Natural Environment Valuation Online Tool

Technical Documentation

Version 1.0, June 2019

Chapter 6b: Water Quantity & Quality Emulation

Land, Environment, Economics and Policy (LEEP) Institute

University of Exeter

Corresponding authors

Nathan Owen n.e.owen@exeter.ac.uk

Brett Day brett.day@exeter.ac.uk

Amy Binner a.r.binner@exeter.ac.uk

Authors

Ian Bateman

Greg Smith

Patrick Collings

Louis Haddrell

Lorena Luizzo

Carlo Fezzi

Collaborators

Forestry Commission / Forest Research

UCL

JNCC

University of Aberdeen

Introduction

The report on water quantity and quality modelling in NEVO outlined how the Soil & Water Assessment Tool (SWAT) (Neitsch et al., 2011) was calibrated for 253 river basins across England, Scotland and Wales. This report describes why it is necessary in NEVO to have a fast-running version of the SWAT model, known as a statistical emulator, and outlines the methodology for building such an emulator.

In the NEVO tool, we wish to use the calibrated SWAT models to investigate the impact of land cover change on water quantity and quality. If a user makes a change to the land cover in a catchment, we need to re-run the SWAT model for this configuration behind the scenes, before returning the output to the user on the NEVO tool. However, re-running the SWAT models in this way poses a number of issues. Firstly, the SWAT models are slow to run and in some cases the user may have to wait some time before getting results returned to them. NEVO has been designed to be a fast-running decision support tool so this is not ideal. For example, running the entire River Thames catchment under a land cover change for the time period 2020-2059 takes approximately 10 minutes. Secondly, a more technical issue is that the SWAT model runs via Windows executable files which are not suited to the Linux server architecture of NEVO. These executable files also depend on a large number of text files which would take up space on the server. Therefore, for the purposes of the NEVO tool we would benefit from replacing the SWAT model with a fast-running alternative, known as a statistical emulator. This report outlines our methodology for building a statistical emulator of a calibrated SWAT model.

The key concept of emulation is that a slow-running model is replaced with an alternative model which is orders of magnitude faster, for a reduction in the prediction accuracy. However, the loss in prediction accuracy can be controlled with enough information about the model itself, and the associated uncertainty can be quantified (Smith, 2013). There are many types of emulator, ranging from simple linear regression models to more complex formulations such as Gaussian process emulators (Rasmussen and Williams, 2006; O'Hagan, 2006) or polynomial chaos expansions (Ghanem and Spanos, 1990; Xiu and Karniadakis, 2002). In this work, we will keep it simple and use various types of linear regression models. However, the complexity of the SWAT model means that we will combine the emulator with other techniques such as principal component analysis (for dimension reduction) and mixture experiments (for land cover variables).

SWAT Model: An Overview

In this section we will give a short overview of the SWAT model, focussing on aspects of the model which are relevant for building a statistical emulator. More detailed information about SWAT can found in the report on water quantity and quality modelling in NEVO.

Catchments and subcatchments

Figure 1 shows an example SWAT catchment, in this case the Mires Beck in East Riding of Yorkshire. The catchment is split into 25 subcatchments, with the river flowing downstream until it exits the catchment at the outlet point of subcatchment 25. For the purposes of this report we distinguish between two types of subcatchment: *first order* subcatchments, which have no subcatchments upstream; and *second order* subcatchments, which take input from upstream subcatchments (which could be first or second order themselves). First and second order subcatchments are shown in yellow and pink respectively in Figure 1.

In the NEVO tool, 253 catchments have been calibrated and split into subcatchments, resulting in a total of 6444 subcatchments.

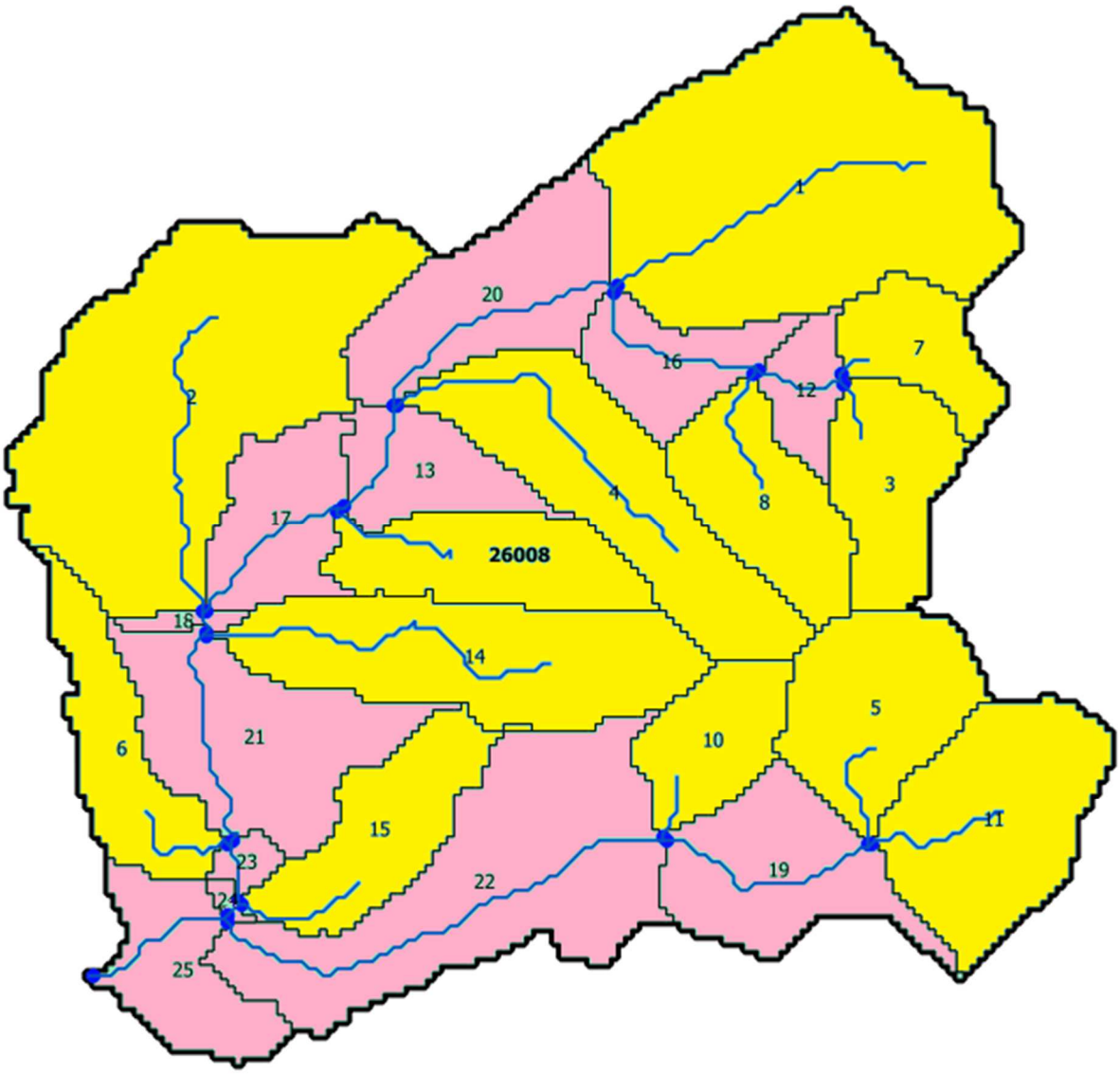


Figure 1: An example of a SWAT catchment (number 26008, Mires Beck in East Riding of Yorkshire), divided into 25 subcatchments. First and second order subcatchments are shown in yellow and pink respectively. The river network is shown as a blue line, with outlet points of each subcatchment shown as blue circles.

Hydrological outcomes

In a SWAT simulation of a catchment, a wide range of hydrological outcomes are recorded for each subcatchment. For more information, see the SWAT documentation (Neitsch et al., 2011). In this work, we have focussed on the following outcomes:

- Average daily streamflow (m^3/s).
- Organic nitrogen (kg N).
- Organic phosphorus (kg P).
- Nitrate (kg N).
- Ammonium (kg N).
- Nitrite (kg N).
- Mineral phosphorus (kg P).
- Dissolved oxygen (kg O_2).

