

## The Marine Virtual Laboratory (version 2.1): enabling efficient ocean model configuration

Peter R. Oke<sup>1</sup>, Roger Proctor<sup>2</sup>, Uwe Rosebrock<sup>1</sup>, Richard Brinkman<sup>3</sup>, Madeleine L. Cahill<sup>1</sup>, Ian Coghlan<sup>4</sup>, Prasanth Divakaran<sup>5</sup>, Justin Freeman<sup>5</sup>, Charitha Pattiaratchi<sup>6</sup>, Moninya Roughan<sup>7</sup>, Paul A. Sandery<sup>5</sup>, Amandine Schaeffer<sup>7</sup>, and Sarath Wijeratne<sup>6</sup>

<sup>1</sup>CSIRO Oceans and Atmosphere, Hobart, Australia

<sup>2</sup>University of Tasmania, Hobart, Tasmania, Australia

<sup>3</sup>Australian Institute of Marine Science, Townsville, Australia

<sup>4</sup>Water Research Laboratory, School of Civil and Environmental Engineering, UNSW Australia, Sydney Australia

<sup>5</sup>Bureau of Meteorology, Melbourne, Australia

<sup>6</sup>University of Western Australia, Perth, Australia

<sup>7</sup>School of Mathematics, UNSW, Sydney, Australia

*Correspondence to:* Peter R. Oke (peter.oke@csiro.au)

**Abstract.** The technical steps involved in configuring a regional ocean model are analogous for all community models. All require the generation of a model grid, preparation and interpolation of topography, initial conditions, and forcing fields. Each task in configuring a regional ocean model is straight-forward - but the process of downloading and reformatting data can be time-consuming. For an experienced modeller, the configuration of a new model domain can take as little as a few hours - but for an inexperienced modeller, it can take much longer. In pursuit of technical efficiency, the Australian ocean modelling community has developed the Web-based MARine Virtual Laboratory (WebMARVL). WebMARVL allows a user to quickly and easily configure an ocean general circulation or wave model through a simple interface, reducing the time to configure a regional model to a few minutes. Through WebMARVL, a user is prompted to define the basic options needed for a model configuration, including the: model, run duration, spatial extent, and input data. Once all aspects of the configuration are selected, a series of data extraction, reprocessing, and repackaging services are run, and a “take-away bundle” is prepared for download. Building on the capabilities developed under Australia’s Integrated Marine Observing System, WebMARVL also extracts all of the available observations for the chosen time-space domain. The user is able to download the take-away bundle, and use it to run the model of their choice. Models supported by WebMARVL include three community ocean general circulation models, and two community wave models. The model configuration from the take-away bundle is intended to be a starting point for scientific research. The user may subsequently refine the details of the model set-up to improve the model performance for the given application. In this study, WebMARVL is described along with a series of results from test cases comparing WebMARVL-configured models to observations and manually-configured models.

It is shown that the automatically-configured model configurations produce a good starting point for scientific research.

## 1 Introduction

Increasingly, model simulations supported by data are becoming fundamental to marine research. Data alone is usually too sparse to provide the complete picture of time and space variations common to the marine environment. Likewise, models by definition are an approximation to the real world and need to be informed by data. Tools combining models and observations offer the researcher or manager ways to optimise and synergise the benefits offered by both data streams.

Setting up models and bringing together the necessary data for initialisation, forcing and validation is a time-consuming and frequently laborious activity, often taking months to achieve a starting point for scientific investigation. Previous developments have established tools for efficiently configuring specific models. For example, Penven et al (2007) describe a system for configuring a model for the Regional Ocean Model. Real-time applications, including support for defence services and search and rescue, motivated the development of tools enabling the rapid configuration and deployment of regional ocean forecasts Rosebrock et al (2015). Building on this capability, a new infrastructure - called the Web-based Marine Virtual Laboratory (WebMARVL; [www.marvl.org.au](http://www.marvl.org.au); [portal.marvl.org.au](http://portal.marvl.org.au)) - has been developed. A schematic diagram showing the services and options offered by WebMARVL is presented in Figure 1. Automated and semi-automated components of the WebMARVL workflow increases functionality and significantly reduces the startup time of any project involving simulation studies of the ocean.

WebMARVL allows a user to quickly and easily configure an ocean general circulation or wave model through web-based portal. The user is prompted to select the model, temporal extent of the model run, spatial extent, and input data. WebMARVL currently supports three community ocean general circulation models, and two community wave models, namely:

- Regional Ocean Modelling System (ROMS; Shchepetkin and McWilliams (2005));
- Modular Ocean Model (MOM; Griffies (2009));
- Sparse Hydrodynamic Ocean Code (SHOC; Herzfeld (2009));
- Simulating WAVes Nearshore mode (SWAN; Booij et al (1999)); and
- WaveWatch III (WW-III; Tolman (2002)).

The temporal extent of the model runs are limited to historical periods - not currently permitting forecasts.

Once all aspects of the configuration are selected, a series of services automatically perform the data extraction, reformatting, and repackaging. WebMARVL produce a “take-away bundle” that

contains all of the input fields and forcing fields in the correct format to be immediately used to run the chosen model (i.e., ROMS or MOM etc). The take-away bundle also includes all of the observations, exploiting capabilities developed under the Australian Integrated Marine Observing System (IMOS; [www.imos.org.au](http://www.imos.org.au); Hill et al (2010) Proctor et al (2010) Hidas et al (2015)), in the chosen time-space domain. The observations are intended to be used for model assessment and/or data assimilation. The take-away bundle is available to the user via direct download from WebMARVL. If a user is to manually undertake each step of gathering topographic data, ocean input data (from a global ocean model run or reanalysis, for initial conditions and nesting), surface forcing data (from an atmospheric model run or reanalysis), and observational data - and then write a series of scripts to “massage” the data into appropriate formats, the process of setting up a model can take days for an expert and months for a non-expert. By contrast, it takes 5-10 minutes, depending on the computer load, to generate a take-away bundle using WebMARVL. The model configuration from the take-away bundle is intended to be a starting point. It is expected that the expert user will refine the model set-up (e.g., mixing schemes, adjustments to topography or coastlines etc.) to improve the model performance for the given application.

The WebMARVL developments are motivated to deliver benefits to the marine community by enabling:

- efficient configuration of a range of ocean and wave models;
- efficient model inter-comparisons;
- assessment of the sensitivity studies, comparing model results with different model parameterisations (e.g., mixing schemes) and configurations (e.g, oceanic and atmospheric forcing);
- efficient configuration of ensemble prediction (in a historical sense) - to help quantify uncertainty - using a single model with different input data, or using multiple models;
- comprehensive model evaluation through model-observation comparisons;
- downscaling of large scale model runs, including climate projections and seasonal forecasts;
- and
- efficient access to and consolidation of observational data for a selected time-space domain.

In this paper, we describe WebMARVL in more detail in section 2, followed by results in section 3, and an example of the data generated by the take-away bundle in section 4. The results presented below typically compare output from manually-configured models to WebMARVL-configured models.

Some test cases also compare WebMARVL-configured models to observations. We seek to demonstrate that the WebMARVL-configured model runs are credible - providing a good foundation for scientific research. We expect that an expert modeller can generate a take-away bundle using WebMARVL, and subsequently refine the model setup to produce model results of suitable quality to

permit high-quality scientific research. Indeed, it is shown in this paper that the automatically-  
90 configured model configurations produce credible results. The paper concludes with a short discussion and summary in section 5.

