

[Click here to view linked References](#)

## **Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe**

Chloé Prodhomme<sup>1</sup>, Francisco Doblas-Reyes<sup>1,2,3</sup>, Omar Bellprat<sup>1,3</sup>, Emanuel Dutra<sup>4</sup>

<sup>1</sup> Catalan Institute of Climate Sciences (IC3), Barcelona, Spain

<sup>2</sup> Catalan Institute for Research and Advanced Studies (ICREA), Barcelona, Spain

<sup>3</sup> Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS),  
Barcelona, Spain

<sup>4</sup> European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK

Submitted to Climate Dynamics

Corresponding author: Chloé Prodhomme, [chloe.prodhomme@ic3.cat](mailto:chloe.prodhomme@ic3.cat)

25th October 2015



1 **Abstract**

2 *Land surfaces and soil conditions are key sources of climate predictability at the*  
3 *seasonal time scale. In order to estimate how the initialization of the land surface affects*  
4 *the predictability at seasonal time scale, we run two sets of seasonal hindcasts with the*  
5 *general circulation model EC-Earth2.3. The initialization of those hindcasts is done*  
6 *either with climatological or realistic land initialization in May using the ERA-Land re-*  
7 *analysis. Results show significant improvements in the initialized run occurring up to the*  
8 *last forecast month. The prediction of near-surface summer temperatures and*  
9 *precipitation at the global scale and over Europe are improved, as well as the warm*  
10 *extremes prediction. As an illustration, we show that the 2010 Russian heat wave is*  
11 *only predicted when soil moisture is initialized. No significant improvement is found for*  
12 *the retrospective prediction of the 2003 European heat wave, suggesting this event to*  
13 *be mainly large-scale driven. Thus, we confirm that late-spring soil moisture conditions*  
14 *can be decisive in triggering high-impact events in the following summer in Europe.*  
15 *Accordingly, accurate land-surface initial conditions are essential for seasonal*  
16 *predictions.*

## 17 1. Introduction

18 In the context of global warming and the associated increasing number of extreme events, such  
19 as heat waves, droughts and floods, the predictability at seasonal time scale of extreme  
20 temperature and precipitation events appears to be crucial for climate services, adaptation and  
21 risk management (Challinor et al. 2005; García-Morales and Dubus 2007; Thomson et al. 2006).  
22 The feasibility of seasonal prediction largely rests on the existence of slow, and predictable,  
23 variations in the ocean surface temperature, sea ice, soil moisture and snow cover, and how  
24 the atmosphere interacts and is affected by these boundary conditions (Shukla and Kinter  
25 2006). Ocean anomalies associated with El Niño - Southern Oscillation (ENSO) and other  
26 ocean phenomena, soil moisture, snow, and ice cover should be taken into account when  
27 initializing the predictions (Balmaseda et al. 2008; Balmaseda and Anderson 2009).  
28 Unfortunately, less information is available about the state of the climate components other than  
29 the atmosphere (Balmaseda et al. 2007; Saha et al. 2010). Due to this source of initial-condition  
30 uncertainty, but also other limitations as model inadequacy, and lack of appropriate  
31 computational resources, the ability to make predictions on time scales longer than two weeks is  
32 still limited (Palmer et al. 2005; Palmer et al. 2005b; Lee et al. 2011).

33 However, in the past years due to increase of the resolution (Fosser et al. 2014; MacLachlan et  
34 al. 2014), the development of better initialization products (Guemas et al. 2014; Balmaseda et  
35 al. 2009; Balsamo et al. 2015; Dee et al. 2009), and the improvement of model physics (Hourdin  
36 et al. 2013; Frenkel et al. 2012) the skill of climate predictions at seasonal and longer time  
37 scales has improved (Doblas-Reyes et al. 2013).

38 Despite the global improvement of seasonal prediction, our ability to forecast temperature and  
39 precipitation in some regions such as Europe remains relatively low. On one side, the  
40 pronounced warming trend since the 1980s is well captured by most of the seasonal  
41 retrospective forecasts over Europe, which provides significant skill for two meter temperatures  
42 (t2m, hereinafter) in this region (Doblas-Reyes et al. 2006). On the other side, the skill of  
43 predicting the variability around the warming trends is much lower (Weisheimer et al. 2011).  
44 This is mainly because the climate variability over Europe is controlled by a variety of  
45 mechanisms, such as, the North Atlantic Oscillation (NAO, Rogers 1997; Rodwell et al. 1999),  
46 the anomalous frequency of a set of weather regimes (Reinhold and Pierrehumbert 1982;  
47 Cassou et al. 2005; Wang et al. 2011), complex teleconnections with the Arctic (Cohen et al.  
48 2014) and with the tropics (Kutiel and Benaroch 2001; Shaman and Tziperman 2011; Behera et  
49 al. 2012), and the coupling between the atmosphere and the land surface (Fischer et al. 2007;  
50 Orsolini, Yvan and Kvamstø 2009; Wang et al. 2011). All these processes are not properly

51 represented in coupled models, which could explain the poor skill over Europe (Seneviratne et  
52 al. 2010; Kim et al. 2012; Scaife et al. 2011). In an early study, Schär et al. (1999) had shown  
53 the existence of a soil-precipitation feedback over Europe. Later on, soil has been shown to  
54 influence precipitations, temperature and extreme temperature over Europe; (Fischer et al.  
55 2007; Douville 2010; Seneviratne et al. 2006, 2010, 2013; Quesada et al. 2012; Bellprat et al.  
56 2013).

57 For instance, Seneviratne et al. (2010) described the soil moisture-temperature coupling  
58 feedback loop in which, when an anticyclonic anomaly is present over Europe the soil moisture  
59 content will either amplify or moderate the surface temperature response. If the soil is moist  
60 (energy limited regime) the available surface energy will preferentially dissipate into latent heat  
61 fluxes and dampen surface heating. Conversely, when the soil is dry (soil moisture limited  
62 regime) more energy is available for sensible heating, inducing an increase of near-surface air  
63 temperature (Seneviratne et al. 2010; Hirschi et al. 2011).

64 As soil moisture partly controls the occurrence of warm events over Europe, a correct  
65 initialization of soil moisture content might be essential to correctly forecast summer extreme  
66 temperatures. This problem was studied by the global land-atmosphere coupling experiment  
67 (GLACE) intercomparison project (<http://gmao.gsfc.nasa.gov/research/GLACE>). The first phase  
68 (GLACE-1) focused on predictability that arises from soil moisture anomalies and determined  
69 the geographical regions where soil moisture exerts a significant influence on surface air  
70 temperature and precipitation (hot spots) of land-atmosphere coupling (Koster et al. 2004). The  
71 second phase (GLACE-2) focused on forecast quality, and assessed the impact of accurate  
72 soil-moisture initialization on actual skill using a multimodel approach (Koster et al. 2011). The  
73 multimodel mean in GLACE-2 indicates a significant soil-moisture contribution to surface  
74 temperature forecast skill in summer with forecast times of up to two months over North and  
75 South America (Koster et al. 2010, 2011). While Europe was not then found as a main region of  
76 improvement when soil moisture is initialized the GLACE project, numerous other studies have  
77 found an impact of soil moisture initialization in Europe (Douville 2010; van den Hurk et al. 2010;  
78 Materia et al. 2014).

79 In the present study, the predictability associated with soil moisture at seasonal time scales is  
80 revisited with a focus on Europe. The originality of the present study resides in three different  
81 aspects. First, the experiments described in this manuscript cover a long period of 30 years  
82 instead of the 10 used in GLACE-2. Second, the forecast time has been extended up to 4  
83 months, which is longer than most of the GLACE-2 experiments (Koster et al. 2004, 2010).  
84 Finally, the initialization of the soil moisture has been performed using the new ERA-Land

85 reanalysis (Balsamo et al. 2015), which is expected to provide a good and consistent estimate  
86 of soil-moisture initial conditions.

87 The paper is structured as follows. In section 2, the EC-Earth2.3 forecast system, the  
88 experimental set up and the forecast quality assessment methods, as well as the definition used  
89 for “mid-extreme” events, are described. Section 3 illustrates the impact of the soil initialization  
90 on temperature and precipitation skill at the global scale and for extremes over Europe. Section  
91 4 describes in detail the role of soil moisture for the two case studies of summer 2003 and  
92 summer 2010. Finally, section 5 offers a summary and the future prospects of the work.

## 93 **2. Model and data description**

### 94 **2.1 The EC-Earth2.3 forecast system**

95 The seasonal hindcast experiments are conducted using the EC-Earth2.3 forecast system  
96 (Hazeleger et al. 2011). EC-Earth2.3 consists of three model components, the Integrated  
97 Forecasting System (IFS) cycle 31r1 for the atmosphere, NEMO2 for the ocean and LIM2 for  
98 the sea ice. The model resolution chosen for the atmosphere is a spectral triangular truncation  
99 at a wavenumber 159 and for the computation of physical processes reduced Gaussian grid  
100 N80, which corresponds to a mesh resolution of around 120 km in the mid-latitudes, with 62  
101 layers in the vertical. EC-Earth uses the H-TESEL (TESEL for Tiled ECMWF Scheme for  
102 Surface Exchanges over Land) scheme for the land surface (van den Hurk et al. 2000), which  
103 includes an improved representation of hydrology over the TESEL scheme, in agreement with  
104 more recent IFS cycles (Balsamo et al. 2009). The model has four active soil layers extending to  
105 a depth of 2.89 meters, without considering capillary rise of groundwater or horizontal exchange  
106 of soil water. The oceanic component is NEMO (Madec 2008) using the ORCA1 horizontal  
107 resolution (which is  $1^\circ$  although with a highly irregular, tripolar grid) and 42 vertical levels. The  
108 LIM2 sea-ice model is coupled to the ocean (Fichefet and Maqueda 1997). All model  
109 components are coupled through the Ocean Atmosphere Sea Ice Soil version 3 (OASIS3;  
110 Valcke 2006) coupler.

### 111 **2.2 Experimental set up**

112 To assess the impact of a realistic land-surface initialization on sub-seasonal and seasonal  
113 forecasts two seasonal hindcast experiments have been performed. A 10-member, four-month  
114 long hindcast experiment has been performed over the period 1981-2010 with start dates the  
115 first of May of each year. The ocean, sea-ice and atmospheric components are initialized with  
116 ORAS4 (Balmaseda et al. 2013), IC3 sea-ice analysis (Guemas et al. 2014) and ERA-Interim

117 (Dee et al. 2011), respectively. In the INIT experiment the land surface is initialized with the soil  
118 moisture and temperature and snow from ERA-Land (Balsamo et al. 2015), which provides  
119 consistent land surface conditions to the forecast system since both share the same land-  
120 surface model version. The ensemble is constructed by using atmospheric singular vectors and  
121 the five ocean analyses available from ORAS4. The CLIM experiment initializes the land surface  
122 using the climatology of ERA-Land for the corresponding start date, this being the only  
123 difference between INIT and CLIM. With this set up, the impact of the land-surface initialization  
124 can be isolated from all the other factors that influence the forecast quality in climate  
125 forecasting.

### 126 **2.3 Forecast quality assessment**

127 The objective of the present study is to assess how the land-surface initialization affects  
128 different aspects of the forecast quality of summer precipitation and temperature, with a specific  
129 focus over Europe.

130 For 500-hPa geopotential height, t2m, precipitation and sea level pressure data from the ERA-  
131 Interim reanalysis have been used as reference (Dee et al. 2011). For precipitation, the 0 to 12  
132 hour forecasts have been used. For soil moisture the ERA-Land reanalysis product is used  
133 (Balsamo et al. 2015).

134 The skill has been estimated using the correlation of the ensemble mean and the mean  
135 anomaly spatial correlation coefficient (MACC, hereafter). We use the Student distribution with  
136 N degrees of freedom to estimate the significance level of correlation, N being the effective  
137 number of independent data calculated following the method of von Storch and Zwiers (2001).  
138 The significance of the difference between two correlations is estimated using the methodology  
139 of Steiger (1980), which takes into account the dependence from sharing the same observations  
140 in both correlation coefficients. . In addition, the two methods to assess the significance of  
141 correlation and the significance of the difference of two correlations takes into account the  
142 independent number of data, which is necessary given the serial correlation typical of the time  
143 series considered. As there is no standard method to assess the significance of the MACC and  
144 difference between two MACC, we estimated their significance with a bootstrap of 100 random  
145 drawings, following the methodology of Masson and Mimack (1992). The drawings are done  
146 over the members (random selection of the members with repetition) and over the space  
147 (bootstrap by square blocks over the considered region). The block size is estimated by  
148 estimating the independent number of data on the longitude and latitude dimensions.

149 As we need to assess the contribution of the trend to the skill, we have compared the correlation  
150 and the MACC calculated on “raw” and detrended data. The detrended values are the residual  
151 of the regression on the global mean two meter temperature (GMT, hereafter) of the concerned  
152 variables; the observations are regressed on the observed GMT and both experiments are  
153 regressed on their simulated GMT (van Oldenborgh et al. 2013).

