Hydrol. Earth Syst. Sci. Discuss., 12, 10559–10601, 2015
www.hydrol-earth-syst-sci-discuss.net/12/10559/2015/
doi:10.5194/hessd-12-10559-2015

Hydrology and Earth System Sciences
Open Access
Discussions

Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations

P. Lopez Lopez[1,2], N. Wanders[2,3], J. Schellekens[1], L. J. Renzullo[4], E. H. Sutanudjaja[2], and M. F. P. Bierkens[2,5]

[1]Deltares, Delft, the Netherlands
[2]Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, the Netherlands
[3]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA
[4]CSIRO Land and Water, Canberra, ACT, Australia
[5]Unit Soil and Groundwater Systems, Deltares, Utrecht, the Netherlands

Received: 24 September 2015 – Accepted: 3 October 2015 – Published: 16 October 2015

Correspondence to: P. Lopez Lopez (patricia.lopez@deltares.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The coarse spatial resolution of global hydrological models (typically > 0.25°) limits their ability to resolve key water balance processes for many river basins and thus compromises their suitability for water resources management, especially when compared to locally-tuned river models. A possible solution to the problem may be to drive the coarse resolution models with locally available high spatial resolution meteorological data as well as to assimilate ground-based and remotely-sensed observations of key water cycle variables. While this would improve the resolution of the global model, the impact of prediction accuracy remains largely an open question. In this study we investigate the impact of assimilating streamflow and satellite soil moisture observations on the accuracy of global hydrological model estimations, when driven by either coarse- or high-resolution meteorological observations in the Murrumbidgee river basin in Australia.

To this end, a 0.08° resolution version of the PCR-GLOBWB global hydrological model is forced with downscaled global meteorological data (from 0.5° downscaled to 0.08° resolution) obtained from the WATCH Forcing Data methodology applied to ERA-Interim (WFDEI) and a local high resolution gauging station based gridded dataset (0.05°). Downscaled satellite derived soil moisture (from approx. 0.5° downscaled to 0.08° resolution) from AMSR-E and streamflow observations collected from 23 gauging stations are assimilated using an ensemble Kalman filter. Several scenarios are analysed to explore the added value of data assimilation considering both local and global meteorological data.

Results show that the assimilation of soil moisture observations results in the largest improvement of the model estimates of streamflow. The joint assimilation of both streamflow and downscaled soil moisture observations leads to further improvement in streamflow simulations (20 % reduction in RMSE).

Furthermore, results show that the added contribution of data assimilation, for both soil moisture and streamflow, is more pronounced when the global meteorological data

are used to force the models. This is caused by the higher uncertainty and coarser resolution of the global forcing.

We conclude that it is possible to improve PCR-GLOBWB simulations forced by coarse resolution meteorological data with assimilation of downscaled spaceborne soil moisture and streamflow observations. These improved model results are close to the ones from a local model forced with local meteorological data. These findings are important in light of the efforts that are currently done to go to global hyper-resolution modelling and can help to advance this research.

1 Introduction

In recent decades, a number of large-scale hydrological and land-surface models have been developed to quantify the global water cycle components, to analyse the human influence on the global water balance, to study climate change impact on water resources and to assess global hydrological extremes, such as drought and flood risk. VIC (Liang et al., 1994, 1996), WaterGAP (Alcamo et al., 2003), ORCHIDEE (d'Orgeval et al., 2008), HTESSEL (Balsamo et al., 2009), JULES (Best et al., 2011), PCR-GLOBWB (Van Beek et al., 2011), SURFEX-TRIP (Decharme et al., 2010, 2013) and W3RA (van Dijk, 2010; van Dijk et al., 2014) are some examples of large-scale hydrological models recently applied to provide water resources assessment over continental to global domains, but their coarse spatial resolution, typically between 0.5–1°, limits their ability to resolve key water balance processes for many river basins (Lanza et al., 1997; Wu and Li, 2009) and thus compromises their suitability for water resources management, especially when compared to locally-tuned hydrological models.

A possible solution to the problem may be to drive these original coarse resolution models with high resolution meteorological data. Several meteorological forcing datasets at a global scale are available, including the European Centre for Medium-Range Weather Forecasts – EMCWF ERA-Interim – global atmospheric reanalysis data (Dee et al., 2011), the Climatic Research Unit Time Series – CRU TS – (Mitchell

and Jones, 2005), the NASA reanalysis Modern-Era Retrospective Analysis for Research and Applications – MERRA – (Rienecker et al., 2011) and the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data – WFDEI – (Weedon et al., 2014). They are the result of integrating Bayesian merging of the available earth observations, in situ datasets and models to construct consistent large-scale meteorological time series. Some recent scientific efforts are conducted to improve the quality and availability of these datasets, for example increasing their spatial and temporal resolution (Cannon, 2011; Ebtehaj and Foufoula-Georgoiu, 2013; Atkinson, 2013). The use of high spatial resolution meteorological data would indirectly improve the resolution of the large-scale model, producing higher accuracy discharge estimates. However, when models that are designed for coarse spatial resolution are used at smaller spatial scale issues may arise with the representation of field scale processes (e.g. lateral flow). At the moment, more research is required to understand the gain that can be obtained using this higher spatial resolution forcing data for uncalibrated global hydrological models at finer spatial resolutions.

Another approach to bridge the gap between the different spatial scales is to assimilate ground-based and remotely-sensed observations of key water cycle variables. Higher resolution satellite data contain information at finer spatial resolution and could be used to correct for sub-optimal model performance at these finer resolutions. Multiple studies have used data assimilation techniques to obtain the best possible estimate of the hydrological system status, merging the strengths of hydrological modelling and observations and mitigating their respective weaknesses (Moradkhani, 2008; Clark et al., 2008; van Dijk et al., 2014). Among the sequential and variational data assimilation methods, the ensemble Kalman filter (Evensen, 2003) has arguably emerged as the most popular choice for assimilation into land surface and hydrological models. The various individual components of the water cycle, such as surface water (Vrugt et al., 2006; Rakovec et al., 2012), soil moisture (van Dijk et al., 2014; Wanders et al., 2014a), snow water (Sun et al., 2004; Moradkhani, 2008) and groundwater (Zaitchik

et al., 2008; Tangdamrongsub et al., 2015), which influence the hydrological system in different ways, can be assimilated into the model.

Soil moisture assimilation has been considered to improve model estimates, due to its key role in the terrestrial water cycle and its responsibility for the partitioning of precipitation between surface water and storage through infiltration. Several studies have assimilated soil moisture data (Draper et al., 2011; Chen et al., 2011; Wanders et al., 2014b) both based on ground soil moisture measurements and remotely sensed satellite soil moisture products from remote observation systems, such as ASCAT (Naeimi et al., 2009), SMOS (Kerr et al., 2012) and AMSR-E (Dorigo et al., 2010). On the other hand, surface water information has often been used for data assimilation frameworks (Vrugt et al., 2006; Rakovec et al., 2012) because discharge provides integrated information of all hydrological states, which is often very effective in improving model simulations. However, the risk of an integrated observation is that in some scenarios accurate simulations could be obtained by adjusting the wrong states variables.

An additional improvement could be made by the assimilation of downscaled or disaggregated satellite soil moisture observations into a particular land surface model (Merlin et al., 2006; Sahoo et al., 2013). Recently, new soil moisture products of higher spatial resolution have been released (Gevaert et al., 2015) but their impact on hydrological model predictions has not yet been explored. For example, they could be used to correct for incorrectly observed (or interpolated) precipitation patterns, which directly affect the input uncertainty into the model.

Moreover, improved results can be obtained by assimilation of multiple observational dataset of different parts of the hydrological cycle into the hydrological model (Barrett and Renzullo, 2009; Reichle et al., 2014). For example, the joint assimilation of discharge and soil moisture could results in an improved understanding of the runoff generation mechanisms and increase the quality and quantity of information incorporated to the model system. However, the added value of this type of joint assimilation procedures is largely unknown and should be further investigated (Lee et al., 2011).

Many data assimilation experiments have been set up in conjunction with local-scale hydrological models and the benefit of data assimilation for large-scale models remains largely an open question. In this context, it is interesting to analyse whether the accuracy of large-scale hydrological models can be improved and become more comparable with locally calibrated model estimates if satellite observations are assimilated. Understanding the potential gain of assimilating satellite observations into large-scale models is a relevant research opportunity and may have potential benefits for water resources management (Van Dijk and Renzullo, 2011). For example, in regions without or low quality meteorological observations the use of large-scale models in combination with satellite data assimilation could improve our understanding of the available water resources.The primary goal of the present study is to investigate the impact of assimilating streamflow and satellite soil moisture observations on the accuracy of global hydrological model estimations, when driven by either coarse- or high-resolution meteorological observations. The Murrumbidgee river basin in the southeast of Australia was chosen as a case study for the investigation because of the variety of land uses in the area, the high level of monitoring available for a large number of relatively unimpaired catchments, and the extensive body of previous studies observing and describing the hydrologic patterns across the basin (Renzullo et al., 2014; van Dijk and Renzullo, 2011). Eight data assimilation scenarios were considered in which discharge and soil moisture observations were either independently or jointly used, and the forcing data were obtained from either local or global data sets. In this context, comparison of the eight scenarios with the locally calibrated model estimates provides insight into how the estimations of global hydrological models driven by global forcing data can come closer to local-scale model predictions.

