Edinburgh Research Explorer

Widespread Persistent Extreme Cold Events Over SouthEast China: Mechanisms, Trends, and Attribution

OPEN ACCESS

JGR Atmospheres

Key Points:
- Change in the frequency of cold waves over South-East China during the past decades is not detectable yet
- CMIP5 and CMIP6 models can reproduce the observed dynamics of these events
- Cold waves are expected to decrease due to the dominant impact of greenhouse gases in the next decades

Correspondence to:

N. Freychet,
nicolas.freychet@ed.ac.uk

Widespread Persistent Extreme Cold Events Over South-East China: Mechanisms, Trends, and Attribution

N. Freychet[1] [iD], **S. F. B. Tett[1]** [iD], **A. A. Abatan[1,2]**, **A. Schurer[1]** [iD], and **Z. Feng[1]** [iD]

[1]School of Geosciences, University of Edinburgh, Edinburgh, UK, [2]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

Abstract South-East China usually experiences mild winter temperatures but can also be impacted by large-scale and long-lasting cold events. Here, we investigate how this category of events has changed during the past 6 decades and why, using both observation and climate model ensembles. We first show that models can reproduce the observed largescale dynamics associated with cold events. A dynamical advection of cold air is mostly responsible not only to trigger cold surges but also we found strong feedback from shortwave radiation, decreasing due to enhanced cloud cover. Recent trends are more difficult to evaluate. Models show a large internal variability and intermodel spread, while observation also indicates large uncertainties in its trend due to high internal variability. Our analysis indicates that even if cold events are likely to have been reduced due to greenhouse gases, trends cannot be attributed with high confidence to any anthropogenic signal alone. However, in the next few decades, the frequency of long-lasting cold events are expected to quickly reduce due to the emerging greenhouse gases signal and possibly weakening aerosol effect over South-East China. We also found that most of the trends are due to the change in the mean temperature. Thus, indirectly, there is no clear evidence that the change in cold event frequency is due to a change in the circulation.

1. Introduction

Winter widespread persistent extreme cold events (WiPECEs) have important socioeconomic impacts in China (Peng & Bueh, 2011). In the past few decades, these events have caused great destruction to health, agriculture, and social amenities (Peng & Bueh, 2011; Zhang & Qian, 2011; Qian et al., 2017). For example, the 2008 and 2016 cold surges in eastern China caused heavy freezing rain and icy conditions that led to power outage in many regions as a result of damaged electricity transmission lines (Sun & Zhao, 2010). The impact of extreme cold events on winter wheat yields is also significant (Powell & Reinhard, 2016). Health impacts of cold temperatures are well known (Chen et al., 2017; Luo et al., 2018; Ma et al., 2013). Wu et al. (2013) showed, for four subtropical cities in China, that longer exposure to cold events had larger impact than similar duration hot events. A case study for Shanghai from Ma et al. (2011) indicated that during a cold spell hospital admission related to cardiovascular and respiratory problems increased by 32%–38%, versus an increase of 2%–8% during a similar duration heat wave. Chen et al. (2017) found an association between extreme low temperature and hemorrhagic stroke admission. Cold effects are also linked to a significant increase of patients with fracture (Du et al., 2013).

A large part of the research community has focused much attention on the variability and spatial distributions of trends in extreme hot temperature events in China as a result on the ongoing anthropogenic warming, while less effort have been devoted to widespread persistent extreme cold temperature events (Qian et al., 2017; Zhou et al., 2009). Most of these studies and the references therein focused on case studies without reference to the climatology of such events. Although Peng and Bueh (2011) examined the extensive and persistent extreme cold events for the period 1951–2009, these authors only used a station observational dataset and did not examine the large-scale circulation features of their classified extreme cold events. Recently, Liao et al. (2020) highlighted the different dynamical configuration for the events of 2008 and 2016. Zuo et al. (2015) also indicated that cold surges were more likely to be persistent during February. The recently released European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis version 5 (ERA5, Hersbach, 2018) data set opens a further opportunity to examine WiPECE over China. In addition, the sixth phase of the Coupled Model Intercomparison Project (CMIP6) also provides updated climate
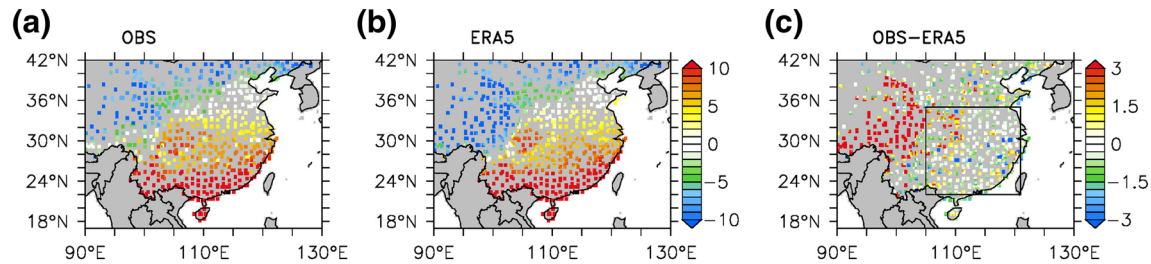
Figure 1. T2M winter means (in °C) during the 1979–2017 period for (a) OBS, (b) ERA5, and (c) difference OBS-ERA5. Black box in (c) indicates the South-East China domain used for this study.

model outputs to evaluate past and future trends of WiPECE, to be compared with the previous versions of the models (CMIP5).

South-East China (defined here as 22–35 N/105–121 E, black box in Figure 1) is heavily populated (Li, Tian, et al., 2018) and highly vulnerable to impacts of climate extremes (e.g., Freychet et al., 2020; Li, Tian, et al., 2018; Sparrow et al., 2018). In this region, winters are usually characterized by low-to-mild temperatures with average daily temperature of 7.5°C for 1979–2017. Populations in this region are poorly acclimatized to cold weather and so not well prepared to endure severe cold surges. The Elderly population is especially vulnerable to long persistent cold surges (Chen et al., 2019). Adaptation measures at household and community levels to such cold events are poor with, for example, a lack of efficient heater and poor insulation, especially in old buildings, which may amplify the vulnerability of the population (Wang et al., 2016). In addition, the healthcare system is usually less prepared for the surge of demand for emergency services during the cold spells with, for example, a high nonresponse rate in the 2008 cold spell in South China (Xiong et al 2010).

The objective of this study is to understand the mechanisms driving WiPECE, how these events have changed and are expected to change over the 21st century, combining observation, reanalysis, and model outputs. The reliability of models to simulate correctly the dynamics of cold events and their trend is also investigated. This study is organized such that methodology and data are presented in Section 2. Results are exposed in Section 3, and concluding remarks provided in Section 4.

2. Materials and Methods

2.1. Datasets

For this study, two-model ensembles from the fifth and sixth Coupled Model Intercomparison Project are considered (CMIP5, Taylor et al., 2012; and CMIP6, Eyring et al., 2016) and evaluated using two reference datasets: The reanalysis ERA5 and an observational station network (OBS). The large ensemble from the Community Earth System Model (CESM, Hurrell et al., 2013) is used to compute the model internal variability for attribution methods. Each data set is described below.

2.1.1. Observational Network (OBS)

A network of 756 stations across China (Li & Yan, 2009) from the China Meteorological Administration is used. It provides daily mean surface temperature at 2 m (T2M) from 1950 to 2017. Reliability of records between 1950 and 1960 is questionable as this corresponds to a transition period when the network increased from less than 100 stations to more than 700 (see Figure 2 in Li et al., 2009). Thus, only T2M for the 1960–2017 period is used in this study. Measurements have been carefully homogenized to remove biases due to instrument changes or station relocation. OBS is gridded to the same grid as ERA5 by averaging all stations available at each grid cell (and masking points without stations, Figure 1a) so both reference datasets can be compared.

