Lightning over the Boreal Zone: Skill Assessment for Various Land-Atmosphere Model Configurations and Lightning Indices

Jonas Mortelmans[1,1], Michel Bechtold[1,1], Erwan Brisson[2,2], Barry Hugh Lynn[3,3], Sujay Kumar[4,4], and Gabrielle J.M. De Lannoy[5,5]

[1]KU Leuven
[2]Goethe University Frankfurt
[3]Hebrew University of Jerusalem
[4]NASA GSFC
[5]KULeuven, Department of Earth and Environmental Sciences

November 30, 2022

Abstract

Current lightning predictions are uncertain because they either rely on empirical diagnostic relationships based on the present climate or use coarse-scale climate scenario simulations in which deep convection is parameterized. Previous studies demonstrated that simulations with convection-permitting resolutions (km-scale) improve lightning predictions compared to coarser-grid simulations using convection parameterization for different geographical locations but not over the boreal zone.

In this study, lightning simulations with the NASA Unified-Weather Research and Forecasting (NU-WRF) model are evaluated over a 955x540 km2 domain including the Great Slave Lake in Canada for six lightning seasons. The simulations are performed at convection-parameterized (9 km) and convection-permitting (3 km) resolution using the Goddard 4ICE and the Thompson microphysics (MP) schemes. Four lightning indices are evaluated against observations from the Canadian Lightning Detection Network (CLDN), in terms of spatiotemporal frequency distribution, spatial pattern, daily climatology, and an event-based overall skill assessment. Concerning the model configuration, regardless of the spatial resolution, the Thompson scheme is superior to the Goddard 4ICE scheme in predicting the daily climatology but worse in predicting the spatial patterns of lightning occurrence. Several evaluation metrics indicate the benefit of working at a convection-permitting resolution. The relative performance of the different lightning indices depends on the evaluation criteria. Finally, this study demonstrates issues of the models to reproduce the observed spatial pattern of lightning well, which might be related to an insufficient representation of land surface heterogeneity in the study area.

Lightning over the Boreal Zone: Skill Assessment for Various Land-Atmosphere Model Configurations and Lightning Indices

Jonas Mortelmans[1], Michel Bechtold[1], Erwan Brisson[2], Barry Lynn[3], Sujay Kumar[4], Gabriëlle De Lannoy[1]

[1]Department of Earth and Environmental Sciences, KU Leuven, Heverlee, B-3001, Belgium
[2]Centre National de la Recherche Scientifique, Toulouse, France
[3]Department of Earth Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 9190401, Israel; Weather It Is, Ltd., Jerusalem 9134401, Israel
[4]Hydrological Science Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Key Points:

- The NU-WRF modeling framework is run at two resolutions to predict lightning over the boreal zone for the first time.
- The simulations at the convection-permitting resolution yield more accurate lightning predictions.

Corresponding author: Jonas Mortelmans, `Jonas.mortelmans@kuleuven.be`

Abstract

Current lightning predictions are uncertain because they either rely on empirical diagnostic relationships based on the present climate or use coarse-scale climate scenario simulations in which deep convection is parameterized. Previous studies demonstrated that simulations with convection-permitting resolutions (km-scale) improve lightning predictions compared to coarser-grid simulations using convection parameterization for different geographical locations but not over the boreal zone.

In this study, lightning simulations with the NASA Unified-Weather Research and Forecasting (NU-WRF) model are evaluated over a 955x540 km^2 domain including the Great Slave Lake in Canada for six lightning seasons. The simulations are performed at convection-parameterized (9 km) and convection-permitting (3 km) resolution using the Goddard 4ICE and the Thompson microphysics (MP) schemes. Four lightning indices are evaluated against observations from the Canadian Lightning Detection Network (CLDN), in terms of spatiotemporal frequency distribution, spatial pattern, daily climatology, and an event-based overall skill assessment. Concerning the model configuration, regardless of the spatial resolution, the Thompson scheme is superior to the Goddard 4ICE scheme in predicting the daily climatology but worse in predicting the spatial patterns of lightning occurrence. Several evaluation metrics indicate the benefit of working at a convection-permitting resolution. The relative performance of the different lightning indices depends on the evaluation criteria. Finally, this study demonstrates issues of the models to reproduce the observed spatial pattern of lightning well, which might be related to an insufficient representation of land surface heterogeneity in the study area.

1 Introduction

The boreal zone consists of a mosaic of different land cover types, mainly forests and peatlands, both storing large amounts of carbon (Turetsky et al., 2015; Scharlemann et al., 2014). One of the natural features shaping the boreal landscape is wildfire (Bowman et al., 2009). Several studies indicate that lightning is the major source of ignition of wildfires in the boreal zone (Turetsky et al., 2015). It is proposed that lightning may increase due to global warming (Flannigan et al., 2013; Loisel et al., 2021; Veraverbeke et al., 2017; Krawchuk et al., 2009; Wotton et al., 2010), threatening the carbon pools above (forests) and below (peatlands) the ground by possibly shifting wildfire regimes.

Until the last decade, most lightning predictions are challenged by (i) the coarse-scale resolution of climate simulations in which the critical process of deep convection is parameterized and the detailed representation of land-atmosphere processes is lacking (Prein et al., 2015; Weisman et al., 1997), and (ii) the use of empirical relationships between uncertain atmospheric variables and lightning, based on the present climate. However, in the last decade, the focus of lightning simulations shifted from the coarse-scale (100 - 10 km) global and regional models to convection-permitting models, operating at a spatial resolution of less than 4 km (Prein et al., 2015). These finer resolution models allow for deep convection to be resolved explicitly, resulting in an improved representation of most convection related processes (Brisson et al., 2016; Prein et al., 2015; Lucas-Picher et al., 2021). At the fine resolution, convection parameterization schemes become obsolete and other processes contributing to deep convection, such as microphysical (MP) processes, and their formulations become more important (Adams-Selin et al., 2013). The finer spatial resolution of convection-permitting models also allows to represent more accurately the effect of land surface heterogeneities in the modeled land-atmosphere interactions (Vanden Broucke & Van Lipzig, 2017).

Lightning is the result of a process known as non-inductive charging (Reynolds et al., 1957; Takahashi, 1978). This mechanism implies electric charge separation due to rebounding collisions between graupel particles and cloud ice crystals in the presence of supercooled liquid water (Mason & Dash, 2000). This process mainly occurs when there is high convective activity in the area. It is thus not surprising that the estimation of lightning occurrence via lightning indices, is based on atmospheric variables that control convective activity (Finney et al., 2018; Romps et al., 2014). All proposed lightning indices have in common that they are diagnostic in nature and, thus, strongly depend on the accuracy of the representation of the relevant atmospheric input variables. This representation is expected to improve when working at a finer resolution (Brisson et al., 2016). However, to date the difference between working at a convection-permitting and convection-parameterized resolution is barely investigated with a focus on lightning indices and a systematic evaluation of different lightning indices is lacking completely over the boreal zone.

This study aims to answer the following questions with a focus on a study domain in the Canadian boreal zone: (i) What is the difference in performance between lightning simulations at the convection-permitting and convection-parameterized resolution? (ii) Since various atmospheric model processes are better resolved at convection-permitting resolution, what is the impact of the MP scheme on lightning indices? (iii) Since no lightning index was specifically developed for the boreal zone, which commonly used lightning index performs best in predicting lightning? To answer these questions, the NASA Unified-Weather Research and Forecasting (NU-WRF) framework is run with four different model configurations using two generally well-performing MP schemes, the Goddard 4ICE scheme (W. K. Tao et al., 2014) and the Thompson scheme (Thompson et al., 2008), at both a convection-parameterized and convection-permitting resolution. Four established lightning indices are diagnosed from the different atmospheric simulation outputs and evaluated against lightning observations.

This paper is organized as follows. In section 2, the model configurations, lightning indices, and evaluation procedures are discussed in detail. Section 3 presents and discusses the simulation results and the evaluation against observations. Lastly, in section 4, the main conclusions of this study are given and research needs are discussed.

2 Methodology

2.1 Study Domain and Period

The study domain is chosen in a region with dominantly forests and peatland, around the Great Slave Lake in Canada, shown by the red rectangle of approximately 550,000 km^2 in Figure 1a. This area is characterized by frequent lightning with critical importance for wildfire ignition (Veraverbeke et al., 2017). The study domain is embedded within two nested simulation domains, using the limited-area approach. The outer (WRF9, full extent of Figure 1) and inner (WRF3, white rectangle of Figure 1) domains have a spatial resolution of 9 and 3 km and a temporal resolution of 36 and 12 s, respectively. The double nested scheme follows the recommendations for spatial spin-up as described in Brisson et al. (2016); Prein et al. (2013) and allows for a study area within WRF3 for which convection-parameterized (9 km) and convection-permitting (3 km) model simulations are performed and compared.

Figures 1b-e show the local topography, Moderate Resolution Imaging Spectroradometer (MODIS) land use, and National Centers for Environmental Prediction (NCEP) surface albedo and greenness fraction of the study domain at 3 km resolution. This data is used, among others, as input by the land component of the coupled land-atmosphere simulations.

The simulations cover six lightning seasons, i.e. the months June through August for the years 2015 through 2020 as these months are known for their high lightning activity in the region of interest (Burrows & Kochtubajda, 2010). For this study, these years are chosen because of an improved network of lightning sensors. For the spin-up of the land-component, a 10-year long cold-start spin-up was defined before the start of the lightning season in 2015. The spin-up for the following lightning seasons were started from the end of the spin-up of the previous lightning season. To also provide a short-term spin-up for the coupled L-A run, the model is started 17 days before the actual start of the lightning season. This 17-day period is considered sufficient for the spin-up of the atmospheric model (Z. Tao et al., 2020).

