Hexagon@University of Bergen, Norway

Csaba Anderlik
IT-department
University of Bergen
Bergen, Norway
e-mail:csaba.anderlik@uib.no

Ingo Bethke
Uni Research Climate
Uni Research, Bergen, Norway
e-mail:ingo.bethke@uni.no

Mats Bentsen
Uni Research Climate
Uni Research, Bergen, Norway
e-mail:mats.bentsen @uni.no

Alok Kumar Gupta
Uni Research Climate
Uni Research, Bergen, Norway
e-mail:alok.gupta@uni.no

Jon Albretsen
Institute for Marine Research
Bergen, Norway
e-mail:jon.albretsen@imr.no

Lars Asplin
Institute for Marine Research
Bergen, Norway
e-mail:lars.asplin@imr.no

Michel d. S. Mesquita
Uni Research Climate, Uni Research
Bjerknes Centre for Climate Research
Bergen, Norway
e-mail: michel.mesquita@uni.no

Saurabh Bhardwaj
The Energy and Resources Institute
New Delhi, India
e-mail: saurabhbhardwaj.81@gmail.com

Laurent Bertino
Nansen Environmental and Remote Sensing Center
Bergen, Norway
e-mail: laurent.bertino@nersc.no

***Abstract*—Hexagon, the current High Performance Computing (HPC) resource at the University of Bergen, Norway is approaching its end of life. This article highlights some of the scientific results in the field of Climate Modelling obtained using this exceptional resource, a Cray XE6m-200 machine.** *(Abstract)*

I. INTRODUCTION

Hexagon originally was installed at the University of Bergen (UiB), Norway in January 2008 [1]. The initial configuration was a Cray XT4 distributed memory system consisting of 1388 nodes interconnected with a high-bandwidth low latency switch network (SeaStar2). Each node hosted a quad-core AMD Opteron chip with 4 GB memory (128 nodes had 8GB memory). The file system was Lustre and the system had a central storage facility with 288TB capacity. These characteristics might not seem very impressive to today's standards, however at the time of the installation the machine was very powerful (and got the highest ranking ever for a Norwegian HPC facility in top500: 66 (https://www.top500.org/list/2008/11/). The machine with its 5550 cores enabled Norwegian scientists to scale up their applications to a much higher number of cores and thus to pursue scientific problems that were out-of-reach earlier.

Due to the high operational cost of a supercomputer the life span of such a facility is usually 3-4 years. It takes about that time before one would get significantly more compute power for the same cost (less electricity per CPU power). Hexagon received a new “brain” in March 2012. Racks, power supplies, cabling and chassis backplanes were reused but the XT4 compute blades were replaced with XE6 versions. Service nodes (I/O and login) were upgraded with “XIO” blades. This is the current Cray XE6m-200 HPC installation at UiB. Beside the quadrupling of the number of cores (22272) the most significant upgrades were the Cray Gemini network chips and the AMD Opteron “Interlagos” CPUs.

The system has been part of the Norwegian national HPC infrastructure and been very popular in the user community for its performance, ease-of-use and reliability. The stability of the machine had a great impact on its maintenance costs, and allowed for a relatively small group of engineers to provide a very high level of service and support to the users of the machine.

The applications that benefit the most from this type of facility are big MPI based programs that require a lot of network and memory bandwidth. The research groups that match this category at UiB come from climate, ocean and other geophysical sciences. In the following sections a few use cases will be presented in more detail.

We would like to acknowledge the following people who have had a strong role in making Hexagon a success story: Petter Erling Bjørstad (Professor, Head of the Institute of Informatics, UiB), Klaus Johannsen (Research Director, Uni Research Computing), Erik Thorsnes (Senior Engineer, Group leader, Uni Research Computing), Alexander Oltu (Senior Engineer, Scientific Computing, IT-department, UiB) and Lorand Janos Szentannai (Engineer, Scientific Computing, IT-department, UiB).