154 All the verification, as well as part of the plotting, have been done using the version 2.1.1 of the  
155 R-based `s2dverification` package ([http://cran.r-  
156 project.org/web/packages/s2dverification/index.html](http://cran.r-project.org/web/packages/s2dverification/index.html)).

157 Contrary to the common evaluation in seasonal forecasting, where seasonal means of the  
158 variables are analyzed, the skill of daily extremes is also evaluated in this paper. To estimate  
159 the daily extremes, we follow the same methodology as Pepler et al. (submitted), which was  
160 inspired by Hamilton et al. (2012) and the CECILIA EU project definitions ([http://www.cecilia-  
161 eu.org/index.htm](http://www.cecilia-eu.org/index.htm)). The extremes have been calculated using  $T_x$  and  $T_n$ , the daily maximum and  
162 minimum temperature, respectively, estimated from the 6 hourly  $t_2m$ .

163 The first set of extremes are the monthly 90th and 10th percentile of  $T_n$  and the 90th percentile  
164 of  $T_x$ , named hereafter  $q_{10}$  and  $q_{90}$  of  $T_n$ , and  $q_{90}$  of  $T_x$ , respectively. For the second set of  
165 extremes, the climatological 90th and 10th percentile of  $T_x$  and  $T_n$  are estimated using data  
166 from all years between 1981 and 2010. This is done separately for the ERA-Interim data and for  
167 the hindcasts. The frequency of days and nights in a month over and under the corresponding  
168 climatological percentile are then estimated. To summarize, the present study will focus on six  
169 of these variables:

- 170 -  $q_{10}$  of  $T_n$  for each month and the percentage of nights in a month under the  
171 climatological value of the  $q_{10}$  of  $T_n$ , also called number of cold nights.
- 172 -  $q_{90}$  of  $T_n$  for each month and the associated number of nights in a month over the  
173 climatological value of the  $q_{90}$  of  $T_n$ , also called number of warm nights.
- 174 -  $q_{90}$  of  $T_x$  for each month and the associated number of days in a month over the  
175 climatological value of the  $q_{90}$  of  $T_n$ , also called number of warm days.

176 The two first variables,  $q_{10}$  of  $T_n$  and the number of cold nights, correspond to cold extremes  
177 while the other four variables are related to warm extremes.

## 178 3. Results

### 179 3.1 Impact of land-surface initialization during boreal summer

180 Figure 1 illustrates the skill of the EC-Earth2.3 system for predicting land  $t_2m$  and precipitation  
181 using the correlation between the ensemble-mean prediction and the observational reference.



182 The results of the CLIM experiment are used as a benchmark. As in most state-of-the-art  
183 forecast systems, EC-Earth2.3 shows high skill for t2m over land almost everywhere except  
184 over some areas where the observational reference might not be trustworthy (Doblas-Reyes et  
185 al. 2013). Statistically significant correlations appear mainly in tropical regions. In contrast, the  
186 predictions exhibit lower skill for precipitation, except over a few regions such as those  
187 neighboring the Pacific basin and sub-Saharan Africa. An important part of the skill in both  
188 temperature and precipitation is linked to ENSO (Landman and Beraki 2012; Phelps et al. 2004;  
189 Doblas-Reyes et al. 2013) whose teleconnections over land is well reproduced by the model in  
190 most of the relevant areas (Fig. S1).

191 The use of a realistic initialization of soil variables (snow, soil moisture and soil temperature)  
192 such as the one used in the INIT experiment compared to the one used in CLIM has generally a  
193 positive impact on the skill of seasonal mean t2m (Fig. 1c). Nevertheless, only very few of the  
194 positive changes are statistically significant at the 95% confidence level (black dots), which is  
195 the likely result of the small differences and the reduced sample size of the experiment, an  
196 aspect that is limited by the observational data available to reliably initialize the hindcasts. The  
197 impact of land-surface initialization on the precipitation skill is patchy, although with a tendency  
198 to show positive differences in correlation. There is no area with a significant decrease of  
199 correlation, whereas a few areas show an important increase of skill (Fig. 1d). The patterns of  
200 improvement cannot be simply described by a modification of the ENSO teleconnections over  
201 land in INIT compared to CLIM (Fig. S1), because they are very similar in both experiments, and  
202 an alternative explanation is needed.

203 It has to be borne in mind that our study considers longer forecast time scales than the GLACE-  
204 2 experiment. For instance, no improvement in seasonal skill over the Great Plains emerges in  
205 Fig. 1 compared to previous studies. However, consistently with previous studies (Koster et al.  
206 2004, 2010, 2011; van den Hurk et al. 2010), there is an important improvement of skill in June  
207 (second forecast month) over the United States, which disappears in July and August (Figs. S2,  
208 S3, S4).

209 In order to quantify precisely the impact on skill seen on Fig. 1, Fig. 2 shows scatter plots of the  
210 difference of correlation between INIT and CLIM against the correlation in CLIM for both  
211 precipitation and t2m in different regions. Figure 2 shows the improvement due to the soil  
212 initialization for temperature prediction: 65.3% of the land points have a positive impact (Fig.  
213 2d). Nevertheless, the correlation difference between INIT and CLIM is significant only in very  
214 few cases (red dots), with no significant negative difference (dark blue dots). In general, in all  
215 regions, more improvements (positive differences, points where the skill is significant in INIT but

216 not in CLIM and statistically significant negative correlation decrease) than degradations  
217 (negative differences, points where the skill is significant in CLIM but not in INIT and statistically  
218 significant negative correlation decrease) are found. This comprehensive analysis shows that  
219 the land-surface initialization has on average a positive impact on the temperature skill when  
220 large regions are considered.

221 Conversely, for precipitation no clear improvements are visible on Fig. 2h: on one side 53.9% of  
222 the grid points have an increased skill in INIT. On the other side, more points are located in the  
223 bottom-right quadrant than in the top-right quadrant, which suggests that more degradations  
224 than improvements occur in the areas where CLIM has skill.

225 Figure 2 shows that, for both temperature and precipitation, the lower the skill in CLIM is, the  
226 stronger the improvements in INIT are. In addition, the degradation in INIT tends to occur when  
227 the skill is already positive in CLIM. This suggests that the land-surface initialization brings skill  
228 to regions where the forecast system has no skill, but it can also negatively perturb the system  
229 in regions of high skill, suggesting that the large-scale signal can be perturbed by soil moisture  
230 initialization. This can be partly explained by the biases in the soil initialization products  
231 (Balsamo et al. 2015) and by the initial shock and drift of the soil variables in the forecasts  
232 (Dirmeyer 2005; Materia et al. 2014). Furthermore, model inadequacies in the representation of  
233 the land and/or land atmosphere coupling might explain the decrease of skill in INIT. Error  
234 compensations may take place in CLIM, in other words, CLIM may have skill in some region for  
235 the wrong reasons. In this case, a better representation of the soil state might in some region  
236 lead to a decrease of skill. An illustration of possible error compensations can be seen over  
237 North-Western South America, where the relation between ENSO and t2m is reversed  
238 compared to the observed one (Fig. S1) while still CLIM has a high t2m skill in this region (Fig. 1  
239 and 3).

240 Another factor that can explain the difference in skill between CLIM and INIT is their  
241 representation of the recent temperature trends. In fact, recent trends can explain a large part of  
242 the seasonal forecast temperature skill (Doblas-Reyes et al. 2006). Figure 3 is similar to Fig. 1,  
243 but this time the correlation has been computed using the residuals of the regression of the  
244 temperature and precipitation fields on the GMT. The comparison of Figs. 1a and 3a illustrate  
245 the important contribution of the trend in the skill of temperature. An important part of the t2m  
246 skill is related to the trend, especially over Europe where most of the skill in CLIM is related to  
247 the trend (Doblas-Reyes et al. 2006, 2013). Conversely, Figs. 1 and 3 suggest that there is  
248 almost no impact on the skill of precipitation from the temperature trend.

249 For both precipitation and t2m, the impact of the land-surface initialization remains very similar  
250 when the effect of the global-mean temperature is removed (Figs. 1b, 3b). This result, and the  
251 inspection of the regression coefficients, suggests that the land-surface initialization affects only  
252 marginally the representation of the temperature trend, consistently with the results of Jaeger  
253 and Seneviratne (2010). The comparison between Figs. 1c and 3c (see also Fig. S5) gives a  
254 hint that the skill improvement in INIT compared to CLIM is slightly stronger when the trend is  
255 removed.

256 As most seasonal forecast systems, EC-Earth2.3 shows widespread skill in seasonal-mean t2m  
257 and relatively low skill for precipitation forecasts. An important part of the skill for forecasting  
258 t2m is linked to the warming trend. The soil moisture initialization leads to a general  
259 improvement of t2m skill and to a lesser extent of precipitation skill, occurring mainly in regions  
260 where the skill is low in the CLIM experiment. This improvement remains robust when the  
261 global-mean trend effect is removed. The rest of the paper focuses on Europe, a region where  
262 soil moisture has been shown to have a strong impact, an aspect that is also evidenced in our  
263 experiments (Figs. 1, 2; Jaeger and Seneviratne 2010; Hirschi et al. 2011; Quesada et al. 2012;  
264 Douville 2010).

### 265 **3.2 Summer skill over Europe**

266 The previous section showed that Europe is one of the regions with the largest impact of the  
267 land-surface initialization. However, all the results described concentrate on seasonal averages  
268 of temperature and precipitation. Instead, various studies have demonstrated that soil moisture  
269 plays an important role in the occurrence of extreme warm events (Jaeger and Seneviratne  
270 2010; Hirschi et al 2011; Hamilton et al. 2012). The prediction of extreme events is highly  
271 relevant to society (Wang et al. 2009). Hence, any skill improvements on this aspect might have  
272 a larger impact than the more traditional result of the increase in seasonal mean skill. This  
273 section focuses on the predictability of “seasonal extremes” or “daily extremes” as defined in  
274 Hamilton et al. (2012), Eade et al. (2012) and Pepler et al. (submitted). The extreme variables  
275 considered, which were selected because they are the most relevant in summer, are classified  
276 in two categories (see Section 2.3):

- 277 - The warm extremes: q90 of Tx, number of warm days, q90 of Tn and number of warm  
278 nights
- 279 - The cold extremes: q10 of Tn and number of cold nights

280 Figure 4 shows the correlation of the ensemble-mean predictions of CLIM for the JJA (one-  
281 month lead time) seasonal mean for the different extreme variables. The correlation for the

282 individual months is provided in the supplementary material (Figs. S6, S7 and S8, for June, July  
283 and August, respectively).

284 Consistent with previous studies, the pattern of extreme temperature skill tends to be similar to  
285 that of the mean temperature (Figs. 3a and 4a-f; Hamilton et al. 2012; Eade et al. 2012). The  
286 skill is also similar for all the variables inside the two groups of extreme variables (Figs. 4 a-d, e-  
287 f). The similarity is found also when undetrended anomalies are considered (Figs. 1a, S9a-f;  
288 Pepler et al., submitted). However, as for the mean temperature, the skill is lower for all extreme  
289 variables when the correlation is calculated on detrended anomalies. However, some regional  
290 differences appear. The skill of the CLIM experiment for the warm extreme variables in the  
291 Mediterranean region is slightly higher than for the seasonal-average t2m skill, while the skill of  
292 the cold extreme variables tends to be higher than the mean t2m skill in eastern and northern  
293 Europe (Figs. 4 a-f, 3a).

294 The correlation changes in INIT with respect to CLIM are very similar for all the extreme warm  
295 variables (Fig. 4g-j). Substantial improvements are found over the Mediterranean region, central  
296 Europe and Scandinavia for extreme warm variables (Figs. 4g-j), which are areas of low skill in  
297 CLIM (Fig. 4a-d). For the extreme cold variables, the soil moisture initialization leads to a weak  
298 improvement over the Mediterranean region and Western Europe and a strong degradation in  
299 northeastern Europe (Fig. 4k-l) that might be linked to the different behaviour of the snow  
300 melting in the two experiments. These patterns obey to a strong intraseasonal evolution of the  
301 skill improvement (Figs. S6-S8), with the skill decrease in northeastern Europe occurring mainly  
302 in June and the skill increase in western Europe in July, especially for the warm extremes.

