Feeding your models with real business data

Magne Myrtveit
ModellData AS
N-5120 Manger, Norway
Phone: +47 5637 4009
Fax: +47 5637 3500
Internet: magmy@modeld.no

(ABSTRACT)

In making models of real world problems or real world processes, the need for transfer of data between model and external files, applications or other devices becomes important. This paper illustrates the use of automatic data transfer in model construction and in process control.

In a typical model construction and verification process, real data plays an important role. Once a model is made of the target system, real data is fed into the model, and the model output is compared to the output of the real system. The result of the comparison is used as a basis for refining model structure, model parameters, or both.

The Dynamic Data Exchange (DDE) object and the Archive object of POWERSIM may be used in a model to import and export data during simulation. This greatly reduces the amount of work involved in checking models against reality. DDE can also be used to set up a data analysis, for example, by launching a spreadsheet or a statistical program. In combination with the multiple runs feature of POWERSIM, a series of model structures and model parameters can be tested with real data without manual intervention.

In a research project aiming at developing a model of the biological process of fish growing in a fish farm, the above tools have allowed for extensive use of data from fish farms in the process of model construction and verification.

The basic feedback loop of a process control system, involves measurement (import of data), comparison between measured values and desired values, followed by a controlling action (output of data) aiming at closing the gap between actual and desired process state. The data transfer protocols of POWERSIM make it possible to implement process control systems as POWERSIM models, and linking the models to the physical process. A range of input/output devices for the Windows operating system supports DDE or Net-DDE, and may therefore be used directly together with POWERSIM.

Keywords: Dynamic Data Exchange (DDE), model verification, process control, data transfer, communication

Tools for modeling real businesses

A variety of tools are used in describing, measuring, storing, retrieving, transferring, analyzing and presenting business data.

Databases and spreadsheets are commonly used to store current and past business data. Statistics can be produced based on time series (history) of data. System behavior can be described using time graphs and time tables that may be produced easily using a variety of tools, including word processors, presentation software, spreadsheets, and database front-ends. System structure (relationships, cause and effect) is often described as static maps (organization charts, flow charts, etc.). For simple systems, differential mathematics can be used to describe system structure and infer system behavior. Using statistical methods (models), projections may be made into the future, for example by extrapolation. Even within the field of systems thinking, a variety of tools are used (Lannon-Kim, 1994).
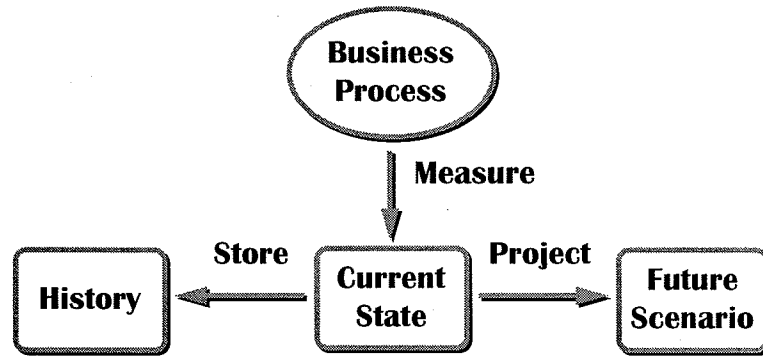
Simulators represent a valuable extension to the set of business tools. A main use of simulation models is to describe a system (for example a problem, an organization or a process), and to carry out simulations mirroring the behavior of the system. The dynamic nature of simulation models adds a new dimension business modeling.

In verifying a model, recorded history is compared to simulated output from the model. If there is a close match between model behavior and actual history, the model will most likely be a valid description of the real system.