II. HEXAGON A WORKHORSE FOR NORWEGIAN CLIMATE MODELLING

Changes in climate involve a complex interplay of physical, chemical and biological processes of the atmosphere, land surface, cryosphere and ocean. Earth System Models (ESMs) are today the only means for aggregation and compression of all information necessary to quantify dynamics and impacts of ongoing climate change and are emerging as the dominant modelling tool for numerical simulations addressing e.g. anthropogenic versus natural climate change, climate variations with time scales ranging from years to millennia, and climate predictability. There is a continues need for model development and multi-model diversity in order to provide best possible future assessments with reliable uncertainty estimates.

The Norwegian Earth System Model ([2],[3],[4]) is a national model infrastructure that serves for integrating new process knowledge of the Earth system and as a laboratory for exploring the dynamics of climate. The establishment of NorESM, which coincided with the initial employment of Hexagon Cray XT4, brought together complementary modelling groups from Norway with predominantly specializations in atmospheric and ocean science. NorESM builds on the Community Earth System Model [5] while including important national contributions, such as advanced representations of chemistry-aerosol-cloud processes and an isopycnic ocean component with biogeochemistry formulations adapted to isopycnic coordinates.

The original NorESM is configured on global 2° atmospheric and 1° ocean grids with 26 and 53 levels, respectively, giving a model state size of O(108). During integration, the model components run in a mixed sequential/concurrent mode and exchange information through a coupler (**Figure 1**a). The system is further parallelized using horizontal domain decomposition for all components (**Figure 1**b) and additional vertical decomposition for the atmospheric component. The model uses Parallel netCDF [6] for fast I/O and is configurable as a pure MPI or optionally hybrid MPI/OpenMP application. Within constraints and considering the complexity of parallelization problem, NorESM shows good scalability on Hexagon, particularly for the ocean component and very good parallel I/O performance (**Figure 1** c). Depending on complexity and resolution, individual NorESM simulations on hexagon have used between ~100 and ~3000 processing cores.

During the period 2008 to 2016, Hexagon hosted more than 80% of all NorESM integrations, totaling a consumption of 200 million CPU hours with exponential increase over time (**Figure 1** d-e). In October 2016 we carried out a survey to map the past, current and planned use of NorESM. This revealed about 50 national users distributed over 7 institutions (see **Figure 1** f). In the survey the participants were asked to list project involvements where NorESM was/is developed or used for doing specific experiments and/or assessing NorESM output. A total of 65 project involvements were reported and the funding source distribution is shown in **Figure 1** g. The total requested grant size for all projects amounts to about 1491 MNOK (~175 million USD) and is distributed among funding sources as shown in **Figure 1** h. At the Nordic level, NorESM is used in e.g. Nordic Council of Ministers/NordForsk projects CRAICC and eSTICC and there is an active collaboration established on atmospheric chemistry-aerosols-clouds with University of Stockholm and University of Helsinki and on ocean processes with University of Copenhagen.

NorESM contributed to the latest Intergovernmental Panel for Climate Change assessment report (IPCC AR5,[7]) - which constitutes the knowledge basis of the Paris Agreement - through participation in the Coupled Model Intercomparison Project phase 5 (CMIP5,[8]). By February 2017, the output generated with NorESM on Hexagon had been used in more than 500 peer-reviewed publications (https://cmip-publications.llnl.gov). This demonstrates the enormous international interest and visibility by contributing to such model intercomparisons.

