Title	Clouds and Radiation Processes in Regional Climate Models Evaluated Using Observations Over the Ice-free Arctic Ocean
Authors	Jun Inoue, Kazutoshi Sato, Annette Rinke, John J. Cassano, Xavier Fettweis, Günther Heinemann, Heidrun Matthes, Andrew Orr, Tony Phillips, Mark Seefeldt, Amy Solomon, Stuart Webster
Citation	Journal of Geophysical Research: Atmospheres, 126(1), 1-25, 2020
Issue Date	2020-12-1
Type	Journal Article
URL	https://doi.org/10.1029/2020JD033904
Right	
Textversion	publisher

JGR Atmospheres

RESEARCH ARTICLE
10.1029/2020JD033904

Key Points:

- Discrepancies remain in surface energy budget (SEB) and cloud properties despite constrained boundary conditions
- Models underestimate occurrence of unstable stratification with low-level water clouds
- Biases in hydrometeor partitioning essentially contribute to limited skill in SEB simulation

Supporting Information:

- Supporting Information S1

Correspondence to:
J. Inoue,
inoue.jun@nipr.ac.jp

Citation:
Inoue, J., Sato, K., Rinke, A., Cassano, J. J., Fettweis, X., Heinemann, G., et al. (2021). Clouds and radiation processes in regional climate models evaluated using observations over the ice-free Arctic Ocean. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD033904. https://doi.org/10.1029/2020JD033904

Received 16 SEP 2020
Accepted 19 NOV 2020

Clouds and Radiation Processes in Regional Climate Models Evaluated Using Observations Over the Ice-free Arctic Ocean

Jun Inoue[1], Kazutoshi Sato[2], Annette Rinke[3], John J. Cassano[4,5], Xavier Fettweis[6], Günther Heinemann[7], Heidrun Matthes[3], Andrew Orr[8], Tony Phillips[8], Mark Seefeldt[4], Amy Solomon[4,9], and Stuart Webster[10]

[1]Meteorology and Glaciology Group, National Institute of Polar Research, Tachikawa, Japan, [2]School of Earth, Energy and Environmental Engineering, Kitami Institute of Technology, Kitami, Japan, [3]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany, [4]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [5]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO, USA, [6]Department of Geography, University of Liége, Liége, Belgium, [7]Department of Environmental Meteorology, University of Trier, Trier, Germany, [8]British Antarctic Survey, National Environmental Research Council, Cambridge, UK, [9]Physical Sciences Division, NOAA Earth System Research Laboratory, Boulder, CO, USA, [10]Met Office, Exeter, UK

Abstract The presence of clouds in the Arctic regulates the surface energy budget (SEB) over the sea-ice surface and the ice-free ocean. Following several previous field campaigns, the cloud-radiation relationship, including cloud vertical structure and phase, has been elucidated; however, modeling of this relationship has matured slowly. In recognition of the recent decline in the Arctic sea-ice extent, representation of the cloud system in numerical models should consider the effects of areas covered by sea ice and ice-free areas. Using an in situ stationary meteorological observation data set obtained over the ice-free Arctic Ocean by the Japanese Research Vessel *Mirai* (September 2014), coordinated evaluation of six regional climate models (RCMs) with nine model runs was performed by focusing on clouds and the SEB. The most remarkable findings were as follows: (1) reduced occurrence of unstable stratification with low-level cloud water in all models in comparison to the observations, (2) significant differences in cloud water representations between single- and double-moment cloud schemes, (3) extensive differences in partitioning of hydrometeors including solid/liquid precipitation, and (4) pronounced lower-tropospheric air temperature biases. These issues are considered as the main sources of SEB uncertainty over ice-free areas of the Arctic Ocean. The results from a coupled RCM imply that the SEB is constrained by both the atmosphere and the ocean (and sea ice) with considerable feedback. Coordinated improvement of both stand-alone atmospheric and coupled RCMs would promote a more comprehensive and improved understanding of the Arctic air-ice-sea coupled system.

1. Introduction

Clouds in the Arctic have a multifaceted role within the Arctic climate system. Solid precipitation on sea ice acts as a thermal modulator by changing the surface albedo, the thermal conductivity, and the roughness length for momentum at the top of the sea ice, whereas precipitation into open water provides freshwater input to the ocean (Vihma et al., 2016). Regarding radiation from clouds in the Arctic, the surface energy budget (SEB) is constrained seasonally by cloud properties (Intrieri et al., 2002; Shupe & Intrieri, 2004), while Arctic clouds are influenced by components of the SEB, in particular the sensible heat flux (SHF), latent heat flux (LHF), and horizontal heat and moisture transports. Although appropriate representation of Arctic clouds in numerical models has remained a challenge over the previous three decades (e.g., Curry et al., 1996), several issues regarding their representation have been revealed through use of new satellite products and general circulation models. For example, these issues concern underestimation of the cloud-top albedo, which causes positive downward shortwave radiation (SWD) model biases (English et al., 2014), negative cloud liquid water model biases and surface albedo adjustment to achieve a credible Arctic sea ice mean thickness (Kay et al., 2016), and significant intermodel differences in low-level cloud amount associated with lower tropospheric stability and cloud microphysical schemes (Taylor et al., 2019).

Field campaigns in the Arctic have provided great opportunities for evaluating numerical models, especially high-resolution regional climate models (RCMs). Data from the Surface Heat Budget of the Arctic Ocean (SHEBA) project (Uttal et al., 2002), which was a year-long field program over the ice-covered Beaufort Sea during 1997–1998, were used in the Arctic Regional Climate Model Intercomparison Project (ARCMIP; Curry & Lynch, 2002). The ARCMIP contributed to identifying how RCMs are limited both in reproducing the observed SEB over sea ice and in connecting their cloud microphysics and radiation processes (Inoue et al., 2006; Rinke et al., 2006; Wyser et al., 2008). The model errors in the sea-ice SEB affects lower boundary layer variability and is influenced by clouds and radiation (Rinke et al., 2006). The accumulated error of the SEB can affect the onset of melting of sea ice (Inoue et al., 2006). The reproducibility of the cloud phase (i.e., liquid or solid) is one of the key elements for better representation of the interactions between cloud and radiation (Wyser et al., 2008). More than 15 years after the ARCMIP, the Arctic Coordinated Regional Downscaling Experiment (Arctic CORDEX), which formed part of a wider CORDEX program (Giorgi et al., 2009), focused on the new state of the Arctic since the SHEBA campaign. Using data collected in the Arctic by a Swedish icebreaker during summer and autumn 2014, Sedlar et al. (2020) evaluated 10 model runs from six different RCMs using different cloud parameterizations and other settings. It was concluded that the distributions and errors in the representations of clouds and radiation were similar to those reported following the ARCMIP studies, although understanding of the processes and model development had only improved marginally. Despite improvement in model resolutions and physical parameterizations based on the latest field campaigns, determination of an essential logic with which to resolve the SEB disagreements in association with clouds and radiation in RCMs remains unclear. The use of the same lateral boundary conditions makes model evaluation effective within the CORDEX framework. However, in the Arctic, surface boundary conditions also strongly affect the lower troposphere because each model treats ice thickness, snow depth, skin temperature, and albedo differently. Uncertainties in near-surface meteorological parameters and surface heat fluxes that originate from the sea-ice surface could amplify and/or compensate each other, making it difficult to evaluate the relationship between the SEB and clouds/radiation. Therefore, a more straightforward setup for intercomparison of RCMs is required to elucidate the fundamental processes that could substantially improve RCM performance.

The new state of the Arctic Ocean is characterized by the distribution of ice-covered areas and ice-free areas. Shipping along certain coastal routes has increased following the reduction in ice cover, for example, the Northern Sea Route (NSR). In navigating the NSR, skillful forecasts of both weather and sea ice are indispensable. However, the predictability of the free-drift of sea ice depends heavily on the predictability of the atmosphere (Inoue et al., 2015; Ono et al., 2016), and uncertainty regarding surface winds influences the skill of wave forecasts (Nose et al., 2018). As extreme wind events associated with cyclones and the periphery of anticyclonic systems can rapidly change the ocean/sea-ice states, improvements in high-resolution modeling and skillful atmospheric forecasts would represent fundamental advancements (Inoue, 2020). One of the primary scientific targets of the Arctic CORDEX was low-level cloud because visibility over the NSR is often reduced owing to low stratus cloud (fog). Thus, improvement in the representation of low-level clouds in RCMs, which has long remained a challenge, would also be an essential contribution in this field.

Sea surface temperature (SST) near the marginal ice zone is generally low owing to the input of meltwater; however, during late autumn, SST can be sufficiently high relative to the cold air mass advected from perennial sea ice, resulting in substantial air-sea heat transfer and the formation of convective cloud (Inoue & Hori, 2011). In comparison with the SHEBA era, cloud base height during autumn has increased owing to the recent expansion of the ice-free area (Sato et al., 2012). Increased cloud base height during autumn has two principal effects on the sea surface energy budget (SSEB). The first is a reduction of downward longwave radiation (LWD) due to the decrease in temperature at the cloud bottom, which accelerates sea surface cooling. A change of cloud type from persistent stratus cloud to stratocumulus with reduced cover could also contribute to reduction in the magnitude of LWD. The second effect is the opposite of the first; LWD at the surface is increased because the increase in ice-free area results in more intense upward SHF and LHF (i.e., air mass modification), even if the cloud bottom height were elevated. This contradictory issue has not been comprehensively evaluated using observations or models. In addition, it could be considered that delayed onset of freezing due to the declining extent of sea ice would change the role of low-level clouds in the SSEB. Therefore, evaluation of the SSEB linked to clouds and radiation would be desired. For example, feedback processes between cloud/radiation and the SSEB could be better identified if the effect

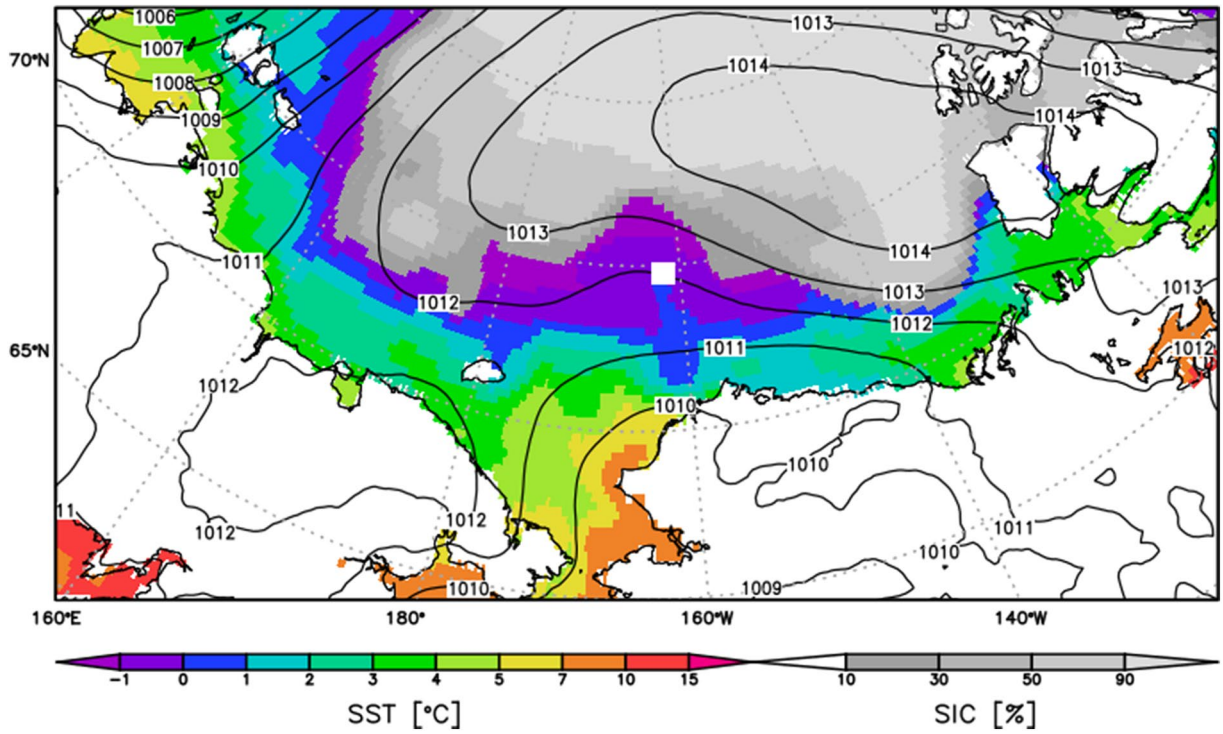


Figure 1. Position of the stationary point of the Research Vessel *Mirai* at 74.75°N, 162.00°W (white square). Monthly mean sea level pressure (contours), ice concentration (gray shading), and sea surface temperature (colored shading) for September 2014, as indicated by ERA5 reanalysis.

of variability of the sea-ice surface is excluded. The thermodynamic response to the SSEB in the freezing season connects directly to sea-ice formation as a coupled system, which represents a developing field for evaluation of coupled RCMs.

The Japanese ice-strengthened Research Vessel (RV) *Mirai* undertook an Arctic research cruise in September 2014. This cruise was characterized by stationary point radiosonde and conductivity-temperature-depth observations conducted during September 7–25 at 74.75°N 162.00°W (Figure 1; Inoue, 2014; Inoue et al., 2018). The other meteorological observations performed around the point included radiation measurements and ceilometer observations. This stationary point observation data set is considered suitable for investigating temporal changes in the atmospheric field during the transitional period from autumn to winter. It should be noted that formation and advection of sea ice were not observed during this period. These data have previously been used to investigate the responses of ocean turbulent mixing and biogeochemical processes to surface wind forcing over the ice-free area have been investigated (Kawaguchi et al., 2016; Nishino et al., 2019).

There are two advantages to evaluation of RCMs for the case of the ice-free situation using the data obtained during this cruise. The first is that there was no change in the location of the observations. The middle of September is a period of transition from the melting season to the freezing season, which can be seen as the change of sign of the SSEB. Furthermore, stationary point observations are useful for analysis of seasonal change not only of the surface and lower troposphere but also of the upper troposphere and lower stratosphere where large positive potential vorticity anomalies exist. Skillful representation of potential vorticity anomalies in numerical models improves the positioning and strength of Arctic cyclones (Inoue et al., 2013; Yamazaki et al., 2015), which can result in better representation of the SSEB through improved modeling of surface winds and air temperature. The second advantage is the simpler bottom boundary condition (i.e., the open ocean) of the RCMs that does not need to consider the response to sea-ice surface variability through different parameterizations among the RCMs (e.g., ice thickness, snow on the ice, albedo, and melt pond fraction). The presence of sea ice is a natural component of the Arctic climate system. However, from the perspective of evaluation of RCMs, the differences in the treatment of sea ice among the various

Table 2
Mean Bias Errors (MBEs) and Root Mean Square Errors (RMSEs) for the Surface Pressure (P; hPa), Air Temperature (T; °C), and Wind Speed (U; m/s) Among the RCMs

Model	P_{MBE}	P_{RMSE}	T_{MBE}	T_{RMSE}	U_{MBE}	U_{RMSE}
CAFS-bsl	−0.1	1.1	−1.2	2.0	−0.2	2.1
CAFS-ini	−1.2	1.6	−2.2	2.7	−0.3	1.7
CCLMi	0.7	0.9	−0.1	1.0	0.0	1.2
CCLM5	1.1	1.2	−0.2	0.8	0.1	1.3
HIRHAM-v1	0.0	0.4	−0.1	0.9	0.0	0.9
HIRHAM-v2	0.0	0.4	−0.1	0.9	0.1	1.0
MAR	−1.2	1.4	−0.4	1.3	−1.4	2.0
METUM	0.0	0.4	−0.1	0.8	−0.2	1.1
WRF	−0.3	1.4	−2.1	2.5	−0.8	1.6
ERA5	0.2	0.3	−0.1	0.6	−0.2	1.1

Abbreviation: RCM, regional climate model.

