Investigating the impact of data uncertainty on the estimation of catchment nutrient fluxes.

Charlotte E.M. Lloyd[1], Jim E. Freer[1], Adrian L. Collins[2], Penny J. Johnes[1], Gemma Coxon[1] and Hampshire Avon DTC team

[1] University of Bristol, Bristol, UK, [2]Rothamsted Research - North Wyke, Devon, UK. Charlotte.Lloyd@bristol.ac.uk

Introduction

The introduction of in-situ quasi-continuous monitoring of water quality provides the means to improve the characterisation of pollutant behaviour and gain new understanding of hydrological and biogeochemical processes occurring within catchments. However these data are not without uncertainties. To date studies have focused on examining uncertainties in nutrient data and how they impact on routinely used metrics such as nutrient load estimation, uncertainties in flow data are largely ignored. In addition to this, having high temporal resolution data is often considered the 'truth' and used as a benchmark from which to assess other lower resolution data sets.

Aim - To analyse a suite of high temporal resolution data sets generated from in-situ sensor networks within an uncertainty framework, including discharge and nutrient uncertainty, to provide robust estimates of nutrient fluxes from catchments impacted by intensive agricultural production practices.

Data and methodology

Data:

- Stage height and velocity were monitored at 15 min resolution at all sites and discharge calculated.
- Nitrate-N and TP were measured at 30 min (high) resolution using a field sensor (Nitratax and Phosphax) at Brixton Deverill
- Nitrate-N and TP at daily resolution using ISCO autosamplers followed by lab analysis at all sites.

Methodology:

1. Uncertainty in flow was calculated using the relationship between stage and discharge.
   - measurement uncertainty was determined using a stable period, usually winter/spring (figures 2 and 3).
2. Uncertainty in the field sensor data was determined by comparing with paired daily lab data.
3. Uncertainty in the lab data was calculated using repeated analysis of a range of standard solutions.
4. Statistics of all the errors were used in a 1st-order autoregressive model to generatel 100 iterations of the data sets including errors.
5. The replicate data sets were used to calculate nutrient loads (examples shown for two field sites).

