Supporting Information for

Enhanced coastal shoreline modelling using an Ensemble Kalman Filter to include non-stationarity in future wave climates

Raimundo Ibaceta[1], Kristen D. Splinter[1], Mitchell D. Harley[1] and Ian L. Turner[1]

1 Water Research Laboratory, School of Civil and Environmental Engineering UNSW Sydney, NSW 2052, Australia.

Contents of this file

Introduction

This supporting information provides details on the generation of synthetic shoreline scenarios and on the dual state-parameter EnKF algorithm.

Text S1.

Synthetic scenarios with the ShoreFor model

Ten shoreline timeseries each spanning 20-years at 3-hourly sampling interval were generated using ShoreFor, forced by a set of synthetic wave records based on real observations and specifically designed to characterize seasonal, storm and mixed seasonal-storm wave climates. The expectation is that parameter variability will be modulated by interannual (i.e. ~5-10 years), inter-decadal and beyond (i.e. trends) timescales responding to larger-scale climate teleconnection patterns (e.g. ENSO, PDO) and longer-term trends in wave climate. For instance, Figure S1a shows the dimensionless fall velocity (Ω) from an inshore wave record located in Southeast Australia, which is modulated by ENSO (Ranasinghe et al., 2004) and potentially IPO. Figure S1b shows the ShoreFor model free-parameter c^{a} derived from Splinter et al., (2014) parametrizations and 5-year running $\overline{\Omega}$ magnitudes. A climate-driven interannual (period ~ 10 years) signal is evident, which also has a negative trend and potentially associated with present trends of average wave climate for this site (Hemer et al., 2013). Four synthetic shape functions were defined to represent differing modes of wave climate variability effects in model parameters: a time-invariant (Shape 1), a linear negative trend (Shape 2), a sinusoidal function with a representative period of 10 years (Shape 3) and a step-wise function (Shape 4). These four shapes were mixed together with increasing degree of complexity and differing wave climates (Figure S1c). Figure S2 to Figure S4 show these ten generated timeseries of synthetic shorelines/parameters.

Text S2.

Dual State-Parameter Ensemble Kalman Filter

The Dual State-Parameter EnKF presented in Pathiraja et al., (2016) is employed in this study. A flowchart of the EnKF algorithm is shown in Figure S5 and explained here.

Consider a dynamical system described at any time t by a vector of model parameters θ_t and vector model states (i.e. shorelines) x_t. In the EnKF, both system states and parameters at time t are random variables, represented by ensembles of states $\{x_t^i\}_{i=1:n}$ and parameters $\{\theta_t^i\}_{i=1:n}$, each with n ensemble members. For the present application we set n = 50 for computational efficiency as n>50 showed no increase in model skill. In the following, (-) and (+) indicate updated and background (or prior) ensembles, respectively. For each realization of the algorithm, the Dual State-Parameter EnKF is implemented as follows:

1. *Create a background parameter ensemble.* If t = 0, generate initial estimates of states/parameter ensembles. At any other given time, generate a prior parameter ensemble following Equation (1).

$$\theta_{t+1}^{i-} = g\left(\theta_t^{i+}\right) \text{ for } i = 1:n \quad (1)$$

where g is a function called *parameter evolution model* to be selected by the user. The method of Xiong et al., 2019 is employed, whereby the parameter evolution model simplifies to a noise-inflation model of the type:

$$g\left(\theta_t^{i+}\right) = \theta_t^{i+} + q^i \text{ for } i = 1:n \quad (2)$$

where q^i is white noise (process-noise) for each i-th ensemble member. The magnitude of process-noise reflects the level of parameter inflation and filter capabilities to track time-varying parametrizations. For the present application where $\theta = [c^a, c^e, \phi]$, we generate q using $Var(q) = [1^{-6}, 5^{-7}, 1^{-4}]$. Text S3 and Figure S6 show the sensitivity of process-noise magnitude on parameter estimation.

