Clim Dyn
DOI 10.1007/s00382-012-1565-z

Evaluation and response of winter cold spells over Western Europe in CMIP5 models

Y. Peings · J. Cattiaux · H. Douville

Received: 30 July 2012 / Accepted: 9 October 2012

Abstract This paper is dedicated to the analysis of winter cold spells over Western Europe in the simulations of the 5th phase of the Coupled Model Intercomparison Project (CMIP5). Both model biases and responses in a warming climate are discussed using historical simulations and the 8.5 W/m^2 Representative Concentration Pathway (RCP8.5) scenario, respectively on the 1979–2008 and 2070–2099 periods. A percentile-based index (10th percentile of daily minimum temperature, Q10) with duration and spatial extent criteria is used to define cold spells. Related diagnostics (intensity, duration, extent, and severity as a combination of the former three statistics) of 13 models are compared to observations and suggest that models biases on severity are mainly due to the intensity parameter rather than to duration and extent. Some hypotheses are proposed to explain these biases, that involve large-scale dynamics and/or radiative fluxes related to clouds. Evolution of cold spells characteristics by the end of the century is then discussed by comparing RCP8.5 and historical simulations. In line with the projected rise of mean temperature, "present-climate" cold spells (computed with the 1979–2008 10th percentile, Q10P) are projected to be much less frequent and, except in one model, less severe. When cold spells are defined from the future 10th percentile threshold ("future-climate" cold spells, Q10F), all models simulate a decrease of their intensity linearly related to the seasonal mean warming. Some insights are given to explain the inter-model diversity in the magnitude of the cold spells response. In particular, the snow-albedo feedback is suggested to play an important role, while for some models changes in large-scale dynamics are also not negligible.

Keywords Climate extremes · Cold spells · European climate · CMIP5 models · Models evaluation

Y. Peings · J. Cattiaux · H. Douville
CNRM-GAME, Météo-France and CNRS, Toulouse, France

Y. Peings (✉)
CNRM/GMGEC/VDR, 42 avenue Gaspard Coriolis,
31057 Toulouse Cedex 01, France
e-mail: yannick.peings@cnrm.meteo.fr;
yannick.peings@meteo.fr

1 Introduction

Extreme climate events can have high socio-economic impacts (Mc Michael et al. 2003; McGregor et al. 2005). They have been investigated in many studies using either observations (Easterling et al. 2000; Frich et al. 2002; Alexander et al. 2006) or climate models (Hegerl et al. 2004; Meehl and Tebaldi 2004; Tebaldi et al. 2006). In particular, cold extremes show a significant decrease during the last decades (Christidis et al. 2005; Brown et al. 2008) consistently with the observed global warming attributed to the human-induced greenhouse gases by the Intergovernmental Panel on Climate Change (IPCC 2007). Climate projections for the twenty-first century confirm this trend with, for instance, frost days being less frequent over Europe (Heino et al. 1999; Meehl et al. 2004). Some studies suggest that future cold extremes, though less frequent, could be as severe as present ones at the global scale (Vavrus et al. 2006; Kodra et al. 2011). However, it is not clear whether these changes will be driven by changes in mean temperature (Räisänen and Ylhäisi 2011) and/or in the shape of the daily temperature distribution (Ballester et al. 2010).

Here we focus on the Western Europe region, where cold spells can have strong consequences on populations (Huynen et al. 2001). This region has experienced particularly cold temperatures during the three last winters

Published online: 27 October 2012

(especially in December 2010, January 2011, and February 2012), which raised some questions about the fate of such cold events in a warming climate (Cattiaux et al. 2010; Guirguis et al. 2011). de Vries et al. (2012) recently discussed in details cold spells statistics over Western Europe, both for present-day and future climate. They concluded that cold spells could be $\sim$5 °C warmer in a future climate and found a change both in the mean and the variance of the wintertime temperature distribution. They partly explained this result by a reduction of the mean temperature zonal gradient due to the higher warming of land compared to ocean. Other studies suggested that the strong reduction in the Arctic sea-ice could be responsible for cold extreme temperatures over Europe through the propagation of Rossby waves (Petoukhov and Semenov 2010; Overland and Wang 2010).

The recent availability of global climate simulations from the 5th phase of the Coupled Model Intercomparison Project (CMIP5) is a new opportunity to assess how cold spells could evolve until the end of the twenty-first century. An important prior step is to check the ability of the new generation of coupled ocean–atmosphere General Circulation Models (GCMs) to simulate cold spells accurately. In this study, we evaluate the cold spells characteristics over Western Europe in a subset of 13 CMIP5 models by using present-day and future climate simulations. The questions we address are the following:

- how well do CMIP5 models represent observed cold spells ?
- how do cold spells characteristics evolve during the twenty-first century under a severe greenhouse gas concentration scenario ?
- how are cold spells biases/responses related to their climatological biases/sensitivity ?

The paper is organized as follows. Section 2 includes a description of observations and models, as well as some details about the cold spells definition and statistics. Section 3 and 4 present the main results of our analysis, respectively for current and future cold spells. Their characteristics are summarized for each model and some possible mechanisms are proposed to explain the models biases/responses. Finally, a summary of our results and future potential prospects are provided in the conclusion.

2 Data and methodology

2.1 Datasets for observations and models

Observed near surface temperature data are taken from the European Climate Assessment and Dataset (ECA&D) project which provides a 0.5° gridded dataset (E-OBS v3.0, Haylock et al. 2008) of daily minimum, maximum and mean temperatures derived from in situ measurements at European stations (http://eca.knmi.nl/dailydata/). This dataset (hereafter referred to as EOBS) is used on the 1979–2008 period to evaluate CMIP5 models. NCEP2 (Kalnay et al. 1996) sea level pressure (SLP) reanalysis and ERSST-v3 (Smith et al. 2008) observed monthly mean sea surface temperatures (SST) are also used. Finally, the NASA/GEWEX Surface Radiation Budget (SRB, Whitlock et al. 1993) dataset, only available over the 1984–2007 period, is used as a reference for net radiation (Rnet), total cloud cover fraction, cloud radiative forcing (CRF) and clear-sky albedo (ALBCS). The CRF is used to diagnose the influence of clouds on the surface energy budget during cold spells. It is computed as the difference between the all-sky and clear-sky net radiative fluxes.

Concerning models, 13 GCMs of the CMIP5 database have been selected based on their data availability at the time of our study (see Table 1 for list of models). Present-day climate is characterized in models by using the historical (referred as HIST in the following of the study) and AMIP (Atmospheric Model Intercomparison Project) simulations. Historical simulations are twentieth century coupled ocean–atmosphere simulations forced by observed natural and anthropogenic forcings. AMIP simulations are atmospheric simulations driven by observed monthly mean sea surface temperatures and similar radiative forcings (except climatological rather than time-dependent anthropogenic aerosols). These additional experiments only cover the 1979–2008 period and are here used to assess the SST contribution to cold spells biases in coupled simulations. For future climate, we focus on the end of the twenty-first century by using the 8.5 W/m^2 Representative Concentration Pathway scenario over the 2070–2099 period (hereafter referred to as RCP8.5). Since the HIST simulations end in 2005, the first member of the historical run was concatenated with its corresponding RCP8.5 member to evaluate each model over the same 1979–2008 period as in the AMIP simulations.

All datasets have been interpolated on a common medium T127 horizontal grid (1.4° $\times$ 1.4°) to ease intercomparison between observations and models.

