

# JGR Atmospheres

## RESEARCH ARTICLE

10.1029/2023JD039420

### Key Points:

- A novel machine learning approach is used to calculate 250 years of maximum freeze depth for Northern Hemisphere seasonally frozen ground
- Under the SSP5-8.5 scenario, the maximum freeze depth in the Northern Hemisphere will decrease by 42% by 2099
- The maximum freeze depth decreases the most in the snow climate zone

### Correspondence to:

X. Peng,  
pengxq@lzu.edu.cn

### Citation:

Chen, C., Peng, X., Frauenfeld, O. W., Chu, X., Chen, G., Huang, Y., et al. (2024). Simulations and prediction of historical and future maximum freeze depth in the Northern Hemisphere. *Journal of Geophysical Research: Atmospheres*, 129, e2023JD039420. <https://doi.org/10.1029/2023JD039420>

Received 9 JUN 2023  
Accepted 30 JAN 2024

## Simulations and Prediction of Historical and Future Maximum Freeze Depth in the Northern Hemisphere

Cong Chen<sup>1</sup> , Xiaoqing Peng<sup>1,2</sup> , Oliver W. Frauenfeld<sup>3</sup>, Xinde Chu<sup>4</sup>, Guanqun Chen<sup>1</sup>, Yuan Huang<sup>1</sup>, Xuanjia Li<sup>1</sup> , Guangshang Yang<sup>1</sup>, and Weiwei Tian<sup>1</sup> 

<sup>1</sup>Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China, <sup>2</sup>National Cryosphere Desert Data Center, Lanzhou, China, <sup>3</sup>Department of Geography, Texas A&M University, College Station, TX, USA, <sup>4</sup>Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan, China

**Abstract** The maximum annual freeze depth (MFD) is a primary indicator of the thermal state of frozen ground, affecting ecosystems, hydrological processes, vegetation growth, infrastructure, and human activities in cold regions. It is thus important to quantify the past, present, and future spatial and temporal variability of MFD at the hemispheric scale. We develop a data-driven MFD simulation method within a machine learning framework, integrating MFD observations from meteorological stations and several environmental predictors, to analyze past and future scenarios in the Northern Hemisphere (NH). Based on ERA5 reanalysis estimates and historical to future CMIP6 scenarios, the NH MFD averaged 133 cm (ERA5) and 131 cm (CMIP6) during 1981–2010, and will vary 81–112 cm during 2015–2100 depending on the emission scenario. During 1950–2013, MFD decreased by 0.37 cm/a (ERA5) versus 0.22 cm/a (CMIP6), and is projected to decrease 0.16–0.69 cm/a by 2100. During 1981–2010, MFD decreased by an average of 19.1% (ERA5) and 13.9% (CMIP6), with a net change of −17 cm (ERA5) and −13 cm (CMIP6). Depending on the emission scenario, MFD will decrease 11% (−12 cm) to 42% (−19 cm) between 2015 and 2099 relative to the 1981–2010. Warming, increased moisture, warmer cold seasons, warmer warm seasons, shallower snow depths, and increased vegetation cover all lead to a reduction in MFD. The results from this novel machine learning approach provide useful insights regarding the fate of future frozen ground changes.

**Plain Language Summary** Seasonally frozen ground covers approximately half of the exposed ground surface in the Northern Hemisphere and is found in areas of intense human activity. There, the moisture retention and occurrence of freezing and thawing significantly impact agricultural production and infrastructure. Maximum freeze depth is a key indicator of the status of seasonally frozen ground. We simulate and predict the spatial distribution of maximum freeze depth at the hemispheric scale and quantify the variability of maximum freeze depth over past and future periods. Depending on the choice of future emission scenario, average maximum freeze depth in the Northern Hemisphere will decrease by 11% (−12 cm) to 42% (−19 cm) between 2015 and 2099, relative to the base period (1981–2010).

## 1. Introduction

Seasonally frozen ground (SFG) is soil that is frozen in winter and thaws completely in summer, and the maximum annual freeze depth (MFD) is an important indicator of the thermal state of SFG. As climate warms, the thickness of the active layer over permafrost increases, ground temperature rises, and permafrost rapidly degrades. Similarly, SFG is also undergoing significant changes, characterized by continuous decreases of MFD, delays in the onset of soil freeze, and an overall decrease in the length of the freezing period. These changes reflect the comprehensive impact of changes in temperature, soil moisture, and other soil characteristics. The occurrence of these changes in turn affects regional hydrologic processes, the surface energy budget, and biogeochemical processes. The reduction of MFD leads to a redistribution of surface energy and soil moisture, which affects ecohydrological processes throughflow production, groundwater-surface water interactions, soil moisture conditions, among other implications (Cheng & Wu, 2007; Jin et al., 2009; Yang et al., 2010). In addition, the soil's seasonal freeze-thaw cycles affect the moisture retention capacity in agricultural areas, where the soil freezes and stores water near the surface during winter and subsequently provides moisture for seeds the following spring, and the freeze-thaw cycle inhibits evaporation and alleviates drought in early spring (Baiz et al., 2008; Maximov et al., 2008; Sugimoto et al., 2002). The MFD also determines the length of the growing season and the yield of the

crops, thus impacting local agriculture. Additionally, the construction of infrastructure projects is complicated by frozen soil, with freezing-thaw cycles leading to uneven deformation of roadbeds, which seriously affect the safety of traffic (Wu et al., 2016; Zhang et al., 2008). Therefore, MFD is an important consideration in the design of the built environment in cold regions, and the determination of MFD in the planning of infrastructure projects is an important prerequisite to ensure the stability of the frozen ground.

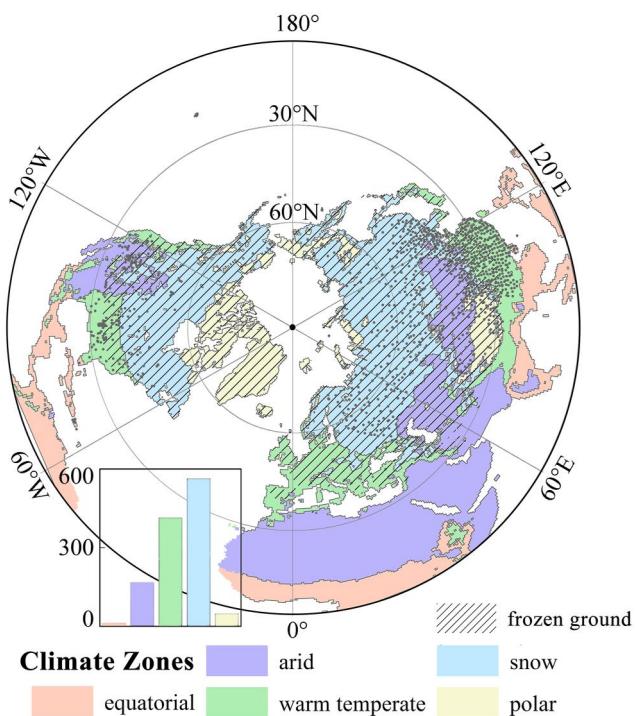
Changes in SFG have been quantified using a variety of approaches. As with permafrost, observations and simulations are the two primary tools for estimating SFG and its changes. The most direct way to observe the MFD of SFG is via borehole measurements, however, such measurements are difficult to obtain, costly, and limited to the point scale. Observational data are highly accurate, but long time series are sparse. In recent years, model simulations have become a common approach for regional-scale MFD estimates. One such approach is based on empirical models such as Kudryavtsev's formula (Kudryavtsev et al., 1977), Nelson's formula (Nelson, 1986; Nelson & Outcalt, 1987), and Stefan's equation (Harlan, 1978). The Stefan equation has a simple structure and high computational efficiency, and is a common approach used to calculate MFD. The Stefan solution has been applied to determine past and future MFD in Eurasia, finding decreases of  $-0.49\text{ cm/decade}$  during 1850–2005. Under a high emission scenario, soil freezing depth could decline as much as  $-4.58\text{ cm/decade}$  (Peng, 2017). However, the uncertainty of the required parameters and driving variables is large, thus leading to errors in the estimates. Physical models, generally coupled with terrestrial or hydrological models, can also be applied to simulate and predict MFD, but their accuracy and computational efficiency is still poor (Walvoord & Kurylyk, 2016; Westermann et al., 2016; Zhang et al., 2017). Statistical models and machine learning are alternate approaches, particularly when the relationship between variables is difficult to describe explicitly. Estimation of MFD is achieved by establishing the statistical relationship between MFD and environmental variables. For example, Luo et al. (2020) established the multivariate linear relationships between MFD and latitude, elevation, temperature, and precipitation, to simulate the spatial distribution of MFD on the Tibetan Plateau from 1960 to 2014. The mean MFD averaged 143.5–153.6 cm, and decreased by an average of 4.9 cm per decade. Wang et al. (2019) estimate the distribution of permafrost and the maximum thickness of seasonally frozen ground for a 2003–2010 baseline and for the future (2040s and 2090s) using statistics and machine learning algorithms. The most dramatic decrease is projected for the southwestern Tibetan Plateau, exceeding 50 cm in the 2090s compared to the baseline period. The projections of future frozen ground changes can enable policy makers to develop appropriate mitigation and adaptation strategies prior to the anticipated changes (Guo & Wang, 2016).

Rapid global warming has led to a highly unstable thermal state of frozen ground and current research is limited by focusing primarily on specific regions, lacking simulations and comprehensive assessments of MFD at hemispheric scales. Therefore, the goal of this study is to apply reanalysis observations and climate model output in conjunction with a machine learning algorithm to, for the entire NH, (a) construct a MFD model, (b) map the spatial distribution of the MFD, and (c) project the response of MFD to future climate change, ultimately providing a 250-year record of the changes in MFD. Because changes in MFD are important indicators of contemporary and future climate, our comprehensive simulations are important to inform environmental change, ecological protection, agricultural and livestock production, and engineering applications for the built environment of NH cold regions.

## 2. Data and Methods

### 2.1. Observed MFD

Site data for MFD are used to train and cross-validate the NH MFD estimates. Site observations include soil temperature, air temperature, precipitation, and snow depth. Daily soil temperature data are used to calculate the potential MFD, as in previous studies (Peng et al., 2017, 2020). A total of 1,220 stations covering the NH are collected for China, Russia, Finland, USA, and Canada (Figure 1). In China there are 635 stations from the China Meteorological Administration (<http://data.cma.cn/en/>) with daily observations starting in 1951. Russia has 264 stations with daily observations, recorded from 1963 on, which we downloaded from the Russian Research Institute of Hydro-meteorological Information -World Data Center (<http://meteo.ru/data/>). The United States has 208 sites with daily observations starting in 1993 through 2021, from the United States Department of Agriculture's Natural Resources Conservation Service (<https://www.nrcs.usda.gov/wps/portal/wcc/home/snowClimateMonitoring/soilClimateConditions/>). Canada has 32 stations from the Oak Ridge National Laboratory Distributed Active Archive Center (<https://doi.org/10.3334/ORNLDAAC/1335/>) with daily observation for



**Figure 1.** The spatial distribution of station locations and the five main climatic zones, with stippling denoting the frozen ground region of the NH. The histogram shows the number of stations in each climate zone, as derived from the Köppen-Geiger classification system.

1996–2014. Finland has a total of 81 sites, with daily data for 2000–2021 from the Finnish Meteorological Institute (<https://litdb.fmi.fi/>). Soil temperature observations at the sites ranged from 0 to 3.2 m in depth. We linearly interpolate the soil freeze depth as the 0°C contour based on soil temperatures throughout the soil profile, and the maximum daily value of each cold season is considered as the MFD for that year (Chen et al., 2022; Frauenfeld et al., 2005).

## 2.2. CMIP6 Output

We use the output for five variables: air temperature, precipitation, snow depth, surface downwelling shortwave radiation, and leaf area index. Of the suite of models developed by various international research institutions and participating in phase 6 of the Coupled Model Intercomparison Project (CMIP6), 22 of those models (Table 1, <https://esgf-node.llnl.gov/projects/cmip6/>) provide these five variables and will therefore be used here. These model outputs are provided in different spatial resolutions but cover the same period from 1850 to 2100. Output from five CMIP6 experiments will be used: a historical experiment for 1850–2014 and projections corresponding to four emission pathways for 2015–2100. The historical climate simulation experiment is an evolution of historical climate from 1850 to 2014 driven by external forcing based on actual observations. For future projections, the CMIP6 models use a framework that combines Shared Socioeconomic Pathways (SSPs) and Representative Concentration Pathways to describe various potential future development of society. The selected pathways are the four “Tier 1” scenarios suggested by O’Neill et al. (2014): SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. The CMIP6 model output was regressed to a common resolution of  $0.5^\circ \times 0.5^\circ$  using bilinear interpolation.

## 2.3. ERA5-Land Data

The European Centre for Medium-Range Weather Forecasts (ECMWF) fifth-generation monthly reanalysis dataset (ERA5) is used as input for the NH MFD model and includes hourly climate variables such as air temperature, precipitation, snow depth, radiation, and leaf area index. ERA5 is a reanalysis dataset providing a consistent view of the water and energy cycles at surface level during several decades (Hersbach et al., 2020). The ERA5-Land version has an equivalent native grid of  $0.1^\circ$  latitude  $\times 0.1^\circ$  longitude, and we downloaded the dataset at this resolution. It contains a detailed record from 1950 onwards. The overlap with the CMIP6 model historical experimental is from 1950 to 2014, a total of 65 years.

## 2.4. Distribution of Non-Permafrost Areas in the NH

Zhang et al. (2003) mapped the distribution of SFG and intermittently frozen ground in the NH based on the mean monthly air temperature (MMAT) by establishing the relationship between the near surface soil freeze/thaw status and MMAT, that is, the relationship between MMAT and monthly total number of days of soil freezing at 5 cm depth. According to this, the extent of SFG is recognized as region north of the 0°C isotherm of MMAT obtained from a surface air temperature climatology. The extent of the intermittently frozen ground is determined as the zone between 0 and 5°C isotherms of MMAT. In permafrost regions, not all areas are actually underlain by permafrost, but are instead SFG. For a comprehensive SFG assessment, our MFD simulations mask the NH non-frozen soil areas, but do include permafrost areas as potential MFD (Figure 1).

