

C. Rüdiger, J.P. Walker, J.D. Kalma, G.R. Willgoose, and P.R. Houser (2004), **Streamflow data assimilation: A study on nested catchments**, in *Proceedings of the 2nd international CAHMDA workshop on: The Terrestrial Water Cycle: Modelling and Data Assimilation Across Catchment Scales*, edited by A.J. Teuling, H. Leijnse, P.A. Troch, J. Sheffield and E.F. Wood, pp. 140–143, Princeton, NJ, October 25–27

Streamflow data assimilation: A study on nested catchments

C. Rüdiger¹, J.P. Walker¹, J.D. Kalma², G.R. Willgoose³, and P.R. Houser⁴

¹*Dept. of Civil and Environmental Engineering, University of Melbourne, Parkville, Australia*

²*School of Engineering, University of Newcastle, Callaghan, Australia*

³*School of Geography, University of Leeds, Leeds, UK*

⁴*Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, USA*

Soil moisture is an important variable in land surface modelling with a significant impact on climate prediction, but the areas shown to have the greatest potential impact are typically also the most densely vegetated. While much work has been concentrated on the assimilation of remotely sensed surface soil moisture observations to constrain land surface model predictions of soil moisture, the use of these measurements is limited to areas of low-to-moderate vegetation. This work proposes to contribute to soil moisture prediction in those densely vegetated areas through the assimilation of streamflow observations. The potential for this approach is demonstrated for a semiarid catchment in a synthetic twin experiment.

Introduction Climate model results are strongly dependent on the initial soil moisture conditions predicted by a land surface model. Moreover, it has been shown that correct initialisation of soil moisture content in areas with dense vegetation cover, such as the Sahel, the Amazon, and south-east Asia, has the greatest potential for positively influencing the predictability of precipitation (*Koster et al.*, 2000).

Previous work has shown that initial conditions on root zone soil moisture content can be accurately predicted when the near-surface soil moisture observations that are available from remote sensing are assimilated into a land surface model (e.g., *Walker and Houser*, 2001). However, the approach is limited to areas with low-to-moderate vegetation cover, as dense vegetation masks the remotely sensed soil moisture signal. Thus, we seek to improve soil moisture prediction in densely vegetated catchments through the assimilation of observed streamflow. As streamflow is an integrated measure of soil moisture content and rainfall events hours, days, or even weeks in the past, implementation of an assimilation scheme to account for this time-lag requires careful consideration.

This paper uses a brute-force implementation of the variational assimilation approach in a synthetic study to demonstrate that this approach to soil moisture initial condition retrieval is feasible. This study is the first step towards a field-based multi-catchment study.

Models The assimilation of streamflow data for the retrieval of initial soil moisture content is addressed in this paper through a synthetic data assimilation study. First, a land surface model is used to generate a “true” data set that provides both the surface soil moisture “observations” and the evaluation data. The initial conditions and land surface forcing data are then degraded in two individual experiments in order to obtain control results.

The land surface model used in this study is the Catchment Land Surface Model (CLSM) of *Koster et al.* (2000). Its framework includes an explicit treatment of sub-catchment soil moisture variability and its effect on runoff and evaporation. Consideration of both spatial distribution of the water table depth and non-equilibrium conditions in the root zone leads to the definition of three bulk moisture prognostic variables (catchment deficit, root zone excess and surface excess) and a special treatment of moisture transfer between them. Using these three prognostic variables, the

catchment may be divided into regions of stressed, unstressed and saturated soil moisture regimes, and the soil moisture profile calculated.

Model runoff is only produced when the current soil moisture content exceeds field capacity and is routed instantaneously to the catchment outlet. To counter this, an inter-catchment routing scheme has been introduced into the model to separately route the surface and subsurface runoff generation throughout the catchment. Surface and subsurface routing is undertaken using a digital elevation model of the catchment and an approximation to the Manning's equation; $V = cS^{0.5}$, where V is velocity, S is surface slope and c is a parameter fitted individually for surface, subsurface, and streamflow conditions to observed streamflow data for the specific catchment. While this parameter depends on the surface and flow conditions of the individual pixels, we assume a single uniform value for the hillslope and streamflow runoff routing components, respectively. Due to their size, travel time in most catchments is short, and losses due to evaporation were therefore considered negligible.

The variational data assimilation approach is based on minimising an objective function over an assimilation window, rather than sequentially using individual observations. In our application we use the Bayesian nonlinear regression suite (NLFIT) of *Kuczera* (1983), which is based on the shuffled complex evolution method of *Duan et al.* (1992), to perform this optimisation. The optimisation is achieved by changing the initial soil moisture state variables until the best fit between model predicted and observed streamflow for a given assimilation window is achieved.

The length of the assimilation window was kept to one month for the results presented in this paper, while the input parameters necessary for the calculations within NLFIT (Box-Cox λ , Box-Cox K , autoregressive parameter) were initially estimated and later manually adjusted. The initial assumption being that the data were normally distributed.

