

2013

Derivation of a Fuzzy National Phosphorus Export Model using 84 Irish Catchments

Ahmed Nasr

Technological University Dublin, ahmed.nasr@tudublin.ie

Michael Bruen

UCD, michael.bruen@ucd.ie

Follow this and additional works at: <https://arrow.tudublin.ie/engschcivart>



Part of the [Environmental Engineering Commons](#)

Recommended Citation

Nasr, A., and Bruen, M., 2013. "Derivation of a fuzzy national phosphorus export model using 84 Irish catchments". *Science of the Total Environment*, vol. 443, pp. 539-548. doi:10.1016/j.scitotenv.2012.10.063

This Article is brought to you for free and open access by the School of Civil and Structural Engineering (Former DIT) at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, vera.kilshaw@tudublin.ie.

DERIVATION OF A FUZZY NATIONAL PHOSPHORUS EXPORT MODEL USING 84 IRISH CATCHMENTS

Ahmed Nasr¹, Michael Bruen^{2,*}

1 School of Civil & Building Services Engineering, College Of Engineering & Built Environment, Dublin Institute of Technology, Bolton Street, Dublin 1, Ireland

2 Centre for Water Resources Research, University College Dublin, Newstead Building, Richview, Belfield, Dublin 4, Ireland

* Corresponding author: Tel. +35317163212, Fax: +35317163297, E-mail: michael.bruen@ucd.ie

Abstract

Implementation of appropriate management strategies to mitigate diffuse phosphorus (P) pollution at the catchment scale is vitally important for the sustainable development of water resources in Ireland. An important element in the process of implementing such strategies is the prediction of their impacts on P concentrations in a catchment using a reliable mathematical model. In this study, a state-of-the-art adaptive neuro-fuzzy inference system (ANFIS) has been used to develop a new national P model capable of estimating average annual ortho-P concentrations at un-gauged catchments. Data from 84 catchments dominated by diffuse P pollution were used in developing and testing the model. Six different split-sample scenarios were used to partition the total number of the catchments into two sets, one to calibrate and the other to validate the model. The k-means clustering algorithm was used to partition the sets into clusters of catchments with

similar features. Then for each scenario and for each cluster case, 11 different models, each of which consists of a linear regression sub-model for each cluster, were formulated by using different input variables selected from among six spatially distributed variables including phosphorus desorption index (PDI), runoff risk index (RRI), geology (GEO), groundwater (GW), land use (LU), and soil (SO). The success of the new approach over the conventional lumped, empirical, modelling approach was evident from the improved results obtained for most of the cases. In addition the results highlighted the importance of using information on PDI and RRI as explanatory input variables to simulate the average annual ortho-P concentrations.

Keywords: diffuse phosphorus pollution; ortho-Phosphate, empirical modelling approach; fuzzy sub-sets; neuro-fuzzy model; k-means fuzzy clustering

1. Introduction

Like many countries, Ireland has continually improved its agricultural output to meet local demand for food as well as for exports. Using the land in any intensive agricultural activity can cause adverse impacts on the environment unless appropriate measures to reduce these impacts are put in place such as water storage facilities (e.g. De Martino et al., 2012; De Paola and Ranucci, 2012). Most of the reviews of water quality in Ireland revealed that diffuse transport of phosphorus (P) by surface and sub-surface flows from agriculture soil to the receiving waters is one of the major environmental problems (e.g. Lehane and O'Leary, 2012; McGarrigle et al., 2002; Toner et al., 2005). Soluble P in a form of ortho-P is readily available for plants and always leads to eutrophication in Irish Rivers and as a result there is a need for a catchment based management strategy that encapsulates all elements contributing to the loss of this form of P (Hutton et al., 2008). The EU Water Framework Directive (WFD) (EEC, 2000) provides the legal grounds required to develop and enforce such a management strategy. It mandates a thorough investigation to predict the impacts which will be produced by each possible management alternative.

The Three Rivers project (MCOS, 2002) was one of the early and detailed studies conducted in Ireland with the aim of developing catchment based monitoring and management systems for the Boyne, Liffey and Suir catchments. A related project was the Lough Derg/Lough Ree Project (KMMP, 2001) which addressed the same objectives in the Three Rivers project. In addition to the valuable management plans developed by these two projects, an important database required for modelling diffuse P loads was generated. Daly and Mills (2006) utilised some of this data to develop an empirical model

described in Appendix I to estimate the annual average ortho-P concentrations from diffuse sources at the outlet of a catchment. Using a number of spatial variables derived from land use, soil type, stocking densities, fertiliser P use and soil P levels Daly and Mills (2006) developed a series of linear regression models relating different combinations of these variables with the average annual ortho-P concentration. From these models they chose the best empirical model as the one which provided the best calibration. In the analysis Daly and Mills (2006) used data from 84 different catchments in Ireland which have been selected on the basis of the following criteria: (i) nested catchments were avoided; and (ii) diffuse pollution represented the main contributor of phosphorus to the stream. Starting with all variables and using a backward-steps regression procedure they eliminated the variables that had no significant effect on the linear regression model. In their final model only two variables were deemed significant and retained in the linear regression equation. These two variables were phosphorus desorption index (PDI) and runoff risk Index (RRI). Due to its simplicity and parsimony this type of empirical export coefficient modelling approach to predicting diffuse contaminant loads has been used widely (e.g. Daly and Coulter, 2000; Davies and Neal, 2007; Johnes, 1996; Lek et al., 1999; Meynendonckx et al., 2006; Su et al., 2006). Usually models of this type do not incorporate in their structure any representation of the actual physical processes involved in the mobilisation and transport of P but instead they seek to establish a numerical link between the contaminant load and the catchment characteristics which influence it.

Most of the simple empirical annual average P export coefficient models are multiple linear regression models that linearly relate the predictors, which are in most cases land use types, with the predictand representing the water quality parameter under consideration such as P soluble reactive concentration (e.g. McGuckin et al., 1999). However, the diffuse P transport process is in fact highly non-linear and the driver variables are in general not limited to land use types but they rather extend to include other predictors (e.g. soil P sorption capacity, residual soil P amounts, density of livestock, climate). Therefore non-linear models are extremely needed in order to accurately express the non-linearity in the process. Furthermore the available observed P concentration data only covers a small number of years and this necessitates the use of Monte Carlo simulation techniques to generate synthetic long time series that can be used in uncertainty analysis (e.g. McFarland and Hauck, 2001). Thus it will be useful if the candidate model can intrinsically account for the uncertainty in the data using the available actual data through an embedded modelling mechanism without the need for an external procedure to generate synthetic data used in uncertainty analysis. The non-linearity and the self-uncertainty modelling features can jointly be found in the fuzzy inference modelling systems (Jang, 1993). These models have been successfully implemented in a number of diffuse pollution modelling studies. For instance Schärer et al. (2006) used a fuzzy decision tree to estimate P export at a catchment scale and also Shrestha et al. (2007) modelled nitrate dynamics in a catchment using a hybrid deterministic–fuzzy rule based model.

The aim in this study is to broaden and strengthen the empirical modelling approach by employing an adaptive neuro-fuzzy inference system (Jang, 1993) to develop a new empirical P export model. In developing this model the available catchment data is first partitioned into a number of clusters based on similarities in their characteristics. A hypothesis is made here that there is a physical basis for the clustering and that if a separate P export model is calibrated for each cluster, a better prediction of diffuse P loads in a catchment would be obtained by combining the outputs of all the cluster models in proportion to the catchment's membership weighting for each cluster. This means that the models developed for each cluster contribute to the diffuse P loads prediction in a catchment depending on the degree by which this catchment belongs to the cluster. The newly developed model is intended to be used as a predictive tool at a catchment level across all the River Basin Districts in Ireland and also with the view that an analogous approach can be used in other countries. For direct comparison with previous models, the new model has been developed and tested with the same data used by Daly and Mills (2006) in their model.

