Groningen long term subsidence forecast

NAM-report	EP202008201822, unrestricted
Version 0	Original version
Version 1	updated parameter values for model uncertainty
Version 2	Appendix B added describing the effect of multiple reservoir realisations on the model uncertainty

Summary

After the discovery of the Groningen gas field in 1959, land subsidence due to gas production was recognized as a potential threat to water management in the Province of Groningen. During the 60 years of gas production from the field, the cemented reservoir sandstone has compacted with tens of centimetres due to the decrease of gas pressure. This was caused by a transfer of the overburden weight from the pore pressure to the grains of the reservoir sandstone, resulting in both a linear elastic and non-linear (visco)-plastic compaction of the reservoir. The resulting land subsidence has been monitored frequently over the past 60 years by optical levelling measurements and, more recently, by InSAR and GNSS (Global Navigation Satellite System, GPS).

Moreover, the decrease of gas pressure and the resulting compaction has led to induced earthquakes. To mitigate seismicity induced by the gas production, the minister of Economic Affairs and Climate Policy decided to minimize and end production from the Groningen field as soon as feasible, expected mid-2022. This will stop further depletion of the main reservoir layer and hence subsidence of the field at large. However, it does not imply that further compaction and subsidence will cease immediately. Both visco-plastic compaction and pressure redistribution will lead to moderate amounts of further subsidence above the gas field. Another source for further compaction can come from pressure decline in the aquifers that are laterally connected to the gas field. Because there are no wells in these aquifers, the pressure cannot be measured directly and can only be estimated from calibration and/or inversion of the geodetic data.

A statistical method is adopted that investigates the compaction and subsidence in and above the Groningen gas field and its laterally connected aquifers with 4 major objectives: invert for the most probable aquifer realisation and a set of RTCiM values, test multiple correlations to the rock compressibility and calculate the model uncertainty. The method uses a Bayesian stochastic approach via Markov Chain Monte Carlo. The inversion for all parameters is a computer intensive process and not feasible for all possible subsidence scenarios. Therefore a 5-step workflow is set up to deliver the above described objectives at reasonable computation times. After the execution of the first four steps we retrieved:

- the most likely reservoir realisation,
- a set of spatially recalibrated Cm grids that still reflect the compressibility trends as shown by the prior Cm grids,
- RTCiM parameter values that temporarily and spatially best fit to the subsidence measurements,
- A quantification of the uncertainty driven by the model.

Using this subsidence scenario and its uncertainty, subsidence forecasts were made based on the latest production scenarios.

Contents

1 Introduction

The Groningen gas field is the largest gas field in Europe with a static GIIP of around 2900 BCM (Ref 2). The gas is being produced from the Permian Slochteren Formation (ROSL). A large part of the Groningen province lies below the sea level, an area that requires active water management. Significant compaction and subsidence due to gas production can be regarded as a threat to constant groundwater levels with possible consequences for land usage and infrastructure like dikes and bridges. Large additional subsidence volumes require additional efforts from water boards to preserve the usage of the land as it is today.
This subsidence threat has long been recognised, which resulted in many studies. These studies all had the objective to match a geomechanical model to the available geodetic data. The history matched model is used for forecasting the future subsidence.
The most recent publications on this topic can be found in Ref 3 and Ref 19.

Besides the impact on groundwater levels, reservoir compaction has also been correlated to the occurrence of induced seismicity (Ref 20). Although induced seismicity is mainly caused by poro-elastic effective stress changes, it is observed that the highest frequencies and magnitudes of earthquakes occur in areas where the compaction is highest.

1.1 Geological setting

The Groningen gas field is situated in the northern part of the Netherlands and produces gas from the Slochteren Sandstone (ROSL) reservoir that consists of a quartz-rich cemented sandstone. The sandstone is part of the Permian Rotliegend Group (RO), a clastic lithostratigraphic unit that is situated at a depth of around 2900m. Its thickness varies from 70m in the south-eastern part of the field to 240m in the northwest. The source of the gas originates form the Carboniferous Westphalian coals, while the seal is formed by shales from the Permian Ten Boer Member (ROCLT) and the evaporites from the Permian Zechstein Group (ZE).

The Groningen structure is characterised by a NNW-SSE trending intrabasinal high, formed mainly by normal faulting during the Mesozoic. It has an overall horst geometry, lying between the Ems Graben to the east and the Lauwerszee Trough to the west (Figure 1).

1.1.1 Study area

The study area is being defined by the Groningen gas field and the most important laterally connected aquifers surrounding the field (Figure 2). Of special interest is the “Southern Lauwerszee Trough (zuidelijke Lauwerszee Trog)” aquifer located between the smaller fields of Vries and Roden and Groningen gas field. The city of Groningen is located above this area. In the north, the “Mowensteert” aquifer is connected to the Groningen gas field causing possible subsidence in the Waddenzee. The Rysum aquifer is connected to the eastern part of the field, while the aquifer between Anneveen and Groningen is connected to the Groningen field in the south. The white coloured aquifers like the Goldhoorn aquifer to the east of the Groningen have no connection to the gas field due to large offset faults disallowing for lateral gas or fluid flow and hence pressure communication.

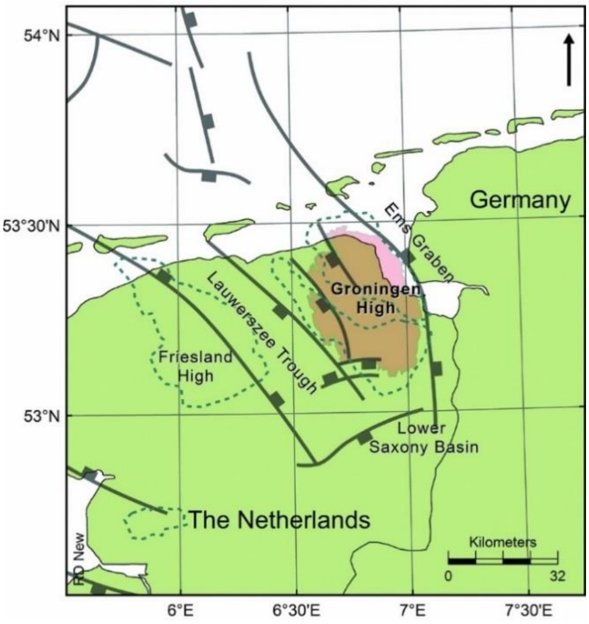


Figure 1; Regional setting and structural elements of the Groningen region. The Groningen gas field is indicated in red; green dashed lines indicate Dinantian carbonate platforms (Ref 4).

Figure 2; overview of the most important lateral aquifers, attached to the Groningen field. In red the field names that are mentioned in the text.

The Rotliegend layer is characterized by a dense fault network as noticeable in the semblance map that is shown in Figure 3.

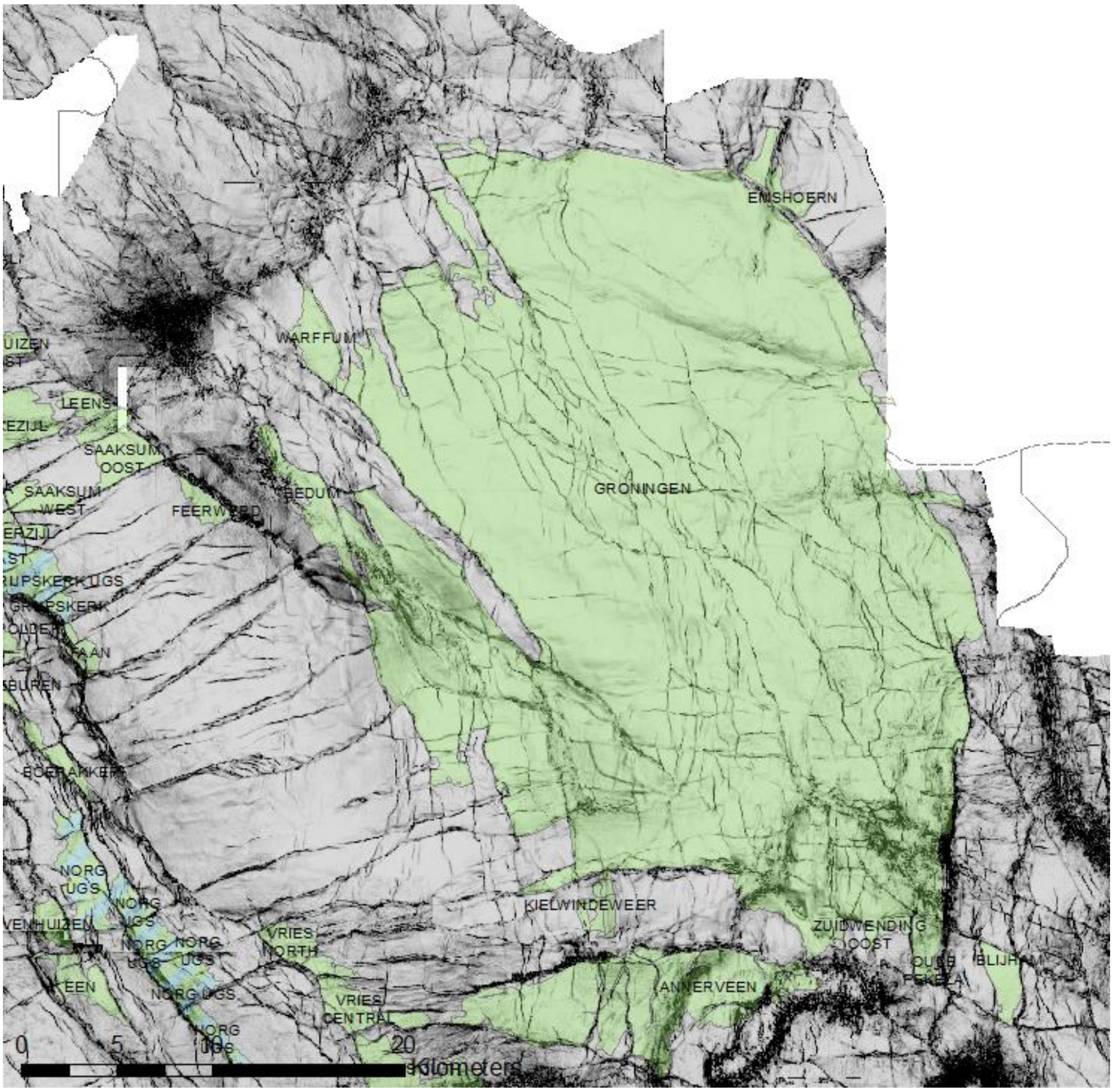


Figure 3; Groningen field and surrounding smaller fields projected on a seismic semblance map that reveals the faults at the top of the Rotliegend

1.2 Production history and forecast

The Groningen field produces gas since 1963. Its development involved around 300 wells and 26 production locations (of which 17 are still in production). The field recovery factor to date (Q3 2020) is around 76%. Figure 4 plots the entire gas production history for the Groningen field with the expected forecast according to the latest estimate “GTS-raming 2020 for an average year”. Since the earthquake in Huizinge in 2012 (of magnitude 3.6) production restrictions have been imposed. In September 2019 the Minister of Economic Affairs and Climate Policy decided to bring down Groningen production as soon as possible, with no significant production volumes expected after 2022 (Ref 9). Every year before 1st of October the Minister takes a “Vaststellingsbesluit” regarding production and operational strategy for the next gas-year[1]. It is expected that this new production

---

[1] A gas-year runs from 1st of October to 30th of September. The next gas-year therefore starts on 1-10-2020.

profile will be in line with the latest estimate advised by Gasunie Transport Services ("GTS-raming 2020").

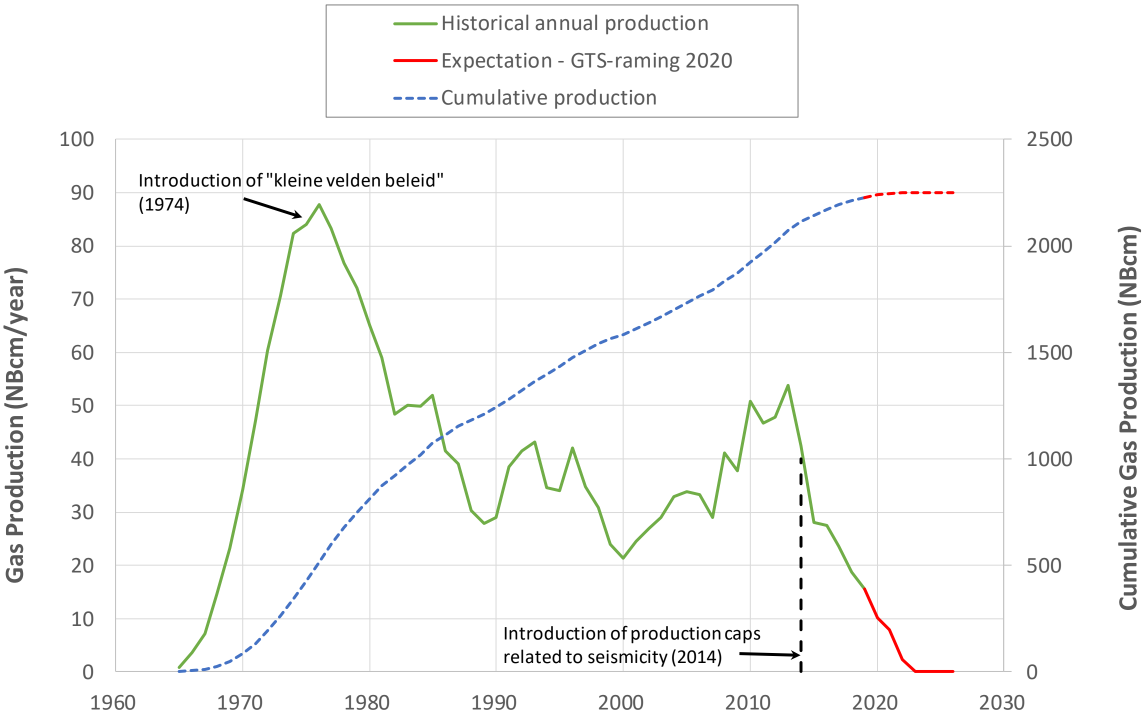


Figure 4; Historical and forecasted Groningen gas production with annotation of major events impacting production

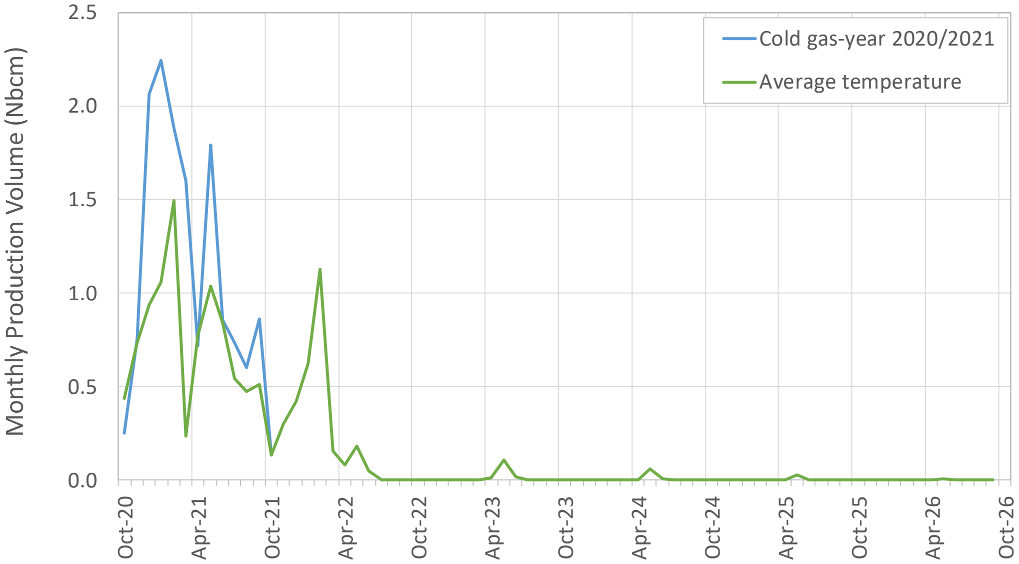


Figure 5; Monthly production profiles based on "GTS-raming" 2020 for average years (base case) as well as a cold gas-year scenario (sensitivity)

The demand for Groningen gas depends mainly on ambient temperature. GTS has provided Groningen gas demand profiles for average, cold, and warm reference years. The expectation profile in Figure 4 is based on an average temperature profile. This forecast is used as a base case for the subsidence forecast, with a cold gas-year 2020/2021 as a sensitivity. The monthly production profiles for these two scenarios (that only differ in gas-year 2020/2021) are plotted in Figure 5.

Two operational strategies have been proposed to the minister for the coming gas year (Ref 11). It is expected that operational strategy 2 will be selected, based on the minister's "Ontwerp-vaststellingsbesluit Groningen gasveld 2020-2021". Therefore, the distribution of production volume over the various clusters in the field is in line with this new strategy, in which clusters Bierum, Eemskanaal, and Siddeburen will be closed in from October 2020 onwards. The volume distribution over the clusters is handled in the forecast by a surface network model coupled to the MoReS dynamic reservoir model.

The latest dynamic reservoir model (V6) is used (as described in Ref 12). Within the boundaries of the Groningen field, the match between modelled and measured pressure (SPG data) is good, with a root-mean-square error of around 2 bar. For both the history-match and forecast period, annual reservoir pressure grids are exported from the reservoir model to be used in the subsidence model.

While there is ample pressure data available across the Groningen gas field, pressures in the surrounding aquifers are much more uncertain due to lack of pressure data. The connected aquifers (see Figure 2) are to a large extent covered by the Groningen reservoir model, but the V6 model only gives a single pressure realisation (constrained by measured pressures in the gas field). In order to obtain a range of pressure forecasts for the aquifers covering a larger uncertainty, additional Groningen reservoir model versions are used.

The model strategy for the Lauwerszee aquifer is different The V6 model includes pseudo-aquifer wells to mimic aquifer depletion in the Lauwerszee aquifer due to the gas fields to the West (Roden, Vries, Pasop and Faan) but this gives only one possible aquifer realisation and the model does not extend all the way to the neighbouring gas fields where pressure data exists. Therefore, box-models have been constructed to cover the area across the Lauwerszee aquifer between the Groningen gas field and the gas fields to the West. The aquifer pressure realisations are further described in Section 3.2.

1.3 Brief overview of subsidence forecast in the past

The first forecasts of the subsidence above the Groningen field originate from 1971 with a maximum value of 1 meter. These predictions were solely based on oedometer core sample tests. Nowadays the oedometer tests have been replaced by uni-axial pore pressure depletion tests that give more reliable and lower estimates of the rock compressibility. In later years, geodetic levelling campaigns collected the first subsidence data and these showed values that were much lower than forecasted by the models in 1971. Calibration of a linear elastic compaction model to these data led to a new forecast of the subsidence at the end of field life. This model predicted a maximum of only 30 cm, less than half of the previous forecast. In 1985 a new compaction model, the rate type compaction model, doubled this number to a maximum value of 65 cm (Ref 14). The parameter values of the RTCM were mainly based on core experiments. Subsequent geodetic measurements showed values again lower than predicted by the RTCM and Ref 17 recommended to restore the linear relationship between depletion and compaction for the Groningen gas field, leading again to lower values of the forecasted subsidence.

Mainly based on the observation of delayed subsidence above the Ameland field, NAM developed a new diffusional compaction model called the "time-decay" model (Ref 6). This model was also used for calculating the compaction in the Dutch fields, documented in the 2015 report on the North Netherland subsidence (Ref 18).

The RTCM was reintroduced by the LTS study on the Ameland subsidence (Ref 23) and it was decided after 2015 to use the same compaction model for Groningen compaction. Subsidence forecasts for the Groningen field using a RTCM were firstly published in the 2016 plan (Ref 15). After the ministerial decision in 2018 to minimize and cease production from the field as soon as possible, a new forecast was published in 2020 reflecting the latest GTS forecast (Ref 8). This scenario was based on a stop of production in 2030. Contour plots of the subsidence in 2030 from Ref 8 can be found in Figure 6.

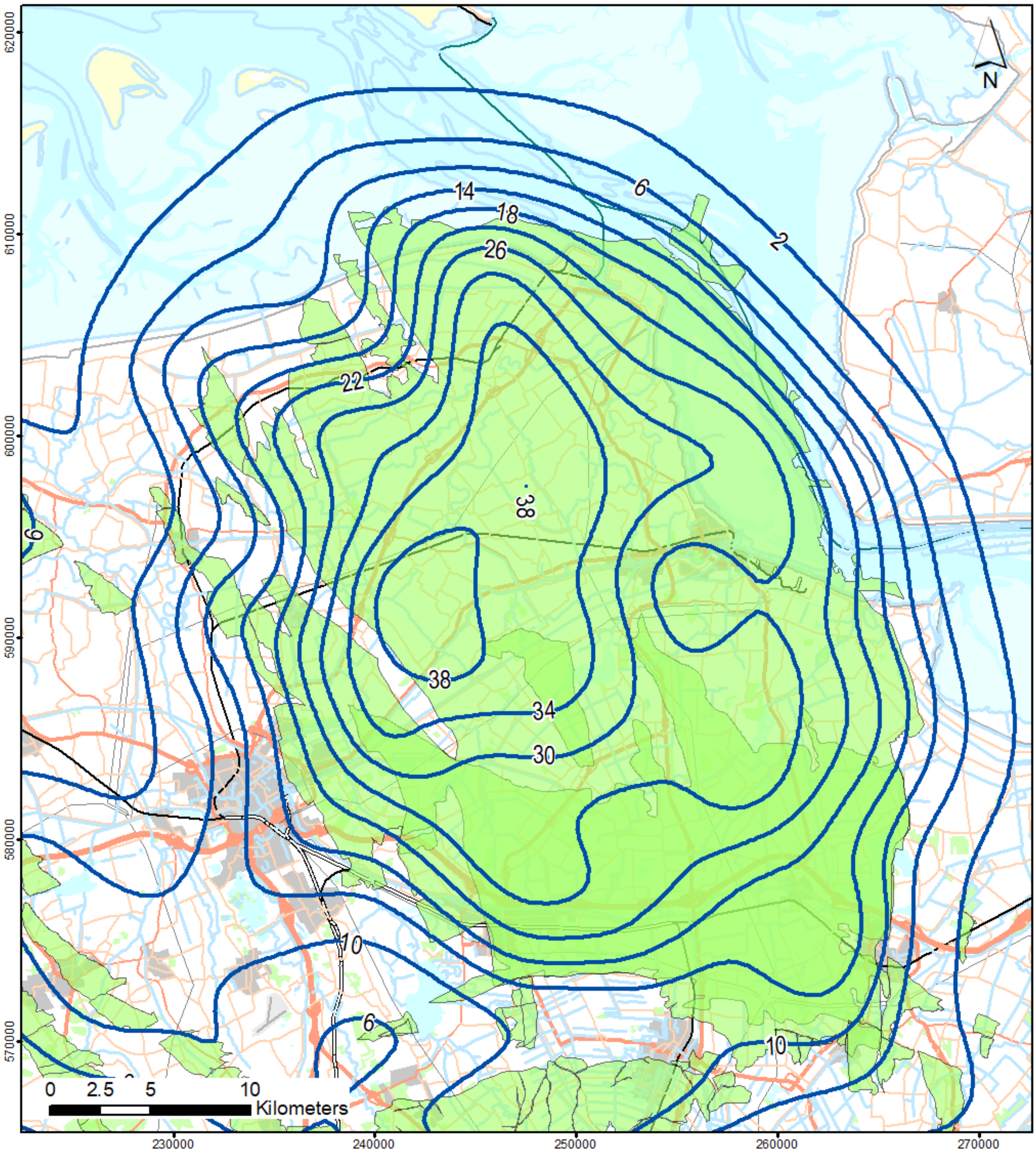


Figure 6; subsidence forecasts as presented in (Ref 8) for the year 2030.

2 Methodology

There are multiple possibilities to calibrate a model to measured data. In the past, deterministic models were manually fitted to the data, without a quantification of the uncertainty. A stochastic approach would circumvent this drawback where data is used to find both the optimal match and its uncertainty. Ref 23 describes a statistical workflow were subsidence model scenarios are selected based on the goodness-of-fit to the geodetic data using a Bayesian approach. Fast analytical models were used that enabled the investigation of around 1 million possible scenarios. Similar fast models are also used in the statistical methodology for forecasting the subsidence above the Groningen field. This method is described in detail by Ref 1.

This chapter provides a summary of this work and describes in more detail the implementation of the workflow for the Groningen gas field. The method is based on Bayesian inference with the aim to estimate model parameters of the RTCiM model (see section 3.4.1 for more information on this compaction model) as well as the error other than measurement noise. The error that is defined as Σ_{emp} in Ref 1 can be regarded as the variability due to the misfit of the model to the geodetic data and has a simple variance structure. The covariance matrix Σ_{emp} is calculated with only having non-zero values on the diagonal of the matrix and uses the following equation to describe the error at a given location:

$$\Sigma_{emp} = \sigma_0 + \alpha\mu_i$$

where μ_i is the modelled displacement (double difference) for location i estimated from the statistical method. The parameters σ_0 and α are estimated by the method to be the same over all locations. This means that the Σ_{emp} increases with higher values for the modelled displacement, provided that a positive value of alfa is estimated.

The variability of the geodetic data (Σ_{geod}) has a defined, more complex, variance-covariance matrix of errors. For a description of the geodetic data, see section 3.1.The total error of the model is described by Σ_{comp} which is the sum of the aforementioned error components. Based on this error model, the probability is calculated using a defined set of model parameters that is confronted to or matched with the data. A Metropolis-Hastings algorithm, which is a specific form of a Markov Chain Monte Carlo, is used to propose different sets of parameters and to reject and accept these based on their likelihood.

The benefit of this statistical workflow is the formal quantification of the Σ_{emp} and when summed to the Σ_{geod} also the quantification of the prediction interval. This workflow is therefore different than the workflow that was adapted for the Long-Term Subsidence study in the Wadden Sea (Ref 23) where the model uncertainty was estimated a-priori. More details on the methodology and defined uncertainty parameters can be found in Ref 1.

Next to the uncertainty of the compaction model and its parameters, another source of uncertainty is the pressure evolution in the gas field and laterally connected aquifers.

For the gas field, after almost 60 years of production, the pressure depletion is well known. Different vintages of reservoir models were history matched using a large set of pressure and production data coming from hundreds of wells that penetrate the reservoir. However, the pressure depletion of connected aquifers is less well understood and constrained, because of the lack of well penetrations and therefore pressure measurements in the aquifers. This condition requires a likelihood metric for the pressure depletion realisations in the aquifers, which is provided by the statistic methodology. Multiple aquifer realisations with different pressure depletions are therefore considered, combined with the pressure history in the gas field that is history matched to the available data by the Groningen V6 reservoir model.

In the Southwest, the area between Groningen and Vries-Roden-Pasop, aquifer models were built (five models, see also Figure 15) which were history matched on the boundaries of the model, to the west on Vries, Roden, Pasop and to the east on the Groningen gas field using pressure information. These models are called “box-models” in this document. For the other aquifers, to the north, east and south, different vintage of reservoir models with a large variation in pressure realisations were used. Details on the box-models and the number of realisations can be found in respectively sections 3.2.4 and 4.2.

It needs to be noted that the reservoir model of Groningen (and more specifically the latest V6 model) calculates both the pressures in the Permian (Rotliegend) and the Westphalian (Carboniferous) sandstones. Since only a small fraction of the gas is produced from the Carboniferous, only one pressure realisation (the V6 pressure realisation) for this formation is considered in all calculations, i.e. only the uncertainty of Rotliegend depletion scenarios is investigated. The statistical workflow is applied to each Rotliegend depletion realisation for the derivation of both the RTCiM parameters as well as the Σ_{emp}.

The statistical methodology is embedded in a workflow that is designed to retrieve most information at acceptable computational time.
This workflow was designed around the following questions:

a) What physical/reservoir parameters can be used best to create prior spatial Cm maps, i.e. which parameters are well correlated with the compaction coefficient? (A uniform compaction grid can be used to test the added value of prior subsurface information).
b) What is the most likely aquifer realisation using the geodetic data?
c) How to reduce/minimise local (spatial) residuals between model and geodetic data?
d) Once the best spatial match to the geodetic data is achieved, how to improve the temporal match?

To address these questions, a workflow was created consisting of the steps described below and visualised in Figure 7. More details for each step will be provided in Chapter 4.

1. Generation of the prior Cm grids and calculate a first set of RTCiM and uncertainty parameters:
    a. Correlation Cm - porosity. The porosity is used as provided in the V6 reservoir model
    b. Correlation Cm - slowness. A velocity map (3D seismic cube) is used to estimate Cm, applying a relation based on compaction monitoring results in the ZRP-3 well, where a distributed strain sensing cable is installed.
    c. No correlation, assuming a uniform Cm for the whole model area.

    For these spatial maps the most likely RTCiM parameters and uncertainty parameters are found using the MCMC method converging to the scenario with the lowest value for the negative log likelihood (Ref 1). An important characteristic of the applied calibration method is that model uncertainty parameters can be explicitly estimated. To obtain optimal estimates of the model uncertainty, it is key that the uncertainty in the geodetic data is correctly captured in the (full) covariance matrix. In this study the levelling data are directly used in the calibration workflow, because of the large history since the start of production in the 1960s and the availability of a proper uncertainty description. For validation, the results of the calibration are subsequently compared to the InSAR and GNSS data. Details on the geodetic input will be provided in section 3.1. In step 1, only benchmarks above the gas field are used in the calibration, because of the well constrained pressure input in the workflow coming from the V6 model.

2. Using the obtained results for the uncertainty (Σ_{emp}) and RTCiM parameters from step 1, the objective of step 2 is to test the aquifer realisations in combination with the spatial Cm grids. Combining the defined box-models and legacy reservoir models, all possible aquifer realisations are tested to obtain a value for the negative log likelihood. The computational effort is relatively small because of the fixed values for the (Σ_{emp}) and RTCiM parameters. A pressure realisation consists of the gas pressure part from the V6 reservoir model (see section 3.2), the V6 pressure realisation for the DC and a combination of possible ROSL aquifer realisations for the different aquifer areas. The benchmarks above the aquifers are used to assess the subsidence fit between model and measurements. The goodness of fit is expressed by the value for the negative log likelihood. The best realisation for each of the possible spatial Cm grids is selected, resulting in three subsidence scenarios.
3. To reduce local higher residuals in some of the areas in the three subsidence scenarios, step 3 uses an inversion scheme to give more weight to the double differences in these areas. The spatial Cm maps of step 1 are used as a prior input into the inversion scheme. Step 3 results in 3 new spatial compressibility maps tailored to each of the three subsidence scenarios.
4. In step 3, local spatial mismatches are resolved. In step 4 attention is given to the temporal signal from the measurements including a new assessment of the Σ_{emp}. New values for the RTCiM model are derived after the application of the MCMC statistical workflow and using the new spatial Cm maps from step 3.
5. Step 5 executes the forward calculation to obtain the forecast for the most likely subsidence scenario including its uncertainty. The estimated parameters for the Σ_{emp} provide the uncertainty of the forecast, i.e. the confidence interval.

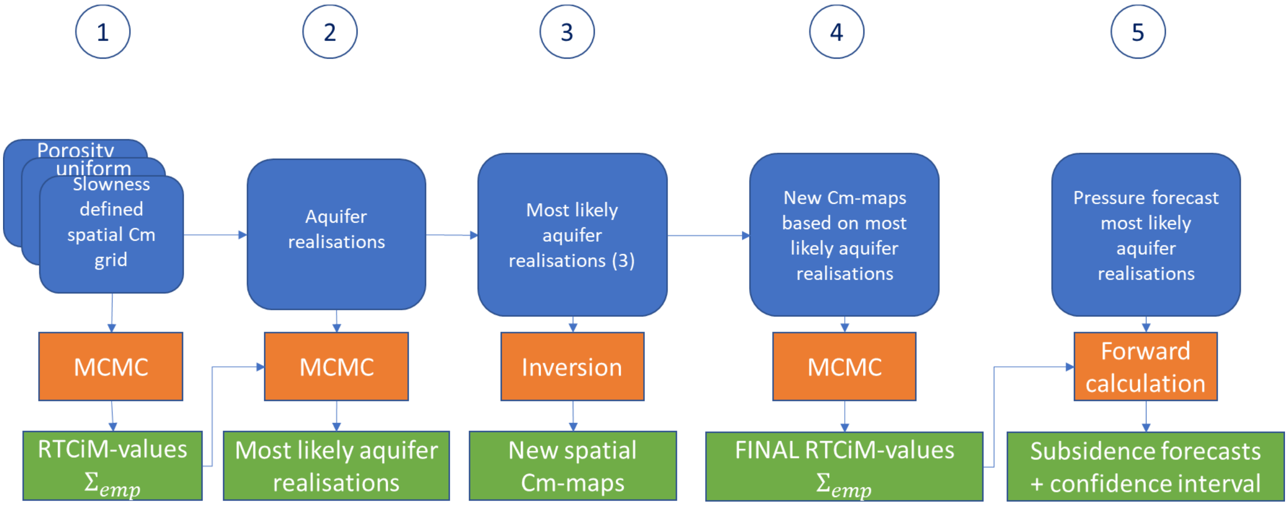


Figure 7; Scheme of the 5-step workflow. Input of each step in blue, the orange boxes visualize the calculation method and in green, the results per step.

3 General input to the model

3.1 Geodetic input

As explained in chapter 2, the levelling data have been directly used in the calibration workflow. The calibration results have subsequently been validated with InSAR and GNSS data.

Figure 8 shows the selected benchmarks from 18 levelling campaigns spanning from 1964 to 2018. After network adjustment of the individual campaigns, spatio-temporal double-differences have been formed to serve as input to the modelling workflow. This subsection describes the preparation of the geodetic dataset, addressing stochastic modelling (3.1.1), outlier handling (3.1.2) and formation of double differences (3.1.3).

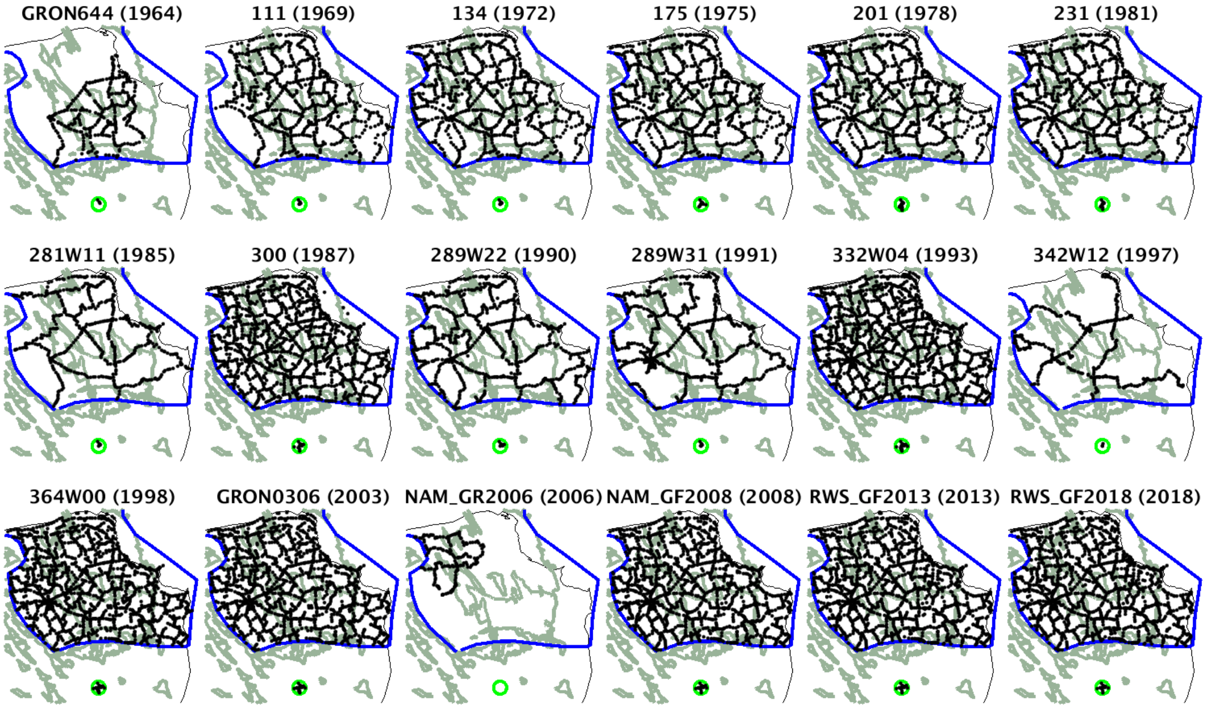


Figure 8; Levelling campaigns that have been used in this study. From the data inside the blue polygon further sub selections have been made to exclude areas affected by unmodelled processes like salt mining. Where possible, also some benchmarks close to the village of Gasselte (green circle) have been added to serve as convenient reference point in spatial differencing.

3.1.1 Stochastic modelling

The uncertainty of geodetic measurements is described in terms of a variance-covariance matrix. This approach has been proposed in the Long Term Subsidence study of the Wadden Sea region (LTS-I; Ref 25) and successfully applied in its follow-up studies (LTS-II; Ref 23). The stochastic model comprises two components: measurement noise and idealisation noise.

Measurement noise is the uncertainty of the measurement itself, resulting from the limited precision of determining height differences between perfectly defined benchmarks. For this component, a common parametric model has been applied to the average from forth and back levelling of a section:

$$\sigma_{\text{meas}}^2 = \begin{cases} a^2, & s < (a/b)^2 \\ b^2 \cdot s, & s \geq (a/b)^2 \end{cases},$$

where s is the levelled distance, a = 0.1 mm and b (in $\mathrm{mm}/\sqrt{km}$) has been estimated from the misclosures of the individual campaigns. Typically, b is between 0.75 and 1.00 $\mathrm{mm}/\sqrt{km}$. Measurement noise has been propagated through network adjustment, resulting in one fully populated covariance matrix per campaign.

Idealisation noise describes displacements of the measurement points due to shallow movements like building settlement or soil compaction. In the LTS-studies, two components have been discriminated (Ref 23):

- *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 where 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 of the shallow subsurface to ground water level changes.

Whilst the significance of the temporal component is evident, the existence of a spatio-temporal component is arguable. In LTS-II, the latter has been calibrated in a mining-free area in the province Friesland. These parameters however overestimated the idealisation noise in the western part of Ameland, from where no gas is produced. Hence, for Ameland, the spatio-temporal component has been included in LTS-II with the rationale to also partly account for geomechanical model imperfections. In the methodology of the current workflow, however, this is not desirable because the model imperfection is explicitly estimated in the workflow. Therefore, the modelled idealisation noise is limited to the temporal component.

As in LTS-II, the parametric model of Houtenbos and Kenselaar (2001; Ref 26) has been applied. The contribution of idealisation noise to the variance of a spatio-temporal double-difference reads:

$$\sigma_{\text{ideal}}^2 = 2\sigma_t^2 \Delta t^{p_t} \ ,$$

where Δt is the time difference, σ_t^2 = 0.16 $\mathrm{mm}/\Delta t^{p_t}$ and p_t = 1.24. Computation of covariances is described in detail in Ref 27.

3.1.2 Outlier handling

Outliers in geodetic data must be addressed to prevent contamination of the study result. In LTS-II (Ref 23), a relatively high level of outlier removal was necessary, because the methodology was very sensitive to outliers. The risk of a biased outcome could sufficiently be controlled by only allowing spatio-temporal smoothness as a prior assumption and showing visually that data adaptions in critical areas were limited.

In this study, outlier handling is less critical due to the wealth of available data and the lower susceptibility of the methodology to outliers by introducing model uncertainty in the calibration workflow. Therefore, outlier identification can be less strict (and thus more conservative) than in LTS-II. The approach pursued here is based on the recent stability analysis from 2019 (Ref 28), where benchmarks are subdivided into three classes: *assumedly stable*, *assumedly partly stable* and *assumedly not stable*. “Stable” means in this context that the associated measurements are reliable, i. e., they are neither suffering from autonomous benchmark movement nor are they subject to identification errors or disturbances. While “stable” benchmarks are included and “unstable” benchmarks are excluded from the dataset, “partly stable” benchmarks are handled according to Figure 9.

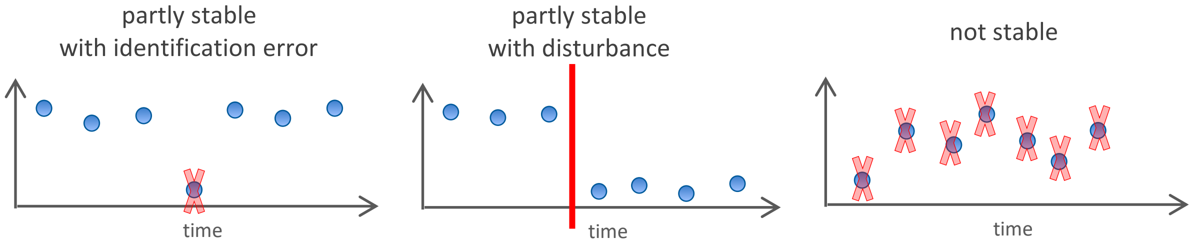


Figure 9; Examples for time series from benchmarks classified as partly stable or not stable. Identification errors are removed, and time series are split in two at disturbances. From benchmarks classified as not stable, all observations are removed.

Like in LTS-II (Ref 23), another less sensitive outlier detection approach, first proposed in LTS-I (Ref 25), has been applied as a redundant additional check: a prior spatio-temporal model (here: a geodetic model based on smoothness constraints) is subtracted from time series of height differences between pairs of benchmarks. The resulting residuals are tested in a geodetic framework. A largely relaxed threshold accounts for the uncertainty of the potentially imperfect model assumptions. There were no additional outliers found by this approach.

Finally, some auxiliary analysis was applied to reveal the most likely candidates of still undetected outliers. The time series of the benchmarks in question were inspected visually. As a result, one single benchmark was additionally excluded from the dataset.

3.1.3 Differencing

With the adjusted levelling networks, the corresponding covariance matrices and the idealisation noise model as input, the software CUPiDO (Connecting Undifferenced Points in Deformation Observations; Ref 29) has been used to create double difference observations by differencing the spatial height differences from the levelling networks in time. To better facilitate time series interpretation, CUPiDO has been steered such that the underground benchmark 000A2080 in Gasselte is selected as reference benchmark for as many double differences as possible (84 %). For the rest, two other reference benchmarks outside the sphere of influence of gas production have been chosen. Theoretically, the situation in a stable area is not a requirement, but since the stochastic model is not perfect, choosing a reference point with zero modelled subsidence mitigates residual biases.

Figure 10 visualises the standard deviations of all double differences. In this plot, a couple of effects can be observed:

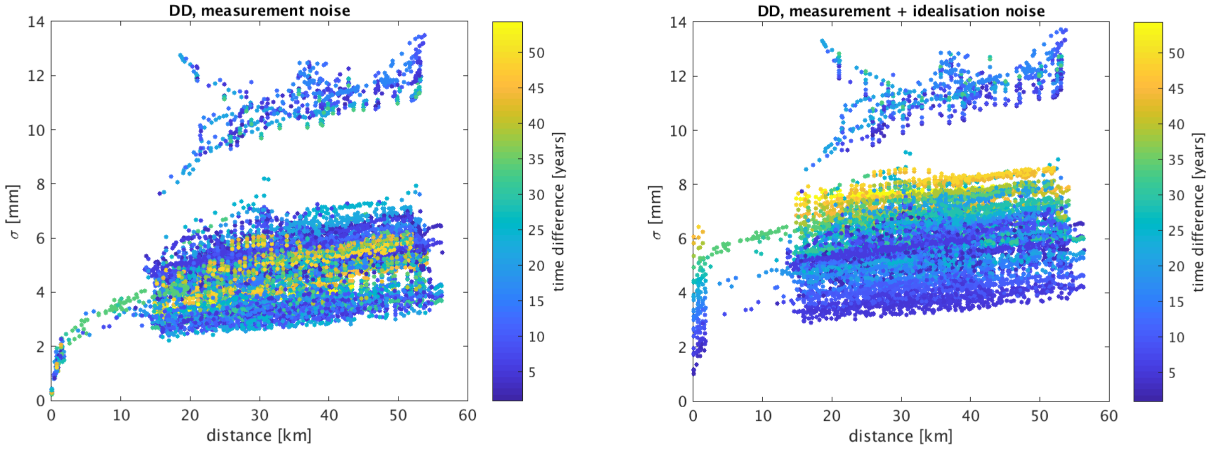


Figure 10; Standard deviations of double differences. Left: only measurement noise, right: measurement noise plus idealisation noise.

- In accordance with the measurement noise model, standard deviations grow with the square root of distance between benchmarks.
- In accordance with the idealisation noise model, standard deviations grow with the time between the two differenced campaigns.
- There are only few double differences spanning spatial distances between 2 km and 15 km. This can be explained with the isolated localisation of the underground benchmark 000A2080 in Gasselte (green circle in Figure 8), which serves as reference benchmark for most of the double differences.
- There is an isolated group of double differences with standard deviations above 8...10 mm. These are all differenced with respect to campaign 281W11 (1985). For this campaign, a relatively high level of measurement noise was modelled (b = $1.60\ \mathrm{mm}/\sqrt{km}$). This was considered the most appropriate way to cope with an unusually high number of loop misclosures in the observation data that were still too small to be considered outliers.

3.2 Description of pressure scenarios for the various lateral connected aquifers

For the pressure prediction within the boundaries of the Groningen gas field the V6 reservoir model (Ref 12) is used, based on the latest production data and production forecast as described in Section1.2. The full field model is also used to calculate pressures in the Möwensteert, Rysum, and South aquifers that are connected to the Groningen gas field (Figure 2). To reflect the larger uncertainty in the aquifer pressures, a range of pressure forecasts is obtained for the aquifers using older versions of the Groningen model.

Table 1 gives an overview of the model versions used and their key features.

Table 1; Overview of Groningen full field reservoir model versions used in this subsidence study

Model	**GFR2013**	**V2.5**	**V4**	**V5**	**V6**
Versions used	Low/Base/High	Low/Base/High	Base case	Base case	Base case
Main original use	ARPR 2013 WP2013	HRA Nov 2015 WP2016	HRA June 2017	HRA Nov 2017 HRA 2018 HRA 2019	HRA 2020
History Match to data up to	31-12-2012	31-12-2015	31-12-2016	30-04-2017	31-12-2017
Data used in HM	SPG/RFT/PNL	+ Subsidence	+ CITHP-to-BHP	+ Time-lapse gravity data	+ Strain data ZRP-3A
Key features/updates w.r.t. previous versions		Extended grid, refined fault mapping	Updated top structure, Local Grid Refinement	Additional fault seal factors to improve match	Carboniferous included, Gas-in-aquifer
		Refined saturation, PVT and well models	Daily production data and timesteps from 2011	ZRP-3 SCAL data included	Refined ZRP-3 SCAL data
			Porosity model from seismic inversion, Closed-loop rock compressibility modelling	2nd iteration seismic inversion	Depletion of Lauwerszee aquifer, uncertainty in porosity inversion

Pressures at the start of each calendar year are exported from MoReS on a 3D grid with respective coordinates (corresponding to the MoReS grid block centres) without any averaging. The V6 model is the only version that includes the Carboniferous reservoir, which is modelled as a separate layer in the subsidence model.

