Dynamic Time Warping as an Alternative Pairwise Coordination Assessment

Novelty Addition to Entry 6 - Module 8

Group 6: Djourdan Johnson, Jacuqot Qiu, Lotte Michels, Nawat Nawati Azhati, Nuo Xu, Xuelin Wei

1 Measuring Pairwise Coordination

Analyzing coordination involves capturing the ways in which components of a system change together over time. There is a plethora of methods that can be employed to quantify such pairwise interactions (Cliff et al., 2023). One such method is the Hilbert transformation, that was demonstrated in the previous e-portfolio entry. Although this is a classical approach for pairwise coordination analysis, various recent papers point out that it can not account well for noise in a signal (Tao, 2024; Wodeyar et al., 2023). The MSPB dataset (Zhu et al., 2024) that is used in this project, however, may contain quite some noise, as found in the first entry of this e-portfolio.

With this in mind, we will further examine pairwise coordination in the beehive signals using a method that has not been covered in the Complex Systems course: Dynamic Time Warping (DTW). Originally proposed in 1978 by Sakoe and Chiba, DTW remains widely applied and actively researched across domains. Moreover, it has been proposed as a noise- and non-stationarity resistant approach compared to the Hilbert transformation (Tao, 2024). In this entry, we will briefly illustrate the underlying concept of DTW and demonstrate its applicability to the MSPB dataset.

The code used for the computations and visualizations in this entry can be found here.

2 Background

DTW was designed to measure the similarity of two time-series based on their optimal alignment (Sakoe and Chiba, 1978). By first temporally matching two signals, the algorithm accounts for inter-signal differences regarding length, lag and phase. Thus, DTW can find pairwise interactions regardless of temporal incongruities, in contrast to traditional similarity metrics such as a straight-forward euclidian distance computation (Mishra, 2021).

The underlying notion of DTW is best explained using an exemplary visualization. All DTW outcomes and visualizations in this e-portfolio entry are generated using the dtw-python package (Giorgino, 2009). Figure 1 displays the alignment returned by the DTW algorithm for the raw beehive temperature and humidity time-series. The diagonal lines indicate the point-to-point mapping to match the signals in such a way that their cumulative distances are minimized.

DTW computes this minimization by setting up a 'cost matrix' and finding the 'cheapest' warping path through this matrix from the bottom left corner to the upper right corner.

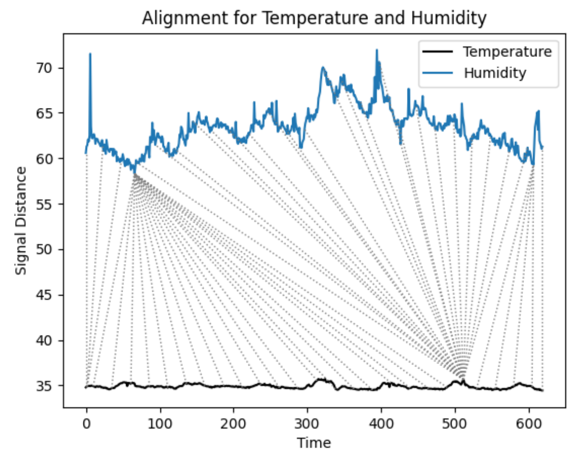

Figure 1: DTW alignment result between the raw temperature and humidity signals. For clarity purposes, the signals are displayed for one week of data (August 8 - August 14).

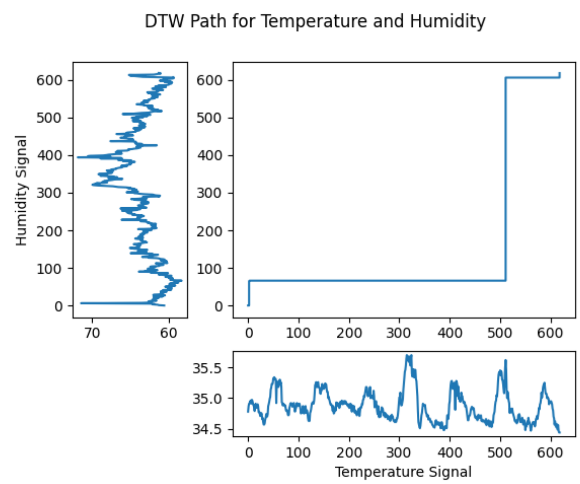

Figure 2: Warping path between the temperature (horizontal) and humidity (vertical) signals corresponding to Figure 1.

