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. 2.4

Report on the application of the fully coupled SCM to test cases based on Arctic YOPP IOPs

Submission of Deliverable

Work Package	WP2 Enhanced weather and climate models
Deliverable No	2.4
Deliverable title	Report on the application of the fully coupled SCM to test cases based on Arctic YOPP IOPs
Version	1
Status	Final
Dissemination level	PU - Public
Lead Beneficiary	10 - SU
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	31 October 2019
Delivery Date	31 October 2019
Coordinating author	Gunilla Svensson (gunilla@misu.su.se)
Contributing authors	Kerstin Hartung (SU and Ludwig-Maximilians-Universität Munich), Jareth Holt (SU) and Michael Tjernström (SU)

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 new tool, the Atmosphere-Ocean Single-Column Model (AOSCM), developed within APPLICATE and demonstrated in D2.1, is further developed to become an even more versatile tool for understanding model behaviour and aid model development. The enhancements of the model are mostly on the technical side with the goal to be able to run well-defined cases that can be compared with process observations. Procedures on how to set up cases have been defined and are described in a document appended to this deliverable.

Observational process level data from the Arctic Ocean obtained in expeditions with the ice-breaker Oden is described and merged into files designed with the purpose of being easily accessible for model development. These observations are from three summer expeditions and provide data in the vertical column that is represented by the AOSCM. The methodology is developed in an international setting and will be utilized for data from the ongoing expedition MOSAiC that will provide data for a whole year October 2019 – October 2020.

An experimental protocol is developed for the AOSCM and tested for a warm air advection case observed during the Oden expedition ACAS in 2014. The protocol covers perturbations of model setup and forcing and consists of 480 simulations of the three-day period. The perturbation analysis reveals that the net energy available at the surface during this period vary between 30 and 130 Wm^{-2}. The results are most sensitive to the advection of moisture but substantial changes are seen when using all forcing and modifying other parameters such as the sea-ice properties, time step, model version etc. The methodology of using the AOSCM and combining it with observations are now mature enough to be expanded to cover the whole set of observational periods.

This work contributes to Objective #2 of WP2 “Develop innovative methods, using observations and a variety of model configurations, to facilitate parameter optimization for physical processes in coupled model systems for NWP and climate”.

1. INTRODUCTION

1.1. Background and objectives

Improving model performance through parameterisation development is a difficult and tedious task as newly developed and implemented schemes seldom increase the skill of a complex model immediately. The task at hand usually involves untangling the model behaviour in terms of existing compensating errors, unbalanced physics representations and numerical issues before the benefit of adding new more advanced physically based improved schemes is revealed. The work reported here aims to facilitate improved understanding of the model behaviour in the parameterisation development procedure. This work consists of three major components: 1) further development of the Atmosphere-Ocean Single-Column Model (AOSCM) that was demonstrated in D2.1; 2) organisation of observational data to be used for process evaluation to understand and improve parameterisations and coupling methods and 3) provide one example of how the tool currently is used. The reported work is part of Task 2.2.1.

1.2. Organisation of this report

Section 2 describes the ingredients needed to be able to set up a test case for the AOSCM. It discusses organisation of appropriate observational data sets, capability developments of the AOSCM, how to create forcing and setup specific cases with the AOSCM. Section 3 proposes a methodology to investigate sensitivities to forcing, model setup and parameter choices for a specific test case. The methodology is under development and here we present results from an interesting period picked from one of the three observational datasets that we are currently working with. Conclusions and an outlook for the continued work in Task 2.2.1 are found in Section 4.

2. Developing test cases

2.1 Observational datasets within APPLICATE

Observational datasets play a crucial role for the understanding of processes that are to be described as parameterisations in models as well as evaluation and verification of model performance. Implemented model parameterisations are based mostly on observational data from the mid- and low latitudes where adequate data exists. It is thus important to inventory existing process-resolving data and to facilitate their efficient usage. For the advancement of model development for the coupled atmosphere-ocean sea ice system it is of extra interest when data from all three components are co-located. There are very few such datasets from the central Arctic Ocean. Before the MOSAiC campaign such data during winter is available from the Surface Heat Budget of the Arctic Ocean (SHEBA; e.g. Uttal 2002) expedition. Limited data is also available from the Norwegian Young sea ICE (N-ICE2015; e.g. Cohen et al. 2017) expedition. For summer conditions the situation is somewhat better. For APPLICATE we have prepared three datasets targeted for use in model development, from expeditions on the Swedish icebreaker Oden (see Figure 1): *Arctic Summer Cloud Ocean Study* (ASCOS, e.g. Tjernström et al. 2012; 2014) in 2008, *Arctic Clouds in Summer Experiment* (ACSE, e.g. Tjernström et al. 2015; 2019, Sotiropoulou et al. 2016) in 2014 and the Arctic Ocean 2018 expedition (AO2018; Vuellers et al, manuscript in preparation) in 2018. We describe the datasets and how we are supporting a more efficient usage through systematic organisation of the data.

Organizing observational data in a way that is usable by a broad community is far from trivial. APPLICATE researchers have been very active in defining a standard protocol for data collected at supersites for process studies that consider FAIR principles, semantics, file format etc. This standard protocol also applies to model data. The development draws on experiences from CMIP (Coupled Model Intercomparison Project), especially CFMIP (Cloud Feedback Model Intercomparison Project) sites (cfSites) and GEWEX Global Atmosphere System Study (GASS) and is further developed within the WWRP Polar Prediction Project (PPP) sub-project YOPPsiteMIP (see D8.9). YOPPsiteMIP organises Arctic supersite data and model data during the YOPP SOPs, and the interest here is the Arctic sites during SOP-NH1/2 and the upcoming MOSAiC dataset.

When the work in Task 2.2.1 was defined, the MOSAiC expedition was planned for 2018/2019. As the expedition shifted by one year and the organisation of observational data within YOPPsiteMIP takes time (Utqiaġvik formerly Barrow is almost ready to be published in the new format), we have utilized field data collected from expeditions with the ice breaker Oden as well as the land station Sodankylä in northern Finland. Since Task 2.2.1 continues to the end of the project, analysis and simulations based around the MOSAiC supersite will be the main activity during the coming year.

Arctic expedition datasets with the ice-breaker Oden for APPLICATE

Sampling strategy

Common for all three experimental datasets discussed here is an attempt to monitor state, motion and processes in a vertical column including the upper ocean, sea-ice and troposphere with multiple instruments and sensors.

Under the ice profiles of temperature and salinity were observed with different forms of CTD instruments and microstructure probes. Profiles of temperature through the sea ice and the surface temperature are observed with thermocouple instruments and IR thermometers. Near the surface we monitor as many components of the surface energy budget as possible: broadband solar and infrared radiation, eddy-covariance fluxes of momentum, sensible and latent heat. Close to the surface, standard weather station meteorological quantities (pressure, temperature, atmospheric moisture, and wind speed and direction) are sampled as well as visibility and precipitation intensity. For the atmospheric column we performed 6-hourly soundings with free flying balloons. Vertical profiles of clouds were gathered by vertically pointing Doppler cloud radars. Vertically pointing lidar also provided cloud information and aerosol backscatter. Thermal profiles, as well as integrated liquid water and water vapor, are derived from scanning microwave radiometers. Winds were estimated from a variety of wind profiling radars and scanning lidar

The exact instrumentation and details in the spatial and temporal cover is somewhat different from different expeditions.

Figure 1. Top left: *Oden during ASCOS with micro-meteorology masts and ocean site in foreground.* **Top right**: *Foredeck of Oden during ACSE, with cloud and wind radars, scanning lidar and microwave radiometer and the bow mast for eddy covariance fluxes in the background.* **Bottom**: *Oden viewed from the top of the bow mast during AO2018, with similar instruments on the foredeck (scanning cloud radar and lidar, and radiometer) and a suite of atmospheric observations on the top 7th deck.*

ASCOS featured a three-week ice drift, with Oden moored to and drifting with the ice. Eddy-covariance fluxes were measured from a mast below the ice and quasi-continuous microstructure ocean profiling of thermal and turbulence structure down to ~400 m was conducted. Incoming, reflected and transmitted spectrally-resolved solar radiation was also measured. A tethered balloon for profiling was deployed on the ice along with a sodar for wind measurements. All these observations were performed on masts or other installations on the ice. A 449MHZ wind profiling radar on-board also measured wind speed and direction. The aforementioned observations are only available during the quasi-stationary three-week ice drift; all other observations, including cloud profiles from a K-Band radar are available for the entire expedition. No lidar was used except for a ceilometer lidar also detecting backscatter profiles.

ACSE was navigating quasi-continuously, first along the Russian shelf break and later further off shore, between Tromsö and Barrow (now Utqiaġvik), stopping only briefly for research stations. Hence there were no on-ice measurements; all observations were carried out on-board. For ocean profiling relatively frequent CTD casts are available but no continuous observations were made. Continuous eddy-covariance surface fluxes were measured from a bow mast on the ship, but only incoming broad-band radiation could be measured. A motion-stabilized W-Band Doppler cloud radar (with vertical range limited to 6km) as well as a 3D scanning lidar for profiling of aerosols and VAD-winds, complementing a 449MHZ wind profiling radar, were deployed.

