Improved weather forecasting using neural network emulation for radiation parameterization

Hwan-Jin Song[1] and Soonyoung Roh[1]

[1]National Institute of Meteorological Sciences, Korea Meteorological Administration

November 22, 2022

Abstract

In this study, a neural network (NN) emulator for radiation parameterization was developed for the use of an operational weather forecasting model in the Korea Meteorological Administration. The development of the NN emulator was based on large-scale training sets and 96 categories (longwave–shortwave, months, land–ocean, and clear–cloud). As the radiation parameterization was replaced by the NN emulator, a 60-fold speedup for the radiation process was achieved, with a decrease of 87.26% in the total computation time. The accuracy of the NN emulator was strictly evaluated through comparison with the results obtained from the infrequent use of the original radiation scheme with the same computational cost. The mean errors of the NN radiation emulator were significantly reduced by 21–34% compared with the infrequent method. The combination of using the NN radiation emulator and applying it infrequently provided an additional speedup of up to 36-fold, corresponding to 2180 times speedup compared with the control run, without a significant reduction in accuracy. The optimized structure for the radiation emulator designed in this study also showed universal robustness even in the use of limited training sets with incomplete coverage. In conclusion, the NN radiation emulator in this study provides benefits regarding both accuracy and computational cost, making it useful for improving weather forecasting modeling.

Improved weather forecasting using neural network emulation for radiation parameterization

Hwan-Jin Song and Soonyoung Roh*

National Institute of Meteorological Sciences, Korea Meteorological Administration, Jeju-do, Republic of Korea

Submitted to Journal of Advances in Modeling Earth Systems

(7 May 2021)

** Corresponding author's address*
Soonyoung Roh
National Institute of Meteorological Sciences,
63568, Seogwipo-si, Jeju-do, Republic of Korea
E-mail: cozet109@gmail.com
Phone: +82-64-780-6757

Abstract

In this study, a neural network (NN) emulator for radiation parameterization was developed for the use of an operational weather forecasting model in the Korea Meteorological Administration. The development of the NN emulator was based on large-scale training sets and 96 categories (longwave–shortwave, months, land–ocean, and clear–cloud). As the radiation parameterization was replaced by the NN emulator, a 60-fold speedup for the radiation process was achieved, with a decrease of 87.26% in the total computation time. The accuracy of the NN emulator was strictly evaluated through comparison with the results obtained from the infrequent use of the original radiation scheme with the same computational cost. The mean errors of the NN radiation emulator were significantly reduced by 21–34% compared with the infrequent method. The combination of using the NN radiation emulator and applying it infrequently provided an additional speedup of up to 36-fold, corresponding to 2180 times speedup compared with the control run, without a significant reduction in accuracy. The optimized structure for the radiation emulator designed in this study also showed universal robustness even in the use of limited training sets with incomplete coverage. In conclusion, the NN radiation emulator in this study provides benefits regarding both accuracy and computational cost, making it useful for improving weather forecasting modeling.

Keywords: neural network, radiation, emulator, speedup, WRF, RRTMG

Plain Language Summary

Numerical weather forecasting model requires a lot of computational resources based on supercomputers. In an attempt to significantly reduce computational cost, emulator studies have been actively conducted. The ultimate goal of most emulator studies is to reduce the computational cost while replicating the accuracy of the control run. However, this study suggests a novel approach that can improve both accuracy and speed for radiation emulator. This is possible through comparison with the operational method based on the infrequent use of the radiation scheme. To do this, large-scale datasets and a highly classified training strategy were adopted. We also suggest that the combination of an NN radiation emulator and its infrequent use can produce significant speedup by as much as thousands of times while maintaining accuracy. This study will shed light on a new research direction for the development of radiation emulator based on numerical weather and climate models.

1. Introduction

Recent advances in artificial intelligence (AI) techniques have provided challenges beyond developing theory-based numerical weather–climate prediction models (Reichstein et al., 2019; Hutson, 2020). The post-processing of numerical model outputs is the most typical example of AI application to numerical weather and climate forecasting (Krasnopolsky and Lin, 2012; Rasp and Learch, 2018; Ham et al., 2019; Scher and Messori, 2019). The application of AI to data assimilation in numerical weather prediction (NWP) models (Boukabara et al., 2019; Cho et al., 2020; Lee et al., 2020) is also an important in the field of weather forecasting. The development of AI emulators (or surrogate models) to replace various processes within the NWP model has been recently attempted for applications such as dynamics (Dueben and Bauer, 2018; Scher, 2018), representation of sub-grid processes with convective parameterization (Brenowitz and Bretherton, 2018; Gentine et al., 2018; O'Gorman and Dwyer, 2018; Rasp et al., 2018; Yuval and O'Gorman, 2020), planetary boundary layer (Wang et al., 2019), and radiation (Chevallier et al., 1998, Chevallier et al., 2000, Krasnopolsky et al., 2005; Krasnopolsky et al., 2010; Belochitski et al., 2011; Krasnopolsky et al., 2012; Pal et al., 2019; Roh and Song, 2020; Belochitski and Krasnopolsky, 2021).

This study focuses on emulator studies for radiation processes. Although longwave (LW) and shortwave (SW) processes can be elaborately represented by a line-by-line radiative model (e.g., Clough et al., 1992; 2005), fast radiative transfer models with corrected-k methods (Mlawer et al., 1997; Iacono et al., 2008; Pincus et al., 2019) are commonly used, owing to the benefit of computational cost. Early studies based on shallow neural network (NN) with a single hidden layer were developed in the framework of the radiative transfer model (Chevallier et al., 1998) or its application to data assimilation with respect to the update of the initial field (Chevallier et al., 2000). The radiative transfer for TOVS (RTTOV) has been utilized using multiple linear regressions since 1999, and it has been widely used in data assimilation in the NWP model (Saunders et al., 2018). Recent studies on radiative transfer modeling have extended the application of various AI techniques, including multiple linear regression, deep neural network (DNN), adaptive network-based fuzzy inference systems, and convolution neural network (CNN) for radiation processes over a clear sky (Liu et al., 2020; Ukkonen et al., 2020; Bilgic and Mert, 2021; Veerman et al., 2021) and 3-dimensional cloud radiative effects (Meyer et al., 2021). As these studies do not utilize repetition by time integration, such as in the numerical forecast model, errors by emulation do not accumulate.

The radiation process also plays a key role in numerical weather–climate prediction models. In numerical models, it takes the form of radiation parameterization. However, the radiation treatment is still expensive for the numerical model; thus, efforts have been made to replace the radiation parameterization with the NN approximation (named the radiation emulator). This presents a challenge because the approximation error of the radiation emulator can rapidly increase during the long-term integration process inside the numerical model (Krasnopolsky et al., 2008). Nevertheless, the improvement in computational cost can significantly accelerate the numerical forecast model, demanding expensive supercomputer calculations. This acceleration can be very beneficial in the event of severe weather, in which urgent notification can lead to the protection of human life and property. As a pioneering study, Krasnopolsky et al. (2010) presented an impressive result, showing that the NN emulator for the Rapid Radiative Transfer Model for General Circulation Models (RRTMG; Iacono et al., 2008) parameterization can improve the computational cost by 16–60 fold (LW to SW), approximately 30-fold for a day, in comparison to the RRTMG scheme, resulting in a 20%–25% reduction in the total computational cost for the National Centers for Environmental Prediction (NCEP) Climate Forecast System model. A follow-up study for the

NCEP Global Forecast System by Krasnopolsky et al. (2012) further reported a 20 to 100 (LW to SW) time speedup and 15%–18% reduction in total computational cost. Recently, Roh and Song (2020) focused on cloud-resolving simulations at fine temporal (20 s) and horizontal (250 m) scales, demonstrating a 20–100-fold speedup using NN emulators for the RRTMG-K parameterization (Baek, 2017). This resulted in an 82%–86% reduction in the total computational cost. This is an interesting result, considering that most previous studies were based on climate simulations under coarse temporal and horizontal resolutions. As the long-term and large-scale (e.g., the entire hemisphere) average is typically considered in climate simulations, errors caused by the emulator can be expressed in reduced form.

According to the comparison results for various AI methods, Belochitski et al. (2011) found that LW emulators based on the approximate nearest neighborhood, classification and regression trees, and random forest methods caused increases in root mean square error (RMSE) of 84%, 40%, and 20%, respectively, compared with the NN method with 80 neurons. Pal et al. (2019) reported that a DNN radiation emulator can produce 8–10 fold speedup and 90%–95% accuracy in the Super-Parameterized Energy Exascale Earth System Model (SP-E3SM) from the United States Department of Energy (DOE); however, they did not provide a specific reduction in the total computational cost. Liu et al. (2020) showed that the use of a CNN emulator could reduce the RMSE of clear-sky LW cooling rates by 41%–59%, compared with the DNN emulator with three hidden layers, whereas the CNN resulted in approximately 10-fold slowdown in contrast to the 10.88-fold speedup for the DNN (i.e., the CNN was approximately 100 times slower than the DNN). We can see that the use of deep hidden layers or more complex structures (Pal et al., 2019; Liu et al., 2020) may not always produce better performance compared with the NN with a single hidden layer (in terms of speedup), although it offers a variety of possibilities for optimization. Belochitski and Krasnopolsky (2021) also noted the risk of developing radiation emulators for the DNN pertaining to the control of complexity and nonlinearity.

Another approach to improve the speed of radiation parameterization is to perform the radiation scheme less often than the time step of the NWP model. Although the infrequent approach is popularly adopted in operational weather forecasting, numerical errors can accumulate over time in interaction with other processes (Xu and Randall, 1995; Pauluis and Emanuel, 2004; Pincus and Stevens, 2013). In particular, a fixed radiation process within the time step of radiation parameterization can induce considerable errors around sunrise and sunset when solar radiation changes rapidly. Roh and Song (2020) insisted that the frequent use of NN radiation emulator in the NWP model should provide benefits in both speedup and accuracy compared to the operational method based on infrequent use. The improvement of accuracy by the emulator is quite interesting because most emulator studies aim to mimic the original parameterization, and this cannot overcome the original scheme. In fact, results for 300-neuron emulators showed reductions of 28.7% and 20.5% in RMSE for LW and SW fluxes, respectively, compared with the infrequent method in which the radiation scheme was called every 20th time step. Additionally, the 56 neuron-based emulator results showed benefits of both five-fold greater speedup (i.e., 20→100 times) and reduced RMSEs of 3.6%–22.8%. However, as the results of Roh and Song (2020) were limited to idealized cloud simulations, whether the radiation emulator can offer benefits in both speedup and accuracy has yet to be verified in real weather forecasting.

The Korean Peninsula is a typical area with a unique precipitation mechanism in the world, as part of a humid monsoon environment, leading to lower accuracy of precipitation forecasts for the area (Song and Sohn, 2015; Song and Sohn, 2018, Song et al., 2019; Song et al., 2020). To improve weather forecasting over the Korean Peninsula, we developed an NN radiation emulator for RRTMG-K parameterization under the framework of the Korea Local

Analysis and Prediction System (KLAPS; Kim et al., 2002), which is an operational short-range weather forecast model of the Korea Meteorological Administration (KMA). This study is significant as the first attempt to improve real-time weather forecasting using an NN radiation emulator. In contrast to climate simulation, the accuracy and stability requirements for weather forecasting are very high. This study also addresses new attempts to optimize the structure of the input–output variables and training sets. Furthermore, the accuracy of the developed radiation emulator was evaluated in comparison with the infrequent use of original radiation scheme with the same computational cost. No similar attempt has been demonstrated in the literature for radiation emulators, except the ideal simulation implemented by Roh and Song (2020). Lastly, this study suggests the possibility of additional speedup while maintaining accuracy through the infrequent use of an NN radiation emulator.

