H2020 - Research and Innovation Action

APPLICATE

Advanced Prediction in Polar regions and beyond: Modelling, observing system design and LInkages associated with a Changing Arctic climaTE

Grant Agreement No: 727862

Deliverable No. 1.2

Provision of process-focused and user-relevant metrics through ESMValTool

Submission of Deliverable

Work Package	**WP1 Weather and climate model evaluation**
Deliverable No	D1.2
Deliverable title	Provision of process-focused and user-relevant metrics through ESMValTool
Version	V1
Status	Final
Dissemination level	**PU - Public**
Lead Beneficiary	**1 - AWI**
Contributors	☒ 1 – AWI ☒ 2 – BSC ☐ 3 – ECMWF ☐ 4 – UiB ☐ 5 – UNI Research ☐ 6 – MET Norway ☐ 7 – Met Office ☒ 8 – UCL ☒ 9 – UREAD ☐ 10 – SU ☐ 11 – CNRS-GAME ☐ 12 – CERFACS ☐ 13 – AP ☐ 14 – UiT ☐ 15 – IORAS ☐ 16 – MGO
Due Date	**01 November 2017**
Delivery Date	30 November 2017

This project has received funding from the European Union's Horizon 2020 Research & Innovation programme under grant agreement No. 727862.

Table of Contents

EXECUTIVE SUMMARY

The aim of APPLICATE is *to develop enhanced predictive capacity for weather and climate in the Arctic and beyond, and to determine the influence of Arctic climate change on Northern Hemisphere mid-latitudes, for the benefit of policy makers, businesses and society.* In order to place confidence in weather and climate predictions it is essential that weather and climate models are able to represent key processes in the atmosphere, oceans, land and cryosphere. In addition, a key goal of APPLICATE is user engagement and providing relevant information for policy-makers, businesses and society. To address these goals, Deliverable 1.2 aims to develop process-based metrics and diagnostics for the Arctic (Task 1.2.1) and co-develop metrics and diagnostics that are relevant for the stakeholders and users of APPLICATE outputs (Task 1.2.2).

To ensure that the evaluation of weather and climate models within APPLICATE is aligned with the wider scientific community, the process-based and user-relevant metrics and diagnostics have been implemented in the open-source ESMValTool software. Metrics and diagnostics were chosen to address gaps in the current capability of ESMValTool and to address APPLICATE User Group requirements. These include ocean and sea ice process-based metrics and diagnostics and user relevant diagnostics and metrics for the shipping, energy and fishing sectors.

Through Deliverable 1.2, four APPLICATE partners have further developed their capability to use and develop the ESMValTool software. Developing this capability is essential for the networking effect and wider uptake of ESMValTool at a community level. Achieving Deliverable 1.2 is an important step for APPLICATE as it sits on the critical path for the next tasks in WP1. In particular Deliverable 1.2 will enable the evaluation and assessment of the CMIP5 and CMIP6 climate models in 2018 and 2019 in Task 1.3. The metrics and diagnostics developed in WP1 will also be used to assess the improvement of weather and climate models in WP2.

As well as the activity in Deliverable 1.2, additional diagnostics and metrics will be developed later in the project. In particular, metrics and diagnostics used to describe and evaluate the linkages between the Arctic and mid-latitude atmospheric circulation (especially weather and climate extremes) will be developed in APPLICATE WP1 towards the end of 2018. Additionally, the ongoing process of user engagement within APPLICATE WP7 will undoubtedly present new opportunities to co-develop metrics and diagnostics to inform other economic sectors (e.g. strong winds for local communities and the shipping and infrastructure sectors).

1. INTRODUCTION

1.1. Background and objectives

The aim of APPLICATE is *to develop enhanced predictive capacity for weather and climate in the Arctic and beyond, and to determine the influence of Arctic climate change on Northern Hemisphere mid-latitudes, for the benefit of policy makers, businesses and society.* Within APPLICATE, WP1 will develop advanced metrics and diagnostics that will be used to observationally constrain weather and climate models. In order to observationally constrain weather and climate model, a key task is to develop a series of metrics and diagnostics that are relevant for APPLICATE. More specifically, Deliverable 1.2 focuses on developing process-based metrics and diagnostics for the Arctic (Task 1.2.1) and co-developing metrics and diagnostics that are relevant for the stakeholders and users of APPLICATE outputs (Task 1.2.2).

A key part of WP1 is that the metric and diagnostic development work is being provided within the framework of the ESMValTool software tool. ESMValTool (Earth System Model eValuation Tool; Eyring et al., 2016; https://www.esmvaltool.org/) is an open-source, community tool for evaluating weather and climate models using adequate metrics and diagnostics (see section 2.1 for further details). Developing metrics and diagnostics through a community software tool will align work in APPLICATE WP1 with activity within the wider scientific community and enable the legacy of APPLICATE.

One of the first activities in WP1 has been to carefully consider the definition and terms of reference for model evaluation within APPLICATE. This has been described in the APPLICATE Model Assessment Plan, which was developed as part of D1.1 (see Deliverable Report D1.1 for more details). A specific part of the Model Assessment Plan was to define the meaning of metrics and diagnostics (see section 2.2 for further details).

The overall aim of Deliverable 1.2 is *to provide process-focused and user-relevant metrics and diagnostics through ESMValTool*. This will be achieved by the following objectives:

1. Develop a series of metrics and diagnostics to assess the ability of weather and climate models to represent processes important for the Arctic, particularly processes targeted in APPLICATE's model developments efforts in WP2.

2. Co-develop a series of metrics and diagnostics with users interested in the impacts and opportunities of weather prediction, climate variability and climate change in the Arctic and their effects on Northern Hemisphere weather.

3. Metrics and diagnostic in developed in Deliverable 1.2 will be provided through the open source ESMValTool software package.

1.2. Organisation of this report

The report is organised as following:

Section 2 Methods and Tools: This section describes the methods and tools used in Deliverable 1.2. This includes an outline of the ESMValTool software tool, a summary of the definitions of metrics and diagnostics defined in Deliverable 1.1, the criteria used to prioritise development of metrics and diagnostics and a description of the engagement and co-development within Deliverable D1.2.

Section 3 Results: This section described the metrics and diagnostics implemented in ESMValTool in detail. Examples developed by APPLICATE partners are provided for both processed-based and user-relevant metrics and diagnostics.

Section 4 Summary and Outlook: This section provides a summary of the results and outcomes of Deliverable 1.2. It also provides an outlook on further development of metrics and diagnostics within APPLICATE and the wider research community.

2. METHODS AND TOOLS

Section 2.1 describes the ESMValTool software tool in more detail, section 2.2 discusses the definitions of metrics and diagnostics, section 2.3 describes the criteria used to prioritise development and section 2.4 outlines the engagement and co-development of the metrics and diagnostics for Deliverable D1.2.

2.1 Description of ESMValTool

ESMValTool (Earth System Model eValuation Tool) is a community tool for evaluating metrics and diagnostics from Earth System Models (ESMs). ESMValTool allows for routine comparisons of single or multiple models, either against predecessor versions and/or observations. ESMValTool is an open-source, community effort open to both users and developers encouraging open exchange of diagnostic source code and evaluation results. In addition, ESMValTool is becoming an important software tool that is being developed within a number of other H2020 project (e.g. CRESCENDO, PRIMAVERA). Similarly, the metrics developed in APPLICATE Deliverable D1.2 are being provided through ESMValTool. In this regard, the uptake and development of ESMValTool is aligned with model assessment goals of APPLICATE and the project’s commitment to open data. This will ensure that the evaluation of weather and climate models that occurs within APPLICATE will reach the wider scientific community.

Although there are many advantages to using ESMValTool there have been a number of challenges as well. In particular there has been a steep learning curve in terms of APPLICATE partners developing ESMValTool software. In addition, as ESMValTool is an open-source tool it means that its timeframe for development often does not coincide with the timeframe for APPLICATE, which can lead to difficulties in terms contributing code to a rapidly developing software project. Nonetheless, the activity within Task 1.2.1 and Task 1.2.2 has resulted in four APPLICATE partners having built capacity to develop and contribute code to the ESMValTool project.

2.2 Metrics and Diagnostics in the APPLICATE Model Assessment Plan

Developing metrics and diagnostic to evaluate weather and climate models is a key objective of APPLICATE WP1. However, terms such as metrics and diagnostics are often loosely defined within the weather and climate science communities. To address this, the APPLICATE Model Assessment Plan was developed in Deliverable 1.1 to i) propose unambiguous *definitions* for terms such as “metric”, “diagnostic” or “constraint” and to ii) *frame* the development of metrics in APPLICATE by proposing a set of criteria that would make such metrics desirable, attractive and useful.

In the APPLICATE Model Assessment Plan, diagnostics and metrics were defined as:

Diagnostics are quantities derived from geophysical data sets. The definition proposed by Knutti et al. (2010) suggests that diagnostics are exclusively derived from model output; our definition is somewhat larger and also includes observational references and reanalyses.

The sea ice extent retrieved from satellite observations of sea ice concentration, the strength of the snow-albedo feedback in a reanalysis or the average eddy kinetic energy of the atmosphere in a coupled climate model over the North Atlantic are all examples of such diagnostics. As such, a diagnostic is a tool to simplify complex information that lives in a high-dimensional physical, temporal, probabilistic space, into something much more easily to digest like maps, time series or histograms. *User-relevant diagnostics* are a particular type of diagnostic tailored for the ever-growing community of users of climate data such as the insurance sector, governments, the tourism industry and more broadly stakeholders.

