

2009
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
6.15
6.1
6.05
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

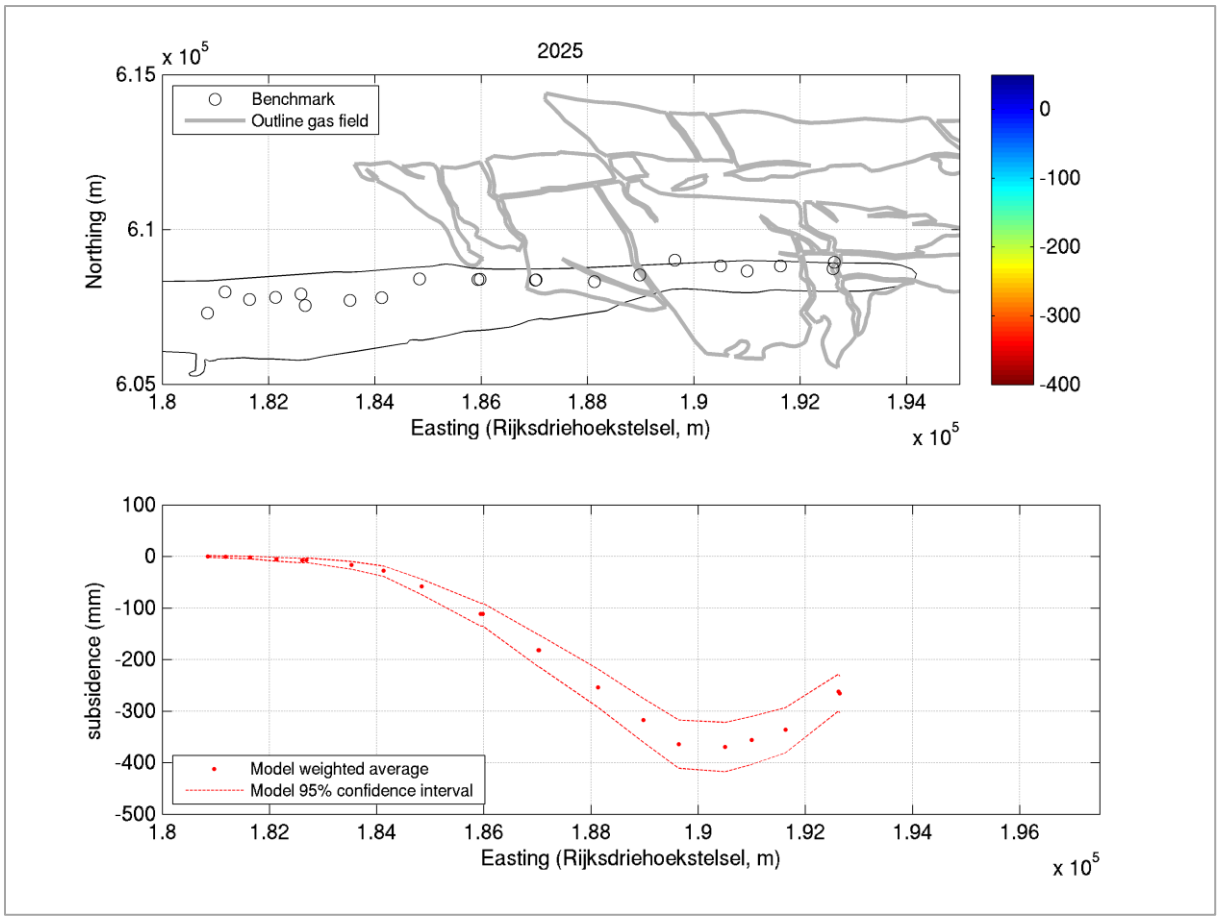

subsidence (mm)
100
0
-100
-200
-300
-400
-500
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