2.1.1.1. ERA5 Reanalysis

The latest version of the European Centre for Medium-Range Weather Forecasts, ERA5 (Hersbach, 2018), provides hourly output at 0.25° resolution, currently during the satellite period (1979 to present). Based

on these outputs, daily mean surface temperature is computed and then spatially masked where gridded observation is not available to match OBS grid points (Figure 1b). OBS and ERA5 have similar winter temperatures patterns and small differences, especially over South-East China (Figure 1c). ERA5 dynamical variables are also used to analyze the circulation during cold events.

2.1.1.2. CMIP5/CMIP6 Model Ensembles

Ensembles of CMIP5 and CMIP6 are selected, based on the availability of daily T2M outputs during the historical period using all forcings (ALL). This leads to a selection of 27 CMIP5 and 10 CMIP6 individual models, with some having several members (Table 1).

Models outputs are first considered for the 1960–2016 historical period to compare with the same period from OBS (and ERA5 during the 1979–2016 period). As the historical simulations stop in 2005 (2014) for CMIP5 (CMIP6), it is extended using results from the RCP4.5 (SSP245) projection scenario. WiPECE (see below) is computed for each individual model and members. When computing the multimodel ensemble mean, models with several members are first averaged so each model has the same weight in the multimodel mean. WiPECE long-term trend projections (to 2,100) are analyzed using RCP4.5 (SSP245) for CMIP5 (CMIP6). These are considered as medium-emission scenarios, thus results should be in the range of what one can reasonably expect from the current policy decisions to reach Paris agreement +2 °C target.

Similar to ERA5, dynamical variables are also extracted from the models to analyze their ability to reproduce the right circulation patterns associated with WiPECE.

For the attribution part of this study, two other scenarios are used: Natural forcing only where the anthropogenic impact is removed (NAT, with 13 and 8 individual models for CMIP5 and CMIP6, respectively) and Greenhouse Gases only where only the anthropogenic historical changes in GHG are considered (GHG, with 11 and 8 individual models for CMIP5 and CMIP6, respectively). Table 1 summarizes the models used for each scenario.

2.1.2. CESM Large Ensemble

Simulations from the Community Earth System Model (Hurrell et al., 2013) are used to quantify the internal variability of the models. T2M is extracted from 35 members from the historical (all forcings) large ensemble (CESM-LE; Kay et al., 2015). WiPECE is computed for each individual member for the 1960–2005 period and the covariance of this multimember ensemble is used to quantify internal variability.

2.2. WiPECE Index Computation

There are several definitions of extreme cold events in China (e.g. Peng & Bueh, 2011; Qian et al., 2017; Zhang & Qian, 2011). This study adopts the following definition. The 9-days window running mean (centered on a calendar day) of the 2-m temperature (T2M) is first computed (Rx9D_T2M). Second, the daily winter 10th percentile threshold index (RX10P_T2M) is computed from Rx9D_T2M at each grid point and for each calendar day, using the 1981–2010 baseline period. Finally, WiPECE is defined when Rx9D_T2M is lower than the daily RX10P_T2M for at least 50% of land points in South-East China domain. The 50% threshold criterion ensures that large-scale events are accounted for. With these combined criteria, both temporal and spatial persistences are considered, thus focusing on the most damaging events at regional scale.

One would expect about 10% of winter days to be below the 10% threshold, so if all persistent cold extremes in southern China were widespread one would expect 342 days using our criteria (10% of winter days). Of the 3,420 winter days for the 1979–2017 period, there are 273 days that satisfy the selection procedure using ERA5 reanalysis mean temperature. This accounts for about 80% of the events, suggesting the majority of cold extremes in south China are widespread events. Applying the same identification procedure to the OBS, we identify 267 (78%) WiPECE days, which is very close to the ERA5 value and confirms the good

Table 1
Individual Model Names for CMIP5 and CMIP6 Ensemble Along With the Number of Members Available for Each Individual Forcing (All Historical Forcings, ALL, Natural Forcing Only, NAT, Greenhouse Forcing Only, GHG, Other Anthropogenic Forcing, OA, and All Anthropogenic Forcing, ANT)

CMIP5	ALL	NAT	GHG	OA	ANT
BCC-CSM1-1	1	1	1	1	1
BCC-CSM1-1-M	1	-	-	-	-
BNU-ESM	1	1	1	1	1
CMCC-CM	1	-	-	-	-
CMCC-CMS	1	-	-	-	-
CNRM-CM5	1	6	6	1	1
ACCESS1-0	1	-	-	-	-
ACCESS1-3	1	-	-	-	-
CSIRO-MK3-6-0	10	5	5	5	5
EC-EARTH	1	-	-	-	-
INMCM4	1	-	-	-	-
IPSL-CM5A-LR	4	3	3	3	3
IPSL-CM5A-MR	1	3		1	
IPSL-CM5B-LR	1	-	-	-	-
HadGEM2-CC	1	-	-	-	-
HadGEM2-ES	1	4	4	1	1
MPI-ESM-LR	3	-	-	-	-
MPI-ESM-MR	3	-	-	-	-
GISS-E2-R	2	-	-	-	-
CCSM4	3	3	2	2	2
NorEMS1-M	1	1	-	-	-
GFDL-CM3	3	3	3	3	3
GFDL-ESM2G	1	-	-	-	-
GFDL-EMS2M	1	1	1	1	1
CESM1-BGC	1	-	-	-	-
CESM1-CAM5	1	1	-	1	-
Total	47	32	26	20	18

CMIP6	ALL	NAT	GHG	OA	ANT
CanESM5	10	10	10	10	10
CNRM-CM6-1	6	3	3	3	3
CNRM-ESM2-1	3	-	-	-	-
EC-EARTH3-VEG	1	-	-	-	-
IPSL-CM6A-LR	1	6	6	1	1
HadGEM3-GC31-LL	1	4	4	1	1
UKESM1-0-LL	5	-	-	-	-
MRI-ESM2-0	1	3	3	1	1
CESM2	1	3	3	1	1
CESM2-WACCM	1	-	-	-	-
Total	30	29	29	17	17

The total indicates the total of individual simulation for each category.

agreement of these two reference datasets. Models indicate a mean of 235 cold days during the same period (ranging from 184 days to 277 days for the lowest and highest numbers, respectively). This is slightly lower than in OBS and ERA5, indicating that models produce less clustered cold days.

2.3. Detection-Attribution Method

To analyze the role of anthropogenic activity on the recent trend in WiPECE (1960–2005) the detection–attribution (DA) method of Ribes et al. (2017) is used. Software is available at https://github.com/rafaelcabreu/attribution and is explained in the study by De Abreu et al. (2019). This method combines different sources of uncertainties (such as internal variability, model uncertainty, and observational uncertainty) to attribute the role of different forcing to explain the observed changes. Individual forcing experiments considered in this study are GHG, NAT, and ALL. Two other cases are derived from these three forcings: Other anthropogenic forcing (OA) corresponding to ALL-NAT-GHG (mainly anthropogenic aerosols); and anthropogenic forcing (ANT) corresponding to ALL-NAT. Finally, 45 years from preindustrial forcing simulations (piControl) are used to estimate the internal variability of CMIP5 and CMIP6 trends.

For this method, WiPECE trends are based on 5-year means for the 1960–2005 period (as most of the NAT simulations stopped in 2005 for CMIP5).

To estimate observation uncertainties, here, we simply use the mean difference between ERA5 and OBS during the 1979–2005 period and consider this error constant (i.e., we assume observation error to be time-independent).

