Blue-Green Systems

Development of the data-driven models for accessing the impact of design variables on heavy metal removal in constructed wetlands

Jiadong Zhang [iD] [a,*], Veljko Prodanovic **IWA**[a], Anna Lintern[b] and Kefeng Zhang **IWA**[a]

[a] Water Research Centre, School of Civil and Environmental Engineering, UNSW Sydney, High Street, Kensington, NSW 2052, Australia
[b] Department of Civil Engineering, Monash University, Wellington Road, Clayton, VIC 3800, Australia
*Corresponding author. E-mail: jiadong.zhang@unsw.edu.au

[iD] JZ, 0000-0003-2100-0037

ABSTRACT

Constructed wetlands are a type of green infrastructure commonly used for urban stormwater treatment. Previous studies have shown that the various design characteristics have an influence on the outflow heavy metal concentrations. In this study, we develop a Bayesian linear mixed model (BLMM) and a Bayesian linear regression model (BLRM) to predict the outflow concentrations of heavy metals (Cd, Cu, Pb and Zn) using an inflow concentration (C_{in}) and five design variables, namely media type, constructed wetland type (CWT), hydraulic retention time, presence of a sedimentation pond (SedP) and wetland-to-catchment area ratio (Ratio). The results show that the BLMM had much better performance, with the mean Nash–Sutcliffe efficiency between 0.51 (Pb) and 0.75 (Cu) in calibration and between 0.28 (Pb) and 0.71 (Zn) in validation. The inflow concentration was found to have significant impacts on the outflow concentration of all heavy metals, while the impacts of other variables on the wetland performance varied across metals, e.g., CWT and SedP showed a positive correlation to Cd and Cu, whereas media and Ratio were negatively correlated with Pb and Zn. Results also show that the 100-fold calibration and validation was superior in identifying the key influential factors.

Key words: cross-validation, green infrastructure, nature-based solutions, Sponge city, stormwater

HIGHLIGHTS

- Data-driven models were developed to predict the heavy metal performance in constructed wetlands.
- The impact of variables was assessed using posterior distributions of the regression coefficients.
- The inflow concentrations had a significant positive impact on outflow concentrations for all metals.
- The most influential variables are varied for metals.
- Multiple-fold cross-validation is needed to robustly identify the key variables.

1. INTRODUCTION

Urban stormwater is a significant factor contributing to the declining health of waterways (McGrane 2016). Heavy metals are types of harmful contaminants of concern in urban stormwater due to their toxicity to aquatic species (Kivaisi 2001). As a result of anthropogenic activities, e.g., usage of automobile and metallic structures, heavy metals accumulate on impervious surfaces and then can be conveyed by urban stormwater into receiving waterways (Sakson *et al.* 2018). Therefore, stormwater management approaches, such as water-sensitive urban design systems, also known as green infrastructure (GI), sustainable urban drainage systems or low-impact development, are being used in urban areas as an integrated solution to treat urban stormwater before it is released into receiving waterways (Roy *et al.* 2008; Li *et al.* 2020; Winston *et al.* 2020).

Constructed wetlands (also known as artificial wetlands) are among the most popular solutions for urban stormwater treatment (Barton & Argue 2007). Constructed wetlands predominantly consist of a sedimentation basin and a macrophyte zone containing media and vegetation. As urban stormwater flows through the sedimentation basin and the macrophyte zones, pollutants are removed through physical (e.g., sedimentation and filtration), chemical and biological (e.g., root zone uptake and microbial degradation) processes (Vymazal 2005; Midhun *et al.* 2016). Previous studies have demonstrated the high removal rates of heavy metals from urban stormwater, with average removal rates of 57, 65, 65 and 52% for Cd, Cu, Pb and Zn, respectively

(Birch *et al.* 2004; Lucas *et al.* 2015). However, the performance can vary significantly depending on the wetland design characteristics and operational conditions (Birch *et al.* 2004; Lucas *et al.* 2015).

There are three types of constructed wetlands: Horizontal Free-Surface Flow constructed wetland (HFSF), Horizontal Subsurface Flow constructed wetland (HSSF) and Vertical Subsurface Flow constructed wetland (VSSF) (Lucas *et al.* 2015). Owing to a number of different structural designs within each type, it is not clear which wetland type should be the preferred choice for heavy metal removal (Khan *et al.* 2009; Rahman *et al.* 2020). Generally, in constructed wetlands, it was found that the presence or absence of a sedimentation pond impacts the removal of particulate heavy metals, as the sedimentation pond provides an opportunity for the particulate metals to settle out of suspension (DAFF 2021). Furthermore, the design size of constructed wetlands, i.e., the ratio of system size to catchment impervious area, is a crucial design factor in a number of wetland guidelines across the world. The recommended value for this ratio in the USA is >2% (US EPA 2000) and is 1–5% in the UK (Ellis *et al.* 2003). Lucas *et al.* (2015) suggested that the ratio values lower than these recommended values result in lower metal removal efficiency. Hydraulic retention time (HRT) has varied impacts on the system performance. The recommended HRT is usually 24 h (Halcrow/UPRC 2000), and when combined sewer overflows are present, it cannot exceed 48 h, as this might result in clogging, which reduces the pollutant removal performance of the wetland (Uhl & Dittmer 2005). Media rich in organic matter with higher adsorption capacity are expected to have higher heavy metal removal efficiency (Singh & Chakraborty 2020). The ability of plants to uptake heavy metals is variable depending on the plant and the metal (Adams *et al.* 2013). For instance, the Cu removal rate can vary from 48.28% removal (*Phragmites australis*) to 99.10% (*Iris pseudacorus*) removal in systems with different vegetations (Wang *et al.* 2014).

Similarly, operational conditions can also influence the system performance. Yadav *et al.* (2010) found that higher inflow concentrations often lead to higher outflow concentrations due to the inability of the wetland to treat the high metal concentrations. Previous studies have found that other operational conditions (e.g., rainfall depth, infiltration rate, antecedent dry period length and evapotranspiration) generally influence the pollutant removal performance of GI (e.g., biofilters; Zhang *et al.* 2021). The processes, however, are quite complicated as they are also subject to various environmental conditions (e.g., pH, redox and temperature) (Schmitt *et al.* 2015; Walaszek *et al.* 2018).

It is clear that there is a complex interaction between several key designs and operational elements used in constructed wetlands that influence their heavy metal (and other pollutants) removal performance. This makes the problem of wetland design optimisation challenging for standard mesocosm studies, since they cannot highlight which elements have the greatest effect on effluent water quality. Computational modelling is an approach that can simulate the system performance based on an understanding of the underlying pollutant removal processes of constructed wetlands. Currently, there is a lack of models that can accurately predict heavy metal removal processes in constructed wetlands. The Stormwater Water Management Model is an example of a commonly used model which represents constructed wetlands and contaminants as blocks with the conceptual model (Kincanon & McAnally 2004). Tomenko *et al.* (2007) have reviewed other less complex models and have identified disadvantages for them, e.g., the first-order decay models usually oversimplified the interaction of constructed wetland modelling, whereas the Monod-type models (i.e., an extension of first-order decay model) could not always assess the parameters directly.

