Hydrology and Earth System Sciences

Operational hydrological data assimilation with the recursive ensemble Kalman filter

H. K. McMillan[1], **E. Ö. Hreinsson**[1], **M. P. Clark**[2], **S. K. Singh**[1], **C. Zammit**[1], **and M. J. Uddstrom**[1]

[1]National Institute of Water and Atmospheric Research, Christchurch, New Zealand
[2]National Center for Atmospheric Research, Boulder, Colorado, USA

Correspondence to: H. K. McMillan (h.mcmillan@niwa.co.nz)

Abstract. This paper describes the design and use of a recursive ensemble Kalman filter (REnKF) to assimilate stream-flow data in an operational flow forecasting system of seven catchments in New Zealand. The REnKF iteratively updates past and present model states (soil water, aquifer and surface storages), with lags up to the concentration time of the catchment, to improve model initial conditions and hence flow forecasts. We found the REnKF overcame instabilities in the standard EnKF, which were associated with the natural lag time between upstream catchment wetness and flow at the gauging locations. The forecast system performance was correspondingly improved in terms of Nash–Sutcliffe score, persistence index and bounding of the measured flow by the model ensemble. We present descriptions of filter design parameters and explanations and examples of filter behaviour, as an information source for other groups wishing to assimilate discharge observations for operational forecasting.

1 Introduction

Many river systems are susceptible to floods that occur rapidly after rainfall events; meaning that effective flood warning systems must use numerical weather prediction (NWP) forecasts to provide sufficient warning time. By coupling NWP and hydrological models, river flows may be forecast several days into the future. When hydrological models are used in such a capacity, forecast errors may occur due to uncertainties in model structure (e.g. unmodelled processes), model parameters, initial conditions (e.g. catchment wetness state determined from previous model runs) and input data (e.g. rainfall and temperature forecasts). Error accumulation can be reduced by data assimilation, a generic technique to combine model predictions with observations, balancing the uncertainty in each, to make the best estimate of the current system state. In hydrology, the difference between modelled and observed flows, together with a quantification of the errors in each, can be used to determine an optimal update to model states and hence to improve future forecasts.

Data assimilation in hydrological models is a relatively recent advance, but one which has been enthusiastically taken up, with various approaches being used: refer to recent reviews by Reichle (2008) and Liu et al. (2012), and references therein. The observations used to update model states can include river flows (Seo et al., 2003), soil moisture (Brocca et al., 2010; Flores et al., 2012), snow-covered area and snow water equivalent (Clark et al., 2006; Zaitchik and Rodell, 2009; Andreadis and Lettenmaier, 2006), and satellite observations of discharge (Neal et al., 2009; Andreadis et al., 2007). Sequential use of different observation types is also possible. For example, Aubert et al. (2003) assimilated streamflow and soil moisture observations. The model states that are updated include in-channel water volume (Ricci et al., 2011), soil water (Lee et al., 2011), groundwater (Zhou et al., 2011; Clark et al., 2008) and snow water equivalent (De Lannoy et al., 2012; Slater and Clark, 2006).

Data assimilation methods used in hydrology include the Kalman filter (and its variants), particle filters and variational methods. These alternative methods and the assumptions required for their optimality are reviewed by Liu and Gupta (2007) and Liu et al. (2012). Kalman filters are the most common: their method alternates a "prediction step" where the model is advanced one step forward in time, and

an "update step" where the current observation(s) is assimilated. The Kalman filter assumes that the model is linear, and uses the relative covariances of model state errors and observation errors to define an optimal "Kalman gain" by which the model is shifted towards the observations. The method is optimal if model states and observations are multivariate normal, and processes are linear. Many extensions have been proposed to support nonlinear processes. These include the extended Kalman filter, which replaces the linearised model with a series of local linear approximations; but this approach may be inaccurate if higher-order model derivatives are significant. The ensemble Kalman filter (EnKF; Evensen, 1994) replaces the linearised model with an ensemble of model realisations. This both allows for nonlinear models and removes the need to propagate the error covariances, which are replaced by sample covariances direct from the ensemble. However, linearisation is still assumed during the update step. The EnKF has found favour in hydrology, where highly nonlinear problems are common.

Particle filters are similar to ensemble Kalman filters in that they rely on an ensemble of model realisations ("particles"). However, instead of updating system states in the particles, the filter merely changes the weights assigned to each particle. This approach makes few assumptions, requiring neither linearisation in prediction or update steps, nor assumptions of Gaussian errors common to Kalman filter approaches. The limitation of particle filters is that a very large number of particles is needed if the model state is high dimensional. Variational approaches are less common in hydrology, as they generally require an adjoint model, which is difficult to derive, but have the potential to be much less computationally expensive than Kalman filter methods for complex systems. Variational approaches allow for nonlinearity through a series of local, linear approximations, requiring similar assumptions to the extended Kalman filter.

In practice, strict criteria such as local linearity or Gaussian errors may not be met, but filtering algorithms can still be applied successfully. For example, Weerts and El Serafy (2006) compared the performance of the EnKF and particle filter, and found that the EnKF was most robust in cases of high uncertainties in model structure or inputs, or low number of ensemble members. Regularisation or heuristic methods are frequently employed to improve filter performance in suboptimal conditions, and can often be framed as a Bayesian update, whereby prior knowledge of the system is taken into account. For example, Kalman gains may be smoothed or approximated by a constant (Sorenson and Madsen, 2004); reducing the spatial support of updates is also common (Houtekamer and Mitchell, 1998, 2001). Particle filters often suffer from degeneracy (i.e. many particles have negligible weights), and various re-sampling or regularisation methods may be used to preserve the particle distribution (e.g. Liu and Gupta, 2007; Noh et al., 2011).

Data assimilation is being incorporated into operational flow forecast systems. In France, Météo-France runs an ensemble streamflow prediction system that assimilates streamflow data to update soil moisture in a distributed model. The system uses the best linear unbiased estimator (BLUE) (Thirel et al., 2010a,b). The same weather ensemble prediction system is used to drive a lumped soil-moisture-accounting type rainfall-runoff model at Cemagref (now Irstea), which uses discharge to update the routing store (Randrianasolo et al., 2010). Komma et al. (2008) showed how the EnKF could be applied operationally in the Kamp catchment in Austria. Their system updated runoff directly, with soil moisture and storage reservoirs updated using a heuristic similarity method. Seo et al. (2009) report on the operational procedure used at the US National Weather Service where data assimilation is carried out using a variational method. Generic data assimilation tools including EnKF are also available, which have been coupled with some popular hydrological models such as HBV and SOBEK (Weerts et al., 2010).

A significant challenge for data assimilation in hydrology is the natural time lag between catchment state and river flow (i.e. the time of concentration), especially in large catchments. This means that an update of distributed model states at the same time as the flow observations may not be physically realistic and can lead to under- or over-shoot in flow correction at later timesteps (Mendoza et al., 2012). This problem led some authors to reject updates to model states having this characteristic, e.g. groundwater stores, in favour of fast flow pathways such as surface runoff or in-channel water volume (Randrianasolo et al., 2010; Berthet et al., 2009). Rakovec et al. (2012) apply the EnKF and allow both routing stores and distributed states to be updated, but only consider the instantaneous covariances, which leads to significant updates only to the routing stores. In the Ovens River catchment, Australia, Li et al. (2011) found that updating soil moisture states resulted in a lagged response in discharge leading to a poorer model performance, but also slower degradation of the forecast accuracy.

A number of solutions have been proposed to account for catchment time lag. These typically build on the standard data assimilation methods listed previously. Noh et al. (2011) propose a lagged regularised particle filter (LRPF), which propagates particles forward over an extended time period to aggregate model response prior to the observation. Standard particle filter methods are then used to update the particles and weights at a pre-determined, fixed lag time. Tests using the distributed Water and Energy transfer Processes model showed that the LRPF performed similarly to a non-lagged particle filter, but was less sensitive to assumptions regarding process noise. Smoothing methods are defined as those which "batch-process" a set of observations, to find the model states that minimise aggregated error over the time window. Smoothing methods are used in reanalysis problems, where availability of "future" data (past the analysis time) is assumed, and implicitly incorporate the relationship between observations and historical model states. Refer to

McLaughlin (2002) and Liu and Gupta (2007) for further discussion of smoothing vs. filtering problems. Smoothing methods include the ensemble Kalman smoother (Evensen and van Leeuwen, 2000; Ravela and McLaughlin, 2007), the fixed lag Kalman smoother (Cohn et al., 1994; a more efficient version of the smoother, which limits historical data to a fixed interval prior to the observation time), and variational methods, e.g. Seo et al. (2009) who test a fixed-lag variational method to update states in the Sacramento soil moisture accounting model.

Pauwels and De Lannoy (2006) developed an extension to the EnKF, the retrospective ensemble Kalman filter, which updates model states preceding an observation. The approach was designed for a model with unit hydrograph routing, and allows updates to all states contributing to the present observation via the unit hydrograph. Their application of this approach to discharge assimilation using the HBV model (Pauwels and De Lannoy, 2009) showed that only a marginal improvement in results was obtained over a non-assimilating model. This was because the method requires the hydrological model to be re-initialised at the observation time minus the lag time, which allowed model error to accumulate, and override the benefits of the updated initial conditions.

In this paper, we build on the concept from Pauwels and De Lannoy (2006) that flow observations can be used, within an EnKF framework, to update model states prior to the observation. This approach allows for the natural lag time of the catchment. However, our method is designed to update model states, from the maximum lag time up to the time of latest observation, through iterative application of the EnKF. This method differs from that of Pauwels and De Lannoy (2006, 2009) who use a single update of the model state vector that has been augmented with all model states during the lag period. Our aim is to address the problem of model error accumulation by providing updated states as close as possible to the start of the forecast period. We named this method the recursive ensemble Kalman filter (REnKF).