Each of these outcomes is recorded in a daily time series of length T. Throughout this report, this time series will be denoted using the vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)$ regardless of hydrological outcome. For simplicity, we will consider $\boldsymbol{y}$ as average daily streamflow. The procedure for the other outcomes is similar apart from a few small details, and these will be discussed where necessary.

For most of the above outcomes, there are actually two measures: those transported into the subcatchment, and those transported out of the subcatchment. We will refer to these as "in" and "out" variables and denote the time series as $\boldsymbol{y}_I$ and $\boldsymbol{y}_O$ respectively. In terms of Figure 1, we think of the "out" variables being measured at the outlet points of each subcatchments (shown as blue circles). We think of the "in" variables as being measured at the point at which the river enters the subcatchment, or if it is a first order subcatchment, the source of the river.

The "in" component does not exist for ammonium and nitrite, and the methodology is adapted accordingly to account for this.

Land use component

As the SWAT simulation runs from the top of the catchment to the bottom, it is not the case that the "out" variables from upstream subcatchments combine together to form the "in" variables in downstream subcatchments. Denote the "in" and "out" variables for subcatchment j as $\boldsymbol{y}_I^{(j)}$ and $\boldsymbol{y}_O^{(j)}$ respectively. Consider subcatchments 5, 11 and 19 in Figure 1. Subcatchments 5 and 11 are first order and combine together to form the second order subcatchment 19. It is not the case that $\boldsymbol{y}_I^{(19)} = \boldsymbol{y}_O^{(5)} + \boldsymbol{y}_O^{(11)}$. In fact, we have identified an additional component, $\boldsymbol{y}_L^{(19)}$, defined as:

$$\boldsymbol{y}_L^{(19)} = \boldsymbol{y}_I^{(19)} - \boldsymbol{y}_O^{(5)} - \boldsymbol{y}_O^{(11)}$$

We will call this the "land use" variable and denote it $\boldsymbol{y}_L$. It is the contribution to the outcome from the land in a particular subcatchment, excluding the input from upstream. For flow, you can imagine that the "land use" component is simply the flow entering the stream via precipitation in the subcatchment flowing over the land. Rearranging this equation, we can see that the "in" variable is simply the "land use" variable plus the "out" variables from the upstream subcatchments:

$$\boldsymbol{y}_I^{(19)} = \boldsymbol{y}_L^{(19)} + \boldsymbol{y}_O^{(5)} + \boldsymbol{y}_O^{(11)}$$

More generally, for a downstream (second order) subcatchment and two upstream (first or second order) subcatchments this can be written as:

$$\boldsymbol{y}_I^{(downstream)} = \boldsymbol{y}_L^{(downstream)} + \boldsymbol{y}_O^{(upstream\ 1)} + \boldsymbol{y}_O^{(upstream\ 2)} \quad (1)$$

The majority of second order subcatchments in the 253 NEVO basins have two upstream subcatchments, but there are rare examples with three. In this case, we simply add the "out" variable from the third upstream subcatchment to Equation (1).

Note that for first order subcatchments, $\boldsymbol{y}_I = \boldsymbol{y}_L$ since there are no upstream subcatchments.

Finally, the "land use" component is not relevant for ammonium and nitrite.

Land covers

As documented in the report on water quantity and quality modelling in NEVO, each SWAT subcatchment has been divided into the proportions of 13 land covers: water, urban, rangeland, forest, pasture, generic agricultural, winter wheat, barley, oilseed rape, potatoes, sugar beet, corn and oats. In terms of NEVO functionality, we do not allow the water or urban land covers to be

changed. Therefore, in this work we build an emulator to represent the hydrological response to the remaining 11 land covers.

Design of Experiments

Two emulators per subcatchment

We propose to build an emulator for each subcatchment, rather than the catchment as a whole. As we will demonstrate, in this way we are able to accurately mimic the behaviour of SWAT downstream through a catchment as a result of land cover change. For each subcatchment, the emulator must tell us the response of the SWAT model hydrological outcomes under a land cover change, and how these change as we move downstream within the subcatchment. The output from the subcatchment emulator must be able to act as inputs for downstream subcatchments, imitating the way a river flows downstream.

The existence of the "land use" component means that we actually require two emulators for each subcatchment:

1. *Land use emulator*. For each outcome we model the relationship between the "land use" variable and the land uses in the subcatchment (expressed as proportions).
2. *Instream emulator*. For each outcome we model the relationship between the "out" variable and the "in" variable, equivalent to the instream processes in the SWAT model.

Using these two emulators we can mimic the full behaviour of SWAT from the top of the catchment to the bottom as follows. We start with all first order subcatchments. Given their land uses we can predict the "land use" variables using the land use emulator. Since these are equal to the "in" variables for first order subcatchments, we can then use the instream emulator to predict the "out" variables. We now move on to second order subcatchments. Again, given the land uses we can predict the "land use" variables using the land use emulator. For second order subcatchments we must use Equation (1) to obtain the "in" variables, combining the "land use" variables and the "out" variables from upstream. Using the "in" variables we can then predict the "out" variables using the instream emulator. We continue in this fashion to the bottom of the catchment.

Experimental data

To build each emulator we need to obtain data from the model by running at different settings. The emulator then interpolates the given data and allows us to make a prediction of what the model would have given at another setting. In terms of the SWAT model, we must run every subcatchment for a range of configurations of the land uses, recording the responses of the hydrological output for each.

Recall that we wish to model the hydrological response as a function of 11 land covers. Denote the land covers as: x_1 (rangeland), x_2 (forest), x_3 (pasture), x_4 (agriculture), x_5 (wheat), x_6 (barley), x_7 (oilseed rape), x_8 (potatoes), x_9 (sugar beet), x_{10} (corn) and x_{11} (oats), or the vector $\boldsymbol{x} = (x_1, x_2, \dots, x_{11})$. Denote the extra two (fixed) land covers as: x_{12} (water) and x_{13} (urban). An experimental design is a set of n configurations of the land uses, denoted as an $n \times d$ matrix $\boldsymbol{X} = (\boldsymbol{x}^1, \boldsymbol{x}^2, \dots, \boldsymbol{x}^n)'$.

The land covers x_j, $j = 1, \dots, 13$, are defined as proportions of the total subcatchment area. For simplicity, assume for now that $x_{12} = x_{13} = 0$. Then, it must be the case that in each subcatchment the remaining 11 land covers sum to 1, or:

$$\sum_{j=1}^{11} x_j = 1$$

This is called a mixture experiment (Scheffé, 1958; Cornell, 2011) and there are a range of suitable experimental designs and associated models. In this work we use a simple experimental design called a simplex lattice design (Scheffé, 1958). A simplex lattice design of level l in d variables consists of all the possible combinations of the points $x_j = 0, \frac{1}{m}, \frac{2}{m}, \ldots, 1$ for $j = 1, 2, \ldots, d$. A simplex lattice design of level l in d variables contains $n = (l + d - 1)!/(l!\,(d - 1)!)$ points. For example, a simplex lattice design of level $l = 2$ in $d = 3$ variables consists of the 6 points:

$$\boldsymbol{X} = ((1,0,0), (0,1,0), (0,0,1), (0.5, 0.5, 0), (0.5, 0, 0.5), (0, 0.5, 0.5))'$$

In this work, we use a simplex lattice design of level $m = l$ in $d = 11$ variables (land covers). This design consists of 66 points and is shown in Figure 2 (black circles).

In reality, the proportion of the water (x_{12}) and urban (x_{13}) land covers does not have to be zero. In this case, we simply scale the experimental design $\boldsymbol{X}$ by a factor of $1 - (x_{12} + x_{13})$.