2. Data and Methods

The dynamic and physical processes of the current operational KLAPS in the KMA are based on the Advanced Research Weather Research and Forecasting (WRF-ARW) model (Skamarock et al., 2019) version 3.9.1.1, as well as recent physics updates achieved through the development of the Korean Integrated Model (KIM) system (Hong et al., 2018). These updated physics schemes are available for WRF versions later than version 4. In this study, we considered operational configurations of KLAPS, such as the RRTMG-K radiation (Baek, 2017) and WRF double moment 7-Class (WDM7) microphysics (Bae et al., 2019), Shin and Hong planetary boundary layer (Shin and Hong, 2015), KIAPS Simplified Arakawa–Schubert (SAS) cumulus (Kwon and Hong, 2017), Unified Noah land surface model (Tewari et al., 2004), and revised MM5 Monin–Obukhov surface layer (Jiménez et al., 2012). The systemic biases for the WDM microphysics scheme, which were reported by Lei et al. (2020), were corrected in this study. Although KLAPS also includes a local data assimilation system, it was replaced with the ECMWF Reanalysis 5 (ERA5) data (Hersbach et al., 2020) in the

WRF Preprocessing System (WPS). As the data assimilation of KLAPS is not used, we will call the used numerical weather forecasting model "WRF" hereafter. The ERA5 datasets include 37 pressure and single-level data with hourly intervals and $0.25° \times 0.25°$ horizontal resolution. The domain for the WRF simulation consists of 234×282 with a horizontal resolution of 5 km around the Korean peninsula and 39 vertical layers (or 40 levels) up to 50 hPa. The WRF control run was integrated during one day every 20 s for both the time step (dt) and radiation time step (radt). Thus, total simulations are accumulated with 4,320 time steps. This study focuses on the RRTMG-K parameterization, which primarily computes vertical heating rates and radiative fluxes over the LW spectrum with 256 g points for 16 bands and the SW spectrum with 224 g points for 14 bands using a two-stream correlated-k method and optimized Monte Carlo independent column approximation. When the RRTMG-K scheme was used at each time step (i.e., the same as the control run), it was responsible for 88.63% of the total computational cost in the WRF control run. As the computational cost of the SW process during the day is approximately 3.72-fold greater than that of LW, the daytime simulation is 4.72 times slower than that at night. For a similar reason, the computational time increases around the summer solstice with a longer daytime period but decreases around the winter solstice, affected by the change in solar zenith angle.

This study examines 48-day cases that consist of two extreme heavy precipitation cases and two non-precipitating events in each month over the Korean peninsula. Between 2009 and 2018, the two selected precipitation events corresponded to the maximum and second maximum cases for daily precipitation in South Korea, whereas two non-precipitating cases were randomly selected. These cases can represent two extreme polarizing events that are considered difficult problems in machine learning, resulting in various atmospheric conditions despite small subsets. For example, outgoing LW radiation is generally characterized by minimum negative values for heavy precipitation events and maximum negative values for clear areas. The training data based on 48 cases were integrated at each 20 s time step and recorded in 10 min intervals. The training data are produced through internal modification of RRTMG-K and related codes, not in final outputs that may be affected by other processes, to extract accurate input and output data. Training sets are further divided by month, land/ocean, and clear/cloud, in addition to LW and SW. The LW and SW parameterizations are already separated in the original scheme, and SW is only used during the daytime when it is defined as a positive solar zenith angle. The monthly separation confines the climatological monthly range of the input and output variables. Additionally, the physical variable characteristics around the surface can be significantly different between land and ocean. As radiative characteristics over cloud sky are very complicated as compared with those over clear sky, the separation of clear sky (i.e., zero cloud fractions at all levels) and cloud sky is physically reasonable. In fact, the data assimilation of satellite radiances in the NWP model is generally considered clear sky only because satellite data assimilation for cloud areas remains a challenging topic (Hong et al., 2018; Saunders et al., 2018; Hersbach et al., 2020). As previous radiation emulator studies did not distinguish between clear and cloudy areas in the NN training step, cloud fraction profiles were used as inputs, even for clear sky; this is a waste in terms of computational efficiency. In summary, a total of 96 categories (12 months, land/ocean, clear/cloud, and LW and SW) are used in the training sets. Each training set consists of three million input–output pairs (corresponding to 1% of the total data). Such an attempt at optimization is noble, as no similar approach has been attempted in literatures for radiation emulators. Additionally, approximately 1.44 billion data records for LW and SW are used in this study, 720 times more than the 200,000 used in Krasnopolsky et al. (2010; 2012), implying that the NN approximation of the current radiation emulator is more mature. This is essential for achieving operational-level accuracy beyond the research level.

As shown in Table 1, the inputs for the RRTMG-K emulator comprise 193 variables, including the following: vertical pressure, vertical temperature, vertical water vapor mixing ratio, vertical ozone mixing ratio, vertical cloud fraction, longitude, latitude, surface elevation, skin temperature, and surface emissivity (LW only), and solar constant (G) multiplied by the cosine solar zenith angle (cosθ) and surface albedo (SW only). Among the input variables, cloud fraction profiles have a significant influence on the LW and SW radiation processes. For example, strong cooling and heating areas are found above the cloud top for LW and SW, respectively, as shown by Roh and Song (2020). For a clear sky, the number of input variables is decreased to 161 as cloud fraction profiles are excluded. Geographical data, such as longitude, latitude, and elevation, are not included as input variables in the RRTMG-K but are added to realistically reflect the regional characteristics of the NN training. In particular, surface elevation is an important variable that affects actual vertical heights in relation to the terrain-following hybrid sigma pressure vertical coordinate of the WRF model (Skamarock et al., 2019). Other redundant constant variables, such as trace gases and aerosols, as well as microphysics variables, are excluded from the input variables to avoid possible uncertainties that they can cause. A total of 42 output variables for LW and SW were considered, such as heating rate profiles in 39 layers, total sky upward fluxes at the top of the atmosphere (TOA) and the surface, and total downward flux at the surface. Clear sky fluxes are excluded from the outputs because they are idealized clear skies and do not affect other variables. Note that there is no downward LW flux at the TOA, and the downward SW flux at the TOA can be parameterized by G×cosθ. G×cosθ is the primary driver of the SW radiation processes associated with the solar cycle.

The NN software based on a single hidden layer (Krasnopolsky, 2014) was utilized to develop the radiation emulator in this study. Previous radiation emulators (e.g., Krasnopolsky et al., 2005; Krasnopolsky et al., 2010; Belochitski et al., 2011; Roh and Song, 2020;

Belochitski and Krasnopolsky, 2021) have been developed using this software (or similar version for old literatures). Here, no tuning for hyperparameters (e.g., batch size, learning rate, activation function, regularization, normalization for inputs and outputs, and weight initialization) was performed, except the number of neurons. All settings of the hyperparameters used in this study are based on the default configuration of Krasnopolsky (2014); these will remain in future works. For non-linear relationships among given inputs and outputs, neural networks can provide an approximated solution (here, we used a hyperbolic tangent function for the activation function). Owing to NN training, weight and bias coefficients, which relate the inputs to the hidden layer and hidden layer to outputs, are obtained. Finally, the radiation emulator based on the obtained weight and bias coefficients completely replaced the original RRTMG-K parameterization in the WRF simulation (called NN-WRF in this study). When the NN results were applied to the WRF model, NN outputs were not produced over the range of min/max of the training set outputs to prevent errors caused by the extrapolation. As given by Krasnopolsky et al. (2010), numerical complexity can be expressed in the form of $k \times (n + m + 1) + m$, where k, n, and m are the numbers of hidden neurons (neurons hereafter), input variables, and output variables, respectively. Thus, if a large number of neurons are considered, the accuracy may be enhanced; however, negative effects on the computational cost result owing to the proportional relationship between the number of neurons and numerical complexity. We expect that more optimization between hidden layers and neurons based on the DNN in future work can bring further improvements in accuracy; however, this is beyond the scope of this study.

After several empirical tests, 90 neurons were selected to target a 60-fold speedup compared to the original RRTMG-K parameterization. The 60-fold speedup corresponds to approximately twice the speedup of Krasnopolsky et al. (2010). Additional experiments using different neurons may be necessary; however, this is not the principal concern of this study.

The speedup of the radiation emulator was calculated under a single processor configuration during a one-day simulation for 48 cases. In other words, the speedup indicates the ratio of computation time for radiation processes (for LW and SW) when the RRTMG-K code (i.e., module_ra_rrtmg_swk.F in the WRF model) is completely replaced by the NN emulator (new module_ra_rrtmg_swk.F). Notably, the RRTMG-K LW code (module_ra_rrtmg_lwk.F) is a subroutine of the SW code (i.e., the input–output array in the LW code is completely shared with the SW code). In this study, two heavy precipitation and two clear-sky dominant events are considered in each month for the entire training dataset (2009–2018). The difference in speedup between clear-sky dominant and heavy precipitation events is not significant within 1%, in contrast to the 17% speedup for the cloud area in Krasnopolsky et al. (2010). Therefore, fewer input variables are considered in this study for clear sky by excluding the cloud fraction, contributing to a 14% reduction in numerical complexity compared with the cloud area. Therefore, distinguishing between clear and cloudy areas is considered an effective way to develop a radiation emulator. This has not been attempted in previous radiation emulator studies, which completely replace radiation parameterization within the numerical weather–climate models, although it has been commonly adopted in data assimilation studies, such as Chevallier et al. (2000). Consequently, the speedup for the radiation process of 60.90-fold (29.86-fold for night and 78.09-fold for daytime) contributes to the decrease in the total computation time of 87.26% (or eight times speedup) in the WRF model. This reduction is relatively large compared with the 20%–25% reduction in Krasnopolsky et al. (2010), implying that this study can pertain to situations in which radiation parameterization is vital for the entire model.

3. Results

Figure 1 represents the accuracy of the NN training for the LW/SW heating rate and flux in terms of monthly RMSEs. Although the NN training tends to converge an optimized solution for all given input–output pairs, the accuracy may vary depending on the output variables. Vertical heating rates for 39 layers and 3 fluxes are displayed together for LW and SW in the figure. Land/ocean results, as well as clear/cloud results, are combined in these RMSEs. Note that a fractional land area of 45.30% and an annual mean clear area fraction of 35.88% are considered in this study. The RMSE results can be highly affected by the presence of clouds in the training cases. As September is characterized by the lowest cloud area fraction (47.47%), the RMSEs for LW in September are thought to be the smallest of the year. In contrast, larger RMSEs for LW were found in January and December. The RMSEs of SW tend to increase in boreal summer because the solar zenith angle increases toward summer solstice, in contrast to the lower errors of the winter season. As the cloud area fraction in June (51.92%) is relatively lower than that in April–May and July–August (66.07–78.97%), the uncertainty of SW in June is relatively small despite the high solar zenith angle. Figure 2 shows the RMSEs of the training sets on land/ocean, as well as clear/cloudy. Although there is no significant difference in SW between land and ocean, the RMSE for LW over land is 13–16% higher than that over the ocean. It can be understood that the high variability of surface temperature and emissivity over land increases the uncertainty of LW. Among the three types of categories (month, land–sea mask, clear–cloud mask), the separation between clear and cloud areas had the greatest impact on optimizing the training accuracy. The RMSEs of heating rates (fluxes) over the cloud area are approximately 17.58 times and 20.25 times (10.60 times and 22.01 times) larger for LW and SW, respectively, indicating that the NN approximation for clouds is highly uncertain.

