esd-2020-37: "A new view of heat wave dynamics and predictability over the Eastern Mediterranean" by Assaf Hochman, Sebastian Scher, Julian Quinting, Joaquim Pinto and Gabriele Messori.

Point by point response to Reviewer 1:

Reviewer 1: In this paper, the authors employ an approach from dynamical systems theory to quantify the (intrinsic) predictability of atmospheric states based on reanalysis data during cold and hot extremes over the Eastern Mediterranean. This is complemented with GEFS reforecasts, which are used to infer forecast uncertainty, or practical predictability. While the distinction and investigation of practical and intrinsic predictability is not new (e.g. Melhauser & Zhang, 2012), I am not aware of any comparable publications in the context of heatwaves. In addition, a simple Lagrangian model is used to reveal the origin of near-surface air masses during hot and cold extreme events. The resulting paper is nicely structured, not too lengthy and certainly an interesting read. I only have two minor comments and a few additional comments, questions and suggestions, as the manuscript is well written and understandable.

Response: Thank you for the positive feedback. We plan on addressing all of the Reviewer's comments in the revised version of the manuscript as described below.

Reviewer 1: In Melhauser & Zhang (2012), the classic Lorenz (1969) paper is cited multiple times; whereas intrinsic predictability is first defined as "the extent to which prediction is possible if an optimum procedure is used", it is then also related to knowledge of the 'atmospheric state'.

Now, here, the authors do not cite any study when they claim that "As opposed to the practical predictability, the intrinsic predictability only depends on the characteristics of the atmosphere itself." Even though I understand that authors might want to stick to historical definitions, to me, it makes only little sense to limit the forecasting system to the atmosphere. An increasing amount of evidence shows that the Earth's surface does not only supply the atmosphere with heat and moisture, but, to some extent, also exerts control over it (e.g. Koster *et al.*, 2010; Dirmeyer *et al.*, 2018). Knowledge of the land (and ocean) surface state thus implies improved predictability up to sub-seasonal timescales (thanks to, e.g., soil memory).

To be clear, I believe the focus on the atmosphere in this study makes sense, but I would still like to question why this 'intrinsic' predictability should be purely atmospheric by any means. To me, it seems like Lorenz (1969) emphasized the knowledge of all governing equations as well as observing the initial state, and I do not see why this would not include other components of the Earth System.

Response: We agree with the Reviewer's comment. Indeed, the atmospheric state depends not only on the atmosphere, but is influenced by interactions with land and ocean. However, it is important to note that in many – albeit certainly not all – cases these interactions influence the atmosphere at time scales longer than those we consider in our analysis (e.g., Entin *et al*., 2000), and act as a seasonal-scale preconditioning to extremely high summer temperatures (e.g. the mechanism discussed in the Zampieri *et al*. (2009) study the Reviewer cites in his second comment). We plan on rephrasing the sentences related to this notion. We will specifically clarify that, while we take a predominantly atmospheric perspective, centered on synoptic timescales, there is ample evidence of the importance of surface interactions with other "spheres" of the climate system, not least for controlling atmospheric predictability (such as in the experiments described in the Koster *et al*. (2010) study the Reviewer pointed us to). In this respect, we will read and reference the relevant literature suggested by the Reviewer.

Reference:

Entin JK, Robock A, Vinnikov KY, Hollinger SE, Liu S, Namkhai A. 2000. Temporal and spatial scales of observed soil moisture variations in the extra tropics. *Journal of Geophysical Research* **105(D9)**: 11865– 11877. https://doi.org/10.1029/2000JD900051

Reviewer 1: Concerning the 2010 heatwave analysis: while it is interesting to show that the air parcels tracked back in time were warmer than on average even 10 days in the past in this specific case, Bieli *et al.* (2015), e.g., already found that high temperatures in the Balkans area tend to be primarily the result of high starting temperatures combined with extensive descent, enabling strong adiabatic heating. The authors also attempt to explain why particularly the metrics calculated on SLP differ so strongly from those calculated on Z500. I am using this opportunity to refer to my previous comment here – perhaps the fact that the evolution of the atmospheric state closer to the surface tends to follow less of a clear pattern, or "the larger spread in dynamical systems properties across the different heat waves for SLP than for Z500", is partially caused by interactions with the land surface. Naturally, these interactions predominantly affect the lowermost parts of the troposphere, and to provide an example, it can actually be seen (if the different units are accounted for) in Fig. 12 of Zampieri *et al.* (2009) that unusually dry soils affect SLP more than Z500 in a modelling experiment. I thus think that the 2010 heatwave part needs a bit more attention, as currently, the main message is that the single case is similar to the climatology with respect to Z500 evolution during heatwaves, but highly different in terms of SLP, and in my opinion, this is not explained sufficiently (see also comments above and concerning L. 296 below).