To give an example, the warping path corresponding to the alignment in Figure 1 is included in Figure 2.

Horizontal and vertical segments in a warping plot indicate one signal holding a value while the other progresses. Thus, humidity holds a value while temperature progresses until around the 500th timepoint, after which these roles switch.

3 Pre-Processing Steps and Parameter Settings

Figure 1 highlights an important aspect of the MSPB dataset: the time-series are heterogeneous. Specifically, temperature is measured in degrees Celsius; humidity is measured in relative percentage; beehive activity is measured using Fourier transform power values. These different units exhibit varying means and standard deviations, as can be seen in Figure 1.

Rather than being dominated by such unit characteristics, coordination analyses should reflect only signal shape or temporal structure. As such, two important pre-processing steps need to taken before DTW implementation: 1) similar to the preparations for the Hilbert transform seen previously, signals need to be centered around a zero mean and 2) signals need to be standardized. In our code, the StandardScaler function from the sklearn package is used (Pedregosa et al., 2011). This functionality performs both the centering and scaling steps to the inputted time-series.

Additionally, prior to applying DTW, it is important to consider constraints on the alignment path to ensure meaningful temporal correspondences. Figure 1 shows that some signal measurements are connected over a temporal lag of more than 400 timepoints, which represents over 4 days in the MSPB dataset (see the first e-portfolio entry for a dataset description). In the context of beehive research, it does not make to consider pairwise interactions between temperature, humidity and colony activity over such an extended time period (Zhu et al., 2024).

To restrict the temporal deviation between aligned time points, a windowing function can be applied. Conventionally, we apply a Sakoe-Chiba window (Sakoe and Chiba, 1978). This functionality can be used to define a fixed bandwidth around the diagonal of the cost matrix within which the warping path must lie. This bandwidth if determined by the window size parameter. We set this parameter to 48 timepoints to avoid alignments connected over more than half a day in time, which is a more intuitive timespan in the context of beehives (Zhu et al., 2024).

Figures 3 and 4 show the DTW results for the temperature and humidity signals after these pre-processing and parameter setting steps.

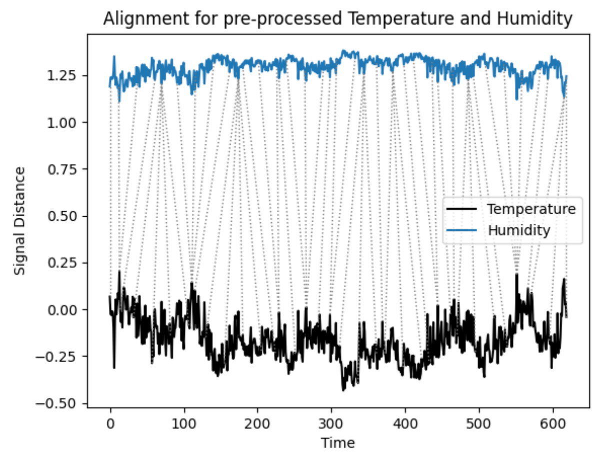

Figure 3: Sakoe-Chiba DTW alignment result between the pre-processed temperature and humidity signals. Matching Figure 1, the signals are displayed for one week of data (August 8 - August 14).

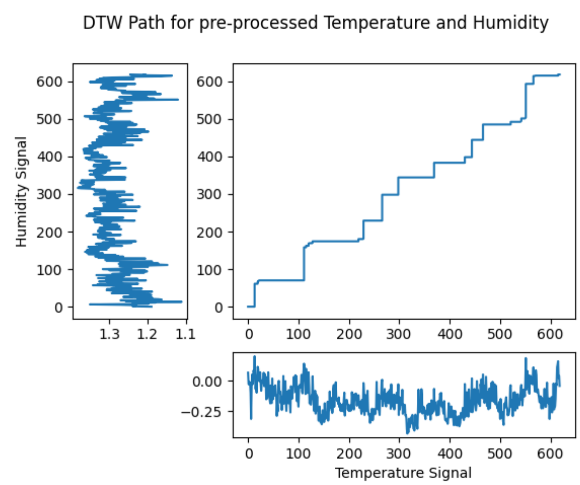

Figure 4: Sakoe-Chiba DTW warping path between the pre-processed temperature (horizontal) and humidity (vertical) signals corresponding to Figure 3.

