DE GRUYTER OPEN

Statistical approach for the estimation of watershed scale nitrate export: a case study from Melen watershed of Turkey

Muhammed Ernur Akiner[1*], Atilla Akkoyunlu[2]

[1]Akdeniz University, Turkey
Vocational School of Technical Sciences, Environmental Protection and Control
[2]Bogazici University, Turkey
Department of Civil Engineering

*Corresponding author's e-mail: ernurakiner@yahoo.com

Abstract: Nutrient pollution such as nitrate (NO_3^-) can cause water quality degradation in rivers used as a source of drinking water. This situation raises the question of how the nutrients have moved depending on many factors such as land use and anthropogenic sources. Researchers developed several nutrient export coefficient models depending on the aforementioned factors. To this purpose, statistical data including a number of factors such as historical water quality and land use data for the Melen Watershed were used. Nitrate export coefficients are estimates of the total load or mass of nitrate (NO_3^-) exported from a watershed standardized to unit area and unit time (e.g. $kg/km^2/day$). In this study, nitrate export coefficients for the Melen Watershed were determined using the model that covers the Frequentist and Bayesian approaches. River retention coefficient was determined and introduced into the model as an important variable.

Introduction

Protection of water sources needs prior research on the determination of possible sources of pollution. To this purpose, using the Frequentist and Bayesian approaches nitrate export coefficient modeling of the Melen Watershed is dealt with during this research. Modelling the export coefficients is a convenient way to analyze the effects of diffuse pollution in a research area. Moreover, refering to Fu (2012), retention of nutrients in the water body was also taken into consideration. Export coefficients are usually determined with help of load measurements at an outlet of a subwatershed where there is a single dominant land use (Brigault and Ruban 2000, Zobrist and Reichert,2006). In order to estimate the export coefficients it is assumed that the export coefficients for the same land use category are the same in all subwatersheds.

Notable researchers regarding the export coefficients such as Rast and Lee (1983), Dillon and Kirchner (1975) used the same/similar approach. They observed concentrations of nutrients in streams (e.g. mg/L) then converted that value to load by multiplying with discharge (e.g. m^3/s). Usually, they had daily flow rates and monthly or quarterly nutrient concentrations that are interpolated to account for days. De Klein and Koelmans (2011) calculated inputs to surface water and exports for 13 lowland river catchments in Western Europe, on a monthly basis. The catchments varied in size (21 to 486 km^2), while annual in-stream retention ranged from 23 to 84% for N. A novel calculation method is presented that quantifies monthly exports from lowland rivers based on an annual load to the river system. The agreement between calculated values and calibration data was high (N: $r^2 = 0.93$; $p < 0.001$). Validation of the model also showed good results with model efficiencies for the separate catchments ranging from 31 to 95% (average 76%). This indicates that exports of nitrogen on a monthly basis can be calculated with few input data for a range of West European lowland rivers (De Klein et al. 2011). Wickham et al. (2005) found that land cover is the main driver of nutrient export, and regional variation is insignificant. Wickham et al. (2008) noticed that the variances of N and P concentrations among different land uses within ecological regions were respectively six and three times greater than the variance among different ecological regions. Though, an effect of different ecological regions is less significant compared to different land use compositions. Studies on distinct geographical regions definitely lead to different results. Most of the studies that deal with nutrient export were conducted in the USA. Current research could extend it to Europe.

There are also some analytical models to estimate the pollutant removal efficiency for surface waters. Yang et al. (2014) successfully applied a screening-level modeling approach to estimate nitrogen loading in Tippecanoe River watershed. Screening-level is not statistical but an advanced analytical model consists of both hydrological and water quality parameters. Although screening-level model is a very comprehensive approach, it is not very favorable if there is no wide range of spatial and temporal observed data for a variety

of hydrological and water quality parameters. However statistical methods may provide reliable results closest to analytical ones without needing complicated data. A typical example of this situation is the research of Zobrist and Reichert (2006) in Swiss Watersheds, where they applied the Bayesian approaches successfully to estimate nitrate export coefficients.

The Frequentist and the Bayesian techniques were applied in order to estimate export coefficients for the Melen Watershed in Turkey. Instead of calculating the contributions of subwatersheds individually, whole watershed was considered for the estimation of total load at the outlet of the Melen Watershed using the calculated nutrient export coefficients. The frequentist approach has the goal of extracting information from data only, without relying on the prior knowledge (Ramirez and Sanz, 2013). In contrast to the Frequentist approach, the Bayesian approach has the goal of combining prior knowledge with data to optimally use both sources of information. Success of the Bayesian approach is directly proportional to the sufficiency of data for acquiring the prior information about estimands (Essahale et al. 2010, Lee 2012). The Bayesian analysis was conducted using the IBM SPSS AMOS software and posterior information about land use based export coefficients was obtained through the MCMC (Markov Chain Monte Carlo) method. Estimated land use based nutrient export coefficients are in kg/km^2/day unit. In addition, monthly river retention values of nitrate in all subwatersheds of the Melen Watershed were estimated. This information was used in order to predict nitrate export coefficients appropriately. The results show that the frequentist approach gives closer estimates to the observed values compared to the Bayesian approach.

