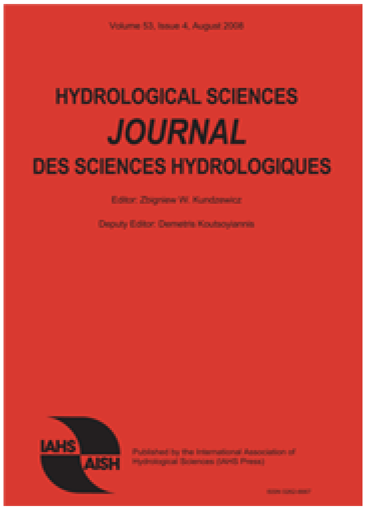


Estimation of phosphorus loads with sparse data for agricultural watersheds in the Czech Republic

Journal:	*Hydrological Sciences Journal*
Manuscript ID:	HSJ-2009-0054
Manuscript Type:	Original Article
Date Submitted by the Author:	20-Oct-2009
Complete List of Authors:	Urbanova, Tereza; T.G.Masaryk Water Research Institute, PRI, Department of Water Ecosystems Protection; Czech University of Life Sciences in Prague, Department of Landscape and Improvement Vogel, Richard; Tufts University Fiala, Daniel; T.G.Masaryk Water Research Institute, PRI, Department of Water Ecosystems Protection Rosendorf, Pavel; T.G.Masaryk Water Research Institute, PRI, Department of Water Ecosystems Protection

Estimation of phosphorus loads with sparse data for agricultural watersheds in the Czech Republic

TEREZA URBANOVA[1,3], RICHARD M. VOGEL[2], DANIEL FIALA[1] & PAVEL ROSENDORF [1]

[1] T.G. Masaryk Water Research Institute, Public Research Institution, Podbabská 2582/30, Praha 6, 160 00 Czech Republic, tereza_urbanova@vuv.cz

[2] Department of Civil and Environmental Engineering, Tufts University, Medford, MA, 02155, USA, Richard.vogel@tufts.edu

[3] Department of Land Use and Improvement, Faculty of Environmental Science, Czech University of Life Sciences Prague, Kamýcká 129, Praha 6 - Suchdol, 165 21, Czech Republic

Abstract Agricultural watersheds in the Czech Republic are one of the primary sources of nonpoint source phosphorus (P) loads in receiving waters. Since such nonpoint sources are in headwater catchments, streamflow and P concentration data are sparse. We show how very short daily streamflow and P concentration records can be combined with nearby longer existing daily streamflow records to result in reliable estimates of daily and annual P concentrations and loads. Maintenance of Variance streamflow record Extension methods (MOVE) are employed to extend short streamflow records. Constituent load regressions are used to predict daily P constituent loads from streamflow and other time varying characteristics. Annual P loads are then estimated. Resulting annual P load estimates ranged from 0.21 to 95.4 kg year^{-1} with a mean value of 11.77 kg year^{-1}. Similarly annual P yield estimates ranged from 0.01 to 0.3 kg ha^{-1}year^{-1} with an average yield of 0.07 kg ha^{-1}year^{-1}. We document how short records of daily streamflow and P concentrations can be combined with a national network of daily streamflow records in the Czech Republic to arrive at meaningful and reliable estimates of annual P loads for small agricultural watersheds.

Key words total phosphorus; agricultural watersheds; MOVE; regression, correlation, load

INTRODUCTION

Phosphorus plays a significant role in surface water eutrophication, since it often limits the growth of freshwater phytoplankton (Correll 1998), namely harmful cyanobacterial blooms which present the most serious manifestation of eutrophication. Substantial reduction of phosphorus loads will be required to meet the high water quality standards set by the Water Framework Directive of the European Union. While various control measures have been successfully adopted for point sources of pollution the regulation of non-point sources is much more challenging due to difficulties associated with the identification of source areas, quantification of pollution and application of appropriate control measures. Thus increasing attention should focus on reduction of non-point pollution.

Agricultural land use is a significant non-point source of phosphorus. Phosphorus is transported mostly by surface or subsurface runoff in the dissolved and particulate phase. An important source of phosphorus is the soil itself though the content of soil phosphorus varies among soil types and depends on the underlying geological formation. The natural content of soil phosphorus is further increased by fertilizer and manure application. Transport of phosphorus can be accelerated by inappropriate agricultural practices due to increased soil erosion and/or due to excessive application of fertilizer or manure or applications at the wrong time and/or place. An overview of phosphorus sources and transport from agricultural land is given by Hart et al. (2004) and Hansen et al. (2001).

A long record of soil testing in the Czech Republic indicates a significant relationship between fertilizer application and available soil P content. Intensive agriculture, promoted in less fertile regions and accompanied by large scale application of fertilizers especially in the 1970s and the 1980s, resulted in an increase of available soil P throughout the country (Central Institute for Supervising and Testing in Agriculture 2009). Accordingly, after the political change in 1989 which was followed by an economical transformation, the consumption of P fertilizers decreased until now up to 25% of the mean consumption prior to 1989 (Klement & Sušil, 2009). The consequent decrease of available soil P was not reported until the mid-1990s but the later results of soil testing showed a decrease of available soil P on arable lands by an average of 5 mg kg^{-1} between 1993-1998 and 1999-2004 soil testing cycles (Klement & Sušil, 2005) and further decrease by an average of 7 mg kg^{-1} between 1999-2004 and 2003-2008 soil testing cycles (Klement & Sušil, 2009). Despite these trends of available soil P levels described above and improvement of point sources of phosphorus, eutrophication of rivers and reservoirs remains a serious water quality problem.