RCMs mask the pure interaction between cloud/radiation and the SSEB, making evaluation of the cloud microphysics and radiation processes difficult. For example, among the RCMs considered in the Arctic CORDEX, sea-ice concentration (SIC) was prescribed by a satellite product (one model), ERA-Interim reanalysis data (four models), and a coupled model (one model). As recent evaluation of several SIC products has revealed that SIC in the ERA-Interim data tends to be underestimated (Nose et al., 2020), any resulting errors in the atmospheric parameters modeled by the RCMs associated with this might result in a poor comparison with the in situ observations (e.g., near-surface air temperature and its vertical structure), which are influenced by the real SIC. In fact, in the case of the Arctic CORDEX over the ice, the skin temperature deviates with a warm bias of up to 2 K among the various RCMs (Sedlar et al., 2020). The same type of problem would be associated with the differences in the parameterizations of albedo and skin temperature of sea ice.

As outlined above, the use of stationary point observation data in an ice-free situation offers new insight into the fundamental understanding of the relationship between the SSEB and cloud/radiation for evaluation of RCMs. To investigate the temporal evolution of the SSEB associated with cloud characteristics, including radiation, evaluation of RCMs at a stationary point in an ice-free situation represents an ideal setup and a new approach. Therefore, this study evaluated the same RCMs as investigated in Sedlar et al. (2020), and used the observation data set collected over the ice-free Arctic Ocean at the stationary point for comparison with the results derived in an ice-covered situation.

2. Models, Experimental Setup, and Observations

2.1. RCMs and Model Experimental Setup

Detailed descriptions of the models, including relevant references, are summarized in the tables in Sedlar et al. (2020). This study uses the same model outputs (except for one model) as Sedlar et al. (2020), but the evaluation location is different. Here, only the main differences among the RCMs relevant to this study are highlighted. Five atmospheric RCMs (i.e., CCLM, HIRHAM5, MARv3.11, MetUM, and WRF-CU) and one coupled RCM (CAFS) were used in this study. It should be noted that the large-scale dynamic constraints and other setups are very different among the RCMs, as summarized in Table 1. CCLM was forced by two different boundary conditions: ERA-Interim (Dee et al., 2011; hereafter, CCLMi) and ERA5 (Hersbach et al., 2020; hereafter, CCLM5). HIRHAM5 produced two model outputs: one used the cloud microphysics scheme by Sundqvist et al. (1989) (hereafter, HIRHAM-v1), while the other used the scheme by Tompkins (2002) (hereafter, HIRHAM-v2). CAFS has also two model outputs: one is the baseline simulation (hereafter, CAFS-bsl) while the other is tuned to the autumn freeze-up (hereafter, CAFS-ini) by increasing the ice nucleating particles from 0.16 (CAFS-bsl) to 1.0 L^{-1} (CAFS-ini). The upper troposphere/stratosphere is nudged in the CAFS simulations, suggesting that the lower troposphere/ocean/sea ice freely evolves over the simulation, and by the period of this evaluation, can be significantly different from the other RCMs due to accumulating errors since the initialization of the simulation on July 1. As the CAFS incorporates its own coupled system, the observation point in the model was partially covered by sea ice (SIC: 10%–80%), suggesting that the SST might be expected to be lower than the observations and that the air temperature would have a cold bias near the surface (Figure S1 and Table 2). The modeled sea ice would also be expected to influence other SSEB components in the CAFS; however, the results from the CAFS are included in this paper because the cloud microphysics scheme it uses (Morrison et al., 2009) is the same as used in the WRF-CU (hereafter, WRF). The MARv3.9 model output used in Sedlar et al. (2020) had a bug in the spectral constrained for the lower stratosphere, impacting the MAR results for a large integration domain. Therefore, the output from MARv3.11 is used in this study. The main improvements between MARv3.10 (Tedesco & Fettweis, 2020) and MARv3.11 (hereafter, MAR) are the inclusion of a blowing snow module (not used here) and a correction of the bug.

Table 1
Names, Abbreviations, Forcing, Vertical and Horizontal Resolutions, Large-Scale Dynamic Constraints, Cloud Microphysics Scheme of the Regional Climate Models Used in the Study, and Model References

Model (Inst.)	Name	Forcing	Vertical/horizontal resolutions	Large-scale dynamic constraints	Cloud scheme	Model references
CAFS (NOAA)	CAFS-bsl	ERA-Interim	L42, ~10 × 10 km	Grid point nudging above 500 hPa (T, U, V, Q)[a]	Morrison et al. (2009)[c,d]	Intrieri et al., in preparation
	CAFS-ini	ERA-Interim	L42, ~10 × 10 km	Grid point nudging above 500 hPa (T, U, V, Q)[a]	Morrison et al. (2009)[c,e]	Intrieri et al., in preparation
CCLM (U Trier)	CCLMi	ERA-Interim	L60, 0.125° (~15 × 15 km)	Forecast mode (reinitialized at 18 UTC, 6-h spin-up)	Doms et al. (2013)	Gutjahr and Heinemann (2018)
	CCLM5	ERA5	L60, 0.125° (~15 × 15 km)	Forecast mode (reinitialized at 18 UTC, 6-h spin-up)	Doms et al. (2013)	Gutjahr and Heinemann (2018)
HIRHAM5 (AWI)	HIRHAM-v1	ERA-Interim	L40, 0.5° (~50 × 50 km)	Grid point nudging at all levels (T, U, V, Q)[a]	Sundqvist et al. (1989)	Sommerfeld et al. (2015)
	HIRHAM-v2	ERA-Interim	L40, 0.5° (~50 × 50 km)	Grid point nudging at all levels (T, U, V, Q)[a]	Tompkins (2002)	Sommerfeld et al. (2015)
MARv3.11 (U Liége)	MAR	ERA-Interim	L24, 25 × 25 km	Spectral nudging above 8 km (T, U, V)[b]	Gallée (1995)	Tedesco and Fettweis (2020)
MetUM (Met Office)	METUM	ERA-Interim	L70, 0.225° (~ 25 × 25 km)	Forecast mode (reinitialized at 00 and 12 UTC, 12-h spin-up)	Wilson et al. (2008)	Orr et al. (2014)
WRF-CU (U Colorado)	WRF	ERA-Interim	L40, 25 × 25 km	Grid point nudging top 20 levels (T, U, V)[b]	Morrison et al. (2009)[a]	Powers et al. (2017)

[a]Nudged every time step. [b]Nudged every 6 h. [c]Double moment cloud microphysics scheme. [d]Ice nuclei concentration: 0.16 L^{-1}. [e]Ice nuclei concentration: 1.0 L^{-1}.

The computational advancement from the ARCMIP (Rinke et al., 2006) can be found in the horizontal and vertical resolutions: horizontal resolution is finer than 25 km, except for the HIRHAM, and vertical resolution ranges between 40 and 70 layers, except for the MAR (24 layers). The increased number of lower-tropospheric layers makes it possible to better compare the vertical structure of meteorological profiles, including cloud properties. All RCMs have seven or more model layers below the 500-m level (the CCLM has the most, i.e., 12 layers).

ERA-Interim reanalysis data forced the six RCMs as lateral boundary conditions over the Arctic CORDEX domain (https://cordex.org/domains/region-11-arctic/; an additional run was performed using the CCLM, as mentioned above). The bottom boundary conditions, in particular SST, were updated daily by the ERA-Interim data. Similarly, the SIC was addressed in the same manner (except for CCLM, which used AMSR2 SIC and CAFS, which incorporated its own SIC in the coupled system). This study did not focus on the ice-covered area; however, a description of the treatment of sea ice by the RCMs is described in Sedlar et al. (2020).

Surface heat and radiation fluxes for the comparison among the RCMs and the observations were prepared as hourly outputs at the closest grid point of the ship as 1-h averaged variables. Other variables in the RCMs are instantaneous values every hour, whereas 1-h averaged values were used in the observations to reduce the observational spikes. Comparison was made using observations acquired from 00:00 UTC on September 7 to 00:00 UTC on September 25, 2014, when the RV *Mirai* stationary point observations were performed (see Section 2.2). Near-surface variables for comparison with the ship observations (e.g., air temperature at the height of 21 m and winds at 25 m above sea level: asl) were interpolated linearly using the first two or three lower layers in each RCM.

2.2. Observations Over the Ice-free Chukchi Sea

During September 7–25, 2014, stationary point observations were performed onboard the RV *Mirai* over the ice-free Chukchi Sea at 74.75°N 162°W (Figure 1). The position of the ship was approximately 150 km from the ice edge. The Beaufort high-pressure system occasionally brought a relatively cold air mass from the perennial sea-ice area to the stationary point that was never covered with new ice nor drifted sea ice.

Throughout the study period, SST remained near 0°C. A detailed description of the cruise can be found in Inoue (2014), and the data that are free for use can be sourced at http://www.godac.jamstec.go.jp/darwin/cruise/mirai/mr14-05/e (JAMSTEC, 2015).

In this study, the following data were used for evaluation of the output of the RCMs. A Vaisala HMP155 humidity and temperature probe, together with an R. M. Young 43408 Gill aspirated radiation shield, was used to measure temperature and humidity on the starboard and port side of the compass deck (21-m asl). Winds were observed using an R. M. Young 05106 mechanical wind sensor positioned on the foremast (25-m asl).

Measurements of SWD and LWD were obtained using Eppley PSP and PIR sensors positioned on the foremast (25-m asl). NOAA installed a similar (more recent calibration) radiation system on the roof of the navigation deck, which produced data with quality superior to that of the ship's radiometers despite occasional shading by the C-band radome. It was found that the SWD data obtained by ship's PSP had a positive gain bias of approximately 10% and an offset value of 2.9 W/m^2. Therefore, the SWD values used in this study were adjusted for this gain bias and offset. To calculate the upward shortwave and longwave radiation (SWU and LWU, respectively), the values of sea surface albedo and emissivity were assumed as 0.055 and 0.97, respectively, which represent typical open ocean conditions (Fairall et al., 1996). SST was monitored underwater at 5-m depth using an underway surface water monitoring system. NOAA also installed high-frequency instruments for obtaining surface turbulent heat fluxes (e.g., a Gill Wind Master Pro model 1561 for wind and air temperature and a LI-COR model 750 for water vapor). As the data were not acquired continuously owing to frequent maintenance, the data were not used in this study. Instead, estimation of the same fluxes based on surface meteorological data using the NOAA COARE 3.5 bulk algorithm (Edson et al., 2013) was used because the values of the fluxes determined by the eddy covariance method and the bulk method agreed well (Inoue, 2014). The flux data, including the radiation data collected by NOAA, are available for download from the following location: ftp1.esrl.noaa.gov/psd3/cruises/Mirai_MR14_05/Mirai/Scientific_analysis/flux/procturb/total/ (last access on June 18 2020).

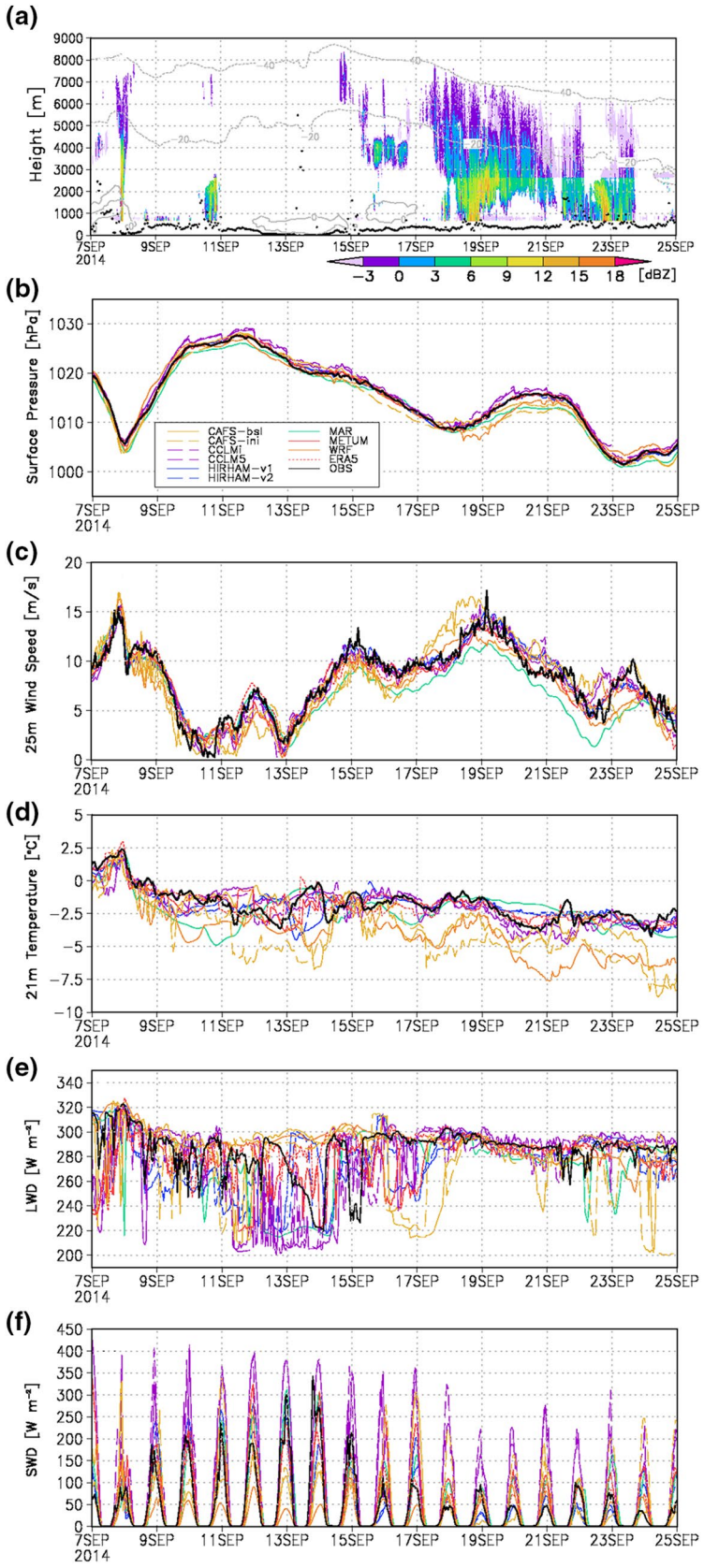
Three-hourly radiosonde launches were performed using a Vaisala DigiCORA III (ver.3.64) SPS 311 sounding system, RS92-SGPD GPS radiosondes, and an ASAP auto balloon launcher. All data were fed into the Global Telecommunication System via the Japan Meteorological Agency immediately after each observation was completed. This means that most of reanalysis products would have assimilated those data. Cloud measurements were performed using HYdrometeor VIdeo Sonde (HYVIS) observations, which collected sequential images of cloud particles as the atmospheric profiles were obtained by the GPS radiosonde (Meisei RS-06G). The HYVIS transmitter has two CCD video cameras: a close-up camera and a microscopic camera. The size of the former (latter) is 7.00 × 5.25 mm (1.200 × 0.9 mm). The minimum particle size detectable is approximately 10 μm. The number of HYVIS observations conducted at the stationary point was six. To determine the cloud ceiling, a Vaisala ceilometer (CL51) continuously monitored the cloud base height with 10-m vertical resolution. A scanning C-band (5,370 MHz) weather radar (Toshiba Co.) monitored precipitating systems by repeating a volume scan with 17 plan position indicators (PPIs) at 6-min intervals. In this study, PPI data at the elevation of 21° were used to establish the presence of precipitating clouds.

