MODELING A COMPLEX ENVIRONMENTAL SYSTEM: THE LAKE BALATON STUDY

L. Somlyódy

August 1981
WP-81-108

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE AUTHOR

László Somlyódy is the leader of the Balaton Case Study, Resources and Environment Area, International Institute for Applied Systems Analysis, Laxenburg, Austria (on leave from the Research Centre for Water Resources Development (VITUKI), Budapest, Hungary).

PREFACE

One of the principal projects of the Task on Environmental Quality Control and Management in IIASA's Resources and Environment Area is a case study of eutrophication management for Lake Balaton, Hungary. The case study is a collaborative project involving a number of scientists from several Hungarian institutions and IIASA (for details see WP-80-187).

This paper primarily considers the methodological framework of the Lake Balaton Case Study, using it as an example of how a system characterized by complexity and uncertainty can be modelled. The suggested approach is supported by several examples and provides an account of both developments since the publication of WP-80-187, and also necessary research for the project's completion.

CONTENTS

INTRODUCTION

The eutrophication of lakes, a typical, unfavorable manifestation of the past few decades, is a consequence of the increase in the amount of nutrients (such as phosphorus and nitrogen compounds) reaching water bodies. This increase is closely related to the generally rapid development of industry, agriculture, and tourism within the watershed or in short, to a change in the infrastructure of the region. Eutrophication, an in-lake phenomenon, the origin of which lies outside the lake, causes unpleasant consequences (e.g., a rise in algal biomass, water discoloration, taste and odor problems, bacterial contamination, etc.) which can greatly limit the use of the lake's water for recreation, water supply, etc., and lead to a drastic change in the ecosystem.

The phenomenon of eutrophication and mathematical modeling of it have been quite well explored for deep lakes, but only to a lesser extent for shallow lakes. Here, due to the low depth and the generally strong wind action, stratification rarely occurs. The dynamics are more complex and are faster; consequently, shallow lakes are much more affected by changes in environmental factors and show less consistent patterns from year to year. The recognition of this gap in knowledge led IIASA's Resources and Environment Area to initiate research on the eutrophication of shallow lakes. Lake Balaton in Hungary, the largest lake in Central Europe, which has exhibited the unlikely signs of eutrophication, was chosen to be one of the two case studies. The case study is being carried out in cooperation with the Hungarian Academy of Sciences and with the participation of several Hungarian institutions.

The general features of the problem from a methodological standpoint are as follows: (i) The system composed of the water body and the watershed is large in a physical sense; therefore, *understanding and managing* it cannot be accomplished solely on an experimental basis. Modeling is absolutely necessary and should be in harmony and interaction with in situ and laboratory measurements and data collection, respectively. (ii) There are strong interactions among biological, chemical, and hydrophysical in-lake processes; furthermore, between the watershed and water body. (iii) There are several stochastic influences (e.g., meteorology, hydrology). (iv) The data available are often inadequate and scarce. Uncertainty consequently plays an important role. (v) The vicinity of the lake is the major tourist resort in Hungary, so there is a strong economic interest in a practical solution to the problem. In other words, the management of the system is of primary interest.

From these characteristics it follows that the lake and its region form a complex environmental system; the problems related to it require a systems analytical approach. The aim is to handle the problem in both a research and management context, that is, to better our understanding on the scientific level and then to utilize this knowledge for working out optimal control strategies for improving the water quality of the lake. The elaboration of such a model or set of models is not an easy task and meets serious methodological difficulties. In this respect, the methodologies developed within IIASA's Resources and Environment Area (see Beck in this issue [4]) could provide support, but at the same time, the case study itself should serve to answer in a wider context, the methodological questions addressed; another reason why the case study was initiated.

The objective of this paper is to illustrate with the example of Lake Balaton, how such a complex system can be understood and managed in its entirety, i.e., together with the related methodological questions. The paper is organized as follows: first, the main characteristics of the system and the modeling approach are outlined (Sections 2 and 3). In Section 4, the individual steps of the analysis are illustrated through examples. These will be related to the sediment-water interaction, spatial mass exchange, nutrient loading problem, and the lake eutrophication model. In most of the examples the influence of uncertainties and stochastic effects will be accounted for. At the end of this section, it is shown how the lake water quality model can be incorporated into the management framework.

It has to be mentioned that the study has not yet been completed. Therefore, in some cases, reference will be made to results which are preliminary only.

2. MAJOR CHARACTERISTICS OF THE SYSTEM

The lake and its watershed are illustrated in Figure 1. The length of the lake is 78 km, the average width around 8 km (surface area nearly 600 km^2) and the average depth 3.1 m. The major inflow of the lake is the River Zala at the southwestern end of the lake which drains half of the total catchment area ($\sim$ 5800 km^2). There is a single outflow at the other end of the lake, Siófok, through a control gate. The mean residence time of water is about 2 years.

The fluctuation in the water's temperature is high. There is a relatively long ice-covered period (around two months), while the temperature in summer may exceed 25° C. Concerning the chemical composition of the water, the high calcium carbonate content and pH value (8.3 to 8.7) should be mentioned. Wind action is important (there are about 80 "stormy" days in a year) resulting in favorable oxygen conditions and a permanent back and forth motion (seiche) along the lake and a complicated circulation pattern. Wind strongly influences sedimentation and release of various materials from the sediment layer (its organic material content is relatively low).

In recent years, remarkable changes have been observed in the water quality due to the rapid increase in tourism, sewage discharges, fertilizer use, and other factors. The algal biomass (algae is the most important primary producer in this case) increased by a factor of 10 when compared with the past 15-20 years. The trend in primary production is similar and at the most polluted western basin, peaks of up to 13.6 g C/m^2d were observed, a hypertrophic value [12]. In short, the average lake conditions moved from mesotrophic to eutrophic, thus endangering the use of the lake for recreational purposes, the prime water use in this case.

Phosphorus plays a dominant role in the eutrophication of the lake (the lake is phosphorus-limited most of the time). Thus, both from the point of view of understanding and managing the system, tracing the phosphorus compounds in the lake and on the watershed is of primary interest. The total phosphorus load of the lake is around 1000 kg/d [15], half of which is available for algal uptake. The load has many components: 33% is derived from sewage, 27% from diffuse sources, 22% is related to runoff processes in the dircet vicinity of the lake, while the contribution of atmospheric pollution is 18%. The ratio of sewage discharges in the available load is higher; only the sewage released in the recreational area (Figure 1) accounts for 36% of the available load. This direct load varies quite a lot, following the fluctuations in population due to tourism, and has a 2-4 time higher value in summer than during the off-season. The load distribution along the lake is approximately uniform; however the volume related value is twelve times higher at the Keszthely Bay (Figure 1) than at the other end of the lake, due to differences in the volume of the four main basins. This fact is also reflected in the pronounced longitudinal gradient of various water quality parameters, e.g., for Chlorophyll-a the ratio of the maximum and minimum values ranges between 4 and 20 [29]. The gradient observed at the same time indicates that the strong wind action and the mixing associated with it are still not sufficient for leveling out the spatial nonuniformities.

From an analysis of the data it is clear that there is not only a critical state of the water quality at Keszthely Bay, but also a spreading deterioration process which extends towards other areas of the lake where the water quality is still good. Thus action is urgently required from the view of the entire lake.

Based on hydrologic and water quality considerations, the lake was divided into four basins, as indicated in Figure 1. The application of the principle of segmentation proved to be a useful tool for modeling, data collection, and data handling.

Concerning data, extensive records are available on hydrology and meteorology. Regular water quality monitoring started ten years ago, in two network systems consisting of 9 and 16 spatial sampling points, respectively (10-20 measurements per year), but irregular data are also available back to the early sixties. Several other in situ and laboratory measurements were also taken (primary production, extinction, sediment-water interaction, velocity, etc.). A survey was done on the nutrient load between 1975-79, which involved 20 tributaries and 27 sewage discharge points [15] (indicated in Figure 1). On the major tributary, daily observations were made during this period [16].

On the basis of data collected by Hungarian institutions recently, more often in interaction with modeling--the "Lake Balaton Data Bank" was created at IIASA, serving as a starting point for the modeling work.

For further details on the case study, the reader is referred to [29, 30].

3. THE MODELING APPROACH

3.1 The previous sections demonstrated the complexity of the water quality problem. Here, methodological questions will be discussed. For the purpose of illustration the in-lake processes will be considered. From this example, conclusions will be drawn, leading to the modeling approach to be adopted.

The water quality of a lake is the result of several physical, chemical, and biological processes. The development of a model is generally based on the appropriate combination of knowledge gained from theory and measurements, respectively (the so-called theoretical and measurement knowledge [8]). Depending on the solidity of the theory of various processes, quite different approaches may be employed. For example, in most of the hydrodynamic applications ("hard" science, see [3]) the model structure is basically determined by the partial differential equations (PDEs) of continuity and momentum (exceptions may be caused by not well defined or unknown boundary conditions, see Section 4.1). Within the domain of water quality, the theoretical background of biology and chemistry ("soft" sciences) is less satisfactorily established and model development should be supported more extensively by the measurement knowledge. The development is generally based on testing hypotheses, and uncertainty plays an important role [9, 14, 30]. Under such conditions, steps such as model structure identification, parameter estimation and the analysis of error propagation [2] are inevitably required--techniques which are available mostly for ordinary differential equations (ODEs) only.

In the majority of the water quality models, the description of processes of both of the categories mentioned is necessitated; the dilemma then is [26]: how should the modeling procedure of biology and chemistry (ODE structure with interactive use of data) be combined with the "precise" treatment of hydrodynamics, the PDE structure of which a priori excludes the application of most of the techniques required for the cognition of phenomena of the other group and consequently the exploration of the entire process?

The answer seems to be relatively easy, namely one should start elaborating detailed models for both groups, and keeping in mind the main features of the other group (time and length scales [11], etc.), simplify them through a sensitivity analysis (aggregation on the base level). Then the same procedure should be repeated after coupling the models (aggregation on the higher level) which should then establish the relative importance of various processes and *harmony among the methodologies* employed.

From this conclusion, an *off-line* modeling approach follows, which avoids the direct coupling of all the detailed submodels. The procedure starts with a resonable *decomposition* of the complex structure into smaller, more tractable units which are accessible for separate and detailed studies. This is followed by *two* different kinds of *aggregations*, the aim of which is to preserve and integrate only essentials, but ruling out the unnecessary details. In contrast to the paper by Beck [4], the procedure of decomposition and subdivision has a different meaning here. In the course of submodel development, experiments of "isolated" character are certainly required, but then the coupled, aggregated ("smaller") model is validated against its detailed ("larger") version and data on the higher level (a tedious task which is rarely documented in the literature). It is also stressed that for such a complex system one "small" model is as unrealistic as one "large" model (accordingly, the solution cannot be looked for in this contrast, i.e., "larger" or "smaller" model--the question rasied by Beck [4]). Thus, a procedure which is required progresses through the subsequent development of *detailed* and *simplified models* by maintaining only the aggregated knowledge on the higher level.

Each modeling step is associated with model and data *uncertainties*; consequently, the aggregation is required in this respect as well: another guideline in our approach.

The tactics of decomposition are especially important if the objective of the study involves such different levels as the scientific understanding of a system and the decision making related to the same problem: a situation that we face here.