Field sites

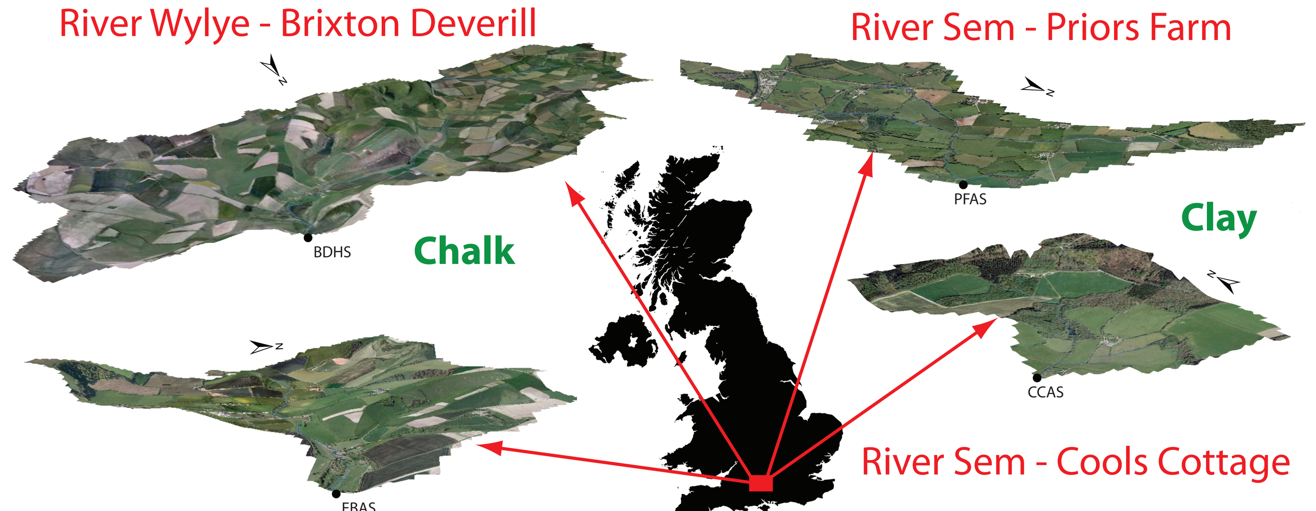


Figure 1: Schematics of field sites on the River Wylye at Brixton Deverill (BDHS), River Ebble at Ebbesbourne Wake (EBAS) and River Sem at Priors Farm (PFAS) and Cools Cottage (CCAS).

Stage-discharge relationships

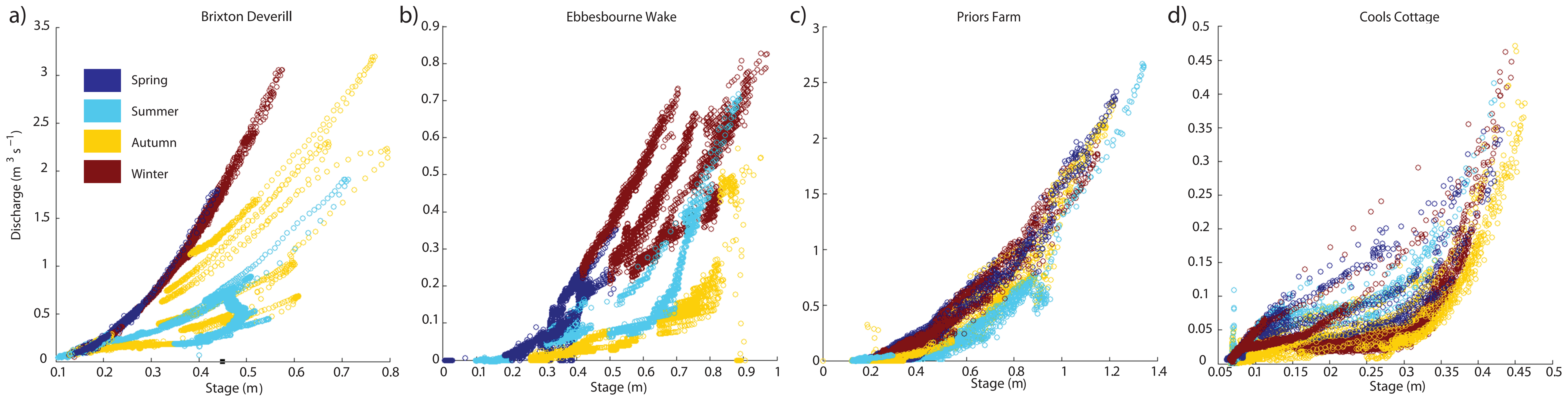
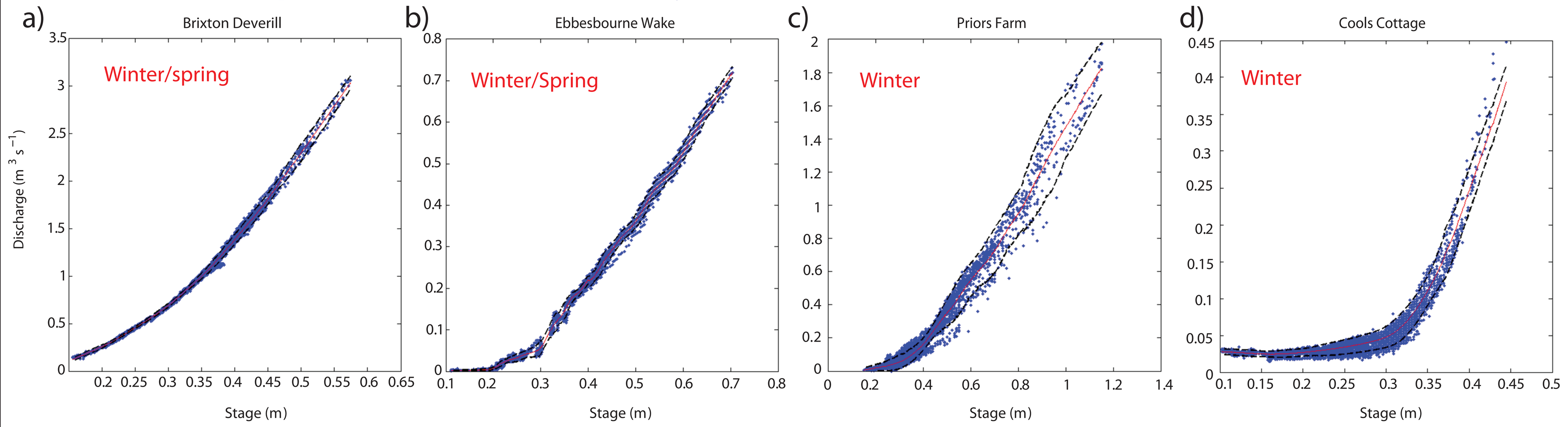


