Responses to referee 1 comments about the article
"*Spatial structures of emerging hot & dry compound events over Europe from 1950 to 2023*"
By Schmutz, Vrac, François, Bulut

Summary

The objective of this article is to propose a novel methodology for quantifying the emergence of statistically significant compound events and then apply this approach to extreme temperature and drought conditions across Europe estimated using ERA5 reanalysis. The authors define a Period of Emergence to quantify whether the probability that the dependence between compound events is temporarily statistically significant (Period of Emergence) or permanently statistically significant (Time of Emergence; an existing method within time series analysis). They then compare these probabilities under two assumptions: that the dependence structure between compound hazards does not change over time, and that this dependence structure does change over time. Finally, they quantify the relative contribution of each component (the drought index, the temperature index, and their dependence structure) to the Time of Emergence and Period of Emergence over time. They compare findings across five regions in Europe. They make an R package available for researchers to conduct similar analyses on any compound event pair. They explore the impact of two signals on the Time of Emergence and Period of Emergence for compound event probability. One, in black, is the probability of compound event if we take the dependence structure between temperature and drought to be constant over time. The other, in blue, is the probability of compound event if we take the dependence structure between temperature and drought to change over time. When either signal temporarily exceeds a confidence interval of noise, call it a period of emergence. When either signal permanently exceeds a confidence interval of noise, call it a time of emergence. Note that the signal is weakened under the assumption of constant dependence structure and is strengthened under the assumption of changing dependence structure.

Overall, this manuscript is well-written and provides a compelling argument for the inclusion of this novel methodology into compound event research. They find that a statistically significant compound event signal emerges earlier in time than previously recognized when 1) the dependence structure is allowed to change over time and 2) a period of emergence is quantified. They quantify how the contribution to these compound events has changed over time, allowing for regional and location-specific applications to compound events. This research contributes to the field of compound and multi-hazard research because it expands our definition of compound events using a novel statistical approach that can be readily applied to any event pair. The manuscript follows a logical structure and identifies potential approaches for future research that would enable more than two hazards to be simultaneously analyzed for their time-varying dependence. There are few clarifying, structural, and editorial changes needed prior to accepting this manuscript. Comments and suggestions are provided below and organized into major and minor.

Response:

We appreciate the time and effort that the reviewer dedicated to providing feedback on our manuscript and are grateful for the insightful comments on our paper. We have incorporated most of the suggestions. Those changes are highlighted within the manuscript. Please see below, in blue, for a point-by-point response to the comments and concerns.

Major comments

- Generally lacking helpful descriptions of each figure, see minor comments for suggestions.

Response

We have taken all the reviewer's comments into account. We now think that the figures are better described. See the answers to each one in the "minor comments" section.

- Section 5.1 – I'm missing a takeaway from this section. What do we now understand if we consider period of emergence and / or changing dependence structure between compound events. Simply that some areas are more at risk of compound hazard than previously expected? That more compound events can be attributed to a dependence structure or when a period of emergence is considered? How does this affect researchers and exposed communities? How can researchers, insurance providers, community planners, etc. use the results of this study to improve outcomes from future compound events? Can we link this kind of analysis with research exploring the drivers of compound events?

Response:

The reviewer is right, our manuscript was initially missing a clear takeaway, which is essential to highlight the main focus of the study. We have now included a concise takeaway at the end of the section (on lines 452–456 of the revised manuscript), as follows:

> *All in all, this paper identifies hotspots of emerging hot and dry compound events (CEs) across Europe. Most regions show a permanent and significant increase in CE probability, while some areas such as Ireland and parts of Eastern Europe exhibit a significant recent and old decrease respectively. In most cases, increases in CE frequency are primarily driven by changes in the heat index, whereas decreases are mainly linked to variations in the drought index. Importantly, the study demonstrates that the statistical dependence between heat and drought is a key factor in detecting the emergence of compound events.*