2. *Generate a set of noisy observations*. The typical accuracy of the shoreline data (defined as R) is used to generate a set of perturbed observations. We assume that shoreline error characteristics are described by a white-noise signal (e.g. Long & Plant, 2012). Therefore, stochastic perturbations are aggregated to the observations vector to generate a set observations ensemble.

$$y_{t+1}^i = y_{t+1}^o + \varepsilon_{t+1}^i, \text{ for } i = 1:n \quad (3)$$

$$\varepsilon_{t+1}^i \sim N\left(0, \Sigma_{t+1}^{y^o y^o}\right) \quad (4)$$

where y_{t+1}^o is the raw observation and $\Sigma_{t+1}^{y^o y^o}$ denotes the observation error covariance matrix (or R^2). For the case of a shoreline model forced by waves, wave data (u_{t+1}^i) could also be perturbed to account for forcing uncertainty in the ensemble characterization (e.g.

derived from inaccuracies in wave modelling or instrumentation). However, we set forcing error to zero here as the focus is on shoreline variability compensated by parameter changes only.

3. *Generate simulated observations using background parameters*. The current state and parameter ensemble is used to calculate simulated observations:

$$\hat{x}_{t+1}^{i} = f\left(\hat{x}_{t}^{i+}, \theta_{t+1}^{i-}, u_{t+1}^{i}\right) \text{ for } i = 1 : n \quad (5)$$

$$\hat{y}_{t+1}^{i} = h\left(\hat{x}_{t+1}^{i}, \theta_{t+1}^{i-}\right) \text{ for } i = 1 : n \quad (6)$$

where f are the model equations (ShoreFor), $\hat{y}_{t+1}^{i}$ are the i=1:n ensemble members of the model simulation generated from the background parameter, h is an operator used to convert model states to observed variables and u_{t+1}^{i} is the forcing vector (e.g. waves)

4. *Kalman Update of Parameters*. The Kalman update equation and the covariance between simulated observations and parameters are used to update the background parameter ensemble:

$$\theta_{t+1}^{i+} = \theta_{t+1}^{i-} + K_{t+1}^{\theta}\left(y_{t+1}^{i} - \hat{y}_{t+1}^{i}\right) \text{ for } i = 1 : n \quad (7)$$

$$K_{t+1}^{\theta} = \Sigma_{t+1}^{\theta\hat{y}}\left[\Sigma_{t+1}^{\hat{y}\hat{y}} + \Sigma_{t+1}^{y^{o}y^{o}}\right]^{-1} \quad (8)$$

Where K_{t+1}^{θ} is the Kalman Gain, $\Sigma_{t+1}^{\hat{y}\hat{y}}$ denotes the covariance matrix of the simulated observations and $\Sigma_{t+1}^{\theta\hat{y}}$ is a matrix of the cross-covariance between background parameters $\left\{\theta_{t+1}^{i-}\right\}_{i=1:n}$ and simulated observed variables $\left\{\hat{y}_{t+1}^{i}\right\}_{i=1:n}$.

5. *Generate simulated observations using updated parameters*. Now the updated model parameter ensemble $\left\{\theta_{t+1}^{i+}\right\}_{i=1:n}$ and the model equations (ShoreFor) are used to generate the background state ensemble $\left\{x_{t+1}^{i-}\right\}_{i=1:n}$.

$$x_{t+1}^{i-} = f\left(x_{t}^{i+}, \theta_{t+1}^{i+}, u_{t+1}^{i}\right) \text{ for } i = 1 : n \quad (9)$$

$$\tilde{y}_{t+1}^{i} = h\left(x_{t+1}^{i-}, \theta_{t+1}^{i+}\right) \text{ for } i = 1 : n \quad (10)$$

with $\tilde{y}_{t+1}^{i}$ the simulated observed variable using updated parameters.

