On the fit of statistical and the $k$-$C^*$ models to projecting treatment performance in a constructed wetland system

A.O. BABATUNDE, Y.Q. ZHAO*, R.J. DOYLE, S.M. RACKARD, J.L.G. KUMAR and Y.S. HU

Centre for Water Resources Research, School of Architecture, Landscape and Civil Engineering, University College Dublin, Belfield, Dublin 4, Ireland

ABSTRACT

The objective of this study was to assess the suitability of statistical and the $k$-$C^*$ models to projecting treatment performance of constructed wetlands by applying the models to predict the final effluent concentrations of a pilot field-scale constructed wetlands system (CWs) treating animal farm wastewater. The CWs achieved removal rates (in g/m$^2$.d) ranging from 7.1-149.8 for BOD$_5$, 49.8-253.8 for COD and 7.1-47.0 for NH$_4$-N. Generally, it was found that the statistical models developed from multiple regression analyses (MRA) were stronger in predicting final effluent concentrations than the $k$-$C^*$ model. However, both models were inadequate in predicting the final effluent concentrations of NO$_3$-N. The first-order area-based removal rate constants ($k$, m/yr) determined from the experimental data were 200.5 for BOD$_5$, 80.1 for TP and 173.8 for NH$_4$-N and these indicate a high rate of pollutant removal within the CWs.

*Address correspondence to Yaqian Zhao, Centre for Water Resources Research, School of Architecture, Landscape and Civil Engineering, University College Dublin, Belfield, Dublin 4, Ireland; Tel. +353-1-7163215; Fax. +353-1-7163297; e-mail: yaqian.zhao@ucd.ie
Keywords Alum sludge; constructed wetlands; first-order, $k$-C* model, regression analysis

INTRODUCTION

Constructed wetlands systems (CWs) are treatment modules systems designed to mimic natural processes involving wetland vegetation, soils, and their associated microbial assemblages in treating wastewater. In recent times, the need to enhance or achieve specific treatment objectives which are often more stringent has led to the emergence of engineered wetland systems. These systems are essentially CWs, but with a more controlled environment.\[1\] Constructed and engineered wetland systems have gained widespread popularity as a wastewater treatment technology due to the fact that they are low-cost, easy-to-operate, efficient and robust.\[2\] They have been utilized globally for the treatment of various types of wastewaters including municipal wastewater from small communities, industrial and high-strength wastewater from agricultural activities. Contaminants in these systems are removed through a combination of physical, chemical, and biological processes including sedimentation, precipitation, adsorption to soil particles or specialized substrates, assimilation by plant tissue and microbial transformations and interactions.

The basis of design for these systems is usually first-order reaction model integrated with a plug-flow (PF) assumption. Although the flow in CWs is generally intermediate between PF and completely mixed, it is acknowledged that the first-order model with PF assumptions provides a conservative design estimate.\[3\] However, while this model is simple and widely used, it fails to adequately characterize the complex processes that occur in CWs.\[4\] For instance, the findings of numerous tracer studies in literature indicate that CWs typically do not possess ideal PF hydraulics. Furthermore, it is noted that current design procedures fail to incorporate atmospheric interactions such as precipitation, evaporation, and transpiration and these variables produce a secondary hydraulic regime that may influence retention times and invalidate steady state theoretical models.\[5, 6, 7\]
Several researchers have proposed more sophisticated CWs performance models, which simulate non-ideal hydraulics, either using a tanks in series (TIS) or plug flow with dispersion (PFD) approach. These models, however, have not been adopted by practitioners due to their complexity and amount of data required to properly use them. In addition, the proposed models still do not incorporate unsteady external hydraulic loading, which can have large effects on CWs performance. Recently, Rousseau et al. reviewed current CWs design approaches and concluded that the first-order PF model remains the best available method, despite its obvious deficiencies. This agrees with the submissions of Stein et al. who noted that first-order models are the most widely used for CWs design. Furthermore, Son et al. remarked that data availability and applicability have meant that simple regression and kinetic models have continued to be used to demonstrate CWs performance. In recent times, a modified first-order kinetic model often called the $k$-C* model was proposed by Kadlec and Knight and the model has been widely applied in CWs design. The $k$-C* model allows for a background or residual concentration ($C^*$), a non-zero lower limit that is approached asymptotically in lieu of the first-order model approach to zero.

On the other hand, statistical models including those derived from multiple regression analyses (MRA) have also been found useful for simplified description and analysis of CWs performance, and their accuracy in predicting such performances has been demonstrated. They are widely used in assessing CWs performance as they provide a means of understanding their treatment process/mechanism although this is limited. In this study, we evaluated the fit of statistical models developed from multiple regression analyses and the $k$-C* models in projecting the treatment performance of a pilot field-scale CWs treating animal farm wastewater by comparing the predicted final effluent concentrations using these models with those observed from the field trial. The CWs reported here, which has been developed in our research group, is unique in that it employs a by-product, alum sludge (the most widely
generated aluminium-coagulated drinking water treatment residual worldwide, as the main substrate primarily to enhance phosphorus (P) removal and also for biofilm attachment.\textsuperscript{14,15,16} Although MRA and $k$-C* models have their limitations, the availability and applicability of data from this unique pilot field-scale CWs were tested and discussed.

**MATERIALS AND METHODS**

**Pilot field-scale constructed wetland system**

The pilot field-scale CWs is located on an animal farm in Newcastle, Co Dublin, Ireland. The set-up and operation of the CWs has been described in Zhao et al.\textsuperscript{16} In brief, the CWs is a four-stage equal sized cell system (each 108cm×94cm×105cm) operated in vertical down-flow mode with a hydraulic loading rate of 0.29 m\textasciicircum{3}/m\textasciicircum{2}.d. The four cells are all interlinked by submersible pumps. The pumps are connected to a digital electronic timer which regulates the flow in the CWs according to a programme schedule. Dewatered alum sludge cake was collected fresh from the industrial filter press of the sludge dewatering unit of a drinking water treatment plant in South-West Dublin, Ireland where aluminium sulphate is used as coagulant. The characteristics of the dewatered alum sludge cake and its suitability for use as a substrate in a CWs has been well studied and reported.\textsuperscript{14} The dewatered alum sludge cake was then used as the main substrate in the CWs. The configuration of the CWs is as follows: 10cm of 10mm gravel at the bottom to serve as base support, 65cm of the dewatered alum sludge cakes to serve as the main substrate layer and 10cm of 20mm gravel to serve as the distribution layer. Each cell of the CWs was planted with common reeds, *phragmites australis*. The CWs was fed with wastewater collected from the secondary holding tank on the farm. However, in order to allow the system and the reeds time to establish, the wastewater from the farm was sometimes diluted with tap water to lower the concentration.