2.4. Trends in OBS and Models

2.4.1. Estimating the Trends

Least-square linear regression is often used to analyze how a signal is changing with time but the high interannual variability of WiPECE leads to large uncertainties when using this method, especially because it can be more sensitive to extrema of the time-series as WiPECE frequency is non-Gaussian. Here, another method is used, based on Theil-Sen (1968). The mean estimate of the trend with this method is expected to be less sensitive to outliers of the time-series and the 5%–95% confidence interval also provides an accurate estimation of uncertainties. For model ensemble mean trends in Figures 6 and 9, the confidence interval is computed by bootstrapping the ensemble (thus providing an estimate of the uncertainty related to the sampling). Finally, uncertainties in trends for the best estimates from the DA method (Figure 8) are computed by generating a 1,000 multivariate normal random time-series based on the best estimates and covariance matrix from the DA.

2.4.2. Removing the Effect of Mean Temperature Changes

To investigate the role of mean temperature changes on the trend of WiPECE, and eventually to isolate the role of dynamical change, an approach similar to Freychet et al. (2017) is used as follows. First, at each grid point, the winter mean T2M is removed from daily temperatures for each individual year. Then the 10th percentile is computed based on these anomalies and all the processes described in the WiPECE index computation part are repeated using this detrended temperature. The result indicates how WiPECEs are changing or not if the yearly T2M was constant. This is an indirect way to detect a potential change in circulation, or temperature variability that could lead to changes in WiPECE.

2.5. Taylor Diagram Uncertainties

To evaluate model performance in reproducing the spatial pattern associated with WiPECE, Taylor diagrams are used (Taylor, 2001) which allow a quick comparison of the spatial correlation and standard deviation of the signal between models and the reference ERA5. To consider the model uncertainties due to their internal variability, we used an autocorrelation test as followed. An individual model (CMIP6 CanESM5) with 10 members (only for surface temperature and sea level pressure) is selected. One of
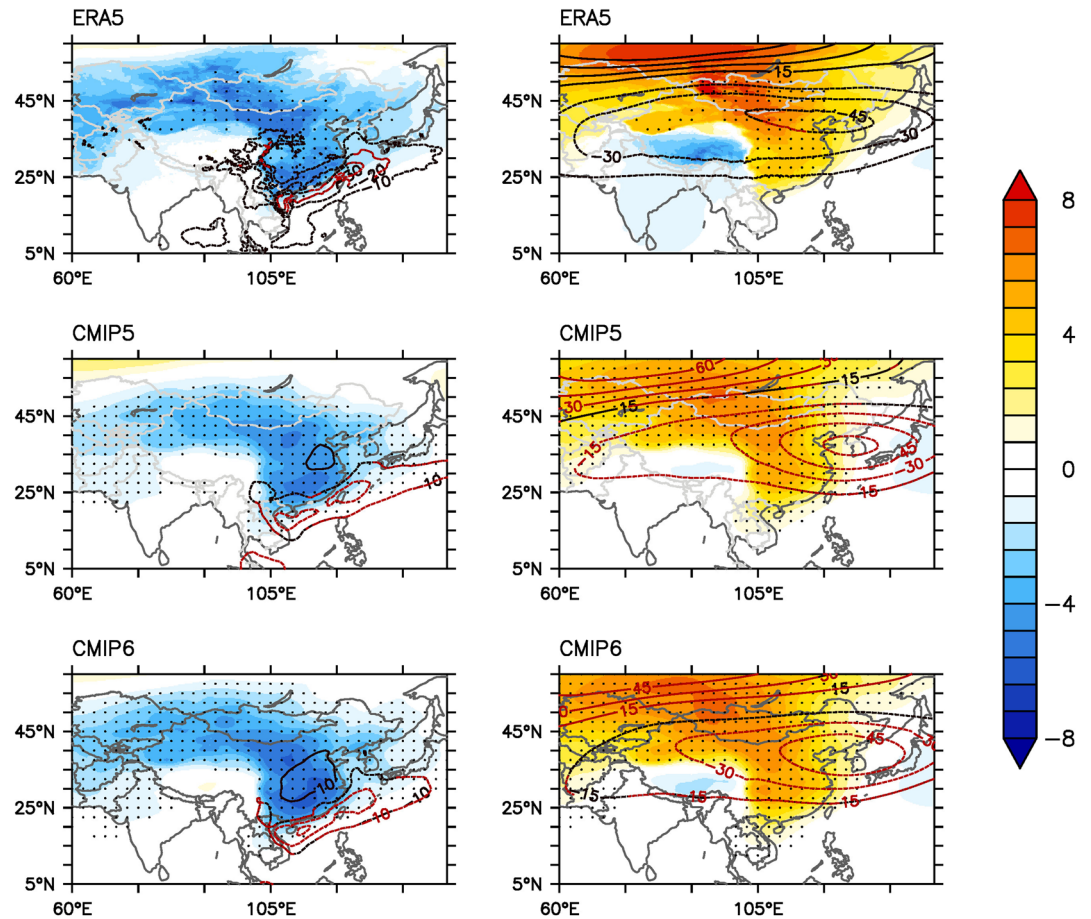
Figure 2. Composites of anomalies during WiPECE events for: (left) T2M (shading, in °C) and surface shortwave radiation (contours, in W.m^{-2}, with an interval of 10 W.m^{-2}); (right) sea-level pressure (shading, in hPa) and 500 hPa geopotential height (contours, in m, by step of 15 m). The same scale is used for both MSL and T2M although units are different. Solid (dashed) contours indicate positive (negative) anomalies. CMIP5 and CMIP6 plots show the ensemble mean of individual model-mean composites. Strippling (for MSL and T2M) or red contours (for Z500 or SSR) indicate where 90% of the cases agree on the sign of the anomaly (either 90% of cold events for ERA5, or 90% of mean individual composites for CMIP5 and CMIP6).

the members is used as a reference (instead of ERA5) and all other members are plotted against it. The process is then repeated with another member as a reference, and so on until looping through all the ensemble. Finally, the maximum error from this test is retained as being the range of internal variability of a model (i.e., the reasonable error a model could get and still be considered as consistent with the reference).

3. Results

3.1. Dynamics of Cold Events

3.1.1. Mean Dynamical Patterns

The dynamics related to WiPECE events are analyzed by compositing: for each WiPECE event, the circulation composite is computed by removing the corresponding calendar day 1981–2010 9-day running mean climatology. Then composites are averaged over the 1979–2017 period for each individual member and model, and for ERA5. Due to limitation of daily dynamical outputs for models and to analyze as many mod-

els as possible, only three variables are considered: sea-level pressure (MSL), 500 hPa geopotential height (Z500), and net shortwave surface radiation (SSR).

In ERA5, WiPECE are associated with a large mid-troposphere trough extending all over China and a surface high pressure indicating an advance of cold air from Siberian regions (Figure 2), confirming previous findings such as Bueh et al. (2011) or Song et al. (2016). Changes in SSR indicate a decrease in solar radiation especially over the South China Sea, likely due to an increase in cloud cover from the advection of continental cold air over the sea where it rapidly saturates and forms clouds. This pattern is however changing rapidly during the events (Figure 3). Before the event, the SSR anomalies cover all South-East China domain, indicating more cloudy condition, in relation with the decreasing temperatures. But during the event, this anomaly is pushed toward the coastal area, and a positive anomaly tends to develop over the region after the middle of the cold spell, indicating that this part of the domain is under clear sky conditions (i.e., drier conditions). This corresponds to the advance of the cold-dry air from the North. Both model ensembles reproduce the observed patterns well, although their magnitudes are slightly larger for MSL and Z500. They have the correct SSR signal over the South China Sea. This signal is less clear over land, and nonsignificant for both ERA5 and models.

