WP 2

Deliverable Report 2

Protocols and Benchmarks for Maximizing the Performance of Water Quality Seasonal Hindcasts

March 2018

Authors: Don Pierson, M. D. Frías, Sixto Herrera, Maialen Iturbide, Leah Jackson-Blake, Andrew French, Daniel Mercado, Tadhg Moore, Marc Scheibel, Eleni Teneketzi, and Leon Van Der Liden

Table of Contents

Introduction

An overarching goal of WATExR is to produce seasonal forecasts of lake and reservoir water quality. To this aim, both the climate seasonal forecasts and the lake/watershed models should be properly linked and the resulting workflow calibrated in order to obtain the expected skill and uncertainties of the resulting seasonal forecast system. This deliverable tries to summarize the procedures proposed within the WATExR project to make this calibration – verification, focusing in the different pieces/elements of the defined workflow:

- First, lake and watershed models should be calibrated by comparing simulated estimations, forced with measured inputs (meteorological data, stream discharge etc.), with measured data during an appropriate calibration period. Calibrated models should then be verified by comparing model predictions to observations for an independent historic period.

- Second, once models have been calibrated and verified, the retrospective climate seasonal forecasts, known as hindcasts, should be used as inputs to the models in order to obtain the equivalent hindcast from the lake and watershed models. To this end, these models should be run in the same way as the climate models that produce seasonal forecasts, that is, they should be initialized at every initialization month considered by the climate forecast system used as input and produce forecasts for the next few months. At the same time, the probabilistic nature of the seasonal forecast should be also considered, by running these models for each member of the ensemble obtained from the climate seasonal forecast system.

- Finally, the skill and uncertainties related with the resulting seasonal forecasts should be evaluated by comparing the obtained hindcast and measured data for a common period, if possible covering at least 30 years.

For descriptive purposes, this deliverable presents some examples of the calibration – verification workflows and model output associated with these is also provided for several sites and models within the WATExR project. The models being used in this study (Table 1) are all ultimately concerned with predicting the water quality or ecology of the lakes and reservoirs participating as case studies in this project. In most study sites, these models can be configured as a chain of watershed and lake/reservoir models, which in turn are often further divided, for purposes of testing and calibration, into their physical and biogeochemical components (Fig 1).

Table 1 Different models being used by the WATExR consortium, which include both process-based and empirical models, and lake and watershed models of differing levels of complexity. Model calibration involves adjusting multiple model parameters to the optimal values that provide the best correspondence between model output and measured values. The methods (computer algorithms/programs) used to produce model calibration are listed. Countries involved are Sweden(SE) Ireland (IE) Denmark (DK) Norway(NO), Spain (ES), Germany (DE) and Australia (AU).

Process-based Lake Models			
Model	Sites used	Calibration Method	Notes
GOTM	SE IE DK NO ES AU	Optimization (ACPy)	General Ocean Turbulence Model (Burchard 2002) Hydrodynamic - coupled with FABM
SELMA	SE IE NO AU	ACPy	Simple Ecological Lake Model Water quality coupled with FABM Developed by PROGNOS
PC Lake	DK ES	ACPy	Lake Water quality model (Janse 1997) coupled with FABM
DOMcast	IE NO	ACPy	DOC specific water quality coupled with FABM Developed by PROGNOS
CE Qual W2	DE		2D hydrodynamic and water quality (Cole and Wells 2002)
GLM	DE		1D hydrodynamic and water quality lake model (Hipsey et al. 2019) FABM like coupling
MyLake	NO		1D hydrodynamic and water quality.lake model (Saloranta and Andersen 2007)

Process-based Watershed Models			
SWAT	SE DK ES	Swat CUP	Soil Water Assessment Tool (Arnold and Fohrer 2005) Semi distributed and saleable
INCA	IE	Custom	Integrated Catchment model Semi distributed and scalable (Jackson-Blake et al. 2016)
SimplyP	NO	MCMC (emcee)	Simplified catchment hydrology and phosphorus model, semi-distributed
mHM	ES DE	Shuffled Complex Evolution	The mesoscale Hydrologic Model Fully distributed
Talsim NG	DE	Manual	Talsim Next Generation (SYDRO Consult GmbH); catchment model that combines hydrological modelling and reservoir operational rules
Empirical models			
Eel migration	IE		
Bayesian Belief Network to simulate lake variables	NO	Set up using observed and modelled catchment and lake data. Alternative Bayesian method used to predict lake variables of interest.	

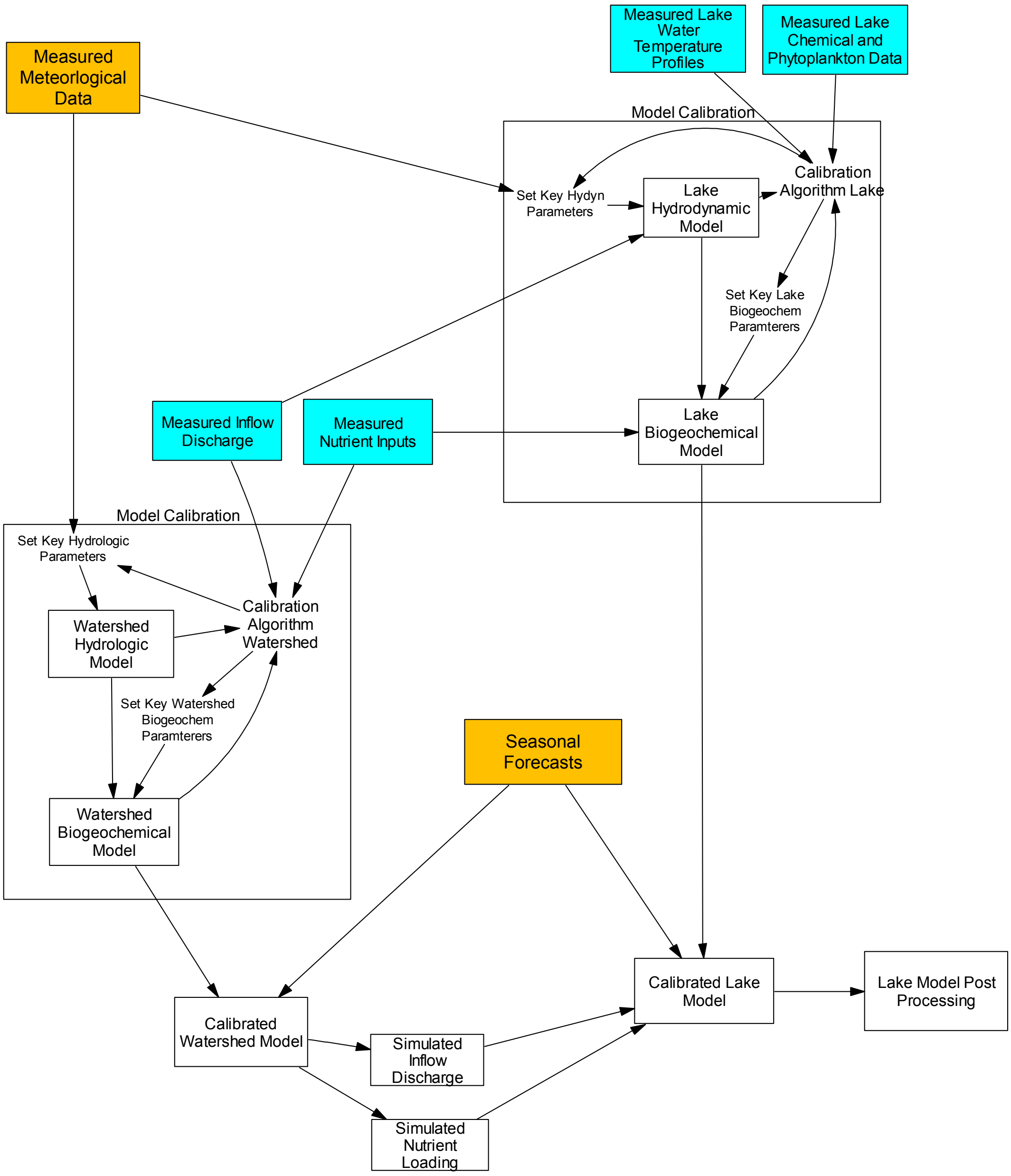

Figure 1 Illustration of the coupling of a generic watershed and lake/reservoir model, and the way in which measured data is used in model calibration. The output of the watershed model becomes one source of input to the lake/reservoir model. Both classes of models may be sub-divided into physical and biogeochemical components, since these are often calibrated separately, using data from different sources. Orange boxes represent the climate data that are used to force the models. In the top observed meteorological data are used during calibration and in the bottom seasonal climate data are used to make retrospective predictions. Blue boxes represent the measured data that are needed to judge model performance during the calibration process.

Sources of Uncertainty

In regards to the model forecasts that WATExR will produce, there are several sources of uncertainty that can in turn affect the range of uncertainty of seasonal predictions.

- Uncertainties of the seasonal forecast at location of the case study
- Appropriateness of model structure
- Uncertainty in the model parameterization
- Uncertainty in the model state at the onset of the simulation forecast

Seasonal forecasts are of probabilistic nature (see e.g. deliverable 2.2) as provided by an ensemble of perturbed forecasts. As a result, each member of the ensemble produced by the climate seasonal forecast system should be used as input of the lake and watershed models (Fig 1) obtaining an ensemble to which the same probabilistic approach should be applied. Uncertainty in the model forcing therefore, should be combined with other impact model uncertainties.

