Temporal Dynamics and Change

Lotte Michels

1 Introduction

For this Complex Systems project, we apply complexity science methods to investigate the dynamics of a beehive. We regard a beehive as a complex system, because it consists of many individual bees collectively producing complex behaviors, such as foraging, honey production, intrusion defense and temperature and humidity regulation. These behaviors emerge without any central control and dynamically change in response to environmental factors (Research and Consortium, 2010; Zhu et al., 2024).

The Multi-modal Sensor Dataset with Phenotypic trait measurements from honey Bees (MSPB) provides several longitudinal measurements in beehives (Zhu et al., 2024). In this first entry of our e-portfolio, various exploratory data techniques will be applied to create an initial understanding of the temporal dynamics within these data. Specifically, the various types of change, stationarity and stability within a singular beehive will be taken under the loop.

2 MSPB Dataset Description

We will analyze parts of the Multi-modal Sensor Dataset with Phenotypic trait measurements from honey Bees (MSPB). Zhu et al. (2024) created this dataset by monitoring 53 beehives in Quebec between April 2020 and April 2021, collecting both longitudinal time-series data and phenotypic annotations.

The time-series consists of temperature, humidity and audio measurements. These signals were recorded every 15 minutes over one year of time. Temperature was measured inside the beehive chambers in degrees Celsius and humidity was measured inside the beehive chambers as a percentage. Moreover, audio recordings were automatically preprocessed to continuously compute the beehive 'power' of the hive. Conventionally, Zhu et al. (2024) use beehive power as a proxy for colony activity: a higher power score reflects a stronger buzzing, indicating an increase of bee presence in the hive. Hive power is measured in decibels (dB) and computed using the fast Fourier transform features covering frequencies between 122 Hz and 515 Hz. Please refer to the paper by Zhu et al. (2024) for an elaboration on the mathematical power formulas.

The phenotypic annotations are beehive characteristics that are sampled either once or a discontinuous and limited amount of times. The traits include number of brooding cells in the hive, honey production, defense performance, hygiene performance and pesticide infestation level.

For this portfolio entry, we discuss and demonstrate complexity science analysis methods of temporal dynamics and patterns of change in the context of the MSPB dataset. We thus restrict ourselves to the time-series signals in the dataset (hive power, temperature and humidity) and investigate these for one single selected beehive: hive number 02204, located in the Coté apiary in Quebec.

3 Temporal Signatures

3.1 Types of Change in the MSPB Dataset

The behavior of a complex system over time is called its temporal pattern (Butner, 2018). A temporal pattern is not just the sum of the system's constituents, rather it emerges from non-linear interactions between the individual time-series that occur within the system. Such time-series can display various types of change over time. Specifically, by means of visual exploratory inspection, a time-series can be assigned one of three temporal signatures (Butner, 2018):

1. **First-order change**: a signal shows constant, linear or curvilinear change over time.
2. **Second-order change**: a signal shows periodic behavior with a constant, changing or oscillating frequency component.
3. **Third-order change**: a signal shows chaotic and seemingly unpredictable behavior that heavily depends on the initial system conditions.

As a first step in the analysis of the MSPB time-series, we can explore which types of change are featured in the data. This dataset quickly reveals the scale-dependent nature of time-series: while temporal signatures may be obscured when viewing data over extended periods (e.g. multiple months), zooming in to shorter time spans (e.g. a week or a day) can uncover structures that are otherwise difficult to detect, or vice versa.

