Description and Evaluation of the NOAA Experimental Coupled Arctic Forecast System (CAFS) Model

Amy Solomon[1,2], Christopher J. Cox[2], Janet M. Intrieri[2],
P.O.G. Persson[1,2], Gijs de Boer[1,2], Matthew D. Shupe[1,2], Mimi Hughes[2],
Antonietta Capotondi[1,2]

[1]Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA
[2]NOAA Earth System Research Laboratory, Physical Sciences Laboratory, Boulder, CO, USA

Description and Evaluation of the NOAA Experimental Coupled Arctic Forecast System (CAFS) Model

Amy Solomon, Christopher J. Cox, Janet M. Intrieri, P. O. G. Persson, Gijs de Boer, Matthew D. Shupe, Mimi Hughes, Antonietta Capotondi

Contents

List of Tables

List of Figures

Abstract

Rapid change in Arctic sea ice requires actionable forecast information for making decisions governing marine and coastal community safety in this challenging environment. In response, the NOAA Physical Sciences Laboratory (PSL) has provided experimental, daily forecasts of Arctic weather and sea ice conditions to stakeholders through the Coupled Arctic Forecast System (CAFS) model since 2016. The model includes dynamical ocean, sea ice, land, and atmospheric models, coupled with a flux coupler. The CAFS model (here, CAFS) is initialized with GFS boundary conditions and satellite-derived sea ice concentration and sea surface temperatures. The model is run daily to produce pan-Arctic, 0- to 10- day forecasts of sea ice, oceanic and atmospheric fields. Alaska-specific regional products, sea ice drifts across the Arctic in support of the Sea Ice Drift Forecast Experiment (SIDFEx), and special-request meteograms and time-height cross-sections for field campaign support (e.g., for the 2020 MOSAiC year-long drift expedition) are made available through a website designed in collaboration with National Weather Service-Alaska Region (NWS-AR) partners.

To characterize and understand the predictive skill of coupled ice, ocean and atmospheric processes, CAFS has been evaluated using Arctic observations obtained from ships, buoy platforms, land observatories, and satellites from 2015 to present. In this paper, we describe the CAFS model and present analyses of sea ice dynamic and thermodynamic processes observed during the MOSAiC campaign (2019-2020). This work contributes to ongoing collaborations with the NWS to evaluate the performance of modeling products, techniques in the Arctic in support of development of the Unified Forecast System, and multi-model international collaborations, etc.

Introduction

Over the past several decades, dramatic Arctic environmental changes have become the "new normal" in the Arctic region. According to recent Arctic Report Cards (e.g., **Osborne et al. 2018; Richter-Menge et al. 2019; Thoman et al. 2020; Moon et al. 2021**), average annual temperature above 60° N and Arctic surface air temperatures have warmed at more than twice the rate of the rest of the planet. Rapid and significant physical environmental changes in the Arctic, such as increases in permafrost thaw and decreases in sea ice extent and thickness (e.g., **SWIPA 2017; IPCC 2019, ARC 2020**) have created significant community, economic, and security concerns.

The satellite record has documented a decreasing trend in annual pan-Arctic sea ice extent since 1979, resulting in annual minimum extents two standard deviations below the 1981-2010 average every summer since 2007 (**Serreze and Stroeve 2015**). Winter maximum (March) and summer minimum (September) 2021 sea ice extents were less extreme compared to the last couple of years, but the 15 lowest minimum extents have all occurred in the last 17 years (**Meier et al. 2023**). Summer 2021 saw the second-lowest amount of older, multi-year ice since 1985, and the post-winter sea ice volume in April 2021 was the lowest since records began in 2010. The amount of multiyear sea ice, based on available data since 1985, reached its second lowest level by the end of summer 2021, sea ice thickness was lower than recent years, and volume was at record low (since at least 2010) in April 2021.

Regionally, the Alaskan Arctic sector has also experienced record low sea ice extents since 2018, with the Bering Sea ice dramatically reduced during the 2018/2019 winter and the Beaufort and Chukchi Seas (north and northeast of Alaska, respectively) experiencing rapid ice loss in spring 2019. By the end of summer 2019, the Chukchi Sea reached one of the lowest ice extents in the satellite record for the region (**Thoman and Walsh 2019**). Consequently, the Bering and Chukchi Seas are now typically characterized by more young, thin first-year ice (FYI) and less old, thick multi-year ice (MYI - ice that survives at least one melt season; **Perovich et al. 2016**). There is an increasing trend in the date of "ice-covering" in the Beaufort and Chukchi Seas, with ice-over now typically two to three weeks later than was typical in the 1980s (**Thoman and Walsh 2019**). These longer periods of open water, decreased ice extent area, thinner ice, and warmer upper ocean temperatures, have set the stage for increased maritime access, use (military, commercial, and tourism), and future economic interests (**Adams and Silber 2017; Smith and Stephenson 2013**). Reduced sea ice also raises concerns over increasing coastal hazards in Alaska, in particular beach run-up and erosion during storms (e.g., **Overeem et al. 2011, Barnhart et al. 2016, Fang et al. 2018**).

The safe execution of maritime activities in remote, challenging, and hazardous Arctic waters requires accurate forecasting of environmental conditions to support search and rescue, oil spill

response, commercial fishing, transportation, community, emergency and ecosystem management, environmental research, tourism, and the energy and mineral industries, to name a few. Increased activity in the Alaskan Arctic sector increases the responsibility of U.S. national agencies to develop focused and skillful forecasts. In response, the National Oceanic and Atmospheric Administration (NOAA) supports the development and improvement of forecast models capable of adequately representing Arctic processes on both regional and global scales.

Accurate representations of physical processes governing atmosphere, ice, ocean and land, as well as the complex interactions between them, are critical for improving forecasts over a wide range of temporal (short-range to decadal) and spatial (local to global) scales. This is particularly true for forecasts of sea ice, which impacts most activities at the ocean-atmosphere interface. Use of coupled atmosphere-sea ice-ocean-land models is especially necessary for understanding the ice response to forcing from atmospheric storms, ocean currents, and the annual regional energy cycle. Correct representation of these system interactions requires accurate simulation of atmospheric winds and boundary layer structure, upper ocean heat content, mixed-phase clouds, and sea ice characteristics and their influence on surface energy balances, ice motion, and internal feedback mechanisms (REFS).

In this paper, we describe a regional coupled modeling system for forecasting the Arctic environment, including sea ice, on the weather time-scale. For our purposes, weather-scale is defined here as the 0- to 10- day time range during which the ice responds directly to quickly changing atmospheric conditions (**Mohammadi-Aragh et al. 2018**) which, in turn, are superimposed on slower energy forcings from the annual cycle and the ocean dynamic and thermodynamic states. By its very nature, the movement, growth, and melt of sea ice is driven by a combination of atmospheric and oceanic processes. Forecast models struggle to produce reliable sub-seasonal forecasts due to inaccurate representation of the complex interactions between different system components, where small errors in one process can produce large errors in the interactions.

In response to a general need for improvement of Arctic forecasting, the NOAA Physical Sciences Laboratory (PSL) modified an existing regionalized climate model, the coupled Regional Arctic System model (RASM, **Maslowski et al. 2012**), to serve as a research tool for assessing coupled physical processes driving the evolution of sea ice on the weather-scale. The scope of this modeling effort was originally conceived in 2015 as a tool for understanding how sea ice evolution impacts lower atmospheric and upper-oceanic structures. Quickly, however, the project goals were extended to deliver daily, 0- to 10- day forecasts and products to NOAA's National Weather Service-Alaska Region (NWS-AR) sea ice forecasters and to the ship-based Sea State field program (**Thomson et al. 2018; Rogers et al. 2018)** in Fall 2015. This routine forecast production provided an opportunity to compare model output to ship-based measurements of the Marginal Ice Zone (MIZ), and to inform the sorts of observations and

physics potentially required to support the future NWS operational, global, coupled modeling framework. Additionally, daily delivery of products to NWS-AR partners provided an iterative environment supporting both assessment of model skill and determination of coupled forecast usefulness and utility in an operational setting. In this manner, daily forecast delivery and performance assessment has been part of PSL model development for three consecutive fall freeze-up seasons (2015-2017). Starting in February 2018, the forecast output was extended to include year-round, daily forecasts to provide model performance assessment over the complete annual cycle (2018-2022).

In the next section, we present details of the coupled ice-ocean-atmosphere model including descriptions of individual model components and initialization fields. In **Section 3**, we provide examples of model hindcasts, validate the skill of these hindcasts against observations from ships, satellites, buoys, and land stations, and provide insight into key parameters for accurate simulation of sea ice evolution. In **Section 4** we outline current and future applications and stakeholders, including information on how this effort is informing development of the Unified Forecast System (UFS) –a global, coupled comprehensive Earth modeling and data assimilation system to be used in NOAA operations and by the research community. Finally, we summarize the results and provide examples of future work in **Section 5**.

2. Coupled Arctic Forecast System

The main objective of the Coupled Arctic Forecast System (CAFS) modeling project is to improve predictions of Arctic sea ice evolution on the weather-scale (0- to 10- days) by: identifying critical (mesoscale to synoptic) coupled, physical processes; characterizing process-level model deficiencies; and improving model representation of those key processes. In this section, we describe the model's individual components, and provide information on model configuration and initialization fields. The model version described here was "frozen" in fall 2019 to provide consistent Arctic forecasts throughout the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC; **Shupe et al. 2022**) campaign year (October 2019-September 2020). Comparing CAFS forecasts with the hallmark, ice-ocean-atmosphere MOSAiC measurements will be the emphasis of our continuing, coupled process-oriented modeling studies.

2.1 CAFS – Origin, Description and Model Components

CAFS was adapted from the Regional Arctic System Model (RASM) to provide regional forecasts on the weather time-scale. RASM is a limited-area (including all Arctic drainages, northern hemisphere storm tracks, sub-polar ocean circulations.), fully coupled, ice-ocean-atmosphere-land model (**Maslowski et al. 2012; Cassano et al. 2017**) which includes the

following components configured for the pan-Arctic region: Weather and Research and Forecasting (WRF) atmospheric model (**Skamarock et al. 2008**); Parallel Ocean Program (POP2) ocean model **(Smith et al. 2010)**; Los Alamos National Laboratory (LANL) Community Ice Model (CICE5; **Hunke et al. 2013**); and the Variable Infiltration Capacity (VIC) land hydrology model (**Hamman et al. 2018**). The ocean and sea ice models used in RASM are regionally configured versions of those used in the National Center for Atmospheric Research (NCAR) Community Earth System Model (CESM), with WRF replacing the CESM Community Atmospheric Model (CAM) and VIC replacing the NCAR Community Land Model (CLM). The components are coupled using a regionalized version of the CESM flux coupler (CPL7 REF?), which includes modifications important for resolving the sea ice pack's inertial response to transient (i.e., weather-scale) events **(Roberts et al. 2014)**.

The NOAA PSL version of RASM, renamed CAFS, was modified for short-term, weather-scale forecasts by:
- Limiting the horizontal domain to the Arctic Ocean and surrounding coastlines (**Fig. 1**);
- Improving the atmospheric horizontal grid spacing to 10-km, thereby explicitly representing more mesoscale atmospheric processes and running all components at a similar resolution;
- Initializing oceanic, sea ice, and atmospheric components with observations, operational analyses and reanalyses, and;
- Replacing the VIC land model with the NCAR CLM with horizontal grid spacing to 10-km consistent with the atmospheric model grid.