Study Area

The Melen Watershed is located in Western Black Sea region of Turkey (Figure 1). It has the 2437 km^2 area (Ozturk et al. 2008). The Melen Watershed is bounded by the Bolu Mountains to the east, the Sakarya Province to the west, the Orhan Mountains to the north, and the Abant Mountains to the south. The Melen Watershed provides fresh drinking water to most of Istanbul.

According to the Ministry of Forestry and the Water Works of Turkey, regarding the pollution status, the Melen River Basin should be dealt with primarily.

The Melen River Basin is currently under the threat of land based pollution. Sumer et al. (2001) revealed in their research that its water can be classified as water class number 2 out of 4. Since 2001, settlements and the population in the watershed have increased. As far as it is known there are no agricultural or urban best management practices applied in the region. Therefore, a significant decrease in the water quality of the river in the future is expected.

Two main rivers are located in the Melen Waterhed. These are the Buyuk Melen and the Kucuk Melen rivers (Figure 1). The government constructed a water regulator close to the outlet of the Buyuk Melen River. Fresh water is pumped to Istanbul with a 150 km long pipe. Protection of water quality in the Melen Watershed is also vital for Istanbul's drinking water quality.

In Figure 1, red bullets indicate sampling sites for data gathered from the State Hydraulic Works of Turkey (DSI, 2011). This data covers crucial information about the historical trend of the pollution in the Melen Watershed. Also, pink bullets indicate sampling sites for data measured by İstanbul Technical University (Ozturk et al. 2008) from different locations in the Melen Watershed, including headwater subwatersheds.

Materials and Methods

The simplest export coefficient model assumes that average diffuse loads can be estimated by a sum of export coefficient terms regarding the different specific land use types (Zobrist and Reichert 2006). Measurements in Melen Watershed were conducted weekly, sometimes daily as long as the weather conditions were suitable, or twice a month. Hence monthly average values were calculated using the mean value of these measurements. These values are the "monthly average daily values". Temporally averaged (monthly average daily values) loads from a number of subwatersheds can be quantified using Equation (1):
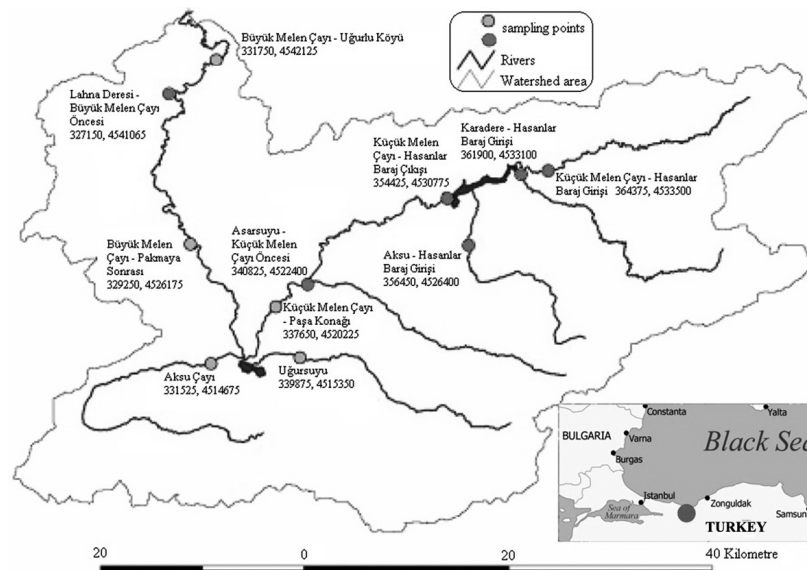

Fig. 1. The Melen Watershed, its rivers, and coordinates of sampling points in WGS84 Datum UTM coordinate system 36N

$$L_{i,j} = \sum_{j=1}^{m} \sum_{i=1}^{n} \left(E_{i,j} \cdot A_{i,j} \right) \cdot \left(1 - R_j \right) \tag{1}$$

Where j is the index labelling the subwatershed, i is the index labelling the land use types (10.66% meadows and pastures, 0.66% lakes and rivers, 36.02% agricultural area, 51.25% forest, and 1.46% urban area), $L_{i,j}$ is the average load at the outlet of subwatershed j as predicted by the model, $E_{i,j}$ is the nutrient export coefficient of i land use type, $A_{i,j}$ is the area of the i land use in j^{th} subwatershed (Figure 2), R is the percent river nutrient retention coefficient.

The watershed is delineated into discrete subwatersheds for enabling the modeling to represent the spatial heterogeneity in the catchment. The delineation of the Melen Watershed was carried out based on a Digital Elevation Model (DEM) created in 10 m. × 10 m. resolution by both digitizing topographical map sheets and modifiying the available vector maps. Created DEM was imported to Arcview grid format with proper projection (UTM – Zone 36 N – WGS84 Datum).

