Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Modeling bee hive dynamics: Assessing colony health using hive weight and environmental parameters

Jürgen Degenfellner [a,*], Matthias Templ [b]

[a] Zurich University of Applied Sciences, Institute of Physiotherapy, Switzerland
[b] University of Applied Sciences and Arts Northwestern Switzerland, School of Business, Switzerland

ARTICLE INFO

Keywords:
Machine learning
Colony hive health
Colony health forecasting
Precision bee keeping

ABSTRACT

Our state-of-the-art methods study hive weight and predict bee health and colony condition using advanced machine learning tools trained with unlabeled data. By integrating methodologies such as signal extraction, similar trend monitoring, principal component analysis, and MM-Regression, our goal is to translate hive weight fluctuations into predictive insights for future hive monitoring systems. In particular, signal extraction methods are used to obtain an interpretable signal and to detect level shifts, exploratory analysis is used to visually detect dissimilar weight trajectories from nearby hives, historical data are used to robustly predict hive weights, and to trigger an alarm when predictions and actual observations differ significantly. Our study shows how these methods can be successfully used to analyze and predict hive weights and underscore the need for future research to accumulate labeled data and to adopt a holistic perspective, incorporating a wider spectrum of influences of hive weight simultaneously.

1. Introduction

Honey bees, particularly the western honey bee (Apis mellifera), play an essential role in our ecosystem, with around 30% of North American food resulting from bee pollination (Anon, 2022).

A typical colony fluctuates between 20,000 bees in winter and 50,000 in summer and exhibits intriguing behaviors, particularly from May to August during nectar flow and pollen availability. Each bee, weighing around 90 mg, can carry significant amounts of nectar and pollen, contributing to the colony's collection of up to 30 kg of pollen annually. Interestingly, the weight of the hive drops at night due to evaporation as the worker bees rest, and only 15% of the bees can simultaneously collect nectar and pollen. The productive period of bees lasts 10–20 days and a single colony can generate up to 200,000 collecting bees. These flights depend on temperatures being above 10–12 degrees centigrade, as bees cannot fly below these temperatures (Tautz and Heilmann, 2007).

Understanding the health and productivity of a bee colony traditionally required intrusive and labor intensive hive inspections. It is against this backdrop that the concept of remote, non-intrusive monitoring of bee hives has emerged as an appealing alternative that allows for continuous assessment of hive health and behavior, reducing both the labor and potential harm associated with traditional inspection methods (Pejić et al., 2022; Alleri et al., 2023).

The realm of wireless sensory networks for the analysis of bee hives is growing. The advances in sensor technology, now smaller, cheaper, and more precise, are amplifying its use in bee research and beekeeping. The simplified connectivity to computers and the Internet allows for continuous, remote, and low-labor monitoring of bee colonies. Once equipped with sensors, hives can be observed undisturbed, even during winter or stressful conditions, when hive inspections are not recommended. After connecting the sensors to the hive, data from the hive could be streamed to the cloud to continuously measure the weight of the hive, the temperature inside and outside the hive, the humidity inside the hive, acoustics or "buzzing" inside the hive, and flight activity inside and around the hive (Hadjur et al., 2022a).

In previous years, a variety of hive data analysis methods have been applied. Ziegler et al. (2022) used a vector autoregressive (VAR) model. Meikle et al. (2006, 2018) used detrended hourly bee hive data and simple linear regression slopes to model hive weights, Human et al. (2006) applied cosinor analysis to half-hour averages to evaluate whether relative and absolute humidity changes exhibited circadian rhythmicity and whether that rhythmicity differed between hives. Bencsik et al. (2011) proposed a non-invasive method to predict bee swarming behavior days in advance, using a transducer to monitor hive vibrations and apply specific weighting factors to those signals. Chen et al. (2012) developed an imaging system for non-intrusive monitoring of honeybee activity at the hive entrance, using infrared technology, computer image

* Corresponding author.
E-mail addresses: juergen.degenfellner@zhaw.ch (J. Degenfellner), matthias.templ@fhnw.ch (M. Templ).

https://doi.org/10.1016/j.compag.2024.108742
Received 31 October 2023; Received in revised form 1 February 2024; Accepted 10 February 2024
Available online 15 February 2024

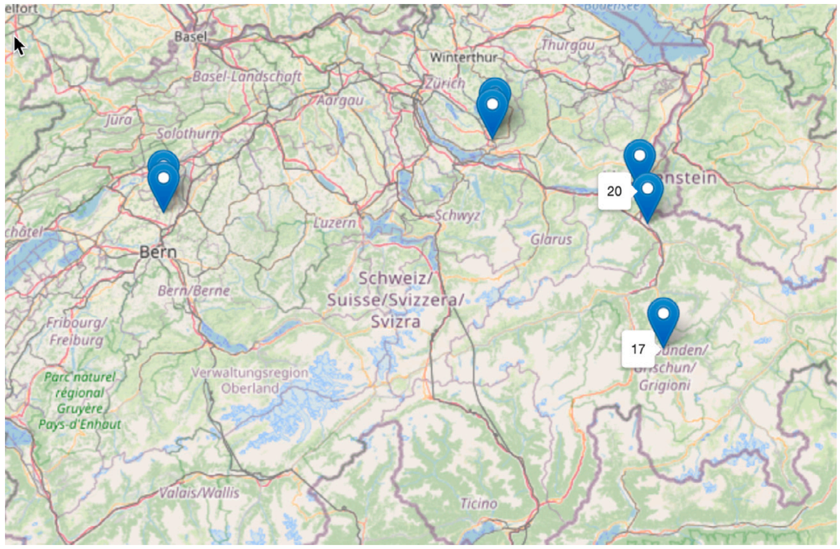

Fig. 1. Geographical locations of seven hives in Switzerland (created with leaflet (Cheng et al., 2023)). Locations 17 and 20 are annotated. These are at higher sea levels (915 m and 955 m) compared to the other hives from HiveWatch (max. 627 m in location 15).

processing, and circular character encoding tags on bees. The system accurately recognized and identified individual bees (with accuracy rates of 98% and 86%, respectively). Edwards-Murphy et al. (2016) used various data from a beehive, including levels of CO_2, O_2, and other gases, along with temperature, humidity, and acceleration, to describe internal conditions and colony activity which allowed for a classification of ten states of the hive with 95. 38% precision and also led to the development of an algorithm that could predict short-term rain based on the parameters of the hive.

Recently, Arias-Calluari et al. (2023) modeled how bee foraging and food processing activities affect global hive weight through a set of ordinary differential equations and showed how to estimate the parameters of this model from measurements on a single day. Gounari et al. (2022) analyzed five years of beehive weight data to assess how weather conditions affect bee productivity in an east Mediterranean island, finding that temperature and water-related parameters impact beehive productivity. Czekonska et al. (2023) used boosted regression trees (BRT) to analyze changes in hive weights that depended mainly on the observation period and the air temperature. Robustillo et al. (2022) used static and dynamic vector autoregressive models and linear and non-linear regression models to predict the internal conditions of the bee hives, finding that the dynamic vector autoregressive model generally provided the best predictions with a feasible computational cost. Kulyukin et al. (2022) investigated the relationship between beehive weight and bee traffic using a two-sensor system and found varying correlations across six bee colonies. The strength of the correlation was stronger over longer time intervals and exact traffic counts did not affect the correlation more than the spread and mean of the traffic. (Ramsey et al., 2020) and others have improved the prediction of swarms from previously several days (Bencsik et al., 2011) to up to 30 days in advance using within-hive vibration data. Research has shown that not only the weights of hives are important but also, for example, the tracing of bees within a colony. However, this refers to different approaches.

Many different commercial hive monitoring systems are on offer (HiveWatch, 2023; Arnia, 2023; Monitoring, 2023; OSBeehives, 2023; SolutionBee, 2023; Ntawuzumunsi et al., 2021; Cecchi et al., 2021). Terenzi et al. (2019) showed how one could build such a system by measuring the weight and sound of the hive. In our contribution, we focus on only measuring the weight of the bee hive and obtaining meteorological information from nearby weather stations. This requires fewer interventions with sensors in a hive and presents a cost-effective solution. However, there is a question about the precision and predictive power of models in predicting the condition of a bee colony.

The data were primarily unlabeled, except for certain swarming events. We adopted a conservative approach, aiming to detect level shifts and instances where the hive weights diverged from the expected values. Integrating both of these aspects algorithmically into a hive monitoring system can be achieved with reasonable effort. This study underscores the challenges and opportunities associated with modern methods to pinpoint events or irregularities in hive weight measurements using unlabeled data.

Main objectives:. As mentioned above, evaluating the well-being and productivity of a bee colony typically involves invasive and time-consuming examinations of the hive. Furthermore, beekeepers face various difficulties when it comes to labeling historical data related to events such as swarming, robbing, or external factors that impact the beehive. These challenges include time limitations, resource-intensive requirements, the need for expertise in data analysis to interpret weight data, and the consideration of multivariate dimensions. This is because the effects of such events can be subtle and easily mistaken for normal variations caused by weather or other natural factors.

Therefore, the main objectives of this study are to avoid these obstacles and to explore and demonstrate effective methodologies for future hive monitoring systems for unlabeled data, to detect unexpected and potentially detrimental changes in bee hives. Since raw beehive data can include measurement errors and outliers, we aim for more precise hive monitoring using robust statistical methods for signal extraction, shift detection, and alarming. The primary goal is to explore and demonstrate effective methodologies for future hive monitoring systems.

Enhancements and outline:. The literature is extended by (cf. Fig. 2)

- introducing methods to extract interpretable signals from beehive weight data and the detection of level shifts. Signal extraction methods and similar trend monitoring (STM) are used in combination with robust estimation, e.g., with repeated median (RM) regressions;
- interpreting biplots from a principal component analysis to detect swarming, robbing and disease by observing behavioral variations from multiple hives and

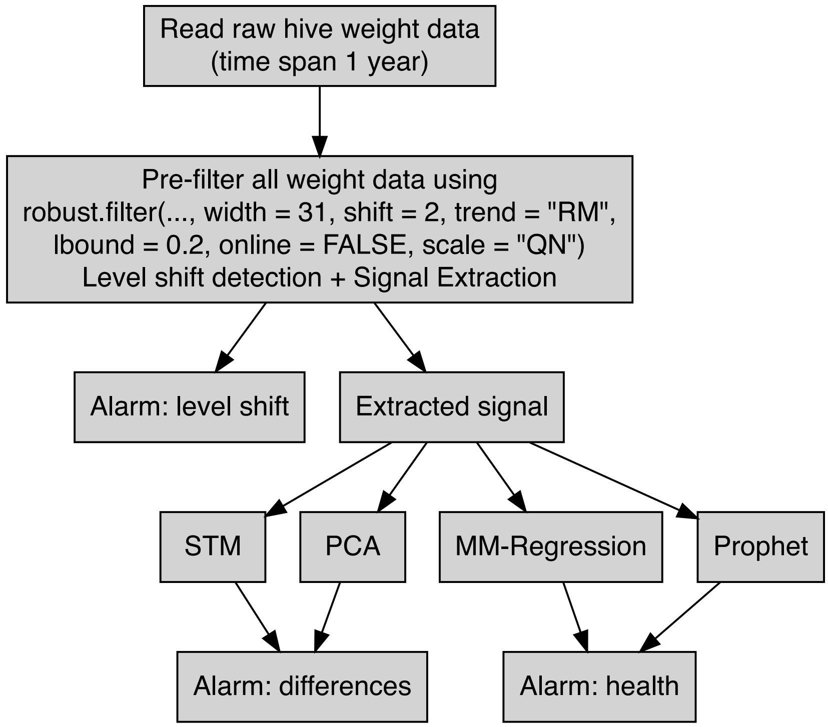

Fig. 2. Overview of the workflow and methods with hive weight data and how they relate to each other (created with DiagrammeR (Iannone, 2023)). For our analysis, 1 year hive weight data was available.

- triggering alarms when predictions of actual hive weights differ significantly from those observed. External information in the form of meteorological variables is used to model the weights of the hives over time. New and important is the use of robust estimation techniques to reduce the influence of outliers. The use of robust methods is essential and improves the quality of the model.

Section 2.1 introduces the data, and Section 2.2 introduces the newly proposed methods. Section 3 shows detailed results for two hives. In the discussion section (Section 4), the new methods and results are discussed in light of the existing literature and the possibilities and problems are mentioned. In addition, references are made to further work.

2. Materials and methods

2.1. Data sets

Data sets were obtained from HiveWatch (2023) from seven sites in Switzerland (Fig. 1, Table 3) with 3 (location 15) to 8 hives in one location. Hives in the same location allow for within-apiary comparison as they share the same environmental conditions, while comparing hives from different locations can indicate site differences. Since the hives were selected from HiveWatch customers, we had a convenient data set from a statistical point of view.

