SEVENTH FRAMEWORK PROGRAMME
THEME 6: Environment (including Climate Change)

Adaptive strategies to Mitigate the Impacts of Climate Change on European Freshwater Ecosystems

Collaborative Project (large-scale integrating project)
Grant Agreement 244121
Duration: February 1st, 2010 – January 31st, 2014

Deliverable 5.1: The REFRESH Common Modelling Framework for the Demonstration Catchments

Lead contractor: **UREAD**
Other contractors involved: **Alterra, AU, BCAS, Bioforsk, Deltares, METU, MLURI (JHI), NERC, NIVA, SLU, SYKE, TRENT, UB, UU-BIO, UDE, UPAT.**

Due date of deliverable: **Month 28 (May 2012)**
Actual submission date: **February 2013**

Work package: 5
Contributors: Alterra, AU, BCAS, Bioforsk, Deltares, METU, MLURI (JHI), NERC, NIVA, SLU, SYKE, TRENT, UB, UU-BIO, UDE, UREAD, UPAT.
Estimated person months: 19

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013) Dissemination Level (add X to PU, PP. RE or CO)		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	X

Abstract

This document describes the REFRESH Common Modelling Framework which is designed to promote best modelling practice in the project and produce outcomes that are comparable between demonstration sites. The Common Modelling Framework is a set of guidelines for the model applications and describes how model performance will be evaluated and scenario outcomes assessed. This document also records the detailed work plan to address the three objectives of the integrated biophysical modelling and the integration between the work-packages focused on scenario generation (work package 1), experimentation in rivers, lakes and wetlands (work packages 2, 3 and 4) and the socio-economic assessments (work package 6). It is intended that this report should be a living document to record modelling practice within the REFRESH project until the end date of 31 January 2014.

Document change control

Date	Author	Version	Comments
09 Dec 2011	Attila Lazar	0.a	Document creation
12 Feb 2012	Andrew Wade	0.b	Edit and update
06 Mar 2012	Andrew Wade	0.c	Document edit
12 Mar 2012	Andrew Wade	0.d	Document edit – response to Reading review
13 Mar 2012	Andrew Wade	0.e	Document edit – response to Reading review
13 Mar 2012	Andrew Wade	0.e.1	Document minor edit
02 Apr 2012	Andrew Wade Ahti Lepistö	0.f	Document edit – update following Third Project meeting
04 Apr 2012	Andrew Wade	0.g	Converted from .docx to .doc for upload to REFRESH filestore
20 July 2012	Andrew Wade	0.h	Inclusion of partner comments and site specific updates. Specifically comments from JHI, NIVA, Bioforsk, Reading, Deltares.
01 Aug 2012	Andrew Wade	0.i	Incorporation of further comments from Ahti Lepisto and Richard Skeffington
08 Aug 2012	Andrew Wade	0.j	Incorporation of check-list for modelling procedure which also forms the template for the write-up of the model applications.
10 Aug 2012	Andrew Wade	0.k	Issued to Project Co-ordinator for checking.
17 Aug 2012	Andrew Wade	0.l	Incorporation of Project Co-ordinator comments
21 Aug 2012	Josef Hejzlar	0.l_JH	Document edit
21 Aug 2012	Dennis Trolle	0.l_JH_DT	Document edit
11 Sep 2012	Andrew Wade	0.m	Document issued for external review by JCR.
13 Feb 2013	Andrew Wade	0.n	Document edit based on JCR comments and discussions at WP5 meeting, September 2012 ad intervening period.
14 Feb 2013	Andrew Wade	1.0	Issued as a deliverable

Table of Contents

1. Introduction

The REFRESH Common Modelling Framework – a theoretical ideal put into practice

The aim of the biophysical modelling work within the REFRESH project is to simulate the baseline flow and nitrogen (N) and phosphorus (P) concentrations at eight demonstration sites across Europe using state-of-the-art hydrological and water quality models, and then to use these models to investigate the likely changes in the flow and N and P concentrations under integrated scenarios of climate, land-use, deposition and water use change. These model applications will take the science further by seeking to determine how the changes in flow and water quality will affect the freshwater biota, and by evaluating the cost effectiveness of measures to mitigate water pollution and adapt to climate change. Specifically there are three objectives:

- to improve the coupled representation of rivers, lakes and wetlands in catchment-scale biophysical models;
- to link catchment-scale, process-based hydrochemical models with measures of ecological impacts, obtained from the literature and the experimental work done in REFRESH, to determine how environmental change will affect indicators of ecological status relevant to the Water Framework Directive and Habitats Directive;
- to integrate the outcomes from the biophysical modelling into cost effectiveness assessments to better determine the most appropriate measures to reduce catchment nitrogen and phosphorus pollution when set against the background of environmental change and food and water security.

To achieve these three objectives and thereby deliver the overall aim, it is necessary to apply and test hydrochemical models at each of the eight REFRESH demonstration sites. These eight modelling studies should be based on best practice and, ideally, done in a consistent way to allow ready comparison of the modelled outcomes and assessments of uncertainty. In addition, planning is required regarding how each of the three objectives, noted above, will be delivered.

This document describes the REFRESH Common Modelling Framework (CMF) which is designed to promote best modelling practice in the project and produce outcomes that are comparable between demonstration sites (sections 2, 3 and 4). This document also records the plans to address the three objectives outlined above (sections 5 and 6). It is intended that this report should be a living document to record modelling practice within the REFRESH project until the end date of 31 January 2014.

The REFRESH CMF is based on the modelling process developed under the EuroHarp project and the broader catchment biophysical modelling literature (STOWA, 1999; Silgram and Schoumans, 2004; Refsgaard et al., 2007). The framework was further developed through discussions at REFRESH workpackage 5 workshops in Antalya and Aberdeen and Sitges. The outcomes of these three meetings are recorded in the minutes of each which are stored on the REFRESH website (http://www.refresh.ucl.ac.uk/).

The REFRESH CMF is detailed in section 2. A check-list for the model applications following the REFRESH CMF is provided in section 3 to the help the REFRESH modellers monitor task completion and to serve as a **common template for the write-up** of each model application.

Given the complexity of applying biophysical models to eight different sites across Europe due to differences in hydrological and water quality issues, climate, geology, soils, land cover, land management, ecosystem type and data availability, it has so far proved impractical to apply exactly the same method at each demonstration catchment due to resource constraints. Furthermore, there are differences in modelling practice compared with previous projects because of the need to develop and apply appropriate modelling tools to address the aim and three objectives. To manage the model applications in a pragmatic way to account for these differences, the REFRESH Common Modelling Framework is thought of as theoretical ideal. The necessary departures from it and the reasons for these, for each of the eight demonstration catchments, are recorded in section 4 of this document.

The intention is that the output from the applications of the process-based biophysical models will be used as input to empirical relationships that link the ecological response to flow and nitrogen or phosphorus concentrations, to determine the subsequent effect on the ecological structure, function and biodiversity of each site (section 5). To represent a multi-stressor environment, Bayesian Belief Networks will be developed to represent the relationships between the abiotic and biotic habitat and the ecological indicators. The output from the process-based models will also be linked to these Bayesian Belief Networks where possible. Specifically, planned model enhancements to improve the representation of rivers, lakes and wetlands in process-based dynamics models are given in sections 5.2.1 and 5.2.2. Ideas to link process-based biophysical models with ecological impacts are given in section 5.2.3. The ecological impact models are innovative as they integrate expert knowledge to summarise the multiple controls on stream macroinvertebrate biodiversity, riparian plant communities and stream algal growth. In addition, the procedure to link the outcomes of biophysical models with Cost Effectiveness Assessment is given in section 6.