AO2018 was similar to ASCOS in that the main portion was a month-long ice camp. Limited CTD and microstructure measurements under the ice are available from another project and possibly spectrally-resolved incoming, reflected and transmitted radiation as well. Eddy-covariances, broad-band radiation and surface temperature are available from the ship for the whole expedition and from an installation on the ice for the ice drift. Tethered soundings are also available from the ice. A scanning K-Band Doppler cloud radar was deployed on the ship only during the ice drift, and a 3D scanning Doppler lidar measured wind profiles (VAD) through the expedition.

2.2 Capability enhancements of the AOSCM

We have further developed the AOSCM to become a more versatile tool since the first version presented in Hartung et al. (2018). Substantial work has gone into improving the control over the initial state and the forcing during the simulation. The two main developments are procedures on how to initialize the coupler and the sea-ice state. The AOSCM is also updated with the option of using the most recent version of OpenIFS (based on IFS version 46r3) while still being back-compatible with the earlier version (based on IFS version 40r1). The capability to use ERA5 (C3S 2017) as initial and forcing conditions instead of ERA-Interim (Dee et al. 2011) is another significant improvement. The newer reanalysis is available with hourly data and the same vertical resolution as the current standard version of IFS. Thus, the degradation in quality from interpolating onto a different vertical grid and between the 6-hourly fields is much reduced.

2.3 Procedure on how to set up an AOSCM case

The AOSCM can (in principle) be applied at any location on Earth. It can also be run using idealized settings, or more loosely based on observations (see Hartung et al., 2015). However, here we mainly discuss cases at a certain single location. Setting up cases that are as consistent as possible with the three-dimensional model creates an environment where sensitivity tests and parameterization development can be done utilizing process-resolving observational data (see Section 2.1).

When applied over a land point, the atmospheric column of the AOSCM is coupled to the land surface. When an ocean point is of interest, the atmospheric column can be run applying forcing at the lower boundary condition or coupled to the ocean model. In both cases, the vertical profiles of temperature and humidity is needed for the initial state and when coupled to the ocean, initial vertical profiles of temperature, salinity and chlorophyll as well. The most complex setup is when a sea ice covered location is chosen, since the sea ice model utilizes five sea-ice categories.

There are a number of steps to go through when setting up a case. Details for each of the steps below are provided in the Appendix, along with more discussion of the setup, input files, and variables for each model component. The basic steps are:

1. Choose a location and time period. This is typically based on a measurement campaign. A compromise between the exact campaign location and the points for which we can find data is needed. Auxiliary files for the ocean model must be updated to reflect the actual latitude, longitude, and bathymetry.
2. Get atmospheric reanalysis data. Software is available to extract all relevant variables from the ERA5 reanalysis. We also use certain surface fluxes from the 4D-Var product to create ocean forcing and coupler initial conditions files.
3. Get oceanic reanalysis data. Both the ORAS4 and ORAS5 (Ocean ReAnalysis System) datasets can provide initial conditions for the ocean model. The ORAS5 product also provides some surface fluxes and sea ice data.
4. Get a restart file for the sea ice model. We currently extract sea ice initial conditions from the restart files generated by full EC-Earth simulations or from a ocean only run forced with atmospheric reanalysis. We are exploring other options for adapting ice data as initial conditions.
5. Create input files for each component. This step is crucial to building a consistent case study. This step involves:
   - Adjusting the sea ice concentration, sea surface temperature, and possibly sea ice albedo in the atmospheric input files to match the ocean/ice data;
   - Using the ERA5 data and 4D-Var fluxes to create ocean forcing files; and
   - Creating initial conditions for the coupler variables by combining 4D-Var downward fluxes, atmospheric near-surface conditions, sea ice conditions (concentration, temperature, and albedo), and sea surface temperature with bulk formulae to calculate the required net fluxes.
6. Incorporate observational data. If available, observational data can be blended with the preceding to create hybrid input files. For the atmospheric component, this requires translating from height or pressure to hybrid-sigma coordinates. For the oceanic component, the translation of temperature and pressure depends on the equation of state chosen.

After following these steps, a case is built which can serve as the default setup and from which controlled perturbation simulations can be created.

3. Example case study with the AOSCM

Here, we provide one example of a study with the AOSCM based on a case from one of the field datasets presented in Section 2.1. Two other periods, selected from the other two datasets presented in Section 2.2. are in focus for being studied using the AOSCM. The ice camp part of ASCOS (three weeks) is used as the testbed for the parameterisation development work reported on in D2.3. Test so far have been in atmosphere-only simulations but with the newly developed capabilities to properly initialise the sea ice, tests in coupled mode follow next, work that will be reported on in D2.7. To increase the statistics for the sensitivity studies, the ice camp part of AO2018 is also used.

The example we choose to report on here is further developing the case reported on in D2.1 and is currently documented in a manuscript (Hartung et al., in prep). We analysed the model sensitivities to a range of parameters for the extreme warm air advection period during the ACSE campaign (see Section 2. 1 and Tjernström et al., 2016). This period was also studied with LES (Sotiropoulou et al., 2018) and in idealized way with the AOSCM (Hartung et al., 2018). We focus on parameter sensitivities because the most optimal parameter settings are not always known beforehand, and the AOSCM allows us to run multiple setups with relatively small computational effort. The results help discern which parameters are most crucial to the system's dynamics during the experiment.

All simulations are initialised on August 1, 06 UTC 2014 and have a simulation time of 72 h. The model atmosphere has 137 vertical levels and, if present, the ocean has 17 levels to a depth of about 41 m. In coupled simulations the ocean initial profile is obtained from ORAS4 analysis and chlorophyll for August from SeaWiFS. Restart files are compiled based on information from the first timestep and the ERA5 forecast for variables not available in the reanalysis (e.g. turbulent fluxes). The first 12 h are treated as spin-up but are still analysed separately as they also contain information about the performance of the system.

Sensitivity to the following perturbations are included so far:

- Atmospheric model version, either OIFS cy40r1 (which is presented in Hartung et al, 2018) or the more recent cy43r3 (2 cases).
- Compilation with default or adapted surface conditions (2 cases). The adapted conditions include a roughness length increased to 0.03 m to match the observed friction velocity (similar to Sotiropoulou et al., 2018) and a surface conductivity increased from 58 to 1010 $Wm^{-2}K^{-1}$ to enforce direct coupling with the surface (i.e. the skin temperature of the atmospheric model is directly equal to the sea-ice temperature passed from the coupler; see Deliverable 2.2 for more information).
- Model and coupling time step, either 450 or 900 s (2 cases).
- Atmospheric forcing and initial conditions from the ERA-5 high-resolution, hourly control run. One sensitivity experiment sounding data to modify the temperature and moisture initial profiles in order to better capture the observed cloud evolution (2 cases).
- Different combinations of atmospheric forcing terms (6 cases). Each of the 4 forcing components (horizontal advection of heat, moisture, and horizontal momentum as well as vertical advection) are here enabled or disabled independently of the others. In this way we can assess the importance of each term, particularly the advection of moisture and heat as in Sotiropoulou et al. (2018). The 6 cases include all forcings being applied (the default); all forcings except one of the four applied; and all forcings without both temperature and moisture advection.
- Sea ice properties either from a default thickness distribution or a global EC-Earth simulation. The former approach uses a sea ice thickness of 1 m, a sea-ice concentration of 100 % and a snow thickness of 0 m. In the latter approach, points are chosen to cover a wide range of sea-ice concentrations (50 to 80 %) in the general vicinity of the icebreaker Oden (160 ^{o}E, 75 ^{o}N) during the period of interest (5 cases).

In total, 480 simulations are analysed using every combination of perturbations: model version (2), surface conditions (2), time step (2), atmospheric initial conditions (2), atmospheric forcing terms (6), and sea ice concentrations (5). The results are summarized as averages over the initial 12 hours and the remaining 60h such as all simulations can easily be compared.

3.1 Results

Sotiropoulou et al. (2018) focused on the additional energy that the surface received due to the clouds present in the very warm and moist air that was advected over the sea ice. This study illustrated the importance of advection in maintaining the clouds and thus the enhancement of the surface fluxes. When forced the same idealized way as the LES, the AOSCM results also depends strongly on the advection but exhibits quite different evolution (not shown).

Figure 2 shows the cloud evolution (liquid water content) for the simulated period for all advective forcing (standard setup, see Figure 3) as well as setups when various advection terms are turned off for the ACSE case built using reanalysis (following the procedure described in Section 2.3). The initial conditions and advection at the location supports two layers of low clouds, one with a top at about 200 m above sea level and one at about 850 m. From the perturbation simulation when removing the moisture advection, it is clear that the upper layer is more dependent on advection. The case with no sensible heat advection shows much more cloudiness and when the vertical advection is removed, the clouds are decreasing in amount and liquid water content. Some of the cases also have clouds at higher elevation later in the simulation (not shown).

