PART TWO: DESIGN OF DEDICATED CAMPAIGNS

8. STATIONARY HYDRAULIC CONDITIONS

In Sect.5 we saw how the method of characteristics works, and how the data can, and should, be *collected on a characteristic line*, when such a method is adopted. This requirement implies a rigorous organization of the campaign, as the sampling is Lagrangian. In fact, the samples must be collected at the corresponding scheduled times, at each sampling station. In order to set up this schedule, it is necessary to know in advance that particular characteristic line corresponding to the time-space velocity field $v(t,l)$ that will be present during the campaign itself. In fact, the schedule must be available prior to the execution of the campaign. Consequently, a direct on-line measurement of the velocity during the campaign itself does not represent a feasible solution. On the contrary, a forecast of river velocity $v(t,l)$ a few days ahead-of-time must inevitably be made. Obviously, forecasting a time-varying velocity is rather difficult, if not impossible. Therefore, it is of great advantage to carry out the campaign in the presence of steady-state hydraulic conditions, since, in that case, the forecast reduces to a space-varying velocity pattern $v(l)$ (notice, however, that once calibrated, the model could also be used to simulate unsteady hydraulic conditions). Therefore one must try hard to meet such conditions during the campaign. But exactly when, and for how long must the hydraulic conditions be in steady-state? In order to answer this question, it is worth identifying a criterion for computing what we call the *time intervals of required stationarity*.

Under steady-state hydraulic conditions the characteristic line defined by equations (2a,b) can be given the following integral form:

$$t(l) \;=\; \int_{l_0}^{l} \frac{1}{v(s)}\,ds \;+ t_0 \qquad (25)$$

where s is the current spatial coordinate and t_0 and l_0 denote respectively the initial time and the initial section of the campaign.

The function $t(l)$ allows to define the sampling schedule, as it establishes the instant $t_k=t(l_k)$ at which the sample at the generic station l_k must be collected.

Obviously, steady-state hydraulic conditions cannot hold forever. On the other hand, it is only required that no perturbations affect the measurements on the characteristic line. Operationally, we need to define the time interval $[t_l', \; t_l'']$ (called from now on the *interval of required* *stationarity*), for each section l, during which a generic cause of hydraulic perturbations (e.g. a tributary or a diversion dam), located there, must not act in order to guarantee that steady-state conditions are met by *travelling on the characteristic line*.

An approximate solution to this problem can be obtained by remembering that hydraulic perturbations travel at two different velocities c_1 and c_2 (corresponding to the two families of characteristic lines of the Saint Venant hyperbolic differential equations of motion for mono-dimensional flows, e.g. see *Li* [1972b]):

$$c_1 = v + (gy)^{1/2} \tag{26a}$$

$$c_2 = v - (gy)^{1/2} \tag{26b}$$

where g is gravity acceleration and y is the mean depth of the river (ratio between the cross-sectional area and the surface width of the section). In sub-critical flows (the most common), c_1 is negative, that is, the perturbations can also travel upstream. We refer to this situation in the following.

The graphical solution to the posed problem is depicted in Figure 5. Note that the initial l_0 and final l_L sections of the stretch of river under consideration must be thought of as possible causes of perturbations, as the hydraulic boundary conditions, specified therewith, constitute the interfaces with the upstream and downstream worlds. In Figure 5, curve (c) is the biochemical model characteristic line (solution of equations (2a,b)) corresponding to the stationary hydraulic conditions considered. Curve (a) corresponds to the path of a *fast perturbation* (the velocity of which is given by equation (26a)), that travels downstream, and affects curve (c) only in the terminal section l_L. Curve (b) corresponds to the path of a *slow perturbation* (see equation (26b)), that travels upstream, and affects curve (c) only in the initial section l_0. Let us denote the time instants, by using t_0'' and t_L', in which such perturbations pass through the initial (l_0) and final (l_L) sections respectively. Fast perturbations which originate in the initial section later than t_0'' do not meet curve (c) in the defined stretch. Hence, from the instant t_0'' on, stationarity is no longer required in the initial section. An analogous reasoning holds for a generic cause of perturbations, located at section l, from the corresponding time instant t_l''. On the other hand, every fast perturbation which originates in the initial section earlier than t_0' (coincident with the instant t_o at which the campaign begins) does not meet curve (c) because it travels with a velocity c_1 (equation (26a)) higher than the river velocity v. Analogously, slow perturbations which originate or pass through l earlier than t_l', do not meet curve (c). It follows that the sought interval of required stationarity, relative to a generic cause of perturbations located in section l, is actually the interval $[t_l', t_l'']$ defined in Figure 5.

The above-defined intervals are however larger than those which are strictly needed because, in their determination, it is assumed that the perturbations do propagate along the whole river stretch. An evident exception is constituted by the presence of hydraulic disconnections (like river-bed jumps) that can interrupt the upstream propagation of slow perturbations. In this case, the construction of Figure 5 must be modified as shown in Figure 6, where D denotes a section where a disconnection is located.