Even valid models can fail significantly in predicting the future of a business. One reason for this is that for most business processes there will be a number of uncertain factors that will influence the system. We therefore have to make assumptions about external factors when making simulations into the future. Fed with different sets of assumptions, a simulation model can produce quite different results. This makes it possible to get answers from the model to what-if questions, and test consequences of assumptions.
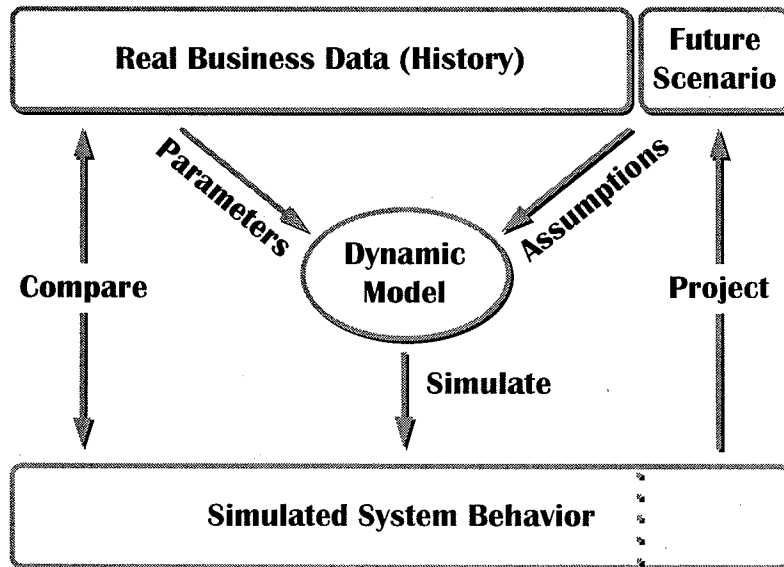
Figure 1: Past, current and projected business data

Figure 2: Transfer of data to and from simulation model

A dynamic model is a virtual world

Creating a simulation model can be looked upon as creating a virtual world representing parts of the real world. We can think of modeling as a goal seeking loop where the real world is the goal, the current model is the actual value, and the modeling process is the corrective action. (The balance icon is used to denote a balancing feedback loop.)
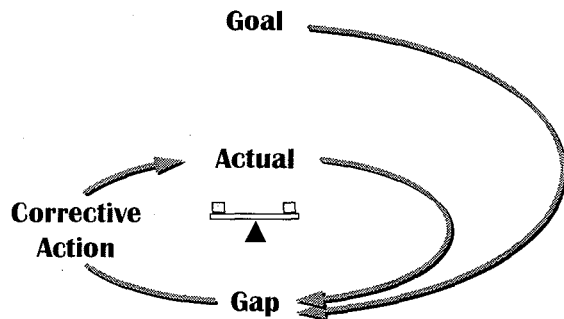
Goal   **Real World**

Actual   **Virtual World**

Corrective Action   **Modeling**

Gap   **Gap**
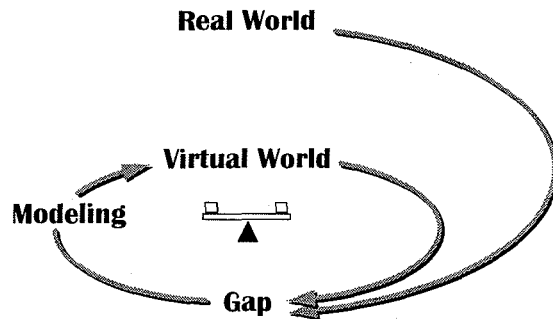
Figure 3: Balancing loop   **Figure 4: Creating a virtual world**

If we change the names in Figure 3 a little, we get a picture of the modeling process, as shown in Figure 4.

In determining the gap between Virtual World (model) and Real World we may use qualitative and quantitative measurements.

Qualitative comparisons between model and real world

Qualitative comparisons between model structure and relationships in the real world are of great value in the conceptualization of a problem. The field of systems thinking has developed numerous tools for analyzing a problem and developing a simulation model describing the problem and its behavior. Using causal loop diagramming, for example you may create a high-level map of any dynamic problem. Or you may try to fit your problem into one of the system archetypes that have been identified.