In this study we introduce a methodology for quantification of annual total phosphorus (TP) loads in receiving waters associated with agricultural watersheds. We describe a procedure for estimation of constituent loads in rivers which have only been subject to sparse measurements of flows and concentrations. To avoid the influence of point sources, all study sites were located in agricultural headwater watersheds without any point source and permanent habitation. For such small watersheds, long periods of hydrologic and water quality data are not normally available. High frequency water quality and quantity sampling is often not feasible because it requires significant financial and personnel resources. Instead, we introduce a cost effective methodology which transfers hydrologic information from nearby gaged watersheds, using regression methods. The idea is to develop a relationship between daily streamflow observations at gaged watersheds and daily streamflow observations at watersheds with sparse records and then to use that relationship along with another relationship between streamflow and constituent concentrations to enable extension of very short records of phosphorus and streamflow measurements into records useful for planning purposes.

For small watersheds, most of the phosphorus is transported by surface runoff processes during rainfall events. Thus P concentration data should be collected at both fixed intervals and during rainfall-runoff events (Toor et al. 2008), which is – similar to high frequency sampling – demanding in terms of time, finance and work organization. Robertson & Roerish (1999) question the role of additional storm samples when estimating annual loads on small watersheds because they can result in positive bias. Instantaneous concentrations measured during storm events are often two or more orders of magnitude higher than average daily concentrations. Storm sampling strategies during extreme events are useful for calibration of watershed models and serve well for the description of temporal patterns in water quality, but should be treated carefully if ones' interest is only in estimation of unbiased annual loads (Robertson & Roerish 1999; Robertson 2003).

Typically, for small agricultural watersheds both water quality and flow records are either unavailable, or extremely sparse. In such cases it is possible to transfer hydrologic information from nearby sites with long-term flow records and/or with high frequency sampling by employing the cross-correlation between short and long records. Such information transfer techniques can by used to both fill-in missing observations and to extend short flow records (Salas 1993).

The primary goals of this study were: 1) to obtain reliable unbiased estimates of annual phosphorus loads for small agricultural watersheds in Central Europe 2) to demonstrate the utility of regression methods for extending very short daily streamflow records using information transfer techniques by exploiting the cross-correlation with nearby long-term streamflow records 3) to demonstrate the utility of regression methods for estimation of daily phosphorus loads from short samples of P concentrations combined with extended streamflow series for small watersheds with sparse flow and concentration data.

METHODS

Study sites and data collection

Fourteen watersheds distributed around the Czech Republic were studied using 20 sampling sites (Fig. 1). The watersheds exhibit only agricultural land uses with primarily arable land and none of the watersheds contain point or diffuse sources of pollution. The watersheds range in size from 16 to 575 ha with a mean watershed area of 198.6 ha. The watersheds represent various agricultural regions and the most commonly occurring soil types in the Czech Republic.

Daily P concentrations and streamflow discharges were measured at monthly intervals at the study sites during 2007-2008, which yields 24 observations of P and streamflow at each site. Samples were analyzed to determine TP concentration. In order to extend these short streamflow records, long term daily streamflow data at nearby gages was obtained from the national monitoring network operated by the Czech Hydrometeorological Institute. Watershed boundaries were delineated in upstream areas above the sampling sites in ArcGIS 9.1 using contour line vectors derived from 1:10,000 maps. Watershed areas were then calculated in order for us later on to define TP yields (kg ha^{-1}year^{-1}) from computed annual P loads and watershed area. Statistical analyses were carried out using the SPSS statistical package.