Figure 1 visualizes the scale-dependent nature of temporal structures. For example, plotted for the whole data collection period, the temperature signal shows third-order change: this signal features chaotic, non-linear and non-periodic change, with a trend downwards toward the end of the sequence. In contrast, the temperature plot for a randomly selected week within August reveals a clear second-order type of change: the signal shows seven peaks and valleys, one for each day. This daily structure is even more defined in the plot made

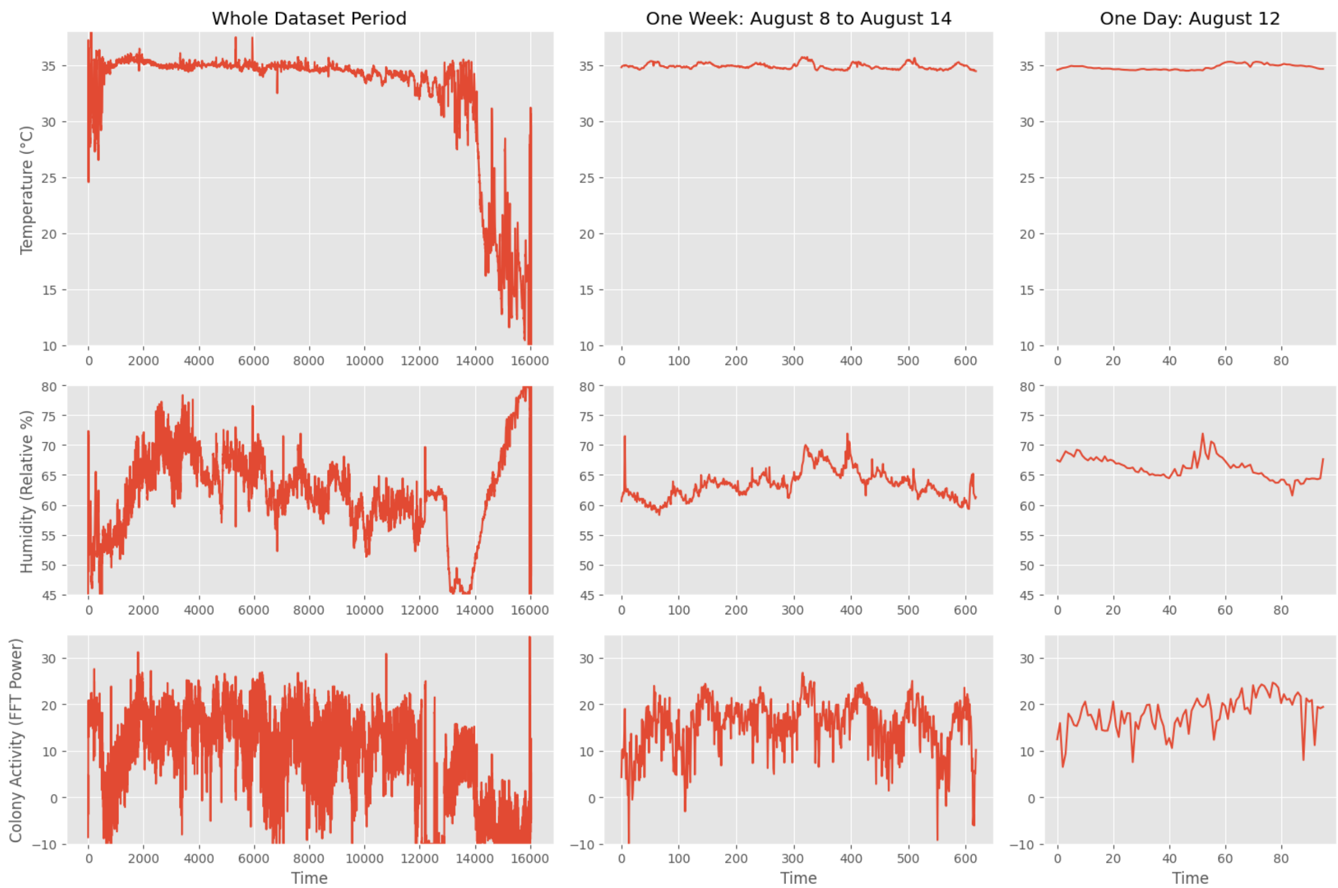

Figure 1: Raw time-series from MSPB beehive nr. 02204 plotted over varying time series: temperature (top row), humidity (middle row) and activity (bottom row) with respect to the full measurement period (left column), one randomly selected week (middle column) and one randomly selected day within that week (right column).

for a randomly selected day in the August week. The oscillatory structure from the weekly plot, however, is no longer observable in the daily temperature plot.