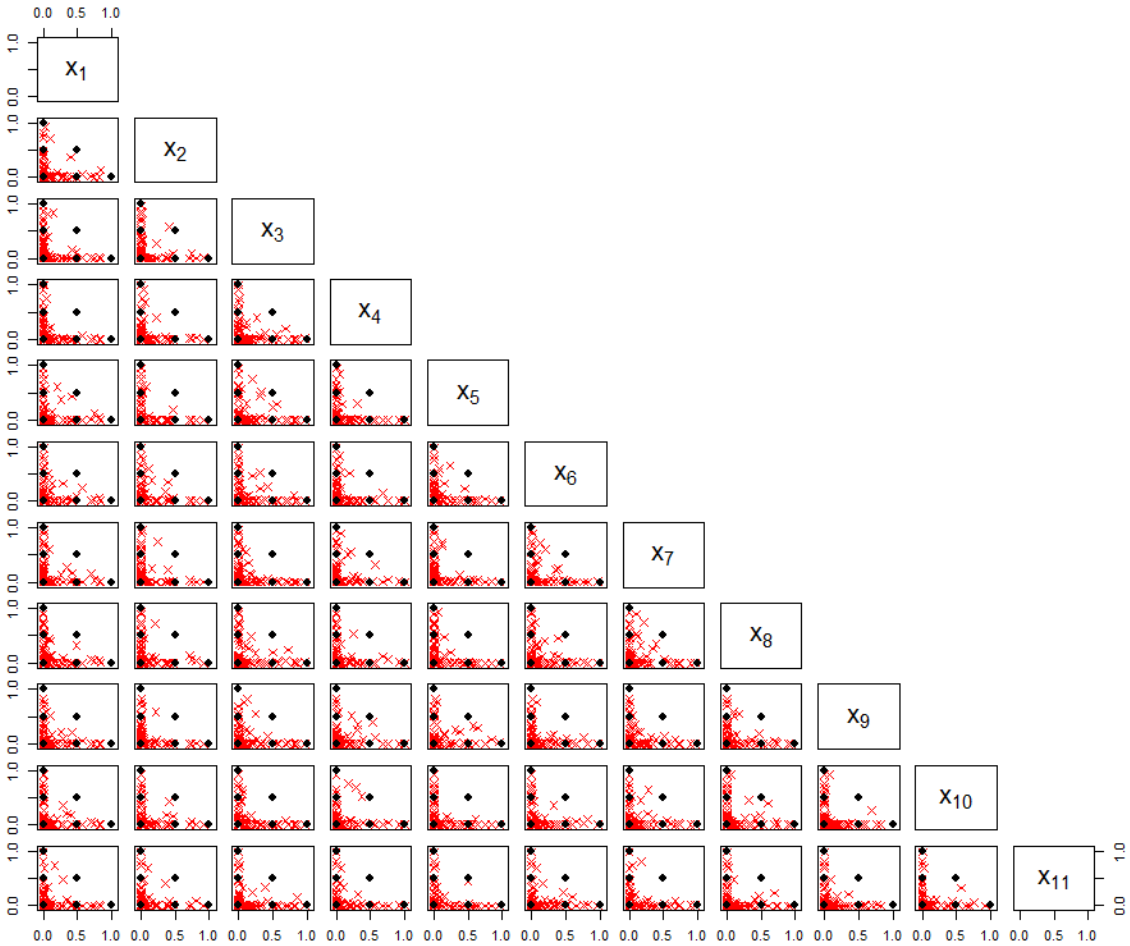


Figure 2: Experimental design (black circles) and validation design (red crosses) for building and validating the subcatchment emulators respectively. Design points are proportions of the 11 land covers of the total subcatchment area.

Validation data

Once an emulator has been built, it is important to validate it to test whether it is doing a good job of replicating the model. One way of doing this is to collect an additional set of model runs, known as a validation design. The emulator can be used to predict the model output at the validation design points, which can be compared to the corresponding model 'truth' in a number of ways to assess emulator accuracy. It is important for the validation design points to be distinct from the experimental design points to fully test the emulator.

In this work, we randomly generate a set of $m = 100$ mixture points – i.e., proportions which sum to one – in $d = 11$ variables. The validation design is denoted as an $m \times d$ matrix $\widetilde{\boldsymbol{X}} = (\widetilde{\boldsymbol{x}}^1, \widetilde{\boldsymbol{x}}^2, \dots, \widetilde{\boldsymbol{x}}^m)'$ and is shown in Figure 2 (red crosses).

Running a SWAT experiment

When running a SWAT subcatchment at a land cover configuration $\boldsymbol{x}^i = \left(x_1^i, x_2^i, \dots, x_{11}^i\right)$, each of the above set of hydrological outcomes return "in" and "out" components as daily time series of length T. The "in" and "out" components for a land configuration i will be denoted $\boldsymbol{y}_I^i = \left(y_{I,1}^i, y_{I,2}^i, \dots, y_{I,T}^i\right)$ and $\boldsymbol{y}_O^i = \left(y_{O,1}^i, y_{O,2}^i, \dots, y_{O,T}^i\right)$ respectively. From the "in" and "out" time series we can derive the "land use" component $\boldsymbol{y}_L^i = \left(y_{L,1}^i, y_{L,2}^i, \dots, y_{L,T}^i\right)$ using Equation (1).

The size of T depends on how long the SWAT simulation is for. For the purposes of the NEVO tool we focus on the four decades: 2020-2029, 2030-2039, 2040-2049, and 2050-2059. Due to leap years, the number of days $T = 3653$ in decades 2020-2029 and 2040-2049 and $T = 3652$ in decades 2030-2039 and 2050-2059.

For a single decade, a SWAT subcatchment is run for the $n = 66$ settings of the land uses in the experimental design $\boldsymbol{X} = (\boldsymbol{x}^1, \boldsymbol{x}^2, \dots, \boldsymbol{x}^{66})'$, and the $m = 100$ settings in the validation design $\widetilde{\boldsymbol{X}} = (\widetilde{\boldsymbol{x}}^1, \widetilde{\boldsymbol{x}}^2, \dots, \widetilde{\boldsymbol{x}}^{100})'$. For each single hydrological outcome, matrices of the "in", "out" and "land use" components are returned. For the experimental design these will be denoted as the $66 \times T$ matrices $\boldsymbol{Y}_I = (\boldsymbol{y}_I^1, \boldsymbol{y}_I^2, \dots, \boldsymbol{y}_I^{66})'$, $\boldsymbol{Y}_O = \left(\boldsymbol{y}_O^1, \boldsymbol{y}_O^2, \dots, \boldsymbol{y}_O^{66}\right)'$ and $\boldsymbol{Y}_L = (\boldsymbol{y}_L^1, \boldsymbol{y}_L^2, \dots, \boldsymbol{y}_L^{66})'$ respectively. For the validation design, these are similarly denoted as the $100 \times T$ matrices $\widetilde{\boldsymbol{Y}}_I, \widetilde{\boldsymbol{Y}}_O$ and $\widetilde{\boldsymbol{Y}}_L$.

Emulator Methodology

Land use emulator

We have identified that SWAT has a "land use" component which is directly driven by the land uses in the subcatchment. In other words, for each subcatchment there exists a relationship $\boldsymbol{y}_L = f(\boldsymbol{x})$ (excluding ammonium and nitrite which do not have a "land use" component). The land use emulator is an approximation to this relationship: $\boldsymbol{y}_L \approx \hat{f}(\boldsymbol{x})$. To construct this approximation we will use the data contained in the SWAT experimental run for a subcatchment, namely the experimental design, $\boldsymbol{X}$, and the "land use" component of a hydrological outcome, $\boldsymbol{Y}_L$.

Recall that $\boldsymbol{Y}_L$ is a matrix of size $100 \times T$, where $T = 3652$ or $T = 3653$ depending on the decade. To make the modelling exercise easier we will reduce the dimension of $\boldsymbol{Y}_L$. In this work, we use a well-known dimension reduction method known as principal component analysis. For ease of presentation in this section we drop the L subscript of $\boldsymbol{Y}_L$ so that the "land use" component is denoted $\boldsymbol{Y}$. We wish to find a lower dimensional representation of $\boldsymbol{Y}$, denoted $\widehat{\boldsymbol{Y}}_K$, which is of size $100 \times K$ where $K \ll T$. A procedure for doing so is as follows (Joliffe, 2011; Wilkinson, 2011):

1. Subtract the column means $\boldsymbol{m}$ from $\boldsymbol{Y}$ to obtain $\widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{m}$.