## 2 WebMARVL

The development of WebMARVL was largely funded under the Australian National Collaborative Research Infrastructure Strategy, through the National eResearch Collaboration Tools and Research  
95 (NeCTAR; [www.nectar.org.au](http://www.nectar.org.au)) program. NeCTAR is a program motivated to establish research software infrastructure to promote and support collaborative research outcomes between Australian researchers. Consequently, the initial version of WebMARVL is restricted to Australian-based researchers, with user authentication provided securely via the Australian Access Federation (AAF; Figure 2). This means that researchers with logins at Australian Universities (e.g., UTAS), and Aus-  
100 tralian research organisations (e.g., CSIRO) simply use their institutional login name and password to log into WebMARVL. An incomplete list of the eligible universities is evident in Figure 2. This aspect of WebMARVL can easily be changed to include international users (e.g., by including AAF Virtual Home, [vho.aaf.edu.au](http://vho.aaf.edu.au)) with a modest additional development.

After logging in, a user is prompted to select the model (i.e., ROMS, MOM etc) from the list of  
105 supported models (see section 1), and is prompted to generate a model grid (Figure 3). The initial version of WebMARVL is restricted to the generation of rectangular grids, so a user effectively needs to specify the corners of the model grid, the angle of rotation, and the grid resolution. This can be done manually - by typing in the desired reference points etc - or by a graphical user interface. The user can zoom in and out of the map in the interface to carefully define the grid of choice.  
110 WebMARVL also offers the option to import files containing your own model grid and bathymetry (for ROMS and SHOC only at this stage).

The user is then prompted to define the temporal extent of the model run using a sliding bar (Figure 3). As the user modifies the temporal extent, the choice of input data - including the surface forcing data and ocean data - automatically updates. WebMARVL maintains a number of data prod-  
115 ucts that can be used for surface forcing and ocean data. The available atmospheric data sets, to be used for surface forcing, include:

- archived fields from analyses and nowcasts using Australia's operational global and regional atmospheric prediction systems (ACCESS-G and ACCESS-R; Puri et al (2010)); and
- the NCEP/NCAR atmospheric reanalysis (NCEP1 Kalnay et al (1996)).

120 The global wave forcing available through WebMARVL is a reanalysis of AUSWAVE Durrant et al (2013), a global configuration of WW-III Tolman (2002). The available ocean data sets, to be used for initial conditions and nesting, include:

- version 3p5 of BRAN (BRAN3p5 Oke et al (2013));
- archived nowcasts and analyses from OceanMAPS Brassington et al (2007), Australia’s operational short-range ocean forecast system;
- the UK Forecasting Ocean Assimilation Model (FOAM Blockley et al (2012)); and
- the Mercator Ocean GLORYS reanalysis Ferry et al (2007).

Of these ocean data sets, OceanMAPS and FOAM are archives of nowcasts produced operationally by the Australian BoM and the UKMet Office; and others are from model runs performed under Bluelink, a partnership between CSIRO, the BoM and the Royal Australian Navy Schiller et al (2009), or provided by the Mercator consortium. As the user adjusts the temporal domain, only the data sets that span the full requested temporal extent are available for selection. Once the user selects the input data, the request is submitted to WebMARVL, and a series of scripts are run automatically to perform the following steps:

- data verification - checking that the chosen data sets have the mandatory variables for the chosen model;
- data extraction - including the forcing data and observations;
- reformatting - to put the netcdf files in the format required by the chosen model, including variable and dimension names, grid constructions etc; and
- repackaging - to collate all of the required data into a single, easy-to-download package.

Data extraction is achieved through a combination of local files and OPenDAP. Note that the data extraction step, listed above, includes interpolation of fields for some models. For applications of ROMS and MOM, fields are interpolated onto the model grids using bi-linear interpolation horizontally and linear interpolation vertically, using a spherical map projection. For applications of SHOC, SWAN, and WW-III, fields from the original parent grids are supplied to the model. In addition, for applications to SWAN, fields are bi-linearly interpolated along the open boundaries to provide boundary forcing. All calculations by WebMARVL are performed on a virtual machine, exploiting readily available open source software (e.g., GNU compilers, netcdf libraries and operators, gridgen). After all WebMARVL tasks are complete, the user can access their take-away bundle directly from the web portal. The user is then free to use and modify the data in the take-away bundles as needed for their specific application.

### 3 Results

In this section, we present results from a series of test cases (Figure 4) performed by researchers from different Australian research institutions for a range of applications. The MARVL project is a community project - with developers working with users (ocean modellers) to develop a tool that is fit for

purpose. Part of the project has involved testing by users in realistic scenarios, with feedback provided to the developers. The results presented below are examples of the test cases performed under this project during the development of WebMARVL. This includes one test case using SHOC, one using MOM, two using ROMS, and two using SWAN. Although WebMARVL supports applications with WW-III, a test case for WW-III is not presented in this paper. The results presented here typically compare output from manually-configured models to WebMARVL-configured models. Some test cases also compare WebMARVL-configured models to observations. We seek to demonstrate that the WebMARVL-configured model runs are credible. We expect that an expert modeller can generate a take-away bundle using WebMARVL, and subsequently refine the model setup to produce model results of suitable quality to permit high-quality scientific research.

### 3.1 Southern Great Barrier Reef - SHOC

Motivated to better understand what factors affect upwelling and cross-shelf exchange on the Southern Great Barrier Reef (SGBR), particularly near Heron Island, a manually- and WebMARVL-configured SHOC run is performed for a domain over the SGBR (Figure 4) for the period November 2008 to March 2009. This period spans a series of events when cold water is observed in temperature records obtained from in situ moorings to be uplifted onto the reef, representing strong cross-shelf exchange. Results from BRAN3p5 Oke et al (2013), with  $\sim 10$  km grid spacing, do not show this same uplift as observed. To determine the extent to which reanalysis represents the dominant processes, a high-resolution regional model is configured to see if it could represent the uplift. The model grid is 2 km in both configurations. The manually-configured run is forced at the surface by wind stress and bulk heat fluxes using fields from ERA-Interim, and the model initial conditions and boundary fields are derived from BRAN3p5. The WebMARVL-configuration is forced at the surface by wind stress and bulk heat fluxes using fields from ACCESS-R, and ocean fields from BRAN3p5. Unlike the BRAN3p5 runs, the regional configuration of SHOC includes tidal forcing at the boundaries. Neither configuration includes freshwater fluxes.

Results are shown in Figure 5, showing initial conditions (panels a-b) and model fields after 18 days of integration (panels c-d), during a strong upwelling/uplifting event. Note that although the model resolution is the same in the manually-configured and the WebMARVL-configured model, the grids are slightly rotated relative to each other. Due to this rotation the bottom topography along the two cross-shelf transects are slightly different at the offshore end.

Recall that the goal of these comparisons is to establish the credibility of WebMARVL. To this end, we note that the initial conditions are comparable in both configurations (Figure 5a-b). In part, the differences in initial conditions is because the WebMARVL-configured run is initialised 4 days earlier to allow the tides to somewhat equilibrate. But note that WebMARVL does not generate tide boundary conditions. The sub-surface model fields during the upwelling/uplifting events are also very similar (Figure 5c-d) - both showing sub- $22^\circ$  temperatures uplifted onto the reef (around

50 m depth). However, the near surface temperatures are quite different, with the WebMARVL-configured run showing warmer surface temperatures, owing to known biases in the fluxes derived from ACCESS-R.

195 Based on these results, we conclude that the  $\sim 10$  km resolution model lacks sufficient resolution to represent the uplift of cold water onto the reef over the SGBR. Additionally, the comparisons demonstrate that for this case, the WebMARVL-configured model reproduces results that compare well to results from a manually-configured model with comparable resolution and forcing.

### 3.2 Southeastern Australia - ROMS

200 Motivated to assess WebMARVL in a region of complex topography in the presence of a strong western boundary current - namely the East Australian Current (EAC) - a 1-km resolution configuration of ROMS is applied to a region off Coffs Harbour ( $\sim 30^\circ\text{S}$ ; Figure 4) in the vicinity of the Solitary Island Marine Park. This region is of particular interest because it is the focus of several national and state programs (e.g., National Environmental Science Programme, [www.environment.gov.au/science/nesp](http://www.environment.gov.au/science/nesp)) and is the location of both state and federal marine reserves. The region is also  
205 covered by an array of land-based High Frequency (HF) radars - returning high spatial and temporal resolution observations of surface currents.