The goal of the modeler is to constrain the range of uncertainty to as great an extent as possible, thereby increasing the value of the forecast. For a forecast to be valuable it must predict a range of conditions that would support a change in water management, reservoir operations, or public outreach. This range cannot be so broad that it essentially encompasses the historical range of seasonal conditions, rather it must encompass a range of conditions that is unique sub-set of the seasonal range. To reduce uncertainty in the simulated results the focus must be on the remaining three factors that affect model parameterization and the initial state of a model. lThese factors are the subject of this deliverable report. The goal here is to document the modeling protocols that can be used to make models as close to the ideal perfect model as possible

All models are a necessary simplification of the world, and all models will put different emphasis on the way they choose to describe the world. Model structure is concerned with first how model developers decided to describe the world and secondly the equations and relationships used to describe the processes included in the model. WATExR places and emphasis on developing usable tools based on previously developed models, so the effects of model structure on simulated output are largely controlled by choosing the most appropriate previously developed model (i.e. SELMA vs PC Lake Table 1), rather than modifying the model itself. However it is worth remembering that sometimes it is not just calibration, but the underlying model structure that needs to be adjusted in order to improve model predictions. One case study site (Norway) will explore model structural uncertainty in WATExR, and will choose the model with lowest structural uncertainty for use in forecasting.

Model Calibration

Given WATExR reliance on previously developed models, it is the process of model calibration that is most important for improving model performance in order to produce the most realistic forecasts. And it is the use of hindcast simulations that allow the effects of model calibration to be quantified. Model calibration involves several steps.

1) Identifying the model state variables that can be used for model calibration. These are ones that can be compared to actual measurements of the state variable or a proxy measurement of the state variable. For example simulations of stream discharge can be compared to actual measurements of stream discharge, while simulations of algal chlorophyll concentration can be compared to actual measurements of extracted chlorophyll or in situ measurements of chlorophyll fluorescence – a proxy measurement. Calibration is therefore dependent on the availability of measured data that serve as a comparison to the simulated output of comparable values. Since the goal of the calibration is to match the measured values as closely as possible there can be an implicit belief that the measured data are correct, which is of course not the case, as measurements also have their own sources of uncertainty. This uncertainty in the measured data can be taken into account in certain automated calibration routines (e.g. many Bayesian methods).

2) Ideally, all uncertain model parameters would be included in model calibration. However, frequently models tend to have too many parameters for this to be done meaningfully. Instead, only a selection can be calibrated. Identifying the model parameters that affect model output of key model state

variables, can at times be a trivial task that is obvious from an understanding of the model structure and equations. At other times it is not as obvious, and it can be necessary to undertake model sensitivity analysis (varying parameters while observing changes in the simulated state variable) in order to identify the most important parameters to optimize. It is also important to note that models can be separately optimized in regards to different simulated state variables using an iterative model workflow (Fig 2), or they can be optimized in regards to all state variables of interest in a single step.

3) Adjustment of model parameters to obtain the best fit between simulated and measured state variables. It is possible to manually and systematically vary model parameters, and to observe the relationship between simulated and measured model state variables. This can be done by both visual examination of the variable time series, or by comparison of metrics of model fit. By repeating this process many times an optimal set of model parameters can be chosen. Manual calibrations can provide greater insight into model behavior, but also can be time consuming and be somewhat subjective. Furthermore, when moving beyond simple models to more complex water quality models many parameters may need to be adjusted and multiple state variables will need to be compared to measured data. In such cases model calibration can require 1000s of simulations and automated methods are preferred. There are several different automated calibration methods used in the WATExR project (Table 1). Many are optimization algorithms, which automatically vary the model parameters, run the model, and calculates a metric of model fit, and then repeat this process until the metric of fit can no longer be improved. The methods differ most in the algorithms used to search for the optimal parameter set. Typical many thousand simulations will be run to obtain the final calibration of a water quality model. The two most commonly used auto-calibration programs in WATExR were ACPy which is described at (http://bolding-bruggeman.com/portfolio/acpy/) and the SWAT watershed model calibration program SWAT CUP (Abbaspour 2012). It is often common to perform the parameter optimization as set of sequential steps i.e. first calibrating the physical model structure and later the biogeochemical

components of the model (Fig 1). A Bayesian approach to model calibration is used in the Norwegian site, where an MCMC algorithm is used to not just find an optimum parameter set, but map out the probability density over the whole posterior parameter space. Results from this kind of algorithm will then be used to estimate parameter-related uncertainty in model simulations. Example workflows for the different WATExR sites are given as examples below.

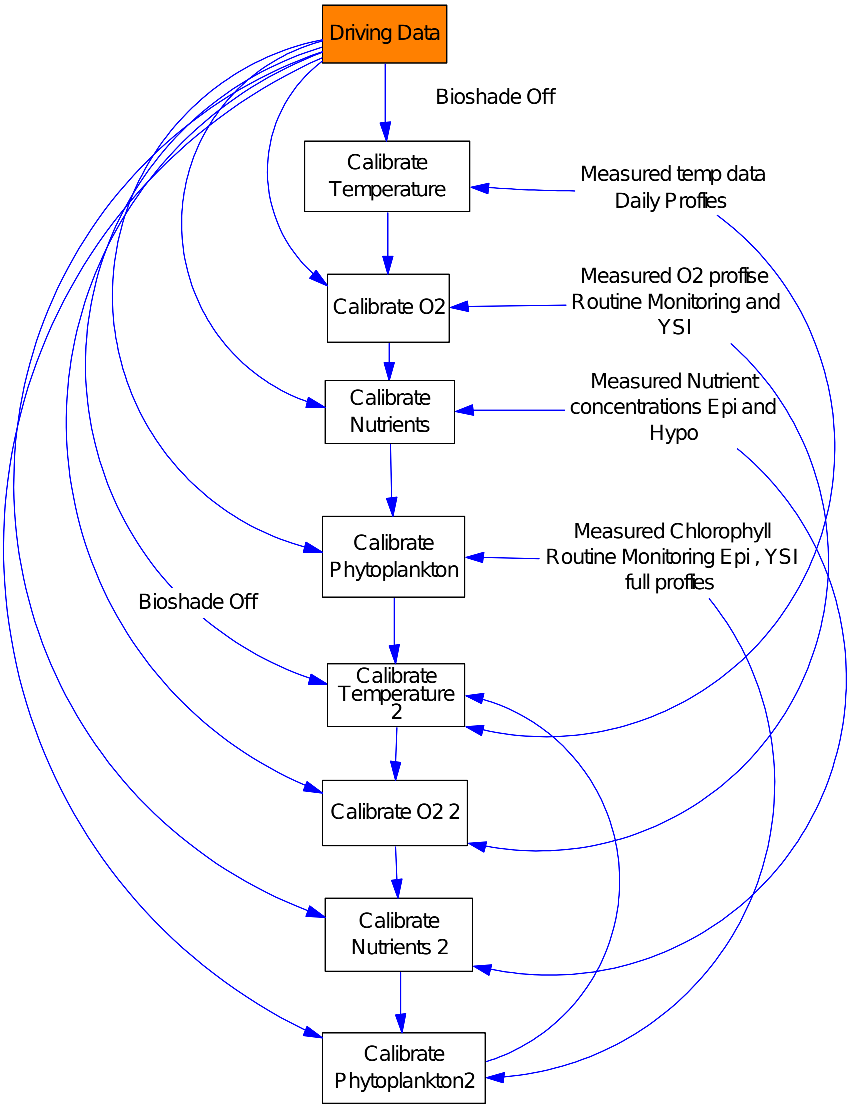

Figure 2 Example of the sequential steps of model calibration used to calibrate the GOTM – SELMA model at Lake Erken. The lake hydrothermal model is first calibrated by comparison with measured lake water temperature data. Following this the biogeochemical model parameters affecting lake dissolved oxygen concentration, lake nutrient concentration and lake chlorophyll concentrations are adjusted. Following each step the range over which the model parameters can be adjusted for that step are restricted to +/- 10% of the optimal values. Following the first set of calibrations the process is iteratively repeated after turning on feedback between the simulated water quality and the light extinction coefficient.

Model Verification under Observed Conditions.

The purpose of the verification simulation is to provide a test of model performance outside of the period used to perform the model calibration as described above. As with the calibration, for the verification it is necessary to have both model forcing data and measured values of model state values that can be compared to those simulated by the model. For the verification the model parameters are fixed to the calibrated values. Models are driven by observed meteorological data for a historic period independent of that used to calibrate the data. The model is run and simulated state values are compared to measured values. This provides an indication of the skill of the watershed/lake models, before introducing uncertainty associated with the seasonal model forecasts.

There are several measures of model fit that are commonly used to judge the fit between simulated and observed values of the model state variables. Ones commonly used in WATExR are the root mean square error (RMSE), the mean absolute error (MAE), the Pearson correlation coefficient (r), the Nash Sutcliff Efficiency (NS).

$$RMSE = \sqrt{\sum \frac{(-obs)^2}{N}}$$

$$MAE = \sum \frac{|-obs|}{N}$$

$$r = \frac{\sum (-\acute{\square})(obs - \acute{obs})}{\sqrt{\sum (-\acute{\square})^2 \sum (obs - \acute{obs})^2}}$$

$$NS = 1 - \sum \square$$

These are calculated for both the calibration and verification simulations. The expectation is that the values of these metrics of model fit will be similar in both the calibration period and the verification period. When this is the case it indicates that the model has been calibrated to the general conditions of the study site and its climate, and that the calibrated parameters can be used to represent the response of the study site to a different set of seasonal weather or future climate conditions. This also implicitly assumes a stationary climate- which is not the case, but which may be a reasonably approximation over the decade or less time periods used for calibration and verification, and also over the seasonal time scales used for WATExR forecasts.