Consequently, the mean RMSE results for the LW and SW heating rates (0.46 K day^{-1} and 0.17 K day^{-1}) appear to be slightly improved than those (1.02 K day^{-1} and 0.49 K day^{-1}) in Roh and Song (2020), although the 90 neurons used in this study is smaller than the 300 neurons used in previous studies. As there is no change in the internal parameters of the NN between Roh and Song (2020) and this study, the advanced accuracy can be interpreted as the result of the decreased uncertainty in the coarse horizontal resolution. In other words, the NWP simulations at 5 km are easier than those at 0.25 km. In the 5 km simulation in this study, the RMSEs for LW and SW heating rates appear to be comparable with those from the 100 km results (0.49 K day^{-1} and 0.20 K day^{-1}) of Krasnopolsky et al. (2010) and the 25 km results (0.52 K day^{-1} and 0.26 K day^{-1}) of Krasnopolsky et al. (2012). Considering the smoothing effect of the 25-fold larger grid size, the results for the training and test sets imply more advanced accuracy in this study. It is also the consequence of 720 times more training data and optimization based on 96 categories (months, land/sea, and clear/ cloud). In conclusion, the RMSE results in Figure 1 represent the maximum performance of the NN radiation emulator. In the process of integrating the NWP model, the error can be greatly amplified.

Figure 3 displays the 12 h forecast fields of fluxes and precipitation for the Typhoon SANBA event (September 17, 2012) between the WRF control run (Figures 3a–c) and WRF simulations with NN radiation emulator results (NN-WRF; Figures 3d–f). In this event, the maximum area-averaged daily precipitation over the Korean Peninsula was recorded for 2009–2018. According to Figure 3, areas with low LW upward flux at the top (LWUPT) and high SW upward flux at the top (SWUPT) are found in relation to clouds that are widely distributed around the typhoon. The LWUPT tends to be smaller owing to the lower cloud top temperature in the clouded sky, whereas it is large under a clear sky because surface signals with high temperature are directly delivered at the TOA. In contrast, the SWUPT increased owing to the presence of clouds with respect to the increased reflection of solar radiation. Despite the 12 h forecast results of 2160 applications with a radiation time step of 20 s, the NN-WRF results present a similar pattern as those from the WRF control run. However, it fails to accurately predict cloud and precipitation patterns. Considering that the NWP results are very sensitive to initial perturbations, this problem is thought to be challenging to overcome using the NN radiation emulator. Figure 4 is an example of a clear-sky dominant non-precipitation case as well as an autumnal equinox date with the same length of night and day (September 23, 2011). In the absence of clouds, LWUPT and SWUPT tend to be determined primarily by surface temperature, surface emissivity (LW only), and surface albedo (SW only). For such a case, the NN approximation can be quite accurate. We can confirm that the NN-WRF can more accurately simulate radiation processes for clear cases than cloud cases with a high degree of uncertainty.

Statistical results for a total of 48 cases are given in Figure 5 as the form of RMSE distribution at each 5 km grid and hourly scale. The spatial distribution of RMSEs was obtained by comparing the NN-WRF results with the WRF control run. Note that the current NN radiation emulator corresponds to a 60-fold speedup compared with the original RRTMG-K parameterization. Therefore, the accuracy of NN-WRF should be compared with the 60-fold reduced use of the RRTMG-K scheme (hereafter, WRF60), providing the same computation cost for a fair comparison. That is, the radiation scheme is used each time step (20 s) for the WRF control run and NN-WRF but is applied at an interval of 1200 s for the WRF60. In general, the RMSE results for LW and SW fluxes over land are much larger than those over ocean in relation to a smoother property over the sea surface. Note that the skin temperature over the ocean is the sea surface temperature, which is not coupled with atmospheric simulation in this study. Compared with WRF60, NN-WRF produces more realistic distributions with lower RMSEs for the LW flux, SW flux, and skin temperature (Figure 5). The remaining high-RMSE areas appear to be uncertainties induced by clouds. As the information of location (longitude, latitude, and elevation) was also utilized as an input variable during the NN rain, it is difficult to identify regional bias in the RMSE distribution, except for high mountain areas over the Kaema Plateau and the northeastern part of China.

For the high mountain region, large variability at the surface can affect much of the lower troposphere, and it can lead to a large error.

The evaluation results for a total of 48 cases are given in the form of a time series in Figure 6. All simulations were integrated over one day from midnight (00:00 LST), and thus, the NN radiation emulators were applied in the order of nighttime before sunset (LW), LW and SW during the daytime, and nighttime after sunrise (LW). The range of the one-day forecast is sufficient for the use of the short-range forecast in KLAPS. The accuracy of LW flux and precipitation tends to be reduced in the latter part of the prediction, whereas the SW flux and skin temperature results are characterized by large errors during the daytime with respect to the diurnal cycle of the sun. During the initial period between 2 h and 5 h, WRF60 produced better results for LW flux and skin temperature than NN-WRF. In this regime, the infrequent use of radiation schemes is thought to not be critical for reducing the accuracy. In the case of WRF60, errors resulting from the infrequent use of the radiation scheme are thought to be amplified in the latter part of the forecast after 6 h, whereas the increase in error is relatively limited in the NN-WRF. After 6 h, the RMSEs of NN-WRF for fluxes and skin temperatures were much lower than those of WRF60. This difference after 6 h could be due to the additional use of SW during the daytime. Overall, the NN-WRF produced lower RMSEs by 32–34% for LW and SW fluxes and 21% for skin temperature compared with WRF60 (Table 2). The improvements in accuracy are the largest by 48% in LWUPB. Notably, the skin temperature is not a direct output variable in the RRTMG-K scheme; rather, it is indirectly affected by the heating rate change around the surface. Because precipitation is more indirectly affected by thermal changes from the radiation process but is basically produced by microphysics parameterization, the improvement in accuracy for the NN-WRF for precipitation is limited to 4%.

As the NN radiation emulator and infrequent use of radiation schemes are independent of each other for the speedup of the radiation process, the combination of the NN radiation emulator and its infrequent use could enable further improvements in computational speed. Figure 7 shows the statistical results further showing total RMSEs of LW flux, SW flux, skin temperature, and precipitation according to the combination of NN radiation emulator and infrequent method by factors of 3, 6, 18, 30, and 36 (NN-WRF3, NN-WRF6, NN-WRF18, NN-WRF30, and NN-WRF36). Note that the WRF60 and NN results are the same as the RMSEs shown in Table 2. Because the WRF60 and NN radiation results were 60 times faster than those of the WRF control run, the experiments for NN-WRF3, NN-WRF6, NN-WRF18, NN-WRF30, and NN-WRF36 correspond to 180, 360, 1080, 1800, and 2160 times faster than the WRF control run. Because numerical errors from the NN radiation emulator and its infrequent use are added and accumulated, errors are generally expected to increase. However, the infrequent use of radiation emulators can partly contribute to the reduction in the fundamental error of the radiation emulator because the frequency of using the emulator is a major factor that can amplify the error. Thus, for the combination case of radiation emulator usage and infrequent method, the increase in numerical error is thought to have been partially compensated and, thus, relatively limited. Consequently, Figure 7 shows that the combination of the NN radiation emulator and infrequent method produces both additional speedup (from 180 to 2160 times) and accuracy improvements. The RMSE errors for LW flux, SW flux, skin temperature, and precipitation were improved by 13–31%, 13–34%, 4–21%, and up to 3%, respectively. Because the final goal in most emulator studies is an accurate reproduction of the control run, with improvements in speed, the emulator cannot exceed the accuracy of the control run. However, this study suggests the possibility of further improvement in both accuracy and speed in relation to radiation emulator frequency. This concept was first proved in an ideal case by Roh and Song (2020), and this study further contributes to the robust demonstration of NWP results based on various real cases. This benefit in relation to accuracy and speed can contribute greatly to improving the operational NWP system. In conclusion, this study will shed light on new research directions on the development of radiation emulators in terms of their frequency of use.

The development of a universal radiation emulator is a challenging topic because radiation emulators have fundamental uncertainty in relation to their dependency on training sets (Boukabara et al., 2020; Belochitski and Krasnopolsky, 2021). This study further investigates the possible impact of training sets on the current radiation emulator results. Herein, we consider new training sets independently from one year (2019) and the same validation sets based on 48 cases used in the previous analysis. Notably, the 24 cases used in this study were for extreme events with maximum and second maximum daily precipitation in each month for the period of 2009–2018. Thus, training sets based on one year (2019) will not be sufficiently representative of extreme flood events at the 10 y scale. Krasnopolsky et al. (2008) also noticed that these far-corner events with rare frequencies can induce significant errors in the NN radiation emulator. We expect such a representation error with extreme values to be mitigated when more training cases and proper sampling procedures are considered. However, as the NN software used in this study (Krasnopolsky, 2014) did not provide parallel processing for fast training, such as batch size, several days were consumed for each training set when three million training sets were considered. Consequently, it was difficult to use training sets of much greater than three million. In the future, this will need to be technically improved with a more advanced training procedure based on a graphics processing unit (GPU). Nevertheless, Figure 8 shows that the monthly RMSE distributions for the test sets (48 cases between 2009 and 2018) appear to be relatively reasonable. Notably, the structure of the input–output variables and the number of training sets (3 million × 96 sets) are the same as those in Figure 1. In relation to the lack of representation for new training sets based on 2019, the mean RMSEs for the test sets (Figure 8) are 28%–33% larger than those in Figure 1. The uncertainty with representation error can be amplified more when it is applied to the WRF simulation, and the validation results in Figure 9 provide a response to this concern. Compared with Figure 7, the mean RMSEs for the NN radiation emulator are actually increased by 25%, 18%, 21% for LW flux, SW flux, and skin temperature, respectively. However, these results still maintained relatively low RMSEs than WRF60, by 15% for LW flux, 22% for SW flux, and 5% for skin temperature. No significant improvement was observed for precipitation, which can be regarded as an indirect output. The effects of using the NN radiation emulator infrequently were also examined. For a speedup of up to 36 fold, lower RMSEs for LW and SW fluxes were maintained compared with WRF60. For skin temperature, the maximum available speedup while maintaining a lower RMSE is slightly reduced by the 18-fold speedup. These radiation emulator results suggest that improvements in both accuracy and speed can be robustly confirmed, even when training sets with incomplete coverage are used. Despite the issue of representation, the optimized input–output structure and categorized training strategy for the radiation emulator in this study are considered to have contributed to maintaining reasonable performance for the validation set, including extreme events. However, for future work, we expect that the consideration of more large-scale training sets and appropriate sampling techniques would produce better performance for radiation emulators by reducing the error due to representation.

4. Summary and Conclusions

To improve the accuracy and speed of the radiation process, the RRTMG-K radiation emulator was developed in this study for the use of the operational KLAPS (or WRF) model, which is a mesoscale weather forecasting model in the KMA. NN training with 90 neurons was performed for large-scale training sets, which consisted of 161 (clear area) or 193 (cloud area) inputs, 42 outputs, and 3 million data points for 96 categories (LW and SW, 12 months, land/ocean, clear/cloudy). We considered a 48-day case, which consisted of two extreme heavy precipitation cases and two non-precipitating cases for each month over the Korean peninsula. The NN training provided weight and bias coefficients, which were inserted into the radiation emulator within the WRF model. Consequently, the RRTMG-K parameterization was completely replaced by the NN radiation emulator. The WRF simulations for both the control run and emulator were integrated for one day with a time step and radiation time step (radt) of 20 s under the domain over the Korean peninsula, with a grid composed of 234 × 282 points representing a 5 km horizontal resolution and 39 vertical layers. The developed emulator based on 90 neurons produced a 60-fold speedup in comparison with the RRTMG-K scheme, resulting in a decrease of 87.26% in the total computation time in the WRF model.