In practice, the design of the campaign demands that some agreements must be made with the managers of hydraulic facilities (dams, derivation channels, etc.) in order to be sure that no variations of flow rates and hydraulic levels occur in the interval $[t_l', t_l'']$. Moreover, attention should be paid to meteorological forecasts in order to avoid variations to the flow rate of some tributary again during the same interval. Measurements of hydrometric levels and of the flow rates relative to all the possible causes should be taken, during the campaign, in order to verify a posteriori the actual occurrence of stationary conditions during the campaign. Finally, it is of interest to observe that when, for whatever reason, direct verification of the stationarity of a particular cause of perturbation is not possible, indirect verification can be carried out by checking the hydrometric level variations in a different section over a suitable interval of time. This interval of time can be easily derived on the basis of a construction analogous to that of Figure 6.

9. RIVER VELOCITY FORECAST

The sampling schedule must be set up a few days prior to the execution of the campaign in order to organize the operational teams. To solve equations (2a,b) it is thus necessary to produce a forecast $\hat{v}(l)$ of the velocity pattern v(l) that will be present during the campaign itself.

The easiest case is when one can reasonably assume that stationary hydraulic conditions will hold from the forecast time up to the end of the campaign. The most direct approach is then to measure the mean cross-sectional velocity during a pre- campaign at a sufficiently high number of sections such that the current velocity pattern v'(l) can be inferred; the forecast will then simply be: $\hat{v}(l)=v'(l)$. This trivial solution is, however, not applicable when the river stretch is too long and/or characterized by too high a flow rate, since both these features generally imply that velocity measurements require a significantly longer period to be completed. To overcome this difficulty one can measure the transport velocity (instead of the mean velocity) by means of floaters or tracers. Obviously, the drawback is that the transport velocity is, in general, different from the mean velocity. This solution, therefore, should be applied only when the ratio between the two velocities is known somehow.

Often, it is not possible to guarantee (or to take advantage of) such a long period of stationarity. Therefore, in a situation of shorter hydraulic stationarity, the best solution is to forecast the river flow rate pattern $\hat{Q}(l)$ [m^3/s] and to transform it into the sought velocity pattern. The pattern $\hat{Q}(l)$ can be computed by solving the continuity equation (which has only one unknown, the flow rate, under stationary conditions):

$$\frac{d}{dl}\hat{Q}(l) = \hat{S}_q(l) \tag{27a}$$

with initial condition

$$\hat{Q}(l_0) = \hat{Q}_0 \tag{27b}$$

where $\hat{S}_q(l)$ [m^3/s] and $\hat{Q}_0$ [m^3/s] are the forecasts of the water inflow pattern $S_q(l)$ and of the initial flow Q_0, respectively, provided that such forecasts are easily obtainable.

As far as the transformation of $\hat{Q}(l)$ into $\hat{v}(l)$ is concerned, the procedure conceptually perhaps most satisfactory is to first compute the water level $\hat{h}(l)$ (by solving the momentum balance equation), and then to compute the velocity pattern $\hat{v}(l)$ on the base of its very definition, i.e. by means of the following formula:

$$\hat{v}(l) = \hat{Q}(l) / A(\hat{h}(l), l) \tag{28}$$

where A(h, l) is the cross-sectional area [m^2] as a function of water level h and location l. This procedure requires, of course, a lot of information (in particular the geometry of river sections and the roughness coefficient) and is, for this reason, generally quite costly.

Alternatively, and more simply, the pattern $\hat{Q}(l)$ can be transformed into $\hat{v}(l)$ by means of a more empirical procedure. The river stretch is partitioned into a suitable number (say n) of hydraulically homogeneous (and possibly hydraulically disconnected) reaches, where both the flow rate and the velocity can be assumed to be approximately space constant. A velocity-flow rate relationship $v_i=v_i(Q_i)$ then has to be determined for each of the n reaches: either experimentally (via non-linear regression on a set of velocity - flow rate couples previously measured under different stationary hydraulic conditions), or theoretically (by using the known hydraulics laws, together with geometric information). Finally, the sought pattern is approximated by the step- wise function $\hat{v}(l)=v_i(\hat{Q}_i)$, with l at reach i and i=1,...,n.

The adopted solution is, in practice, often a combination of the previous approaches. This is true mainly as a result of keeping the cost of the model as low as possible.

10. AN APPLICATION EXAMPLE: THE GARZA STREAM

The Garza stream is located in northern Italy. Its small catchment of about 74 km^2 is mainly rain fed. A model of a stretch, about 30 km long, was required in order to evaluate the effectiveness of alternative treatment plant locations. In fact, the river is highly polluted because of industrial paper mills located in the upper part (0 to 12.5 km), and municipal and industrial waste discharges from the town of Brescia. A BOD-DO model was selected to carry out this study as the main pollutants are biodegradable organic substances.