Impact models driven by seasonal climate forecasts

Following model calibration and validation to historical data the next step of model validation used in the WATExR project is to run the calibrated watershed and lake models driven by the retrospective seasonal climate forecasts.. As a result this will produce a hindcast of seasonal forecasts for the variables of interest for each case study. Since in the hindcasting mode, we are working with archived previous forecasts, the forecasts can then be compared to actual measured values of the model state variables in order to judge the predictive power of the seasonal forecasts. A second level of model validation can be made by forcing the watershed and lake models with measured meteorological data during the hindcasting period (e.g. the EWEMBI dataset) and then comparing simulated state variables based on simulations using measured meteorological data with the simulated state variables based on simulations forced with the seasonal climate forecasts. Comparison of these two hindcasting methods will provide an estimate of the uncertainty in the forecast simulations related to the model uncertainty and the forecast uncertainty. Forecast uncertainty is further quantified by evaluating the variability between simulations run with 15 different plausible meteorological time series ie the ensemble member of the seasonal forecast.

An example of the seasonal hindcasting method is shown in Fig 3.

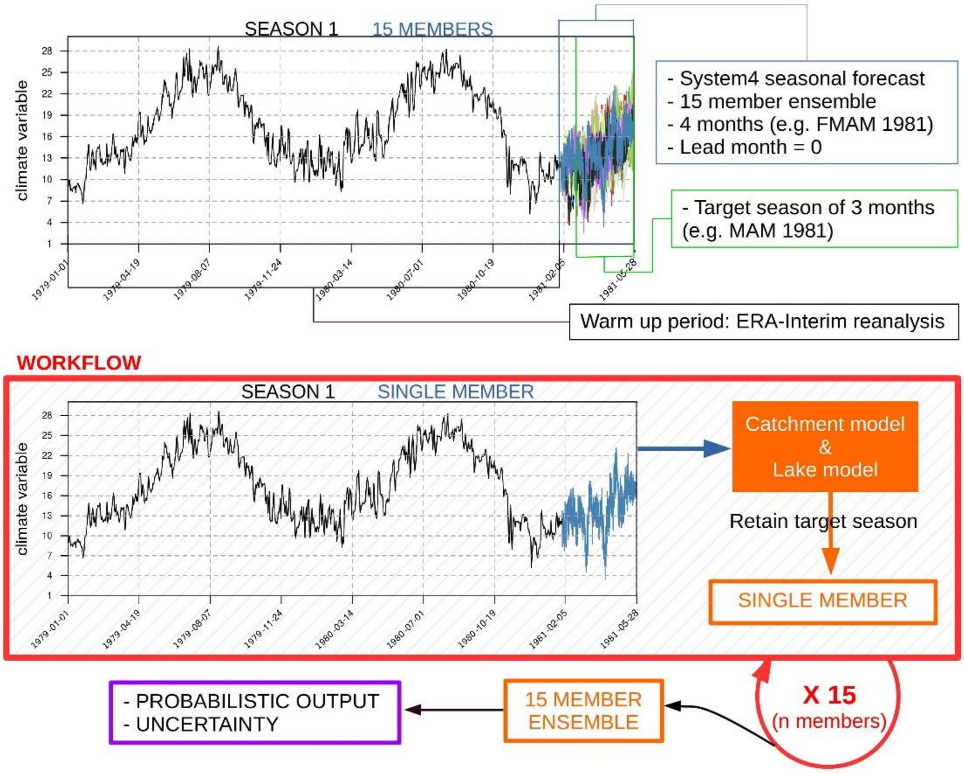

Figure 3 Overview of seasonal hindcasting method. Bottom shows workflow for running each ensemble member of the seasonal forecast system. This is accomplished by first spinning up the impact model using reanalysis data, and then running the individual ensemble forecast. The spread of the seasonal forecast ensemble is shown in the upper portion of the figure.

Examples of Calibration and Hindcast Simulations

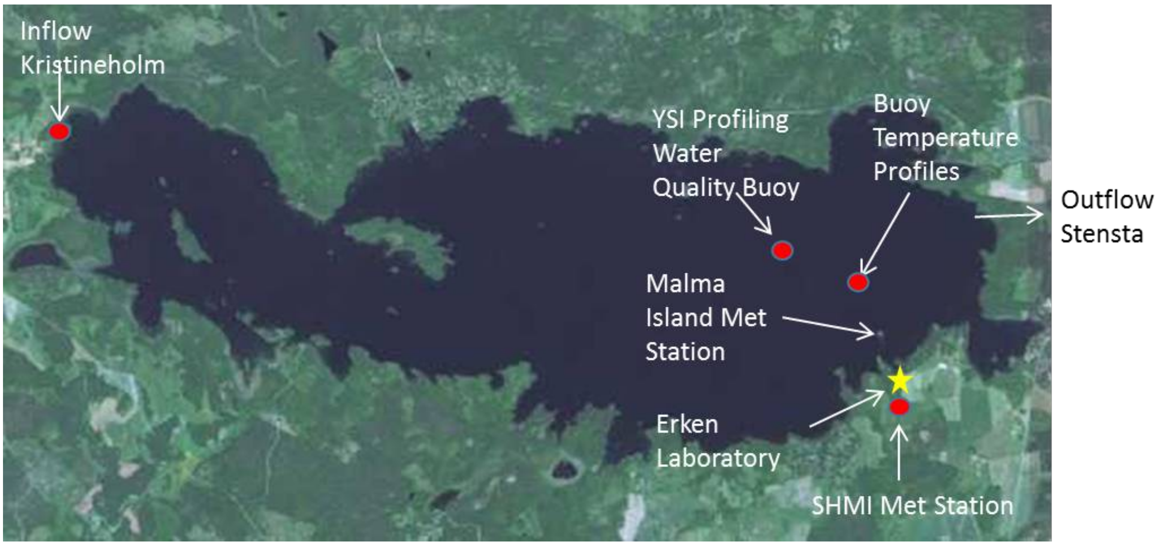

Figure 4 Lake Erken showing the sites where model forcing data and calibration data are collected.

Lake Erken – Hydrothermal and dissolved oxygen model calibration

Lake Erken is a moderately eutrophic lake located in east-central Sweden near the Baltic coast (59.8 N 18.6 E). The lake has a surface area of 24 km^2, an average depth of 9 m, a maximum depth of 21 m and a water residence time of approximately 7 years. The results presented here are of the calibrations of the GOTM hydrothermal model and the coupled SELMA biogeochemical model based on a stepwise calibration scheme (Fig 2) using the ACPy auto-calibration program (Table 1). These calibrations correspond to the first two steps in the calibration workflow shown in Figure 2.

Both temperature and oxygen calibrations were made using data collected between 2006 and 2016. Most meteorological data needed to drive the model have been collected by an automated data logger system placed on Malma Island ~500 m offshore from the Erken laboratory (Fig 4). This near lake location ensures that the data collected will be representative to the climate directly impacting the lake. Starting in 2006 the Swedish Metrologic and Hydrologic Institute (SMHI) installed an automated weather station near the Erken laboratory. These data can be used to supplement those collected on Malma Island. Also, since 2006 measurements of the inflow and outflow discharges have also been collected by automated systems. These data allow a lake water balance to be constructed, and therefore allow for the most realistic GOTM simulations, where lake level varies seasonally, and where nutrient loads can accurately be estimated. Measurements of model state variables used to calibrate the model were made using a combination of automated and manual water quality measurements. Water temperature profiles were collected every 30 min during the ice free periods using a buoy based string of 30 thermocouples spaced at 0.5 m intervals. Oxygen data was measured at approximately 2 week intervals as part of the Erken Laboratory routine monitoring program. These data were merged with more frequent hourly oxygen profiles collected during 2015-2016 by a YSI profiling system.

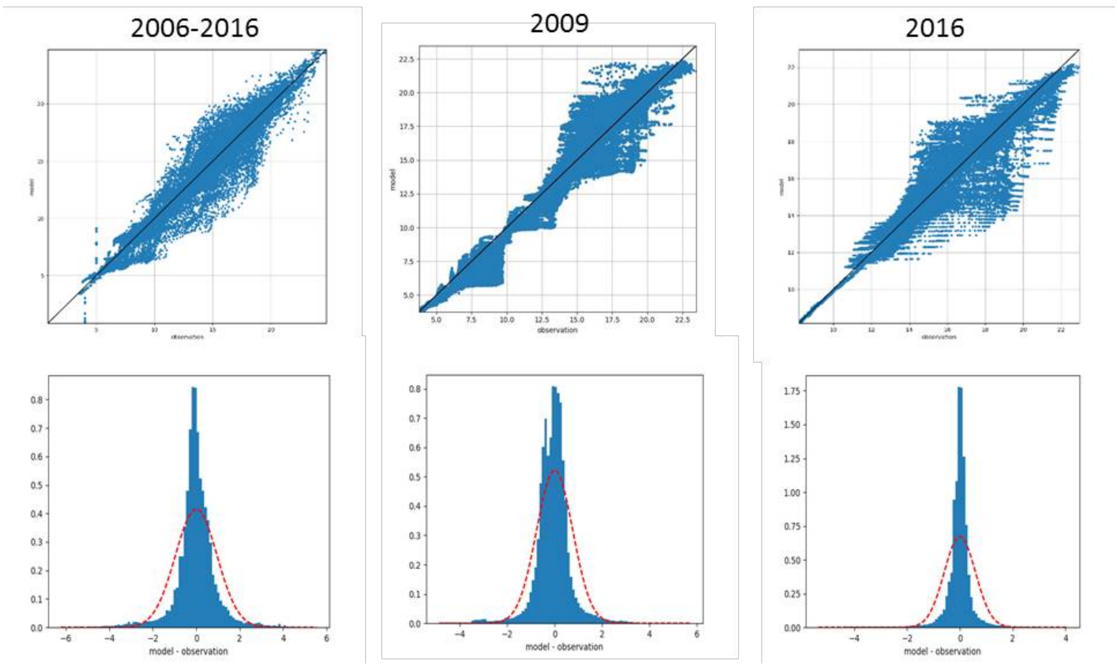

Figure 5. Top plots of the measured (x axis) vs. simulated temperature (y axis) for all points simulated during the entire calibration period, and during two specific years 2009 and 2016. Bottom histograms of the difference between the simulated and measured water temperatures.