2.1.1. Beehive weight data

The weight of a hive is measured with a load cell-equipped weighing platform. The measurement range is 0 to 200 kg, with a resolution of 5 g and a measurement cycle of 60 s. The size of the weighing platform is customized for the size of a beehive. The weighing device works autonomously with a battery change interval of 1–2 years. A mobile phone connection SIM card is integrated, as well as an antenna to transmit data to a server. As already mentioned, the records are not labeled. Even events such as swarming or robbing are not labeled in the data.

2.1.2. Meterological information

HiveWatch sources weather-related variables using the Darksky API (Rudis, 2017) (one of many other external sources for the same weather data). They include the following measurements: the temperature outside the bee hive, relative humidity (abbr. rH, inside), humidity (outside the hive), cloud cover, precipitation intensity, precipitation probability, air pressure, visibility, dew point, wind speed, wind bearing, latitude and longitude. In Table 4, the variables of the HiveWatch data sets are explained, and Table 5 shows the hives in each location.

As hive research is still in its infancy, much of the available data is inefficient in terms of accuracy and completeness. The application of methods is therefore also based on the information available and the completeness of the information, which varies depending on the hive and location. It is important to note that 2019 was a very poor year for honey in Switzerland Bauernzeitung (2019). It was only good in the mountains, such as at locations 17 and 20 which were at sea levels of 915 m and 955 m and 51 km apart (distance calculated with geosphere (Hijmans, 2022)). Hence, we focus our analysis on these locations to demonstrate the methods. Location 21 was used to depict a visually larger level shift (Fig. 3).

2.2. Statistical analysis

2.2.1. Overview

Fig. 2 gives an overview of the workflow and methods with the hive weight data and how they relate to each other.

It shows that different methods are used for different purposes. When using raw data, significant fluctuations can hinder effective visualization and may have a large impact on different alarming methods. To address this, signal extraction (prefilter step in Fig. 2) can be applied as a technique to smooth the data and aid in its visualization. For monitoring systems, this smoothing can act as a preliminary filter to mitigate the impact of extreme data points before triggering alerts. For the following types of alarming (cf. Fig. 2) and associated methods, typically the extracted signal is used after robust filtering.

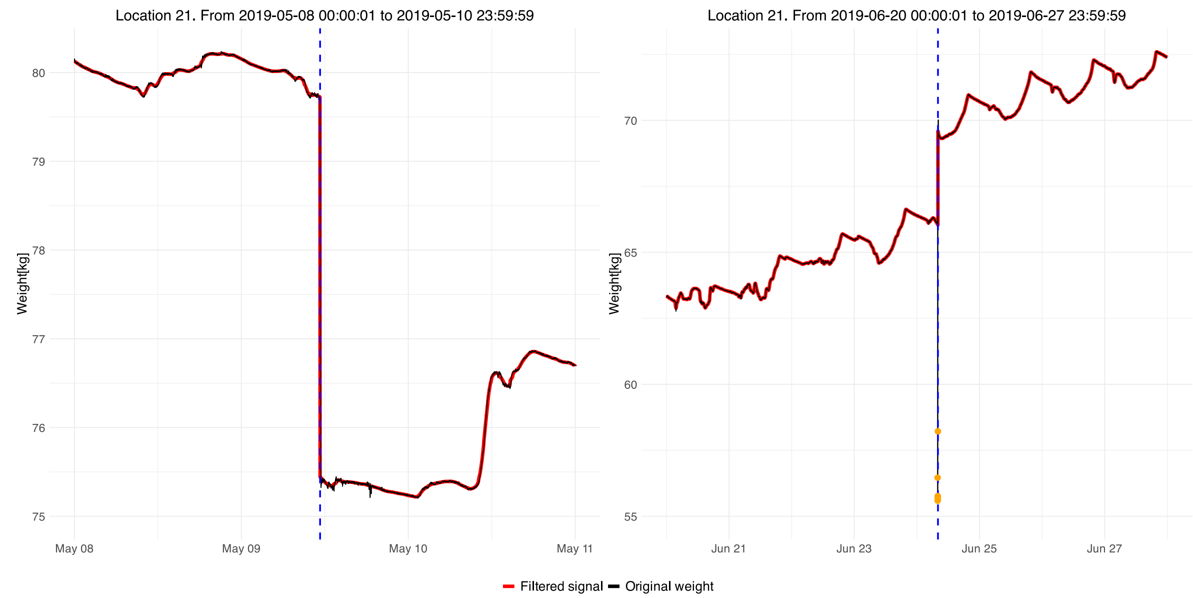

Fig. 3. Two examples of signal extraction from hive weight data. A window width of 31 (time steps) was used for the filtering algorithm. The blue dashed lines indicate level shifts that were automatically detected. **Left**: A level shift (downwards) is visible in the weight data and is detected. **Right**: A level shift (upwards) and an outlier-patch (orange dots) occur (sharp weight drop) during a foraging phase. Both are detected and filtered smoothly (red line).

Alarms for level shifts: The aim is to detect level shifts such as those caused by robbing, swarming, beekeeper interventions, hive inspections, or animal interference, see Fig. 2 (Alarm: level shift).

Signal extraction using robust filtering: In addition to previously mentioned pre-processing step to extract a signal, such pre-filters such as the one used in this work, can be already used for detecting level shifts.

Alarms for differences in hives Available data from different hives are compared and alarms are triggered when a hive behaves differently to the majority of hives. Two methods are suggested, with the second being of explanatory nature.

Similar trend monitoring (STM): The objective is to identify and notify about changes between hives. Robust filtering (see above) ensures that the comparison between hives is not solely influenced by level shifts, but instead emphasizes systematic differences between hives.

Principal component analysis (PCA): The goal is to visually examine the fluctuations in weight of several beehives in order to detect any hives that exhibit weight patterns that deviate from the majority of the hives. The purpose is to identify beehives that might be experiencing health problems, such as a decrease in weight due to infestation by Varroa mites.

Alarms for *health* indication: These alarms serve as an indicator of potential health problems in the beehive. In both suggested approaches, the covariates time, temperature, dew point, cloud cover, humidity, air pressure, wind speed, visibility, and precipitation intensity were included in the models.

MM-regression: This method is employed to activate alerts when there is a substantial difference between the weight of the beehive and the expected beehive weights.

Prophet: While an alarm is also triggered in the same manner, when predictions are too distinct from observed weights, the kind of method differs completely from MM-regression by using Fourier series to provide a flexible model.

2.2.2. *Signal extraction using robust filtering*

Signal extraction (Borowski et al., 2015; Borowski, 2013; Anon, 2023) was utilized as a preliminary step to the actual analysis, with the aim of extracting an interpretable signal. The same algorithms identified level shifts.

Given the high variability and level shifts in the data, possibly caused by external forces such as the beekeeper's intervention, it was necessary to extract the main signal. This signal represents a robust estimate of the hive weight or noise level, derived from the data provided every 1 or 5 min. We applied methods originally designed for analyzing data from intensive care unit (ICU) patients to hive data, employing various filtering methods. The challenge lay in striking a balance in the decision-making process between identifying a level shift and normal data reflecting bee behavior, and between distinguishing an outlier and a legitimate data point. The use of robust level shift detection allowed us to obtain an adequate signal for further analysis while keeping false-positives to a minimum.

The model and assumptions for our analysis are described by the following equations (see Fried Fried, 2004):

$$Y_t = \mu_t + E_t + \sum_{k=1}^{K} \sum_{i=0}^{\delta_k - 1} \omega_{k,i} 1_{\left\{\sum_{j=1}^{k} T_{j+i}\right\}}(t) \tag{1}$$

$$\mu_t = \eta_t + \sum_{l=1}^{L} \nu_l 1_{\left[\sum_{j=1}^{l} s_{j,\infty}\right]}(t) \tag{2}$$

The robust trend approximation is given by:

$$Y_{t+i} = \mu_t + i\beta_t + E_{t+i} + r_{t+i}, \; i = -m, \dots, m, \tag{3}$$

μ_t and β_t indicate the level and slope within the window, respectively. E_{t+i} is a random noise variable with a median of zero and a variance of σ_t^2. Meanwhile, r_{t+i} symbolizes the approximation error with a magnitude constrained by $\varepsilon\sigma_t$. The window moves at time points from $i = -m$ to $i = m$. The time series model explained herein handles outliers and level shifts. Random errors at time point t, denoted as E_t, are generated from a distribution with a median of zero and variance σ_t^2. The potentially infinite numbers of outlier patches and level shifts are represented by K and L, respectively. The times between subsequent outliers and level shifts are indicated by $T_j, S_j, j \in \mathbb{N}$. The duration of the kth outlier patch is δ_k with $\delta_k = 1$ representing an isolated outlier. The size of the ith outlier in the kth patch is denoted by $\omega_{k,i}$, and ν_l represents the size of the lth level shift. The model treats errors as white noise, disregarding correlations. The observation equation only accommodates additive outliers by adding a constant to the undisturbed value $\mu_t + E_t$. To incorporate substitutive outliers, which describe pure measurement artifacts, Y_t is replaced by a constant at the corresponding time points. This model assumes that η_t and σ_t vary smoothly over time, with the data residing on a fixed grid. Despite its complexity, this model finds specific applications in intensive care monitoring. However, extraction of μ_t can be challenging if outlier patches occur in short time lags or close to a level shift, underlining the need to draw the healthcare professional's attention to potential shifts or outlier patches.

A range of methods are available for robust approximation of the signal within a time window, such as the median (MED), repeated median regression (RM), least trimmed squares regression, and least median of squares regression (LMS).

For our data sets, we chose RM, since there were no large differences in filter quality compared to the other methods mentioned, and it is a well-known method that proved itself in practice, see, e.g., Fried (Fried, 2007, 2011, 2006). In order to estimate local variability robustly (within one-time window), there are (among others) the following methods available: Median absolute deviation about the median (MAD), Rousseeuw's and Croux's Q_n (QN) and S_n (SN) scale estimators (Rousseeuw and Croux, 1993) and length of the shortest half (LSH). We used Q_n, again because no relevant filtering quality differences were observed and Q_n has good properties (Rousseeuw and Croux, 1993).

2.2.3. Similar trend monitoring (STM)

We applied STM (Borowski et al., 2015), a robust technique for real-time tracking of pairwise coherences in nonstationary multivariate data streams, to detect deviating foraging or weight loss behavior. The STM is built around the concept that two variables are considered coherent at a particular time t if their current trends align. The current trend for each univariate data stream is continuously estimated by applying robust repeated median (Siegel, 1982) regression lines within moving time windows, encompassing the most recent observations. Given the potential for abrupt shifts and evolving trends in the data streams, the selection of the window width is critically important to ensure accurate trend estimation. As such, STM employs a recently devised method for determining the window widths at every time point, so that the presumption of a linear trend is warranted. The STM's monitoring statistic is computed as the difference between the RM estimations of the current slopes of two streams, which is then divided by an estimate of the standard deviation of this difference. The STM function we used is based on the mSCARM (*multivariate slope comparing adaptive repeated median*) filter and on the *similar slope monitoring*-statistics (SSM), see (Borowski, 2013; Anon, 2023). The SSM statistic $S_t(i,j)$ is defined as:

$$S_t(i,j) := S_t(\mathbf{Y}_{t,n_t(i)}, \mathbf{Y}_{t,n_t(j)}) = \frac{D_t(i,j)}{\sqrt{\widehat{Var}[D_t(i,j)]}}, k = 1, \dots, K \tag{4}$$

$$\text{whereas } D_t(i,j) := \hat{\beta}(\mathbf{Y}_{t,n_t(i)}) - \hat{\beta}(\mathbf{Y}_{t,n_t(j)}) \text{ with} \tag{5}$$

$$\mathbf{Y}_{t,n_t}(i) := (Y_{t-n_t(i)+1}(i), \dots, Y_t(i)), \mathbf{Y}_{t,n_t(j)} := (Y_{t-n_t(j)+1}(j), \dots, Y_t(j)) \tag{6}$$

The smaller the realization of $S_t(i,j)$, the more similar the two time series are.

For practical purposes in real-time analysis, an upper limit c_{bound} can be defined. If the absolute value of $s_t(i,j)$ does not exceed this upper limit, the two time series are considered similar:

$$|s_t(i,j)| = |s_t(j,i)| \le c_{bound} \tag{7}$$

Our data exhibited two types of disturbances: obvious outliers resulting from calibration, measurement errors, or human intervention, and more subtle disturbances that raised questions about the use of signal extraction. To handle these disturbances, we employed robust level shift detection and signal extraction techniques. For shift detection, we used *robust.filter* from the package robfilter (Fried et al., 2022) with the following parameters: width = 31, shiftd = 2, trend = "RM", lbound = 0.2, online = FALSE, scale = "QN". Specifically, shiftd = 2 indicates the factor with which the current scale estimate is multiplied with for shift detection. The default is shiftd = 2 corresponding to a 2 σ rule.

