



The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction



A. Kunnath-Poovakka^{a,*}, D. Ryu^a, L.J. Renzullo^b, B. George^c

^a Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia

^b CSIRO Land and Water, P.O. Box 1666, Canberra, ACT 2601, Australia

^c Integrated Water and Land Management Program, ICARDA, P.O. Box 2416, Cairo, Egypt

ARTICLE INFO

Article history:

Received 7 September 2015

Received in revised form 8 February 2016

Accepted 11 February 2016

Available online 22 February 2016

This manuscript was handled by Tim R.

McVicar, Editor-in-Chief, with the assistance of Di Long, Associate Editor

Keywords:

Evapotranspiration (ET)

Soil moisture (SM)

Streamflow

Calibration

Remotely sensed (RS) data

SUMMARY

Calibration of spatially distributed hydrologic models is frequently limited by the availability of ground observations. Remotely sensed (RS) hydrologic information provides an alternative source of observations to inform models and extend modelling capability beyond the limits of ground observations. This study examines the capability of RS evapotranspiration (ET) and soil moisture (SM) in calibrating a hydrologic model and its efficacy to improve streamflow predictions. SM retrievals from the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and daily ET estimates from the CSIRO MODIS ReScaled potential ET (CMRSET) are used to calibrate a simplified Australian Water Resource Assessment – Landscape model (AWRA-L) for a selection of parameters. The Shuffled Complex Evolution Uncertainty Algorithm (SCE-UA) is employed for parameter estimation at eleven catchments in eastern Australia. A subset of parameters for calibration is selected based on the variance-based Sobol' sensitivity analysis. The efficacy of 15 objective functions for calibration is assessed based on streamflow predictions relative to control cases, and relative merits of each are discussed. Synthetic experiments were conducted to examine the effect of bias in RS ET observations on calibration. The objective function containing the root mean square deviation (RMSD) of ET result in best streamflow predictions and the efficacy is superior for catchments with medium to high average runoff. Synthetic experiments revealed that accurate ET product can improve the streamflow predictions in catchments with low average runoff.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Significant research has been done in the past to develop efficient calibration algorithms to attain reliable streamflow predictions at gauged catchments. Calibration of hydrologic/land surface models is carried out usually using streamflow since it gives time-integrated information about water output from the catchment. However, lack of streamflow observations and other ground data in the vast majority of areas makes hydrological model calibration a difficult task. Particularly, accurate estimation of runoff at ungauged catchments is a growing concern for the hydrologic community (Sivapalan, 2003; Wagener and Montanari, 2011).

Alternative approaches to the conventional streamflow-based calibration for ungauged locations include estimation of parameters from prior information on catchment physical characteristics (e.g., soil hydraulic properties and vegetation properties)

(Atkinson et al., 2003; Koren et al., 2003) and regionalization. For example, in parameter regionalization, a model is calibrated for a number of gauged watersheds and the model parameters are derived from a regression relationship between the watershed characteristics and parameters (Abdulla and Lettenmaier, 1997; Jakeman et al., 1992; Parajka et al., 2007; Post et al., 1998; Sefton and Howarth, 1998; Viney et al., 2009; Wagener and Wheater, 2006; Wagener et al., 2004; Zhang et al., 2011). Yadav et al. (2007) modified the approach by regionalizing flow characteristics and by incorporating uncertainty in the regressed estimates. However, the drawbacks of these approaches have been reviewed previously (Beven, 2000; Wagener and Montanari, 2011; Wagener and Wheater, 2006); uniqueness in topography, geology, vegetation features and anthropogenic modification for the watersheds within similar climatic region can make individual watersheds respond to input in widely different manners (Beven, 2000). Wagener and Wheater (2006) stated that through regionalization, model structural uncertainty is transferred to the optimized parameters, which can result in biased calibration. To explore methods that can

* Corresponding author.

overcome these issues associated with the regionalization, we investigate the efficacy of RS data in calibration and its utility as an alternative source of observational data to constrain our model parameter estimates for ungauged locations.

Typically, RS data provides spatially distributed land surface parameters with a regular temporal repeat across much of the globe, and the volume and quality of data have greatly increased over the past decades providing extensive datasets for hydrologic modelling. The RS data has been widely used in land use and land cover classification (e.g., soil and land use pattern) and the estimation of vegetation indices, surface SM and evapotranspiration (ET) (Asner et al., 2003; Chen and Cihlar, 1996; Congalton, 1991; Gupta et al., 2008; Kustas et al., 1994; Otlé et al., 1989; Owe et al., 2008; Sucksdorff and Otle, 1990). Resulting land surface products have enhanced overall capability of hydrologic modelling from local to global scales. For example, several studies demonstrated that annexation of RS land surface temperature, SM, leaf area index (LAI) and fraction vegetation cover data improves hydrologic modelling (Andersen et al., 2002; Corbari and Mancini, 2013; Crow and Ryu, 2009; Garcia-Quijano and Barros, 2005; Parajka et al., 2006; Renzullo et al., 2008; Silvestro et al., 2015; Sutanudjaja et al., 2014; Zhang and Wegehenkel, 2006). LAI estimated by the MODerate resolution Imaging Spectrometer (MODIS) onboard Terra and Aqua satellites can be used with the Penman–Monteith equation to produce 8-day composite ET (Cleugh et al., 2007; Leuning et al., 2008; Zhang et al., 2008). Satellite-derived products, such as ET, SM and vegetation biomass, have become important parts of hydrological modelling.

ET retrieved from AVHRR (Advanced Very High Resolution Radiometers) (Nemani and Running, 1989; Taconet et al., 1986) and MODIS (Cleugh et al., 2007; Guerschman et al., 2009; Leuning et al., 2008; Zhang et al., 2008) satellites and microwave SM retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) (Njoku et al., 2003; Owe et al., 2008), Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2001) of European Space Agency (ESA), and Advanced Scatterometer (ASCAT) (Bartalis et al., 2007) have widely been used in hydrology. Immerzeel and Droogers (2008) calibrated the Soil Water Assessment Tool (SWAT) model with MODIS ET using Gauss–Marquardt–Levenberg (GML) algorithm which increased correlation between simulated and observed ET resulting in improved streamflow predictions. Using a two-parameter surface conductance (G_s) model, optimized by catchment water balance estimates of ET (precipitation minus runoff), Zhang et al. (2008) generated 8-day composite ET (E_{RS}) from MODIS LAI. The E_{RS} was useful in estimating long-term runoff using the concept of water balance. But the study did not address the capability of E_{RS} in calibrating rainfall–runoff models. Zhang et al. (2009) concluded that multi-objective calibration of SimHyd model with streamflow and E_{RS} produced better daily and monthly runoff compared to calibration with streamflow alone. Droogers et al. (2010) showed that optimization of the Soil–Water–Atmosphere–Plant (SWAP) model using satellite-derived actual ET can predict irrigation demand with acceptable accuracy.

Hydrological application of RS SM has been primarily on its assimilation into land surface models to improve profile SM and other outputs linked to the soil (Crow and Van den Berg, 2010; Han et al., 2012; Pauwels et al., 2001; Reichle and Koster, 2005; Renzullo et al., 2014). Recent studies have shown that assimilation of satellite SM can improve streamflow predictions of rainfall–runoff models (Alvarez-Garreton et al., 2014, 2015; Crow and Ryu, 2009; Parajka et al., 2006; Pauwels et al., 2001). Similarly, the assimilation of a Soil Wetness Index (SWI) derived from ASCAT into a continuous distributed hydrologic model resulted in improved discharge predictions (Brocca et al., 2010, 2012).

Very few studies have incorporated satellite SM in model calibration. Campo et al. (2006) used SM retrievals from European Remote Sensing (ERS) scatterometer signals to optimize the parameters of soil dynamics in distributed hydrologic model. Parajka et al. (2009) used ERS scatterometer signals to calibrate semi-distributed hydrologic model along with runoff. Although both studies demonstrated improved streamflow predictions, Campo et al. (2006) study was limited to the areas with no or sparse vegetation while the study by Parajka et al. (2009) showed deterioration in SM predictions. Zhang et al. (2011) calibrated the Australian Water Resource Assessment landscape model (AWRA-L) with streamflow, NOAA-AVHRR LAI and TRMM-MI (Tropical Rainfall Measuring Mission – Microwave Imager) SM using multi-objective criteria. Even though the study resulted in marked improvement of LAI and SM, the improvement of streamflow was marginal. Sutanudjaja et al. (2014) calibrated a physically based large-scale coupled groundwater–land surface model, called PCR-GLOBWB-MOD, using SWI derived from ERS scatterometer and discharge data. The resulting optimum parameter set predicted discharge, SM and groundwater dynamics with acceptable accuracy.

Challenges in using remote sensing data for model calibration and data assimilation are discussed comprehensively in Van Dijk and Renzullo (2011). One of the major challenges is that the information content of RS observations varies based on transient vegetation (Barrett and Renzullo, 2009; Campo et al., 2006) and topography (Parajka et al., 2009). Appropriate specification of errors, as part of the retrieval process, is critical for effective calibration and data assimilation (Alvarez-Garreton et al., 2014). Suitability of coarse-resolution remote sensing data for small-scale studies is also a concern.

Previous studies show that RS data can be used to optimize the model parameters under certain conditions (Kunnath Poovakka et al., 2013; Mohanty, 2013; Zhou et al., 2013). Zhang et al. (2008) suggest that RS ET and SM can be used in the calibration of rainfall–runoff models to improve runoff estimations in ungauged catchments. A calibration scheme which relies solely on remote sensing data will be greatly beneficial in modelling at ungauged catchments, especially if it can be demonstrated to result in improved estimation compared with uncalibrated model. In this study, Microwave SM retrievals from the AMSR-E and daily estimates of ET from CSIRO MODIS ReScaled potential ET (CMRSET) model are used to calibrate a hydrologic model using 15 different objective functions considering various combinations of RMSD and correlation of ET and SM. The Shuffled Complex Evolution (SCE) calibration algorithm is used to calibrate a grid-based hydrologic model modified from the Australian Water Resource Assessment – Landscape (AWRA-L) model. Main research objectives of this study are to investigate (1) how effective RS ET and SM are in calibrating the model and (2) how biases in RS ET affects prediction of calibrated models.

2. Materials and methods

2.1. Study catchments

The efficacy of calibrating AWRA-L with RS ET and SM is assessed for 11 catchments in eastern Australia. First, the calibration method is developed and tested in the Loddon River catchment at Newstead, Victoria, Australia and then the method is evaluated in 10 catchments along eastern Australia (Fig. 1). The Loddon River catchment contains the Wombat flux tower station upstream and is one of the hydrologic reference stations maintained by the Australian Bureau of Meteorology (BoM). The observed ET and SM data from the Wombat station are used to evaluate the satellite data.

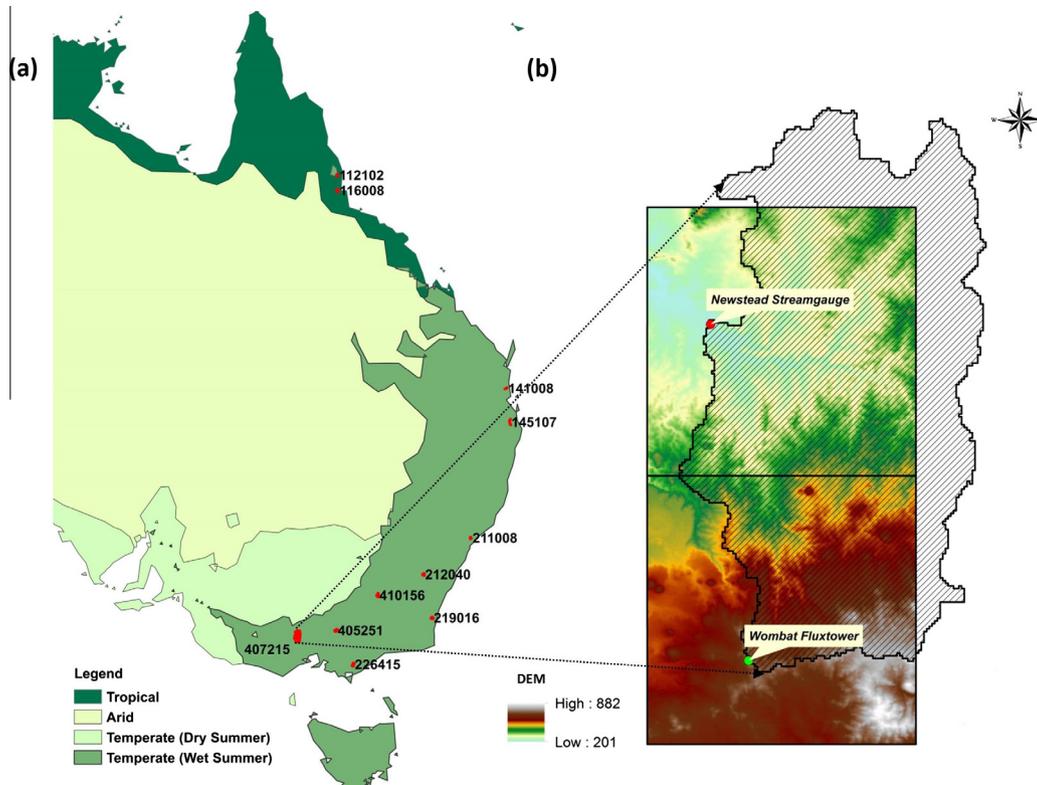


Fig. 1. (a) Climatic zones of eastern Australia identifying the study catchments (by catchment ID) used in this study; (b) the Loddon River catchment at Newstead with overlaid AWRA grids cells and digital elevation model.