Among the 96 categories for NN training, the separation between clear and cloudy areas greatly contributed in optimizing the speed and accuracy of the emulators. The RMSEs over the cloud area in the NN training were 10.60–22.01 times larger than those over the clear sky. The mean RMSE results for the training sets were 0.46 K day^{-1} and 0.17 K day^{-1} (3.34 W m^{-2} and 16.34 W m^{-2}) for LW and SW heating rates (LW and SW fluxes), respectively. To evaluate the accuracy, we analyzed the mean RMSEs for 48 cases. The NN-WRF results produced considerably lower RMSEs of 32–34% for LW and SW fluxes and 21% for skin temperature compared with the WRF60, corresponding to the same 60-fold speedup based on the infrequent use of the radiation scheme. Furthermore, the combination of the NN radiation emulator and the infrequent method exhibited both accuracy and speed improvements compared with the operational method (WRF60). For example, the infrequent use of the NN radiation emulator produced results 180, 360, 1080, 1800, and 2160-fold faster than the control run, and the RMSEs from the emulators were still lower than those of the WRF60.

Finally, this study examined the impact of representation errors on training sets to develop a universal radiation emulator. The first year of 2019 was considered an independent new training period; however, 48 cases were considered in validation sets, including extreme heavy precipitation events in the 10-year (2009–2018) scale. Although the resulting RMSEs increased by 18–25% (except for precipitation) in relation to the imperfect coverage of training sets, they were still more accurate than those obtained by the operational method (WRF60). Despite the representation error, the optimized input–output structure and categorized training strategy in this study were considered to have contributed in maintaining reasonable performance, even for extreme flood events.

We emphasize that the radiation emulator developed in this study is a first attempt to improve meso-scale weather forecasting based on real cases in terms of both accuracy and speedup; most previous radiation emulator studies were based on climate simulations, and their ultimate goals were confined only to increasing the speedup. Additionally, the categorized NN training (months, land/ocean, clear/cloudy) deserved recognition for its novelty in applying the NN radiation emulator to operational NWP systems beyond the research level. The combination of NN radiation emulator and its infrequent application also provides remarkable results, contributing to a speedup of greater than 1000-fold, which was not reported in previous radiation emulator studies associated with numerical weather–climate models. As the computational speed of the NN radiation emulator, along with its infrequent use, is fast enough, the use of a more complex structure for the radiation emulator is also possible. This sheds light on a new research direction for the radiation emulator, considering that the current development of radiation emulators with numerical weather–climate models has been limited by mostly NN (partially DNN) due to computational cost. However, we acknowledge the limitations of this study such as the use of various machine learning techniques in addition to the NN based on the single hidden layer, the optimization of hyper-parameters, and the use of more large-scale training sets and advanced sampling techniques. Consideration of these issues will lead to improved performance of the radiation emulator in the future.

Acknowledgements

We are extremely grateful to Vladimir Krasnopolsky and Alexei Belochitski (NOAA) for providing the NN software used in this study and their sincere advice at the beginning of the study. We also acknowledge the dedicated efforts of Vijay Tallapragada and Hyesook Lee for supporting international joint research between NOAA and KMA in 2018 to develop physics emulators for numerical models. This study used the NN sorftware provided by doi:10.7289/v5qr4v2z (Krasnopolsky, 2014). The WRF model used in this study was obtained from https://www2.mmm.ucar.edu/wrf/users. This work was funded by the KMA Research and Development Program "Development of AI techniques for weather forecasting" under Grant (KMA2021-00121).

References

Bae, S. Y., S.-Y. Hong, and W.-K. Tao (2019), Development of a single-moment cloud microphysics scheme with prognostic hail for the Weather Research and Forecasting (WRF) model, *Asia-Pacific Journal of Atmospheric Sciences,* 55, 233–245, doi:10.1007/s13143-018-0066-3.

Baek, S. (2017), A revised radiation package of G-packed McICA and two-stream approximation: Performance evaluation in a global weather forecasting model, *Journal of Advances in Modeling Earth Systems,* 9, 1628–1640, doi:10.1002/2017MS000994.

Belochitski, A., P. Binev, R. DeVore, M. Fox-Rabinovitz, V. Krasnopolsky, and P. Lamby (2011), Tree approximation of the long wave radiation parameterization in the NCAR CAM global climate model, *Journal of Computational and Applied Mathematics*, 236, 447–460, doi:10.1016/j.cam.2011.07.013.

Belochitski, A., and V. Kranopolsky (2021), Robustness of neural network emulations of radiative transfer parameterizations in a state-of-the-art General Circulation Model, *Atmosphere,* https://arxiv.org/abs/2103.07024z (in revision).

Bilgic, H. H., and I. Mert (2021), Comparison of different techniques for estimation of incoming longwave radiation, *International Journal of Environmental Science and Technology,* 18, 601–618, doi:10.1007/s13762-020-02923-6.

Boukabara, S.-A., V. Krasnopolsky, J. Q. Stewart, E. S. Maddy, N. Shahroudi, and R. N. Hoffman (2019), Leveraging modern artificial intelligence for remote sensing and NWP: Benefits and challenges, *Bulletin of the American Meteorological Society,* 100, ES473–ES491, doi:10.1175/BAMS-D-18-0324.1.

Boukabara, S.-A., V. Krasnopolsky, S. G. Penny, J. Q. Stewart, and A. McGovern (2020), Outlooks for exploiting artificial intelligence in the earth and environmental sciences, *Bulletin of the American Meteorological Society,* doi:10.1175/BAMS-D-20-0031.1.

Brenowitz, N. D., and C. S. Bretherton (2018), Prognostic validation of a neural network unified physics parameterization, *Geophysical Research Letters,* 45, 6289–6298. doi:10.1029/2018GL078510.

Chevallier, F., F. Chéruy, N. A. Scott, and A. Chédin (1998), A neural network approach for a fast and accurate computation of a longwave radiative budget, *Journal of Applied Meteorology,* 37, 1385–1397, doi:10.1175/1520-0450(1998)037.

Chevallier, F., J.-J., Morcrette, F. Chéruy, and N. A. Scott (2000), Use of a neural-network-based long-wave radiative-transfer scheme in the ECMWF atmospheric model, *Quaterly Journal of the Royal Meteorological Society,* 126, 761–776, doi:10.1002/qj.49712656318.

Cho, D., C. Yoo, J. Im, and D.-H. Cha (2020), Comparative assessment of various machine learning-based bias correction methods for numerical weather prediction model forecasts of extreme air temperatures in urban areas, *Earth and Space Sciences,* 7, e2019EA000740, doi:10.1029/2019EA000740.

Clough, S. A., M. J. Iacono, and J.-L. Moncet (1992), Line-by-line calculation of atmospheric fluxes and cooling rates: Application to water vapor, *Journal of Geophysical Research,* 97, 15761–15785, doi:10.1029/92JD01419.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown (2005), Atmospheric radiative transfer modeling: a summary of the AER codes, *Journal of Quantitative Spectroscopy and Radiative Transfer,* 91, 233–244, doi:10.1016/j.jqsrt.2004.05.058.

Dueben, P. D., and P. Bauer (2018), Challenges and design choices for global weather and climate models based on machine learning, *Geoscientific Model Development,* 11, 3999–4009, doi:10.5194/gmd-11-3999-2018.

Gentine, P., M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis (2018), Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters,* 45, 5742–5751, doi:10.1029/2018GL078202.

Ham, Y.-G., J.-H. Kim, and J.-J. Luo (2019), Deep learning for multi-year ENSO forecasts, *Nature,* 573, 568–572, doi:10.1038/s41586-019-1559-7.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. Rosnay, I. Rozum, F. Vamborg, S., Villaume, and J.-N. Thépaut (2020), The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society,* 146, 1999– 2049. doi:10.1002/qj.3803.

Hong, S.-Y., Y. C. Kwon, T.-H. Kim, J.-E. Kim, S.-J. Choi, I.-H., Kwon, J. Kim, E.-H. Lee, R.-S. Park, and D.-I. Kim (2018), The Korean Integrated Model (KIM) system for weather forecasting, *Asia-Pacific Journal of Atmospheric Sciences,* 54, 267–292, doi:10.1007/s13143-018-0028-9.

Hutson, M. (2020), AI shortcuts speedup simulations by billions of times, *Science,* 367, 6479, doi:10.1126/science.367.6479.728.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *Journal of Geophysical Research,* 113, D13103, doi:10.1029/2008JD009944.

Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante (2012), A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, 140, 898–918, doi:10.1175/MWR-D-11-00056.1.

Kim, Y.-S., O.-R. Park, and S.-O. Hwang (2002), Realtime operation of the Korea Local Analysis and Prediction System at METRI, *Asia-Pacific Journal of Atmospheric Sciences,* 38, 1–10.

Krasnopolsky, V. M., M. S. Fox-Rabinovitz, and D. V. Chalikov (2005), New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of longwave radiation in a climate model, *Monthly Weather Review*, 133, 1370–1383, doi:10.1175/MWR2923.1.

Krasnopolsky, V. M., M. S. Fox-Rabinovitz H. L. Tolman, and A. A. Belochitski (2008), Neural network approach for robust and fast calculation of physical processes in numerical environmental models: Compound parameterization with a quality control of larger errors, *Neural Networks,* 21, 535–543, doi:10.1016/j.neunet.2007.12.019.

Krasnopolsky, V. M., M. S. Fox-Rabinovitz, Y. T. Hou, S. J. Lord, and A. A. Belochitski, (2010), Accurate and fast neural network emulations of model radiation for the NCEP coupled Climate Forecast System: Climate simulations and seasonal predictions, *Monthly Weather Review*, 138, 1822–1842, doi:10.1175/2009MWR3149.1

Krasnopolsky, V. M., and Y. Lin (2012), A neural network nonlinear model ensemble to improve precipitation forecasts over continental US, *Advances in Meteorology,* 2012, 649450, doi:10.1155/2012/649450.

Krasnopolsky, V. M., A. A. Belochitski, Y. T. Hou, S. J. Lord, and F. Yang (2012), Accurate and fast neural network emulations of long and short wave radiation for the NCEP

Global Forecast System Model, NCEP/NWS, NOAA, office note 471, https://www.emc.ncep.noaa.gov/officenotes/newernotes/on471.pdf.

Krasnopolsky, V. M. (2014), NCEP neural network training and validation system: Brief description of NN background and training software, Environment Modeling Center, NCEP/NWS, NOAA, doi:10.7289/v5qr4v2z.

Kwon, Y. C., and S. Hong (2017), A mass-flux cumulus parameterization scheme across gray-zone resolutions, *Monthly Weather Review*, 145, 583–598, doi:10.1175/MWR-D-16-0034.1.

Lee, A., B.-J. Sohn, E. Pavelin, Y. Kim, H.-S. Kang, R. Saunders, and Y.-C. Noh (2020), Assessment of cloud retrieval for IASI 1D-var cloud-sky assimilation and improvement with an ANN approach, *Weather and Forecasting,* 35, 1363–1380, doi:10.1175/WAF-D-19-0218.1.

