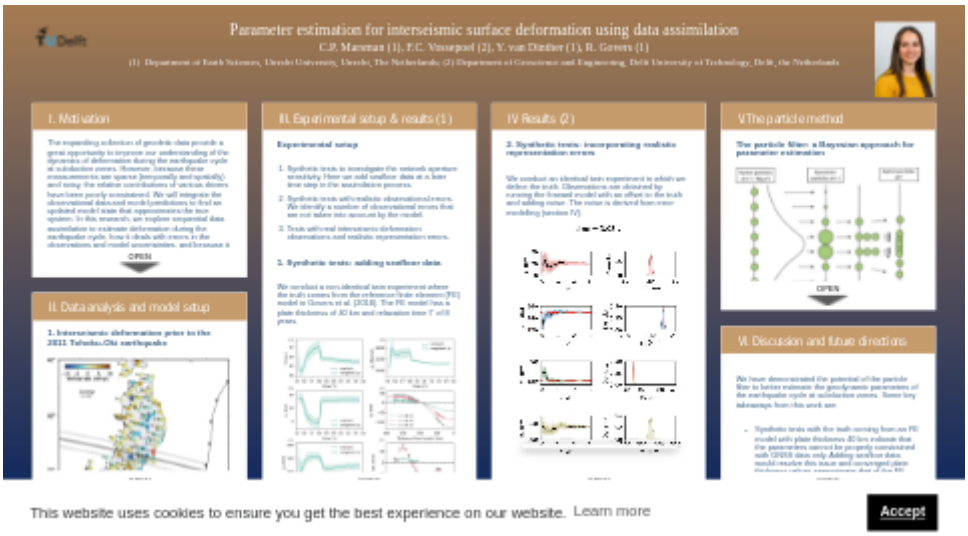
Parameter estimation for interseismic surface deformation using data assimilation



C.P. Marsman (1), F.C. Vossepoel (2), Y. van Dinther (1), R. Govers (1)

(1) Department of Earth Sciences, Utrecht University, Utrecht, The Netherlands; (2) Department of Geoscience and Engineering, Delft University of Technology, Delft, the Netherlands

I. MOTIVATION

The expanding collection of geodetic data provide a great opportunity to improve our understanding of the dynamics of deformation during the earthquake cycle at subduction zones. However, because these measurements are sparse (temporally and spatially) and noisy, the relative contributions of various drivers have been poorly constrained. We will integrate the observational data and model predictions to find an updated model state that approximates the true system. In this research, we explore sequential data assimilation to estimate deformation during the earthquake cycle, how it deals with errors in the observations and model uncertainties, and because it can quantify the uncertainties of the process parameters and geodynamic system states.

⇒We explore data assimilation - and particularly the particle method - in a context of a conceptual mechanical model to estimate model parameter values and their uncertainties, based on noisy interseismic surface displacements.