ERA5 data (Hersbach et al., 2020) were also utilized as an additional reference data source (1-h product) because it assimilates many satellite-based brightness temperatures for cloud water, column water vapor, and cloud top temperature (Yao et al., 2020). In addition, ERA5 uses a prognostic microphysics scheme that separates cloud water, cloud ice, rain, and snow as prognostic variables (Sotiropoulou et al., 2015).

3. Results

3.1. Near-Surface Conditions

During the stationary point observation, one frontal precipitating system associated with a low pressure system passed on September 8 causing strong winds exceeding 15 m/s (Figures 2a–2c). After that event, a relatively calm situation with intermittent sunny conditions persisted under the Beaufort high pressure system until around September 14. The latter half of the observation period was characterized by moderate-strong winds of around 10 m/s and cold air advection from the sea-ice area with low-level clouds, because the observation point was located at the southwestern edge of the high pressure system where the pressure gradient was relatively strong (Figure 1). In addition, a precipitating system extending from the


(a)
Height [m]
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
7SEP 2014
9SEP
11SEP
13SEP
15SEP
17SEP
19SEP
21SEP
23SEP
25SEP
−3 0 3 6 9 12 15 18 [dBZ]
(b)
Surface Pressure [hPa]
1030
1020
1010
1000
CAFS−bsl
CAFS−ini
CCLM1
CCLM5
HIRHAM−v1
HIRHAM−v2
MAR
METUM
WRF
ERA5
OBS
(c)
25m Wind Speed [m/s]
20
15
10
5
0
(d)
21m Temperature [°C]
5
2.5
0
−2.5
−5
−7.5
−10
(e)
LWD [W m⁻²]
340
320
300
280
260
240
220
200
(f)
SWD [W m⁻²]
450
400
350
300
250
200
150
100
50
0

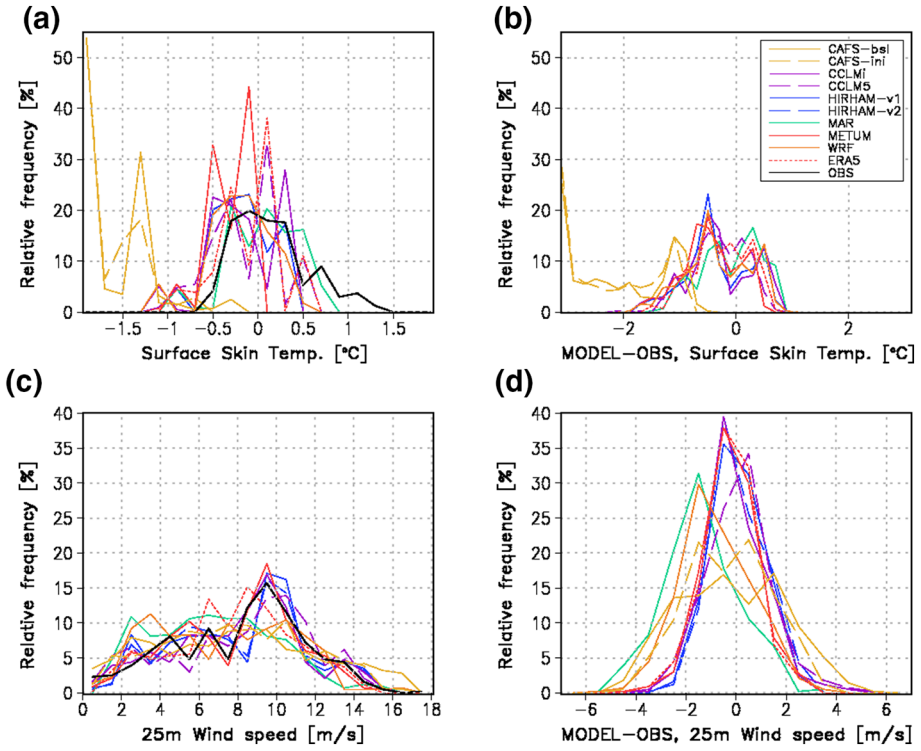


Figure 3. Relative frequency distributions of (a) surface skin temperature (°C; 0.2°C bins), (b) modeled surface skin temperature minus observation difference (°C), (c) surface wind (m/s; 1 m/s bins), and (d) modeled surface wind speed minus observation difference (m/s).

lower troposphere to the upper troposphere dominated during this latter period (Figure 2a). Therefore, during the entire study period, several interesting meteorological conditions were recorded.

The synoptic variability of surface pressure is reproduced well by the RCMs (Figure 2b). Intermodel spread of the surface wind is within ±1 m/s (Figure 2c). A few models (e.g., MAR) tend to underestimate the surface wind by more than 1 m/s regardless of the synoptic variabilities, which would lead to underestimation of the turbulent heat fluxes. Air temperature at 21-m height (Figure 2d) is characterized by more inter-model differences in comparison with the surface pressure. For example, the CAFS and WRF both have significant cold bias (note that the CAFS has partial ice cover at this position). During the sunny and calm situations of September 11–14, some RCMs deviate by 1°C from the observations. During the latter half of the period, the temporal variability in air temperature is relatively small, although some positive bias is evident on September 20 (e.g., MAR). Overall, the deviations of the modeled wind and temperature from the observations (summarized in Table 2) are smaller than found over sea ice, as reported in Sedlar et al. (2020).

One of the main reasons for the small deviations in the surface meteorological parameters between the RCMs and the observations is the quasi-stationary condition of SST (Figures 3a and 3b). The relative frequency distribution of SST shows that for 75% of the period SST fell within the range of 0 ± 0.5°C (Figure 3a), which is a different setup from the ice-covered condition. Skin surface temperature (i.e., no SST output but assumed as SST) in the RCMs also varies within the same range of 0 ± 0.5°C (Figures 3a and S1a) except for the coupled model (CAFS with partial ice cover), which means that the bottom boundary conditions for the SSEB were almost the same among the RCMs.

Figure 2. (a) Time-height cross-section of radar reflectivity obtained by the onboard scanning weather radar (colored shading; dBZ), air temperature obtained by radiosondes (gray dotted contours; °C), and cloud base height obtained by the ceilometer (back line). Time series of hourly observed (black) and modeled (colors) (b) surface pressure (hPa), (c) wind speed (m/s) at 25 m above sea level, (d) air temperature (°C) at 21 m above sea level, (e) downward longwave radiation (W/m^2), and (f) downward shortwave radiation (W/m^2). Radar data are based on 6-min scans at elevation of 21.0° with pulse width change (long and short modes) at the height of approximately 2,580 m. Owing to the effects of ship mast and sea clutter, data below the height of 645 m were masked.

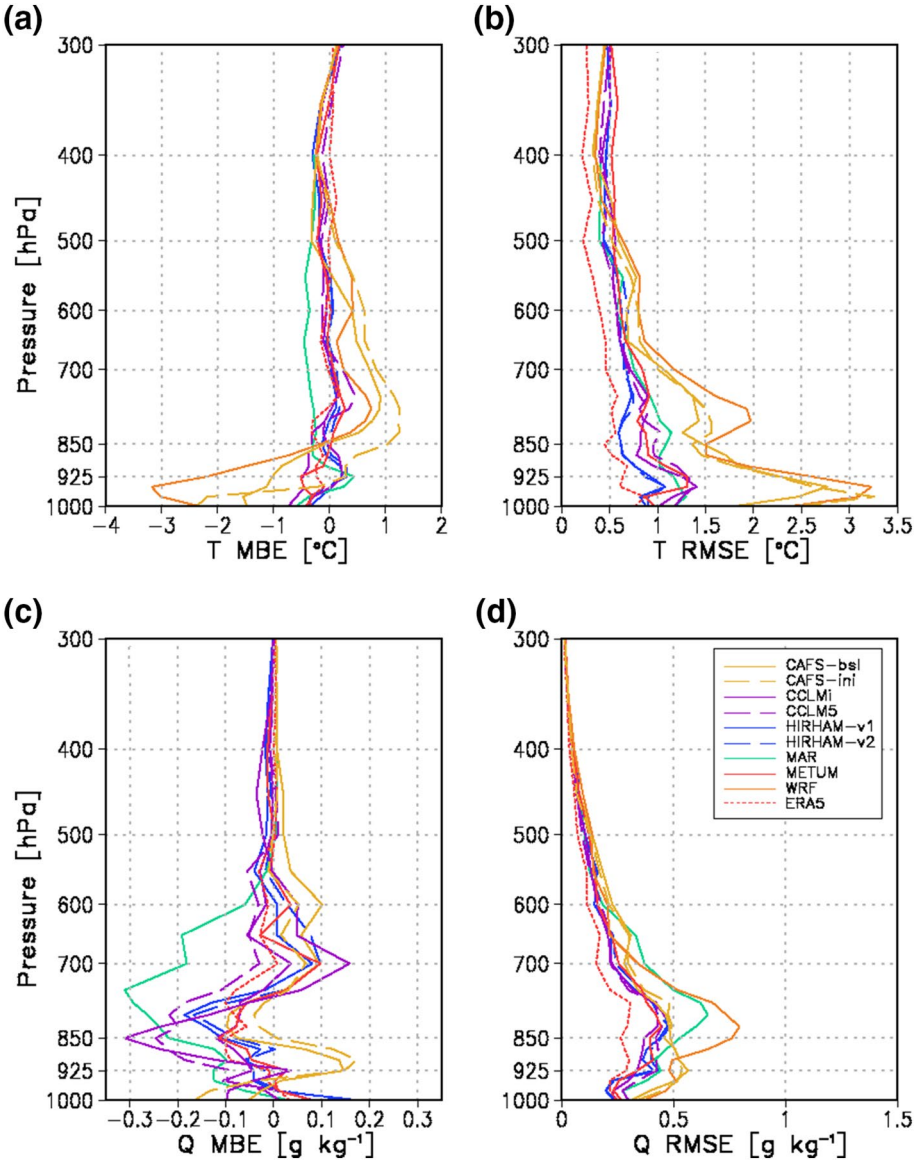


Figure 4. (a and b) Vertical temperature (°C) and (c and d) specific humidity (g/kg) profile statistics, where the mean bias error (MBE) is shown in panels (a and c) and the root mean square error (RMSE) is shown in panels (b and d). Model outputs and observations were interpolated linearly onto the 27 standard pressure levels. Reanalysis error profiles are included for ERA5 (dotted red line).

3.2. Vertical Structures

The vertical structure of the mean bias errors (MBEs) and the root mean square errors (RMSEs) for air temperature and specific humidity are shown in Figure 4. It can be seen that the MBE in air temperature is slightly negative (i.e., around −0.2°C) throughout the entire troposphere in most models (Figure 4a). The exception is the WRF that has a large negative MBE from the surface to 875 hPa (i.e., upto −3.2°C). This large cold bias was also found in the ice-covered case (Sedlar et al., 2020). A moderate negative MBE in the lower troposphere is also found in CAFS owing to the partial ice cover described in the previous section and Figure S1. The altitude range of this negative MBE is different between the CAFS-bsl and CAFS-ini. A negative MBE throughout the mid-troposphere is produced by the MAR. Note that the MAR also has a much larger vertical velocity (up to three times larger) in this region compared to the other models (not shown). Variation in the vertical structure of the RMSE is small (i.e., within 1.5°C) in the lower troposphere

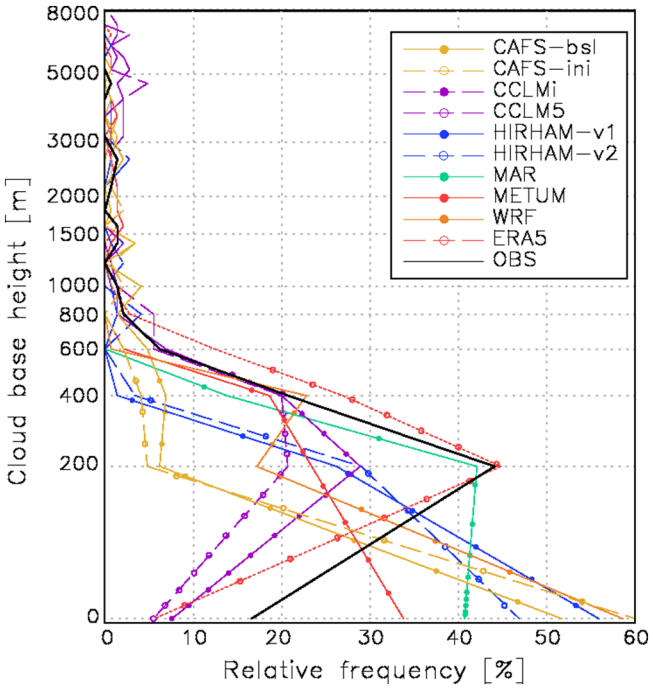


Figure 5. Relative frequency distribution of modeled (colored lines) and observed (black line) cloud base height (m). Vertical resolution is a 200-m bin in which each model has more than two layers within ±100 m of the level. Dots indicate the original model levels below the height of 500 m for each model.

(except for the CAFS and WRF; Figure 4b). The most significant spread in the RMSE is found at around 950 hPa, where low-level clouds exist. An additional aspect is that the vertical profiles of air temperature and specific humidity in most of the RCMs tend to converge towards uniform values at 500 hPa and above, which is the expected behavior because the RCMs use nudging toward the forcing data for the top half of the atmosphere (Table 1). The RMSE in the upper troposphere is still 0.5°C because the RMSE of the forcing data (ERA-Interim) is also around 0.4°C (not shown) which is larger than that in ERA5 (0.25°C).