Lei, H., J. Guo, D. Chen, and J. Yang (2020), Systematic bias in the prediction of warm-rain hydrometeors in the WDM6 microphysics scheme and modifications, *Journal of Geophysical Research: Atmospheres,* 125, e2019JD030756, doi:10.1029/2019JD030756.

Liu, Y., R. Caballero, and J. M. Monteiro (2020), RadNet 1.0: exploring deep learning architectures for longwave radiative transfer, *Geoscientific Model Development*, 13, 4399–4412, doi:10.5194/gmd-13-4399-2020.

Meyer, D., R. J. Hogan, P. D. Dueben, and S. L. Mason (2021), Machine learning emulation of 3D cloud radiative effects. Journal of Advances in Modeling Earth Systems, https://arxiv.org/abs/2103.11919 (in revision)

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *Journal of Geophysical Research,* 102, 16663–16682, doi:10.1029/97JD00237.

O'Gorman, P. A. and J. G. Dwyer (2018), Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events, *Journal of Advances in Modeling Earth Systems*, 10, 2548–2563. doi:10.1029/2018MS001351.

Pauluis, O., and K. Emanuel (2004), Numerical instability resulting from infrequent calculation of radiative heating, *Monthly Weather Review,* 132, 673–686, doi:10.1175/1520-0493(2004)132.

Pal, A., S. Mahajan, and M. R. Norman (2019), Using deep neural networks as cost-effective surrogate models for Super-Parameterized E3SM radiative transfer, *Geophysical Research Letters,* 46, 6069–6079, doi:10.1029/2018GL081646.

Pincus, R., and B. Stevens (2013), Paths to accuracy for radiation parameterizations in atmospheric models, *Journal of Advances in Modeling Earth Systems,* 5, 255–233, doi:10.1002/jame.20027.

Pincus, R., E. J. Mlawer, and J. S. Delamere (2019), Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models, *Journal of Advances in Modeling Earth Systems,* 11, 3087–3089, doi:10.1029/2019MS001621.

Rasp, S., and S. Lerch (2018), Neural networks for postprocessing ensemble weather forecasts, *Monthly Weather Review*, 146, 3885–3990, doi:10.1175/MWR-D-18-0187.s1.

Rasp, S., M. S. Pritchard, and P. Gentine (2018), Deep learning to represent subgrid processes in climate models, *Proceedings of the National Academy of Sciences of the United States of America,* 115, 9684–9689, doi:10.1073/pnas.1810286115.

Reichstein, M., G. Camps-Valls, B. Stevens M. Jung, J. Denzler, N. Carvalhais, Prabhat (2019), Deep learning and process understanding for data-driven earth system science, *Nature,* 566, 195–204, https://doi.org10.1038/s41586-019-0912-1.

Roh, S., and H.-J. Song (2020), Evaluation of neural network emulations for radiation parameterization in cloud resolving model, *Geophysical Research Letters,* 47, e2020GL089444, doi:10.1029/2020GL089444.

Saunders, R., J. Hocking, E. Turner, P. Rayer, D. Rundle, P. Brunel, J. Vidot, P. Roquet, M. Matricardi, A. Geer, N. Bormann, and C. Lupu (2018), An update on the RTTOV fast radiative transfer model (currently at version 12), *Geoscientific Model Development,* 11, 2717–2737, doi:10.5194/gmd-11-2717-2018.

Scher, S. (2018), Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning, *Geophysical Research Letters,* 45, 12616–12622, doi:10.1029/2018GL080704.

Scher, S., and G. Messori (2019), Weather and climate forecasting with neural networks: using GCMs with different complexity as study-ground, *Geoscientific Model Development,* 12, 2797–2809, doi:10.5194/gmd-2019-53.

Shin, H. H., and S. Hong (2015), Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions, *Monthly Weather Review*, 143, 250–271, doi:10.1175/MWR-D-14-00116.1.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, Z. Liu, J. Berner, W. Wang, J. G. Powers, M. G. Duda, D. M. Barker, and X.-Y. Huang (2019), A description of the Advanced Research WRF model version 4, *NCAR Technical Notes*, doi:10.5065/1DFH-6P97.

Song, H.-J., S. Kim, S., Roh, and H. Lee (2020), Difference between cloud top height and storm height for heavy rainfall using TRMM measurements, *Journal of the Meteorological Society of Japan,* 98, 901–904, doi:10.2151/jmsj.2020-044.

Song, H.-J., B. Lim, and S. Joo (2019), Evaluation of rainfall forecast with heavy rain types in the high-resolution Unified Model over South Korea, *Weather and Forecasting,* 34, 1277–1293, doi:10.1175/WAF-D-18-0140.1.

Song, H.-J., and B. J. Sohn (2015), Two heavy rainfall types over the Korean peninsula in the humid East Asian summer environment: A satellite observation study, *Monthly Weather Review*, 143, 363–382, doi:10.1175/MWR-D-14-00184.1.

Song, H.-J., and B. J. Sohn (2018), An evaluation of WRF microphysics schemes for simulating the warm-type heavy rain over the Korean peninsula, *Asia-Pacific Journal of Atmospheric Sciences,* 54, 1–12, doi:10.1007/s13143-018-0006-2.

Tewari, M., F. Chen, W. Wang, J. Dudhia, M. LeMone, K. Mitchell, M. Ek, G. Gayno, J. Weigel, and R. Cuenca (2004), Implementation and verification of the unified Noah land surface model in the WRF model, *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction,* American Meteorological Society, Seattle, WA, 11 – 15 Jan.

Ukkonen, P., R. Pincus, R. J. Hogan, K. P. Nielsen, and E. Kaas (2020), Accelerating radiation computations for dynamical models with targeted machine learning and code optimization, *Journal of Advances in Modeling Earth Systems,* 12, e2020MS002226. doi:10.1029/2020MS002226.

Veerman M. A., R. Pincus, R. Stoffer, C. M. van Leeuwen, D. Podareanu, and C. C. van Heerwaarden (2021), Predicting atmospheric optical properties for radiative transfer computations using neural networks, *Philosophical Transactions of the Royal Society A,* 379, 20200095. doi:10.1098/rsta.2020.0095.

Wang, J., P. Balaprakash, and R. Kotamarthi (2019), Fast domain-aware neural network emulation of a planetary boundary layer parameterization in a numerical weather forecast model, *Geoscientific Model Development,* 12, 4261–4274, doi:10.5194/gmd-12-4261–2019.

Xu, K.-M., and D. A. Randall (1995), Impact of interactive radiative transfer on the macroscopic behavior of cumulus ensembles. Part I: radiation parameterization and sensitivity tests, *Journal of the Atmospheric Sciences,* 52, 785–799, doi:10.1175/1520-0469(1995)052.

Yuval, J., and P. A. O'Gorman (2020), Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, *Nature Communications,* 11, 3295, doi:10.1038/s41467-020-17142-3.

Lists of Tables and Figures

Table 1. List of input and output variables for neural network longwave (LW) and shortwave (SW) emulators. Note that the number of input variables decreases by 161 for the case of clear sky.

Table 2. Evaluation results for WRF simulations with 60-fold reduced usage (WRF60) and NN radiation emulator results (NN-WRF) compared with those from the WRF control run. Statistics in table represent pattern correlation and the root mean squared error (RMSE) at hourly and 5 km scales for 48 cases.

Figure 1. Monthly root mean squared error (RMSE) of (a) LW heating rate, (b) SW heating rate, (c) LW flux, and (d) SW flux for NN training results. Heating rates with 39 layers and three fluxes are displayed together in the figure. Land and ocean results, as well as clear and cloud, are averaged in the figure.

Figure 2. Same as Figure 1 but for land and ocean as well as clear and cloud.

Figure 3. Spatial distributions of 12 h forecast LW and SW upward fluxes at the top of the atmosphere (LWUPT and SWUPT) and hourly precipitation for the Typhoon SANBA event (September 17, 2012) between the WRF control run and NN radiation emulator (NN-WRF).

Figure 4. Spatial distributions of 12 h forecast LW and SW upward fluxes at the top of the atmosphere (LWUPT and SWUPT), and skin temperature for a clear-sky dominant event (September 23, 2011) between the WRF control run and NN radiation emulator (NN-WRF).

Figure 5. Root mean square error (RMSE) distributions of LW and SW fluxes and skin temperature between WRF simulations for 60-fold reduced usage (WRF60) and NN radiation emulator (NN-WRF) compared with those from the WRF control run. The average RMSEs from three LW and SW fluxes are presented.

Figure 6. Time series of the root mean square error (RMSE) of (a) LW flux, (b) SW flux, (c) skin temperature, and (d) precipitation between WRF simulations 60-fold reduced usage (WRF60) and NN radiation emulator (NN-WRF) compared with those from the WRF control run.

Figure 7. Statistical results showing the root mean square error (RMSE) of (a) LW flux, (b) SW flux, (c) skin temperature, and (d) precipitation for WRF simulations according to the combination of NN radiation emulator (NN-WRF) and its infrequent usage, compared with the 60-fold reduced usage of the radiation parameterization (WRF60). In the x-axis, e NN-WRF3, NN-WRF6, NN-WRF18, NN-WRF30, NN-WRF36 represent additional speedup improvements of NN radiation emulator through the 3, 6, 18, 30, and 36-fold reduced usage. Those emulator results are 180, 360, 900, 1080, 1800, 2160 times faster than those of the WRF control run, as the NN radiation emulator is already 60 times faster than the control run.

Figure 8. Same as Figure 1 but for results based on NN training from independent one year while keeping the same test sets used in Figure 1.

Figure 9. Same as Figure 7 but for results based on NN training from independent one year while keeping the same validation cases as Figure 7.

Table 1. List of input and output variables for neural network longwave (LW) and shortwave (SW) emulators. Note that the number of input variables decreases by 161 for the case of clear sky.

Inputs	#
Vertical pressure	1–39
Vertical temperature	40–78
Vertical water vapor	79–117
Vertical ozone	118–156
Vertical cloud fraction	157–188
Longitude	189
Latitude	190
Surface Elevation	191
Skin temperature (LW)	192
Surface emissivity (LW)	193
Cosine solar zenith angle multiplied by solar constant (SW)	192
Surface albedo (SW)	193
Outputs	**#**
Vertical total sky heating rate (LW, SW)	1–39
Total sky longwave upward flux at the top (LWUPT)	40
Total sky longwave upward flux at the bottom (LWUPB)	41
Total sky longwave downward flux at the bottom (LWDNB)	42
Total sky shortwave upward flux at the top (SWUPT)	40
Total sky shortwave upward flux at the bottom (SWUPB)	41
Total sky shortwave downward flux at the bottom (SWDNB)	42

Table 2. Evaluation results for WRF simulations with 60-fold reduced usage (WRF60) and NN radiation emulator results (NN-WRF) compared with those from the WRF control run. Statistics in table represent pattern correlation and the root mean squared error (RMSE) at hourly and 5 km scales for 48 cases.