Metrics (used interchangeably with performance metrics in this document) are quantitative measures of agreement between a simulated and observed quantity which can be used to assess the performance of individual models (Knutti et al., 2010). Thus, metrics reflect the agreement of a diagnostic from a system with respect to the same diagnostic computed from a reference. More precisely, a metric maps a diagnostic to a single real number, given a reference. Metrics are inherently attached to the notion of “distance” in geometry. Ideally, they should be defined according to a set of axioms too (such as positivity, triangle inequality, symmetry, nullity). Several types of metrics can be distinguished. These include *standard error metrics, predictability metrics, forecast quality metrics* and *process-based metrics*.

In addition to the definitions of metrics and diagnostics the Model Assessment Plan provided an underpinning framework that describes the desirable criteria for metrics and diagnostics. This framework has been named the “CRISTO” framework, an acronym that describes the six criteria, which are Completeness, Rationale, Interpretability, Stability, Transparency and Observability.

1) **Completeness**: An ensemble of metrics should be as *complete* as possible, meaning that the metrics should together cover all relevant aspects for which the system is to be evaluated.
2) **Rationale**: A metric should always be defined with a clear scope in mind, and according to a scientific question clearly stated *a priori*.
3) **Interpretability**: Metrics should always be accompanied with supporting information: a short description, a figure or an animation.
4) **Stability**: Metrics should be stable with respect to internal variability and interannual variability in the system assessed. In addition, it shouldn’t be affected too much by uncertainty in the reference.
5) **Transparency**: Metrics should be fully reproducible so that anybody is free to verify the steps leading to the final result.
6) **Observability**. A good metric should be derived from diagnostics that are easily observable.

Adoption of the CRISTO framework will ensure that the assessment of models, reanalyses and prediction systems is conducted in APPLICATE following strict scientific standards.

2.3 Selecting metrics and diagnostics for development

In addition to the CRISTO framework described in the APPLICATE Model Assessment Plan and outlined above, the following additional criteria were used to help select metrics and diagnostics for development in Task 1.2:

1. Relevance to the wider scientific community and external users: Metrics and diagnostics developed within D1.2 must be scientifically relevant to wider scientific community and/or relevant quantities that are of interest to Arctic stakeholders (and specifically the stakeholders engaged with APPLICATE, e.g. shipping, Arctic communities, fishing, etc.)

2. Relevance to APPLICATE: Metrics and diagnostics must be relevant to the specific aims of APPLICATE, i.e., they must focus on Arctic processes, linkages between the Arctic and lower latitudes and/or be user relevant. In addition, there is a focus in APPLICATE on both weather and climate timescales and this may influence the choice of metrics and diagnostics.

3. Novelty in ESMValTool: Metrics and diagnostic should not have been previously implemented within ESMValTool (although it desirable that metrics and diagnostic should have been used previously in the scientific literature).

4. Feasibility: There are a number of considerations regarding the feasibility of implementing the metrics and diagnostics. These considerations include i) that metrics and diagnostics are well defined, ii) that metrics are observable and can be evaluated and iii) that weather and climate model output exists to calculate metrics and diagnostics.

2.4 Co-development of metrics and diagnostics for development

The selection of metrics and diagnostics in Deliverable 1.2 required discussion and co-development with internal and external users. This included internal discussion (e.g. with other WPs within APPLICATE) and external engagement (e.g. with the ESMValTool development team and the APPLICATE user group). The following describes the process to co-develop these metrics and diagnostics:

Task 1.2.1 Process-based Metrics and Diagnostics: A number of key stakeholders were identified with regards to the development of metrics and diagnostics for Task 1.2.1. External stakeholders included the ESMValTool development team. To facilitate discussion the ESMValTool co-ordinator (Veronika Eyring, DLR) was invited to the 2017 APPLICATE General Assembly and serves on the APPLICATE Advisory Board. Internally, discussions were held with APPLICATE WP2 researchers focussing on weather and climate model development. A number of key priorities were identified in these discussions. In particular, it was felt that APPLICATE could address the relatively small number of ocean metrics and diagnostics in the current version of ESMValTool. A second priority was felt to be increasing the number of sea ice metrics and diagnostics.

Task 1.2.2 User-relevant Metrics and Diagnostics: The process of engagement for Task 1.2.2 involved APPLICATE WP7 and the APPLICATE User Group that has been assembled as part of WP7. In the first online meeting and the follow-up Arctic Circle roundtable meeting with representatives of the APPLICATE User Group, WP7 discussed priority areas where enhanced forecasts could inform Arctic stakeholders' decision-making. A priority topic identified in these preliminary discussions was:

1. New routes for shipping through the Arctic. Container shipping across the Arctic is becoming a reality, but there is still an important question on the navigability of the Arctic Ocean. Shipping companies need enhanced climate forecast to make decisions about when is the perfect period for shipping, which route to select, how much time it will take and how much icebreaker assistance the journey through the Arctic will need.

Additional topics discussed during the Arctic Circle Assembly that APPLICATE WP7 has recently attended were:

2. Sustainable management of fishing and food security in the Arctic. Fishing is strongly affected by the changes in ocean temperature and salinity. As ocean temperatures increase in the Arctic region, new activities – such as new shipping routes, deep-sea mining, oil and gas exploitation, and tourism – will present multiple stresses for fish stocks. Enhanced climate prediction could support sustainable fishery practice.

3. Renewable energy. Moving towards low-carbon energy sources, such as wind turbines, is a key consideration for many Arctic countries and communities. Essential to understanding the viability of renewable energy for the energy sector is an assessment of the impact of climate variability and climate change on low-level winds and quantities such as renewable wind capacity.

These three topics are translated in the following metrics: sea ice free regions, ocean temperatures and renewable wind capacity. Additional topics have been discussed within the APPLICATE User Group (e.g. strong winds) and these will be followed up with discussions in the forthcoming User Group meeting that will take place during the APPLICATE General Assembly in January 2018 (and in the subsequent user engagement activities). WP7 will continue discussions with Arctic stakeholders regarding priority topics where enhanced climate prediction can inform their activities and support sustainable development of the Arctic.

3. RESULTS AND DISCUSSION

Section 3 describes the metrics and diagnostics developed as part of D1.2. Section 3.1 describes the process-based metrics and diagnostics and Section 3.2 describes user-relevant metrics and diagnostics.

3.1 Process-based metrics and diagnostics implemented in ESMValTool

As described in section 2, discussions with the ESMValTool co-ordinator, ESMValTool development team and researchers from APPLICATE WP2 had identified ocean and sea ice metrics and diagnostics as key priorities for development. Table 1 lists the process-based metrics and diagnostic that have been developed for Task 1.2.1 and Deliverable 1.2. All metrics and diagnostics have been developed in the ESMValTool framework.

Table 1. List of process-based metrics and diagnostics in Task 1.2.1

Name	Metric or Diagnostic	Partner
Hovmoller diagram of temperature and salinity for predefined ocean region	Diagnostic	AWI
Hovmoller diagram of temperature and salinity anomalies for predefined ocean region	Diagnostic	AWI
Mean vertical temperature and salinity profile for predefined regions	Diagnostic	AWI
Spatial distribution of temperature and salinity at original model levels	Diagnostic	AWI
Spatial distribution of ocean temperature at the mean Atlantic water core level	Diagnostic	AWI
Atlantic water core depth	Metric	AWI
Atlantic water temperature	Metric	AWI

Spatial difference between climatology and simulated temperature and salinity at different levels.	Diagnostic	AWI
T/S diagram for predefined ocean region	Diagnostic	AWI
Vertical temperature and salinity sections along predefined paths	Diagnostic	AWI
Spatial distribution of the mean ocean currents speed at different levels	Diagnostic	AWI
Vertical transect of temperature and salinity for arbitrary sequence of points	Diagnostic	AWI
Sea ice feedback diagnostics	Diagnostic	UCL

The following subsections described in more detail the metrics and diagnostics developed in ESMValTool for the ocean (3.1.1) and sea ice (3.1.2).

3.1.1 Ocean diagnostics

As mentioned in Section 2, ESMValTool already has a good selection of metrics and diagnostics related to the atmospheric part of the coupled model systems. There is a relative lack of ocean and sea ice related metrics and diagnostics, and consequently we decide to put most of our efforts in to filling in this gap. Our main focus was the Arctic Ocean, however most of the diagnostics are implemented in a way that can be easily expanded to other parts of the World Ocean. At present all diagnostics and metrics are applied to CMIP5 20th century coupled historical simulations (period 1970-2005) so that comparison to available climatologies and observations is possible. In the ESMValTool framework similar analysis can be easily performed for scenario simulations as well as for the upcoming CMIP6 simulations. Only a subset of CMIP5 models were used in particular because our analysis is currently limited to z coordinate models. When the vertical interpolation of sigma coordinate models to z levels is implemented in ESMValTool it will be straightforward to expand the analysis to models that use different vertical coordinates.

The development of ocean diagnostics are largely based on the experience gained during CLIVAR Coordinated Ocean-ice Reference Experiments - Phase II (CORE II). The ocean metrics and diagnostics partly reproduce the well tested and widely accepted Arctic Ocean related diagnostics described in Ilıcak et al., (2016). The ocean related metrics and diagnostics can be divided in to four main types described below:

a. Vertical distribution of temperature and salinity

The vertical structure of temperature and salinity (T and S) in the ocean model is a key diagnostic that is used for ocean model evaluation. Realistic T and S distributions mean that models properly represent dynamic and thermodynamic processes in the ocean. Different ocean basins have different hydrological regimes so it is important to perform analyses of vertical TS distribution for different basins separately. The basic diagnostic in this sense is mean vertical profiles of temperature and salinity over some basin averaged for a relatively long period of time. Fig. 1 shows the mean (1970-2005) vertical ocean potential temperature distribution in the Eurasian Basin of the Arctic Ocean. In addition to individual vertical profiles for every model, we also show the mean over all participating models and a similar profile from climatological data (PHC3).