This paper describes an implementation of the REnKF at NIWA (National Institute of Water and Atmospheric Research) for operational flood forecasting in New Zealand, using the distributed model TopNet running over a variety of catchments. The aim of this paper is to present (1) a thorough exploration of the filter algorithm, set up decisions and parameter choices, (2) quantification of the performance of the REnKF in a variety of catchments, including comparison with the EnKF and (3) investigation of filter behaviour including error parameter sensitivity, ensemble spread, perturbation size and magnitude of updates by lag time, location and model state. Our intention is that, by describing our filtering method and assessing its performance in an operational setting, we will provide an information source for other groups wishing to assimilate discharge observations for operational forecasting.
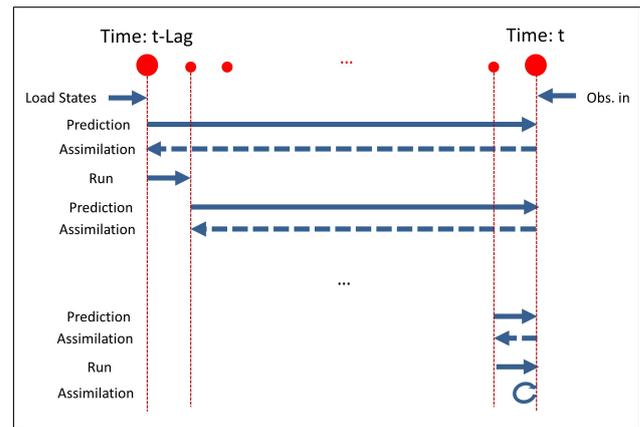

Fig. 1. Schematic showing REnKF algorithm to assimilate an observation at time t.

2 REnKF implementation

2.1 Algorithm

Our implementation of the REnKF runs as a wrapper around the EnKF as implemented by Clark et al. (2008). Following Clark et al. (2008), we allow the filter to update distributed states of soil storage, aquifer storage and surface (routing) storage. The REnKF cycles through prior model states and updates each one based on a streamflow observation at the current timestep. In this iterative approach, the earliest states (greatest time before the observation time) are updated first, and the model is re-run to calculate the updated streamflow predictions; then the states one timestep closer to the current are updated, and the model again re-run, etc. Complete details of the algorithm are given in Appendix A, and the algorithm is represented as a schematic in Fig. 1.

As a recursive application of the EnKF, the REnKF is optimal under the same conditions as the EnKF, i.e. linear approximation of the update step, and Gaussian distributions of the model and observation errors. Following Clark et al. (2008) we process flow observations in log space, to improve the Gaussian approximation of errors. By construction, the distributions of model and observation errors are not independent after the first EnKF iteration for each observation time (because some information from the observation is propagated via the model during the prediction step). However, as with other regularisation strategies, we aim for a practical method that addresses the problems with excessive correction in filter implementations that occur when the catchment lag time is not considered.

2.2 Assimilation parameters

The REnKF requires a number of parameters in order to execute the EnKF updating step: these are described here. Refer to Sect. 4 (System results) for details on the choice of

parameter values. As with all data assimilation schemes, the EnKF optimises the model behaviour according to the trade-off between errors in the model and errors in the observed data. Therefore specification of the magnitude of these two error sources is necessary for the statistical optimality of the filter.

2.2.1   Observation error

Errors in streamflow observations derive from errors in river stage measurement, and errors in the rating curve used to transform stage to discharge. The latter include errors in stage and velocity gaugings, assumption of a particular form of stage–discharge relationship, extrapolation beyond the maximum gauging and cross-section change (McMillan et al., 2010; Di Baldassarre and Montanari, 2009; Westerberg et al., 2011). The EnKF assumes that errors in streamflow are normally distributed. We follow Clark et al. (2008) who improved filter performance by transforming observed and modelled streamflow to log space before computing the Kalman gain. Hence, the standard deviation of the observed error is assumed proportional to the log discharge, with a proportionality constant ε_{obs} that must be specified (Eq. 1):

$$\left[\log(q_{true}) - \log(q_{obs}) \right] \sim N\left(0, \ \left[\varepsilon_{obs} \cdot \log(q_{obs}) \right]^2 \right). \quad (1)$$

In addition, the ensemble square root filter variant of the EnKF is used, which uses a modified Kalman gain to remove the need to perturb the observations (for further details see Clark et al., 2008). This is advantageous as the observation perturbation magnitude was found to be a very sensitive parameter (Moradkhani et al., 2005).

2.2.2   Model error

The EnKF quantifies model error by using the variance of streamflow predictions from an ensemble of model realisations. Model errors can be caused by uncertainties in model inputs, model structure and parameter values. The number of ensemble members must be specified, and depends on the number of states to be updated: we used 50 members, which was found to be sufficient using the same model structure in a catchment larger than any tested here (Clark et al., 2008). The ensemble is created using stochastic perturbations of forcing and state variables: in this case we perturb the forcing precipitation depth and state variables for soil moisture and depth to the water table. For a detailed description of the approach to perturbing the ensemble, refer to Clark et al. (2008).

The perturbations are parameterized as fractional error parameters for precipitation (ε_p), change in soil moisture between subsequent timesteps (ε_s), and baseflow (ε_z). Errors are assumed to follow a uniform distribution $U[-\varepsilon_i \cdot q_i, +\varepsilon_i \cdot q_i]$ where ε_i is the error parameter corresponding to flux q_i. The derived errors are then applied to the precipitation depth, soil moisture state variable, and depth to water table (by inverting the baseflow equation in the latter case):

$$p' \sim p + U\left[-\varepsilon_p \cdot p, \quad +\varepsilon_p \cdot p \right] \quad (2a)$$

$$s_t' \sim s_t + U\left[-\varepsilon_s \cdot (s_t - s_{t-1}), \quad +\varepsilon_s \cdot (s_t - s_{t-1}) \right] \quad (2b)$$

$$z' \sim -m \left\{ \log \left[\frac{q_b + U\left[-\varepsilon_z \cdot q_b, \quad +\varepsilon_z \cdot q_b \right]}{K_0 \cdot m} \right] + \lambda \right\}. \quad (2c)$$

By specifying the perturbations in this way, model errors are permitted to be large when model fluxes are large (e.g. during storm events) and small when model fluxes are small (e.g. during drier periods). This approach differs from some previous studies where error variances are defined to be temporally constant (Reichle et al., 2002; Crow and Van Loon, 2006), but corresponds more closely to the modeller's conceptualization of error magnitudes, for example that rainfall errors can be approximated by a multiplicative error term (McMillan et al., 2011).

2.2.3   Spatial and temporal correlation of model error

In our application of the REnKF, stochastic perturbations are applied in each subcatchment of the watershed model, and at each model timestep. The perturbations are correlated in space and time to represent spatial dependency of forcing errors and temporal persistence of model errors. The correlation is introduced by using Gaussian random fields parameterised by the decorrelation time τ and the correlation length L: these parameters are required for each perturbation (rainfall, soil moisture, depth to water table). The method is described by Clark et al. (2008) and uses techniques from Clark and Slater (2006) and Evensen (2003). Figure 2 shows an example of the precipitation perturbation patterns from the Grey catchment where the spatial correlation length is set to 10 km vs. 65 km. Details of correlation parameter choices for all the test catchments are given in Sect. 4.1.

The perturbation correlations play a part in controlling the spread of the flow ensemble at the catchment outlet. Where perturbations are sustained in time, the ensemble spread is increased due to the memory of the catchment. Where spatial correlation length is large compared to the catchment size, many subcatchments undergo perturbations of the same sign, which increases ensemble spread. The last effect is also influenced by the geometry of the river network.

2.2.4   Lag time

The lag time is an REnKF-specific parameter that determines the maximum timestep difference between a streamflow observation and a prior model state that can be updated during the assimilation of that observation. Lag time can be conceptualised as the time of concentration of the catchment, i.e. the maximum time taken for rainfall occurring at the headwaters to affect streamflow at the gauging site. An initial estimate of the lag time could therefore be made a priori, although our experience showed that similar lag times performed well across a wide variety of catchments (refer to Sect. 4.1).
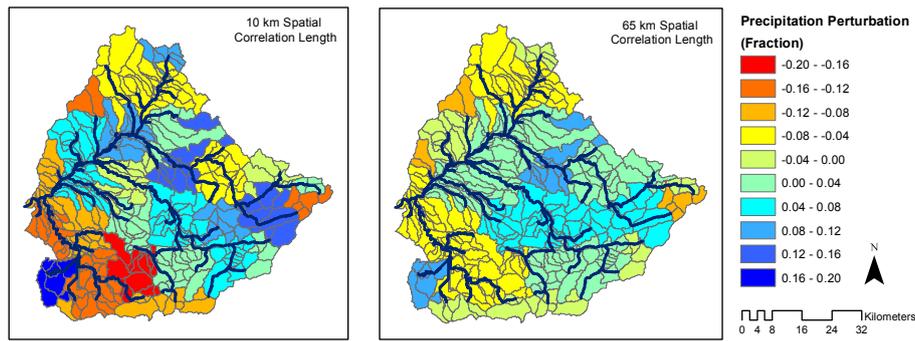
Fig. 2. Example maps of fractional precipitation perturbation across the Grey catchments with spatial error correlation of 10 km (left panel) and 65 km (right panel).
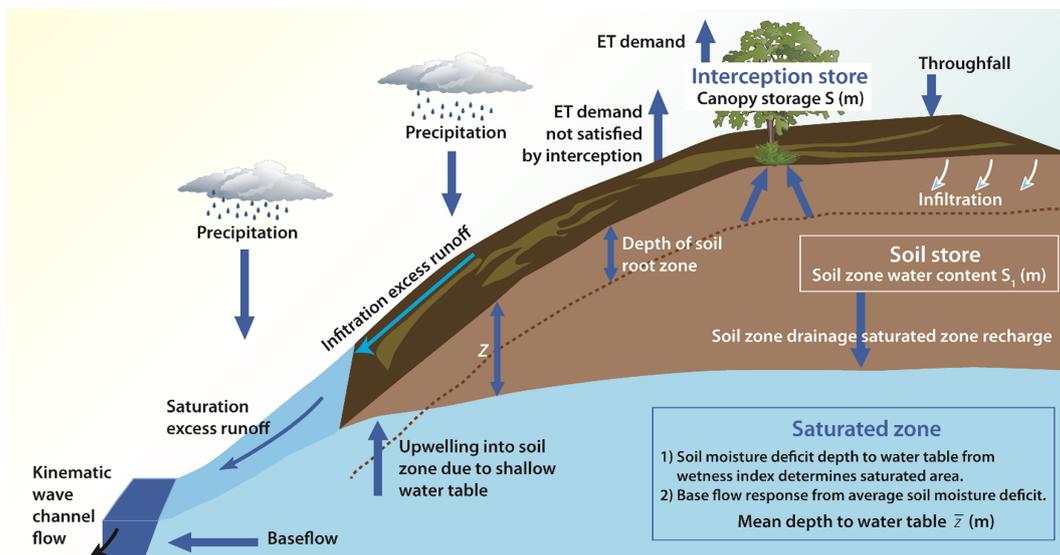

Fig. 3. Schematic representation of the water balance component of TopNet (adapted with permission from Bandaragoda et al., 2004).

