Improving uncertain nutrient load estimates for Lake Balaton

A. Clement
Department of Sanitary and Environmental Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary

Abstract Annual nutrient loads have been estimated for Lake Balaton over three decades. Tributaries may transport about half of the loads into the lake. The contribution of diffuse sources may reach two thirds of the total load. Biweekly/monthly water quality monitoring on small inflows (0.01 m³/s–0.3 m³/s range) results in a high uncertainty of load estimates. This paper evaluates the degree of uncertainties by using analytical expressions of sampling theory. Load-flow relationships were derived for five streams and annual total phosphorus load was predicted by four load estimation methods. A seasonal regression model, based upon the evaluation of historical set of observed phosphorus loads, appeared best to refine load estimates on small inflows. Correction frequently led to load estimates that exceeded uncorrected loads by a factor of two to three. Since the dynamics of the watercourses determined the errors of load estimates, stratified sampling is needed to decrease the uncertainties.

Keywords Lake Balaton; load estimates; monitoring; phosphorus; sampling; uncertainties

Introduction During the last two decades, large-scale eutrophication management measures have been taken in Lake Balaton that had undergone a rapid eutrophication during the 1970s (Herodek, 1986). Development of the wastewater infrastructure significantly decreased the contribution of point sources, and thus, the share of diffuse sources increased to 60–70% (Jolánkai, 1997). Tributaries transport a considerable portion of the diffuse loads to the lake. The largest tributary, the Zala River (mean flow is 7 m³/s), drains nearly half of the watershed of Lake Balaton (5776 km²). Mean flow rate of the 50 other perennial and intermittent streams varies between 0.01 m³/s–0.3 m³/s (Figure 1). Loads have been monitored daily at the mouth section of the Zala River since 1975. However, biweekly/monthly water quality observations in the small inflows result in a high uncertainty of load estimates (Somlyódy and van Straten, 1986; Jolánkai, 1997). Most of the annual flow of these streams is associated with some flood events that pass within several hours or days depending on the size and type of the watersheds. Water quality sampling coincides with these flood waves accidentally. Concentrations may exceed the average values by orders of magnitude after heavy storms. Annual nutrient loads have been estimated regularly during the last three decades. Despite the wide range of factors causing potential uncertainties in load estimates, the “routine” method of estimating tributary loads to Lake Balaton is to multiply the measured concentration with the sampling flow. Future water quality management of Lake Balaton should focus on reducing diffuse loads (Istvánovics and Somlyódy, submitted). Thus, refinement of nutrient load estimates is crucial to correct management decisions.

Problems associated with infrequent sampling are well known from water quality monitoring practice (Bodo and Unny, 1983; Somlyódy and van Straten, 1986; Preston et al., 1989; Preston and Bierman, 1992; USEPA, 1999). Tributary loading is a continuous process that can be represented as the product of concentration and flow over time. Flow can be measured continuously. However, water quality sampling data are typically available on a
more limited basis. In order to improve the accuracy of load estimates, sampling frequency should be increase. Commonly routine sampling data may exhibit a bias towards low streamflow rates (Dolan et al., 1981). Therefore, in cases where pollution events are relatively brief (e.g. small and intermittent creeks), sampling periods have to be short. Since concentration measurements are expensive, a practical balance must be maintained between an acceptable level of error and cost (Preston and Bierman, 1992). However, load estimates can be improved considerably by “generating” the load values from more frequent flow data using statistical methods (Richards, 1989). Available loading estimators can be grouped into three main categories including averaging, ratio and regression methods (Dolan et al., 1981; Bodo and Unny, 1983; Miertschin, 1986; Preston et al., 1989; Vries and Klavers, 1994).

Unfortunately limited information has been found about the performance of these methods in the case of event-responsive systems like tributaries of Lake Balaton. The objective of the present study is to assess uncertainties of load estimates for small inflows of the lake.

Materials and methods
Continuously flow data and biweekly total phosphorus (TP) and total nitrogen (TN) concentration measurements of eight of the small inflows of Lake Balaton were used in the study (Table 1; source: National Water Quality Monitoring Data base).