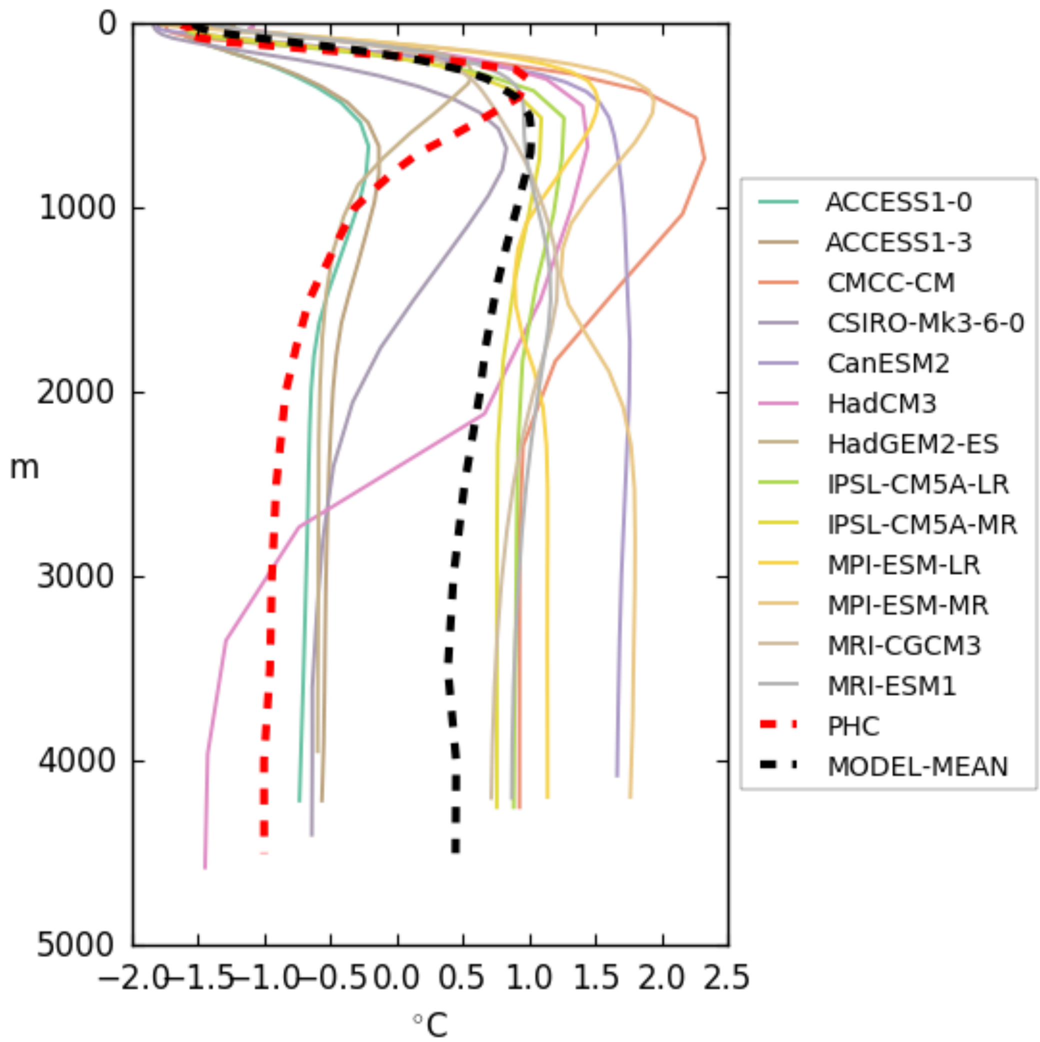

Fig. 1 Mean (1970-2005) vertical potential temperature distribution in the Eurasian basin for participating CMIP5 coupled ocean models, PHC3 climatology (dotted red line) and multi-model mean (dotted black line).

The characteristics of vertical TS distribution can change with time, and consequently the vertical TS distribution is an important indicator of the behaviour of the coupled ocean-sea ice-atmosphere system in the North Atlantic and Arctic Oceans. One way to evaluate these changes is by using Hovmoller diagrams. We have created Hovmoller diagrams for two main Arctic Ocean basins – Eurasian and Amerasian with T and S spatially averaged on a monthly basis for every vertical level. Examples are shown in Fig. 2. This diagnostic allows the temporal evolution of vertical ocean potential temperature distribution to be assessed. In addition, we also provide Hovmoller diagrams of T and S anomalies relative to the first time step of the analyzed time range.

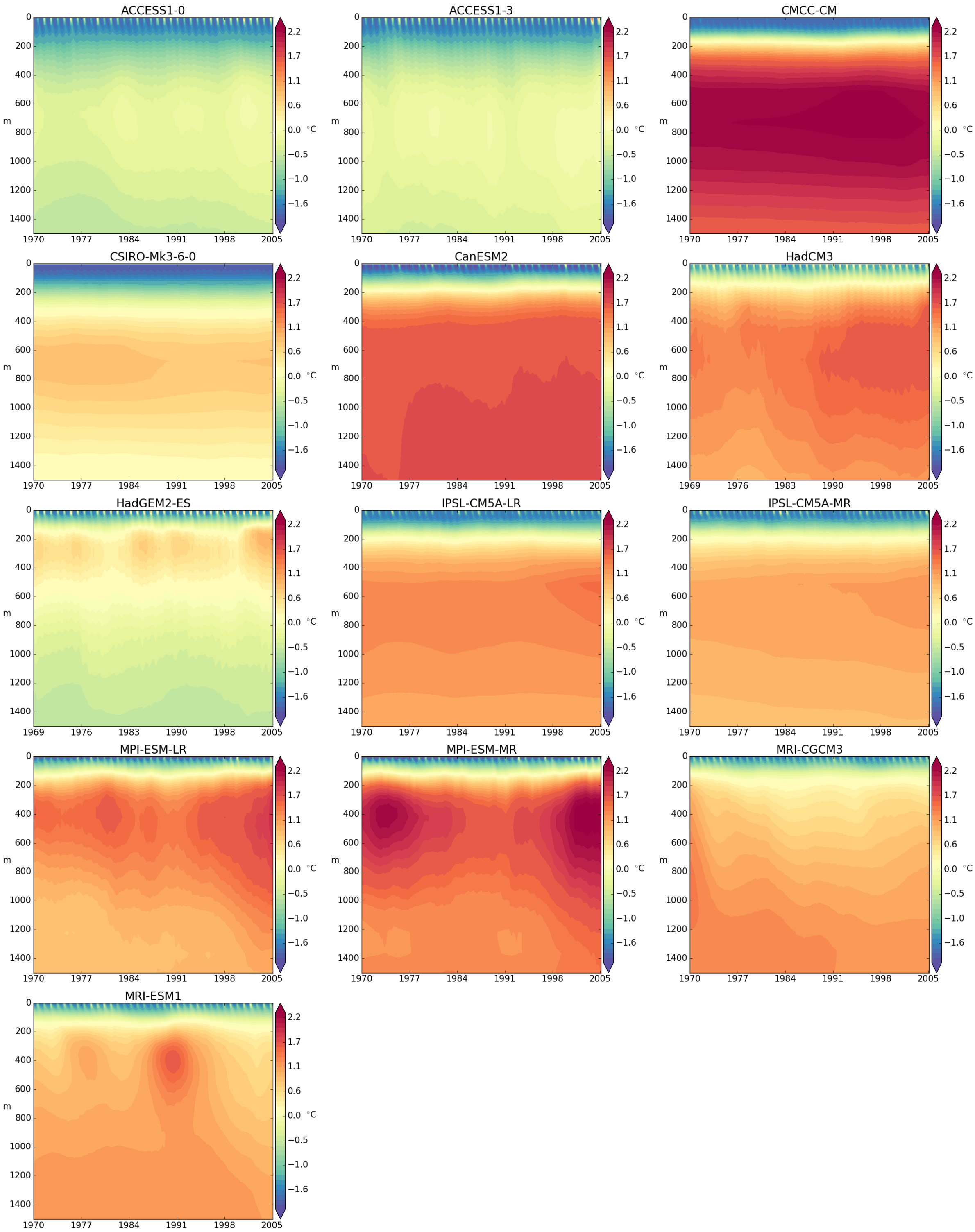

Fig. 2 Hovmoller diagram of monthly spatially averaged potential temperature in the Eurasian Basin of the Arctic Ocean for selected CMIP5 climate models (1970-2005).

T-S diagrams combine temperature and salinity, which allows for the analysis of water masses and their potential for mixing. The lines of constant density for specific ranges of temperature and salinity are shown on the background of the T-S diagram. The dots on the diagram are individual grid points from specified region at all model levels within user specified depth range. The depths are colour coded. Examples of the mean (1970-2005) T-S diagram for Eurasian Basin of the Arctic Ocean shown in Fig. 3.