3.2 The application of this approach for the Lake Balaton problem is explained with the help of Figure 2, which shows the framework of the research [30]. The first decomposition that directly comes to mind is the distinction between lake and watershed, since as mentioned before, the water quality problem lies in the lake, but the causes, and practically all control possibilities are to be found in the watershed. Next, the various units should be separated and the essential results put together on a successively higher level of integration. The procedure involving five strata will be discussed in greater detail for the Lake Eutrophication Model (LEM), with reference to models now being elaborated. The parallels in the Nutrient Loading Model (NLM) can be found through Figure 1.

Stratum 5

First, those segments of the lake should be isolated which can be considered approximately uniform from the viewpoint of water quality (complete mixing inside each unit) and from the factors influencing them. The objective of the models on this stratum is to describe the algal dynamics and nutrient cycling for all the segments, involving both the water body and the sediment, since the latter is a sink and source of various materials and their interaction plays an important role in shallow lakes (also from the point of view of management, after a reduction in load, the new equilibrium of the lake will be determined by the nutrient release of the sediment). These kinds of models based on the mass conservation principle and formulated through a set of nonlinear ODEs are well-known in the literature [21]. In the frame of the present study, three submodels, BEM, BALSECT, and

SIMBAL were developed (see [13, 19, 31], with respect to their comparison [30]) which differs basically in the number of state variables (between 4 and 7) and essential parameters (10-17), as well as in the mathematical formulation of various processes and in the parameter estimation technique adopted. It is noted here that some of the parameters can be derived from further isolation up to a lower level with appropriately designed experiments. As examples, the estimation of algal growth parameters from vertical primary production measurements [32] and the study of wind induced sediment-water interaction (see Section 4.1) may be mentioned.

To end the discussion on stratum 5, it is stressed that several steps of the analysis are based on intuition (starting from the segmentation to the formulation of various biochemical processes), thus the inclusion of data with their uncertainties is imperative.

Stratum 4

On the next level the segment-oriented biochemical and sediment models are coupled by involving mass in- and outflows at the boundaries of the units. For this purpose, a hydrodynamic-transport model can be used. In light of the experiences gained from the study of the Great Lakes [7], it was decided not to use a coupled more-dimensional hydrodynamic-transport model incorporating the submodels of the lower stratum: the gain in information is not proportional to the increase in complexity; furthermore it causes methodological difficulties as explained in Section 3.1. Here again, an off-line technique is applied. The basic assumption is that it is sufficient to subdivide the lake in a longitudinal direction only. This is supported by the riverine shape of the lake and the presumably extensive transversal mixing, since the prevailing wind direction is nearly perpendicular to the lake's axis (the description of the shoreline effects is not the objective here). Consequently the parallel development of an unsteady three-dimensional (3D) and one-dimensional (1D) hydrodynamic model was decided on. The results gained (for details see [22] and [27] showed that the two models could be equally calibrated against dynamic water level data and suggested that the cross-sectionally averaged discharge, Q(t), can be properly derived from the much simpler 1D model which also allowed an uncertainty analysis on the wind data, a methodologically remarkable aspect (Section 4.2). The 1D version captures only the rather rapid seiche motion on the lake (convection) but not the different kinds of backflows and circulations: their mixing influence should be derived from the 3D model. In fact, a longitudinal dispersion coefficient, D_L, can be approximately calculated from the velocity field by the method similar to that employed for rivers [10]. Based on the initial experiences [23] it seems to be sufficient to perform the computations for some typical stormy events and afterward to correlate D_L to wind parameters. The relationship $D_L[W(t)]$ gained, will allow replacement of the coupled 3D hydrodynamic-transport model by a 1D version. Thus the submodels of stratum 5 will be incorporated into a set of longitudinal dispersion equations on stratum 4. This aggregation is achieved by a sensitivity analysis solely on the hydrodynamics. The second aggregation (see Section 3.1) can be arrived at with the use of the coupled dispersion biochemical model currently being elaborated, which allows a sensitivity study of all the processes involved.

It should be mentioned here (see Section 2) that provisionally, in all three biochemical models four segments (see Figure 1) are assumed; their coupling is based on hydrologic throughflow and a wind influenced mass exchange process described globally. Since the model structure based on ODEs has many advantages, one of the objectives of the study on the 1D coupled model is whether the four basins concept can be maintained or not.

Stratum 3

The involvement of mass exchange among segments as described before will result in the Lake Eutrophication Model (LEM) (Figure 2) which has several forcing functions, such as solar radiation, water temperature, wind, etc. (natural or *uncontrollable factors*) and the *nutrient load*. Since the latter is the only factor to be controlled, it plays a distinguished role; however, less modeling work was done on it in the frame of the case study. This can be explained by the relatively high contribution of the sewage load (modeling is basically not needed here because of the nature of the problem) and the limited amount of watershed data available for non-point source modeling. A thorough data collection and the derivation of a nutrient balance for the whole lake were preferred (for details see [15]), the results of which were already summarized in Section 2. This study also involved an uncertainty analysis in relation to the unobserved contribution of floods to the load (Section 4.3). The research allowed the derivation of the temporal and spatial pattern of the load components in a descriptive fashion, both for LEM and the Water

Quality Management Model (WQMM) on stratum 2. In fact, at the first stage of lake model development the modeling of any of the driving functions is not necessarily required; both for calibration and validation, historical data can be employed. For planning purposes the situation is different, therefore the *stochastic effects* of both the load and meteorology were involved through Monte-Carlo simulation (Section 4.4).

<u>Stratum 2</u>

The objective of WQMM is to generate alternative management options and strategies (the effect of these being expressed through NLM which should be used here in a planning mode) and to select from among these alternatives, on the basis of one or more objectives. Both water quality and expenses can be used as objective functions or constraints, and quite often their weighting is required. Frequently the load can replace the lake's water quality in the optimization in which case LEM is used merely to check the reaction of the lake and WQMM may have a simpler structure. Admittedly however, the first version is more obvious because of the nature of the problem. This formulation however leads to the dilemma: how should a complex dynamic model be incorporated into the optimization framework [28]?

At this step aggregation is also needed. This starts with the selection of certain *water quality indicators* characterizing the large scale and long-term behavior of the system serving as a basis for decision making. Different parameters (yearly peak, different averages, duration of critical concentrations, frequency distributions, etc.) of typical water quality components (primary production, algal biomass, Chlorophyll-a, etc.) can be used as indicators. Subsequently the dynamic model LEM can be used in terms of indicators established, I, under reduced loading conditions or in another way under several loading scenarios, L. Since the definition of indicators introduces temporal averaging, it is expected that the lake's response will be less complex compared to the dynamic simulation and a simple, direct I(L) type relationship can be found. If such a solution has already been attained, LEM could be replaced by I(L) in WQMM; an essential aggregation (see Section 4.5).

Among the management alternatives, only the two most important options are mentioned here: (i) tertiary treatment (point source load reduction), (ii) establishing reservoirs (consisting of two segments serving for the removal of both particulate and dissolved nutrient forms, respectively [29]) at the mouth of rivers which are the recipients of point and non-point source pollutants. The optimization should then be based on the trade-off between the two basic alternatives, with respect to their locations (e.g., regional versus local treatment) and the spatial variation of the lake's water quality.

Stratum 1

For the sake of completeness it has to be mentioned that WQMM could be thought of as being a part of a regional development policy model forming the top of the pyramid, a field which is beyond the scope of this study.

4. ILLUSTRATION OF THE DIFFERENT STEPS OF THE APPROACH

4.1 Wind Induced Sediment Water Interaction (Stratum 5)

For studying the sediment-water interaction in lakes, several approaches are possible (see, for example [24]). In this study, yet another method was chosen [25], in recognition that when eutrophication is considered, more than just the physical processes should be examined. Daily measurements were taken for 6 months, at the mid-point (depth H = 4.3 m) of the Szemes basin (Basin 2, Figure 1). The measurements involved Secchi depth, temperature, suspended solids (SS), Chlorophyll-a, and phosphorus fractions at different vertical locations. Wind velocity and direction were recorded continuously, from which hourly averages were calculated. The objective of the first part of the analysis was to describe the dynamics of the suspended solids as a function of wind. This then allowed for a characterization of the temporal changes in the light conditions, the deposition, and resuspension of other particulate material and to some extent, also the releases of dissolved components. Here only the behavior of SS will be reported.

The analysis started from a simplified transport equation for describing the temporal and vertical changes of the average SS concentration in the basin, neglecting inflow and outflow. It was recognized however, that the problem had an undefined boundary condition at the bottom, z = H, [25]

$$wc - E\,\frac{\partial c}{\partial t} = \phi_d - \phi_e \quad , \tag{1}$$

where c is concentration, w is settling velocity, E is vertical eddy viscosity, and ϕ_d and ϕ_e are the fluxes of deposition and resuspension, respectively. In fact, one of the objectives of the measurements was to formulate the boundary condition. From the observations made, it appeared that the temporal changes governed the system (see Figure 3a for the depth integrated values and wind speed). The c(z) vertical profiles were quite uniform, except close to the bottom where the expected, but sudden increase could be observed. Accordingly, it was decided not to determine the unknown boundary condition from the PDE formulation (a rather tedious porcedure), but to integrate the turbulent diffusion equation along the depth, and use the ODE derived, which thus directly involves the boundary condition itself.

In order to carry out this step, *hypotheses* were needed for the fluxes ϕ_d and ϕ_e. The deposition was characterized by its P probability (expressing which portion of the particles reaching the interface would remain there):

$$\phi_d = Pw\tilde{c} \quad , \tag{2}$$

(here, the wavy line indicates depth averaged value), while ϕ_e by an empirical relationship [18]

$$\phi_e = k\rho_w \frac{\rho_s}{\rho_s - \rho_w} w_e \quad , \tag{3}$$

where ρ_s and ρ_w are sediment and water densities, and w_e is entrainment velocity. To find w_e, the concept of energy transformulation between potential and turbulent kinetic energies often employed for stratified lakes was adopted [6]. Accordingly, under simplified conditions:

$$w_e \sim \frac{1}{H} W^n \quad , \tag{4}$$

where the power depends on the Richardson number.

Using these *hypotheses*, the depth integrated transport equation takes the following form [25]:

$$\frac{d\tilde{c}}{dt} = -K_1\tilde{c} + K_2W^n \quad , \tag{5}$$

where K_1 and K_2 comprise on the one hand parameters listed in part before, being approximately constant for a given situation, and unknown coefficients on the other hand, derived from the hypotheses (Equations 2 to 4). Consequently, the structure of the model should be identified and the parameter values, K_1, K_2, and n, estimated from measurements. The feasibility of Equation (5) can be appreciated from Figure 3a, which clearly shows the influence of the wind velocity on the concentration. However, a simple regression between the W and SS is not precise enough; the involvement of SS in a time series fashion improves it, thus suggesting the influence of settling and deposition.

First a deterministic estimation technique was adopted to derive the unknown coefficients which resulted in realistic values but without proving the correctness of the hypotheses (a posteriori model structure identification, see Beck in this issue [4]).

For this purpose, as a second step, the Extended Kalman Filter (EKF) method was applied [2, 5]. For the power n a value near to 1 was derived which corresponded to the small Richardson number [26]. Subsequently n was fixed to 1 since in this case the physical interpretation of the results is more obvious. The recursive estimation started from the estimates of the deterministic technique. The results are illustrated in Figure 3a. As is apparent, the agreement between observations and model calculation is reasonably good, and the parameters become approximately constant after the first 40-50 days (Figure 3b), proving that the *model structure* is *correct* [2, 4]. Some slight parameter changes can be observed at the end of the period; this may be caused, e.g., by the exclusion of inflow-outflow processes (or by other phenomena such as algal blooms). This suggests that the isolation of subprocesses is generally not complete. From the analysis, a realistic order of magnitudes follows for all the essential physical quantities; in this connection see [26].