3.2.1 Möwensteert aquifer

There are several faults separating the Möwensteert aquifer from the North part of the Groningen gas field. In general, the degree of pressure communication in this area and the level of connectivity to the Groningen field is uncertain. The suite of historic model vintages captures an appropriate wide range of uncertainty in the Möwensteert aquifer (Figure 11).

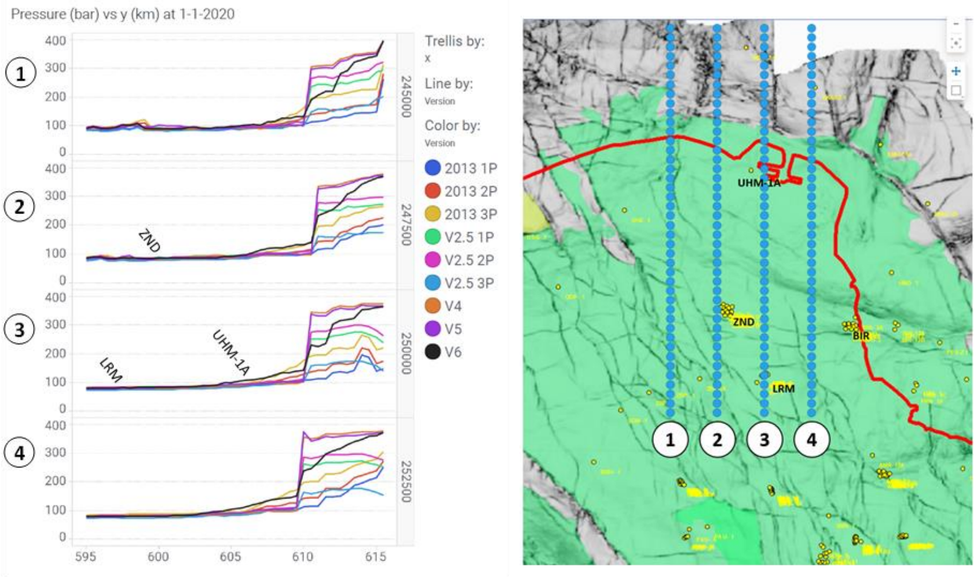


Figure 11; Slochteren reservoir pressure profiles on 1-1-2020 along four S-N lines into the Möwensteert aquifer for various model vintages. 1P/2P/3P denote low/base/high case models.

3.2.2 Rysum aquifer

RFT data from 1989 is available for the RYSM-Z1C well located in the far East of the field (Figure 12), in the Dutch-German "common area". This data showed the Slochteren reservoir to be depleted in line with the Groningen field, with some differential depletion in the water leg (up to 10 bar). The RYSM-Z1C RFT data was used in the history-match process for all model vintages in Figure 12. No pressure information is available for German wells further East. Towards the East, the spread in modelled pressures in the suite of models increases (Figure 12) but, because of the RYSM-Z1C data point, is not as large as for the Möwensteert aquifer.

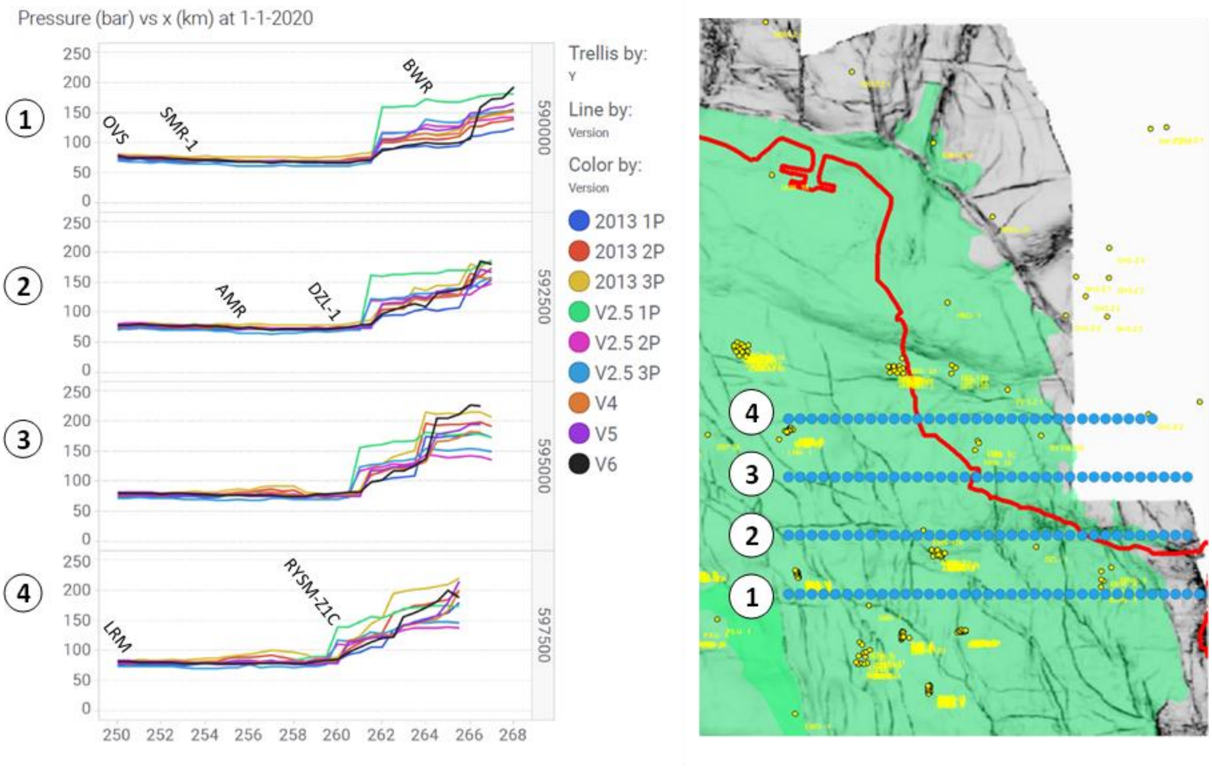


Figure 12; Slochteren reservoir pressure profiles on 1-1-2020 along four W-E lines into the Rysum aquifer for various model vintages. 1P/2P/3P denote low/base/high case models.

3.2.3 South aquifers

The South aquifers span the region between the Groningen field and the Annerveen field to the South. The gas fields Annerveen-Veendam (well ANV-1) and Zuidwending-East (well OPK-4) on the South-East side of Groningen are included in the Groningen model from version V2.5. These fields follow the Groningen pressure decline with some delay, which is matched in the model by a group of (partly sealing) faults. The suite of models captures a range of possible aquifer pressures (Figure 13).

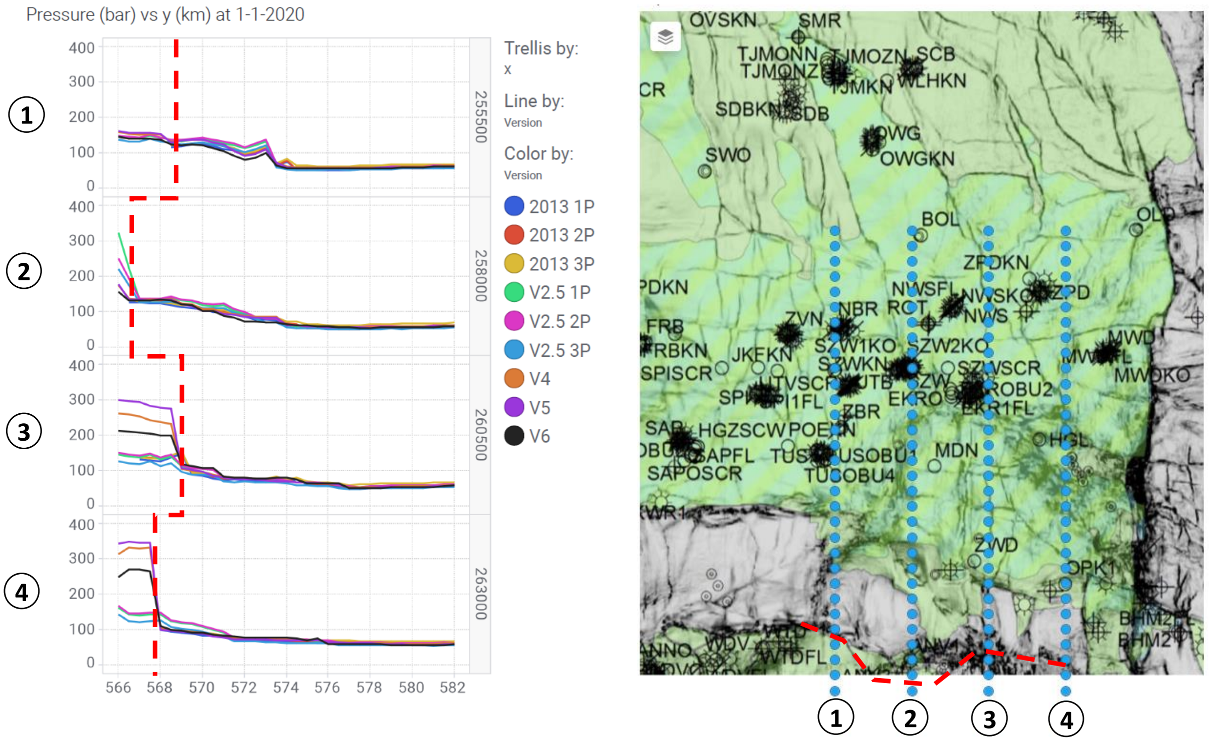


Figure 13; Slochteren reservoir pressure profiles on 1-1-2020 along four S-N lines from the Southern aquifers for various historic model vintages. 1P/2P/3P denote low/base/high case models.

3.2.4 Southern Lauwerszee aquifer (box-models)

The Groningen gas field connects at its South-West boundary to the Southern Lauwerszee aquifer ("Zuidelijke Lauwerszee Trog" aquifer in Figure 2). To the West of this aquifer there are several gas fields operated by NAM. Figure 14 is a semblance map showing the entire Groningen gas field and a large area to the West. Several large W-SW to E-NE running faults are present that based on pressure/production behaviour in the bounding gas fields and estimated fault throws are expected to be sealing. Therefore, the Southern Lauwerszee area bounding the Groningen field can be divided into several blocks that are not in direct pressure communication. Five box-models are designed to cover the area from Vries-Central in the South to the Pasop field further North. Though the exact orientation of the faults varies, the models are assumed to be in the same direction such that they are aligned without gaps in between. The idealised rectangular boxes are highlighted in Figure 15. The box-models are one-dimensional, meaning pressure gradients perpendicular to the boundary faults are neglected. At both ends of the rectangular model the pressure is constrained by either measured pressure data or pressures taken from the V6 Groningen model (Table 2).

Table 2; Description of data used to constrain pressure at box-model boundaries during history-match period

Box-model	West boundary	East boundary
Pasop	PSP-1, PSP-2 and PSP-3 SPG data (1991-2013)	TBR-1 and TBR-4 SPG data (1967-2016)
ROD100	ROD-101 and ROD-102 SPG data (1971-2011)	HRS-2 and EKL-13 SPG data (1981-2017)
ROD201	ROD-201 SPG data (1977-2013)	Annual pressure points KHM-1 of V6 model
VRSN	VRS-4, VRS-7 and VRS-8 SPG data (1985-2016)	Annual pressure points from border of Groningen V6 model
VRSC	VRS-2 and VRS-5 SPG data (1970-2019)	KWR-1A and KWR-2 SPG data (1997-2017)

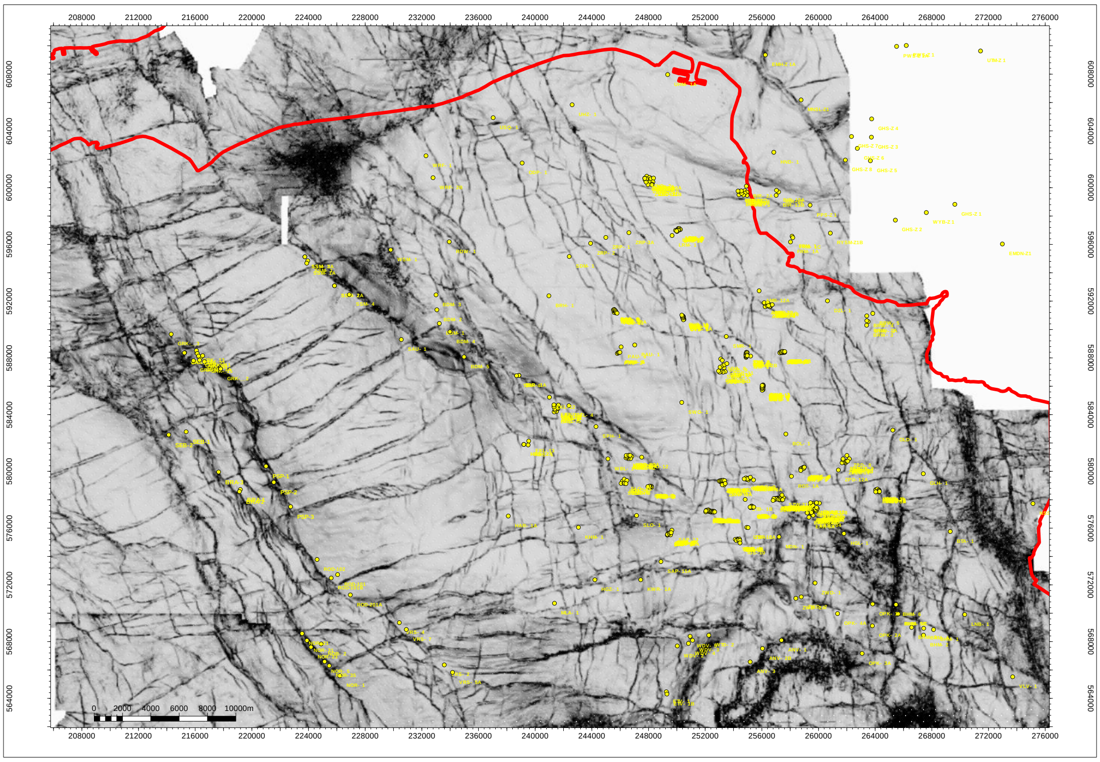


Figure 14; Semblance map of Groningen field and surrounding areas. Gas fields are indicated in green. The yellow dots represent well locations.

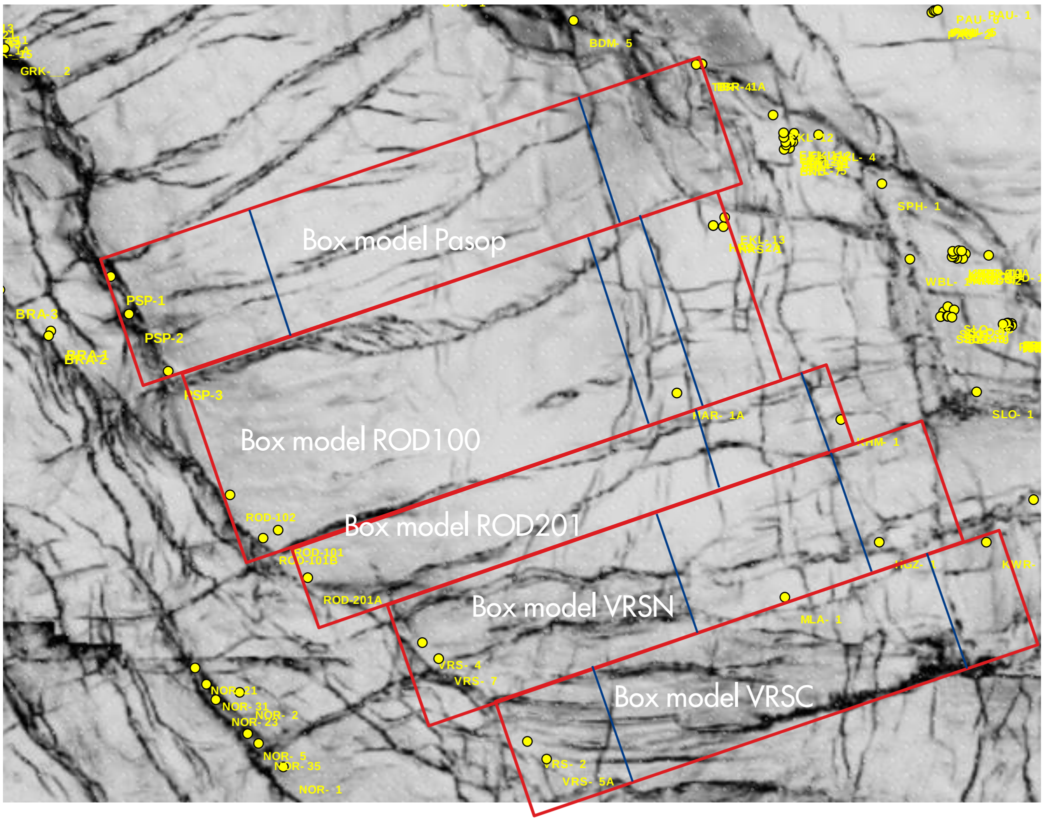


Figure 15; Detail of semblance map in Figure 14.

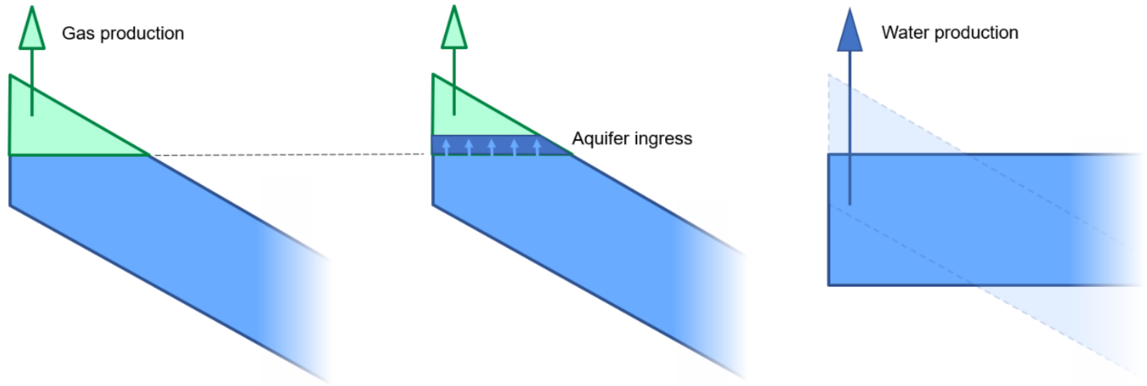


Figure 16; Schematic of the simplifying assumptions for the box-modelling. Gas accumulations in the fault bound structural highs are produced, leading to aquifer ingress within the gas accumulation (GWC rise). This behaviour is mimicked in a flat, one-dimensional aquifer model by producing an equivalent amount of water to impose the measured pressure trend from the producing gas well.

Only a single water phase is considered in the box-models which are assumed to be flat. The simplifying assumptions are visualised in Figure 16. All measured pressures used to constrain the models are converted to a reference depth of 3100 m TVDSS. This is the datum level for the Pasop field. Pressures in all box-models are defined at this depth. Note that the datum level for the various gas fields varies between 2875 mTVDSS (Groningen) and 3280 mTVDSS (Kielwindeweer). The initial GWC varies from 2996 mTVDSS (Ten Boer) to 3303 mTVDSS (Kielwindeweer).

Each box-model has two internal faults roughly corresponding to the locations of the largest S-SE to N-NW running faults (Figure 22). Incorporation of an additional fault in the VRSC model was tested but gave similar results in terms of possible pressure ranges in the aquifer. For the history-match period a collection of 1000 models is run for each box, with varying values for the permeability and two fault seal factors. Permeability can vary over three orders of magnitude (from 0.1 to 100 mD) and fault transmissibility ranges from fully open (seal factor of 1) to largely baffling with a seal factor (multiplication factor of transmissibility across the fault) of 0.001. Each model has a single well at (or near) each model boundary corresponding to the approximate locations of the wells listed in Table 2. The water production rate in these wells is adjusted such that the pressure constraints are met. A check is made whether the range of obtained water production volumes is reasonable in relation to the produced gas volumes. From the set of 1000 models, a sub-set of models that give a good pressure match is selected. Within this set, five models are chosen that cover the full range of possible pressures in the aquifer. For these five models, pressure forecasts have been generated. These are briefly discussed below for each box-model.

3.2.4.1 Pasop box-model

The Pasop box-model has a length of 19 km and width of 4 km. Up to 1-1-2015 the pressure on the West boundary is constrained by a weighted average of pressure data from PSP-1, PSP-2, and PSP-3, after which water production from the Pasop side of the model is stopped. Gas production from the Pasop field stopped a year earlier. On the East model boundary TBR-4 (and a few early TBR-1) SPG pressure data points are used to constrain the model during the history-match period. From 1-1-2020 onwards annual TBR-4 bottom-hole pressures from the V6 model forecast are used as constraint. As a result, the pressure in each selected scenario is the same on the Groningen side, whereas within the aquifer the selected cases cover a wide range of pressures (Figure 17). Historic SPG pressure data in TBR-4 is well matched with the Groningen V6 model.

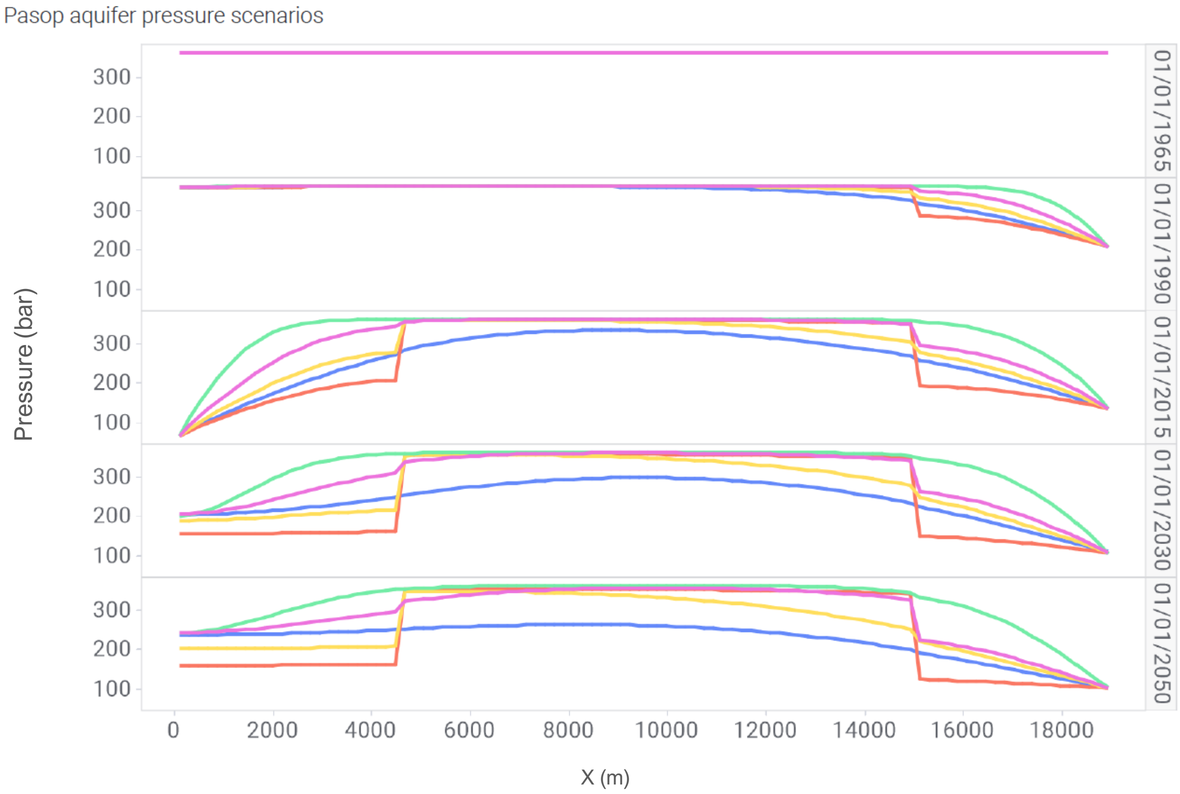


Figure 17; Pressure profiles from Pasop (left) to ten Boer (right) for five selected Pasop box-model scenarios at various time instants (from top to bottom: 1965, 1990, 2015, 2030, 2050).

3.2.4.2 ROD-100 box-model

The ROD-100 box-model has a length of 17 km and width of 6 km. On the West side of the box-model water production is constrained such that pressures matching ROD-101 and ROD-102 SPG data are obtained. At 1-1-2010 water production from the ROD-100 side is stopped. Gas production from ROD-100 wells stopped in 2003. The last SPG pressure data point from 2011 serves to select suitable models. Very recently, a new SPG has been taken in ROD-102. This data was not used to constrain the models but serves as an additional check. The pressure forecasted by the five selected models is only somewhat (2 to 16 bar) higher than the pressure measured.

On the Groningen side, SPG data of HRS-2 and EKL-13 is used to constrain model production during the history-match period. From 2020 onwards, annual V6 model pressures for HRS-2 are used. Figure 18 shows the pressure profiles across the aquifer for five selected scenarios ranging from maximum pressure (minimum depletion) in green to minimum pressure (maximum depletion) in red.

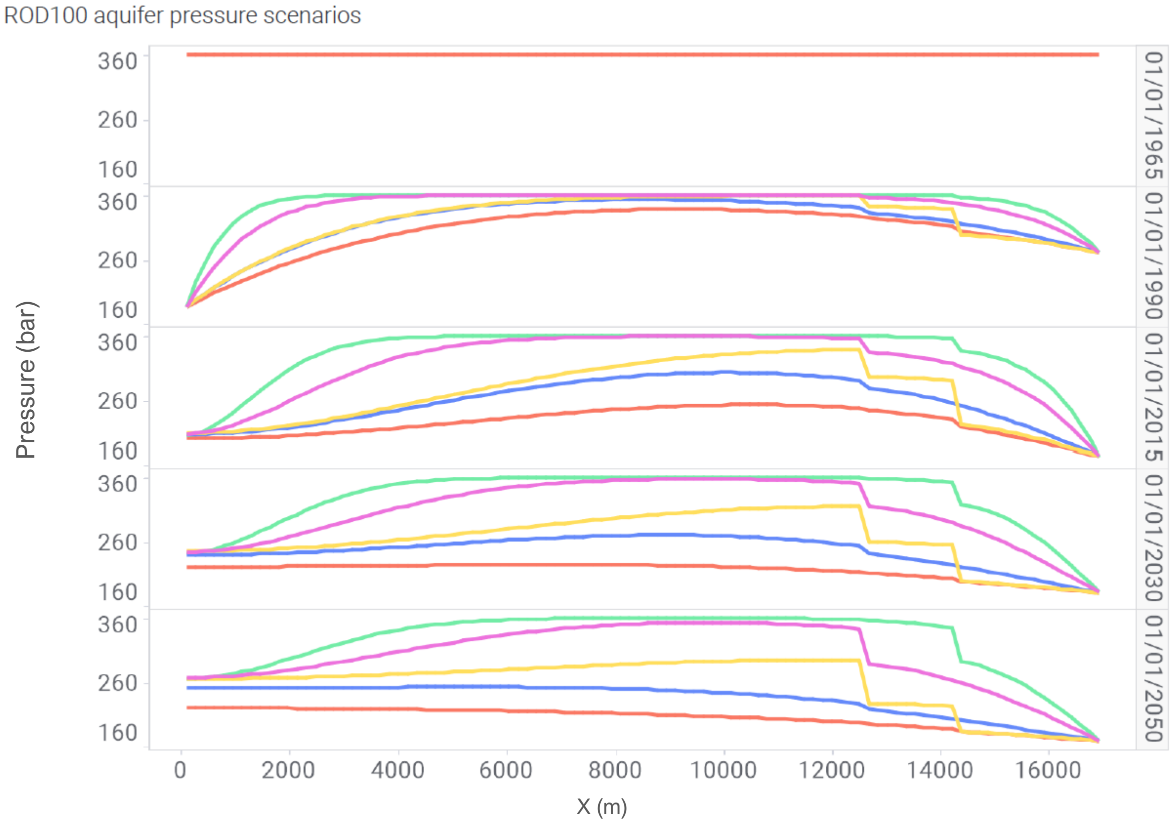


Figure 18; Pressure profiles from Roden-100 (left) to the Groningen field (right) for five selected ROD-100 box-model scenarios at various time instants (from top to bottom: 1965, 1990, 2015, 2030, 2050).

3.2.4.3 ROD-201 box-model

The ROD-201 box-model has a length of 17 km and width of 2.5 km. ROD-201 did not produce any gas after August 1986, but regular pressure data is available up to 1999. Water production from the West side of the box-model is constrained such that this pressure data is matched. From 1-1-2000 model water production related to ROD-201 gas production is stopped. An additional pressure data point in 2013 serves to check the history-match quality of the selected runs. The 5 selected runs deviate between -5 and + 7.5 bar from the SPG data point.

On the Groningen side, annual V6 model pressures for KHM-1 are used to constrain water production from the box-model. There is no SPG data available prior to 1990. The V6 model matches the later

KHM-1 SPG data points (between 1990 and 2015) reasonably well. Figure 19 shows the pressure profiles across the aquifer for the five selected scenarios.

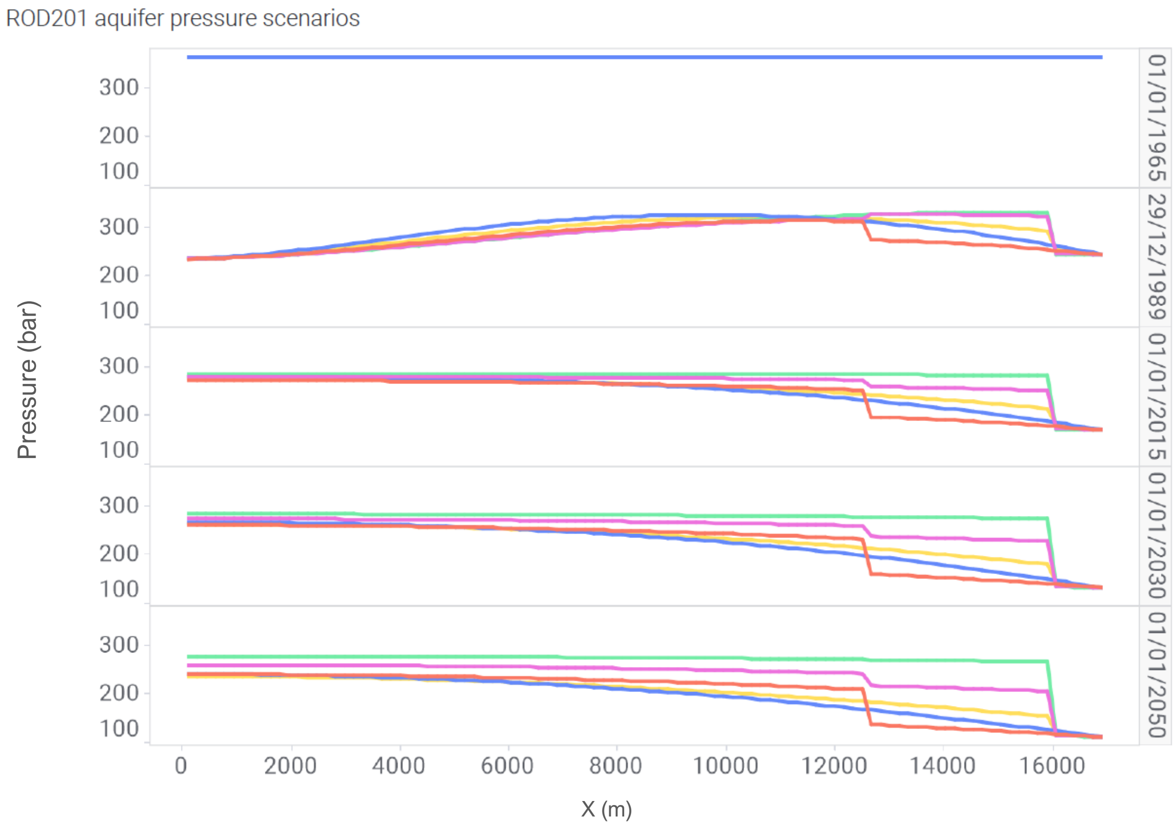


Figure 19; Pressure profiles from Roden-201 (left) to the Groningen field (right) for five selected ROD-201 box-model scenarios at various time instants (from top to bottom: 1965, 1990, 2015, 2030, 2050).

3.2.4.4 VRSN box-model

The VRSN box-model has a length of 17 km and width of 3.85 km. During the history-match period up to mid-2019 SPG data of VRS-4, VRS-7, and VRS-8 is used to constrain water production from the West side of the box-model. In the forecast, this production is stopped. VRSN wells have not produced since Jan-2019. On the Groningen side, the pressure is matched to V6 model pressure data. Figure 20 shows the pressure profiles across the aquifer for the five selected scenarios, from the maximum pressure (minimum depletion) scenario in green to the minimum pressure (maximum depletion) scenario in red.

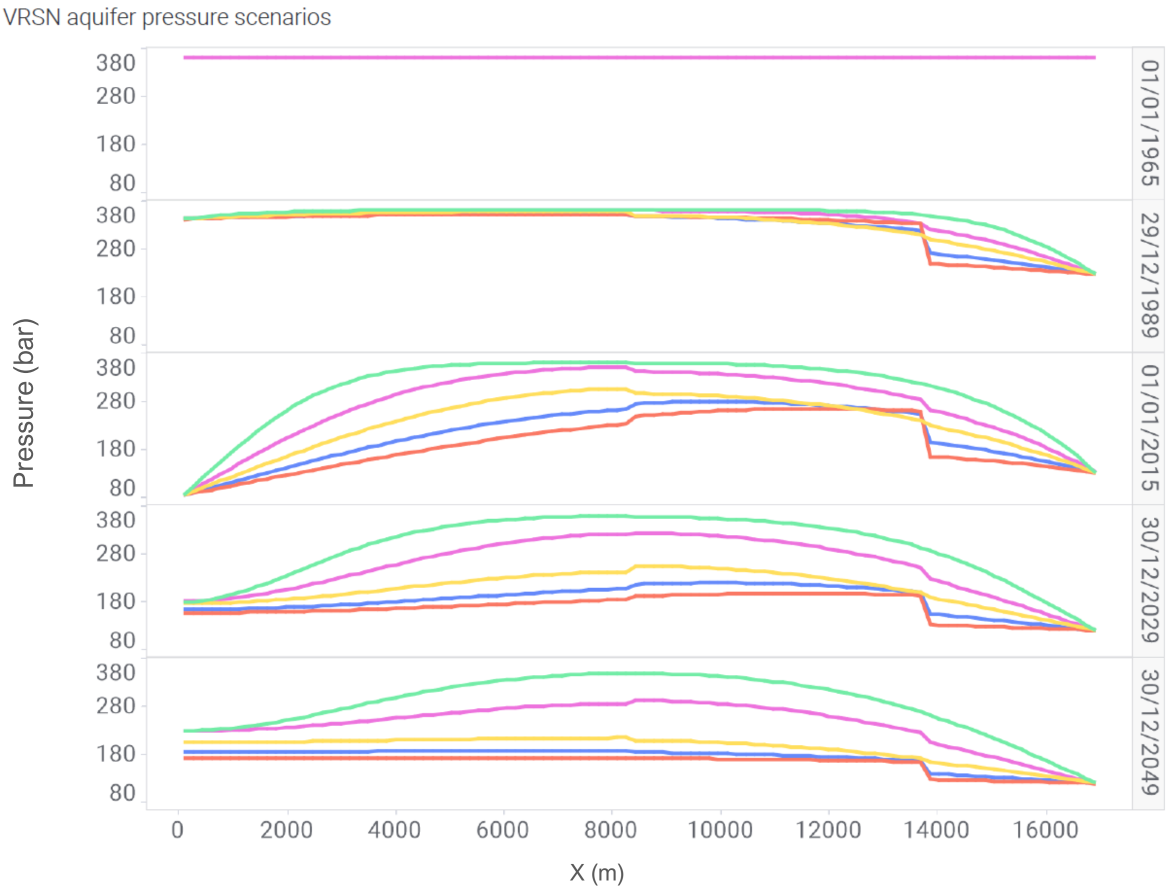


Figure 20; Pressure profiles from Vries-Noord (left) to the Groningen field (right) for five selected VRSN box-model scenarios at various time instants (from top to bottom: 1965, 1990, 2015, 2030, 2050).

3.2.4.5 VRSC box-model

The VRSC box-model has a length of 16 km and width of 3.63 km. During the history-match period SPG data from VRS-2 and VRS-5 on the West side and KWR-1 and KWR-2 on the East side is used to constrain water production from the model. In the forecast, model water production from Vries-Central is kept at a constant rate until mid-2025 when production from Vries-Central is expected to stop. This production rate is based on the last water rate required in the history-match period to obtain a pressure match and varies between the five selected models within a factor 2. On the Groningen side, again V6 model pressures (this time for KWR-2) are used as forecast constraint. Figure 21 shows the pressure profiles across the aquifer for the five selected scenarios.

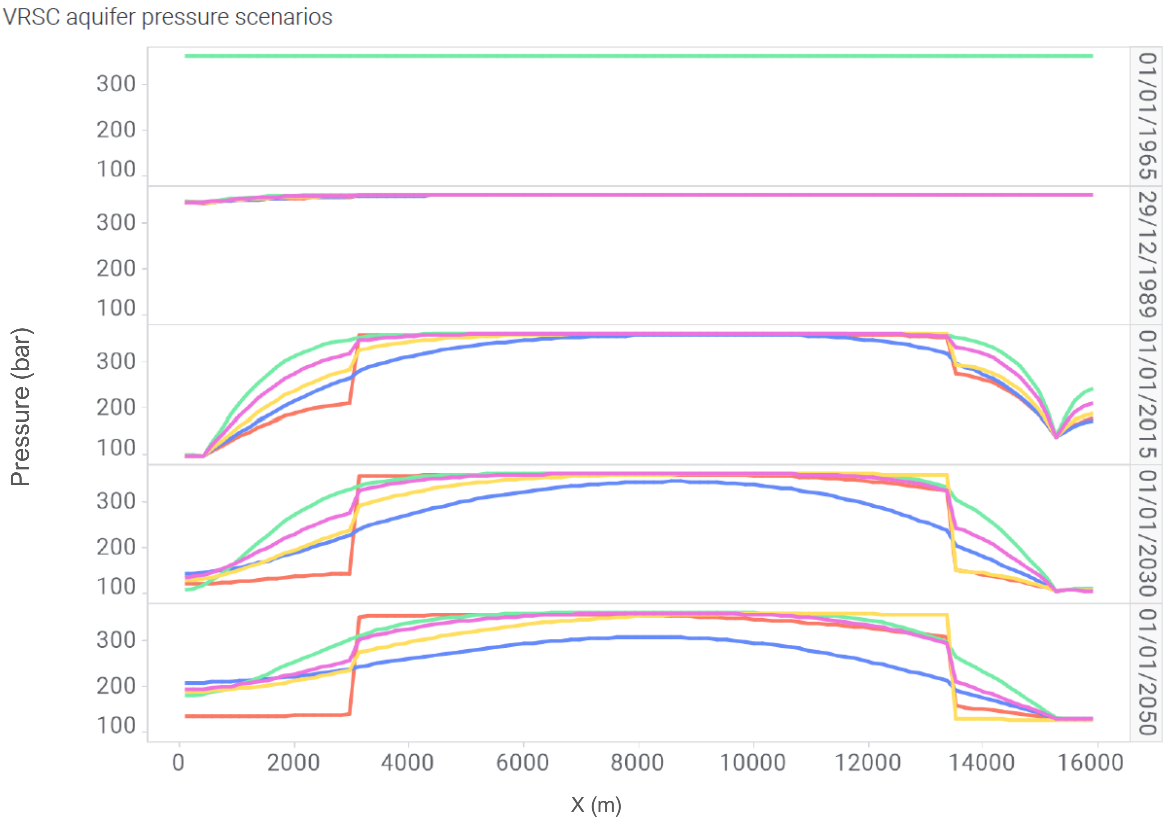


Figure 21; Pressure profiles from Vries-Central (left) to Kielwindeweer (right) for five selected VRSC box-model scenarios at various time instants (from top to bottom: 1965, 1990, 2015, 2030, 2050).

An overview of all box-models is shown in Figure 22, now for all time steps. Each graph shows the pressure in time, where time is defined on the vertical axis. The horizontal axis in each graph shows the distance from west to east. The colour of the figures represents the pressure. In these graphs, the first year, 1967, defines the initial pressure of the box-models. The pressures in the year 2020 show for instance that the pressure in the centre of most of the box-models is at its initial state. The rows in Figure 22 show the different box-models with the northern box-model on top and the most southern box-model at the bottom. The columns show the different pressure realisations per box-model from min to max.

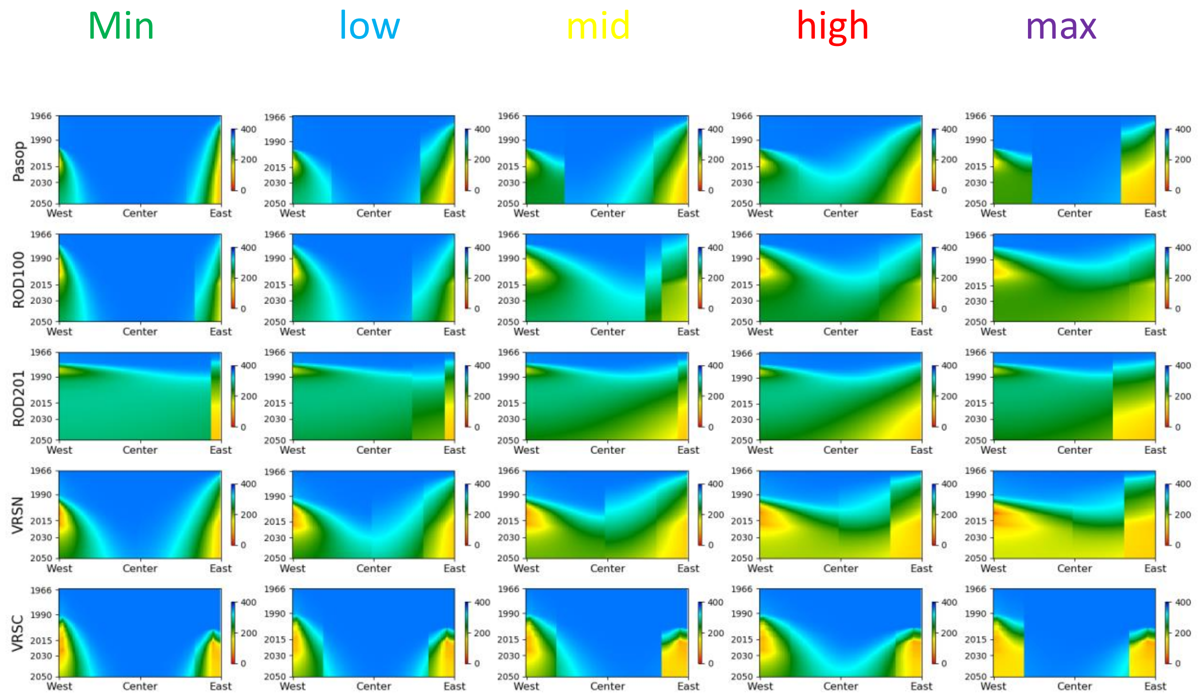


Figure 22; Box-models covering the pressure behaviour of the southwestern aquifer. From top to bottom the northern to southern box-models and from left to right the Min to Max pressure realisations. Time is presented on the vertical axis of each graph while distance from West to East is found on the horizontal axis.

3.3 Thickness of the compacting layer

A reservoir thickness map, shown in Figure 23, is retrieved from the V6 reservoir model and is used as a net thickness input for the subsidence calculations. The net thickness is based on log information from numerous wells and seismic attributes. The number of well penetrations is much higher for the Rotliegend (ROSL) sandstones than for Carboniferous (DC) sandstones. Therefore, a detailed thickness map is constructed for the ROSL while a uniform thickness map is created for the DC.

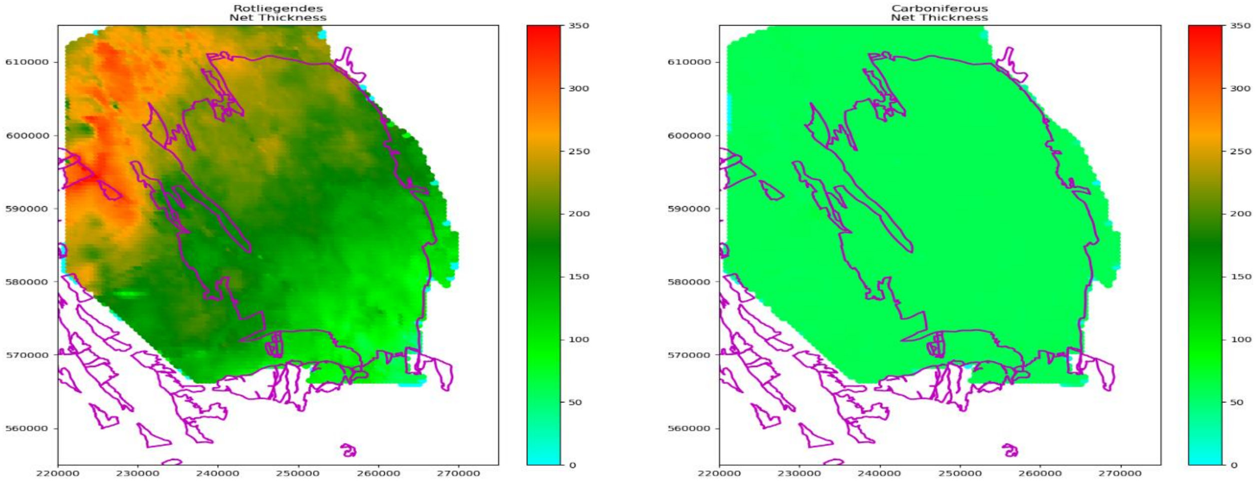


Figure 23; Net reservoir thickness of the V6 reservoir model used for the subsidence calculations. Left: net thickness for the ROSL. Right: Net thickness for DC.