In terms of specific humidity, a dry MBE can be found in some models in the lower troposphere, in particular above 850 hPa (i.e., CCLM, HIRHAM, and MAR; Figure 4c). It would be worth investigating the moisture transport from the surface through LHF and lower tropospheric stability (LTS) as an indicator of convective motion in the cloud layer (discussed later). As the two HIRHAM runs show the same MBE profiles, the differences in the cloud microphysics schemes are not the cause of the dry MBE. Comparison of the CCLM5 and CCLMi runs reveals that the differences in the MBE are small, suggesting that the choice of the lateral boundary conditions is not the source of the dry bias, although the positive bias above 850 hPa is reduced in CCLM5. The RMSE of specific humidity in the lower troposphere (Figure 4d) is generally below 0.5 g/kg with two peaks at 925 and 830 hPa, where deviations in the heights of the cloud base and top might occur. Overall, as expected, the variation in the vertical structure of the MBE and RMSE in air temperature and specific humidity over the ice-free ocean is smaller than over sea ice (Sedlar et al., 2020). It should be noted that the sign of the MBE in specific humidity is opposite (i.e., negative over the open ocean and positive over sea ice) despite the same negative MBE in air temperature in the lower troposphere, suggesting that there might be differences in the behavior of cloud formation and/or precipitation between the two bottom boundary conditions.

Measurements from the shipboard ceilometer during the study period showed that cloud occurrence at 200 ± 100 m and in the surface–500-m layer was more than 40% and 80%, respectively (Figure 5), which means that boundary layer clouds over the open ocean were dominant. It should be noted that the frequency of occurrence of cloud lower than the 100-m ceiling was less than 20% partly. One of the reasons for the surface cloud-free layer was the cold air advection during the latter half of the period with stratocumulus. The modeled cloud base height among the RCMs, which is defined as the lowest model level at which cloud water/ice mixing ratio is greater than 0.01 g/kg (i.e., cloudy condition), can be categorized into two groups. One group has steady increase in frequency toward the sea surface (i.e., CAFS, HIRHAM, METUM, and WRF); the second group has peak frequency between the 200- and 500-m levels (i.e., CCLM, MAR, and ERA5), similar to the observations. The cold bias in air temperature in the lower troposphere in the CAFS and WRF (Figure 4a) is consistent with the bias in cloud ceiling in these models. However, other models cannot be explained by the temperature and moisture biases. This means that other processes such as the cloud scheme and SSEB might be responsible for determining cloud base height. An outlier in cloud base height is evident in the middle of the troposphere in the CCLM runs. Only ERA5 reproduces the observed frequency well. From the perspective of vertical resolution, higher vertical resolution near the surface does not always provide better representation of the cloud ceiling (e.g., MAR).

3.3. Surface Energy Fluxes

As half the period spent at the stationary point observation was nighttime, it is natural that the relative frequency distribution of the observed SWD is biased toward lower values (Figure 6a); however, it is worth considering the logarithmic slope of the frequency. The WRF fails to reproduce the cases when SWD is larger than 70 W/m^2 partly because of its excessive simulation of optically thick low level clouds (Figure 5).

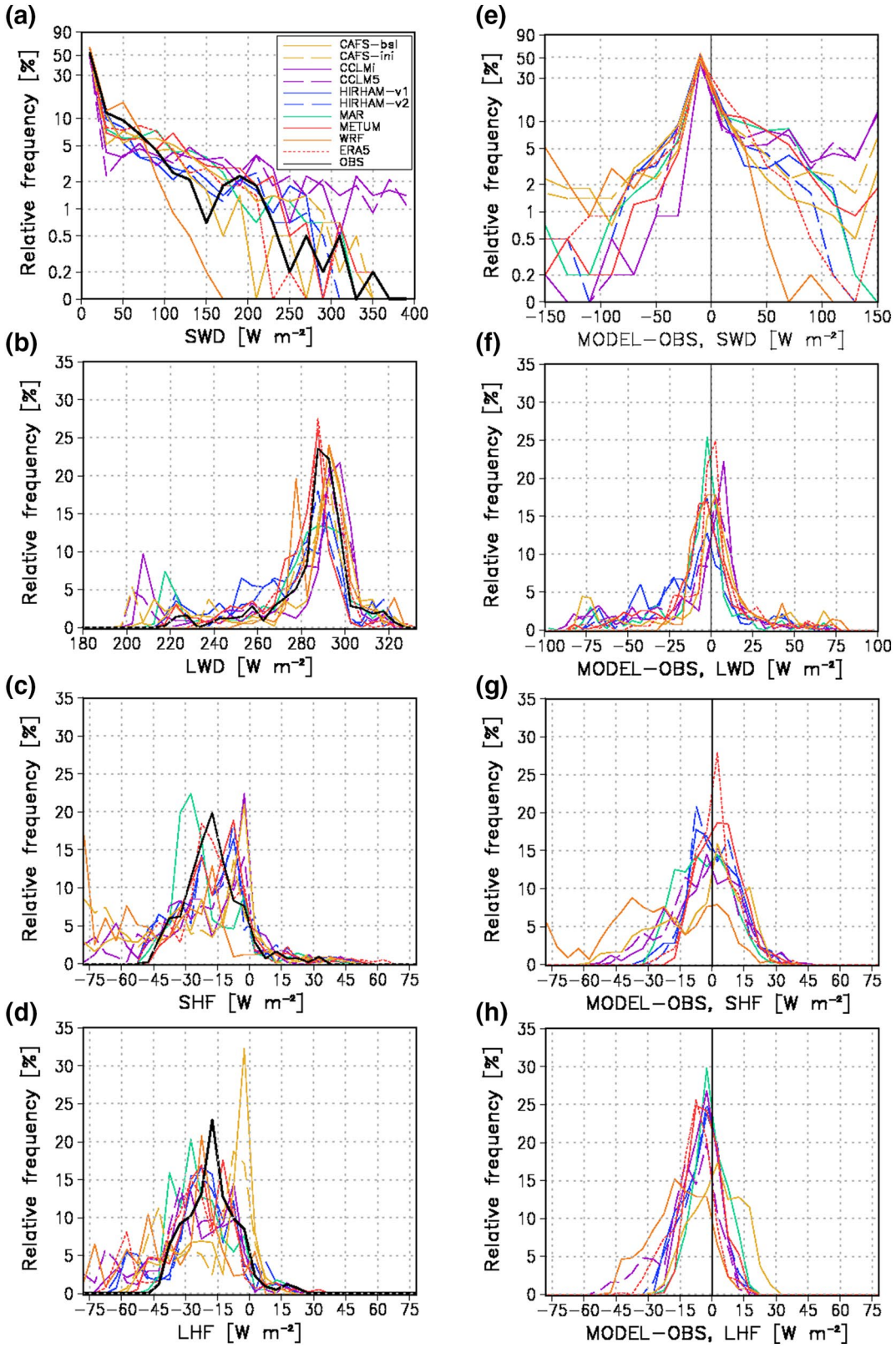


Figure 6. Relative frequency distributions (%) of observed (black line) and modeled (colored lines) flux components (W/m^2): (a) downward shortwave radiation (SWD; 20 W/m^2 bins), (b) downward longwave radiation (LWD; 5 W/m^2 bins), (c) sensible heat flux (SHF; 5 W/m^2 bins), (d) latent heat flux (LHF; 5 W/m^2 bins), and model minus observation relative frequency distributions of (e) SWD, (f) LWD, (g) SHF, and (h) LHF.

The CAFS-bsl and ERA5 underestimated it at the range between 200 and 300 W/m^2. On the other hand, the CCLMi and CCLM5 tend to overestimate SWD, in particular at the range larger than 330 W/m^2. This could be attributed to the SWD from the less amount of cloud water, which is discussed later. The two HIRHAM runs, WRF, CAFS-bsl, and METUM all underestimate the occurrence of SWD values larger than around 250 W/m^2. Note that the METUM overestimates the frequency of occurrence of SWD values in the range 100–200 W/m^2, resulting in a bias of 7 W/m^2 (Table 3).

The relative frequency in observed LWD becomes markedly high between 280 and 300 W/m^2, accounting for more than half the total occurrence (Figure 6b). This means that a cloudy situation was dominant, which is consistent with the observational results obtained using the ceilometer (Figure 5). The lower value of LWD at around 225 W/m^2 with 2% relative frequency indicates that clear-sky conditions when any hydrometeors mixing ratio equals zero also occurred. The relative frequency of LWD in the CCLMi has a maximum at 295 W/m^2, which is 10 W/m^2 larger than the observed peak value. Another feature of the CCLMi run is the second peak at around 200–210 W/m^2, which indicates that the condition of a clear sky and/or upper-level clouds occurred frequently, although such conditions were not observed that frequently (the MAR has the same feature). Furthermore, the upper-cloud ceiling in the CCLM supports these results (Figure 5). Nevertheless, the values around 200–230 W/m^2 from all the RCMs and the observations reflect the occurrence of some clear-sky events. The HIRHAM runs successfully reproduce the frequent LWD value at around 290 W/m^2; however, the frequency is underestimated by nearly 10%. Conversely, the frequency is relatively high around 250–280 W/m^2, which might reflect the lower temperature of the cloud bottom and/or a smaller amount of cloud water. Although the WRF has the two peaks around 295 and 280 W/m^2, mainly resulted from the difference in the cloud base heights (discussed later in Figure 9a), the first peak (24%) was very close to the observation. The MAR underestimates the frequency in the range of 280–300 W/m^2 but overestimates it in the lower range of LWD. Solid and liquid precipitation near the surface would contribute to lower values of LWD (discussed later). The CAFS successfully reproduces the frequency distribution of the LWD, while the METUM underestimates the higher values of the observed peak (i.e., it has a narrower spread), resulting in a bias of −8 W/m^2 (Table 3). Overall, the performance of ERA5 is superior to that of the RCMs because of the many satellite-derived cloud products it assimilates (e.g., Yao et al., 2020).

Both the SHF and the LHF show heat loss from the ocean to the atmosphere (i.e., upward) during most of the study (~90% of the period; Figures 6c and 6d). This means that the majority of the period could be identified as a well-mixed regime owing to the existence of a relatively cold and dry air mass moving from the ice-covered region to the ice-free ocean. The frequency distributions of modeled SHF and LHF show the same pattern as the observations; however, a peak in the frequency of SHF is can be seen between −20 and −5 W/m^2 in half of the RCMs. The MAR overestimated the surface cooling by SHF with the peak frequency at −25 W/m^2 despite a weak wind bias (Figures 3c and 3d), while the WRF overestimated it with the peak frequency at −80 W/m^2 by a cold temperature bias (Figure 4a). The peak frequency in the CCLM runs are different from each other, although both have similar biases in air temperature and winds (Figures 3c and 4a). The values lower than −50 W/m^2 are apparent in the CAFS, CCLM, and WRF, partly due to the negative bias in the air temperature (Figure 4a). The peak frequency in LHF can be found between −30 and −10 W/m^2. The CAFS runs have the peak frequency between −10 and 0 W/m^2 due to the ice cover. However, most models also overestimated the frequency at the range smaller than −50 W/m^2. Overall, the HIRHAM and METUM were broadly able to reproduce the observed conditions; however, it was characterized by a relatively small overestimate of upward LHF, and an underestimate of upward SHF, which therefore largely cancel (Table 3).

3.4. Sea Surface Energy Budget

The observed SSEB shows a general loss of heat from the sea surface (during ~70% of the total period) with a peak value between −70 and −50 W/m^2 (Figure 7a). This heat loss could be attributed mainly to SHF and LHF cooling (~−35 W/m^2 in total) and net longwave radiation (LWN; LWN = LWD − LWU = 290 − 315 = −25 W/m^2). Surface cooling by SHF and LHF is typical over the open ocean during this season; however, the magnitude of LWN depends heavily on cloud status (it should be noted that SST remained almost constant during this period). Therefore, the degree of representation of LWN under well-mixed conditions is crucial for skillful prediction of the SSEB. Considering this situation, the relative frequency of the modeled SSEB shows that the net heat loss is a common feature in all RCMs. Most RCMs, however, overestimate the degree

Table 3
Mean Bias Errors (MBEs; W/m^2) and Root Mean Square Errors (RMSEs; W/m^2) for SWD, LWD, SHF, LHF, and SEB Among the RCMs

Model	SWD_{MBE}	SWD_{RMSE}	LWD_{MBE}	LWD_{RMSE}	SHF_{MBE}	SHF_{RMSE}	LHF_{MBE}	LHF_{RMSE}	SEB_{MBE}	SEB_{RMSE}
CAFS-bsl	−13	60	−3	34	−5	20	3	12	−21	55
CAFS-ini	5	67	−9	36	−16	28	−1	12	−25	59
CCLMi	108	132	−9	31	−3	17	−7	13	19	76
CCLM5	110	134	−8	29	−5	16	−11	17	17	76
HIRHAM-v1	−8	41	−11	29	1	12	−5	10	−21	44
HIRHAM-v2	−9	32	−8	23	1	11	−5	10	−18	40
MAR	3	32	−9	26	−4	13	−5	9	−13	43
METUM	7	43	−8	24	4	10	−3	8	1	39
WRF	−28	63	6	21	−25	36	−16	20	−59	81
ERA5	−7	38	0	21	1	10	−9	13	−12	37

Abbreviations: LHF, latent heat flux; LWD, downward longwave radiation; RCM, regional climate model; SEB, surface energy budget; SHF, sensible heat flux; SWD, downward shortwave radiation.

of the net heat loss by 10–30 W/m^2 throughout the entire period. Both CCLM runs tend to have a large amplitude of the diurnal cycle of the net heat loss throughout the entire period partly because of the positive bias in the SWD in the daytime (from the time series shown in Figure 7c). During daytime, overestimation of SWD (Figure 6a) contributes to much heat input into the ocean, while less LWD (Figure 6b) at night contributes to overestimation of the net heat loss. This might be attributable to the representation of clouds. The WRF overestimates the net heat loss after September 16 which coincides with the occurrence of frequent precipitation (Figure 2a) and thus colder clouds. The frequency of a large negative difference of more than −50 W/m^2 exceeds 50% (Figure 7b). This is derived mainly from the larger SHF and LHF (Figures 6c and 6d) and smaller SWD (Figure 6a) associated with the negative bias of near-surface air temperature (Figure 4a) and much cloudiness (Figure 5). The two HIRHAM runs match the variability of the observations reasonably well, although the amplitude of the diurnal variability of the SEB is sometimes mismatched (Figure 7c).