The Loddon River catchment at Newstead (1028 km²) is located in temperate wet summer climatic region with an average annual rainfall of 700 mm. Loddon River at Newstead is a big catchment compared to other ten. Therefore, the calibration is performed at a grid resolution of 0.25° × 0.25°. Most of the catchment is occupied in two grid cells of 0.25° × 0.25° resolution mesh grid of Australia (Fig. 1b). The lower grid cell is mostly covered by the Wombat State Forest with dominant tree species *Eucalyptus obliqua*, *Eucalyptus radiata* and *Eucalyptus rubida*, whereas the upper grid cell is mainly grasslands.

Ten evaluation catchments chosen are shown in Fig. 1a and their details are summarized in Table 1. As they are small catchments in comparison with the Loddon River catchment, calibration is performed at a finer grid resolution of 0.05° × 0.05°.

2.2. Simplified AWRA-L model

The AWRA-L is a grid-based hydrologic model developed by the Commonwealth Scientific and Industrial Research Organization

(CSIRO) and BoM, Australia (Renzullo et al., 2014; Van Dijk, 2010; Vaze et al., 2013). It is a one-dimensional model calculating the flows and stores of water at the land surface on a daily time step. Ground-based (e.g., from stream gauges and flux towers) and remote sensing observations have been used in the model development and for defining parameters to ensure consistency in model estimates with the observable components of water states and fluxes (Van Dijk, 2010).

Compared with other hydrologic/land surface models, AWRA-L requires relatively small number of forcing variables for predicting major water balance components such as streamflow, ET and SM. The input forcing variables of the model are daily precipitation, daily minimum and maximum temperature and solar radiation. The model consists of three unsaturated soil layers, a groundwater store and a separate routing water store where surface and subsurface flows join. In the AWRA-L version 0.5 lateral redistribution of water between grid cells is not considered.

The original version of AWRA-L has two Hydrologic Response Units (HRU) per each grid cell: deep-rooted and shallow-rooted

Table 1
Summary of 11 study catchments.

Station ID	Gauging station	River/creek	Basin	Area (km ²)	Climatic region	Avg runoff (mm/day)
407215	Newstead	Loddon Rv	Loddon Rv	1028	Temperate wet summer	0.04
112102	Upper Japoonvale	Liverpool Ck	Johnstone	78	Tropical	5.10
116008	Abergowrie	Gowrie Ck	Herbert	124	Tropical	2.64
141008	Kiels Mountain	Eudlo Ck	Maroochy	62	Temperate wet summer	0.96
145107	Main Rd Br	Canungra Ck	Logan-Albert	101	Temperate wet summer	0.52
211008	Avondale	Jigadee Ck	Macquarie Tuggerah Lakes	55	Temperate wet summer	0.38
212040	Pomeroy	Kialla Ck	Hawkesbury River Basin	96	Temperate wet summer	0.04
219016	Cobargo	Narira Rv	Bega River Basin	92	Temperate wet summer	0.21
226415	Traralgon South	Traralgon Ck	Latrobe	128	Temperate wet summer	0.30
405251	Ancona	Brankeet Ck	Goulburn	121	Temperate wet summer	0.16
410156	Book Book	Kyeamba	Murrumbidgee River	145	Temperate wet summer	0.05

vegetation. The fraction of each HRU and its land-cover-specific parameters are determined based on the fraction of tree coverage estimated from the historical AVHRR vegetation dynamics in 1981–2006 (Donohue et al., 2009). Parameters such as groundwater drainage coefficient and drainage fraction at field capacity are calculated using the Budyko's dryness index, which is the ratio of mean annual precipitation to mean annual potential evapotranspiration. Numerical solutions to Richard's equation and Brooks–Corey soil water relationships are used to describe soil–water drainage in AWRA-L (Van Dijk and Marvanek, 2010). A Penman–Monteith formulation of potential ET is used, and augmented by a Priestly–Taylor formulation when wind, air pressure and vapour pressure observations are not available (Van Dijk, 2010). The total ET from the model is the sum of transpiration and evaporation from rainfall interception, soil and groundwater (evaporation from ground-water saturated areas). AWRA-L simulates the vegetation cover changes in response to the soil water availability, and equilibrium leaf biomass is calculated by considering the hypothetical leaf biomass obtained when maximum transpiration rate is equal to the maximum root water uptake.

To simplify the parameterisation of AWRA-L for the investigation, the model is modified in the current application to have only one HRU per grid cell. That is, each grid cell is classified as either shallow or deep-rooted vegetation depending on the tree fraction and parameters are selected accordingly. If the tree fraction is greater than 50% of a grid cell, its vegetation cover is considered as deep-rooted else as shallow-rooted. A simplified conceptual diagram of simplified AWRA-L model is shown in Fig. 2. Unlike many other models, thickness of soil layers is not fixed in the original AWRA-L; indicative soil layer thickness varies with respect to water content at field capacity of that layer (S_zFC). A topsoil layer thickness of 50 mm is defined to approximate the sensing depth of AMSR-E. For this study, the depth of top, shallow and deep soil layers are fixed at 50 mm, 250 mm and 1200 mm by specifying the parameter for available water content at field capacity, S_zFC , as:

$$S_zFC = \frac{FC}{100} * Depth \quad (1)$$

where FC is the field capacity of soil in percentage and $Depth$ is the soil layer depth.

2.3. Data

Daily time series of precipitation, minimum and maximum temperature and solar radiation from January 2003 to December 2010 derived from the Australian Water Availability Project (AWAP) are used as input forcing (<http://www.csiro.au/awap/>). AWAP data are available at a grid resolution of $0.05^\circ \times 0.05^\circ$ for across Australia.

RS ET and surface SM are used for model calibration. Daily estimates of ET from CMRSET based on MODIS surface reflectance (Terra satellite) in visible, near-infrared and shortwave infrared range (Guerschman et al., 2009), and microwave SM retrievals from the C- and X-band brightness temperatures of AMSR-E version 5.0 (Owe et al., 2008; Su et al., 2013) are employed. Actual ET derived from MOD43B4 MODIS product is available from January 2001 to December 2013.

The original AWAP and CMRSET data with $0.05^\circ \times 0.05^\circ$ (approximately $5 \text{ km} \times 5 \text{ km}$) resolution over Australia is aggregated to the model grid scale $0.25^\circ \times 0.25^\circ$ for calibration in the Loddon River catchment at Newstead. For ten small catchments, AMSR-E SM, which represents the volumetric water content of top 1–2 cm of soil at a resolution of $0.25^\circ \times 0.25^\circ$, is downscaled to $0.05^\circ \times 0.05^\circ$.

Daily streamflow records from the BoM website (<http://www.bom.gov.au/water/hrs/index.shtml>) at the Loddon River at Newstead (gauge identifier: 407215) is used for evaluating the efficacy of the calibration scheme in predicting streamflow. Daily ET and SM measures from the Wombat flux tower station for the period 2010–2013 are used to assess the accuracy of

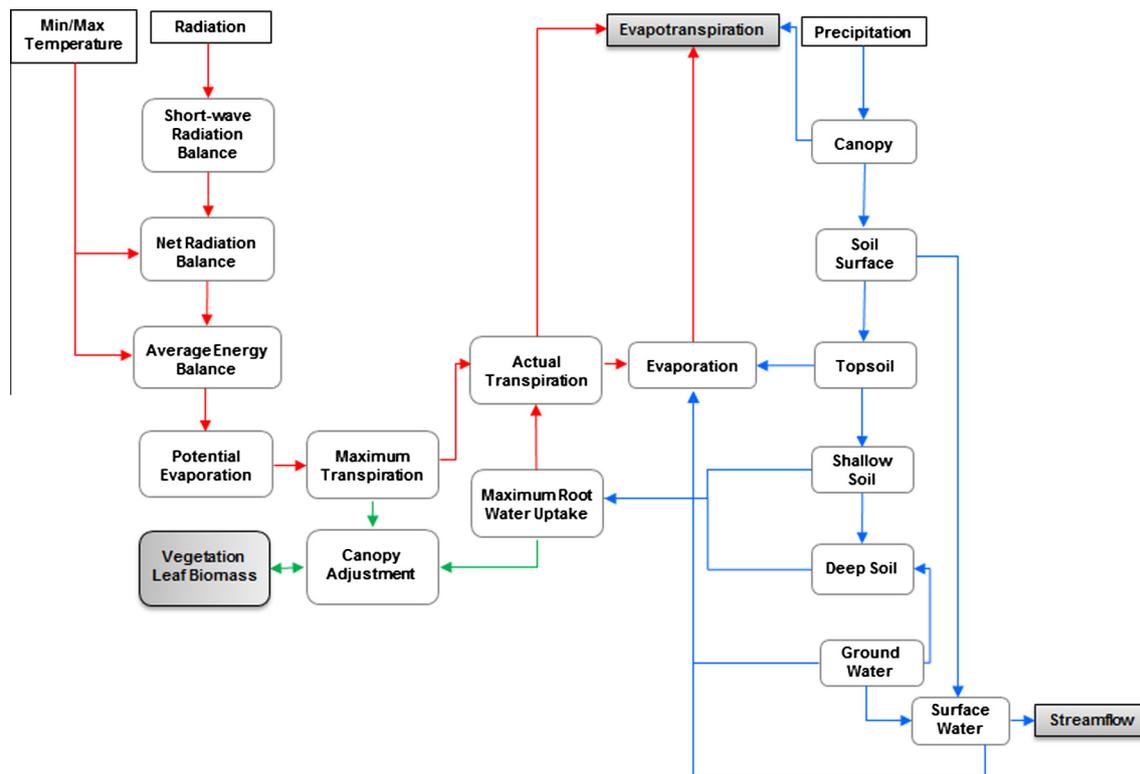


Fig. 2. Schematic diagram of AWRA-L single HRU model (modified from Van Dijk (2010)).

satellite data (<http://ozflux.org.au/monitoringsites/wombat/index.html>). Streamflow records for other catchments were downloaded from respective state water monitoring portal (<https://www.dnrm.qld.gov.au/water/water-monitoring-and-data/portal>, <http://real-timedata.water.nsw.gov.au/water.stm>, <http://data.water.vic.gov.au/monitoring.htm>).

2.4. Sensitivity analysis

Our simplified AWRA-L model has 34 parameters, of which, 26 are land-cover-specific parameters. In order to reduce the number of calibration parameters, a subset of parameters most sensitive to ET, SM and streamflow is chosen using the variance-based Sobol' sensitivity. In this algorithm total variance of model response is disaggregated to the contributions from individual parameters and parameter interactions (Sobol', 1990). Parameter sensitivity and the interaction sensitivity are normalized by the total output variance to calculate a sensitivity coefficient between 0 and 1.

In this study, sensitivity analysis for ET, SM and streamflow are performed separately for all the parameters. *RMSD* and the linear correlation coefficient (*R*) are used as the evaluation criteria for checking the contribution of each parameter variance on main output, as they constitute the objective functions tested in this study. The parameters to be calibrated are chosen based on the overall parameter sensitivities to ET, SM and streamflow. Upper and lower bounds of the parameters are adopted from Van Dijk (2010) and continuous uniform distributions of parameters within the bounds are sampled to test the sensitivity. Sensitivity tests are conducted for the period 2003–2007.