Briefly, WebMARVL is used to configure ROMS with 1 km horizontal resolution, 14 vertical levels, with surface forcing from ACCESS-R, and ocean initial conditions and forcing from BRAN3p5.

210 The model is run for the period April to July 2012.

When compared to results from BRAN3p5, the WebMARVL-configured ROMS run produced alongshore currents that are more baroclinic (with 50% stronger alongshore currents) and with more variability in surface temperature (ROMS ranging from  $11.2\text{--}26.7^\circ\text{C}$ ; and BRAN3p5 ranging from  $18\text{--}25.5^\circ\text{C}$ ), due to better-resolved processes.

215 The model results (Figure 6a) agree reasonably well with observations from different platforms, including satellite sea surface temperature (with differences of  $0.5\text{--}0.8^\circ\text{C}$ ), moorings (with upper ocean temperature differences of  $0.8^\circ\text{C}$  and correlations of 0.86; and velocity differences of  $0.11\text{--}0.26$  m/s, with correlations of 0.41-0.45), and with HF radar-derived surface velocities. In particular, the WebMARVL-configured ROMS run shows very detailed sub-mesoscale variability. Figure 6a  
220 shows an example of the EAC instabilities at the inshore edge. These frontal vortices and eddies are important for driving vertical dynamics and ultimately nutrient supply and biological activity. Both the modelled and observed velocity fields derived from HF radars show such sub-mesoscale eddy structures on the edge of Coffs Harbour (Figure 6b,c; see also Figure 3 in Roughan et al (2015)). Although the precise location of the eddy differs, its size and intensity compare qualitatively  
225 well. This test case demonstrates that a WebMARVL-configured high-resolution ROMS run can realistically reproduce both the mesoscale and sub-mesoscale variability in a complex region of strong currents. Analysis of the WebMARVL-configured ROMS is ongoing, and promises to provide

important insights into the cross-shelf processes, upwelling and connectivity in the Solitary Island Marine Park area off eastern Australia.

### 230 3.3 Southwestern Australia - ROMS

To assess the performance of a WebMARVL-configured model for a very high-resolution application, WebMARVL is used to configure ROMS with  $\sim 500$  m horizontal resolution, 13 vertical levels, with surface forcing from ACCESS-R, and ocean initial conditions and forcing from BRAN3p5. The model is set up to cover part of the Perth Canyon and Rottnest continental shelf, off southwestern  
235 Australia (Figure 4). The model is run for the period July to October 2011. At the time of testing, WebMARVL did not provide fields of evaporation and precipitation. For this case however, the users supplemented the WebMARVL take away bundle with these fields from ERA-Interim. The model is also forced at the boundary with tidal forcing, using eight primary harmonic constituents (specified by the user - not from WebMARVL).

240 Model results are compared to near-surface velocities from HF radar and sub-surface velocities from a moored ADCP. Results of comparisons between modelled and observed fields are presented in Figure 7. This includes a comparison of daily-averaged modelled and observed near-surface velocities on 13 and 16 July 2011 - when there is a distinct cyclonic eddy present on the shelf edge (Figure 7a-b). These comparisons show that the model realistically reproduces this dominant feature,  
245 with good agreement in the location, strength, and size of the eddy. A comparison of time-series of the velocity at 10 m depth, measured by a moored ADCP, from the IMOS Rottnest Island National Reference Station, is presented in Figure 7c-d. These comparisons show very good agreement between the observed and modelled velocities.

The results from this test case demonstrate that with some additional refinement of the WebMARVL-  
250 configured very high resolution ROMS, model results in excellent agreement with observations can be achieved. Fields from these model runs are being analysed in detail to examine the dynamics of deep water overflows that frequently occur off Western Australia due to winter cooling Pattiaratchi et al (2011).

### 3.4 Southeastern Australia - MOM

255 To assess the performance of WebMARVL for setting up a high-resolution configuration of MOM4p1, a 2-km resolution model off southeastern Australia (Figure 4) is set up. The model has 51 vertical levels and is nested within OceanMAPS and forced at the surface with atmospheric fields from ACCESS-G. The fields presented in Figure 8 show the initial conditions - representing an interpolation from the  $\sim 10$ -km resolution grid of OceanMAPS, to the 2-km resolution grid of the regional  
260 domain. Also shown in Figure 8 are the model fields after 16 days of simulation. Evident in Figure 8 is that the high-resolution model adds significant detail to the fields - with the generation of fine-scale filaments and sub-mesoscale eddies. Although this demonstration doesn't assess the skill of



the high-resolution configuration, it does demonstrate that the WebMARVL can be used as a starting point for a regional configuration of MOM.

### 265 3.5 Southeastern Australia - SWAN

To assess the suitability of WebMARVL to configure a very high-resolution regional wave model, two domains are considered. One domain is off Sydney ( $\sim 33.75^\circ\text{S}$ ), and the second is off Coffs Harbour ( $\sim 30.3^\circ\text{S}$ ). The approximate locations of the model domains are shown in Figure 4. The Sydney test case is motivated to assess the performance of the WebMARVL-configured SWAN model near  
270 the coast, in shallow water. The Coffs Harbour test case is motivated to assess the performance farther offshore, in deeper water.

The Sydney configuration has 50 m horizontal resolution, spanning a  $\sim 13 \times 22$  km rectangular grid offshore of the northern beaches of Sydney. Two WebMARVL configurations are tested for this domain: one forced by 12 km-resolution ACCESS-R winds (hereafter MARVL SWAN: BoM)  
275 and one forced with 20 km-resolution NCEP winds (hereafter MARVL SWAN: NCEP). These runs were compared to a manually-configured SWAN model by Cardno (2011) Cardno (2011), hereafter Manual SWAN: OEH (NSW Office of Environment and Heritage). MARVL SWAN: BoM is forced with winds from ACCESS-R and nested inside AUSWAVE Durrant et al (2013). Manual SWAN: OEH is forced with winds from the NCEP/NCAR atmospheric reanalysis and nested inside a series  
280 of four coarser grids (a SWAN “transfer” grid and state, national and global configurations of WW-III). Established in 2011, the Manual SWAN: OEH model is carefully configured and tuned for the NSW coastline (including Sydney and Coffs Harbour regions) including adjustments to forcing NCEP/NCAR wind speeds within 100 km of the NSW coast. Time series of observed significant wave height approximately 1 km offshore in 21 m of water are compared with Manual SWAN: OEH  
285 model time series at the same location and MARVL SWAN: BoM from the nearest available wet cell (1.4 km seaward of the wave buoy) in Figure 9. Although the WebMARVL configured SWAN run produces results that are qualitatively consistent with observations (i.e., generally following the “shape” of wave episodes measured by the wave buoy), the automatically-configured model significantly under-predicts peak wave heights for individual storm events in shallow water. Similar  
290 performance was observed with the MARVL SWAN: NCEP model (not shown). The key reason for this difference is errors in topography and the position of the model coastline which is generally located several kilometres offshore of the true coastline (i.e., 0 m contour). WebMARVL extracts topography from the Australian Bathymetry and Topography Grid Whiteway (2009) produced by Geoscience Australia, which has resolution of 9 arc seconds ( $\sim 280 \times 280$  m), and is too coarse to be  
295 useful for very high-resolution applications in shallow water, close to the coastline in this region.