2. Using a single value decomposition or eigendecomposition of $\widehat{\boldsymbol{Y}}$, obtain a column matrix of eigenvectors, $\boldsymbol{V}$, and a diagonal matrix of eigenvalues, $\boldsymbol{\Lambda}$. The eigenvalues satisfy $\lambda_1 \geq \cdots \geq \lambda_T$.
3. Select the first K eigenvectors and eigenvalues to explain a proportion P of variability in the data, by solving for K: $\sum_{k=1}^{K} \lambda_k / \sum_{k=1}^{T} \lambda_k \geq P$. In this work we set $P = 0.99$. Denote the first K eigenvectors using the matrix $\boldsymbol{V}_K$.
4. Project the data by computing $\widehat{\boldsymbol{Y}}_K = \widehat{\boldsymbol{Y}}\boldsymbol{V}_K$. The columns of $\widehat{\boldsymbol{Y}}_K$ are the principal components, denoted $\hat{y}_1, \ldots \hat{y}_K$.
5. Transform back to the original dimension of the data by computing $\boldsymbol{Y}_K = \widehat{\boldsymbol{Y}}_K \boldsymbol{V'}_K + \boldsymbol{m}$.

Using this approach for all the subcatchments, we found that we could greatly reduce the dimension of T. Across all subcatchments, values of K ranged between 1 and 53, with a median value of 7.

By using principal component analysis, we have effectively reduced the dimension of $\boldsymbol{y}_L$ to a set of principal components $\hat{y}_k, k = 1, \ldots, K$. We are now interested in the relationships between the principal components and the land uses in the subcatchment, or in other words, the relationships $\hat{y}_k = f_k(\boldsymbol{x}), k = 1, \ldots, K$. The land use emulators now take the form $\hat{y}_k \approx \hat{f}_k(\boldsymbol{x}), k = 1, \ldots, K$.

It remains to select the parametric form of the emulator $\hat{f}_k(\boldsymbol{x})$. Recall that the vector $\boldsymbol{x}$ is a set of land cover proportions constrained to sum to 1, and that we used a simplex lattice design of level $l = 2$ from the mixture experiments literature. This design accommodates the following quadratic regression model (Scheffé, 1958; Cornell, 2011):

$$\hat{f}_k(\boldsymbol{x}) = \sum_{i=1}^{11} \alpha_{i,k} x_i + \sum_{i<j} \sum_{j}^{11} \alpha_{ij,k} x_i x_j \quad (2)$$

This polynomial regression model has the same number of coefficients as the number of points in the simplex lattice design. As such, there are no degrees of freedom and the resulting polynomial surface interpolates the points in the experimental design. The coefficients are estimated from the data using ordinary least squares with the fitlm function in MATLAB (MATLAB 2017b).

Given a new land cover configuration $\boldsymbol{x}^* = (x_1^*, x_2^*, \ldots, x_{11}^*)$ we can predict the value of the principal components as $\hat{y}_k^* \approx \hat{f}_k(\boldsymbol{x}^*), k = 1, \ldots, K$. Arranging the predicted principal components in a vector $\widehat{\boldsymbol{y}}_K^* = (\hat{y}_1^*, \ldots, \hat{y}_K^*)$, we transform back to the original dimension of the data by computing $\boldsymbol{y}_K^* = \widehat{\boldsymbol{y}}_K^* \boldsymbol{V'}_K + \boldsymbol{m}$.

Instream emulator

With the emulator for the "land use" component complete, we move on to studying the relationship between the "in" and "out" components. This relationship can be thought of as the instream processes of the SWAT model, or the function $\boldsymbol{y}_O = g(\boldsymbol{y}_I)$. As before we seek to approximate this relationship with an emulator, $\boldsymbol{y}_O \approx \hat{g}(\boldsymbol{y}_I)$.

Recall that both $\boldsymbol{y}_O$ and $\boldsymbol{y}_I$ are time series of length T, and denote their value at some time $t = 1, \ldots, T$ as $y_{O,t}$ and $y_{I,t}$ respectively. We simply use a lagged regression model of the form:

$$y_{O,t} \approx \beta_0 + \beta_1 y_{I,t} + \beta_2 y_{I,t-1} \quad (3)$$

That is, the "out" component at time t is a linear function of the "in" component at times t and $t-1$. The coefficients are estimated using all the available data in the matrices $\boldsymbol{Y}_I$ and $\boldsymbol{Y}_O$, again using the fitlm function in MATLAB.

We have found that this simple relationship gives accurate results for average daily stream flow, organic nitrogen, organic phosphorus, and nitrate. For mineral phosphorus and dissolved oxygen,

performance was improved by adding “in” components (plus one lag term) for all the other hydrological outcomes to Equation (3). Finally, given that the “in” components do not exist for ammonium and nitrite, their “out” components are regressed against the “in” components (plus one lag term) for all the other hydrological outcomes.

Results

In this section we demonstrate the above methodology applied to the average daily streamflow of subcatchment 1 of the River Taw in the decade 2030-2039. This particular subcatchment is chosen as an example and results are similar for all other subcatchments.

Land use emulator

In Figure 3 we have plotted the “land use” component of average daily streamflow in the decade 2030-2039 under two extreme settings of the land cover in the subcatchment: 100% pasture and 100% forest. We can clearly see that the flow differs under these different land covers – under the 100% forest scenario the peaks of flow are much smaller and the response in the days following the peaks is smoother. In Figure 4, average daily streamflow is also plotted for this subcatchment but we have zoomed in on the first 500 days of the decade to more clearly see the difference between the two land cover scenarios.

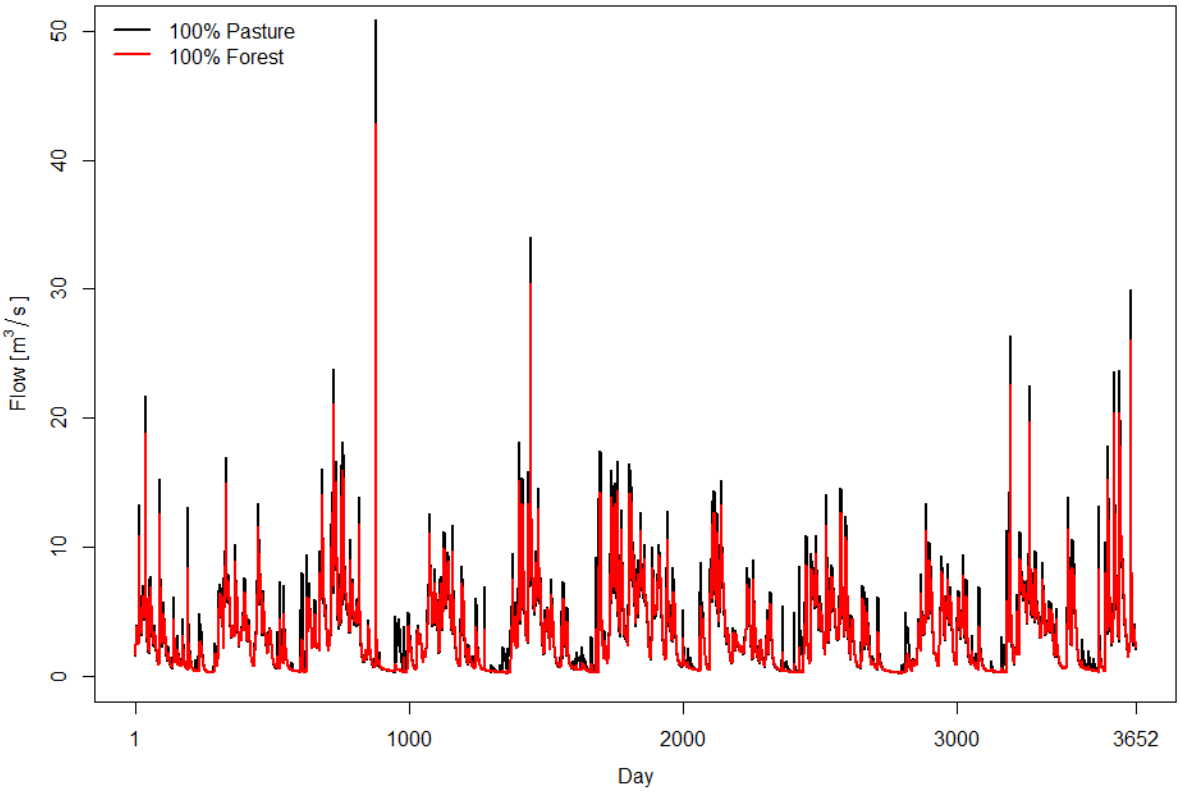


Figure 3: Average daily streamflow for subcatchment 1 of the River Taw in the decade 2030-2039 under two settings of land cover in the subcatchment: 100% pasture (black) and 100% forest (red).

