CHAPTER 5.

DISCUSSION AND CONCLUSIONS

5.1. The SALMO-OO simulation library as a tool for ecosystem analysis and management

The SALMO-OO model is a generic model and has been designed to simulate a wide variety of lake conditions. Generic models often pose the problem of generality versus accuracy, where generic models tend to forfeit a degree of accuracy in order to simulate realistic, general conditions (Levins, 1966). The SALMO-OO model in its original form simulates a variety of lake conditions very well, whilst still maintaining an acceptable level of accuracy. By applying a library of process models, at this stage for phytoplankton growth and grazing processes, it is anticipated that the scope of the model can be broadened and the accuracy of simulating key state variables, such as phosphate concentration, phytoplankton biomass and algal functional group dynamics, can be improved. The simulation library introduces an additional validation measure in order to improve the simulation of specific lake ecosystem by the use of expert knowledge. To further improve the accuracy of any deterministic model built on expert knowledge additional computational technologies, such as evolutionary algorithms or parameter optimisation techniques, can be included in the development of hybrid decision support systems, which have become popular with environmental scientists and managers.

5.1.1 Outcomes of the experiments using the phytoplankton growth and grazing simulation library for lakes with different environmental conditions

The SALMO-OO simulation library has been developed with a selection of phytoplankton growth and grazing process models that have been taken from literature models developed for specific lake systems. Three phytoplankton growth and three grazing process models are included within the simulation library. In order to improve the simulations of key state variables for various lakes it was necessary to compare the simulation of different combinations of growth and grazing process functions with those process functions from the original SALMO-OO model. The experimental procedure for analysing all relevant combinations can be a tedious and time-consuming process for a non-expert, such as an environmental manager. An ideal simulation model is one that is simple to use, is informative and can give worthwhile predictions (Steel, 1997). Therefore, in order to develop a simple decision support system for the SALMO-OO model, categories based on trophic state and mixing conditions were created and suitable model structures were found from the simulation library to improve the simulation of lakes within each category. At least two lakes were tested for each category. The model structures for each category are based on the realistic simulation of phosphate concentration, zooplankton and phytoplankton biomass and algal functional group dynamics. The categories were selected as an improvement to the original SALMO-OO modelling results. In all cases improvements to the simulation of phytoplankton biomass and algal functional group dynamics in particular were achieved by replacing the original SALMO-OO growth and grazing process models with those from the simulation library. Table 5.1 summarises the findings presented in section 4.2 for each lake category.

Table 5.1. Summary of generic model structures found by the SALMO-OO simulation library for different categories of lakes and reservoirs based on trophic state and mixing conditions. **AB** – Arhonditsis and Brett (2005); **HJ** - Hongping and Jianyi (2002); **CL** - CLEANER Model - (Park *et al.*, 1974; Scavia & Park, 1976).

Trophic State	Mixing Conditions	Best combination from SALMO-OO simulation library	Validation data sets
Eutrophic and Hypertrophic	Dimictic	Growth CL and Grazing HJ	Bautzen reservoir, Germany Lake Arendsee, Germany
	Warm Monomictic	Growth AB and Grazing AB	Lake Hartbeespoort, South Africa Lake Roodeplaat, South Africa Lake Klipvoor, South Africa
Mesotrophic	Dimictic	Growth CL and Grazing AB	Saidenbach reservoir, Germany Weida reservoir, Germany
Oligotrophic	Dimictic	Growth CL and Grazing CL	Lake Stechlin, Germany Lake Soyang, South Korea

It is apparent that the growth model from CLEANER contributes most frequently to the improved results for each state variable analysed. The CLEANER growth model has a unique formulation for the calculation of photosynthesis, which may contribute to the success of this growth function. Phytoplankton photosynthesis (PHO*i*) is described by a maximum photosynthesis rate (PHOMAX*i*), modified by suboptimal conditions (Ut) and a temperature limitation function. The subscript *i* denotes the different phytoplankton functional groups (*i* =1 diatoms; *i* = 2 green algae; *i* =3 blue green algae). The combined limitations of light and nutrients are represented as a normalised factor (Ut) that is mathematically analogous to the inverse of the total effect of electrical resistors in parallel (Park *et al.*, 1974; Scavia & Park, 1976).

$$PHO_i = PHOMAX_i * Ut * PHOT_i \tag{1.1}$$

$$Ut = \frac{N}{\sum_{i}^{n}\left(\frac{1}{f(U)_i}\right)} \quad \text{Therefore,} \quad Ut = \frac{3}{\left(\left(\frac{1}{f(PHOL_i)}\right)+\left(\frac{1}{f(PHOP_i)}\right)+\left(\frac{1}{f(PHON_i)}\right)\right)} \tag{1.2}$$

The mean resistance construct was developed specifically for CLEANER, where *n* is the number of limiting functions and is used to normalise the total limitation term (Ut). $f(U)_i$ are the normalised individual limitation functions. The CLEANER model developers state that if no nutrient is limiting, the function is totally limiting (Park *et al.*, 1974; Scavia & Park, 1976). Their reason for preferring this new construct is intuitive. The authors believe it is reasonable to assume that adaptation and species replacement in a natural assemblage will moderate the limiting effect of any particular nutrient or combination of nutrients. Therefore, this construct may better represent the actual limitation process at the ecosystem level. The CLEANER model developers examined various formulations for the limitations of light and nutrients on photosynthesis, including the commonly used multiplicative and minimum constructs. The developers found in addition to the mean resistance construct the minimum function produced excellent results, however, experiments with the multiplicative function, seemed to limit photosynthesis more severely than is actually observed in nature (Park *et al.*, 1974; Scavia & Park, 1976).

Three types of photosynthesis formulations are accounted for within the phytoplankton simulation library:

1. The multiplicative function
2. The minimum function, otherwise commonly known as Liebig's Law of the Minimum construct, and
3. The mean resistance construct developed for CLEANER.

The multiplicative function is simply the maximum photosynthesis rate multiplied by sub-optimal conditions, which generally includes limitations due to light intensity, water temperature and different nutrient concentrations (equation 1.3). This construct is the most commonly used to calculate the photosynthesis rate of phytoplankton (Table 5.2). Both SALMO-OO and the growth model from Hongping and Jianyi (2002) use the multiplicative function to calculate photosynthesis.

$$\mathrm{PHO}i_{T,E,H} = PHOMAXi * \mathrm{PHOL}i * \mathrm{PHOT}i * \mathrm{POHP}i * \mathrm{PHON}i \tag{1.3}$$

The minimum function, based on Liebig's Law of the Minimum, is used to a lesser degree then the multiplicative function, but is still a popular method of calculating the photosynthesis rate for phytoplankton. Liebig's Law of the Minimum states that the yield of any organism is determined by the abundance of a substance that in relation to the needs of the organism is least abundant in the environment (Jorgensen, 1994). In many cases the minimum function is used to determine the limiting nutrient only rather then the single limiting resource, but in some cases light intensity is also included as shown in Table 5.2. The temperature limitation function is often excluded from the minimum construct, as temperature will influence the growth rate even if light or nutrients are limiting (Zonneveld, 1998). The growth model from Arhonditsis and Brett (2005) only uses the minimum function to determine which nutrient is the most limiting (equation 1.4). However, within their lake model silicon and carbon are also considered as nutrients affecting phytoplankton metabolism and are thus included in the minimum function, whereas within the SALMO-OO simulation library only phosphate and nitrogen state variables are considered.

$$\mathrm{PHO}i_{T,E,H} = PHOMAXi * \mathrm{PHOL}i * \mathrm{PHOT}i * \min\left\{\mathrm{PHOP}i, \mathrm{PHON}i\right\} \tag{1.4}$$

Table 5.2. List of various modellers/author's that have elected to use either the multiplicative or minimum function for the calculation of gross phytoplankton photosynthesis rate.