For water temperature a total 6 parameters were calibrated. These included parameters affecting wind induced mixing, surface water heat exchange, and the absorption of heat by the water. In general we found that approximately 4000 model iterations lead to stable solution that gave consistent sets of optimized model parameters. Results of the calibrations are shown in Figures 5-6.

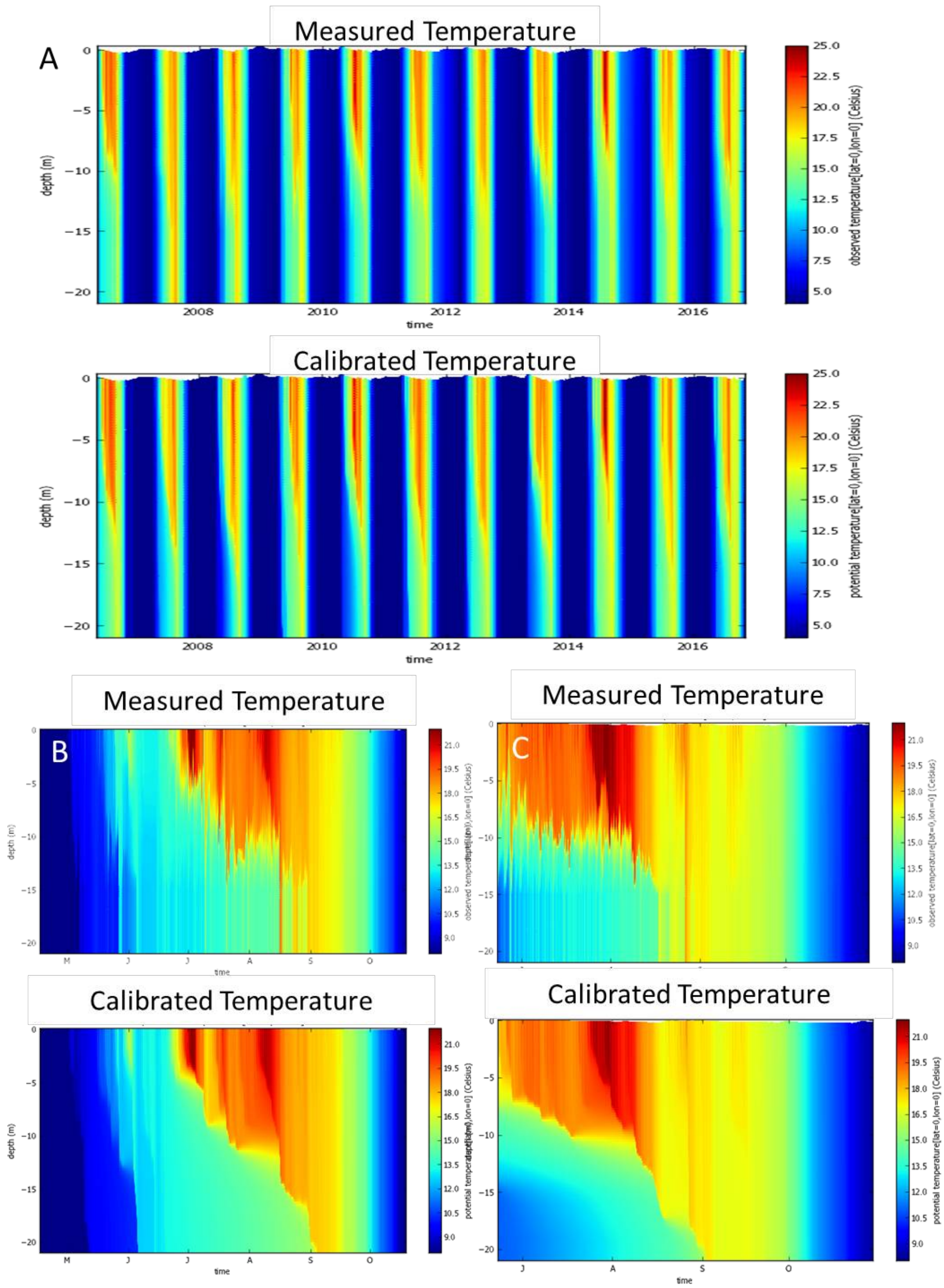

Figure 6. Comparison of measured and simulated water temperatures with temperature data plotted as heat maps. A) 2006-2016 daily data time step. B and C) more detailed plots of data from 2009 and 2016.

For biogeochemical simulations of dissolved oxygen an additional nine parameters were calibrated. These included parameters that affect the sinking, burial, erosion, and decomposition of detrital organic matter as well as the exchange of oxygen between the atmosphere and the upper water column.

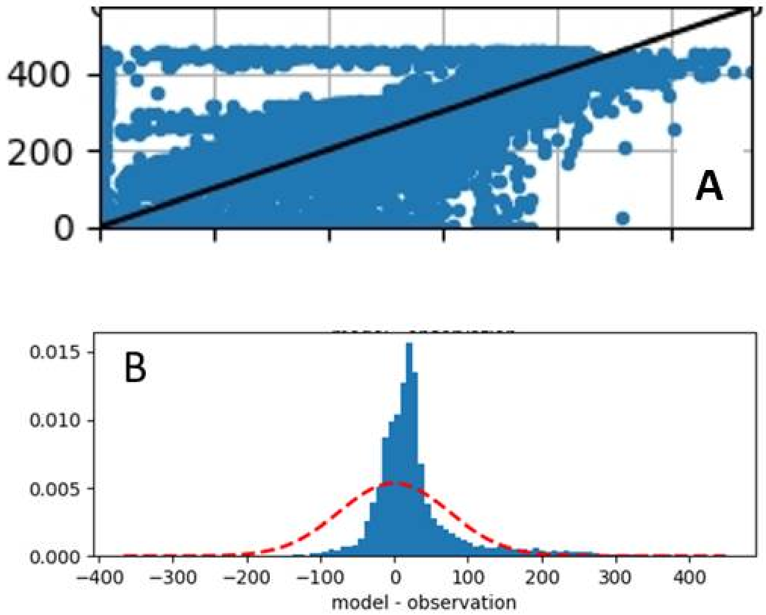

Figure 7 A) relationship between simulated and observed oxygen concentrations and B) the distribution of the differences between simulated and observed oxygen concentrations. Simulations are made with GOTM-SELMA following optimization using ACPy. Dissolved oxygen concentrations are in mmole m^{-3}

The results of the first step of oxygen calibration (step 2 Fig 2) are shown in Figure 7. These are still preliminary and are expected to improve as the model calibration proceeds through several more iterations. A more detailed examination of the oxygen data (Fig 8) shows that oxygen in the surface layer are well simulated. There is a slight tendency for model underestimation in the mid-water column. The addition of the nutrient and phytoplankton calibration steps are expected to improve the results by providing more biomass that will decompose and reduce the mid-water column oxygen concentrations.

