18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Evaluation of alternative model-data fusion approaches for retrospective water balance estimation

Van Dijk, A.I.J.M. [1], and L.J. Renzullo [1]

[1] *CSIRO Land and Water/Water for a Healthy Country Flagship, CSIRO-Bureau of Meteorology Water Information Research and Development Alliance*
Email: albert.vandijk@csiro.au

Water resources observation and prediction systems are being developed in the Australian Bureau of Meteorology to produce water information services, and will include rolling water balance estimation. A prototype Australian Water Resources Assessment Model (AWRAM) has been developed, and the nation-wide coverage, currency, accuracy, and consistency required means that remote sensing plays an important role. This paper tests and discusses alternative methods of blending models and observations. Integration of on-ground and remote sensing data into land surface models typically involves state updating through model-data assimilation techniques. By comparison, retrospective water balance estimation and hydrological scenario modelling to date has mostly relied on non-sequential parameter estimation against stream flow observations, and has made little use of satellite earth observation. The most appropriate model-data fusion approach for a continental water balance estimation system will need to consider the trade-off between accuracy gains when using more sophisticated synthesis techniques and additional observations, and the computational overheads this incurs. This trade-off was investigated using relatively simple but well-performing lumped models of seasonal vegetation dynamics and catchment hydrology that are implemented in the prototype AWRAM, while formal inter-comparison experiments to assess alternative component model paradigms and structures are underway.

The performance of different model-data fusion (MDF) approaches was evaluated using flux tower ET measurements at four sites in Australia together with satellite observations of soil moisture over the corresponding area (AMSR-E passive microwave instrument). These observations, rather than hydrometric observations (e.g. streamflow), were chosen because of the more direct relationship they have with the site water balance over shorter time scales. Satellite-observed vegetation vigour (MODIS Enhanced Vegetation Index, EVI) was the assimilated variable. The MDF techniques tested include non-sequential estimation of model parameters (calibration against EVI, ET or both) and scaling of rainfall inputs, as well as sequential updating of leaf area index or soil moisture content using the ensemble Kalman filter. Non-sequential parameter estimation did not appear to provide much benefit compared to using prior parameter estimates, suggesting that the model parameterisation was comparatively robust and parameter values spatially invariant, at least when compared to errors in forcing data. A combination of parameter estimation and state updating did lead to improvements in some aspects of evaluation; reducing the apparent error in monthly evapotranspiration by 1% and in monthly top soil moisture content by 12%, respectively, when compared to using a priori parameter estimates. However it was also about three orders of magnitude more computationally intensive. Rainfall input adjustment was only tested in a relatively crude, non-sequential manner but results were encouraging, and appear to be a promising candidate for sequential approaches.

Keywords: *model data assimilation, remote sensing, hydrology, land surface modelling*

1. INTRODUCTION

Water resources observation and prediction systems are being developed to produce water information services as part of the Australian Bureau of Meteorology's new statutory role. It will include rolling water balance estimation to underpin national water accounts, and water resources assessments that interpret current water resources availability and trends in a historical context. A prototype Australian Water Resources Assessment Model (AWRAM) has been developed (Van Dijk *et al.*, in review). The nation-wide coverage, currency, accuracy, and consistency required means that remote sensing plays an important role, along with in-situ observations.

There are different approaches to blending models and observations (Barrett *et al.*, 2008). Some common choices include: [1] the observations that are to be blended in and its implications for the model structure (e.g. coupled water-energy balance simulation to assimilate land surface temperature observation; dynamic vegetation cover development simulation to assimilate vegetation observations); [2] estimation of model simulated states (stores, fluxes) *versus* model inputs (e.g. rainfall scaling) or model parameters; [3] sequential approaches (i.e. updating model estimates at each time step an observation is available) *versus* non-sequential approaches (i.e. single set of parameter estimates that minimise observation-model mismatches over a period); [4] the sophistication of methods used, from very simple (e.g. direct insertion or statistical correction) to more complex and computationally intensive methods requiring generation of ensembles (e.g. particle or ensemble Kalman filters), many iterations in a search for an optimum solution (global search algorithms), or calculation of the model Jacobian with respect to the target variable (e.g. variational methods). Integration of on-ground and remote sensing data into land surface models in atmospheric applications typically involves sequential assimilation approaches. Streamflow forecasting typically involves direct insertion or bias correction, although more sophisticated assimilation approaches are increasingly being used. By comparison, water balance estimation and hydrological scenario modelling to date has mostly relied on non-sequential optimisation of (time invariant) model parameters.