The four-stage CWs was operated in series. Influent and final effluent samples were collected from the CWs weekly and analysed for COD (both total and soluble COD,
(sCOD)), BOD₅ (Lovibond OxiDirect apparatus, Lennox, UK), TP (Ascorbic method, Clesceri et al.¹⁷), PO₄-P, TN (Persulfate method, Clesceri et al.¹⁷), NH₄-N, NO₃-N, NO₂-N, SS and Turbidity (Hach turbidity meter 2100N IS). Except where indicated, all the water quality parameters were analysed using a Hach DR/2400 spectrophotometer according to its standard operating procedures. Furthermore, real time measurements of oxidation-reduction potential (ORP), temperature and pH were obtained from a YSI multi-parameter probe inserted in the influent tank and in each stage of the CWs.

**Correlation and regression analysis**

Prior to the multiple regression analysis, correlation analysis was conducted between the water quality parameters to investigate any relationships present. This was done to give a greater understanding of suitable parameters to be used in the MRA. MRA was then performed to test the relationship between each of BOD₅, COD, TP, NH₄-N and NO₃-N and other parameters. These parameters were selected based on their relative importance in assessing water quality and the possibility of being able to predict their concentration in the final effluent from data obtained in real time, or from other parameters that are cheaper, easier and/or quicker to analyse. The MRA analysis was performed for each of the dependant variables using a combination of one to five predictor variables. The aim was to determine the optimum model for each combination of input variables by estimating BOD₅, COD, TP, NH₄-N and NO₃-N with as few inputs variables as possible in order to reduce associated errors and costs. A 95% confidence interval was specified for the regressions. The models were tested by examining the adjusted $R^2$, the significant F value, the p values and the mean absolute scaled error (MASE), as shown in Eq. (1) & (2).

$$q_i = \frac{\sum_{i=1}^{n} e_i}{n - 1}$$

$$Y_i = \sum_{i=2}^{e} [Y_i - Y_{i-1}]$$

(1)

(2)
where \( q_t \) is the scaled error of sample point \( t \), \( e_t \) is the residual of sample point \( t \), \( n \) is the sample size and \( Y_i \) is the observed value at step \( i \).

\[
MASE = \text{mean}(|q_t|)
\]

(2)

**Fitting \( k\)-C* model to experimental data**

Further to the MRA, the \( k\)-C* model, as shown in Eq. (3) \(^{[10]}\), was fitted to the experimental data. First-order area-based removal rate constants \((k)\), assuming removal to non-zero background concentrations \((C^*)\), were estimated for BOD\(_5\), COD, TP, NH\(_4\)-N, NO\(_3\)-N, TN and SS.

\[
-k_y q y = \ln \frac{C_o - C^*}{C_i - C^*}
\]

(3)

where \( k \) is the first order area-based removal rate constant \((\text{m/yr})\), \( C_o \) is the outlet concentration \((\text{mg/l})\), \( C_i \) is the inlet concentration \((\text{mg/l})\), \( C^* \) is the irreducible background concentration \((\text{mg/l})\), \( q \) is the hydraulic loading rate \((\text{m}^3/\text{m}^2.\text{day})\), and \( y \) is the fractional distance from inlet to outlet. The equation was rearranged for the fitting procedure, Eq. (4).

\[
C_o = C^* + (C_i - C^*) \exp\left(\frac{-k_y}{q}\right)
\]

(4)

The values of the irreducible background concentrations, \( C^* \), were adopted from the lowest effluent concentration observed for each pollutant in the CWs similar to the approach used by Trang et al.\(^{[18]}\) It is noted that very little information exists on suitable values for \( C^* \) and the values adopted in this study (shown later on in Table 5) are intermediate. In the literature, a range of \(1.7 \leq C^* \leq 18.2\) mg/l BOD\(_5\) with a mean of 9.9 was reported by Stein et al.\(^{[9]}\) and Kadlec and Knight\(^{[10]}\), respectively, while Shepherd et al.\(^{[19]}\) reported a range of \(23 \leq C^* \leq 450\) mg/l COD in a CWs treating winery wastewater.
RESULTS