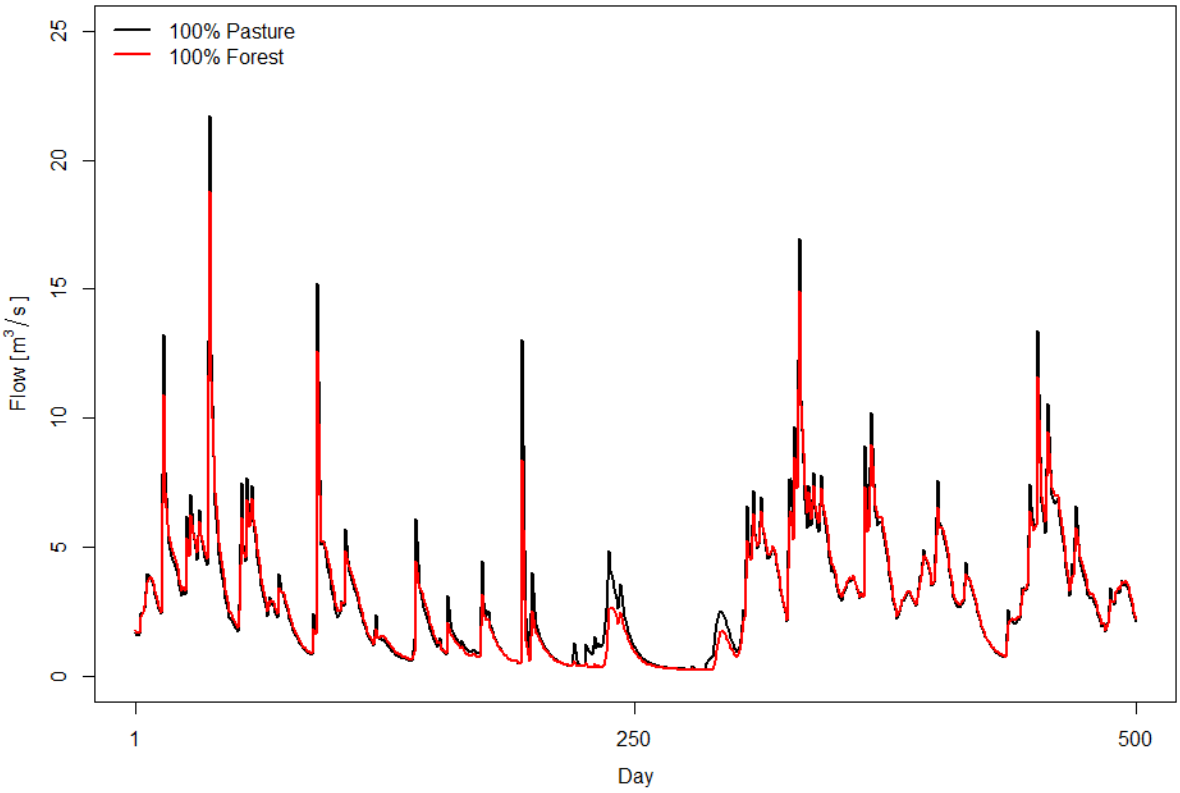


Figure 4: Average daily streamflow for subcatchment 1 of the River Taw for 500 days in the decade 2030-2039 under two settings of land cover in the subcatchment: 100% pasture (black) and 100% forest (red).

Recall that the dimension of the "land use" component data for the 66 land cover experiments is reduced using principal component analysis. For the case of average daily stream flow for subcatchment 1 of the River Taw, it was found that $K = 7$ principal components could explain over 99% of the original data. The 7 columns of $\boldsymbol{V}_K$ are used to project the original data onto a reduced dimensional space known as the principal components. The columns of $\boldsymbol{V}_K$ are also of the dimension of the original data (a 3653 day time series) and hence can be plotted to spot trends in the principal components. They are shown in Figure 5. The mean value for each day across the 66 experiments has already been subtracted, so the columns of $\boldsymbol{V}_K$ show the biggest additional sources of variability in the time series, in order from largest to smallest. We can see that in general, the magnitude of the columns of $\boldsymbol{V}_K$ decrease to reflect this fact. Each time series has its own particular form, picking up various peaks and troughs of the original average daily streamflow data.

After projecting the data along the columns of $\boldsymbol{V}_K$, we are in position to fit the above described regression model to the principal components $\hat{y}_1, \ldots \hat{y}_K$. Once the coefficients of the regression have been estimated, we can test the fit of the emulator in the "principal component space" using the validation data. A plot of the emulator predictions against the principal components for the validation data is shown in Figure 6. For all principal components, the majority of the points lie close to the $y = x$ line indicating a very good emulator fit.

We can also test the fit of the emulator in the original space of the data (i.e. a time series of average daily streamflow) after reconstruction. Such a test will contain errors not only from the fit of the emulator, but also from the reconstruction using a reduced set of principal component. Since it is hard to visualise the goodness of fit of the reconstructed emulator time series to the original time series, we choose to analyse some summary statistics of the average daily streamflow. Specifically, we study the mean and median flow over the decade, as well as the 5th and 95th percentiles of flow.

The emulator predictions of these summary statistics are shown in Figure 7. It is clear that the emulator and the principal component reconstruction provide an accurate prediction of the mean and median flow, but the tails of flow distribution as shown but the 5th and 95th percentiles.