Synthetic Experiments To demonstrate the feasibility of the proposed approach, a set of synthetic experiments have been undertaken for a subcatchment of the Goulburn River experimental catchment in SE Australia (*Rüdiger et al.*, 2003). Observed meteorological forcing data from five weather stations located within and surrounding the Goulburn River catchment were used as input to the model. In this application averages for the five weather stations were applied uniformly throughout the catchment for each forcing parameter. The data set used comprised a one-year period. Forcing data used by CLSM are temperature, wind speed, precipitation, specific humidity, and long and short wave downward solar radiation. While most of these forcing data are observed by the weather station, radiation observations were not available and radiation data were therefore obtained from the Global Data Acquisition System (GDAS) model.

Soil and vegetation properties in the subcatchment were assumed to be spatially uniform and estimated as the dominant value for each parameter. Vegetation data and greenness index were obtained from $0.25^\circ \times 0.25^\circ$ global vegetation maps, while soil type and properties were taken from the a digitised version of the Australian Soils Atlas (*Northcote et al.*, 1968). Topographic data are incorporated in the model from compound topographic index calculated from a DEM with a resolution of 250 m.

Using the input data described above, CLSM was spun-up for a one-year period on the aforementioned forcing data, by repeated simulation until convergence of the initial conditions to an equilibrium state was achieved. The output from the subsequent simulation with these initial conditions was then assumed to be the "true" data. This was used to provide both the one-hourly streamflow observation data and the soil moisture time series evaluation data. The assimilation experiments were run for one month of this period during which two significant runoff events took place.

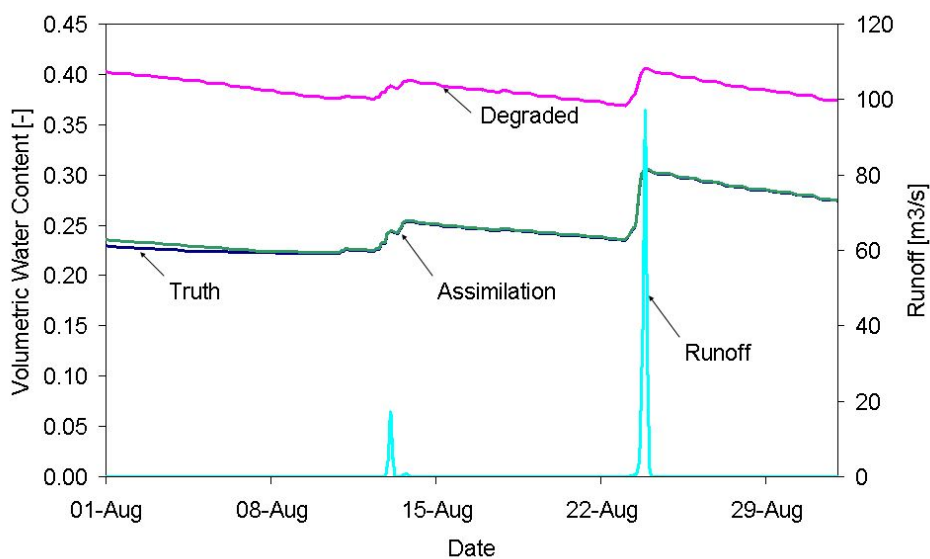


Figure 4.3: Results for experiment 1 assimilation run showing profile soil moisture content and observed streamflow.

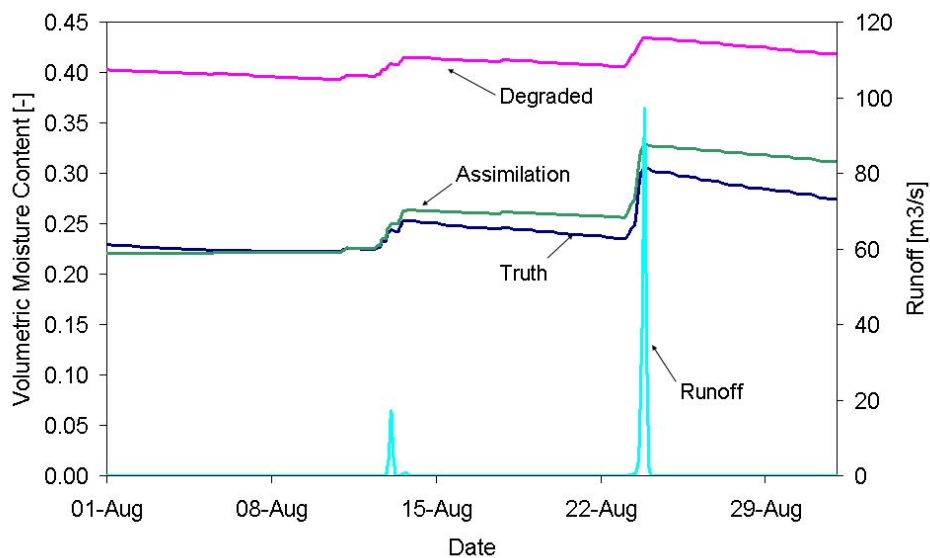


Figure 4.4: Results for experiment 2 assimilation run showing profile soil moisture content and observed streamflow.