3 NZ operational system for NWP and flood forecasting

The aim of this work was to develop and evaluate an operational system for flood forecasting, with streamflow data assimilation provided by the REnKF. The system requires a regional NWP model for New Zealand (NZ) coupled to a hydrological model. An overview of this system is provided here.

3.1 Numerical weather prediction

The regional NWP model NZLAM (NZ Local Area Model) forecasts the atmospheric state over New Zealand for 48 h ahead. NZLAM is a local implementation of the UK Met Office Unified Model System, and derives its lateral boundary conditions from the global Unified Model. NZLAM is warm cycled on a 6 h basis, using initial conditions from previous runs. The model resolution is currently 12 km, using a rotated lat/long grid with 324 × 324 points in the horizontal and 70 levels in the vertical, up to 80 km height. NZLAM

incorporates a full 3DVAR data assimilation system (Lorenc et al., 2000) of observations from land, ship, and upper air stations, drifting buoys, aircraft, and satellites. Model forecasts of meteorological variables including precipitation and temperature are provided at an hourly time step.

3.2 Hydrological model

NZLAM forecasts provide input to a hydrological model, which simulates soil moisture, groundwater levels and river flow over the period of the NWP forecasts. We use the TopNet model (Fig. 3), which combines TOPMODEL concepts of sub-surface storage controlling the dynamics of the saturated contributing area and baseflow recession (Beven and Kirkby, 1979; Beven et al., 1995) with a kinematic wave channel routing algorithm (Goring, 1994). This approach leads to a model that can be applied over large watersheds using smaller sub-basins within the large watershed as model elements (Ibbitt and Woods, 2002; Bandaragoda et

al., 2004). Complete model equations are provided by Clark et al. (2008). TopNet is routinely used for hydrological modelling applications in New Zealand, and uses nationally available information on catchment topography, physical and hydrological properties. This information is derived from a digital river network (River Environment Classification; Snelder and Biggs, 2002), 30 m digital elevation model, and land cover and soil databases (Land Cover Database; Land Resource Inventory; Newsome et al., 2000). For the applications described here, Strahler 3 subcatchments of typical size $10\,\mathrm{km}^2$ are used.

TopNet uses seven calibrated parameters for each subcatchment. To reduce the dimensionality of the parameter estimation problem, initial values for the parameters were estimated from the sources described. The spatial distribution of the parameters was then preserved, and the values were adjusted uniformly using a spatially constant set of parameter multipliers. Calibration used precipitation and climate data from Tait et al. (2006) who interpolated data from over 500 climate stations in New Zealand across a regular $0.05°$ latitude/longitude grid (approximately $5\,\mathrm{km} \times 5\,\mathrm{km}$). The precipitation data were bias-corrected using a water balance approach (Woods et al., 2006). These data are provided at daily time steps, and are disaggregated to hourly data before use in the model. The parameter values used here are those in current use for the operational forecasting system. The calibration methods varied by catchment according to the responsible hydrologist, and consisted of a semiautomatic method using either Monte Carlo simulation (two catchments), or the ROPE (RObust Parameter Estimation) calibration method (Bardossy and Singh, 2008) (five catchments) to obtain a small ensemble of possible parameter sets. This was followed by review by a hydrologist to determine a single preferred set based on visual inspection of the model simulation results. Note that parameter values are not perturbed during the assimilation. Instead, we allow for model error by perturbing the state variables.

3.3 Streamflow assimilation

The hydrological model is run on a 6 h cycle to mirror the NWP forecast input. Each time a new NZLAM forecast becomes available (0, 6, 12, 18 h), the hydrological model is run forward for 48 h to provide river flow forecasts to the end user with as little latency as possible. Real-time hourly telemetered observations of streamflow then become available to the system. At forecast time +5 h (i.e. 5, 11, 17, 23 h), the hydrological model is re-run in assimilation mode. The REnKF uses available observations to update the model states, and hence provide optimal initial conditions ready for the next forecast run in the 6 h cycle.
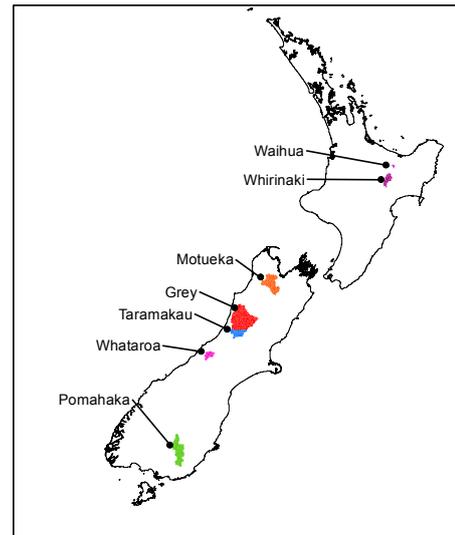

Fig. 4. Location of the seven test catchments within New Zealand.

3.4 Forecast delivery

The weather and hydrological forecasts are made available to end users through the environmental forecasting tool "Eco-Connect", which includes a web delivery application (Uddstrom, 2011). Graphical model output includes maps of NWP output, and location-specific forecasts for climate variables, rainfall and flow. The flow forecasts show observed flow up to current time, followed by predicted flow for a 48 h future time period. Flow forecasts from previous NZLAM cycles can be viewed to check for consistency between forecasts. The ensemble spread can optionally be superimposed on the forecast, as an indication of model error.
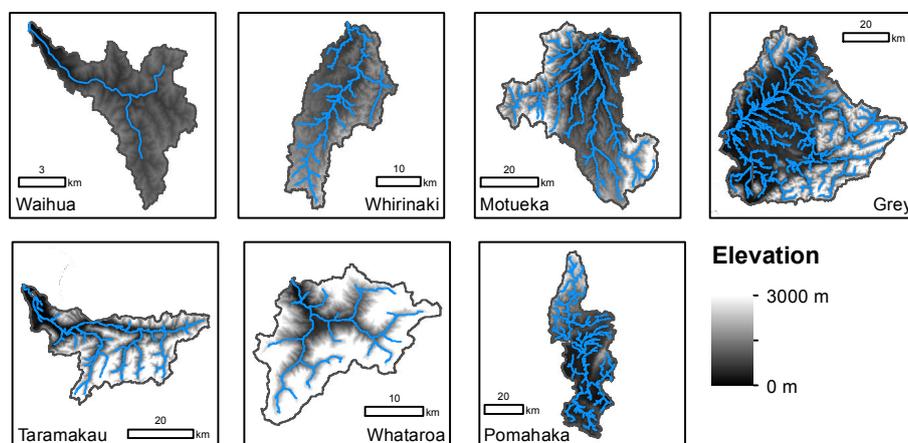
3.5 Test catchments

The operational system currently includes seven catchments that are not influenced by hydropower operations and hence maintain natural flows. All these catchments were used to test the REnKF data assimilation method. Figure 4 shows the location of these catchments within NZ, and Fig. 5 shows a closeup of each catchment including major rivers and elevation. The catchments span a diverse range of topography, land cover and climate: these physical characteristics are summarised in Table 1.

3.6 Summary of forecast performance tests

We performed a series of tests on the REnKF in operational forecasting mode, using the seven catchments described in the previous Sect. 3.5. Section 4 shows the results of these tests, and they are summarised here for clarity. In each case, we compared the REnKF against three benchmarks: (1) no assimilation, (2) free-running ensemble and (3) standard EnKF. The forecasts were tested for lead times between

Table 1. Catchment characteristics.

Name		Waihua	Whataroa	Whirinaki	Taramakau	Motueka	Pomahaka	Grey
Area [km^2]		50	450	510	880	1760	1870	3820
Elevation range [m]		80–760	65–3090	195–1240	10–2240	70–1840	40–1480	15–1945
Average stream slope [°]		1.31	5.67	1.12	1.97	1.41	0.49	0.84
	Forest	97	34	93	43	64	9	65
	Pasture	1	0	4	4	18	73	12
Land cover [%]	Scrub/tussock	2	22	3	35	17	17	18
	Rock/ice/other	0	44	0	18	1	1	5
Mean annual snow water equivalent [m]		0.0	0.9	0.0	0.2	0.0	0.0	0.1
Mean annual precipitation [m]		1.83	11.4	1.61	6.29	1.67	0.96	3.95
Mean annual runoff [m]		1.14	9.35	0.90	5.72	1.07	0.49	3.17

Fig. 5. Close-up view of each of the test catchments, with the river network and elevation shown.

0–6 and 42–48 h, in 6 h increments, and the forecast performance was quantified in terms of Nash–Sutcliffe score, persistence index, and percentage of time that the flow measurement lies within the ensemble bounds. Further, we tested the sensitivity of ensemble spread (for state variables and forecast flow) and model performance to the REnKF fractional error parameters. We then used a variety of graphical methods to investigate the update behaviour of the filter across dimensions of space, observation time and time lag.