Regarding the turn from the positive to negative daily SSEB, the amplitude of daily SWD would be a primary factor during this period. The observations show that the daily SSEB turned to a negative value of −20 W/m^2 on September 15 (Inoue, 2014; i.e., the onset of prefreezing indicated by the black triangle in Figure 7c). From this day, the daily variability of the SSEB became small, partly because of the persistent low-level clouds and partly because of the seasonal decrease in SWD. Half the models simulate the onset of prefreezing 1 week ahead. This discrepancy can be attributed to reduced SWD (i.e., CAFS and WRF), and considerable SHF and LHF cooling (CAFS, MAR, and WRF). On the other hand, in both CCLM runs, the onset of prefreezing is delayed a few days partly due to the positive bias in the SWD (Table 3). It should be noted that although the METUM represents SSEB reasonable well, this is attributed to compensation of errors in the energy budget (Figure 6), that is, the errors in SWD and LWD largely cancel (Table 3), as do the errors in SHF and LHF, so both the incoming radiative flux and upward turbulent flux are both correct (although for the wrong reasons).

4. Discussion

4.1. Relationship Between LWN and LTS

The heat loss from the sea surface with clouds was the prevailing situation during the latter half of the study period. As SWD was no longer the primary factor of the SSEB during this period, LWN (particularly LWD) became the primary driver determining the sign of the SSEB. The amount of LWD depends on the state of the boundary layer clouds that is often constrained by LTS, which is defined as the difference in equivalent potential temperature between two levels (Wood & Bretherton, 2006). In this study, LTS was calculated based on the values at 925 hPa and 21-m asl. The observed relative frequency distribution between the LWD and LTS recorded a maximum value in a well-mixed regime (i.e., unstable stratification where LTS < −1 K) with LWN in the range between −30 and −20 W/m^2 (black dot in Figure 8). This situation indicates a rela-

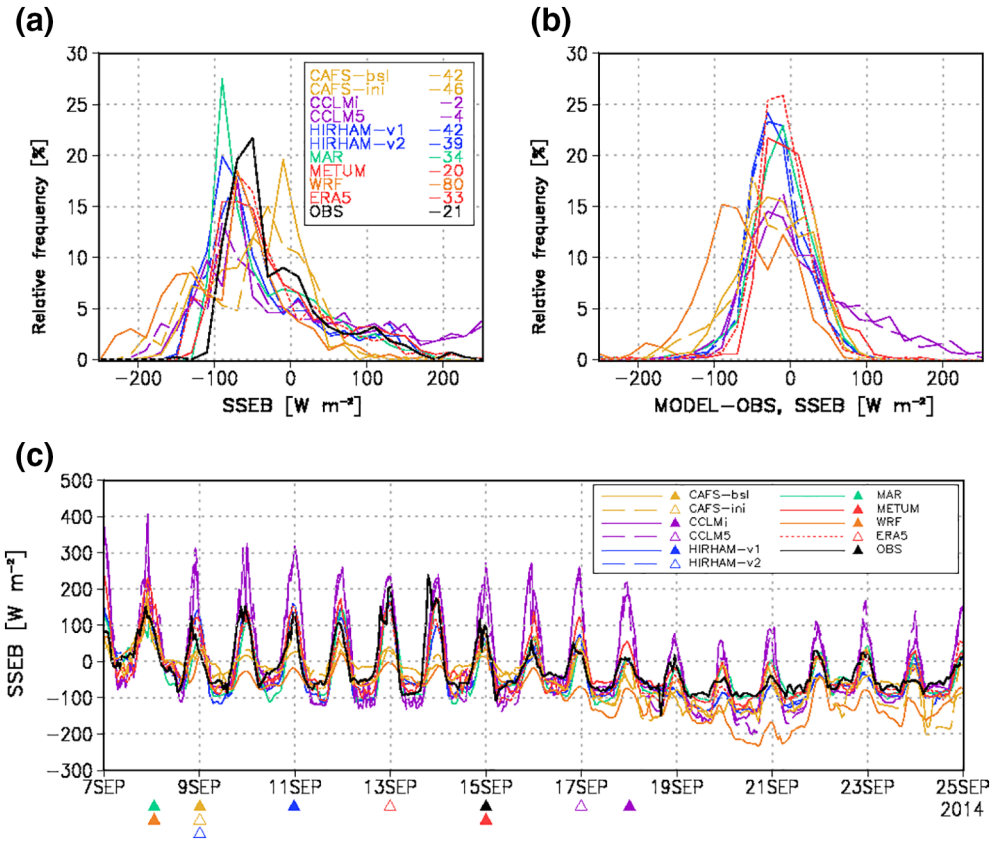


Figure 7. (a and b) The same as in Figure 6 but for the surface energy budget (SEB; W/m^2; 20 W/m^2 bins). (c) Time series of the SEB, where positive (negative) SEB values, indicate sea surface gain (loss) of energy. Each triangle in (c) indicates the date when the daily mean SEB first turned negative. The numbers in (a) indicate the mean SEB values during the period.

tively cold air mass over a warm ocean (LTS is negative) with some clouds (LWD is larger than in clear-sky conditions; Figure 6b). Thus, LWD would be determined mainly by the cloud phase (i.e., liquid and/or solid), temperature of the cloud base, and cloud base height.

Both HIRHAM runs show a nearly neutral regime (−1 to 0 K) with a long-tailed LWN distribution toward −70 W/m^2 (Figures 8f and 8g). This means that stronger longwave cooling in comparison with the observations would be attributed to less representation of boundary layer clouds (e.g., less cloud cover, less cloud liquid water, lower temperature of cloud base, and/or presence of an upper-cloud base). The difference in the cloud microphysics in the HIRHAM is found in the sharpness of the LWN frequency distribution. Specifically, the HIRHAM-v2 based on the Tompkins scheme has a narrow LWN range and allows for a few events under unstable stratification (LTS < −1 K; Figure 8g). Despite overestimation of the occurrence of the stable regime (LST > 0 K) in the both CCLM runs, the model reproduces the nearly neutral regime (−1 to 0 K in both CCLM runs) and the well-mixed regime (LST < −1 K in the CCLM5) with a smaller variability of LWN, that is, between −35 and −15 W/m^2 (Figures 8d and 8e). The finest vertical resolution in the CCLM (12 layers below 500-m height; 18 layers below 1,000-m height) with finer horizontal resolution (0.125 × 0.125°) might contribute to better representation. The MAR (Figure 8h) is the only model that produces a more well-mixed regime (LTS < −2 K), which is also apparent in the observations. The peak frequency and the range of LWN also agree with the observations; however, the reason would be different from the HIRHAM. This is because vertical motion in the MAR is more than three times of magnitude stronger than in the other models (figure not shown), which suggests that the MAR favors relatively unstable conditions regardless of the presence of clouds. Consequently, the MAR produces much more snow than cloud ice (i.e., nearly zero, as discussed later). The CAFS-bsl has a narrow range of LWN with a maximum relative

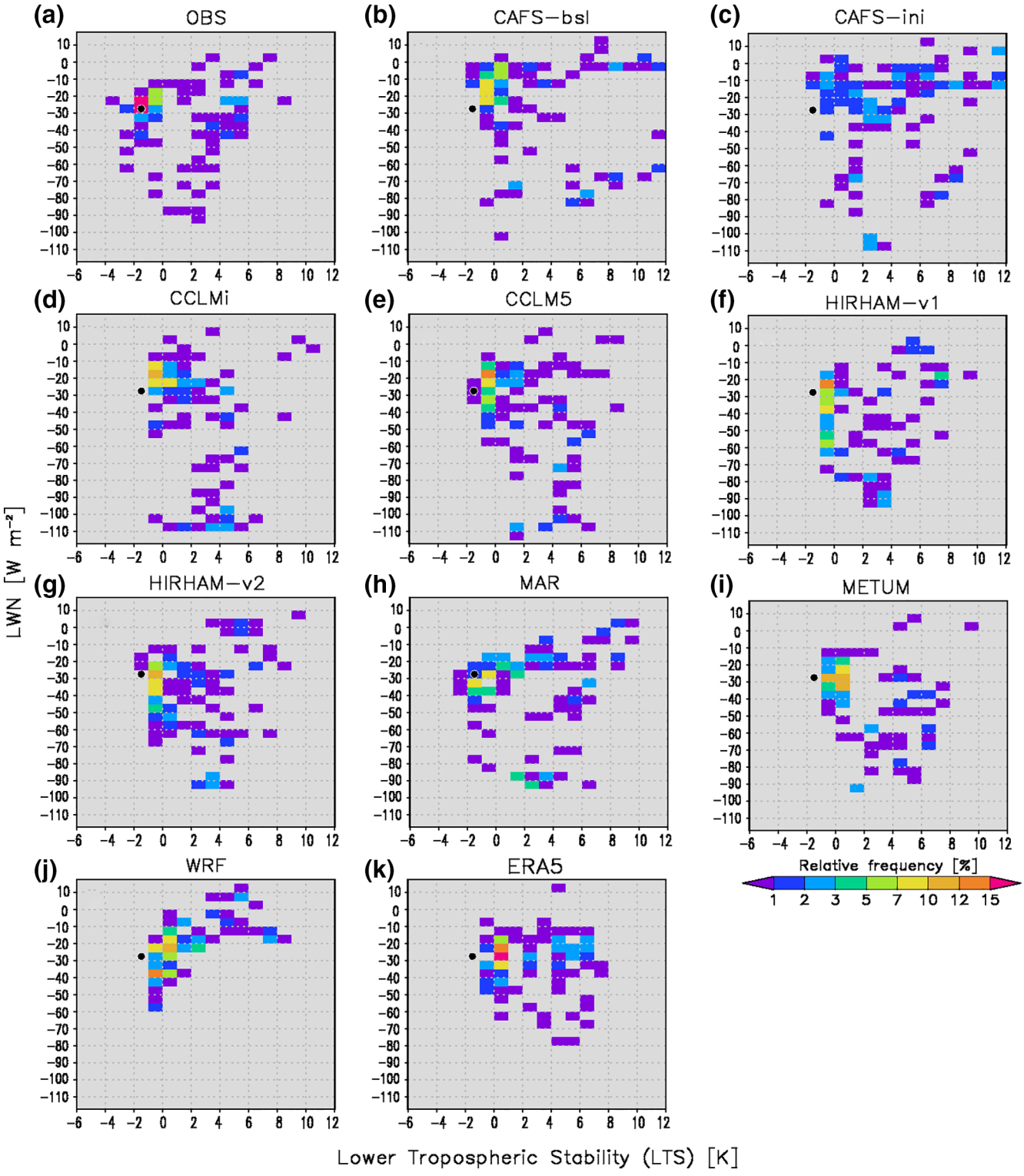


Figure 8. Relative frequency distributions showing the relationship between surface net longwave radiation (LWN; W/m^2) and lower tropospheric stability (LTS; K) for (a) observations, (b) CAFS-bsl, (c) CAFS-ini, (d) CCLMi, (e) CCLM5, (f) HIRHAM-v1, (g) HIRHAM-v2, (h) MAR, (i) METUM, (j) WRF, and (k) ERA5. Statistics are valid for the period from 00:00 UTC on September 7 to 00:00 UTC on September 25. The black dot in each panel indicates the place where the observation recorded the maximum frequency.

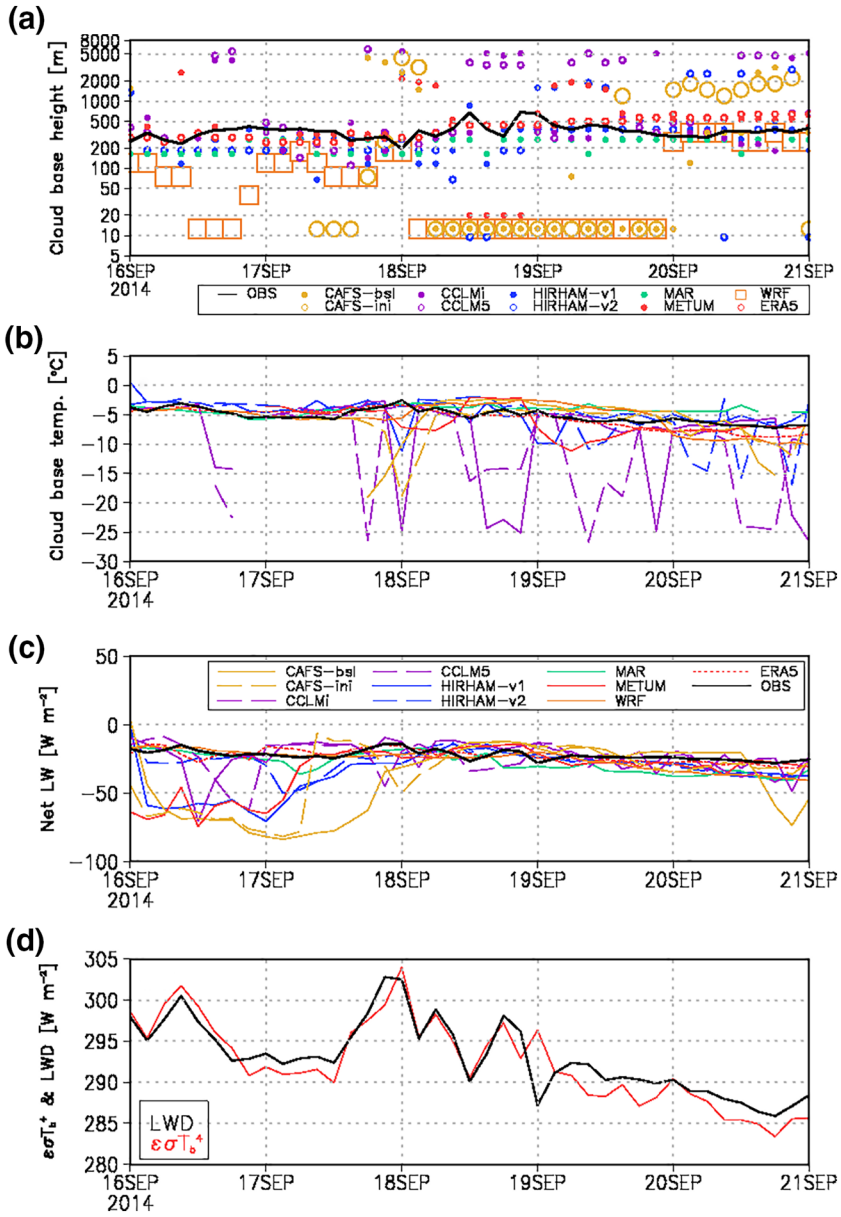


Figure 9. Time series of observed (black line) and modeled (colored lines) of (a) cloud base height (m), (b) cloud base temperature (°C), (c) net longwave radiation (W/m^2), and (d) observed downward longwave radiation (LWD; black line, unit: W/m^2) and LWD estimated based on the cloud base temperature ($\varepsilon\sigma T_b^4$; red line).