## 2.5. Climate Zone Data

To provide further context for the distribution and variability of MFD, the Köppen-Geiger climate classification system is used to categorize the NH into climatic zones and statistically assess the MFD for the different climate zones. The Köppen-Geiger climate classification (<http://koeppen-geiger.vu-wien.ac.at/>) is based on temperature, precipitation, and vegetation. The global land area is divided into five main climate zones, that is, equatorial, arid,

**Table 1**  
*Description of CMIP6 Models Used in This Study*

Model	Spatial resolution (lat × lon)	Country	Experiment				
			Historical	SSP1-2.6	SSP2-4.5	SSP3-7.0	SSP5-8.5
BCC-CSM2-MR	160 × 320	China	✓	✓	✓	✓	✓
BCC-ESM1	64 × 128	China	✓	-	-	-	-
CAS-ESM2-0	128 × 256	China	✓	✓	✓	✓	✓
CESM2	192 × 288	United States	✓	-	-	-	-
CESM2-WACCM	192 × 288	United States	✓	✓	✓	✓	✓
CMCC-CM2-HR4	192 × 288	Italy	✓	-	-	-	-
CMCC-CM2-SR5	192 × 288	Italy	✓	✓	✓	✓	✓
CMCC-ESM2	192 × 288	Italy	✓	✓	✓	✓	✓
EC-Earth3-CC	256 × 512	European Union	✓	-	✓	-	✓
EC-Earth3-Veg	256 × 512	European Union	✓	✓	✓	✓	✓
EC-Earth3-Veg-LR	160 × 320	European Union	✓	✓	✓	✓	✓
FGOALS-g3	80 × 180	China	✓	✓	✓	✓	✓
GFDL-ESM4	180 × 288	United States	✓	✓	✓	✓	✓
GISS-E2-1-G	90 × 144	United States	✓	-	-	-	-
GISS-E2-1-H	90 × 144	United States	✓	-	-	-	-
GISS-E2-2-G	90 × 144	United States	✓	-	-	-	-
GISS-E2-2-H	90 × 144	United States	✓	-	-	-	-
KIOST-ESM	96 × 192	Korea	✓	✓	✓	-	✓
NorESM2-LM	96 × 144	Norway	✓	✓	✓	✓	✓
NorESM2-MM	192 × 288	Norway	✓	✓	✓	✓	✓
SAM0-UNICON	192 × 288	Korea	✓	-	-	-	-
TaiESM1	192 × 288	Taiwan, China	✓	✓	✓	✓	✓

*Note.* A check mark (✓) indicates that output from that experiment was available, while a dash (-) indicates that the experiment did not have output available for the particular scenario.

warm temperate, snow, and polar climate (Figure 1). This paper masks the non-permafrost zone, so the zonal statistics do not include the equatorial climate.

## 2.6. Soil Texture Data

A comprehensive, gridded Global Soil Dataset for use in Earth System Models (GSDE) was developed at Sun Yat-sen University's Land-Atmosphere Interaction Research Group. GSDE provides soil information including soil particle-size distribution, organic carbon, nutrients, etc. GSDE is based on the Soil Map of the World and various regional and national soil databases, including soil attribute data and soil maps (<http://globalchange.bnu.edu.cn/research/soilw>). The resolution is 30 arc-seconds (approximately 1 km at the equator). Four soil attributes are used in this paper, namely, sand content (% of weight), silt content (% of weight), clay content (% of weight) and gravel content (% of volume).

## 2.7. Methods

Twelve variables that affect the thermal state of soil were selected as inputs for the training model and simulations. Because MFD is an annual variable, these variables are applied at the annual time scale: average annual air temperature, annual cumulative precipitation, annual air freezing index, annual air thawing index, average annual snow depth, maximum annual snow depth, average annual surface downwelling shortwave radiation, average annual leaf area index, and soil texture (clay content, gravel content, sand content, and silt content). Although there is likely autocorrelation in these environmental variables that may affect the

accuracy of the model, we did not correct for this in our analysis. Air temperature is considered to be the most direct factor affecting the MFD of the soil (Frauenfeld, 2004; Frauenfeld & Zhang, 2011), and the freezing/thawing index is a measure of the duration and magnitude of the cold/warm season, indicating the depth and intensity of the freeze-thaw action (Liu et al., 2016; Streletskiy et al., 2015; Wang et al., 2015; Wu et al., 2016). Although MFD is an annual variable, it is nonetheless influenced by seasonal climatic variability that is inherently included and integrated by the freezing and thawing indices. The annual freezing/thawing index is calculated from the monthly average temperature (Frauenfeld et al., 2007). The air freezing and thawing indices are the cumulative temperature of months with air temperatures below or above 0°C, expressed in °C-days. The freezing index is calculated from 1 July to 30 June of the following year, while the thawing index is calculated from 1 January to 31 December of each year as follows:

$$FI = \sum_{i=7}^{M_F} \overline{T_i} | \cdot D_i (\overline{T_i} < 0)$$

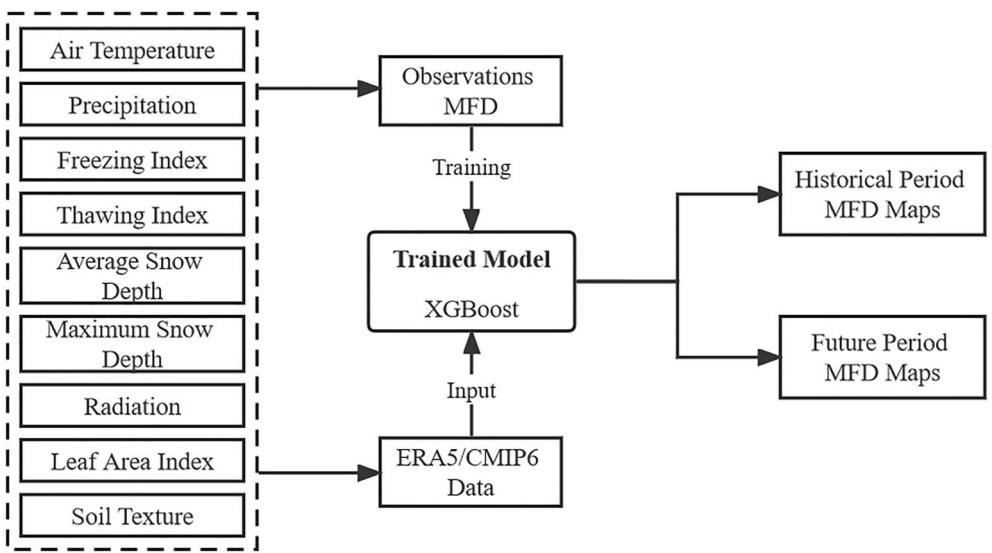
$$TI = \sum_{i=1}^{M_T} \overline{T_i} | \cdot D_i (\overline{T_i} > 0)$$

where FI is the freezing index, TI the thawing index,  $\overline{T_i}$  is the average temperature in the  $i$ th month multiplied by  $D_i$ , the number of days in the  $i$ th month. The months with negative monthly average temperatures from July to June of the following year are summed to obtain the annual FI. Similarly, TI is the sum of monthly accumulated temperatures for months with positive average monthly temperatures.

Precipitation determines the redistribution of surface water and groundwater, and thus the frozen ground's thermal state. Solar radiation maintains the atmospheric and surface temperatures and is the main source of heat (Ala-Aho et al., 2021). Different types of land cover such as snow, vegetation, and soil types directly affect the surface energy balance (Frauenfeld & Zhang, 2011; Luo et al., 2009; Luo & Lü, 2009; Shiklomanov, 2012; Yang et al., 2019). The magnitude of the snow depth has two types of effects on soil temperature including insulation and cooling (Frauenfeld, 2004; Zhang, 2005). The average annual snow depth and the maximum annual snow depth are calculated according to the snow year (July 1 of the year to June 30 of the following year). Vegetation affects the freezing depth of the soil by influencing the net radiation at the surface (Chang et al., 2012; Swann et al., 2010). Soil texture determines the thermophysical properties of the soil in addition to the thermal and hydraulic conductivity (Nelson, 1987).

Xtreme Gradient Boosting (XGBoost), a machine learning algorithm, was used to construct a NH MFD model based on station observations of MFD as training data, and ERA5-Land reanalysis data and CMIP6 model output data were regressed to simulate the past and predicted future spatial distribution of MFD in the NH (Figure 2). XGBoost is a novel gradient boosting decision tree algorithm proposed by Chen and Guestrin (2016). Its core idea is to gradually improve the predictive ability of a model through gradient boosting (Chen & Guestrin, 2016). The observed MFD is divided into a training set and a test set at an 8:2 ratio, which is used for training and evaluating the model, respectively. Glaciers and lakes in the NH were excluded from all simulations. It is assumed that the soil texture will not change in the future. Linear regression was used in the calculation of multi-year trends for time series and climate sensitivity analyses, with a 95% statistical significance level.

In the model training process, ten-fold cross-validation was used and run 300 times to achieve the optimal model. Root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination ( $R^2$ ) were selected as evaluation metrics. The smaller the RMSE and MAE, the higher the prediction accuracy of the model.  $R^2$  reflects the degree of fit between the predicted and measured values, and the closer the  $R^2$  is to 1, the better the model goodness-of-fit and the higher the prediction accuracy. The model errors shown in Table 2 are calculated for the test set. The model based on the ERA5 reanalysis had the smallest errors and the highest accuracy for the simulation of MFD. Among the 22 CMIP6 models, CAS-ESM2-0 had the smallest MFD errors and MFD based on the EC-Earth3-CC model had the largest errors. The 22 models' average RSME is 33.15 cm, MAE is 22.96 cm, and  $R^2$  is 0.81.



**Figure 2.** Flow chart of MFD simulation prediction in the NH.

### 3. Results

#### 3.1. Spatial Distribution of MFD

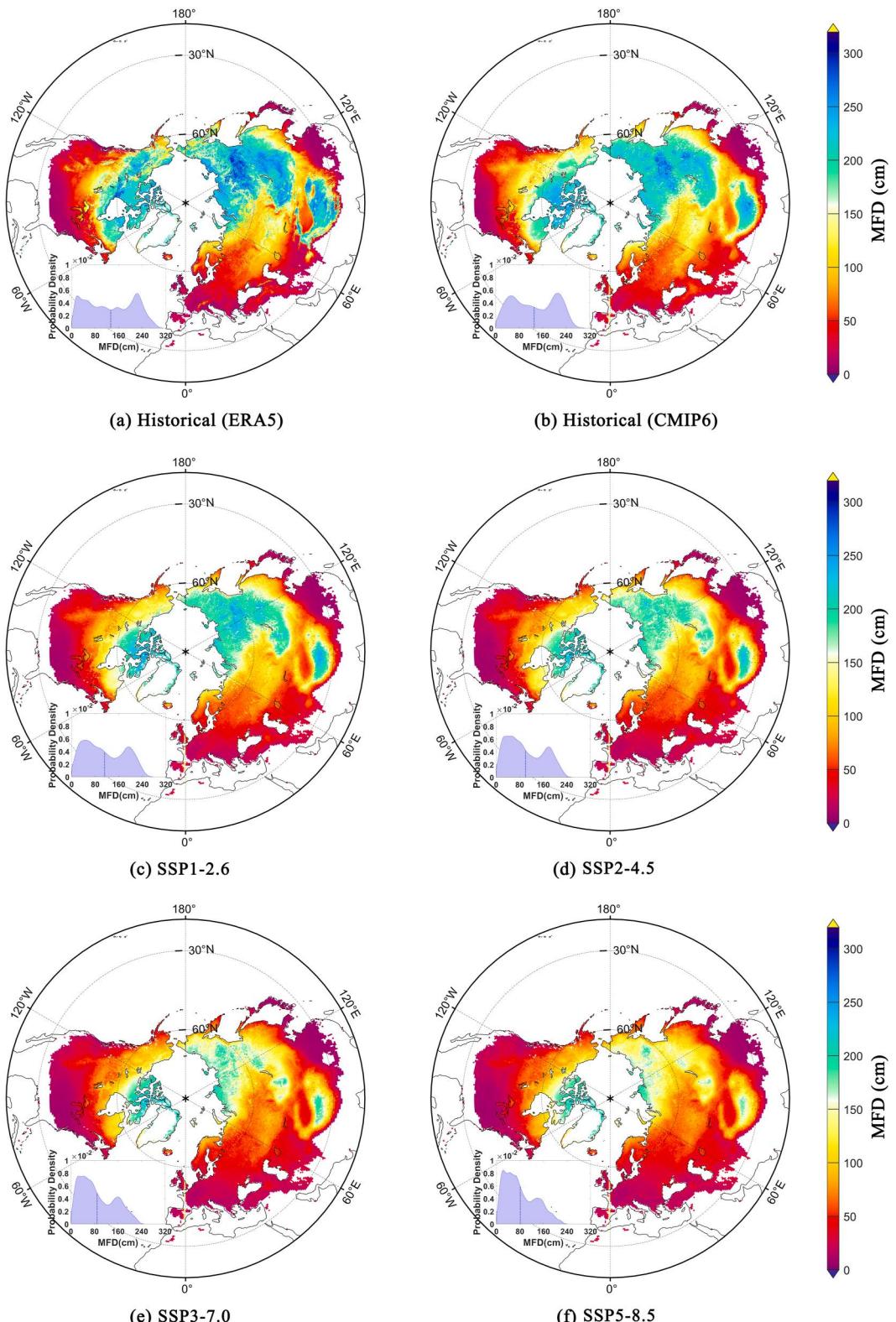
Thirty-year average MFD was simulated for the historical period (1981–2010) and for four future scenarios (2070–2099). In both the historical and future periods, MFD shows expected decreases from high to low latitudes and increases from low to high elevations, with greatest values in the circumpolar region and the Tibetan Plateau (Figure 3). The probability density function insets show that there are two peaks of MFD in the historical period, at about 50 and 210 cm, respectively. It reflects the distribution of MFD in the NH, with MFD concentrated around 50 cm at low latitude and low elevation areas and around 210 cm at high latitude and high elevation areas. Similar peaks are evident in the future, but the peak at 210 cm shifts toward much lower values and, overall, there are also greater probabilities toward shallower freeze depths. During the historical climatology, the average NH MFD was 133 cm (ERA5) and 131 cm (CMIP6), with greatest regional MFD at high latitudes (e.g., Siberia) and high elevations (e.g., the Tibetan Plateau) reaching 312 cm (ERA5) and 288 cm (CMIP6). Based on SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, MFDs will decrease to 112 cm, 100 cm, 89 cm, and 81 cm, respectively, during the end-of-century climatology, with greatest regional MFD values of only 261 cm, 240 cm, 237 cm, and 225 cm, respectively.