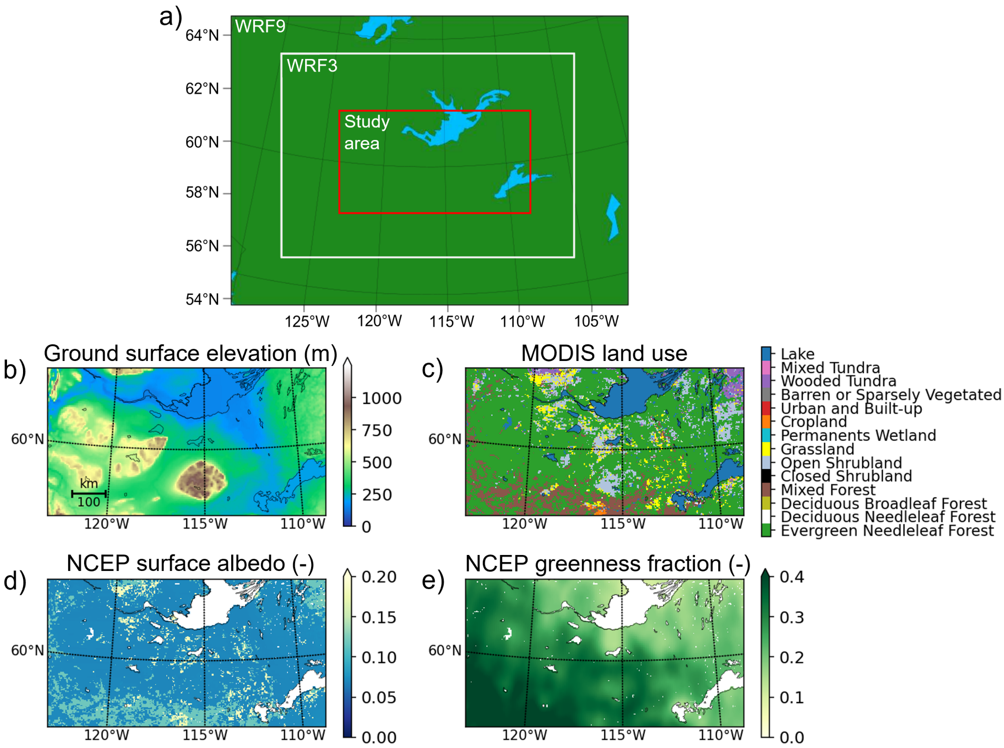

Figure 1. (a) The NU-WRF nested domains and study area (red) and its (b) topography (m), (c) MODIS land use, (d) NCEP surface albedo, and (e) NCEP greenness fraction that is used as input for the NU-WRF simulations.

2.2 Coupled Land-Atmosphere Model Configuration

The NU-WRF model is one of the leading state-of-the-art coupled land-atmosphere models that allows simulations at both convection-parameterized and convection-permitting resolution. It is an observation-driven modeling system that integrates aerosol, cloud, precipitation, and land processes at spatial resolutions of 1 – 25 km (Peters-Lidard et al., 2015). The NU-WRF model combines the National Center for Atmospheric Research (NCAR) Advanced Research WRF (ARW) (Skamarock et al., 2008) dynamical core atmospheric model with the Goddard Space Flight Center (GSFC) Land Information System (LIS) (Kumar et al., 2006, 2008) for the land component. LIS integrates the use of high-resolution satellite data, advanced land surface models (LSMs), and high-performance computing tools at high resolution. The LIS framework has multiple LSMs, one of which is the Noah-MP LSM (Niu et al., 2011). Table 1 summarizes the most important configuration options used in this study and the input datasets along with their original resolution.

Table 1. Overview of key model configuration and spatial input datasets.

Component	G4ICE - 9 km	G4ICE - 3 km	THOM - 9 km	THOM - 3 km
Land surface model	Noah-MP version 3.6			
Surface layer drag coefficient	Chen97			
Land use	MODIS including lake category (1 km)			
Topography	Global Multi-resolution Terrain Elevation Data 2010 (30 arcsec)			
Surface albedo	NCEP_Native (0.144°)			
Greenness fraction	NCEP_Native (0.144°)			
Microphysics	Goddard 4ICE	Goddard 4ICE	Thompson	Thompson
Planetary boundary layer	Mellor-Yamada-Janjic			
Cumulus parameterization	Grell-Dévényi	N/A	Grell-Dévényi	N/A
Longwave and shortwave radiation	Goddard 2017 radiation scheme			
Meteorological forcing	MERRA-2 (0.625°x0.5°)			

2.2.1 Land Surface

In this study, the Noah-MP version 3.6 is used as the LSM. This model has improved land physics compared to the standard Noah LSM, i.e. with advancements in the physics for surface fluxes, skin temperature, and runoff (Niu et al., 2011). The complete configuration settings can be found in the configuration file that is provided in the open data repository (see 'Open Research' section). The used settings are the default for Noah-MP, as outlined in the source code, with the exception of the surface layer drag coefficient. Various studies demonstrated the high sensitivity of surface energy fluxes to the choice of surface layer drag coefficient, with region-specific performance differences (e.g. Niu et al. (2011); Yang et al. (2011)). Tests with the default setup including the Monin-Obukhov similarity scheme (Brutsaert, 1982) resulted in a strong underestimation of simulated lightning occurrence for all indices. Tests with the Chen97 scheme (Chen et al., 1997) resulted in a more realistic number of lightning occurrences for both convection-parameterized. As indicated by Yang et al. (2011), the surface layer drag coefficient is the most important factor for modeling land skin temperature. Therefore, one can conclude that with the given model setup, the Chen97 scheme results in a better representation of the land skin temperature, and consequently surface energy fluxes. Given the importance of surface energy fluxes for deep convection, Chen97 was subsequently used for all simulations.

Because the default NU-WRF sea surface temperature (SST) data input derived from microwave and infrared sensors (Wentz et al., 2016) was found to not be reliable for the Great Slave Lake in our study area, we used SST data input from the Group for High Resolution SST (GHRSST) level 4 SST daily analysis (Hoyer et al., 2014; Danish Meteorological Institute, Center for Ocean and Ice, 2007). This SST data is then used for those grid cells that are identified as lakes by the land use classification (Table 1) to calculate surface energy fluxes (NASA, 2020).

2.2.2 Atmosphere

The choices for the atmospheric MP, planetary boundary layer (PBL), and cumulus parameterization scheme are based on 13 papers on the use of WRF or NU-WRF to model convection (Blake et al., 2017; Fierro et al., 2013; Gharaylou et al., 2020; Iguchi et al., 2017; Madala et al., 2014; Santanello et al., 2013; W. K. Tao et al., 2016; Z. Tao et al., 2020; Wong et al., 2013; Lang et al., 2014; Comin et al., 2018; Dawn & Satyanarayana, 2020; Gilliland & Rowe, 2007). The literature study indicated that the highest sensitivity and uncertainty of the atmospheric simulations was related to the choice of the MP scheme. To address this issue, two MP schemes, the Goddard 4ICE (W. K. Tao et al., 2014) and Thompson (Thompson et al., 2008) MP schemes, are used and compared in this study. These two schemes are among the most commonly used MP schemes and are both proven to be able to represent the atmospheric processes, such as deep convection, in both temperate and arctic regions (Lang et al., 2014; He & Loboda, 2020). Since the boreal region is geographically and climatologically located between these two regions, we decided that these two MP schemes are likely to best represent the atmospheric processes in the boreal zone. To our knowledge, no literature on the boreal region comparing different MP schemes exists. In terms of PBL and cumulus parameterization, literature did show one clear superior option in combination with the selected MP schemes. Both MP schemes proved to perform especially well in combination with the Mellor-Yamada-Janjic PBL scheme (Mellor & Yamada, 1982). This scheme outperforms other PBL schemes in the representation of thunderstorms (Madala et al., 2014). For the convection parameterization of WRF9 (9 km), the Grell-Dévényi cumulus ensemble (Grell & Dévényi, 2002) is used.

The use of two different MP schemes for two spatial model resolutions results in four different model configurations: (i) the Goddard 4ICE MP scheme at 9 km (G4ICE - 9 km), (ii) the Goddard 4ICE MP scheme at 3 km (G4ICE - 3 km), (iii) the Thompson MP scheme at 9 km (THOM - 9 km), and (iv) the Thompson MP scheme at 3 km spatial resolution (THOM - 3 km). Lightning occurences are diagnosed from each of these four model configurations using four different lightning indices, resulting in 16 numerical experiments.

2.3 Lightning Indices

Some lightning indices can be used to determine lightning flash densities directly, whereas others provide a lightning probability that then needs to be converted into lightning flash densities. Several lightning indices exist with various levels of complexity, spanning from the approach of Price and Rind (1992), based on the convective cloud top height, to those approaches based on the evolving electric field in storms, such as the lightning index of Fierro et al. (2013).

In this study, four different lightning indices are compared: (i) the Lightning Potential Index (LPI) of Yair et al. (2010), (ii) the lightning threat (LT) of McCaul et al. (2009), (iii) the Price and Rind (1992) index based on maximal updraft velocity (PR92W), and (iv) the product of CAPE and convective precipitation rate (CAPExP) developed by Romps et al. (2014). All indices are diagnosed from the hourly NU-WRF output.

2.3.1 LPI

The LPI (J kg^{-1}) is an empirical index that is based on ice fractions and supercooled liquid water mixing ratios in the region between 0 and -20 °C. In this temperature range, the noninductive mechanisms that involve the collision of ice and graupel particles are most effective, because they require the presence of super-cooled liquid water to have charge separation due to the rebounding collisions between graupel and cloud ice crystals (Saunders, 2008). This index does not directly estimate the flash density but is a measure of the potential for charge generation and separation that leads to lightning (Yair et al., 2010; Lynn & Yair, 2010). It is calculated from the vertical updraft velocity and the mixing ratios of liquid water, cloud ice, snow, and graupel.

$$LPI = 1/V \int\int\int \epsilon w^2 dxdydz \tag{1}$$

where V is the volume of air in the layer between 0 and -20°C; w is the vertical updraft velocity (m s^{-1}); dx and dy are the horizontal, and dz the vertical dimensions of the grid cell (m); and ϵ is a dimensionless number between 0 and 1:

$$\epsilon = \frac{2(Q_i Q_l)^{0.5}}{(Q_i + Q_l)} \tag{2}$$

where Q_l is the total liquid water mass mixing ratio (kg kg^{-1}); and Q_i is the ice fractional mass mixing ratio (kg kg^{-1}), which is defined as:

$$Q_i = q_g \left[\left(\frac{(q_s q_g)^{0.5}}{(q_s + q_g)} \right) + \left(\frac{(q_i q_g)^{0.5}}{(q_i + q_g)} \right) \right] \tag{3}$$

where q_i, q_g, and q_s are the mass mixing ratios for cloud ice, graupel, and snow, respectively (all in kg kg^{-1}).