2.2 Duration/spatial extent of present-day cold extremes and definition of cold spells

Cold days are first identified at each grid cell by using the 10th percentile (Q10) of the daily minimum temperature (Tmin) as a threshold. The observed Q10 is shown in Fig. 1a with the domain of study covering Western and Central Europe between 42 N/55 N and −10 W/25E (hereafter referred to as the "WE domain"). Spatial and temporal criteria are then necessary to define cold spells at

Table 1 Summary of WECS statistics for EOBS and historical simulations of CMIP5 models

Acronym	Center	Model	Frequency	Severity	Intensity	Duration	Extent
EOBS	/	/	10.2	−54.7	−8.7	10.3	0.56
CNRM	CNRM and CERFACS, France	cnrm-cm5	9.5	−47.8	−9.1	9.6	0.52
IPSL	IPSL, France	ipsl-cm5a-lr	9.1	−75.3	−10.9	10.7	0.59
BCC	BCC, China	bcc-csm1-1	9.4	−70.6	−9.9	10.3	0.60
CCCMA	CCCMA, Canada	canesm2	8.6	−48.2	−8.5	10.1	0.53
CSIRO	CSIRO and QCCCE, Australia	csiro-mk360	9.5	−59.5	−10.5	10.8	0.50
GFDL	NOAA-GFDL, USA	gfdl-esm2 m	10.5	−45.0	−7.7	9.8	0.54
INM	INM, Russia	inmcm4	9.8	−76.1	−11.6	12.2	0.49
MIROC	AORI, NIES and JAMSTEC, Japan	miroc5	11.0	−50.2	−7.8	11.1	0.54
HAD	MOHC, UK	hadgem2	12.1	−56.8	−8.6	12.4	0.48
MPI	MPIM, Germany	mpi-esm-lr	10.9	−58.2	−8.1	11.6	0.55
MRI	MRI, Japan	mri-cgcm3	8.5	−56.2	−8.4	10.3	0.55
NCAR	NCAR, USA	ccsm4	9.2	−57.6	−8.1	11.1	0.56
NCC	NCC, Norway	noresm1-m	10.6	−54.1	−8.5	11.0	0.52

Frequency is expressed in % (of DJFM days), severity in °C day, intensity in °C, duration in days and extent in fraction of the WE domain. Period: 1979–2008

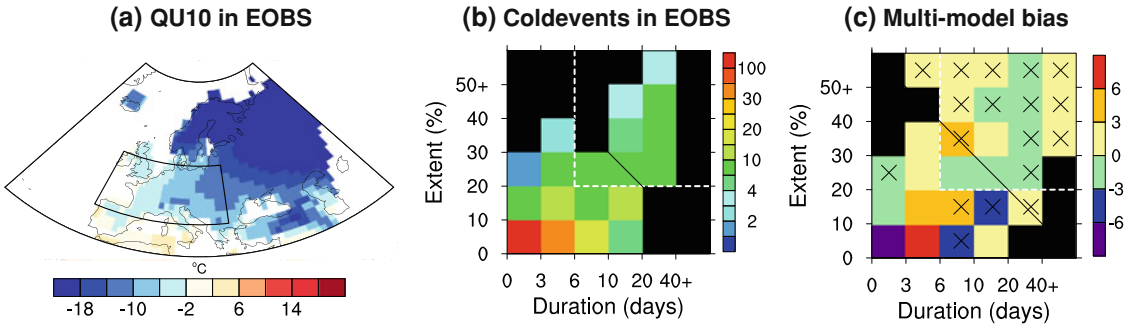

Fig. 1 **a** 10th percentile of Tmin over Europe in EOBS and WE domain; **b** Number of cold events (one or more grid point below the Q10) for EOBS, according to their extent and duration. Dashed white lines indicate the events defined as cold spells; **c** Same but for multi-model bias in historical simulations (HIST-EOBS). Crosses indicate regions where the sign consistency between models is higher than 80 %. Period: 1979–2008; Season DJFM

the regional scale. They are necessarily somewhat arbitrary and depend on the severity of cold spells we want to focus on. For this choice, it is useful to have an idea of the temporal and spatial characteristics of cold events over our domain of interest. Figure 1b shows the number of cold extremes in EOBS over the WE domain, as a function of both duration and spatial extent intervals. A cold extreme is basically one or more consecutive days with at least one grid point below the Q10. It explains the great number in the bottom left corner of the plot, as 1 day with only one grid point below the Q10 is considered as a cold extreme. As expected, there is a fairly linear relationship between duration and extent of the cold extremes, the longest events generally corresponding to the most extended ones. The dashed white line depicts the major events defined as cold spells in the following of the study: we require that events cover at least 20 % of the WE domain and last more than 6 consecutive days. This definition retains 36 events on the 1979–2008 period in EOBS observations. To validate our methodology, we have verified that it captures the main cold spells recorded in France since 1979 as for example the January 1987 cold wave (French national meteorological service, personal communication).

The same diagnostic has been applied onto the HIST simulations over the same period. The number of selected cold spells ranges between 29 and 39 depending on the model, with a multi-model climatology of 33 i.e. slightly more than one cold spell per year as in observations. The ensemble mean multi-model biases as a function of extent and duration are shown in Fig. 1c. To test their significance, crosses have been superimposed onto the pixels where at least 80 % of the models are in agreement with the sign of the ensemble-mean bias. The biases are moderate and sign consistency is high for the cold spells (white dashed line). The same diagnostic computed from AMIP simulations gives a multi-model average of 35 cold spells

(from 28 to 45) and quite similar results for the bias distribution as a function of duration and extent.

Our cold spell definition (using Q10 and 6 consecutive days for duration criterion) is consistent with the recommendations of the Expert Team on Climate Change Detection and Indices (ETCCDI) (Zhang et al. 2011). For each model, the calculation is made with the simulated Q10 rather than the observed one. It means that the simulated cold spells are defined independently of the climatological biases. Our season of interest runs from December to March (DJFM) and we use the seasonal Q10 instead of the calendar day Q10, which means that cold spells are defined in the same way regardless of the wintertime day.

2.3 Characteristics of cold spells

Several statistics averaged over the 1979–2008 period are used to describe WE cold spells (referred to as WECS in the following of the study):

- *Intensity:* average of Tmin anomaly during WECS. Daily anomalies are computed by removing a 15-day moving window climatology. For each WECS day, a spatial average of anomalies on (not necessarily adjacent) cold pixels is computed. Then, a temporal average is done on WECS days. Unit: °C.
- *Duration:* number of days of WECS (by construction $\geq$6 days). Unit: days.
- *Extent:* extent of WECS, expressed as the fraction of the WE domain (by construction $\geq$0.2).
- *Severity:* expressed for each WECS as the product of the 3 previous statistics (severity = intensity $\times$ duration $\times$ extent). It is similar to a degree-day index (accumulation of °C below daily climatology) weighted by the fraction of the domain which is covered by the WECS (0.2–1). Unit: °C days

In addition, the climatological frequency of WECS is computed as the percentage of winter days when the WE domain experiences cold spell conditions. Though cold days are defined on a Q10 threshold basis, it is not necessarily 10 % because of the duration and extent criteria in the regional cold spell definition.

2.4 Monthly cold spell severity index (CSSI) and composites analysis

To discuss the relationship between the cold spells biases/responses and the mean climate, a regional and monthly cold spell severity index (CSSI) has been constructed by summing the daily area-averaged WECS intensities for each month. Intensity, duration and extent are thus taken into account in this synthetic index and months with a high CSSI are characterized by a strong occurrence of WECS. Composites are then calculated by selecting the months for which the CSSI is beyond the upper decile (12 beyond the 120 months of the 30-year periods). These "high CSSI composites" are helpful to determine the large-scale and/or local processes associated with WECS. The significance of composites are determined by using a Student *t* test. We use monthly data for this analysis because of the lack of daily data in the CMIP5 database for some variables and/or models. By using models for which both daily and monthly data are available, we verified that the composites were not too dependent on the use of monthly instead of daily data. In order to test the sensitivity of our results to the index choice, we compared the CSSI to the commonly-used cold spell duration index (CSDI, Alexander et al. 2006). The CSDI is defined as the annual count of days from cold events with Tmin < Q10 during at least 6 days. The CSSI definition is close to the CSDI but integrates additional information about the intensity and severity of WECS. Overall, we find that inter-annual time series of these two indices are highly correlated over 1979–2008 (R = 0.98), which suggests that our results do not depend on the index choice in a significant manner. We thus decided to focus on the more-integrative CSSI in this paper.