2.2.4. Principal component analysis (PCA)

PCA and visual representation using biplots were used for the exploratory and interpretative analysis of data from more than one hive to detect swarming, robbing, or disease by observing behavioral variations in multiple hives.

PCA was used for data from more than one beehive. We hypothesized that events of interest, such as swarming, robbing, or diseases, could be detected by observing deviant behavior in one or two hives, as visualized in the biplot of the PCA (see Figs. 9, 10, 11, 12). A biplot in principal component analysis is a graphical representation that combines both the selected score variables and the corresponding loadings of these variables in a reduced-dimensional space, typically in $\mathbb{R}^2$. It allows us to visualize the relationship between variables and observations simultaneously. We used the angle between the variables (vectors) in the biplot as a measure of arrhythmic behavior and an indicator of hives deviating behavior (see, e.g., Fig. 9). Note that different amplitudes are not well recognized with this visualization technique.

2.2.5. Robust MM-regression

The prediction of the actual weight of the hive was carried out with robust linear regression (Maronna et al., 2006) using historical data to trigger an alarm when predictions and actual measurements differ significantly.

It turned out that robust MM-regression outperformed classical and also non-linear approaches like generalized additive models to predict weight change behavior within days. This approach allowed us to account for potential outliers in the data while maintaining the robustness of our predictions. The advantage of MM-regression is that outliers are automatically detected and down-weighted to have a bounded influence on our estimates. For data without outliers, we obtain almost identical results to those obtained with ordinary least squares regression. For data with outliers, we get reliable results and outlier detection is automatically part of the process.

To introduce robust estimators, we start with our linear regression model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$, with $\mathbf{y}$ the response vector with n observations, $\mathbf{X}$ the design matrix of dimension $p \times n$, $\boldsymbol{\beta}$ the p regression coefficients, and ϵ the n residuals $\hat{\epsilon}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_q x_{iq})$. In ordinary least squares estimation, the solution is obtained by minimizing the sum of squared residuals. The general idea of MM-regression is to replace the residual squares $\hat{\epsilon}_i^2$ with another function of the residuals $\rho(\hat{\epsilon}_i)$ that limits the influence of outliers and even make the residuals *scale-free*. Here ρ is a symmetric function of the residuals with minimum at 0. The robustness properties derive back to ρ and to its derivative $\psi = \rho'$, because for optimization, the derivative is set to 0.

The iteratively reweighted least squares algorithm (IRLS) gives a solution of the problem

$$\hat{\boldsymbol{\beta}}_S = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} s\left(\hat{\epsilon}_1(\boldsymbol{\beta}), \dots, \hat{\epsilon}_n(\boldsymbol{\beta})\right)$$

$s\left(\hat{\epsilon}_1(\boldsymbol{\beta}), \dots, \hat{\epsilon}_n(\boldsymbol{\beta})\right)$ is defined as the solution of

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{\hat{\epsilon}_i(\boldsymbol{\beta})}{s_\epsilon(\boldsymbol{\beta})}\right) = K$$

The derivative of the well-known Tukey's biweight function is used as the ρ function. For modeling the hive weights, we model the hive weights with variables temperature, pressure, wind speed, precipitation intensity, cloud cover, dew point, and time of day. Using our model estimates and the explanatory data for the new day, robust prediction

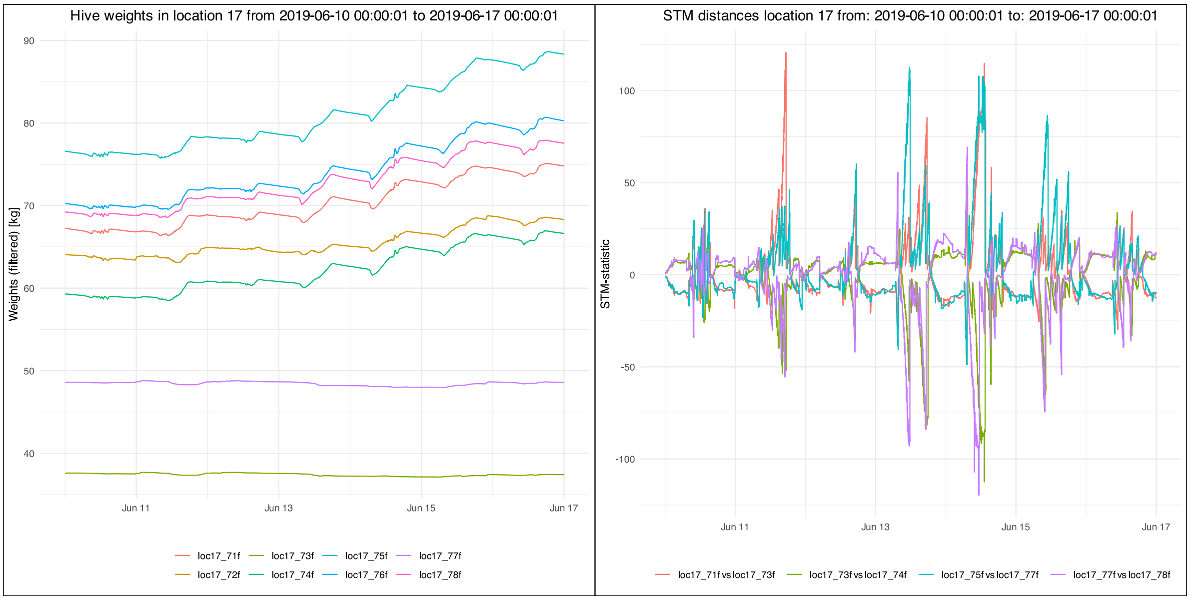

Fig. 4. **Left**: The graph displays the filtered (absolute) weights of hives situated at location 17, spanning the period from June 10th, 2019 to June 17th, 2019. **Right**: The associated measures of distance, derived from STM statistics, are presented for the identical time span. To enhance clarity in visualization, only hive pairs that registered distance measures surpassing 110 at least once have been plotted. Note, there is a total of $\binom{8}{2} = 28$ distinct pairs of hive distances.

intervals for hive weights are then estimated. A larger deviation of the predictions was reported compared to the actual measurements. The alarm threshold was that the actual weight of the hives was outside the prediction intervals. MM-Regression was called using *lmrob* (Salibian-Barrera and Yohai, 2006) in the R package robustbase (Rousseeuw et al., 2009) with the parameters method = "MM" and cov = ".vcov.w".

2.2.6. Facebook's prophet

In addition, we used Prophet (Taylor and Letham, 2021), a time series forecasting model with covariates designed to handle the common features of business time series and applied it to detect if actual hive weights deviate from predictions. It was developed by Facebook and is available as open-source software in Python and R (Taylor and Letham, 2021). The model is designed to have intuitive parameters that can be adjusted without knowing the details of the underlying model, making it accessible to analysts without specific expertise in time series modeling. It is based on a decomposable time series model with three main components: trend, seasonality, and holidays.

Trend: Prophet automatically detects changes in trends by selecting change points from the data; Seasonality: Prophet uses Fourier series to provide a flexible model. It allows for fitting seasonal patterns that change more quickly, albeit with increased risk of overfitting; Holidays and Events: Prophet allows the analyst to provide a custom list of past and future events, identified by the unique name of the event or holiday. The impact of a particular holiday on the time series is often similar year after year, so it is important to incorporate it into the forecast.

The model equation is given as:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (8)$$

where: $g(t)$ represents the trend function which models non-periodic changes; $s(t)$ represents periodic changes (e.g., weekly and yearly seasonality); $h(t)$ represents the effects of holidays that occur on potentially irregular schedules over one or more days; and ε_t represents the error term accounts for any unusual changes not accommodated by the model.

The model is designed to handle the common features of business time series such as multiple strong seasonalities, trend changes, outliers, and holiday effects. It also allows for manual review and adjustment, enabling reliable and practical forecasting of business time series. The command *prophet* in the package Prophet (Taylor and Letham, 2021) was called using the parameter daily.seasonality = TRUE (others were set to default).

According to Tautz (Tautz and Heilmann, 2007), the most active period of the bee year is approximately from Mai to August, depending mainly on the availability of pollen and nectar. Hence, we focus our analysis examples on this data range. We tried to analyze typical hive weight behaviors in multi-hive systems in an exemplary fashion to showcase the possibilities of the methods.

The data were basically unlabeled with regard to all other than (some) swarming events. We chose a conservative approach and tried to detect level shifts and events in which the hive weights deviated from the predicted weights. Both can be algorithmically included in a future hive monitoring system with reasonable effort.

Software: We used the R package robfilter (Fried et al., 2022) for signal extraction. This package includes methods for robust trend approximation, such as the least median of squares (LMS) and the repeated median (RM), which are high breakdown point methods for regression and scale estimation. For principal component analysis and biplots we used the standard tools in the stats package of R. Robust MM-regression was made with the R package robustbase (Rousseeuw et al., 2009) and its function lmrob (Salibian-Barrera and Yohai, 2006), and the alarming is self-implemented. We additionally used the implementation of Prophet from the R package Prophet (Taylor and Letham, 2021, 2018).

3. Results

3.1. Signal extraction and similar trend monitoring (STM)

Fig. 3 shows two examples of robust filtering for weights described in Section 2.2.2.

As the data sets were not labeled by beekeepers, we cannot exactly quantify the correctly classified shifts and the corresponding false-negative and false-positive rates. However, as can be seen in Fig. 3, the shifts were generally well detected and we could not find any situation where a large shift was not detected by this approach.

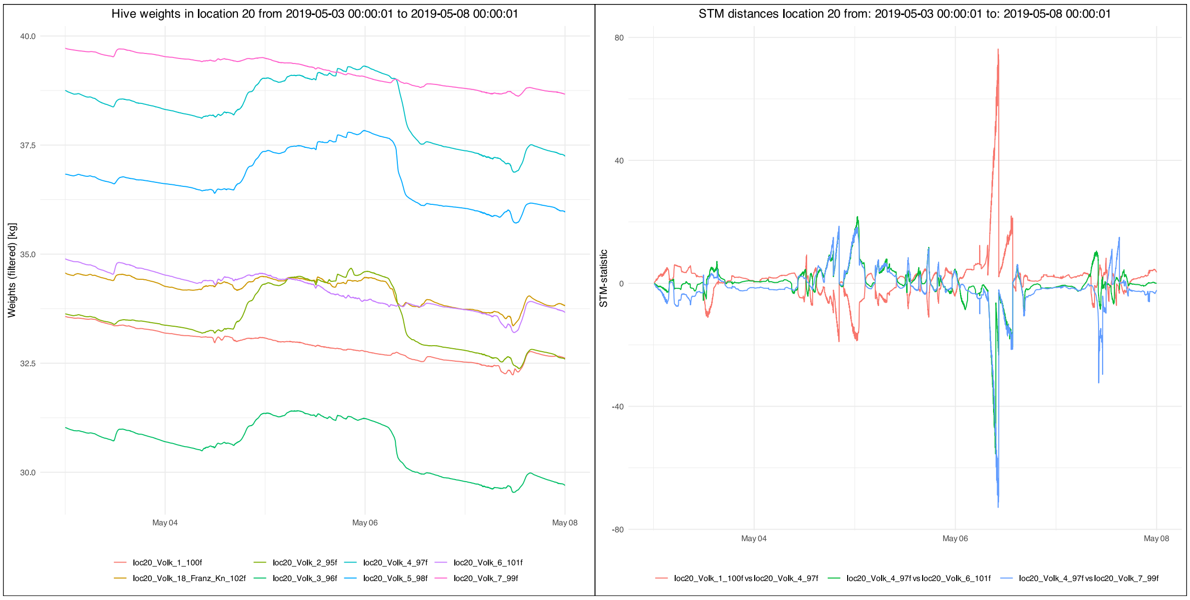

Fig. 5. **Left**: The filtered (absolute) weights of hives at Location 20, tracked over the period from May 3rd, 2019 to May 8th, 2019. **Right**: Corresponding measures of distance (via STM statistics) over the same time frame. For the sake of clarity, only those hive pairs that registered distance measures exceeding 50 at least once have been plotted. In total, there would be $\binom{8}{2} = 28$ distinct hive pair distances.

Fig. 4 presents the filtered (absolute) hive weights from location 17 on the left and the corresponding distance measures, or STM-statistics, on the right. The patterns suggest that the weights of hives 73 and 77 follow a comparable trajectory. Similarly, the remaining six hives also appear to demonstrate parallel weight patterns. Note that the distances for hive 73 or 77 in combination with other hives – but not both – are plotted. This observation aligns with the interpretative essence of the STM statistic: smaller values indicate a higher degree of similarity between the vectors.