3.4 Compaction model

As pointed out in section 1.3, the legacy type of compaction models and subsidence forecasts that result from these models varied in time. The simplest model, used at the start of the production, assumed a linear relationship of compaction with the thickness of the reservoir layer, the depletion and the compressibility of the rock. After the year 2000, a delay of the subsidence at the start was observed above many gas fields in the Netherlands (Ref 23), which required a more complex compaction model. A bi-linear model was adopted by NAM after the year 2000 that addresses the delay at the start of the production by applying two linear branches. In the bi-linear model the early phase of the compaction is described by a low compressibility that switches to a higher compressibility value at a certain pressure drop or level of depletion resulting in an accelerated compaction and subsidence. This simple model mimics an elasto-plastic law like the Cam-clay model. The Cam-Clay model is recently validated by laboratory experiments on Groningen core (Ref 5). A true (time-independent) elasto-plastic compaction model, however, cannot explain the delayed response of the subsidence that is observed at the end of field life. This response is very noticeable from the subsidence measurements above the Ameland gas field (Ref 23). Again, several explanations for this observation are plausible, one of them being the time dependent or visco-plastic behaviour of the sandstone itself. Other mechanisms are delayed pressure response in the aquifers or a visco-plastic deformation of the overburden.

The delay of the subsidence at the start and end of the production can be explained by a diffusional process originating from the pressure behaviour or compaction behaviour. This type of behaviour is described by Ref 6 and used by NAM for making subsidence forecasts in the period 2012-2016. After this period NAM switched to the application of the *rate type compaction model* (RTCM) in the subsidence forecasts for Groningen and other fields.

De Waal (Ref 21) firstly proposed a rate type compaction model to describe the compaction behaviour of cemented sandstone. In his model, the time dependent compaction (strain) of a sandstone is a function of the loading rate of the rock. This model originates from soil mechanics principles but is applied as well to describe the stress strain response of a sandstone plug in the laboratory. Ref 22 described the application of this model to laboratory experiments on Rotliegend sandstone core material. Improvements by TNO (Ref 22) on the original work of Ref 21 led to the definition of the isotach (i) formulation of the rate type compaction model (RTCiM), which was also implemented by NAM (e.g. Ref 23).

The RTCiM model gives the best match to core experiments when compared to any of the other compaction models. Another advantage of the RTCiM model is that it spans a wide range of temporal behaviours to reflect the possible visco-plastic behaviour of the sandstone. The RTCiM model is versatile and can behave more linear with depletion or exhibit time decay and temporal characteristics depending on the values of the parameters. This choice concurs as well with the findings of the LTS-II research (Ref 23).

3.4.1 RTCiM implementation in the study

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 (Ref 24):

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. 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 true 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.

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.

3.4.2 Validating the RTCiM model to the stress-strain curves of the core plug tests

In 2015 the Zeerijp-3A well was drilled in the part of the field that experiences the highest compaction levels. Core was retrieved from large parts of the Rotliegend and Carboniferous. Among various types of tests, core plugs from these cores were drilled and tested by uniaxial-strain pore pressure depletion tests. The compressibility value is the main objective of the test but the test procedure documented by Ref 13 allows as well for validating the RTCiM model and for retrieving the parameter values that best fit to the stress-strain curves.

An example of the stress-strain plot vs time is shown in Figure 24. The sample is firstly loaded to in-situ stress and pore pressure conditions followed by a stabilization period. Then the core plug was depleted under uniaxial strain boundary conditions. The depletion interval contains three steps with multiple loading-unloading events. Finally, after depletion, the core plug was held under constant pressure conditions for a prolonged time were the creep of the sample was assessed.

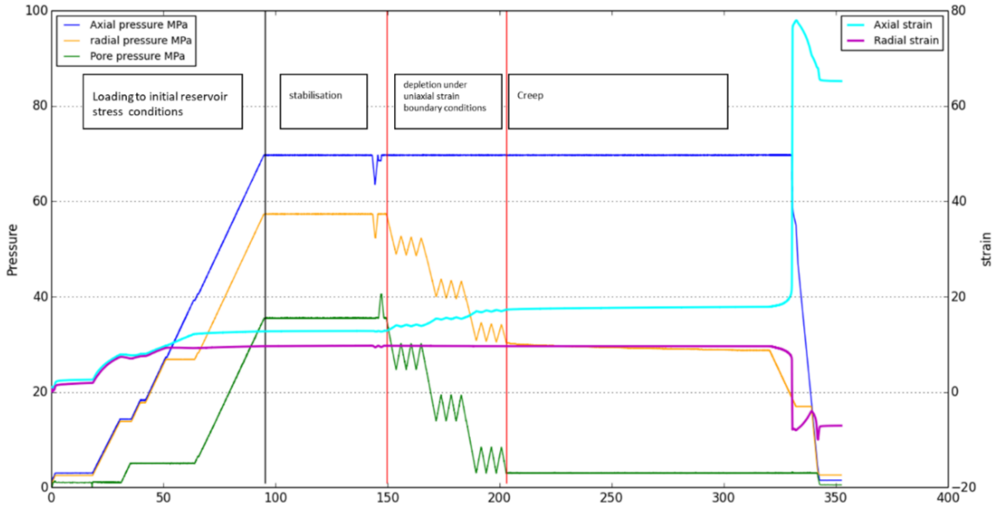


Figure 24; curves for different test variables for sample ZRP-3A_123D (from Ref 13)

Although not specifically designed for this purpose, the stress strain path of the depletion and creep phase in these tests can be used to derive the RTCiM parameters. The fit of the model to the complex stress strain curve can be regarded as a validation step of the RTCiM.

The calibration of the core plug test to a RTCiM model was done using a Latin Hyper Cube (LHC) Monte Carlo method. 10000 combinations of RTCiM parameters were generated, each with a different set of parameter values. The strain was calculated for each combination using the same stress path as used in the core plug experiment. For all combinations the Root Mean Square (RMS) between the measured and modelled strain was calculated and the parameter set with the lowest RMS value was selected.

The stress-strain curve of the core plug test was calibrated in two different ways. The first uses the part of the curve that consists of both the depletion and creep part (left graph, Figure 25) while the second uses only the depletion part (right graph, Figure 25).

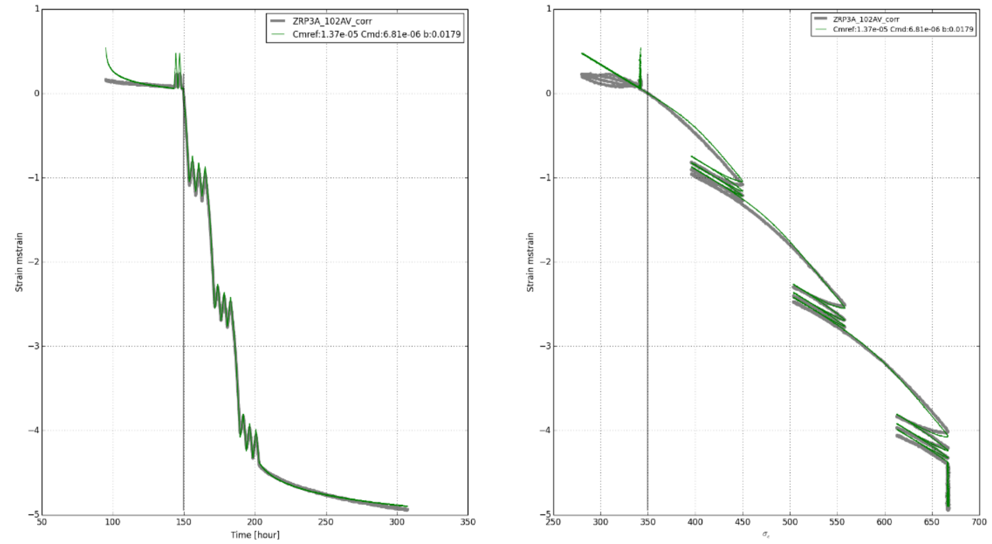


Figure 25; Comparison between modelled (green line) and measured (gray line) stress/strain path. The left figure shows the strain vs time, the right the strain vs effective Stress. The vertical line at T=150 and S at 350 indicates the start of the simulation (=start of depletion phase).

The results of the RTCiM calibration are shown in Table 3 and compared with the results from the uniaxial strain tests documented by Ref 13. Note the similarity between the values of the labtest $C_{m\ tot}$ and $C_{m_{ref}}$ and the Mean low Cm values and C_{m_d}. Both ways of calibration led to values that are quite similar, providing confidence in the values that were derived from calibration.

Table 3; Results of the lab tests and RTCiM calibration.

	Results Lab tests				Results RTCiM calibration				
Coreplug	Depth [m] Ah	Por (Hg) [%]	$C_{m\ tot}$ [10^{-5} bar^{-1}]	Mean low Cm [10^{-5} bar^{-1}]		$C_{m_{ref}}$ [10^{-5} bar^{-1}]	C_{m_d} [10^{-5} bar^{-1}]	$C_{m_d}/C_{m_{ref}}$ [-]	B [-]
ZRP-3A 38 CV	3661.74	19.14	0.928	0.56	Depletion part only	0.935	0.57	0.61	0.0067
					Depletion plus creep	0.931	0.48	0.51	0.0084
ZRP 3A_39CV	3640.85	25.38	1.56	0.87	Depletion part only	1.51	0.94	0.62	0.012
					Depletion plus creep	1.54	0.76	0.49	0.014
ZRP-3A_102AV	3561.64	22.45	1.37	0.605	Depletion part only	1.37	0.58	0.42	0.015
					Depletion plus creep	1.37	0.68	0.5	0.018

3.5 The influence model

The influence model translates the compaction of the reservoir into surface subsidence. Studies (Ref 23) into the subsidence above the Ameland field concluded that a thick salt layer above the reservoir significantly impacts the temporal behaviour of the subsidence. Compared to the Groningen gas field, the Ameland field is a relatively small field where compaction leads to stress arching in the overburden, changing the shear stress above the reservoir. These shear stresses cause creep deformation in the salt resulting in a smaller and more profound subsidence bowl. Due to its large size, stress arching is nearly absent above the Groningen field and therefore it is assumed in this study to neglect the salt creep. Still the salt can result in a steeper edge of the subsidence bowl, a phenomenon that already was recognised and described by Ref 7. We adopted therefore the influence model as described by Ref 7, that combines a half-space model with a rigid basement using a value of 7 km for the rigid basement and a Poisson's ratio of 0.2. The effect of this model results in a change of the bowl shape that is independent with time.

3.6 Definition of the calibration areas

Different areas for calibration were defined in the various steps (see chapter 4). Outlines of these areas are visualised in Figure 26. The first step, where a first set of RTCiM values is calculated, uses only the benchmarks above the gas field within the blue polygon in the figure. The benchmarks within the red polygon are selected for the aquifer realisations in step 2. These aquifer areas are indicated by the green areas in the figure. For the spatial Cm grid calculation, step 3, the same benchmarks within the red polygon were used. The small green dots show the locations of the benchmarks while the small blue dots in the regular pattern show the centre points of the subsidence model grid.

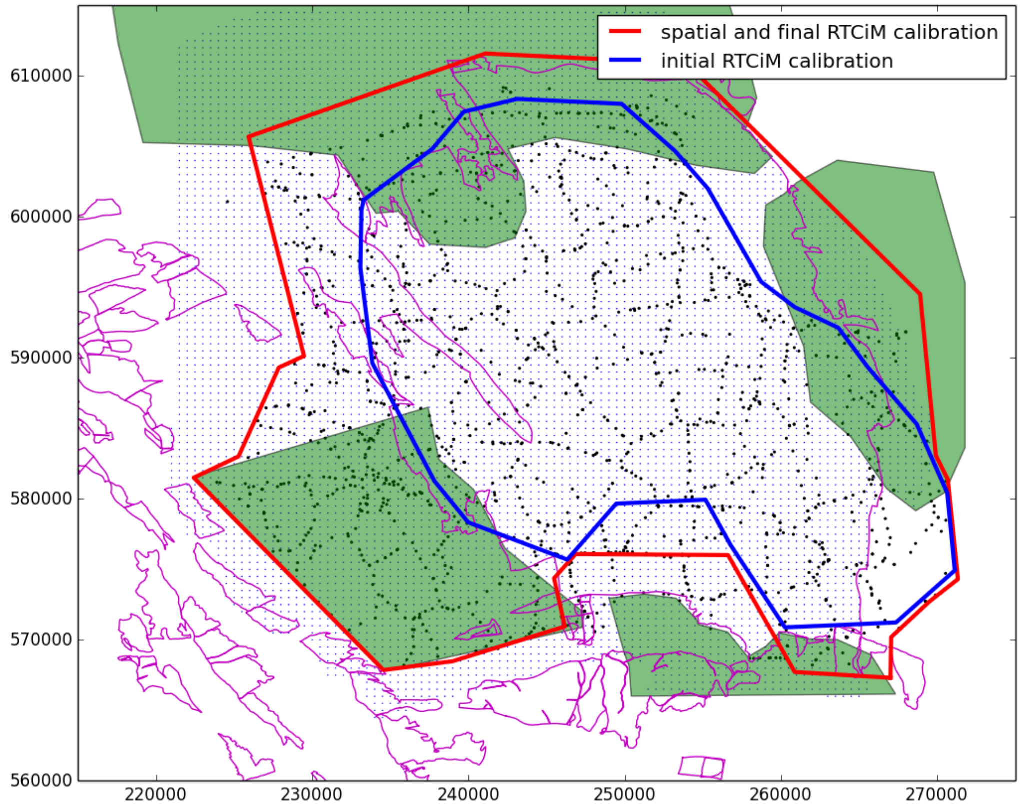
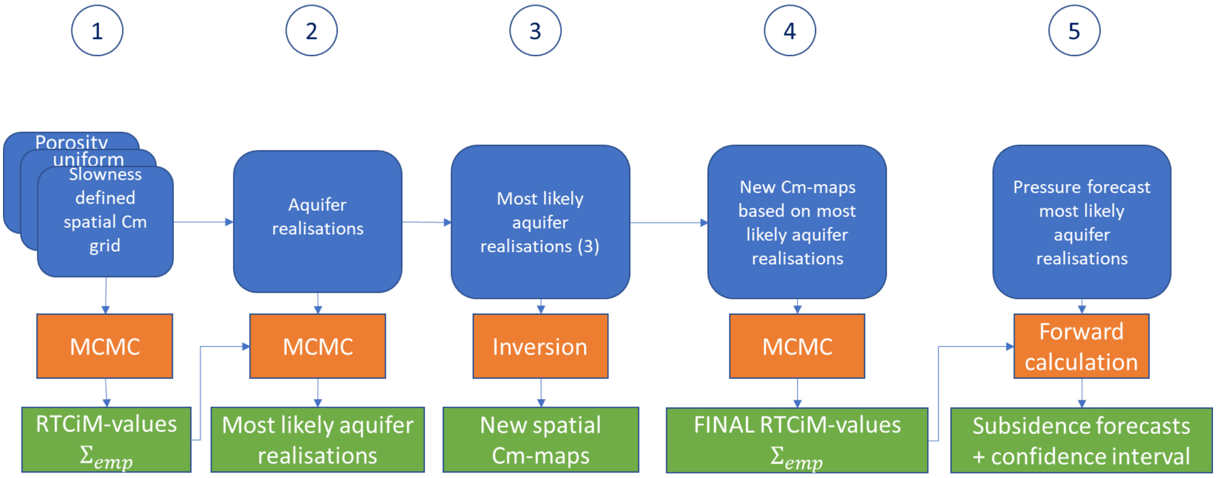


Figure 26; Calibration areas defined for the various steps in the workflow

4 Specific input and results per step

Chapter 2 described the summary and concept of the workflow that has been followed. For convenience, Figure 7 has been copied below to provide guidance to the structure of chapter 4.



4.1 Step 1, Generation of the prior Cm grids and calculate a first set of RTCiM parameters

Rock compressibility is dependent on the loading rate. Compaction calculations for (loading) periods of decades require using the static compressibility (Cm), which is generally deducted from core plug experiments, available only at sparse locations in the field. To create a Cm map over the field, correlation of Cm with a mapped (reservoir) property can be used.

In most previous NAM studies on subsidence, a relationship was used between porosity and static Cm. Figure 27 shows a cross-plot of Cm values against the measured porosity in Rotliegend core plugs. A third order polynomial trendline through these datapoints has the following expression:

$$C_m = 273 \cdot por^3 - 68.72 \cdot por^2 + 9.85 \cdot por + 0.21 \qquad [10^{-5}\,\text{bar}^{-1}],\ por\ [-]$$

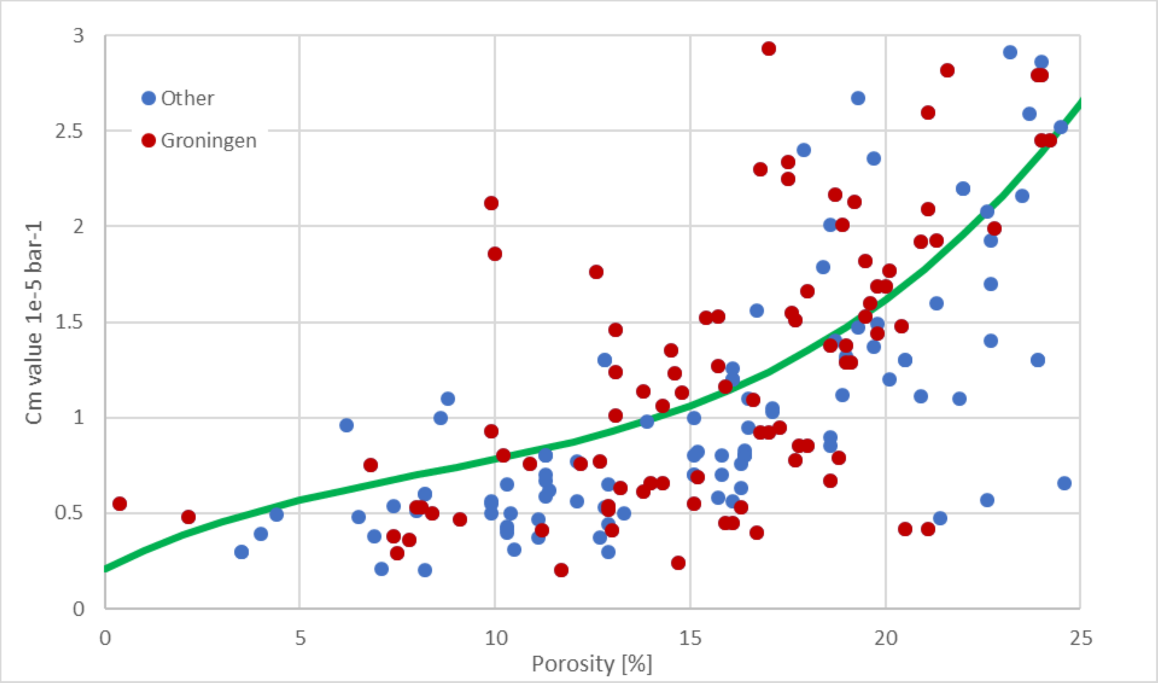


Figure 27; Cm versus porosity from core plug experiments of the Rotliegend sandstone. The green line shows the Cm-porosity relation, the red dots show the experiments from core plugs on the Groningen field and the blue ones show the results for other fields in the North Netherlands.

The correlation with porosity is an assumption and not the only possibility to create a prior Cm map. Therefore, also two other spatial distributions were used. The simplest of the two is a uniform grid. The uniform grid was chosen to check the impact of importing possible prior subsurface knowledge. A prior value of $1 \cdot 10^{-5}$ bar^{-1} is used in step 1 for the uniform grid and the map is presented in Figure 29.The other is correlation with the compressional velocity.

A recent study (Ref 30) on the distributed strain sensor that is installed in the Zeerijp-3A well investigates the linear regression of multiple log data with the compressibility derived from strain measurements. A constant pressure depletion is assumed based on the well calibrated V6 reservoir model to derive compressibility from strain. The study concludes that the acoustic information from both the compressional and shear sonic fits best to the compressibility based on the residual standard error and R^2 value (also known as the coefficient of determination).

The relation between the compressional slowness and the C_m found through a linear regression is:

$$C_m = 0.042 * DTC - 2.433$$

Where DTC is the compressional slowness, measured in µs/ft, and Cm is given in $10^{-5} \cdot$bar^{-1}.

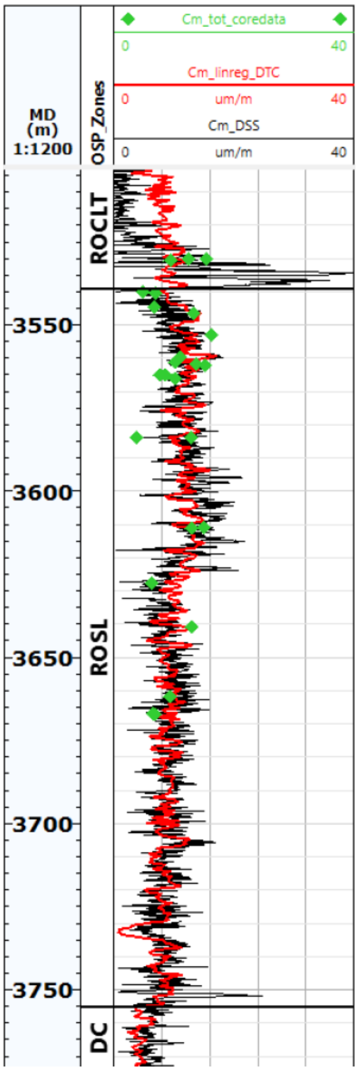


Figure 28; correlation between the velocity and compressibility in the ZRP-3A well. Black: Cm values estimated from DSS. Red: derived from the compressional sonic using the suggested linear regression. Green dots: Cm values from core plugs. Units used for Cm are µm/m/bar for all three data sets (from: Ref 30).

The benefit of the correlation to the slowness is that seismic velocities can be used and added to the sonic data in the wells to create a Cm map. Therefore, a velocity map was extracted from the Groningen seismic cube and converted to a prior Cm map applying the correlation found in the ZRP-3A well.

All maps are visualized in Figure 29. The figure shows in the first row the maps for the measured properties slowness and porosity and on the second row the corresponding compressibility maps derived from these properties. The compressibility maps are used as prior input for both the C_{m_d} and $C_{m_{ref}}$. In the MCMC calculations, multipliers to these maps are estimated.

A uniform Cm grid with a Cm of 0.3 10^{-5} bar^{-1} was used for the Carboniferous (DC) layer. The value is based on recent core plug measurements (Ref 31). Uniaxial experiments show values with an average of 0.3 $10^{-5}bar^{-1}$. This uniform DC grid has been used in the first step and is combined with the three possible Cm maps for the ROSL sandstones.

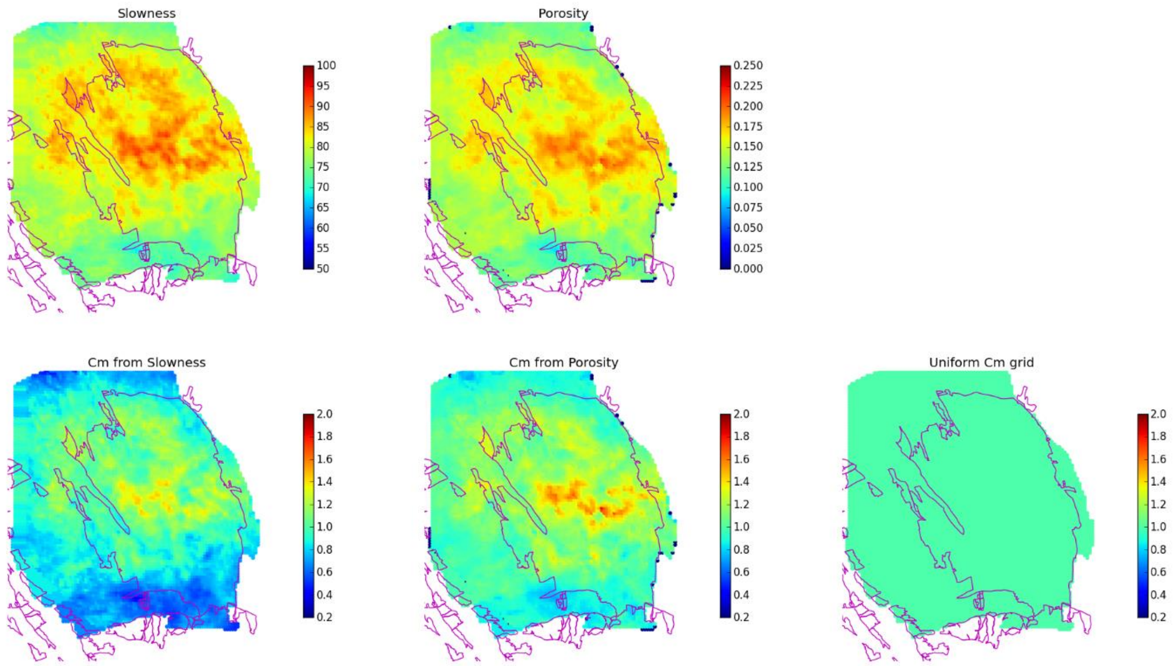


Figure 29; Top row: the slowness (μsec/ft) and porosity [-]. Bottom the ROSL Cm maps calculated using the slowness, porosity and uniform Cm map: Cm [10^{-5}bar^{-1}].
For the Carboniferous (DC) a uniform Cm map with value 0.3 10^{-5}bar^{-1} was used.

4.1.1 RTCIM parameters matched to the benchmarks above the gas field

In step 1, a set of starting values for both the RTCiM and the uncertainty parameters is defined. The Markov Chain Monte Carlo (MCMC) procedure will search for the parameters that best fit to the data.

Figure 30 shows how the different parameters change their values throughout the iterations. The example in this figure shows a scenario with 9000 iterations. Clearly visible is the steep change in parameter values in the first 200 iterations, which is reflecting the burn-in period. The end of this period is marked by a transition into a plateau (after around 1000 iterations) with small variations of the parameter values. Small variations are controlled by the settings of the thresholds in the MCMC procedure (Ref 1).

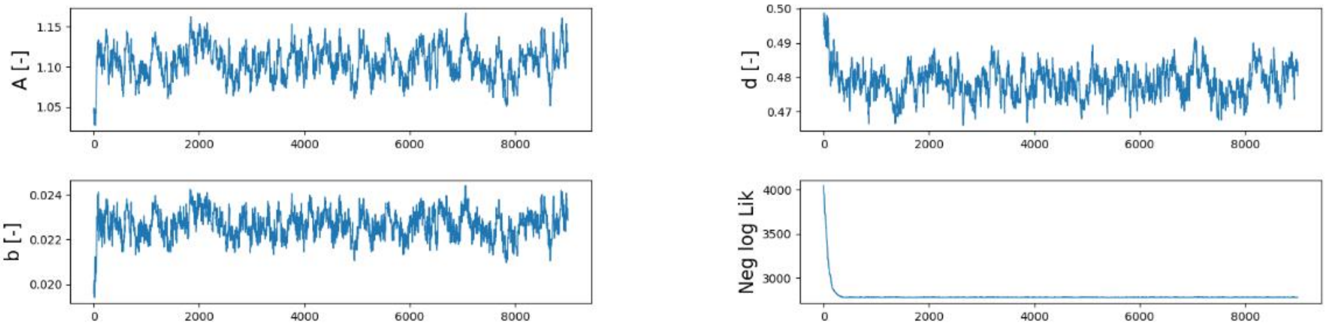


Figure 30; Example of a trace that visualises the Markov Chain Monte Carlo. A,b,d are RTCiM parameters and the Negative log likelihood is the metric of the goodness of fit.

For definition:

A: multiplier to the Cm map to obtain the value for $C_{m_{ref}}$
d: multiplier to the Cm map to obtain the value for C_{m_d}
b: b-values in the RTCiM
σ_0: intercept of the linear equation for Σ_{emp}
α: slope of the linear equation for Σ_{emp}
NLL: Negative log likelihood

The MCMC in the first step is designed for 9000 iterations where 1000 iterations were used for the burn-in period. The double-difference measurements were selected within the blue polygon in Figure 26. The reason for the limited area is that the historical evolution of the gas pressure is well history matched by the reservoir model and therefore provides more certainty to a correct description of the compaction behaviour of the reservoir. The result of the MCMC calculation in step 1 is tabulated in Table 4. Note that the RTCiM parameters A and d are multiplication factors to the various prior Cm maps. Figure 31 shows the histograms of these parameters for each of the Cm maps, using the iterations after the burn-in period.

Table 4; Mean values for the RTCiM and uncertainty parameters after step 1

	Cm grid		
Solved parameters	Uniform	Slowness	Porosity
A [-]	1.06	0.90	0.81
d [-]	0.48	0.43	0.40
b [-]	0.022	0.021	0.020
σ_0 [cm^2]	0.67	0.74	0.70
α [cm]	0.071	0.081	0.072
NLL [-]	2555	2800	2613

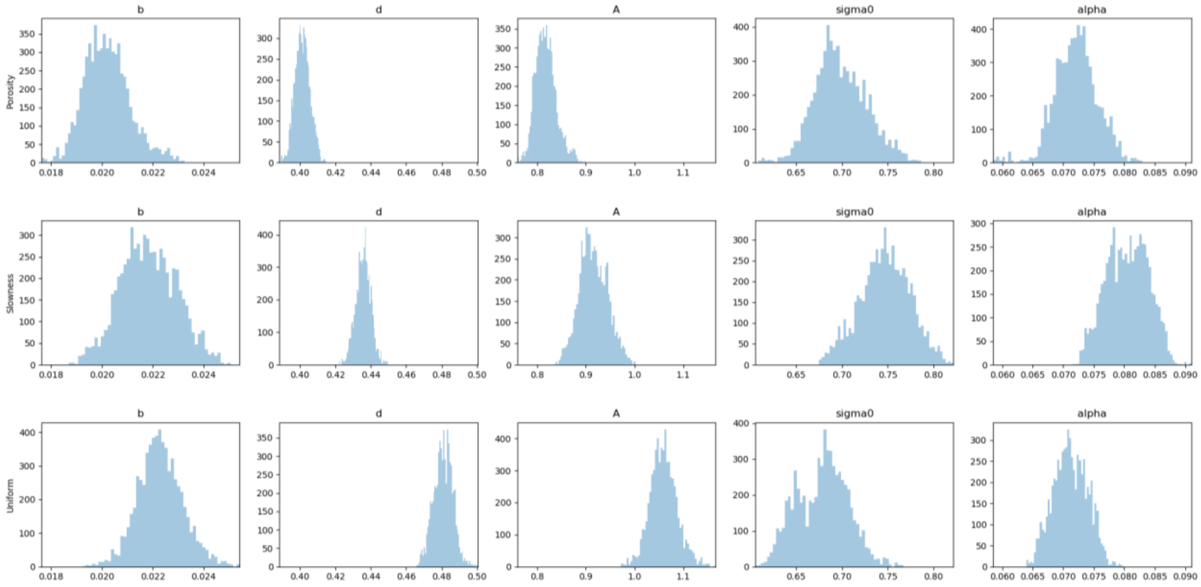


Figure 31; Histograms of the output parameters found for the different Markov Chain Monte Carlo runs. Each row shows the results for a different Cm map, the Y-axis represent the frequency count of the variable in the MCMC.

A prediction interval is defined by the sum of the parameter uncertainty, Σ_{emp} and Σ_{geod}. The contribution of the parameter uncertainty with respect to the Σ_{emp} is insignificant (< 5%) and is therefore not considered. A 95% prediction interval should therefore cover around 95% of the measurement values. Figure 32 shows, for the uniform Cm map, on the left the prediction interval in blue for all predicted double differences that are ranked according to the median modelled displacement. The prediction interval is compared to the measurement values and coverage can be calculated. A 90.5% coverage has been calculated; close to the theoretical value.

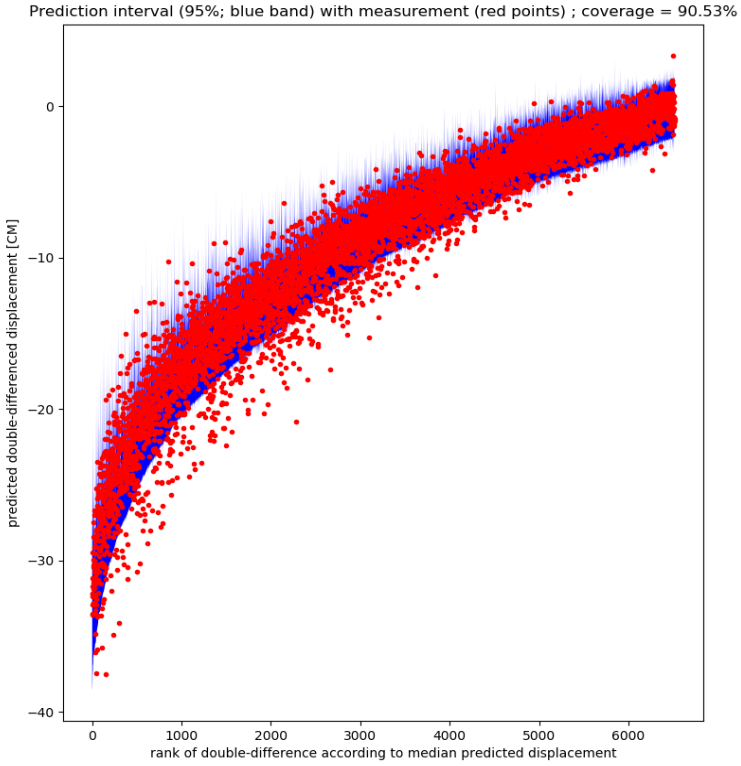


Figure 32; Prediction interval and coverage (uniform Cm map).

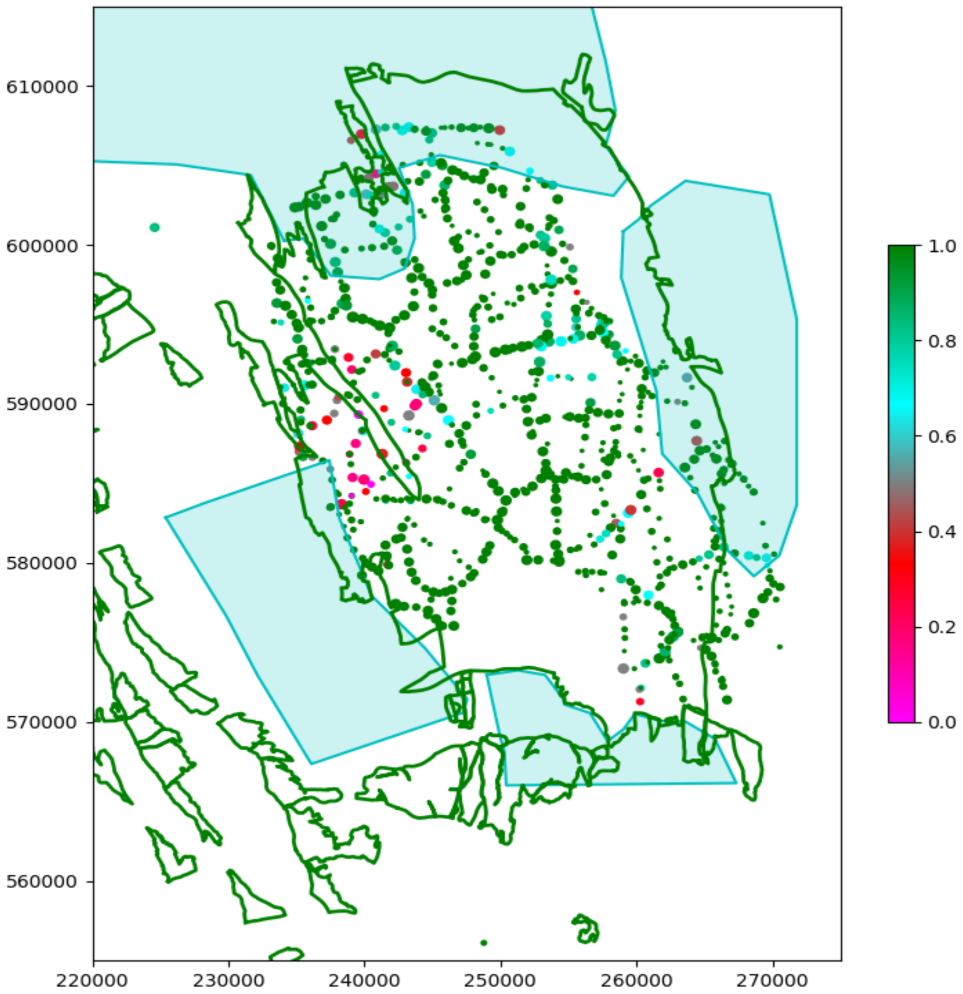


Figure 33; Coverage per benchmark. Size of the dot indicates the number of double differences that are linked to the benchmark. Colour indicates the fraction of double differences that are inside the prediction interval, green means more double differences inside the prediction interval.

Figure 33 shows a map of the benchmarks that were used in the calculation. The dot size represents the number of double differences to this benchmark. The colour of the dot indicates for each benchmark the fraction of the amount of measured double differences that fall within the prediction interval. This is defined as the coverage. Green means that all double differences that are related to the benchmark are within the prediction interval, while the pink colour means that none of the double differences are within the prediction interval. The white space in the south with no measurements indicates that the area was deselected from the calculation because of interference with the subsidence being caused by the salt mining activities in that area. The blue areas show the locations of the aquifers, discussed in step 2 (next paragraph).

A cluster of pink benchmarks is apparent in the eastern part of the field indicating a possible misfit of the model to the data. A possible explanation for the misfit could be that locally the Cm is higher than the value used in the model. This type of local issues will be addressed in steps 3 and 4.

4.2 Step 2, selection of the aquifer realisations

At the start of step 2 we have a calibrated set of RTCiM and uncertainty parameters for each of the Cm grids based on the pressure evolution for the gas and aquifer areas from the V6 reservoir model. In step 2, the negative log likelihood (NLL) of the pressure realisations from the V6 and other models, like the box-models, is calculated using the measurements above each individual aquifer area. This is done for each aquifer depletion scenario and for each Cm grid. Because of the prohibitive amount of computing time (multiple days per realisation), no further calibration is performed for the RTCiM and uncertainty parameters.

4.2.1 Definition of the pressure realisation of the box-models

The NLL of each reservoir pressure realisation is calculated using the set of RTCiM and uncertainty parameters of step 1. Each realisation has defined pressure profiles for the gas field and the aquifers. The gas field pressure is the same in all scenarios, i.e. the history matched pressure of the V6 model. The aquifer pressure profiles are varied in each realisation. Five box-models, with each having five pressure realisations, are provided to deliver the pressure realisations in the Southwestern aquifer area (see section 3.2.4. and Figure 22). The number of permutations is in total 3125 plus an extra permutation by the pressures of the V6 model in this area instead of the box-models (visualised in Figure 34). 3126 permutations are evaluated for each Cm grid. The realisations are sorted per row where each row reflects a box-model shown on the map in Figure 40 (outlined by the yellow rectangles).

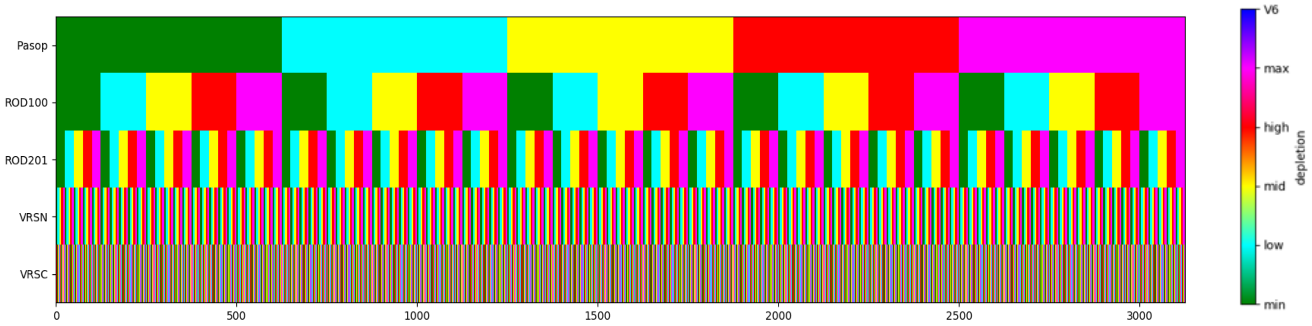


Figure 34; Selected aquifer realisations. Each row shows the naming of the box-model. Each column shows the combination of pressure profiles for the specific box-models (rows). The colours indicate relatively the amount of depletion. The V6 realisation is not visible at this scale.

4.2.2 Results for the box-models

The NLL value per permutation for the SW aquifer for the slowness grid is shown in Figure 35. In the top graph, the x-axis represents all the permutations, visualised again in the lower part of the figure for convenience. The columns in the bottom figure can be associated with the value in the graph above.

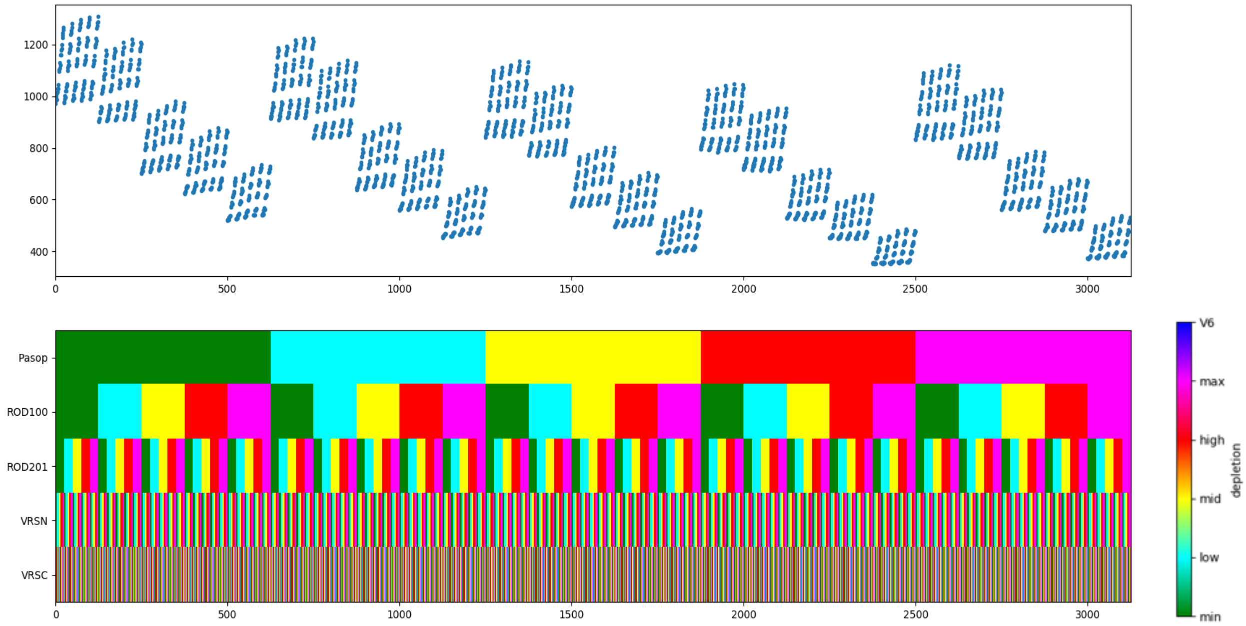


Figure 35; Top: NLL per aquifer permutation for the SW aquifer and slowness grid. The x-axis shows the number of the permutations. Bottom: graph showing the corresponding permutation in colour, with the colours indicating relatively the amount of depletion. The V6 realisation is not visible at this scale.

The same exercise is executed as well for the porosity and uniform Cm grid. Subsequently the probabilities, derived from the NLL, are sorted from high to low probability value, up to a cumulative value of 95% and shown in Figure 36, Figure 37 and Figure 38 for respectively the slowness, porosity and uniform Cm grids. From the 3126 realisations, 18 realisations comprise the 95% cumulative probability for the slowness Cm grid, 5 for the porosity defined Cm grid and only 1 for the uniform Cm grid. For all grids it can be concluded that the execution of the MCMC method for the two most northern realisations (PSP and ROD100) resulted in the selection of the same reservoir realisation (high and max depletion case), while for the other three box-models there is more variability between the grids but with a tendency of selecting the realisations with less depletion.

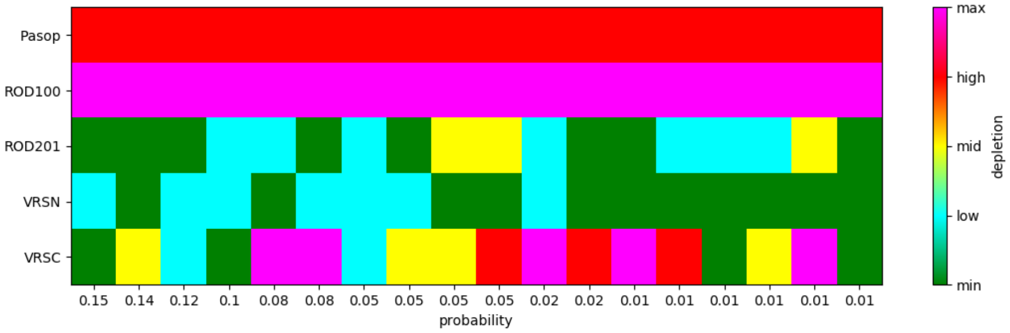


Figure 36; Probability for the SW aquifers based on the slowness Cm grid. The x-axis shows the probability of the realisations sorted from a high to low value.

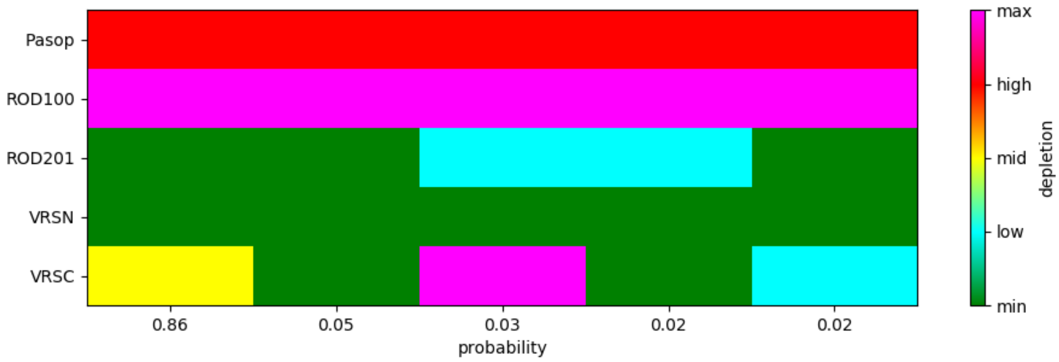


Figure 37; Probability for the SW aquifers based on the porosity Cm grid. The x-axis shows the probability of the realisations sorted from a high to low value.