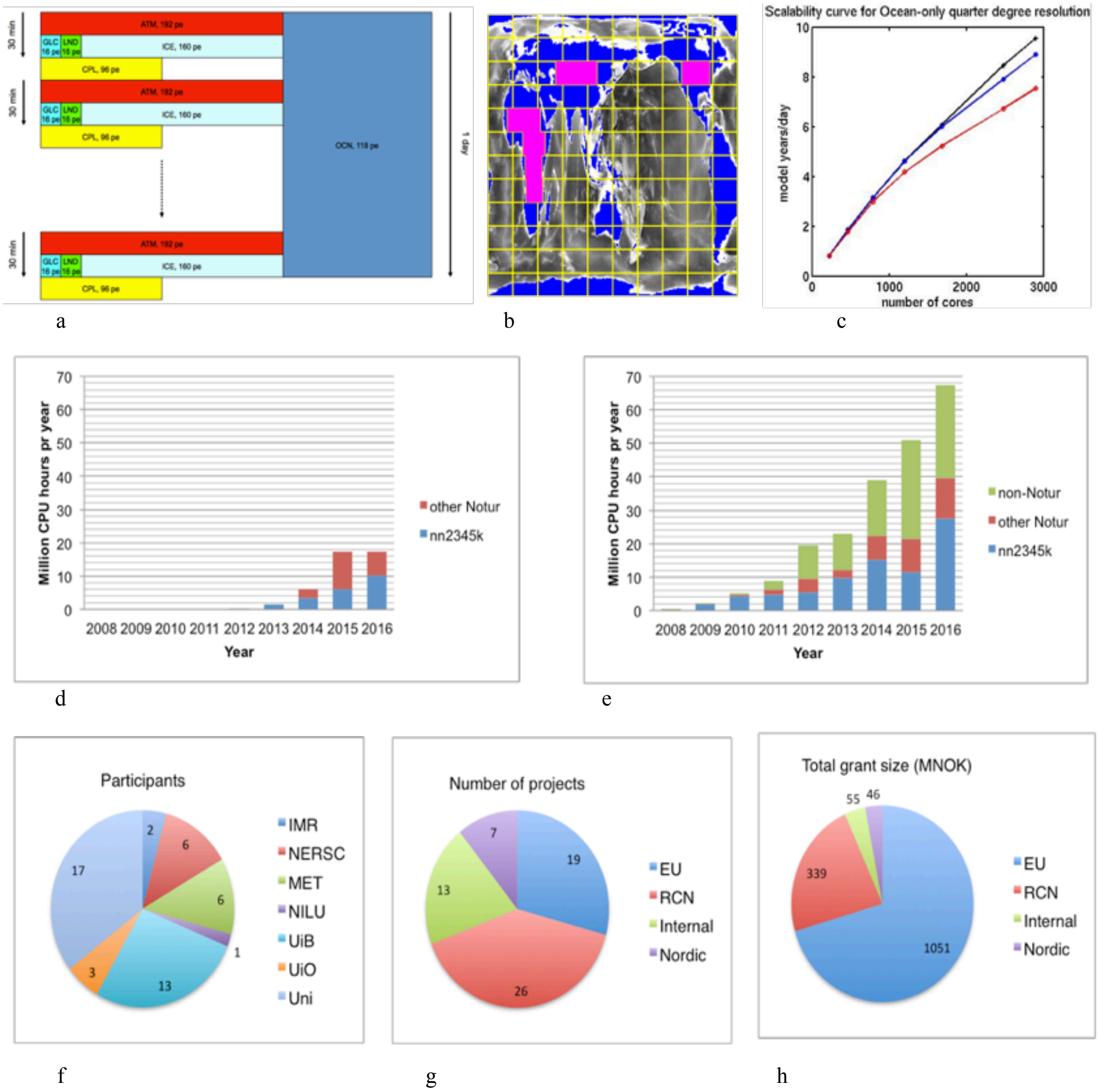

__Figure 1__: __a__, Diagram explaining the mixed sequential/concurrent integration of components. __b__, Domain decomposition of ocean component. __c__, Scaling of NorESM2's 1/4° ocean component using parallel I/O (black) versus sequential I/O (blue) for writing default output versus parallel I/O for writing 21 GB per simulation day enhanced output (red). Numbers do not include processors assigned to sea ice and land components. __d__, NorESM's CPU usage on Hexagon. __e__, NorESM's CPU usage on other national facilities (vilje). __f__, User survey participants sorted according to institutes. __g__, Projects that use NorESM sorted according to funding source. __h__, Funding amount of projects that use NorESM sorted according to funding source.

In summary, Hexagon has been the workhorse of Norwegian climate modelling activities with demonstrable scientific and societal impacts. With preparations for the upcoming CMIP6 [9] underway, climate modelling enters now a new phase that constitutes a step change in computational demands. As result of increased complexity and resolution, NorESM's successor NorESM2 will require 20 times as much resources, making it necessary to find a bigger horse.

III. The NorKyst800 – a current model system for the Norwegian Coast

Aquaculture is a vital industry in Norway, and a sustainable production is a precondition for long-term development and growth where the measure of sustainability is the natural ecosystem. It is important to ensure a clean marine environment and good production locations for aquaculture, with minimum impact from long-distance transport emissions and pollution from more local sources. To assist in the development and improvement of the aquaculture production and to find healthy adaptations to the marine environment and biological diversity, knowledge on ocean currents is crucial.