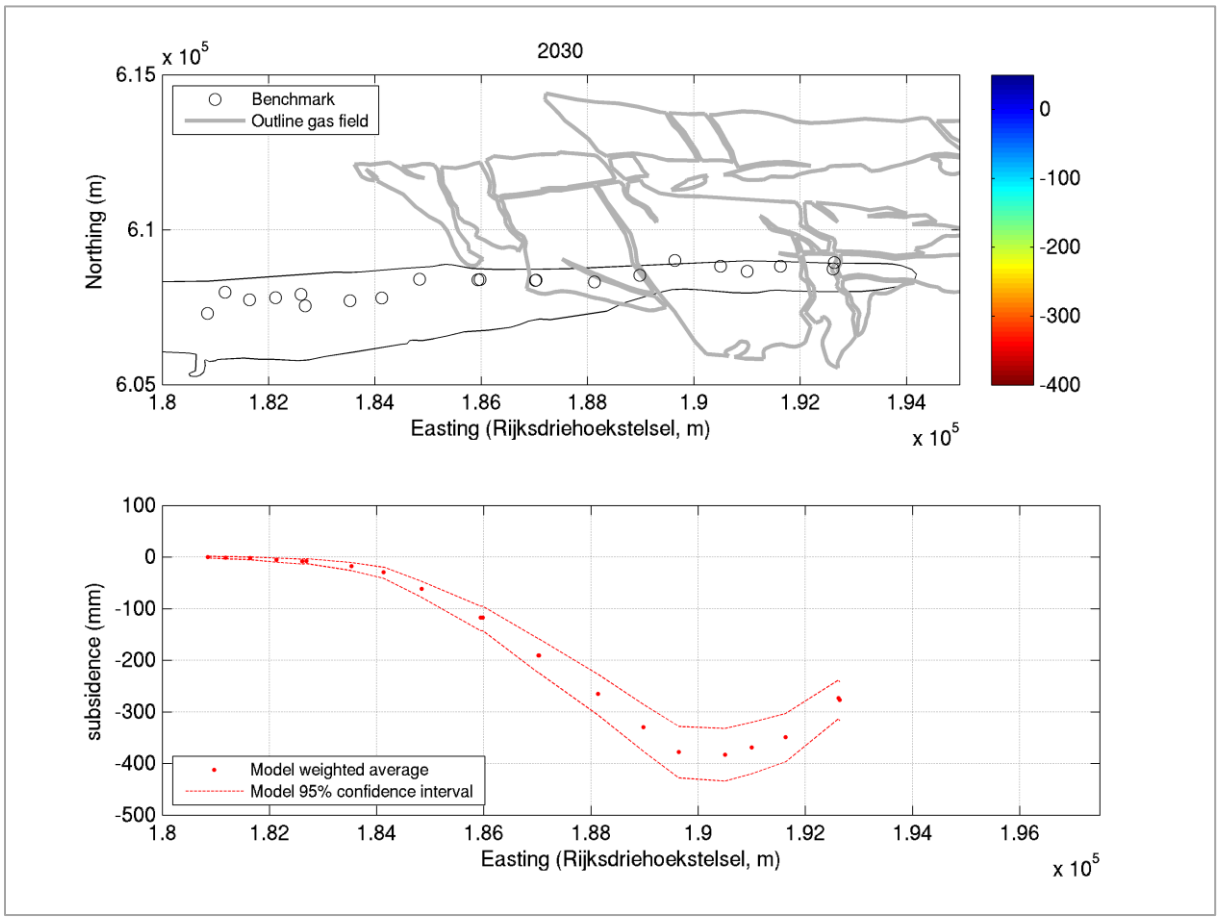

2011
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

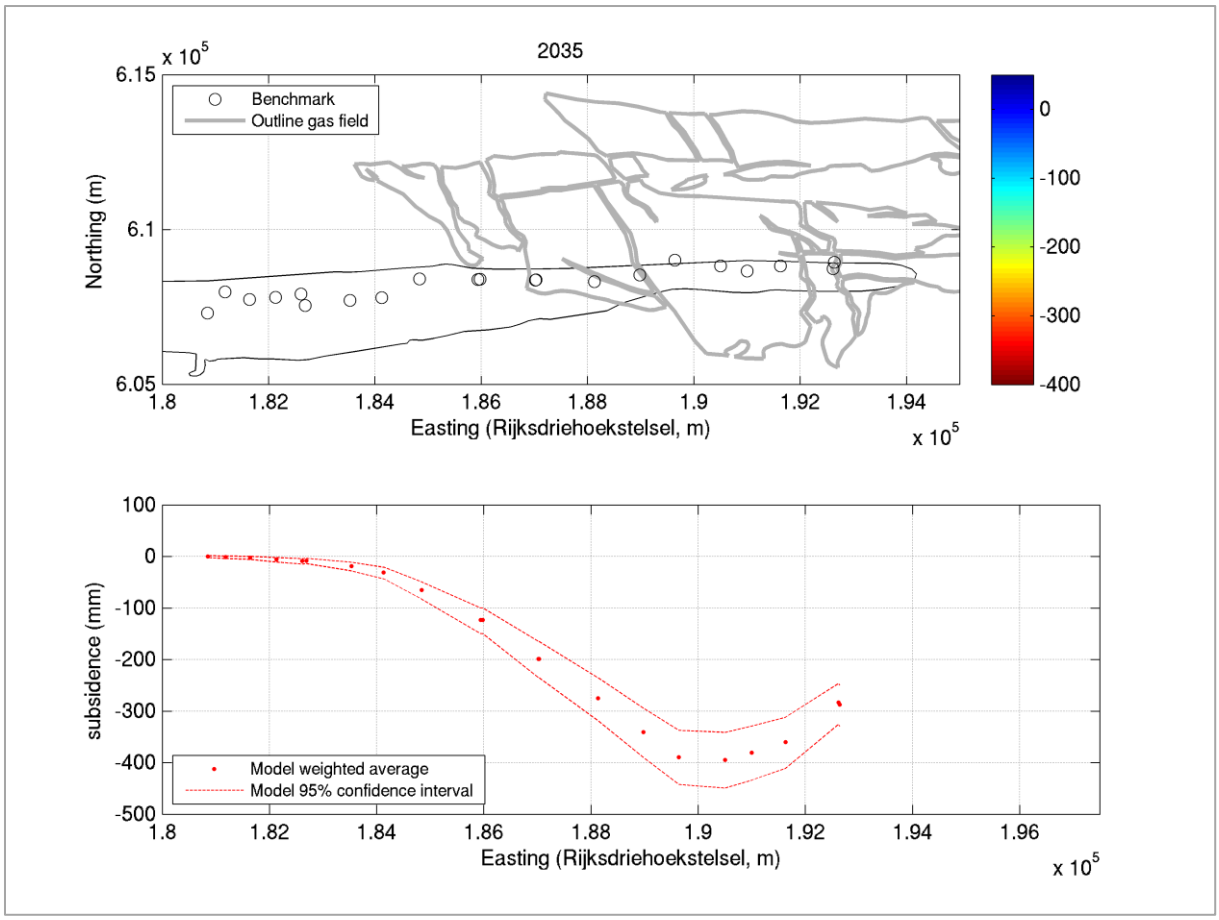

2014
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

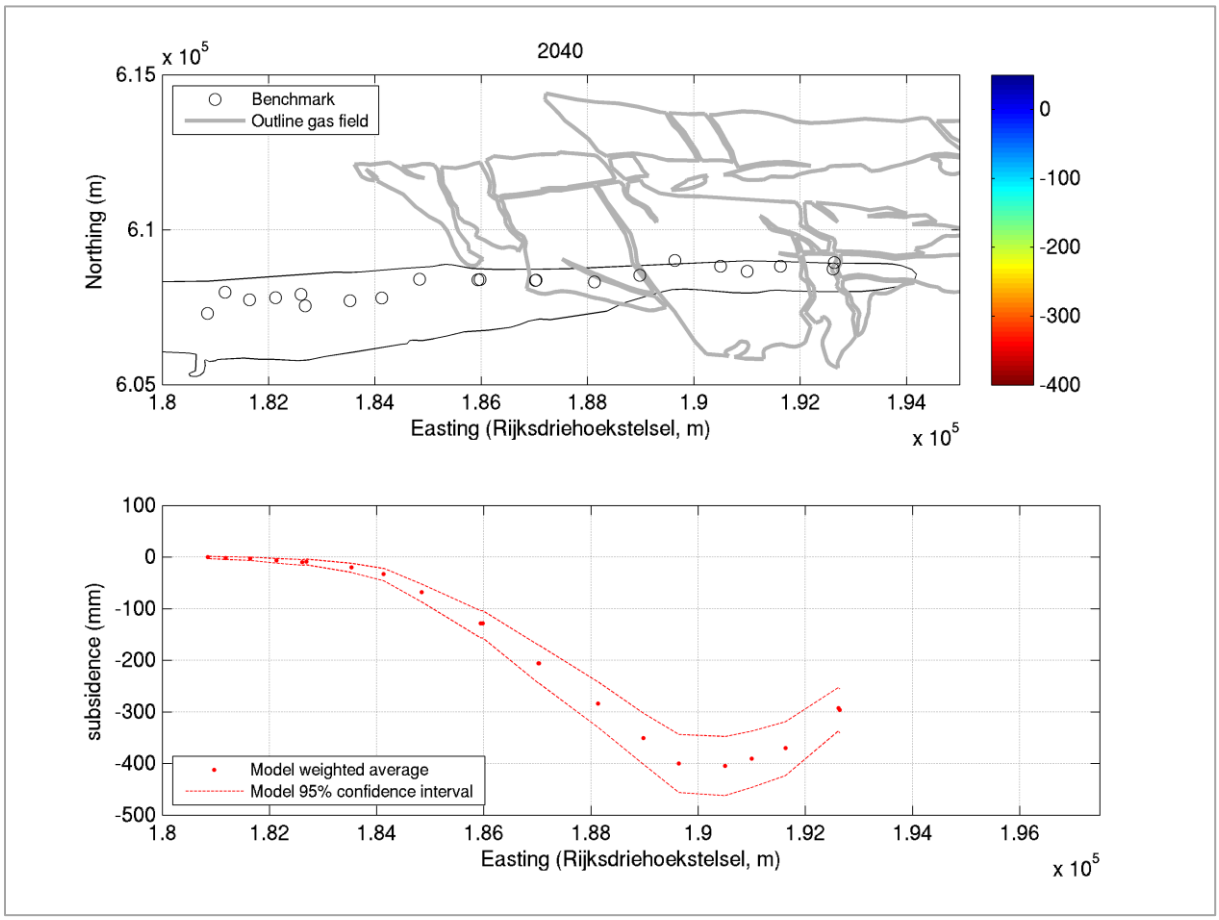

2017
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

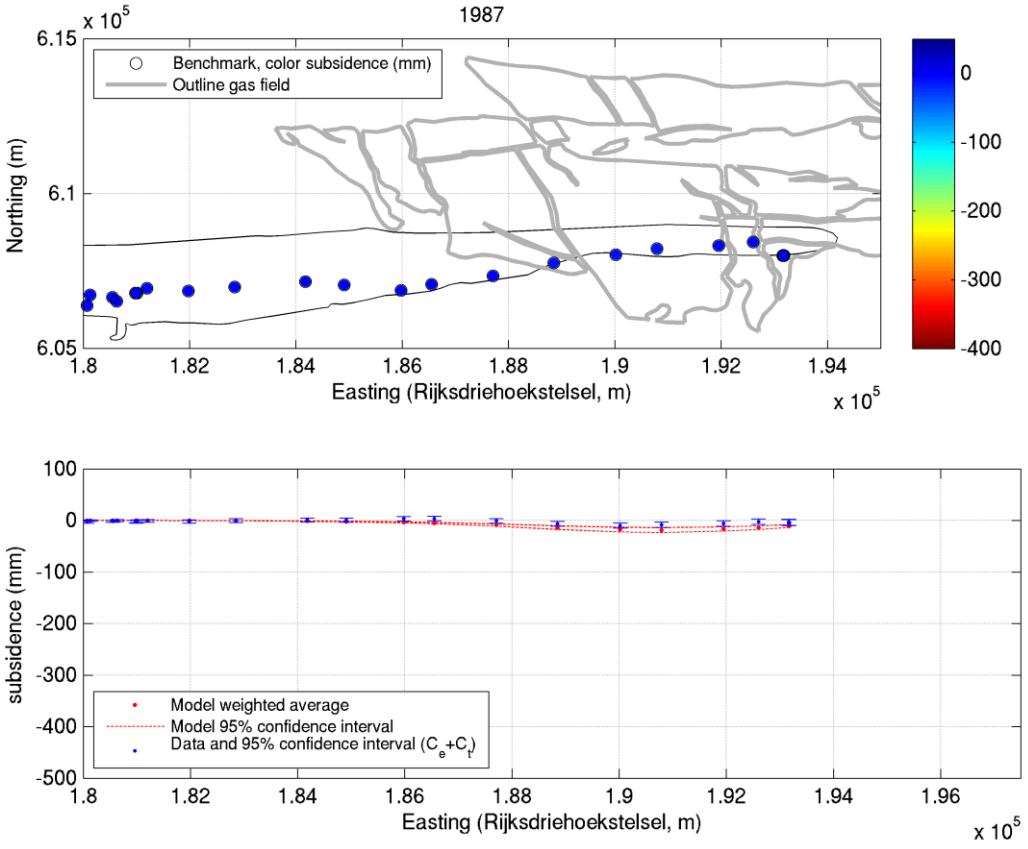

2020
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

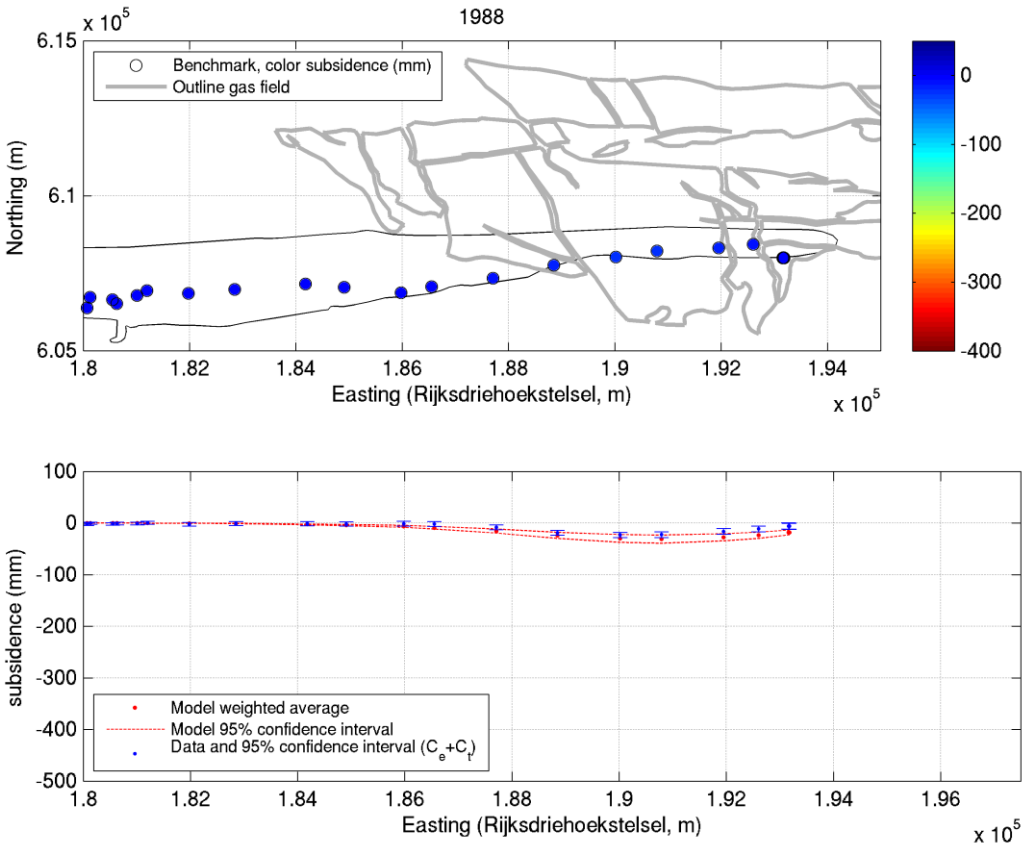

2025
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.96

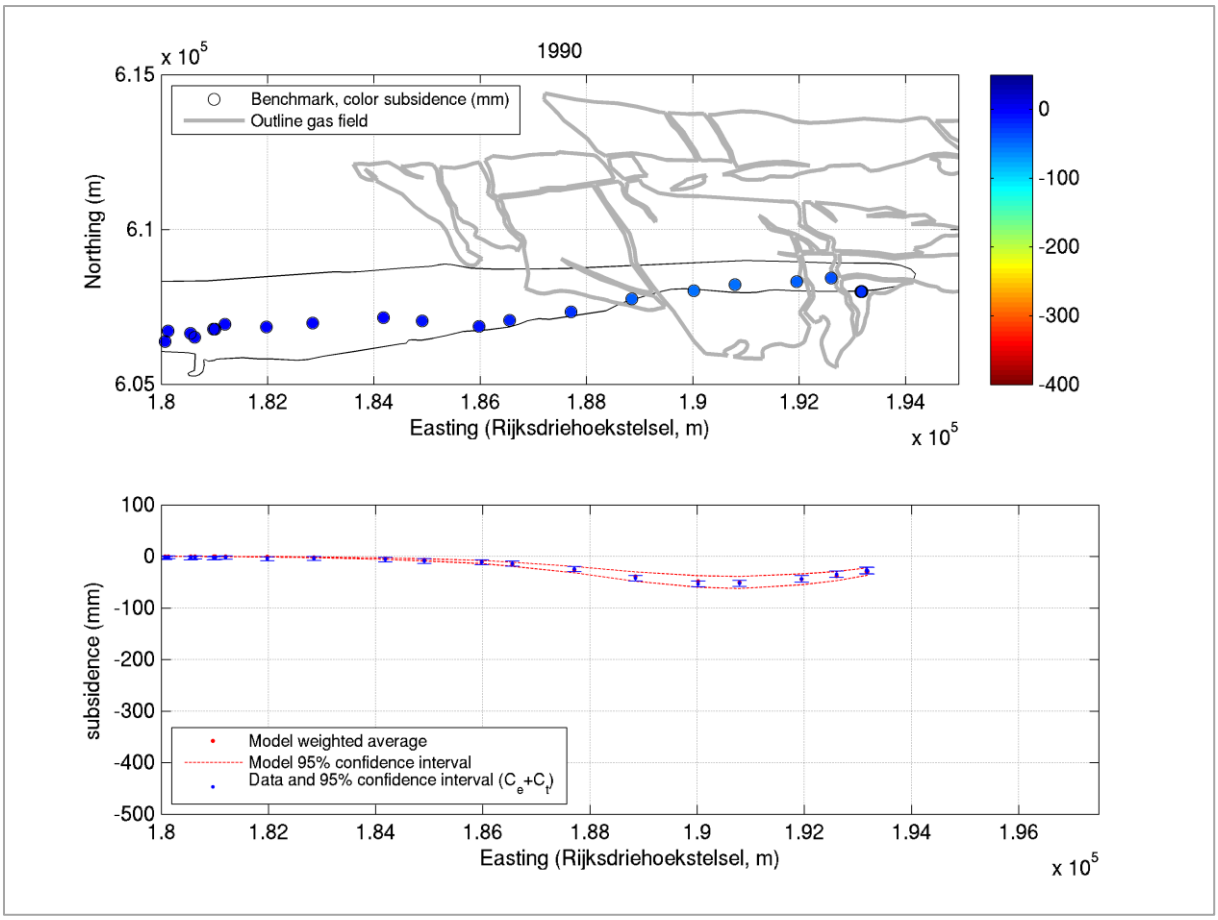

2030
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.96

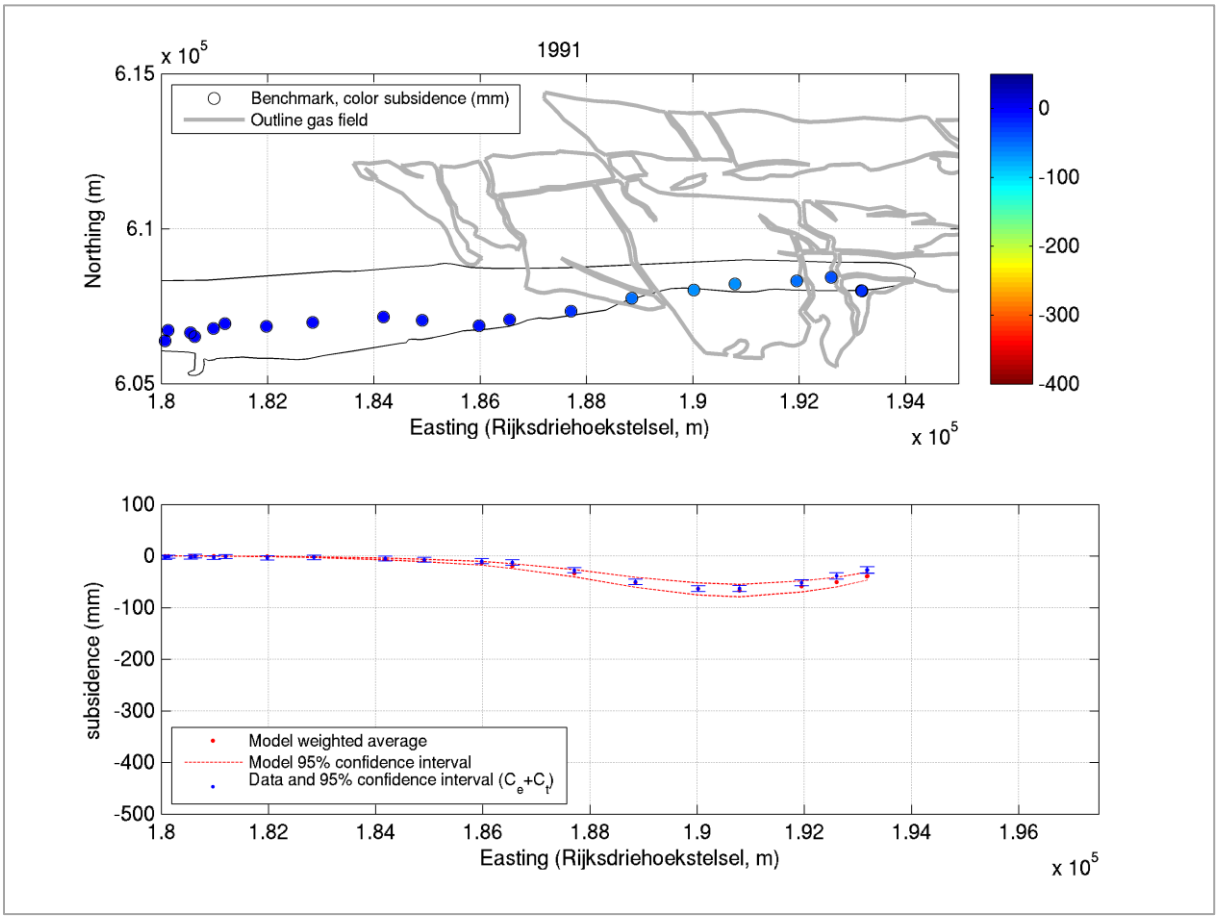

2035
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

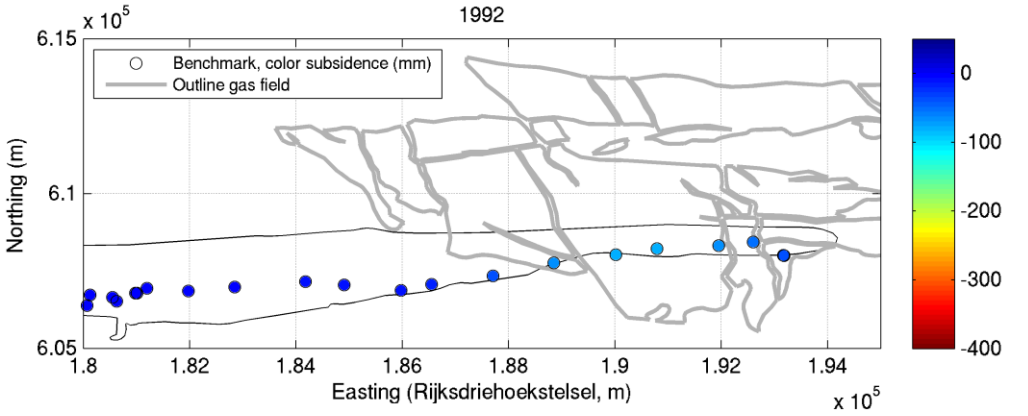

2040
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

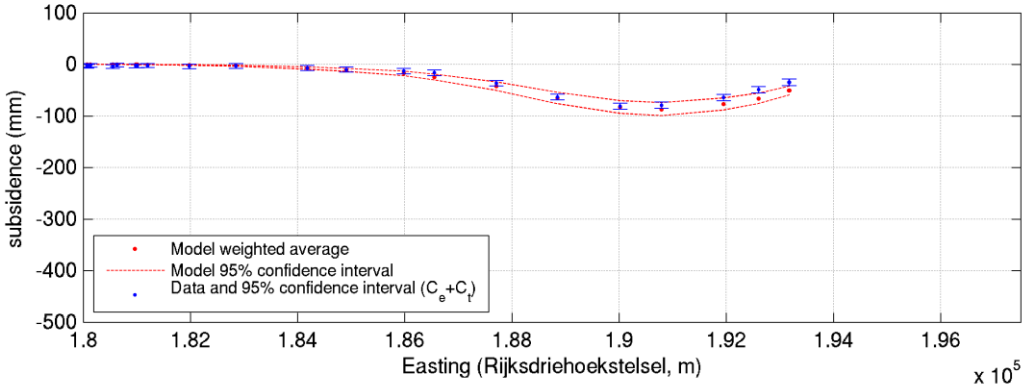

1987
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

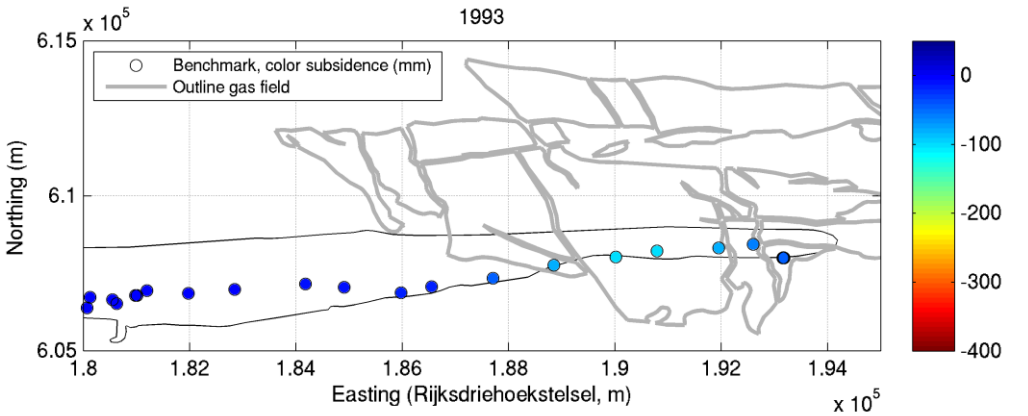

1988
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

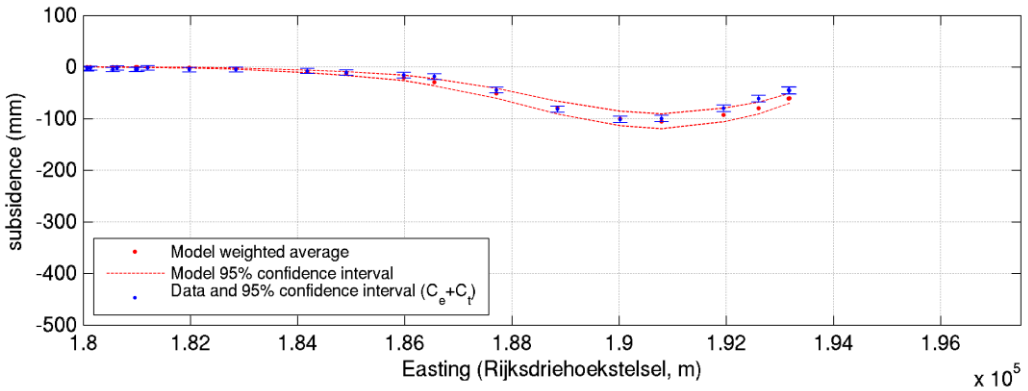

1990
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

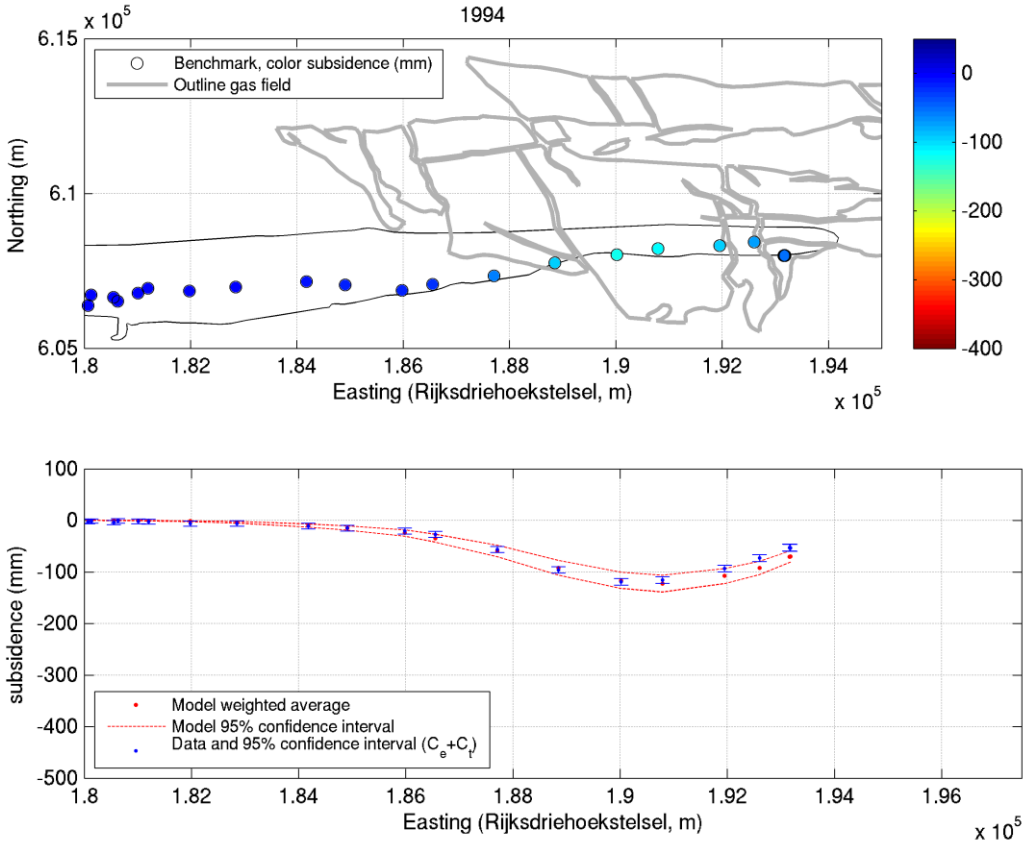

1991
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

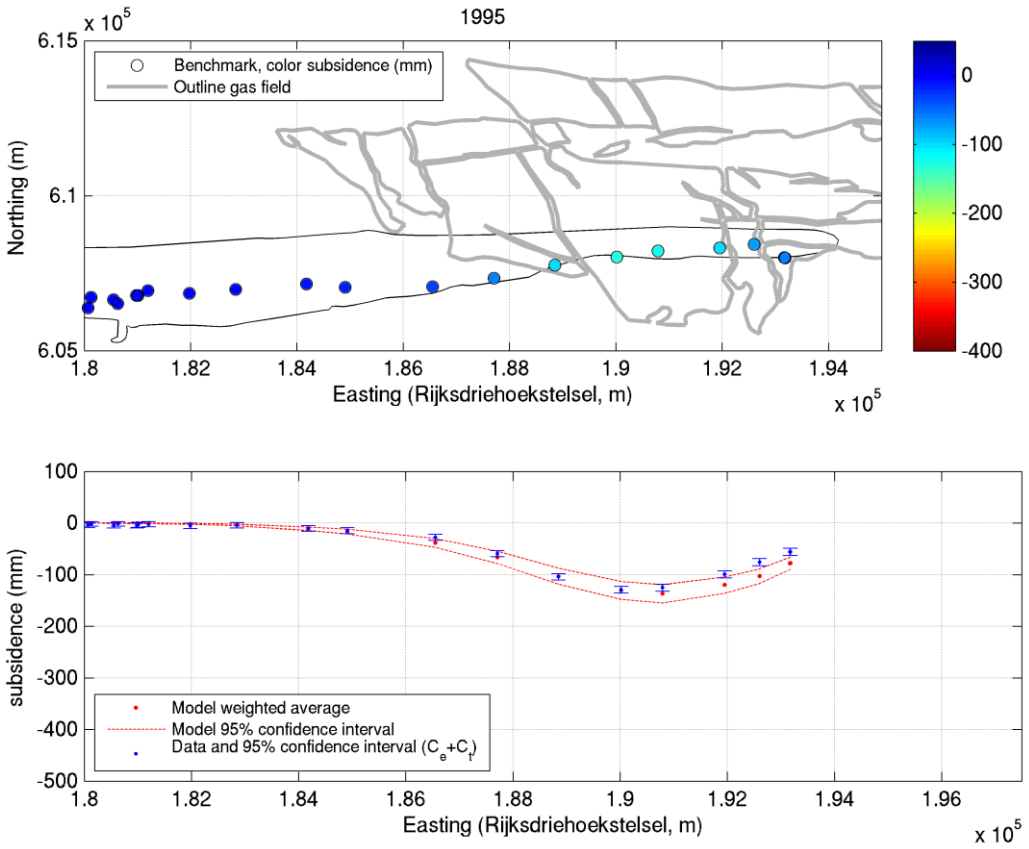

1992
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

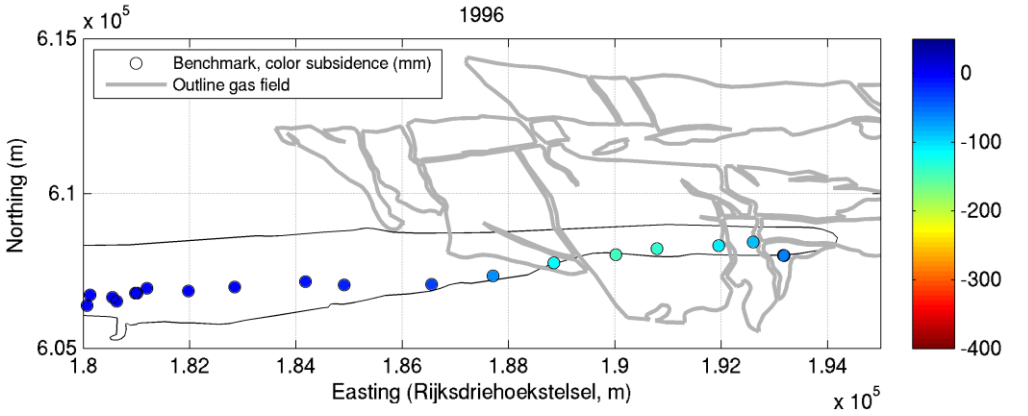

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

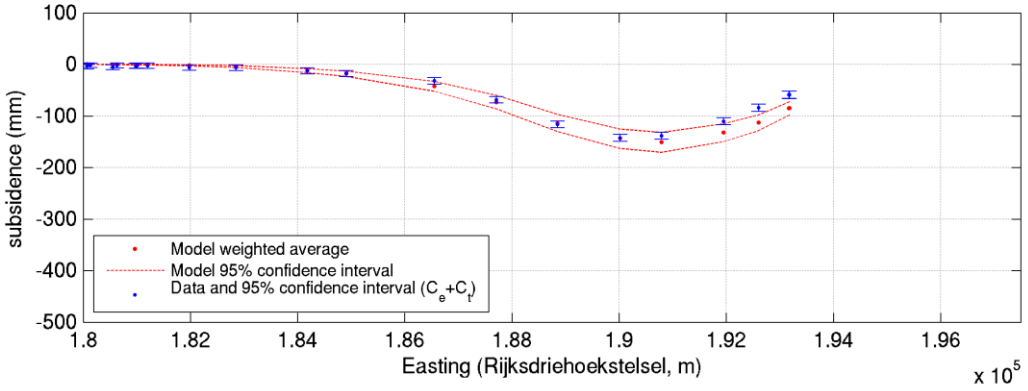

1993
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

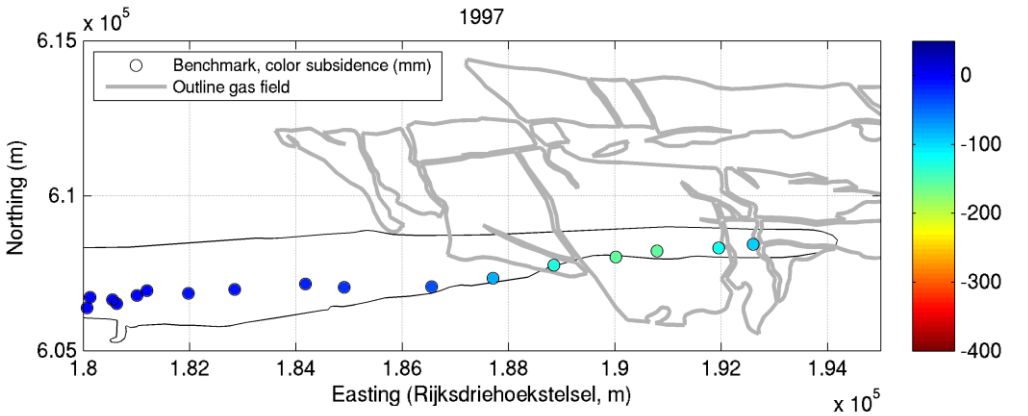

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

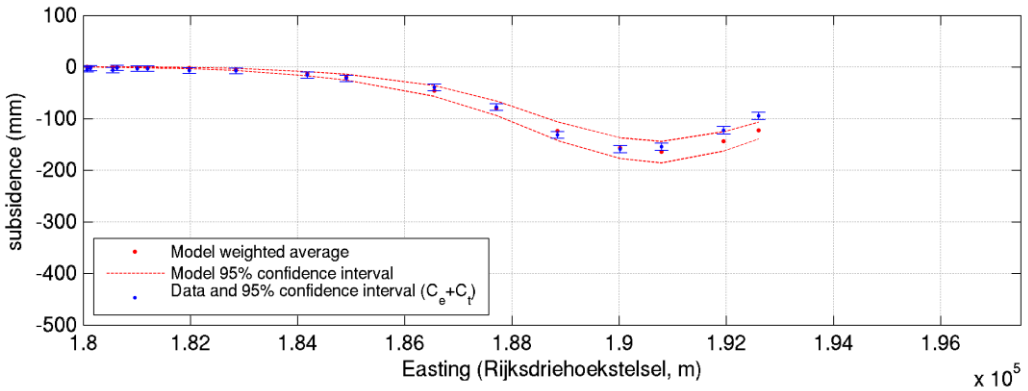

1994
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

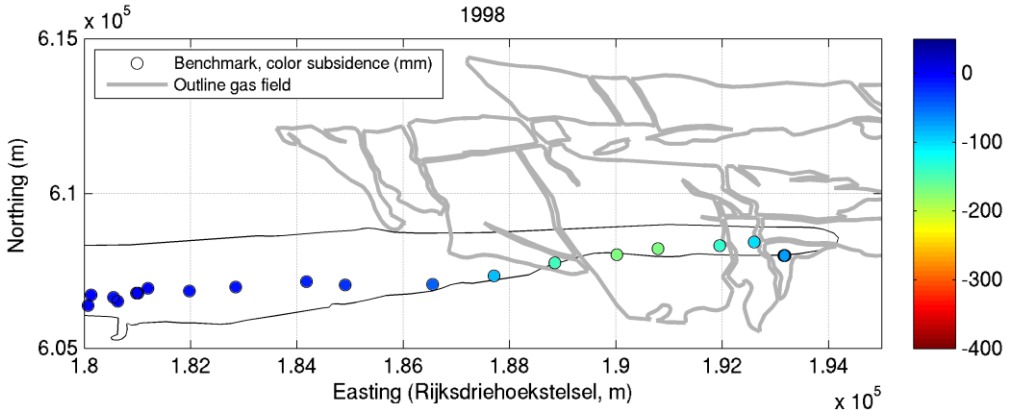

1995
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

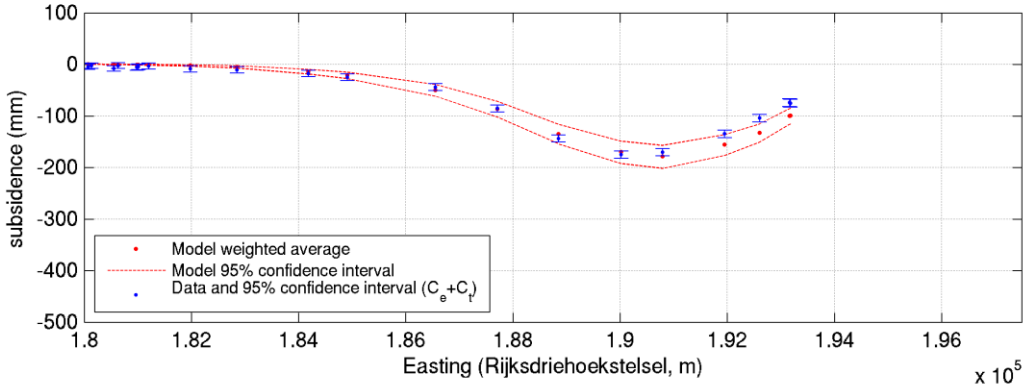

1996
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

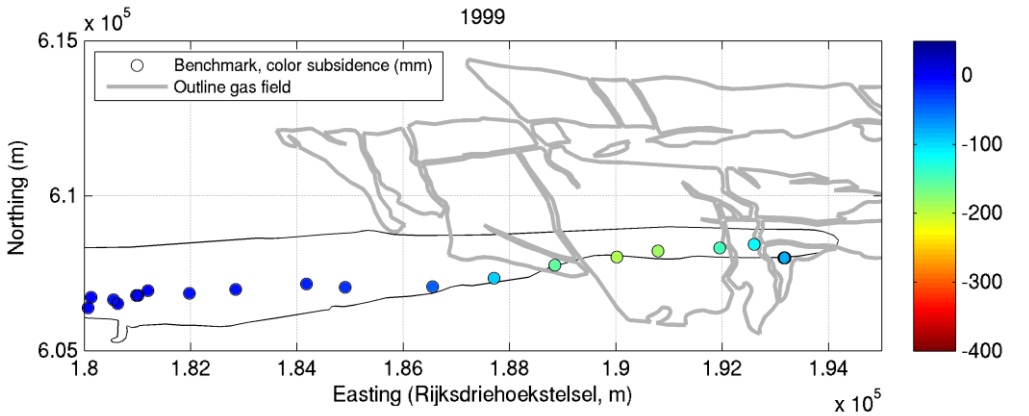

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

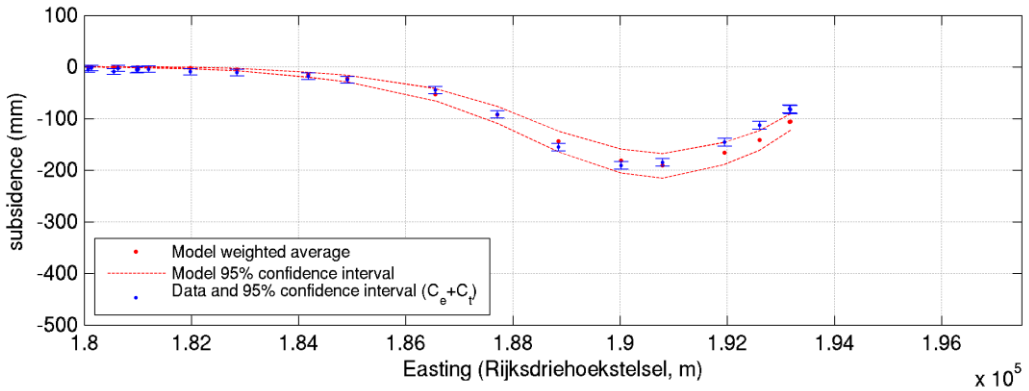

1997
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

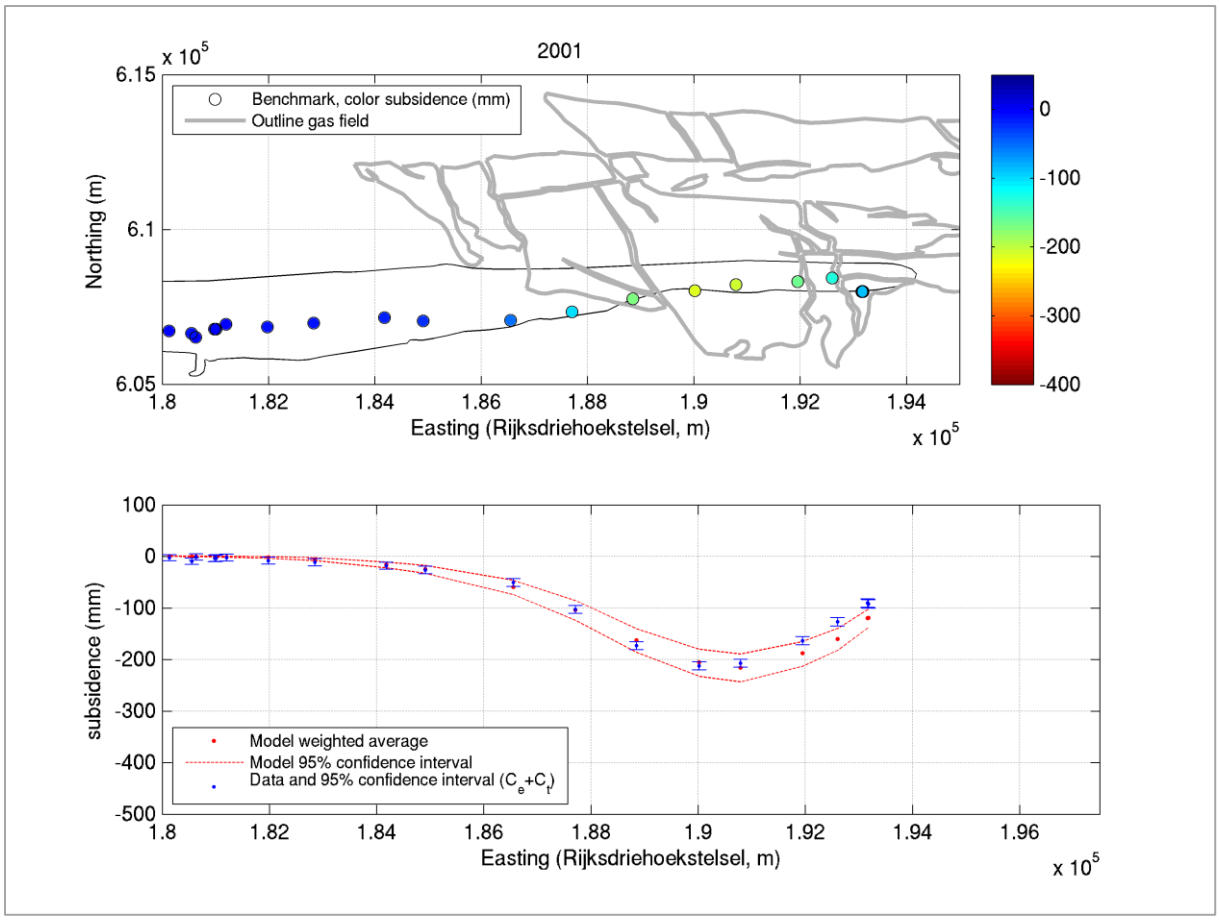

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

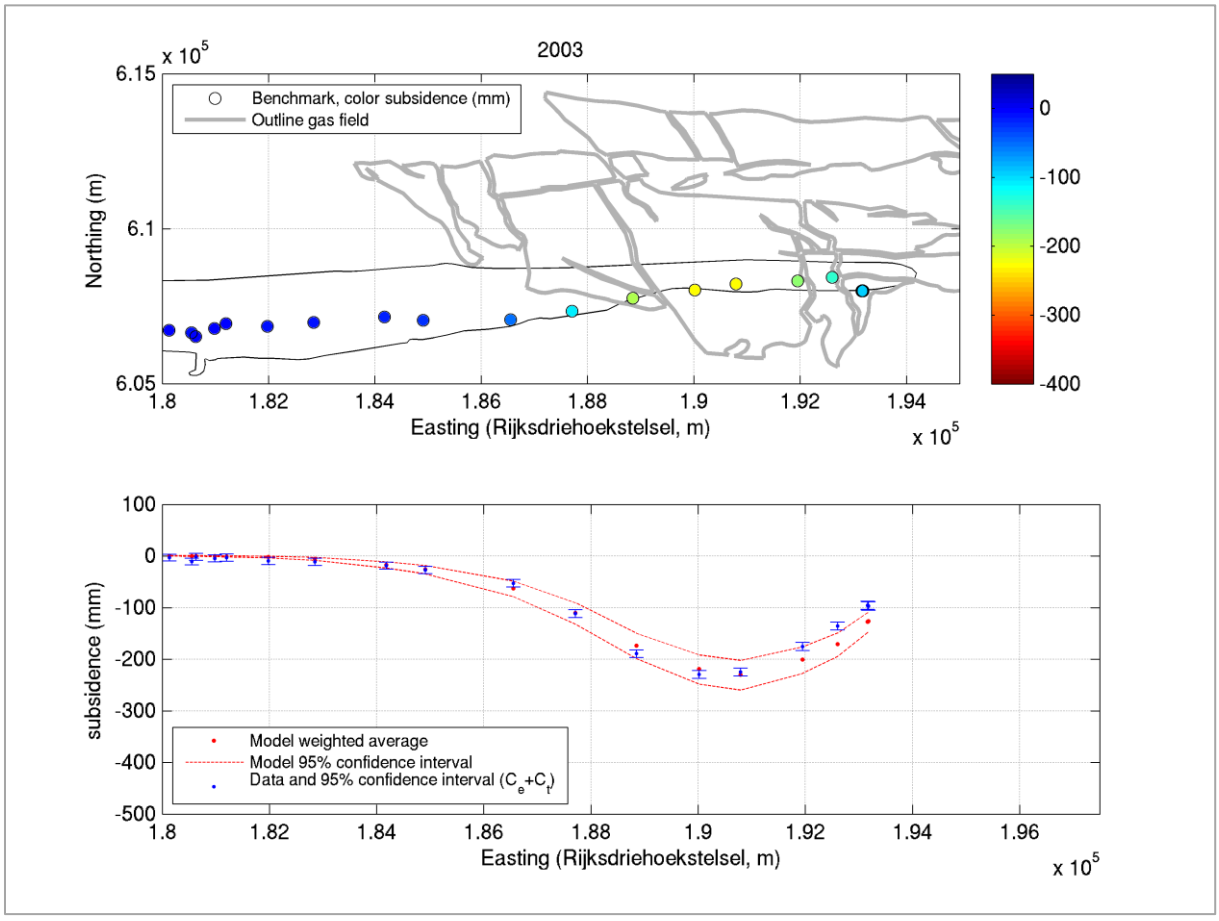

1998
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

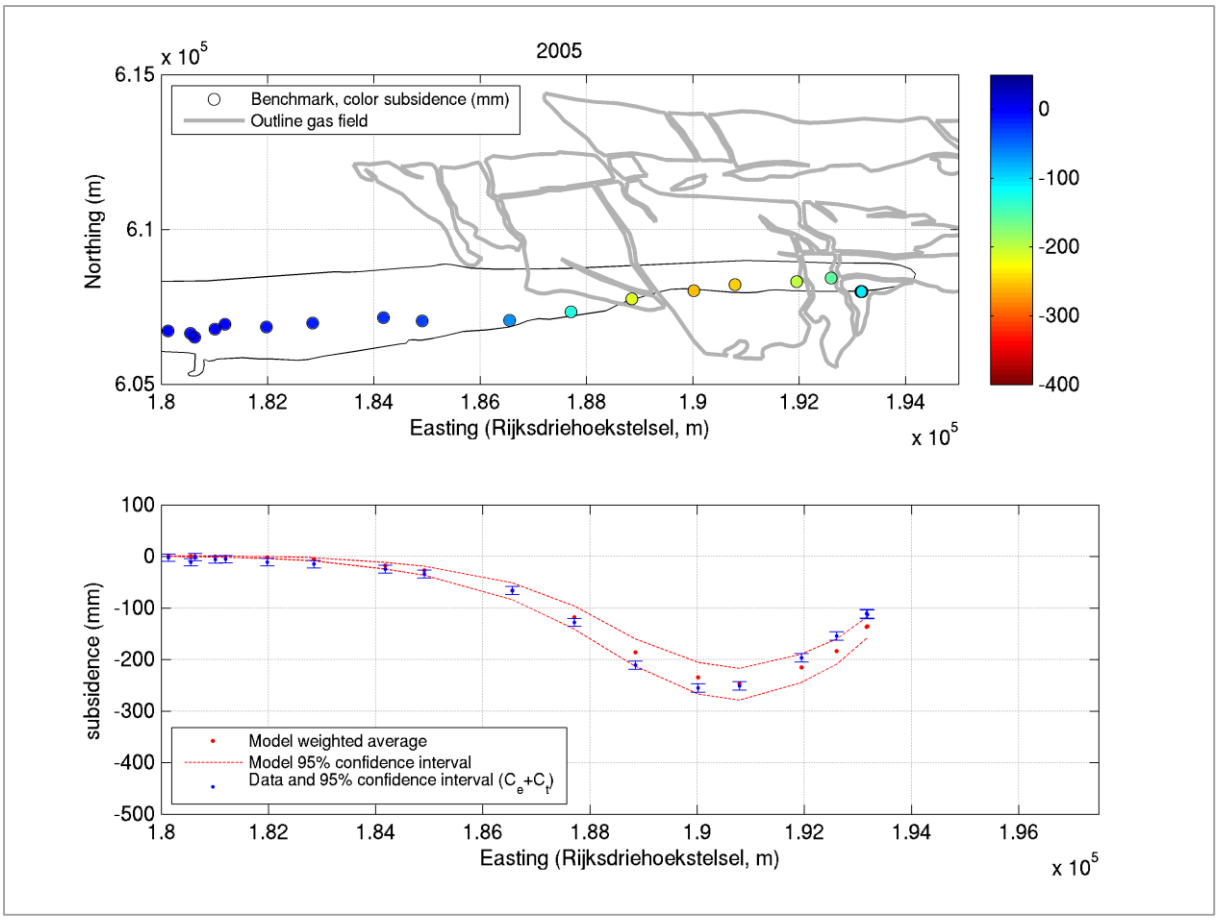

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

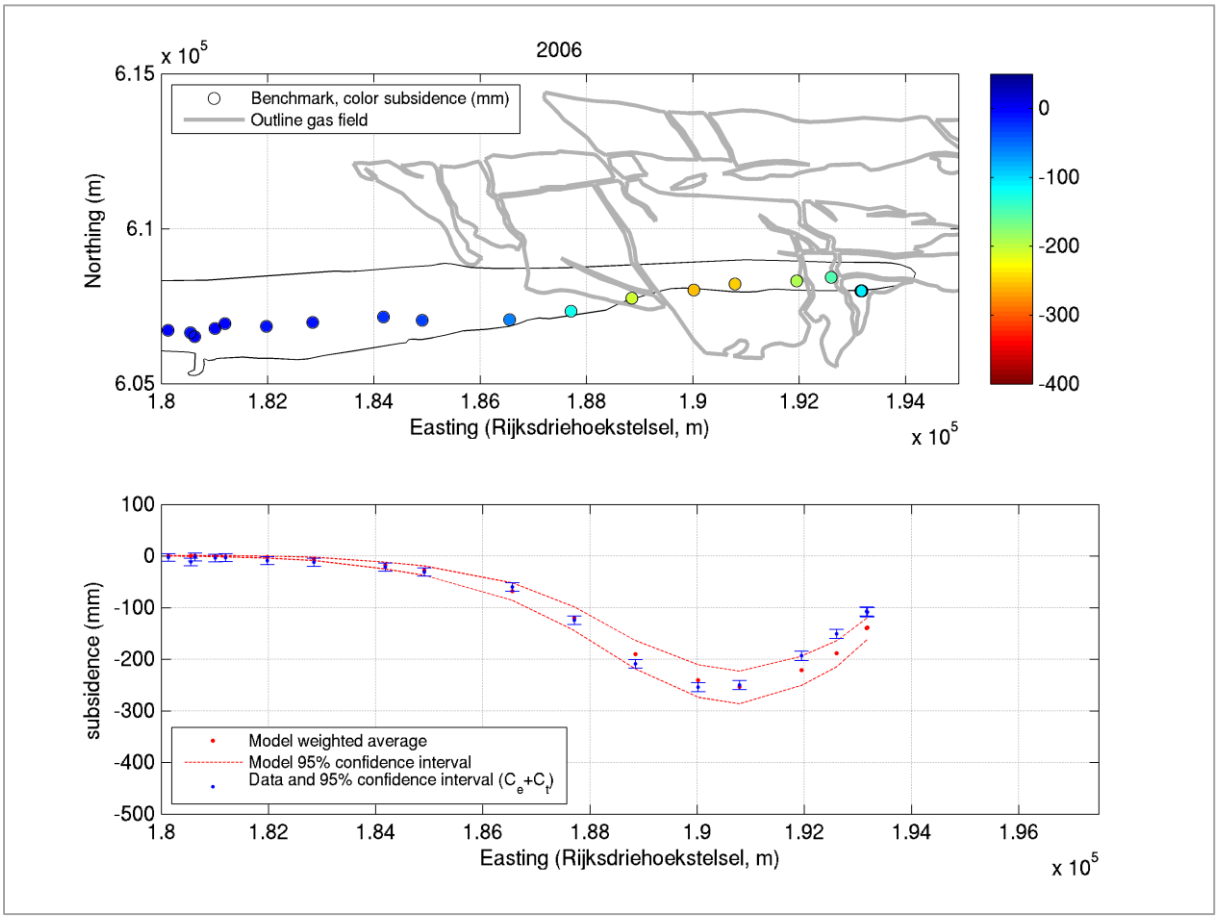

1999
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

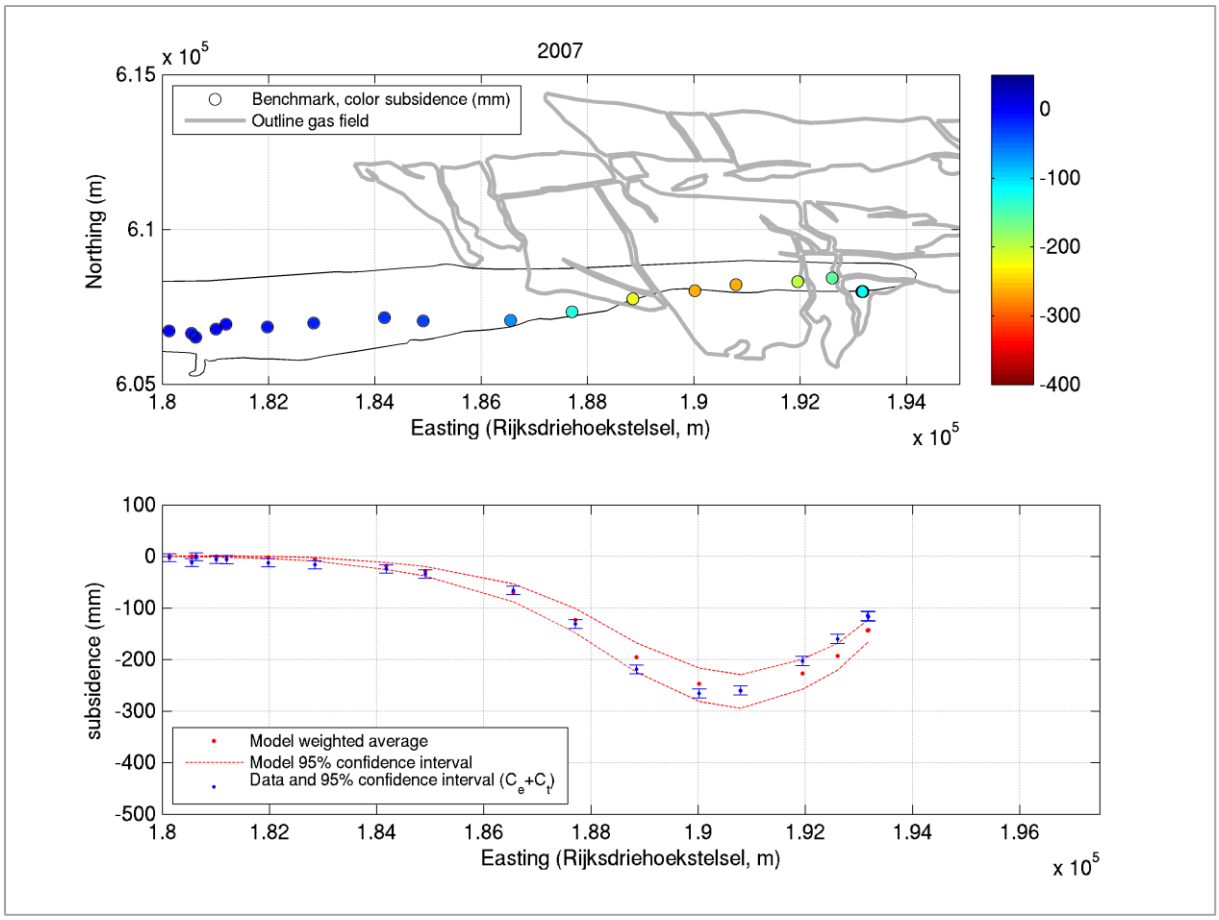

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

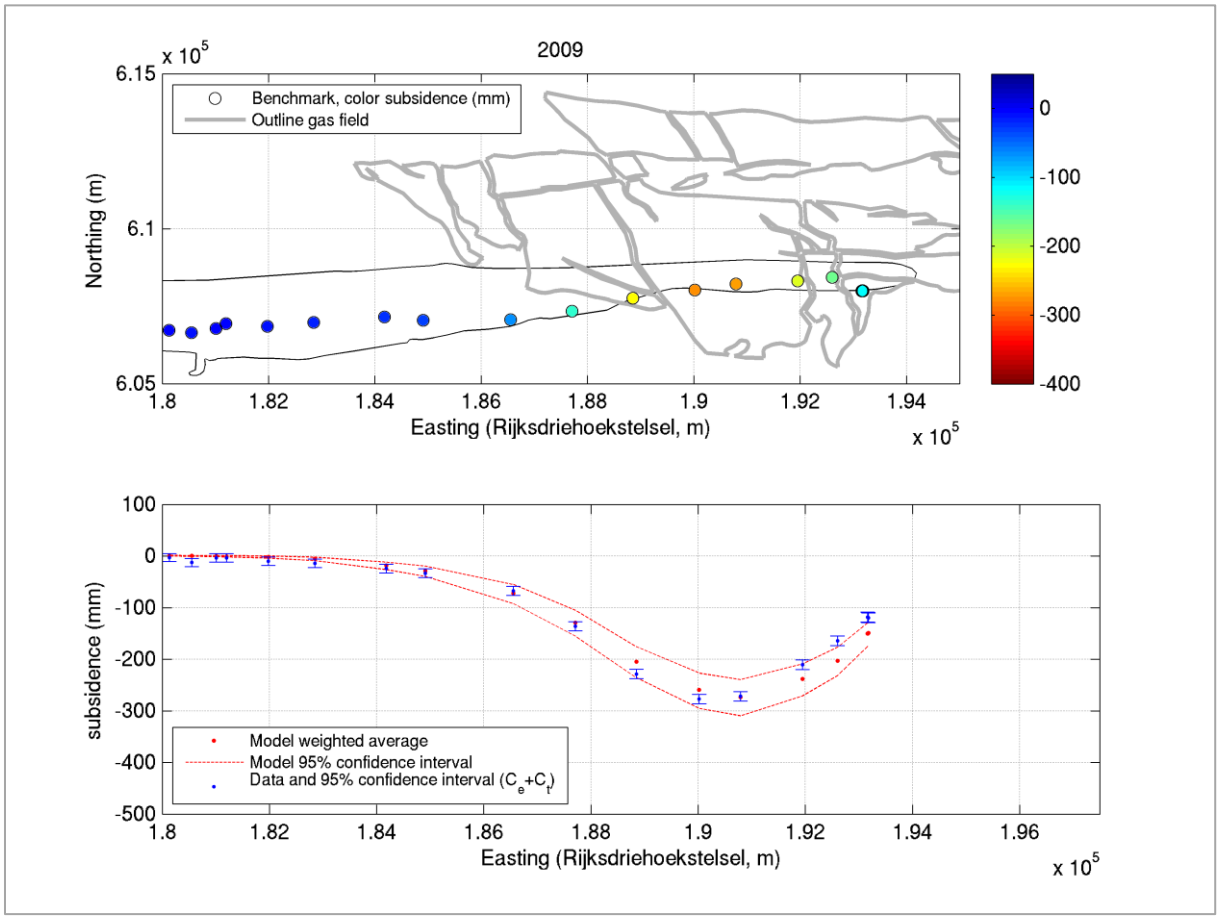

2001
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

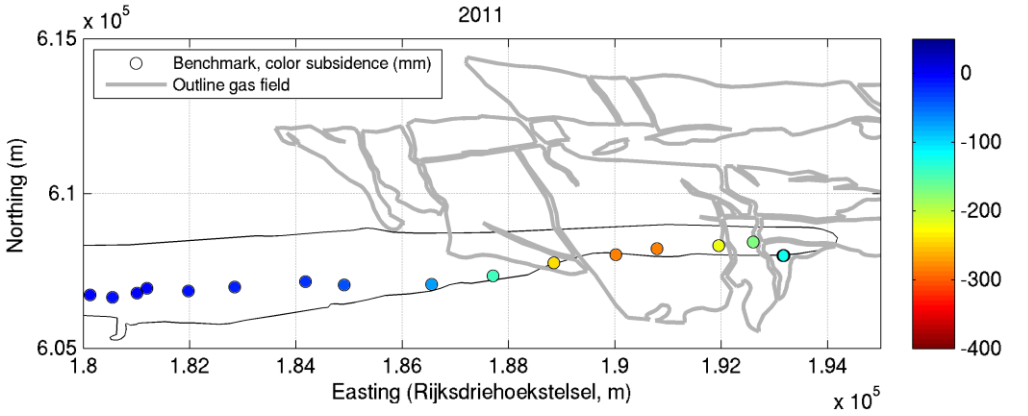

2003
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

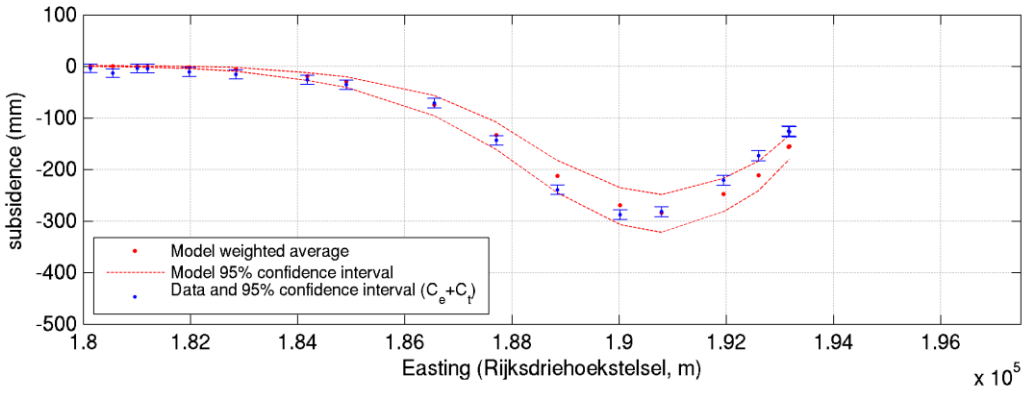

2005
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

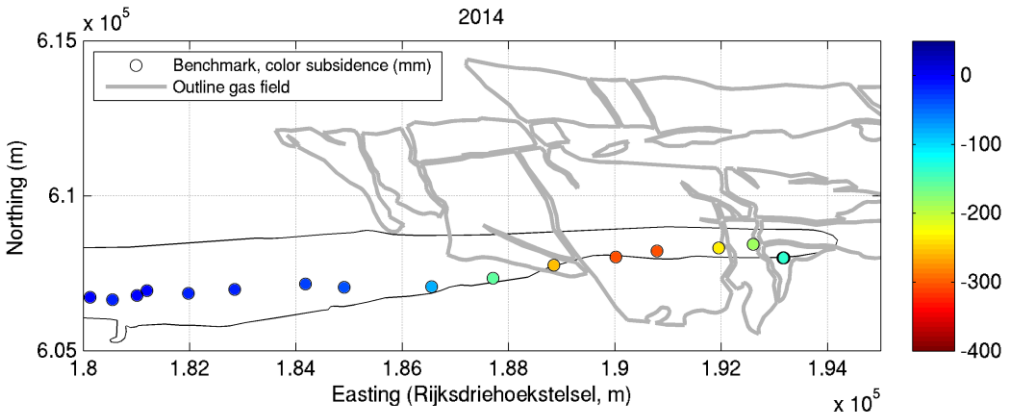

2006
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

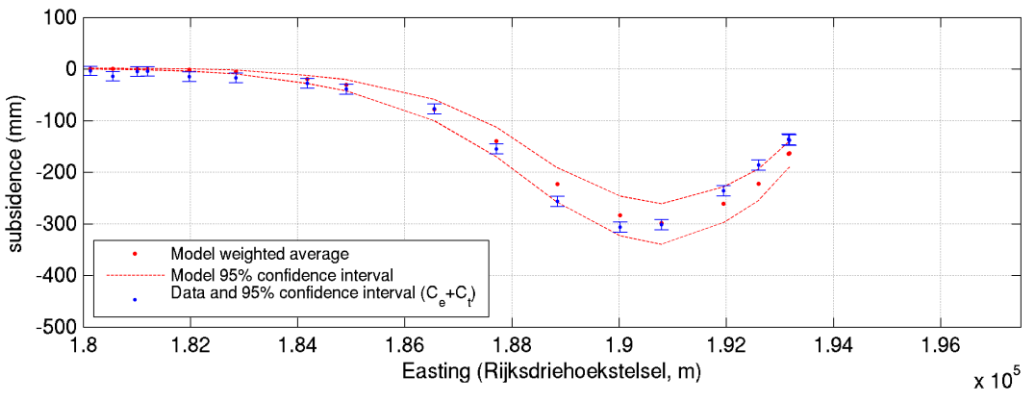

2007
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

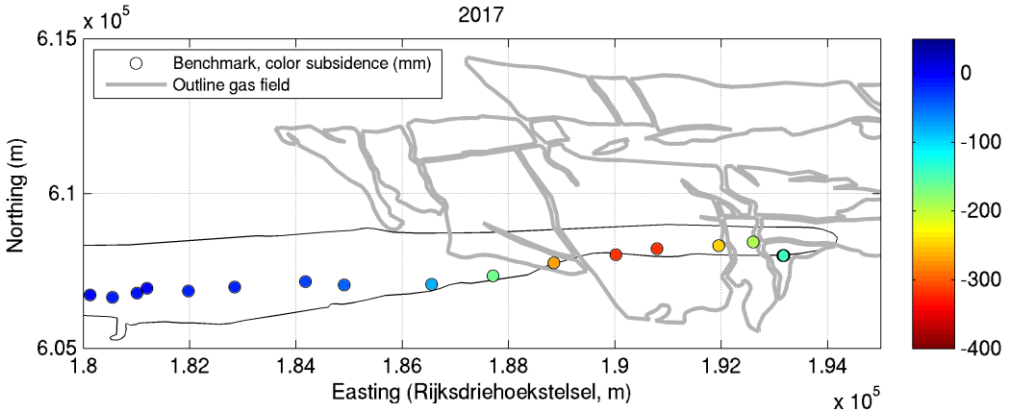

2009
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.96

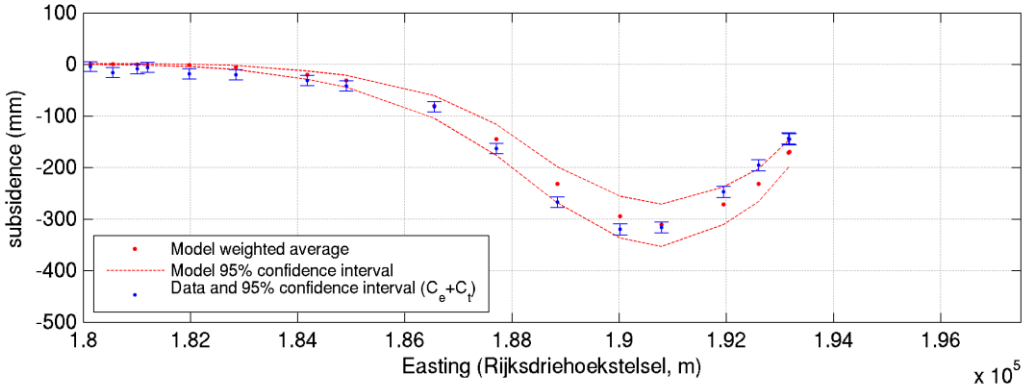

2011
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5

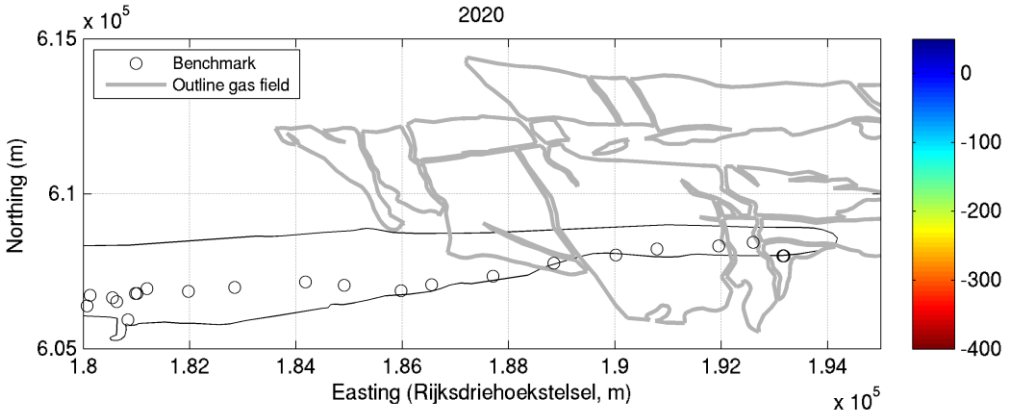

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

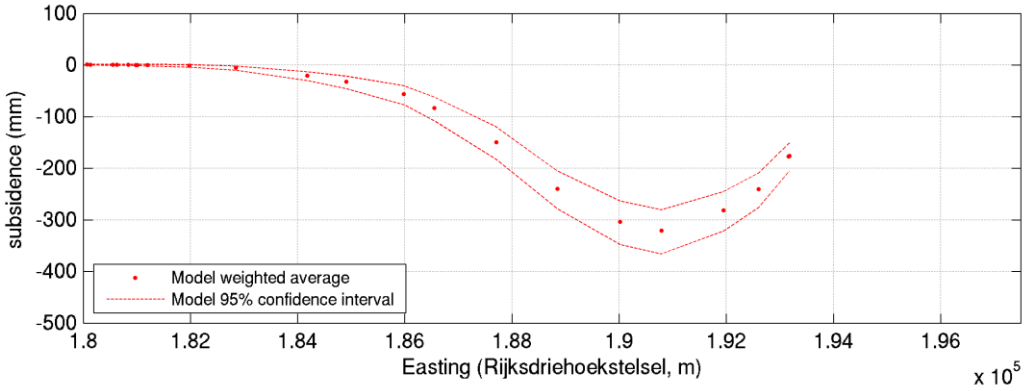

2014
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark, color subsidence (mm)
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5

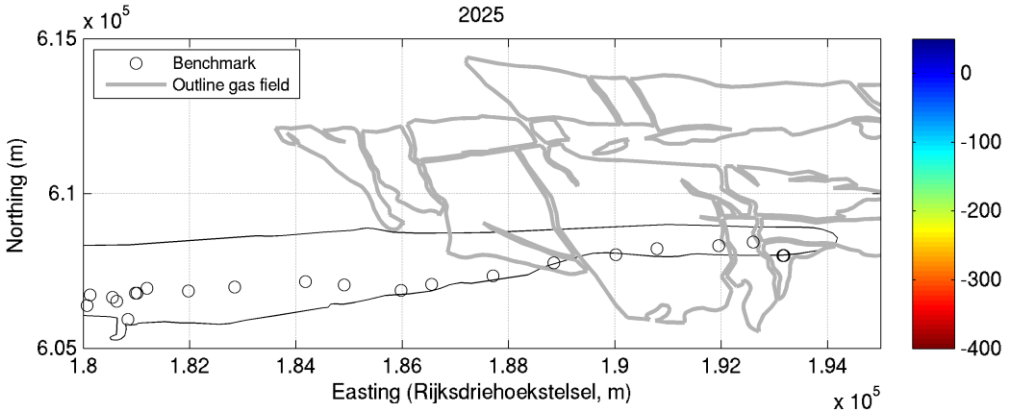

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

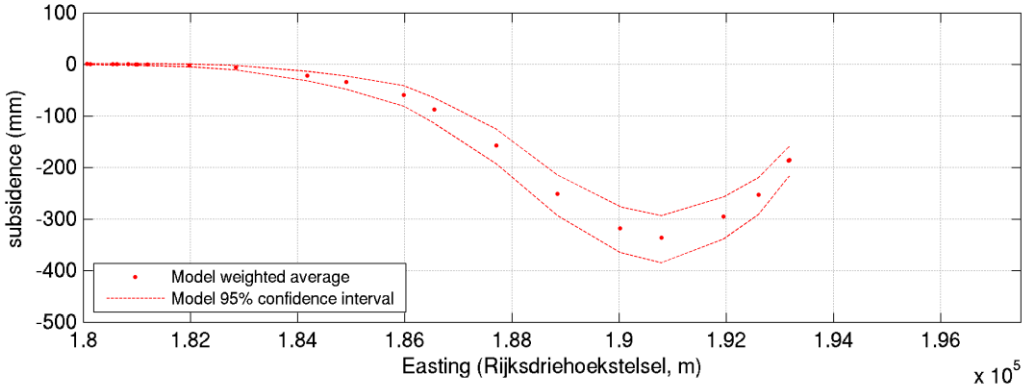

2017
x 10^5
Benchmark, color subsidence (mm)
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

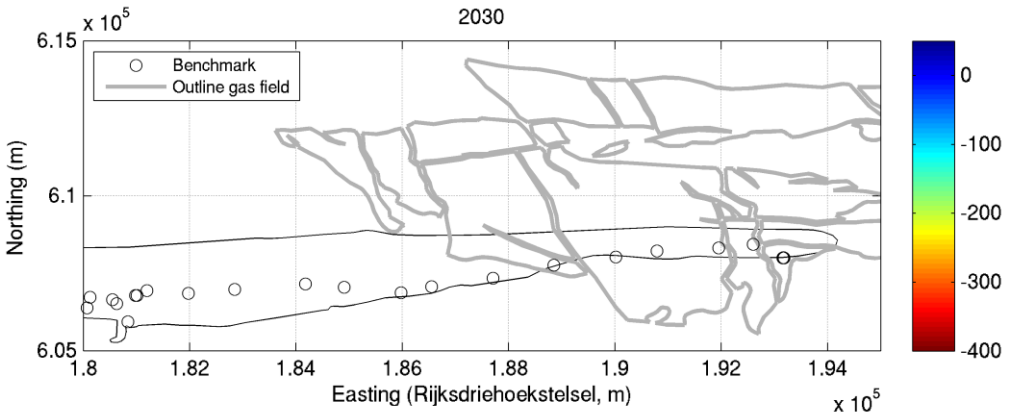

subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
Easting (Rijksdriehoekstelsel, m)
x 10^5

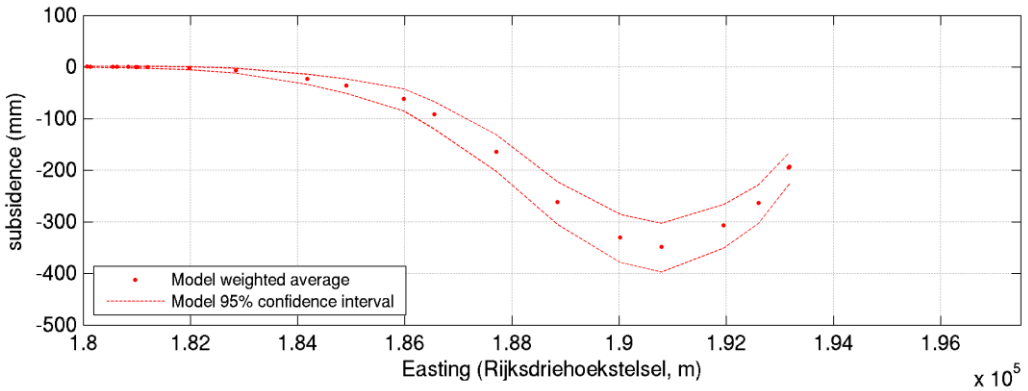

2020
x 10^5
Benchmark
Outline gas field
Northing (m)
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

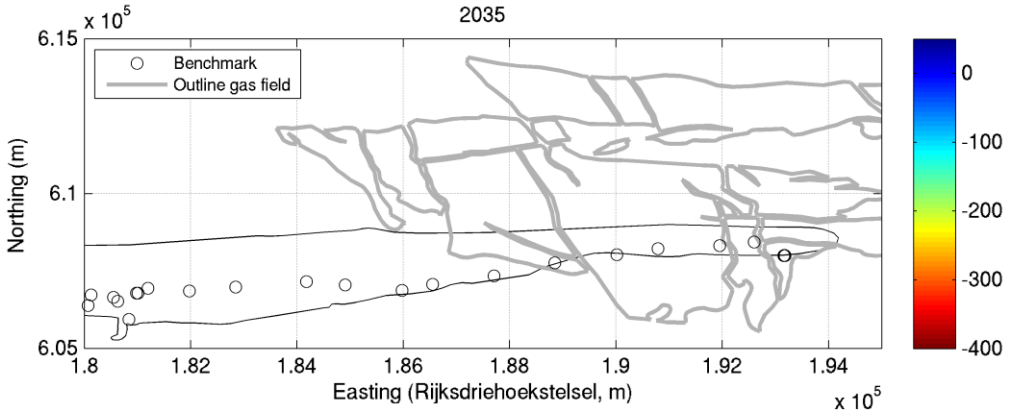

subsidence (mm)
Model weighted average
Model 95% confidence interval
Easting (Rijksdriehoekstelsel, m)
x 10^5