2. Estimation of nutrients loads using catchment characteristics

The level of nutrients, including phosphorus (P) and nitrogen (N), in a stream is usually an indicator of the situation in its upland catchment. Therefore in situations where diffuse pollution is significant it is always possible to obtain some estimate of nutrient levels from empirical models conditioned on catchment characteristics. The particular catchment characteristics which result in a robust model may not be known in advance and hence a trial and error procedure is usually followed to determine the best catchment characteristics. The relationship between the nutrient loads and the catchment

characteristics in the export coefficient models (e.g. McGuckin et al., 1999) is always described by a first order multiple linear regression model as follows:

$$L = b_0 + \sum_{k=1}^{nvar} b_k x_k \quad (1)$$

where L nutrient load;

x_k the value of the k^{th} catchment characteristic;

$nvar$ total number of the catchment characteristics;

b_0 constant term of the linear regression model;

b_k coefficient of the k^{th} catchment characteristic of the linear regression model.

The total number of terms in the linear regression model is equal to the total number of catchment characteristics which have been included in the model plus one. The constant term and the coefficients, (i.e. the model parameters) are estimated using the least squares parameter estimation method. To obtain reliable estimates for the parameters it is always recommended to use data from as many sites as possible. However it is also recommended to select sites from a homogenous region where similar catchment characteristics (e.g. phosphorus desorption index (PDI), runoff risk index (RRI) (see Appendix II for further explanation of PDI and RRI), soil types, land use types, geology, aquifer types) prevail so that the resulting model would be a better representation, but only of that region. Hence it is not advisable to use such a model in regions outside the one used in estimating its parameters.

Here a new approach has been developed to produce a class of model that can be more readily applied in heterogeneous regions. The approach is based on fuzzy inference

systems already used extensively in hydrological and water quality modelling (e.g. Chen et al., 2006; Dixon, 2005; Haberlandt et al., 2002, Jacquin and Shamseldin, 2006; Marce et al., 2004; Nayak et al., 2004). These modelling systems integrate the outputs from a number of sub-models to estimate a single overall output. Each sub-model can be considered as representative of a specific region type where the catchment behaviour is assumed homogeneous. The data used in developing the model are for the same 84 catchments used by Daly and Mills (2006) to develop their national P model. Such a national model is a tool of extreme importance in managing diffuse P pollution at a catchment level in each River Basin District in Ireland. The newly developed model is aimed at providing an improved, albeit more complex, alternative national P model. The model is tested by using part of the data set to calibrate the model parameters and the remaining part to validate the performance of the resulting model.

3. New neuro-fuzzy national P export model

Using a single general equation to estimate diffuse P loads from catchment characteristics may work well for a single homogeneous region but may not give good predictions outside of this region. Thus its use for the whole of Ireland is questionable. The reason for this is the wide variability in the behaviour of the catchments used to derive the equation. If among those catchments there is a dominant cluster of catchments with a homogenous condition then this cluster would influence strongly the parameter estimation process. The estimated parameters would fit well for catchments in this cluster while its performance for other catchments may not be as good. It is possible to improve the model performance if a separate model is defined for each cluster of distinct homogenous catchments. However, when grouping the catchments into a number of

clusters there will always be overlaps between these clusters because some catchments may have features in common with more than one cluster and may be difficult to assign to a single cluster. In our proposed neuro-fuzzy approach, a catchment does not have to be a member of only one cluster, but is assigned a membership weighting relating to all clusters. Higher weighting implies stronger association between the catchment and that cluster. Our hypothesis is that “if a separate P export model is calibrated for each cluster then a better diffuse P loads prediction in a catchment can be obtained by combining the outputs of all the cluster models in proportion to the catchment’s membership weighting for each cluster”. This means that the models developed for each cluster contribute to the diffuse P loads prediction in a catchment depending on the degree by which this catchment belongs to the cluster. The newly developed neuro-fuzzy national P model uses an annual time step in simulation and its structure is illustrated in Fig. 1. The mathematical form of this model, which here describes the relationship between the average annual concentrations of ortho-P (resulting from diffuse P loads) and physical characteristics for catchment i , is as follows:

$$orthoP_i = \left(\frac{\sum_{j=1}^{nc} w_j * O_j}{\sum_{j=1}^{nc} w_j} \right)_i = \left(\frac{\sum_{j=1}^{nc} w_j * \left(b_{0j} + \sum_{k=1}^{nvar} b_{kj} * x_k \right)}{\sum_{j=1}^{nc} w_j} \right)_i \quad (2)$$

where $ortho-P_i$ average annual ortho-P concentration at the outlet of catchment i ;

nc total number of catchment clusters;

w_j weight given to the linear regression sub-model of the j^{th} cluster;

O_j output of the linear regression sub-model of the j^{th} cluster which represents the average annual ortho-P load contributed by this cluster;

$nvar$ total number of independent variables defining the catchment characteristics used in the linear regression sub-model of each cluster;

x_k value of the k^{th} catchment characteristic;

b_{0j} constant term in the linear regression sub-model of the j^{th} cluster;

b_{kj} coefficient of the k^{th} catchment characteristic in the linear regression sub-model of the j^{th} cluster.

Based on the above-mentioned hypothesis the model is a weighted average of a number of linear regression sub-models. The number of linear regression sub-models is equal to the number of clusters (nc) sufficient to represent homogeneous groupings of the catchments, i.e. each cluster consists of a number of catchments with similar characteristics. In addition, each cluster is represented by a centre point with properties which are assumed to be representative of all catchments in the cluster. The k-means clustering algorithm (Hartigan and Wong, 1979) described in the next section is used to assign the catchments into nc clusters and also to calculate a centre vector of the spatial variables for each cluster. In addition, a standard deviation vector for each cluster can be calculated using the resulting centre vector and the vectors of the spatial data for all catchments in the cluster.

By assuming that each cluster is a fuzzy set then it is possible to estimate the degree by which a catchment belongs to a cluster with a membership function. In the current model the most widely used Gaussian function was employed for this purpose (e.g. Jacquin and Shamseldin, 2006; Nasr and Bruen, 2008). It has two parameters, the location or the

centre vector of the cluster (c) and the scale or the standard deviation of the cluster (σ), while the spatial data vector (x) is the input variable. The form of the Gaussian function is as follow:

$$w = f(x) = e^{-\frac{\|x-c\|^2}{\sigma^2}} \quad (3)$$

This function gives the weights (w) for Eqn. 2 which determine the contribution of a sub-model to the overall estimation of the ortho-P concentration. Once the weights (w) for each cluster are fixed, Eqn. 2 becomes a linear regression model relating the spatial variables with the ortho-P concentration. The parameters of the linear model ($b_0, b_1, b_2, \dots, b_{nvar}$) for all clusters can then be found using the least square method.

4. k-means clustering algorithm

Clustering is the partitioning of a data set into sub-sets (clusters), so that the data in each sub-set (ideally) share some common characteristics. Some defined distance measure such as the Euclidean distance is often used to determine proximity of the data in a cluster. The k-means clustering algorithm (Hartigan and Wong, 1979) is one of the simplest unsupervised learning algorithms for this partitioning when the number of clusters (k) is known a priori. The number of clusters is normally determined based on the amount and characteristics of the data which is used in calibrating the model. Using many clusters will result on a complex model with many parameters and this requires large amount of data to obtain these parameters with any confidence. As will be described later the current model was tested with two and three clusters only due to the

limited amount of data which does not allow the testing of models with more than three clusters.