The Garza case study must be regarded as a limiting boundary for model applicability owing to the very small spatial scale which implies very low flows (significantly lower than 1 m^3/s) and high sensitivity to even very small waste inflows. In fact, as is apparent in the following, it was necessary to acquire a very detailed knowledge and description of the system in order to be able to predict its behavior. It is because of the difficulties associated with the Garza modeling process that it was selected as a testing example of the previously illustrated methodology.

As seen in Sect. 6.2 a river can be modeled by a plug-flow model if the conditions stated by the Extended Dobbins' Criterion (EDC) are satisfied; namely: $kD/v^2 < 0.01$ and $\omega < 0.01\ v^2/D$, where, in the case of a BOD-DO model, k is the greater of the deoxygenation and reaeration rate coefficients (see Sect. 6.4), and ω is the highest frequency still significant in the power spectrum of the river load functions. To check the satisfaction of these conditions, an experiment was undertaken, first of all, to estimate the longitudinal dispersion coefficient D by means of the impulse release of a chloride saline solution, and the measurement of consequent downstream concentrations as functions of time (the method is presented in *Day* [1975]). The estimate supplied the value D=0.9 [m^2/s]. The literature [*EPA*, 1985, p. 111] revealed that, in relation to the Garza mean depth (0.03-0.3 [m]), the reaeration coefficient may be in the range 1-100 [d^{-1}], and the deoxygenation coefficient in the range 0.004-4.24 [d^{-1}]. Then, by assigning to v the value of the mean velocity (0.35 [m/s]) in the whole stretch, the result is $kD/v^2=0.008$: i.e. the first condition is satisfied. From the same figures it turns out that ω must be lower than 117.5 [d^{-1}], which is equivalent to saying that the period of the highest harmonic still significant in the river load spectrum must be greater than 15 minutes. If the Garza had been a medium spatial scale system the conclusion would have hence been that daily waste fluctuations would not at all have prevented the application of a plug-flow model. In our case, however, no guarantee was available. While the daily pattern of a medium-scale sewerage system (which constitutes a typical load to the Garza stream) certainly does not contain harmonics with periods lower than 15 minutes, the same is not a priori true for the small -scale waste inflows that act on the Garza stream. (pH, dissolved oxygen concentrations and conductivity data collected for a period of several hours at a representative station showed, indeed, sudden variations certainly due to impulse- like discharges). From time to time, the frequencies associated with such loads were certainly much higher than the threshold ω= 117.5 [d^{-1}]. Moreover, it is not at all infrequent to find the presence of sporadic, impulse-like discharges that, owing to the small spatial scale of the system, can significantly affect its quality.

Nevertheless, as already shown in Sect. 3 and as will be even more evident in the following, the collection of all the data required to fully simulate nonsteady- state (aperiodic) conditions using a dispersion model is absolutely impossible. We hence decided to go on, notwithstanding, with the application of a plug-flow model in order to organize the data collection campaigns according to the method of characteristics. The final judgement is left to the results obtained in model development. For this reason, as well as for the very small spatial scale, the presented case study lies on the boundaries of model development orthodoxy. Perhaps it is, however, well representative of the way one often has to deal with real-world problems.

More than 70 of the sampling stations (where a width-depth cross-sectional integrated sample was collected to produce the measurements of the quality indicators: BOD, COD and DO concentrations, temperature T and pH) and hydraulic sections (where the flow rate and the hydrometric levels were measured) were selected. The sampling stations were located on the basis of the following criteria:

- to cover the entire river stretch (about 30 km);
- to ensure that all the quality data were representative of the average cross- section concentration (according to the hypothesis of spatial mono-dimensionality). This implies that the sampling stations were not located short distances downstream of inflow points (tributary or discharge); and that the samples were depth and width integrated over the section;

- to measure all the significant load inputs (tributaries and discharges);
- to allow the indirect estimate of distributed unrecorded loads: hence they were located typically upstream and downstream of reaches where the presence of a distributed source was suspected on the base of the anthropic and hydrogeological information on the area.

The hydraulic sections were located on the basis of the following criteria:
- to measure the flow rate of all the relevant water inflow inputs and (tributaries and discharges) and outputs), in order to determine the inflow term $S_q(l)$, and the loads **P** by multiplying (see equation (1b)) the flow rate by the concentration of each compound;
- to allow the determination of the actual flow $Q(l)$ and the checking (via the continuity equation) of the measured inflow term $S_q(l)$ (both $Q(l)$ and $S_q(l)$ are needed, together with $v(l)$, to integrate the model equations, see Sect. 3);
- to allow the determination of the actual velocity $v(l)$ in order to compute the actual characteristic line;
- to allow the a posteriori verification of the stationarity of the actual hydraulic conditions.

After a preliminary (partial) data collection campaign aimed at verifying local collection problems, five complete one-day-long campaigns were carried out in order to allow the estimation of the parameters that express the reaction rates in WODA (see Appendix). Thus the campaign dates were selected to guarantee the presence of different hydrological, thermal and load conditions. Table 2 indicates the characteristic elements of the five campaigns.