2 Approach, materials and methods

2.1 Study basin: Murrumbidgee river, Australia

The selected study area was the Murrumbidgee river basin (84 000 km^2) located in south-east Australia, specifically in the south west of New South Wales (Fig. 1). The Murrumbidgee river is the second largest river in the Murray–Darling system, flowing for a distance of approximately 1600 km. Elevations range from over 1900 m in the East to less than 50 m on the Western plain. Forest and woodland coverage dominate in the East, with pasture and cropping in the central region, and increasing grassland to the West. The Western plain is dominated by clay-loam soils and with decreasing clay content in the middle and eastern region (Peischl et al., 2012). The climate in the catchment is one of the most diverse in New South Wales, with an average annual rainfall that ranges from 1700 mm yr^{-1} in the higher elevations of the Snowy Mountains in the east, to less than 350 mm yr^{-1} on the Western plain. Average reference evapotranspiration varies from less than 1000 mm yr^{-1} in the south-east, to over 1800 mm yr^{-1} in the West. Mean annual flow increases from less than 45 $m^3 s^{-1}$ in the upstream tributaries to approximately 125 $m^3 s^{-1}$ in the mid-Murrumbidgee (Green et al., 2011).

2.2 Hydrological models

The simulations for two distributed hydrological models, i.e. the local OpenStreams wflow_sbm and the global PCR-GLOBWB, were performed for the period 2007–2010. The period 2000–2007 was used to "spin up" the models. The local and large-scale models are described in detail in the following two sub-sections.

2.2.1 Local-scale model: OpenStreams wflow_sbm

The local-scale hydrological model employed in this study was the OpenStreams wflow_sbm model (Schellekens, 2014). This is a distributed model derived from topog_sbm simple bucket model developed by Vertessy and Elsenbeer (1999). The

OpenStreams wflow_sbm model (OSWS) is programmed in the PCRaster-Python environment (Wesseling et al., 1996; Karssenberg et al., 2010) and it is publicly available through the OpenStreams project (https://github.com/openstreams/wflow). The defined spatial resolution used in this study was $0.01^\circ \times 0.01^\circ$ (approx. 1 km × 1 km) and the temporal resolution was daily. A schematic representation of OSWS is given in Fig. 2.

OSWS model structure consists of three main routines: (i) rainfall interception (schematized by the Gash model – Gash, 1979), (ii) soil processes (schematized by the topog_sbm model) and (iii) river drainage and overland flow (modelled by the kinematic wave routing over a drainage network).

The water enters each model cell from precipitation to the canopy interception storage or snow storage. The remaining liquid water infiltrates into the soil. At the same time, water is taken from the soil through evapotranspiration (based on soil water content and vegetation type). The water exchange into the soil considers two vertical soil layers, the unsaturated zone (UZ) and the saturated zone (SZ), based on topog_sbm structure. Total runoff is the sum of the direct runoff, the melt water that does not infiltrate into the soil and the baseflow (lateral subsurface flow from the saturated zone). This total runoff is conducted along the river network as discharge with kinematic wave routing.

The OSWS model was calibrated for the Murrumbidgee river basin using observations from in situ streamflow gauges (BoM, 2015) for the time period 1990–2010. These gauges are different from those considered in all the data assimilation scenarios to ensure an independent verification.

2.2.2 Large-scale model: PCR-GLOBWB

The large-scale hydrological model employed in this study was PCR-GLOBWB (Van Beek and Bierkens, 2009; Van Beek et al., 2011). Similar to OpenStreams wflow_sbm, PCR-GLOBWB is essentially a leaky-bucket type of model applied on a cell-by-cell basis. PCR-GLOBWB is coded in the PCRaster-Python environment. A spatial resolution

of 0.08° (approx. 10 km × 10 km at the equator) and a daily temporal resolution were used in this study. A schematic representation of PCR-GLOBWB is given in Fig. 3.

For each time step and cell, PCR-GLOBWB calculates the water balance components, including the water storage in three vertical soil layers (0–5, 5–30 and 30–150 cm) and one underlying groundwater reservoir, as well as the water exchange between the layers (percolation, capillary rise) and between the top layer and the atmosphere (rainfall, evapotranspiration and snowmelt). Sub-grid variability is taken into account considering the variations of elevation, land cover, vegetation and soil. The total runoff of a cell consists of direct runoff (saturation excess surface runoff), non-infiltrating melt water, interflow (lateral drainage from the soil profile) and baseflow (groundwater runoff from the lowest linear reservoir). The simulated runoff is routed along the river network based on the Simulated Topological Networks (STN30; Vörösmarty et al., 2000). Water abstraction and consumptive water use (domestic, industrial, livestock, irrigation) and reservoir management are included.

In contrast to the local-scale model, PCR-GLOBWB was not calibrated for the study basin. Hydrological model parameters were derived from vegetation, soil properties or geological information and estimated at a global scale.

2.3 Datasets

2.3.1 Meteorological forcing data

The forcing data required to drive both hydrological models are precipitation, temperature and reference evapotranspiration (ET). Two types of forcing data were used in this study: local forcing data, representing the best available data, and global forcing data representing a lower spatial resolution data set but one which is available globally.

Local precipitation and temperature data were obtained from the gridded datasets generated by the Australian Bureau of Meteorology under the Australian Water Availability Project (AWAP) (Jones et al., 2009). The data are derived from station-level daily precipitation and air temperature, and interpolated on a 0.05° × 0.05° grid covering the

Australian continent. These data represent high resolution meteorology in this study which, we argue, will provide the modelling benchmark results.

Global precipitation and temperature data were obtained from the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data – WFDEI – (Weedon et al., 2014). The daily global precipitation and temperature data were provided at a spatial resolution of 0.5° × 0.5°. To obtain finer spatial resolution climate maps, from 0.5° × 0.5° grid to 0.08° × 0.08° grid, a downscaling procedure was applied based on a linear regression analysis (Sutanudjaja et al., 2011). It makes use of precipitation and temperature lapse rates derived from the 10' CRU-CL2.0 climatology dataset (New et al., 2002).

Local and global reference evapotranspiration (ET) were obtained through Hamon method (Allen et al., 1998; Lu et al., 2005).

Figure 4 shows the daily mean precipitation, temperature and reference evapotranspiration for the study time period (2007–2010) for both forcing data sets. Aside from the resolution difference, precipitation ranges from higher values in the mountainous regions of the catchment, with increasing variance in elevation and rainfall, to lower values in the western plain. On the contrary, temperature, and subsequently reference evapotranspiration decrease from West to East.

Local and global forcing data show some differences in their spatial distribution and magnitude. Each climate variable shows similar spatial distribution across the various resolutions (rows in Fig. 4), with larger variations in the high elevation areas. However, local temperature magnitude differs in 3–4 °C with the global temperature, which is also reflected in the reference evapotranspiration. The downscaled global forcing data show a similar pattern to the global data with increasing resolution.

Table 1 shows the catchment daily mean values of the climate forcing variables for each year individually. Global temperature deviates from local approximately 3–4 °C, which is also reflected in reference evapotranspiration, but less pronounced. Local and global precipitation, by contrast, is very similar for each year.

2.3.2 Soil moisture data

Soil moisture observations retrieved from AMSR-E (Advanced Microwave Scanning Radiometer – EOS) brightness temperatures were provided by the Vrije Universiteit Amsterdam (VUA) in collaboration with NASA. AMSR-E is a multi-frequency passive microwave radiometer that uses C- (6.9 GHz) and X-band (10.65 and 18.7 GHz) radiance observations to derive near-surface soil moisture via the LPRM radiative transfer model (Owe et al., 2008; De Jeu et al., 2008). In the present study, C-band AMSR-E data reported on a regular ∼ 50 km global grid with an observation depth of 2 cm and a revisit time of 1–3 days was considered (Owe et al., 2008). Brightness temperatures from C-band AMSR-E were downscaled using the smoothing filter-based modulation technique. In this technique, brightness temperatures from the C-band (approx. 50 km × 50 km) are adjusted based on data from the Ka-band (approx. 10 km × 10 km). From the downscaled C-band AMSR-E brightness temperatures, soil moisture on a $0.08^{\circ} \times 0.08^{\circ}$ spatial resolution is estimated. To improve the quality of the final soil moisture products a precipitation mask is applied (Gevaert et al., 2015).