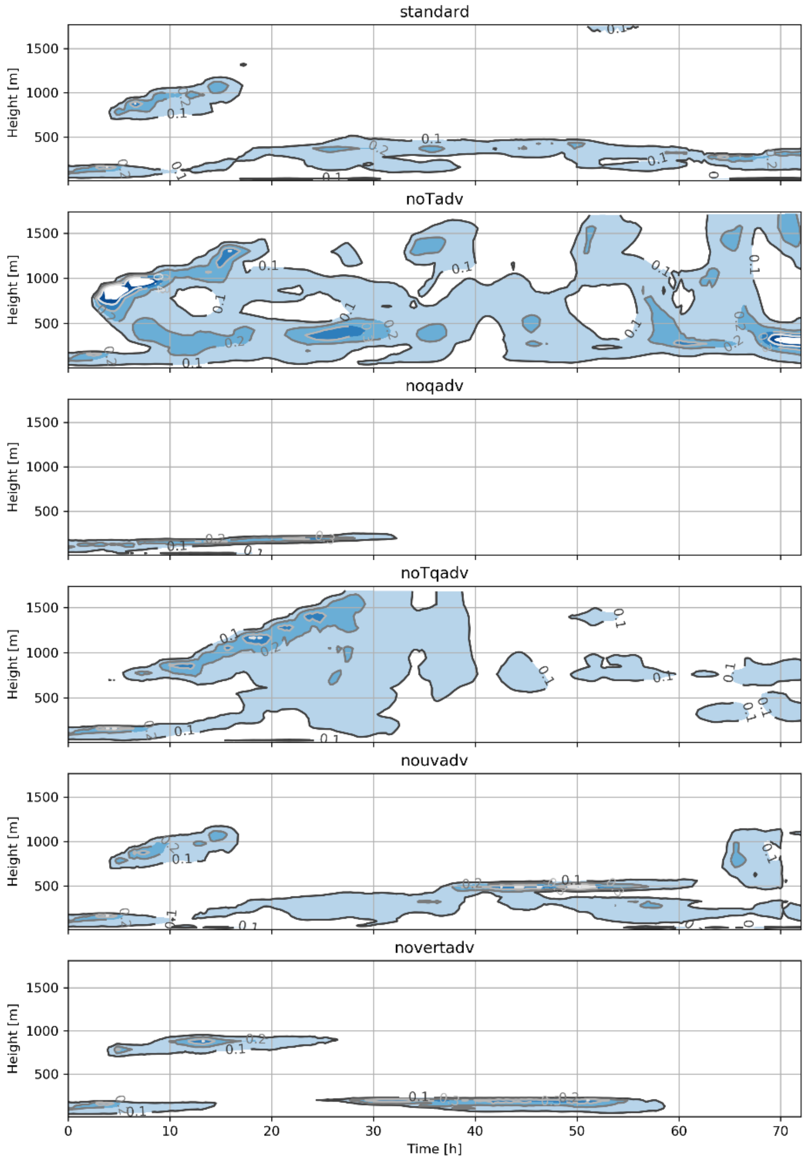

Figure 2. Time height cross section of liquid water content (g kg^{-1}) for the 72 hours of simulation for the ACSE case. The top panel shows the result using all forcing from ERA5, dt=900s and oifs40r1. The lower panels show results when various forcing fields are omitted, in turn temperature, moisture, temperature and moisture, momentum and vertical wind.

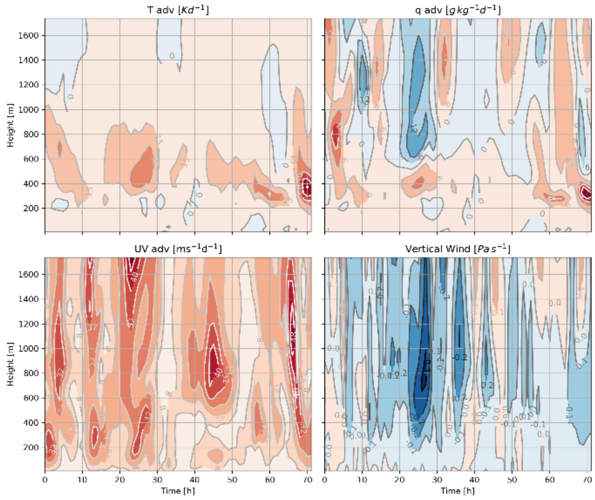

Figure 3. Time height cross sections of the forcing fields from ERA5 for the ACSE simulation, horizontal advection of temperature, humidity, momentum and vertical advection.

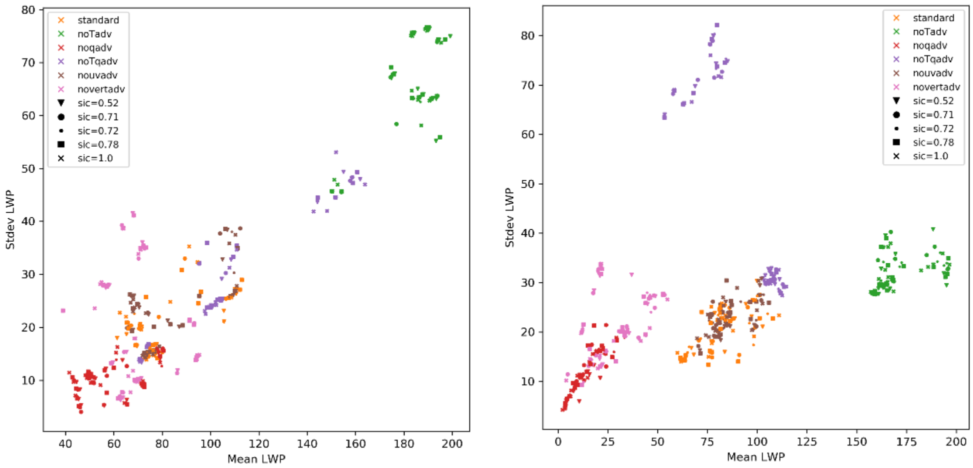

Figure 4. Simulated LWP and variability (standard deviation) over lowest 1000 m for first 12 h (left) and consecutive 60 h (right). The colours and marker indicate type of simulation, see legend.

As illustration of the kind of analysis that can be done, here we examine the variability found in the perturbation simulations and how that influences the liquid-water path (LWP), the low-level stability and the surface energy budget (SEB). After the spin-up period (right-hand side of Figure 4) the clearest difference in mean values of the LWP comes from turning off temperature advection, which also increases the variability of the LWP, likely related to the decreased stability (Figure 5). Temperature advection (Figures 3) is strongest just below 1 km but still mostly located above the model-determined boundary layer height (always below 300 m for all simulations, not shown). Moisture advection is mainly positive inside the boundary layer but varies in sign and magnitude above around 500 m (Figure 3). The case without moisture advection has an overall lower LWP, as well as less variability. The simulations with the highest variability at the later times (right panel Figure 4) are all with the updated IFS version (not shown).

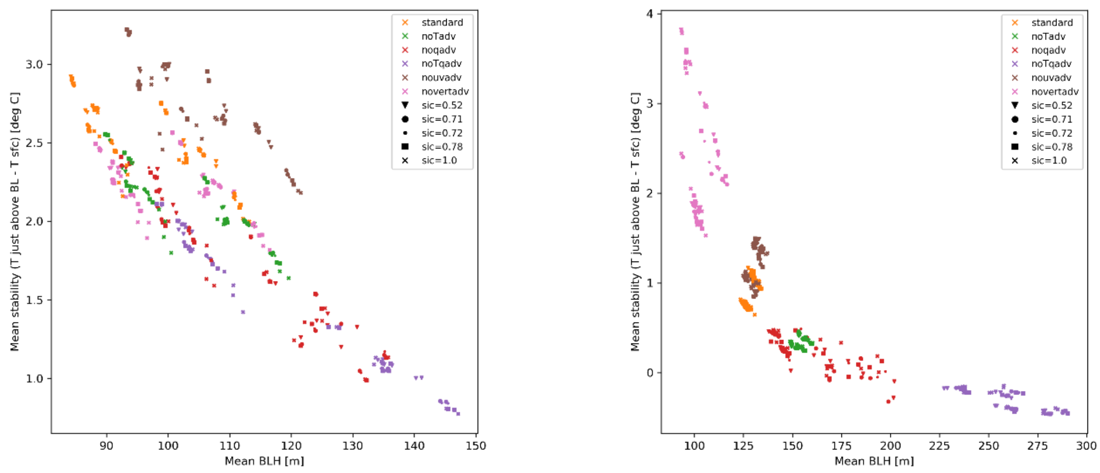

Figure 5. Simulated low-level stability (temperature difference between surface and just above the model-diagnosed boundary layer height) versus mean boundary layer height (m) for first 12 h (left) and consecutive 60 h (right). The colours and marker indicate type of simulation, see legend.

The lower/level stability (Figure 5), compared to the case with default forcing, is similarly reduced in setups with no moisture advection and becomes negative (unstable atmosphere) if both temperature and moisture advection are turned off at the later times.

If neither heat nor moisture is advected, the mean LWP is of similar order of magnitude as without heat advection. Turning off vertical advection tends to reduce the maximum LWP variability and leads to overall lower mean LWP that can be explained by reduced condensation due to less cooling due to removed vertical motions (Figure 3). Moreover, in the absence of vertical advection, the stability of the lower atmosphere is even further increased (Figure 5), over a depth of around 100 m to about 2-4 K (compared to the default value of around 1 K with a boundary-layer depth of around 125 m). For both the LWP and the stability measure, no easily discernible dependency of the results based on the sea-ice concentration are found (Figures 4 and 5).