2.5. Model calibration

The present study investigates the utility of RS ET and SM to calibrate model in the absence of streamflow measurement. Calibration is conducted on daily model run of AWRA-L for the period 2003–2007. Multiple objective functions are defined by combining *RMSD* and *R* in ET and SM denoted by *a*, *b*, *c* and *d* as:

$$a = NRMSD_{ET}; b = 1 - R_{ET}; c = 1 - R_{SM}; d = NRMSD_{SM} \quad (2)$$

where $NRMSD_{ET}$ and $NRMSD_{SM}$ represent the normalized root mean square deviation between observed and simulated ET and SM. *RMSD* is normalized by dividing it by the range of observed values as:

$$NRMSD = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (V_{obs} - V_{sim})^2}}{(\max(V_{obs}) - \min(V_{obs}))} \quad (3)$$

where V_{obs} and V_{sim} are the observed and simulated variables for the time period *n*. Due to the different vertical support and inherent systematic bias existing between the model (AWRA-L) and satellite (AMSR-E) SM (Reichle and Koster, 2004), AWRA-L SM is rescaled using a variance matching method (Brocca et al., 2010) to the observed space before calculating *RMSD*. R_{ET} and R_{SM} are the linear correlation coefficients for ET and SM, respectively. We have set up 15 objective functions using the individual error metrics *a–d* and all the possible combinations of them. The individual metrics, *a–d*, vary between 0 and 1.

The SCE-UA global optimization method is adopted to derive the optimal set of parameters for AWRA-L. In SCE-UA, calibration starts with a population of points sampled randomly from a feasible parameter space. The points are divided into different complexes and each complex is updated to evolve through a statistical process called simplex. Shuffling of the population and reassigning of the points are carried out periodically. As the analysis progresses, the entire population try to converge towards the global optimum (Duan et al., 1992). A total of 15 different calibrations are repeated with 21 different initial parameter sets. Mean

and standard deviation of the calibrated parameters are used to summarize the final parameters. Monthly surface and subsurface runoffs generated using calibrated parameters are compared with the monthly observed streamflow for evaluation. Calibrated model predictions are also compared with predictions of the original version of AWRA-L and our simplified version calibrated with streamflow which serve as benchmarks.

3. Results

3.1. Sensitivity analysis

The Sobol' parameter sensitivity to streamflow, ET and SM without any objective function and in terms of *RMSD* and the linear correlation is shown in Fig. 3 for two grid cells in Loddon Catchment. A subset of nine parameters showed higher sensitivity to at least one of those three outputs. The nine parameters (see Table 2) include four land cover specific parameters (*Tgrow*, *Tsenc*, *hveg* and *Ud0*), two parameters defining soil water drainage (*beta* and *FdrainFC*), one parameter representing soil evaporation (*FsoilEmax*), interception (*S_sls*) and transpiration (*Vc*). Land cover specific parameters such as *Tgrow*, *Tsenc* and *hveg* showed large influence on ET and SM predictions whereas they showed negligible influence to streamflow. Streamflow was sensitive to *Ud0* mainly in grid cells with deep-rooted vegetation. In order to reduce the complexity of calibration, we fixed four vegetation parameters based on the information available from vegetation maps and some SCE-UA calibrations done with all nine parameters. After the nine-parameter calibration, it was found that the optimized values of the four vegetation parameters can be logically in disagreement with each other, due to reciprocity between vegetation parameters. Thus, the parameter *hveg* is fixed using the global height of vegetation map of Simard et al. (2011). *Tgrow*, *Tsenc* and *Ud0* are also fixed referring to the optimized values of the original model (Van Dijk, 2010).

The five parameters selected for calibration control main components of model water balance such as soil water drainage, soil evaporation, interception loss and root water uptake. Parameters *beta* and *FdrainFC* effectively play the opposite roles in the soil water process: low value of *beta* increases the groundwater drainage whereas low value of *FdrainFC* decreases it. Soil evaporation is constrained by *FsoilEmax*. Large interception loss is associated with high *S_sls* while the transpiration loss is parameterised by *Vc*. Both ET and SM were sensitive to most parameters; however streamflow exhibited noticeable sensitivity to only *beta* and *S_sls*. Since SM and ET are sensitive to similar set of parameters, calibration using only SM or ET can impart positive or negative effect on the other variable.

3.2. Model calibration

The calibrated model predictions at the Loddon River catchment were compared with the predictions given by the original AWRA-L version 0.5 ('control case 1' marked by the dotted lines) (Van Dijk, 2010) and with predictions of the simplified model calibrated with streamflow for the calibration period 2003–2007. In order to understand calibration results better, we evaluated the relative performance of the calibrated model using the 15 objective functions against the control case. A calibration yielding predictions with low *RMSD* and high correlation is considered ideal. The equation for relative performance (\check{r}) of *RMSD* and *R* was generated in such way that positive deviation from control case represents perfect model, which helps visualize the results.

$$\check{r}_{RMSD} (\%) = \frac{(RMSD_{Control Case} - RMSD_{Simplified AWRA-L})}{RMSD_{Control Case}} * 100 \quad (4)$$

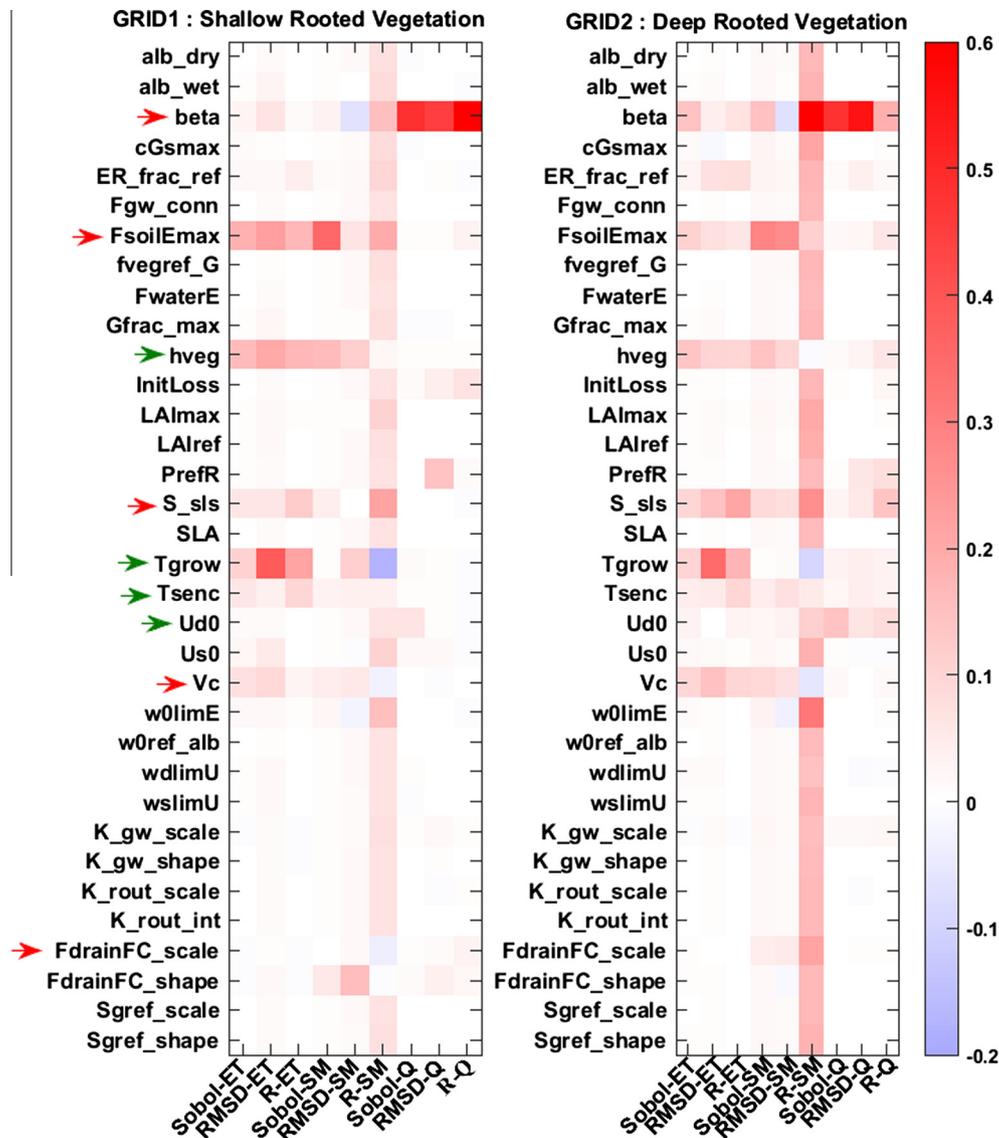


Fig. 3. Sensitivity test results from Sobol's algorithm without any objective function and with objective functions such as *RMSD* and linear correlation for ET, SM and streamflow. The red arrow in figure represents the parameters selected for calibration. The green arrow represents the parameters fixed with available information. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
List of sensitive parameters. First five parameters are selected for calibration.

Parameters	Description	Range
<i>beta</i>	Coefficient describing rate of hydraulic conductivity increase with water content	1–14
<i>FsoilE_{max}</i>	Soil evaporation scaling factor when soil water supply is not limiting evaporation	0.2–1
<i>FdrainFC_{scale}</i>	Drainage fraction at field capacity scaling factor	0.2–5
<i>S_{sls}</i>	Specific canopy rainfall storage capacity per unit leaf area	0.03–0.8
<i>Vc</i>	Vegetation photosynthetic capacity index per unit canopy cover	0.05–1
<i>Ud0</i>	Maximum root water uptake rates from deep soil	0.1–7
<i>Tgrow</i>	Characteristic time scale for vegetation growth towards equilibrium	20–1000
<i>Tsenc</i>	Characteristic time scale for vegetation senescence towards equilibrium	10–200
<i>hveg</i>	Height of the vegetation canopy	0.1–50

$$\check{Y}_R (\%) = \frac{R_{\text{Simplified AWRA-L}} - R_{\text{Control Case}}}{R_{\text{Control Case}}} * 100 \quad (5)$$

Results of RS based calibration are compared with AWRA-L version 0.5 and predictions of simplified model calibrated with streamflow in Sections 3.2.1 and 3.2.2.

3.2.1. Control Case 1 – AWRA-L model version 0.5

Fig. 4 presents the scatter plot between *RMSD* and correlation of streamflow, ET and SM. Each plus sign in the figure corresponds to the mean *RMSD* and *R* for the parameters calibrated with an arbitrary choice of 21 different initial search positions. For most objective functions tested, improvement (or deterioration) in *RMSD* in streamflow (*RMSD_Q*) was paired with deterioration

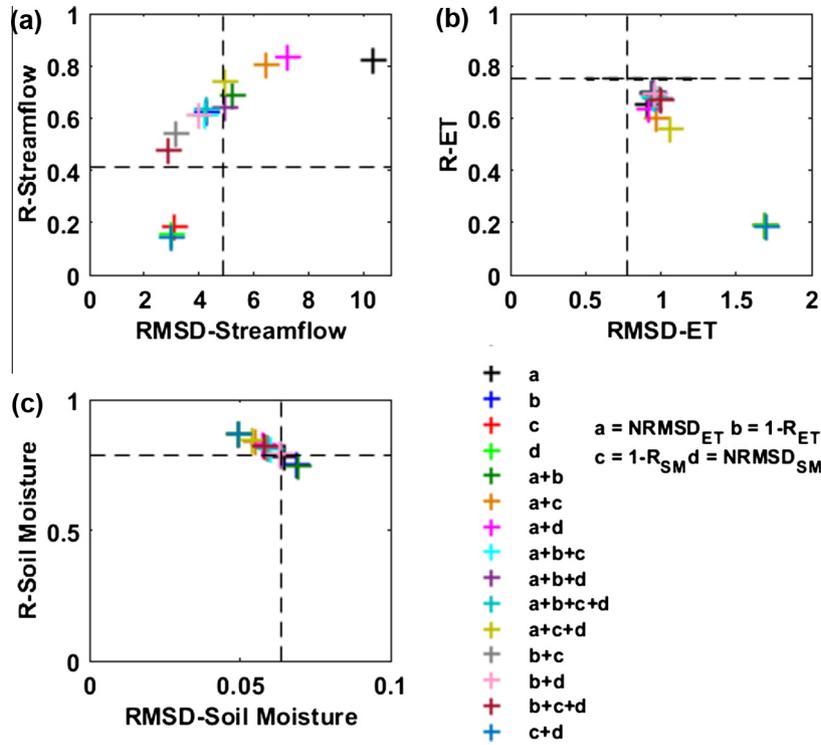


Fig. 4. Scatter plot of *RMSD* and *R* between observed and calibrated model predictions of (a) streamflow, (b) ET and (c) SM. The black dotted line in the graph represents the *RMSD* and *R* of respective variables given by optimized parameters of original AWRA-L model version 0.5.