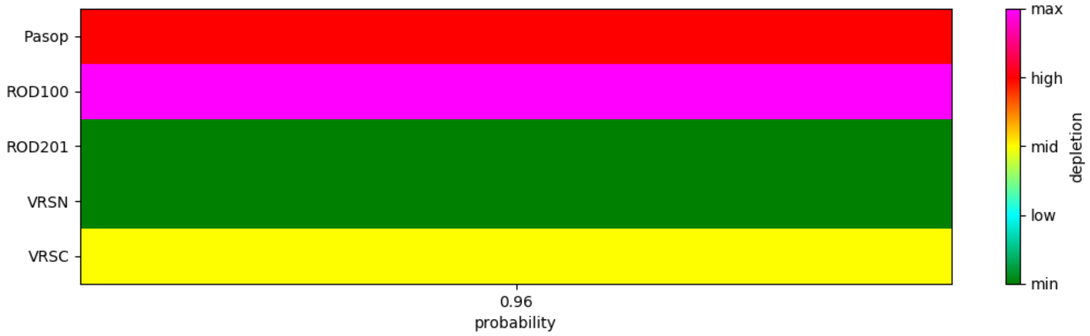


Figure 38; Probability for the SW aquifers based on the uniform Cm grid. The x-axis shows the probability of the realisations sorted from a high to low value.

The depletion profiles as shown in Figure 22 for the various box-models are again shown in Figure 39 but now with the results, shown in previous figures, projected on top of it. The box-models are ordered from min to max depletion in the columns and from north to south in the rows. The depletion patterns are described in section 3.2.4.
Per Cm grid, the letters in the figure are positioned on the most probable aquifer realisation that was selected, where U stands for uniform, P for porosity and S for slowness. The combination of these selected realisations per box-model forms the selected aquifer permutation that will be used in step 3. The selected permutation for the uniform and porosity Cm grid is the same. The coloured arrows show the ranges of possible realisations within the 95% cumulative probability per box-model. This is not given for the uniform grid because only one permutation resulted already in a probability of 96%.

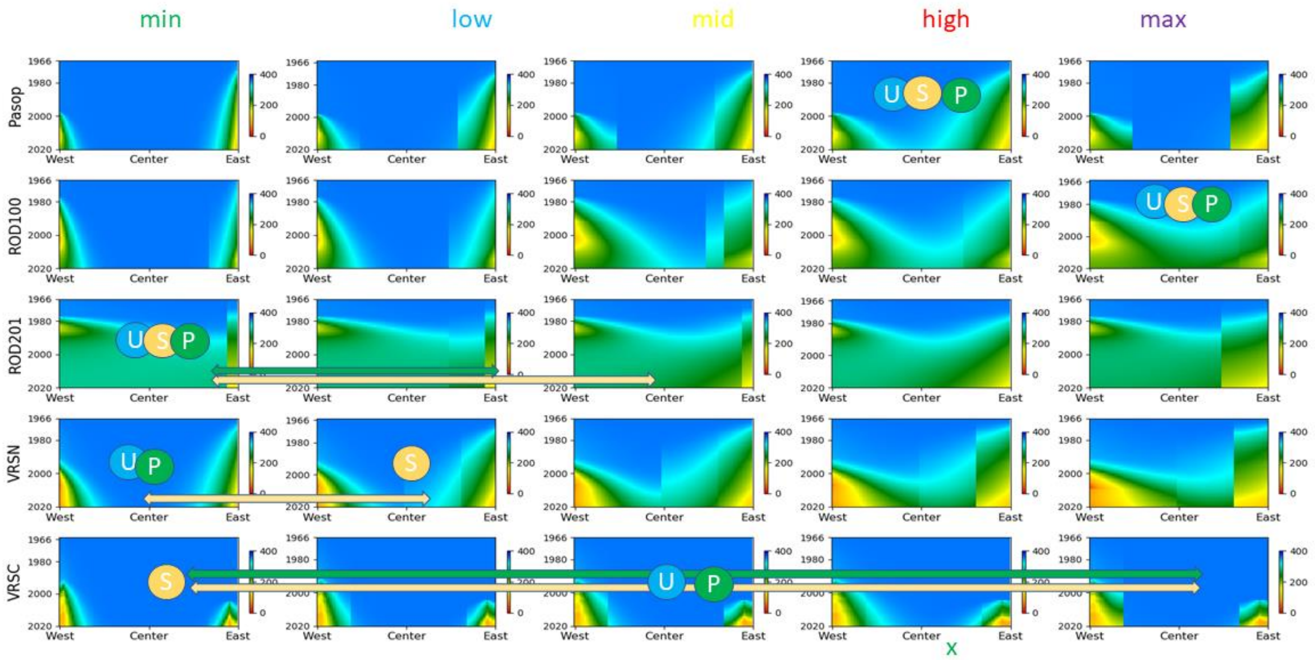


Figure 39; Aquifer realisations with the results plotted on top of it. Letters for the various grids (Porosity, Slowness and Uniform) are plotted on the most probable realisation while horizontal arrows indicate the ranges of possible realisations within the 95 cumulative probability band for the slowness and porosity Cm grids.

4.2.3 Other aquifers

For the definition of the pressure variation in the Möwensteert, Rysum and the South aquifers, 9 older versions of the Groningen reservoir model are used (paragraph 3.2) showing a large spread of possible pressures in both time and space.

The pressure realisations consist of the gas part from the V6 model, a compilation of pressure profiles for the box-models and one of the possible pressure profiles for either of the Möwensteert, Rysum and South aquifers. Figure 40 shows how the aquifer areas (in purple) are stitched to the gas bearing part of the V6 model (in green). There is an overlap with the gas area to enable a smooth transition between the older versions of the Groningen reservoir model and the latest V6 model.

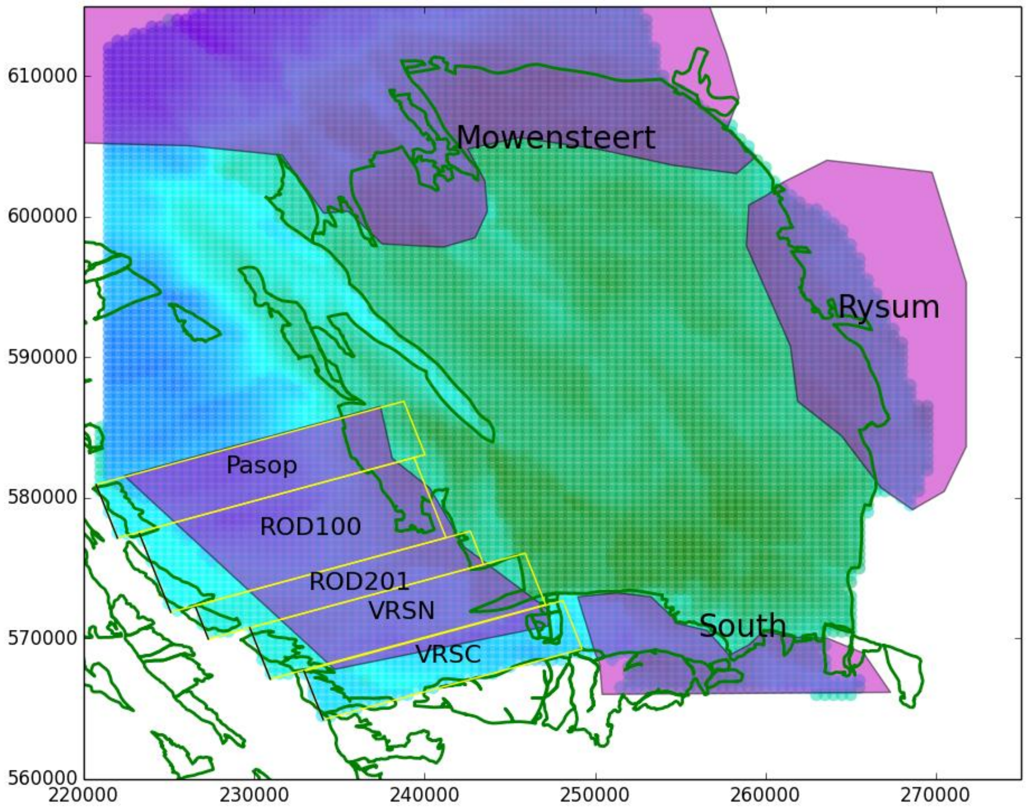


Figure 40; map of the Groningen pressure area. Gas fields are bounded by a green outline and aquifer areas are shown by the purple areas. The yellow boxes in the Southwestern aquifer are the box-models. The dots in the background indicate the total V6 reservoir model area.

Parts of the box-models extend outside the V6 reservoir model grid. For these areas, values for the Cm, depth and thickness are the average values of the box-model inside the Groningen model. In case of the Möwensteert, Rysum and South aquifers the V6 model is large enough to cover necessary aquifer areas. This means that the values for the depth, thickness and Cm are fixed for all scenarios.

Because of the numerous wells and pressure measurements in the gas field and lack of this information in the aquifers, the aquifers are treated as separate calibration areas. Besides the calculation of an overall value for the NLL, an additional NLL value is calculated for each aquifer per realisation. The pressures for the DC are not varied.

The results for the values of the NLL are presented Figure 41 for each of the areas and used Cm grid. The figure shows, for all Cm grids, that the V6 model performs best in the Möwensteert and Rysum aquifer areas. Only a few benchmarks were available for the NLL calculation above the South aquifer, which is visible in the low value of the NLL. It is concluded that the depletion calculated by the V6 model, matches best with the subsidence data and is therefore selected as the realisation that provides the pressure input for the aquifer areas other than the southwestern aquifer.

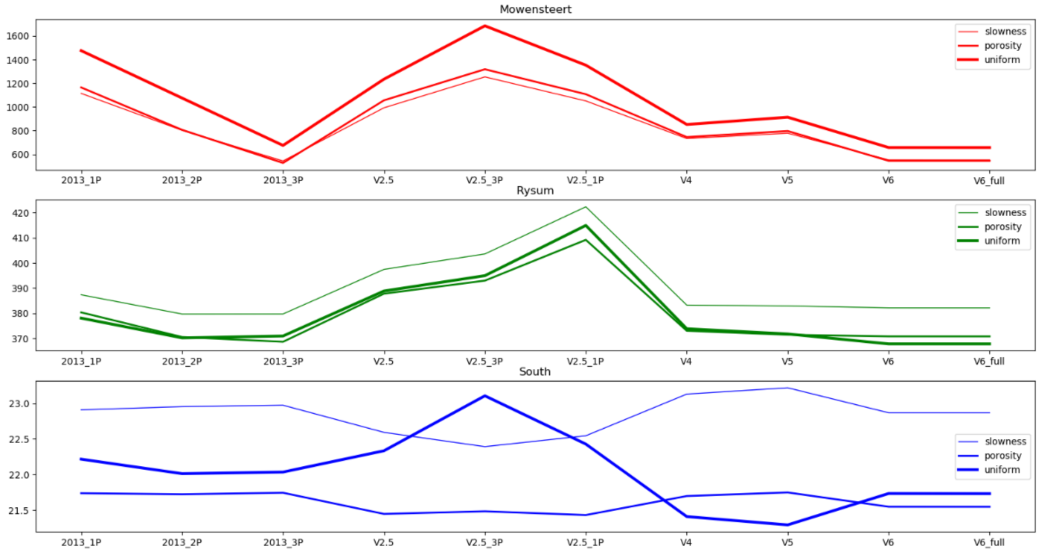


Figure 41; Negative Log Likelihood per aquifer realisation for the other aquifers.

A picture of the most likely and chosen pressure realisation for the uniform and porosity based Cm grids is presented in Figure 42. The colours show the depletion in the year 2020.

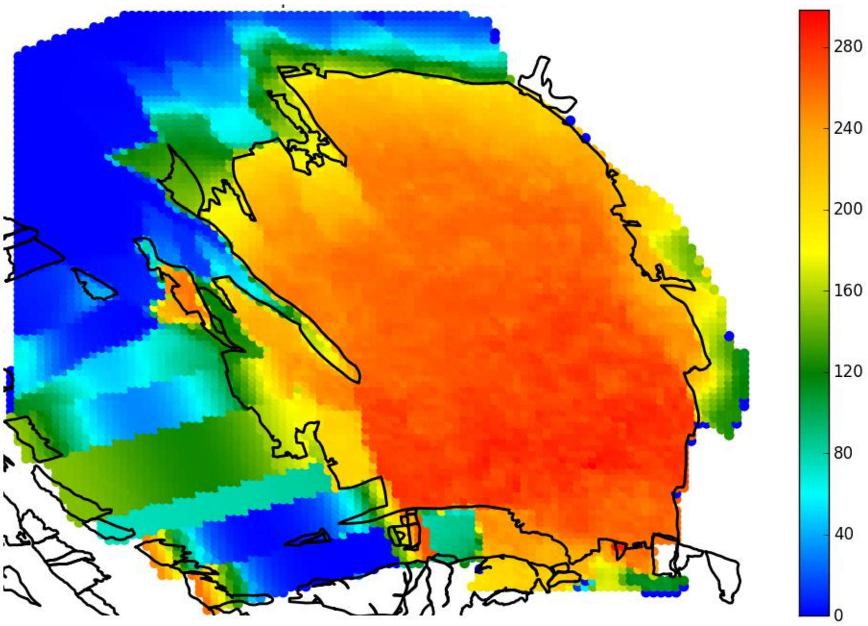


Figure 42; Pressure depletion [bar] for the selected aquifer realisation (uniform and porosity-based Cm grids)

4.3 Step 3, inversion to obtain optimal spatial Cm grids

With the RTCiM and uncertainty parameters from step 1 and chosen pressure realisation in step 2, step 3 focusses on the local reduction of residuals. The same inversion method that is described by Ref 3 will be used in this step. This method uses both the information from the subsidence measurements and the prior uniform, slowness and porosity-based Cm grids.

An inversion to subsurface compaction requires a smoothing or regularisation method to avoid bullseyes in the Cm grid during the inversion. Ref 3 describes how a prior porosity-based Cm grid is used to serve that role. A penalty function is defined that controls the importance or weight of the Cm grids, and therefore “smoothness” to the prior Cm grid, in the inversion procedure. A lower penalty value means a smaller role of the Cm grid during the inversion, resulting in a larger local variation of Cm values (possibly more bullseyes) but a better fit to the measured subsidence data. The effect of the penalty function on the posterior Cm grid is illustrated in Figure 43 for the uniform Cm grid. A low value of the penalty factor results in a map with small spatial clusters of very high and low Cm values. A penalty factor of 100 will give values in the Cm grid that are very close to the value of the initial grid.

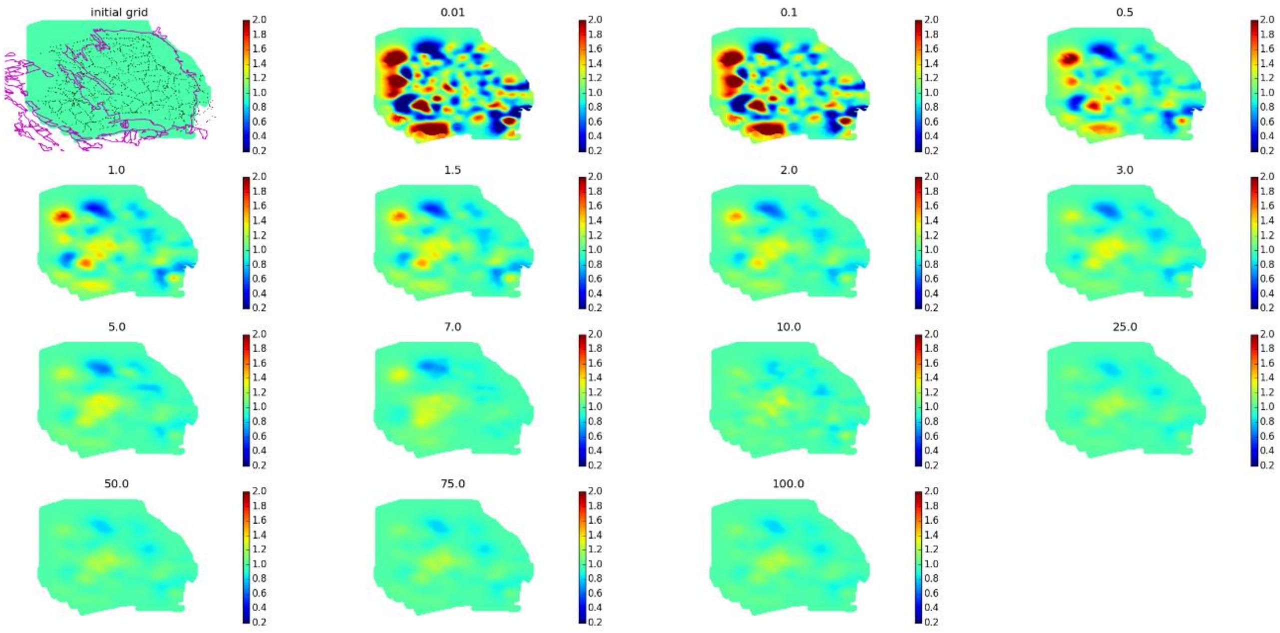


Figure 43; Resulting spatial Cm grid after inversion to the measured data using the uniform grid as prior Cm grid. The colour indicates the Cm value, the number above the graph is the penalty value that is used in run. The figure top left is the prior initial Cm grid.

The aim is now to find a subsurface posterior Cm grid that still looks plausible, where the word plausible can be quantified by evaluating the RMS and the Negative Log Likelihood as a function of the penalty factor. The RMS value is defined by:

$$\mathrm{RMS} = \sqrt{\mathrm{n}^{-1}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{DDg}-\mathrm{DDm})_{\mathrm{i}}^{2}}$$

With DDg being the measured double differences and DDm the modelled double differences. The functions for the misfit will likely show a curved response with low values at low values for the penalty factor and high values for the metrics of misfit when high values for the penalty function are used. The point with the highest gradient can be selected as an optimum for a balance between the fit to the data and a grid that still honours the patterns and values shown by the prior Cm grids.

A set of inversion runs were set-up for each Cm grid with different values for the penalty factor. The results of these runs are visualised in Figure 44. The vertical blue line shows the point close to the highest gradient at a value for the penalty factor of 5. At this value, it also apparent that the posterior slowness-based Cm grid shows the highest values for both the RMS and NLL meaning the poorest fit to the data. The posterior uniform grid shows the best fit to the measured data at this stage.

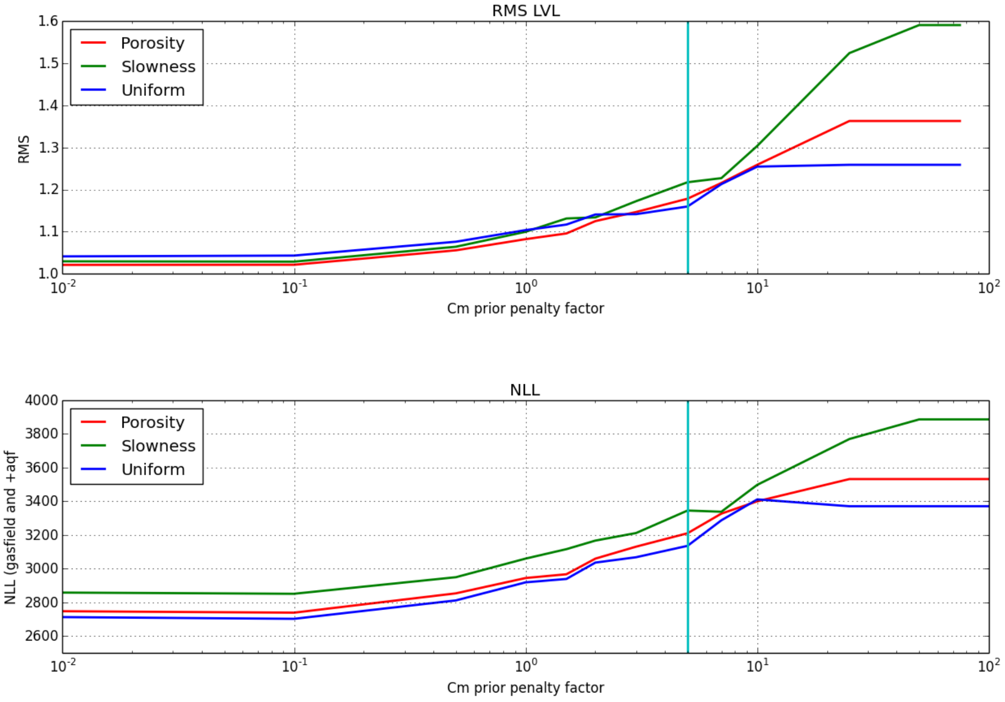


Figure 44; RMS (top graph) and NLL (bottom graph) as a function of the penalty factor for each Cm grid. The vertical line is positioned at a penalty factor of 5.

4.3.1 Results Spatial Cm calibration

Figure 45 shows the effect of the penalty factor value of 5 on all the prior Cm grids. The top row in this figure shows the posterior cross-plots of the modelled displacements against measured displacements with the RMS value and standard deviation mentioned in the title of each plot. For comparison to the posterior results on the third row, the second row shows the prior Cm grids. The bottom row presents the cross-plots of the inverted Cm values versus the initial Cm values. It can be concluded that the RMS values of both the posterior uniform and porosity-based Cm grids are close to each other with the porosity-based grid having a somewhat smaller value of the standard deviation. Also, it is concluded that the Cm patterns that are visible in the prior Cm grids are reflected by the posterior Cm grids. The aim of honouring these patterns has been successfully delivered in this step 3.

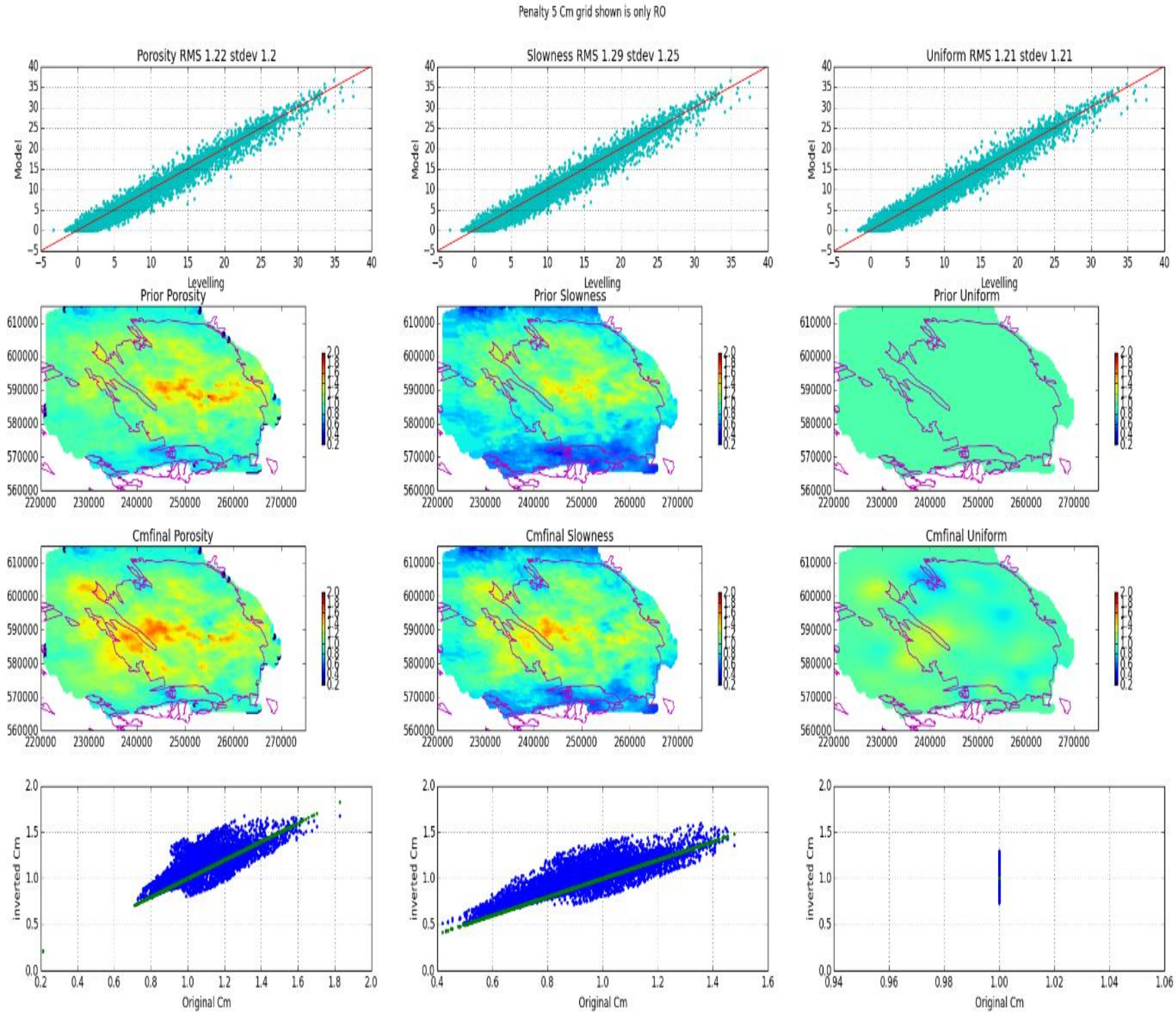


Figure 45; Results for the spatial Cm inversion using a value for the penalty factor of 5. The columns indicate the prior Cm grid. The top row shows a cross-plot of the measured and modelled subsidence. The second row the prior Cm grids, the third row the resulting spatial Cm grids and the bottom row the original vs the inverted Cm values.

4.4 Step 4, improve the temporal fit to the data, by adjusting the RTCiM parameter values

The values for the uncertainty and RTCiM parameters from step 1 were used in all previous steps. In step 4, a new iteration based on the MCMC methodology is performed to recalculate these parameters to ensure the optimal spatial (step 3) and temporal fit to the data. In addition, also the benchmarks above the aquifer areas are used in step 4. The red boundary shown in Figure 26 defines the calibration area of step 4.

The remaining input in the calculation consists of the posterior Cm grids that were calculated in step 3 and the pressure grids from the pressure realisation that was selected in step 2.

The histograms of the parameter values are shown in Figure 46. This figure presents both the histograms of the parameter values found in step 1 and the newly obtained parameter values in step 4. It is concluded that the values of the RTCiM (A,d,b) obtained in this step are close to the values found step 1. However the distribution is in general smaller and higher after step 4 with a slightly higher value for the A and d that affect the $\mathrm{C}_{\mathrm{m}_{\mathrm{ref}}}$ and $\mathrm{C}_{\mathrm{m}_{\mathrm{d}}}$ and a slightly higher value of the parameter b pointing to a slower plastic strain response of the compaction to the pressure change.

The fit to the measurements is significantly improved after step 4 with a large reduction of both the σ_0 and α.

The results after step 4 are also shown in Table 5. The pre and post NLL values cannot be compared with each other as the calibration in step 4 uses a larger area and therefore more benchmarks. From these results we conclude that the NLL value for the porosity-based Cm grid is lowest. This Cm grid will be used in the final step 5 where the forecasts will be presented.

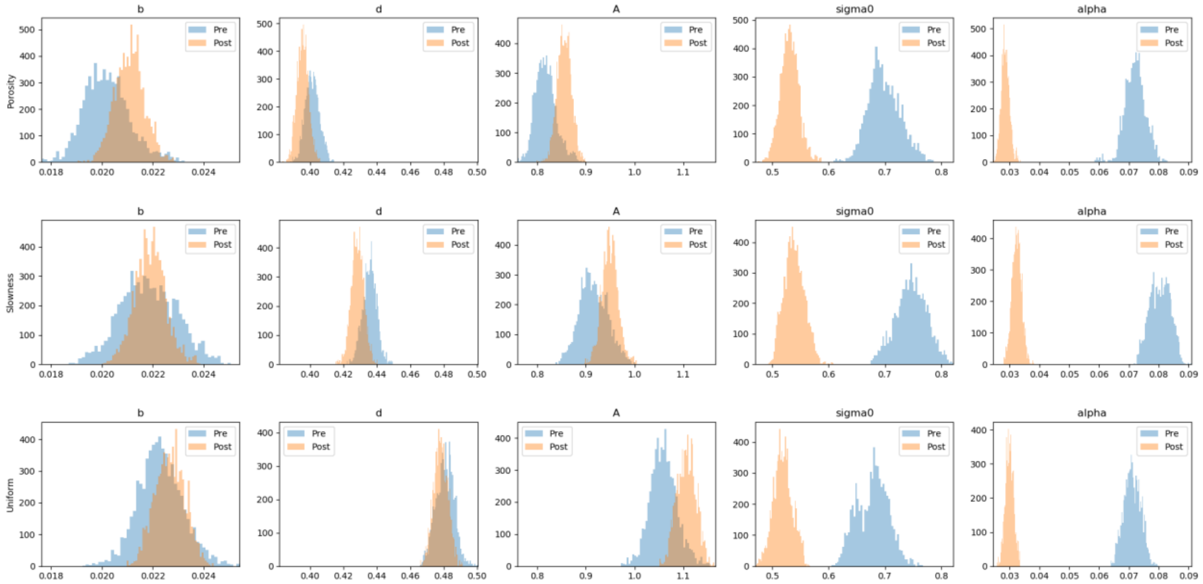


Figure 46; Histogram of the resulting RTCiM and uncertainty parameters. The new parameters after step 4 (Post) are compared with the parameters found in step 1 (Pre).

Table 5; Mean values for the RTCiM and uncertainty parameters after step 4 with in brackets the results of step 1

	Cm grid		
Solved parameters	Uniform	Slowness	Porosity
A [-]	1.11 (1.06)	0.95 (0.90)	0.85 (0.81)
d [-]	0.48 (0.48)	0.43 (0.43)	0.40 (0.40)
b [-]	0.023 (0.022)	0.022 (0.021)	0.021 (0.020)
σ_0 [cm^2]	0.52 (0.67)	0.54 (0.74)	0.53 (0.70)
α [cm]	0.03 (0.071)	0.032 (0.081)	0.029 (0.072)
NLL [-]	2780 (2555)	2894 (2800)	2725 (2613)

The selected subsidence scenario uses the selected pressure profile from step 2, the porosity-based Cm grid that was defined in step 3 and the RTCiM and uncertainty parameter values retrieved in step 4. A coverage plot is also calculated for this scenario and shown by the left plot in Figure 47. The coverage is now 92%, a value that improved when compared to the results in step 1. More important is the improvement of the coverage per benchmark (map on the right in Figure 47). No clusters of benchmark with a poorer coverage are observed. The adjusted Cm grid in step 3 contributed most to this improvement compared to a similar map presented in Figure 32, showing the coverage calculated in step 1.

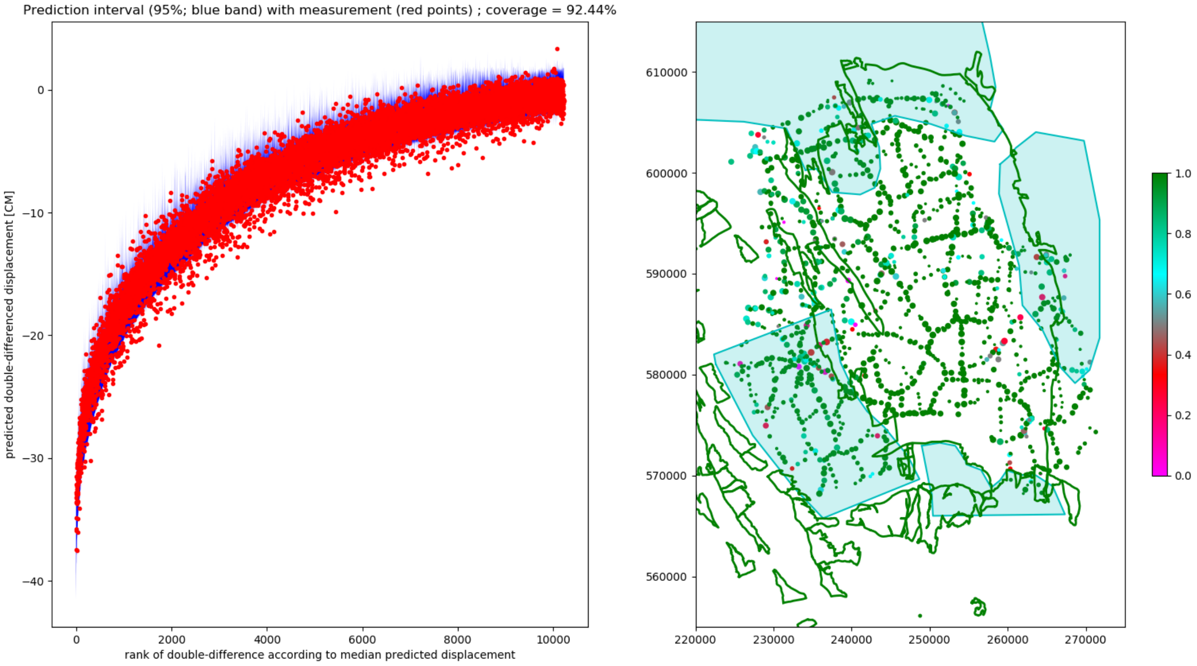


Figure 47; Prediction interval and coverage (left graph). Right picture: coverage per benchmark. The size of the dot indicates the number of double differences that are linked to the benchmark. The colour indicates the fraction of double differences that are inside the prediction interval. Dark green means that all double differences linked to the specific benchmark are inside the prediction interval.

The standard deviation for the model uncertainty is the square root of the variance Σ_{emp} , and shown in Figure 48.

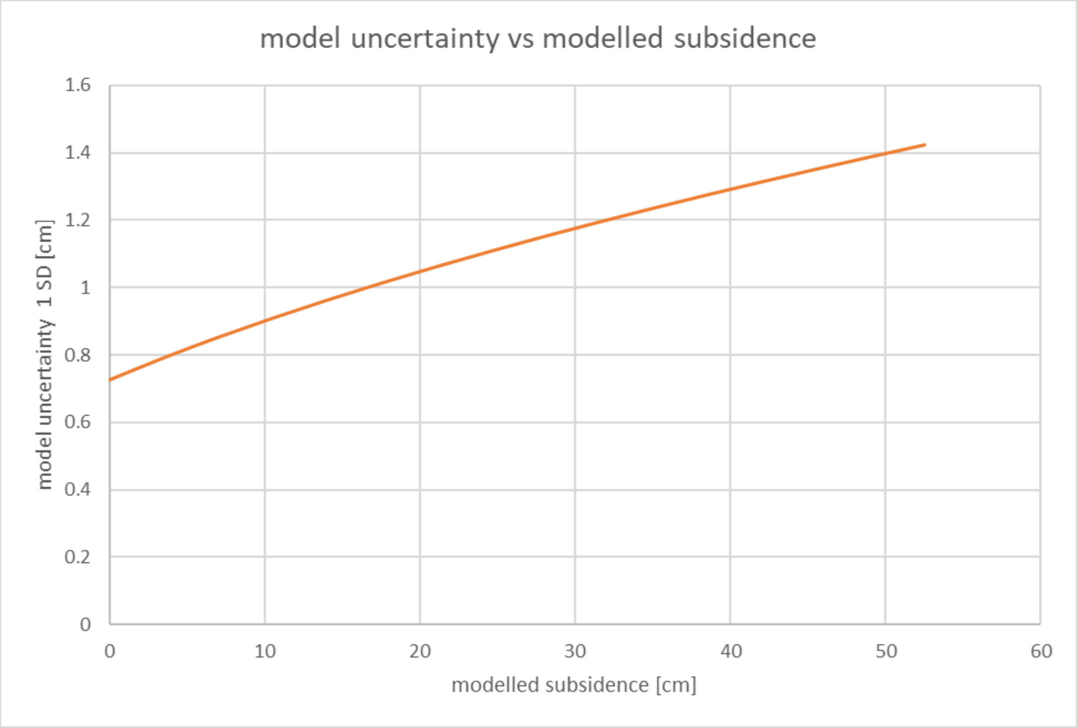


Figure 48; model uncertainty vs modelled subsidence

To visualise the match between modelled and measured subsidence over time since the first levelling surveys, several plots are presented in Figure 49 for selected benchmarks above the gas field. The plots show modelled and measured subsidence at various benchmark locations over the Groningen field. Figure 50 shows the time series for a selection of benchmarks above the aquifers. The figures include the InSAR time series (median values from measurement points in a 500 m radius) from the subsequently overlapping satellite missions Envisat, Radarsat-2 and Sentinel-1.

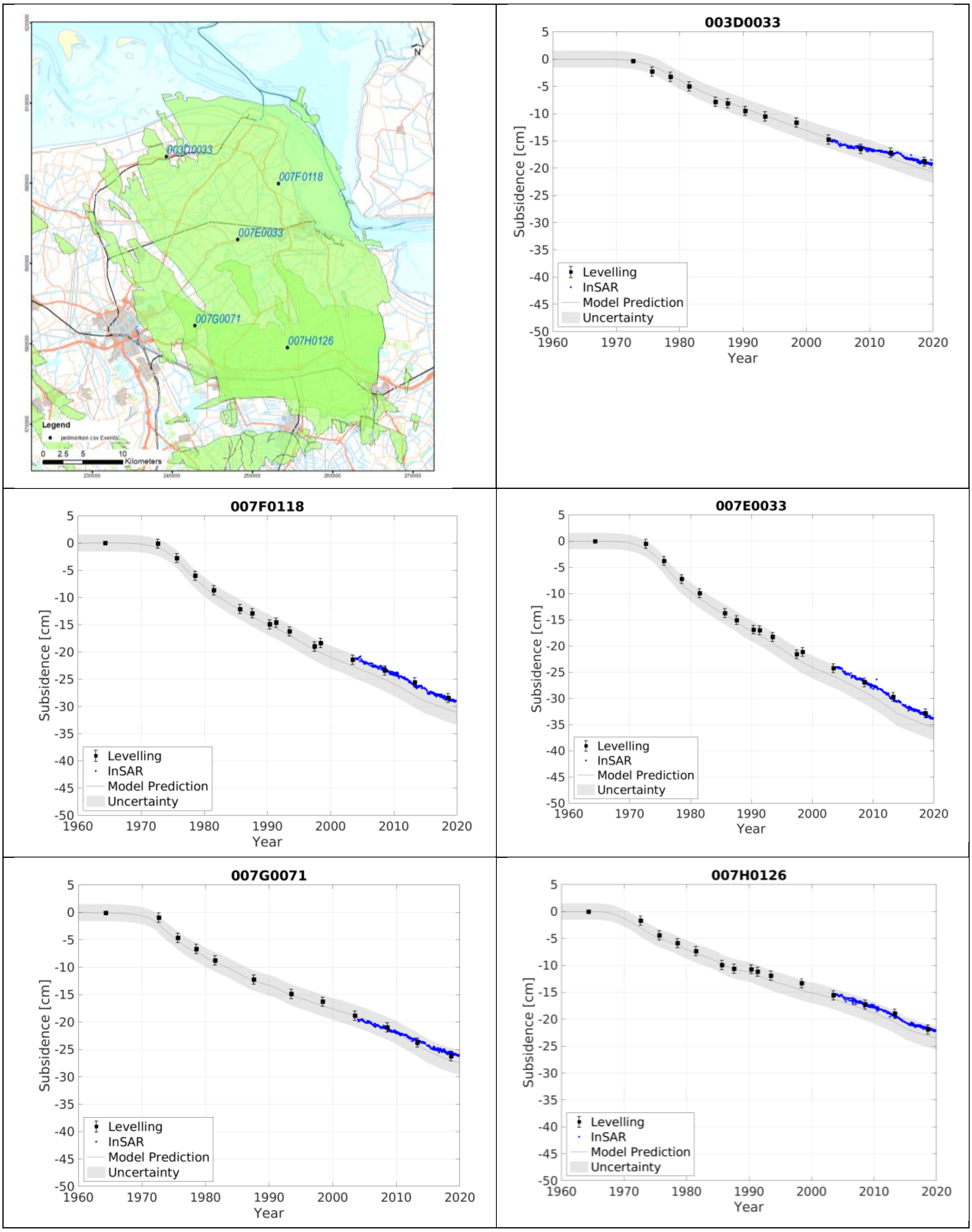


Figure 49; Subsidence at benchmark locations: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty, the blue dots are the InSAR measurements.

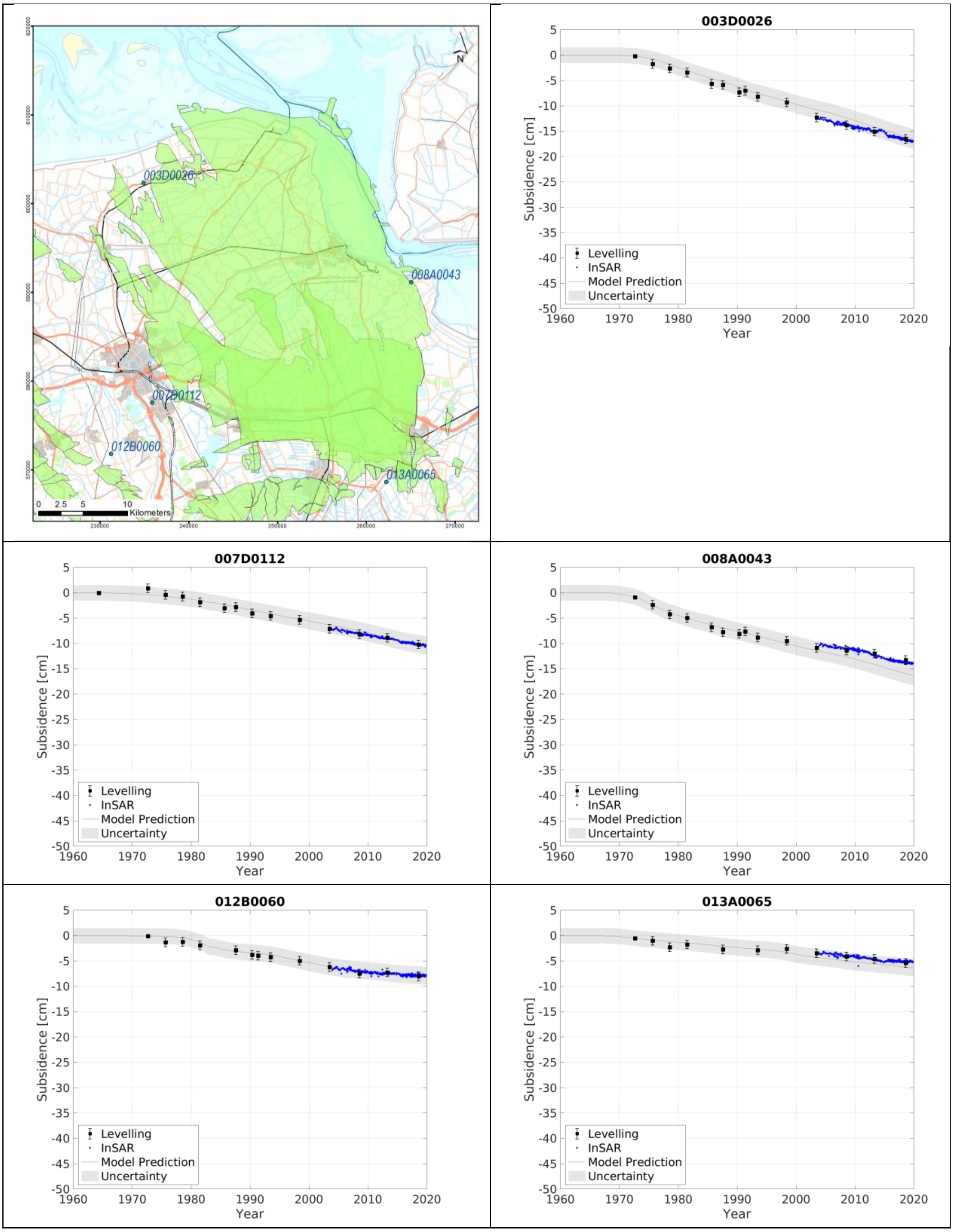


Figure 50; Subsidence at benchmark locations above the aquifers: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty, the blue dots are the InSAR measurements.

Similar plots, but for a shorter time interval, of the modelled subsidence compared to the geodetic measurements have been prepared for the GNSS points above and around the Groningen field. The map shown in Figure 51 presents the locations of these points. The abbreviations in the labels refer to the time plots in subsequent figures. A selection of points above the field is presented in Figure 52, while the locations above the aquifers are presented in Figure 53. All time series plots at the GNSS locations can be found in Appendix A. The figures include the InSAR measurements and have been used to validate the results of the calibration workflow. The InSAR, GNSS and levelling measurements match within the accuracy levels of the techniques, and hence have a similar match to the modelled subsidence. More information on the use of the InSAR technique for subsidence monitoring and the agreement with other geodetic measurement techniques can be found in Ref 32 and Ref 33.