If you choose to use causal loops to describe your problem, studying feedback loops can shed light on important characteristics of the model's behavior. And if you use archetypes to describe your problem, the typical behavior of each archetype can be used to make rough statements about basic behavior of your real world problem (Senge et al. 1994).

Quantitative comparisons between model and real world

A simulation model can be compared to a hypothesis about a phenomenon. By comparing model behavior to the real world we can determine the validity of our hypothesis. A model that corresponds well to the phenomenon it describes, can be used as a theory (see Figure 5).

In the range between scientifically verified theories and vague statements there is room for a broad variety of models. Based on the intended use of a model, the level of necessary accuracy will vary. As an example, a model for educational purposes can be useful even if it does not give

100% accurate output. In the business world, however, we often encounter situations where even small variations in model output are of importance.
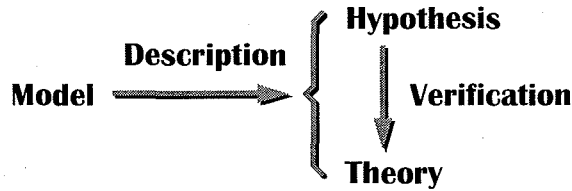

Figure 5: Models describe hypotheses and theories (verified hypotheses)

The final adjustments of a model normally involves parameter tuning. In Figure 6 we have added labels to the arrows from Real World and Virtual World to Gap. One set of data is obtained by gathering statistics about the real world, while another set is produced as output from a simulation run.

The sets of measured and simulated data can be compared to compute a numerical gap between the real world and the model. The gap can be computed in different ways. A simple solution is to compute the difference between actual and simulated data. Another approach is to compute the sum of the squares of the differences between actual and simulated data. The latter approach may be useful when several variables are compared, or when time series are compared. (The term *time series* is used when we refer to a set of values for a variable at different times.)
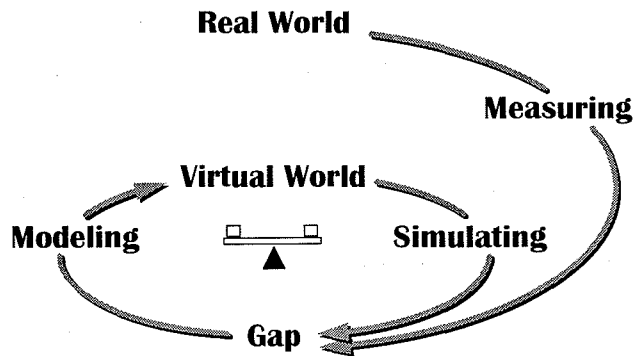
In Figure 7 we have made a table with a hypothetical actual time series [5, 7, 3, 4] and a simulated series [4, 5, 2, 3]. The differences between each element of the actual and the simulated series are computed and squared. By adding together the numbers in the final row we get 7, which is a measurement for the gap between actual and simulated values.

Figure 6: Comparing simulation output to real world measurements

Time	1	2	3	4
Actual (A)	5	7	3	4
Simulated (S)	4	5	2	3
Difference (D=A-S)	1	2	1	1
D*D	1	4	1	1