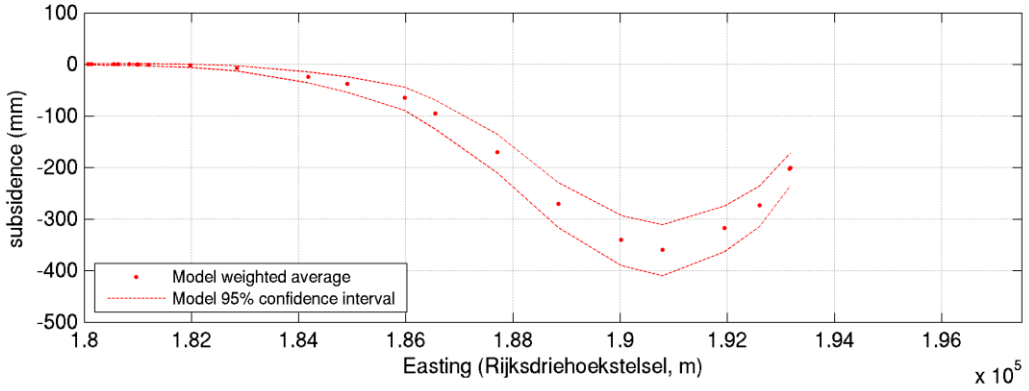

2025
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5

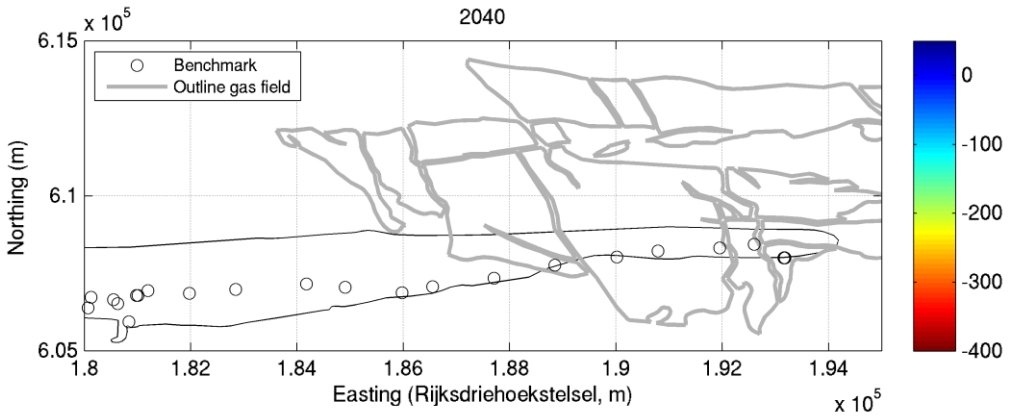

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

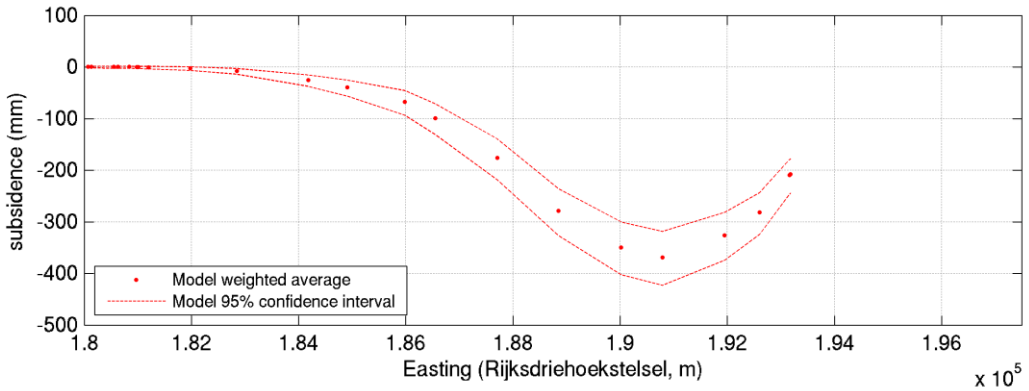

2030
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
0
-100
-200
-300
-400
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5

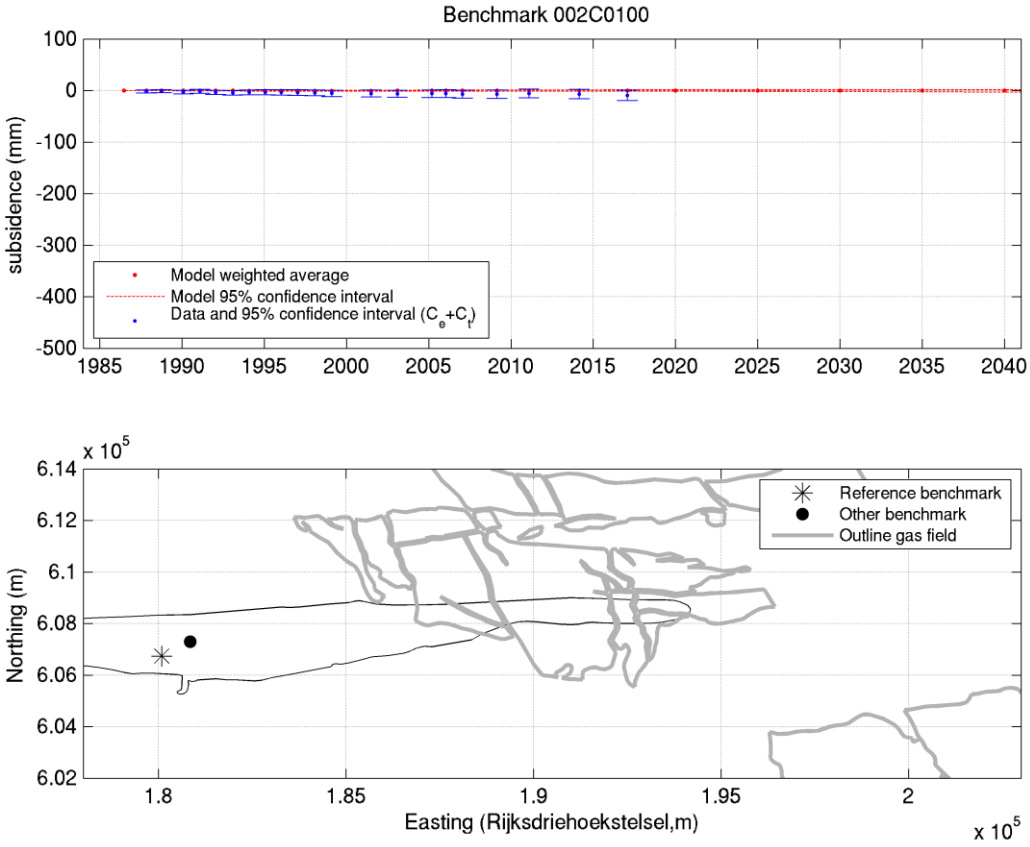

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

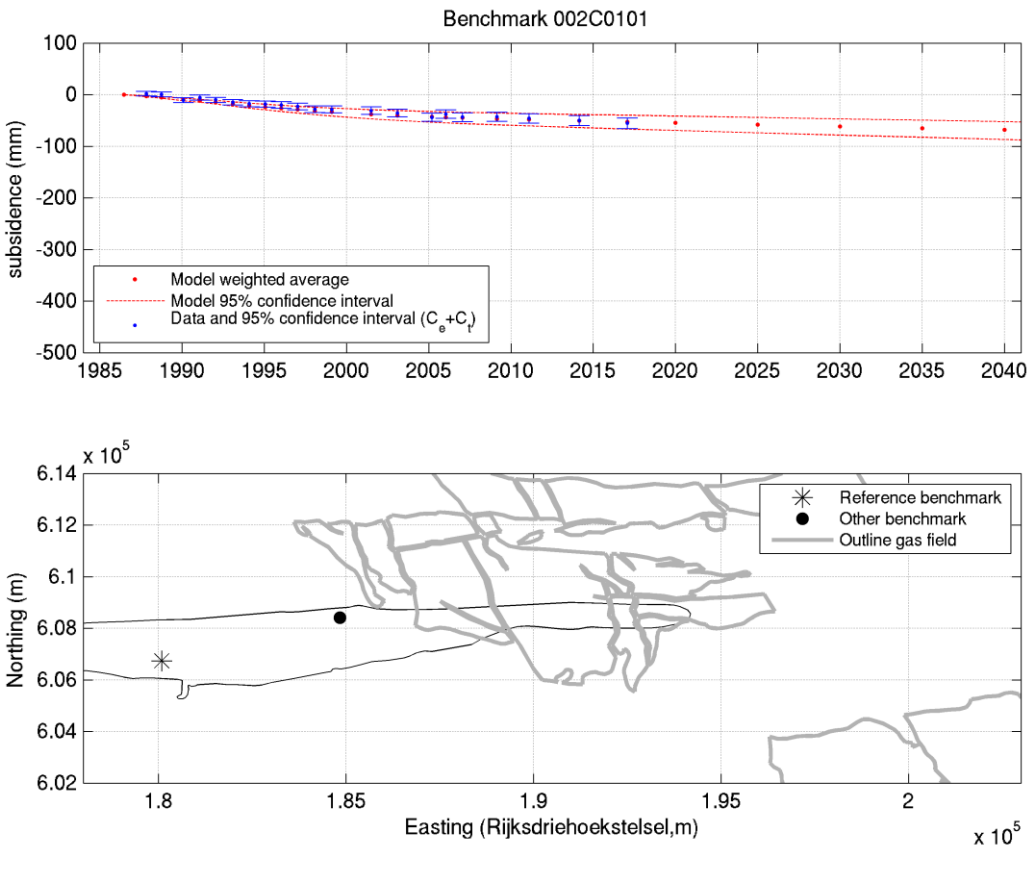

2035
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

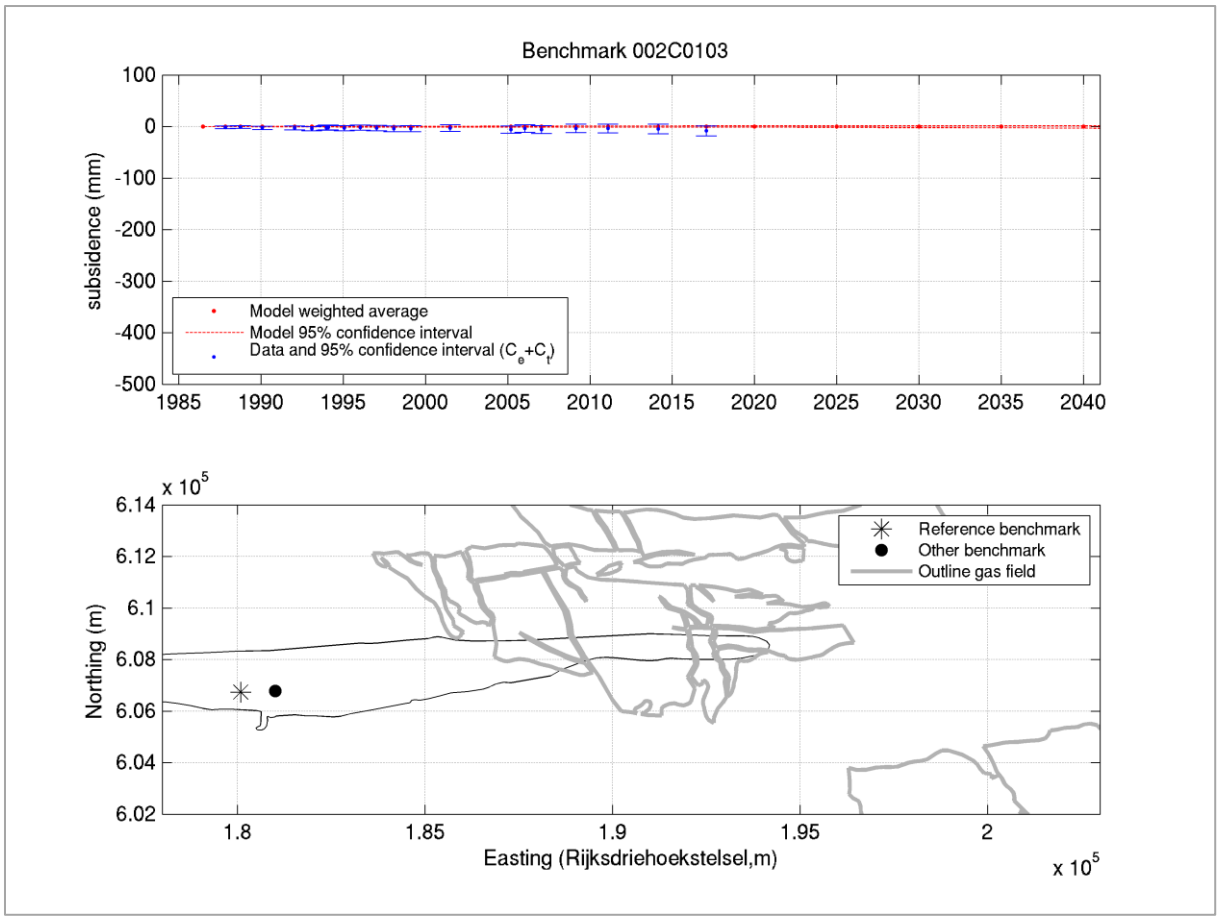

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

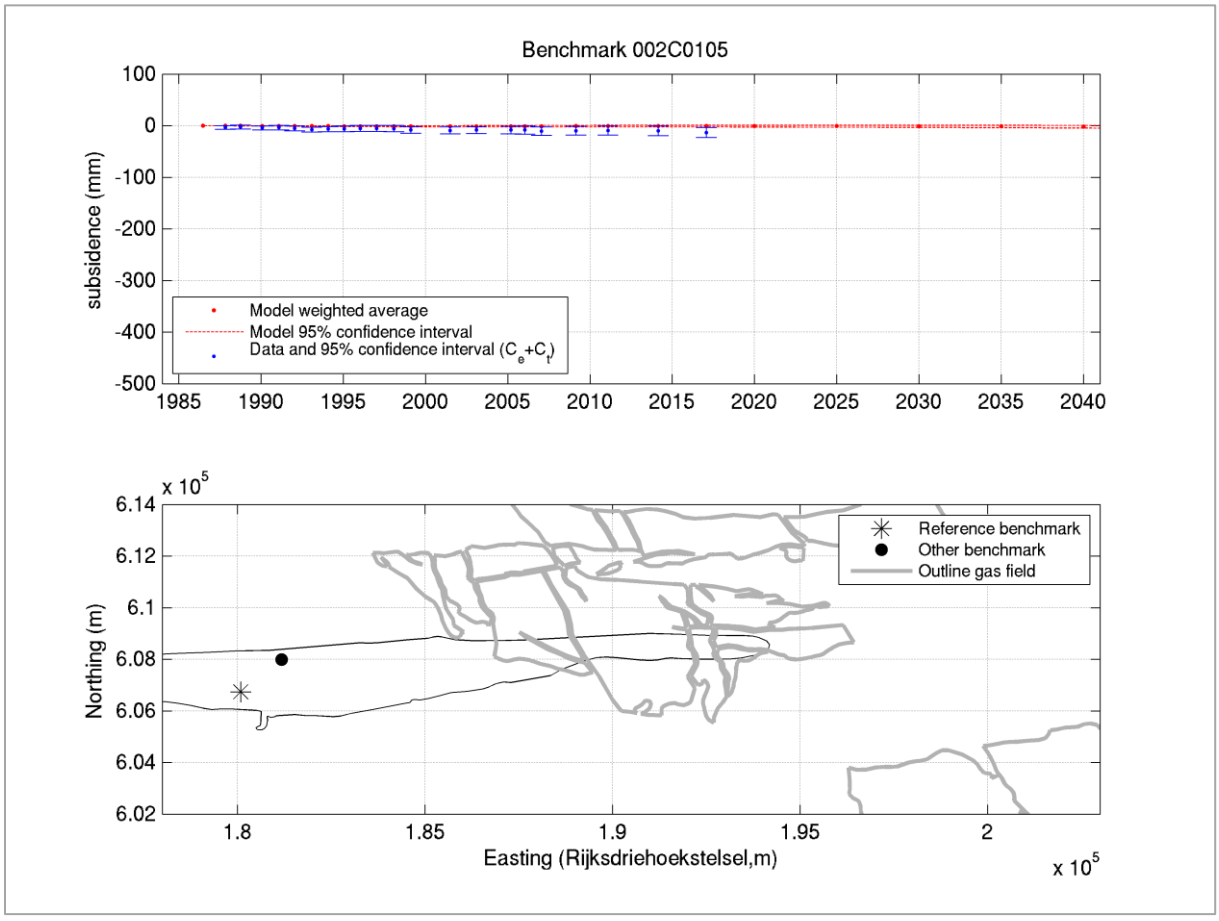

2040
x 10^5
6.15
6.1
6.05
Northing (m)
Benchmark
Outline gas field
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Easting (Rijksdriehoekstelsel, m)
x 10^5
0
-100
-200
-300
-400

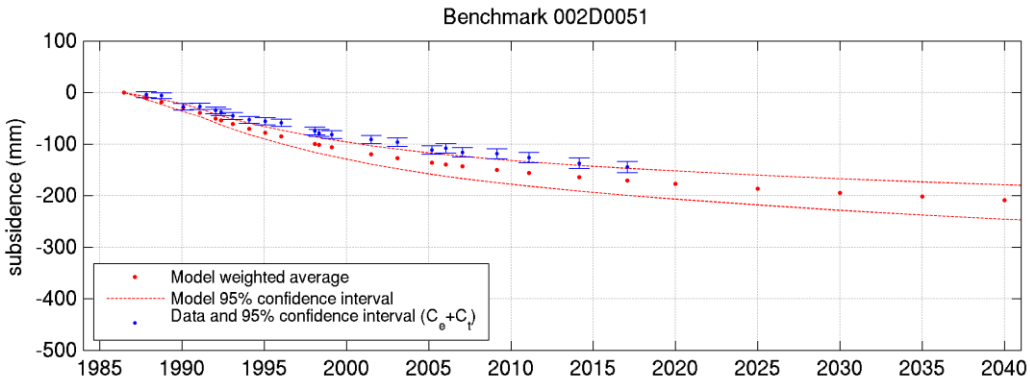

100
0
-100
-200
-300
-400
-500
subsidence (mm)
Model weighted average
Model 95% confidence interval
1.8
1.82
1.84
1.86
1.88
1.9
1.92
1.94
1.96
Easting (Rijksdriehoekstelsel, m)
x 10^5