2. The REFRESH Common Modelling Framework

The REFRESH Common Modelling Framework consists of six steps: model setup, calibration, testing, sensitivity analysis, uncertainty analysis and scenario analysis (Figure 1). In this case, the rationale for the development and application of the biophysical models is provided in the Description of Work. In line with the work of Oreskes et al. (1994) it is agreed that no biophysical modelling approach can strictly be validated, but instead models can be evaluated according to whether they provide a reasonable representation of a set of observations or not. Furthermore, given the sources of uncertainty associated with models and their use, the REFRESH modelling group strongly adheres to the concept of models as 'learning tools' which are used alongside experimental and monitoring evidence to consider plausible ecosystem responses, in terms of flow, water quality, ecological structure and function, and biodiversity to future environmental change (Wade et al., 2008).

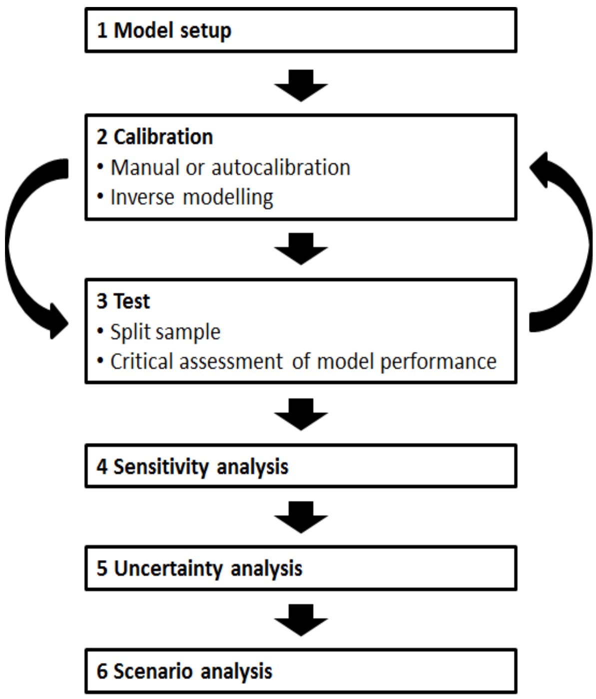

Figure 1. The REFRESH Common Modelling Framework

A model journal (i.e. a logbook) will be used to record all modelling activities and decisions made for each demonstration site. Such a model journal is an essential record-keeping tool, allows the re-assessment and continuation of the current study at a later stage and will help keep the modelling activities transparent. The model journal should contain information about the sources/types of data used in the study, estimates of all model parameters, and a detailed methodology of the model set-up and model calibration. An example of a model journal can be found in STOWA (1999).

2.1. Model setup

Each catchment-scale, process-based model application will characterise the key land cover types, atmospheric deposition, effluent inputs, water abstractions and the driving climate variables, namely precipitation (or a derivative, the hydrologically-effective rainfall) and air temperature. This set up in terms of the conceptual model applied and the geometry of the model will be specified in each report for the model applications to each demonstration site. In particular, it is necessary to specify the units of land cover and soil-type accounted for and the connections between these. The models will be run at a daily time step and make best use of observed flow and water quality observations for calibration and testing. The geometry of the applied models must allow differentiation between hydrochemical responses within the demonstration site.

A novel aspect of REFRESH is the use of chaining models to simulate the impacts of environmental change on flow and nutrient delivery from rivers to wetlands and/or lakes, and then to determine the effect on the wetland and lake ecosystems in terms of ecological structure, function and biodiversity. When chained models are used then the set-up, calibration and testing of each component model in that chain will need to be described.

2.2. Model calibration

2.2.1. Manual calibration

In the first instance, model calibration will be done manually to gain an understanding of model behaviour. Iterative calibration between the flow and biogeochemical components of process-based models tends to lead to better model performance when judged using goodness of fit criteria and this will be done (McIntyre et al. 2005; Medici et al. 2010; section 2.3.2).

Process-based, catchment-scale models of flow and water quality are complex and automatic calibration is useful for searching through parameter ranges to find suitable combinations that give rise to acceptable model behaviours. Where automatic calibration routines are used, the modeller should test some of the parameter sets to check the model is giving the right results for the right reasons. This may be done in the form of a sensitivity analysis (section 2.4.1).

2.2.2. Inverse modelling and auto-calibration

Inverse modelling is being trialled in the set-up of the INCA-P model in the Vansjø-Hobøl catchment in Norway. Selected parameters used by INCA-P will be estimated through a Bayesian inference scheme, where each parameter is given a prior distribution and a posterior distribution is estimated using a novel MCMC algorithm highly suitable for complex models (MCMC-DREAM, Vrught et al., 2008; 2009). Essentially this makes the application deviate from Figure 1 by using an automated algorithm in Matlab to perform step 2 and 3 simultaneously, i.e. a form of autocalibration of the model. The Norwegian application of INCA-P will therefore also deviate from the steps delineated below in section 2.3, model testing, as it includes positing formal likelihoods in the Bayesian analysis. The MCMC-DREAM algorithm also allows for using posterior distributions and model runs to run scenarios, in which the estimated parameter uncertainty will also be included in the simulated futures, and outputs will be presented as probability distributions and not as single time-series or numbers.

The use of PEST is also being trialled for use with INCA-N and INCA-P. The former is based on the River Kennet system and the latter is based on an application in Finland. Both DREAM and PEST will be used to explore the benefits of adding more geometrical units (e.g. land cover and soil types) to the model structure and determine to what extent adding more data reduces uncertainty.

2.3. Model testing

2.3.1. Split-sample testing

The observed flow and water quality datasets should be split both in terms of time period (*split-in-time*) or spatial location (*split-in-space*) to allow calibration at one set of locations for one time period and model testing for a second time-period or second set of locations. The *split-in-space* model calibration and testing may be done by calibrating a model to a subset of sites on the main river channel and testing against the remaining sites with observations. Given the available resources within the project, it is not feasible to calibrate the model for one sub-catchment and then test the model for another, adjacent sub-catchment at all eight demonstration sites.

An initial test of model performance will not be done as in EuroHarp (Schoumans, 2003). REFRESH and EuroHarp have different aims and objectives. EuroHarp aimed to study the structure and behaviour of several nitrogen and phosphorus models by assessing their sensitivity to available data, their performance and thus the applicability of different models on the same study areas. REFRESH does not aim to rank available models, but aims to use one or two process-based models intensively as learning tools to assess the likelihood of flow, water quality and ecological changes under scenarios of environmental change.

Where possible, the model calibration and testing will be done by two modellers. The first modeller calibrates the model, and when satisfied with the results, provides both the results and the calibrated parameter set to the second modeller. The second modeller runs the model test with the calibrated parameter set and assesses the model results, identifies areas of good model performance and problematic issues (e.g. missed extremes, poor baseflow simulation, etc.), and thus gives a critical model performance review to the first modeller. This independent expert opinion will be used by the first modeller to improve the model performance in a re-run of the model for the calibration period. This interaction between the first and second modeller can take place several times, until both modellers agree that the model has potentially reached optimal performance for the issue being studied (Figure 1). The results and the final parameter set and the conclusions on the system behaviour will be recorded. At the same time, the two modellers should summarise what has been learnt from the calibration and test phase, such as: did the results improve with subsequent calibration and re-testing; what were the critical processes in the study area; which processes would need further field-based investigation; has the model achieved the modelling objectives; what aspects of the river-system does the modelling framework represent well, or poorly? Where model calibration and testing by two modellers is impractical due to insufficient resources or in-house expertise, then emphasis shall be placed on the application of multiple models to assess structural uncertainty. When applying an advanced autocalibration tool (e.g. an MCMC algorithm for the Vansjø-Hobøl catchment), the focus will be put on expert-based selection of realistic parameter set(s).