4 System results

4.1 Flow forecasting results

To test the REnKF flow forecasting, the system was run in hindcast mode for the water year 1 April 2011–31 March 2012. This relatively short time period was used in order to limit model running time (a 1 yr run for the Grey catchment took approximately 9 days using an Intel Xeon CPU E5540, 2.53 GHz). The hydrological model runs on a 6 h cycle, with a 48 h forecast produced at each cycle. Therefore to create a continuous series for performance assessment, we concatenated the first 6 h of each forecast. The

system was run in three modes: (1) free running ensemble, no assimilation, (2) EnKF assimilation and (3) REnKF assimilation. As explained in Sect. 3.3, for each 6 h window the assimilation has been completed up to the start time of the forecast, but the model ensemble is free-running during the 6 h forecast period. The results from the REnKF assimilation for all catchments are shown in Fig. 6. The flow simulations are shown for an example period of 2 months containing some of the larger flow events of the water year. Rank histograms are also given, showing the quantile location of the measured flow value within the model ensemble. The simulation parameters are given in Table 2: parameters were kept constant across catchments apart from correlation length, which was adjusted with catchment size. Correlation times, lengths and fractional errors were reduced for the Whirinaki where ensemble spread was otherwise too great. Lag time was set at 12 h, which our experiments showed provided good results for a wide range of catchment sizes, after comparison of lag times from 0–24 h (not shown). The exception was in a catchment with pumice soils and hence a very damped response (not included in the seven catchments used here), where a lag time of 24 h improved filter behaviour. This suggests that
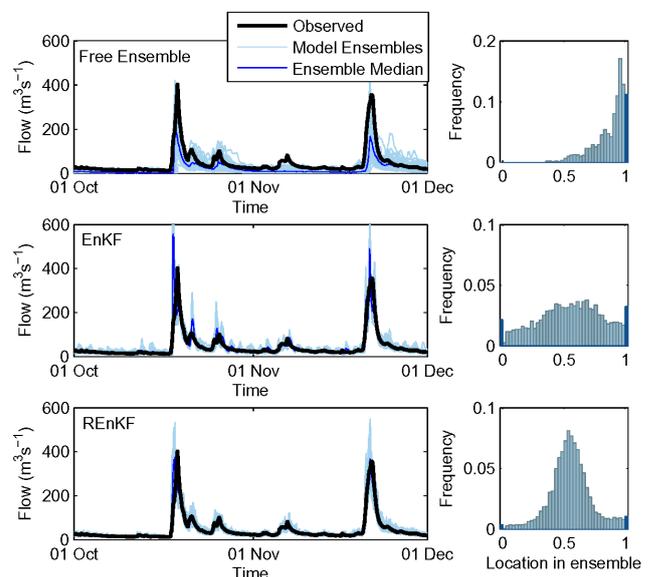
Table 2. REnKF model parameters used for each catchment.

Catchment	Decorrelation times (h)			Decorrelation lengths (km)			Fractional error			Lag (h)
	Precip	Soil	Baseflow	Precip	Soil	Baseflow	Precip	Soil	Baseflow	
Waihua	24	120	0.2	10	10	10	0.2	0.4	0.2	12
Whirinaki	1	24	0.2	2	1	1	0.2	0.4	0.05	12
Motueka	24	120	0.2	50	50	50	0.2	0.4	0.2	12
Grey	24	120	0.2	65	65	65	0.2	0.4	0.2	12
Taramakau	24	120	0.2	30	30	30	0.2	0.4	0.2	12
Whataroa	24	120	0.2	25	25	25	0.2	0.4	0.2	12
Pomahaka	24	120	0.2	75	75	75	0.2	0.4	0.2	12
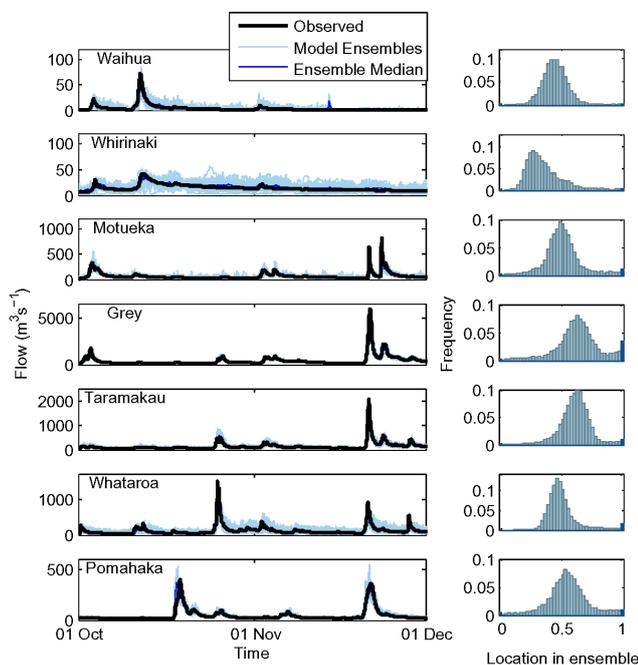
Fig. 6. REnKF simulations for all catchments, showing (left panels) ensemble simulations for a 2 month period 1 October 2011–1 December 2011 (right panels) rank histograms of location of measured flow within the model ensemble, calculated for the year 1 April 2011–31 March 2012. Dark bars at locations 0/1 show observations outside the model ensemble.

Fig. 7. Simulations for the Pomahaka catchments showing (left panels) ensemble simulations for a 2-month period 1 October 2011–1 December 2011, (right panels) rank histograms of location of measured flow within the model ensemble, calculated for the year 1 April 2011–31 March 2012. Dark bars at locations 0/1 show observations outside the model ensemble. The rows represent (upper panels) free running ensemble with no assimilation, (middle panels) EnKF assimilation, (lower panels) REnKF assimilation.

optimal lag time might be controlled by catchment soil and geological characteristics rather than size.

The results of the simulations in terms of (1) Nash–Sutcliffe (NS) score of ensemble median and (2) percentage of time the flow measurement lies within the ensemble bounds are given in Tables 3 and 4 respectively. A graphical example of the differences between the free running ensemble, EnKF and REnKF, and their corresponding rank histograms, is given for the Pomahaka catchment in Fig. 7. The first column ("single run") of Table 3 shows the NS score of each model when run deterministically. The models are running in validation mode, i.e. outside of the time period

used for calibration, and the performances vary considerably, providing a wide and challenging range of test conditions for the REnKF. Table 3 also shows that the NS scores for the "ensemble" run (i.e. without assimilation) can be considerably worse than the underlying "single run" model. This is the result of drift in the free-running ensemble. We discuss this point and its implications later in this section.

There are several points to note resulting from these simulations. The performance measures in Tables 3 and 4 show that the REnKF provides consistently good performance across all catchments. This is true even when the original model had poor NS performance. The consistency of the REnKF is in contrast to the EnKF assimilation without time
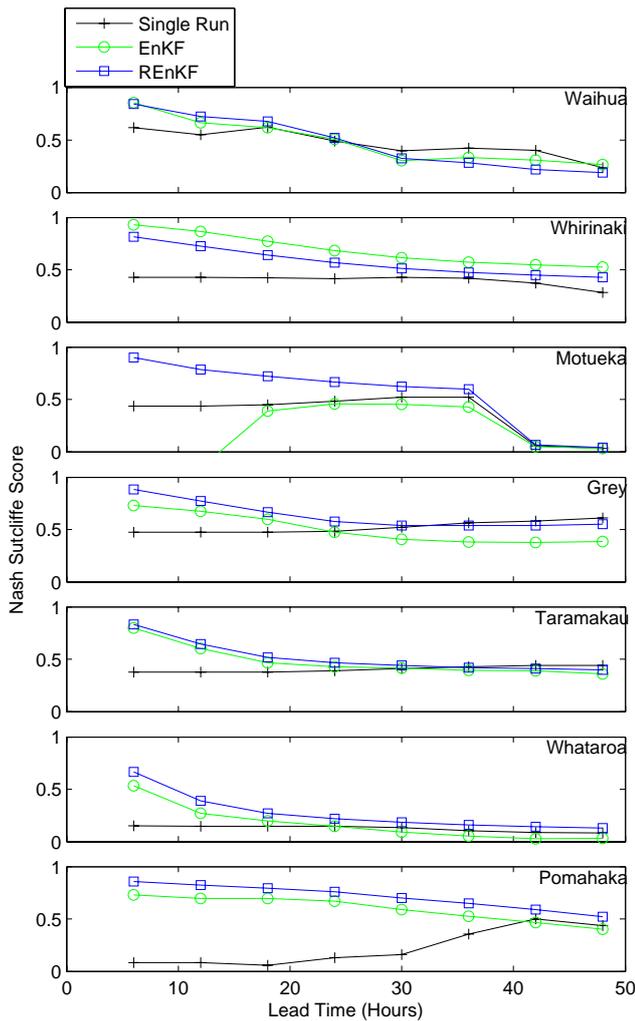
Fig. 8. Nash–Sutcliffe scores by catchment for single model run (no assimilation), EnKF and REnKF, for lead times 6–48 h, calculated for time period 1 April 2011–31 March 2012.

Table 3. Nash–Sutcliffe score for single model run and ensemble median (free-running, EnKF and REnKF), calculated for time period 1 April 2011–31 March 2012.

Catchment	Single run	Free ensemble median	EnKF median	REnKF median
Waihua	0.6167	0.5846	0.8973	0.8251
Whirinaki	0.4277	0.0051	0.9338	0.8307
Motueka	0.4379	0.2786	−1.1640	0.8982
Grey	0.4747	−0.5613	0.7302	0.8828
Taramakau	0.3759	−1.1020	0.7970	0.8305
Whataroa	0.1499	−0.1190	0.5353	0.6649
Pomahaka	0.0836	0.2116	0.7298	0.8583

Table 4. Percentage of time flow measurement lies within ensemble bounds, calculated for time period 1 April 2011–31 March 2012.

Catchment	Ensemble	EnKF	REnKF
Waihua	0.9993	0.9961	0.9943
Whirinaki	0.9933	0.9999	0.9965
Motueka	0.9928	0.9450	0.9857
Grey	0.5295	0.8969	0.9628
Taramakau	0.3876	0.9714	0.9902
Whataroa	0.9869	0.9161	0.9826
Pomahaka	0.8878	0.9462	0.9857

(Nash–Sutcliffe score) by catchment, as a function of lead time, using the same forecast period as previously used. Typical results are for the REnKF to outperform both the EnKF and no-assimilation run, up to a lead time of 24 h. After 24 h, accumulating model errors begin to overwhelm the REnKF advantage of improved initial conditions, and the performances tend to converge: at these longer lead times, the best performing methods are REnKF, three catchments; no-assimilation, three catchments; EnKF, one catchment. As with the 0–6 h lead time, we find that the REnKF method is the most reliable across different catchments and different lead times.