The blending of remote sensing data into water balance models is still in early stages of development. To determine an appropriate model-data fusion approach for continental water balance estimation, the trade-off between computational overhead and the gain in accuracy needs to be considered. This paper investigates this trade-off by trialling some alternative model-data fusion approaches, using a landscape hydrological model and satellite-based estimates of soil moisture and vegetation properties for four Australian sites.

2. MODEL

The prototype AWRAM uses relatively simple but well-performing lumped models of catchment water balance (OCCAM) and seasonal vegetation dynamics (EGG) as interim solutions (Van Dijk *et al.*, in review). Formal inter-comparison experiments to assess alternative component model paradigms and structures are underway and will lead to future improvements.

2.1. OCCAM water balance model

Van Dijk *et al.* (in review) analysed daily streamflow data for more than 284 headwater catchments (50−2000 km^2); interpolated station rainfall and climate data; and

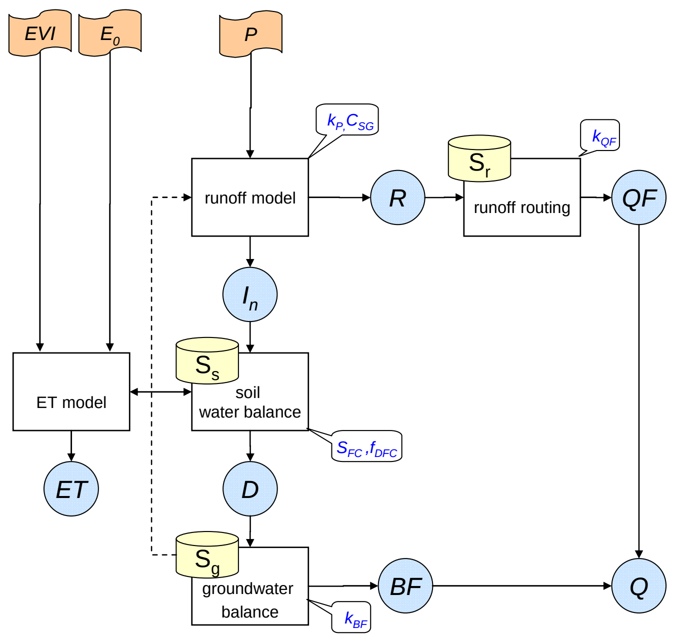


Figure 1. Illustration of the OCCAM model structure. Shown are model components (rectangles), model input time series (pink), the six free model parameters (see text). Internal fluxes and outputs (blue circles) are storm runoff (*R*), quick flow (*QF*), net infiltration (I_n), evapotranspiration (*ET*), drainage (*D*), baseflow (*BF*), total streamflow (*Q*). Storages (yellow) are soil water (S_s), groundwater (S_g) and runoff (S_r).

observations of vegetation vigour by the MODIS sensor aboard the Terra and Aqua satellites. Streamflow data were processed to obtain estimates of baseflow and storm flow (or quick flow). For each component of the water balance several candidate model structures were derived from theory and literature. Separate analyses were performed to develop model components describing baseflow dynamics, surface runoff generation, evapotranspiration (ET) and soil water drainage (Figure 1). Alternative model structures were compared to the data, and Akaike's Information Criterion (Akaike, 1970) was used to determine the structures representing the optimal trade-off between the number of free parameters and explanatory power. The composite model was named OCCAM (for 'Optimum Complexity CAtchment Model').

The model has three input data sources, all required at daily time step: the Enhanced Vegetation Index (EVI), potential ET (E_0) and precipitation (P). Excluding the ET model component, the model has six 'free' model parameters that are allowed to vary between catchments. These describe: infiltration excess runoff (parameter k_P), saturation overland flow (C_{Sg}), storm flow recession (k_{QF}), soil water storage at field capacity (S_{FC}), drainage rate at field capacity (f_{DFC}), and baseflow recession (k_{BF}). Analyses were performed to assess how accurately these parameters can be estimated based on covariance with catchment attributes and spatial correlation, and predictive regression equations were developed accordingly.