2.3.2. Judging the modelled outcomes

Visual inspection of the fit of the simulated data to that observed, and a comparison of process fluxes with published values, should be used throughout the calibration and testing processes to ensure that parameter adjustments contribute to realistic model outputs. Published values for nitrogen and phosphorus fluxes associated with soil and in-stream processes are collated in Whitehead et al., 1998; Wade et al., 2002; Rankinen et al. 2004 and Wade et al., 2008. These published values can be used to compare with the annual flux estimates produced by INCA-N and INCA-P and other models to verify behaviour.

As part of model testing, each model application must be assessed using goodness of fit criteria to determine if the results are a reasonable fit to the data and modelled behaviour is sensible. The Nash-Sutcliffe coefficient and the Root Mean Squared Error paired-tests (simulated vs. observed values) will be as used as a basis to assess the model performance in terms of both streamwater concentrations and loads (Nash and Sutcliffe, 1970). In addition, statistics describing the volumetric water balance and pollutant loads will be compared with measured data. Water quality parameters are affected by both biotic and abiotic processes, thus predictions are often less accurate than for the flow (e.g. Wade et al., 2001). Thus, it is recommended to do un-paired performance tests for water chemistry measures. For details of these paired and un-paired tests see Silgram and Schoumans (2004). Recent work by Jolliff et al. 2009 will be considered for innovative ways to illustrate model skill.

The cumulative frequency distributions for the observed and the modelled data should also be reported. This is an 'unpaired' test which does not take account of the temporal pattern of the data series, but can be useful for comparing the percentage of time a critical value is exceeded (e.g. Q_{95}). Given the complex controls on flow and water chemistry, the model simulation of an observed time series is an exacting performance test. The assessment of whether a model can reproduce mean and extreme behaviours, in terms of cumulative frequency distributions, allows consideration of overall system behaviour. The Nash-Sutcliffe and Root Mean Squared Error paired test can have low values because of only slight differences in magnitude and/or timing of streamwater concentration peaks between simulated and observed time series.

The individual performance tolerance limit to reject a model application will be judged on a case-by-case basis. It will be left to the modellers' discretion to judge model fit based on the temporal resolution and quality of the observed flow and water quality data. The simulation of hydrochemistry together with hydrology can help constrain the behaviour of the hydrological component of the model to provide plausible explanations of the system flow pathways (Birkel et al., 2011; Medici et al., 2010; 2012).

The principal quantitative criterion to be used during calibration and testing for assessing how well the modelled series fits the observed data is the Nash-Sutcliffe coefficient of determination (Aitken, 1973):

Nash-Sutcliffe criterion, N-S is given by

$$N\text{-}S\ criterion = \frac{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2 - \sum_{i=1}^{n}\left(O_i - P_i\right)^2}{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2} \qquad [1]$$

where O_i is the i^{th} observed data point, P_i is the i^{th} modelled data point commensurate with the i^{th} observed data point, $\bar{O}$ is the mean of the observed data series, and n is the number of data points. The maximum value that can be achieved is one, and in this case the model successfully reproduces all of the variation about the mean of the series. Model calibrations should, therefore, attempt to achieve the maximum N-S value possible. For systems with a limited number of data points and a high variance, this number may be significantly less than one. The N-S criterion can have negative values, in which case the model performs less well than if you simply estimated each value of the observed time series using the mean of the observations.

In addition to the N-S criterion, the two paired tests of model performance should also be reported for the calibrated models:

Root Mean Squared Error, RMSE:

$$RMSE = \frac{100}{\overline{O}}\sqrt{\frac{\sum_{i=1}^{n}\left(P_i - O_i\right)^2}{n}} \qquad [2]$$

Relative Error, RE:

$$RE = \frac{100}{n}\sum_{i=1}^{n}\frac{(O_i - P_i)}{O_i} \qquad [3]$$

Whilst goodness of fit criteria will be reported, due to the uncertainties associated with environmental monitoring and modelling, it is also essential to consider model performance in terms of the issue being studied. As such, the modeller(s) working on each application should take a view, and report on, which aspects of the system's behaviour the model does, or does not, simulate well. In this project, extremes are important in terms of low- and high-flow conditions and in terms of protracted periods without water. Thus, it is recommended a model's ability to characterise low and high flow periods and the associated chemical conditions will be assessed, in addition to the extended periods of low flow, seasonal hydrological and hydro-chemical dynamics and long-term trends. Where possible the RMSE will be used for model optimisation in DREAM and PEST.

Once model calibration is complete, then the parameter values derived for different geometric units (e.g. land cover or soil type or hydrological response unit) will be compared to determine which units may behave in the same way and where there are key differences. However, for the purpose of these model applications, geometrical units will not be combined because of the need to assess how each may respond to projected environmental change.

Due to insufficient water quality observations, the biophysical models will not be calibrated by introducing observations in a step-wise manner to assess how the model parameters might vary through time. Namely the conditions of stationarity are assumed to exist. It is acknowledged that split sample calibration and testing assumes stationarity also. These assumptions must be noted in deliverables and articles. Parameter sets from the different applications of the same model but to different sites will be collated for comparison and the determination of recommended ranges for future studies. Differential split sample testing, for example looking at wet and dry periods, will not be done since this method is limited in this case by insufficient water quality data and this approach may result in model calibrations biased to particular conditions or the generation of two, or more, parameter sets that will need to be combined for scenario analysis.

2.4. Sensitivity and robustness testing

2.4.1. Sensitivity analysis

Sensitivity analysis must be done for all model applications where this has not been done previously for a particular model. There are now at least three studies of the parameter sensitivity of the INCA-N model (McIntyre et al., 2005; Rankinen et al., 2006; Shahsavani and Grimvall, 2009, 2011) and the sensitivity of the old version of INCA-P has been explored also (Dean et al., 2009), but not for the new version. Thus it is advised that, at least a simple, manual local-sensitivity analysis, should be done for the new version of INCA-P at six of the demonstration sites where the model output will be used: Dee, Louros, Pyhäjärvi, Thames/Thame, Vansjø-Hobøl and Vlatava-Rimnov.

A local sensitivity analysis is one in which only one parameter is changed at once and the response of the simulation results is analysed. In a global sensitivity analysis, many or all the parameters may be changed and assessment made of the effect of the parameter vector on the model results (Global sensitivity analysis: e.g. Spear and Hornberger, 1980; GLUE methodology: Beven and Freer, 2001). Given model behaviour is a function of the whole set of parameters and interactions of parameters, then in the REFRESH Project, a global sensitivity analysis will be used for process-based models where possible. One suggested method is the Generalised Likelihood Uncertainty Estimator (GLUE) which is based on the idea that there is no optimum parameter set and all that can be done is a comparison of the likelihood of different parameter sets giving acceptable behaviours, as judged by an objective function. Another suggested method is that of Spear and Hornberger, 1980 which is a Monte Carlo-based method like GLUE, where the model is run many times with randomly selected parameter values, and the parameter settings are segmented based on whether the simulation results were acceptable or unacceptable. With the Spear and Hornberger approach, these segmented parameter vectors are then further assessed with statistical techniques, such as the Kolmogorov-Smirnov test, which measures if there is a statistically significant difference between the parameter values that give acceptable and unacceptable model behaviours (Spear 1970). This information can be used to identify parameter ranges that give acceptable, as opposed to unacceptable, modelled outcomes and the parameters can be ranked in order of sensitivity. The outcomes of this method, in terms of narrowed ranges for acceptable parameters, can be used to inform model calibration. The selected behavioural criteria (e.g. Nash-Sutcliffe coefficient, seasonal minimum and maximum, tolerance interval around the observations, etc.) for parameter vector segmentation depend on the simulated variable type and the available observations, will thus be identified on a case-by-case basis.