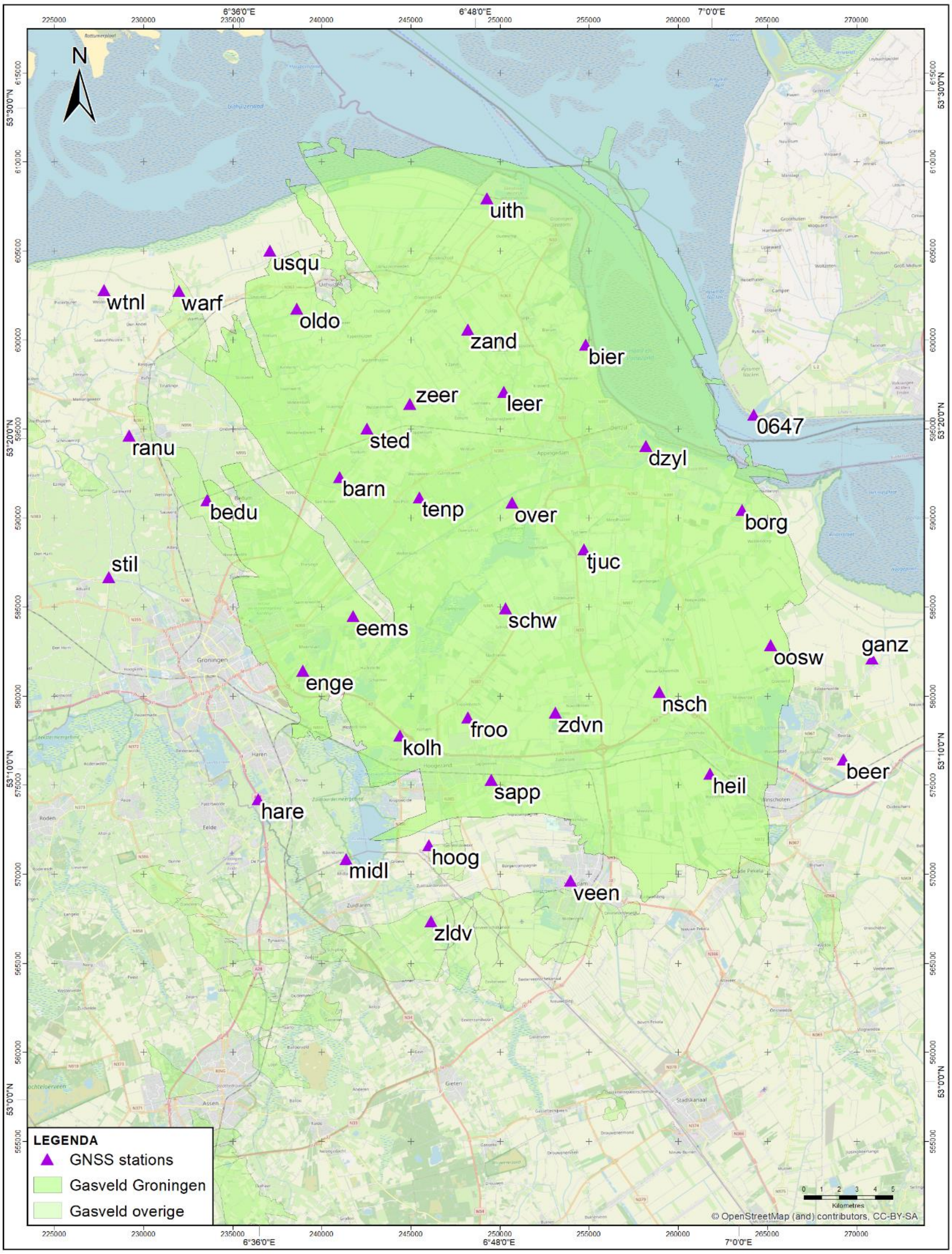


Figure 51; Overview of the GNSS stations above and around the Groningen gas field.

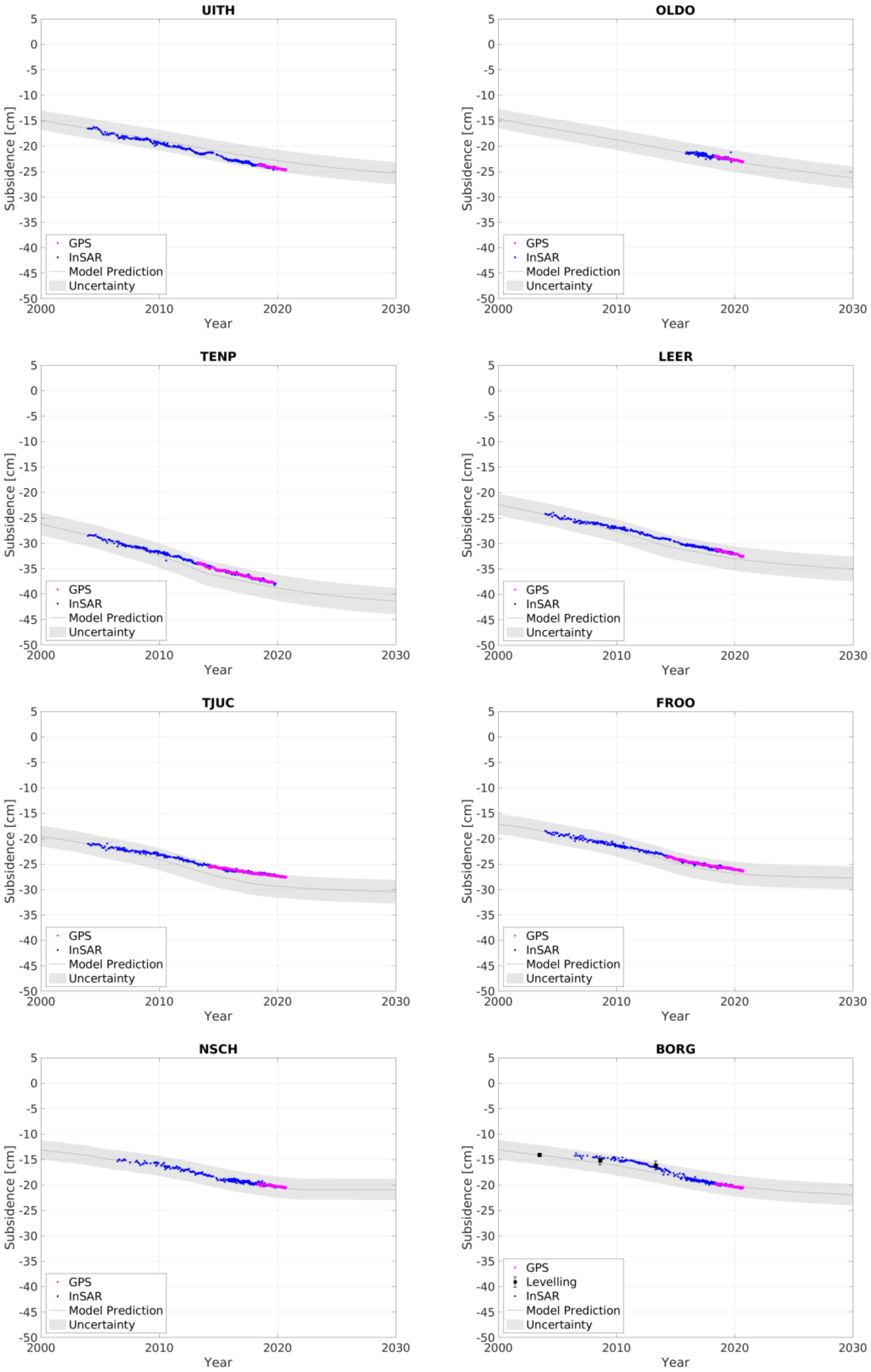


Figure 52; Subsidence on locations above the Groningen gas field: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty (95% confidence interval), the blue dots are the InSAR measurements.

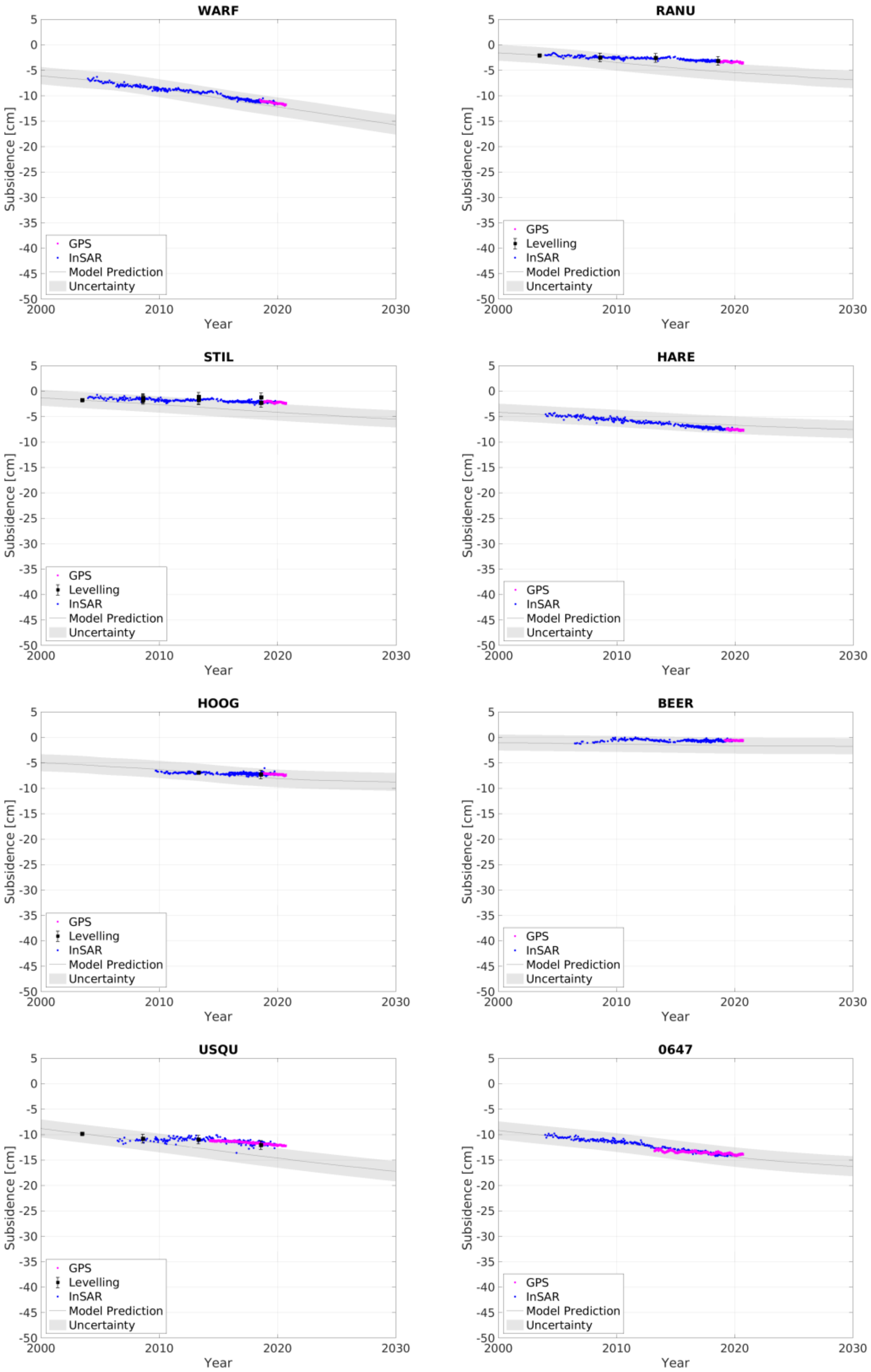


Figure 53; Subsidence on locations above aquifers: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty (95% confidence interval), the blue dots are the InSAR measurements.

It is concluded that most of the measurements are within the model bandwidth. The selected subsidence scenario that is calibrated with the levelling data is validated by the GNSS and InSAR data.

The map in Figure 54 shows the contours for the modelled subsidence between 1972 and 2018 for this scenario. The contours can be compared to the measured values on the benchmarks, shown by the labels.

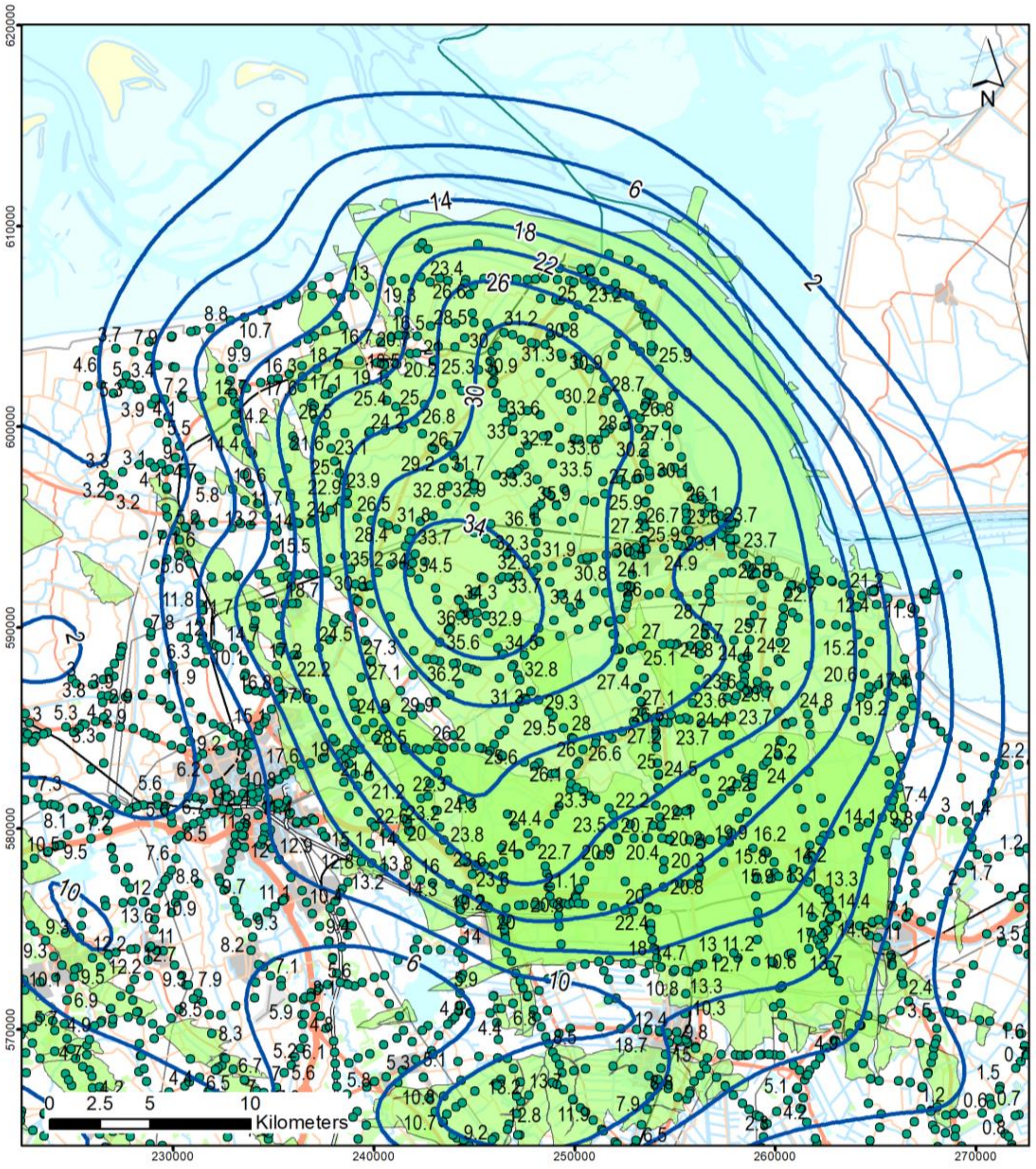


Figure 54; Contours (solid blue line) of the modelled subsidence between 1972 and 2018 compared to the measurements (green dots with value label) spanning the same period. All values are in cm.

4.5 Step 5, subsidence forecasts

In this section the current status and expected development of subsidence up to 2080 is presented using the subsidence scenario and its uncertainty from step 4. Subsidence caused by gas production and aquifer depletion from the Groningen field is combined with the effects from ongoing gas production from neighbouring fields as published in the subsidence “status rapport” for these fields (Ref 34). Most of the subsidence has already occurred over the last 60 years. Subsidence forecast for the years 2030, 2050 and 2080 are presented in Figure 55, Figure 56 and Figure 57. The contours present the forecasted subsidence while the colours represent the P95 uncertainty. This uncertainty results from the Σ_{emp} as derived in step 4 and basically predicts that 95% of the future mean values of the measured subsidence falls in this interval.

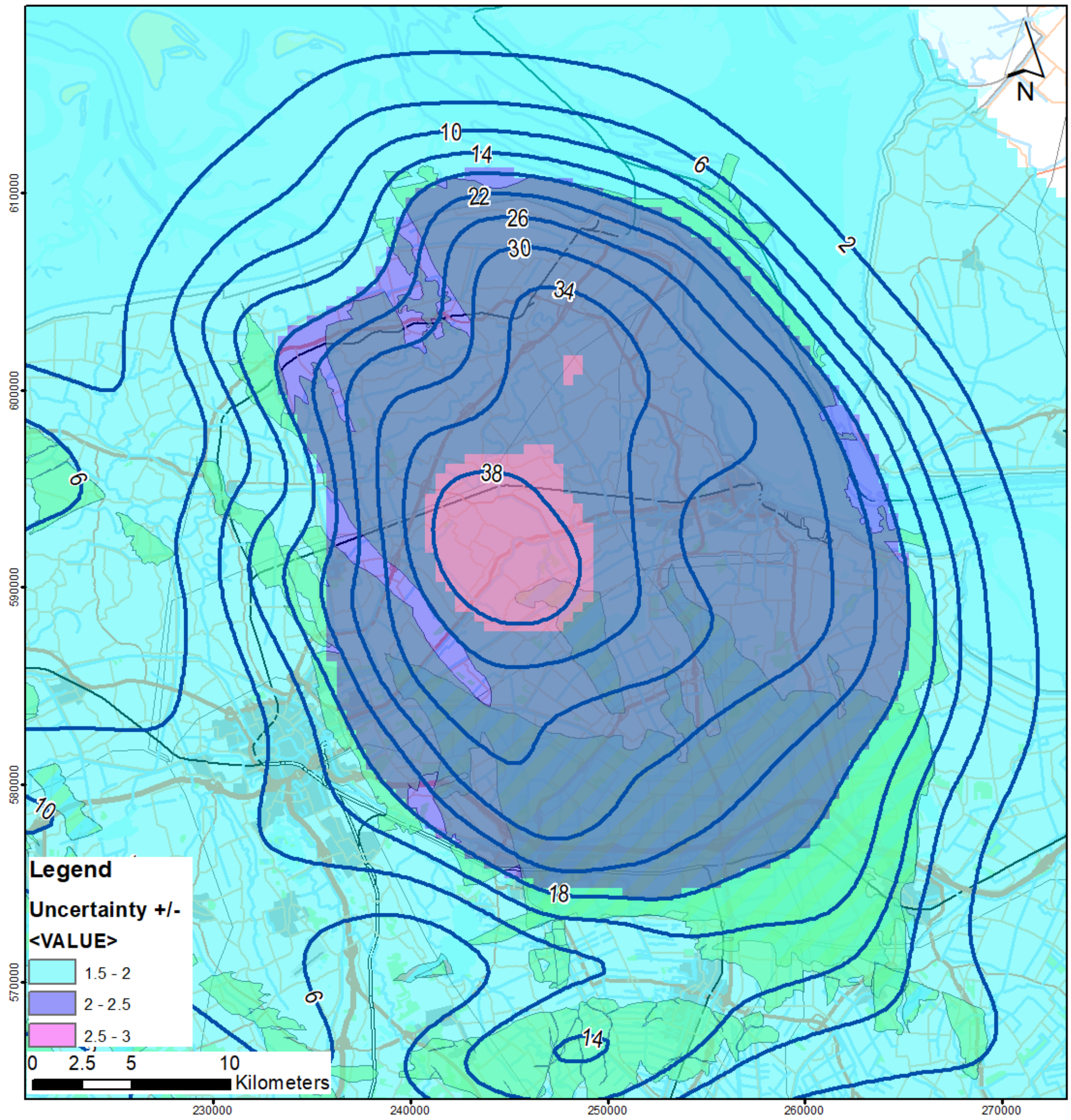


Figure 55; Forecasted subsidence in cm for 2030 (contours). The colours indicate the P95 uncertainty (± in cm).

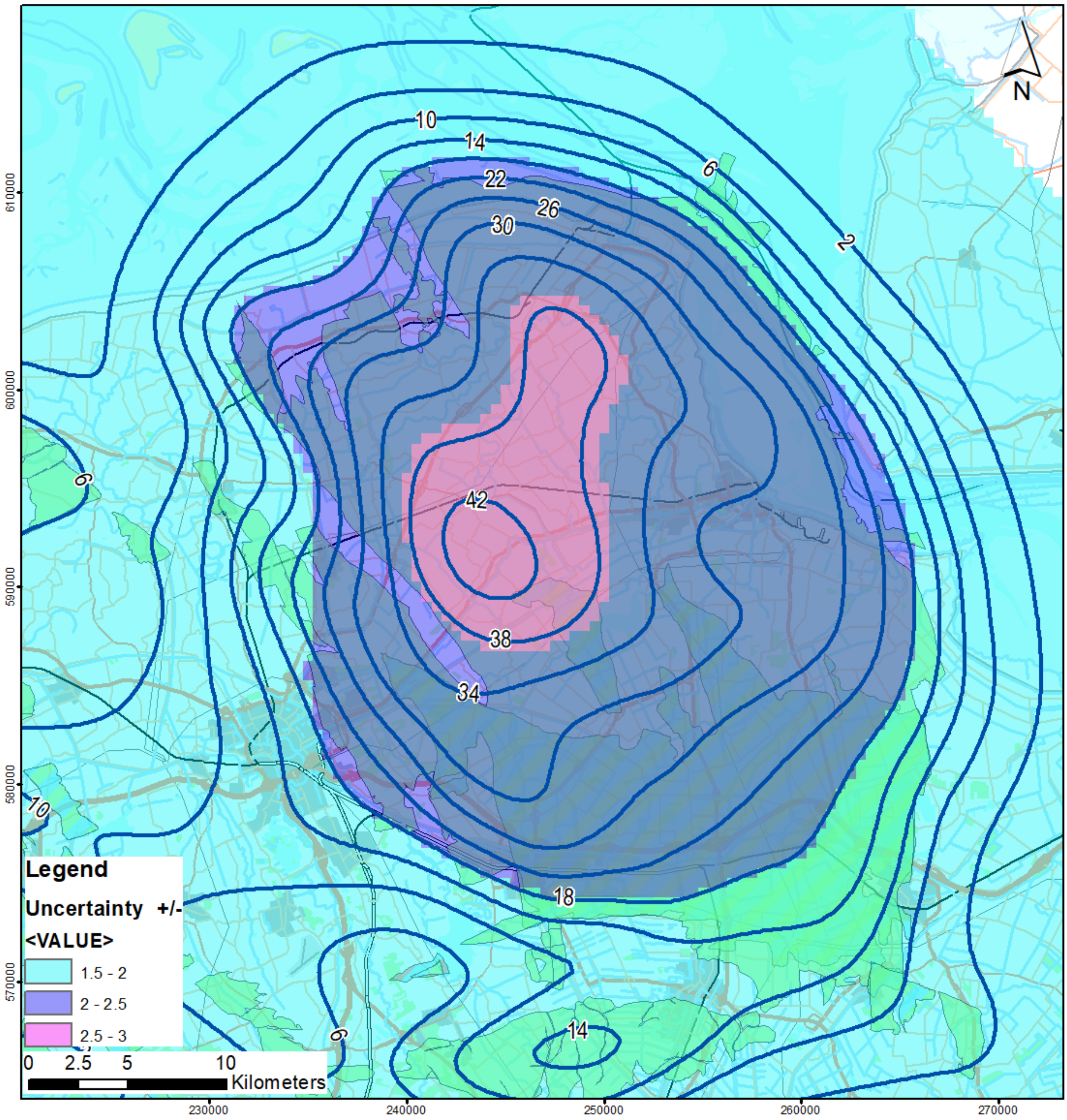


Figure 56; Forecasted subsidence in cm for 2050 (contours). The colours indicate the P95 uncertainty (± in cm).

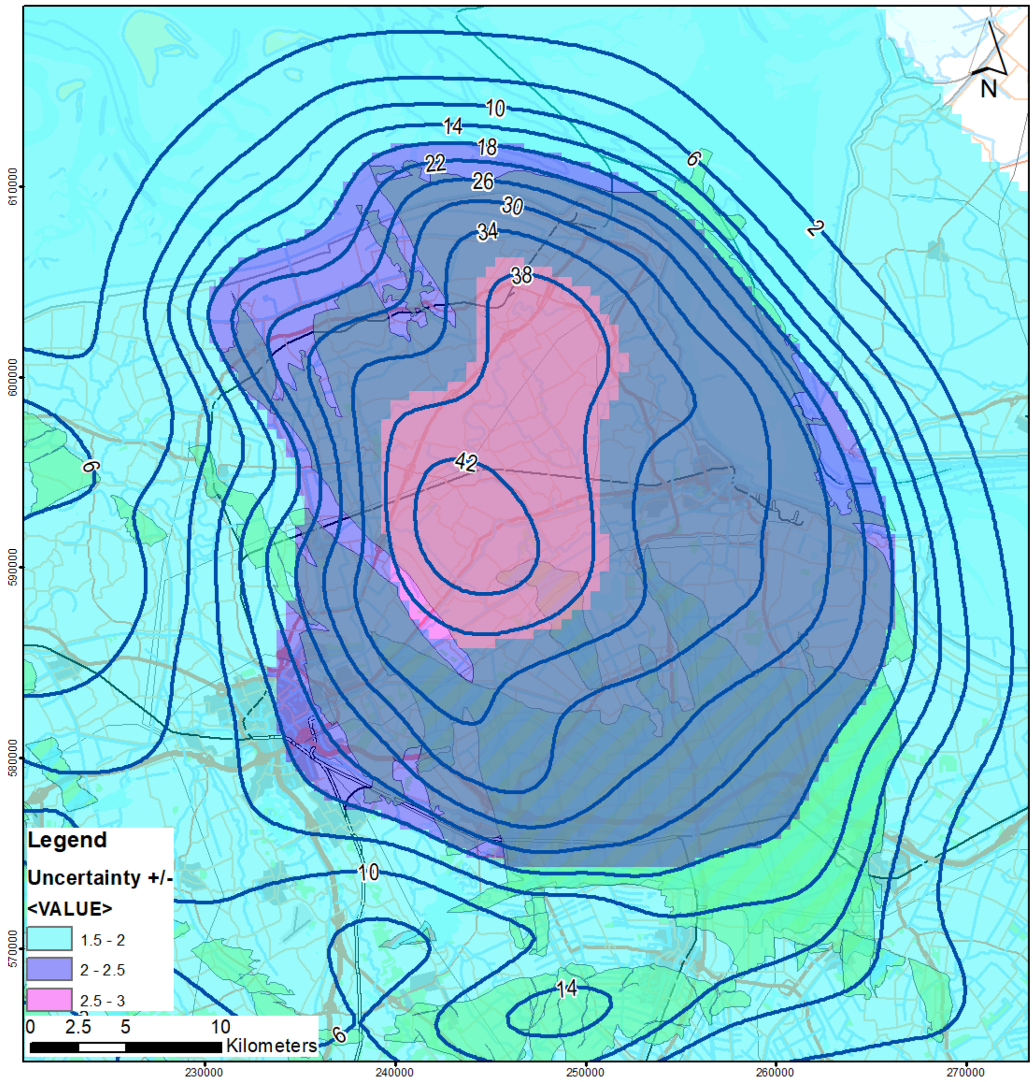


Figure 57; Forecasted subsidence in cm for 2080 (contours). The colours indicate the P95 uncertainty (± in cm).

Finally, Figure 58 and Figure 59 show both the historical and forecasted subsidence till 2080 for the selected benchmarks above (Figure 58) and around the Groningen gas field (Figure 59).

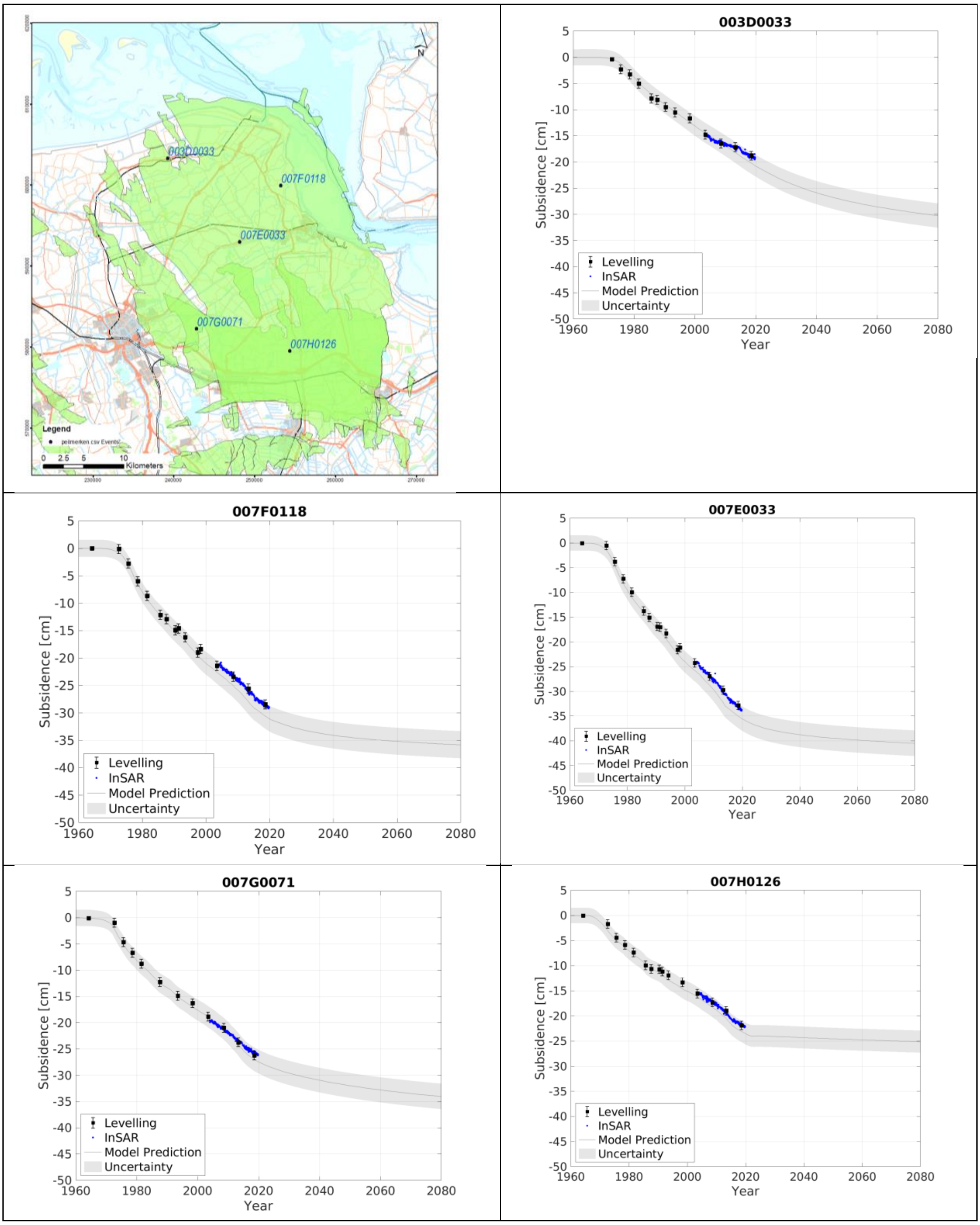


Figure 58; Subsidence at benchmark locations till 2080: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty, the blue dots are the InSAR measurements.

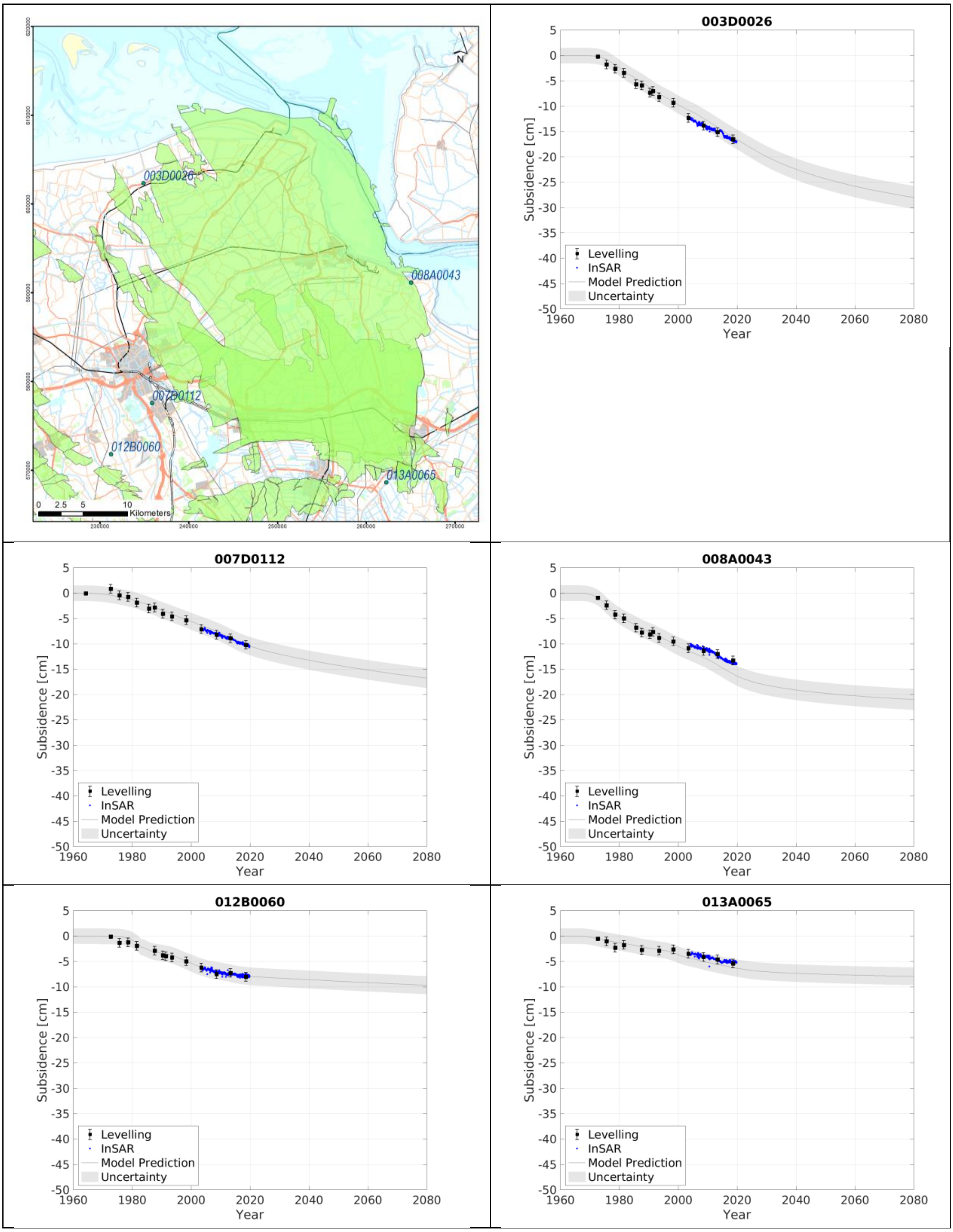


Figure 59; Subsidence at benchmark locations above the aquifers till 2080: dark grey line is the predicted subsidence, grey is the P95 confidence interval, black dots are levelling measurements plus uncertainty, the blue dots are the InSAR measurements.

5 Conclusions

A statistical method is adopted that investigates the compaction and subsidence in and above the Groningen gas field and its laterally connected aquifers with 3 major objectives: (1) invert for the most probable aquifer realisation and a set of RTCiM values, (2) test multiple correlations to the rock compressibility and (3) calculate the model uncertainty. The method uses a Bayesian stochastic approach via Markov Chain Monte Carlo.

The inversion for all parameters is a computer intensive process and technically not feasible for all possible subsidence scenarios. Therefore a 5-step workflow is set up to deliver the above described objectives at reasonably low computation times. After the execution of the first four steps we retrieved:

- the most likely reservoir realisation,
- a set of spatially recalibrated Cm grids that still reflect the compressibility trends as shown by the prior Cm grids,
- RTCiM parameter values that temporarily and spatially best fit to the subsidence measurements,
- A quantification of the uncertainty driven by the model.

Using this subsidence scenario and its uncertainty, subsidence forecasts (step 5) were made based on the latest production scenarios.

6 References

Ref 1: Bierman S., & Towe, R. (2020) Statistical methodology for fore casting of subsidence above the Groningen gas field For the Rate Type Compaction isotach Model (RTCiM) for reservoir compaction. Shell report SR.20.00973

Ref 2: De Jager, J., & Visser, C. (2017) Geology of the Groningen field – an overview. Netherlands Journal of Geosciences, 96(5), S3-S15. doi:10.1017/njg.2017.22

Ref 3: Van Eijs, R., & Van der Wal, O. (2017). Field-wide reservoir compressibility estimation through inversion of subsidence data above the Groningen gas field. Netherlands Journal of Geosciences, 96(5), S117-S129. doi:10.1017/njg.2017.30 https://www.cambridge.org/core/journals/netherlands-journal-of-geosciences/article/fieldwide-reservoir-compressibility-estimation-through-inversion-of-subsidence-data-above-the-groningen-gas-field/D5B31E109E9F1687770E50D447E0E43F

Ref 4: Kortekaas, M., & Jaarsma, B. (2017) Improved definition of faults in the Groningen field using seismic attributes. Netherlands Journal of Geosciences, 96(5), S71-S85. doi:10.1017/njg.2017.24

Ref 5: Pijnenburg, R., Verberne, B., Hangx, S., & Spiers, C. (2018) Deformation Behaviour of Sandstones From the Seismogenic Groningen Gas Field: Role of Inelastic Versus Elastic Mechanisms. Journal of Geophysical Research: Solid Earth. 10.1029/2018JB015673.

Ref 6: 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.

Ref 7: Geertsma, J. & 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.

Ref 8: NAM (2020) Seismic Hazard and Risk Assessment Groningen Field update for Production Profile GTS - raming 2020. NAM-report EP202003201727

Ref 9: Ministerie van Economische Zaken en Klimaat (2019) “Definitief vaststellingsbesluit Groningen gasveld 2019-2020,". https://www.rijksoverheid.nl/onderwerpen/gaswinning-in-groningen/documenten/publicaties/2019/09/10/bijlage-6---definitief-vaststellingsbesluit-groningen-gasveld-2019-2020.

Ref 10: NAM (2018) Production Optimisation 2018.

Ref 11: NAM (2020) Operationele Strategieën voor het Gasjaar 2020-2021, EP202002207545

Ref 12: NAM (2018) Groningen Dynamic Model Update 2018 – V6, EP201809202872.

Ref 13: Hol, S., van der Linden, A., Bierman, S., Marcelis, F. & Makurat, A. (2018) Rock Physical Controls on Production-induced Compaction in the Groningen Field. Sci Rep 8, 7156 (2018). https://doi.org/10.1038/s41598-018-25455-z

Ref 14: NAM (1985) Bodemdaling als gevolg van aardgasonttrekking in de provincie Groningen. https://eu001-sp.shell.com/sites/AAAAB0343/Documents/Bodemdaling%20als%20gevolg%20van%20aardgasonttrekking%20in%20de%20Provincie%20Groningen,%20Herziene%20voorspellingen.pdf#search=11%2E391%20bodemdaling%20groningen

Ref 15: NAM (2016) Winningsplan Groningen Gasveld 2016. EP201604259068 https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/nl/c7f20ac1-1818-4f2f-82b0-b347fa2455aa

Ref 16: NAM (2018) Assessment of Subsidence based on Production Scenario “Basispad Kabinet”. EP201806209337. https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/5a472ec3-8983-4ced-9ddd-0f8a54c962a5

Ref 17: Schutjens, P.M.T.M., van Hasselt J.P., van Kooten, J.F.C. & Loosveld, R.J.H. (1990) Overzicht van recent onderzoek uitgevoerd op het KSEPL naar compactie en bodemdaling in Groningen. RKER.89.206 https://eu001-sp.shell.com/:b:/r/sites/AAFAA4377/Global%20Library%20Non-confidential%20Reports/RKER.89.206%20-%20Overzicht%20van%20recent%20onderzoek%20uitgevoerd%20op%20het%20KSEPL%20naar%20compactie%20en%20bodemdaling%20in%20Groningen_/RKER.89.206.PDF?csf=1&e=Prf8kD

Ref 18: NAM (2015) Bodemdaling door aardgaswinning. Statusrapport 2015 en prognose tot het jaar 2080. https://nam-onderzoeksrapporten.data-app.nl/reports/download/bodemdaling/nl/03ff38de-1870-4d6a-988b-e6bcc595f6ef

Ref 19: Van Thienen-Visser, K., & Fokker, P. (2017). The future of subsidence modelling: Compaction and subsidence due to gas depletion of the Groningen gas field in the Netherlands. Netherlands Journal of Geosciences, 96(5), S105-S116. doi:10.1017/njg.2017.10

Ref 20: Bourne, S.J., Oates, S.J., van Elk, J. & Doornhof, D., (2014). A seismological model for earthquakes induced by fluid extraction from a subsurface reservoir. Journal of Geophysical Research 119 (12): 8891–9015. doi: 10.1002/2014JB011663

Ref 21: 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

Ref 22: Pruiksma, J.P., Breunese, J.N., Thienen-Visser, K., & De Waal, H. (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.

Ref 23: NAM (2017) Ensemble Based Subsidence application to the Ameland gas field – long term subsidence study part two (LTS-II) continued study. https://nam-feitenencijfers.data-app.nl/download/rapport/2a2da56c-faee-4453-a8a7-c64c43b418a6?open=true

Ref 24: TNO (2017) Ensemble-based Subsidence Interpretation and Prediction ESIP: Technical Reference Manual. TNO report-R11278

Ref 25: Samiei-Esfahany, S. & Bähr, H. (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.

Ref 26: Houtenbos, A.P.E.M. & Kenselaar, F. (2001) Peilmerk hoogte variaties – stochastische analyse van peilmerkbeweging in Nederland. TUDelft, 2001.

Ref 27: Samiei-Esfahany S. & Bähr, H. (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.

Ref 28: NAM (2019) Stabiliteitsanalyse van waterpaspeilmerken 2019 – Toelichting methodiek en toepassing op de meetregisters "Noord-Nederland, Groningen en Waddenzee 2018" en "Zuid-Holland 2017". Nederlandse Aardolie Maatschappij, 12 juli 2019, EP20190621048

Ref 29: van Leijen, F., Samiei-Esfahany, S., van der Marel, H., & Hanssen, R (2017) A standardized approach for the integration of geodetic data for deformation analysis, 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, 2017, pp. 957-960.

Ref 30: Kole, P., Cannon, M., Tomic, J. & Bierman, S. (2020) Analysis of first four years of in-situ strain data in Zeerijp-3A. Shell report EP201908210907

Ref 31: van der Linden, A.J., Marcelis, F.H.M., Hol, S. & El Azouzi, K. (2020) Mechanical compression testing Carboniferous underburden material from the Zeerijp-3A well, Groningen Field, The Netherlands. Shell report SR.20.00670

Ref 32: NAM (2019): Meetregister bij het meetplan Noord-Nederland, Groningen en Waddenzee 2018 – Addendum InSAR. Nederlandse Aardolie Maatschappij, 19 juli 2019, EP201907205139.

Ref 33: NAM (2020): Gebruik van InSAR in de meetregisters Noord-Nederland en Groningen; Analyse beschikbaarheid en betrouwbaarheid. Nederlandse Aardolie Maatschappij, 29 mei 2020, EP202005202312.

Ref 34: NAM (2020) Bodemdaling door aardgaswinning. Statusrapport 2020 en prognose tot het jaar 2080 (in preparation).

Appendix A Subsidence time series GNSS stations

This Appendix contains the geodetic deformation timeseries in the period 2003-2020 at the GNSS stations (from 2013), including the InSAR time series (median values from measurement points in a 500 m radius) from the subsequently overlapping satellite missions Envisat, Radarsat-2 and Sentinel-1, and the subsidence model including its uncertainty.

Note: the Veendam time series is outside the subsidence model uncertainty bound due to additional subsidence resulting from salt mining activities.