The ET estimation component has no free parameters as such; instead a single set of seven parameters was optimised across all catchment data available. The seven parameters include: two parameters are related to the contribution of the aerodynamic component of total evaporative energy; two for the relationship between EVI and surface conductance; two for the effect of atmospheric vapour demand on surface conductance; and one defines the maximum amount of near-surface water available for evaporation from the soil and wet canopy. EVI was chosen as a predictor of ET in preference to derived products (e.g. leaf area index or vegetation fractional cover); there are plant physiological arguments to assume that EVI provides a better integrated measure of canopy scale surface conductance. An example is the known relationship between leaf chlorophyll content (greenness) and stomatal density (surface conductance per unit LAI). A previously published ET estimation method using a similar approach (Guerschman *et al.*, 2009) showed a bias of plus or minus ~15% over longer periods when compared to flux tower and catchment water balance estimates of ET.

2.2. EGG dynamic vegetation model

The Equilibrium Greenness Growth (EGG) model used here simulates vegetation dynamics by calculating the EVI that could be sustained given soil moisture availability. The equilibrium EVI is determined by considering the hypothetical EVI leading to a maximum transpiration rate ($E_{t,max}$) that equals maximum root water uptake (U_{max}) under transient soil moisture conditions. The vegetation moves towards this equilibrium state with a prescribed degree of inertia. The model can include one or more land cover types, each defined by their fractional cover and properties. Currently, two land cover types are considered: deep- and shallow-rooted vegetation. The primary output is EVI although this is internally linked to leaf area index and canopy fractional cover.

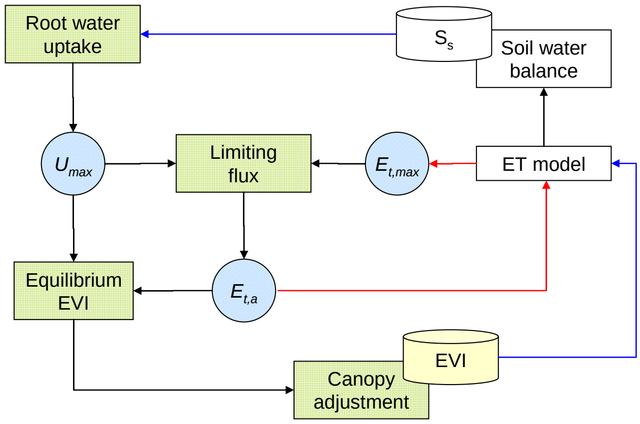


Figure 2. Illustration of the Equilibrium Greenness model. Symbols as in previous figure, with additional green elements showing the model components that are internal to the Equilibrium Greenness model. The red and blue arrows show feedbacks between the models within and between time steps, respectively.

Unlike streamflow data, EVI data to calibrate and constrain the dynamic vegetation model are available with near complete temporal and spatial coverage, the need for parsimony in parameterisation is less severe, and the number of parameters was allowed to be larger to increase flexibility in calibration. There are seven parameters for each land cover type, describing the relationship between soil water content and root water uptake (three parameters), between EVI, canopy cover, and LAI (two parameters), and describing the relative rates of canopy adjustment (two parameters).