2.4.2. Robustness analysis

The sensitivity analysis should also include a robustness analysis, where extreme input and parameter values (time series and/or parameter values) are used to see how the model responds. This can result in a crash of the model which helps to explore the model behaviour, but can also highlight scenarios which are unlikely, but still possible. Determining the likelihood of these extreme cases is difficult. Such situations have to be assessed manually to determine if the model behaviour is reasonable. The results of the robustness analyses are expected to be valuable for policy makers and catchment managers, because they can widen perspectives and illuminate key issues that “may otherwise be missed” (Mahmoud et al. 2009).

2.5. Uncertainty analysis

The uncertainty analysis extends the global sensitivity analysis, by running the acceptable ‘global sensitivity analysis’ parameter sets again, and recording the simulated values. These results will be statistically aggregated to draw the band of uncertainty around the calibrated model results. This uncertainty should be compared with the simulation results of the ‘robustness analysis’. This uncertainty analysis should be done for simulated flow and streamwater nitrogen and phosphorus concentrations. In addition to this, part of the Norwegian application will utilize Bayesian inference scheme (see section 2.2.2).

To deal with structural uncertainty, there will be multiple model applications at the following demonstration sites:

- Dee, UK (INCA-N and StreamN);
- Font del Regas- Arbúcies-Fuirosos, La Tordera, Spain (N catchment models);
- Skuterud-Vansjø-Hobøl, Norway (INCA-N and –P);
- Vltava-Rimov, Czech Republic (INCA-N and P, export co-efficient);
- Yläneenjoki/Pyhäjärvi, Finland (MyLake and LLR lake models);
- and if resources allow, Lake Beyshir and catchment, Turkey (SWAT, INCA-N and P).

Applying more than one model to a study area allows the result and parameter settings to be compared to learn more about the underlying processes of the system and gaps in the scientific (field-based) understanding and model representation. The REFRESH project was not designed specifically to compare the performance of different model structures, though to some extent this will be done at, at least, four sites. The basis of the comparison will be the annual loads and extreme flows and concentrations.

2.6. Scenario analysis

The output from three Global Circulation Model-Regional Climate Model combinations derived during the ENSEMBLES project will be used: ECHAM5-KNMI, HadRM3P-HadCM3Q0 and SMHIRCA-BCM. ECHAM5-KNMI was chosen because it performed the best during ENSEMBLES climate modelling testing and the modelled projections are close to the ENSEMBLES average. HadRM3P-HadCM3Q0 represents one extreme in the ensemble producing warm, dry summers in comparison with other combinations whilst SMHIRCA-BCM represents the other extreme, being relatively cold and wet. The A1B emission scenario will be used since this was used in ENSEMBLES. By 2050-2060, there is little deviation in climate model output due to emission scenario, hence using a single scenario is acceptable. Monthly precipitation, temperature and potential evaporation data have already been uploaded onto the REFRESH web-site for each of the eight European demonstration sites. These data cover the period from 1960 to 2060. Climate model daily precipitation and temperature data have also been uploaded for each of the eight sites. It was agreed that the climate model control period be 1961-1990. This is the period for which outputs of the climate model have been compared with observations to check model robustness in the ENSEMBLES project.

The baseline period for the catchment flow and water chemistry modelling is 1981-2010 and the scenario period, 2031-2060. In the Description of Work it is unclear if the project is projecting to 2050 or 50 years hence (to 2060). Using the period 2031-2060 covers both 2050 and 2060. The modelled outcomes for the scenario period, 2031 and 2060, will be compared with the baseline period, 1981-2010.

Daily potential evaporation data will be uploaded to the web-site for each demonstration site. For those sites with lakes, then additional climate model outputs will be provided: wind speed, wind direction, etc. However, it should be noted that these should be interpreted with great caution and these data are likely to be unreliable. Maria Shahgedanova from the University of Reading will also provide advice on bias correction of precipitation and temperature data at each of the eight demonstration sites. A note on bias correction has been added to the REFRESH website.

Future land cover was generated for the model scenarios through a local interpretation of the four storylines (global-local, economic-sustainable) as outlined in the Special Report on Emissions Scenarios. An example was presented in the Antalya meeting for the Yläneenjoki catchment. To help interpret the likely future land cover then this should be done, in each demonstration catchment, in conjunction with colleagues from WP6 and stakeholders.

Long term sulphur and nitrogen (ammonium and nitrate; total, wet and dry) deposition scenarios were generated for the eight demonstration sites and selected sub-catchments. The core dataset used to generate the deposition scenarios was provided by EMEP (European Monitoring and Evaluation Programme, http://www.emep.int/), using measured and modelled emissions, meteorological data and a chemical transport model developed at the Meteorological Synthesizing Centre - West (MSC-W). The data is based on observations until 2009, with 2020 scenarios generated in August 2011 based on the Current Legislation Estimate (CLE), Gothenburg Protocol. As such, these represent the most up-to-date scenarios available at present. Forested areas receive higher dry deposition inputs than open areas, so EMEP data was processed at James Hutton Institute, Aberdeen to reflect current and projected land use (according to the land use scenarios) within each demonstration catchment. Phosphorus deposition is assumed to be negligible.

The water use scenarios were derived for each catchment based on expert knowledge of modellers, socio-economists (in Wp6) and stakeholders. These scenarios will be checked for consistency with those from the SCENES Project, but given that SCENES is working at the European scale it is not expected that the scenarios will be consistent between scales though a good reason for the deviation is required. It is expected that there will be a relationship between the water- and land-use scenarios.

The climate, land use, deposition and water use scenarios will be combined to provide four ‘combined SRES scenarios’. The climate scenarios will be run alone through the catchment models first, and then the other four scenarios will be run separately to determine the impact of land cover, deposition and water use change in addition to climate for each SRES scenario.

In summary, fifteen scenarios at most will be run with the numerical models by using 3 different climate simulators (or GCM-RCM combinations). Thus 5 scenarios will be based on each climate simulator output. These are:

- a single SRES emission scenario. This will be the one which represents average conditions (A1B). It is not necessary to run several emission scenarios because the timeframe of the REFRESH project goes up to 2060 only. There is no large difference in the effects on climate of different emission scenarios by 2060.
- four scenarios considering different land cover, deposition and water usage conditions and the climate.

This analysis of the scenario outcomes will include, as a minimum, a comparison of estimated annual loads and the Q5, Q50 and Q95 flows and the C5, C50 and C95 concentrations. These outcomes will be compared between the scenario and baseline periods.

If resources allow, then a sensitivity/robustness/uncertainty analysis will be done whereby each model will be run with sequential changes in precipitation and temperature to investigate a wider range of possible hydrological, hydro-chemical and ecological responses beyond the range of the climate simulations derived from the three GCM-RCMs. The purpose of this exercise is to see if critical precipitation and temperature thresholds exist that will cause major shifts in ecological structure, function and biodiversity.

Again, if resources allow, then a re-run of the uncertainty analysis with the scenario set-ups will be done. Also those aspects of uncertainty not included in the assessment should be noted in deliverables and articles describing the model applications and modelled outcomes.

3. Modelling procedure check list and template for write-up

Table 1 provides a check list for the REFRESH modellers working at the demonstration sites. This check list also provides a template for the write-up of the model applications with the intention that a common format will aid the synthesis of the model applications and simulation outcomes.