6. *Kalman Update of States*. Finally, the states are updated using the Kalman equation for correlated measurement and process noise (Pathiraja et al., 2016). Note that the standard Kalman Equations (Equations 7 and 8) are not used because they assume independent errors between background and observations, which is no longer valid in the dual parameter-state filter, since the simulated states were generated from

observations from the same time step. Thus, state updating is carried out considering the potential correlation between background noise and observations.

$$x_{t+1}^{i+} = x_{t+1}^{i-} + K_{t+1}^{x}\left(y_{t+1}^{i} - \tilde{y}_{t+1}^{i}\right) \text{ for i = 1 : n} \quad (11)$$

$$K_{t+1}^{x} = \left[\Sigma_{t+1}^{x\tilde{y}} + \Sigma_{t+1}^{\varepsilon_x y^o}\right]\left[\Sigma_{t+1}^{\tilde{y}\tilde{y}} + \Sigma_{t+1}^{\varepsilon_{\tilde{y}} y^o} + \left(\Sigma_{t+1}^{\varepsilon_{\tilde{y}} y^o}\right)^{T} + \Sigma_{t+1}^{y^o y^o}\right]^{-1} \quad (12)$$

$$\varepsilon_{x_{t+1}}^{i} = x_{t+1}^{i-} - \hat{x}_{t+1}^{i},\ \varepsilon_{\tilde{y}_{t+1}}^{i} = \tilde{y}_{t+1}^{i} - \hat{y}_{t+1}^{i} \quad (13)$$

Where $\Sigma_{t+1}^{\varepsilon_x y^o}$ is the covariance between $\left\{\varepsilon_{x_{t+1}}^{i}\right\}_{i=1:n}$ and the observations and $(\,)^T$ is the transpose operator. $\Sigma_{t+1}^{x\tilde{y}}$ is the cross-covariance matrix between simulated observed variables $\left\{\tilde{y}_{t+1}^{i}\right\}_{i=1:n}$ and states $\left\{x_{t+1}^{i-}\right\}_{i=1:n}$. For more details on this step and its derivation refer to the Appendix in Pathiraja et al. (2016).

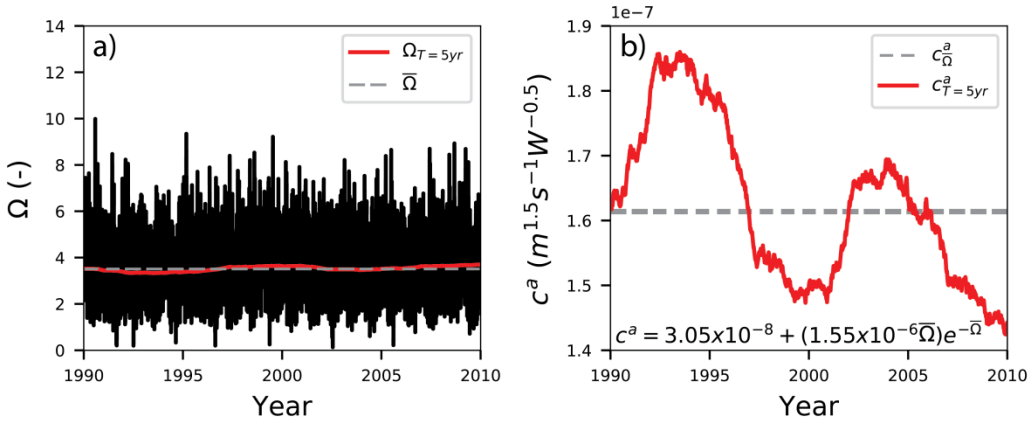
Text S3.

Algorithm sensitivity to the magnitude of process-noise (q)

The magnitude of process-noise reflects the level of parameter inflation and the filter capabilities to track time-varying parametrizations. For the present application where θ = $[c^a, c^e, \phi]$, we set $Var(q) = [1^{-6}, 5^{-7}, 1^{-4}]$ for all model runs. The algorithm sensitivity to this magnitude is presented for Scenario 10 (dt = 1 day, R = 1 m) by varying $Var(q)$ by one and two larger/lower order of magnitudes in process-noise by:

1. Setting $Var(q_\emptyset)$ magnitude unchanged and decreasing $Var(q_{c^a})$, $Var(q_{c^e})$ magnitudes (Figure S6a)
2. Setting $Var(q_{c^a}), Var(q_{c^e})$ magnitude unchanged and decreasing $Var(q_\emptyset)$ magnitude (Figure S6b)
3. Setting $Var(q_\emptyset)$ magnitude unchanged and increasing $Var(q_{c^a})$, $Var(q_{c^e})$ magnitudes (Figure S6c)
4. Setting $Var(q_{c^a})$, $Var(q_{c^e})$ magnitude unchanged and increasing $Var(q_\emptyset)$ magnitude (Figure S6c)

In general, results show that reducing process-noise magnitude results in slowed convergence in parameter-estimation for the three parameters (Figure S6a,b). In turn, this slowed convergence results in the miss-specification of the remaining parameters. In contrast, larger magnitudes of process noise (Figure S6c,d) do not slow the parameter convergence but incorporate higher frequency variability beyond the time-scales of interest.



c)

Scenario	1	2	3	4	5	6	7	8	9	10
Wave climate	Seasonal	Storm	Seasonal	Storm	Seasonal	Storm	Seasonal	Storm	Mixed	Mixed
Shape 1	$c^{a,e}$, Φ	$c^{a,e}$, Φ	c^{e}, Φ	c^{e}, Φ	Φ	Φ	Φ	Φ	$c^{a,e}$	
Shape 2			c^{a}	c^{a}			c^{a}	c^{a}		c^{a}
Shape 3					$c^{a,e}$	$c^{a,e}$	$c^{a,e}$	$c^{a,e}$		$c^{a,e}$
Shape 4									Φ	Φ

Figure S1. a) 20-year dimensionless fall velocity timeseries from an inshore wave record potentially modulated by climate drivers. b) c^a obtained from the parametrization proposed in Splinter et al., (2014) for the 20-year period (grey dashed line) and using a 5-year running mean of the dimensionless fall velocity. c) Synthetic scenarios generated from different wave climates and combinations of parameter variability (Shape 1 to Shape 4).