4 Dynamic Time Warping Results

After determining the pre-processing steps and parameter settings, the DTW algorithm was applied on the complete time-series from the MSPB dataset in a pairwise fashion: temperature against humidity, power against humidity, and temperature against power. The warping results are displayed in Figure 5.

Based on these alignments, the distance between two signals can be computed. The distance feature returned by the dtw-python package is the accumulation of the euclidean distances between the matched signal points (Giorgino, 2009). This can be normalized by dividing by the length of the signals, which results in an average distance between two signals. Normalized distance is 1.25 the temperature-humidity alignment, 2.33 for the power-humidity alignment and 0.96 for the temperature-power alignment. Out of the three pairwise comparisons, the temperature and power signals thus seem to be most correlated.

However, it is hard to interpret a degree of similarity or coordination based merely on the distance metric. Lower distance values indicate better alignment, but there's no inherent threshold for what counts as "similar" or "coordinated" as the meaning of distance values is often context-dependent.

Some solutions have been proposed to get a better sense of signal coordination based on DTW outcomes. For instance, Stent (2024) described that an interpretable similarity score could be computed by dividing the cumulative distances by

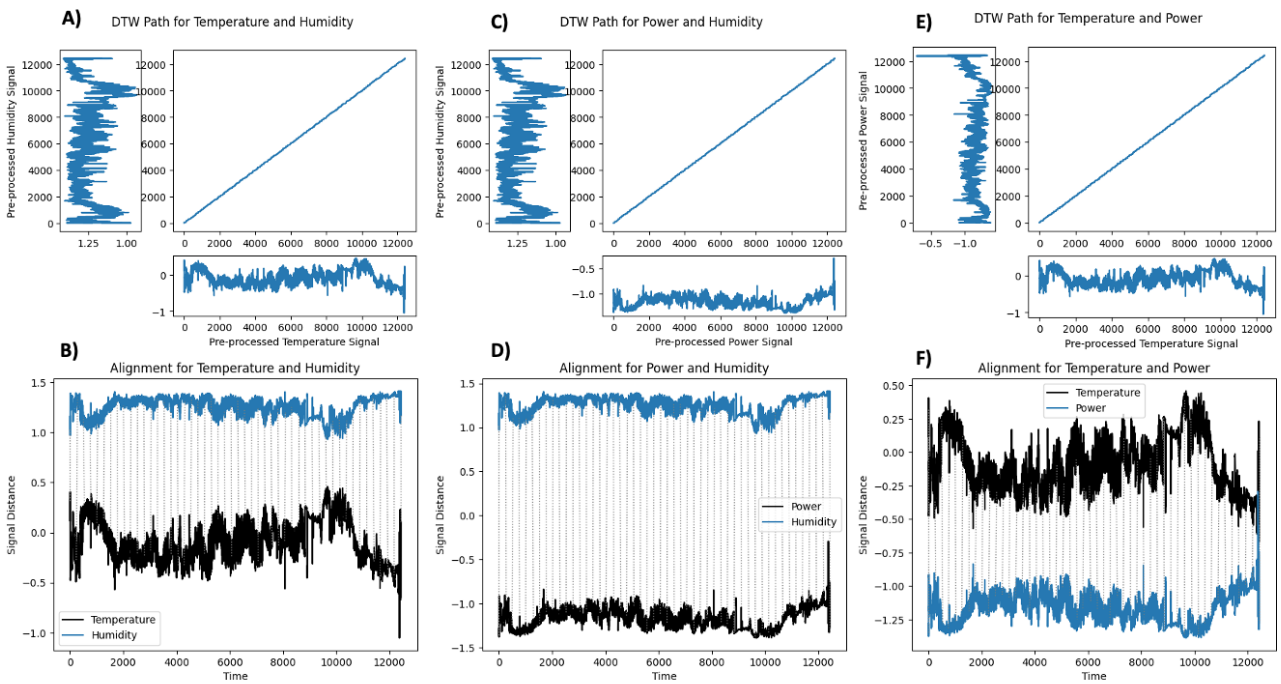

Figure 5: DTW Results for the full temporal signals in the MSPB dataset. A-B: Alignment for temperature and humidity, C-D: Alignment for power and humidity, E-F: Alignment for temperature and power.