Each subwatershed has a contribution on the total load at the outlet of the Melen Watershed. In order to quantify their contribution, flow path of the whole watershed has to be specified. Figure 3 shows the flow path or the direction of the flow at the watershed.

The retention and loss of nutrients in river systems can have significant detrimental consequences on downstream water quality (Donohue et al. 2005, Hogan et al. 2012, Vsetickova et al. 2012, Izagirre et al. 2013). Peterson et al. (2001) examined the nitrogen removal efficiencies of headwater streams from all over the United States. They found that the smaller the stream (lower the order), the higher the efficiency for removal of nitrogen (N) is. This is because the water is in greater contact with various biofilm surfaces in smaller streams. On
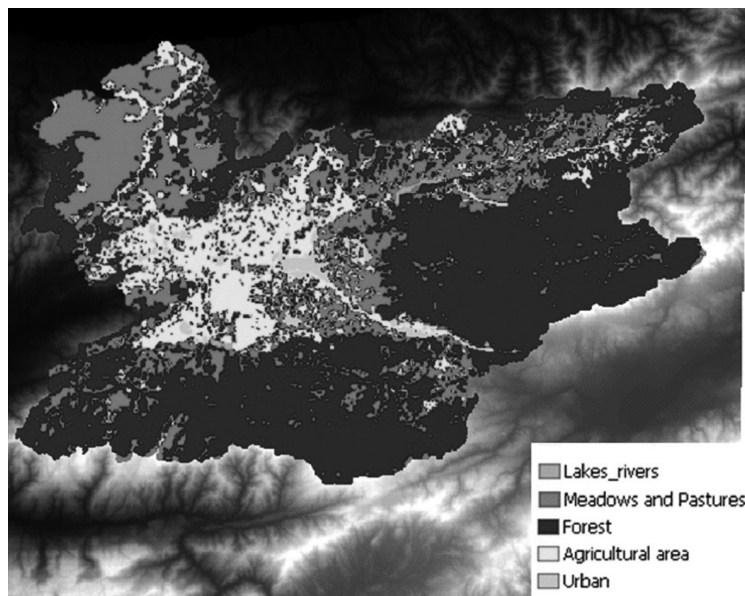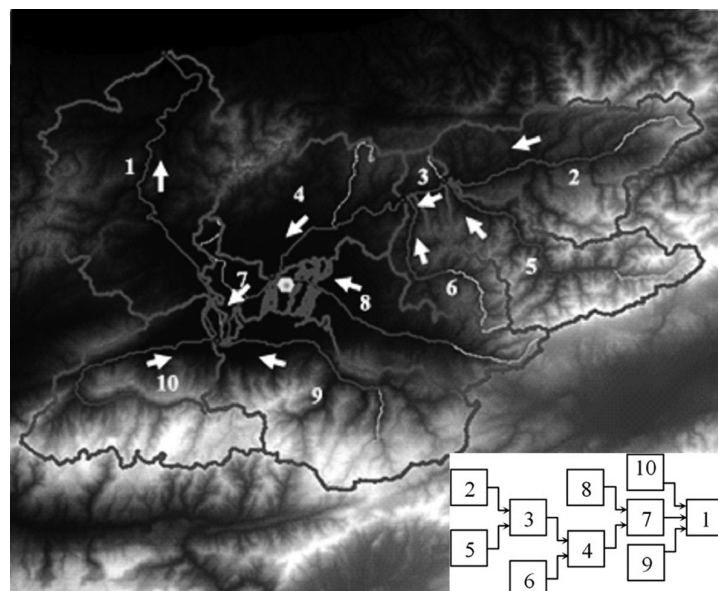

Fig. 2. Land use map for the Melen Watershed

Fig. 3. Subwatersheds, rivers, and flow path of the Melen Watershed

average, dissolved inorganic nitrogen (both NH_4^+ and NO_3^-) is removed at a rate of 64% per kilometer of a stream. The small size of the stream ensures a large amount of water-sediment contact, which removes nitrogen from runoff via nitrification and denitrification by bacteria in the sediments (Peterson et al. 2001).

According to De Klein and Koelmans (2011), monthly retention of nitrogen is estimated from surface water area specific runoff (Equation (2)). Nitrate retention was calculated using the Equation (2) assuming that percent retention of NO_3^- is almost equal to percent retention of total N in the rivers of Melen Watershed.

$$R_i = 0.0246 \left(\frac{Q_i}{SW} \right)^{-0.57} \qquad (2)$$

Where Q_i is the average (monthly) discharge (m³/s), SW is the total area of surface water in the catchment (ha), R_i is the retention fraction (-), i is the index for month (-).