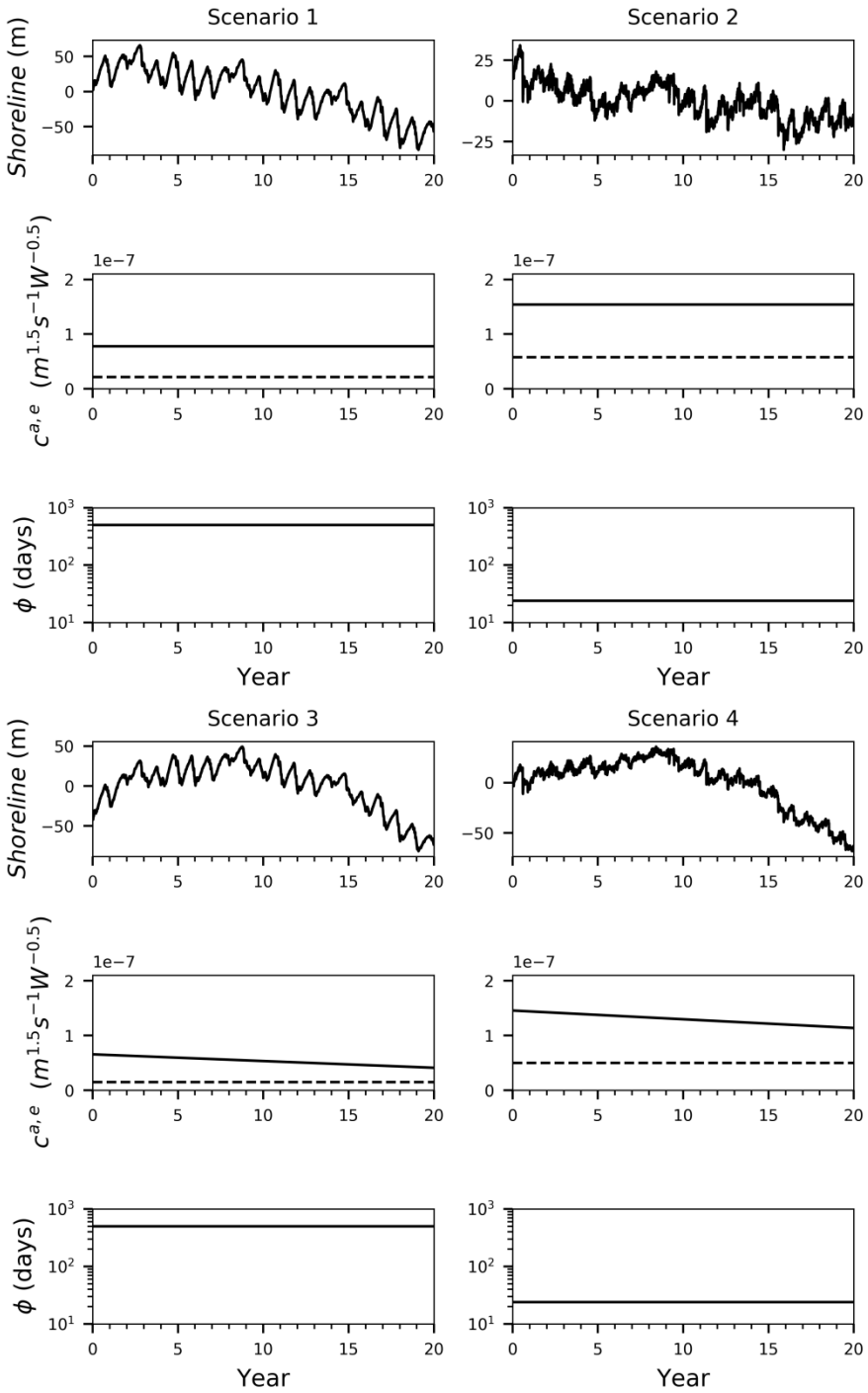


Figure S2. Synthetic shorelines and parameter timeseries for Scenarios 1, 2, 3 and 4. For each scenario, top panel : shoreline timeseries, middle panel : c^a (continuous line) and c^e (dashed line) and bottom panel : ϕ

