

1 An approach for coupling higher and lower levels in marine ecosystem models and its application to
2 the North Sea

3 SUGGESTED NEW TITLE: *Couplerlib*: A Metadata driven library for the integration of multiple models
4 of higher and lower trophic level marine systems with inexact functional group matching

5
6 Jonathan Beecham*[!], Jorn Bruggeman⁺, John Aldridge*, Steven Mackinson*

7 *Cefas Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk, NR33 0HT, United Kingdom.

8 ⁺ Bolding & Burchard ApS, Strandgyden 25, DK-5466 Asperup, Denmark.

9 [!]Corresponding Author – Jonathan.Beecham@cefas.co.uk, Telephone +44 1502 524541, Fax +44
10 1502 513865

11
12 Abstract

13 End to end modelling is a rapidly developing strategy for modelling in marine systems science and
14 management. However problems remain in the area of data matching and sub-model compatibility.

15 A mechanism and novel interfacing system (Couplerlib) is presented whereby a physical/
16 biogeochemical model (GOTM-ERSEM) that predicts dynamics of the lower trophic level (LTL)
17 organisms in marine ecosystems is coupled to a dynamic ecosystem model (Ecosim) that predicts
18 food-web interactions among higher trophic level (HTL) organisms. Coupling is achieved by means of
19 a bespoke interface which handles the system incompatibilities between the models and a more
20 generic Couplerlib library which uses metadata descriptions in extensible mark-up language (XML) to
21 marshal data between groups, paying attention to functional group mappings and compatibility of
22 units between models. In addition, within Couplerlib, models can be coupled across networks by
23 means of socket mechanisms.

24 As a demonstration of this approach, a food web model (Ecopath with Ecosim, EwE) and a
25 physical/biogeochemical model (GOTM-ERSEM) representing the North Sea ecosystem were joined
26 with Couplerlib. The output from GOTM-ERSEM varies between years dependent on oceanographic

1 and meteorological conditions. Although inter-annual variability was clearly present, there was
2 always the tendency for an annual cycle consisting of a peak of diatoms in spring, followed by (less
3 nutritious) flagellates and dinoflagellates through the summer resulting in an early summer peak in
4 the mesozooplankton biomass. Pelagic productivity, predicted by the LTL model, was highly seasonal
5 with little winter food for the higher trophic levels. The Ecosim model was originally based on the
6 assumption of constant annual inputs of energy and, consequently, when coupled, pelagic species
7 suffered population loss over the winter months. By contrast, benthic populations were more stable
8 (although the benthic linkage modelled was purely at the detritus level so this stability reflects the
9 stability of the Ecosim model). The coupled model was used to examine long term effects of
10 environmental change, and showed the system to be nutrient limited and relatively unaffected by
11 forecast climate change, especially in the benthos. The stability of an Ecosim formulation for large
12 higher trophic level food webs is discussed and it is concluded that this kind of coupled model
13 formulation is better for examining the effects of long term environmental change than short term
14 perturbations.

15 Highlights

16 A generic coupler was developed for linking higher and lower trophic level models.

17 A linked GOTM-ERSEM / EwE model was developed using calibration from two North Sea
18 Ecosystem studies.

19 The EwE model was modified to take into account the seasonal dynamics of the LTL model.

20 The spring diatom bloom and hence silicate availability was a critical driving factor in regulating
21 HTL dynamics.

22 Examination of the long term consequences of the model under climate change forecasts
23 revealed a relatively small change using the EwE food web model.

1 The couplerlib approach provides a means of decoupling ecological and geochemical model
2 construction and integration from the software development process, provides data consistency
3 checking and conversion of units and functional group assignment.

4 Keywords

5

6 End to End Modelling, Food webs, Model Coupling, Metadata, Climate Change, North Sea

7

1. Introduction

End to end modelling is becoming a hot topic in marine systems (Rose *et al.* 2010) primarily because of the need, implied by regulatory frameworks such as the Marine Strategy Framework Directive (MSFD) (http://ec.europa.eu/environment/marine/index_en.htm) for monitoring and management of the marine systems to take into account distal effects of any deliberate or accidental anthropogenic changes to parts of the marine ecosystem, and modelling is used to predict how indicators, such as critical species biomasses might relate to ecosystem change. To this end a number of international projects such as Meece (www.meece.eu) have been set up to produce integrative modelling systems in order to assess the likely impact of anthropogenic change. However, the problem which remains is that the models that are used in an end to end modelling system are themselves complicated with many parameters. Joining such models end to end has been described as ‘Putting lipstick on a pig’ (Rose 2012), and an appeal is sometimes made to start with simpler models (Fulton *et al.* 2003).

Given the choice between constructing end to end model systems *de novo* and joining existing models (usually written by different teams, often in different languages or even running on separate machines) together ‘Frankenstein Style’ or even combining multiple models at source level, the former has a number of obvious advantages. The multiple models (or multiple components of an integrated modelling system) can benefit from a unified design process, can use consistent co-ordinate system and model entity representation and can benefit from unified data input, output, visualisation, control and validation. Most importantly, the models can be aligned in terms of whether they are strategic models for exploring general principles, or tactical models designed to explore specific aspects of the system with a high degree of detail and domain specific knowledge (Levins 1966). Joining models with widely differing focus will typically result in a combined model

1 with only broad results for a narrow domain, and yet will still require a large amount of data to
2 calibrate it. On the other hand, we are often forced to work with combining existing model because
3 of the existence of code, with updates and a user community, the investment of time in the lengthy
4 processes of calibrating and validating these models. Combining models by formally combining code
5 can improve the reliability and intelligibility of the model and make validation easier, but combines
6 the disadvantage of unsympathetic algorithms having to work together with an inability to easily
7 incorporate code updates to the original sources. For most modelling projects sociological aspects
8 of the modelling process will have an influence on the strategy of end to end modelling employed.

9 In many cases the objective of the end to end modelling approach includes a high degree of
10 prediction for example it is one of the aims of the EU Framework programme 7 project Meece
11 (www.meece.eu). However although it is possible to model phytoplankton blooms and short term
12 fish stocks management fairly reliably, there are a number of less tractable problems, including
13 estimating the recruitment of young fish to the adult stock. It is a difficult function to quantify in
14 models because it is dependent not only on the size of the parent stock, but also on interactions of
15 the lower trophic level components of marine ecosystems, notably zooplankton, that are poorly
16 understood and which are highly dependent on physical driving factors such as temperature,
17 irradiation and wind (Cushing 1996). Zooplankton modelling has always been the weakest link in the
18 end to end chain and how such processes govern fish recruitment and population variability is still
19 poorly understood (Minto et al., 2008).

20 Coupled physical-biogeochemical models such as GOTM-ERSEM (Burchard et al. 1999; Burchard et
21 al., 2006; Baretta et al., 1997; Vichi et al. 2003) and POLCOMS-ERSEM (Lewis and Allen 2009;
22 Blackford et al. 2004) that predict changes in primary production and zooplankton abundance as
23 outcomes of hydrodynamic and biogeochemical processes, are similarly not without limitation, an

important one being their inability to capture top-down trophic impacts from higher grazers such as fish, which are not explicitly included.

Although lower trophic level components can quite easily be represented (albeit simply) in multi-species fisheries models such as Ecopath with Ecosim (EwE) (Christensen et al. 2005), representation of how environmental drivers influence the dynamics of higher trophic species is more challenging. In EwE, this is usually achieved using simple production and consumption forcing patterns, which gloss over biogeochemical processes such as how the dynamics of the different nutrient components determine the limitations of phytoplankton growth. With the knowledge that environmental changes can have drastic impacts on lower trophic level components such as phytoplankton (Richardson and Schoeman, 2004) and that fishing has strong impacts on the abundance and structure of fish stocks (e.g. Lotze et al., 2006), investigating how ecosystems respond to combined pressures necessitates that processes at lower and higher levels are linked. This is the basis of the present trend towards end-to-end modelling, whereby the connection of physics to fish and fisheries is made (Travers et al. 2007; Libralato and Solidoro 2009; Rose et al. 2010). The end-to-end modelling of Kearney (2012) links the NEMURO lower trophic level model with a 23 component EwE model, achieves ecosystem stability, intra and inter annual variation in biomass of higher and lower trophic level species, though it does not encapsulate all the processes that lead to variation in fish populations, especially recruitment variation. However it does achieve a class leading degree of correspondence with observations.

Whilst the realisation of effective end to end modelling systems for prediction will depend on the development of precise sub-models for some of these critical components, their development will also depend upon consistent systems for model integration and data exchange. Coupled LTL-HTL model systems should allow for iterative exchange of data among models, thus capturing important feedback processes within an ecosystem. Thus they should enable us to investigate ecological issues,

such as how changes in phytoplankton abundance in shelf seas might be attributed simultaneously to abiotic drivers such as nutrient inputs and changes to summer stratification or a reduction of predation on zooplankton by fish. Such complex interactions can only be understood through an end-to-end approach that combines physics, lower trophic levels and higher trophic levels.