In addition, the reviewer rightly pointed out the importance of clarifying the relevance of this work for different sectors. We have addressed this helpful suggestion in the beginning of the *Perspectives* section (lines 458–473 of the revised paper), where we now emphasize the study's potential applications for scientific research, exposed communities and decision-makers:

> *The emergence of a climate signal is a valuable impact indicator, as it marks the point when the signal exceeds natural variability—beyond which human and ecological systems are typically adapted. Compound events, especially those combining heat and drought, have severe impacts on society, water resources, and agriculture (Ribero et al.). Therefore, identifying when, for how long, where, and to what extent the occurrence of these high-impact events has changed is of great interest to many communities. Stakeholders and farmers can benefit from such studies in designing*

adaptation strategies. For instance, Lesk et al. (2022) emphasize the importance of developing agricultural adaptation policies tailored to different spatial and temporal scales of hot and dry events.

This type of research is also valuable for climate physicists, hydrologists, and agronomists, who aim to understand the physical and physiological mechanisms underlying significant increases or decreases in these compound events. We analyzed how changes in statistical dependence affect the emergence of compound events. While it's well established that there is a physical link between heat and drought, particularly through land–atmosphere feedbacks, (Miralles et al., 2014), it is essential to further explore how changes in statistical dependence reflect underlying physical processes. In terms of atmospheric drivers, Ionita et al. (2021) demonstrated that hot and dry events across Europe are frequently associated with atmospheric blocking patterns. It would be valuable to connect statistical contributions with physical drivers: how might the spatial patterns of the contributions (of the three components forming the CE) relate to soil characteristics, oceanic influences, or atmospheric dynamics?

We just want to highlight that the submitted article does not look at changing dependence structure between compound events, but instead the dependence between the variables that constitute a single compound event, in this study hot and dry events. This notion of dependence between variables constituting the CE is specified in line 136 ("*dependence between variables*") and line 156 ("*dependence between hydroclimatic variables"*) of the originally submitted article.

- Section 5.2 – how would your analysis change if you considered aggregating your indices at different temporal resolutions? Ie. daily, weekly, monthly max values? Min values? What is the sensitivity of your analysis to your chosen variables and chosen resolutions? Would the application of this method to climate model simulations affect the confidence interval of noise?

Response

The method developed in this study is generic and can be applied to any bivariate compound event. Modifying the variables, spatial scale, or temporal resolution will naturally lead to different compound event probabilities, and consequently, different outcomes in terms of emergence and component contributions.

We used the 6-month Standardized Precipitation Evapotranspiration Index (SPEI6) to represent drought, as it captures the water balance over six months and is well-suited for identifying agricultural drought (Ionita and Nagavciuc, 2021). The analysis was performed at a monthly time scale to match the temporal resolution of this index. We have added this clarification on lines 120-125 of the revised paper:

The [SPEI] is based on the difference between precipitation and evapotranspiration, reflecting the water balance. The SPEI indicator, considered to be one of the best for *drought monitoring (e.g., Ionita and Nagavciuc, 2021; Blauhut et al., 2016), is used for the present study as it combines the advantages of SPI, with its variety of timescales, and those of PDSI with the consideration of temperature evolution. Depending on the chosen accumulation period, it can capture the different types of drought (Ionita and Nagavciuc, 2021). Then the analysis is performed at a monthly time scale to match the temporal resolution of the SPEI.*

The reviewer highlighted an interesting avenue for future work: exploring how changing variables affects the outcomes. We have incorporated this perspective on lines 474-479:

"The method developed in this study is generic and applicable to any bivariate compound event. Changing the variables, spatial scale, or temporal resolution will naturally change the event probabilities and, in turn, highlight different patterns of emergence and contribution. Further analysis of other bivariate events, such as extreme precipitation and wind, or alternative representations of hot and dry conditions (e.g., minimum temperature with precipitation, or the number of hot days with SPEI3), across different spatial and temporal scales, could yield valuable insights for future research tailored to the needs of societal stakeholders."