More recently, data-driven models have been used to investigate and predict the pollutant removal performance of GI (Fang *et al.* 2021; Nguyen *et al.* 2021; Zhang *et al.* 2021). The strength of the data-driven model is that it can establish the relationships between removal efficiency and system characteristics with a limited understanding of the underlying removal processes and without quantifying the parameters that define these removal processes. For example, both Tomenko *et al.* (2007) and Nguyen *et al.* (2022) have used data-driven models – artificial neural network (ANN) to predict biochemical oxygen demand (BOD5) effluent concentration in constructed wetlands with high accuracy (i.e., mean $R^2 = 0.997$ and 0.925 for the proposed ANN models in the two studies). A regression model has been conducted to analyse the trends for TP (total phosphorus), TN (total nitrogen), BOD5 and TSS (total suspended solids) removal performance in constructed wetlands and design factors with R^2 from 0.49 to 0.62 (Son *et al.* 2010). The Bayesian modelling approaches have also been used to understand the key design characteristics and elements that impact the treatment performance of GI (Liu *et al.* 2017), including constructed wetlands (e.g., Nguyen *et al.* 2018). Main *et al.* (2017) have applied the Bayesian method to predict the neonicotinoid concentrations in constructed wetlands. The Bayesian approaches

are commonly used for data-driven modelling due to their computational efficiency, flexibility of analysis data type, ability to clearly quantify parameter uncertainty and ability to assign a non-informative prior distribution for parameters that we have little *a priori* information and data about (Kazembe *et al.* 2008; Liu *et al.* 2017). However, models that predict heavy metal removal in constructed wetlands are still lacking. In addition, we also lack an understanding of which design characteristics and operational conditions need to be optimised to enhance the heavy metal treatment performance in constructed wetlands.

This study aims to develop the data-driven models to predict the heavy metal removal performance of constructed wetlands and investigate wetlands' key design characteristics and operational condition. The data-driven models were developed using a Bayesian modelling approach to create a detailed understanding of the relationship between the outlet concentration of each metal and wetland characteristics. They were then tested using a systematic approach to assess the model performance. Using the developed models, we also investigated the impact of constructed wetland design characteristics and inflow concentrations on the heavy metal outlet concentration.

2. METHODOLOGY

2.1. Data collection

A systematic literature review was performed to collect data on constructed wetlands' heavy metal treatment performance in field-scale and pilot-scale systems. This systematic review was conducted following the method proposed by Pickering & Byrne (2014). A combination of keywords – 'metals', 'urban stormwater' and 'constructed wetland' was used to search for the relevant literature in ScienceDirect, Web of Science and Google scholar. The initial screening resulted in 207 papers (on 21 May 2021), for which the main body, appendixes and other linked sources were reviewed. Based on the literature review, five design variables (media type, constructed wetland type (CWT), HRT, presence of a sedimentation pond (SedP) and wetland-to-catchment area ratio (Ratio)) and inflow and outflow concentrations have been selected as input variables for this study (Table 1). It is noted that although the type of plants could influence the performance, it was not included as different studies often used similar plants and it was difficult to differentiate between their treatment performance levels. From the initial screening, a total of 12 papers (Supplementary Material, Table S-1) passed the pre-designed requirements, i.e., the paper must have all of the required data/variables as shown in Table 1. In cases when the Ratio was missing, it was calculated manually using the catchment area, imperviousness and the constructed wetland surface area. In cases when the C_{in} was missing, the C_{out} and removal efficiency were used to estimate the C_{in}. Attempts were made to collect the raw data (i.e., the data collected from each event), but there were cases that the paper only reported mean values. These mean values were still included to gather more data points for the models. Individually, 47, 67, 62 and 75 datasets (that included all seven

Table 1 | Summary of selected variables

	Variable name	Unit	Data type	Description
Design variables	Ratio	–	Numerical	The ratio of constructed wetland surface area to the contributing catchment impervious area
	SedP	–	Categorical	Presence of sedimentation pond: '1' for presence, '0' for absence
	Media	–	Categorical	Type of media: '0' for soil, gravel and sand, '1' for the presence of other materials[a]
	HTR	Hours (h)	Numerical	Hydraulic retention time: the average time the stormwater remains in the system
	CWT	–	Categorical	Constructed wetland type: '0' for HFSF, '1' for HSSF, VSSF or Hybrid[b]
Other	C_{in}	μg/L	Numerical	Inflow concentrations
Parameters	C_{out}	μg/L	Numerical	Outlet concentrations

[a]Soil, gravel and sand were the most commonly used materials for constructed wetland media; other materials mean that the commonly used materials were mixed with other materials including the volcanic gravel, pumice, quartz stone, organic materials and wood chips.
[b]HFSF, Horizontal Free-Surface Flow constructed wetland; HSSF, Horizontal Subsurface Flow constructed wetland; VSSF, Vertical Subsurface Flow constructed wetland; Hybrid, two or more types of constructed wetlands used.

parameters) were collected for Cd, Cu, Pb and Zn, respectively (statistical summary of the data is presented in Supplementary Material, Table S-2).

2.2. Data pre-processing

All categorical data are organised into two groups. SedP was classified into yes/no categories (Yes = 1/No = 0). For media types, '0' was assigned to media types made of soil, sand, gravel or the combination of those materials which are most commonly used (Rahman *et al.* 2020), '1' was given to media types containing organic matter, volcanic gravel, pumice, quartz stone and/or wood chips, which have been reported to influence the system performance (Marchand *et al.* 2010; Lucas *et al.* 2015). When considering the effect of CWT, due to the limited water filtration through the soil (Parde *et al.* 2021) in the HFSF, it was regarded as an individual category (assigned as '0'), whereas HSSF, VSSF and hybrid constructed wetland were assigned as '1'.

Most collected data did not follow the normal distribution. Thus, we transformed the data using the Box–Cox transformation to increase the symmetry of the datasets. This was followed by *z*-score standardisation to make the data dimensionless. This approach was also used by other similar studies for data pre-treatment (e.g., Liu *et al.* 2017).

2.3. Model structure and implementation (Bayesian model approach)

In this study, we assume that each heavy metal outflow concentration in the *i*th site, $C_{\text{out},i}$, follows a normal distribution, with mean μ_i and variance σ^2 (Equation (1)):

$$P(C_{\text{out},i} \mid \mu_i, \sigma^2) \sim N(\mu_i, \sigma^2) \tag{1}$$

σ^2 represents within-site variation, which includes the measurements' error and variability in concentration levels within the site. The mean concentration μ_i is modelled using the following two different approaches:

1. *BLRM*:

$$\mu_i = \alpha + \sum_{n=1}^{6} \beta_n * \text{characteristics}_n(i) \tag{2}$$

where μ_i is a linear function of different variables (including the five design variables, and C_{in} as shown in Table 1). α is the regression intercept, and β_n is the regression parameter (slope) for the *n*th variable.