We also tested the model performances using the persistence index (Kitanidis and Bras, 1980). Similar to the Nash–Sutcliffe, the persistence index compares model performance against a benchmark, where a score of 1 represents perfection; 0 is no better than the benchmark; scores less than 0 are worse than the benchmark. In this case, the benchmark is the last flow observation. Results from the REnKF, EnKF and no-assimilation forecasts at all lead times are given in Fig. 9. For very short lead times (e.g. 6, 12 h), scores can be negative. This reflects the nature of data assimilation, which produces a prediction that is a compromise between observed and model prediction flow values, based on their relative errors. At longer lead times, scores are positive. As with the Nash–Sutcliffe performance measure, the
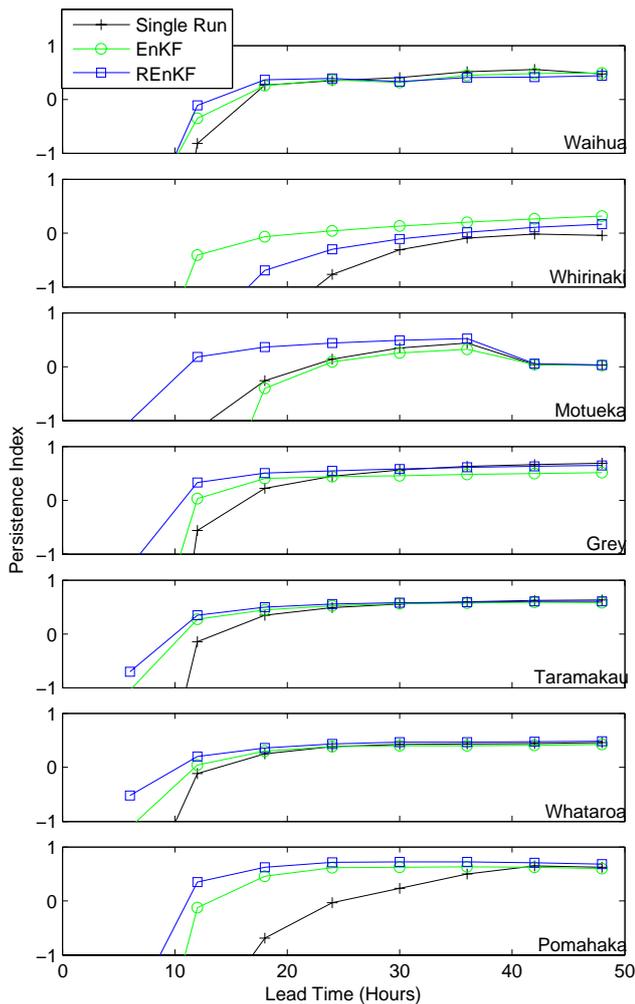
lag, which in two out of seven cases gives a higher NS score than the REnKF, but in other cases is significantly worse. The reasons are illustrated in Fig. 7, which shows that the EnKF can cause spikes and instabilities in the flow ensemble. We will return to the reason and nature of these instabilities in Sect. 4.4. Hence, although Fig. 7 shows that the rank histogram is flatter for the EnKF than the REnKF (which is typical of all catchments: not shown), there is a higher percentage of time where the observations are outside the bounds of the EnKF ensemble (refer to the dark outer bars of the rank histograms in Fig. 7, and numerically for all catchments in Table 4).

We tested the performance of each simulation for a range of lead times. The lead times used were from 0–6 to 42–48 h, in 6-h intervals. Additional mitigation measures may be possible if reliable operational forecasts are available at extended lead times. Figure 8 shows model performance

Fig. 9. Persistence index scores by catchment for single model run (no assimilation), EnKF and REnKF, for lead times 6–48 h, calculated for time period 1 April 2011–31 March 2012. Graphs are truncated at a score of −1 to improve readability.

REnKF typically outperforms the EnKF and no-assimilation run.

    The spread of the ensemble clearly varies between catchments. Empirically, we found a narrower spread relative to flow magnitude was typical of larger catchments, even when using similar assimilation parameters (Table 2). This may be caused by the cumulative effect of perturbations in many subcatchments tending to be smaller than those in one subcatchment, as individual perturbations cancel each other out, even where spatially correlated perturbations are used. We also note a trade-off in ensemble spread, where a wide spread at low flow values is required in order that the high flow spread is sufficiently wide to enable the filter to correct for the large model errors that can occur. This behaviour might be mitigated by increasing the dependence of the perturbation size on the flow magnitude in the REnKF filter design. Because our system is designed mainly for flood warning

purposes, we chose assimilation parameters that gave a good ensemble spread during flood events. When a sufficient ensemble spread is achieved, our results showed that this can degrade model performance under free-running conditions. For example, this is shown as bias in the rank histogram for the free-running ensemble for the Pomahaka (Fig. 7, upper row). This situation could occur where there are gaps in the observations available for model updates. Therefore the reliability of telemetered observations should be considered when setting ensemble spread. The degradation of performance is caused by asymmetries, thresholds or nonlinearities in the model, which can cause ensemble drift, such as the lower bound of zero flow, or nonlinearity of response when the water table reaches the surface. We return to these points in Sects. 4.2 and 4.3, where the model sensitivity and update behaviour are discussed in more detail.

4.2 Sensitivity to error parameters

During setup of our operational flow forecasting system, we carried out sensitivity experiments to determine the role and preferred values for each parameter of the REnKF system. Examples of our results are shown here to demonstrate to the reader the significance of each parameter, and to provide guidance for future implementations of similar systems.
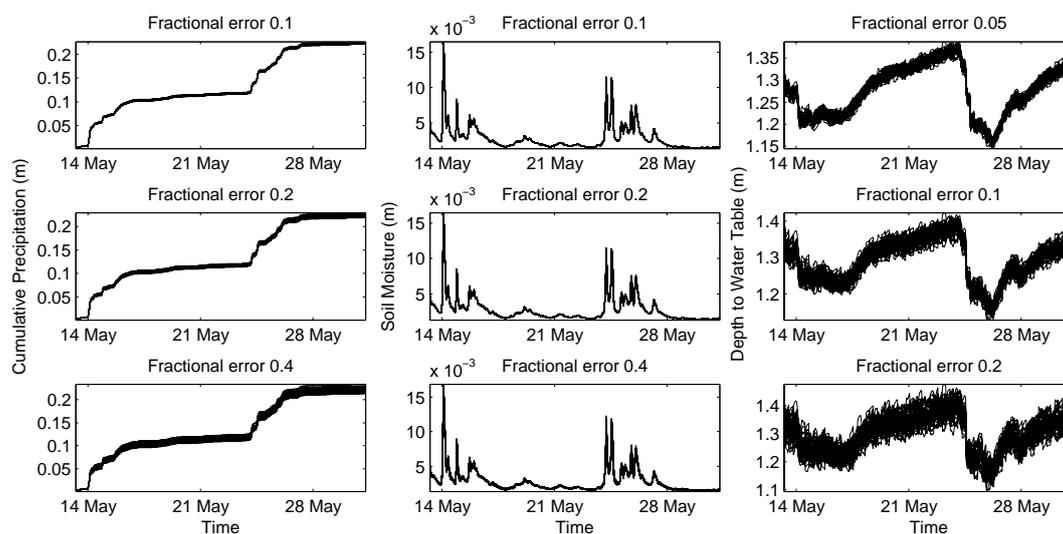
    The fractional error parameters ε_{p} (precipitation), ε_{s} (soil moisture) and ε_z (depth to water table) control the spread of the model ensemble in terms of state variables and hence the ensemble spread in flow predictions (refer to Eqs. 2a–c). An example for the Grey catchment is shown in Fig. 10.

    Figure 10 shows that the change in ensemble spread is greatest for depth to water table, and smaller for precipitation and soil moisture. These results are mirrored in the effect on the flow ensemble spread. Figure 11 shows an example where the model without assimilation (dotted line) underpredicts the observed flow (solid line). Changing the soil moisture fractional error parameter ε_{s} (left column) has little effect on the ensemble spread (grey lines) and the model ability to correct the simulated flow values. Similarly, changing the precipitation parameter ε_{p} has little effect (not shown). In contrast, changing the depth to water table parameter ε_z (right column) has a large effect on the ensemble spread and consequently on the model predictive ability.

    More generally, we tested the effect of the fractional error parameters on ensemble spread and model performance over a time period of 6 months for three selected catchments of varying size: Waihua, Grey and Pomahaka (Table 5). The results confirm that fractional error parameters for precipitation and soil moisture have negligible effect on ensemble spread or model performance. Water table fractional error has a substantial effect on ensemble spread: increasing the error parameter by a factor of 4 increases the spread by a factor in the range 2.5–3.5. The consequent effect on model performance differs by catchment: for the Grey, the Nash–Sutcliffe score increases with increased ensemble spread; for

Table 5. The effect of fractional error parameters on ensemble spread and Nash–Sutcliffe performance score under the REnKF, for selected catchments during a 6 month simulation period 31 March 2011–31 September 2011.

	Catchment	Precipitation fractional error			Soil moisture fractional error			Water table fractional error		
		0.1	0.2	0.4	0.1	0.2	0.4	0.05	0.1	0.2
Ensemble spread	Waihua	8.1	8.4	8.2	8.1	8.2	8.4	2.4	4.3	8.4
$(m^3\,s^{-1})$	Grey	182.4	182.3	182.7	181.5	181.6	182.3	73.5	110.3	182.3
	Pomahaka	30.9	31.0	31.1	30.4	30.5	31.0	12.4	18.3	31.0
Nash–Sutcliffe	Waihua	0.88	0.87	0.87	0.87	0.87	0.87	0.88	0.90	0.87
score	Grey	0.91	0.91	0.91	0.91	0.91	0.91	0.85	0.89	0.91
	Pomahaka	0.91	0.91	0.91	0.90	0.91	0.91	0.90	0.91	0.91

Fig. 10. Ensemble spread during an example REnKF model run for (left panels) cumulative precipitation, (middle panels) soil moisture and (right panels) depth to water table, with increasing fractional error parameter. Series shown for the Grey catchment, average values over all subcatchments.

the Waihua and Pomahaka, little change is observed. We suggest that these differences reflect the success of the ensemble spread in approximating model error: where the ensemble spread was not previously sufficient to capture all observed values (Grey; also refer to the rank histograms in Fig. 6), an increased spread improves the performance. Where the ensemble spread was already sufficient (Waihua, Pomahaka), an increased spread has little effect on the performance, or may begin to degrade performance as the spread is further increased.