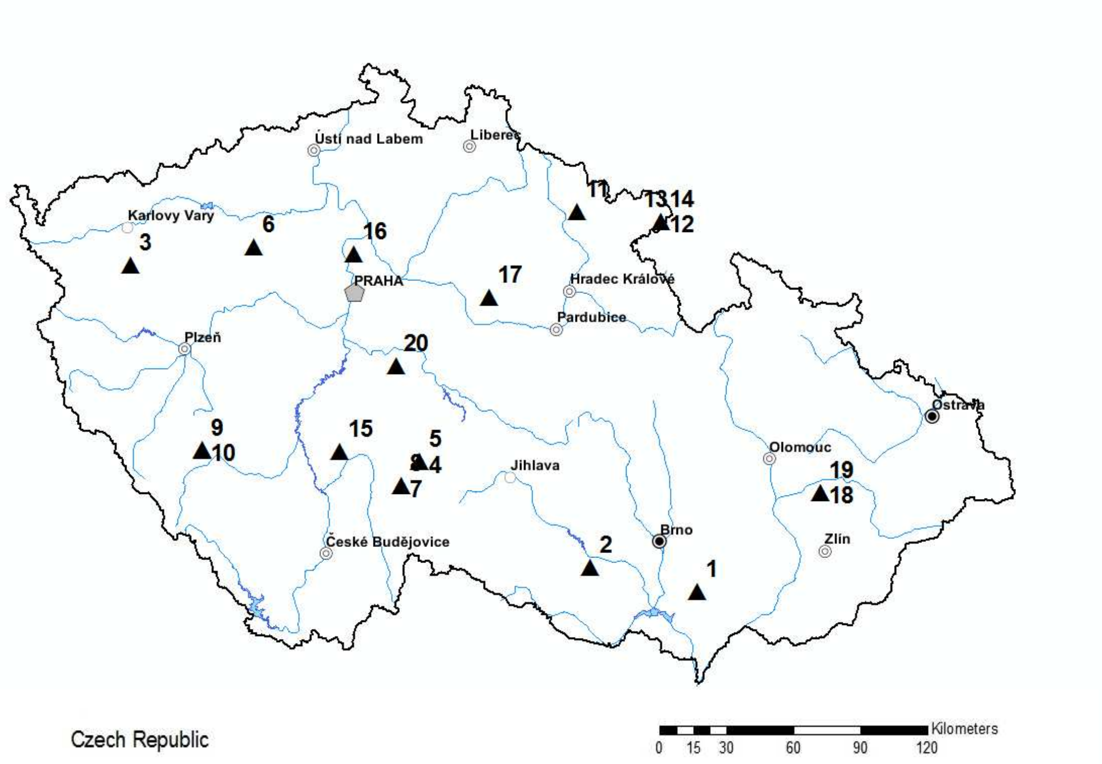


Fig. 1 Location of study sites in the Czech Republic

Extension of Short Daily Streamflow Records

In this section, the goal is to use streamflow record extension methods to create complete daily flow series over the period 2007-2008 from the 24 values of daily streamflow samples collected monthly at each study site. The idea is to exploit the cross-correlation between the short flow records and nearby long-term index records. In order to find the best index station for each site we computed the Pearson product-moment correlation coefficient between the logarithms of daily streamflow at the study sites and the various long term potential index gages in the vicinity of each site. The index station is chosen as that gaging station which exhibits the highest correlation between the logarithms of the corresponding flow series.

Once an index gaging station is chosen, we employ the maintenance of variance extension (MOVE) method for extending streamflow records. This method assumes that the standardized logarithms of the daily flows are equal at both sites, so that

$$\frac{y_t - \hat{\mu}_y}{\hat{\sigma}_y} = \frac{x_t - \bar{x}}{s_x} \qquad (1)$$

Where y_t and x_t are the logarithms of the daily flows at the short and long record sites, respectively, $\bar{x}$ and s_x are the sample mean and standard deviation of the long x record and $\hat{\mu}_y$ and $\hat{\sigma}_y$ are specialized estimators of the mean and standard deviation of the short record designed specifically, to generate minimum variance and unbiased estimates of the extended values of daily streamflow at the short record site.

Various MOVE techniques were first suggested by (Hirsch 1982) and later improved slightly by (Vogel & Stedinger 1985; Grygier et al. 1989 and others) to reduce the bias and variance in estimates of the mean and variance of the flows at the short record site and to obtain a reasonable and unique extended streamflow record.

In general, all MOVE methods can be described by rewriting equation (1) as

$$y_t = a + bx_t \qquad (2)$$

The principle of MOVE methods is to derive estimates of the model coefficients *a* and *b*, in (2) that transform the long record *x* data into estimates of *y* in such a way that the theoretical moments of the estimated *y* values equal their true, but unknown values. In this study we used the MOVE3 introduced by (Vogel & Stedinger, 1985) with the model coefficients, *a* and *b*, reported by both Vogel & Stedinger (1985) and Salas (1993).

The goodness-of-fit of the resulting daily streamflow extensions were assessed using the Nash-Sutcliffe model efficiency (NSE) (Nash & Sutcliffe, 1970). The NSE coefficient ranges from -∞ to 1, with higher values generally indicating a better fit. Values of NSE<0 indicate cases in which the extended sequence is a worse representation than if one were to simply report every extended streamflow using the sample mean of the short streamflow record.

Phosphorus Load Regression Model Development

Phosphorus concentrations and daily streamflows are often approximately described by a lognormal distribution and the logarithms of concentration and daily streamflow is often well approximated by a bivariate normal distribution (Clarke 1990; Vogel et al. 2005). Since loads are the product of concentration and flow, they may also be approximated by a lognormal distribution. Therefore regression analysis between loads and flows is usually performed after logarithmic transformation of both variables.

Regression models of TP load for each site were developed assuming a linear relationship between the logarithm of loads and streamflow. This simple log-linear model can be further improved by accounting for nonlinearities, seasonality, censored data, time trends, and residual serial correlation (Helsel & Hirsch 1992; Cohn 1995). Vogel et al. (2005) showed that the correlation between the logarithms of load and flow is always greater than the correlation between the logarithm of concentration and flow. This is termed spurious correlation (Kenney 1982) and results because load is the product of flow and concentration, thus load is functionally related to flow. Others have noted that although this increased correlation is spurious it may be useful in augmenting and extending a short record of load. The following multivariate linear regression was fit at each site:

$$\ln(L) = \beta_0 + \beta_1\ln(Q) + \beta_2\sin(\omega T) + \beta_3\cos(\omega T) \qquad (3)$$

where L is the total phosphorus load in mg day^{-1}, Q is daily discharge in dm^3 s^{-1}, T is Julian day, $\varpi = 2\pi/365$ and β_0 , β_1, β_2 and β_3 are regression coefficients. The sin and cos terms account for the seasonal variation in behaviour of phosphorus loads. Other terms may be added to (3) such as a trend term, however it was found that none led to model improvements.