2. *BLMM*:

$$\mu_i = (\alpha + \text{site}_i) + \sum_{n=1}^{6} \beta_n * \text{characteristic } s_n(i) \tag{3}$$

$$P(\text{site}_j \mid 0, \sigma_a^2) \sim N(0, \sigma_a^2)$$

where $j = 1, 2, \ldots, 12$

In the BLMM model, an additional *site*$_j$ parameter was introduced to account for the variability across different sites, and it was assumed to follow normal distribution with mean 0 and variance σ_a^2. An extra data column – 'site' – was then added to the dataset in order to implement the extra variable '*site*$_j$', and 1–12 were assigned for 12 individual sites (i.e., 12 individual papers). Thus, disparate linear regression lines were produced for different sites with individual intercepts. The model can potentially inform us that whether accounting for site variability is needed.

For both models, α was the regression intercept. The magnitude of the regression coefficients (i.e., β_n) indicated the importance of influential variables. The four-model parameters ($\alpha, \beta_n, \sigma^2, \sigma_a^2$) have been assumed independent to each other. Regarding the prior distributions in previous studies, non-informative uniform priors were assigned to σ^2, σ_a^2 as no negative value should be assigned to standard deviations; non-informative normal distributions were assigned to α, β_n for numerical data (Liu *et al.* 2017). Our preliminary testing showed that the models were not able to converge when using binomial distributions for the regression coefficients of categorical variables; therefore, a normal distribution was still assigned to the categorical data in this study. The same approach (i.e., using normal distribution for the regression coefficients of categorical variables) was also

implemented by the other previous studies (Liao *et al.* 2019).

$$\alpha \sim N(0, 0.01); \beta_n \sim N(0, 0.01);$$

$$\sigma \sim \text{Uniform}(0, 100); \sigma_a \sim \text{Uniform}(0, 100) \tag{4}$$

2.4. Model implementation

The samples from posterior probability distributions were obtained using the Markov Chain Monte Carlo with the Gibbs sampling. The 'JAGs' package was used in the R environment. The model was simulated with four chains and 50,000 iterations. The first 5,000 iterations were abandoned as a 'warm-up' period to allow the convergence of the Markov chain (Rhat value of approximately 1). To reduce the effect of auto-correlation, only every second iteration was adopted when estimating the posterior distributions.

2.5. Model testing and application

2.5.1. *N*-fold calibration and validation

The model was tested by the *N*-fold calibration and validation approach, which is a standard method to observe the robustness of model performance (Zhang *et al.* 2021). In this study, 100-fold calibration and validation (with a 70–30 split of data randomly for calibration and validation) was used, following the suggestions given in Fang *et al.* (2021) when the model is applied to the highly variable data collected from various literature works. Model fit was assessed using the Nash–Sutcliffe efficiency (NSE) (i.e., a potentially reliable statistic that has been widely used) (Nash & Sutcliffe 1970). A deviance information criterion (DIC), a Bayesian criterion for model comparison that is tried to find a balance between model complexity and data fitness (Spiegelhalter *et al.* 2002), was also estimated.

$$NSE = \frac{\sum\limits_{i=1}^{n} (OBS_i - PRE_i)^2}{\sum\limits_{i=1}^{n} (OBS_i - \overline{PRE_i})^2} \tag{5}$$

where *OBS* is the observation, *PRE* is the prediction and *n* is the number of data points.

2.5.2. Impact of design characteristics and operational condition on the heavy metal removal performance

The posterior distribution of the regression coefficients was used to assess the impact of the independent variables on the performance of constructed wetlands in removing heavy metals (in terms of C_{out}). This was done using the following two approaches: (i) using the full dataset to have an overall view of the impact and (ii) using the 100-fold calibration and validation results, to investigate the variability of the impacts, i.e., the posterior distribution from each fold was generated to see whether the impact/distribution varied across multiple runs. This was done to investigate whether the impact of the variables (evaluated using the posterior distribution of the regression coefficients) could be done just on the full dataset, or should be through multiple-fold calibration and validation.

3. RESULTS AND DISCUSSION

3.1. Model performance

The results from the 100-fold calibration and validation of the two models are presented in Figure 1, and the NSE and DIC values are summarised in Table 2. Based on the calibration, the BLMM was found to have better performance in simulating C_{out} of all heavy metals, i.e., Cd (mean NSE $\pm$ standard deviations $= 0.74 \pm 0.06$), Cu (0.75 ± 0.06), Pb (0.51 ± 0.15) and Zn (0.74 ± 0.05) (Table 2). This is on average around 61% higher than the BLRM. When validating the models, we had lower NSEs, especially for the BLRM (Figure 1). For Pb, the mean NSE for the BLMM and the BLRM dropped to 0.28 and -2.24, respectively, and the standard deviation increased to ± 0.45 (three times larger) and ± 4.56 (45.6 times larger), respectively, when compared to calibration (Table 2). Similarly, poorer validation of data-driven models using the literature data was observed when predicting heavy metal outflow concentrations by biofilters (Fang *et al.* 2021). However, the validation of Zn was more satisfactory using the BLMM, with a minor drop for mean and an increase in standard deviation observed (i.e., NSE$_{validation} = 0.71 \pm 0.08$). Similarly, a larger drop in mean NSE$_{validation}$ of Zn was observed for
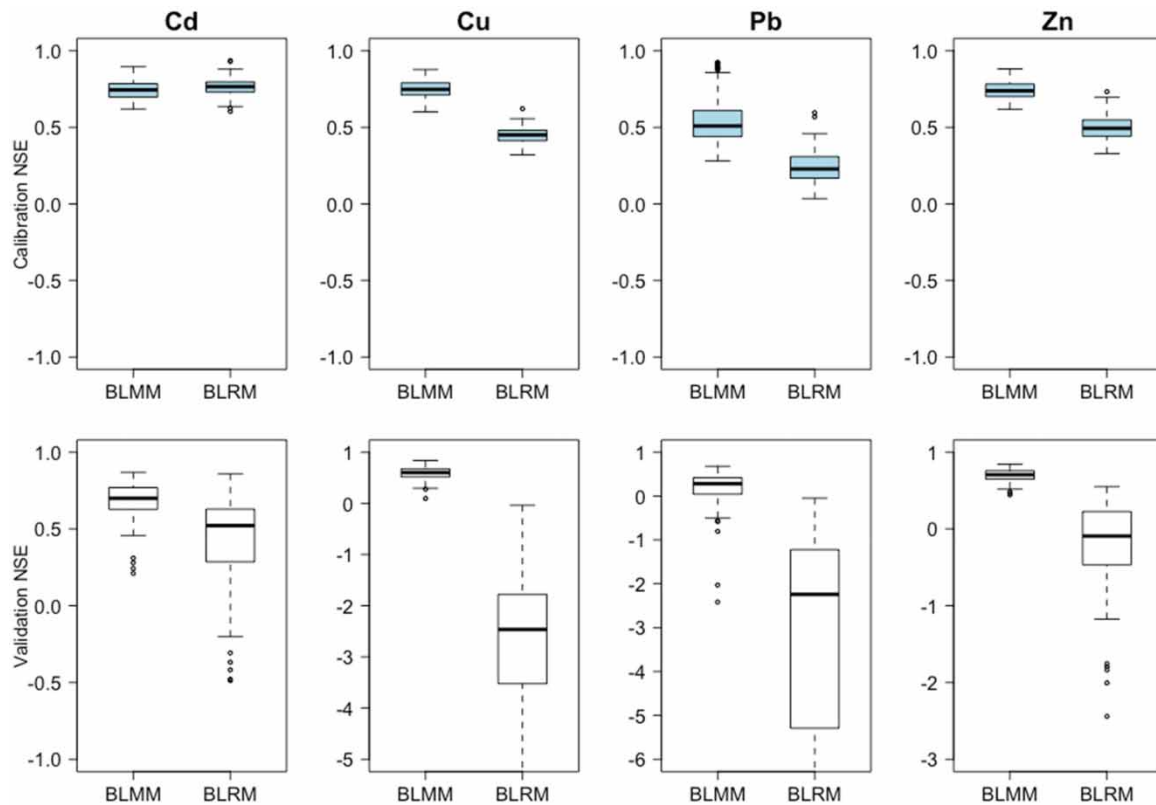
Figure 1 | Box plots of NSE values generated by the 100-fold calibration and validation using the BLMM and the BLRM. Small circles stand for outliers and the black bold lines stand for the median NSE values. The interquartile ranges from 25th to 75th percentile values.