2.3.2 McCaul Lightning Threat

The McCaul Lightning Threat (LT) (flashes (5min$\cdot$gridbox)$^{-1}$) is a linear combination of (i) the upward fluxes of precipitating ice hydrometeors in the mixed-phase region at the -15°C level and (ii) the vertical integral of cloud ice, graupel and snow, as follows:

$$LT = 0.95 k_1 (w q_g)_m + 0.05 k_2 \int \rho (q_g + q_s + q_i) dz \tag{4}$$

where $k_1 = 0.042$; $k_2 = 0.20$; w is the vertical updraft velocity (m s^{-1}); ρ is the air density (kg m^{-3}); and q_i, q_g, and q_s are the cloud ice, graupel and snow mixing ratio, respectively. The subscript m indicates the -15 °C level.

2.3.3 PR92W

The PR92W index (flashes (min)$^{-1}$) is based on the relation between maximum updraft velocity (w_{max}; in m s^{-1}) and the number of flashes per minute:

$$PR92W = c\, 5 \cdot 10^{-6}\, w_{max}^{4.54} \tag{5}$$

where c is a calibration factor used to generalize the original equation from a 5 km spatial resolution to all possible resolutions (Price & Rind, 1994). The calibration factor c is defined as:

$$c = 0.97241\, e^{0.048203R} \tag{6}$$

where R is the grid cell area in squared degrees. Price and Rind (1994) state that this calibration factor does not depend on the latitude or longitude. These relatively simple relations have been shown to perform relatively well at different spatial resolutions (ranging from 1 - 36 km) by several studies (Ushio et al., 2001; Yoshida et al., 2009; Barthe et al., 2010; Wong et al., 2013) and were for a long time the most frequently used lightning indices.

2.3.4 **CAPE $\times$ P**

The last lightning index used in this study is the product of convective available potential energy (CAPE, in J kg^{-1}) and the convective precipitation rate (P, in kg (m^{2}s)$^{-1}$), expressed in flashes (m^{2}s)$^{-1}$, as developed by Romps et al. (2014). This product is a good proxy for lightning distribution over land when multiplied with a constant of proportionality (η/E) to convert it to a flash density:

$$\text{CAPE} \times \text{P} = \eta/E \cdot \text{CAPE} \cdot \text{P} \tag{7}$$

where η/E consists of the dimensionless conversion efficiency η and the energy discharge per flash E (in J). Romps et al. (2014) found that CAPExP correlates best with the observed lightning with η/E equal to $1.3 \cdot 10^{-11} \mathrm{J}^{-1}$ (Romps et al., 2014).

Since the NU-WRF model only simulates convective precipitation rate when a convective parameterization scheme is used (only activated for 9 km simulations), the convective precipitation rate used for the CAPExP index in this study is determined based on the method described in Churchill and Houze (1984). They defined convective cores as grid cells with twice the rainfall rate of the background (2 grid cells in each direction) average or any grid cell with a rain rate of >20 mm h^{-1}. The grid cells directly surrounding the convective center are also considered convective regions. To keep the results of the two resolutions comparable, this method was used to determine convective precipitation for both the 3 and 9 km simulations.

2.4 Evaluation

2.4.1 Lightning Observations

The observational data used for evaluation is provided by the Canadian Lightning Detection Network (CLDN). The data covers the area between 58 – 66°N and 108 - 125°W for the years 2015 – 2020. Earlier data are available but subject to large biases due to the use of older sensor technology and therefore not used. The dataset consists of individual flashes measured with a spatial and temporal precision of 0.0001° (approximately 5 m) and 10^{-3} s, respectively. A classification into cloud-cloud and cloud-ground lightning is also provided.

For further evaluation, the observational data is converted to an hourly flash density at the two different resolutions. The flashes are summed regardless of the type (cloud-cloud or cloud-ground) since the different lightning indices used in this study do not differentiate between types of flashes.

2.4.2 Rescaling of Lightning Indices

The lightning indices derived from all experiments are rescaled to the observations. This is on the one hand needed to convert the LPI, which represents the potential for lightning to occur, to a flash density and on the other hand to allow a conistent comparison across lightning indices. We followed the two-step procedure as described in Brisson et al. (2021). First, the excessive small flash densities are eliminated so that the total sum (in both space and time) of modeled lightning flashes equals the total sum of observed lightning flashes. In a second step, a linear function is derived to relate the model output to the observed flash densities. Note that each lightning index is rescaled by a single linear model for the entire domain which makes overfitting issues very unlikely given the large sample size. Note that the approach is applicable to climate change scenarios since the same linear equation could be used to rescale future predictions without altering the climate change signal as demonstrated in Brisson et al. (2021).

2.4.3 Evaluation: Precipitation and Surface Energy Fluxes

The ability of the model to accurately simulate surface energy fluxes is key to the quality of lightning predictions. Therefore, the modeled LH and SH as well as total precipitation patterns of the different model configurations are evaluated in a first step. This is done by comparing the spatial patterns of the 6-year summer averages of the precipitaion, LH, and SH modeled by the different model configurations against MERRA-2 reanalysis data.

2.4.4 Evaluation: Spatiotemporal Frequency Distribution of Lightning

To evaluate the frequency distribution of the modeled lightning indices to that of the observations, the spatiotemporal probability density functions (PDF; not shown) are calculated for each index. The PDFs are then compared to the observations using the Perkins skill score (PSS) (Perkins et al., 2007) from a rescaled flash density of 0.1 flashes (h km^2)$^{-1}$ onward. The PSS measures the common area between two PDFs as follows:

$$\text{PSS} = \sum_{i=1}^{n} min(z_s(i), z_0(i)) \tag{8}$$

where i is the bin index; n is the total number of bins; and $z_o(i)$ and $z_s(i)$ are the relative frequencies of a given bin from the observations and model, respectively. The PSS is a measure for non-linearities between two datasets. A PSS of 1 means that two PDFs are identical, while a value < 0.7 indicates that the two PDFs differ significantly according to Perkins et al. (2007).

2.4.5 Evaluation: Spatial Pattern and Diurnal Climatology of Lightning

To evaluate the different experiments in terms of their capability to simulate lightning in space and time, the 6-year average spatial patterns of daily flash densities and the 6-year averaged diurnal cycle are computed. The results are compared with the CLDN observations by means of the spatial Pearson correlation coefficients (R), for each of the experiments. For the diurnal cycle, the 6-year averaged diurnal cycle is for each grid cell in space evaluated against the CLDN observations. For consistency, the results of the 3 km experiments are regridded to match the resolution of the 9 km simulations.

2.4.6 Evaluation: Event-Based Skill Assessment of Lightning

Lastly, the predictive performance of the different experiments is also evaluated on an 'event-by-event' (i.e. grid cell per grid cell and time step per time step) basis by using the following evaluation metrics: (i) the probability of detection (POD, equation 9), (ii) the critical success index (CSI, equation 10), (iii) the bias (equation 11), and (iv) the success ratio (SR, equation 12). These four metrics can be easily presented together in a single "performance diagram" as shown in Roebber (2009). The POD, CSI and SR all range between 0 and 1, with 1 a perfect score. The bias, defined here as the ratio of the observed events and the predicted events, ranges from 0 to infinity, but still has 1 as a perfect score. Because of the way Roebber (2009) designed their performance diagram, predictions that are further to the top-right are more accurate than points closer to the bottom-left. Additionally, ideal predictions are located on the bias=1 diagonal. The metrics are calculated as follows,

$$\text{POD} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

$$\text{CSI} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{10}$$

$$\text{bias} = \frac{\text{TP} + \text{FP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{SR} = 1 - \frac{\text{FP}}{\text{TP} + \text{FP}} \tag{12}$$

where TP, FP, TN, and FN are the true positive (hits), false positive (false alarms), true negative (correct negatives), and false negative (misses) lightning predictions.

To calculate these four skill metrics, the model output and the observations are aggregated to a 72x72 km^2 grid and a 6-hourly time interval. This aggregation was done to reduce the effects of the temporal and spatial inaccuracy of the atmospheric inputs on the skill of the lightning indices. The performance diagram proposed by Roebber (2009) provides a summary of both spatial and temporal simulation performance.