Figure 2: Stage-discharge curves split by season for a) Brixton Deverill, b) Ebbesbourne Wake, c) Priors Farm and d) Cools Cottage field sites.

- All the field sites showed strong seasonal variation in the relationship between stage and discharge.
- Brixton Deverill shows a stable relationship through winter and spring months, but is very dynamic during summer and autumn months when weed growth is rapid, causing backwatering effects.
- Ebbesbourne Wake shows shifts relating to changes in channel characteristics, through the winter each big storm event causes a shift in the relationship.
- Priors Farm and Cools Cottage data show that the relationship is dynamic with season in the clay catchments also - possibly linked with vegetation growth and changes to the bank shape.

Discharge uncertainty



Subsets of data were chosen where the relationship was stable.

A continuous sequence of data was selected.

Non-parametric loess regression technique was used to assess the relationship (Figure 3).

Figure 3: Stage-discharge uncertainty at each of the four field sites, where the blue dots are the observations, the red line the result of the Loess regression fit and the black dashed lines represent 2 s.d. away from this fit.

Residuals were examined to determine their s.d. and the autocorrelation in the errors. A 1st-order autoregressive model was used to generate multiple set of errors accounting for the autocorrelation and heteroscadasticity of discharge uncertainty.

1st order autoregressive model

$$q_k = \alpha q_{k-1} + \sqrt{1-\alpha^2}W_k$$

where q_k is the error at time k, is α temporal autocorrelation and W_k is random white noise at time k.