The hemispheric-average MFD during the historical climatology (1981–2010) for the individual 22 models (Figure 4) ranged from 122 cm (GFDL-ESM4) to 140 cm (NorESM2-LM), with NorESM2-MM comparing best to the ERA5 estimate of 133 cm. At end of the century, the hemispheric-mean MFD for the 13 models with available SSP1-2.6 output (Figure 4) will range from 98 cm (CMCC-ESM2) to 123 cm (NorESM2-LM). For the SSP2-4.5 scenario, 14 models show MFD will range from 87 cm (CMCC-CM2-SR5) to 109 cm (NorESM2-MM). The SSP3-7.0 scenario indicates an average MFD based on 12 models that will range from 75 cm (EC-Earth3-Veg) to 99 cm (TaiESM1). Finally, the hemispheric-average MFD for 14 models will range from 66 cm (EC-Earth3-Veg) to 95 cm (KIOST-ESM), if the SSP5-8.5 scenario holds true by the end of the century.

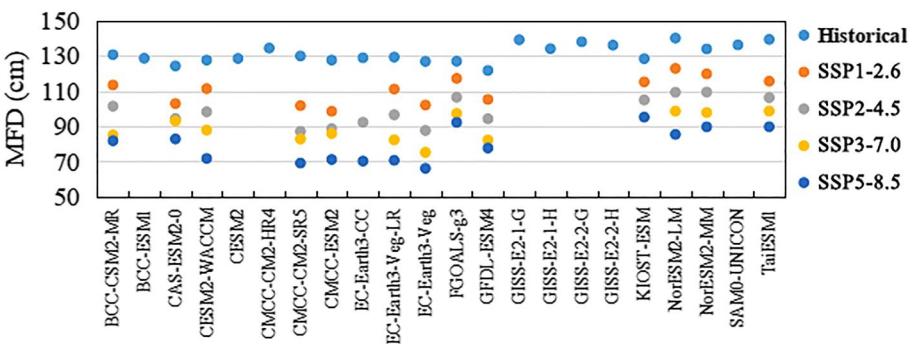
**Table 2**

*The Validation Metrics of MFD for Each Model Simulation*

Data	RMSE (cm)	R <sup>2</sup>	MAE (cm)
BCC-CSM2-MR	34.21	0.79	23.83
BCC-ESM1	32.85	0.81	23.11
CAS-ESM2-0	31.09	0.82	21.80
CESM2	33.48	0.80	23.62
CESM2-WACCM	32.65	0.81	22.64
CMCC-CM2-HR4	31.73	0.82	22.29
CMCC-CM2-SR5	32.53	0.81	22.74
CMCC-ESM2	32.66	0.81	22.63
EC-Earth3-CC	35.51	0.78	24.08
EC-Earth3-Veg	34.07	0.79	23.55
EC-Earth3-Veg-LR	33.36	0.80	23.36
FGOALS-g3	31.62	0.83	22.22
GFDL-ESM4	33.86	0.80	23.33
GISS-E2-1-G	32.97	0.81	22.75
GISS-E2-1-H	32.82	0.82	22.84
GISS-E2-2-G	33.04	0.81	22.41
GISS-E2-2-H	32.34	0.82	22.05
KIOST-ESM	33.76	0.79	23.40
NorESM2-LM	35.12	0.79	24.08
NorESM2-MM	33.17	0.81	22.92
SAM0-UNICON	33.28	0.81	22.57
TaiESM1	33.28	0.80	22.96
Multi-Model Mean	33.15	0.81	22.96
ERA5	30.59	0.83	20.87



**Figure 3.** Spatial distribution of average annual MFD for the historical period (1981–2010) based on (a) ERA5 and (b) the CMIP6 multi-model mean, and for the future scenarios (2070–2099) based on (c) SSP1-2.6, (d) SSP2-4.5, (e) SSP3-7.0, and (f) SSP5-8.5.

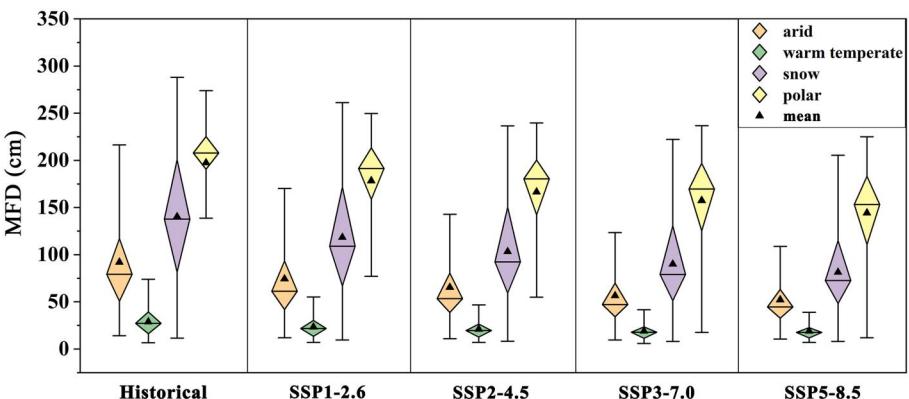


**Figure 4.** Hemispheric average MFD for the individual 22 CMIP6 models.

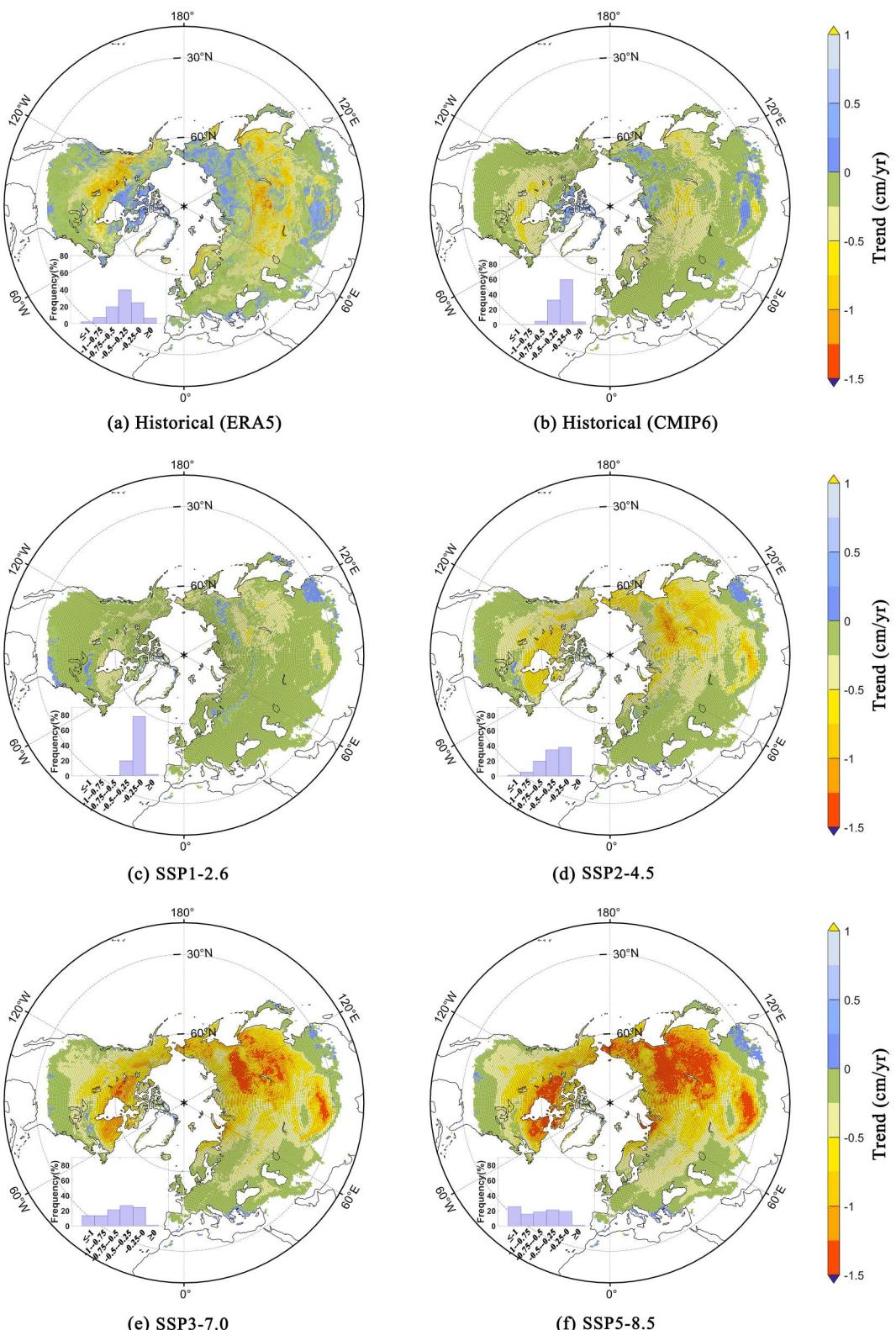
For a clearer understanding of the spatial distribution of MFD across the NH in the historical and future periods, we divided MFD variability according to climatic zones. The resulting distribution shows that under both historical (1981–2010) and future periods (2070–2099), the MFD of the four major climatic zones, ranked from large to small, are the polar climatic zone, snow climatic zone, arid climatic zone, and warm temperate climatic zone (Figure 5). For the historical to future scenarios of SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, the climate-zone average MFD sequentially decreases. The average MFD in the warm temperate climate region during the historical period was 29 cm and will range 19–23 cm in the future, depending on the climate scenario. In the arid climate region, MFD was 92 cm historically, and will decrease to 52–74 cm in the future. The average MFD in the snow climate region used to be 140 cm, with projections ranging 81–118 cm. Finally, in the polar region, MFD was 197 cm but will decrease to 144–178 cm.

### 3.2. Trends of MFD

To understand the temporal changes in MFD, long-term trends for the NH were calculated for the historical period (1950–2013) and the future (2015–2099) based on the multi-model means for the four emission scenarios. Most of the regions show a significant decreasing trend, most strongly in high-latitude and high-elevation regions (Figure 6). Especially in the multi-model averaged SSP5-8.5 scenario, MFD rapidly decreases. While decreases in MFD were potentially between  $-0.22 \text{ cm/a}$  (CMIP6) and  $-0.37 \text{ cm/a}$  (ERA5) during the historical period, future trends range from  $-0.16$  to  $-0.69 \text{ cm/a}$ . During 1950–2013, 60% of the NH had MFD trends between 0 to  $-0.25 \text{ cm/a}$ , and 32% of the hemisphere ranged between  $-0.5$  to  $-0.25 \text{ cm/a}$ . In the future (2015–2099) under the SSP1-2.6 scenario, 78% of the NH's MFD will see changes ranging 0 to  $-0.25 \text{ cm/a}$ , and 20% of the area will be between  $-0.5$  to  $-0.25 \text{ cm/a}$ , with the decreasing trend of MFD in some areas being lower. For the SSP2-4.5 and SSP3-7.0 scenarios, the decreases in MFD will be significantly larger compared to the historical period. In the SSP5-8.5 scenario, the decrease in MFD in more than 25% of the NH will be greater than  $-1 \text{ cm/a}$ .



**Figure 5.** Historical (1981–2010) and future (2070–2099) regional average MFD for the different climatic zones.



**Figure 6.** Spatial distribution of MFD trends for the historical period (1950–2013) based on (a) ERA5 and (b) CMIP6, and (c)–(f) the four future scenarios (2015–2099). Stippled regions indicate statistically significant trends (95% level).

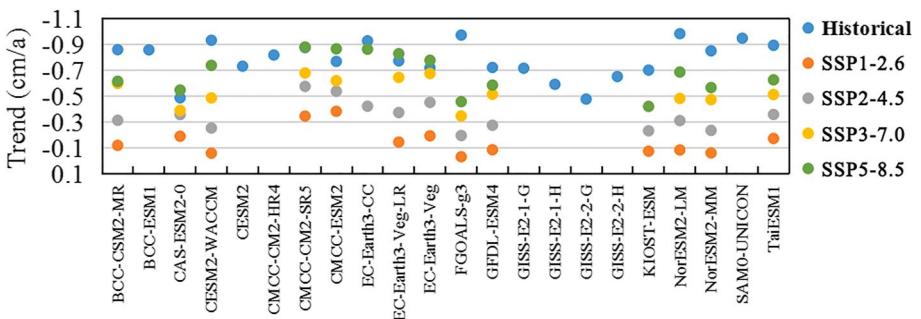


Figure 7. Hemispheric average MFD for the individual 22 CMIP6 models.

The spatial distribution of long-term trends in MFD for each CMIP6 model in the historical period (1981–2010) show that the SFG in the transition zone along the permafrost margins experienced the most significant changes in terms of the strongest decreases in MFD. The 22 models indicate regional average trends ranging from  $-0.48 \text{ cm/a}$  (GISS-E2-2-G) to  $-0.98 \text{ cm/a}$  (NorESM2-LM), with BCC-ESM1 having the greatest trends of up to  $-4.92 \text{ cm/a}$ . In the future period (2015–2099), the hemispheric-average MFD decreases for the 13 SSP1-2.6 model scenarios will range from  $-0.10 \text{ cm/a}$  (FGOALS-g3) to  $-0.44 \text{ cm/a}$  (CMCC-ESM2). For SSP2-4.5, the hemispheric-average MFD for 14 will range from  $-0.28 \text{ cm/a}$  (FGOALS-g3) to  $-0.63 \text{ cm/a}$  (CMCC-CM2-SR5). The NH-average MFD for the 12 models under the SSP3-7.0 scenario is projected to decrease  $-0.42 \text{ cm/a}$  (FGOALS-g3) to  $-0.73 \text{ cm/a}$  (CMCC-CM2-SR5). The greatest NH trends, under the SSP5-8.5 scenario based on 14 models, will be  $-0.51 \text{ cm/a}$  (KIOST-ESM) to  $-0.93 \text{ cm/a}$  (EC-Earth3-CC). The maximum MFD changes for the SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 climate scenarios will be  $-2.06 \text{ cm/a}$  (CMCC-ESM2),  $-2.56 \text{ cm/a}$  (CMCC-CM2-SR5),  $-2.97 \text{ cm/a}$  (NorESM2-LM), and  $-3.50 \text{ cm/a}$  (CESM2-WACCM), respectively (Figure 7).