The current CAFS configuration includes the dynamical POP2 ocean model, the LANL CICE5.3 sea ice model, the NCAR CLM4.5 land model, and the Advanced Research WRF Version 3.6 atmospheric model. The individual model components are coupled using the NCAR Flux Coupler (CPL7) model. Additional modifications include: activation of the Morrison bulk double-moment cloud microphysics scheme for droplets and frozen hydrometeors in WRF that allows both size and number of hydrometeors to vary in response to environmental conditions **(Morrison et al. 2009)**; updating the NCAR CAM radiation schemes to use hydrometeor effective radii calculated in the microphysical scheme (**Collins et al. 2006**); and use of the revised Fifth-Generation Penn State-NCAR Mesoscale Model (MM5) surface layer scheme (with modified similarity function for stable conditions) (**Grell et al. 1994**) and the diagnostic, non-local Yonsei University (YSU) planetary boundary layer scheme (**Hong 2010; Hu et al. 2013**). Originally the CAFS domain covered the Arctic basin and surrounding coastal regions but was extended in 2018 to include the Bering Strait (to provide expanded forecast area guidance to NWS-AR) and Fram Strait (to provide sea ice forecasts for a MOSAiC drift track

intercomparison project (**Goessling et al. 2020**)). **Figure 1** illustrates the CAFS extended domain with an overlay of the MOSAiC Drift track.
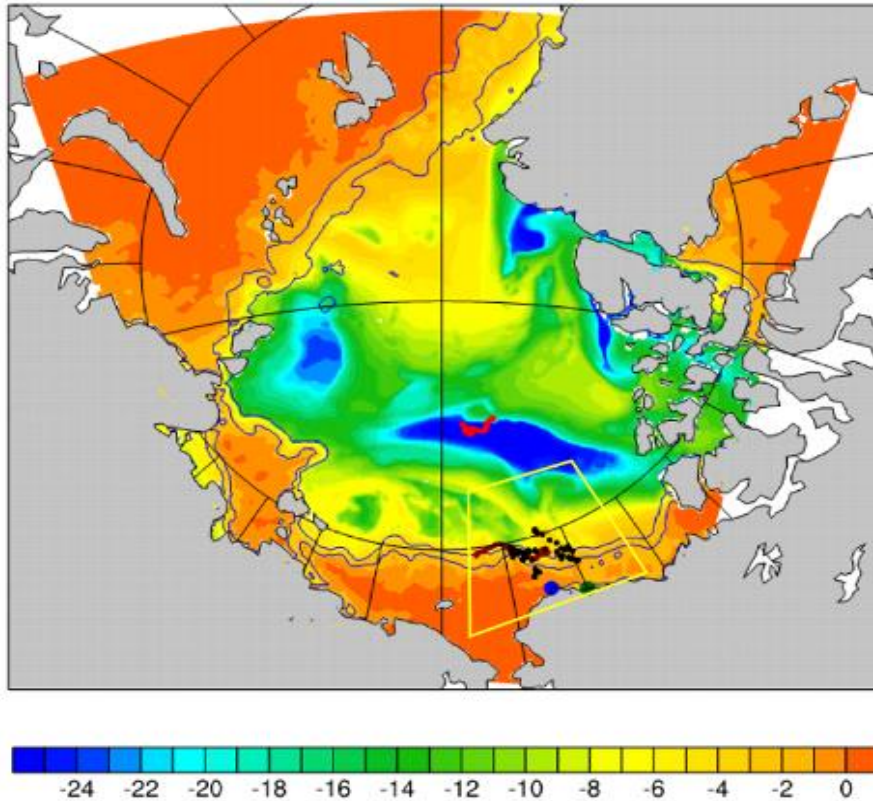
Figure 1. Map of CAFS domain. Overlaid on an example September SST map are locations of the main validation data sets used in this paper: Drift tracks of the Central Arctic buoy (red line) and MIZ buoy (brown line); Sikuliaq locations during Sea State (black dots) with the Sea State domain (yellow trapezoid); land observatories on the North slope of Alaska at Utqiagvik (blue dot) and Oliktok Point (green dot).

2.2 CAFS – Model Configuration

CAFS is run daily using the 0000 UTC analyses to produce 0- to 10- day forecasts and 6-hourly sea ice, atmosphere and ocean forecast guidance products. Currently, CAFS forecasts are initialized at the beginning of each daily run with the NOAA Global Forecast System (GFS) deterministic atmosphere forecasts (GFS analyses are used to initialize the hindcasts), a satellite-derived sea ice concentration (SIC), and the NOAA day/night blended satellite Sea Surface Temperature (SST) product. A Sea Ice Thickness (SIT) product from the European Space

Agency's (ESA) Cryosat-2 satellite was used to initialize the model in spring 2018 and has been running cycled thereafter. All individual model components are summarized in **Tables 1-4**.

Table 1. CAFS Model atmosphere component and specifications.

Weather Research and Forecasting	WRF-ARW 3.6
Horizontal Grid Spacing **# Grid Points**	10-km 384x432
Grid Configuration **# Vertical Levels**	Polar Stereographic There are 40 geopotential height levels and 39 pressure levels. T, U, V are on pressure levels and W is on geopotential levels. The first geopotential levels is ~23 m and the first pressure level is ~12 m
Time Step **Coupling Frequency**	WRF: 30-s Coupler: 30-min step
Lateral Boundary Conditions	0.5° GFS 3-hourly forecasts of T, P, WV
Initialization	T, P, U-V winds, & wv initialized with GFS analysis; Hydrometeor mass and number concentration initialized with values from the 24-h forecast from the previous day's forecast
Longwave/Shortwave Radiation	CAM
Surface Layer	Revised-MM5
Boundary Layer	YSU
Cloud Microphysics	Morrison Double-Moment Scheme
Convective Parameterization	Grell 3D
Information passed from WRF to Coupler	SLP, P, T, Tpot, specific humidity, air density, winds at lowest model level; Sfc LW flux; Sfc direct & diffuse SW, net SW flux; Precipitation type; Height of bottom atmospheric level
Information passed from Coupler to WRF	Area-weighted latent, sensible, momentum fluxes; T(Sfc & 2-m), 2-m specific humidity, 10-m winds; Sfc roughness length, wind stress; Upward LW sfc, direct & diffuse albedo; Sfc snow water equivalent; Evaporation; Land, ocean, ice fractions; Sfc stability; Air/sea exchange coefficient; Friction velocity
Nudging	None

Table 2. CAFS Model ice component and specifications.

Los Alamos Sea Ice	CICE 5.3
Horizontal Grid Spacing **# Grid Points**	Rotated Sphere shared with POP 420 x 464
Grid Configuration **# Ice Categories**	Rotated Sphere 7 Ice Thickness Categories
Time Step **Coupling Frequency**	CICE: 15-min Coupler: 30-min step
Initialization	SIT initialized with ESA CryoSat2 in February 2018 then free-running; Forecasts: Reinitialized daily with NESDIS Blended Day/Night SST and AMSR2 SIC; Hindcasts: Reinitialized daily with NASA MUR SIC and SST
Ice Rheology	EVP / no grounded ice
Information passed from CICE to Coupler	Turbulent fluxes for WRF BL & sfc schemes; Sfc albedo for WRF radiation scheme; Net heat, fresh water, and salt fluxes; Penetrating SW radiation; Ice-ocean stress for ocean model
Information passed from Coupler to CICE	P, T, Tpot, specific humidity, air density, winds at lowest atmos model level; Sfc LW down flux; Sfc direct & diffuse SW, net SW flux; Precipitation type; Height of bottom atmospheric level; Freezing/melting potential; Ocean sfc currents; Sea surface slope; Ocean sfc salinity, SST

Table 3. CAFS Model Ocean component and specifications.

Parallel Ocean Program	POP Version 2.1
Horizontal Grid Spacing **# Grid Points**	Rotated Sphere shared with CICE 420 x 464
Grid Configuration **# Vertical Levels**	B-grid 45
Time Step Coupling Frequency	POP: 15-min Coupler: 30-min step
Spin-up/Relaxation/Restarts	Ocean spun-up with strong restoring to TOPAZ July 2014 monthly mean & spun-up with GFS 6-11 July 2017 forcing (repeated 4 times). Relaxation to TOPAZ 2014 MM T/S 2 points from coastline below surface (to keep T & salinity from drifting in the absence of runoff); Relaxation to TOPAZ 2014 MM T/S 15 points along N/S boundaries; Ocean restarts from previous day's forecast
Ocean inflow/outflow	Not included in the configuration
Information passed from POP to Coupler	Ocean sfc currents; Sea sfc salinity, SST; Sea surface slope; qflux
Information passed from Coupler to POP	Net heat, fresh water, and salt fluxes; Penetrating shortwave radiation; Ice-ocean stress fluxes; Ice fraction; Ice & river runoff flux

Table 4. CAFS Model Land component and specifications.

NCAR Community Land Model	CLM 4.5
Horizontal Grid Spacing # Grid Points	10-km 384x432
River Discharge/Outflow	Not included in this configuration
Grid Configuration # Vertical Levels	Polar Stereographic 15
Time Step Coupling Frequency	30-min Coupler: 30-min step
Initialization	Cold start
Restart	Land restarts from previous day's forecast
Information passed from CLM to Coupler	Turbulent fluxes & wind stress for WRF BL & sfc schemes; Sfc albedo for WRF radiation scheme
Information passed from Coupler to CLM	P, T, Tpot, specific humidity, air density, winds at lowest atmos model level; Sfc LW down flux; Sfc direct & diffuse SW & net SW flux; Precipitation type; Height of bottom atmospheric level

2.2.1 Lateral Boundary Conditions

The GFS, NOAA's operational global weather forecast model, is produced by the NWS National Centers for Environmental Prediction (NCEP) and used daily by operational forecasters across the U.S. CAFS forecasts use the global, 0.5° GFS, 3-hourly forecast fields of pressure (P), temperature (T), horizonal winds (U), and water vapor (QV) for lateral forcing at the boundaries to produce higher-resolution (10-km) regional forecasts. GFS biases therefore may affect CAFS performance, but such impacts are not specifically addressed in the current study. Ocean lateral boundaries at the gateway regions are relaxed to a coupled, ocean-sea ice data assimilation system developed for the Arctic (TOPAZ4; **Xie et al. 2017; Sakov et al. 2012**) using 2017 monthly temperature and salinity fields.

2.2.2 Initial Conditions

Although CAFS does not have direct data assimilation capabilities, GFS assimilates atmospheric, land, ocean, and ice observations available through the Global Telecommunication System (GTS). GFS initializes sea ice conditions using satellite analyses; however, the atmosphere-only model does not evolve sea ice during the forecast time period (0-16 days) and is therefore inadequate for sea ice forecasting guidance needs. Additionally, GFS does not include initial

cloud property fields, so CAFS hydrometeor mass and number concentrations are initialized with fields from the first day of the previous day's hindcast to reduce spin-up time required to develop mature clouds in CAFS.

To develop skillful sea ice forecasts, a model must be initialized with accurate sea surface temperature (SST) and sea ice states (**Allard et al. 2018; Posey et al. 2015**). CAFS is initialized each day with the Advanced Microwave Scanning Radiometer (AMSR2) sea ice concentration (SIC) data (**Spreen et al. 2008**). The AMSR2 radiometer is a conically-scanning passive microwave radiometer system that measures in seven frequency bands ranging between 6.925 GHz and 89.0 GHz at both horizontal and vertical polarizations. The 6.25 km AMSR2 SIC information is interpolated to the CAFS horizontal grid. The atmospheric model grid resolution is 10-km but the ocean is on a non-uniform grid which is ~9-10 km.

CAFS ocean surface temperatures are initialized with a daily, satellite-derived NOAA day/night blended product (**Maturi et al. 2017**). NOAA's National Environmental Satellite Data and Services (NESDIS) generates a daily operational 0.05° global high-resolution satellite-based SST analyses with a nominal product latency of 1 to 3 hours from product generation. This analysis combines SST data from U.S., Japanese and European geostationary infrared imagers, and low-Earth orbiting infrared (U.S. and European) SST data, into a single high-resolution 5-km product.

CAFS forecasts use initial sea ice thickness (SIT) conditions from ESA's CryoSat-2 and Soil Moisture and Ocean Salinity (SMOS) satellites on 14 February 2018 and are thereafter initialized-- with fields from the first day of the previous day's forecast. Comparisons with the NASA ICESat-2 laser altimeter freeboard and CryoSat-2 observations from May 2019 were consistent with CAFS to within ~15%. For example, all of the main SIT features such as the higher values within the Canadian Archipelago and the Beaufort MYI "arm" features (pers. comm., **Farrell 2019**) were present in the CAFS thickness fields.

3. CAFS Model Validation

CAFS forecasts and hindcasts were extensively validated with a variety of observations obtained from oceanic, ice, and land locations. Hindcasts were evaluated to identify model biases that limit forecast skill and identify the model physical processes that cause hindcasts to drift away from observations. In this section, we present a selection of hindcasts validated against buoy observations, satellite measurements, and concurrent in situ observations to demonstrate model skill in predicting atmospheric state variables, wind and boundary layer structures, synoptic features, surface radiation terms, cloud microphysics, and ice and ocean properties. Model hindcasts are compared to observations from: the ONR sponsored Sea State campaign (Oct-Nov

2015); land-based atmospheric observatories from two Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) Program observatories (2017, 2018) located on the North Slope of Alaska (Oliktok Point and Utqiaġvik, Alaska; see **Uttal et al. 2015**); buoys in the Beaufort and Chukchi Seas (2015, 2017, 2019); and satellite-based instruments (2015, 2018, 2019). The validation datasets, which include both vertically-resolved observations (e.g., profiles of ocean and atmospheric properties including clouds), as well as point observations (e.g., surface flux components, sea ice characteristics), are described in more detail below.