the length of the best warping path found by DTW. The resulting metric will lie between 0 (identical signals) and 1 (no similarities were found at all). Applied to the beehive DTW alignments, this approach returns similarity scores of 0.01, 0.02 and 0.01 respectively. These DTW similarity scores thus indicate very high pairwise interactions within the beehive.

Stent (2024) does give a disclaimer, however, that these similarity metrics may be very informative by themselves, but should be considered alongside the graphical representations above (graphs B, D and F within Figure 5). These indeed do indicate some signal interactions, but stating the signals are structurally near to identical may be a far stretch, especially in plots B and F.

The bottom plots in Figure 5 also display another interesting aspect of the beehive time-series, namely that their relative distances and interactions may differ over time. This highlights another pitfall of the global distance or similarity metrics discussed so far: these measures assume stationary coordination over time. However, in many complex systems, pairwise interactions are not stationary. Examining coordination over time may thus be more informative than computing a single, aggregate statistic.

This issue is increasingly being addressed, for instance by Likens and Wiltshire (2020). These authors argue for the use of windowed analysis methods, in their case applied to multiscale synchrony. However, sliding windows can also be applied in conjunction with DTW (Linke et al., 2020; Meszlényi et al., 2017). The next section will therefore provide a more fine-grained pairwise coordination assessment, applying windowed DTW to explore pairwise interactions over time.

5 Windowed Dynamic Time Warping Results

In windowed DTW, a temporal window is slid over the time-series in a step-wise fashion. The DTW algorithm is applied at each step on the signal parts that fall within the window, returning a warping path and (normalized) distance metric for only that specific segment of the time-series. Thus, a sequence of distance values is obtained that can illustrate how pairwise interactions between two signals change over time.

This approach has been succesfully applied in previous research, for instance by Linke et al. (2020). Their experiment involved windowed DTW to recognize functional connectivity anomalies in autistic subjects, outperforming other traditional pairwise interaction methods (such as Pearson correlation). Specifically, the authors used the windowed DTW distance values to retrieve signal similarity scores over time, following a computational approach proposed by Meszlényi et al. (2017). This approach entails that the sequence of distance values are multiplied by -1 and subsequently centered around zero. Next, it is proposed that resulting scores lower than zero indicate 'below average similarity' and resulting scores higher than zero indicate 'above average similarity'.

Following these guidelines, we applied windowed DTW to the time-series in the MSPB dataset. We created a temporal window of 96 timepoints, covering one day in the data; We made the informed assumption that the pairwise interactions between the beehive signals remained stationary over the time-course of one day (Zhu et al., 2024). Next, the DTW window was slid it over the temperature, humidity and power signals with a step size of one timepoint. At each step, DTW was implemented as delineated in the previous entry sections. Finally, following the guidelines by Meszlényi et al. (2017), the

