



**HAL**  
open science

# Heat extremes in Western Europe are increasing faster than simulated due to missed atmospheric circulation trends

R. Vautard, J. Cattiaux, T. Happé, J. Singh, R. Bonnet, C. Cassou, D. Coumou, F. D'Andrea, Davide Faranda, E. Fischer, et al.

► **To cite this version:**

R. Vautard, J. Cattiaux, T. Happé, J. Singh, R. Bonnet, et al.. Heat extremes in Western Europe are increasing faster than simulated due to missed atmospheric circulation trends. 2023. hal-04266530v2

**HAL Id: hal-04266530**

**<https://hal.science/hal-04266530v2>**

Preprint submitted on 22 Apr 2023 (v2), last revised 31 Oct 2023 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Heat extremes in Western Europe are increasing faster than simulated due to missed atmospheric circulation trends

R. Vautard[1], J.Cattiaux[2], T. Happé[3], J. Singh [4], R. Bonnet[1], C. Cassou[5], D. Coumou[3,1,8], F. D'Andrea[6], D. Faranda[7], E. Fischer [4], A. Ribes[2], S. Sippel [4], P. Yiou[7]

[1] Institut Pierre-Simon Laplace, CNRS, Université Paris-Saclay, Sorbonne Université, France

[2] Centre National de Recherches Météorologiques, Université de Toulouse, Météo-France, CNRS, Toulouse, France.

[3] Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[4] Institute for Atmospheric and Climate Science, ETH Zurich, Zürich, Switzerland

[5] Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, CNRS UMR 5318, Toulouse, France

[6] Laboratoire de Météorologie Dynamique, IPSL, CNRS, Paris, France

[7] Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay and IPSL, 91191 Gif-sur-Yvette, France

[8] Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

## Abstract

Over the last 70 years, extreme heat has been increasing at global scale [1,2], with a rapid rate in several regions including Western Europe [3]. Climate models broadly capture heat trends globally [1], but exhibit systematically weaker extreme heat trends than observations in Western Europe [4-6], together with a weaker summer warming [7,8]. The causes of this mismatch, confirmed here by the analysis of 273 latest generation coupled climate simulations, among which only a handful of them overpass observed trends, are not well understood. Here we use a circulation analogue approach [9,10] to identify the dynamical contribution to daily maximum temperature trends [11-12], and show that a substantial fraction (0.8°C [0.2°-1.4°C] of 3.4°C per global warming degree) of the trend is due to circulation changes, largely due to increases in southerly flows over Western Europe. Their rapid increase in frequency (+43% per global warming degree

33 [10%-76%] since 1950) and persistence are largely underestimated in the 32 climate  
34 model flow simulations analyzed, as well as their overall dynamical contributions to  
35 temperature trends. The few simulations reaching the observed warming trends in  
36 extreme heat have weaker and non-significant dynamical changes, indicating  
37 compensating biases in dynamical and thermodynamical trends. These model biases in  
38 circulation trends can be due to a systematically underestimated or erroneous  
39 representation of the circulation response to external forcing, or to a systematic  
40 underestimation of interdecadal variability, or both. The former implies that future  
41 projections are too conservative, the latter that we are left with deep uncertainty  
42 regarding the pace of future summer heat in Europe: the current strong trend could  
43 weaken or increase in future decades. This calls for caution when interpreting climate  
44 projections of heat extremes over Western Europe, in particular in view of adaptation to  
45 heat waves.

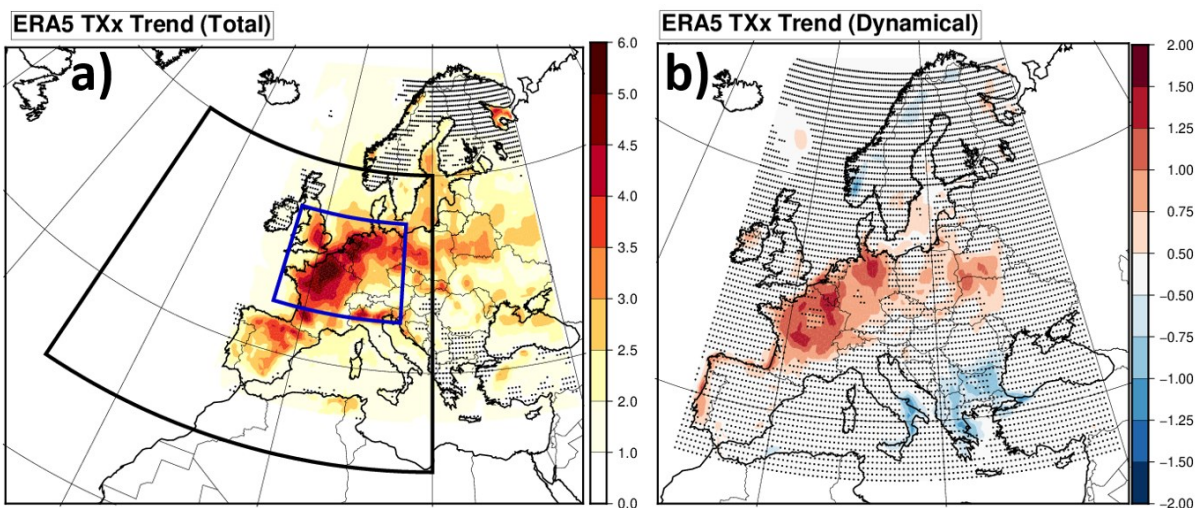
46  
47 Summer temperatures and heat extremes in Western Europe have warmed much faster than  
48 elsewhere in the mid-latitudes over the last two decades [3,5]. As a consequence, several  
49 unprecedented heatwaves took place in the last 20 years. In 2003, the full summer season mean  
50 temperature was unprecedented in Europe [13]. Northwestern Europe was hit by record  
51 temperatures in 2018 [14,15]. In 2019, two short (3-day) but intense heat waves saw all-time  
52 temperature records broken in many places, associated with a rapid northward advection of  
53 Saharan air [6]. All-time records were broken again in 2022, with temperatures above 40°C  
54 reaching far north (eg. Brittany, U.K.) [16]. Unprecedented, and even record-shattering  
55 extremes are plausible in climate projections [17], but the pace of their increasing magnitude  
56 in Western Europe is generally not predicted by these climate models [4,18].

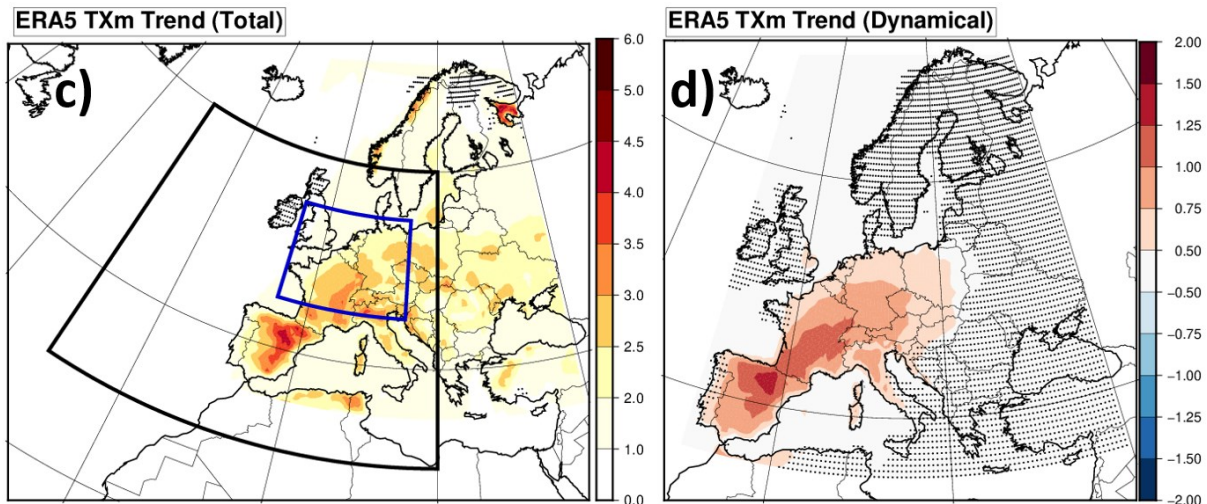
57  
58 Here we focus on summer (JJA) maximum and mean of daily maximal temperatures (resp.  
59 denoted hereafter TXx and TXm for simplicity), and the regional amplification of their trends  
60 relative to the global temperature trend. Trends in TXx and TXm are calculated over the 73-  
61 year 1950-2022 period using a linear regression with the Global mean Surface Air Temperature  
62 (GSAT, see methods section) from ERA5, and are expressed in °C per global warming degree  
63 (GWD). As shown in Figure 1 and Extended Data Fig. 1, both ERA5 reanalyses [19] and E-  
64 OBS interpolated observations [20] exhibit trends reaching more than 5°C/GWD for TXx in  
65 northern France and Benelux. Over the limited area spanning 5W-15E; 45N-55N (blue box,  
66 called hereafter “Western Europe”), the land area-average TXx trend is 3.4°C/GWD for ERA5

67 and E-OBS [2.4 - 4.3°C/GWD]. It exceeds the more moderate TXm trends by about 40% for  
68 ERA5 (2.4°C/GWD [1.7 - 3.0°C/GWD] and 30% for E-OBS (2.6°C/GWD [1.9 -  
69 3.3°C/GWD]). These rapid warming trends are exceptional on a global scale: The 20°x10°  
70 Western Europe region has the highest TXx (all year round) trend of all regions of the same  
71 size around the globe between 75°S and 75°N shifted by steps of 5° (including sea points),  
72 even though this region was originally not selected for this purpose.

73  
74 A variety of processes have been proposed for explaining these overproportional warming  
75 trends with respect to global temperature change. For mean summer temperatures, changes in  
76 mean atmospheric circulation [7,21], changes in aerosol [22] and changes in early summer soil  
77 moisture [23] and related feedbacks were considered for explaining (part of) the trends. For  
78 extreme heat, the increase in the frequency and persistence of split midlatitude jet states over  
79 the last 40 years, possibly associated with the reported weakening of the mean summer zonal  
80 circulation [24], can explain about a third of the amplified trend in heatwave intensity [3].  
81 Changes in atmospheric circulations around Europe that favor heat were also emphasized [25],  
82 in particular a positive trend in a dipole structure with a low pressure over the Eastern Atlantic  
83 [26-27] and a high pressure over the Mediterranean extended towards central Europe [28]. Yet,  
84 no increasing trend was found in blocking over Scandinavia that has led to the 2018 heat wave  
85 [14,29]. Moreover, reported changes in Rossby waves are not robust and are sensitive to their  
86 exact definition [30]. In addition, variability of summer temperatures has been shown to be  
87 large in Central Europe [31]. Thus, while several studies have hinted at a potential role of  
88 dynamical changes in amplifying European heat waves, a systematic analysis is lacking,  
89 including also how models simulate these changes.

90





92  
93 **Figure 1:** ERA5 temperature trends relative to the global warming level ( $^{\circ}\text{C}/\text{GWD}$ ), for

94 summer **Maximum daily Tmax (TXx)** (top row) and **summer Mean Tmax (TXm)**, bottom row).