The principal challenges of coupling LTL models with HTL models include reconciling differences in how the models handle and represent important processes at different time and spatial scales (Rose et al. 2010). Whilst it is vitally important to test and understand how model behaviour is influenced by choices in the level of detail (one or many groups), the groups that are represented (size, age, bulk biomass) and the choice of time and spatial scales of processes (e.g. capturing seasonality), there is also a need to be pragmatic. This means making choices that enable the models to be coupled and tested, either through comparison with empirical data or by their ability to generate plausible, testable hypothesis that are consistent with understanding.

The paper describes a methodology for coupling LTL and HTL models, which has been developed by the authors. This 'Couplerlib' approach is generic, allowing for exchange of information between separate LTL and HTL models. In effect Couplerlib is a glue layer between different models, consisting of a library of routines for checking data consistency (in terms of names of groups, chemical elements and units) between models, carrying out data conversions and providing network protocols. It is not a universal coupler, which would be very difficult to achieve given the potentially enormous range of languages and calling conventions that models may use. Instead, the user is required to provide a compatible interface for use with Couplerlib this being the 'front end' used to input required information to setup and perform a model run.

In the specific example presented here, Couplerlib connects two models representing the dynamics of the North Sea ecosystem; GOTM-ERSEM (Burchard et al., 2006), a physical-biogeochemical model of the LTLs, and the food-web EwE model (Mackinson and Daskalov 2007) of the HTLs. The two

models are downloadable at www.gotm.net and www.ecopath.org respectively. In this case the EwE front end serves as the primary interface for Couplerlib, through which a graphical front-end for GOTM-ERSEM is called for specification of the LTL model run parameters.

To facilitate a clear understanding and promote further development of the methodological approaches, we describe the technical process of linking EwE to GOTM-ERSEM. Specific problems and solutions to overcoming them are discussed.

2. Methodology

This section describes the design principles and system methodology, when applied to linking a biogeochemical model, ERSEM, to a food web model Ecosim. The scope of Couplerlib is wider than this because it is a generic coupler which uses metadata to describe the model linkages. Information about source code and implementation is given in Appendix A.

2.1 Principles of Model Coupling

Couplerlib, in association with suitable calling routines, can be used as part of a model coupling in a number of different ways::

- a) Direct Coupling (used to couple ERSEM and the GETM physics model) : Coupled models share a memory space and are able to read and write directly into it. This requires the models to have the same way of storing basic data information for integer and floating point variables, which can present a number of problems. C / C++ and FORTRAN, for example have different array storage (row major and column major) which may require a glue layer to convert data order. Coupling between data that is stored alternately between models will require some form of indirection (i.e. pointers or arrays indexed via

an index variable). Also Visual Basic .net as used by EwE uses references to managed objects to store array and class information and Native Fortran and C/C++ used by the LTL models use simple pointers to a block of memory. Direct coupling would require a pointer to be passed to a reference which is not allowed in .net languages, so a conversion layer must be provided. Furthermore work direct coupling even using a single language not allow any form of data conversion such as between units, grid resolutions or where functional groups are merged between models, so can be used only when the data is in the same format in the different models (the case for GOTM and ERSEM by not EwE and ERSEM). Despite these limitations it is an attractive form of coupling because of its speed and simplicity, for example there is direct linking between the GOTM and ERSEM models in this example. Couplerlib can be used for checking data and providing diagnostic information during its initiation but not used for passing data between models.

b) Managed Coupling (used for the examples here) . Managed Coupling uses a management layer between two models which is capable of data conversion, data validation, reorganisation, splitting and merging of data. It is possible to couple models with very different internal representations by incorporating the necessary glue layers within the management layer. The ERSEM to EwE coupling described here uses managed coupling. Typically as well and data being transferred between models via the management layer, there is some form of control data which can specify how the data exchange is controlled. The Couplerlib system presented here is such a management layer, which stores control information via a system of XML files (Fig. 1).

c) Networked Coupling (used experimentally in development). This is an extension of managed coupling which can operate between processes and across the web. In order

to achieve this, there need to be two copies of couplerlib with communication processes between each. Typically one of the processes is the master with capabilities for operator interaction and the rest are slaves operating under a system of messages between them. The Atlantis system (atlantis.cmar.csiro.au, Fulton 2010) also implements a networked coupling system between network components but is both data and control central, whereas Couplerlib is control central and data distributed. Networked coupling would be needed to link an ERSEM model running on a Linux cluster to eWe on a standalone PC. Networked coupling can be used on the same machine where it is virtualised running Linux and windows or across processes and between processes on the same machine. Couplerlib is typically run from single or synchronous threading so the results of multiple parallel threads are brought together before conversion to another model. This is seldom a bottleneck where HTL models, specifically Ecosim are concerned because they typically have a long timestep of over a day, compared to a few minutes in LTL models.

d) Offline Coupling (used in subsequent development of 3D models) uses some form of file to store data so that multiple models can run, store data and allow a second model to process this model independently. Because of the slow nature of file reading Couplerlib implements offline coupling using netCDF files as a one-way only coupling. Whilst limited in not being bidirectional, it is often used for short term models with limited effects of HTL components on the nutrient levels.

Couplerlib can be used to link models one way, in which case the values of quantities in the source model overwrite the respective quantities in the destination model, or bidirectionally (not for offline coupling) in which case quantities are passed back from the destination model and re-overwrite the values in the source model. In this way calculation is shared between models and double accounting needs to be avoided, by for example allowing

consumption and not production in the destination model and consumption only in the source model.

2.2 Couplerlib Design and Specifications

A critical feature of Couplerlib is its ability to specify multiple models and interfaces between them and perform validation on these interfaces, only validated interfaces may link and exchange data. The principle behind this is that multiple coupling mechanisms may be provided, only some of which may be allowed in each situation. An example would be the provision of data in multiple types of grid and non-grid spatial format, which might require a conversion layer between the models, by specifying multiple models we can allow code for multiple conversion libraries to be incorporated within a model.

Existing approaches that have linked lower and higher trophic levels using single species and individual based approaches (for example Rose et al. 2007; Travers et al. 2009), define the linkages before the simulation is run. However, in connecting ERSEM to Ecosim we are confronted by the problem that the constituent components of Ecosim models are explicitly specified at runtime and ERSEM output is specified by means of a user invoked compiler. In this case it is necessary to develop a coupling mechanism where metadata information that is generated at runtime can still be cross-matched and validated. The mechanism as shown in fig. 1 is for the user to specify a linkage in terms of models, variables and quantities (metadata) in xml which is parsed by the Couplerlib system to build the linkages between models.

For linking to EwE, which is Windows only, managed coupling is only possible within a windows environment. The data store in Couplerlib is global and persistent, and is referenced by an enumeration (i.e. a list of items which is converted to a numeric sequence) of model and interface.

1 In principle a single Couplerlib can be used to link multiple models with multiple connections for
2 different types of data between them, however this was restricted to two (intake and offtake) for
3 this work. However, care must be taken to ensure that a frame of data (data for single model and
4 interface combination) is only read after it is completely written; otherwise the data may be
5 corrupted. Because programming libraries for multithreaded systems have various ways to signal
6 events and suspend program execution on a thread, these mechanisms need to be provided by the
7 interface to Couplerlib. The strength of Couplerlib is that once this interface code is written it can be
8 used over and over again for different Ecosim models. This threaded mode of interfacing also allows the
9 graphical interface of both models to run simultaneously.

10 When Couplerlib operates in networked mode, each process or machine is required to have its own
11 Couplerlib. The networked Couplerlibs are initially synchronised by reading metadata descriptions of
12 all the models being linked from files that will be available on all machines as a Uniform Resource
13 Identifier (Berners-Lee et al. 2005). At run time the networked Couplerlibs are synchronised on a
14 need-to-know, object request basis (the receiving Couplerlib specifies what data the transmitting
15 Couplerlib needs to send) even though validation has been carried out during the initial model
16 loading phase. Data from a model is written to the local copy of Couplerlib and the remote copy
17 synchronises its own copy when an instruction is received to fetch data. In addition to transferring
18 data, the networked Couplerlib exchanges diagnostics and model output to be displayed remotely
19 and controls the synchronisation of data.

20 To ensure that coupled models are synchronised, Couplerlib uses the Berkeley Sockets Mechanism,
21 specifically blocking sockets. Sockets form the basis of the internet protocol and are used to ensure
22 that networked services are connected from right source to destination (i.e. two ends that have the
23 same socket number) and can be assigned a number by the user or by the operating system. By
24 assigning each a station number, multiple models can connect using networked Couplerlibs.

2.3 Metadata information exchange and specification

Couplerlib is a coupling system which uses a metadata description of the model linkage, combined with emitted metadata from the component models and a dictionary of functional group and unit definitions to form a linkage between models. It allows linkages to be described in terms of a functional level of abstraction, without the user needing to understand computer representation (such as indexing of arrays) and in many cases without altering code.

Both EwE and GOTM-ERSEM are reconfigurable modelling frameworks which can emit metadata to describe their composition. In the case of GOTM-ERSEM the output data to be emitted is partly specified by a user specified list of variables which are then mapped into the main data array (known internally as the CC array) at source level, with parameter values being loaded from a FORTRAN namelist. The physics/LTL model GOTM has an optional front end GUI interface containing a parameter editor and the model specification can be adjusted during the data editing phase before the core GOTM program starts up. GOTM-ERSEM has an internal array of metadata which describes the model's internal arrays in terms of full name, abbreviated name and unit dimensions, which is created for each instantiation of the source code at the same time as the CC array and describes internal as well as output variables. The Ecopath, Ecosim and Ecospace programs also have a main data array where metadata is stored, although dimensional information is specified for all data items together. This data array can be changed at runtime by means of the editor built into EwE v6, if organisms are added or deleted, for example.