3.1 Validation of Coupled Ice-Ocean Properties

The ONR Sea State campaign collected measurements in the Chukchi Sea region from the light icebreaker research vessel (R/V) *Sikuliaq* during the 2015 freeze-up period. The *Sikuliaq* transected various ice (and temperature) conditions during its 6-week cruise –from open water (where temperatures were relatively warm), to the MIZ (where temperatures were significantly colder), and up to the edge of the MYI pack as the fall freeze-up of first year ice evolved. In essence, the ocean froze under the ship during the course of the Sea State campaign, as described by **Thomson et al. (2016)** and **Persson et al. (2016).** The extensive suite of Sea State measurements used to assess model performance included those obtained from the: ocean [SST, underway Conductivity-Temperature-Depth (uCTD)]; atmosphere [basic meteorological parameters such as near-surface wind, pressure, temperature; boundary layer stability; turbulent fluxes of sensible and latent heat, and momentum; radiative fluxes; 4 times daily rawinsondes]; and ice [SIT, SIC, type]. Over 4300 uCTD's and 165 rawinsondes were launched during the campaign to characterize the coupled atmosphere-ocean system. Neither rawinsonde nor surface data were sent to the GTS, so these are independent of GFS analyses.

To facilitate assessment of CAFS performance, model error, and predictability skill as a function of forecast hour, model output from the four surrounding points were interpolated to the exact location of the ship at the validation time. The bar-whisker plots in **Figure 2** illustrate model performance compared to oceanic temperature profiles from uCTDs (blue) and atmospheric temperature profiles from rawinsondes (red). These plots illustrate model error of the coupled ocean-atmosphere state at the location of the German Icebreaker *Polarstern* over the yearlong campaign for forecast lead times of 12-h, 24-h, 180-h and 240-h. Observational profiles, which extend vertically from ~150 m below the ocean surface to ~120 mb in the upper atmosphere, capture important features and transitions that need to be captured by the model to ensure accurate projection of ice properties. This includes the structure of the ocean halocline (~50-m below ocean surface), the ocean mixed layer (~0 - 25-m depth), the ice surface, the atmospheric surface layer (lowest ~20 - 50-m), the atmospheric boundary layer (~975 - 900 mb), the free troposphere, and the tropopause (~250 mb). Bars and whiskers are shown for all atmospheric and ocean model levels, with white dots (It would be good to increase the diameter of the white dots or to give them another color (e.g., green or bright red) to make them visible. They are hard to

see) indicating median values and black dots indicating mean values. …the oceanic comparisons matched uCTDs (162) to the nearest-in-time (within 12-h) rawinsoundings.

An examination of the upper two panels in **Figure 2**, reveals temperature errors throughout the column of less than ~0.5 ° C (black dots) up until 36-h forecast times (not shown). Notable exceptions include larger deviations evident at the three major temperature transition zones: the tropopause (+/-1 ° C out to 180-h); the boundary layer air temperatures (+2-3 ° C after 180-h); and the ocean halocline (~-0.5 ° C across all lead times). Mismatches between the height of the simulated and observed tropopause are likely related to a combination of errors in the GFS analyses, potential model thermodynamic biases, the CAFS modeled height of the tropopause (**Wilson et al. 2011**) and model-measurement resolution differences at the tropopause.

Errors in the ocean surface temperatures are small (< -0.25 ° C across all lead times) however, larger errors are apparent in simulated temperatures around the ocean halocline (~ -50 m) where Root Mean Square Error (RMSE) values remain nearly constant (~ 0.5 ° C across all lead times) indicating that the errors at depth are associated with errors in defining the depth of the mixed layer. Overall, the oceanic temperature errors remain steady across the different lead times due to daily initialization of the ocean surface satellite-derived SST's, as well as, slower evolution and mixing of oceanic temperatures, in general. Earlier versions of the model allowed the ocean to free-run with error at initialization from accumulating biases. This issue ultimately led us to rely on a physically-consistent SST/SIC initialization pair that was initialized at the start of each forecast time.

It is generally known that NWP models lose the greatest predictability and skill within the 5- to 7-day range (REF?) and this is consistent with CAFS results, as illustrated in the difference between the 120- and 240- hours profiles (lower 2 panels of **Fig. 2**). This is especially evident in the atmospheric boundary layer (925 mb and below), where temperature bias error in the lowest 2-km grows rapidly after 5 days, from roughly +1.5 ° C at 180 h to +2.5 - 3 ° C at 240 h. This warm bias in comparison to observations indicates that the model is unable to maintain observed boundary layer stratification and rapidly evolves into a less-stable state

An analysis of CAFS hindcast skill over all forecast lead times (every 6-h), shows RMSE of temperature growing from approximately 1 ° C to 4 ° C starting at day-5 within the boundary layer and from around 1 ° C to 3 ° C at the tropopause after day-5.5 (not shown). RMSE of ocean temperature remains fairly constant across all forecast lead times (~1 ° C throughout all

depths except for 2 ° C at the halocline). RMSE of temperatures in the free troposphere (~850 to 275 mb) varies from 1 ° C and grows to around 2.5 ° C starting on day-7.5.
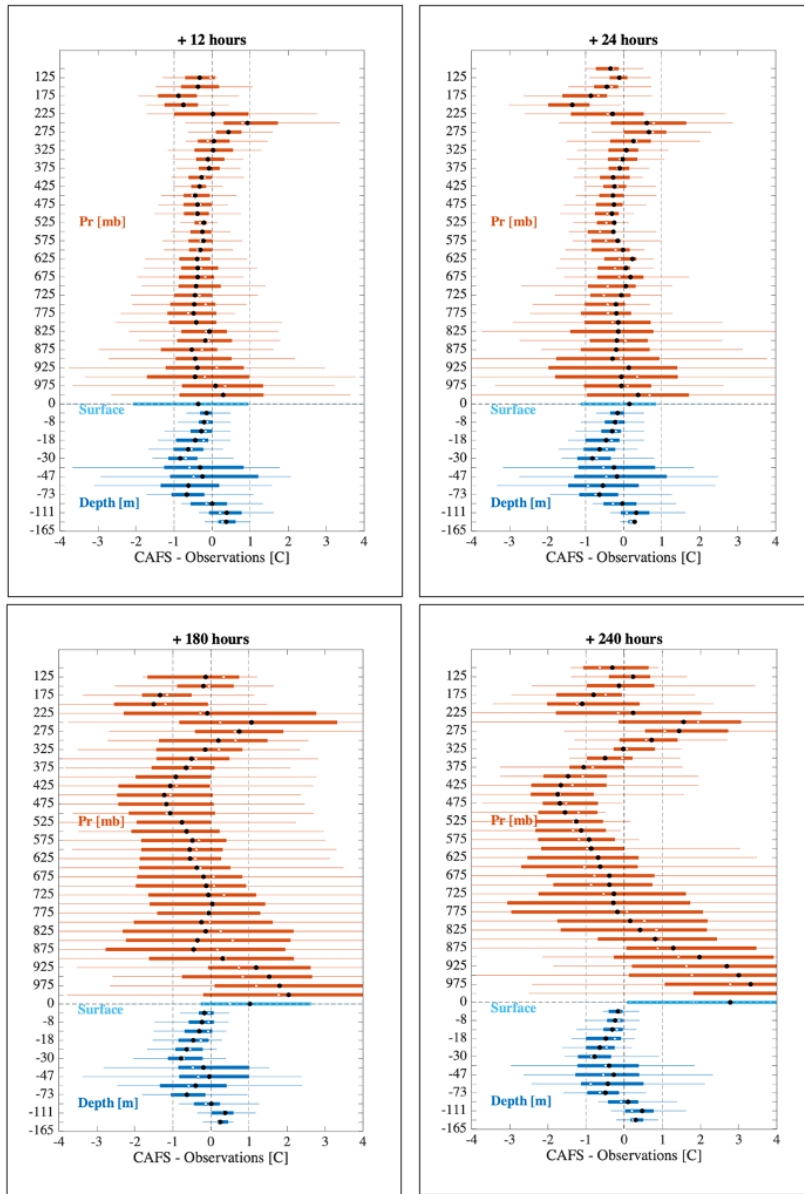

Figure 2. Combined ocean CTD and atmospheric rawinsondes temperature (° C) error profiles from the location of the *R/V Sikuliaq* over the 6-week campaign at forecast lead times of 12-h, 24-h, 180-h and 240-h. Bar and whiskers are shown for all atmosphere and ocean model levels, with thin lines indicating minimum and maximum values, and thick lines indicating 25% and 75% values. Black dots indicate median values and the white dots indicate mean values.

To understand how CAFS handles complex ice-ocean-atmosphere coupled processes, model output was compared to surface meteorological quantities (**Figs. 3a-d**), as well as various surface

flux parameters including radiative up- and downwelling longwave and shortwave fluxes (LW$_u$, LW$_d$, (**Fig. 3e**), SW$_u$, SW$_d$), sensible heat flux (H$_s$), and latent heat flux (H$_l$) measured during the Sea State campaign.
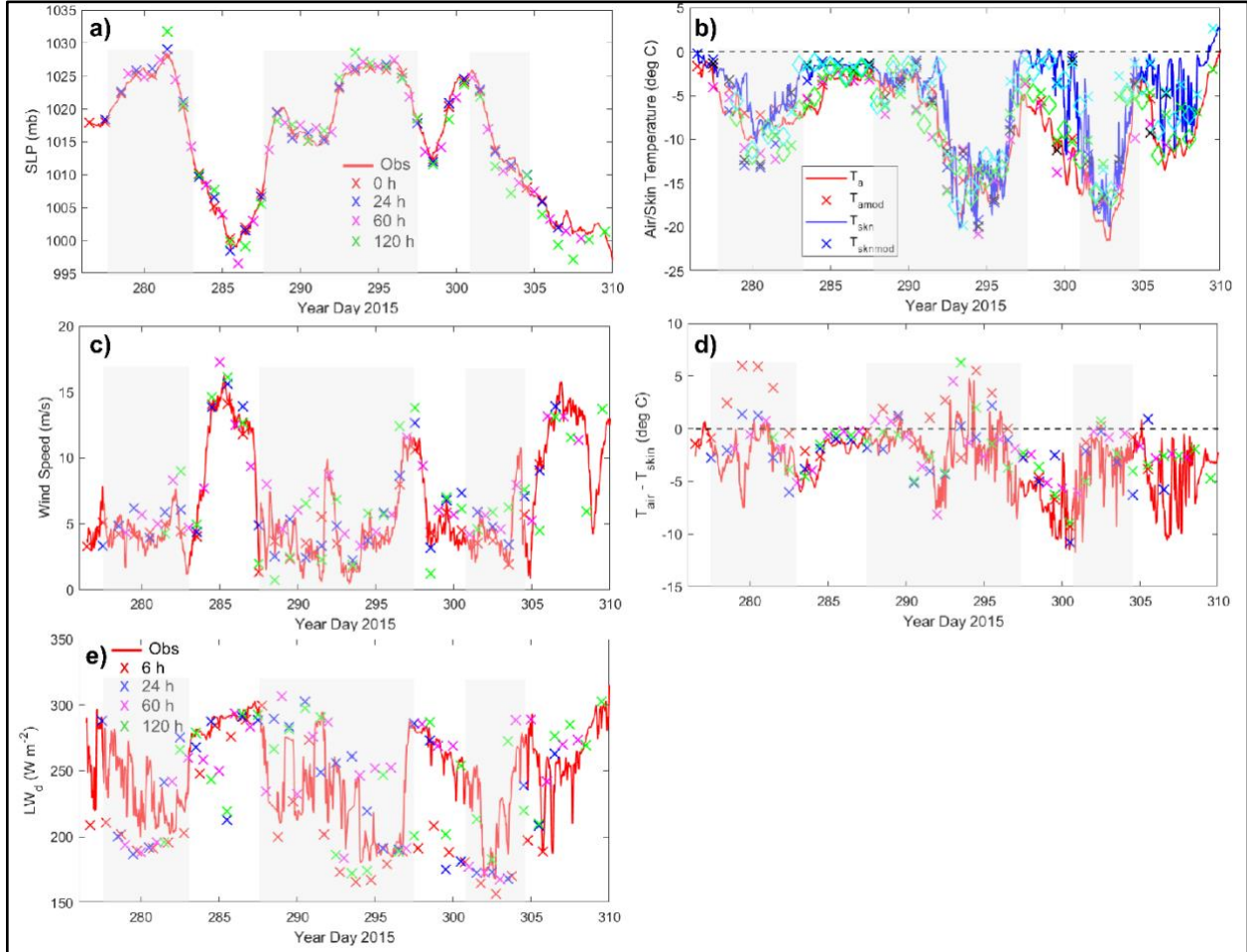

Figure 3. Comparisons between observed (solid lines) and model output from hindcast lead times ("x") for: a) Sea Level Pressure (mb); b) air (red) and skin (blue) temperature (°C); c) wind speed (ms-1); d) surface layer stability (Tair – Tskin; °C); and e) downwelling longwave radiation (LWd; Wm-2). Model lead times are 0-h [red; except 6-h in e)], 24-h (blue), 60-h (magenta), and 120-h (green), so each observation time includes multiple points from the sequential daily initializations. Observed data are 1-h running means and model output was interpolated to the observation heights (e.g., 15.4 m for Tair, 16.5 m for winds). The times when the ship was away from the ice edge are shaded.