The humidity signal also shows a third-order temporal signature for the full data period: the plot shows chaotic, non-linear and non-periodic changes with varying trends over time. For the weekly plot, some sharp peaks and valleys may indicate an oscillatory trend, suggesting second-order change. Such peaks and valleys are visualized more clearly in the daily plot for humidity. However, focusing purely on this daily graph, one might argue that these peaks exhibit arbitrary trends and classify them as a third-order type of change.

Regarding the beehive activity/power signal, the full data plot appears very noisy and may reflect third-order change, but it could be argued that some global second-order trend is present: an overall oscillatory wave may be recognized, with a downwards trend towards the end of the data period. Zooming into one week of data reveals a more refined, but still noisy, oscillatory pattern. This indicates second-order change throughout a week in the data. Zooming into one day within this week, the power signal exhibits third-order change again. Zhu et al. (2024) explain that small fluctuations may occur in the power signal due to environmental noises (like human speech, rain or wind) picked up by the audio recordings. This may explain why the power signal looks noisier than the humidity and temperature time-series.

Overall, the plots in Figure 1 - mostly those for the temperature signal - thus demonstrate the importance of exploring time-series on various time scales in order to confidently label them with the correct temporal signature. Depending on the research goal, some time scales may be more interesting than others. For our e-portfolio, we will analyze the full time-series within the MSPB dataset. This allows us to also explore seasonal honeybee behaviors (Research and Consortium, 2010), for instance when looking at pairwise coordination in entry 6. Keeping the scale-dependent nature of the beehive time-series in mind, we will regularly zoom in on smaller time periods within the data to account for local effects in our consecutive analysis.

3.2 Downsampling Effects

Time-series signals are sequences of measurements taken at a constant rate in time: the sampling rate. Choosing a suitable sampling rate is crucial to capture the actual temporal patterns within a signal (Butner, 2018). If sampling rate is too low, undersampling may occurs. In this case, informative high-frequency dynamics are lost and the underlying temporal signature cannot be correctly identified. On the other hand, a very high sampling rate may lead to oversampling, in which case the proximity of measurements can make it difficult to