Errors associated with infrequent sampling can be analyzed statistically (Cochran, 1962). Consider a finite population (time series) of “N” uncorrelated elements. The exact sample mean is $\bar{Y}_N$ and the standard deviation is $\sigma_N$. If we take a random sample of “n” elements (1 $\leq$ n $\leq$ N) without replacement, the expectation of the sample mean $E(\bar{Y}_N)$ gives an unbiased estimate of $\bar{Y}_N$. Assuming Gaussian distribution, the relative error of the estimate ($\Delta y / \bar{Y}_N$) at 95% confidence level ($t = 1.96$) is

$$\alpha = t \frac{\sigma_N}{\bar{Y}_N} \left( \frac{N-n}{Nn} \right)^{1/2}.$$  \hspace{1cm} (1)

Equation (1) shows that the standard deviation/mean ratio determines the number of samples required to reach a predetermined accuracy (see also Sherwani and Moreau, 1975; Sanders et al., 1983; USEPA, 1999). Since the distribution of flow and water quality variables is typically skewed, for example, lognormal, the above analytical expression is
approximate. In an earlier study (Clement and Buzás, 1999) we compared the errors of annual runoff and load estimates obtained by Monte Carlo analysis with those derived from Cochran’s analytical expression. For the Zala River and the Danube, despite a slight underestimation of the error, the analytical method yielded satisfactory results.

In a number of small inflows of Lake Balaton flow is registered continuously \((N = \text{some hundreds to thousands samples/yr})\) and the annual mean flow \((\bar{Q}_N)\) may be determined with reasonable accuracy, whereas the frequency of water quality sampling is much lower \((n = 26 \text{ samples/yr})\). Further considerations are required to establish the link between flow variability and load estimate error. The \(L_N / \sigma_{LN} \) ratio of load \((L)\) was approximated by the \(\sigma_{ON} / \bar{Q}_N \) ratio of flow \((\bar{Q})\) as follows (Clement and Buzás, 1999):

\[
\frac{\sigma_{LN}}{L_N} \equiv \frac{\sigma_{ON}}{\bar{Q}_N} \beta \quad \text{and} \quad \beta = \frac{\sigma_{Ln}}{\sigma_{Qn} / \bar{q}_n},
\]

where \(L_N\) is the accurate annual main load (which is unknown), \(I_N\) is the annual mean load derived from discrete observations \((I_i = q_i c_i\) ), and \(\bar{q}_n\) is the mean sample flow. The proportionality factor, \(\beta\), can be obtained from \(n < N\) water quality samples. Equations (1) and (2) allow us to determine the error of annual load estimates or to calculate the sampling frequency needed to reach the required accuracy.

Finally the improvement of load estimates of the small inflows was attempted. The performances of four load calculation methods were analyzed, as follows: (1) measured concentration multiplied by the sample flow \((I_N)\), this corresponds to the “routine” method; (2) the unbiased ratio estimator including a bias adjustment term (Cochran, 1962), where the mean load is basically a product of the mean flow \((\bar{Q}_N)\) and the flow-weighted mean concentration derived from discrete observations; (3) regression based on load-flow relationships obtained from historical sampling data; and (4) the “cluster” method. In the latter the observed range of flow is subdivided into clusters. Average loads of TP are determined for each flow cluster, and time series of load is generated from daily flow data.

**Error of load estimates in small inflows of Lake Balaton**

The \(\sigma_{ON} / \bar{Q}_N\) ratio of flow was calculated for eight streams using Eq. 1 (Table 1). The values showed wide variation from inflow to inflow and year by year. From the analysis of some major Hungarian rivers (Clement and Buzás, 1999), we found that the \(\sigma_{ON} / \bar{Q}_N\) ratio decreases with increasing mean flow. In contrast to this empirical relationship, the mean flow and the \(\sigma_{ON} / \bar{Q}_N\) ratio were not interrelated in the case of the small inflows. The possible explanation is that large rivers aggregate and smooth specific differences of their