3 Results and Discussion

3.1 Precipitation and Surface Energy Fluxes

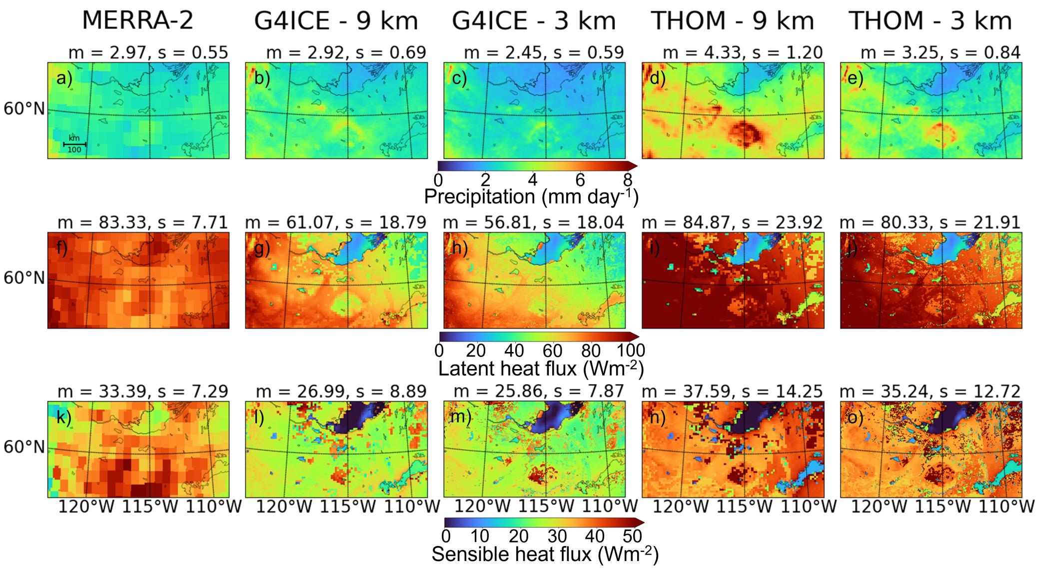

Figure 2. 6-year averaged summer precipitation (a-e), latent heat flux (f-j) and sensible heat flux (k-o) estimated by MERRA-2 and the 4 different model configurations (G4ICE - 9 km, G4ICE - 3 km, THOM - 9 km and THOM - 3 km). The spatial mean (m) and standard deviation (s) are provided at the top of each figure.

Figure 2 shows the 6-year (2015-2020) averaged daily precipitation, LH, and SH from the NU-WRF simulations compared with MERRA-2 data. MERRA-2 (first column) has a known yearly average positive bias for the boreal summer LH of 20 W m^{-2} and globally overestimates the SH by 6 Wm^{-2}, on average (Draper et al., 2018). Although these biases were not determined specifically for our study domain, they here serve as inidication of a possible bias.

Compared to the MERRA-2 data, the G4ICE - 9 km simulation shows a similar mean precipitation (2.92 and 2.97 mm day^{-1} for G4ICE - 9 km and MERRA-2, respectively), whereas the G4ICE - 3 km simulation shows a slight underestimation (2.45 mm day^{-1}). Both the THOM - 9 km and THOM - 3 km simulations show an overestimation of the spatially averaged daily precipitation rate, with 4.33 and 3.25 mm day^{-1}, respectively. The THOM simulations mainly show a higher precipitation rate around topographic structures in the center and western part of the study domain (see Figure 1b), likely associated with orographic precipitation (Smith, 1979). This sensibility to topography is also seen for the G4ICE simulations and MERRA-2 but is less pronounced. Orographic clouds are primarily composed of supercooled liquid water. Therefore, the representation of this microphysical variable by the MP scheme will greatly influence the presence of orographic precipitation (Sarmadi et al., 2019). Since the G4ICE MP is a single moment scheme and the THOM MP is a double moment scheme, both schemes simulate hydrometeors differently (Dawn & Satyanarayana, 2020). Sarmadi et al. (2019) also found that differences in simulated precipitation are mainly due to uncertainties in the physical processes of the model. Thus, the difference in simulated hydrometeors explains part of the observed differences in precipitation.

Both MP schemes produce a southwest-northeast gradient of decreasing LH, while for SH, the gradients disagree. The G4ICE simulations show a southwest-northeast gradient of decreasing SH, while the opposite is observed for the THOM simulations. For both surface energy fluxes, MERRA-2 shows a more uniform pattern of generally higher values. The differences are strongest over the Great Slave Lake where both MP schemes produce lower surface energy fluxes. Despite their diverging spatial gradients, the spatial means of the LH of both THOM simulations (84.87 and 80.33 W m^{-2} for 9 km and 3 km resolution, respectively) are similar to that of the MERRA-2 (83.33 W m^{-2}), while those of the G4ICE simulations are significantly lower (61.07 and 56.81 W m^{-2}, respectively). Similar conclusions can be made for the average SH, with the G4ICE simulations showing lower averages compared to both MERRA-2 and the THOM simulations. The fact that both LH and SH are lower for G4ICE, indicates that this MP scheme reflects more radiation than the THOM scheme. This consequently results in lower surface temperatures (not shown here).

3.2 Lightning Simulations

3.2.1 Spatiotemporal Frequency Distribution

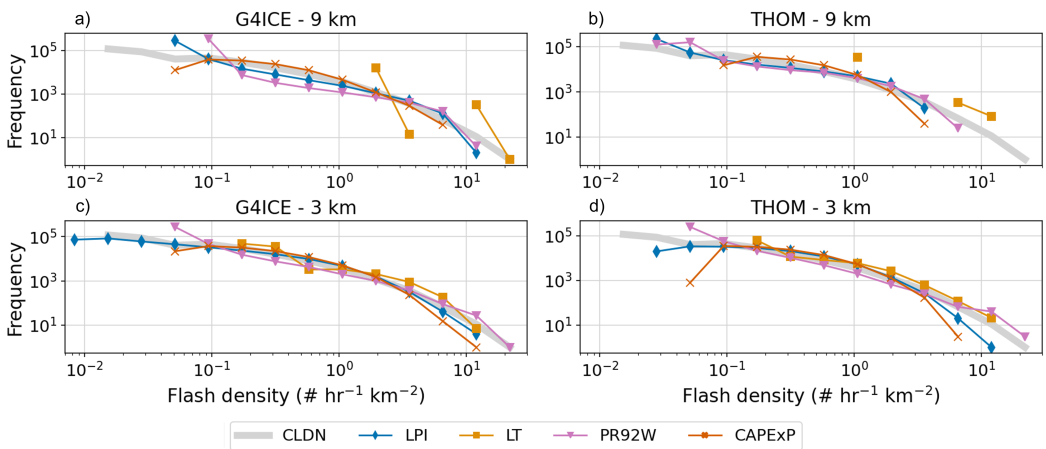

Figure 3. Spatiotemporal frequency distributions of the domain average hourly flash rates for each model configuration as observed (gray), as represented by the LPI, LT3, PR92W, and CAPExP parameterizations after linear rescaling to match the observations.

Figure 3 compares the spatiotemporal frequency distribution of the observed hourly flash density (in gray) and the model output for each experiment. The PSS values of these frequency distributions are presented in Table 2. The time dependent graphs in Figure 3 and the tabular data of PSS in Table 2 together show a clear improvement at 3 km compared to 9 km for all experiments, with the exception of PR92W for the THOM MP (PSS of 0.88 and 0.87 for 9 and 3 km, respectively), and CAPExP for the G4ICE MP (with a PSS value of 0.94 for both resolutions). The largest improvements can be seen for the LT index, with PSS values increasing from 0.02 to 0.65 and from 0.08 to 0.78 for the G4ICE and THOM simulations, respectively. Note that the PSS is only calculated for flash densities higher than 0.1 flashes h^{-1} km^{-2}, as the lower values were influenced by the cutoff value applied during the rescaling (see section 2.4.2).

Both MP schemes fail to provide an accurate prediction for LT, with only three or four different predicted flash densities at 9 km. For the other indices, the G4ICE - 3 km model configuration provides the highest PSS values. This model configuration also captures both the lowest as the highest flash densities better than the other configurations, as shown in Figure 3. This superiority of the convection-permitting resolution can also be seen for the THOM MP scheme.

Table 2. PSS values for the different model configurations and indices.

Index	G4ICE - 9 km	G4ICE - 3 km	THOM - 9 km	THOM - 3 km
LPI	0.89	0.93	0.89	0.91
LT	0.02	0.65	0.08	0.78
PR92W	0.81	0.88	0.88	0.87
CAPExP	0.94	0.94	0.90	0.91

3.2.2 Spatial Patterns

The spatial pattern of the 6-year average flash density of the CLDN observations (Figure 4a-d) shows a southwest-northeast gradient of decreasing lightning. The main cluster of lightning for the CLDN observations also shows agreement with the greenness fraction shown in Figure 1. This cluster is not seen in the simulation output. For the G4ICE simulations, the main area of lightning is in the southern half of the study domain, with relatively few flashes in the northeastern corner. The main cluster for the THOM simulations has shifted towards the center of the domain, especially for LPI and CAPExP. A comparison of the spatial R values shown in Figure 5 shows that the G4ICE - 3 km configuration (R = 0.43 - 0.57), is superior to the other model configurations in terms of predicting the spatial pattern of the CLDN observations, with the exception of the CAPExP index, which performs better at a coarser resolution. Even though this model configuration also fails to predict the exact cluster of the observed lightning occurrences, it does show the southwest-northeast gradient of decreasing lightnig that is also in the CLDN observations. THOM - 3 km also shows higher correlations than THOM - 9 km, except for CAPExP (R = 0.26 and 0.25 at 9 km and 3 km, respectively). Romps et al. (2018) found that CAPExP performed very well at a 0.5° grid, indicating that this index might work better at a coarser resolution. Brisson et al. (2021) state that CAPExP shows problems at convection-permitting resolutions because explicitly resolving convection leads to a null CAPE if there is convective precipitation. Prein et al. (2015), on the other hand, found that convection-permitting resolutions improved the representation of extreme precipitation and summertime convection. Both are directly used by the CAPExP index. This better representation of convection at the finer resolution also explains the better performance of the indices that depend on the maximum vertical updraft velocity, as convection determines the theoretical maximum updraft velocity (Bao & Sherwood, 2019).