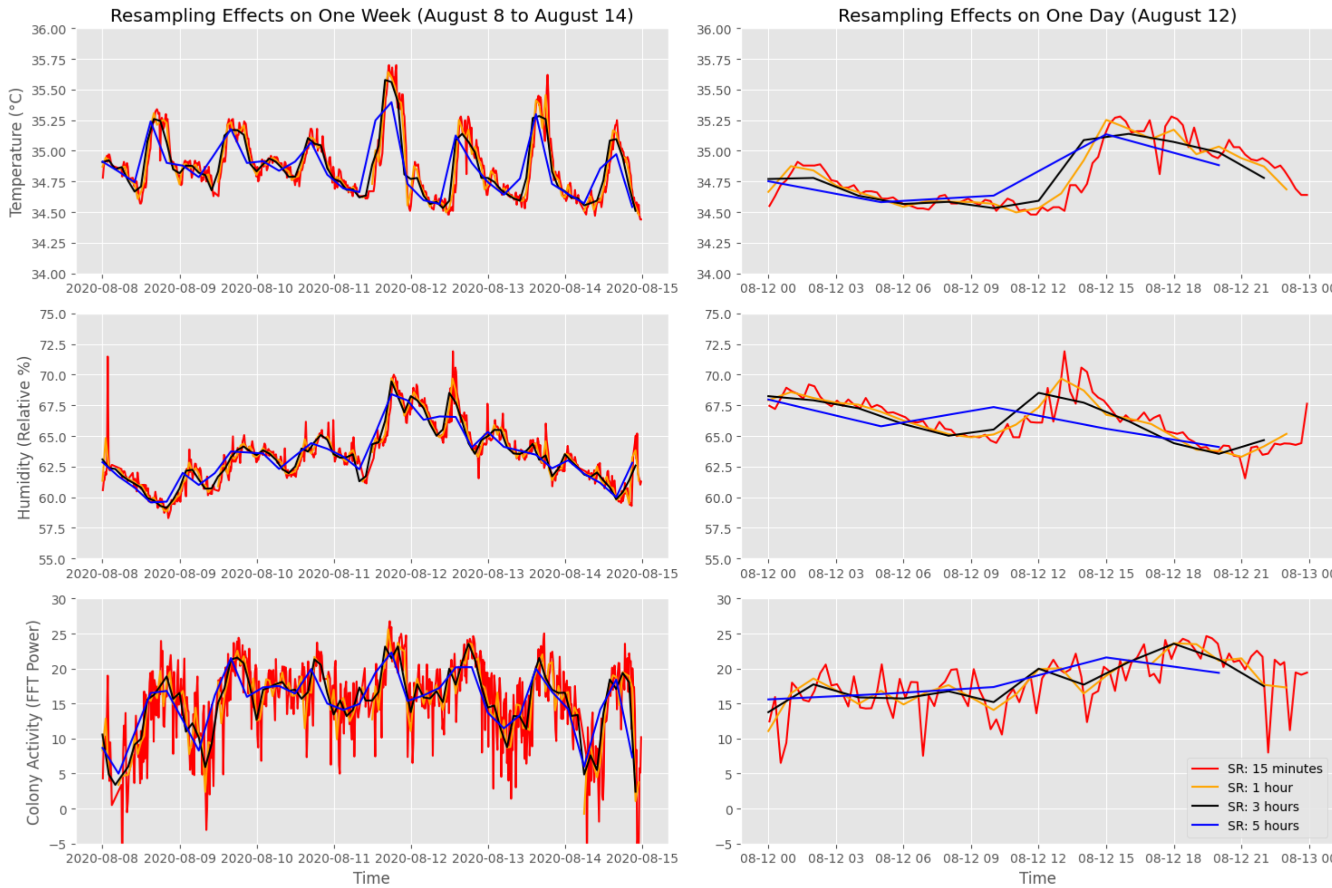

Figure 2: Sampling effects on the temperature, humidity and hive power signals of various sampling rates (SRs range between 15 minutes and 5 hours). For clarity, downsampling effects are displayed over the timecourses of one week (left) and one day (right).

detect patterns when the signal is plotted. In general, the higher the order of the temporal signature you hypothesize for your signal of interest, the more observations you want to make (Butner, 2018).

The time-series in the MSPB dataset are recorded every 15 minutes. Their values can be downsampled to check if the temporal signatures identified in these signals are susceptible to changes in sampling rate. Downsampling entails that the temporal resolution of a signal is reduced by grouping values into larger time intervals. For instance, 15 minute recordings can be converted to hourly averages. Downsampling is often applied to smooth out signal noise, but it can also lead to a loss of fine-grained trends in the signal.

We examined the downsampling effects for three different temporal resolutions: 1 hour, 3 hours and 5 hours. The results on the weekly and daily plots (which make resampling effects most clear) are displayed in Figure 2. These graphs show that down-sampling removes value variations within small time-scales, but puts more emphasis on the global signal trends. Whereas the order and complexity of the signal attenuate in the daily plots with decreasing temporal resolution, the temporal signatures and trends remain mostly identifiable in the weekly plots.

4 Time-Series Stability

4.1 Stability of Signal Values

To further investigate change, we can look at the distribution of the values and residuals in a time-series. Histograms or density plots (smoothed histograms) of raw values give an indication of the stable values that a signal is centered around, which will form peaks in the plot (Butner, 2018). These values represent the states that a signal tends to center around. The narrower a density plot is around a peak, the closer the signal stays to this value. Conversely, the wider a density plot is, the more the signal tends to deviate from this central value. Density plots can show more than one peak, in which case the signal features multiple stable states: this is called multistability. Alternatively, if a density plot shows no notable peaks, this may indicate an unstable temporal course of the signal (lack of stable states).