According to Berkeley Earth (2020), there are three change-points in the trend of global mean temperature during the historical 1850–2022 period, around 1910, 1940, and 1970. To highlight the MFD changes in the context of these change-points, we divide our record into the same three time periods: 1911–1940, 1941–1970, and 1971–2000. The NH MFD initially decreased at a rate of  $-0.20 \text{ cm/a}$ , then deepened by  $0.19 \text{ cm/a}$ , followed by a strong decrease of  $-0.34 \text{ cm/a}$  (Figure 8). The greatest MFD trends in the NH increased from  $-0.65 \text{ cm/a}$  at the beginning of the 20th century to  $-2.27 \text{ cm/a}$  at the end of the 20th century, primarily in the transition zone at the edge of the permafrost region.

Figure 9 shows the NH average trend of MFD for the long-term time series (1850–2099). In the historical period (1850–2013), MFD in the NH decreases at a rate of  $-0.05 \text{ cm/a}$ . The MFD calculated based on ERA5 data (1950–2020) decreases at a rate of  $-0.22 \text{ cm/a}$ . In the future period (2015–2099), the MFD shows a significant

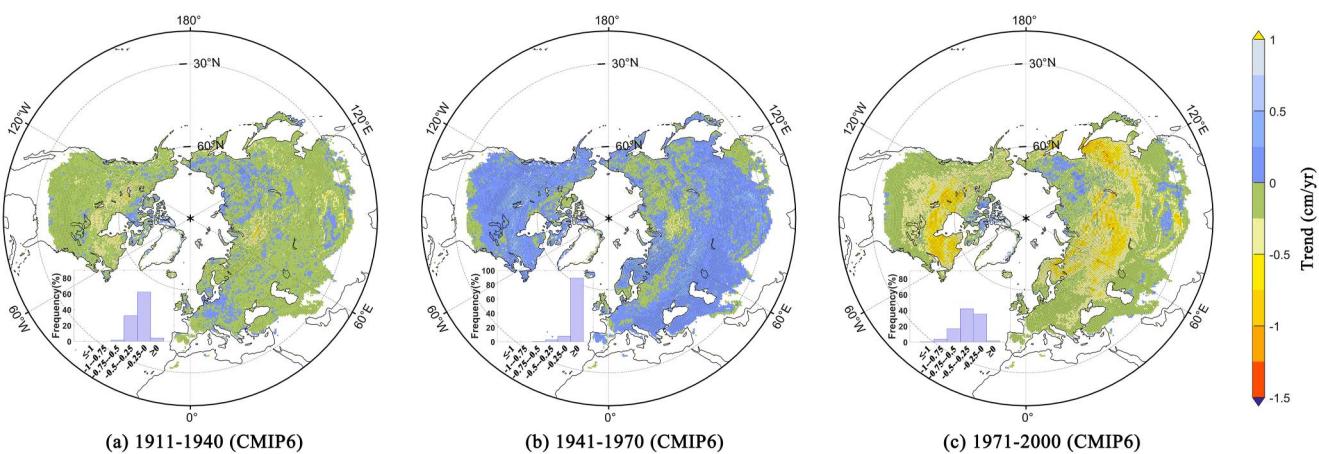
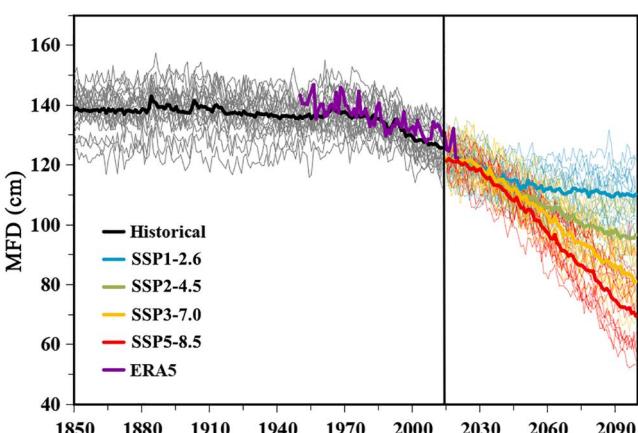


Figure 8. Multi-model mean trends in MFD under different historical periods. Stippled regions indicate statistically significant trends (95% level).



**Figure 9.** Long-term MFD variability in the NH. The thin lines represent the individual models during the historical (gray) and future (colors) periods.

and at low elevations in both the historical and future periods (Figure 11). This indicates that under global warming conditions, the soil in low elevation and low latitude areas will no longer freeze (MFD decreases to 0 cm). In the past 30 years, 60% of the regions have experienced MFD reductions by less than 15%. By the end of the century, MFD in the NH will be reduced by an average of 11%–42% depending on the emission scenario.

Similarly, the net changes in MFD in the different climatic zones indicate that the amount of MFD changes during the historical and future periods are greatest in the snow and polar climatic zones, and lowest in the arid and warm temperate climatic zone (Figure 12). From the historical to future SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 scenarios, the net MFD reduction in the warm temperate and arid climate region will be 3.8–9.3 cm and 11–42 cm, respectively. In the snow and polar climate regions, net MFD changes will range 15–67 cm and 15–67 cm, respectively, depending on the emission scenario.

## 4. Discussion

### 4.1. Potential Uncertainties

The spatial and temporal MFD characteristics in the NH were quantified based on station observations, reanalysis data, and climate model output. The XGBoost machine learning algorithm was applied to simulate and predict MFD, with an average RMSE and MAE of 33.15 and 22.96 cm, respectively, for the multi-model (22 models) MFD average. For our analyses, these errors correspond to the deviations between the CMIP6 model output and the meteorological variables measured at the stations. The MAE and RMSE is calculated between the CMIP6 model data and the station observations of average annual air temperature, annual cumulative precipitation, average annual snow depth, respectively (Figures 13–15). The RMSE of the annual mean temperature in the CESM2-WACCM model is the smallest with 2.5°C and the MAE is 1.8°C. The largest RMSE and MAE are for the GISS-E2-2-H model, at 5.1 and 4.2°C, respectively. The mean RMSE of the multi-model average is 3.5°C. The RMSE for annual cumulative precipitation ranged between 239 mm (EC-Earth3-Veg) and 468 mm (CMCC-CM2-SR5), and the MAE ranged between 173 and 345 mm. The annual mean snow depth in the CESM2-WACCM model has the smallest RMSE of 5.0 cm and MAE of 2.7 cm. The accuracy of these climate variables affects the simulation and assessment of the NH MFD.

An additional error source comes from the coarse resolution of the data. This study applies  $0.5^\circ \times 0.5^\circ$  data for the simulations, while field observations apply at the point scale. This mismatch in the resolution of the data affects the simulation accuracy of MFD. Nonetheless, although there are uncertainties in simulating and projecting MFD at the hemispheric scale, this study effectively represents the spatial distribution and temporal variability of MFD across the NH. Improving the simulation accuracy of MFD would necessitate increasing field monitoring and collecting more comprehensive observational data including soil parameters.

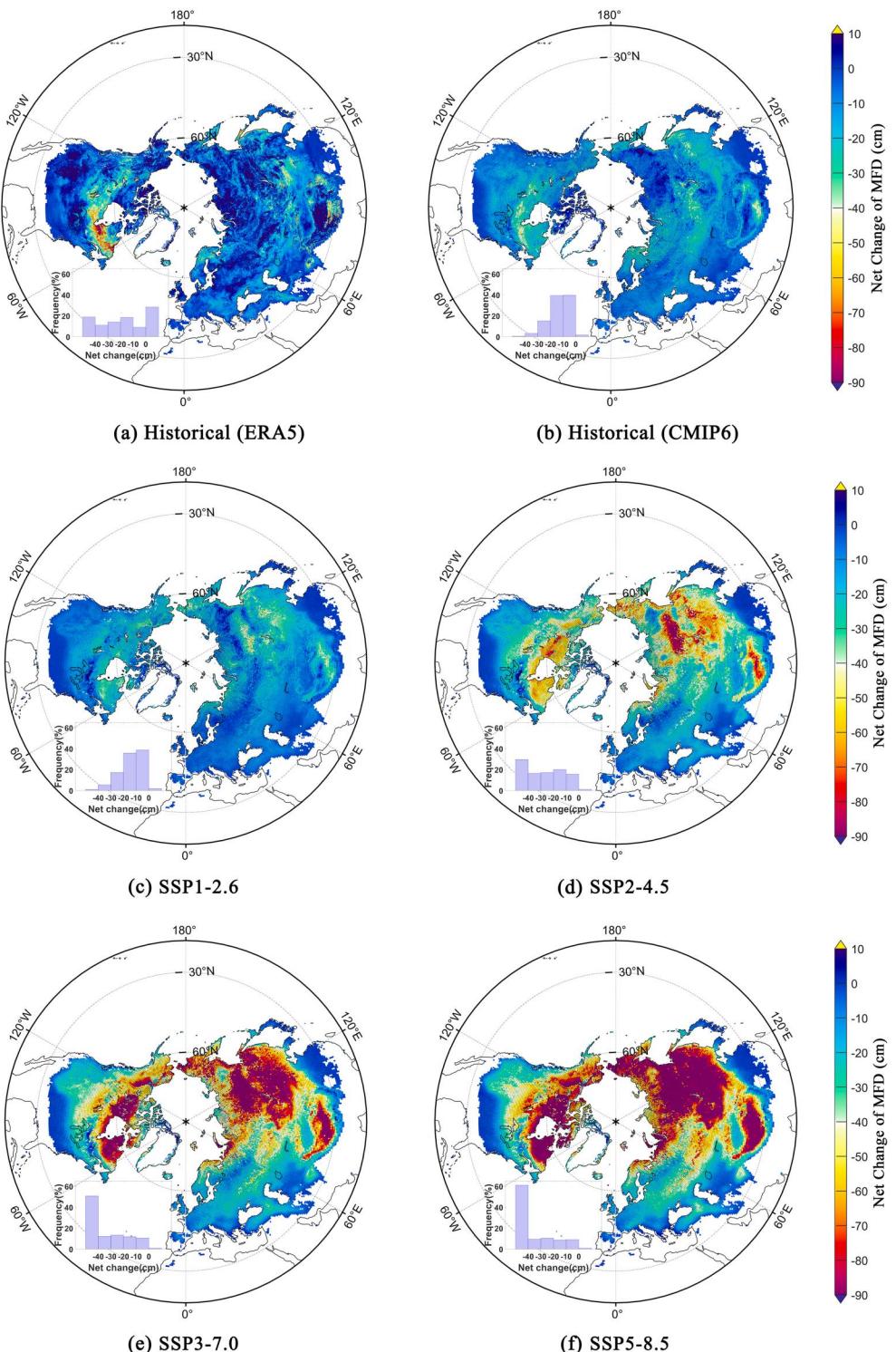
The uncertainties mentioned above result in potential errors in the MFD model, which affect the simulations and prediction accuracy of the MFD across the NH. The errors between the measured MFD at all stations and the

decrease with the rate of change from slow to fast:  $-0.15$  cm/a (SSP1-2.6),  $-0.35$  cm/a (SSP2-4.5),  $-0.53$  cm/a (SSP3-7.0), and  $-0.68$  cm/a (SSP5-8.5).