Performance evaluation

The composition of the wastewater varied over time and the range of pollutant concentration in the influent wastewater to the CWs were BOD₅ (31-968 mg/l), COD (124-1634 mg/l), PO₄-P (2.8-60 mg-P/l), TN (16-273 mg-N/l) and SS (25-633 mg/l). It is noted that the characteristics of the source wastewater varied greatly over time in concert with seasonal changes and farming operations. The mean monthly removal efficiencies achieved in the CWs was determined to range from 56.6%-83.5%, 35.6%-84.2%, 11.2%-77.5%, 48.5%-92.5%, 75.4%-93.8%, 73.0%-96.5% and 46.3%-83.3% for BOD₅, COD, TN, NH₄-N, TP, P and SS respectively. In most cases, the removal efficiencies were generally lowest at the start-up of the CWs. However, removal efficiencies for TP and PO₄-P were high from the beginning in this study. Table 1 shows the loading and removal rates obtained from the CWs during the monitoring period. It is particularly noted that the CWs was able to efficiently remove PO₄-P both at low loading rates of 2.8 g-P/m².d when 2.6 g-P/m².d (93%) was removed, and high loading rates of 9.5 g-P/m².d when 9.2 g-P/m².d (97%) was removed. The system also performed well at removing BOD₅ both at low loading rates of 12 g-BOD₅/m².d when 7.5 g-BOD₅/m².d (95.8 g-BOD₅/m².d was removed and at high loading rates of 201 g-BOD₅/m².d when 149.8 g-BOD₅/m².d was removed. However, the removal of COD seemed to be more suited to higher loading rates. This may suggest that at lower COD loadings, internal release within the CWs has a more profound effect on COD removal. For instance, at a loading rate of 376.3 g-COD/m².d, 253.8 g-COD/m².d (68%) was removed compared to 49.8 g-COD/m².d (49%) removed at a lower loading rate of 102 g-COD/m².d. A possible reason could be that the loading rate is not high enough to treat. A similar observation was made in the case of a CWs treating piggery wastewater effluent where it was found that the concentration of residual-P increased along the path of the CWs. However, the exact causes of this observed phenomenon are still unclear. The same trend was observed for the removal of sCOD where...
at higher loading rates of 304.1 g-sCOD/m².d, 234.8 g-sCOD/m².d was removed compared to
14.1 g-sCOD/m².d removed at a lower loading rates of 57 g-sCOD/m².d. The removal of TN
fluctuated greatly but it appeared to perform better as the CWs became more established. The
system was able to efficiently remove NH₄-N both at low loading rates of 11 g-NH₄-N/m².d when 7.1 g-NH₄-N/m².d was removed, and high loading rates of 51.1 g-NH₄-N/m².d when 47
g-NH₄-N/m².d was removed.

[INSERT TABLE 1 HERE]

Correlation analysis

The results of the correlation analysis are presented in Table 2. It shows that the effluent
concentrations of BOD₅ are strongly related to the influent concentrations of BOD₅ with a
correlation coefficient, R, of 0.794. This suggests that although the CWs was efficient in
BOD₅ removal, the influent BOD₅ loading impacts on the final effluent BOD₅ concentration
that is obtained. The influent concentrations of COD and sCOD also correlate well with the
effluent concentrations of BOD₅ with R values of 0.715 and 0.634, respectively. The effluent
concentrations of COD are closely related to both the influent sCOD and SS with R values of
0.636 and 0.667, respectively. This also shows that both the influent concentrations of sCOD
and SS impacts on the final effluent concentration of COD that is obtained. The concentration
of TP in the effluent correlates very well with other water quality parameters such as influent
concentrations of BOD₅, COD and sCOD which produced R values of 0.795, 0.802 and
0.791, respectively. The effluent concentrations of NH₄-N did not correlate well with other
water quality parameters, but did exhibit negative correlation with both influent temperature
and pH with R values of -0.624 and -0.561, respectively. It is known that both pH and
temperature affect NH₄-N removal processes, and this is also the case in CWs.[21]

The correlation between influent BOD₅ and effluent PO₄-P and effluent BOD₅ and
influent PO₄-P is 0.420 and 0.594 respectively, while the correlation between influent COD
and effluent TP and effluent COD and influent TP is 0.802 and 0.511 respectively. Except for
the correlation between influent COD and effluent TP, all the correlations are of medium
strength. Although in the system, different processes account for the removal of organics
(BOD$_5$ and COD - through microbial degradation) and P and TP (mainly through adsorption
but not exclusively), the significance of the correlations found is that it can indirectly indicate
the loading on the system. For instance, the strength of the wastewater used was characterized
from its BOD$_5$ values based on which appropriate dilution was carried out to generate the
influent into the system. Consequently, a medium correlation between influent BOD$_5$ and
effluent PO$_4$-P would suggest that at high influent BOD$_5$ values (which would imply
comparatively higher organic loading rate on the system since the hydraulic loading rate is
fairly constant), a high value of effluent PO$_4$-P value can be expected as the system is
considerably loaded. Similar explanation can be given for the correlation between COD and
TP.

Nitrification, which is usually defined as the biological oxidation of ammonium-nitrogen
to nitrate-nitrogen with nitrite-nitrogen as an intermediate in the reaction sequence, is
favoured at high values of both temperature and pH and this could explain the negative
relationship observed. Effluent concentrations of NO$_3$-N displayed strong correlations with
TN and NH$_4$-N concentrations in the influent with R values of 0.621 and 0.632, respectively.
The effluent PO$_4$-P concentrations did not show any strong associations with any other
parameters. In the CWs understudy, P is primarily removed by adsorption onto the alum
sludge substrate through the ligand exchange mechanism and this has been shown in our
previous study.$^{[22]}$

\[\text{[INSERT TABLE 2 HERE]}\]
MRA for predicting final effluent concentration of selected pollutants

The results of the optimization of input variables for predicting final effluent concentrations for the selected pollutants are presented in Table 3 while regression models adjudged to be the best in predicting the final effluent concentrations for the selected pollutants are presented in Table 4. The results are presented below as they relate to each specific pollutant.

**BOD$_5$**

From the regression analysis (see Table 3), it was found that the combination of predictor variables that returned the best adjusted $R^2$ of 0.86 and the lowest MASE of 0.783 was that of influent BOD$_5$, NH$_4$-N, PO$_4$-P and COD. From the significant F value of $3.1 \times 10^{-8}$ it can be seen that the null hypothesis can be rejected as the value is less than the 0.05 confidence interval set during regression. All the p values for the predictor variables are also below the threshold value of 0.05. The model is shown in Table 4. Except for the model constructed using an input combination of three independent variables (i.e. influent concentration of BOD$_5$, NH$_4$-N and PO$_4$-P) which gave an adjusted $R^2$ value of 0.837 and a MASE value of 0.994, all the other models had MASE values above 1 even though their adjusted $R^2$ values ranged from 0.760-0.857. It is worth noting that the model constructed for predicting final effluent BOD$_5$ effluent concentrations using just one predictor variable, which was the BOD$_5$ of the influent, gave the least adjusted $R^2$ of 0.76 and the worst MASE value of 1.628. This suggests that although there may be reasonable correlation between the influent and effluent BOD$_5$ concentrations, the latter cannot be solely used to predict the former and they may be sources of organic matter release within the CWs.