Soil moisture observations from AMSR-E were compared to the unsaturated zone layer of OSWS and the first of the three vertical layers constituting the soil profile in each grid cell of PCR-GLOBWB. To match the remotely sensed soil moisture observations to the statistics of corresponding hydrological model states for soil water, different strategies can be followed, such as linear or minimum–maximum (MM) matching (Brocca et al., 2011), mean-standard deviation (μ-σ) matching (Draper et al., 2009) and cumulative distribution function (CDF) matching (Reichle and Koster, 2004). In this study, a linear rescaling method was used. The converted satellite soil moisture values θ_{new} (in $m^3\,m^{-3}$) used for assimilation were calculated as

$$\theta_{new} = I_{min} + \left(\frac{I_{max} - I_{min}}{\theta_{max} - \theta_{min}}\right)(\theta - \theta_{min}) \quad (1)$$

where I_{max} and I_{min} are the field capacity and the wilting point of the modelled soil moisture values at each grid cell [$m^3\,m^{-3}$] and θ_{max} and θ_{min} are the maximum and minimum of AMSR-E satellite soil moisture values at the respective grid location [–].

2.3.3 Discharge data

Discharge observations were provided by the Bureau of Meteorology (BoM) and the Commonweath Scientific and Industrial Research Organisation (CSIRO), under the Water Information R+D Alliance (WIRADA). A total of 23 discharge monitoring stations with daily observations in the Murrumbidgee river and its main tributaries was available for the period January 2007 to December 2010. To ensure an independent evaluation of model simulations after the assimilation, a split sample approach of streamflow stations was used (e.g. Lee et al., 2012; Rakovec et al., 2012; Wanders et al., 2014b). The discharge of 10 stations was used for assimilation into the large-scale hydrological model, the remaining 13 stations were used for evaluation. Assimilation and evaluation stations were selected such they are equally distributed over the catchment and are situated both in small tributaries and the main Murrumbidgee river. Figure 1 shows a map with the discharge locations. Table 2 summarizes some key hydrological data.

2.4 Data assimilation

2.4.1 Ensemble Kalman filter

The Ensemble Kalman filter (EnKF) is a sequential data assimilation method evolved from the standard Kalman filter (Evensen, 1994) that has been used previously for assimilation of observations into land surface and hydrological models (Chen et al., 2011; Draper et al., 2011; Wanders et al., 2014a; Tangdamrongsub et al., 2015). It is a Monte Carlo based approach that integrates an ensemble of model states forward in time to represent the error statistics of the model estimate when observations are

assimilated (Burgers et al., 1998; Evensen, 2003). The state equation in a discrete form is given as

$$x_{t+1} = f\left(x_t, F_{t+1}, p, \varepsilon_t\right) \tag{2}$$

where f are the dynamical model equations that represents the hydrological processes in the system, x_t is the model state at time t, F_t is the forcing at time t (e.g. precipitation and temperature), p are the model parameters and ε_t is the model error.

To assimilate observations into the hydrological model, the already mentioned observations, downscaled remotely sensed AMSR-E soil moisture and discharge, can be linearly described as

$$y_t = H_t x_t + \epsilon \tag{3}$$

where y_t is the observations vector, H is the observation model or operator that relates the model states x_t to the observations y, and ϵ is the random noise with a zero mean and an error given by the observations error covariance matrix R_t.

The EnKF calculates the analysis at each time t, x_t^{a}, of the model forecast, x_t^{f}, as

$$x_t^{\mathrm{a}} = x_t^{\mathrm{f}} + K_t\left[y_t - H_t\left(x_t^{\mathrm{f}}\right)\right] \tag{4}$$

where K_t is defined as the Kalman gain

$$K_t = P_t H_t^{\mathrm{T}}\left(R_t + H_t P_t H_t^{\mathrm{T}}\right)^{-1} \tag{5}$$

with H_t^{T} the transpose matrix of the observation model at time t and P_t the state error covariance matrix of the model prediction calculated from the spread between the different ensemble members given as

$$P_t = \frac{\sum_{n=1}^{N}\left(x_n^{\mathrm{f}} - \overline{x^{\mathrm{f}}}\right)\left(x_n^{\mathrm{f}} - \overline{x^{\mathrm{f}}}\right)^{\mathrm{T}}}{N-1} \tag{6}$$

where $\overline{x^{\mathrm{f}}}$ the ensemble average of model simulations and N the number of ensemble members considered.

2.4.2 Assimilating soil moisture and discharge observations

In this study, the EnKF was applied to update state variables of the large-scale hydrological model, PCR-GLOBWB, on each daily time step using downscaled remotely sensed AMSR-E soil moisture and discharge observations. We used 100 ensemble members and all the observations were calculated and assimilated at each 0.08° model grid cell for each day that observations are available. The EnKF has been implemented in the PCRaster modelling environment (Karssenberg et al., 2010).

Eight different data assimilation scenarios with PCR-GLOBWB were inter-compared and compared to the OSWS estimates without any data assimilation. The data assimilation (DA) scenarios are described in Table 3, indicating the meteorological forcing and the observations used in each scenario. Simulations forced with local meteorological data are denoted with L and simulations forced with global meteorological data are denoted with G. Independent assimilation of discharge (L_Q and G_Q) and soil moisture (L_SM and G_SM) were investigated, as well as the joint assimilation of both observation types (L_SM_Q and G_SM_Q).

In the EnKF, to account for model and observations uncertainty, stochastic noise can be introduced in model forcing data, parameters, soil moisture and discharge observations. For the local and global meteorological forcing, the precipitation was perturbed with additive Gaussian white noise with standard deviation of 10 % of the nominal value (Adam and Lettenmaier, 2003; Hijmans et al., 2005). The errors were assumed to be spatially uncorrelated. For the assimilation of the satellite soil moisture data, spatial information on the measurement error covariance R (Eq. 5) was required. The structure of R was determined from estimates of Wanders et al. (2012) over Spain, obtained by using high-resolution modelling of the unsaturated zone. The average standard error of AMSR-E is 0.049 $\mathrm{m^3\,m^{-3}}$. The error covariance between the discharge observations

was set to zero while the standard error for the discharge observations was assumed to be 10 % of the discharge. It was additionally assumed that the covariance between the satellite soil moisture observations and discharge observations equals zero.

2.5 Evaluation

The impact of assimilating discharge and soil moisture observations into the large-scale hydrological model PCR-GLOBWB compared with the locally calibrated model OSWS, is separately analysed: firstly, on soil moisture estimates and secondly, on discharge estimates. A common regular $0.08^\circ \times 0.08^\circ$ grid (approx. 10 km) was adopted for the inter-comparison of the two different resolution hydrological models estimates. For this purpose, OSWS estimates were upscaled with a linear resampling from $0.01^\circ \times 0.01^\circ$ (approx. 1 km) to $0.08^\circ \times 0.08^\circ$ (approx. 10 km).

Results were produced for each of the 23 locations listed in Table 2. For practical reasons, the following section includes results for a limited number of evaluation locations only, both in the Murrumbidgee river and its tributaries. This combination thus comprises stations with varying sizes of contributing area.

To understand and inter-compare the performance of the different data assimilation scenarios described in Table 3, an extensive evaluation was carried out, including the calculation of various evaluation metrics, such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Pearson's correlation coefficient (r) and Nash Sutcliffe efficiency (NSE – Nash and Sutcliffe, 1970).

3 Results

3.1 Impact of assimilation on soil moisture estimates

The time series of simulated soil moisture for 0–5 cm for each data assimilation scenario and downscaled AMSR-E observations for the time period January 2008–May 2009 at 410057 gauging station (mountainous region) are shown in Fig. 5.

The use of global forcing data produces a different dynamic response of soil moisture estimates compared to the local forcing data. This fact is due to the discrepancies between both meteorological datasets, which govern the water exchange processes between the top layer and the atmosphere (precipitation, evapotranspiration and temperature).

Even though global precipitation quite accurately depicts the overall character of the precipitation (daily mean values of the local and global precipitations show similar spatial distributions and magnitudes – see Sect. 2.3.1), the global precipitation misses specific rainfall events at particular days and locations due to its lower resolution. This is especially important for warm season precipitation and regions in mountainous terrain (e.g. 410057 gauging station in Fig. 5), which are dominated by convective storms. The differences in precipitation are reflected in soil moisture estimates of both hydrological models and their impact is higher in PCR-GLOBWB estimates. When PCR-GLOBWB is forced with local data, soil moisture estimates produce patterns with a more accurate description of the small-scale variability of the observations in time. Whereas, when global forcing is used, soil moisture results in a smoother estimation of the observations.