In Fig. 5, the filtered (absolute) weights of eight hives from Location 20 are presented (on the left). It is possible to distinguish two distinct clusters based on the trajectories of the hive weights. Hives numbered 99–101 exhibit a noticeable decrease in weight, while hives numbered 95–98 demonstrate a weight increase around May 5th to 6th. This observation is corroborated by the STM statistics displayed on the right side of the figure. Larger dissimilarities are discernible between, but not within, these two clusters.

3.2. Shift detection

3.2.1. Location 17

In the monthly overview plots (refer to Fig. 17), there are several discernable shifts. For example, the pronounced shift in May is accurately detected across all hives. An example of this can be seen in Fig. 6, which shows the hive labeled as 'loc17_78'. Here, an upward shift is observed at 14:21:21 on May 17, 2019, most likely due to feeding or adding elements to the hive.

Furthermore, all other major level shifts are accurately identified. These include shifts on June 7, 2019, July 3, 2019, and August 8, 2019, as well as three separate instances on August 12, 2019, and another on August 27, 2019.

The hive labeled as '17_72' exhibits an anomalous pattern of behavior in June. This hive experienced numerous noticeable shifts during this period. A significant upward shift was recorded at 12:10:12 on June 6, 2019. This spike in weight, most likely attributable to rainfall (with a recorded precipitation intensity of 0.31), appears to be a false-positive, as the surge seems to have occurred naturally, rather than due to any undesirable hive activities. The other hives in location 17 showed the same shift (except hives 73 and 77 that were inactive; see 17). The hives in location 20 showed different behavior and now shifts were detected probably due to less rainfall at that time (precipitation intensity of 0.07).

Another upward shift was observed at 08:09:01 on June 7, 2019, during which there was a notable weight fluctuation. Given the scale of the variation, it is plausible to interpret this incident as an intervention by a beekeeper, suggesting that the detection was accurate.

On June 9, 2019, at 10:05:27, a substantial decrease in weight – approximately 2.5 kg – was detected. According to HiveWatch, this drop is indicative of a swarming event. Another upward shift recorded at 14:57:19 on June 14, 2019, showed a weight gain of less than 1 kg. This rise is most likely a false-positive due to its similarity to other hives in the location (Fig. 7).

Finally, two abrupt weight drops were identified at 13:24:52 on June 20, 2019, and 13:23:43 on June 21, 2019. According to the insights from HiveWatch, these decreases were correctly identified as swarming events. These instances underscore the complex and dynamic nature of hive activities, making it essential to monitor and interpret hive weight data with a discerning eye.

3.2.2. Location 20

The behavior exhibited by this location is marginally more chaotic compared to location 17. Several prominent level shifts are visible (Fig. 18), all of which have been accurately identified, mirroring the correct detections at location 17.

As an example, we applied shift detection to hive 101 (see Fig. 8). Two distinct shifts were detected at 11:08:00 on May 16, 2019, and 11:47:33 on May 23, 2019. Upon detailed examination of these abrupt changes, it is evident that these shifts were accurately identified as level shifts (but without known cause because of lack of labeling).

3.3. Exploratory analysis with biplots

3.3.1. Location 17

Fig. 9 illustrates an exploratory principal component analysis (PCA) that suggests similarity between hives 73 and 77, as evidenced by their nearly parallel trajectories. On the contrary, the remaining hives seem to display a more synchronous pattern according to the biplot.

However, hive 72 deviates slightly, demonstrating a smaller slope until the end of June, which is more evident in a broader context (see Fig. 17, upper right).

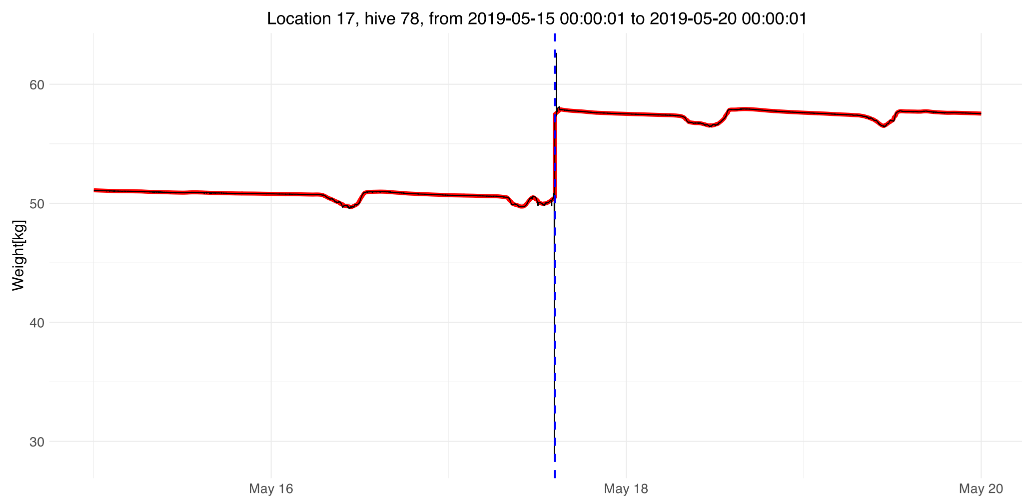

Fig. 6. Filtered (absolute) weights of hive '17_78' located at site 17. A notable upward shift is detected at 14:21:21 on May 17, 2019, which is represented by a dashed blue line in the figure.

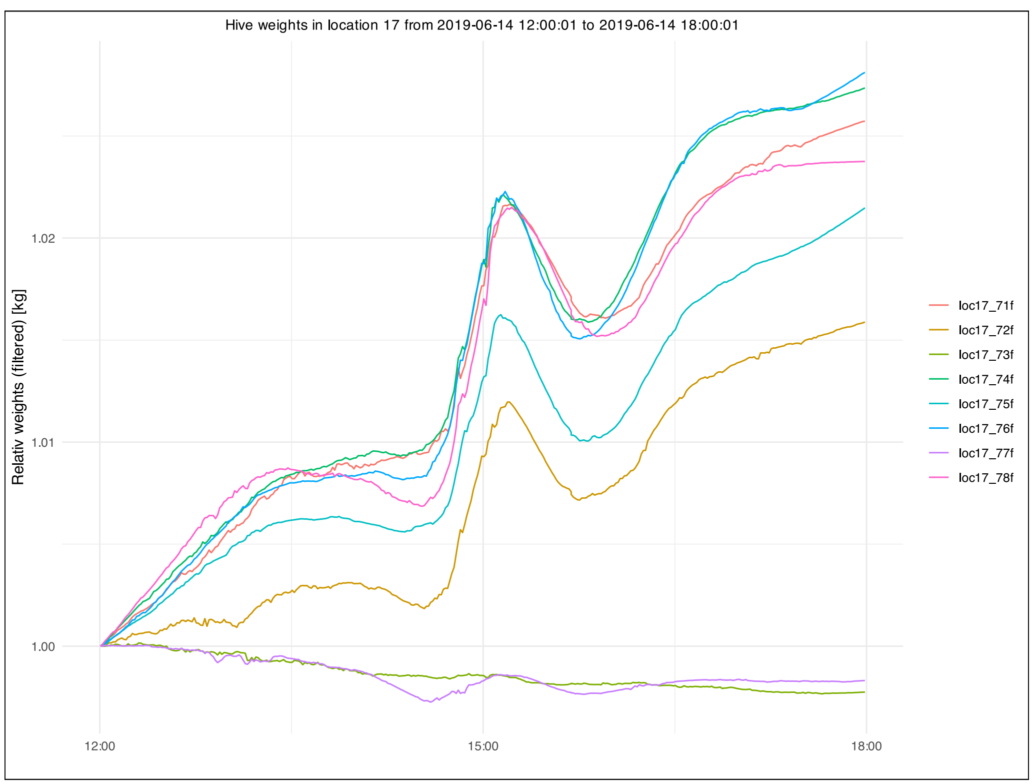

Fig. 7. Filtered (relative) weights in location 17 on June 14, 2019.

A separate instance of a biplot is presented in Fig. 10. Note that the month of August is displayed (while June is depicted in Fig. 9). Here, hive 75 exhibits a faster weight loss rate compared to the other hives. This differential rate is discernible in the biplot, which reveals a greater degree of variation for this particular hive.

These visualizations underscore the diverse trajectories of different hives and highlight the value of exploratory PCA in revealing these nuanced differences, thereby aiding in a more comprehensive understanding of hive dynamics.

3.3.2. Location 20

Fig. 11 reveals that the depicted hives can be roughly classified into two clusters characterized by similar variations. More specifically, hives 99–101 are observed to undergo a decrease in weight, whereas hives 95–98 experience a weight increase around May 5th to 6th. Positioned between these two clusters, hive 102 exhibits only a slight weight increment, as illustrated in the biplot on the right.

Fig. 12 portrays the asynchronous behavior of hive 101 in comparison with the other hives. Notably, its weight decreases while all the

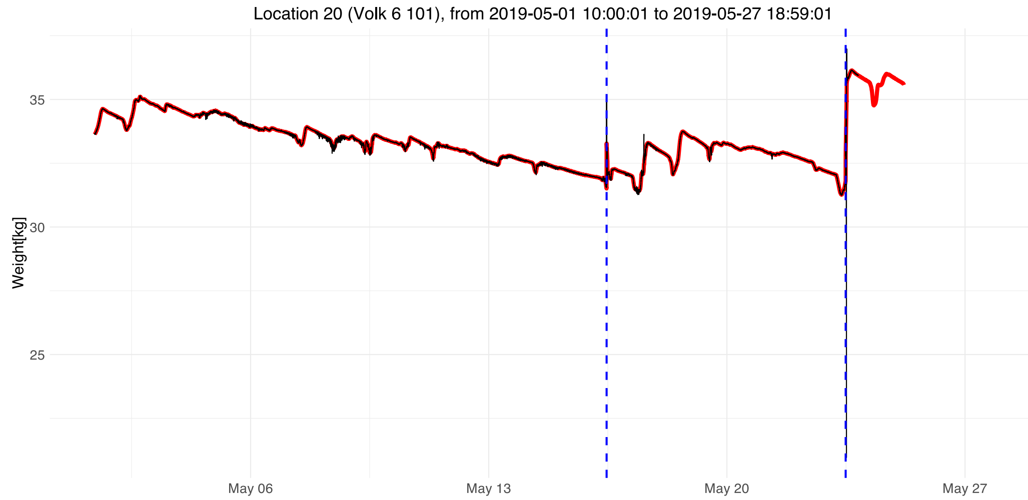

Fig. 8. Filtered (absolute) weights of hive 6 101 at location 20. Notably, there are discernible shifts in the hive weight at two specific timestamps: the first on May 16th, 2019 at 11:08:00, and the second on May 23rd, 2019 at 11:47:33. These shifts are highlighted with blue dashed lines for ease of reference.

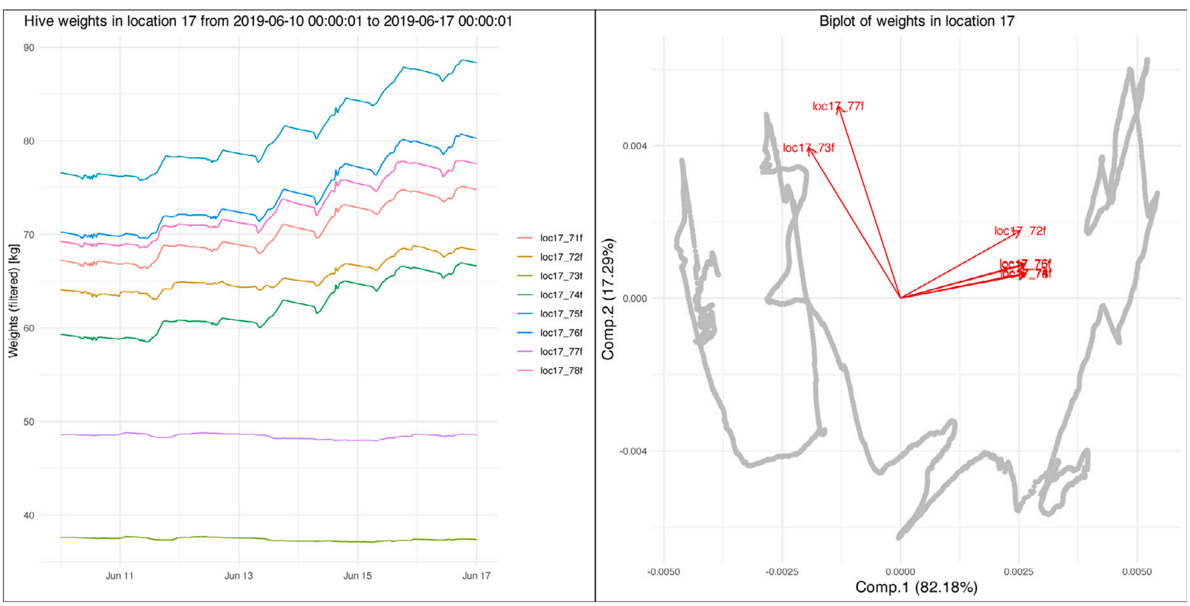

Fig. 9. Filtered (absolute) hive weights (left) and biplot (right) for hives in location 17 in June 2019.

other hives either maintain or increase their weight to varying degrees, an observation distinctly visible in the right-side biplot. Hive 99 was not included in the analysis due to its negligible activity levels.