For all four model configurations, LT and PR92W show clear line structures that are not seen for any of the other indices. Both indices depend strongly on the maximal vertical updraft velocity (see equations 4 and 5). Among all other indices, only LPI is dependent on the updraft velocity, but also strongly depends on other atmospheric variables (see equation 1). A map of the maximal updraft velocity for each grid cell (not shown here) shows very similar line-like structures as those seen for LT and PR92W, supporting this strong dependence on this atmospheric variable. One weather phenomenon that is known to go along strong updrafts is a squall line, a line of thunderstorms that forms along a cold front and is characterized by frequent lightning and strong updrafts (Newton, 1950). Another explanation is so-called topographic convergence. The observed line struc-

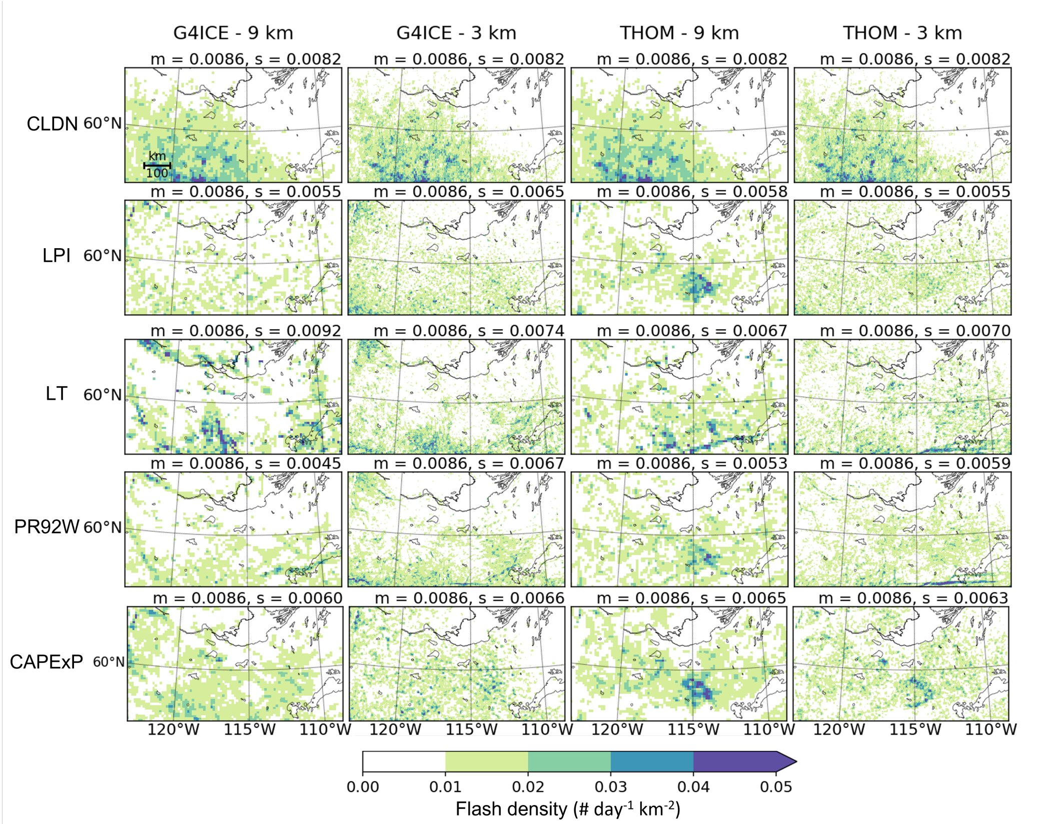

Figure 4. 6-year average total daily flash density for the months of June through August of 2015-2020 for the observations (CLDN), LPI, LT, PR92W, and CAPExP parameterizations and for each model configuration (G4ICE - 9 km, G4ICE - 3 km, THOM - 9 km and THOM - 3 km). The spatial mean (m) and standard deviation (s) are provided at the top of each figure.

tures correlate well with the location of terrain height differences, as shown in Figure 1. Due to the elevated terrain, the wind flow is forced to go up and around the topographic structure, which can cause convective uplift near complex topography (Barthlott et al., 2006).

The CAPExP index shows a very low flash density of < 0.01 flashes h^{-1} km^{-2} for the THOM simulations over and around Lake Athabasca in the southeast, making it stand out from the surroundings. While this is also true for the G4ICE model configurations over the lake, these experiments do show lightning at the edge of the lake. This can be explained by the difference in relative LH and SH over this lake compared to the surrounding land for both MP schemes (Figure 2). For the THOM experiments, both LH and SH over this lake are smaller than for the surrounding land. This implies that there is less energy for the formation of strong thunderstorms over the lake as compared to the land (Beringer & Tapper, 2002). For the G4ICE simulations, the LH of the lake is higher than that of the surrounding land and SH is only slightly smaller. In absolute values, the fluxes over the lake are higher for the G4ICE than for the THOM experiments. Since CAPE is highest when both LH and SH are high, this all leads to more predicted thunderstorms around Lake Athabasca for the G4ICE simulations.

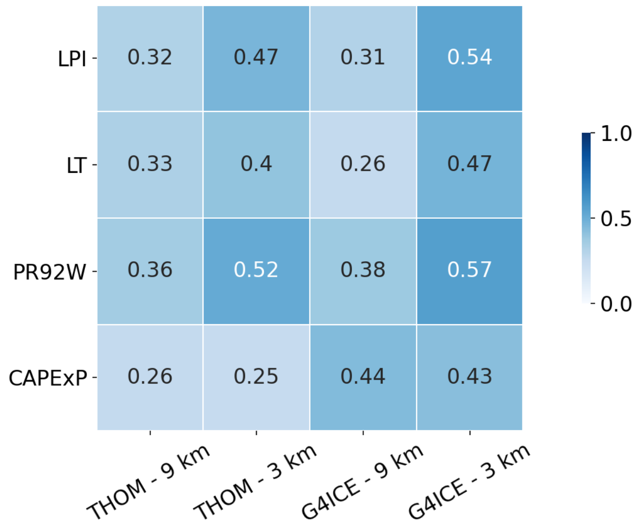

Figure 5. Spatial Pearson correlation coefficients between the 6-year summer average of each lightning index and the CLDN observations for the 4 model configurations.

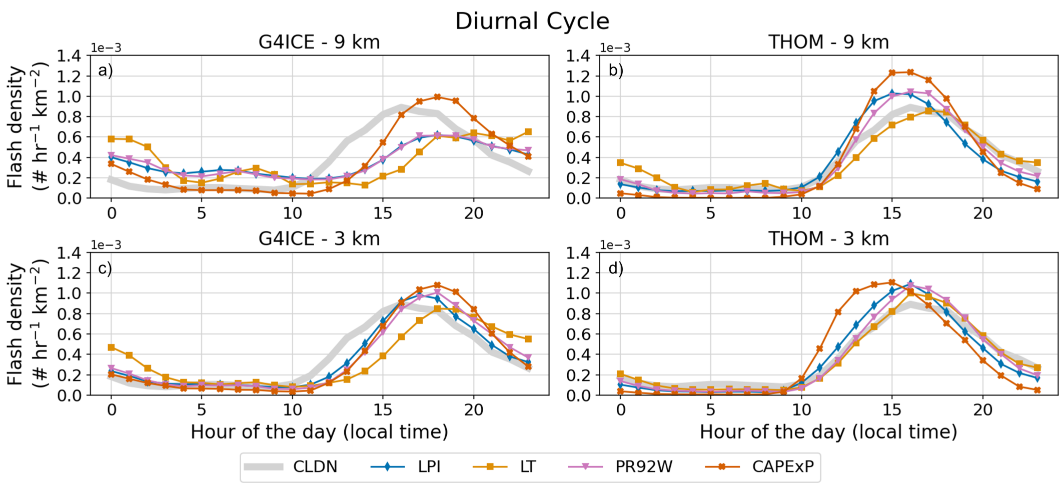

Figure 6. Six-year average diurnal cycle of the domain average flash density for each experiment, diagnosed by the LPI, LT, PR92W, CAPExP indices, and compared to (gray) observations.

3.2.3 Diurnal Climatology

Figure 6 shows the diurnal cycles of the lightning indices for each experiment together with that of the CLDN observations (in gray). All diurnal cycles exhibit a peak in the afternoon around 3-5 pm local time, which matches with the peak in the reference observations. Both THOM experiments better predict the diurnal cycle than the G4ICE experiments, especially at the coarser resolution and early in the morning. This better performance is also shown by the higher temporal R values (see Figure 7).

The diurnal cycles of the THOM simulations show a very clear peak in the afternoon, starting 1 or 2 hours before (for CAPExP and LPI of THOM - 3 km) to 1 hour after (for LT of THOM - 9 km) the observed peak. The G4ICE simulations, on the other

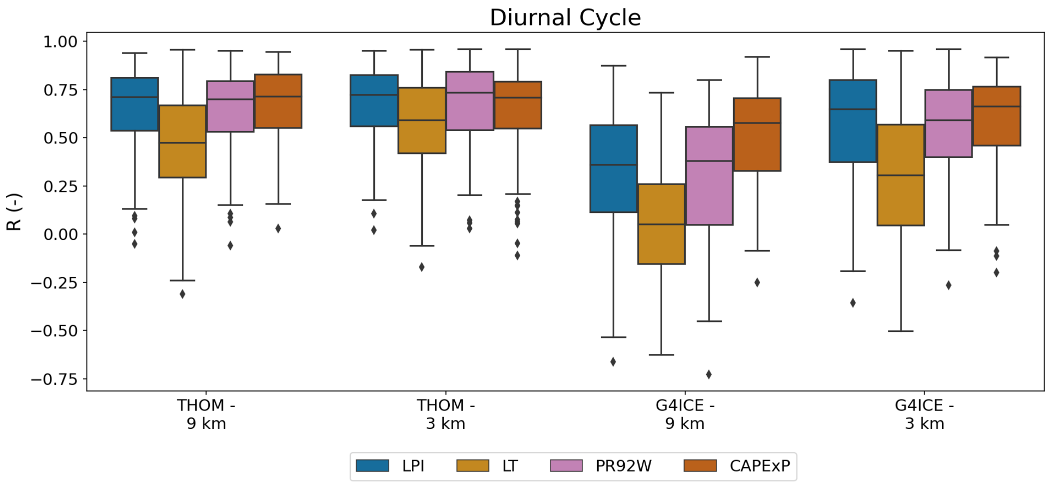

Figure 7. Temporal Pearson correlation coefficients (R) between simulated and observed (CLDN) six-year averaged diurnal cycle of flash density calculated for each grid cell, and grouped by lightning index and experiment.

hand, generally peak one to three hours after the observed peak. These differences between the simulated and observed peak are reduced for the 3 km simulations, with the exception of CAPExP for the THOM - 3 km experiment, for which the peak is now shifted to approximately 2 hours before the observed peak. In terms of absolute values, the CAPExP index for the THOM - 9 km experiment greatly overestimates the observed peak, whereas for the G4ICE - 9 km experiment, all indices, except CAPExP, underestimate the peak. The diurnal cycles in Figrue 6 and the Pearson correlation coefficients in Figure 7 show the best agreement between CAPExP and the observed diurnal cycle, confirming the findings of Romps et al. (2018) that CAPExP is good at capturing the timing of the observed diurnal cycle. However, for the THOM - 3 km simulations, the LPI performs better. For all experiments, LT is the least capable of capturing the diurnal cycle of the observations.