95 The raw trend (left panels) is compared to the estimated dynamical contribution to these trends  
96 (right panels), obtained by replacing daily temperatures by those of best circulation analogues  
97 with a thermodynamic correction (see Methods). The areas highlighted are: **(black box)** the  
98 area used to calculate the anomaly correlation of 500 hPa streamfunction for the definition of  
99 analogues; the Western Europe focus area **(blue box)**, where maximal daily temperature trends  
100 are averaged in this study. Dotted points show areas where statistical significance of trends is  
101 less than 95% (two sided). The statistical test uses a 2-sigma rule for the regression coefficient,  
102 accounting for the total number of well-separated analogues (see Methods).

103  
104 We used a method based on circulation analogues to assess the role of dynamical changes in  
105 the **TXx and TXm** trends (see the methods section for a full description). Regional atmospheric  
106 circulation patterns are characterized by their 500 hPa streamfunction over the domain shown  
107 in Fig. 1a (black box). We identify circulation analogues for a given day by searching for other  
108 summer dates (JJA months) with similar anomaly structures, measured by the spatial anomaly  
109 correlation coefficient (ACC). A set of dates with circulation analogues allows us to calculate  
110 statistics conditionally to a given circulation [9,10,32,33], or to assess the role of dynamical  
111 changes in circulation-conditioned variables [11,12].

112

113 In order to estimate the contribution of dynamical changes to TXx and TXm trends (called  
114 hereafter the “dynamical TXx and TXm trends”), we replace each daily temperature field by  
115 the temperature field from a different day that had the best analogue circulation. In the absence  
116 of long-term trends in circulation, this is equivalent to shuffling the temperature time series  
117 while keeping the dynamics, thereby creating a trend-free “analogue temperature time series”.  
118 In the presence of long-term circulation trends, the trend in the analogue temperature time  
119 series comes from the changes in circulations (e.g. an increase in circulations favorable to heat,  
120 or vice versa). Replacement by analogues should in principle remove thermodynamical effects  
121 from global warming. As global warming is not homogeneous across the time period, and to  
122 ensure analogue regional temperatures represent a given global warming level, we further apply  
123 a correction by scaling all analogue temperatures to a reference year for global warming (2022)  
124 (see Methods). We verified that results were similar in both cases (with and without scaling,  
125 not shown).

126

127 The dynamical TXx trend (Fig 1b) is generally positive over Western Europe and reaches about  
128 1.5°C/GWD in several areas. The dynamical TXm trend is found to exceed 1°C/GWD over  
129 Southwestern Europe (Fig 1d). Over Western Europe, the average TXm and TXx dynamical  
130 trends are respectively 0.74°C/GWD [0.26-1.21°C/GWD] and 0.79°C/GWD [0.24-  
131 1.35°C/GWD]. For E-OBS the dynamical trends are 0.78°C/GWD [0.27-1.29°C/GWD] and  
132 0.86°C/GWD [0.29-1.43°C/GWD] for TXm and TXx respectively.

133

134 We verify these findings on the dynamical contributions to extreme temperatures trends with a  
135 second method, called “dynamical adjustment” [34] : The method uses a spatial circulation  
136 field (here: z500 for consistency with previous studies) as a proxy in order to estimate the  
137 contribution of circulation to temperature variability. Here, we use ridge regression, a linear  
138 regression technique that regularizes the coefficients of the high-dimensional circulation  
139 predictors [35], and we subsequently evaluate the dynamical contribution of z500 to the  
140 Western Europe TXx trends and averaged results over Western Europe (see method details in  
141 the Methods section). Results are consistent with the analogue approach (Extended Data Figure  
142 2), although with a slightly weaker dynamical TXx trend of 0.56 °C/GWD.

143

144 To test the sensitivity of our results to the analogue domain, we performed sensitivity  
145 experiments by extending and reducing the domain by 10° longitude and 5° latitude (leaving  
146 about 2/3 or more of the domain common with the reference one). The dynamical trend is

147 significant and within  $0.5^{\circ}\text{C}/\text{GWD}$  and  $0.9^{\circ}\text{C}/\text{GWD}$ , except when reducing the domain towards  
148 the North-Eastern part ( $20\text{W}-20\text{E};35\text{N}-60\text{N}$ ), (dynamical tendency reduced to  $0.38^{\circ}\text{C}/\text{GWL}$ ) a  
149 probable consequence of the key role of the upstream part of the pattern.

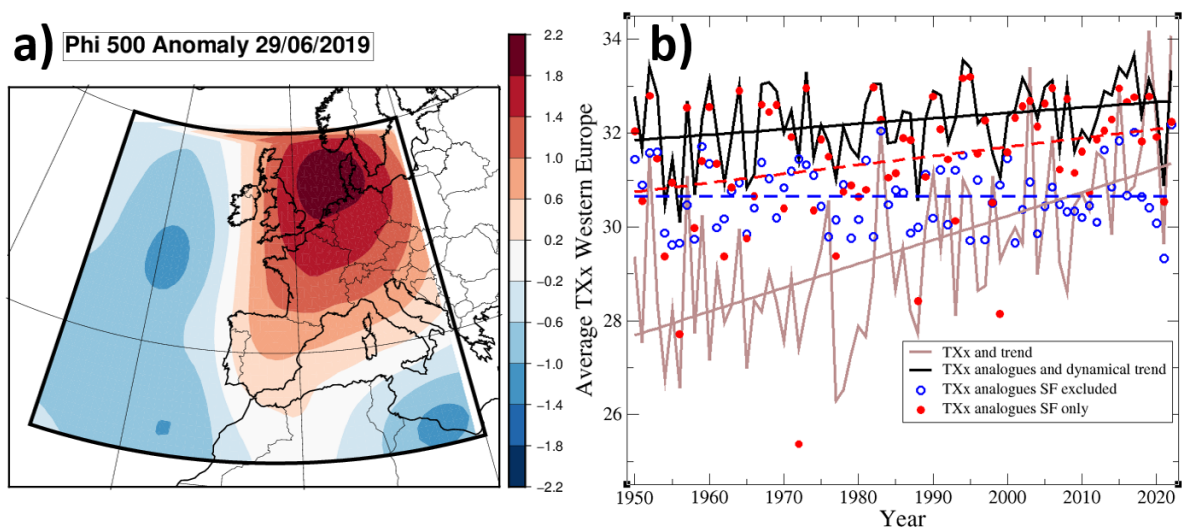
150

151 Further, we investigate the specific streamfunction patterns associated with summer maximum  
152 extreme temperatures over central France [ $1.5\text{E};46.5\text{N}$ ] – i.e., a region where the TXx  
153 dynamical trend is large (see Fig. 1). We select the reference date (29/06/2019) for which the  
154 streamfunction pattern (Fig. 2a) has a maximal average ACC (0.59) with other streamfunction  
155 patterns occurring each year when maximal temperature (TXx) is reached at this grid point, so  
156 it is most representative of those “TXx days”. We find that about 15% of the summer days in  
157 total have an ACC larger than 0.5 with the 29/06/2019 pattern, and that 53 out of 72 other TXx  
158 patterns also correlate by more than 0.5. For the sake of simplification, we will refer this class  
159 of patterns as the “Southerly Flow” patterns (SF), since almost all of the patterns bear a positive  
160 west-east streamfunction gradient (eg. 99% of patterns when considering the gradient between  
161  $15^{\circ}\text{W}$  and  $5^{\circ}\text{E}$  at  $50^{\circ}\text{N}$ ), inducing southerly flows over the Western margin of Europe. This  
162 pattern also includes a strong anticyclonic component over Central Europe, which induces  
163 increased radiation and potential land-atmosphere feedbacks if persistent. As another example,  
164 the outstanding temperatures in London on 19/07/2022 were also obtained with a similar  
165 circulation pattern (ACC=0.81 with 29/06/2019) (not shown). To assess sensitivity to the  
166 reference pattern we also repeat all calculations with the 10 most representative TXx patterns  
167 (Extended data Figure 3) in the above sense. In these other cases, the frequency of associated  
168 correlated flows is within the 10-20% range.

169

170 To check how the SF days contribute to the dynamical trend, we recalculated the dynamical  
171 trend excluding the SF days: we removed SF days from the time series, calculated the analogue  
172 temperatures of remaining days, the resulting yearly TXx, and recalculated the dynamical  
173 trend. We also did the opposite operation by keeping only SF days in the time series. On  
174 average over Western Europe (Figure 2b), the dynamical TXx trend without SF patterns  
175 becomes insignificant over Western Europe ( $0.08^{\circ}\text{C}/\text{GWD}$  on average over Western Europe),  
176 while the SF-only TXx dynamical trend is both high and statistically significant ( $1.3^{\circ}\text{C}/\text{GWD}$ ).  
177 Similar results are found when using a different reference date among the 10 most  
178 representative patterns (not shown). Dynamical TXx trends over Western Europe can therefore  
179 largely be explained by changes in the characteristics of SF patterns. First, their frequency has  
180 increased by 43% [10%;76%] per GWD (52% with time between 1950 and 2022) (see

181 **Extended Data Table 1**). Second, the number of “events” (one event is defined here as a set of  
 182 consecutive days) per year and their mean persistence have increased (see **Extended Data**  
 183 **Figure 4**). The persistence of SF patterns has increased by about **24% along the period [-1%,**  
 184 **+50%]** as a function of GWD. Such changes all give more chance, within a season, to reach  
 185 the high end of the conditional temperature distribution. Other characteristics may also have  
 186 changed (eg. amplitude) but were not investigated here. **Significant frequency increases are**  
 187 **also found for at least the 10 most representative patterns of Extended Data Figure 3, with rates**  
 188 **in the range of 35% to 55% (Extended Data Table 1).**  
 189



190  
 191 **Figure 2:** (left) Streamfunction anomaly of the 29/06/2019; (right) yearly time series of the  
 192 Western Europe average of TXx (brown), the TXx of the analogue time series, averaged over  
 193 Western Europe and **using** the 3 best analogues (**black** curve) (see Methods), and the  
 194 corresponding time series obtained by excluding (resp. including only) Southerly Flow (SF)  
 195 pattern dates before calculating the analogue TXx values (blue circles, resp. red circles). The  
 196 sets of dates (SF dates or **SF excluded** dates) within a year over which the yearly maximum is  
 197 sought are therefore complementary. In each case, analogues are calculated using the full set  
 198 of patterns (i.e. for **SF excluded** dates, analogues may contain SF patterns). Linear trends for  
 199 all series are also shown, with the same color as the series. **The dashed trends are for SF-only**  
 200 **or SF-excluded cases.**

201  
 202 Note that SF is not the only flow pattern changing, and not all patterns associated with TXx  
 203 days have an increasing frequency or persistence. For instance, the 23/07/2021 pattern,  
 204 corresponding with summer TXx in central France for 2021, shows no particular evolution



205 (Extended data Fig. 4). Our results are also consistent with the increase in occurrence and  
206 persistence of the specific class of double jet circulations explaining a large fraction of  
207 European heat extremes [3], and about half (i.e., much more than the mean probability, 15%)  
208 of double-jet days are found within the SF days.

209

210 The representation of summer TXx and TXm trends have also been analyzed for a large number  
211 of CMIP6 model simulations (273 simulations in total for 36 models) (see Methods section for  
212 data processing). Over Western Europe, almost all CMIP6 simulations fail to simulate the  
213 observed strong TXx trends, as seen in Figure 3a, plotting the percentage of simulations with  
214 larger trends than observed, for each grid point. These differences are less pronounced for TXm  
215 (Fig. 3b) but the number of runs reaching the ERA5 trend remains small here too (10-20% in  
216 large parts of South-Western Europe). There are also other land areas outside Western Europe  
217 where the CMIP6 simulations are mostly above the observed warming TXx trend (i.e. Sahara,  
218 Northern Scandinavia, Southern Balkans). This suggests that there is no general  
219 underestimation of extreme heat trends over all regions (or land regions), and that the Western  
220 Europe case is quite specific. However, understanding these regional discrepancies across the  
221 globe is beyond the scope of this article.