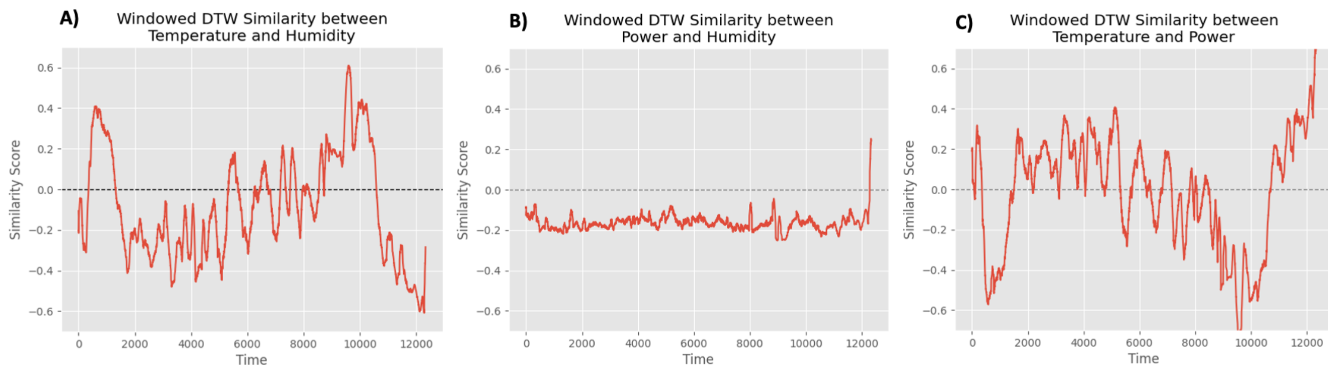

Figure 6: Signal similarity scores over time, obtained by Windowed DTW. Values above zero (dotted baseline) reflect 'above average' similarity. A: pairwise interactions over time between temperature and humidity, B: pairwise interactions over time between power and humidity, C: pairwise interactions over time between temperature and power.

obtained sequence of normalized distance scores were negated and demeaned. The resulting similarity scores over time are presented in Figure 6.

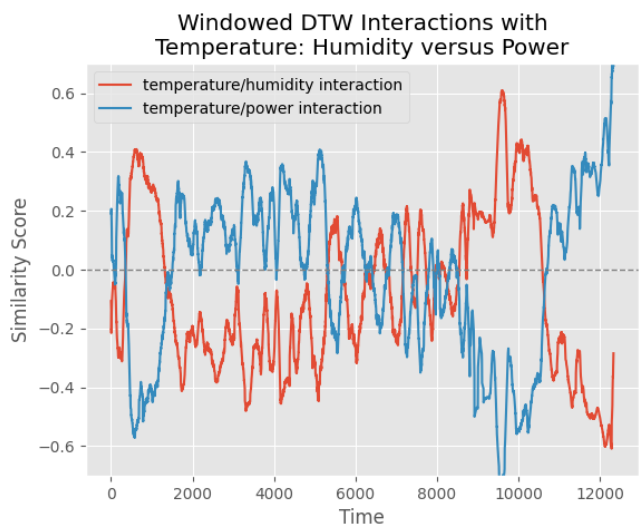

Figure 7: A combined plot of the windowed DTW similarity signals for both the temperature-humidity and temperature-power interactions.

The graphs in Figure 6 suggest that pairwise interactions change over time. Whereas the plots suggest no notable similarities between power and humidity, pairwise interactions seem to fluctuate for the temperature signal a lot more. The temperature-humidity and temperature-power pairings sometimes reach similarity scores of around 0.4, which indicates note-worthy coordination between these signals.

Interestingly, the windowed similarity scores for temperature-humidity and temperature-power seem to alternate over time. This becomes more clear in Figure 7. The fact that these similarity curves alternate over time suggests dynamic shifts in coordination between temperature and the two other variables.

These dynamic shifts may be interpreted within the context of seasonal activities within honeybee colonies, as described by the Mid-Atlantic Apiculture Research and Extension Consortium (2010). Transitioning into the summer or winter period, beehives show emergent thermo-moisture regulatorions which keeps the temperature within the beehive more or less around 33 degrees Celsius, which is best for brooding. These transition periods are visible in the large red peaks in Figure 7. For instance, the peak around timepoint 10.000 represents the months of September and October, during which honeybees will form a tight cluster to maintain core temperature in their hive (Research and Consortium, 2010). Colony activity may then be low, while humidity and temperature inside the hive are more tightly coupled due to condensation from respiration and minimal hive fanning behavior.

In contrast, during the summer period (in between timepoints 2000 to 6000), honeybees show active foraging behavior including the gathering of water for the regulation of beehive temperature (Research and Consortium, 2010). This active thermoregulation is thus reflected in an increase in pairwise interaction between colony temperature and power, visible by the blue peaks in Figure 7. This phenomenon was also discovered when investigating the cross-recurrence between temperature and hive power in entries 4 and 5 (on modules 6 and 7).