The Coffs Harbour configuration has 100 m horizontal resolution, spanning a  $\sim 39 \times 31$  km rectangular grid offshore off Coffs Harbour. Two WebMARVL configurations are again tested - one forced by 12 km-resolution ACCESS-R winds (hereafter MARVL SWAN: BoM) and one forced

with 20 km-resolution NCEP winds (hereafter MARVL SWAN: NCEP). These runs are compared to a another manually-configured SWAN model forced with adjusted winds from the NCEP/NCAR atmospheric reanalysis (hereafter Manual SWAN: OEH). Like the application off Sydney, the location of the model coastline is misaligned with the true coastline position.. This leads to errors like those described above near the coast (not shown). Time series of modelled and observed significant wave height about 12 km offshore, in 72 m of water, are compared in Figure 10. These results show that the WebMARVL-configured model results are in good agreement with observations at an offshore site in deep water under ambient wave conditions (significant wave heights less than 3 m). These comparisons demonstrate that the automated model configurations produce overall results that are comparable to those produced by the manually developed SWAN model in this area over an extensive period of time.

The test cases presented here lead to the conclusion that the WebMARVL-configured wave model is suitable for operational climate science modellers (i.e., relatively deep water applications, say  $> 50$  m depth), but is not recommended for engineering design and/or shallow water applications. The primary reasons for this conclusion are due to limitations in the topography and less reliable performance under extreme storm wave conditions (infrequent events). The topography that currently underpins WebMARVL is too coarse ( $\sim 280$  m resolution) for very-high resolution (50-100 m grids) near the coast – where the location of the coastline is often poorly represented in the automatically-configured models. However, in order to overcome this limitation, WebMARVL also enables the user to upload their own topography if necessary.

#### 4 Take-away bundles

To give the reader an idea of the data that is typically provided in the take-away bundles produced by WebMARVL, we provide some technical details of one of the examples above. For this demonstration, we report on the take-away bundle associated with the MOM example (Example 3.4). Files generated for MOM include:

- grid\_spec.nc: a netcdf data file containing information about the model grid and topography;
- mom.nc: a netcdf data file containing four-dimensional fields of temperature, salinity, velocity and sea-level that are used to force the model (i.e., boundary forcing and optionally sponge forcing); and three-dimensional (horizontal dimensions and time) fields of wind stress, heat flux components, freshwater flux components (excluding river forcing);
- ocean\_temp\_salt.res.nc: a netcdf data file containing three-dimension (horizontal and vertical dimensions) of temperature, salinity, velocity and sea-level that are used to initialise the model;
- data\_table: a text file containing logical flags, variable names, and parameters to define the surface forcing for the model and to define what output the model will generate;

- ocean\_solo.res: a text file containing information about the model start date, run length and calendar type.

335 Note that the mom.nc file, listed above, contains many fields that are commonly stored in separate files for MOM runs. These includes the sponge files, the restore files, the force files. In addition to these files, the model requires a series of namelists to be defined to specify the options defining the model’s numerics, forcing, etc. These namelists are generally defined in a run script that is prepared by the user.

## 340 **5 Discussion and summary**

A new web-based tool, called WebMARVL, has been developed to increase the efficiency of configuring a regional ocean model. The development of WebMARVL was undertaken by CSIRO and UTAS, and tested by a cross-section of the Australian coastal ocean modelling community. Tests of the system were performed by researchers from seven different research organisations, using five  
345 different community models (i.e., SHOC, ROMS, MOM, SWAN, and WW-III). The time taken to configure any of the chosen regional models using WebMARVL is just a few minutes. This contrasts to the traditional manual approach, taking days to months - depending on the proficiency of the modeller.

In this paper, we describe the technical steps that are automated by WebMARVL to configure  
350 each regional application. We provide a series of inter-comparisons between established, manually configured models, and between models and observations. Each WebMARVL-configured model set-up is not intended to deliver an optimised model configuration. It is intended to efficiently deliver a “good starting point” for further model development and refinement. The examples presented in this study demonstrate the credibility of WebMARVL - producing model configurations that are a good  
355 starting point for scientific research. WebMARVL is freely available to researchers associated with Australian research organisations, and there are plans to make it available globally. It is anticipated that by delivering improved efficiency to the coastal ocean modelling community, WebMARVL will become a valuable research tool - helping experienced, and student modellers to more quickly get to the heart of a scientific study.

## 360 **6 Code Availability**

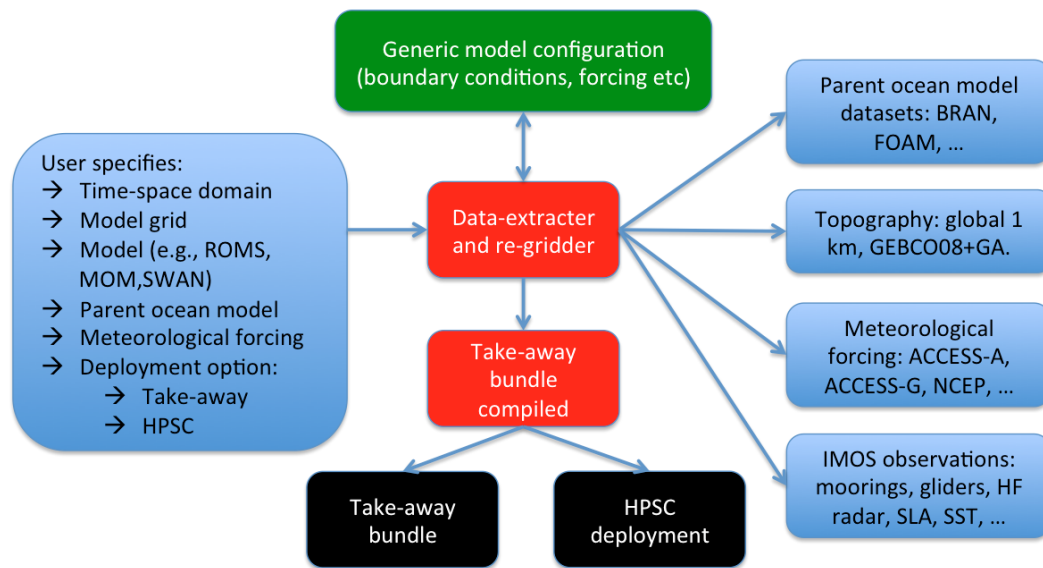
MARVL software (version 2.1) is open source code that is available under an MIT license. However, software for all of the underpinning models are available under open source. ROMS software (version 3.4, revision 633) is available from [www.myroms.org](http://www.myroms.org). MOM software (version 4p1) is available from [mom-ocean.org/web](http://mom-ocean.org/web). SHOC software (version 1.1, revision 4718) is available from [www.emg.cmar.csiro.au](http://www.emg.cmar.csiro.au). SWAN software (version 40.91AB) is available from [swanmodel.sourceforge.net](http://swanmodel.sourceforge.net).  
365

Gridgen software is available from [anaconda.org/timcera/gridgen-c](https://anaconda.org/timcera/gridgen-c). Netcdf Operators software is available from [nco.sourceforge.net](https://nco.sourceforge.net). GNU compilers are available from [gcc.gnu.org](https://gcc.gnu.org).

*Acknowledgements.* Funding for this research was provided by National eResearch Collaboration Tools and Research (NeCTAR; sub-contract VL206), UTAS and CSIRO. Data was sourced from the Integrated Marine  
370 Observing System (IMOS) - IMOS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative. Wave buoy observations were collected and provided by Manly Hydraulics Laboratory (MHL). Wave buoy observations and the manual configured SWAN models are owned the NSW Office of Environment and Heritage (OEH). The authors also acknowledge J. Luick and R. Morrison, who contributed to the MARVL project as testers.