Multiplicative function for gross photosynthesis	**Liebig's Law of the Minimum function for gross photosynthesis**
$\mathrm{PHO}i_{T,E,H} = PHOMAX_i * \mathrm{PHOL}_i * \mathrm{PHOT}_i * \mathrm{POHP}_i * \mathrm{PHON}_i$	$\mathrm{PHO}i_{T,E,H} = PHOMAX_i * \min \{\mathrm{PHOL}_i, \mathrm{POHP}_i, \mathrm{PHON}_i\} * \mathrm{PHOT}_i$
Arhonditsis & Brett, (2005)	Canu *et al.* (2004)
Bonnet & Poulin (2002)	Chen *et al.* (2002)
Canale *et al.* (1976)	Drago *et al.* (2001)
Childers & McKellar Jr. (1987)	Hamilton & Schladow (1997)
Hongping & Jianyi (2002)	Menshutkin *et al.* (1998)
Imboden & Gächter (1978)	Robson & Hamilton (2004)
Jayaweera & Asaeda (1996)	Rukhovets *et al.* (2003)
Kmet & Straskraba (1989)	Sagehashi *et al.* (2000)
Krivtsov *et al.* (1998)	Scavia (1980)
Lima *et al.* (2002)	
Mesple *et al.* (1995)	
Norberg & DeAngelis (1997)	
Parker (1968)	
Benndorf (1979)	
Recknagel (1980)	
Thebault & Salencon (1993)	
Varis (1993)	
Varis (1984)	
Yang *et al.* (2000)	

The multiplicative and minimum functions are two vary common equation types to calculate phytoplankton photosynthesis rates, however, the CLEANER mean resistance equation for photosynthesis rate was found to give the most improvement to the simulation of phytoplankton biomass, when linked to any of the three grazing models included in the simulation library, which incidentally is formulated in a similar manner. Only the simulation of the South African lakes, which are all considered as hypertrophic and warm monomictic, improved by the addition of the growth and grazing models from Arhonditsis and Brett (2005), rather than the CLEANER growth model. Interestingly, out of all the published lake and water quality models that I searched the CLEANER mean resistance construct has never be used outside of the CLEANER family of models.

The three literature models included in the SALMO-OO simulation library offer a variety of mathematical formulations for the calculation of phytoplankton photosynthesis, respiration and grazing by zooplankton. Each model is based on strong and rigorous scientific principles that have been in use for many years in many different lake models (Eppley, 1972; Jorgensen, 1994; Reynolds, 1993; Riley & Stefan, 1988; Steel, 1997; Steele, 1962; Straskraba & Gnauck, 1985; Zonneveld, 1998). Each process model has been extensively validated, and shown to perform realistically and accurately within the models from which they have been taken. Therefore, we can be confident that the assumptions, theory and outputs of these process models can be relied upon to perform as they were designed.

There are many different modelling techniques and paradigms available to conceptualise and problem solve for a natural system, and as a result there are a variety of model structures to choose from to simulate different lake conditions. However, the difficulty in dealing with higher-level problems is the complexity involved in such models, which are difficult to define through a single, consistent system of paradigms (Villa, 2001). Many researchers acknowledge that more than one modelling paradigm or technique is often needed to capture the minimal description of a complex system, to address the complexity and multidisciplinary nature of environmental problems and to recognise the cross-scale effects with different phenomena operating at different scales (Dale & Swartzman, 1984; Riley & Stefan, 1988; Steel, 1997; Swartzman, 1979; Villa, 2001; Zonneveld, 1998). Thus, integrating a greater choice of modelling options, whether they are in the form of hybridised technologies or alternative model structures, into a single framework can assist in the development of more informative, "smarter" models and make the computerised decision-making process easier. There are many modelling studies that have been published where an individual or a group of researchers either review different mathematical formulations of certain processes, such as the mathematical modelling of the photosynthesis-irradiance curve, or present the results of a model development project where several alternative structures have been tested and critiqued to find the optimum model structure to represent the system or problem under study (Canale *et al.*, 1976; Dale & Swartzman, 1984; Park *et al.*, 1974; Riley & Stefan, 1988; Scavia & Park, 1976; Sequeira *et al.*, 1991; Swartzman, 1979; Todorovski, 2003; Zonneveld, 1998). The simulation library offers a single platform to explore model structure and behaviour for both cases, and uniquely, for a wide variety of different lake conditions.

5.1.2. The SALMO-OO simulation library as a validation tool

During the developmental stage of any computer model, the validation process is crucial in perfecting the accuracy of the models' outputs, both quantitatively and qualitatively. The phytoplankton simulation library has been designed as a validation tool to improve the prediction of different lake conditions by the SALMO-OO model. By replacing the original SALMO phytoplankton growth and grazing process functions with those in the simulation library a better approximation of a particular lake's dynamics can be achieved. However, the selection of the best model structure for each lake dataset is a subjective process that relies upon expert knowledge of lake ecosystem interactions. This presents possible problems in "differences of opinion" in which model structure is actually the best available for each lake category. A procedure of analysis was developed to minimise these decision discrepancies by relying on several criteria for the selection of the best model structure. These criteria are summarised in Figure 5.1 (a more detailed explanation can be found in section 3.6.3):

A major emphasis of the selection criteria is the importance on qualitative analysis of the simulation results. To make an informed decision as to the best results for a given lake it was necessary to graphically view observed and simulated outputs to determine whether or not the results adhered to the current expert knowledge of lake ecosystem dynamics and if the model was able to describe significant lake conditions (i.e. major algal blooms or high nutrient concentrations). Quantitative analysis of the model outputs with the measured data from each lake did not always give very informative answers to how well the model was performing and how adequately the model described lake dynamics. The quantitative methods used to analyse model performance were linear regression (r^2) and root mean square error (RMSE). Originally, only r^2 was calculated to assess model performance, as this "goodness-of-fit" statistic is commonly used as a model performance yardstick in the ecological modelling field. However, in many cases such unrealistic discrepancies between the r^2 and the qualitative results occurred that I included the RMSE statistic as an additional statistical means of testing model performance.