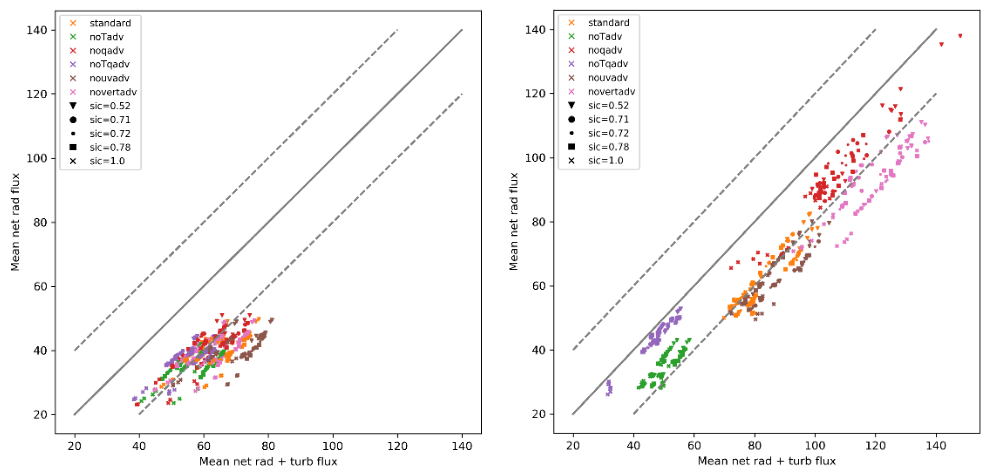

Figure 6. Net surface radiative fluxes against sum of net radiative and turbulent fluxes for first 12 h (left) and consecutive 60 h (right). Full line is diagonal (no turbulent fluxes) and dashed lines denote a contribution of the turbulent fluxes of ±20 Wm^{-2} to the SEB.

The variations in cloud field affect the shortwave and longwave fluxes substantially (not shown) which has consequences for the amount of energy that is available at the surface for melting sea-ice, which was observed to be rapid during these days. Figure 6 shows the net radiative fluxes plotted against the net energy available from the atmosphere, i.e. when the additional contribution from turbulent fluxes to the SEB is included. It is clear from the left panel that the available energy in the early part of the simulation is a net positive value of about 60 W m^{-2}, for which the turbulent fluxes contribute with about one third. In the default setup the turbulent fluxes are about 20 W m^{-2} and always positive, thus adding energy to the surface. The variability in the total available energy at the surface is large in the simulations, from about 30 to 130 W m^{-2}. The turbulent fluxes are of similar magnitude most of the time and even if temperature advection or horizontal momentum advection are turned off. The additional energy through turbulent fluxes is reduced in both setups without moisture
advection, getting closer to zero when both temperature and moisture advection are removed. Increased turbulence energy contribution to the SEB is found in the absence of vertical motion.

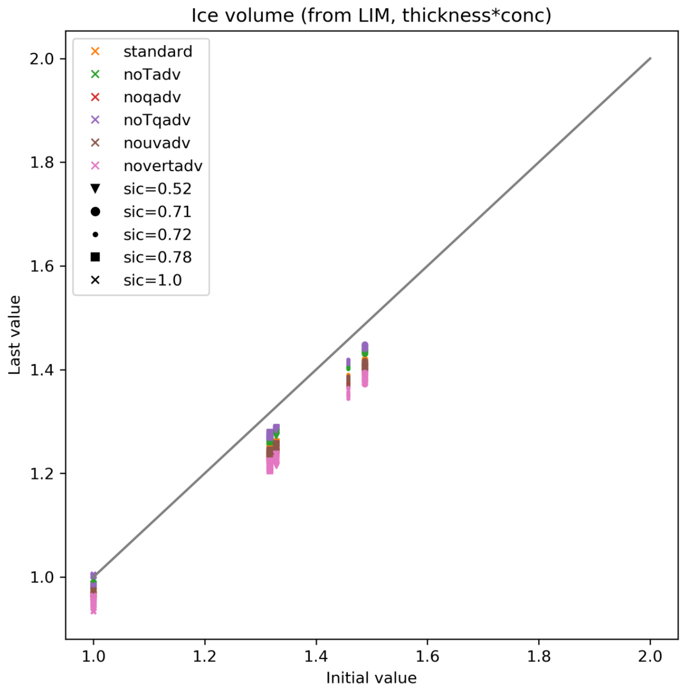

Figure 7. Evolution of ice volume during simulations.

As the energy is converging at the surface, the sea ice is evolving (Figure 7). Regardless of initial thickness or concentration, similar amount of sea ice is lost during the simulated 72 hours. Most sea ice is lost for the case of no vertical advection and turning off heat or heat and moisture advection reduces the sea-ice melt.

4. CONCLUSIONS AND OUTLOOK

Through further development, the coupled atmosphere-ocean single-column model (AOSCM) has become an even more versatile tool. There have been improvements in terms of technical capabilities as well as an update to a more recent model cycle. Procedures for setting up AOSCM cases have been developed and documented, especially how to initialise the ocean and sea ice. Improvements can be seen in the forcing as ERA5 is available. As it has the same vertical resolution and higher temporal resolution, there is less ambiguity due to interpolation issues of how the forcing act.

An inventory of observational datasets collected during expeditions with the Swedish icebreaker Oden is provided. Data from this supersite is process resolving and access to data from three expeditions have been made easier as they are rewritten following the format used within the WWRP Polar Prediction Project subproject YOPPsiteMIP. The upcoming MOSAiC data will follow the same format as well as other Arctic supersites that participate in the YOPPsiteMIP effort. These data files will be part of the WWRP PPP legacy available for model evaluation and development.

We present a case study on a summer event when warm air was advected over sea ice. A range of perturbations was made such that 480 simulations were produced. The analysis framework relates various processes to each other to sort out what is important for process representation. The analysis focuses on the cloud and its impact on the lower-level stability and the effect of the atmosphere on the surface energy budget, which will continue with information drawn from observations.

It is clear from the analysis that the advection plays a very important role for the evolution of the processes for this case. This poses a challenge, as it is not easy to assess if the applied forcing is adequate. This is most likely the case for most situations in the Arctic as it is a region with a net loss of energy to space and thus is highly influenced of advection of heat and moisture from lower latitudes. This means that developing parameterisations in idealised frameworks is problematic and thus a framework where the large-scale advection can be separated from the local processes are of great value. Thus, the AOSCM provides a more robust environment to test how sensitive the physical parameterisations are to the environmental conditions as they can easily be perturbed.

Presented here is just one of several test cases that are currently being modelled using the AOSCM. Most notably is the simulation of the entire period of the ASCOS ice camp for which the developments in terms of turbulent mixing in clouds (see D2.3) are tested. More investigations on parameters and procedures for the atmosphere-ocean coupling are also tested and will continue. How to use the AOSCM in Lagrangian mode, i.e. to follow the airmass, or sea-ice, will be further developed and investigated with the aid of MOSAiC observations. The results from these experiments will be published in peer-reviewed papers and reported on in D2.7.

5. REFERENCES

C3S, 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS). URL: https://cds.climate.copernicus.eu/cdsapp#!/home.

Dee, D.P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F., 2011:The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society,* **137,** 553–597.

Cohen,L., S. R. Hudson, V. P. Walden, R. Graham, M. A. Granskog, 2017: Meteorological conditions in a thinner Arctic sea ice regime from winter to summer during the Norwegian Young Sea Ice expedition (N-ICE2015). *J. Geophys. Res*, **122**. 7235-7259, DOI: 10.1002/2016JD026034

Hartung, K., Svensson, G., Struthers, H., Deppenmeier, A.-L., and Hazeleger, W., 2018: An EC-Earth coupled atmosphere-ocean single-column model (AOSCM) for studying coupled marine and polar processes, *Geosci. Model Dev.,* **11,** 4117–4137.

Hartung et al., 2019: Understanding processes during a warm-air advection event with a coupled 1D-model. *Manuscript in preparation.*

Sotiropoulou, G., Tjernström, M., Savre, J., Ekman, A. M. L., Hartung, K., and Sedlar, J., 2018: Large Eddy Simulation of warm-air advection and air-mass transformation in the summer Arctic. *Quart. J. Roy. Meteor. Soc.,* 144 (717), 2449-2462.

Sotiropoulou, G., M. Tjernström, J. Sedlar, P. Achtert, B. J. Brooks, I. M. Brooks, P. O. G. Persson, J. Prytherch, D. J. Salisbury, M. D. Shupe, P. E. Johnston, D. Wolfe, 2016: Atmospheric conditions during the Arctic Clouds in Summer Experiment (ACSE): Contrasting open-water and sea-ice surfaces during melt and freeze-up seasons. *J. Clim.*, **29**, 8721-8744, doi: 10.1175/JCLI-D-16-0211.1.

Tjernström, M., P. Achtert, M. D. Shupe, J. Prytherch, J. Sedlar, B. J. Brooks, I. M. Brooks, P. O. G. Persson, G. Sotiropoulou, D. J. Salisbury, 2019: Arctic summer air-mass transformation, surface inversions and the surface energy budget, Journal of Climate, 32, 769-789, https://doi.org/10.1175/JCLI-D-18-0216.1

Tjernström, M., M. D. Shupe, I. M. Brooks, P. O. G. Persson, J. Prytherch, D. Salisbury, J. Sedlar, P. Achtert, B. J. Brooks, P. E. Johnston, G. Sotiropoulou and D. Wolfe, 2015: Warm-air advection, air mass transformation and fog causes rapid ice melt, *Geophys Res. Let.*, **42**, 5594–5602, doi:10.1002/2015GL064373.