Individual model performances are analyzed with Taylor diagrams (Figure 4). Most individual models are able to capture the right pattern signals in Z500 and MSL, with correlations above 0.8, although the magnitude of this signal for Z500 is less well-captured compared with MSL. SSR patterns have lower correlations due to the inability of models to simulate its decrease over land. There is no clear improvement between individual CMIP5 and CMIP6 models for the dynamics fields although the CMIP6 ensemble mean is slightly closer to ERA5. T2M correlations are also improved in CMIP6 models, indicating a better representation of WiPECE temperature spatial features in the new generation of models. *If a specific dataset (model) was perfectly consistent with the observations then the correlation would be one and the standard deviation identical to that observed. However, because of internal variability, the signal cannot be expected to be perfectly consistent. To evaluate whether any discrepancy could be explained by internal variability an autocorrelation test was performed (see methods), although only on the T2M and MSL results due to data availability. Thus, based on this test (red dashed lines in* Figure 4*), ensemble means are found consistent with observation (i.e., they are in the range of the internal variability).*

Based on this analysis, it is reasonable to conclude that models can reproduce WiPECE for the right dynamical reasons. Difference in SSR estimate over land could be due to the way models parametrize clouds (especially because of changing conditions from increased cloud cover to clear sky, illustrated in Figure 3), but this is a specific question that would require independent work. It is also found that CMIP6 ensemble mean is overall slightly better than CMIP5 mean. However, given the limited number of models for CMIP6, the observed improvement should not be considered as a meaningful conclusion.

3.1.2. Heat Budget

To complete the dynamical analysis, a heat budget is performed with ERA5. Composites are computed in the same way as described above but are also calculated for each day before and after each WiPECE (up to 15 days after and before the center of WiPECE event). Several radiation and heat transport variables are used and depicted in Figure 5. There are two important processes controlling the temperature drop. First, anomalous negative heat flux convergence, due to the advection of cold dry air from the North, is visible up to 8 days before (relative to the center of WiPECE). Second, as mentioned before, a decrease in net shortwave radiation, indicating an increased albedo (either change in cloud cover or snow), is also visible before WiPECE, although it is partly compensated by an increase in longwave radiation. The sum of these terms leads to a sharp decrease in the total energy budget and thus in temperature. This negative energy budget persists long enough that temperature drops below its climatological 10th percentile for several days. Temperature starts rising slowly after the center of WiPECE, when both radiation and heat transport reverse to normal or even slightly positive and the total energy budget becomes positive. Changes in surface latent and sensible heat fluxes are small. The change in P-E budget is also negligible in terms of heat budget (corresponding to a maximum anomaly of 0.026 W.m^{-2}).

This analysis highlights the importance of dynamical cold air transport and shortwave reduction, with both playing an important role in triggering and maintaining cold anomalies.
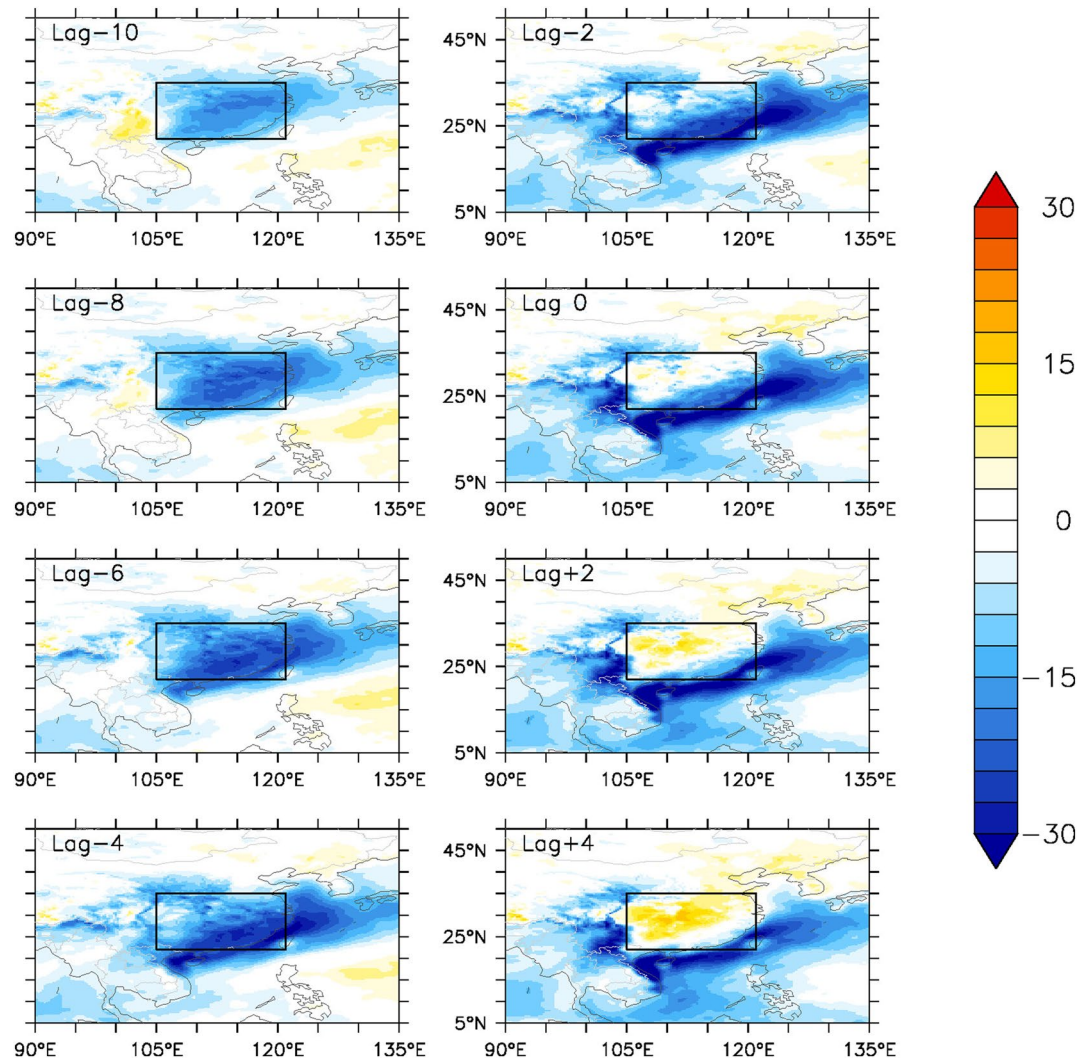
Figure 3. Anomalies of SSR (W.m^{-2}) from 10 days before (Lag−10) to 4 days after (Lag+4) the center of a cold wave (Lag 0).

3.2. Recent Trends and Attribution

The mean frequency and trend in WiPECE is now analyzed, first using the period 1979–2017 as it is common to ERA5 and other datasets. The mean number of cold days per year during this period is about 7 days in OBS and ERA5 (Figure 6a). It can vary from 0 to more than 25 days but more than half of the years have less than 10 WiPECE days. There is a good agreement between ERA5 and OBS for the mean and variability of WiPECE days per year, confirming the reliability of the reanalysis to reproduce the observations. This average frequency is less than our minimum threshold of 9 days. This is due to the spatial extent criteria: Even if at each grid point, there is a persistent cold (9 days or more), the number of days where these events overlap at a large spatial scale may be shorter (here typically 7 days). It is also noticeable that the distribution is highly asymmetric around the mean, with a hard limit on the low side (0 days) and long tail on the high side.