Stepwise linear regression selection procedures were employed to determine, whether the seasonal factors led to improvements in the model. Regression performance was evaluated by using residual diagnostics to assure that model residuals were homoscedastic and approximately normally distributed. In addition, influence statistics were used to identify and eliminate outliers. Helsel & Hirsch (1992) summarize methods for development of such multivariate models including a complete discussion of the model diagnostics employed in this study.

Once developed, load regression models for each site of the form shown in (3) were used in combination with the extended streamflow records to estimate a complete daily series of TP loads over the period 2007-08 for all sites. Loads were then corrected for logarithm transformation bias introduced from the retransformation of the power-law model (Ferguson 1986) by multiplying the resulting loads in real space by a bias correction factor (BCF) as suggested by Cohn (1995) and others. See Vogel et al. (2005) for further background on the behaviour of the BCF for such load regression models. The estimated daily loads were summed to calculate annual loads.

Finally, both MOVE and regression methods were cross-validated using a single year of data for estimating model parameters and the following year for validation.

Values of streamflow and TP load for second year were predicted and compared with actual values to determine goodness-of-fit of the models.

RESULTS AND DISCUSSION

Streamflow Record Extension

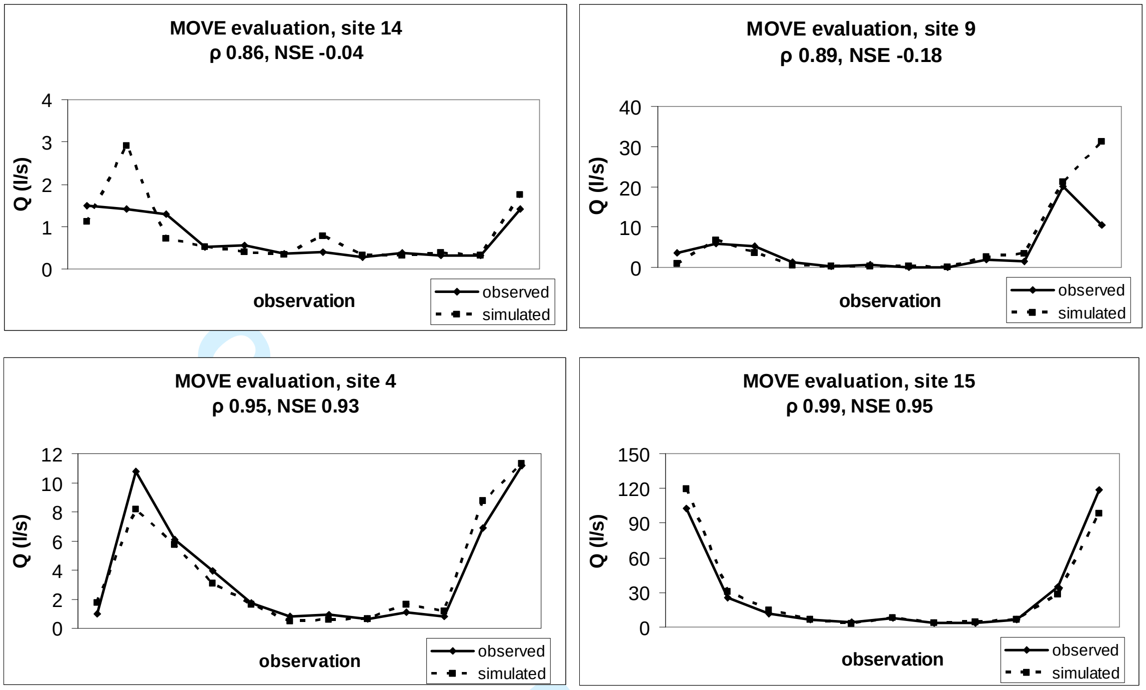
As is often the case in practice, only a single daily streamflow measurement was available for each month at each site. These sparse daily streamflow records were extended by filling in missing values using the MOVE3 information transfer method from nearby gages (see Vogel & Stedinger, 1985 and Salas 1993). The first step was to choose an index gage for each study site using Pearson´s correlation ρ between the logarithms of the daily streamflows at the study sites and various potential index gages. The long record gage y exhibiting the largest correlation with each short record site x was chosen as the index gage for that particular short record site. Table 1 documents the correlations which ranged from 0.86 to 0.99 with an average value of 0.91 for the 15 study sites. For the remaining 5 sites there were no index gages with correlations above 0.8, which is considered the minimum correlation needed to transfer information (Vogel & Stedinger, 1985).

Once an index station is chosen, the MOVE3 method was performed to extend the daily flow records at each site. To evaluate the MOVE3 method, we report the performance of MOVE3 for estimating the 24 reported values of daily streamflow which were available for each site. Here we use the Nash-Sutcliffe Efficiency (NSE) as a measure of MOVE3 model performance with the results reported in Table 1. As expected, the values of NSE are proportional to the values of ρ. Values of NSE ranged from -0.18 to 0.95 with a mean value of 0.66. Only two sites exhibited such poor values of NSE equal to -0.04 and -0.18, that we expect the MOVE3 method to perform poorly at those two sites.