frequency of 7%–10% under a nearly neutral situation (LTS > −1 K; Figure 8b). In addition, the CAFS must react very differently because of its own coupled system. For example, the cold bias of the air temperature in the lower troposphere results from the lower bias in SST (Figure 3a), while the SST is easily cooled by the turbulent heat fluxes (Figures 6c and 6d), presumably due to the shallow ocean mixed layer although the observed ocean mixed layer extended to the depth of at least 20 m (Kawaguchi et al., 2016). The increase in the ice nucleating particles in the CAFS-ini promotes more stable conditions with a weaker LWN–LTS relationship because ice clouds dominate (Figure 8c). The WRF has a cold bias in the lower troposphere; however, despite overestimation of the net heat loss by turbulent heat fluxes (Figures 6c and 6d), the range of LTS is close to neutral (Figure 8j). The relatively lower LWN (−35 W/m^2) can be attributed to the lower air temperature at the cloud base. The METUM and ERA5 favor the neutral or stable condition (LTS > −1 K), although the LWN range and peak frequency are relatively consistent with the observations (Figures 8i and

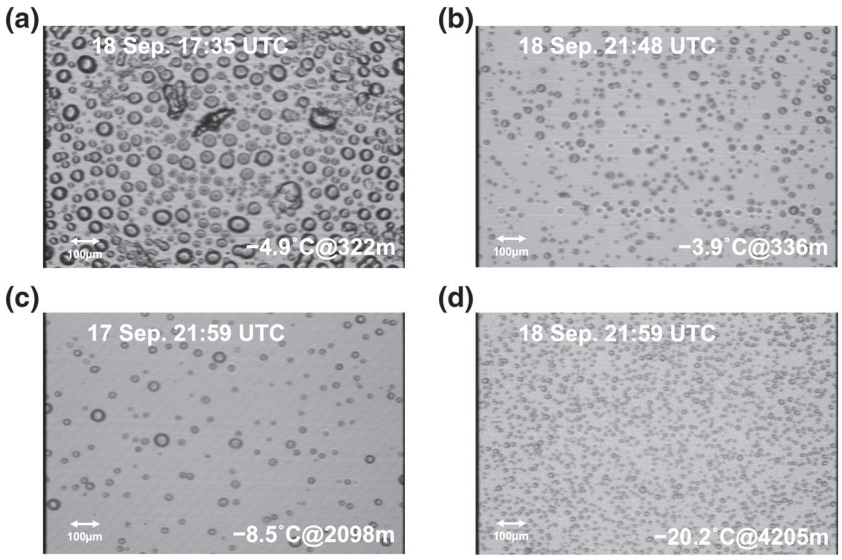


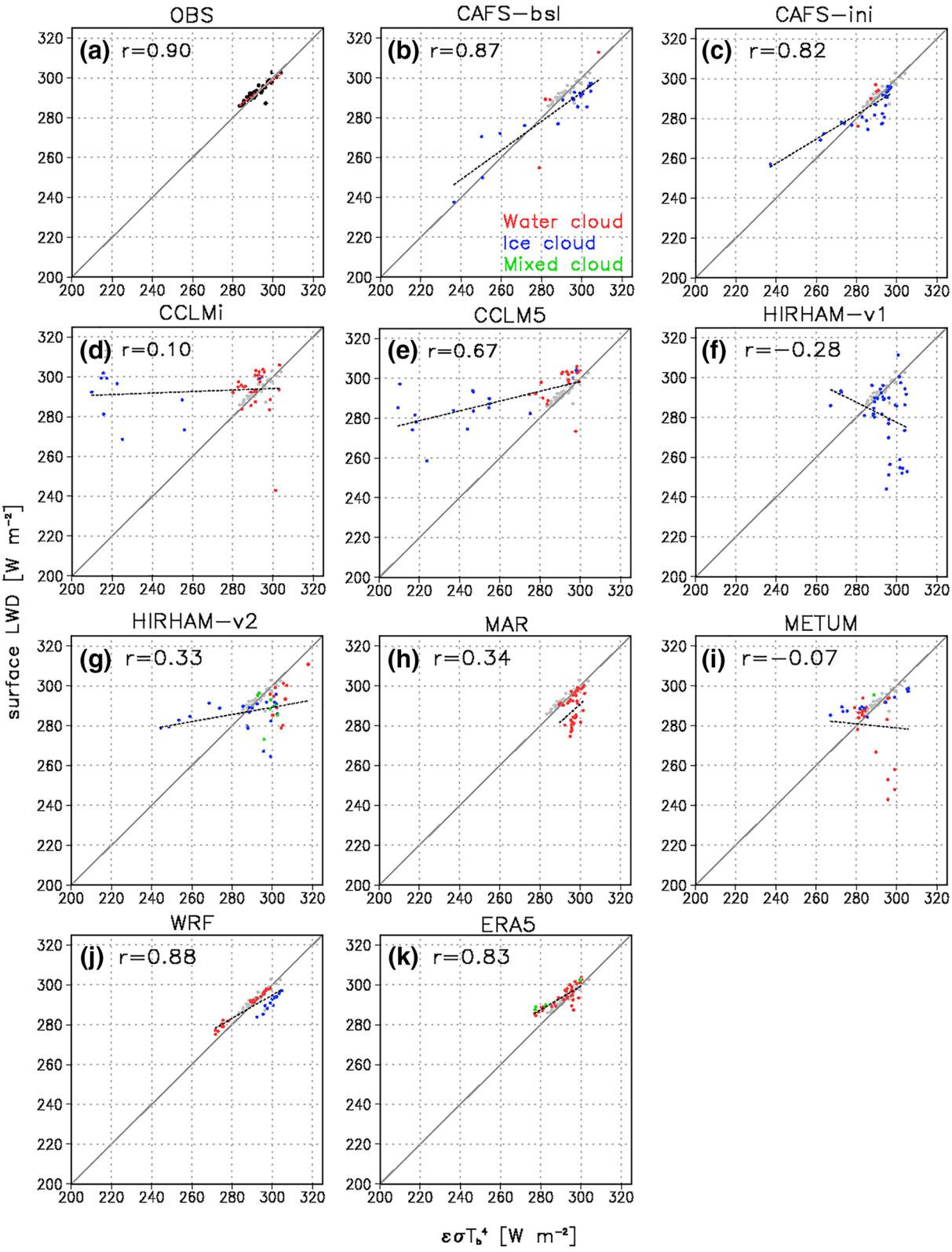
Figure 10. HYVIS images taken during 17–18 September, 2014: (a and b) in the lower troposphere near cloud base and (c and d) in the mid-troposphere. Air temperature and height are indicated in the lower-right corner of each panel. The size of each image is 1.2 × 0.9 mm.

8k). Overall, most RCMs fail to reproduce the well-mixed regime (LTS < −1 K), which is the same pattern revealed for the ice-covered condition (Sedlar et al., 2020).

4.2. Relationship Between Cloud Base Height/Temperature and LWD

Based on the results of the onboard ceilometer from September 16 to 21, cloud base height stays broadly the same with increased variability (Figure 9a), which was partly due to modification of the air mass by advection of the cold air over the relatively warm sea surface (i.e., SSEB < 0 W/m^2, as seen in Figure 7c). This situation also coincides with an increase in precipitating clouds (Figure 2a). The period is characterized by a decrease in cloud base temperature (Figure 9b) and a decrease in LWD (Figure 9d). The HYVIS image taken at 17:35 UTC on September 18 reveals that cloud water (with a minimal amount of cloud ice) was dominant in the boundary layer cloud at 322 m under −4.9°C (Figure 10a: the image size is 1.2 × 0.9 mm; typical radius of cloud water droplets is in the range of 20–30 μm). A different situation with smaller-sized cloud water droplets was found at 21:48 UTC on the same day (Figure 10b). The amount of LWD estimated by $\varepsilon\sigma T_b^4$ (where ε is the emissivity (= 1.0), σ is the Stefan–Boltzmann constant (= 5.67×10^{-8} W m^{-2} K^{-4}), and T_b is the temperature at cloud base observed by the radiosondes) agrees well with the onboard measurement of LWD (Figures 9d and 11a: the correlation coefficient is 0.90). It suggests that the emissivity at the cloud base is nearly 1.0 and that cloud water dominated during the latter half of the period (September 16–21).

Cloud base height in the RCMs, which is defined as the lowest model level at which cloud water/ice mixing ratio is greater than 0.01 g/kg, ranges from the surface to the 600-m level during the study period (Figure 9a). The clouds in the CAFS and WRF with a cold bias in air temperature in the lower troposphere dominate just above the sea surface during September 18–20. After September 20, when the air temperature decreased to −2.5°C or below (Figure 2d), the cloud base height increased up to the 300-m level as observed. This well-mixed regime might break the decoupled boundary layer in the CAFS and WRF as indicated by the cold bias near the surface (~925 hPa) and the warm bias aloft (~800 hPa; Figure 4a). Although the planetary boundary layer (PBL) scheme is different between the CAFS (Hong et al., 2006) and the WRF (Nakanishi & Niino, 2006), the combination of each PBL scheme and the same cloud microphysics scheme (Morrison et al., 2009) might cause the decoupled boundary layer with the cold bias near the surface. In the CCLM, upper level clouds occasionally dominate (Figure 9a), as seen in the relative frequency in the cloud base height (Figure 5). This would result in underestimation of the cloud base temperature and LWN. The MAR slightly underestimates cloud base height by 100–200 m, resulting in the warmer cloud base temper-


OBS
(a) r=0.90
CAFS−bsl
(b) r=0.87
Water cloud
Ice cloud
Mixed cloud
CAFS−ini
(c) r=0.82
CCLMi
(d) r=0.10
CCLM5
(e) r=0.67
HIRHAM−v1
(f) r=−0.28
HIRHAM−v2
(g) r=0.33
MAR
(h) r=0.34
METUM
(i) r=−0.07
WRF
(j) r=0.88
ERA5
(k) r=0.83
surface LWD [W m⁻²]
εσT_b⁴ [W m⁻²]

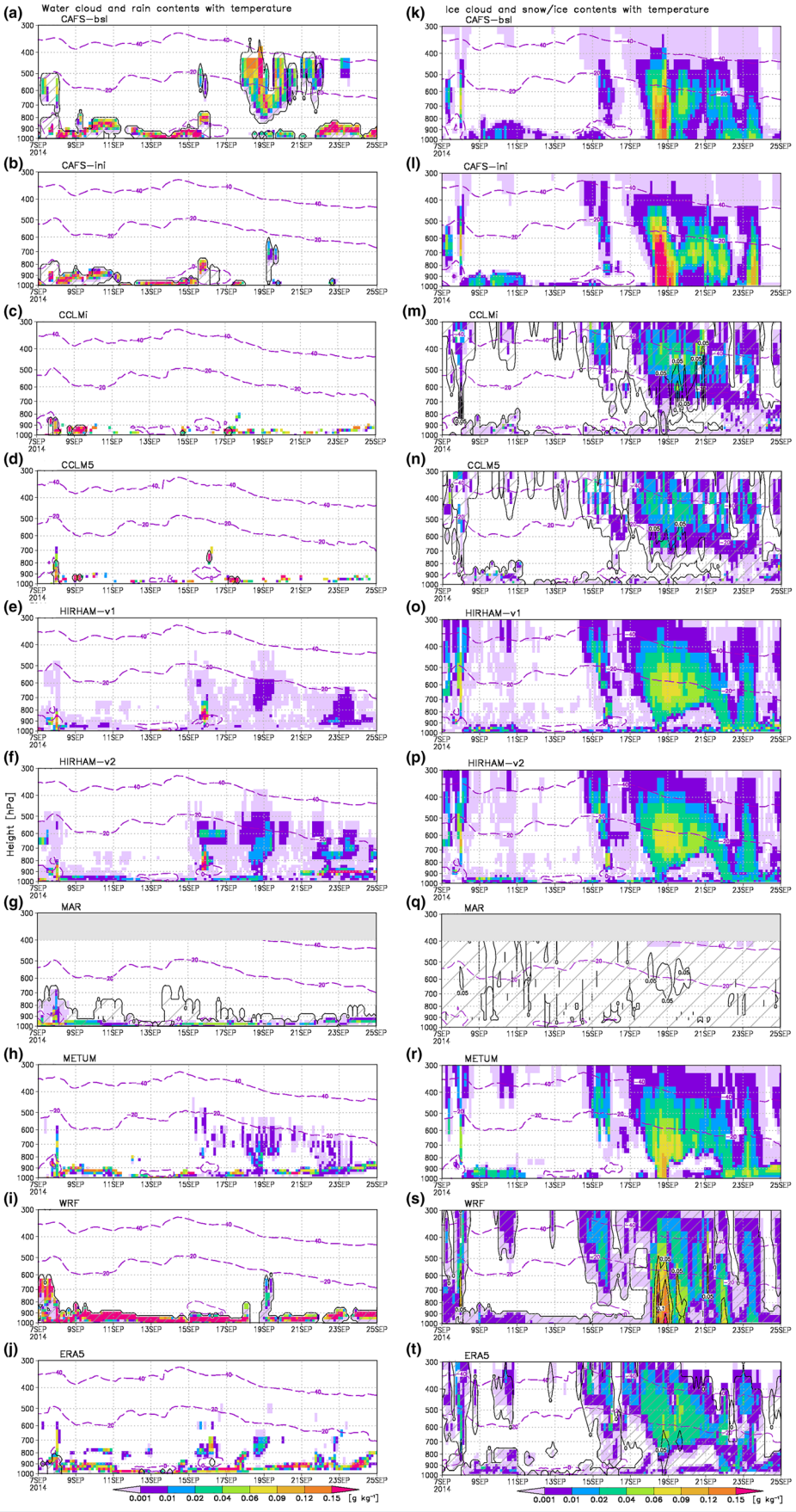
ature (Figure 9b). The representation of LWN in the HIRHAM-v2 is superior to that in the HIRHAM-v1 (Figure 9c) until September 17, whereas there is frequent underestimation of cloud base temperature after September 19 because of the disagreement of cloud base height. The METUM overestimates cloud base height after September 19, resulting in a lower cloud base temperature and reduced LWD.

The high correlation between the observed LWD and estimated LWD_{T_b} ($;\varepsilon\sigma T_b^4$) based on cloud base temperature T_b (Figure 11a) indicates that LWD is determined primarily by the temperature of the water at the bottom of the cloud, as indicated by the HYVIS image (Figures 10a and 10b). This linear relationship between the observed LWD and LWD_{T_b} would give the first-order approximation of the cloud-radiation relationship for cloud water among the RCMs, although other hydrometeors (e.g., rain, snow, and graupel) in the different model layers might influence the LWD. The CAFS and WRF hold the $\mathrm{LWD} - \mathrm{LWD}_{T_b}$ relationship in their system (Figures 11b, 11c, and 11j); however, a cold bias in the lower troposphere would influence the LWD via the lower cloud ceiling with a warmer cloud base temperature and/or additional emission from the overestimation of precipitation (it should be noted that the WRF has snowfall at the surface one order of magnitude larger than the other models: Figure 12s). Comparison of the two HIRHAM runs reveals v2 is better than v1 in terms of the frequent occurrence of liquid or mixed-phase clouds (red and green dots in Figures 11f and 11g), although the relationship remains problematic for ice clouds (blue dots), suggesting that the Tompkins scheme works well in maintaining the $\mathrm{LWD} - \mathrm{LWD}_{T_b}$ relationship. This can be seen in the time-height cross-section of the modeled clouds (discussed later in Figures 12f and 12p). Both CCLM runs underestimate LWD_{T_b} (Figures 11d and 11e) owing to the frequent occurrence of upper-cloud ice instead of lower-cloud water, as seen in Figures 5, 9a, 12c, and 12d (i.e., gray body clouds; $\varepsilon < 1$). It is difficult to conclude whether the forcing of the CCLM5 (i.e., ERA5) is suitable because the relationship of ERA5 (Figure 11k), which is in good agreement with the observations (Figure 11a), is significantly different from the CCLM5 (it should be noted that the CCLM5 uses ERA5 with 1-h temporal resolution, which is a different model setup to the CCLMi). The MAR (Figure 11h) partly reproduced the relationship; however, the half of events indicate that the surface LWD (280 W/m^2) is smaller than the LWD_{T_b} (290 W/m^2). The unrealistic solid precipitation in the MAR might cause this discrepancy (discussed later in Figure 12q). The poor agreement between the METUM and the observations (Figure 11i) and the underestimate of LWD_{T_b} is possibly because the amount of liquid clouds is typically underrepresented in METUM simulations (Gilbert et al., 2020), resulting in too much ice.