Lastly, before committing to these findings, surrogate tests were conducted to assess the robustness of the DTW similarity results and preclude any spurious outcomes. The applied surrogate testing methods are described in more detail in Entry 5/Module 7 of this ePortfolio. The surrogate test results are displayed in Figure 8. First, data shuffling was applied and repeated 20 times. The windowed DTW results on the surrogate signals show no significant pairwise interactions. In addition, segment shuffling was applied. Segments of 24 hours were created to maintain the daily cycles typically displayed by honeybeehive time-series as seen in Entry 1/Module 2 and described by Zhu et al. (2024). Again, segment shuffling was repeated 20 times and no significant pairwise interactions came up. Thus, both surrogate tests indicate that the findings from windowed DTW in Figure 6 and 7 are robust.

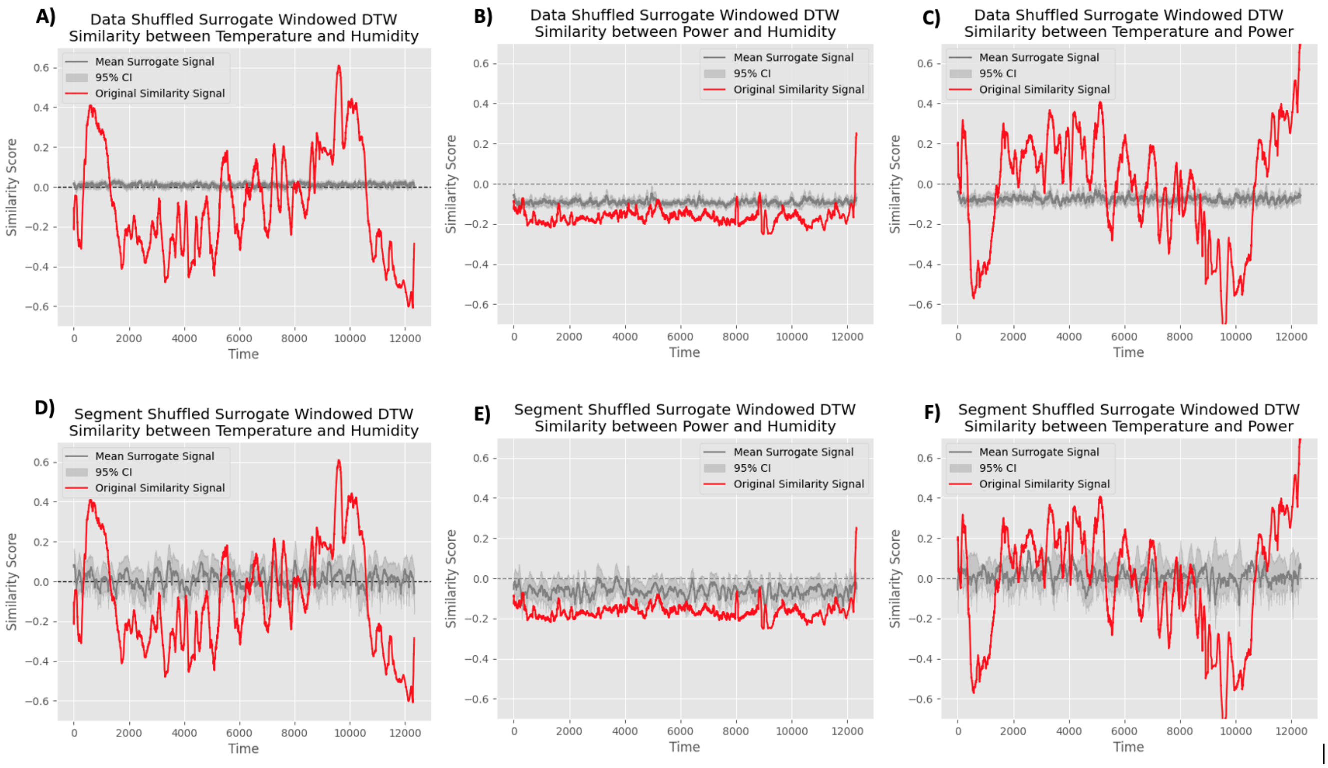

Figure 8: Surrogate test results on the windowed DTW findings. The gray markers indicate the mean and 95% confidence interval after 20 rehearsals. The original temperature-humidity and temperature-power interactions are significantly different from the surrogate distributions (graphs A-D and C-F).