**COD**
The most efficient model developed for predicting final effluent concentrations of COD consists of four input variables (see Table 3). The model gave an adjusted $R^2$ value of 0.711. This model also has a significant statistical relationship as the F test provides a significant F value of $1.5 \times 10^{-08}$ which rejects the null hypothesis. The p values for the model shown in Table 3 can be seen to be satisfactory. However, although the MASE value for the model is 1.825 which is greater than 1, it was the lowest compared to the MASE values obtained for the other models. The model is presented in Table 4.

**TP**

The most efficient model for predicting influent TP concentrations consisted of two input variables as shown in Table 3. Of all the models constructed for predicting the final effluent concentration of TP, it gave the highest adjusted $R^2$ value of 0.801 and the lowest MASE value of 1.620. The model is presented in Table 4.

**NH$_4$-N**

The model developed using a combination of three input variables was adjudged to be the best in predicting final effluent concentrations of NH$_4$-N (see Table 3). The model gave an adjusted $R^2$ value of 0.696 and had the lowest MASE of 1.111 compared to the other models. The significant F value for the model was found to be $5.4 \times 10^{-6}$, which is lower than the threshold value of 0.05. Both the predictor variables of influent temperature and SS recorded p values that were low enough to be considered significant. The model is presented in Table 4.

**NO$_3$-N**
The models developed for the prediction of final effluent concentrations of NO$_3$-N were not as strong as models developed using other parameters as dependant variables such as BOD$_5$ as can be seen from the adjusted $R^2$ values in Table 3. Although the models gave adjusted $R^2$ values ranging from 0.613-0.669, the MASE values were quite high. The best adjudged model based on the adjusted $R^2$ and MASE value is presented in Table 4.

$k$–$C^*$ model

Using Eq. 3, the $k$ values (in m/yr) for each pollutant were calculated to be 200.5, 271.8, 80.1, 173.8, 618.5, 271.4 and 223.2 for BOD$_5$, COD, TP, NH$_4$-N, NO$_3$-N, TN and SS respectively (see Table 5). The data was then fitted to Eq. 4, which yielded estimates with $R^2$ values of 0.574, 0.350, 0.368, 0.005, 1.6x10$^{-4}$, 0.390 and 0.682 for BOD$_5$, COD, TP, NH$_4$-N, NO$_3$-N, TN and SS respectively. The MASE values for these models can also be seen in Table 5. The results show that the $k$–$C^*$ model fitted well for predicting final effluent concentration of BOD$_5$, COD, TN TP and SS but performed poorly for NH$_4$-N and NO$_3$-N.

Residual analysis

Fig. 1(a & b) shows the residual analysis from predicting final effluent BOD$_5$ concentrations using the MRA and the $k$–$C^*$ model, respectively. For Fig. 1(a), it can be seen that the data points are randomly dispersed about zero for all concentrations, therefore, no pattern is discernable and so the model is reliable. However, in the case of Fig. 1(b), there is a definite positive skew observed and so the model assumptions are not correct. It can therefore be said that the model for predicting the final effluent BOD$_5$ concentrations using MRA is more reliable than that of the $k$–$C^*$ model. Fig. 1(c & d) shows the residual analysis for estimating final effluent COD concentrations using the MRA and $k$–$C^*$ model, respectively.
From Fig. 1(c), it can be seen that the MRA model has uniform dispersion about the line $R=0$. This indicates that the model is acceptable.

However, similar to the performance of the $k-C^*$ model in predicting final effluent BOD$_5$ concentrations, the performance of $k-C^*$ model in predicting final effluent COD concentrations displayed a definite positive skew as can be seen in Fig. 1(d). This arises when the model assumptions are not met, indicating that the model is not as suitable for predicting the final effluent concentrations of COD as the MRA model. The residual analysis for the prediction of final effluent TP concentrations using the MRA and the $k-C^*$ model is presented, respectively, in Fig. 1(e & f). Fig. 1(e) shows that the MRA model again displays a random scattering of points above and below the line $R=0$ signifying that the model is adequate for predicting final effluent TP concentrations. However, the residual plot for $k-C^*$ model (Fig. 1(f)) displays a negative skew, again indicating that it is not as appropriate as the MRA model. Fig. 1(g) shows the residual analysis for the prediction of final effluent NH$_4$-N concentrations using MRA. This plot shows a random scattering of residuals about the line $R=0$ and so indicates that model is adequate in predicting the final effluent NH$_4$-N concentrations. However, for the kinetic model (Fig. 1(h)), it shows a definite positive to negative skew.

This pattern indicates that the model is not as accurate in predicting final effluent NH$_4$-N concentrations as the MRA model. In the case of NO$_3$-N, the residual plots for the MRA and $k-C^*$ models as shown in Fig. 1 (i & j) respectively, indicate that both models cannot be considered accurate enough to predict final effluent concentrations of NO$_3$-N. However, the reason why both models failed in predicting NO$_3$-N is still unclear. The MRA model shows a definite pattern in that as the predicted value increases, the residuals increase in magnitude in both the positive and negative directions, creating the cone shape. On the other hand, the $k-$
C* model displays an obvious positive skew which is unacceptable as it implies that the model assumptions are false.

**DISCUSSION**

Generally, the performance of the CWs in removing key pollutants in the wastewater can be considered as excellent. In all cases, the treatment efficiency obtained was comparable to, or higher than, the performances obtained in conventional and other CWs. The system under trial was able to efficiently remove PO$_4$-P at both high and low loading rates. The removals of SS, BOD$_5$ and COD were also generally efficient with a trend of gradual improvement in removal rates as the system matures. This would indicate that the system was emerging from the start-up stage and that the biomass was maturing. However, the reason that the system is so efficient in P removal is due to the abundance of aluminium ions in the dewatered alum sludge which enhances P adsorption from the wastewater through the ligand exchange mechanism.\[22\] In relative terms, these removal efficiencies can be considered as excellent and showing good promise. Vymazal\[21\] reported that average removal efficiencies in vertical subsurface flow CWs were 44.6% for TN and 59.5% for TP. Therefore, it can be seen that the respective performance efficiencies obtained in this study are considerably higher than the reported averages and this indicates the potential of the CWs described in this study for providing enhanced wastewater treatment by CWs. Furthermore, an interesting feature of the performance worth noting is the ability of the CWs to achieve, concurrently, high-rate removal of organics (BOD$_5$, COD) and nutrients (emphasis on P). It is well known that it is often a challenge to achieve concurrent high removal efficiencies for P and organic matter in CWs. CWs are usually efficient in organics (BOD$_5$, COD) reduction, but the corresponding removal efficiencies for nitrogen (N) and P are often low.\[2\]. The four-stage pilot field-scale
CWs described in this study has demonstrated that by using alum sludge as substrate, high P removal efficiency can be achieved alongside high removal efficiency for organics.