II. DATA ANALYSIS AND MODEL SETUP

1. Interseismic deformation prior to the 2011 Tohoku-Oki earthquake

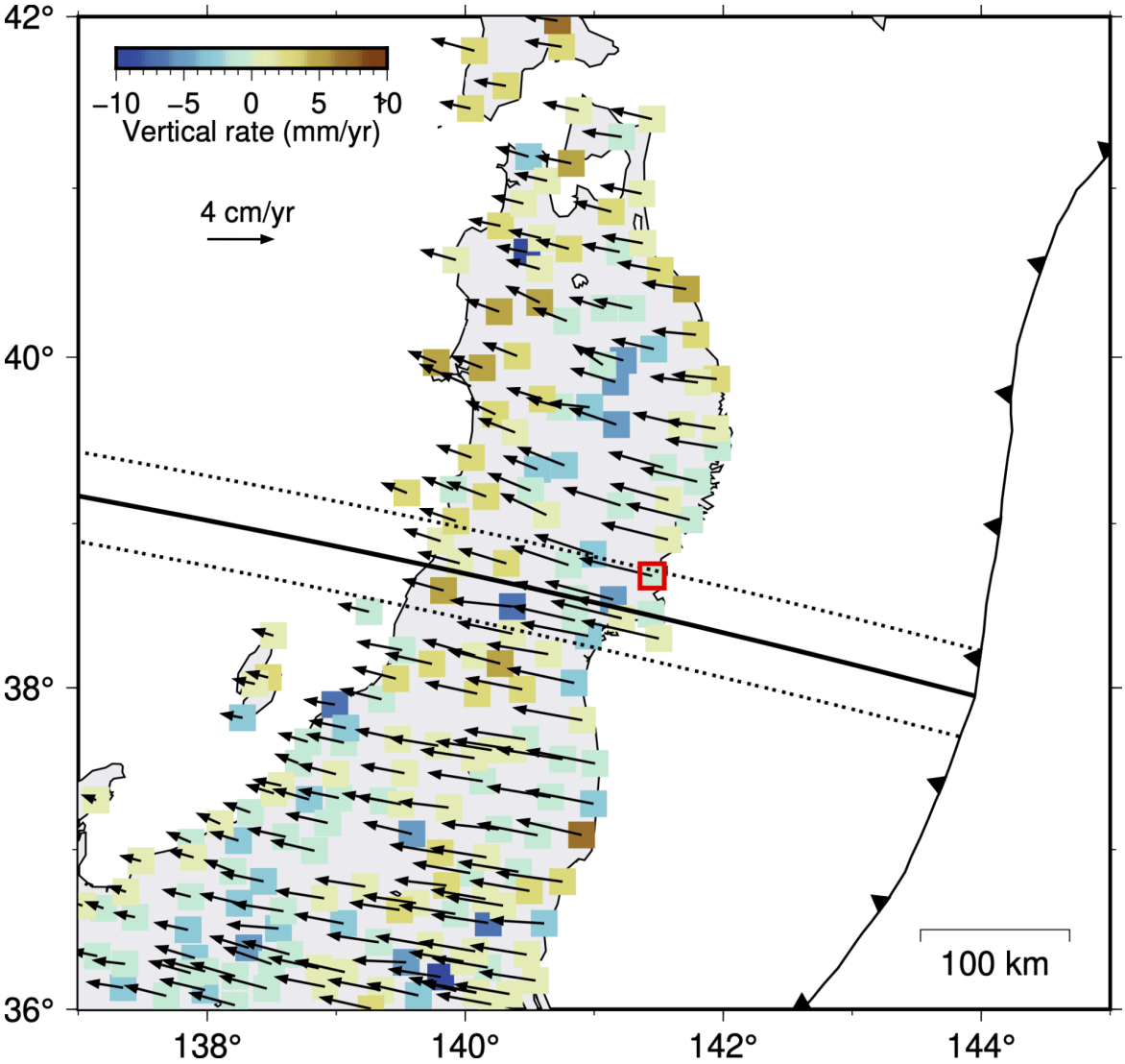


Figure 1. *Vertical and horizontal velocities in Northeastern Japan between January 1997 and May 2000. The solid black line coincides with the largest pre-seismic coupling fraction perpendicular to the trench following Loveless and Meade (2010). The dotted lines show the upper and lower bounds of the swath profile; 30 km north and south of the solid line. Black vectors indicate the horizontal GNSS velocities, whereas colored squares denote the magnitude of the vertical velocities.*

We assimilate vertical surface displacements in Northeast Japan prior to the 2011 Tohoku-Oku earthquake. We select daily measurements (Nakagawa, 2009) between January 1997 and May 2000 processed in ITRF2008 (Altamimi, 2011), operated by the Geospatial Information Authority (GSI) of Japan.

We can estimate the linear trend due to fault locking and plate convergence by means of a least-squares fit to the positional model:

$$u^i(t) = a^i + b^i t + c^i \cos(2\pi t) + d^i \sin(2\pi t) + e^i \cos(4\pi t)$$

$$+ f^i \sin(4\pi t) + \sum_{k=1}^{K} g^i \theta(t_k - t),$$

where the superscript i denotes the ith station, u is the displacement, a is a constant, b is the slope representing the interseismic strain accumulation, $c - f$ are annual and semi-annual seasonal coefficients, g is the change in deformation due to a sudden event, θ is the Heaviside function, and t_k is the time of a sudden event.

The linear component is isolated for the east-, north- and up components. The horizontal and vertical velocities are depicted in Figure 1, showing a landward motion and a flexural pattern for the vertical velocities along the swath.

2. Strong correlation between residual time series

We analyze the GNSS residuals, defined as the difference between the least-squares fit and the raw displacements. The residual time series are highly correlated (Figure 2). The strong spatial coherence is caused by many systematic errors in the GNSS network, which are commonly named common-mode errors (CME's).