303 To better understand the intraseasonal evolution of the impact of the soil initialization on the  
304 skill, the MACC calculated over Europe (20°W70°E-25°N75°N) for CLIM and INIT and the  
305 difference between the MACC in both experiments are displayed in Fig. 5. In May (first month of  
306 the forecast), Fig. 5a shows that the skill for predicting the mean and the cold extremes is high  
307 (up to 0.7), while the skill for the warm extremes is substantially lower (around 0.25). For the  
308 CLIM experiment, the skill of all the variables decreases along the forecast time and reaches  
309 almost zero in July (Fig. 5a). In INIT, as in CLIM, the skill sharply decreases between May and  
310 June, but remains almost constant at ~0.1 for all variables, which is statistically significant at  
311 95% but not high enough to be considered useful in term of seasonal forecasting (Fig. 5a).  
312 Hence, the positive impact of the soil initialization over Europe is more obvious a few weeks  
313 after the forecasts have been initialized and is found for all the variables considered. This can  
314 be better observed in Fig. 5b, which displays the difference of MACC between the two  
315 experiments. There is almost no difference for the variables in May, while in June the cold

316 extremes and the mean t2m exhibit a negative impact of land-surface initialization and a positive  
317 or neutral impact for the warm extremes. The decrease in skill of the cold extremes in June is  
318 related to the important decrease of skill in central Europe (Figs. 6, S5), which occurs for all  
319 variables but is stronger for the cold extremes. As in previous cases, the degradation of skill due  
320 to the land initialization happens for regions and periods where the skill is high in CLIM (Figs.  
321 4e-f, 5, S3). Conversely, in July and August, when the skill is low in CLIM (Fig. 5a), we observe  
322 an important improvement in INIT for all variables, especially for the warm extremes (Fig. 5b).  
323 In spite of the positive impact of the land initialization over Europe, different regions experience  
324 a different impact. Figure 6 shows the correlation difference for the mean t2m and the extreme  
325 variables averaged in some of the regions defined in Christensen and Christensen (2007). In  
326 two regions of low skill in the CLIM experiment (Scandinavia and eastern Europe; Fig. 4), the  
327 land-surface initialization has a positive impact for all variables and during the complete  
328 forecast length (Fig. 6d, f, g). In the Alps and Mediterranean area, despite a degradation of skill  
329 during one forecast month, the skill is generally higher in INIT than in CLIM. In the three other  
330 regions considered, France, central Europe and the Iberian Peninsula, the results are less clear  
331 with improvement for some variables occurring simultaneously to degradation of other variables.  
332 No statistically significant differences can be found, except for the number of warm days in  
333 eastern Europe and for the number of warm nights in Scandinavia.  
334 In summary, the impact of the land-surface initialization is generally positive on predictions of  
335 both the mean t2m and extreme temperature variables and is slightly stronger for the warm than  
336 for the cold extremes. The improvements last the whole forecast length. However, the results  
337 vary from one region to another, and might be associated with the correct prediction of a few  
338 events. An analysis of the impact on two of the most relevant events recorded recently over  
339 Europe might help interpreting these results.

#### 340 4. Predictions of the European summers of 2003 and 2010

341 Dry soils seem to have played a key role in the development of the 2003 and 2010 heat waves  
342 over Western Europe and Russia (Weisheimer et al. 2011; Quesada et al 2012; Fischer 2014).  
343 The CLIM and INIT experiments allow investigating the soil contribution to these events and to  
344 understand their role in determining the seasonal forecast skill.  
345 Figures 7 and 8 illustrate the summer 2003 and 2010 events from observational estimates and  
346 their representation in both INIT and CLIM. The left column shows the observed anomalies for  
347 five variables: t2m, precipitation, 500-hPa geopotential height (z500, hereafter), sea level  
348 pressure (SLP, hereafter) and vertically integrated soil moisture. Dots are used to mark the

349 areas where the anomalies are higher than the climatological upper quintile for t2m, z500 and  
350 SLP and are lower than the climatological lower quintile for precipitation and soil moisture. The  
351 CLIM and INIT results are displayed in the central and right columns, respectively. Instead of  
352 displaying ensemble-mean anomalies, which usually are seriously damped when compared to  
353 the reference, the forecast odds are computed from the ensemble. The odds are the ratio  
354 between the probability for the anomalies to be in the upper quintile, the interquintile range or  
355 the lower quintile, and the climatological probability of these three categories (respectively 20%,  
356 60% and 20%). Each point is attributed to the category corresponding to the highest odds ratio.  
357 If the point is attributed to the interquintile range or if there is no category assigned (the  
358 categories with two highest odds ratio have an equal value) the point is drawn in white. If the  
359 point is attributed to the lower/upper quintile category, the corresponding odds ratio is plotted  
360 with the left/right color scale. The odds ratio is a useful way of representing the signal in a  
361 probabilistic way because it gives an estimate of how anomalous the probability of the event is  
362 (i.e. the number of times it can occur above its climatological frequency) independently of the  
363 baseline. These figures allow visualizing how the hindcasts predict the extreme quintile  
364 categories for each point.

365 For the 2003 heat wave, Fig. 7 confirms the occurrence of the warm and dry event over  
366 Western Europe in 2003. A blocked regime is visible in the geopotential height, with negative  
367 anomalies over north-eastern Europe and positive anomalies over the North Atlantic and  
368 Western Europe (Fig. 7g; Garcia-Herrera et al. 2010). The blocking regime is also clearly visible  
369 on SLP, except over Western Europe where, consistently with García-Herrera et al. (2010) and  
370 Fischer et al. (2007b), the heat low mechanism takes place.

371 Both INIT and CLIM are able to forecast with high probability this warm and dry anomaly over  
372 Western Europe (Fig. 7 b, c, e, f). A successful prediction of the 2003 heat wave has previously  
373 been achieved with retrospective forecasts presented in Weisheimer et al. (2011), where the  
374 authors highlighted the crucial role of the land surface for the correct prediction of this event. An  
375 initial dry anomaly in spring has further been discussed to have been pre-requisite for the  
376 development of the 2003 heat wave (Fischer et al. 2007b; Ferranti and Viterbo, 2006). The fact  
377 that both experiments are able to forecast the 2003 heat wave is hence surprising and suggests  
378 that the exceptional high temperatures in 2003 may be largely a consequence of a strong  
379 dynamical forcing. This is supported further by the fact that, in spite of starting from  
380 climatological initial conditions, the CLIM experiment develops a high probability of extremely  
381 low soil moisture over the Mediterranean and Western Europe. This result is consistent with the  
382 studies of Feudale and Shukla (2011a, 2011b), which suggest oceanic conditions to be a major

383 driver of the heat wave. However, the soil moisture, precipitation and temperature are  
384 forecasted with higher probabilities in INIT than in CLIM (Fig. 7 b, c, e). Moreover, the spatial  
385 pattern of the observed anomalies is better reproduced in INIT than in CLIM. For instance, the  
386 dipole structure of temperature and precipitation between north-eastern and western Europe, a  
387 characteristic of a blocking regime, is, in contrast to CLIM, reproduced more realistically in INIT.  
388 These differences between the two experiments suggest that soil moisture plays a role in  
389 maintaining the blocking regime over Europe and for the occurrence or maintenance of the  
390 baroclinic anomalies of the heat low mechanism over western Europe, consistently with the  
391 studies of Fischer et al. (2007b) and Miralles et al. (2014).

392 In the case of the 2010 heat wave Fig. 8a, d, g and m show the occurrence of the warm and dry  
393 event over Russia in 2010 associated with a dry soil moisture anomaly and an anticyclone over  
394 Russia. This warm and dry anomaly associated to high sea level pressure is substantially higher  
395 (or lower for soil moisture and precipitation) than the climatological higher quintile for all  
396 concerned variables, consistently with Dole et al. (2011). Unlike 2003, for the summer 2010  
397 event no heat low mechanism takes place associated with the anticyclone and warm and dry  
398 anomalies over Russia, although the z500 anomaly is shifted with respect to the SLP anomaly.  
399 Figure 8 shows that CLIM is not able to predict with probabilities substantially different from the  
400 climatological ones the extreme characteristics of the 2010 Russian heat wave for none of the  
401 considered variables, except for the soil moisture anomaly. Conversely, in INIT, high  
402 probabilities for warm and dry anomalies are found in Eastern Europe (Fig. 8c, f). Figure 8i  
403 shows that INIT predicts relatively well the z500 anomalies, indicating that soil moisture  
404 initialization might have a feedback on the atmospheric circulation. Nevertheless as for 2003,  
405 the SLP pattern of anomalies is not reproduced correctly in INIT nor in CLIM (Fig. 8k, l). The  
406 anomalies of temperature, precipitation are misplaced compared to the observational reference  
407 (Fig. 8c, f).

408 In order to better understand how the soil initial conditions can affect the predictability of the  
409 event, figure 9 shows for both 2003 and 2010, the soil moisture anomalies, with respect to the  
410 daily climatology calculated over 1981-2010, for May 1st. Previous studies have suggested that  
411 the 2003 spring was possibly drier than usual (Fischer et al. 2007), however more recent  
412 analyses have shown that it was actually likely close to climatology. For instance, in a  
413 catchment in Northeastern Switzerland with measurements of whole surface water balance  
414 (including soil moisture and evapotranspiration), a recent study (Seneviratne et al. 2012) has  
415 shown that soil moisture was not particularly low prior to June 2003. This result was confirmed  
416 more broadly for a large part of Central Europe in another study (Whan et al. 2015) based on a

417 newly derived soil moisture dataset (Orth and Seneviratne 2015). However, Fig. 9 shows that  
418 according the ERA-Land product, the soil moisture over western Europe in 2003 exhibit a large  
419 dry anomaly over the whole western Europe at the beginning of May. During the course of May,  
420 soil moisture dries over western Europe and recovers at the end of the month and then  
421 decreases during the whole summer. This behaviour is similar to the one described in Whan et  
422 al. (2015) based on the soil dataset of Orth and Seneviratne (2012). The two products have the  
423 same evolution during summer 2003 but ERA-Land does have larger soil anomalies than the  
424 other product in both June and May.

425 In CLIM, logically, the soil initial condition is very close to 0, while in INIT the simulation starts to  
426 a dryer state, however probably due to the interpolation errors (from T511 to T106) and the drift  
427 of the first time steps, the first day is less dry than the observed state. The progressive drying of  
428 soil during summer is well reproduced by the two simulations (due to the ensemble averaging  
429 the evolution is smoother in the simulation than in the reanalysis). Independently of the initial  
430 condition of soil, the successful forecast of temperature and precipitation leads to the correct  
431 evolution of the soil during the summer. However, it is not clear from the present experiment to  
432 know if in July or August, after the strong drying of June (Fig. 9b and Whan et al. 2015), the soil  
433 conditions are important, additional experiments with more start dates would be needed.

434 In 2010, INIT starts with a dry anomalies equivalent to the observed one, while again CLIM  
435 starts from 0 (Fig. 9d). Conversely to 2003, the model is unable to forecast the evolution of the  
436 soil moisture during the forecast, while the ERA-Land shows a drastic drying during summer,  
437 INIT and CLIM keep the same anomaly. So while in INIT, the dry conditions will allow the heat  
438 wave to develop the neutral condition in INIT will inhibit its development.

439 To summarize, it seems that the 2003 warm event was predictable even without the correct  
440 initialization of the land surface, consistently with the studies of Feudale and Shukla (2011a,  
441 2011b). The atmospheric and ocean conditions are enough to generate the dry soil moisture  
442 anomalies (Fig. 7h and 9b). This last feature shows that the atmospheric circulation was  
443 predictable by the model even without the correct soil-moisture initial condition. It hence  
444 suggests that the anticyclonic circulation over Europe was driven by the large scale conversely  
445 to what has been suggested by previous studies (Garcia-Herrera et al. 2010). However, it  
446 appears that the soil moisture is also an important factor for the occurrence of the 2003 heat  
447 wave. First, the precipitation and temperature are better predicted when the soil moisture is  
448 initialized (Fig. 7b-c, e-f). Moreover, the results of both experiments show that the soil moisture  
449 has a feedback on the atmospheric circulation. Finally, the soil moisture seems important to  
450 simulate the cold and moist anomalies in Eastern Europe, which are occurring with the heat



451 wave over Western Europe. Conversely, the 2010 heat wave is predictable only when the soil  
452 moisture is adequately initialized suggesting that the dry soil moisture anomalies at the  
453 beginning of spring might have been crucial for the development of the heat wave over Russia.

## 454 5. Summary and conclusions

455 While European climate is hardly predicted by coupled models, many studies have shown the  
456 essential role played by soil moisture in this region (Schär et al. 1999, 2004; Fischer et al. 2007;  
457 Douville 2010; Seneviratne et al. 2006, 2010, 2013; Quesada et al. 2012). In the framework of  
458 the GLACE project, the role of soil-moisture initialization has been assessed at sub-seasonal  
459 time scales (Koster et al. 2010, 2011; van den Hurk et al. 2010). Nevertheless, fewer studies  
460 evaluated how soil-moisture initialization can affects skill, especially over Europe, at longer time  
461 scales (Douville 2010).

462 The present study aims to assess the added value of land-surface initialization for seasonal  
463 forecasts. Two sensitivity experiments, consisting in 30 years of 10-member ensemble  
464 hindcasts of 4 month length have been run. Both sensitivity experiments have been carried out  
465 with the EC-Earth2.3 forecast system initialized in the same way for ocean, atmosphere and sea  
466 ice. The difference between both sensitivity experiments resides in the initialization of the soil.  
467 While the soil temperature, moisture and snow are initialized with ERA-Land in INIT, these  
468 variables are initialized with a climatology of ERA-Land in CLIM.