Generally the steps in implementing the k-means clustering algorithm can be summarised as follows::

- (1) The main idea is to start with some initial choice of positions for the k centroids, one for each cluster. These initial centroids should be chosen carefully because different starting locations generate different results. They should be as far away from each other as possible, given the data set. In the current model the available catchments were randomly divided into k clusters and then the centroid of each clusters was initially defined as the mean values of the catchment characteristics for these catchments
- (2) The next step is to take each catchment and associate its catchment characteristics with the nearest centroid.
- (3) At this point k new centroids are calculated as the points which represent the centres of gravity of the new clusters resulting from the previous step.
- (4) Steps (2) and (3) are repeated until the change in the k centroids is insignificant.

In essence the algorithm aims at minimising an objective function, in this case a squared error function (J) in the following form:

$$J = \sum_{l=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_l\|^2 \quad (4)$$

where $\|x_i^{(j)} - c_i\|^2$ is a distance measure between a data point $x_i^{(j)}$ and the cluster centre c_i ;

n total number of the data points.

5. Application of the neuro-fuzzy national P model

Data from 84 different catchments in Ireland (Fig. 2) were used to develop the neuro-fuzzy national P model. These were split into two sets, the first of which was used for model calibration and the second for validation. In the calibration phase, the centre vector, the standard deviation vector, and the linear model parameters were calculated for each cluster using the calibration data set. Then, the second data set was used to verify the performance of the neuro-fuzzy national P model. The strategy of splitting the data set (84 catchments) into two parts for calibration and validation is important for the credibility of the resulting model. Here we have divided the available data in calibration and validation sets in six different ways and taken the mean of the results. Two of these cases have an equal number of catchments (42) in both calibration and validation sets and four cases have 63 catchments in the calibration and 21 in the validation set. Taking the mean of the results is more robust than taking the results from any single case.

6. Variables in the neuro-fuzzy national P export model

In the neuro-fuzzy national P export model the dependant variable, the annual average ortho-P concentration for a particular catchment, can be estimated from values of two or more indices representing phosphorus desorption index (PDI), runoff risk Index (RRI), Geology (GEO), Groundwater (GW), Land use (LU), and Soil (SO). The PDI and RRI described in Appendix II were introduced by Daly and Mills (2006) in their national P

model to quantify the potential risk of P loss from soil by the desorption process and the transport of P by surface runoff respectively. They found a strong correlation between both indices and ortho-P concentrations and hence they have been included in the current neuro-fuzzy national P model. Each index was obtained by calculating an area weighted average of risk categories defined subjectively for each soil type in a catchment (see Appendix II). The calculated values of the two indices were considered to be one of four risk degrees. These include (i) Low (ii) Moderate (iii) High; and (iv) Very High. The range of values in each degree was arbitrarily defined by Daly and Mills (2006). The maps of the other spatial variables in each catchment show different categories for each variable distributed over the catchment area. Any category which occupies less than 10% of the area in any of the 84 catchments was ignored. Table 1 summaries the categories of the spatial variables included in the model.

The frequency distribution of the observed average annual ortho-P concentrations in the 84 catchments is shown in Fig. 2. Some statistics calculated from the data are also presented in the same figure. The average annual ortho-P concentrations in the 84 catchments are found to be varying from a low of 0.004 mgP/l to a peak of 0.12 mgP/l; i.e. the range is 0.116 mgP/l. Most of the catchments have average annual ortho-P concentration between 0.012 mgP/l and 0.024mg/l.

7. Formulation of different neuro-fuzzy national P export models

Because of the limited number of catchments with sufficient data available to calibrate the linear model parameters for all cluster sub-models in the neuro-fuzzy national P model described above, only 2 and 3 clusters per catchment have been investigated to

date. The amount of data available places an upper limit on the number of parameters which can be calibrated. This in turn places an upper limit on the number of catchment characteristics (independent variables) and the number of clusters that can be used. Using many catchment characteristics each with a number of categories is not practicable if the resulting total number of parameters for the linear regression sub-models (for the case of 2 and 3 clusters) is larger than the number of data points in each of the six calibration-validation scenarios. Here we limit the investigation to a maximum of 3 clusters and also investigate a range of combinations from 2 to 4 independent variables (Table 2). When sufficient information from more than the 84 catchments used in the current study becomes available, a wider range of model structures can be tested. Running the k-means algorithm on the calibration data for the 2 clusters case has produced 21 catchments in each cluster. Then for the 3 clusters case it placed 12 catchments in cluster 1, 17 catchments in cluster 2 and 13 catchments in cluster 3. A scatter plot between the PDI versus the RRI was used to show the relative location of the centroid of each cluster with respect to the other catchment members of this cluster. Figures 4 and 5 show this scatter plot for the 2 and the 3 clusters cases respectively. No attempt was made in this study to analyse for geographic relationships between the catchments in each cluster.

8. Analysis of Results

The 11 candidate neuro-fuzzy national P export models listed in Table 2 were assessed for the 2 clusters and 3 clusters cases on the basis of the coefficient of correlation (R^2) between modelled and measured average annual ortho-P concentrations for both calibration and validation data sets. The results are shown in Table 3. For each model, the no-clustering case with a structure similar to the one of Daly and Mills (2006) was also

tested and presented as a base case against which the performance of the models with clustering is compared. Thus both the differences in performances achieved by adding additional explanatory variables and/or model complexity can be determined.

Calibration: The calibration R^2 values range from 0.43 for the no-cluster model to 0.86 for one of the 3-clusters cases (Table 3). Note the use of 2 clusters always gave R^2 values better than the no-cluster case and that 3-clusters were better than 2 clusters for all models. Compared to model_1 the performance of all other models showed an improvement in both the 2 and 3 clusters cases. The R^2 value for model_11 with 3 clusters is the best. Although in many modelling cases, it might be expected that the more complex model should do better than the simpler model in calibration, for the type of model considered here, the more complex model does not necessarily contain the simpler model as a special case, because the clustering may be different. Hence it is possible for the simpler model to perform better in calibration than a more complex one.

Validation: As shown in Table 3 the validation results are not as good as the calibration results, with 0.56 being the best value. In general the results for the 2 clusters were better than the 3 clusters case in 6 cases, were worse in 2 and equal in one case. However the 2 clusters case was better than the no-clusters case for 9 of the 11 cases, supporting the use of clustering. As for which are the best combinations of independent variables, model_3, which has a Groundwater index (GI) in addition to the RRI and PDI variables of model_1 was the best for the no cluster, 2 clusters and 3 clusters cases.

9. Discussion

The neuro-fuzzy national P export model is expected to provide a powerful tool which can facilitate the prediction of the annual amounts of diffuse source P from a catchment using only the catchment characteristics as inputs. Such predication is required during the design of any management plan to reduce the amount of P loss from land to water. The model has been calibrated using data from 84 catchments from different regions in Ireland. A split sample technique has been used, in which some data is used in calibrating the model and the remaining data used to validate the calibrated model. This independent validation result is important to judge the possibility of generalising the use of the model for predictions in other catchments not included in the calibration.

Calibration of the models has been performed for six different random divisions of the available data into calibration and validation sets. For each of the six calibration-validation scenarios 11 different models have been formulated to determine from among the six spatial variables (PDI, RRI, GEO, GW, LU, and SO) the appropriate ones which can be used as predictors to the average annual Ortho-P concentrations. All models have been run for the cases of 2 and 3 clusters or sub-models. Then performance of each model was assessed based on the mean of R^2 values for the six scenarios. Generally the results of R^2 during calibration indicated that for all models the use of 3 clusters is better than 2 clusters. This finding is expected since the use of 3 clusters or sub-models increases the number of parameters in the model and this in turn increases the degree of freedom in the model and hence better calibration results can generally be expected. Nevertheless the use of many parameters may not result in a good performance during validation if the model has been over-parameterised.