Figure 7: Computing the gap between actual and simulated data

If the actual values are imported into the simulation model, the gap may be computed during simulation, like this:
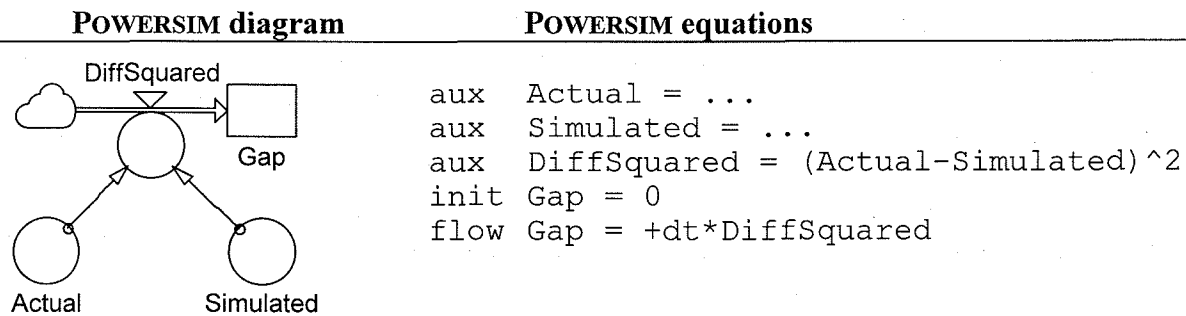
POWERSIM diagram	POWERSIM equations


```
aux   Actual = ...
aux   Simulated = ...
aux   DiffSquared = (Actual-Simulated)^2
init  Gap = 0
flow  Gap = +dt*DiffSquared
```

Figure 8: Computing the gap as part of a simulation

The definitions of *Actual* and *Simulated* will vary from model to model.

Importing data into the simulation

Values of external factors (observed or assumed) influencing a model are often stored in data files. The same is true for historical business data that we may want to include in our simulation for comparison with simulated vales.

For small data sets it may be OK to type the data into the simulation model, for example as parameters to a table look-up function. In a real business situation, however, we may want to experiment with different data sets, and each set of data may be quite large. This means that typing data into the simulation model becomes inflexible, time consuming, and error prone. We should therefore look for methods of automatic data transfer between a simulation model and the environment.

POWERSIM includes mechanisms for accessing external data files directly, or indirectly, via communication with applications managing the data (for example spreadsheets).
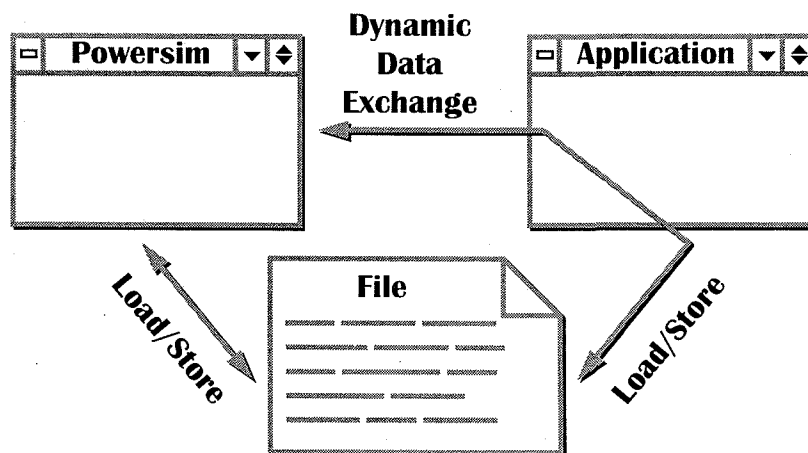

Figure 9: Data transfer with files and applications

Importing data from files
POWERSIM can read text files directly into a simulation using the archive object. Setting up a transfer is done by creating an archive object in the model, and connecting a data file to one or more parameters (scalar variables or array elements) of the model.

Text files with data to be imported into POWERSIM can be created in several ways. You may, for example, use the Windows text editor Notepad or any other application that is able to store files as plain text. Most spreadsheets will allow you to choose Text as output format when storing a spreadsheet using the File Save As command. The same is the case for most word processors.

POWERSIM can read data either as time series or as single values. As an example, we may enter the time series *Actual* that was used in the previous chapter, into a text file ACTUAL.TXT. The data can be organized in the file as displayed in Figure 10.

A tab character is used as column separator. Options to the archive object are used to match the format of the imported data, for example choice between horizontal or vertical orientation, presence of time series row or column, presence of row or column headers, choice of column separator, numerical format.