The statistical models developed for predicting final effluent concentrations of selected key water quality parameters using MRA were found to be promising. The MRA model for predicting final effluent BOD\textsubscript{5} concentrations was exceptionally good and this is reiterated by the p values of each predictor variable and their asterisks as can be seen in Table 4 coupled with the fact that the MASE < 1. This indicates that the model is quite good for predicting the final effluent BOD\textsubscript{5} concentrations and also, the predictions are reliable. However, in the case of COD, although the best MRA model for predicting final effluent COD concentrations gave an adjusted R\textsuperscript{2} of 0.711, the MASE >1 and this suggests that while the model is strong at predicting final effluent COD concentrations, it may contain some errors. This is also the case for the MRA model for predicting final effluent TP concentration which gave an adjusted R\textsuperscript{2} of 0.801, indicating that over 80% of all possible outcomes are covered by the model. However, the MASE >1 also indicates that the prediction may contain some errors. Similar performance was also observed in the MRA model for predicting final effluent concentration of NH\textsubscript{4}-N in the CWs. An adjusted R\textsuperscript{2} of 0.696 was obtained but the MASE >1 (1.111) which indicates that the model is strong at predicting the final effluent concentration, but it may contain some errors. In the case of NO\textsubscript{3}-N, the model seems unreliable with a MASE of 5.419.

The values of the first order area-based removal rate constant, k (in m/yr), obtained in the study are high and they compare very well with the range of values found in the scientific literature. In vertical flow CWs treating high-strength synthetic wastewater in Thailand, Kantawanichkul et al.\cite{23} found k values of 49.8, 30.1 and 13.5 m/yr for COD, TKN and TP, respectively. Other k values (in m/yr) reported in literature range from 8–95 (BOD\textsubscript{5}), 22–30 (COD), 3.6–24 (TN and TKN) and 4.9–84 (TP).\cite{23, 24, 25, 26} However, it is the opinion of the authors that these values have limited use for comparison and caution is urged in drawing firm conclusions from them. The estimated removal rate constants and apparent background
concentrations have been shown to depend strongly on input water quality as well as hydraulic loading rates (HLR)[4,27] For instance, the $k$ values reported by Kantawanichkul et al.[23] were obtained at an HLR of 80 mm/d which is two orders of magnitude lower than the HLR used in this study. Also, Stein et al.[9] stated that $C^*$ is relatively more important than $k$ when using the $k$–$C^*$ model and it was suggested that the rate parameter $k$ may be less important than the residual concentration, particularly when predicting effluent organic matter values. However, as previously noted, very little information exists in literature regarding suitable values for $C^*$ particularly for different types of wastewater. Nonetheless, the $k$ values obtained in this study imply the high rate of pollutant removal in the CWs. However, the $k$ value obtained for nitrate-nitrogen is quite high and cannot be relied upon. Similar observation has been made by Kadlec and Wallace.[1]

From Table 5, it can be seen that the $k$–$C^*$ model was best fitted to $BOD_5$ and SS removal in the CWs with $R^2$ values of 0.574 and 0.682 and MASE values of 0.644 and 0.680 respectively, indicating that the values obtained are quite reliable. The MASE values for COD and TN were also <1. Overall, the MRA models were found to be more apt in predicting the final effluent concentrations of pollutants in CWs than the $k$–$C^*$ model. It should however be noted that the CWs in this study applied alum sludge as substrate, which could give different treatment characteristics when compared to other CWs particularly in terms of P removal. Therefore caution is urged in generalising the results.

CONCLUSIONS

1. Statistical models developed from multiple regression analyses were found to be strong in predicting final effluent concentrations for the constructed wetland system in this study. In particular, the selected model for predicting final effluent concentration
of BOD$_5$ gave an adjusted R$^2$ value of 0.860 between predicted and observed data and a mean absolute scaled error <1 (0.783).

2. The $k$-C* model was also fitted to the experimental data and the first order area-based removal rate constant, $k$ (m/yr), for BOD$_5$, COD, TP, TN and SS was calculated to be 200.5, 271.8, 80.1, 271.4 and 223.2 respectively.

3. Comparison of the regression and $k$-C* models shows that the regression models were much stronger in predicting final effluent concentrations in the constructed wetland system.

Acknowledgements

Authors gratefully acknowledge financial support received for this field study from: (1) Enterprise Ireland under the Proof of Concept Scheme (project No. PC/2007/0308); (2) the Irish state Department of Agriculture, Fisheries and Food under the Research Stimulus Fund (project No. RSF 07-528). The UCD farms, Lyons Estate, Newcastle, Co. Dublin, Dr Edward Jordan and Mr Michael Hegarty are all sincerely thanked for their support for the field work.

REFERENCES

[1] Kadlec, R.H.; Wallace, S. Treatment Wetlands, 2nd edition. CRC Press, Raton, FL, 2009

[2] Park, W.H. Integrated constructed wetland systems employing alum sludge and oyster shells as filter media for P removal. Ecol. Eng. 2009, 35, 1275-1282.

[3] Knight, R.L.; PayneJr. V.W.E.; Borer, R.E.; ClarkeJr, R.A.; Pries, J.H. Constructed wetlands for livestock wastewater management. Ecol. Eng. 2000, 15, 41-55

[4] Kadlec, R.H. The inadequacy of first-order treatment wetland models. Ecol. Eng. 2000, 15, 105–119.