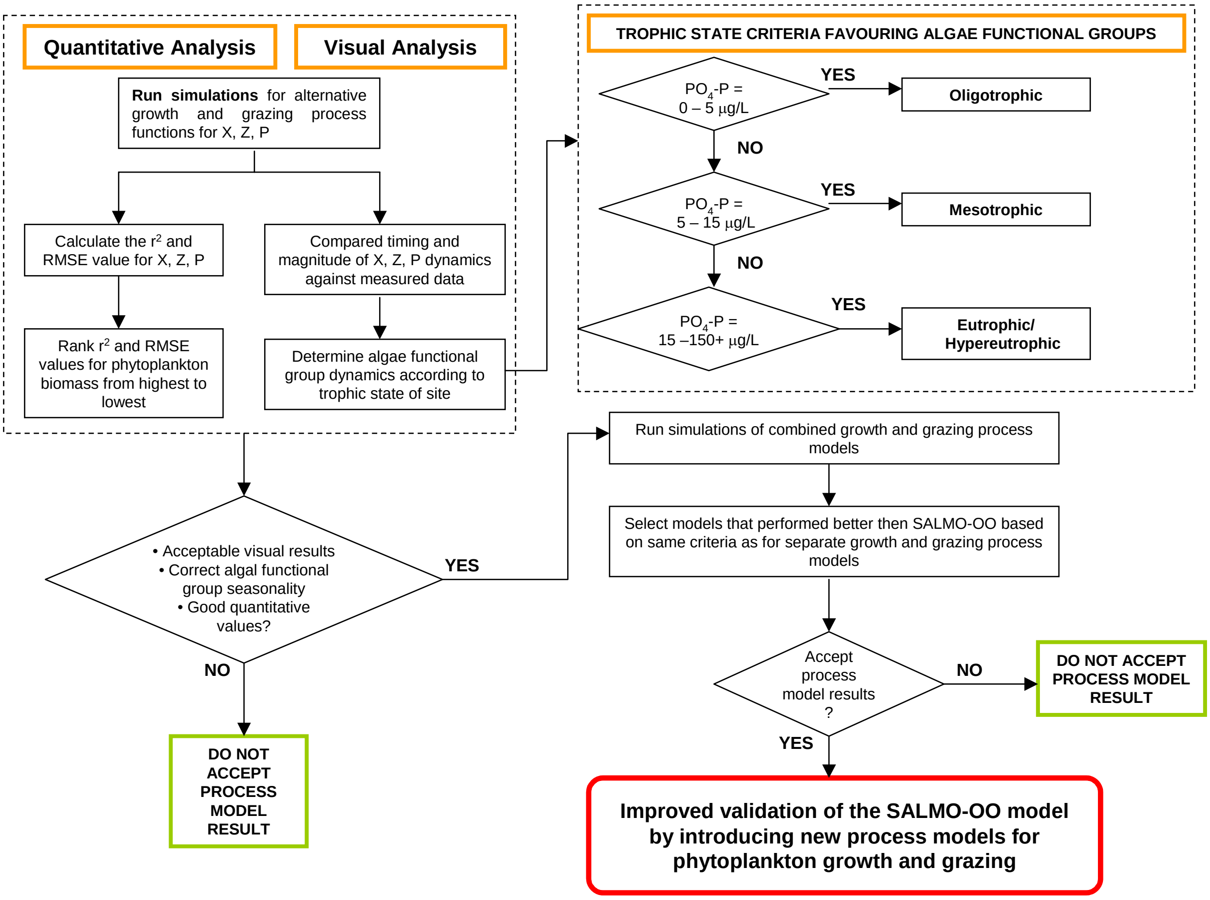

Figure 5.1. Selection criteria and assessment procedures for analyses of alternative growth and/or grazing model structures to improve the validation of the SALMO-OO model.

For example, the phytoplankton simulation results for Bautzen Reservoir given by the combination of the growth model from CLEANER and the grazing model from Hongping & Jianyi (2002) produced an r^2 value of 0.15; a significant improvement compared to the phytoplankton results produced by the original SALMO-OO growth and grazing functions ($r^2 = 0.0013$). However, statistically an r^2 value of 0.15 is a very poor result and indicates that the model has only a 15% chance of successfully explaining the variance between observed and simulated values of phytoplankton biomass in Lake Bautzen, although, when the results are analysed qualitatively the model does describe phytoplankton dynamics very well, particularly compared to the original SALMO-OO results (Figure 5.2). The improvement in the timing of the phytoplankton spring peak predictions is more accurate compared to SALMO-OO and the phytoplankton functional groups are realistic and consistent with what is expected to occur in a eutrophic lake such as Bautzen Reservoir, with a clear distinction in species succession during summer. Thus, for a deterministic model this result for phytoplankton biomass is considered to be an excellent achievement.

In addition, it was found that the RMSE statistic gives a better quantitative assessment, particularly when comparing the results of different model combination. The RMSE for phytoplankton predictions for Bautzen Reservoir is 8.89, which is an improvement upon the phytoplankton RMSE given by the original SALMO-OO growth and grazing functions (RMSE = 10.32). Therefore, it is clear that the combination of the CLEANER growth model and the grazing model from Hongping & Jianyi (2002) reduces the deviance between observed and simulated outputs. Regardless of the statistical results, the visual assessment of simulated outputs against measured data gives the most information about how well the model is performing, in what areas does the model perform poorly and the degree of improvement between the original SALMO-OO model and the results given by the inclusion of alternative growth and grazing process models. The use of r^2 and RMSE assessments alone cannot perform these functions.

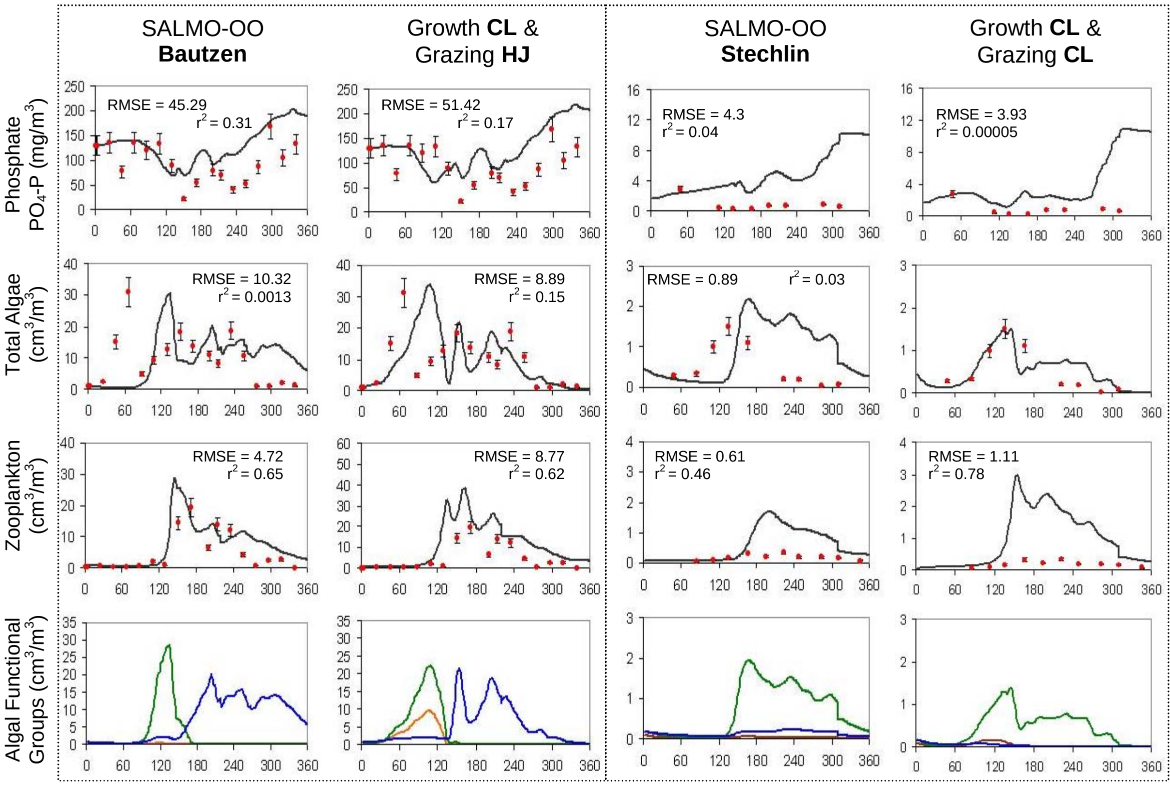

Figure 5.2. SALMO-OO simulation library results for Bautzen Reservoir and Lake Stechlin. Comparisons are made with the simulations from the SALMO-OO model without any changes to the growth or grazing process equations. X-axis is in days. ▬ Blue-green Algae; ▬ Green Algae; ▬ Diatoms; ● Measured data with standard deviation bars of 15%. **HJ** - Hongping and Jianyi (2002); **CL** - CLEANER Model - (Park *et al.*, 1974; Scavia & Park, 1976).