Table 2 - Data gathering campaign on the Garza stream.

Campaign n.	Date	River stretch [Km]	Instream hydro station #	Biochemical station #	Inflow/outflow station #
1	July 24, 1991	20.3 – 30.5	4	6	7
2	October 30, 1991	0.0 – 16.2	5	12	34
3	October 31, 1991	15.6 – 30.5	6	9	33
4	March 11, 1992	0.0 – 17.6	5	11	38
5	March 12, 1992	20.3 – 30.5	6	10	31

The presence of stationary hydraulic conditions was guaranteed by selecting a period, on the basis of the procedure illustrated in Sect. 8, far enough away from notable meteorological events. In addition, a great deal of work was carried out to make suitable agreements with the managers of the numerous controllable hydraulic devices that are located in the Garza catchment. In particular, a few small family-sized artisan factories still exist which exploit the main stream and some tributaries to move water wheels that may produce transient phenomena. At the same time, the stream is highly exploited for irrigation purposes, with a number of small, individually managed, irrigation canals which can suddenly change the water regime of the stream, by diverting more than 50% of the flow.

Finally, due attention was paid to selecting days that were far from week-ends and other periods of interruption to industrial work activity. This was done in order to obtain data that were representative of the general load condition, and in order to operate closely enough under stationary biochemical conditions. This latter condition would not have been necessary if the data had been collected exactly on the actual characteristic line, i.e. the characteristic line corresponding to the hydraulic conditions actually present during the campaign. However, since it is impossible to guarantee that the actual characteristic line is followed perfectly, the presence of nearly stationary biochemical conditions helps in reducing the size of possible errors.

In order to determine the sampling schedule, a forecast of the stream velocity was produced a few days prior to the execution of the campaigns. This was done by partitioning the river stretch into 9 reaches and simply measuring the transport velocity v_i' in each reach i through dye experiments (hence $\hat{v}_i=v_i'$ for i=1,...,9).

By solving equation (25) with the forecasted velocity, the sampling schedule was defined and the operational teams organized: two teams (comprising two operators) collected water samples, and the third team (comprising three operators) measured the flow rates and velocities. The schedule was adjusted on-line on the basis of the mean velocity obtained as a by-product of the flow rate measurements during the campaign.

As an example, Table 3 shows the data from the third campaign (October 31, 1991) that covered the downstream portion of the Garza stream. The data have already been checked and integrated (i.e. missing data have been reconstructed) for the calibration phase using the procedure described in the following paragraph.

Preliminary analysis of the data

An accurate analysis of the collected data was undertaken on the basis of the following criteria:

Hydrological criteria

a) No significant variations of flows and hydrometric levels should have occurred in the intervals of required stationarity corresponding to the actual characteristic line. The data affected by hydraulic transients were simply disregarded.

b) The water balance must have been verified between two consecutive stream measurements: any deviation was inquired into in order to determine (if possible) the most probable cause: was it due to an unknown concentrated inflow? or to a distributed inflow? When both cases seemed improbable, we assumed that the cause of the lack of balance was a measurement error in one of the terms. In any case the consequent actions are specified below.

Quality criteria

c) By definition, the COD value must always be higher than the corresponding BOD value. In addition, the ratio between the BOD and COD concentrations, for each sample, should not vary significantly along the stream (since the load composition was expected to be the same all over the catchment). When one of these conditions was violated, the pair of measurements was disregarded.

d) When an evident increase in the BOD and COD concentrations appeared between two consecutive instream stations, and this increase was not imputable to known and measurable loads, then an unrecorded BOD was introduced. Its cause was assumed to be the unmeasured municipal or agricultural fluxes (see below) and its value was estimated on the basis of the number of equivalent inhabitants.

Points b) and d) deserve further insight. Even if the campaign design had been as accurate as possible, not everything could have been measured. In particular, owing to budget constraints, the checking of some small drains and discharge facilities (e.g. sewerage outlets), that are normally dry (inactive), was omitted during the campaign. Such checking would have required an additional team (remember that measurements must follow a tight schedule). Thus we cannot completely exclude the possibility that some of them were active during the campaign. In addition, some drains reach the Garza stream in an urban area, through which the stream is piped. Hence, it was simply impossible to check whether they were active or not, and, if so, to measure their flow rates and quality. As a result, when the balance between two measurements of the stream flow rate was unsatisfied (see criterion b) and revealed a deficit of inflow, it was considered a sign that uncontrolled discharges (in the stretch concerned) could have been active. Since the balance check was done the day following the campaign, a short inquiry was possible in the majority of cases to determine, with reasonable probability, which of the drains or discharges, if any, were active. A flow rate value was inputted such as to close the gap (see Example 2 below). In some cases there happened to be either no uncontrolled drains or discharges in the unbalanced stretch (see Example 3 below), or it was not possible to conclude which one of them was active. The third possibility was that the balance was negative. In these cases the balancing inflow (in an algebraic sense) was assumed to be distributed. The reader may be surprised at such a detailed analysis and may claim that the unbalance may have been due simply to measurement errors. It actually may have been. However, even in this latter case, the unbalanced flow must be included. In the opposite case in fact, given that under stationary conditions the flow rate equation is simply an integrator, the error in the computed flow rate would accumulate while proceeding downstream. On the contrary, by inserting the balancing term, the error is confined to the stretch where it occurred. Then, if we have to include a balancing term anyway, why not try to locate it in its most probable location? This is the rationale used in our procedure (the resulting balancing terms appear in italics in Table 3).