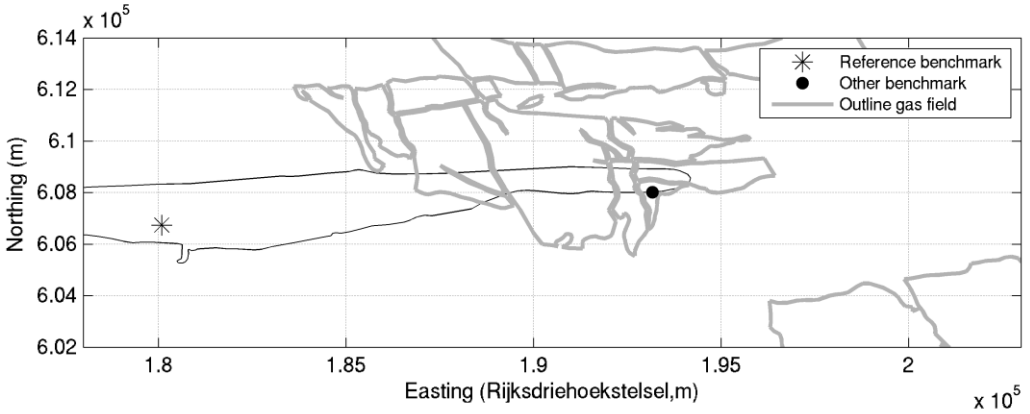

Benchmark 002C0100
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

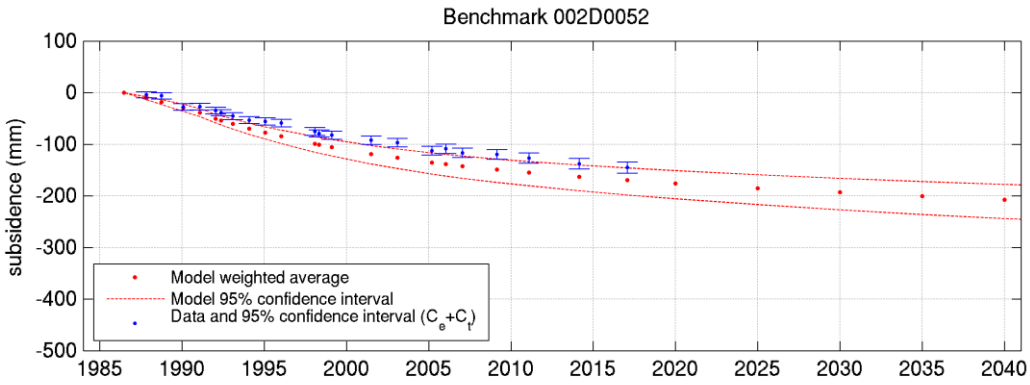

Benchmark 002C0101
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

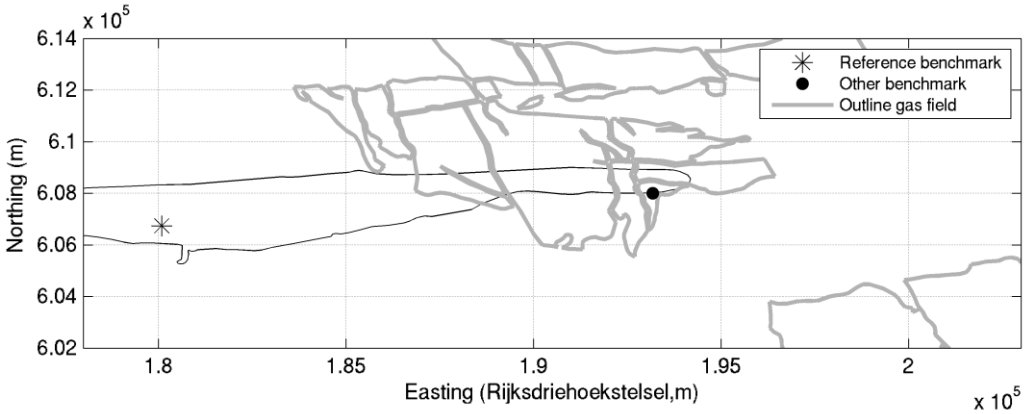

Benchmark 002C0103
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

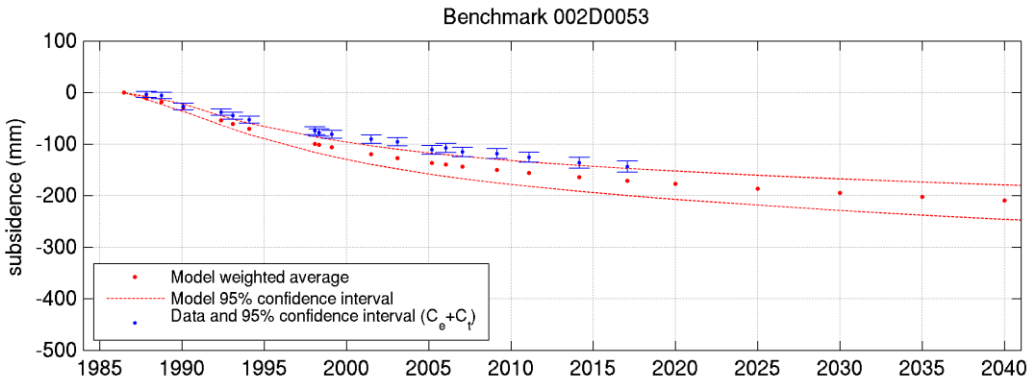

Benchmark 002C0105
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

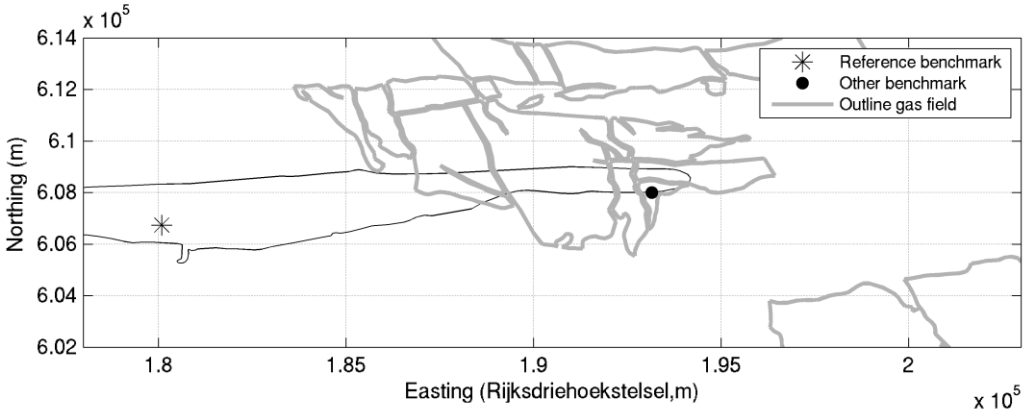

Benchmark 002D0051
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

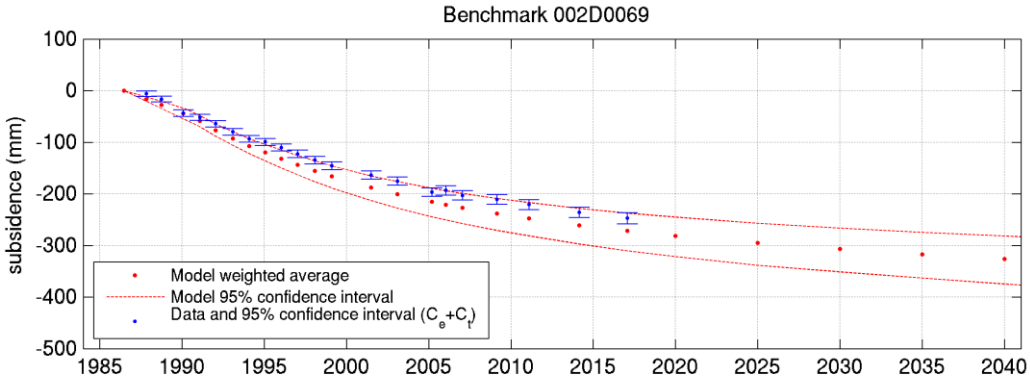

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

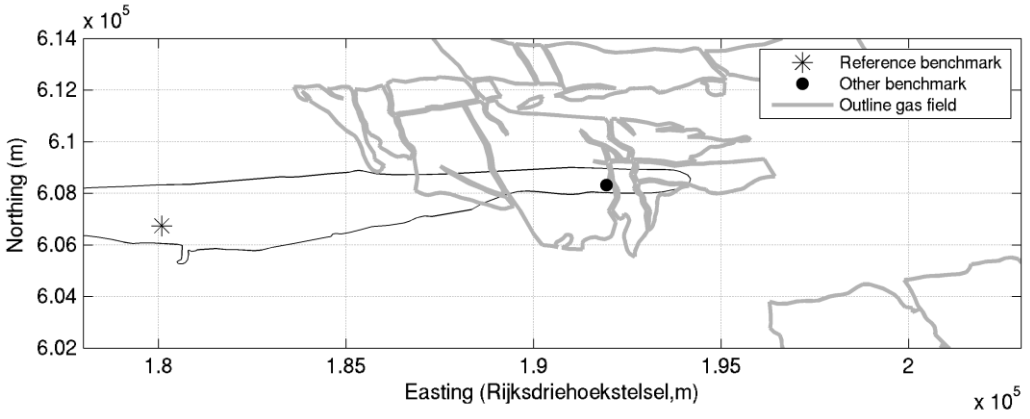

Benchmark 002D0052
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

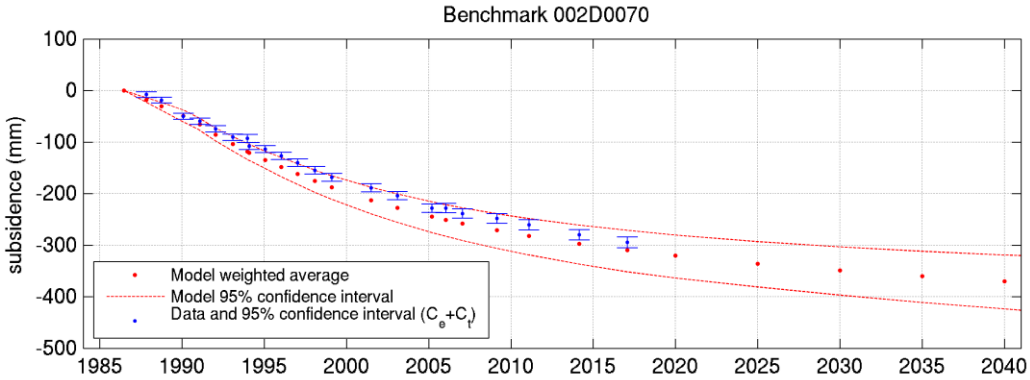

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

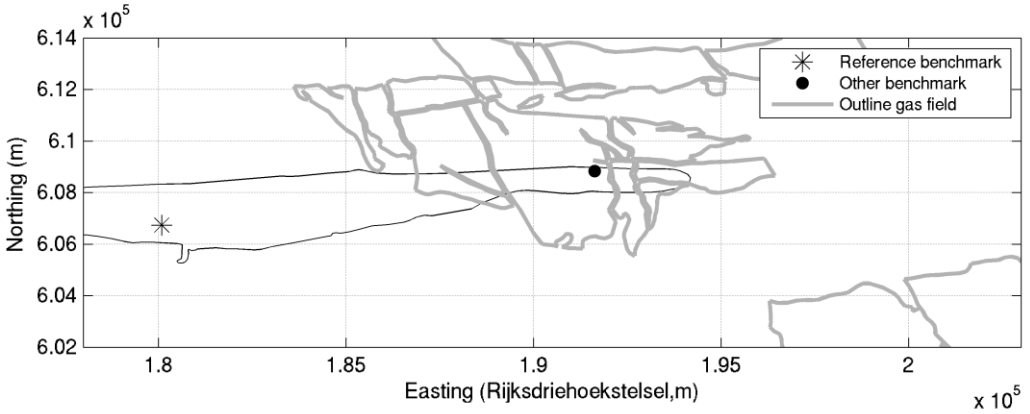

Benchmark 002D0053
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

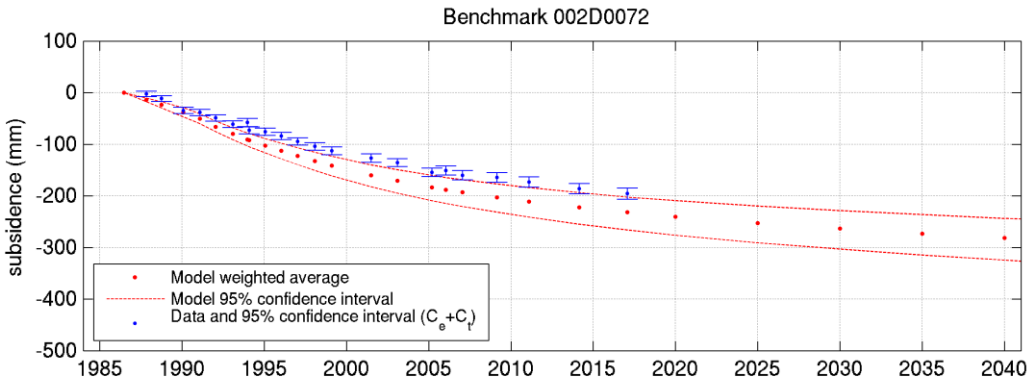

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

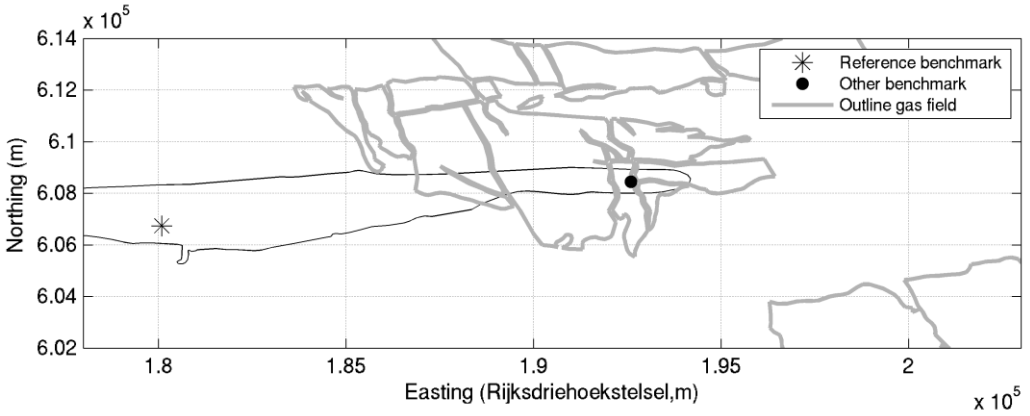

Benchmark 002D0069
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

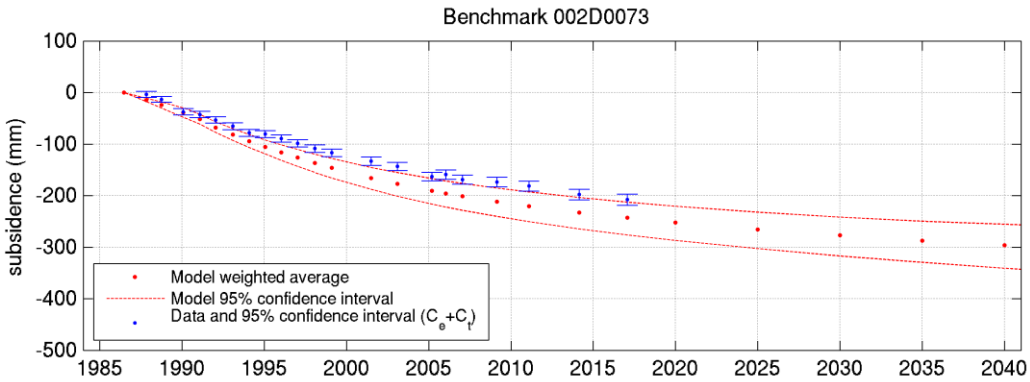

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

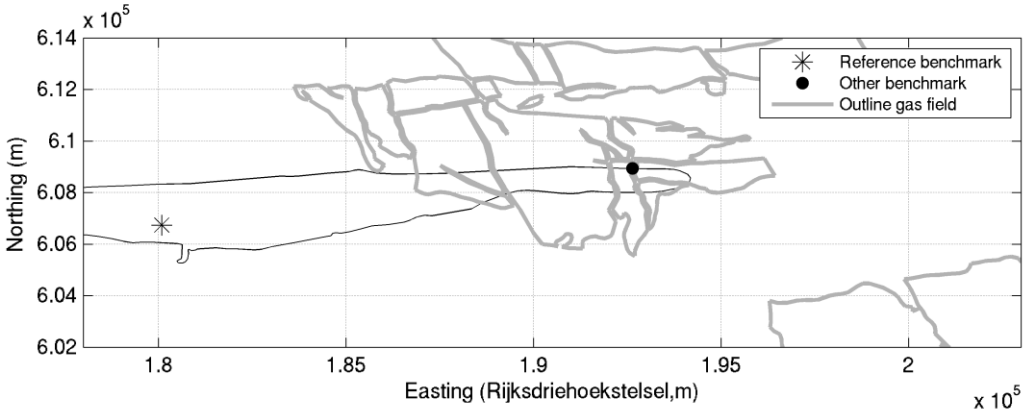

Benchmark 002D0070
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

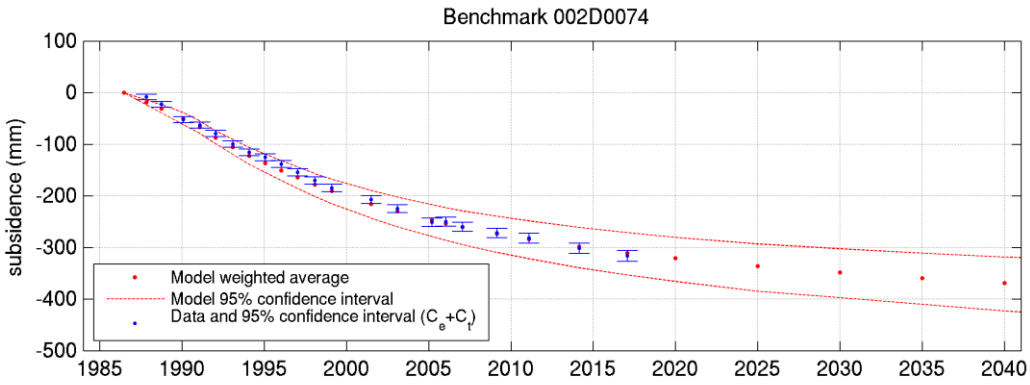

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

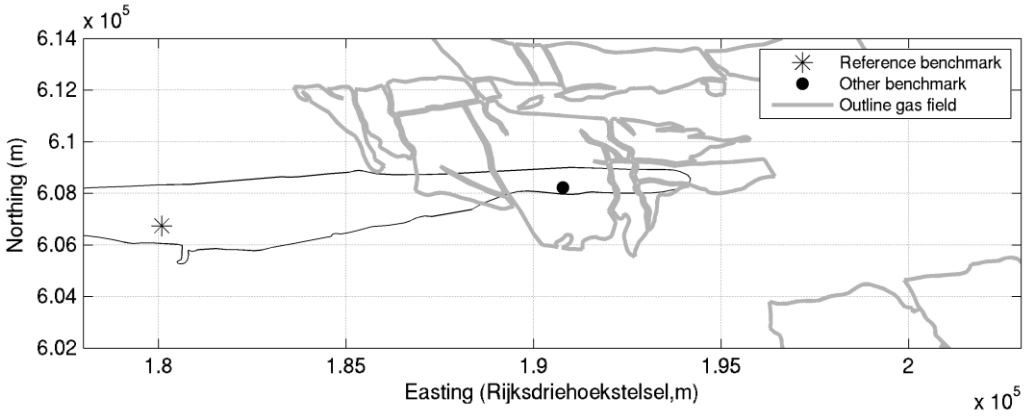

Benchmark 002D0072
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

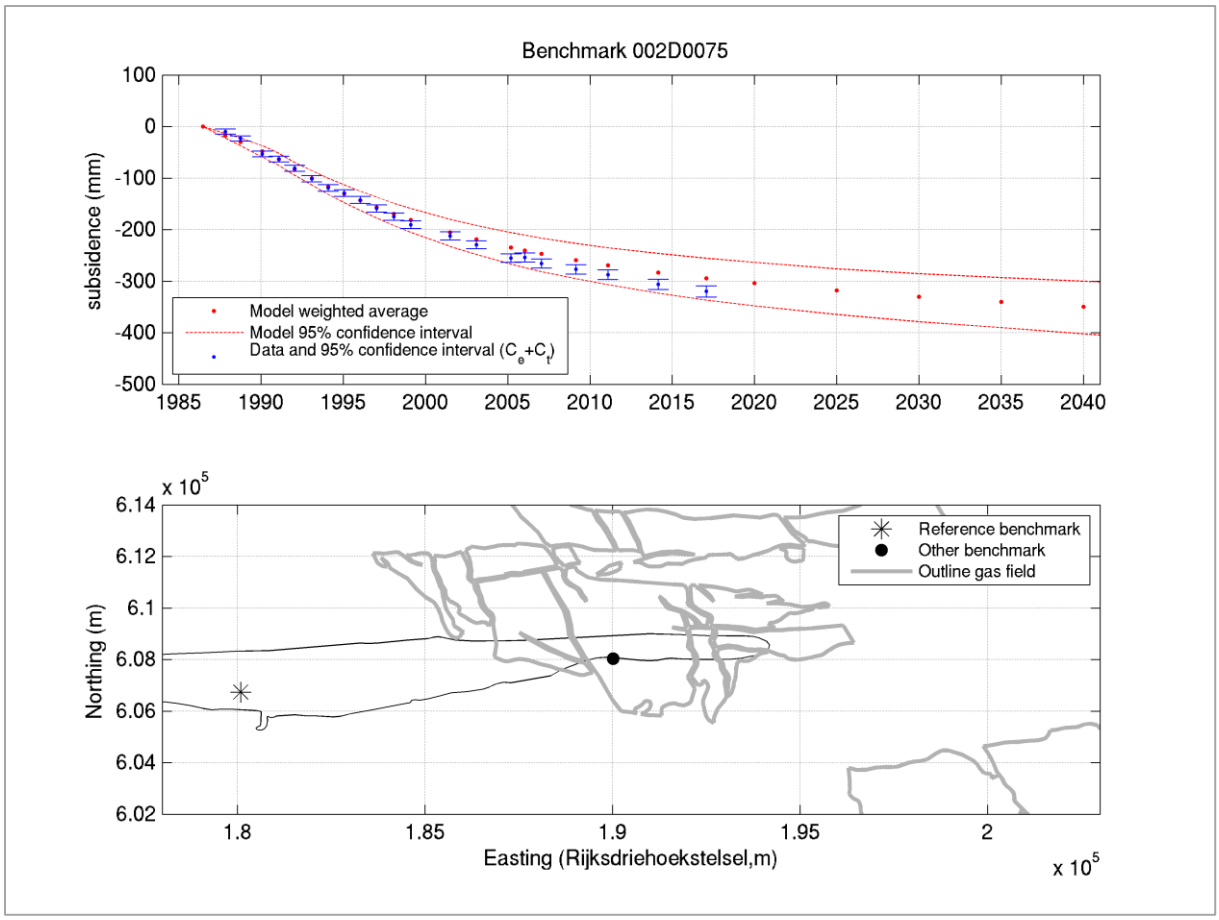

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

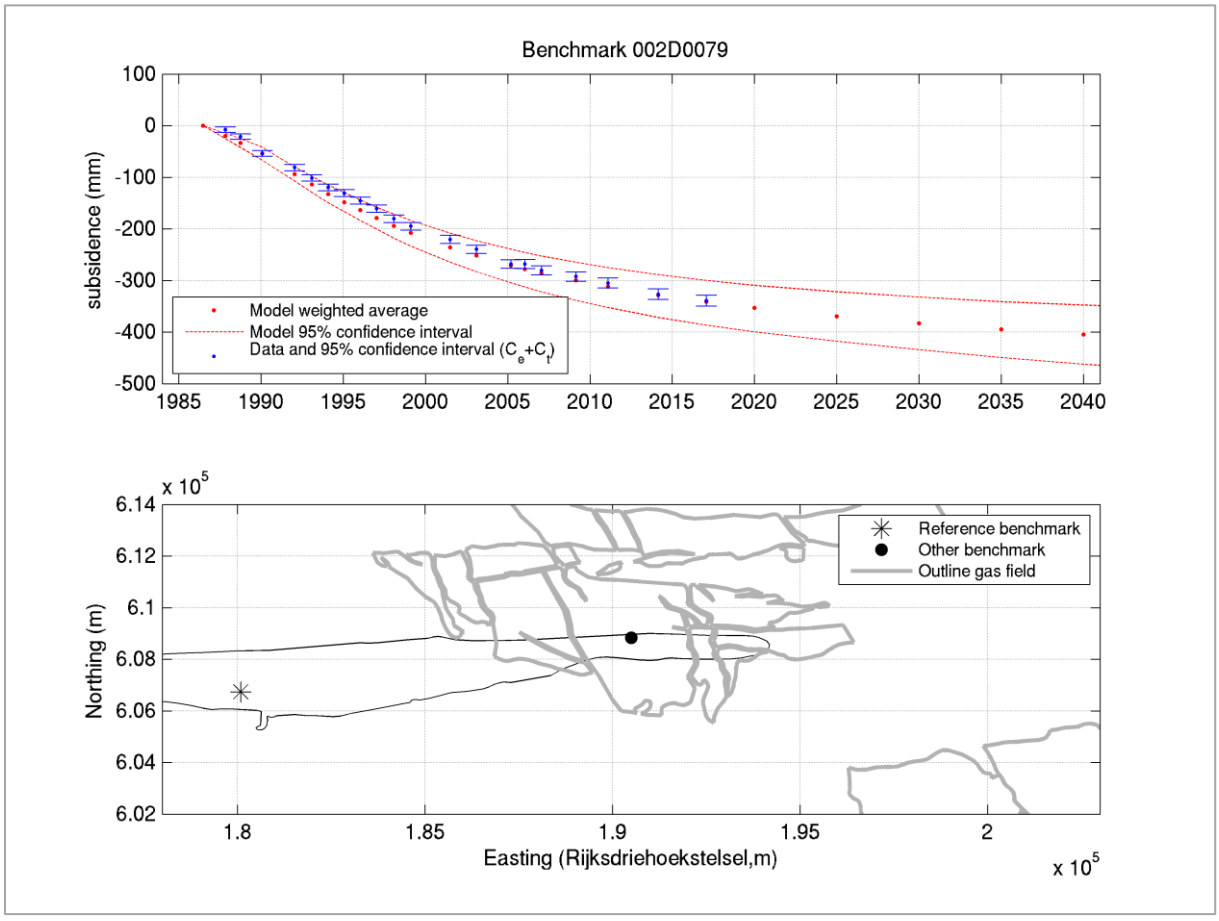

Benchmark 002D0073
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

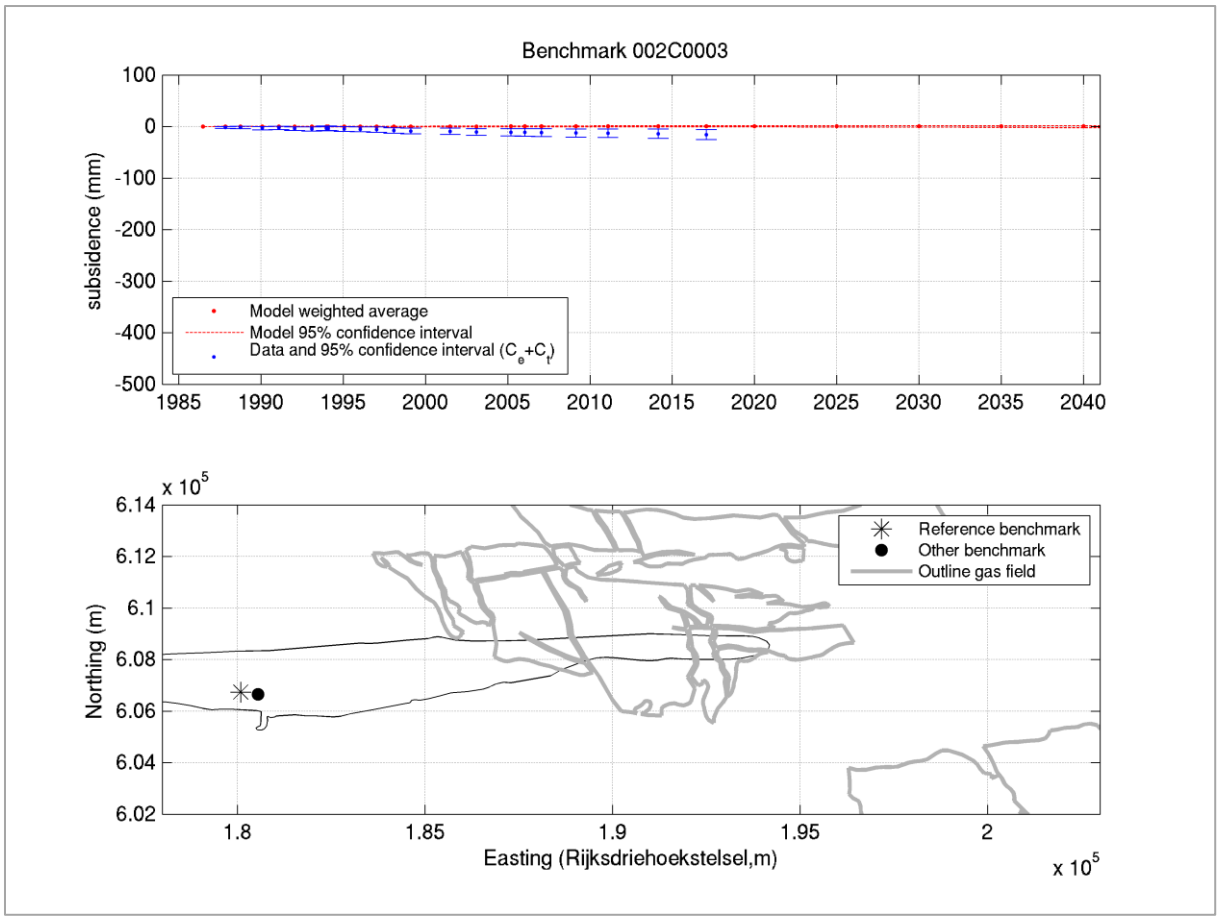

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

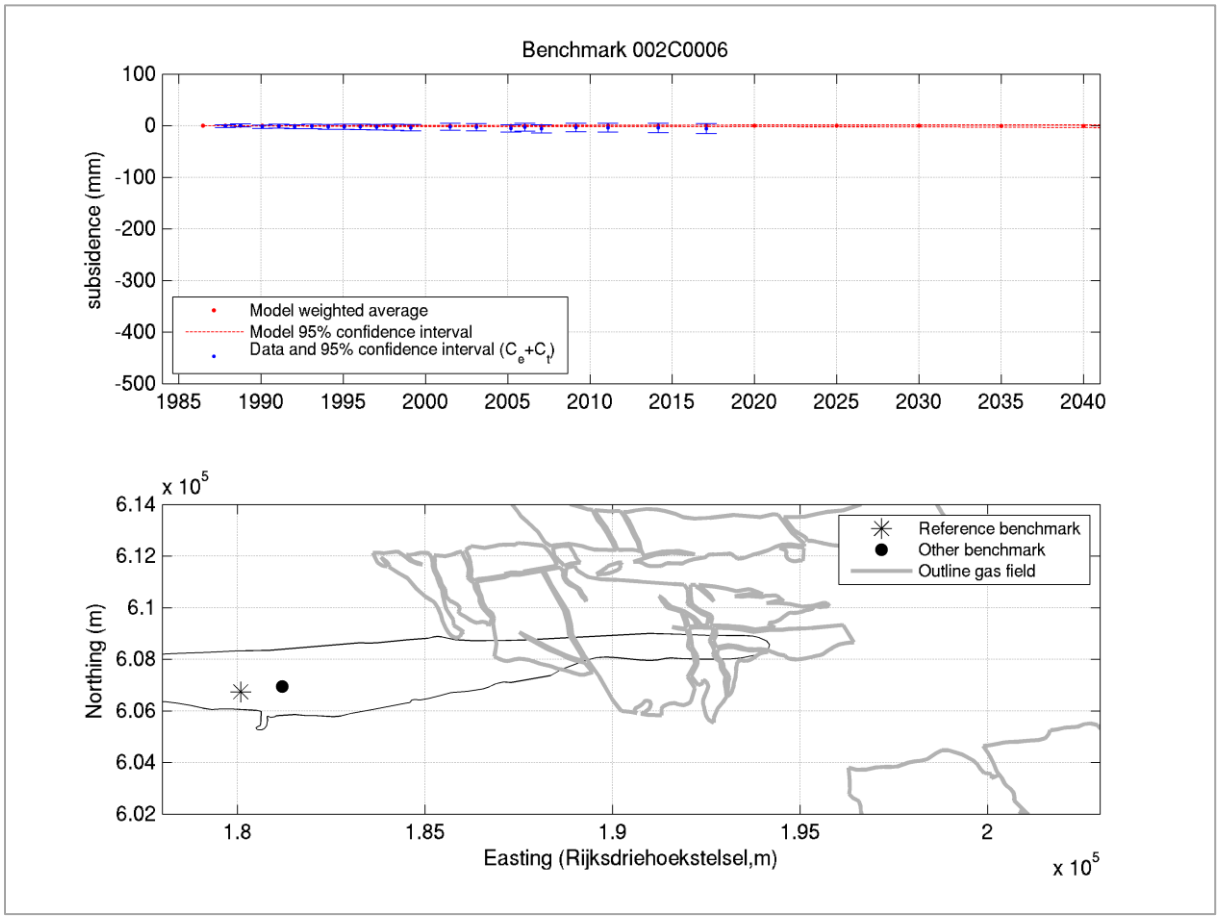

Benchmark 002D0074
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

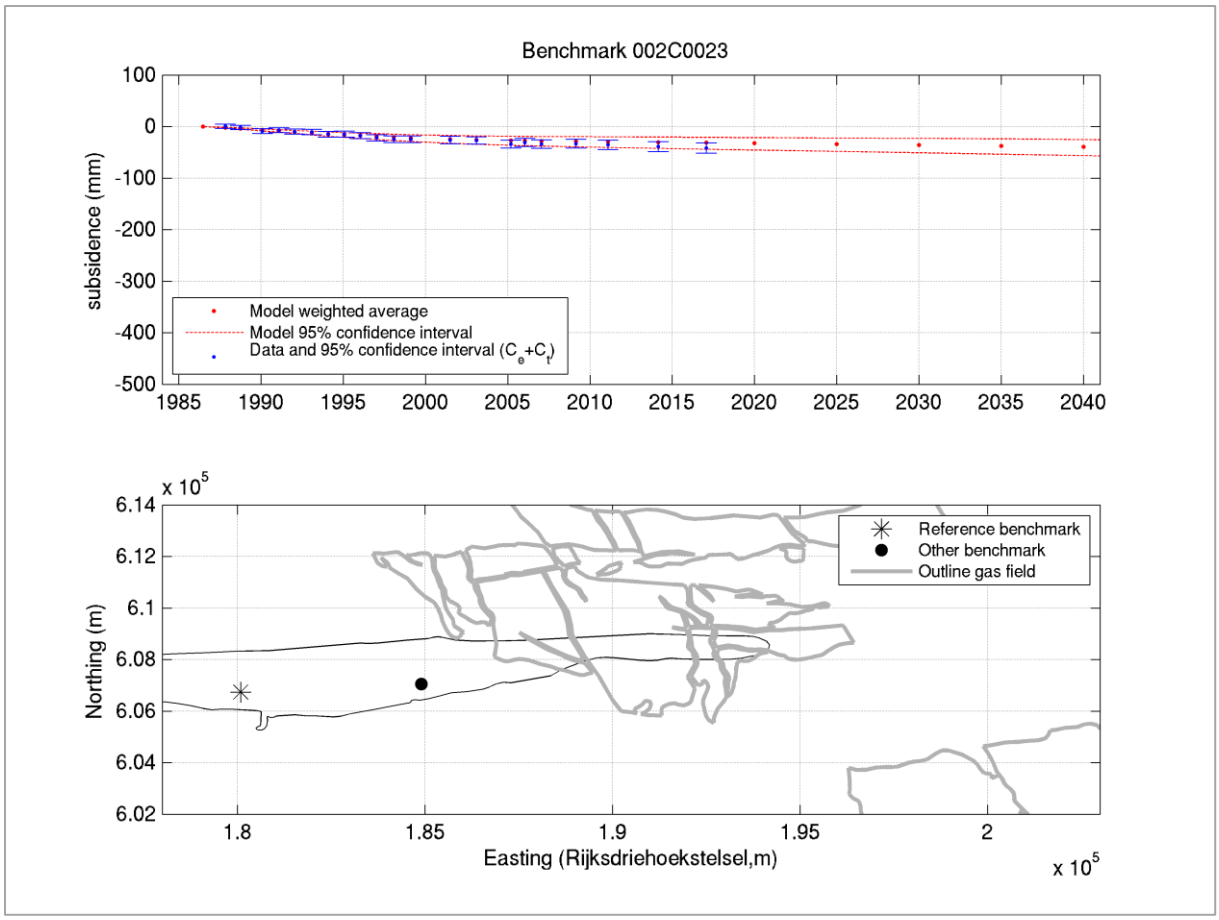

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

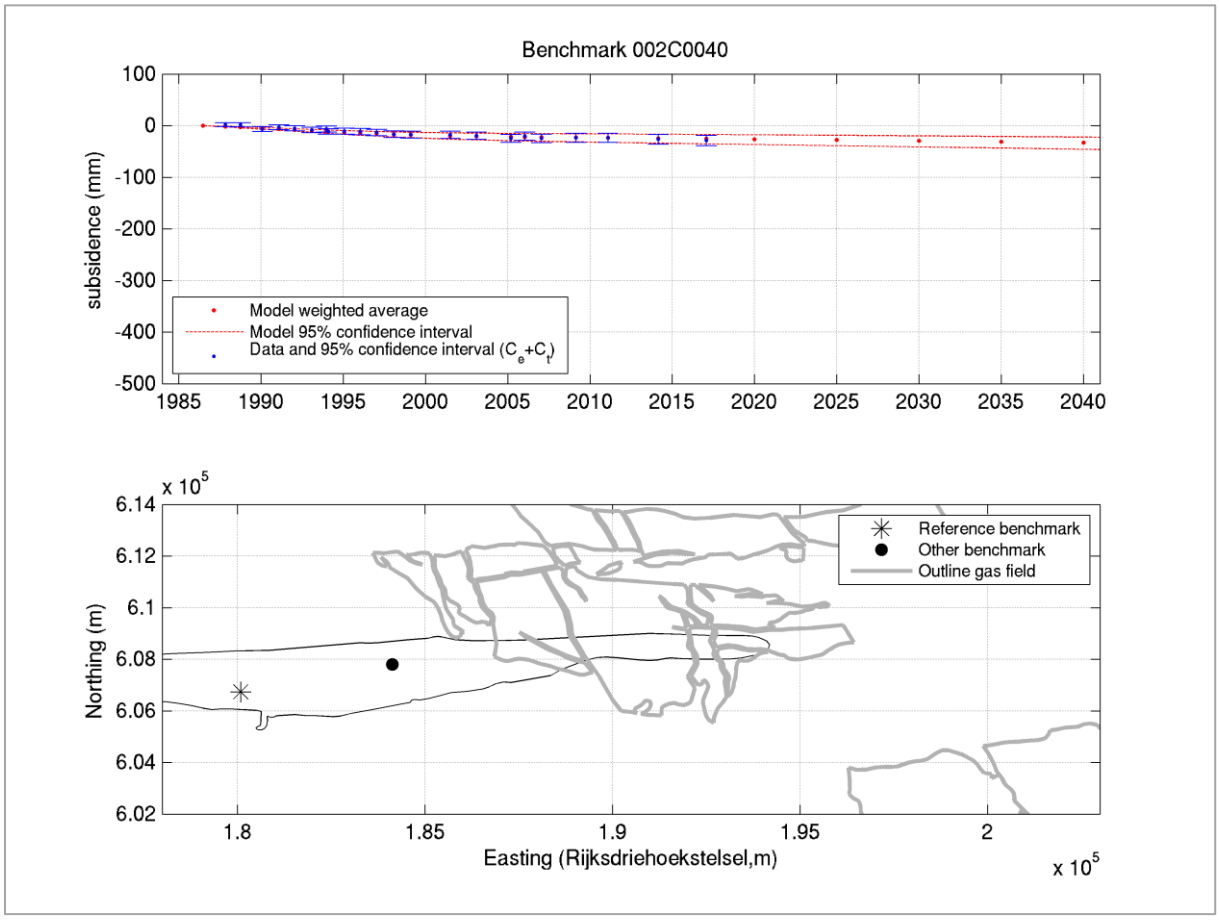

Benchmark 002D0075
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

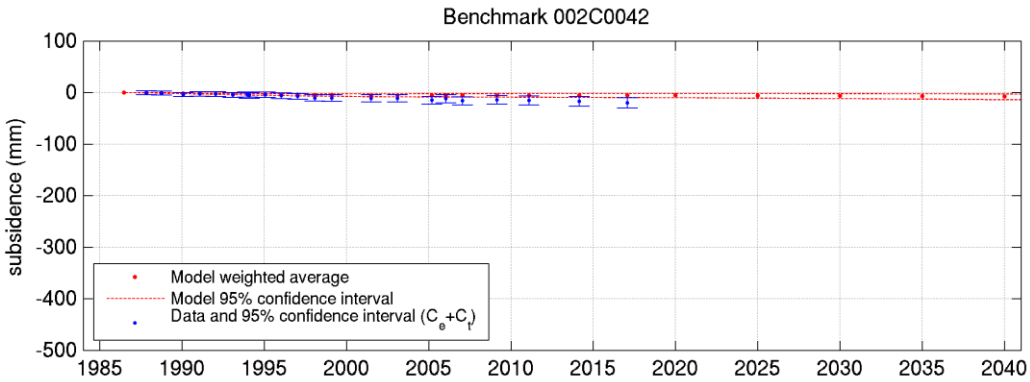

Benchmark 002D0079
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

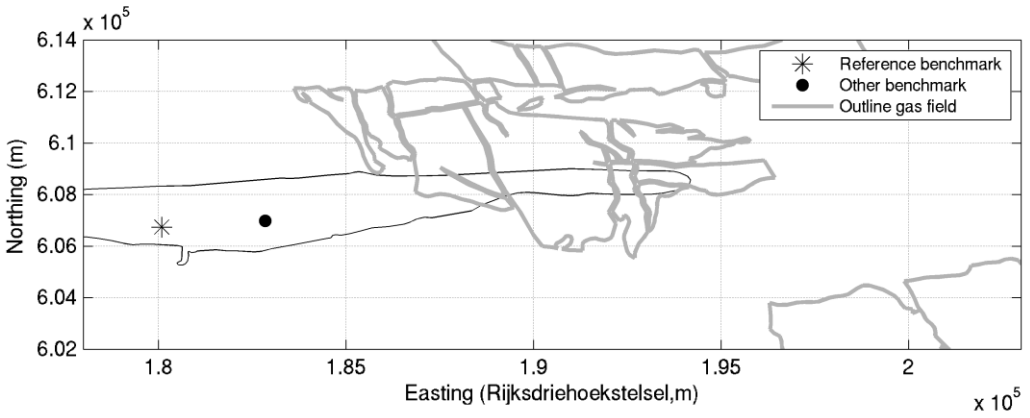

Benchmark 002C0003
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

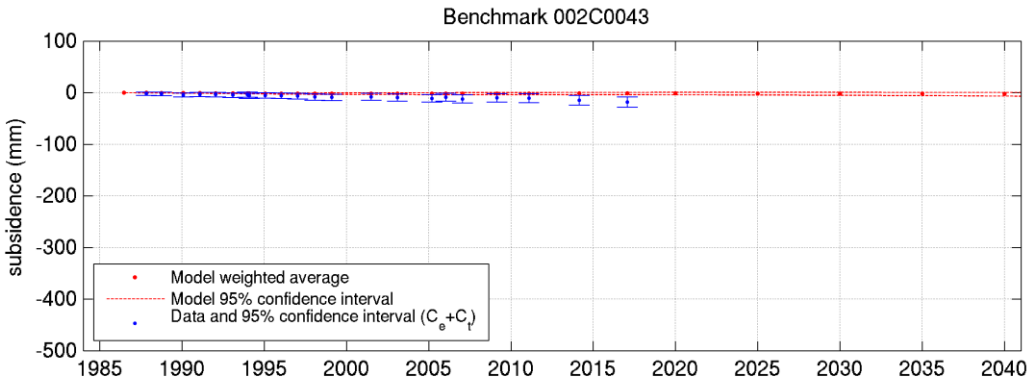

Benchmark 002C0006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

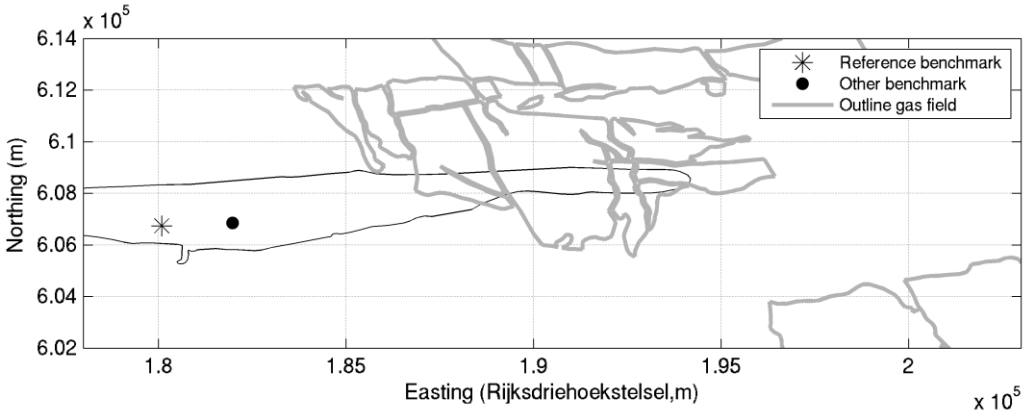

Benchmark 002C0023
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

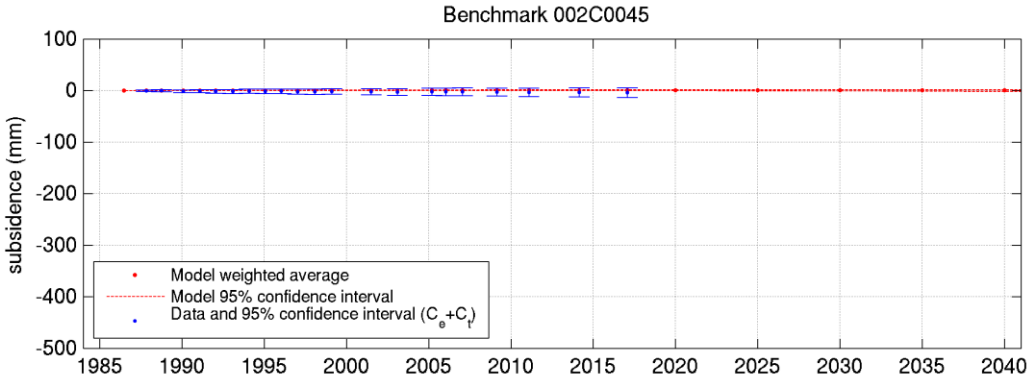

Benchmark 002C0040
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

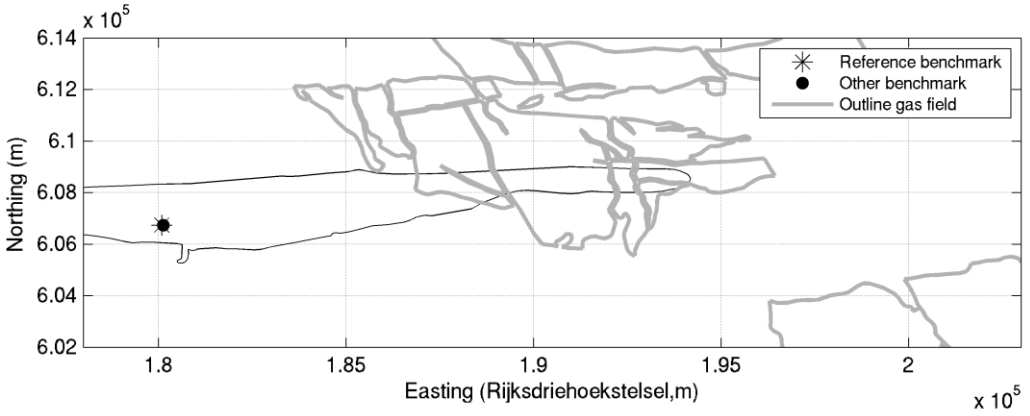

Benchmark 002C0042
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

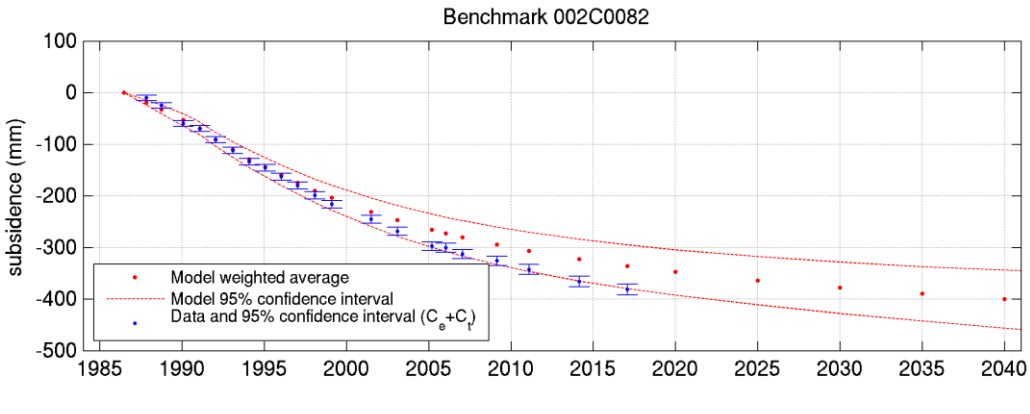

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

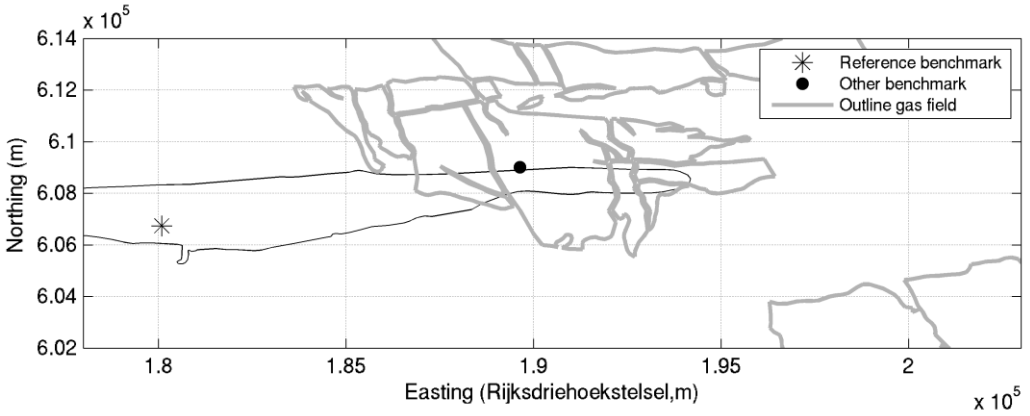

Benchmark 002C0043
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

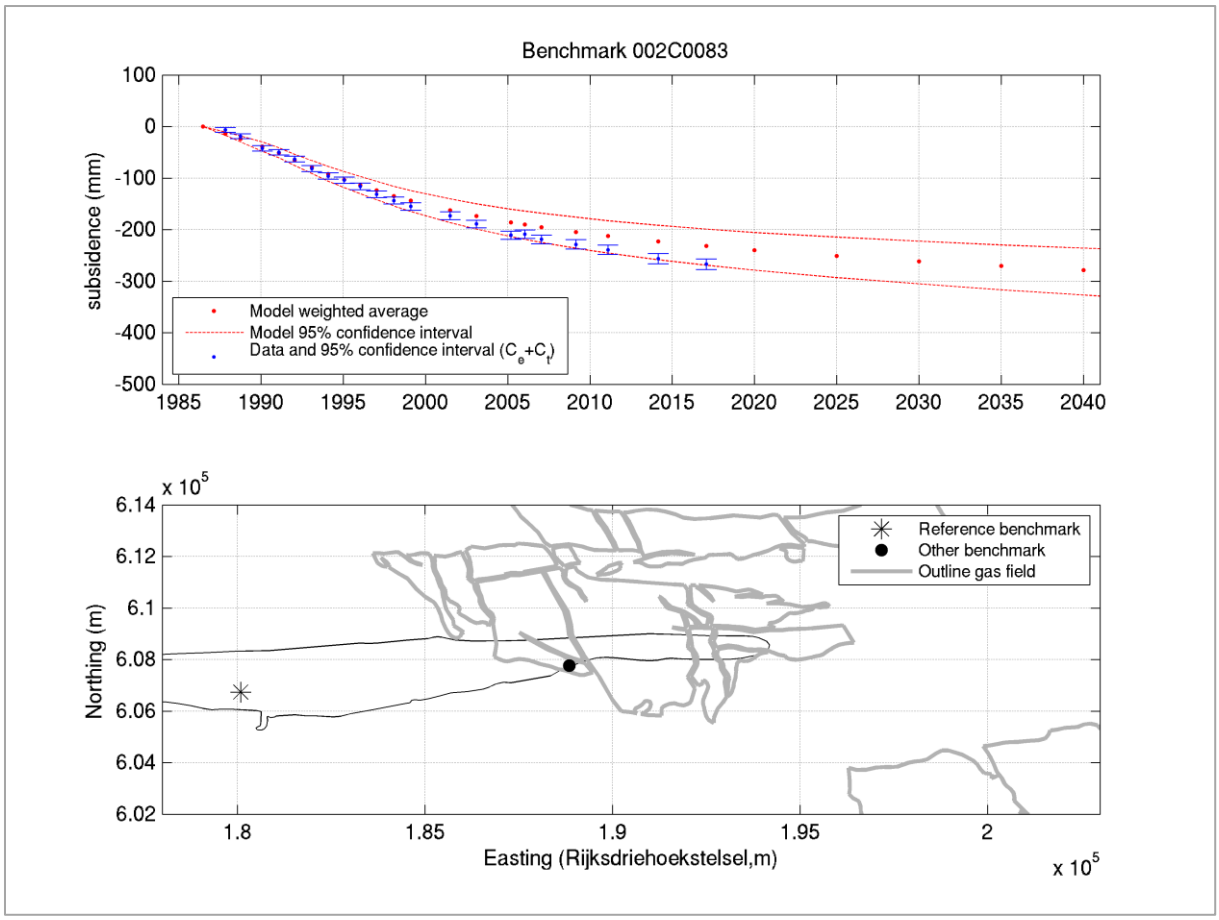

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

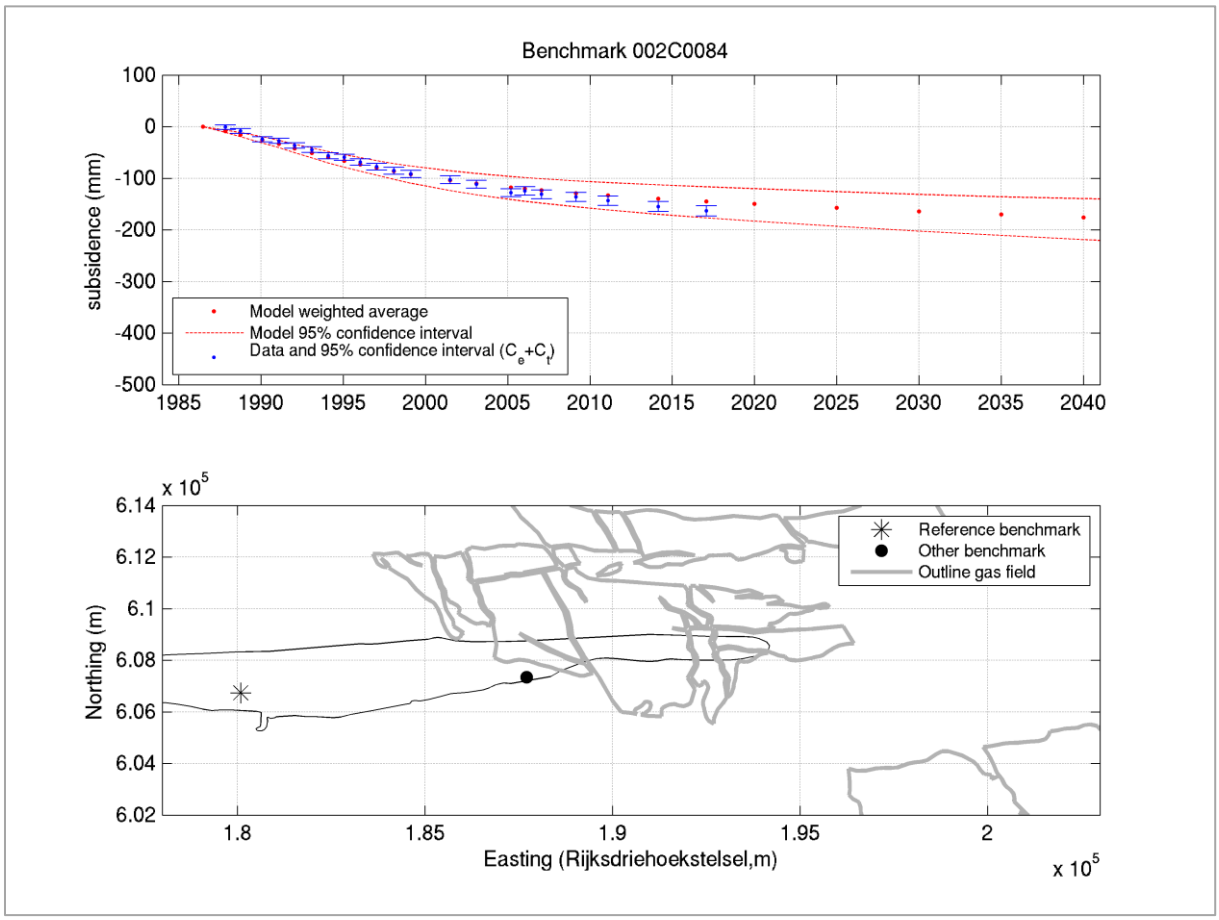

Benchmark 002C0045
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

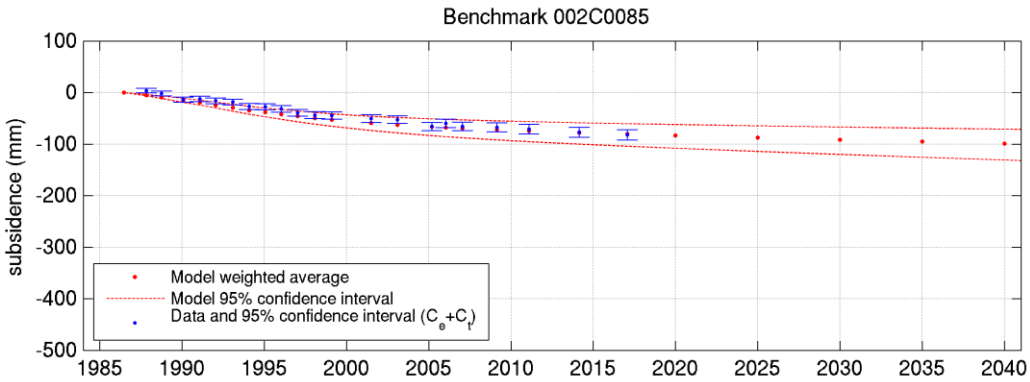

x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

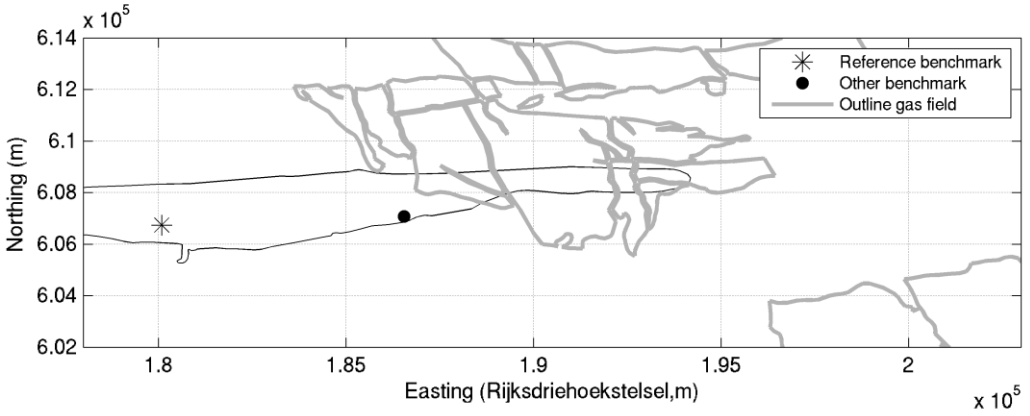

Benchmark 002C0082
subsidence (mm)
100
0
-100
-200
-300
-400
-500
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

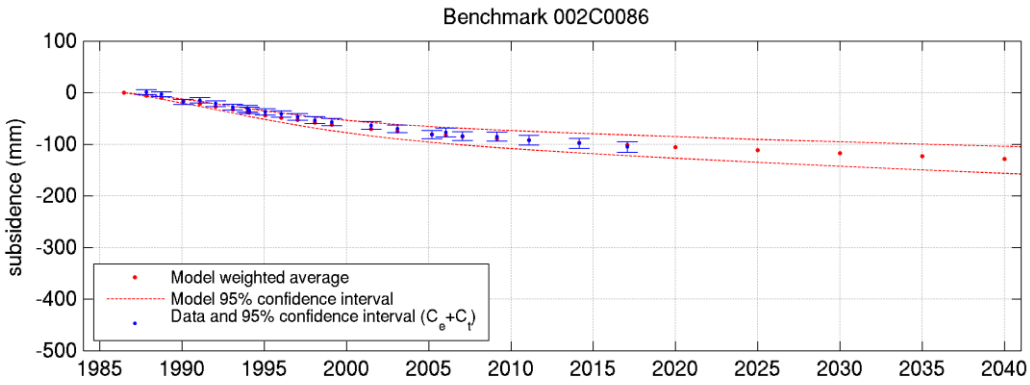

x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel,m)
x 10^5
Reference benchmark
Other benchmark
Outline gas field