3. DATA

The coupled OCCAM-EGG model only requires precipitation and E_0 as inputs. Interpolated station rainfall data are available through the SILO service. Priestley-Taylor E_0 was produced using interpolated climate station data of temperature and shortwave radiation, and albedo climatology derived from remote sensing (see e.g. Guerschman *et al.*, 2009 for details on these data). Both were available as daily data at 0.05° resolution. The period 2000−2006 was used in all analysis. The assimilated variable in this study was EVI calculated from 1km MODIS reflectance data using the processing methods described in Guerschman *et al.* (2009). Soil moisture of the top few centimetres of soil (θ, in m^3 m^{-3}) can be retrieved using passive microwave remote sensing of soil moisture. However these retrievals are relatively noisy, of relatively coarse resolution (>25 km), and the signal source depth is not clearly defined (but most likely <5 cm). These features make assimilation more challenging and therefore at this stage soil moisture observations were only used for evaluation. The θ data used was retrieved from AMSR-E brightness temperatures using methods described in Owe *et al.* (2008) and made available by VU Amsterdam (see also Draper *et al.*, 2009). Daily flux tower ET estimates were available from four sites: Howard Springs (HoSp) in open forest savannah near Darwin, NT (August 2001−August 2006; J. Beringer, Monash University, Melbourne); Kyeamba (Kye) in open grazing land in southern NSW (January −December 2005, J. Walker, Melbourne University); Tumbarumba (Tumb) in wet open sclerophyll forest in southern NSW (February 2002−April 2005; R. Leuning, CSIRO); and Virginia Park (ViPa) in open woodland savannah in Queensland (July 2001−February 2003; R. Leuning, CSIRO). The time series of EVI and θ were derived by taking the pixels covering the flux tower location.

4. METHODS

4.1. Prior parameter estimation

The six free OCCAM catchment parameters were estimated from long-term average rainfall and PET statistics using regression equations (Van Dijk *et al.*, in review). A single set of seven ET model component parameters was calibrated against long-term average ET estimates derived from average rainfall and streamflow for catchments with sufficient data available (N=199). Fractions of deep- and shallow-rooted vegetation were estimated from 20-m resolution 'woody vegetation cover' data based on Landsat TM mapping (NFI, 1997). The seven parameters of the EGG model for both cover types were derived by visual calibration of parameter values against EVI observations for the flux tower sites and a subset of catchments. Subsequently, the median of visually fitted values across sites and catchments was used as the prior parameter estimate (that is, values were the same for all four sites). These prior parameter estimates can not be seen as fully independent from the EVI observations, but they are independent from the flux tower ET observations.

4.2. Model-data fusion approaches

Nine approaches were tested as described in Table 1. During the visual fitting process, it was found that six out of the fourteen of EGG parameters had more influence on simulated states and fluxes than the remainder. Only these six parameters were therefore used to test alternative parameter estimation approaches (PV, PKL, PKS, PE, PEV). The adjustment of rainfall estimates (PP) was treated as a non-sequential parameter estimation problem: a single parameter was used to linearly scale the daily rainfall time series. The objective function chosen for optimisation was the Nash-Sutcliffe Model Efficiency (NSME) in explaining 16-day EVI averages, except for PE and PEV. For the former, the objective function used was the NSME for monthly average ET, whereas for the latter, the product of the NSMEs for EVI and ET was used. Optimisation was done by random hypercube sampling in three steps, starting with

Table 1. Model data fusion approaches evaluated.

AP	*A priori* parameter estimates (identical for all sites).
PV	non-sequential estimation of the six most sensitive EGG parameters by calibrating against EVI.
PP	non-sequential estimation of a rainfall scaling coefficient by calibrating against EVI.
KL	sequential updating of leaf area index (LAI) using prior parameter values (note that LAI and EVI are directly related in the model and therefore effectively amount to the same).
KS	sequential updating of root zone soil moisture content (Ss) using prior parameter values.
PKL	PV followed by sequential updating of LAI.
PKS	PV followed by sequential updating of Ss.
PE	non-sequential estimation of the six most sensitive vegetation parameters, by calibration against flux tower ET observations.
PEV	non-sequential estimation of the six most sensitive vegetation parameters, by calibration against EVI observations and flux tower ET observations.

the full feasible parameter range and in subsequent steps narrowing down to 50% and 20% of the feasible parameter range around the best estimates. The number of random samples in each step was set at 3^n (i.e. 128 draws for n=6 parameters). In all cases, the model was 'spun up' for the first two years of the period.