Regarding model simulations, they will likely result in different distribution fits compared to reanalysis data, leading to different confidence intervals for natural variability. We recommend referring to François and Vrac (2023), who investigated the Time of Emergence (ToE) of compound events using simulations. Their study spans a longer period (1871–2100) and relies on 13 climate models to estimate ToE, either by computing the median or by pooling the contributing variables. We have reformulated the second-to-last paragraph of the Perspectives section, lines 506-511 of the revised article:

The ERA5 reanalysis dataset is analysed for the present case study in order to detect past changes. However the period is only available from 1950 to 2023. Using simulations from climate models can shed new lights on multivariate detection and attribution framework. Simulations provide access to longer datasets, enabling the analysis of extended PoE trends. They offer insight into how climate models represent natural variability compared to reanalysis data. Analysing simulations can be used to evaluate CMIP6 models ability to retrieve compound events emergence. If simulations detect the right frequency of PoE without the right main driver, the reliability of the model may be weakened.

Minor comments

Line

- 55 needs a citation for ‘most damaging impacts’.

Response:

Zscheischler et al. (2020) underscore the damaging impacts of compound events, noting in the key points of their paper that “Compound events — a combination of multiple drivers and/or hazards that contribute to societal or environmental risk — are responsible for many of the most severe weather-related and climate-related impacts.”

We have added the reference to Zscheischler, et al. (2020) in the manuscript after the mention of “the most damaging impacts” one line 52 of the revised article :

> *“As compound events contribute to the most damaging impacts (Zscheischler., et al., 2020), the question of multivariate emergence has recently arisen.”*

- 60-61 explain why covariance matrix is not appropriate for CE.

Response:

Using a covariance matrix to model dependence often implies a Gaussian assumption, as it captures only linear relationships. However, compound event analysis focuses on tail dependencies, which are typically non-linear and poorly represented by Gaussian models (Hao and Singh, 2016). This makes the covariance matrix not appropriate for CE. Hao and Singh (2016), Abatzoglou et al., (2020), and Tootoonchi et al., (2022) highlight the limitations of using Gaussian assumption and advocate for the use of copulas, in particular for compound event studies.

We added an explanation on lines 59–64 of the revised article :

> *“However the dependence between the variables is assumed to be Gaussian, i.e., fully characterized by a covariance matrix. This implies that only linear relationships are captured, and any nonlinear or asymmetric dependence — particularly in the extremes — are ignored. This approach is not well suited for analyzing compound extremes, where the focus lies on understanding how variables co-occur specifically in the tails of the distributions (e.g., Zscheischler et al., 2020; Tootoonchi et al., 2022). Capturing such behavior requires more flexible dependence structures beyond the Gaussian framework (Hao and Singh, 2016).”*

- Paragraph starting on 48 is very difficult to follow, consider rewriting from line 55 to the end, potentially break up into a few paragraphs, remove excessive detail and streamline.
    - I missed the explanation as to why a gaussian assumption of the dependence between variables is not appropriate for compound extremes.

Response:

This question relates to the previous one. You can find a rewording of the whole paragraph on lines 48-66 of the revised paper, below:

> *Emergence of multivariate events remains largely underexplored. The great majority of studies analyses the emergence at a global scale and for univariate variables: mainly temperature (diffenbaugh et al., 2011, Mahlstein et al., 2011, Hawkins et al., 2012) and precipitation (Giorgi et al., 2009, Fischer et al., 2014, Murphy et al., 2023), but also drought index (Osso et al., 2022}, fire weather index (Abatzoglou et al., 2019), sea level (Yu et al., 2014) and biogeochemical cycle (Keller et al., 2014). As compound events contribute to the most damaging impacts (Zscheischler et al., 2020), the question of multivariate emergence has recently arisen. The understanding of past and future changes of compound events occurrence~~s~~ is of great importance for adaptation planning.. The goal is now to detect when a multivariate distribution is statistically different from a baseline period. Williams et al., (2007) used the standardized Euclidean distance to quantify the differences between two climates in the 20th and 21st~~th~~ centuries. Mahony et al. (2017) adapted this metric to take the covariance between variables into account with the Mahalanobis distance. The latter, transformed into percentiles of the chi distributions, is called "sigma dissimilarity" and has been used to identify multivariate climate departures (Abatzoglou et al., 2020, Mahony et al., 2018). However the dependence between the variables is assumed to be Gaussian, i.e., fully characterized by a covariance matrix. This implies that only linear relationships are captured, and any non-linear or asymmetric dependence — particularly in the extremes — are ignored. This approach is not well suited for analyzing compound extremes, where the focus lies on understanding how variables co-occur specifically in the tails of the distributions (Zscheischler. et al, (2020), Tootoonchi. et al (2022). Capturing such behavior requires more flexible dependence structures beyond the Gaussian framework (Hao and Singh, 2016). That is why, François and Vrac, (2023) defined a time of emergence (ToE) applicable to multivariate events, as the year from which the compound event probability (the signal) is always out of the natural variability. In this method, the signal is quantified with bivariate copula, which allows a modeling of a not Gaussian dependence.*

- 254 componen is missing a t; perhaps remove ‘even’ before ‘negative’, or find another way to highlight this otherwise counterintuitive statement.

Response:

We added a “t” line 259, and we removed “even”. We rephrased the sentence on lines 258-260 of the revised paper :

> *“Concerning the dependence, as the probability associated to the change of this component decreases until the lower bound of the natural variability, evolving in the opposite direction to p, its contribution during PoE-up is negative”.*

- 271 Perhaps outline this argument as a counterfactual. When I first read this section, I thought your argument was that it is safe to assume that dependence can be kept constant. But I see that you are setting up the argument that this assumption is invalidated through this modeling framework and that you indeed have similar findings to Wang et al., 2021 in the prior sentence.

Response:

Wang et al. (2021) emphasize the role of stronger dependence in increasing the occurrence of hot and dry events. In this context, our study investigates how changes in dependence can affect the emergence of compound events probability. The aim of section 3.2 is to examine the influence of dependence changes on the signal emergence by exploring what happens if the copula parameter is constant or evolves. We incorporated some clarification at the end of the paragraph on lines 276-277 of the revised article:

> *"Then, this section investigates how changes in dependence can affect the emergence of compound events probability."*

- 289 please elaborate on the line 'not considering the dependence evolution would advance the signal emergence'. Does this suggest that signal emergence occurs earlier in the record under the assumption of constant dependence? I think rewording this sentence could provide needed clarity.

Response:

Exactly: the sentence "*Not considering the dependence evolution would advance the signal emergence*" means that signal emergence occurs earlier under the assumption of constant dependence. We rephrased the sentence using these words for more clarity, on line 296 of the revised article:

> *"Hence, for Vilnius, the emergence occurs earlier under the assumption of a constant dependence."*

- 415 please include something like (78% of the grid points observe a time of emergence prior to the end of the temporal record) for clarity.

Response:

We added this clarification to the sentence, on line 425-426 of the revised article:

> *The signal permanently emerged above the natural variability in the majority of the area (78% of the grid points show a time of emergence prior to the end of the temporal record).*

- 420 for my own interest, do any points experience a period of emergence, but then return to within the noise confidence interval? If so, where do those locations occur?

Response:

This is an interesting question. This is related to whether there are any grid points that exhibit a PoE and no ToE.

Every grid point experiences a period of emergence (either PoE-low or PoE-up). However, not every grid point exhibits a significant change persisting until the end of the dataset and that is precisely one of the key findings of this study.