6 Conclusion

To summarize, Dynamic Time Warping (DTW) can be applied as a noise- and non-stationarity resistant approach towards pairwise coordination analysis. DTW can be used to align signals of different units, given that these signals are centered and scaled. Moreover, for DTW to make meaningful alignments between two sequences, it is important set a contextually relevant point-to-point matching bandwidth. DTW can then be used to obtain an aggregate signal distance and signal similarity score. However, these metrics are very much context-dependent, making them hard to interpret, and do not take temporal coordination variations into account. We demonstrated that windowed DTW may be a more suitable approach towards examining the pairwise coordination over time in complex systems such as a beehive. For the beehive that is investigated throughout our ePortofolio, notable pairwise coordinations for the temperature-humidity and temperature-power alignments were found. Interestingly, these interactions alternated each other, indicating dynamic coordination shifts in the beehive over time. These dynamic shifts may reflect typical seasonal behaviors shown by honeybees over the year. Surrogate tests confirmed the robustness of these findings.

References

[Cliff et al.2023] Oliver M. Cliff, Annie G. Bryant, Joseph T. Lizier, Naotsugu Tsuchiya, and Ben D. Fulcher. 2023. Unifying pairwise interactions in complex dynamics. *Nature Computational Science*, 3(10):883–893, 9.

[Giorgino2009] Toni Giorgino. 2009. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *DOAJ (DOAJ: Directory of Open Access Journals)*, 8.

[Likens and Wiltshire2020] Aaron D Likens and Travis J Wiltshire. 2020. Windowed multiscale synchrony: modeling time-varying and scale-localized interpersonal coordination dynamics. *Social Cognitive and Affective Neuroscience*, 16(1-2):232–245, 9.

[Linke et al.2020] A.C. Linke, L.E. Mash, C.H. Fong, M.K. Kinnear, J.S. Kohli, M. Wilkinson, R. Tung, R.J. Jao Keehn, R.A. Carper, I. Fishman, and R.-A. Müller. 2020. Dynamic time warping outperforms Pearson correlation in detecting atypical functional connectivity in autism spectrum disorders. *NeuroImage*, 223:117383, 9.

[Meszlényi et al.2017] Regina J. Meszlényi, Petra Hermann, Krisztian Buza, Viktor Gál, and Zoltán Vidnyánszky. 2017. Resting State FMRI Functional Connectivity Analysis using dynamic time warping. *Frontiers in Neuroscience*, 11, 2.

[Mishra2021] Abhishek Mishra. 2021. Time series similarity using dynamic time warping -Explained. 12.

[Pedregosa et al.2011] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay.

2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[Research and Consortium2010] Mid-Atlantic Apiculture Research and Extension Consortium. 2010. Seasonal Cycles of Activities in Colonies, 6.

[Sakoe and Chiba1978] H. Sakoe and S. Chiba. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics Speech and Signal Processing*, 26(1):43–49, 2.

[Stent2024] Mark Stent. 2024. Dynamic Time Warping - Mark Stent - Medium. 11.

[Tao2024] Xiyue Tao. 2024. Dynamic time warping based instantaneous frequency estimation for rotating machineries under non-stationary operating conditions. *Journal of Physics Conference Series*, 2815(1):012044, 8.

[Wodeyar et al.2023] Anirudh Wodeyar, Francois A Marshall, Catherine J Chu, Uri T Eden, and Mark A Kramer. 2023. Different methods to estimate the phase of neural rhythms agree, but only during times of low uncertainty. *bioRxiv (Cold Spring Harbor Laboratory)*, 1.

[Zhu et al.2024] Yi Zhu, Mahsa Abdollahi, Ségolène Maucourt, Nico Coallier, Heitor R. Guimarães, Pierre Giovenazzo, and Tiago H. Falk. 2024. MSPB: a longitudinal multi-sensor dataset with phenotypic trait measurements from honey bees. *Scientific Data*, 11(1), 8.