As can be observed in Figure 3, for one month in the middle of the total period, no measurements were available, so the model was used for prediction. The appropriateness of the model is also illustrated by the fact that after getting new data, the parameter values did not change. This second period served for validation, following the identification and calibration procedure.

The study, which underlines the definite need to combine both theoretical and measurement knowledge, resulted in two basic achievements: (i) the estimation of the unknown boundary condition of a transport problem (which was then also solved by using an implicit finite difference method but the "submodel" was not maintained for the complex study as the vertical changes are not essential in this case from the point of view of eutrophication), and (ii) the description of the processes of deposition and resuspension through an ODE which can be easily incorporated into the biochemical submodel with a similar ODE structure.

In addition to the wind induced interaction discussed here, the sediment biochemistry is also of major importance, a field where further research is required.

4.2 Application of a 1D Hydrodynamic Model (Stratum 4)

The objective of the model has already been explained. The complete one-dimensional equation of momentum and continuity was solved by using a conventional implicit finite difference scheme [27]. For the matrix inversion, an effective decomposition technique was developed, resulting in economic computations [27]. Dynamic input is the longitudinal component of the wind force, while the output is the water level and streamflow rate at each cross section (Δx = 2000 m). The two parameters of the model (drag coefficient and bottom friction) were calibrated on the basis of the work of Muszkalay [20]. From the data of nearly ten years of observations, he derived empirical relationships between some typical wind parameters of a storm and the corresponding maximum denivellation of water surface and velocity at the Tihany peninsula (Figure 1), respectively. For validation, more than ten historical events were selected. The results of one example are shown in Figure 4. This event can be characterized by a long-lasting longitudinal wind followed by smaller shocks of different directions (Figure 4a). It is apparent that the agreement between measured and simulated water levels is satisfactory (Figure 4b). The discharge shows a striking oscillation between -2000 and 3000 m^3/s (Figure 4c: associated with the seiche phenomenon) which is higher by two orders of magnitude than the hydrologic throughflow and may cause considerable fluctuation in the volume of the basins (Figure 1). The seiche phenomenon alone will certainly not cause mixing and its influence can be judged only by the method described in Section 3.2. Still, its effect can be illustrated from another angle: at a given location the rapid seiche motion will cause an oscillation of various constituents within a day, which strongly depends also on the longitudinal gradient. This may result in quite a critical *error* in the concentration determined through *instantaneous sampling*, a recognition which will allow definition through the model of an uncertainty range of historical measurements. To find a more satisfactory agreement than in the previous example is often impossible. The reason is quite simple: a small error in the wind direction may cause a drastic change in the wind force, if the direction is far from the longitudinal one. In fact, there are many kinds of *uncertainties* in the *wind direction*, such as random fluctuation (turbulence), the influence of hills on the northern side of the lake, which cause nonuniformities in the wind field, measurement errors, etc. Figure 5 illustrates the case (transversal wind conditions). A deterministic simulation did not prove acceptable. Bearing in mind the possible role of uncertainties, a random component was subsequently added to the wind direction (Gaussian distribution, zero mean, 12.5° standard deviation: a modest value) and a Monte Carlo simulation was performed. Figure 5, which shows the results of all the 100 runs, does not require detailed discussion: it shows the extreme sensitivity to input data uncertainty (compared to this the parameter sensitivity is negligible) and illustrates how difficult it is to validate a deterministic model (to a lesser extent this is also true for a more-dimensional model). This behavior also suggests that some time averaging should be introduced for the transport model part. According to our analysis, the mean and variance of the flow rate time series on a daily basis shows limited sensitivity only. Since generally one is not interested in short-term concentration changes (the character of the sampling problem mentioned before is different) this allows use of the mean value for convection in the dispersion model, while the effect behind the variance can be incorporated to the dispersion coefficient.

The example suggests that although a deterministic hydrodynamic model can hardly be verified in a strict sense, the same can be done for a transport model in a water quality study by filtering out the influence of uncertainties.

4.3 The Nutrient Load under Uncertainty and Stochasticity (Stratum 3)

In accordance with the nutrient loading estimate done [15], more than 40% of the total phosphorus (TP) load reaches the lake through tributaries. As is well known, floods play a decisive role in the yearly total transport, their contribution ranging generally between 70-90%. This fact is in most of the cases not reflected in the monitoring strategy; generally one or two observations made monthly at the mouth of the river are available. Thus the influence of floods remains unobserved. The infrequent data collection is characteristic for 19 of the 20 tributaries of Lake Balaton, while for the major pollution source, the River Zala which accounts for more than half of the tributary load, daily measurements were performed (1975-79, [16]).

Since from this data set the "accurate" load for different averaging periods (such as a month or year) can be derived, it allows one to study the *error* caused by *scarce observations*. The procedure is a straightforward Monte Carlo type technique which starts with a random selection on the detailed data set following the sampling strategy of the other tributaries and calculates the load of the period in question. After making a sufficient number of random selections the statistical parameters of the load can be determined. The results for the long-term monthly average load (on the basis of a four year long observation period and 200 data selections) are illustrated in Figure 6. The choice of a month was made for two reasons: (i) being in possession of two monthly observations for a period of several years only the monthly average load can be estimated at best; (ii) from a sensitivity study on LEM [28], it turned out that it is sufficient to use this load type as a forcing function. As can be seen from Figure 6, which shows the mean and extreme values, as well as the domain of ± standard deviation, the error is quite high and its fluctuation follows the change in the mean value. On the basis of this study and a similar analysis for the yearly averages, the annual load of other tributaries was corrected (here the similarities of major subwatersheds were also examined, but the correction certainly has an extrapolative character) and a random component was added to the monthly average load component [28].

To incorporate the *stochastic influence* of the hydrologic regime a regression analysis was done on the data set of River Zala. It was found that the monthly average TP load correlated satisfactorily with the corresponding stream flow rate. Accepting the statistics of the monthly average stream flow from long-term observations [1] the load can be calculated in a stochastic fashion. Figure 7 shows the characteristics of the load pattern for 1976-79 and the 90% confidence levels derived. For illustrating the influence of the hydrologic regime an event of low probability in July 1975, is likewise indicated.

As a final output of the research outlined in this section a load scenario generator was developed for the whole lake, which accounted for both uncertainty and stochasticity, discussed above. For further details see [28].

It is noted here that using historical data, a similar analysis was made on climatic (uncontrollable) factors, which allowed the water temperature and daily radiation to be generated in harmony with each other, in a random fashion [28]. Thus, *future scenarios* can be generated for all the essential forcing functions of the lake model--an essential tool for *planning purposes*.

4.4 The Lake Eutrophication Model (Stratum 3)

The preliminary results gained with the simplest model, SIMBAL [31], developed for Lake Balaton are given below. The model is a phosphorus cycle model, that is, all the state variables (the essentials are two algal groups, detritus, and dissolved inorganic phosphorus) are expressed in terms of phosphorus, for the four basins indicated in Figure 1 (see Section 3.2). A Monte Carlo simulation is incorporated into the model to find areas in parameter space where the model produces results fully within specified boundaries drawn around the data to account for data uncertainty (see, e.g., Section 3.2) and thus easily applicable to test various hypotheses ([31] and also [9, 14]).

Among the calibration runs, results for the phytoplankton phosphorus, PPP, for the four basins are given in Figure 8 (as 1977 forcings data was used) together with the corresponding observation variable, Chlorophyll-a (basin average values). It is pointed out that Chlorophyll-a and PPP cannot be directly compared to each other; however, since a more or less linear measurement equation is expected among them, PPP should follow the pattern of Chlorophyll-a: a trend which can be generally observed. For illustration the standard deviation around the trajectory for Basin 2 estimated through the Monte Carlo simulation is also indicated (*parameter uncertainty*). Further discussion on the calibration and model improvement required can be found in [31].

For management purposes historical data cannot be used. Either some critical, unfavorable environmental conditions should be introduced or the model should be considered stochastic through input data which are basically random variables when future planning is in question. Here the latter approach was adopted and the generators outlined in the previous section coupled to the lake model. Two essential results for Basin 1 are presented in Figures 9 and 10.

In the first case, uncertainties caused by natural factors were considered and the 1977 load was maintained. The summary of 100 Monte Carlo runs (mean, ± standard deviation, and the extremes of PPP) suggests the relatively large sensitivity of the lake's water quality to meteorological factors and explains the essential year to year changes observed in the behavior of the lake even when the load remained unaffected. The second case (Figure 10) involved the *random generation* of both *natural* and *controllable* factors. While for the previous example the specific 1977 load was adopted, here the mean load of the input generator was derived from data for the period 1975-79 (Section 4.3). This is the reason why the average trajectory shown in Figure 10 differs from that shown in Figure 9. The inclusion of load randomness had an obvious influence: the range of uncertainty of the water quality simulation results became much wider.

Compared to the role of parameter uncertainty (Figure 8), the meteorological factors represent the same order of magnitude, while the contribution of load to the model uncertainty is twice as high. It is stressed that in the frame of the present example, no load reduction was employed. Thus Figure 10 shows in which domain the water quality may range under the present conditions, since the changes in trend are already relatively small from year to year. Control alternatives will certainly modify not only the mean load but also the related uncertainties (e.g., the smoothing effect of a retention pond or the uncertainty of the effectiveness η, of a given management option (Figure 11), an issue which should be taken into account on the level of decision making.

4.5 The Incorporation of the Lake Model into WQMM (Stratum 2)

The question that we are planning to answer here was raised in Section 3. In the course of the analysis outlined subsequently, the different parameters of the PPP(t) distribution were selected as water quality indicators characterizing the algal behavior (see 3.2), and deterministic simulations were performed with the dynamic lake model, SIMBAL, under reduced loading conditions. It turned out from the study that the lake's response expressed in terms of the yearly average or peak of PPP is quite linear on the load in a wide range [28] (the same experiences were also gained with the two other models). It is noted that a similar linearity is expressed for total phosphorus, TP, by several empirical models [33]. However, for shallow lakes, TP may not properly characterize the process of eutrophication since this component is strongly influenced by wind induced interaction at the bottom (Section 4.1).

The recognition of the linearity leads to an important *aggregation*: the *dynamic lake model* can be *replaced* at the level of WQMM (see Figure 2) by a simple *linear equation* (see Figure 11):

$$\underline{c} = \underline{c}_0 + \underline{\underline{A}} \; (\underline{L} - \underline{L}_0) \quad , \tag{6}$$

or

$$\Delta\underline{c} = \underline{\underline{A}} \; \Delta\underline{L} \quad . \tag{7}$$

Here $\underline{L}_0$ and $\underline{c}_0$ are the initial load and concentration "vectors" respectively, defined by the number of uniform segments assumed in LEM (at present N = 4, Figure 1); $\Delta\underline{L}$ represents the reduction of load achieved by various control alternatives, while $\Delta\underline{c}$ is the corresponding long-term response of the lake (expressed in terms of water quality indicators) characterizing the new equilibrium of the system. The elements of the $\underline{\underline{A}}$ matrix maintain the essence of LEM: again a transition from a "large" model to a "smaller" on a higher level of the hierarchy of Figure 2. The first remarkable feature of equations (6) and (7) is that they clearly preserve the influence of subprocesses on the lower strata and show the subsequent steps of aggregation. The elements of the main diagonal are determined primarily by biochemical processes and sediment-water interaction (stratum 5), while the other elements mainly express the influence of hydrodynamics and associated mass exchange (stratum 4), showing that a management action at the region of the i^{th} segment will affect the water quality of other segments as well (Figure 11). It is briefly noted that through LEM (stratum 3), the uncertainties discussed in the previous section can also be included in equation (6).