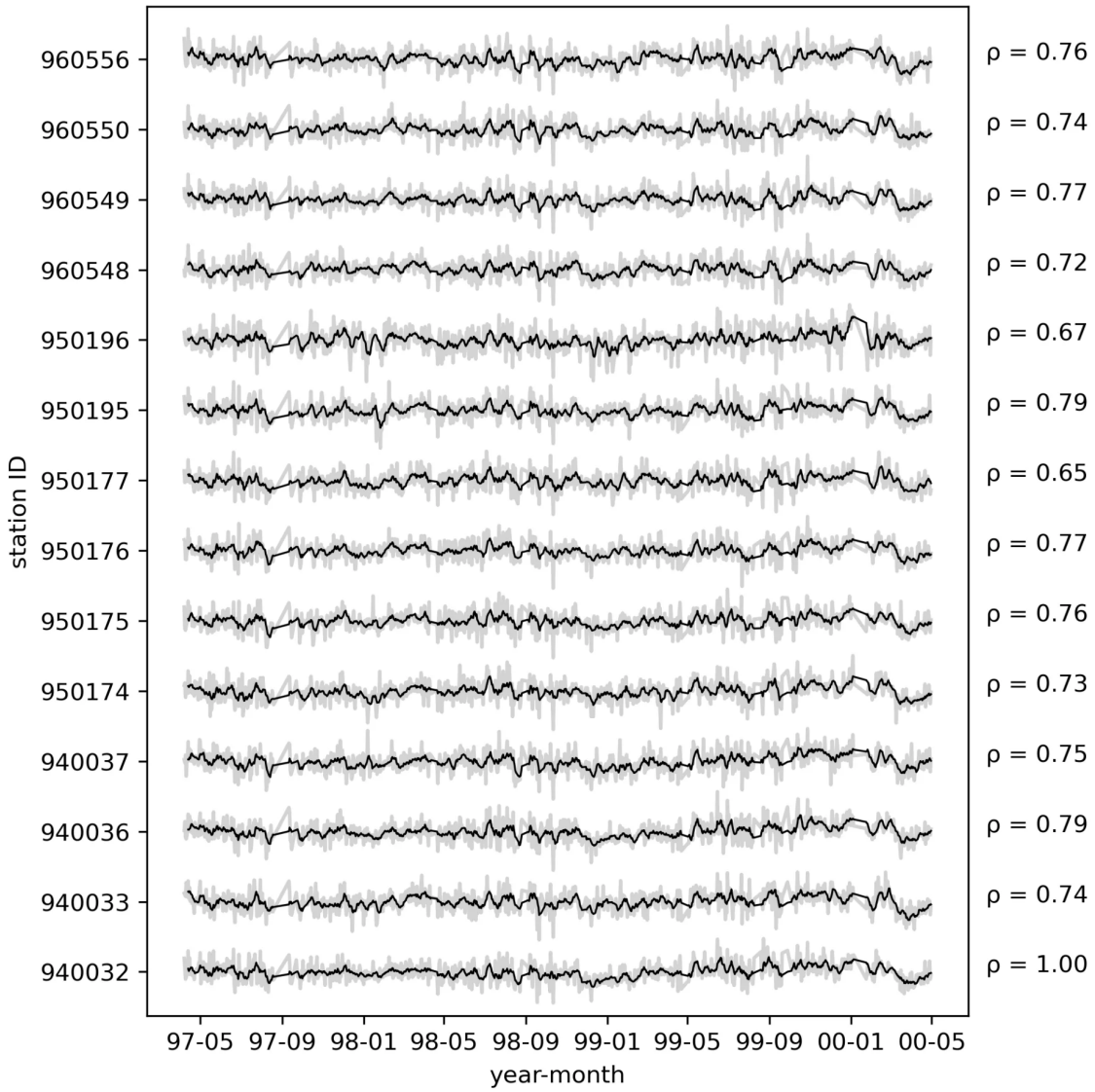


***Figure 2**. Residual time series of the GNSS stations within the swath profile in Figure 1. The numbers on the right hand side are the Pearson correlation coefficient between the station time series and the time series of station 940032.*

3. Error modelling

Discrepancies between our forecast models and geodetic observations are attributed to observation errors, which include measurement errors and representation errors. As representation errors we define:

1. Seasonal effects (annual and semi-annual trends)
2. Steps (resulting from instrumental changes and coseismic displacement)
3. Common-mode error (the systematic errors in the GNSS network)
4. Noise (instrumental)

We can take these errors into account using the particle method. Figure 3 shows these errors at station 940032. To identify the common-mode error (CME) we apply an Empirical Orthogonal Function (EOF) analysis (Preisendorfer, 1988) to find spatial and temporal variability of the residual time series (i.e., the so-called modes). We seek the spatio-temporal structure (i.e., the first mode) contributing the most to the variance in the residual time series, which we define as the systematic common-mode error. The second mode explains most of the remaining variance, and so on. We assume that the sum of the remaining modes is (instrumental) noise.