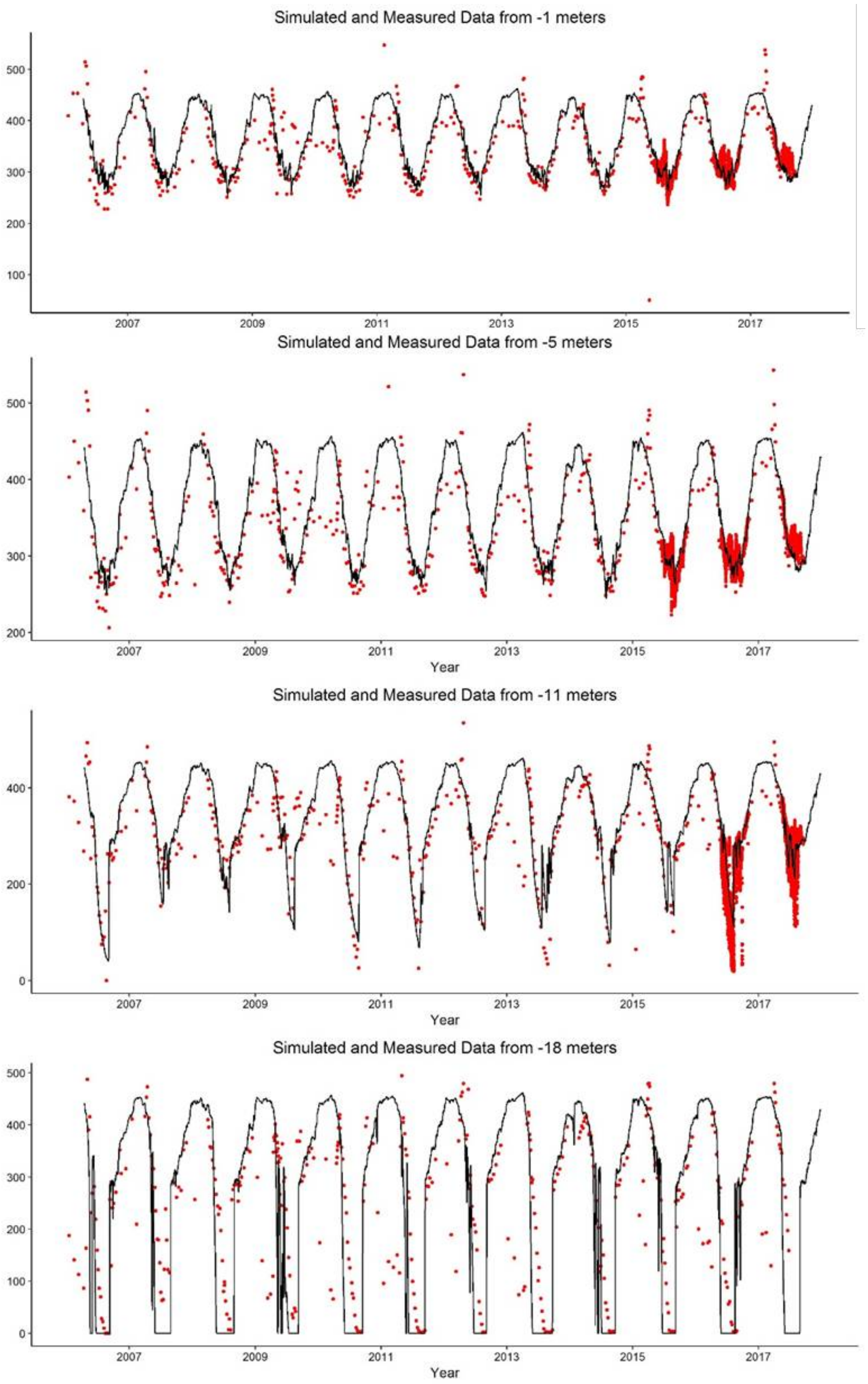

Figure 8 Simulated (solid line) and measured (points) concentrations (mmole m-3) of dissolved oxygen at four depths in the water column of Lake Erken.

Hydrological model calibration for Ter river basins associated to Sau Reservoir.

The Ter river basin has a area of 1680 m2 and is the main source of water for our case study in Spain, the Sau Reservoir. Accordingly, a hydrological modeling was implemented as a part of the workflow. The selected model was mHM v5.9 (Samaniego et al, 2010; Kumar et al, 2013). The model was calibrated between 1990-1999 and validated between 2000-2009 using EWEMBI (0.5º global observational database) as meteorological forcing data. To set the parameters that better represent discharge, an autocalibration procedure was implemented using the Shuffled Complex Evolution optimization algorithm and 1.0 - 0.5*(NSE+lnNSE) function (used to optimized both at the same time, maximum and minimum streamflows). 9 main set of parameter values were optimized: interception, snow-melt, soil moisture, runoff, potential evaporation, interflow, percolation, routing conditions and geology. The final setup for the model is presented in Figures 9 and 10 for the calibration and validation periods, respectively.

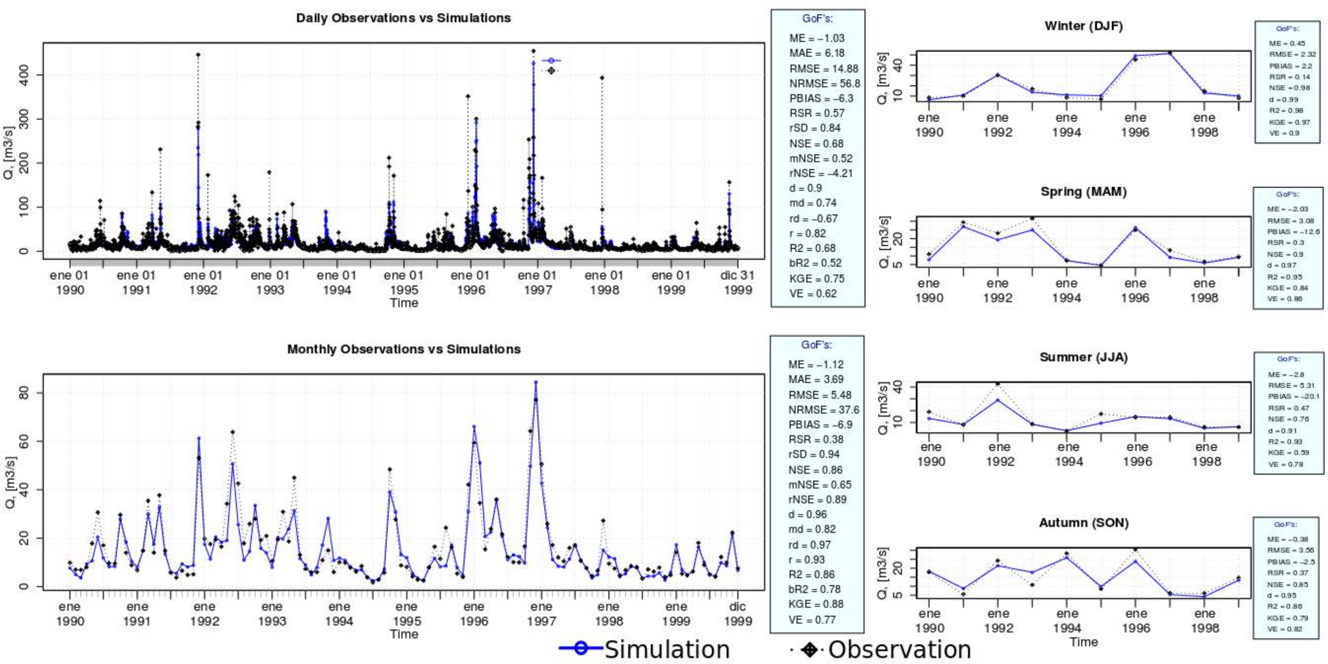

Figure 9. Hydrologic model calibration results for the Ter river basin Measures of model fit are calculated on simulated daily mean ??? and mean seasonal estimations of river discharge.

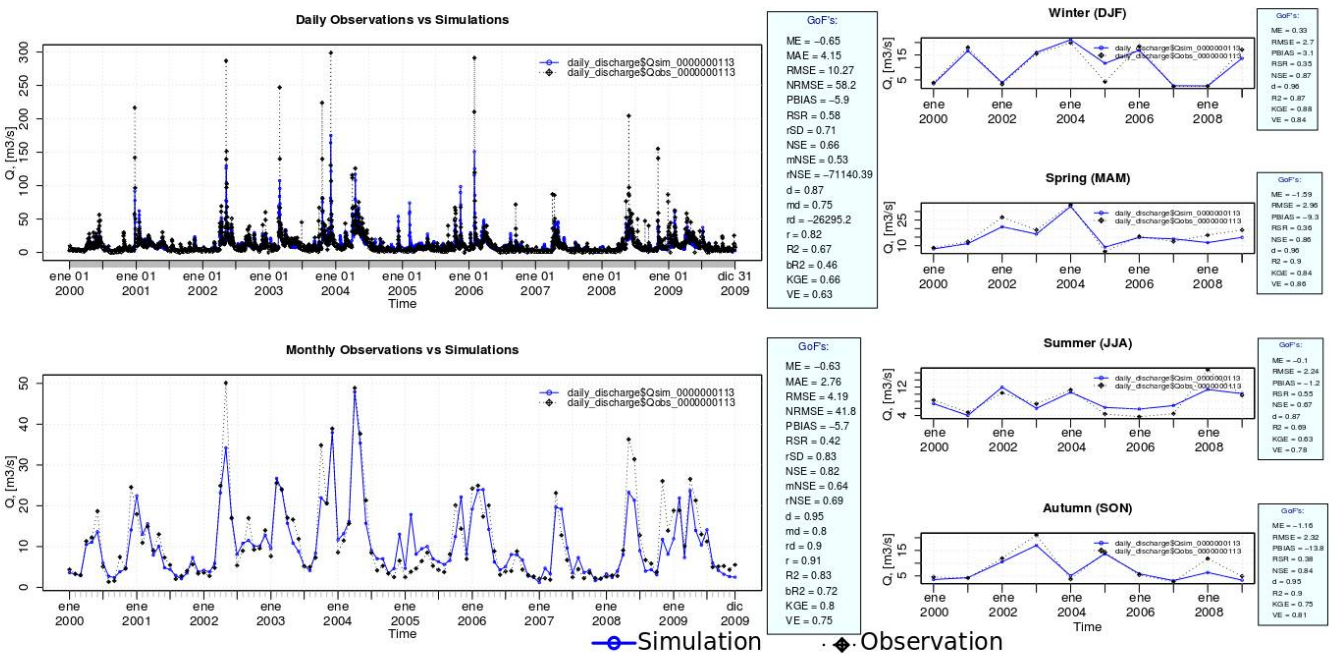

Figure 10. Hydrologic model validation results for the Ter river basin

Mt. Bold Hydrologic and hydrodynamic model calibration

Mt. Bold Reservoir is a moderately eutrophic reservoir located in the Lofty Ranges outside Adelaide in South Australia (35.12 S 138.70 E). The reservoir has a maximum depth of 41.4m which fluctuates seasonally. It has two inflows; the Onkaparinga River and the Echunga Creek. The Onkaparinga River also receives water which is pumped from the Murray River. This water is added to ensure the reservoir water level does not drop too low as this reservoir supplies drinking water to the inhabitants of Adelaide. The modeling workflow (Fig. 11) will require a hydrology model to forecast the inflows coming into the reservoir and a hydrodynamic model to model the water level and thermal structure of the reservoir.