BOD and DO concentrations must be associated with each one of these terms, since their presence does affect the balance of compounds. When the type of source was identified, we adopted an a posteriori estimate of its quality characteristic. On the contrary, when no cause was identified, the BOD and DO values were set equal to the BOD and DO instream concentrations as computed by the model in the inflow section. The rational of this assumption is so as not to perturb the compound concentrations.

Following a detailed study of hydrological and sewerage system maps of the Garza catchment, the catchment was partitioned into two types of sub-catchments. The first type of sub-catchment is drained by a measured inflow (tributary, sewage discharge, or a mixture of both). The second type of sub-catchment is bounded by sub-catchments of the first type and discharges into a stretch of the stream that is limited by two inflow sections. The load brought by sub-catchments of the first type reaches the stream as a point source whose value is known (measured inflow rate times measured BOD concentration). The load which is potentially contributed by the i-th sub- catchment of the second type reaches the stream as a non-point source, whose intensity is expected to be proportional to the number E_i of equivalent inhabitants that are located in it. More precisely, at each point l of the reach (of length L_i) that drains the i-th sub-catchment of the second type, the intensity of the distributed load is proportional to $B_u(l)=E_i/L_i$, by an unknown constant of proportion (u) to be estimated (see Appendix). The values of E_i were determined by cross- correlating the cadastral map with census data. The resulting function $B_u(l)$ is shown in Figure 7. $B_u(l)$ must be thought of as an indirect measure of the potential load, not of the actual load. In fact, since the campaign was carried out in a period of no rain, it is not said that the potential load would have actually reached the stream: it may have been stored somewhere (e.g. the majority of this load is due to cow and pig excrement that reaches the stream only when the cowsheds and pigsties are cleaned). We decided then to consider the unrecorded BOD load of the i-th reach only when a couple of consecutive instream BOD (and COD) measurements revealed an increment that could not be explained by the presence of measured inflows.

With reference to Table 3, let us make some examples that better clarify our procedure:

Example 1. Both the BOD and COD instream concentrations increase from sec. 44 to sec. 47ex. The increase is very small and in principle attributable to measurement errors. The occurrence of a synchronous increase makes, however, quite unlikely the possibility of a measurement error. Furthermore, also the concentrations of other pollutants indicators (NO_3, NH_4) increase in the same reach. Since it turns out (see Fig. 7) that a potential unrecorded BOD load of about 550 [eq. in./km] exists in the same reach, this source was considered to be active during the campaign.

Example 2. Between sec. 46 and 49: the hydrologic balance showed a deficit inflow of 15 [l/s]. Although this amount could be explained simply as a measurement error, the synchronous increase of BOD and COD concentrations from sec. 47ex and 49, supports the presence of a discharge. A survey confirmed that the 47bs drain was actually active. We therefore introduced an inflow of 15 [l/s] in sec. 47bs, for which the BOD and DO concentrations should be specified. However, since these concentrations were not measured, we considered it more reasonable to assume the presence of the distributed unrecorded BOD load of Figure 7. Accordingly, the BOD concentration of the new discharge was set to zero. The DO concentration was assumed to be 8.0 [mg/l], which is a value such as not to perturb the computed instream DO concentration (according to what has already been stated). In addition, a sensitivity analysis revealed that these assumptions do not significantly affect the parameter estimates.

Example 3. In the 57A-59 reach (2.2 km long) the hydrologic balance showed a deficit of an inflow rate of 240 [l/s]. This is certainly not a measurement error and is probably due to distributed drainage from the irrigation network. A distributed inflow of 109 [l/s km], between the two stations, has hence been included. The quality of this inflow was characterized by a BOD of 3 [mg/l], i.e. nearly clean water, in agreement with measurements previously taken in the same period of the year on some of the agricultural drains, and by a DO concentration of

7.6 [mg/l], equal to the mean stream concentration. Again the a posteriori sensitivity analysis showed that these choices were not critical.

Example 4. The increase of BOD and COD concentrations from sec. 52 to sec. 58 might be explained by the presence of the Naviglio Grande load (sec. 53). Therefore the sources of the unrecorded BOD load in the reach were assumed not to have been active during the campaign.