Task	Description	Date completed
1	*Model setup*	
1.1	Model input data collated	
1.2	Model geometry defined	
1.3	Observations for testing collated	
2	*Calibration*	
2.1	Manual calibration	
2.2	Auto-calibration[1]	
2.3	Inverse modelling[1]	
3	*Testing*	
3.1	Split sample test done	
3.2	Calibration and test performance assessed objectively	
4	*Sensitivity and robustness testing*	
4.1	Sensitivity analysis[2]	
4.2	Robustness analysis	
5	*Uncertainty analysis*	
6	*Scenario analysis*	
6.1	ECHAM5-KNMI climate scenario	
6.2	ECHAM5-KNMI climate plus four land cover/N deposition/water use scenarios	
6.3	HadRM3-HadCM3Q0 climate scenario	
6.4	HadRM3-HadCM3Q0 climate plus four land cover/N deposition/water use scenarios	
6.5	SMHIRCA-BCM climate scenario	
6.6	SMHIRCA-BCM climate plus four land cover/N deposition/water use scenarios	

Table 1. The template for reporting the application of the REFRESH Common Modelling Framework at each of the eight demonstration sites. The intention is that this template will also serve as a check-list for the model applications. [1]Manual calibration is the minimum requirement.[2]A Sensitivity analysis is not required for INCA-N as this has been done before.

4. Application of the REFRESH Common Modelling Framework

This section describes any major departures from the theoretical ideal of the Common Modelling Framework and summarises any additional work that has been done to support the model developments and applications at each of the eight demonstration sites.

4.1. La Tordera, Spain

4.1.1. The application of multiple models to address structural uncertainty

In the Fuirosos, a sub-catchment of La Tordera in Spain, three models of flow and nitrate and ammonium dynamics were developed and tested to determine if additional model complexity gives a better capability to model the hydrology and nitrogen dynamics of a small Mediterranean forested catchment, or if the additional parameters cause over-fitting (Medici et al. 2012). Three nitrogen-models of varying hydrological complexity (LU4-N, LU4-R-N and SD4-R-N) were considered. For each model, general sensitivity analysis (GSA) and Generalized Likelihood Uncertainty Estimation (GLUE) were applied, each based on 100,000 Monte Carlo simulations. The results highlighted the most complex structure (SD4-R-N) as the most appropriate, providing the best representation of the non-linear patterns observed in the flow and streamwater nitrate concentrations between 1999 and 2003. The 5% and 95% GLUE bounds for the SD4-R-N model, obtained considering a multi-objective approach, provide the narrowest GLUE band for streamwater nitrogen, which suggests increased model robustness, though all models exhibit periods of inconsistent good and poor fits between simulated outcomes and observed data. The SD4-R-N model simulated satisfactorily the first dry hydrological year, which neither the LU4-N nor the LU4-R-N could reproduce. Thus the results confirm the importance of the riparian zone in controlling the short-term (daily) streamwater nitrogen dynamics in the Fuirosos, but not the overall flux of nitrogen from the catchment. The results also show that as the complexity of a hydrological model increases, over-parameterisation may occur, but conversely additional process representation may lead to additional acceptable model simulations. Water quality data help constrain the hydrological representation in process-based models. Increased complexity was justifiable for modelling river-system hydrochemistry.

4.1.2. Simulation of hydrology and water chemistry for the Arbúcies

The Arbúcies is one of the main tributaries of La Tordera (catchment size 112 km^2), and will be the focal point for the hydrochemical and ecological model applications in Spain. Initially, the intention was to use one or several of the N-models developed by Medici et al. 2010, as these address some of the important features of Mediterranean hydrology and nitrogen dynamics (presence of a deep aquifer, presence of a riparian zone, infiltration in the stream bed and sub-surface flow, etc.). However, these models were developed for a small headwater stream, and applying them on a larger catchment with a longer in-stream travel time and significant in-stream processes would require large modifications of the model.

Instead, the hydrochemical modelling will be performed with the new hydrological model PERSiST (Pan-European Runoff Simulator for Solution Transport), which will be fully integrated with INCA-N, enabling a simultaneous calibration of the hydrology and nitrogen dynamics. The model calibration can be tested with both split-in-time and split-in-space data sets, for the latter the data from the nested catchment sampling in the headwater of Font de Regas can be used. The sensitivity and uncertainty analysis can be performed in collaboration with Swedish University of Agricultural Sciences (Martyn Futter). Uncertainty and scenario analyses will be performed according to the outline.

4.2. Lake Beysehir and catchment, Turkey

An eco-hydrological catchment model and a series of lake models will be used to describe effects of climate, land use and water management on hydrology and water quality in the tributaries as well as Lake Beysehir itself. By chaining these models, the potential effects of a series of future scenarios of climate and land use will be simulated for Lake Beysehir.

4.2.1. Simulation of stream flow and nutrient loads to Lake Beysehir

For the Beysehir catchment, the SWAT model (Arnold et al., 1998) will be applied to describe stream flow and nutrient loads of the main tributaries to the lake. Inverse modelling will be used to perform calibration of model parameters through the SUFI2 approach, followed by a global parameter sensitivity analysis also by SUFI2 (Abbaspour et al. 2007). Split-sample test (both in space and time) will be performed, based on the data collected through an intensive monitoring program, including stream flow and nutrient loads of the main tributaries to the lake for a two year period starting in April 2010.

4.2.2. Simulation of ecological dynamics of Lake Beysehir

For Lake Beysehir itself, a series of aquatic ecological models will be applied, including DYRESM-CAEDYM, PCLake and PROTECH (see Mooij et al. 2010 for comparison). DYRESM-CAEDYM will mainly be used to generate thermodynamic input to PCLake and PROTECH (including water temperature at different depths), and the latter two models will be calibrated manually to reflect the observed phytoplankton dynamics in the lake.

4.2.3. Deviation from the modelling ideal for the Beyshir case

While several aquatic ecological models will be applied to Lake Beysehir itself, enabling evaluation of structural uncertainty, additional alternative eco-hydrological models will not be applied to the Beysehir catchment due to the resource constraints of the REFRESH project.

4.3. River Thames and River Kennet, UK

4.3.1. The application of multiple models to address structural uncertainty

In the Thames catchment, the INCA-N and INCA-P models will be applied for the main catchment and in the Thame, Kennet and Lambourn tributaries. INCA-Chalk, a version of INCA-N that includes the vertical movement of nitrate down the unsaturated zone, will also be applied in the River Kennet (Jackson et al., 2007). The Thame contrasts with Lambourn since the former drains clay and the latter chalk. Hence in combination the model applications will test model performance in different geological settings and across a range of catchment areas from 100 to 10 000 km^2. Comparison of model structures and parameter sets will show if the simulation of the long-term transportation of nitrate gives a different result to the simpler INCA-N model, and that the differences in model calibration relate to the different geological settings in a physically reasonable way.

4.4. Vltava, The Czech Republic

A system of catchment and lake models will be used to describe effects of climate, land use, losses from agricultural production, pollution by wastewaters/aquacultures, and water management on hydrology and water quality in the Vltava River network and in several reservoirs that are situated at the headwater tributaries and at the downstream part of the Vltava River. In reservoirs, also their hydrology, hydrodynamics, trophic state, phytoplankton growth, and water quality for water supply will be modelled. The catchment modelling will be done with the INCA-N and INCA-P models with hydrological inputs from the precipitation-runoff model (IHACRES, HSPF). The INCA-simulated water quality in the runoff from catchment will be used as the input to the reservoir ecosystem model (CE-QUAL-W2), together with outputs from several auxiliary empirical sub-models that describe other needed input variables for the reservoir model, i.e., inflow temperature, reservoir operation schemes for different reservoir filling and flow conditions, and concentration of DOM.

4.4.1. Structural uncertainty testing

Structural uncertainty of this system of models will be examined via modelling at different spatial scales (from small subcatchments with homogeneous land use via medium subcatchments with well monitored inputs from natural, municipal, and agricultural sources to large complex catchments) and via comparison with empirical models. The empirical models will include (i) the apportionment of measured nutrient river loads among runoffs from forest, agricultural, and urban areas (export co-efficient approach), municipal wastewater discharges, and atmospheric deposition with considering retention processes in the river network for the catchment nutrient modelling and (ii) empirical retention models (Vollenweider type) for the modelling of nutrients and trophic state in reservoirs.