Model ensembles are close to observed datasets and within the range of variability although the multimodel mean is closer to 6 days per year. There is no clear difference between historical CMIP5 and CMIP6 ensembles. Individual models from both ensembles show more variable results. They are all able to capture the asymmetry of the distribution but some models underestimate the mean by 3 days and most of the models tend to have shorter tails, that is, they simulate shorter events compared with observation.
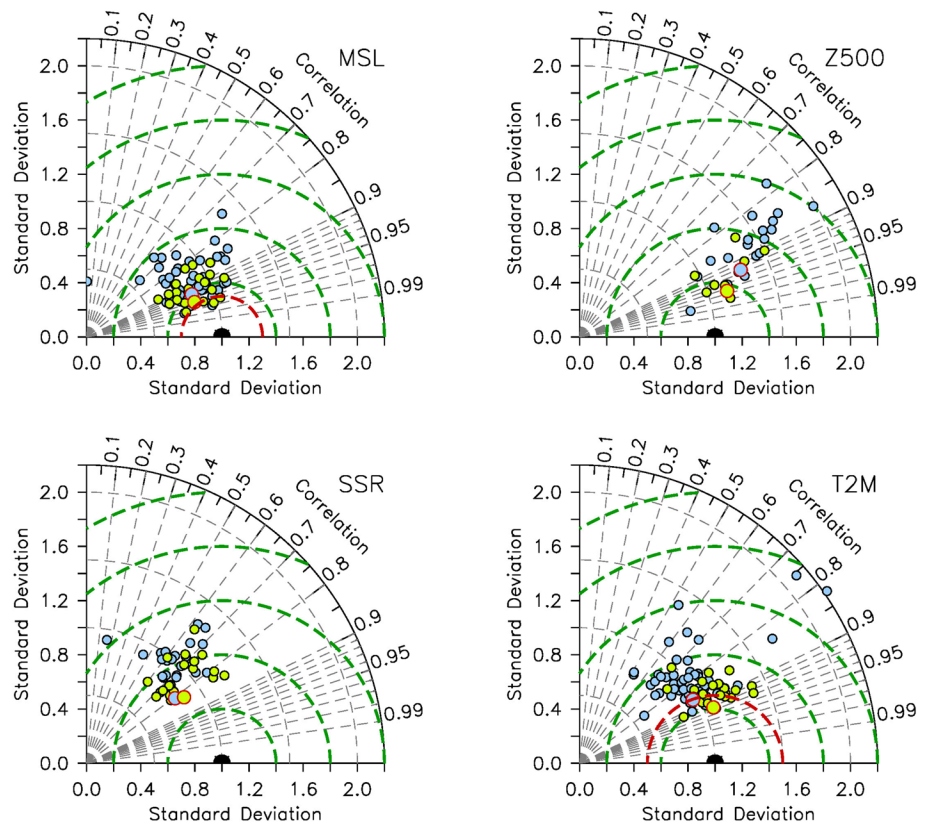
Figure 4. Taylor diagrams for sea-level pressure (MSL), 500 hPa geopotential height (Z500), shortwave solar radiation (SSR), and 2m atmospheric temperature (T2M), based on composites of anomalies during WiPECE events. CMIP5 (blue) and CMIP6 (green) individual models are displayed by black contoured circles and ensemble means are shown by red contoured circles. Reference is ERA5 (black dot). Dashed green lines show the Root Mean Square Error with an interval of 0.4, and red dashed line indicates the internal model variability based on an autocorrelation test from a model with 10 individual members (see Method).

The evolution of WiPECE days per year is now analyzed (Figure 6b). ERA5 and OBS have a very good agreement during their common period (1979–2016). They both show large peaks in 1983 (almost 40 days) and 2005 (about 30 days). Before 1979, OBS shows several high WiPECE frequency years, indicating an overall colder period. Based on these time-series, the mean trend is a decrease of about 1 day per decade (Figure 6c,
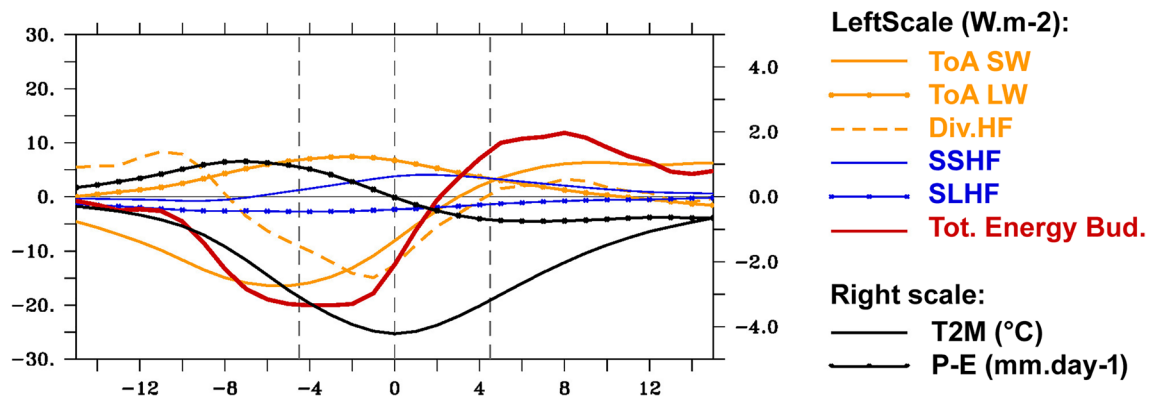

Figure 5. Time-lag energy budget during WiPECE displayed as composites of anomalies averaged over the studied region. Energy budget is decomposed as: net top of the atmosphere shortwave (ToA SW) and longwave (ToA LW) radiations, total vertically integrated energy flux divergence (Div.HF), surface sensible (SSHF), and latent (SLHF) heat flux. The sum of all previous flux is displayed in red as Total Energy Budget. All the fluxes are expressed in W.m^{-2} (left scale). The evolution of T2M (in °C, right scale) and precipitation minus evaporation (P − E, mm.d^{-1}, right scale) are also displayed in black.
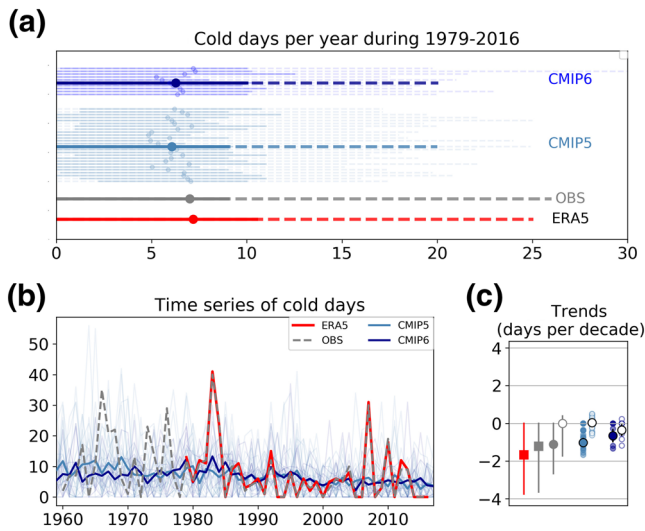
(a)

(b)

(c)

Figure 6. (a) Mean (dots) number of WiPECE days per year with the 25–75th and 5–95th percentiles (horizontal solid and dashed bars respectively) computed from the 1979–2016 period, for ERA5 (red), OBS (gray), CMIP5 (light blue) and CMIP6 (dark blue) models. For models with multiple simulations, simulations are pooled together to compute the mean and percentiles. CMIP5 and CMIP6 ensemble means are indicated by darker symbols and are computed by pooling together all WiPECE events from each individual member. (b) Time-series of number of cold days per year for ERA5, OBS and CMIP5 and CMIP6 ensemble means (thick lines) and individual models (thin lines). (c) Theil-Sen (Sen, 1968) trend estimates (in days per decade) in the number of cold days per year computed from 1960 to 2005 (circle symbols) or 1979–2005 (square symbols) time-series displayed in (b). For ERA5 and OBS, error bars show the 5%–95% confidence interval. For CMIP5 and CMIP6 individual model trends are shown by contourless circles and ensemble mean by black contoured circle. Errors in the ensemble mean trends show the 5%–95% confidence interval and are computed by bootstrapping the ensembles (see Methods). Finally, empty circles in (c) indicate trends after removing the interannual changes in T2M (see Methodology).

