Wu, L., Curceac, S., Atkinson, Peter, Milne, A. E. and Harris, P. 2021. A case study on the effects of data temporal resolution on the simulation of water flux extremes using a process-based model at the grassland field scale. *Agricultural Water Management.* 255. https://doi.org/10.1016/j.agwat.2021.107049

A case study on the effects of data temporal resolution on the simulation of water flux extremes using a process-based model at the grassland field scale

Lianhai Wu [a,*], Stelian Curceac [a], Peter M. Atkinson [b, c, d], Alice Milne [e], Paul Harris [a]

[a] Rothamsted Research, Department of Sustainable Agriculture Sciences, North Wyke EX20 2SB, Devon, UK.
[b] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK.
[c] Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
[d] State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China.
[e] Rothamsted Research, Department of Sustainable Agriculture Sciences, Harpenden AL5 2JQ, UK

* Corresponding author: Lianhai Wu

Email: lianhai.wu@rothamsted.ac.uk

Abstract

Projected changes to rainfall patterns may exacerbate existing risks posed by flooding. Furthermore, increased surface runoff from agricultural land increases pollution through nutrient losses. Agricultural systems are complex because they are managed in individual fields, and it is impractical to provide resources to monitor their water fluxes. In this respect, modelling provides an inexpensive tool for simulating fluxes. At the field-scale, a daily time-step is used routinely, however, it was hypothesized that a finer time-step will provide more accurate identification of peak fluxes. To investigate this, the process-based SPACSYS model that simulates water fluxes, soil carbon and nitrogen cycling as well as plant growth with a daily time-step was adapted to provide sub-daily simulations. As a case study, the water flux simulations were checked against a 15-minute measured water flux dataset from April 2013 to February 2016 from a pasture within a monitored grassland research farm, where the data were up-scaled to hourly, 6-hourly and daily. Analyses were conducted with respect to model performance for: (a) each of the four data resolutions, separately (15-minute measured versus 15-minute simulated; hourly measured versus hourly simulated; etc.); and (b) at the daily resolution only, where 15-minute, hourly and 6-hourly simulations were each aggregated to the daily scale. Comparison between measured and simulated fluxes at the four resolutions revealed that hourly simulations provided the smallest misclassification rate for identifying water flux peaks. Conversely, aggregating to the daily scale using either 15-minute or hourly simulations increased accuracy, both in prediction of general trends and identification of peak fluxes. For the latter investigation, the improved identification of extremes resulted in 9 out of 11 peak flow events being correctly identified with only 2 false positives, compared with 5 peaks being identified with 4 false positives of the usual daily simulations. Increased peak flow detection accuracy has the potential to provide clear field management benefits in reducing nutrient losses to water.

Key words: SPACSYS; extreme flows; North Wyke Farm Platform; scale effects; grassland;

1 Introduction

Flooding in the UK puts more than 5 million people in 2.4 million properties at risk each year (Environment Agency, 2009). Projected changes to rainfall patterns (Watts and Anderson, 2016) may exacerbate the existing risks posed by flooding. Flash flooding or surface water flooding, defined as those flood events where the rise in water is either during or within a few hours of the rainfall that produces the rise, is one of the most common types of flooding in the UK. The utilised agricultural area, of which almost 60% is permanent grassland, covers 71% of the total land of the UK (Department for Environment, Food and Rural Affairs, 2019). Water fluxes or surface runoff generated from agricultural land can contribute significantly to local floods and nutrient losses that cause water pollution. Flooding of farmland is likely to become more frequent in some areas under projected climate change (Brown et al., 2016), although intriguingly, studies have found increases in precipitation extremes do not necessarily mean increases in flood magnitude, due to decreased soil moisture at storm onset and reduced storm durations (Sharma et al., 2018; Wasko et al., 2019). Further, soil erosion is accelerating due to more intense rainfall, leading to the loss of valuable topsoil and the pollution of watercourses (Morison and Matthews, 2016).

Accurate forecasting of water runoff (or water fluxes) from agricultural land is, therefore, not only a vital component of flood early-warning systems, but also for associated management strategies for nutrient loss and water pollution. Water fluxes from the soil surface are controlled by soil properties. Long-term hydrological studies have shown that sandy Alfisols can generate higher runoff compared to clayey Vertisols (Pathak et al., 2013), and a greater risk of flooding on clay soils has been reported (Charlton et al., 2010). The wetness of the soil before a precipitation event (Merz and Plate, 1997) and soil compaction also affect water fluxes. Farm machinery and livestock (Adimassu et al., 2019; Alaoui et al., 2018; Newell Price et al., 2012) can cause serious compaction and so exacerbate flood risk. Natural events, particularly long and intense precipitation events (Archer and Fowler, 2018), and land cover variation (Dadson et al., 2017; Keesstra et al., 2018) also affect flux.

Agricultural systems are complex because they are generally managed at the field scale and each field has its own unique set of soil conditions and topology. Monitoring water surface fluxes in fields is costly both in time and financially. In this respect, modelling provides an effective tool for simulating or forecasting water fluxes. The SPACSYS model (Wu et al., 2007) is one such process-based model. It is a field scale and weather-driven dynamic simulation model. Since it was first published in 2007, it has been developed to provide added functionality (Bingham and Wu, 2011; Liu et al., 2013; Wu et al., 2019; Wu et al., 2015). The model can simulate the interactions of soil carbon (C), nitrogen (N) and phosphorus (P), plant growth and development, water re-distribution and heat transformation in agricultural fields. The model has been used to investigate several issues including resource use efficiency by crops (Wu et al., 2009), greenhouse gas (GHG) emissions (Abalos et al., 2016; Perego et al., 2016), the responses of cropping/grassland systems to environmental change (Wu et al., 2016) and the forecasting of crop yield and stocks of C and nutrients (Zhang et al., 2016) under various climatic and soil conditions.

The SPACSYS model has been developed to investigate not only temporal dynamics, but also within-field spatial variation in processes such as water runoff, using a linked, grid-based approach (grid-to-grid) (Liu et al., 2018). As in all previous implementations of SPACSYS, and common to many agriculture-focused models (Ahuja et al., 2002), a daily time-step was used. However, model predictions of water flux did not increase in accuracy when considering grid connectivity. We hypothesise, that a finer time-step might provide this improvement instead; not only in the grid-to-grid model, but also in the (non-grid-to-grid) standard model, as investigated here. Although not demonstrated within this study, increasing the accuracy of water flux simulations should implicitly increase the accuracy of associated SPACSYS simulations, such as those for nutrient loss that use predicted water flux in their calculation.

For our case study, we used measured 15-minute water flux data from one field (or sub-catchment) of the North Wyke Farm Platform (NWFP). The NWFP is a systems scale research facility in the south-west of England for investigation of the sustainability of lowland ruminant production systems (Orr et al., 2016). South-west England has a relatively wet climate where the greatest rainfall is in winter and the driest times are between April to July. August tends to show an increase in rainfall over July and starts the inexorable rise in rainfall into autumn and early winter. More recently, the number of flood events has increased (Stevens et al., 2016), mostly in the autumn and winter months; all as a likely consequence of increased surface water runoff (Palmer and Smith, 2013).

For this study, the NWFP’s 15-minute water flux data were up-scaled to hourly, 6-hourly and daily data and the SPACSYS model was adapted to provide corresponding downscaled simulations at 15-minute, hourly and 6-hourly resolutions (in addition to its usual daily output). This provided four measured water flux datasets and four simulated water flux datasets over a study period of 34 months (April 2013 to February 2016). Simulations were generated using the same field management practices and parameter configurations. These rich water flux datasets enabled investigation of the effects of temporal scale on model performance not only in terms of extreme water runoff, which is the study focus and provides it's novelty, but also in terms of general trends.

2 Materials and Methods

2.1 Model description

The SPACSYS model includes a plant growth and development component, an N cycling component, a C cycling component, a P cycling component, plus a soil water component that includes representation of water flow to field drains as well as downwards through the soil layers, together with a heat transfer component. The equations to quantify such different processes have been described elsewhere (Liu et al., 2013; Wu et al., 2019; Wu et al., 2007; Wu et al., 2015). Here, only the processes influencing directly the soil water component are presented.

For SPACSYS, the Richard's equation for water potential and Fourier's equation for temperature are used to simulate water and heat fluxes, which are inherited from the SOIL model (Jansson, 1998). If the water content in a layer rises above a specified value a proportion is held in macropores such that rapid downward water movement takes place due to gravitational forces alone. Water flow from the soil profile to a drainage pipe occurs when the ground water table is above the bottom level of the pipe and the soil below the ground water table is saturated. The Hooghoudt drainage flow equation with modification is adopted for the subsurface drainage flow.

The main processes concerning plant growth in SPACSYS are plant development, assimilation, respiration, root growth and development, water uptake, nutrient uptake, biological N fixation for legume plants and partitioning of photosynthate and nutrients from uptake estimated with various mechanisms implemented in the model. N cycling coupled with C cycling covers the transformation processes for organic matter and inorganic N. The main processes and transformations causing size changes to mineral N pools are mineralization, nitrification, denitrification including N gaseous emission and plant N uptake. P cycling is linked to other components such as the plant component, heat transformation and the water cycle. Organic P is subdivided into certain sub-pools with different forms which are connected with transformation rates.