Sequential updating was undertaken using the Ensemble Kalman filter (EnKF). Four variations were tested (KL, KS, PKL and PKS). Prior estimates were used for the uncertainty in EVI observations (estimated at ±0.05 from high frequency variation in the data) and model EVI estimates (somewhat arbitrarily estimated at ±0.1). In all cases, 100 ensemble members were used and updating was done each time a new EVI observation became available (typically every 16 days, but depended on cloud cover). It is noted that the PE and PEV strategies are only feasible where flux tower ET measurements are available and therefore not a realistic approach for continental application. They were included only as 'control' cases providing an indication of the performance that can be achieved where such observations are available, and perhaps to identify any model structural issues.

4.3. Evaluation

The performance of alternative strategies was evaluated by comparison against the daily values and monthly averages of observed EVI, ET and θ. A seven-day median filter was applied to both observed and model simulated θ to remove some of the noise in the former caused by nonaligned footprints, atmospheric contamination and retrieval model uncertainty. Indicators of estimation accuracy used were (1) the absolute and relative bias between period average values (2) the fraction of unexplained variance (FoUV; the complement of correlation coefficient, $1-r^2$), and (3) the standard error of estimate (SEE).

5. RESULTS

5.1. Indicators of performance

The agreement between observed and estimated EVI, ET and θ is illustrated in Figure 3, showing average relative bias (but without considering the direction of bias), FoUV and SEE averaged across the four sites. The prior model parameter estimates (AP) provide very reasonable estimates of EVI (when compared to PV and PEV), but more importantly also provide good estimates of ET (when compared to PE and PEV). Parameter optimisation increases model performance for the observation used for calibration, which was expected because the objective function NSME is directly related to SEE, whilst bias and FoUV are likely to be correlated. Compared to the prior parameter estimates (AP), calibration against one observation slightly deteriorates performance in estimating the other: calibration against EVI (ET) using approach PV (PE) leads to greater error in ET (EVI). Rainfall scaling (PP) does not produce better estimates of any of the observations on average (but see further on).

The sequential approaches (KL, KS, PKL, PKS) use EVI observations in updating and therefore for these EVI agreement is not an objective indicator of performance. If prior parameters are used (KL, KS), state updating produced considerably worse ET estimates, but better θ estimates. If PV parameters are used, LAI updating (PKL) produces a slight improvement in SEE and FoUV for ET (from 5.8 to 5.4 mm/mo and from 0.13 to 0.12, respectively) but a greater bias (from 10 to 17%). Soil moisture estimates are better than for any of the other approaches. Updating of S_s (PKS) appears less effective than updating LAI (PKL) on all counts.

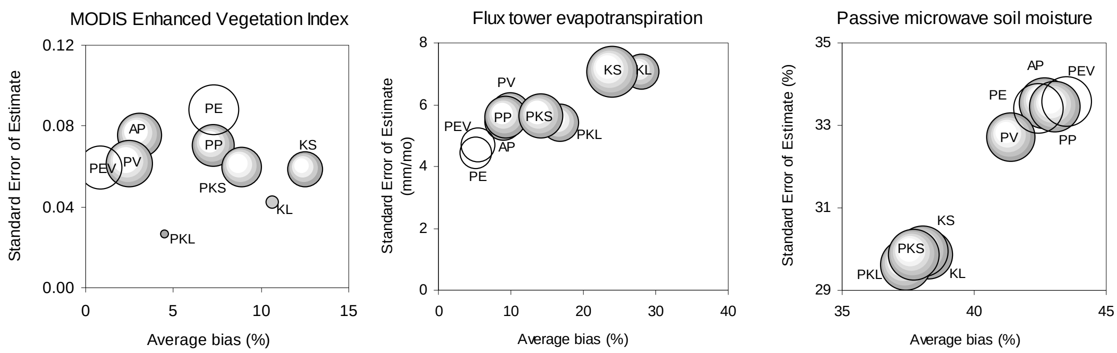
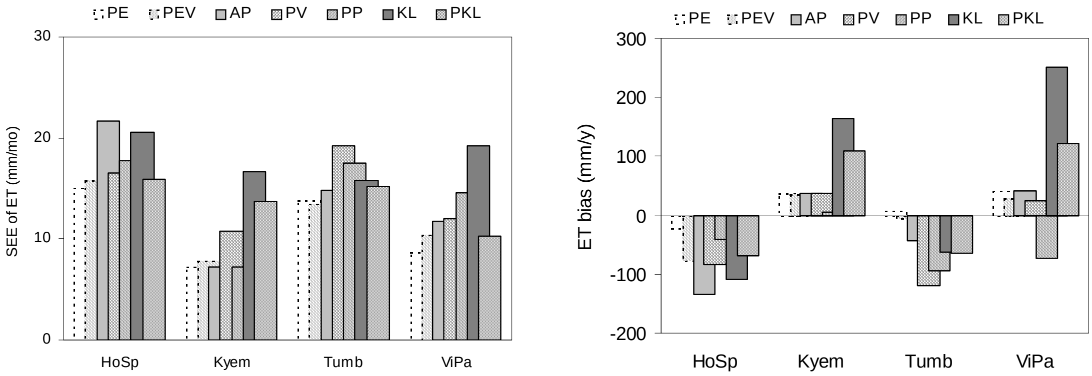