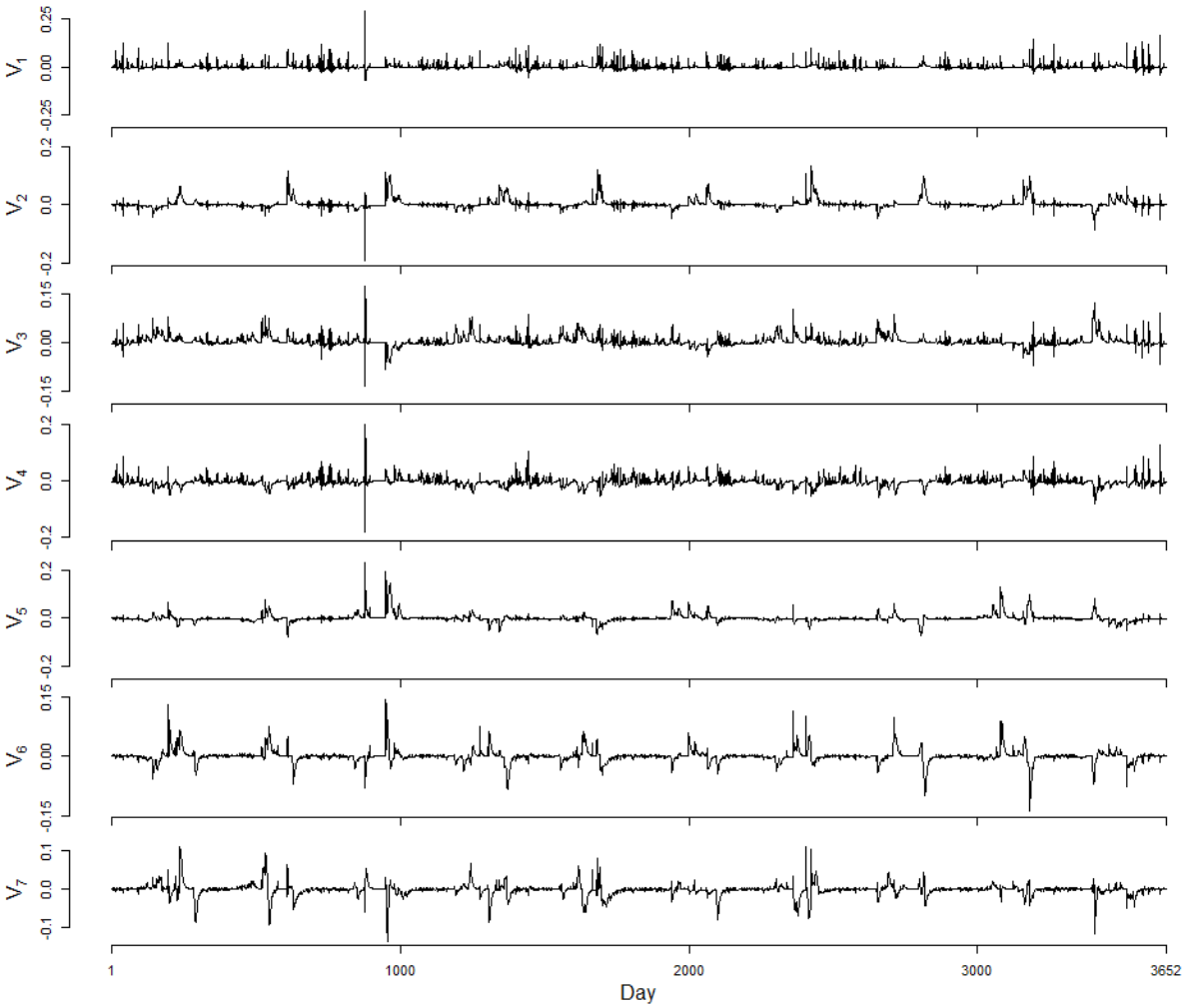


Figure 5: The column vectors of the V_K matrix for subcatchment 1 of the River Taw, along which the data is projected to form the principal components.

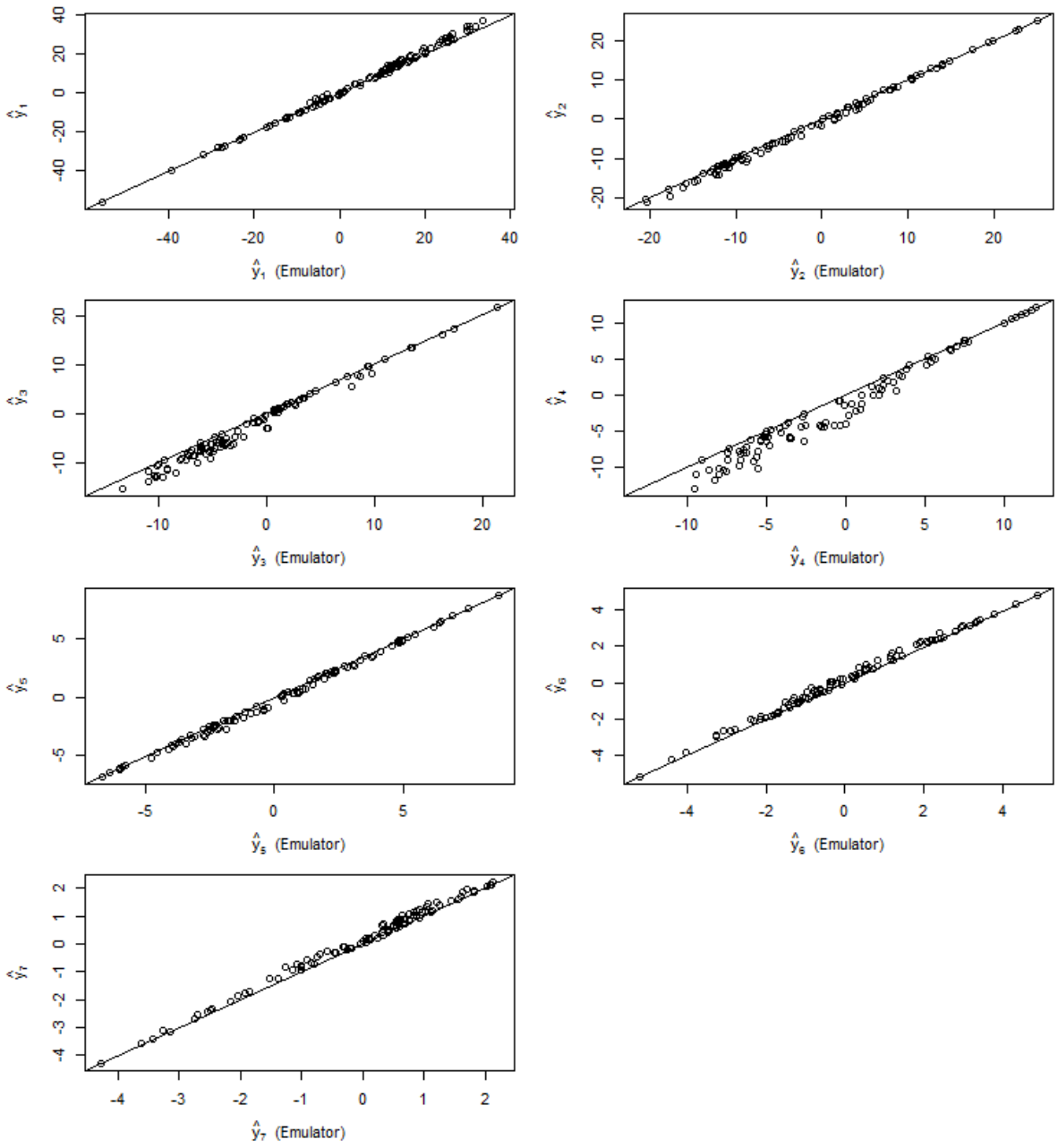


Figure 6: A plot of prinicipal components $\hat{y}_1, \ldots, \hat{y}_K$ against the corresponding emulator predictions at the validation data, for subcatchment 1 of the River Taw.