(or improvement) in correlation (R_Q). Nevertheless, 7 out of 15 objective functions performed better than the original model (Fig. 4). Ideally, the best objective function is the one producing small *RMSD* and high correlation against observed streamflow compared with the control case ('control case' marked by the dotted lines), falling in the upper left quadrants of Fig. 4. The objective function $NRMSD_{ET}$ (a) or $NRMSD_{ET} + NRMSD_{SM}$ (a + d) yielded high correlation but large *RMSD* values for streamflow. Considering the direct contribution of ET to the catchment water balance, this may indicate biases (mean annual or seasonal) in the RS ET. Objective functions examining correlation and *RMSD* of SM give good results in terms of *RMSD* but exhibited the lowest correlation. Relative performance of calibrated models for prediction of streamflow, ET and SM from the control case is shown in Fig. 5.

With regard to ET predictions (Fig. 5), no objective function generated improved ET in terms of both *RMSD* and correlation compared to the ET predictions of original model. Even though the original AWRA-L model was designed to provide a good estimate of streamflow, the parameterisation of the model resulted in good ET predictions as well. This gives an insight to the limitation of the simplified single-HRU model. In the original model, each grid cell can have two HRUs defined by their tree cover fraction (i.e., deep-rooted or shallow-rooted vegetation). The simplified model, however, has one HRU per grid cell, determined by tree cover fraction (>50% tree cover is defined as deep-rooted vegetation). During testing we observed that, unlike SM and streamflow, ET was mainly sensitive to vegetation parameters such as T_{grow} , T_{senc} and $hveg$ (Fig. 3). It is therefore likely that the poor prediction of ET is due to the single-HRU model's limited representation of spatial heterogeneity over the larger grid cells ($0.25^\circ \times 0.25^\circ$). SM predictions of the simplified model were similar to those of the control case. All objective functions, other than the exclusive ET-based objective functions, showed some improvement from the control case, while ET-based calibrations showed minor deterioration.

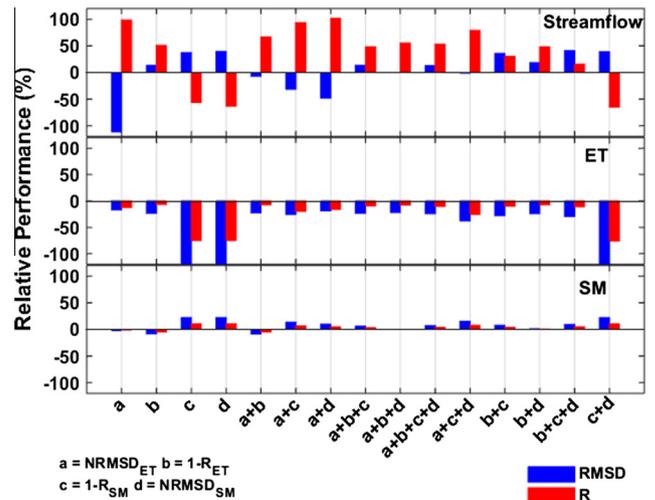


Fig. 5. Relative performance of *RMSD* and *R* between observed and calibrated model predictions of streamflow (top panel), ET (middle panel) and SM (bottom panel) with respect to those of original AWRA-L model version 0.5. Positive values for both *RMSD* and *R* represent improvement.

Solely SM-based calibration improved $RMSD_{SM}$ by 22% and R_{SM} by 10%.

Objective function composed of $NRMSD_{ET}$ or $1 - R_{ET}$ (or both) resulted in the best *RMSD* and correlations for ET predictions while the results for SM predictions remained relatively unchanged for all the objective functions. Objective function $NRMSD_{ET}$ overestimated the streamflow mainly during high-flow periods (Fig. 6a). Calibration reduced the *RMSD* between modelled and observed ET by increasing the soil water drainage to groundwater with a high value of $F_{drainFC}$, the drainage fraction. Low value of $F_{soilE-max}$ reduced the soil evaporation to minimum (Table 3). Finally

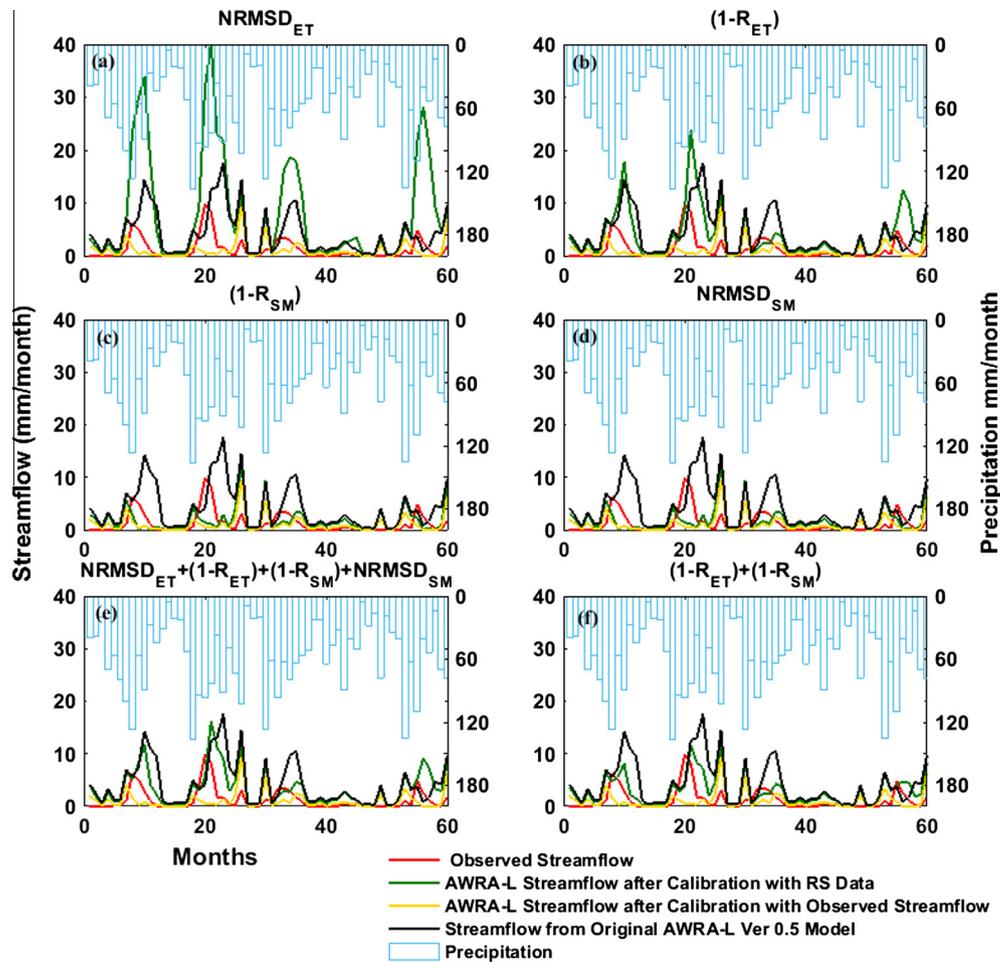


Fig. 6. Streamflow predictions for different objective functions. (a) $NRMSD_{ET}$, (b) $1 - R_{ET}$, (c) $1 - R_{SM}$, (d) $NRMSD_{SM}$, (e) $NRMSD_{ET} + (1 - R_{ET}) + (1 - R_{SM}) + NRMSD_{SM}$ and (f) $(1 - R_{ET}) + (1 - R_{SM})$.

Table 3
Median of optimized parameter values for different objective functions.

Objective functions	Grid cells/HRU	Beta	FsoilEmax	FdrainFC_scale	S_sls	Vc
Streamflow Calibration	–	13.90	0.2	4.29	0.78	0.05
AWRA-L Version 0.5	HRU 1	4.5	0.2	0.0685	0.1	0.35
	HRU 2	4.5	0.5	0.0685	0.1	0.65
$NRMSD_{ET}$	Grid 1	6.75	0.22	3.81	0.04	0.05
	Grid 2	7.52	0.20	2.87	0.03	0.28
$1 - R_{ET}$	Grid 1	6.69	0.20	4.19	0.03	0.13
	Grid 2	13.95	0.20	2.72	0.03	0.56
$1 - R_{SM}$	Grid 1	13.80	0.96	0.20	0.37	0.99
	Grid 2	13.36	0.83	0.23	0.61	0.97
$NRMSD_{SM}$	Grid 1	13.55	0.96	0.20	0.37	0.98
	Grid 2	13.05	0.87	0.23	0.62	0.97
$NRMSD_{ET} + (1 - R_{ET}) + (1 - R_{SM}) + NRMSD_{SM}$	Grid 1	1.10	0.25	0.30	0.03	0.10
	Grid 2	11.60	0.35	0.35	0.03	0.33
$(1 - R_{ET}) + (1 - R_{SM})$	Grid 1	1.09	0.25	0.23	0.03	0.11
	Grid 2	13.30	0.38	0.33	0.03	0.41

high groundwater discharge and surface runoff ended in a highly overestimated streamflow. However, this calibration resulted in highest correlation to streamflow. In comparison to this, calibration with only R_{ET} (Fig. 6b) improved the $RMSD_Q$ by increasing the transpiration loss (high value of Vc). Calibrations that combine correlation and $NRMSD_{ET}$ (a + b) enhanced R_Q and provided $RMSD$ of almost equal value to the original AWRA-L model.

Three of 15 objective functions (c, d and c + d) are exclusively based on SM and they resulted in similar performance in ET and streamflow predictions: namely, large underperformance of ET prediction, degraded R_Q , and improved $RMSD_Q$. Calibration with SM enhanced the $RMSD$ and R of top-layer SM by increasing the interception loss with a high value of s_sls (canopy storage capacity for unit leaf area), which reduced surface runoff. Most water

reaching the surface and soil was lost as ET to atmosphere due to high $F_{soilEmax}$ (maximum soil evaporation fraction) and V_c (photosynthetic capacity index). High β and small $F_{drainFC}$ decreased groundwater drainage and thereby reduced the groundwater discharge. The low surface runoff and groundwater discharge resulted in streamflow predictions with very low $RMSD$ (2.95 mm). Unlike other calibrations, objective functions including either $RMSD_{SM}$ or R_{SM} ended in poor R_Q (e.g., $R_Q = 0.18$ for c). High amount of transmission losses reduced or abandoned the groundwater contribution to stream and quick surface runoff becomes the sole conferrer for streamflow giving an impression of lag in streamflow predictions (Fig. 6c and d). This also indicates that slow groundwater flows play a major role in matching peak runoff.

Calibration results clearly demonstrate how each parameter adapts to redistribute total water among different components of model water balance based on the objective function (Table 3). Vegetation plays an important role in governing interception, evaporation and soil water drainage (Andersen et al., 2002). The high value of s_{sls} and V_c in grid 2 comparing to grid 1 for most of the optimization experiments contributed to high interception and transpiration loss in grid with deep-rooted vegetation. Reciprocity between the parameters was also noted during calibration. β and $F_{drainFC}$ can adjust themselves to control the groundwater flow, and both $F_{soilEmax}$ and V_c control the total ET output.

The overestimated runoff in calibration using objective functions a and $a + d$ was due to increased baseflow as a result of high groundwater drainage (high value of $F_{drainFC}$). When $RMSD$ decreased, the baseflow component also decreased due to enhanced transmission losses. Some cases, (e.g., c, d, c + d, and b + c + d) exhibited shifted runoff peaks through high transmission loss and zero baseflow, resulting in decreased correlation. The calibration parameters behaved in the opposite ways in exclusive ET and SM based calibrations. None of the calibration experiments resulted in improvements in both $RMSD$ and R for all the three variables in comparison with control case. Therefore, the choice of objective function is subjective depending on the final goal of calibration. If time of peak flow/correlation in runoff is important, a or $a + d$ will be preferred. Calibration with $b + c + d$ or $b + c$ or purely SM-based calibrations will be the best option when quantity of water is important over peaks.