Experiments	WRF60	NN-WRF
Speedup of radiation	60	60.9039
Reduced computation time	87.1528%	87.2642%
LW flux [W m^{-2}]	0.9858, 8.1067	0.9935, 5.5495
LWUPT	0.9695, 10.7497	0.9868, 7.1105
LWUPB	0.9969, 4.4654	0.9992, 2.3017
LWDNB	0.9911, 9.1050	0.9944, 7.2362
SW flux [W m^{-2}]	0.9692, 48.9695	0.9865, 32.1263
SWUPT	0.9584, 61.4614	0.9821, 40.2748
SWUPB	0.9802, 10.9657	0.9904, 7.6515
SWDNB	0.9691, 74.4815	0.9869, 48.4527
Skin temperature [K]	0.9989, 0.5105	0.9993, 0.4018
Precipitation [mm]	0.9408, 0.5315	0.9455, 0.5112

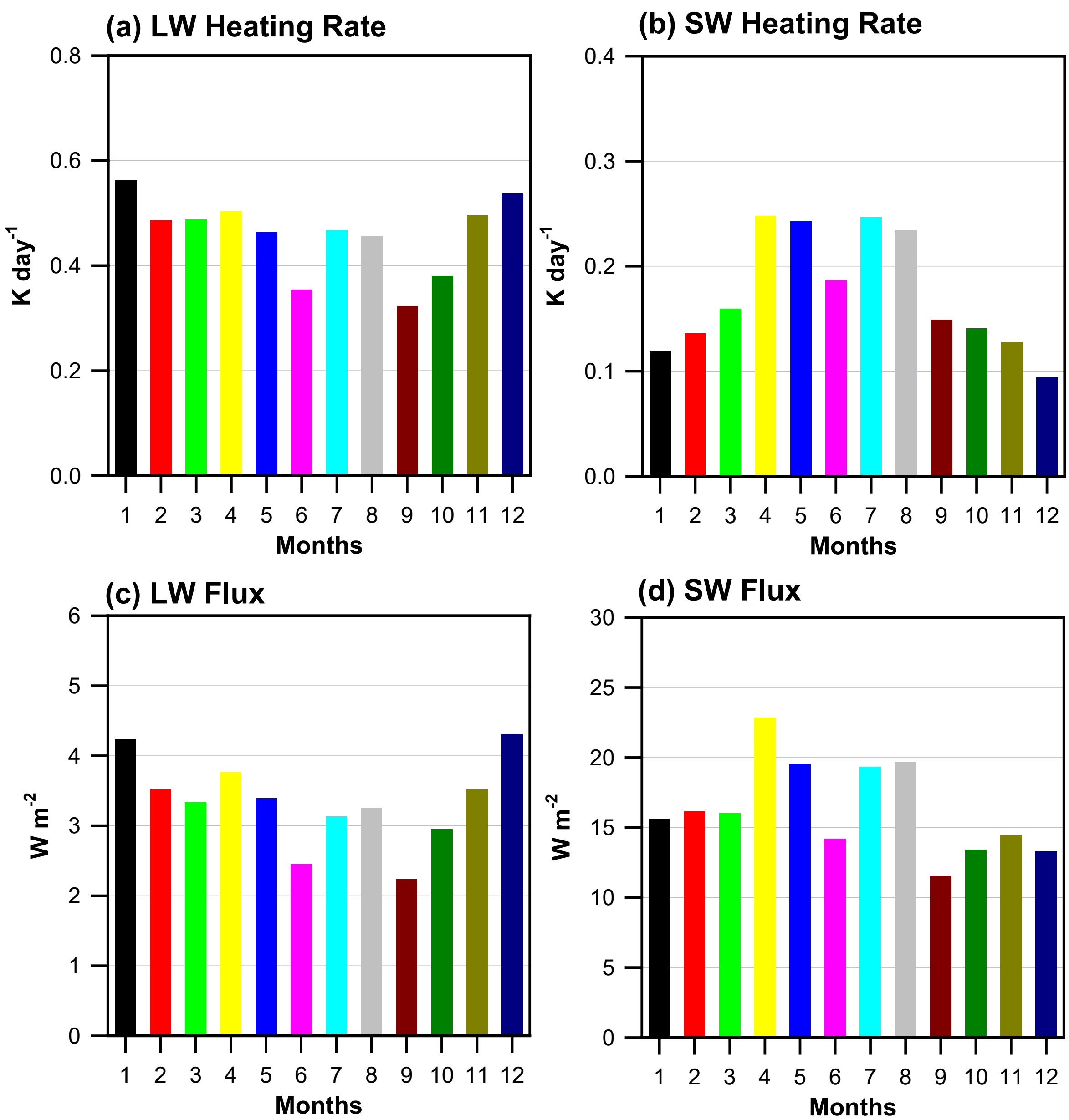

Figure 1. Monthly root mean squared error (RMSE) of (a) LW heating rate, (b) SW heating rate, (c) LW flux, and (d) SW flux for NN training results. Heating rates with 39 layers and three fluxes are displayed together in the figure. Land and ocean results, as well as clear and cloud, are averaged in the figure.

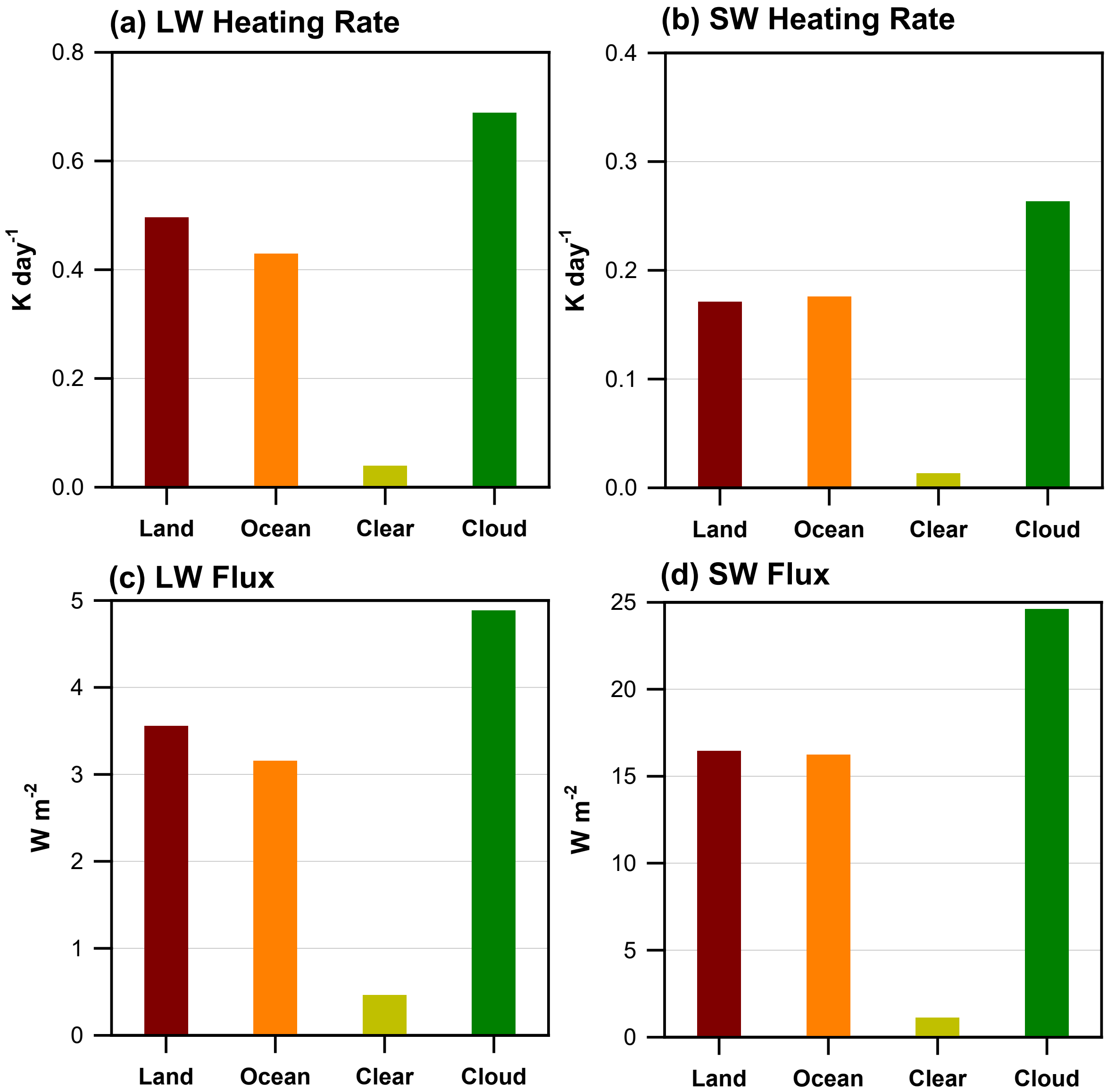

Figure 2. Same as Figure 1 but for land and ocean as well as clear and cloud.

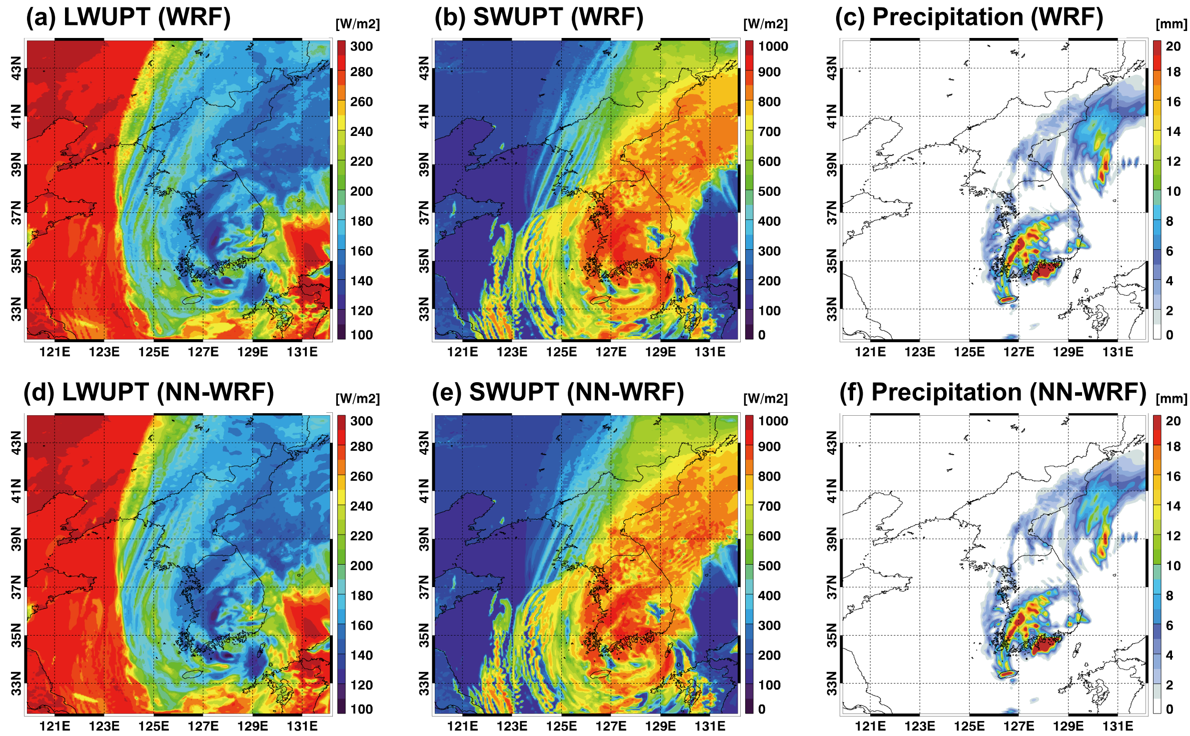

Figure 3. Spatial distributions of 12 h forecast LW and SW upward fluxes at the top of the atmosphere (LWUPT and SWUPT) and hourly precipitation for the Typhoon SANBA event (September 17, 2012) between the WRF control run and NN radiation emulator (NN-WRF).