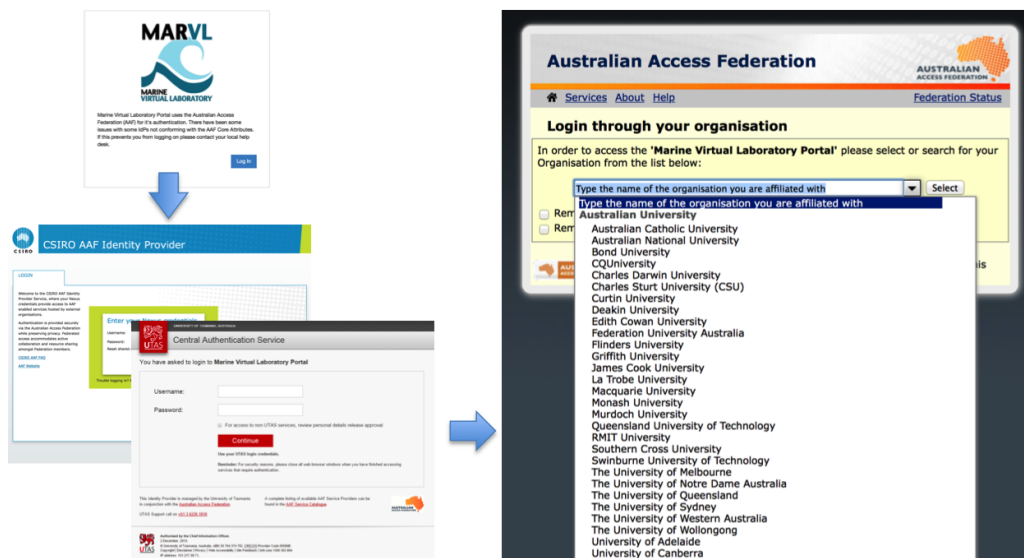
## 375 References

- Blockley EW, Martin MJ, Hyder P (2012) Validation of FOAM near-surface ocean current forecasts using Lagrangian drifting buoys. *Ocean Science* 8:551–565, doi:doi:10.5194/os-8-551-2012
- Booij N, Ris RC, Holthuijsen LH (1999) A third-generation wave model for coastal regions: 1. model description and validation. *Journal of Geophysical Research: Oceans* 104(C4):7649–7666, doi:10.1029/98JC02622,   
380 <http://dx.doi.org/10.1029/98JC02622>
- Brassington GB, Pugh TF, Spillman C, Schulz E, Beggs H, Schiller A, Oke PR (2007) Bluelink development of operational oceanography and servicing in Australia. *J Research Pract Inf Tech* 39(2):151–164
- Cardno A (2011) NSW coastal waves: Numerical modelling. final report. In: Final Report, Prepared for Office of Environment and Heritage (NSW), September, p 122pp
- 385 Durrant TH, Greenslade DJ, Simmonds I (2013) The effect of statistical wind corrections on global wave forecasts. *Ocean Modelling* 70:116–131
- Ferry N, Remy E, Brasseur P, Maes C (2007) The Mercator global ocean operational analysis system: Assessment and validation of an 11-year reanalysis. *Journal of Marine Systems* 65(1-4):540–560, doi:10.1016/j.jmarsys.2005.08.004
- 390 Griffies SM (2009) Elements of MOM4p1, GFDL Ocean Group Technical Report 6. Tech. rep., NOAA/Geophysical Fluid Dynamics Laboratory
- Herzfeld M (2009) Improving stability of regional numerical ocean models. *Ocean Dynamics* 59(1):21–46, doi:10.1007/s10236-008-0158-1
- Hidas MG, Proctor R, Atkins N, Atkinson J, Besnard L, Blain P, Bohm P, Burgess J, Fruehauf D, Galibert G,   
395 Hoenner X, Hope J, Jones C, Mancini S, Pasquer B, Nahodil D, Reid K, Tattersall K (2015) Infrastructure and workflows for australia’s integrated marine observing system. *Ocean Dynamics* submitted, this issue
- Hill K, Moltmann T, Proctor R, Allen S (2010) The Australian integrated marine observing system: Delivering data streams to address national and international research priorities. *Marine Technology Society Journal* 44(6):65–72
- 400 Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu J, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo K, Ropelewski C, Wang J, Leetmaa A, Reynolds RW, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* 77(3):437–471
- Oke PR, Sakov P, Cahill ML, Dunn JR, Fiedler R, Griffin DA, Mansbridge JV, Ridgway KR, Schiller A (2013)   
405 Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modelling* 67:52–70, doi:10.1016/j.ocemod.2013.03.008
- Pattiaratchi C, Hollings B, Woo M, Welhena T (2011) Dense shelf water formation along the south-west Australian inner shelf. *Geophysical Research Letters* 38(10)
- Penven P, Marchesiello P, Debreu L, Lefevre J (2007) Software tools for pre- and post-processing of oceanic   
410 regional simulations. *Environmental Modelling Software*, 23:660–662
- Proctor R, Roberts K, Ward B (2010) A data delivery system for IMOS, the Australian Integrated Marine Observing System. *Advances in Geosciences* 28(28):11–16
- Puri K, Xiao Y, Sun X, Lee J, Engel C, Steinle P, Le T, Bermous I, Logan L, Bowen R, et al (2010) Preliminary results from numerical weather prediction implementation of ACCESS. *CAWCR Research Letters* 5:15–22



**Figure 1.** Schematic outlining the steps in WebMARVL. User requirements are specified through the WebMARVL interface. The system then selects the grid, the model(s), the parent model(s) for initialisation and boundary conditions, the topography and meteorological forcing. The system subsequently interpolates the inputs to the specified grid requirements. Depending on the deployment option (a) the files are bundled up as a take-away package, or (b) deployed to a High Performance Super Computer (HPSC).

- 415 Rosebrock U, Oke PR, Carroll G (2015) An application framework for the rapid deployment of ocean models in support of emergency services: Application to the MH370 search. In: Environmental Software Systems. Infrastructures, Services and Applications, Springer, pp 235–241
- Roughan M, Schaeffer A, Suthers IM (2015) Chapter 6 - sustained ocean observing along the coast of southeastern australia: Nsw-imos 2007-2014. In: Weisberg YLKH (ed) Coastal Ocean Observing Systems, Academic Press, Boston, pp 76 – 98, doi:<http://dx.doi.org/10.1016/B978-0-12-802022-7.00006-7>
- 420 Schiller A, Meyers G, Smith N (2009) Taming australia’s last frontier. American Meteorological Society Bulletin 90(4):436–440
- Shchepetkin A, McWilliams JC (2005) The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling 9(4):347–404, doi:10.1016/j.ocemod.2004.08.002
- 425 Tolman HL (2002) User manual and system documentation of WAVEWATCH-III version 2.22, vol 133. NOAA/NWS/NCEP/MMAB Technical Note 222
- Whiteway TG (2009) Australian bathymetry and topography grid. Geoscience Australia Record 21:pp46



**Figure 2.** Screen-shots of sample WebMARVL login screens.

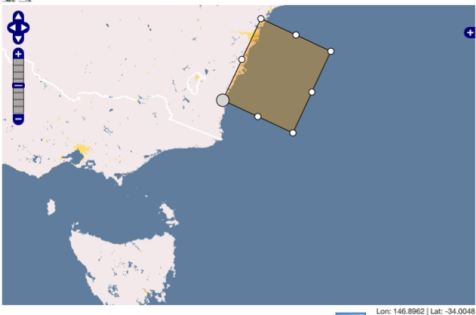
Model Run Specification Fill out the form fields and click next until all tabs are complete.

Previous Next

Model Grid Temporal Extent and Forcing

### Grid Definition

Origin is large white circle, Longitude and Latitude are WGS84



Lon: 146.8962 Lat: -34.0048

Forcing Datasets

Type: Shape: geographic-rectangular

Origin: X (Longitude): 149.996 Y (Latitude): -36.1

Size: X (N): 60 Y (N): 70

Spacing: X (dphi degrees): 0.0452 Y (dlambda degrees): 0.0452

Rotation: Angle (degrees): 335

Update Grid

Temporal Extent and Forcing Data

Set the temporal extent and select one forcing dataset from each group.