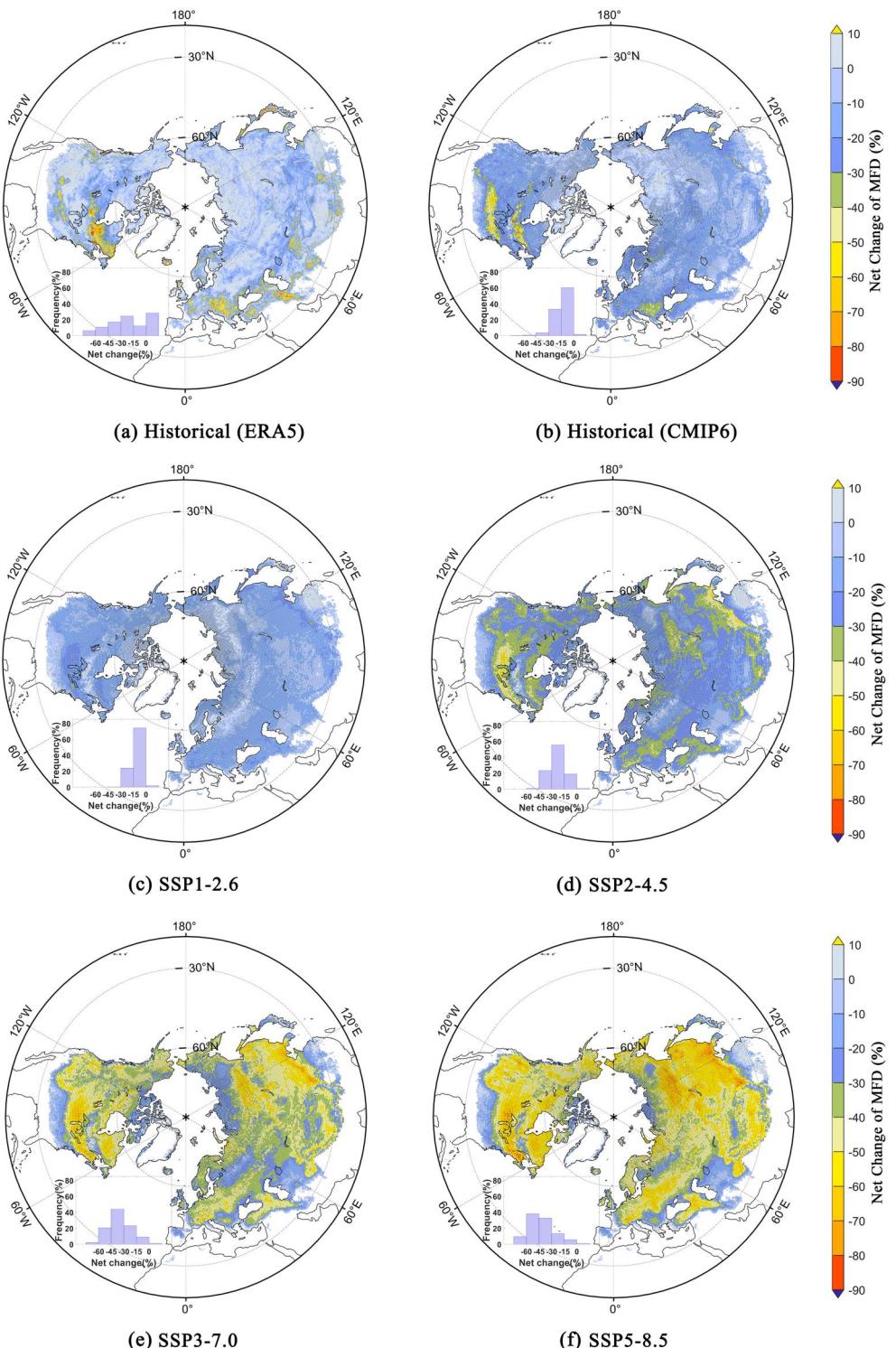
### 3.3. Net Change of MFD

To better analyze the magnitude of MFD changes in the NH during the historical and future periods, the net change in MFD for the past (1981–2010) and future (2015–2099) are calculated (Figure 10). The net change thus indicates how much the MFD has changed during 2015–2099 relative to 1981–2010. Net differences in MFD in the NH over the past 30 years are  $-17$  cm (ERA5) and  $-13$  cm (CMIP6). For the climate scenario projections corresponding to SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, the regional average net changes over the next 85 years will be  $-14$  cm,  $-30$  cm,  $-46$  cm, and  $-58$  cm, respectively. The magnitude of MFD change is more significant at high latitudes and elevations, for example, the greatest reduction is a maximum magnitude of up to 200 cm under the SSP5-8.5 scenario.

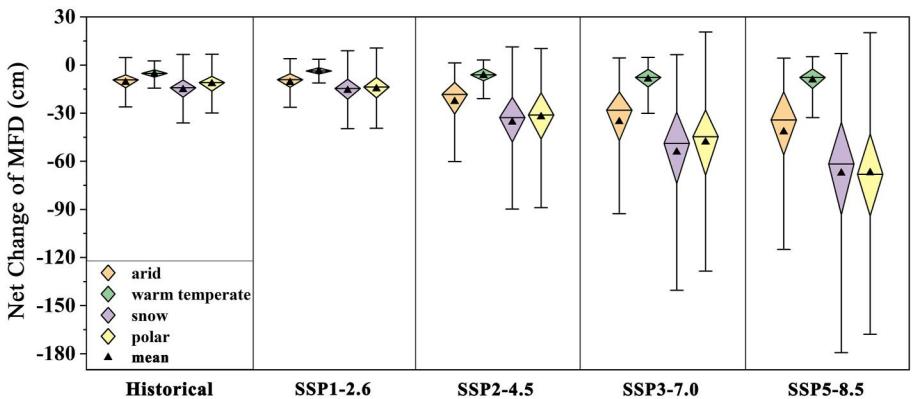
Expressing the net changes across the NH as a percentage relative to the 1981–2010 base period, the percent change in MFD is larger in low latitudes



**Figure 10.** Net change in MFD in the NH for the past (1981–2010) and future (2015–2099). The inset shows the frequency distribution of the net change in MFD in the NH. Stippled regions indicate statistically significant trends (95% level).



**Figure 11.** Percentage change in NH MFD for the past (1981–2010) and future (2015–2099) periods, respectively, relative to the base period (1981–2010). The inset shows the frequency distribution of the net change in MFD in the NH. Stippled regions indicate statistically significant trends (95% level).



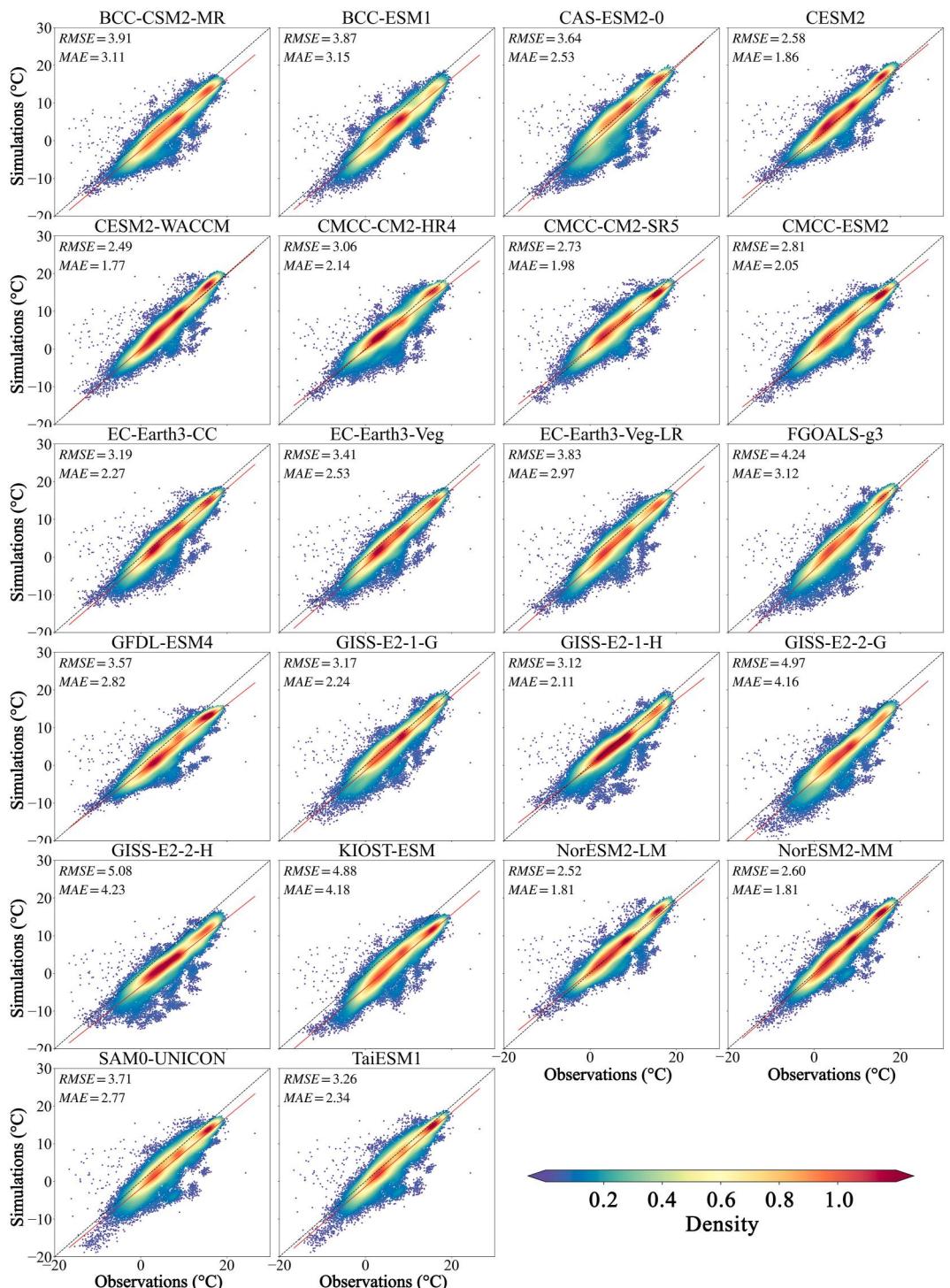
**Figure 12.** Historical (1981–2010) and future (2015–2099) regional average net change in MFD for the different climatic zones.

simulated values based on the 22 CMIP6 models (Figure 16) indicate that the RMSE of the GISS-E2-2-H model is the smallest, with 34.3 cm, and the MAE is 23.2 cm. The largest RMSEs and MAEs are for the BCC-CSM2-MR model, at 40.4 and 27.2 cm, respectively.

#### 4.2. Sensitivity to Climate Change

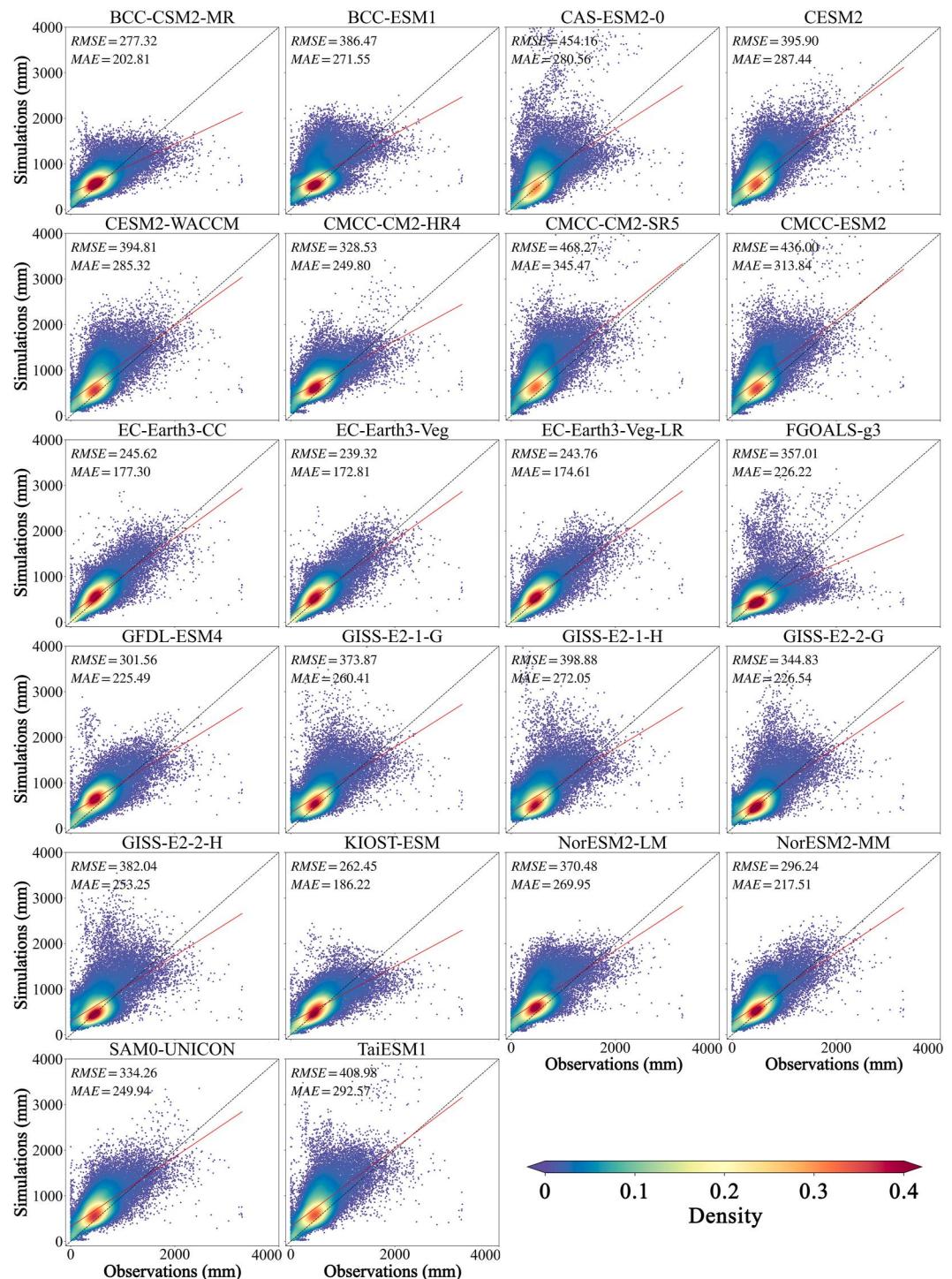
To understand the sensitivity of MFD to climate in the NH, the changes are statistically assessed and discussed. Using 1981–2010 as the base period, we determine the association between MFD and climate variables under the four future (2015–2099) scenarios (Figures 17 and 18). The variables are annual mean air temperature, annual cumulative precipitation, annual mean snow depth, annual maximum snow depth, annual freezing index, annual thawing index, annual mean surface downwelling shortwave radiation, and average annual leaf area index, and these are discussed and compared for the different climatic zones. For both the entire NH and the different climatic zones, the fluctuation of climate factors from SSP1-2.6 to SSP5-8.5 scenarios increases sequentially, reflecting a significant decrease in MFD. At the hemispheric scale, the sensitivity of MFD to climate factors other than radiation is more consistent for the four future scenarios, with the sensitivity (averaged for the four future scenarios) being  $-7.87 \text{ cm}^{\circ}\text{C}$  for air temperature,  $-0.40 \text{ cm/mm}$  for precipitation,  $0.44 \text{ cm}/10^{\circ}\text{C}\cdot\text{d}$  for freezing index,  $-0.42 \text{ cm}/10^{\circ}\text{C}\cdot\text{d}$  for thawing index,  $3.87 \text{ cm/cm}$  for average snow depth,  $3.17 \text{ cm/cm}$  for maximum snow depth,  $4.77 \text{ cm}/\text{W}\cdot\text{m}^{-2}$  for radiation, and  $-0.82 \text{ cm}/10^{-2}$  for leaf area index.