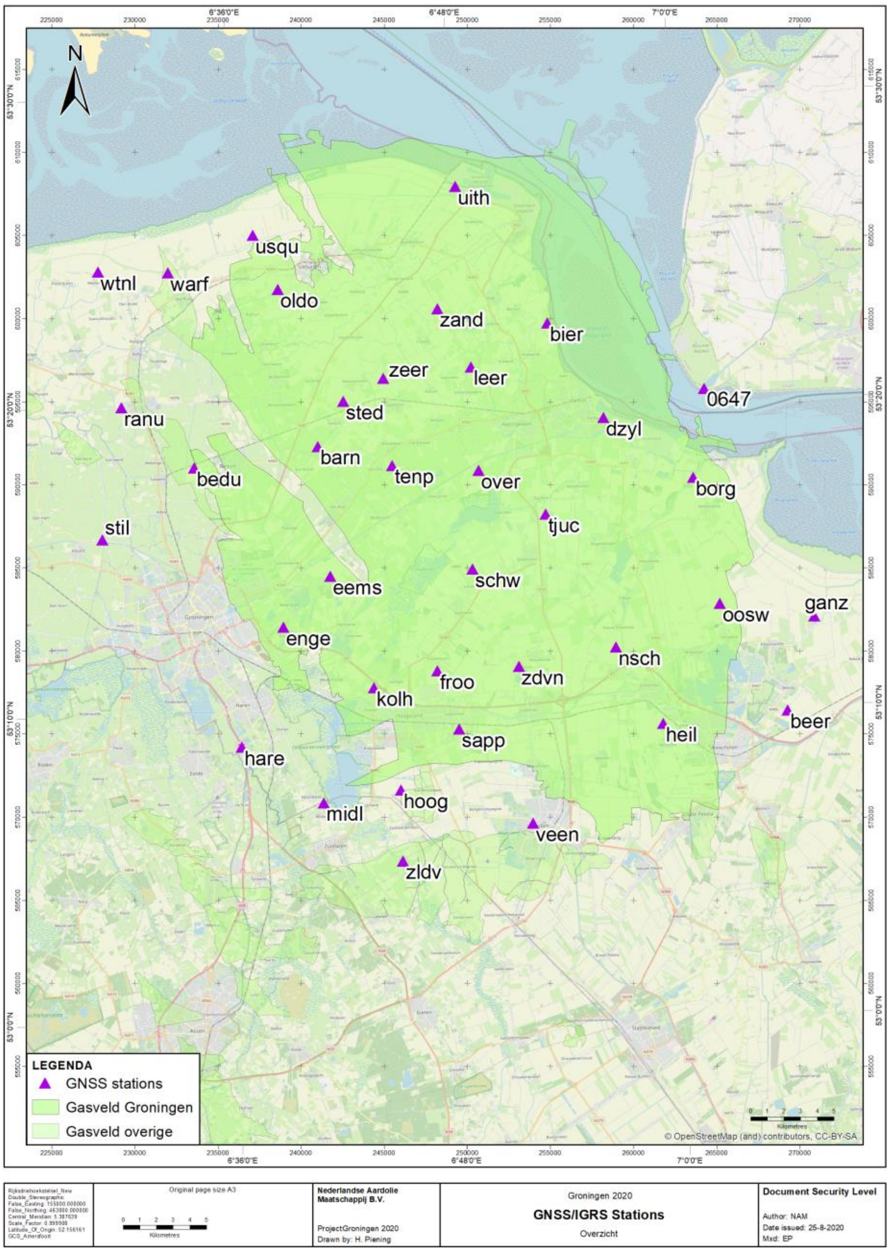
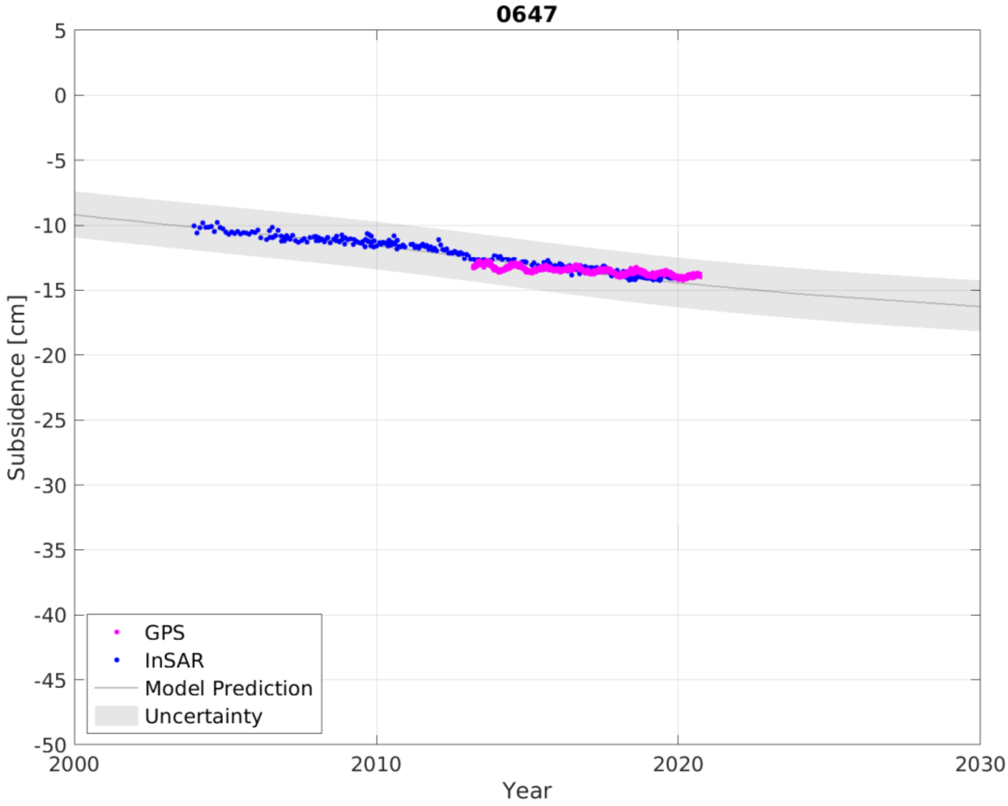
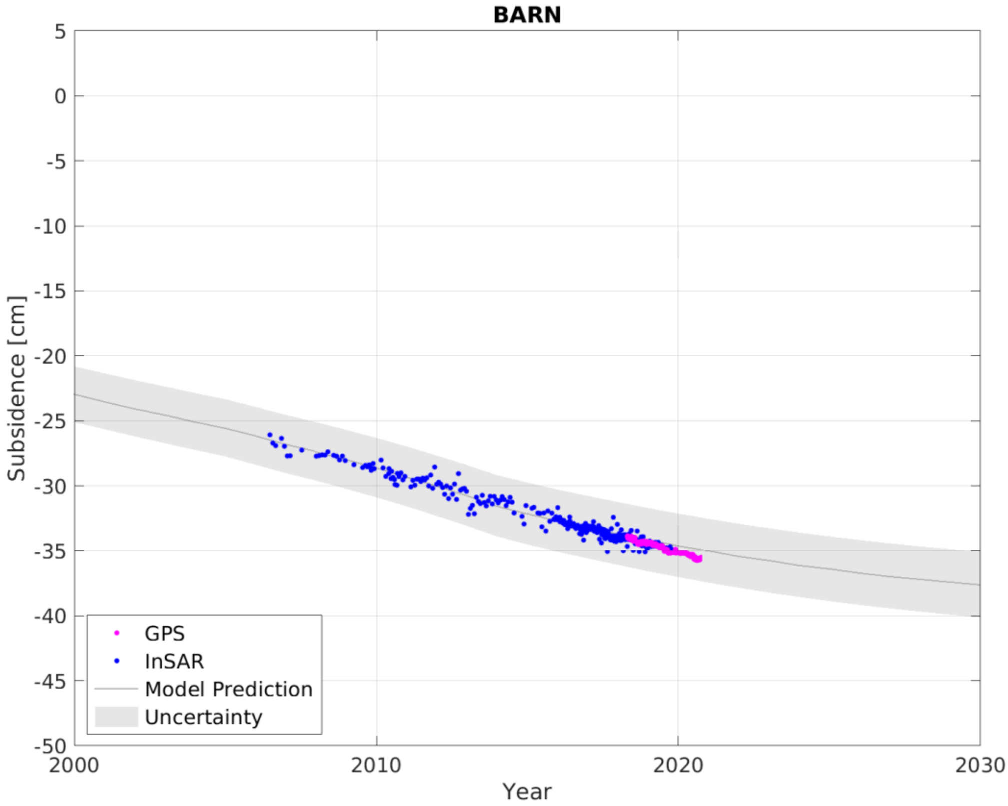
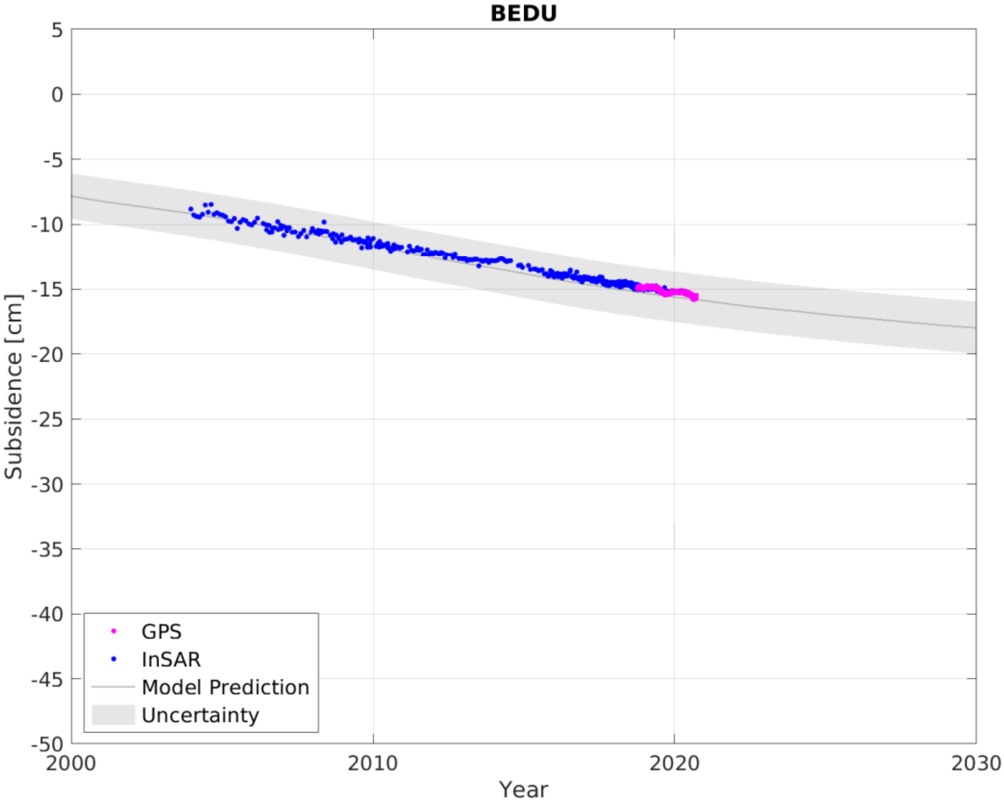
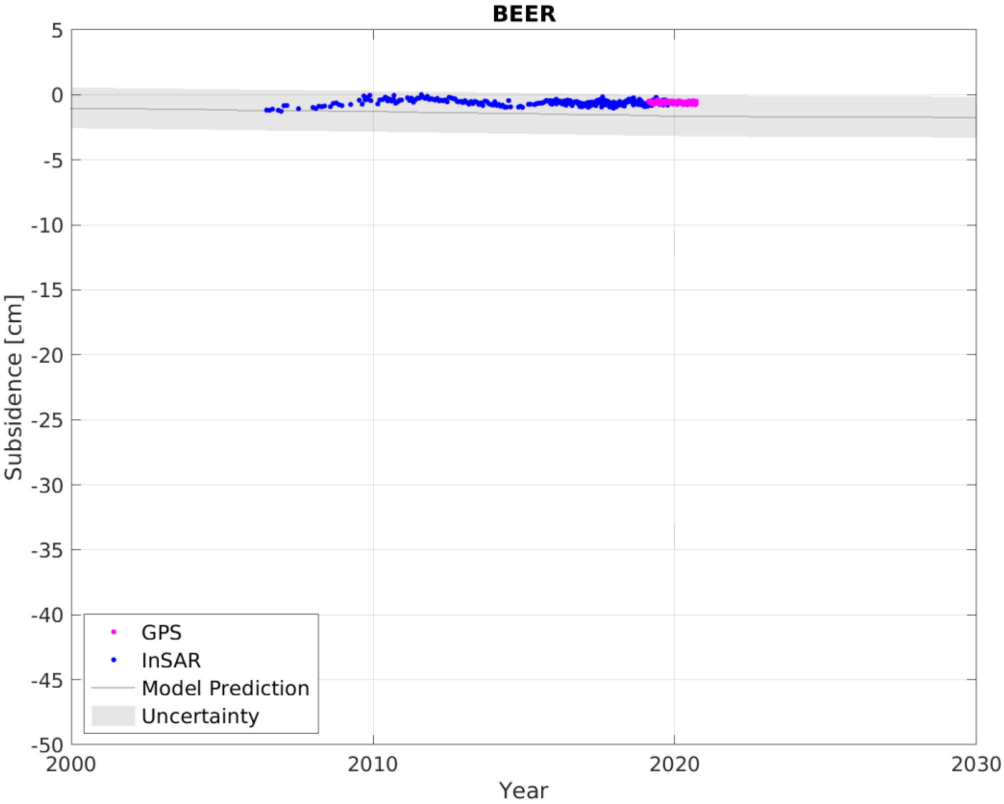


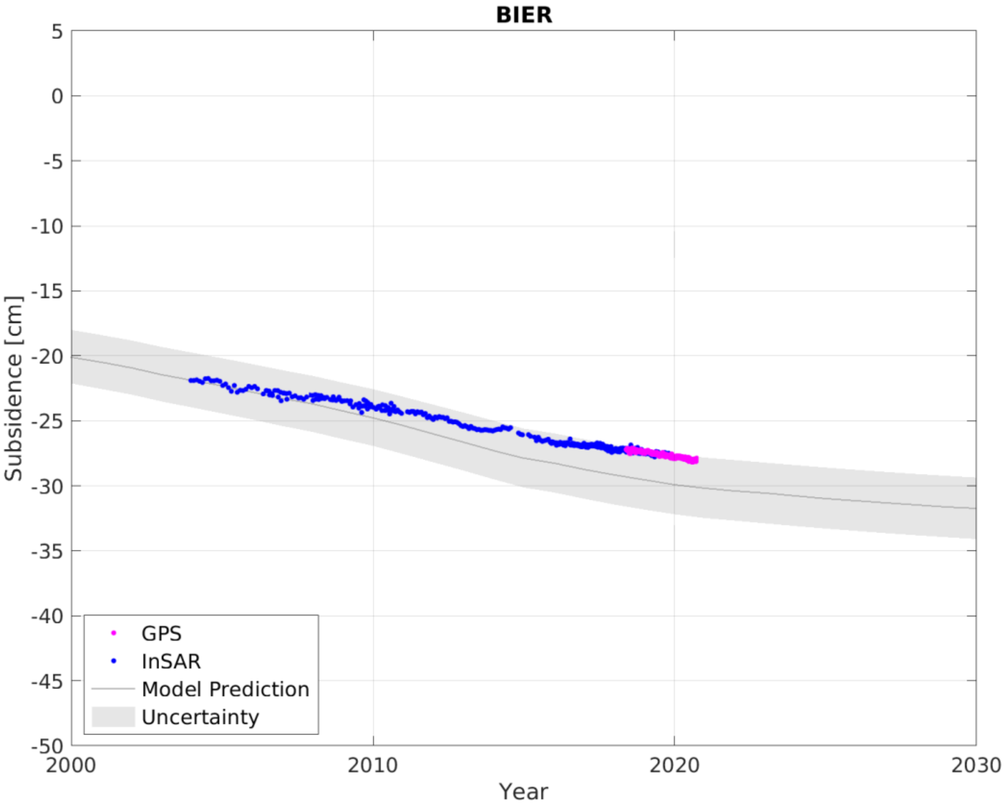
Figure 60; Overview Groningen GNSS stations.

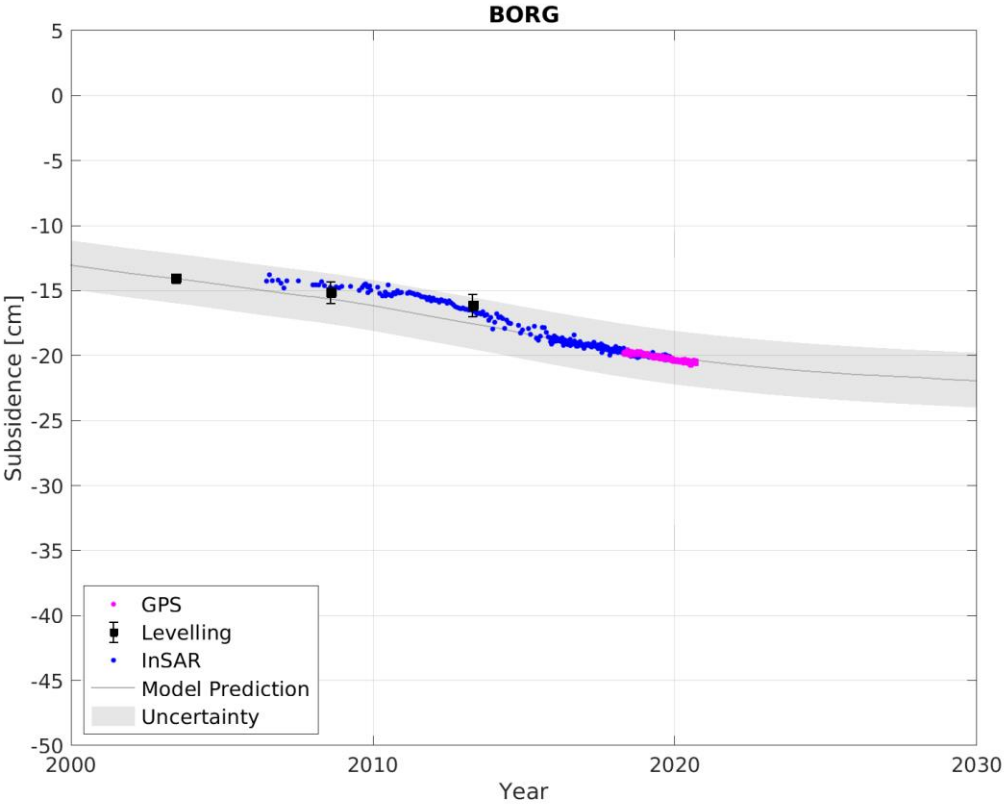

0647
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

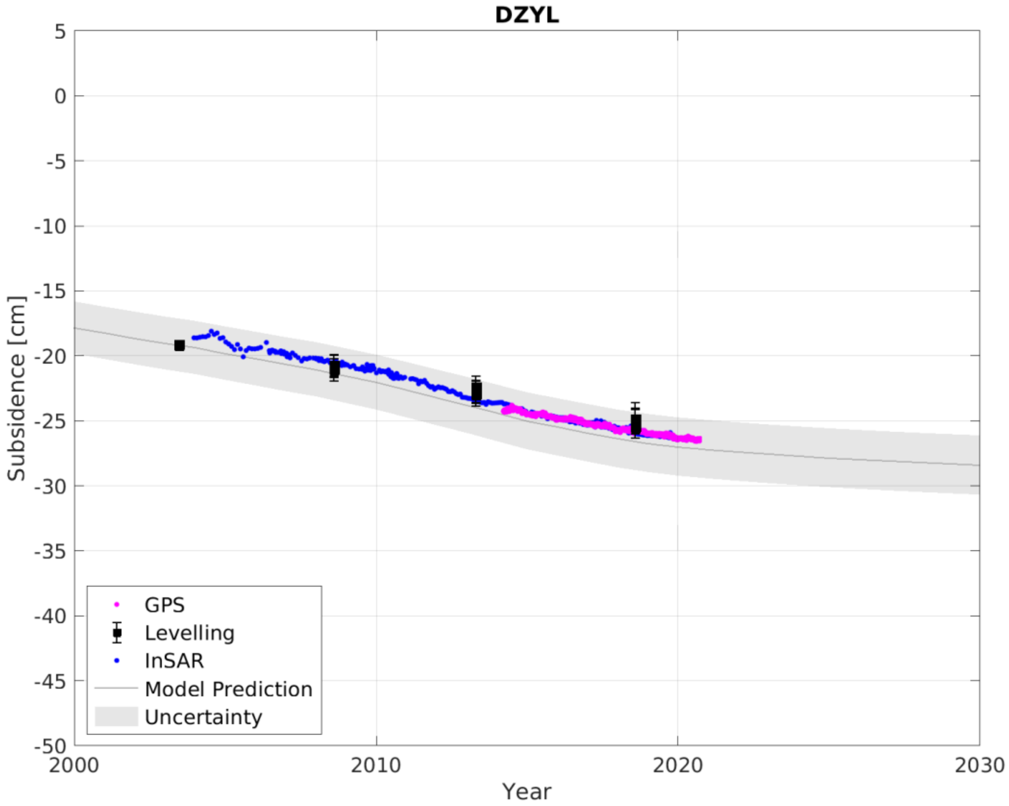

BARN
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

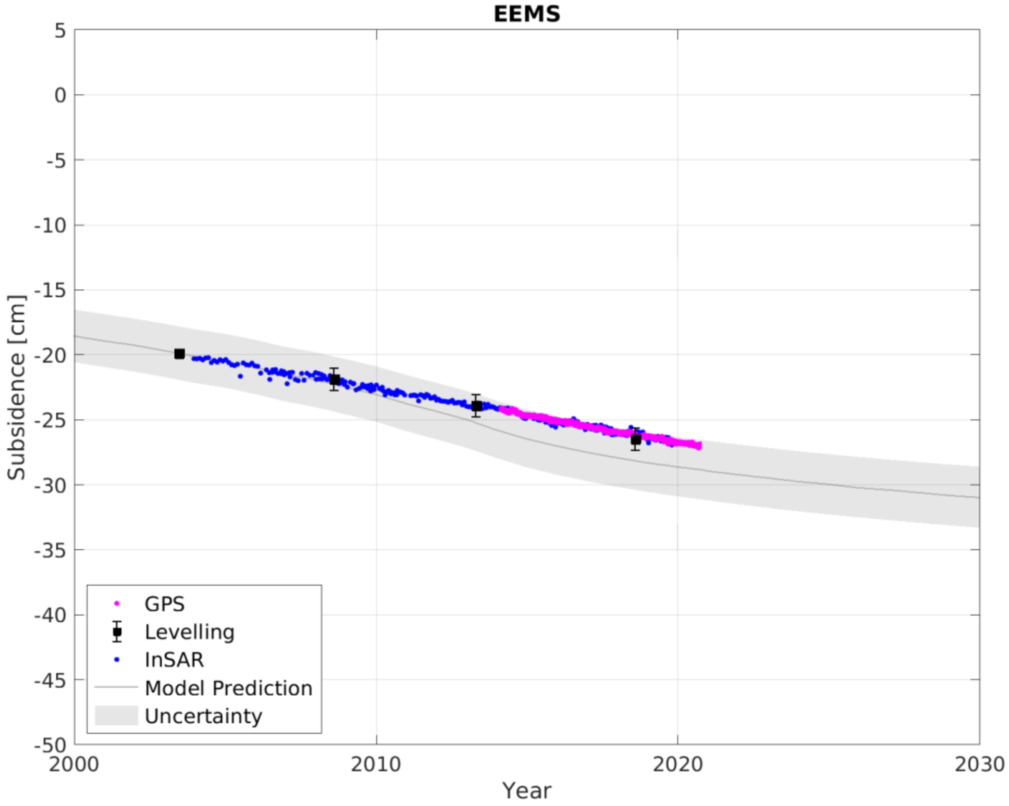

BEDU
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

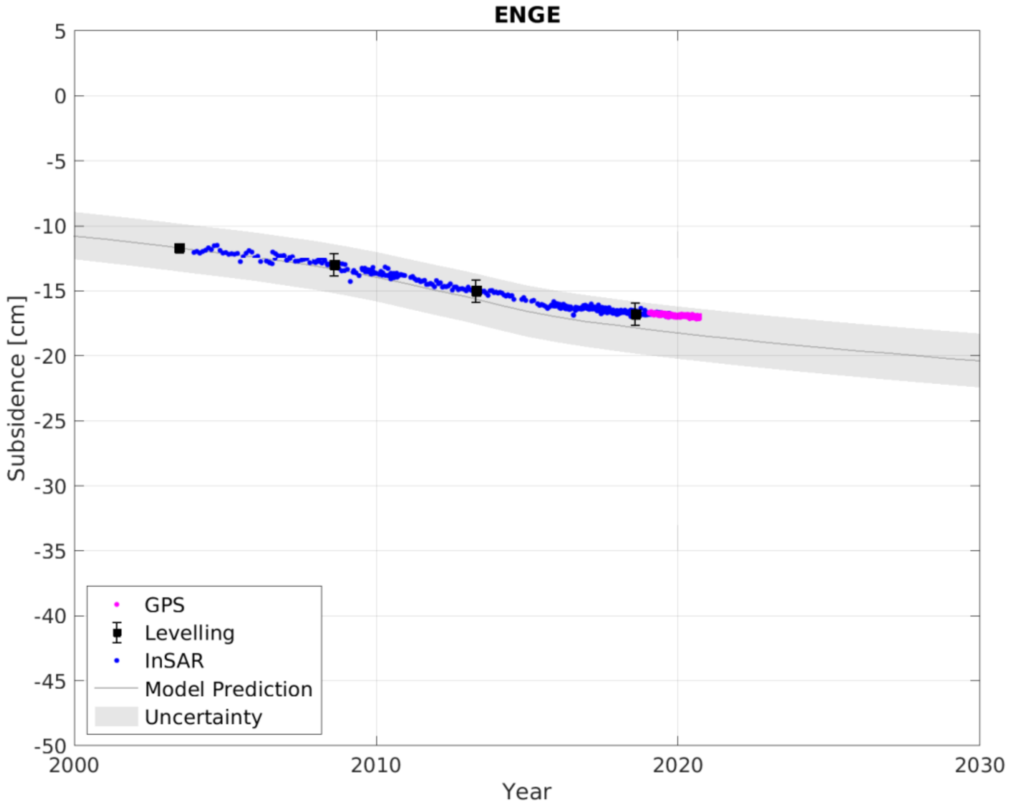

BEER
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

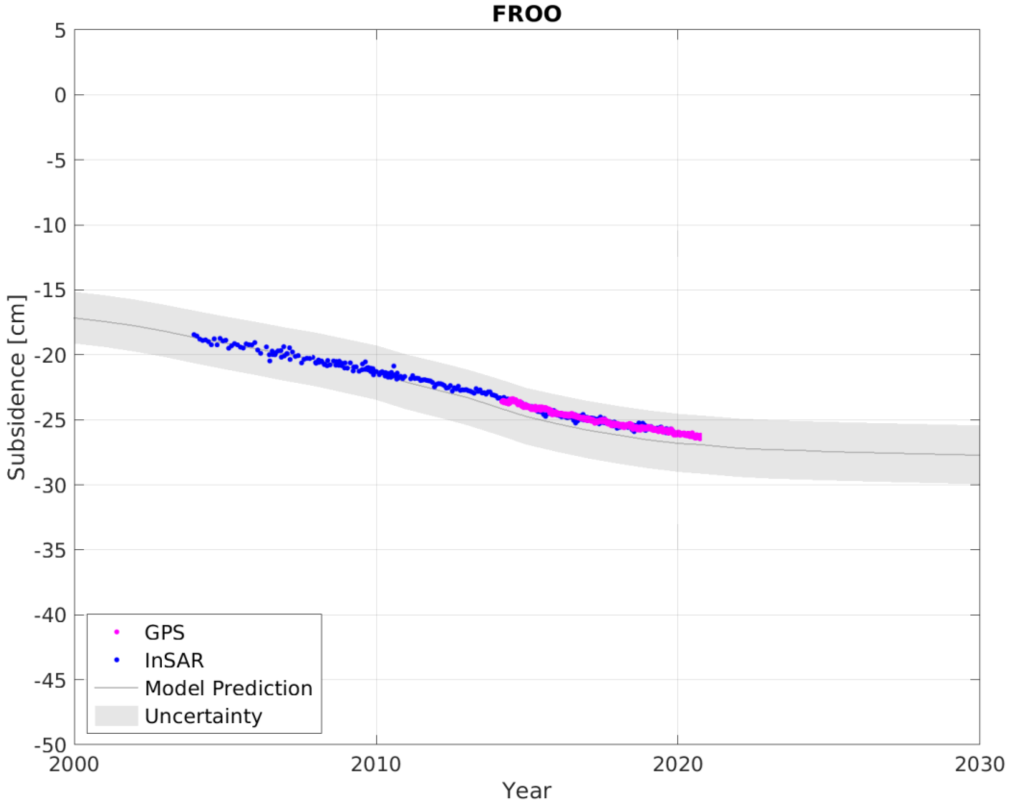

BIER
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

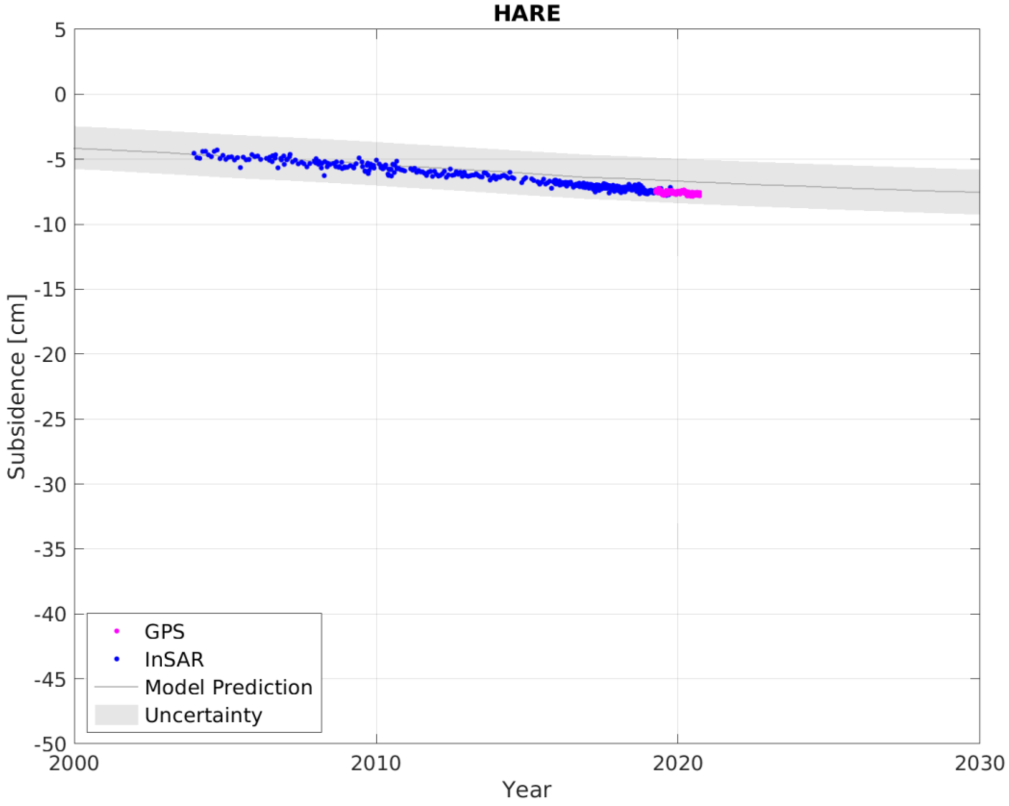

BORG
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

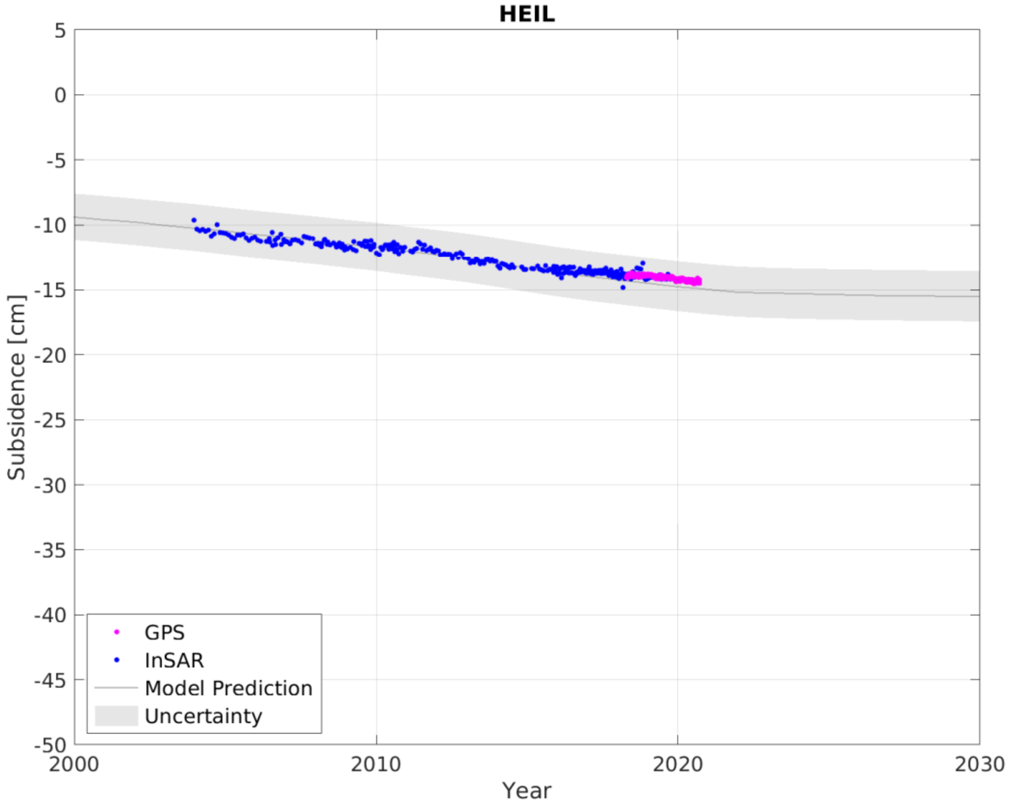

DZYL
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

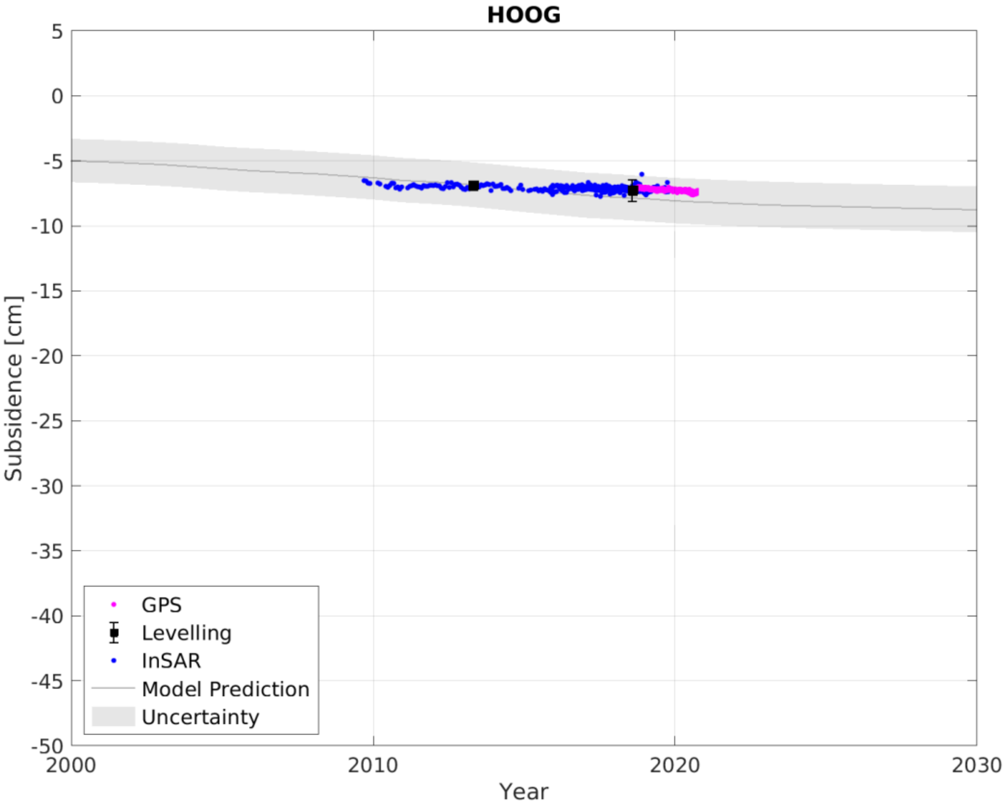

EEMS
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

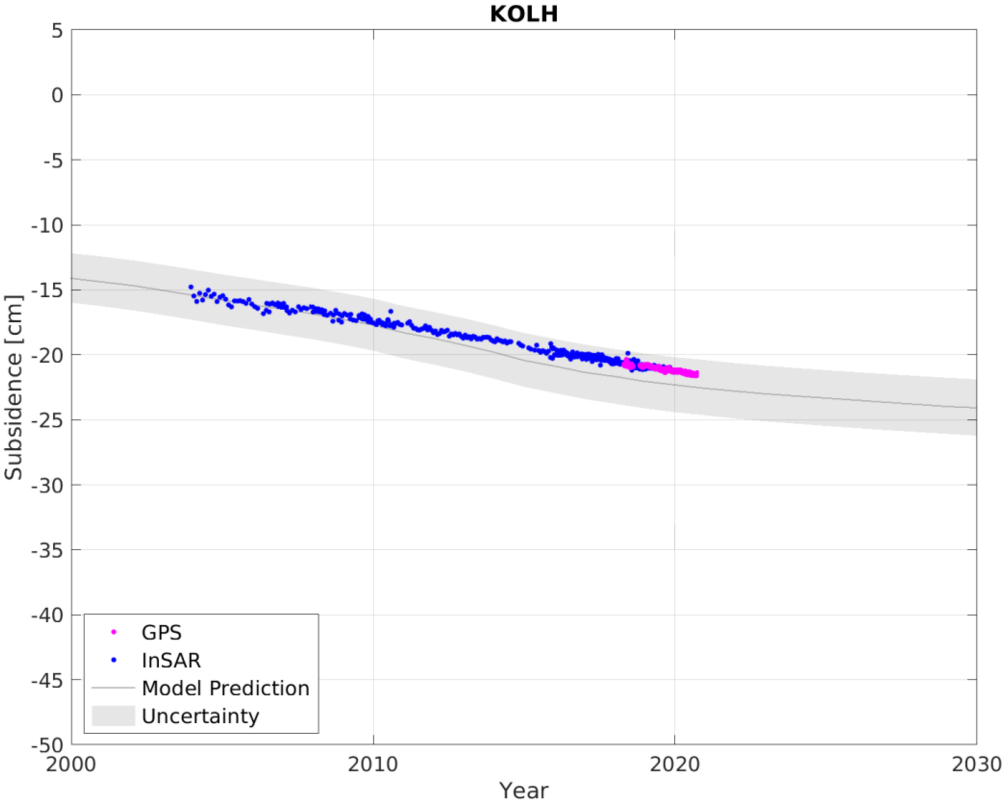

ENGE
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Year
2000
2010
2020
2030
GPS
Levelling
InSAR
Model Prediction
Uncertainty

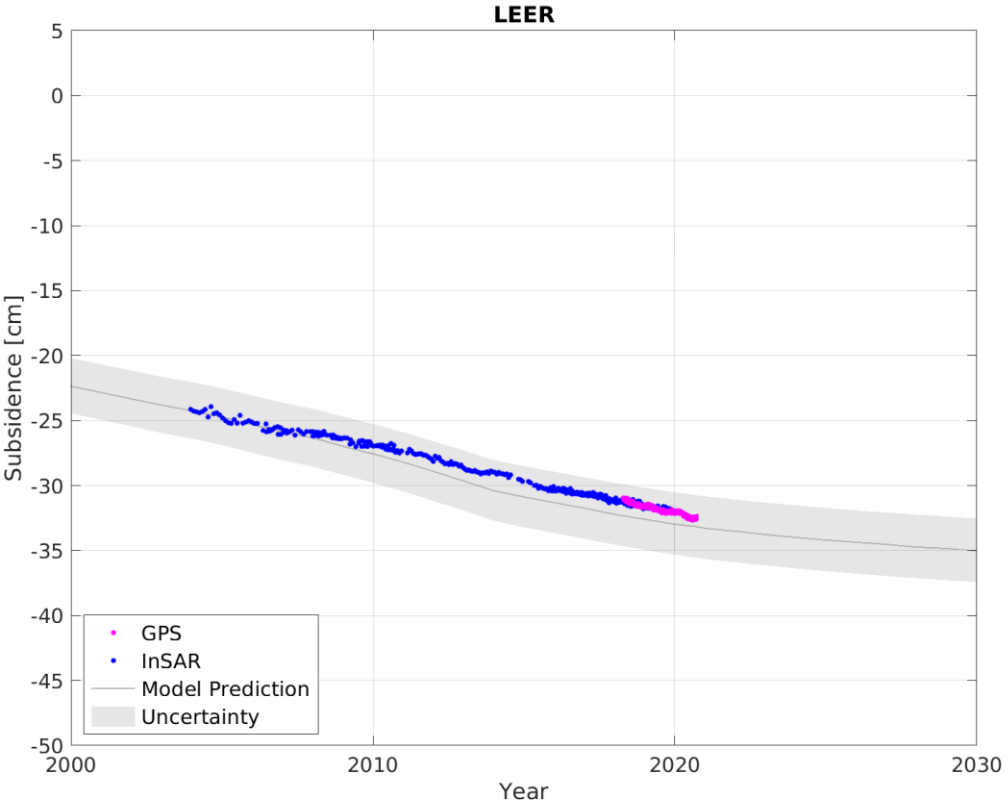

FROO
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Year
2000
2010
2020
2030
GPS
InSAR
Model Prediction
Uncertainty