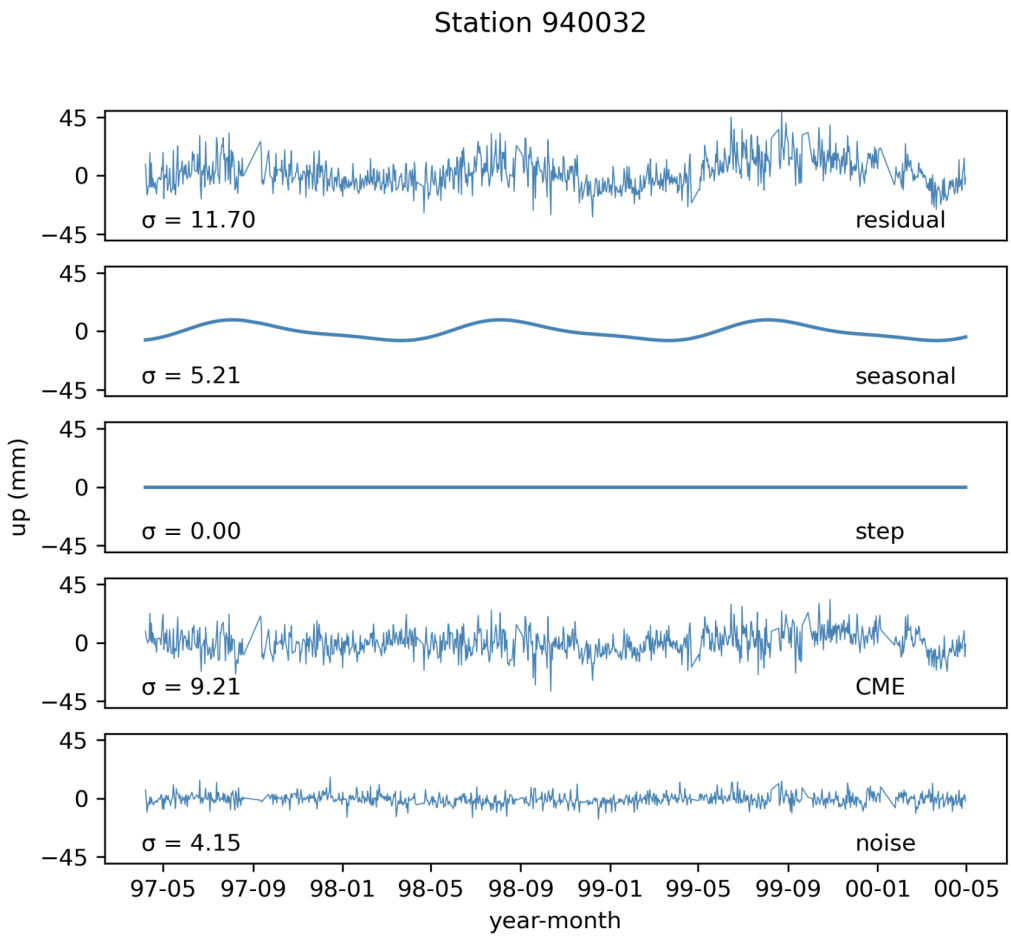


***Figure 3**. Residual time series and associated errors, including seasonal effects, steps, CME and noise.*

4. Model setup

Interseismic flexure of the overriding plate is well-known to precede megathrust earthquakes. We use a two-dimensional elastic flexure model as a conceptual context for exploring assimilation of GNSS vertical deformation observations.

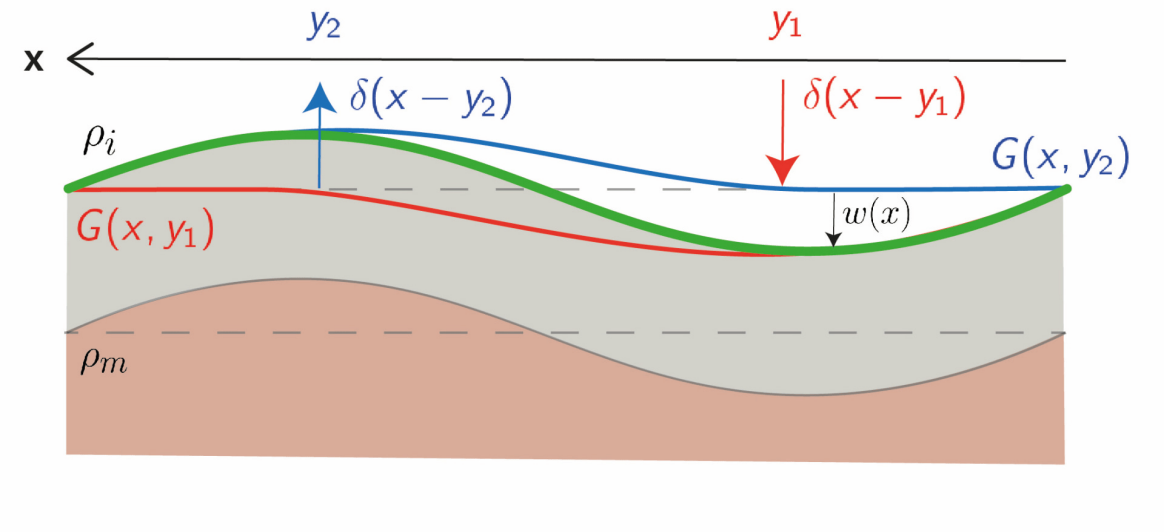


Figure 4. *2D Elastic flexure model setup. An infinite elastic plate is subjected to two line-loads* $\delta(x - y_1)$ *and* $\delta(x - y_2)$ *at* y_1 *and* y_2*, respectively. The elastic plate is underlain my asthenospheric mantle, resulting in a hydrostatic restoring force* $\rho_m g w$ *and driving surface force* $\rho_i g w$*.* ρ_m *and* ρ_i *are the mantle/asthenospheric and water/air densities, respectively,* g *is the gravitational acceleration and* w*is the vertical displacement. The green line is the integrated result of the responses by the two line-loads.*

III. EXPERIMENTAL SETUP & RESULTS (1)

Experimental setup

1. Synthetic tests to investigate the network aperture sensitivity. Here we add seafloor data at a later time step in the assimilation process.