The sensitivity of MFD to climate change varies in the different climatic zones (Tables 3 and 4). The sensitivity of MFD to air temperature is lowest in the warm temperate climate, with an average of  $-2.08 \text{ cm}^{\circ}\text{C}$  for the four future scenarios, where the MFD is inherently low due to the milder climate. The sensitivity of MFD to air temperature is greatest in the snow climate, with an average of  $-9.07 \text{ cm}^{\circ}\text{C}$ . In this climate region, SFG is widely distributed, and the freeze-thaw cycle near the surface is the most sensitive to climate change, with rising temperatures directly affecting ground temperature. This is followed by the arid climate ( $-7.31 \text{ cm}^{\circ}\text{C}$ ) and the polar climate ( $-6.70 \text{ cm}^{\circ}\text{C}$ ), where the environment is harsh and permafrost persists all year round, buffering the reduction of seasonal freezing depth due to the presence of permafrost. When air temperature rises by  $1.5^{\circ}\text{C}$ , MFD is projected to decrease by an average of 10.63 cm in the warm temperate climate region, 4.45 cm in the arid climate region, 13.05 cm in the snow climate region, and 6.25 cm in the polar climate region. When air temperatures rises by  $2^{\circ}\text{C}$ , MFD is projected to decrease by an average of 14.28 cm (warm temperate climate), 5.49 cm (arid climate), 17.59 cm (snow climate), and 9.60 cm (polar climate). The sensitivity of MFD to precipitation is most significant in the arid climate zone, with an average of  $-0.52 \text{ cm/mm}$ . Precipitation increases the water content of near-surface soils, requiring more latent heat for soil freezing and thus accelerating the decrease in freezing depth (Luo et al., 2020). This is followed by the snow climate zone with an average of  $-0.44 \text{ cm/mm}$ . A small amount of rainfall increases the temperature of the soil surface layer, which reduces the freezing depth. Finally, in the polar and warm temperature climate, the sensitivity to precipitation is  $-0.29 \text{ cm}/\text{mm}$  and  $-0.08 \text{ cm}/\text{mm}$ , respectively. Air temperature most directly affects the MFD of the soil, while the freezing index (the cumulative effect of negative temperatures) is also a key factor in determining the MFD (Frauenfeld, 2004; Frauenfeld & Zhang, 2011). The sensitivity of MFD to the freezing index is the greatest in the warm



**Figure 13.** Comparison of observations and simulations of air temperature. The colorbar displays the spatial density value of scattered points, that is, the degree of density. The red solid line is the linear fitted line. The black dashed line is the 1:1 line.

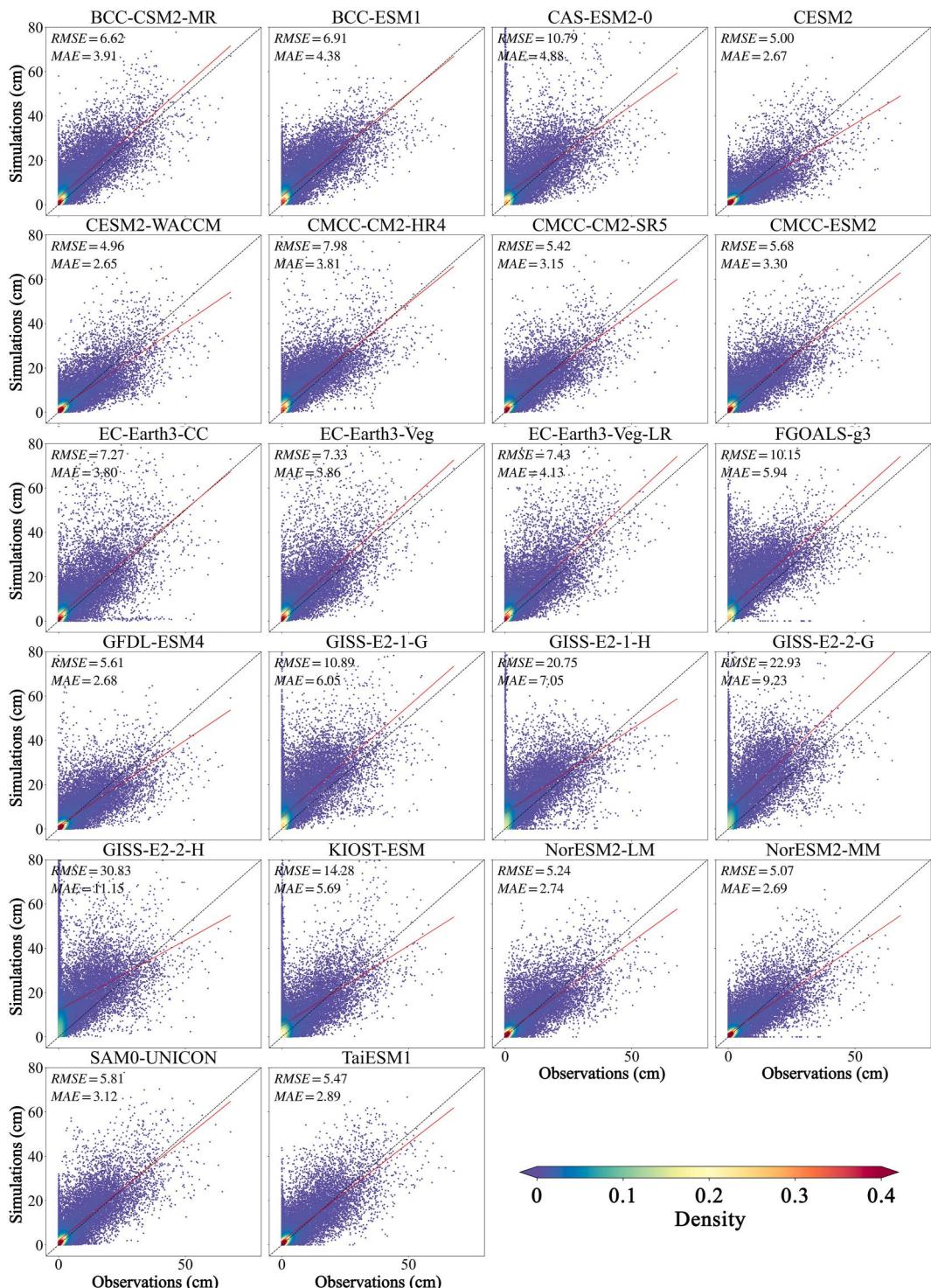
temperate climate zone with an average of  $1.16 \text{ cm}/10^\circ\text{C}\cdot\text{d}$ , followed by arid climate with an average of  $0.82 \text{ cm}/10^\circ\text{C}\cdot\text{d}$ . For regions with high mean annual temperatures, the magnitude of the cold season directly determines the MFD. The sensitivity of MFD to the freezing index is relatively small in the snow and polar climates,  $0.49 \text{ cm}/10^\circ\text{C}\cdot\text{d}$  and  $0.25 \text{ cm}/10^\circ\text{C}\cdot\text{d}$ , respectively. On the contrary, thawing index is the accumulation of positive



**Figure 14.** Comparison of observations and simulations of precipitation. The colorbar displays the spatial density value of scattered points, that is, the degree of density. The red solid line is the linear fitted line. The black dashed line is the 1:1 line.

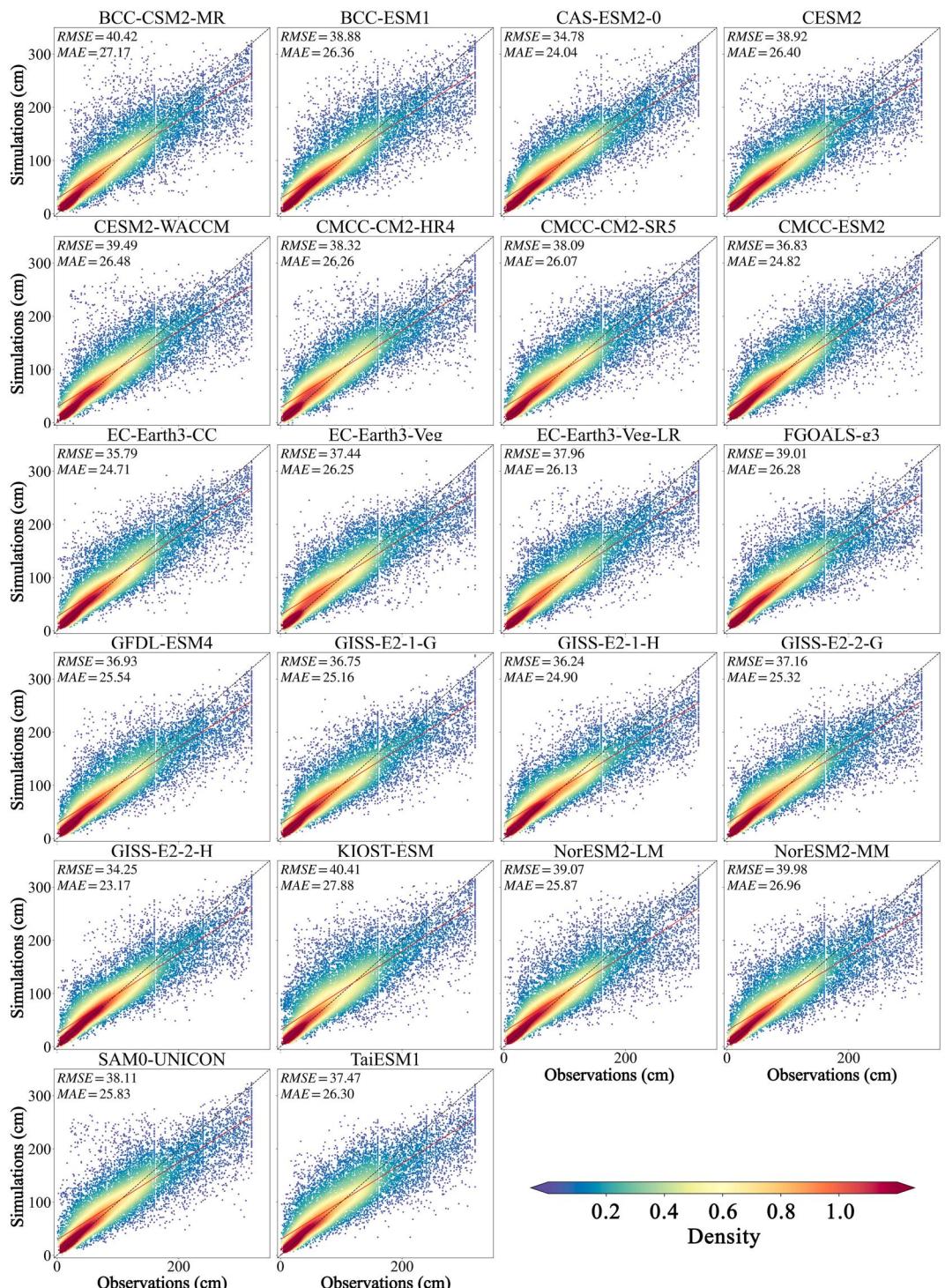
temperature, and affects the thermal state of frozen ground. The sensitivity of MFD to the thawing index is largest in the polar climate, with an average of  $-0.68 \text{ cm}/10^\circ\text{C}\cdot\text{d}$ , followed by the snow climate, with an average of  $-0.50 \text{ cm}/10^\circ\text{C}\cdot\text{d}$ .

The sensitivity of MFD to snow depth in the different climatic zones also reveals a pattern such that in the arid climate with low snowfall, snow replenishes soil moisture near the surface while also providing a cooling effect



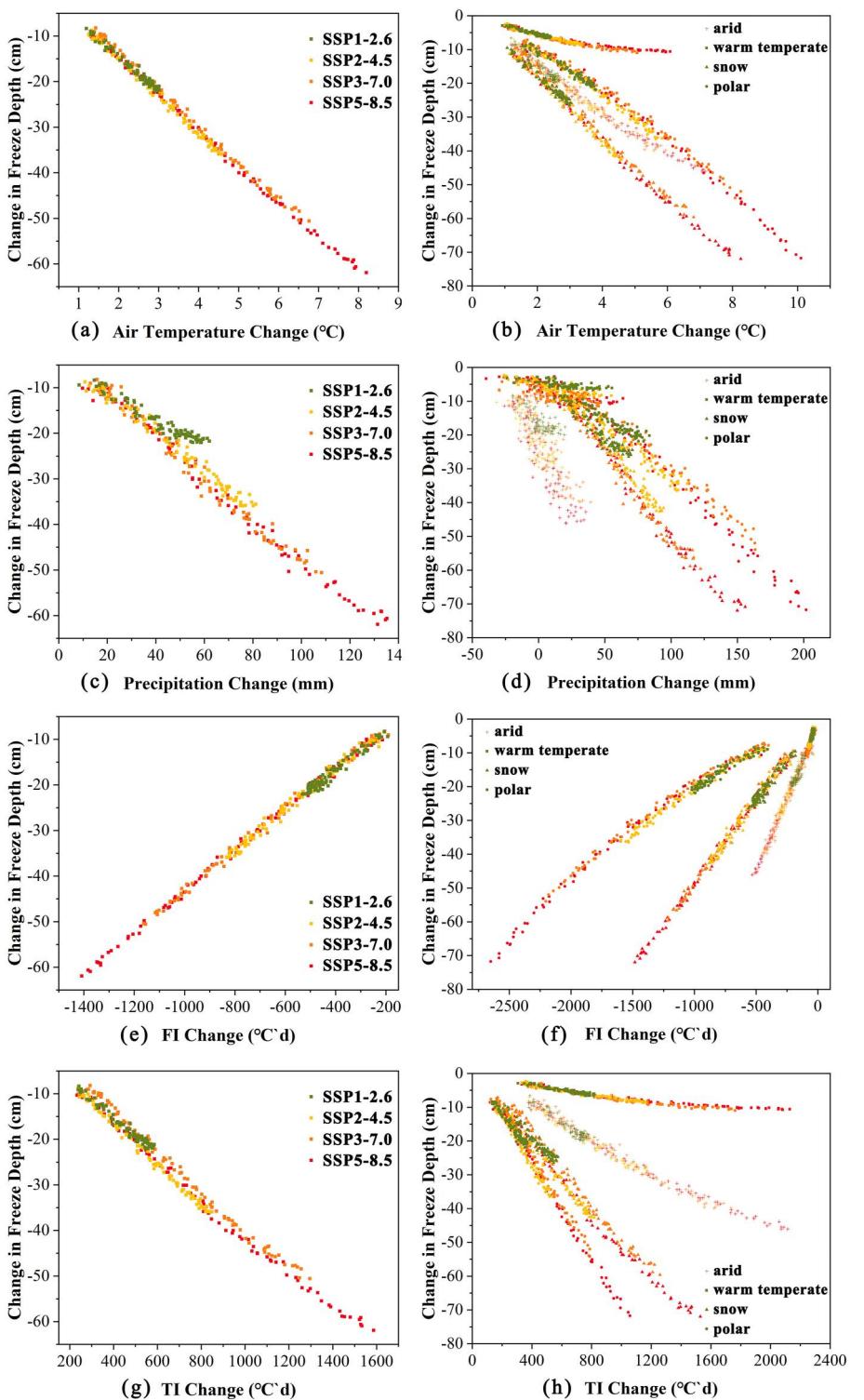
**Figure 15.** Comparison of observations and simulations of average snow depth. The colorbar displays the spatial density value of scattered points, that is, the degree of density. The red solid line is the linear fitted line. The black dashed line is the 1:1 line.