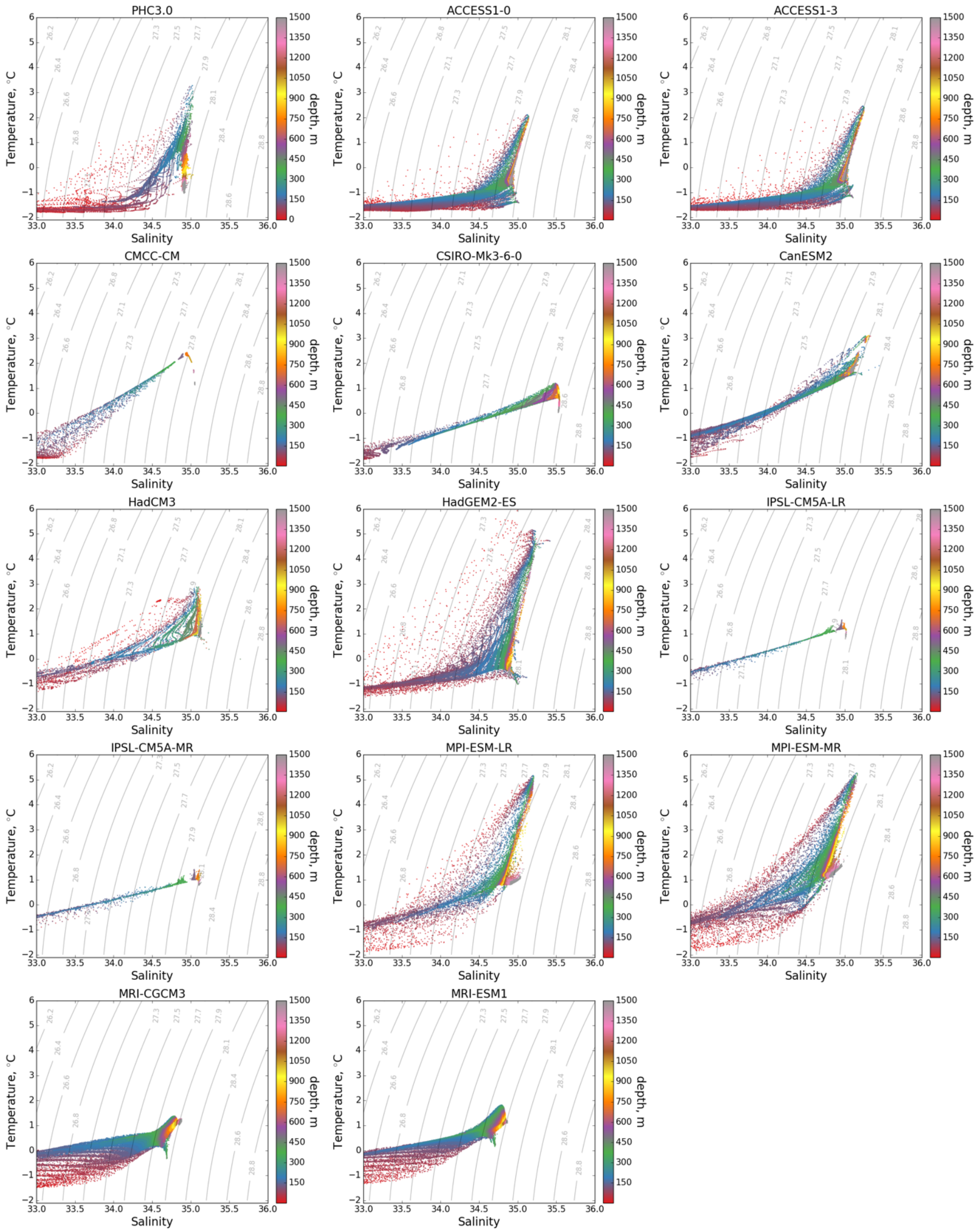

Fig. 3 Mean (1970-2005) T-S diagrams for Eurasian Basin of the Arctic Ocean for selected CMIP5 models and PHC3.0 observations.

b. Spatial distribution of temperature and salinity

The spatial distribution of basic oceanographic variables characterises the properties and spreading of ocean water masses. For coupled models, capturing the spatial distribution of oceanographic variables is especially important in order to correctly represent the ocean-ice-atmosphere interface.

We have implemented plots with spatial maps of temperature, salinity and current speeds at original model levels. For temperature and salinity, we have also implemented spatial maps of model biases from the observed climatology. For the model biases, values from the original model levels are linearly interpolated to the climatology and then spatially interpolated from the model grid to the regular PHC (climatology) grid. Resulting fields (Fig. 4) show model performance in simulating spatial distribution of temperature and salinity.

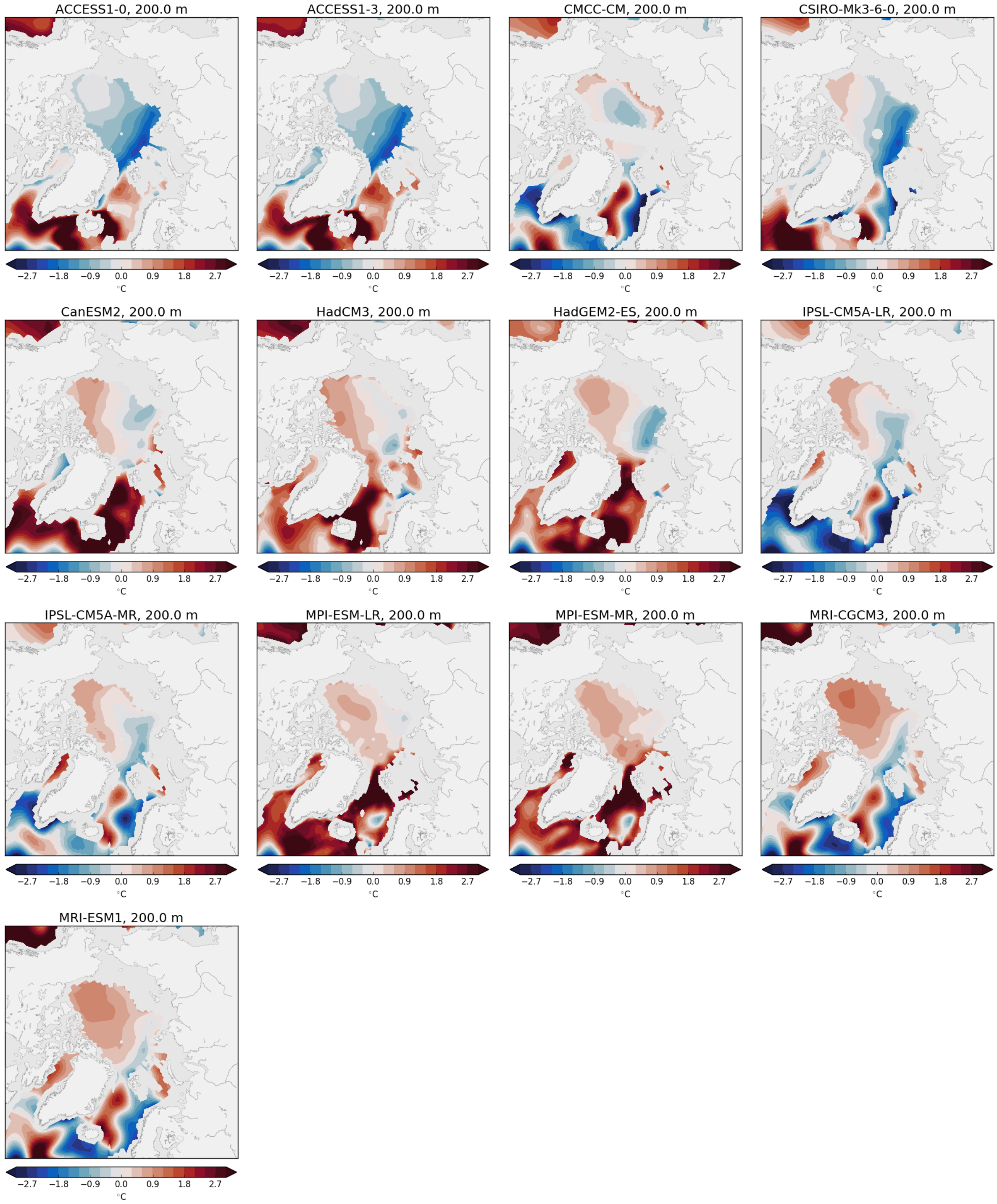

Fig. 4 Mean (1970-2005) potential temperature bias at 200m relative to PHC climatology.

c. Vertical transects

Vertical transects through arbitrary sections are important for the analysis of the vertical distribution of ocean water properties and especially useful when exchange between different ocean basins is evaluated. We have implemented diagnostics that allow for the definition of an arbitrary ocean section by providing set of points on the ocean surface. For each point, a vertical profile on the original model levels is interpolated. All profiles are then connected to form a transect. The great-circle distance between the points is calculated and used as along-track distance.

One of the main use cases is to create vertical sections across ocean passages. Fig. 5 shows examples of such a section through the Fram Strait, one of the main regions of water exchange between Arctic and the North Atlantic.