Table 2 | Summary of the mean and standard deviations for both the BLMM and the BLRM for NSE and DIC values based on the 100-fold calibration and validation

| Metal | Model | NSE | | DIC | |
		Calibration	Validation	Calibration	Validation
Cd	BLMM	0.74 ± 0.06	0.70 ± 0.13	56.3 ± 8.8	23.5 ± 8.6
	BLRM	0.77 ± 0.06	0.52 ± 0.76	60.1 ± 7.8	29.8 ± 38.1
Cu	BLMM	0.75 ± 0.06	0.60 ± 0.13	88.9 ± 14.2	40.3 ± 10.8
	BLRM	0.45 ± 0.05	−2.46 ± 1.89	113.3 ± 6.3	120.5 ± 36.5
Pb	BLMM	0.51 ± 0.15	0.28 ± 0.45	114.2 ± 26.5	45.7 ± 47.4
	BLRM	0.23 ± 0.10	−2.24 ± 4.56	123.6 ± 10.3	78.4 ± 43.8
Zn	BLMM	0.74 ± 0.05	0.71 ± 0.08	100.5 ± 12.7	39.6 ± 9.0
	BLRM	0.49 ± 0.08	−0.09 ± 0.73	122.2 ± 8.7	63.2 ± 32.4

the BLRM (i.e., dropped 0.58), but it was much smaller than the drop produced by Pb. This further confirms the necessity of using the multiple-fold calibration and validation to ensure the robustness of the models.

Overall, the best fits were found for Cd, Cu and Zn for the BLMM, with the mean NSE$_{validation}$ values of 0.70, 0.60 and 0.71, respectively. The ranges of NSE values were also small and all positive, with only a few outliers observed. The BLRM did not perform well for all heavy metals, especially for Cu, Pb and Zn, which had negative mean NSE values with high variability (Figure 1). Previous studies have also reported the variability in model performance across different pollutants when using numerical modelling approaches to assess constructed wetlands (Dharmasena *et al.* 2021), e.g., the values of R^2 for TSS, TN and TP were 0.434, 0.124 and 0.490, respectively.

The predictive errors, represented by the DIC, further proved that the BLMM was better than the BLRM in terms of model complexity and model fitness (maximum mean DIC for the BLMM and the BLRM = 114.21

and 123.59, respectively). In addition to that, the prediction accuracy of individual heavy metal outflow concentration followed the same result from the NSE comparison, whereby Pb had the worst performance for both models (i.e., maximum mean DIC). The better performance of the BLMM indicates that more variabilities were produced between different sites, and our proposed intercept adjustment factor (i.e., site$_j$) successfully countered part of caused variabilities. Thus, the BLMM was used to explore the impact of variables on the wetland performance in the next section.

3.2. Impact of variables on the wetland performance

Figure 2 presents the posterior distributions of regression coefficients generated from the calibration of the full dataset. Specifically, the posterior distribution of regression coefficients of inflow concentration (C_{in}) showed that the C_{in} had a positive impact on the outlet concentration (C_{out}) with a small variance, i.e., the mean and standard deviation values of regression coefficients for Cd, Cu, Pb and Zn are 0.41 ($\pm$ 0.12), 0.37 ($\pm$ 0.09), 0.47 ($\pm$ 0.20) and 0.15 ($\pm$ 0.08), respectively. These findings have a good agreement with a previous study by Yadav *et al.* (2010) who found out that the number of metal ions increased at higher C_{in}. Studies on stormwater bioretention systems (which function very similar to VSSF) also found that the higher C_{in} of heavy metals can lead to increased C_{out} (Fang *et al.* 2021).