3 Cold spells in present-day climate

The WECS statistics (see Sect. 2.3) for observations and each of the HIST/AMIP simulations are reported in Table 1 and Fig. 2b, c where model values are represented as relative departures (in %) from observations. Figure 2a also compares frequencies of WECS in HIST and AMIP simulations against EOBS observations.

3.1 WECS statistics in HIST

HIST frequencies are in good agreement with observations for all models (differences lower than 2 %, which corresponds to about 2.5 wintertime days, Fig. 2a). The biases on severity are strongly model dependent (Fig. 2b). Their significance is low (only significant at the 90 % confidence level for IPSL and INM according to a Student *t* test) due to the low significance of duration and extent biases compared to intensity ones. Intensity (red crosses) is the main contributor to severity biases. Duration (green squares) and extent (blue circles) are of secondary importance, but in some cases they can strengthen or cancel the intensity errors. Most of the models overestimate the duration of WECS, while their spatial extent is generally underestimated. MRI is the model that exhibits the better agreement with the observed statistics (except for frequency, Table 1).

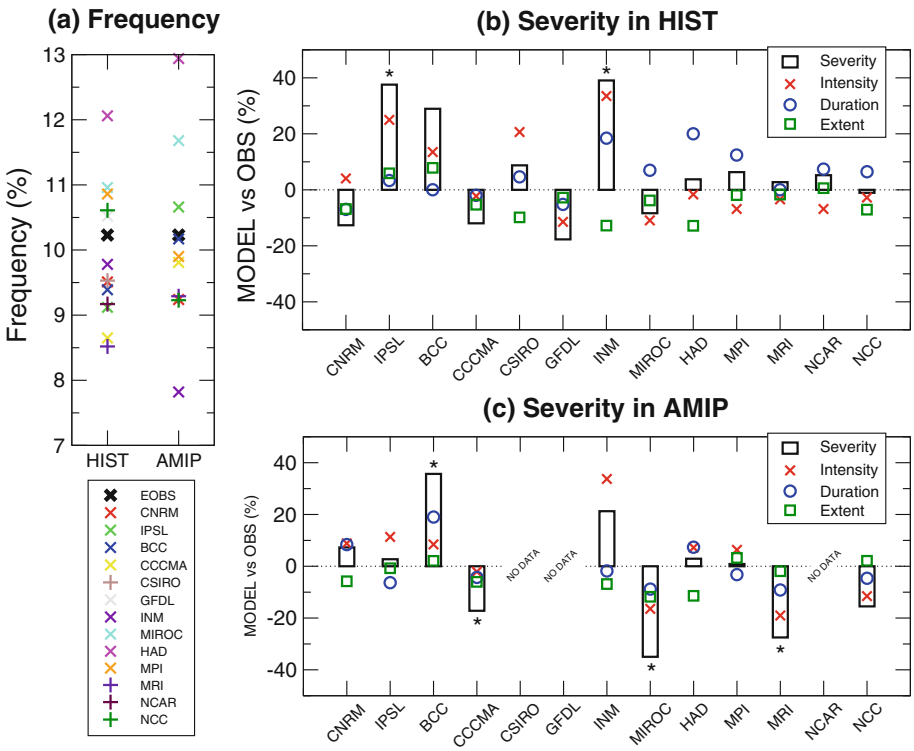

Fig. 2 **a** Frequency of cold spell days in winter (DJFM) for EOBS and CMIP5 models; **b** Biases of WECS statistics for historical simulations, expressed as the departure from EOBS (%). Stars indicate the 90 % confidence level for severity differences based on a Student *t* test; **c** Same as b but for AMIP simulations. Period: 1979–2008

3.2 WECS statistics in AMIP

The same analyses than in previous section are shown for AMIP simulations in Fig. 2c. Note that the daily Tmin from AMIP was not available for CSIRO, GFDL and NCAR at the time of the study. For some models, the biases are quite different from those found for HIST, suggesting the importance of SST and ocean–atmosphere coupling on WECS representation. This is especially true for MRI and IPSL which simulate much less severe WECS than in HIST. For IPSL, this is an improvement while for MRI it leads to significantly underestimate the WECS severity. Interestingly, these models also have the most pronounced cold bias of SST over the North Atlantic basin in HIST (−4.2° for IPSL and −4° for MRI between 35°N and 60°N). Cattiaux et al. (2011) have shown that North Atlantic SST can play a role on wintertime temperature extremes over Europe, through the heat and water vapor advection by the westerlies, which alter both turbulent fluxes and downward longwave radiation. Given their cold SST biases, these effects are probably underestimated in the IPSL and MRI coupled models so that the WECS severity decreases when observed SST are prescribed in the atmospheric GCMs.

3.3 Geographical distribution of Tmin anomalies during WECS

The results presented in Fig. 2 are representative of the whole WE domain. To get an idea of the spatial distribution of WECS anomalies in observations and models, Fig. 3a (EOBS) and b (multi-model bias) show composites of Tmin derived from the monthly CSSI (see Sect. 2.4 for details about methodology). In observations, the cold anomaly associated with WECS is maximum in the north-east of the WE domain and spreads over Northern Europe (Fig. 3a). In models, a north–south dipole is found on Tmin biases during WECS, with an underestimation (overestimation) of WECS intensity over Northern Europe (Southern Europe) (Fig. 3b). These anomalies are in good agreement among models (high sign consistency), suggesting that the same processes may be responsible for individual model biases.

3.4 Analysis of large scale dynamics

In this section, we investigate the potential role of large-scale dynamics on WECS biases. Figure 3c, d show high CSSI composites of SLP to identify the large-scale circulation anomalies associated with the WECS over Europe and the corresponding model biases. In the NCEP2 reanalysis, they are associated with a high pressure centred over Iceland and a low pressure over the Mediterranean region. This pattern is reminiscent of the NAO- pattern and limits the advection of oceanic flow by westerly winds, causing strong cooling over Western/Northern Europe. Actually, this monthly composite is a mix of two predominant synoptic weather regimes associated with cold temperatures over Europe: the negative phase of the NAO

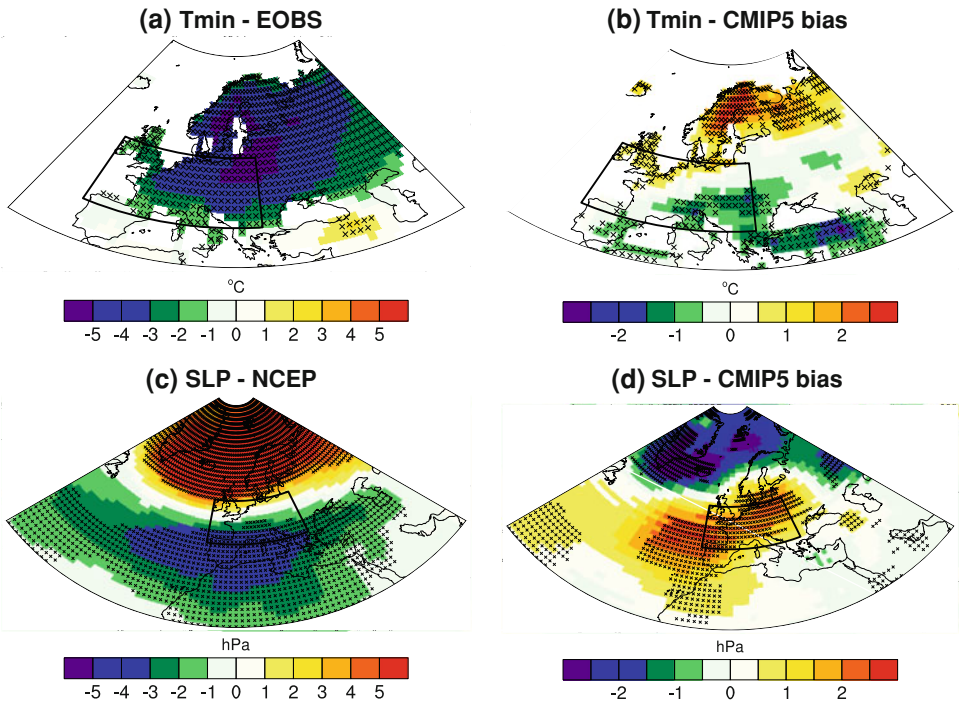

Fig. 3 **a** High CSSI composites of Tmin anomaly in observations; **b** Multi-model bias for historical simulations (HIST-EOBS); **c**, **d** same as **a**, **b** for SLP. On **a**, **c**, stipples indicate regions where composites are significant to the 95 % confidence level. On **b**, **d**, stipples indicate regions where the sign consistency between models is higher than 80 %. Period: 1979–2008