The second remarkable feature of equation (6), and this is of primary importance on stratum 2, is that its linear structure allows its direct involvement in an optimization framework: a solution that we were looking for (see Section 3.2).

At the end of this section, it is noted that optimization of this kind is not an easy task. First of all each $L_{0,i}$ element is composed of various types of nutrients (sewage, diffuse sources, etc.) which in a general case has different spatial locations within the corresponding region. Thus $L_{0,i}$ itself is a vector (in other words, several homogeneous segments of the watershed may belong to a given i^{th} uniform segment of the water body). Second, there are several options for loading reduction and the same ΔL_i could be arrived at by different control strategies. In addition, just as with the linkage of the c_i elements through in-lake mass exchange processes (Figure 11), the $L_{1,i}$ components are coupled through realization of the desired management strategy implemented in the watershed (e.g., regional treatment and the related sewerage system). All these are features of the problem that should be accounted for at the level of WQMM. In fact, such an approach is also being elaborated within the case study [17].

5. CONCLUDING REMARKS

In the paper a system characterized by two major factors was considered: (i) complexity due to the multifarious interactions of in-lake and watershed processes, respectively, and different levels of interest such as scientific understanding and decision making, and (ii) the uncertainty of various sources. The objective is also twofold, not only to study the specific issues of this system, but also to answer general methodological questions for systems with similar features.

The modeling strategy adopted is an off-line procedure where the complex structure is decomposed into units, tackled separately in detail, but preserving and coupling only the essentials on the higher level of the modeling hierarchy. The method is characterized by several steps of aggregation (based on sensitivity analyses in order to find the relative importance of various subprocesses), and validation (detailed models against field data and the aggregated models against their detailed versions and data on the higher level).

A harmony is also needed with regard to the methodologies employed; none of them should exclude the application of other techniques required after the coupling and they should allow the proper combination of theory and observations in the model. Uncertainty should be involved at each step and also aggregated simultaneously with the off-line construction of the complex model.

Admittedly the coupling, aggregation, and validation of the linked model are the crucial points in the development: it is not certain whether the simplifications planned can be performed and whether a model of realistic structure and complexity can be arrived at on the higher level.

Although the study is not yet finished and future research is certainly required (the role of sediment, completion of coupling the segment models and that of the water quality management part) the individual examples introduced show that the modeling principles outlined here can be successfully adopted for the different strata and suggest that this is also valid for the entire model incorporating the level of decision making. It is believed, too, that other systems characterized by complexity and uncertainty can be treated in a similar way; thus models of reasonable size and structure can be gained.

REFERENCES

1. S. Baranyi, Hydrological Conditions of Lake Balaton, in "Summary of the research results related to Lake Balaton", edited by S. Baranyi, VIZDOK, Budapest (in Hungarian), pp. 128-177, (1979).

2. M.B. Beck, System Identification, Estimation, and Forecasting of Water Quality, Part I: Theory. International Institute for Applied Systems Analysis, WP-79-31, p. 123. (1979).

3. M.B. Beck, Hard or Soft Environmental Systems? Ecological Modelling, 11, pp. 233-251 (1981).

4. M.B. Beck, Identifying Models of Environmental Systems' Behaviour, J. of Mathematical Modelling (1981) (forthcoming).

5. M.B. Beck and L. Somlyódy, Identification of a model for wind induced sediment-water interaction in Lake Balaton. International Institute for Applied Systems Analysis, WP-81- (1981) forthcoming.

6. S. Bloss and D.R.F. Harleman, Effect of wind-mixing on the thermocline formation in lakes and reservoirs. Massachusetts Institute of Technology, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Report No. 249, p. 68, (1979).

7. F.M. Boyce, A.S. Fraser, E. Halfon, D. Hyde, D.C.L. Lam, W.M. Schertzer, A.H. El-Shaarawi, T.J. Simons, K. Willson, D. Warry and T.J. Simons. Assessment of Water Quality Simulation Capability for Lake Ontario. Scientific Series No. 111. Inland Water Directorate, National Water Research Institute, Canada Center for Inland Waters: Burlington, Ontario, p. 220.

8. P. Eykhoff, System identification - parameter and state estimation, Wiley, Chichester, p. 555 (1974).

9. K. Fedra, G. van Straten and M.B. Beck, Uncertainty and arbitrariness in ecosystems modelling: a lake modelling example, Ecological Modelling, 13, pp. 187-210 (1981).

10. M.B. Fischer, J. Imberger, E.J. List, R.C.Y. Koh and N.H. Brooks, Mixing in Inland and Coastal Waters, Academic Press, p. 483 (1979).

11. D.R.F. Harleman and P. Shanahan, Aspects of Wind-driven Circulation and Mixing in Eutrophication Studies of Lake Balaton. Proceedings of the Second Joint MTA/IIASA Task Force Meeting on Lake Balaton Modeling. G. van Straten, S. Herodek, J. Fischer, and I. Kovacs (Editors). MTA VEAB, Veszprem, Vol. 1, pp. 50-67 (1980).

12. S. Herodek and G. Tamas, The Primary Production of Phyto-Plankton in the Keszthely-basin of Lake Balaton in 1973-74, Annual. Biol. Tihany Vol. 42, pp. 175-190 (1975).

13. S. Herodek, P. Csaki and T. Kutas, The BEM Modeling Approach I-III, Proceedings of the Second Joint MTA/IIASA Task Force Meeting on Lake Balaton Modeling. G. van Straten, S. Herodek, J. Fischer, and I. Kovacs (Editors). MTA VEAB, Veszprem, Vol. 1, pp. 81-103 (1980).

14. G.M. Hornberger and R.C. Spear, Eutrophication in Peel Inlet-I. Problem-defining Behaviour and a Mathematical Model for the Phosphorus Scenario. Water Research 14, pp. 29-42 (1980).

15. O. Joó, Data for the Eutrophication of Lake Balaton and Considerations related to Control Activities. Proceedings of the Second Joint MTA/IIASA Task Force Meeting on Lake Balaton Modeling. Veszprem, Vol. 2, pp. 139-166 (1980).

16. G. Jolànkai and L. Somlyody, Nutrient Loading Estimate for Lake Balaton. International Institute for Applied Systems Analysis, CP-81-21, p. 56 (1981).

17. L.B. Kovács, E. Dobolyi and F. Inotay, Background for developing a Sewage Management Model for the Lake Balaton Region (in Hungarian), Computer and Automation Institute of the Hungarian Academy of Sciences (SZTAKI) Working Paper MO/17, p. 27 (1980).

18. D.C.L. Lam and M. Jaquet, Computation of physical transport and regeneration of phosphorus in Lake Erie, Journal of Fisheries Research Board of Canada, Vol. 33, No. 3, pp. 550-563 (1976).

19. A. Leonov, Simulation and Analysis of Phosphorus Transformation and Phytoplankton Dynamics in Relation to the Eutrophication of Lake Balaton, Proceedings of the 2nd State of the Art Conference in Ecological Modelling, Liège, Belgium and also International Institute for Applied Systems Analysis WP-80-88, p. 30 (1980).

20. L. Muszkalay, Water Motions in Lake Balaton. In a summary of research results for Lake Balaton. Edited by S. Baranyi VIZDOK, Budapest (in Hungarian), pp. 34-128 (1979).

21. D. Scavia and A. Robertson, Perspectives on Lake Ecosystem Modeling, Ann Arbor Science, p. 326 (1979).

22. P. Shanahan, D.R.F. Harleman and L. Somlyódy, Modeling Wind-Driven Circulation in Lake Balaton, International Institute for Applied Systems Analysis, WP-81-7, p. 166 (1981).

23. P. Shanahan, Application of a Coupled Hydrophysical-Biochemical Eutrophication Model to Lake Balaton. R.M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Ph.D. thesis, in preparation (1981).

24. T.P. Sheng and W. Lick, The Transport and Resuspension of Sediments in a Shallow Lake, in Journal of Geophysical Research, Vol. 84, No. C4, April, pp. 1809-1826 (1979).

25. L. Somlyódy, Preliminary Study on Wind-Induced Interaction between Water and Sediment for Lake Balaton (Szemes Basin). In Proceedings of the Second Joint MTA/IIASA Task Force Meeting on Lake Balaton Modeling, pp. 26-50 (1980).

26. L. Somlyódy, Water Quality Modeling: A Comparison of Transport Oriented and Biochemistry Oriented Approaches, General Report. International Conference on Numerical Modeling of River, Channel and Overland Flow for Water Resources and Environmental Applications, Bratislava. March 4-8, p. 40 (1981).

27. L. Somlyódy and M. Virtanen, Application of a One-dimensional Hydrodynamic Model to Lake Balaton. International Institute for Applied Systems Analysis, CP in preparation (1981).

28. L. Somlyódy and J. Eloranta, The Use of a Lake Eutrophication Model in the Management Framework. International Institute for Applied Systems Analysis, WP in preparation (1981).

29. G. van Straten, G. Jolánkai and S. Herodek, Review and Evaluation of Research on the Eutrophication of Lake Balaton - A Background Report for Modeling. International Institute for Applied Systems Analysis, WP-79-13, p. 94 (1979).

30. G. van Straten and L. Somlyódy, Lake Balaton Eutrophication Study: Present Status and Future Program. International Institute for Applied Systems Analysis, WP-80-187, p. 89 (1980).

31. G. van Straten, Analysis of Model and Parameter Uncertainty in Simple Phytoplankton Models for Lake Balaton. Proceedings of the 2nd State of the Art Conference in Ecological Modelling, Liege, Belgium, and also International Institute for Applied Systems Analysis, WP-80-139, p. 28 (1980).

32. G. van Straten and S. Herodek, Estimation of Algal Growth Parameters from Vertical Primary Production Profiles, papers submitted to Ecological Modelling and also International Institute for Applied Systems Analysis, CP-81-19, p. 35 (1981).

33. R.A. Vollenweider, Advances in Defining Critical Loading Levels for Phosphorus in Lake Eutrophication, Mem. Ist. Ital. Idrobiol., 33, pp. 53-83 (1976).