2. Synthetic tests with realistic observational errors. We identify a number of observational errors that are not taken into account by the model.

3. Tests with real interseismic deformation observations and realistic representation errors.

1. Synthetic tests: adding seafloor data

We conduct a non-identical twin experiment where the truth comes from the reference finite element (FE) model in Govers et al. (2018). The FE model has a plate thickness of 40 km and relaxation time τ of 8 years.

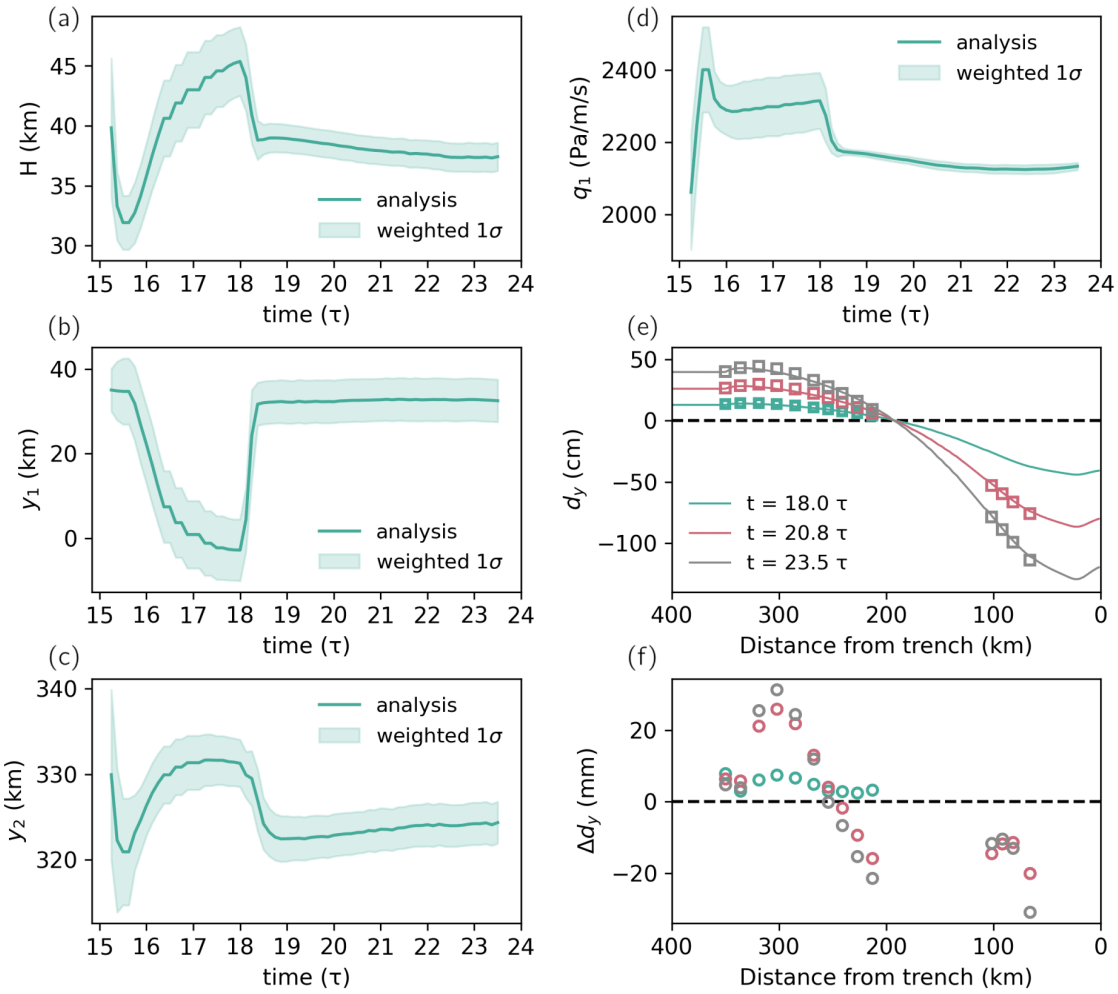


***Figure 5**. Assimilation results of uniform plate thickness model with seafloor data added at $t = 18\tau$. Panels a, b, c and d show the parameter development for the plate thickness, the first point-load location, the second point-load location and the first point-load magnitude, respectively. Panel e shows the development of the deformation in which solid lines are the truth (coming from the FE model in Govers et al. (2018)), and squares denoting the PF analysis at the data locations. Panel f shows the difference between the truth and the PF analysis. Note that the color definitions in (f) are the same as in (e).*

Figure 5 shows that once seafloor data is added to the assimilation process, all parameters converge towards new values (panels a-d). The plate thickness converges towards the truth with smaller uncertainties, which indicates that seafloor data improves the performance of the particle filter.