and the Scandinavian blocking (Cassou 2008). Figure 3d suggests that the amplitude of these large-scale circulation patterns is underestimated by the multi-model, which could explain the underestimated intensity of WECS over Northern Europe. This result is consistent with the North Atlantic weather regime analysis of Cattiaux et al. (2012), which however indicates that large scale dynamics only explains a limited fraction of European temperature biases in CMIP5 models.

3.5 Analysis of the surface radiative budget

Besides large-scale circulation, local thermodynamic and radiative processes can also contribute to errors in mean and/or extreme temperatures. To assess the influence of radiative biases, Fig. 4 shows high CSSI composites of surface net radiation (Rnet) and surface cloud radiative forcing (CRF) for observations and multi-model bias. SRB observations show positive Rnet anomalies during WECS (Fig. 4a) resulting from reduced upward longwave radiation associated with cold surface temperature during WECS (not shown). Cloud cover anomalies during cold spells also contribute positively to the surface energy balance as shown by the increased CRF (Fig. 4c). The latter effect mainly arises from solar radiation (Fig. 4e) due to a weaker total cloud cover in the north of the WE domain (not shown). These results indicate that dynamically driven WECS are tempered by both longwave and shortwave negative feedbacks according to the satellite observations.

These feedbacks are not well captured by the CMIP5 multi-model which shows strong biases in the high CSSI composites of Rnet, indicating a wrong sign of the Rnet signal during WECS (Fig. 4b). This deficiency is consistent with a too weak CRF (Fig. 4d) that is mainly related to the longwave (not shown) rather than shortwave (Fig. 4f) cloud radiative effect. This radiative bias helps to maintain the cold anomaly at the surface and suggests that the greenhouse effect of high-level clouds is underestimated by the models during WECS. To go further, it would be interesting to determine whether this bias comes from a lack of high-level clouds in models. However, this diagnosis is not available in the CMIP5 database for simulations of interest. Besides clouds and their radiative properties, some CMIP5 models also show shortwave biases that are related to snow cover and/or snow albedo biases during WECS (not shown). However, there is no agreement between the models for the signals associated with surface conditions.

In summary, and despite a large inter-model diversity, we found some common biases in large-scale dynamics and surface radiation among our subset of CMIP5 models in some regions. Over Northern Europe and the northern part of our WE domain (mainly British Islands), dynamical biases are predominant and lead to an underestimation of WECS intensity. In the rest of our WE domain, radiative biases take over dynamics biases and the underestimation of negative cloud feedbacks leads to overestimate the WECS intensity.

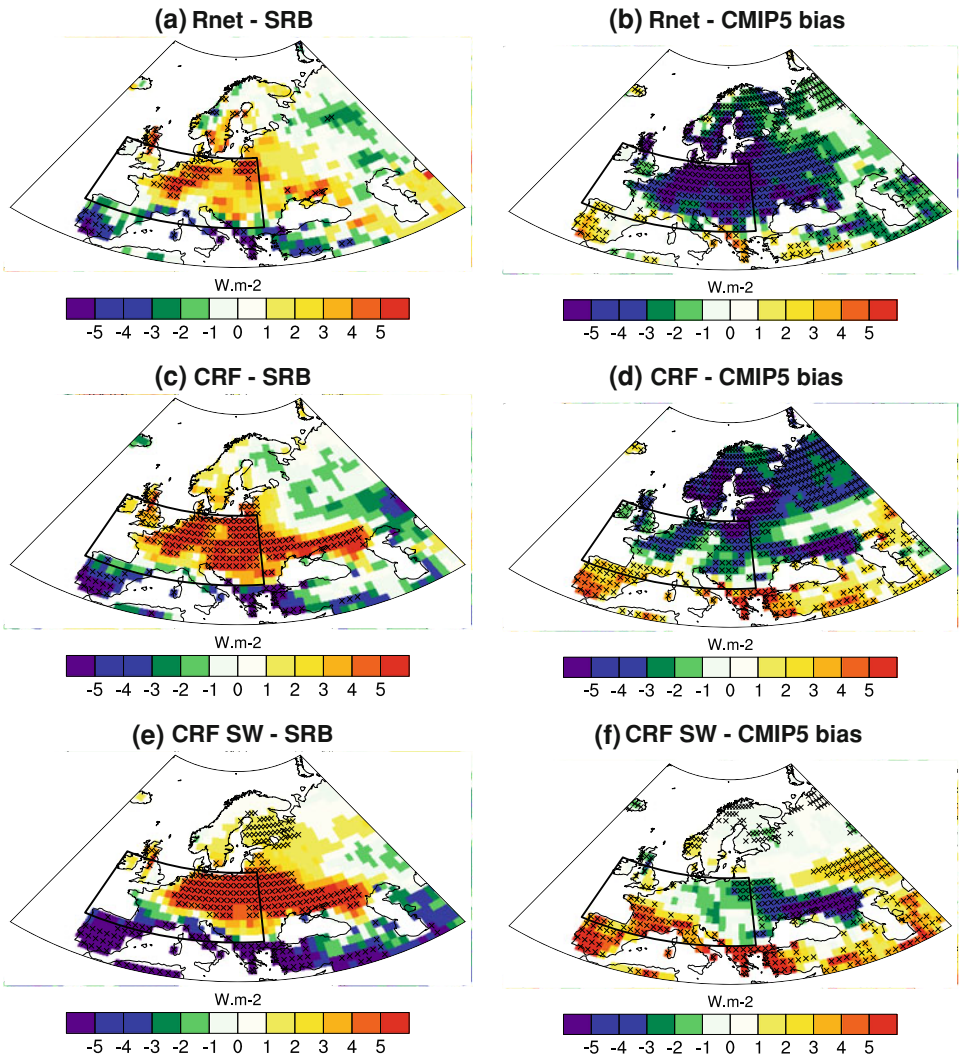

Fig. 4 **a** High CSSI composites of net radiative flux anomaly in observations; **b** Multi-model bias for historical simulations (HIST-EOBS); **c**, **d** same as **a**, **b** for cloud radiative forcing; **e**, **f** same as **a**, **b** for the shortwave component of CRF. On **a**, **c** and **e**, stipples indicate regions where composites are significant to the 95 % confidence level. On **b**, **d** and **f**, stipples indicate regions where the sign consistency between models is higher than 80 %. Period: 1984–2007 (SRB period)

4 Cold spells in RCP8.5 scenarios

Figure 5 depicts the same statistics as in Fig. 2 but for RCP8.5, with two different Q10 thresholds used to identify WECS. First, Q10 from HIST (Q10P) is used to describe how characteristics of "present-day" WECS will change over the twenty-first century. Secondly, Q10 from RCP8.5 is used (Q10F) to characterize "future" WECS based on a shifted Tmin distribution due to climate change. In Fig. 5b, c, statistics are expressed as relative differences (in %) from HIST (2070–2099 minus 1979–2008).