Consequently the location in terms of numerical indexation within the data arrays of the two models is not fixed at runtime. Indeed there is no guarantee that the necessary data for a linked model run will even be available once the simulation starts, since the user may remove a needed component from their model instance. The first stage of model coupling is for each model to provide a specification of the data available for linking, which it provides between model loading / editing and

the start of the run. Each model provides this information in XML format. The data is supplied in a hierarchical manner, suitable for reading with the Document Object Model library (Le Hégaret et al., 2009) with the following hierarchy:

```
<Model>
  <Description> Author, System, Languages
  <Interface>
    <Time Information> Periodicity, Start and Stop Times
    <Grid Information> Dimensions, Start and end, Resolution
    <DataItems> of type Phytoplankton, Zooplankton, Consumers, Resources
      <Name>
      <Symbol>
      <Chemical Consituent> Carbon, Nitrogen, Silicate, Phosphorous
      <Flux Direction> Concentration, Intake, Offtake, Mortality, Production
      <Units>
    </DataItems>
  </Interface>
</Model>
```

Information on the dimensions of values is included in the metadata specification. If the dimension of a coupled variable differs between coupled models, such as density of a nutrient per unit volume, versus density per unit area, Couplerlib checks in its conversion dictionary to find an appropriate dimensional converter and then applies the relevant conversion ratio to all values transferred between the models (in this case multiplying or dividing by water column depth).

The model specification is produced by combining a template file, which is an XML file describing the basics of each model (such as language and system), with variable-specific metadata provided dynamically by the model after loading /editing. The specifications for the coupled models are combined by use of a coupler specification which is also an XML file; this specifies the models to be linked as a number of interfaces. The coupler crosschecks for each interface in turn to see whether the connection is permissible. Permissible interfaces occur when the model and system information are permissible, the spatial and temporal dimensions are in range for the data, and there is a

1 registered correspondence between all functional groups specified in the interface; i.e. the groups
2 have the same name, or there is a correspondence between groups in the system dictionary (e.g.
3 group Z4 in ERSEM corresponds to Herbivorous mesozooplankton in our EwE model) or a
4 correspondence between multiple groups (e.g. P1 to P5 in ERSEM sum to phytoplankton in our EwE
5 model) . So for each organism in the interface there must be a way of obtaining all nutrient
6 concentrations and the time and area of the data must be consistent. The functional groups that are
7 checked are only those which are specified as actually needing to be linked in the interface
8 specification.

9
10 Alongside the XML files used for specifying the rules of the interface, schema (XSD) files can be used
11 to validate the vocabulary and patterns of any human written XML file. For example the XML file will
12 specify the axes used in the grid, whilst an XSD file will specify that a three dimensional grid will
13 require three axis dimensions each of which will need a start co-ordinate, a length and a resolution.
14 The XSD file is loaded into a XML editor such as EditiX (www.editix.com) prior to creating or loading
15 an XML file. It can also be located online via a URL so may reside in a cross-institutional repository.
16 Its use is optional but will reduce run time errors in Couplerlib.

17
18 Couplerlib can be extended to the case where the functional group values are structured in some
19 way (e.g. age or size). Additional metadata will define the number and specification of age / size
20 classes. However further conversion code would need to be supplied when one model outputs as
21 size or age structured and only a single biomass values is needed.

22
23 However, it will often be the case that there is a disparity between the metadata produced by the
24 two models to be coupled. The simplest is the use of different names by the two models, other
25 situations arise where there is an inexact match of the functional group definitions between the two

models: for example, the version of ERSEM used here has 5 different phytoplankton groups whereas the EwE North sea model has a single phytoplankton class. It is acceptable, within EwE to model the relative offtake of all the phytoplankton groups as a single value if the predation pressure on bulk phytoplankton in EwE can be assumed to apply equally to all phytoplankton groups, but the production and consumption of phytoplankton within GOTM –ERSEM vary between phytoplankton groups. In each timestep the predation of phytoplankton is returned as a fixed fraction of input phytoplankton and then this fractional reduction is applied as a proportion to all phytoplankton groups in the ERSEM model. Consumption of phytoplankton directly by higher trophic level groups (e.g. Fish Larvae) is a relatively unimportant energy flow in the EwE models we have studied with only $3.2 \text{ g m}^{-2} \text{ year}^{-1}$ out of $60.8 \text{ g m}^{-2} \text{ year}^{-1}$ total predation not being consumed by zooplankton so this inaccuracy has little effect on the zooplankton mediated energy flows. The final case of disparity is where units of measurement vary between models, which occurs in this case as ERSEM expresses plankton concentrations in molar concentration units of C, N, P and Si and the EwE North Sea model in terms of tonnes per square kilometre of total biomass. These different conflicts are dealt with by a final XML file called the dictionary file. This specifies mappings between groups and units, unit conversions and synonyms. Furthermore items in the dictionary can be specified in terms of their context (models, interface and organism name). This is useful for exceptions to normal rules, such as greater wet-mass to dry mass ratios for jellyfish compared to most organisms. The dictionary is searched when an exception is found between organism names, or unit names, during the coupling process. The strength of the dictionary approach is that the model names, units and conversion definitions are stored outside code in a repository which can be subject to human scrutiny without need to write code and is by default platform and model independent (but a specific model context can also be specified)

2.4 Use of Graphical User Interfaces (GUIs) – model front ends

1
2 Both the EwE models and GOTM (and hence GOTM-ERSEM) have GUIs. In the case of EwE the
3 interface is written in the .net programming environment for windows, and whilst programmed in
4 Visual Basic, it uses an object structure which is fully accessible using Visual C++ and C#. The EwE
5 system is runtime extendable using plugins which use the .net interface system to extend the object
6 model of the underlying EwE system. Plugins are dynamically loaded as DLLs (executable code
7 program extensions, which must be called from another executable). This plug-in architecture is
8 used to couple EwE, via Couplerlib, to the GOTM-ERSEM front end. The GOTM interface, however, is
9 written in Python, which is an interpreted language: the interface consists of a set of scripts and
10 modules that require a Python interpreter to run. Accordingly an interpreter is embedded within
11 the EwE plugin, together with python.net (Pythonnet.sourceforge.net): a .net compatible API that
12 provides access to the Python interpreter.

13
14 The coupler plugin has a visual interface to select the directory where the coupler description XML
15 files described previously are contained. This launches the Python interpreter to run the scripts
16 which collectively make up the GOTM-ERSEM front end. These scripts enable the GOTM and ERSEM
17 configuration files to be manipulated at run time and also for output data to be visualised. In turn,
18 the Python-based front end communicates with the FORTRAN core of GOTM and ERSEM via a C
19 wrapper layer – F2Py (Peterson 2007). To enable coupling with external models, this layer has been
20 expanded with interfaces for read-write access to relevant FORTRAN data objects, in particular the
21 arrays that describe the values of biogeochemical state variables. Couplerlib is available to parse
22 metadata and carry out the kind of managed data conversion described in the previous section. The
23 net result of these various conversion and API layers is that meaningful data objects in GOTM-ERSEM
24 are available from within EwE and vice versa.

2.5 The LTL and HTL models

The 1D GOTM-ERSEM (Burchard et al., 2006) lower trophic level model was set up to represent a site in the Oyster Grounds (a muddy sand site immediately south of the Dogger Bank at 54°24' N, 4°3' E) with appropriate water depth (40m), tidal velocities (M_2 amplitudes of 30cm s^{-1}) and meteorological forcing. This parameterisation simulates the annual dynamics of a summer stratified site that is broadly representative of average behaviour over a large area of the central North Sea. Preliminary validation of this model is reported in (Van Der Molen et al 2012, in press).

The North Sea EwE ecosystem model (Makinson & Daskalov, 2007) is a spatially averaged representation of the biomass of food-web components over the whole North Sea (ICES div IVa-c). It consists of 66 functional groups, which are both pelagic and benthic. However, in order to be consistent with ERSEM, this 66 was extended by two new detritus groups (separating particulate organic matter (POM) into pelagic and benthic components and adding a new faecal POM group). Making 68 in total. The Lower trophic level groups used by both EwE and GOTM-ERSEM for pelagic and benthic components are listed in table 1.