IV. RESULTS (2)

2. Synthetic tests: incorporating realistic representation errors

We conduct an identical twin experiment in which we define the truth. Observations are obtained by running the forward model with an offset to the truth and adding noise. The noise is derived from error modelling (section IV).

[VIDEO] https://res.cloudinary.com/amuze-interactive/video/upload/vc_auto/v1637920769/agu-fm2021/E5-AB-E7-AF-10-C7-6E-0D-13-90-10-F3-16-76-70-7C/Video/video_elurt2.mp4

***Video 1**. Development of parameters and their kernel density function resulting from synthetic data. In the time series (left panels) the black lines are the weighted mean, the darker colors are the 16/84 percentile of the ensemble spread and the lighter colors are the 5/95 percentile of the ensemble spread. The dashed red lines denote the pre-defined truth of the parameter.*

Video 1 shows that all parameters converge towards the pre-defined truth. Over time, the uncertainty associated with those model parameters decrease and we see a narrow density distribution. Figure 6 shows that there is a strong correlation between the two parameters H and q1 (panel c) where the truth has a larger wavelength. In case of low wavelengths (Htruth = 5 km), the correlation disappears as we do not allow H to be lower than 5 km.

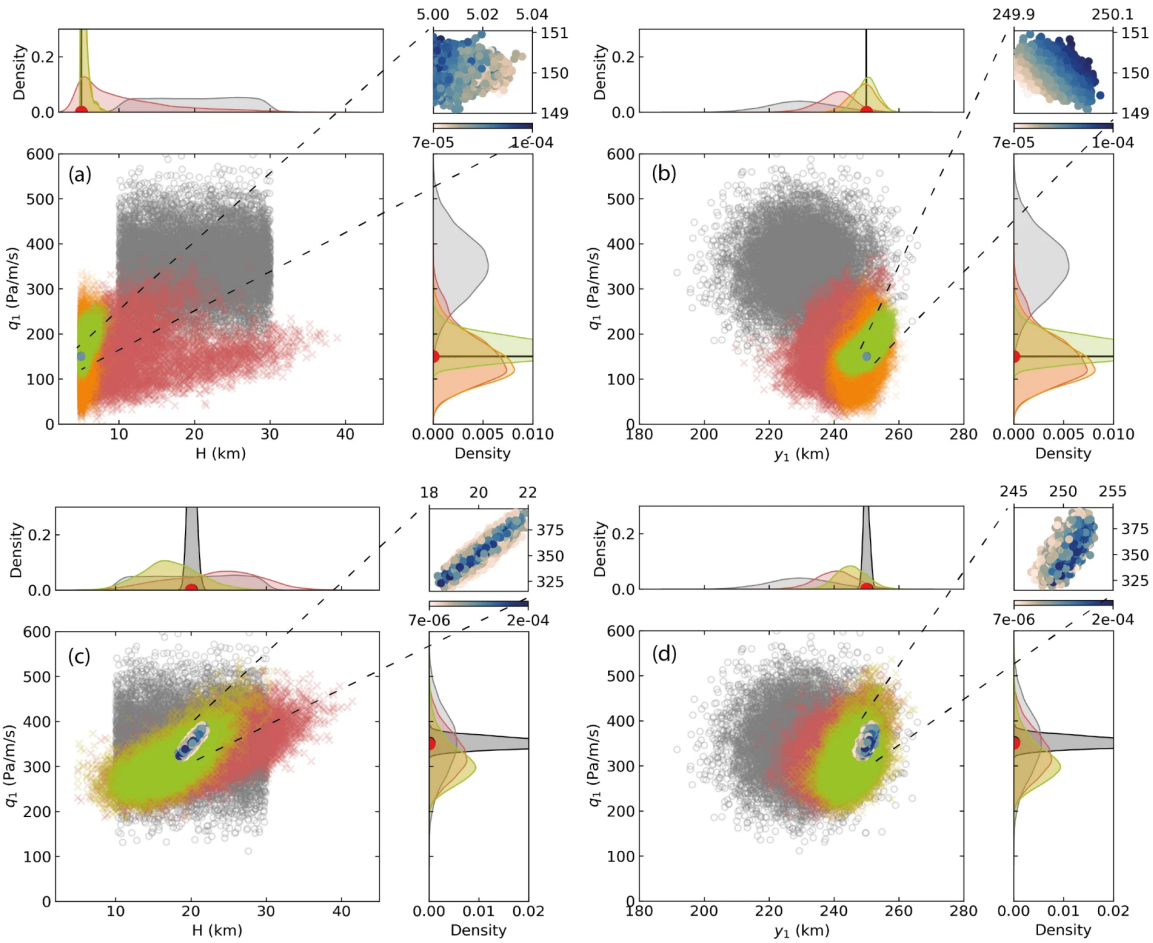


***Figure 6**. Scatter and kernel density function diagrams of the parameter pairs (a,c) H and q1 and (b,d) q1 and y1. (a,b) results have low truth values for H and q1 (low wavelengths) and (c,d) results have larger truth values for H and q1 (high wavelength). The grey, red, orange and green scattered dots are the particle*

distribution at times t=0, t=0.20, t=0.34, and t=0.52 years. The inset shows the parameter distribution at the end of the simulation (t=3.0 years), where colors denote the particle weight. The red dot in the density plots denotes the pre-defined truth.