2.2 The North Wyke Farm Platform

The study site is located in south-west England, at the NWFP, Rothamsted Research, Okehampton, Devon (50°46’10’’N, 30°54’05’’W). For the period 1985-2015, the mean annual temperature in North Wyke ranges between 6.8 and 13.4 °C, the mean annual rainfall is 1033 mm and the climate is classed as cool temperate. The platform is a 63 ha systems-based experimental facility divided into 15 hydrologically isolated sub-catchments across three 21 ha farmlets with five sub-catchments in each. At the time of this study, all three farmlets were used solely for grazing livestock research (sheep and cattle) where each farmlet was operating under different sward management strategies: no re-seeding (permanent pasture); re-seeded monoculture; and re-seeded legume mix. The platform monitors routinely water runoff and water chemistry in each of the 15 sub-catchments, together with other primary data collections (e.g. greenhouse gas emissions, livestock performance) so that each farming system can be described, contrasted and compared according to its level of sustainability (Orr et al., 2016). Datasets are freely available from https://www.rothamsted.ac.uk/north-wyke-farm-platform, including those used in this study.

2.3 Model configuration

For this study, we focused on water fluxes for one sub-catchment in the permanent pasture farmlet called 'Golden Rove'; a single field that has been under permanent pasture since the outset of the platform in 2010 (Fig. Figure 1). The soil class for this field is primarily Halstow, which comprises a slightly stony clay loam topsoil (approximately 36% clay) that overlies a mottled stony clay (approximately 60% clay), derived from underlying Carboniferous Culm rocks (Harrod and Hogan, 2008). The study field also has a smaller, but not insignificant area of Denbigh-Cherubeer soil class. In the simulations, the soil type was ignored.

To mimic the grazing system, daily grass intake and excretion of sheep and cattle in the field was estimated before running the simulations (Carswell et al., 2019; Wu et al., 2016). Soil physical and chemical properties of the field were adopted from a previous study of the same field (Wu et al., 2016). The temporal frequency for the measured water fluxes (l s^{-1}) from a NWFP water flume has been 15-minutes since the outset of the platform's setup in October 2012. However, meteorological measurements at the same 15-minute resolution were only initiated from 30 April 2013. Thus, to ensure consistency in the frequency of the driving variables and the water flux as an output variable, simulations also started from 30 April 2013. An end-date of 15 February 2016 was chosen to give an interrupted data collection time period of 34 months. A longer time period would entail having significant periods of missing data due to instrument failure (i.e. there were no measurements on water flux between 15 February and 24 October 2016). A previous scale-focused study, analysing the measured 15-minute water fluxes together with aggregations at 30-minute, hourly, 3-hourly, 6-hourly, 12-hourly and daily resolutions, indicated that 15-minute, hourly, 6-hourly and daily resolutions is sufficient to communicate all key outcomes adequately (Curceac et al., 2020). Thus, the same four temporal resolutions were adopted for the model simulations of this study.

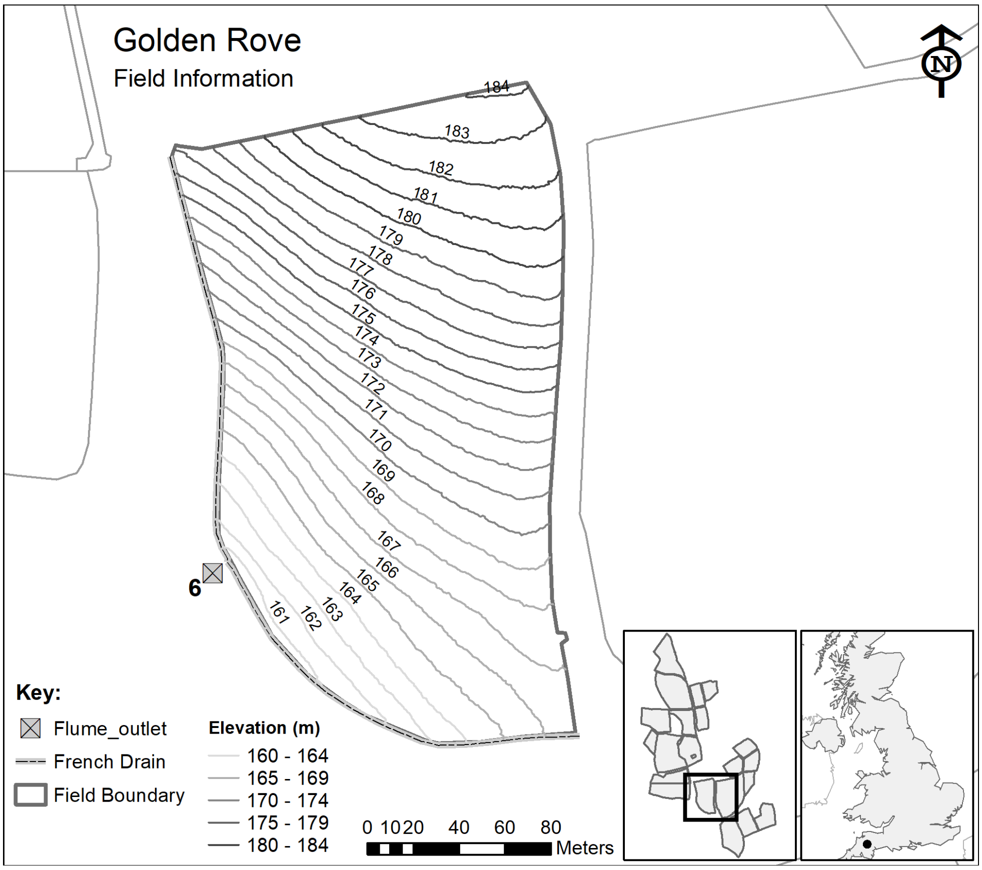

Figure 1. Details of the NWFP sub-catchment selected for this study (sub-catchment number 6 of 15, consisting of a single field called Golden Rove).

2.4 Statistical analysis

Two distinct sets of statistical analyses were conducted with respect to model performance and data resolution: (a) model performance for each of the four data temporal resolutions, separately (i.e. 15-minute measured versus 15-minute simulated; hourly measured versus hourly simulated; 6-hourly measured versus 6-hourly simulated; daily measured versus daily simulated); and (b) model performance conducted at the daily temporal resolution only, where 15-minute, hourly and 6-hourly simulations were each aggregated to the daily scale. The latter analyses provide valuable insights into the worth of using fine temporal resolution data to increase the accuracy of daily simulations, especially with respect to the accurate identification of extremes. This is important as many process-based models in the literature simulate only at a daily time-step (e.g. Del Grosso et al., 2009).

2.4.1 Model performance graphics

Model performance graphics consist of time-series plots, density plots and scatterplots of measured and simulated datasets. For the latter, the ideal 1:1 line, a linear regression fit, and a non-linear regression fit (i.e., a Loess smoother fit; Cleveland, 1979) are given where he estimated intercept and slope parameters from the linear fit should equal zero and one for perfect model simulations, respectively. Results (p-values) from a linear hypothesis test are reported comparing this ideal model with the estimated model using a finite sample F test (see Fox, 2016). The non-linear regression provides added insight into where the simulated values tend to over- or under-predict (e.g., at measured low or high values, respectively). Time-series plots for the errors (i.e. measured minus simulated data) are also given.

2.4.2 Model performance indices

To further assess the accuracy of the simulations, six accuracy indices were calculated: the mean absolute error (MAE), the normalized root mean square error (NRMSE), the percentage bias (PBIAS), the Nash-Sutcliffe efficiency (NSE), the index of agreement (*d*) and the Kling-Gupta efficiency (KGE), as given in Table 1.

Table 1. Accuracy indices formulae, where $\hat{z}_i$ are the simulated values, z_i are the measured values, $\bar{z}_i$ is the mean of the measured values, r is the Pearson product-moment correlation coefficient (between measured and simulated) and σ is the standard deviation.

Index form	Index formula	Min.	Max.	Ideal
Error	$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} \lvert \hat{z}_i - z_i \rvert$	0	∞	0
Error	$\mathrm{NRMSE} = 100\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{z}_i - z_i)^2}}{z_{max} - z_{min}}$	0	∞	0
Error	$\mathrm{PBIAS} = 100\frac{\sum_{i=1}^{N}(\hat{z}_i - z_i)}{\sum_{i=1}^{N} z_i}$	0	∞	0
Agreement	$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(\hat{z}_i - z_i)^2}{\sum_{i=1}^{N}(z_i - \bar{z}_i)^2}$	-∞	1	1
Agreement	$d = 1 - \frac{\sum_{i=1}^{N}(\hat{z}_i - z_i)^2}{\sum_{i=1}^{N}(\lvert \hat{z}_i - \bar{z}_i \rvert + \lvert z_i - \bar{z}_i \rvert)^2}$	0	1	1

Agreement	$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{\hat{z}}}{\sigma_z} - 1\right)^2 + \left(\frac{\bar{\hat{z}}}{\bar{z}} - 1\right)^2}$	$-\infty$	1	1

2.4.3 Simulation accuracy of measured peaks

To investigate model accuracy in simulating water flux peaks, a threshold at the 99th percentile of each measured water flux dataset was used to identify peak flows. Model simulations were then assessed to determine if they could similarly exceed this threshold coinciding with a measured exceedance. Incidences of correct peak flow simulations, false negatives (simulation does not exceed threshold when measured flow does), false positives (simulation exceeds threshold when measured flow does not) and corresponding Kappa values are reported. The Kappa statistic provides a measure of agreement beyond the level of agreement expected by chance alone. General guidelines for Kappa values are as follows: less than 0.2 slight agreement, 0.2 to 0.4 fair agreement, 0.4 to 0.6 moderate agreement, 0.6 to 0.8 substantial agreement, greater than 0.8 almost perfect agreement, and equal to 1 perfect agreement.