It is important to note that the statement that the Naviglio Grande load could explain the increase in BOD concentrations is only a logical deduction: it is not based on quantitative assessments. In fact, a quantitative assessment would be on the one hand impossible in this phase, since it would require the specification of the BOD decay rate (that has not been estimated yet), and on the other hand unfair, since it is exactly on the closure of this type of balance that the parameter estimation is based.

Parameter estimation

The parameter estimation of the BOD-DO WODA model (see Appendix) was carried out by a deterministic least squares approach. That is, the Sum of Squares Deviations (SSD) between model output and measured concentrations has been minimized. The minimization was carried out by an iterative multi-dimensional search procedure available in the WODA package (based on *Zangwill* [1967]). The data (*dataset)* of the first campaign (July 24, 1991), turned out to be lacking the most from the point of view of load determination, and therefore they were kept for model validation, while the calibration was performed on the four remaining datasets.

In order to better comprehend what follows, one should note that the second (October 30, 1991), and the fourth (March 11, 1992) campaigns covered only the upper stretch of the Garza stream, which is characterized by rather rapid flow and low light exposure. The first (July 24, 1991), the third (October 31, 1991), and the fifth (March 12, 1992) campaigns covered, instead, the downstream stretch. This is characterized by a much more uniform flow, high light exposure, the presence of agricultural fertilizers and organic loads, and by the considerable consequent presence of algae (both macro and microphytes).

From an analysis of all the available data (not reported here for reasons of space), it was first noted that there was a considerable influence by algal activity on the oxygen balance. Hence a calibration session was carried out by using BOD data only, in order to estimate only the BOD parameters, namely k_{11} and u (see Appendix, the first one defines the deoxygenation rate, and the second one the pro capita BOD contribution). The effects of the unrecorded BOD load was analyzed by comparing the results of two independent estimates: the estimate of k_{11} alone (when u is set to zero) (*Est. 1a*), and of both k_{11} and u (*Est. 1b*). The following results were obtained:

Est. 1a

Estimated parameters: k_{11}

Considered datasets: 2, 3, 4, 5

Number of available BOD data: 40

Notes: it is assumed u=0 (no unrecorded load)

Results: k_{11} = 4.7 [d^{-1}]

SSD_{BOD} = 445

Est. 1b

Estimated parameters: k_{11} and u

Considered datasets: 2, 3, 4, 5

Number of available BOD data: 40

Notes:

Results: k_{11} = 6.0 [d^{-1}]; u $\cong$ 88.2 [gr BOD/eq.in. day]

SSD_{BOD} = 219

It can be noted that the performance obtained by *Est. 1b* is significantly better with respect to *Est. 1a*, since its SSD is nearly 50% of the previous one. The value of u is slightly higher than usual (65-75 gr BOD/eq.in. day). This could be the consequence of a slight underestimating of the unrecorded load, or just the consequence of a less than perfectly correct usage of the equivalence coefficients used to compute the number of equivalent inhabitants. Nevertheless, nothing more can be said on the base of the available information. The value obtained for the deoxygenation rate at 20 [$^{\circ}$C], k_{11} = 6.0 [day^{-1}], lies near the upper boundary of the literature range [*EPA*, 1985, p.. 152]. Figure 8 shows the BOD simulation of the 2, 3, 4, 5 datasets used in the *Est 1b*, in relation to the BOD parameterization thereby obtained. This parameterization is the one adopted in what follows.

The estimate of the DO parameters was then carried out, namely, the parameters k_{22} , k_{23} and k_{24} (see Appendix) were considered. They represent, respectively, the reaeration rate at 20

°C, the algal respiration rate and the peak photosynthetic oxygen production rate. On the basis of on-field assessment, the algal activity was assumed to be present only downstream of km 23.5 (i.e. k_{23} and k_{24} were set to zero upstream of that section. Furthermore, this means that algae could have affected only datasets number 1, 3, and 5, which cover the downstream stretch of the Garza). The DO data in the fifth dataset show a strong over-saturation in the downstream stretch (see Figure 10) which reveals (together with the values of other quality indicators) the presence of an eutrophication phenomenon. Therefore, we decided to carry out two different estimations: one for non-eutrophic conditions (dataset 2,3,4), and one for eutrophic conditions (datasets 5). For the first estimation the results were as follows:

Est. 2
Estimated parameters: k_{22} , k_{23} and k_{24}
Considered datasets: 2, 3, 4
Number of available DO data: 30
Notes: algal activity is assumed downstream of 23.5 km.
Results: k_{22} = 24.8 [d^{-1}];
k_{23} = - 0.4 [mg/l h]; k_{24} = 1.2 [mg/l h];
SSD_{DO} = 23.9

The results of *Est. 2* confirm the presence of algal activity. The values obtained for the above three parameters lie at the limits of the literature range [*EPA*, 1985]. General performance was rather satisfactory: Figure 9 shows the corresponding DO simulations.