3.2.4 Overall Event-Based Skill Assessment

In the previous sections, the performance of the different experiments and lightning indices was evaluated either for a strong aggregation in space (long-term spatial pattern) or in time (diurnal cycle, frequency distribution), and on the basis of flash densities. Here we present an additional assessment on an event-by-event basis using the predictive skill scores (POD, CSI, SR, and bias) that do not take into account the absolute flash densitites but only the presence or absence of lightning. It is important to note that the bias used here is not following the conventional definition of bias in statistics, i.e. it is not the difference between the predictions and the observations, but rather the ratio of total (in both space and time) predicted grid cells with lightning occurrence over the total observed grid cells with lightning. By this definition, the bias does not take the temporal and/or spatial mismatch between the model and observations into account. This mismatch is represented in the POD, CSI, and SR.

Figure 8 represents the performance diagram as described in Roebber (2009), showing the POD versus the SR. The blue contours represent the CSI and dashed lines the bias. Points further to the top right indicate a better overall accuracy for that experiment. Figure 8 shows that, by comparing for each index the different model configurations, the 3 km simulations (green) systematically have a higher POD, SR, and CSI. The

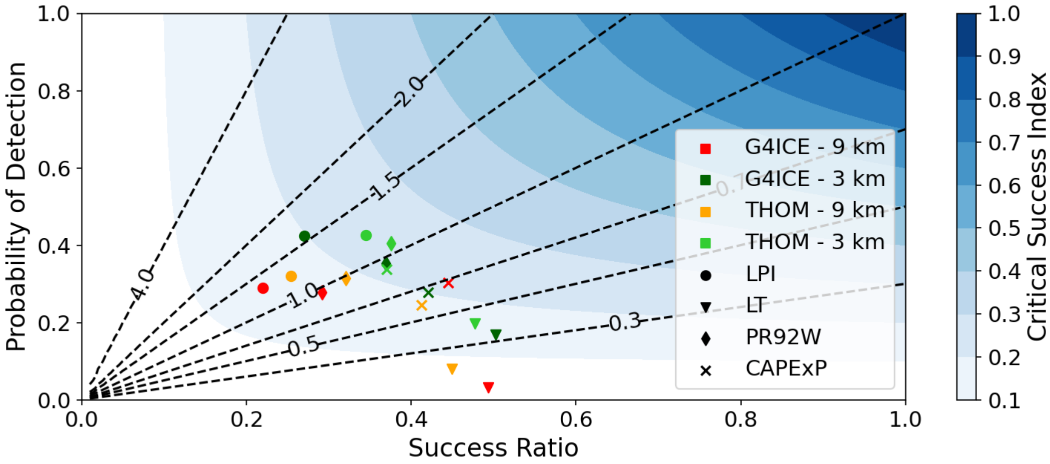

Figure 8. Performance diagram of the POD on the y-axis versus SR on the x-axis. The blue contours represent the CSI and the dashed lines represent the bias. The skills were calculated on a 72x72 km^2 grid and aggregated over six hours. Different colors indicate different experiments, while different markers indicate the different lightning indices.

bias seems less influenced by the model configuration, but rather depends on the index. This indicates that the indices a bias > 1 predict a lot of lightning occurrences, whereas those with a bias < 1 predict less lightning occurrences in general, independent of the model configuration.

For all four model configurations, PR92W has a bias very close to 1, indicating that it predicts a similar amount of lightning occurrences in space and time as the observations. The LPI (only for the G4ICE - 3 km experiment) shows the highest POD, but has a rather low SR. The LT, on the other hand, has the highest SR for all model configurations, but also the lowest POD. It is noticeable that for all lightning indices, there are clear clusters of the datapoints for the four model configurations. These clusters are somewhat expected as the natural tendency of an index to predict more/less lightning would not change much with another MP or resolution. The different clusters can be ranked in order of decreasing POD, bias, or increasing SR, all resulting in the same order: (i) LPI, (ii) PR92W (bias of approximately 1), (iii) CAPExP, and (iv) LT.

The high bias (> 1) of the LPI, indicates that this index systematically predicts more grid cells with lightning than observed after calibration. The LT, on the other hand, has a bias < 1, indicating that it predicts less lightning occurrences than observed. By predicting much less lightning occurrences than observed, the POD is naturally rather low and the SR high, as there are not many mismatches if not a lot of data is available. For the LPI, the higher POD can also be explained by the high bias: if more lightning is predicted, the probability of detecting observed lightning is naturally higher. But this high bias has the trade-off that there are more false alarms, and thus it has a lower SR. PR92W and CAPExP are in between the LPI and LT in terms of bias, POD and SR. With the bias of the PR92W almost equal to 1 for all model configurations, and SR, POD and CSI relatively good, especially for the 3 km simulations, this index shows a lot of potential over this study area. The CAPExP has more variability in terms of the skill scores, strongly depending on the model configuration.

4 Conclusion

Global warming might enhance lightning activity in the boreal zone and lead to more wildfire ignitions. Therefore, there is a need for reliable lightning estimates at the different time scales of weather and climate simulations. In this paper, the NU-WRF model is evaluated with two MP schemes: (i) the Goddard 4 ICE scheme, and (ii) the Thompson scheme, both operating at a convection-parameterized (9 km) and convection permitting (3 km) horizontal resolution, leading to four model configurations: (i) G4ICE - 9 km, (ii) G4ICE - 3 km, (iii) THOM - 9 km, and (iv) THOM - 3 km. These configurationss are first compared in their capability to simulate energy fluxes throughout the domain, using MERRA-2 data as a reference. To diagnose lightning flash densities from the model output, four lightning indices (LPI, LT, PR92W, and CAPExP) are used, resulting in 16 lightning predictions. These are evaluated for their capability to model the frequency distribution, spatial pattern and diurnal cycle of CLDN observations. Additionally, four predictive event-based skill scores are compared for all combinations of model configurations and lightning indices. The main results can be summarized as follows.

1. For the evaluation of the precipitation and the surface energy fluxes, the THOM simulations show a similar spatially averaged LH and SH as MERRA-2, which however has a known positive bias for both LH and SH. The G4ICE simulations do not show this bias and result in a spatial average of both LH and SH that is approximately 20 and 6 W m^{-2} lower than those of MERRA-2.
2. For the spatial pattern, no model configuration predicts the observed cluster of high lightning occurrence in the southwestern part of the domain. Only the G4ICE - 3 km simulations are capable of reproducing the southwest-northeast gradient of decreasing lightning that is seen for the observations (with higher spatial R values than other for configurations). In general, the convection-permitting resolution is superior to the convection-parameterized resolution, except for the CAPExP index, which is a lightning index that is known to perform better at a coarser resolution (Romps et al., 2018).
3. For the diurnal cycle, the THOM MP scheme performs better than G4ICE. CAPExP is superior to the other lightning indices, except for the THOM - 3 km setup, and LT performs the worst. A clear benefit of using the finer resolution is only seen for the G4ICE experiments.
4. The event-based skill scores, represented in a performance diagram, support that the convection-permitting modeling at 3 km resolution leads to a generally higher performance for all model configurations and lightning indices.
5. No MP scheme is found to be superior for all evaluated aspects. Whereas the THOM MP scheme results in a better timing for most indices, the G4ICE scheme results in a better predicted spatial pattern.
6. No lightning index is found to be superior for all evaluated aspects.

Based on those results, we conclude that diagnosing lightning indices from the output of a convection-permitting model seems to be beneficial. However, we emphasize that this conclusion is only valid for the applied model configurations, that is, using another MP, PBL, or cumulus parameterization scheme might lead to a different conclusion.

Furthermore, it is important to note that the performance of all lightning indices strongly depends on the capability of the atmospheric model to accurately represent the required input parameters for the lightning indices. An error in the model representation of e.g. updraft velocity will lead to an error in the PR92W index. Our finding that no lightning index was generally superior to another is thus to be seen in light of the atmospheric model skill. The performance of the lightning indices may depend differently on forecast skill. Since the model skill of a specific forecast is not well known beforehand, decisionmakers and forest managers might consider to use an ensemble of different lightning indices for the estimation of lightning probability.

Based on the results from this study, we could also make some suggestions for future work. A first suggestion is based on the relative poor representation of the spatial pattern of lightning. An improved representation of the boreal land mosaic, consisting mainly of forests and peatlands, might be needed in the land surface modeling component (Bechtold et al., 2019; Qiu et al., 2018; Melton et al., 2019), since peatlands and forests substantially differ in their partitioning of energy fluxes (Helbig et al., 2020). A second suggestion, directed to future lightning predictions in different climate scenarios is to not use a linear rescaling to make the different lightning indices comparable. Instead, it might be better to calibrate the different parameters of the different lightning indices in order to better account for non-linearities.