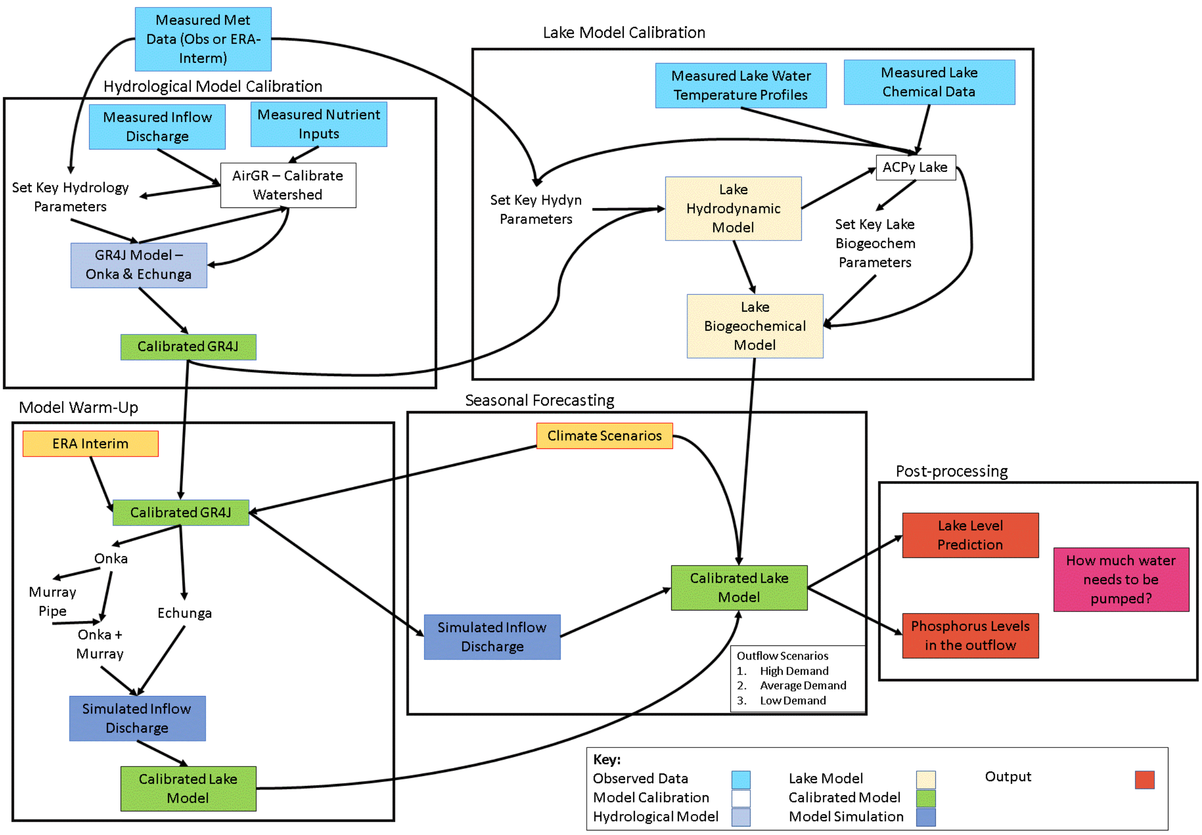

Figure 11. Schematic showing the steps involved in calibrating the hydrological and lake models. Then it shows how the model warm-up period is run followed by the seasonal forecasts.

The GR4J hydrological model was calibrated using measured daily flow data from both catchments. The model is forced by precipitation, evapotranspiration and mean air temperature. Figure 12 shows the calibration period for the Echunga Creek and it shows how the simulated water flow and inputs to the reservoir match up with the observed data.

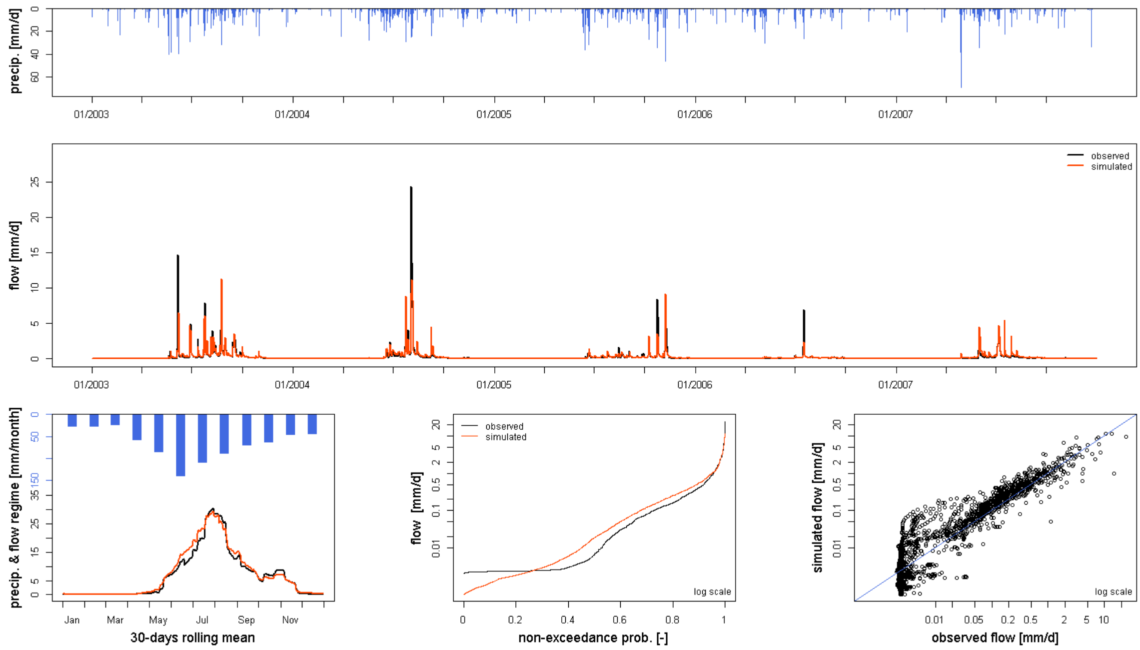

Figure 12. Plots showing precipitation inputs, and time series of modelled and observed flow data. The bottom panel shows rolling 30 day average of modelled and observed flows, the probability of non-exceedance and a scatter plot of observed vs modelled flow for the Echunga Creek.

The reservoir will be modeled using GOTM (General Ocean Turbulence Model Burchard 2002) which will simulate variations in reservoir hydrodynamics and water level. This will allow predictions to be made of the expected storage within the reservoir which could in turn be used to predict how much water will be needed from the Murray River to keep the reservoir at a desired level. The hydrodynamic model will be dynamically coupled with SELMA (Simple ecological model for aquatic systems) using FABM (Framework for Aquatic Biogeochemical Models) which will allow simulation of phosphorus levels within the reservoir. Currently Mt. Bold Reservoir is responsible for the majority of the phosphorus load to Happy Valley Reservoir, especially providing labile phosphorus in the Summer. This contributes to the frequent growth of problematic concentrations of pelagic Cyanobacteria and presents challenges for the water treatment plant fed from Happy Valley Reservoir.

Burrishoole catchment – Quantification of fish migration phenology.

The Burrishoole catchment (~ 100 km^2) drains into the Northeast Atlantic through Clew bay on the west coast of Ireland (53° 56’N, 9° 35’ W). The catchment is situated in the Nephin Mountain range and experiences a temperate oceanic climate with typically mild winters (2 – 4 °C minimum air temperatures) and cool summers (15 – 20 °C maximum air temperature). Burrishoole contains approximately 474 ha of wetted area (449 ha lacustrine, 25 ha fluvial), which is occupied by diadromous fishes (European eel (*Anguilla anguilla*), Atlantic salmon (*Salmo salar*) and sea trout (*Salmo trutta*)). Full fish trapping facilities at the only two freshwater outflows from the catchment have been in place since 1970, affording the Marine Institute with a multi-decadal complete census of daily downstream (adult eel, and juvenile salmon and sea trout smolts) and upstream (adult salmon and sea trout) fish movements.

Stakeholder interests, questions and expectations

The Marine Institute is interested in the timing of three primary migration events that are closely associated with environmental conditions during the northern meteorological seasons: (i) catadromous silver eel spawning migration (autumn), (ii) salmon smolt seaward migration (spring) and (iii) sea trout smolt seaward migration (spring). The Marine Institute have historically conducted analyses of daily fish census data to quantify long term trends (e.g., in silver eel abundance (Poole et al., 2018), and salmon and trout survival (de Eyto et al., 2016)) and have explored statistical associations between daily environmental variability and fish movements (e.g., silver eel (Sandlund et al., 2017), Atlantic salmon (Byrne et al., 2003), and sea trout (Byrne et al., 2004)). Under the framework of WATExR, the Marine Institute would like to know the extent to which daily counts of migrating fishes are predictable; and, by extension, the extent to which seasonal meteorological projections provide model forcing/driving data that are useful for predicting migration phenology (i.e., in terms of start, peak and end date of migration periods).

Methods

To address the questions posed by the stakeholder, we have defined a modeling workflow (Fig 13) that is partitioned into three evaluation procedures that are herein described in terms of one fish species (European eel), but a similar approach may apply to other models for eels and other species for which more than one theoretically plausible model may be applied: (i) an evaluation of a model that quantifies associations between environmental conditions and daily counts including its predictive performance using held-out test data; (ii) an evaluation of a model that quantifies migration phenology (start, peak and end dates of migration period) from a time series of daily counts that span an entire migration season; and (iii) an evaluation uncertainty propagation for the entire workflow that predicts migration phenology from a seasonal forecast at the daily time step. All model building, calibration and testing was carried out using R (R Core Team, 2018), using packages glmmTMB (Magnusson et al., 2018) for regression modeling, bbmle (Bolker and R Development Core Team, 2017) for distribution modeling, and DHARMa (Hartig, 2019) for statistical validation.