4.1 WECS statistics with Q10P threshold

As expected, "present-day" cold spells (hereafter referred to as Q10P WECS) are much less frequent in RCP8.5 for all models (Fig. 5a). Frequencies are less than 2 % of wintertime days, with the most important change found in CSIRO (none Q10P WECS at the end of the century). This result is consistent with the simulated global warming and with previous studies on the subject (Vavrus et al. 2006; Kodra et al. 2011; Räisänen and Ylhäisi 2011). An interesting point is to determine whether these rare events would be as severe as present-day ones. Statistics of severity are shown in Fig. 5b for Q10P WECS. In all except one model (GFDL), the severity of Q10P WECS is strongly reduced in RCP8.5 (from −10 % to almost −75 % of the HIST severity). This is mainly related to a shortening of WECS events (blue circles) and to a lesser extent to the reduction of intensity (red crosses). Response in spatial extent is more model-dependent with either an increase or a decrease (green squares). These differences in severity are significant at the 90 % confidence level for 8 models among 12.

4.2 WECS statistics with Q10F threshold

Figure 5c shows the same statistics for Q10F WECS. By design, frequencies are now close to 10 % of the wintertime days. As expected from the mean temperature

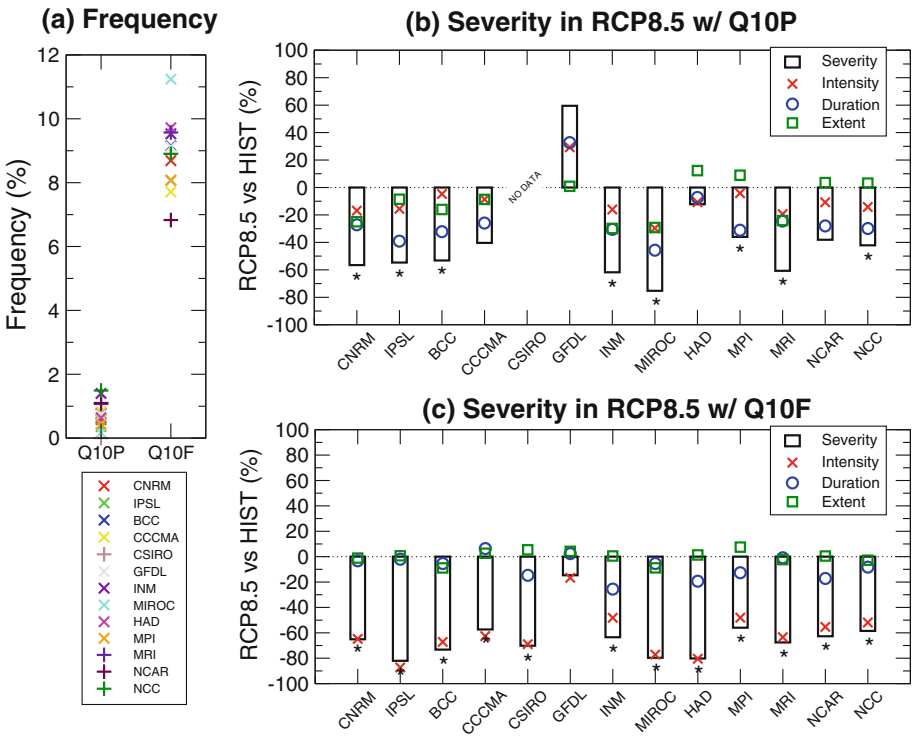

Fig. 5 **a** Frequency of cold spell days in winter (DJFM) for Q10P and Q10F WECS; **b** Response of WECS statistics for Q10P in RCP8.5 simulations, expressed as the departure from HIST (%). Stars indicates the 90 % confidence level for severity differences based on a Student *t* test; **c** Same as b but for the Q10F threshold. Period: 2070–2099

increase by the end of the twenty-first century, severity decreases drastically for all models except GFDL (from −60 to −80 %). Though for some models some differences are found in the duration and extent parameters, this decrease is mainly due to a weaker intensity of WECS. In line with the results of de Vries et al. (2012), the duration parameter does not exhibit any significant change in all models. The extent parameter is even less sensitive to global warming. Conversely, all differences (except for GFDL) are significant at the 90 % confidence level for both severity and intensity parameters. The weak sensitivity of GFDL, that is discussed in Sect. 4.4, is consistent with its relatively low change of mean temperature over Europe in RCP8.5 (+2.5° C in average over the WE domain, which represents the lower range of changes predicted by the models).

4.3 Multi-model response of large-scale circulation and surface radiation for Q10F WECS

To discuss the possible contribution of large-scale dynamics to changes in WECS (using the Q10F threshold in RCP8.5), Fig. 6 shows high CSSI composites of Tmin and SLP over North Atlantic and Europe. The decrease of the WECS intensity over the whole WE domain (Fig. 6b) is associated with a weakening of the NAO- pattern (Fig. 6d). Models are in good agreement over the Mediterranean Sea where they exhibit an anticyclonic anomaly that favors the advection of warm air from north Africa and south Atlantic over the WE domain. This fairly robust dynamical response tends to decrease WECS intensity at the end of the twenty-first century.

Figure 7 shows the same composite analysis for surface net radiation (Rnet), total and shortwave cloud radiative forcings (CRF and CRFSW), as well as clear-sky albedo (ALBCS). ALBCS is indeed a good indicator of the snow cover/albedo feedback, which is supposed to play a significant role in climate scenarios (Qu and Hall 2006; Levis et al. 2007). The Rnet response (Fig. 7b) is in sign agreement with the Tmin change depicted in Fig. 6b, with a strong increase in the amount of radiation absorbed at the land surface. Conversely, the CRF response (Fig. 7d) shows a weakening of the absorbed radiation during WECS that mitigates the surface warming. Over the northeastern part of the WE domain, this effect is mainly related to a decrease of the shortwave radiation (Fig. 7f) associated with an increase of the total cloud cover fraction (not shown). Over the southern part of the WE domain, the decrease in longwave radiation (not shown) is predominant and suggests a decrease of high-level clouds that limits the greenhouse effect. Thus, the response of cloud cover processes, related to changes not only in total cloud cover but also possibly in cloud properties and vertical distribution, has a negative contribution on the surface energy budget in average over the WE domain.

Concerning surface processes, Fig. 7h depicts a strong decrease in clear-sky surface albedo at the end of the RCP8.5 scenarios. It highlights the effect of the snow cover