Using spatial variables other than the PDI and RRI as predictors has been examined by comparing the value of R^2 for model_1 with the values of the other models for the 2 and 3 clusters cases. The calibration results suggested that there is a benefit for the model in additional spatial variables as well as PDI and RRI. However, the validation results do not always show the same trend except for few models and this suggests that adding more variables to the PDI and RRI in a model may result only in a slight improvement. The variety in performance in the validation results emphasises the variability in the degree by which the spatial variables influence the processes which affect the mobilisation and transportation of P from land to water.

Figures 6 and 7 compares the observed average annual ortho-P concentration with the estimated values for the models which resulted in the best R^2 values during validation for the cases of 2 and 3 clusters respectively. In the two figures the points which represent the actual and the estimated values are closely scattered around the 1:1 line and this indicates a reasonable matching between the observed and the estimated values. However, an underestimation of all values larger than 0.05 mgP/l by the models is noticeable.

To investigate the usefulness of the neuro-fuzzy national phosphorous model, the best model for each scenario was compared with a model that used same input variables and had a structure similar to the Daly and Mills (2006) model. The comparison of the models was based on the R^2 values (shown in Table 4). It is obvious from the table that, in all scenarios, the neuro-fuzzy national phosphorus model was better than the original Daly

and Mills (2006) model. However, Table 4 also shows that in all scenarios, except for Scenario 4, no single neuro-fuzzy model order was found to be the best for both calibration and validation.

10. Conclusions

The concept of fuzzy modelling was applied to develop a national model of annual average ortho-P concentrations using catchment characteristics as independent variables. Data from 84 catchments from Ireland were used in developing and testing the new model. The k-means clustering algorithm has been used to determine 2 and 3 clusters of similar catchments. For each clusters case 11 different models have been formulated by using different input variables selected from among 6 candidate spatial variables (PDI, RRI, GEO, GW, LU, and SO). The following conclusions can be drawn from the results of the model application:

- (1) The new fuzzy clustering model performs better than the no cluster case at predicting the annual average ortho-P concentrations at a catchment level. Such a model is quite general and can be used in a wide range of applications related to the implementation of the WFD in Ireland. For instance in assessing any proposed land use management option to minimise the P loss in a catchment the model can be used to estimate the P load in the catchment under current land use conditions and thereafter to predict the change in P load post the implementation of the proposed land use management option in the catchment. This could inform economic analyses on the effectiveness of land use change measures. The model can also assist in identifying the most critical combinations of land use and soil type from the point of view of P export.

- (2) The best calibration results were obtained for the more complex models (i.e. many spatial variables) and those using 3 clusters. However the validation results indicated that the best models mostly have 2 clusters and fewer (2 or 3) spatial variables.
- (3) PDI and RRI are the essential variables but not the limited variables in predicting annual average ortho-P concentrations. The use of other spatial variables, particularly groundwater (GW), can improve the prediction and also their use is recommended if the resulting model is to be used for studying the effect of different catchment management options. In fact the model with only PDI, RRI and GW performs best in validation, regardless of the number of clusters used.

Acknowledgements

The authors would like to acknowledge the Environmental Research Centre of the Environmental Protection Agency (Ireland) for the financial support and for making its GIS database available to the study. A number of organisations are also acknowledged for providing the data of the catchments used in the study. These include Teagasc, the Three Rivers Project, and the Lough Derg/Lough Ree project. In particular, the data provided by Dr. Karen Daly of Teagasc was essential for this study, which could not have been done without this data.

References

Chen SH, Lin YH, Chang LC, Chang F.J. The strategy of building a flood forecast model by neuro-fuzzy network. *Hydrol. Processes* 2006, 7:1525-1540.

CORINE (Co-ORDination of INformation on the Environment). Land Cover Project, The Nomenclature. Commission of the European Communities 1989; Directorate-General Environment Nuclear Safety and Civil Protection, Rue de la Loi 200, B-1049 Brussels, Belgium.

Daly K, Mills P. Eutrophication from agricultural sources: Relating catchment characteristics to phosphorus concentrations in Irish Rivers - Final Report. Environmental Protection Agency; 2006, P O Box 3000, Johnstown Castle, Wexford, Ireland.

Daly K, Styles D. Eutrophication from Agricultural Sources – Phosphorus Chemistry of Mineral and Peat Soils in Ireland. Final Report. Environmental Protection Agency 2005, P O Box 3000, Johnstown Castle, Wexford, Ireland.

Daly K., Coulter B. National phosphorus model. In: Tunney H, Coulter B, Daly K, Kurz I, Coxon C, Jeffrey D, Mills P, Kiely G, Morgan G, editors. Quantification of phosphorus loss from soil to water. Synthesis Report, R&D Report Series no. 6. Environmental Protection Agency; 2000, P O Box 3000, Johnstown Castle, Wexford, Ireland.

Davis H, Neal C.. Estimating nutrient concentrations from catchment characteristics across the UK. *Hydrol. Earth Syst. Sci.* 2007, 11(1): 550-558.

De Martino G, De Paol, ., Fontan, N, Marin, G, Ranucci A. . Experimental assessment of level pool routing in preliminary design of floodplain storage. *Sci. Total Environ.* 2012, 416(1): 142-147.

De Paola F, Ranucci A. Analysis of spatial variability for stormwater capture tank assessment. *Irrig. Drain.* 2012, online library, DOI: 10.1002/ird.1675.

Dixon B. Applicability of neuro-fuzzy techniques in predicting groundwater vulnerability: a GIS-based sensitivity analysis. *J. Hydrol.* 2005, 309: 17-38.

EEC (European Economic Community). The council of 23 October 2000 establishing a framework for Community action in the field of water policy (2000/60/EC). Official Journal of the European Communities 2000, L327:1-72.

Gardiner MJ, Radford T. Soil Associations of Ireland and their land use potential (Explanatory Bulletin to Soil Map of Ireland 1980). An Foras Taluntais; 1980, 19 Sandymount Avenue, Dublin 4, Ireland.

Haberlandt U, Krysanova V, Bardossy A. Assessment of nitrogen leaching from arable land in large river basins: Part II. Regionalisation using fuzzy rule based modelling. Ecol. Modell. 2002, 150: 277-294.

Hartigan JA, Wong MA. Algorithm AS 139: A k-means clustering algorithm. Appl. Stat. 1979, 28: 100-108.

Hutton SA, Harrison SSC, O'Halloran J. Literature Review: An evaluation of the role of forests and forest practices in the eutrophication and sedimentation of receiving waters. Western River Basin District Publication; 2008, Water Framework Directive, Ireland.

Jacquin AP, Shamseldin AY . Development of rainfall-runoff models using Takagi-Sugeno fuzzy inference systems. J. Hydrol. 2006, 329: 154-173.

Jang JSR. ANFIS architecture. In: Jang JSR, Sun CT, Mizutani E, editors. Neuro-fuzzy and Soft Computing, Prentice Hall, New Jersey; 1993.

Johnes PJ. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. J. Hydrol. 1996, 183: 323-349.

KMMP (Kirk McClure Morton and EG Pettits). Lough Leane Catchment Monitoring and Management System: Final Report. Kirk McClure Morton, Consulting Engineers; 2001, Dublin, Ireland.

Lehane M, O'Leary B. Ireland's Environment 2012 – An Assessment. Environmental Protection Agency; 2012, P.O. Box 3000, Johnstown Castle Estate, Wexford, Ireland.

Lek S, Guiresse M, Giraudel J. Predicting stream nitrogen concentration from watershed features using neural networks. *Water Res.* 1999, 33(16): 3469-3478.