4.3. Partitioning Cloud Water and Ice

This section discusses the vertical differences in the fractions of cloud water and ice (Figures 12a–12t) to promote further understanding of the several biases elucidated in previous sections. One of the remarkable biases among the RCMs is the cold bias in the boundary layer in the CAFS and WRF (Figure 4a). The partial sea-ice cover in the CAFS might promote this temperature bias; however, a large amount of cloud water in the lower troposphere (Figures 12a, 12b, and 12i) would increase radiative cooling at the cloud top, resulting in turbulent mixing in the boundary layer and further condensation as cloud water (Inoue et al., 2005). Both models use the double-moment scheme (Morrison et al., 2009), which solves the mixing ratio and number concentration for each hydrometeor. The correlation between the modeled LWD and LWD_{T_b} is highest among the RCMs (0.88 in the WRF and 0.87/0.82 in the CAFS-bls/-ini: Figures 11b, 11c, and 11j), suggesting that cloud-radiation interaction is reproduced better by the double-moment scheme in comparison with a single-moment scheme. The impact of the difference in ice nucleating particles between the CAFS-bsl and CAFS-ini on the evolution of cloud systems can be found after September 17 when the cold and dry air mass from the perennial ice dominated (Figures 12k and 12l). The cloud water in the CAFS-ini almost does not exist (Figure 12b). Instead of this, the cloud ice increased from the surface to the mid-troposphere.

Figure 11. Relationships between downward longwave (LWD) radiation at the surface and that estimated based on the cloud base temperature (LWD_{T_b}) for (a) observation, (b) CAFS-bsl, (c) CAFS-ini, (d) CCLMi, (e) CCLM5, (f) HIRHAM-v1, (g) HIRHAM-v2, (h) MAR, (i) METUM, (j) WRF, and (k) ERA5. Observation data are also indicated in each panel as gray dots. The letter *r* in the upper-left corner and the black dotted line in each panel represent the correlation coefficient and linear regression line, respectively. Different colored dots indicate the cloud phase. Statistics are valid for the period from 00:00 UTC on September 16 to 00:00 UTC on September 21.


(a) Water cloud and rain contents with temperature
(k) Ice cloud and snow/ice contents with temperature
CAFS−bsl
CAFS−ini
CCLMi
CCLM5
HIRHAM−v1
HIRHAM−v2
MAR
METUM
WRF
ERA5
Height [hPa]
7SEP 2014
9SEP
11SEP
13SEP
15SEP
17SEP
19SEP
21SEP
23SEP
25SEP
0.001 0.01 0.02 0.04 0.06 0.09 0.12 0.15 [g kg⁻¹]

The MAR is the only model that did not produce cloud ice except for the level colder than −40°C (Figure 12q). In the microphysics scheme of the MAR, described in Gallée (1995), homogeneous freezing of cloud water and homogeneous deposition of water vapor to form cloud ice are only allowed in conditions under −40°C. Vertical velocity in the MAR is more than three times of magnitude stronger than in the other models, which presumably accelerates the conversion from cloud ice to snow (in fact, snow exists in the MAR, as indicated by the hatching in Figure 12q). Underestimation of cloudiness in the MAR has also been found in an experiment over Greenland (Fettweis et al., 2017). The negative wind speed bias in the lower troposphere (Figure 3c) might result from less radiative cooling at the cloud top and less turbulent mixing (including momentum).

In both CCLM runs, fewer events with persistent low-level clouds in the liquid phase were found in comparison with the other models (Figures 12c and 12d), although this model setup is the finest in terms of horizontal (15 km) and vertical (12 layers below the 400-m level) resolutions (Figure 5). The snow distribution in the CCLM (hatching in Figures 12m and 12n) would be consistent with the weather radar observations (Figure 2a); however, the conversion of cloud water to cloud ice and snow might be too strong. Sedlar et al. (2020) also highlighted the possibility of failure in the interaction between the relative sparsity of low-level cloud ice and persistent cloud water.

Comparison of the two HIRHAM runs reveals the differences in the cloud microphysical schemes (Figures 12e, 12f, 12o, and 12p). The HIRHAM-v2 with the Tompkins scheme reproduces the cloud water near the sea surface more frequently as a mixed-phase cloud system. This characteristic was also found in the case of the ice-covered region (Sedlar et al., 2020). The other feature of this run is that mid-tropospheric clouds (e.g., September 16, 18–19, and 22–23) are found as supercooled cloud droplets. According to the HYVIS soundings at around 22:00 UTC on September 17 and 18, cloud water was detected at 2,100-m asl on September 17 and at 4,200-m asl on September 18 (Figures 10c and 10d). The typical cloud droplet radius size is 10 μm. Therefore, the HIRHAM-v2 can be considered superior to v1 in terms of reproducing cloud water in a cold situation (e.g., <−10°C), although the amount of cloud water is not known quantitatively.

The METUM also simulates relatively less low-level clouds in the liquid phase, consistent with its aforementioned bias in cloud phase (Gilbert et al., 2020). The time-height variability is closest to ERA5 (Figure 12j). Although the vertical distribution of clouds was not observed, the snow distribution in ERA5 (Figure 12t) throughout the entire troposphere is consistent with the weather radar observations (Figure 2a), suggesting that the cloud microphysics scheme and/or the assimilated satellite products perform well in representing ice clouds and precipitating clouds in the mid- and upper troposphere.

5. Summary

The principal issues found in this study are summarized as follows:

- A striking result was that even in an ice-free ocean, the variability of LWD is poorly explained by the modeled cloud bottom characteristics (Figures 9 and 11: height, temperature, and phase). Some models succeeded in representing persistent boundary layer clouds with relatively small amounts of liquid water (i.e., HIRHAM, MAR, and METUM), while other models produced considerably more cloud water with a low-temperature bias in the lower troposphere (i.e., CAFS and WRF). Reduced cloud occurrence was also found in some cases (CCLM). Based on such variability of cloud-radiation interaction, the SSEB also deviated to some degree (Figure 7). In particular, situations with less cloud occurrence (CCLM), less cloud water amount (HIRHAM), and much heat loss by turbulent heat fluxes (CAFS, MAR, and WRF) were found to cause greatest discrepancy from the observations.
- More than 80% of the observed cloud events were accompanied by a low ceiling (below 400 m) with a frequent peak height at 200-m asl (Figure 5). This means there is frequent occurrence of a cloud-free layer just above the sea surface; however, such a vertical structure is not reproduced in the RCMs, except in the CCLM and MAR. One of the reasons is that most models favor neutral and/or stable conditions, al-

Figure 12. Time-height cross-sections of modeled cloud water (colors) and rain (hatching enclosed by contours; left-hand column) for (a) CAFS-bsl, (b) CAFS-ini, (c) CCLMi, (d) CCLM5, (e) HIRHAM-v1, (f) HIRHAM-v2, (g) MAR, (h) METUM, (i) WRF, and (j) ERA5. The figures in the right-hand column (k–t) are the same but for cloud ice (colors) and snow (hatching enclosed by contours). Note that the CAFS, HIRHAM, and METUM did not produce outputs of rain and/or snow. Dotted lines indicate the air temperature.

though the most frequently observed LWN cooling at the sea surface is in the range of −30 to −20 W/m^2 with 35% frequency under unstable stratification (Figure 8a). The CCLM5, HIRHAM-v2, and MAR were partly able to reproduce the relatively unstable conditions with an appropriate LWN range although the stronger vertical velocity in the MAR remains as an issue; the finer vertical and horizontal resolutions (CCLM5) or adequate cloud water representation in the lower layer (HIRHAM-v2) might contribute to the better representation of the LWN–LTS relationship.

- Comparison of the two HIRHAM runs with different cloud microphysics schemes highlighted that cloud water is reproduced better by the Tompkins scheme used in the HIRHAM-v2 (Figure 12f). Although the amount of cloud water in the lower boundary layer was smaller in comparison with that in the CAFS and WRF, this scheme adequately stimulated the supercooled cloud water in the mid-troposphere in a −20°C environment (Figures 10d and 12f). The other RCMs have other specific features, for example, absence of cloud ice (MAR), less cloud water (CCLM and METUM), considerable cloud water (CAFS and WRF). The conversion processes from cloud ice/water to solid/liquid precipitation might be responsible for these discrepancies, as well as biases in the determination of cloud phase (Furtado & Field, 2017). From the perspective of the moments for cloud schemes, the double-moment scheme (Morrison et al., 2009) applied in the CAFS and WRF was found superior in reproducing the relationship between LWD and cloud water characteristics (height and temperature of cloud base: Figures 9 and 11).
- With respect to the number concentration of ice nucleating particles between the CAFS-bsl (0.16 L^{-1}) and CAFS-ini (1.0 L^{-1}), the LWN–LTS relationship was well reproduced in the CAFS-bsl (Figure 8b) because the cloud water has a vital role for determining the LWD in this period. However, under a colder situation over the ice-covered area, the CAFS-ini could simulate the clouds adequately, as reported in Sedlar et al. (2020). In this sense, the ice nucleating particles should be carefully tuned for the target seasons/locations.
- In the METUM the partitioning of cloud water and ice is diagnosed using a single-moment scheme (Wilson et al., 2008), which typically overestimates cloud ice (Figure 12r) and underestimates cloud liquid (Figure 12h) (Gilbert et al., 2020) – evident in the poorly reproduced LWD − LWD$_{T_b}$ relationship (Figure 8i). These biases are consistent with the findings that the METUM underestimated LWD (Figure 6b). However, it is notable that despite this the METUM also overestimated SWD (Figure 6a), which we speculate could be due to a deficit of supercooled liquid upper layer clouds, which can also reflect incoming SW (Barrett et al., 2017). Overall, the representation of SSEB (Figure 7) in the METUM was reasonable because the biases in downward radiation flux and upward turbulent flux more-or-less compensated for each other (Table 3).
- Despite the many differences between the CCLMi and CCLM5, it was difficult to explain why the choice of boundary conditions (ERA-Interim or ERA5) might have led to those differences (e.g., SWD in Figure 6a; LTS in Figures 8d and 8e; and cloud base temperature in Figure 9b). The ERA5 originally shows the best performance for most parameters; however, although it assimilates many satellite products related to clouds (Yao et al., 2020), the CCLM5 does not always provide better performance in comparison with the CCLMi. Analysis of the use of different large-scale dynamic constraints (e.g., grid point nudging) rather than the forecast mode and/or different intervals of updating the lateral boundary conditions (every hour for the CCLM5 and every 6 h for the CCLMi) might provide further understanding.
- The CAFS is the only model that has its own Arctic coupled system. For example, the CAFS has a remarkable cold bias in air temperature (Figure 4a) due to the cold bias in surface skin temperature (Figure 3a) associated with the partial ice cover (10%–80% in SIC: Figure S1). Additional condensation of cloud water in the lower troposphere (Figures 12a and 12b), which strengthens cloud-top radiative cooling and weakens SWD at the surface, would present feedback for additional surface cooling. However, oceanic stratification and its variability could also affect the SST and lower troposphere. This suggests that once reproduction of the cloud system fails, the entire air-(ice)-sea coupled system could deviate considerably from reality.
- Overall, ERA5 performs well despite the coarser horizontal and vertical resolutions than the RCMs. Data assimilation of satellite data might be highly useful to represent cloud-radiation relationships. More importantly, the microphysics scheme in ERA5 (Forbes & Ahlgrimm, 2014) treats liquid and ice clouds separately as prognostic variables, which removes the diagnostic temperature-dependent liquid/ice phase split from the previous scheme (the threshold was −23°C). Therefore, this scheme increases the chance of the supercooled liquid cloud to exist between 0°C to −38°C. Considering the overestimation of cloud

water in CAFS and WRF (Figures 12a, 12b, and 12i) and underestimating them in CCLM, HIRHAM, MAR, and METUM (Figures 12c–h) against ERA5 (Figure 12j), there is room for improvement in representing supercooled clouds as a part of mixed-phase clouds.

6. Conclusion

Treatment of sea ice remains a critical problem regarding intercomparison of RCMs because there are many unique setups in each model (e.g., ice thickness, albedo, and snow depth). Conversely, the ice-free condition, which is an emerging part of the new state of the Arctic Ocean, is the simplest form of boundary conditions for use in RCMs; however, consideration of such an approach has not yet been adopted. Based on this new perspective, the six coordinated RCMs with nine model runs, which form part of the Arctic CORDEX activities, were evaluated using in situ observations obtained by the RV *Mirai* over the ice-free ocean in September 2014. The main target of this work was the SSEB and cloud-radiation interaction. Most RCMs adequately reproduced the temporal evolution of basic near-surface meteorological parameters, with a smaller spread among the RCMs (Figure 2) than found in previous studies of the ice-covered condition; however, many issues associated with clouds and SSEB still exist as summarized above (e.g., cloud phase).

Ever since the ARCMIP era based on the SHEBA field campaign data (Inoue et al., 2006; Rinke et al., 2006; Wyser et al., 2008), representation of the Arctic cloud system in RCMs has remained a challenge. However, the choice of the boundary conditions (e.g., most advanced reanalyzes, satellite products), several epochal parameterizations for clouds and other processes, and the constant evolution of models with high-performance computing could mark a milestone in development of the understanding of the air-ice-sea Arctic coupled system. The development of a standalone atmospheric model remains fundamental for estimation of the benchmarks of the current models, while the development of a coupled model will be critical for establishing an improved earth system model for predicting future climate change and variability. The Multidisciplinary drifting Observatory for the Study of Arctic Climate expedition (MOSAiC Science Plan Writing Team, 2016), which is a 1-year-long expedition into the central Arctic, must be a significant step toward attaining essential understanding of the Arctic climate system and beyond.