Table 1 Pearson´s correlation coefficient ρ for correlation between logarithms of flows at study sites and nearby index gages and Nash-Sutcliffe Efficiency of the model performance in streamflow simulation

Site	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
ρ	0.95	0.87	0.90	0.95	0.94	0.86	0.96	0.88
NSE	0.92	0.26	0.75	0.93	0.81	0.78	0.91	0.86
Site	**9**	**10**	**11**	**12**	**13**	**14**	**15**	
ρ	0.89	0.87	0.99	0.91	0.91	0.86	0.93	
NSE	-0.18	0.68	0.95	0.79	0.72	-0.04	0.75	

Vogel & Stedinger (1985) and others show that information transfer gains of MOVE methods are highly dependent on the strength of correlation (value of ρ) and on the length of the short record. Information transfer gains generally increases as either ρ increases and/or the number of observations in the short record increases. Vogel & Stedinger (1985) document that as long as ρ>0.8, information gains are to be expected, and that is the case for 15 sites considered in this study. Figure 2 compares values of daily flow predictions using the MOVE3 method with the 12 observed values for two of the sites which had the highest NSE values and for two of the sites which had the lowest NSE values.



MOVE evaluation, site 14
ρ 0.86, NSE -0.04

MOVE evaluation, site 9
ρ 0.89, NSE -0.18

MOVE evaluation, site 4
ρ 0.95, NSE 0.93

MOVE evaluation, site 15
ρ 0.99, NSE 0.95

Fig. 2 Comparison of estimated daily streamflow using the MOVE3 method with observed daily streamflow for the two strongest and two weakest cases with reported Pearson´s correlation coefficient (ρ) and Nash-Sutcliffe Efficiency (NSE)

The index gages are generally located on larger rivers than the study sites thus one might expect quite different hydrologic behaviour among these rivers, yet we were able to find a suitable index gage with sufficient correlation (ρ above 0.8) for 15 small watersheds out of 20 considered in this study. Perhaps this is in part due to the fact that the national monitoring network of stream gages is relatively dense and operates in all regions of the Czech Republic.

Regression models of daily phosphorus loads

The primary goal of this study is to develop reliable estimates of annual TP loads for all of the study sites considered in this study for the study period. Regression models for estimation of TP loads were developed for each site using the streamflow and TP load observations during 2007 and 2008. Table 2 summarizes the models for each site including independent variables included, adjusted R^2, standard errors of estimate in %, and the final regression equations. The p-values for all model coefficients were always below 0.01 except for a few values of the seasonal coefficients which whose p-value fell between 0.01 and 0.05. The adjusted R^2 values range from 0.60 to 0.98 with values above 0.9 in more than half of the cases. In all cases, the model residuals ε were found to be approximately normally distributed.

Table 2 Regression models for daily total phosphorus loads in (mg day^{-1}) for each study site as a function of daily discharge Q (dm^3 s^{-1}) and Julian day T.

Site	Variables in the model	R^2	adj.R^2	SEE%	regression equation
1	lnQ, cos(ωT)	0.893	0.884	73.3	lnL = 9.311 + 0.993 lnQ - 0.814 cos(ωT) + ε
2	lnQ, cos(ωT)	0.746	0.731	57.3	lnL = 7.575 + 1.143 lnQ - 0.745 cos(ωT) + ε
3	lnQ	0.953	0.950	56.3	lnL = 7.692 + 0.912 lnQ + ε
4	lnQ, cos(ωT)	0.787	0.769	59.7	lnL = 8.266 + 1.101 lnQ - 0.482 cos(ωT) + ε
5	lnQ, sin(ωT), cos(ωT)	0.825	0.801	35.8	lnL = 7.373 + 1.254 lnQ - 0.421 cos(ωT) - 0.313 sin(ωT) + ε
6	lnQ, T	0.634	0.600	43.3	lnL = 7.600 + 2.153 lnQ + 0.002 T + ε
7	lnQ	0.738	0.728	76.1	lnL = 7.349 + 1.091 lnQ + ε
8	lnQ, sin(ωT)	0.924	0.911	44.2	lnL = 7.421 + 1.176 lnQ - 0.502 sin(ωT) + ε
9	lnQ, cos(ωT)	0.923	0.915	76.0	lnL = 7.904 + 0.965 lnQ - 0.708 cos(ωT) + ε
10	lnQ	0.955	0.953	53.0	lnL = 7.492 + 0.988 lnQ + ε
11	lnQ, sin(ωT)	0.952	0.947	31.3	lnL = 8.442 + 1.247 lnQ - 0.273 sin(ωT) + ε

12	lnQ, sin(ωT)	0.960	0.953	56.5	$\ln L = 9.213 + 1.222 \ln Q - 0.594 \sin(\omega T) + \varepsilon$
13	lnQ	0.952	0.949	32.9	$\ln L = 9.094 + 1.045 \ln Q + \varepsilon$
14	lnQ, sin(ωT)	0.985	0.984	12.2	$\ln L = 9.255 + 1.061 \ln Q - 0.112 \sin(\omega T) + \varepsilon$
15	lnQ	0.668	0.652	73.4	$\ln L = 8.034 + 0.757 \ln Q + \varepsilon$

Fig. 3 illustrates two examples of the relationship between the logarithms of daily TP load and the logarithms of daily streamflow at two sites in which streamflow was the only independent variable chosen but which exhibit a wide range of goodness of fit. These two cases (sites 3 and 15) correspond to a range of R^2 equal 0.67 and 0.95.