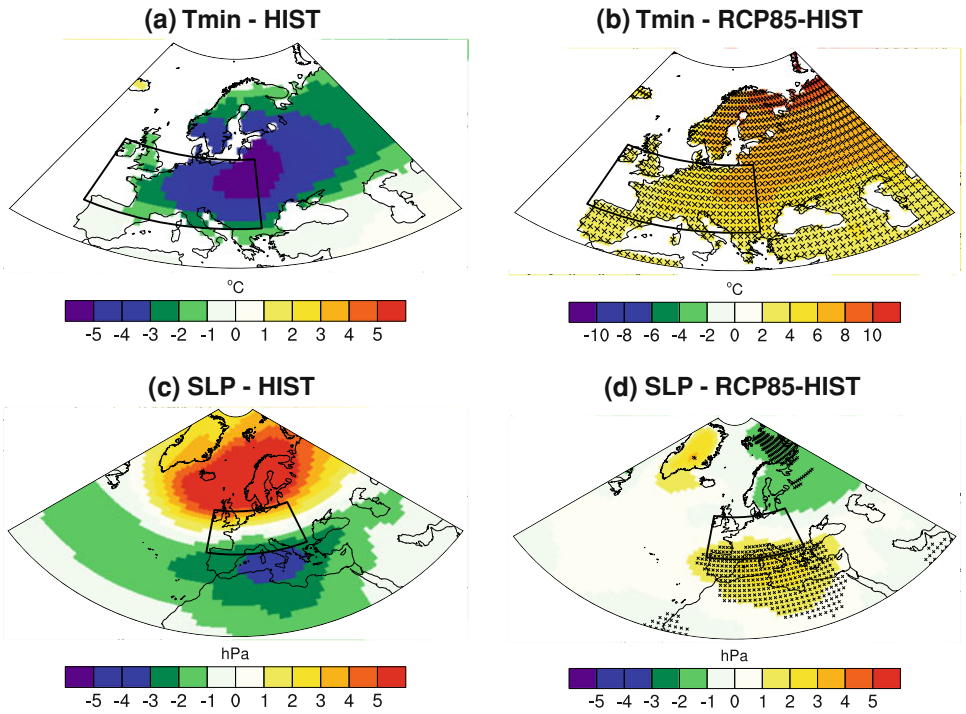

Fig. 6 **a** High CSSI composites of Tmin anomaly in HIST (multi-model mean); **b** Multi-model response in RCP8.5 (RCP8.5-HIST) by using Q10F; **c**, **d** same as a and b but for SLP. On **b**, **d**, stipples indicate regions where the sign consistency between models is higher than 80 %. Periods: 1979–2008 (HIST) and 2070–2099 (RCP8.5)

retreat in response to global warming, not only for seasonal mean temperatures but also during Q10F WECS. Indeed, models with available snow outputs (8 models) simulate on average 4 % of snow cover over the WE domain in RCP8.5, instead of 17 % in HIST. Thereby, unlike the multi-model CRF response found over Europe, the snow-albedo feedback locally enhances the surface warming over the WE domain as well as over Eastern Europe.

4.4 Focus on individual model responses

To quantify more precisely the cause for the decrease of WECS intensity in individual models, Fig. 8 gathers scatter plots of WECS intensities versus some climatic variables. The relationship between the WECS intensity and the mean change of Tmin over the domain is clear ($R = -0.84$, $p < 0.05$). Models showing the strongest decrease of WECS intensity are those with the strongest seasonal mean warming (Fig. 8a). While correlations with CRF and ALBCS changes are not significant (Fig. 8b, c, $R = 0.43$ and $R = 0.34$), the potential relationship between changes in ALBCS versus WECS intensity is broken by two "outliers". Indeed, when GFDL and IPSL are removed from the list of models, the correlation between the intensity response and the decrease of ALBCS becomes significant ($R = 0.69$, $p < 0.05$). Interestingly, GFDL (IPSL) is among the models showing the weakest (strongest) response in both WECS intensity and mean temperature (Fig. 8a).

Possible causes for these extreme responses of GFDL and IPSL are suggested in Fig. 9 with SLP and SST composites computed as in Fig. 7. In line with the sustained WECS intensity, the GFDL atmospheric pattern favors polar air advection over Western Europe (Fig. 9a) with positive SLP anomalies over Greenland and negative anomalies over Scandinavia. This response of large-scale dynamics is not found in the other models and is associated with a particular SST pattern (Fig. 9c). Strikingly, a cold anomaly is found southward of Greenland and Iceland coasts. Such a regional exception to global warming was highlighted in several studies and could arise from a slowdown of the North Atlantic thermohaline circulation that mitigates the warming over Western Europe (Vellinga and Wood 2002; Dai et al. 2005). Since GFDL is the model with the strongest cold anomaly of SST over this region, that might explain its limited change in WECS intensity in the RCP8.5 scenario.

IPSL exhibits a very different SLP pattern (Fig. 9b) with a strong anticyclonic anomaly centered over the Mediterranean Sea. This anomalous circulation tends to favor the advection of warm air from North Africa over the WE domain and provides an explanation for the strong decrease of WECS intensity in this model. Again, the pattern of the SST warming could contribute to this atmospheric response. Indeed, the Mediterranean SST warms strongly in IPSL, as illustrated in Fig. 9d (around +4 K between RCP8.5 and HIST). Moreover, as reported in Fig. 8d, a fairly close relationship is found between the decrease in WECS intensity and the rise of Mediterranean SST in RCP8.5 ($R = -0.67$, $p < 0.05$). Further analyses would however be required to understand whether the

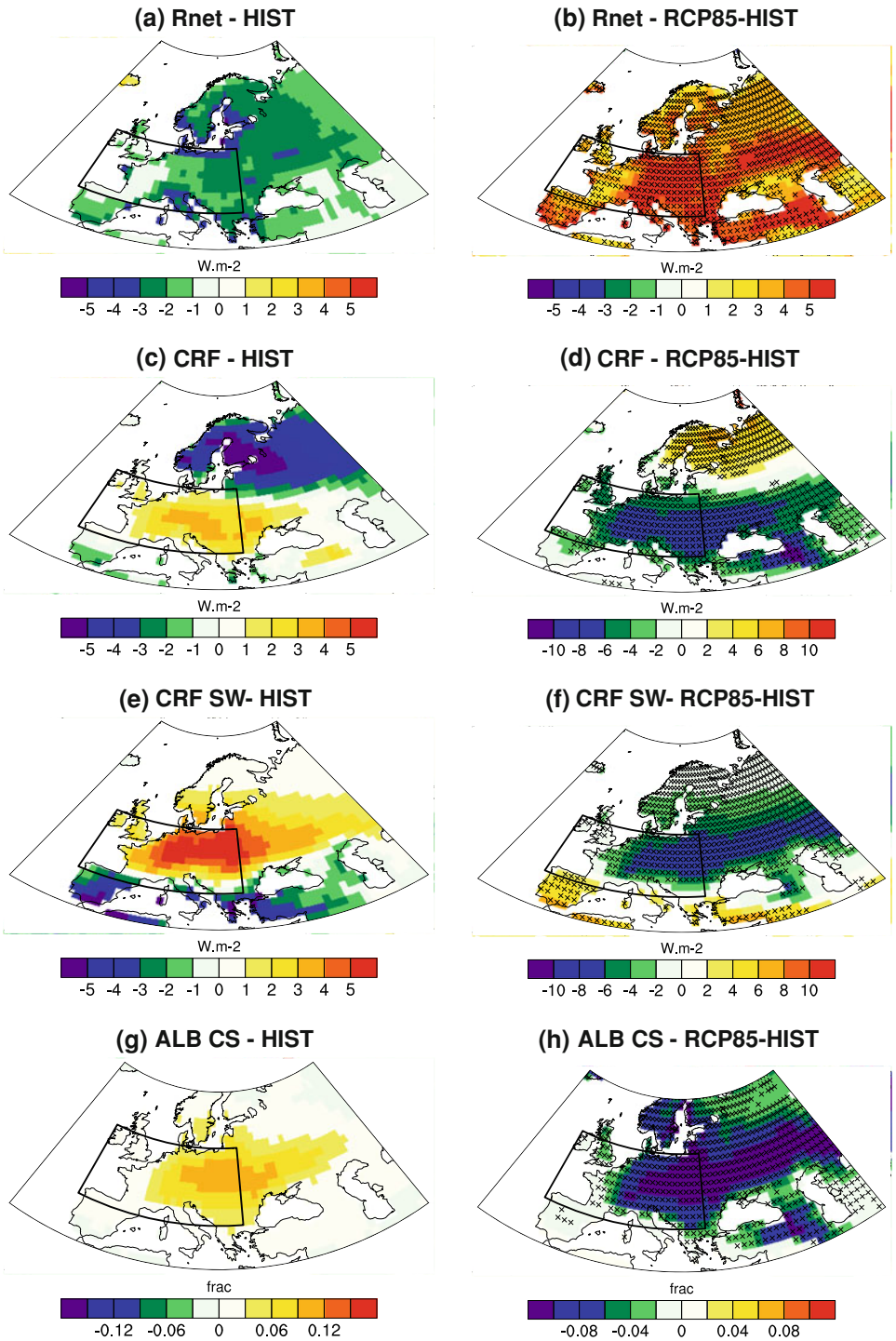

Fig. 7 **a** High CSSI composites of net radiative flux anomaly in HIST (multi-model mean); **b** Multi-model response in RCP8.5 (RCP8.5-HIST) by using Q10F; **c**, **d** same as a and b but for cloud radiative forcing; **e**, **f** same as a and b but for the shortwave component of CRF; **g**, **h** same as a and b but for clear-sky albedo. On **b**, **d**, **f** and **h**, stipples indicate regions where the sign consistency between models is higher than 80 %. Periods: 1984–2007 (HIST) and 2070–2099 (RCP8.5)

Mediterranean SST plays a passive or active role in the response of WECS to global warming in CMIP5 models.