Specifically, 78% of the grid points show a ToE-up (colored points in Figure 3a), and 4% show a ToE-low (colored points in Figure 3b). This means that 18% of the grid points experience neither a ToE-up nor a ToE-low. The 18% correspond to the white areas in Figure 3a, excluding the few colored pixels visible in Figure 3b. These regions include, for instance, Great Britain, parts of Norway, and northern Russia. We revised the first paragraph of Section 4.1 in the Results to incorporate this point, lines 314-324 of the revised paper :

> *Spatial patterns of emergence are presented Figure 3. Hot and dry events occurrences emerged in most part of Europe and North Africa. 78% of the grid points show a ToE-up (colored points in Figure 3a) and 4% show a ToE-low (colored points in Figure 3b). In other words, 18% of the grid points show no ToE-up or ToE-low, suggesting that at these locations, the signal ultimately returns within the range of natural variability. The 18% correspond to the white areas in Figure 3a (no significant permanent rise), excluding the few colored pixels visible in Figure 3b (significant decrease). The soonest ToE-up occurred in the region MA-IB, in Scandinavia and in Russia, which means that they have been affected by a significant increase of compound hot and dry events for several decades (Figure 3a). In the whole MA-IB area, ToE-up is detected very early, even before 1970, in contrast to Scandinavia where CE occurrence do not evolve similarly over the region. In Ireland and in the eastern region (EAST), the probability has significantly decreased until the end of the study period, as shown by the coloured pixels Figure 3b. Results of ToE can be compared to risk ratio (RR) values (Figure 3 in Supplementary). In MA-IB, where ToE is early, RR is the highest above 30. However in Scandinavia, RR is lower but ToE happened as early.*

- Table 2. Please elaborate on this description. Include units, describe each of the four contributions, etc.

Response:

We added the unit % in the table. The new caption is:

> *Table 2: Spatial average of each component (the heat index Tmax, the drought index S, and the dependence C) contribution during upper-PoE and lower-PoE (respectively Contrib_Tmax, Contrib_S, Contrib_C). Contrib_int is the residual term.*

- Figure 1. It would be helpful to visualize the signal prior to smoothing. Is that possible to include in a figure or even in the supplement?

Response:

The graph below compares the original unsmoothed signal (thin line) with the smoothed signal (thick curve) for the Vilnius point, which is used as an illustrative example in the methodology. In this case, computing periods of emergence on the original signal would result in three PoE-up, including one lasting only one year.

We have included the following graph in Supplementary Figure 1 and added the following sentence in the revised manuscript on lines 202-203:

> *This smoothing is necessary, as illustrated in the supplementary material (Fig. S1), where the unsmoothed signal leads to the detection of a 1-year PoE.*