3. Tests with real interseismic deformation observations and realistic representation errors

Tests with real data showed that parameter convergence could not be achieved unless additional constraints were applied. Video 2 shows the development of the model parameters where we do not allow the plate thickness to be smaller than 5 km.

[VIDEO] https://res.cloudinary.com/amuze-interactive/video/upload/vc_auto/v1638270006/agu-fm2021/54-9D-89-F7-09-3E-BB-AC-A4-36-03-3F-A0-D2-F6-59/Video/video2_real_vuxawt.mp4

***Video 2**. Development of parameters and their kernel density function resulting from real data. In the time series (left panels) the black lines are the weighted mean, the darker colors are the 16/84 percentile of the ensemble spread and the lighter colors are the 5/95 percentile of the ensemble spread. The red lines in the H plots denote the 5 km minimum constraint applied to the plate thickness.*

V. THE PARTICLE METHOD

The particle filter: a Bayesian approach for parameter estimation

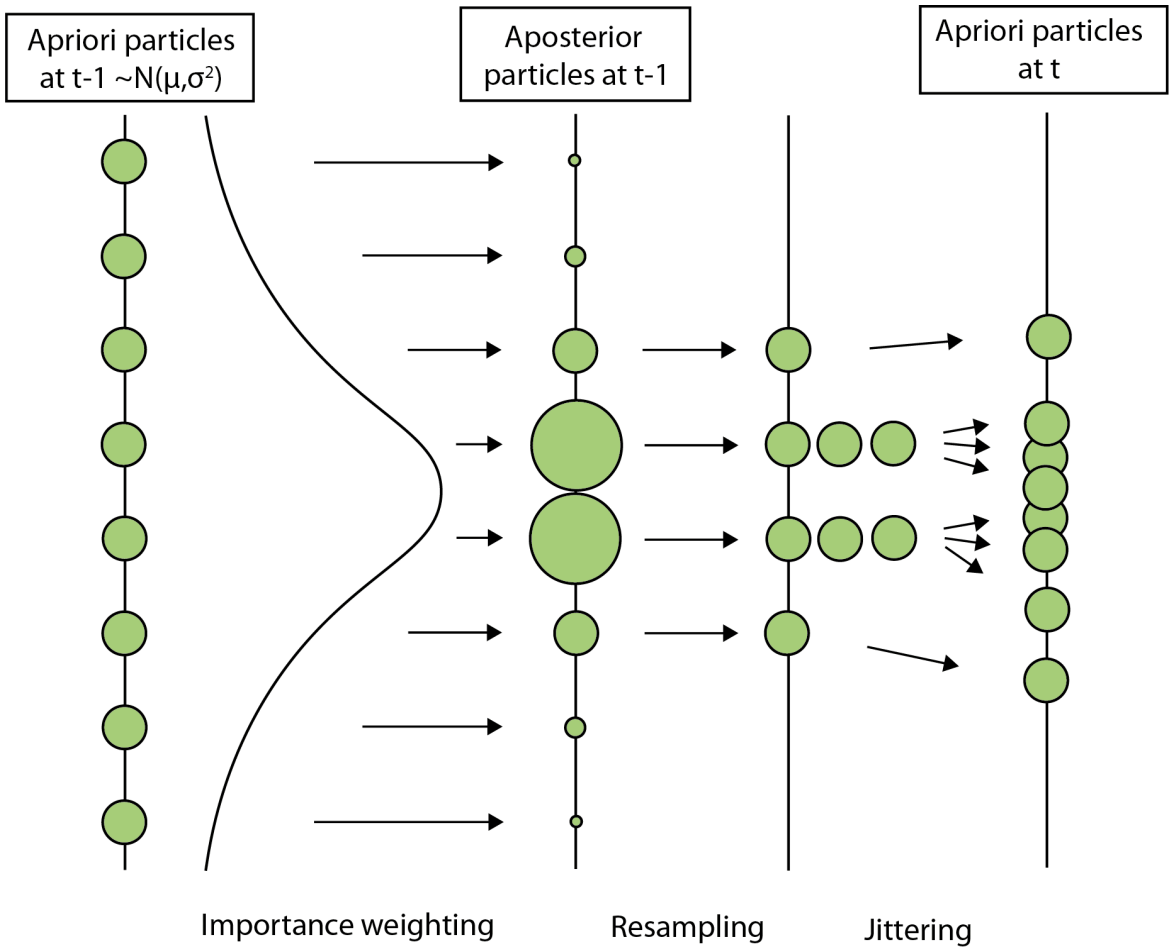


***Figure 7**. The particle method and importance resampling. A priori particles are drawn from a prior distribution (e.g., a normal distribution) and are assigned equal weight. Subsequently, we run a set of forward models (particles), each with a different set of parameter values. Each particle is weighted. In the resampling step, particles with low weight are removed and particles with high weight are duplicated, so that the total ensemble size stays constant. A small perturbation ("jittering") is applied to ensure that the particles are well-distributed.*

The particle method applies a Bayesian inversion to approximate the unknown model state using previous measurements. The particle method uses a Monte Carlo approach to represent the model state probability distribution using a finite number of samples, also called particles.