Tjernström, M., and 43 co-authors, 2014: The Arctic Summer Cloud Ocean Study (ASCOS): Overview and experimental design, *Atmos. Chem. Phys.*, **14**, 2823–2869, doi:10.5194/acp-14-2823-2014

Tjernström, M., C. E. Birch, I. M. Brooks, M. D. Shupe, P. O. G. Persson, J. Sedlar, T. Mauritsen, C. Leck, J. Paatero, M. Szczodrak and C. R. Wheeler, 2012: Meteorological conditions in the Central Arctic summer during the arctic summer cloud ocean study (ASCOS), *Atmos. Chem. Phys.*, **12**, 6863–6889, doi:10.5194/acp-12-6863-2012.

Uttal, T., and Coauthors, 2002: Surface Heat Budget of the Arctic Ocean. *Bull. Amer. Meteorol. Soc.*, **83**, 255-275.

6. ACRONYMS

AOSCM atmosphere-ocean single-column model
MOSAiC Multidisciplinary Drifting Observatory for the Study of Arctic Climate
PPP Polar Prediction Project
SCM single-column model
WWRP World Weather Research Program
YOPPsiteMIP Year of Polar Prediction supersite Model Intercomparison Project

7. ANNEXES

Instructions to set up a case with the AOSCM

The attachment is the latest draft of a document with instructions on how to set up a case with the AOSCM. The document is developed to help new users of the AOSCM to properly define all aspects of the model and the forcing. It will be published on the AOSCM pages within the EC-Earth development portal.

Setting up AOSCM cases

Jareth Holt, MISU, October 2019

This document describes how to build cases for the EC-Earth Atmosphere-Ocean Single Column Model (AOSCM, Hartung et al., 2018). The AOSCM couples the OpenIFS atmospheric (based on IFS cycles 40r1 and 43r3 available) with the NEMO3.6-LIM3 ocean and sea ice model. It is designed for studying the dynamics of the coupled system, with a focus on the air-sea ice and sea ice-ocean interfaces. It is also designed for exploring the effects of changes in e.g. sea ice representation or boundary layer turbulence parameterizations.

Because it is a single-column model, the AOSCM requires a set of initial, boundary, and forcing conditions for each simulation. For example, the atmospheric temperature tendency is (Hartung et al., 2018, eqn. 3):

$$\frac{\partial T}{\partial t} = -\dot{\eta}\frac{\partial T}{\partial \eta} + \frac{R_m T}{c_p p}\omega + F_T + P_T + \frac{T_{ref} - T}{\tau}.$$

Here, η is the vertical coordinate in OpenIFS (a hybrid-sigma pressure coordinate) with vertical velocity $\dot{\eta}$, whereas $\omega = D_t p$ is the pressure velocity. Both of these velocities are external to the model and must be specified at every timestep if vertical advection is included. In addition, F_T represents horizontal temperature advection, and T_{ref} is a reference temperature profile to relax towards (often to help stabilize the single-column behavior). If included, both of these quantities are external as well. Only P_T, representing the parameterization of purely vertical turbulent and radiative processes, is completely internal to the AOSCM.

This document is organized into two parts. The first part describes the quantities incorporated in each model component: input variables, forcing, prognostic vs. diagnostic variables, etc. This part also describes sources for some of these variables where appropriate. The second part describes our current methods for building up case studies. These methods seek to make atmosphere-only, ocean-only, and coupled model runs comparable to each other and to observations. This section ends with a discussion on some of the key issues we face in standardizing these input files.

Summary

Our current approach to building cases for the EC-Earth Atmosphere-Ocean Single Column Model is as follows.

1. **Choose a location and time period.** This is typically based on a measurement campaign. A compromise between the exact campaign location and the points for which we can find data is needed. The namelist for NEMO must be updated to reflect the actual latitude, longitude, and bathymetry.

2. **Get atmospheric reanalysis data.** Software is available to extract all of the ASCM variables from the ERA5 reanalysis. We also use certain surface fluxes from the 4D-Var product to create ocean forcing and coupler initial condition data.

3. **Get oceanic reanalysis data.** Both the ORAS4 and ORAS5 reanalyses can provide initial conditions for the ocean model. The ORAS5 product also provides some surface fluxes and sea ice data.

4. **Get a LIM3 restart file for sea ice.** We currently extract sea ice initial conditions from the restart files generated by full EC-Earth simulations. We are exploring other options for adapting ice data as initial conditions.

5. **Create input files for each component.** This step is crucial to building a consistent case study. This step involves:

   - Adjusting the sea ice concentration, sea surface temperature, and possibly sea ice albedo in the atmospheric input files to match the ocean/ice data;
   - Using the ERA5 data and 4D-Var fluxes to create ocean forcing files; and
   - Creating initial conditions for the coupler variables by combining 4D-Var downward fluxes, atmospheric near-surface conditions, sea ice conditions (concentration, temperature, and albedo), and sea surface temperature with bulk formulae to calculate the required net fluxes.

6. **Incorporate observational data.** If available, observational data can be blended with the preceding to create hybrid input files. For the atmospheric component, this requires translating from height or pressure to hybrid-sigma coordinates. For the oceanic component, the translation of temperature and pressure depends on the equation of state chosen.

1 Model component requirements

Each of the model components has a set of control parameters as well as initial, boundary, and forcing conditions. We will first describe each of these components separately.

1.1 Atmosphere input data

1.1.1 Input file `scm_in.nc`

The initial, boundary, and forcing conditions are all given in a single input file, `scm_in.nc`. The name of this input file is fixed; the model will always read whatever is in the run directory as `scm_in.nc`, though these files can be managed by run scripts. All of the variables are expected to be given over a time period that completely covers the simulation time (including after the last timestep). This is true both for prognostic quantities, which are only used as initial conditions, as well as climatological/surface parameters that should be constant. Some quantities are optional and will be given default values.

The vertical structure of OpenIFS is a hybrid-sigma grid. The pressure at the top edge of grid cell k is

$$p_k^e(t) = A_k + B_k \cdot p^s(t)$$

where p^s is the surface pressure and A, B are fixed arrays. The grid indexing starts at the top of the atmosphere, with $A_0 = B_0 = 0$ and ends at the surface with $A_N = 0, B_N = 1$. Here, N is the number of grid cells, and there are configurations with $N = 60$, 91, and 137. The pressure at the center of grid cell k is

$$p_k^c(t) = \frac{1}{2}(p_k^e(t) + p_{k+1}^e(t)), \qquad k = 0, 1, \ldots, N-1.$$

In the IFS documentation, the collection of p^c is referred to as the 'full' grid, on which the majority of variables are specified. The edge points p^e is the 'half' grid, reserved primarily for vertical velocities and fluxes.

The quantities in `scm_in.nc` can be divided into a number of groups as follows.

- **Grid description**: These quantities describe the vertical and temporal grid of the input data.
  - `time, second, date`: Time axis for the input file. `time` and `second` are identical and typically the number of seconds since the start of the simulation. However, `time` is used by `su1c.F90` whereas `second` is used by `surip1c.F90`, so both are required. `date` is the date as a YYYYMMDD integer.
  - `nlev, nlevp1`: Number of vertical levels for the model full and half grid, respectively. `nlev` should be 60, 91, or 137, and `nlevp1` = `nlev` + 1.
  - `nlevs`: Number of soil/sea ice vertical levels. While read from the input file, the model expects this number to be 4.
- **Necessary climatological parameters**: These parameters characterize the surface and depend only on time.
  - `lat, lon`: Latitude and longitude in degrees North and degrees East.
  - `ps`: Surface pressure in Pa. Since an SCM does not model horizontal mass convergence, changes in surface pressure are a model input.
  - `lsm`: Land-sea mask, with 1 denoting complete land and 0 denoting complete ocean/sea ice.
  - `high_veg_type, low_veg_type`: Index for the type of vegetation, from 1-20. See ECMWF (2017b, chpt. 8) for a list of types. High and low vegetation types can be specified separately.
  - `high_veg_cover, low_veg_cover`: Fraction of the grid cell covered by the vegetation types, from 0-1.
  - `mom_rough, heat_rough`: Roughness lengths for momentum and heat in m. These apply only to the land surface; the roughness lengths over ocean and sea ice are calculated internally.
  - `albedo`: Average broadband albedo of the surface, from 0-1. Note however that ocean, sea ice, and snow albedos are calculated internally.
  - `orog, sdor, isor, anor, slor`: Parameters describing the orography. `orog` is the mean height of the surface in m above sea level. The other quantities are nondimensional and relate to the form drag. They specify, in order: the sub-grid scale standard deviation of orography; the anisotropy of orographic features, with 0 for circular and 1 for ridges; the angle (in radians from eastward) of the features; and slope of the features, with 0 for flat surfaces and 1 for 45° slopes. **NB**: The standard deviation of orography is listed as dimensionless, but should actually be in m. See the entry for `sdfor` below.
  - `sea_ice_frac`: Fraction of the grid box covered by sea ice, from 0-1.
  - `open_sst`: Surface temperature of open (not sea ice-covered) seawater in K. This quantity can be either a pure forcing, or corrections may be applied for different layer effects; see ECMWF (2017b, chpt. 8.10) for details.