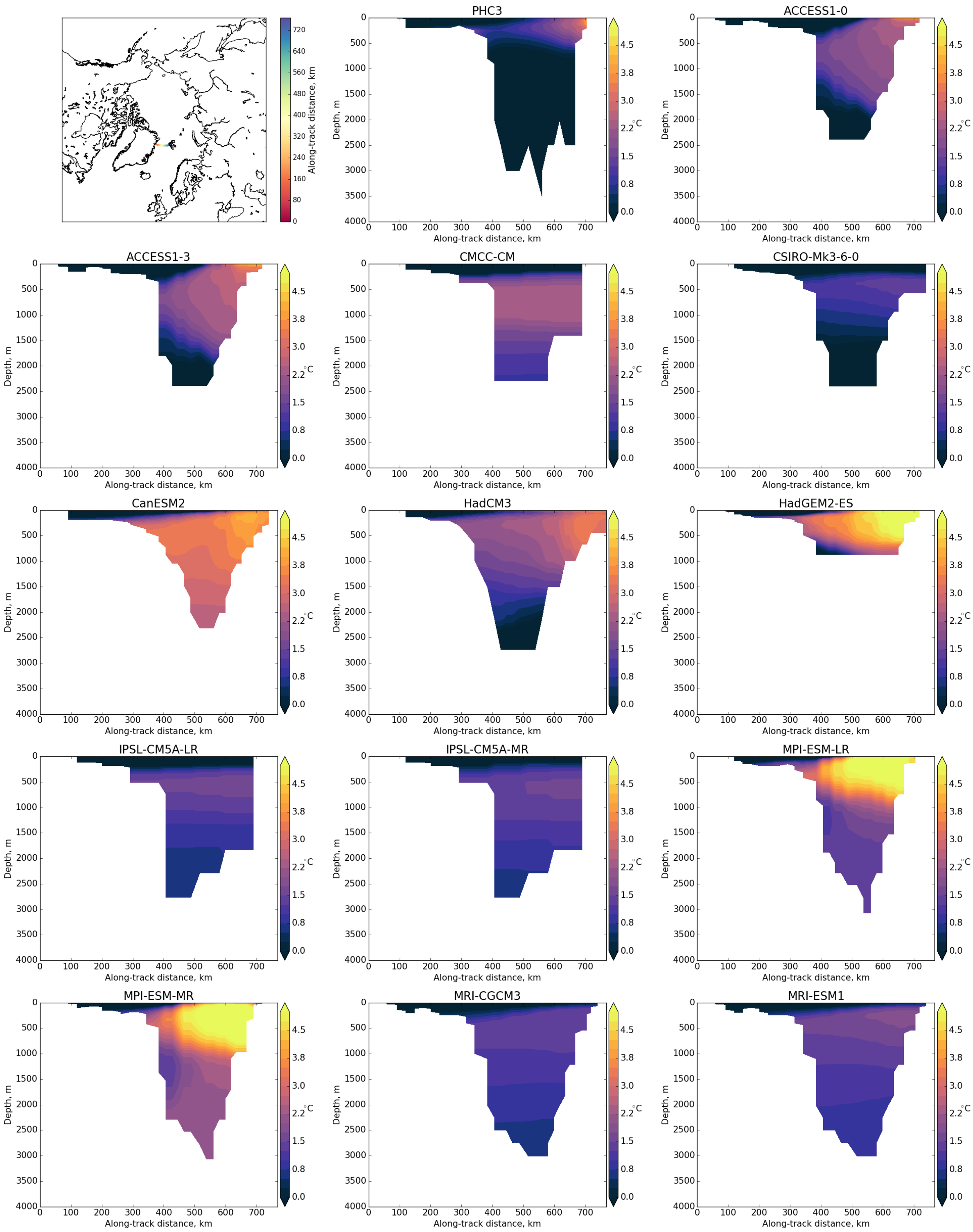

Fig. 5 Mean (1970-2005) potential temperature across the Fram strait. PHC3 is climatology.

We also have implemented a transect that follows the pathway of the Atlantic Water following Ilıcak et al., (2016).

d. Atlantic water characteristics in the Arctic Ocean

This set of diagnostics and metrics is specific for the Arctic Ocean. Atlantic water is a key water mass of the Arctic Ocean and its proper representation is one of the main challenges in Arctic Ocean modelling. We have created two metrics by which models can be easily compared in terms of Atlantic water simulation. The temperature of the Atlantic Water core is calculated for every model as the maximum potential temperature between 200 and 1000 meters depth in the Eurasian Basin. The depth of the Atlantic Water core is calculated as the model level depth where the maximum temperature is found in Eurasian Basin (Atlantic water core temperature). Fig. 6 shows examples of Atlantic Water core temperature metric.

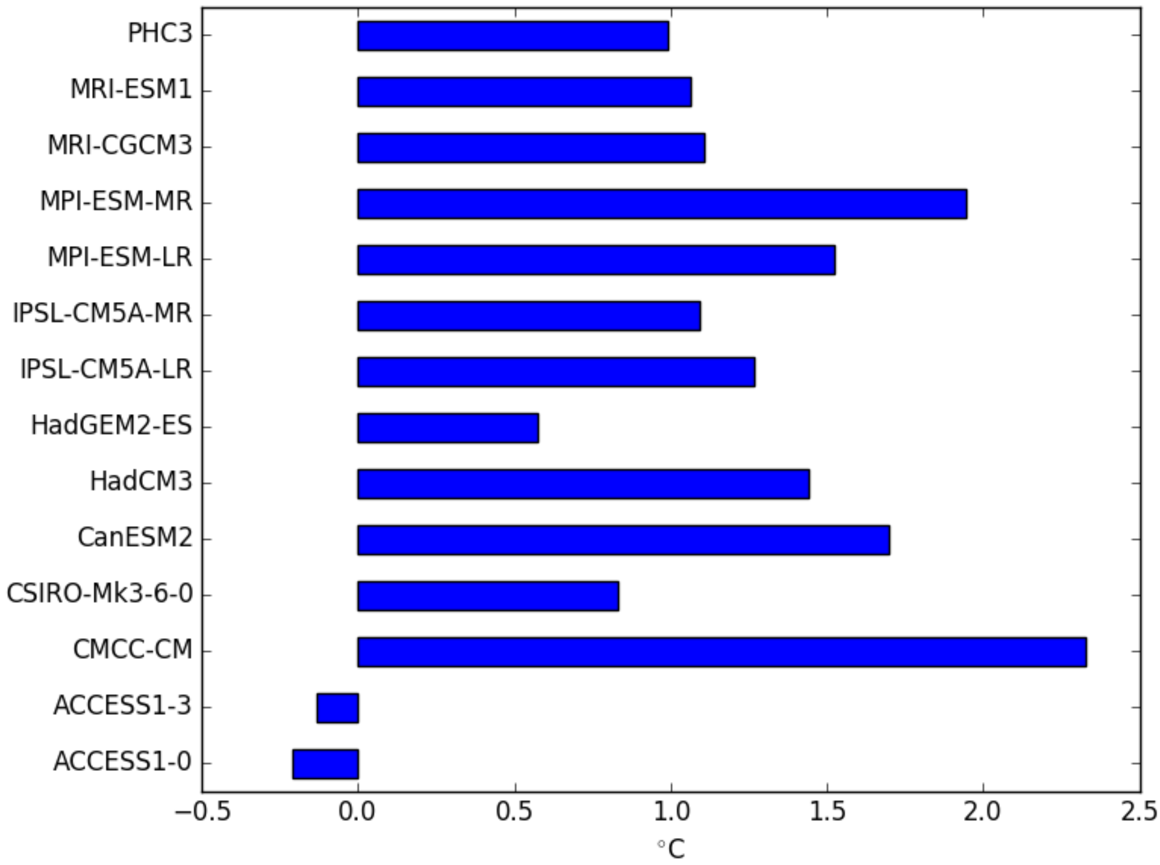

Fig. 6 Mean (1970-2005) Atlantic Water core temperature. PHC3 is an observed climatology.

In order to evaluate the spatial distribution of Atlantic water in different climate models we also provide diagnostics with maps of the spatial temperature distribution at the model's Atlantic Water depth.

e. Ocean metrics and diagnostics web interface

To improve the user experience when using the APPLICATE diagnostics package we have combined all ocean diagnostics and metrics in a web based interface. For each diagnostic and metric a thumbnail and short description is provided. For the mean vertical profiles over predefined ocean basins, both static and interactive versions of plots are provided. These plots are relatively busy and the interactive web interface allows a subset of models to be selected or allows the user to concentrate on a specific depth range. Examples of the web interface are shown in Fig. 7.