3 Results

3.1 Model performance for each of the four data temporal resolutions, separately

Comparisons between the measured and simulated water flux rates at different temporal resolutions are shown in Fig. Figure 2. Visually, it appears that simulations of daily and 6-hourly water fluxes tend to under-predict the measured data, often missing high peaks, while simulations of 15-minute and hourly data possibly tend to over-predict. However, the scatterplots of the measured and simulated data, together with the ideal 1:1 line, a linear regression fit, and a Loess smoother fit (Fig. Figure 3) present a more complete picture. Simulations for all four temporal resolutions clearly tend to over-predict, with the level of over-prediction increasing as the resolution increases. Over-prediction is shown with each linear regression fit lying below the 1:1 line; and increasingly so, as the resolution increases. All linear regression fits were found to be significantly different to the 1:1 line, each with *F*-test *p*-values < 0.0001.

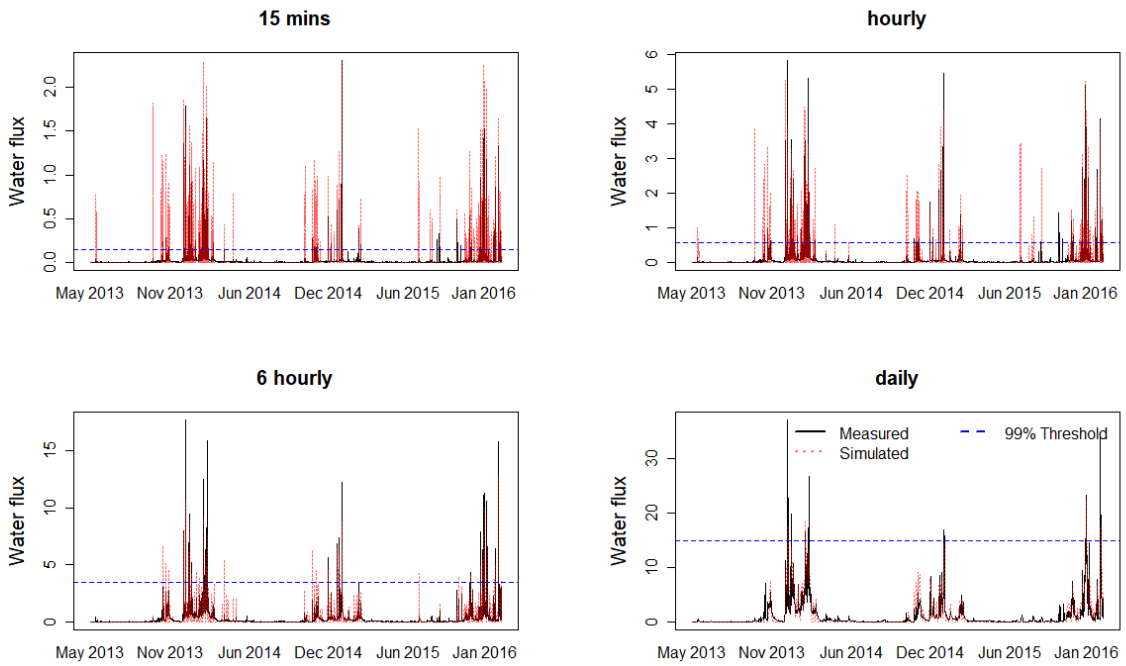

Figure 2. Time-series plots for measured and simulated water flux data (not aggregated) for 15-minute, hourly, 6-hourly and daily data (in units of mm 15min^{-1}, mm h^{-1}, mm 6h^{-1} and mm d^{-1}, respectively). All plots are shown with a threshold at the 99th percentile of measured data (at 0.138 mm 15min^{-1}, 0.553 mm h^{-1}, 3.45 mm 6h^{-1} and 14.9 mm d^{-1}, respectively).

For all four temporal resolutions, the tendency to over-predict decreases at the largest measured water fluxes, as shown by the concave behaviour of the loess smoother fit (Fig.

Figure 3), with daily simulations tending to under-predict at very large fluxes, thus, missing extreme events that may cause flooding and associated nutrient and sediment losses. Clearly, 'smoothing bias' increases as temporal resolution decreases. The 15-minute simulations maintain the variation shown in the measured data (i.e. observations range from 0 to 2.306 mm $15min^{-1}$ while simulations range from 0 to 2.310 mm $15min^{-1}$), while the daily simulations do not (i.e. observations range from 0 to 36.97 mm d^{-1} while simulations only range from 0 to 22.20 mm d^{-1}). As each 'simulation-to-observation' comparison is on a different scale, it is not useful to present further model fit diagnostics, such as error and agreement indices.

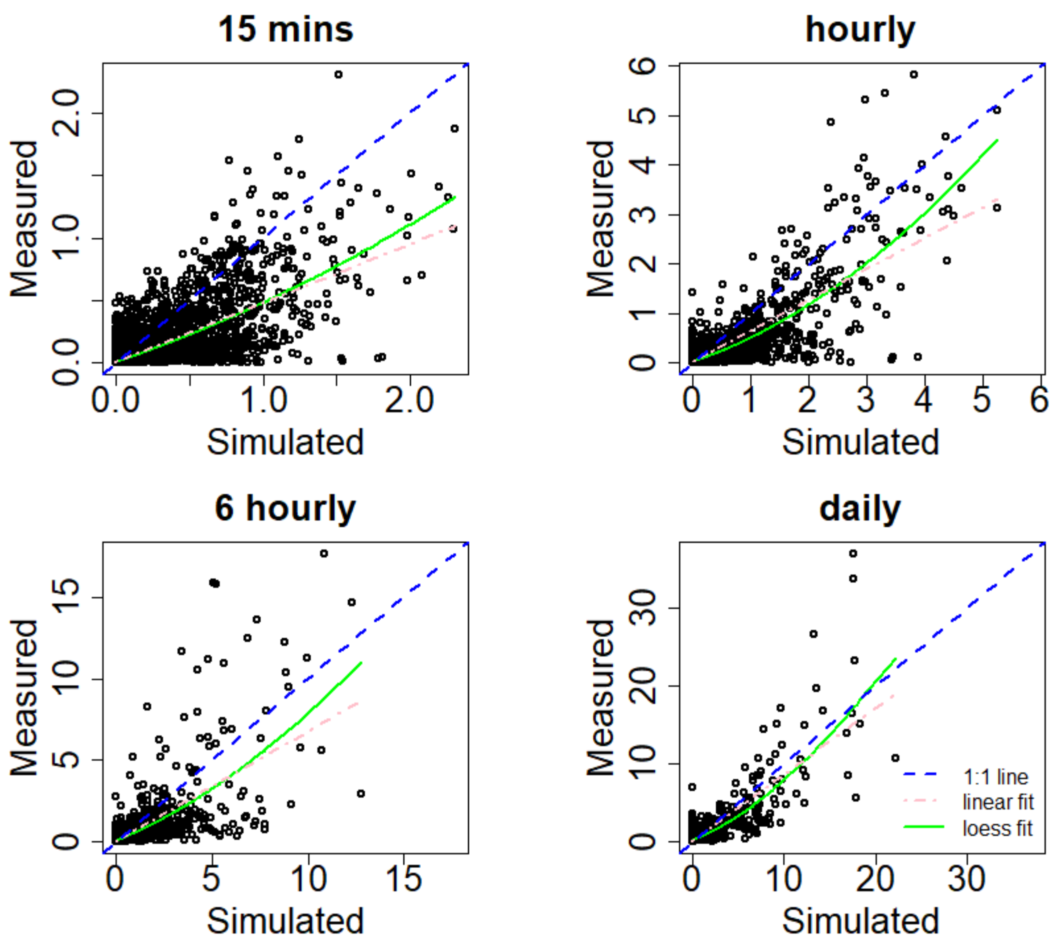

Figure 3. Scatterplots of the measured and simulated data (not aggregated) for 15-minute, hourly, 6-hourly and daily data. Scatterplots are shown with the 1:1 line, a linear regression fit and a loess smoother fit. Units are in mm 15min^{-1}, mm h^{-1}, mm 6h^{-1} and mm d^{-1}, respectively.

3.2 Model performance for simulations aggregated to the daily scale

Comparisons between the measured and simulated water flux rates aggregated to the same daily scale are shown in Fig. Figure 4. There are clear instances of both over- and under-prediction for all four daily outputs. The scatterplots (Fig. Figure 5) of daily measured and daily simulated data from different aggregations, together with the 1:1 line, a linear fit, and a loess smoother fit, again provide a clear visualisation of the relations in the time-series plots. Simulations for all three aggregations to daily (15-minute, hourly and 6-hourly) again tend to over-predict (as their linear fits lie below the 1:1 line), but this over-prediction is broadly similar across the four datasets, and not as great as that found with the unaggregated data, above. The 6-hourly aggregations appear to be the least accurate. Again, all linear regression fits were found to be significantly different to the 1:1 line, each with *F*-test *p*-values < 0.0001.