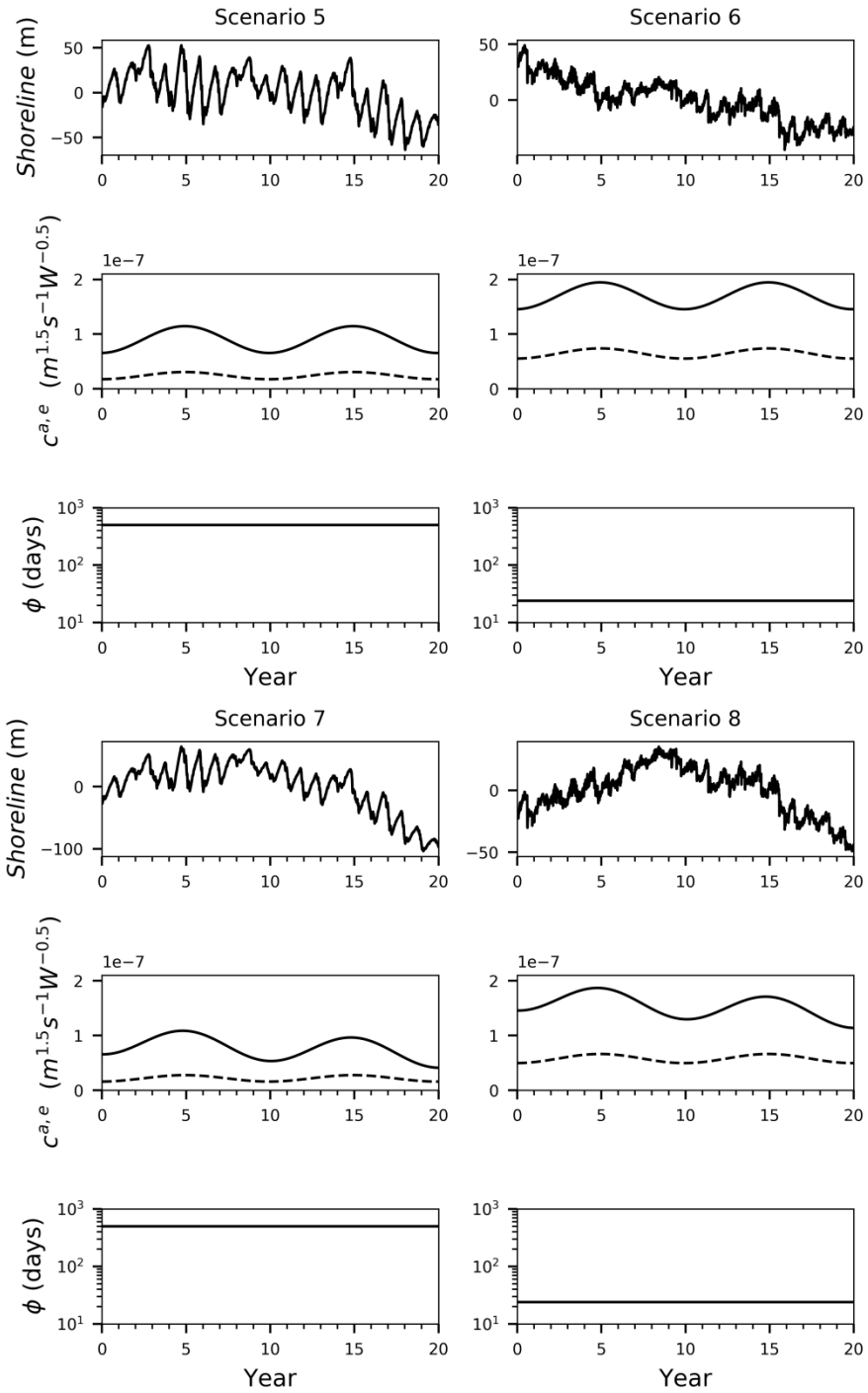


Figure S3. Synthetic shorelines and parameter timeseries for Scenarios 5, 6, 7 and 8. For each scenario, top panel : shoreline timeseries, middle panel : c^a (continuous line) and c^e (dashed line) and bottom panel : ϕ

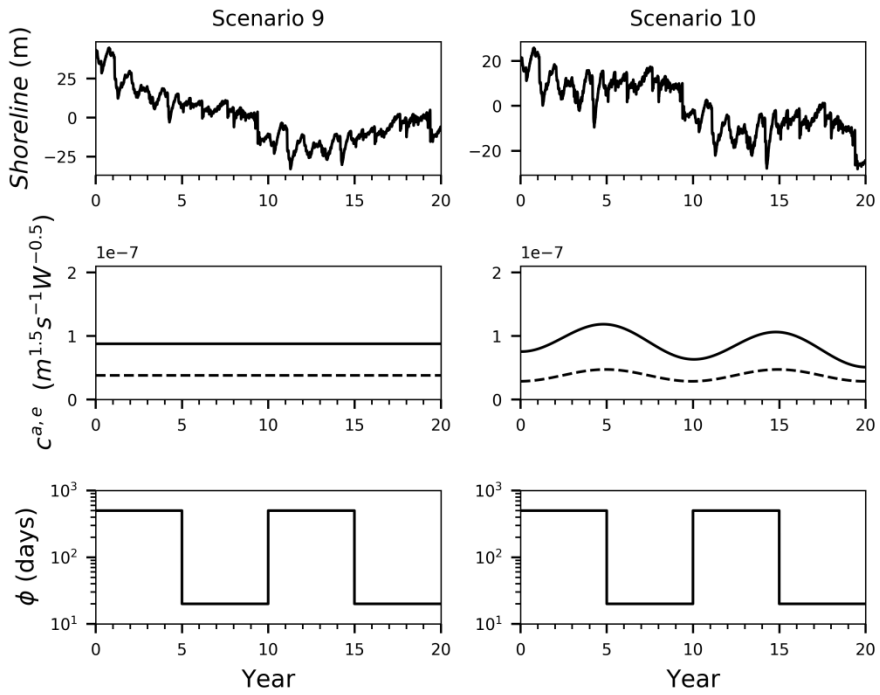


Figure S4. Synthetic shorelines and parameter timeseries for Scenarios 9 and 10. For each scenario, top panel : shoreline timeseries, middle panel : c^a (continuous line) and c^e (dashed line) and bottom panel : ϕ