Response: Thank you for this comment and for pointing us to some useful references we had overlooked. We agree that the discussion of the 2010 heat wave could be extended with respect to the differences between the dynamical systems analysis close to the surface and at 500 hPa. We specifically plan to discuss the possible influence air-sea interactions may have on the dynamical systems metrics at surface level. Soil moisture, on the other hand, may not be an important factor for controlling heat waves over the South-Eastern part of the Mediterranean, but rather more important in the North-Eastern areas (Zittis *et al*., 2014). It may however be possible that low soil moisture in the regions where the air parcels originate influence the intensity of Eastern Mediterranean heat waves. We will add a discussion of this point. We will further read the references suggested by the Reviewer and refer to them in the revised version of the manuscript.

Reference:

Zittis G, Hadjinicolaou P, Lelieveld J. 2014. Role of soil moisture in the amplification of climate warming in the eastern Mediterranean and the Middle East. *Climate Research* 59: 27-37. https://doi.org/10.3354/cr01205

Reviewer 1: L. 36: Feeling the heat**,** 2018 (commamissing?)

Response: A comma will be added.

Reviewer 1: L. 49: Saaroni and Ziv,2000 (spacemissing?)

Response: A space will be added.

Reviewer 1: L. 126: If the CSI comprises the boundary layer height, then why is it absent from the equation below (L. 130)? Also, in the cited Saaroni *et al.* (2017), the atmospheric boundary layer height itself is barely mentioned, but rather the (height of the) persistent marine inversion. I thus recommend slightly editing (or shortening) this part to further enhance the consistency and clarity of the text.

Response: We agree with the Reviewer and will reformulate this part of the manuscript according to Saaroni et al. (2017).

Reviewer 1: L. 212: Is there any reason for this choice, i.e. initializing trajectories between the surface and 90 hPa above, other than simplicity? To me, it seems more intuitive to always track the air masses within the atmospheric boundary layer back in time, whose height may vary from day to day, and tends to be (positively) anomalous particularly during heatwaves (this might not be the case in the study area of interest, but was certainly true for the 'epicenter' of the 2010 Russian heatwave; see, e.g. Miralles *et al.*, 2014). However, considering that Dayan *et al.* (2002) demonstrated onlylittle synoptic-scale influence on summer ABL heights compared to the distance to the coast, the usage of a constant layer to be tracked backward might be entirely justified, but perhaps the authors can still elaborate on their choice.

Response: Thank you for this comment. According to recent studies, the planetary boundary layer height in Israel during summer is ~600 – 900m above the surface (Uzan *et al.*, 2016; 2020). Assuming that the pressure decreases by 1 hPa every 8m height difference, 90 hPa corresponds to about 720m, thus 90hPa can be considered a reasonable choice.

We have further evaluated the trajectories that start from 90hPa above the surface and higher (Figure R1). We find that at these levels the picture is somewhat different. Air parcels associated with heat waves are located over the Mediterranean Sea and east of Israel prior to the heat wave (Figure R1a). Air parcels associated with cool days originate mostly over the Atlantic and Europe and are transported downstream (Figure R1b). This further indicates that our trajectory analysis successfully identifies air parcels in the boundary layer, which have a different path than those in the free troposphere. Moreover, the thermodynamic properties of the air parcels do not show any noticeable differences between heat waves and cool days (Figure R1c-f). We will elaborate on the choices we made with regards to the back trajectories' analysis as described above in the revised version of the manuscript.