The usage of the particle filter is shown in Figure 7. The basic steps of the particle consists of

1. initialization;
2. a time update (running the forward model);
3. a measurement update (assigning weights to the particles based on their fit to the observational data);
4. resampling and jittering (i.e., adding a small perturbation).

VI. DISCUSSION AND FUTURE DIRECTIONS

We have demonstrated the potential of the particle filter to better estimate the geodynamic parameters of the earthquake cycle at subduction zones. Some key takeaways from this work are:

- Synthetic tests with the truth coming from an FE model with plate thickness 40 km indicate that the parameters cannot be properly constrained with GNSS data only. Adding seafloor data would resolve this issue and converged plate thickness values approximate that of the FE model.
- Synthetic tests conducted in an identical-twin setting with realistic representation errors, with plate thickness 5 or 20 km, using GNSS data, were able to converge towards the truth.
- Experiments with real data and realistic representation errors do not converge towards constant parameter values without applying constraints. This is likely due to low wavelength solutions being able to provide a better fit to the noisy observations. Seafloor data should provide constraints to avoid such low wavelength solutions.

Future work

Even though the synthetic tests show remarkable convergence and identify parameter uncertainties and correlations, tests with real data reveal that additional measures are needed. The geodynamic model fails to properly approximate the location of the locked zone. A more realistic and complex model is suggested to obtain better approximations of the true system. Future work will focus on the usage of an FE model to include horizontal surface displacements and explore the need for additional types of measurements and constraints.

DISCLOSURES

C.P. Marsman was funded by the Dutch Research Council (NWO), grant number ALW.GO.2019.001

ABSTRACT

Geodetic data provide an opportunity to improve our understanding of the processes and parameters controlling the dynamics of deformation during the earthquake cycle at subduction zones. However, the observations contain noise and are temporally and spatially sparse, whereas dynamical models are unequivocally imperfect. Also, the relative contributions from various drivers of surface deformation are poorly constrained by independent observations. Some drivers may be static or vary slowly in time (e.g., plate motion), whilst others vary significantly during the earthquake cycle (e.g., viscoelastic relaxation). Data assimilation combines prior estimates of dynamical models, with the likelihood of observations into posterior estimates of the state evolution and time-independent parameters of a physical process. We explore the usefulness of data assimilation by using a particle filter to estimate the (spatially variable) elastic thickness of the overriding plate and the extent of the locked zone. We assimilate vertical interseismic surface displacements into a 2D elastic flexural model. The particle filter uses a Monte Carlo approach to represent the state probability distribution by a finite number of realizations ("particles"). We use sequential importance resampling to preserve particles statistically close to the observations and duplicate and perturb them. Synthetic experiments demonstrate that the particle filter effectively estimates 1D elastic thickness from synthetic observations. However, elastic thickness estimates for models with a landward increase in plate thickness show larger uncertainty near the coast as the sensitivity of surface displacements reduces with increasing plate thickness. Interestingly, the effectiveness of the elastic thickness estimation is highly sensitive to network aperture, including GPS/A. Assimilation of interseismic vertical velocities prior to the 2011 Tohoku-Oki earthquake yields estimates of upper plate thickness that agree with previous studies. However, results of the locked zone extent are not as expected, which could be due to missing physics in the relatively conceptual model. These results demonstrate the potential of the particle filter to better understand the geodynamic process parameters of the earthquake cycle at subduction zones.

REFERENCES

Altamimi, Z., Collilieux, X., & Métivier, L. (2011). Itrf2008: an improved solution of the international terrestrial reference frame. Journal of Geodesy,85(8), 457-473.

Govers, R., Furlong, K. P., van de Wiel, L., Herman, M. W., & Broerse, T. (2018). The geodetic signature of the earthquake cycle at subduction zones: Model constraints on the deep processes. Reviews of Geophysics,56(1), 6-49.

Loveless, J. P., & Meade, B. J. (2010). Geodetic imaging of plate motions, slip rates, and partitioning of deformation in japan. Journal of Geophysical Research: Solid Earth,115(B2).

Nakagawa, H. (2009). Development and validation of geonet new analysis strategy (version 4).J. Geogr. Surv. Inst.,118, 1-8.

Preisendorfer, R. (1988). Principle component analysis in meteorology and oceanography. Elsevier.