5 Conclusion

In this paper, an original methodology is developed to define cold spells over Western Europe (WECS). We use a percentile-threshold index (10th percentile of Tmin, Q10) to retain cold days and some duration and extent criteria to define the regional cold spells (respectively 6 days and 20 % of the WE domain). Based on this index, WECS statistics in historical simulations of 13 CMIP5 models have been compared to EOBS observations over the 1979–2008 period. While most models are able to simulate realistic WECS features, model-dependent as well as

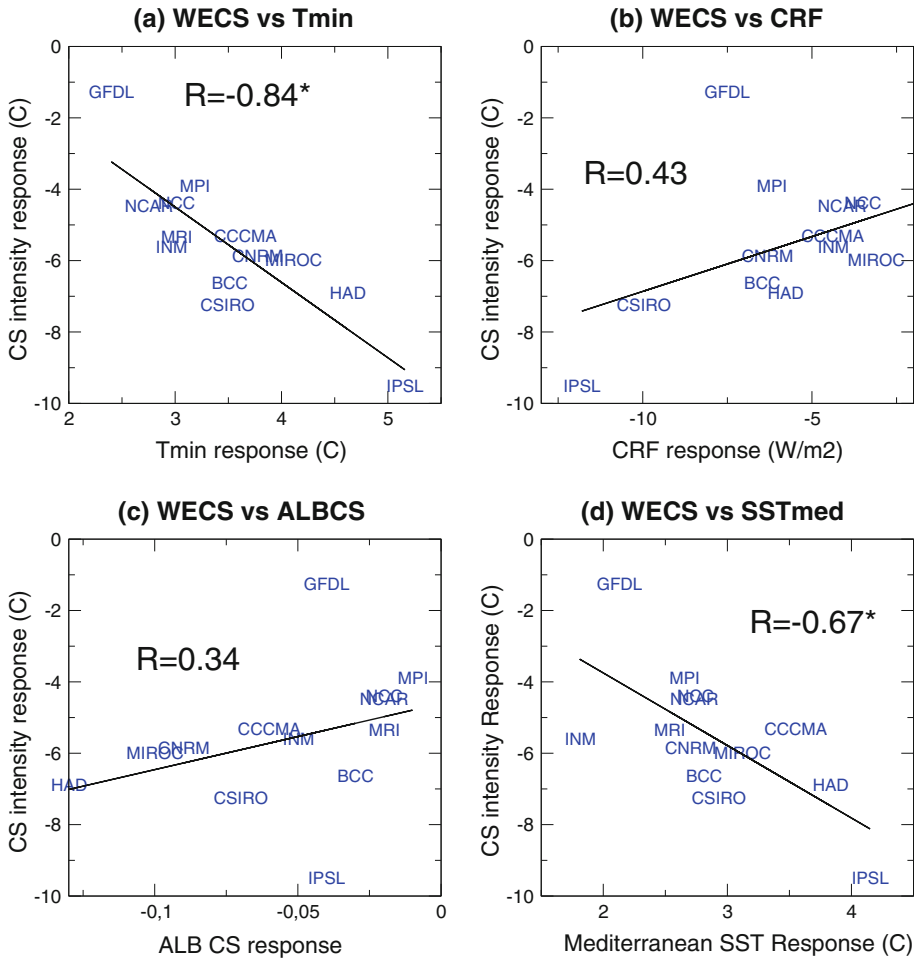

Fig. 8 Scatterplot of changes in WECS intensity versus seasonal mean changes (RCP8.5-HIST) in: **a** Tmin averaged over the WE domain; **b** same as a but for cloud radiative forcing; **c** same as a but for clear-sky albedo; **d** same as a but for Mediterranean SST. *Stars* indicate that correlation is significant to the 95 % confidence level

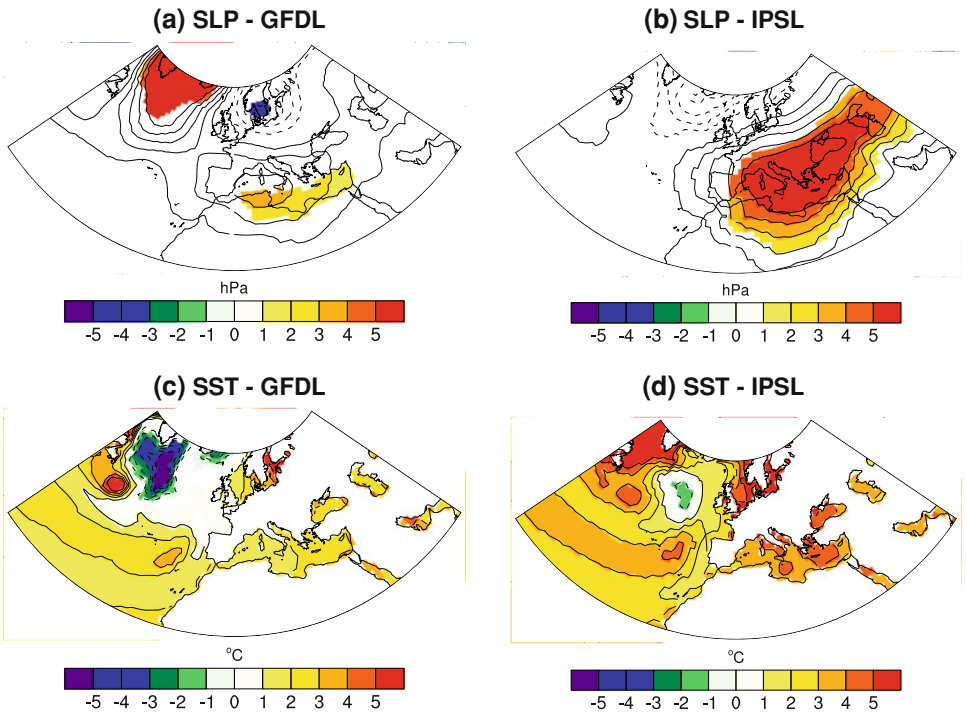

Fig. 9 **a** Model response (RCP8.5-HIST) of high CSSI composites of SLP anomaly for GFDL; **b** same as a but for IPSL. **c**, **d** same as a and b but for SST. Significant values at the 95 % confidence level are shaded. Period: 2070–2099

common biases are identified on the intensity, duration and extent of these events. Among these three parameters that determine WECS severity, intensity has the most important contribution on the model biases.

A preliminary analysis of the possible causes for such biases has highlighted the role of ocean–atmosphere coupling and SST errors, with some models simulating more realistic WECS in AMIP mode. Moreover, some models share common radiative biases that lead to underestimate the amount of energy absorbed at the land surface and overestimate the magnitude of Tmin anomalies during WECS. However, this cold bias is offset over Northern Europe by the dynamics bias that cause an overestimation of oceanic air advection.