Reference

Uzan L, Egert S, Alpert P. 2016. Ceilometer evaluation of the eastern Mediterranean summer boundary layer height – first study of two Israeli sites. *Atmospheric Measurement Techniques* **9**: 4387–4398. https://doi.org/10.5194/amt-9-4387-2016

Uzan L, Egert S, Khain P, Levi Y, Vadislavsky E, Alpert P. 2020. Ceilometers as planetary boundary layer height detectors and a corrective tool for COSMO and IFS models. *Atmospheric Chemistry and Physics* 20: 12177-12192. https://doi.org/10.5194/acp-20-12177-2020

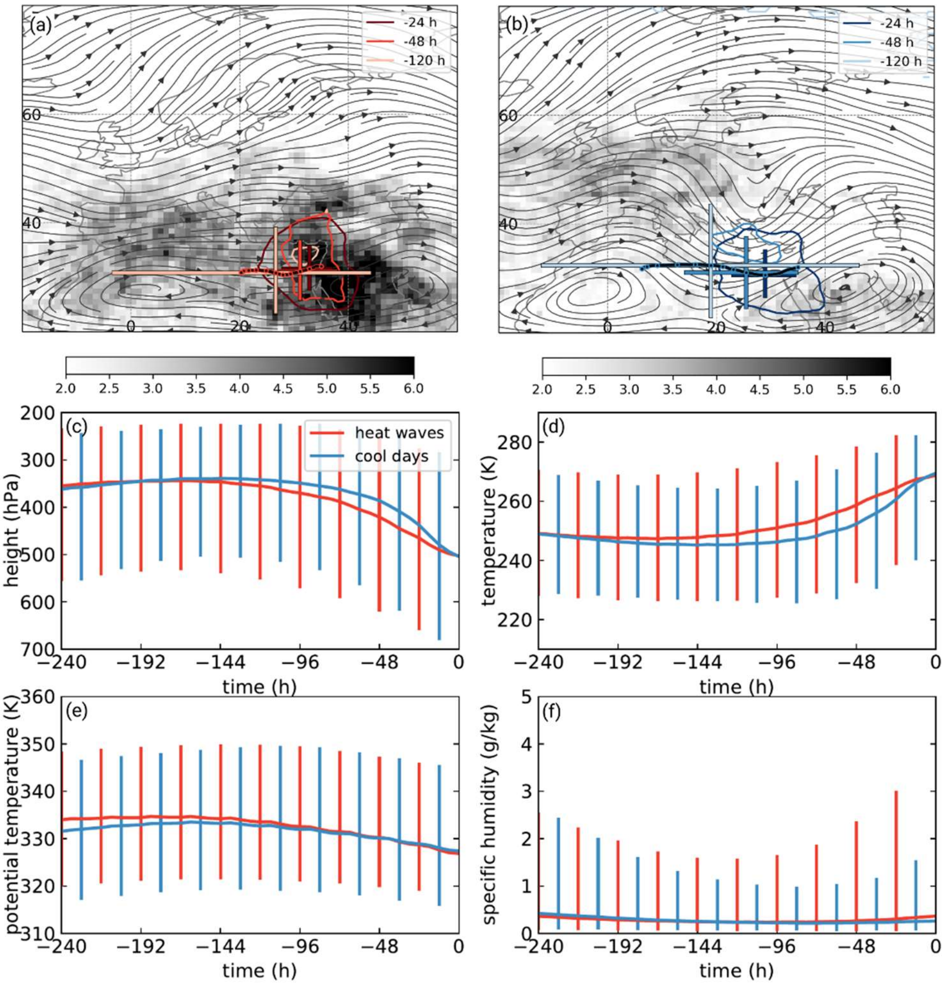


Figure R1 Same as Figure 2, but for trajectories initialized from 90hPa above the surface and higher with the uncertainty around the median trajectories of 25 – 75% of the trajectories.

Reviewer 1: L. 240: It is quite interesting that the specific humidity does not increase nearly as much in the last 48 hours prior to arrival during heatwaves as for cold extremes, but is this only a consequence of different 'inflow', i.e. more trajectories over the Mediterranean Sea? Of course, this cannot be gauged solely by a visual comparison of Figs. 2a & 2b (in which, to me, the trajectory densities shortly before arrival do not seem to differ much), but I would also suspect that additional factors are at play – such as, e.g., enhanced convective activity (and hence moistening of the troposphere).