469 The comparison of those two experiments for summer (June-to-August average, one-month  
470 lead time) shows that land-surface initialization has a positive impact on temperature skill and  
471 also, to a lesser extent, on precipitation, which is consistent with previous studies (Koster et al.  
472 2004, 2010; Douville et al. 2010; Materia et al. 2014). This improvement is robust whether the  
473 warming trend is considered or not. At regional scale, particularly over Europe, the skill  
474 improves in a similar way for t2m and a set of associated extreme variables. The improvement  
475 occurs up to the last forecast month, which contrasts with the results described in van den Hurk  
476 et al. (2010), who found that the improvement goes up to six weeks over Europe. As they found  
477 in their study, land initialization can degrade the skill during the second month of the forecast,  
478 while important improvements occur at longer forecast times.

479 Land initialization is also crucial for the prediction of the 2010 heat wave over Russia. The  
480 prediction of the 2010 event is successful only when the soil moisture is initialized, showing that  
481 the dry conditions preceding the heat wave were decisive in the occurrence of the event.  
482 Conversely, the 2003 European heat wave is predicted by both experiments, with either a  
483 climatological or a realistic land-surface initialization, suggesting that the event was driven by

484 the large-scale atmospheric circulation. The slightly better skill of the INIT experiment for this  
485 event still suggests a positive feedback of dry soil on temperature, consistently with Weisheimer  
486 et al. (2011).

487 This study shows an improvement of temperature skill when land is initialized, Nevertheless,  
488 while initializing realistically soil moisture improves skill in regions and for variables of low skill,  
489 for regions and variables with high skill, the land-surface initialization could lead to a skill  
490 degradation, consistently with the findings of Materia et al. (2014). This means that the land-  
491 surface initialization can increase the skill in regions where the original forecast system has no  
492 skill but, at the same time, can negatively perturb the large-scale signal or the local conditions  
493 in regions of positive skill. A better knowledge of the interaction between the large-scale  
494 circulation and the local land-atmosphere coupling as well as an evaluation of the role of the  
495 soil-moisture drift on the temperature anomalies simulated is needed to understand the skill  
496 degradation. This assessment will require an inspection of the daily evolution of different  
497 variables, such as temperature, soil moisture, precipitation and fluxes. The comparison of  
498 different soil initialization products and different initialization techniques, such as for example the  
499 one known as anomaly initialization, could also help better understanding the processes  
500 involved. It is also important determining to what measure the findings of the current study are  
501 model dependent. The authors have plans to perform the analysis in a multi-model framework.  
502 An interesting result of the study is the ability to predict the 2003 European heat wave even  
503 without realistic land-surface initialization from May, suggesting that there is a role for the large-  
504 scale circulation. Deeper analysis are needed to confirm the robustness of this result, first large  
505 discrepancies seems to exist between different dataset suggesting that different soil product  
506 should be tested for the initialization. Moreover, a large drying which occurs at the beginning of  
507 June, new simulation would be needed to know the influence of this drying on the heat wave.  
508 The INIT and CLIM simulation will be extended in this purpose. With the help of those extended  
509 simulations, the authors will analyze the possible remote forcing of the blocking events over  
510 Europe in 2003 and the land-atmosphere feedbacks, which took place that summer.

## 511 Acknowledgements

512 The research leading to these results has received funding from the EU Seventh Framework  
513 Programme FP7 (2007-2013) under grant agreements 308378 (SPECS), 282378 (DENFREE)  
514 and 607085 (EUCLEIA), and from the Spanish Ministerio de Economía y Competitividad  
515 (MINECO) under the project CGL2013-41055-R. We acknowledge the s2dverification R-based  
516 package (<http://cran.r-project.org/web/packages/s2dverification/index.html>). We also thank

517 ECMWF for providing the ERA-Land initial conditions and computing resources through the  
518 SPICCF Special Project.  
519

520

## 521 **References**

522 Balmaseda M, Anderson D, Vidard A (2007) Impact of Argo on analyses of the global ocean.  
523 *Geophys Res Lett* 34:L16605. doi: 10.1029/2007GL030452

524

525 Balmaseda M, Vidard A, Anderson DLT (2008) The ECMWF Ocean Analysis System: ORA-S3.  
526 *Mon Weather Rev* 136:3018–3034. doi: 10.1175/2008MWR2433.1

527

528 Balmaseda M, Anderson D (2009) Impact of initialization strategies and observations on  
529 seasonal forecast skill. *Geophys Res Lett* 36:L01701. doi: 10.1029/2008GL035561

530

531 Balmaseda M, Alves O, Arribas A, et al. (2009) Ocean initialization for seasonal forecasts.  
532 *Oceanography* 22:154–159.

533

534 Balmaseda M, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis  
535 system ORAS4. *Q J R Meteorol Soc* 139:1132–1161. doi: 10.1002/qj.2063

536

537 Balsamo G, Beljaars A, Scipal K, et al. (2009) A Revised Hydrology for the ECMWF Model:  
538 Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast  
539 System. *J Hydrometeorol* 10:623–643. doi: 10.1175/2008JHM1068.1

540

541 Balsamo, G., et al (2015) ERA-Interim/Land: a global land surface reanalysis data set, *Hydrol.*  
542 *Earth Syst. Sci.*, 19, 389-407, doi:10.5194/hess-19-389-2015 2015.

543

544 Behera S, Ratnam J V., Masumoto Y, Yamagata T (2012) Origin of extreme summers in  
545 Europe: the Indo-Pacific connection. *Clim Dyn* 41:663–676. doi: 10.1007/s00382-012-1524-8

546

547 Bellprat O, Kotlarski S, Lüthi D, Schär C (2013) Physical constraints for temperature biases in  
548 climate models. *Geophys Res Lett* 40:4042–4047. doi: 10.1002/grl.50737

549

550 Cassou C, Terray L, Phillips A (2005) Tropical Atlantic influence on European heat waves. *J*  
551 *Clim* 2805–2811.

552

553 Cohen J, Screen J a., Furtado JC, et al. (2014) Recent Arctic amplification and extreme mid-  
554 latitude weather. *Nat Geosci* 7:627–637. doi: 10.1038/ngeo2234  
555

556 Challinor AJ, Slingo JM, Wheeler TR, Doblas-Reyes FJ (2005) Probabilistic simulations of crop  
557 yield over western India using the DEMETER seasonal hindcast ensembles. *Tellus A* 57:498–  
558 512. doi: 10.1111/j.1600-0870.2005.00126.x  
559

560 Christensen JH, Christensen OB (2007) A summary of the {PRUDENCE} model projections of  
561 changes in {E}uropean climate by the end of this century. *Clim Change* 81:7–30.  
562

563 Dee DP, Berrisford P, Poli P, Fuentes M (2009) ERA-Interim for climate monitoring. *ECMWF*  
564 *Newsletter* 119: 5–6.  
565

566 Dee DP, Uppala SM, Simmons a. J, et al. (2011) The ERA-Interim reanalysis: configuration and  
567 performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. doi:  
568 10.1002/qj.828  
569

570 Dirmeyer P (2005) The land surface contribution to the potential predictability of boreal summer  
571 season climate. *J Hydrometeorol* 6:618–632. doi: 10.1175/JHM444.1  
572

573 Doblas-Reyes FJ, Hagedorn R, Palmer TN, Morcrette J-J (2006) Impact of increasing  
574 greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys Res Lett* 33:L07708.  
575 doi: 10.1029/2005GL025061  
576

577 Doblas-Reyes FJ, García-Serrano J, Lienert F, et al (2013) Seasonal climate predictability and  
578 forecasting: Status and prospects. *Wiley Interdiscip Rev Clim Chang* 4:245–268. doi:  
579 10.1002/wcc.217  
580

581 Dole R, Hoerling M, Perlwitz J, et al. (2011) Was there a basis for anticipating the 2010 Russian  
582 heat wave? *Geophys Res Lett* 38:1–5. doi: 10.1029/2010GL046582  
583

584 Douville H (2010) Relative contribution of soil moisture and snow mass to seasonal climate  
585 predictability: A pilot study. *Clim Dyn* 34:797–818. doi: 10.1007/s00382-008-0508-1  
586

587 Eade R, Hamilton E, Smith DM, et al. (2012) Forecasting the number of extreme daily events  
588 out to a decade ahead. *J Geophys Res Atmos* 117:D21110. doi: 10.1029/2012JD018015  
589

590 Feudale L, Shukla J (2011) Influence of sea surface temperature on the European heat wave of  
591 2003 summer. Part I: an observational study. *Clim Dyn* 36:1691–1703. doi: 10.1007/s00382-  
592 010-0788-0  
593

594 Feudale L, Shukla J (2011) Influence of sea surface temperature on the European heat wave of  
595 2003 summer. Part II: a modeling study. *Clim Dyn* 36:1705–1715. doi: 10.1007/s00382-010-  
596 0789-z  
597

598 Fichet T, Maqueda Morales MA (1997) Sensitivity of a global sea ice model to the treatment of  
599 ice thermodynamics and dynamics. *J Geophys Res* 102:12609. doi: 10.1029/97JC00480  
600

601 Fischer EM, Seneviratne SI, Lüthi D, Schär C (2007) Contribution of land-atmosphere coupling  
602 to recent European summer heat waves. *Geophys Res Lett* 34:L06707. doi:  
603 10.1029/2006GL029068  
604

605 Fischer EM, Seneviratne SI, Vidale PL, et al (2007b) Soil Moisture–Atmosphere Interactions  
606 during the 2003 European Summer Heat Wave. *J Clim* 20:5081–5099. doi: 10.1175/JCLI4288.1  
607

608 Fischer EM (2014) Climate science: Autopsy of two mega-heatwaves. *Nat Geosci* 7:332–333.  
609 doi: 10.1038/ngeo2148  
610

611 Fosser G, Khodayar S, Berg P (2014) Benefit of convection permitting climate model  
612 simulations in the representation of convective precipitation. *Clim Dyn* 44:45–60. doi:  
613 10.1007/s00382-014-2242-1  
614

615 García-Herrera R, Díaz J, Trigo RM, et al. (2010) A Review of the European Summer Heat  
616 Wave of 2003. *Crit Rev Environ Sci Technol* 40:267–306. doi: 10.1080/10643380802238137  
617

618 García- Morales M, Dubus L (2007) Forecasting precipitation for hydroelectric power  
619 management: how to exploit GCM's seasonal ensemble forecasts. *Int J Climatol* 1705:1691–  
620 1705. doi: 10.1002/joc

621  
622 Guemas V, Doblas-Reyes FJ, Mogensen K, et al. (2014) Ensemble of sea ice initial conditions  
623 for interannual climate predictions. *Clim Dyn*. doi: 10.1007/s00382-014-2095-7  
624  
625 Hamilton E, Eade R, Graham RJ, et al. (2012) Forecasting the number of extreme daily events  
626 on seasonal timescales. *J Geophys Res* 117:D03114. doi: 10.1029/2011JD016541  
627  
628 Hazeleger W, Wang X, Severijns C, et al. (2011) EC-Earth 8.2: description and validation of a  
629 new seamless earth system prediction model. *Clim Dyn* 39:2611–2629. doi: 10.1007/s00382-  
630 011-1228-5  
631  
632 Hirschi M, Seneviratne SI, Alexandrov V, et al. (2011) Observational evidence for soil-moisture  
633 impact on hot extremes in southeastern Europe. *Nat Geosci* 4:17–21. doi: 10.1038/ngeo1032  
634  
635 Hourdin F, Grandpeix J-Y, Rio C, et al. (2012) LMDZ5B: the atmospheric component of the  
636 IPSL climate model with revisited parameterizations for clouds and convection. *Clim Dyn*  
637 40:2193–2222. doi: 10.1007/s00382-012-1343-y  
638  
639 Jaeger EB, Seneviratne SI (2010) Impact of soil moisture–atmosphere coupling on European  
640 climate extremes and trends in a regional climate model. *Clim Dyn* 36:1919–1939. doi:  
641 10.1007/s00382-010-0780-8  
642  
643 Kim H-M, Webster PJ, Curry J a. (2012) Seasonal prediction skill of ECMWF System 4 and  
644 NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter. *Clim Dyn* 39:2957–  
645 2973. doi: 10.1007/s00382-012-1364-6  
646  
647 Kutiel H, Benaroch Y (2002) North Sea-Caspian Pattern (NCP)—an upper level atmospheric  
648 teleconnection affecting the Eastern Mediterranean: Identification and definition. *Theor Appl*  
649 *Climatol* 28:17–28. doi: 10.1007/s704-002-8205-x  
650  
651 Koster RD, Suarez MJ, Liu P, et al (2004) Realistic Initialization of Land Surface States: Impacts  
652 on Subseasonal Forecast Skill. *J Hydrometeorol* 5:1049–1063.  
653