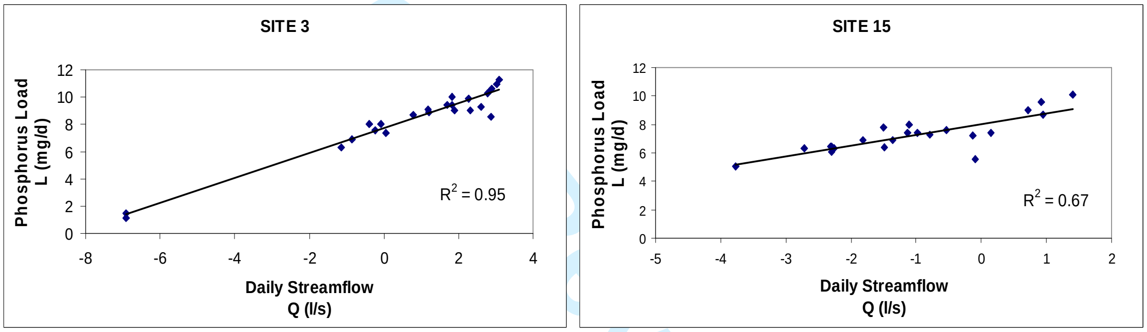


Fig. 3 Relationship between daily streamflow Q and daily phosphorus load L (2007-2008) at two sites with high and low goodness of fit.

Estimation of Annual Phosphorus Loads

The regression equations reported in Table 2 were then used to compute daily TP loads given the extended records of mean daily streamflows generated using the MOVE3 methodology. Daily TP loads were corrected for bias and summed into annual loads. In some cases, summer droughts resulted in periods of zero streamflows at a few study sites which poses a challenge because a logarithmic transformation is used in the TP load regressions. Therefore annual loads were corrected by subtraction of daily loads which fell on days with observed zero streamflows.

Annual TP loads ranged from 0.21 kg year^{-1} to 95 kg year^{-1} and corresponding annual TP yields ranged from 0.01 to 0.66 kg ha^{-1}year^{-1}. Although we are aware of the importance of rainstorm events in phosphorus transport to streams and possible underestimating of loads, we decided for the purpose of this study not to include records from additional sampling which was simultaneously carried out at study sites using passive automatic samplers or co-operating staff in the field because this could result in a positive bias in annual load estimates as described previously. Results obtained by additional sampling during extreme events will be utilized in ongoing research to quantify the possible bias in annual TP load estimates, to capture seasonal variability in TP loads and to identify critical periods for watershed management.

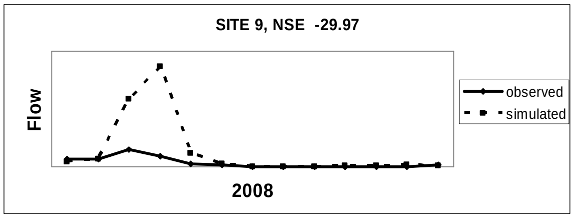
Model cross validation

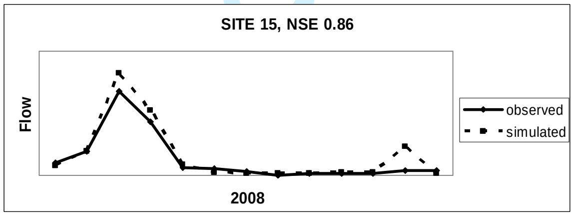
Here we describe the use of a split-sample cross validation experiment for evaluating the credibility of resulting TP load estimates. Performance of regression based estimates of TP loads and MOVE methods for streamflow record extension were validated by using data from 2007 to estimate model parameters and those fitted models were than used to estimate streamflow and TP load values for 2008. Estimated values were then compared with actual values measured in 2008 and Nash-Sutcliffe Efficiency (NSE) is reported for each site in tab 3.

Table 3 Nash-Sutcliffe Efficiency (NSE) for MOVE and regression cross-validation of streamflow and phosphorus load estimates for 2008

Site	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
NSE (streamflow)	-21.72	-0.89	0.78	0.80	0.74	0.33	0.27	0.12
NSE (P loads)	0.11	0.18	0.83	0.24	0.58	-3.2	0.19	0.35
Site	**9**	**10**	**11**	**12**	**13**	**14**	**15**	
NSE (streamflow)	-29.97	-11.46	-4.25	0.59	0.62	0.72	0.86	
NSE (P loads)	0.50	0.45	0.87	0.75	0.73	0.93	0.65	

Values of NSE ranged from -29.97 to 0.86 with a mean value of 0.33 for the MOVE method for streamflow estimates and from -3.2 to 0.93 with a mean value of 0.5 for regression method for the TP load estimates. Fig. 4 compares values of daily flow computed using the MOVE3 method with actual values measured in 2008 and Figure 5 compares values of TP load predictions using the regression methods with actual values measured in 2008. The cases which had the highest, the lowest and the average NSE values are reported.