The impact of the SedP on the system performance was highly variable, demonstrated by the largest 95% credible intervals (compared to all other variables), spanning across the zero line of 'no effect' (i.e., the vertical dashed line in Figure 2). The exception was for Cu, where a pure positive impact was found (mean $\pm$ standard deviations
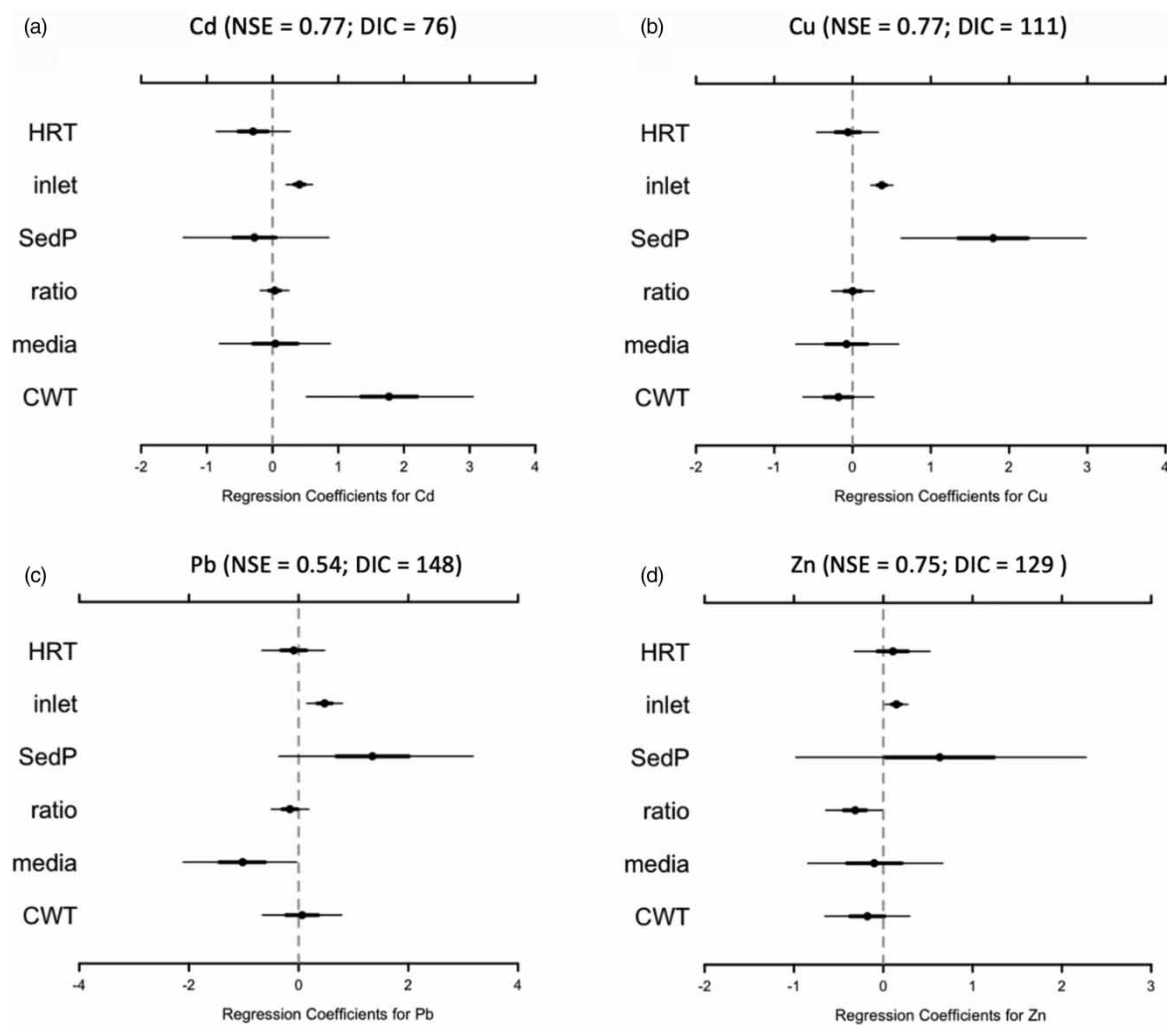

Figure 2 | Posterior distributions of regression coefficient plots for (a) Cd, (b) Cu, (c) Pb and (d) Zn – results from the calibration on the full dataset. Short thick lines are 50% posterior credible intervals, the thin lines indicate 95% credible intervals and the dots represent the mean values.

of coefficient = 1.8 ± 0.75). This shows that the SedP could increase the outlet concentration of Cu. Previous studies show mixed findings regarding the partitioning of Cu in dissolved and particulate form in stormwater, e.g., Maniquiz-Redillas & Kim (2014) observed a higher fraction of Cu in dissolved form, while Huber *et al.* (2016) found more Cu presented in particulate form. Thus, it is possible that Cu can transform between two forms, and accumulated Cu in the sedimentation pond can be resuspended during flow events and cause it to leach out of the system, e.g., a significant drop of Cu removal efficiency can be observed in the largest rainfall (from 56–86% to 21%) (Birch *et al.* 2004).

The posterior distributions of the regression coefficients of HRT show very little impact on the heavy metal outlet concentrations, as the mean regression coefficient values were very close to zero (i.e., −0.30 for Cd, −0.06 for Cu, −0.09 for Pb and 0.10 for Zn), and the credible intervals are across the zero line of 'no effect'. The literature also reported mixed impacts of HRT on the performance of constructed wetlands. Chen (2009) and Baum *et al.* (2021) have shown that, in general, a higher HRT can lead to improved system performance (i.e., lower outlet concentrations). However, very long HRT (over 48 h) might lead to poor performance gradually due to clogging in subsurface wetlands. Negative removal efficiency has been observed for heavy metals in saturated or blocked systems (i.e., −97 to 97% for Cu) (Lucas *et al.* 2015; Upadhyay *et al.* 2016). In this study, the HRT varied considerably in the raw data for all metals (i.e., min–max was 0.4–127.2 h), which could be the possible cause for having negligible effects.

The impact of the Ratio on the heavy metal outlet concentrations was very small, with low values of mean regression coefficients, which also crossed the zero line of 'no effect'. The exception was found for Zn, where a significant negative impact was observed (mean ± standard deviations of coefficient = −0.32 ± 0.19; Figure 2). This indicates that a larger system size can lead to lower outflow concentrations. Commonly, the Ratio should be designed, so that the constructed wetland can retain a certain amount of runoff and pollution reduction (Maillard *et al.* 2011; Lucas *et al.* 2015). In this study, the Ratio values did not vary significantly, i.e., the mean ± standard deviations values for the Ratio of Cd, Cu, Pb and Zn were 2 ± 0.6%, 2 ± 0.8%, 2 ± 0.7% and 2 ± 0.7%. Carleton *et al.* (2001) and Lucas *et al.* (2015) have explored in their study that the larger Ratio could increase the removal efficiencies for all four heavy metals; however, oversized systems may experience prolonged dry conditions that will harm the system health.

It was found that the subsurface flow and hybrid wetlands had better Cu and Zn removal, demonstrated by the significant negative impacts of CWT on their removal, i.e., the 50% credible intervals are all negative side (Figure 2). Previous studies also found that the subsurface flow wetlands or hybrid systems generally show better system performance in removing heavy metals (e.g., Marchand *et al.* 2010; Parde *et al.* 2021), probably due to an enhanced adsorption of heavy metals as the water flows through the media. However, the opposite results were observed for Cd, where the HFSF appeared to be better, demonstrated by the significant positive impacts (i.e., mean ± standard deviations of coefficient = 1.77 ± 0.96; Figure 2). No previous studies have directly indicated the better performance of HFSF for Cd removal over other wetland types; however, many studies have found that Cd achieved the highest removal compared with other heavy metals (i.e., 91.9 and 99.6%) in HFSF wetlands (Kanagy *et al.* 2008; Khan *et al.* 2009). In terms of Pb removal, CWT had no clear impact, as the mean regression coefficient value was very close to zero (0.06 ± 0.44; Figure 2), and the regression lines crossed the zero line of 'no effect'. This is probably due to the fact that sedimentation is the dominant process for particulate heavy metal removal, and Pb is mostly in particulate form (i.e., Pb had a higher particulate fraction than Cu and Zn) (Mosley & Peake 2001; Liu *et al.* 2007). Thus, Pb can be readily removed irrespective of water filtration through soil (e.g., Aswad *et al.* 2020).