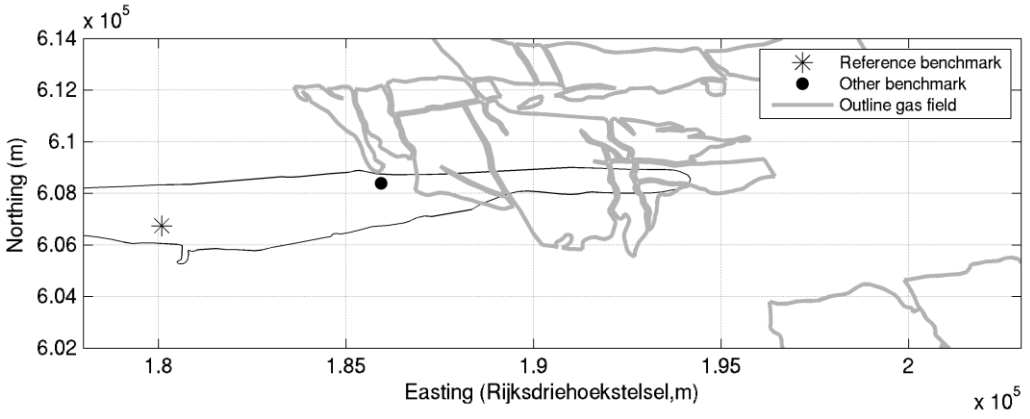

Benchmark 002C0083
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

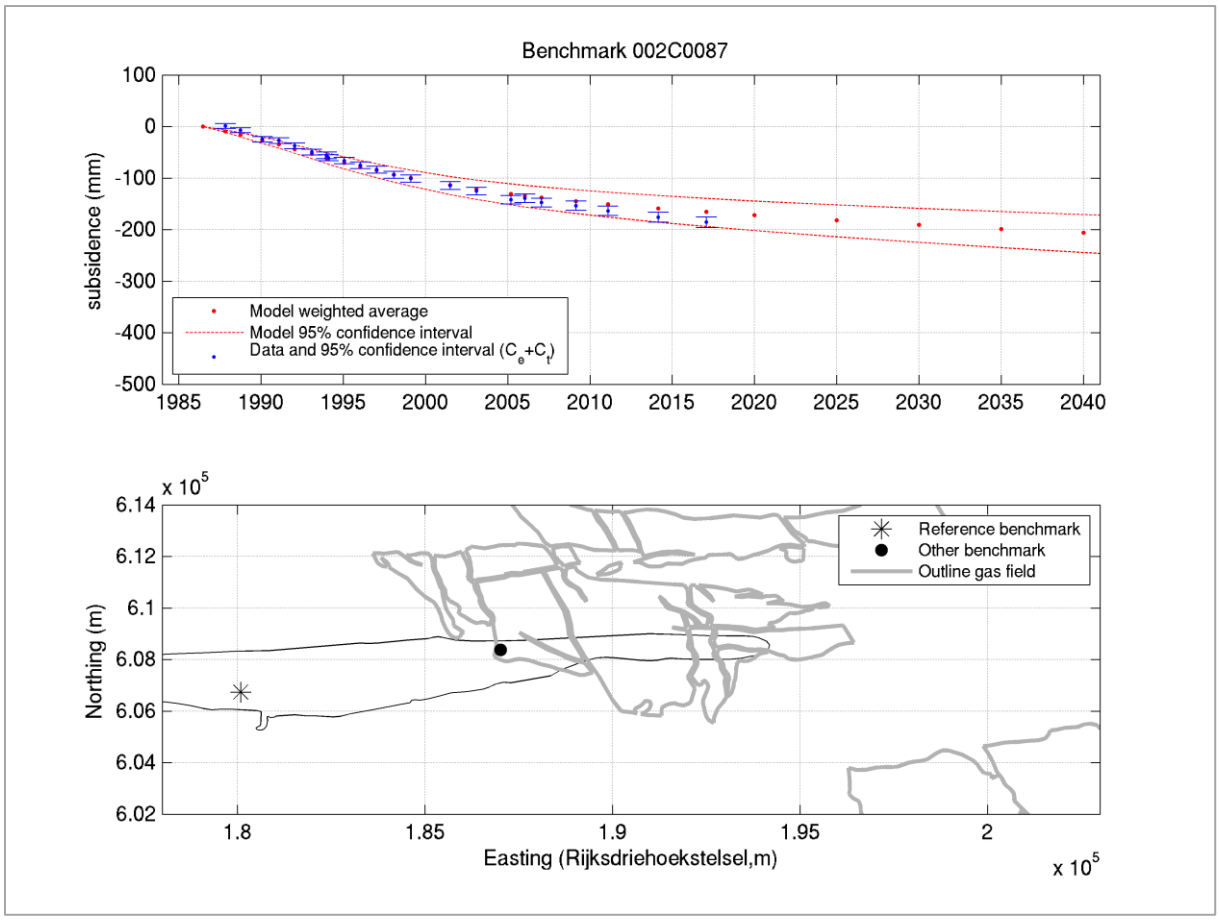

Benchmark 002C0084
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

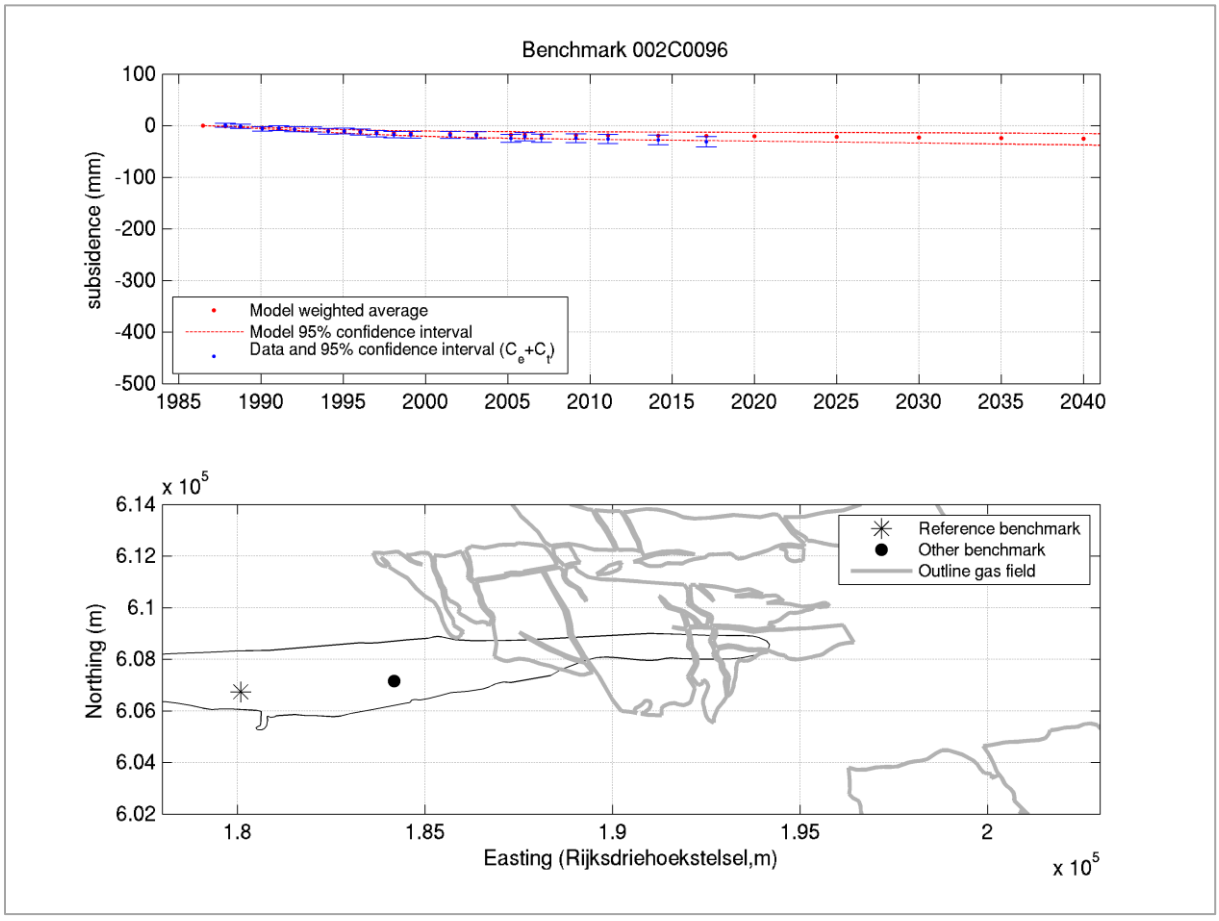

Benchmark 002C0085
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

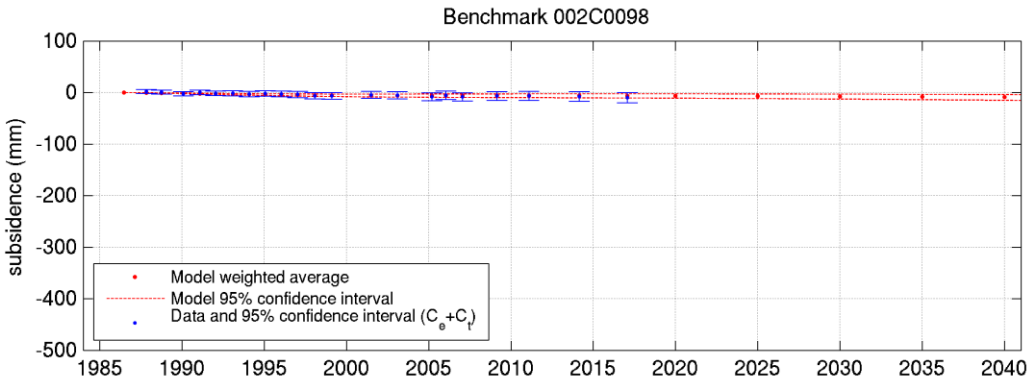

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

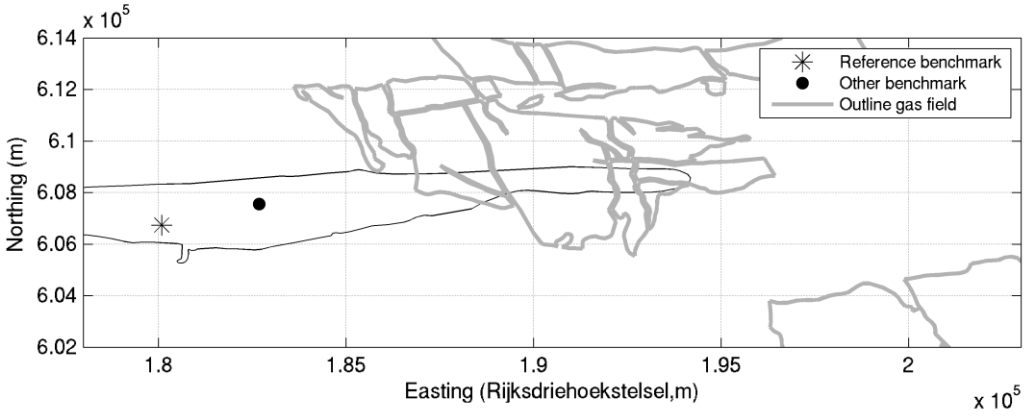

Benchmark 002C0086
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

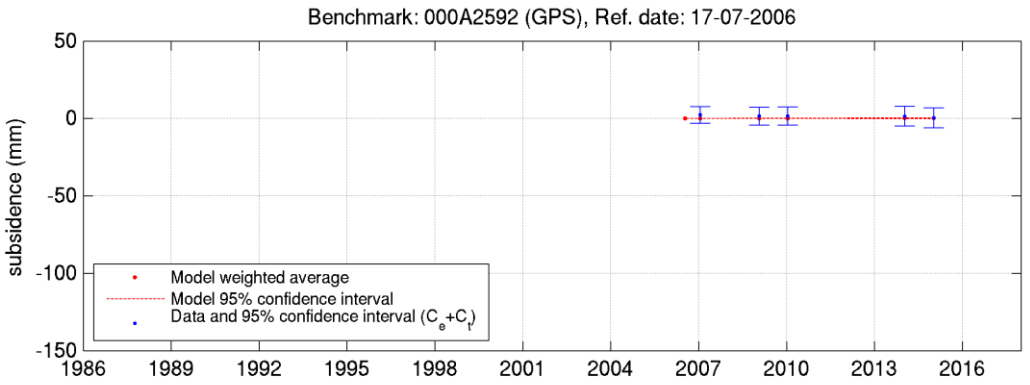

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

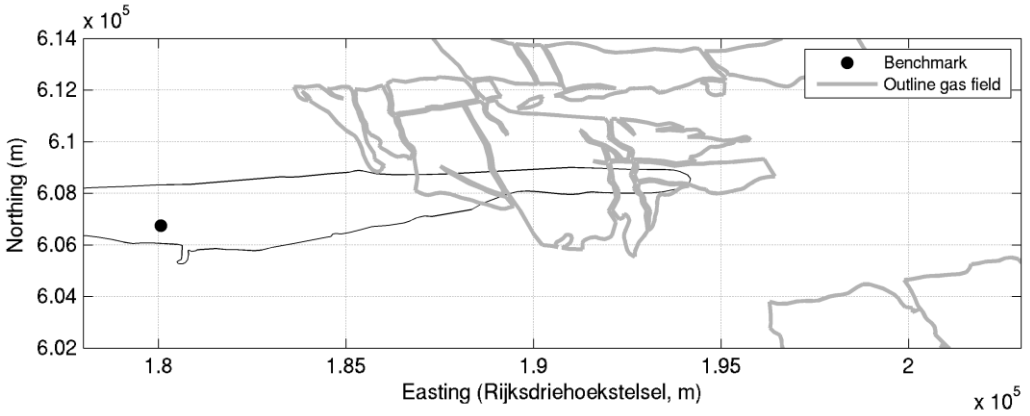

Benchmark 002C0087
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

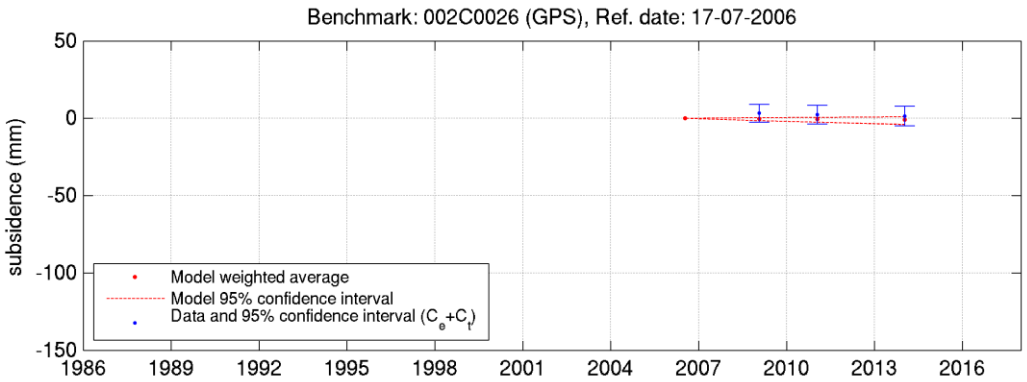

Benchmark 002C0096
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

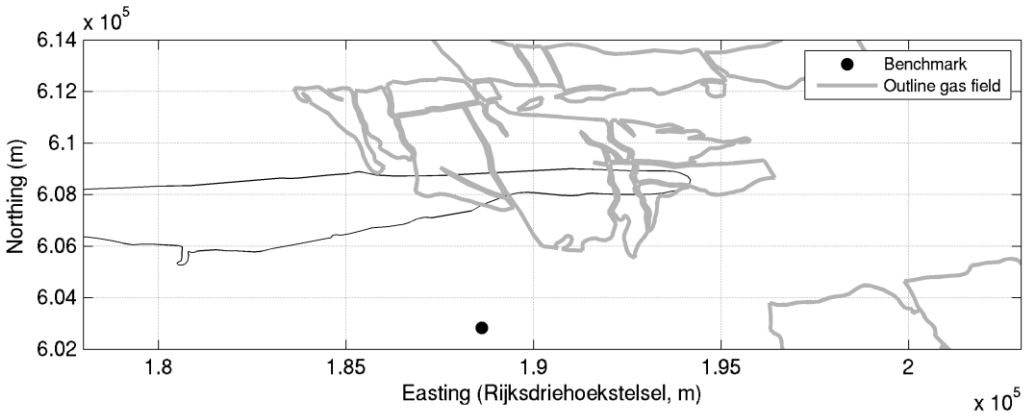

Benchmark 002C0098
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

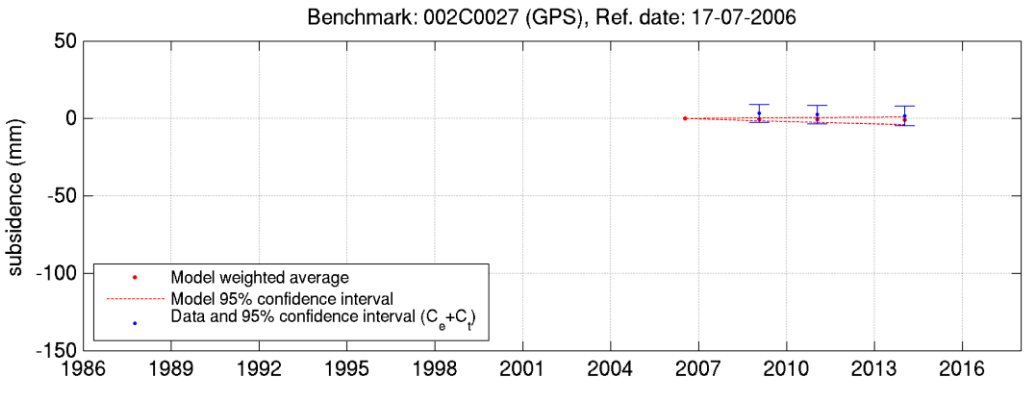

x 10^5
Northing (m)
Reference benchmark
Other benchmark
Outline gas field
Easting (Rijksdriehoekstelsel,m)
x 10^5

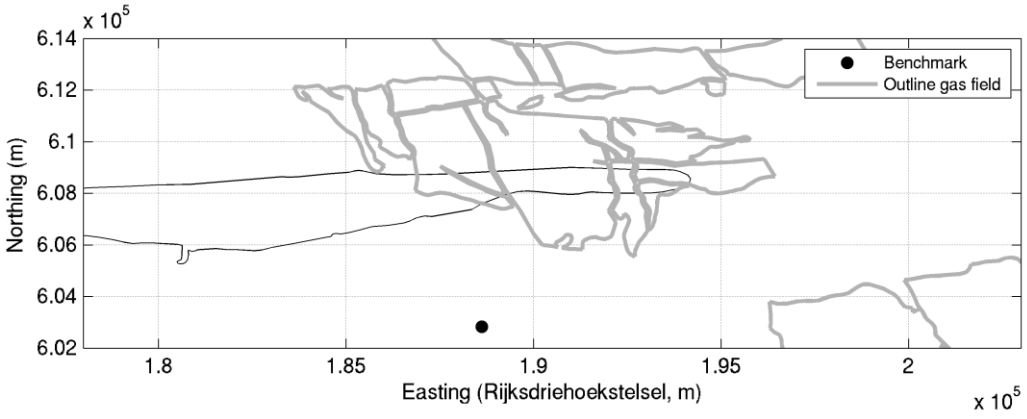

Benchmark: 000A2592 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

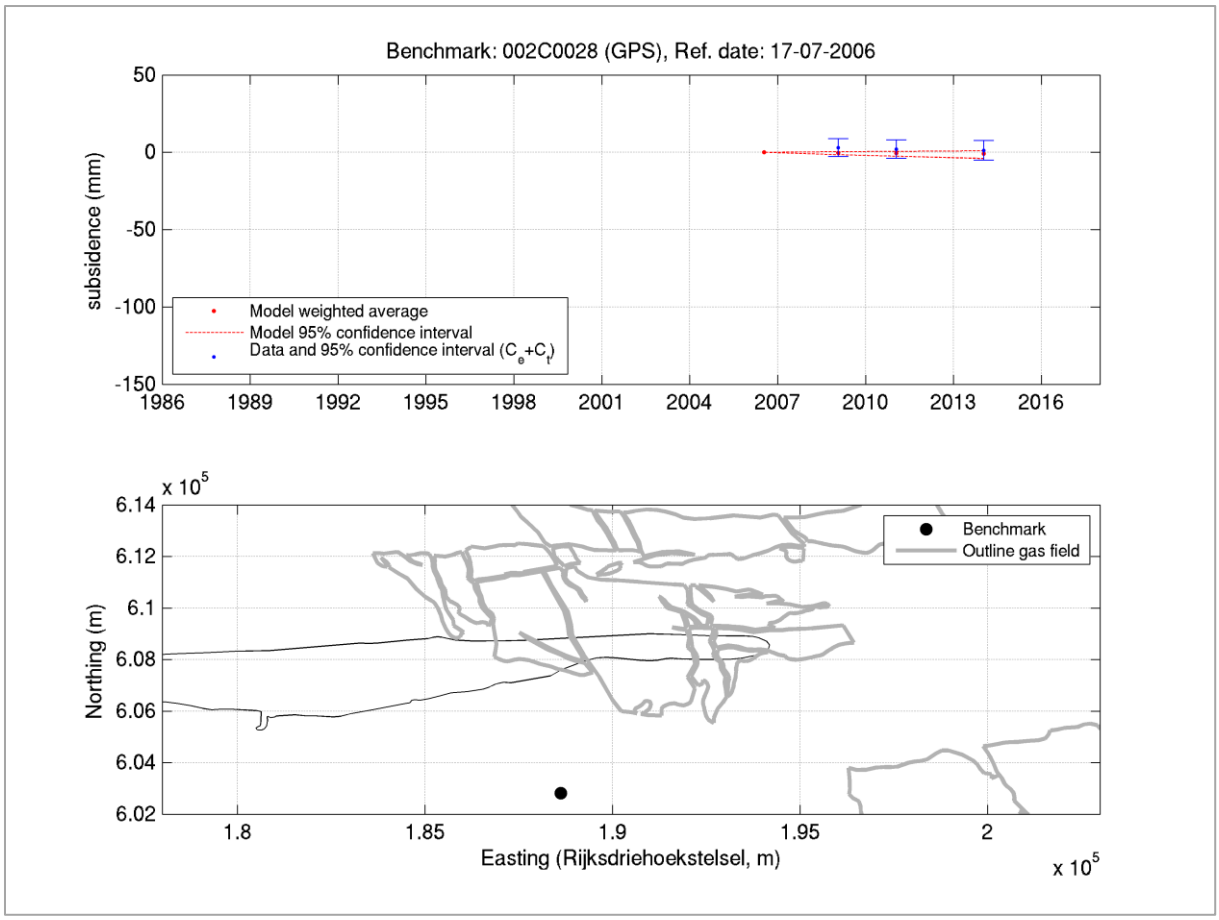

x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

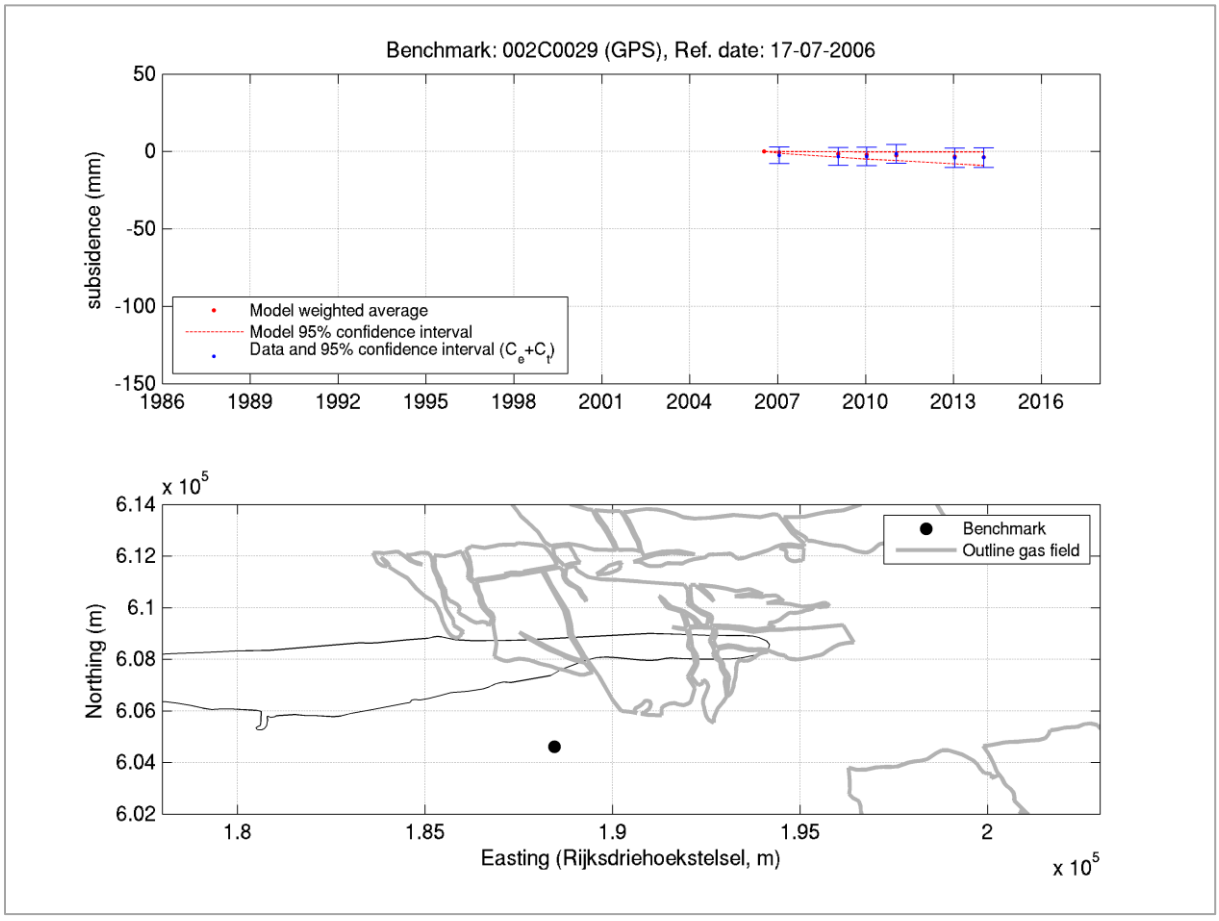

Benchmark: 002C0026 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

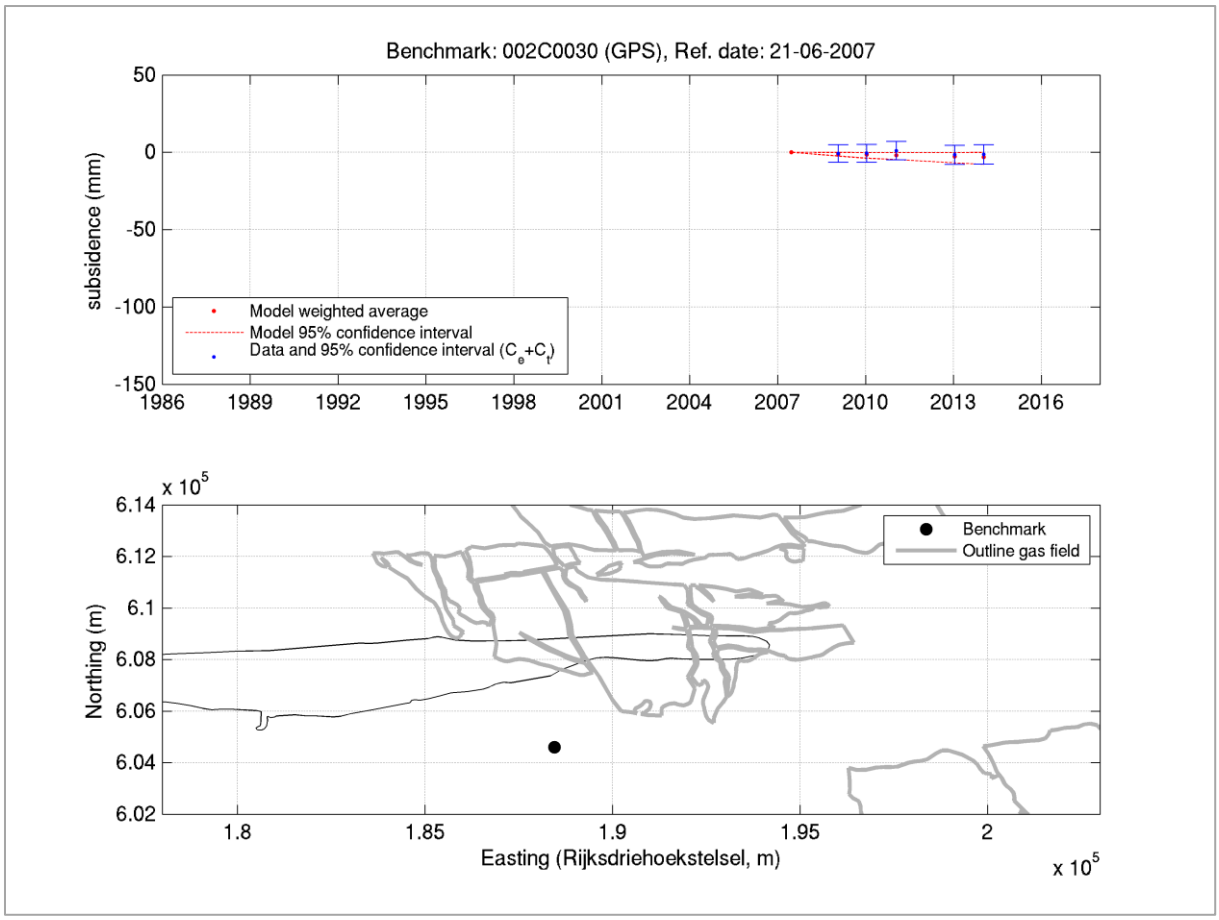

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

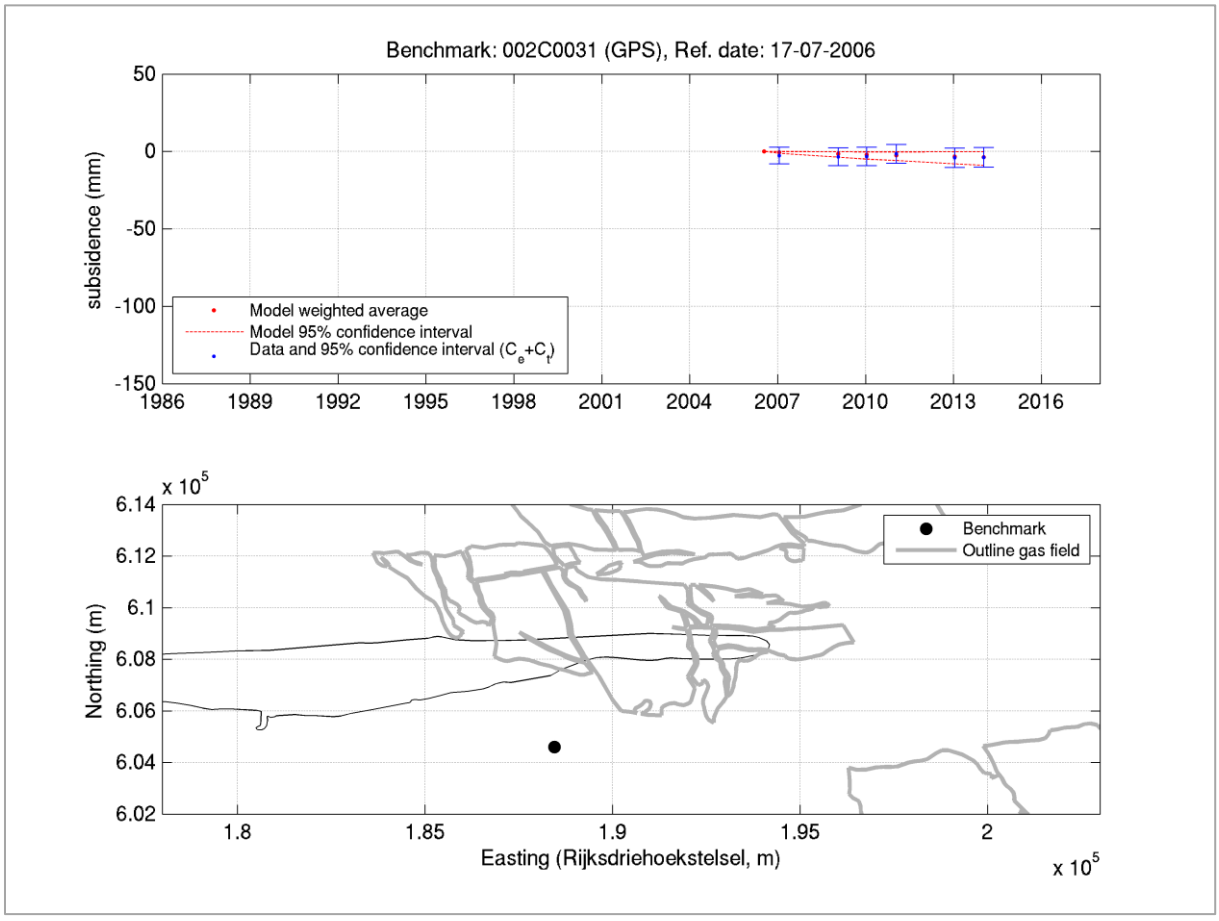

Benchmark: 002C0027 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

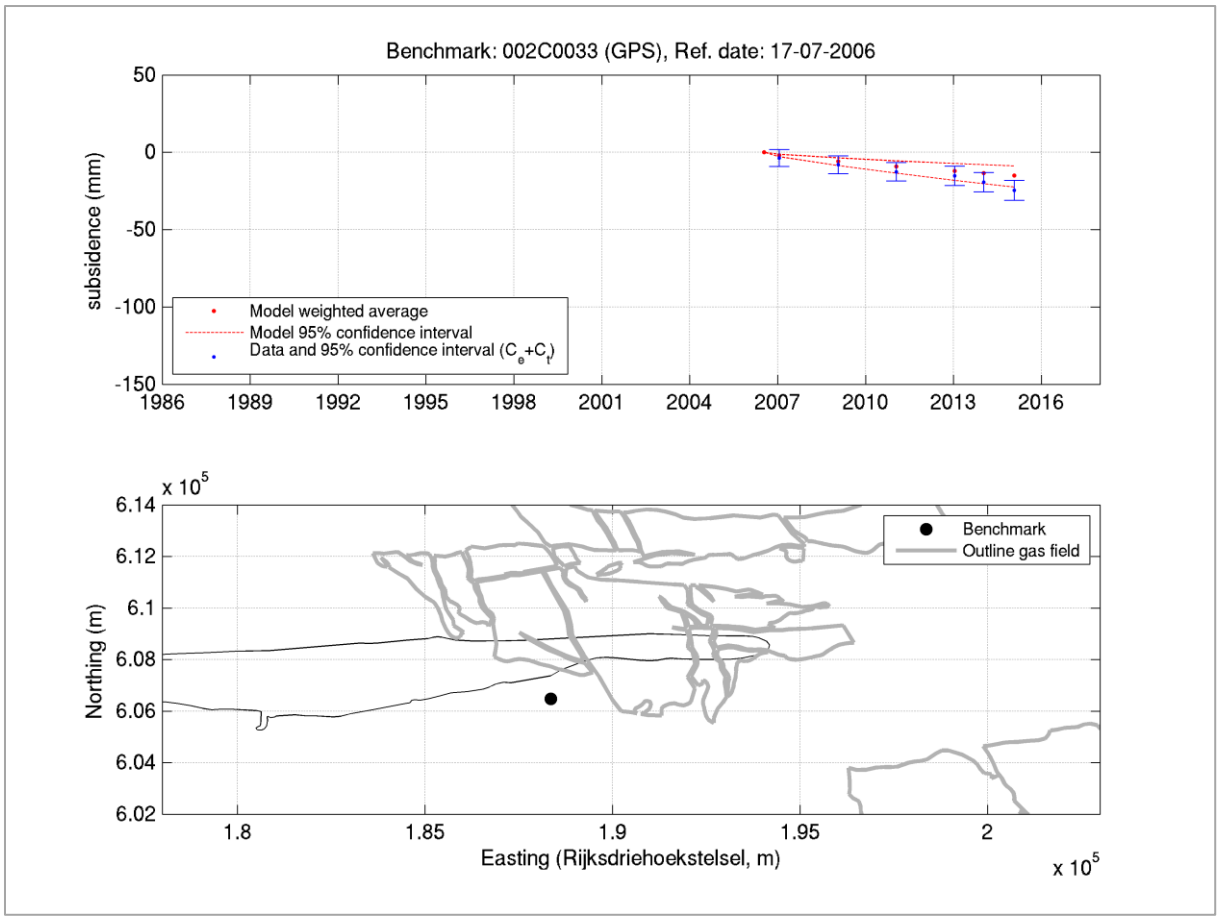

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

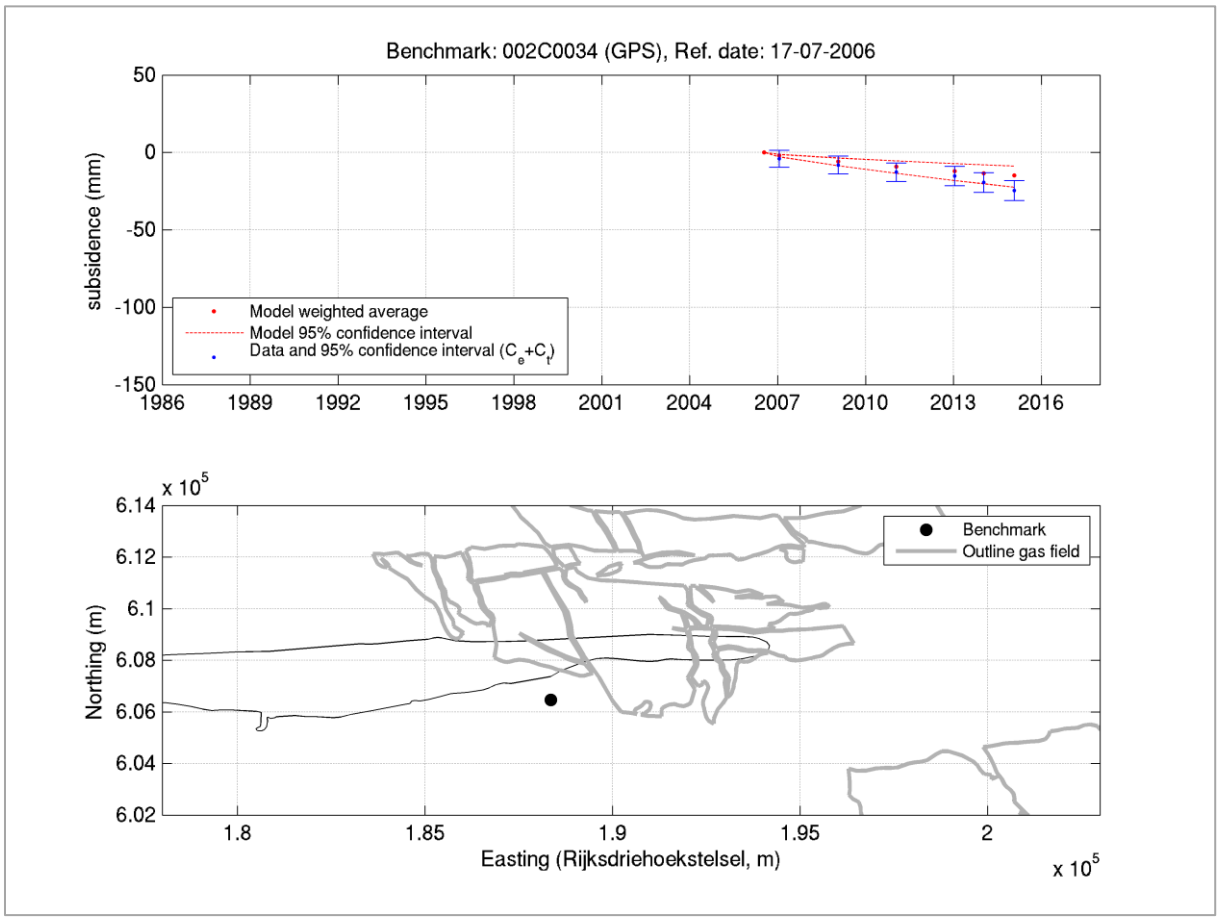

Benchmark: 002C0028 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

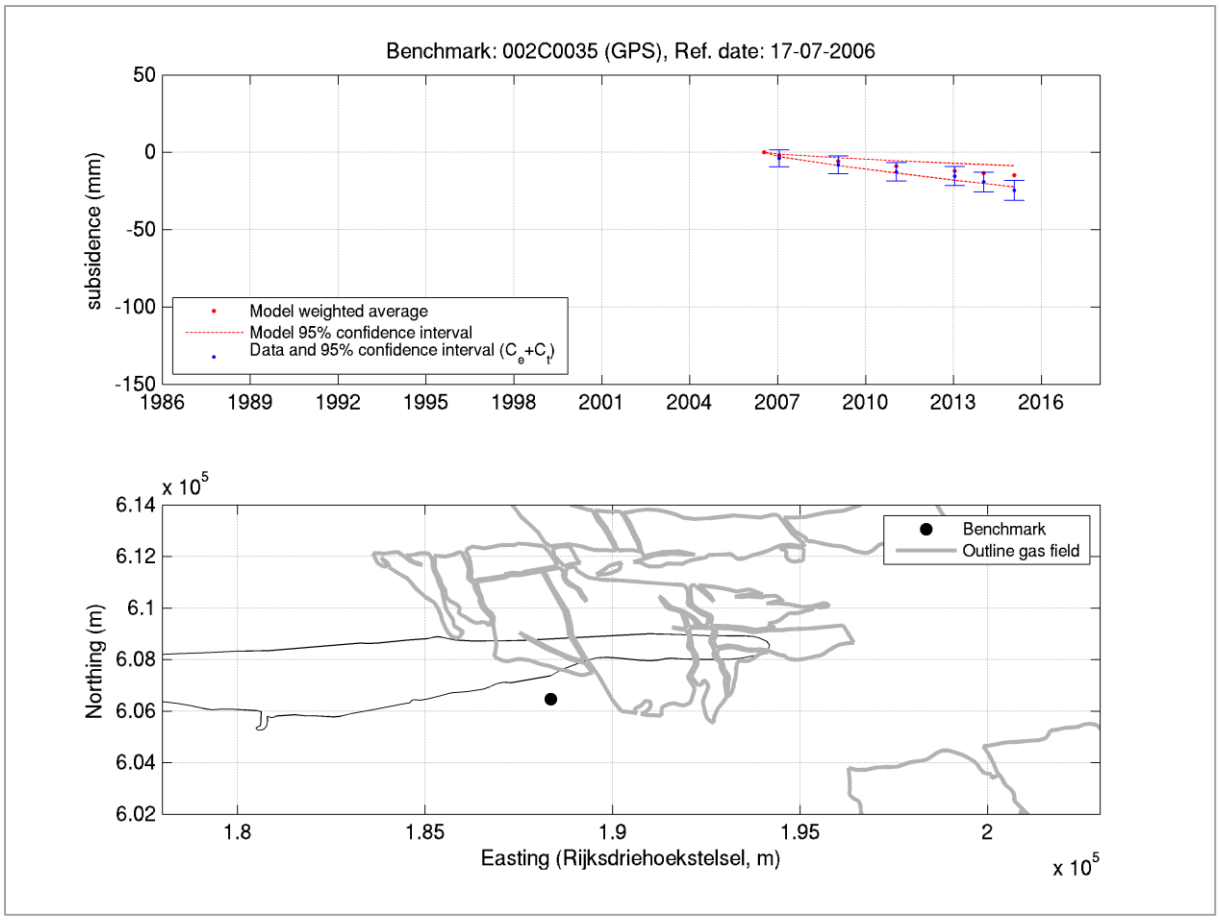

Benchmark: 002C0029 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

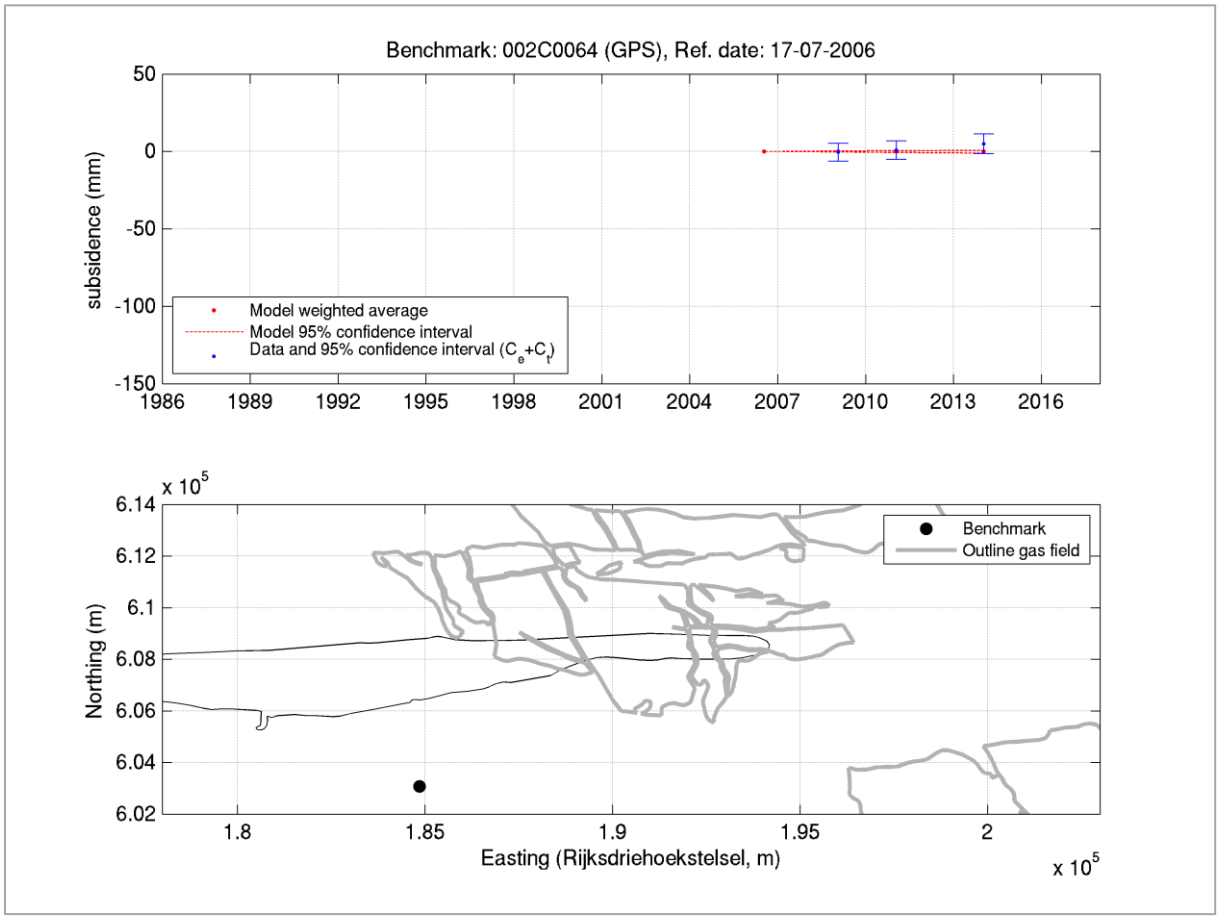

Benchmark: 002C0030 (GPS), Ref. date: 21-06-2007
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