Marce R, Comerma M, Garcia JC, Armengol J. A neuro-fuzzy modelling tool to estimate fluvial nutrients loads in watershed under time-varying human impact. *Limnol. Oceanogr.: Methods* 2004, 2: 342-255.

McFarland AMS, Hauck L. Determining nutrient export coefficients and source loading uncertainty using in-stream monitoring data. *J. Am. Water Resour. Assoc.* 2001, 37(1): 223-236.

McGarrigle M., Bowman JJ, Clabby KJ, Lucey J, Cunningham P, MacCarthaigh M, Keegan M, Cantrell B, Lehane M, Clenaghan C, Toner PF. Water quality in Ireland 1998 - 2000. Environmental Protection Agency; 2002, P.O. Box 3000, Johnstown Castle Estate, Wexford, Ireland.

MCOS (M.C. O'Sullivan and Company Ltd.). Three Rivers Project: Final Report. Consulting Engineers, 2002, Dublin, Ireland.

McGuckin SO, Jordan C, Smith RV. Deriving phosphorus export coefficients for CORINE land cover types. *Wat. Sc. Tech.* 1999, 39(12): 47-53.

Meynendonckx J, Heuvelmans G, Muys B, Feyen J. Effects of watershed and riparian zone characteristics on nutrient concentrations in the River Scheldt Basin. *Hydrol. Earth Syst. Sci.* 2006, 10: 913-922.

Nasr A, Bruen M. Development of neuro-fuzzy models to account for temporal and spatial variations in a lumped rainfall-runoff model. *J. Hydrol.* 2008, 349: 277-290.

Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS. A neuro-fuzzy computing technique for modelling hydrological time series. *J. Hydrol.* 2004, 291, 52-66.

Schärer M, Page T, Beven K. A fuzzy decision tree to predict phosphorus export at the catchment scale. *J. Hydrol.* 2006, 331: 484-494.

Shrestha RR, Bárdossy A, Rode M. A hybrid deterministic-fuzzy rule based model for catchment scale nitrate dynamics. *J. Hydrol.* 2007, 342: 143-156.

Su B, Kazama S, Katayama S, Sawamoto M. Influence of catchment characteristics on nitrogen concentration in largely forested streams during the base flow period. *Hydrol. Processes*, 2006, 20: 4391-4403.

Toner P, Bowman J, Clabby K, Lucey J, McGarrigle ML, Concannon C, Clenaghan C, Cunningham P, Delaney J, O'Boyle S, MacCarthaigh M, Craig M, Quin R. Water quality in Ireland 2001 - 2003. Environmental Protection Agency; 2005, P.O. Box 3000, Johnstown Castle Estate, Wexford, Ireland.

Appendices

Appendix I: Summary of Daly and Mills (2006) model

- A dataset of 84 sub-catchments were used to build a linear regression model that relates the flow weighted average ortho-P (fwOrtho-P) with a number of catchment characteristics.

- Each sub-catchment represents water quality monitoring station with daily mean flow and phosphorus data covering the period between 1998 and 2001. Also in all sub-catchment point source pollution was not known to cause an influence on water quality.
- Sources of data for the catchment characteristics include: Digital Elevation Model (DEM), Soil Map, General Soil Map, Land Use/Cover Map, GIS layer of National Soil Test P (STP), District Electoral Divisions (DED) Map, Central Statistics Office (CSO) data, Habitats Indicators Map
- Correlations was firstly performed between fwOrtho-P number of variables and the results of coefficient of correlation (R^2) are given in the table below:

Variable	R^2	Source of data
% of sub-catchment mapped as acidic soil	0.36	Soil Map
% of sub-catchment mapped as peat soil	-0.31	Soil Map
Phosphorus Desorption Index (PDI)	0.42	Soil Map and STP GIS layer
Runoff Risk Index (RRI)	0.41	General Soil Map
Un-improved Pasture	0.53	Land Use/Cover Map
Improved Pasture	0	Land Use/Cover Map
Soil Phosphorus Test category 1 (SP1) representing Morgan's P 0–6 mg/l	-0.38	STP GIS layer
Soil Phosphorus Test category 2 (SP2) representing Morgan's P 6–10 mg/l	0.32	STP GIS layer
Soil Phosphorus Test category 3 (SP3) representing Morgan's P above agronomic values	0	STP GIS layer
Soil Phosphorus Test Index (SPI)	0.29	STP GIS layer
Livestock Unit Density (LUD)	0	DED Map and CSO Data
Fertiliser P input and central statistics data;	0	DED Map
% of sub-catchment areas with Topographic Wetness Index (TWI) >12	0	DEM
% of sub-catchment areas with Dry Grass	0.43	Habitats Indicators Map

% of sub-catchment areas with Wet Grass	0	Habitats Indicators Map
% of sub-catchment areas with Grass Peat	0	Habitats Indicators Map

- A backwards-stepwise regression model was computed for fwOrtho-P concentrations using all of the available catchment data, namely, the land-cover classes, PDI, RRI, soil P, livestock density and fertiliser P indices. The final step in the regression model retained Unimproved Pasture, Arable, SPI and PDI that accounted for 41.4% of the variation in the fwOrtho-P data.
- Full report describing the model can be accessed in the following web-link: http://www.epa.ie/downloads/pubs/research/water/epa_eutrophication_from_agricultural_sources_ertdi42_final.pdf

Appendix II

Phosphorus Desorption Index (PDI)

Daly and Styles (2005) conducted a study to derive a desorption weightings for the peat and mineral Irish soils based on phosphorus sorption isotherms analysis carried out on a number of soil samples. The samples ranged in properties such as %OM and pH over a range of Morgan's Soil P Test values. The result of this study was used by Daly and Mills (2006) in order to define risk rank for the peat and the mineral soils based on desorption rates over similar ranges of Soil P Test and sorption capabilities. Thus, mineral soils with high sorption capacities and desorption rates were ranked as highest risk whilst peat soils were ranked as lowest risk. Further analysis on the mineral soils indicated that non-calcareous soils displayed the highest sorption capacities and the highest desorption rates compared to calcareous mineral soils. Following this analysis desorption in non-calcareous and calcareous soils was expressed relative to lowest desorption values in peat soils and calculated as a ratio to generate a phosphorus desorption index (PDI) that could

be used to weight each soil group in terms of risk of P loss by desorption. The risk weights are defined arbitrary as follows;

Soil categories	Risk weights
Non-calcareous mineral	3.2
Calcareous mineral	1.9
Peat	1

Runoff Risk Index (RRI)

The percentage of gley in a soil has been used as a detrimental factor to the potential runoff risk. Therefore threshold levels for the percentage of gley in a soil were decided upon and soils were divided into runoff risk categories and weighted against each other in terms of potential runoff risk based on values shown in the table below. The weightings were derived subjectively and are not based on measured data. For each sub-catchment an area-weighted runoff risk index (RRI) was generated by multiplying the area of each category by its assigned weight.