2.6 The biology of the coupling

The GOTM-ERSEM representation of the physics and primary production and the EwE representation of the higher trophic levels are coupled via biomass of pelagic plankton groups and nutrients returned to the water column via detritus (Table 1). Omnivorous zooplankton (Z4 in ERSEM), microzooplankton, bacteria and phytoplankton (Z5, Z6, B and P1-P5 in ERSEM) are chosen as the coupling groups. The most significant of these links is with omnivorous zooplankton, because

in shelf seas it forms the principal pathway that connects the energy from the lower trophic food web with the higher trophic level consumers. The link between omnivorous and carnivorous mesozooplankton is handled in Ecosim rather than ERSEM (with predation by carnivorous Zooplankton turned off in ERSEM). This modification (carried out by the authors and Piet Rudaji author of ERSEM) was to increase the stability of Carnivorous Zooplankton populations, otherwise there was a tendency for either Carnivorous Zooplankton to go extinct over the winter months or for over consumption of Phytoplankton late in the year depending on the coefficients of predation of the two mesozooplankton groups. Nutrients are returned to ERSEM from Ecosim within the particulate, dissolved and faecal detritus components, which required adding the faecal component in Ecosim. The 1-D version of GOTM-ERSEM captures the biogeochemical processes occurring in the water column, describing the plankton ecosystem from incoming solar radiation, via phytoplankton to zooplankton. The EwE model assumes that the biomass of all functional groups is distributed evenly throughout the model domain.

The coupling is by two unidirectional Couplerlib interfaces. ERSEM passes to EwE the biomass of the pelagic functional groups, while the amount of offtake by predators as a proportion of biomass is returned to ERSEM, as are the detritus fluxes which are added to the ERSEM detritus pools. The version of EwE used has been specially adapted to run at a timestep of a single day, rather than a month. This is to prevent the accumulation of consumption over time long time steps leading to apparent overgrazing of omnivorous zooplankton. Internally GOTM and ERSEM run the North Sea model at a timestep of 10 minutes. The feeding information of the HTL components on the LTL components is returned to ERSEM as an offtake proportion by which the relevant LTL quantities are reduced. The dynamics of the benthic components are run separately in each model and thus shadow each other. Detritus is returned to ERSEM, but benthic groups are not fed into EwE nor updated in ERSEM. The EwE component is affected by the temporal variability of the detritus, but

not by the immediate variation of benthic conditions input experienced by the ERSEM model (because these are not explicitly coupled). Consequently the variability of the species which consume benthic components within EwE is less than those that feed heavily via the pelagic route. Which was observed as rapid annual oscillations of pelagic components of short lifetime and asymptotic increase in solely benthic components such as epibenthos. The benthos in the HTL model is therefore dominated by the top down higher trophic level interactions, improvement of the bottom up influences is dependent on improvements of the LTL benthic model within ERSEM that are actively being researched in the LTL modelling community.

ERSEM explicitly describes the concentrations of individual chemical elements (carbon, nitrogen, phosphorous and, for some functional groups, Silicon). However EwE is a biomass only model (the ratios of carbon, nitrogen and phosphorous are not considered. Consequently Couplerlib must estimate the stoichiometric ratios of EwE model groups, such as detritus, when transferring biomass based values from EwE models back to GOTM-ERSEM, so as to conserve biomass. It achieves this by having two detritus pools excretion which matches the C:N:P ratios of ingested nutrients and decay which starts out assuming a Redfield ratio and then adjusts to remove old biomass in the Redfield ratio and accumulated biomass in its new ratio. This is in accordance with the findings of Frigstad et al. (2011) that P concentration in non-autotrophs could vary seasonality by around 50% over Redfield but was averaged over the course of a year to Redfield or slightly above. ERSEM predicted variations of 100% or more in P composition in Zooplankton in the most extreme conditions.

Care must be taken to ensure dimensional correctness, for example when EwE models are expressed in terms of wet biomass, whereas in ERSEM they are expressed as g C/m^2 . When the dimensions of values are included in the metadata specification, Couplerlib will check in its conversion dictionary to

find an appropriate dimensional converter and then apply the relevant conversion ratio to all values transferred between the models.

2.7 The coupling process in action

The main sequence of inter-model calls and data communication is illustrated in fig. 2. It can be seen that there are six model components. At the ends of the stack of components are the EwE and GOTM models themselves (ERSEM is implemented as a library extension to GOTM but is interfaced through GOTM). Furthermore there is a front end interface to GOTM, which is written in Python and uses the Qt graphical interface components (Bruggeman & Bolding 2007). Although the use of this component is optional, it provides useful on line data editing, file management and visualisation facilities, so the decision was made to couple the models via the front end to keep the existing functionality.

The six most important points of interface between the core and the GOTMlink plugin are indicated in fig. 2, but there are additional points for documentation, identification and visual interaction with the plugin. Control passes from the EwE core to the plugin at those specific points where additional functionality can be provided by the GOTMlink interface. When the two models are coupled in the same process running on the same machine the two sides of the model (EwE plus its plugin and GOTM plus its GUI and Interface) operate on separate threads; that is to say, there are separate sequences of operations for the EwE and GOTM models. The EwE and GOTM interfaces signal directly when a step in one model is complete and the second model should take over, with execution of the first model suspended in a wait state. Using separate threads is essential where the component models are themselves multithreaded and may need to keep other threads such as the Graphical User Interface or network code running – EwE is like this, GOTM is not, but can be embedded within a multithreaded server, such as a network daemon. Data transfer is carried out by

1 means of Couplerlib, which provides marshalling of the data using external metadata specifications.

2 However the two interfaces can exchange basic information about model status, time to run for, and
3 the like through shared memory.

4
5 In networked mode the GOTM and EwE sides cannot communicate directly since they do not share
6 common memory. Instead everything goes via Couplerlib. The EwE and GOTM sides have their own
7 copies of Couplerlib. The two Couplerlibs will exchange information using TCP/IP sockets so that the
8 receiving end of a data transfer makes a copy from information requested from the sender

9 Couplerlib. Information used to synchronise the models and give information on model state such as
10 waiting, finished and error conditions for various operation is sent via a separate pair of sockets and
11 diagnostic information is sent to the Couplerlib which is attached to the graphical console (the EwE
12 GOTMlink plugin in this case).

13
14 The GOTMlink plugin includes an interface to the model coupling facilities. There are four user
15 initiated operations: GOTM model load, Edit, Run and Visualize. Only the run call operates beyond
16 the GOTM GUI into GOTM itself. Instantiation of the EwE model is controlled through the EwE GUI.
17 When the EwE and GOTM models are loaded they export their specifications to separate XML files.
18 These are made available, together with the user provided interface specification and dictionary to
19 Couplerlib. Data, such as the timestep, and the start and stop times are transferred from GOTM to
20 EwE sides (so they can be available before EwE produces its XML metadata). The bulk of the data is
21 transferred via Couplerlib. The critical call to Couplerlib is the CheckIF call, which loads all the
22 metadata specifying the interface, and cross checks this against the specifications of the component
23 models. Providing this is consistent, the GOTM and EwE threads are synchronised to ensure that the
24 Ecosim module has been started, resulting in the EcoSimRunInitialized call being sent from the EwE

core to the plugin. Both component models can then use GetIFAddress to locate the position of the data that will be transferred in the Couplerlib data store.

The main loop consists of repeated calls to the GOTM library, requesting simulation for a particular length of time, equal to the EwE time step. After every call the GOTM array holding the values of biogeochemical state variables is accessed and its values are stored in Couplerlib. The EwE part of the model will wait for these data to be written, and then fetch the values it requires, using Couplerlib to carry out any conversion between units or between functional groups. After updating the values of coupled variables in EwE in this fashion computation continues with the EwE calculations for a single time step. When complete, updated EwE data are written to a different frame of the Couplerlib repository and GOTM must synchronise and then read the data. The changes in biogeochemical state variables during the EwE time step are output to Couplerlib as (dimensionless) relative changes, which can be applied directly across all vertical layers of the GOTM-ERSEM water column without violating mass conservation. Finally if EwE has finished (the Simulation Ending Interface call is made) GOTM can be made to terminate early and cleanup called to ensure data are written to output files.

2.8 EwE model re-parameterisation

The very first attempt to link the two models was not successful. Pelagic groups that feed off mesozooplankton in EwE overexploited the mesozooplankton during the winter period causing the death of the zooplankton populations and a consequent reduction in pelagic productivity. This was a result of the Ecosim model having being calibrated primarily on summer plankton levels to set the consumption to biomass level, albeit annualised i.e. the seasonality in plankton food production was not captured in EwE (winter zooplankton levels in ERSEM drop to around 1% of summer ones). In

practice consumption is very much lower in winter months when the plankton population is lower but also the basal metabolic rate of the predatory fish is reduced by around a factor of 2.4 compared to summer (Clark & Johnson 1999). In addition to the reduced metabolism which follows an exponential temperature to metabolism law, there may be behavioural changes that reduce prey consumption still further. For example sandeels, which were observed to be one of the major predators of mesozooplankton switch to being benthic between November and April (Engelhard et al. 2008) and so have zero consumption of plankton in winter months. To account for these phenomena, it was necessary to add a function to reduce predation of pelagic sources in winter months. Two forcing functions were used – one a purely metabolic model using a $Q_{10}=2.0$ rule of temperature, with maximum temperature having a multiplier based on the original EwE model of 1.0, the other fitted the observations on sandeel populations of Englehard et al. (2008). In addition metabolic costs were reduced to 0.3 for adults and 0.0 for juveniles (this latter figure reflects the seasonal production of juveniles starting in the spring, and their migration from spawning grounds, so there will not be juveniles in the area modelled but the population still needs to be modelled and EwE allows only a fixed value for metabolic costs). The background mortality of the zooplankton in ERSEM was reduced to a nominal value to take account of the fact that most mortality would be due to predation which is accounted for in the EwE model. Interestingly, the level of zooplankton predicted by the ERSEM model peaked at around the value used to calibrate EwE. All of these changes represent a degree of bringing the data to the model and compensating for known limitations of a single model by dampening the damaging effects of assumptions in the underlying model (specifically the low winter zooplankton in ERSEM). In the discussion we will suggest how this compensation may be addressed in the future.