Streamflow data output from the model is only available for the catchment outlet and is not available for ungauged tributaries within the catchment. Hence, streamflow output is only a lumped value for the whole upstream catchment. Similarly, soil moisture values are catchment averages of the surface, root zone and profile soil moisture content.

Two synthetic experiments were undertaken to demonstrate the assimilation of streamflow data for soil moisture retrieval. First, only the initial conditions of the three prognostic soil moisture states (catchment deficit, surface excess, and root zone excess) were arbitrarily degraded (experiment 1) while the forcing and observation data were perfect; initial conditions were degraded arbitrarily, so that an extreme wet condition was created. Second, both the soil moisture initial conditions and the forcing data were artificially degraded (experiment 2); the precipitation was increased by 20% and the incoming solar radiation was decreased by 33%. In this way we were able to explore the effect of erroneous forcing data on the assimilation results. The impact of this was to create an artificially wet catchment simulation.

Figures 4.3 and 4.4 show the catchment average profile soil moisture content for the assimilation and no assimilation runs compared to the “true” data and the “observed” streamflow, for experiment 1 and 2 respectively. In experiment 1 there is a good agreement between the retrieved and truth soil moisture data, while the results from no assimilation continue to overestimate the true soil moisture amount. While there is a small difference between the true and retrieved initial soil moisture states, this is quickly corrected when a streamflow runoff event takes place. Although these results were anticipated, given that we used a perfect model with perfect observations, it still demonstrates that there is a significantly strong relationship between the streamflow prediction and profile soil moisture content, even under semi-arid conditions.

Results from experiment 2 show a similar improvement in soil moisture prediction when streamflow observations are assimilated, compared with the no assimilation run. However, due to the dramatic increase in precipitation and decrease in radiation compared to the truth, the soil moisture prediction, even with assimilation, tends to predict a soil moisture content which is too wet towards the end of the assimilation window. To counter this, the assimilation attempts to make the initial conditions as dry as possible, but is constrained by the residual soil moisture content set in the model.

Both experiments showed good results for the retrieval of catchment average soil moisture profiles and adequate results for the root zone moisture content. However, retrieval of surface soil moisture was problematic for experiment 2. In this case, NLFIT was unable to find an optimum value for the initial surface excess prognostic variable. This is due to the small influence of the initial surface excess states on the total runoff, and the fact that only minor changes were produced in the objective function, even for major changes of the initial prognostic state value.

Conclusion This study has demonstrated using synthetic experiments that streamflow data assimilation has the potential to improve model prediction of soil moisture. Using a variational data assimilation approach, initial soil moisture states were retrieved by “calibrating” the model streamflow prediction to observed streamflow records. As residual soil moisture and porosity set upper and lower limits on the initial soil moisture storage, it was not possible to correctly predict the soil moisture time series over long time periods in the presence of large water balance errors. It is proposed that this problem may be resolved through shorter event based assimilation windows.

Acknowledgements This project is funded by an ARC Discovery Grant (DP 0209724). A visit by the principal author to NASA Goddard Space Flight Center was sponsored by a Postgraduate Overseas Research Exchange Scholarship from the University of Melbourne. The invaluable help and insight to the catchment land surface model provided by Randy Koster and Sarith Mahanama is greatly appreciated.

Bibliography

- Duan, Q., S. Soorooshin, and V. Gupta, Effective and efficient global optimisation for conceptual rainfall-runoff models, *Water Resour. Res.*, 28, 1015–1031, 1992.
- Koster, R., M. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, A catchment-based approach to modeling land surface processes in a general circulation model, 1. Model structure, *J. Geophys. Res.*, 105(D20), 24,809–24,822, 2000.
- Kuczera, G., Improved parameter inference in catchment models, *Water Resour. Res.*, 19(5), 1151–1162, 1983.
- Northcote, K., et al., Atlas of Australian soils, sheets 1 to 10. with explanatory data, CSIRO Aust. and Melbourne University Press, Melbourne, 1968.
- Rüdiger, C., R. Davidson, H. Hemakumara, J. Walker, J. Kalma, G. Willgoose, and P. Houser, Catchment monitoring for scaling and assimilation of soil moisture and streamflow, in *Proceedings of the international congress on modelling and simulation (MODSIM), 14–17 July, 2003*, pp. 386–391, Modelling and Simulation Society of Australia and New Zealand, Townsville, Australia, 2003.
- Walker, J., and P. Houser, A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *J. Geophys. Res.*, 106(D11), 11,761–11,774, 2001.