Temporal Extent: 2010-03-01 - 2010-06-09 (100 days)

Zoom Reset

Atmosphere: 2010-08-04 - 2013-01-02 (882 days)

access-a-surface

ncap 2009-08-18 - 2013-01-01 (1232 days)

Ocean: 2009-09-11 - 2013-01-09 (1216 days)

foam1-an

ofam2-1-an 2012-10-21 - 2012-12-26 (66 days)

ofam2-an 2011-11-02 - 2012-11-06 (372 days)

Submit

Run Name: Jervis Bay

Model: shoc\_wf

Grid: Type: geographic-rectangular D-Lambda: 0.0452 D-Phi: 0.0452 Rotation: 335 N: 60 Nj: 70 X-Origin: 149.996 Y-Origin: -36.1

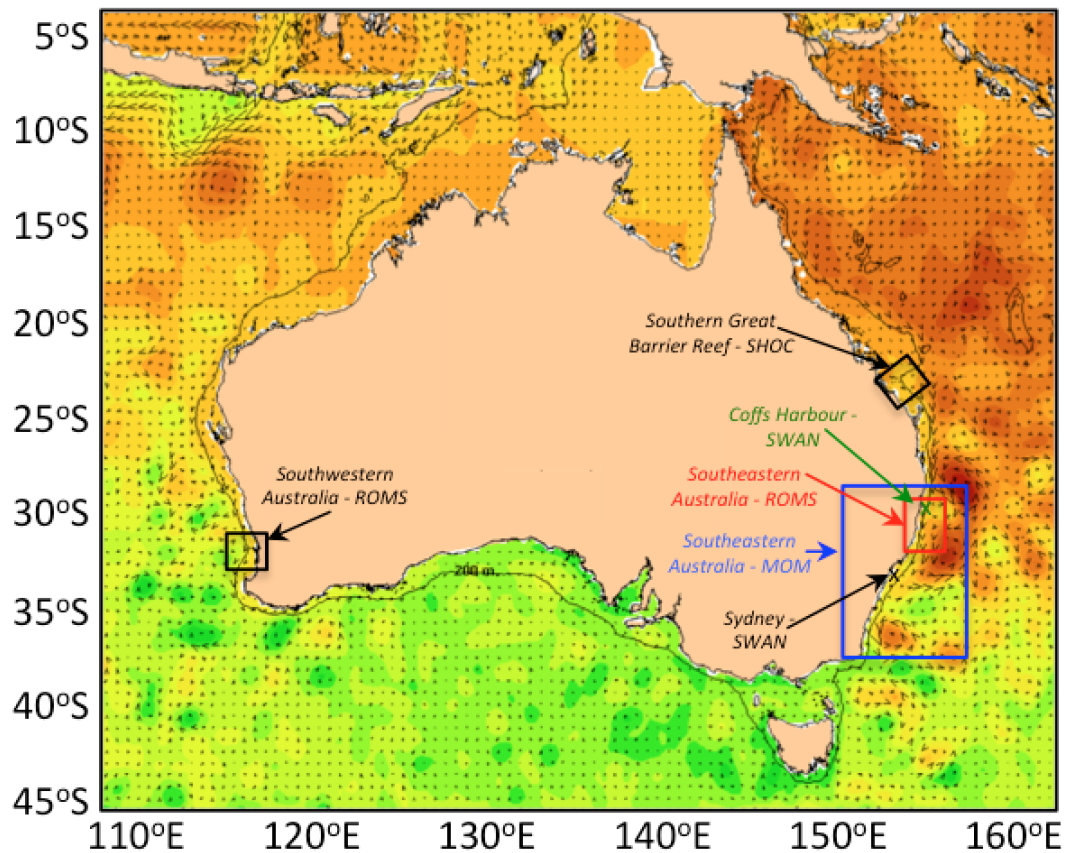
Temporal Extent: Start: 2010-03-01 00:00 +00 Duration: 30 days

Stream	Dataset	Variable Name	Description
global-atmos	ncap	mslp	Mean Sea Level Pressure
global-atmos	ncap	wind_u_wind_v	U direction wind speed
global-ocean	foam1-an	salt	salinity
global-ocean	foam1-an	eta	seasurface elevation
global-ocean	foam1-an	u,v	Eastward current
global-ocean	foam1-an	temp	temperature

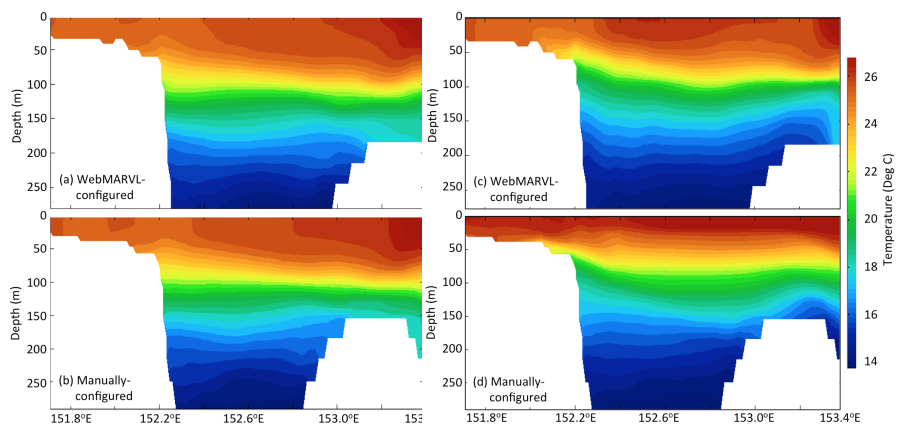
Submit Run

**Figure 3.** Screen-shots of WebMARVL showing the steps of (top-to-bottom) configuring a grid, setting the temporal extent and selecting input data, and submitting a job.