<table>
<thead>
<tr>
<th>Water course</th>
<th>Mean flow at the mouth (m³/s)</th>
<th>Catchment area (km²)</th>
<th>(\sigma_Q)</th>
<th>(\alpha_{TP})</th>
<th>(\sigma_{TN})</th>
<th>(\alpha_{max})</th>
<th>(\sigma_{Q,mean})</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burnót-creek</td>
<td>0.14</td>
<td>82</td>
<td>1.08</td>
<td>±41%</td>
<td>±65%</td>
<td>±44%</td>
<td>-44%</td>
<td>-20%</td>
</tr>
<tr>
<td>Eger-creek</td>
<td>0.19</td>
<td>361</td>
<td>1.58</td>
<td>±60%</td>
<td>±82%</td>
<td>±55%</td>
<td>-56%</td>
<td>-22%</td>
</tr>
<tr>
<td>Fuzfoi-creek</td>
<td>0.02</td>
<td>9.5</td>
<td>1.45</td>
<td>±54%</td>
<td>±64%</td>
<td>±54%</td>
<td>-52%</td>
<td>-45%</td>
</tr>
<tr>
<td>Jamai-creek</td>
<td>0.08</td>
<td>59</td>
<td>0.89</td>
<td>±30%</td>
<td>±47%</td>
<td>±33%</td>
<td>-31%</td>
<td>-20%</td>
</tr>
<tr>
<td>Koröshegyi-creek</td>
<td>0.09</td>
<td>34</td>
<td>0.87</td>
<td>±33%</td>
<td>±54%</td>
<td>±26%</td>
<td>-39%</td>
<td>-31%</td>
</tr>
<tr>
<td>Örvényesi-creek</td>
<td>0.05</td>
<td>20</td>
<td>0.57</td>
<td>±22%</td>
<td>±44%</td>
<td>±23%</td>
<td>-15%</td>
<td>-5%</td>
</tr>
<tr>
<td>Tapolca-creek</td>
<td>0.24</td>
<td>39</td>
<td>0.38</td>
<td>±14%</td>
<td>±21%</td>
<td>±16%</td>
<td>+10%</td>
<td>-5%</td>
</tr>
<tr>
<td>Tétves-creek</td>
<td>0.19</td>
<td>101</td>
<td>1.33</td>
<td>±51%</td>
<td>±64%</td>
<td>±60%</td>
<td>-50%</td>
<td>-43%</td>
</tr>
</tbody>
</table>
numerous subwatersheds, whereas the flow regime of small watercourses sensitively reflect specific features of their watersheds. For example, Tetves- and Eger-creaks springing from hilly regions have more variable flow than that of the Jamai- and the Köröshegyi- creeks, which have flatter watersheds. The permanent springs and the large wetland type watershed near the mouth balance the flow regime of the Tapolca-creek.

In order to characterize the dynamics of watercourses related to the variability of the watersheds characteristics, I performed multiple regression analysis using a few GIS-based parameters that are presumably capable of catching this variability. The following expression yielded the best fit to the data ($r^2=0.88$):

$$\sigma_Q / \bar{Q}_N = 0.74S - 86.6SR + 2.19AU + 1.26RQ + 0.95.$$  

(3)

where $S$ is the proportion of the catchment area with more than 3º slope to the whole catchment (%), $SR$ is the specific runoff (the mean flow at the mouth divided by the catchment area, m$^3$ s$^{-1}$ ha$^{-1}$), $AU$ is the ratio of urban area (%), and $RQ$ is the ratio of the base flow and the annual mean flow.

In order to determine the error of annual load estimates, the $L / \sigma_{LN}$ ratio was calculated from Eq. (2) for the TN and TP loads. The 95% probability error of the annual mean flow ($\alpha_Q$) and loads ($\alpha_{TP}$ and $\alpha_{TN}$) estimated from biweekly samplings were derived from Eq. (1) ($N = 365$ to $2000$, corresponding to the number of flow measurements in each year and $n = 26$). The results are shown in Table 1.

One can check the results obtained on the basis of purely theoretical considerations by calculating the difference of the annual mean flow estimated from continuous and biweekly data (Table 1). The highest observed difference ($\alpha_{Q,max}$) reasonably approximated the 95% probability error derived from Eq. (1). Since the distribution of flow is skewed towards high flows, infrequent observations are less likely to fall into the range of higher flows. As a consequence, the actual runoff was underestimated in the majority of cases (Table 1). Since diffuse loads mainly coupled with precipitation and the relationship is relative strong between the load and flow (see later), the annual loads are likely underestimated, too.