222

223 When averaging TXx trends over the Western Europe region above defined, only 4 of the 273  
224 individual runs analyzed (members of 3 models out of 36, ACCESS-ESM1, NorESM2-LM and  
225 KIOST-ESM) have a larger trend than the observations. The strong TXx trends observed  
226 correspond to the ~98-99th percentile of the overall CMIP6 distribution and could, from a  
227 statistical standpoint, be interpreted as consistent with Western Europe witnessing a very  
228 unlikely phase of internal interdecadal variability. However, in the five large model ensembles  
229 that were at our disposal (eg. ACCESS-ESM1-5, CanESM5, IPSL-CM6-LR, MIROC6, MPI-  
230 ESM1-2-LR), only ACCESS-ESM1-5 has a few members for which TXx warms as rapidly as  
231 observed (Figure 3c), but this ensemble strongly overestimates the TXm trend (Figure 3d).  
232 Hence, this ensemble does not correctly estimate the daily maximal temperature distribution as  
233 observed in ERA5.

234

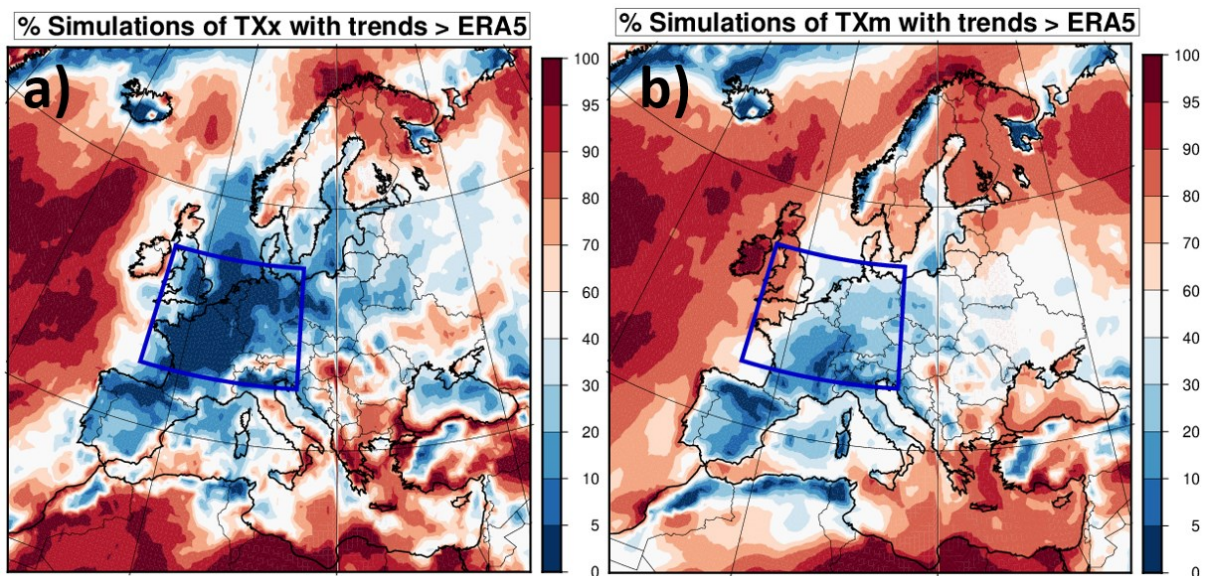
235 Our results are qualitatively robust to the way trends are calculated. We estimated trends  
236 relative to time instead of GWD, and to each model initial-condition ensemble mean GWD  
237 instead of individual member GWD. In the first (resp. second) case, 9 (resp. 5) simulations  
238 (from 4 different models) slightly exceed the ERA5 TXx trend. Trends relative to time allowed

239 in particular two members of CanESM5 to reach observations thanks to the strong global  
240 warming (about 1.7°C since 1950), while the regional response to global warming (the regional  
241 trend as a function of GWD is about twice weaker than in ERA5.

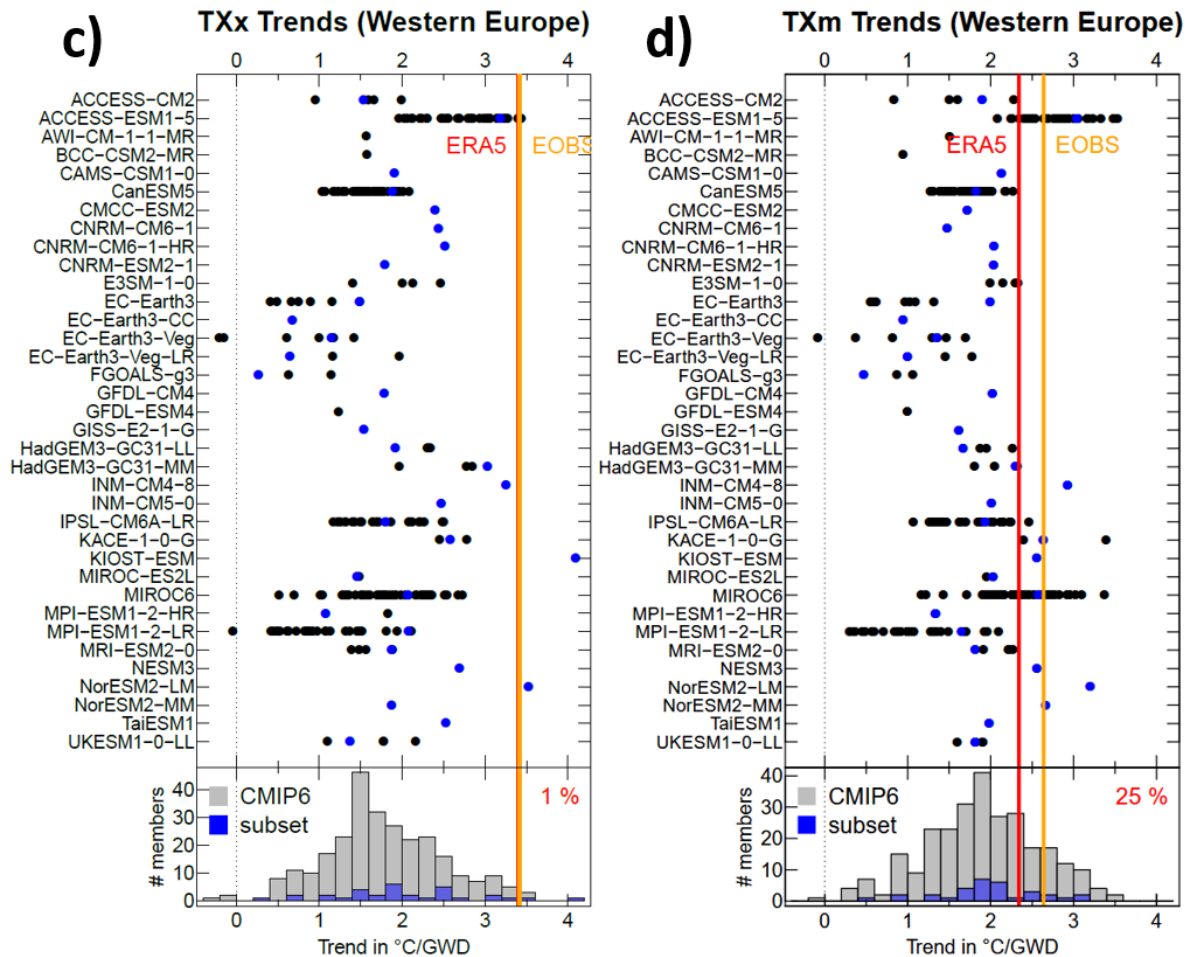
242

243 We also implemented a multiple testing procedure, the False Discovery Rate [36-38], to test  
244 the significance of the result in Western Europe. Under the hypothesis that "models are  
245 indistinguishable from reality", the rank of the observed TXx and TXm trends in the  
246 distribution of members is uniform and there can be regions over which the observation falls  
247 outside the model range only by chance. Extended Data Figure 5 shows that even taking into  
248 account the multiple nature of the test, Western Europe is among the regions where the  
249 mismatch between observed and simulated TXx trends is significant at the 95% confidence  
250 level in the sense of the FDR procedure, while no significant mismatch is found in this region  
251 for TXm trends.

252



253



254

255 **Figure 3:** Comparison between ERA5 and 273 CMIP6 simulations of trends in **Maximum**  
 256 **summer TX (TX<sub>x</sub>, left panels) and Meansummer TX (TX<sub>m</sub>, right panels)** in °C/GWD  
 257 represented in different ways; top panels: percentage of simulations with a trend larger than  
 258 ERA5 at each grid point; bottom panels: representation of trends for model ensembles (dots)  
 259 and observations (red and orange lines) after averaging over Western Europe (5°W to 15°E ;  
 260 45°N-55°N); blue dots represent the simulations that were analyzed with the analogue  
 261 approach. Histograms at the bottom of the figure summarize the overall distribution of the **TX<sub>x</sub>**  
 262 **(left) and TX<sub>m</sub> (right)** trends across the 273 simulations considered, together with the (blue)  
 263 part analyzed with the analogue approach. Percentages of simulations with a trend larger than  
 264 ERA5 are indicated in top right corners.

265

266 **In addition to the underestimated trends,** climate simulations do not capture the dynamical  
 267 changes underlying these temperature extreme changes. We applied the analogue analysis to  
 268 the first available realization for each model (to save computing and data handling and  
 269 processing burden) for which 500 hPa wind fields were available (32 simulations in total). This  
 270 set was found to be rather representative to the overall simulation distributions, see Figure 3a-

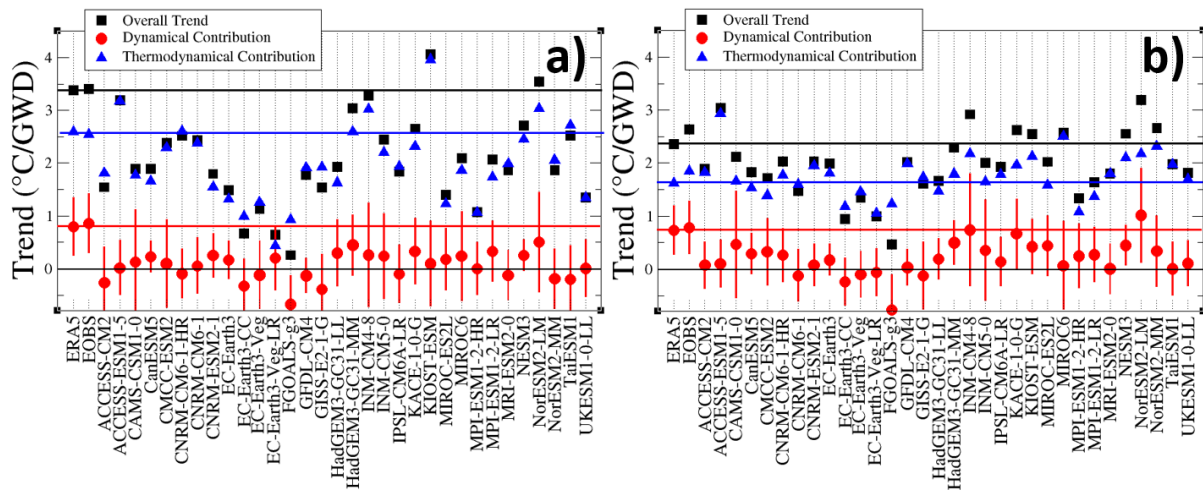
271 b histograms) regarding TXx trends, with 2 models (KIOST-ESM and NorESM2-LM)  
272 exceeding the observed trend, and 3 models with trends only slightly below (ACCESS-ESM1-  
273 5, HadGEM3-GC31-MM and INM-CM4-8). None of their dynamical temperature TXx trends  
274 are statistically significant nor reach the amplitude of the observed one over Western Europe  
275 (Figure 4a). Those simulations that do capture the observed overall TXx trend likely suffer  
276 from compensating biases between thermodynamical trends (too high) and dynamical trends  
277 (too low), as shown in Figure 4a. Some confidence intervals however encompass the ERA5  
278 and E-OBS TXx trends, leaving the possibility that variability explains the mismatch. For 12  
279 models, we could have access to new ensemble members that provided maximal overall TXx  
280 trend over Western Europe among available simulations. However these new members still  
281 underestimate the dynamical trend and the SF frequency changes, and the dynamical trend was  
282 only increased in 7 cases (Extended Data Table 2). This shows that the model underestimation  
283 of the dynamical trend does not (or weakly) depend on the member. On average over Western  
284 Europe, for TXm, a few models do have dynamical trends comparable to or larger than  
285 observations, but most have lower trends (Figure 4b).