Nevertheless, some objective functions were providing improved SM and streamflow predictions and ET close to the control case. The objective function $b + c$ considering correlation of ET and SM, gave about 36% improvement in $RMSD_Q$ and 30% improvement in R_Q with slight improvement (3–7%) in both $RMSD_{SM}$ and R_{SM} (Fig. 5). Similarly the objective function $a + b + c + d$ provided good prediction of streamflow in terms both $RMSD$ (13% improvement) and correlation (52% improvement) with about 5% increase in $RMSD_{SM}$ and R_{SM} . For both calibrations correlation of evapotranspiration remained almost unchanged with 25–30% deterioration in $RMSD$ (Fig. 5). This can be considered as one of the best objective function as it provides good improvement in streamflow without major diminution in ET.

3.2.2. Control Case 2 – simplified AWRA-L model calibrated with streamflow

The objective function ‘1-NSE’ (Nash–Sutcliffe Efficiency) was used to calibrate simplified AWRA-L against streamflow. Relative performance of the AWRA-L model calibrated with RS ET and SM in comparison with prediction of the model calibrated with streamflow is shown in Fig. 7. Almost all the objective functions provided very high correlation with streamflow in comparison with the control case, however no objective function resulted in a reduced $RMSD$. Since the objective function in control case use NSE , it is obvious that it will improve mainly $RMSD$

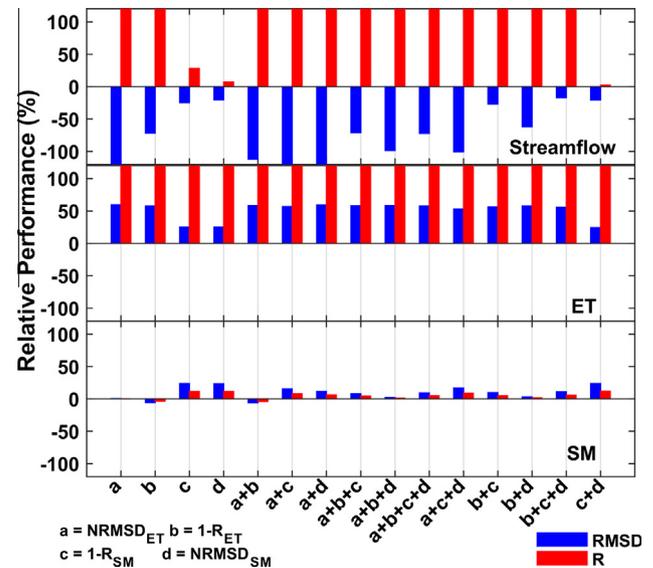


Fig. 7. Relative performance of $RMSD$ and R between observed and calibrated model predictions of streamflow (top panel), ET (middle panel) and SM (bottom panel) with respect to predictions of the simplified AWRA-L model calibrated with streamflow. Positive values for both $RMSD$ and R represent improvement.

(2.47 mm/month) than correlation (0.14) of streamflow. Thus, the streamflow prediction of the control case looked similar to prediction of streamflow from SM based calibrations (Fig. 6c and d).

Overall, improved correlation in ET coincides with increase in correlation of streamflow. $NRMSD_{ET}$ -based objective function can be useful when the timing of peak flows is an important criterion for calibration. ET is more relevant for monthly estimates of streamflow as it has direct impact on baseflow. Accurate specification of the evaporative loss term is essential to get accurate discharge prediction. Studies have already suggested that proper pre-processing of derived satellite products is critical for the efficacy of calibration (Van Dijk and Renzullo, 2011). Although the calibration with $NRMSD_{ET}$ produced best prediction of evapotranspiration, comparison of ground observation of ET with CMRSET in Fig. 11 reveals that CMRSET underestimated ET in the study area over more than 50% of the period of 2010–2012. Guerschman et al. (2009) suggest an error in CMRSET of 10–20%. It appears that the overestimated streamflow by calibrated model is due to the underestimated ET used for calibration. In Section 3.3.1, we perform a synthetic experiment to investigate the impact of biases (either positive or negative) in RS ET on calibrated streamflow predictions.

Results have shown that the calibration experiments do not result in improved ET prediction compared with the control case 1. This may be the result of modification of the AWRA-L model from two to one HRU. Uncertainties associated with the input and calibration data can also result in erroneous prediction. To better understand the effect of uncertainty associated with model structure and input forcing, a set of synthetic experiments was conducted, discussed in detailed in Section 3.3.2.

3.2.3. Application in other catchments

To assess how the AWRA-L model calibrated by RS data performs elsewhere, we applied the same calibration scheme to ten small catchments (Fig. 1) in eastern Australia. $RMSD$ and R of resulted streamflow predictions are compared with control case 1 (AWRA-L model version 0.5) and control case 2 (simplified AWRA-L model calibrated with streamflow). The catchments 112102, 116008, 141008 have relatively high average runoff as they have a daily average streamflow equal to or greater than

1 mm/day during the calibration period. Catchments 145107, 211008 and 219016 are treated as catchments with medium average runoff (daily average flow 0.3–1 mm/day), and the remaining catchments (212040, 226415, 405251 and 410156) as low average runoff catchments (daily average flow lower than 0.3 mm/day) (Table 1).

Fig. 8 presents the relative performance of correlation and *RMSD* of streamflow with respect to the control case 1 for different catchments. For catchments with high average runoff, calibration with *NRMSD_{ET}* resulted in the best streamflow predictions. With the exception of catchment 145107, ET-based calibration provided improved or equally good predictions of streamflow for medium flow catchments, and there are other multi-objective calibrations that give better streamflow predictions as well (e.g., *b + c*, *a + c + d*). Low flow catchments behaved in similar manner to the Loddon River catchment, i.e., improvement in correlation of streamflow with decrement in *RMSD* for objective function *NRMSD_{ET}*. Some other objective functions (e.g., *a + c + d*, *b + c*, *b + c + d*) performed better than *NRMSD_{ET}* calibration in low flow catchments.

Relative change in performance of *RMSD* and *R* between the observed and calibrated model predictions of streamflow with respect to *RMSD* and *R* of simplified AWRA-L model calibrated with streamflow is shown in Fig. 9. Streamflow predictions for the objective function *NRMSD-ET* provided similar prediction as the control case 2 in the high flow catchments (0–5% decrement). As seen in the control case 1 except for catchment 145107, *NRMSD_{ET}* was the best for streamflow predictions in medium flow catchments as well. In case of low flow catchments, multi-objective calibration performs better than *NRMSD_{ET}* calibration.

Streamflow predictions for the objective function *NRMSD_{ET}* for all ten catchments are shown in Fig. 10. Calibrated model predictions and the control case 2 were very similar for high flow catchments and medium flow catchments. Control case 1 was slightly underestimated in catchments 116008 and 141008. In the catchment 145107, calibrated model slightly overestimated streamflow. Streamflow was always overestimated in the low flow catchments, but showed good correlation with the observed streamflow and this result is similar to the result obtained in the Loddon River catchment (average daily runoff of 0.04 mm/day). As seen in the Loddon River catchment, underestimated RS ET can be a possible reason for overestimation of streamflow in low average flow catchments.

3.3. Synthetic experiments

3.3.1. Synthetic ET data

A synthetic experiment was set up to understand the influence of bias in observed ET on streamflow estimation. For the experiment, synthetic truth optimum parameter set (θ^{truth}) was generated and run the AWRA-L model. The predicted streamflow and evapotranspiration are considered ‘truth’ (Eq. (6)). Synthetic observations of ET (ET^{obs}) to calibrate the model have been generated by perturbing truth predictions of ET (ET^{truth}) considering error model for ET. From Fig. 11, presenting the error between Wombat flux tower ET with CMRSET for the period from 2010 to 2012, it is evident that error model is not following any particular pattern. A synthetic set of ET observations was generated by adding bias (b_t) and white noise (ω_t) to the ‘truth’ ET. A maximum bias of

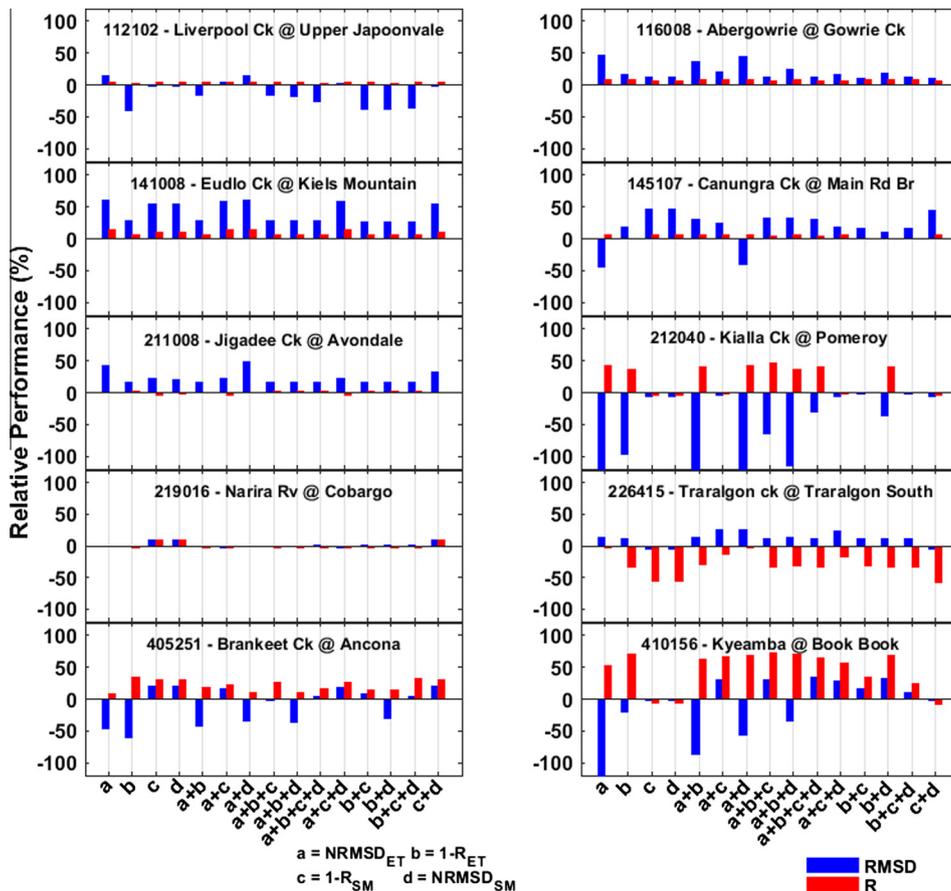


Fig. 8. Relative performance of *RMSD* and *R* between observed and calibrated model predictions of streamflow with respect to *RMSD* and *R* of original AWRA-L model version 0.5 for ten catchments. Positive values for both *RMSD* and *R* represent improvement.

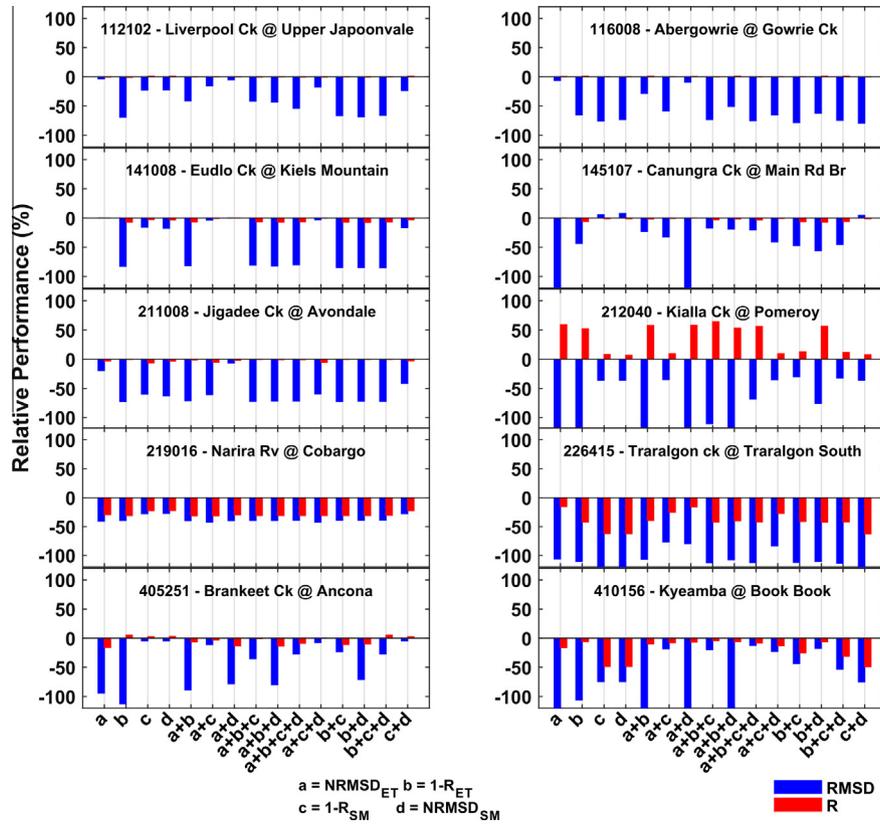


Fig. 9. Relative performance of *RMSD* and *R* between observed and calibrated model predictions of streamflow with respect to *RMSD* and *R* of simplified AWRA-L model calibrated with streamflow for ten catchments. Positive values for both *RMSD* and *R* represent improvement.