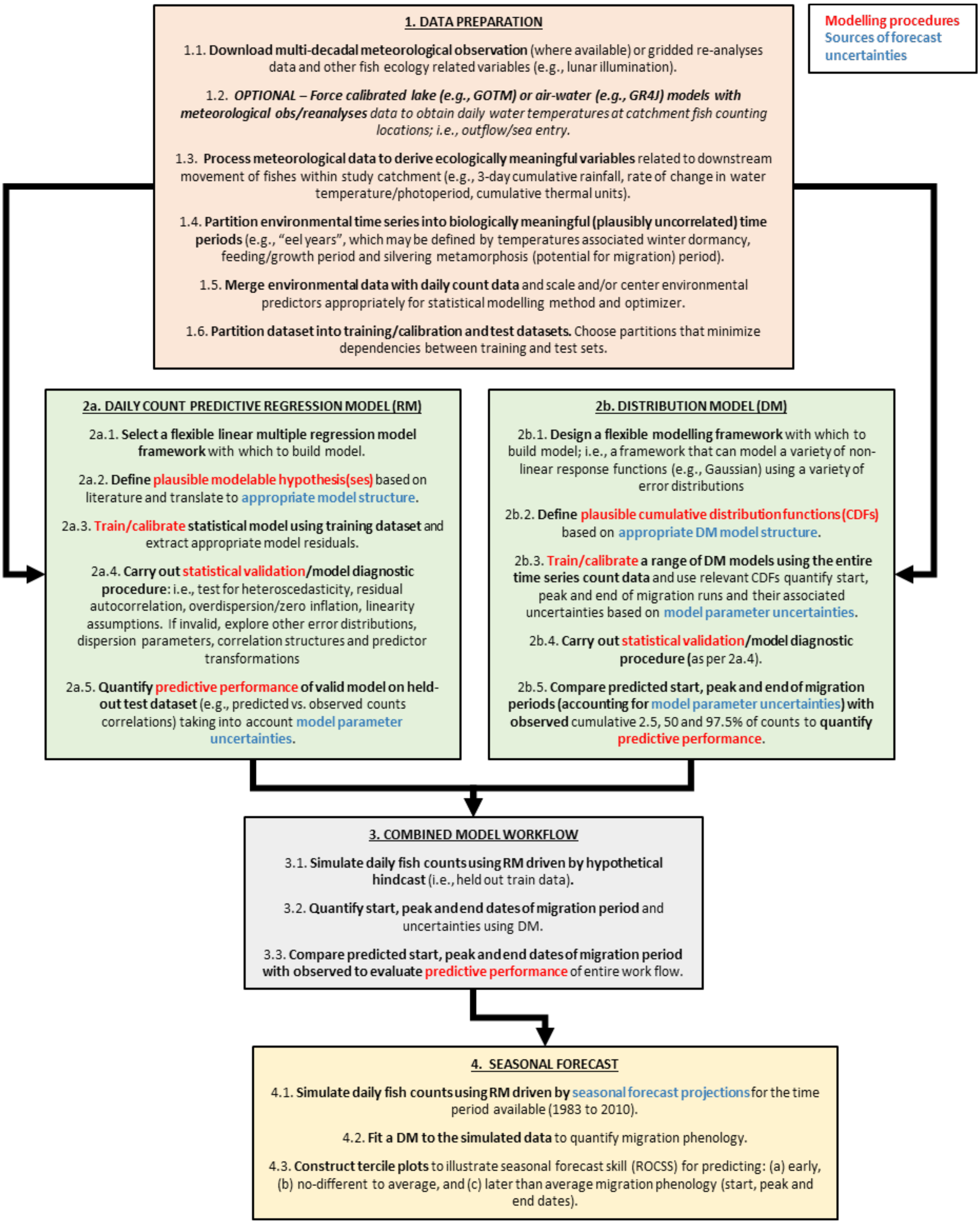

Figure 13. Modeling workflow and sources of uncertainties for the seasonal forecasting of fish migration phenology in the Burrishoole catchment.

Preliminary results

The figures presented here illustrate some regression model and distribution model statistical validation procedures and evaluations of model predictive performance prior to testing with seasonal meteorological projections (i.e., up to the end of steps **2a** and **2b** in the workflow Fig 13).

Regression model

We used a generalized linear mixed effect model framework to model daily eel counts. Calibrations were made using daily eel count data collected between 1993 and 2013 (training data), while testing data was held-out for the period 1970 to 1992 (excluding eel count data for the years 1978, 1984 and 1989, during which flooding around the fish traps lead to an incomplete census). Complete water temperature data required to calibrate the regression model were collected adjacent to the fish traps. Meteorological data (total daily rainfall) were obtained from records taken from a local automated weather station. Photoperiod data were collected using the R package geosphere (Hijmans, 2017), and lunar illumination data were collected using the R package lunar (Lazaridis, 2014). Model statistical validation and predictive performance evaluation (Fig 14) was carried out in accordance with 2a.4. described in Fig 13.

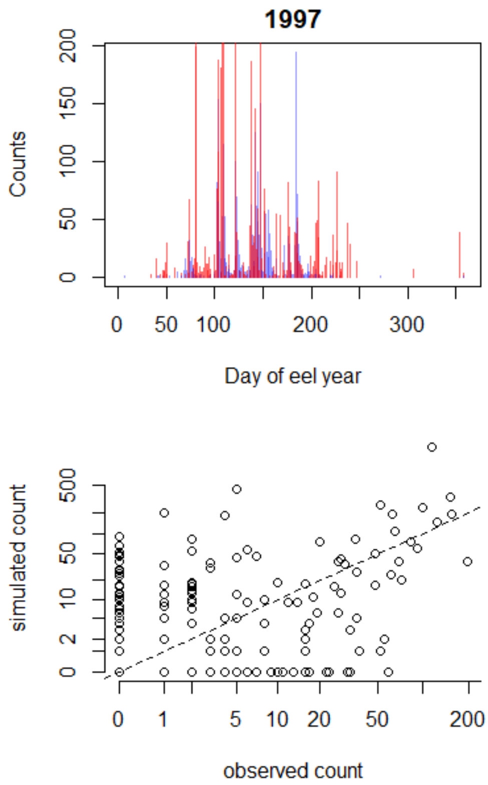

Figure 14. Example of simulated counts plotted alongside observed data for zero-inflated generalized linear mixed effect model. Vertical blue lines are observations and semi-transparent red lines are simulated counts. The dashed line illustrates x = y. The model structure was composed of six parameters distributed within conditional and zero inflation mixture (three parameters in each) and included all possible interactions. The negative binomial error distribution (and log link) was used in this instance. No additional dispersion parameter or correlation structure (e.g., autoregressive AR(1)) (beyond an inherent compound symmetry structure implied by the random intercept of “eel thermal year”) were implemented at this stage, but remain a focus of ongoing work.

Distribution model

We explored the quantification of fish migration phenology using a range of non-linear distribution models (e.g., Gaussian, Weibull) and error distributions (e.g., Poisson, negative binomial). Cumulative distribution functions were not appropriately defined for all our applications of functions (e.g., the t-distribution), but nevertheless we have applied empirical estimates in the meantime. The statistical validity of our choices of distribution model varied between non-linear functions, though this did not necessarily correlate with predictive performance of phenology (e.g., Weibull appearing valid for the vast majority of years (not shown), but only appearing to predict the end of the migration period with useful accuracy; Figs 4 and 5) - we emphasize that these results are preliminary demonstrations of our modeling workflow.