We then considered the eutrophic conditions of the fifth dataset (March 12, 1993). The results were as follows:

Est. 3
Estimated parameters: k_{23} and k_{24}
Considered dataset: 5
Number of available DO data: 10
Notes: the algal activity is assumed only downstream of 23.5 km; k_{22} is set at the value of *Est.* 2
Results: k_{23} = -1.0 [mg/l h]; k_{24} = 6.15 [mg/l h];
SSD_{DO} = 2.3

Figure 10 (curve a) shows the simulation corresponding to the parameterization obtained in *Est. 2* (non-eutrophic conditions), and (curve b) shows the simulation obtained on the basis of the above estimate.

It is evident that the presence of abnormal algal activity was certainly caused by an eutrophication phenomenon which may have been the consequence of significant rainfalls that occurred at the beginning of March, and which could have washed off the chemical and organic agricultural fertilizers applied in late February.

The possibility of an overestimate of the BOD load on the Naviglio Grande input (sec. 53, 20.35 km) is analyzed in Figure 11. Such an overestimate cannot be excluded because the sample was taken at an underground check-point in the sewerage system where there could have been local accumulation phenomena. By reducing the Naviglio Grande load it can be noted that the immediate downstream BOD and DO concentrations would be better explained.

As a conclusion to the calibration phase, we proposed using the set of parameters obtained in the *Est. 1b* and *Est. 2* for normal conditions, while using those obtained in *Est. 1b* and *Est. 3* for eutrophic conditions (March and Summer).

Validation

The campaign on July 24, 1991 was used to assess the validity of the model. Figure 12 (curve a) shows the simulation obtained with the parameters previously calibrated (*Est. 1b* and *3*). Curve b, corresponds to an increased value in the parameter k_{24} (k_{24} = 8.5 [mg/l h]). As is apparent, the data are much better explained. Such an increase seems to be justified by observing that eutrophication phenomena should have been even stronger than those detected in March owing to the much higher summer temperature.

The estimates of the algal parameters k_{23} and k_{24} have also been validated by a *dark-bottle* experiment conducted on a water sample collected at 26.59 km in February 1991. (In the dark-bottle experiment, half of the sample was put in a transparent bottle and the remaining half in a dark bottle. The transparent bottle was left exposed to the daily sunshine cycle for five days. DO concentrations in both the bottles were then measured at the beginning and end of the period. The difference between the DO variation measured in the transparent bottle and that in the dark-bottle, divided by the elapsed time, is an estimate of the average daily oxygen production D). It turned out D =4.3 [mg/l day]. The value of D should correspond to the integral over 24 hours of the net algal production rate, as defined by equation (A3d) in the Appendix. The value of k_{24} can then be estimated given an a priori estimate of k_{23} (or vice versa). Thus, by assuming the value for k_{23} obtained in the *Est. 2*, one obtains k_{24} = 1.6 [mg/l h] from the experimental data, which is sufficiently close to the value of 1.2 [mg/l h] obtained in *Est. 2*.

Finally, it can be verified that the Dobbins' Criterion is actually satisfied by the estimates of k_{11} and k_{22}, but we have no elements with which to express a judgment on the satisfaction of the EDC.

CONCLUSIONS

In this second part of the paper we have described the basic steps for the design of a data collection campaign, which is aimed at parameter estimation in a plug flow model. Data must be collected along a characteristic line. A river velocity forecast is necessary to determine in advance the characteristic line and hence the sampling schedule. The practical possibilities available to determine that forecast sufficiently in advance have been shown schematically. As this forecast can only be produced easily for stationary hydraulic conditions, an operational criterion aimed at identifying and verifying such conditions has been derived. Finally, an example of the application of the proposed methodology to the Garza stream has been described in detail.

FIGURE CAPTIONS

Fig. 5. Graphic construction of the intervals of required stationarity.

Fig. 6. Graphic construction of the intervals of required stationarity when hydraulic disconnections (D) are present.

Fig. 7. The Garza unrecorded BOD load expressed as the number of equivalent inhabitants per kilometer. The dotted load is actually active only when the cowsheds and pigsties are cleaned.

Fig. 8. Simulation of the dataset used in *Est. 1b* with the parameterization obtained there.

Fig. 9. Simulation of the dataset used in *Est. 2* with the parameterization thereby obtained.

Fig. 10 Simulation of the eutrophic dataset (March 12, 1992 campaign): (a) with the parameterization of *Est. 2*, and (b) with the parameterization of *Est. 3*.

Fig. 11 Simulation of the fifth dataset (March 12, 1992 campaign): analysis of a possible overestimation of the load in Sec. 53 (20.35 km).

Fig. 12 Simulation of the July 24, 1991 campaign (first dataset) with the parameters obtained in *Ests.1b and 3* (continuous line), and with the same parameters, but at a higher value (8.5 [mg/l h]) of k_{24} (dashed line).