3.4. Alarming for health indication

3.4.1. Location 17

In the weight prediction analyses of hive 75 (Fig. 10 left), we used MM-Regression and Prophet. The observed weight drop (with unknown cause) in hive 75 is detected by both methods, as shown in Fig. 13 (MM-Regression) and Fig. 14 (Prophet), respectively. Hive 75 experiences a steeper weight decrease compared to nearby hives. However, there is a notable distinction in the timing of detection. MM-Regression ($R^2 = 0.997$ on training data) demonstrates an earlier detection, which can be attributed to its narrower prediction intervals compared to Prophet. Table 1 contains the regression results on the training data. *Time* has a near-zero value since it is measured every minute compared to the covariates which are measured every hour. All variables show a negative association with hive weight.

Table 1
Robust regression results for filtered (absolute) weights on training data in Fig. 13.

Coefficient	Estimate	95% CI
Time	−0.000	[0.000, 0.000]
Temperature	−0.003	[−0.005, −0.002]
Cloud.Cover	−0.004	[−0.014, 0.005]
Dew.Point	−0.015	[−0.017, −0.013]
Visibility	−0.003	[−0.003, −0.002]
Preciptensity	−0.578	[−0.586, −0.569]
Pressure	−0.002	[−0.005, 0.001]
Wind.Speed	−0.004	[−0.004, −0.003]

3.4.2. Location 20

For hive 7_99 at location 20, both MM-Regression and Prophet were employed to anticipate weight changes, which can be indicative of unique events such as swarming. As previously demonstrated with hive 75 at location 17, these methods were employed with a brief span of training data, which were then used to generate predictions over the subsequent hours.

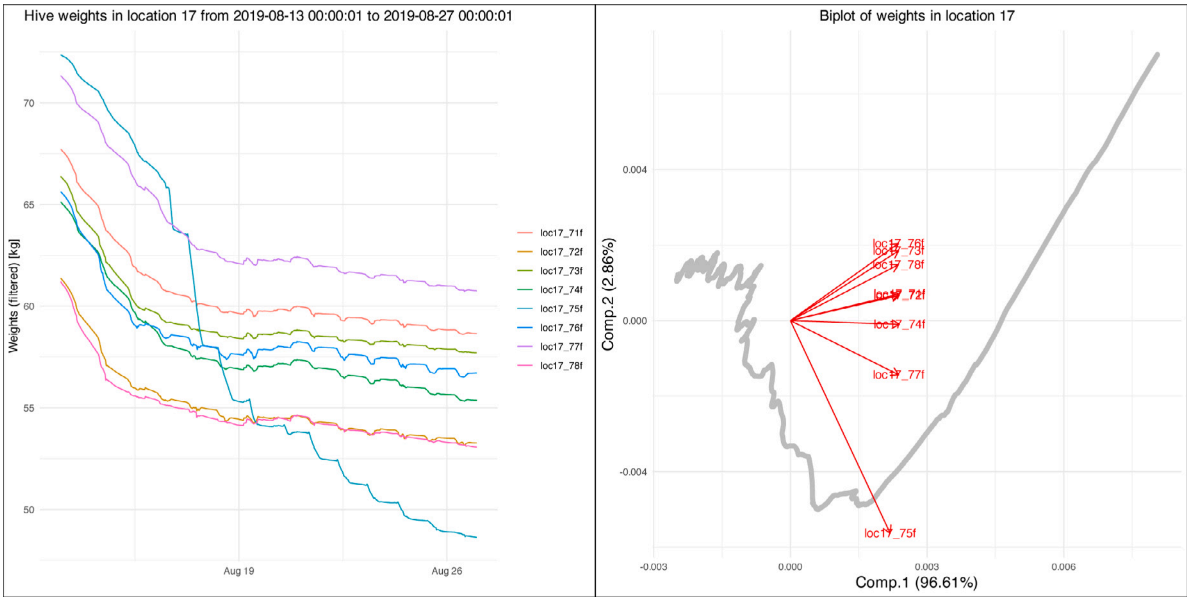

Fig. 10. Filtered (absolute) hive weights (left) and biplot (right) for hives in location 17 in August 2019.

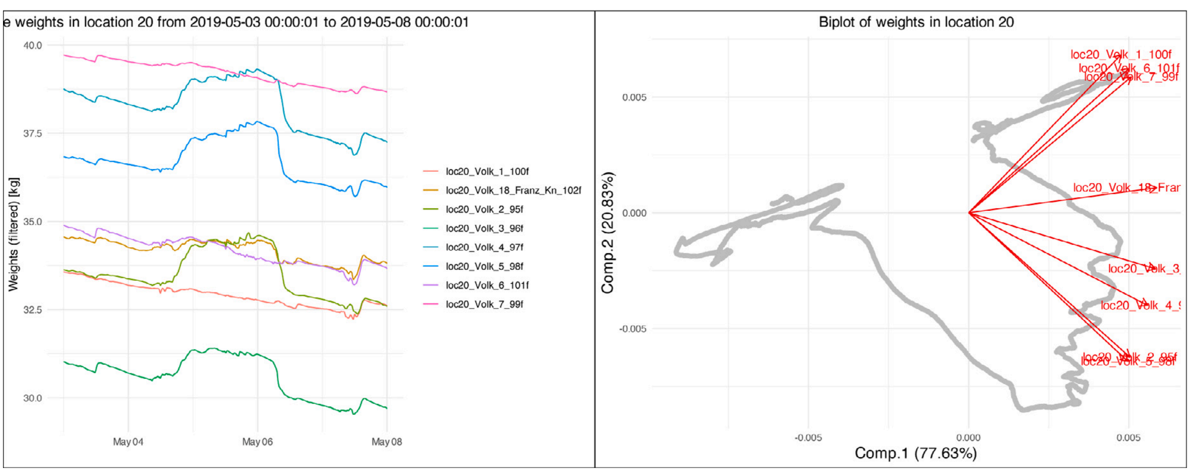

Fig. 11. Filtered (absolute) hive weights (left) and biplot (right) for hives in location 20 in May 2019.

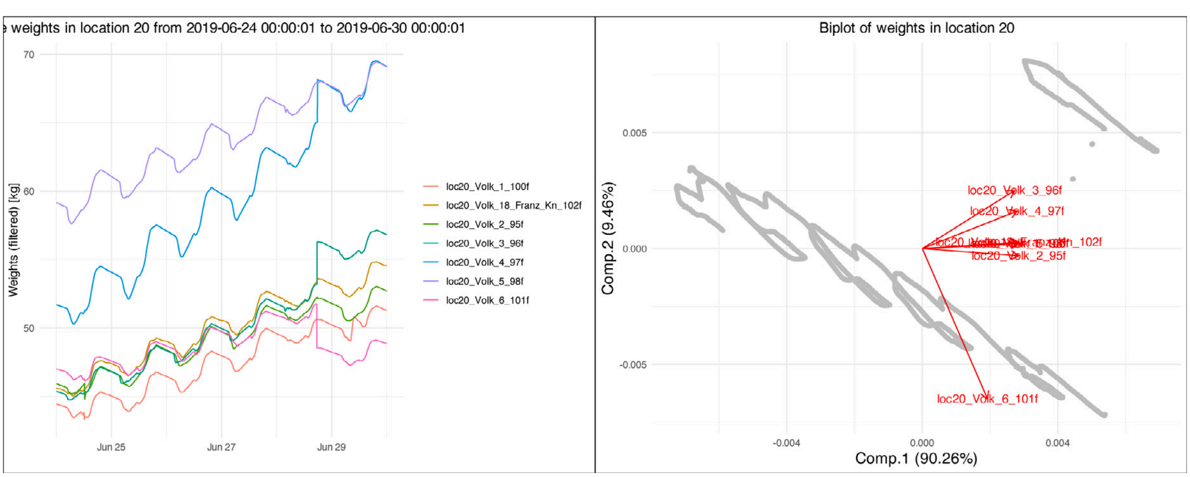

Fig. 12. Filtered (absolute) hive weights (left) and biplot (right) for hives in location 20 in June 2019. Hive 99 was excluded from the analysis since there was negligible activity.

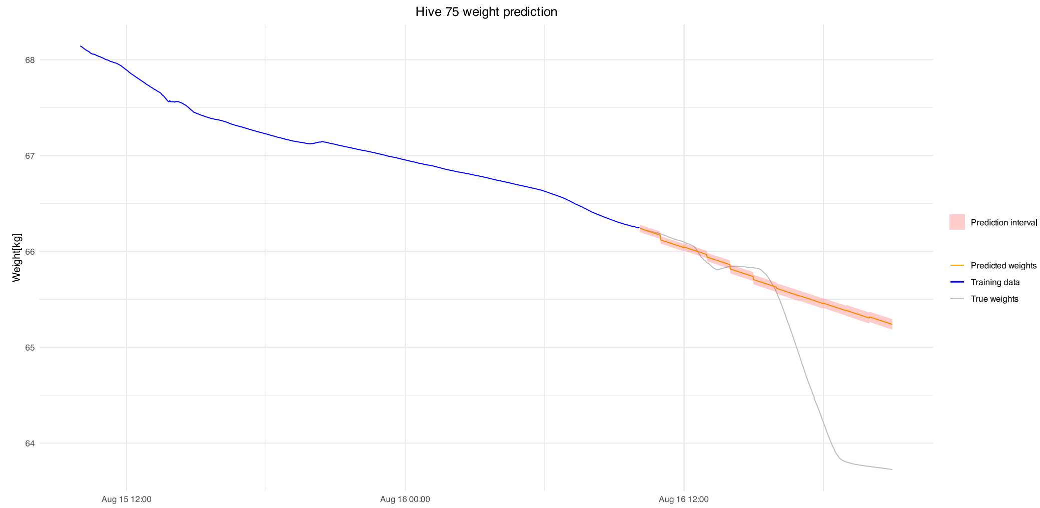

Fig. 13. Hive 75 (location 17) is loosing weight faster compared to the hives nearby. The plot shows prediction intervals using MM-Regression. Deviation of the true weights (gray) from the prediction could be used to trigger an alarm.

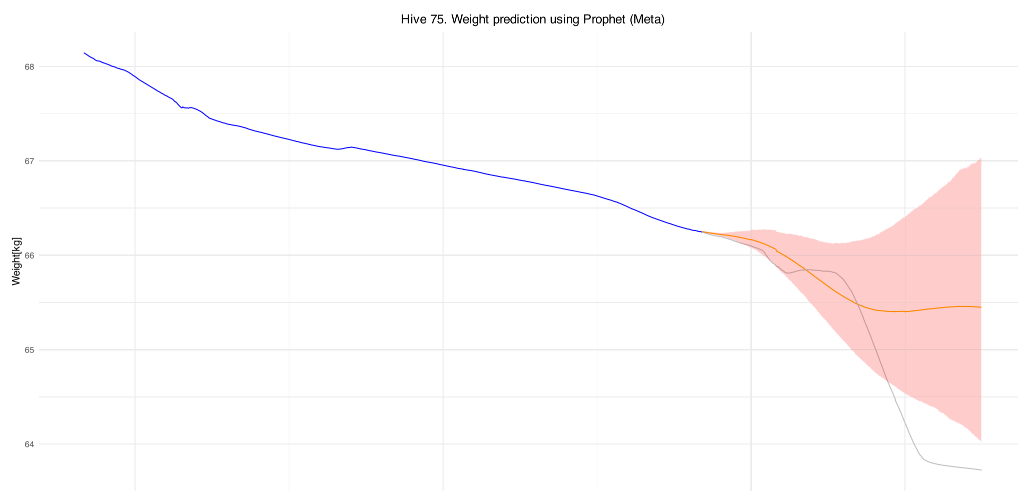

Fig. 14. Hive 75 (location 17) is loosing weight faster compared to the hives nearby. The plot shows prediction intervals using Prophet. Deviation of the true weights (gray) from the prediction could be used to trigger an alarm.

Table 2
Robust regression results for filtered (absolute) weights on training data in Fig. 15.