3 Results

3.1 One dimensional two way coupling between LTL and HTL.

Here the results from the application of Couplerlib to coupling the LTL and HTL models of the North Sea are described. Results are produced separately but concurrently by the EwE and GOTM-ERSEM parts of the model.

Within ERSEM, the dynamics of Zooplankton are constrained by the dynamics of 5 phytoplankton groups, the most important of which are, diatoms flagellates and dinoflagellates (fig 3a, 3b, 3c resulting LTL data before aggregation in EwE). Two years of the 40 year run are examined in detail: 1970 and 1971. There is an annual dynamic, with silicon dependent diatoms peaking in March – April and then being limited by Silicon availability. The dynamics into the summer are then determined by the growth first of flagellates and dinoflagellates that lead to the gradual reduction in the zooplankton peak. The transition between a phytoplankton community dominated by diatoms and one dominated by Dinoflagellates was observed in the North Sea EwE model calibration exercise (Mackinson & Daskalov 2007). However there is also considerable annual variation and seasonal variation in diatom and dinoflagellate numbers during the summer period.

In fig 3d the modelled omnivorous zooplankton dynamics are indicated (they are subject to our best guesses as to, in particular, the seasonal mortality of zooplankton, which is a challenge to estimate (Daewel et al. 2014)). Between 1970 and 1971 there is a considerable difference in the timing of the peak as a result of differences in the diatom bloom. These differences are likely to have profound, though not completely understood effects on recruitment on pelagic fish in the following year.

An examination of the effect on a number of functional groups is shown in fig. 4. Each graph shows the changes in fish population numbers away from the calibrated levels at the start of the simulation (based on the 1990 calibration data) as a result of coupling to the LTL model, the transient effects of the coupling swamp the inter-annual variation to be described later . A significant rebalancing

between functional groups occurs as a result of the introduction of the seasonal lower trophic levels. That the coupled model should give different answers is not surprising, given that i) there is a switch from a non-seasonal model to a seasonally driven one ii) the ERSEM model is calibrated for a single point, the EwE model is for the whole North Sea iii) It has been shown that for large models small changes in the growth of a single population can have drastic consequences for the population of the North Sea (Rossberg 2013). iv) Communities are subject to periodic shifts in species within a guild e.g. Sardine and Anchovy cycles (Chavez et al. 2003). Given these differences between the models we cannot expect a congruency between coupled and uncoupled models, nor can we predict the likely population of a single species based on meteorological data alone, because it is hard to estimate recruitment in a mixed species environment. Nevertheless a stable long-term equilibrium may be found, as the trophic network with Ecosim readjusts. There is a considerable rebalancing between pelagic groups, for example Herring numbers decline (probably a result of the difficulties of modelling this migratory fish in a static model with no space, no larval dynamics and limited seasonality). We believe that with extensive reworking of the ERSEM and EwE models and extensive recalibration we can produce a useful operating model to examine how changes in the physical and lower trophic level environment may further affect the ecosystem balance between multiple species.

3.2 Long Term Effects of Changes in the physical environment on Fish Biomass

The coupled model was run using synthesised weather data based on the IPCC climate change scenarios A1B (which is a rapid global industrial growth scenario with a mixture of fossil and non-fossil fuels resulting in around 3.2 degrees of warming) and B1 (with limited industrial growth and an emphasis on CO₂ reduction resulting in 2.4 degrees of warming) (Nakicenovic et al. 2000). The nutrient scenario was either high (representing the level of nutrients found in the eutrophic Southern North Sea) or low (with half the levels of N, and P) (Table 2). The use of these scenarios

reflects the time frame over which this model development occurred but the nutrient levels are levels chosen for sensitivity analysis and not agreed scenario estimates.

The populations of marine species affected by changes in the environment are grouped in seven blocks: Microbes (plankton and bacteria), higher level invertebrates (shrimps, crabs, bivalves etc.), small and medium pelagic fish (herring, mackerel, sand eels etc.), Demersal fish (cod, haddock etc. but not flatfish), Flatfish (sole, plaice etc. excluding skates & rays), sharks (including skate and rays) and a block containing marine mammals and birds. The populations of those species relative to the current temperature and nutrient scenario (baseline) in the four scenarios for temperature and nutrients are shown in fig. 5. It may be seen that the system is very much nutrient limited, especially in pelagic species (where there is a proportional relationship between nutrients and biomass). In large benthic species and mammals the relationship is less than linear (i.e a 30% reduction in biomass for a 50% reduction in nutrients). The effect of warming is relatively small and greatest for pelagic species and higher trophic level predators such as sharks, and negligible for the microbes. The reason for diminution of nutrient effect at higher trophic level effect is that nutrients limit the microbial populations whilst increased temperature increases rates of turnover of microbes and hence nutrients, and that there is strong negative feedback on predated populations which are predated at a higher rate when predator populations increase.

4 Discussion/conclusions

Whilst end to end modelling, from physics to ecosystem has been characterised as an important procedure in developing an understanding of marine ecosystems, it is one beset by numerous challenges, technical, scientific and organisational. Couplerlib is designed to address some of the former, which include incompatibilities of system, language, metadata descriptions and terminology, units, spatial and temporal scale, and some of the deepest system problems such as incompatible

1 methods of data storage. In an ideal world we would design end to end modelling systems in a top-
2 down co-ordinated manner and the Atlantis model comes closest to this. In other cases some of the
3 technical incompatibilities can be overcome by translating whole models and directly interfacing
4 them to existing models. However heavy duty model refactoring is expensive and can break the link
5 between development teams working on the core modelling solutions. Couplerlib is proposed as a
6 more workable solution, whose main strength is its ability to operate over different coupling
7 strengths from direct coupling to loose coupling via netCDF files. Its main advantage is that the links
8 are made explicit via metadata rather than programming, that simple reconciliation of groups and
9 data can be carried out and that it can work across languages and systems including across
10 networks.

11
12 The main disadvantages of couplerlib (and of the model coupling approach in general) are a complex
13 install and that it still requires interface code to integrate into the models to be coupled. The most
14 difficult parts of this project, where getting the threading interaction between the two models
15 GOTM/ERSEM and EwE, which was a specialist programming task involving waits and signalling
16 between processes, which occurs in the glue layer between couplerlib and the model code. This is a
17 task that only has to be done once for each modelling tool, so that the majority of model users
18 would not have to alter code. In a large multi-team project it is likely that the team will include
19 coders and ecological modelling specialist who can work in their specific domain.

20
21
22 The technical innovation of the coupling processes and the solutions to the ecological problems that
23 emerged has exposed both technical and ecological issues of reconciling the two models, the key
24 points of which are discussed below.

1 In principle Couplerlib is designed to handle coupling of any groups among models, but in this
2 example only the plankton groups are coupled. The reason for this lies in the effort/benefit ratio, the
3 argument being that plankton groups form the main energy conduits through which environmental
4 changes get translated in to changes in secondary productivity, and poor correspondence between
5 the benthic components of the two North Sea models means a lot of effort would be spent for little
6 marginal gain. Having the dynamics of the benthic components in each model run separately is taken
7 as providing a useful point of comparison.

8
9
10 The example coupling, links a 1D LTL model calibrated using parameters from the North Sea ‘Oyster
11 Grounds’ with the HTL food web model whose parameters are spatially averaged over the whole
12 North Sea. This discrepancy is related to the more strategic nature of EwE – long term changes in
13 ecosystem composition versus the more tactical nature of ERSEM – response to environmental
14 transients. While the parameters used to represent each model differ to in their degree of spatial
15 specificity, we do not see this giving rise to conceptual or technical issues that cannot be reconciled.
16 First, the Oyster Grounds is a site close to the tidal mixing front that divides summer stratified with
17 the mixed region to the south. Although it simply cannot capture fully the diversity between a mixed
18 site in the southern North Sea and a deep strongly stratified location in the northern North Sea, it is
19 nevertheless representative of the dynamics of a large area of the central North Sea. Second, both
20 models utilise many rate and state parameters from a much a wider area of the North Sea or are
21 indeed specified even more generally. The internal consistency of each of the separate models
22 might thus imply that each could be considered as rather general representations of the North Sea
23 ecosystem. From a technical point of view, the practical arguments for coupling these non-spatially
24 resolved models are compelling; the coupling allowing proof of concept on the technical methods
25 and early identification of key ecological issues before extending to spatially resolved models. We

1 believe that the main issues have been resolved and plans are in place to tackle coupling spatially
2 resolved models, knowing full well the new issues will emerge, such as capturing the dynamics of
3 annual migrations for some species. On the other hand running ERSEM as a spatial model and
4 aggregating over space would help scale up the model.