APPENDIX

WODA is a modeling support system (MSS) [*Kraszewski and Soncini-Sessa*, 1986]. The model adopted in WODA is an extension of the Dobbins' model [*Dobbins*, 1964]: it is a BOD-DO model which describes the BOD decay along a stretch due to bacterial degradation and sedimentation, as well as the variations of DO concentrations due to photosynthetic oxygen production, natural and artificial aeration, and oxygen consumption in the water and in the sediments. The extension with respect to the Dobbins' model consists of the explicit formalization of the dependence of the reaction rates on the flow rate and temperature, and in the introduction of a specific parameter which allows one to take the distributed unrecorded load into account.

The equations which constitute the model are now presented briefly. The hydrological submodel comprises two equations, the continuity equation:

$$dQ/dl = \mathbf{S_q} \qquad \text{(A1a)}$$

where S_q is assumed to be the sum of concentrated Q_c (impulse function of the space l(ô)) and distributed Q_d (piecewise constant function of the space l(ô)) inflows (or outflows); and, the momentum balance equation that under stationary conditions can be given the empirical form (conceptually valid only for uniform flow):

$$v = w_1 Q^{w2} \qquad \text{(A1b)}$$

where w_1 and w_2 are piece-wise constant functions of the space $l(\hat{o})$, and are thus defined by a couple of values in each reach. These values can be estimated in WODA (least square estimate) on the basis of couples of measurements of the velocity v and of the flow Q in the reach.

Instead of including a complete thermal sub-model, WODA assumes that the solution is obtained independently, and it assumes accordingly that an exogenous piece-wise constant function $T(l(\hat{o}))$ is given to specify the thermal condition. This function can be determined, for instance, by suitably interpolating field measurements.

The BOD and DO equation is of the form (2c), that is reproduced here for the reader's benefit

$$d\mathbf{p}/d\tau = - \mathbf{S_q}(t,l)\,\mathbf{p} + \mathbf{S} \qquad \text{(A2a)}$$

where $l=l(\hat{o})$ is the characteristic line solution of equations (2a,b) and **p** is the vector of the BOD (b) and DO (d) concentrations. The source term **S** is modeled as

$$\mathbf{S} = \begin{vmatrix} -K_1 b + (B(l) + u\, B_u(l))/A \\ \\ - K_3 b + K_2\,(d_3(T)-d) + D(l)/A \end{vmatrix}$$

where $d_s = d_s(T)$ is the oxygen saturation concentration; $(B(l)+uB_u(l))$ and $D(l)$ are the corresponding external loads (i.e. the components of the vector **P**) and K_1, K_2, K_3 are reaction rates (could be functions of space), while u is a constant parameter. The BOD decay rate K_1 has the following form:

$$K_1 = k_{11}\, h_1^{(T-20)} + k_{12} \qquad \text{(A3a)}$$

where k_{11}, k_{12} and h_1 are parameters. The first term on the right side of (A3a) is the de-oxygenation rate (that depends heavily on the temperature), while the second term takes into account the effects of sedimentation. To be consistent, the de-oxygenation rate K_3 is defined by:

$$K_3 = k_{11}\, h_1^{(T-20)} \qquad \text{(A3b)}$$

The re-oxygenation rate K_2, which depends on temperature and on flow rate, is assumed to be of the form:

$$K_2 = k_{22}\, h_2^{(T-20)}\, Q^{k21} \qquad \text{(A3c)}$$

where k_{21}, k_{22} and h_2 are parameters.

The load B(*l*) in (A2b) is assumed to be recorded (i.e. exogenously given) and is the sum of a concentrated $B_c(l)$ load (an impulse-wise function computed along the characteristic line *l*(ô)) and of a distributed B_d load (a piece-wise constant function). The term $B_u(l)$ is a piece-wise constant function, that represents the known (or guessed) pattern of the unrecorded distributed load. As in the case studies in this paper, this pattern can be, for instance, the pattern for the equivalent inhabitants obtained by some census data. Notice that what is important is to identify the spatial distribution of this load and not its absolute magnitude, as this is taken into account by the scale factor u which is to be estimated. The functional problem of identifying non-point sources (the unrecorded distributed load) is thereby seen as a parametric problem.

The load term D(*l*) is the sum of three terms that represent the oxygen load introduced by external sources (for instance natural or artificial falls and aerators), concentrated and distributed water inflows, and a term D_f that represents the net photosynthetic oxygen production rate:

$$D_f = k_{23} + k_{24} F(t) \qquad \text{(A3d)}$$

where *F*(t(ô)) represents the daily fluctuation of photosynthetic activity (zero at night and sinusoidal during the day between sunrise and sunset) and k_{23} and k_{24} are parameters.

In conclusion, WODA has nine parameters (which are to be estimated), six of which, namely k_{11}, k_{12}, k_{21}, k_{22}, k_{23}, k_{24} (see equations (A3a), (A3c) and (A3d)), can be further refined by describing them with suitable piece-wise constant functions of the space (defined of course by a higher number of parameters).

The two parameters h_1 and h_2 (see equations (A3a) and (A3c)) are the only ones that can reasonably be assumed to be measurable in a laboratory test. They, however, can be identified from field data as well, when they are sufficiently numerous.