286

287 We also calculated the thermodynamical trend obtained as a residual by subtracting the  
288 dynamical trend from the total trend and reported the result in Figure 4 (blue triangles). This  
289 shows that climate models exhibit thermodynamical contributions that are broadly consistent  
290 with ERA5, albeit with significant spread across models. TXx thermodynamical trends can  
291 both be underestimated or overestimated across models, but there is a tendency for an  
292 underestimation of TXx thermodynamical trends, and a general agreement for TXm trends.  
293 This analysis clearly shows that dynamical changes are mainly responsible for the systematic  
294 mismatch between modeled and observed temperature trends.

295

296



297  
298  
299

300 **Figure 4:** Total trend (black), dynamical contribution (red) and thermodynamical contribution  
 301 from ERA5 (blue line), for the mean summer temperature ( $TX_x$ , left panel) and yearly maximal  
 302 temperature ( $TX_m$ , right panel), averaged over Western Europe. The thermodynamical  
 303 contributions are simply calculated as residual by subtracting the dynamical trend from the  
 304 total trend. 95% Confidence intervals are shown for the dynamical trends.

305

306 Climate simulations realistically simulate the climatological mean frequency of the SF patterns  
 307 (range from 13.5% to 17.5%, not shown). However, the rapid observed increase in frequency  
 308 of this flow field (+43%/GWD [10%-76%]) is only captured by one of the simulations  
 309 (NorESM2-LM, , and largely underestimated by the others (Extended Data Table 1). The  
 310 NorESM2-LM simulation, however, also largely underestimates the change in frequency of  
 311 more than half of the other representative patterns (Extended Data Table 1). This is the prime  
 312 reason why models fail to reproduce the over-proportioned trends in heat extremes.

313

314 Overall, our results show that CMIP6 simulations underestimate the rapid observed warming  
 315 of extreme heat over Western Europe. The analysis of atmospheric circulation changes shows  
 316 that there is a large dynamical contribution to this observed trend, which is underestimated by  
 317 climate models, explaining a large part of the discrepancy in trend between models and  
 318 observations. By contrast, models and observational trends are broadly consistent in terms of  
 319 the thermodynamic contribution to the trend. We cannot rule out other sources of systematic  
 320 uncertainties such as lack of homogeneity of reanalyses, in particular for circulation patterns,

321 or inaccuracies in the aerosol and land use forcing changes that would translate in systematic  
322 model trend biases.

323

324 Determining the cause of the mismatch between models and observations is critical to assess  
325 whether the large observed warming **TXx** trend is likely or unlikely to continue. If due to a  
326 wrong forced dynamical regional response – models underestimate the forced response to  
327 greenhouse gases – then this mismatch is expected to remain and even strengthen in the future,  
328 as global warming increases. If related to unforced internal variability **[39,40]** – internal  
329 variability simulated by models is too small **[41]** – then the mismatch is expected to decrease  
330 in the future, but the term of this decrease is unknown and could be years or decades, leaving  
331 the fate of Western Europe heatwaves in large uncertainty.

332

333 Here we have shown that the observed extreme temperature trends for Western Europe are not  
334 captured by CMIP6 **simulations**, due to underestimated dynamical trends. Similar conclusions  
335 were found for wintertime weather over Europe **[42]**. Further research needs to determine the  
336 causes of the mismatch between simulated and observed **heat** trends, whether this is due to  
337 uncaptured internal variability or missing (dynamical) forcing/processes. Either way, our  
338 results call for caution when using climate model projections for adaptation and resilience  
339 plans.

340

341

## 342 Methods

### 343 Calculation of dynamical contributions to mean and extreme summer temperature

344 **trends:** The method used to estimate dynamical contribution to the change in one variable  
345 follows the conceptual framework developed in Vautard et al. (2016), with a different  
346 implementation here. It is based on the estimation of the change in the variable solely due to  
347 the changes in regional upper-air circulations. For instance, even without extra heating from  
348 radiative and diabatic processes, an increase in the frequency of southerly flows in Western  
349 Europe would induce a mean regional warming. An increase in anticyclonic conditions would  
350 similarly lead to increased radiation and thus temperature. This can also lead to a cooling if  
351 increasingly frequent circulations are linked to cooler temperatures (eg. in Northerly winds).  
352 To estimate this dynamical effect of changing circulations on temperatures, we need to  
353 carefully remove any thermodynamical effect of climate change.

354

355 We assume that daily temperature  $T$  (which can be mean, minimum or maximum daily  
356 temperature, and in the current article will be maximum temperature) has a distribution at a  
357 given location or grid point which depends on the atmospheric circulation and on other  
358 processes, including global warming. We then assume a decomposition into:

359

$$360 \quad (1) \quad T = \langle T|X \rangle_{GWD} + T' \quad ,$$

361

362 where  $X$  is the 500 hPa streamfunction anomaly, characterizing the atmospheric circulation  
363 (simultaneous to the temperature),  $GWD$  stands for the global warming degree,  $\langle T|X \rangle_{GWD}$  is  
364 the average daily maximum temperature conditioned to the circulation, assumed to be  
365 dependent on  $GWD$ , and  $T'$  is a fluctuation. This circulation-conditioned temperature includes  
366 not only advection effects (i.e. from cooler/warmer regions), but also all processes linked to  
367 the circulation (subsidence in anticyclone, increased radiation, surface-atmosphere feedbacks,  
368 ...), so the overall dynamical trend includes all underlying processes tied to the dynamical  
369 conditions. In order to remove thermodynamical effects due to climate change, we scale all  
370 temperatures to a reference warming level. For this, we assume that the circulation-conditioned  
371 mean temperature depends linearly on the global warming level, so the decomposition can be  
372 written:

373

$$374 \quad (2) \quad T = \langle T|X \rangle_{ref} + b(X) \cdot (GWD - GWD_{ref}) + T' \quad ,$$

375  
376  
377  
378  
379  
380  
381  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408

where *ref* refers to a reference global warming level, taken here as that of 2022, so all changes are expressed relative to 2022. The coefficient  $b(X)$  represents the mean warming rate conditioned to the circulation  $X$ , which includes thermodynamical effects of the climate change response – it is therefore assumed that the amount of warming depends on the circulation type. Assuming one can calculate  $b(X)$  and  $GWD$ , all daily temperatures are then scaled to the reference level with the following thermodynamical correction:

$$T_s = T - b(X) \cdot (GWD - \underbrace{GWD}_{382}^{D_{ref}}) \quad (3)$$

The dynamical contribution to any temperature trend constructed from daily temperatures (eg. here TXm, TXx) can then be calculated from the  $T_s$  time series, because changes with  $GWD$  are only through the changes in the frequency of occurrences of  $X$  for given  $GWD$ s. Trends should also not depend on the particular time  $T_s$  values are drawn as long as they occur simultaneously to a streamfunction anomaly which is similar to that encountered in the same sequence order as that of the series. Hence to increase statistical robustness and remove any residual link to the specific order of temperatures, we replace  $T_s$  temperatures by those occurring in circulations  $X$  along the time series. This has the advantage of “randomizing” the timing of analogues and providing multiple realizations to calculate dynamical trends. A new temperature analogue series is created by replacing each daily with that of the best circulation analogue, then another new series is made with the second best analogue, etc... (see below for practical analogue calculation). From each of these analogue time series, TXm and TXx are recalculated for each year, then averaged across analogues, and a regression with  $GWD$  is calculated at each grid point, together with its confidence interval, (plus or minus twice the standard error of the regression coefficient). To keep analogue quality high, we limit the number of time series to 3. To calculate time series of averages over Western Europe land, we apply the  $0.5^\circ \times 0.5^\circ$  land mask of E-OBS and average over the grid points included in [-5W - 15E ; 45N - 55N]

**Estimation of yearly GWD :** In practice,  $GWD$  is calculated as a moving centered 5-year average of the global temperature with available data, for reanalyses and models, accounting for series ends in ERA5 (i.e. for 1950, taking into account an average only over 1950 to 1952, and for 2022 an average over 2020 and 2021). The 2022 value is then subtracted to all values, so  $GWD$  is 0 in 2022, and generally negative before.



409

410 **Selection of circulation analogues :** In practice, circulations are characterized by the 500 hPa  
411 streamfunction over the  $[-30\ 20^{\circ}\text{E} ; 30\ 60^{\circ}\text{N}]$  domain. Analogs of a given circulation are  
412 characterized by anomaly correlation coefficient (ACC) between streamfunction fields. For  
413 each summer day, we collect the best analogues (highest ACCs), and impose that they remain  
414 spaced by 6 days or more within a season, and self-analogues are not considered. This is done  
415 by successively testing fields in descending order of the ACC, and skipping days not respecting  
416 the separation with previously selected fields.

417

418 **Calculation of the circulation-conditioned thermodynamical trend  $b(X)$  :** To calculate  
419  $b(X)$ , we also use analogue circulations, in a different way than above: For each summer day  $d$   
420 of the 1950-2022 period, we estimate  $b(X(d))$  using a regression of each raw temperature  $T(d)$   
421 (before thermodynamical correction) associated with a large set of best analogue circulations  
422 of  $X(d)$  found between 1950 and 2022 with the  $GWD$  values of their respective year. We use  
423 the best 1% summer analogues (67 days) with the same spacing of at least 6 days. 99% of the  
424 worst of these 67 analogues across all summer days have  $\text{ACC} > 0.5$ , 65% have  $\text{ACC} > 0.7$ .  
425 Imposing a quality criterion on analogues such as  $\text{ACC} > 0.7$  or more would leave days with an  
426 insufficient number of analogues for regression.

427

428 **Dynamical adjustment:** Dynamical adjustment is used as a second, alternative technique to  
429 estimate the influence of circulation-induced temperature trends. This method relies on the idea  
430 that temperature variability can be decomposed into a component that is driven by circulation-  
431 induced variability, and a residual, thermodynamical component. The “thermodynamical”  
432 component is expect to contain a forced signal as well as any other unexplained variability or  
433 feedbacks [43]. Most applications of this technique characterize circulation-induced  
434 temperature variability using a proxy variable such as geopotential height [34,35,44,45].  
435 Dynamical adjustment techniques typically rely on linear methods such as variants of linear  
436 regression or circulation analogue techniques.