FIGURE CAPTIONS

Figure 1. Major characteristics of the system
K - Keszthely, T - Tihany, S - Siófok
I...IV typical basins of the lake
-.- boundary of the catchment
=== Boundary of the recreational area
• sewage discharges in the region

Figure 2. The method of decomposition: hierarchy of models
(1) submodels for uniform segments and
(2) coupling of the submodels

Figure 3a. Identification and parameter estimation of a model for wind induced sediment-water interaction for Lake Balaton: recursive estimate of the suspended solids concentration; * Observations

Figure 3b. Recursive parameter estimates for the sediment-water interaction model

Figure 4. Simulation of a historical event: longitudinal wind conditions

Figure 5. The influence of wind data uncertainty on the discharge at the Tihany peninsula (T = 0 corresponds to 8.7.1963,
• discharge derived from measurements [19]

Figure 6. Monthly average TP load: uncertainty caused by scarce observations (River Zala, 1976-79); 3 - mean value; 4 and 2 - ± standard deviation; 5 and 1 - extreme values

Figure 7. Influence of the hydrologic regime on the monthly average load, River Zala; o 90% confidence level

FIGURE CAPTIONS CONTINUED

Figure 8. Results from SIMBAL. Comparison of field data for four basins (left) and average model for runs satisfying the behaviour definition (right). Adopted from van Straten [28, 29]

Figure 9. The influence of meteorologic factors on the water quality. Basin 1

3 - mean value, 4 and 2 - ± standard deviations

5 and 1 - extreme values

Figure 10. The combined influence of uncertainty and stochasticity in the meteorology and loading, respectively, on the water quality. Basin 1

3 - mean value, 4 and 2 - ± standard deviations,

5 and 1 - extreme values

Figure 11. Incorporation of the lake model in WQMM

i - uniform segments in the lake

$\leftrightarrows$ - interaction among basins through mass exchange and hydrologic throughflow

$\leftrightarrow$ - interaction among load components through control activities

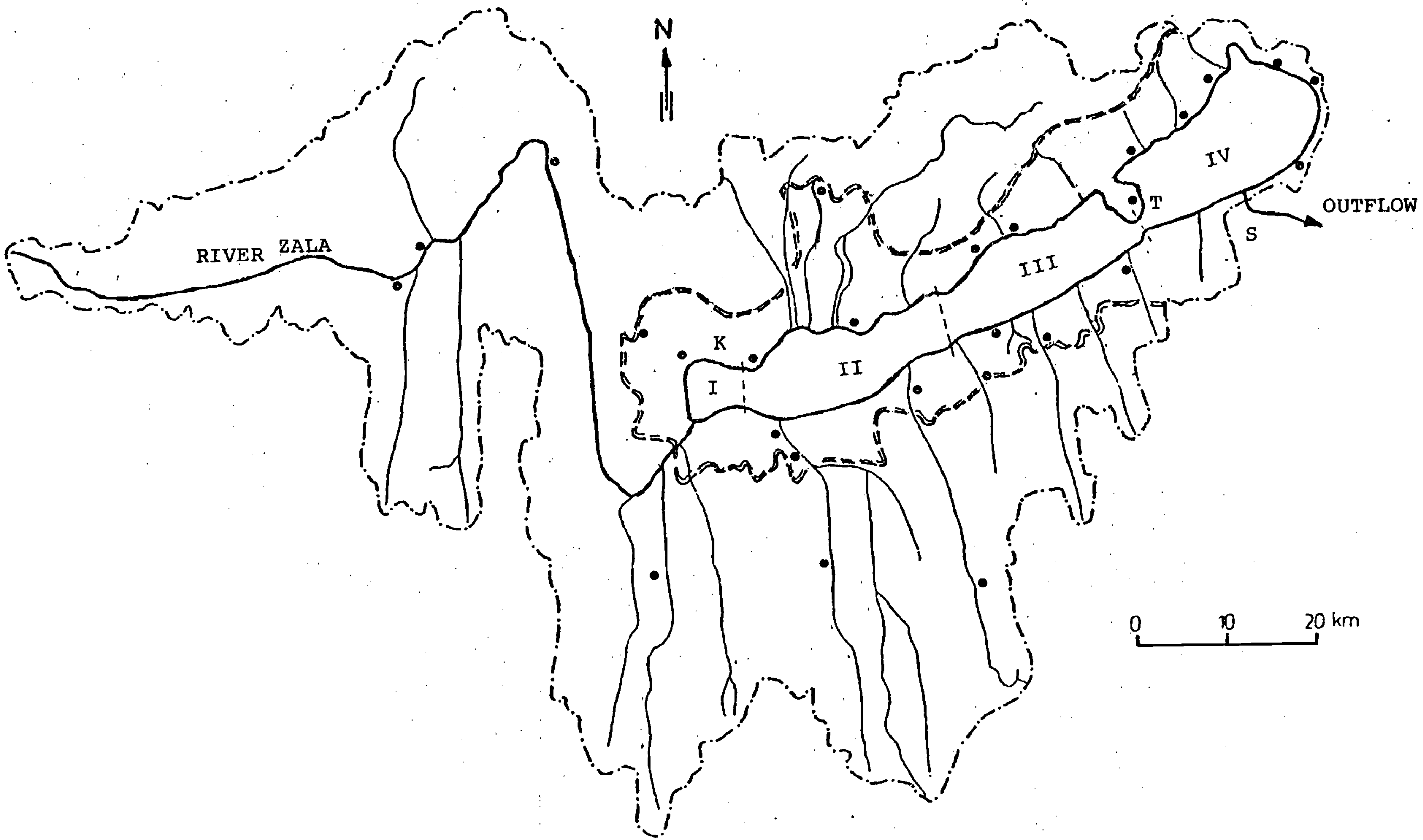

Figure 1. Major characteristics of the system
K - Keszthely, T - Tihany, S - Siófok
I...IV typical basins of the lake
-.- boundary of the catchment, === boundary of the recreational area
• sewage discharges in the region

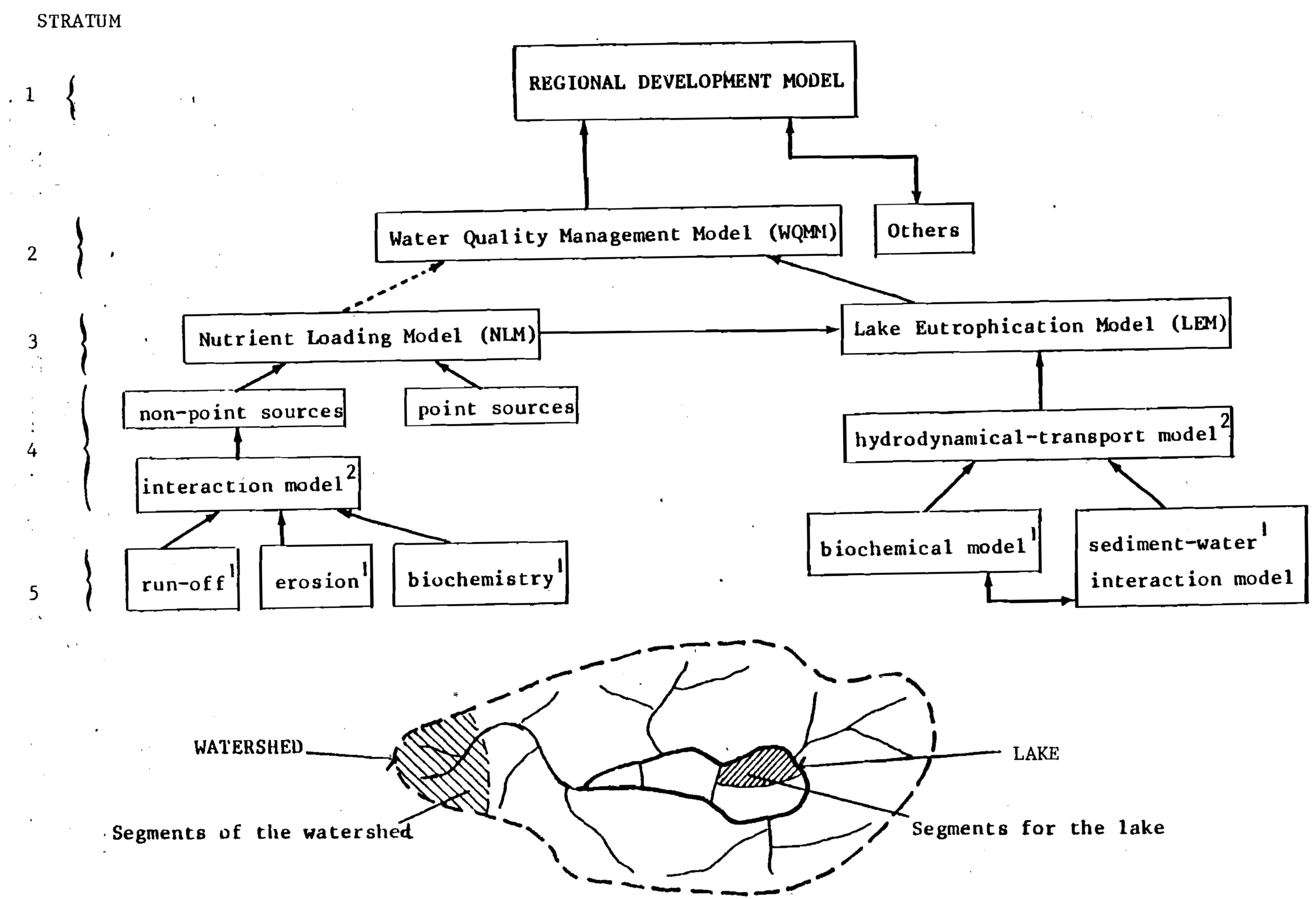

Figure 2. The method of decomposition: hierarchy of models
(1) submodels for uniform segments and 2) coupling of the submodels

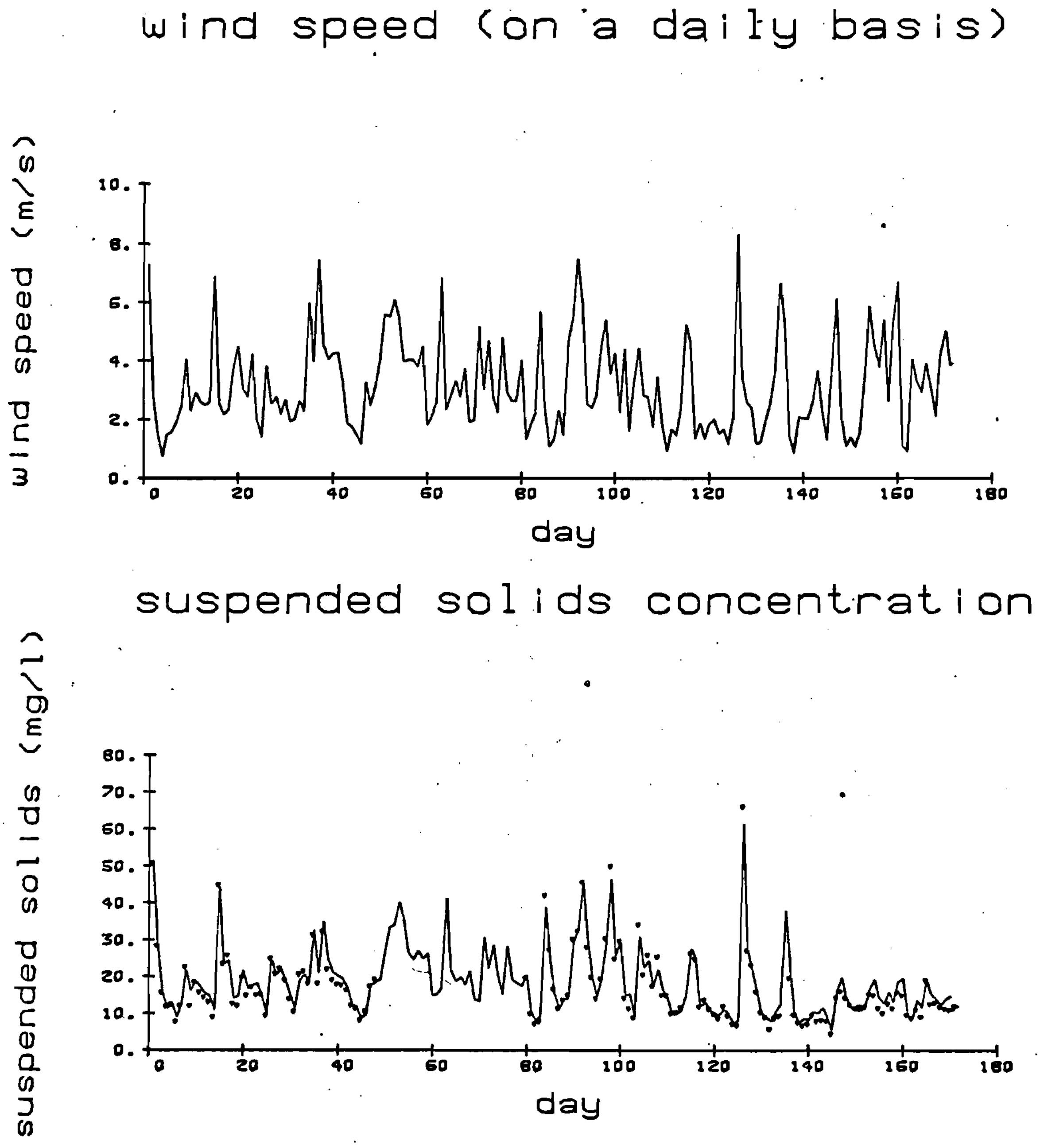

Figure 3a. Identification and parameter estimation of a model for wind induced sediment-water interaction for Lake Balaton: recursive estimate of the suspended solids concentration