Statistical methods are used to do predictions that are as close as possible to the observed values. One of the most preferred methods of parameter estimation for distribution fitting is the Maximum Likelihood Method (Law and Kelton 1991). Maximum Likelihood Estimation (MLE) seeks the parameter values that are most likely to produce the observed distribution (Gardner 2012, Meer et al. 2013, Nichols et al. 2013). The basic goal of using MLE is to determine the parameters that maximize the probability or likelihood of the sample data.

MLE consists of three steps. These are: specifying the likelihood function, taking derivatives of likelihood with respect to the parameters, setting the derivatives equal to zero and solving for the parameters.

When assuming normally and independently distributed stochastic errors that add to the results of the deterministic function given by the Equation (1), the likelihood functions of the export coefficient model for loads become (Equation (3)):

$$f(x \mid E, \sigma_x^2) = \frac{1}{\sigma_x^n \sqrt{2\pi}^n} \exp\left\{ -\sum_{j=1}^{n} \frac{(x - L_j(E))^2}{2\sigma_x^2} \right\} \qquad (3)$$

Where $x = (x_1, \ldots, x_n)$ is the vector of loads of the watershed, σ_x is the standard deviation of the normal distribution of loads or concentrations around the deterministic model results, and L_j is the average load at the outlet of the watershed according to the Equation (1).

In the frequentist approach, the parameter estimates are determined by maximization of the likelihood function shown in the Equation (3) into which measurements are substituted for the argument describing the outcomes (Equation (4)):

$$\left(\hat{E}, \hat{\sigma}_x^2 \right) = \underset{E, \sigma_X^2}{\mathrm{argmax}} \, f_X\left(x \mid E, \sigma_X^2 \right) \qquad (4)$$

Where *argmax* stands for the argument of the maximum, that is to say, the set of points of the given argument for which the value of the given expression attains its maximum value. In this equation, $\hat{E}$ are the estimates of the export coefficients E

as introduced in Equation (1), $\hat{\sigma}_x^2$ is the estimate of the variance of the additive stochastic error term σ_X^2, f_x is the likelihood function of the model, and x values are the measured loads at the watershed outlet.

The frequentist approach requires a well-defined maximum of the likelihood function in order to provide unique results (Cowan 2012). For the normal distribution; N(μ,σ²) or N(E,σ²) has probability density function and the likelihood as seen in Equations (5)-(6) respectively. In Equation (6) $\bar{x}$ is the sample mean.

$$f(x \mid E, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{ -\frac{(x - L(E))^2}{2\sigma^2} \right\} \qquad (5)$$

$$f(x_1, \ldots, x_n \mid E, \sigma^2) = \prod_{i=1}^{n} f(x_i \mid E, \sigma^2) =$$

$$= \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ -\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{2\sigma^2} \right\} \qquad (6)$$

This distribution has two parameters (E, σ²), so we can maximize the likelihood over both parameters, $L(E, \sigma) = f(x_1, \ldots, x_n \mid E, \sigma)$. Logarithm is continuously increasing function over the range of likelihood. The values which maximize the likelihood will also maximize its logarithm (Cowan 2012). Maximizing logarithm requires less algebra. Log likelihood is differentiated with respect to E and equated to zero as follows (Equation (7)). Similarly we differentiate the log likelihood with respect to σ and equate to zero (Equation (8)).

$$\frac{\partial}{\partial E} \ln\left[\left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ -\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{2\sigma^2} \right\} \right]$$

$$= \frac{\partial}{\partial E}\left[\ln\left(\frac{1}{2\pi\sigma^2} \right)^{n/2} + \left\{ -\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{2\sigma^2} \right\} \right] = 0 \qquad (7)$$

$$\frac{\partial}{\partial \sigma} \ln\left[\left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ -\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{2\sigma^2} \right\} \right]$$

$$= \frac{\partial}{\partial \sigma}\left[\ln\left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} + \left\{ -\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{2\sigma^2} \right\} \right] \qquad (8)$$

$$= -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - L(E))^2}{\sigma^3} = 0$$

Maximum likelihood estimator calculated from the above two derivatives for $\theta = (E, \sigma^2)$ is symbolized as seen in Equation (9):

$$\hat{\theta} = \left(\hat{E}, \hat{\sigma}^2\right) \qquad (9)$$

In contrast to the Frequentist approach, the Bayesian approach has the goal of combining prior knowledge with data to use both sources of information optimally (O'Reilly et al., 2012). Prior knowledge on parameter values has to be formulated as a prior probability density, $f_{prior}(E, \sigma_x^2)$, and is then updated to the posterior density by applying Equation (10). The constant of proportionality is calculated by normalization of the posterior density. This technique has the advantage of still being applicable if the parameters are not identifiable from data. In the case of poor identifiability, the posterior distribution is not much different from the prior (Gelman 2006, Morris et al. 2012). In the case of high information content of data, it is typically much narrower. The disadvantage of this technique is that use of prior information introduces a subjective element into data evaluation procedure.

$$f_{post}\left(E, \sigma_x^2 | x\right) \propto f_x\left(x | E, \sigma_X^2\right) \cdot f_{prior}\left(E, \sigma_x^2\right) \qquad (10)$$

In the Frequentist inference, any given experiment is considered as one of an infinite sequence of possible repetitions of the same experiment with statistically independent results (Everitt, 2006). The independent and identically distributed observations $(x_1, ..., x_n)$ come from the sampling distribution $f(X | \theta)$ where θ is the fixed parameter value.