Eight different data assimilation scenarios were investigated, with global and local forcing data (Table 3). In the OL scenario, no data is assimilated into the model to correct intermediate model states. In this first scenario, soil moisture observations are underestimated when local forcing is used. The assimilation of discharge observations, Q scenario, results in similar soil moisture estimates compared to the OL scenario, as expected. This similarity is caused by the assimilation procedure which constrains the model to follow the discharge observations directly affecting on groundwater and routing processes, which are a poor reflection of the upper soil moisture content from 0 to 5 cm. In contrast, the assimilation of soil moisture observations in SM and SM_Q scenarios produces a reduction of the negative bias, improving the soil moisture estimates, especially when local forcing is used.

Figure 6 shows the impact of each data assimilation scenario on the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Pearson's correlation coefficient (r). Results of the catchment daily mean values are shown. Each histogram shows the evaluation metric on the vertical axis vs. the data assimilation scenarios on the horizontal axis. Figure 6 consists of a matrix of multiple panels, with rows showing the three considered evaluation metrics and columns showing local and global forcing data.

From this figure, some general observations can be made. Evaluation results show differences between the results from local and global forcing of the models. For example, the use of local instead of global forcing produces a decrease of 4 and 2 % in RMSE and MAE, respectively and an increase of 6 % in r, when the OL scenario in PCR-GLOBWB is considered.

The large-scale model, PCR-GLOBWB, without data assimilation shows a poorer performance than the locally calibrated model OSWS on soil moisture predictions. Assimilating discharge observations (Q scenario) does not lead to an improvement on soil moisture estimates, whereas downscaled AMSR-E soil moisture observations assimilation (SM scenario) results in an increase of r and a decrease of RMSE and MAE. Therefore, after assimilating soil moisture observations, evaluation results of PCR-GLOBWB and OSWS are closer to each other. For example, percent differences in RMSE between both models are reduced from 22 % (L_OL and L_w) to 16 % (L_SM_Q and L_w) and MAE from 14 % (L_OL and L_w) to 16 % (L_SM_Q and L_w).

Additionally, boxplots of the catchment daily mean values are included in Fig. 7 considering local (upper panel) and global (lower panel) forcing. The assimilation of soil moisture observations leads to an evident improvement in the statistical distribution of PCR-GLOBWB soil moisture estimates, reducing the differences in dispersion with the observations.

Assimilating soil moisture observations and forcing the model with high spatial resolution meteorological datasets impacts the quality of soil moisture estimates with PCR-GLOBWB to a similar extent. Results indicate that the highest improved performance is

achieved when their combination occur, i.e. soil moisture observations are assimilated into a model driven by local forcing data (L_SM and L_SM_Q scenarios).

3.2 Impact of assimilation on streamflow estimates

The simulated and observed streamflow estimates at 410057 gauging station are shown in Fig. 8. From this figure, it is clear that the peaks in streamflow are poorly estimated by PCR-GLOBWB, whereas OSWS is able to capture them with higher accuracy, independently of the forcing data used. This is most probably explained by the higher resolution and the calibration of model parameters for the study.

Differences between local and global forcing data (see Sect. 2.3.1) are reflected in differences in streamflow estimates from both models. When global data is used, evapotranspiration is lower; hence a higher amount of water is introduced into the models, resulting in higher streamflow estimates. By assimilating discharge and soil moisture observations, intermediate hydrological processes, including groundwater state, percolation and surface runoff among others, are corrected and errors in forcing data are reduced to improve streamflow predictions (e.g. November 2008). The best performance is achieved when soil moisture and discharge data are jointly assimilated.

To further analyse and quantify the influence of each data assimilation scenario on streamflow estimates, the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Pearson's correlation coefficient (r) and Nash Sutcliffe efficiency (NSE) were calculated and included in Fig. 9.

The highest r and NSE and the lowest RMSEs and MAEs are obtained when models are forced with local meteorological data. The use of global forcing data leads to a reduction in performance, which is more significant for the large-scale than for the local-scale model. Without assimilation, forcing PCR-GLOBWB with local data (L_OL) instead of global data (G_OL) results in an increase of 80 % in r and a decrease of 70 % in RMSE and 72 % in MAE on average. OSWS also improves its streamflow estimates but to a lesser degree, with increases of 7 % in r and decreases of 28 % in RMSE and 43 % in MAE (L_w and G_w).

Both observations assimilation, discharge and soil moisture separately (Q and SM), improve streamflow models estimates. The highest improvement is achieved when both observations are assimilated into the model (SM_Q), as expected. The improvement is higher when soil moisture observations (SM) are assimilated than the case of discharge assimilation alone (Q). For example, in terms of r G_SM scenario results in an increase of 20 % and G_Q scenario of 5 % relative to G_OL scenario at 410107 gauging station. Some possible explanations could be the finer resolution of AMSR-E soil moisture observations and/or the basin hydrological features, which characterize it as a catchment mainly driven by direct runoff, where the highest contribution to the total runoff comes from the upper soil layer and not from the groundwater zone.

The largest improvements were found at gauging stations in the main channel of the Murrumbidgee river, such as station 410001 where assimilating soil moisture and discharge observations increases r from 0.56 to 0.79 and decreases RMSE and MAE from 49.54 to 28.77 $m^3 s^{-1}$ and from 43.85 to 10.76 $m^3 s^{-1}$, i.e. comparing the L_OL and L_SM_Q scenarios. This improvement is more significant when the model is forced with global data than with local data. At station 410001, RMSE varies from 157.62 to 29.01 $m^3 s^{-1}$, MAE from 148.29 to 10.88 $m^3 s^{-1}$ and r from 0.26 to 0.72 when G_OL and G_SM_Q scenarios are compared.

Boxplots of streamflow estimates are included in Fig. 10. The results clearly show that, compared to the observed streamflow, the median values of PCR-GLOBWB streamflow without data assimilation are very biased. The greatest amount of spread is observed between OL scenarios and observations, across all stations. Every data assimilation scenario (Q, SM and SM_Q) shows improvement in the statistics of streamflow, correcting not only its median value, but also the overestimation.

In contrast with the soil moisture evaluation results, Figs. 8–10 indicate that for streamflow the use of local forcing provides a larger improvement than assimilating soil moisture and discharge observations. Using a global model with local forcing and assimilating satellite soil moisture data yields reasonably good streamflow predictions along the main river of this catchment. Moreover, also on the main channel, a global

model with global forcing may still yield reasonable results as long as both discharge data and soil moisture data are assimilated.

4 Discussion

The joint assimilation of discharge and downscaled satellite soil moisture observations produces the largest improvement on PCR-GLOBWB streamflow estimates (20 % reduction in RMSE). These results agree with the findings made by Wanders et al. (2014a) in the Upper Danube and it is also in line with the expectations, where more observations lead to a better constrained model simulation.

A relevant finding of this study is the positive impact of assimilating soil moisture observations on the streamflow estimates, compared to the independent discharge assimilation. The variable effectiveness of soil moisture assimilation has been previously reported in literature. Whereas some studies found improvements (Draper et al., 2011; Wanders et al., 2014b), others obtained mixed or unsatisfactory results (Crow et al., 2005). The scale of soil moisture observations, the dominant runoff processes in the study basin and the model structure and parameters uncertainties may partly explain this variability. In this particular study, the novel use of a finer spatial resolution satellite soil moisture product together with the climate and hydrogeological characteristics of the catchment could be a possible explanation of the positive impact of soil moisture assimilation. Renzullo et al. (2014) assimilated satellite soil moisture observations from multiple sensors (ASCAT and AMSR-E) obtained at scales coarser than the model (AWRA-L) resolution. In the present study, the scale of soil moisture observations coincides with the model scale. Moreover, runoff in the Murrumbidgee river basin is mainly dominated by direct runoff processes, with reduced contribution from the groundwater zone (Green et al., 2011). These catchment conditions, together with their representation in the model structure are most likely responsible for the added value of assimilating soil moisture.

For the assimilation of AMSR-E soil moisture observations, spatial information on the measurement error covariance R (Eq. 5) was based on results from previous studies over Spain (Wanders et al., 2012). Our philosophy was to set the AMSR-E errors to realistic values determined and validated in previous studies, so that all the required information was already available. The determination of AMSR-E observations uncertainty specifically over the Murrumbidgee river basin with physically-based modelling or in-situ soil moisture measurements (Su et al., 2013) could be further investigated.

Meteorological data play a key role in soil moisture and discharge model estimates. The various model and data assimilation options were evaluated under both high- and low-resolution meteorological forcing. In general, the higher spatial resolution of the local forcing data results in better model predictions (PCR-GLOBWB and OSWS) of both soil moisture and streamflow. The coarse resolution of the global forcing could lead to failures to detect extreme rainfall events or differences at specific regions, such as mountainous areas. Recent studies have developed several downscaling procedures, e.g. geostatistical methods of blending satellite and gauge data (Chappell et al., 2013), that could improve model predictions when global forcing is used.