4.4.2. Deviations from the modelling ideal

There will be essentially no deviations from the REFRESH Common Modelling Framework' at the large-scale level of modelling. At the level of small and medium size catchments the split-in-time model calibration/validation is not done in all cases because the available measured data series are mostly too short (3 years and less).

4.5. Vansjø-Hobøl, Norway

The Vansjø-Hobøl catchment will be modelled at three different spatial scales; a small scale application of INCA-P and INCA-N to the Skuterud catchment (manual calibration); a medium scale application of INCA-P to the river Hobøl (autocalibration methods) and a full scale application of INCA-P linked to MyLake set up for the whole catchment.

4.5.1. Deviations from the modelling ideal for Skuterud

The Skuterud catchment is a well-studied, reference catchment of the Norwegian Agricultural Environmental Monitoring Programme (JOVA). Considering the size of the catchment, which is 4.5 km^2, and benefiting from the detailed information available from the JOVA and from the results of the plot-scale experiments on soil and P losses carried out in the catchment, the INCA model applications for Skuterud will deviate from the modelling ideal in that:

- no sub-catchments will be defined in the modelling procedure, and, consequently, no split-in-space model calibration will be applied;

- an advanced land use specific calibration will be performed by tuning the land-use specific parameters such that the calculated losses would fit the measured values and/or those, derived from expert's estimates;
- the manually calibrated parameter set will further be used as an input and reference background during medium- and full-scale calibration of the INCA-P model for the Vansjø-Hobøl catchment.

4.5.2. Deviation from the modelling ideal for the Hobøl river and the whole catchment.

The main application of INCA-P for the Hobøl river alone as well as the coupling of INCA-P to MyLake will be performed in a Bayesian Framework where complex Markov Chain Monte Carlo procedures are applied. This entails that these applications will deviate from the modelling ideal in that:

- they will not include singular model output; i.e. all relevant output will be in the form of distributions, and therefore a single baseline (or scenario) outcome will not be reported;
- since the outcomes are distributions, compliance issues are not questions of either/or, but to which degree we have confidence in compliance;
- only some of the measures that have been proposed from WP6/WP5 collaboration will be included in the modelling, as some of the mitigation options are not easily implemented in the modelling code.

4.6. Lake Pyhäjärvi and River Yläneenjoki, Finland

The catchment of the Lake Pyhäjärvi will be modelled at different spatial scales; a medium scale application of INCA-N both to the Yläneenjoki and Pyhäjoki catchments, and a full scale application of INCA-P for Yläneenjoki linked to MyLake set up for the whole Lake Pyhäjärvi. LLR probabilistic model will also be used for the lake.

4.6.1. The application of multiple models to address structural uncertainty, together with advanced spatial studies of the lake water quality

Lake Load Response (LLR) is a steady-state probabilistic model incorporating three sub-models. These are a N and P retention model which allows the calculation of lake N and P concentrations given the loads; a hierarchical linear regression model which predicts chlorophyll-a given calculated N and P concentrations; and a logistic regression model which predicts phytoplankton biomass. The LLR model thus converts N and P loads or concentrations from process-based catchment models (INCA-N and INCA-P) into estimates of chlorophyll-a or phytoplankton biomass, both in the present state and in the future climates. LLR is available from web-site *http://lakestate.vyh.fi/cgi-bin/frontpage.cgi?kieli=ENG* and Ahti Lepistö or Olli Malve from SYKE can provide advice.

MyLake is a process-based 1d model for simulation of daily vertical distribution of lake water temperature and thus stratification, and phosphorus-phytoplankton dynamics.

These two lake model applications are compared and supported by spatial water quality studies of the lake (e.g. cyanobacteria pigments and chl-*a*). Two relatively new research methods have been used to assess the spatial water quality components of Pyhäjärvi: i) transect measurements from a moving boat and ii) remote sensing data based estimates (Lepistö et al., 2010).

4.6.2. Deviation from the modelling ideal for Yläneenjoki/Pyhäjärvi site

There is a plan to test the output from a Global Circulation Model-Regional Climate Model scenario ECHAM5-KNMI, using two different bias correction methods. The focus in the CC scenario work will be in ECHAM5-KNMI scenario (with extremes only if resources allow), due to a complex model chain of WSFS model - INCA-P model – MyLake model.

4.7. River Dee, UK

4.7.1. Application of multiple models

In the Dee catchment the STREAM-N model will be applied in addition to INCA-N for at least one sub-catchment (Tarland Burn). STREAM-N is a semi-distributed hydrology and nitrate model that captures the characteristics of land use and physical structure of catchments by carrying out calculations of the water and N balance on a relatively fine grid (100m resolution in this example). Calculated flows and N fluxes from the grid network are routed to the stream network by linking rate parameters with topographic data (Dunn et al., 2007). STREAM-N has an integral hydrological model, but makes simplifications of the N cycle compared with the representation in INCA-N. A sensitivity analysis of STREAM-N will be undertaken on the Tarland Burn sub-catchment of the Dee. Calibration of STREAM-N will be based on interpretation of MC simulations varying parameter values across broad ranges in random combinations, and using a suite of objective functions to assess goodness of fit. Simulations from STREAM-N will be cross-compared with INCA-N to evaluate the significance of model structure in controlling the results of the model analyses.

In the upland sub-catchment of the River Gairn the Model for Acidification of Groundwater in Catchments (MAGIC) will be applied to assess the dynamics of C and N in the system. Contrary to models applied at other sites, MAGIC runs on a monthly time-step as it is concerned with evaluating long-term changes in biogeochemistry linked to climatic conditions and atmospheric deposition. The various scenarios for atmospheric deposition will be assessed in greater depth within the River Gairn application, as the simulated stream response is much more likely to be sensitive to the changes in this upland situation.

At the scale of the whole river basin the Nitrogen Risk Assessment Model for Scotland (NIRAMS) will be applied. NIRAMS operates in a similar manner to STREAM-N but using data at a coarser spatial resolution.

No scenarios for water resources will be applied to the River Dee. Although the main stem of the river is used for public water supply abstraction, the remainder of the catchment is largely unmanaged in relation to water resources.

4.8. River Louros, Greece

For the Louros, there are two major obstacles to successfully apply the outlined model framework. First, the hydrology in the area is very complex. Most of the bedrock in the area consists of permeable and fractured karstic rock, underlain by an impermeable layer, often with topography different from the surface. Consequently, the catchment boundaries cannot be drawn from the topography, and may at best be approximated from a geological map. Fractures and fissures in the bedrock also make the water residence time unpredictable, and both belowground streams and springs are abundant in the area. Furthermore, a large part of the lowland area is connected to the estuary by drainage channels controlled by water gates, and the direction of flow may differ between seasons in these areas. The second obstacle is that the only available flow record is the outflow of the hydroelectric dam. This flow is in part regulated by the number of turbines operated by the power plant (i.e. by electricity demand), and thus, there are no available data describing the natural runoff. Chemical observations (N and P) are scarce, but still provide some data for model calibration. Information on agricultural practises, irrigation and fertiliser application rates are however excellent. Given the limited hydrological data, the aim of the PERSiST hydrological modelling will be to give the correct water balance (the runoff coefficient is known). For the chemical modelling (INCA-N and INCA-P), the aim will be to broadly simulate mean N- and P-concentrations, to relate fertiliser application rates to in-stream concentrations.

As the temporal dynamics of nutrient concentrations is not the focus, a split-time model test is not meaningful. There could potentially be an opportunity for a split-space model testing, by setting up the model for the neighbouring river of Acheron.

A limited sensitivity/uncertainty analysis can be performed, based on the parameters that control the long term mean nutrient concentrations. Scenario analyses can be performed as outlined in section 2.6.

