

## Transformation of measurements into input data for a river model

K. Beaugelin-Seiller and C. Brochier

*Institute for Protection and Nuclear Safety, DPRE/SERLAB, Laboratory of Environmental Modelling, CE Cadarache, bâtiment 159, BP. 1, 13108 Saint-Paul-lez-Durance, France*

**Abstract.** Before running a calculation code, it is necessary to collect and select the required entry data, data being a fundamental element upon which reasoning as well as research are based. In order to be rigorous, these practical elements should be managed in parallel with the modelling activities, aiming at common targets. For example, such a view is necessary for building relevant experimental designs. Indeed such methods require the collection of measurements through formalised protocols meeting the needs of the code. Such measurements could then define, for given space and duration, relevant data sets to "feed" the code. However, the calculation codes are frequently supplied *a posteriori* on the basis of measurements collected for other purposes. At first, the corresponding raw information must be collected and then adapted to the model needs, meaning a transformation of measurements into elaborated data. At this stage three kinds of problem exist: access to data, quality of data and quantity of available data. These issues are exposed here as part of the CASTEAUR project (simplified calculation of radionuclides transfer in receiving waterways), actually developed at the IPSN.

### 1. INTRODUCTION

In the field of the environment, in the largest acceptance of this word, the use of modelling is today largely widespread, whatever the aim may be. Consecutively to this approach, a lot of calculation codes have been developed. But using such a tool requires indispensable knowledge in order to determine its parameters and entry data values. Acquisition and selection of appropriate information, far from innocuous, condition both numerical results and their interpretation. In a rigorous process, experimentation, modelling and data acquisition would be lead together, according to a common view. In practice, experiments and models are often more or less closely linked: the parameter values are then well adapted to the calculation needs. But the entry data is rarely collected for the modelling purpose. Most often, the model user exploits information already acquired for other needs. He does then at first localise and collect the most adapted data: this brings up the question of data accessibility. Generally, they are measurements which require a minimum of transformation to be adapted to the model requirements. But a good adaptation calls for interest to two new kinds of problem: quality and quantity of available information. During the development of the CASTEAUR (simplified calculation of radionuclides transfer in receiving waterways) prototype, these difficulties were met and solved when testing the code on a real case. This paper presents the corresponding approach.

### 2. CONTEXT

#### 2.1. The CASTEAUR concept

CASTEAUR aims to assess the impact on fluvial ecosystems of both accidental and routine radioactive releases [1]. Derived from a phenomenological approach [2], this operational tool is structured over a simplified representation of the hydrographic network, associated with four other sequentially integrated models. Each of them is relative to a science necessarily drawn in the transfer of radionuclides in

hydrosystem, e.g. hydraulics, sedimentology, ecology and radioecology. Four kinds of radioactive releases can be considered, from a single pulse to a continuous pollution. The ecosystem is described by six components: water, suspended and deposited matters, primary producers, first order consumers and fish; the third last compartments constitute a simplified food web. Finally, the whole concept was formalised in a computerised prototype, which gives the possibility to mix the different kind of releases, pollutants and ecosystem components.

## 2.2. Data requirements

To test the most complicated configuration of the code, it was necessary to collect information for the five models. To make easier the data search [3], it was decided to work on a well-known trio, concerning the river (Rhône river, France), the releases (routine liquid effluents of Marcoule plant in 1993) and the radioactive nuclide ( $^{137}\text{Cs}$ ).

The description of the hydrographical net, divided in as many reaches as necessary in order to be homogeneous, needs the following geometric characteristics for each reach: length, width, depth, slope and bank angle.

The hydraulics model requires per reach the average flow on the study period and, if known, two specific coefficients (Strickler coefficient and longitudinal diffusion coefficient). But these last data can be determined by the code.

To solve the sedimentological aspects, the nature of suspended matters, e.g. mineral or phytoplanktonic, must be specified, with its average load value, in correspondence with the flow, its settling velocity and its shear stress for deposit.

Biological characterisation concerns only the last level of the food web, the fish. With regards to the studied species, the growing rate, the feeding rate and the dietary, spread between macrobenthos, zooplankton and juvenile fish, have to be informed.

The final assessment of radionuclide concentrations in the different parts of the hydrosystem uses radioecological parameters, which are mainly accumulation and depuration kinetics.

The step following the identification of the nature of the necessary data is their collect, which is exposed in the next paragraph.