Time	Actual
1	5
2	7
3	3
4	4

Figure 10: Text file holding time series to be imported into POWERSIM

If we want to import the data in Figure 10 into our simulation model, we define a constant variable *Actual* to receive the data. We then open the archive object to make a connection between the constant and the text file ACTUAL.TXT. Before simulating, we should define the time horizon for the simulation to be from 1 to 4, as this is the range for which we have data. We can now run the simulation.

In order to verify that the variable Actual takes on the values stored in the text file, we can create a time table object and make Actual a parameter of the time table. The resulting diagram is displayed in Figure 11.
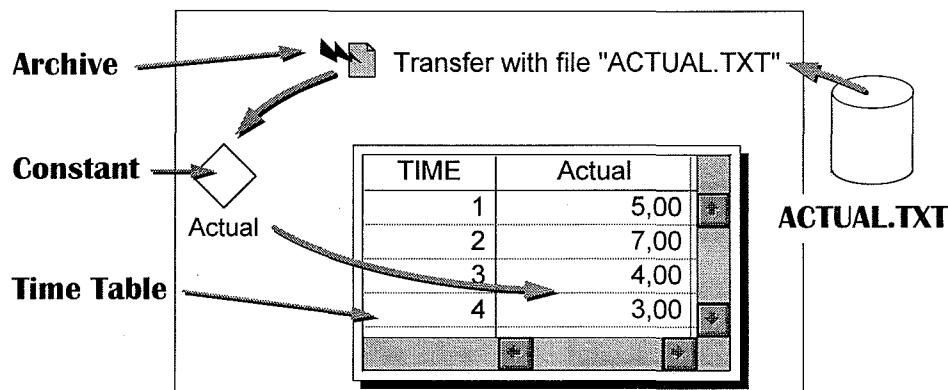

Figure 11: Diagram with constant, archive and time table

Importing data from applications

Dynamic Data Exchange (DDE) is the name of a standard protocol for data transfer between MS-Windows based applications. Most industry strength applications – including POWERSIM, MS-Word, MS-Excel, and Lotus 123 – support DDE.

DDE is based on a client server architecture, where a client application can ask a server application for services. POWERSIM's DDE object defines the client end of a DDE link, allowing POWERSIM models to transfer data to or from a DDE server application.

	A	B
1	1	5
2	2	7
3	3	3
4	4	4

Figure 12: Spreadsheet with time series

As an example, let us enter the data in the file ACTUAL.TXT into a spreadsheet as displayed in Figure 12.

In order to obtain data from a DDE server (for example a spreadsheet) we must inform the DDE object about:

Server – the name of the server application, for example Excel
Topic – the name of the document where the data is stored, for example ACTUAL.XLS
Item – the part of the document where the data is located, for example R1C1:R4C2

The DDE object in POWERSIM is almost identical to the archive object. The main difference is that File in the archive object is replaced by Server, Topic and Name in the DDE object.

Exporting data from the simulation

By changing the transfer to *To File* for the archive object and *To Application* for the DDE object, data can be transferred from a simulation to a file or another application, respectively.

It is possible to use several archive and DDE objects in one model, allowing for transfer of data in both directions. As an example, while data is read from one column of a spreadsheet, simulation output can be written back into another column.

The possibility to store simulation results in a file or into documents managed by other applications, opens for storage, transfer, analysis, and presentation of simulation output using applications of your choice.

Closing the circle of continuos improvement

Creating a micro world may serve several objectives, including;

- gaining insight into the real world through the modeling process,
- training people with management flight simulators based on running simulations,
- designing strategies for improving the real world using experimental simulation.

POWERSIM is designed to support all of the above activities. In Figure 13 we display how a continuos improvement process can be implemented.

