Review of Hochman *et al.*, 2020

A New View of Heat Wave Dynamics and Predictability over the Eastern Mediterranean

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.

- 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  Now, here, the authors do not cite any study when they claim that “As opposed to the practical predictability, the intrinsic predictability only depends on the characteristics of the atmosphere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.
- 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).

Further comments

- L. 36: Feeling the heat**,** 2018 (comma missing?)
- L. 49: Saaroni and Ziv,2000) (space missing?)
- 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.
- 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 only little 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.
- 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).
- 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.

- 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.
- 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.
- 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).
- L. 305: (Quandt *et al*., 2019)**,** played (comma missing?)
- L. 322 (+ 365, 511): K**ue**ne ⇒ Keune, typo.
- 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 initialisation times.
- 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.

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).