To collect information of the physical environment is difficult because of the complex dynamics on numerous length scales starting at

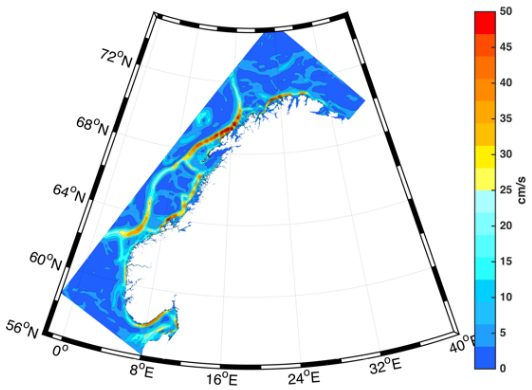

***Figure 2**: Map showing the extension of the NorKyst800 model. The colors denote the long-term average of surface current speed. The mean current patterns are mainly directed northwards, and the color scale denotes speed in cm/s.*

O(100m) and since the geographical extension is so large. Norway has a long and complicated coast with numerous fjords and islands. Along the coast flows the Norwegian Costal Current which has great impact on water exchanges with and within fjords, and reproducing a realistic physical state with a numerical hydrodynamical model requires high resolution, both for the numerical solver and for the forcing data (atmospheric conditions, river discharge etc.). Collecting data from observations of current and hydrography is important, but too costly to fulfil the actual need. Thus, a combination of observations and numerical current modelling is the most cost-efficient and best approach to get the necessary environmental information.

The NorKyst800 model system [10] is a compromise between geographical extension of the model domain and sufficient resolution. The whole Norwegian coast is possible to simulate (**Figure 2**), and horizontally we use 2600x900 grid points. Vertically we use 35 terrain-following levels in our standard experiments,

however an increase of vertical levels will be preferred when computers attain increased capacity. We have used the Regional Ocean Current Model System (ROMS) as the hydrodynamical model embedded in the NorKyst800 system

([10], [11], or see http://myroms.org). ROMS is an open-source, state-of-the-art, three-dimensional, free-surface, primitive equation numerical model using a generalized terrain-following s-coordinate in the vertical.

The NorKyst800 model system has been tested widely on Hexagon to configure a standard set up. The standard model applies 1024 processors for about 20 days (~500.000 core hours) to complete a one-year simulation. A more than 20-years long hindcast archive has been produced and will be used in several projects and applications requiring high-resolution current fields from the Norwegian coast. In addition, the NorKyst800 model is a necessary step to execute even higher resolution fjord models (typically 160m resolution), and several computer demanding efforts have resulted in valuable outcomes within different scientific disciplines (as e.g. described in [12], [13], [14] and [15]). Finally, the computational efforts on Hexagon have also assisted the Norwegian Meteorological Institute in implementing the same model in their operational suite providing daily high-resolution ocean current forecasts.

IV. Hexagon provides insights into climate change in India

Since 2009, a series of four collaborative projects have looked at climate change science, variability and its impacts in India: a) the Norwegian Framework Agreement Phase I funded by Norwegian Ministry of Foreign Affairs, which focused on capacity building and regional climate modelling; b) NORINDIA, which emphasized the study of the hydrological impacts of climate change [16], funded by the Norwegian Research Council; c) C-ICE, an ongoing project which studies how changes of sea ice around Antarctica could affect the Indian Summer monsoon, also funded by the Norwegian Research Council and the Indian Ministry of Earth Sciences; and d) PREPARE, a new project under Norwegian Framework Agreement Phase - II, which studies climate extremes in the future and intraseasonal variability, funded by Norwegian Ministry of Foreign affairs. All of these projects have had one thing in common: they have used Hexagon.