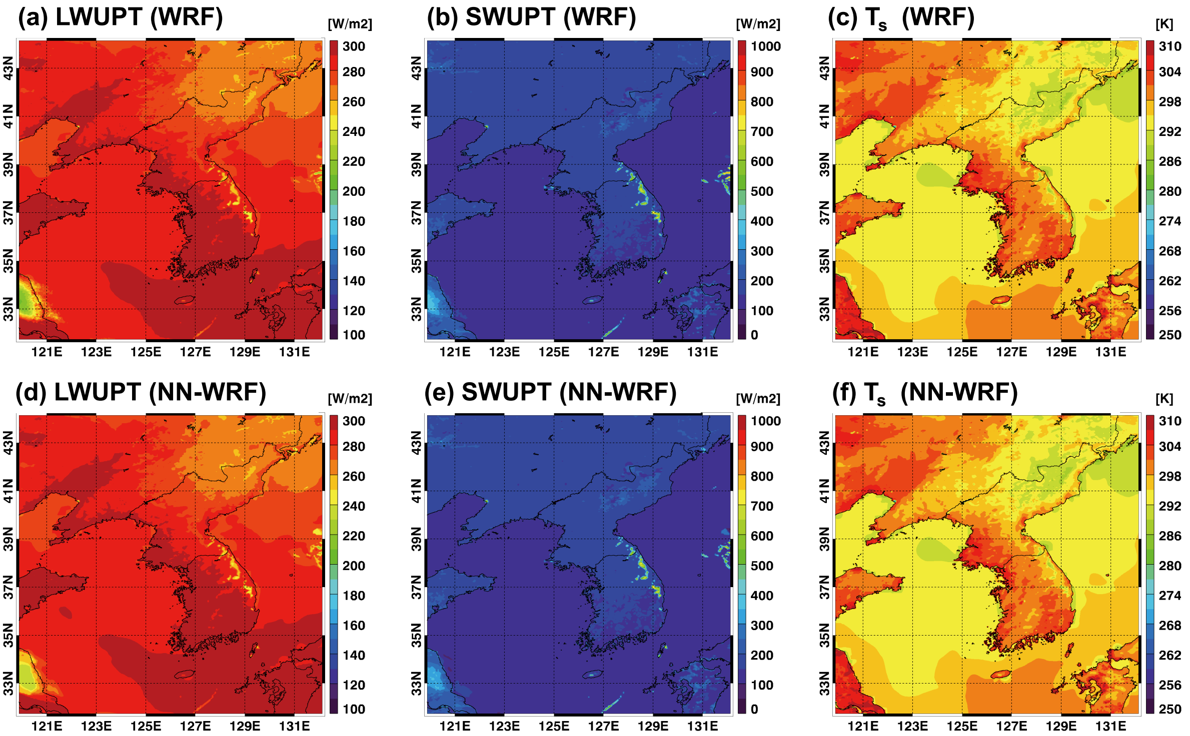

Figure 4. Spatial distributions of 12 h forecast LW and SW upward fluxes at the top of the atmosphere (LWUPT and SWUPT), and skin temperature for a clear-sky dominant event (September 23, 2011) between the WRF control run and NN radiation emulator (NN-WRF).

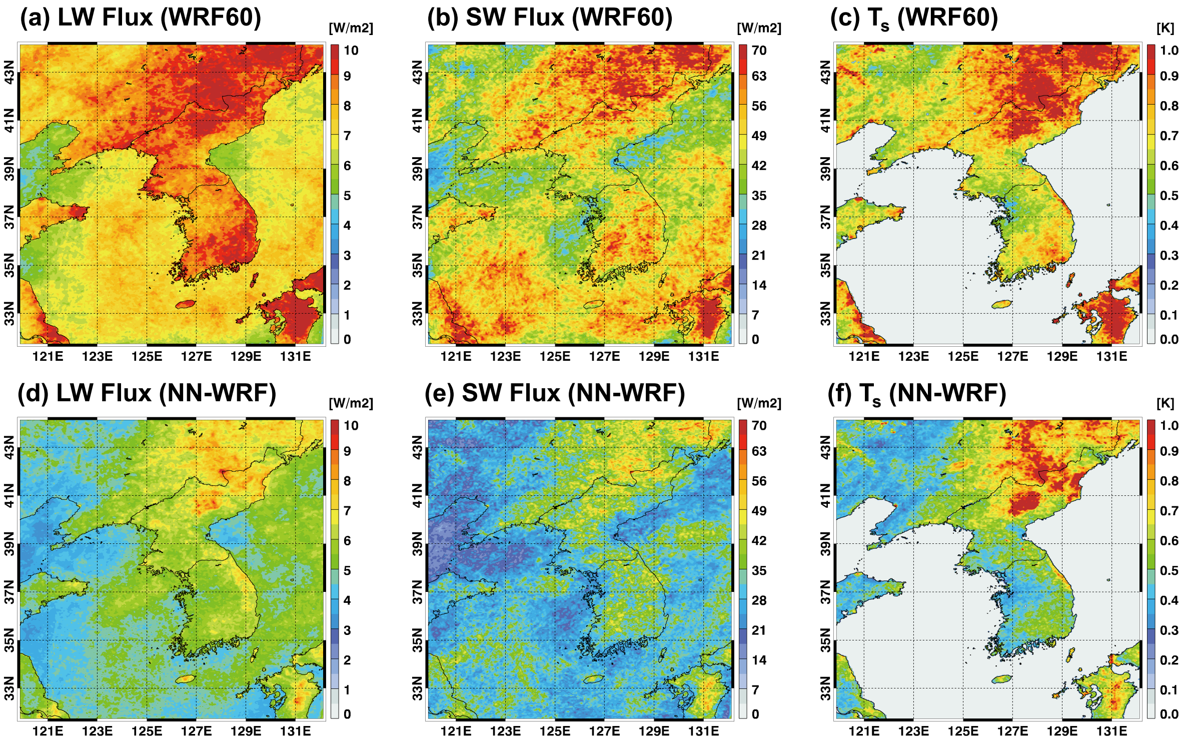

Figure 5. Root mean square error (RMSE) distributions of LW and SW fluxes and skin temperature between WRF simulations for 60-fold reduced usage (WRF60) and NN radiation emulator (NN-WRF) compared with those from the WRF control run. The average RMSEs from three LW and SW fluxes are presented.

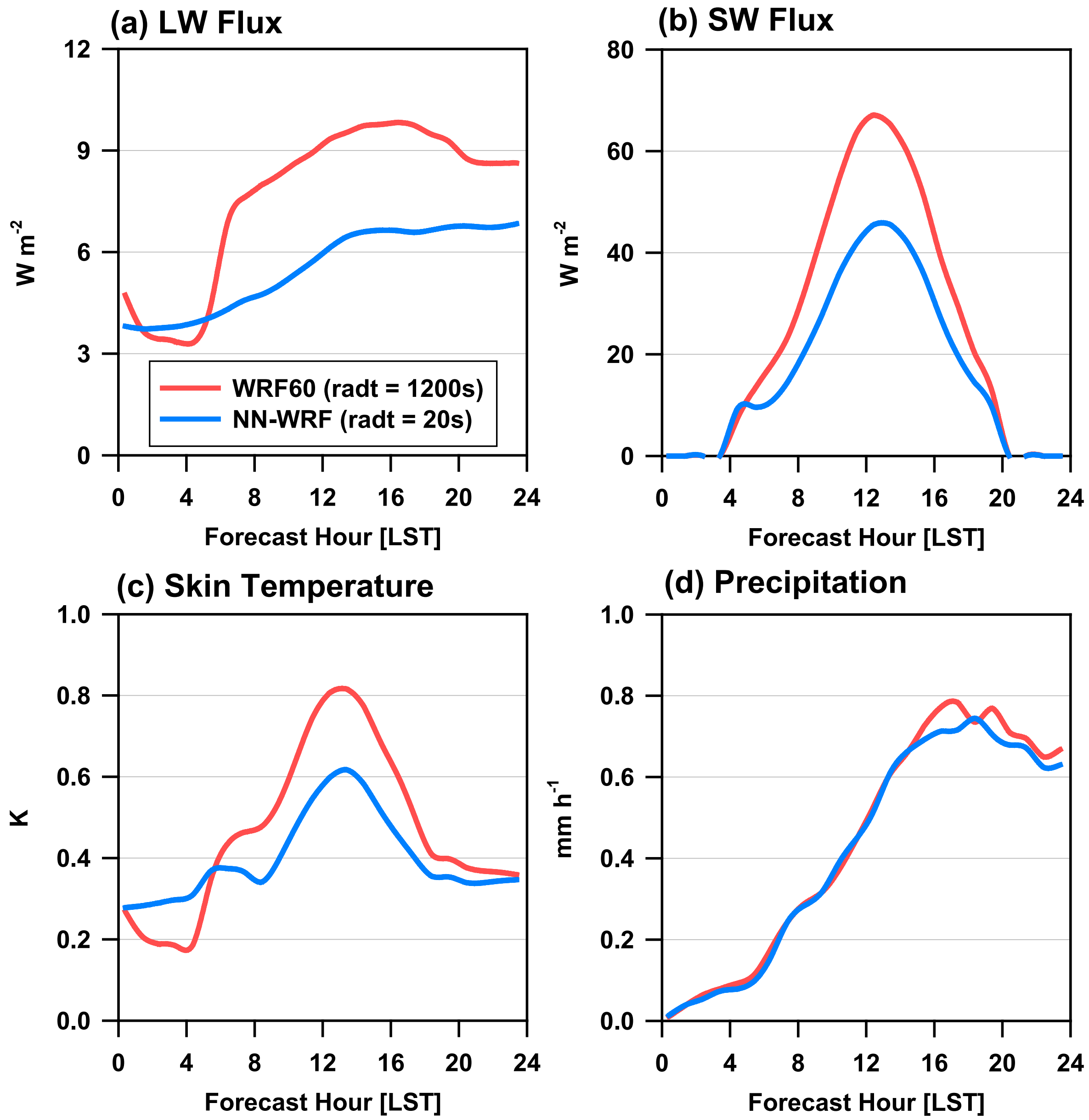

Figure 6. Time series of the root mean square error (RMSE) of (a) LW flux, (b) SW flux, (c) skin temperature, and (d) precipitation between WRF simulations 60-fold reduced usage (WRF60) and NN radiation emulator (NN-WRF) compared with those from the WRF control run.