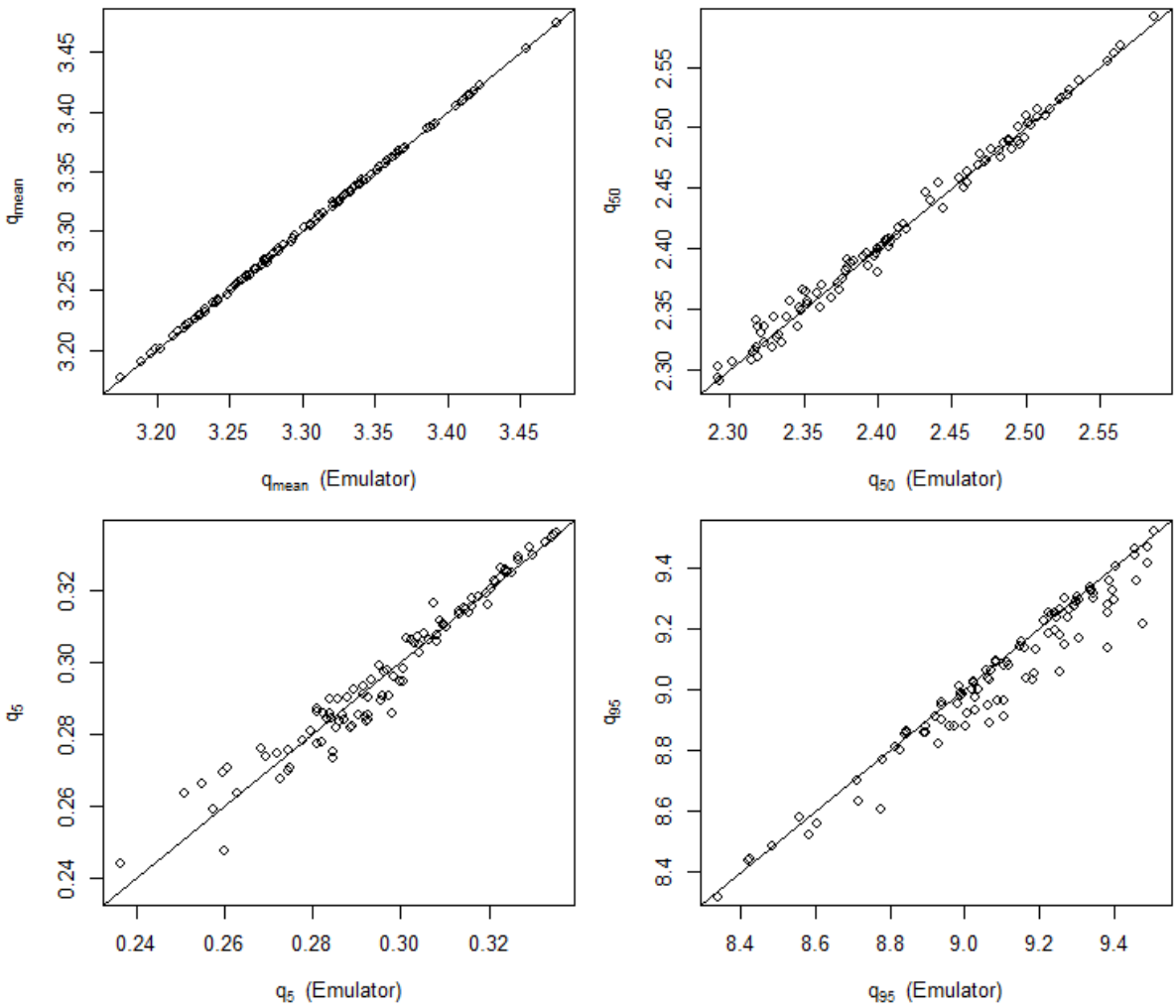


Figure 7: A plot of the mean, median, 5th and 95th percentiles of average daily streamflow against the corresponding land use emulator predictions at the validation data, for subcatchment 1 of the River Taw.

Instream emulator

Recall that the instream emulator is a relationship between the "in" and "out" components of a hydrological outcome in a subcatchment. The proposed methodology to capture this relationship is a simple lagged regression model with the "out" component as the dependent variable, and the "in" component (both on the same day and at a one-day lag) as independent variables.

To fit this model we need to prepare an appropriate dataset. For the land use emulator, each subcatchment was run at 66 different settings for the land cover, with the "land use" component recorded for each. Since these 66 settings for the land cover are likely to give the full range of hydrological outcomes, we also record the "in" and "out" components for fitting the instream emulator. Using all available data, for each hydrological outcome we prepare a dataset with the "out" component and "in" components (same day and with one-day lag) on a given day and for a given land use setting. We estimate the coefficients of the lagged regression model using all available data.

To assess the goodness of fit of the instream emulator we again use the validation data. Under a land cover setting from the validation data, we take the "in" component and the lagged "in" component and use the instream emulator to predict the corresponding "out" component. We can then compare the prediction to the observed "out" component from the SWAT model. A plot of this

comparison for average daily streamflow for subcatchment 1 of the River Taw is shown in Figure 8. It is clear that the simple lagged regression model provides an accurate relationship between the “in” and “out” component in this case. Similar results were obtained for all other subcatchments and other hydrological outcomes.

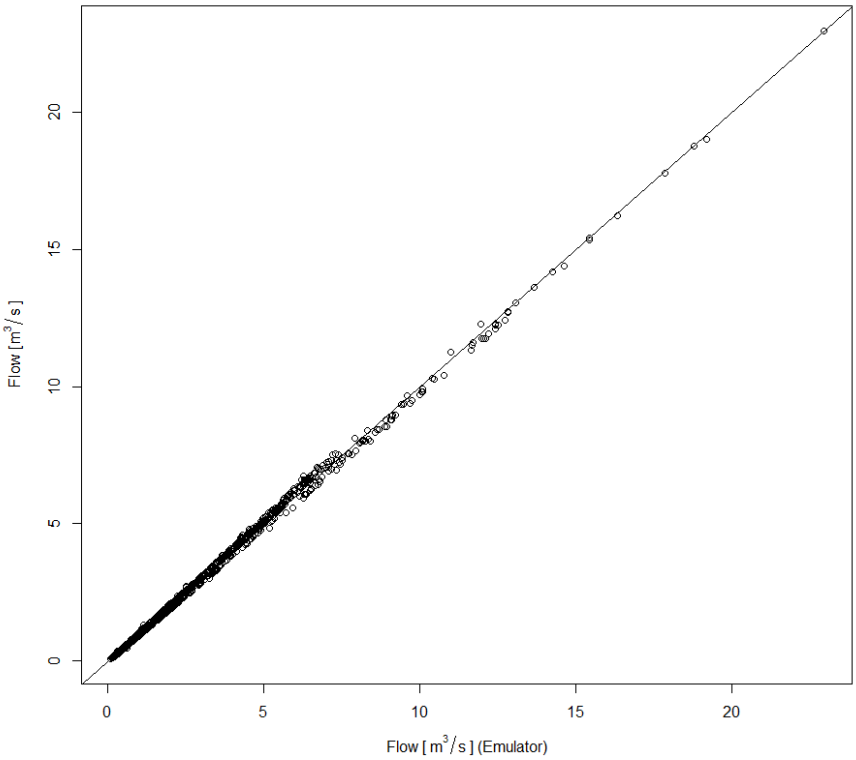


Figure 8: A plot of the "out" component of average daily streamflow from the SWAT model and as predicted by the instream emulator, at the validaiton data for subcatchment 1 of the River Taw.

Summary

In this report, we have shown how the SWAT model can be accurately approximated with a set of fast running emulators. The emulators were built for each subcatchment of a calibrated SWAT catchment. In terms of the NEVO tool, this has resulted in 6444 subcatchment emulators across 253 calibrated catchments. The set of hydrological outcomes predicted by the emulators includes: average daily streamflow, organic nitrogen, organic phosphorus, nitrate, ammonium, nitrite, mineral phosphorus and dissolved oxygen.

To accurately represent a river from the top of a catchment to the bottom, we identified a set of components for each hydrological outcome and developed different emulator methodology for each. The SWAT model records an “in” and “out” component for a hydrological outcome. Simply put, the “in” component is that measured as the river enters the subcatchment and the “out” component is that measured as the river exits the subcatchment. However, as the river flows downstream starting with first order subcatchments and passing through second order subcatchments, we identified an additional “land use” component which directly responds to the land cover in the subcatchment. To capture this complexity we proposed two kinds of emulator for each subcatchment: a land use emulator to model the relationship between the land covers and the “land use” component; and an instream emulator to model the relationship between the “in” and “out” components within the river channel. Using these two emulators, and adding together the

appropriate components as we move through subcatchments, we were able to accurately mimic the SWAT model.

The land use emulator comprised two components. Firstly, the dimension of the "land use" component itself was reduced using principal component analysis. The result of this analysis, the principal components, were then directly linked in the second stage to the land cover settings in the subcatchment using a quadratic regression model from the field of mixture experiments. Given a new land cover settings, the regression model predicted the principal components, which could then be reconstructed into a full time series for a hydrological outcome.

The instream emulator was a more simple relationship between the "in" and "out" components of a hydrological outcome. We used a lagged regression model to relate the "out" component to the "in" component on the same day and at a one-day lag.

In terms of the speed up the emulator provides over the SWAT model we revisit our example of the River Thames given in the Introduction. Using the emulator, running the entire River Thames catchment under a land cover change for the time period 2020-2059 takes approximate 1 minute – around 10 times faster than the SWAT model.

References

Cornell, J.A., 2011. *Experiments with mixtures: designs, models, and the analysis of mixture data* (Vol. 403). John Wiley & Sons.

Ghanem, R. and Spanos, P.D., 1990. Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics*, *57*(1), pp.197-202.

Jolliffe, I., 2011. *Principal component analysis* (pp. 1094-1096). Springer Berlin Heidelberg.

MATLAB and Statistics Toolbox Release 2017b, The MathWorks, Inc., Natick, Massachusetts, United States.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R. and Williams, J.R., 2011. *Soil and water assessment tool theoretical documentation version 2009*. Texas Water Resources Institute.

O'Hagan, A., 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, *91*(10-11), pp.1290-1300.