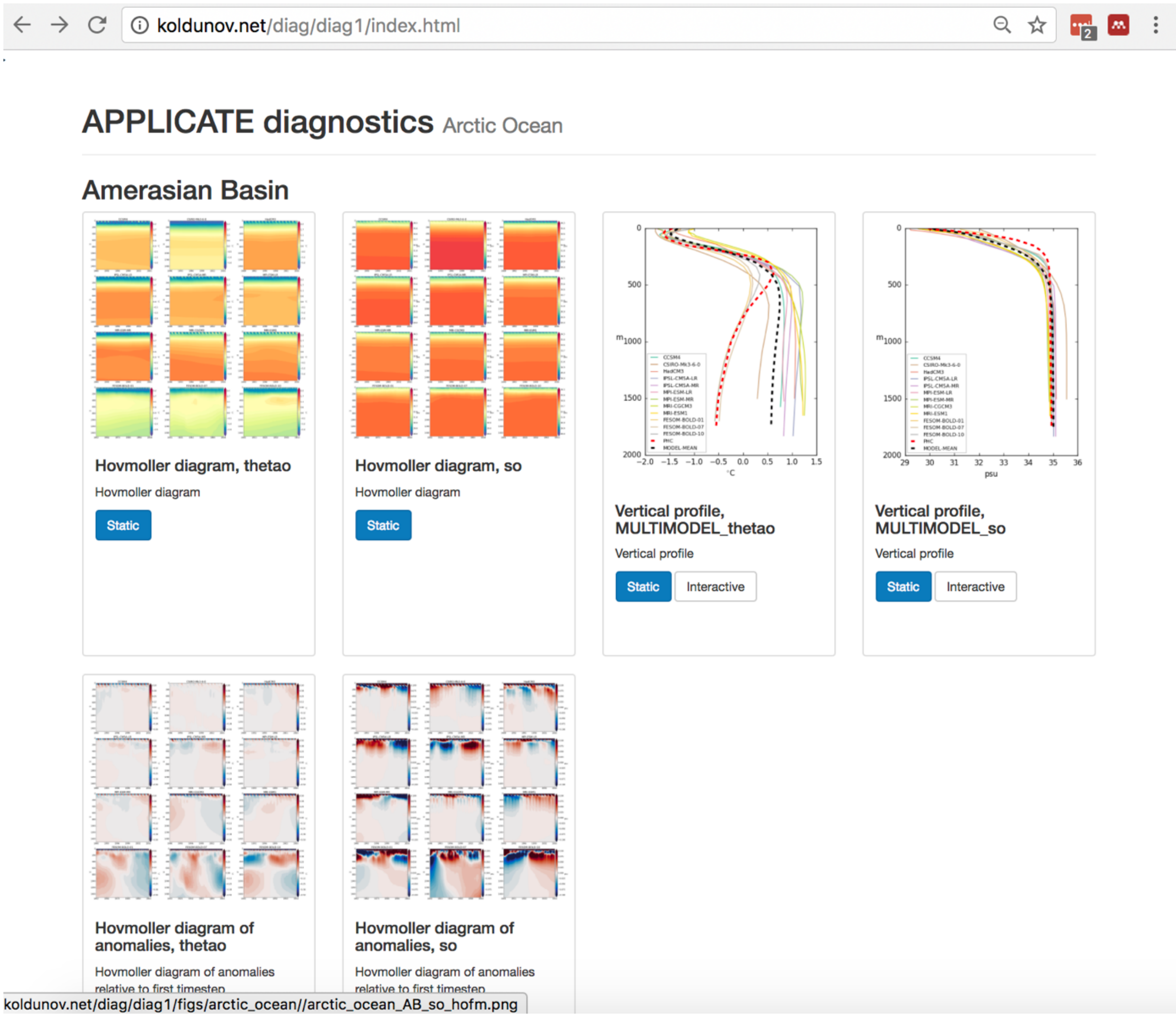

Fig. 7 Examples of APPLICATE ocean metrics and diagnostics web interface.

3.1.2 Sea ice feedback diagnostics

Sea ice has long been recognized as a key player of polar climate. Its central role in Arctic Amplification and its possible influences on lower latitude weather regimes explain why its recent and future evolutions have attracted much attention in the scientific community. While all climate models generally project that Arctic sea ice decline will continue over the 21st century, the magnitude of these changes remain uncertain. Some models predict ice-free conditions as early as 2030 under a “business as usual emission” scenario, others don’t even foresee this possibility before 2100. It is therefore of high priority to (1) understand the physical reasons leading to the spread in current sea ice projections and (2) develop appropriate diagnostics and metrics to evaluate the simulation of sea ice in current climate models.

We have developed a process-oriented approach to quantify the way that CMIP5 models simulate the two basic features of sea ice seasonality: the processes of thermodynamic melt and growth. We here focus on the process of sea ice growth (the diagnostic relevant to the process of sea ice melt is under development). Our diagnostic, the “ice formation efficiency” (IFE) estimates the ability of a model to recover a summer sea ice volume anomaly during the next growing season. This diagnostic is related to the well-known negative ice growth-ice thickness feedback, which states that thin ice grows faster than thick ice. We found that the process of sea ice growth is more efficient (i.e., IFE is more negative) when the ice is thin (Fig. 8).

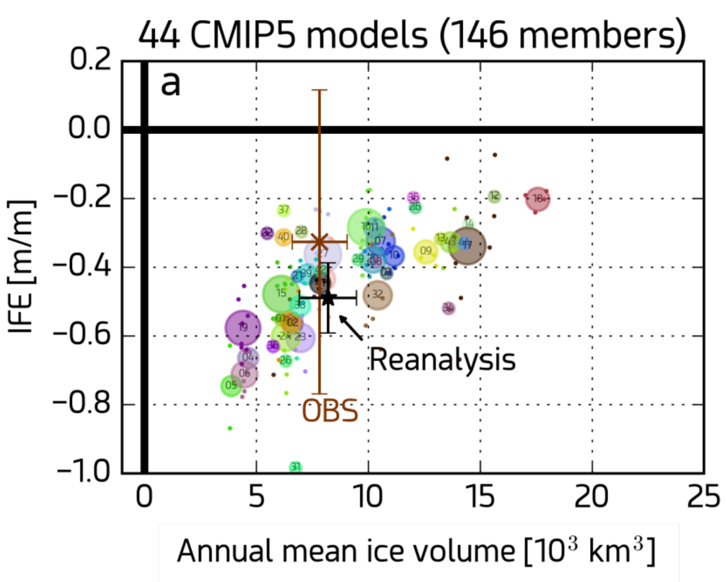

Fig. 8. Ice Formation Efficiency (IFE, defined as the regression between summer-to-winter volume production and summer sea ice volume, all computed north of 80°N) versus the annual mean sea ice volume north of 80°N. All quantities are evaluated on 1955-2004 for 44 CMIP5 models. Individual members are shown as small dots and ensemble means as circles (larger circles if more members). Estimate from the PIOMAS reanalysis (Schweiger et al., 2011) and from observations (Kwok et al., 2009; EUMETSAT, 2015) are also shown (2003-2008).

In turn, we found that the strength of the process of sea ice growth (IFE) controls the year-to-year variability and change of sea ice in CMIP5 models. For example, we found that the stronger the IFE, the weaker the sea ice volume persistence (R = –0.59, p-value < 0.0001) and the weaker the 1955-2004 loss in sea ice volume (R = –0.32, p-value = 0.02). These relationships form the basis of 'emergent constraints', as they indicate a direct connection between the seasonal processes of growth and melt, and the inter-annual behavior of sea ice in the models. Given the strong dependence of IFE on the mean state, our analysis suggests that simulating the annual mean sea ice thickness realistically is a necessary condition to trust climate predictions and projections from a model.

Our developments are the topic of an article to be submitted. The function computing the IFE has been successfully introduced in the ESMValTool (private Github) and will be transferred to the public Github as soon as the paper is accepted for publication.

3.2 User-relevant metrics and diagnostics implemented in ESMValTool

As mentioned in Section 2, user engagement through the APPLICATE User Group has identified a number of priority area for development of metrics and diagnostics for ESMValTool. Currently, there is very little user-relevant information beyond basic meteorological variables (temperatures, rainfall etc…) so the work in Task 1.2.2. will extend ESMValTool capabilities. The discussion with users in Task 1.2.2 will be continued through ongoing engagement with the APPLICATE User Group (e.g. the next User Group meeting at the APPLICATE General Assembly in January 2018). Table 2 outlines the metrics and diagnostics developed for the Task 1.2.2.

Table 2. List of user-relevant metrics and diagnostics in Task 1.2.2

Name	Relevant Sectors	Metric or Diagnostic	Partner
Sea ice free regions	Shipping	Metric	BSC
Ocean temperatures	Fishing, Blue Growth	Diagnostic	AWI
Renewable Wind Capacity	Energy	Diagnostic	UREAD

The following subsections described in more detail the user relevant metrics and diagnostics developed in ESMValTool.

3.2.1 Sea Ice Free Regions for the Shipping Sector

The Earth Sciences Department at BSC has focused on implementing a diagnostic describing sea ice-free regions in ESMValTool. Understanding the potential location and seasonality of sea ice free regions is of paramount importance for shipping companies crossing the Arctic (and directly mentioned by Uwe Pahl, the previous captain of the German research icebreaker RV Polarstern, on the Polar Prediction Matters blog). The metric generally used to represent ice-free regions is the sea ice edge, which is often defined by the 15% contour of sea ice concentration. This metric has been already included in the ESMValtool, as illustrated in Figure 9. A complementary metric proposed by Goessling et al. (2016) for model evaluation is the integrated ice-edge error, which is defined as the area where the forecast and the “truth” disagree on the ice concentration being above or below 15%. This user-relevant metric will be the next one to be implemented.

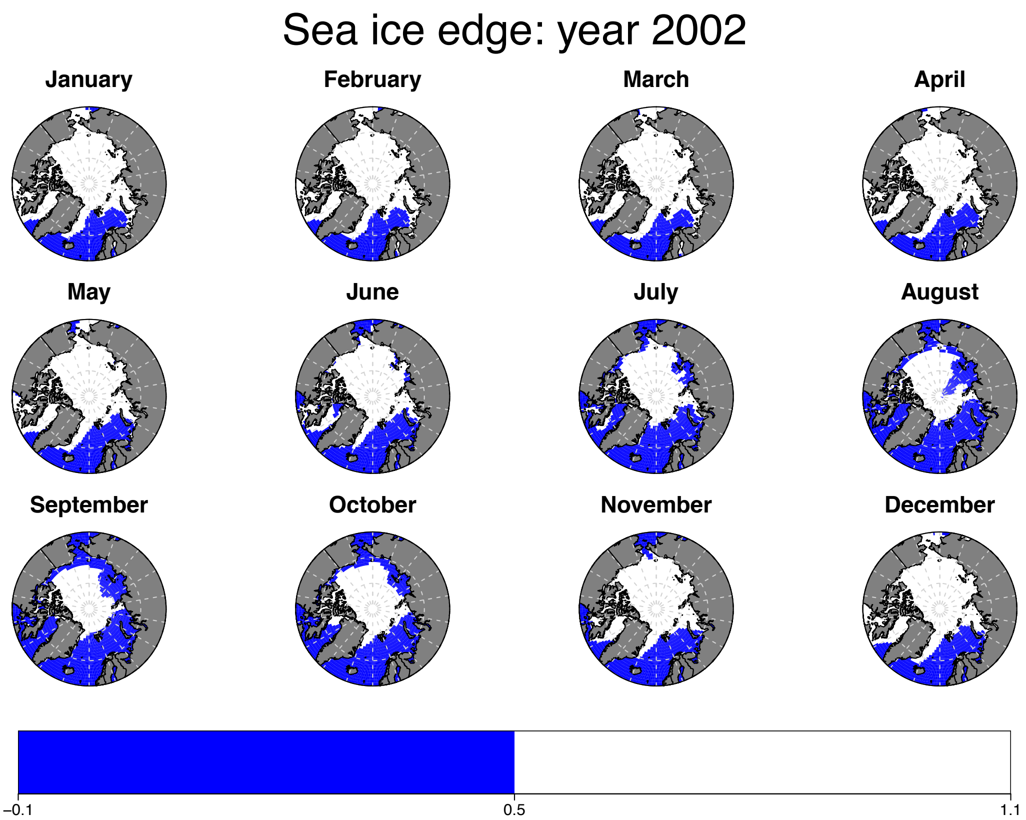

Fig. 9 Example of the representation of sea ice edge in the ESMValTool for all the months in a given year (2002) in a climate simulation with the EC-Earth climate model.