Conversely, r^2 values can give misleading results to suggest that a models prediction are highly accurate, however, when compared qualitatively the results are quite poor. The zooplankton simulation results for Lake Stechlin, given by the substitution of the growth and grazing models from CLEANER, give an r^2 of 0.78, which is an excellent result for a deterministic model (Figure 5.2). However, visually the model simulation shows an over prediction in zooplankton biomass by several-fold. The SALMO-OO results for zooplankton biomass also produce a relatively high r^2 value of 0.46, however, the model also grossly over predicts zooplankton biomass. Another misleading issue is that the r^2 value produced by the CLEANER growth and grazing model combination is significantly better then the r^2 value produced by the original SALMO-OO growth and grazing functions, yet the simulation results show that the alternative model structure produces a larger degree of over prediction. In this case, the use of the RMSE statistic gives a clearer analysis of model performance. The RMSE for zooplankton biomass predictions given by SALMO-OO (RMSE = 0.61) is considerably lower then the RMSE produced by the CLEANER growth and grazing model combination (RMSE = 1.11). This trend is reflected in the visual results, as the zooplankton biomass predictions by SALMO-OO are much closer to the measured data than the simulation given by the alternative model structure (Figure 5.2).

The validation of complex process based models is an important activity to test the models accuracy and to determine areas of improvement. Rykiel (1996) gives a concise definition of model validation *as a demonstration that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.* This definition indicates that the model is acceptable for use, not that it embodies any absolute truth, or even that it is the best model available. Validation is often thought of as simply a comparison of simulated data with data obtained by observation and measurement of the real system. However, such a test cannot demonstrate the logical validity of the model's scientific content (Rykiel, 1996). Nevertheless, the r^2 and RMSE are two popular measures to determine model accuracy, in regards to how well the simulated outputs fit the measured data from the real system. The calculation of both r^2 and the RMSE statistic involves analysing the error associated with the predicted and measured data. Therefore, error estimates allow inferences to be made about the reliability of comparisons between observed and predicted data sets (Huffman, 1997). However, not all sources of error can be identified or accounted for, and will always influence the results of any model away from the true observations that are given by the environment that is being simulated.

Most ecological models are strongly non-linear, and parameter values and environmental data are subject to considerable uncertainty (Huffman, 1997; Murray & Parslow, 1999). There are several sources of error that can give misleading results about the relationships between the model outputs and the system being simulated by the model. These include estimation error for initial conditions, parameter values and input data, which will cause errors to propagate through the model of the system. Similarly, sampling errors occur from collecting field data and those errors associated with laboratory experiments, which will affect the accuracy of parameter values and validation data sets. The occurrence of estimation and sampling errors indicates that the time series sampled for comparison to model outputs is only one possible realisation of the output from the real system (Jansen, 1998; Loehle, 1997). Therefore, parameter and time-series data uncertainties are difficult to eliminate as the error is caused by the natural variation between the system modelled and by the estimation error (Jansen, 1998).

The lack of suitable data not only increases the assumptions within the model but limits the confidence of the modelling results (van Tongeren, 1995). It is often reported in ecological modelling publications that the failure of the model to predict or represent the target site can be greatly attributed to the lack of appropriate data, particularly for model validation testing (Angelini & Petrere Jr., 2000; Drago *et al.*, 2001; Krivtsov *et al.*, 1998; Salençon, 1997). The data sets used to calibrate and validate the SAMO-OO simulation library were mostly collected as part of monitoring programs to assess the changing conditions of the lake's ecosystem. For example, the data collected for the assessment of the South African lakes was conducted by professionals experienced with water quality data collection techniques, and we are confident that the quality of this data is the best that can be currently achieved (Van Ginkel, *pers. com.* 2006). However, Lake Stechlin data was also collected by experienced data collectors, but less data was collected. Therefore, with less data values for a given year to compare to the daily time series outputs given by SALMO-OO, a less confident assessment of the models ability to predict the dynamics of Lake Stechlin occurs. These types of data limitation can cause discrepancies with the statistical comparisons, and requires extensive pre-processing and interpolation to enable the data to be used in modelling experiments. These limitations and uncertainties are often undesirable, but unavoidable.

Resolution or structural errors can occur when significant factors are not included in the model due to cost, size or knowledge restrictions (Huffman, 1997; Jansen, 1998; Loehle, 1997). The effect of structural errors cannot be suitably quantified as the "structurally correct" model is unknown which is often the case with natural systems (Barlund & Tattari, 2001; Jansen, 1998). Some error or bias can occur when data is manipulated for the preparation of different analytical methods, for example, interpolation of missing data points or transformations of data to a more flexible format for statistical analysis. For more detailed discussion on uncertainty analysis see Bacsi & Zemankovics (1995); Barlund & Tattari (2001); Cale *et al.* (1983); Gardener *et al.* (1982); Haag & Kaupenjohann (2001); Hakanson (1999); Jansen (1998); Jorgensen (1994); Marsili-Libelli (1992); Oreskes *et al.* (1994); Parysow & Gertner (1999); Parysow *et al.* (2000); van Tongeren (1995); Wallach & Genard (1998).

There are several methods that may be usefully applied to determine uncertainties and to partition errors in a model. These include sensitivity analysis, uncertainty analysis and structural analysis (Elliott *et al.*, 2000; Mayer & Butler, 1993; Power, 1993). Sensitivity analysis determines the effects of varying parameter values on model outputs and determines how robust the model's structure is (Wallach & Genard, 1998). Those parameters that cause significant changes in the model's behaviour should be estimated with the greatest accuracy. Uncertainty analysis is similar to sensitivity analysis, but takes into account specifically the uncertainty in input and parameter values on the model outputs. Therefore, assuming that the distribution of the inputs and the parameters are known, we can sample from those distributions and generate resulting output variable distributions (Wallach & Genard, 1998). This is often achieved using the Monte Carlo method, which randomly or systematically scans the range of possible parameter values to identify those parameter sets giving an unacceptable simulation result (Barlund & Tattari, 2001).