Weerts and El Serafy (2006) note that specifying the error model is the most difficult part of applying a Kalman or particle filter. Moradkhani et al. (2005) used a fractional error of 0.1 for all state variables but found that the ensemble spread was relatively insensitive to this value, compared to the observed value perturbation. Clark et al. (2008) used fractional errors of 0.2 (rainfall), 0.1 (soil moisture) and 0.05 (depth to water table), noting that it is difficult to attribute total model error to individual sources based

only on streamflow observations. Data assimilation applications favour over-estimation of the ensemble spread, to avoid excessive reliance on model predictions over observed values (Crow and van Loon, 2006). This effect can be especially severe during recession periods where model predictions converge.

Our results showed that, using the EnKF setup of Clark et al. (2008), the fractional error of depth to water table is the most sensitive parameter. We hypothesize that this is due to the different methods used to perturb the model states: for precipitation and soil moisture the flux is perturbed, whereas depth to the water table is modified by perturbing the baseflow and then inverting the baseflow equation. The latter approach changes the store volume as well as the flux. An additional investigation (not shown) demonstrated that if we forced a larger soil moisture ensemble spread, model flow predictions worsened. A possible reason is that empirical correlations between model soil moisture and flow do not indicate a causal relationship. Instead, they could have a common
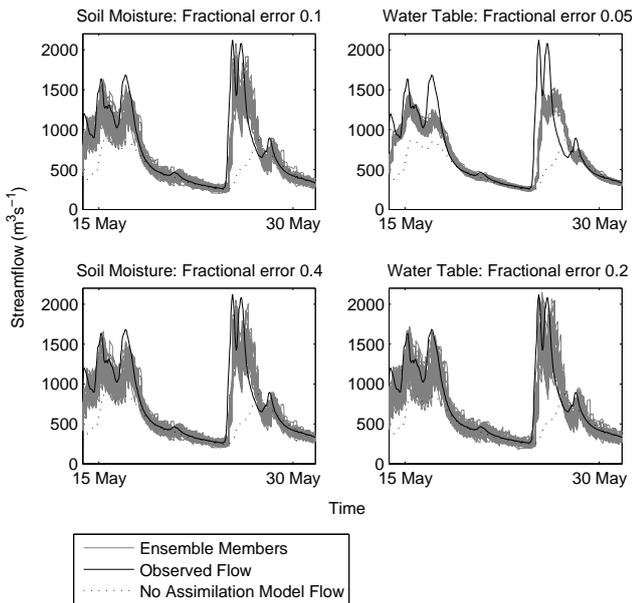
Fig. 11. Streamflow ensemble spread during an example REnKF model run showing a comparison of fractional error parameters for (left panels) soil moisture and (right panels) depth to water table. Series shown for the Grey catchment at the gauging location. Ensemble members are shown in grey, observed streamflow in black, modelled streamflow without assimilation as dotted line.

Fig. 12. Size of REnKF update to model state "depth to water table" for the Pomahaka catchment, by observation time and lag time. Updates are shown for the gauging station reach (upper panel) and a selected headwater reach (middle panel). Observed and modelled flows at the observation times are shown for comparison (lower panel).

cause (rainfall depth), while elevated flows are controlled by the water table depth via saturation excess flow. Crow and Van Loon (2006) similarly found that assimilation of surface soil moisture observations to constrain deeper root-zone soil moisture in a land surface model could degrade model performance when model error assumptions were not appropriate.

4.3 Update behaviour

We selected some large rainfall events to investigate the update behaviour of the filter, i.e. where and when updates were made to model states. It is difficult to visualise filter update behaviour since it is high dimensional: two spatial dimensions (subcatchment location), observation time, time lag and update size. Each of these applies to each model state to be updated. For clarification, the update at observation time T_O, lag time T_L, refers to the change made to the model state at time $T_O - T_L$ as a result of the information gained from the observation at time T_O.

By plotting the observation time against lag time T_L, and colouring according to update size, we can examine (1) which flow observations lead to the greatest model state updates and (2) at what lag times these updates occur. Figure 12 shows an example for the Pomahaka catchment, for updates to the water table depth in two catchments: one at the gauging site and one in the headwaters. From this figure we see that the gauging catchment undergoes greater updates than the headwater catchment. This can be expected since
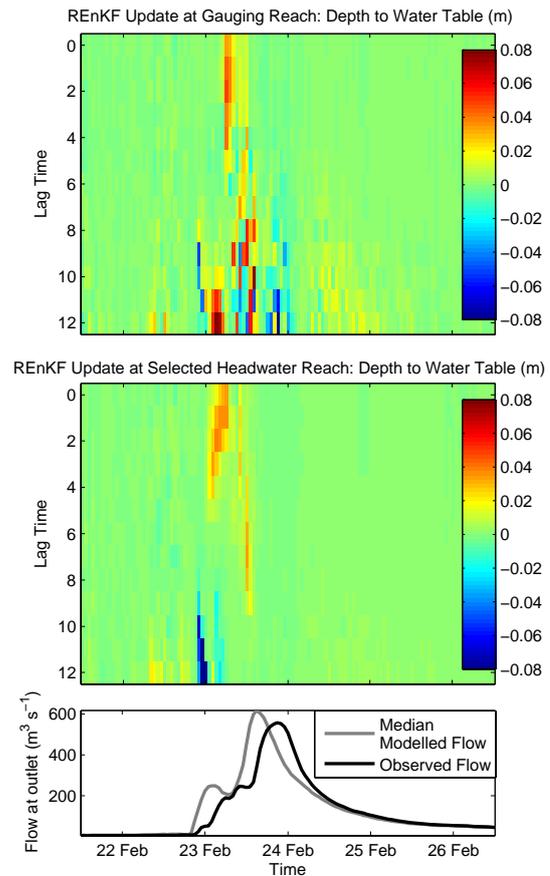
correlations between the water table depth in a subcatchment and flow at the gauge will be greater where the locations of these two measurements are closer. We also see that larger updates are often made at the maximum time lag allowed by the filter (12 h in this case). This reflects the design of the filter whereby the "first pass" (at the maximum lag) removes gross model errors, and then the model simulation is progressively fine-tuned at subsequent (smaller) lag times.

The dependence of model update size on lag time was compared in the general case for the Pomahaka catchment (Fig. 13), by taking an average over all observation times and all subcatchments (absolute update is used so that positive and negative updates do not cancel each other). This confirms the pattern whereby update size increases with lag time. However, the updates to depth to water table and routed flow storage are relatively greater at shorter lag times than the updates to soil moisture storage. A comparison can also be made of the size of updates between the different model states. The size of the soil moisture update is up to two orders
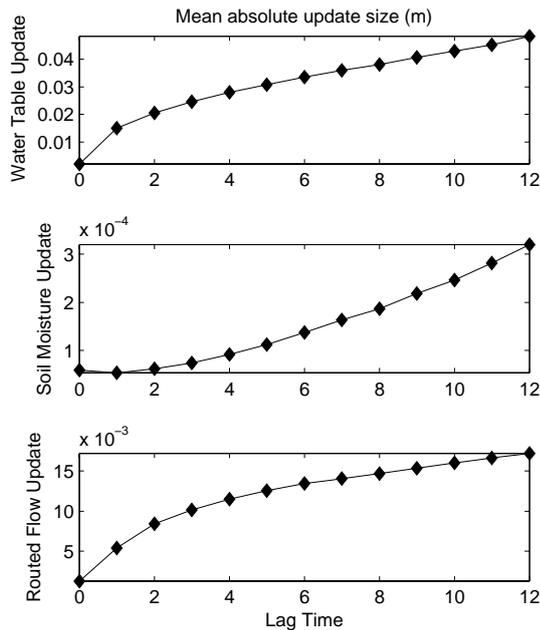
Fig. 13. Comparison of size of REnKF update by model state and lag time for the Pomahaka catchment, during the flood event shown in Fig. 8. Update values shown are the mean value over time and subcatchments.

of magnitude smaller than the updates to routed flow or water table storages. This reflects both smaller correlations between soil moisture and streamflow, and smaller ensemble spread. However, as noted in the section "Sensitivity to error parameters", increasing the ensemble spread and hence the update size tended to degrade the model forecast. The updates to the water table are greatest, and approximately three times the magnitude of the updates to routed flow. This behaviour may be partly dependent on the hydrological model structure; for example in this case the TOPMODEL formulation means that water table depth controls saturation excess flow, which is an important contribution to storm flow volumes.

To understand spatial patterns of updates, we also plotted the size of update in each subcatchment. Figure 14 shows an example, again from the Pomahaka, just prior to the flood event illustrated in Fig. 12. The figure shows the updates for different lag times (i.e. different time before observation), relating to the same observation time of 06:00 a.m. LT, 23 February 2012. Updates are typically largest at the largest lag time, but updates are made throughout the catchment at all lag times. At the shortest lag times, updates tend to be concentrated along the main stem of the river network (rivers of Strahler order 4 and above are shown on the figure). This type of spatial organisation is consistent with our expectation that only subcatchments able to quickly affect flow magnitude would be updated at short lag times.

4.4 Numerical artefacts under the EnKF

The flow forecasting results above showed that one reason for improved performance of the REnKF over the EnKF was due to artefacts in the model behaviour under the EnKF. These could take the form of "spikes" in the forecast flow (Fig. 15a), or oscillatory behaviour (Fig. 15b). Spike-type artefacts were found by Mendoza et al. (2012), using a similar implementation of the TopNet model with EnKF data assimilation for a Chilean basin.

Our investigation has led us to hypothesise that the same underlying mechanism is responsible for both such artefacts. As described in the introduction, each catchment has a natural lag time between rainfall and the associated rise in river flow. This means that subcatchments distant from the gauging site may have model state variables (soil moisture, depth to water table, surface storage) that are not strongly correlated with the flow measured at the gauge at the same timestep. At times the correlation can become negative, even strongly negative, according to the random perturbations applied. Hence, for example, a model that predicts too high a flow may try to correct that error during the filtering step by adding, rather than subtracting, water from the groundwater store in the distant subcatchments. The additional water causes a spike flow response in subsequent timesteps. This is the situation shown in Fig. 15a. In less severe situations, the correlation may be positive, but not representative of the stronger connection between current model state in remote subcatchments and future flow at the gauge. In this situation, oscillatory behaviour (Fig. 15b) could be produced before the system settles to the observed flow value. The oscillating updates to the model state "depth to water table" that cause this behaviour are shown in Fig. 16. In both cases, the distortion occurs because the observation data overwhelm the update time step, and drive the update towards itself. Both these types of behaviour are resolved by the REnKF, which both allows for a time lag between cause (change in water storage) and effect (change in flow), and essentially "tests" the Kalman filter update by running the model forward in time to recalculate the simulated flow.