(a)

(b)

Figure 7. (a) as Figure 6a but based on 5-years mean anomaly time-series. (b) distribution of cold days per year (anomalies relative to the 1960–2005 mean) for the CESM large ensemble where individual member results are pooled to compute the distribution.

either for 1979–2016 or the whole 1960–2016 period), close to the value of 0.9 day per decade found by Shi et al. (2018) for their cold spell duration index. The large uncertainties are due to the high interannual variability. It is noticeable that the recent 2016 winter event (Qian et al., 2017) does not appear in the WiPECE index as it lasted less than 9 days in most of the places. Some parts of China also experienced more severe reductions in cold events, as for example, Guangzhou where a decreasing trend of 1.77 days per decade was found (Zhang et al., 2017) although there definition of cold spells was different.

Model ensemble means are not expected to reproduce the same variability as the averaging process removes internal variability. Moreover, CMIP models are fully coupled, thus the ocean signal is not expected to correspond year to year to the actual observed signal. Despite these aspects, both ALL model ensembles are within the range of observation and are able to reproduce the decreasing trend over the 1960–2016 period. Individual models interannual variability is consistent with OBS and ERA5 and model trends are in the range of observation errors. The high interannual variability leads to large model spread and observational error on trends. This internal variability is greatly influenced to the variability of the East Asia winter monsoon, which experienced a weakening in the late 1980s and a reamplification in the early 2000s (Wang & Chen, 2014; Wang et al., 2009), consistent with the WiPECE signal.

To analyze the role of anthropogenic forcing in model ensembles, WiPECE signals are analyzed with the different scenarios described in the methodology. Here trends are based on 5-years mean values and for the 1960 to 2005 period to cover the same period for each ensemble (NAT ending in 2005 for CMIP5). The multiyear averaging reduces variability and allows more focus on the long-term changes and forced variability. It also leads to visually more Gaussian-like distributions (shown in Figure 7), a requirement for the attribution method. However, the model distribution, on the basis of a K-S test, is not consistent with a Gaussian distribution (tested with the CESM large ensemble, with a *p*-value result close to 0 leading to reject the null hypothesis that the two samples were drawn from the same distribution). This remains a challenge to apply attribution methods to extreme events (Hegerl et al., 2006) and results below are discussed with caution.

Based on 5-years mean time-series, the OBS trend is highly uncertain (Figure 8, compared to Figure 6c). Temporal smoothing leads to less interannual variability but also less sampling over the period, thus similar uncertainties to evaluate the trends. This is echoed by model internal variability (evaluated from piControl runs). The spread between models encompass the observed trend and is of the same order as the observational uncertainty. Simply based on these facts, it cannot be ruled out that the recent observed trend in WiPECE is simply due to internal variability. This is confirmed by DA results (Table 2), with about a 10% probability that the observed signal is explained by internal variability.

Individual forcing results bring more insight on the contribution from different factors. GHG trends are the closest to the observed signal and show the most confident results with almost all models having negative trends, although most of those are weaker than the observed best estimate. Results are less clear for ALL and ANT, especially for the CMIP6
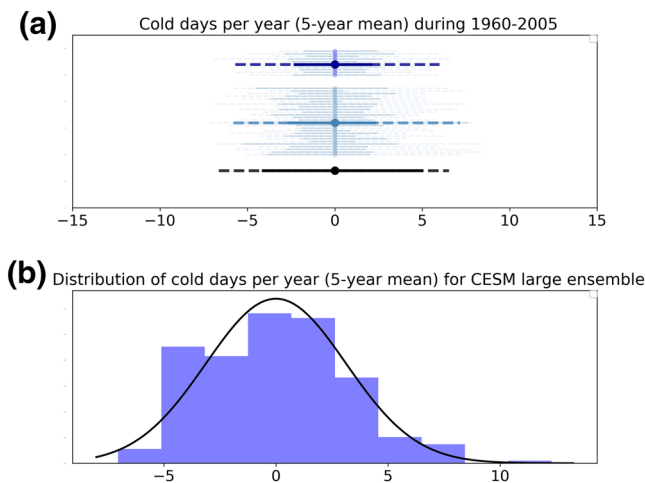
Table 2
Hypothesis Testing P-Value for Individual Forcings From CMIP5 and CMIP6 Ensemble Using De Abreu et al., 2019 Method.

	Internal variability	NAT	GHG	ALL	Combined NAT + GHG + OA	OA	ANT
CMIP5	0.12	0.31	0.49	0.65	0.89	0.42	0.47
CMIP6	0.12	0.32	0.69	0.37	0.84	0.07	0.33

Individual forcings correspond to: Natural forcing only (NAT), Greenhouse Gases only (GHG), all historical forcing (ALL) and residual forcing (OA, mainly aerosols). A test is also performed by combining GHG + NAT + OA models to re-create an ensemble similar to ALL but only with models available for both GHG and NAT and using errors from each sub-ensemble to compute the covariance matrix. The anthropogenic signal only is built with ALL-NAT.
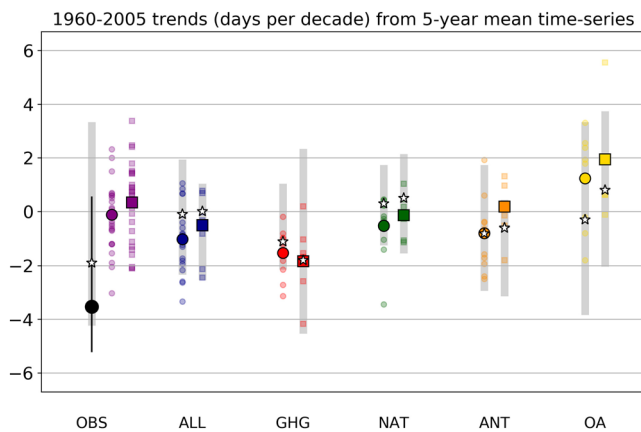

Figure 8. 1960–2005 trends estimate from Theil-Sen method and based on 5-years mean WiPECE values for OBS (black symbol) and different forcings of CMIP5 (circle symbols) and CMIP6 (square symbols). Ensemble mean trend is shown by the large dark symbol and individual model trends are displayed by light symbols. For OBS, the vertical black bar indicates the 5%–95% confidence interval from the Theil-Sen method. For each case, white star symbols indicate the Theil-Sen trends based on the best fit of the DA method and the gray bar is the 5%–95% confidence interval estimated from 1,000 resampling. Purple symbols show results from a 45 years section of PiControl simulations, indicating the internal variability of the models.
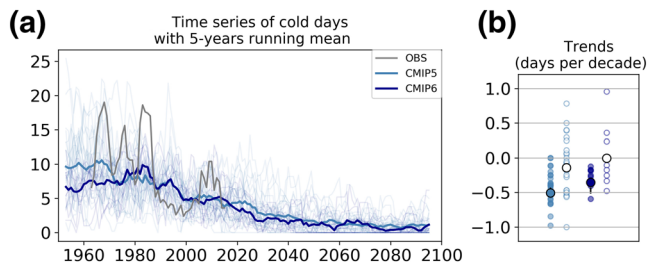

Figure 9. As Figures 6b and 6c but including long term projections for CMIP5 (with RCP4.5 scenario from 2006) and CMIP6 (with SSP245 scenario from 2015). Trends are computed for the 2005–2065 period.

ensemble. This indicates that other contributions in the models, such as anthropogenic aerosol emissions, tend to mask or counter act GHG effects (as displayed with OA trends).

Attribution results reflect these results with larger confidence on the role of GHG to explain observation. However, each individual forcing result also shows a potential contribution and none can be excluded with good confidence, including NAT. This is the major point from this analysis: Due to large internal uncertainties and intermodel spread, none of the individual forcing hypotheses can be rejected. Even in OBS, the signal is barely detectable if we consider the uncertainty (internal variability) associated with its trend.