In this instance, ‘smoothing boas’ increases as aggregation resolution decreases, where simulations for 15-minute and hourly aggregations both increase the variation shown in the measured daily data (i.e. 0 to 36.97 mm d^{-1}); with 15-minute daily aggregations ranging from 0 to 38.94 mm d^{-1} and hourly daily aggregations ranging from 0 to 39.64 mm d^{-1}. Conversely, the 6-hourly daily aggregations and the daily simulations reduce variation with the 6-hourly daily aggregations ranging from 0 to 31.70 mm d^{-1} and the (unaggregated) daily simulations ranging from 0 to 22.20 mm d^{-1}. In summary, daily simulations based on component 15-minute and hourly aggregations have the potential to identify peak water fluxes (and, thus, flood events) and predict their magnitudes more accurately, relative to 6-hourly aggregations and (unaggregated) daily simulations.

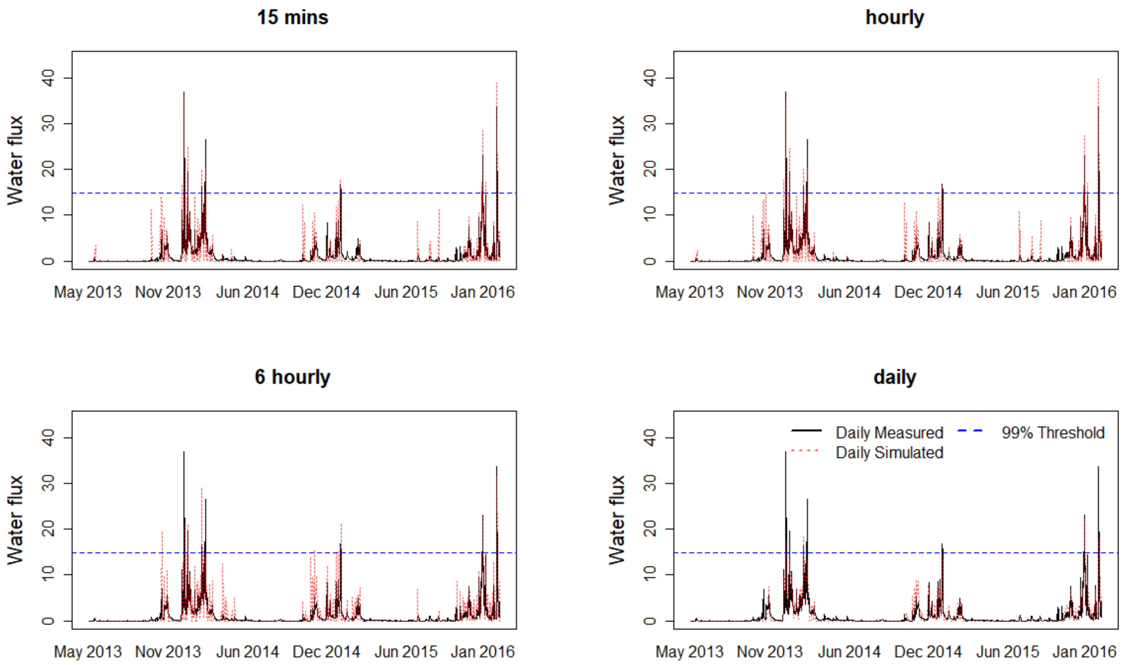

Figure 4. Time-series plots for daily measured and daily simulated water flux data (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily). All units in mm d^{-1}. All plots are shown with a threshold at the 99th percentile of measured data (14.90 mm d^{-1}).

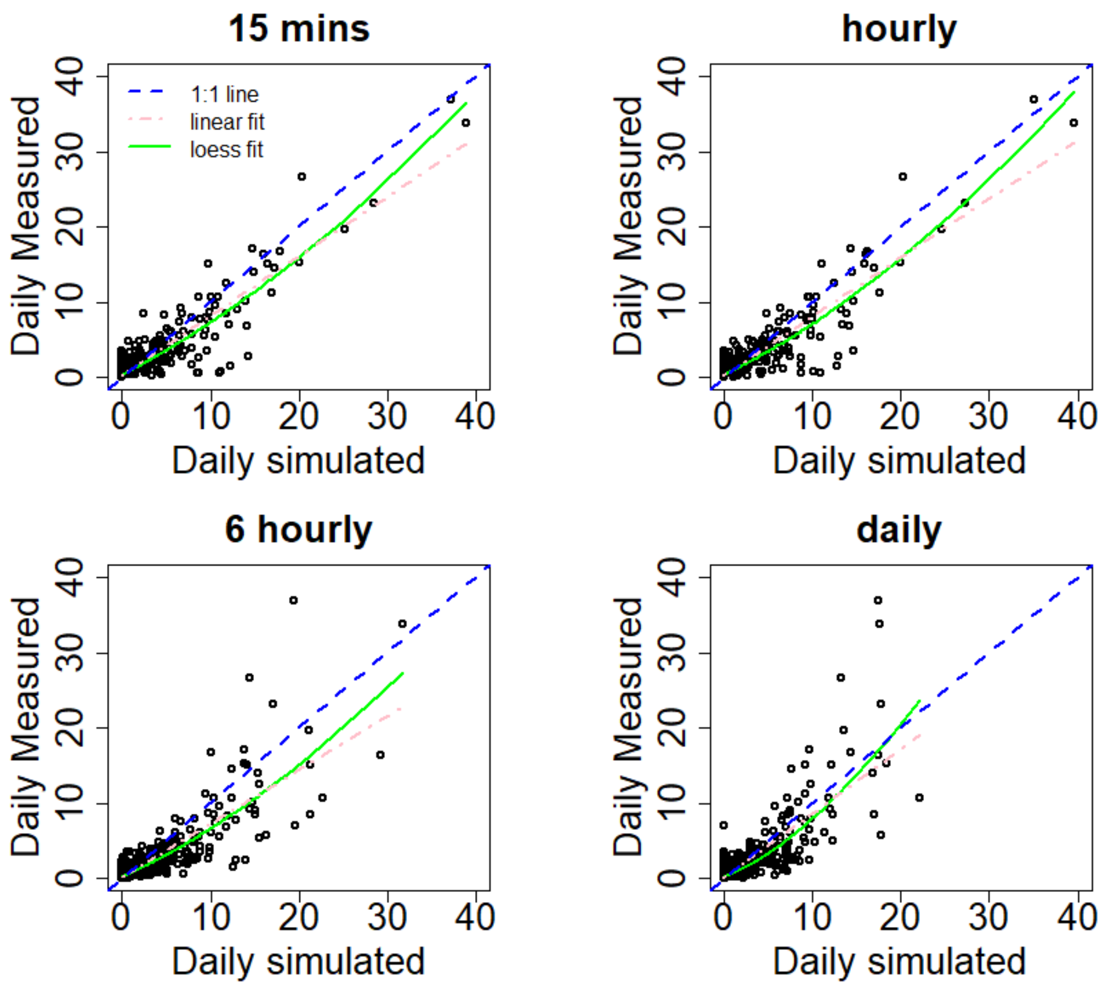

Figure 5. Scatterplots of the daily measured and daily simulated water flux (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily). Scatterplots are shown with the ideal 1:1 line, a linear regression fit and a loess smoother fit. All units in mm d^{-1}.

Further clarity on bias is provided in the density plots for the measured and simulated data (Fig. Figure 6). Here, daily simulations based on 15-minute and hourly aggregations have a lower density at small daily water fluxes than that found with the measured data, while the 6-hourly aggregations and (unaggregated) daily simulations have a higher density at small daily water fluxes. This is combined with a longer tail in the density curve for the 15-minute and hourly aggregations, as each can simulate large daily water fluxes, while the 6-hourly aggregation and (unaggregated) daily simulations do not have this property.