- **Optional climatological parameters**: These quantities do not have to be given but allow for more details of the surface to be specified.
  - soty: Soil type, as an index between 1 and 7. See ECMWF (2017b, chpt. 11.9) for the list of types; defaults to 2 ('medium' texture).
  - laih, lail: Leaf area index in $m^2 m^{-2}$ for high and low vegetation, respectively. Defaults to 0.
  - aluvp, aluvd, alnip, alnid: Specific albedos to use for parallel (p) and diffuse (d) radiation in the UV (uv) and near-infrared (ni) bands. If missing, the general albedo quantity is used.
  - sfc_ice_albedo: Surface sea ice albedo from 0-1. The default value is -1.0; this quantity should only be missing if there is no sea ice.
  - sdfor: Standard deviation of filtered subgrid-scale orography in m. This quantity has been computed by ECMWF to include scales between 3 and 22 km. If missing, sdor is used instead; thus, sdor should have the same units of m.
- **Necessary prognostic variables**: These variables must be in the input file as initial conditions. They are given on 'full' model levels. They will also serve as reference profiles if relaxation is enabled for the simulation (only applies to wind and temperature).
  - u, v: Zonal and meridional velocity in $m\,s^{-1}$.
  - t: Temperature in K.
  - q: Water vapor mixing ratio in $kg\,kg^{-1}$.
  - ql: Cloud liquid water mixing ratio in $kg\,kg^{-1}$.
  - qi: Cloud ice water mixing ratio in $kg\,kg^{-1}$.
  - cloud_fraction: Cloud area fraction, from 0-1.
- **Optional prognostic variables**: Two prognostic variables are allowed to be missing from the file: the mixing ratios of rain qr and snow qsn in $kg\,kg^{-1}$. If missing, they are initialized to zero, i.e. there is no precipitation at the start of the simulation.
- **Multi-level ground variables**: These are properties of the ground or sea ice that have 4 layers. The layer depths are given in ECMWF (2017b), chapter 8.5 for soil and 8.9 for sea ice.
  - t_soil: Temperature of the soil in K.
  - q_soil: Soil moisture content in m of water equivalent.
  - t_sea_ice: Temperature of sea ice in K.
- **Necessary surface prognostic variables**: These single-valued surface quantities are calculated prognostically by the model but require initial conditions.
  - t_skin: Skin temperature of the surface in K.
  - q_skin: Water content of the skin reservoir, in m of water equivalent.
  - snow: Depth of snow on the surface in m.

- **Optional surface prognostic variables**: The snow surface variables are allowed to be missing: `t_snow` (K), `albedo_snow` (0-1) and `density_snow` (in $\mathrm{kg\,m^{-3}}$). More accurately, if `t_snow` < 100.0 (including if it is not filled after initialization), then `t_snow` takes the top value of `t_soil` and the albedo and density are set to default values of 0.88 and 100.0, respectively.

- **Surface fluxes**: Externally-supplied surface sensible and latent heat fluxes (`sfc_sens_flx`, `sfc_lat_flx`), each in $\mathrm{W\,m^{-2}}$. The SCM has the capability to use flux-specified boundary conditions rather than computing surface turbulent exchange, but this is not the default or recommended behavior. Regardless, these variables must be specified.

- **Forcing variables**: These quantities force the tendency equations, and hence have to be specified at every timestep. All of them are on the 'full' model grid except for `etadotdpdeta`, which is on the 'half' grid. **NB**: These variables must always be present in `scm_in.nc`, since they are first read in by the model and then reset to 0 if a given forcing is turned off for that simulation.

  - `ug, vg`: Zonal and meridional geostrophic velocities in $\mathrm{m\,s^{-1}}$. These quantities are used to represent the pressure gradient forcing.
  - `omega`: Vertical pressure velocity in $\mathrm{Pa\,s^{-1}}$. This is only used in calculating adiabatic heating in the temperature tendency, not vertical advection.
  - `uadv, vadv, tadv, qadv`: Advective tendencies for zonal and meridional velocities ($\mathrm{m\,s^{-2}}$), temperature ($\mathrm{K\,s^{-1}}$), and water vapor ($\mathrm{kg\,kg^{-1}\,s^{-1}}$). These quantities reflect the horizontal winds and gradients.
  - `etadotdpdeta`: Vertical coordinate velocity in $\mathrm{Pa\,s^{-1}}$. This is the velocity used in calculating vertical advection, and is related primarily to the change in surface pressure.

The ERA reanalyses, ERA-Interim and ERA5, use IFS as the atmospheric model and thus have the variables and format required for the SCM. All of the above quantities are listed in the ERA5 catalogue at https://apps.ecmwf.int/data-catalogues/era5 with two caveats. The first caveat is minor: the recorded values of `etadotdpdeta` ("Eta-coordinate vertical velocity") may be missing a term. This pressure velocity is given by (ECMWF, 2017a, eqn. 2.9)

$$\dot{\eta}\frac{\partial p}{\partial \eta} = -\frac{\partial p}{\partial t} - \int_0^{\eta} \nabla \cdot \left(\mathbf{v}_H \frac{\partial p}{\partial \eta}\right) d\eta = \dot{p}_1 + \dot{p}_2.$$

This quantity is necessarily 0 at the top ($\eta = 0$) and bottom ($\eta = 1$) of the atmosphere. However, in some of the available datasets, only $\dot{p}_2$ was saved. Fortunately, the first term is easily calculated, since this velocity is defined on the model 'half' grid and pressure is only related to surface pressure:

$$\dot{p}^e_{1,k} = \frac{\partial}{\partial t}\left(A_k + B_k p^s(t)\right) = B_k \frac{dp^s}{dt} = B_k \dot{p}^e_{1,N}.$$

Since $\dot{p}(\eta = 1) = \dot{p}^e_N = 0$, we have $\dot{p}^e_{1,N} = -\dot{p}^e_{2,N}$. Therefore we can get the corrected velocity directly from the given one:

$$\dot{p}_k = \dot{p}'_k - B_k \dot{p}'_N$$

where $\dot{p}$ is now the corrected velocity and $\dot{p}'$ is the uncorrected one.

The second caveat to using the public ERA5 data for `scm_in.nc` is that quantities involving gradients – the geostrophic velocities and advective tendencies – are unavailable. These could be calculated from the gridded data using finite differences. However, this would be inconsistent with the model's spherical harmonics formulation of the dynamics. This can also introduce numerical noise into the simulation, to the detriment of the stability and accuracy of the SCM.

For this reason, we currently set up cases using the procedure outlined in `https://confluence.ecmwf.int/display/OIFS/How+to+extract+Single+Column+Model+forcing+data+from+ECMWF+analysis`. This requires access to the `ecgate` system and the `ifs` group due to the use of proprietary software for spherical harmonics transformations. Using the `getscmdata` script described at that page, all variables for running the OpenIFS SCM can be downloaded for a single date and model grid point into a GRIB file, which can then be converted to netCDF using the `eccodes` software suite. Work is currently underway to expand these scripts to 1) avoid proprietary software and remove the need for `ifs` group access, and 2) get data for collections of dates and multiple grid points simultaneously, making the retrieval more efficient and complying with the MARS retrieval guidelines.

There is one modification to the `getscmdata` script that we recommend. In the standard IFS output, the `albedo` variable is the climatological mean which is only applied over land surfaces. The snow, ocean, and sea ice albedos are calculated internally. Instead, the 'forecast albedo' (`fal` variable) provides the albedo calculated by the model. However, this variable is the average albedo for the grid point, including both sea ice and the solar zenith angle-dependent ocean albedo. Deriving the appropriate values for `sfc_ice_albedo` from this needs to be done with care.

Even if the forecast albedo is extracted, IFS (up to at least cycle 43r3) calculates sea ice albedo internally based only on the date (with bare, melting ice in summer and dry, snow-covered ice in winter) and not on the state of the ice. Instead of using the standard reanalysis, one can also use the 4D-Var reanalysis, which includes some adjustment of surface parameters in its assimilation. Using this albedo, which was tuned so that the atmospheric fields would better match observations, may be more reflective of the actual state of the ice.

1.2 Ocean input data

The NEMO single-column model is actually just a special case of the NEMO model with its domain restricted to a 3×3 grid. This grid, rather than a 'true' single column, is used so that the NEMO core can run as-is, calculating all the necessary quantities for horizontal gradients. Since the values at all points on this grid should be the same, the resulting horizontal gradients (and hence horizontal transport and diffusion) will be zero, effectively giving single-column results. Keep in mind that certain vector components (for velocity and stress) are defined on a grid offset from other variables (temperature, salinity). These quantities need to have their own 3×3 grid coordinates as well.

Unlike OpenIFS, the input data for NEMO is divided into initial conditions and forcing. In both cases, there are few constraints on the file format. In particular, the file and variable names used are defined in the NEMO namelist. Provided below are our suggestions for these names.