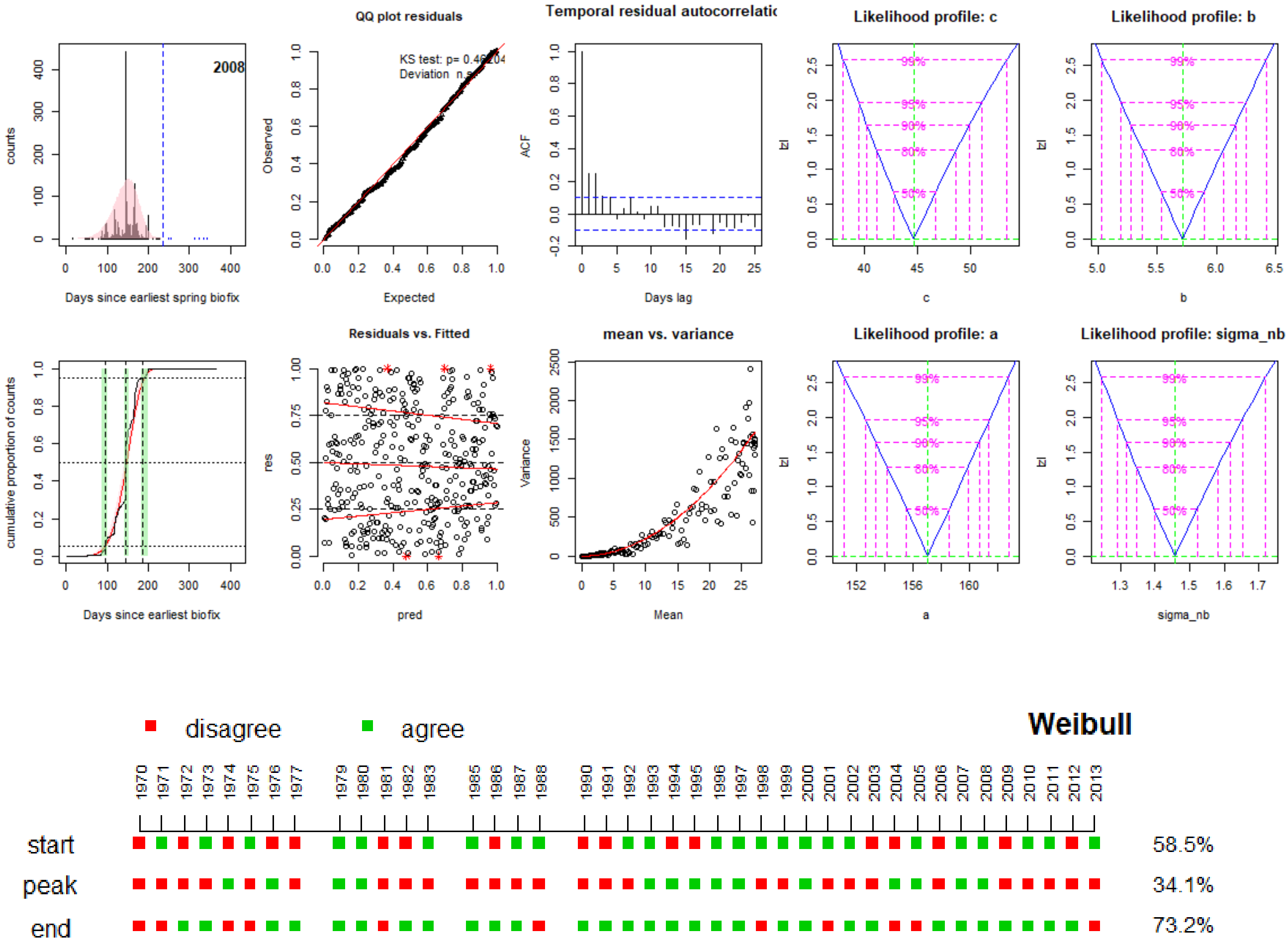

Figure 154. Summary plot of statistical validation and predictive performance of a three-parameter Weibull distribution (a, b, c) (plus negative binomial dispersion parameter, $1/\text{sigma\_nb}^2$) for quantifying the phenology of the 2008 European silver eel spawning migration. Topleft plot shows the observed counts (black bars) and a bootstrapped 95% (percentile) prediction interval (pink transparency) derived from data simulated from the model.

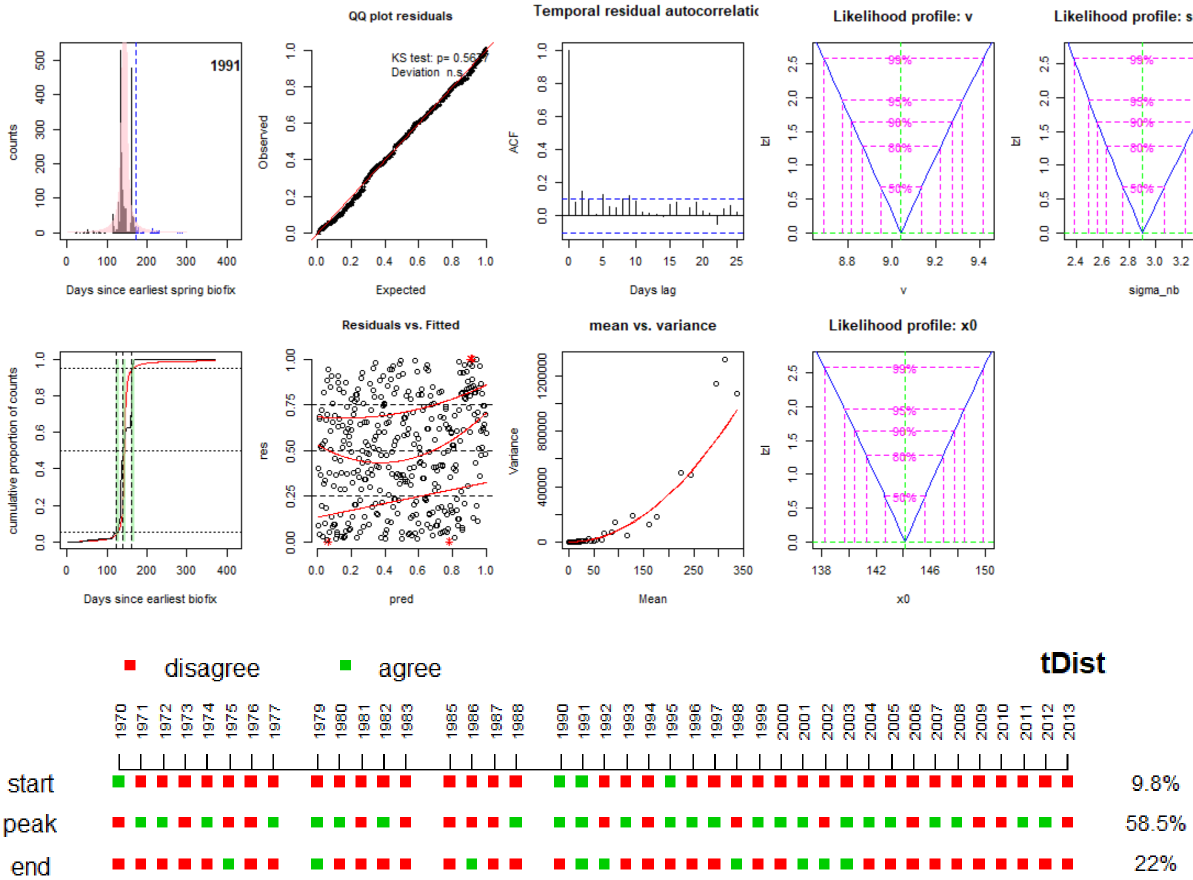

Figure 16. Summary plot of statistical validation and predictive performance of a two parameter (v, x0) t-distribution (plus a negative binomial dispersion parameter, 1/sigma_nb^2) for quantifying the phenology of the 1991 European silver eel spawning migration. Topleft plot shows the observed counts (black bars) and a bootstrapped 95% prediction interval (pink transparency) derived from data simulated from the model.

Discussion and recommendations

The results presented herein are preliminary and do not include seasonal phenology forecasts. Nevertheless, these results illustrate a feasible workflow for prediction of fish migration phenology at Burrishoole, which has in the case of the panmictic European eel, the potential for geographic extrapolation to other catchments in Europe.

Notable modeling limitations encountered to date include limited computing capacity for model complexity/no. regression parameters in models that require a correlation structure. We also note that the range of distribution models used to quantify migration phenology of the European eel might simply have too light (thin) distribution tails to account for the sporadic migration of silver eels at large temporal distances from the peak of the migration period and are therefore susceptible to leverage of “extreme outliers”.

The next stage for this case study will be to demonstrate the complete workflow up to operational seasonal forecasts.

References

Arnold, J. G., and N. Fohrer. 2005. SWAT2000: current capabilities and research opportunities in applied watershed modelling. Hydrological Processes **19**:563-572.

Burchard, H. 2002. Applied turbulence modelling in marine waters. Springer Science & Business Media.

Cole, T. T., and S. A. Wells. 2002. CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.1. Instruction Report EL-2002-1., U.S. Army Engineering and Research Development Center, Vicksburg, MS.

Hipsey, M. R., L. C. Bruce, C. Boon, B. Busch, C. C. Carey, D. P. Hamilton, P. C. Hanson, J. S. Read, E. de Sousa, M. Weber, and L. A. Winslow. 2019. A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). Geosci. Model Dev. **12**:473-523.

Jackson-Blake, L. A., A. J. Wade, M. N. Futter, D. Butterfield, R. M. Couture, B. A. Cox, J. Crossman, P. Ekholm, S. J. Halliday, L. Jin, D. S. L. Lawrence, A. Lepistö, Y. Lin, K. Rankinen, and P. G. Whitehead. 2016. The INtegrated CAtchment model of phosphorus dynamics (INCA-P): Description and demonstration of new model structure and equations. Environmental Modelling & Software **83**:356-386.

Janse, J. H. 1997. A model of nutrient dynamics in shallow lakes in relation to multiple stable states Shallow Lakes’ 95 Hydrobiologia **342/343** 1-8.

Saloranta, T. M., and T. Andersen. 2007. MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. Ecological Modelling **207**:45-60.

Bolker, B., R Development Core Team, 2017. bbmle: Tools for General Maximum Likelihood Estimation.

Byrne, C.J., Poole, R., Dillane, M., Rogan, G., Whelan, K.F., 2004. Temporal and environmental influences on the variation in sea trout (Salmo trutta L.) smolt migration in the Burrishoole system in the west of Ireland from 1971 to 2000. Fish. Res. 66, 85–94. https://doi.org/10.1016/S0165-7836(03)00146-2

Byrne, C.J., Poole, R., Rogan, G., Dillane, M., Whelan, K.F., 2003. Temporal and environmental influences on the variation in Atlantic salmon smolt migration in the Burrishoole system 1970–2000. J. Fish Biol. 63, 1552–1564.

de Eyto, E., Dalton, C., Dillane, M., Jennings, E., McGinnity, P., O’Dwyer, B., Poole, R., Rogan, G., Taylor, D., 2016. The response of North Atlantic diadromous fish to multiple stressors, including land use change: a multidecadal study. Can. J. Fish. Aquat. Sci. 73, 1759–1769. https://doi.org/10.1139/cjfas-2015-0450

Hartig, F., 2019. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.

Hijmans, R.J., 2017. geosphere: Spherical Trigonometry.

Lazaridis, E., 2014. lunar: Lunar Phase & Distance, Seasons and Other Environmental Factors.

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Martin Maechler, Koen van Bentham, Ben Bolker, Mollie Brooks, 2018. Generalized Linear Mixed Models using Template Model Builder.

Poole, W.R., Diserud, O.H., Thorstad, E.B., Durif, C.M., Dolan, C., Sandlund, O.T., Bergesen, K., Rogan, G., Kelly, S., Vøllestad, L.A., 2018. Long-term variation in numbers and biomass of silver eels being produced in two European river systems. ICES J. Mar. Sci. https://doi.org/10.1093/icesjms/fsy053

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sandlund, O.T., Diserud, O.H., Poole, R., Bergesen, K., Dillane, M., Rogan, G., Durif, C., Thorstad, E.B., Vøllestad, L.A., 2017. Timing and pattern of annual silver eel migration in two European watersheds are determined by similar cues. Ecol. Evol. 7, 5956–5966. https://doi.org/10.1002/ece3.3099