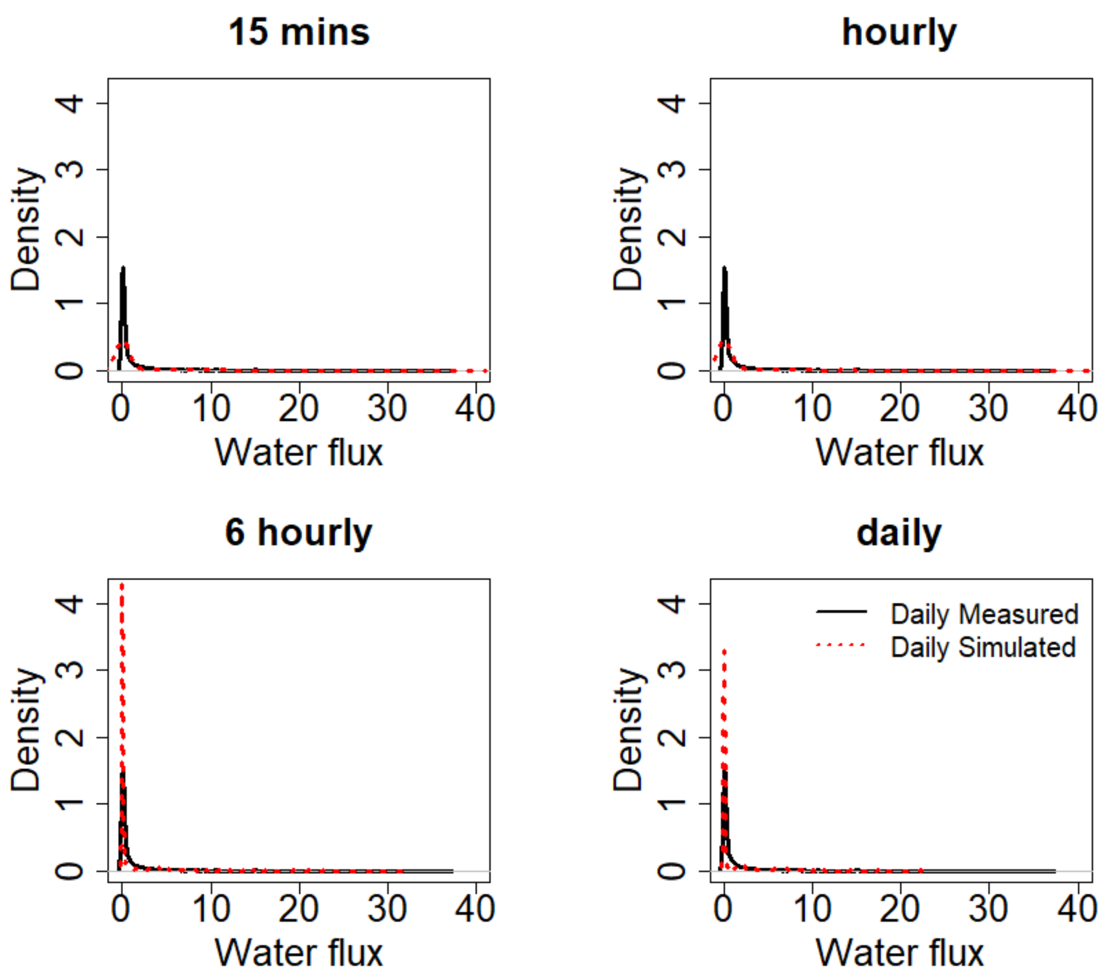

Figure 6. Density plots for daily measured and daily simulated data (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily). All units in mm d^{-1}.

The error indices (MAE, NRMSE and PBIAS) are reported for each daily aggregation in Fig. Figure 7, where the 15-minute and hourly aggregations clearly perform more accurately than the 6-hourly aggregation and (unaggregated) daily simulations. Errors (i.e. residuals) are also reported over the study time period in Fig. Figure 8, where errors tend to be larger for the daily simulations based on the 6-hourly aggregation and the (unaggregated) daily simulations. Interestingly, the 6-hourly aggregation consistently is the least accurate, including being less accurate than the (unaggregated) daily simulations.

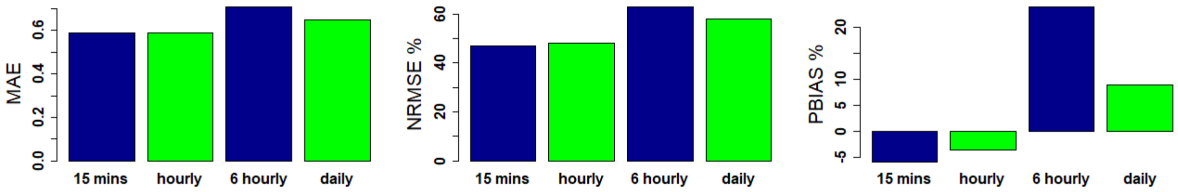

Figure 7. Error (MAE, NRMSE, PBIAS) indices with respect to daily measured and daily simulated data (with 15-minute, hourly, 6-hourly data aggregated to daily).

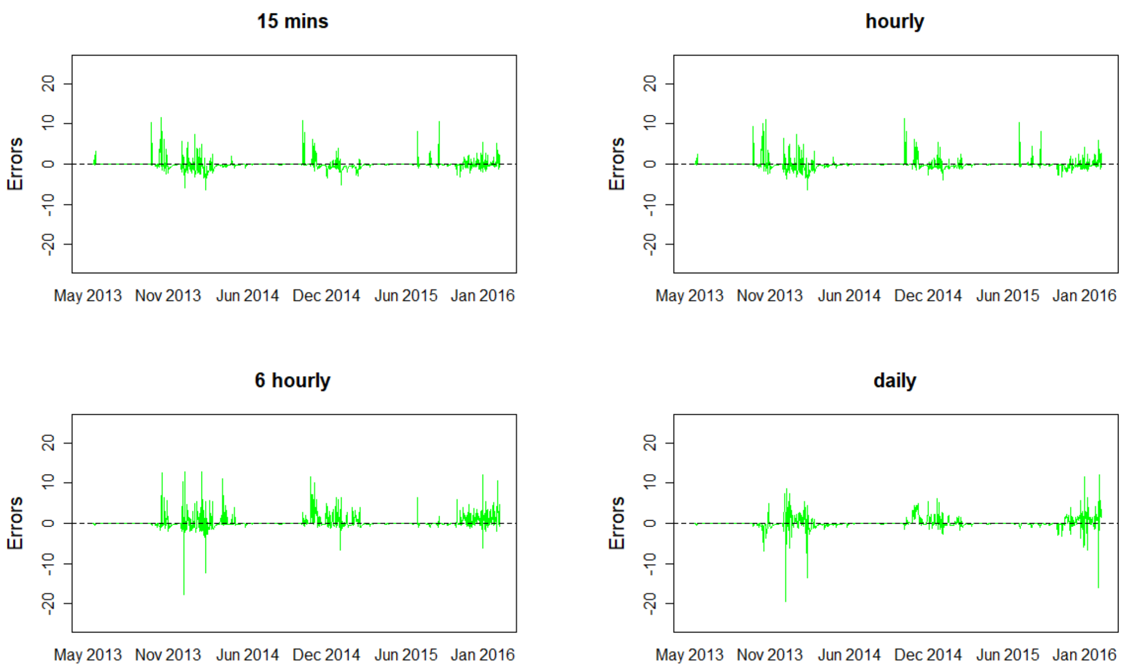

Figure 8. Time-series of errors (simulated minus measured data) aggregated to the daily temporal resolution. All units in mm d^{-1}. Positive errors represent over-prediction by the model.

Agreement indices (NSE, d and KGE) are reported for each daily aggregation in Fig. Figure 9, where again the 15-minute and hourly aggregations perform more accurately than the 6-hourly aggregation and (unaggregated) daily simulations (although daily simulations perform relatively well according to the KGE index). From the given accuracy diagnostics, it is not immediately apparent whether daily simulations based on 15-minute or hourly aggregations are the most accurate, and as such, both appear to increase the accuracy relative to SPACSYS' daily simulations. Again, the 6-hourly aggregation is the least accurate.

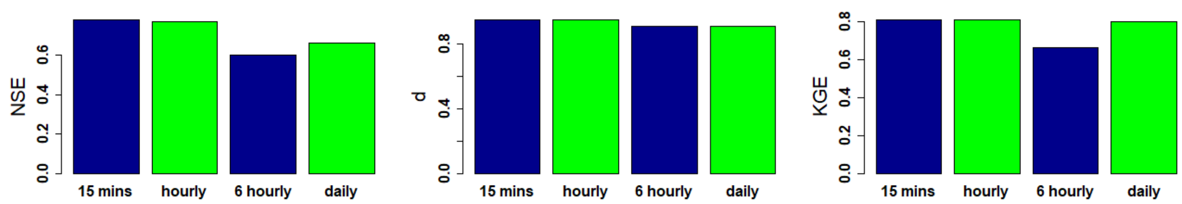

Figure 9. Agreement (NSE, *d*, KGE) indices with respect to daily measured and daily simulated data (with 15-minute, hourly, 6-hourly simulations aggregated to daily).

3.3 Simulation of measured peaks

To investigate the ability of the model to simulate and identify water flux peaks, the 99^{th} percentile of each measured water flux dataset was used as a threshold to identify peak flows, as highlighted in Figs. Figure 2 and Figure 4 (the dashed blue line). Incidences of correct peak flow simulations, false negatives, false positives and the resultant Kappa values are given in Table 2. It appears that hourly simulations provide the largest correct classification rate (Kappa = 0.553) for the unaggregated approach, but with only moderate success in identifying measured peak flows (as a promising 92% identification rate is tempered by a poor mis-identification rate). Conversely, aggregating to the daily scale using either 15-minute or hourly simulations was able to provide much greater agreement in identifying measured peak flows at the daily scale, with each identifying 9 out of 11 peak flow events correctly, coupled with only 2 false positives (Kappa = 0.816 in both cases). This level of agreement was far greater than that found through directly simulating the daily data, which provided only moderate success in identifying measured peak flows (Kappa = 0.495). Again, the 6-hourly aggregation is the least accurate with a relatively high number of false positives (simulated flow exceeds the threshold when measured flow does not).

Table 2. Accuracy at peak water fluxes according to simulation resolution. Peaks taken at 99th percentile of measured data (see the dashed blue line in Figs. Figure 2 and Figure 4).