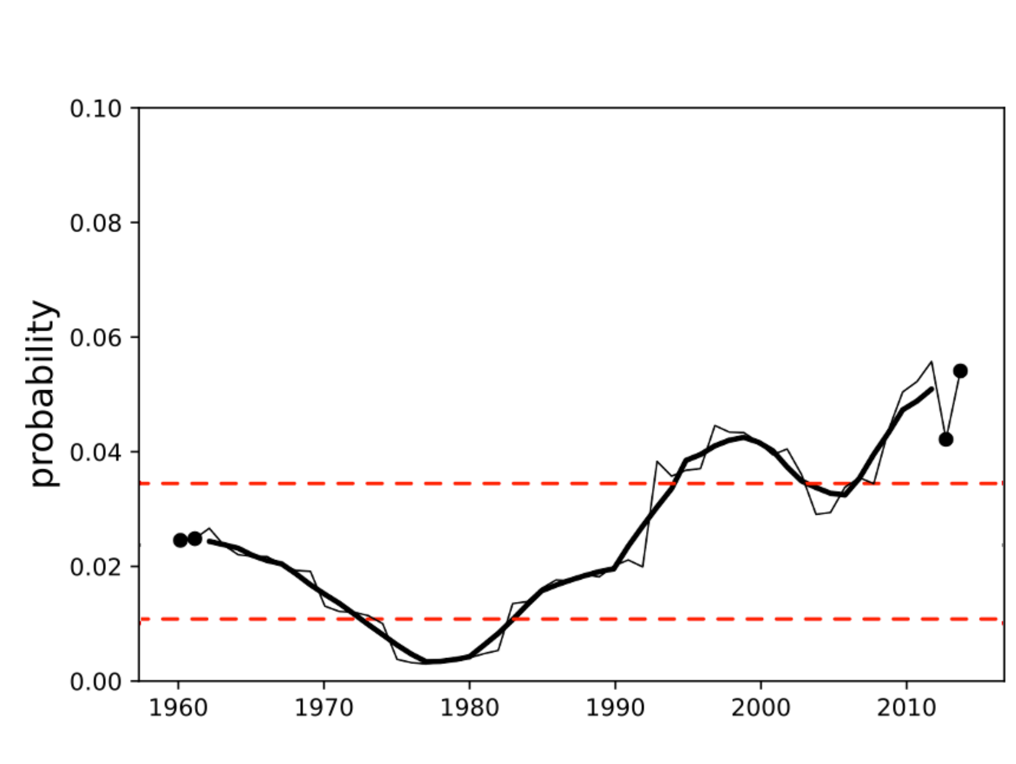

Figure S1 : Compound event probability signal in Vilnius. The thick curve represents the smoothed signal (5-year window), and the thin line the original (unsmoothed) signal. The natural variability is given by the two dotted red lines.

- Figure 2. Explain what the difference between the red and black arrows is. Is it appropriate to compare regions of different spatial scales, covering different proportions of land and ocean area?

Response:

The purpose of defining the five regions is to group areas that exhibit similar behaviors in terms of emergence, component contributions, and the influence of dependence. This makes the interpretation of results more straightforward. Since the regions are used solely as an analytical aid, we did not account for their spatial extent when computing any metrics. Indeed, if we were analyzing spatial aggregates, the size of each region would be relevant. The reviewer can find an extract from the original paper below on lines 296-298:

> *“The goal of this part is to shed light on spatial patterns of emergence over Europe and north Africa. To showcase a variety of mechanisms, five areas will be studied in greater depth. The latter were selected because signals tend to exhibit similar behaviors within each one.”*

To illustrate the typical behavior within each of the five regions, we selected one representative point per region, indicated by red crosses on the figure 2. We then analyzed the temporal evolution at these points. The five graphs in Figure 6 correspond to these five red crosses. The two black crosses (on the figure 2) mark the locations associated with the two graphs presented in Figure S9. This explanation can be found on the original title of the figure 2:

> *“The red (black) crosses are the points specifically examined for the temporal analysis (in Supplementary).”*

To clarify, no ocean or sea data were included in the analysis. Grey areas on the maps indicate regions that were not studied. These grey areas may be more clearly visible in Figure S2, which shows the bivariate threshold used in this study.

- Figure 3. Are you able to show with different colors / hatch marks the four cases where white color is? It’s currently difficult to decipher what white means across the region. It would also be helpful to give a short summary of your interpretation of this figure.

<u>Response:</u>

We understand the reviewer’s point. We initially tried to distinguish the different cases individually, but the result became difficult to read. To improve clarity, we chose to use white areas to indicate the absence of the metric (with the specific metric identified in each map title). In maps (a) and (b), white areas indicate the absence of ToE-up and ToE-low, respectively. This means that in those regions, the signal has not shown any recent significant increase or decrease.
In maps (c) and (d), white areas indicate the absence of PoE-up and PoE-low respectively, meaning the signal has never exhibited a significant increase or decrease over the entire study period.
A more detailed caption for the figure 3 is provided below :

> *Figure 3: Maps of emergence features, when the signal varies below (right panels) or above (left panels) the natural variability. Date of (a) upper-ToE and (b) lower-ToE. Total duration of (c) upper-PoE and (d) lower-PoE. Maps (a) and (b) highlight the spatial patterns of time of emergence. In most of Europe, the probability signal has emerged above the range of natural variability by the end of the dataset. In contrast, some regions show a significant recent (Ireland) and old (Eastern Europe) decrease in CE probability. White areas in these maps indicate the absence of ToE-up (a) and ToE-low (b). Maps (c) and (d) identify hotspots of periods of emergence, showing where*

significant increases (PoE-up) or decreases (PoE-low) have occurred over the study period. White areas in these maps indicate the absence of PoE-up (c) and PoE-low (d).