Also, these projects were made possible through the collaboration with TERI (The Energy and Resources Institute) in New Delhi. TERI Researchers were given access to Hexagon and a series of high resolution simulations were conducted using the Weather Research and Forecasting (WRF) model and the Norwegian Earth System model. They were also involved in an online training program of learning how to use WRF [17], as well as on how to run WRF on Hexagon - both courses run on m2lab.org

Some of the climate runs were made in a 'tropical channel' mode (Fig. 1), where the model domain consists of the boundaries above and below a certain latitude and no side boundaries. This configuration allows for the interaction of weather systems from the extra-tropics through the north-and-south boundaries. Also, the generated waves are allowed to propagate around the tropical belt more naturally - as in the real world and in global models, which improves predictability.

In NORINDIA, we have used a different type of set-up: that of coupling WRF to an advanced hydrological modelling system. This makes it possible to conduct high resolution simulations of the atmosphere, which feed into a high-resolution hydrological system. Output from such configuration can then be used to assess water availability in selected basins in India ([18],[19]). All in all, Hexagon has contributed greatly to the advance of the understanding of climate change in India.

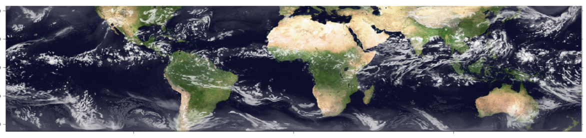

Figure 3*: Li, L., Gochis, D., Sobolowski, S., and Mesquita, M.d.S. (2017) Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model. Journal of Geophysical Research - Atmospheres, accepted.*

V. Operational Oceanography

Operational oceanography is traditionally CPU-hungry in the sense that it requires both high horizontal resolution to resolve the ocean mesoscale eddies (typically 50 to 300 km wide, these are essential for forecasting the drift of objects at the ocean surface, such as oil spills or search and rescue operations) and also expensive data assimilation methods able to constrain complex 3-dimensional ocean currents using satellite observations at the surface and sparse in-situ profiles from autonomous buoys. The availability of computing resources has thus always been - and still remains - a major bottleneck for reaching the goals of operational oceanography.

The TOPAZ system is the Norwegian contribution to international operational oceanography and constitutes the Arctic Marine Forecasting Center in the European Copernicus Marine Services. Before Hexagon, NERSC was able to run the TOPAZ system in near real time forecast mode using an ensemble Kalman Filter [20] with the HYCOM model applied at 12 km resolution, but without any biological component. The introduction of the machine has triggered the production of an Arctic reanalysis [22],[23] for the physical ice-ocean system, but also the production of a biogeochemical reanalysis with the coupled HYCOM-NORWECOM model, assimilating both physical observations and satellite ocean color data [24] .The forecast and re-analysis products have become the main Arctic contributions of the European Copernicus Marine Services in 2015, see http://marine.copernicus.eu for more information. Figures 1 and 2 show examples of sea ice coverage in the Arctic summer 2016 and the corresponding map of ocean primary production showing that some of the production grows close to the ice edge. Before the physical, biogeochemical and data assimilation modules were integrated on Hexagon, it was not possible to produce those maps consistently, based on the same input data.
Such ocean reanalyses are important as reference databases in order to judge whether or not the present situation is usual or not and in order to appreciate different aspects of recent climate change, which is especially fast in the Arctic Ocean, see for example the case of ocean fluxes to the Arctic in [25]. The production of reanalyses also represents an invaluable testbed for assessing the stability of data assimilation methods.
NERSC has recently started to work on a 6 km resolution of the TOPAZ system and a NeXt generation Sea Ice Model NeXtSIM [26] that are both expected to better meet the needs of the scientific and operational users in the Arctic.
as a footnote.

A. Acknowledgment

L. Bertino acknowledges a grant of CPU from the Norwegian supercomputing project number nn2993k and the NordForsk Center of Excellence EmblA.

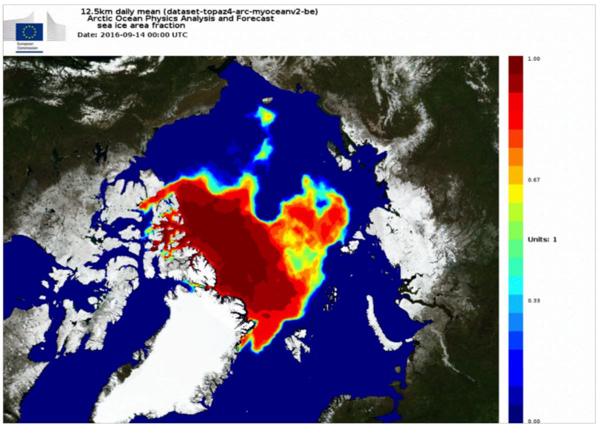

Figure 4: Example of sea ice forecast on the 14th September 2016 from the TOPAZ system. Red is full ice coverage.