10% of average synthetic truth ET and error term follows a normal distribution with zero mean and standard deviation of ground observed ET was added.

$$\widehat{ET}_t^{truth} = h(X_t, \theta^{truth}) \quad (6)$$

$$ET_t^{obs} = \widehat{ET}_t^{truth} + b_t + \omega_t \quad (7)$$

Model was calibrated by each realization of synthetically generated ET observations using $NRMSD_{ET}$ as objective function. Monthly streamflow prediction for each calibration was plotted and analyzed against the synthetic truth predictions (Fig. 12). As expected, a negative bias in ET results in overestimated streamflow whereas positive ET bias in underestimated streamflow and was more visible during the peak-flow periods. It was noted that small changes in daily ET has notable effect on streamflow predictions. Bias in streamflow was high when a negative bias is added to ET than positive bias, which illustrates that calibration with an underestimated ET can yield greater bias in streamflow (Fig. 13).

The results of synthetic experiments are consistent with the real calibration experiment with $NRMSD_{ET}$ (Fig. 6a). Underestimated satellite ET is the key reason for overestimation of streamflow. The predicted streamflow is higher than the observation only during heavy flow periods in real experiments as well. Thus, from this experiment we can conclude that a high quality ET product can improve the predictions of streamflow. This strong dependence between ET and streamflow give insight to new research area, in which observed streamflow can be used for correcting RS ET.

3.3.2. Forcing and model structural error parameterization

Another set of synthetic experiments were performed to understand the effect of forcing error and model structural error on

calibrated AWRA-L results. The errors in model predictions mainly arise from input forcing, model parameters, initial state variables and model structure. These errors are usually quantified by incorporating unbiased synthetic noise to forcing variables, models states or parameters. Lognormal multiplicative error model for rainfall is widely used to model forcing uncertainty (Alvarez-Garreton et al., 2015; Crow et al., 2011; DeChant and Moradkhani, 2012; Li et al., 2014). In this study we perturbed rainfall data to quantify input uncertainty and assumed that error due to model structure and model parameters are accumulated in SM state as in most of the data assimilation studies (Alvarez-Garreton et al., 2015; Li et al., 2014; Ryu et al., 2009).

We adopted multiplicative error model for rainfall (P) data.

$$P' = \varepsilon_p P \quad (8)$$

$$\varepsilon_p \sim \ln N(1, \sigma_p^2) \quad (9)$$

where ε_p follows a lognormal distribution with mean μ and standard deviation σ_p . μ was set be 1 to create an unbiased rainfall and σ_p was empirically defined as 0.25 (25% error) as in study by DeChant and Moradkhani (2012). For each day top soil moisture state was updated by adding zero mean gaussian distribution with standard deviation (σ_s) 3 mm/day

$$\varepsilon_s \sim N(0, \sigma_s^2) \quad (10)$$

The model was calibrated against synthetic truth ET (ET^{truth}) and SM (SM^{truth}) generated using synthetic truth parameter set employed in Section 3.3.1. The model has tested for all the 15 objective functions set up for this study. Fig. 14 depicts the *RMSD* and *R* between calibrated model predictions and synthetic truths. *RMSD* for streamflow varies around 2.5 (mm/month) with high correlation (close to 0.8) for all the objective functions except for $a + d$.

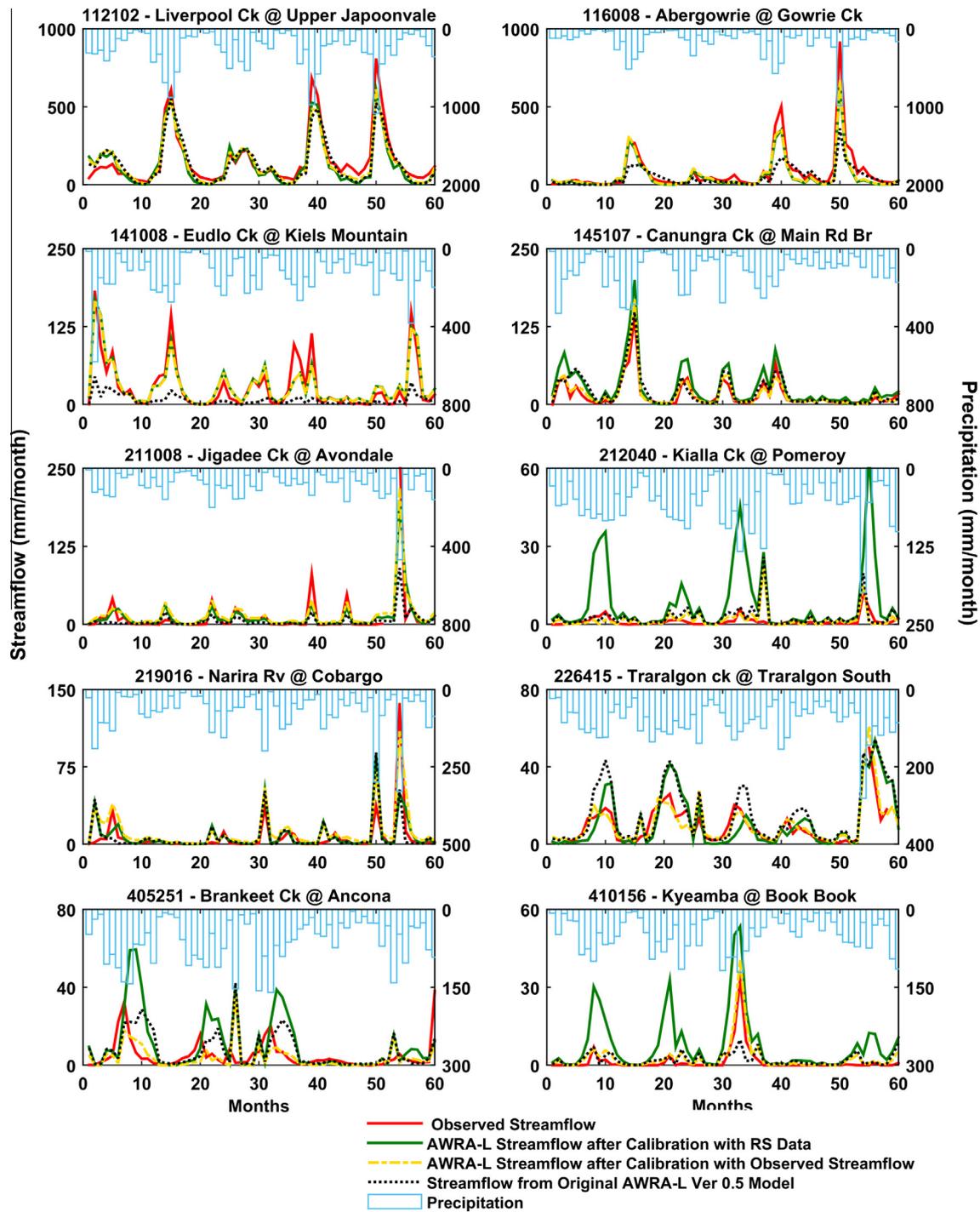


Fig. 10. Streamflow predictions for the objective function $NRMSD_{ET}$ in ten small catchments.

As seen in real calibration experiments, ET predictions improved for objective functions formed on $RMSD_{ET}$ and R_{ET} . All the experiments resulted in increased R_{SM} , however, lowest $RMSD_{SM}$ was attained only for objective functions using SM. These results revealed that addition of forcing error and model structural error to calibration affects streamflow in similar manner for most of the experiments whereas, quality of ET and SM predictions varies based on the objective function used.

Comparing the fifteen different objective functions, the objective functions using $RMSD$ and correlation in ET (e.g., a , b , $a + b$, $a + b + d$) performs better than the other objective functions after

inclusion of input and model errors to calibration. Besides the improvement in streamflow, those calibrations resulted in improved ET predictions. This indicates that calibration based on ET is less vulnerable to the errors in calibration. Study also affirms that ET is more pertinent in providing accurate monthly or long term streamflow predictions.

4. Discussion

Calibrating hydrologic models with ET and SM has important implications for streamflow prediction in ungauged catchments.

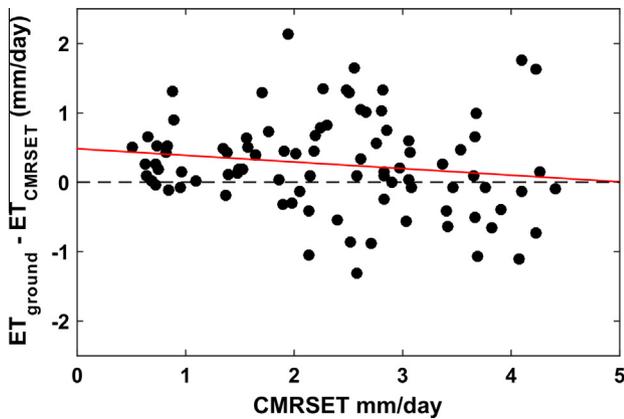


Fig. 11. Difference between flux tower ET and the CMRSET.

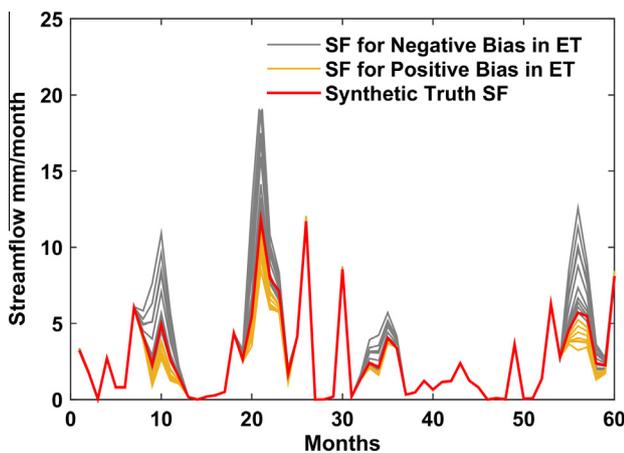


Fig. 12. Streamflow (SF) predictions for synthetic ET observations.

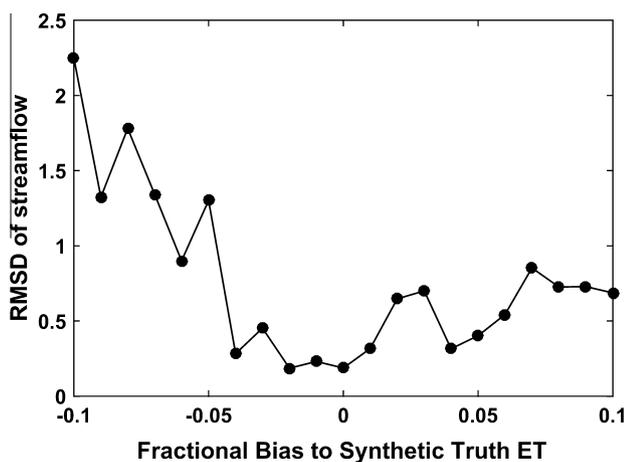


Fig. 13. RMSD between synthetic truth streamflow and streamflow predictions of the model after calibration to synthetic ET observations.