5 Conclusions

In this paper we described the design of a recursive ensemble Kalman filter (REnKF), which iteratively updates model states prior to an observation. We then demonstrated how the REnKF was implemented for an operational flow forecasting system. We described the filter design decisions facing the user, and described the filter parameters and their effect on the filter behaviour. The performance of the system was compared with a deterministic model, with a free-running ensemble and with the standard EnKF, which does not account for lag times between model states and streamflow at the catchment outlet.
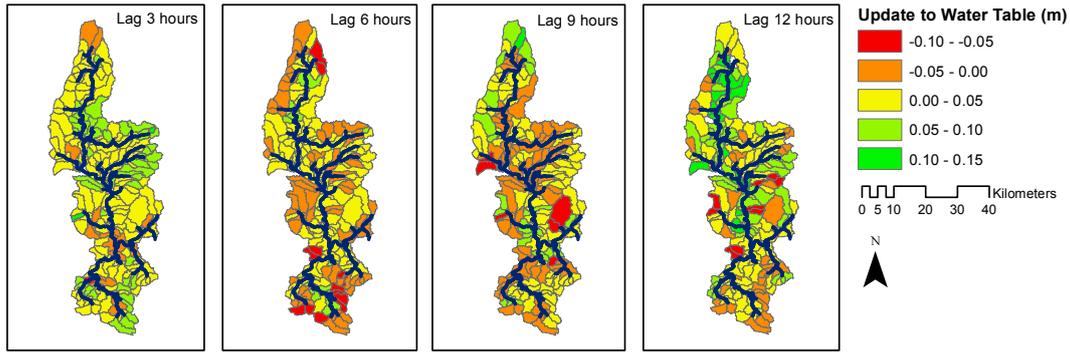
Fig. 14. Size of REnKF update to model state "depth to water table" for the Pomahaka catchment, by subcatchment and lag time, at observation time 23 February 2012, 06:00:00 LT.
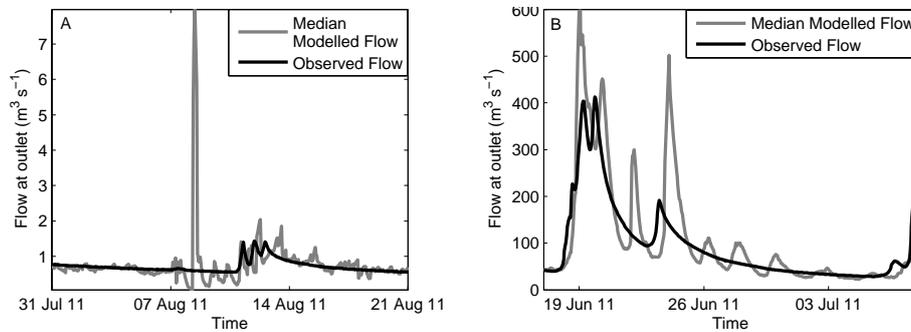

Fig. 15. (A) Spike in modelled flow for the Waihua River when using the ensemble Kalman filter. **(B)** Oscillations in modelled flow for the Motueka River when using the ensemble Kalman filter.
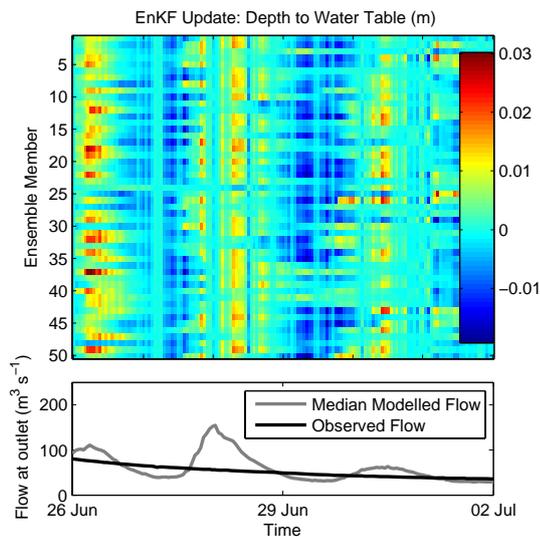

Fig. 16. Updates to depth to water table for each ensemble member for a sample reach in the Motueka. Modelled and measured flow are shown for comparison. Lags between updates and flow response cause oscillations during this period.

Following our investigation, we can make several comments and recommendations for future users of the REnKF or similar systems in a hydrological context. For the filtering steps we used the implementation of the EnKF described by Clark et al. (2008), which perturbs model precipitation, soil moisture and depth to water table. We found that, of these, the error parameters controlling perturbations of depth to water table were the most sensitive in terms of model ensemble spread and performance. Similarly, the filter updates the soil moisture storage, depth to water table and surface (routing) storage, and the updates to the depth to water table were largest. Updates to the surface storage were within the same order of magnitude. In contrast, soil moisture storage updates were small, and experiments to increase the ensemble spread and update size were shown to degrade model performance. This result suggests that a future implementation of the REnKF could reasonably remove the soil moisture storage from the model states to be perturbed or updated. This would increase the efficiency of the filter, and is unlikely to reduce forecast quality.

Our results showed that, in cases where large model errors could occur (including due to error in forecast rainfall), then a large ensemble spread in modelled streamflow was required for good REnKF performance. This could be achieved most directly by increasing the fractional error parameter or

correlation length of the depth to water table perturbations. The decorrelation time was less important since this time is typically longer than the time between model updates. However, the need for a large ensemble spread during high flow periods had to be balanced against creating too wide a spread at low flows, and a possible deterioration of performance where the free-running ensemble was used.

In all, we found a significant improvement in model forecasts when using the REnKF, which was able to overcome the instabilities of filter behaviour found with the EnKF. Across seven diverse catchments, the REnKF scored consistently highly against the performance measures of NS score for the ensemble median, and percentage time that the observed flow fell within the ensemble bounds. We would therefore recommend the REnKF concepts to other research groups wishing to use telemetered streamflow data for assimilation into a hydrological model.

Appendix A

REnKF algorithm

The main features of our implementation of the REnKF algorithm involve storing prior forcing data and model states for later retrieval during the assimilation process. The algorithm also writes model states and associated fluxes to output files after they have been altered for the last time, as the model steps forward in time past the point at which those steps could be further altered by the recursive assimilation scheme.

An important parameter of the REnKF is the lag time, N (refer to Sect. 2.2.5). This parameter controls the maximum timestep difference between a streamflow observation and a prior model state that can be updated during the assimilation of that observation. To describe the REnKF process in more detail, assume that a simulation has reached a timestep t (t refers to the end of the timestep interval). At this timestep, the forcing data for the previous N timesteps have been stored in temporary memory as well as model states at timestep $t - N$. The recursive assimilation process for this timestep, t, is as follows:

1. save forcing data for current timestep in temporary memory;

2. load model states at time step $t - N$ from temporary memory;

3. run model from time $t - N$ to t to get the model estimate of the streamflow observations at time t;

4. set a recursive timestep counter j equal to zero, used in the following steps;

5. load model states at timestep $t - (N - j)$;

6. assimilate streamflow observation at time t, using the EnKF to update model states at time $t - (N - j)$;

7. if $j = 0$, write model states to output files at time $t - N$;

8. if $j = 1$, store model states to temporary memory needed at step (ii) for next timestep;

9. run the model through one timestep from time $t - (N - j)$ to time $t - (N - j) + 1$;

10. if $j = 0$, write model fluxes for timestep $t - (N - j) + 1$ to output files;

11. store model states at time $t - (N - j) + 1$ to temporary memory to use in the next iteration at (v);

12. run the model from timestep $t - (N - j) + 1$ to t to get the model estimate of the streamflow observation at t;

13. repeat steps (v) to (xii) for $j = 1, 2, ..., N - 1$;

14. for $j = N$, repeat (vi).

REnKF assimilation for streamflow observation at timestep t is completed. Advance model from t to $t + 1$. The process is then repeated for all remaining timesteps. This implementation increases the number of model single-timestep executions by a factor of $F(N) = N(N + 3)/2$, which gives $F(12) = 90$ and $F(24) = 324$. It is therefore clear that this implementation of the REnKF is computationally intensive.

References

Andreadis, K. M. and Lettenmaier, D. P.: Assimilating remotely sensed snow observations into a macroscale hydrology model, Adv. Water Resour., 29, 872–886, doi:10.1016/j.advwatres.2005.08.004, 2006.

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., and Alsdorf, D. E.: Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a rasterbased hydrodynamics model, Geophys. Res. Lett., 34, L10403, doi:10.1029/2007gl029721, 2007.

Aubert, D., Loumagne, C., and Oudin, L.: Sequential assimilation of soil moisture and streamflow data in a conceptual rainfallrunoff model, J. Hydrol., 280, 145–161, doi:10.1016/s0022-1694(03)00229-4, 2003.

Bandaragoda, C., Tarboton, D. G., and Woods, R.: Application of TOPNET in the distributed model intercomparison project, J. Hydrol., 298, 178–201, doi:10.1016/j.jhydrol.2004.03.038, 2004.

Bárdossy, A. and Singh, S. K.: Robust estimation of hydrological model parameters, Hydrol. Earth Syst. Sci., 12, 1273–1283, doi:10.5194/hess-12-1273-2008, 2008.

Berthet, L., Andréassian, V., Perrin, C., and Javelle, P.: How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments, Hydrol. Earth Syst. Sci., 13, 819–831, doi:10.5194/hess-13-819-2009, 2009.

Beven, K. J. and Kirkby, M. J.: A physically based variable contributing area model of basin hydrology, Hydrol. Sci. Bull., 24, 43–69, 1979.

Beven, K. J., Lamb, R., Quinn, P., Romanowicz, R., and Freer, J.: TOPMODEL, in: Computer Models of Watershed Hydrology, edited by: Singh, V. P., Water Resour. Publ., Highlands Ranch, Colorado, 627–668, 1995.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010.

Clark, M. P. and Slater, A. G.: Probabilistic quantitative precipitation estimation in complex terrain, J. Hydrometeorol., 7, 3–22, 2006.