3.2.2. Ocean Temperatures for the Fishing Sector

In order to provide aggregated information on environmental characteristics relevant for fisheries and aquaculture, AWI have developed and implemented in ESMValTool two basic diagnostics. Diagnostics can be easily adjusted to fit requirements for different marine species.

The first diagnostics are monthly mean plots of the mean, minimum and maximum value of certain ocean variables (e.g. temperature, salinity, etc.) in certain depth ranges for specific regions (e.g. sea or exclusive economic zone). The plots also contain information about limits of tolerance for the species and its optimal range. Fig. 10 demonstrate an example of this diagnostic for the temperatures in the first 100 meters of the North Sea with tolerances and optimal ranges for Atlantic cod *Gadus morhua* (Righton et al., 2010) shown as red and blue shadings respectively.

Another diagnostic provides information on thermal/salinity habitat availability. It is based on the mean temperature/salinity values for certain periods and provides histograms and probability density functions of waters with different temperatures/salinities for certain depth ranges over specific regions. This diagnostic shows the relative amount of waters with different properties available for species to occupy. Comparisons to climatology (PHC) also allow the quality of model simulations for key fisheries and aquaculture regions to be evaluated. Fig. 11 demonstrate example of this diagnostic for mean (1970-2005) temperature in the first 100 meters of the North Sea.

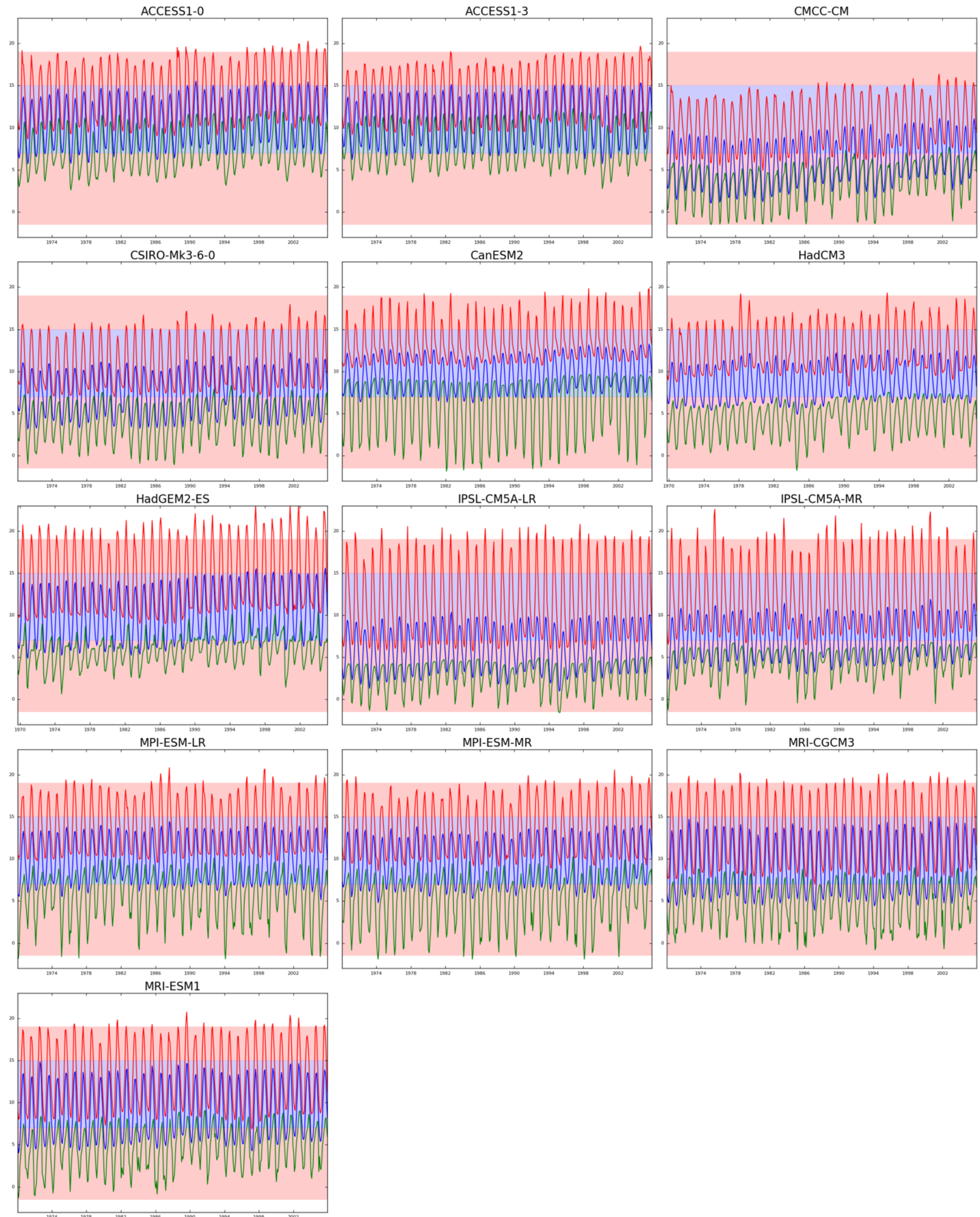

Fig. 10. Monthly mean minimum (green), mean (blue) and maximum (red) temperature in the first 100 meters of the North Sea with tolerance (red shading) and optimal (blue shading) ranges for Atlantic cod Gadus morhua (Righton et al., 2010).

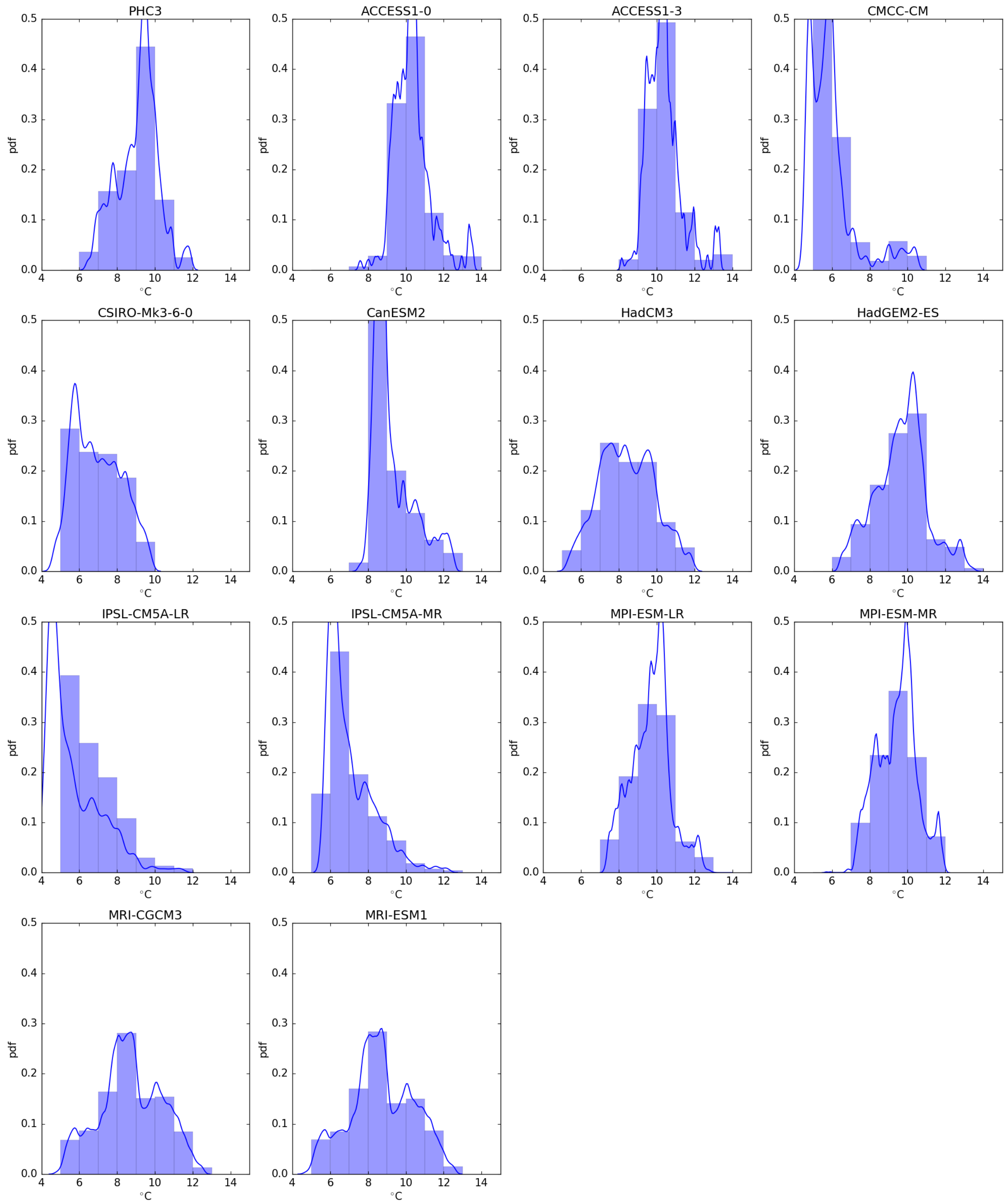

Fig.11 Histogram and probability density function for mean (1970-2005) temperature in the first 100 meters of the North Sea. PHC3 is climatology.