The MCMC is a simulation technique that computes posterior values of interest by sampling from posterior distributions (Huang and Yu 2010, Konomi et al. 2013). The Bayesian posterior distribution is obtained by the MCMC method. This is beneficial for multi-parameter models where it is hard to have algebraic solution. The MCMC algorithms are computational tools that allow for the generation of random numbers from the posterior distribution $\pi*(\theta | X)$ using the numerator of the expression in Equation (11) without calculating the integral in the denominator (Lele et al. 2007).

$$\pi*(\theta | X) = \frac{f(x_1, x_2, ..., x_n | \theta)\pi(\theta)}{\int f(x_1, x_2, ..., x_n | \theta)\pi(\theta)d\theta} \qquad (11)$$

Where $f(x_1, x_2, ..., x_n | \theta)$ is the likelihood, $\pi*(\theta | X)$ is the posterior, and $\pi(\theta)$ is the prior distribution.

Parameter estimation depends on the frequency distribution of data. The main goal is to find which distribution fits data better. For this aim, the Kolmogorov-Smirnov's goodness of fit test was performed, since it is applicable even for a sample consisting of a small number of observations. According to the results of this test, best-fitting distributions were identified. Additionally, there should be a single term (e.g. mean or μ) that consists of all desired parameters (Forest E_{For}, agricultural E_{Agr}, meadows E_{Mea} and residential area (E_{Res}) export coefficients) to take its derivative. Hence, it is not necessary to deal with a distribution that has complex multiplicative terms. While selecting the appropriate distribution for the Frequentist

Approach, this situation was also taken into consideration. Setting the derivative of the likelihood function with respect to a single term to zero and solving for the unknown term leads to four equations with the aforementioned four unknown parameters. In this study, solutions for four linearly independent equations for four unknowns were obtained using the Direct Search optimization package of the Maple 15 Pro.

Results and Discussion

According to the Kolmogorov-Smirnov's goodness of fit test and the consideration above, the frequency distribution of the nitrate load data at the outlet of the Melen Watershed is defined as the Inverse Gaussian. The Inverse-Gaussian distribution of x_i is described by two characteristics, a mean $\mu > 0$ and precision $\lambda > 0$. See Equation (12) for the probability density function.

$$f(x; \mu, \lambda) = \left[\frac{\lambda}{2\pi x^3}\right]^{1/2} \exp\frac{-\lambda(x - \mu)^2}{2\mu^2 x}, \ 0 < x < \infty \quad (12)$$

Where $\mu > 0$ is the mean. $\lambda > 0$ is the shape parameter, variance is given by μ^3 / λ.

Maximum likelihood is the Frequentist but it is also a part of the Bayesian inference. Table 1 shows estimations for NO_3^- export coefficients (kg/km²/day) using the Frequentist approach as the mean monthly average daily value for each year. By applying the procedure shown in Equations (5)–(9), considering that the distribution is the inverse Gaussian, unknown export coefficient parameters can be calculated. For each year between 1995 and 2006, export coefficients for different land use types were calculated.

Estimated nutrient export coefficients are in kg/km²/day unit. In other words, every day, export coefficient times kg of nitrate load is emitted per km². Retention in the water body was taken into account for the estimation of export coefficients. Briefly, values in Table 1 multiplied by the corresponding land use area and (1 – retention coefficient) give us the daily nitrate load estimation in kg per km². Overall estimation for nitrate load was calculated using mean value of the predicted export coefficient values (Mea=0.759; Agr=2.749; For=0.606; Res=1.678).

The key issue that differs the Bayesian estimation from the Frequentist approach is to use the prior information about estimands. Headwater subwatersheds are not affected by other subwatersheds. Headwater subwatersheds of the Melen Watershed are subwatersheds 2, 5, 6, 8, 9 and 10 (see Figure 3). It is necessary to use the observed data from headwater subwatersheds in order to define prior distributions of the land use based export coefficients. This prior information helps us to see what is the distribution of export coefficient frequencies, what is their mean, standard deviation, etc. Sometimes use of high level prior information is crucial. To this purpose, the usual method of getting this prior information is to have sampling stations in such an area where a single land use is dominated. More precisely, if it is required to observe a prior distribution for agricultural area nutrient export coefficient (Agr or E_{Agr}), we need to sample in an area that is agriculturally dominated.