1.2.1 Initial conditions file

The NEMO initial conditions file must contain the following variables in order for the model to run:

- `time_counter, time_counter_bnds`: The time dimension for the data. Typically, `time_counter` is a size-1 array (the middle of the timestep) and `time_counter_bnds` is size-2 (the beginning and end of the timestep). The simulation should start at the value of `time_counter`.
- `nav_lat[_i], nav_lon[_i]`: The latitude and longitude of the grid. These arrays define the physical location of the 3×3 grid. The central point is the location of the 'single column'. As mentioned above, variables defined on a grid offset from temperature and salinity require their own grid of `(nav_lat, nav_lon)`, typically specified with a numerical suffix (`_2,_3`).
- `depth[t,u,v]`: The depths for the model grid. While all of these variables share the same vertical grid here, they would not in the full NEMO model and thus each requires its own depth coordinate.

The physical variables used by NEMO depend somewhat on the equation of state (EOS) chosen. This is controlled by the namelist variable `nn_eos` in `nameos`. The choices are:

- `nn_eos=-1`, polyTEOS10-bsq. This uses a version of the most recent equation of state, in a polynomial form optimized for Boussinesq ocean models such as NEMO.
- `nn_eos=0`, polyEOS80-bsq. This uses an older but very common EOS, also in polynomial, Boussinesq form.
- `nn_eos=1`, S-EOS. This is a simplified EOS, but still allows for some effects of cabbeling and thermobaricity.

Knowing these options for the EOS, the initial condition variables in NEMO are:

- `votemper`: Temperature variable in °C, either conservative temperature for polyTEOS10 or potential temperature for the others.
- `vosaline`: Salinity variable. For polyTEOS10, this is absolute salinity in $\mathrm{g\,kg^{-1}}$; for the others, it is practical salinity in psu.
- `vozocrtx, vomecrty`: Zonal and meridional velocities in $\mathrm{m\,s^{-1}}$. More specifically, these are the velocities in the x- and y-directions on the model grid. Therefore the velocities have to be adjusted if the SCM grid is not oriented along latitudes and longitudes. (Compare to the ORAS variables `vozocrte, vomecrtn`.)

If the velocities are not included among the initial conditions, the simulation starts from rest. If the temperature and salinity are missing, the default state is one of constant 35.5 psu salinity and a temperature profile typical of the tropical ocean.

1.2.2 Forcing

Forcing data for NEMO is divided amongst a few sources. The primary forcing files consist of the atmospheric variables required to calculate bulk fluxes. (Currently, the AOSCM has only been tested with the CORE bulk formulation.) These files contain the following:

- wndwe, wndsn: Near-surface winds in the zonal and meridional directions in $\mathrm{m\,s^{-1}}$. The wind height is specified by the namelist variable rn_zu in namsbc_core.
- qsr, qlw: The downward shortwave and longwave radiative fluxes in $\mathrm{W\,m^{-2}}$.
- tair, humi: The near-surface air temperature in K and specific humidity in $\mathrm{kg\,kg^{-1}}$. The height of these variables is specified by rn_zqt.
- prec, snow: The rates of total precipitation (rain and snow) and of snow only in $\mathrm{kg\,m^{-2}\,s^{-1}}$.

Another forcing to provide to NEMO is the chlorophyll content of the upper ocean. The treatment of the chlorophyll input is controlled by namtra_qsr. Time-varying surface chlorophyll concentrations, in $\mathrm{(g\,Chl)\,L^{-1}}$, can be provided and will be used to calculate the penetrative solar radiation through the upper ocean. Note that this data will only be used if nn_chdta is 1 or 2; if it is 0, a constant value of 0.05 $\mathrm{(g\,Chl)\,L^{-1}}$ is used.

The last forcing files used by NEMO are maps of the magnitudes of the and M2 (principal lunar semidiurnal) and K1 (lunar diurnal) tidal components. Tidal data is important for the generation of internal waves. Since the propagation and breaking of these waves affects mixing in the abyssal ocean, the tidal components are necessary for modeling deep ocean processes on long timescales. The files distributed with the AOSCM (M2rowdrg.nc, K1rowdrg.nc) provide this information for the global ocean, and are unlikely to need altering for any AOSCM case study.

1.3 Sea ice input data

The LIM sea ice model uses the framework of an ice thickness distribution. The sea ice is divided into 5 thickness categories, with bounds of .6, 1.3, 2.2, and 3.8 m. (The first category is ice thinner than 0.6 m; the last category is ice thicker than 3.8 m.) Each category covers a given area, and the ice state is defined by the total volume, salt content, and enthalpy of the ice and any overlying snow. From these extensive variables, average intensive variables (mean thickness, temperature, and salinity) can be calculated.

The ability to specify sea ice initial conditions was recently added to the AOSCM. The format of the initial conditions file is the same as the restart files output by LIM. However, these files include many variables related to stress and advection which are irrelevant in the single-column formulation. The key variables needed are:

- a_i_htc[icat]: The concentration of sea ice in thickness category icat (1-5). This is the fraction of the grid cell covered by ice in this category, and has units of $\mathrm{m^2\,ice/m^2\,ocean}$.
- v_i_htc[icat], v_s_htc[icat]: The volume of ice and snow, respectively, in each category. The units are $\mathrm{m^3\,ice/m^2\,ocean}$. Thus, the total volume of ice in the grid cell is the area of the grid cell times the sum of the v_i. The average thickness of ice within a category is v_i_htc[icat] / a_i_htc[icat] and should fall within the thickness bounds for that category.
- smv_i_htc[icat]: Total salt content of ice in each category, in units of $\mathrm{(g\,salt/kg\,ice)/(m^3\,ice/m^2\,ocean)}$. The bulk salinity of each category is smv / a, and these salinity units are roughly equivalent to PSU.
- tempt_il[ilay]_htc[icat]: The total enthalpy of the sea ice in thickness category icat and layer ilay. The ice is divided internally into a number of layers (typically 2 or 4) for calculating diffusive

heat fluxes. The units of this enthalpy are $J/(m^2$ ocean). Dividing `tempt / a` gives the specific enthalpy of melting of the ice, in units of $J/(m^3$ ice).

- `tempt_sl[ilay]_htc[icat]`: The total enthalpy of snow overlying thickness category `icat`. Currently, LIM only has one snow layer, but this variable naming convention has been adopted anticipating an upgrade to a multi-level snow scheme. The units are the same as for `tempt_il`.

- `t_su_htc[icat]`: Surface (skin) temperature of the ice or snow in each thickness category in K. This variable may not be strictly necessary as an initial condition (this remains to be tested) but knowing whether the surface is initially melting or not may be important for the atmospheric fluxes at the first timestep.

Other variables present in a LIM restart file are related to velocities, stresses, advective tendencies, or global budgets. We are in the process of testing the importance of each to the behavior of the AOSCM.

1.4 Coupler requirements

The OASIS3-MCT coupler is used by EC-Earth to manage how variables are transferred between the atmosphere and ocean/sea ice models. This general coupler is able to handle models with different grids and temporal resolutions, including serial or concurrent coupling.

The following files are required by OASIS in order to run a coupled simulation:

- `data/oasis/cf_name_table.txt`: A list of the indices, CF-convention names, and units for all the variables that are available for coupling.

- `data/oasis/rmp_O1CD_to_ASCM.nc, rmp_ASCM_to_O1CD.nc`: NetCDF files describing how values from the ocean grid (`O1CD`) relate to values on the atmospheric grid (`ASCM`) and *vice versa*.

- `data/oasis/rstas.nc, rstos.nc`: Restart files containing initial conditions for certain variables, described in detail below.

- `namelists/namcouple`: Namelist file for parameters controlling the behavior of OASIS.

For the namelist file, we provide the script `namelists/oasis.sh` to generate `namcouple` dynamically, based on variables set either in the `config-run.xml` configuration file or the coupled model run script. Because this is a single-column model, the `rmp` grid files are very simple and unlikely to change from case to case. Similarly, the `cf_name_table` should always be the same, though it may require updating if attempting to use newer versions of the model components. Therefore, out of these OASIS-specific files, only `rstas` and `rstos` need to be created for each case.

The restart file `rstas.nc` describes the forcing *of* the atmosphere *on* the ocean. That is, its variables are all fluxes, defined positively for downward fluxes (from the atmosphere). Its variables can be grouped as follows:

- `A_TauX_oce, A_TauY_oce, A_TauX_ice, A_TauY_ice`: Zonal and meridional components of the wind stress on the open ocean and sea ice.

- `A_Precip_liquid, A_Precip_solid`: Instantaneous liquid (rain) and solid (snow) precipitation rates.

- `A_Evap_total`, `A_Evap_ice`: Instantaneous evaporation rates, both total and ice-only (sublimation).
- `A_Qs_oce`, `A_Qs_ice`, `A_Qs_mix`, `A_Qns_oce`, `A_Qns_ice`, `A_Qns_mix`: The solar (`Qs`) and non-solar (`Qns`) fluxes, specified over ocean only, sea ice only, and averaged over both (`mix`).
- `A_dQns_dT`: The derivative of the non-solar flux over sea ice with respect to sea ice surface/skin temperature. This quantity plays a key role in solving the implicit equations determining the surface boundary condition and surface fluxes.
- `A_Runoff`, `A_Calving`: The fluxes of water and ice into the ocean from river runoff and the calving of icebergs from land-fast ice.