The above considerations indicate that anthropogenic impact on WiPECE recent trend is still difficult to attribute as the signal barely emerges from the internal variability, consequently it is not possible to carry out a robust detection of a forced signal. Still, at least in the model simulations, it can be seen that forcing from GHGs reduces the occurrence of WiPECE (due to warmer temperatures) but will have been largely offset by other external forcings.

Finally, the question of potential changes in circulation is addressed as WiPECE trends could be due to either a change in the mean temperature or a change in the frequency of the dynamics leading to these cold events (as analyzed in Section 3.1), or perhaps both. To do so, the mean winter temperature in each model and OBS is first detrended at each grid point and then WiPECE events are re-computed based on these detrended temperatures. Trends in WiPECE obtained with this methodology are almost zero (Figure 6c, empty symbols). Thus, the observed decrease in cold events, even if still hard to detect, has been mainly due to a change in mean temperature. Results from this indirect method do not indicate a contribution from changing circulation. This point would require confirmation from a more careful analysis based directly on the analysis of circulation indices, which is beyond the scope of this study.

3.3. End-of-Century Projected Changes

Long-term evolution of WiPECEs are investigated for the end-of-century projection using RCP4.5 and SSP245 scenarios for CMIP5 and CMIP6 ensembles respectively. As expected in a warming world the frequency of cold events reduces (Figure 9). In opposition to recent trends analyzed previously, in future projections the forced signal becomes much stronger, and by the end of the century has emerged from internal variability leaving a clearly detectable signal. This corresponds to a dominant GHG
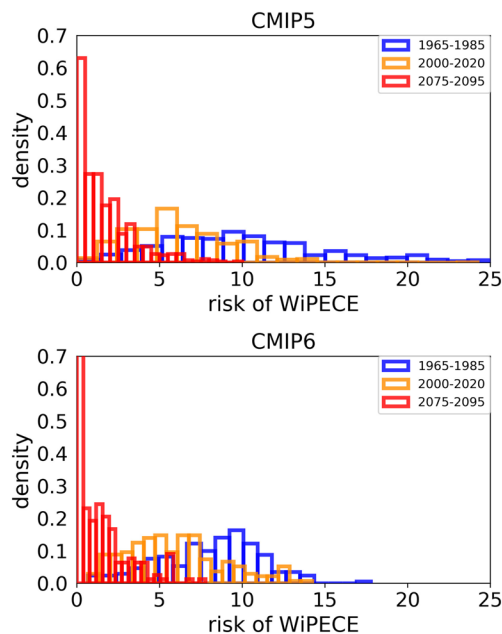
Figure 10. Distribution of the WiPECE probability (%) of WiPECE per winter day based on 5-year running means, for CMIP5 (historical simulation extended with RCP4.5 after 2005) and CMIP6 (historical simulation extended with SSP245 after 2014) ensemble computed for three different periods: 1965–1984, 2000–2019, and 2075–2094. Density histograms are computed from pooling together each individual model during each specific period. The risk corresponds to the chance (in %) that each day and location belongs to a WiPECE.

signal on temperatures along with a potential decreasing impact of other forcing such as aerosols.

Ensemble means indicate a downward trend in WiPECE days until the middle of 21st century, followed by a stabilization around 1 or 2 days per winter. If we consider that models tend to underestimate the observed trend (as mentioned previously) this reduction could occur even faster. After the middle of the century, individual models still show year to year variability with peaks up to 5 days. Expressed as a risk (Figure 10), this translates by a change from currently 6%–7% chance that a day belongs to a WiPECE to a 0%–2% chance at the end of the century.

Thus, even if WiPECE frequency is expected to strongly reduce, this type of event will still occasionally occur. According to model projections, cold surges will not disappear and temperatures in South-East China will still reach below their current 10% climatology coldest values for many days. To investigate the role of a potential long-term circulation change in the future, the same methodology as in the previous section was applied. Results (Figure 9b) indicate again that without changes in the mean temperature the frequency of WiPECE stays fairly constant (empty symbols, to compare with filled symbols). Thus, as this detrend signal is close to zero, there is no clear indication that WiPECE long-term trend could be due to a circulation change although this is an indirect conclusion and more precise analysis on the circulation change would be needed.

4. Conclusions

In this study, we focused on persistent large-scale cold events (Winter widespread Persistent Extreme Cold Events, WiPECE) over South East China, investigating their dynamics and recent trends, with a focus on how well models can reproduce the observed signal.

We analyzed the dynamical pattern associated with these events using ERA5 reanalysis. We showed that a strong advection of cold air from the North and a decrease in shortwave radiation over South East China were both responsible for the persistent drop of temperatures.

Historical simulation of global climate models (CMIP5 and CMIP6) showed good skills to reproduce the dynamical patterns identified in ERA5 although the change in shortwave radiation is less well captured.

We then analyzed the frequency of WiPECE days in both ERA5 and an observational network (OBS). Both datasets indicate large interannual variability, ranging from 0 to 25 WiPECE days in a year. We found that models were able to reproduce correctly the asymmetry of the observed distribution and its variability.

From 1960 to 2005, observations and models suggested a decrease in WiPECE days of about 1 day per decade. However, we couldn't attribute this trend with good confidence due to a large internal variability (in models and observations). Models overall seemed to underestimate the magnitude of the observed trend. Our analysis indicates that greenhouse gases largely drive this decreasing trend while aerosol impact (and other residual factors) tends to mitigate it. Some models are more sensitive to aerosols leading to competing effects with GHG and unclear trends in WiPECE. Based on a DA analysis, we found that none of the single forcing results could be excluded with good confidence. Each single forcing shows at least some probability of explaining the observed recent trend largely due to the high degree of internal climate variability. This also shows the limit of DA methods to analyze extreme events over a limited period, when the signal-to-noise ratio is low (in both observation and models).

We extended our analysis to long-term projections and confirmed, as many studies did, that the risk of WiPECE is expected to reduce by the end of the century but it should not disappear completely. We found that the main driver of change was the increase in winter-mean temperatures reducing the frequency of

long-lasting cold temperatures below the historical values. This suggests, indirectly, that circulation changes and changes in temperature variances are not significant drivers of projected change in WiPECE. Large dynamical advection of cold air should still occur at the same frequency. If the population is less prepared to endure cold surges, because more focus is made on adapting to warmer temperatures, this could lead to an increased vulnerability of the South-East China population.

Finally, we could not find clear differences between CMIP5 and CMIP6 ensemble performances to reproduce the observed signals though, in this study, the number of CMIP6 models was much lower than the number of CMIP5 models. Thus, as more CMIP6 model results become available, this lack of clear differences needs further verification.

Data Availability Statement

All data used in this study are publicly available. CMIP5 (Taylor et al., 2012) and CMIP6 (Eyring et al., 2016) can be downloaded for example from https://esgf-node.ipsl.upmc.fr/project/. CESM Large Ensemble (Kay et al., 2015) data are available at http://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html. Homogenized temperature data for China (Li & Yan, 2009) are available through the Chinese Meteorological Administration website (https://data.cma.cn/en/).

References

Bueh, C., Xian-Yue, F. U., & Zuo-Wei, X. (2011). Large-scale circulation features typical of wintertime extensive and persistent low temperature events in China. *Atmospheric and Oceanic Science Letters*, *4*, 235–241.

Chen, J.-H., Jiang, H., Wu, L., Liao, X., Lu, Y., Tao, X.-Q., et al. (2017). Association of ischemic and hemorrhagic strokes hospital admission with extreme temperature in Nanchang, China. *Journal of Clinical Neuroscience*, *43*, 89–93.