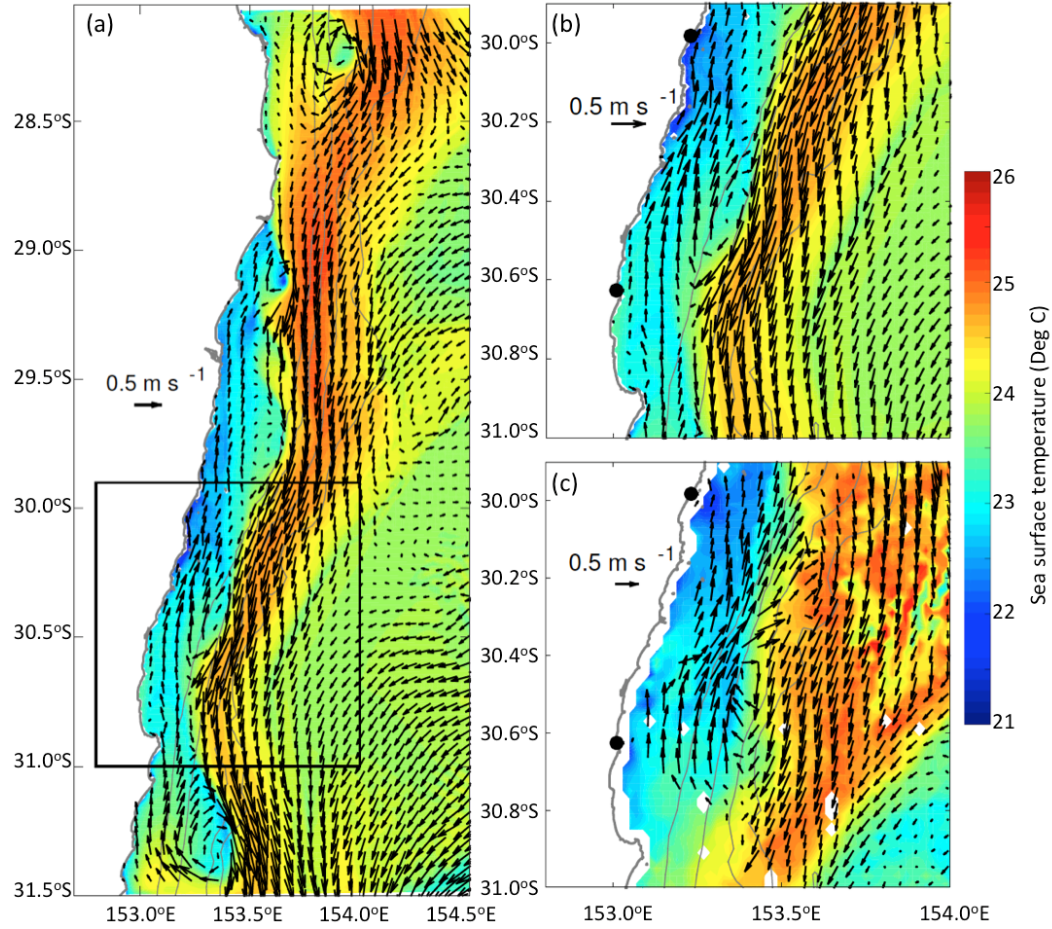




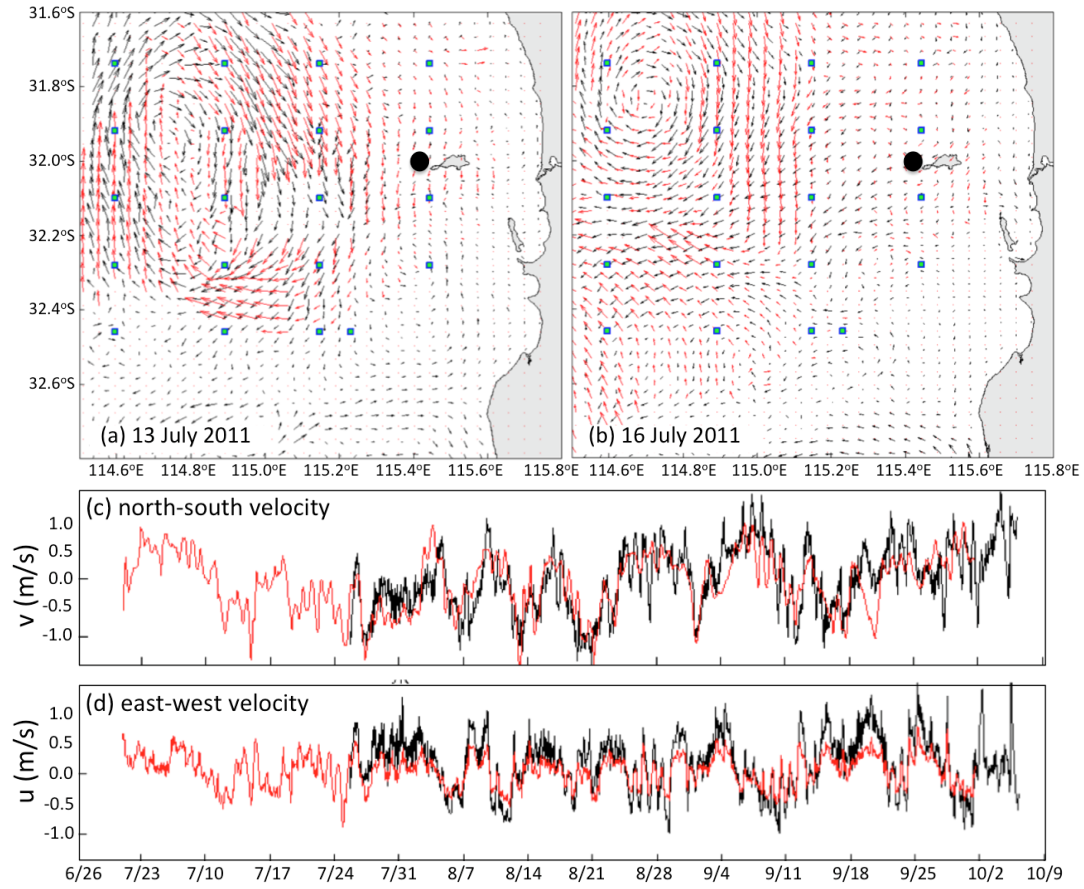
**Figure 4.** Map of Australia, showing the sea-level (colour) and geostrophic surface velocities (from [oceancurrent.imos.org.au](http://oceancurrent.imos.org.au)), with the approximate locations of each model domain used for testing. Labels, arrows, and boxes are colour-coded for clarity off eastern Australia.



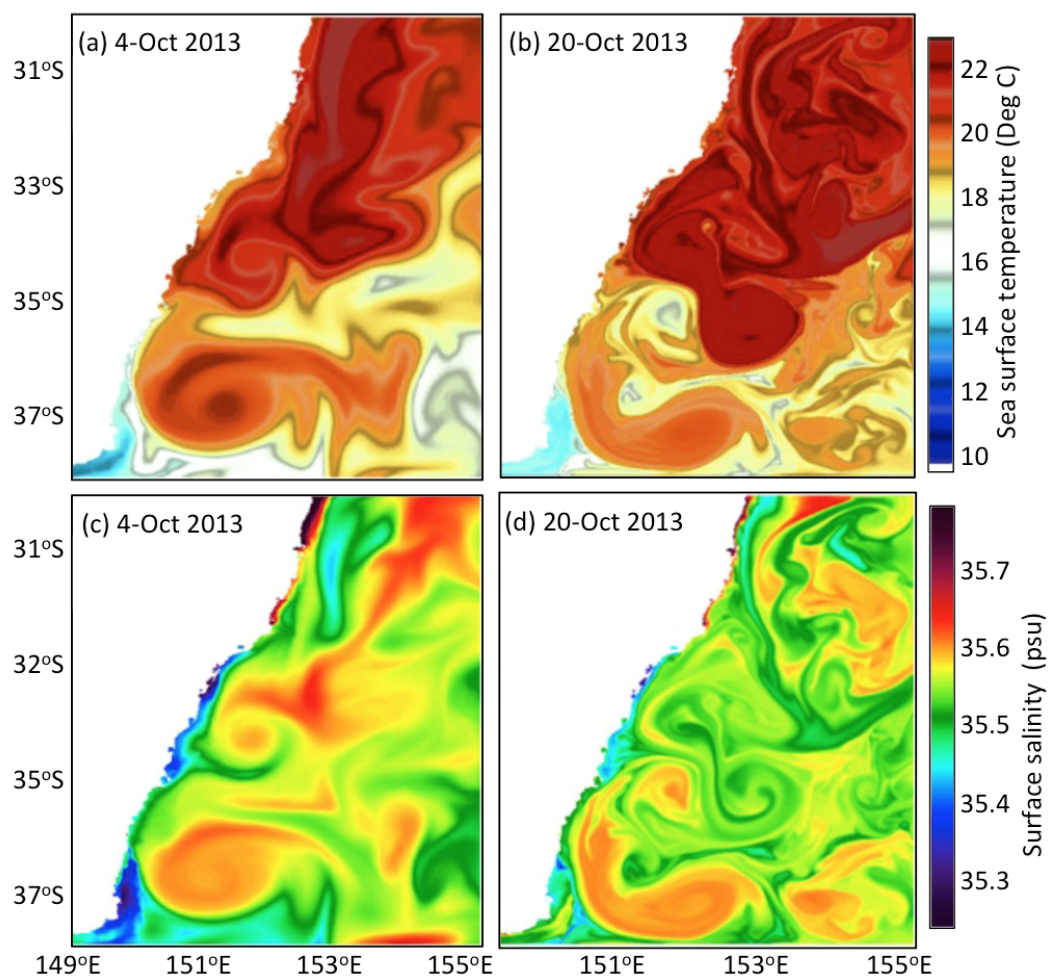
**Figure 5.** Comparison of a temperature section off the Southern Great Barrier Reef (Figure 4) from the SHOC model, using (a,c) a WebMARVL configuration and (b,d) a manual configuration for (a-b) 1 December 2008 and (c-d) 18 December 2008.