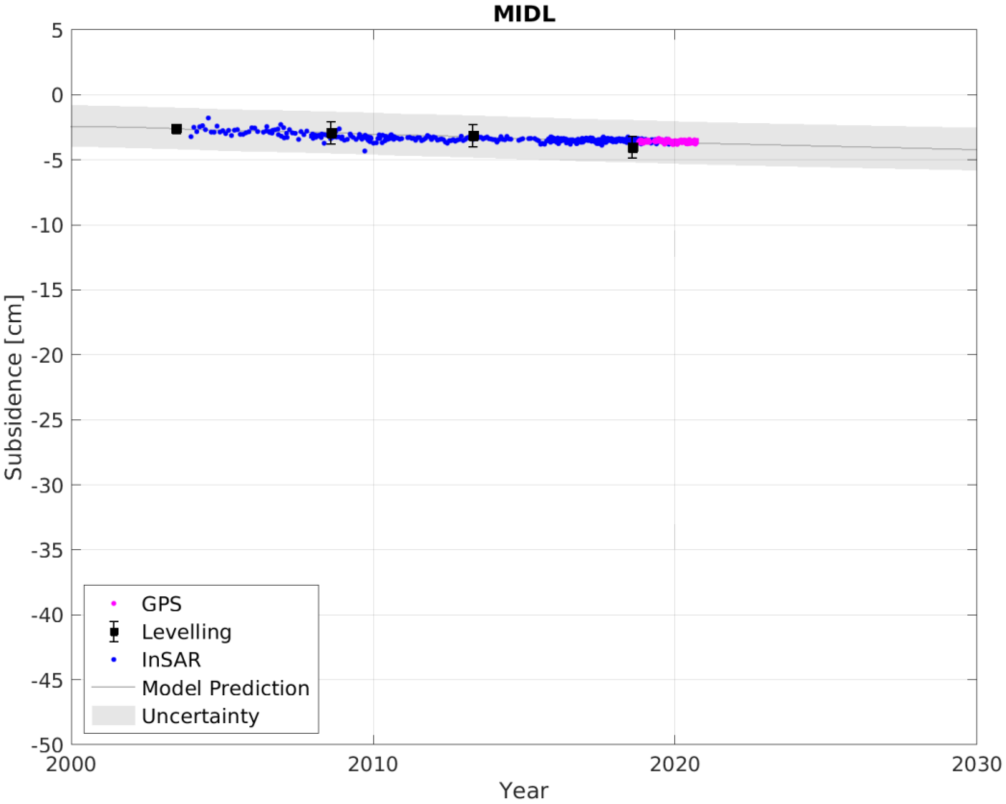

HARE
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

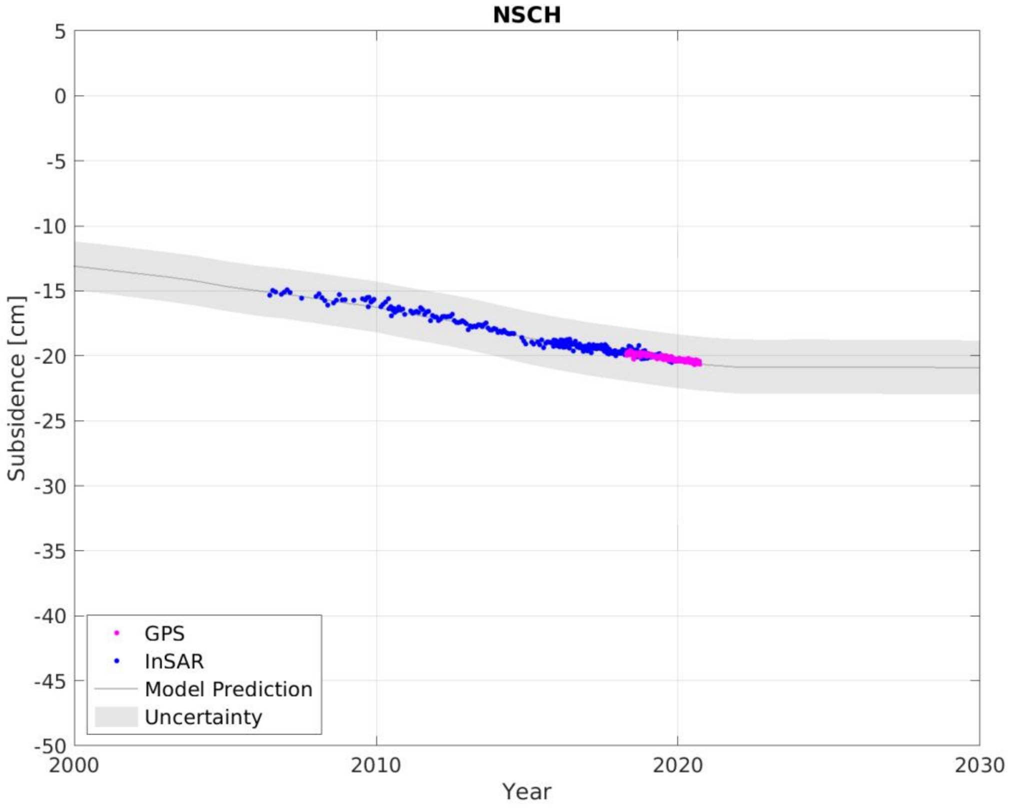

HEIL
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

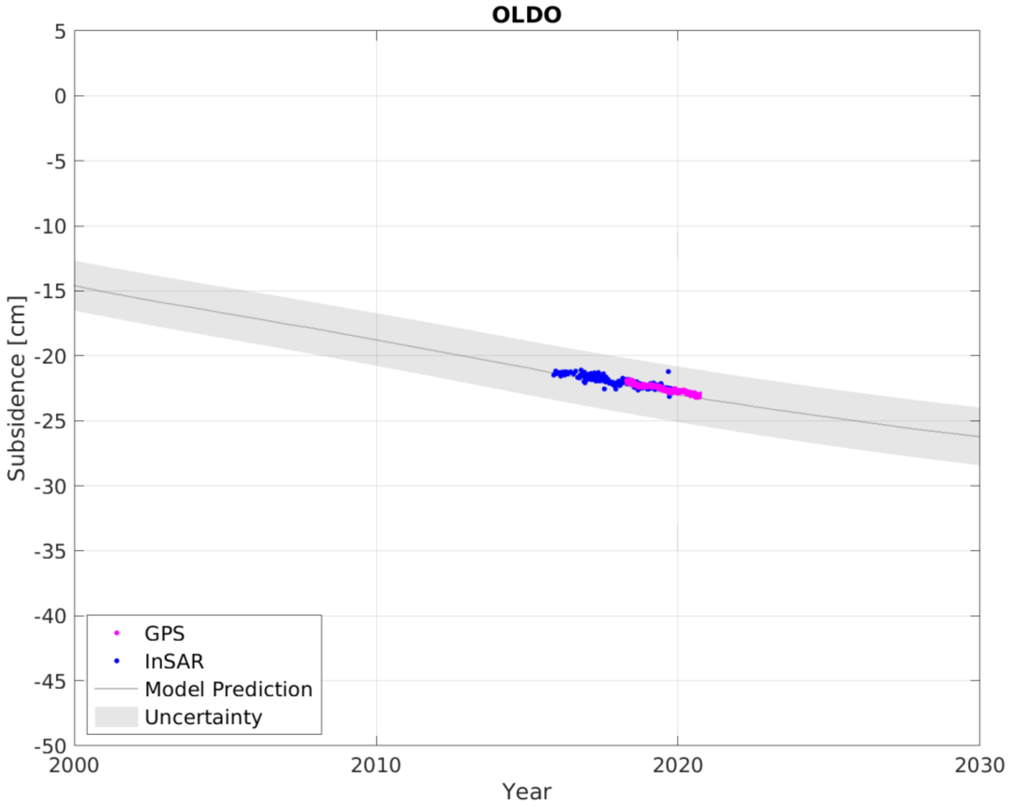

HOOG
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

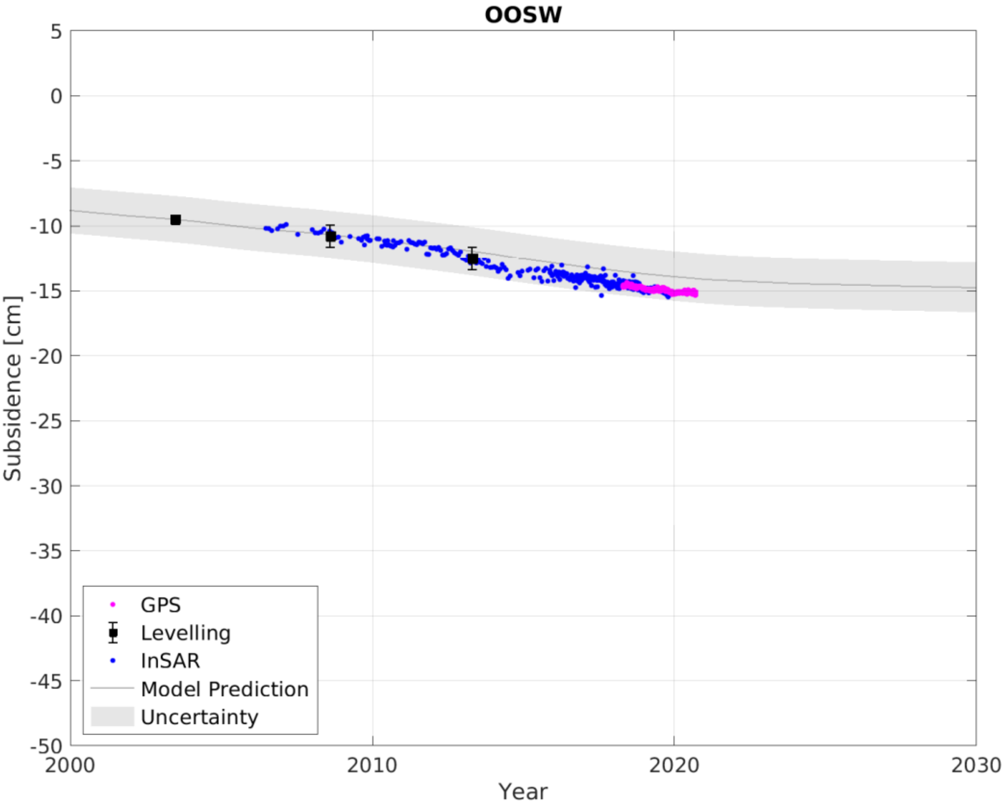

KOLH
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

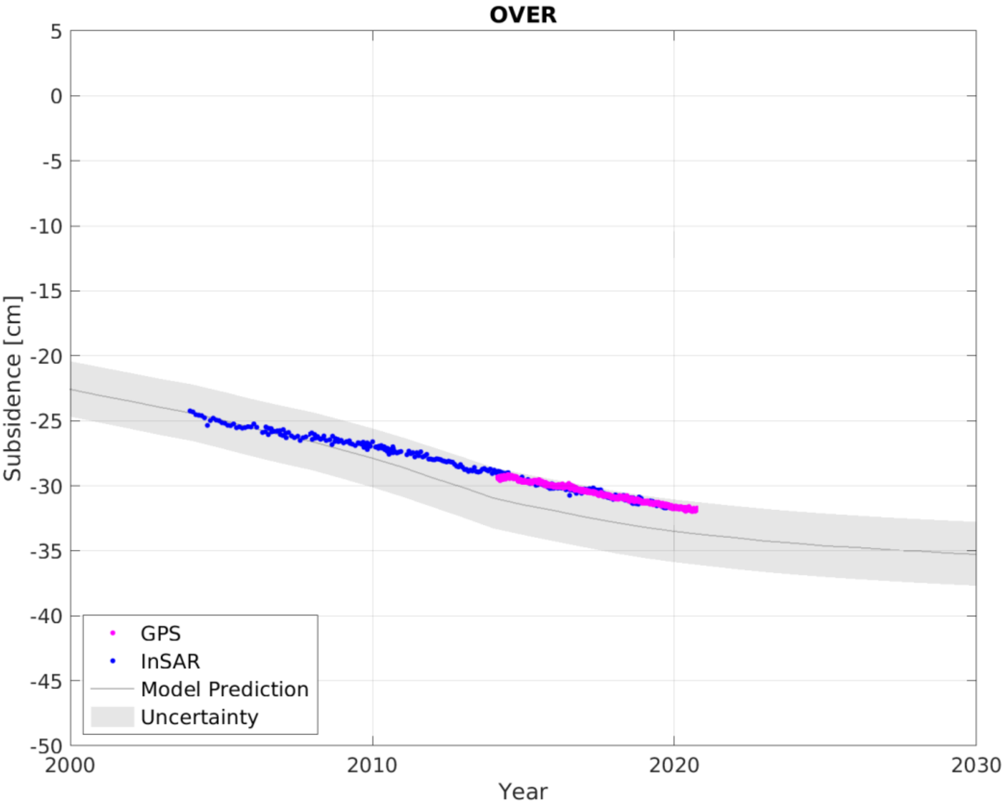

LEER
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

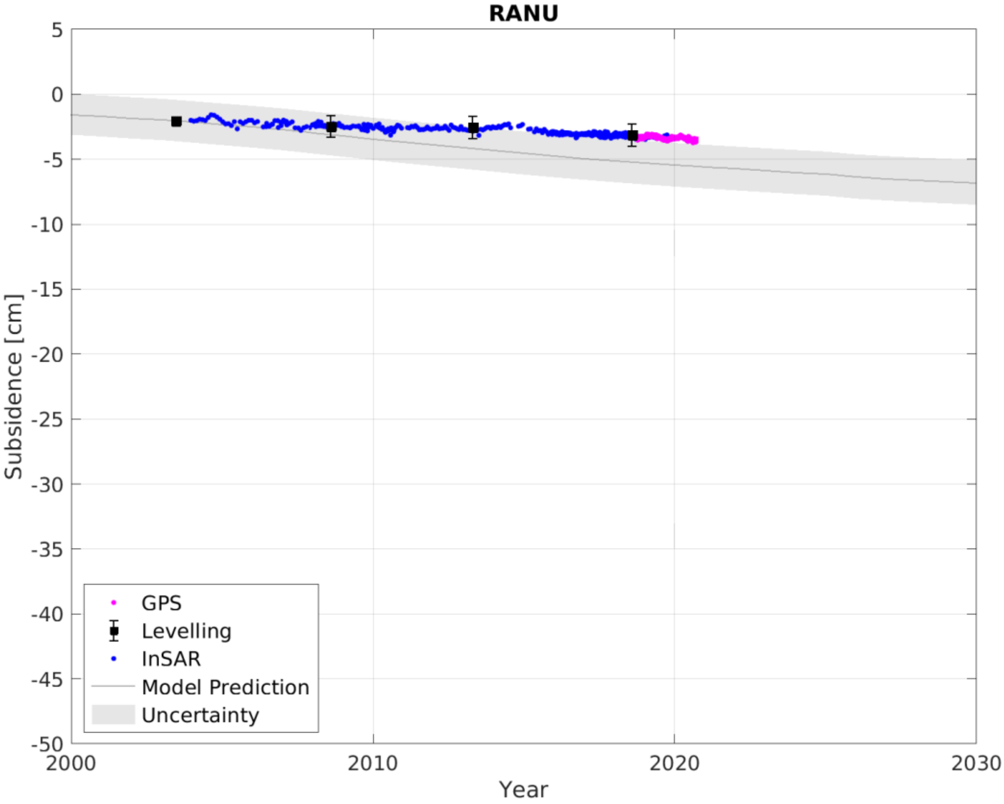

MIDL
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

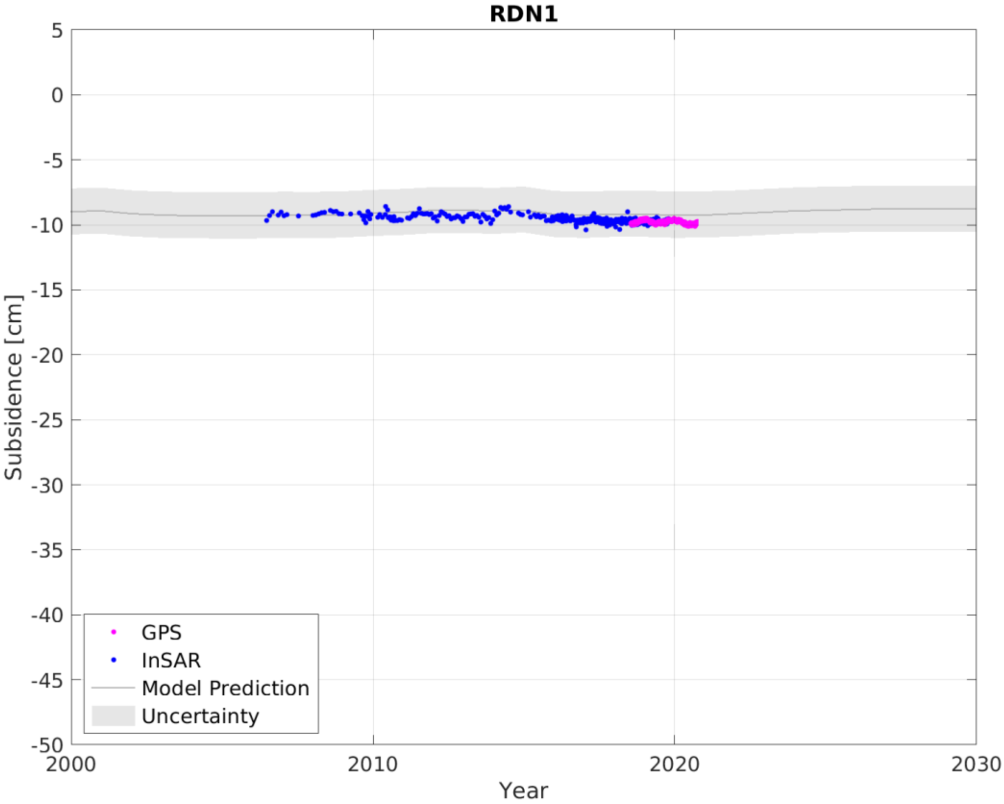

NSCH
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

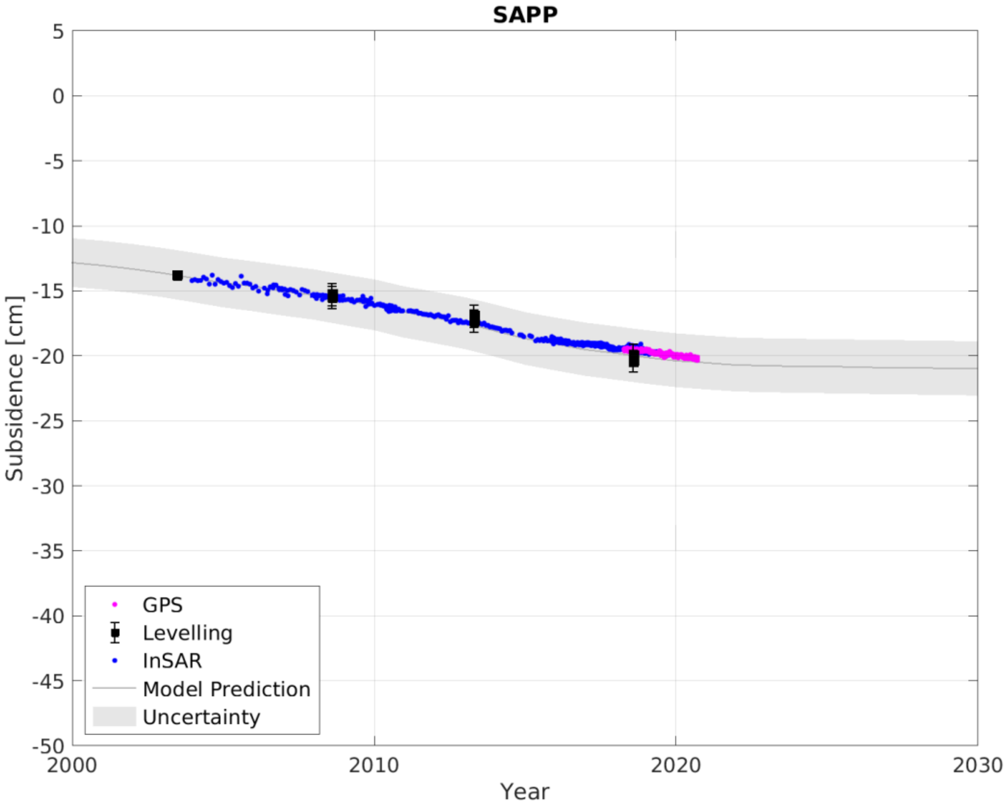

OLDO
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

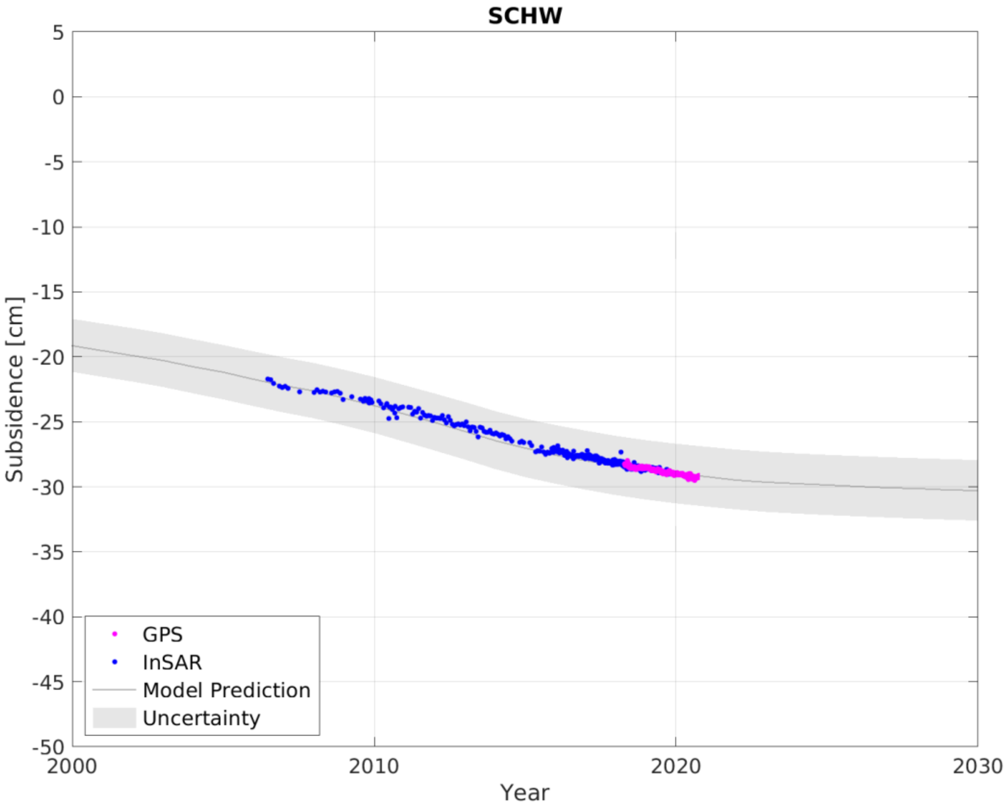

OOSW
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

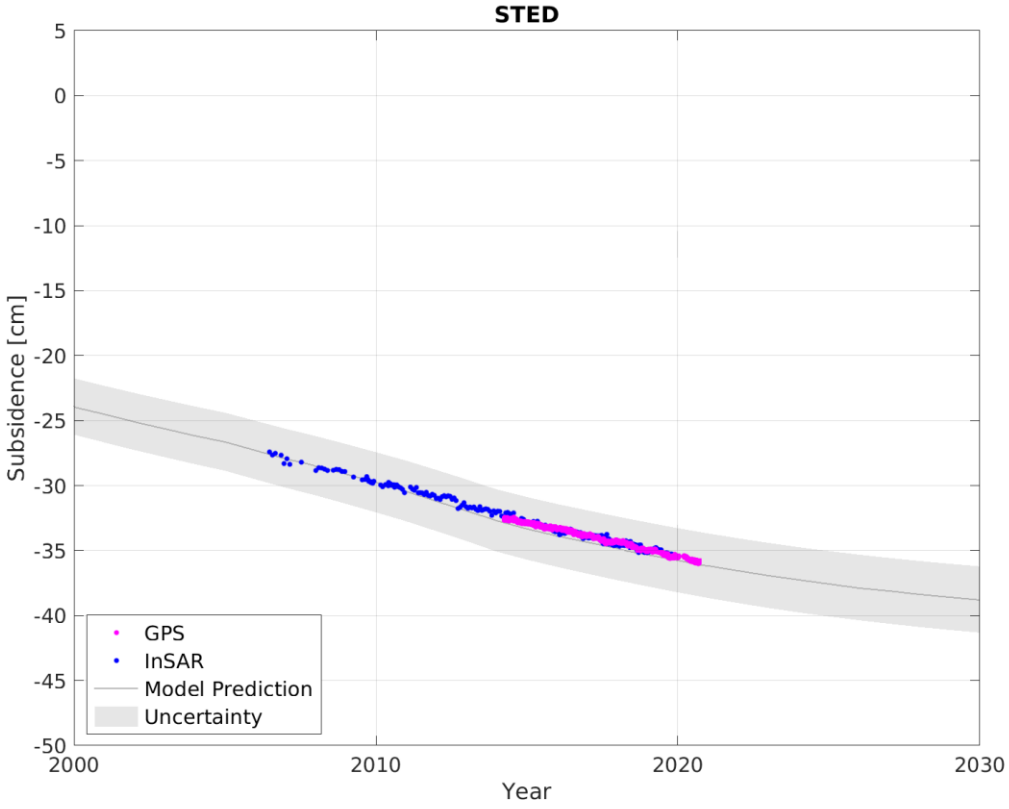

OVER
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

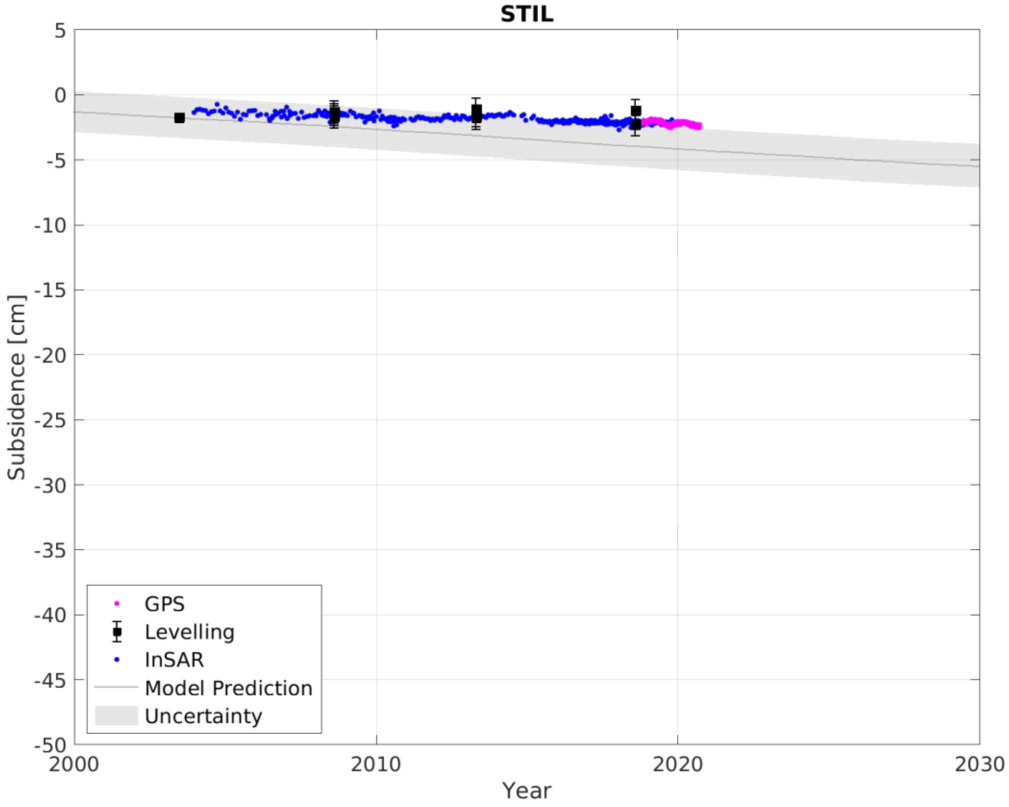

RANU
Subsidence [cm]
Year
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
GPS
Levelling
InSAR
Model Prediction
Uncertainty

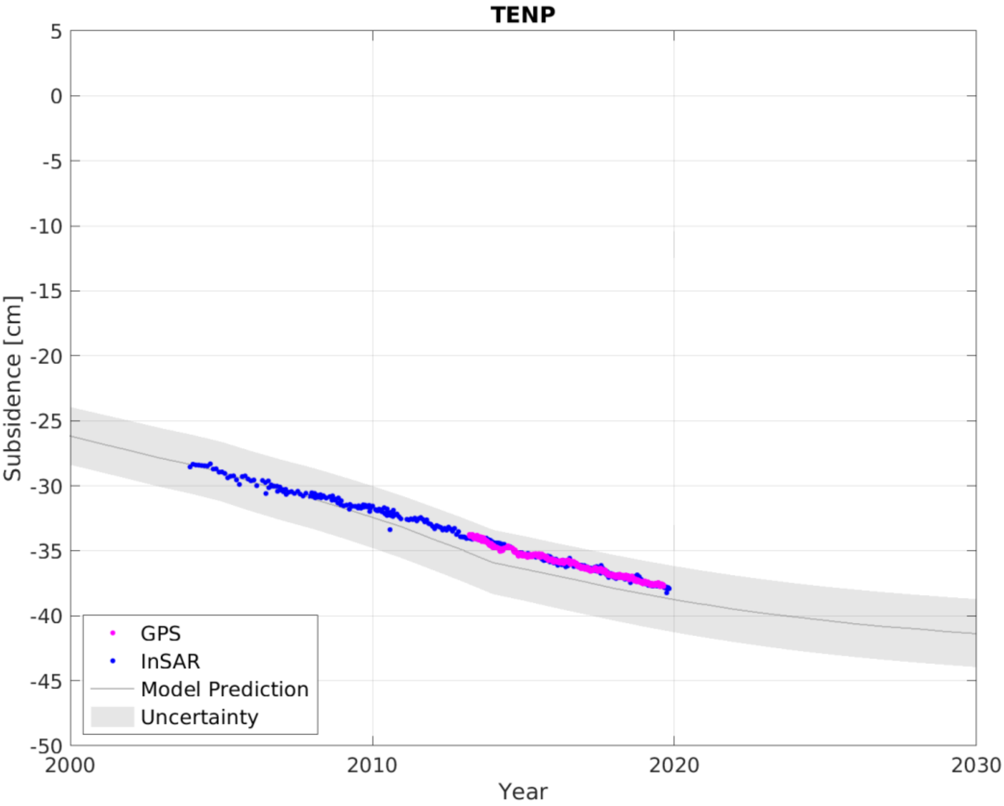

RDN1
Subsidence [cm]
Year
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
GPS
InSAR
Model Prediction
Uncertainty

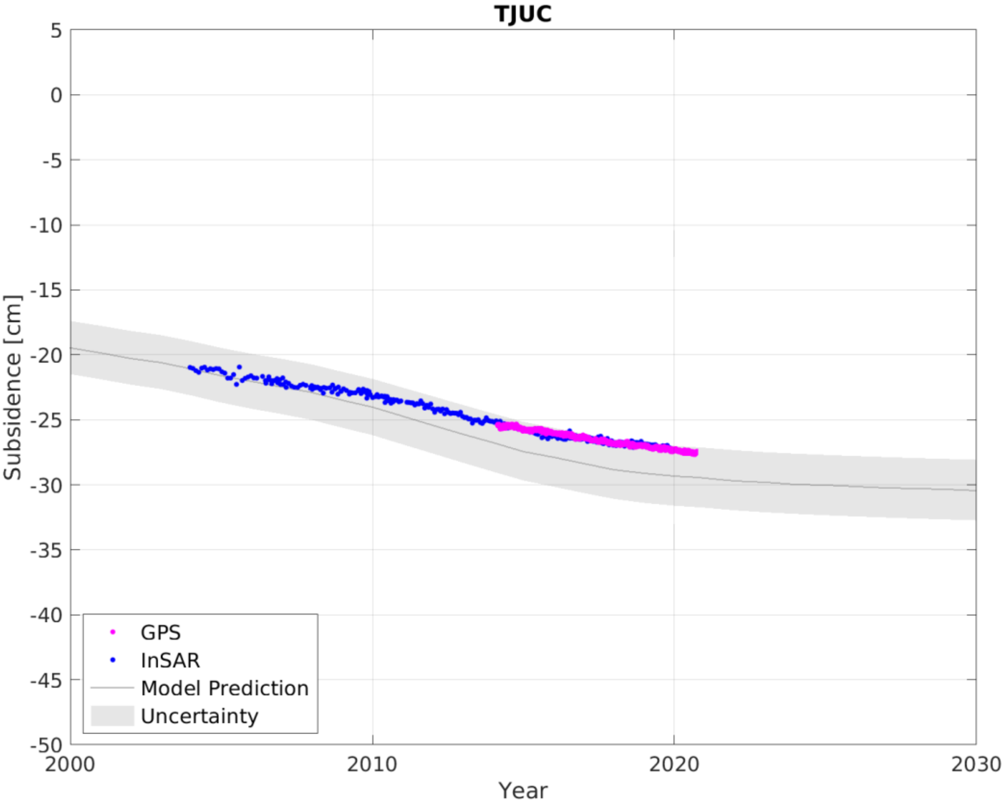

SAPP
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

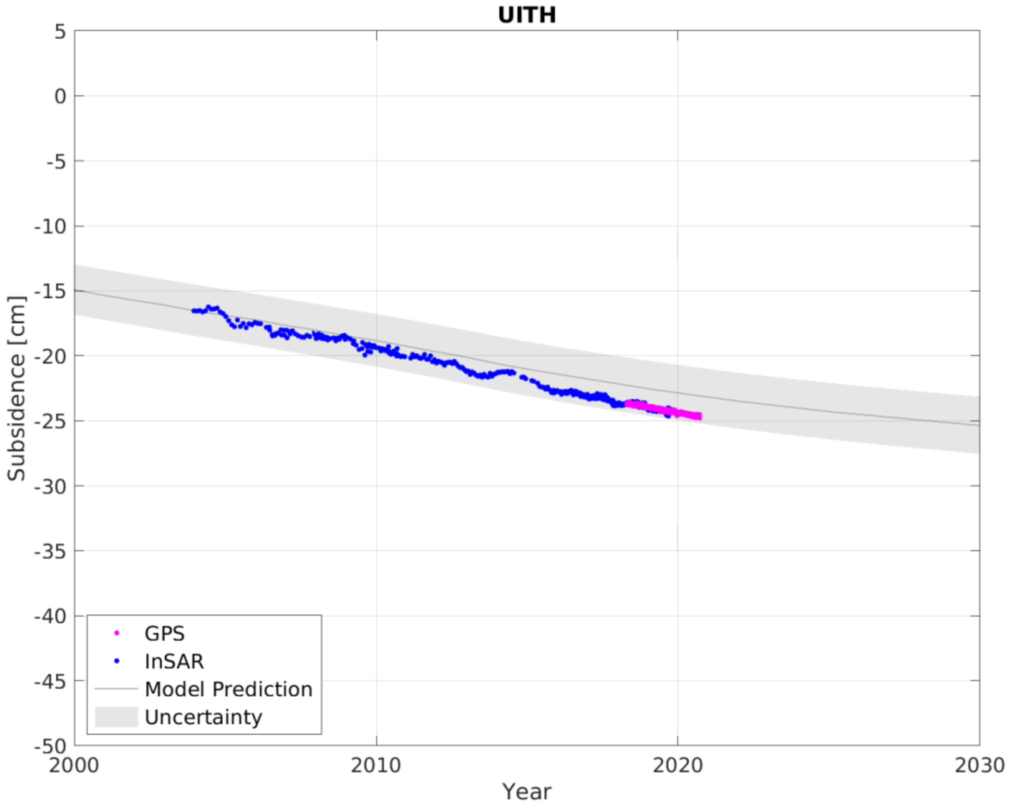

SCHW
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Subsidence [cm]
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

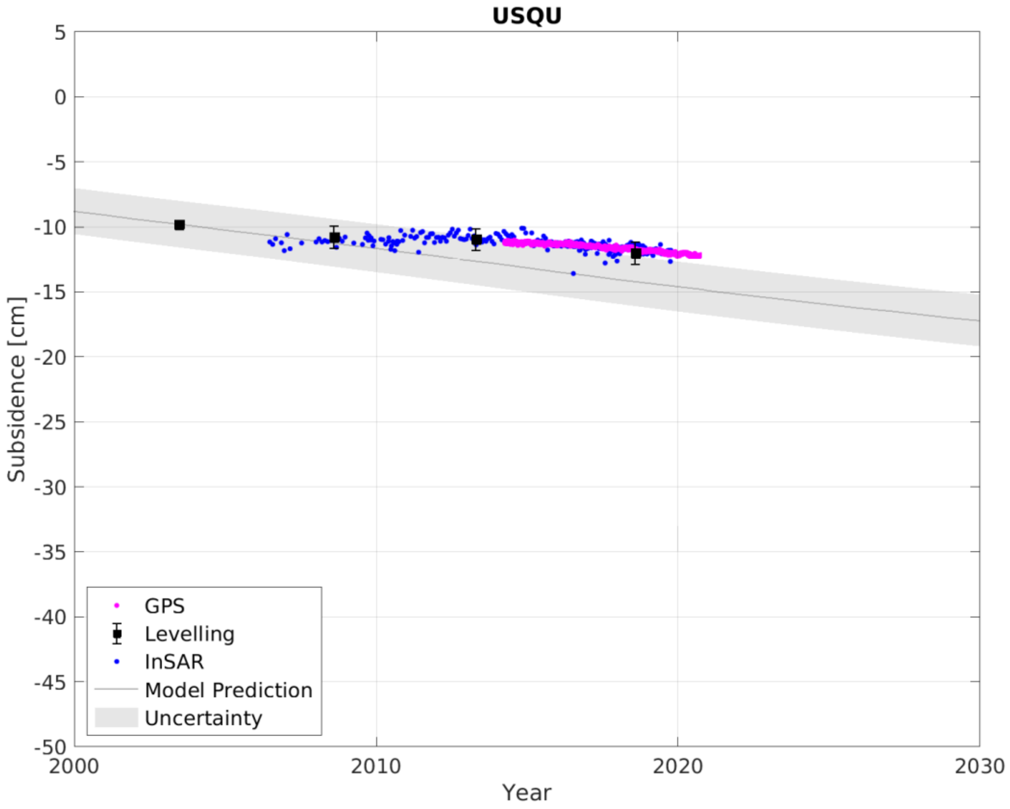

STED
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

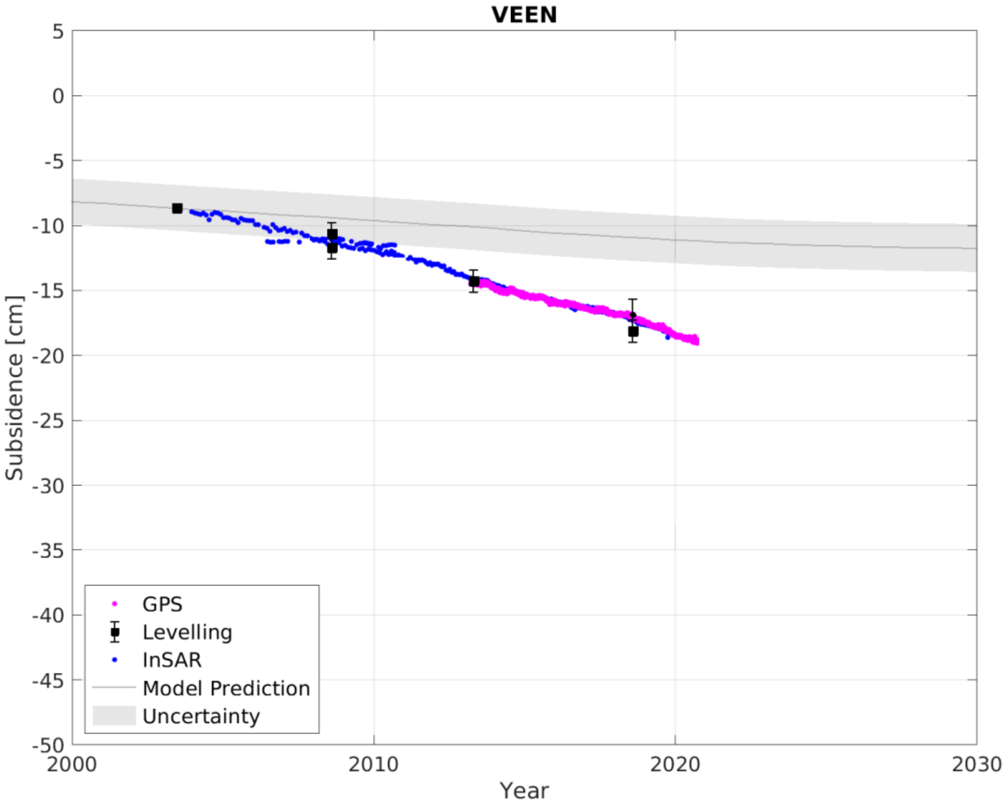

STIL
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

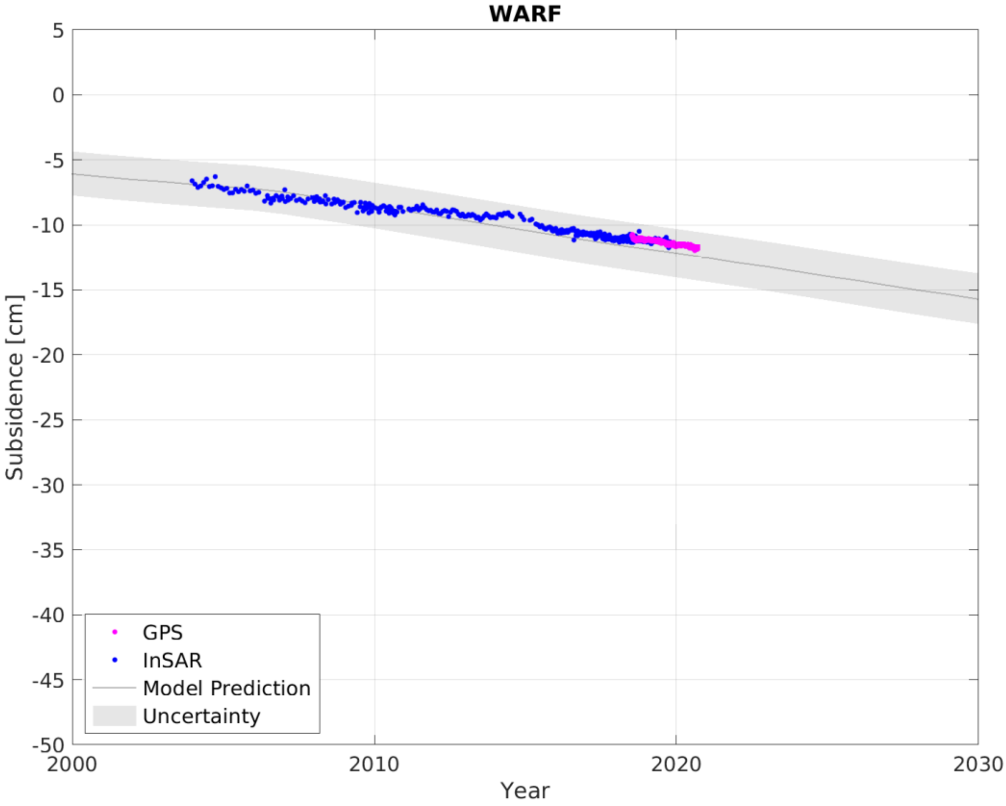

TENP
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

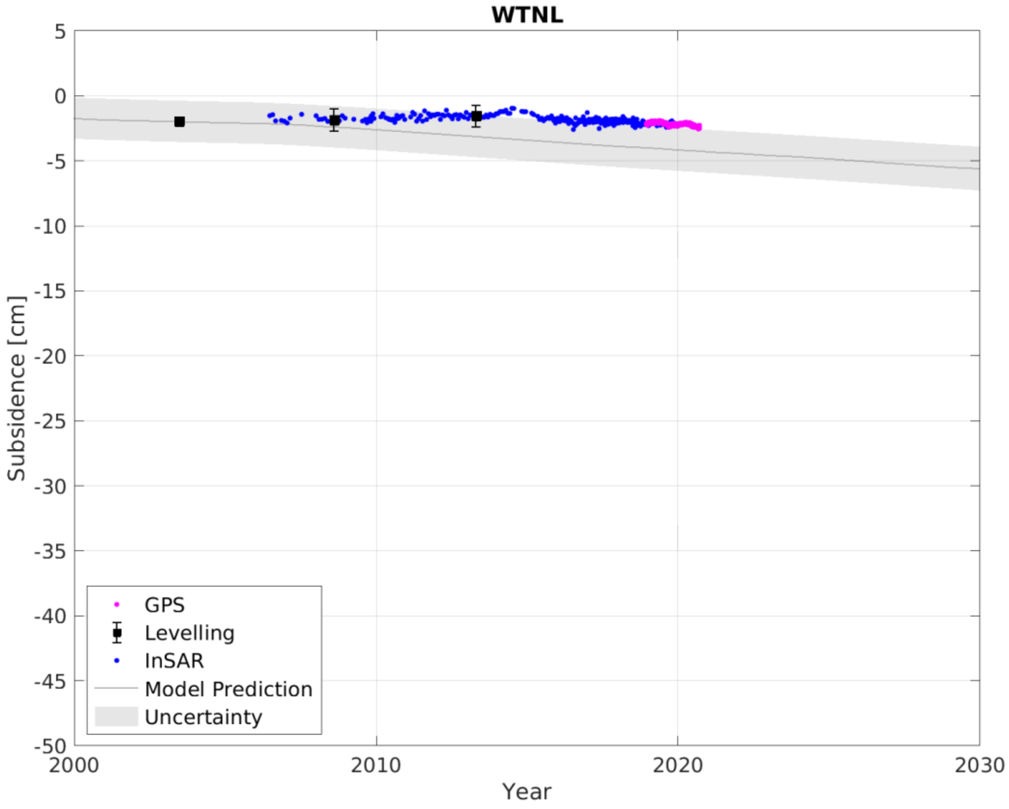

TJUC
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

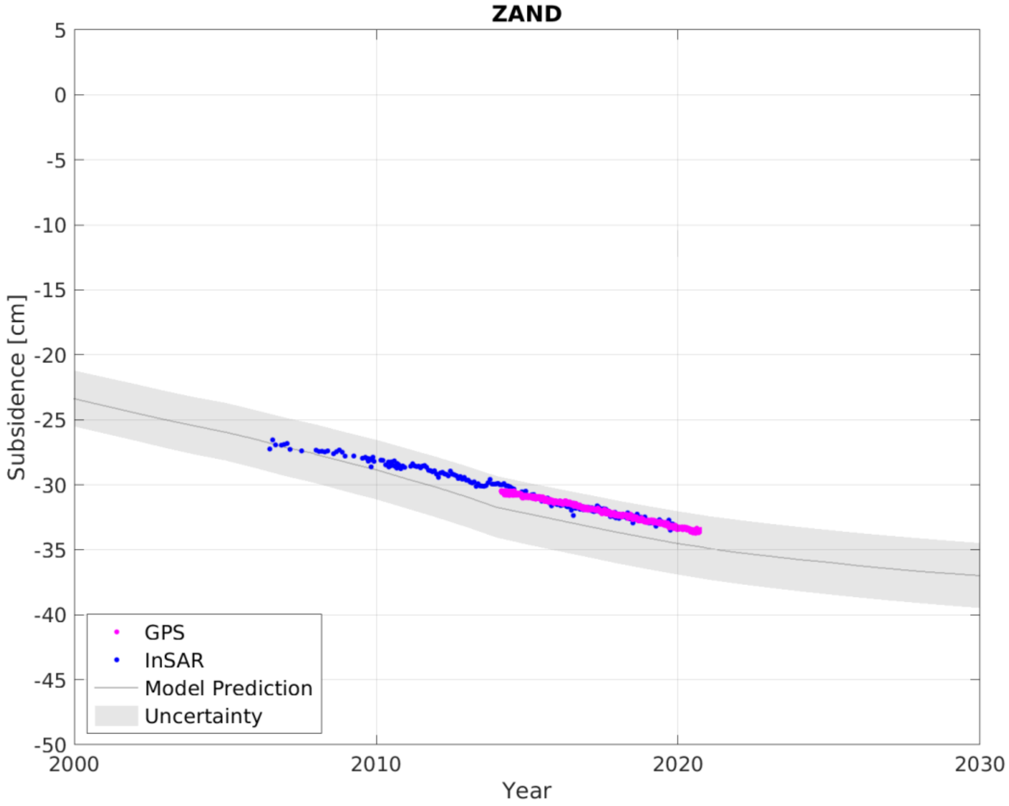

UITH
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

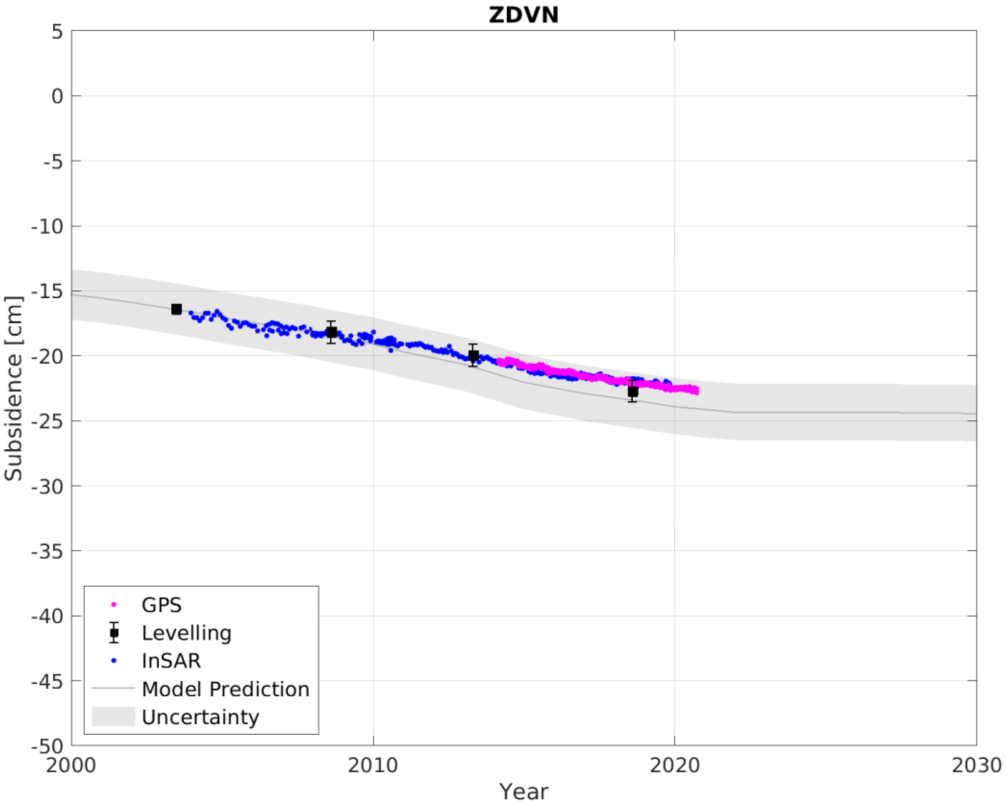

USQU
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

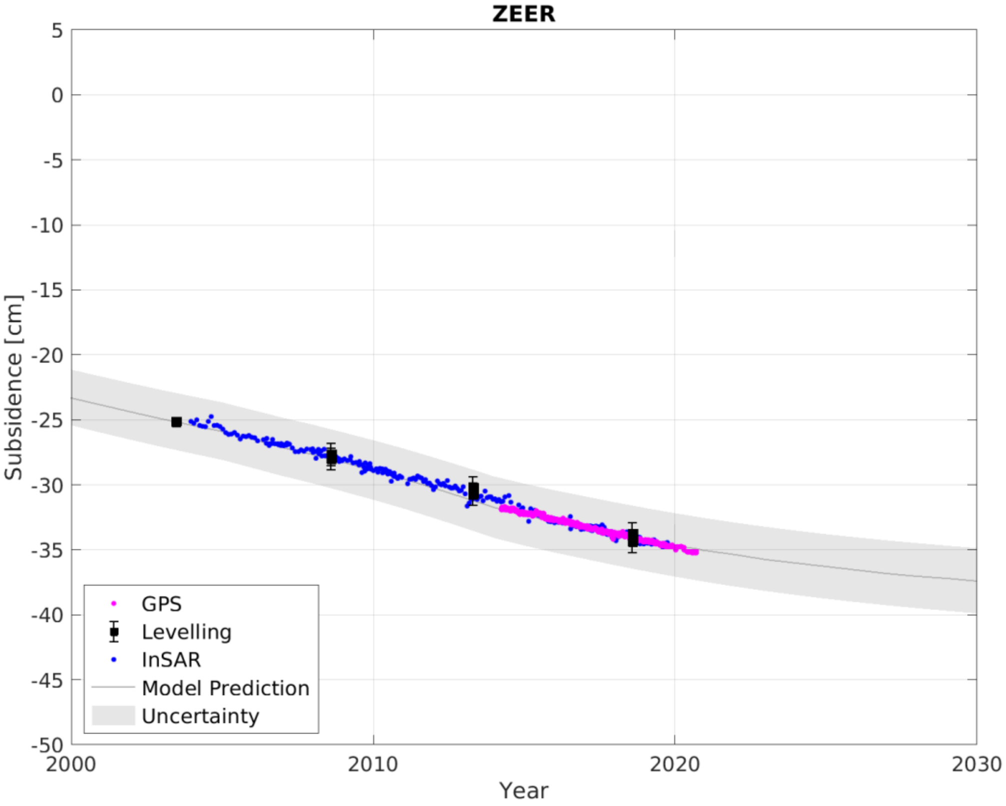

VEEN
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty

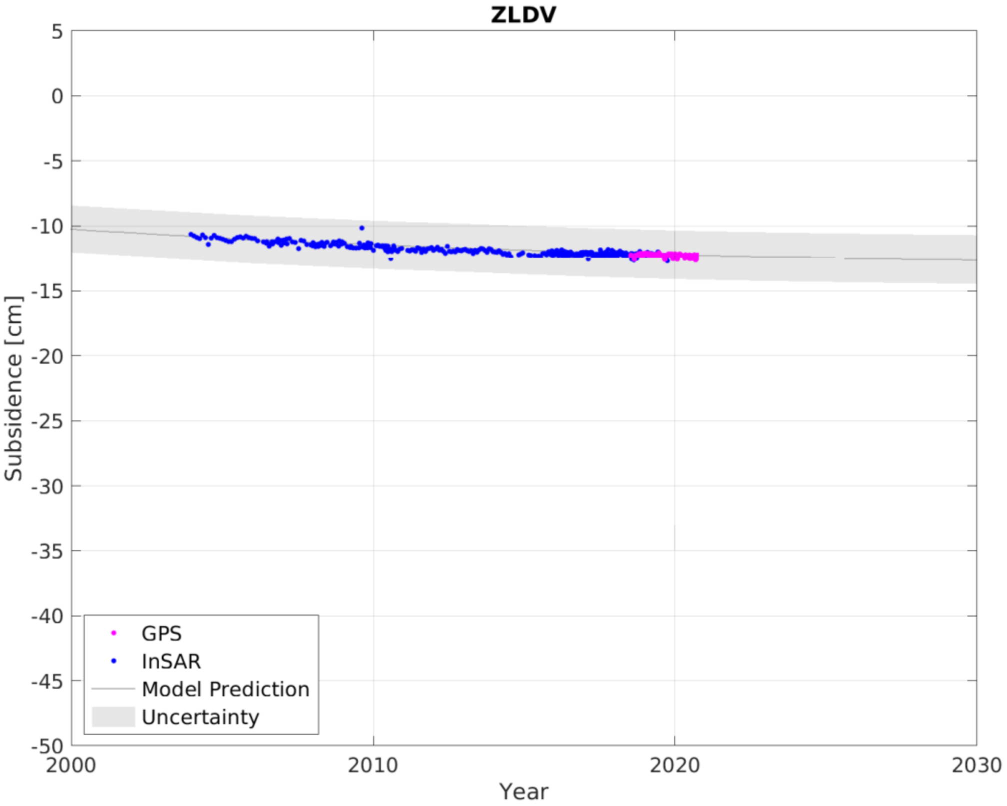

WARF
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty


WTNL
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty


ZAND
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty


ZDVN
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty


ZEER
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
Levelling
InSAR
Model Prediction
Uncertainty


ZLDV
Subsidence [cm]
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
2000
2010
2020
2030
Year
GPS
InSAR
Model Prediction
Uncertainty

Appendix B The effect of multiple reservoir realisations on the modelled subsidence uncertainty

Introduction

This appendix addresses the effect of aquifer uncertainty in addition to the sources of uncertainty that are described in the main report. The main report predicts subsidence for only the most probable aquifer realisation and subsidence scenario, selected by the geodetic data in these areas. The benefit of this choice is the selection of only one set of geomechanical parameter values combined with the quantification of the model uncertainty. In this appendix the subsidence and the model uncertainty are statistically calculated using multiple aquifer realisations to assess the impact on the model uncertainty.