Simulation resolution	**Sample size**	**Measured Peaks**	**Correctly Simulated**	**False Negative**	**False Positive**	**Kappa**
Unaggregated						
15-minute	97920	980	759 (77)*	221	1224	0.506
hourly	24480	245	225 (92)	20	335	0.553
6-hourly	4080	41	32 (78)	9	52	0.503
daily	1020	11	5 (45)	6	4	0.495
Aggregated to daily						
15-minute	1020	11	9 (82)	2	2	0.816

hourly	1020	11	9 (82)	2	2	0.816
6-hourly	1020	11	6 (55)	5	9	0.455
daily	1020	11	5 (45)	6	4	0.495

* Value in brackets shows a percentage of correctly simulated peaks to measured peaks.

4 Discussion

4.1 Model performance

4.1.1 Unaggregated data

The statistical analyses for model performance suggested that the SPACSYS model simulates the general trend of water fluxes at the four different temporal resolutions reasonably well (Figs. Figure 2 and Figure 3). All simulations tended to over-predict water flux, however, and only simulations at the finest resolutions maintained the variation in the measured data. The accuracy of water flux peak simulations varied among the four resolutions (Table 2). Almost 92% of the measured peaks over the simulated period were modelled correctly at an hourly resolution, the resolution with the smallest misclassification rate. However, this was tempered by a high rate of predicting peaks that did not exist. A previous statistical analysis of peak flows at different scales from a different NWFP sub-catchment (similar in size to the one used here), modelled and simulated by a Generalized Pareto distribution, also showed the greatest agreement at the hourly resolution (Curceac et al., 2020).

4.1.2 Aggregated to Daily

When simulations at a finer temporal resolution were aggregated to a daily rate, the simulation results using both the 15-minute and hourly aggregations showed the greatest accuracy broadly equally, both in the prediction of general trends (Figs. Figure 4 to Figure 9) and the identification of peak flows (Table 2). This demonstrates clearly that the daily simulation of water fluxes with the SPACSYS model, informed by finer temporal resolution data, can increase simulation accuracy. This result is an important advance relative to previous SPACSYS studies, which only used a daily time-step, and which similarly used sub-catchments of the NWFP as the study site (Liu et al., 2018).

The simulation results using both the 15-minute and hourly aggregations showed the greatest accuracy broadly equally. When the simulation time step is getting longer, average precipitation intensity might be weaker, which causes simulated water fluxes smoother and the deviations with monitored water fluxes larger. Given the complexity of the soil-water processes that operate across the field, it is not surprising to see substantial variation around the 1-1 line when predicting flow at 15-minute to 6-hour scales due to inherent variability. We would expect to see a similar phenomenon with any similar process-based model. What is important is that when these fine-scale predictions are aggregated we get substantial improvements to daily predictions.

4.2 Results in context and their generalisation

Results are consistent with other studies that similarly showed that differences in the (unaggregated or aggregated) time-step have the greatest impact on runoff simulation accuracy relative to other factors, some of which, also investigated changes in spatial resolution (i.e. aggregating over different spatial units) (Choi et al., 2018; Huang et al., 2019; Jeong et al., 2010; Kavetski et al., 2011; Merz et al., 2009). Thus, the value of using aggregated fine temporal resolution simulations to increase the accuracy of daily simulations can be said to hold generally for other process-based models provided the hydrological process is described appropriately. However, it does not follow that daily simulation accuracy will continue to increase as the temporal resolution of the aggregated data becomes finer. This study found aggregating hourly simulations to daily to be just as accurate as aggregating

15-minute simulations to daily, while aggregating 6-hourly simulations to daily performed less well than the usual daily simulations.

The temporal resolution for process-based models should be chosen carefully to balance between capturing all important processes, the study objectives and data availability. For our study, with flooding as context, the identification of water flux extremes in a grassland field (or small sub-catchment) with a heavy clayed soil, is viewed as *the* important process, more so than capturing broad trends in water flux. It is well-known that running a model at a finer resolution, then aggregate, will increase the prediction accuracy in a broad sense (see above). What has received less attention in the literature is the effects of temporal resolution on a model's ability to capture extremes (e.g. see Schaller et al., 2020, in the context of streamflow). In this respect, our study has found daily peak flows to be more accurately identified using aggregations of simulations at finer resolutions, than using coarse daily simulations directly. Of course, measurement at a finer resolution comes at a cost and this needs to be balanced with associated improvements in model accuracy. In this instance, this interplay is simple to resolve since aggregating to the daily scale using both 15-minute and hourly simulations were equally as accurate, meaning measuring at an hourly interval would be sufficient for the case study site.

The appropriate temporal resolution to simulate water fluxes using a hydrological model depends on hydro-climatological and geophysical characteristics, and the scale of the process. It has been suggested that an appropriate temporal resolution could be between 12 hours for middle-sized upstream areas and 48 hours for a complete river basin (Booij and Tran, 2005). As the size of the field for this study is < 4 ha, the indicated hourly resolution appears reasonable. Observed and projected changes in the UK's climate suggest an increase in heavy rain events and wetter winters (Committee on Climate Change, 2017), where some UK regions will be more affected than others. This will inevitably change agricultural management practice and land use across the UK. Taking as an example the grazed pasture of this study, introducing a deep-rooting grass suited to its heavy clay soils (Macleod et al., 2013) and/or the mechanical loosening of topsoil (Newell-Price et al., 2011) would reduce water runoff, whereas conversion to an arable crop (e.g. wheat) would provide its own set of water runoff influences. Such changes would alter the characteristics of the water fluxes generated, as the field's soil properties will change, meaning the determination of an appropriate resolution to simulate water fluxes may also change from the hourly resolution suggested here. This is analogous to other hydrological studies where, for example, different overflow designs in roof drainage structures have markedly variable responses to rainfall intensity increases (Verstraten et al., 2019).

4.3 Inputs that impact hydrological model performance

Key model input variables such as precipitation can determine the impacts of simulation time-steps on the performance of hydrological models; for example, the duration and temporal variability of a precipitation event in relation to the rainfall–runoff lag time (Ficchì et al., 2016). A multiple-day precipitation event is the main cause of continuous runoff events and related peaks. For example, for this study, there was almost an unbroken measured precipitation period from 14 December 2013 to 5 March 2014, which brought a total of 541 mm of precipitation, 78% of which was measured as surface runoff (i.e., measured water flux). Study simulations showed 70, 70, 81 and 85% as water fluxes over the period at the 15-minute, hourly, 6-hourly and daily resolutions, respectively. Previous studies showed that wetter soils had less capacity to store water, resulting in greater runoff volumes (Huang et al., 2017; Kibet et al., 2014; Zehe et al., 2010). Both observations and simulations in this study confirmed this finding. Conversely, for a single day event, a measured 92% of 40.2 mm daily precipitation was discharged on 23 December 2013. The simulations generated 99, 99, 90 and 44% water losses at the 15-minute, hourly, 6-hourly and daily resolutions, respectively. Thus, almost all of the precipitation contributed to the water loss on this day, where only the daily-scale simulation did not capture this. However, although heavy rainfall is necessary to generate water fluxes, it is not a sufficient condition for a higher surface runoff rate to occur (Ledingham et al., 2019). For example, there was about 25 mm precipitation on 14 May 2013 and on 13 August 2015, but both the simulations and the observations (at all four resolutions) did not show apparent water fluxes. Further, a daily precipitation of 4 mm on 27 February 2015 generated a measured 120% water loss, together with simulated values of 48, 148, 125 and 75% water loss at the 15-minute, hourly, 6-hourly and daily resolutions, respectively.

The generation of water fluxes not only depends on the intensity of precipitation, but also surface coverage, topology and soil physical properties of the field. In hydrology, lag time, defined as the time difference between the peak runoff and mass centre of rainfall excess (Hall, 1984), is usually used to determine a runoff rate. Although the SPACSYS model does not use this parameter to estimate the surface runoff rate, it uses the Richard’s equation to calculate soil water redistribution where soil hydraulic conductivity, saturated water content and plant uptake play critical roles in water infiltration and consequently surface runoff. A trial study on the spatial variation of soil hydraulic conductivity (unpublished data) in a NWFP field, nearby to the study field, highlighted clear within-field variation, partially because of compaction caused by grazed animal movement. However, the soil physical properties used in the simulations were estimated based on soil texture, and at the field-scale only. The error and uncertainty introduced by this approach are likely to be transferred to the errors in simulating infiltration and surface runoff rates. To improve model simulation accuracy, soil physical properties as core information should be provided wherever possible.

4.4 Further considerations of scale

The processes controlling water fluxes operate across a range of spatial and temporal scales, and the time-series that are recorded represent an aggregation of these effects. For example, effects of evapotranspiration will dominate at annual scales whereas more local impacts of precipitation will manifest at finer scales (Rust et al., 2014). As noted above, the characteristics of the study area, (e.g. size, soil condition and topography) will impact the dominant scales of variation and hence the frequency at which it is most appropriate to model or measure water fluxes. Rust et al. (2014) presented an analysis which aimed to determine whether the process-based model they studied captured the scale-dependent variation measured in catchment runoff. They analysed model residuals using wavelet-based signal processing methods and found that although their model captured broadly the scale-dependent variation in the data, fine scale variation was always under-predicted. Our results shed light on their observation as we confirm that if the scale of the model prediction is not sufficiently fine then model-damping will result in an under-prediction of extreme events.