Temperature observations at 15.4 m height on the ship's bow tower show that the ship spent four time periods near the ice edge (warmer temperatures) and three periods within the growing first-year ice region (colder temperatures) (**Fig. 3b**). During several of the ice-edge periods, lower

sea-level pressure (SLP) and larger pressure gradients (**Fig. 3a)** were observed, producing stronger winds **(Fig. 3c)**. More details of the atmospheric and ice conditions during Sea State are provided by **Persson et al. (2018)**. CAFS captured synoptic-scale temporal/spatial variability reasonably well, with only small errors in SLP (**Fig. 3a)** and only slightly larger errors in temperature and wind speed (**Figs. 3b and 3c)**. Forecast (hindcast) errors as a function of forecast time (not shown) demonstrate that SLP bias decreases from near 0 hPa for 12- to 36-h forecasts to -0.8 hPa between 60- and 120-h forecasts, with a mean bias of -0.4 hPa. RMSE increases from 0.5 hPa at 12-h to 2 hPa near 100- and 120-h. Wind speed bias and RMSE are approximately +0.8 ms^{-1} and 2.2 ms^{-1}, respectively, with only a small increase in RMSE with increasing forecast time. The bias and RMSE of air temperature (T_a) average +0.6 °C and 2.8 °C, while the bias and RMSE of skin temperature are around -1.1 °C and 3.6 °C, respectively. Hence, in the mean, the CAFS-generated surface layer along the ship track is too stable throughout each hindcast, since the skin temperature is too cold and the air temperature too warm. This manifests as a positive bias in stability in CAFS in **Figure 3d**.

Downwelling longwave radiation (LW_d) provides a measure of the degree of cloudiness, with high values linked to cloudy conditions and low values connected to clear sky conditions. **Figure 3e** shows that CAFS had some success in predicting sky conditions, though there were also many periods with large discrepancies. The mean bias in LW_d for 1- to 5-day hindcasts was extremely small (+0.3 W m^{-2}), though the large mean RMSE of 39.8 Wm^{-2} reflects timing offsets in otherwise correct simulations. This RMSE value is about half of the upper limit of ~80 Wm^{-2}, given by the difference between cloudy and clear conditions (see **Fig. 3e**) if the model were incorrect at every validation time. The sensible heat flux (H_s) is also an important component of the energy budget in this environment. While not shown in **Figure 3**, H_s has a bias and RMSE of -1.9 Wm^{-2} and 32.4 Wm^{-2}, respectively, where the small negative mean bias is consistent with the stability bias in the CAFS-simulated surface layer, while the large RMSE reflects impacts of clouds and ice conditions on H_s.

In general, models struggle to capture features of the complex Marginal Ice Zone where the fine-scale ice features cannot be adequately captured by satellite initialization fields. As shown in **Figure 4**, LW_d (cloud) hindcasts have significant errors as do hindcasts of SIC near the ship. **Figure 4a** reveals that CAFS predicts SIC reasonably well during the 3 periods when the ship was in the interior of the ice, where observed SIC's were close to 100% and the predicted SIC's were around 80-98%. However, during periods when the ship was near the ice edge with observed SIC varying rapidly (spatially and temporally) between 0-100%, CAFS often significantly underestimated (e.g., Year Day 283-287) or overestimated (e.g., Year Day 305-308) SIC. A scatterplot of difference between simulated and observed SIC as a function of the observed SIC (**Fig. 4b**) shows a broad error scatter within the 100% error range, and that the model prediction errors are not concentrated near the desired 0% value, except possibly for observed SIC >88%. Note that SIC at the initial time in the model ("O" in **Fig. 4b)** is

underestimated for observed SIC of < 75%. Hence, in all but the highest SICs, simulated values are hindered by poor initialization, resulting in poor forecasts for this single point validation.
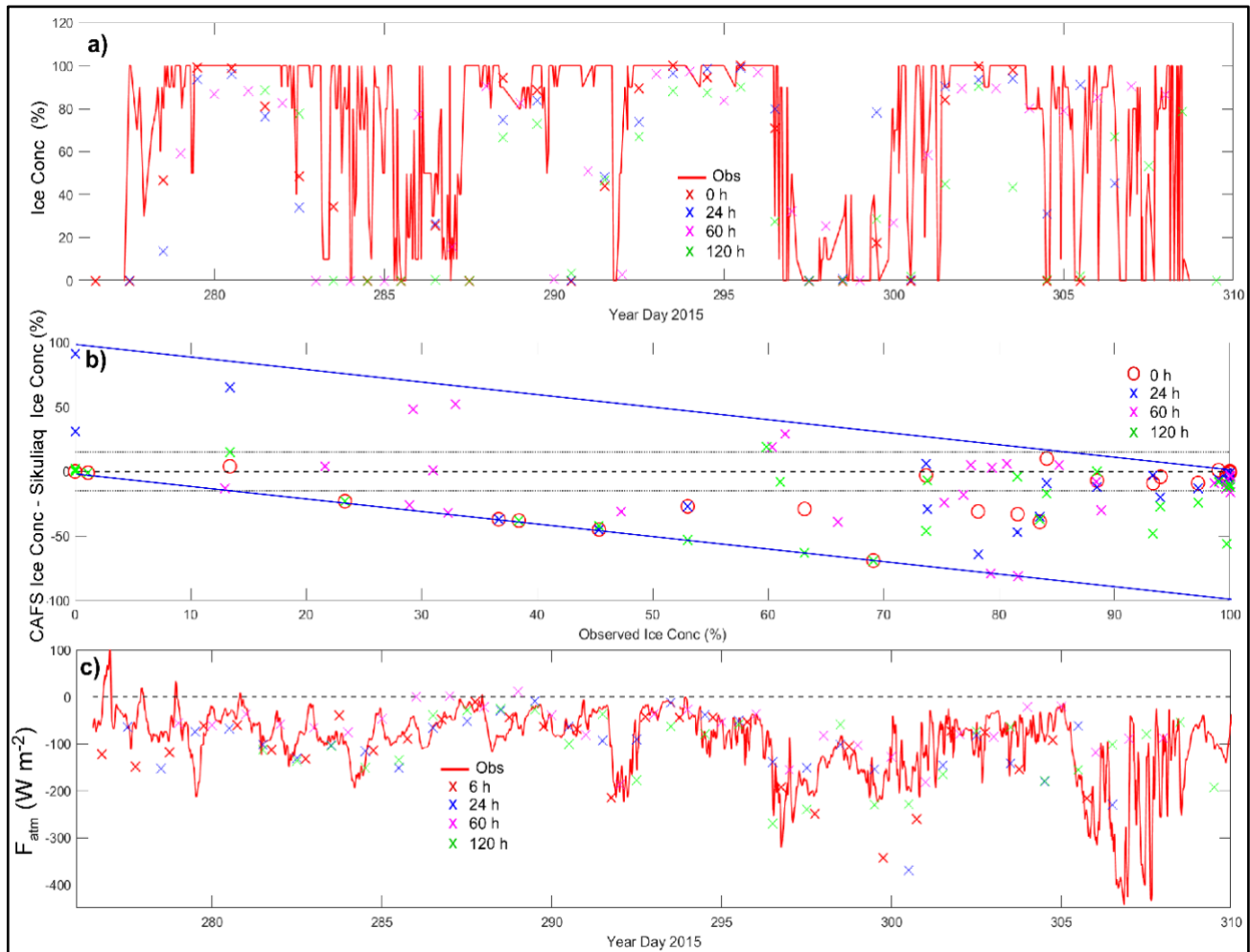

Figure 4. a) Time series of observed (red line) and modeled ice concentration for 0 h (red x), 24-h (blue x), 60-h (magenta x), and 120-h (green x) lead times; b) scatterplot of modeled minus observed ice concentration as a function of observed ice concentration for 4 model lead times; and c) time series of observed (red line) and modeled Fatm for 4 model lead times. Modeled values are from hindcasts. The blue lines in b) bracket the possible range of difference points.

This poor prediction of SIC near the ice edge leads to poor predictions of air-surface temperature differences and, therefore, turbulent heat flux (H_s and H_l) near the ice edge. As a result of the erratic but significant errors in LW_d (cloudiness) and turbulent heat flux (ice concentration) near the ice edge, evaluation of net atmospheric energy flux ($F_{atm} = SW_{net} + LW_{net} - H_s - H_l$) **(Fig. 4c)**

shows some periods where the model values compare reasonably well to the observations (e.g., Year Day 280-285) and others where F_{atm} was simulated quite poorly (e.g., Year Day 297-301; Year Day 306-308). The mean 120-h bias in the model F_{atm} is only 9.4 Wm^{-2}, while the mean RMSE is large at 72.1 Wm^{-2} for these hindcasts, reflecting the model's difficulties with predicting cloud cover and sea ice concentration near the ice edge.

To examine the sea ice forecasting success of CAFS hindcasts on a more regional basis, regional observed (AMSR2 satellite) and modeled SIC fields were compared (**Fig. 5**). In general, observed and modeled SIC fields diverged locally during the 5-day simulations initialized on each day (**Figs. 5a-c**), but the model performed reasonably well when examining the regional statistics. **Figure 5d** shows that CAFS produced similar amounts of ice growth over 5-day periods during Sea State as observed with AMSR2 but was unable to reproduce smaller areas of ice loss (through advection or melt) during the few instances that this was observed.