5 Conclusions

The study investigates the influence of discharge and soil moisture assimilation on the accuracy of large-scale model (PCR-GLOBWB) predictions, when driven by meteorological forcing datasets of high- and coarse-resolution, compared with local-scale model (OSWS) estimations.

Results show that the assimilation of soil moisture observations results in the largest improvement of the model estimates of streamflow. The joint assimilation of both streamflow and downscaled soil moisture observations leads to further improvement in streamflow simulations (20 % reduction in RMSE).

In general, the higher spatial resolution of the local forcing data results in better models predictions of both soil moisture and streamflow. The added value of using

higher spatial resolution forcing data is more significant for PCR-GLOBWB than for OSWS. When the impact on model accuracy of assimilating observations and forcing the models with higher spatial resolution data are compared, the latter leads to a more substantial improvement of streamflow predictions.

Furthermore, results show that the added contribution of data assimilation, for both soil moisture and streamflow, is more pronounced when the global meteorological data are used to force the models. This is caused by the higher uncertainty and coarser resolution of the global forcing.

The greatest benefit is obtained when local coarse resolution forcing data are used in combination with streamflow and soil moisture observations assimilation into the large-scale hydrological model, PCR-GLOBWB.

In conclusion, the present research study shows that data assimilation of high resolution soil moisture succeeds in resolving short-comings that exist nowadays in global hydrological models and can partly overcome the difference in model performance between a large-scale hydrological model driven by coarse resolution forcing data and a local-scale model forced with higher resolution meteorological data. Moreover, it demonstrates that further investments and improvements in remotely sensed observations, especially in soil moisture products can benefit large-scale hydrological model predictions and bring these closer to those obtained from local-scale hydrological modelling.

Acknowledgement. This research received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 603608, “Global Earth Observation for integrated water resource assessment”: eartH2Observe. The work of N. Wanders was supported by NWO (GO-AO/30) and Rubicon grant (825.15.003). We would like to thank CSIRO in Canberra, ACT, Australia for their active collaboration and constant engagement in the present research project. The authors also want to thank Anouk Gevaert for providing the downscaled AMSR-E soil moisture observations. Geert Sterk is acknowledged for productive comments and discussions on this investigation.

References

Adam, J. C. and Lettenmaier, D. P.: Adjustment of global gridded precipitation for systematic bias, J. Geophys. Res.-Atmos., 108, 4257, doi:10.1029/2002JD002499, 2003.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrolog. Sci. J., 48, 317–337, 2003.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper 56, FAO, Rome, 300 pp., 1998.

Atkinson, P. M.: Downscaling in remote sensing, Int. J. Appl. Earth Obs., 22, 106–114, 2013.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the integrated forecast system, J. Hydrometeorol., 10, 623–643, 2009.

Barrett, D. J. and Renzullo, L. J.: On the efficacy of combining thermal and microwave satellite data as observational constraints for root-zone soil moisture estimation, J. Hydrometeorol., 10, 1109–1127, 2009.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R .L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geosci. Model Dev., 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.

BoM – Bureau of Meteorology: Australian Government, Water Data Online, available at: http://www.bom.gov.au/waterdata/, last access: 25 September 2015.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe, Remote Sens. Environ., 115, 3390–3408, 2011.

Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, Mon. Weather Rev., 126, 1719–1724, 1998.

Cannon, A. J.: Quantile regression neural networks: implementation in R and application to precipitation downscaling, Comput. Geosci., 37, 1277–1284, 2011.

Chappell, A., Renzullo, L. J., Raupach, T. H., and Haylock, M.: Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia, J. Hydrol., 493, 105–114, 2013.

Chen, F., Crow, W. T., Starks, P. J., and Moriasi, D. N.: Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, Adv. Water Resour., 34, 526–536, 2011.

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

Crow, W. T., Bindlish, R., and Jackson, T. J.: The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall–runoff partitioning, Geophys. Res. Lett., 32, L18401, doi:10.1029/2005GL023543, 2005.

Decharme, B., Alkama, R., Douville, H., Becker, M., and Cazenave, A.: Global evaluation of the ISBA-TRIP continental hydrological system, Part II: Uncertainties in river routing simulation related to flow velocity and groundwater storage, J. Hydrometeorol., 11, 601–617, 2010.

Decharme, B., Martin, E., and Faroux, S.: Reconciling soil thermal and hydrological lower boundary conditions in land surface models, J. Geophys. Res.-Atmos., 118, 7819–7834, 2013.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, 2011.

De Jeu, R. A. M., Wagner, W., Holmes, T. R. H., Dolman, A. J., Van De Giesen, N. C., and Friesen, J.: Global soil moisture patterns observed by space borne microwave radiometers and scatterometers, Surv. Geophys., 29, 399–420, 2008.

d'Orgeval, T., Polcher, J., and de Rosnay, P.: Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes, Hydrol. Earth Syst. Sci., 12, 1387–1401, doi:10.5194/hess-12-1387-2008, 2008.

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture datasets, Hydrol. Earth Syst. Sci., 14, 2605–2616, doi:10.5194/hess-14-2605-2010, 2010.

Draper, C. S., Walker, J. P., Steinle, P. J., de Jeu, R. A., and Holmes, T. R.: An evaluation of AMSR-E derived soil moisture over Australia, Remote Sens. Environ., 113, 703–710, 2009.

Ebtehaj, A. M. and Foufoula-Georgiou, E.: On variational downscaling, fusion, and assimilation of hydrometeorological states: A unified framework via regularization, Water Resour. Res., 49, 5944–5963, doi:10.1002/wrcr.20424, 2013.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res.-Oceans, 99, 10143–10162, 1994.

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

Gash, J. H. C.: An analytical model of rainfall interception by forests, Q. J. Roy. Meteorol. Soc., 105, 43–55, 1979.

Gevaert, A. I., Parinussa, R. M., Renzullo, L. J., van Dijk, A. I. J. M., and de Jeu, R. A. M.: Spatio-temporal evaluation of resolution enhancement for passive microwave soil moisture and vegetation optical depth, Int. J. Appl. Earth Obs., doi:10.1016/j.jag.2015.08.006, in press, 2015.

Green, D., Petrovic, J., Moss, P., and Burrell, M.: Water resources and management overview: Murrumbidgee catchment, NSW Office of Water, Sydney, 2011.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25, 1965–1978, 2005.

Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., and Bierkens, M. F.: A software framework for construction of process-based stochastic spatio-temporal models and data assimilation, Environ. Modell. Softw., 25, 489–502, 2010.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., and Delwart, S.: The SMOS soil moisture retrieval algorithm, IEEE T. Geosci. Remote, 50, 1384–1403, 2012.

Lanza, L. G.; Schultz, G. A., and Barrett, E. C.: Remote sensing in hydrology: some downscaling and uncertainty issues, Phys. Chem. Earth, 22, 215–219, 1997.

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

Lee, H., Seo, D.-J., Liu, Y., Koren, V., McKee, P., and Corby, R.: Variational assimilation of streamflow into operational distributed hydrologic models: effect of spatiotemporal scale of adjustment, Hydrol. Earth Syst. Sci., 16, 2233–2251, doi:10.5194/hess-16-2233-2012, 2012.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res., 99, 14415–14415, 1994.

Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L model: evaluation and modification, Global Planet. Change, 13, 195–206, 1996.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M.: A comparison of six potential evapotranspiration methods for regional use in the southeastern United States, J. Am. Water Resour. As., 41, 621–633, 2005.

Merlin, O., Chehbouni, A., Boulet, G., and Kerr, Y.: Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data, J. Hydrometeorol., 7, 1308–1322, 2006.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Moradkhani, H.: Hydrologic remote sensing and land surface data assimilation, Sensors, 8, 2986–3004, 2008.

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W.: An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, IEEE T. Geosci. Remote, 47, 1999–2013, 2009.

Nash, J. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, 1970.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Clim. Res., 21, 1–25, 2002.

Owe, M., De Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res.-Earth, 113, F01002, doi:10.1029/2007JF000769, 2008.

Peischl, S., Walker, J. P., Rüdiger, C., Ye, N., Kerr, Y. H., Kim, E., Bandara, R., and Allahmoradi, M.: The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment, Hydrol. Earth Syst. Sci., 16, 1697–1708, doi:10.5194/hess-16-1697-2012, 2012.

Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating fre-

quency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435–3449, doi:10.5194/hess-16-3435-2012, 2012.
Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture. Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938, 2004.
Reichle, R. H., De Lannoy, G. J., Forman, B. A., Draper, C. S., and Liu, Q.: Connecting satellite observations with water cycle variables through land data assimilation: examples using the NASA GEOS-5 LDA S, in: The Earth's Hydrological Cycle, Springer, the Netherlands, 577–606, 2014.
Renzullo, L. J., van Dijk, A. I. J. M., Perraud, J. M., Collins, D., Henderson, B., Jin, H., and McJannet, D. L.: Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment, J. Hydrol., 519, 2747–2762, 2014.
Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, J. Climate, 24, 3624–3648, 2011.
Sahoo, A. K., De Lannoy, G. J., Reichle, R. H., and Houser, P. R.: Assimilation and downscaling of satellite observed soil moisture over the Little River Experimental Watershed in Georgia, USA, Adv. Water Resour., 52, 19–33, 2013.
Schellekens, J.: OpenStreams wflow documentation release 1.0RC1, Deltares, available at: http://wflow.readthedocs.org/en/latest/ (last access: 25 September 2015), 2014.
Su, C. H., Ryu, D., Young, R. I., Western, A. W., and Wagner, W.: Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia, Remote Sens. Environ., 134, 1–11, 2013.
Sun, C., Walker, J. P., and Houser, P. R.: A methodology for snow data assimilation in a land surface model, J. Geophys. Res.-Atmos., 109, D08108, doi:10.1029/2003JD003765, 2004.
Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin, Hydrol. Earth Syst. Sci., 15, 2913–2935, doi:10.5194/hess-15-2913-2011, 2011.
Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.: Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, Hydrol. Earth Syst. Sci., 19, 2079–2100, doi:10.5194/hess-19-2079-2015, 2015.
Van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Tech. Rep., Department of Physical

Geography, Utrecht University, Utrecht, the Netherlands, available at: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf (last access: 28 February 2011), 2009.
Van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Global monthly water stress: I. Water balance and water availability, Water Resour. Res., 47, W07517, doi:10.1029/2010WR009791, 2011.
van Dijk, A. I. J. M.: The Australian water resources assessment system, Version 0.5, available at: http://www.clw.csiro.au/publications/waterforahealthycountry/2010/wfhc-awras-evaluation-against-observations.pdf (last access: 8 December 2015), 2010.
van Dijk, A. I. J. M. and Renzullo, L. J.: Water resource monitoring systems and the role of satellite observations, Hydrol. Earth Syst. Sci., 15, 39–55, doi:10.5194/hess-15-39-2011, 2011.
van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y., and Tregoning, P.: A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble, Hydrol. Earth Syst. Sci., 18, 2955–2973, doi:10.5194/hess-18-2955-2014, 2014.
Vertessy, R. A. and Elsenbeer, H.: Distributed modeling of storm flow generation in an Amazonian rain forest catchment: effects of model parameterization, Water Resour. Res., 35, 2173–2187, 1999.
Vörösmarty, C. J., Fekete, B. M., Meybeck, M., and Lammers, R.: A simulated topological network representing the global system of rivers at 30 min spatial resolution (STN-30), Global Biogeochem. Cy., 14, 599–621, 2000.
Vrugt, J. A., Gupta, H. V., Nualláin, B., and Bouten, W.: Real-time data assimilation for operational ensemble streamflow forecasting, J. Hydrometeorol., 7, 548–565, 2006.
Wanders, N., Karssenberg, D., Bierkens, M., Parinussa, R., de Jeu, R., van Dam, J., and de Jong, S.: Observation uncertainty of satellite soil moisture products determined with physically-based modeling, Remote Sens. Environ., 127, 341–356, doi:10.1016/j.rse.2012.09.004, 2012.
Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrol. Earth Syst. Sci., 18, 2343–2357, doi:10.5194/hess-18-2343-2014, 2014a.
Wanders, N., Bierkens, M. F., de Jong, S. M., de Roo, A., and Karssenberg, D.: The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models, Water Resour. Res., 50, 6874–6891, doi:10.1002/2013WR014639, 2014b.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-Interim reanalysis data, Water Resour. Res., 50, 7505–7514, 2014.

Wesseling, C. G., Karssenberg, D. J., Burrough, P. A., and Deursen, W.: Integrating dynamic environmental models in GIS: the development of a dynamic modelling language, Trans. GIS, 1, 40–48, 1996.

Wu, H. and Li, Z.-L.: Scale issues in remote sensing: a review on analysis, processing and modeling, Sensors, 9, 1768–1793, 2009.

Zaitchik, B. F., Rodell, M., and Reichle, R. H.: Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi River basin, J. Hydrometeorol., 9, 535–548, 2008.

Table 1. Catchment daily mean (μ)/standard deviation (σ) of precipitation (P), temperature (T) and reference evapotranspiration (ET) in the Murrumbidgee river basin for 2007–2010.

	P (mm day^{-1})		T (°C)		ET (mm day^{-1})	
	Local μ/σ	Global μ/σ	Local μ/σ	Global μ/σ	Local μ/σ	Global μ/σ
2007	1.38/3.22	1.36/3.19	19.73/7.05	16.02/6.66	3.98/2.13	3.18/1.62
2008	1.24/3.19	1.26/3.32	18.89/6.65	15.18/6.24	3.88/2.06	3.00/1.51
2009	1.16/2.48	1.15/2.33	19.82/7.46	16.07/6.76	4.08/2.33	3.24/1.83
2010	2.35/5.46	2.45/5.19	18.50/6.91	15.03/6.46	3.69/2.04	2.95/1.56

Table 2. Hydrometeorological and geographical information of analysed catchments at Murrumbidgee river basin.

Station number	Station name	Basin area (km^2)	Outlet location		Mean flow ($m^3 s^{-1}$)
			Longitude	Latitude	
Assimilation					
410088	Goodradidgee River at Brindabella	419.66	148.73	−35.42	3.719
410062	Numeralla River at Numeralla School	691.38	149.35	−36.18	1.164
410024	Goodradigbee River at Wee Jasper (Kashmir)	1050.60	148.69	−35.17	5.805
410044	Muttama Creek at Coolac	1058.49	148.16	−34.93	1.072
410033	Murrumbidgee River at Mittagang Crossing	1809.84	149.09	−36.16	2.228
410761	Murrumbidgee River below Lobbs Hole Creek	9332.28	149.10	−35.54	5.463
410130	Murrumbidgee river at D/S Balranald weir	28 651.21	143.49	−34.67	10.113
410023	Murrumbidgee river at D/S Berembed weir	34 133.07	146.84	−34.88	38.449
410021	Murrumbidgee river at Darlington Point	37 804.78	146.00	−34.57	29.899
410040	Murrumbidgee river at D/S Maude weir	43 110.97	144.30	−34.48	16.888
Evaluation					
410107	Mountain Creek at Mountain Creek	140.54	148.84	−35.03	1.078
410705	Molonglo River at Burbong	350.22	149.31	−35.34	0.591
410048	Kyeamba Creek at Ladysmith	350.30	147.53	−35.20	2.41
410038	Adjungbilly Creek at Darbalara	390.89	148.25	−35.02	1.494
410734	Queanbeyan River at Tinderry	557.73	149.35	−35.61	0.893
410057	Goobarragandra River at Lacmalac	559.77	148.35	−35.33	4.964
410026	Yass River at Yass	1226.98	148.91	−34.84	1.198
410091	Billabong Creek at Walbundrie	2859.77	146.72	−35.69	2.113
410050	Murrumbidgee River at Billilingra	3353.91	149.13	−35.98	3.836
410001	Murrumbidgee River at Wagga Wagga	39 856.21	147.37	−35.10	59.638
410005	Murrumbidgee River at Narrandera	45 321.40	146.55	−34.76	48.074
410078	Murrumbidgee River at Carrathool	69 854.30	145.42	−34.35	27.057
410136	Murrumbidgee River at D/S Hay Weir	73 241.50	144.71	−34.52	25.243

Table 3. Data assimilation scenarios including abbreviations, forcing data, hydrological model and assimilated observations.

Identifier	DA scenarios		
	Forcing data	Hydrological model	Assimilated observations
L_OL	Local (AWAP)	PCR-GLOBWB	Open Loop (None)
L_Q			Discharge stations
L_SM			AMSR-E soil moisture
L_SM_Q			Discharge stations and AMSR-E soil moisture
L_w		OpenStreams wflow_sbm (OSWS)	None
G_OL	Global (WFDEI)	PCR-GLOBWB	Open Loop (None)
G_Q			Discharge stations
G_SM			AMSR-E soil moisture
G_SM_Q			Discharge stations and AMSR-E soil moisture
G_w		OpenStreams wflow_sbm (OSWS)	None

Table A1. Evaluation results of the catchment daily means of soil moisture estimates.