5 Conclusions

For the grassland study site, the adapted process-based model (SPACSYS) could adequately simulate the trends in measured water fluxes and identify their extremes. At a daily time-step, model accuracy increased when simulations were run at finer temporal resolutions, specifically 15-minute and hourly, and then aggregated to daily (a coarse output resolution commonly used in field-scale agricultural settings). Aggregating using 6-hourly simulations was less accurate. For the study site, which constitutes a field of a grassland research farm platform (NWFP), simulation of water fluxes at an hourly resolution is likely optimal since use of the 15-minute resolution did not increase prediction accuracy or the ability to identify extremes in flow further. Therefore, for modelling purposes, monitoring frequency could be reduced safely to hourly from the current 15-minute resolution.

Results provide information not only for the NWFP experiment, but also and indirectly, the UK grassland farming regions that its outputs upscale to (Pulley and Collins, 2019). Study results are crucial in relation to meeting the increasing demand for reliable simulation-based runoff forecasts at daily and sub-daily resolutions, where accurate knowledge of peak discharge and stage are essential not only for flood protection, but also to help increase the forecast accuracy of associated emissions such as nutrient or sediment loss, that each use water flux as a component. Further research is called for in specifying the temporal resolution amongst the wide range of field-scale hydrological/agricultural models currently applied. This needs to be coupled with linked changes in climate and land use to increase model forecast accuracy and to optimise data acquisition schemes on farms generally.

6 Acknowledgements

This research was funded by the BBSRC Institute Strategic Programme grant, “Soils to Nutrition” (BBS/E/C/000I0330, BBS/E/C/000I0320), the BBSRC National Capability grant for the North Wyke Farm Platform (BBS/E/C/000J0100) and a PhD studentship funded by Rothamsted Research and Lancaster University.

References

Abalos, D., Cardenas, L.M., Wu, L., 2016. Climate change and N_2O emissions from South West England grasslands: a modelling approach. Atmos. Environ., 132: 249-257. DOI:https://doi.org/10.1016/j.atmosenv.2016.03.007

Adimassu, Z., Alemu, G., Tamene, L., 2019. Effects of tillage and crop residue management on runoff, soil loss and crop yield in the Humid Highlands of Ethiopia. Agric. Syst., 168: 11-18. DOI:https://doi.org/10.1016/j.agsy.2018.10.007

Ahuja, L.R., Ma, L., Howell, T.A., 2002. Agricultural System Models in Field Research and Technology Transfer. CRC Press, Boca Raton.

Alaoui, A., Rogger, M., Peth, S., Blöschl, G., 2018. Does soil compaction increase floods? A review. J. Hydrol., 557: 631-642. DOI:https://doi.org/10.1016/j.jhydrol.2017.12.052

Archer, D.R., Fowler, H.J., 2018. Characterising flash flood response to intense rainfall and impacts using historical information and gauged data in Britain. J. Flood Risk Manag., 11: S121-S133. DOI: https://doi.org/10.1111/jfr3.12187

Bingham, I.J., Wu, L., 2011. Simulation of wheat growth using the 3D root architecture model SPACSYS: validation and sensitivity analysis. Eur. J. Agron., 34: 181-189. DOI:https://doi.org/10.1016/j.eja.2011.01.003

Booij, M.J., Tran, H.M., 2005. Appropriate scales for hydro-climatological variables in the Red River basin. In: Wagener, T. et al. (Eds.), Proceedings of symposium S6 held during the Seventh IAHS Scientific Assembly. IAHS publication. IAHS Press, pp. 325-332.

Brown, I., Thompson, D., Bardgett, R., Berry, P., Crute, I., Morison, J., Morecroft, M., Pinnegar, J., Reeder, T., Topp, K., 2016. UK Climate Change Risk Assessment Evidence Report: Chapter 3, Natural Environment and Natural Assets. Report Prepared for the Adaptation Sub-committee of the Committee on Climate Change, London.

Carswell, A.M., Gongadze, K., Misselbrook, T.H., Wu, L., 2019. Impact of transition from permanent pasture to new swards on the nitrogen use efficiency, nitrogen and carbon budgets of beef and sheep production. Agric. Ecosyst. Environ., 283: 106572. DOI:https://doi.org/10.1016/j.agee.2019.106572

Charlton, M.B., Bailey, A., Arnell, N., 2010. Water for Agriculture – Implications for Future Policy and Practice Report for the Royal Agricultural Society of England.

Choi, Y.S., Shin, M.-J., Kim, K.T., 2018. Preliminary study of computational time steps in a physically based distributed rainfall–runoff model. Water, 10: 1269. DOI: https://doi.org/10.3390/w10091269

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc., 74: 829-836. DOI:https://10.1080/01621459.1979.10481038

Committee on Climate Change, 2017. UK Climate Change Risk Assessment 2017. Synthesis Report., London.

Curceac, S., Atkinson, P.M., Milne, A., Wu, L., Harris, P., 2020. An evaluation of automated GPD threshold selection methods for hydrological extremes across different scales. J. Hydrol., 585: 124845. DOI:https://doi.org/10.1016/j.jhydrol.2020.124845

Dadson, S.J., Hall, J.W., Murgatroyd, A., Acreman, M., Bates, P., Beven, K., Heathwaite, L., Holden, J., Holman, I.P., Lane, S.N., O'Connell, E., Penning-Rowsell, E., Reynard, N., Sear, D., Thorne, C., Wilby, R., 2017. A restatement of the natural science evidence concerning catchment-based 'natural' flood management in the UK. Proc. R. Soc. A, 473: 20160706. DOI:https://doi.org/10.1098/rspa.2016.0706

Del Grosso, S.J., Ojima, D.S., Parton, W.J., Stehfest, E., Heistemann, M., DeAngelo, B., Rose, S., 2009. Global scale DAYCENT model analysis of greenhouse gas emissions and mitigation strategies for cropped soils. Glob. Planet. Change, 67: 44-50. DOI:https://doi.org/10.1016/j.gloplacha.2008.12.006

Department for Environment, Food and Rural Affairs, 2019. Agriculture in the United Kingdom 2018, London.

Environment Agency, 2009. Flooding in England: A National Assessment of Flood Risk, Bristol, UK.

Ficchì, A., Perrin, C., Andréassian, V., 2016. Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events. J. Hydrol., 538: 454-470. DOI:https://doi.org/10.1016/j.jhydrol.2016.04.016

Fox, J., 2016. Applied Regression Analysis and Generalized Linear Models (Third Edition). SAGE Publications Inc., Califonia.

Hall, M.J., 1984. Urban Hydrology. Elsevier Applied Science Publishers, Lodon.

Harrod, T.R., Hogan, D.V., 2008. The soils of North Wyke and Rowden, North Wyke Research, North Wyke, Devon.

Huang, J., Kang, Q., Yang, J.X., Jin, P.W., 2017. Multifactor analysis and simulation of the surface runoff and soil infiltration at different slope gradients. IOP Conf. Ser. Earth Environ. Sci., 82: 012019. DOI:https://doi.org/10.1088/1755-1315/82/1/012019

Huang, Y., Bárdossy, A., Zhang, K., 2019. Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data. Hydrol. Earth Syst. Sc., 23: 2647-2663. DOI:https://doi.org/10.5194/hess-23-2647-2019

Jansson, P.-E., 1998. Simulation model for soil water and heat conditions. Description of the SOIL model. Division of Agricultural Hydrotechnics Communications 98:2, Swedish University of Agricultural Sciences, Uppsala.

Jeong, J., Kannan, N., Arnold, J., Glick, R., Gosselink, L., Srinivasan, R., 2010. Development and integration of sub-hourly rainfall–runoff modeling capability within a watershed model. Water Resour. Manag., 24: 4505-4527. DOI:https://doi.org/10.1007/s11269-010-9670-4

Kavetski, D., Fenicia, F., Clark, M.P., 2011. Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. Water Resour. Res., 47: W05501. DOI: https://doi.org/10.1029/2010WR009525

Keesstra, S., Nunes, J., Novara, A., Finger, D., Avelar, D., Kalantari, Z., Cerdà, A., 2018. The superior effect of nature based solutions in land management for enhancing ecosystem services. Sci. Total Environ., 610-611: 997-1009. DOI:https://doi.org/10.1016/j.scitotenv.2017.08.077

Kibet, L.C., Saporito, L.S., Allen, A.L., May, E.B., Kleinman, P.J.A., Hashem, F.M., Bryant, R.B., 2014. A protocol for conducting rainfall simulation to study soil runoff. J. Vis. Exp.: e51664. DOI:https://doi.org/10.3791/51664

Ledingham, J., Archer, D., Lewis, E., Fowler, H., Kilsby, C., 2019. Contrasting seasonality of storm rainfall and flood runoff in the UK and some implications for rainfall-runoff methods of flood estimation. Hydrol. Res., 50: 1309-1323. DOI:https://doi.org/10.2166/nh.2019.040

Liu, Y., Li, Y., Harris, P., Cardenas, L., Dunn, R.M., Sint, H., Murray, P., Lee, M., Wu, L., 2018. Modelling field scale spatial variation in water run-off, soil moisture, N_2O emissions and herbage biomass of a grazed pasture using the SPACSYS model. Geoderma, 315: 49-58. DOI:https://doi.org/10.1016/j.geoderma.2017.11.029

Liu, Y., Wu, L., Watson, C.A., Baddeley, J.A., Pan, X., Zhang, L., 2013. Modeling biological dinitrogen fixation of field pea with a process-based simulation model. Agron. J., 105: 670-678. DOI:https://doi.org/10.2134/agronj2012.0412

Macleod, C.J.A., Humphreys, M.W., Whalley, W.R., Turner, L., Binley, A., Watts, C.W., Skot, L., Joynes, A., Hawkins, S., King, I.P., O'Donovan, S., Haygarth, P.M., 2013. A novel grass hybrid to reduce flood generation in temperate regions. Sci. Rep., 3: 1683. DOI:https://doi.org/10.1038/srep01683

Merz, B., Plate, E.J., 1997. An analysis of the effects of spatial variability of soil and soil moisture on runoff. Water Resour. Res., 33: 2909-2922. DOI:https://doi.org/10.1029/97WR02204

Merz, R., Parajka, J., Blöschl, G., 2009. Scale effects in conceptual hydrological modeling. Water Resour. Res., 45: W09405. DOI: https://doi.org/10.1029/2009WR007872

Morison, J.I.L., Matthews, R.B., 2016. Agriculture and forestry climate change impacts summary report. DOI:https://nerc.ukri.org/research/partnerships/ride/lwec/report-cards/agriculture/

Newell-Price, P., Chambers, B., Whittingham, M., 2011. The alleviation of grassland compaction by mechanical soil loosening. Defra Project BD5001 Final Report.