In general, CAFS-simulated surface radiation and turbulent fluxes are in reasonable agreement with Sea State observations after accounting for discrepancies expected from errors in ice edge location and SIC, and for errors in the surface-layer stability described previously. When mean fluxes are plotted as a function of relative distance to the ice edge, over-ice and over-water fluxes are in reasonable agreement after accounting for stability errors (not shown). Fluxes near the ice edge, however, contain larger errors because of error in representing ice edge location and sea ice properties. Surface radiation quantities are more erratic, likely because of misrepresentation of clouds in CAFS. Additionally, although smaller-scale dynamic structures such as low-level jets (e.g., **Guest et al. 2018**) and turbulent flux gradients at the ice edge were resolved in CAFS, differences in timing or strength were attributed to inexact location of the initialized ice edge and ice thickness. Resolving smaller scale features represented by point measurements in the highly

variable MIZ is a challenge since the resolution of satellite-derived initialization fields (typically 3.0 - 12.5 km) does not capture the complex reality of the ice edge.
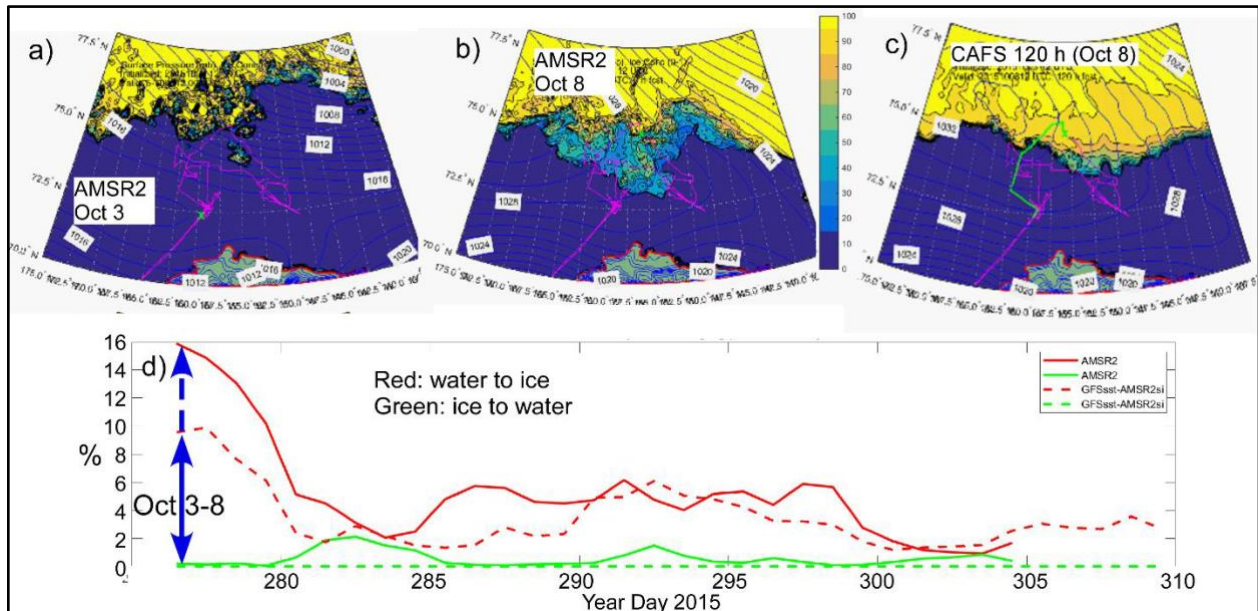

Figure 5. Observed and modeled ice concentration 5-day changes within the Sea State domain. Shown are sample AMSR2 ice concentration fields from Oct 3 and Oct 8 (a & b), and the forecast ice concentration field on October 8 (c) initialized 120 h earlier using the October 3 AMSR2 field in a) as initial condition. Panel d) shows the percentage of pixels within the shown Sea State domain that change from water to ice (red) and from ice to water (green) during the 5 days (120-h) after the given time for AMSR2 (solid) and the CAFS model hindcasts (dashed). The October 3 (Year Day 276) to October 8 (Year Day 281) observed and modeled changes are shown by the double blue arrow. 15% ice concentration is used as the ice/no-ice threshold. The Sea State ship track is shown in each panel (magenta line), with the track for the given day marked in green.

3.2 Validation of Near-Surface Atmospheric and Sea Ice Properties using Buoys

Measurements of atmospheric surface air temperature (SAT), sea level pressure (SLP), 2-m air temperature and ice motion from the International Arctic Buoy Programme (IABP) were used to further validate CAFS hindcasts. The U.S. component of the IABP network (USIABP) is managed by the University of Washington and the U.S. National Ice Center (NIC). The USIABP provides real-time data that are uploaded to the GTS and therefore available for assimilation into operational weather forecasts (e.g., GFS). In this section, we provide an overview of work completed over the past 2 years to assess CAFS model performance using drifting buoys. Assessing CAFS performance in a range of SICs is important for understanding model representation of coupled processes, as illustrated by the validations in **Section 3a**. In **Figure 1,**

two IABP buoy tracks (2015) are shown (brown and red lines) overlaid onto an example fall season CAFS surface temperature (° C) field – one buoy in high concentration sea ice in the "Central Arctic" (here, CA buoy, reporting SLP and T_{2m}) and one buoy in the Marginal Ice Zone (here, MIZ buoy, reporting SLP and SAT). Comparison of CAFS performance statistics in these two environments illustrates some of the model's performance characteristics.

Figure 6 shows Taylor diagrams (**Taylor 2001**) of CAFS hindcasts and persistence forecasts (from 1 October – 5 November 2015, concurrent with the Sea State campaign) at 0-, 1-, 3-, 5-, 7-, and 10-day forecast lead times of SLP and T_{2m} (top row) and SLP and SAT (bottom row) compared to observations taken by the CA and MIZ buoys depicted in the **Figure 1** drift tracks. Taylor diagrams graphically represent how accurately a model simulates the real-world system by mapping pattern correspondence between simulated and observed fields through a single point that combines information on the correlation coefficient and ratio of the standard deviations of the two patterns.

Together, these statistics provide a summary illustration of how well the model reproduces observations by quantifying the forecast skill in terms of three statistics: Pearson correlation coefficient; standard deviation; and bias. The Pearson correlation coefficient is mapped along the azimuthal angle; the normalized standard deviation is proportional to the radial distance from the origin; and the bias is defined by the symbol size as indicated in the legend located in the upper left of each panel. Forecasts that match best with the pattern of observations will have the highest correlation, lowest bias and be clustered around the location of the graph marked REF on the x-axis (Correlation=1.0, Normalized Standard Deviation=1.0). CAFS hindcasts, as a function of lead time (by color in the upper right legend of each panel), are indicated by the #1 above each symbol and persistence forecasts are indicated by the #2 above each symbol. The lowest bias for any different lead time is denoted by a solid square (which for "persistence" is the 0-day forecast in all panels). In general, note that many of the model points in **Figure 6** are clustered along the dashed arc showing the low standard deviation and indicating that the amplitude of the pattern variations compares very well. However, as expected, the correlation

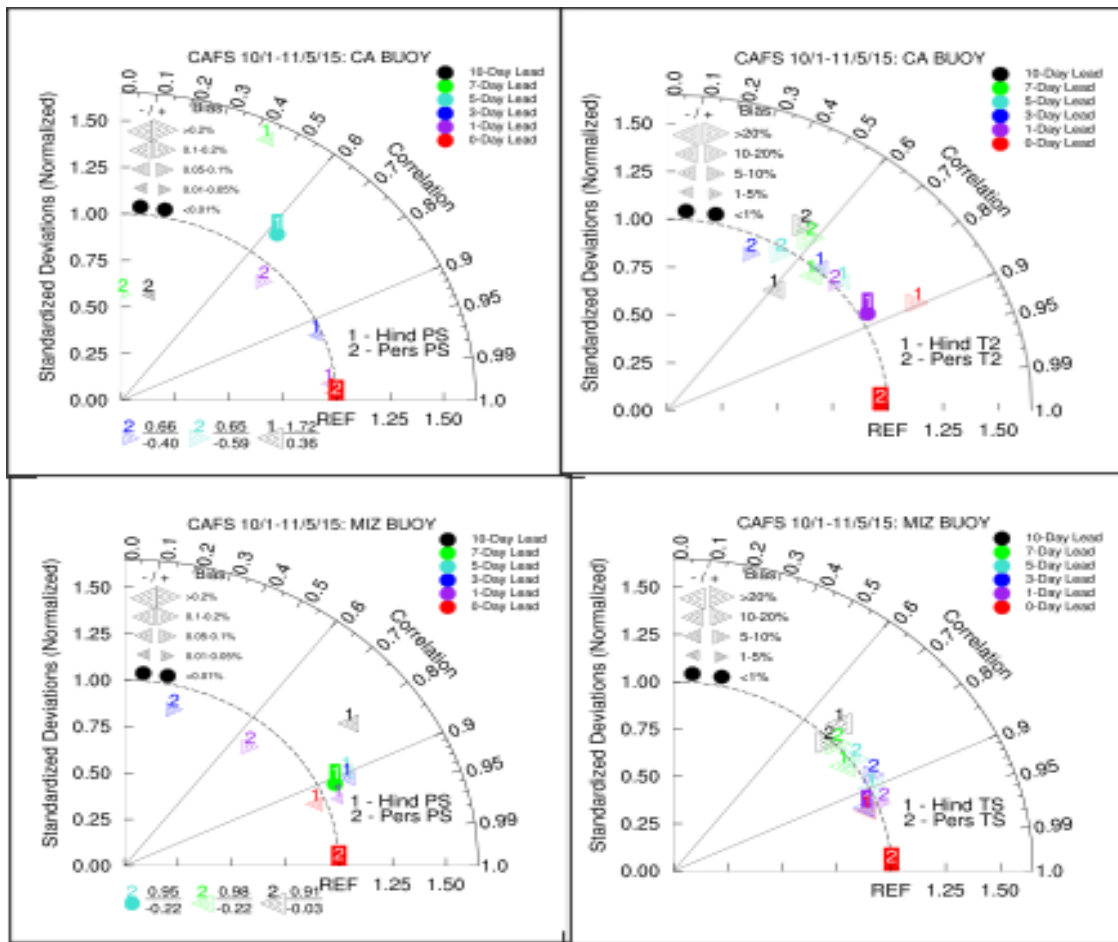decreases with increasing hindcast lead time.

Figure 6. Taylor diagrams of 10-day CAFS hindcasts during the Sea State period (1 Oct – 5 Nov 2015) compared to observations obtained from IABP CA buoy and MIZ buoy (tracks shown in Fig. 5). Each diagram shows standard deviation (radial axis), Pearson's correlation coefficient (azimuthal axis), and bias (legend in upper left) of the hindcasts as a function of lead time (#1 above each symbol) as well as for a persistent forecast (#2 above each symbol). Panel: A) Surface pressure (hPA) along the CA drift track; B) 2m temperature (°C) along the CA drift track; C) Surface pressure (hPa) along the MIZ drift track; D) Surface air temperature (°C) along the MIZ drift track. The lowest bias for the range of lead times (0-, 1-, 3-, 5-, 7-, 10-day) shown is denoted by a solid square.

In summary, Taylor diagrams allow us to visualize forecast skill and investigate sources of error compared with a persistence benchmark. For example, the fact that CAFS hindcast SLP in the central Arctic beats persistence but rapidly drifts away from observations may be due to errors in

model physics, and T_{2m} in the central Arctic (and also SLP in the MIZ) is not accurately initialized, indicating that better initial conditions are needed to improve forecast skill.

In addition to these Taylor diagrams, we present a regional comparison of CAFS-modeled sea ice area (total area covered by sea ice based on AMSR2 satellite sea ice concentration) for the Chukchi and Beaufort Seas area (defined as 180°-125° W) during the 2017 and 2018 freeze-up seasons (defined as 1 November – 5 December). **Figure 7** shows RMSE of 35 individual, 10-day CAFS forecasts (red dots) and persistence forecasts (black dots) as a function of lead time. In both 2017 and 2018 forecast years, CAFS beats a persistence forecast within 2 days. CAFS error growth is less than 20% between a 2-day and 10-day lead time, while persistence forecast error grows by 220% over the same lead times. That is, during the freeze-up season, CAFS grows ice within the domain while a persistence forecast doesn't, though the locations of ice growth may differ from observations (e.g., **Fig. 4**). The freeze-up season is the time of year when a coupled model will do best against persistence forecasts.
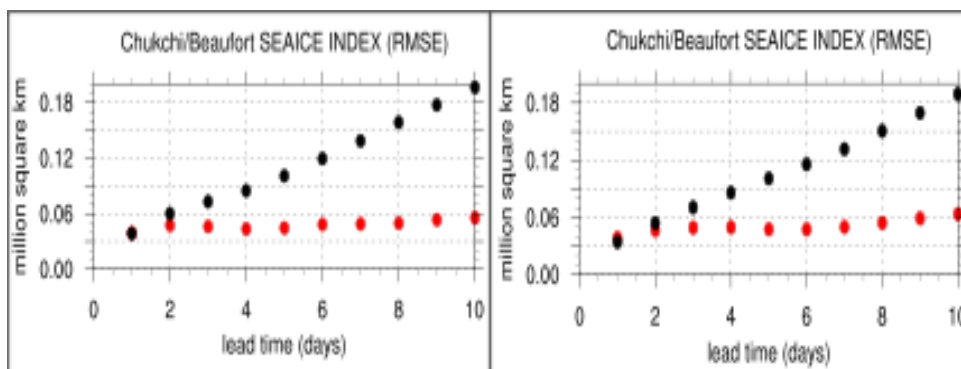

Figure 7. RMSE of 35 10-day CAFS hindcasts (red dots) and persistence forecasts (black dots) of sea ice area as a function of forecast lead time for 2017 (left) and 2018 (right) freeze-up (1 November to 5 December) in the Beaufort and Chukchi Seas, defined as 180-125°W.