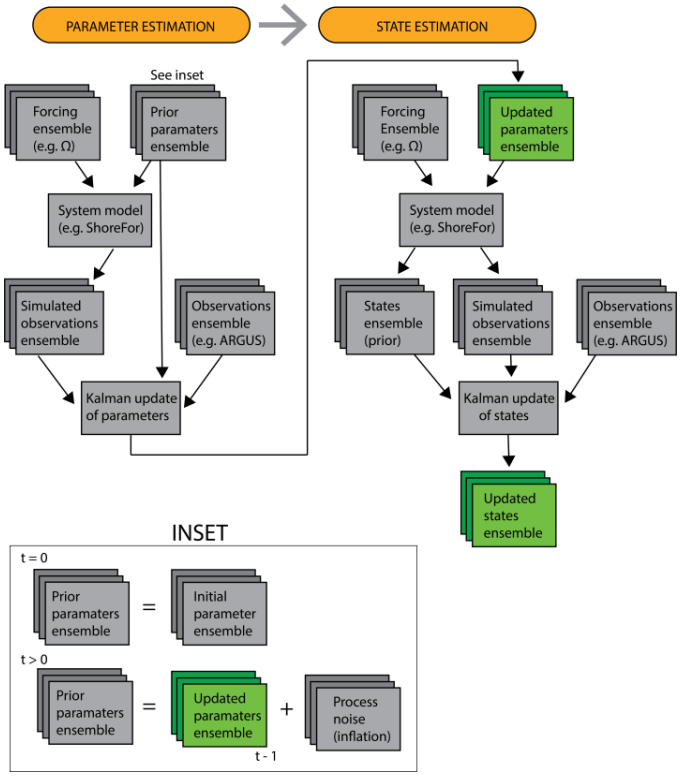


Figure S5. Flowchart of the dual state-parameter EnKF

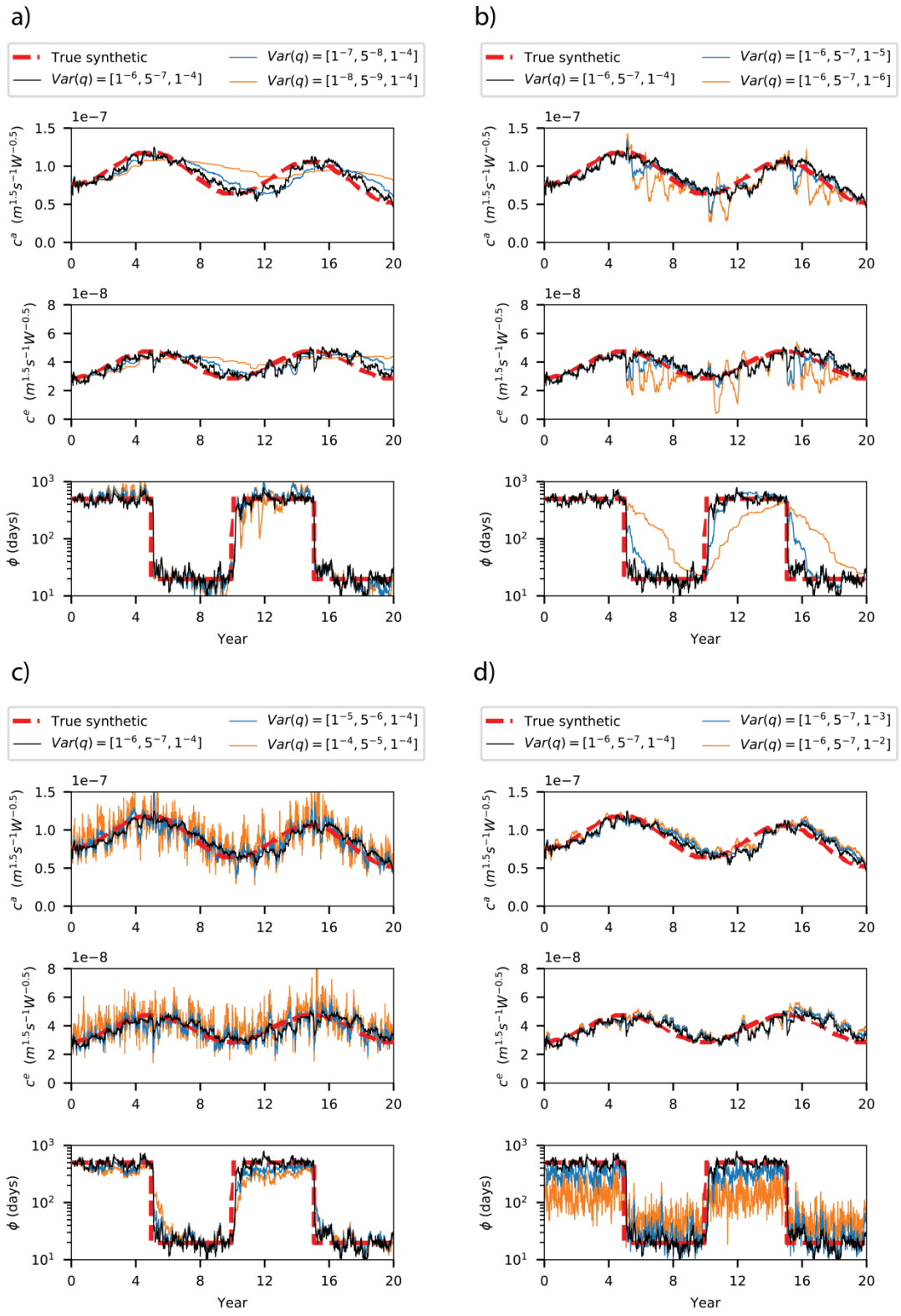


Figure S6. Algorithm sensitivity to different magnitudes of process noise. Black lines are obtained using the magnitude of process-noise adopted for every model run in this work. Red

dashed lines are the true synthetic value and the remaining continuous lines indicate the effect of different magnitudes of process noise (see legends).

References

Hemer, M. A., Fan, Y., Mori, N., Semedo, A., & Wang, X. L. (2013). Projected changes in wave climate from a multi-model ensemble. Nature Climate Change, 3(5), 471–476. https://doi.org/10.1038/nclimate1791

Long, J. W., & Plant, N. G. (2012). Extended Kalman Filter framework for forecasting shoreline evolution. Geophysical Research Letters, 39(13), n/a–n/a. https://doi.org/10.1029/2012GL052180

Pathiraja, S., Marshall, L., Sharma, A., & Moradkhani, H. (2016). Hydrologic modeling in dynamic catchments: A data assimilation approach. Water Resources Research, 52(5), 3350–3372. https://doi.org/10.1002/2015WR017192

Ranasinghe, R., McLoughlin, R., Short, A., & Symonds, G. (2004). The Southern Oscillation Index, wave climate, and beach rotation. Marine Geology, 204(3–4), 273–287. https://doi.org/10.1016/S0025-3227(04)00002-7

Splinter, K. D., Turner, I. L., Davidson, M. A., Barnard, P., Castelle, B., & Oltman-Shay, J. (2014). A generalized equilibrium model for predicting daily to inter-annual shoreline response. Journal of Geophysical Research: Earth Surface, 119, 1936–1958. https://doi.org/10.1002/2014JF003106

Xiong, M., Liu, P., Cheng, L., Deng, C., Gui, Z., Zhang, X., & Liu, Y. (2019). Identifying time-varying hydrological model parameters to improve simulation efficiency by the ensemble Kalman filter: A joint assimilation of streamflow and actual evapotranspiration. Journal of Hydrology, 568(November 2018), 758–768. https://doi.org/10.1016/j.jhydrol.2018.11.038