Acknowledgments

J. Mortelmans thanks the Research Foundation - Flanders (FWO) for funding this research (FWO.G095720N). The computer resources and services used in this work were provided by the High Performance Computing system of the Vlaams Supercomputer Center, funded by the Research Foundation - Flanders (FWO) and the Flemish Government. We thank Environment Canada for their generous permission to use Canadian Lightning Data Network data as observational reference for the model output. We thank Carlos Cruz from the NU-WRF team for technical support, and Alexander Gruber and Nicole Van Lipzig for valuable feedback on the conducted research and the manuscript.

References

Adams-Selin, R. D., Van Den Heever, S. C., & Johnson, R. H. (2013). Impact of graupel parameterization schemes on idealized bow echo simulations. *Monthly Weather Review*, *141*(4), 1241–1262. doi: 10.1175/MWR-D-12-00064.1

Bao, J., & Sherwood, S. C. (2019). The Role of Convective Self-Aggregation in Extreme Instantaneous Versus Daily Precipitation. *Journal of Advances in Modeling Earth Systems*, *11*(1), 19–33. doi: 10.1029/2018MS001503

Barthe, C., Deierling, W., & Barth, M. C. (2010). Estimation of total lightning from various storm parameters: A cloud-resolving model study. *Journal of Geophysical Research Atmospheres*, *115*(24), 1–17. doi: 10.1029/2010JD014405

Barthlott, C., Corsmeier, U., Meißner, C., Braun, F., & Kottmeier, C. (2006). The influence of mesoscale circulation systems on triggering convective cells over complex terrain. *Atmospheric Research*, *81*(2), 150-175. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0169809505002577` doi: https://doi.org/10.1016/j.atmosres.2005.11.010

Bechtold, M., De Lannoy, G. J. M., Koster, R. D., Reichle, R. H., Mahanama, S. P., Bleuten, W., ... Tiemeyer, B. (2019). PEAT-CLSM: A Specific Treatment of Peatland Hydrology in the NASA Catchment Land Surface Model. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2130–2162. doi: 10.1029/2018MS001574

Beringer, J., & Tapper, N. (2002). Surface energy exchanges and interactions with thunderstorms during the Maritime Continent Thunderstorm Experiment (MCTEX). *Journal of Geophysical Research Atmospheres*, *107*(21), AAC 3–1–AAC 3–13. doi: 10.1029/2001JD001431

Blake, B. T., Parsons, D. B., Haghi, K. R., & Castleberry, S. G. (2017). The structure, evolution, and dynamics of a nocturnal convective system simulated using the WRF-ARW model. *Monthly Weather Review*, *145*(8), 3179–3201. doi: 10.1175/MWR-D-16-0360.1

Bowman, D. M. J. S., Balch, J. K., Artaxo, P., Bond, W. J., Carlson, J. M., Cochrane, M. A., ... Pyne, S. J. (2009). Fire in the earth system. *Science*, *324*(5926), 481–484. doi: 10.1126/science.1163886

Brisson, E., Blahak, U., Lucas-Picher, P., Purr, C., & Ahrens, B. (2021). Contrasting lightning projection using the lightning potential index adapted in a convection-permitting regional climate model. *Climate Dynamics*, *57*(7-8), 2037–2051. doi: 10.1007/s00382-021-05791-z

Brisson, E., Van Weverberg, K., Demuzere, M., Devis, A., Saeed, S., Stengel, M., & van Lipzig, N. P. M. (2016). How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics? *Climate Dynamics*, *47*(9-10), 3043–3061. doi: 10.1007/s00382-016-3012-z

Brutsaert, W. (1982). *Evaporation into the Atmosphere: Theory, History, and Applications* (1st ed.). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-1497-6

Burrows, W. R., & Kochtubajda, B. (2010). A decade of cloud-to-ground lightning in canada: 1999–2008. part 1: Flash density and occurrence. *Atmosphere-Ocean*, *48*(3), 177-194. doi: 10.3137/AO1118.2010

Chen, F., Janjić, Z., & Mitchell, K. (1997). Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Boundary-Layer Meteorology*, *85*(3), 391–421. doi: 10.1023/A:1000531001463

Churchill, D. D., & Houze, R. A. a. (1984). Development and Structure of Winter Monsoon Cloud Clusters On 10 December 1978. *Journal of the Atmospheric Sciences*, *41*(6), 933–960. doi: 10.1175/1520-0469(1984)041⟨0933:dasowm⟩2.0.co;2

Comin, A. N., Schumacher, V., & Justino, F. (2018). Impact of different microphysical parameterizations on extreme snowfall events in the Southern Andes. *Weather and Climate Extremes*, *21*(January), 65–75. Retrieved from `https://doi.org/10.1016/j.wace.2018.07.001` doi: 10.1016/j.wace.2018.07.001

Danish Meteorological Institute, Center for Ocean and Ice. (2007). *GHRSST Level 4 DMI_OI Global Foundation Sea Surface Temperature Analysis (GDS version 2)* [dataset]. NASA Physical Oceanography DAAC. doi: https://doi.org/10.5067/GHGDM-4FD02

Dawn, S., & Satyanarayana, A. N. V. (2020). Sensitivity studies of cloud microphysical schemes of WRF-ARW model in the simulation of trailing stratiform squall lines over the Gangetic West Bengal region. *Journal of Atmospheric and Solar-Terrestrial Physics*, *290*, 105396. doi: https://doi.org/10.1016/j.jastp.2020.105396

Draper, C. S., Reichle, R. H., & Koster, R. D. (2018). Assessment of MERRA-2 land surface energy flux estimates. *Journal of Climate*, *31*(2), 671–691. doi: 10.1175/JCLI-D-17-0121.1

Fierro, A. O., Mansell, E. R., Macgorman, D. R., & Ziegler, C. L. (2013). The implementation of an explicit charging and discharge lightning scheme within the wrf-arw model: Benchmark simulations of a continental squall line, a tropical cyclone, and a winter storm. *Monthly Weather Review*, *141*(7), 2390–2415. doi: 10.1175/MWR-D-12-00278.1

Finney, D. L., Doherty, R. M., Wild, O., Stevenson, D. S., MacKenzie, I. A., & Blyth, A. M. (2018). A projected decrease in lightning under climate change. *Nature Climate Change*, *8*(3), 210–213. doi: 10.1038/s41558-018-0072-6

Flannigan, M., Cantin, A. S., De Groot, W. J., Wotton, M., Newbery, A., & Gowman, L. M. (2013). Global wildland fire season severity in the 21st century. *Forest Ecology and Management*, *294*, 54–61. doi: 10.1016/j.foreco.2012.10.022

Gharaylou, M., Farahani, M. M., Mahmoudian, A., & Hosseini, M. (2020). Prediction of lightning activity using WRF-ELEC model: Impact of initial and boundary conditions. *Journal of Atmospheric and Solar-Terrestrial Physics*, *210*, 105438. doi: 10.1016/j.jastp.2020.105438

Gilliland, E. K., & Rowe, C. M. (2007). A comparison of cumulus parameterization schemes in the WRF model. *87th AMS Annual Meeting*(Arakawa).

Grell, G. A., & Dévényi, D. (2002). A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters*, *29*, 34–38. doi: https://doi.org/10.1029/2002gl015311

He, J., & Loboda, T. V. (2020). Modeling cloud-to-ground lightning probability in Alaskan tundra through the integration of Weather Research and Forecast (WRF) model and machine learning method. *Environmental Research Letters*, *15*(11). doi: 10.1088/1748-9326/abbc3b

Helbig, M., Waddington, J. M., Alekseychik, P., Amiro, B., Aurela, M., Barr, A. G., ... Schulze, C. (2020). The biophysical climate mitigation potential of boreal peatlands during the growing season. *Environmental Research Letters*, *15*(10). doi: 10.1088/1748-9326/abab34

Hoyer, J. L., Le Brogne, P., & Eastwood, S. (2014). A bias correction method for Arctic satellite sea surface temperature observations. *Remote Sensing of Environment*, *146*. doi: https://doi.org/10.1016/j.rse.2013.04.020

Iguchi, T., Tao, W. K., Wu, D., Peters-Lidard, C., Santanello, J. A., Kemp, E., ... Loikith, P. (2017). Sensitivity of CONUS summer rainfall to the selection of cumulus parameterization schemes in NU-WRF seasonal simulations. *Journal of Hydrometeorology*, *18*(6), 1689–1706. doi: 10.1175/JHM-D-16-0120.1

Krawchuk, M. A., Cumming, S. G., & Flannigan, M. D. (2009). Predicted changes in fire weather suggest increases in lightning fire initiation and future area burned in the mixedwood boreal forest. *Climatic Change*, *92*(1-2), 83–97. doi: 10.1007/s10584-008-9460-7

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., ... Sheffield, J. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling and Software*, *21*(10), 1402–1415. doi: 10.1016/j.envsoft.2005.07.004

Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T., ... Houser, P. R. (2008). A land surface data assimilation framework using the land information system: Description and applications. *Advances in Water Resources*, *31*(11), 1419–1432. doi: 10.1016/j.advwatres.2008.01.013

Lang, S. E., Tao, W. K., Chern, J. D., Wu, D., & Li, X. (2014). Benefits of a fourth ice class in the simulated radar reflectivities of convective systems using a bulk microphysics scheme. *Journal of the Atmospheric Sciences*, *71*(10), 3583–3612. doi: 10.1175/JAS-D-13-0330.1

Loisel, J., Gallego-Sala, A. V., Amesbury, M. J., Magnan, G., Anshari, G., Beilman, D. W., ... Wu, J. (2021). Expert assessment of future vulnerability of the global peatland carbon sink. *Nature Climate Change*, *11*(1), 70–77. doi: 10.1038/s41558-020-00944-0

Lucas-Picher, P., Argüeso, D., Brisson, E., Tramblay, Y., Berg, P., Lemonsu, A., ... Caillaud, C. (2021, 11). Convection-permitting modeling with regional climate models: Latest developments and next steps. *WIREs Climate Change*, *12*. doi: 10.1002/wcc.731

Lynn, B., & Yair, Y. (2010). Prediction of lightning flash density with the WRF model. *Advances in Geosciences*, *23*(February), 11–16. doi: 10.5194/adgeo-23-11-2010

Madala, S., Satyanarayana, A. N. V., & Rao, T. N. (2014). Performance evaluation of PBL and cumulus parameterization schemes of WRF ARW model in simulating severe thunderstorm events over Gadanki MST radar facility - Case study. *Atmospheric Research*, *139*, 1–17. doi: 10.1016/j.atmosres.2013.12.017

Mason, B., & Dash, J. G. (2000). Charge and mass transfer in ice-ice collisions: Experimental observations of a mechanism in thunderstorm electrification. *Jour-*

nal of Geophysical Research, *105*, 10185-10192.