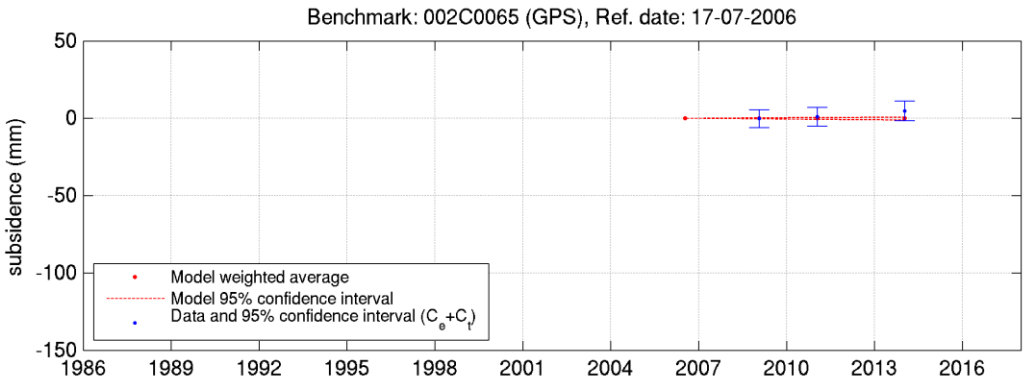

Benchmark: 002C0031 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

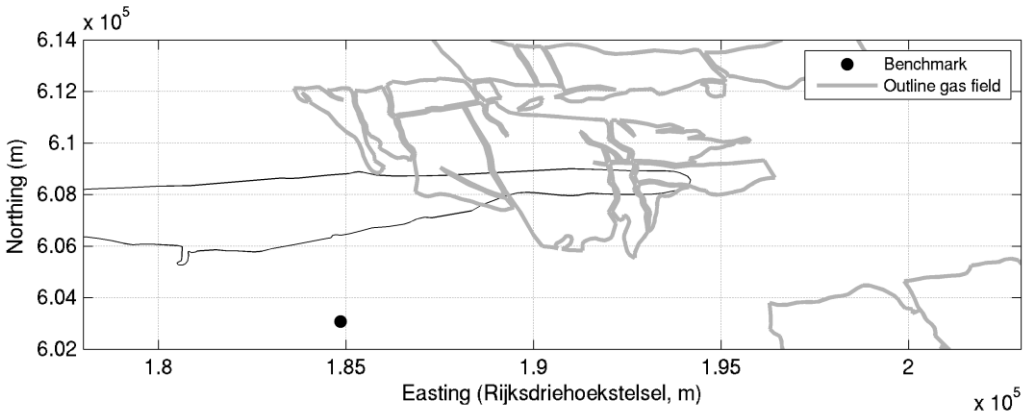

Benchmark: 002C0033 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

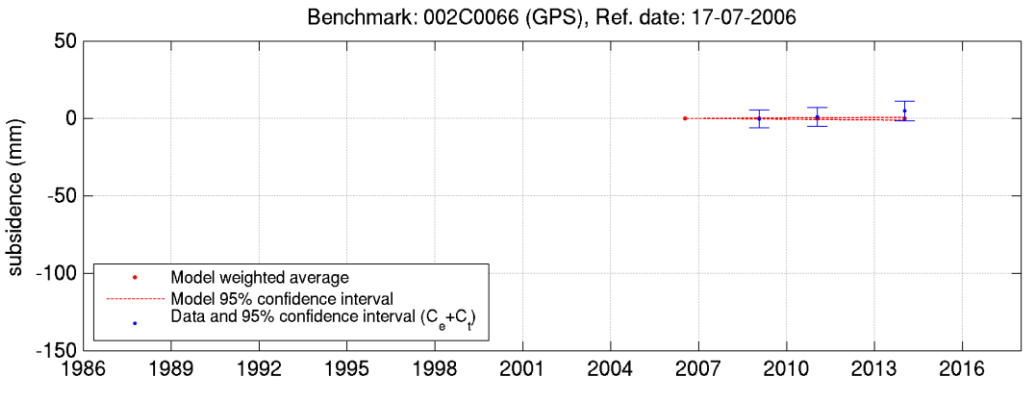

Benchmark: 002C0034 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

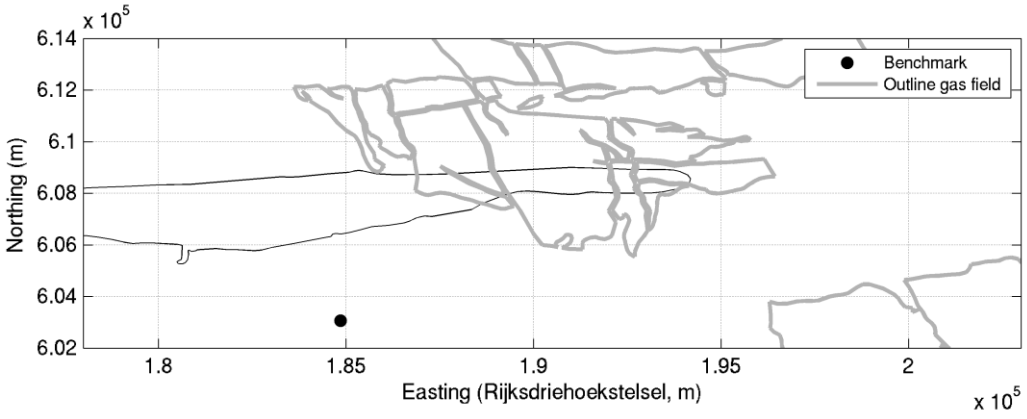

Benchmark: 002C0035 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

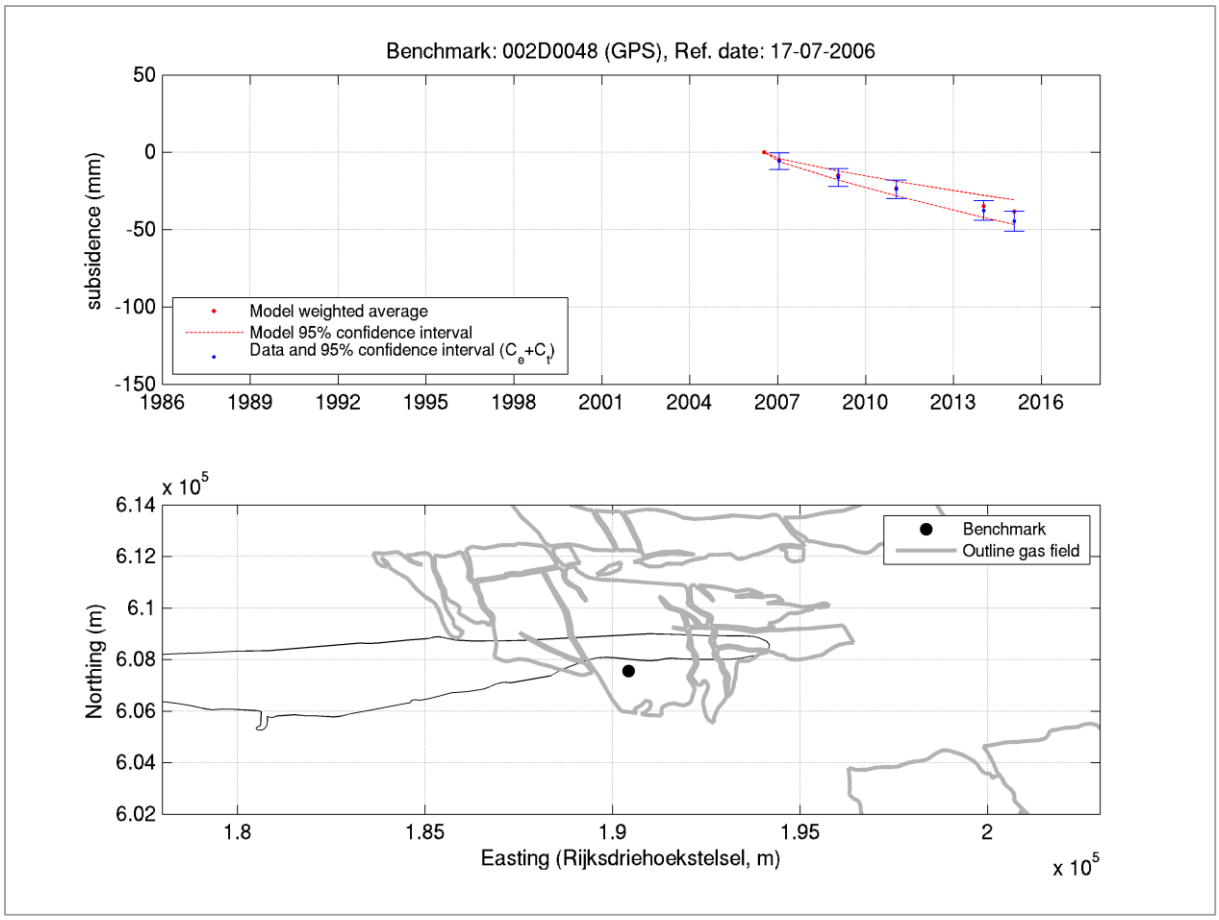

Benchmark: 002C0064 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

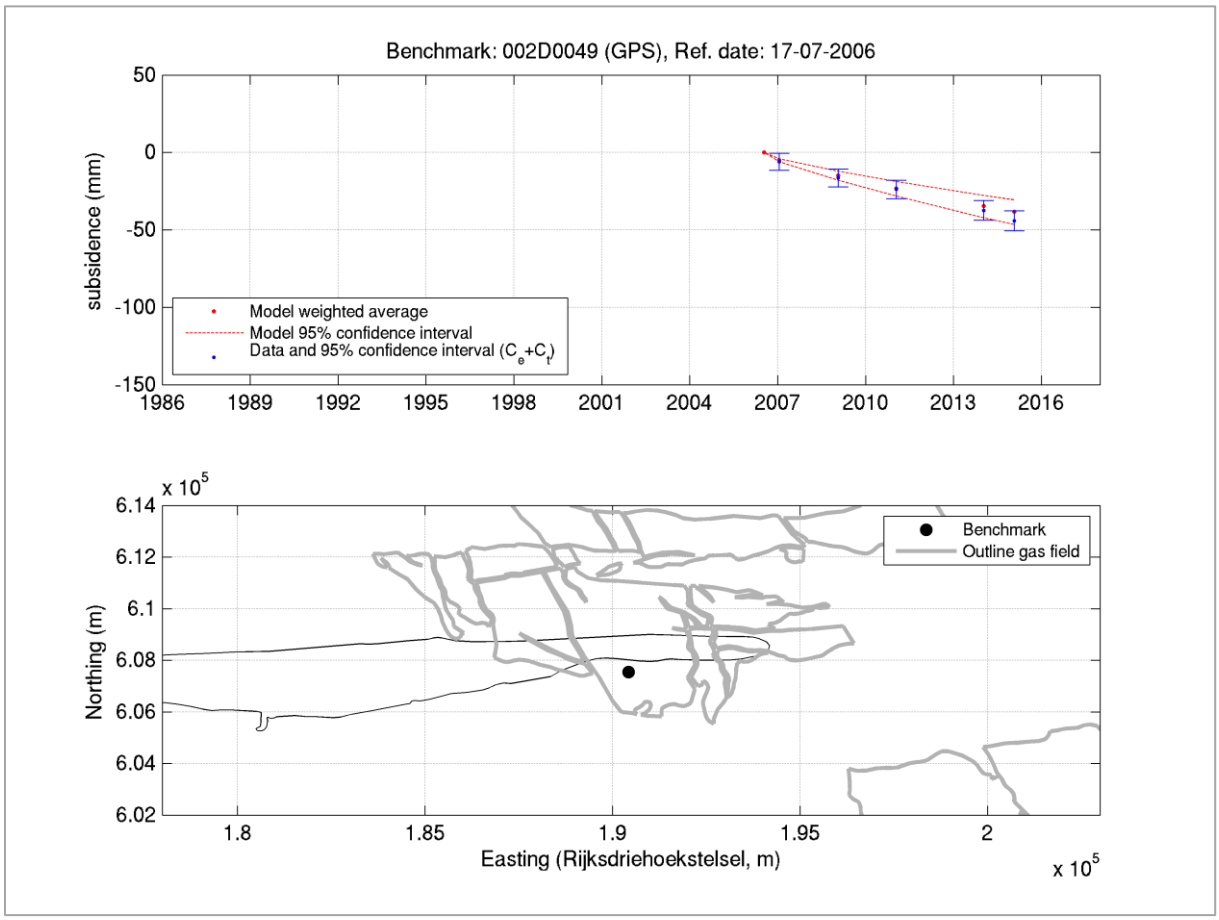

Benchmark: 002C0065 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

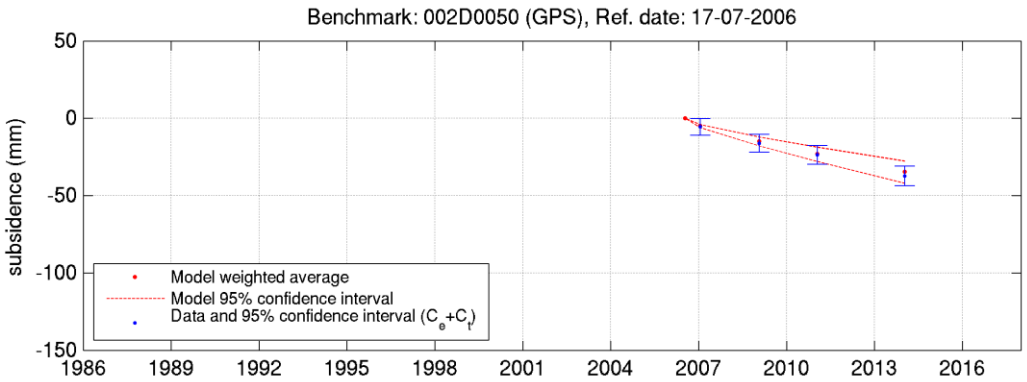

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

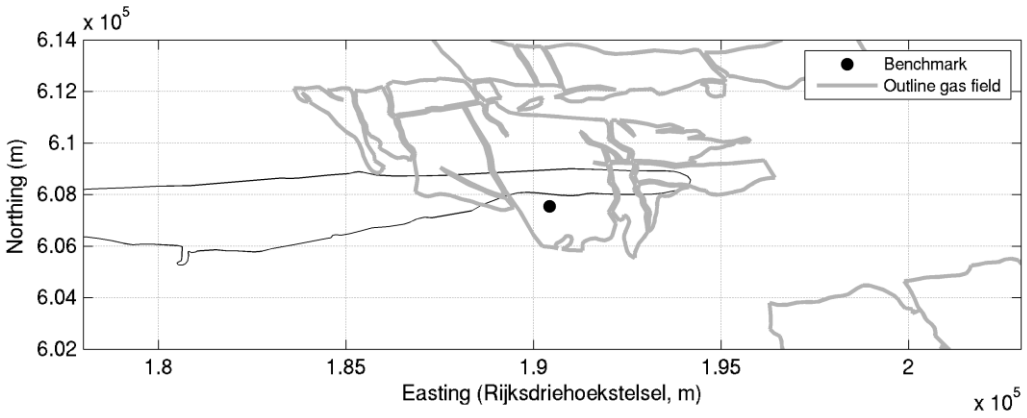

Benchmark: 002C0066 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)

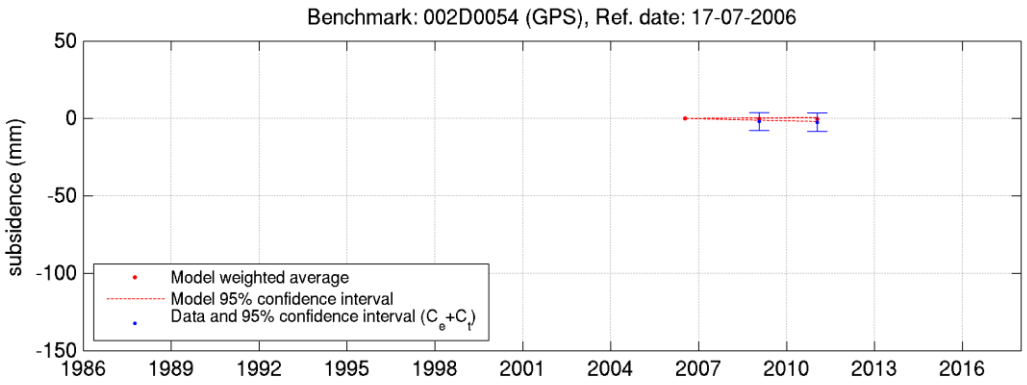

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

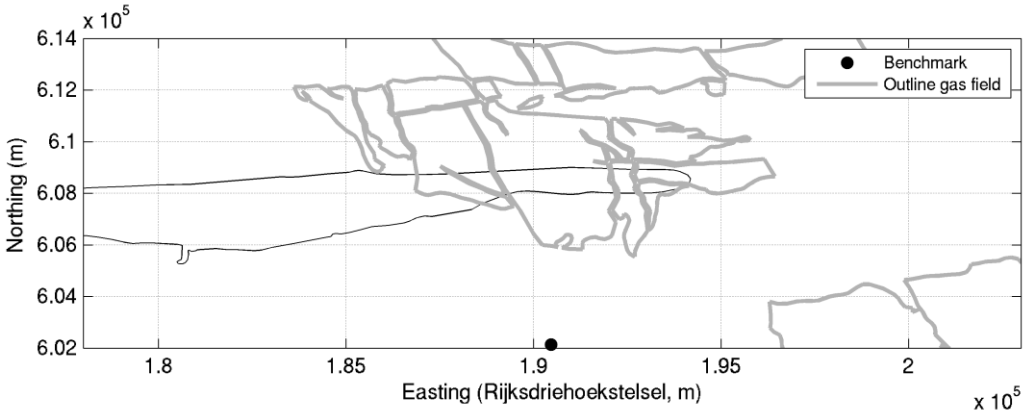

Benchmark: 002D0048 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

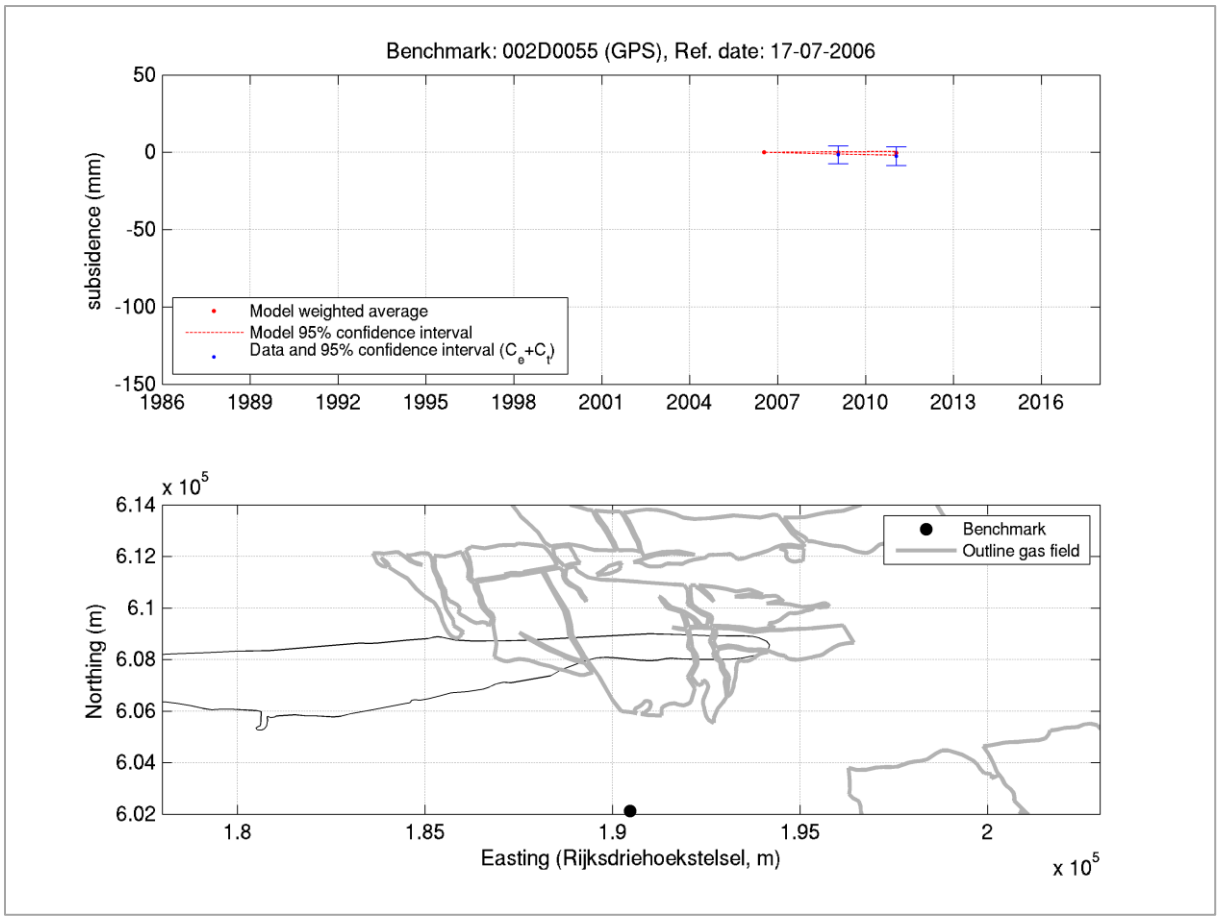

Benchmark: 002D0049 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

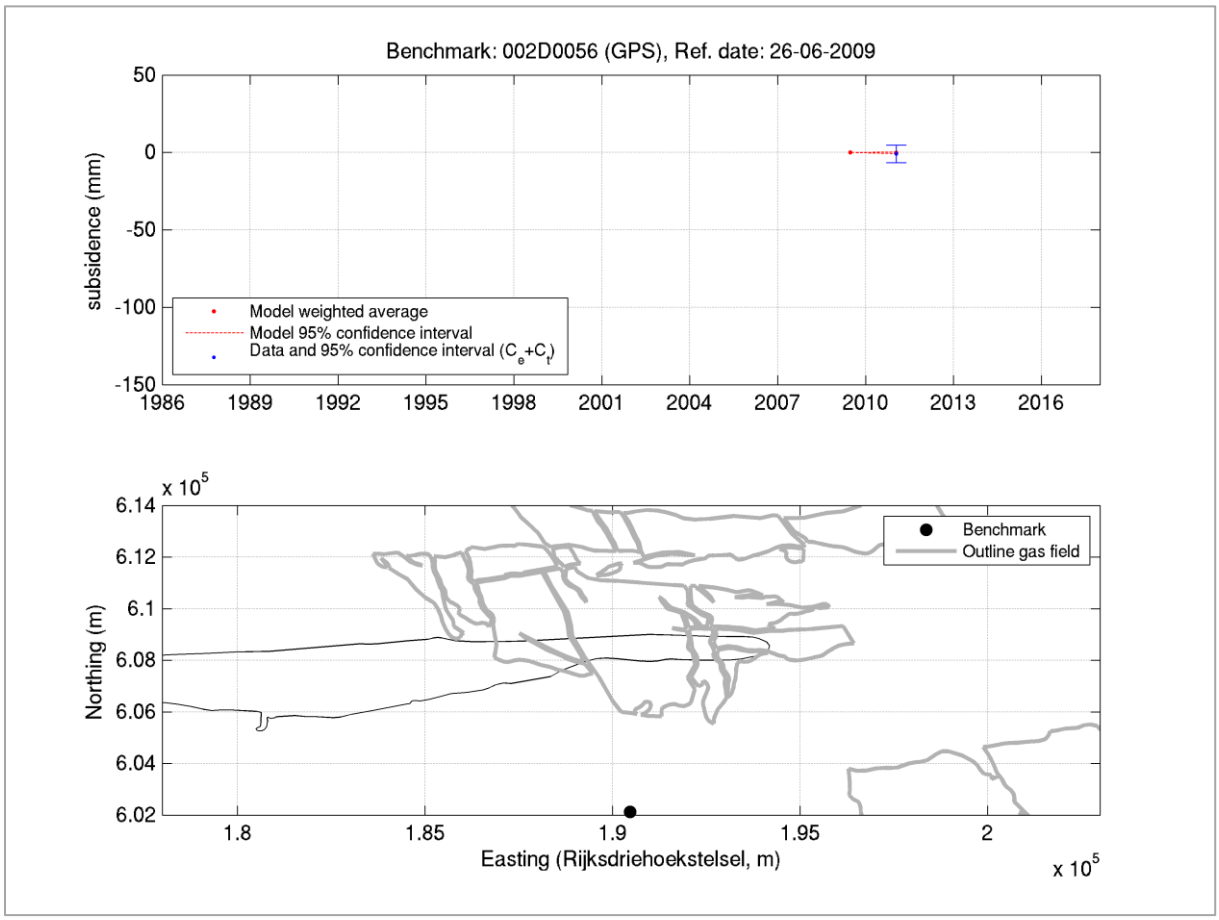

Benchmark: 002D0050 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

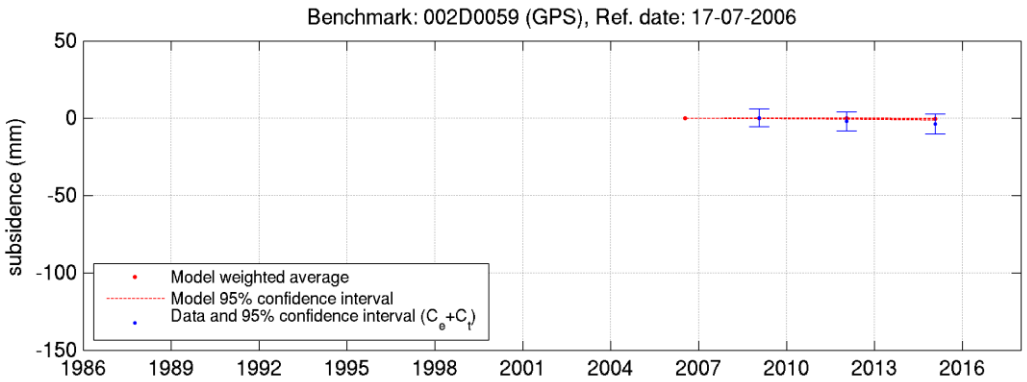

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

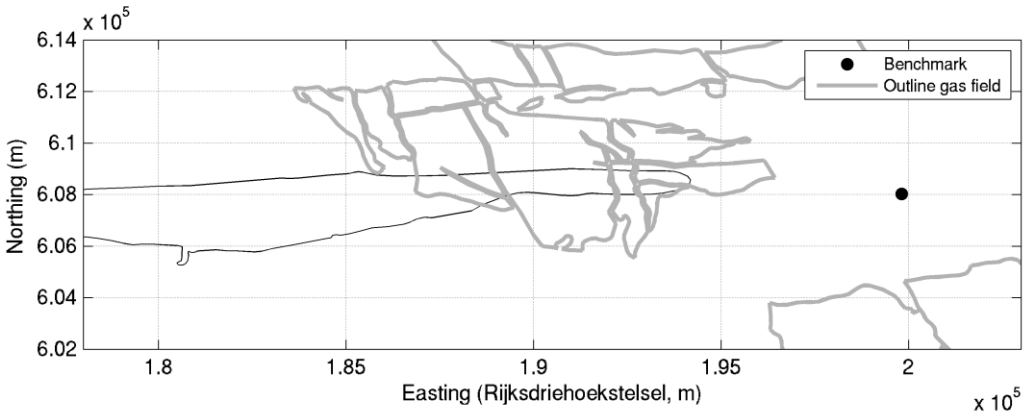

Benchmark: 002D0054 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

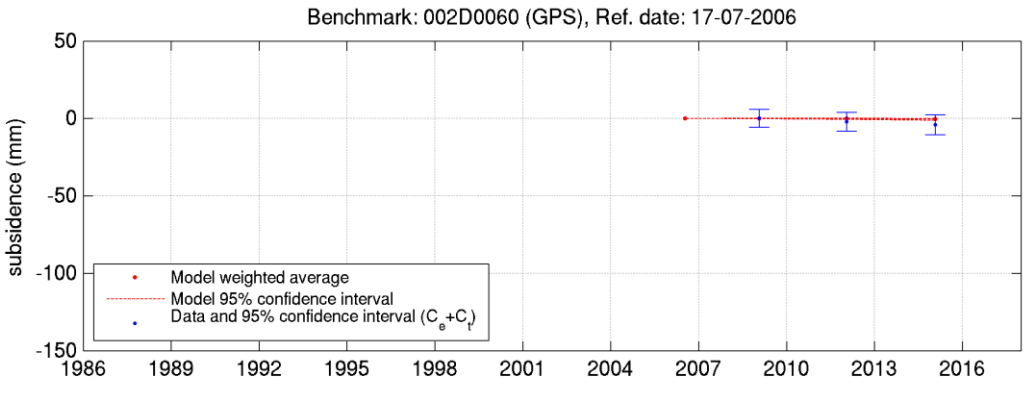

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

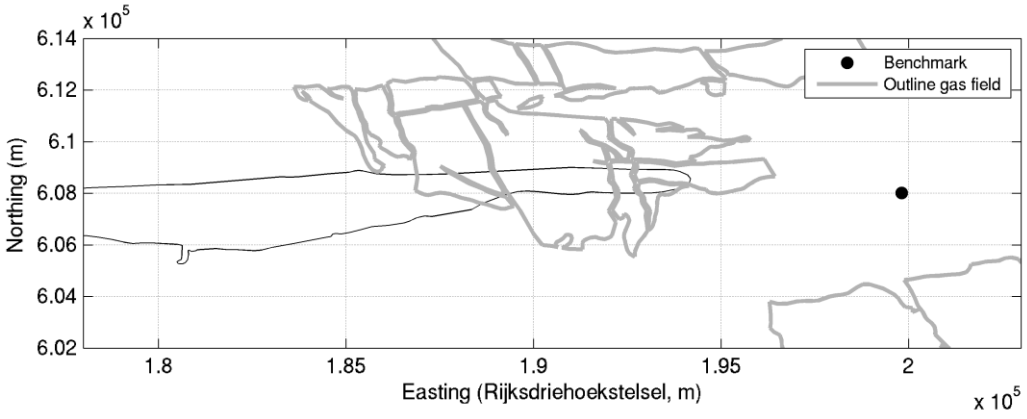

Benchmark: 002D0055 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

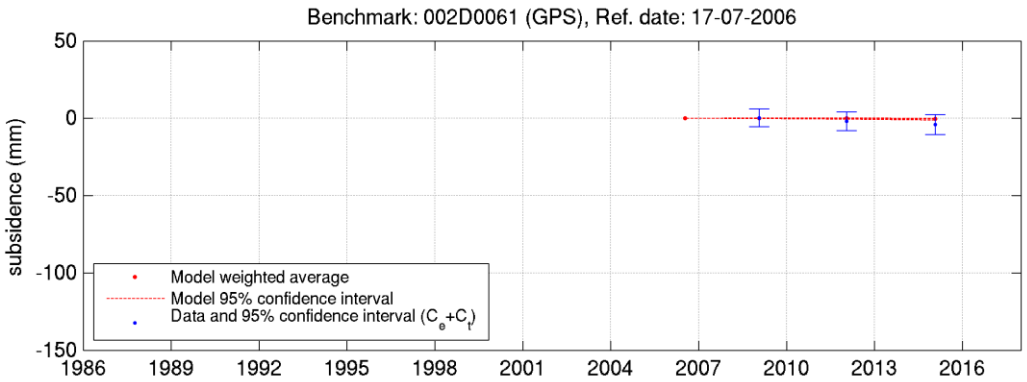

Benchmark: 002D0056 (GPS), Ref. date: 26-06-2009
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