Each of the arrows in Figure 13 involve actions that must be carried out in order to close the loop. POWERSIM's data transfer capabilities are important in importing measurements from the real world into a simulation. Advanced modeling capabilities like arrays and co-models help make realistic models (Myrtveit 1994). For creating custom interfaces to micro worlds POWERSIM contains authoring capabilities (Davidsen, Myrtveit 1994), and can also be integrated with custom interface
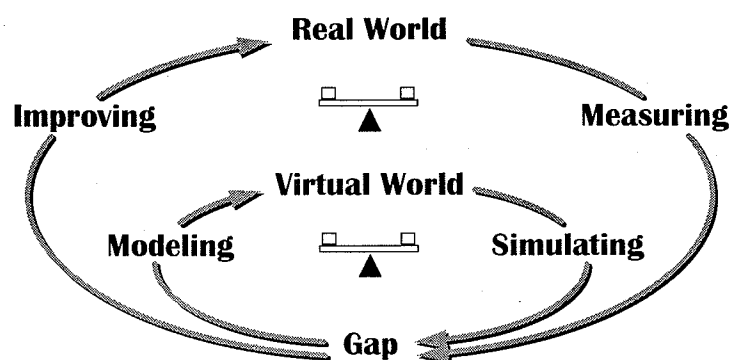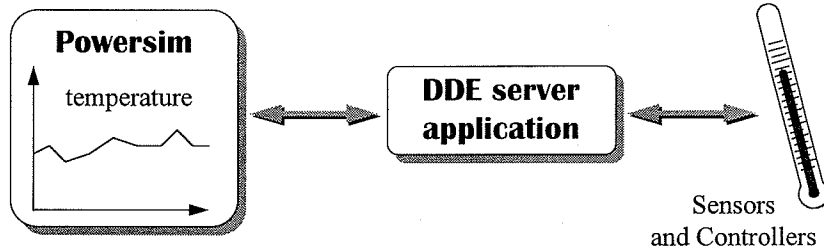
Figure 13: Simulation as a basis for continuos improvement

applications using both DDE and Application Programming Interface (API) calls from languages like C, C++ and Visual Basic (Myrtveit, Vavik 1995).

Using DDE to control a process

A DDE link can be automatic ("hot") or manual ("cold"). An automatic link will transfer data whenever the source data is altered. This can be used to interact with a model during simulation. It is, for example, possible to set up a DDE connection between POWERSIM and Excel where a POWERSIM variable is updated whenever the value of a given cell is changed in

Figure 14: Connecting a model to the external world

the spreadsheet. The opposite connection is also possible, i.e. that a spreadsheet cell is updated continuously during a simulation based on a POWERSIM variable.

Input/output equipment for lab experiments or process control are often shipped with software components defining a DDE interface to the equipment. Using such interfaces POWERSIM models can be connected to the physical world.

Example – a virtual fish farm

The fish farming industry is important along the entire cost of Norway. The world's largest aqua research laboratory is located here, with access to lots of research data and skilled researchers. A project was launched together with the Norwegian environmental authorities to develop a better understanding of the relationship between production, economy and wastes to the environment.

A team was put together with people from two research laboratories, a fish feed producer and myself. It soon turned out that the research communities had access to advanced knowledge in isolated areas such as influence of light, currents, temperature, oxygen feed and feed composition on fish growth. At the same time there was an obvious need for combining the fragmented knowledge into a unified whole. We therefore choose to create a virtual world *aquaSIM*™ portraying the fish farming system. The resulting model would be verified using existing research data. If necessary new data would be obtained from the research stations and co-operating fish farms. Finally, the model would be presented in two versions; one for research purposes and one for planning purposes.

The model was designed around the use of co-models, where two different interfaces were used; one for the research model and another for the planning version.

A micro world for planners

For pedagogical purposes the planner's micro world was designed to visualize differences between alternative production methods. Each production decision was made a parameter to the *aquaSIM* model. By varying the parameters, different scenarios can be played out.

By interacting with slide bars, number objects and buttons in the interface model (main) the planner can set different parameters in the reference model and the alternative fish farm model (the two *aquaSIM* co-models in the bottom of Figure 15).