* Observations

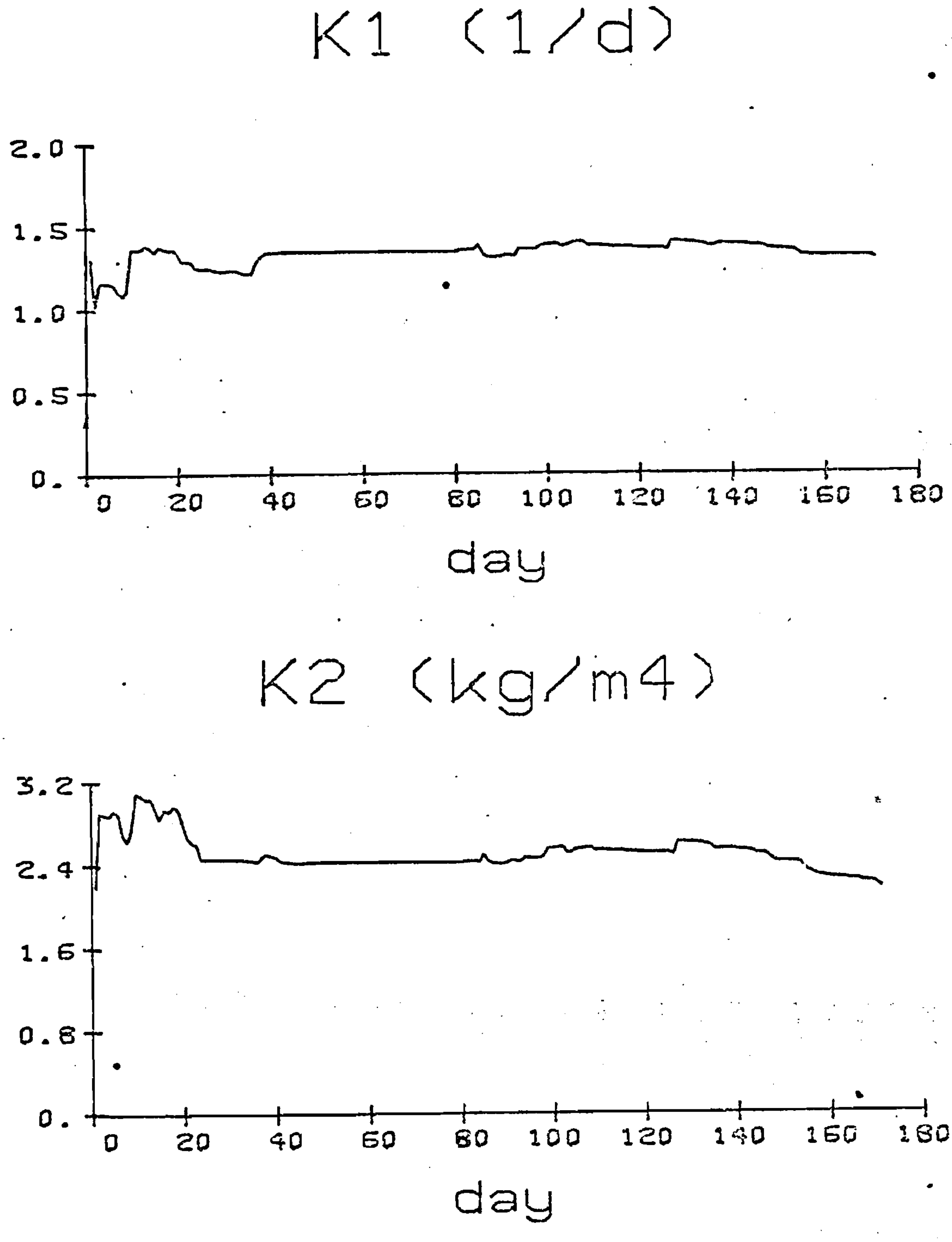

Figure 3b. Recursive parameter estimates for the sediment-water interaction model

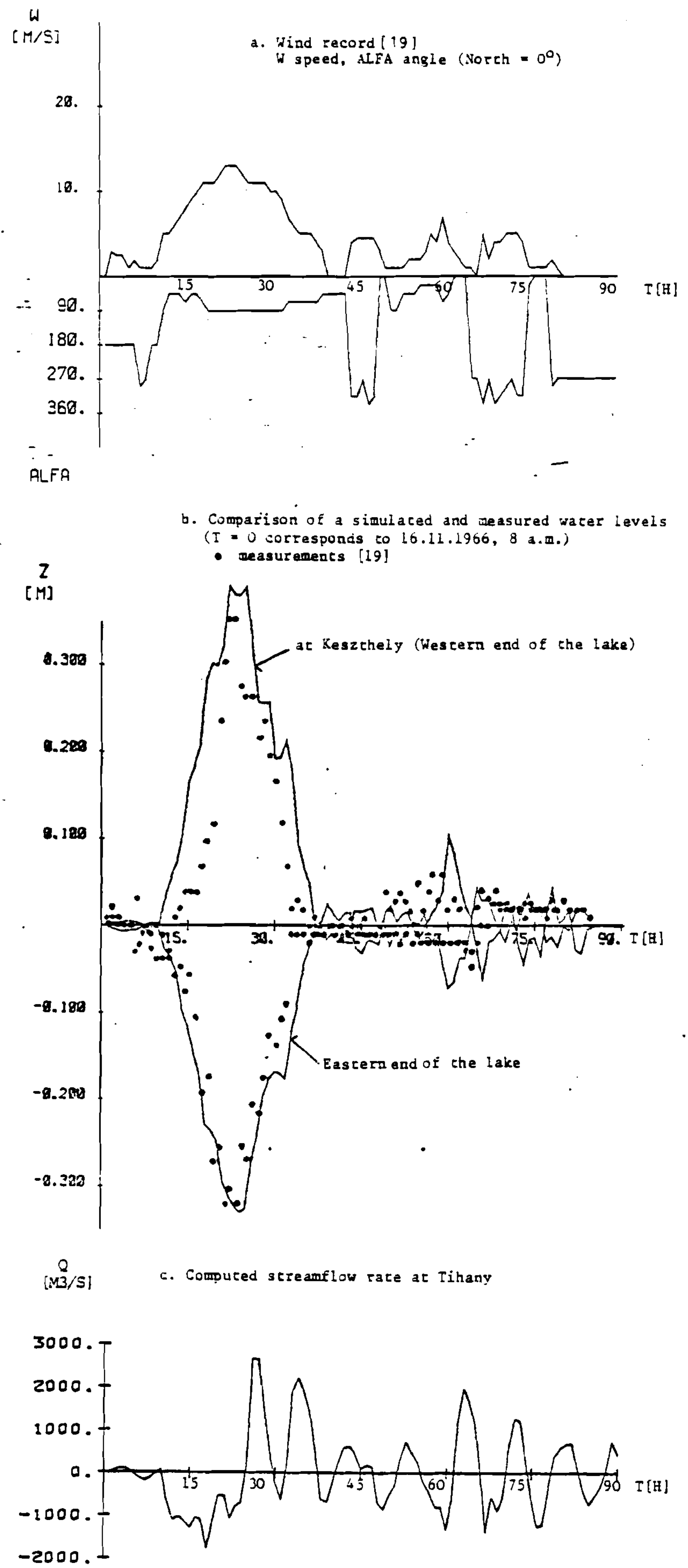

Figure 4. Simulation of a historical event: longitudinal wind conditions

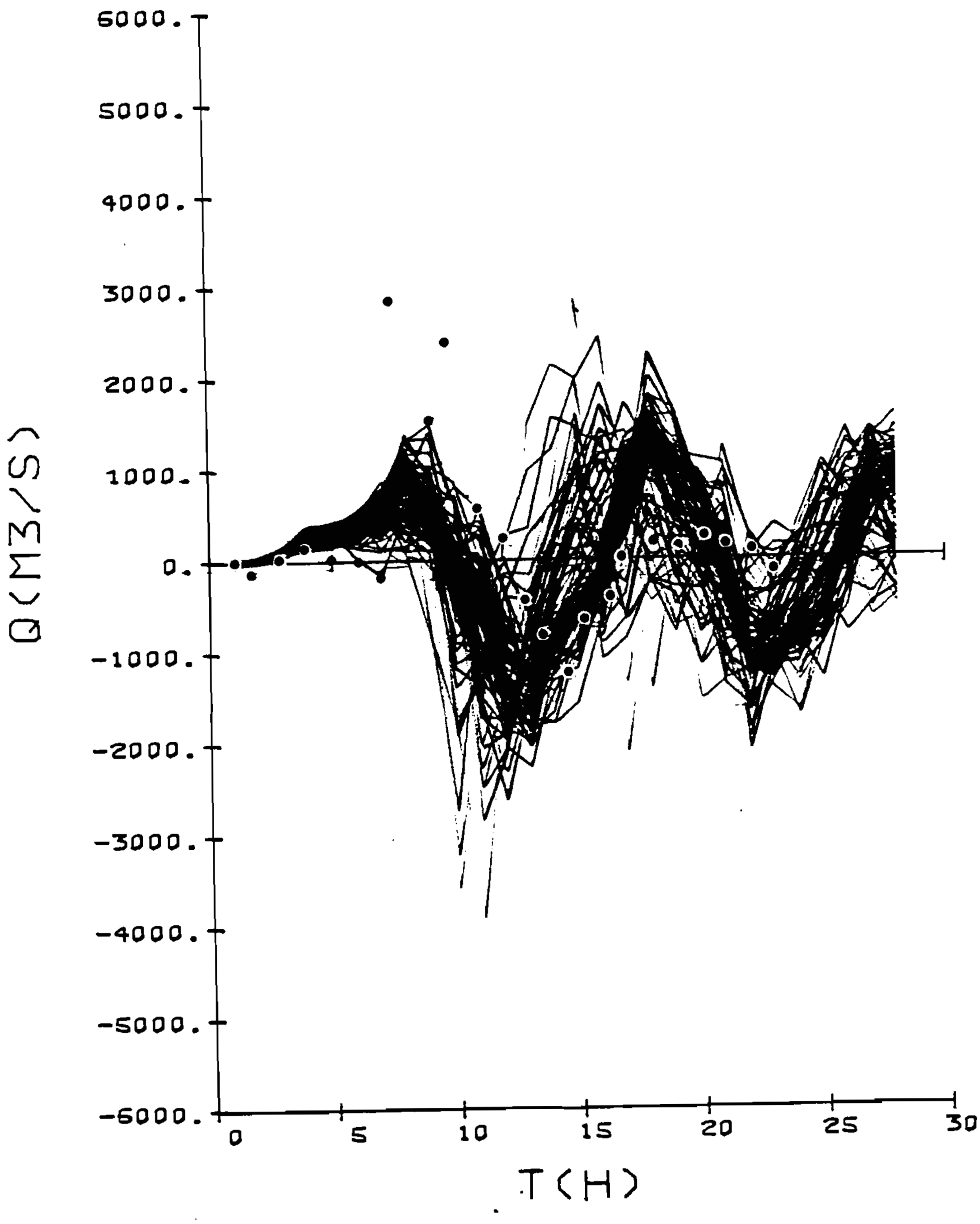

Figure 5. The influence of wind data uncertainty on the discharge at the Tihany peninsula (T = 0 corresponds to 8.7.1963, 8 a.m.)
• discharge derived from measurements [19]