Response: Thank you very much for this suggestion. Our additional analysis reveals that for most of the time, the portion of terrestrial back-trajectories is similar for heat waves and cool days (Figure R2). It is only at about 72 to 24 hours prior to the events that the portion of terrestrial (marine) back-trajectories is lower (higher) for cool days than for heat waves (Figure R2). This is in line with the evolution in specific humidity along the trajectories, which increases more strongly for the cool days than for the heat waves (Fig. 2f). The increase of moisture is most likely related to the passing of the air masses over the Mediterranean Sea. As a caveat, this analysis does not account for local processes such as the convective activity mentioned by the Reviewer. We will revise the text accordingly and consider adding Figure R2 as a Supplementary Figure.

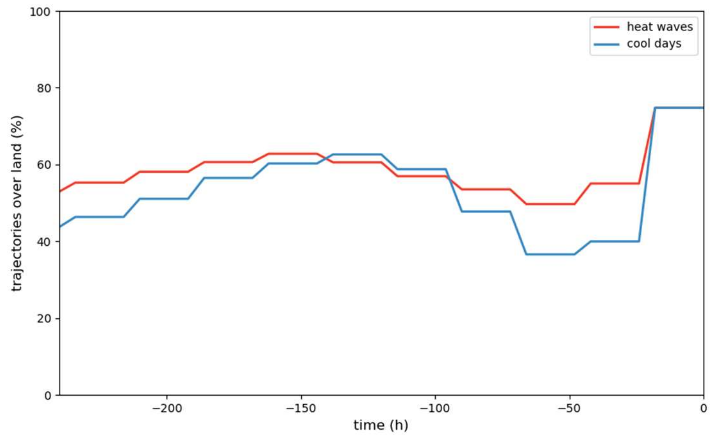


Figure R2 The portion of trajectories over land for heat waves (red line) and cool days (blue line).

Reviewer 1: L. 264: "The build-up towards this type of event is characterized by an increase in θ (decrease in persistence) and a decrease in d (Fig. 4d)." This comment also concerns the Methods section; I think the authors provide a good overview of the two dynamical system metrics, but perhaps it would be helpful to explicitly state that, as explained by Moloney *et al*. (2019), more (expanding) dimensions around the state of interest (or degrees of freedom, I suppose) imply less predictability. Or, in other words, lower d suggests higher predictability – this is actually stated as such on L. 335 in the Summary, but as far as am I concerned, not before. Perhaps it would seem a bit confusing to edit the sentence I am quoting above (L. 264), as in this example (Fig. 4d), persistence decreases, yet predictability as gauged by the local dimension d increases. Still, I believe I am not the only reader who would appreciate a bit more guidance throughout the manuscript.

Response: We will clarify these points and provide a more intuitive qualitative interpretation of the metrics earlier in the manuscript, focusing on their relation with the intrinsic predictability. To that end, we have completely re-structured Sect. 2.3 to provide both a clearer intuitive explanation of the metrics and a more detailed description of the mathematical background to ensure the reproducibility of our results.

Reviewer 1: L. 280: "The pattern somewhat resembles the temporal evolution of d computed on SLP (cf. Fig. 5e and Fig. 4c), but stands in stark contrast to the pattern computed on Z500". While I agree that there is a stark contrast to the pattern shown in Fig. 4a (Z500), I find this resemblance a bit difficult to spot, as the peaks in d and msl spread appear to be shifted by about one day. Is there any obvious reason for this? Also, I imagine this would look different for shorter or longer lead times, so come to think of it, why not 24 hours less or more? This is not a request to repeat the entire analysis for different lead times (probably out of scope anyways), but I am just curious if the authors looked into this and if so, how much this choice even matters in the first place.

Response: Thank you for this comment. Since the spread and error are computed every 24 hours and the dynamical systems metrics are instantaneous in time (local in phase-space) and computed from the 6-hourly data, we believe that a shift of up to 24-hours may be reasonable. We have further tested a 24-hour shorter lead time and found that the peak in the spread is at about 0 h, resembling the peak in *d* computed on SLP (Figure R3, left panel). We will discuss this point briefly in the revised text, as we have indeed not described in detail the sensitivity of our results to the choice of lead-time.