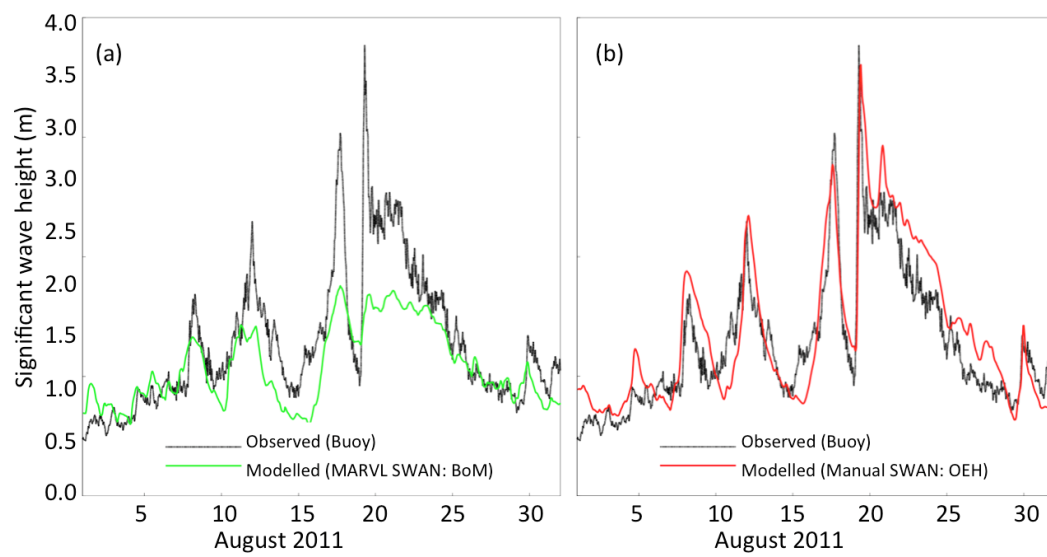
**Figure 6.** Snap-shots of (a,b) modelled and (c) observed surface velocity and SST for 15 May 2012. Model results are from a 1-km resolution WebMARVL-configured ROMS run off southeastern Australia (Figure 4), and observations are from 6-day composite satellite AVHRR SST and land-based HF radar measurements (the location of the HF radar sites are shown by black dot in panels (b) and (c)). The coastline and 100-, 200-, 1000-, 2000-m and 5000-m isobaths are shown in grey. For clarity, velocity vectors are only plotted every 5th grid point in panel (a), every 4th grid point in panel (b), and every 3rd point for the HF radar data in panel (c).



**Figure 7.** Comparisons between observed (HF radar; red) and WebMARVL-configured ROMS (black) surface velocity (a-b) and time-series of observed (ADCP; red) and WebMARVL-configured ROMS (black) at 10 m depth at the ADCP location, off the coast of Western Australia. The location of ADCP is denoted by the black dots in panel (a) and (b).

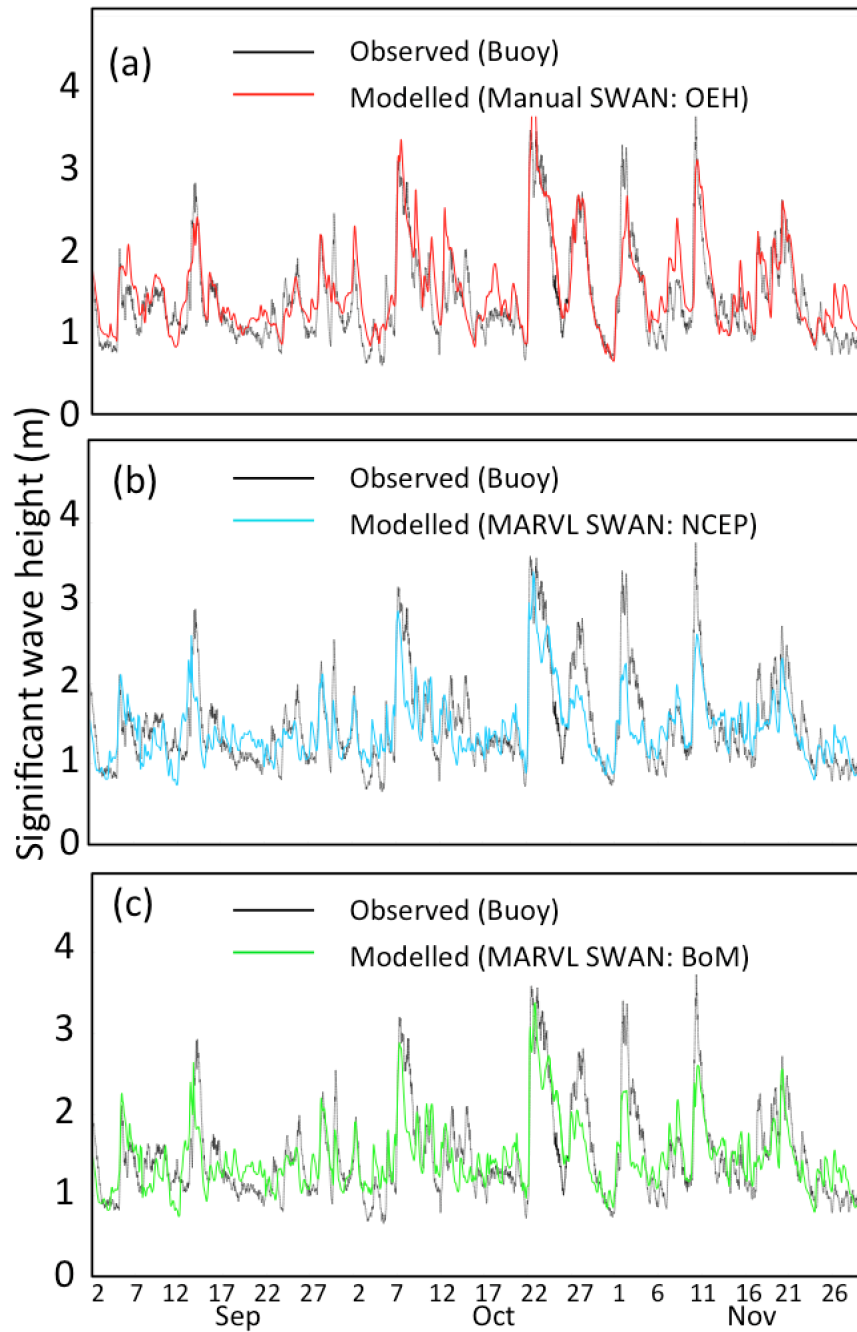


**Figure 8.** Snapshots of (a-b) sea surface temperature and (c-d) surface salinity for the (a,c) initial conditions, interpolated from a  $0.1^\circ$ -resolution model (OceanMAPS) and (b,d) the model fields after 16 days of evolution by a 2-km resolution model.



**Figure 9.** Time-series of modelled (colour) and observed (black) significant wave height off Sydney (~1 km offshore, in 21 m of water), for two different model configurations - (a) MARVL SWAN: BoM and (b) Manual SWAN: OEH.





**Figure 10.** Time-series of modelled (colour) and observed (black) significant wave height off Coffs Harbour (~12 km offshore, in 72 m of water), for three different model configurations - Manual SWAN: OEH (red, a), MARVL SWAN: NCEP (blue, b) and MARVL SWAN: BoM (green, c).