The media type has the strongest negative impact on the Pb outlet concentration (−1.05 ± 0.63), while no impacts were found for the Cd (0.04 ± 0.52), Cu (−0.08 ± 0.40) and Zn (−0.10 ± 0.46) (the regression lines crossed the zero line of 'no effect'; Figure 2). This shows that adding other materials (e.g., the volcanic gravel, pumice, quartz stone, organic materials and wood chips) could lead to better performance in removing Pb. Studies have also shown that adding organic materials to soil could be a beneficial approach in enhancing heavy metal treatment, by cation exchange, complexation and precipitation (Yadav *et al.* 2012; Saeed *et al.* 2021). Specifically, Vymazal *et al.* (2010) summarised the reasons as follows: Pb may bind strongly to organic matters, carbonate is an effective sink for Pb and Pb is preferred to be complexed to organic matters in either rhizosphere solutions or on the root surface.
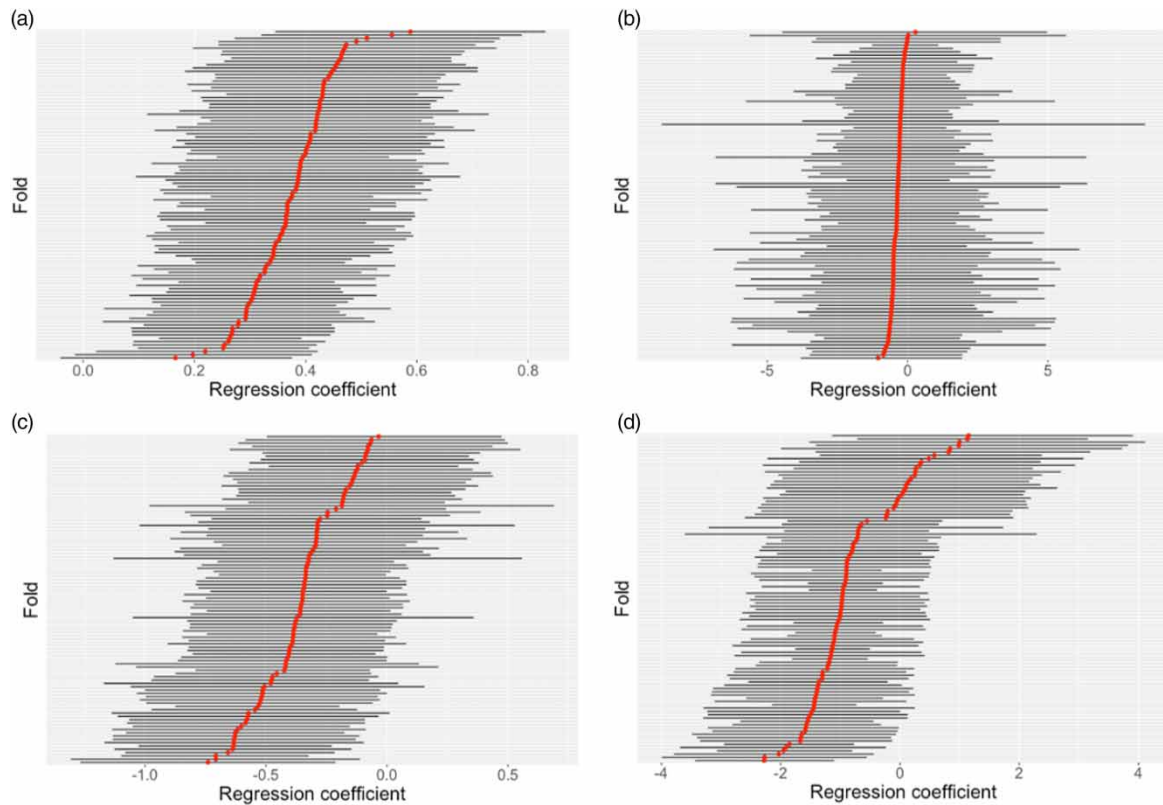
Figure 3 | Change in the distributions of regression coefficients for a specific variable and heavy metal across the 100-fold calibration and validation. (a) Regression coefficient of C_{in} for Cu, (b) Regression coefficient of SedP for Cd, (c) Regression coefficient of Ratio for Zn, and (d) Regression coefficient of media for Pb.

3.3. Variability in the impacts of variables: results from the 100-fold calibration and validation

The distributions of selected variables for specific heavy metals across the 100-fold calibration and validation are presented in Figure 3. The full distributions of regression coefficients of inflow concentrations (C_{in}) for Cu are all on the positive side, i.e., regression coefficient means and 95% credible intervals were solely positive, except for two-fold that had negligible negative values at 5% confident intervals (i.e., −0.01 and −0.04) (Figure 3(a)). Similar results were found for other metals (shown in Figures S-2–S-5 in the Supplementary Material). This has good agreement with the results further generated by the model using the full dataset (as in the previous section, which used the full dataset). A good agreement was also found on the impact of sedimentation ponds on the Cd removal between two approaches, i.e., using the full dataset/100 – fold calibration and validation – both showing no effects (Figure 3(b)).

Figure 3(c) shows that all the mean values of the regression coefficient of Ratio for Zn were negative (−0.74 to −0.03); however, in 17 of the 100 calibration and validation cases, the 50% credible intervals of posterior distribution of the coefficients crossed the zero line of 'no effect'. This is inconsistent with the findings from the full dataset simulation, which show pure significant negative effects (mean regression coefficient = −0.32 ± 0.19; Figure 2). This shows that the posterior distribution of the regression coefficients of Ratio may change from pure negative values to also cover the positive ranges. Similarly, inconsistent findings were observed between the full dataset approach and the 100-fold approach for the impact of media on Pb removal in constructed wetlands (Figure 3(d)). In these cases, checking the impact of variables through the *N-fold* calibration and validation approach is preferable.

4. CONCLUSIONS

This study was performed to investigate the possibility of using the BLMM and the BLRM to explore the relationships between independent variables (i.e., design variables and inflow concentration) and outlet concentrations of four heavy metals (i.e., Cd, Cu, Pb and Zn) in a constructed wetland. Better performance was found by the BLMM

than the BLRM, indicating the need to consider the variance generated from different sites. The assessment of the posterior distribution of regression coefficients shows that C_{in} has a consistently positive impact on C_{out} for all heavy metals, while the impacts of design variables (e.g., HRT, Ratio, CWT, SedP and media type) varied for different heavy metals, i.e., CWT and SedP showed a positive correlation to Cd and Cu, whereas media and Ratio were negatively correlated with Pb and Zn. Results also show that the 100-fold calibration and validation was superior in identifying the key influential factors.

It is expected that the model developed in this study could be used by practitioners to design future wetlands and/or understand how existing wetlands are working in terms of heavy metal removal. In addition, the under-standing of the key design characteristics and operational conditions that influence the heavy metal removal will help guide practitioners and stormwater managers on the optimal design of constructed wetlands for pollutant removal performance.