Monte Carlo simulation provides another possibility to check the applicability of the analytical method. From the continuous flow series of the Tetves-creek, $n=26$ data were selected randomly. Figure 2(a) presents the distribution of relative errors obtained in 1000 repetitions. The relative error at 95% probability level was 0.58. That is slightly higher than the analytical error (0.51, Table 1). This is in agreement with our earlier results on large rivers (Clement and Buzás, 1999).

Equation (1) allows the calculation of sampling frequency required to achieve a desired level of precision in the estimation of annual loads. Provided that a 15% error is allowed,
even daily sampling would be insufficient in the highly dynamic inflows of Lake Balaton. In addition to this, a fixed number of samples result in a seasonally and annually variable accuracy in accordance with the temporal variability of the standard deviation/mean ratio (Figure 2b). Since the necessary sampling size is proportional to the variance, which may differ considerably over time, sampling frequency will need to be greater during periods of greater variance (Sherwani and Moreau, 1975). An effective means of treating such populations is the use of stratified sampling and to separate event and non-event periods (Cochran, 1962; Bodo and Unny, 1983). Although the latter is easily achievable when evaluating historical data, it remains unknown at the time of sampling. Consequently, automated sampling devices must be applied that set the number of samples to the level or to the rate of flow, and thus ensure an adjustment of the sampling frequency to the expected variance.

Improving load estimates of small inflows

In order to improve TP load estimates in small inflows, the load-flow relationships of five creeks were evaluated. The period to be analyzed must be specified with great care. On the one hand, too short a time series is likely to exclude measurements of extreme flood events (Jolánkai et al., 1999). On the other hand, historical data may truncate the time series and mask significant recent changes in the watershed (e.g. land use, fertilizer application etc.). In order to avoid these distortions, the available load series were shortened year by year starting at the beginning of the observations (1975) until no substantial changes occurred in the mean TP load. Finally, 150 to 300 data pairs corresponding to the last 8–15 years were left in the various inflows.

The relationship between flow and TP load varies at different time scales. Although diffuse pollution increases rapidly when runoff starts and precipitation determines the dynamics of transport (Jolánkai et al., 1999), seasonal differences are considerable. Short, heavy storms in summer penetrating through the vegetation cover onto the dry soil obviously generate different flow-load relations than quiet autumn rainfalls. For many streams the greatest concentrations of suspended sediment and other pollutants (e.g. particulate P) occur during spring following snowmelt (Haneda and Matsumoto, 1983; USEPA, 1999). Four main characteristic periods were distinguished:

i. the summer period (June-August);
ii. the vegetation period with the exception of summer month;
iii. the winter and early spring period (November-April) excluding the beginning of snow melting;
iv. and the assumed beginning of snow melting based on data of water and air temperature.

The regression analysis was performed on seasonally data using the formula suggested by Jolánkai et al. (1999):

\[ l_i = a_0 + a_1 q_i + a_2 q_i^2, \]  

(4)

where \( a_0 \) might be considered to loads from point sources, the second term might reflect the natural background load, and the third one refers to diffuse loads. The results showed that distinguishing of characteristic periods with different runoff features considerably improved the performance of the regression model. The slopes indicate that precipitation induces higher P concentrations during the summer than in other periods. Dilution by intense precipitation is evident in the autumn and the winter. The relationship between the load and flow was highly variable in both space and season. The weakest correlation was found in the Tapolca-creek with the highest share of loads from point sources. The best correlation was obtained in the Tetves-creek, where erosion considerably contributes to TP load. Figure 3 shows the results for the Eger-creek. Since most of the excess P comes from soil erosion during high runoff, in accordance with the results of Somlyódy and Straten (1986), the regression model could be improved considerably by including the
concentration of suspended solids. Unfortunately, the SS concentration is measured biweekly in the small inflows preventing the application of the improved model.