654 Koster RD, Mahanama SPP, Yamada TJ, et al (2010) Contribution of land surface initialization  
655 to subseasonal forecast skill: First results from a multi-model experiment. *Geophys Res Lett*  
656 37:L02402. doi: 10.1029/2009GL041677  
657

658 Koster RD, Mahanama SPP, Yamada TJ, et al (2011) The Second Phase of the Global Land–  
659 Atmosphere Coupling Experiment: Soil Moisture Contributions to Subseasonal Forecast Skill. *J*  
660 *Hydrometeorol* 12:805–822. doi: 10.1175/2011JHM1365.1  
661

662 Landman W a., Beraki A (2012) Multi-model forecast skill for mid-summer rainfall over southern  
663 Africa. *Int J Climatol* 32:303–314. doi: 10.1002/joc.2273  
664

665 Lee S-S, Lee J-Y, Ha K-J, et al. (2010) Deficiencies and possibilities for long-lead coupled  
666 climate prediction of the Western North Pacific-East Asian summer monsoon. *Clim Dyn*  
667 36:1173–1188. doi: 10.1007/s00382-010-0832-0  
668

669 MacLachlan C, Arribas A, Peterson KA, et al. (2014) Global Seasonal Forecast System version  
670 5 (GloSea5): a high resolution seasonal forecast system. *Q J R Meteorol Soc.* doi:  
671 10.1002/qj.2396  
672

673 Madec G (2008) NEMO ocean engine. Note du Pole de modelisation, Institut Pierre-Simon  
674 Laplace (IPSL) No 27. ISSN No 1288-1619  
675

676 Mason S, Mimmack G (1992) The use of bootstrap confidence intervals for the correlation  
677 coefficient in climatology. *Theor Appl Climatol* 45:229–233.  
678

679 Materia S, Borrelli A, Bellucci A, et al (2014) Impact of Atmosphere and Land Surface Initial  
680 Conditions on Seasonal Forecasts of Global Surface Temperature. *J Clim* 27:9253–9271. doi:  
681 10.1175/JCLI-D-14-00163.1  
682

683 Miralles DG, Teuling AJ, Heerwaarden CC Van (2014) Mega-heatwave temperatures due to  
684 combined soil desiccation and atmospheric heat accumulation. *Nat Geosci* 7:345–349. doi:  
685 10.1038/ngeo2141  
686



687 Orsolini YJ, Kvamstø NG (2009) Role of Eurasian snow cover in wintertime circulation: Decadal  
688 simulations forced with satellite observations. *J Geophys Res* 114:D19108. doi:  
689 10.1029/2009JD012253  
690

691 Rene Orth, Seneviratne SI (2015) Introduction of a simple-model-based land surface dataset for  
692 Europe. *Environ Res Lett* 10:044012. doi: 10.1088/1748-9326/10/4/044012  
693

694 Palmer TN, Shutts GJ, Hagedorn R, et al. (2005) Representing Model Uncertainty in Weather  
695 and Climate Prediction. *Annu Rev Earth Planet Sci* 33:163–193. doi:  
696 10.1146/annurev.earth.33.092203.122552  
697

698 Palmer TN, Doblas-Reyes FJ, Hagedorn R, Weisheimer A (2005b) Probabilistic prediction of  
699 climate using multi-model ensembles: from basics to applications. *Philos Trans R Soc Lond B*  
700 *Biol Sci* 360:1991–8. doi: 10.1098/rstb.2005.1750  
701

702 Pepler AS, Diaz L, Karthik Kashinath K, et al, The ability of a multi-model seasonal  
703 forecasting ensemble to forecast the seasonal distribution of daily extremes. *Weather*  
704 *and Climate Extremes* (submitted)  
705

706 Phelps M, Kumar A, O'Brien J (2004) Potential predictability in the NCEP CPC dynamical  
707 seasonal forecast system. *J Clim* 17:3775–3785.  
708

709 Quesada B, Vautard R, Yiou P, et al. (2012) Asymmetric European summer heat predictability  
710 from wet and dry southern winters and springs. *Nat Clim Chang* 2:736–741. doi:  
711 10.1038/nclimate1536  
712

713 Reinhold B, Pierrehumbert R (1982) Dynamics of weather regimes: Quasi-stationary waves and  
714 blocking. *Mon. Weather Rev.*  
715

716 Rodwell M, Rowell D, Folland C (1999) Oceanic forcing of the wintertime North Atlantic  
717 Oscillation and European climate. *Nature* 398:25–28.  
718

719 Rogers J (1997) North Atlantic storm track variability and its association to the North Atlantic  
720 Oscillation and climate variability of northern Europe. *J Clim* 10:1635–1647.

721

722 Saha S, Moorthi S, Pan H-L, et al. (2010) The NCEP Climate Forecast System Reanalysis. Bull  
723 Am Meteorol Soc 91:1015–1057. doi: 10.1175/2010BAMS3001.1

724

725 Scaife A a., Copsey D, Gordon C, et al. (2011) Improved Atlantic winter blocking in a climate  
726 model. Geophys Res Lett 38:L23703. doi: 10.1029/2011GL049573

727

728 Schär C, Lüthi D, Beyerle U, Heise E (1999) The soil-precipitation feedback: A process study  
729 with a regional climate model. J Clim 12:722–741.

730

731 Seneviratne SI, Lüthi D, Litschi M, Schär C (2006) Land-atmosphere coupling and climate  
732 change in Europe. Nature 443:205–9. doi: 10.1038/nature05095

733

734 Seneviratne SI, Corti T, Davin EL, et al (2010) Investigating soil moisture-climate interactions in  
735 a changing climate: A review. Earth-Science Rev 99:125–161. doi:  
736 10.1016/j.earscirev.2010.02.004

737

738 Seneviratne S, Wilhelm M, Stanelle T, et al. (2013) Impact of soil moisture- climate feedbacks  
739 on CMIP5 projections: First results from the GLACE- CMIP5 experiment. Geophys Res Lett  
740 40:5212–5217. doi: 10.1002/grl.50956

741

742 Seneviratne SI, Lehner I, Gurtz J, et al. (2012) Swiss prealpine Rietholzbach research  
743 catchment and lysimeter: 32 year time series and 2003 drought event. Water Resour Res 48:1–  
744 20. doi: 10.1029/2011WR011749

745

746 Shaman J, Tziperman E (2011) An Atmospheric Teleconnection Linking ENSO and  
747 Southwestern European Precipitation. J Clim 24:124–139. doi: 10.1175/2010JCLI3590.1

748

749 Shukla J, Kinter III JL (2006) Predictability of seasonal climate variations: A pedagogical review.  
750 Predictability in Weather and Climate, 306-341.

751

752 Steiger (1980) Tests for comparing elements of a correlation matrix. doi: 10.1037/0033-  
753 2909.87.2.245

754

755 Valcke S (2006) OASIS3 user guide (prism\_2-5). PRISM support initiative report 3, pp 64  
756  
757 van den Hurk BJJM, et al (2000) Offline validation of the ERA40 surface scheme. European  
758 Centre for Medium-Range Weather Forecasts Technical Memorandum 295. Available from  
759 <http://nwmstest.ecmwf.int/publications/library/do/references/show?id=83838>  
760  
761 van den Hurk BJJM, Doblas-Reyes F, Balsamo G, et al (2010) Soil moisture effects on seasonal  
762 temperature and precipitation forecast scores in Europe. *Clim Dyn* 38:349–362. doi:  
763 10.1007/s00382-010-0956-2  
764  
765 van Oldenborgh GJ, Doblas Reyes FJ, Drijfhout SS, Hawkins E (2013) Reliability of regional  
766 climate model trends. *Environ Res Lett* 8:014055. doi: 10.1088/1748-9326/8/1/014055  
767  
768 von Storch H, Zwiers FW (2001) *Statistical Analysis in Climate Research*. Cambridge University  
769 Press, 484 pp.  
770  
771 Wang A, Bohn TJ, Mahanama SP, et al. (2009) Multimodel Ensemble Reconstruction of  
772 Drought over the Continental United States. *J Clim* 22:2694–2712. doi:  
773 10.1175/2008JCLI2586.1  
774  
775 Wang E, Xu J, Jiang Q, Austin J (2008) Assessing the spatial impact of climate on wheat  
776 productivity and the potential value of climate forecasts at a regional level. *Theor Appl Climatol*  
777 95:311–330. doi: 10.1007/s00704-008-0009-5  
778  
779 Wang G, Dolman a. J, Alessandri a. (2011) A summer climate regime over Europe modulated  
780 by the North Atlantic Oscillation. *Hydrol Earth Syst Sci* 15:57–64. doi: 10.5194/hess-15-57-2011  
781  
782 Whan K, Zscheischler J, Orth R, et al. (2015) Impact Of Soil Moisture On Extreme Maximum  
783 Temperatures In Europe. *Weather Clim Extrem* 1–11. doi: 10.1016/j.wace.2015.05.001  
784  
785 Weisheimer A, Doblas-Reyes FJ, Jung T, Palmer TN (2011) On the predictability of the extreme  
786 summer 2003 over Europe. *Geophys Res Lett* 38. doi: 10.1029/2010GL046455  
787

788 **Figure captions**

789 **Fig. 1: a** Correlation of the ensemble mean t2m averaged in JJA (one-month lead time) in the  
790 CLIM experiment. The dots mark the areas where the correlation is significant at the 95%  
791 confidence level. **b** Same as **a**, but for precipitation. **c** Difference of correlation of the ensemble  
792 mean between the INIT and CLIM experiments for the t2m in JJA. The dots mark the areas  
793 where the difference of correlation is significant at the 95% confidence level. **d** Same as **c** but  
794 for precipitation.

795  
796 **Fig. 2: a** Scatter plot of the difference of correlation of the JJA t2m (one-month lead time)  
797 between INIT and CLIM against the correlation of the JJA t2m in CLIM over the the Northern  
798 Hemisphere land grid points. The numbers in the corners correspond to the percentage of grid  
799 points in the respective quadrants. The grey dots correspond to the values in the grid points  
800 where neither the correlation in CLIM, INIT, nor the difference of correlation between INIT and  
801 CLIM is significant. The black dots represent the points where the correlation is significant at  
802 95% confidence level in both CLIM and INIT, the orange dots the points where the correlation is  
803 significant at 95% confidence level in INIT but not in CLIM, the light blue dots to the points  
804 where the correlation is significant at 95% confidence level in CLIM but not in INIT and the red  
805 (dark blue) dots to the points where the correlation difference is significantly positive (negative)  
806 at 95% confidence level. **b** Same as **a**, but in the tropics. **c** Same as **a**, but in the Southern  
807 Hemisphere (without Antarctica). **d** Same as **a**, but over the whole globe. The **e**, **f**, **g** and **h**  
808 panels show the equivalent results for precipitation.

809  
810 **Fig. 3:** As in Fig. 1, but with the correlation computed using the residual of the regression of the  
811 temperature and precipitation anomalies on the global mean temperature.

812  
813 **Fig. 4:** Correlation of the ensemble mean in JJA (one-month lead time) in CLIM, for **a** q90 of Tx,  
814 **b** number of warm days, **c** q90 of Tn, **d** number of warm nights, **e** the q10 of Tn and **f** number of  
815 cold nights. The dots mark the areas where the correlation is statistically significant with a 95%  
816 confidence level. Difference of correlation between INIT and in CLIM in JJA, for **g** q90 of Tx, **h**  
817 number of warm days, **i** q90 of Tn, **j** number of warm nights, **k** the q10 of Tn and **l** number of  
818 cold nights. The dots mark the areas where the difference of correlation is significant at 95%  
819 confidence level and the correlation has been computed using the detrended anomalies.

820  
821 **Fig. 5: a** Mean spatial anomaly correlation coefficient (MACC) calculated for the ensemble-  
822 mean hindcasts of CLIM (plain line) and INIT (dotted line) over the land in Europe (10°W40°E-

823 35°N75°N) for the monthly mean t2m (black), the q90 of Tx (red), the q90 of Tn (pink), the q10  
824 of Tn (purple), the number warm days (orange), the number of warm nights (green) and the  
825 number of cold days (light blue). The MACC is calculated on detrended anomalies. The solid  
826 (open) dots mark the values significant at 95% level in INIT (CLIM), estimated with a bootstrap  
827 over of 100 drawings. **b** Same as **a** but for the difference of the MACC between INIT and CLIM.  
828 None of the difference of MACC is significant at 95% level, estimated with a bootstrap of 100  
829 drawings.

830

831 **Fig. 6: a** Difference of correlation between the INIT and CLIM experiments for the temperature  
832 variables averaged in the Iberian Peninsula region (10°W3°E-36°N44°N). **b** Same as **a**, but for  
833 France (5°W5°E-44°N50°N), **c** central-Europe (2°W16°E-48°N55°N), **d** Scandinavia (5°E30°E-  
834 55°N70°N), **e** the Alps (5°E15°E-44°N48°N), **f** the Mediterranean area (3°E25°E-36°N44°N) and  
835 **g** Eastern Europe (6°E30°E-44°N55°N).