Rasmussen, C.E. and Williams, C.K., 2006. *Gaussian processes for machine learning* (Vol. 2, No. 3, p. 4). Cambridge, MA: MIT Press.

Scheffé, H., 1958. Experiments with mixtures. *Journal of the Royal Statistical Society: Series B (Methodological)*, *20*(2), pp.344-360.

Smith, R.C., 2013. *Uncertainty quantification: theory, implementation, and applications* (Vol. 12). Siam.

Wilkinson, R. D., 2011. Bayesian calibration of expensive multivariate computer experiments. In: *Large-scale Inverse Problems and Quantification of Uncertainty*, Chapter 10 (L. Beigler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, L. Tenorio, B. van Bloemen Waanders & K. Willcox, eds). John Wiley & Sons, New Jersey, USA, pp. 195–215.

Xiu, D. and Karniadakis, G.E., 2002. The Wiener--Askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, *24*(2), pp.619-644.

Appendix - Calibrating the Water Quality Model

Motivation

The water quality outputs presented in NEVO are derived from models developed using the Soil and Water Assessment Tool. These models have been calibrated using data on flow but have not been calibrated with regards to their estimates of nutrient concentrations, which we will link through to non-use valuations of changes in water quality.

Our initial exploration into the predicted nutrient concentrations produced by the NEVO water quality model highlighted a potential need for further calibration. The phosphorus levels being predicted by our SWAT models seem to be relatively higher than expected, whilst our predicted nitrogen levels were relatively low. Consequently, changes in land use where predicted to lead to negligible changes in overall WFD classification as the high phosphorus levels prohibited changes in WFD classification.

This calibration exercise aims to explore whether the SWAT models need to be calibrated to align with observed nutrient concentrations.

Data

To conduct the calibration we collected data on observed pollutant levels from the Environment Agency (EA) monitoring data (2017), river network information from OS Rivers and Water Framework Directive catchment shapefiles.

Pollutant	Units	Available points	Points in river	WFD catchments without interior point
Organic N	mg/l	114	114	8369
Nitrate (NO3)	mg/l	6313	6266	5120
Nitrite (NO2)	mg/l	6873	6810	5103
Phosphorus	mg/l	653	484	8056
Dissolved Oxygen	mg/l	6686	6244	5127

Methodology

To calibrate our models we needed to develop a method for linking EA monitoring data to our NEVO sub-basins (derived using SWAT). We begin by calibrating our 2020 baseline NEVO sub-basin predictions using the 2017 EA monitoring data.

Step 1

The first step is to link the EA monitoring points to EA defined WFD catchments. Where a WFD catchment contained one or more monitoring point it was assigned a pollutant value equal to the average of the observed pollutant levels within the catchment.

Step 2

For those WFD catchments which did not contain any monitoring points we adopted a network approach. Using the OS Rivers dataset we constructed a routing network, this identifies source points (the beginning of rivers), river nodes (where sections of rivers join), and river stretches.

EA monitoring points were added to the network as Points of Interest, and the nearest node on the network to each monitoring point was identified. If a monitoring point was further than 5km from the river network it was dropped from the set of points used to calibrate the model. These points are likely to be located in lakes, ponds and small tributaries and as such are not likely to correlate well with water quality levels in the NEVO river network.

A network routing approach was used to identify, for each WFD catchment not containing an EA monitoring point, the closest node on the river network with associated EA monitoring point data. Where this node was within a 10km along the river network the WFD catchment was assigned a pollutant value equal to the observed pollutant levels related to the node.

Step 3

As the SWAT defined sub-basins used in NEVO do not perfectly overlap with the WFD catchments used by the environment agency we calculate the proportion of the NEVO sub-basin that falls into each of the WFD catchments. We then calculate the weighted sum of the observed pollutant levels using these proportions as the weightings. Where a WFD catchment has not been assigned any EA monitoring point data the weightings are adjusted proportionally so that they sum to 1.

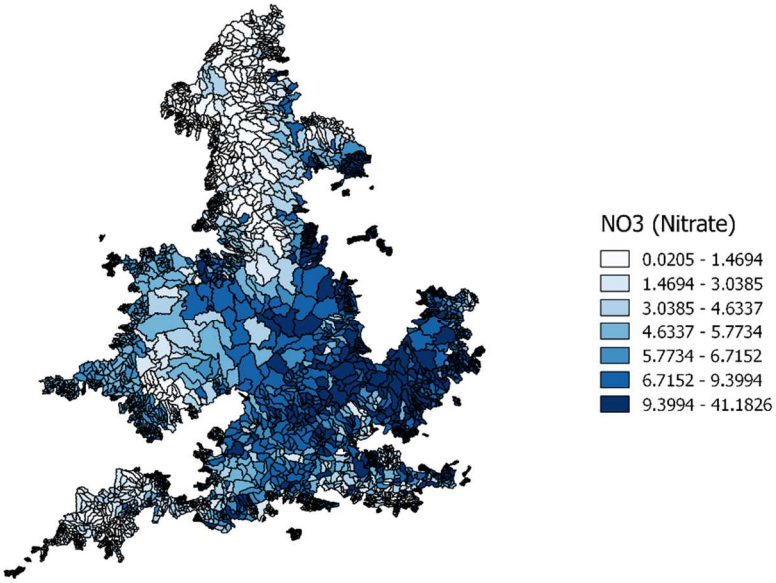
Step 4

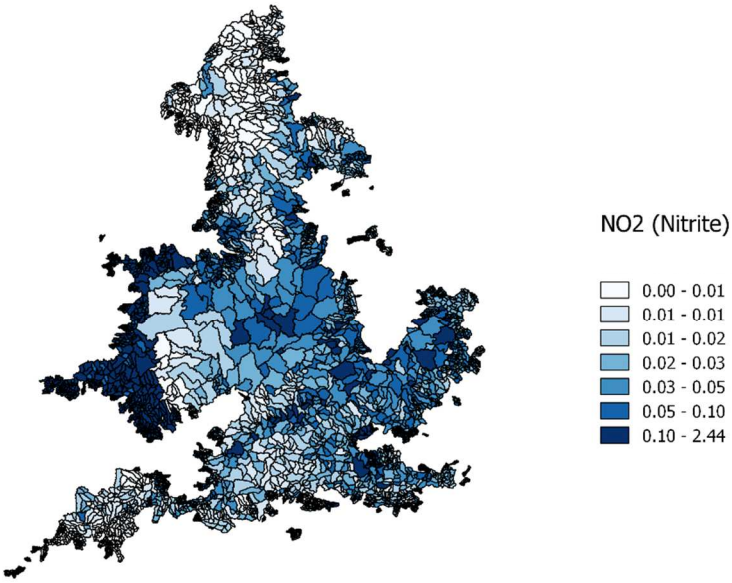
At the end of Step 3 we are able to review which NEVO catchments have been assigned weighted observed pollutant data and which have not. Where a NEVO catchment has not been assigned data this means that it does not overlap with any WFD catchment that either contains EA monitoring data or is within 10km along the river network of a node with EA monitoring data. In these cases we adopted a distributional approach to calibrate the catchment. First we calculate where this catchment falls in terms of percentiles of the predicted nutrient concentrations from the NEVO model. Secondly, for each catchment we calculate the equivalent percentiles nutrient concentration amongst the NEVO catchments that have been assigned EA monitoring data.

Step 5

The final step in the calibration process is to evaluate the adjustment required to translate the NEVO model outputs. We do this be calculating a sub-basin specific intercept transformation for the 2020 NEVO output data and applying this to model predictions in each decade. We then check that the model outputs under the most extreme land use changes conform to logical rules, for example nutrient concentrations cannot fall below zero.

Mapping the outputs





Organic N

- 0 - 0.065
- 0.065 - 0.3558
- 0.3558 - 0.7052
- 0.71 - 1.09
- 1.09 - 3.40

Phosphorus

- 0.0027 - 0.0037
- 0.0037 - 0.0266
- 0.0266 - 0.0643
- 0.0643 - 0.1233
- 0.1233 - 0.2369
- 0.2369 - 0.4011
- 0.4011 - 4.0900