Newell Price, B., Chambers, B., Whittingham, M., 2012. Characterisation of Soil Structural Degradation Under Grassland and Development of Measures to Ameliorate its Impact on Biodiversity and other Soil Functions. Defra Project BD5001 Final Report.

Orr, R., Murray, P., Eyles, C., Blackwell, M., Cardenas, L., Collins, A.L., Dungait, J., Goulding, K., Griffith, B., Gurr, S., Harris, P., Hawkins, J., Misselbrook, T., Rawlings, C., Shepherd, A., Sint, H., Takahashi, T., Tozer, K., Whitmore, A.P., Wu, L., Lee, M., 2016. The North Wyke Farm Platform: effect of temperate grassland farming systems on soil moisture contents, runoff and associated water quality dynamics. Eur. J. Soil Sci., 67: 374-385. DOI:https://doi.org/10.1111/ejss.12350

Palmer, R.C., Smith, R.P., 2013. Soil structural degradation in SW England and its impact on surface-water runoff generation. Soil Use Manag., 29: 567-575. DOI:https://doi.org/10.1111/sum.12068

Pathak, P., Sudi, R., Wani, S.P., Sahrawat, K.L., 2013. Hydrological behavior of Alfisols and Vertisols in the semi-arid zone: Implications for soil and water management. Agric. Water Manag., 118: 12-21. DOI:https://doi.org/10.1016/j.agwat.2012.11.012

Perego, A., Wu, L., Gerosa, G., Finco, A., Chiazzese, M., Amaducci, S., 2016. Field evaluation combined with modelling analysis to study fertilizer and tillage as factors affecting N_2O emissions: a case study in Po valley (Northern Italy) Agric. Ecosyst. Environ., 225: 72-85. DOI:https://doi.org/10.1016/j.agee.2016.04.003

Pulley, S., Collins, A.L., 2019. Field-based determination of controls on runoff and fine sediment generation from lowland grazing livestock fields. J. Environ. Manage., 249: 109365. DOI:https://doi.org/10.1016/j.jenvman.2019.109365

Rust, W., Corstanje, R., Holman, I.P., Milne, A.E., 2014. Detecting land use and land management influences on catchment hydrology by modelling and wavelets. J. Hydrol., 517: 378-389. DOI:https://doi.org/10.1016/j.jhydrol.2014.05.052

Schaller, N., Sillmann, J., Müller, M., Haarsma, R., Hazeleger, W., Hegdahl, T.J., Kelder, T., van den Oord, G., Weerts, A., Whan, K., 2020. The role of spatial and temporal model resolution in a flood event storyline approach in western Norway. Weather and Climate Extremes, 29: 100259. DOI:https://doi.org/10.1016/j.wace.2020.100259

Sharma, A., Wasko, C., Lettenmaier, D.P., 2018. If precipitation extremes are increasing, why aren't floods? Water Resour. Res., 54: 8545-8551.

DOI:https://doi.org/10.1029/2018wr023749

Stevens, A.J., Clarke, D., Nicholls, R.J., 2016. Trends in reported flooding in the UK: 1884–2013. Hydrolog. Sci. J., 61: 50-63. DOI:https://doi.org/10.1080/02626667.2014.950581

Verstraten, L., Wasko, C., Ashford, G., Sharma, A., 2019. Sensitivity of Australian roof drainage structures to design rainfall variability and climatic change. Build. Environ., 161: 106230. DOI:https://doi.org/10.1016/j.buildenv.2019.106230

Wasko, C., Sharma, A., Lettenmaier, D.P., 2019. Increases in temperature do not translate to increased flooding. Nat. Commun., 10: 5676. DOI:https://doi.org/10.1038/s41467-019-13612-5

Watts, G., Anderson, M., 2016. Water climate change impacts report card 2016. DOI:https://nerc.ukri.org/research/partnerships/ride/lwec/report-cards/water/

Wu, L., Bingham, I.J., Baddeley, J.A., Watson, C.A., 2009. Modeling plant nitrogen uptake using three-dimensional and one-dimensional root architecture. In: Ma, L., Ahuja, L.R., Bruulsema, T.W. (Eds.), Quantifying and understanding plant nitrogen uptake systems modeling. CRC Press, Boca Raton, FL, pp. 197-218.

Wu, L., Blackwell, M., Dunham, S., Hernández-Allica, J., McGrath, S.P., 2019. Simulation of phosphorus chemistry, uptake and utilisation by winter wheat. Plants, 8: 404. DOI:https://doi.org/10.3390/plants8100404

Wu, L., McGechan, M.B., McRoberts, N., Baddeley, J.A., Watson, C.A., 2007. SPACSYS: integration of a 3D root architecture component to carbon, nitrogen and water cycling - model description. Ecol. Model., 200: 343-359. DOI:https://doi.org/10.1016/j.ecolmodel.2006.08.010

Wu, L., Rees, R.M., Tarsitano, D., Zhang, X., Jones, S.K., Whitmore, A.P., 2015. Simulation of nitrous oxide emissions at field scale using the SPACSYS model. Sci. Total Environ., 530–531: 76-86. DOI:http://dx.doi.org/10.1016/j.scitotenv.2015.05.064

Wu, L., Zhang, X., Griffith, B.A., Misselbrook, T., 2016. Sustainable grassland systems: A modelling perspective based on the North Wyke Farm Platform. Eur. J. Soil Sci., 67: 397-408. DOI:https://doi.org/10.1111/ejss.12304

Zehe, E., Graeff, T., Morgner, M., Bauer, A., Bronstert, A., 2010. Plot and field scale soil moisture dynamics and subsurface wetness control on runoff generation in a headwater in the Ore Mountains. Hydrol. Earth Syst. Sci., 14: 873-889. DOI:https://doi.org/10.5194/hess-14-873-2010

Zhang, X., Sun, N., Wu, L., Xu, M., Bingham, I.J., Li, Z., 2016. Effects of enhancing soil organic carbon sequestration in the topsoil by fertilization on crop productivity and stability: evidence from long-term experiments with wheat-maize cropping systems in China. Sci. Total Environ., 562: 247-259. DOI:https://doi.org/10.1016/j.scitotenv.2016.03.193

Figure captions

Figure 1. Details of the NWFP catchment selected for this study (catchment number 6 of 15, consisting of a single field called Golden Rove).

Figure 2. Time-series plots for measured and simulated water flux data (not aggregated) for 15-minute, hourly, 6-hourly and daily data (in units of mm $15min^{-1}$, mm h^{-1}, mm $6h^{-1}$ and mm d^{-1}, respectively). All plots are shown with a threshold at the 99th percentile of measured data (at 0.138 mm $15min^{-1}$, 0.553 mm h^{-1}, 3.45 mm $6h^{-1}$ and 14.9 mm d^{-1}, respectively).

Figure 3. Scatterplots of the measured and simulated data (not aggregated) for 15-minute, hourly, 6-hourly and daily data. Scatterplots are shown with the 1:1 line, a linear regression fit and a loess smoother fit. Units are in mm $15min^{-1}$, mm h^{-1}, mm $6h^{-1}$ and mm d^{-1}, respectively.

Figure 4. Time-series plots for daily measured and daily simulated water flux data (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily). All units in mm d^{-1}. All plots are shown with a threshold at the 99th percentile of measured data (14.90 mm d^{-1}).

Figure 5. Scatterplots of the daily measured and daily simulated data (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily).

Scatterplots are shown with the ideal 1:1 line, a linear regression fit and a loess smoother fit. All units in mm d^{-1}.

Figure 6. Density plots for daily measured and daily simulated data (with the first three plots having data aggregated from: 15 minutes to daily; hourly to daily; 6 hourly to daily). All units in mm d^{-1}.

Figure 7. Error (MAE, NRMSE, PBIAS) indices with respect to daily measured and daily simulated data (with 15-minute, hourly, 6-hourly data aggregated to daily).

Figure 8. Time-series of errors (simulated minus measured data) aggregated to the daily temporal resolution. All units in mm d^{-1}. Positive errors represent over-prediction by the model.

Figure 9. Agreement (NSE, *d*, KGE) indices with respect to daily measured and daily simulated data (with 15-minute, hourly, 6-hourly simulations aggregated to daily).