Finally, we describe results of sea ice dynamic quantities as compared to IABP buoy observations and in comparison to other operational forecast centers. The IABP observations were incorporated in near real-time as part of the World Meteorological Organization's (WMO) Year of Polar Predictions (YOPP; **WMO WWRP 2016**) Sea Ice Drift Forecast Experiment (SIDFEx) project (**Goessling et al. 2020**). The SIDFEx project is an international community effort to compare daily drift forecasts (at lead times from days to a year), from a variety of modeling centers (including NOAA) to the locations of select IABP buoys, as well as the

location of the MOSAiC campaign's *Polarstern* icebreaker during the year-long drift deployment (2019-2020).

The systematic assessment of drift forecasting capabilities allows for improvements related to physical understanding of sea ice and can help identify and resolve model shortcomings, predictability limits, and interaction between sea ice and boundary layer physics in the atmosphere and ocean. It has also provided an opportunity to compare CAFS with other drift model products from international modeling centers in Europe (ECMWF), Canada (Environment Canada), the United Kingdom (UKMet), Germany (DWD), and Norway (MetNo) to name a few. The comparisons can be found online and provide forecast skill assessment over a range of ice conditions (https://www.polarprediction.net/key-yopp-activities/sea-ice-prediction-and-verification/sea-ice-drift-forecast-experiment/).

3.3 Validation of Properties using Land Station Data

Atmospheric measurements obtained from two DOE ARM sites located along the North Slope of Alaska (NSA), Utqiaġvik (formerly, Barrow) and Oliktok Point (OLI) (**de Boer et al. 2018**) were also used to compare CAFS-simulated fields to measurements and to assess biases of surface pressure, 2-m temperature, wind speed and direction, surface water vapor, and radiative quantities (**ARM 2018a; 2018b**) for the time periods spanning 1 November through 5 December in 2017 and 2018. Observations are 1-min averages which were then averaged across 20 minutes to approximate correspondence to the "spatial scale" included within the model's 10-km grid spacing. A selection of time series for several lead times (0.25-, 2-, 5-, 10-day) and the corresponding Taylor diagrams for 2-m temperature (° C) and the downwelling LW radiation term (W m^{-2}) are shown in **Figures 8-9**, respectively.

Additional Taylor diagrams of surface pressure (hPa), 2-m water vapor (g kg^{-1}), and 10-m wind speed (ms^{-1}) are presented in **Figure 10** to summarize hindcast statistics. Each of the Taylor diagram distributions contains approximately 32,000 comparison points that use 1-min observations +/- 2 hours from the hindcast time, with the exception of the 10-day forecast (where there is no associated +2 hours).

Figure 8 shows example T_{2m} time series at a selection of forecast lead times (from Oliktok Point, AK during Fall 2017) and the corresponding summary Taylor diagram (for both Utqiaġvik and Oliktok Point, Alaska during Fall 2017 and 2018). In general, CAFS captures the temperature structures well from the initial time period through day-5 with correlations between 0.95 and 0.8.

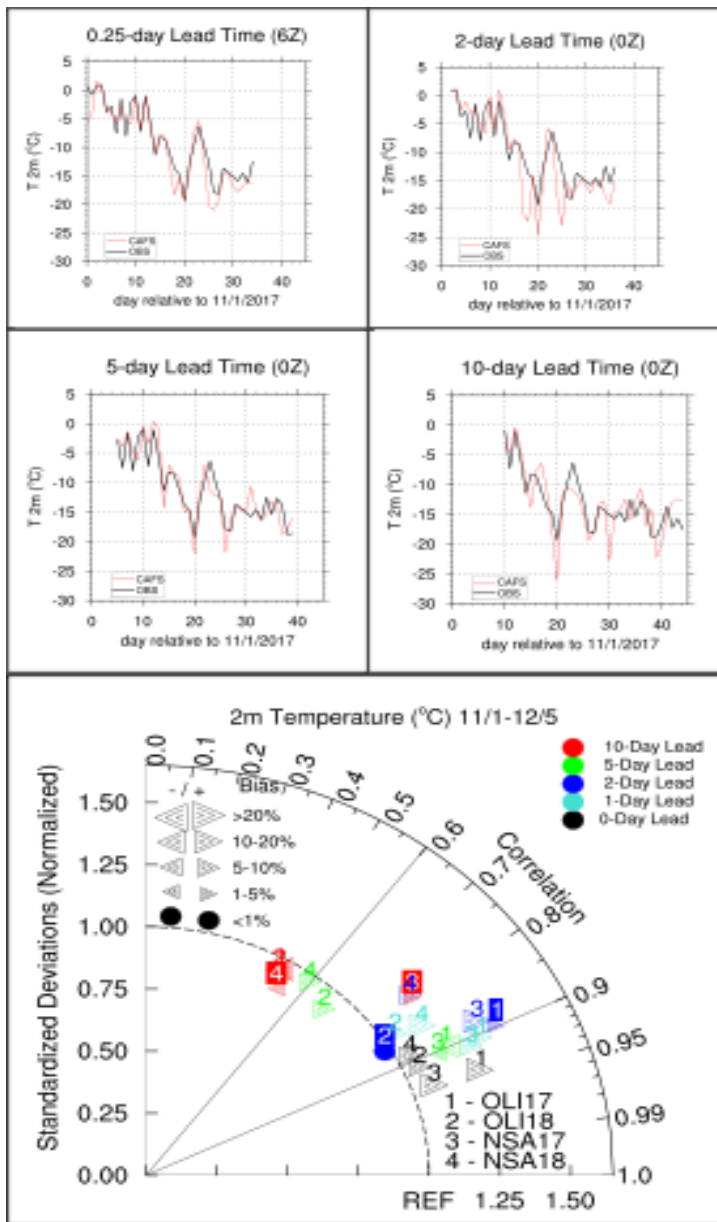Differences in the 10-day lead time time-series relate to the lower correlation (~0.5) shown in the Taylor diagram.

Figure 8. Time series of 2-m temperature (°C) from measurements at the DOE North Slope of Alaska observatory at Utqiaġvik, AK (black) and the CAFS model (red) for 0.25-, 2-, 5-, and 10-day lead times. Taylor diagram for T_{2m} at Utqiaġvik and Oliktok Point, Alaska.

Figure 9 depicts the downwelling longwave time series and summary Taylor diagrams as an example of a more process-oriented model skill metric. As in the previous discussion, forecast skill falls off after day-5 and both upwelling (not shown) and downwelling radiative terms have lower correlations between the 0- to 5-day lead times (~0.8 – 0.5). The model tends to get the

clear sky periods correct (lower values of LW$_d$) but there is an offset in the LW$_d$ during cloudy periods, which may correspond to modeled clouds being too thick or too warm. Hence, the standard deviations are larger than 1.0.
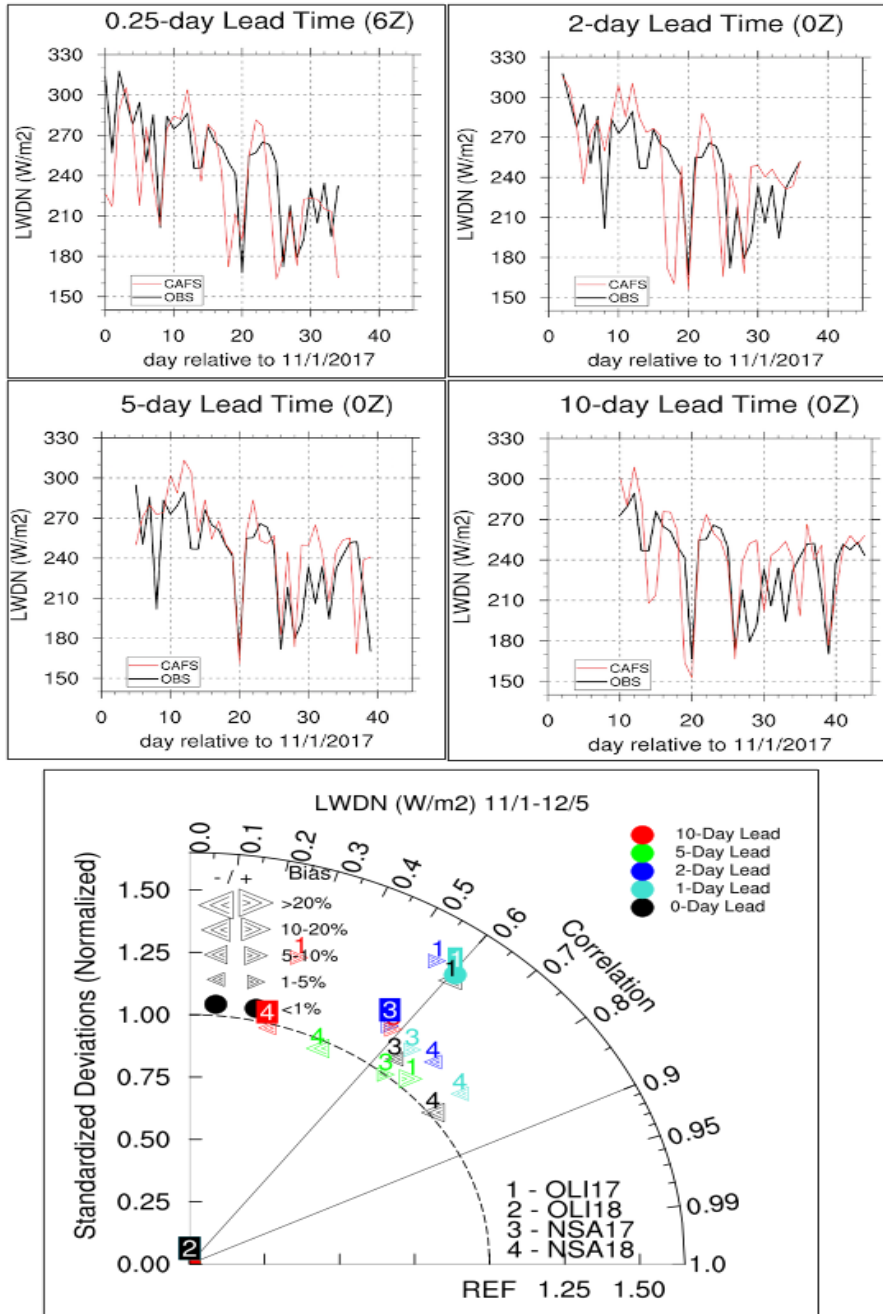

Figure 9. Time series of downwelling 926 longwave radiation (Wm-2) from measurements at the DOE North Slope of Alaska observatory at Utqiaġvik, AK (black) and the CAFS model (red) for 0.25-, 2-, 5-, and 10-day lead times during fall 2017. Taylor diagram for LWdn at Utqiaġvik (NSA) and Oliktok Point (OLI), Alaska.

Lastly, we present Taylor graph comparisons of the CAFS model surface meteorology quantities (**Figs. 10 and 11**). CAFS forecasts of surface pressure (**Fig. 10, top panel**) show excellent skill as we have already described in previous sections with modeled values consistently close to the observations. Note the tight clustering of points with a progression of correlation values along the normalized standard deviation arc equal to 1.0 from >0.99 toward lower values with increasing hindcast lead time, with 0.87 at day-10. The largest drop in skill is between the 5- and 10- day forecast from ~0.93 to 0.87, as would be expected with initial value NWP. This expected result is in part tied to using the GFS fields for initialization which are based on the GFS data assimilation process incorporating the available observations. expected result is in part tied to using the GFS fields for initialization which are based on the GFS data assimilation process incorporating the available observations.