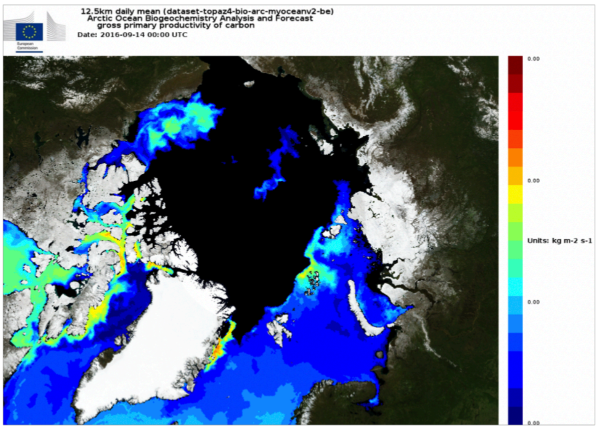

Figure 5: Surface primary production on the same day as in Figure 4. Yellow and red areas are the most productive on that day.

VI. References

[1] Thorsnes, E.2012: Hexagon, The UIB Supercomputer – has a new brain. Meta Magazine, 2012-2

[2] Bentsen, M., I. Bethke, J. B. Debernard, T. Iversen, A. Kirkevåg, Ø. Seland, H. Drange, et al. "The Norwegian Earth System Model, NorESM1-M – Part 1: Description and Basic Evaluation of the Physical Climate." Geoscientific Model Development 6, no. 3 (May 24, 2013): 687–720. doi:10.5194/gmd-6-687-2013.

[3] Iversen, T., M. Bentsen, I. Bethke, J. B. Debernard, A. Kirkevåg, Ø. Seland, H. Drange, et al. "The Norwegian Earth System Model, NorESM1-M – Part 2: Climate Response and Scenario Projections." Geoscientific Model Development 6, no. 2 (March 22, 2013): 389–415. doi:10.5194/gmd-6-389-2013.

[4] Tjiputra, J. F., C. Roelandt, M. Bentsen, D. M. Lawrence, T. Lorentzen, J. Schwinger, Ø. Seland, and C. Heinze. "Evaluation of the Carbon Cycle Components in the Norwegian Earth System Model (NorESM)." Geoscientific Model Development 6, no. 2 (March 4, 2013): 301–25. doi:10.5194/gmd-6-301-2013.

[5] Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall, 2013: The Community Earth System Model. Bull. Amer. Meteor. Soc., 94, 1339–1360, doi:10.1175/BAMS-D-12-00121.1.

[6] Li, Jianwei, Wei-keng Liao, Alok Choudhary, Robert Ross, Rajeev Thakur, William Gropp, and Rob Latham. "Parallel netCDF: A Scientific High-Performance I/O Interface." Eprint arXiv:cs/0306048, June 11, 2003. http://arxiv.org/abs/cs/0306048.

[7] Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, B. and Midgley, B.M., 2013. IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.

[8] Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl. "An Overview of CMIP5 and the Experiment Design." Bulletin of the American Meteorological Society 93, no. 4 (April 1, 2012): 485–98. doi:10.1175/BAMS-D-11-00094.1.

[9] Eyring, Veronika, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization." Geoscientific Model Development 9, no. 5 (May 26, 2016): 1937–58. doi:10.5194/gmd-9-1937-2016.

[10] Albretsen J, Sperrevik AK, Staalstrøm A, Sandvik AD, Vikebø F & Asplin L (2011) NorKyst-800 Report No. 1 - User Manual and Technical Descriptions. Fisken og Havet 2/2011, Institute of Marine Research.

[11] Asplin L, Johnsen IA, Sandvik AD, Albretsen J, Sundfjord V, Aure J & Boxaspen KK (2014) Dispersion of salmon lice in the Hardangerfjord. Mar. Bio. Res. 10, 216-225.