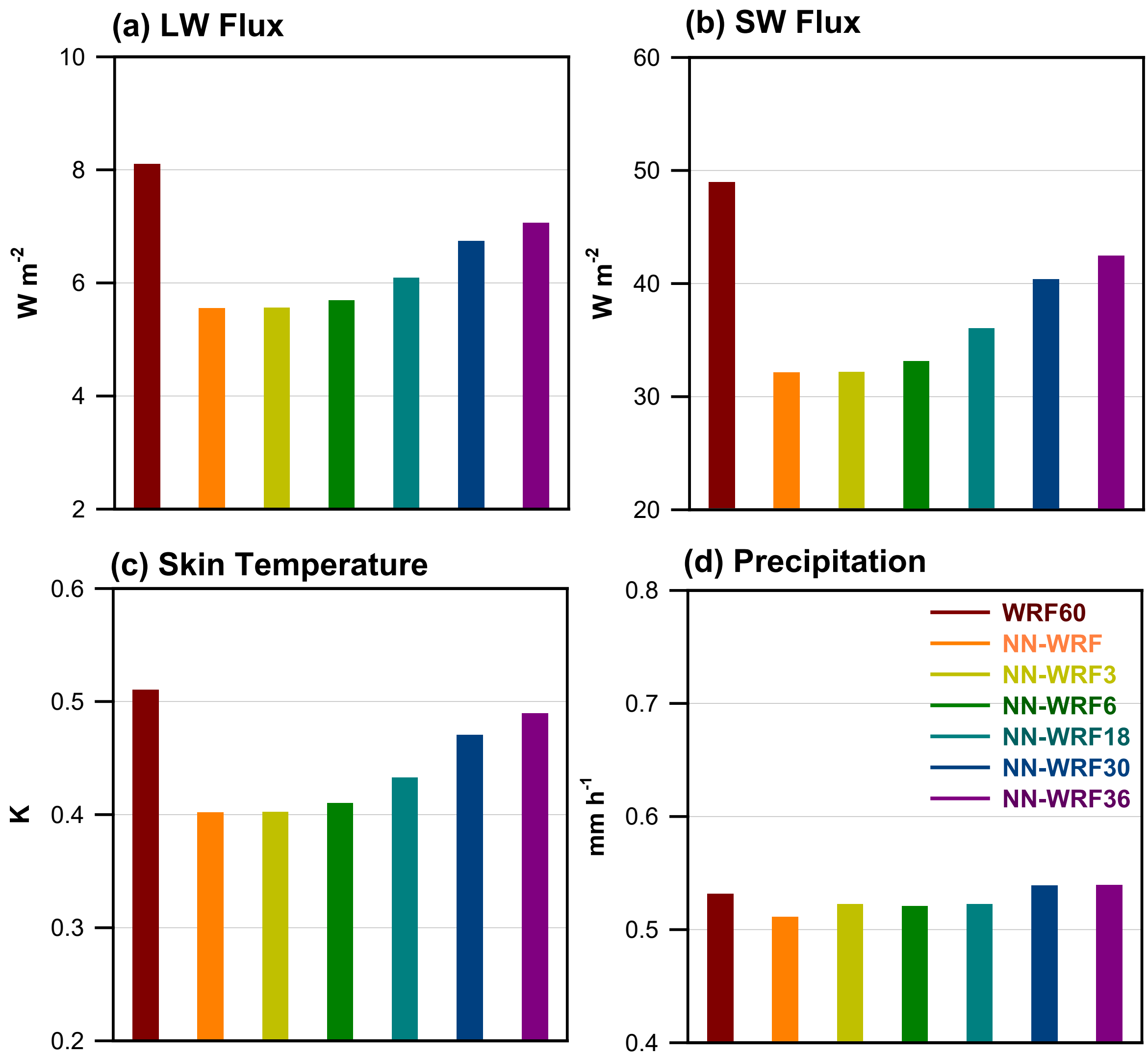

Figure 7. Statistical results showing the root mean square error (RMSE) of (a) LW flux, (b) SW flux, (c) skin temperature, and (d) precipitation for WRF simulations according to the combination of NN radiation emulator (NN-WRF) and its infrequent usage, compared with the 60-fold reduced usage of the radiation parameterization (WRF60). In the x-axis, e NN-WRF3, NN-WRF6, NN-WRF18, NN-WRF30, NN-WRF36 represent additional speedup improvements of NN radiation emulator through the 3, 6, 18, 30, and 36-fold reduced usage. Those emulator results are 180, 360, 900, 1080, 1800, 2160 times faster than those of the WRF control run, as the NN radiation emulator is already 60 times faster than the control run.

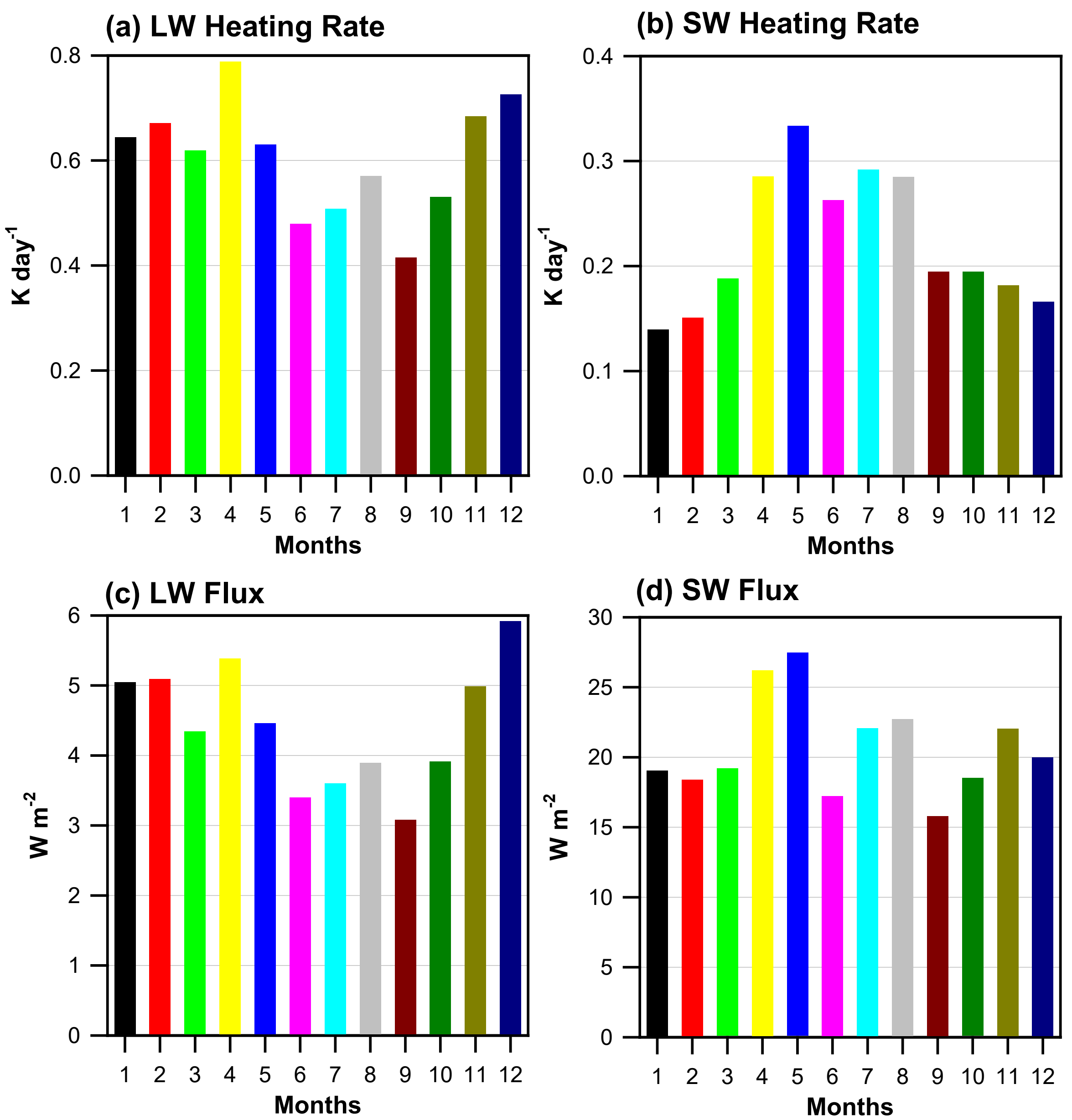

Figure 8. Same as Figure 1 but for results based on NN training from independent one year while keeping the same test sets used in Figure 1.

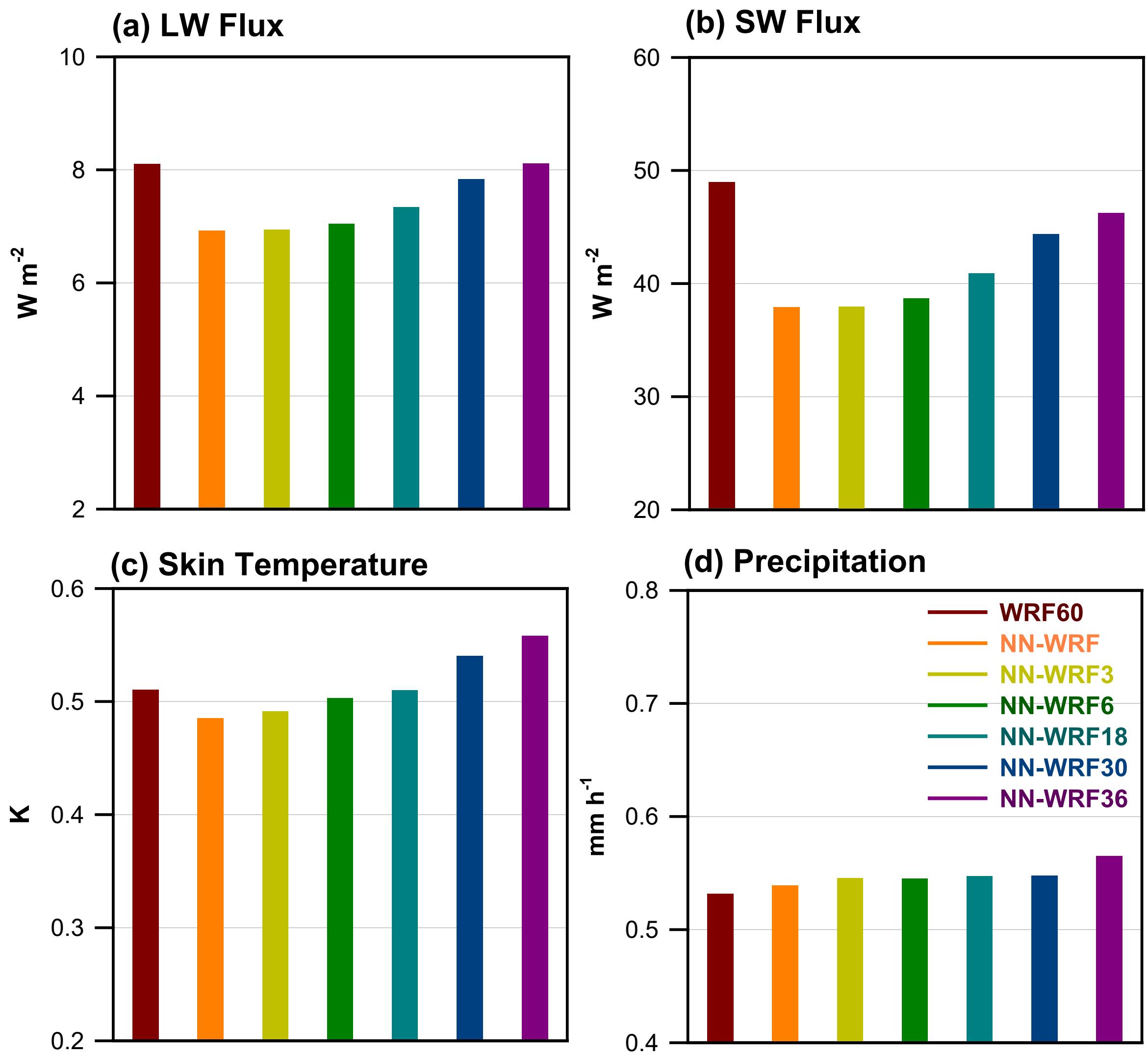

Figure 9. Same as Figure 7 but for results based on NN training from independent one year while keeping the same validation cases as Figure 7.