The results of this study, however, should be regarded as preliminary and future work should be done to further validate the findings using more data collected from the literature or through field monitoring. Additionally, the extension of the methodology might be achieved by introducing the Bayesian hierarchical model to explore the influences where the data for operational conditions are available (e.g., rainfall depth and antecedent dry weather period) due to the fact that the heavy metal removal mechanisms could be affected by operational conditions as well.

ACKNOWLEDGEMENT

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Adams, A., Raman, A. & Hodgkins, D. 2013 How do the plants used in phytoremediation in constructed wetlands, a sustainable remediation strategy, perform in heavy-metal-contaminated mine sites? *Water and Environment Journal* **27**, 373–386. https://onlinelibrary.wiley.com/doi/10.1111/j.1747-6593.2012.00357.x.

Aswad, Z., Ali, A. & Al-Mhana, N. 2020 Vertical subsurface flow and free surface flow constructed wetlands for sustainable power generation and real wastewater selective pollutants removal. *Journal of Engineering and Sustainable Development* **24**, 91–102.

Barton, A. B. & Argue, J. R. 2007 A review of the application of water sensitive urban design (WSUD) to residential development in Australia. *Australasian Journal of Water Resources* **11** (1), 31–40.

Baum, P., Kuch, B. & Dittmer, U. 2021 Adsorption of metals to particles in urban stormwater runoff-does size really matter? *Water* **13**, 309.

Birch, G. F., Matthai, C., Fazeli, M. S. & Suh, J. Y. 2004 Efficiency of a constructed wetland in removing contaminants from stormwater. *Wetlands* **24** (2), 459.

Carleton, J. N., Grizzard, T. J., Godrej, A. N. & Post, H. E. 2001 Factors affecting the performance of stormwater treatment wetlands. *Water Research* **35** (6), 1552–1562.

Chen, M. 2009 Study on the heavy metals removal efficiencies of constructed wetlands with different substrates. *Journal of Water Resource and Protection* **01**, 22–28.

DAFF 2021 *Sediment Basins*. Department of Agriculture, Fisheries and Forestry, Queensland, Australia. Available from: https://wetlandinfo.des.qld.gov.au/resources/static/pdf/management/sediment-basins-factsheet-230114-v1.pdf (accessed 31 July 2021).

Dharmasena, T., Chua, L. H. C., Barron, N. & Zhang, H. 2021 Performance assessment of a constructed wetland using a numerical modelling approach. *Ecological Engineering* **173**, 106441.

Ellis, J. B., Shutes, R. B. E. & Revitt, D. M. 2003 *Constructed Wetlands and Links with Sustainable Drainage Systems*. Environment Agency, Bristol. Available from: https://eprints.mdx.ac.uk/6076/1/SP2-159-TR1-e-p.pdf.

Fang, H., Jamali, B., Deletic, A. & Zhang, K. 2021 Machine learning approaches for predicting the performance of stormwater biofilters in heavy metal removal and risk mitigation. *Water Research* **200**, 117273.

Halcrow/UPRC 2000 *Final Field Studies Report*. Protection of the Water Environment Using Balancing Facilities. Halcrow & Urban Pollution Research Centre, Middlesex University, Environment Agency Thames Region.

Huber, M., Welker, A. & Helmreich, B. 2016 Critical review of heavy metal pollution of traffic area runoff: occurrence, influencing factors, and partitioning. *Science of The Total Environment* **541**, 895–919.

Kanagy, L. E., Johnson, B. M., Castle, J. W. & Rodgers, J. H. 2008 Design and performance of a pilot-scale constructed wetland treatment system for natural gas storage produced water. *Bioresource Technology* **99** (6), 1877–1885.

Kazembe, L. N., Chirwa, T. F., Simbeye, J. S. & Namangale, J. J. 2008 Applications of Bayesian approach in modelling risk of malaria-related hospital mortality. *BMC Medical Research Methodology* **8** (1), 6.

Khan, S., Ahmad, I., Shah, M. T., Rehman, S. & Khaliq, A. 2009 Use of constructed wetland for the removal of heavy metals from industrial wastewater. *Journal of Environmental Management* **90** (11), 3451–3457.

Kincanon, R. & McAnally, A. S. 2004 Enhancing commonly used model predictions for constructed wetland performance: as-built design considerations. *Ecological Modelling* **174** (3), 309–322.

Kivaisi, A. K. 2001 The potential for constructed wetlands for wastewater treatment and reuse in developing countries: a review. *Ecological Engineering* **16** (4), 545–560.

Li, Y., Deletic, A. & McCarthy, D. T. 2020 Copper-zeolite integrated stormwater biofilter for nutrient removal – the impact of intermittent wetting and drying conditions. *Blue-Green Systems* **2** (1), 352–363.

Liao, S.-J., Marshall, J., Hazelton, M. L. & French, N. P. 2019 Extending statistical models for source attribution of zoonotic diseases: a study of campylobacteriosis. *Journal of The Royal Society Interface* **16** (150), 20180534.

Liu, J., Dong, Y., Xu, H., Wang, D. & Xu, J. 2007 Accumulation of Cd, Pb and Zn by 19 wetland plant species in constructed wetland. *Journal of Hazardous Materials* **147** (3), 947–953.

Liu, S., Ryu, D., Western, A., Webb, A., Lintern, A., Waters, D. & Thomson, B. 2017 *Modelling the Impact of Land Use and Catchment Characteristics on Stream Water Quality Using a Bayesian Hierarchical Modelling Approach in the Great Barrier Reef Catchments*.

Lucas, R., Earl, E. R., Babatunde, A. O. & Bockelmann-Evans, B. N. 2015 Constructed wetlands for stormwater management in the UK: a concise review. *Civil Engineering and Environmental Systems* **32** (3), 251–268.

Maillard, E., Payraudeau, S., Faivre, E., Grégoire, C., Gangloff, S. & Imfeld, G. 2011 Removal of pesticide mixtures in a stormwater wetland collecting runoff from a vineyard catchment. *Science of The Total Environment* **409** (11), 2317–2324.

Main, A. R., Fehr, J., Liber, K., Headley, J. V., Peru, K. M. & Morrissey, C. A. 2017 Reduction of neonicotinoid insecticide residues in Prairie wetlands by common wetland plants. *Science of The Total Environment* **579**, 1193–1202.

Maniquiz-Redillas, M. & Kim, L.-H. 2014 Fractionation of heavy metals in runoff and discharge of a stormwater management system and its implications for treatment. *Journal of Environmental Sciences* **26** (6), 1214–1222.

Marchand, L., Mench, M., Jacob, D. L. & Otte, M. L. 2010 Metal and metalloid removal in constructed wetlands, with emphasis on the importance of plants and standardized measurements: a review. *Environmental Pollution* **158** (12), 3447–3461.

McGrane, S. J. 2016 Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review. *Hydrological Sciences Journal* **61** (13), 2295–2311.

Midhun, G., Divya, L., George, J., Jayakumar, P. & Suriyanarayanan, S. 2016 Wastewater treatment studies on free water surface constructed wetland system. In: *Integrated Waste Management in India: Status and Future Prospects for Environmental Sustainability* (Prashanthi, M. & Sundaram, R. eds). Springer International Publishing, Cham, pp. 97–109.