Figure 3. Indicators of estimation uncertainty for alternative model-data fusion approaches for three observation types. Smaller circles closer to the bottom left corner indicate desired performance; the area of the circles is proportional to the unexplained variance, varying between (a) 11–61%, (b) 11–17% and (c) 60–62% (note the truncated axes in this panel). Acronyms are described in the text; the shaded circles indicate approaches feasible at continental scale.



a) b)

Figure 4. Indicators of estimation uncertainty for five of the eight alternative model-data fusion approaches compared to observed monthly ET: (a) unexplained variance, (b) annualised average bias between estimate and observation.

5.2. Differences between sites

Agreement in ET for the four sites separately is illustrated in Figure 4. Results generally reflect those discussed earlier. Overall, the common set of prior parameter estimates produced remarkably good performance, suggesting that the calibrated model parameters are relatively invariant between sites. Calibrating model parameters against observed EVI rather than prescribing values (PV vs. AP) does not always reduce SEE and bias. Similarly, LAI updating does not always lead to improved ET estimates (PKL vs. PV); even less so when prior parameter estimates are used (PKL vs. AP). Biases tend to have equal sign within a site regardless of the method. Comparing across sites, rainfall input scaling (PP) appeared to result in generally similar or perhaps even slightly better performance than vegetation parameter optimisation (PV).

6. DISCUSSION

6.1. Merit of alternative MDF approaches

The apparently modest gains achieved through any of the model-data fusion approaches when compared with prior parameter estimates was surprising. The influence of errors in the forcing data and errors in the model structure should be considered here, and there may also have been issues of representativeness between the site measurements and the satellite pixel. Calibrating the model parameterisation generally led to somewhat better agreement with the target observations, but deterioration against other observations. This may indicate a degree of ‘over fitting’ to compensate errors and bias in the observations. It does not explain why state updating with prior parameter estimates performed worse than state updating with calibrated parameter values. Overall, most approaches showed similar biases when compared to ET observations. State updating that brought modelled EVI in closer agreement with observations also improved agreement with θ slightly. We speculate that this is primarily because of errors in the gridded rainfall inputs. As the model accounts for soil water, such errors can have a considerable influence on the magnitude and timing of simulated EVI, ET and θ under water limited conditions. Differences in estimated versus observed rainfall and ET indeed showed some consistency in biases: average SILO rainfall varied by −13%, +4%, −16% and −2% from site rainfall (at HoSp, Kyem, Tumb and ViPa, respectively), whereas model ET varied by −4%, +1%, −15% and −17% from observations. The relatively good performance of rainfall scaling (PP) in reproducing ET estimates, and in particular the smaller bias for some of the sites, may provide some support for the importance of rainfall errors. However, the resulting adjusted rainfall estimates were generally worse than the SILO estimates; varying by +25%, −3%, −33% and −29% from tower observations. This is not surprising considering the rather crude uniform scaling method applied. Rainfall input adjustment was only tested in a relatively crude, non-sequential manner but results were encouraging, and would appear a promising candidate for sequential approaches, particularly when considering the additional information about rainfall contained in remotely sensed soil moisture and land surface temperature dynamics.