The density plots for the raw temperature, humidity and power signal values are displayed in the top row of Figure 3. For all three signals, these plots indicate multistability. The temperature plot shows a sharp peak around 35 degrees celsius. This peak makes sense in the context of typical beehive phenomena: thermoregulation, a well-known emergent honeybee colony characteristic, keeps the beehive temperature more or less constant between 33 to 38 degrees (Research and

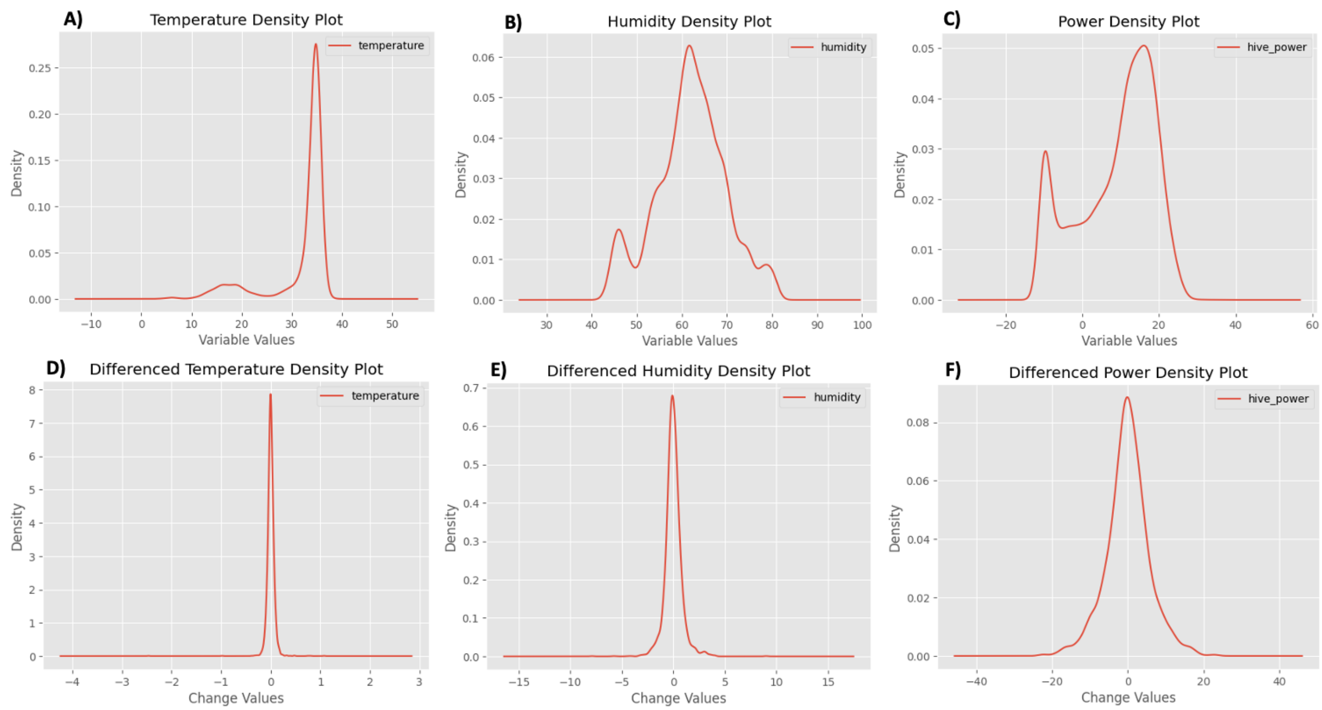

Figure 3: Density plots visualizing the stability of temporal signals and their change. A-C: density plots of the raw temperature, humidity and power values. D-F: density plots of the residuals within the respective signals.

Consortium, 2010). Effectively controlling this temperature level has been found to increase honey production and brood development and decrease bee mortality rates (Research and Consortium, 2010; Zhu et al., 2024). Similarly, the optimum levels for beehive humidity have been found to lie between 50% and 60% (Zhu et al., 2024). This is also reflected in the main peak in plot B in Figure 3.