836

837 **Fig. 7: a** Observed anomalies of t2m for 2003 JJA (one-month lead time) mean (K). The dots  
838 indicate the area where the anomaly is in the upper quintile (estimated over 1981-2010). **b** Odds  
839 in CLIM for t2m. The odds are the ratio between the probability for the anomalies to be in the  
840 upper quintile, the interquintile range or the lower quintile and with the climatological probability  
841 of these three categories (20%, 60% and 20%, respectively). Each point is attributed to the  
842 category corresponding to the highest odds ratio. If the point is attributed to the interquintile  
843 range or if there is no category assigned (the categories with two highest odds ratio have an  
844 equal value) the point is drawn in white. If the point is attributed to the lower/upper quintile  
845 category, the corresponding odds ratio is plotted with the left/right color scale. **c** Same as **b**, but  
846 for INIT. **d** Observed anomalies of precipitation for 2003 JJA mean (mm/day). The dots indicate  
847 the area where the anomaly is in the lower quintile for the 1981-2010 period. **e** same as **b**, but  
848 for precipitation. **f** Same as **c**, but for precipitation. **g, h, i** same as **a, b, c**, but for geopotential  
849 height at 500 hPa (m). **j, k, l** same as **a, b, c**, but for monthly mean of 6 hourly SLP (hPa). **m, n,**  
850 **o** same as **d, e, f**, but for the vertically integrated volume fraction of water in soil (m<sup>3</sup>/m<sup>3</sup>).

851

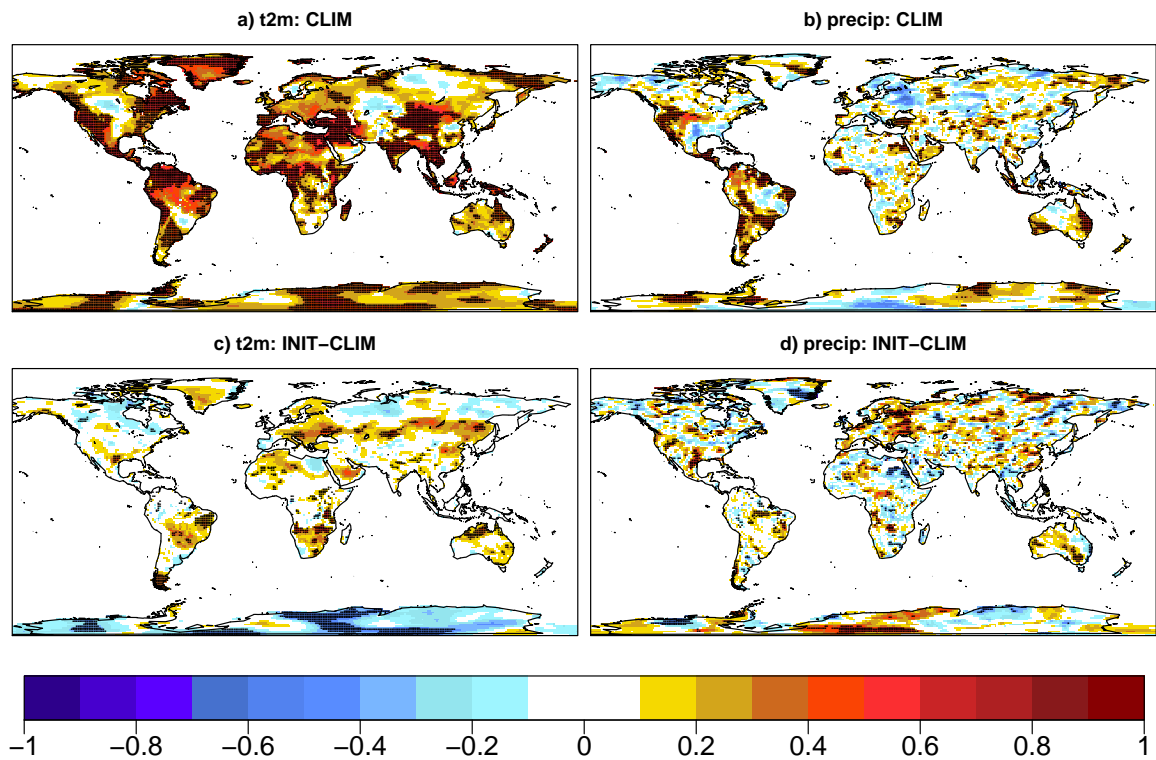
852 **Fig. 8:** Same as Fig. 7, but for JJA (one-month lead time) 2010.

853

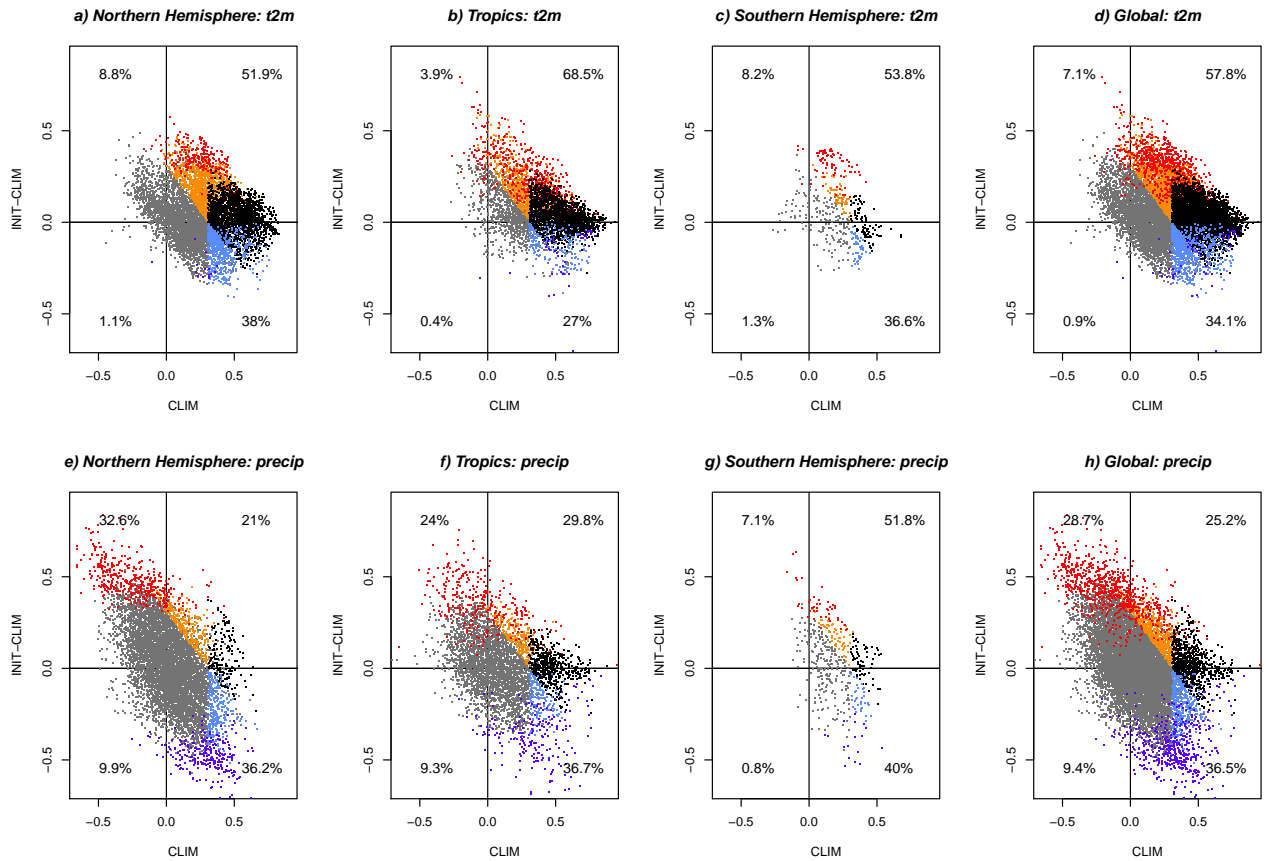
854 **Fig.9: a** Standardized anomalies with respect to the daily climatology computed over 1981-2010  
855 of ERA-Land for May 1st 2003. **b** Evolution of the daily anomalies of summer 2003 averaged in  
856 the black box of a (5W20E-43N55N) in black for ERA-Land, in blue for the ensemble mean of  
857 CLIM and in red for the ensemble mean of INIT. **c** Same as a but for May 1st 2010. **d** Same as

858 b but for the box drawn on c (25E55E-45N60N) during summer 2010. For all the panel the unit  
859 is  $\text{m}^3/\text{m}^3$ .

## Figures

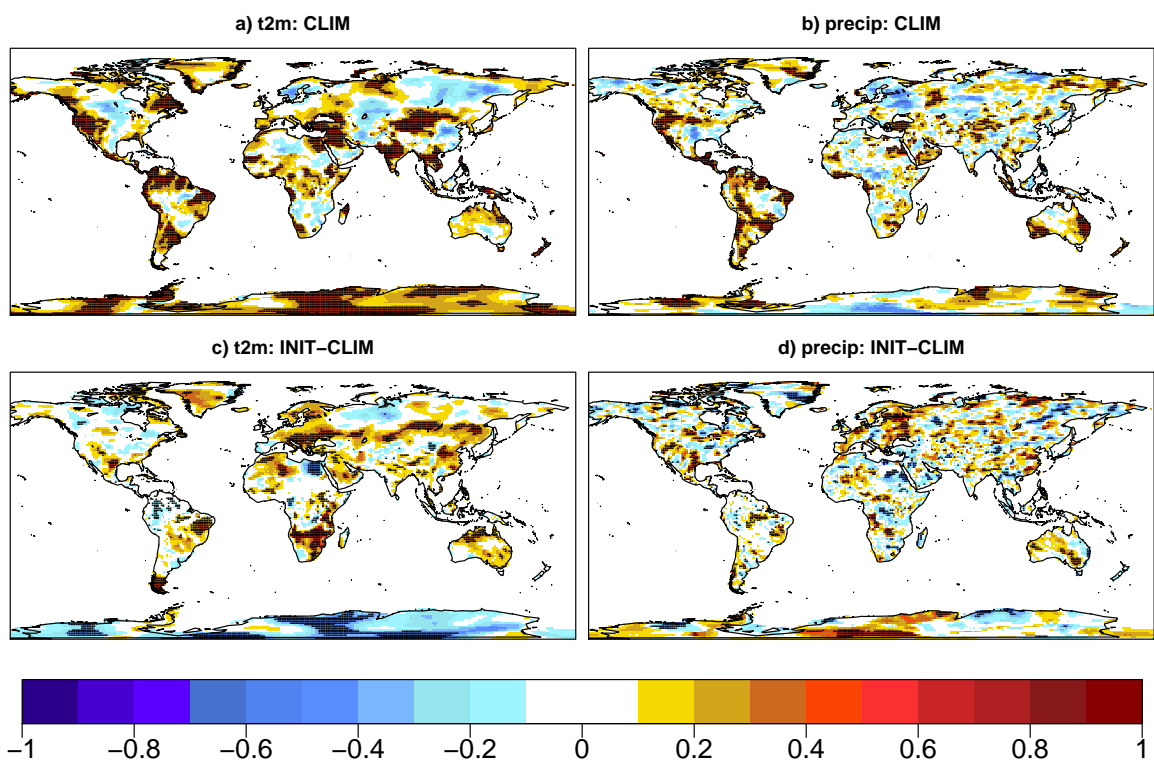


**Fig. 1:** **a** Correlation of the ensemble mean t2m averaged in JJA (one-month lead time) in the CLIM experiment. The dots mark the areas where the correlation is significant at the 95% confidence level. **b** Same as **a**, but for precipitation. **c** Difference of correlation of the ensemble mean between the INIT and CLIM experiments for the t2m in JJA. The dots mark the areas where the difference of correlation is significant at the 95% confidence level. **d** Same as **c** but for precipitation.