Chen, J., Yang, J., Zhou, M., Yin, P., Wang, B., & Liu, J., et al. (2019). Cold spell and mortality in 31 Chinese capital cities: Definitions, vulnerability and implications. *Environment International*, *128*, 271–278.

De Abreu, R. C., Tett, S. F. B., Schurer, A., & Rocha, H. R. (2019). Attribution of detected temperature trends in Southeast Brazil. *Geophysical Research Letters*, *46*, 8407–8414. https://doi.org/10.1029/2019GL083003

Du, Y.-D., Wang, X.-W., Yang, X.-F., Ma, W.-J., Ai, H., & Wu, X.-X. (2013). Impacts of climate change on human health and adaptation strategies in South China. *Advances in Climate Change Research*, *4*. https://doi.org/10.3724/SP.J.1248.2013.208

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*, LLNL-JRNL-736881.

Freychet, N., Tett, S., Wang, J., & Hegerl, G. (2017). Summer heat waves over eastern China: Dynamical processes and trend attribution. *Environmental Research Letters*, *12*(2), 024015

Freychet, N., Tett, S. F. B., Yan, Z., & Li, Z. (2020). Underestimated change of wet-bulb temperatures over East and South China. *Geophysical Research Letters*, *47*, e2019GL086140. https://doi.org/10.1029/2019GL086140

Hegerl, G. M., Allen, N., Bindoff, N., Gillett, D., Karoly, X., Zhang, F., & Zwiers, F. (2006). Climate change detection and attribution: Beyond mean temperature signals. *Journal of Climate*, *19*, 5058–5077. https://doi.org/10.1175/JCLI3900.1

Hersbach, H., & Coauthors (2018). *Operational global reanalysis: Progress, future directions and synergies with NWP*. European Centre for Medium Range Weather Forecasts. https://doi.org/10.21957/tkic6g3wm

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The community Earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*, 1339–1360.

Kay, J., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The community Earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*, 1333–1349.

Liao, Z., Zhai, P., Chen, Y., & Lu, H. (2020). Differing mechanisms for the 2008 and 2016 wintertime cold events in southern China. *International Journal of Climatology*, *40*(11), 4944–4955. https://doi.org/10.1002/joc.6498

Li, M., He, B., Guo, R., Li, Y., Chen, Y., & Fan, Y. (2018). Study on population distribution pattern at the county level of China. *Sustainability*, *10*, 3598. https://doi.org/10.3390/su10103598

Li, C., Tian, Q., Yu, R., Zhou, B., Xia, J., Burke, C., et al. (2018). Attribution of extreme precipitation in the lower reaches of the Yangtze River during May 2016. *Environmental Research Letters*, *13*, 014015.

Li, Z., & Yan, Z.-W. (2009). Homogenized daily mean/maximum/minimum temperatures series for China from 1960-2008. *Atmospheric and Oceanic Science Letters*, *2*, 237–243. https://doi.org/10.1080/16742834.2009.11446802

Li, Q., Zhang, H., Liu, X., Chen, J., Li, W., & Jones, P. (2009). A mainland China homogenized historical temperature dataset of 1951–2004. *Bulletin of the American Meteorological Society*, *90*, 1062–1065.

Luo, Y., Li, H., Huang, F., Van Halm-Lutterodt, N., Xu, Q., Wang, A., et al. (2018). The cold effect of ambient temperature on ischemic and hemorrhagic stroke hospital admissions: A large database study in Beijing, China between years 2013 and 2014—Utilizing a distributed lag non-linear analysis. *Environmental Pollution*, *232*, 90–96.

Ma, W., Xu, X., Peng, L., & Kan, H. (2011). Impact of extreme temperature on hospital admission in Shanghai, China. *The Science of the Total Environment*, *409*, 3634–3637.

Ma, W., Yang, C., Chu, C., Li, T., Tan, J., & Kan, H. (2013). The impact of the 2008 cold spell on mortality in Shanghai, China. *International Journal of Biometeorology*, *57*, 179–184.

Peng, J.-B., & Bueh, C. (2011). The definition and classification of extensive and persistent extreme cold events in China. *Atmospheric and Oceanic Science Letters*, *4*, 281–286. https://doi.org/10.1080/16742834.2011.11446943

Powell, J. P., & Reinhard, S. (2016). Measuring the effects of extreme weather events on yields. *Weather and Climate Extremes*, *12*, 69–79.

Qian, C., Wang, J., Dong, S., Yin, H., Burke, C., Ciavarella, A. (2017). Human influence on the record-breaking cold event in January of 2016 in eastern China. *Bulletin of the American Meteorological Society*, *99*(1), S118–S122. https://doi.org/10.1175/BAMS-D-17-0095.1

Ribes, A., Zwiers, F. W., Azaïs, J., & Naveau, P. (2017). A new statistical approach to climate change detection and attribution. *Clim Dyn*. *48*, 367–386. https://doi.org/10.1007/s00382-016-3079-6

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, *63*, 1379–1389.

Shi, J., Cui, L., Wen, K., Tian, Z., Wei, P., & Zhang, B. (2018). Trends in the consecutive days of temperature and precipitation extremes in China during 1961–2015. *Environmental Research*, *161*, 381–391.

Song, L., Wang, L., Chen, W., & Zhang, Y. (2016). Intraseasonal variation of the strength of the East Asian trough and its climatic impacts in boreal winter. *Journal of Climate*, *29*, 2557–2577.

Sparrow, S., Su, Q., Tian, F., Li, S., Chen, Y., Chen, W., et al. (2018). Attributing human influence on the July 2017 Chinese heatwave: The influence of sea-surface temperatures. *Environmental Research Letters*, *13*, 114004.

Sun, J., & Zhao, S. (2010). The impacts of multiscale weather systems on freezing rain and snowstorms over southern China. *Weather and Forecasting*, *25*, 388–407.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*, 7183–7192. https://doi.org/10.1029/2000JD900719

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*, 485–498.

Wang, L., & Chen, W. (2014). the East Asian winter monsoon: Re-amplification in the mid-2000s. *Chinese Science Bulletin*, *59*, 430–436.

Wang, L., Huang, R., Gu, L., Chen, W., & Kang, L. (2009). Interdecadal variations of the East Asian winter monsoon and their association with quasi-stationary planetary wave activity. *Journal of Climate*, *22*, 4860–4872.

Wang, L., Liu, T., Hu, M., Zeng, W., Zhang, Y., Rutherford, S., et al. (2016). The impact of cold spells on mortality and effect modification by cold spell characteristics. *Scientific Reports*, *6*, 1–10.

Wu, W., Xiao, Y., Li, G., Zeng, W., Lin, H., Rutherford, S., et al. (2013). Temperature–mortality relationship in four subtropical Chinese cities: A time-series study using a distributed lag non-linear model. *The Science of the Total Environment*, *449*, 355–362.

Xiong, W., Li, C., Deng, L., & Li, Q. (2010). Impacts of ice disaster on health and daily life of residents. *Chinese Journal of Public Health*, *26*, 1256–1258.

Zhang, R., Chen, Z. Y., Ou, C. Q., & Zhuang, Y. (2017). Trends of heat waves and cold spells over 1951–2015 in Guangzhou, China. *Atmosphere*, *8*, 37.

Zhang, Z. J., & Qian, W. H. (2011). Identifying regional prolonged low temperature events in China. *Advances in Atmospheric Sciences*, *28*, 338–351.

Zhou, W., Chan, J. C. L., Chen, W., Ling, J., Pinto, J. G., & Shao, Y. (2009). Synoptic-scale controls of persistent low temperature and icy weather over southern China in January 2008. *Monthly Weather Review*, *137*, 3978–3991.

Zuo, Z., Zhang, R., Huang, Y., Xiao, D., & Guo, D. (2015). Extreme cold and warm events over China in wintertime. *International Journal of Climatology*, *35*, 3568–3581.