The SALMO model has been extensively analysed for sources of uncertainty and the errors occurring in the model have been minimised as much as possible as can be achieved for a generic, deterministic model (Benndorf & Recknagel, 1982; Recknagel & Benndorf, 1982). Scenario analysis, sensitivity analysis and parameter optimisation were used to objectively validate the original SALMO model when it was first developed (Recknagel, 1984; Recknagel, 1989; Recknagel & Benndorf, 1982). The SALMO model was validated against four lakes with varying trophic states, from oligotrophic to hypertrophic. Three of the lakes used to test the original model, the reservoirs Bautzen and Saidenbach and Lake Stechlin, were also used to validate the SALMO-OO simulation library. Various scenario analyses, such as biomanipulation and reduction in phosphate loads, were performed to determine the realism and descriptive accuracy of the model. The scenario analyses were able to determine that the SALMO model responded "correctly", therefore, in accordance with what was expected to occur as a result of each scenario analysis for four lakes with different trophic states (Recknagel, 1989).

Nevertheless, there are still sources of error that have had an impact on the development and parameterisation of the SALMO-OO simulation library, and the generation of results for each lake. It is likely that a high degree of error and bias is present in the model, which would cause greater deviances from the goodness-of-fit line, influencing the r^2 values in particular. Even in attempting to limit bias by analysing r^2 for particular seasons corresponding to algal growth, the r^2 values were not significantly improved. Wallach and Genard (1998) acknowledge that there is little that can be done about the uncertainties within model parameters that are notorious in increasing model variance and consequently, affecting the predictive outcome. Deterministic models are conceptualised based on current scientific knowledge about a system and constructed using mathematical functions to describe certain physical, chemical and biological processes. As many of the processes are not completely understood there will always be a certain amount of error and deviation between the mathematical expressions based on our knowledge and what is actually causing certain ecological processes to occur. Therefore, it is difficult to expect a high degree of goodness-of-fit from a deterministic model, compared to other computational modelling techniques such as regression modelling and evolutionary algorithms that develop functions to describe observed data based on patterns in that dataset.

Another method of determining model uncertainty is structural analysis, which is applied to evaluate the importance of various factors by comparing multiple independent models with different formulations (Loehle, 1997). The simulation library provides a similar functionality as the structural analysis concept to the SALMO-OO model. Phytoplankton dynamics are very sensitive to subtle factors that are often not included (or inadequately presented) in a model (Loehle, 1997), usually due to lack of knowledge about certain processes or difficulties in parameterisation. Therefore, to provide additional, flexible functionality to an ecosystem model, in the form of alternative formulations of process that have been traditionally associated with sources of uncertainty and error, gives the user a greater control over the ability of the model to predict phytoplankton dynamics accurately with different lake conditions. Ellner *et al.* (2002) present arguments along similar lines in that the selection of process equations can have a confounding effect upon the goodness-of-fit statistics. For example, Model 1 may fit better then Model 2 because Model 1 makes the right mechanistic assumptions about a processes and Model 2 does not. But it is also possible that Model 2 simply suffers from a poor choice of functional form for a process rate that is not part of Model 1 (type-II instead of type-III functional response etc.). Thus, the choices that a modeller makes on the type of mathematical functions to describe a natural process will cause different responses in the overall model compared to the observed data, and these discrepancies will be reflected in the statistical outputs. These comments from Ellner *et al.* (2002) also support the concept of the SALMO-OO simulation library as a means of additional validation support.

For example, a significant inaccuracy, which occurred with all three German lakes tested by the original SALMO model, was the delay in the prediction of the phytoplankton spring peak. The model was able to simulate the magnitude of the spring peak very well for all lakes, but failed to simulate the appropriate timing and produced peaks that were later compared to the peaks exhibited by the measured data for phytoplankton biomass. These inaccuracies were further explored with the polar coordinate method of sensitivity analysis. The sensitivity analysis showed that the phytoplankton state variable was especially sensitive to the constant parameters for the minimum and maximum photosynthesis rate, respiration rate and the half saturation constant for the uptake of light, particularly during spring. This problem with the original SALMO model has been a cause for further improvements. Applying the phytoplankton simulation library has shown that the timing of the phytoplankton spring peak has improved by using an alternative growth function. Figure 5.2 illustrates the improvements made by the simulation library in the timing of the spring phytoplankton peak for Bautzen Reservoir and Lake Stechlin.

Many ecosystem models within the literature, are validated based on visual, graphical representations between predicted and measured output data and are accepted as successful validations based on visual assessment alone (Collins, 1980; Farnsworth-Lee & Baker, 2000; Recknagel, 1989). However, many other models, particularly regression models, are tested statistically without relying heavily upon visual assessments of model outputs (Asaeda & Van Bon, 1997; Elliott *et al.*, 2000; Mayer & Butler, 1993; Mitsch & Reeder, 1991). Nevertheless, it is also acknowledged in the literature that models should be validated both visually and quantitatively (Elliott *et al.*, 2000; Power, 1993; van Tongeren, 1995). Elliott *et al* (2000) conclude that the use of statistical analysis is equally important in validation assessments as qualitative analysis of the results. However, they believe that this may be misleading in terms of descriptive ecosystem models. They acknowledge that ecosystem models often produce outputs that are slightly out of temporal step with the measured data and that such models would probably be rejected if validation statistics alone were considered (Bacsi & Zemankovics, 1995). However, when they tested several ecosystem models they found, for example, that one validated model produced a good fit through validation statistics, yet failed to predict a spring algal bloom, which was quite significant to the effective management of the system.

One of the main disadvantages of linear regression for model validations is the sensitivity to outliers, which are points that occur far from the fitted line and so produce a large residual (i.e. observed y – predicted $\hat{y}_i$). Outliers may represent erroneous data, or may indicate a poorly fitting regression line. If a point lies far from the other data in the horizontal direction, it is known as an influential observation. The reason for this distinction is that these points may have a significant impact on the slope of the regression line. Therefore, one or two outliers can sometimes seriously skew the results of a least squares regression analysis. The removal of outliers from the data set under analysis can at times dramatically affect the outcomes of the least squares regression analysis. It is not uncommon to remove certain outliers if there is reason to believe that other variables not in the model explain why the outliers are unusual. Alternatively, outliers may suggest that additional explanatory variables need to be brought into the model to explain the occurrence of the outliers if they are important. Apart from removing outliers from the data set, they may be statistically transformed which tends to "pull in" the outliers closer to the other data points near the regression line. Such methods include the square root, logarithmic, and inverse ($y = 1/x$) transforms. However, Krambeck (1995) states that transforming non-linear ecological processes or data to fit linear regression models is a misleading procedure that, however is commonly used by many ecological modellers. By turning a non-linear function into a linear one may give mathematically identical results, but mean nothing in a biological sense. The reason is that ecological data inevitably contains errors and these are transformed too, sometimes creating unexpected results. This problem with linear regression and outliers seems to be a major cause of statistical weakness produced by the SALMO-OO simulation library.

Loehle (1997) also agrees that we cannot solely rely on statistical assessments alone for model validation. He states that using goodness-of-fit measures, such as r^2, to compare model outputs with observed data tends to be problematic for process-based model. This approach can lead to an untruthful indication of inadequate model performance or may lead to calibration of a model against a misleading dataset. However, many ecological modellers seem to rely heavily on a single statistical test to summarise goodness-of-fit and to determine the “correctness” of the model. Thorough testing of a model requires the application of a suite of techniques. No single technique provides adequate information (Loehle, 1997). Loehle (1997) proposes a hypothesis-testing framework for model performance evaluations rather than relying on goodness-of-fit statistical methods. Therefore, the focus is on whether the model differs from reality, not how tightly it fits the data in a regression sense. Precision is properly measured as the width of the bounds on the expected behaviour of the real system. Therefore, when the model is tested for realism, it would be unfair to fault the model for not fitting the expected trend line precisely (as measured by goodness-of-fit statistics) because real data have considerable spread or variations. Rather, we want to know if the model gives results that fall within the bounds we expect for a real ecosystem. Thus, if the model does fall within these bounds then we cannot reasonably say that the model differs in behaviour from that exhibited by the real system.