## 3. ACCESS TO DATA

All the information required is provided by the water management field, except where radioecology is concerned. This last data is that acquired during the experimental studies performed in the IPSN Laboratory of Experimental Radioecology, and for which the conditions of determination are well known. They are then easily accessible and the whole knowledge about their acquisition conditions allows to check their compliance with the code requirements.

### 3.1. Water management in France

The water related data concerns a small part of a well studied and monitored river. But even in this case, a lot of actors are involved, considering the nature of data (biological, physical, hydraulic...), its acquisition periodicity (hourly, daily, weekly...) and its mathematical significance (instantaneous single value, average...). The knowledge about the whole functioning of a river section, and moreover of a river, is widely spread between users of the resource (industrials as well as associations), regulatory bodies, and scientists.

Collection the entire required data set and being sure that it covers all the needs, implies contact of each organism concerned in any way by the Rhône river management. The first stage, then, is an inventory of these information sources illustrated, as the following, for three of the four kinds of information, the fourth, connected to hydrography, does not create problems, on this part of the Rhône river.

### 3.2. Hydraulics

The main information researched is the flow of the river: on the Rhone river, at least four providers of such data are officially listed: the main one is the National Company of the Rhone river (CNR), which is in charge of river development; the second is Electricity of France (EDF), which manages some dams; and the last are two local authorities, the Regional Direction of Environment (DIREN) and the Water Agency Rhone-Corsica-Mediterranean sea (AE-RMC). Access to the CNR data is easy, buying a Web or Minitel connection. The national data base HYDRO [4], created in the sixties and less specific, is supposed to centralise all the flow data of French rivers, from all producers, responsible for the installation of the measurement stations, their maintenance, the data collection and the feeding of the base. Moreover, compatible with Windows only since 1997, this tool is really useful essentially for the most recent years but would be more and more attractive in the future, as ever more peoples involved in environmental activities are becoming aware of its interest. It is used by paying a subscription and purchasing a right to export data.

### 3.3. Sedimentology

In term of sedimentary dynamics, the accessible data through more or less regular *in situ* measurements are the nature of suspended matters (SM) and its load. The other parameters are issued from experimental works and/or bibliographical research.

The routine analysis of suspended matters in rivers carried out in France aims to give its concentration, without specification of mineral or organic part. In parallel, some controls of chlorophyll concentration are made, which give an indication of the phytoplankton abundance. On the Rhone river, two main organisations hold this information in correlation with the flow: the CNR and the AE-RMC, which pass on their data free of charge upon request.

### 3.4. Hydrobiology

The desired data is the dominant species in the study area and its growth rate, its dietary and its feeding rate. The latter information is closely linked with the age of fish, which must also be determined. That requires some biometric records, like weight and length.

Three sources of information exist for the Rhone river fish: the regional agency of the Superior Council of Fishing, the AE-RMC and the regional bureau of the Centre of Agricultural Mechanisation and of Forest and Water Rural Engineering (CEMAGREF). In the aim of this scientific study, the collected data were available free of charge upon request from each organisation. Several databases exist on this subject, but they contain only partial information. From a first resource, the dominant species was identified, with its biometric characteristics, which allow calculation of age and growth rate. A second investigation was necessary to find the corresponding dietary and a third search gives the feeding rate.

## 4. QUALITY OF DATA

### 4.1. Flow data

Four kinds of measures are available: instantaneous data and daily, monthly or yearly average flow. Some of these values are deduced from the first one, but the aggregation method is not well documented. Moreover, several measurements net coexist on the river, and sometimes the location of the measurement is not clearly established. In the HYDRO system, data producers are responsible for the validation and sometimes for the correction of this information. In the CNR database, some data is not validated. So, the quality of data closely depends on the possibilities, the needs and the good will of each producer...

#### **4.2. Suspended matter**

The measurements of suspended matter concentration are normalised. But even when respecting the norm, some practices are left to human appreciation, mainly related to the taking of samples. Most often, as for the chlorophyll determination, this data is acquired during punctual studies, without real sampling plane, especially concerning the frequency and the location of the sampling. It thus gives no follow-up, neither in time nor in space.

#### **4.3 Fish biometry and physiology**

The biometric data, like metric values, are easily well determined. The physiological data, not deduced from the previous, are generally issued from experimental work, and found in the literature. Qualifying their quality, then, is difficult, but their publication can be considered as a quality criteria.