McCaul, E. W., Goodman, S. J., LaCasse, K. M., & Cecil, D. J. (2009). Forecasting lightning threat using cloud-resolving model simulations. *Weather and Forecasting*, *24*(3), 709–729. doi: 10.1175/2008WAF2222152.1

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, *20*, 851. doi: https://doi.org/10.1029/rg020i004p00851

Melton, J., Arora, V., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., & Teckentrup, L. (2019). CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework and site-level performance. *Geoscientific Model Development Discussions*, 1–40. doi: 10.5194/gmd-2019-329

NASA. (2020). NU-WRF Version 10 User ' s Guide.

Newton, C. W. (1950). STRUCTURE AND MECHANISM OF THE PRE-FRONTAL SQUALL LINE. *Journal of Atmospheric Sciences*, *7*(3), 210–222. doi: 10.1175/1520-0469(1950)007⟨0210:SAMOTP⟩2.0.CO;2

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., … Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research Atmospheres*, *116*(12), 1–19. doi: 10.1029/2010JD015139

Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007). Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. *Journal of Climate*, *20*(17), 4356–4376. doi: 10.1175/JCLI4253.1

Peters-Lidard, C. D., Kemp, E. M., Matsui, T., Santanello, J. A., Kumar, S. V., Jacob, J. P., … Zupanski, M. (2015). Integrated modeling of aerosol, cloud, precipitation and land processes at satellite-resolved scales. *Environmental Modelling and Software*, *67*, 149–159. doi: 10.1016/j.envsoft.2015.01.007

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., & Georgievski, G. (2013). Added value of convection permitting seasonal simulations. *Climate Dynamics*, *41*(9-10), 2655–2677. doi: 10.1007/s00382-013-1744-6

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., … Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*(2), 323–361. doi: 10.1002/2014RG000475

Price, C., & Rind, D. (1992). A Simple Lightning Parameterization for Calculating Global Lightning Distributions. , *97*(92), 9919–9933.

Price, C., & Rind, D. (1994). Modeling Global Lightning Distributions in a General Circulation Model. *Monthly Weather Review*, *122*, 1930–1939. doi: 10.1175/1520-0493(1994)122⟨1930:mgldia⟩2.0.co;2

Qiu, C., Zhu, D., Ciais, P., Guenet, B., Krinner, G., Peng, S., … Ziemblinska, K. (2018). ORCHIDEE-PEAT (revision 4596), a model for northern peatland CO2, water, and energy fluxes on daily to annual scales. *Geoscientific Model Development*, *11*(2), 497–519. doi: 10.5194/gmd-11-497-2018

Reynolds, S. E., Brook, M., & Gourley, M. F. (1957). Thunderstorm charge separation. *Journal of Atmospheric Sciences*, *14*(5), 426 - 436. doi: 10.1175/1520-0469(1957)014⟨0426:TCS⟩2.0.CO;2

Roebber, P. J. (2009). Visualizing multiple measures of forecast quality. *Weather and Forecasting*, *24*(2), 601–608. doi: 10.1175/2008WAF2222159.1

Romps, D. M., Charn, A. B., Holzworth, R. H., Lawrence, W. E., Molinari, J., & Vollaro, D. (2018). CAPE Times P Explains Lightning Over Land But Not the Land-Ocean Contrast. *Geophysical Research Letters*, *45*(22), 12,612–623,630.

doi: 10.1029/2018GL080267

Romps, D. M., Seeley, J. T., Vollaro, D., & Molinari, J. (2014). Projected increase in lightning strikes in the united states due to global warming. *Science*, *346*(6211), 851–854. doi: 10.1126/science.1259100

Santanello, J. A., Kumar, S. V., Peters-Lidard, C. D., Harrison, K., & Zhou, S. (2013). Impact of land model calibration on coupled land-atmosphere prediction. *Journal of Hydrometeorology*, *14*(5), 1373–1400. doi: 10.1175/JHM-D-12-0127.1

Sarmadi, F., Huang, Y., Thompson, G., Siems, S. T., & Manton, M. J. (2019). Simulations of orographic precipitation in the Snowy Mountains of Southeastern Australia. *Atmospheric Research*, *219*(January), 183–199. Retrieved from `https://doi.org/10.1016/j.atmosres.2019.01.002` doi: 10.1016/j.atmosres.2019.01.002

Saunders, C. (2008). Charge separation mechanisms in clouds. , 335–353.

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: Understanding and managing the largest terrestrial carbon pool. *Carbon Management*, *5*(1), 81–91. doi: 10.4155/cmt.13.77

Skamarock, W. C., Klemp, J. B., Dudhi, J., Gill, D. O., Barker, D. M., Duda, M. G., . . . Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3. *Technical Report*(June), 113. doi: 10.5065/D6DZ069T

Smith, R. B. (1979). The influence of mountains on the atmosphere. In B. Saltzman (Ed.), (Vol. 21, p. 87-230). Elsevier. doi: https://doi.org/10.1016/S0065-2687(08)60262-9

Takahashi, T. (1978). Riming electrification as a charge generation mechanism in thunderstorms. *Journal of Atmospheric Sciences*, *35*(8), 1536 - 1548. doi: 10.1175/1520-0469(1978)035⟨1536:REAACG⟩2.0.CO;2

Tao, W. K., Lang, S., Zeng, X., Li, X., Matsui, T., Mohr, K., . . . Yang, Y. M. (2014). The Goddard Cumulus Ensemble model (GCE): Improvementsand applications for studying precipitation processes. *Atmospheric Research*, *143*, 392–424.

Tao, W. K., Wu, D., Lang, S., Chern, J. D., Peters-Lidard, C., Fridlind, A., & Matsui, T. (2016). High-resoution NU-WRF simulations of a deep convective-precipitation system during MC3E: Further improvements and comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, *121*, 1278–1305. doi: 10.1002/2015JD023986.Received

Tao, Z., Chin, M., Gao, M., Kucsera, T., Kim, D., Bian, H., . . . Akimoto, H. (2020). Evaluation of NU-WRF model performance on air quality simulation under various model resolutions - An investigation within the framework of MICS-Asia Phase III. *Atmospheric Chemistry and Physics*, *20*(4), 2319–2339. doi: 10.5194/acp-20-2319-2020

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, *136*(12), 5095–5115. doi: 10.1175/2008MWR2387.1

Turetsky, M. R., Benscoter, B., Page, S., Rein, G., Van Der Werf, G. R., & Watts, A. (2015). Global vulnerability of peatlands to fire and carbon loss. *Nature Geoscience*, *8*(1), 11–14. doi: 10.1038/ngeo2325

Ushio, T., Heckman, S. J., Boccippio, D. J., Christian, H. J., & Kawasaki, Z. I. (2001). A survey of thunderstorm flash rates compared to cloud top height using TRMM satellite data. *Journal of Geophysical Research Atmospheres*, *106*(D20), 24089–24095. doi: 10.1029/2001JD900233

Vanden Broucke, S., & Van Lipzig, N. (2017). Do convection-permitting models improve the representation of the impact of LUC? *Climate Dynamics*, *49*(7), 2749–2763. doi: 10.1007/s00382-016-3489-5

Veraverbeke, S., Rogers, B. M., Goulden, M. L., Jandt, R. R., Miller, C. E., Wiggins, E. B., & Randerson, J. T. (2017). Lightning as a major driver of recent large fire years in North American boreal forests. *Nature Climate Change*, *7*(7), 529–534. doi: 10.1038/nclimate3329

Weisman, M. L., Skamarock, W. C., & Klemp, J. B. (1997). The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, *125*(4), 527–548. doi: 10.1175/1520-0493(1997)125⟨0527:TRDOEM⟩2.0.CO;2

Wentz, F., Mears, C., Meissner, T., Ricciardulli, L., Lindsley, R., Manaster, A., ... Wentz, K. (2016). *Monthly Mean Wind Speed Data Set on a 1 degree grid made from Remote Sensing Systems Version-7 Microwave Radiometer Data, V0701* [dataset]. Santa Rosa, CA, USA. Retrieved 2020-10-01, from `www.remss.com`

Wong, J., Barth, M. C., & Noone, D. (2013). Evaluating a lightning parameterization based on cloud-top height for mesoscale numerical model simulations. *Geoscientific Model Development*, *6*(2), 429–443. doi: 10.5194/gmd-6-429-2013

Wotton, B. M., Nock, C. A., & Flannigan, M. D. (2010). Forest fire occurrence and climate change in Canada. *International Journal of Wildland Fire*, *19*(3), 253–271. doi: 10.1071/WF09002

Yair, Y., Lynn, B., Price, C., Kotroni, V., Lagouvardos, K., Morin, E., ... Del Carmen Llasat, M. (2010). Predicting the potential for lightning activity in Mediterranean storms based on the Weather Research and Forecasting (WRF) model dynamic and microphysical fields. *Journal of Geophysical Research Atmospheres*, *115*(4), 1–13. doi: 10.1029/2008JD010868

Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., ... Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research Atmospheres*, *116*(12), 1–16. doi: 10.1029/2010JD015140

Yoshida, S., Morimoto, T., Ushio, T., & Kawasaki, Z. I. (2009). A fifth-power relationship for lightning activity from Tropical Rainfall Measuring Mission satellite observations. *Journal of Geophysical Research Atmospheres*, *114*(9), 1–10. doi: 10.1029/2008JD010370