First of all, the observed data from headwater subwatershed 6 were analyzed since forest area is dominated (91.15%) in this subwatershed. After getting information for forest area nutrient export coefficient (E_{For}), data from headwater subwatersheds

www.czasopisma.pan.pl   PAN   www.journals.pan.pl

Statistical approach for the estimation of watershed scale nitrate export: a case study from Melen watershed of Turkey   49

Table 1. Estimated NO_3^- export coefficients (kg/km²/day) using frequentist approach

	Whole watershed – NO_3^- – Export coefficient (kg/km²/day). Mean "monthly average daily" values for each year			
Year	Meadows Brush and Pasture	Agricultural	Forest	Residential
1995	0.595	2.687	0.505	1.901
1996	0.508	2.084	0.405	1.206
1997	0.708	3.544	0.612	1.995
1998	0.710	3.140	0.554	1.926
1999	0.608	2.033	0.400	1.257
2000	1.220	3.331	0.950	1.968
2001	0.705	2.117	0.605	1.106
2002	0.826	3.077	0.700	1.966
2003	0.804	3.000	0.615	1.903
2004	1.101	3.125	0.815	1.844
2005	0.603	2.182	0.505	1.410
2006	0.726	2.673	0.600	1.653

10, 2 and 8 were consecutively analyzed in order to specify agricultural (E_{Agr}), meadows (E_{Mea}) and residential area (E_{Res}) nutrient export coefficients, respectively. Please notice that the observed data available from these subwatersheds were in a sufficient amount only for nitrate parameter. Thus prior distributions were created for the nitrate export coefficients of each type of land use (Mea, Agr, For, Res). Then the Bayesian estimation was able to start.

Using the IBM SPSS AMOS software (Arbuckle 2009), the Bayesian analysis was conducted and posterior information about land use based nitrate export coefficients was obtained using the MCMC method. Single value for each land use based nitrate export coefficients was estimated using whole monthly data from January 1995 to December 2006. In the AMOS, parameter configuration for the analysis is important (Chenini and Khemiri 2009, Loehlin 2013). During the Bayesian analysis, nitrate loads exported from each type of land use (Mea, Agr, For, Res) were selected as independent variables and total nitrate exported (NO_3^-) at the outlet in kg/km²/day was selected as a dependent variable. Dagum, Gamma, Kumaraswamy and Wakeby distributions were encountered during the Bayesian analysis phase of this study.

Prior information is not always very informative. It is a must to have good agreement between the Frequentist approach and the Bayesian approach with a non-informative prior. Figure 4 shows that the assigned priors are highly informative for the Bayesian estimation. The Bayesian approach does not give close estimations to the Frequentist MLE solution. Posterior distribution is significantly different than the prior and the likelihood. The Bayesian approach gives different estimates for land use based nitrate export coefficients. On the other hand, predicted yearly average nitrate loads (kg/day) using either the Bayesian approach or the Frequentist approach have determination coefficient (R^2) values close to each other (R^2=0.75 for the frequentist approach

and R^2=0.74 for the Bayesian approach). Frequentist approach gives closer estimates to the observed values with respect to the Bayesian approach (Figure 4). A sample application of both the Frequentist and the Bayesian approaches for land use based nitrate export coefficients was shown in detail. Predicted nitrate export coefficients were tabulated in Table 2.

Conclusion

The primary objective of this research is to create a unique nutrient export coefficient model for the Melen River Basin, which has the wide range of land use characteristics. While doing this, retention coefficient and also the effect of the draining upper subwatershed were considered. Two different but related techniques were used for the modeling of nutrient export coefficients. These are so called the Frequentist approach or the maximum likelihood estimation and the Bayesian estimation using the MCMC algorithm. For the latter technique, the AMOS software was used. The Bayesian estimation differs from frequentist approach since it uses the prior information about estimands.