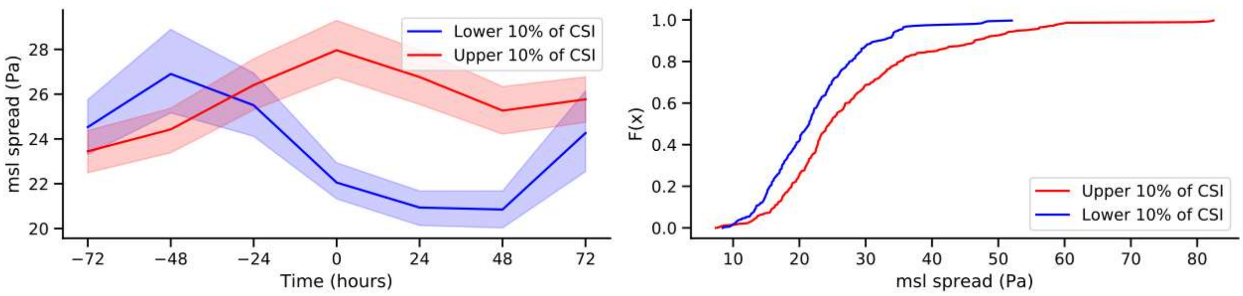


Figure R3 Same as Figure 5e, f but for a 24-hour shorter lead time.

Reviewer 1: L. 296: "We further hypothesize that differences between the single case and the climatology may be related to the relatively small day to day variations during summer over the Eastern Mediterranean (Ziv *et al.*, 2004), which make it challenging to depict the exact onset of a heat wave". Could you please elaborate how the challenges related to defining the onset of a heatwave could contribute to the extreme differences between the 2010 case and the climatology, but only for SLP and not Z500 – this is not obvious to me.

Response: When comparing the climatology of the temporal evolution of *d* and θ for Z500

(Fig. 4a) with the single case (Fig. 6b) it is relatively easy to see that in both there is an increase in both d and θ as the heat wave develops. On the other hand, when comparing the temporal evolution of d and θ for SLP (Fig. 4c with Fig. 6c), one can see that depicting the exact time the heat wave starts is very important for comparison. Still, in both Figures d increases and θ decreases at some point in the chosen time window, but the timing of these trends is shifted between the climatology (Fig. 4c) and the single case (Fig. 6c). We will clarify this in the revised version of the manuscript.

Reviewer 1: L. 303: Concerning the anticyclonic wave breaking, if my understanding is correct, then this can be seen in Fig. 8, as the trough east of the ridge centered over European Russia, clearly visible from Fig. 8b onward, is tilted (southwest-northeast; Davini *et al*., 2012) and advected westward (consistent with the definition given in Quandt *et al*., 2019). I suggest adding a brief description along these lines for readers unfamiliar with the terminology, this would also prevent readers from overlooking Fig. 8 (which, currently, is only mentioned but not discussed in the main text).

Response: Thank you for this suggestion. The Reviewer is correct with regard to the anticyclonic Rossby wave breaking shown in Fig. 8. We will add some more information on this for the reader and discuss Fig. 8 more in detail in the revised version of the manuscript.

Reviewer 1: L. 305: (Quandt *et al*., 2019)**,** played (comma missing?)

Response: A comma will be added.

Reviewer 1: L. 322 (+ 365, 511): K**ue**ne ⇒ Keune

Response: The typo will be corrected in all mentioned lines.

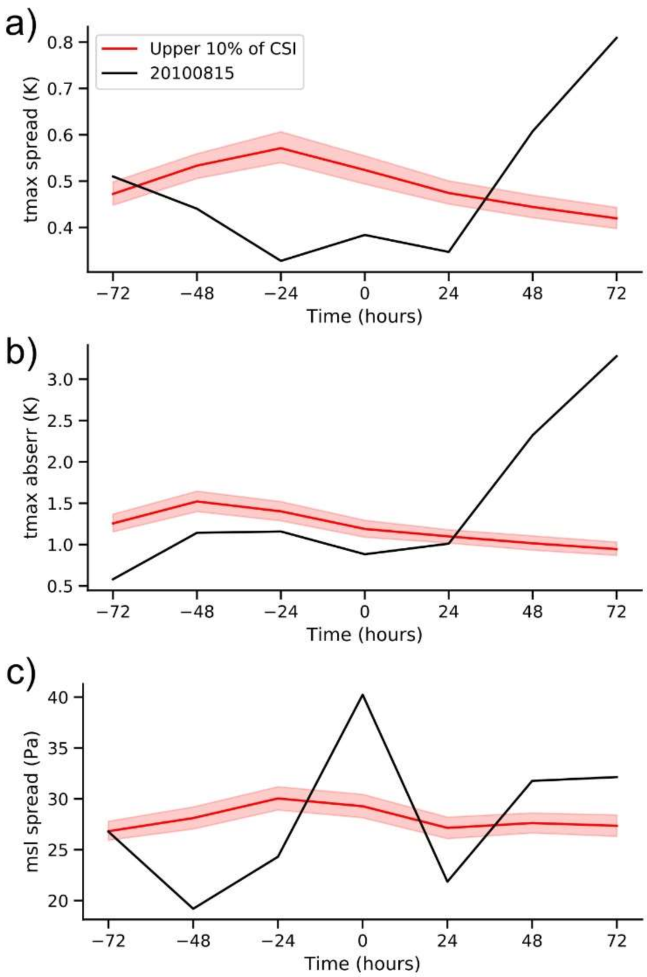
Reviewer 1: Fig. 5: While a few sentences in the Methods explain what is really shown in Fig. 5, I believe the caption might benefit from a small addition, hinting at the fact that results are plotted for their corresponding initialization times.