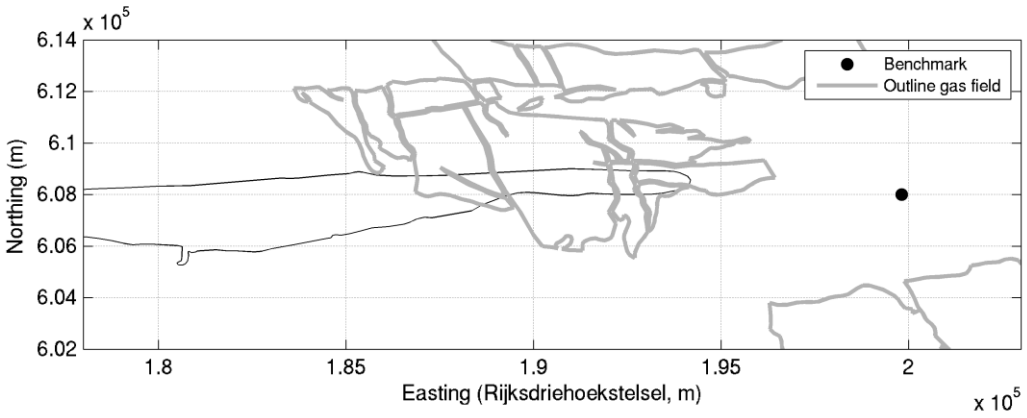

Benchmark: 002D0059 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

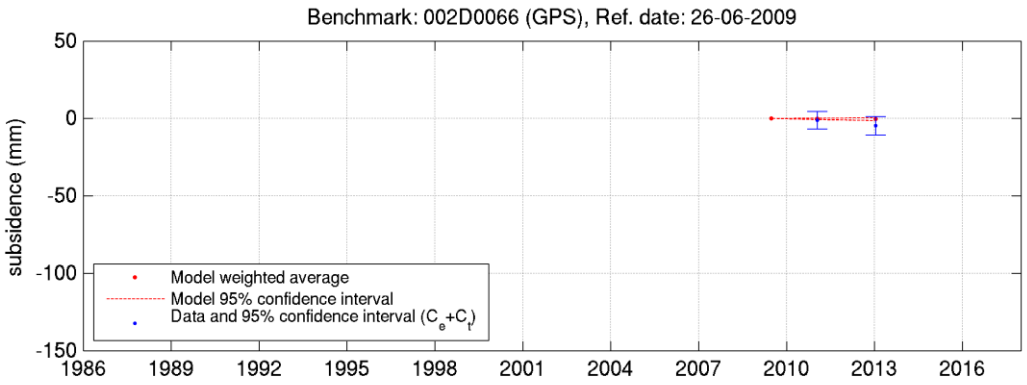

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

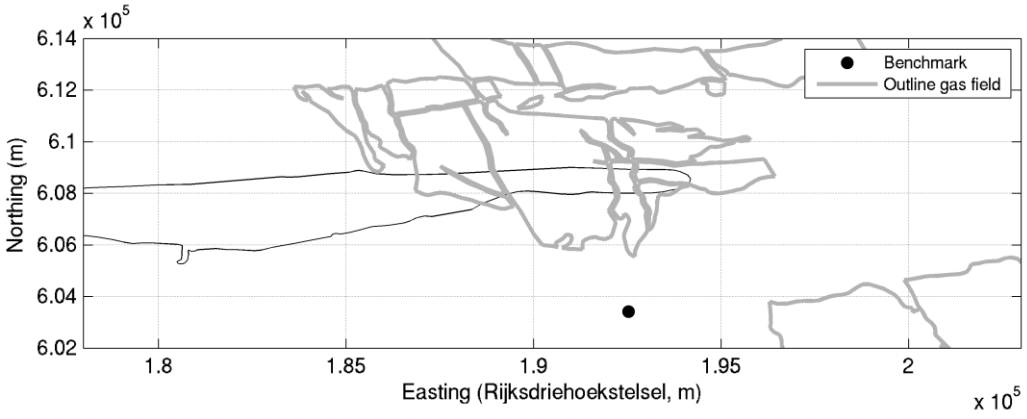

Benchmark: 002D0060 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

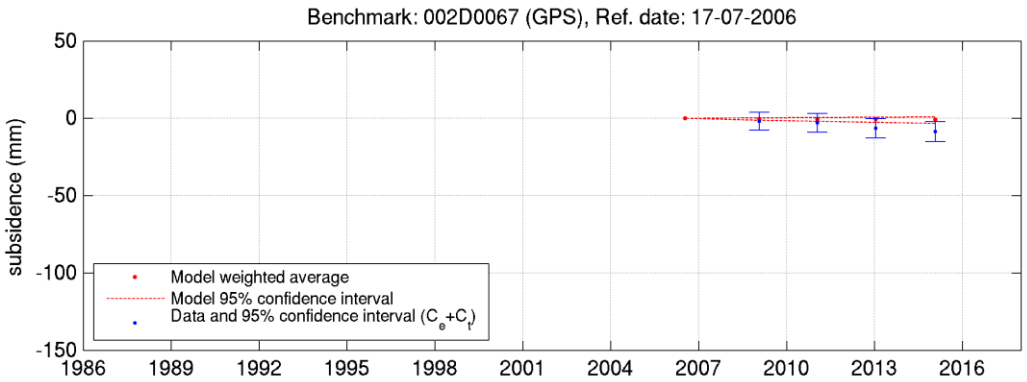

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

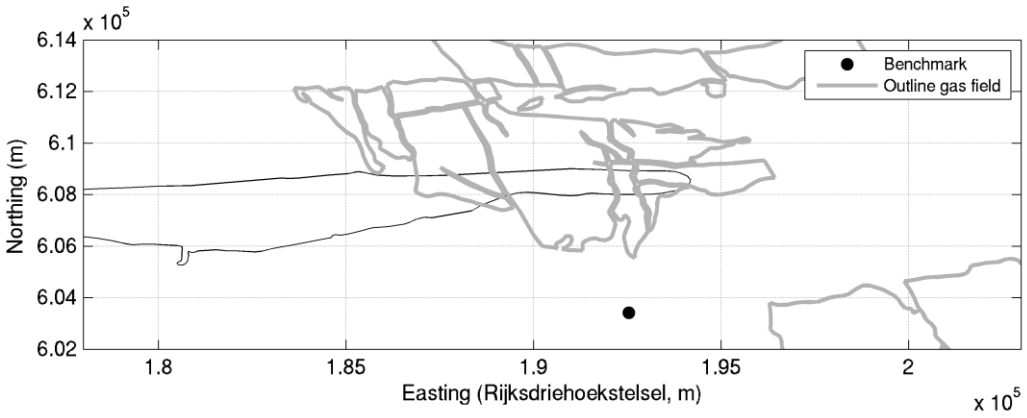

Benchmark: 002D0061 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

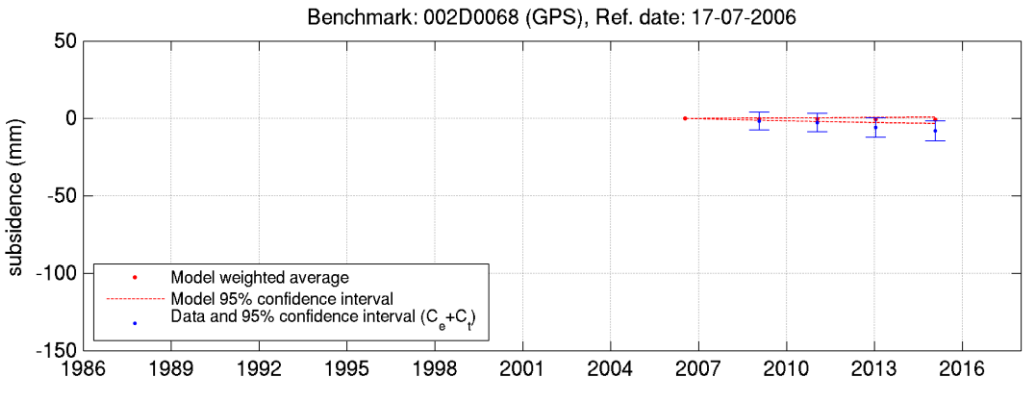

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

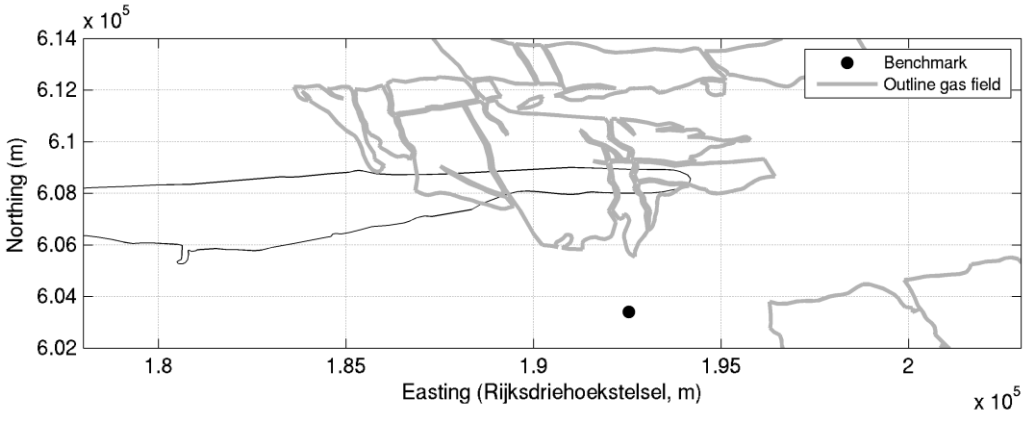

Benchmark: 002D0066 (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

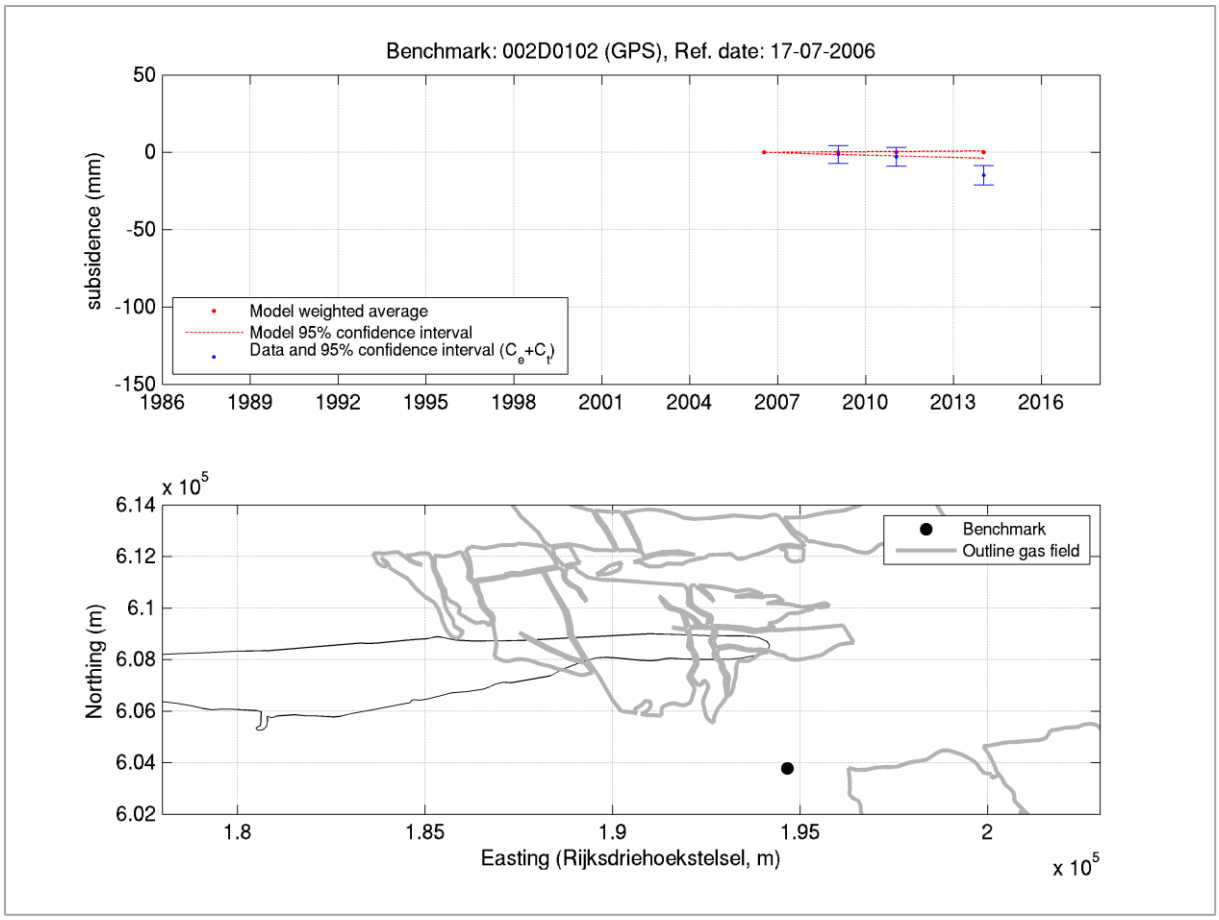

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

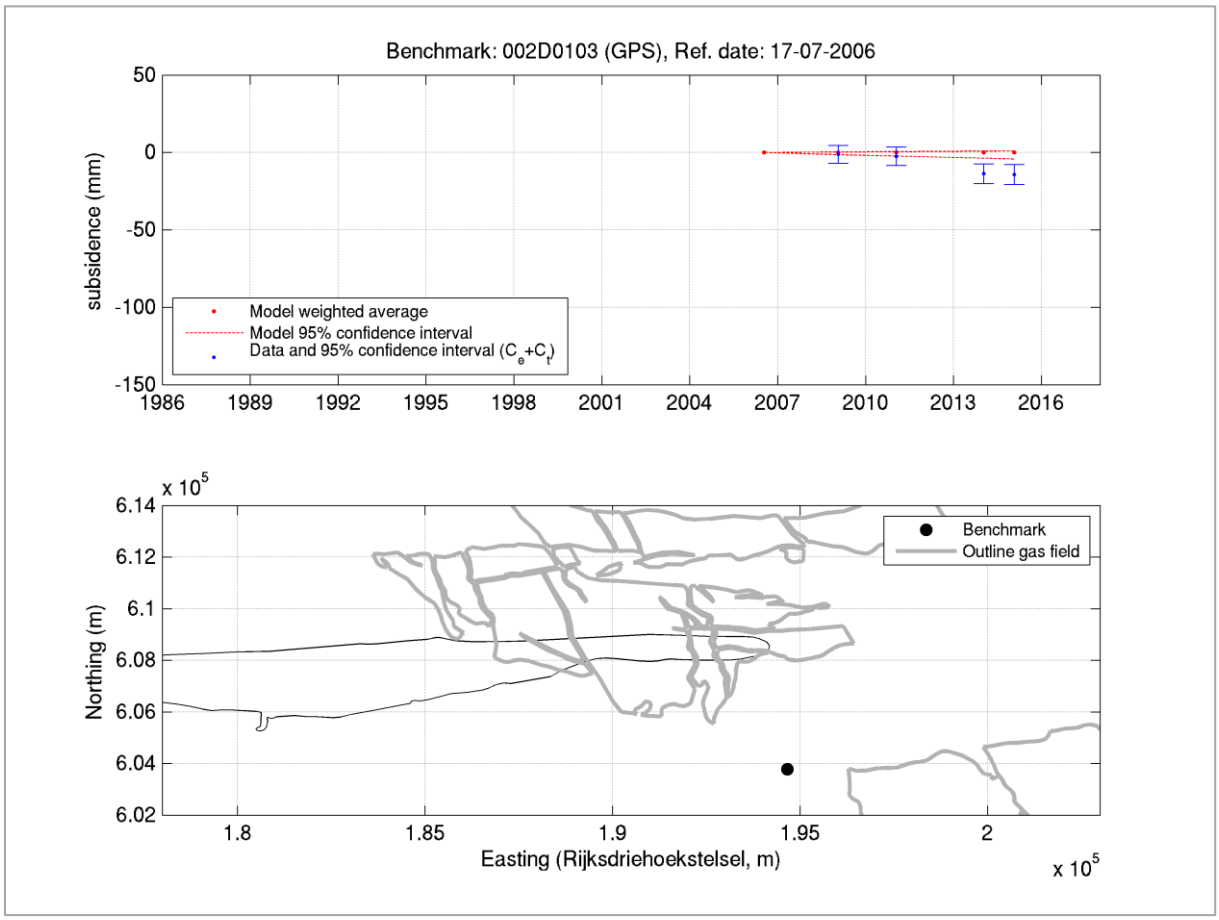

Benchmark: 002D0067 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

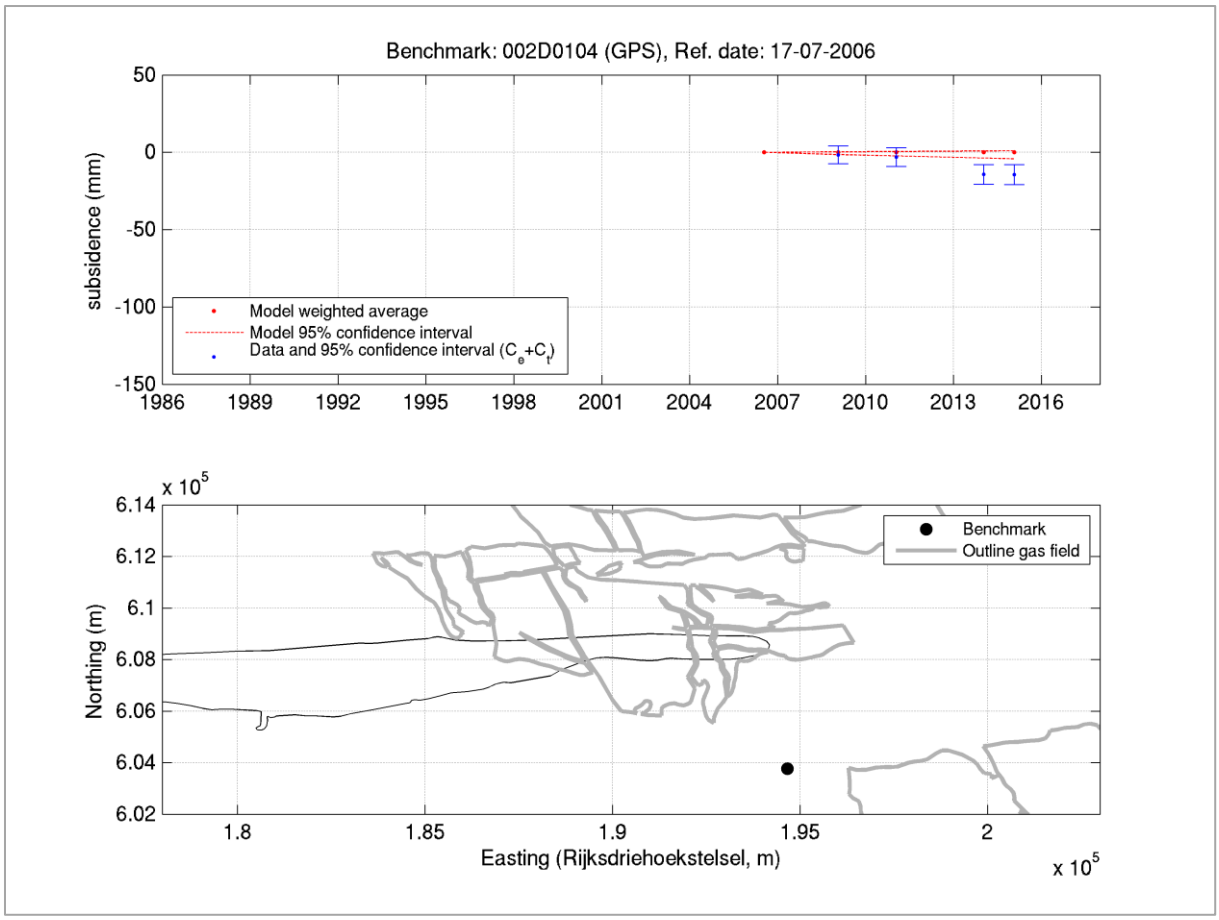

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

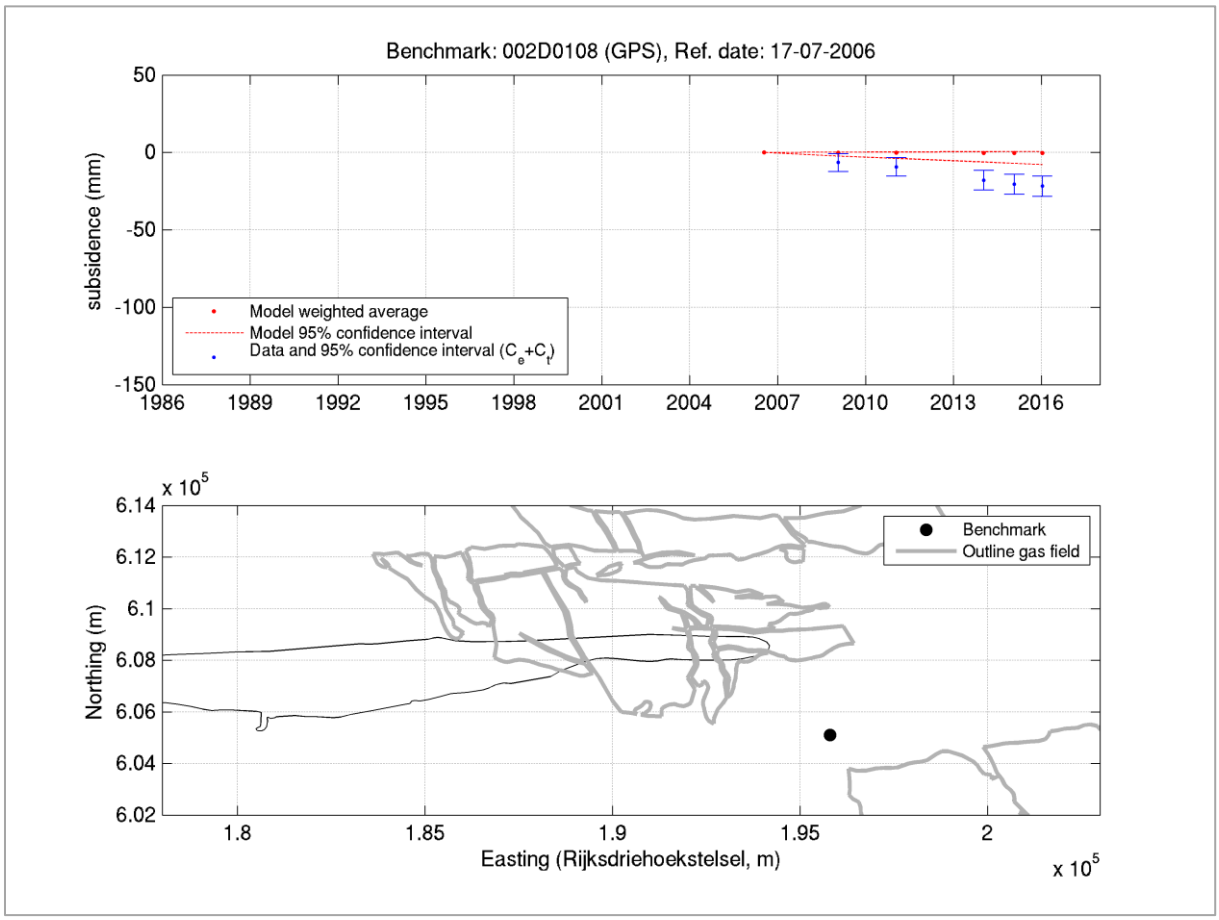

Benchmark: 002D0068 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

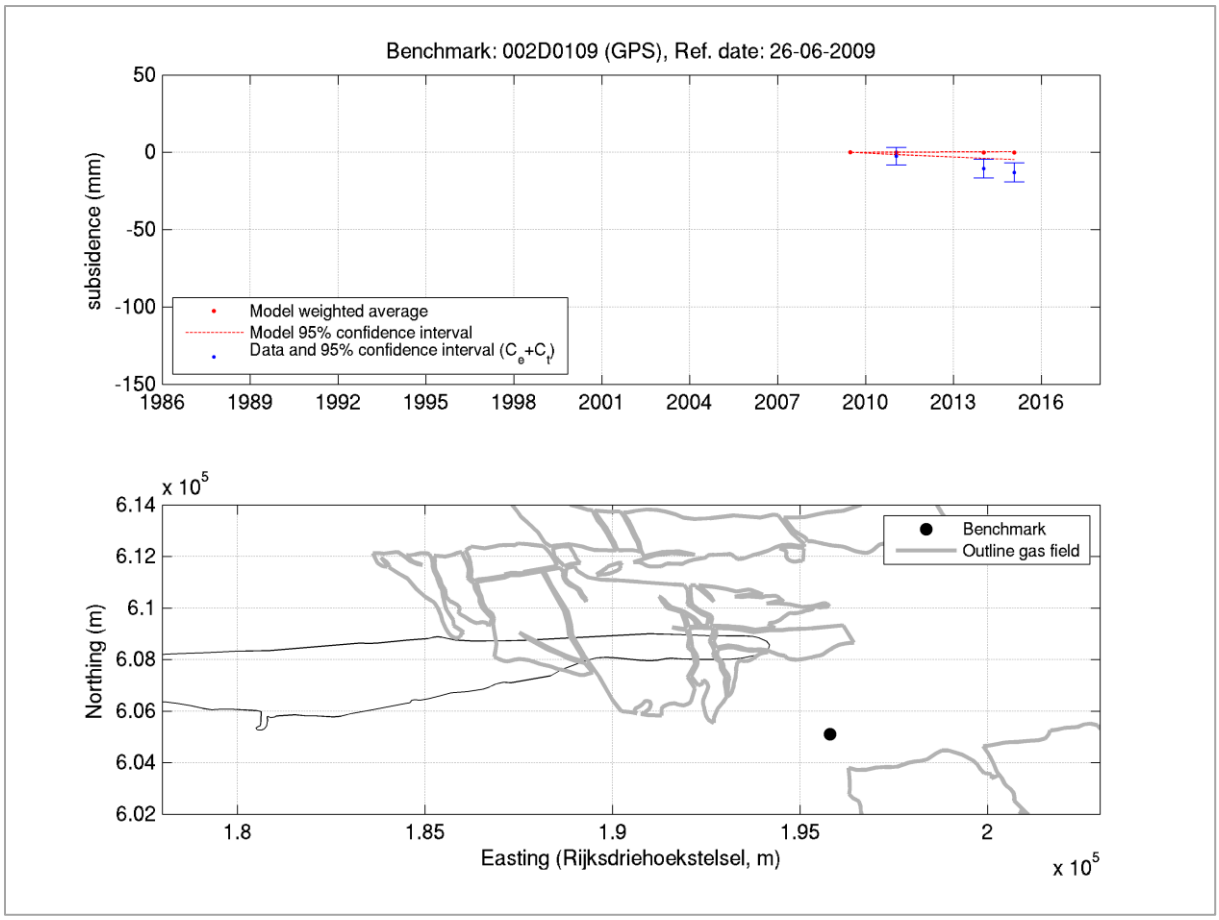

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

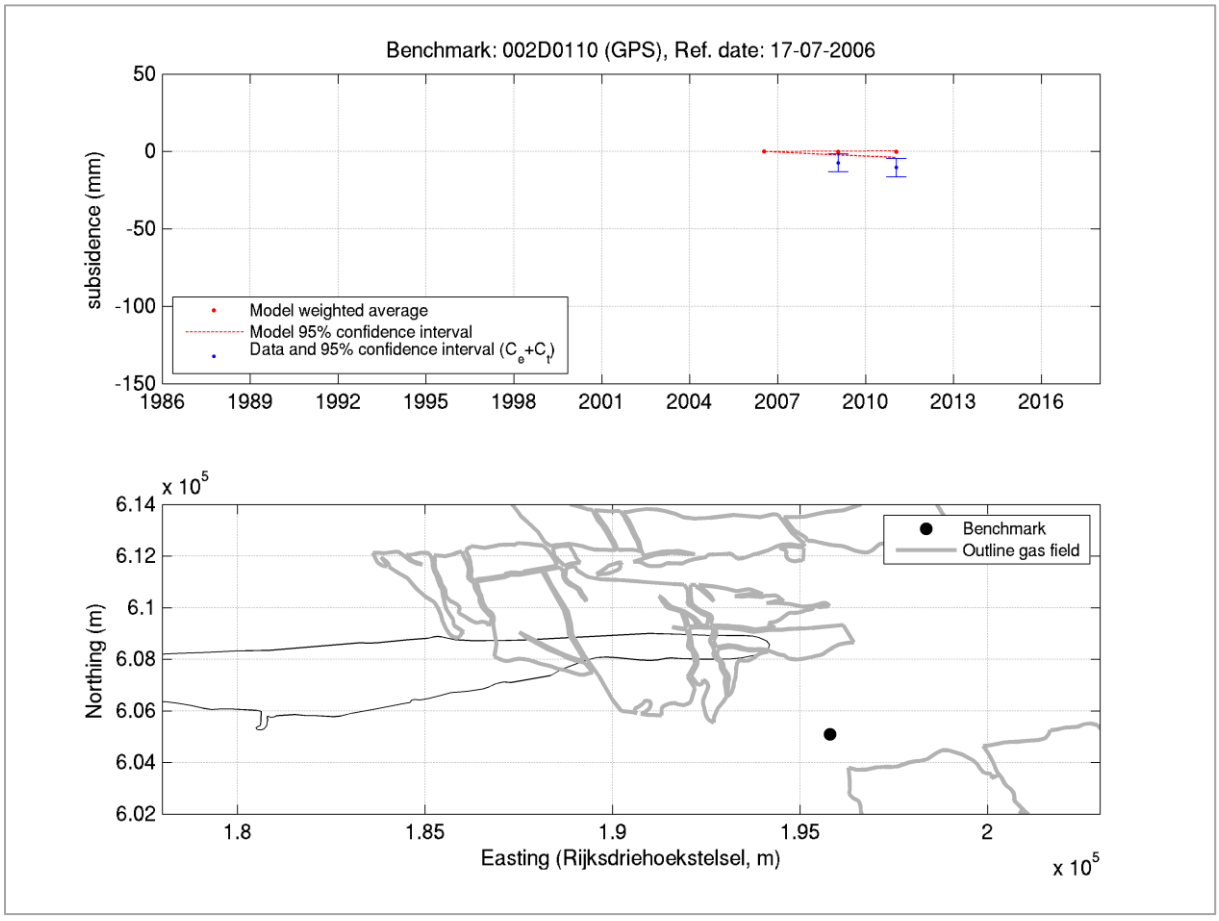

Benchmark: 002D0102 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

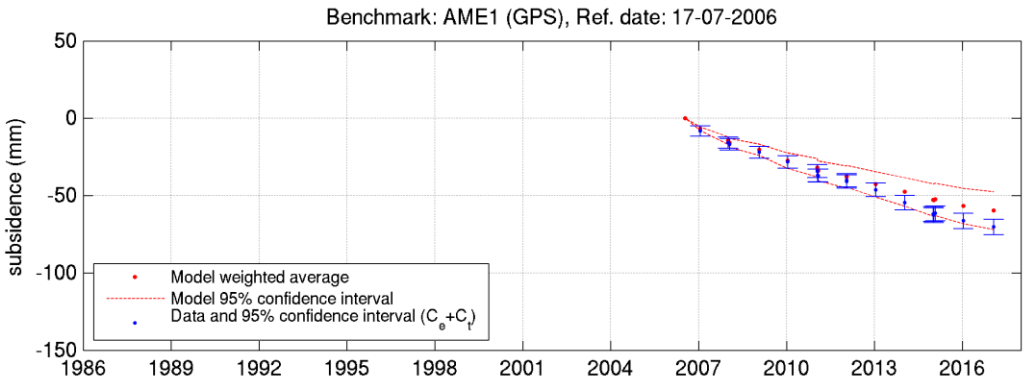

Benchmark: 002D0103 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (Ce+Ct)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

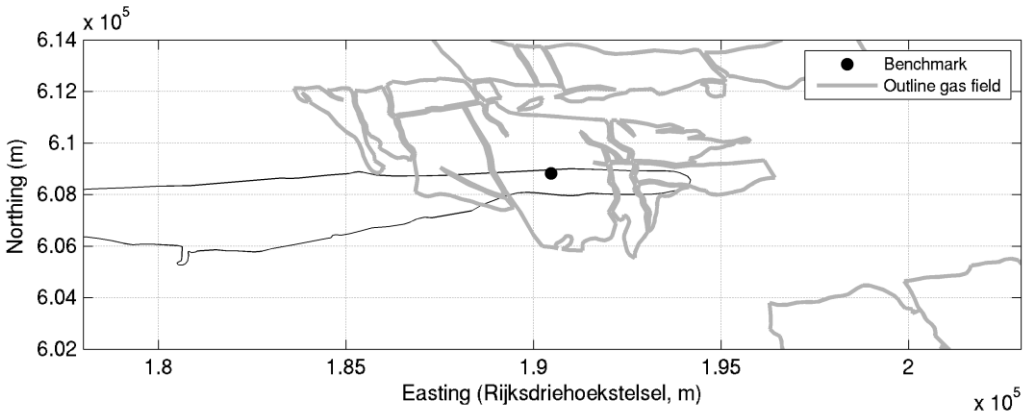

Benchmark: 002D0104 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

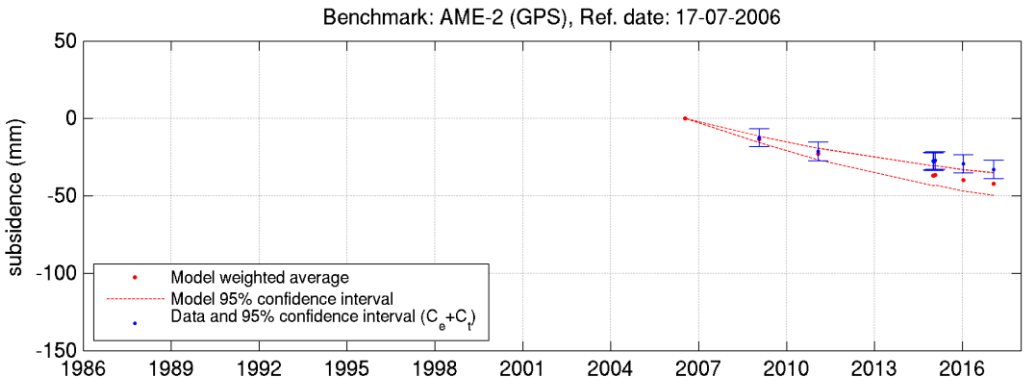

Benchmark: 002D0108 (GPS), Ref. date: 17-07-2006
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

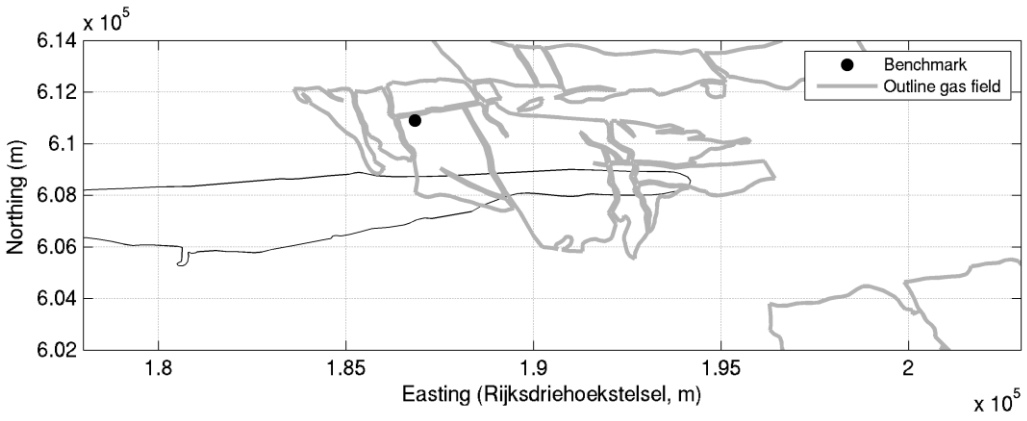

Benchmark: 002D0109 (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

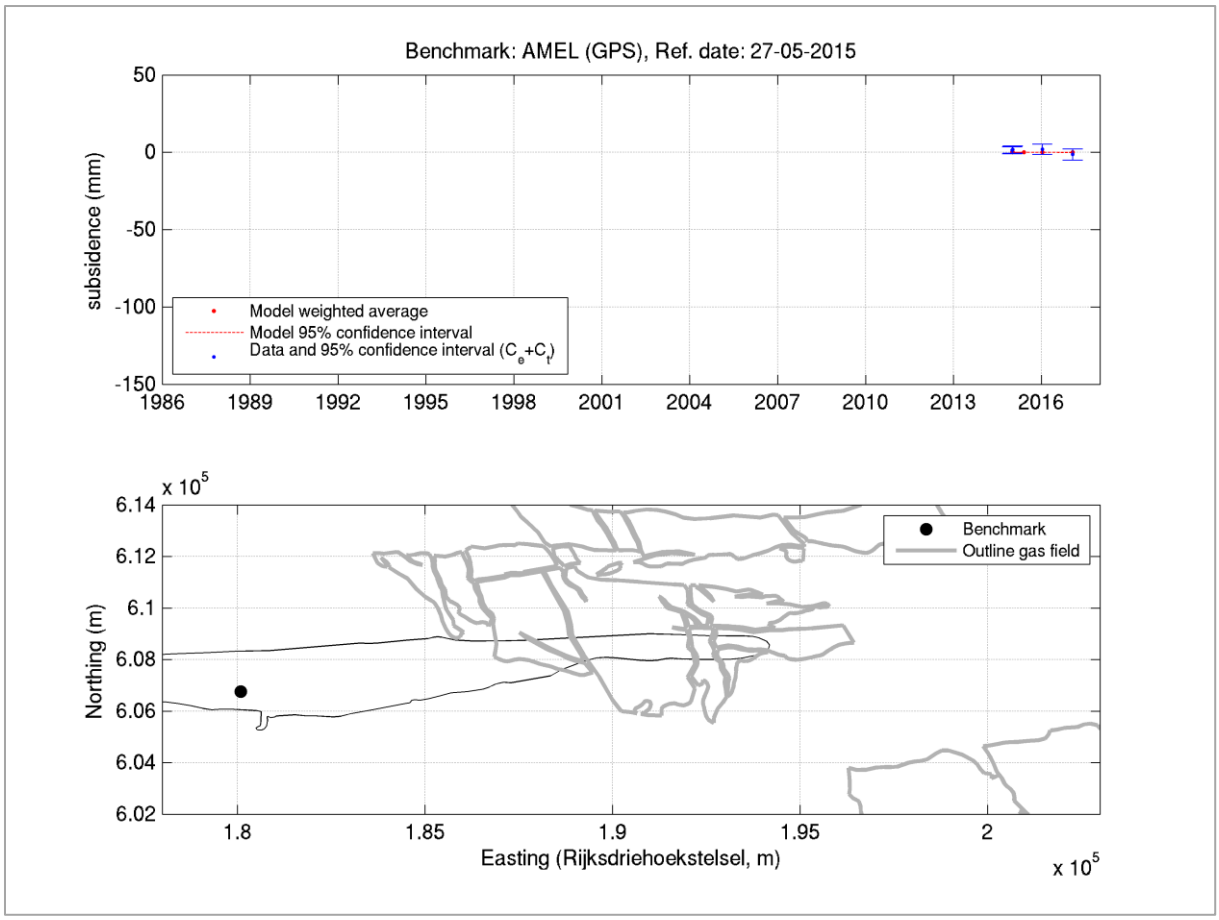

Benchmark: 002D0110 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

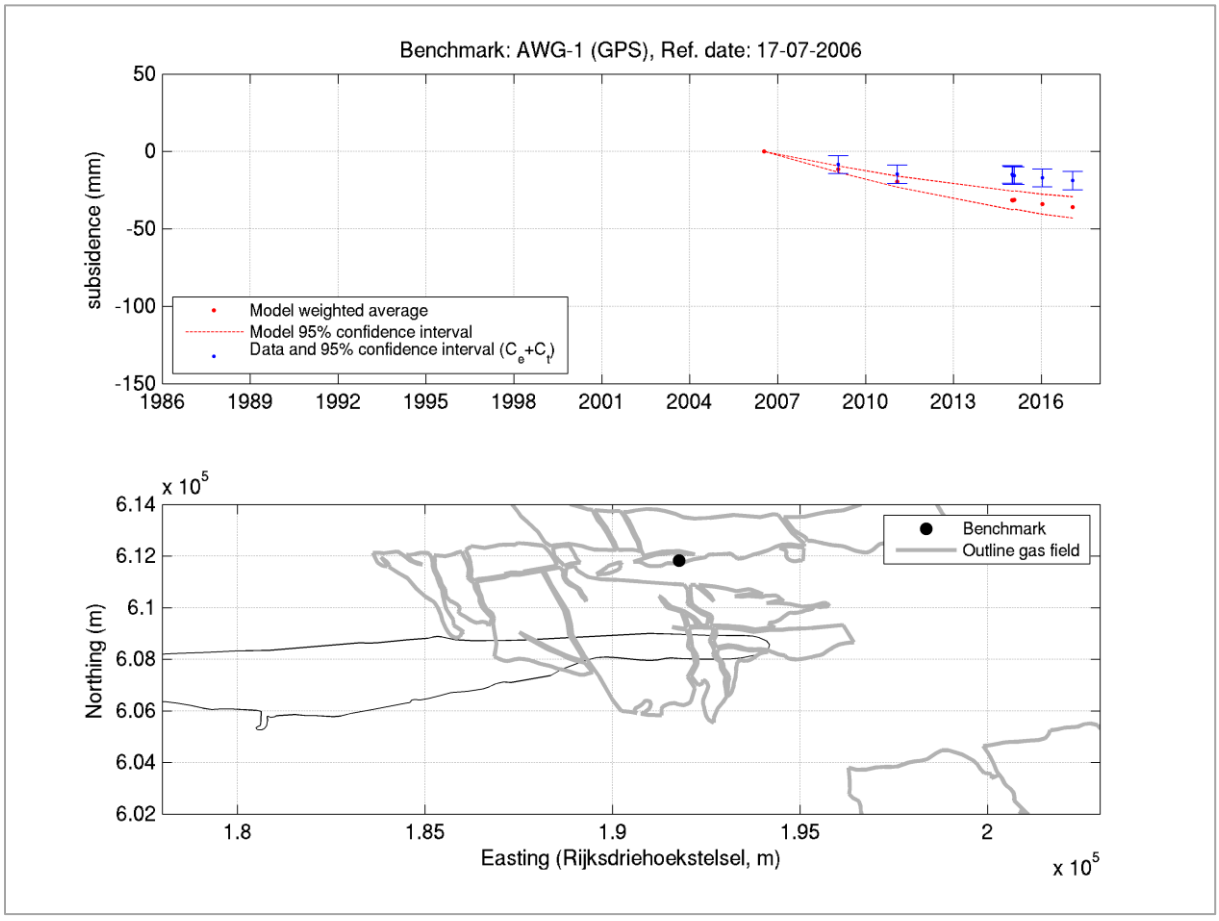

Benchmark: AME1 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

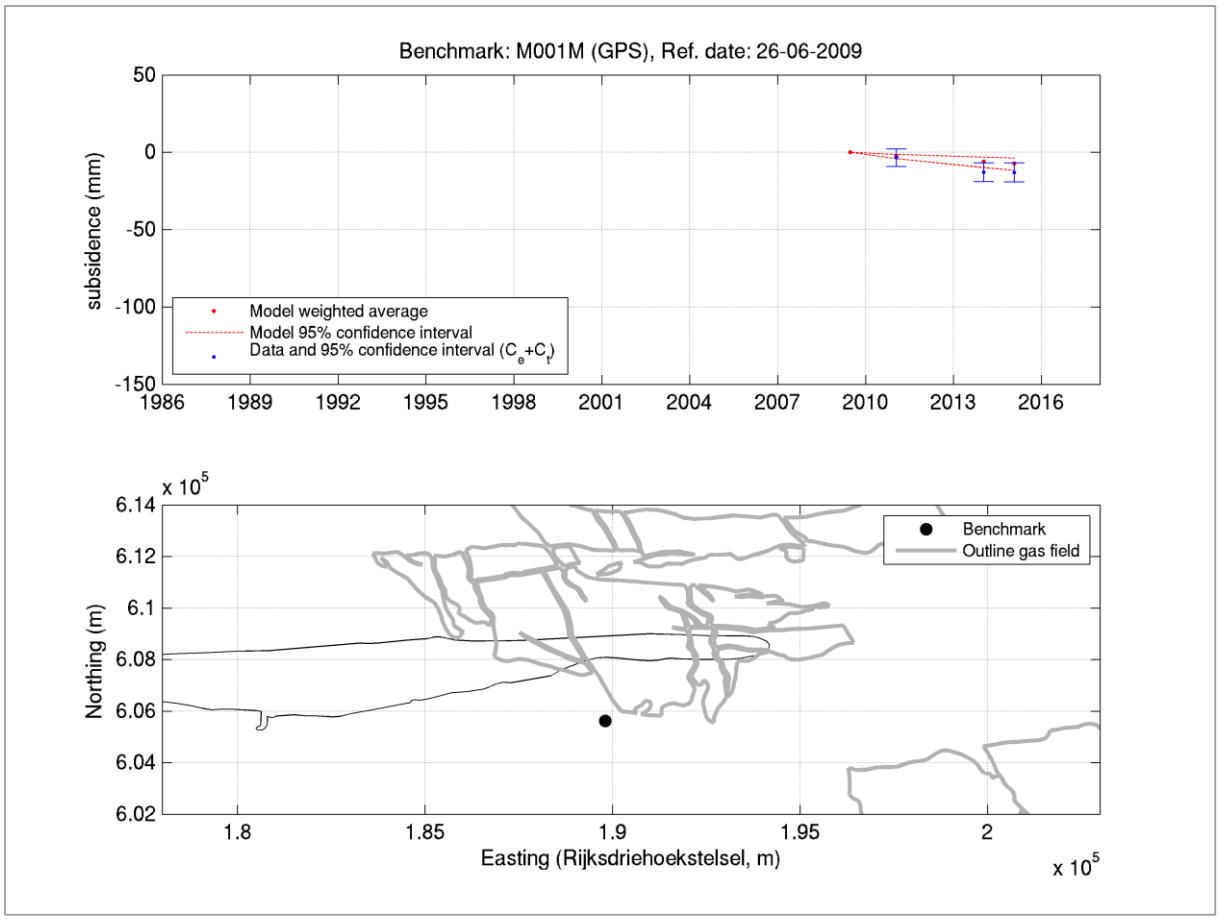

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

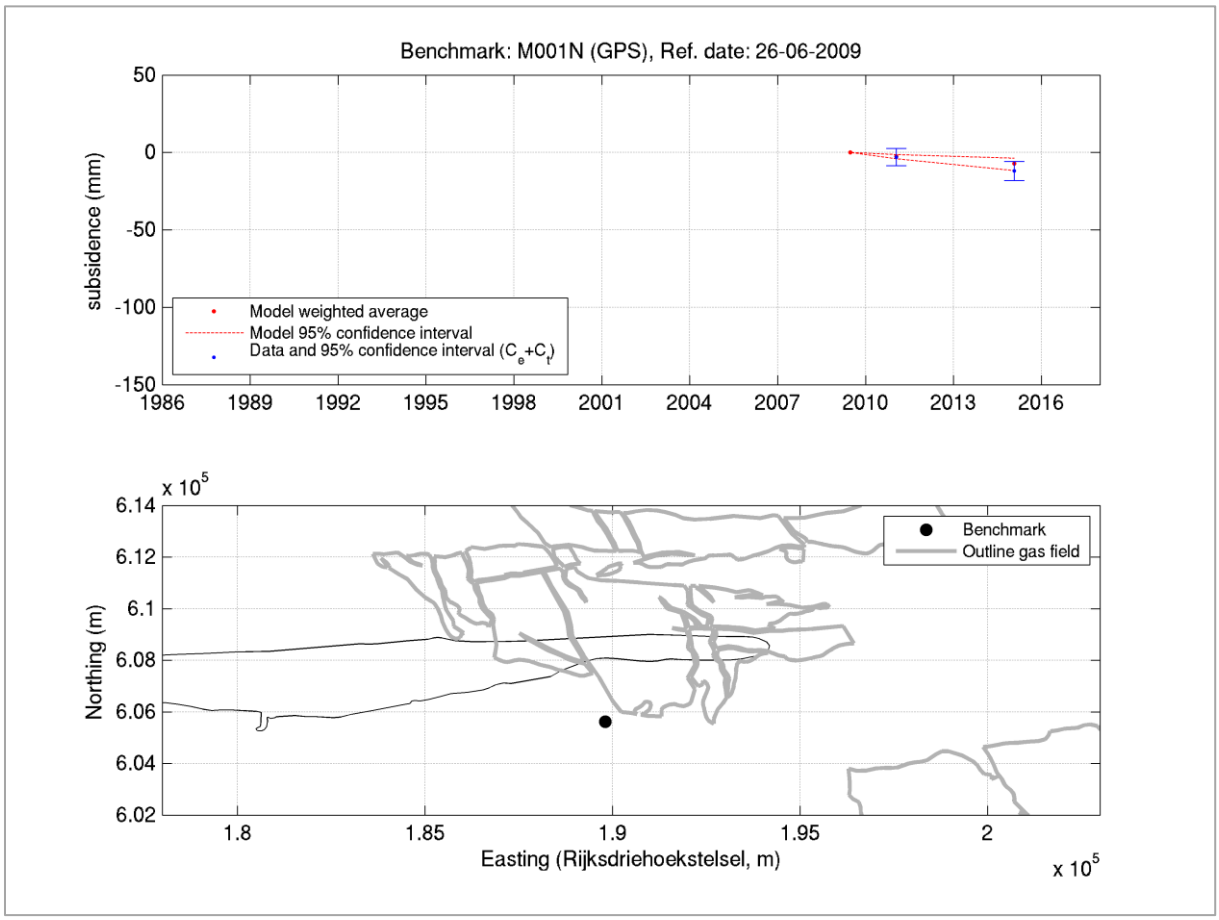

Benchmark: AME-2 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

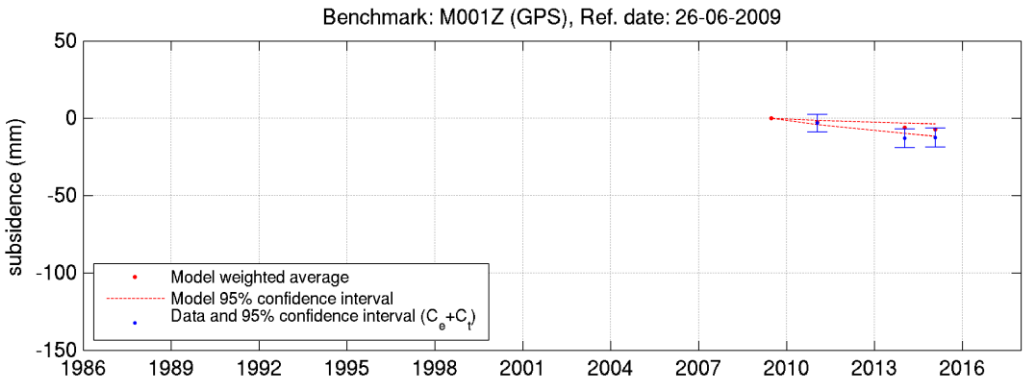

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

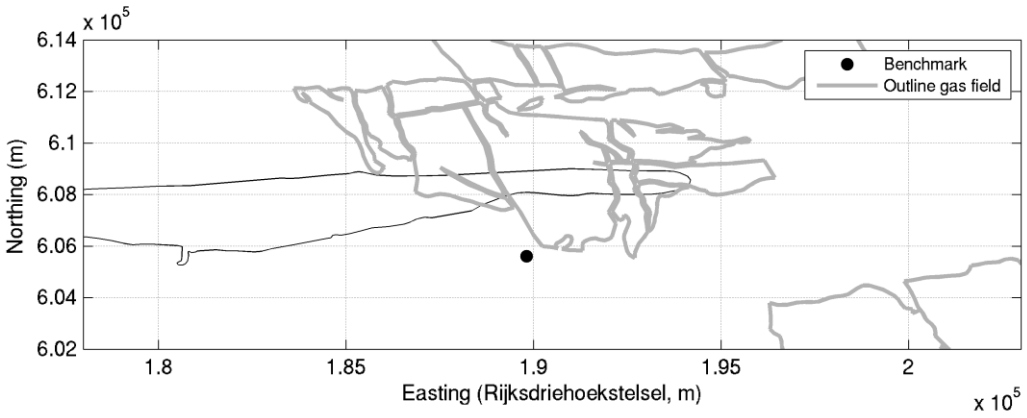

Benchmark: AMEL (GPS), Ref. date: 27-05-2015
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

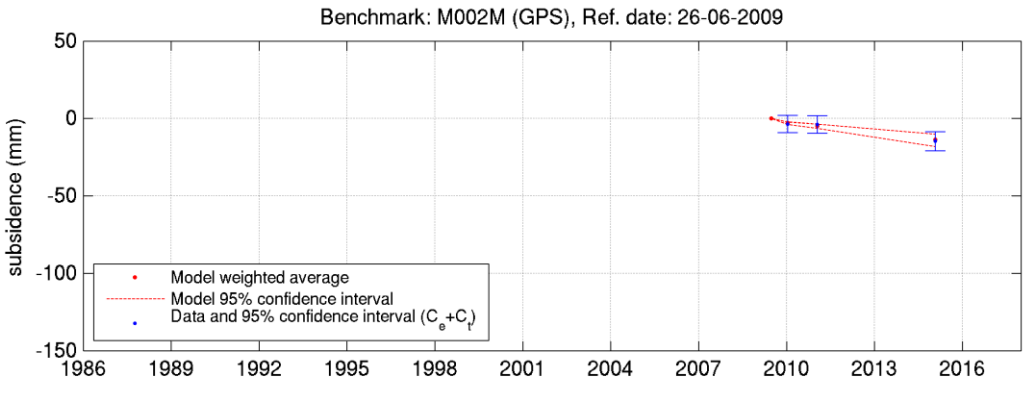

Benchmark: AWG-1 (GPS), Ref. date: 17-07-2006
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

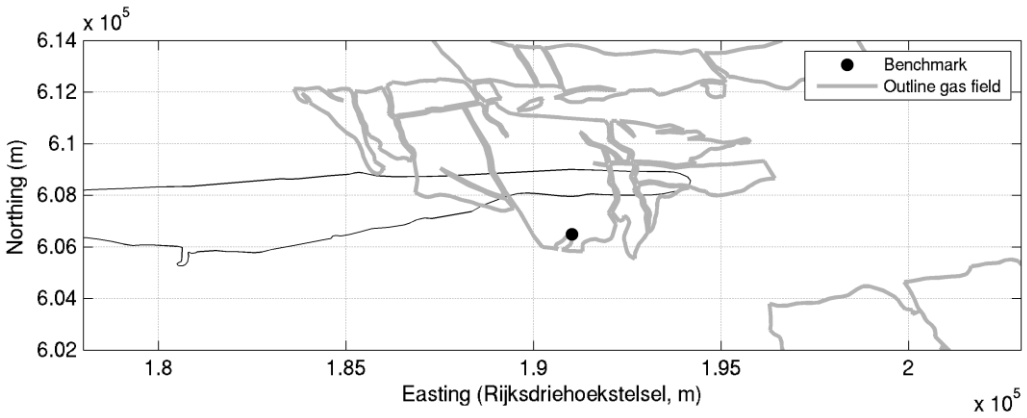

Benchmark: M001M (GPS), Ref. date: 26-06-2009
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

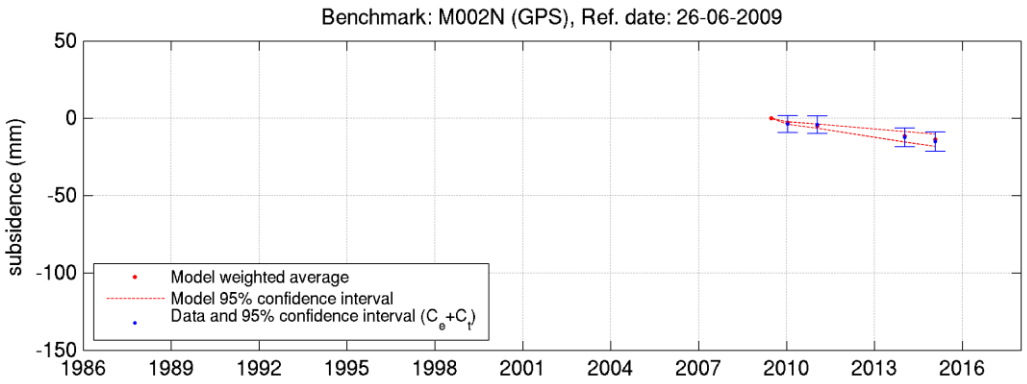

Benchmark: M001N (GPS), Ref. date: 26-06-2009
subsidence (mm)
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
Northing (m)
Benchmark
Outline gas field
Easting (Rijksdriehoekstelsel, m)
x 10^5