[5] Kincanon, R.; McAnally, A.S. Enhancing commonly used model predictions for constructed wetland performance: as-built design considerations. Ecol. Model. 2004, 174, 309-322
[6] Jamieson, R.; Gordon, R.; Wheeler, N.; Smith, E.; Stratton, G.; Madani, A. Determination of first order rate constants for wetlands treating livestock wastewater in cold climates. Environ. Eng. Sci. 2007, 6, 65-72

[7] Son, Y.K.; Yoon, C.G.; Kim, H.C.; Jang, J.H.; Lee, S.B. Determination of regression model parameter for constructed wetland using operating data. Paddy Water Environ. 2010, DOI 10.1007/s10333-010-0211-9

[8] Rousseau, D.; Vanrolleghem, P.; De Pauw, N. Model-based design of horizontal subsurface flow constructed treatment wetlands: a review. Water Res. 2004, 38, 1484-1493

[9] Stein, O.R.; Biederman, J.A.; Hook, P.B.; Allen, W.C. Plant species and temperature effects on the $k$-$C^*$ first order model for COD removal in batch-loaded SSF wetlands. Ecol. Eng. 2006, 26, 100-112

[10] Kadlec, R.H.; Knight, R.L. Treatment Wetlands. Lewis Publishers, CRC Press Inc., Boca Raton, FL, 1996

[11] Tomenko, V.; Sirajuddin, A.; Viktor, P. Modelling constructed wetland treatment system performance. Ecol. Model. 2007, 205, 355-364.

[12] Murray-Gulde, C.L.; Bridges, W.C.; Rodgers, J.H. Evaluating performance of a constructed wetland treatment system designed to decrease bioavailable copper in a waste stream. Environ. Geosci. 2008, 15, 21-38

[13] Tang, X.; Eke, P.E.; Scholz, M.; Huang, S. Processes impacting on benzene removal in vertical-flow constructed wetlands. Bioresour. Technol. 2009, 100, 227-234

[14] Babatunde, A.O.; Zhao, Y.Q.; Burke, A.M.; Morris, M.A.; Hanrahan, J.P. Characterization of aluminium-based water treatment residual for potential phosphorus removal in engineered wetlands. Environ. Pollut. 2009, 157, 2830-2836
[15] Babatunde, A.O.; Zhao, Y.Q.; Zhao, X-H. Alum sludge-based constructed wetland system for enhanced removal of P and OM from wastewater: Concept, design and performance analysis. Bioresour. Technol. 2010, 101, 6576-6579

[16] Zhao, Y.Q.; Babatunde, A.O.; Hu, Y.S.; Kumar, J.L.G.; Zhao, X-H. A two-prong approach of beneficial reuse of alum sludge in engineered wetland: First experience from Ireland. Waste and Biomass Valorization, 2010, 1, 227-234

[17] Clesceri, L.S.; Greenberg, A.E.; Eaton, A.D. Standard Methods for the Examination of Water and Wastewater, 20th Ed., American Public Works Association, Washington, DC, 1998

[18] Trang, N.D.T.; Konnerup, D.; Schierup, H-H.; Chiem, N.H.; Tuan, L.A.; Brix, A. Kinetics of pollutant removal from domestic wastewater in a tropical horizontal subsurface flow constructed wetland system: Effects of hydraulic loading rate. Ecol. Eng. 2010, 36, 527-535

[19] Shepherd, H.L.; Tchobanoglous, G.; Grisner, M.E. Time-dependent retardation model for chemical oxygen demand removal in a subsurface-flow constructed wetland for winery wastewater treatment. Water Environ. Res. 2001, 73, 597-606

[20] Lee, S.; Maniquiz, M. C.; Kim, L-H. Characteristics of contaminants in water and sediment of a constructed wetland treating piggery wastewater effluent. J. Environ. Sci. 2010, 22, 940-945

[21] Vymazal, J. Removal of nutrients in various types of constructed wetlands. Sci. Total Environ. 2007, 380, 48-65

[22] Yang, Y.; Zhao, Y.Q.; Babatunde, A.Q.; Wang, L.; Ren, Y.X.; Han, Y. Characteristics and mechanisms of phosphate adsorption on dewatered alum sludge. Separ. Purif. Technol. 2006, 51, 193-200
[23] Kantawanichkul, S.; Kladprasert, S.; Brix, H. Treatment of high-strength wastewater in tropical vertical flow constructed wetlands planted with Typha angustifolia and Cyperus involucratus. Ecol. Eng. 2009, 35, 238–247.

[24] Noorvee, A.; Repp, K.; Põldvere, E., Mander, Ü. The effects of aeration and the application of the k–C* model in a subsurface flow constructed wetland. J. Environ. Sci. Heal. Pt. A. 2005, 40, 1445–1456.

[25] Öövel, M.; Tooming, A.; Mauring, T.; Mander, Ü. Schoolhouse wastewater purification in a LWA-filled hybrid constructed wetland in Estonia. Ecol. Eng. 2007, 29, 17-26

[26] Konnerup, D.; Koottatep, T.; Brix, H. Treatment of domestic wastewater in tropical, subsurface flow constructed wetlands planted with Canna and Heliconia. Ecol. Eng. 2009, 35, 248–257