on the surface, thus deepening the freezing depth. MFD has the most significant sensitivity to snow depth with an average of 22 cm/cm, followed by the snow climate and warm temperate climate at 9.1 cm/cm and 7.6 cm/cm, respectively. When the snow depth reaches a certain depth, the thermal insulation of snow is optimal, preventing the ground from dissipating heat in winter and increasing the ground temperature while reducing the freezing



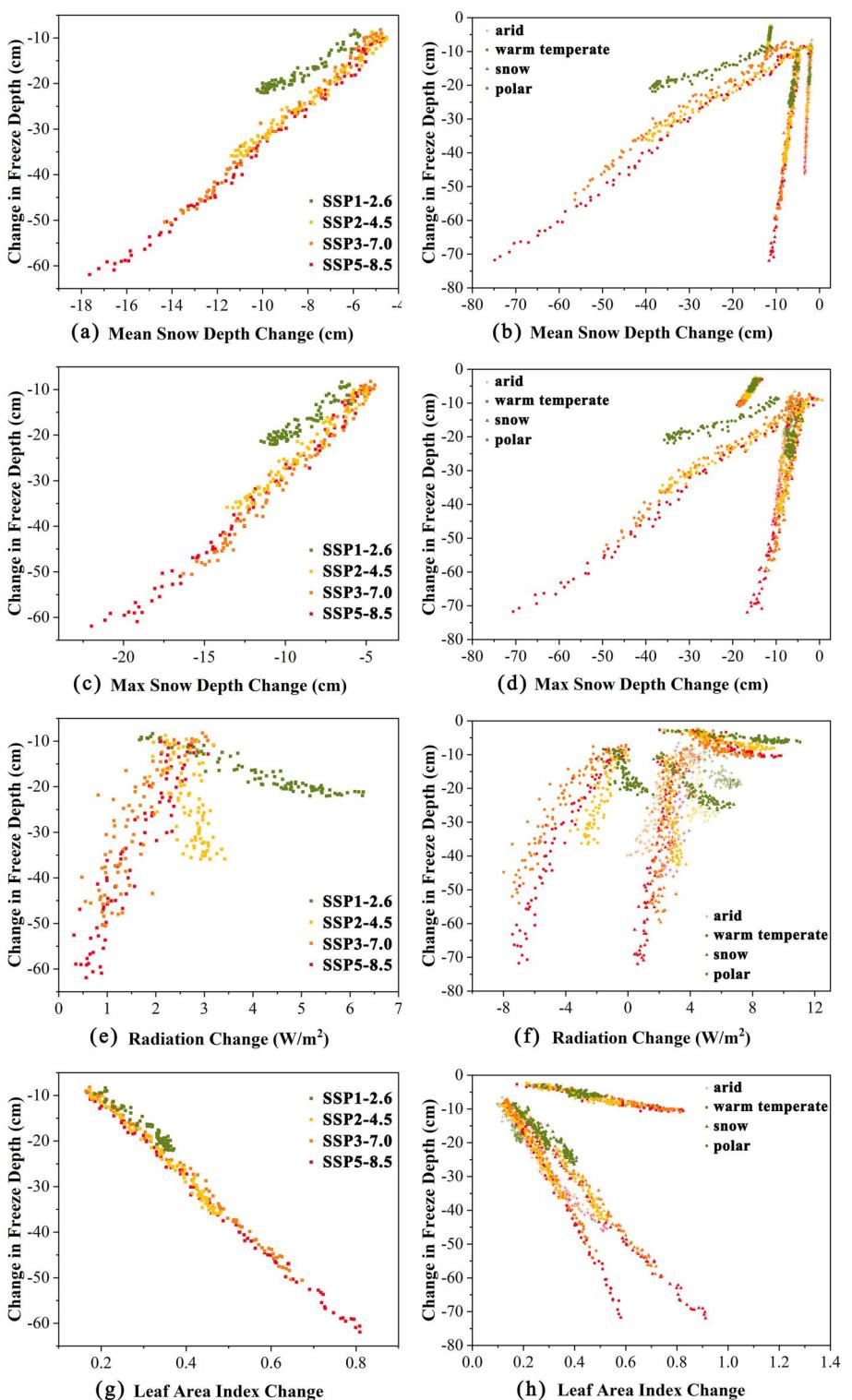
**Figure 16.** Comparison of observations and simulations of MFD. The colorbar indicates the density of the scattered points. The red solid line is the linear fit and the black dashed line is the 1:1 line.

depth (Frauenfeld & Zhang, 2011; Park et al., 2014, 2015; Zhang, 2005). In the polar climate zone with perennial snow, the thermal insulation of snow on the ground is more consistent and its sensitivity is the smallest, at 0.73 cm/cm. Similarly, the sensitivity of the MFD to the maximum snow depth is, in descending order, greatest in the arid climate (6.8 cm/cm), then the snow climate (5.4 cm/cm), warm temperate climate (1.7 cm/cm), and the



**Figure 17.** Sensitivity of MFD to changes in air temperature, precipitation, freezing index (FI), and thawing index (TI).

polar climate ( $0.77 \text{ cm/cm}$ ). The sensitivity of MFD to radiation decreases in the future period (Stanhill & Cohen, 2001), especially manifested in the arid climate, snow climate, and polar climate, where the MFD in these regions seems to be unaffected by radiation. In contrast, in the warm temperate climate zone, radiation increases result in MFD decreases, with a sensitivity of  $-1.1 \text{ cm/W} \cdot \text{m}^{-2}$  on average. An increase in the leaf area index



**Figure 18.** Sensitivity of MFD to change in mean snow depth, maximum snow depth, radiation, and leaf area index.

**Table 3**

Sensitivity of MFD to Change in Annual Mean Air Temperature (AT), Annual Cumulative Precipitation (PR), Annual Freezing Index (FI), and Annual Thawing Index (TI)

Climate Factors		AT	PR	FI	TI
Climate Zone		cm/°C	cm/mm	cm/10°C·d	cm/10°C·d
Arid	SSP1-2.6	-8.00	-0.27	0.76	-0.29
	SSP2-4.5	-7.52	-0.48	0.85	-0.27
	SSP3-7.0	-7.24	-0.60	0.83	-0.26
	SSP5-8.5	-6.47	-0.73	0.86	-0.22
Warm temperate	SSP1-2.6	-2.51	-0.04	0.93	-0.07
	SSP2-4.5	-2.31	-0.08	1.23	-0.07
	SSP3-7.0	-1.94	-0.11	1.21	-0.06
	SSP5-8.5	-1.57	-0.10	1.25	-0.04
Snow	SSP1-2.6	-8.36	-0.31	0.45	-0.43
	SSP2-4.5	-9.49	-0.46	0.49	-0.54
	SSP3-7.0	-9.44	-0.52	0.50	-0.53
	SSP5-8.5	-9.01	-0.48	0.49	-0.49
Polar	SSP1-2.6	-5.87	-0.21	0.22	-0.57
	SSP2-4.5	-6.71	-0.30	0.25	-0.72
	SSP3-7.0	-6.81	-0.32	0.25	-0.72
	SSP5-8.5	-7.40	-0.35	0.28	-0.71

reduces the MFD, and the sensitivity of the MFD to leaf area index is the greatest in arid climates, at  $-112 \text{ cm}/10^{-2}$ . An increase in the leaf area index implies an increase in vegetation cover, which acts as an insulator that is not conducive to soil freezing (Godínez-Alvarez et al., 2009; Piao et al., 2003). Sensitivity in the snow climate and polar climate is still high, at  $-84 \text{ cm}/10^{-2}$  in both zones. The sensitivity of MFD to leaf area index in the warm temperate climate is only  $-6.9 \text{ cm}/10^{-2}$ .

In summary, air temperature is one of the most direct and important factors affecting soil temperature (Liu et al., 2016; Streletskiy et al., 2015; Wang et al., 2015; Wu et al., 2016), it is also one of the most direct factors affecting the depth of soil freezing (Frauenfeld, 2004; Frauenfeld & Zhang, 2011). Changes in precipitation cause changes in soil moisture content, which in turn affects soil heat conduction, and because water has a high specific heat capacity, an increase in soil moisture content reduces soil heat conduction rate. The increase in soil water content therefore slows down the transfer of heat through the soil in order to inhibit warming. The freezing index and thawing index are cumulative values of negative and positive temperature respectively, and are influenced by the same mechanisms as air temperature. During the cold season, the presence of seasonal snow cover affects

**Table 4**

Sensitivity of MFD to Change in Annual Mean Snow Depth (ASD), Annual Maximum Snow Depth (MSD), Annual Mean Radiation (SR), and Average Annual Leaf Area Index (LAI)

Climate Factors		ASD	MSD	SR	LAI
Climate Zone		cm/cm	cm/cm	cm/W·m <sup>-2</sup>	cm/10 <sup>-2</sup>
Arid	SSP1-2.6	14.38	4.04	-2.88	-1.23
	SSP2-4.5	20.47	6.63	-4.79	-1.22
	SSP3-7.0	27.35	8.19	7.84	-1.04
	SSP5-8.5	24.92	8.34	14.28	-0.97
Warm temperate	SSP1-2.6	6.5	1.39	-0.42	-0.03
	SSP2-4.5	7.91	1.75	-0.95	-0.08
	SSP3-7.0	8.67	1.81	-1.64	-0.08
	SSP5-8.5	7.47	1.66	-1.40	-0.09
Snow	SSP1-2.6	7.34	4.01	-3.30	-0.73
	SSP2-4.5	9.10	5.77	-16.36	-0.92
	SSP3-7.0	10.62	6.43	19.03	-0.90
	SSP5-8.5	9.18	5.38	23.17	-0.83
Polar	SSP1-2.6	0.41	0.42	-4.32	-0.27
	SSP2-4.5	0.72	0.75	9.06	-0.75
	SSP3-7.0	0.87	0.94	6.01	-1.14
	SSP5-8.5	0.91	0.95	8.22	-1.21

surface temperatures, which in turn changes soil temperatures (Zhang, 2005). The effect of snow depth on soil temperature manifests itself in two effects: thermal insulation and cooling, depending on the thickness of the snow (Frauenfeld & Zhang, 2011; Park et al., 2014, 2015; Zhang, 2005). Higher latitudes and higher elevations with thicker snow are not conducive to soil freezing, due to the low thermal conductivity of the snow, which prevents the ground from dissipating heat. Lower latitudes and elevations with thinner snow are favorable for soil freezing, due to the high albedo of the snow, which reduces heat absorption from the ground. The amount of radiation determines the amount of heat that can be gained by the ground and thus affects the ground temperature. Soil freezing occurs in the cold season, when greater vegetation cover hinders the diffusion of heat from the surface and acts as an insulating effect, which is not conducive to soil freezing.

Spatial and temporal differences in climate determine the spatial and temporal differences in MFD variability. The occurrence of soil freezing and the magnitude of the MFD are the result of the combined effect of climate variables. The complex hydrothermal exchange between the ground and the atmosphere determines the hydrothermal conditions of the soil, which affects the MFD. Similarly, soil texture also influences the freeze depth of

the soil. Soil texture determines the thermophysical properties of the soil, including its composition, structure, density, porosity, thermal conductivity, etc., which all influence freezing and thawing processes. Topographic variables like slope and aspect also potentially influence MFD, which will be further explored in-depth in future studies. In addition, the autocorrelation of these environmental variables equally affects the accuracy of the model. Although not included here, to improve the accuracy of machine learning simulations, the autocorrelation between variables will be considered in subsequent studies to further improve the model. Finally, future work will assess the ability of our machine learning model to simulate permafrost changes in addition to those in SFG, as the same physical processes govern both environments.

## 5. Conclusions

This study develops a data-driven approach for modeling the MFD of SFG. Based on ERA5 reanalysis data and CMIP6 output, the XGBoost machine learning algorithm is used to simulate and predict the current and future spatial and temporal variability of MFD in the NH. In the context of global warming, these results can provide data to support the study of SFG changes and their impacts.

From the historical period to future projections based on SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, the regional average MFD was around 133 cm (ERA5) and 131 cm (CMIP6) during 1981–2010, and will range 112.08 cm (SSP1-2.6), 99.68 cm (SSP2-4.5), 88.88 cm (SSP3-7.0), and 81.04 cm (SSP5-8.5) in 2070–2099 depending on which emission scenario unfolds. Most of the regions shows substantial MFD decreases, most significant in the high-latitude and high-elevation regions, especially based on SSP5-8.5. During the 1981–2010 period, MFD decreased by an average of 13.16 cm (CMIP6) to 17.24 cm (ERA5), corresponding to 13.93%–19.10%. Ranging from the SSP1-2.6 to the SSP5-8.5 climate scenarios, the MFD in the NH will decrease by 11% (−12.02 cm), 23% (−14.55 cm), 35% (−16.54 cm), and 42% (−18.71 cm), on average, between 2015 and 2099 relative to the base period. The most significant change in MFD corresponds to the snow climate zone, with an average reduction of 43.23 cm.

The sensitivity of MFD to precipitation, snow, and leaf area index is greatest in dry climates, sensitivity to temperature is greatest in snowy climates, and to freezing and thawing index in warm temperate and polar climates, respectively. As shown in Section 4.2, the direct influence on the MFD is due to changes in soil heat fluxes, with warming, increased moisture, warmer cold seasons, warmer warm seasons, shallower snow depths, and increased vegetation cover all leading to a reduction in the MFD. The analysis of soil heat fluxes and the SFG-climate-vegetation interactions need to be further investigated.