Taylor diagrams for 2-m water vapor (Fig. 10 top bottom) and 10-m wind speed (Fig. 11) show a similar and expected reduction in correlation with longer forecast lead time with most 0- to 5-day values between 0.95 - 0.7. However, there are differences in the standard deviation (lower values for wind speeds and higher values for water vapor), and correlations are generally lower. Lower values for wind speed may indicate an inability for the model to represent small-scale

coastal atmospheric features and variations at these two sites. More detailed analysis is needed to clarify the reasons for the differences.
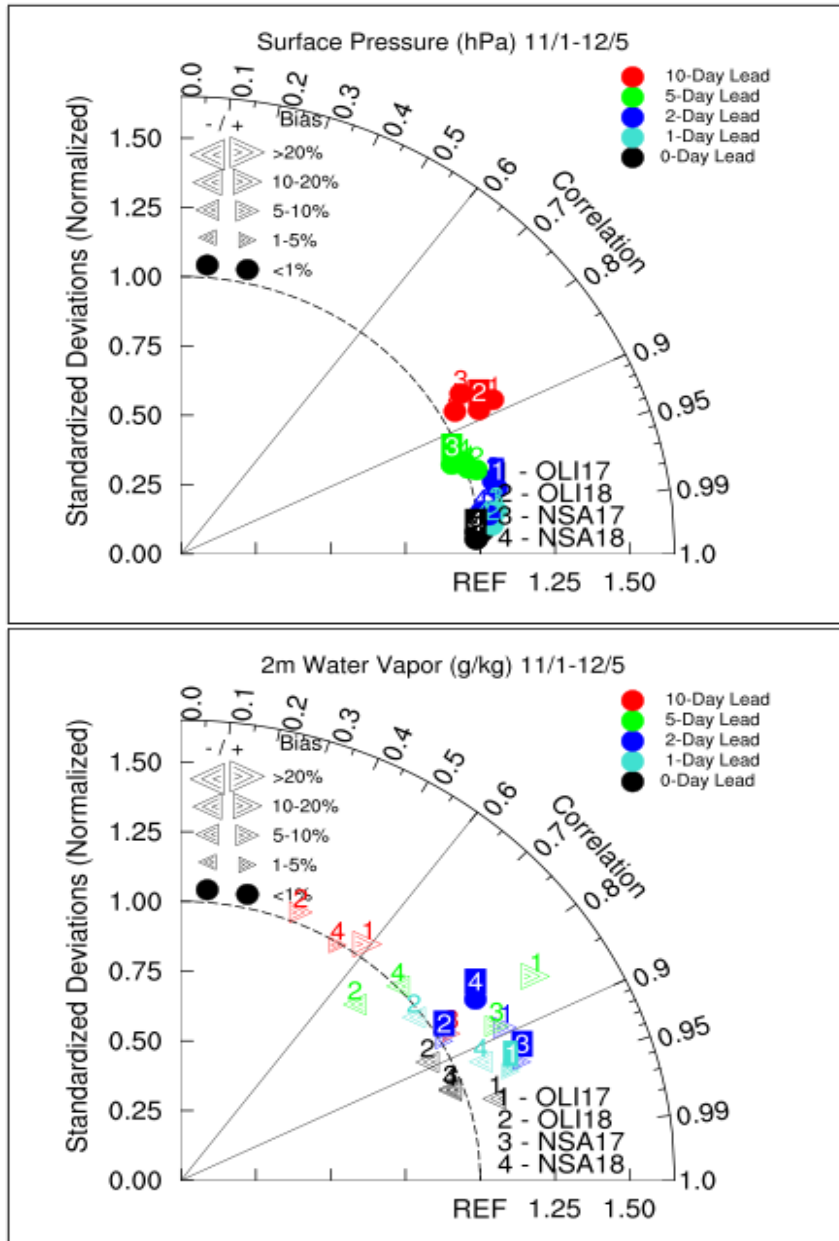

Figure 10. Taylor diagrams of CAFS hindcasts in comparisons to measurements of surface pressure (hPa) – top; 2-m water vapor -lower at Utqiaġvik (NSA) and Oliktok Point (OLI), Alaska for fall 2017 and 2018.
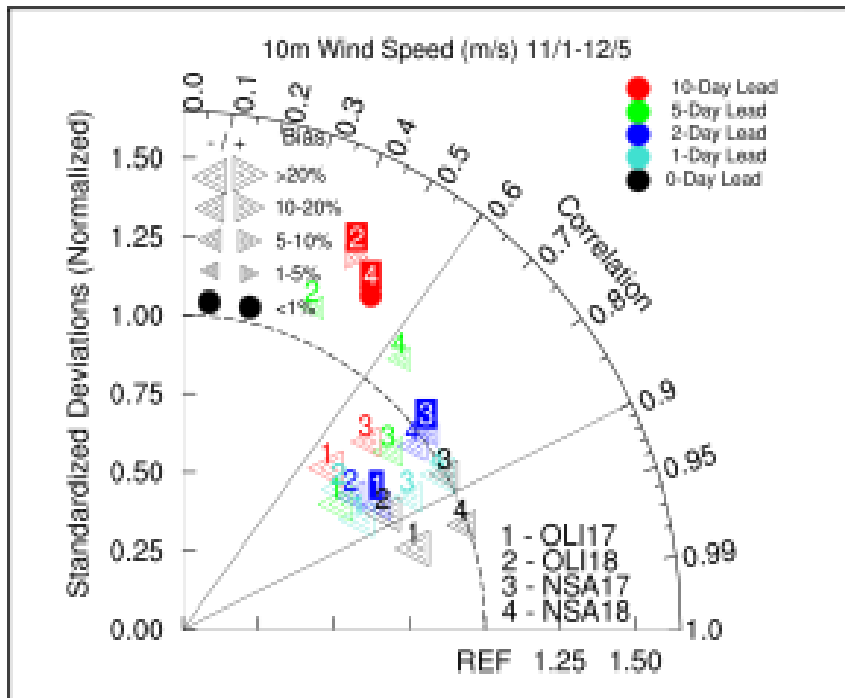
Figure 11. 10-m wind speed (ms⁻¹) – bottom row at Utqiaġvik (NSA) and Oliktok Point (OLI), Alaska for fall 2017 and 2018.

4. Applications

NOAA PSL has provided real time, experimental, daily, 0- to 10- day forecasts of Arctic weather and sea ice evolution information during freeze-up seasons from 2015-2018 and daily forecasts, year-round, starting in February 2018 via a publicly-accessible web interface. The CAFS suite of experimental ice, ocean and atmosphere model products have been used by a variety of stakeholders. At present, the National Weather Service-Alaska Region and the National Ice Center (NIC) incorporate CAFS ice, ocean and atmospheric products as guidance for their 1- to 5- day Arctic sea ice forecasts.

A variety of atmospheric, oceanic, sea ice and coupled forecast products, including animations of the 0- to 10- day forecasts, time-height cross sections, tracer and drift forecasts, meteograms (time series of meteorological quantities for a given location), and archived output files are posted daily to the website (https://www.esrl.noaa.gov/psd/forecasts/seaice/). The CAFS forecast page was developed with the NWS-AR Sea Ice Program (SIP) forecasters to ensure that the most useful guidance products, on the desired higher-resolution Alaska region grid, was easily accessible and in a usable format for inclusion into their GIS platform. Feedback from SIP forecasters has been invaluable given their experience with rapidly evolving Arctic marine

conditions, their interactions with polar operators in a real time environment, and their knowledge of real-world freeze-up processes.

CAFS forecast products have also been used to inform a variety of Arctic field campaigns over the last several years, including: the Navy Ice Experiment (ICEX; 2017, 2018) camps; ONR-sponsored Stratified Ocean Dynamics of the Arctic (SODA; 2018) and Sea State (2015; described in Section 3) research cruises held in the Beaufort and Chukchi Seas; the SIDFEx (described in **Section 3**) and MOSAiC campaigns (**Shupe et al. 2022**) during their year-long drift deployment (2019-2020); and NASA's ARCSIX (https://espo.nasa.gov/arcsix/content/ARCSIX) aircraft campaign (2024).

5. Conclusions and Future Work

Experimental, weather-scale (0- to 10- day) forecasts of Arctic weather and sea ice evolution have been provided to stakeholders using the Coupled Arctic Forecast System fully-coupled, ice-ocean-atmosphere, regional model since 2015. The current configuration of the coupled model includes the WRF atmospheric model, the POP2 dynamical ocean model, the CICE5 sea ice model, and the NCAR land and flux coupler models. The model is initialized each day with satellite sea ice concentration and sea surface temperature data. The CAFS sea ice forecasting project has two distinct objectives: the first is to provide model guidance for short-term forecasts of Arctic sea ice and weather; the second is to characterize process-level model deficiencies to improve model representation of key processes. NOTE: Can we support explanations of how to improve the representations, of what really is responsible for the errors reported and how to correct the model schemes.

Performance metrics, developed with observations from ship-based campaigns, buoys, satellite measurements, and Arctic land stations were shown to discuss hindcast skill relative to lead time and model biases. Model output was compared with rawinsonde profiles (temperature, pressure, relative humidity, and wind speed and direction), uCTD profiles (ocean temperature), surface meteorology quantities (sea level pressure, surface and 2-m temperature, 2-m water vapor, 10-m wind speed and direction, surface radiative fluxes), and satellite observations (sea ice extent and thickness) during the fall freeze-up season were discussed.

The comparisons between CAFS and observations show that, in general, CAFS is a skillful coupled model able to provide useful forecast guidance products to sea ice forecasters. CAFS captures synoptic-scale temporal/spatial variability, as well as, near-surface weather with only small errors in SLP and only slightly larger errors in temperature and wind speed. CAFS shows skill in creating, moving and melting sea ice, with forecast skill beating persistence after 1 day with correlations ~0.9-0.95 until day-2 and ~0.7-0.8 until ~day 5.5-6.5, as compared with satellite and buoy data. Temperature errors throughout the atmosphere column were < ~0.5 ° C,

up to 36-h lead times and errors in the ocean surface temperatures < -0.25 ° C, across all lead times. Exceptions included larger deviations evident at major temperature transition zones: the tropopause (+/-1 ° C); the boundary layer air temperatures (+2-3 ° C after 180-h); and the ocean halocline (~-1.5 ° C).

 As expected, CAFS predictability skill significantly decreases after day-7, especially in the atmospheric boundary layer (925 mb and below), where temperature bias error in the lowest 2-km grows rapidly from roughly +1 ° C at 180-h to +2.5 – 3.25 ° C at 240-h. This warm bias indicates that the model is unable to maintain observed boundary layer stratification and rapidly evolves into a less-stable state.   Process studies to evaluate if this bias is due to the misrepresentation of air-surface interaction processes, issues with the model coupling, representation of clouds, or to other boundary layer and surface processes and parameterizations continue to be evaluated.

Modeled quantities linked to strongly coupled processes such as radiative fluxes, boundary layer stability, and cloud structure exhibited lower correlations, especially in the Marginal Ice Zone regions where the fine-scale ice features are not captured in the satellite initialization fields. In general, CAFS-simulated surface radiation and turbulent fluxes are in reasonable agreement with point observations after accounting for discrepancies expected from errors in ice edge location and SIC, and for errors in the surface-layer stability. Validations using several of the data sets show that initial conditions of ice concentrations and even 2-m temperature are often in error, especially near the ice edge, leading to errors in a number of air-surface interaction processes.

When compared to point measurements, CAFS exhibited some success in predicting sky conditions, cloud structure, and phase, though some periods showed error in correctly simulating the timing and lifetime. Improvements were made after initializing the cloud hydrometeor mass and number with fields from the first day of the previous day's hindcast to reduce spin-up time of cloud microphysical fields, as well as the use of a double-moment microphysical scheme implemented in WRF.

 Ultimately, understanding coupled processes is a major driver behind the CAFS development process, and we will continue to evaluate surface energy exchange against observations of ice, ocean, and atmosphere through comparison with additional data sets. Measurements obtained over an annual cycle from the MOSAiC Drifting Observatory will allow us to further assess and validate the model, address outstanding process questions, and intercompare with other international coupled forecast models. Advancing our understanding of coupled forecast skill and error is an important component of NOAA's continued development of the global Unified Forecast System. In this capacity the regional CAFS model has informed discussions with the

UFS teams to understand processes unique to the Arctic environment and their cumulative influence on sea ice forecasting.

The Arctic is a unique environment and the heterogeneity of the surface conditions drive variability. Arctic cyclones, which transport heat and momentum from lower latitudes and impact navigation though atmosphere, ocean, and sea ice conditions are impacted by small-scale variability such as meso-scale cyclones, tropopause potential vorticity anomalies, and land-sea and ice-ocean temperature and moisture contrasts. Ongoing studies with CAFS reveal that global models with regional, higher-resolution configurations in the Arctic and parameterizations adequate and specific to simulate Arctic boundary layers and mixed-phase cloud systems should be considered for optimized forecast skill.

6. Acknowledgments

Washington U.S. International Buoy Programme; and, the RASM development team especially, Andrew Roberts, Wieslaw Maslowski, and John Cassano.