437

438 Here, we use the spatial pattern of  $z500$  in a relatively large circulation domain over Europe  
439 and the North Atlantic ( $-30$  to  $20^{\circ}\text{E}$ ,  $30$  to  $60^{\circ}\text{N}$ , similar to Fig. 1), following the method  
440 outlined in [45]. However, we introduce some modifications and additional details. We use a  
441 regularized regression technique, called “ridge regression”, which is well-suited to deal with  
442 the large number of circulation predictor grid cells and a relatively short observed record. For

443 TXx, we train our ridge regression model on the 15 warmest days in each summer during 1950-  
444 2021 at each grid cell in the ERA5 reanalysis, resulting in a total of 1080 observations (72  
445 summers and 15 days per summer). Since the z500 field contains information about the lower  
446 troposphere, and is affected by temperature change via thermal expansion, we detrend the  
447 spatial z500 field by subtracting the global average z500 at each time step and over each grid  
448 cell in the circulation domain. Hence, the analysis is based only on relative changes within the  
449 z500 field. To obtain regional estimates of the circulation-induced component of TXx, we  
450 performed an area-weighted average across the grid cells within the study domain.

451

452

453

454 **Acknowledgements**

455

456 This study was partly supported by the European Union’s Horizon 2020 research and  
457 innovation programme under grant agreement No 101003469 (XAIDA project). PY was also  
458 supported by the grant ANR-20-CE01-0008-01 (SAMPRACE). The authors thank Dr. Efi  
459 Rousi for providing the sequences of dates of double-jet days. The authors also thank Atef Ben  
460 Nasser and the ESPRI IPSL data and computing service for their support in handling the large  
461 ensemble of climate simulations.

462

463 **Data availability**

464

465 All analyzes have been conducted using 3 main data sets. The ERA5 reanalysis and the E-OBS  
466 data sets (processed from the <https://climate.copernicus.eu>) has been downloaded, and are  
467 available from the Climate Explorer <https://climexp.knmi.nl> . CMIP6 model simulations are  
468 available from the IPSL ESGF node <https://esgf-node.ipsl.upmc.fr/> .

469

470 **Code availability**

471

472 Codes used in this article develop classical statistical algorithms, and are available upon  
473 request. Most have been archived on:

474 <https://zenodo.org/record/7832152#.ZDsPKnZBxaS>

475

476 **Authors contributions**

477

478 R.V., J.C. and J. S. carried out the statistical analysis. T. H. provided the streamfunction fields  
479 for ERA5 and the calculation method. All other authors contributed to the design of the study  
480 and the interpretation of results. All authors contributed to the writing of the article.

481

482

483

484

## References

485

[1] Seneviratne, S.I., et al. [Masson-Delmotte, et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1513–1766 (2021).

487

488

[2] Robinson, A. et al. Increasing heat and rainfall extremes now far outside the historical climate. *npj Clim Atmos Sci* 4, 45 (2021).

490

491

[3] Rousi, E., Kornhuber, K., Beobide-Arsuaga, G., Luo, F., & Coumou, D. Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia. *Nature communications*, 13(1), 1-11 (2022).

494

495

[4] van Oldenborgh, et al. Western Europe is warming much faster than expected, *Clim. Past*, 5, 1–12, <https://doi.org/10.5194/cp-5-1-2009> (2009).

496

497

498

[5] van Oldenborgh, et al. Attributing and projecting heatwaves is hard: we can do better. *Earth's Future*, 10(6), e2021EF002271 (2022).

500

501

[6] Vautard, R., et al. Human contribution to the record-breaking June and July 2019 heatwaves in Western Europe. *Environmental Research Letters*, 15(9), 094077 (2020).

503

504

[7] Bo e, J., et al. Past long-term summer warming over western Europe in new generation climate models: role of large-scale atmospheric circulation. *Environmental Research Letters*, 15(8), 084038 (2020).

507

508

[8] Ribes, A., et al. An updated assessment of past and future warming over France based on a regional observational constraint. *Earth System Dynamics Discussions*, 1-29 (2022).

510

511

[9] Yiou, P., Vautard, R., Naveau, P., & Cassou, C. Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe. *Geophysical Research Letters*, 34(21) (2007).

514

515

[10] Cattiaux, J., et al. Winter 2010 in Europe: A cold extreme in a warming climate. *Geophysical Research Letters*, 37(20) (2010).

517

518

[11] Vautard, R., & Yiou, P. Control of recent European surface climate change by atmospheric flow. *Geophysical Research Letters*, 36(22) (2009).

520

521

[12] Vautard, R., et al. Attribution of human-induced dynamical and thermodynamical contributions in extreme weather events. *Environmental Research Letters*, 11(11), 114009 (2016).

523

524

- 525 [13] García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., & Fischer, E. M. A review of  
526 the European summer heat wave of 2003. *Critical Reviews in Environmental Science and*  
527 *Technology*, 40(4), 267-306 (2010).  
528
- 529 [14] Yiou, P., et al. Analyses of the Northern European summer heatwave of 2018. *Bulletin of*  
530 *the American Meteorological Society*, 101(1), S35-S40 (2020).  
531
- 532 [15] McCarthy, M., Christidis, N., Dunstone, N., Fereday, D., Kay, G., Klein-Tank, A., ... &  
533 Stott, P. Drivers of the UK summer heatwave of 2018. *Weather*, 74(11), 390-396 (2019).  
534
- 535 [16] [https://www.worldweatherattribution.org/without-human-caused-climate-change-](https://www.worldweatherattribution.org/without-human-caused-climate-change-temperatures-of-40c-in-the-uk-would-have-been-extremely-unlikely/)  
536 [temperatures-of-40c-in-the-uk-would-have-been-extremely-unlikely/](https://www.worldweatherattribution.org/without-human-caused-climate-change-temperatures-of-40c-in-the-uk-would-have-been-extremely-unlikely/)  
537
- 538 [17] Fischer, E.M., Sippel, S. & Knutti, R. Increasing probability of record-shattering climate  
539 extremes. *Nat. Clim. Chang.* 11, 689–695 (2021).  
540
- 541 [18] Lorenz, R., Stalhandske, Z., & Fischer, E. M. Detection of a climate change signal in  
542 extreme heat, heat stress, and cold in Europe from observations. *Geophysical Research Letters*,  
543 46(14), 8363-8374 (2019).  
544
- 545 [19] Hersbach, H., et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal*  
546 *Meteorological Society*, 146(730), 1999-2049 (2020).  
547
- 548 [20] Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., & Jones, P. D. An ensemble  
549 version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research:*  
550 *Atmospheres*, 123(17), 9391-9409 (2018).  
551
- 552 [21] Hoogeveen, J., & Hoogeveen, H. Winds are changing: An explanation for the warming of  
553 the Netherlands. *International Journal of Climatology* (2022).  
554
- 555 [22] Nabat, P., Somot, S., Mallet, M., Sanchez-Lorenzo, A., & Wild, M. Contribution of  
556 anthropogenic sulfate aerosols to the changing Euro-Mediterranean climate since 1980.  
557 *Geophysical Research Letters*, 41(15), 5605-5611 (2014).  
558
- 559 [23] Stegehuis, A. I., et al. Early summer soil moisture contribution to Western European  
560 summer warming. *Journal of Geophysical Research: Atmospheres*, 126(17), e2021JD034646  
561 (2021).  
562
- 563 [24] Coumou, D., Lehmann, J., & Beckmann, J. The weakening summer circulation in the  
564 Northern Hemisphere mid-latitudes. *Science*, 348(6232), 324-327 (2015).  
565
- 566 [25] Terray, L. A dynamical adjustment perspective on extreme event attribution. *Weather and*  
567 *Climate Dynamics*, 2(4), 971-989 (2021).  
568

569 [26] Horton, D. E., et al. Contribution of changes in atmospheric circulation patterns to extreme  
570 temperature trends. *Nature*, 522(7557), 465-469 (2015).  
571

572 [27] Faranda, D., Messori, G., Jézéquel, A., Vrac, M., Yiou, P.. Atmospheric circulation  
573 compounds anthropogenic warming and extreme climate impacts in Europe. *PNAS*, 2023  
574

575 [28] Fery, L., Dubrulle, B., Podvin, B., Pons, F., & Faranda, D. Learning a weather dictionary  
576 of atmospheric patterns using Latent Dirichlet Allocation. *Geophysical Research Letters*, 49(9),  
577 e2021GL096184 (2022).  
578

579 [29] Davini, P., & d’Andrea, F. From CMIP3 to CMIP6: Northern Hemisphere atmospheric  
580 blocking simulation in present and future climate. *Journal of Climate*, 33(23), 10021-10038  
581 (2020).  
582

583 [30] Kornhuber, K., et al. Extreme weather events in early summer 2018 connected by a  
584 recurrent hemispheric wave-7 pattern. *Environmental Research Letters*, 14(5), 054002 (2019).  
585

586 [31] Suarez-Gutierrez, L., Li, C., Müller, W. A., & Marotzke, J. Internal variability in European  
587 summer temperatures at 1.5 C and 2 C of global warming. *Environmental Research Letters*,  
588 13(6), 064026 (2018).  
589

590 [32] Jézéquel, A., Yiou, P., & Radanovics, S. Role of circulation in European heatwaves using  
591 flow analogues. *Climate dynamics*, 50(3), 1145-1159 (2018).  
592

593 [33] Faranda, et al. A climate-change attribution retrospective of some impactful weather  
594 extremes of 2021, *Weather Clim. Dyn.*, 3, 1311–1340, [https://doi.org/10.5194/wcd-3-1311-](https://doi.org/10.5194/wcd-3-1311-2022)  
595 2022 (2022).  
596

597 [34] Deser, C., A. Phillips, M. A. Alexander, and B. V. Smoliak. Projecting North American  
598 climate over the next 50 years: Uncertainty due to internal variability. *J. Climate*, 27, 2271–  
599 2296 (2014).  
600

601 [35] Sippel, S., et al. Uncovering the forced climate response from a single ensemble member  
602 using statistical learning. *Journal of Climate*, 32(17), 5677-5699 (2019).  
603

604 [36] Wilks, D. S. (2006). On “field significance” and the false discovery rate. *Journal of applied*  
605 *meteorology and climatology*, 45(9), 1181-1189.  
606

607 [37] Wilks, D. (2016). “The stippling shows statistically significant grid points”: How research results  
608 are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American*  
609 *Meteorological Society*, 97(12), 2263-2273  
610

611 [38] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and  
612 powerful approach to multiple testing. *Journal of the Royal statistical society: series B*  
613 *(Methodological)*, 57(1), 289-300.

614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644

[39] Qasmi, S., Cassou, C., & Boé, J. Teleconnection between Atlantic multidecadal variability and European temperature: Diversity and evaluation of the Coupled Model Intercomparison Project phase 5 models. *Geophysical Research Letters*, 44(21), 11-140 (2017).

[40] McKinnon, K. A., & Deser, C. Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, 31(17), 6783-6802 (2018).

[41] O'Reilly, C.H., Befort, D.J., Weisheimer, A. et al. Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty. *Commun Earth Environ* 2, 194 (2021). <https://doi.org/10.1038/s43247-021-00268-7>

[42] Blackport, R., & Fyfe, J. C. Climate models fail to capture strengthening wintertime North Atlantic jet and impacts on Europe. *Science Advances*, 8(45), eabn3112 (2022).