The evolution of WECS characteristics at the end of the twenty-first century (2070–2099 period) is then documented using RCP8.5 concentration scenarios. Not surprisingly, WECS from the current period (computed with the present 10th percentile, Q10P) are projected to be much less frequent by all models. Nonetheless, these "present-day" WECS can be more severe as in the late twentieth century in one model, in line with previous results on cold extremes (Vavrus et al. 2006; Kodra et al. 2011). WECS defined from the future 10th percentile threshold (Q10F) show a strong decrease of intensity that scales with the increase of the mean temperature over Europe between present and future. While others scenarios have not been investigated in this study, this linear relationship suggests that in a more moderate scenario (e.g. RCP4.5) the WECS decrease would be weaker in line with a weaker rise of mean temperature. For multi-model ensemble mean, and in contrast with biases, changes in large-scale dynamics and surface radiation contribute the same sign to the response of WECS intensity. The negative contribution of cloud radiative forcing is offset by a decrease of surface albedo, in line with the snow cover retreat simulated in RCP8.5. However, some models (GFDL and IPSL) exhibit particular atmospheric patterns that shape the simulated decrease of WECS intensity. Such patterns could be related to the specific SST response found in the North Atlantic and Mediterranean Sea but further studies will be necessary to clarify this question. Besides, the question of trends of WECS has not been addressed in this paper and would deserve a dedicated work using appropriate statistical analysis over longer periods than the 30-year time slices used in the present study.

In summary, and despite some robust features of the WECS response found in the CMIP5 multi-model, large model uncertainties remain about the frequency and severity of such events at the end of the twenty-first century. Further studies and model improvements are needed for better constraining both regional processes (cloud and snow radiative feedbacks) as well as large-scale dynamics and coupled ocean–atmosphere processes (SST patterns) in global climate models. The presence of potential outliers in this respect also raises the issue of defining objective measures of model performance and combining projections from multiple models.

Acknowledgments The authors are grateful to S. Tyteca at CNRM for helpful CMIP5 data download. They also thank modeling groups, for producing and making available their model outputs, and the WCRP's Working Group on Coupled Modeling responsible for CMIP. Finally, thanks to the two anonymous reviewers for their helpful comments. This work is supported by the SECIF project from the French National Research Agency (ANR) and by the FP7 EUCLIPSE project (grant agreement #244067).

References

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Ambenje P, Rupa Kumar, Revadekar J, Griffiths G (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res 111. doi:10.1029/2005JD006290

Ballester J, Giorgi F, Rodó J (2010) Changes in European temperature extremes can be predicted from changes in PDF central statistics: a letter. Clim Change 98:277–284

Brown SJ, Caesar J, Ferro AT (2008) Global changes in daily extreme temperatures since 1950. J Geophys Res 113:D05115

Cassou C (2008) Intraseasonal interaction between the Madden–Julian oscillation and the North Atlantic oscillation. Nature 455(7212):523–527. doi:10.1038/nature07286

Cattiaux J, Vautard R, Cassou C, Yiou P, Masson-Delmotte V, Codron F (2010) Winter 2010 in Europe: a cold extreme in a warming climate. Geophys Res Lett 37:L20704. doi:10.1029/2010GL044613

Cattiaux J, Vautard R, Yiou P (2011) North-Atlantic SST amplied recent wintertime European land temperature extremes and trends. Clim Dyn 36(11–12):2113–2128. doi:10.1007/s00382-010-0869-0

Cattiaux J, Douville H, Peings Y (2012) European temperatures in CMIP5: origins of present-day biases and future uncertainties. Submitted to Clim. Dyn

Christidis N, Stott P, Brown S, Hegerl G, Caesar J (2005) Detection of changes in temperature extremes during the second half of the 20th century. Geophys Res Lett 32(20). doi: 10.1029/2005GL023885

Dai A, Hu A, Meehl GA, Washington WM, Strand WG (2005) Atlantic *thermohaline* circulation in a coupled general circulation model: unforced variations versus forced changes. J Clim 18:3270–3293

de Vries H, Haarsma RJ, Hazeleger W (2012) Western European cold spells in current and future climate. Geophys Res Lett 39:L04706. doi:10.1029/2011GL050665

Easterling DR, Evans JL, Ya P, Groisman et al (2000) Observed variability and trends in extreme climate events: a brief review. Bull Am Meteor Soc 81(3)

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. Clim. Res. 19:193–212

Guirguis K, Gershunov A, Schwartz R, Bennett S (2011) Recent warm and cold daily winter temperature extremes in the Northern Hemisphere. Geophys Res Lett 38:L17701. doi:10.1029/2011GL048762

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J Geophys Res 113:D20119. doi:10.1029/2008JD010201

Hegerl GC, Zwiers FW, Stott PA, Kharin VV (2004) Detectability of anthropogenic changes in annual temperature and precipitation extremes. J Clim 17:3683–3700

Heino R et al (1999) Progress in the study of climate extremes in northern and central Europe. Clim Change 42:151–181

Huynen MM, Martens P, Schram D, Weijenberg MP, Kunst AE (2001) The impact of heat waves and cold spells on mortality rates in the Dutch population. Environ Health Perspect 109:463–470

IPCC (2007) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC. In: Solomon S, Qin, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Cambridge University Press, Cambridge, p 996

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NMC/NCAR 40-Year reanalysis project. Bull Am Meteor Soc 77:437–471

Kodra E, Steinhaeuser K, Ganguly AR (2011) Persisting cold extremes under 21st-century warming scenarios. Geophys Res Lett 38:L08705. doi:10.1029/2011GL047103

Levis S, Bonan GB, Lawrence PJ (2007) Present-day springtime high-latitude surface albedo as a predictor of simulated climate sensitivity. Geophys Res Lett 34:L17703. doi:10.1029/2007GL030775

Mc Michael AJ, Campbell-Lendrum DH, Corvalan CF, Ebi KL, Githelo A, Scheraga JD, Woodward A (2003) Climate change and human health: risks and responses. World Health Organ, Geneva

McGregor GR, Ferro CA, Stephenson DB (2005) Projected changes in extreme weather and climate events in Europe. Extreme Weather Clim Events Public Health Responses Part 1:13–23. doi:10.1007/3-540-28862-7_2

Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305:994–997

Meehl GA, Tebaldi C, Nychka D (2004) Changes in frost days in simulations of twenty-first century climate. Clim Dyn 23:495–511. doi:10.1007/s00382-004-0442-9

Overland JE, Wang M (2010) Large-scale atmospheric circulation changes associated with the recent loss of Arctic sea ice. Tellus 62A:1–9

Petoukhov V, Semenov VA (2010) A link between reduced Barents-Kara sea ice and cold winter extremes over northern continents. J Geophys Res 115:D21111. doi:10.1029/2009JD013568

Qu X, Hall A (2006) Assessing snow albedo feedback in simulated climate change. J Clim 19: 2617–2630, doi:10.1175/JCLI3750

Räisänen J, Ylhäisi JS (2011) Cold months in a warming climate. Geophys Res Lett 38:L22704. doi:10.1029/2011GL049758

Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). J Clim 21:2283–2296

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GE (2006) Going to the extremes: an intercomparison of model-simulated historical and future changes in extreme events. Clim Change 79:185–211

Vavrus SJ, Walsh JE, Chapman WL, Portis D (2006) The behavior of extreme cold air outbreaks under greenhouse warming. Int J Climatol 26:1133–1147

Vellinga M, Wood RA (2002) The atmospheric response to a THC collapse: scaling relations for the Hadley circulation and the nonlinear response in a coupled climate model. Clim Change 54:251–267

Whitlock CH, Charlock TP, Staylor WF, Pinker RT, Laszlo I, Di Pasquale RC, Ritchey NA (1993) WCRP surface radiation budget shortwave data product description—Version 1.1. NASA Technical Memorandum 107747, National Technical Information Service, Springfield, Virginia

Zhang X, Alexander L, Hegerl GC, Jones PD, Klein-Tank A, Peterson TC, Trewin B, Zwiers F (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. WIREs Clim Change. doi:10.1002/wcc.147