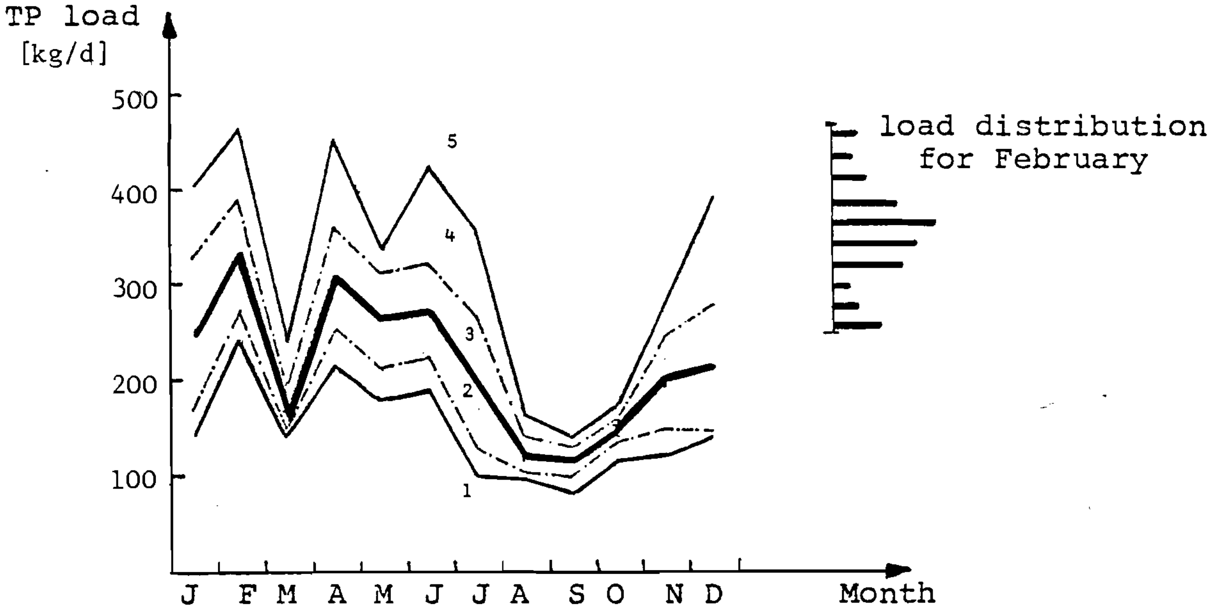

Figure 6. Monthly average TP load: uncertainty caused by scarce observations (River Zala, 1976-79); 3 - mean value; 4 and 2 - ± standard deviation; 5 and 1 - extreme values

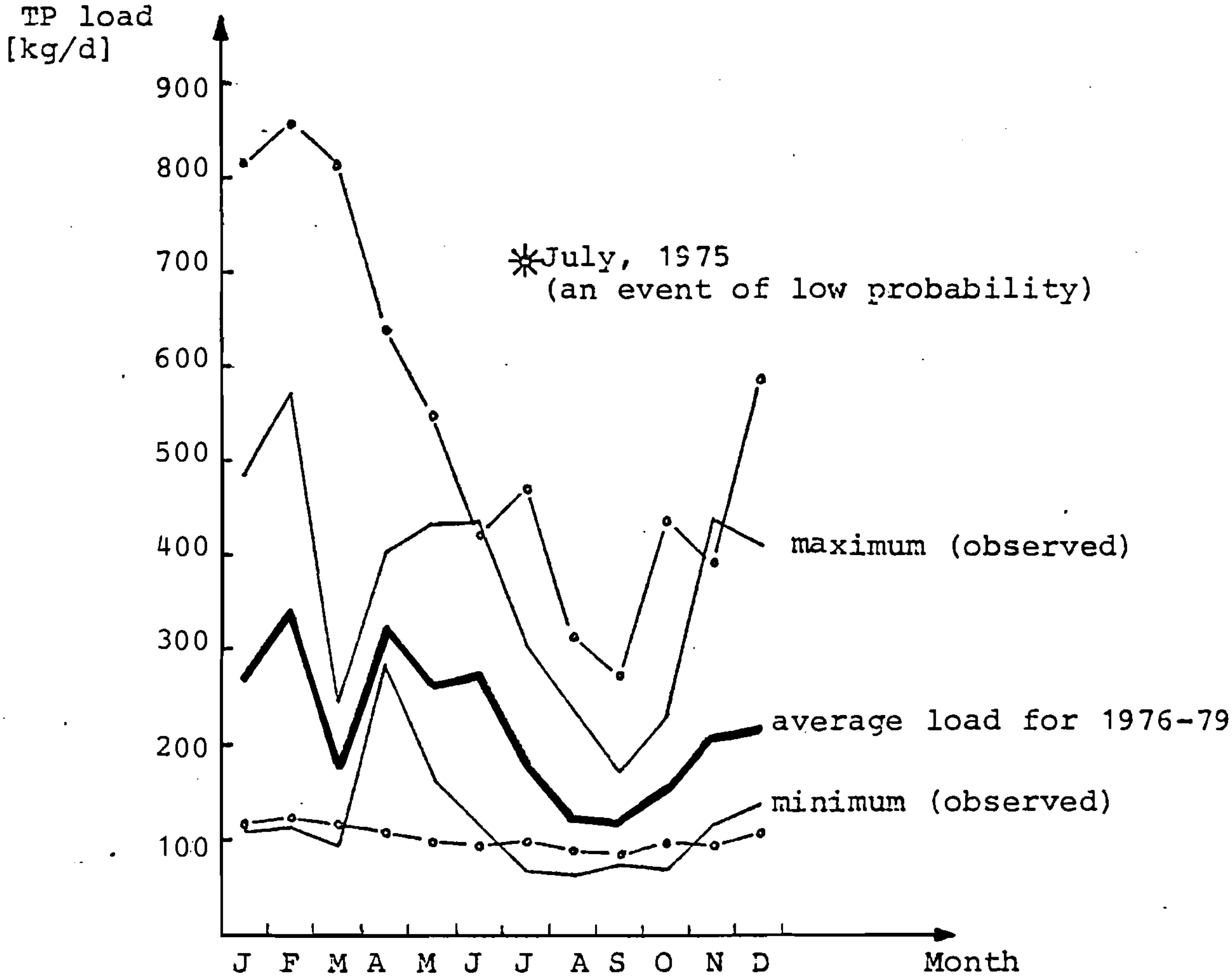

Figure 7. Influence of the hydrologic regime on the monthly average load, River Zala

o 90% confidence level

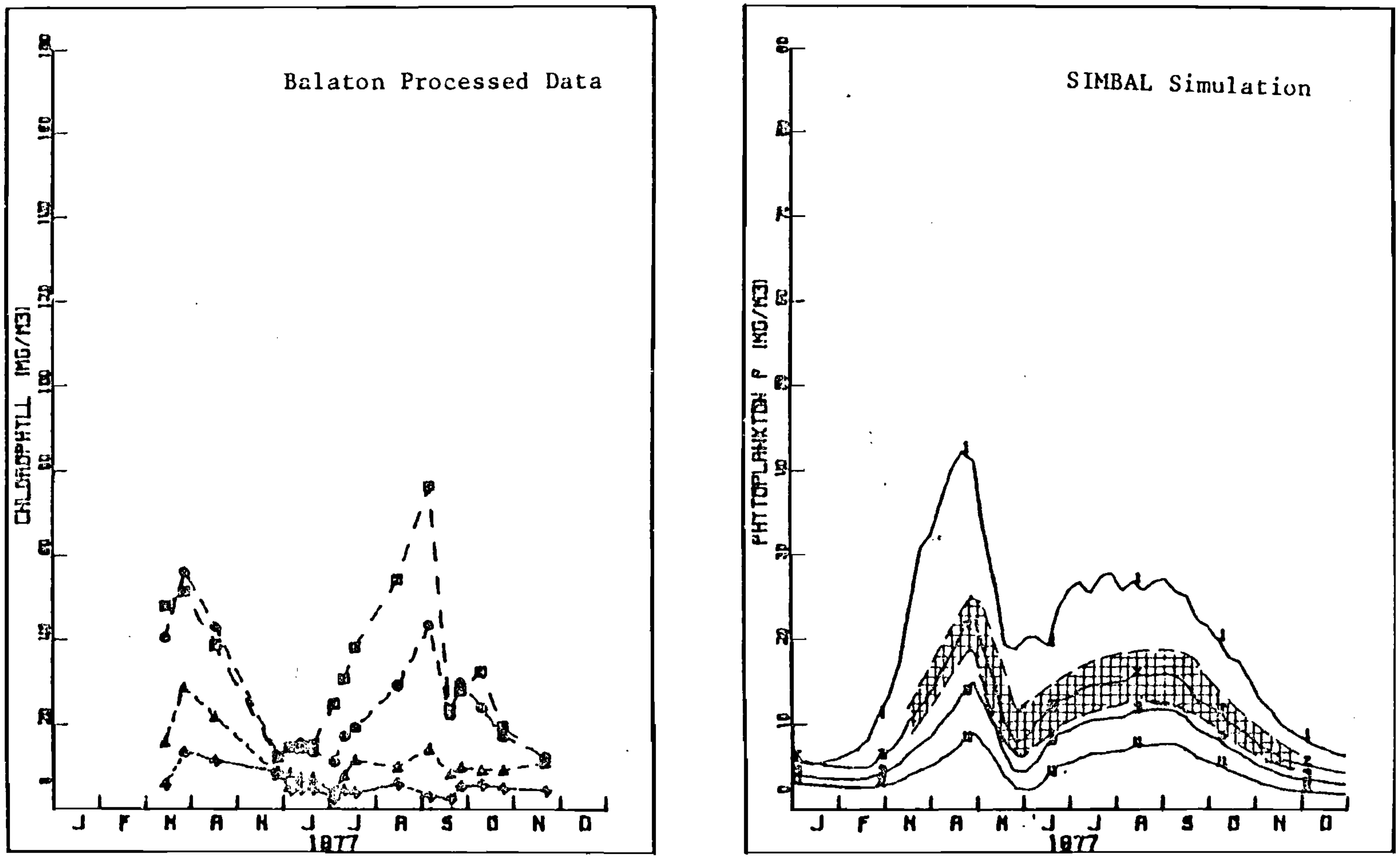

Figure 8. Results from SIMBAL. Comparison of field data for four basins (left) and average model simulations for runs satisfying the behaviour definition (right). Adopted from van Straten [28, 29].