[12] Espeland SH, Albretsen J, Olsen EM & Bodvin T (2015) Modelling drift of pelagic offspring: the importance of egg surveys in providing a realistic model initialization. ICES J. of Mar. Sci., doi:10.1093/icesjms/fsv134.

[13] Haidvogel DB, Arango H, Budgell WP, Cornuelle BD, Curchitser E, Di Lorenzo E, Fennel K, Geyer WR, Hermann AJ, Lanerolle L, Levin J, McWilliams JC, Miller AJ, Moore AM, Powell TM, Shchepetkin AF, Sherwood CR, Signell RP, Warner JC & Wilkin J (2008) Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System. J. of Comp. Phys., 227, 3595-3624.

[14] Johnsen IA, Asplin L, Sandvik AD & Serra-Llinares RM (2016) Salmon lice dispersion in a northern Norwegian fjord system and the impact of vertical movements. Aquacult. Environ. Interact., 8, 99–116.

[15] Samsing F, Johnsen IA, Stien LH, Oppedal F, Albretsen J, Asplin L & Dempster T (2016) Predicting the effectiveness of depth-based technologies to prevent salmon lice infection using a dispersal model. Preventive Veterinary Medicine, 129, 48-57, doi:10.2016/j.prevetmed.2016.05.010.

[16] Sandvik AD, Bjørn PA, Ådlandsvik B, Asplin L, Skarðhamar J, Johnsen IA, Myksvoll M & Skogen MD (2016) Toward a model-based prediction system for salmon lice infestation pressure, Aquacult. Environ. Interact., 8, 527–542.

[17] Shchepetkin AF & McWilliams JC (2005). The Regional Ocean Modeling System: A split-explicit, free-surface, topography following coordinates ocean model, Ocean Modelling, 9, 347-404.

[18] Li, L., Gochis, D., Sobolowski, S., and Mesquita, M.d.S. (2017) Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model. Journal of Geophysical Research - Atmospheres, accepted.

[19] Mesquita, M.d.S., Veldore, V., Li, L., Krishnan, R., Orsolini, Y., Senan, R., et al. (2016). Forecasting India's Water Future - Eos. Eos, 97. http://doi.org/10.1029/2016eo049099

[20] Walton, P. J., Yarker, M. B., Mesquita, M. D. S., & Otto, F. E. L. (2015). Helping to Make Sense of Regional Climate Modeling: Professional Development for Scientists and Decision Makers Anytime, Anywhere. Bulletin of the American Meteorological

Society, BAMS–D–14–00111.1. http://doi.org/10.1175/BAMS-D-14-00111.1

[21] G. Evensen, Data assimilation: the Ensemble Kalman Filter. Springer, 2009 (2nd Ed.)

[22] P. Sakov, F. Counillon, L. Bertino, K. A. Lisæter, P. R. Oke, and A. Korablev, “TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic,” Ocean Sci., vol. 8, no. 4, pp. 633–656, 2012. [Online]. Available: http://www.ocean-sci.net/8/633/2012/

[23] J. Xie, L. Bertino, F. Counillon, K. A. Lisæter, and P. Sakov, “Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 19912013,” Ocean Science, vol. 13, no. 1, pp. 123–144, 2017. [Online]. Available: http://www.ocean-sci.net/13/123/2017/

[24] E. Simon, A. Samuelsen, L. Bertino, and S. Mouysset, “Experiences in multiyear combined state-parameter estimation with an ecosystem model of the North Atlantic and Arctic Oceans using the Ensemble Kalman Filter,” Journal of Marine Systems, vol. 152, pp. 1–17, 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0924796315001190

[25] V. S. Lien, S. S. Hjøllo, M. D. Skogen, E. Svendsen, H. Wehde, L. Bertino, F. Counillon, M. Chevallier, and G. Garric, “An assessment of the added value from data assimilation on modelled Nordic Seas hydrography and ocean transports,” Ocean Modelling, vol. 99, pp. 43–59, mar 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1463500315002462

[26] P. Rampal, S. Bouillon, E. Ólason, and M. Morlighem, “Nextsim: a new lagrangian sea ice model,” The Cryosphere, vol. 10, no. 3, pp. 1055–1073, 2016. [Online]. Available: http://www.the-cryosphere.net/10/1055/201