Nutrient uncertainty

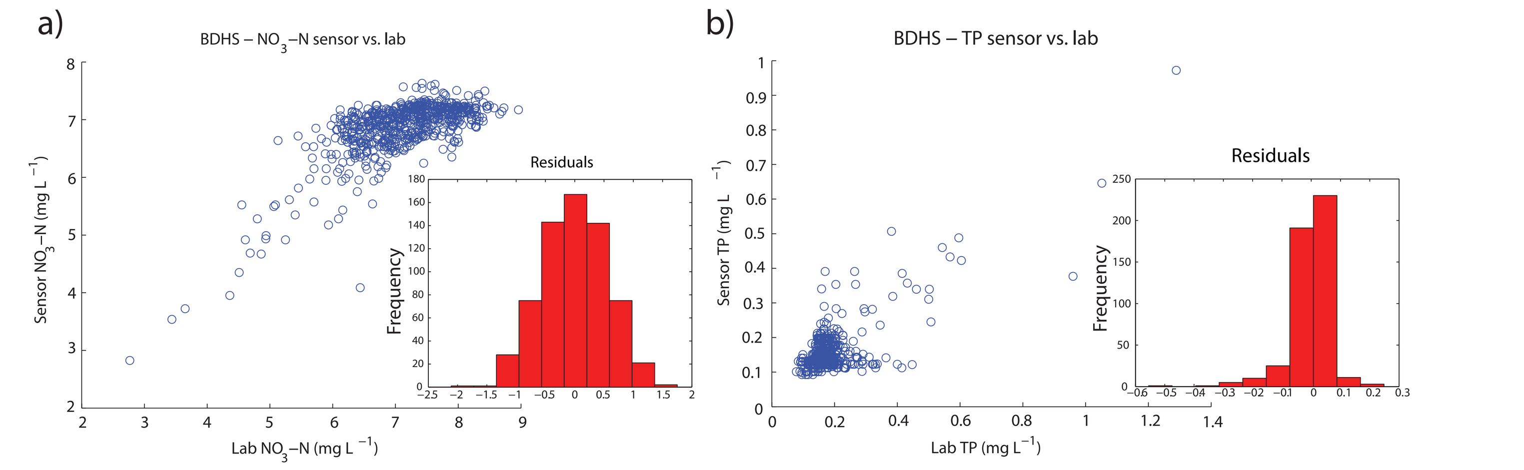


Figure 4: Plots showing the relationship between paired sensor and lab data at Brixton Deverill and the distribution of residuals for a) nitrate-N and b) total phosphorus.

- Field sensor data were validated using paired laboratory data, where the lab data was treated as the 'truth' (figure 4).
- Residuals were found to be autocorrelated but homoscadastic, the errors were calculated using the 1st order autoregressive model accordingly.
- Sites where only laboratory data was available errors were determined by examining repeated lab standards (not shown).
- Lab errors were found to be heteroscadastic but were assumed to be independent, errors were modelled to include the heteroscadasticity.

- The modelled errors were applied to each data set to produce 100 replicate data sets covering the range of uncertainty in the flow, and nutrient data.
- Figure 5 shows the replicate data sets in green and original data sets in blue, for two of the field sites.
- In general ithe sensor data sets had wider uncertainty bounds than the laboratory analysis.
- The replicate data sets were then used to calculate repeated estimates of the nutrient loads (figure 6).