[43] Merrifield, A., F. Lehner, S.-P. Xie, and Deser, C. Removing circulation effects to assess central US land–atmosphere interactions in the CESM large ensemble. *Geophys. Res. Lett.*, 44, 9938–9946 (2017).

[44] Smoliak, B. V., J. M. Wallace, P. Lin, and Fu, Q. Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. *J. Climate*, 28, 1613–1629 (2015).

[45] Saffioti, C., E. M. Fischer, and Knutti, R. Improved consistency of climate projections over Europe after accounting for atmospheric circulation variability. *J. Climate*, 30, 7271–7291 (2017).

[46] Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E., & Knutti, R. Climate change now detectable from any single day of weather at global scale. *Nature climate change*, 10(1), 35-41 (2020).

## 645 **Supplementary Information**

646

### 647 **Observation and model data**

648

649 We used ERA5 reanalysis of daily maximum temperatures and streamfunction fields.  
650 Streamfunction is calculated from u- and v- wind fields at 500 hPa on a T127 Gaussian grid,  
651 and then interpolated on a 1x1 regular grid, following:

652  $u = -\frac{\partial\psi}{\partial y}$ , and  $v = \frac{\partial\psi}{\partial x}$ , where  $\psi$  is the streamfunction,  $u$  is the zonal- and  $v$  the meridional

653 component of the wind fields.

654

655 Surface daily maximum temperatures from ERA5 are interpolated to a 0.5x0.5 grid. We also  
656 used observations from the E-OBS dataset v24e [20] for daily maximum temperature (TX). E-  
657 OBS was initially taken from a 0.25 x 0.25 grid and projected onto the 0.5x0.5 grid. When  
658 considering averages over the selected Western Europe area [5W-15E;45N-60N], data are  
659 masked using the E-OBS land/sea mask (see below).

660

661 Daily maximum temperatures and streamfunction are also calculated from model simulations  
662 including all first members of each CMIP6 model ensemble. In order to increase as much as  
663 possible the estimation of capacity of models to simulate TXx and TXm trends, we used all  
664 possible CMIP6 simulations made available through the ESGF infrastructure. When  
665 considering only TXx and TXm calculations for Figure 3, we used 273 simulations made with  
666 36 different models (see Figure 3). For Figure 4 and the analogue analysis for models, we keep  
667 only 32 models and 1 realization for which we have simultaneous 500 hPa wind fields and  
668 daily maximum temperatures.

669

670 To have an historical time series to be compared with reanalysis or observations, we  
671 concatenate historical and SSP5-8.5 scenarios available (from 2015 to 2022). Initial tests made  
672 with SSP2-4.5 showed that results presented here are insensitive to this choice.