Runoff Risk Class	%Gley in Soil	Risk Weights
RR1	5-10	1
RR2	15-25	2
RR3	50	3
RR4	>75	4

Figure 1. Neuro-fuzzy national P export model

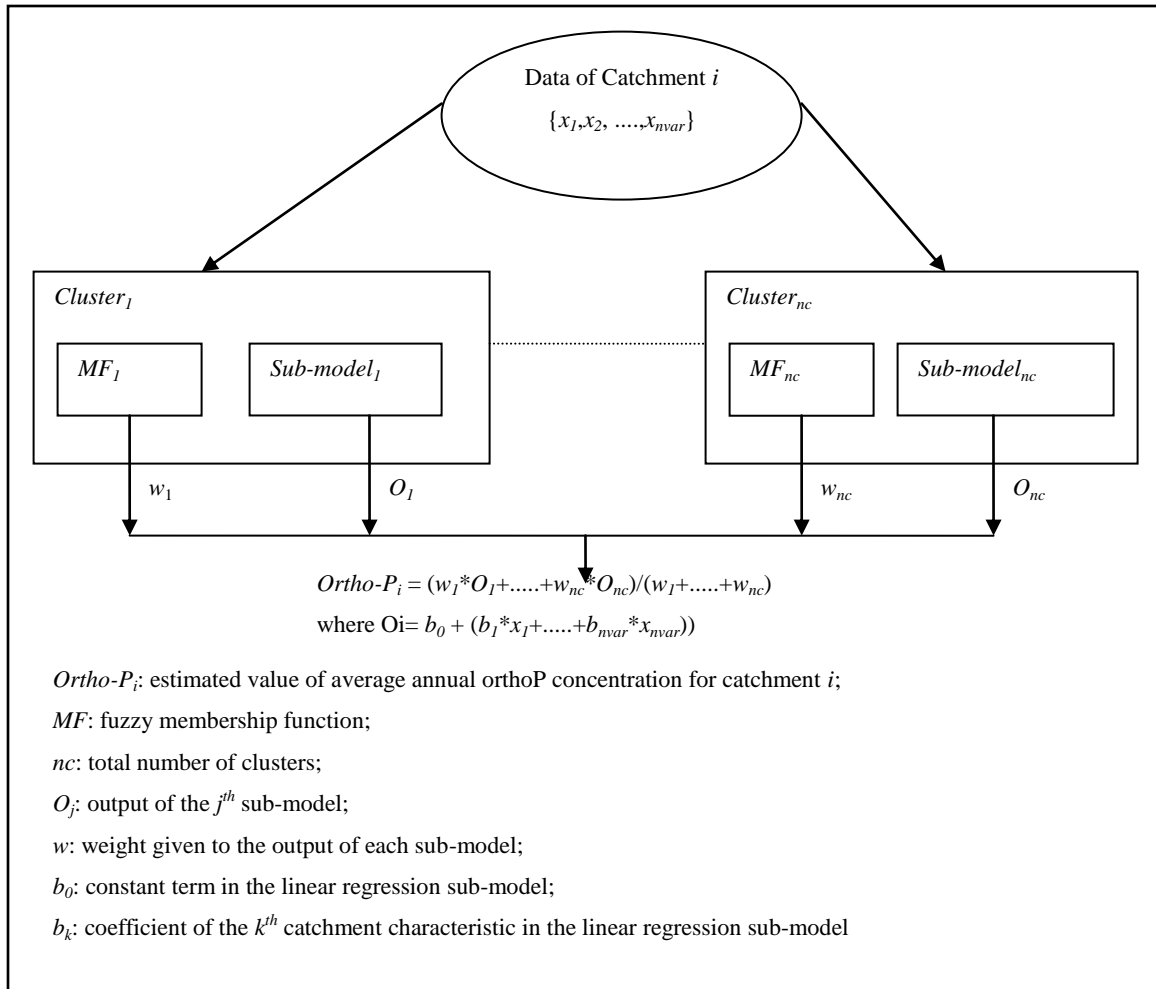


Figure 2. Distribution of the 84 catchments used in the neuro-fuzzy national P export model

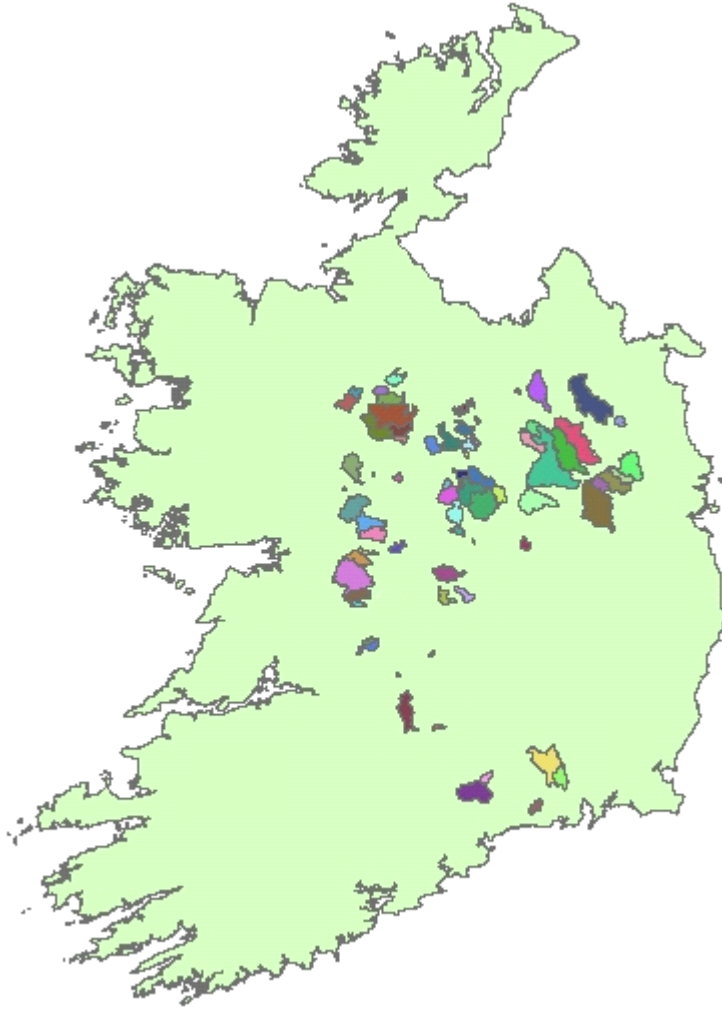


Figure 3. Frequency distribution of the annual average ortho-P concentrations in the 84 catchments used in the neuro-fuzzy national P export model