Coefficient	Estimate	95% CI
Time	0.000	[0.000, 0.000]
Temperature	−0.011	[−0.013, −0.009]
Dew.Point	0.011	[0.008, 0.015]
Cloud.Cover	0.587	[0.567, 0.607]
Pressure	−0.102	[−0.116, −0.088]
Wind.Speed	−0.018	[−0.02, −0.016]
Visibility	−0.061	[−0.063, −0.059]

As shown in Fig. 15, MM-Regression ($R^2 = 0.95$) provides a prediction interval that clearly deviates from the actual weight data (gray) following the swarming event. This deviation can serve as a robust alarm trigger for beekeepers, allowing them to intervene in a timely manner. Table 2 contains the regression results on the training data. *Time* has a near-zero value since it is measured every minute compared to the covariates which are measured hourly. *Temperature*, *Pressure*, *Wind.Speed* and *Visibility* show a negative association with hive weight. *Dew.Point* and *Cloud.Cover* are positively associated. Note, that *Preciptensity* was not included as a predictor here (as opposed to hive 75, Table 1, Fig. 13) because the algorithm did not converge.

In parallel, the Prophet model, illustrated in Fig. 16, also effectively captures the drastic weight drop of hive 7_99. As the model prediction diverges from the actual data, an alert could be raised to signal potential anomalies, such as a swarming event.

Both MM-Regression and Prophet serve as powerful tools for detecting significant anomalies in hive weight data, enabling timely interventions. Despite their different algorithmic bases, both have demonstrated their utility in responding to drastic events, such as swarming across various hive locations.

We cannot provide an exact quantification for false-positives and false-negatives here, as we are working with unlabeled data. It is also important to mention that an alarm does not mean that there is unlimited certainty that bees are sick or that an event such as swarming

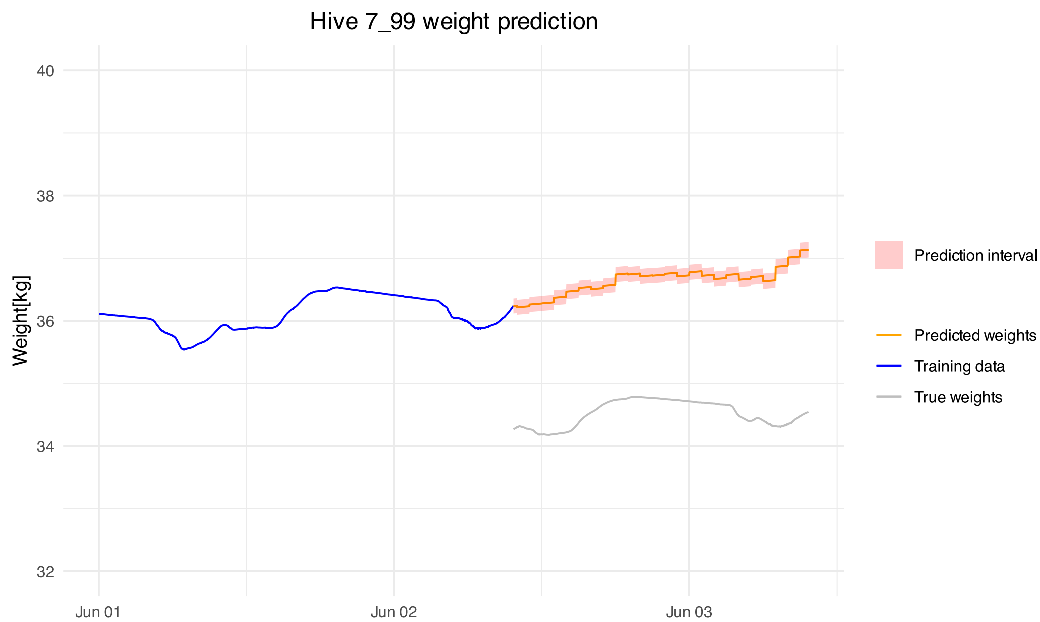

Fig. 15. Swarming event in hive 7_99 (location 20). This hive dropped in weight compared to the other hives in the same location. The plot shows prediction intervals using MM-Regression. Deviation of the true weights (gray) from the prediction could be used to trigger an alarm. Notably, the predicted weights have been adjusted upwards by approximately 0.34 kg to ensure a smoother representation on the graph. This discrepancy between weight data and covariates is attributed to the fact that while weight data is recorded every minute, covariates are updated hourly.

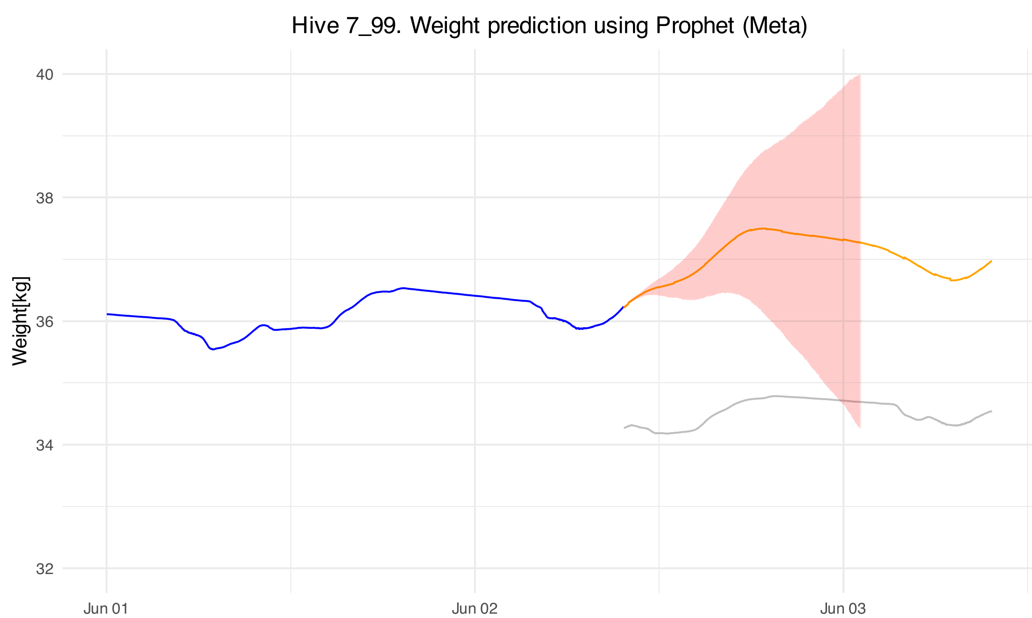

Fig. 16. Swarming event in hive 7_99 (location 20). This hive dropped in weight compared to the other hives in the same location. The plot shows prediction intervals using Prophet. Deviation of the true weights (gray) from the prediction could be used to trigger an alarm.

has occurred. An alarm should prompt the beekeeper to check the hive to ensure the well-being of the colony.

4. Discussion

4.1. General considerations

Continuous monitoring of hive weight provides information about how the hive interacts with the environment without disturbing the colony and provides insight into the state of the colony. Modeling hive weight data is the subject of extensive research. While Hadjur et al. (2022b) provides a comprehensive survey of the methods used for different kinds of beehive data, we discuss our results and methods specifically to selected literature.

Our research was centered on the analysis of complex, unlabeled data obtained from beehives, a common occurrence when evaluating hive weights. As in Meikle (2008, 2016), Meikle et al. (2018), Czekonska et al. (2023) and Holst and Meikle (2018), the methods therefore cannot be formally quantified, for example, in terms of false-negative or false-positive rates. This paper attempts to highlight the problems and possibilities of modern methods for detecting events or atypical occurrences in hive weight measurement of unlabeled data. Exemplary, the success of the individual methods used is evaluated. Therefore, we are expanding the current literature with sophisticated methods for different tasks.

One of the hurdles in the analysis of beehive weight data is creating a model that not only utilizes the information within the data set and external meteorological data, as mentioned by Robustillo et al. (2022), but also incorporates a fresh approach, namely, reducing the impact of data errors and outliers in an automated fashion. The removal of outliers is not sufficiently addressed in the literature, and so far only simple rule-based methods (Brini et al., 2023a; Arias-Calluari et al., 2023; Ziegler et al., 2022) or non-robust methods for temperature data only (Davidson et al., 2022) or multivariate robust distances with the minimum covariance determinant (Davidson et al., 2022) have been used. The latter method is suitable for the detection of multivariate outliers, but not in the context of a multiple modeling approach where leverage points are the main problem, and it is also not suitable for the detection of level shifts because it does not account for external (meteorological) variables that influence the beehive weight.

4.2. Robust methods for data cleaning, signal extraction, and level shift alarming

Various methods have been used and researched to detect and filter outliers based on thresholds and non-robust z-scores on moving windows (Brini et al., 2023b) and/or to detect level shifts in bee hive weight data. For level shift detection, this includes time series analysis to identify sudden changes (level shifts) that may indicate a significant event like a swarming, a disease outbreak, or a major change in foraging behavior (Ramsey et al., 2020; Robustillo et al., 2022). Methods from statistical process control, such as control charts, can be used to monitor bee hive weight data over time (Ziegler et al., 2022). If observed weight data go outside predetermined control limits, it may indicate a significant level shift (or outlier). However, these kinds of methods are often too simplistic since bee hive data are highly non-linear and a lot of different events can happen. Furthermore, change point analysis methods and anomaly detection algorithms can be useful in detecting changes in hive weight, suggesting changes in hive activity or health and data points that deviate significantly from the norm, potentially signaling relevant events in the hive (Davidson et al., 2022).

Also, decision trees, random forests, or neural networks can help predict significant changes based on historical data (Edwards-Murphy et al., 2016; Anwar et al., 2023; Brini et al., 2023a; Robustillo et al., 2022; Davidson et al., 2022; Torky et al., 2023). However, these kinds of methods need labeled data, and standard deep learning methods rely on non-robust loss functions and can thus be highly influenced by outliers (Templ, 2023) that occur frequently in bee hive weight measurements. For the latter issue, the proposed robust signal extraction method might serve as pre-filter.

Instead of using basic thresholds and non-robust moving windows based on z-scores as proposed in Brini et al. (2023b), robust signal extraction methods are not affected by outliers and offer a more advanced approach for data cleaning, including filtering and outlier detection (see Figs. 3, 6, 8). These methods also allow for the detection of level shifts and level shifts in the presence of outliers, as well as the extraction of the primary signal.

Robust signal extraction methods have previously been developed and used in healthcare (Davies, 2004) to trigger alarms when, for example, patients' heart rates change. In our work, these methods were tested and employed with the aim of detecting and distinguishing level shifts, regular observations, and outliers. We are not aware of any literature using such robust methods for hive weight analysis.

Despite the challenges associated with the use of unlabeled data, the results of this analysis demonstrate that the methods successfully and robustly detected level changes in the signal. Therefore, even without knowing the specific anomaly, the method efficiently recognizes the presence of an anomaly and can alert the beekeeper to investigate what is happening in the beehive. The use of robust signal extraction methods is advantageous for detecting true level shifts and defining a level shift when an outlier is present. In addition, without robust signal extraction of the main signal, some of the subsequent methods used for visualization and alarming would potentially trigger many false-positive alarms.

4.3. Alarming with similar trend monitoring and exploratory analysis using biplots

Similar trend analysis was used in Borowski et al. (2015) for robust online-surveillance of pairwise coherences in nonstationary multivariate data streams. The application of STM distances to evaluate the weight of beehives over time is an effective method to identify hives with distinct behaviors, as, for example, seen in Figs. 4 and 5, and can be used for automatic alarming based on the STM distances.

Principal component analysis and corresponding biplots offer another approach. Biplots were used in Miranda et al. (2016) to investigate different external factors that influence bee hive weights using partial least squares regression. We have used biplots in a different way – biplot obtained from principal component analysis – which allows us to examine and compare different hives at the same time. A deviating colony is easily detected. However, this method does not allow to trigger automatic alarms, thus this dimension reduction method serves as an explorative tool only. Figs. 9 to 12 demonstrate effective visualization to detect correlated hives and hives that behave differently over time. If there are additional sensors that can measure different aspects of a beehive, such as the loudness of bees or the humidity in the hive, this can be incorporated into the principal component analysis (PCA) to analyze higher-dimensional signals. By including more of these signals, the visualization becomes more powerful and the benefits of conducting a multivariate exploration with PCA increase, as it allows for the simultaneous examination of multiple signals.