7. Works Cited

Adams, J. D., and G. K. Silber, 2017: 2015 vessel activity in the Arctic. *NOAA Technical Memorandum NMFS-OPR-57*, doi:10.7289/V5/TM-NMFS-OPR-57.

Allard, R. A., S. L. Farrell, D. A. Hebert, W. F. Johnston, L. Li, N. T. Kurtz, M. W. Phelps, P. G. Posey, R. Tilling, A. Ridout, and A. J. Wallcraft, 2018: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system. *Adv. in Space Res.*, **62,** 1265-1280.

Atmospheric Radiation Measurement (ARM) Climate Research Facility, updated hourly. Quality-Controlled radiation product (QCRAD1LONG). 2017-11-01 to 2018-12-05, 71.32 N, 156.52 W and 70.495 N, 149.886 W: North Slope Alaska (NSA) and ARM Mobile Facility (OLI) Oliktok Point, Alaska; AMF3 processed data, 2018a: Compiled by C. Long, Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed at http://dx.doi.org/10.5439/1227214.

Atmospheric Radiation Measurement (ARM) Climate Research Facility, updated hourly. Surface Meteorological Instrumentation (MET). 2017-11-01 to 2018-12-05, 70.495 N 149.886 W: ARM Mobile Facility (OLI) Oliktok Point, Alaska; AMF3 (M1) and 71.323 N 156.609 W: North Slope Alaska (NSA) Central Facility, Barrow AK (C1), 2018b: Compiled by D. Holdridge and J.

Kyrouac. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed at http://dx.doi.org/10.5439/1025220.

Barnhart, K., C. Miller, and I. Overeem, 2016: Mapping the future expansion of Arctic open water. *Nature Clim. Change,* **6,** 280–285. https://doi.org/10.1038/nclimate2848.

Cassano, J. J., and Coauthors, 2017: Development of the Regional Arctic System Model (RASM): Near-surface atmospheric climate sensitivity. *J. Clim.,* 5729-5753, doi:10.1175/JCLI-D-15-0775.1.

Collins, W. D., and Coauthors, 2006: The formulation and atmospheric simulation of the Community Atmosphere Model: CAM3. *J. Climate*, **19 ,** 2144–2161.

de Boer, G., and Coauthors, 2018: A bird's eye view. *Bull. Amer. Meteor. Soc.,* 1197-1212. doi: 10-1175/BAMS-D17-01561.

Fang, Z., P.T. Freeman, C.B. Field, and K.J. Mach, 2018: Reduced Sea Ice Protection Period Increases Strom Exposure in Kivalina, Alaska. *Arctic Science*, **4**, 525-537, https://doi.org/10.1139/as-2017-0024.

Farrell, S., 2019. Pers. Comm.

Goessling, H. F., and Coauthors, 2020: The Sea Ice Drift Forecast Experiment: Rationale, Implementation, Interim Results, and Outlook. In preparation for submission to The Cryosphere.

Grell, G. A., J. Dudhia, J., and D. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5) (No. NCAR/TN-398+STR). University Corporation for Atmospheric Research. doi:10.5065/D60Z716B.

Guest, P., P. O. G. Persson, S. Wang, M. Jordan, Y. Jin, B. Blomquist, and C. Fairall, 2018: Low-level baroclinic jets over the new Arctic Ocean. *J. Geophys. Res.-Oceans*, Sea State special issue, **123**, 4074–4091, https://doi.org/10.1002/2018JC013778.

Hamman, J. J., B. Nijssen, T. J. Bohn, D. R. Gergel, and Y. Mao, 2018: The Variable Infiltration Capacity model version 5 (VIC-5): infrastructure improvements for new applications and reproducibility, Geosci. Model Dev., 11, 3481-3496, https://doi.org/10.5194/gmd-11-3481-2018.

Hong, S. Y., 2010: A new stable boundary layer mixing scheme and its impact on the simulated East Asian summer monsoon, Quart. J. Roy. Meteor. Soc., 136, 1481–1496, doi:10.1002/Qj.665.

Hu, X.-M., P. M. Klein, and M. Xue, 2013: Evaluation of the updated YSU planetary boundary layer scheme within WRF for wind resource and air quality assessments. J. Geophys. Res. Atmos., 118, 10,490-10,505, doi:10.1002/jgrd.50823.

Hunke, E. C., W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. M. Elliott, 2013: CICE: The Los Alamos sea ice model, documentation and software, Version 5.0., *Los Alamos National Laboratory Tech. Rep*. LA-CC-06-012.

IPCC, 2019: Summary for Policymakers. In: IPCC Special Report on the Ocean and Crysophere in a Changing Climate. Portner, O.-H., D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Minterbeck, M. Nicolai, A. Okem, J. Petzhold. B. Rama, N. Weyer, Eds.

Maslowski W., J. Clement Kinney, M. Higgins, and A. Roberts, 2012: The future of arctic sea ice. *Annual Review of Earth and Planetary Sciences*, **40**, 625-654.

Maturi, E., A. Harris, J. Mittaz, J. Sapper, G. Wick, X. Zhu, P. Dash, and P. Koner, 2017: A new high-resolution Sea Surface Temperature blended analysis. *Bull. Amer. Meteor. Soc.,* **98**, 1015–1026, https://doi.org/10.1175/BAMS-D-15-00002.1.

Meier, W. N., Petty, A., Hendricks, S., Kaleschke, L., Divine, D., Farrell, S., Gerland, S., Perovich, D., Ricker, R., Tian-Kunze, X., Webster, M., 2023: NOAA Arctic Report Card 2023: Sea Ice. NOAA technical report OAR ARC 23-06, https://doi.org/10.25923/f5t4-b865.

Mohammadi-Aragh, M., H. F. Goessling, M. Losch, N. Hutter, and T. Jung, 2018: Predictability of Arctic sea ice on weather time scales. Scientific reports, 8.

Moon, T. A., M. L. Druckenmiller, and R. L. Thoman, Eds., 2021: *Arctic Report Card 2021*, doi:10.25923/5s0f-5163.

Morrison, H., and Coauthors, 2009: Intercomparison of model simulations of mixed-phase clouds observed during the ARM Mixed-Phase Arctic Cloud Experiment. II: Multilayer cloud. *Q. J. Roy. Meteorol. Soc.*, **135**, 1003-1019, doi:10.1002/qj.415.

Osborne, E., J. Richter-Menge, and M. Jeffries, Eds., 2018: Arctic Report Card 2018. https://www.arctic.noaa.gov/Report-Card.

Overeem, I., R. S. Anderson, C. W. Wobus, G. D. Clow, F. E. Urban, and N. Matell, 2011: Sea ice loss enhances wave action at the Arctic coast. *Geophys. Res. Letts*., **38**, L17503, doi:10.1029/2011GL048681.

Perovich, D., J. Richter-Menge, B. Elder, T. Arbetter, K. Claffey, and C. Polashenski, 2015: Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice. http://imb.erdc.dren.mil.

Persson, P. O. G., B. Blomquist, P. Guest, S. Stammerjohn, C. Fairall, L. Rainville, B. Lund, S. Ackley, and J. Thomson, 2018: Shipboard observations of the meteorology and near-surface environment during autumn freezeup in the Beaufort/Chukchi Seas. *J. Geophys. Res.-Oceans*, **123**, 4930–4969. https://doi.org/10.1029/2018JC013786.

Posey, P. G., E. J. Metzger, A. J. Wallcraft, D. A. Hebert, R. A. Allard, O. M. Smedstad, M. W. Phelps, F. Fetterer, J. S. Stewart, W. N. Meier, and S. R. Helfrich, 2015: Improving Arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast system. *The Cryosphere*, **9**, 1735–1745, 2015, doi:10.5194/tc-9-1735-2015.

Ralph, F. M., and Coauthors, 2013: The emergence of weather-related test beds linking research and forecasting operations. *Bull. Amer. Meteor. Soc.*, **94**, 1187–1211.

Richter-Menge, J., M. L. Druckenmiller, and M. Jeffries, Eds., 2019: Arctic Report Card 2019, https://www.arctic.noaa.gov/Report-Card.

Roberts, A., A. P. Craig, W. Maslowski, R. Osinski, A. DuVivier, M. Hughes, B. Nijssen, J. Cassano, and M. Brunke, 2014: Simulating transient ice-ocean Ekman transport in the Regional Arctic System Model and Community Earth System Model, *Ann. Glaciol.*, **69**, doi:10.3189/2015AoG69A760.

Rogers, Hu. Shen, Ha. Shen, V. Squire, S. Stammerjohn, J. Stopa, M. Smith, P. Sutherland, and P. Wadhams, 2018: Overview of the Arctic Sea State and Boundary Layer Physics Program. *J.*

Geophys. Res.-Oceans, **123**,
https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2018JC013766.

Sakov, P., F. Counillon, L. Bertino, K. A. Lisæter, P. R. Oke, and A. Korablev, 2012: TOPAZ4:
an ocean-sea ice data assimilation system for the North Atlantic and Arctic, *Ocean Sci.*, **8**, 633–
656, https://doi.org/10.5194/os-8-633-2012.

Serreze, M. C., and J. C. Stroeve, 2015: Arctic sea ice trends, variability and implications for
seasonal ice forecasting. *Phil Trans A* 373, 20140159 doi:10.1098/rsta.2014.0159.

Shupe, M. D., and coauthors, 2022: Overview of the MOSAiC expedition -- Atmosphere.
Elementa: Science of the Anthropocene 10(1). DOI:10.1525/elementa.2021.00060

Skamarock, W., J. B. Klemp, and J. Dudhia, 2008: A description of the advanced research WRF
version 3. NCAR Technical Note NCAR/TN-475 + STR.

Smith, R., and Coauthors, 2010: The parallel ocean program (POP) reference manual: ocean
component of the community climate system model (CCSM) and community earth system model
(CESM). *LAUR-01853* 141 (2010): 1-140.

Smith, L. C., and S. R. Stephenson, 2013: New Trans-Arctic shipping routes navigable by
midcentury. *PNAS* E1191–E1195 doi: 10.1073/pnas.1214212110.

Spreen, G., L. Kaleschke, and G. Heygster, 2008: Sea ice remote sensing using AMSR-E 89
1448 GHz channels *J. Geophys. Res.*,**113**, C02S03, doi:10.1029/2005JC003384.

SWIPA, 2017: Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2017. Arctic Monitoring
and Assessment Programme (AMAP), Oslo, Norway. 269 pp.

Taylor, K., 2001: Summarizing multiple aspects of model performance in a single diagram. *J.
Geophys. Res.*, **106**, doi:10.1029/2000JD900719.

Thoman, R. and J. E. Walsh, 2019: Alaska's changing environment: documenting Alaska's
physical and biological changes through observations. H. R. McFarland, Ed. International Arctic
Research Center, University of Alaska Fairbanks. DOI:10.13140/RG.2.2.24481.15209.

Thoman, R.L., J. Richter-Menge, and M.L. Druckenmiller, Eds., 2020: Arctic Report Card 2020,
https://doi.org/10.25923/mn5p-t549.

Thomson, J., and Coauthors, 2016: Emerging trends in the sea state of the Beaufort and Chukchi Seas. *J. Ocean Modelling*, DOI: 10.1016/j.ocemod.2016.02.009.

Thomson, J., and Coauthors, 2018: Overview of the Arctic Sea State and Boundary Layer Physics Program. *J. Geophys. Res.-Oceans*, **123**, https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2018JC013766.

Uttal, T., and Coauthors, 2015: International Arctic Systems for Observing the Atmosphere (IASOA): An International Polar Year Legacy Consortium, *Bull. Amer. Meteorol. Soc.*, **97**, 1033-1056, https://doi.org/10.1175/BAMS-D-14-00145.1.

Wilson, A. B., D. H. Bromwich, and K. M.Hines, 2011: Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain: Surface and upper air analysis, *J. Geophys. Res.*, **116**, D11112, doi:10.1029/2010JD015013.

WMO WWRP, 2016: WWRP Polar Prediction Project, Implementation Plan for the Year of Polar Prediction (YOPP). WWRP/PPP No. 4., http://polarprediction.net.

Xie, J., L. Bertino, F. Counillon, K. A. Lisæter, and P. Sakov, 2017: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013. *Ocean Sci.,* doi:10.5194/os-13-123-2017.