5
6 Plankton growth in ERSEM is regulated by four elements: carbon (C), whose production depends on
7 light and temperature, nitrogen (N), phosphorous (P) and silicon (Si). With the particular
8 parameterisation chosen for the Oyster ground model, it has been seen that N is not generally
9 limiting and that light, as expected, is only limiting during the winter months, and that Si and P are
10 more limiting in spring. However, since we have fixed the total amount of these elements and the
11 model is closed, these results are a consequence of our particular choice of initial conditions rather
12 than being a true model 'prediction'. The dynamics of silicon dynamics is more interesting. In
13 particular diatoms, which are more 'digestible' by mesozooplankton, bloom in early spring and then
14 are replaced by the flagellates and dinoflagellates as the Si concentration drops, a result consistent
15 with Broekhuizen et al. (1995). The possible biological consequences of this are discussed by Officer
16 & Ryther (1980). One consequence of the Si dependency of diatom production is that Si is only
17 gradually accumulated as a result of inflow from rivers and thus does not change much from one
18 year to the next, so that the total diatom production level is constant from one year to the next
19 although timing of spring bloom will vary from year to year in relation to temperature and light.
20 There is apparent inter-annual variation between the timings of the growth of the different
21 phytoplankton peaks, the importance of which will depend on the ability of zooplankton to
22 incorporate biomass from the different phytoplankton groups. This single point or average model
23 cannot represent the inter-annual variation in phytoplankton bloom location resulting from inter-
24 annual variation in weather combined with ocean currents. For example where plankton bloom

location varies in space from year to year the effect on the higher trophic level species will depend on the ability of that species to move with the plankton bloom location.

Another issue is that of spatial pattern of biomass production across the whole of the North Sea, which can vary considerably across time. For example, data used in a meta-analysis by Broekhuizen et al. (1995) showed considerable variation in summer peak omnivorous zooplankton numbers with populations off the Norway coast (Box 3 in their paper) having a peak between days 100 and 180, whilst populations off Denmark (Box 10) were low until after day 150 and stayed high between day 180 to 270 a late summer bloom. Overall the original ERSEM parameterisation placed too much emphasis on the spring diatom bloom as a source of production and so gave a pattern of productivity which did not lead to a stable ecosystem. The current parameterisation which peaks around day 180 gives a better estimate of mean mesozooplankton biomass, even though the nutrient model which gives rise to it requires some refinement to fit to observed levels.

There are a number of known divergences from observations for the ERSEM model: for carnivorous zooplankton there is under prediction in winter and too great a rise in spring (avoided by moving this group to EwE); and for microzooplankton there is an over estimate of densities (although this may be as a result of errors in the observations) (Broekhuizen et al 1995). Modified EwE parameterisations have been made to correct for these: reducing winter metabolism and setting consumption for herring to be very low in winter and adjusting the maintenance cost of herring downwards, in order to prevent this underestimation leading to broader erroneous trophic effects, as a result of erroneous herring deaths. In the long term improvements to the zooplankton component of ERSEM may remove the need for some of these corrections. In the short term these results raise a number of issues about the dynamics of planktivorous species such as herring. Given the paucity of zooplankton during the winter months, how do they manage to survive over winter? The answers

1 seem to be multifaceted; herring switch prey over the winter months (Blaxter & Hunter, 1982),
2 herring are migratory and may move to extend the season of availability of food (Corten 2000) and
3 herring do seem to exist in a negative energy balance over winter leading to loss of condition. All
4 these aspects are of potentially great interest for researchers and should be incorporated into future
5 versions of the underlying models.

6
7 The timeframe of the constituent models is critical to their interoperability. This has often been
8 characterized in terms of timestep, but reality is much more complex than that. GOTM is a model
9 that makes use of very high resolution temporal data, to represent the effects of short term weather
10 conditions and tides, at a temporal resolution that is sub-hourly. The biological component of ERSEM
11 can respond to these changes in a matter of days because some of the functional groups, e.g.
12 bacteria and diatoms can respond over this timeframe. In the same way the Ecosim model has these
13 components with P/B ratios of 30 per year or more suggesting that the daily timeframe is more
14 appropriate for simulation than the default monthly timeframe. However Ecosim is a model that
15 relies on empirical calibration which is usually annual, although some attempts have been made to
16 make seasonal Ecosim models with quarterly recalibration (Althausen 2003). In the current context it
17 would seem that there is a need for seasonal adjustment of parameter values (especially of P/B
18 ratios, basal metabolic rate and starvation mortality). Data for this calibration are not generally
19 available at high frequency, however.

20
21
22 Differences in biomass and productivity between, often quite closely related species vary
23 considerably between the original North Sea Ecosim model and the coupled model. However both of
24 these results are an encapsulation of a complex spatiotemporal dynamic because there are seasonal
25 fluctuations in the abundance of many of these species. To make a detailed predictive coupled

1 model would require calibration in time and space, to ascertain, for example the changes in diet
2 composition which occur in time and space. When seasonal variation in productivity is included (e.g.
3 by coupling models or adding in seasonal forcing), temporal calibration in Ecopath with Ecosim
4 requires consideration of seasonal effects on long-term equilibrium, and given that Ecosim may have
5 multiple equilibria, coupling may cause instability in equilibria in some circumstances.

6
7 Does the short term behaviour matter when considering long-term environmental change? In terms
8 of the pelagic system the answer would seem very much to be yes. Over-winter survival of
9 everything from small pelagic fish to seabirds seems to depend upon reserves built up in the
10 previous season, and variation in survival and recruitment is generally considered to be influenced
11 by bottom-up constraints on feeding. In the modelling exercise reported here it was found that
12 overwinter survival often depended on sufficiently rapid growth of plankton populations in spring
13 when a fish's metabolism was increasing. It was necessary to ensure that the forcing functions used
14 to limit winter feeding were matched the timing of the spring bloom. Further software development
15 could enable this matching to be automatic. This development could be used to model situations,
16 such as the case of Moroccan sardine populations, where timing between zooplankton peaks and
17 conditions favourable for spawning influences recruitment (Machuo et al 2009). We can use
18 experimental plankton studies to make a sensitive calibration of planktonic response to short term
19 meteorological variation, but our ability to explore the effect of climate change on the whole
20 ecosystem, particularly its higher trophic level components, and to predict typical weather in future
21 years is very limited.

22
23 The addition of a process oriented model such as GOTM-ERSEM to EwE may allow us to express the
24 relationships between the drivers of change and the consequences. However, the relative
25 insensitivity of EwE to short term perturbations begs a number of questions as to its suitability in

1 studies on the likely consequences of climate perturbations to the ecosystem. Firstly the fixed
2 seasonality of Ecosim may not adequately represent the results of shifts in the timing of variable
3 onset events such as plankton blooms, but could be incorporated by adding a run time updated
4 forcing function to the feeding via a plugin within Ecosim. Secondly Ecosim is limited in its ability to
5 represent the effect of environmental variation on transient stages of the life cycle of marine
6 organism such as the pelagic larvae of fish and benthic invertebrates. This is because of the lack of a
7 model of short term drivers, and components for body condition, and the calibration of the multi-
8 stanza model based on a Von Bertalanfy growth curve which is based on the growth curve of sharks
9 and does not represent larval stages well (Urban 2002). The North Sea model calibration,
10 consequently, pays little attention to these stages. Recognising such limitations, we are however
11 faced with a practical dilemma; the need for data and modelling tools that help address the pressing
12 research question about how ecosystem respond to change, and the implications for management.
13 The pragmatic route must be to use the modelling capability we have and use it to help learn how to
14 do better in the future.

15
16 Alongside the issue of time is that of space. For some parameter values there were times of the year
17 when either prey or predators became severely depleted because of availability of prey, but in
18 reality the patchiness of and interconnectedness of areas with different prey levels would lead to
19 rapid recolonisation of the area by new predators and prey. An example of spatiotemporal local
20 severe depletion was that of the Bay of Biscay anchovy in 2005 (Borja et al. 2008). To address this,
21 an Ecospace model (the spatial equivalent of Ecosim) can be linked to the ERSEM model within
22 GETM, the spatialized version of GOTM; this will allow us to model the temporal shift of marine
23 organisms in response to the spatiotemporal dynamics of both marine plankton and fishing
24 pressure. However, spatial biogeochemical models require considerable computational resources to
25 run in most practical situations (typically of the order of a week or more time on a cluster for a 100

year run). The single column GOTM-ERSEM model requires about 90 seconds per year to run on an Intel i5 processor, and combined with the integrated front end, makes it much more suitable for exploratory parameterisation, teaching and prototyping than the larger models. The Ecosim model takes under 1 second per year to run, it has a timestep of 1 day rather than 10 minutes and has only a single depth layer so that it is calculating less than 1/1000 of the number of time and space elements and Couplerlib is a small proportion of the Ecosim total (less than 10%) so is a trivial addition. However it is less suited to the situation where biological component models such as BFM are being called every short timestep across the managed interface. In this case a more direct coupling method such as FABM (Bruggeman & Bolding 2014) is more suitable, though this will not work across languages and systems in the same way Couplerlib does. Couplerlib can still be used for initial metadata validation, however.