	Local			Global		
	RMSE [m^3m^{-3}]	MAE [m^3m^{-3}]	r [–]	RMSE [m^3m^{-3}]	MAE [m^3m^{-3}]	r [–]
w	0.07571	0.05716	0.79865	0.07854	0.06082	0.77603
OL	0.09735	0.07520	0.43071	0.10094	0.07662	0.40715
Q	0.09738	0.07523	0.43041	0.10095	0.07664	0.40652
SM	0.09085	0.06699	0.53452	0.09372	0.06811	0.50230
SM_Q	0.09032	0.06609	0.52085	0.09302	0.06845	0.49623

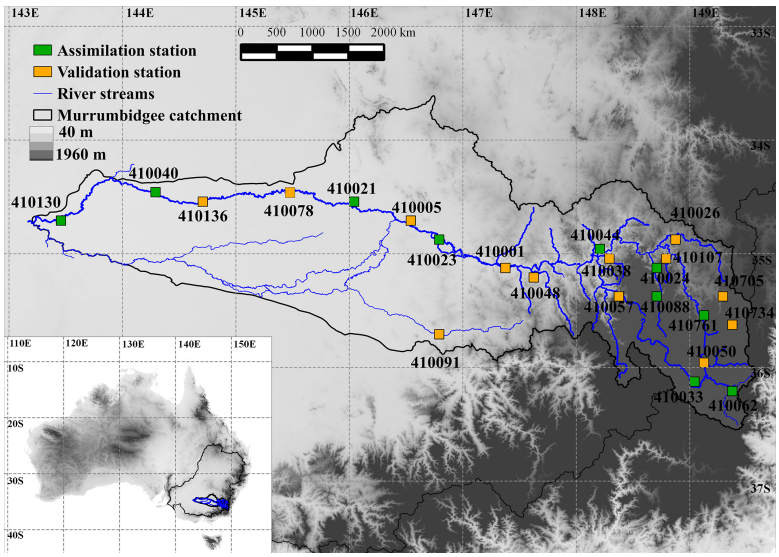

Figure 1. Map of the Murrumbidgee river basin and its location in Australia as part of the Murray–Darling system. Green squares indicate locations for assimilation of streamflow observations and orange squares indicate locations for evaluation of streamflow observations. Each streamflow location is identified with a gauging station number according to BoM (2015).

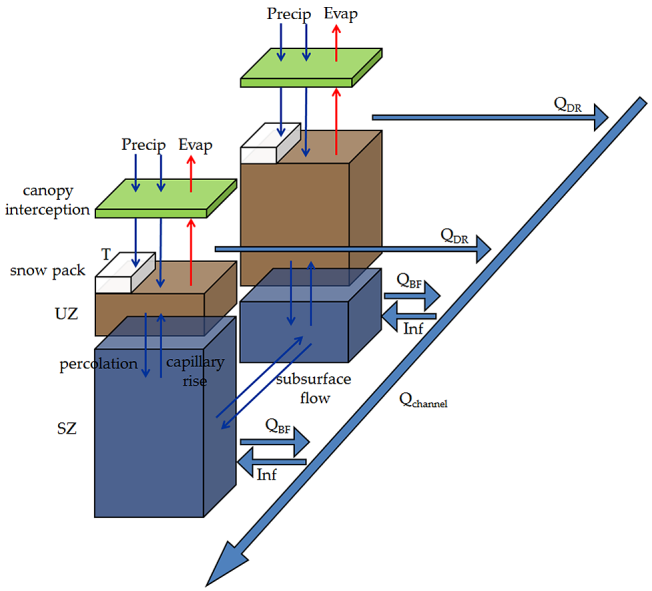

Figure 2. OpenStreams wflow_sbm model structure, adapted from Vertessy and Elsenbeer (1999) and Schellekens (2014). Symbols definition are as follows: Precip, precipitation; Evap, evaporation; T, Temperature; UZ, unsaturated zone; SZ, saturated zone; Qchannel, total runoff; QDR, direct runoff; QBF, baseflow and Inf, water flow from the river channel to the saturated zone.

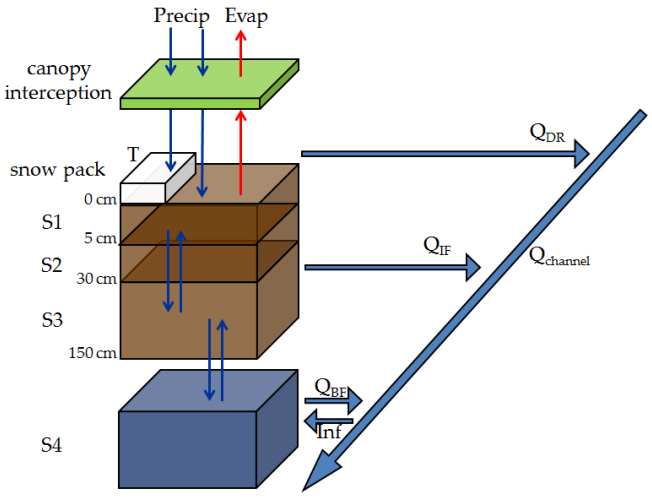

Figure 3. PCR-GLOBWB model structure, adapted from Van Beek et al. (2011). Symbols definition are as follows: Precip, precipitation; Evap, evaporation; T, Temperature; S1, first soil layer; S2, second soil layer; S3 third soil layer; S4, groundwater reservoir; Qchannel, total runoff; QDR, direct runoff; QIF, intermediate flow; QBF, baseflow and Inf, water flow from the river channel to the groundwater reservoir.

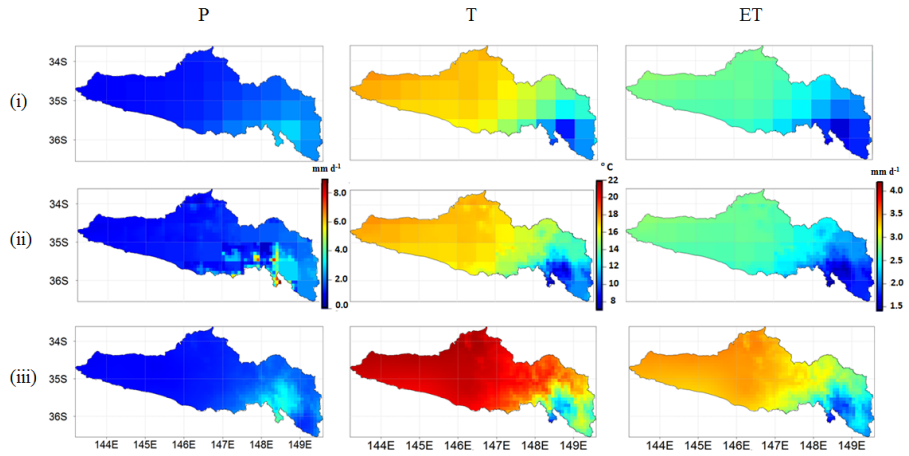

Figure 4. Daily mean precipitation, temperature and reference evapotranspiration for the time period 2007–2010 from the (i) global (0.5° × 0.5°), (ii) downscaled global (0.08° × 0.08°) and (iii) local (0.05° × 0.05°) forcing data sets.

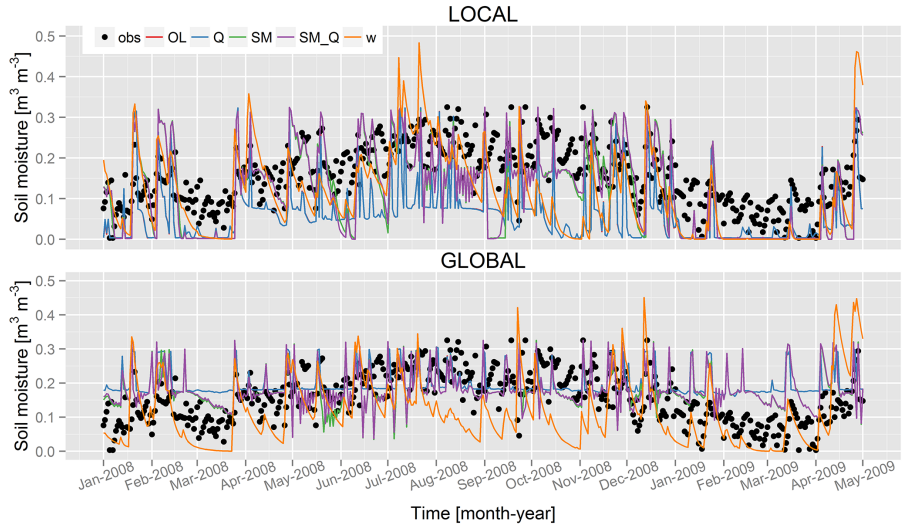

Figure 5. Simulated and observed soil moisture estimates at 410057 gauging station in a tributary of the Murrumbidgee river for the time period January 2008–May 2009. The upper panel shows soil moisture time series when local data is used as model forcing. Soil moisture time series obtained with the global forced models are shown in the lower panel. Each panel contains results for each data assimilation scenario plotted with different colours lines (OL – red, Q – blue, SM – green and SM_Q – purple) and downscaled AMSR-E observations with black points.