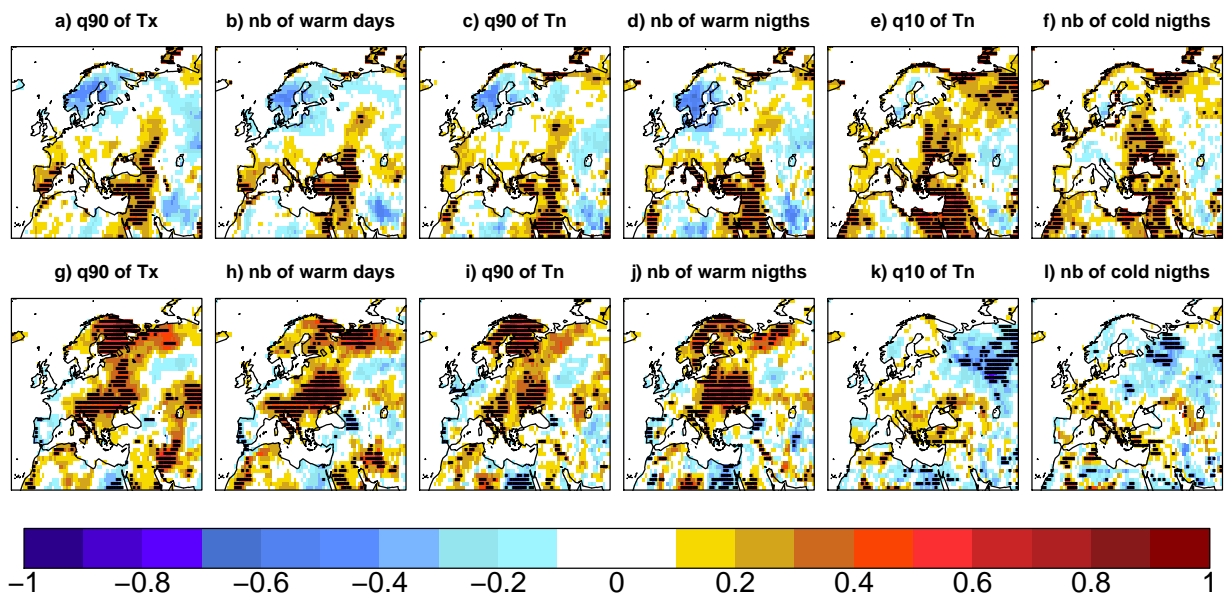


**Fig. 2:** **a** Scatter plot of the difference of correlation of the JJA t2m (one-month lead time) between INIT and CLIM against the correlation of the JJA t2m in CLIM over the the Northern Hemisphere land grid points. The numbers in the corners correspond to the percentage of grid points in the respective quadrants. The grey dots correspond to the values in the grid points where neither the correlation in CLIM, INIT, nor the difference of correlation between INIT and CLIM is significant. The black dots represent the points where the correlation is significant at 95% confidence level in both CLIM and INIT, the orange dots the points where the correlation is significant at 95% confidence level in INIT but not in CLIM, the light blue dots to the points where the correlation is significant at 95% confidence level in CLIM but not in INIT and the red (dark blue) dots to the points where the correlation difference is significantly positive (negative) at 95% confidence level. **b** Same as **a**, but in the tropics. **c** Same as **a**, but in the Southern Hemisphere (without Antarctica). **d** Same as **a**, but over the whole globe. The **e**, **f**, **g** and **h** panels show the equivalent results for precipitation.

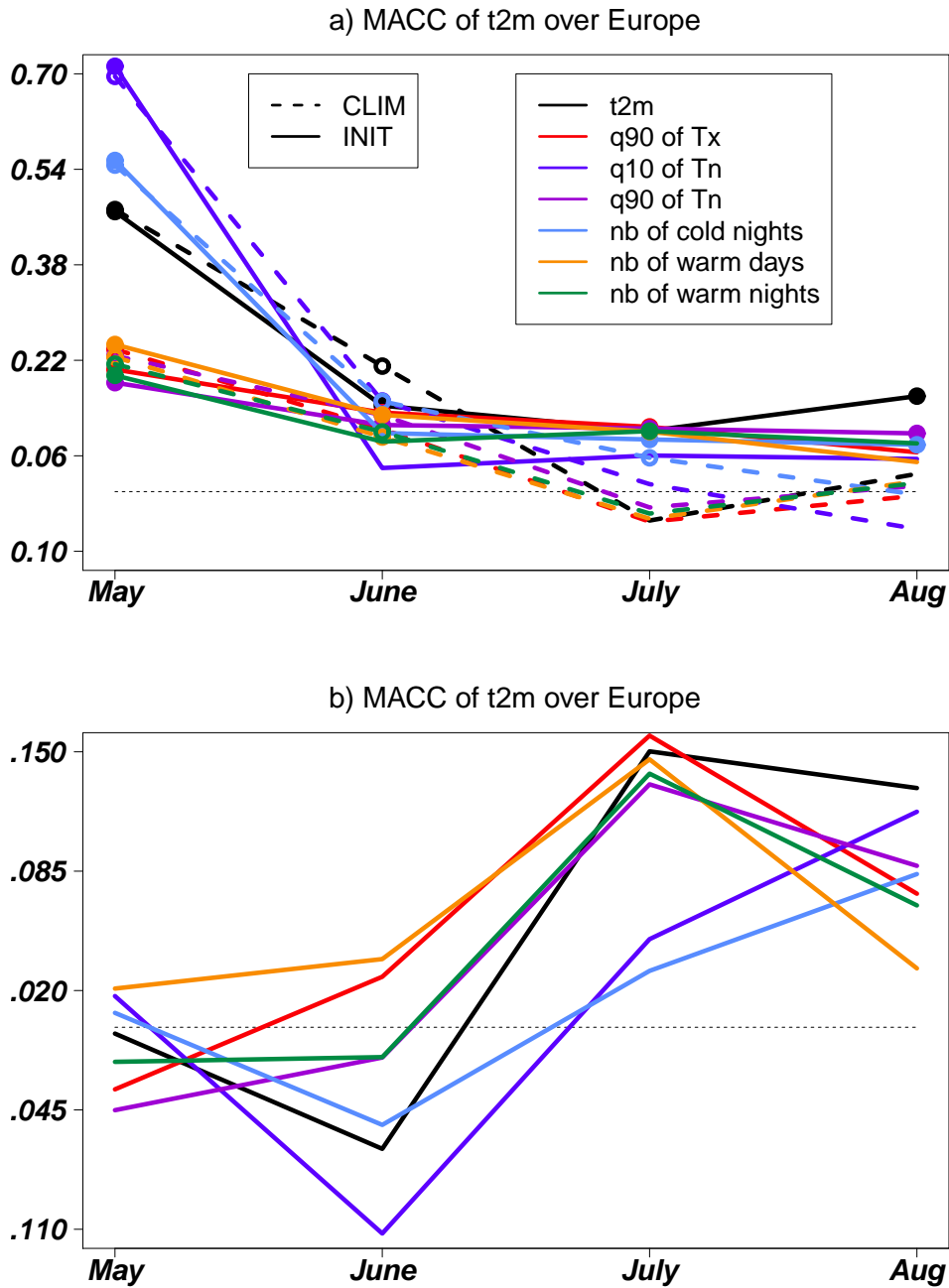




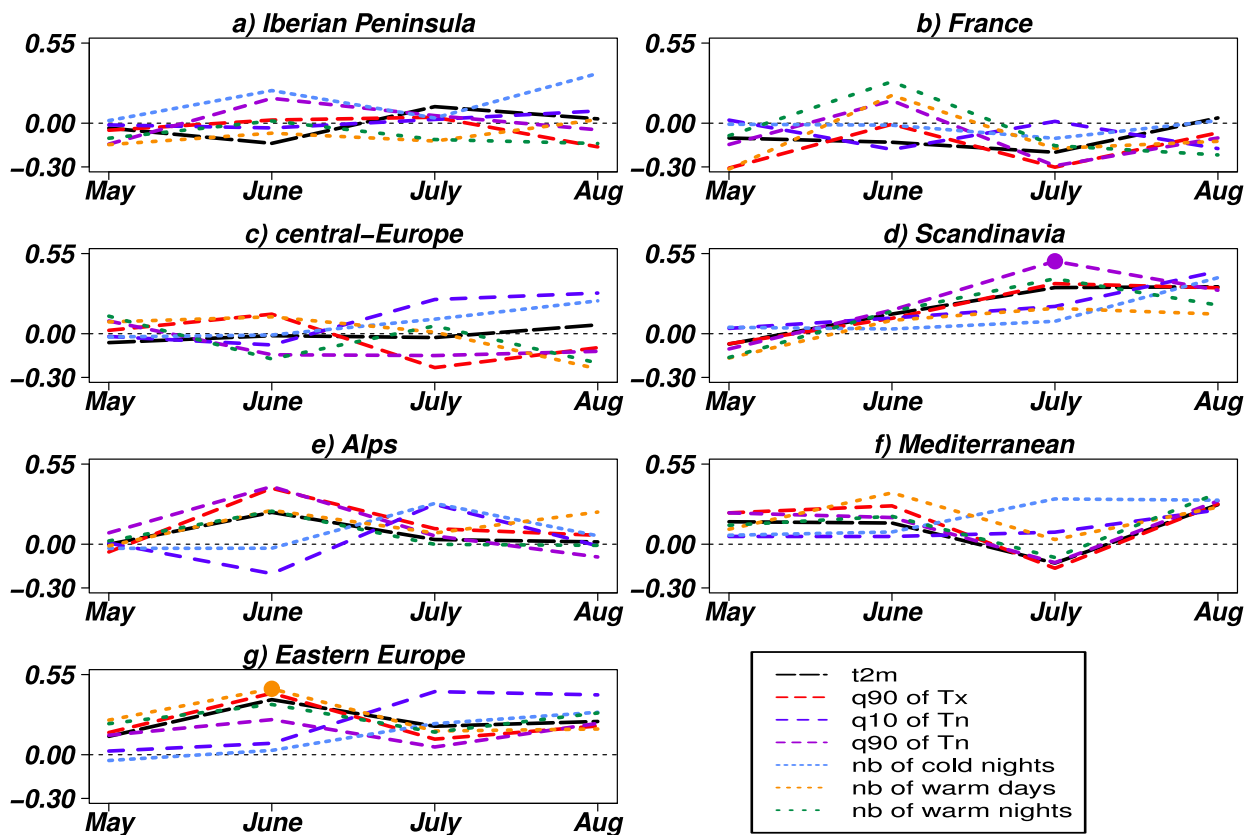
**Fig. 3:** As in Fig. 1, but with the correlation computed using the residual of the regression of the temperature and precipitation anomalies on the global mean temperature.



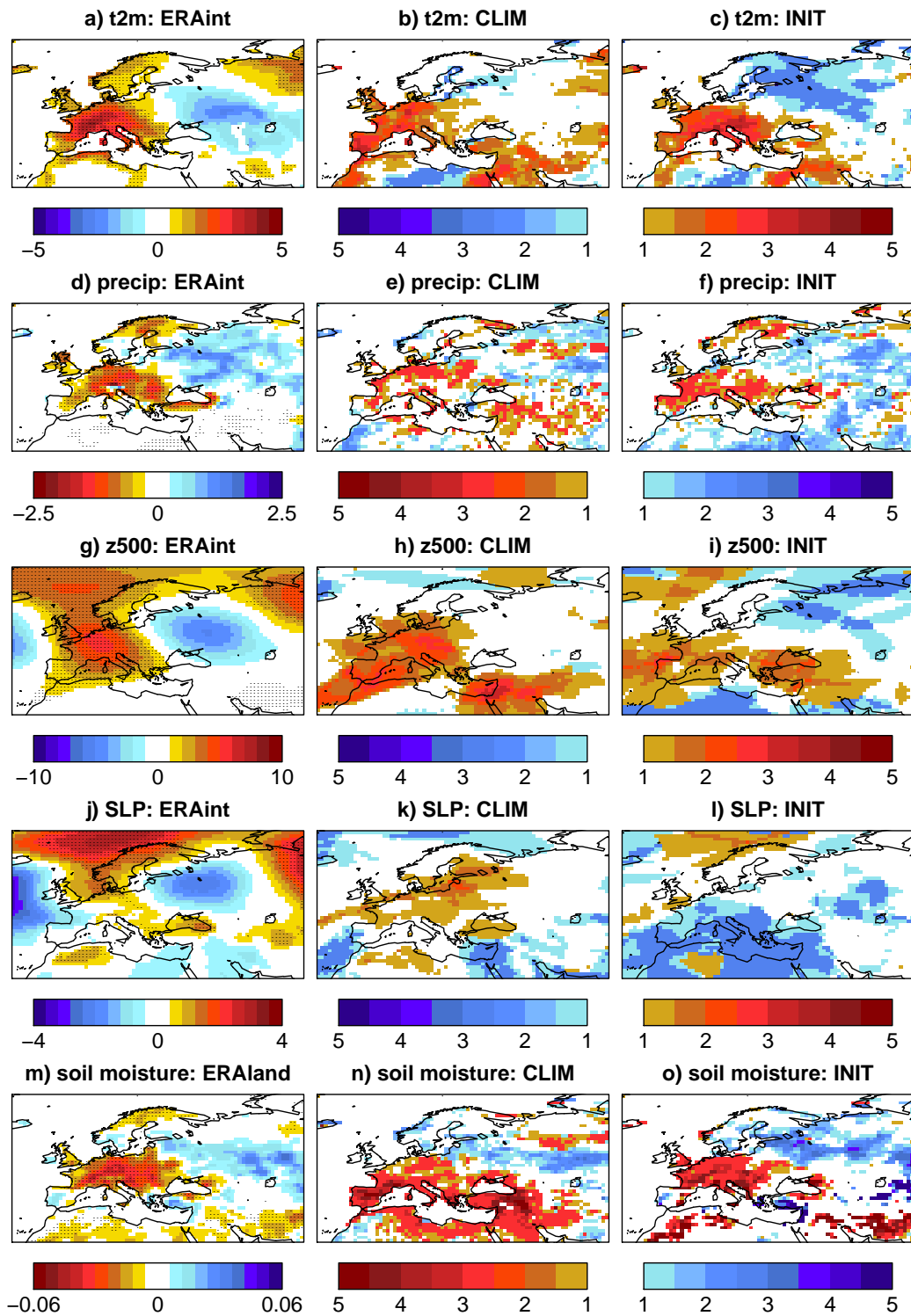
**Fig. 4:** Correlation of the ensemble mean in JJA (one-month lead time) in CLIM, for **a** q90 of Tx, **b** number of warm days, **c** q90 of Tn, **d** number of warm nights, **e** the q10 of Tn and **f** number of cold nights. The dots mark the areas where the correlation is statistically significant with a 95% confidence level. Difference of correlation between INIT and in CLIM in JJA, for **g** q90 of Tx, **h** number of warm days, **i** q90 of Tn, **j** number of warm nights, **k** the q10 of Tn and **l** number of cold nights. The dots mark the areas where the difference of correlation is significant at 95% confidence level and the correlation has been computed using the detrended anomalies.