By running two alternative scenarios concurrently it becomes easy to compare the two simulations by displaying time series from both models in the same reports. As an example, the feed consumption, growth and pollution can be displayed in time graphs where one color (say gray) is used for the reference model and another color (say red) is used for the alternative. In Figure 17 the growth (weight) of fish is compared. The alternative co-model (labeled Actual) uses a feeding intensity of 80% of the reference co-model.

As can be seen from Figure 16 the user may choose between different productions (species of fish), locations (temperature zones) and feed compositions. The effects of these parameters on the model are controlled by parameters.

The actual parameters are stored in external files, which are loaded into the simulation at simulation start-up using the archive object, described earlier.

A micro world for researchers

This application was developed in order to gain insight into the fish farming process. One key issue was to validate the model. This was set up along the lines
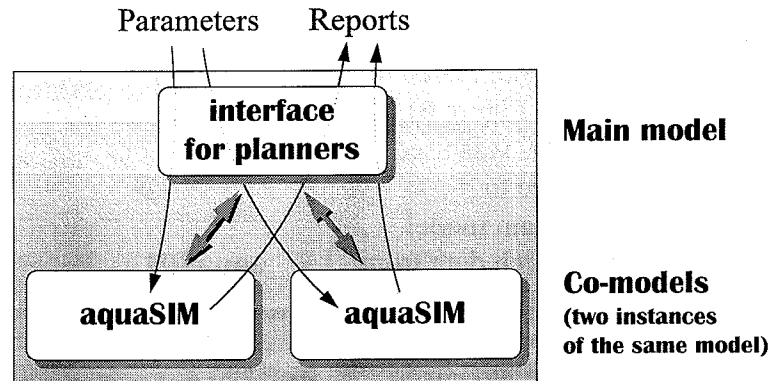

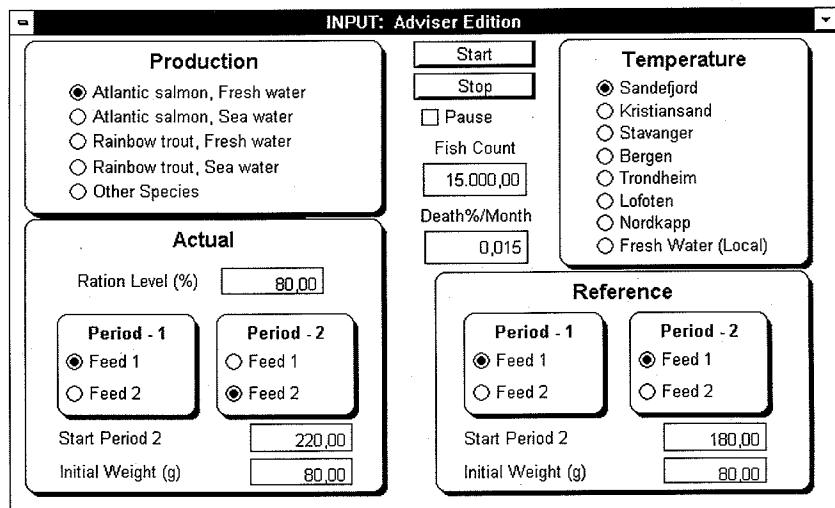**Figure 15: Structure of the planner's fish farm model**