Mosley, L. M. & Peake, B. M. 2001 Partitioning of metals (Fe, Pb, Cu, Zn) in urban run-off from the Kaikorai Valley, Dunedin, New Zealand. *New Zealand Journal of Marine and Freshwater Research* **35** (3), 615–624.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* **10** (3), 282–290.

Nguyen, X. C., Chang, S. W., Nguyen, T. L., Ngo, H. H., Kumar, G., Banu, J. R., Vu, M. C., Le, H. S. & Nguyen, D. D. 2018 A hybrid constructed wetland for organic-material and nutrient removal from sewage: process performance and multi-kinetic models. *Journal of Environmental Management* **222**, 378–384.

Nguyen, X. C., Ly, Q. V., Li, J., Bae, H., Bui, X.-T., Nguyen, T. T. H., Tran, Q. B., Vo, T.-D.-H. & Nghiem, L. D. 2021 Nitrogen removal in subsurface constructed wetland: assessment of the influence and prediction by data mining and machine learning. *Environmental Technology & Innovation* **23**, 101712.

Nguyen, X. C., Nguyen, T. T. H., Le, Q. V., Le, P. C., Srivastav, A. L., Pham, Q. B., Nguyen, P. M., La, D. D., Rene, E. R., Ngo, H. H., Chang, S. W. & Nguyen, D. D. 2022 Developing a new approach for design support of subsurface constructed wetland using machine learning algorithms. *Journal of Environmental Management* **301**, 113868.

Parde, D., Patwa, A., Shukla, A., Vijay, R., Killedar, D. J. & Kumar, R. 2021 A review of constructed wetland on type, treatment and technology of wastewater. *Environmental Technology & Innovation* **21**, 101261.

Pickering, C. & Byrne, J. 2014 The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research & Development* **33** (3), 534–548.

Rahman, M. E., Bin Halmi, M. I. E., Bin Abd Samad, M. Y., Uddin, M. K., Mahmud, K., Abd Shukor, M. Y., Sheikh Abdullah, S. R. & Shamsuzzaman, S. M. 2020 Design, operation and optimization of constructed wetland for removal of pollutant. *International Journal of Environmental Research and Public Health* **17** (22), 8339.

Roy, A. H., Wenger, S. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., Shuster, W. D., Thurston, H. W. & Brown, R. R. 2008 Impediments and solutions to sustainable, watershed-scale urban stormwater management: lessons from Australia and the United States. *Environmental Management* **42** (2), 344–359.

Saeed, T., Alam, M. K., Miah, M. J. & Majed, N. 2021 Removal of heavy metals in subsurface flow constructed wetlands: application of effluent recirculation. *Environmental and Sustainability Indicators* **12**, 100146.

Sakson, G., Brzezinska, A. & Zawilski, M. 2018 Emission of heavy metals from an urban catchment into receiving water and possibility of its limitation on the example of Lodz city. *Environmental Monitoring and Assessment* **190** (5), 281–281.

Schmitt, N., Wanko, A., Laurent, J., Bois, P., Molle, P. & Mosé, R. 2015 Constructed wetlands treating stormwater from separate sewer networks in a residential Strasbourg urban catchment area: micropollutant removal and fate. *Journal of Environmental Chemical Engineering* **3** (4), 2816–2824.

Singh, S. & Chakraborty, S. 2020 Performance of organic substrate amended constructed wetland treating acid mine drainage (AMD) of North-Eastern India. *Journal of Hazardous Materials* **397**, 122719.

Son, Y., Yoon, C., Kim, H., Jang, J.-H. & Lee, S. 2010 Determination of regression model parameter for constructed wetland using operating data. *Paddy and Water Environment* **8**, 325–332.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. 2002 Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B* **64** (4), 583–639.

Tomenko, V., Ahmed, S. & Popov, V. 2007 Modelling constructed wetland treatment system performance. *Ecological Modelling* **205** (3), 355–364.

Uhl, M. & Dittmer, U. 2005 Constructed wetlands for CSO treatment: an overview of practice and research in Germany. *Water Science and Technology* **51** (9), 23–30.

Upadhyay, A. K., Bankoti, N. S. & Rai, U. N. 2016 Studies on sustainability of simulated constructed wetland system for treatment of urban waste: design and operation. *Journal of Environmental Management* **169**, 285–292.

US EPA 2000 *A Handbook of Constructed Wetlands, Volume 5: Stormwater*. United States Environmental Protection Agency, Washington, D.C. Available from: https://yosemite.epa.gov/water/owrccatalog.nsf/ 065ca07e299b464685256ce50075c11a/27d2b544e85d1d8285256d83004fdb0c!OpenDocument (accessed 29 July 2021).

Vymazal, J. 2005 Horizontal sub-surface flow and hybrid constructed wetlands systems for wastewater treatment. *Ecological Engineering* **25** (5), 478–490.

Vymazal, J., Švehla, J., Kröpfeloya, L., Němcová, J. & Suchý, V. 2010 Heavy metals in sediments from constructed wetlands treating municipal wastewater. *Biogeochemistry* **101** (1/3), 335–356.

Walaszek, M., Bois, P., Laurent, J., Lenormand, E. & Wanko, A. 2018 Urban stormwater treatment by a constructed wetland: seasonality impacts on hydraulic efficiency, physico-chemical behavior and heavy metal occurrence. *Science of The Total Environment* **637–638**, 443–454.

Wang, C., Zheng, S.-s., Wang, P.-f. & Qian, J. 2014 'Effects of vegetations on the removal of contaminants in aquatic environments: a review'. *Journal of Hydrodynamics, Series B* **26** (4), 497–511.

Winston, R. J., Arend, K., Dorsey, J. D. & Hunt, W. F. 2020 Water quality performance of a permeable pavement and stormwater harvesting treatment train stormwater control measure. *Blue-Green Systems* **2** (1), 91–111.

Yadav, A. K., Kumar, N., Sreekrishnan, T. R., Satya, S. & Bishnoi, N. R. 2010 Removal of chromium and nickel from aqueous solution in constructed wetland: mass balance, adsorption–desorption and FTIR study. *Chemical Engineering Journal* **160** (1), 122–128.

Yadav, A. K., Abbassi, R., Kumar, N., Satya, S., Sreekrishnan, T. R. & Mishra, B. K. 2012 The removal of heavy metals in wetland microcosms: effects of bed depth, plant species, and metal mobility. *Chemical Engineering Journal* **211–212**, 501–507.

Zhang, K., Liu, Y., Deletic, A., McCarthy, D. T., Hatt, B. E., Payne, E. G. I., Chandrasena, G., Li, Y., Pham, T., Jamali, B., Daly, E., Fletcher, T. D. & Lintern, A. 2021 The impact of stormwater biofilter design and operational variables on nutrient removal – a statistical modelling approach. *Water Research* **188**, 116486.