The restart file `rstos.nc` describes ice/ocean state variables that impact the atmosphere. These are:

- `O_SSTSST`, `O_TepIce`: Surface/skin temperature of open seawater and sea ice in K.
- `O_AlbIce`: Albedo of the sea ice including any overlying snow (0-1).
- `OIceFrc`: Sea ice area fraction/concentration (0-1).
- `OIceTck`, `OSnwTck`: Mean thickness of sea ice and overlying snow in m.

We currently have two strategies for generating appropriate OASIS restart files. The first strategy is to run a very short (single-timestep) atmosphere-only simulation. This will generate surface fluxes consistent with the initial conditions which are then used to fill in `rstas.nc`. We provide with the AOSCM the Python script `compile-rst.py` for creating the restart files from atmosphere-only runs.

The second strategy is to get these fluxes from the ERA5 4D-Var assimilation. The surface fields in this dataset have also been adjusted by the assimilation scheme and are publicly available. In contrast, most surface fluxes are not stored variables of the standard reanalysis. One potential issue with using this data is the assimilation window. Assimilations are started at 0900 and 2100 UTC each day, with 12 hourly forecast steps in between. The hours of 0900 and 2100 can thus be represented by two different forecast steps, and fields will not generally be smooth when changing assimilation windows. The second issue is whether the 4D-Var surface fields are consistent enough with the standard reanalysis to use both datasets together. Better consistency might be achieved by also using 4D-Var for the other atmospheric fields, but these fields are then also subject to the above assimilation window issue. Tests are ongoing to assess 1) how similar the restart files generated in these ways are, and 2) how similar the standard reanalysis and 4D-Var assimilation are in our case studies.

2 Building cases

The previous section described the inputs to each of the model components (atmosphere, ocean, sea ice, and the coupler). In this section we will discuss our strategies for generating these inputs in a way that promotes consistency across atmosphere, ocean, and coupled simulations.

The first step in building a case study is to choose the location and time period. All of our coupled cases so far have been Eulerian (latitude and longitude held constant) but we hope to expand to sea ice-, air mass-, and water mass-following Lagrangian simulations. One air-mass following experiment was for example done for Pithan et al. (2016). Our cases have been based on specific measurement campaigns, so it is natural to use the same time frame as the campaigns. We most often simulate 5-10 days at a time, with 1 month of campaign data to compare to. At the Pacific measurement buoy PAPA the model was reasonable stable over 1-month long simulations but stability tests have yet to be done over the sea-ice and with different types of forcing to see whether or not the AOSCM is stable over longer periods.

The second step is to extract the atmospheric data. The basic data extraction is described in section 1.1. (Another reason to use an IFS grid-point as the location is that the tools for extracting single grid-point data already exist.) At this stage, there is a choice to be made for atmosphere-only runs as to whether to use the default climatological IFS sea ice albedo or provide other albedo values. (In coupled simulations the albedo is calculated as part of LIM and then passed to the atmosphere.) Similarly, there is a choice as to whether initial conditions for the coupler variables will be derived from a short atmosphere-only simulation or extracted from the 4D-Var assimilation. If the latter, 4D-Var surface fluxes at the initial time step have to be extracted as well. This setup choice arises because the default setup of the coupler is lagged: the atmosphere uses data from ocean and sea-ice from the previous time steps.

The third step is to extract ocean data. So far, we have used the Ocean ReAnalysis System (ORAS) products. ORAS4 provides temperature, salinity, and horizontal velocity variables on a $1^\circ \times 1^\circ$ grid. ORAS5 provides many variables on both a native grid (ORCA025) and a $1^\circ \times 1^\circ$ grid. In addition to the above, ORAS5 provides surface values (wind stress, surface temperature and salinity, water and heat flux) and a few sea ice parameters (concentration, mean thickness, and velocity). In the model setup, the primary choice is whether the initial conditions and reference profiles for relaxation will include only the temperature and salinity or the horizontal velocities as well. A second choice is whether the equation of state (EOS) in the AOSCM run will use potential or conservative temperature, and the values in the initial conditions will have to be pre-processed accordingly. A third choice with ORAS5 is whether to try to incorporate the sea ice or surface variables in the case. For example, the sea surface temperature, sea ice concentration, and sea ice thickness (through its effect on albedo) could be used to force the atmosphere-only simulations. The surface fluxes could also be initial conditions for some of the OASIS restart variables.

The fourth step is to extract sea ice data. Since the initial conditions for LIM require data across thickness categories and ice layers, we have not yet used observational data to initialize the model. Instead, we have used restart files from historical, global EC-Earth runs. We are investigating whether and how these restart files can be combined with other data sources, as well as how observations or other sea ice data can be used to initialize LIM.

The last step in building up a case study is determining the coupler initial conditions/restart files. The coupler requires initial conditions for certain surface values, especially fluxes, in order to force the models until such fluxes can be calculated internally. Certain quantities (e.g. sea surface temperature and sea ice albedo) may come directly from the other data sources. Other quantities (e.g. sensible heat flux, evaporation rate) can be estimated from other variables via bulk formulae. However, precipitation and radiative fluxes are particularly difficult to approximate without use of a model. We have so far handled this in two ways. The first way is to run a short (single timestep) atmosphere-only simulation and use the fluxes output from that. The second is to use the ERA5 4D-Var product, in which surface data has been assimilated and surface fluxes are among its outputs. We are in the process of testing the impact of each approach on simulations. Either approach gives all the fluxes needed as initial conditions, but these can be combined with the values mentioned above to hopefully create a more consistent starting point.

When trying to set up consistent case studies, an eminent obstacle is that these disparate data sources can disagree on key variables at their interfaces. A few of the issues we have come across so far are:

- Sea ice concentration. This is a key variable in both the atmosphere and ocean/sea ice models. Since the sea ice concentration is a purely external input to the atmospheric model, one solution is to keep the downward fluxes from the atmosphere the same (LW and SW radiation, precipitation) while scaling the upward fluxes to reflect the ocean/sea ice configuration from the ocean model. This reasoning can also be extended to other sea ice properties, such as albedo, snow amount, or skin temperature. This is not an issue for most implementations of the full EC-Earth model since runs are almost always initialized using a restart file after many years of spin-up time, and sea ice has a chance to equilibriate after an ice-free initial state.

- Location. The tools for working with IFS' native spherical harmonic data can currently extract all the data needed for the SCM at a given point on the Gaussian grid. Thus it has been simplest so far to take the location of the case study as that grid point. Extending these tools to handle non-grid points, especially in a Lagrangian context, is one of our immediate goals. The next question is: Should the ocean/sea ice data be interpolated to match that point, or is using the nearest point on that grid sufficient? How should the interpolation be done? This is related to the next point.

- Ocean thermodynamic variables. The primary thermodynamic variable in an ocean model might be *in-situ* temperature, potential temperature, or conservative temperature. Also, the vertical variable might be depth or pressure, with the other diagnosed from hydrostatic balance. How best to match this data to the SCM configuration, which typically uses potential temperature and depth (pressure implicit) as the thermodynamic variables, is an open question. This matters when interpolating data so that conservation laws of the model are respected and spurious static instabilities are avoided.

- Interface fluxes. The coupler requires initial conditions of some surface fluxes. However, the atmosphere, ocean, and sea ice models all use slightly different formulations for the bulk fluxes. It is unclear whether these differences in the initial conditions for the coupler variables alone make a significant impact on the coupled simulations. Currently, we use the ocean and sea ice formulations when calculating these fluxes, as these are bulk formula. In contrast, surface fluxes in IFS are treated implicitly, and vertical transport throughout the entire atmospheric column (including in-cloud turbulence and convection) has to be considered when calculating the surface fluxes.

3 Acknowledgments

The author acknowledges helpful comments from Kerstin Hartung and funding received from the Horizon 2020 project APPLICATE (Advanced Prediction in Polar regions and beyond: Modelling, observing system design and LInkages associated with a Changing Arctic climaTE) under Grant Agreement No: 727862.

References

ECMWF. Part III: Dynamics and numerical procedures. In *IFS Documentation CY43R3*, number 3 in IFS Documentation. ECMWF, 2017a. URL https://www.ecmwf.int/node/17735.

ECMWF. Part IV: Physical processes. In *IFS Documentation CY43R3*, number 4 in IFS Documentation. ECMWF, 2017b. URL https://www.ecmwf.int/node/17736.

Kerstin Hartung, Gunilla Svensson, Hamish Struthers, Anna Lena Deppenmeier, and Wilco Hazeleger. An EC-Earth coupled atmosphere-ocean single-column model (AOSCM.v1-EC-Earth3) for studying coupled marine and polar processes. *Geoscientific Model Development*, 11(10):4117–4137, 10 2018. ISSN 19919603. doi: 10.5194/gmd-11-4117-2018. URL https://www.geosci-model-dev.net/11/4117/2018/.

Felix Pithan, Andrew Ackerman, Wayne M. Angevine, Kerstin Hartung, Luisa Ickes, Maxwell Kelley, Brian Medeiros, Irina Sandu, Gert-Jan Steeneveld, H. A. M. Sterk, Gunilla Svensson, Paul A. Vaillancourt, and Ayrton Zadra. Select strengths and biases of models in representing the Arctic winter boundary layer over sea ice: the Larcform 1 single column model intercomparison. *Journal of Advances in Modeling Earth Systems*, 8(3):1345–1357, 9 2016. ISSN 19422466. doi: 10.1002/2016MS000630. URL http://doi.wiley.com/10.1002/2016MS000630.