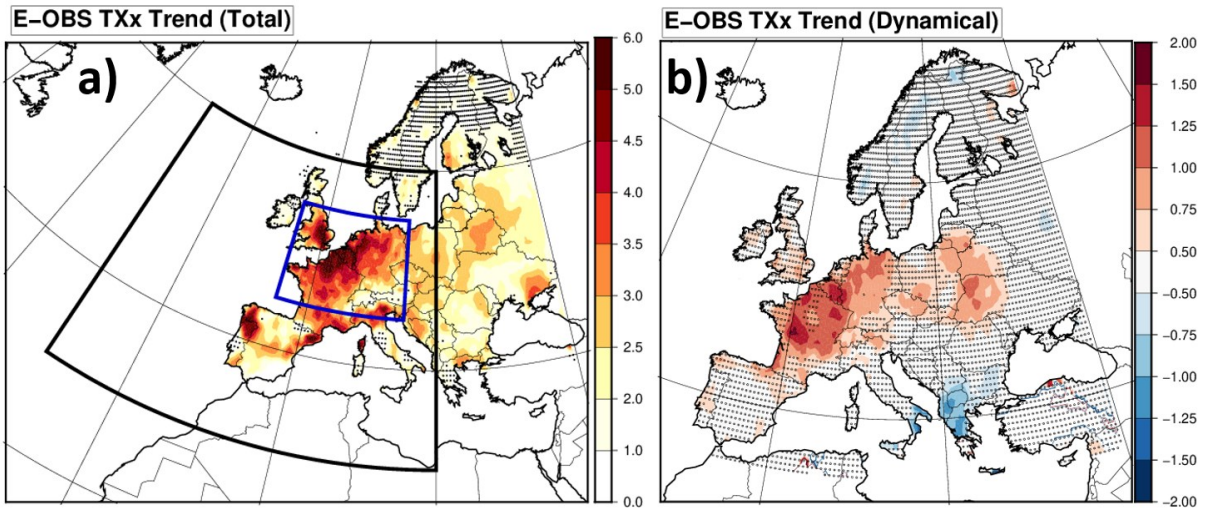
673

674

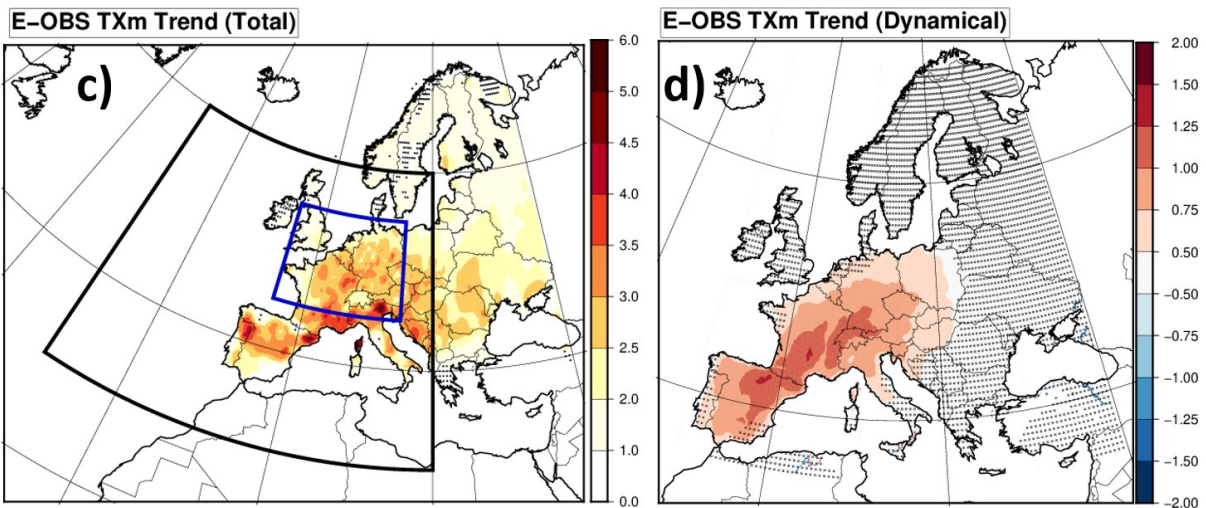


675 Extended Data

676



677



678

679 Extended Data Figure 1: Same as Figure 1 but for E-OBS maximum daily temperatures

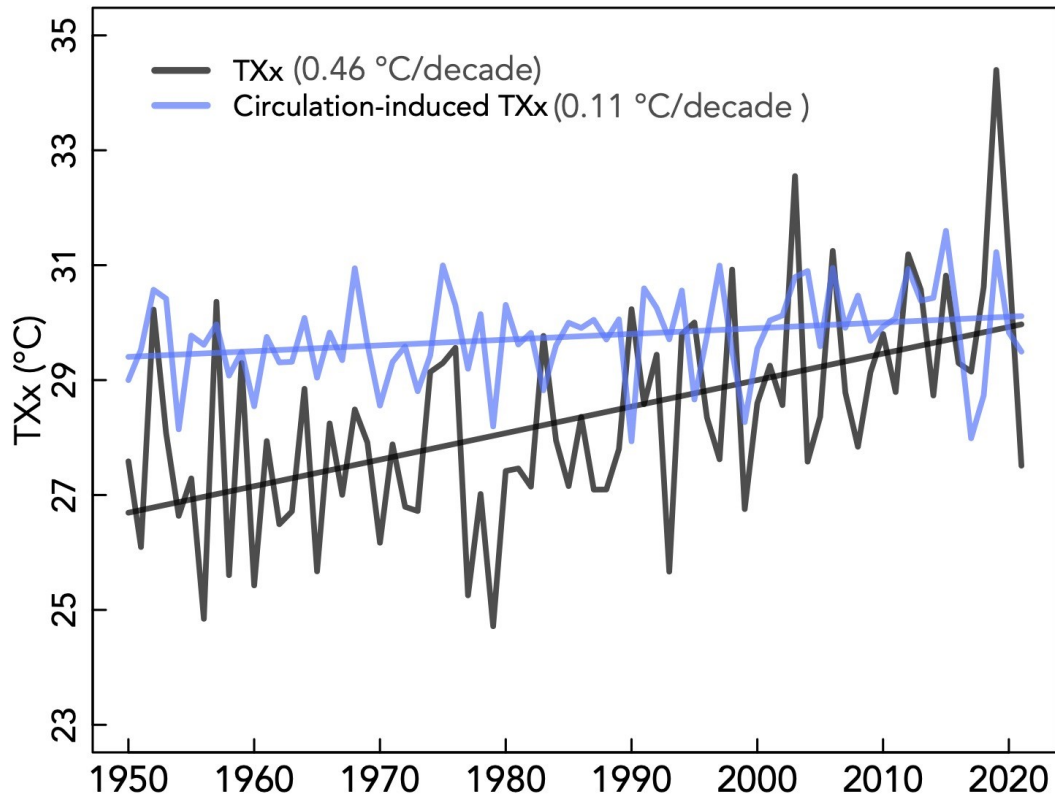
680

681

682

683

684



685

686

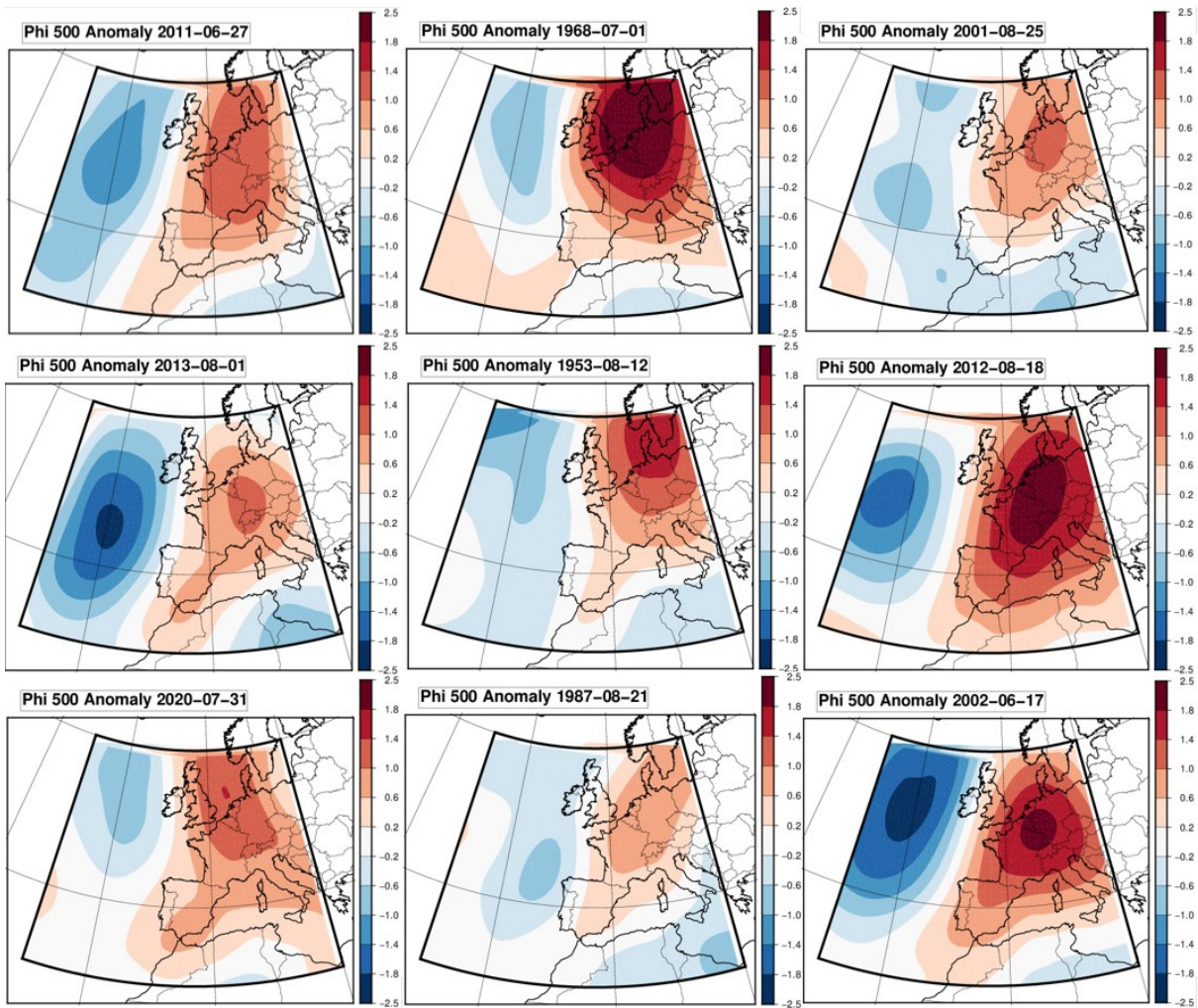
687 **Extended Data Figure 2:** Dynamical contribution to forced trend during 1950-2021 in ERA5

688 reanalysis. Black and blue lines present the area-averaged TXx and circulation-induced TXx

689 over western Europe (5-15E, 45-55N), respectively. The values in parenthesis indicate the trend

690 in the corresponding TXx time series. The trends are estimated based on Sen's slope estimator.

691



692

693 **Extended Data Figure 3:** 500 hPa streamfunction anomalies of the 9 most representative  
 694 circulations, beyond 29/06/2019, when TXX is reached over Central France [1.5E-46.5N], by  
 695 decreasing order of representativeness.