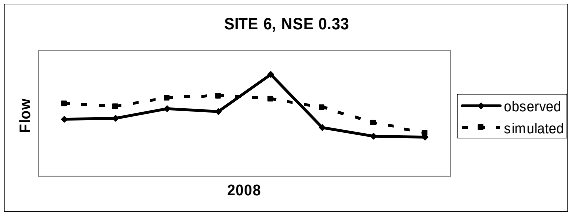


Fig. 4 Comparisons of estimated daily streamflow with observed daily streamflow for the strongest, the weakest and average cases with reported Nash-Sutcliffe Efficiency (NSE) of the MOVE3 method

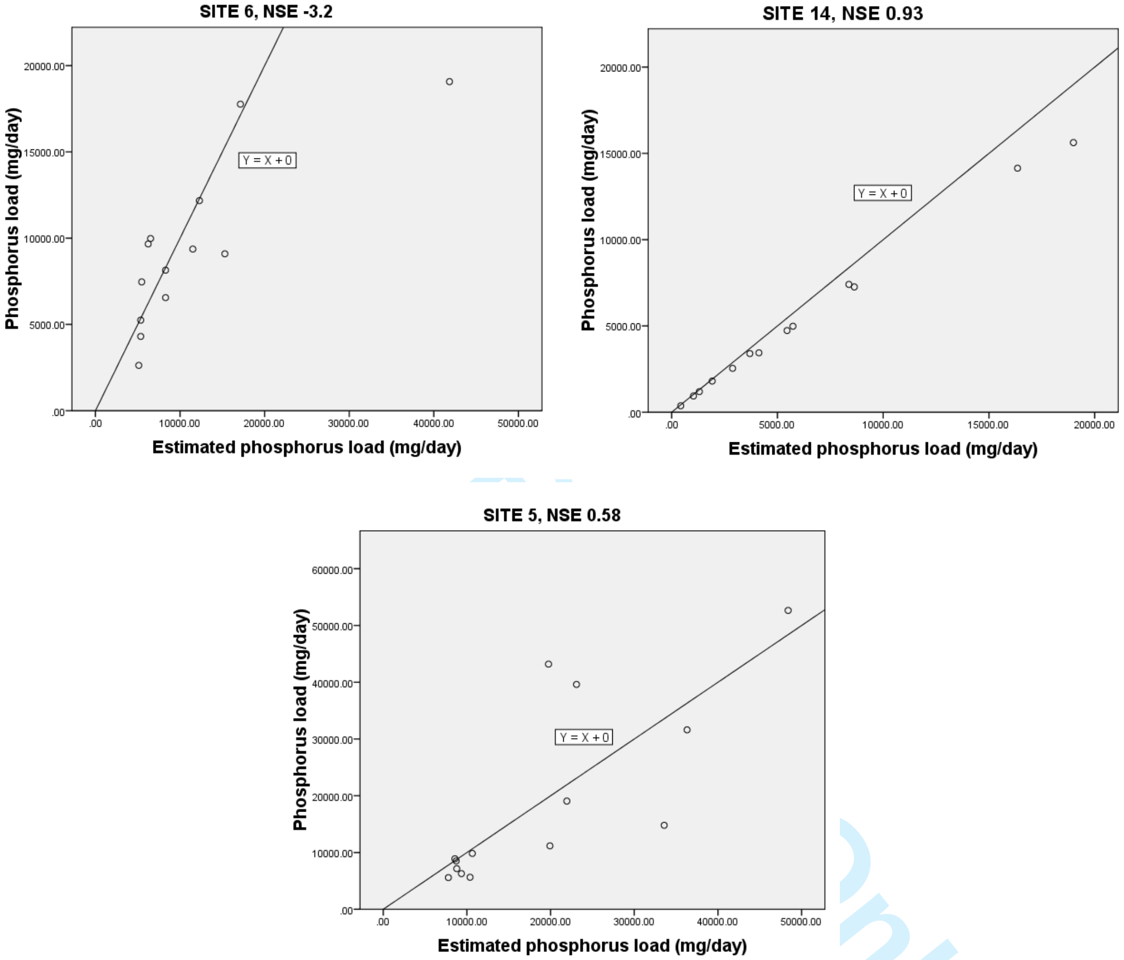


Fig. 5 Comparisons of estimated daily phosphorus loads with observed phosphorus loads for the strongest, the weakest and the average cases with reported Nash-Sutcliffe Efficiency (NSE) of the regression method

Results show relatively good model performance of TP load prediction and streamflow extension even for models based on a single year of observations. However, performance of model predictions is expected to increase with the length of records. In some cases seasonal time parameters entered the regression model based on only a single year of data, however we elected to exclude these parameters because they are likely to reduce the accuracy of resulting TP load prediction. This is in accordance with Haggard et al. (2003) who discussed the effect of seasonal factors on prediction accuracy and suggested sufficient repetition of seasonal cycles when using seasonal time parameters in regression models. The regression model of TP load gives reliable estimates within the range of measured streamflow used for its development. Extrapolation beyond this range may influence the accuracy of predictions. The range of sampled streamflow used to develop the regressions corresponds to 95% of all daily streamflow values estimated for the study period. The remaining 5% (approximately 18 days per year) correspond to extreme events of high streamflows usually associated with soil erosion. Such extreme runoff with soil erosion is not fully captured by this approach, leading to potential underestimation of the annual TP load by as much as 50% or more (see Kalff, 2002). Thus the remaining 5% of daily TP loads corresponding to such extreme events should be considered separately taking into account individual watershed characteristics and results obtained by sampling during extreme events.