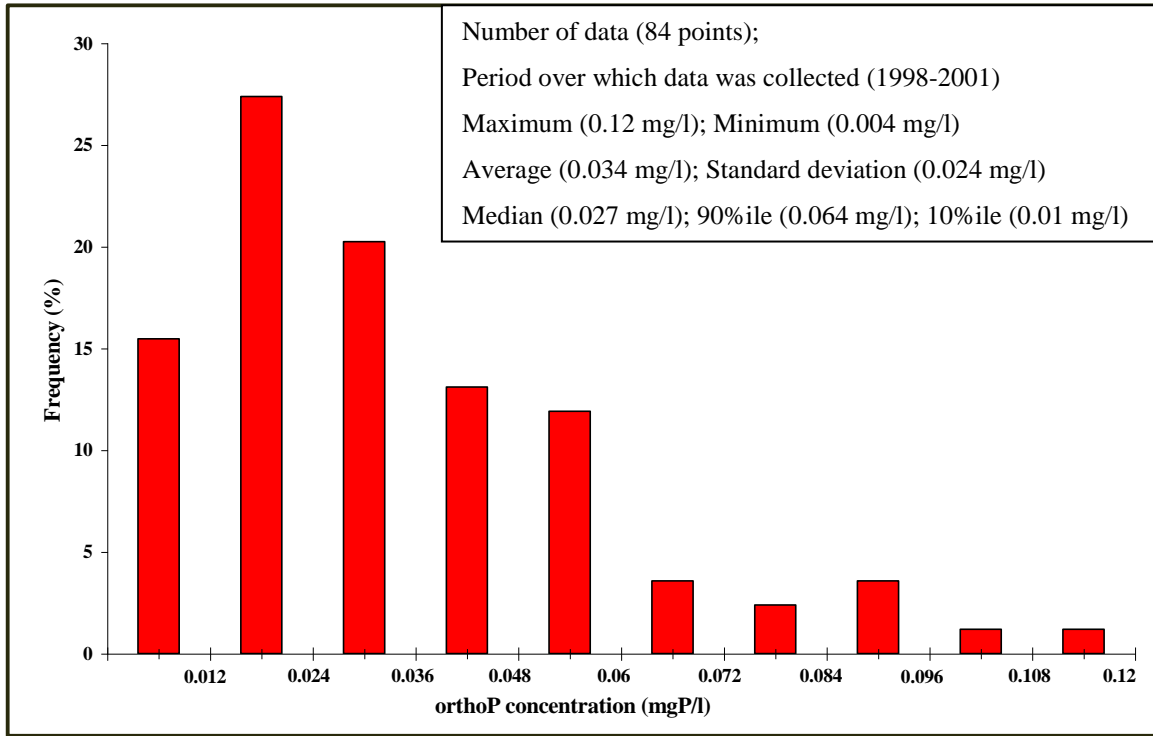


Figure 4. Clustering results for the 2 clusters case

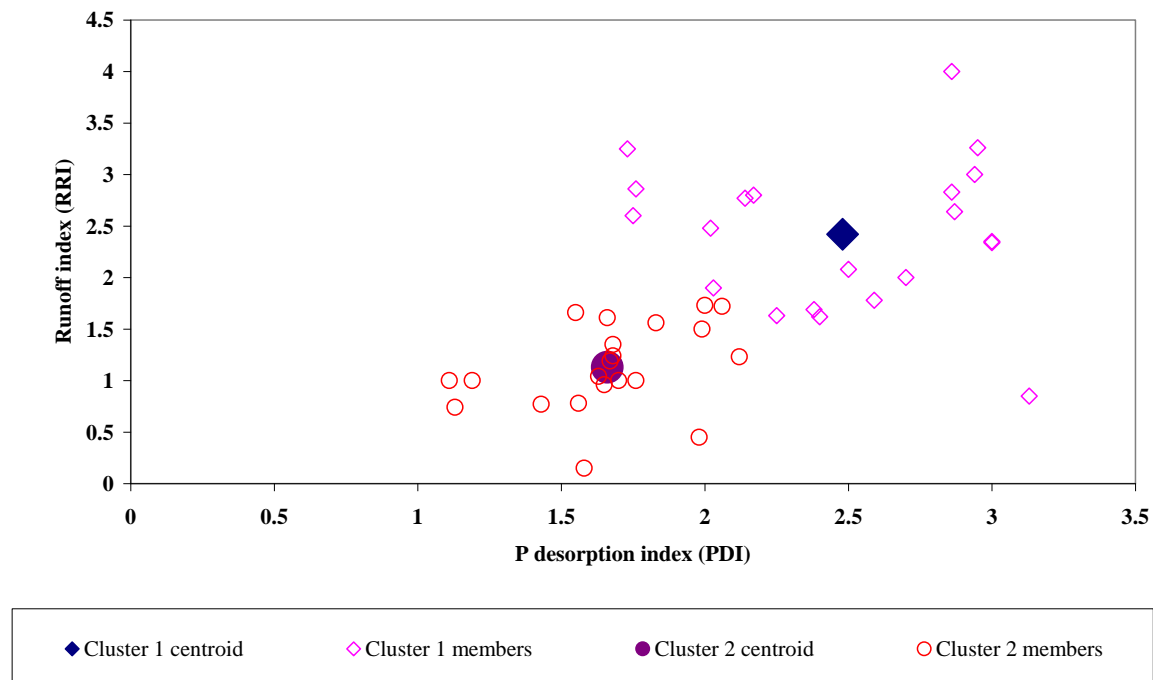


Figure 5. Clustering results for the 3 clusters case

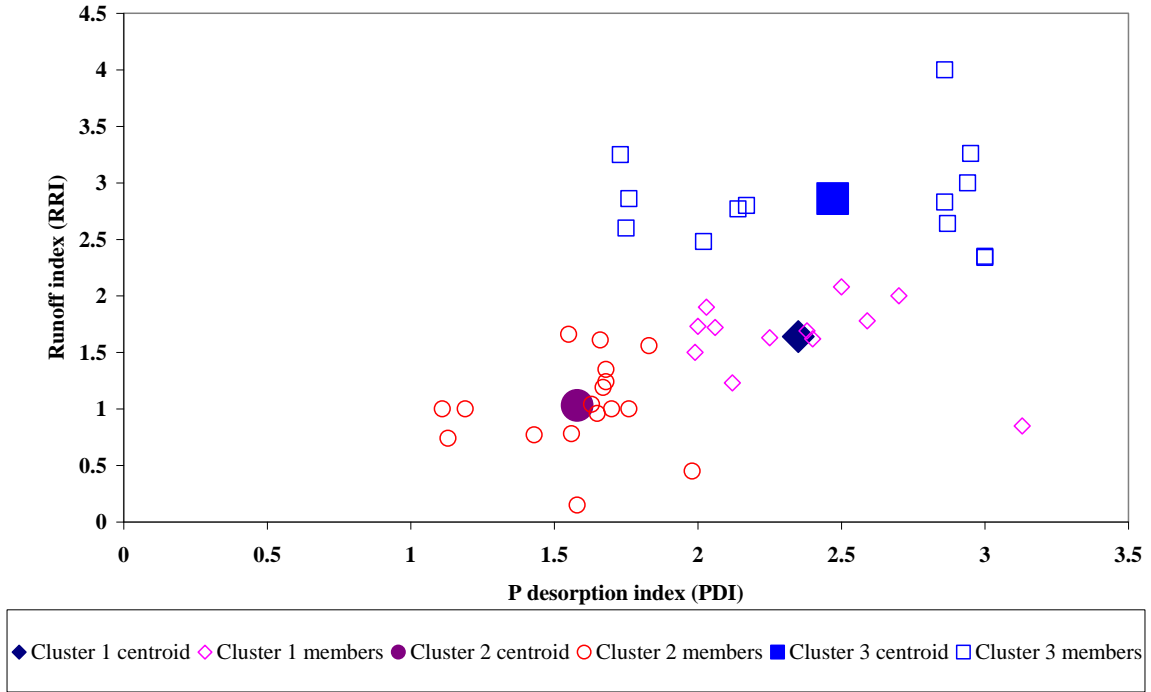


Figure 6. Observed vs Estimated average annual ortho-P concentrations for validation – model_3 of the 2 clusters case