Based on the results, we can conclude that the Frequentist approach gives better estimations with respect to the Bayesian approach. Reliability of the results depends on the quality of the data used. Field works, especially sampling in dominated land use areas, helps to specify more reliable prior distribution of each land use based nutrient export coefficients in order to get more precise estimations, particularly through the Bayesian approach. The Frequentist approach gives convincing results. Results from the Bayesian approach would have been better if there could be a sufficiently large temporal data for headwater subwatersheds. Further studies, which take this issue into account, will need to be undertaken. This study has important findings for developing export coefficient models and it is intended to guide researchers on the subject.
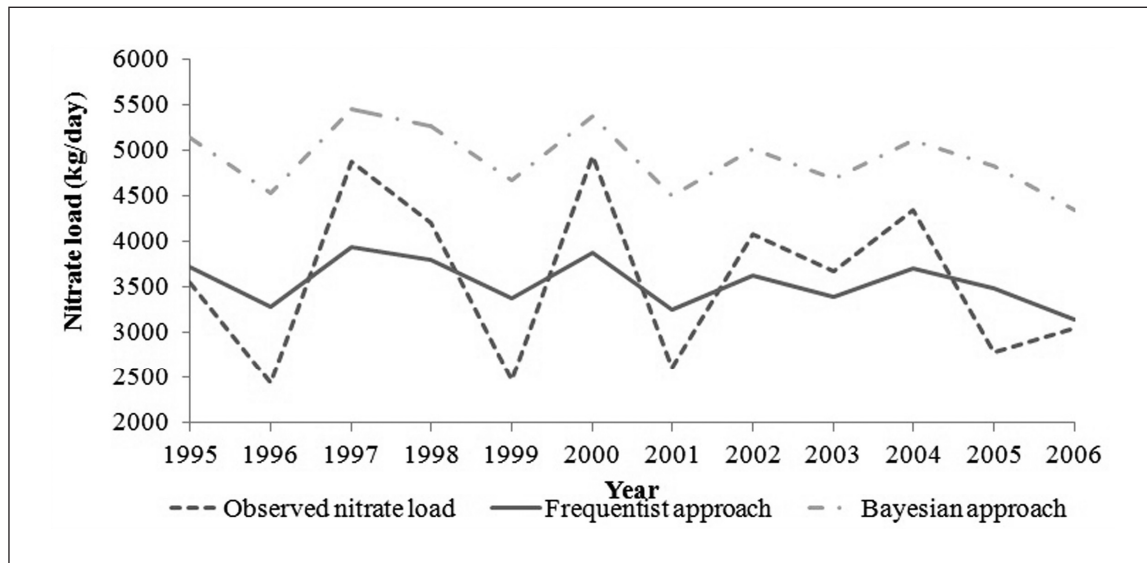
Fig. 4. Observed nitrate loads (kg/day), Frequentist and Bayesian estimations for each year between 1995 and 2006

Table 2. Estimated nutrient export coefficients (kg/km²/day) using Frequentist and Bayesian approaches

	Nitrate (NO_3^-) export coefficients (kg/km²/day)			
	Meadows Brush and Pasture	**Agricultural**	**Forest**	**Residential**
The Frequentist Approach	0.759	2.749	0.606	1.678
The Bayesian Approach	1.611	3.832	1.288	2.462

References

Arbuckle, J.L. (2009). *AMOS 18 User's Guide,* AMOS Development Corporation, USA 2009.

Brigault, S. & Ruban, V. (2000). External phosphorus load estimates and P-budget for the hydroelectric reservoir of Bort-Les-Orgues, France, *Water, air, and soil pollution*, 119 (1–4), pp. 91–103.

Chenini, I. & Khemiri, S. (2009). Evaluation of ground water quality using multiple linear regression and structural equation modeling, *International Journal of Environmental Science and Technology*, 6 (3), pp. 509–519.

Cowan, G. (2012). *Topics in Statistical Data Analysis for HEP*, Lhc Physics, USA, 341.

De Klein, J.J.M. & Koelmans, A.A. (2011). Quantifying seasonal export and retention of nutrients in west European lowland rivers at catchment scale, *Hydrological Processes*, 25 (13), pp. 2102–2111.

Dillon, P.J. & Kirchner, W.B. (1975). The effects of geology and land use on the export of phosphorus from watersheds, *Water Research*, 9, pp. 135–148.

Donohue, I., Styles, D., Coxon, C. & Irvine, K. (2005). Importance of spatial and temporal patterns for assessment of risk of diffuse nutrient emissions to surface waters, *Journal of Hydrology*, 304 (1–4), pp. 183–192.

DSI (State Hydraulic Works), (2011). *1995–2006 Stream Pollution Parameters Measurements in the Melen Watershed*, Ankara, Turkey 2011. (Unpublished report).

Essahale, A., Malki, M., Marín, I. & Moumni, M. (2010). Bacterial diversity in Fez tanneries and Morocco's Binlamdoune River, using 16S RNA gene based fingerprinting, *Journal of Environmental Sciences*, 22 (12), pp. 1944–1953.

Everitt, B.S. (2006). *The Cambridge Dictionary of Statistics*, Third Edition, Cambridge University Press, Cambridge, 432.

Fu, B., Wang, Y.K., Xu, P. & Yan, K. (2012). Modelling nutrient retention function of ecosystem – a case study in Baoxing County, China, *Procedia Environmental Sciences*, 13, pp. 111–121.

Gardner, M. (2012). Improving the interpretation of 'less than' values in environmental monitoring, *Water and Environment Journal*, 26 (2), pp. 285–290.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper), *Bayesian Analysis*, 1 (3), pp. 515–534.

Hogan, D.M., Labiosa,W., Pearlstine, L., Hallac, D., Strong, D., Hearn, P. & Bernknopf, R. (2012). Estimating the cumulative ecological effect of local scale landscape changes in South Florida, *Environmental Management*, 49 (2), pp. 502–515.

Huang, S.J. & Yu, J. (2010). Bayesian analysis of structural credit risk models with microstructure noises, *Journal of Economic Dynamics and Control*, 34 (11), pp. 2259–2272.