Seasonal regressions were used to generate time series of load from the continuous flow record and summarized to obtain annual loads ($L$) for the period 1990–1996. A comparison of the estimated loads with biweekly observations (Figure 4a) showed that the regression model overestimated the load in the range of low flows. Therefore distinguishing periods of low and high flows further refined the method. Loads transported by flood events over the 80% probability level of flow were generated by the regression model, whereas the base load was calculated from the mean of the biweekly observations. Annual loads were summed up from the “base load” and the “event load”.

In order to apply the “cluster” method, the number of flow clusters was established from the evaluation of empirical distribution of the flow time series. The upper cluster corresponded to the flow of 98% probability (considering extreme events) during the examined period. Mean TP concentration and load values were calculated for each flow cluster, then load time series were generated from the continuously flow records (Figure 4a). The share of the annual flow and the generated TP load corresponding to flow clusters were compared (Figure 4b). It was found that in the case of the high variability inflows, nearly 25% of the runoff and 30–50% of the load was associated with extreme flood events. Flows of 80% probability transported half of annual runoff and 50–80% of the TP load. Bodo and Unny established similar results for a small “flashy” stream considering the transport of suspended solids.

Finally, the results of the use of different load estimators are summarized. Figure 5(a) displays the series of calculated annual TP loads for five inflows. Substantial variability is evident among the methods. In the majority of cases, the “routine” procedure and the ratio estimator predicted lower values than other methods. Predictions derived from the load-flow relationship (regression and “cluster” methods) are generally similar to each other. There is no a priori test to control the performance of the estimators. The averaging method applied...
routinely may underestimate the actual loads; however, the theoretical accuracy of the estimates is known (Table 1). It is assumed that the predicted loads (by the ratio, regression and cluster methods) approximate the accurate values. The differences between annual loads estimated from the “routine” and the other methods were calculated and compared with the 95% probability errors (Figure 5b). Occasionally, the underestimation was substantial: predicted loads may exceed the measured one by a factor of two to three (Figure 5). In accordance with the results of the statistical analysis (Table 1), differences, which were found to be the highest in the case of Burnót, Tetves, and Eger creeks, also varied from year to year.

The question arises, how this uncertainty influences the estimated loads of Lake Balaton. Assuming that TP loads were underestimated in each stream to the highest degree, small inflows might transport nearly 200% times more TP. This would correspond to a 5–20% increase in the P load of Lake Balaton relative to the standard estimate based on the “routine” method (Jolánkai, 1997). The unmeasured TP load may be as high as 40–50 tons yr⁻¹, nearly equal to the target load prescribed for 2010. Obviously, these figures represent the upper limit of errors that are unlikely to occur. However, they clearly illustrate that load estimates of sufficient accuracy need much more efforts than biweekly sampling in the small, dynamic inflows. From the viewpoint of water quality management, these results further support the conclusion that an intensified monitoring effort before planning further measures to reduce diffuse loads is needed.

Conclusions

i. The error of load estimates in case of biweekly sampling in small inflows varies between 15–80%;

ii. The uncertainties mainly coupled with the flow regime of the watercourses, which sensitively reflect specific features of their watersheds;

Figure 5  Summary of the results of different load prediction methods: estimated annual TP loads (a), and the differences between annual loads estimated from the “routine” and the other methods (b).
iii. Load estimates can be improved by using different load estimators. The widely used ratio estimator did not yield considerably different loads as compared to the biased “routine” method. Loads computed by procedures based upon the evaluation of historical set of observed phosphorus loads might exceed the measured load by a factor of two to three. The seasonal separation of data showed that distinguishing characteristic periods of different runoff features could considerably improve the performance of the regression model. Experimental measurements of the load during flood events in the major inflows would be sufficient to refine the load-flow relationships.

iv. The “routine” method of estimating nutrient loads to Lake Balaton results in a 5–20% annual underestimation, but this would correspond to 50–150% unmeasured TP load in the eastern bays of the lake.

Acknowledgements

The study was performed in the frame of the Lake Balaton Research Programme of the Hungarian Prime Minister’s Office. I express particular thanks to V. Istvánovics and K. Buzás for their comments on the present paper.

References