5. Integration between WP5 (modelling) and WP2-3-4 (field-based studies)

5.1. Demonstration sites

To help assess the impacts of environmental change on ecological structure and function and biodiversity, the outputs from the catchment hydrology and water quality models will be used with site specific flow and/or water chemistry relationships to determine the response. Partners working on each site will refine the indicator list (i.e. select the main indicator) and search for site specific relationships between that indicator and the physical and chemical environment at each demonstration site. To achieve this, buddy ecologists will be identified for the Thames and Dee work. At other demonstration sites, scientists working in WP2-4 will collaborate with their colleagues in WP5. Those working on the REFRESH reviews, in work packages 2, 3 and 4, are asked to focus on the indicators to be used in the WP5 assessments in the first instance. The typology of each of the eight demonstration sites will be identified, and this will be used to help determine the functional relationships between stressors and ecological indicators as part of the reviews done in work packages 2, 3 and 4.

It should be possible to link some of the REFRESH lake experiments done in WP3 by using the results to help verify the response of lake models to changes in temperature and nutrient loading. Correspondingly, also lake water quality measurements made by WP3 (such as Chl-a) will be utilised to test lake model approaches where the experimental and demonstration sites are the same, such as Lake Pyhäjärvi. For other demonstration sites where experiments are not carried out, such as Lake Vansjø, analysis of existing data in WP3 (Task 3-4) will contribute to modelling ecological impacts by comparing the modelled responses in terms of altered water levels and changes in nutrient concentrations on the ecosystem response with those observed.

Useful information on ecological indicators and functional relationships can be found at http://habitat.deltares.nl. This gives access to the HABITAT tool which is free to use for REFRESH members. Contact Valesca Harezlak at Deltares for advice. The site www.freshwaterecology.info developed by BOKU during Euro-limpacs contains a vast amount of species-level information of ecological traits for macroinvertebrates, fish and diatoms. WP2-4 intend to extend this database to cover riparian species, covering initially beetles, dragonflies and damselflies as part of their review work in REFRESH. Furthermore, the webpage http://www.climate-and-freshwater.info/ (also a product of Euro-limpacs) should be updated with results from the literature reviews in WP2-4.

5.2. Exploratory model developments

The process-based models used in work package 5 will be developed to improve the simulation of coupled river-wetland-lake systems. These models will not be developed and applied at all eight demonstration sites, but instead will make use of the opportunity within the project to build on existing ecological and catchment-scale modelling efforts to take the science forward. In REFRESH, wetlands are defined as the stream-riparian-gradient (SRG), namely the area (laterally) alongside the stream that is limited on the stream end by the lowest summer water table and on the upland end by the highest winter water table (based on the high water mark of rare, once per hundred years, floods).

5.2.1. Wetlands process-based modelling

The Centre for Ecology and Hydrology will continue to test how extensively relationships between wetland plant communities and water table depth, derived for the UK, can be extended across Europe. As a step towards incorporating wetlands in the process-based models, a simple hydrological and nitrogen model will be developed for the Boxford wetland on the River Lambourn to describe bi-directional flow and nitrogen exchange between the river and the riparian wetland.

5.2.2. Enhancements to the INCA models

A new hydrological model, PERSiST will be created. This development is needed to provide estimates for hydrologically effective rainfall (HER) and soil moisture deficit (SMD) for use in the INCA suite of models. PERSiST will be integrated with an enhanced version of INCA-N to simulate the storage and transport of water, nitrate and ammonium in branched river-systems. To date, INCA-N has only been applicable to the main channel of a river catchment. The structure of the PERSiST model will be flexible and able to incorporate a zone representing the wetland interface between land and river. It is intended that the riparian model developed by CEH shall be incorporated in the INCA-N model.

The combined PERSiST and INCA-N model will be tested using data collected in the La-Tordera-Font del Regas sub-catchment. The combined model incorporates a riparian zone. The nitrogen mass-balance is being measured in the Font del Regas sub-catchment. In particular, the nitrogen dynamics are being measured in three riparian zones with different vegetation covers and these are concurrent with streamwater, litterfall, tree biomass and soil nitrogen measurements.

A recently developed version of INCA-P will also be further tested and refined. This model was developed as part of FP6 Euro-limpacs project but will be enhanced so that it is applicable to branched river-systems.

5.2.3. The REFRESH models

WP5, together with ecologists from WP2, 3 and 4 will develop three empirical models to describe multi-stressor relationships using Bayesian Belief Networks:

- river macroinvertebrate model;
- riparian plant community model;
- river macrophyte-algal model.

These three models will be based on Bayesian Belief Networks (BBNs) and will be used to articulate expert beliefs about the dependencies between different variables and derive the implications of these beliefs in terms of a multi-stressor impact on key aspects of the ecological structure, function and biodiversity. The modelled outcomes will be tested against the results of the field-based studies in WPs 2-4 where possible, and the results used to update the probabilities and structures of the BBNs. The BBNs will make use of the WP2-4 reviews. The BBNs will also be used to make probabilistic predictions of the effects of environmental change. In the case of the river algal model, the new INCA model simulation results will be fed directly into the empirical model to derive the impact on flow and nutrient conditions on algal biomass. For the river macroinvertebrate and riparian plant community models, the outcomes of the process-based simulations will be used to define relationships in the BBNs.

A series of meetings was held at the University of Reading between the 28 May and 17 July 2012 to integrate the expert knowledge into the BBNs. Project ecologists worked with the WP5 modellers to develop the functional relationships needed in the empirical models.

A version of the macrophyte-algal dominance model will be applied to the Louros catchment. As the Louros data set is focused on macrophyte biodiversity, the first step will be to simulate the macrophyte communities or dominant macrophyte species.

Macroinvertebrate data has been sampled for the Arbúcies as well as other sites in the Tordera catchment. Macroinvertebrate observations are also available from the experimental flooding and drying experiment site in the Arbúcies. The University of Barcelona (Joan Riera) will process the macroinvertebrate data. The macroinvertebrate REFRESH model developed by Richard Skeffington, Piet Verdonshot and Anna Besse- Lototskaya for a Dutch stream will used as a starting point, and modified for local conditions.

6. Integration between WP5 (modelling) and WP6 (socio-economics)

WP6 will identify possible mitigation actions and in collaboration with biophysical scientists from WP5 will evaluate their effectiveness based on the study site characteristic, model analysis and literature values. These effectiveness values will be used in numerical models to test their effect on the hydro-chemical and ecological conditions of the study area in question. This allows the identification of a possible change in status and thus the likelihood of compliance with the EU Water Framework Directive as a result of the mitigation measure. The simulation results and EU WFD compliance information will be fed back to WP6 partners, who will calculate the cost-effectiveness of the measures and will conclude whether these mitigation measures are proof against climate change.

The following steps were agreed between WP5 and 6 for each of the six common sites (Dee, Thames, Pyhäjärvi/Yläneenjoki, Vansjø-Hobøl, Vlatava, Louros):

Step 1: WP6 recognizes non-compliance issues and defines problems through utilizing monitoring data and consulting national experts and catchment-specific stakeholders

Step 2: Run the WP5 models to calibrate and test and provide baseline modelled flow and hydrochemistry conditions.

Step 3: WP6 identify a series of mitigation measures (options) per non-compliance issue, also through consultation with stakeholders.

Step 4: WP6 provide WP5 with information (e.g. agronomic data through consulting national experts) outlining the implications of each mitigation measure for diffuse pollution

Step 5: WP5 run measures, identified and provided with initial information from WP6, through the catchment-scale models to check compliance with the Water Framework Directive, in terms of ecological status, and determine the efficiency of the measures at the catchment scale. Compliance with other directives and guidelines can also be checked, such as impacts on Natura 2000 sites notified under the Habitats Directive.

Step 6: WP6 assesses the cost effectiveness of mitigation measures which achieve compliance.