In the realm of biogeochemical modelling there are a number of different models (e.g. Nemo-PISCES (Aumont et al. 2003, Madec et al. 1991, ECOSMO (Schrum et al. 2006), Ecoham (Pätsch and Kühn 2008)) which model the dynamics of plankton in response to changes in the physical environment mediated by physical oceanography. The differences between the model stem from the number of plankton groups and the treatment of the benthos. However the treatment of the environment as a three dimensional grid lends itself to a common interface with Ecosim and Ecospace (its 2 dimensional equivalent) though a specification of Couplerlib with 2 dimensional data elements and appropriate metadata specification. Ultimately, however the choice of modelling system will depend on the kind of data available and the parameters collected, because model parameterisation is a time consuming and costly part of any modelling process. The choice of ERSEM for both 1 dimensional and 3 dimensional modelling is as a result of the use of this model for the North Sea area being examined.

1 There is considerable debate about whether the components of end to end models should be simple
2 or complicated. Very often we are faced with models and sub models that are both over
3 complicated, in the sense that they contain too many parameters and components (Fulton et al.
4 2003) and over simplistic in that they contain what can certainly be described as simplifying
5 formulations and may even be seen as dysfunctional formulations (Anderson & Mitra 2009). This
6 sterility of this debate has been acknowledged in a number of forums (e.g. the WGIPEM ICES
7 working group <http://www.ices.dk/community/groups/Pages/WGIPEM.aspx>) along with the general
8 conclusion dating back to William of Occam that things should be as complicated as necessary but
9 no more so. What are needed therefore, are not so much simple models as simple conclusions that
10 are likely to be robust to uncertain small scale and detailed assumptions and limitations in model
11 formulation. We are a long way off capturing all the physical, biological and human conditions that
12 will enable us to predict the cod in the North Sea in 50 years' time, but we at least are starting to
13 have the tools that enable us to integrate important scientific findings within a whole ecosystem
14 context.

15
16 In conclusion, linking two very different models has produced a system which is richly expressive of
17 the kinds of interactions of the different functional groups that form pelagic ecosystems and the
18 causes of dynamics, including temperature, nutrient ratios and biological components of the system.
19 It is not at this stage very precise (Levins, 1966) because it models a spatial area as a single point in
20 space and because of a number of known limitations not just of the underlying ERSEM and EwE
21 models, but also of spatiotemporal variability of marine data. Nevertheless the linked model
22 represents a considerable advance on verbal arguments because it can demonstrate the likely
23 relative magnitude of these effects in a wider context of nutrient dynamics and marine fisheries
24 management.

5 Acknowledgements

The authors would like to thank Karsten Bolding for advice on the use of the GOTM model, Mark Platts for help in the preparation of the Ecosim figures and Piet Ruadij for help on understanding the ERSEM components of the model. Many thanks to Villy Christensen, Carl Waters, Sherman Lai, Joe Buszowski and Jeroen Steenbeek of the EwE team at UBC Vancouver for their aid during my 2 weeks stay in Vancouver and remote support for the plugin.

This work was funded under the Framework 7, Meece program (www.meece.eu) Cefas Seedcorn funding and Defra contract M1202. Useful discussion leading to the development of the model occurred during a workshop of the EasyCo Interreg project (www.project-easy.info).

6 References

- Althausen, L.L. 2003. An Ecopath/Ecosim analysis of an estuarine food web: Seasonal energy flow and response to River-flow related perturbations. MSC Thesis Louisiana State University and Agricultural and Mechanical College.
- Anderson, T.R. & Mitra, A. 2009. Dysfunctionality in ecosystem models: an underrated pitfall? *Prog. Oceanogr.* 54:66-68.
- Aumont, O., E. Maier-Reimer, S. Blain and P. Pondaven, 2003. An ecosystem model of the global ocean including Fe, Si, P co-limitations, *Global Biogeochem. Cycles*, 17:1060.
- Baretta, J.W., Ebanhoh, W. & Ruadij, P. 1997. The European Regional Seas Ecosystem Model (ERSEM) II, *J. Sea Res.* 38:229-483.
- Berners-Lee, T., Fielding, R.T., Masinter, L. Uniform Resource Identifier (URI): Generic Syntax 2005. Internet Society <http://labs.apache.org/webarch/uri/rfc/rfc3986.html>.

- 1 Blackford, J.C., J.I. Allen, Gilbert, F.J. 2004. Ecosystem dynamics at six contrasting sites: a generic modelling
- 2 study, *J. Mar. Sys.* 52:191–215.
- 3
- 4 Blaxter, J.H.S, Hunter, J.R. 1982. The biology of Clupeoid fishes. *Adv. Mar. Biol.* 20:1-223.
- 5
- 6 Borja, A., Fontán, A., Sáenz, J., Valencia, V. 2008. Climate, oceanography, and recruitment: the case of the Bay
- 7 of Biscay Anchovy (*Engraulis sicolus*). *Fish. Oceanog.*, 17: 477-493.
- 8
- 9 Broekhuizen, N., Heath, M.R., Hay, S.J., Gurney, W.S.C. 1995. Modelling the dynamics of the North Sea's
- 10 Mesozooplankton. Netherlands. *J. Sea Res.* 33:381-406.
- 11
- 12 Bruggeman, J, Bolding, K. 2007 GOTM Graphical User Interface
- 13 <http://www.gotm.net/index.php?go=software&page=gui/index>
- 14
- 15 Bruggeman, J, Bolding, K, 2014 A general framework for aquatic biogeochemical models. *Env. Mod. Soft.*
- 16 61:249-265
- 17
- 18 Burchard, H. Bolding K., Villarreal M.R., 1999. GOTM—a general ocean turbulence model. Theory,
- 19 applications and test cases, Tech. Rep. EUR 18745 EN, European Commission.
- 20
- 21 Burchard,H., Bolding K., W. Kuhn W., Meister A., Neumann T.,Umlauf L.,2006, Description of a flexible and
- 22 extendable physical-biogeochemical model system for the water column, *J. Mar. Sys.*, 61, pp 180-211.
- 23 Chavez, F. P., Ryan, J., Lluch-Cota, S.E. & Ñiquen, M, 2003. From anchovies to sardines and back:
- 24 multidecadal change in the Pacific Ocean. *Science*, **299**, 217 - 221,
- 25
- 26 Christensen, V., Walters, C.J. & Pauly, D. 2005. Ecosim with Ecopath: a user's guide, November 2005 edition.
- 27 Report Fisheries Centre, University of British Columbia, Vancouver, Canada.
- 28
- 28 Clark, A. & Johnson, N.M. (1999) Scaling of metabolic rate with body mass and temperature in teleost fish
- 29 *Journal of Animal Ecology*, 68:893–905,

Corten, A. 2000. A possible adaptation of herring feeding migrations to a change in timing of the *Calanus finmarchicus* season in the eastern North Sea. *ICES J. Mar. Sci.*, 57:1261-1270.

Cushing, D.H. 1996. Towards a science of recruitment in fish populations. Kinne, O. ed. Excellence in Ecology, 7. 175 pp. Ecology Institute, Luhe, Alemania.

Daewel, U., Sætre Hjøllø, S., Huret, M., Ji, R., Maar, M., Niiranen, S., Travers-Trolet, M., Peck, M.A., and van de Wolfshaar, K.E. (2014) Predation control of zooplankton dynamics: a review of observations and models *ICES J. Mar. Sci.* 71 (2): 254-271

Engelhard, G. H., van der Kooij, J., Bell, E.D., Pinnegar, J.K., Blanchard, J.L., Mackinson, S. & Righton, D. 2008. Fishing mortality versus natural predation on diurnally migrating sandeels, *Ammodytes marinus*. *Mar. Ecol. Prog. Ser.*, 369:213-227.

Frigstad, H., Andersen, T., Hessen, D.O., Naustvoll, L.-J., Johnsen T.N. and Bellerby, R.G.J. 2011. Seasonal variation in marine C:N:P stoichiometry: can the composition of seston explain stable Redfield ratios. *Biogeosciences* 8:2917-2933.

Fulton, E.A., Smith A.D.M, Johnson, C.R. 2003 Effect of complexity on marine ecosystem models. *Marine Ecol. Prog. Ser.* 253:1-6.

Fulton, E.A. 2010 Approaches to end to end ecosystem models. *J. Mar. Sys.*, 81:171-183.

Kearney, K.A. An analysis of marine ecosystem dynamics through development of a coupled physical-biogeochemical-fisheries food web model. Ph.D. Thesis University of Princetown 2012.

Le Hégaret, P., Whitmer, R., Wood, L. 2009. Document Object Model www.w3c.org/DOM

Lotze, H.K. H. S., Lenihan, B.J., Bourque, R.H., Bradbury, R.G., Cooke, M.C., Kay, S.M., Kidwell M., Kirby, C.H., Peterson, J.B., Jackson, C. 2006. Depletion, Degradation, and Recovery Potential of Estuaries and Coastal Seas. *Science* 312:1806-1809.

Levins, R. 1966. The Strategy of Model Building in Population Biology, *Am. Sci.*, 54:421-431.

Lewis, K. & Allen, J.I. 2009. Validation of a hydrodynamic-ecosystem model simulation with time-series data collected in the western English Channel, *J. Mar. Sys.*, 3:296-311.