Both the temperature and humidity density plots also show smaller peaks, around 18 degrees celsius and 45% and 80% respectively. It appears that these smaller peaks are created by sudden signal fluctuations at the end of the measurement period, as visible in Figure 1. According to Zhu et al. (2024), such sudden fluctuations regarding internal beehive temperature and humidity may reflect the presence of an active queen bee or the occurrence of honeybee swarming behavior.

For the beehive power, the density plot in Figure 3 also indicates two stable states. Most of the signal values are centered around a power value of 18 dB. The plot features an additional peak around -10 dB, which seems to be the cutoff of the power computation (Zhu et al., 2024).

4.1.1 Measurement Errors Effects

In addition to his notions of temporal patterns, Butner (2018) described three notions of variability in a signal:

1. **Perturbations**: disturbances that cause a system to deviate from its stable state(s).

2. **Random Walk**: cumulative noise that causes a drift in the signal trend.

3. **Measurement Error** (ME): imprecisions in signal recordings, possibly introducing biases in the data.

Contrary to the first two notions, ME can change the 'real values' in a sequence and thus affect the signal stability. To assess how susceptible a system is to ME, various degrees of random noise can be added to explore the effects on the system's stable states.

We examined ME effects on the beehive temperature, humidity and power signals by adding random noise with increasing standard deviation units to the time-series, and plotting the resulting density curves against each other. These curves are displayed in Figure 4. For all three signals, the graphs show that the most prominent peaks in the density plots remain, but are flattened for large standard deviation settings. Moreover, the smaller peaks in the original multi-stability density plots quickly attenuate as ME noise is added. This attenuation is most prominent for the -10 dB peak in the power density plot, suggesting this peak may be a bias in the data created by measurement error. Zhu et al. (2024) also described the power signal to be most sensitive to such noise, as the audio recordings used to compute the power signal may pick up on environmental sounds.

4.2 Stability of Change

In addition to the raw signal density plots, we can create similar graphs for the sequence residuals to get an idea of the stability of temporal *change* in a time-series. That is, we compute the differences between consecutive values in a signal and plot these to investigate the distribution of signal fluctuations. If a residual density plot is centered around zero, this suggests mean signal stability. Moreover, a narrow distribution/density plot indicates consistently small value changes between consecutive timesteps, which may suggest signal sta-

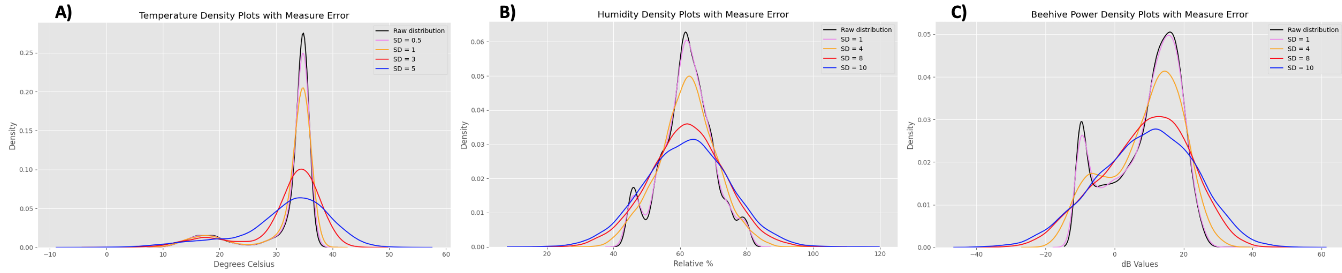

Figure 4: Density plots visualizing signal stability under varying degrees of random noise. For temperature (A), noise standard deviations range from 0.5 to 5. For humidity and power (B and C), noise standard deviations range from 0.5 to 10.

tionarity. On the other hand, flat and spread-out plots display large and variable differences over time and may suggest signal non-stationarity. Lastly, asymmetry in residual density plots may indicate an up- or downwards trend in the signal (for left- or right-skewed distributions respectively). However, it is important to note that conclusions about signal trend and stationarity should be supported and mostly drawn from visual inspection and statistical tests.