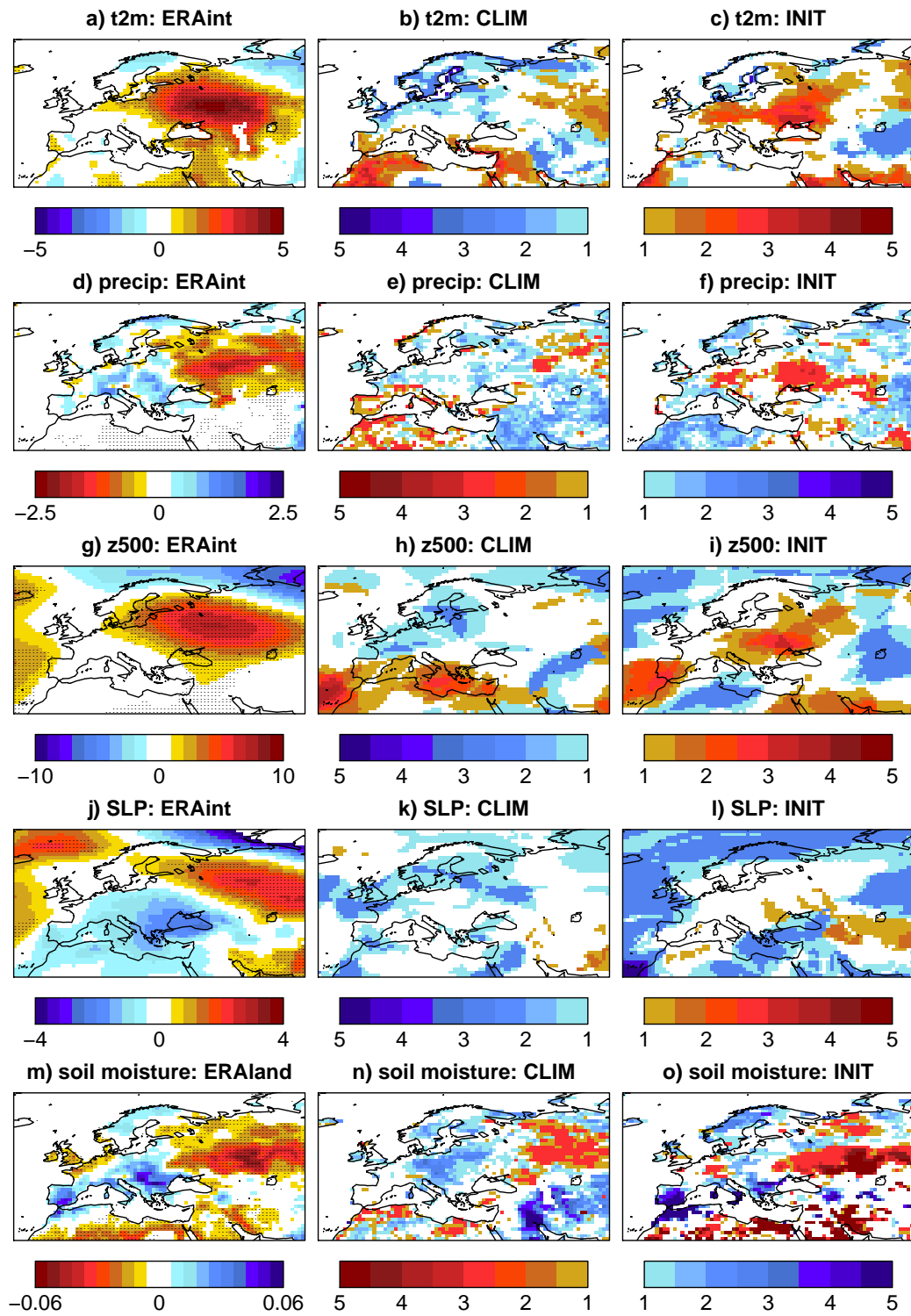
**Fig. 5:** **a** Mean spatial anomaly correlation coefficient (MACC) calculated for the ensemble-mean hindcasts of CLIM (plain line) and INIT (dotted line) over the land in Europe ( $10^{\circ}\text{W}40^{\circ}\text{E}$ - $35^{\circ}\text{N}75^{\circ}\text{N}$ ) for the monthly mean t2m (black), the q90 of Tx (red), the q90 of Tn (pink), the q10 of Tn (purple), the number warm days (orange), the number of warm nights (green) and the number of cold days (light blue). The MACC is calculated on detrended anomalies. The solid (open) dots mark the values significant at 95% level in INIT (CLIM), estimated with a bootstrap over of 100 drawings. **b** Same as a but for the difference of the MACC between INIT and CLIM.



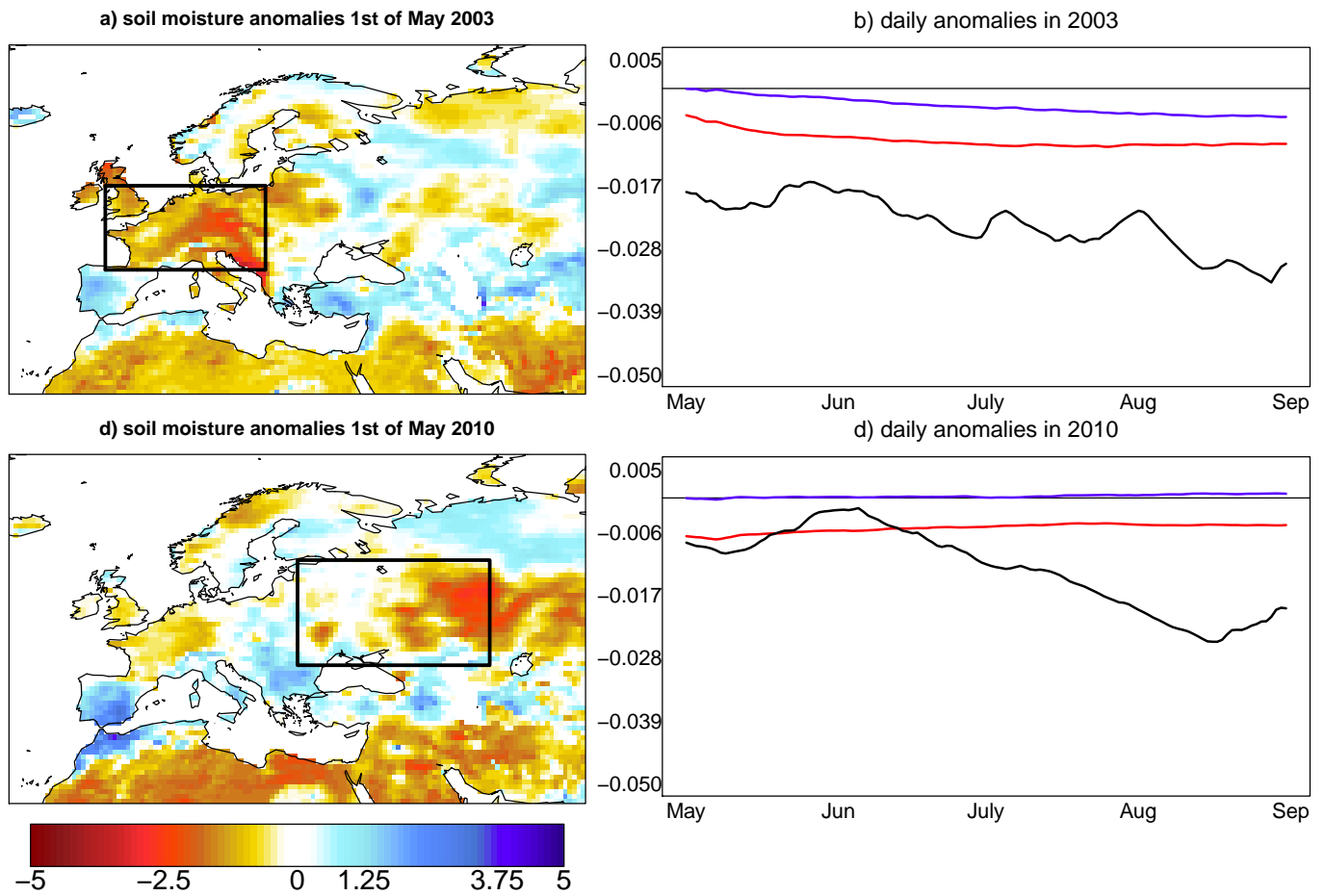
**Fig. 6:** **a** Difference of correlation between the INIT and CLIM experiments for the temperature variables averaged in the Iberian Peninsula region ( $10^{\circ}\text{W}3^{\circ}\text{E}-36^{\circ}\text{N}44^{\circ}\text{N}$ ). **b** Same as **a**, but for France ( $5^{\circ}\text{W}5^{\circ}\text{E}-44^{\circ}\text{N}50^{\circ}\text{N}$ ), **c** central-Europe ( $2^{\circ}\text{W}16^{\circ}\text{E}-48^{\circ}\text{N}55^{\circ}\text{N}$ ), **d** Scandinavia ( $5^{\circ}\text{E}30^{\circ}\text{E}-55^{\circ}\text{N}70^{\circ}\text{N}$ ), **e** the Alps ( $5^{\circ}\text{E}15^{\circ}\text{E}-44^{\circ}\text{N}48^{\circ}\text{N}$ ), **f** the Mediterranean area ( $3^{\circ}\text{E}25^{\circ}\text{E}-36^{\circ}\text{N}44^{\circ}\text{N}$ ) and **g** Eastern Europe ( $6^{\circ}\text{E}30^{\circ}\text{E}-44^{\circ}\text{N}55^{\circ}\text{N}$ ).



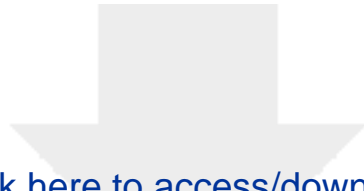
**Fig. 7:** **a** Observed anomalies of t2m for 2003 JJA (one-month lead time) mean (K). The dots indicate the area where the anomaly is in the upper quintile (estimated over 1981-2010). **b** Odds in CLIM for t2m. The odds are the ratio between the probability for the anomalies to be in the upper quintile, the interquintile range or the lower quintile and with the climatological probability of these three categories (20%, 60% and 20%, respectively). Each point is attributed to the category corresponding to the highest odds ratio. If the point is attributed to the interquintile range or if there is no category assigned (the categories with two highest odds ratio have an equal value) the point is drawn in white. If the point is attributed to the lower/upper quintile category, the corresponding odds ratio is plotted with the left/right color scale. **c** Same as **b**, but for INIT. **d** Observed anomalies of precipitation for 2003 JJA mean (mm/day). The dots indicate the area where the anomaly is in the lower quintile for the 1981-2010 period. **e** same as **b**, but for precipitation. **f** Same as **c**, but for precipitation. **g, h, i** same as **a, b, c**, but for geopotential height at 500 hPa (m). **j, k, l** same as **a, b, c** but for monthly mean of 6 hourly SLP (hPa). **m, n, o** same as **d, e, f** but for the vertically integrated volume fraction of water in soil ( $\text{m}^3/\text{m}^3$ ).



**Fig. 8:** Same as **Fig. 7**, but for JJA (one-month lead time) 2010.



**Fig. 9:** **a** Standardized anomalies with respect to the daily climatology computed over 1981-2010 of ERA-Land for May 1st 2003. **b** Evolution of the daily anomalies of summer 2003 averaged in the black box of **a** (5W20E-43N55N) in black for ERA-Land, in blue for the ensemble mean of CLIM and in red for the ensemble mean of INIT. **c** Same as **a** but for May 1st 2010. **d** Same as **b** but for the box drawn on **c** (25E55E-45N60N) during summer 2010. For all the panel the unit is  $\text{m}^3/\text{m}^3$ .



Click here to access/download

**Electronic Supplementary Material**  
sup.pdf