Step 7: Run the scenarios through the catchment-scale models without any measures. This is to be done in a stepwise way, with the climate scenarios done first, and then with each of the four combined climate-land cover-deposition-water use scenarios (to give a total of five scenario outcomes per climate scenario). Compliance with the WFD, HD and impact on Natura 2000 sites will be checked using those ecological indicators identified as relevant for the catchment.

Step 8: Run the models to include the proposed measures found to be effective at step 5 in each of the scenario runs. Compliance with the WFD, HD and impact on Natura 2000 sites will be checked, as will the catchment-scale efficiency of the measures. Comparison of the results from steps 5 and 8 will determine if the measures, effective at step 5, will be more or less effective under climate change.

Colleagues in WP5 and 6, responsible for delivery of each step in the six demonstration sites, have been identified and progress will be monitored by the WP5 and 6 co-leaders.

7. References

Abbaspour, K. C.; Yang, J.; Maximov, I.; Siber, R.; Bogner, K.; Mieleitner, J.; Zobrist, J.; Srinivasan, R. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. Journal of Hydrology, 333, (2), 413-430.

Aitken A P. 1973. Assessing systematic errors in rainfall-runoff models. Journal of Hydrology 20:121-136.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., and Williams, J.R. 1998. Large area hydrologic modeling and assessment part I: model development. Journal of the American Water Resources Association 34(1), 73-89.

Beven KJ, Freer J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. Journal of Hydrology 249: 11–29.

Birkel C, Tetzlaff D, Dunn SM, Soulsby C. Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models. Water Resources Research 47, W02515, doi: 10.1029/2010WR009547.

Dean S, Freer J, Beven K, Wade AJ, Butterfield D. 2009. Uncertainty assessment of a process-based integrated catchment model of phosphorus. Stochastic Environmental Research and Risk Assessment 23(7): 991-1010.

Jackson BM, Wheater HS, Wade AJ, Butterfield D, Mathias SA, Ireson AM, Butler AP, McIntyre N, Whitehead PG. 2007. Catchment-scale modelling of flow and nutrient transport in the Chalk unsaturated zone. Ecological Modelling, 209, 41-52.

Jolliff JK, Kindle JC, Shulam I, Penta B, Friedrichs MAM, Marjorie A, Helber R, Arnone RA. 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment Journal of Marine Systems, 76, 64-82.

Lepistö, A. Huttula, T., Koponen, S. Kallio, K. Lindfors, A., Tarvainen, M. & Sarvala, J. 2010. Monitoring of spatial water quality in lakes by remote sensing and transect measurements. *Aquatic Ecosystem Health & Management* 13(2): 176-184.
http://www.informaworld.com/smpp/content~db=all~content=a922936446~frm=abslink

Mahmoud M, Liu Y, Hartmann H, Stewart S, Wagener T, Semmens D, Stewart R, Gupta H, Dominguez D, Dominguez F, Hulse D, Letcher R, Rashleigh B, Smith C, Street R, Ticehurst J, Twery M, van Delden H, Waldick R, White D and Winter L (2009). "A formal framework for scenario development in support of environmental decision-making." Environmental Modelling & Software 24(7): 798-808.

McIntyre N, Jackson B, Wade AJ, Butterfield D, Wheater HS, 2005. Sensitivity analysis of a catchment-scale nitrogen model. Journal of Hydrology, 315, 71-92.

Medici C, Bernal S, Butturini A, Sabater F, Martin M, Wade AJ, Francés F, 2010. Modelling the inorganic nitrogen behaviour in a small Mediterranean forested catchment, Fuirosos (Catalonia). Hydrol. Earth Syst. Sci. 14, 1–15.

Medici C, Wade AJ, Francés F. 2012. Does increased hydrochemical model complexity decrease robustness? Journal of Hydrology, 440-441, 1-13.

Mooij WM, Trolle D, Jeppesen E, Arhonditsis G, Belolipetsky PV, Chitamwebwa DBR, Degermendzhy AG, DeAngelis DL, De Senerpont Domis LN, Downing AS. 2010. Challenges and opportunities for integrating lake ecosystem modelling approaches. Aquatic Ecology, 44(3), 633-667.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models. Part 1 – a discussion of principles. Journal of Hydrology, 10, 282-290.

Oreskes N, Shrade-Frechette K, Belitz K. 1994. Verification, Validation and Confirmation of Numerical Models in the Earth Sciences. Science 263: 641-646.

Rankinen, K., Lepistö, A. & Granlund, K. 2004. Integrated nitrogen and flow modelling (INCA) in a boreal river basin dominated by forestry: Scenarios of environmental change. *Water, Air & Soil Pollution*, *Focus* **4**: 161-174.

Rankinen K, Karvonen T, Butterfield D. 2006. An application of the GLUE methodology for estimating the parameters of the INCA-N model. Science of the Total Environment 265: 123-139.

Refsgaard JC, van der Sluija JP, Lajer Højberg A, Vanrolleghem PA. 2007. Uncertainty in the environmental modelling process – a framework and guidance. Environmental Modelling and Software 22:1543-1556.

Silgram M, Schoumans O.F. (eds.) 2004. Modelling Approaches: Model Parameterisation, Calibration and Performance Assessment Methods in the EUROHARP Project. EUROHARP report 8-2004, NIVA report SNO 4740-2003, ISBN 82-577-4491-3, Oslo, Norway, 18pp. http://www.wise-rtd.info/sites/default/files/d-2007-09-04-Eurharp_report8.pdf. Last accessed 5 July 2012.

Shahsavani D, Grimvall A. 2009. An adaptive design and interpolation technique for extracting highly nonlinear response surfaces from deterministic models. Reliability Engineering and System Safety 94 (7):1173-1182.

Shahsavani D, Grimvall A. 2011. Variance-based sensitivity analysis of model outputs using surrogate models. Environmental Modelling and Software 26:723-730.

Spear R. 1970. The application of Kolmogorov Renyi statistics to problems of parameter uncertainty in system design, Int. J. Control 11: 771–778.

Spear R, Hornberger GM. 1980. Eutrophication in Peel Inlet II Identification of critical uncertainties via generalized sensitivity analysis, Water Res. 14: 43–49

STOWA, 1999. Smooth Modelling in Water Management, Good Modelling Practice Handbook, STOWA report 99-05, Dutch Department of Public Works, Institute for Inland Water Management and Waste Water Treatment report 99.036, ISBN 90-5773-056-1.

Vrugt JA, ter Braak CJF, Gupta HV, Robinson BA. 2008. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modelling? Stochasitc Environmental Research and Risk Assessment, 23(7), 1011-1026.

Vrugt JA, ter Braak CJF, Diks CGH, Higdon D, Robinson BA, Hyman JM. 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. International Journal of Nonlinear Sciences and Numerical Simulation, 10(3), 273-209.

Wade AJ, Hornberger GM, Whitehead PG, Jarvie HP, Flynn N. 2001a. On modelling the mechanisms that control instream phosphorus and macrophyte dynamics: an assessment of a new model using General Sensitivity Analysis. Water Resources Research 37, 2777-2792.

Wade AJ, Whitehead PG, Butterfield D, 2002a. The Integrated Catchments model of Phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogeneous river systems: model structure and equations. Hydrology and Earth System Sciences, 6, 583-606.

Wade AJ, Jackson BM, Butterfield D. 2008. Over-parameterised, uncertain 'mathematical marionettes' – how can we best use catchment water quality models? An example of an 80 year catchment-scale nutrient balance. Science of the Total Environment 400(1-3): 52-74, doi: 10.1016/j.scitotenv.2008.04.030.

Whitehead PG, Wilson EG, Butterfield D. 1998. A semi-distributed Nitrogen Model for Multiple Source Assessments in Catchments (INCA): Part 1 - Model structure and Process Equations. Sci. Total Environ. 201: 547-558.