The density plots for the residual temperature, humidity and power signal values are displayed in Figure 3. All these residual plots are centered around zero, which indicates that all signals are stable around their means: 31.8 (sd: 6.2), 62.0 (sd: 7.8) and 8.7 (sd: 10.0) respectively. Matching the standard deviation statistics, the residual density plot is most narrow for temperature, followed by humidity and lastly by the power residuals. Similarly, value differences between consecutive measurements are smallest for temperature, followed by humidity and lastly by the power signal.

Based on the analysis displayed in Figure 3, we can thus infer that out of the three time-series, the temperature signal is most stable and the power signal is most unstable. As explained, the narrow density plots may indicate signal stationarity. However, more fine-grained analyses are necessary to draw such conclusions. We will therefore look at stationarity tests in the next section.

5 Statistical Stationarity Analysis

So far we have inspected the dynamics in MSPB time-series with regards to their types of change and stability. The residual density plots from the previous section suggested that the temperature, humidity and power signals may be stationary over time. Stationarity in a sequence entails that the dynamics in a signal do not change over time (Butner, 2018). It is important to assess stationarity in a signal, as some complexity science analyses may make the assumption that signal dynamics remain stable over time. An example would be..., covered in entry... of our e-portfolio.

To empirically assess whether a time-series is stationary, statistical tests can be conducted. Conventionally, we will apply and compare two such methods: the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Dickey and Fuller, 1979; Kwiatkowski et al., 1992).

While they both check for data stationarity, the ADF and KPSS tests have opposite null hypotheses: The ADF test's

Signal	**ADF**		**KPSS**	
	Test (statistic)	P-value	Test (statistic)	P-value
Temperature	False (-1.9)	0.35	False (3.17)	< 0.01
Humidity	**True** (-3.2)	0.02	**False** (2.6)	< 0.01
Power	True (-4.2)	< 0.01	True (0.4)	0.08

Table 1: ADF and KPSS test results for the stationarity within the MSPB time-series in beehive 02204. Contradicting test results are highlighted in bold text.

null hypothesis is that the time series has a unit root (i.e., it is non-stationary), while the KPSS test's null hypothesis is that the series is stationary. So, for the ADF test, a significant test statistic or low p-value suggests the signal is stationary, whereas for the KPSS test, a significant test statistic or low p-value suggests the signal is not stationary. They can be used complementarily to cross-validate conclusions about stationarity. Such cross-validation is important, as the tests may sometimes give contradicting outputs due to slightly different computational sensitivities. This is also demonstrated by Table 1.

Table 1 summarizes the ADF and KPSS test results for the beehive time-series. The dynamics in the beehive power signal appear to be stationary over time. In contrast, both tests classify the temperature signal as non-stationary throughout the measurement period. This finding matches with the signal's visualization in Figure 1. Regarding the humidity signal, the test results are inconclusive. Looking at the signal's plot in Figure 1, we argue the sequence looks some-what non-stationary. As such, we will not assume stationarity for the humidity signal.

6 Conclusion

In this first e-portfolio entry, we introduced the topic of our project and provided a description of the dataset that will be analyzed. Moreover, a variety of exploratory data techniques was applied to get a better idea of the temporal signals (temperature, humidity and colony activity (power)) that will be further taken under the loop in the upcoming e-portfolio entries.

It was found that all time-series display third-order change over the full measurement period, while some second-order oscillatory patterns may be identified by zooming in to a smaller time-scale. Moreover, high resistance to downsampling effects was found for these temporal signatures. Next, it was observed that all time-series tended towards multistability in their signal values. However, an analysis of measurement error effects indicated that this multistability dissipates quickly under randomly added Gaussian noise. In addition, all signals showed zero-centered uni-stability in their residuals, suggesting signal stationarity. However, ADF and KPSS stationarity tests indicated only the beehive power signal to be stationary over time.

References

Jon Butner. 2018. *Quantitative Reasoning Under a Dynamical Social Science*. University of Utah, 1.02 edition, 6.

David A. Dickey and Wayne A. Fuller. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366a):427–431, 6.

Denis Kwiatkowski, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3):159–178, 10.

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

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.