Response: We will expand the caption of Fig. 5 to explain that the results are plotted for initialization time.

Reviewer 1: Fig. 9: I suggest using red colors for the upper 10% of CSI, as in previous figures, and plotting the 2010 heatwave results in black (or any other color than blue) instead.

Response: We will change the Figure according to the Reviewer's suggestion (see below new

Figure 9).



New Figure 9 Forecast spread/skill for the mid-August heat wave centered (0 h) on 15.8.2010 at 12UTC (black line). The mean temporal evolution of the ensemble model spread for Tmax (a), SLP (c) and absolute error for Tmax (b) of forecasts with lead-time 69h, initialized at different time lags with respect to the event, computed every 24 hours. The heat waves (upper 10% of CSI - red line) are displayed for reference. A 95% bootstrap confidence interval for all heatwaves is displayed in shading.

Reviewer 1: References

Bieli, M., Pfahl, S. & Wernli, H. A lagrangian investigation of hot and cold temperature extremes in Europe. Q. J. R. Meteorol. Soc. 141, 98–108 (2015).

Davini, P., Cagnazzo, C., Gualdi, S. & Navarra, A. Bidimensional diagnostics, variability, and trends of northern hemisphere blocking. J. Clim. 25, 6496–6509 (2012).

Dayan, U., Lifshitz-Goldreich, B. & Pick, K. Spatial and structural variation of the atmospheric boundary layer during summer in Israel-profiler and rawinsonde measurements.

J. Appl. Meteorol. 41, 447–457 (2002).

Dirmeyer, P. A., Halder, S. & Bombardi, R. On the Harvest of Predictability From Land States in a Global Forecast Model. J. Geophys. Res. Atmos. 123, 13,111-13,127 (2018).

Koster, R. D. *et al.* Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. Geophys. Res. Lett. 37, 1–6 (2010).

Levi, Y., Shilo, E. & Setter, I. Climatology of a summer coastal boundary layer with 1290-MHz wind profiler radar and a WRF simulation. J. Appl. Meteorol. Climatol. 50, 1815–1826 (2011).

Lorenz, E. N. Atmospheric Predictability as Revealed by Naturally Occurring Analogues. J. Atmos. Sci. 26, 636–646 (1969).

Miralles, D. G., Teuling, A. J., Van Heerwaarden, C. C. & De Arellano, J. V. G. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. Nat. Geosci. 7, 345–349 (2014).

Melhauser, C. & Zhang, F. Practical and intrinsic predictability of severe and convective weather at the mesoscales. J. Atmos. Sci. 69, 3350–3371 (2012).

Moloney, N. R., Faranda, D. & Sato, Y. An overview of the extremal index. Chaos 29, (2019).

Quandt, L. A., Keller, J. H., Martius, O., Pinto, J. G. & Jones, S. C. Ensemble sensitivity analysis of the blocking system over Russia in summer 2010. Mon. Weather Rev. 147, 657–675 (2019).

Saaroni, H., Savir, A. & Ziv, B. Synoptic classification of the summer season for the Levant using an ‘environment to climate’ approach. Int. J. Climatol. 37, 4684–4699 (2017).

Zampieri, M. *et al.* Hot European Summers and the Role of Soil Moisture in the Propagation of Mediterranean Drought. J. Clim. 22, 4747–4758 (2009)

Response: Thank you for providing these references. We plan on citing them where applicable and referring to them in the revised text.