3.2.3. Wind Power Capacity for the Energy Sector

Key requirements for the renewable energy sector are i) information to help inform the placing of wind turbines, and ii) understanding how low-level winds are impacted by climate variability and climate change. Although strong winds are desirable for wind power production, excessively strong winds can lead to wind turbines needing to be shut down for safety reasons. There is therefore an optimal range of winds that maximises wind power production. The dependence of wind power generation upon wind speed is characterised by a "wind power capacity curve" (i.e. the fraction of power that can be generated from a wind turbine's given rating as a function of wind speed).

The wind power capacity diagnostic developed here is based upon the wind power capacity curve and approach described in detail in Cannon et al. (2015). Surface (10m) wind speeds are first adjusted to a 70m hub height using a logarithmic wind profile for neutral stability. A safety shut-down wind speed threshold of 25 ms^{-1} is assumed. The wind power capacity curve is then used to translate the wind speed into a wind power capacity.

Figure 12 shows spatial maps of the annual mean wind power capacity for two CMIP5 models (HadGEM2-ES and GFDL-CM3). Offshore wind turbines can produce up to 95% of rated capacity, but this falls off as the turbine are placed further inland. There are also differences between the two climate models due to their representation of low level winds. These differences are very apparent in the Arctic region. The diagnostic can be applied to additional climate models and climate change simulations to understand how wind power capacity might change in the future. In addition, the wind power capacity curve is very simply implemented in ESMValTool and can be easily altered to represent new turbine technologies.

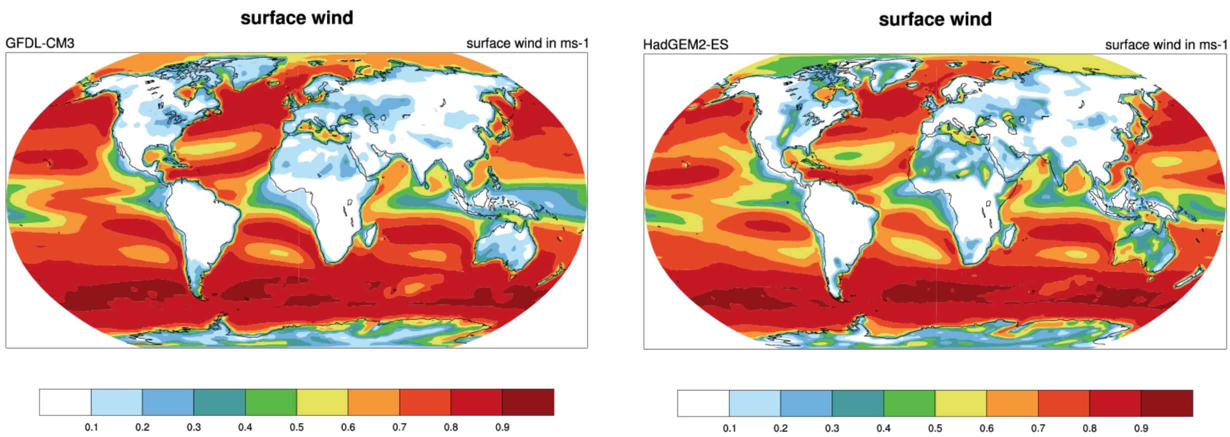

Figure 12. Annual mean (1980-2005) wind power capacity (fraction of power generated from a wind turbine) for GFDL-CM3 and HadGEM2-ES CMIP5 historical simulations. Based on the approach of Cannon et al. (2015).

4. SUMMARY AND OUTLOOK

Section 4.1 provides a summary of the results and outcomes of Deliverable 1.2, while section 4.2 provides an outlook on further development of metrics and diagnostics within APPLICATE and the wider research community.

4.1 Summary

The overall aim of Deliverable 1.2 is to provide process-focused and user-relevant metrics and diagnostics through ESMValTool, an open-source, community tool for evaluating metrics and diagnostics from weather and climate models. The developments in ESMValTool will be used within WP1 of APPLICATE to observationally constrain weather and climate models. The following is a brief summary of the outcomes of Deliverable 1.2:

1. Process-based and user-relevant metrics and diagnostics have been implemented in the open-source ESMValTool software. Metrics and diagnostics were chosen to address gaps in the current capability of ESMValTool and to address APPLICATE User Group requirements. These include ocean and sea ice process-based metrics and diagnostics and user relevant metrics and diagnostics for the shipping, energy and fishing sectors.

2. Through Deliverable 1.2, four APPLICATE partners have further developed their capability to use and develop the ESMValTool software. Developing this capability is essential for the networking effect and wider uptake of ESMValTool at a community level.

3. Deliverable 1.2 is an important deliverable for APPLICATE from a user engagement perspective. The work in Deliverable 1.2 has enabled links to be formed between WP1, WP2 and WP7. In addition, the work has provided a focus within WP1 for external engagement with the WP7 APPLICATE User Group.

4.2 Outlook

The aim of APPLICATE is *to develop enhanced predictive capacity for weather and climate in the Arctic and beyond, and to determine the influence of Arctic climate change on Northern Hemisphere mid-latitudes, for the benefit of policy makers, businesses and society.* An essential aspect of enhancing the predictive capabilities of weather and climate models is evaluating their performance. To ensure that the evaluation of weather and climate models within APPLICATE is aligned with the wider scientific community, the metrics and diagnostics have been provided using the open-source ESMValTool software.

Although Deliverable 1.2 has successfully provided a series of process-based and user-relevant metrics and diagnostics with the ESMValTool software, the following points outline some of the future directions for model evaluation and the development of metrics and diagnostics within APPLICATE:

1. *Weather and Climate Model Evaluation:* Deliverable 1.2 is an important as it sits on the critical path for the next tasks in WP1. In particular Deliverable 1.2 will enable the evaluation and assessment of the CMIP5 and CMIP6 climate models in 2018 and 2019 in Task 1.3. The metrics and diagnostics developed in WP1 will also be used to assess the improvement of weather and climate models in WP2.

2. *Development of metrics and diagnostics:* As well as the activity in Deliverable 1.2, additional diagnostics and metrics will be developed. In particular, metrics and diagnostics used to described and evaluate the linkages between the Arctic and mid-latitude atmospheric circulation (especially weather and climate extremes) will be developed in APPLICATE WP1. Additionally, the ongoing process of user engagement within APPLICATE WP7 will undoubtedly present new opportunities to co-develop metrics and diagnostics to inform other economic sectors (e.g. strong winds for local communities and the shipping and infrastructure sectors).

5. REFERENCES

Cannon, D. J., D. J. Brayshaw, J. Methven, P. J. Coker and D. Lenaghan (2015) Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. Renewable Energy, 75, 767-778. doi:10.1016/j.renene.2014.10.024

Eyring, V. et al., (2016) ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, Geosci. Model Dev. 9, 1747-1802

EUMETSAT Ocean and Sea Ice Satelitte Application Facility. Global sea ice concentration reprocessing dataset 1978-2015 (v1.2, 2015), [Online]. Norwegian and Danish Meteorological Institutes. Available from http://osisaf.met.no

Goessling HF, Tietsche S, Day JJ, Hawkins E, Jung T (2016) Predictability of the Arctic sea ice edge. Geophys Res Lett 43:1642–1650

Ilıcak, M. et al., An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part III: Hydrography and fluxes, Ocean Modelling, Volume 100, April 2016, Pages 141-161, ISSN 1463-5003, doi.org/10.1016/j.ocemod.2016.02.004

Knutti, R., G. Abramowitz, M. Collins, V. Eyring, P.J. Gleckler, B. Hewitson, and L. Mearns, 2010: Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections. In: Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, and P.M. Midgley (eds.)]. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland.

Kwok, R. *et al.*Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008. *J. Geophys. Res.* **114**, C07005 (2009)

Righton, D. A., Andersen, K. H., Neat, F., Thorsteinsson, V., Steingrund, P., Svedäng, H., ... & Wright, P. (2010). Thermal niche of Atlantic cod Gadus morhua: limits, tolerance and optima. *Marine Ecology Progress Series*, *420*, 1-13.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H. Uncertainty in modeled Arctic sea ice volume. *J. Geophys. Res.***116**, C00D06 (2011)

6. ACRONYMS

ESM: Earth System Model

ESMValTool: Earth System Model VALidation TOOL

NWP: Numerical Weather Prediction

CMIP5: Fifth Coupled Model Intercomparison Project

CMIP6: Sixth Coupled Model Intercomparison Project