CONCLUSIONS

In this study we illustrate how short records of daily streamflow and phosphorus concentrations can be combined with a national network of daily streamflow records in the Czech Republic in order to derive meaningful estimates of annual phosphorus loads for small agricultural watersheds. Importantly, we perform a cross-validation experiment which tested the ability of the resulting methodology to estimate annual phosphorus loads during a time period which was not used in the development of the model used to predict streamflows and/or phosphorus loads. Such cross-validation experiments are essential to gaining an understanding of the credibility of a modelling approach. We expect other studies similar to ours could benefit by exploiting regional streamflow information for extension and/or augmentation of streamflow records in the Czech Republic and possibly elsewhere.

Acknowledgements

This study was funded by Ministry of the Environment of the Czech Republic through Institutional research plan MZP0002071101 “Research and Protection of the Hydrosphere – research of relations and processes in the aquatic component of the environment, with particular focus on the impact of anthropogenic pressures and on the sustainable use and protection, including legislative”. Special appreciation is extended to Hlavka foundation, Czech Ministry of Education, Youth and Sports through project ME 897 “Landscape-architectural principles of rural landscape restoration“, and Tufts University for enabling international cooperation. We would also like to thank Dr. Jitka Brzakova and Czech Hydrometeorological Institute for data provision and to Dr. Petr Sklenička and Dr. Ondrej Beranek for their support.

REFERENCES

Central Institute for Supervising and Testing in Agriculture (2009) Vývoj základních agrichemických vlastností zemědělských půd v České republice. Retrieved May 07, 2009, from http://www.zeus.cz/Articles/8931-2-Agrochemicke+vlastnosti.aspx.

Clarke, R. T. (1990) Statistical Characteristics of Some Estimators of Sediment and Nutrient Loadings. *Water Resources Research* **26**(9), 2229 - 2233.

Cohn, T. A. (1995) Recent Advances in Statistical-Methods for the Estimation of Sediment and Nutrient Transport in Rivers. *Reviews of Geophysics* **33,** 1117-1123.

Correll, D. L. (1998) The Role of Phosphorus in the Eutrophication of Receiving Waters: A Review. *Journal of Environmental Quality* **27**, 261-266.

Ferguson, R. I. (1986) River Loads Underestimated by Rating Curves. *Water Resources Research* **22**(1), 74-76.

Grygier, J. C., Stedinger J. R. & Yin, H. B. (1989). A Generalized Maintenance of Variance Extension Procedure for Extending Correlated Series. *Water Resources Research* **25**(3), 345-349.

Haggard, B. E., Soerens, T. S., Green, W. R. & Richards, R. P. (2003) Using regression methods to estimate stream phosphorus loads at the Illinois River, Arkansas. *Applied Engineering in Agriculture* **19**(2), 187-194.

Hansen, N. C., Daniel, T. C., Sharpley, A. N. & Lemunyon, J. L. (2001) The fate and transport of phosphorus in agricultural systems. *Journal of Soil and Water Conservation* **57**(6), 408-417.

Hart, M. R., Quin, B. F. & Nguyen, M. L. (2004) Phosphorus runoff from agricultural land and direct fertilizer effects: A review. *Journal of Environmental Quality* **33**(6), 1954-1972.

Helsel, D. R. & Hirsch R. M. (1992) Statistical methods in water resources. *Studies in Environmental Science* **49**. Elsevier, Amsterdam, The Netherlands.

Hirsch, R. M. (1982) A Comparison of 4 Streamflow Record Extension Techniques. *Water Resources Research* **18**(4), 1081-1088.

Kalff, J. (2002) *Limnology*. Benjamin-Cummings Publishing Company, San Francisco, CA, USA.

Kenney, B. C. (1982) Beware of Spurious Self-Correlations. *Water Resources Research* **18**(4), 1041-1048.

Klement, V. & Sušil, A. (2005) *Porovnání vývoje agrochemických vlastností půd za období 1993-1998 a 1999-2004.* Central Institute for Supervising and Testing in Agriculture, Brno, Czech Republic.

Klement, V. & Sušil, A. (2009) *Porovnání vývoje agrochemických vlastností půd za období 1999-2002 a 2005-2008.* Central Institute for Supervising and Testing in Agriculture, Brno, Czech Republic.

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

Robertson, D. M. (2003) Influence of different temporal sampling strategies on estimating total phosphorus and suspended sediment concentration and transport in small streams. *Journal of the American Water Resources Association* **39**(5), 1281-1308.

Robertson, D. M. & Roerish E. D. (1999) Influence of various water quality sampling strategies on load estimates for small streams. *Water Resources Research* **35**(12), 3747-3759.

Salas, J. D. (1993) Analysis and modeling of hydrologic time series. In: *Handbook of Hydrology* (ed. by D. R. Maidment) 19.1-19.72. McGraw-Hill**.**

Toor, G. S., Harmel, R. D., Haggard, B. E. & Schmidt, G. (2008) Evaluation of regression methodology with low-frequency water quality sampling to estimate constituent loads for ephemeral watersheds in Texas. *Journal of Environmental Quality* **37**(5), 1847-1854.

Vogel, R. M., Rudolph, B. E. & Hooper, R. P. (2005) Probabilistic Behavior of Water-Quality Loads. *Journal of Environmental Engineering* **131,** 1081-1089.

Vogel, R. M. & Stedinger J. R. (1985) Minimum Variance Streamflow Record Augmentation Procedures. *Water Resources Research* **21**(5), 715-723.