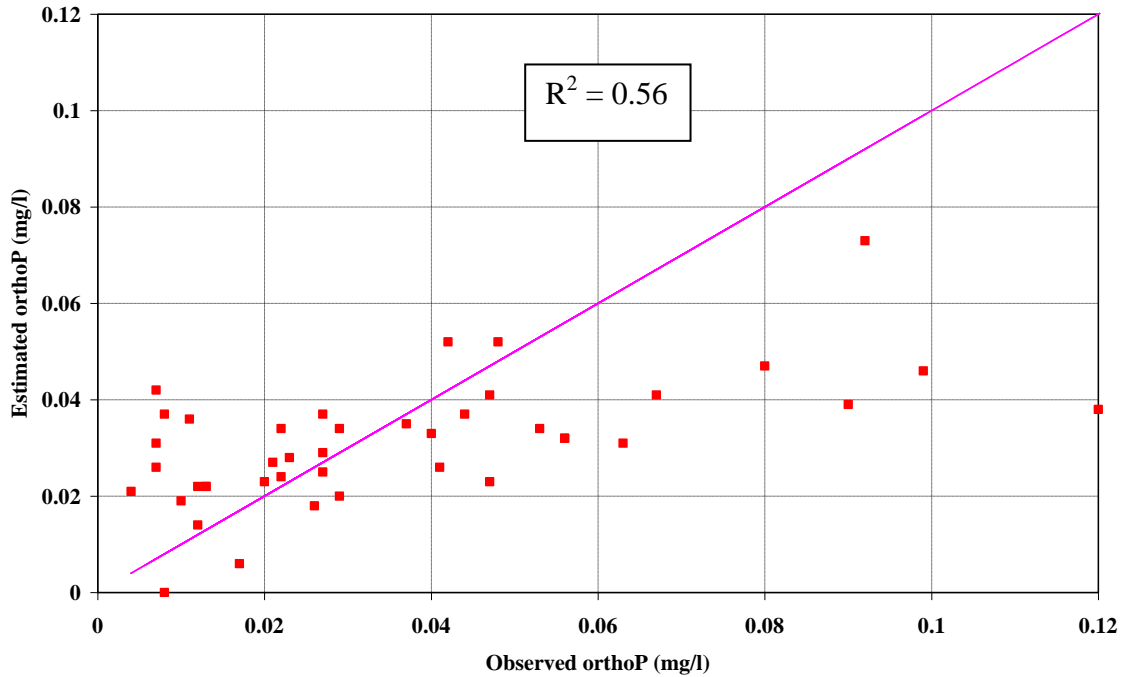


Figure 7. Observed vs Estimated average annual ortho-P concentrations for validation – model_3 of the 3 clusters case

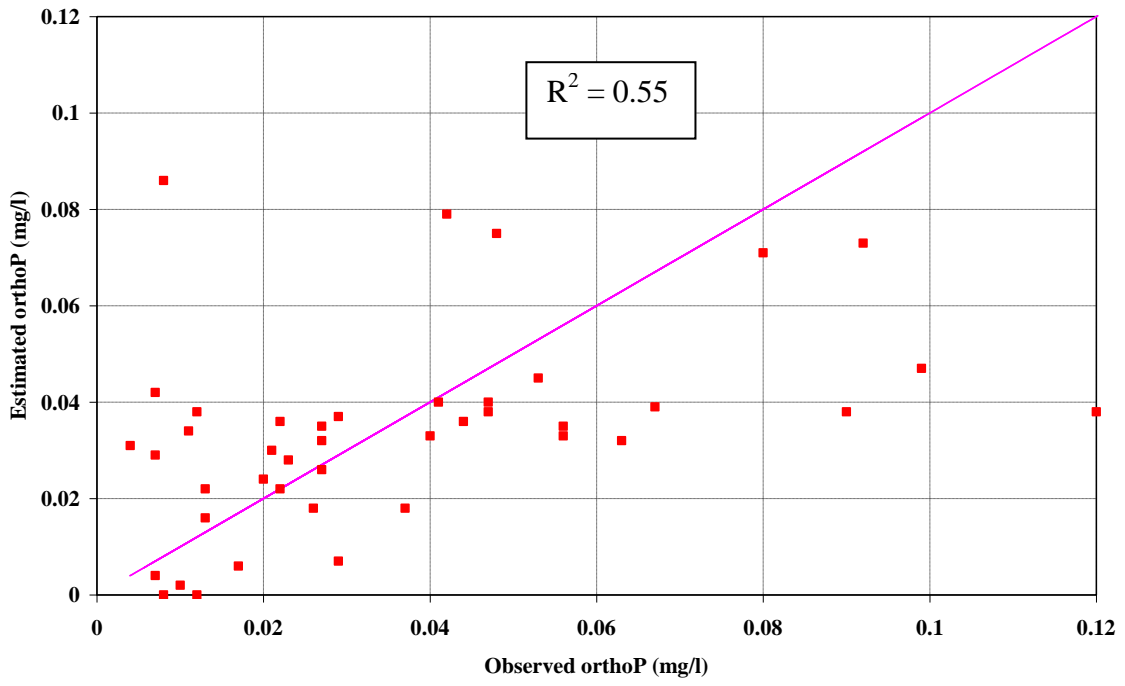


Table 1. Spatial independent variables tested in the neuro-fuzzy national P model

Variable	Categories	Source of Data
P desorption index (PDI)	(1) Low (2) Moderate (3) High	Daly and Mills (2006)
Runoff index (RRI)	(1) Low (2) Moderate (3) High (4) Very high	Daly and Mills (2006)
Geology (GEO)	(1) Sand and Gravels (2) Carboniferous limestone (3) Ordovician (4) Rhyolite	Geology map
Groundwater bodies (GW)	(1) Gravel (2) Karstic (3) Poorly productive bedrock (4) Productive fissured bedrock	Aquifer map
Land use (LU)	(1) Agricultural areas (2) Forest and semi-natural areas (3) Wetlands (4) Artificial surfaces	Derived from land use map (CORINE, 1989)
Soil (SO)	(1) Deep well drained mineral (2) Shallow well drained mineral (3) Deep poorly drained mineral (4) Poorly drained mineral soils with peaty topsoil (5) Peats (6) Miscellaneous	Soil map (Gardiner, and Radford, 1980)

Table 2. Potential neuro-fuzzy national P export models tested

Model	Variables of the model	Number of parameters for each sub-model	Total No of parameters – case of 2 clusters	Total No of parameters – case of 3 clusters
Model_1	PDI + RRI	3	6	9
Model_2	PDI + RRI + GEO	7	14	21
Model_3	PDI + RRI + GW	7	14	21
Model_4	PDI + RRI + LU	7	14	21
Model_5	PDI + RRI + SO	9	18	27
Model_6	PDI + RRI + GEO + GW	11	22	33
Model_7	PDI + RRI + GEO + LU	11	22	33
Model_8	PDI + RRI + GEO + SO	13	26	39
Model_9	PDI + RRI + GW + LU	11	22	33
Model_10	PDI + RRI + GW + SO	13	26	39
Model_11	PDI + RRI + LU + SO	13	26	39

Table 3. R^2 values for 11 candidate neuro-fuzzy models

Model	Calibration			Validation		
	No cluster	2 clusters	3 clusters	No cluster	2 clusters	3 clusters
model_1	0.43	0.48	0.58	0.48	0.43	0.45
model_2	0.45	0.66	0.70	0.47	0.56	0.52
model_3	0.54	0.64	0.73	0.50	0.56	0.55
model_4	0.61	0.71	0.79	0.21	0.28	0.32
model_5	0.54	0.69	0.75	0.33	0.45	0.42
model_6	0.56	0.76	0.79	0.46	0.53	0.45
model_7	0.64	0.74	0.81	0.23	0.28	0.28
model_8	0.56	0.78	0.85	0.37	0.28	0.30
model_9	0.68	0.78	0.83	0.25	0.41	0.24
model_10	0.62	0.75	0.83	0.38	0.45	0.25
model_11	0.63	0.77	0.86	0.22	0.28	0.54

Table 4. Summary of the neuro-fuzzy national phosphorus models which achieved the best R² values vs the Daly and Mills (2006) model

Scenario	Calibration				Validation			
	Neuro-fuzzy national P model			Daly and Mills model	Neuro-fuzzy national P model			Daly and Mills model
	Cluster case	Best Model	R ²	R ²	Cluster case	Best Model	R ²	R ²
1	3	Model_8	0.90	0.73	3	Model_4	0.38	0.32
2	3	Model_11	0.97	0.63	2	Model_3	0.65	0.5
3	3	Model_10	0.84	0.72	3	Model_9	0.60	0.38
4	3	Model_10	0.79	0.63	3	Model_8, Model_10	0.68	0.52, 0.49
5	3	Model_8, Model_10	0.87	0.41, 0.45	2	Model_3	0.76	0.66
6	3	Model_9	0.87	0.70	3	Model_1	0.65	0.56