4.4. Robust regression

While Arias-Calluari et al. (2023) applied a physical model to model bee hive weights, we used a statistical model, which is applied to (unlabeled) training data. Since hive data contains outliers, it is highly recommended to use robust regression methods, such as MM-Regression, with the highest efficiency and robustness. The suggested robust signal extraction methods may not be able to identify those outliers, as the definition of an outlier in a regression context (with a functional linear relationship of predictors and response) is distinct from that in a univariate signal. Thus, we improved the methods of, e.g.,Meikle (2008, 2016), Meikle et al. (2018) or Czekonska et al. (2023) by automatically down-weighting outliers when estimating the main effects and the model. In our analysis, we also considered non-linear methods such as generalized additive models, as already applied by Robustillo et al. (2022) to model hive weight data, but in general the robust MM-Regression performed better. This is not surprising because, first, the model diagnostics show that the model assumptions are satisfied in our model, i.e. there is a linear functional relationship between the hive weights and the design matrix formed. Second, there are many outliers in the data that threaten classical non-robust methods. Regression coefficients in robust regression should be interpreted with caution, as they can change depending on the specific hive and time window used (see Tables 1 and 2). To our knowledge, it has also not yet been studied for unlabeled data whether a hive develops as expected or whether events such as swarming, predation, or health problems occur in the colony, which leads to different weight processes in the hive. It is important to include covariates such as temperature, humidity, and the general weather constellation in the model, as, e.g.,Meikle (2008, 2016), Meikle et al. (2018) and Czekonska et al. (2023) also did, as bee activity is very weather dependent. Our proposed alerting method triggers an alarm when observed values fall outside the model-based prediction intervals (see Figs. 13, 14, 15, 16). Therefore, it can also be seen as a sophisticated control chart tool that takes into account a complex statistical model, covariates, and outliers. Alternatively, the Prophet algorithm can also be used as a time series model with and without covariates.

4.5. Limitations

Our study also has some limitations. First, our primary reliance on unlabeled data presents a significant constraint. Although we devised methods to detect noticeable shifts and events, the lack of labeled data points made it challenging to validate and formally quantify the effectiveness of these methods in terms of, e.g., false-negative or false-positive rates. Second, while our statistical models incorporated various covariates, such as temperature, humidity, and overall weather patterns, other unconsidered environmental or behavioral factors could influence the hive weights. Statistically speaking, the available data sets were convenience data sets and may not be representative of a broader range of hive locations. Nevertheless, we did not generalize to a larger population and the weight predictions were made model-based considering the meteorological covariates specific to the individual hive.

In general, the challenge in building an alarm system is to find the right balance in the decision process between unwanted hive behavior and normal data, an anomalous and a valid data point. Beekeepers should not send signals too often due to desensitization. Too many false-positives would result in ignoring all or most of the alarms. Too few alarms, on the other hand, would render the alarms useless and result in an unnecessary loss of bees. Ideally, the beekeeper has installed a video camera near the hives, providing a visual aid to verify an alarm. Regarding signal extraction and outlier detection, we used the recommended estimation method (repeated median regression, RM) according to Borowski (Borowski, 2013) to estimate the true underlying signal online and offline (as we have done in our research so far). Parameters used for signal extraction, such as window width, shift detection factor (default is a 2 σ rule), or lower bound for scale estimation, should ideally be estimated using labeled data sets to obtain ‘optimal’ performance. The same is true for feature selection in MM-Regression and Prophet (e.g., variables to use for the prediction or window width for training). These would be determined using a training and validation set and later tested on unseen data. Parameters for similar trend monitoring and principal component analysis (among other window width) could also be improved by labeled data, for instance by using cross validation.

Regarding the benefit for beekeepers, we clarify that our research is not a ready-to-use tool. Instead, it serves as an enhancement to the analysis toolbox for hive weight prediction, intended for inclusion in future hive monitoring systems. Our goal was to make a contribution to more robust and precise hive monitoring. To the best of our knowledge, the methods we used to model beehive weights were not used yet (specifically signal extraction, similar trend monitoring, PCA and MM-Regression) and expand the toolbox for hive weight analysis and precision beekeeping.

4.6. Future work

Going forward, beekeepers may undertake the significant task of accurately labeling their data sets. This challenge requires diligent observation, and the complexity is amplified, as interventions, such as inspecting the hive, can inherently influence the bee colony. Consequently, generating well-labeled training data that remain unaffected by beekeeper interventions seems an almost impossible feat. However, once a sufficient volume of labeled data becomes available, we can begin numerical comparisons of methods, expressing outcomes through metrics such as false-negative or false-positive rates. To achieve this, a minimum of around 30 (even much more when using deep learning methods) correctly labeled training data points would be needed for each type of event. One would expect that the performance of such a future hive monitoring system is vastly improved by labeled data. Although demanding, creating labeled data sets seems attainable with enough dedicated effort using all available surveillance (e.g. sufficiently accurate cameras and sound).

Future models for hive weight could consider the many additional factors influencing hive weight simultaneously. Considerations could extend to include nectar, pollen, and propolis flow; internal hive temperature and humidity; noise levels within the hive; brood temperature; the development status of the brood; the number of foragers; the health and status of the queen; as well as the stored amounts of nectar and pollen. By holistically integrating these variables, predictions could be substantially improved and more accurate.

Moreover, the establishment of collaborative data repositories where beekeepers worldwide can upload and share their hive weight data, observations, and notes might be a solution to data scarcity and diversity issues. Such a platform would not only pool a more diverse range of data but also foster global collaboration between researchers and beekeepers, potentially speeding up the data collection and validation process.

4.7. Conclusion

Given the nature of our data, which is unlabeled, we cannot definitively assert that the deviations and level shifts we have identified are indicative of undesirable hive weight behavior. However, robust signal extraction proves to be a valuable tool for estimating the underlying signal amidst noise and significant outliers, applicable in both on-line and off-line scenarios. Although we have only assessed it qualitatively, the level shift detection also performs satisfactorily. MM-Regression and Prophet showed promising results for indicating deviating behavior, including level shifts. We expect further improvement after model calibration and validation in labeled data sets.

We have successfully applied advanced statistical techniques to analyze noisy hive weight data, an approach that, to our knowledge, has not been previously utilized in this field. Although the methods applied are promising, future research is needed to fine-tune and validate the results using labeled data sets.

Abbreviations

The following abbreviations are used in this manuscript:

mSCARM	Multivariate slope comparing adaptive repeated median
PCA	Principal component analysis
SSM	Similar slope monitoring
STM	Similar trend monitoring

Funding

This research received no external funding.

CRediT authorship contribution statement

Jürgen Degenfellner: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Writing – original draft, Writing – review & editing, Validation, Visualization. **Matthias Templ:** Conceptualization, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank Silvio Ziegler and the team from HiveWatch for providing gus with data sets for the analysis.

Appendix

See Figs. 17 and 18 and Tables 3–5.

Table 3
Locations and their geographical coordinates.

Location	Latitude	Longitude
12	47.1	7.5
15	47.3	8.9
17	46.7	9.6
20	47.1	9.5
21	47.3	8.8
24	47.1	7.5
56	47.0	9.5

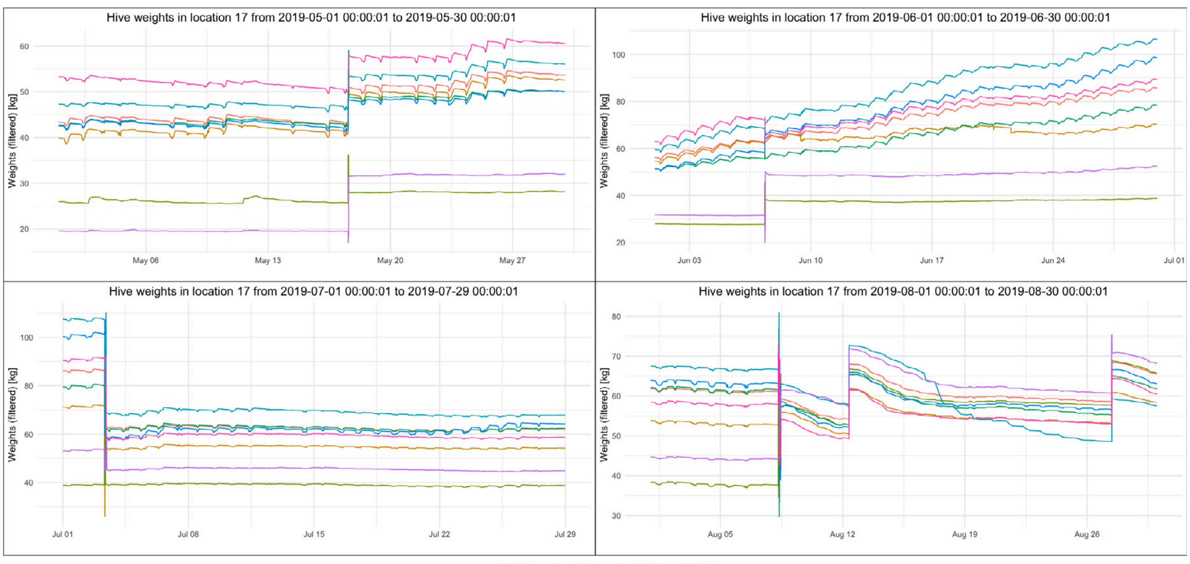

Fig. 17. Overview of filtered (absolute) hive weights in location 17 most active period (May to August).

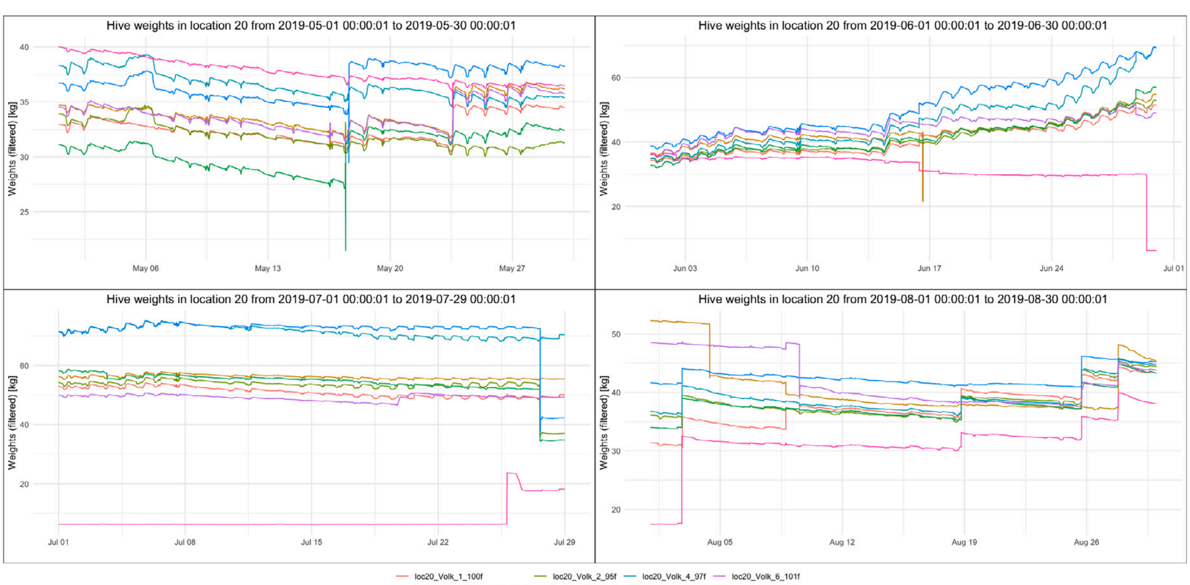

Fig. 18. Overview of filtered (absolute) hive weights in location 20 most active period (May to August).

Table 4
Variables in the HiveWatch data sets (2018–2019).

Variable	Description
Time	Varying time distances, mostly minute-data
Weight	Hive weight in kg
Temperature	Outside temperature in degrees Celsius
Apparent.Temperature	Perceived temperature considering air temperature, humidity, and wind speed
Cloud.Cover	Number between 0 and 1 indicating cloud coverage
Humidity	Relative humidity
Preciptensity	Measures the intensity of rainfall
PrepProbability	Probability of precipitation
Pressure	Air pressure in mbar
Visibility	Distance at which an object or light can be clearly discerned in km
Dew.Point	Temperature to which air must be cooled to become saturated with water vapor
Wind.Speed	Wind speed in km/h
Wind.Bearing	Direction from which the wind is coming, measured in degrees from true north in a clockwise direction
Latitude	Latitude of the hive location
Longitude	Longitude of the hive location

Table 5
Hives per location: HiveWatch (2018–2019)

Location	Hives
12	Volk_15_43, Volk_16_55, Volk_17_62, Volk_18_56, Volk_19_549
15	NewScale_46, NewScale_47, NewScale_48
17	71, 72, 73, 74, 75, 76, 77, 78
20	Volk_1_100, Volk_18_Franz_Kn_102, Volk_2_95, Volk_3_96, Volk_4_97, Volk_5_98, Volk_6_101, Volk_7_99
21	103, Volk_2_104, Volk_3_105, Volk_4_106, Volk_5_107, Volk_6_108, Volk_7_109, Volk_8_110
24	Volk_1_131, Volk_2_132, Volk_3_133, Volk_4_134, Volk_5_135
56	567_454, 568_455, Volk_561_448, Volk_562_449, Volk_563_450, Volk_564_451, Volk_565_452, Volk_566_453

References

Alleri, M., Amoroso, S., Catania, P., Verde, G.L., Orlando, S., Ragusa, E., Sinacori, M., Vallone, M., Vella, A., 2023. Recent developments on precision beekeeping: A systematic literature review. J. Agric. Food Res. 100726.