A widely used method to predict streamflow at ungauged or sparsely gauged catchments is the regionalization of parameters or flow indices (Abdulla and Lettenmaier, 1997; Jakeman et al., 1992; Parajka et al., 2007; Post et al., 1998; Sefton and Howarth, 1998; Viney et al., 2009; Wagener and Wheater, 2006; Wagener et al., 2004; Zhang et al., 2011; Yadav et al., 2007). Many studies described the limitations of parameter regionalization approaches. Limitations include the inappropriateness of transferring some

watershed physical characteristics between potentially unique watersheds (Wagener and Montanari, 2011) and the influence of model structural uncertainty on the optimality of parameters spreading to other watersheds (Wagener and Wheater, 2006). As an alternative to the regionalization, model calibration proposed in this study can provide the optimized parameters for ungauged catchment using RS observation from the same catchment. In control case 1 above, regionalized parameters for the original AWRA-L version 0.5 over Australia are used. Our results demonstrate, for most of the catchments tested, improved streamflow predictions for more than one objective functions based on RS ET and SM in comparison with control case 1.

It is also shown that quality ET for calibration is critical to produce reliable streamflow predictions. ET and streamflow comprise the major output fluxes in the steady-state rainfall–runoff processes. For a given rainfall, increase or decrease in either flux will be reflected on the other. Therefore quality of ET products plays a major role in obtaining accurate streamflow predictions. The accuracy of the satellite products can vary greatly across large regions due to climatic variability and heterogeneity in land surface conditions, such as topography and vegetation cover (Barrett and Renzullo, 2009; Campo et al., 2006; Parajka et al., 2009). For example, previous studies have reported ET is mainly overestimated in wet catchments or wet periods and underestimated in dry catchments or dry periods (Mu et al., 2011). Furthermore, MODIS actual ET product may fail to capture spatial heterogeneity when meteorological input features coarse spatial resolution (Long et al., 2014; Yang et al., 2015). Synthetic experiments conducted in Section 3.2.1 show that underestimated ET used for calibration has more impact on streamflow predictions than overestimated ET. The result implies that the decrement in streamflow due to overestimated ET in wet catchments may not be as high as increment in streamflow in dry catchment due to the underestimated ET. This is consistent with the real calibration based on $NRMSD_{ET}$, high flow/wet catchments yields more accurate streamflow predictions whereas in low flow/dry catchments such as the Loddon River catchment calibration by underestimated ET results in significantly overestimated streamflow. Incorporation of spatially and temporally varying meteorological variables, particularly wind speed, would greatly improve ET estimation, leading to more accurate hydrological modelling (McVicar et al., 2012a, 2012b). In our investigation we adopted the AWRA-L model version 0.5 default day-time average wind speed value of 3.5 m/s across the Australian continent and for all time. This value is higher than that reported by McVicar et al. (2008), however we note that their value represents the average of day and night time wind observations.

Calibration by RS SM appears to work better in low flow catchments. Improvement in SM predictions is not significant as the simulated SM for control cases and AMSR-E SM are in good agreement. Due to the difference between the top surface soil layer of simplified AWRA-L model (5 cm) and the sensing depth of AMSR-E (1–2 cm), rescaling was applied to make them comparable. However, SM rescaling can potentially impart additional uncertainty to SM calibrations by amplifying SM retrieval errors (Draper et al., 2009). Ideally, calibration based on good estimates of ET in low flow catchments would result in more accurate prediction due to the direct reciprocity between ET and streamflow.

For multi-objective calibration it is possible that the optimal set of parameters lies on the Pareto front over varying climate conditions (Gupta et al., 1998). This possibility needs to be further investigated in the future. Lastly, since AWRA-L is a physically based landscape model, the proposed calibration scheme can be transferred to the other compatible hydrologic models. More studies should follow to examine the transferability of this scheme and to explore the use of different RS products for hydrological model calibration.

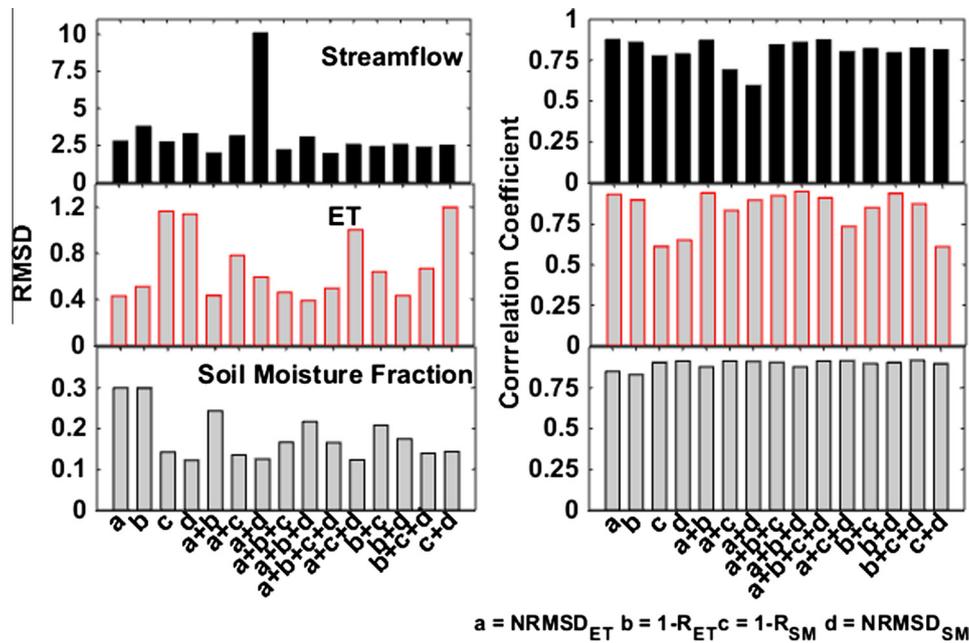


Fig. 14. RMSD and R of streamflow, ET and SM between synthetic truths and predictions of synthetically calibrated AWRA-L model.

5. Conclusion

The current study analyzed the ability of RS ET (CMRSET) and SM (AMSR-E) in calibrating a grid based hydrologic model, AWRA-L. Efficacy of the calibration was evaluated on the accuracy of streamflow predictions. A total of 15 objective functions considering RMSD and correlation of ET and SM were used to calibrate the model.

Most objective functions performed well in the catchments with medium to high average runoff. In particular, calibration based on $NRMSD_{ET}$ provided the best streamflow predictions at high flow catchments, which is as good as prediction from model calibrated with streamflow (control case 2) and better than predictions of the original AWRA-L model version 0.5 model (control case 1). In catchments with low average runoff, $NRMSD_{ET}$ based calibration exhibited best correlation to streamflow however RMSD was high. This is attributed to the underestimated RS ET in dry catchments. Other multi-objective calibrations that performed better in low flow catchments. The objective function $NRMSD$ of ET would be a better choice when timing of peak flow is important. Further comparison of CMRSET with flux tower ET in the dry Loddon River catchment showed that CMSRET was underestimated on most of the days in minimal amount. Synthetic experiments showed that, negatively biased (underestimated) ET has more detrimental effect on the model calibration and subsequent streamflow prediction. Synthetic calibration experiment showed that the ET-based calibration was more resistant to errors in calibration. Based on the synthetic and the real calibration experiments, we conclude that calibration with accurate ET measurements is a robust way to produce accurate streamflow predictions.

This study on the use of RS data for hydrologic model calibration has important implications to the catchments with sparse or no gauging. The calibration performed gives an insight to the scope of RS data in hydrological model calibration. Referring to the strong dependence between the streamflow and evapotranspiration streamflow can be utilized for satellite ET bias correction.

Acknowledgements

This research is financially supported by International Postgraduate Research Scholarships (IPRS), the Australian Postgraduate

Award (APA) and the Carlton Connect Initiative Fund Project 22514, representing collaboration between the Melbourne School of Engineering, The University of Melbourne, and the South Australian Government Department of Environment, Water and Natural Resources. We would like to acknowledge Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BoM) for providing data. We gratefully thank comments and ideas of Dr. Chun-Hsu Su that have significantly improved the quality of the paper. We thank the three anonymous reviewers, whose valuable comments/suggestions helped to improve this manuscript.

References

- Abdulla, F.A., Lettenmaier, D.P., 1997. Development of regional parameter estimation equations for a macroscale hydrologic model. *J. Hydrol.* 197 (1–4), 230–257.
- Alvarez-Garreton, C., Ryu, D., Western, A.W., Crow, W.T., Robertson, D.E., 2014. The impacts of assimilating satellite soil moisture into a rainfall–runoff model in a semi-arid catchment. *J. Hydrol.* 519, 2763–2774. <http://dx.doi.org/10.1016/j.jhydrol.2014.07.041>.
- Alvarez-Garreton, C., Ryu, D., Western, A.W., Su, C.H., Crow, W.T., Robertson, D.E., Leahy, C., 2015. Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: comparison between lumped and semi-distributed schemes. *Hydrol. Earth Syst. Sci.* 19 (4), 1659–1676. <http://dx.doi.org/10.5194/hess-19-1659-2015>.
- Andersen, J., Dybkjaer, G., Jensen, K.H., Refsgaard, J.C., Rasmussen, K., 2002. Use of remotely sensed precipitation and leaf area index in a distributed hydrological model. *J. Hydrol.* 264 (1–4), 34–50. [http://dx.doi.org/10.1016/S0022-1694\(02\)00046-X](http://dx.doi.org/10.1016/S0022-1694(02)00046-X).
- Asner, G.P., Scurlock, J.M.O., Hicke, J.A., 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecol. Biogeogr.* 12 (3), 191–205. <http://dx.doi.org/10.1046/j.1466-822X.2003.00026.x>.
- Atkinson, S., Sivapalan, M., Viney, N., Woods, R., 2003. Predicting space–time variability of hourly streamflow and the role of climate seasonality: Mahurangi Catchment, New Zealand. *Hydrol. Process.* 17 (11), 2171–2193.
- Barrett, D.J., Renzullo, L.J., 2009. On the efficacy of combining thermal and microwave satellite data as observational constraints for root-zone soil moisture estimation. *J. Hydrometeorol.* 10 (5), 1109–1127.
- Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.* 34 (20).
- Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci. Discuss.* 4 (2), 203–213.
- Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., Hasenauer, S., 2010. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrol. Earth Syst. Sci.* 14 (10), 1881–1893.

- Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., Hahn, S., 2012. Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall-runoff modeling. *IEEE Trans. Geosci. Remote Sens.* 50 (7), 2542–2555.
- Campo, L., Caparrini, F., Castelli, F., 2006. Use of multi-platform, multi-temporal remote-sensing data for calibration of a distributed hydrological model: an application in the Arno basin, Italy. *Hydrol. Process.* 20 (13), 2693–2712.
- Chen, J.M., Cihlar, J., 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sens. Environ.* 55 (2), 153–162.
- Cleugh, H.A., Leuning, R., Mu, Q., Running, S.W., 2007. Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sens. Environ.* 106 (3), 285–304.
- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37 (1), 35–46.
- Corbari, C., Mancini, M., 2013. Calibration and validation of a distributed energy-water balance model using satellite data of land surface temperature and ground discharge measurements. *J. Hydrometeorol.* 15 (1), 376–392. <http://dx.doi.org/10.1175/jhm-d-12-0173.1>.
- Crow, W., Van den Berg, M., 2010. An improved approach for estimating observation and model error parameters in soil moisture data assimilation. *Water Resour. Res.* 46 (12).
- Crow, W.T., Ryu, D., 2009. A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. *Hydrol. Earth Syst. Sci.* 13 (1), 1–16. <http://dx.doi.org/10.5194/hess-13-1-2009>.
- Crow, W.T., van den Berg, M.J., Huffman, G.J., Pellarin, T., 2011. Correcting rainfall using satellite-based surface soil moisture retrievals: the Soil Moisture Analysis Rainfall Tool (SMART). *Water Resour. Res.* 47 (8). <http://dx.doi.org/10.1029/2011WR010576> W08521.
- DeChant, C.M., Moradkhani, H., 2012. Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resour. Res.* 48 (4). <http://dx.doi.org/10.1029/2011WR011011>.
- Donohue, R.J., McVicar, T.R., Roderick, M.L., 2009. Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. *Global Change Biol.* 15 (4), 1025–1039. <http://dx.doi.org/10.1111/j.1365-2486.2008.01746.x>.
- 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. *Remote Sens. Environ.* 113 (4), 703–710. <http://dx.doi.org/10.1016/j.rse.2008.11.011>.
- Droogers, P., Immerzeel, W., Lorite, I., 2010. Estimating actual irrigation application by remotely sensed evapotranspiration observations. *Agric. Water Manage.* 97 (9), 1351–1359.
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28 (4), 1015–1031.
- García-Quijano, J.F., Barros, A.P., 2005. Incorporating canopy physiology into a hydrological model: photosynthesis, dynamic respiration, and stomatal sensitivity. *Ecol. Modell.* 185 (1), 29–49. <http://dx.doi.org/10.1016/j.ecolmod.2004.08.024>.
- 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. Hydrol.* 369 (1–2), 107–119. <http://dx.doi.org/10.1016/j.jhydrol.2009.02.013>.
- Gupta, P., Singh, R., Raghuvanshi, N., Dutta, S., Panigrahy, S., 2008. Effect of remotely sensed data on the performance of a distributed hydrological model: case study. *J. Hydrol. Eng.* 13 (10), 939–947.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resour. Res.* 34 (4), 751–763. <http://dx.doi.org/10.1029/97wr03495>.
- Han, E., Merwade, V., Heathman, G.C., 2012. Implementation of surface soil moisture data assimilation with watershed scale distributed hydrological model. *J. Hydrol.* 416–417, 98–117.
- Immerzeel, W.W., Droogers, P., 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol.* 349 (3–4), 411–424. <http://dx.doi.org/10.1016/j.jhydrol.2007.11.017>.
- Jakeman, A., Hornberger, G., Littlewood, I., Whitehead, P., Harvey, J., Bengala, K., 1992. A systematic approach to modelling the dynamic linkage of climate, physical catchment descriptors and hydrologic response components. *Math. Comput. Simul.* 33 (5), 359–366.
- Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.* 39 (8), 1729–1735.
- Koren, V., Smith, M., Duan, Q., 2003. Use of a priori parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. *Water Sci. Appl.* 6, 239–254.
- Kunnath Poovakka, A., Ryu, D., Renzullo, L., Pipunic, R.C., George, B., 2013. Calibration of land surface model using remotely sensed evapotranspiration and soil moisture predictions. *Proceedings of the 20th International Congress on Modelling and Simulation MODSIM2013*, December. <<http://www.mssanz.org.au/modsim2013/L19/poovakka.pdf>>.
- Kustas, W., Perry, E., Doraiswamy, P., Moran, M., 1994. Using satellite remote sensing to extrapolate evapotranspiration estimates in time and space over a semiarid rangeland basin. *Remote Sens. Environ.* 49 (3), 275–286.
- Leuning, R., Zhang, Y., Rajaud, A., Cleugh, H., Tu, K., 2008. A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation. *Water Resour. Res.* 44 (10).
- Li, Y., Ryu, D., Western, A.W., Wang, Q.J., Robertson, D.E., Crow, W.T., 2014. An integrated error parameter estimation and lag-aware data assimilation scheme for real-time flood forecasting. *J. Hydrol.* 519 (Part D), 2722–2736. <http://dx.doi.org/10.1016/j.jhydrol.2014.08.009>.
- Long, D., Longuevergne, L., Scanlon, B.R., 2014. Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. *Water Resour. Res.* 50 (2), 1131–1151. <http://dx.doi.org/10.1002/2013WR014581>.
- McVicar, T.R., Roderick, M.L., Donohue, R.J., Li, L.T., Van Niel, T.G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N.M., Mescherskaya, A.V., Kruger, A.C., Rehman, S., Dinpashoh, Y., 2012a. Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation. *J. Hydrol.* 416–417, 182–205. <http://dx.doi.org/10.1016/j.jhydrol.2011.10.024>.
- McVicar, T.R., Roderick, M.L., Donohue, R.J., Van Niel, T.G., 2012b. Less bluster ahead? Ecohydrological implications of global trends of terrestrial near-surface wind speeds. *Ecohydrology* 5 (4), 381–388. <http://dx.doi.org/10.1002/ecc.1298>.
- McVicar, T.R., Van Niel, T.G., Li, L.T., Roderick, M.L., Rayner, D.P., Ricciardulli, L., Donohue, R.J., 2008. Wind speed climatology and trends for Australia, 1975–2006: capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophys. Res. Lett.* 35 (20), L20403. <http://dx.doi.org/10.1029/2008GL035627>.
- Mohanty, B.P., 2013. Soil hydraulic property estimation using remote sensing: a review. *Vadose Zone J.* 12 (4). <http://dx.doi.org/10.2136/vzj2013.06.0100>.
- Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* 115 (8), 1781–1800. <http://dx.doi.org/10.1016/j.rse.2011.02.019>.
- Nemani, R.R., Running, S.W., 1989. Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *J. Appl. Meteorol.* 28 (4), 276–284.
- Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. *IEEE Trans. Geosci. Remote Sens.* 41 (2), 215–229.
- Ottlé, C., Vidal-Madjar, D., Girard, G., 1989. Remote sensing applications to hydrological modeling. *J. Hydrol.* 105 (3–4), 369–384. [http://dx.doi.org/10.1016/0022-1694\(89\)90114-5](http://dx.doi.org/10.1016/0022-1694(89)90114-5).
- Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res.* 113 (F1) F01002.
- Parajka, J., Blöschl, G., Merz, R., 2007. Regional calibration of catchment models: potential for ungauged catchments. *Water Resour. Res.* 43 (6).
- Parajka, J., Naeimi, V., Blöschl, G., Komma, J., 2009. Matching ERS scatterometer based soil moisture patterns with simulations of a conceptual dual layer hydrologic model over Austria. *Hydrol. Earth Syst. Sci.* 13 (2), 259–271.
- Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., Scipal, K., 2006. Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. *Hydrol. Earth Syst. Sci. Discuss.* 10 (3), 353–368.
- Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., 2001. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. *J. Hydrol.* 251 (1–2), 88–102.
- Post, D., Jones, J., Grant, G., 1998. An improved methodology for predicting the daily hydrologic response of ungauged catchments. *Environ. Modell. Softw.* 13 (3), 395–403.
- Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.* 31 (19). <http://dx.doi.org/10.1029/2004gl020938> L19501.
- Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA catchment land surface model. *Geophys. Res. Lett.* 32 (2).
- Renzullo, L.J., Barrett, D.J., Marks, A.S., Hill, M.J., Guerschman, J.P., Mu, Q., Running, S.W., 2008. Multi-sensor model-data fusion for estimation of hydrologic and energy flux parameters. *Remote Sens. Environ.* 112 (4), 1306–1319.
- Renzullo, L.J., van Dijk, A.I.J.M., Perraud, J.M., Collins, D., Henderson, B., Jin, H., Smith, A.B., McJannet, D.L., 2014. Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. *J. Hydrol.* 519 (Part D), 2747–2762. <http://dx.doi.org/10.1016/j.jhydrol.2014.08.008>.
- Ryu, D., Crow, W.T., Zhan, X., Jackson, T.J., 2009. Correcting unintended perturbation biases in hydrologic data assimilation. *J. Hydrometeorol.* 10 (3), 734–750. <http://dx.doi.org/10.1175/2008JHM1038.1>.
- Sefton, C., Howarth, S., 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *J. Hydrol.* 211 (1), 1–16.
- Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., Boni, G., 2015. Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. *Hydrol. Earth Syst. Sci.* 19 (4), 1727–1751. <http://dx.doi.org/10.5194/hess-19-1727-2015>.
- Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res.: Biogeosci.* (2005–2012) 116 (G4).
- Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrol. Process.* 17 (15), 3163–3170. <http://dx.doi.org/10.1002/hyp.5155>.
- Sobol', I.Y.M., 1990. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie* 2 (1), 112–118.
- Su, C.-H., Ryu, D., Young, R.I., Western, A.W., Wagner, W., 2013. Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. *Remote Sens. Environ.* 134, 1–11. <http://dx.doi.org/10.1016/j.rse.2013.02.016>.

- Sucksdorff, Y., Otle, C., 1990. Application of satellite remote sensing to estimate areal evapotranspiration over a watershed. *J. Hydrol.* 121 (1), 321–333.
- Sutanudjaja, E.H., van Beek, L.P.H., de Jong, S.M., van Geer, F.C., Bierkens, M.F.P., 2014. Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. *Water Resour. Res.* 50 (1), 687–705. <http://dx.doi.org/10.1002/2013wr013807>.
- Taconet, O., Bernard, R., Vidal-Madjar, D., 1986. Evapotranspiration over an agricultural region using a surface flux/temperature model based on NOAA-AVHRR data. *J. Climate Appl. Meteorol.* 25 (3), 284–307.
- Van Dijk, A., 2010. The Australian Water Resources Assessment System. Version 0.5, vol. 3, no. 5.
- Van Dijk, A., Marvanek, S., 2010. Derivation of a Simple Soil Drainage Model. AWRA Background Paper 2010/1.
- Van Dijk, A., Renzullo, L., 2011. Water resource monitoring systems and the role of satellite observations. *Hydrol. Earth Syst. Sci.* 15 (1), 39–55.
- Vaze, J., Viney, N., Stenson, M., Renzullo, L., Van Dijk, A., Dutta, D., Crosbie, R., Lerat, J., Penton, D., Vleeshouwer, J., 2013. The Australian Water Resource Assessment modelling system (AWRA). Proceedings of the 20th International Congress on Modelling and Simulation MODSIM2013, December. <<http://www.mssanz.org.au/modsim2013/L17/vaze.pdf>>.
- Viney, N., Perraud, J., Vaze, J., Chiew, F., Post, D., Yang, A., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, pp. 13–17.
- Wagener, T., Montanari, A., 2011. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resour. Res.* 47 (6). <http://dx.doi.org/10.1029/2010wr009469>.
- Wagener, T., Wheeler, H.S., 2006. Parameter estimation and regionalization for continuous rainfall–runoff models including uncertainty. *J. Hydrol.* 320 (1–2), 132–154. <http://dx.doi.org/10.1016/j.jhydrol.2005.07.015>.
- Wagener, T., Wheeler, H.S., Gupta, H.V., 2004. Rainfall–runoff modelling in gauged and ungauged catchments.
- Yadav, M., Wagener, T., Gupta, H., 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Adv. Water Resour.* 30 (8), 1756–1774. <http://dx.doi.org/10.1016/j.advwatres.2007.01.005>.
- Yang, Y., Long, D., Guan, H., Liang, W., Simmons, C., Batelaan, O., 2015. Comparison of three dual-source remote sensing evapotranspiration models during the MUSOEXE-12 campaign: revisit of model physics. *Water Resour. Res.* 51 (5), 3145–3165. <http://dx.doi.org/10.1002/2014WR015619>.
- Zhang, Y., Chiew, F., Zhang, L., Leuning, R., Cleugh, H., 2008. Estimating catchment evaporation and runoff using MODIS leaf area index and the Penman–Monteith equation. *Water Resour. Res.* 44 (10).
- Zhang, Y., Chiew, F.H.S., Zhang, L., Li, H., 2009. Use of remotely sensed actual evapotranspiration to improve rainfall–runoff modeling in Southeast Australia. *J. Hydrometeorol.* 10 (4), 969–980. <http://dx.doi.org/10.1175/2009jhm1061.1>.
- Zhang, Y., Viney, N., Chiew, F., Van Dijk, A., Liu, Y., 2011. Improving hydrological and vegetation modelling using regional model calibration schemes together with remote sensing data. In: Chan, F., Marinova, D., Anderssen, R.S. (Eds.), Proceedings of the 19th International Congress on Modelling and Simulation MODSIM2011. Modell. Simul. Soc. Aust. and NZ, December. <<http://www.mssanz.org.au/modsim2011/14/zhang.pdf>>.
- Zhang, Y., Wegehenkel, M., 2006. Integration of MODIS data into a simple model for the spatial distributed simulation of soil water content and evapotranspiration. *Remote Sens. Environ.* 104 (4), 393–408. <http://dx.doi.org/10.1016/j.rse.2006.05.011>.
- Zhou, Y., Zhang, Y., Vaze, J., Lane, P., Xu, S., 2013. Improving runoff estimates using remote sensing vegetation data for bushfire impacted catchments. *Agric. For. Meteorol.* 182–183, 332–341. <http://dx.doi.org/10.1016/j.agrformet.2013.04.018>.