- Figure 4. It would also be helpful to give a short summary of your interpretation of this figure.

<u>Response:</u>

A more detailed caption for the figure 4 is provided below :

Figure 4 : Contribution of each statistical component (Tmax and S marginals, and the dependence C) during upper-PoE. Spatial patterns of (a) ContribTmax,up, (b) ContribS,up, (c) ContribC,up and (d) Contribint,up. The last map corresponds to the residual term or interaction term. The sum of the four panels is equal to 100 (unit in %). White areas indicate regions where no PoE-up occurred. The contribution metric allows to identify which component change is most responsible for the significant increase in compound event probability. In most of Europe, Tmax is the dominant driver. S is the main contributor only in Italy, though it remains positive across almost all regions. The dependence component shows regional contrast—it contributes positively in the west, but negatively in Eastern Europe. Finally, the residual term is the dominant factor in the Iberian Peninsula.

- Figure 9. It would also be helpful to give a short summary of your interpretation of this figure.

<u>Response:</u>

A more detailed caption for the figure 9 is provided below :

Figure 9~~.~~ Time of emergence above the natural variability for the 7-2022 hot and dry compound event. (a) Spatial distribution of upper-ToE. White grid cells indicate no upper-ToE detected within the study period (i.e., no significant and persistent increase in CE probability up to 2014). (b–c) Evolution of compound event probabilities at two specific locations, marked by black crosses in panel (a). These points are selected to illustrate two contrasting cases with similar upper-ToE: one with zero probability during the reference period (b), and another with greater natural variability (c). The probability associated with changes of dependence, temperature index Tmax, and drought index S are coloured respectively in green, orange and pink. CE probability when the dependence is constant is shown in blue. The year indicated on the x-axis is the middle of the 20y window. All signals are smoothed using a 5-year window; thus the two first and two last years cannot be used for smoothing, and are plotted with points.

In the same way, we have detailed the title of the other figures :

Figure 5: Contribution of each statistical component (Tmax and S marginals, and the dependence C) during lower-PoE. Spatial patterns of (a) ContribTmax,low, (b) ContribS,low, (c) ContribC,low and (d) Contribint,low. The last map corresponds to the residual term or interaction term. The sum of the four panels is equal to 100 (unit in %). White areas indicate regions where no PoE-low occured. The contribution metric

allows to identify which component change is most responsible for the significant decrease in compound event probability. In most of Europe, S is the dominant statistical driver. Tmax is the main contributor in Italy, Hongry, Ukraine, Greece and Turky. Contribution of the dependence change is mainly positively for lower-PoE, contrasting with the residual term which is highly negative.

Figure 6 : Evolution of hot-dry CE probabilities : when only one statistical component evolves, when the dependence is kept constant and the final signal p, at 5 points located in the 5 areas under study. Locations are given in Fig.02 by red crosses. The probability associated with changes of dependence, temperature index Tmax, and drought index S are coloured respectively in green, orange and pink. CE probability when the dependence is constant p_(X,Y), and the probability signal p are shown in blue and black. The year indicated on the x-axis is the middle of the 20y window. All signals are smoothed using a 5-year window; thus the two first and two last years cannot be used for smoothing, and are plotted with points. These graphs illustrate both the timing and magnitude by which the probabilities from different experiments emerge from natural variability. The vertical lines represent the time of emergence for the signals p (in black) and p_(X,Y) (in blue). In some cases, the interval between these two ToE is large (e.g., panels d and e), while in others it is nonexistent (e.g., panel c).

Figure 8 : Probability during the reference period of the 7-2022 hot and dry event and its associated risk ratio over Europe (a) Probability of the event during the 1950-1969 period and (b) risk ratio (ratio of CE probabilities between the first period 1950-1969 and the last period 1994-2023). White grid cell refers to a null probability during the baseline period.