Clark, M. P., Slater, A. G., Barrett, A. P., Hay, L. E., McCabe, G. J., Rajagopalan, B., and Leavesley, G. H.: Assimilation of snow covered area information into hydrologic and land-surface models, Adv. Water Resour., 29, 1209–1221, doi:10.1016/j.advwatres.2005.10.001, 2006.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., and Uddstrom, M. J.: Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Adv. Water Resour., 31, 1309–1324, doi:10.1016/j.advwatres.2008.06.005, 2008.

Cohn, S. E., Sivakumaran, N. S., and Todling, R.: A fixed-lag Kalman Smoother for retrospective data assimilation, Mon. Weather Rev., 122, 2838–2867, doi:10.1175/1520-0493(1994)122<2838:aflksf>2.0.co;2, 1994.

Crow, W. T. and Van Loon, E.: Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, J. Hydrometeorol., 7, 421–432, doi:10.1175/jhm499.1, 2006.

De Lannoy, G. J. M., Reichle, R. H., Arsenault, K. R., Houser, P. R., Kumar, S., Verhoest, N. E. C., and Pauwels, V. R. N.: Multiscale assimilation of Advanced Microwave Scanning Radiometer-EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado, Water Resour. Res., 48, W01522, doi:10.1029/2011wr010588, 2012.

Di Baldassarre, G. and Montanari, A.: Uncertainty in river discharge observations: a quantitative analysis, Hydrol. Earth Syst. Sci., 13, 913–921, doi:10.5194/hess-13-913-2009, 2009.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics, J. Geophys. Res.-Oceans, 99, 10143–10162, doi:10.1029/94jc00572, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003.

Evensen, G. and van Leeuwen, P. J.: An ensemble Kalman smoother for nonlinear dynamics, Mon. Weather Rev., 128, 1852–1867, doi:10.1175/1520-0493(2000)128<1852:aeksfn>2.0.co;2, 2000.

Flores, A. N., Bras, R. L., and Entekhabi, D.: Hydrologic data assimilation with a hillslope-scale-resolving model and L band radar observations: Synthetic experiments with the ensemble Kalman filter, Water Resour. Res., 48, W08509, doi:10.1029/2011WR011500, 2012.

Goring, D. G.: Kinematic shocks and monoclinal waves in the Waimakariri, a steep, braided, gravel-bed river, in: Proceedings of the International Symposium on waves: Physical and numerical modelling, University of British Columbia, Vancouver, Canada, 336–345, 1994.

Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796–811, 1998.

Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Weather Rev., 129, 123–137, doi:10.1175/1520-0493(2001)129<0123:asekff>2.0.co;2, 2001.

Ibbitt, R. P. and Woods, R.: Towards rainfall-runoff models that do not need calibration to flow data, in: Friend 2002 – Regional Hydrology: Bridging the Gap Between Research and Practice, edited by: van Lanen, H. A. J. and Demuth, S., IAHS Publ., 274, 189–196, 2002.

Kitanidis, P. K. and Bras, R. L.: Real-time forecasting with a conceptual hydrologic model, 2. Application and results, Water Resour. Res., 16, 1034–1044, 1980.

Komma, J., Bloschl, G., and Reszler, C.: Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting, J. Hydrol., 357, 228–242, doi:10.1016/j.jhydrol.2008.05.020, 2008.

Lee, H., Seo, D. J., and Koren, V.: Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states, Adv. Water Resour., 34, 1597–1615, doi:10.1016/j.advwatres.2011.08.012, 2011.

Li, Y., Ryu, D., Wang, Q. J., Pagano, T., Western, A., Hapuarachchi, P., and Toscas, P.: Assimilation of streamflow discharge into a continuous flood forecasting model, in: Risk in Water Resources Management, edited by: Bloschl, G., Takeuchi, K., Jain, S., Farnleitner, A., and Schumann, A., IAHS Publ., 107–113, 2011.

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006wr005756, 2007.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, Hydrol. Earth Syst. Sci., 16, 3863–3887, doi:10.5194/hess-16-3863-2012, 2012.

Lorenc, A. C., Ballard, S. P., Bell, R. S., Ingleby, N. B., Andrews, P. L. F., Barker, D. M., Bray, J. R., Clayton, A. M., Dalby, T., Li, D., Payne, T. J., and Saunders, F. W.: The Met. Office global three-dimensional variational data assimilation scheme, Q. J. Roy. Meteorol. Soc., 126, 2991–3012, doi:10.1256/smsqj.57001, 2000.

McLaughlin, D.: An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering, Adv. Water Resour., 25, 1275–1286, doi:10.1016/s0309-1708(02)00055-6, 2002.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., and Clark, M.: Impacts of uncertain flow data on rainfall-runoff model calibration and discharge predictions, Hydrol. Process., 24, 1270–1284, doi:10.1002/hyp.7587, 2010.

McMillan, H., Jackson, B., Clark, M., Kavetski, D., and Woods, R.: Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models, J. Hydrol., 400, 83–94, doi:10.1016/j.jhydrol.2011.01.026, 2011.

Mendoza, P. A., McPhee, J., and Vargas, X.: Uncertainty in flood forecasting: A distributed modeling approach in a sparse data catchment, Water Resour. Res., 48, W09532, doi:10.1029/2011wr011089, 2012.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28, 135–147, doi:10.1016/j.advwatres.2004.09.002, 2005.

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., and Pappenberger, F.: A data assimilation approach to discharge estimation from space, Hydrol. Process., 23, 3641–3649, doi:10.1002/hyp.7518, 2009.

Newsome, P. F. J., Wilde, R. H., and Willoughby, E. J.: Land Resource Information System Spatial Data Layers, Technical Report, Palmerston North, Landcare Research NZ Ltd., New Zealand, 2000.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S.: Applying sequential Monte Carlo methods into a distributed hydrologic model: lagged particle filtering approach with regularization, Hydrol. Earth Syst. Sci., 15, 3237–3251, doi:10.5194/hess-15-3237-2011, 2011.

Pauwels, V. R. N. and De Lannoy, G. J. M.: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, J. Hydrometeorol., 7, 458–477, doi:10.1175/jhm490.1, 2006.

Pauwels, V. R. N. and De Lannoy, G. J. M.: Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space, Water Resour. Res., 45, W08428, doi:10.1029/2008wr007590, 2009.

Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435–3449, doi:10.5194/hess-16-3435-2012, 2012.

Randrianasolo, A., Ramos, M. H., Thirel, G., Andreassian, V., and Martin, E.: Comparing the scores of hydrological ensemble forecasts issued by two different hydrological models, Atmos. Sci. Lett., 11, 100–107, doi:10.1002/asl.259, 2010.

Ravela, S. and McLaughlin, D.: Fast ensemble smoothing, Ocean Dynam., 57, 123–134, doi:10.1007/s10236-006-0098-6, 2007.

Reichle, R. H.: Data assimilation methods in the Earth sciences, Adv. Water Res., 31, 1411–1418, doi:10.1016/j.advwatres.2008.01.001, 2008.

Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, J. Hydrometeorol., 3, 728–740, doi:10.1175/1525-7541(2002)003<0728:evekff>2.0.co;2, 2002.

Ricci, S., Piacentini, A., Thual, O., Le Pape, E., and Jonville, G.: Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting, Hydrol. Earth Syst. Sci., 15, 3555–3575, doi:10.5194/hess-15-3555-2011, 2011.

Seo, D. J., Koren, V., and Cajina, N.: Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting, J. Hydrometeorol., 4, 627–641, doi:10.1175/1525-7541(2003)004<0627:rvaoha>2.0.co;2, 2003.

Seo, D. J., Cajina, L., Corby, R., and Howieson, T.: Automatic state updating for operational streamflow forecasting via variational data assimilation, J. Hydrol., 367, 255–275, doi:10.1016/j.jhydrol.2009.01.019, 2009.

Slater, A. G. and Clark, M. P.: Snow data assimilation via an ensemble Kalman filter, J. Hydrometeorol., 7, 478–493, doi:10.1175/jhm505.1, 2006.

Snelder, T. H. and Biggs, B. J. F.: Multiscale River Environment Classification for water resources management, J. Am. Water Resour. Assoc., 38, 1225–1239, doi:10.1111/j.1752-1688.2002.tb04344.x, 2002.

Sorensen, J. V. T. and Madsen, H.: Efficient Kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the North Sea and Baltic Sea system, J. Geophys. Res.-Oceans, 109, C03017, doi:10.1029/2003jc002144, 2004.

Tait, A., Henderson, R., Turner, R., and Zheng, X.: Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface, Int. J. Climatol., 26, 2097–2115, doi:10.1002/joc.1350, 2006.

Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S., and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France – Part 1: Description and validation of the assimilation system, Hydrol. Earth Syst. Sci., 14, 1623–1637, doi:10.5194/hess-14-1623-2010, 2010a.

Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S., Regimbeau, F., and Habets, F.: A past discharge assimilation system for ensemble streamflow forecasts over France – Part 2: Impact on the ensemble streamflow forecasts, Hydrol. Earth Syst. Sci., 14, 1639–1653, doi:10.5194/hess-14-1639-2010, 2010b.

Uddstrom, M.: Environmental prediction and innovation, New Zealand Sci. Rev., 68, 37–41, 2011.

Weerts, A. H. and El Serafy, G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403, doi:10.1029/2005wr004093, 2006.

Weerts, A. H., El Serafy, G. Y., Hummel, S., Dhondia, J., and Gerritsen, H.: Application of generic data assimilation tools (DATools) for flood forecasting purposes, Comput. Geosci., 36, 453–463, doi:10.1016/j.cageo.2009.07.009, 2010.

Westerberg, I., Guerrero, J. L., Seibert, J., Beven, K. J., and Halldin, S.: Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras, Hydrol. Process., 25, 603–613, doi:10.1002/hyp.7848, 2011.

Woods, R., Hendrikx, J., Henderson, R. D., and Tait, A.: Estimating mean flow of New Zealand rivers, J. Hydrol. (NZ), 45, 95–100, 2006.

Zaitchik, B. F. and Rodell, M.: Forward-Looking Assimilation of MODIS-Derived Snow-Covered Area into a Land Surface Model, J. Hydrometeorol., 10, 130–148, doi:10.1175/2008jhm1042.1, 2009.

Zhou, H. Y., Gomez-Hernandez, J. J., Franssen, H. J. H., and Li, L. P.: An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering, Adv. Water Resour., 34, 844–864, doi:10.1016/j.advwatres.2011.04.014, 2011.