[27] Kadlec, R.H. Effects of pollutant speciation in treatment wetlands design. Ecol. Eng. 2003, 20 1–16.
Figure 1. Residual analysis from estimation of final effluent concentrations for selected water quality parameters using statistical and the \( k-C^* \) models.
Table 1. Performance summary of the constructed wetland system from Feb.-Dec. 2009

| Parameter | g/m².d | Feb  | Mar  | Apr  | May  | Jun  | Jul  | Aug  | Sep  | Oct  | Nov  | Dec  |
|-----------|--------|------|------|------|------|------|------|------|------|------|------|------|
| BOD₅      | loading| nd   | 12.0 | 13.1 | 31.9 | 89.2 | 69.5 | 89.1 | 68.8 | 119.1| 201.4| 119.3|
|           | removal| nd   | 7.5  | 7.1  | 21.8 | 70.6 | 48.2 | 64.9 | 37.9 | 87.6 | 149.8| 89.0 |
| COD       | loading| nd   | 207.6| 140.1| 102.0| 141.1| 109.8| 189.3| 200.4| 285.0| 376.3| 202.3|
|           | removal| nd   | 53.7 | 58.0 | 49.8 | 93.0 | 67.9 | 151.1| 139.3| 168.2| 253.8| 147.3|
| sCOD      | loading| nd   | 166.8| 67.3 | 57.0 | 82.9 | 65.7 | 109.4| 119.9| 183.1| 304.1| 108.6|
|           | removal| nd   | 36.5 | 15.3 | 14.1 | 46.7 | 35.9 | 83.2 | 78.0 | 129.9| 234.8| 82.7 |
| SS        | loading| 92.1 | 72.3 | 41.7 | 33.6 | 29.3 | 29.8 | 42.1 | 49.6 | 66.7 | 100.5| 78.3 |
|           | removal| 55.2 | 37.4 | 23.7 | 24.3 | 22.9 | 21.4 | 33.9 | 40.0 | 44.0 | 58.9 | 38.7 |
| TP        | loading| nd   | 4.8  | 3.1  | nd   | nd   | nd   | nd   | nd   | 9.6  | 8.6  | 6.4  |
|           | removal| 3.8  | 4.2  | 3.0  | 2.5  | nd   | 2.6  | 4.7  | 6.8  | 8.9  | 7.2  | 4.9  |
| PO₄-P     | loading| 7.4  | 3.9  | 2.8  | 5.2  | 4.5  | 3.8  | 6.9  | 7.8  | 9.5  | 7.9  | 5.3  |
|           | removal| 6.1  | 3.5  | 2.6  | 4.8  | 4.3  | 3.5  | 6.5  | 7.5  | 9.2  | 6.9  | 3.8  |
| TN        | loading| 24.4 | nd   | 25.4 | nd   | 27.7 | 24.7 | 58.3 | 60.9 | 63.1 | 38.2 | 31.4 |
|           | removal| 2.9  | 0.9  | 11.6 | 4.4  | 17.4 | 15.1 | 38.3 | 22.8 | 24.5 | 27.9 | 17.0 |
| NH₄-N     | loading| 23.7 | 21.0 | 11.0 | 13.4 | 18.6 | 17.6 | 37.2 | 47.1 | 51.1 | 31.6 | 26.5 |
|           | removal| 16.5 | 7.5  | 7.1  | 11.7 | 16.6 | 15.4 | 33.6 | 43.5 | 47.0 | 24.8 | 12.9 |

nd not determined
| R value | Influent | Effluent |
|---------|----------|----------|
|         | BOD$_5$  | COD      | sCOD     | PO$_4$-P | TP      | NH$_4$-N | TN      | SS      | NO$_3$-N | Temp | pH     | ORP     |
| BOD$_5$ | 0.794    | 0.715    | 0.634    | 0.594    | 0.556   | 0.452    | 0.369   | 0.534   | 0.156    | -0.114 | -0.254 | -0.369  |
| COD     | 0.341    | 0.581    | 0.636    | 0.373    | 0.511   | 0.159    | 0.041   | 0.667   | -0.183   | -0.591 | -0.541 | 0.279   |
| sCOD    | -0.071   | 0.266    | 0.510    | 0.115    | 0.229   | -0.084   | -0.181  | 0.471   | -0.098   | -0.552 | -0.442 | 0.236   |
| PO$_4$-P| 0.420    | 0.365    | 0.422    | 0.296    | 0.114   | 0.040    | -0.045  | 0.268   | 0.104    | -0.250 | -0.379 | -0.039  |
| TP      | 0.795    | 0.802    | 0.791    | 0.492    | 0.556   | 0.277    | 0.182   | 0.650   | -0.186   | -0.399 | -0.469 | 0.031   |
| NH$_4$-N| 0.104    | 0.259    | 0.303    | 0.039    | 0.033   | 0.188    | -0.118  | 0.430   | -0.001   | -0.624 | -0.561 | 0.354   |
| TN      | -0.206   | 0.029    | -0.081   | 0.169    | 0.084   | 0.558    | 0.588   | -0.031  | 0.079    | -0.011 | -0.122 | 0.432   |
| SS      | 0.499    | 0.729    | 0.716    | 0.538    | 0.666   | 0.284    | 0.032   | 0.855   | -0.208   | -0.624 | -0.583 | 0.194   |
| NO$_3$-N| 0.012    | 0.183    | 0.023    | 0.214    | 0.141   | 0.632    | 0.621   | 0.020   | 0.010    | 0.076  | -0.046 | 0.102   |

Table 2. Correlation matrix of influent and effluent concentrations for the water quality variables.
Table 3. Optimal input variables combination for predicting final effluent concentrations