### **5. QUANTITY OF DATA**

#### **5.1. Flow measurement**

As seen before, at least a daily value is available. The quantity of data is really not a problem in this case.

#### **5.2. Matter information**

The frequency of their collection is irregular in time as well as in space. For example, more than 5000 flow data has been collected in an other study on the Rhone river, when only about 2000 values of SM concentration are available in the same period and the same part of the river [5]. Where the chlorophyll measurement is concerned, this quantity falls to less than 400 data.

#### **5.3. Fish data**

The biometry of the Rhone river fish is well documented, as the CEMAGREF regularly monitors their populations. The physiological aspect is less known, at least for the studied species in this part of the river. Thus, some extrapolations are necessary to extend the knowledge acquired elsewhere [6, 7, 8] to the present case.

### **6. TRANSFORMATION OF MEASUREMENTS INTO INPUT DATA**

#### **6.1. Average flow on the study period**

Wanting an average value on a determined time period, and having daily measurements at one's disposal, the required data is the mean of these records on the considered time. Verifying the consistence of the incriminated measures, eventually through an inter-annual comparison, circumvents the eventual absence of validation by the producer.

#### **6.2. Nature and load of suspended matter**

The treatment of the suspended matters aspect faces two problems, quality and quantity. To solve the quality question, the consistence of the total set of each kind of data is verified, and the aberrant values are turned down. The quantity problem is not solvable: all the consistent values should be used as best possible. The few chlorophyll concentrations are analysed to determine whether the study period corresponds with an algal bloom: this enables the nature of suspended matter to be decided. Another way consists of transforming the pigment concentration into a number of cells per unit of volume and then, knowing the weight of a cell, into a load, which is to be compared with the total load. This one is deduced from the load measurements available, which are mathematically correlated with the flow data. Applied

to the average flow previously determined, this expression gives a load value, considered as the best estimates.

### 6.3. Age of fish

Concerning fish, data may be used directly, assuming ages of the individuals that very classic and validated biometric models [9, 10] enable calculation.

## 7. CONCLUSION

It appears through such approaches that the gap between measurements and models requirements is still wide. Three axes of work have to be explored. First, the data acquisition is still problematic in terms of both quality and quantity: it is necessary to ameliorate the significance of the measurement itself, identifying uncertainty sources and always making sampling plans. In parallel, it would be interesting to revise the finality of measurement nets, including new objectives, related to a widest use of measures. The second point concerns the means used to put information at disposal of users: the database concept must be encouraged and promoted, including the necessity to record sufficient information to recount the story of each measure. The third and last way of thinking is related to the aggregation methods, which should be a function of the use aim: they have to identify measurements that can be used and the manner in which to use them.

All these considerations can easily be carried out if both model designers and data producers are convinced that they have to work together more closely than ever.

## References

- [1] Boyer P., Garnier-Laplace J., Beaugelin-Seiller K., Adam C., El-Ganaoui O. Prototype CASTEAUR : note de présentation générale. Report SERLAB/00-017 (IPSN, Paris, 2000).
- [2] Boyer P. and Garnier-Laplace J. Projet TRANSRIV – Système d'évaluation des transferts au sein des rivières. Report SERE/98-015 (IPSN, Paris, 1998).
- [3] Beaugelin-Seiller K. and Rommens C. Prototype CASTEAUR : enseignement d'un cas-test. Report SERLAB/00-013 (IPSN, Paris, 2000).
- [4] Ministère de l'Aménagement, du territoire et de l'environnement. HYDRO, banque nationale de données pour l'hydrométrie et l'hydrologie. Presentation notice (2000).
- [5] Brochier C. Contribution à l'acquisition, la selection et la mise à disposition des données nécessaires aux codes de calcul TRANSRIV et CASTEAUR. Engineer thesis (ISIM, Montpellier, 2000).
- [6] Hellawell J.M. *Freshwat. Biol.* **1** (1971) 369-387.
- [7] Le Louarn H., Baglinières J.L., Marchand F., Hamonet J.M.. *Bull. Sci. Tech. INRA Dept Hydrobiol.*, (1997)
- [8] Mann R.H.K. *J. Fish. Biol.* **8** (1975) 265-288.
- [9] Bhattacharya C.C. *Biometrics* **23** (1967) 115-135
- [10] Von Bertalanffy L. *Hum. Biol.* **10** (1938) 181-213