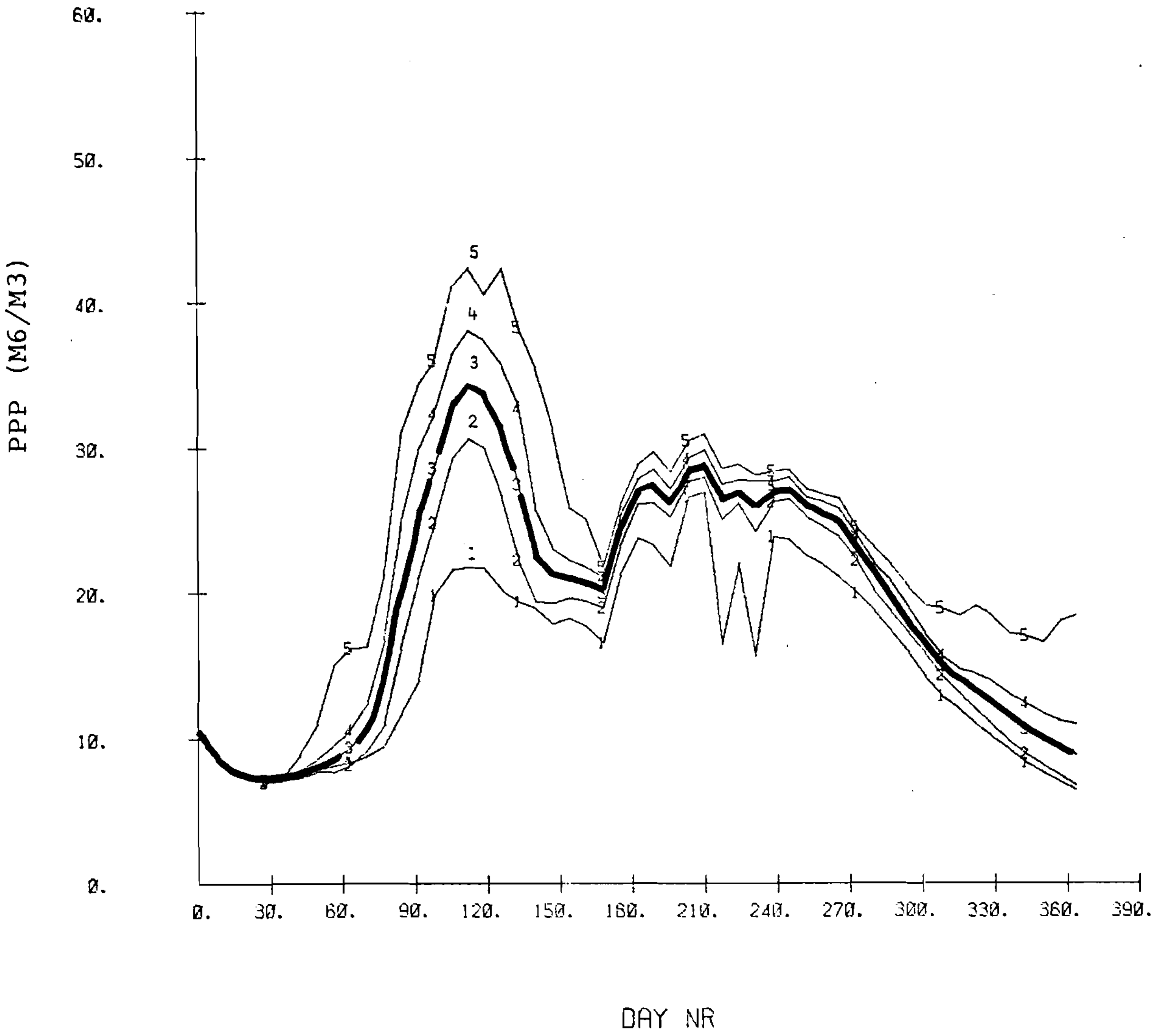

Figure 9. The influence of meteorologic factors on the water quality. Basin 1.
3 - mean value, 4 and 2 - ± standard deviations
5 and 1 - extreme values

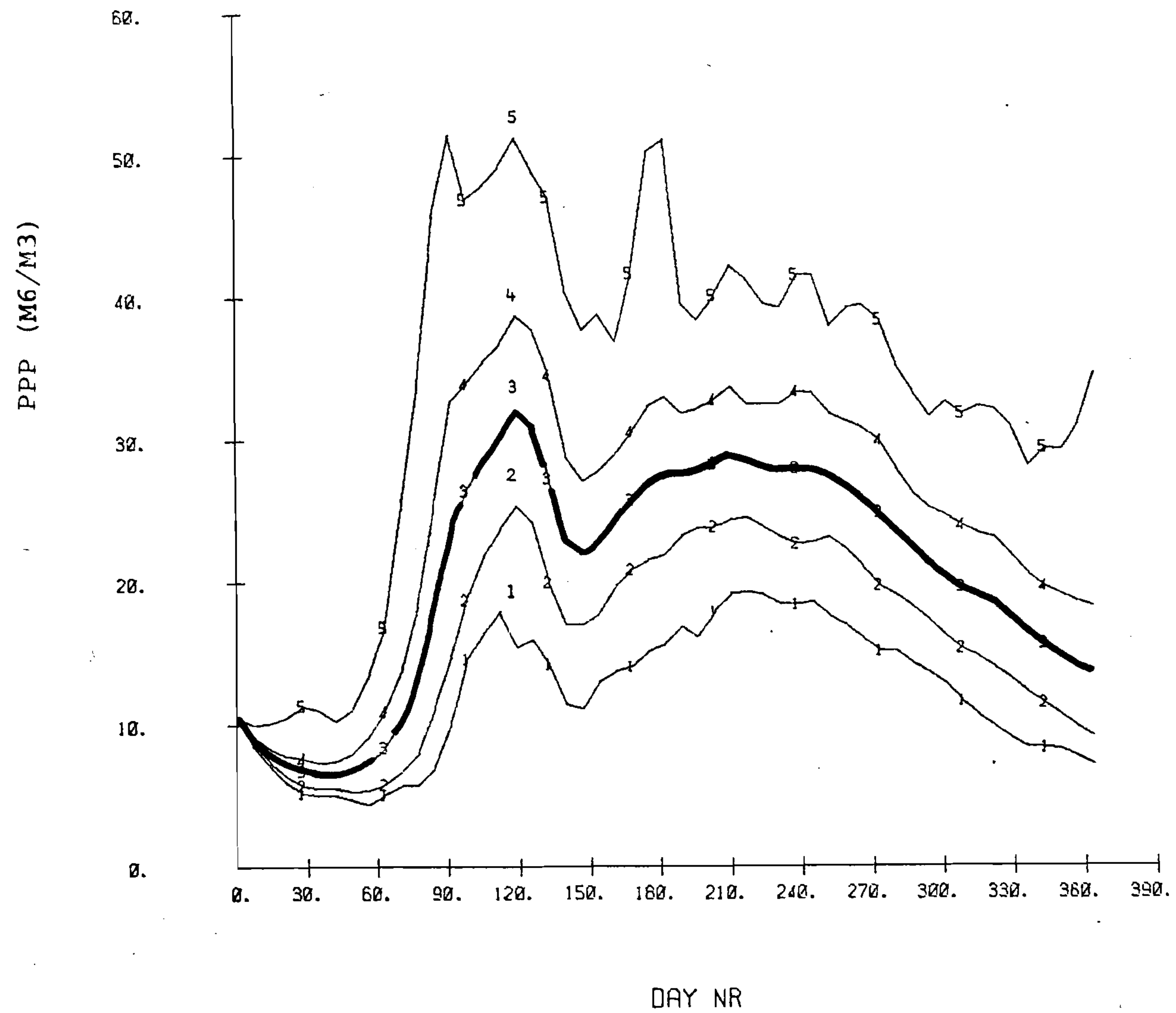

Figure 10. The combined influence of uncertainty and stochasticity in the meteorology and loading, respectively, on the water quality. Basin 1.
3 - mean value, 4 and 2 - ± standard deviations,
5 and 1 - extreme values

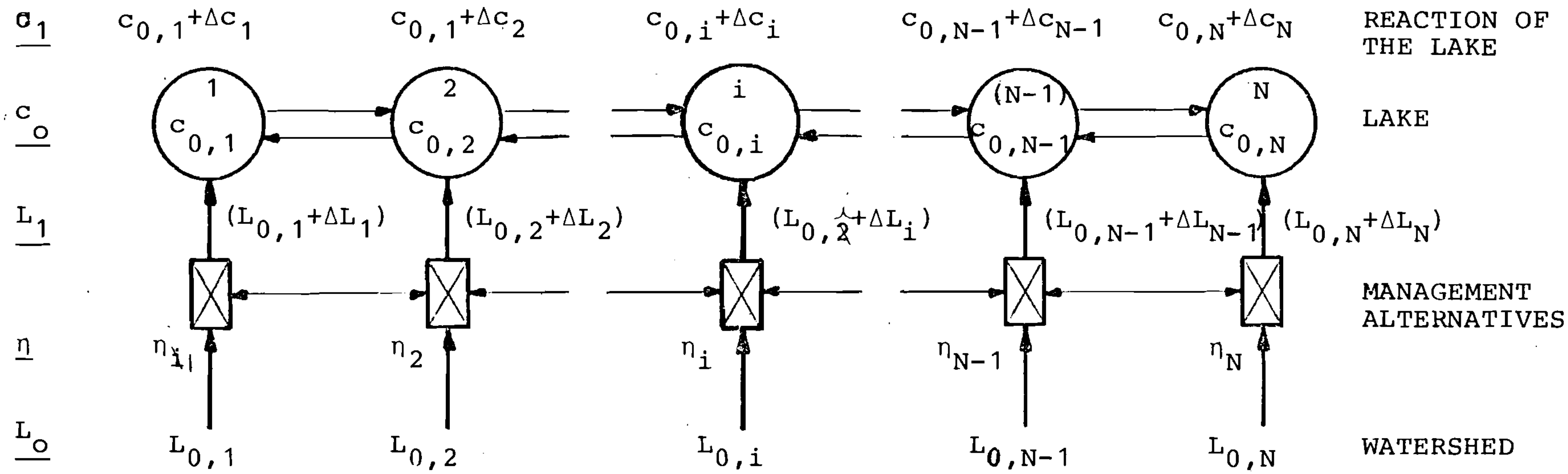

Figure 11. Incorporation of the lake model in WQMM
i - uniform segments in the lake
$\leftrightarrows$ - interaction among basins through mass exchange and hydrologic throughflow
$\leftrightarrow$ - interaction among load components through control activities