Izagirre, O., Argerich, A., Martí, E. & Elosegi, A. (2013). Nutrient uptake in a stream affected by hydropower plants: comparison between stream channels and diversion canals, *Hydrobiologia*, 712(1), pp. 105–116.

Konomi, B.A., Dhavala, S.S., Huang, J.Z., Kundu, S., Huitink, D., Liang, H., Ding, Y. & Mallick, B. K. (2013). Bayesian object classification of gold nanoparticles, *The Annals of Applied Statistics*, 7(2), pp. 640–668.

Law, A.M. & Kelton, W.D. (1991). *Simulation Modeling and Analysis*, McGraw-Hill, Inc., New York 1991.

Lee, P.M. (2012). *Bayesian Statistics: An Introduction*, 4th Edition, Wiley, USA 2012.

Lele, S.R., Dennis, B. & Lutscher, F. (2007). Data cloning: easy maximum likelihood estimation for complex ecological models

using Bayesian markov chain monte carlo methods, *Ecology Letters*, 10, pp. 551–563.

Loehlin, J.C. (2013). *Latent variable models: An introduction to factor, path, and structural equation analysis,* Psychology Press 336.

Mallick, B.K. (2013). Bayesian object classification of gold nanoparticles, *The Annals of Applied Statistisc*, 7(2), pp. 640–668.

Meer, M.H., Horne, J.B., Gardner, M.G., Hobbs, J.P.A., Pratchett, M. & Herwerden, L. (2013). Limited contemporary gene flow and high self-replenishment drives peripheral isolation in an endemic coral reef fish", *Ecology and Evolution,* 3(6), pp. 1653–1666.

Morris, W.K., Vesk, P.A. & McCarthy, M.A. (2012). Profiting from pilot studies: Analysing mortality using Bayesian models with informative priors, *Basic and Applied Ecology*, 14(1), pp. 81–89.

Nichols, J.M., Judd, K.P., Olson, C.C., Waterman, J.R. & Nichols, J.D. (2013). Estimating detection and identification probabilities in maritime target acquisition, *Applied Optics*, 52 (12), pp. 2531–2545.

O'Reilly, J.X., Jbabdi, S. & Behrens, T.E. (2012). How can a Bayesian approach inform neuroscience?, *European Journal of Neuroscience*, 35 (7), pp. 1169–1179.

Ozturk, I., Tanık, A., Cokgor, E.U., Gurel, M., Mantas, E.P., Insel, G. & Ozabali, A. (2008). *Buyuk Melen Watershed Integrated Protection and Water Management Master Plan*, Istanbul Technical University, Turkey 2008.

Peterson, B.J., Wollheim, W.M., Mulholland, P.J., Webster, J.R., Meyer, J.L., Tank, J.L., Marti, E., Bowden, W.B., Valett, H.M., Hershey, A E., McDowell, W.H., Dodds, W.K., Hamilton, S.K., Gregory, S. & Morrall, D.D. (2001). Control of nitrogen export from watersheds by headwater streams, *Science*, 292, pp. 86–90.

Ramirez, J.G. & Sanz, R. (2013). On the limitations of standard statistical modeling in biological systems: a full Bayesian approach for biology, *Progress in Biophysics and Molecular Biology*, http://dx.doi.org/10.1016/j.pbiomolbio.2013.03.008.

Rast, W. & Lee, G.F. (1983 Nutrient loading estimates for lakes, *Journal of the Environmental Engineering Division ASCE*, 109, pp. 502–517.

Sumer, B., Ileri, R., Samandar, A. & Sengorur, B. (2001). Water quality in Buyuk Melen and its branches, *Ekoloji*, 10 (39), pp. 13–18.

Vsetickova, L., Adamek, Z., Rozkosny, M. & Sedlacek, P. (2012). Effects of semi-intensive carp pond farming on discharged water quality, *Acta Ichthyologica Et Piscatoria*, 42 (3), pp. 223–231.

Wickham, J.D., Wade, T.G. & Riitters, K.H. (2008). Detecting temporal change in watershed nutrient yields, *Environmental Management*, 42, pp. 223–231.

Wickham, J.D., Riitters, K., Wade, T.G. & Jones, K. (2005). Evaluating the relative roles of ecological regions and land-cover composition for guiding establishment of nutrient criteria, *Landscape Ecology*, 20 (7), pp. 791–798.

Yang, G., Best, E.P., Whiteaker, T., Teklitz, A. & Yeghiazarian, L. (2014). A screening-level modeling approach to estimate nitrogen loading and standard exceedance risk, with application to the Tippecanoe River watershed, Indiana, *Journal of environmental management*, 135, pp. 1–10.

Zobrist, J. & Reichert, P. (2006). Bayesian estimation of export coefficients from diffuse and point sources in Swiss watersheds, *Journal of Hydrology*, 329 (1), pp. 207–223.