## Data Availability Statement

All CMIP6 model outputs are openly available from phase 6 of the Coupled Model Intercomparison Project (CMIP6, 1995). The ERA5-Land Data are obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF, 2020). The Köppen-Geiger climate classification published by Rubel and Kottek (2010). The soil texture data are openly available from Sun Yat-sen University's Land-Atmosphere Interaction Research Group (Shangguan et al., 2014).

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant 42161160328, 42171120), National Cryosphere Desert Data Center (E01Z790201), and the Fundamental Research Funds for the Central Universities (Izujbky-2023-01, Izujbky-2021-ct13). We thank the anonymous reviewers whose thoughtful comments improved our manuscript.

## References

- Ala-Aho, P., Autio, A., Bhattacharjee, J., Isokangas, E., Kujala, K., Marttila, H., et al. (2021). What conditions favor the influence of seasonally frozen ground on hydrological partitioning? A systematic review. *Environmental Research Letters*, 16(4), 043008. <https://doi.org/10.1088/1748-9326/abe82c>
- Baiz, S., Tighe, S. L., Haas, C. T., Mills, B., & Perchanok, M. (2008). Development of frost and thaw depth predictors for decision making about variable load restrictions. *Transportation Research Record*, 2053(1), 1–8. <https://doi.org/10.3141/2053-01>
- Chang, X.-l., Jin Huijun, 金., Wang Yongping, 王., Zhang Yanlin, 张., Zhou Gangyi, 周., Chu Fuqiang, 车., & Zhao Yumei, 赵. (2012). Influences of vegetation on permafrost: A review. *Acta Ecologica Sinica*, 32(24), 7981–7990. <https://doi.org/10.5846/stxb201202120181>
- Chen, C., Peng, X., Frauenfeld, O. W., Zhao, Y., Yang, G., Tian, W., et al. (2022). Comprehensive assessment of seasonally frozen ground changes in the Northern Hemisphere based on observations. *Journal of Geophysical Research: Atmospheres*, 127(20), e2022JD037306. <https://doi.org/10.1029/2022jd037306>
- Chen, T., & Guestrin, C. (2016). XGBoost. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).
- Cheng, G., & Wu, T. (2007). Responses of permafrost to climate change and their environmental significance, Qinghai-Tibet Plateau. *Journal of Geophysical Research*, 112(F2). <https://doi.org/10.1029/2006jf000631>
- Coupled Model Intercomparison Project. (1995). phase 6 of the coupled model Intercomparison project [Dataset]. World Climate Research Programme Retrieved from <https://esgf-index1.ceda.ac.uk/search/cmip6-ceda/>

- European Centre for Medium-Range Weather Forecasts fifth-generation monthly reanalysis dataset. (2020). ERA5-Land monthly averaged data [Dataset]. European Centre for Medium-Range Weather Forecasts. Retrieved from <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=form>
- Frauenfeld, O. W., & Zhang, T. (2011). An observational 71-year history of seasonally frozen ground changes in the Eurasian high latitudes. *Environmental Research Letters*, 6(4), 044024. <https://doi.org/10.1088/1748-9326/6/4/044024>
- Frauenfeld, O. W., Zhang, T., Barry, R. G., & Gilichinsky, D. (2004). Interdecadal changes in seasonal freeze and thaw depths in Russia. *Journal of Geophysical Research*, 109(D5), D05101. <https://doi.org/10.1029/2003JD004245>
- Frauenfeld, O. W., Zhang, T., & McCreight, J. L. (2007). Northern Hemisphere freezing/thawing index variations over the twentieth century. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 27(1), 47–63. <https://doi.org/10.1002/joc.1372>
- Frauenfeld, O. W., Zhang, T., & Serreze, M. C. (2005). Climate change and variability using European Centre for Medium-Range Weather Forecasts reanalysis (ERA-40) temperatures on the Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, 110(D2), D02101. <https://doi.org/10.1029/2004JD005230>
- Godínez-Alvarez, H., Herrick, J., Mattocks, M., Toledo, D., & Van Zee, J. (2009). Comparison of three vegetation monitoring methods: Their relative utility for ecological assessment and monitoring. *Ecological Indicators*, 9(5), 1001–1008. <https://doi.org/10.1016/j.ecolind.2008.11.011>
- Guo, D., & Wang, H. (2016). CMIP5 permafrost degradation projection: A comparison among different regions. *Journal of Geophysical Research: Atmospheres*, 121(9), 4499–4517. <https://doi.org/10.1002/2015jd024108>
- Harlan, R. L. (1978). Ground thermal regime. In *Geotechnical Engineering for Cold Regions*.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Jin, H., He, R., Cheng, G., Wu, Q., Wang, S., Lü, L., & Chang, X. (2009). Changes in frozen ground in the Source Area of the Yellow River on the Qinghai-Tibet Plateau, China, and their eco-environmental impacts. *Environmental Research Letters*, 4(4), 045206. <https://doi.org/10.1088/1748-9326/4/4/045206>
- Kudryavtsev, V., Garagulya, L., & Melamed, V. (1977). *Fundamentals of frost forecasting in geological engineering investigations (Osnovy Merzlotnogo Prognoza pri Inzhenerno-Geologicheskikh Issledovaniyakh)* Rep. Cold Regions Research And Engineering Lab Hanover NH.
- Liu, Y., Wang, L., Liu, B., & Henderson, M. (2016). Observed changes in shallow soil temperatures in Northeast China, 1960–2007. *Climate Research*, 67(1), 31–42. <https://doi.org/10.3354/cr01351>
- Luo, S., & Lü, S. (2009). Soil thermal conductivity parameterization establishment and application in numerical model of central Tibetan Plateau. *Chinese Journal of Geophysics*, 52(4), 919–928.
- Luo, S., Lü, S., & Zhang, Y. (2009). Development and validation of the frozen soil parameterization scheme in Common Land Model. *Cold Regions Science and Technology*, 55(1), 130–140. <https://doi.org/10.1016/j.coldregions.2008.07.009>
- Luo, S., Wang, J., Pomeroy, J. W., & Lyu, S. (2020). Freeze–thaw changes of seasonally frozen ground on the Tibetan Plateau from 1960 to 2014. *Journal of Climate*, 33(21), 9427–9446. <https://doi.org/10.1175/jcli-d-19-0923.1>
- Maximov, T., Ohta, T., & Dolman, A. J. (2008). Water and energy exchange in East Siberian forest: A synthesis. *Agricultural and Forest Meteorology*, 148(12), 2013–2018. <https://doi.org/10.1016/j.agrformet.2008.10.004>
- Nelson, F. E. (1986). Permafrost distribution in central Canada: Applications of a climate-based predictive model. *Annals of the Association of American Geographers*, 76(4), 550–569. <https://doi.org/10.1111/j.1467-8306.1986.tb00136.x>
- Nelson, F. E. (1987). A frost index number for spatial prediction of ground frost zones. *Arctic and Alpine Research*, 19, 279–288.
- Nelson, F. E., & Outealt, S. I. (1987). A computational method for prediction and regionalization of permafrost. *Arctic and Alpine Research*, 19(3), 279–288. <https://doi.org/10.2307/1551363>
- O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., et al. (2014). A new scenario framework for climate change research: The concept of shared socioeconomic pathways. *Climatic Change*, 122(3), 387–400. <https://doi.org/10.1007/s10584-013-0905-2>
- Park, H., Fedorov, A. N., Zheleznyak, M. N., Konstantinov, P. Y., & Walsh, J. E. (2015). Effect of snow cover on pan-Arctic permafrost thermal regimes. *Climate Dynamics*, 44(9–10), 2873–2895. <https://doi.org/10.1007/s00382-014-2356-5>
- Park, H., Sherstiukov, A. B., Fedorov, A. N., Polyakov, I. V., & Walsh, J. E. (2014). An observation-based assessment of the influences of air temperature and snow depth on soil temperature in Russia. *Environmental Research Letters*, 9(6), 064026. <https://doi.org/10.1088/1748-9326/9/6/064026>
- Peng, X. (2017). *Spatial-temporal variations of seasonally frozen ground and its response to climate change in the Northern Hemisphere* (Doctoral thesis). Lanzhou University.
- Peng, X., Zhang, T., Frauenfeld, O. W., Du, R., Wei, Q., & Liang, B. (2020). Soil freeze depth variability across Eurasia during 1850–2100. *Climatic Change*, 158(3), 531–549. <https://doi.org/10.1007/s10584-019-02586-4>
- Peng, X., Zhang, T., Frauenfeld, O. W., Wang, K., Cao, B., Zhong, X., et al. (2017). Response of seasonal soil freeze depth to climate change across China. *The Cryosphere*, 11(3), 1059–1073. <https://doi.org/10.5194/tc-11-1059-2017>
- Piao, S., Fang, J., Zhou, L., Guo, Q., Henderson, M., Ji, W., et al. (2003). Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999. *Journal of Geophysical Research*, 108(D14). <https://doi.org/10.1029/2002jd002848>
- Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen–Geiger climate classification [Dataset]. *Meteorologische Zeitschrift*, 19(2), 135–141. <https://doi.org/10.1127/0941-2948/2010/0430>
- Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for Earth system modeling [Dataset]. *Journal of Advances in Modeling Earth Systems*, 6(1), 249–263. <https://doi.org/10.1002/2013ms000293>
- Shiklomanov, N. I. (2012). Non-climatic factors and long-term, continental-scale changes in seasonally frozen ground. *Environmental Research Letters*, 7(1), 011003. <https://doi.org/10.1088/1748-9326/7/1/011003>
- Stanhill, G., & Cohen, S. (2001). Global dimming: A review of the evidence for a widespread and significant reduction in global radiation with discussion of its probable causes and possible agricultural consequences. *Agricultural and Forest Meteorology*, 107(4), 255–278. [https://doi.org/10.1016/s0168-1923\(00\)00241-0](https://doi.org/10.1016/s0168-1923(00)00241-0)
- Strellets, D. A., Sherstiukov, A. B., Frauenfeld, O. W., & Nelson, F. E. (2015). Changes in the 1963–2013 shallow ground thermal regime in Russian permafrost regions. *Environmental Research Letters*, 10(12), 125005. <https://doi.org/10.1088/1748-9326/10/12/125005>
- Sugimoto, A., Yanagisawa, N., Naito, D., Fujita, N., & Maximov, T. C. (2002). Importance of permafrost as a source of water for plants in east Siberian taiga. *Ecological Research*, 17(4), 493–503. <https://doi.org/10.1046/j.1440-1703.2002.00506.x>
- Swann, A. L., Fung, I. Y., Levis, S., Bonan, G. B., & Doney, S. C. (2010). Changes in Arctic vegetation amplify high-latitude warming through the greenhouse effect. *Proceedings of the National Academy of Sciences*, 107(4), 1295–1300. <https://doi.org/10.1073/pnas.0913846107>
- Walvoord, M. A., & Kurylyk, B. L. (2016). Hydrologic impacts of thawing permafrost—A review. *Vadose Zone Journal*, 15(6), 1–20. <https://doi.org/10.2136/vzj2016.01.0010>

- Wang, K., Zhang, T., & Zhong, X. (2015). Changes in the timing and duration of the near-surface soil freeze/thaw status from 1956 to 2006 across China. *The Cryosphere*, 9(3), 1321–1331. <https://doi.org/10.5194/tc-9-1321-2015>
- Wang, T., Yang, D., Fang, B., Yang, W., Qin, Y., & Wang, Y. (2019). Data-driven mapping of the spatial distribution and potential changes of frozen ground over the Tibetan Plateau. *The Science of the Total Environment*, 649, 515–525. <https://doi.org/10.1016/j.scitotenv.2018.08.369>
- Westermann, S., Langer, M., Boike, J., Heikenfeld, M., Peter, M., Etzelmüller, B., & Krinner, G. (2016). Simulating the thermal regime and thaw processes of ice-rich permafrost ground with the land-surface model CryoGrid 3. *Geoscientific Model Development*, 9(2), 523–546. <https://doi.org/10.5194/gmd-9-523-2016>
- Wu, Q., Zhang, Z., Gao, S., & Ma, W. (2016). Thermal impacts of engineering activities and vegetation layer on permafrost in different alpine ecosystems of the Qinghai–Tibet Plateau, China. *The Cryosphere*, 10(4), 1695–1706. <https://doi.org/10.5194/tc-10-1695-2016>
- Yang, M., Nelson, F. E., Shiklomanov, N. I., Guo, D., & Wan, G. (2010). Permafrost degradation and its environmental effects on the Tibetan Plateau: A review of recent research. *Earth-Science Reviews*, 103(1–2), 31–44. <https://doi.org/10.1016/j.earscirev.2010.07.002>
- Yang, M., Wang, X., Pang, G., Wan, G., & Liu, Z. (2019). The Tibetan Plateau cryosphere: Observations and model simulations for current status and recent changes. *Earth-Science Reviews*, 190, 353–369. <https://doi.org/10.1016/j.earscirev.2018.12.018>
- Zhang, T., Baker, T. H. W., Cheng, G.-D., & Wu, Q. (2008). The Qinghai–Tibet Railroad: A milestone project and its environmental impact. *Cold Regions Science and Technology*, 53(3), 229–240. <https://doi.org/10.1016/j.coldregions.2008.06.003>
- Zhang, T., Barry, R., Knowles, K., Ling, F., & Armstrong, R. L. (2003). Distribution of seasonally and perennially frozen ground in the Northern Hemisphere.
- Zhang, T., Frauenfeld, O. W., Serreze, M. C., Etringer, A., Oelke, C., McCreight, J., et al. (2005). Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin. *Journal of Geophysical Research*, 110(D16). <https://doi.org/10.1029/2004jd005642>
- Zhang, Y., Cheng, G., Li, X., Jin, H., Yang, D., Flerchinger, G. N., et al. (2017). Influences of frozen ground and climate change on hydrological processes in an alpine watershed: A case study in the upstream area of the Hei'he River, Northwest China. *Permafrost and Periglacial Processes*, 28(2), 420–432. <https://doi.org/10.1002/ppp.1928>