Figure 16: Giving input to two simulations (Actual and Reference) from the Main model
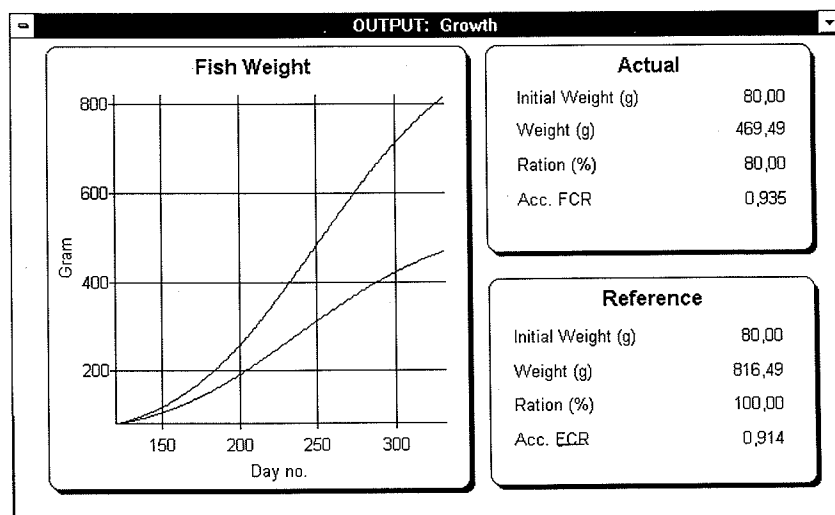

Figure 17: Comparing growth at different feeding levels

described earlier in this paper. A simulation model was fed with research data and the simulated output compared to the real data (see Figure 6).

A main model was used to create an interface to *aquaSIM*. The main model would read research data from a spreadsheet, run *aquaSIM* as a co-model and output simulation results alongside the real data in the spreadsheets.

Figure 18: Importing model parameters from external files using archive object

Since both the researcher's version and the planner's micro world uses *aquaSIM* as the core co-model, any improvement to the research model will automatically be part of the planner's version.

The first stage of the *aquaSIM* project was completed this spring (1995). The researchers give the feedback that they now have a powerful new tool for understanding the complex nature of fish farming. The model building process has uncovered new areas for further investigation. Looking at the initial objective of the project, the simulations show that the most economic way to run a

Figure 19: Structure of research model and connection to real data

fish farm is also the way that causes the least damage to the environment. This corresponds well with research data, and can be visualized in convincing ways using the *aquaSIM* POWERSIM models. The vast amounts of that that was necessary in validating process could not have been used without access to automatic data transfers.
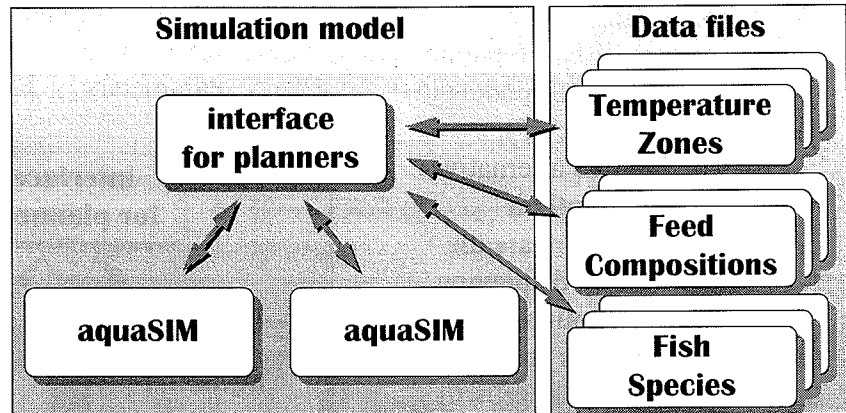
References

Davidsen, P.I., Myrtveit M. 1994. *Der RÜTLI Management Simulator*. Proceeding International System Dynamics Conference, 9–21.

Lannon-Kim, C. 1994. *A beginner's guide to systems thinking*, Pegasus Communications, Inc.

ModellData AS 1994. *POWERSIM 2.0 User's Guide and Reference*

Myrtveit, M. 1994. *Powerful modeling using array variables*. Proceeding International System Dynamics Conference, 173–184.

Myrtveit, M., Vavik L. 1995. *Object Based Dynamic Modeling*. Proceeding Intern. System Dynamics Conference.

Senge, P.M., Roberts, C., Ross, R.B., Smith, B.J., Kleiner, A. 1994. *The Fifth Discipline fieldbook*. Nicholas Brealey Publishing Limited, London.