6.2. Trade-off between performance gain and computational overhead

In this application, the computational overheads of alternative MDF approaches increased from using prior parameter estimates (AP, a single run of ~$2x10^3$ model evaluations), Kalman filter approaches (KL, KS, 100 ensemble members each day producing ~$2x10^5$ model evaluations) and non-sequential approaches (PV, PE, PEV, PP – three times 128 the full period, i.e. ~$8x10^5$ model evaluations), to combined parameter estimation and state updating (e.g. PKL and PKS, in the vicinity of ~10^6 evaluations). It is noted that 100 ensemble members may be towards the upper end of commonly used ensemble size, whereas 348 parameter samples is perhaps towards the lower number of iterations needed when optimising six parameters (though more than necessary when optimising one parameter as for PP). The only approach that reduced SEE in monthly AET and θ (by −1% and −12%, respectively, when compared to method AP) was the combined parameter estimation-state updating approach (PKL) that was also one of the two most computationally intensive.

Whether implementation of this approach is practically achievable depends on the application. It was practical for this study (taking minutes), but a single run for the same period across the Australian continent ($7.7 \cdot 10^6$ km^2) at the 250-m MODIS EVI resolution would create ~$4 \cdot 10^{11}$ calculations per run. This by itself is still within the realms of current computing power, for example some of the more involved weather prediction systems perform tasks of similar magnitude within a few hours on super computers. However the combined parameter estimation and state updating approach tested here would increase computational load by about three orders of magnitude and become impractical. Computing speed will almost certainly continue to increase, and considerable efficiencies can probably be found in practice; this is subject of further research.

7. CONCLUSIONS

Several model-data fusion approaches were tested using a coupled catchment water balance and vegetation dynamics model. The performance of alternative model-data fusion approaches was evaluated using flux tower ET measurements at four sites in Australia and satellite observations of soil moisture over the corresponding area (AMSR-E passive microwave instrument). Satellite-observed vegetation vigour (MODIS EVI) was the assimilated variable. Non-sequential parameter estimation did not provide much benefit over the use of prior parameter estimates, suggesting that the model parameterisation was comparatively robust and parameter values spatially invariant by good approximation. A combination of parameter estimation and state updating did lead to improvements in some aspects of evaluation, but the computational costs were about three orders of magnitude greater. Further research is planned to try and find more computationally efficient ways of achieving a similar result.

ACKNOWLEDGMENTS

This research was funded through the water information research and development alliance between the Bureau of Meteorology and the CSIRO Water for a Healthy Flagship Program. We gratefully acknowledge the assistance of Robert Pipunic, Jeff Walker, Ray Leuning, Lindsay Hutley, Jason Beringer, Juan Pablo Guerschman, Yi Liu and Richard de Jeu in collecting and/or processing the data for this study.

REFERENCES

Akaike, H., (1970) Statistical predictor identification *Ann. Inst. Stat. Math.*, 22, 203–217

Barrett, D.J., Kuzmin, V.A., Walker, J.P., McVicar, T.R., Draper, C., (2008) Improving streamflow forecasting by integrating satellite observations, in situ data and catchment models using model-data assimilation methods. eWater Technical Report., pp. 53. eWater CRC, Canberra, Australia.

Draper, C.S., Walker, J.P., Steinle, P.J., de Jeu, R.A.M., Holmes, T.R.H., (2009) An evaluation of AMSR-E derived soil moisture over Australia. *Rem. Sens. Env.* 113(4), 703-710.

Guerschman, J.-P., Van Dijk, A.I.J.M., Mattersdorf, G., Beringer, J., Hutley, L.B., Leuning, R., Pipunic, R.C., Sherman, B.S., (2009) Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia *J. Hydr.* 369, 107-119.

NFI, (1997) National Plantation Inventory of Australia. National Forest Inventory, BRS, Canberra.

Owe, M., De Jeu, R.A.M., Holmes, T., (2008) Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res.*, 113(F01002).

Van Dijk, A.I.J.M., Guerschman, J.P., Marvanek, S., (in review) Derivation of an Optimum Complexity Catchment Model (OCCAM) from gauging, terrain and satellite data. A report for the Bureau of Meteorology., pp. 100. Water for a Healthy Country Flagship, CSIRO, Canberra, Australia.