| Input Variables | Adj R² | Sig F | P Values | MASE |
|-----------------|--------|-------|----------|------|
| BODs            |        |       |          |      |
| 1               | 0.760  | 2.1x10⁻¹⁰ | 2.2x10⁻¹⁰ | 1.628 |
| 1+2             | 0.857  | 7.6x10⁻¹¹ | 8.9x10⁻¹⁰ | 1.09  |
| 1+2+3           | 0.837  | 1.9x10⁻⁰⁹ | 3.8x10⁻⁰⁸ | 0.994 |
| 1+2+3+4         | 0.860  | 3.1x10⁻⁰⁸ | 0.001    | 0.013 | 0.039 | 0.046 | 0.783 |
| 1+2+3+4+5       | 0.849  | 1.9x10⁻⁰⁹ | 0.006    | 0.021 | 0.053 | 0.149 | 0.993 | 1.352 |
| COD             |        |       |          |      |
| 6               | 0.574  | 7.5x10⁻¹¹ | 7.5x10⁻¹¹ | 5.043 |
| 5+8             | 0.707  | 2.4x10⁻⁰₈ | 0.002    | 0.001 | 2.172 |
| 1+7+8           | 0.679  | 6.6x10⁻⁰₈ | 2.9x10⁻⁰₆ | 8.9x10⁻⁰⁷ | 0.223 | 1.934 |
| 1+7+9+11        | 0.711  | 1.5x10⁻⁰₈ | 2.6x10⁻⁰⁵ | 3.7x10⁻⁰⁹ | 4.7x10⁻⁰⁴ | 0.168 | 1.825 |
| 1+6+7+9+11      | 0.702  | 1.3x10⁻⁰⁷ | 2.9x10⁻⁰⁴ | 0.210 | 3x10⁻⁰⁵ | 0.600 | 0.297 | 1.899 |
| TP              |        |       |          |      |
| 7               | 0.615  | 3.5x10⁻⁰⁹ | 3.5x10⁻⁰⁹ | 3.350 |
| 6+7             | 0.801  | 6.4x10⁻¹⁰ | 0.107    | 1.7x10⁻¹⁰ | 1.620 |
| 1+7+9           | 0.718  | 1.2x10⁻⁰⁷ | 0.069    | 0.052 | 0.102 | 2.545 |
| 1+4+6+7         | 0.668  | 2.2x10⁻⁰⁵ | 0.491    | 0.205 | 0.104 | 0.845 | 2.436 |
| 1+4+6+7+8       | 0.682  | 3.6x10⁻⁰⁵ | 0.917    | 0.147 | 0.096 | 0.338 | 0.181 | 2.310 |
| NH₄-N           |        |       |          |      |
| 6               | 0.604  | 3.9x10⁻⁰⁸ | 3.9x10⁻⁰⁸ | 3.038 |
| 6+7             | 0.649  | 1.5x10⁻⁰⁹ | 7.0x10⁻⁰⁹ | 0.136 | 3.308 |
| 6+7+8           | 0.696  | 5.4x10⁻⁰⁶ | 0.001    | 0.512 | 0.006 | 1.111 |
| 6+7+8+9         | 0.739  | 1.4x10⁻⁰⁵ | 0.057    | 0.202 | 0.005 | 0.805 | 1.331 |
| 6+7+8+9+4       | 0.682  | 4.9x10⁻⁰⁴ | 0.078    | 0.704 | 0.028 | 0.760 | 0.965 | 1.446 |
| NO₃-N           |        |       |          |      |
| 2               | 0.613  | 2.9x10⁻¹⁰ | 2.9x10⁻¹⁰ | 8.625 |
| 2+10            | 0.650  | 3.2x10⁻⁰⁸ | 6.4x10⁻⁰⁴ | 0.053 | 5.861 |
| 2+3+10          | 0.670  | 6.5x10⁻⁰⁷ | 0.002    | 0.222 | 0.045 | 5.419 |
| 2+3+6+10        | 0.652  | 1.7x10⁻⁰⁶ | 0.001    | 0.759 | 0.241 | 0.053 | 6.550 |
| 2+3+6+10+11     | 0.669  | 8.6x10⁻⁰⁵ | 0.011    | 0.593 | 0.221 | 0.312 | 0.372 | 8.950 |

1-BODₛ inf; 2-NH₄-N inf; 3-PO₄-P inf; 4-COD inf; 5-TP eff; 6-Temp inf; 7-sCOD inf; 8-SS inf; 9-pH inf; 10-TN inf; 11-TP inf. Underline indicates best mean absolute scaled error (MASE). inf = influent; eff = effluent
**Table 4.** Optimum models for predicting final effluent concentrations from multiple regression analysis

| Parameter | Model Equation                                                                 | Intercept | p-value |
|-----------|--------------------------------------------------------------------------------|-----------|---------|
| BOD₅      | 0.153(BOD₅ Inf)*** + 0.252(NH₄-N Inf)* - 1.872(PO₄-P Inf)* + 0.058(COD Inf)*    | + 2.940   |         |
| COD       | -0.326(BOD₅ Inf)*** + 0.529(sCOD Inf)*** - 119.198(pH Inf)*** + 0.842(TP Inf)    | + 1019.043|         |
| TP        | -0.078(TEMP Inf) + 0.005(sCOD Inf)***                                            | + 0.147   |         |
| NH₄-N     | -1.866(TEMP Inf)*** - 0.004(sCOD Inf) + 0.065(SS Inf)**                           | + 29.533  |         |
| NO₃-N     | 1.020(NH₄-N Inf)** + 0.980(PO₄-P Inf) - 0.513(TN Inf)*                           | - 7.256   |         |

*= p≤0.05, **= p≤0.01 and ***= p≤0.001
Table 5. First-order area-based removal rate constants ($k$), background concentration ($C^*$) and coefficients of determination ($R^2$) for BOD$_5$, COD, TP, NH$_4$-N, NO$_3$-N, TN and SS removal in the constructed wetland system based on the $k$–$C^*$ model

| Parameter | $C^*$ (mg/L) | $k$ (m/yr) | $R^2$ | MASE |
|-----------|--------------|------------|-------|------|
| BOD$_5$   | 15           | 200.504    | 0.574 | 0.644|
| COD       | 60           | 271.753    | 0.350 | 0.929|
| TP        | 0.3          | 80.134     | 0.368 | 6.216|
| NH$_4$-N  | 1.5          | 173.758    | 0.005 | 2.324|
| NO$_3$-N  | 0.5          | 618.504    | $1.6 \times 10^{-4}$ | 2.090|
| TN        | 10           | 271.440    | 0.390 | 0.693|
| SS        | 5            | 223.184    | 0.682 | 0.680|
Figure 1

(a) Predicted BOD$_5$ Concentrations (mg/L)
(b) Predicted BOD$_5$ Concentrations (mg/L)
(c) Predicted COD Concentrations (mg/L)
(d) Predicted COD Concentrations (mg/L)
(e) Predicted TP Concentrations (mg/L)
(f) Predicted TP Concentrations (mg/L)
(g) Predicted NH$_4$-N Concentrations (mg/L)
(h) Predicted NH$_4$-N Concentrations (mg/L)
(i) Predicted NO$_3$-N Concentrations (mg/L)
(j) Predicted NO$_3$-N Concentrations (mg/L)