Statistical validation is often defined as an objective process, however, whether an ecologist trusts the models capabilities of simulating a system is mainly a complex, subjective decision (Botterweg, 1995; Goodrich, 1992). Levins (1966) discusses three basic requirements of a model: generality, realism and precision. Ideally, all three requirements should be optimised or satisfied when developing realistic, robust models. However, this is a difficult processes and usually one requirement is sacrificed to satisfy the remaining requirements. Thus, it becomes a question of which requirement is the most important, and which requirement should be sacrificed in order to develop a valid model suitable for problem solving. Levins (1966) goes on to discuss that dynamic, time series, ecosystem type models usually sacrifice precision in order to achieve realism and generality, these requirements being the most important to achieve. This view suggests that qualitative results are more informative then quantitative results of a model’s outputs. According to Levins (1966), it is more important that the SALMO-OO simulation library results show clear behavioural differences between lake categories and actually be representative of each lake, than precisely predicting each measured value in the data set.

It is necessary to remember what is in fact important to the modeller and the questions asked from the model. Ecologists may not be concerned with the model being able to simulate all the data to the same level of precision and may wish to focus on the patterns observed (Elliott *et al.*, 2000). For example, if eutrophication control is the objective, it is more important to succeed in predicting major algal bloom events rather then predicting base line algal biomass levels (Harris, 1998). A study conducted by Ellison and Bedford (1995) on the simulation of wetland vascular plant communities, discusses how they developed a vascular plant simulation model, using visual and statistical validation techniques. They concluded that statistically many of the state variables did not predict the measured data well ($r^2 < 0.5$) and normally the model would have been rejected. However, visually the results were representative of the system and predicted corresponding trends present in the measured data. Therefore, they accepted the model's structure and functionality based on these qualitative assessments rather than statistical analysis, as the model's results satisfied the purpose of developing the model for wetland restoration.

The results obtained from the study by Ellison and Bedford (1995) present a similar scenario to the results obtained by SALMO-OO. Ellison and Bedford's study indicates that the purpose of developing an ecosystem model is an important consideration that should be incorporated into the validation process (Botterweg, 1995; Goodrich, 1992). The purpose of the SALMO-OO simulation library is to provide additional functionality to improve the models flexibility and performance to lakes with different trophic states and mixing regimes. Therefore, the SALMO-OO simulation library has to be representative of each lake system defined by the categories. The fact that the statistical analyses demonstrate that the SALMO-OO simulation library gave, in some cases, quantitatively poor results does not diminish the model's abilities to represent different lake environments. The visual validation results still adequately suggest that the SALMO-OO simulation library satisfies the requirements of generality and realism, however, at the expense of precision. This is consistent with the model requirements outlined by Levins (1966) when developing deterministic ecosystem models where generality and realism are the main focus and precision is sacrificed. Levins (1966) further agrees that focussing on generality and realism is more important then being precise when developing dynamic, time series models. Goodrich (1992) also agrees that when statistics are invalid then whether the model fulfils the purpose becomes the critical, deciding issue in acceptance of the model as a management tool.

The debate on the use of qualitative versus quantitative validation techniques for testing the predictive success of a model is complex and unresolved in the ecological modelling field. Certain academics strongly favour the use of statistics in testing a model's performance as it is useful to support subjective assessments and gives confidence in the predictive abilities of the model (Bacsi & Zemankovics, 1995; Mayer & Butler, 1993; Power, 1993). However, statistical assessments such as regression analysis and the RMSE statistic are important to consider, but qualitative decisions are unavoidable when assessing the descriptive nature of dynamic ecosystem models (Harris, 1998; Radford & Blackford, 1996). In fact, Botterweg (1995) suggests that complex deterministic models that simulate several linked processes cannot be validated, but more and more evidence can be collected that supports the status of the model as being correct.

5.2. Generic models for the simulation of lake ecosystems

Generic models tend to be those that are designed to simulate more than one type of functionally equivalent system or organism by using the same model structure, but different parameter and input values to define the details of a particular system or organism (Meyer, 1990). Reynolds and Acock (1997) believe a broader definition for ecological modellers is necessary and in addition to Meyer’s (1990) definition propose several design criteria for generic model development:

1. *Transferability:* A good generic design should be suitable for application to ecosystems or a target group by the use of different model parameters or different modules.
2. *Additivity:* A good generic design must be able to simulate functionally similar, yet different systems, by the addition (or subtraction) of modules.
3. *Separability:* Individual modules should be readily recognised by experts in the field as separate processes of the system under study. The purpose of each module should be readily apparent. Modules that combine several functions not normally considered together may be more difficult to parameterise for a new system.

The SALMO lake ecosystem model was designed to be generic for a range of different lake conditions (e.g. different trophic state, morphometry or climate conditions). The model structure (i.e. the mathematical functions) and constant parameter values are kept unchanged for each simulation, and it is the measured environmental input data that distinguishes between each lake. With the implementation of the object-oriented version of SALMO (SALMO-OO) the modularisation of the model’s structure has been achieved in a transparent manner. This has allowed each state variable to be defined as a module, where all relevant functions and specific parameter values are defined. In regards to point 3 of Reynolds and Acock’s (1997) generic model design criteria, any expert in the freshwater modelling field should be able to recognise and understand the mechanisms of each state variable module simulated by SALMO-OO.

The advantage of generic model design is “economy of effort and understanding” (Reynolds & Acock, 1997). Generic models provide an alternative to the development of *ad hoc* models for each specific ecosystem or organism under study, therefore, reducing the need to build and test new models from scratch. As a consequence generic models are appealing to ecological modellers as only one model structure is developed and familiarised with (Grimm *et al.*, 2004; Reynolds & Acock, 1997), thus, less time is spent at the conceptualisation and building stage of model development. Many researchers who have adopted a generic modelling approach find that generic design offers a flexible platform for alternative hypothesis testing (Bussenschutt & Pahl-Wostl, 2000; Pahl-Wostl & Imboden, 1990; Reynolds & Acock, 1997; Zonneveld, 1998). This has been a key advantage in the development of the SALMO-OO simulation library, which explores alternative hypotheses through the use of different phytoplankton process functions that have different premises. SALMO-OO is the core generic model, and the modularisation of the models structure by means of object-oriented design has allowed the inclusion of alternative model structures to be carried out in an efficient manner by minimising the amount of code to be written and tested. For example, the phytoplankton growth functions are designed with a generic skeleton structure, so that each growth method returns the same output variable that is then used to calculate the total phytoplankton biomass (see Appendix C which gives all the source code for the simulation library).