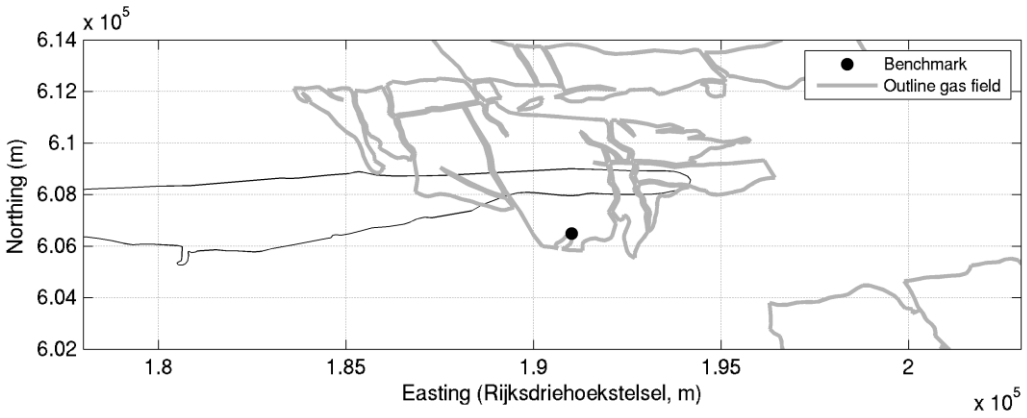

Benchmark: M001Z (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

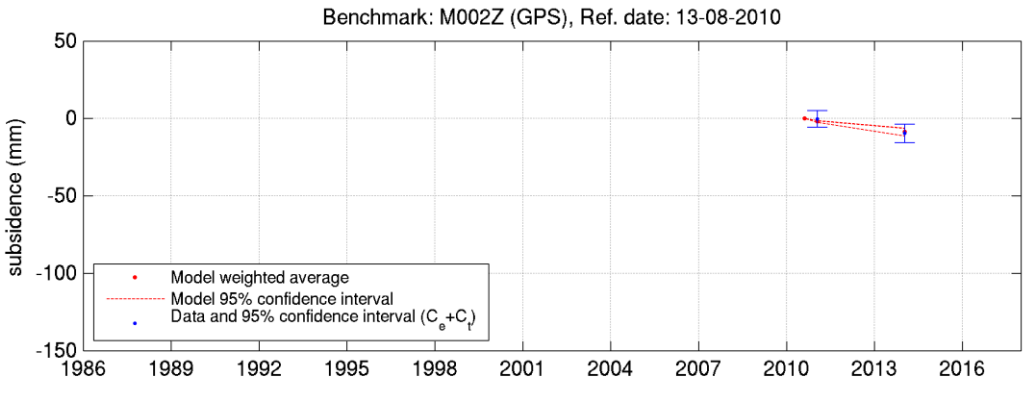

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

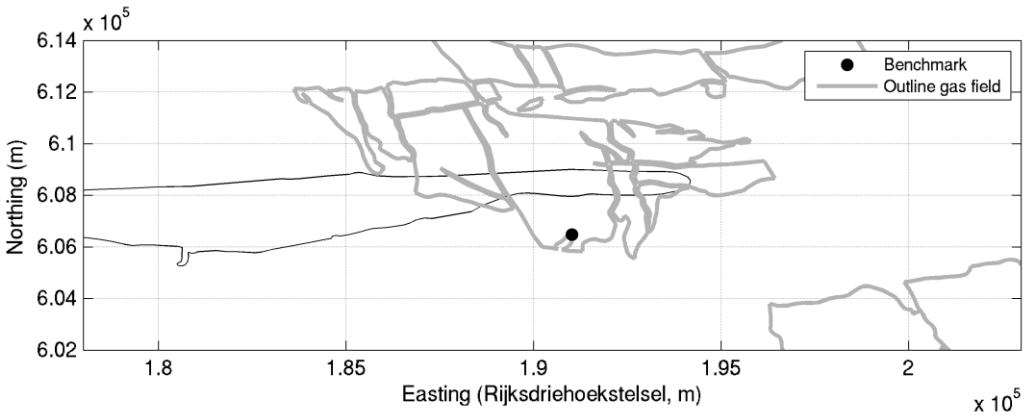

Benchmark: M002M (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

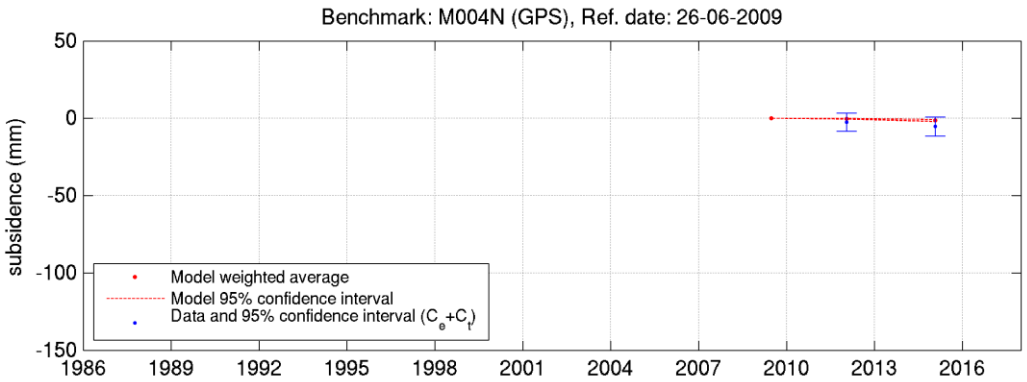

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

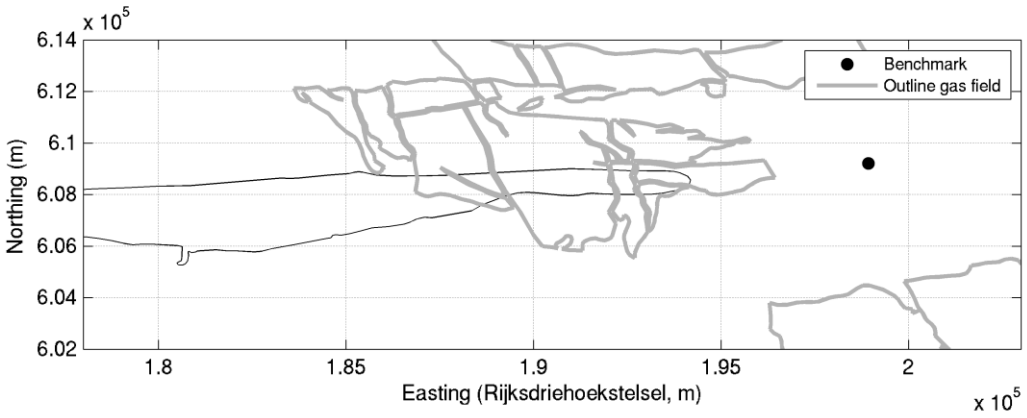

Benchmark: M002N (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

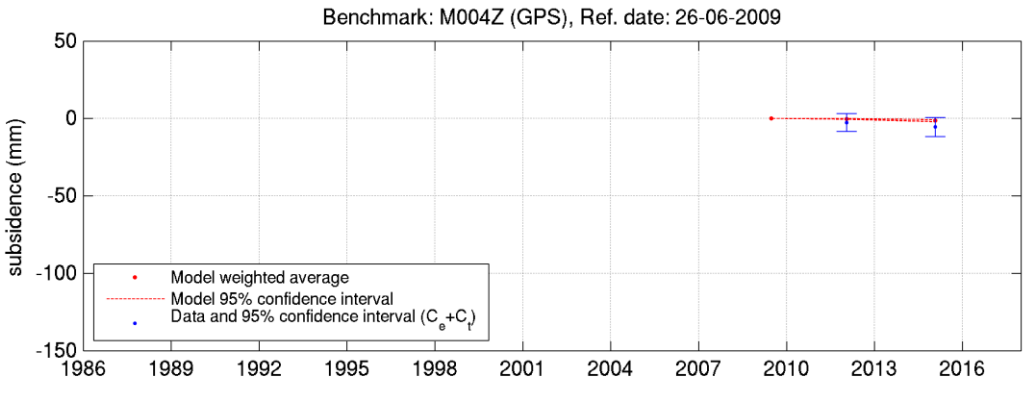

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

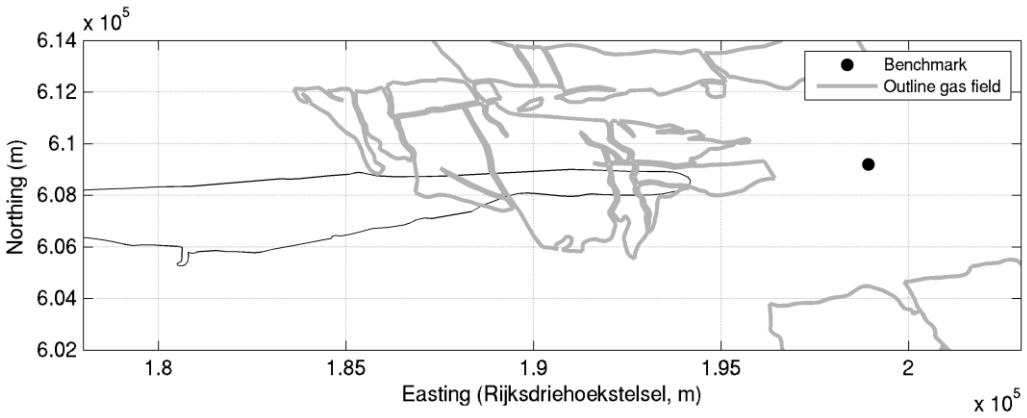

Benchmark: M002Z (GPS), Ref. date: 13-08-2010
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

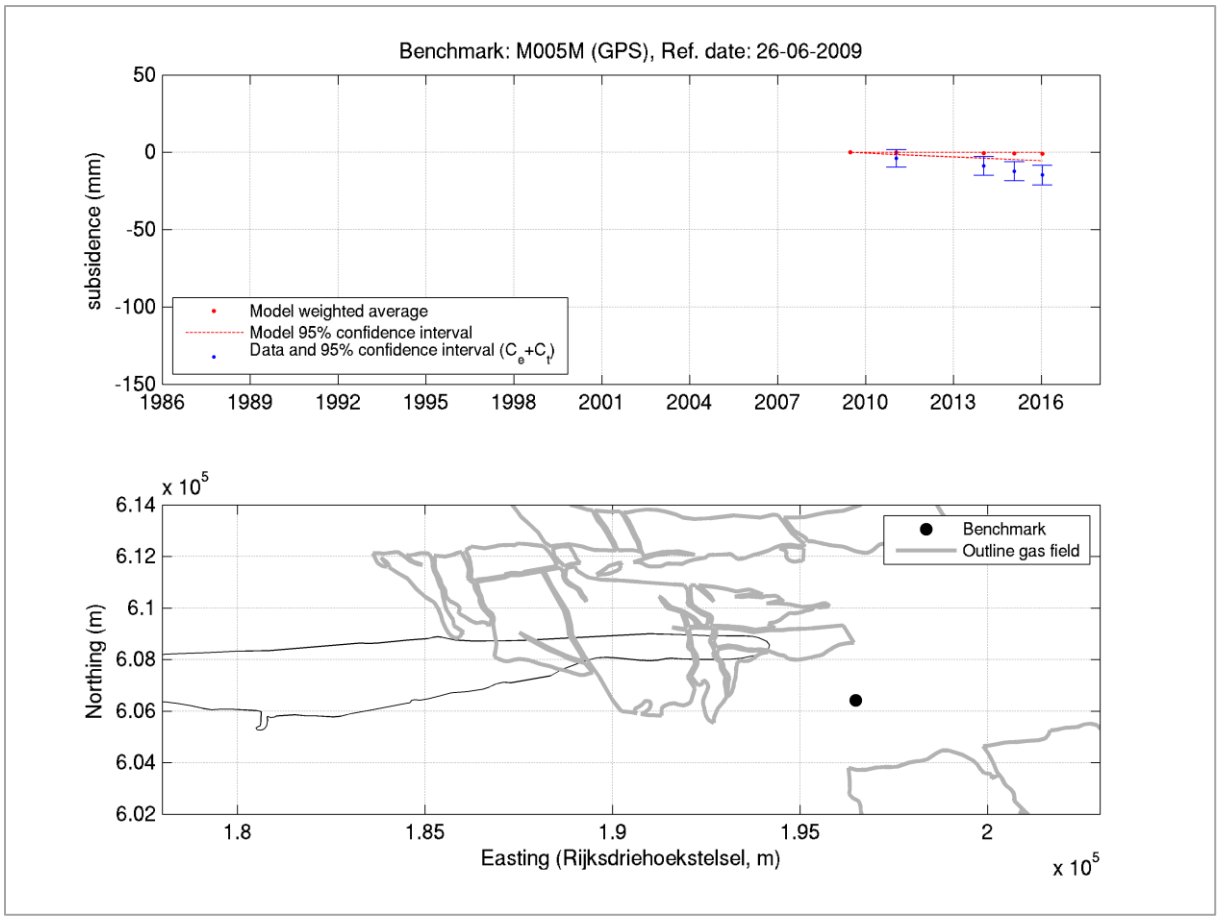

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

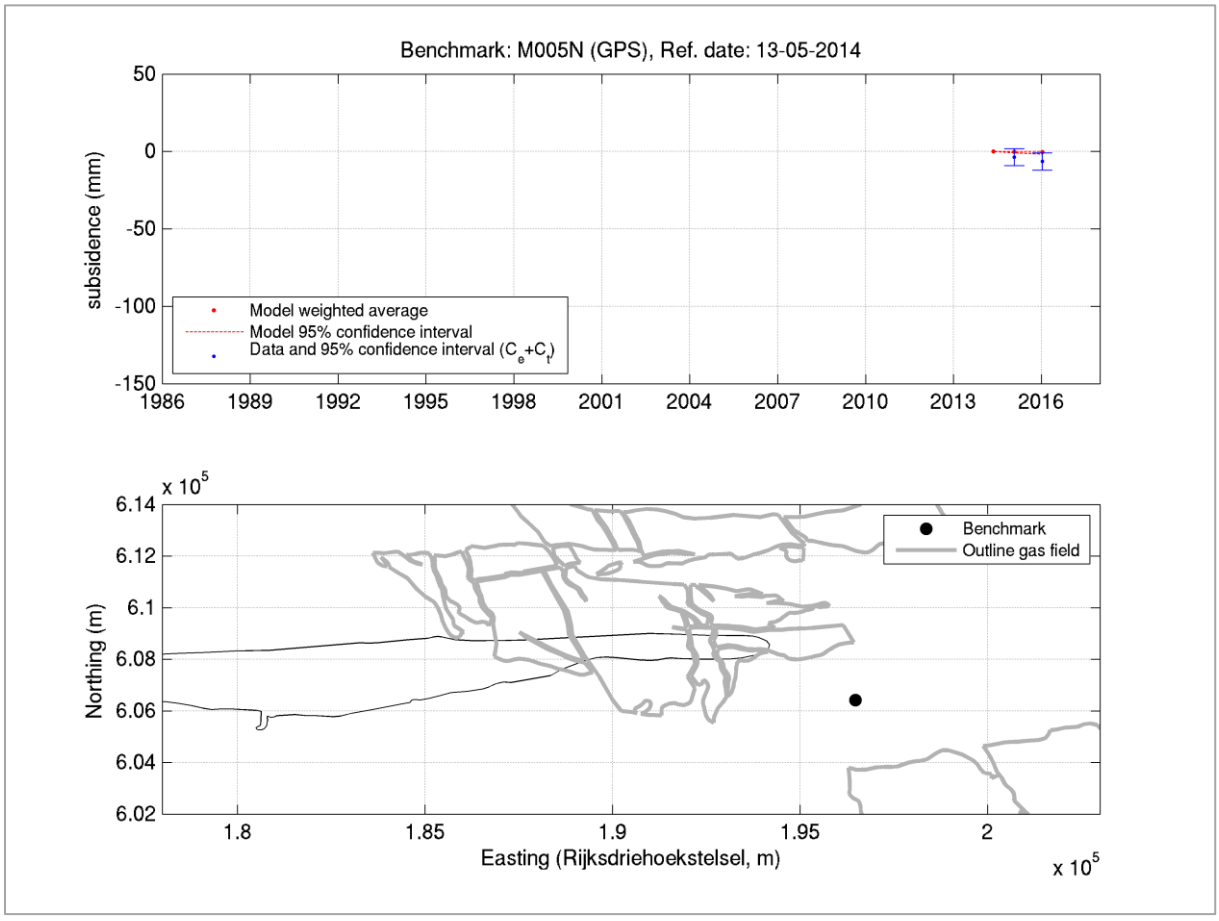

Benchmark: M004N (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

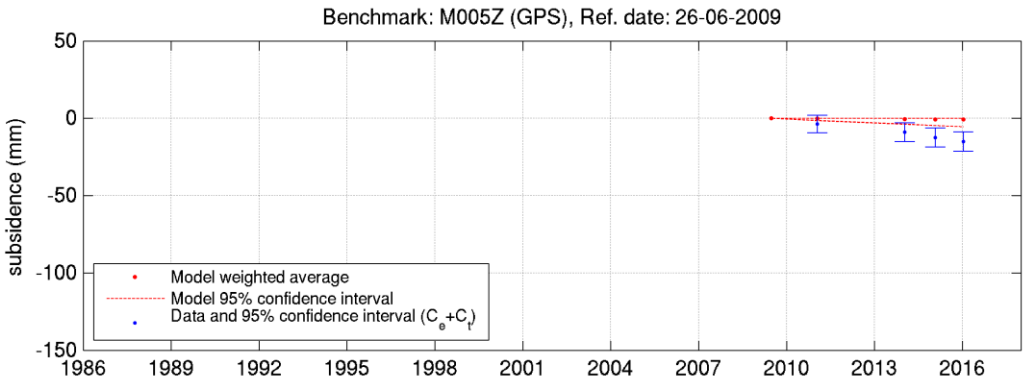

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

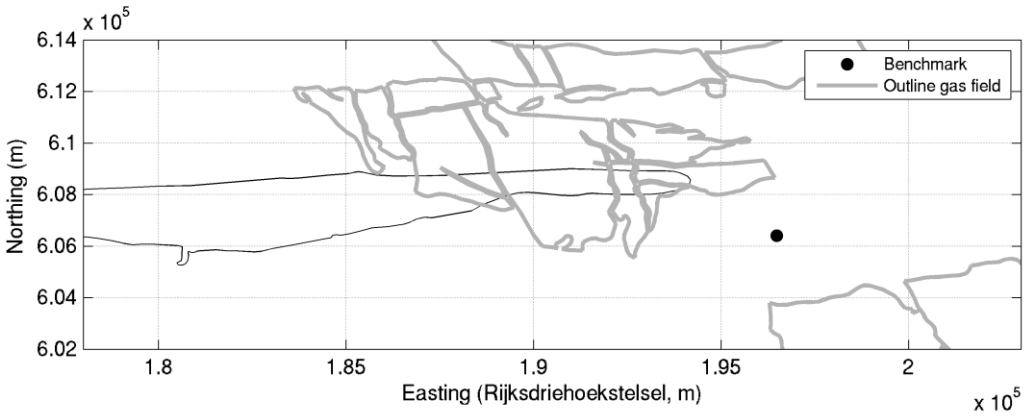

Benchmark: M004Z (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

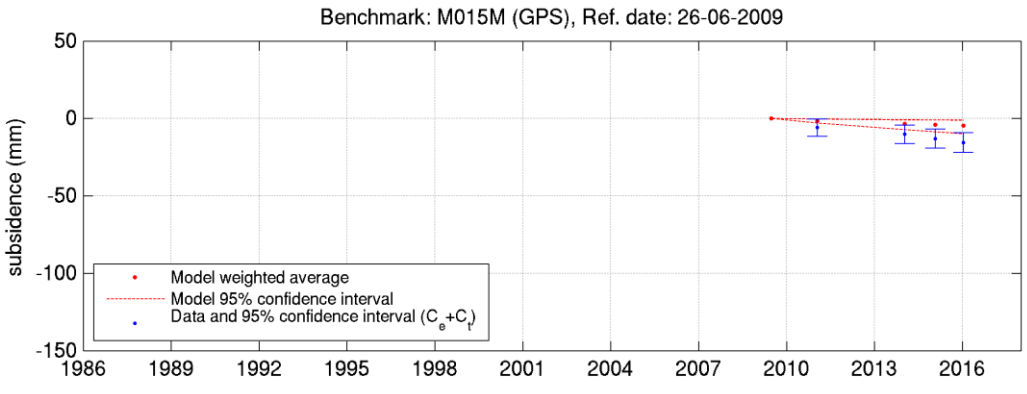

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

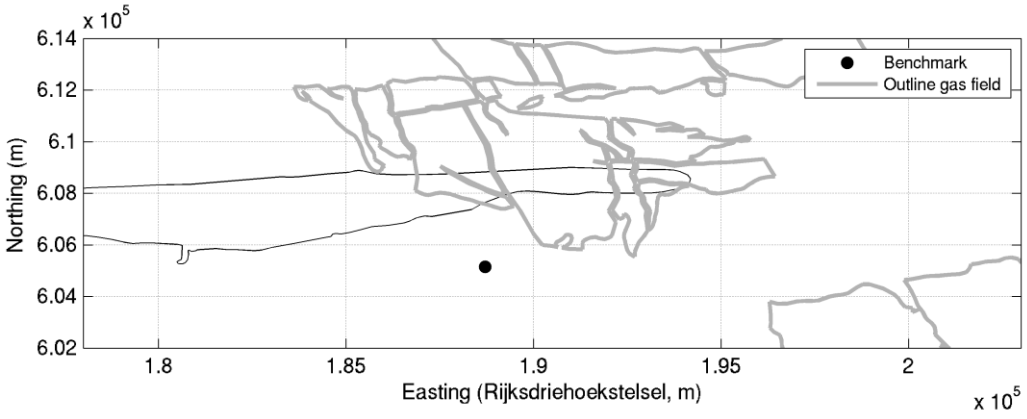

Benchmark: M005M (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

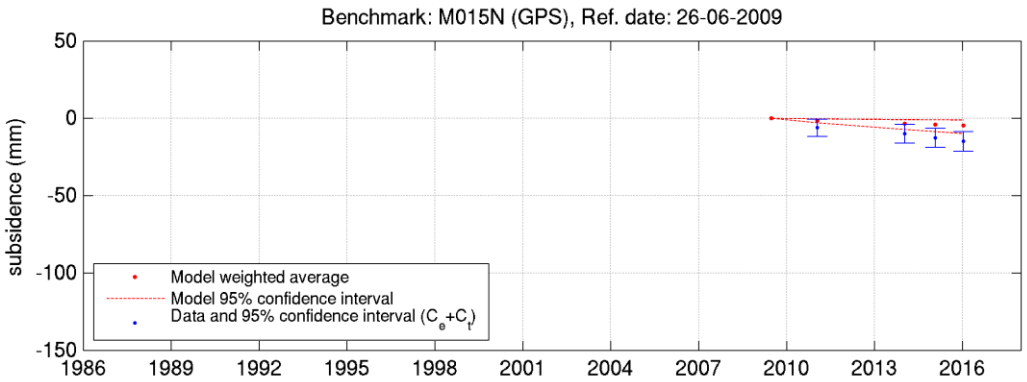

Benchmark: M005N (GPS), Ref. date: 13-05-2014
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)
x 10^5
6.14
6.12
6.1
6.08
6.06
6.04
6.02
Northing (m)
Benchmark
Outline gas field
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5

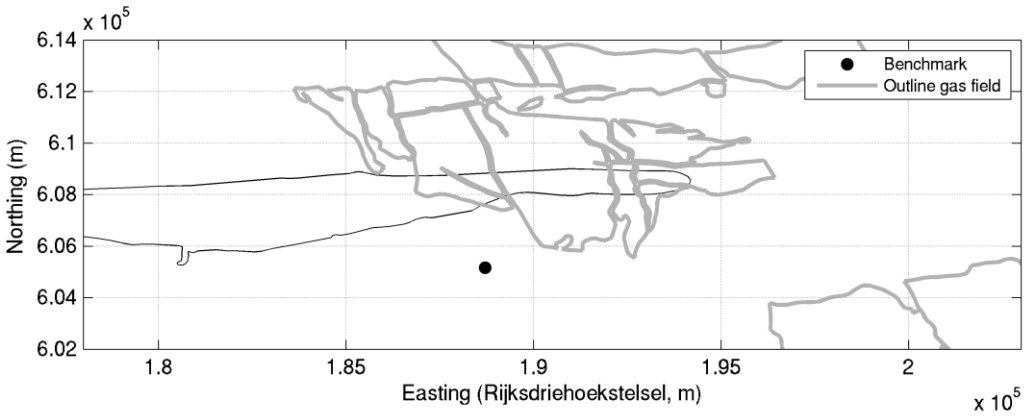

Benchmark: M005Z (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)

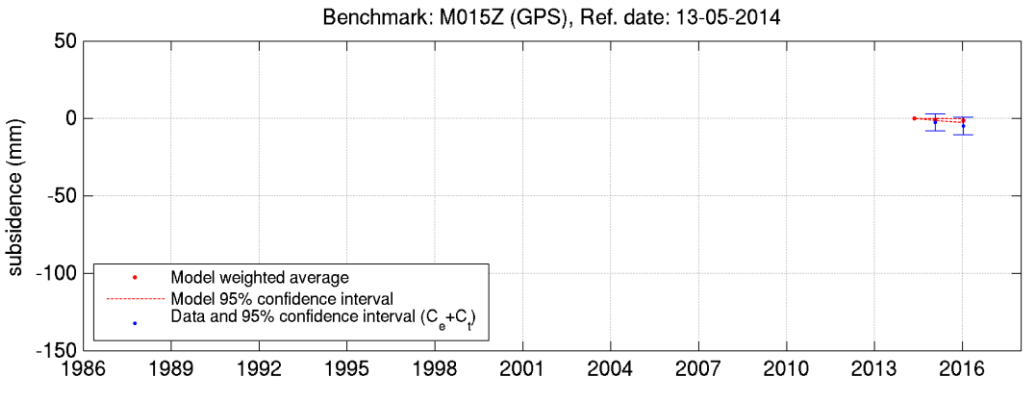

x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

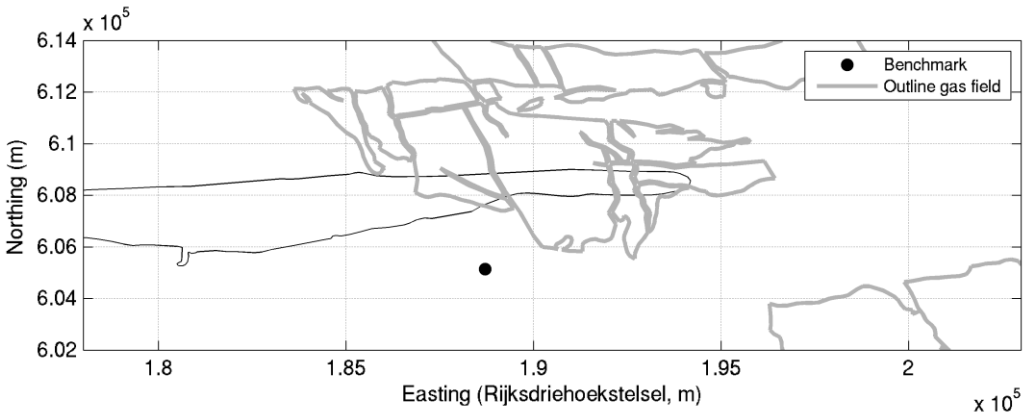

Benchmark: M015M (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)


x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field


Benchmark: M015N (GPS), Ref. date: 26-06-2009
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)


x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field


Benchmark: M015Z (GPS), Ref. date: 13-05-2014
subsidence (mm)
50
0
-50
-100
-150
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
Model weighted average
Model 95% confidence interval
Data and 95% confidence interval (C_e+C_t)


x 10^5
Northing (m)
6.14
6.12
6.1
6.08
6.06
6.04
6.02
1.8
1.85
1.9
1.95
2
Easting (Rijksdriehoekstelsel, m)
x 10^5
Benchmark
Outline gas field

Appendix 3. On the visual interpretation of confidence intervals

For individual parameters, visualisation of their expectable variation is straightforward and can be supported by tools like error bars, box plots or histograms. However, when multiple, eventually correlated parameters are considered at the same time, visualisation involves some pitfalls to be aware of. This applies in particular to some LTS-II results shown in this report, where variance-covariance matrices are used to characterise the expectable variation of double-difference observations. These matrices are generally fully populated, which implies that covariances and thus correlations are nonzero.

Higher-dimensional confidence regions

An issue that arises for the interpretation of any multi-dimensional parameter set can be demonstrated with a simple example. Suppose that X and Y are two random variables with standard deviations $\sigma_X = \sigma_Y = 1$ and a correlation coefficient of $\rho = 0.8$. This implies a covariance of $\sigma_{XY} = \rho\sigma_X\sigma_Y = 0.8$, and the corresponding covariance matrix is:

$$\mathrm{D}\left\{\begin{pmatrix} X \\ Y \end{pmatrix}\right\} = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix},$$

where $\mathrm{D}\{\cdot\}$ denotes the dispersion operator. If both parameters follow a normal distribution, the two-dimensional probability density function is shown in Figure 81.

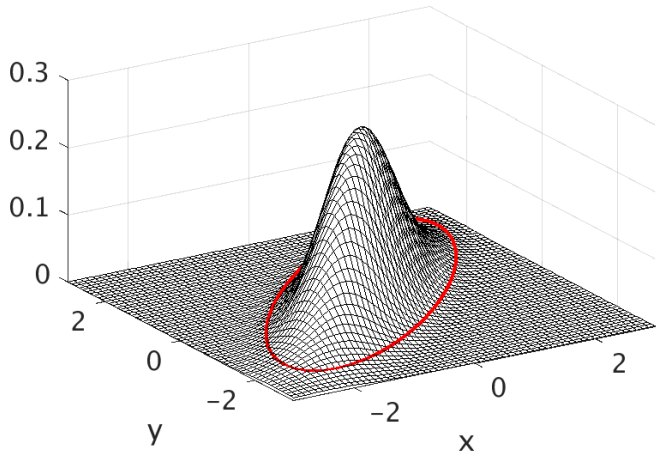


Figure 81. Probability density function of x and y with the 95 % confidence region indicated in red.

A confidence region is any area in in the XY-plane, in which the tuple (X,Y) is located with a certain likelihood. By additionally minimising that area, the shape of confidence region becomes a horizontal cross-section of the probability density function, which is an ellipse in case of normally distributed parameters.

Two-dimensional confidence regions or ellipses, respectively, are an adequate way to visualise the expectable variation of correlated parameters. However, the consideration of only two parameters at a time is often not sufficiently insightful. But if multiple parameters are visualised in one graph, like LTS-II double-differences in time series or profiles, their expectable variation can only be visualised in isolation.

Figure 82 shows a simulation of 50 000 realisations of X and Y represented by blue dots. 95 % of these dots lie within the red confidence ellipse. However, when projecting the ellipse on the x-axis and defining a one-dimensional confidence interval by the extent of the projected ellipse (dashed vertical lines), then 98.6 % of all dots fall into that interval. On the other hand, when defining a one-dimensional confidence interval that contains 95 % of all dots (continuous vertical lines), these are not the same 95 % that fall within the two-dimensional confidence region.

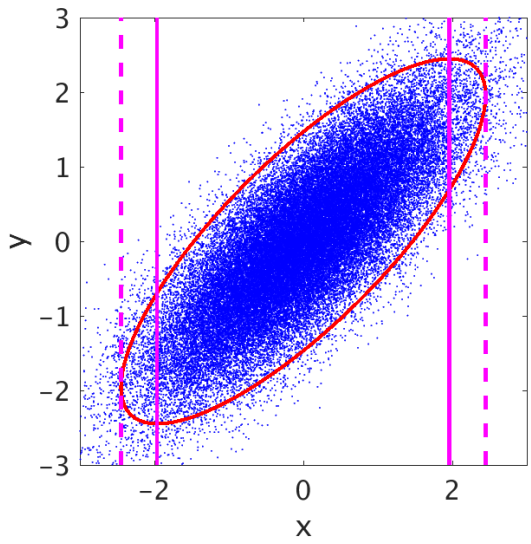


Figure 82. Simulation with 50 000 realisations of X and Y.

Note that this is a specific simulation for two dimensions (i. e., two random variables/parameters). For three parameters, the confidence region would have the shape of an ellipsoid. For more parameters, this would become a hyper-ellipsoid, which cannot be visualised anymore. When increasing the number of parameters, the following two questions arise:

A. What percentage of realisations would contain an n-dimensional hyper-ellipsoid that fits into the one-dimensional 95 % confidence intervals of all the n dimensions?
B. What percentage of realisations would contain a one-dimensional confidence interval that is defined as a projection of the n-dimensional 95 % hyper-ellipsoid onto that one dimension (delimited by the dashed lines in Figure 82)?

In answer to these questions, Figure 83 shows that the percentages quickly approach 0 (question A) and 100 (question B), respectively, if the number of parameters increases.

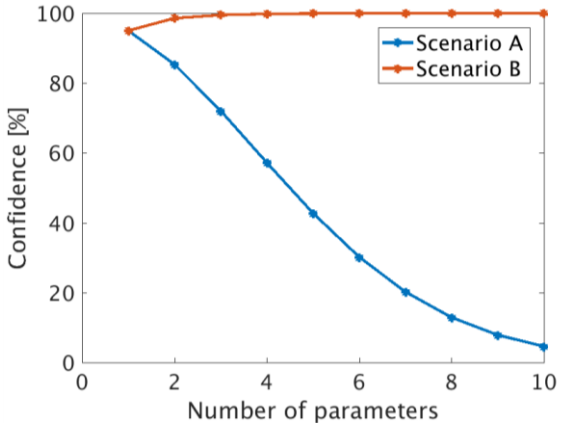


Figure 83. Size of confidence intervals/regions as a function of the number of parameters for the two scenarios described in the text.

Considering this simple example can help with the interpretation of visual comparison of models results with geodetic data. In LTS-II, the number of double-differences, which equals the dimension of the covariance matrix, is 1417. This means that

(A) not all models that fall within the confidence interval of an individual double-difference need to be accepted models.
(B) Likewise, a model that is accepted based on the n-dimensional test statistic does not necessarily have to fall within the one-dimensional confidence interval of an individual double-difference.

Adjusting the confidence intervals in a way that all accepted models fall within would increase their size dramatically and the informative value would be nullified.

Note that this effect is independent of the degree of correlation; it also applies to uncorrelated parameters.

Conditional probabilities for correlated parameters

Interpreting graphs based on correlated data involves another pitfall. Suppose that X and Y in the example above represent two double-difference variables from the LTS-II study visualised in a profile plot. Let I_X and I_Y be the 95% confidence intervals for X and Y, represented by error bars. An intuitive (but false) reasoning would be: If X is outside the 95 % confidence interval, then the probability that Y is also outside is still 0.05. Hence, the probability that both X and Y are outside the 95% confidence interval is 0.05·0.05 = 0.0025:

$$p\{Y \notin I_Y / X \notin I_X\} = 0.05 \qquad p\{X \notin I_X \wedge Y \notin I_Y\} = 0.0025.$$

However, this is only true for uncorrelated variables. If X and Y are correlated, these probabilities evolve differently. This is displayed in Figure 84.

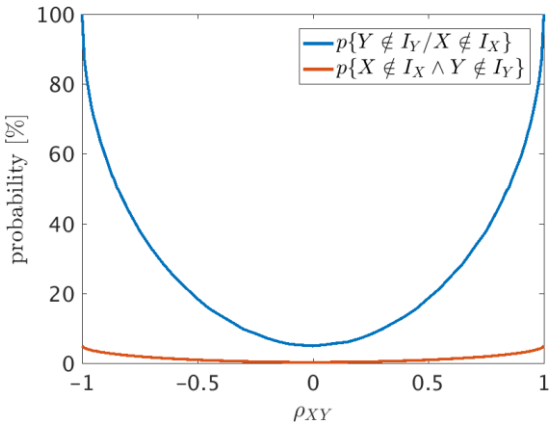


Figure 84. Conditional probability $p\{Y \notin I_Y / X \notin I_X\}$ and joint probability $p\{X \notin I_X \wedge Y \notin I_Y\}$ for two random variables X and Y and their corresponding 95 % confidence intervals I_X and I_Y, respectively, depending on the correlation coefficient ρ_{XY}. The blue line indicates the probability that Y is outside its confidence interval I_Y under the precondition that X is outside its confidence interval I_X. The red line shows the probability for both X and Y being outside their respective confidence intervals.

From the graph, it can be observed that if the correlation coefficient between X and Y increases from zero to one and given that X is outside the 95 % confidence interval, then the likelihood of Y also being outside the 95% confidence interval increases quickly.

To investigate the implications of this effect for LTS-II, normally distributed measurement and idealisation noise is assumed. The idealisation noise is computed with the parameters ON_C (see section 5.3), which includes both a temporal and a spatio-temporal component.

The degree of correlation of levelling double-differences in LTS-II primarily depends on separation in space and separation in time. This can be seen in Figure 85, where pairings of double-differences with common reference benchmarks and epochs are considered.

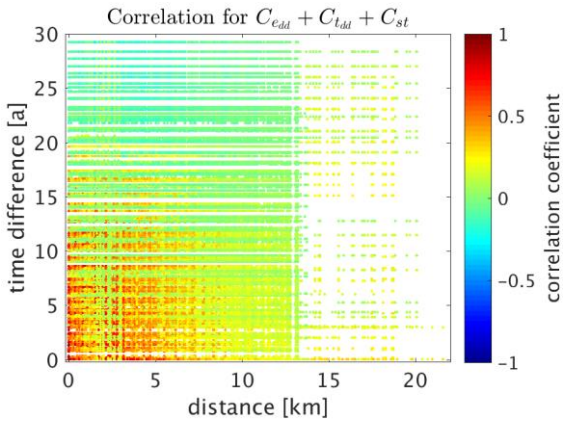


Figure 85. Correlation of all pairings of levelling double-differences used in LTS-II, for which reference benchmark and reference epoch are identical. Consider for instance two double-differences (P_{ref},P_A,t_{ref},t_1) and (P_{ref},P_B,t_{ref},t_2). For these, "distance" is the spatial distance between P_A and P_B, and "time difference" is the time interval between t_1 and t_2.

Especially when time difference or spatial distance are zero, the correlation can be very high. As a matter of fact, this is the case for most visualisation plots. In time series plots for a single benchmark, the spatial distance between double-differences is per definition zero. Profile plots visualise a single epoch, for which the time difference of all double-differences is likewise zero.

Consider for instance the profile plot in Figure 86, for which also the correlation with a selected double-difference (DD_{11}) is shown. All double-differences are significantly correlated, whereas the degree of correlation decreases with distance. Suppose now that DD_{11} as obtained from subsidence modelling would be outside its 95 % confidence interval I_{11}:

$$p\{DD_{11} \notin I_{11}\} = 0.05.$$

Then, the conditional probability that DD_{10} and DD_{12} are outside their confidence intervals as well is:

$$p\{DD_k \notin I_k, k \in \{10,12\}/DD_{11} \notin I_{11}\} = 0.40.$$

Based on the same precondition (DD_{11} being outside its confidence interval), the probability that DD_9, DD_{10}, DD_{12} and DD_{13} are outside is:

$$p\{DD_k \notin I_k, k \in \{9,10,12,13\}/DD_{11} \notin I_{11}\} = 0.11.$$

The probability that all other double-differences (DD_1-DD_{13}) are outside their confidence intervals is still:

$$p\{DD_k \notin I_k, k \in \{1 \dots 10,12,13\}/DD_{11} \notin I_{11}\} = 0.0006.$$

and thus small but a lot larger than it would be for uncorrelated parameters:

$$p\{DD_k \notin I_k, k \in \{1 \dots 10,12,13\}/DD_{11} \notin I_{11}, \text{all DD uncorrelated}\} = 0.05^{12} = 2.4 \cdot 10^{-16}.$$

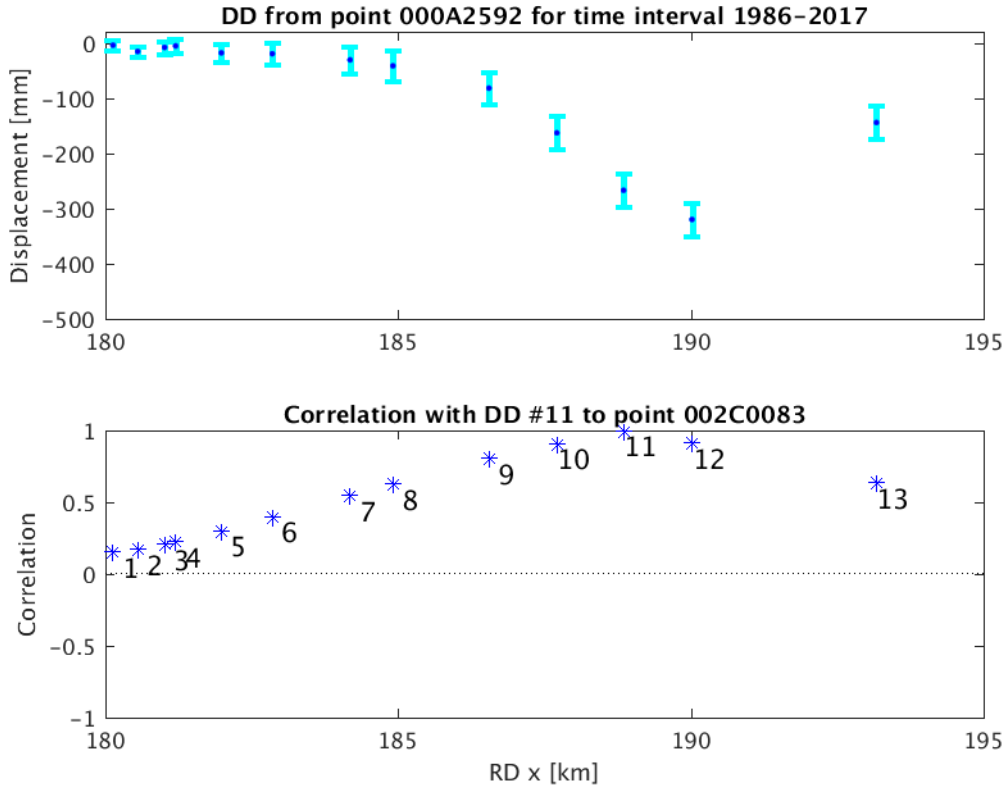


Figure 86. Profile of double-differences with respect to benchmark 000A2592 and the time interval 1986-2017. Some of these double-differences are already contained by the original set obtained from CUPiDO, some other may have been transformed to match the common reference benchmark and epoch. Upper plot: relative displacement with 95 % confidence intervals. Lower plot: correlation with double-difference no. 11 (000A2592-002C0083, 1986-2017), which has a correlation of 1.0 with itself.

The lesson from these considerations for the interpretation of LTS-II results is that a common assumption does not hold for correlated data. All 95 % confidence intervals shown in one plot do not necessarily need to cover 95 % of the subsidence predictions. This applies to both profile plots (where all double-differences are correlated due to temporal congruence) and time series plots (where all double-differences are correlated due to spatial congruence).

Looking at the general picture

An inherent limitation of visualising the confrontation results in two-dimensional plots is that a single plot can never convey all information. The goodness-of-fit of the ensemble members is assessed globally, meaning that there are usually many regions where the fit is excellent or reasonable but also some regions where the fit is poor. A single time series plot or profile plot with a poor fit does not allow for the conclusion that the overall fit is poor. A global assessment is done by considering all double-differences at the same time. This requirement is satisfied by the χ^2 test statistic.

Conclusions

Care must be taken with the interpretation of confidence intervals for multi-parametric models. It can be somewhat unintuitive. Global acceptance of the whole model is not properly reflected in the confidence interval of a single parameter. If parameters are correlated, the likelihood that multiple parameters fall outside their error bars simultaneously is almost certain to be underestimated from looking at a two-dimensional plot only. In the confrontation between the subsidence model and the geodetic data and the computation of the test-statistic, all these correlations are properly considered. This needs to be kept in mind when interpreting plots showing the results from LTS-II.