Acknowledgments
This work was supported by JSPS KAKENHI (grant numbers 24241009, 18H03745, and 18KK0292) and the Arctic Challenge for Sustainability Project (ArCS) (grant number JPMXD1300000000) of the Japanese Ministry of Education, Culture, Sports, Science and Technology to Jun Inoue. The authors are greatly indebted to the officers and crew of the RV *Mirai*. NOAA radiation data onboard the RV *Mirai* was collected and calculated by Dr. Byron W. Blomquist (NOAA). HYVIS data were collected by Drs. Shuichi Mori and Masaki Katsumata (JAMSTEC). Processing of the C-band radar data was performed by Mr. Souichiro Sueyoshi (Nippon Marine Enterprises, Ltd.). The CCLM contribution (Günther Heinemann) was funded by the German Federal Ministry of Education and Research (BMBF) under grants 03F0776D and 03F0831C (CATS). Annette Rinke acknowledges the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project 268020496 TRR 172, within the Transregional Collaborative Research Center "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)3." The WRF-CU contribution was supported by the Regional and Global Model Analysis (RGMA) component of the U.S. Department of Energy's Office of Science, as a contribution to the HiLAT–RASM project. We thank James Buxton MSc from Edanz Group (https://en-author-services.edanzgroup.com/) for editing a draft of this manuscript. The authors also thank the editor and two anonymous reviewers for their valuable and insightful comments to improve the manuscript.

Data Availability Statement

The observational data used in this study can be freely available from the following web address: http://www.godac.jamstec.go.jp/darwin/datatree/e. The integrated RCMs data set used in this study is also available from the Arctic Data archive System (https://doi.org/10.17592/001.2020091601).

References

Barrett, A. I., Hogan, R. J., & Forbes, R. M. (2017). Why are mixed-phase altocumulus clouds poorly predicted by large-scale models? Part 1. Physical processes. *Journal of Geophysical Research: Atmospheres*, *122*(18), 9903–9926. https://doi.org/10.1002/2016JD026321

Curry, J. A., & Lynch, A. H. (2002). Comparing arctic regional climate model. *EOS, Transactions American Geophysical Union*, *83*(9), 87. https://doi.org/10.1029/2002EO000051

Curry, J. A., Schramm, J. L., Rossow, W. B., & Randall, D. (1996). Overview of arctic cloud and radiation characteristics. *Journal of Climate*, *9*(8), 1731–1764. https://doi.org/10.1175/1520-0442(1996)009⟨1731:OOACAR⟩2.0.CO;2

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597.

Doms, G., Förstner, J., Heise, H., Herzog, H.-J., Mironov, D., Raschendorfer, M., et al. (2013). A description of the nonhydrostatic regional COSMO-model. Part II. Physical parameterizations. Retrieved from http://www.cosmo-model.org/content/model/documentation/core/cosmo_physics_5.00.pdf

Edson, J., Jampana, V., Weller, R., Bigorre, S., Plueddemann, A., Fairall, C., et al. (2013). On the exchange of momentum over the open ocean. *Journal of Physical Oceanography*, *43*, 1589–1610. https://doi.org/10.1175/JPO-D-12-0173.1

English, J. M., Kay, J. E., Gettelman, A., Liu, X., Wang, Y., Zhang, Y., & Chepfer, H. (2014). Contributions of clouds, surface albedos, and mixed-phase ice nucleation schemes to arctic radiation biases in CAM5. *Journal of Climate*, *27*(13), 5174–5197. https://doi.org/10.1175/JCLI-D-13-00608.1

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Fettweis, X., Box, J., Agosta, C., Amory, C., Kittel, C., Lang, C., et al. (2017). Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate mar model. *The Cryosphere*, *11*, 1015–1033. https://doi.org/10.5194/tc-11-1015-2017

Forbes, R. M., & Ahlgrimm, M. (2014). On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model. *Monthly Weather Review*, *142*(9), 3425–3445. https://doi.org/10.1175/MWR-D-13-00325.1

Furtado, K., & Field, P. (2017). The role of ice microphysics parametrizations in determining the prevalence of supercooled liquid water in high-resolution simulations of a Southern Ocean midlatitude cyclone. *Journal of the Atmospheric Sciences*, *74*(6), 2001–2021. https://doi.org/10.1175/JAS-D-16-0165.1

Gallée, H. (1995). Simulation of the mesocyclonic activity in the Ross sea, Antarctica. *Monthly Weather Review*, *123*(7), 2051–2069. https://doi.org/10.1175/1520-0493(1995)123h2051:SOTMAI>2.0.CO;2

Gilbert, E., Orr, A., King, J. C., Renfrew, I. A., Lachlan-Cope, T., Field, P. F., & Boutle, I. A. (2020). Summertime cloud phase strongly influences surface melting on the Larsen C ice shelf, Antarctica. *Quarterly Journal of the Royal Meteorological Society*, *146*(729), 1575–1589. https://doi.org/10.1002/qj.3753

Giorgi, F., Jones, C., & Asrar, G. (2009). Addressing climate information needs at the regional level: The CORDEX framework. *World Meteorological Organization Bulletin*, *58*(3), 175–183.

Gutjahr, O., & Heinemann, G. (2018). A model-based comparison of extreme winds in the arctic and around Greenland. *International Journal of Climatology*, *38*(14), 5272–5292. https://doi.org/10.1002/joc.5729

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https://doi.org/10.1175/MWR3199.1

Inoue, J. (2014). *R/V Mirai Cruise Report MR14-05*. Retrieved from http://www.godac.jamstec.go.jp/catalog/data/doc_catalog/media/MR14-05_all.pdf

Inoue, J. (2020). Review of forecast skills for weather and sea ice in supporting arctic navigation. *Polar Science*, 100523. https://doi.org/10.1016/j.polar.2020.100523

Inoue, J., Enomoto, T., & Hori, M. (2013). The impact of radiosonde data over the ice-free arctic ocean on the atmospheric circulation in the northern hemisphere. *Geophysical Research Letters*, *40*, 864–869. https://doi.org/10.1002/grl.50207

Inoue, J., & Hori, M. E. (2011). Arctic cyclogenesis at the marginal ice zone: A contributory mechanism for the temperature amplification? *Geophysical Research Letters*, *38*(12), L12502. https://doi.org/10.1029/2011GL047696

Inoue, J., Kosović, B., & Curry, J. (2005). Evolution of a storm-driven cloudy boundary layer in the arctic. *Boundary-Layer Meteorology*, *117*, 213–230. https://doi.org/10.1007/s10546-004-6003-2

Inoue, J., Liu, J., Pinto, J. O., & Curry, J. A. (2006). Intercomparison of Arctic regional climate models: Modeling clouds and radiation for SHEBA in May 1998. *Journal of Climate*, *19*(17), 4167–4178. https://doi.org/10.1175/JCLI3854.1

Inoue, J., Sato, K., & Oshima, K. (2018). Comparison of the arctic tropospheric structures from the era-interim reanalysis with in situ observations. *Okhotsk Sea and Polar Oceans Research*, *2*, 7–12.

Inoue, J., Yamazaki, A., Ono, J., Dethloff, K., Maturilli, M., Neuber, R., et al. (2015). Additional arctic observations improve weather and sea-ice forecasts for the Northern Sea Route. *Scientific Reports*, *5*, 16868. https://doi.org/10.1038/srep16868

Intrieri, J. M., Fairall, C. W., Shupe, M. D., Persson, P. O. G., Andreas, E. L., Guest, P. S., & Moritz, R. E. (2002). An annual cycle of Arctic surface cloud forcing at SHEBA. *Journal of Geophysical Research*, *107*(C10), SHE 13-1–SHE 13-14. https://doi.org/10.1029/2000JC000439

JAMSTEC (2015). *Mirai mr14-05 cruise data*. Retrieved from http://www.godac.jamstec.go.jp/darwin/cruise/mirai/mr14-05/e https://doi.org/10.17596/0001861

Kawaguchi, Y., Nishino, S., Inoue, J., Maeno, K., Takeda, H., & Oshima, K. (2016). Enhanced diapycnal mixing due to near-inertial internal waves propagating through an anticyclonic eddy in the ice-free Chukchi Plateau. *Journal of Physical Oceanography*, *46*, 2457–2481. https://doi.org/10.1175/JPO-D-15-0150.1

Kay, J., L'Ecuyer, T., Chepfer, H., Loeb, N., Morrison, A., & Cesana, G. (2016). Recent advances in arctic cloud and climate research. *Current Climate Change Reports*, *2*, 159–169. https://doi.org/10.1007/s40641-016-0051-9

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, *137*(3), 991–1007. https://doi.org/10.1175/2008MWR2556.1

MOSAiC Science Plan Writing Team (2016). *Multidisciplinary drifting observatory for the study of arctic climate science plan*. Retrieved from https://www.iasc.info/images/news/MosaicSciencePlan2016.pdf

Nakanishi, M., & Niino, H. (2006). An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, *119*, 397–407. https://doi.org/10.1007/s10546-005-9030-8

Nishino, S., Kawaguchi, Y., Inoue, J., Yamamoto-Kawai, M., Aoyama, M., Harada, N., & Kikuchi, T. (2019). Do strong winds impact water mass, nutrient, and phytoplankton distributions in the ice-free Canada basin in the fall?. *Journal of Geophysical Research: Oceans*, *125*, e2019JC015428. https://doi.org/10.1029/2019JC015428

Nose, T., Waseda, T., Kodaira, T., & Inoue, J. (2020). Satellite-retrieved sea ice concentration uncertainty and its effect on modeling wave evolution in marginal ice zones. *The Cryosphere*, *14*, 2029–2052. https://doi.org/10.5194/tc-14-2029-2020

Nose, T., Webb, A., Waseda, T., Inoue, J., & Sato, K. (2018). Predictability of storm wave heights in the ice-free Beaufort Sea. *Ocean Dynamics*, *68*, 1383–1402. https://doi.org/10.1007/s10236-018-1194-0

Ono, J., Inoue, J., Yamazaki, A., Dethloff, K., & Yamaguchi, H. (2016). The impact of radiosonde data on forecasting sea-ice distribution along the Northern Sea Route during an extremely developed cyclone. *Journal of Advances in Modeling Earth Systems*, *8*, 292–303. https://doi.org/10.1002/2015MS000552

Orr, A., Phillips, T., Webster, S., Elvidge, A., Weeks, M., Hosking, S., & Turner, J. (2014). Met office unified model high-resolution simulations of a strong wind event in Antarctica. *Quarterly Journal of the Royal Meteorological Society*, *140*, 2287–2297. https://doi.org/10.1002/qj.2296

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., et al. (2017). The weather research and forecasting model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, *98*(8), 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1

Rinke, A., Dethloff, K., Cassano, J., Christensen, J., Curry, J., Du, P., et al. (2006). Evaluation of an ensemble of arctic regional climate models: Spatiotemporal fields during the SHEBA year. *Climate Dynamics*, *26*, 459–472. https://doi.org/10.1007/s00382-005-0095-3

Sato, K., Inoue, J., Kodama, Y.-M., & Overland, J. (2012). Impact of arctic sea-ice retreat on the recent change in cloud-base height during autumn. *Geophysical Research Letters*, *39*, L10503. https://doi.org/10.1029/2012GL051850

Sedlar, J., Tjernström, M., Rinke, A., Orr, A., Cassano, J., Fettweis, X., et al. (2020). Confronting arctic troposphere, clouds, and surface energy budget representations in regional climate models with observations. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD031783. https://doi.org/10.1029/2019JD031783

Shupe, M. D., & Intrieri, J. M. (2004). Cloud radiative forcing of the arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle. *Journal of Climate*, *17*(3), 616–628. https://doi.org/10.1175/1520-0442(2004)017<0616:CRFOTA>2.0.CO;2

Sommerfeld, A., Nikiema, O., Rinke, A., Dethloff, K., & Laprise, R. (2015). Arctic budget study of intermember variability using HIRHAM5 ensemble simulations. *Journal of Geophysical Research: Atmosphere*, *120*(18), 9390–9407. https://doi.org/10.1002/2015JD023153

Sotiropoulou, G., Sedlar, J., Forbes, R., & Tjernström, M. (2015). Summer arctic clouds in the ECMWF forecast model: An evaluation of cloud parametrization schemes. *Quarterly Journal of the Royal Meteorological Society*, *142*, 387–400. https://doi.org/10.1002/qj.2658

Sundqvist, H., Berge, E., & Kristjánsson, J. (1989). Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Monthly Weather Review*, *117*, 1641–1657. https://doi.org/10.1175/1520-0493(1989)117<1641:CACPSW>2.0.CO;2

Taylor, P. C., Boeke, R. C., Li, Y., & Thompson, D. W. J. (2019). Arctic cloud annual cycle biases in climate models. *Atmospheric Chemistry and Physics*, *19*(13), 8759–8782. https://doi.org/10.5194/acp-19-8759-2019

Tedesco, M., & Fettweis, X. (2020). Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet. *The Cryosphere*, *14*(4), 1209–1223. https://doi.org/10.5194/tc-14-1209-2020

Tompkins, A. (2002). A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *Journal of the Atmospheric Sciences*, *59*, 1917–1942. https://doi.org/10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2

Uttal, T., Curry, J. A., McPhee, M. G., Perovich, D. K., Moritz, R. E., Maslanik, J. A., et al. (2002). Surface heat budget of the Arctic Ocean. *Bulletin of the American Meteorological Society*, *83*(2), 255–276. https://doi.org/10.1175/1520-0477(2002)083<0255:SHBOTA>2.3.CO;2

Vihma, T., Screen, J., Tjernström, M., Newton, B., Zhang, X., Popova, V., et al. (2016). The atmospheric role in the arctic water cycle: A review on processes, past and future changes, and their impacts. *Journal of Geophysical Research: Biogeosciences*, *121*(3), 586–620. https://doi.org/10.1002/2015JG003132

Wilson, D., Bushell, A., Kerr-Munslow, A., Price, J., & Morcrette, C. (2008). PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. *Quarterly Journal of the Royal Meteorological Society*, *134*, 2093–2107. https://doi.org/10.1002/qj.333

Wood, R., & Bretherton, C. (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of Climate*, *19*, 6425–6432. https://doi.org/10.1175/jcli3988.1

Wyser, K., Jones, C., Du, P., Girard, É., Willén, U., Cassano, J., et al. (2008). An evaluation of arctic cloud and radiation processes during the SHEBA year: Simulation results from eight Arctic regional climate models. *Climate Dynamics*, *30*, 203–223. https://doi.org/10.1007/s00382-007-0286-1

Yamazaki, A., Inoue, J., Dethloff, K., Maturilli, M., & König-Langlo, G. (2015). Impact of radiosonde observations on forecasting summertime arctic cyclone formation: Arctic radiosondes for cyclone forecast. *Journal of Geophysical Research: Atmospheres*, *120*, 3249–3273. https://doi.org/10.1002/2014JD022925

Yao, B., Liu, C., Yin, Y., Liu, Z., Shi, C., Iwabuchi, H., & Weng, F. (2020). Evaluation of cloud properties from reanalyzes over East Asia with a radiance-based approach. *Atmospheric Measurement Techniques*, *13*, 1033–1049. https://doi.org/10.5194/amt-13-1033-2020