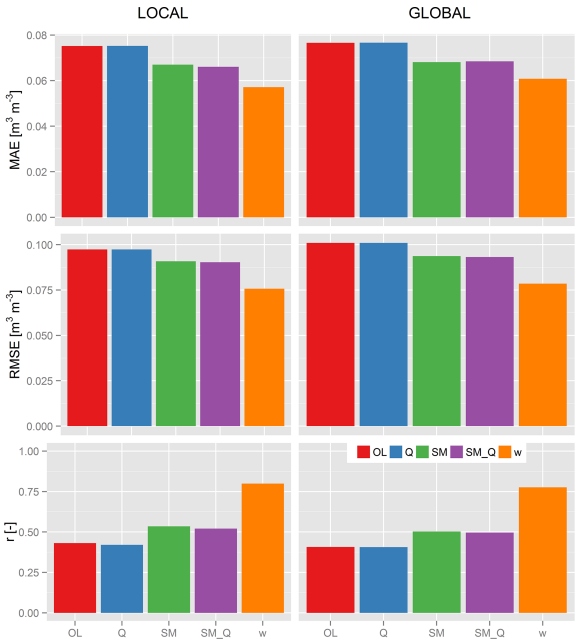

Figure 6. Evaluation results of the catchment daily means of soil moisture in the Murrumbidgee river basin. In the rows, three different evaluation metrics are shown; from top to bottom these are the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE) and the Pearson's correlation coefficient (r). Columns show various forcing data: local and global. (For clarity, the exact values are included in Table A1.)

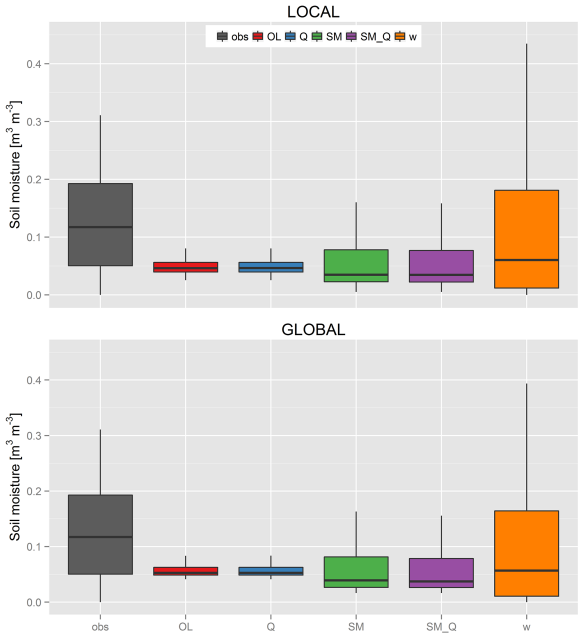

Figure 7. Boxplots of the catchment daily means of soil moisture in the Murrumbidgee river basin. The upper panel shows soil moisture when local data is used as model forcing. Soil moisture obtained with the global forced models is shown in the lower panel. Boxplots of each panel illustrate the first and third quantile ranges (box), the median (dark line) and the maximum–minimum range (whiskers) of soil moisture estimates.

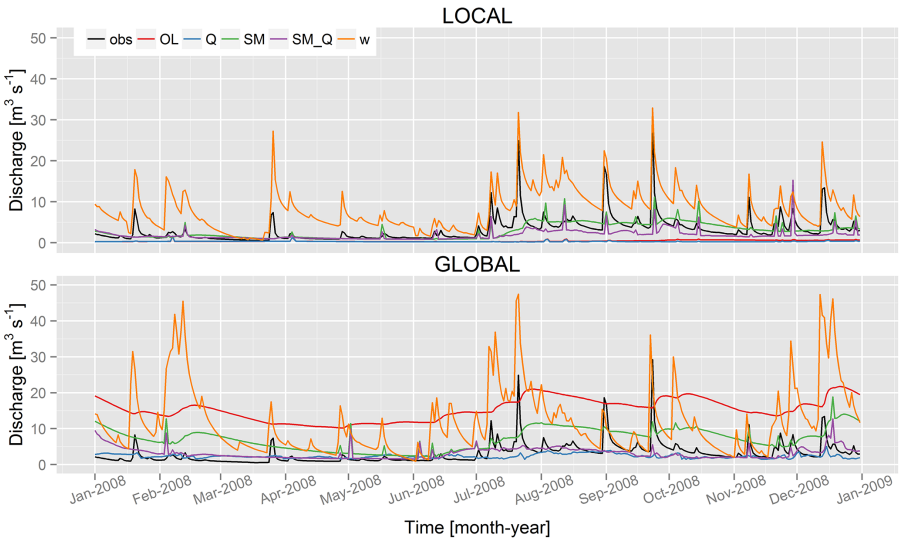

Figure 8. Simulated and observed streamflow estimates at 410057 gauging station in a tributary of the Murrumbidgee river for the time period January 2008–January 2009. The upper panel shows streamflow when local data is used as model forcing. Streamflow obtained with the global forced models is shown in the lower panel. Each panel contains results for each data assimilation scenario and the observed streamflow estimates plotted with different colours lines (OL – red, Q – blue, SM – green, SM_Q – purple and obs – black). The ensemble mean is given for each data assimilation scenario.

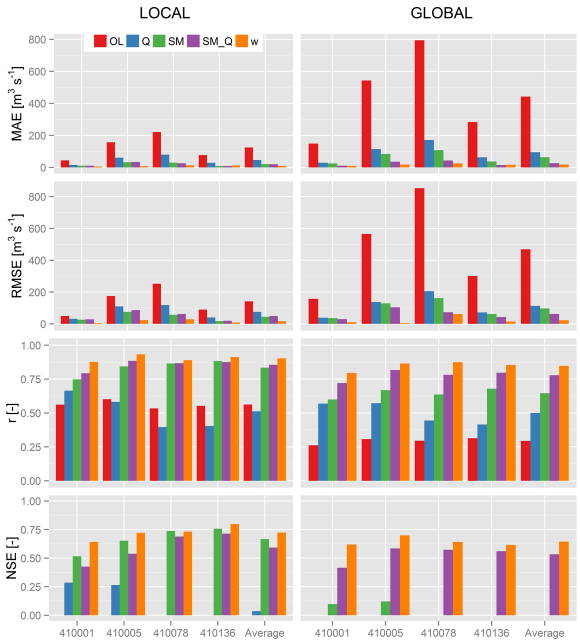

Figure 9. Evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee river. Average values calculated across those locations are shown in the rightmost bar of each histogram. In the rows, four diferent evaluation metrics are shown; from top to bottom these are the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), the Pearson's correlation coefficient (r) and the Nash Sutcliffe efficiency. Columns show various forcing data: local and global.

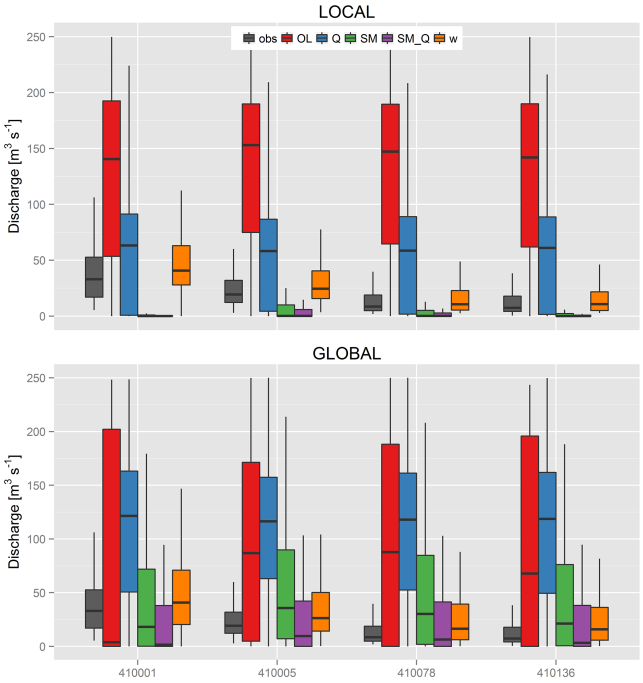

Figure 10. Boxplots of streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee river. The upper panel shows streamflow when local data is used as model forcing. Streamflow obtained with the global forced models is shown in the lower panel. Boxplots of each panel illustrate the first and third quantile ranges (box), the median (dark line) and the maximum–minimum range (whiskers) of streamflow estimates.