Anon, 2022. Economics of bee pollination. https://beespotter.org/topics/economics/. (Accessed 23 June 2023).

Anon, 2023. Fakultät statistik - r pakete und software. https://www.statistik.tu-dortmund.de/1542.html. (Accessed 3 July 2023).

Anwar, O., Keating, A., Cardell-Oliver, R., Datta, A., Putrino, G., 2023. Apis-prime: A deep learning model to optimize beehive monitoring system for the task of daily weight estimation. Appl. Soft Comput. 144, 110546. http://dx.doi.org/10.1016/j.asoc.2023.110546.

Arias-Calluari, K., Colin, T., Latty, T., Myerscough, M., Altmann, E.G., 2023. Modelling daily weight variation in honey bee hives. PLoS Comput. Biol. 19 (3), 1–22. http://dx.doi.org/10.1371/journal.pcbi.1010880.

Arnia, 2023. Arnia - remote hive monitoring. https://www.arnia.co/. (Accessed 22 June 2023).

Bauernzeitung, 2019. Rekordtiefe honigernte in der schweiz. https://www.bauernzeitung.ch/artikel/tiere/rekordtiefe-honigernte-in-der-schweiz-361071. (Accessed 7 July 2023).

Bencsik, M., Bencsik, J., Baxter, M., Lucian, A., Romieu, J., Millet, M., 2011. Identification of the honey bee swarming process by analysing the time course of hive vibrations. Comput. Electron. Agric. 76 (1), 44–50. http://dx.doi.org/10.1016/j.compag.2011.01.004.

Borowski, M., 2013. Echtzeit-Extraktion relevanter Information aus multivariaten Zeitreihen basierend auf robuster Regression (Ph.D. thesis). Technische Universität Dortmund, http://dx.doi.org/10.17877/DE290R-14393.

Borowski, M., Busse, D., Fried, R., 2015. Robust online-surveillance of trend-coherence in multivariate data streams: the similar trend monitoring (stm) procedure. Stat. Comput. 25, 913–928. http://dx.doi.org/10.1007/s11222-014-9462-4.

Brini, A., Giovannini, E., Smaniotto, E., 2023a. A machine learning approach to forecasting honey production with tree-based methods. arXiv:2304.01215.

Brini, A., Giovannini, E.C., Smaniotto, E., 2023b. A machine learning approach to forecasting honey production with tree-based methods. ArXiv abs/2304.01215.

Cecchi, S., Spinsante, S., Terenzi, A., Orcioni, 2021. A smart sensor-based measurement system for advanced bee hive monitoring. Sensors, MDPI 20, 2726. http://dx.doi.org/10.3390/s20092726.

Chen, C., Yang, E., Jiang, J., Lin, T., 2012. An imaging system for monitoring the in-and-out activity of honey bees. Comput. Electron. Agric. 89, 100–109. http://dx.doi.org/10.1016/j.compag.2012.08.006.

Cheng, J., Karambelkar, B., Xie, Y., 2023. leaflet: create interactive web maps with the javascript 'leaflet' library. r package version 2.1.2. https://CRAN.R-project.org/.

Czekonska, K., Lopuch, S., Miscicki, S., Bankowski, J., Szabla, K., 2023. Monitoring of hive weight changes in various landscapes. Apidologie 54 (3), 30. http://dx.doi.org/10.1007/s13592-023-01012-0.

Davidson, P., Steininger, M., Lautenschlager, F., Krause, A., Hotho, A., 2022. Anomaly detection in beehives: an algorithm comparison. In: Ahrens, A., Prasad, R.V., Benavente-Peces, C., Ansari, N. (Eds.), Sensor Networks. Springer International Publishing, Cham, pp. 1–20.

Davies, P., 2004. Robust signal extraction for on-line monitoring data. J. Statist. Plann. Inference 122, 65–78. http://dx.doi.org/10.1016/j.jspi.2003.06.012.

Edwards-Murphy, F., Magno, M., Whelan, P.M., O'Halloran, J., Popovici, E.M., 2016. b+WSN: Smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring. Comput. Electron. Agric. 124, 211–219. http://dx.doi.org/10.1016/j.compag.2016.04.008.

Fried, R., 2004. Robust filtering of time series with trends. J. Nonparametr. Stat. 16, http://dx.doi.org/10.1080/10485250410001656444.

Fried, R., 2006. Repeated median and hybrid filters. Comput. Statist. Data Anal. 50, 2313–2338. http://dx.doi.org/10.1016/j.csda.2004.12.013.

Fried, R., 2007. On the robust detection of edges in time series filtering. Comput. Statist. Data Anal. 52, 1063–1074. http://dx.doi.org/10.1080/10485250410001656444.

Fried, R., 2011. Robust repeated median regression in moving windows with data-adaptive width selection. SFB 823, Discussion Papers, Technische Universität Dortmund.

Fried, R., Schettlinger, K., Borowski, M., 2022. Robfilter: robust time series filters. r package version 4.1.3. https://CRAN.R-project.org/package=robfilter.

Gounari, S., Proutsos, N., Goras, G., 2022. How does weather impact on beehive productivity in a mediterranean island? Italian J. Agrometeorol. (1), 65–81. http://dx.doi.org/10.36253/ijam-1195.

Hadjur, H., Ammar, D., Lefèvre, L., 2022a. Toward an intelligent and efficient beehive: A survey of precision beekeeping systems and services. Comput. Electron. Agric. 192, 106604. http://dx.doi.org/10.1016/j.compag.2021.106604, https://www.sciencedirect.com/science/article/pii/S0168169921006219.

Hadjur, H., Ammar, D., Lefèvre, L., 2022b. Toward an intelligent and efficient beehive: A survey of precision beekeeping systems and services. Comput. Electron. Agric. 192, 106604. http://dx.doi.org/10.1016/j.compag.2021.106604.

Hijmans, R.J., 2022. geosphere: spherical trigonometry. r package version 1.5-18, https://CRAN.R-project.org/package=geosphere.

HiveWatch, 2023. Hivewatch - smart beehive monitoring; hivewatch.ch. https://hivewatch.ch/en. (Accessed 22 June 2023).

Holst, N., Meikle, W.G., 2018. Breakfast canyon discovered in honeybee hive weight curves. Insects, MDPI 9.

Human, H., Nicolson, S., Dietemann, V., 2006. Do honeybees, apis mellifera scutellata, regulate humidity in their nest? Naturwissenschaften 1 (93), 397–401.

Iannone, R., 2023. Diagrammer: graph/network visualization. r package version 1.0.10. https://CRAN.R-project.org/package=Diagrammer.

Kulyukin, V., Tkachenko, A., Price, K., Meikle, W., Weiss, M., 2022. Integration of scales and cameras in nondisruptive electronic beehive monitoring: On the within-day relationship of hive weight and traffic in honeybee (Apis Mellifera) colonies in langstroth hives in Tucson, Arizona, USA. Sensors 22 (13), http://dx.doi.org/10.3390/s22134824, https://www.mdpi.com/1424-8220/22/13/4824.

Maronna, R., Martin, D., Yohai, V., 2006. Robust Statistics: Theory and Methods. Wiley.

Meikle, W., 2008. Within-day variation in continuous hive weight data as a measure of honey bee colony activity. Apidologie 39, 694–707. http://dx.doi.org/10.1051/apido:2008055.

Meikle, W., 2016. Monitoring colony phenology using within-day variability in continuous weight and temperature of honey bee hives. Apidologie 47, 1–14. http://dx.doi.org/10.1007/s13592-015-0370-1.

Meikle, W., Holst, N., Colin, T., Weiss, M., Carroll, M., McFrederick, Q.e.a., 2018. Using within-day hive weight changes to measure environmental effects on honey bee colonies. PLoS One http://dx.doi.org/10.1371/journal.pone.0197589.

Meikle, W., Holst, N., Mercadier, G., Derouané, F., James, R., 2006. Using balances linked to dataloggers to monitor honeybee colonies. J. Apicult. Res. 45 (1), 39–41. http://dx.doi.org/10.1080/00218839.2006.11101311.

Miranda, M.A., Bicout, D.J., Bøtner, A., Butterworth, A., Calistri, P., Depner, K., Edwards, S., Garin-Bastuji, B., Good, M., Gortázar-Schmidt, C., Michel, V., More, S.J., Nielsen, S.S., Raj, M., Sihvonen, L.H., Spoolder, H.A.M., Stegeman, J.A., Thulke, H.-H., Velarde, A., Willeberg, P.W., Winckler, C., 2016. Assessing the health status of managed honeybee colonies (healthy-b): A toolbox to facilitate harmonised data collection. EFSA J. 14, 4578.

Monitoring, B., 2023. Beehive monitoring - remote beehive monitoring. https://www.beehivemonitoring.com/de/. (Accessed 22 June 2023).

Ntawuzumunsi, E., Kumaran, S., Sibomana, L., 2021. Self-powered smart beehive monitoring and control system (sbmacs). Sensors, MDPI 21 (10), 3522. http://dx.doi.org/10.3390/s21103522.

OSBeehives, 2023. Osbeehives - buzzbox hive health monitor & beekeeping app. https://www.osbeehives.com/. (Accessed 22 June 2023).

Pejić, J., Milovanović, M., Božilov, A., Pejić, P., 2022. Impact of the Precision Beekeeping on the Living Environment. In: Facta Universitatis, Series: Working and Living Environmental Protection, pp. 049–061.

Ramsey, M.-T., Bencsik, M., Newton, M.I., Reyes, M., Pioz, M., Crauser, D., Delso, N.S., Le Conte, Y., 2020. The prediction of swarming in honeybee colonies using vibrational spectra. Sci. Rep. 10 (1), 9798. http://dx.doi.org/10.1038/s41598-020-66115-5.

Robustillo, M.C., Pérez, C.J., Parra, M.I., 2022. Predicting internal conditions of beehives using precision beekeeping. Biosyst. Eng. 221, 19–29. http://dx.doi.org/10.1016/j.biosystemseng.2022.06.006.

Rousseeuw, P.J., Croux, C., 1993. Alternatives to the median absolute deviation. J. Amer. Statist. Assoc. 88 (424), 1273–1283. http://dx.doi.org/10.1080/01621459.1993.10476408.

Rousseeuw, P.J., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Mächler, M., 2009. *robustbase*: basic robust statistics. R package version 0.4-5. http://CRAN.R-project.org/package=robustbase.

Rudis, B., 2017. darksky: tools to work with the 'dark sky' 'api'. r package version 1.3.0. https://CRAN.R-project.org/.

Salibian-Barrera, M., Yohai, V.J., 2006. A fast algorithm for S-regression estimates. 15, pp. 414–427.

Siegel, A., 1982. Robust regression using repeated medians. Biometrika 68, 242–244. http://dx.doi.org/10.1093/biomet/69.1.242.

SolutionBee, 2023. Solutionbee - hive scale and beekeeping solutions. https://solutionbee.com/. (Accessed 22 June 2023).

Tautz, J., Heilmann, H.R., 2007. Phänomen Honigbiene, 1st Edition Spektrum Akademischer Verlag, Heidelberg.

Taylor, S.J., Letham, B., 2018. Forecasting at scale. Amer. Statist. 72 (1), 37–45. http://dx.doi.org/10.1080/00031305.2017.1380080.

Taylor, S., Letham, B., 2021. prophet: automatic forecasting procedure. r package version 1.0. https://CRAN.R-project.org/package=prophet.

Templ, M., 2023. Enhancing precision in large-scale data analysis: An innovative robust imputation algorithm for managing outliers and missing values. Mathematics 11 (12), 2729. http://dx.doi.org/10.3390/math11122729.

Terenzi, A., Cecchi, S., Spinsante, S., Orcioni, S., Piazza, F., 2019. Real-time system implementation for bee hives weight measurement. In: 2019 IEEE International Workshop on Metrology for Agriculture and Forestry. MetroAgriFor, pp. 231–236. http://dx.doi.org/10.1109/MetroAgriFor.2019.8909252.

Torky, M., Nasr, A.A., Hassanien, A.E., 2023. Recognizing beehives'health abnormalities based on mobile net deep learning model. Int. J. Comput. Intell. Syst. 16 (1), 135. http://dx.doi.org/10.1007/s44196-023-00311-9.

Ziegler, C., Ueda, R.M., Sinigaglia, T., Kreimeier, F., Souza, A.M., 2022. Correlation of climatic factors with the weight of an apis mellifera beehive. Sustainability, MDPI 14 (9), 5302. http://dx.doi.org/10.3390/su14095302.