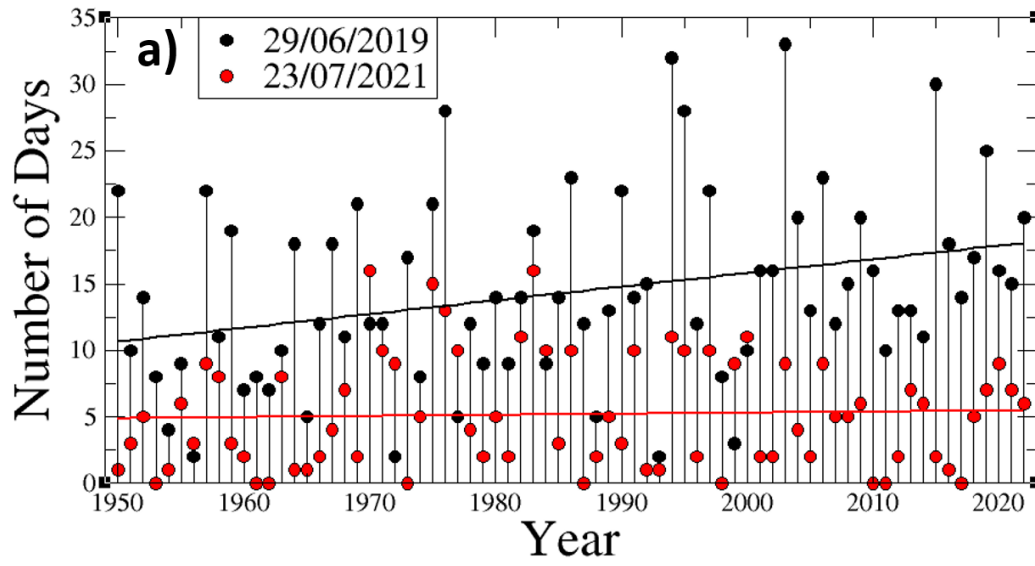
696

697

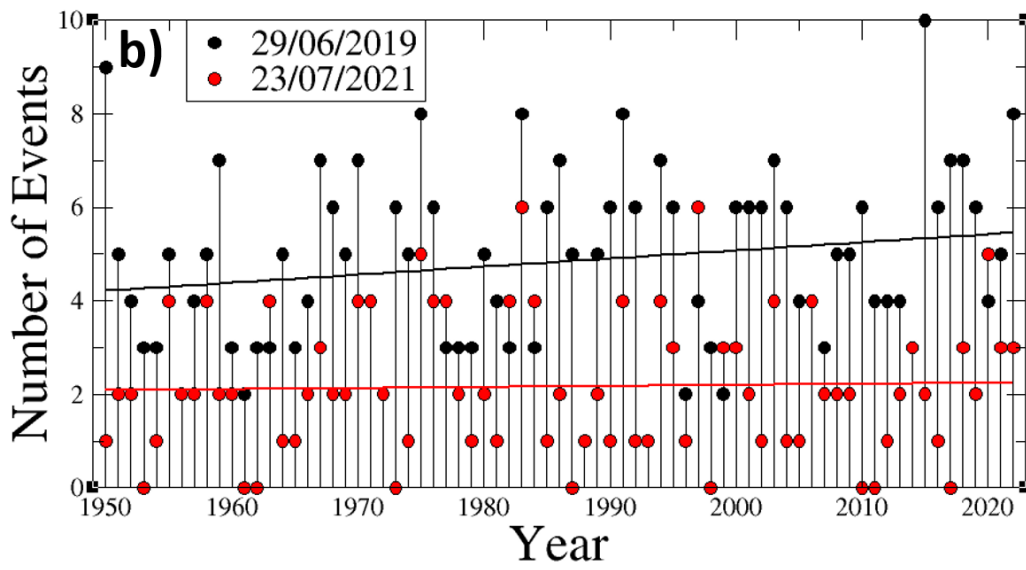
698

699

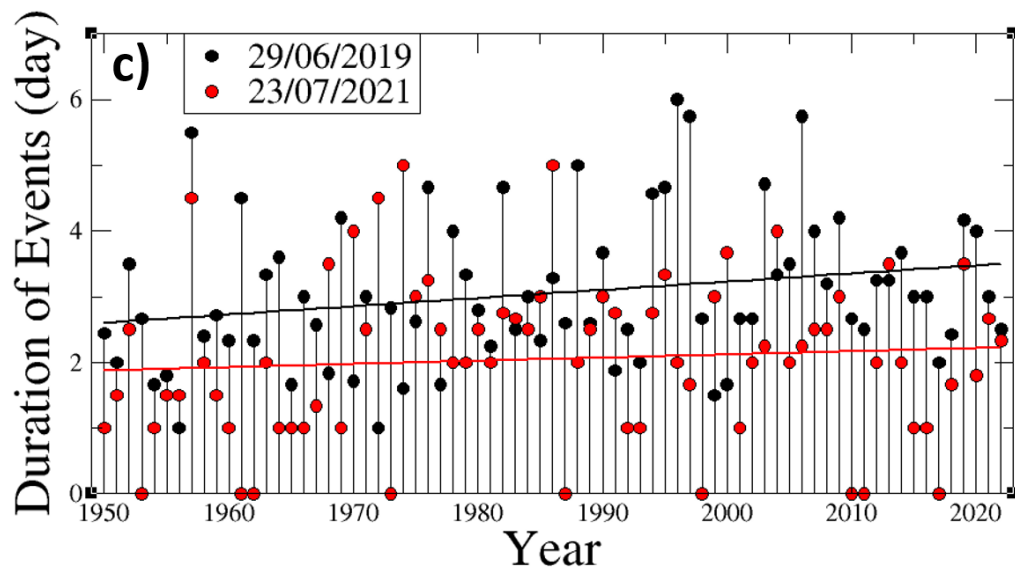
700



701



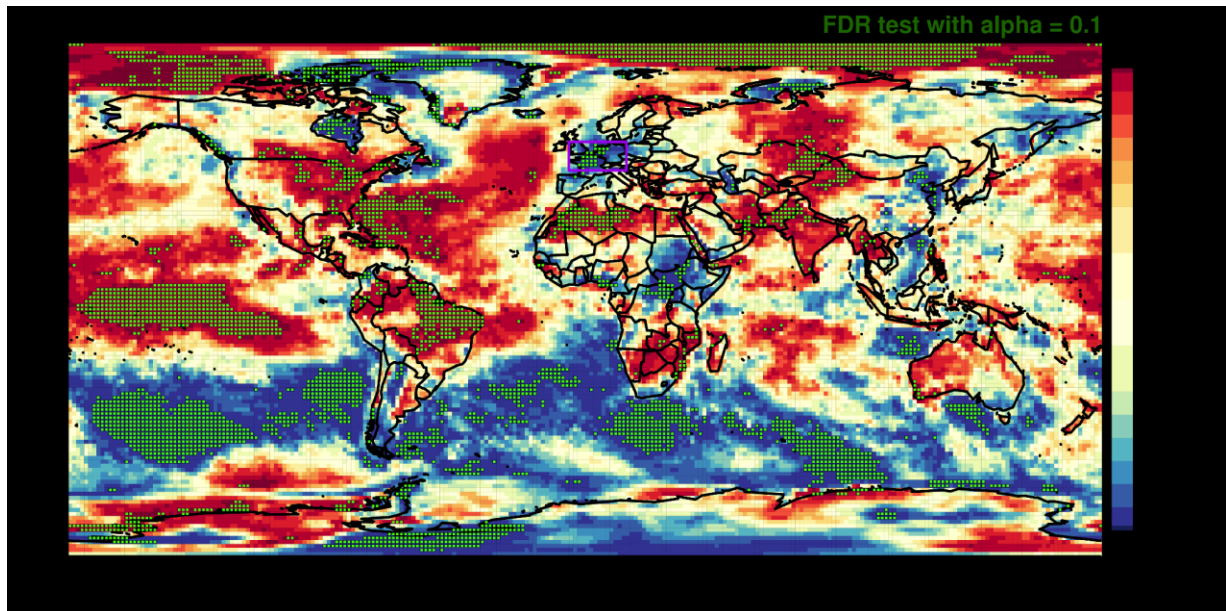
702



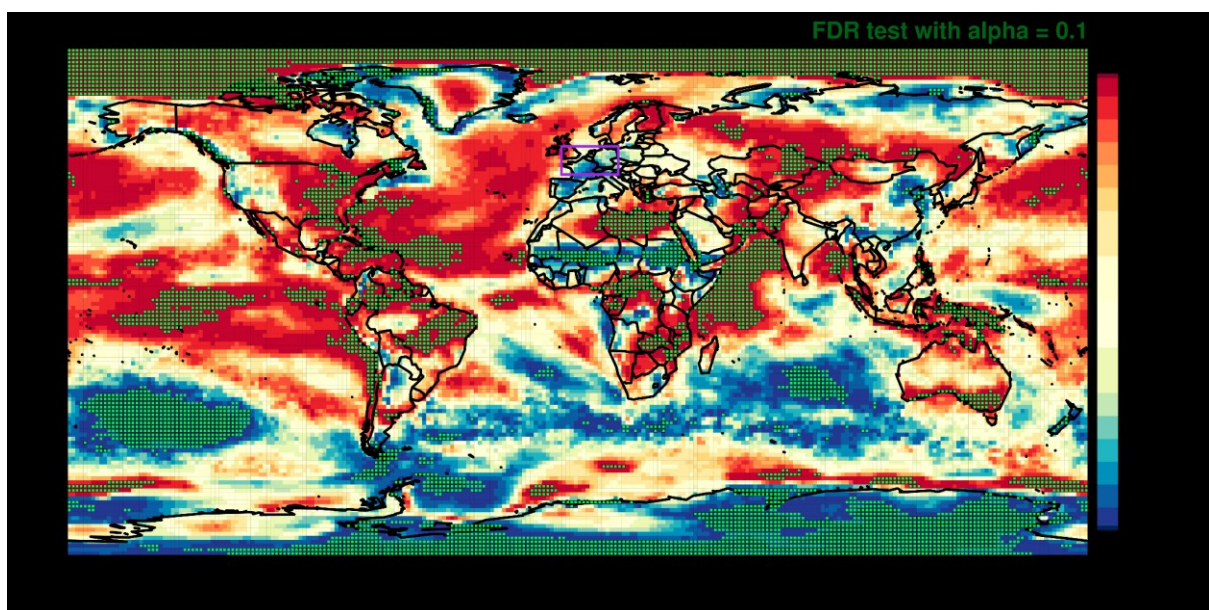
703

704 **Extended Data Figure 4:** Evolution of the yearly number of days (a), number of events (b)  
705 and mean duration of events (c) (0 when no event found) for Southerly Flow patterns (black)  
706 (streamfunction anomalies with an ACC with the 29/06/2019 anomaly greater than 0.5). For  
707 comparison, the figure also shows (in red) the same statistics but for another pattern (not  
708 shown), that of the anomaly of the 23/07/2021, which corresponds to the date of TXx for 2021  
709 in central France.

710



711



712

713

714 **Extended Data Figure 5:** Generalization of Figure 3 at global scale. Percentage of simulations  
715 with a trend larger than ERA5 at each grid point for (top) the annual maximum of TX and  
716 (bottom) the JJA mean of TX. Green stippling indicates grid points where the mismatch

717 between observed and simulated trends is significant at the 95% confidence level in the sense  
 718 of the False Discovery Rate procedure (i.e. a two-sided multiple test with  $\alpha=0.1$ ). The  
 719 Western Europe box is highlighted in magenta. For the top panel, the annual (rather than JJA  
 720 in Figure 3) maximum of TX is used here to capture summer heat extremes in both  
 721 hemispheres; in Western Europe annual or JJA maximum are equivalent.

722  
 723

Model_realization	2019 06/29	2011 06/27	1968 07/01	2001	2013	1953	2012	2020	1987	2002
<b>ERA5</b>	<b>42.9</b>	<b>51.4</b>	<b>34.5</b>	<b>43.9</b>	<b>48.7</b>	<b>35.8</b>	<b>52.5</b>	<b>54.1</b>	<b>55.0</b>	<b>47.8</b>
ACCESS-CM2_r1i1p1f1	-16.4	-26.9	-28.1	-23.3	-18.7	-27.5	-26.8	-29.9	-30.4	-23.2
ACCESS-ESM1-5_r2i1p1f1	7.0	-7.2	1.4	-6.5	-4.8	9.2	-8.3	-2.6	-11.6	-5.8
CAMS-CSM1-0_r2i1p1f1	-1.0	13.6	2.5	0.7	9.6	8.8	14.2	17.7	-0.6	16.5
CanESM5_r1i1p1f1	10.6	0.2	3.0	9.7	-1.4	4.7	-6.7	-0.5	16.2	2.3
CMCC-ESM2_r1i1p1f1	2.6	11.3	6.9	5.9	17.6	11.6	7.9	6.4	22.0	18.9
CNRM-CM6-1-HR_r1i1p1f2	23.2	17.6	27.3	25.8	17.7	26.8	21.6	24.1	29.1	17.0
CNRM-CM6-1_r1i1p1f2	-10.6	3.8	-6.3	-12.9	2.6	-9.4	6.2	6.3	-12.3	7.3
CNRM-ESM2-1_r1i1p1f2	6.0	0.5	4.8	4.9	-2.4	0.6	0.3	-2.0	-6.1	5.8
EC-Earth3_r1i1p1f1	14.2	7.3	6.9	9.6	5.6	4.8	4.9	2.8	17.8	3.4
EC-Earth3-CC_r1i1p1f1	-14.4	-14.8	-17.6	-12.2	-17.1	-12.2	-17.1	-13.0	-4.8	-12.8
EC-Earth3-Veg_r1i1p1f1	-5.3	-10.9	-17.4	-9.5	0.0	-15.2	-10.3	-14.7	-14.8	-14.5
EC-Earth3-Veg-LR_r1i1p1f1	-9.3	-4.3	-1.1	1.8	13.1	-2.6	4.7	-3.1	-14.3	-0.5
FGOALS-g3_r1i1p1f1	-53.1	-48.6	-52.9	-50.0	-45.5	-53.6	-47.6	-57.5	-45.1	-48.1
GFDL-CM4_r1i1p1f1	-1.7	4.1	-8.5	-6.0	-2.3	-9.1	0.3	2.5	-5.6	9.3
GISS-E2-1-G_r1i1p1f2	-16.3	-30.3	-32.1	-29.7	-7.9	-28.7	-14.3	-26.7	-39.3	-28.5
HadGEM3-GC31-LL_r1i1p1f3	15.9	7.4	18.4	11.1	2.9	16.1	9.0	15.9	12.6	14.9
HadGEM3-GC31-MM_r1i1p1f3	13.6	14.2	16.1	17.9	9.6	17.5	3.1	9.5	34.5	13.3
INM-CM4-8_r1i1p1f1	16.4	22.5	6.6	15.5	23.6	8.5	18.8	20.2	6.2	22.5
INM-CM5-0_r1i1p1f1	0.7	-1.3	-4.3	-5.5	-5.1	-6.7	5.8	-7.5	11.1	-2.0
IPSL-CM6A-LR_r1i1p1f1	5.3	-24.1	-16.8	-8.9	-12.5	-9.3	-18.5	-21.6	-19.2	-26.6
KACE-1-0-G_r1i1p1f1	17.0	18.0	19.7	25.9	14.2	25.4	12.1	15.0	34.0	6.8
KIOST-ESM_r1i1p1f1	8.4	7.7	20.3	19.9	-3.4	22.2	-0.4	10.9	23.9	8.1
MIROC-ES2L_r1i1p1f2	13.6	0.2	14.2	24.0	-2.2	24.6	-5.6	11.4	18.9	-8.0
MIROC6_r1i1p1f1	20.1	12.5	23.6	22.9	25.7	14.7	18.3	20.1	7.0	14.9
MPI-ESM1-2-HR_r1i1p1f1	8.0	-0.7	9.4	8.8	6.3	14.8	-3.4	0.5	16.3	1.7
MPI-ESM1-2-LR_r1i1p1f1	2.6	4.1	2.5	2.0	-6.8	2.9	-6.3	-2.4	18.5	-2.3
MRI-ESM2-0_r1i1p1f1	10.7	-7.3	6.5	-0.6	5.9	-1.2	2.9	-9.8	-13.0	-16.3
NESM3_r1i1p1f1	17.8	10.5	14.4	15.1	9.5	14.4	8.7	12.1	23.5	7.7
NorESM2-LM_r1i1p1f1	42.7	24.6	39.0	41.5	11.0	41.8	23.6	26.5	36.2	10.6
NorESM2-MM_r1i1p1f1	-8.7	-22.3	-10.8	-9.6	-25.7	-10.4	-32.0	-32.4	-11.1	-30.5
TaiESM1_r1i1p1f1	-4.2	-3.4	2.7	-5.6	-13.3	3.3	-7.9	2.7	2.0	-0.3
UKESM1-0-LL_r1i1p1f2	-13.1	-19.3	-18.4	-14.4	-20.1	-16.7	-21.6	-21.4	-16.0	-25.3

724  
 725 **Extended Data Table 1:** Models used for the dynamical temperature trend analysis and the  
 726 analysis of Southerly Flows (SF) frequency trends (and reanalysis, first line); the name  
 727 indicates the model and the realization, with ssp585 concatenated after 2014 to historical  
 728 simulation in order to have a continuous time series from 1950 to 2022; Second to last column:  
 729 SF frequency change over 1950-2022 as a function of the GWD. In each column, SFs are  
 730 defined as “ACC with the reference pattern greater than 0.5”, with the reference patterns among  
 731 one of the most representative ones represented in Figure 2a and Extended Data Figure 3.

732

733

Model_realization	TXx trend	TXx dynamical trend	Model_realization with largest TXx trend	TXx trend	TXx dynamical trend
<b>ERA5</b>	<b>3.38</b>	<b>0.79</b>			
<b>E-OBS</b>	<b>3.41</b>	<b>0.86</b>			
ACCESS-CM2_r1i1p1f1	1.55	-0.27	ACCESS-CM2_r4i1p1f1	1.89	0.10
ACCESS-ESM1-5_r2i1p1f1	3.19	0.02	ACCESS-ESM1-5_r34i1p1f1	3.42	0.35
CanESM5_r1i1p1f1	1.89	0.23	CanESM5_r6i1p2f1	2.02	0.46
EC-Earth3_r1i1p1f1	1.49	0.16	EC-Earth3_r4i1p1f1	1.49	0.24
EC-Earth3-Veg_r1i1p1f1	1.14	-0.12	EC-Earth3-Veg_r4i1p1f1	1.42	-0.29
EC-Earth3-Veg-LR_r1i1p1f1	0.65	0.20	EC-Earth3-Veg-LR_r3i1p1f1	1.16	-0.15
FGOALS-g3_r1i1p1f1	0.26	-0.67	FGOALS-g3_r4i1p1f1	0.87	-0.11
HadGEM3-GC31-LL_r1i1p1f3	1.93	0.30	HadGEM3-GC31-LL_r4i1p1f3	2.49	0.02
KACE-1-0-G_r1i1p1f1	2.65	0.33	KACE-1-0-G_r3i1p1f1	2.80	0.44
MPI-ESM1-2-LR_r1i1p1f1	2.06	0.33	MPI-ESM1-2-LR_r3i1p1f1	2.09	0.20
MRI-ESM2-0_r1i1p1f1	1.87	-0.12	MRI-ESM2-0_r5i1p1f1	1.89	0.07
UKESM1-0-LL_r1i1p1f2	1.35	0.01	UKESM1-0-LL_r2i1p1f2	2.25	-0.05

734

735 **Extended Data Table 2: TXx overall and dynamical trends for ERA5, E-OBS, and models**  
736 **for which a member of the ensemble with largest overall trend in available simulations in**  
737 **addition of the first member. Trends are expressed in °C/GWD.**

738

739