Libralato, S and Solidoro, C. 2009. Bridging biogeochemical and food web models for an End-to-End representation of marine ecosystem dynamics: The Venice lagoon case study. *Ecological Modelling* 220, 2960-2971

Machuo, E., Ettahiri, O. Kifani, S., Benazzouz, A., Makaoui, A. & Demarcq, H. 2009 Environmental control of the recruitment of sardines (*Sardina pilchardus*) over the western Saharan shelf between 1995 and 2002: a coupled physical/biogeochemical modelling experiment. *Fish. Oceanog.* **18:287-300**.

Madec, G., M. Chartier, P. Delecluse and M. Crépon, 1991 : A three-dimensional numerical study of deep water formation in the Northwestern Mediterranean Sea, *J. Phys. Oceanogr.*, 21, 9, 1349-1371,

Mackinson, S., Daskalov, G. 2007. An ecosystem model of the North Sea to support an ecosystem approach to fisheries management: description and parameterisation. Sci Ser. Tech. Rep., Cefas Lowestoft, 142: 196pp.

Minto, C., Myers R.A. & Blanchard, W 2008 Survival variability and population density in fish populations. *Nature* 452: 344-347.

Nakicenovic, Nebojsa; Alcamo, Joseph; Davis, Gerald; de Vries, Bert; Fenhann, Joergen; Gaffin, Stuart; Gregory, Kenneth; Grubler, Arnulf; Jung, Tae Yong; Kram, Tom; La Rovere, Emilio Lebre; Michaelis, Laurie; Mori, Shunsuke; Morita, Tsuneyuki; Pepper, William; Pitcher, Hugh M.; Price, Lynn; Riahi, Keywan; Roehrl, Alexander; Rogner, Hans-Holger; Sankovski, Alexei; Schlesinger, Michael; Shukla, Priyadarshi; Smith, Steven J.; Swart, Robert; van Rooijen, Sascha; Victor, Nadejda; Dadi, Zhou. 2000. Special Report on Emissions

Scenarios : a special report of Working Group III of the Intergovernmental Panel on Climate Change.
Cambridge University Press, New York, NY (US)

Officer, C.B, Ryther, J.H. 1980. The importance of silicon in marine eutrophication. *Mar. Ecol. Prog. Ser.* 3:83-91.

Pätsch, J. and Kühn, W., 2008. Nitrogen and carbon cycling in the North Sea and exchange with the North Atlantic - a model study, Part I. Nitrogen budget and fluxes. *Continental Shelf Research*, 28: 767-787.

Peterson, P. 2007 F2Py: Fortran to Python Interface Generator. <http://cens.ioc.ee/projects/f2py2e/>.

Richardson, A.J., Schoeman, D.S. 2004. Climate impact on plankton ecosystems in the Northeast Atlantic. *Science*, 305: 1609-1612.

Rossberg, A.G. 2013 Food webs and biodiversity: Foundations, models, data. Wiley.

Rose, K.A., F.E. Werner, B.A. Megrey, M. Noguchi-Aita, Y. Yamanaka, D.E. Hay, J.F. Schweigert and M.B. Foster. 2007. Simulated herring growth responses in the Northeastern Pacific to historic temperature and zooplankton conditions generated by the 3-dimensional NEMURO nutrient-phytoplankton-zooplankton model. *Ecological Modelling* 202(1-2): 184-195.

Rose, K. A., Allen, J. I., Artioli, Y., Barange, M., Blackford, J., Carlotti, F., Cropp, R., Daewel, U., Edwards, K., Flynn, K., Hill, S. L., HilleRisLambers, R., Huse, G., Mackinson, S., Megrey, B., Moll, A., Rivkin, R., Salihoglu, B., Schrum, C., Shannon, L., Shin, Y-J., Smith, S. L., Smith C.; Solidoro, C.; St. John, M., Zhou, M., 2010. End-To-End Models for the Analysis of Marine Ecosystems: Challenges, Issues, and Next Steps. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*. 2:115–130.

Rose, K.A. 2012 End-to end models for marine ecosystems: Are we on the precipice of a significant advance or just putting lipstick on a pig? *Scientia Marina* 76(1): 195-201

Schrump, C., Alekseeva, I., Stjohn M. 2006 Development of a coupled physical–biological ecosystem model
 ECOSMOPart I: Model description and validation for the North Sea. *J. Mar. Sys.* 61:79-99.

Travers, M., Shin, Y.-J., Jennings, S., Cury, P. 2007. Towards end-to-end models for investigating the effects of
 climate and fishing in marine ecosystems. *Prog. Oceanog.* 75:751–770.

Travers, M., Y.-J. Shin, S. Jennings, E. Machu, J. A. Huggett, J. G. Field, and P. M. Cury. 2009. Two-way
 coupling versus one-way forcing of plankton and fish models to predict ecosystem changes in the Benguela.
Ecological Modelling. 220, 3089-3099

Urban, H.,J. 2002. Modeling growth of different developmental stages in bivalves. *Mar.Ecol.Prog.Ser.* 238:109-
 114.

van der Molen, J., Aldridge, J.N., Coughlan, C., Parker, E.R., Stephens, D., & Ruardij, P 2012, Modelling
 marine ecosystem response to climate change and trawling in the North Sea, Biogeochemistry Special Issue –
 Marine Ecosystem Connections

Vichi, M., Oddo, P., Zavatarelli, M., Coluccelli, A., Coppini, G., Celio, M., Fonda Umani, S., Pinardi, N.
 2003.Calibration and validation of a one-dimensional complex marine biogeochemical fluxes model in different
 areas of the northern Adriatic shelf., *Ann. Geophys.*, 21, 413–436.

Appendix A Source files and availability

Source files within this project have been included as follows:

- The source files for the standard C++ interface to Couplerlib.
- The source files for the Managed interface to Couplerlib which act as a wrapper when called from
 managed code other than C/ C++ (especially Visual Basic from EwE).
- The Source files for the interface between Couplerlib and EwE.
- The xml dictionary, and xml specification files for the GOTM/ ERSEM andEwE models
- Parameter namelists for GOTM and ERSEM.

1 In addition to this the model will require a complete implementation of the GOTM, ERSEM and EwE
2 models, and third party libraries for netCDF reading, a complete Python interpreter (2.5 to 2.7) with
3 Numpy, Matplotlib and PyQt4 together with driving data for GOTM (meteorological data etc.) and the
4 Access database of the relevant EwE model.

5 This material can be supplied on request but the kind of support needed to commission any non-trivial
6 implementation of an end to end model is likely to require considerable time and resources which can
7 best be provided by third party support.

1 Table 1. ERSEM and EwE function groups mapped

EwE Group No	EWE group	ERSEM group	Notes
51	Carnivorous zooplankton	Carnivorous mesozooplankton (Z3) but moved to EwE	Copepods
52	Herbivorous and omnivorous zooplankton (copepods)	Omnivorous mesozooplankton (Z4)	Copepod nauplii, copepodites, microzooplankton
53	Gelatinous zooplankton		
54	Large crabs		
55	Nephrops		
56	Epifaunal macrobenthos (mobile grazers)	Epibenthic predators (Y1)	Cross reference Group
		Filter feeders (epifauna / infauna) (Y3)	
57	Infaunal macrobenthos	Deposit feeders (infauna) (Y2)	
58	Shrimp		
59	Small mobile epifauna (swarming crustaceans)		
60	Small infauna (polychaetes)	Deposit feeders (infauna) (Y2) Infaunal predators (Y5)	
61	Sessile epifauna	Filter feeders (epifauna / infauna) (Y3)	
62	Meiofauna	Meiofauna (Y4)	
63	Benthic microflora (incl. Bacteria, protozoa))	Aerobic benthic bacteria (H1) Anaerobic benthic bacteria (H2)	
64	Planktonic microflora (incl. Bacteria, protozoa)	Microzooplankton (Z5) Heterotrophic flagellates (Z6) Bacteria (B)	nanoflagellates of up to 20 micrometer SED
65	Phytoplankton (autotrophs)	Diatoms (P1) Flagellates (P2) Picophytoplankton (P3) Large phytoplankton (P4) Small diatoms (P5)	

Table 2

Scenario	N Concentration	P Concentration	Warming
A1B1-High Nut.	10 mmol m ⁻³	1.0 mmol m ⁻³	+3.2 C
A1B1 - Low Nut.	5 mmol m ⁻³	0.5 mmol m ⁻³	+3.2 C
B1-High Nut.	10 mmol m ⁻³	1.0 mmol m ⁻³	+2.4 C
B1 - Low Nut.	5 mmol m ⁻³	0.5 mmol m ⁻³	+2.4 C

Figure captions

Fig 1 Outline of the information sources in the form of XML files and model components between two models running in network mode with a twin copies of Couplerlib.

Fig. 2 Sequence of operations when GOTM-ERSEM is linked with Ecopath with Ecosim (EwE) in managed model. Grey box is part of model that is run repeatedly on a daily timestep.

Fig. 3 GOTM-ERSEM prediction for the dynamics of 4 Plankton types – a.) Diatoms, b.) Flagellates c.) Dinoflagellates and d.) Omnivorous mesozooplankton within a water column over a two year period using synthetic meteorological data 1970-1971.

Fig. 4 Dynamics of size commercial fish species under the linked ERSEM –EwE model 1960-2000 simulation.

Fig. 5 Changes in relative biomass of 7 classes of functional groups under differing conditions of warming and nutrient availability (see table 2 for definitions.)