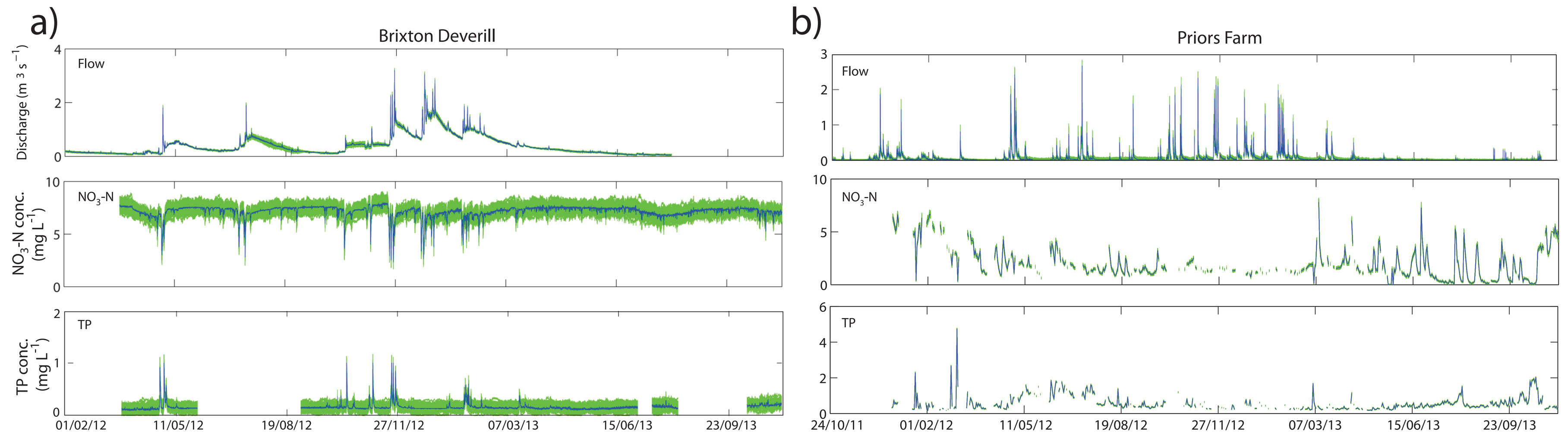


Figure 5: Time series of flow, nitrate-N and total phosphorus at a) Brixton Deverill and b) Priors farm, where green shows 100 realisations of the time series following the error modelling and the blue shows the original data set.

Nutrient loads and implications

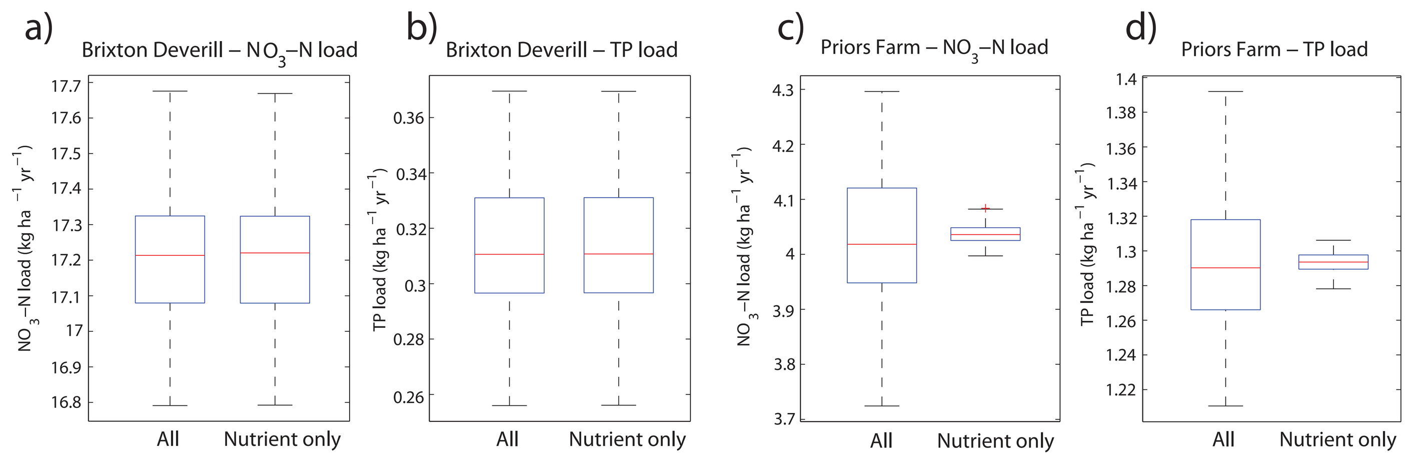


Figure 6: Range of nutrient load estimates calculated from the 100 replicated data sets for a) Brixton Deverill and b) Priors Farm, comparing where both flow and nutrient uncertainty are included (all) and where only nutrient uncertainty is considered.

- At Brixton Deverill the flow uncertainty in is small and therefore the uncertainty in the sensor information is the most influential on the resulting load, ranging from 16.8-17.7 kg ha^{-1} yr^{-1}.
- At Priors Farm, flow uncertainty plays a larger role, increasing the distribution of the load estimates, as well as a small shift in the median load estimate.

This work highlights the importance of not using a single stage-discharge curve but using a velocity-area method for flow calculation in small headwater streams - the seasonal behaviour is extremely dynamic.

It is important to include both flow and nutrient uncertainties in water quality analysis as both can play an important role in producing robust analysis of nutrient behaviour, even when high-temporal resolution data is available.