This is done by slightly amending the calculation scheme used in the main report (Figure 61).

The starting point of the new calculations is step 3, where new reservoir realisations with a lower probability are combined with the reservoir realisation used in the main report. This choice leads as well to the creation of multiple subsidence scenarios.

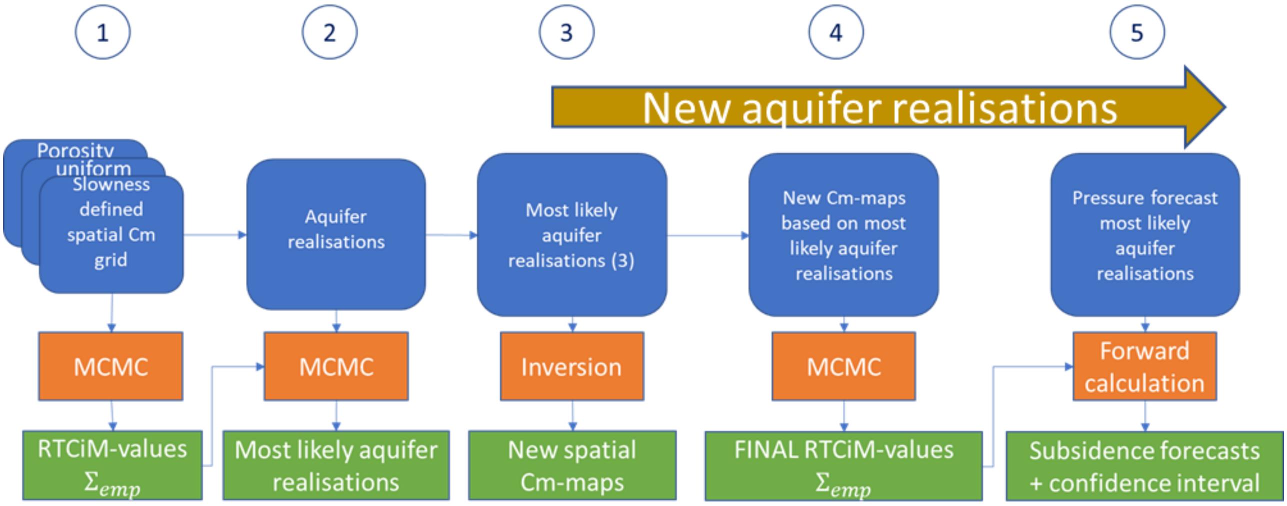


Figure 61; Calculation scheme for the multiple reservoir realisations.

The starting point of the modified workflow is the probability value of the aquifer realisations that were calculated in step 2. All aquifer realisations with a probability value larger than 2 % are selected. New spatial Cm grids are calculated for the selected realisations (step 3) in an inversion scheme using the Cm-porosity grid as prior information. Then a Markov Chain Monte Carlo was run for this realisation to determine the posterior bandwidth of RTCiM and Σ_{emp} parameter values.

In the main report the most likely parameter value set was used for the subsidence and model uncertainty calculation. In the modified workflow, all posterior parameter values from the MCMC calculations are used for forward subsidence and model uncertainty calculations. This is executed for all selected aquifer realisations. From these forward calculations the uncertainty bandwidth is defined by statistically drawing realisations from these aquifer realisations, where the number of drawings per realisation is determined by the probability value of the specific realisation.

Next sections discuss the calculation steps and results in more detail.

Selection of pressure realisation per aquifer area

Calculations have been performed for four lateral aquifer areas: South-West aquifer, Möwensteert in the North, Rysum to the East and the South aquifer (Figure 40). Multiple aquifer realisations were selected for each area, except for Möwensteert where one aquifer realisation has been selected (probability of 100%).

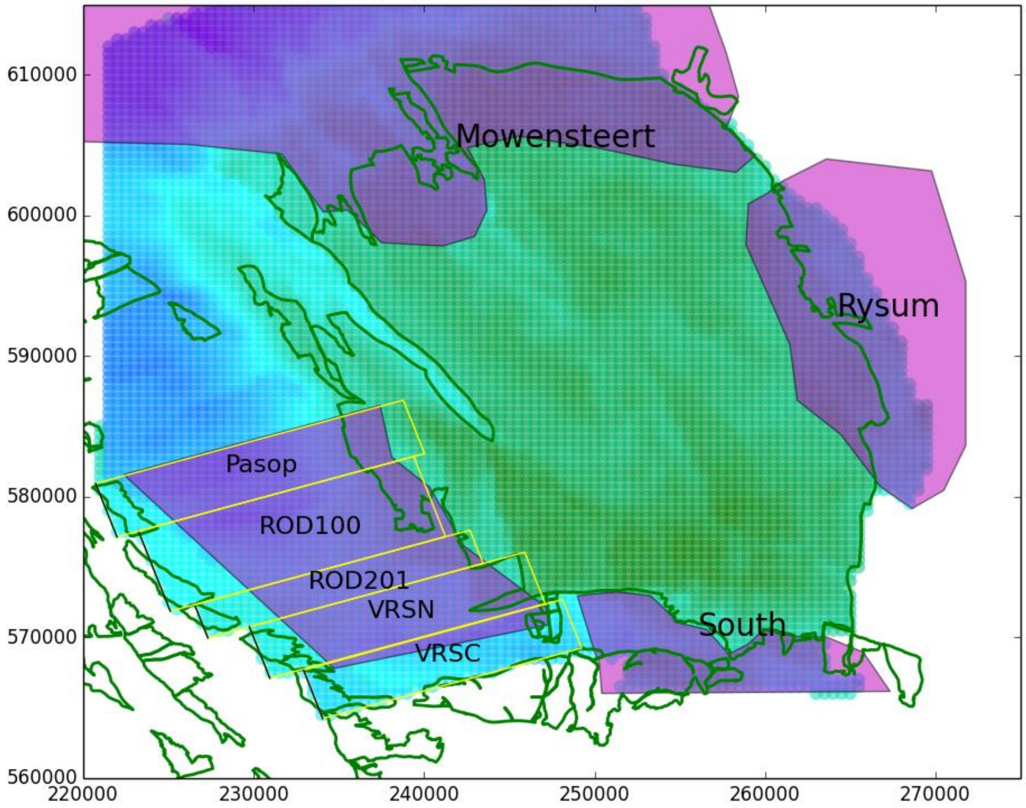


Figure 62; map of the Groningen pressure area. Gas fields are bounded by a green outline and aquifer areas are shown by the purple areas. The yellow boxes in the Southwestern aquifer are the box-models. The dots in the background indicate the total V6 reservoir model area.

The South-West aquifer, between Groningen and Roden/Vries fields, consists of 5 so called box-models: VRSC, VRSN, ROD201, ROD200 and Pasop (Figure 40). Each box-model has 5 pressure realisations ranging from a minimum (min) to a maximum (max) depletion case (see also Figure 64 for the depletion profiles).

Two additional box-model realisations were selected for the statistical analyses. These realisations have a probability of 5 and 3% (Figure 37). The pressures in time of the selected realisations are shown in Figure 64, which is similar to Figure 39 in the main report.

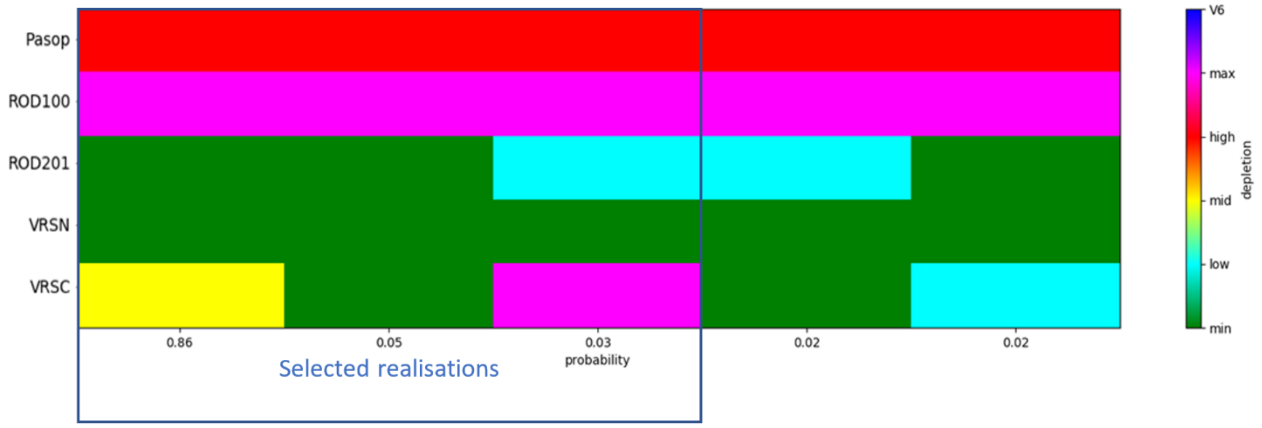


Figure 63; Probability for the SW aquifers. The x-axis shows the probability of the realisations sorted from a high to low value. The box with "selected realisations" contains the realisations that are selected for the uncertainty calculation.

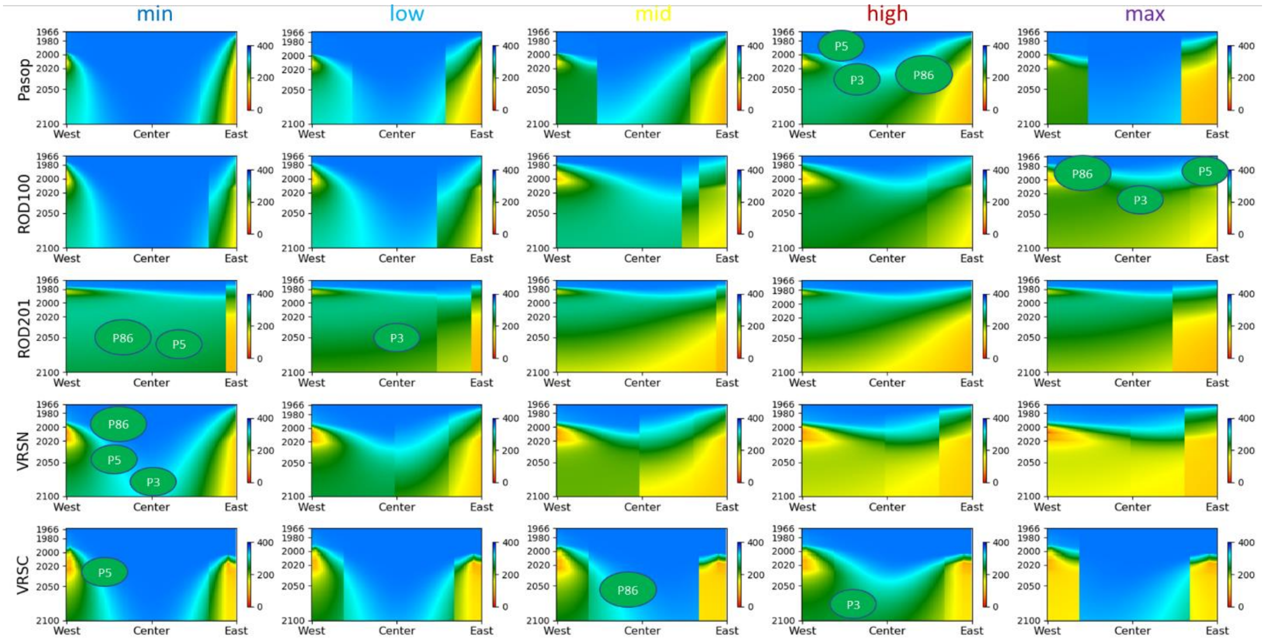


Figure 64; Additional aquifer realisations selected for the South-West aquifer. Each realisation is a combination of depletion (min to max) cases selected from each box-model. The realisations or combinations are indicated by green ovals with the text P86, P5 and P3.

The probability is also calculated for alternative realisations in the other aquifer areas Möwensteert, Rysum and the South.
The recent V6 reservoir model is selected for these areas in the main report. This is based on the low value of the negative log likelihood and the incorporation of the latest geological and reservoir engineering insights.

Table 6 shows the probabilities for the possible realisations in 'Rysum' and 'South'. For Möwensteert, a 100% probability is assigned to the V6 model. The pressures of the V4 model in the South aquifer are almost similar to the pressures that were modelled by the V5 model. Therefore, only the V5 model is taken into account in the subsequent calculations.

Table 6; Probabilities (%) for the lateral aquifer areas 'Rysum' and 'South'

	'Rysum'	'South'
2013_1P	0	0
2013_2P	11	19
2013_3P	74	19
V4	1	20
V5	5	19
V6	9	23

Three runs, in addition to the run reported in the main report (Model 0), were executed to provide the full span of thousands of new subsidence scenarios.

Table 7; Selected model runs, required to compute the results for the aquifer realisations. Note that Model 0 is the model used in the main report.

	Model 0	Model 4	Model P5	Model P3
South West	P86	P86	P5	P3
Möwensteert	V6	V6	V6	V6
Rysum	V6	V5	2013_3P	2013_2P
South	V6	V5	2013_3P	2013_2P

The pressure grids for the aquifer realisations in these runs are merged with V6 model that covers in all runs the gas bearing part of the Groningen field. This process is explained in more detail in the main report. An overview of the depletion in 2020 and 2080 for the different runs is shown in Figure 65. The purple polygons show the areas where the pressure values were taken from the older (than V6) reservoir models.

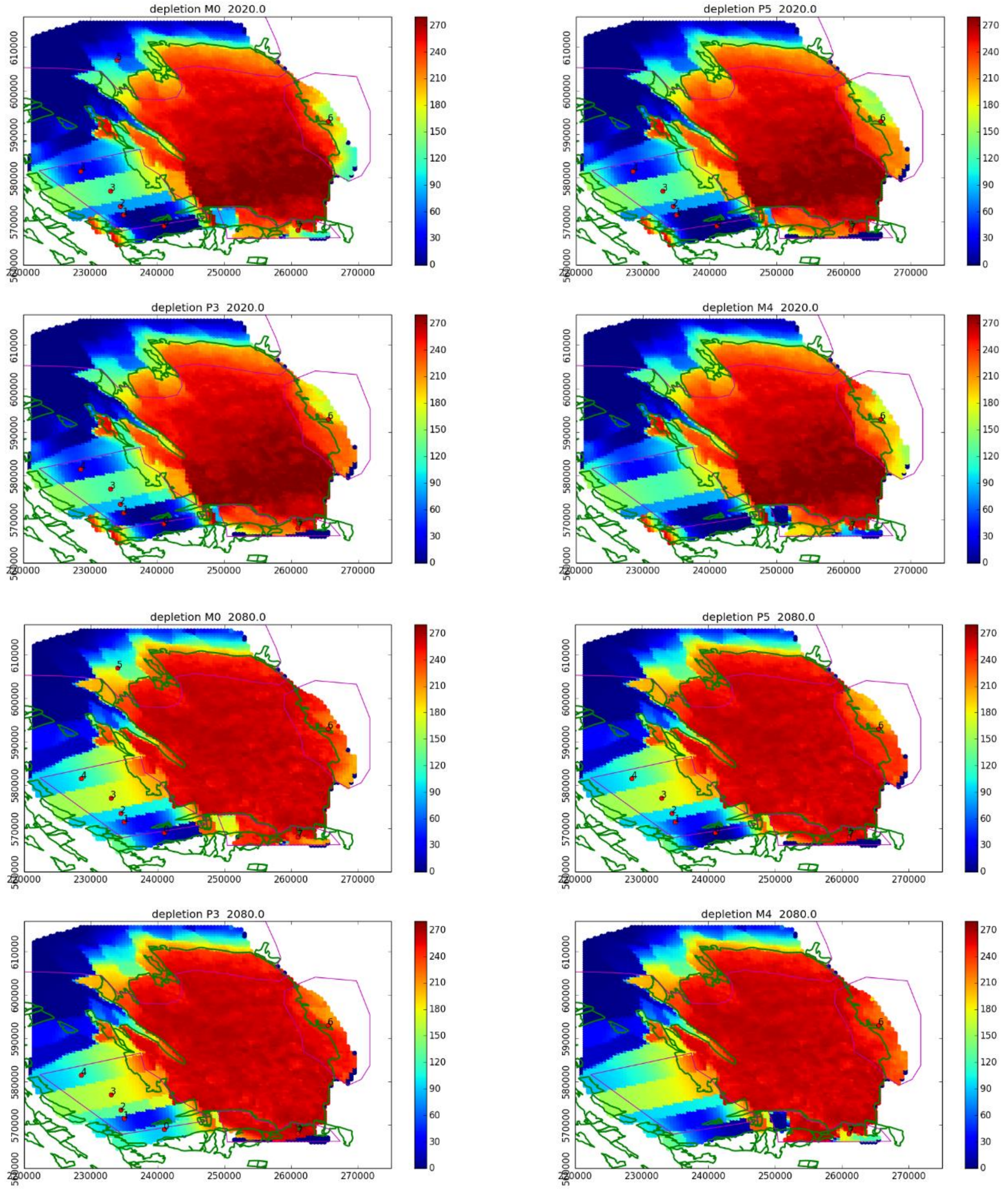


Figure 65, Depletion maps for the selected aquifers in 2020 and 2080. The magenta polygons indicate the extend of the aquifer areas.

Figure 66 shows the pressure versus time for the various reservoir realisations at the locations indicated in Figure 65. These realisations are all used in the statistical analysis.

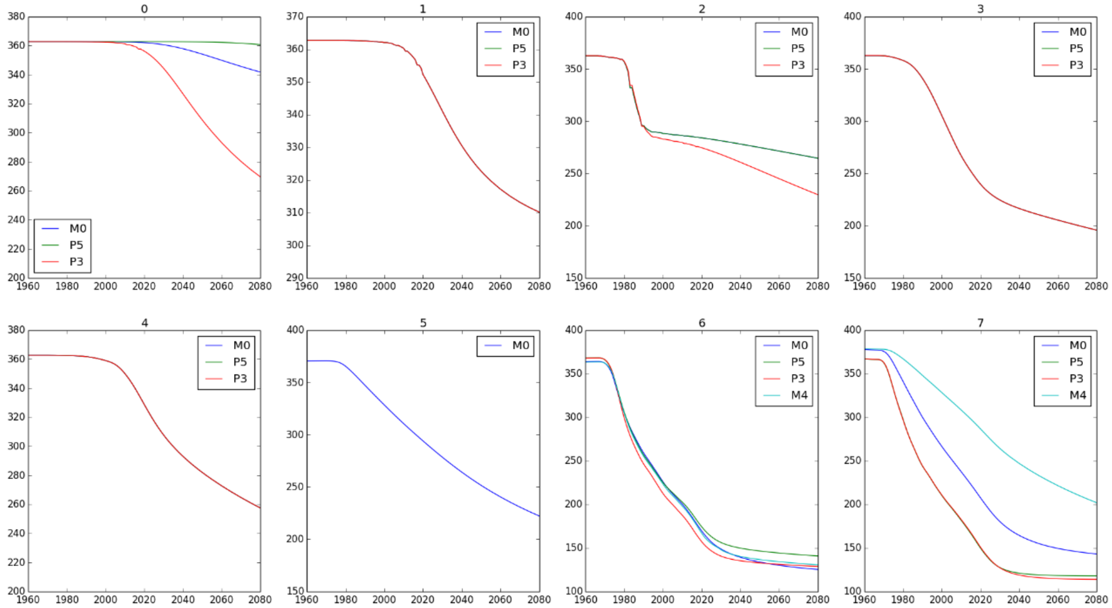


Figure 66; Pressures in time at different aquifer locations (see maps in Figure 65) for the selected realisations .

Calibration to create a new spatial Cm grid

Next step (step 3 in Figure 61) is a spatial Cm calibration for all pressure realisations to the measured subsidence. The posterior spatial Cm grids per realisation show small differences between the realisations and are comparable with the maps that are shown in the main report.

Monte Carlo simulation to define RTCiM parameter sets

Step 4 is similar as in the main report but now being executed for all reservoir realisations. Monte Carlo simulations are performed to define the RTCiM and uncertainty parameter values for each model run. This step results in 8000 subsidence scenarios. The results for the scenario with the lowest Negative Log Likelihood (NLL) value are tabulated in Table 8. The scenario with the lowest NLL values has the highest probability. The posterior distributions of parameter values are visualised in Figure 67. These distributions are used for the draws in the statistical analysis, which is described in the next section.

Table 8; Overview of the most likely posterior set of parameter values.

	Model 0	Model 4	Model P5	Model P3
A [-]	0.85	0.92	0.88	0.93
d [-]	0.4	0.4	0.41	0.41
b [-]	0.021	0.023	0.022	0.23
σ_0 [cm^2]	0.53	0.56	0.54	0.57
α [cm]	0.029	0.033	0.032	0.034

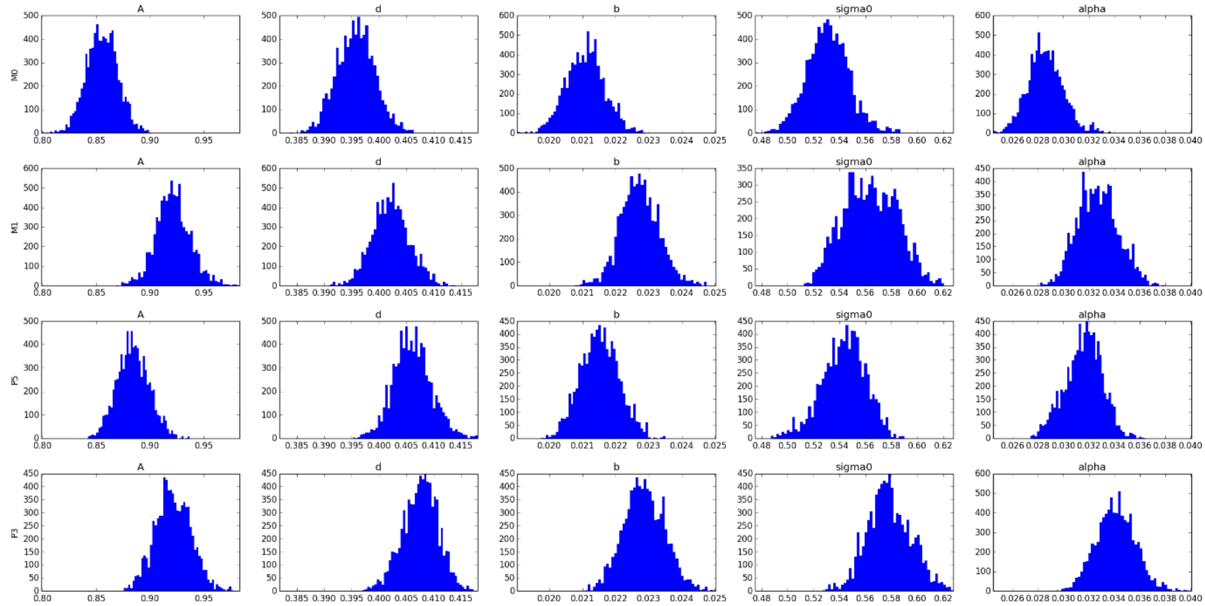


Figure 67; Distributions of RTCiM and model uncertainty parameter values. The rows show the parameter distribution per model run.

Uncertainty calculation for multiple realisations

This section describes the steps that are required to calculate the model uncertainty for the total set of pressure realisations. Firstly, the RTCiM and Σ_{emp} parameter values are drawn from the Monte Carlo run done in the previous step. Note that the burn-in periods of the MCMC runs (1000 iterations) are not included in this process. The forward subsidence is calculated with the RTCiM parameter values that are retrieved for each draw. Subsequently, the model uncertainty Σ_{emp} is calculated using the Σ_o, α and $Subsidence$ values from that draw. This σ_{model_unc} is 1 standard deviation of the model uncertainty distribution.

$$\sigma_{model\_unc} = \sqrt{\Sigma_{emp}} = \sqrt{\Sigma_o + \alpha \cdot Subsidence}$$

The simple example in Figure 68 shows the modelled subsidence for 200 hypothetical benchmarks (with index N). The green dot per benchmark is being calculated as the sum of the value for the subsidence (blue line) and the value for the uncertainty that is drawn from the uncertainty distribution.

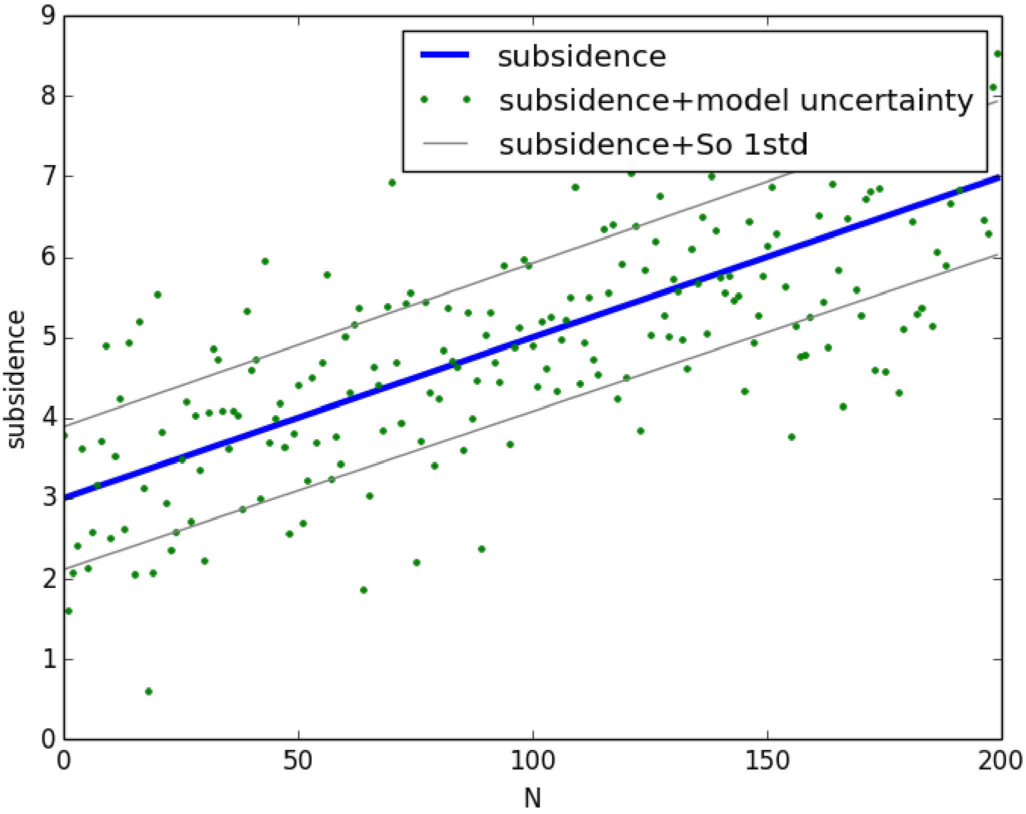


Figure 68; Example of the subsidence and model uncertainty calculation for 200 hypothetical benchmarks.

This method makes it possible to combine the scenarios for multiple reservoir realisations and to create the joint distribution describing total uncertainty above the aquifer areas. To enable this, 3000 scenarios per model run of the subsidence plus its uncertainty were stored for all model runs.

Subsequently from the total of 12000 scenarios, 3000 scenarios were drawn where the selection per reservoir realisation is scaled by the probability value of the specific aquifer realisation. These values are tabulated in Table 9. For example, 25% of the scenarios (750 scenarios) were drawn from the V6 realisation in the South aquifer and combined with 25% draws from the other realisations (V5, 2013_3P, 2013_2P, see Table 9) to build the full uncertainty distribution for the South aquifer. Finally, from the total distribution, the P2.5, P50 and P97.5 is calculated.

Table 9; Probability of each aquifer realisation per model run. The names of the aquifers realisations are listed in the grey cells.

	Model 0	**Model 4**	**Model P5**	**Model P3**	**Model 0**	**Model 4**	**Model P5**	**Model P3**
SW	0.91	0	0.056	0.034	P86	P86	P5	P3
Mowensteert	1	0	0	0	V6	V6	V6	V6
Rysum	0.10	0.05	0.74	0.11	V6	V5	2013_3P	2013_2P
South	0.25	0.25	0.25	0.25	V6	V5	2013_3P	2013_2P

Figure 69 shows the impact of the draws on the joint distribution and joint uncertainty calculation. In the top left figure, the individual model run results are shown for benchmark 012B0060, located above the southwestern aquifer. The middle row in Figure 69 shows the histograms of the modelled subsidence for the individual realisation while the bottom histogram shows the combined distribution. The top right picture shows in pink the solution space between P2.5 and P97.5 with the black line in the middle representing the P50 case. For comparison also the solution space from the single realisation, as documented by the main report, is shown by the top and bottom black lines. The difference in uncertainty by the two approaches is in this case negligible. Figure 69 also shows the small parameter uncertainty in the histograms, indicated by the blue peaks.

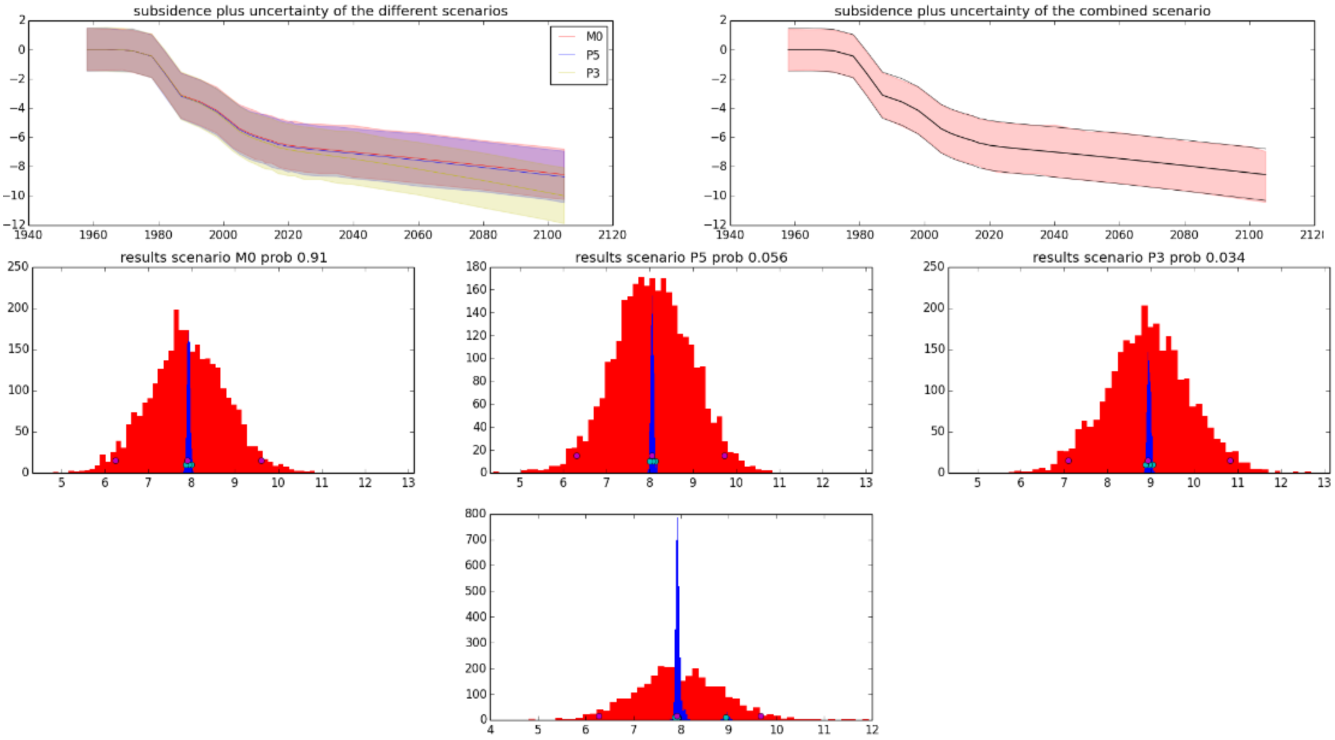


Figure 69; Top row shows the uncertainty bands for the 3 individual realisations (left) and combined analysis (right). The middle row shows the modelled subsidence histograms for the 3 selected model runs for 2080. Red is the model uncertainty and in blue the parameter uncertainty. The dots indicate the P2.5, P50 and P97.5 percentile of the uncertainties in the histogram. The bottom row shows the combined histogram (from the middle row).

The differences in the spatial distribution of the uncertainty (years 2020 and 2080) for the workflow described by the main report (best subsidence scenario for a single reservoir realisation) and the workflow described in this appendix (combined realisations) is presented in Figure 70. The scatter in the right figure is a result of the statistical approach. In the statistical approach the uncertainty slightly increases above the Rysum (in the east) and South aquifers.

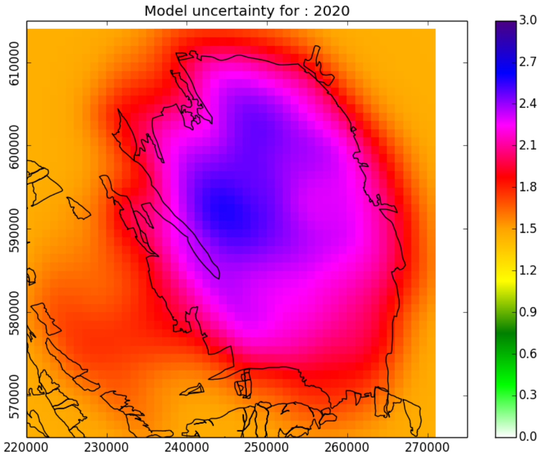
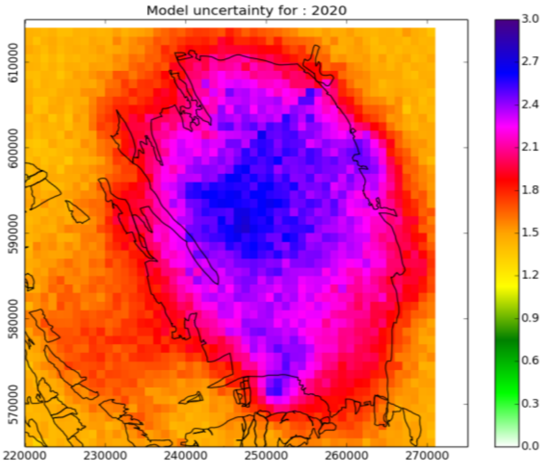
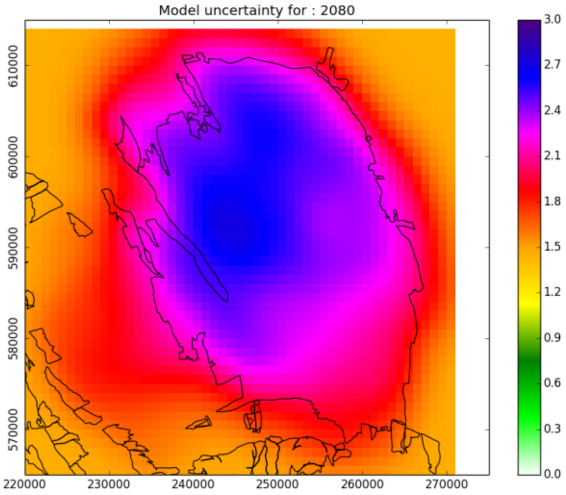
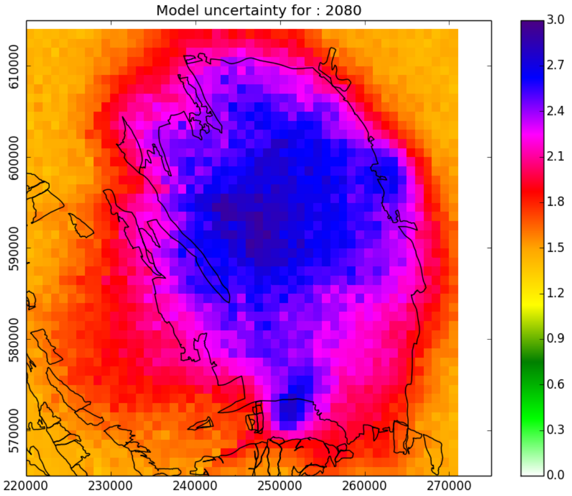
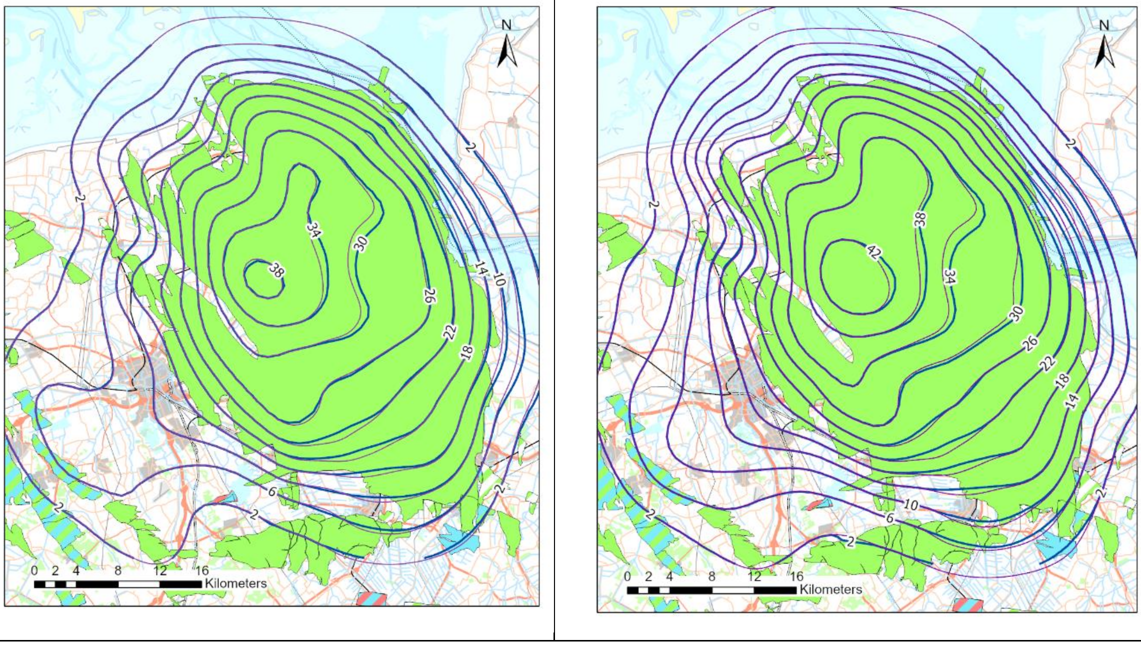
	V6 model only	Combined realisations
2020		
2080		

Figure 70; Model uncertainty ranges for 2020 and 2080. The left column shows the uncertainty (2σ) as presented in the main report and computed for the single subsidence scenario. The right column shows half of the P2.5-P97.5 range which compares to a 2σ range.

Figure 71 shows the difference in the modelled subsidence for the two methods. The purple contours show the subsidence that resulted from the single selected subsidence scenario, following the workflow in the main report. The blue contours represent the P50 value from the statistical workflow as described in this appendix.



Subsidence for 2020

Subsidence for 2080

Figure 71; Subsidence (P50) for 2020 and 2080. The blue contours show the combined realisation and the purple contour the V6 model only.

The comparison indicates that the overall difference in results between the two methods is very small. This is further illustrated by several time series for the benchmarks that were also presented in the main report (Figure 59).

Conclusion

The statistical analysis demonstrates that the incorporation of multiple aquifer realisations does not lead to a significant increase of the model uncertainty. This is mainly caused by the large difference in the probability values for the various reservoir realisations. This results in a small impact on the uncertainty by these additional realisations that are incorporated in the statistical workflow. For the Groningen long term subsidence prediction, it is recommended to adopt the prediction as presented in the main report.

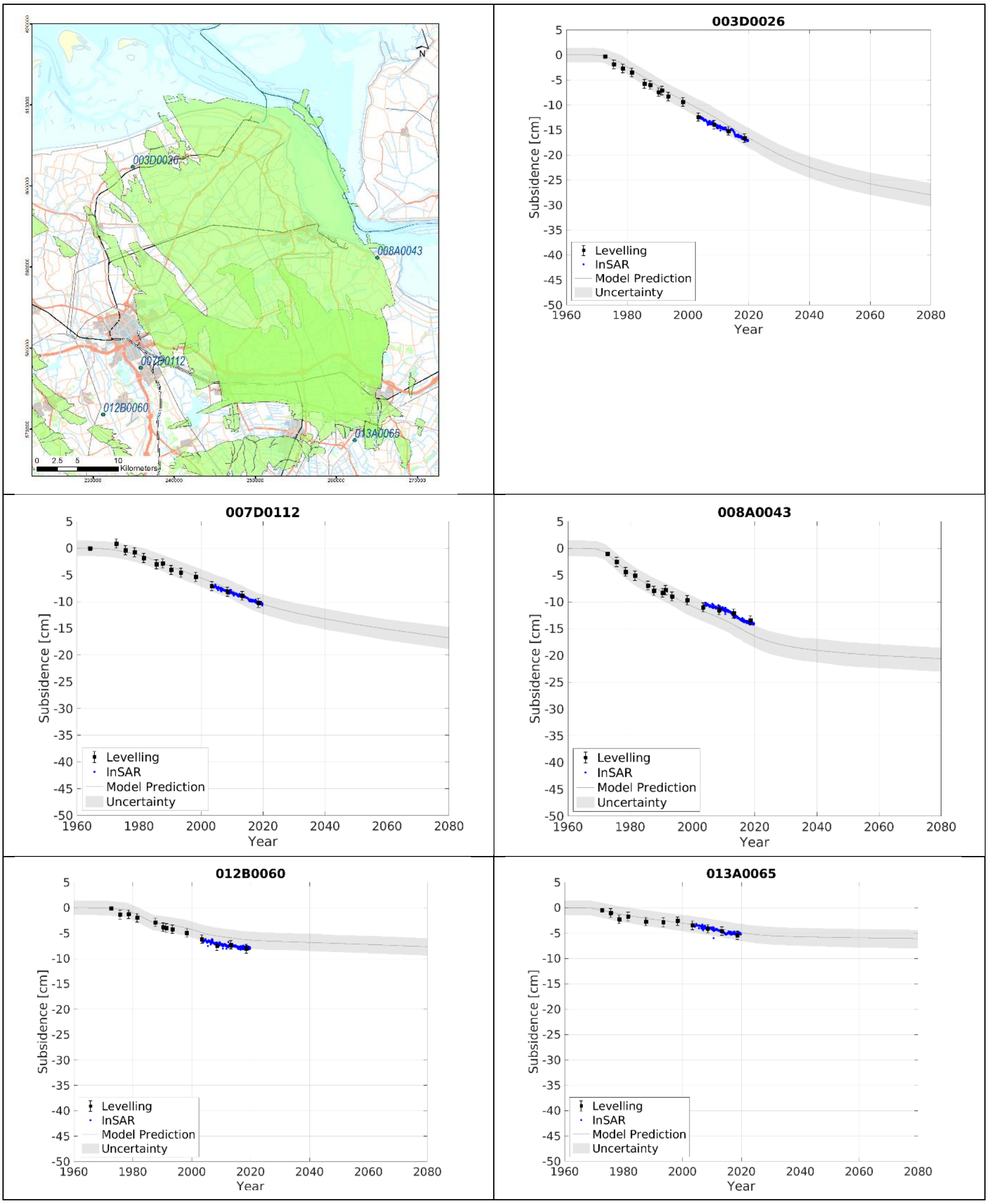


Figure 72; Modelled subsidence based on the statistical analysis at benchmark locations above the aquifers till 2080 compared to the measured subsidence. The dark grey line presents the predicted subsidence (P50). The light grey band shows the P95 confidence interval. Black dots show the levelling measurements plus measurement uncertainty, the blue dots are the InSAR measurements.