Common parameter values and names to both the alternative process models and those from SALMO-OO were used where possible, in keeping with the overall generic design. Therefore, less calibration of parameter values was necessary, which many ecological modellers will agree is a time consuming and error-prone process (Ford, 1999; Hamilton & Schladow, 1997). Eleven new parameters where included in the SALMO-OO simulation library, in addition to those parameters already accounted for in the original SALMO-OO growth and grazing functions. In keeping with the overall generic design objectives, nine of these simulation library specific parameters were kept constant for the simulation of alternative phytoplankton growth and grazing functions. The one compromise that was necessary in order to use the alternative growth and grazing models to improve model performance was to calibrate the phytoplankton optimal respiration rate (RO) and the zooplankton maximum grazing rate (Gmax), as these parameter values proved to be highly sensitive. As a result, the user must manually calibrate both of these parameters in order to fine-tune the simulation library to achieve the most accurate results possible. This is undesirable and contradictory to the points made above, as the main objective of the SALMO-OO model is to have all the parameter values kept constant. However, if the RO and Gmax parameters were kept constant the simulation library would not significantly improve the models performance from that produced by SALMO-OO's original phytoplankton functions. Therefore, this is an area that needs to be improved and tested more thoroughly, and will be discussed further in the conclusion.

Even though generic models are an efficient and economical means of exploring ecosystem behaviour and for testing hypotheses, there are some disadvantages and limitations in adopting a generic modelling approach. Generally, with site-specific models the calibration of parameter values fine-tunes the model to better fit the field measurements of the system that is being simulated. The validation data set is often an independent data set from a similar system. With generic models the initial calibration of parameter values is a more complex exercise, as the parameter values need to be within a range that is suitable for the simulation of a variety of systems. The SALMO-OO model has constant parameter values that are not changed for the simulation of each different lake, nevertheless the model has the ability to realistically simulate lakes with different trophic states, morphometry and mixing conditions. Therefore, there is no additional calibration and validation needed when the model simulates a new lake dataset. Unfortunately, as a result of the generality of parameter values in generic models there is a higher degree of bias and uncertainties inherent in the model structure, which causes the accuracy of the results to be diminished.

As was discussed in the previous section, Levins (1966) states there are three basic requirements of a model: generality, realism and precision. It is difficult to achieve each requirement in a single model, and often one requirement is sacrificed to satisfy the remaining requirements. Therefore, it is a question of which requirement is the most important, and which should be compromised in order to develop a valid model suitable for problem solving. This really becomes an issue of model purpose and the questions being investigated by the model. In the case of the SALMO-OO model, precision has been compromised in order to achieve realism and generality, as is the case for most generic models. Generic models also tend to compromise more on detailed descriptions so that general applicability is achieved (Grimm *et al.*, 2004). For deterministic ecosystem models it is the desirability to understand and investigate ecosystem dynamics and behaviour that causes such models to be developed, rather than to achieve very high accuracy.

For example, Parrott and Kok (2001) developed an investigative generic model to explore the fundamental properties of complex ecosystem networks. The objective of the model was realism rather than accuracy, in order to create a model for the investigation of plant growth and development that was sufficiently generic to be applicable to a wide variety of plants (i.e. herbs, bushes and trees). As a result, the model may not necessarily provide an accurate representation of any particular plant species. When used to simulate the behaviour of a large number of plants belonging to different species, however, it does provide reasonable predictions of total biomass accumulation in an ecosystem, in addition to depicting the major influences of plants on the soil and atmosphere environments. The model has been intentionally developed in this manner to facilitate investigative engineering research, and as such the objective has not been to accurately describe or predict the actual dynamics of any existing system. Thus, the authors found that it was more important to test each of the fundamental relationships programmed in the model, and that the overall system dynamics exhibited expected results as according to ecosystem theory and observed behaviour. Therefore, these factors hold true for many models that are being used for investigating ecosystem dynamics, and by improving the model in these areas then we can be reasonable confident in the usefulness of the model as a decision-support tool (Bussenschutt & Pahl-Wostl, 2000).

Another point that is emphasised by Parrott and Kok (2001) is that such descriptive models that provide understanding rather than accuracy are used as decision support tools rather than decision-making tools (Bussenschutt & Pahl-Wostl, 2000; Grimm *et al.*, 2004). Due to the restriction in the accuracy of such models they are difficult and unreliable in giving precise estimates of key state variables. To be a decision-making tool, SALMO-OO would have to deliver absolute, reliable predictions of the key variables that are of interest in water quality management, which is very difficult given the uncertainties in model parameters and structure. Therefore SALMO-OO is instead a tool for decision support for lake and reservoir management. Thus, the ultimate goal of the SALMO-OO simulation library is to support the ranking of management scenarios and to base decisions on understanding, not on mere numbers. For example, the SALMO-OO simulation library offers a variety of different scenario analysis options to explore ecosystem management and behaviour. A case study for the three South African lakes that were validated in this study has been investigated using the SALMO-OO scenario analysis for phosphate reduction and biomanipulation of zooplankton population dynamics (Recknagel *et al.*, 2006). The application of the 90% reduction in phosphate loads scenario analysis for Lake Hartbeespoort demonstrates a shift in algal species abundances from a hypertrophic state to a mesotrophic state with the dominance of green algae instead of toxic blue-green algae. Similar results were achieved for Lakes Roodeplaat and Klipvoor, although these lakes would remain eutrophic due to the extreme algal blooms occurring in summer. However, the biomanipulation scenario analysis demonstrated that control of blue-green algal blooms was more difficult to achieve when combined with P-load reduction, as this management option reinstated blue-green algal dominance in summer by targeting mainly green algae as a result of stronger grazing by zooplankton. The outcomes of the South African lakes scenario analysis can assist decision makers in tailoring the specific management option to apply to the real systems and to rule out which management option would be less successful, thus supporting which decision should be made.

5.2.1. The benefits of the object-oriented paradigm for generic model development and implementation

Generic models are designed to encompass as many specific situations as possible. The task of the user is to parameterise the generic model and to tailor it to some degree by choosing from alternative modules (Grimm *et al.*, 2004). The SALMO-OO simulation library has been redeveloped from the original FORTRAN program version into an object-oriented software package, with a graphical user interface that is much more user-friendly and informative. The SALMO-OO simulation library has been designed to act as a decision support tool for lake and reservoir management. Thus, the library can be used to improve model validation as best as possible from a knowledge-based approach by selecting alternative mathematical functions. Scenario analysis can then be performed to determine the effect of particular management strategies or environmental degradation on the lake being investigated. The SALMO-OO model is tailored by the inclusion of site-specific measured data for driving variables such as water temperature, solar radiation and water inflow and outflow.

The benefit of the object-oriented paradigm in ecological modelling has been extensively discussed in the Literature Review chapter. The main advantages of the object-oriented paradigm for ecological modelling are the logical modularisation of model components, which provides a more flexible platform for model development, and the suitability of the paradigm to simulate the behaviour of natural entities more realistically due to the use of objects. The object-oriented paradigm is naturally suited as a mode of developing generic models. Reynolds and Acock (1997) present the concepts of modularity and "genericness" as methods of good model design. Adopting the object-oriented paradigm for model development facilitates the modular design component, as the concept of classes and inheritance mechanisms naturally categorises ecological processes or state variables into logical, hierarchical modules. For example, all phytoplankton process functions, specific parameter values and information are encoded into a single class that is only accessed by other objects through special message protocols. This effectively de-clutters the model structure as the objects and processes need only be defined once in the program. Not only is the model easier to navigate around and learn, it reduces the propagation and time needed to fix errors and facilitates the understanding of model logic (Acock & Reddy, 1997; Dawson & Swatman, 1999; Lemmon & Chuk, 1997).

The generic design component allows several types of model formulations with different output variables or assumptions to be investigated within a single model structure, and can often be applied to a variety of conditions that in a traditional setting would require a different model for each case study. The object-oriented paradigm facilitates the development of generic models as the alternative processes functions or modules can be easily added to the models structure, particularly due to the object-oriented concept of inheritance. Therefore, classes (or modules) within a model can be arranged in a hierarchical structure, which inherit properties (such as parameter values or process functions) from the proceeding classes. Thus, only those particular features that have changed within the context of the model need to be added to the new descendent classes (Silvert, 1993). For example, the `AlgaeLibrary` class, which encodes all relevant information about the phytoplankton simulation library, is a descendent of the `Phytoplankton` class, which descends from the `Salmo` super class. The `AlgaeLibrary` class has access to all functions and parameter values that are defined by the `Salmo` and `Phytoplankton` classes. Thus, when defining alternative process functions for phytoplankton only new equations and parameter values need to be defined, rather then having to repeat code that has already been written for previous classes. Inheritance provides a powerful mechanism for organising and structuring simulation models and allows the reuse of a class's behaviour in the definition of new classes (Baskent *et al.*, 2001; Sequeira *et al.*, 1997).

Applying the object-oriented paradigm to a lake ecosystem model was no easy task, considering a new program language had to be learnt and also initially dealing with the abstract concepts of the object-oriented paradigm took some time and effort. Often the luxury of learning new programming languages and design techniques is not readily available to many ecological modellers, which is possibly why the object-oriented modelling approach and the use of the Java program language is still not wide spread, even though the advantages of such an approach are beneficial. Once the hierarchical structure for the SALMO-OO model was completed, designing the simulation library was very straightforward. The methods for the alternative phytoplankton process models were designed based on the original phytoplankton process functions. Common parameter values that are accessed by various objects were encoded into one super class, therefore, this made these parameters easy to find and update. When it was necessary to include additional process functions to the SALMO-OO simulation library, the object-oriented structure of the model facilitated this activity.

However, the object-oriented paradigm can only benefit model development to a certain extent. Thus, good model design begins with the conceptualisation of the model domain and assumptions, how the model will mathematically represent key processes, and what level of spatial and temporal detail is necessary to give meaningful results. These considerations are greatly dependent on the questions being asked of the model, and such questions will also determine the modelling approach to a certain degree. The main area of use for the SALMO-OO simulation library is as a decision support tool for determining management outcomes in freshwater lakes and reservoirs. SALMO-OO is a descriptive, investigative model that gives the user understanding in the behaviours of lake ecosystems. By providing the simulation library as an additional validation toolbox this has improved the overall model performance to give more accurate and realistic results for phytoplankton dynamics.

5.3. Conclusions

To draw the final conclusions from this study let us revisit the aims outlined in the Introduction chapter:

1. Does the application of a phytoplankton growth and grazing simulation library improve the applicability and accuracy of SALMO-OO?

As a generic model SALMO-OO, in its original format, simulates a wide variety of lake conditions and scenarios. With the application of the simulation library of phytoplankton growth and grazing process models, the accuracy of the model to predict key state variables has been improved, not only in the quantitative results, but also in the prediction of key events, such as algal blooms that are important to environmental managers to forecast. Each phytoplankton process model included in the simulation library is based on strong and rigorous scientific principles that have been in use for many years in many different lake models. As a result, each process model has been extensively validated, and shown to perform realistically and accurately within the models from which they have been taken. The main goal of improving the SALMO-OO validation results was based on achieving a more general, realistic and accurate simulation of phytoplankton functional group dynamics. The original SALMO model only simulated two phytoplankton functional groups for a given simulation. The SALMO-OO simulation library is able to simulate three phytoplankton functional groups realistically for different trophic states. Although many of the lake data sets that were tested did not have measured data for phytoplankton functional groups the few lake data sets that we did have data for indicated that the model performed well in the simulation of functional group dynamics and succession through the seasons.

Although process based ecosystem models have been criticised in the past as being useless in giving confident predictions for management scenarios, they benefit the decision-making process by providing a platform where ecosystem structure and behaviour can be explored. Process-based models can also give us an understanding in the gaps that exist in our knowledge of a system. By increasing our knowledge base and our abilities to measure ecological data more accurately these advances will be reflected in the deterministic models that are developed as decision support tools. The SALMO-OO simulation library offers a greater choice of modelling options by providing alternative process model structures in a single user friendly framework.

2. Can generic model structures be found using the SALMO-OO simulation library for lakes with different trophic states, climate conditions or morphometry?

Generic model structures were found for nine lakes based on four categories of trophic state and mixing conditions. Each model structure that was found to best simulate a particular category was not necessarily the best performing structure for an individual lake dataset, but performed best overall for all lakes in each category. The reason for distinguishing lake categories was to improve the models generality and realism, and to simplify the use of the model for environmental management scenarios. Determining a particular model structure from the possible combinations of phytoplankton growth and grazing models was a rigorous and time consuming process. Therefore, by suggesting a particular structure for a type of lake in a simplified category the user would only need to fine-tune the model by calibration of the phytoplankton respiration rate and the maximum grazing rate.

The manual calibration of the phytoplankton respiration rate (RO) and the maximum zooplankton grazing rate (Gmax) is a step in the modelling process that needs to be eliminated. The main concept of SALMO-OO is to be used as a generic model where constant parameter values are kept unchanged for each simulation of each different lake, distinguished only by the environmental input data. However, a drawback of generic models is the sacrifice of precision in order to achieve realism and generality. Therefore, to improve the precision of the SALMO-OO model through the simulation library some calibration was deemed necessary. However, further research in this area is currently taking place, where evolutionary algorithms are being used to replace sensitive constant parameters, such as the maximum photosynthesis rate, the maximum zooplankton grazing rate and the plankton respiration rate, by either automation of optimising parameter values or by replacing the constant parameters with functions evolved from the input data (Cao *et al.*, in press). Preliminary results are very encouraging and have increased the models accuracy to levels acceptable for SALMO-OO to be used as a decision making tool. The parameter optimisation and function evolution modelling steps will be offered as a further validation step in the user interface within the SALMO-OO simulation library software package. Also, the model would be more realistic if the values for RO and Gmax were different for each plankton functional group, as these values in nature would be different for each phytoplankton species. This generalisation was performed in order to achieve some simplicity in the development and testing of the model library. Automation of the calibration process for these variables would solve this problem. Nevertheless, the inclusion of hybrid evolutionary algorithms to find functions based on some environmental variable would reduce the need to calibrate these parameters and functions could be discovered based on each plankton functional group, rather than using a common variable for all phytoplankton species.

The SALMO-OO simulation library provides additional functionality in the form of alternative model structures to improve the models flexibility and performance to lakes with different trophic states and mixing regimes. The simulation library is currently being expanded to include alternative process for zooplankton growth and mortality and a sediment library. The object-oriented paradigm for model design and implementation will allow these additions to SALMO-OO to be completed in a straightforward flexible manner. Other state variables that are being considered for inclusion into SALMO-OO are ODEs for silica, dissolved inorganic carbon and additional phytoplankton and zooplankton functional groups. The adoption of the object-oriented paradigm and the construction of the simulation library as a decision support tool will provide a modelling environmental which is user friendly, transparent and adaptable to a wide variety of lake system.