UNIVERSITÀ DEGLI STUDI FIRENZE

DISEI
DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

Working Papers
Quantitative Methods for Social Sciences

Beyond Average Hive Performance: Tail Risk Measurement in Italian Apiculture with Honey-at-Risk

Alessio Brini, Giacomo Toscano, Ginevra Virginia Lombardi, Maria Elvira Mancino

Working Paper N. 01/2026

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

Beyond Average Hive Performance: Tail Risk Measurement in Italian Apiculture with Honey-at-Risk

Alessio Brini[a] Giacomo Toscano[b] Ginevra Virginia Lombardi[b]
Maria Elvira Mancino[b]

January 12, 2026

[a] *Duke University Pratt School of Engineering, 305 Teer Engineering Building Box 90271 Durham, NC 27708 (USA). E-mail: alessio.brini@duke.edu*

[b] *DISEI, University of Florence, Via delle Pandette 32 50127 Firenze (Italy). E-mail: giacomo.toscano@unifi.it, ginevravirginia.lombardi@unifi.it, mariaelvira.mancino@unifi.it*

Abstract

This paper provides a framework for measuring honey-production risk that complements standard mean-based analyses by explicitly targeting downside tail risk. Using hive-weight data from Italian hives over the period 2021–2024, downside tail risk is quantified through the Honey-at-Risk (HaR) metric, defined as the quantile of aggregate daily production shocks for groups of hives exposed to homogeneous climatic conditions. The aggregate shock distribution is modeled based on a flexible meta-distributional framework that combines hive-specific marginal distributions with a copula for the cluster dependence structure, allowing for improved accuracy in representing heterogeneous marginal behavior and cross-hive dependence. The implementation of HaR documents substantial spatial heterogeneity, with mountain clusters facing not only higher downside tail risk but also higher variability in the magnitude of the risk measure, compared to plains and hills. Moreover, logistic regressions suggest that extreme losses, that is, losses exceeding HaR, are associated with climatic conditions (specifically, with atmospheric temperature), through two distinct effects: average daily temperature is negatively related to the probability of extreme losses, while extreme realizations of daily maximum temperature tend to increase this probability. Taken together, these results position HaR as an operational complement to elasticity-based assessments, enabling the identification of high-risk areas and supporting targeted adaptation measures as well as the design of weather-indexed insurance and compensation schemes.

Keywords: honey production risk, temperature risk, climate change, copulae, extreme events

1 Introduction

Honeybees, specifically the *Apis mellifera* species, are pivotal for human society, as they significantly contribute to crop production through pollination. Indeed, honeybees affect 35% of the global food supply, including the pollination of 70% of the world's most crucial crop species (Tscharntke et al., 2012), and thus have a profound impact on the global economy (Delaplane et al., 2000; Klein et al., 2007; Garibaldi et al., 2014). Moreover, honeybees are notable for the production of honey, a commodity which represents a multi-billion-euro global market: in 2022, the European Union, housing 20 million beehives and producing 218,000 tons of honey, stood as the second-largest producer globally, trailing China[1].

Climatic variability is a primary determinant of honey production. While other key factors, such as pests and diseases, can be mitigated through effective beekeeping practices, the influence of climate, which affects flowering phenology and consequently nectar and pollen availability, remains largely exogenous (van Engelsdorp and Meixner, 2010; Van Espen et al., 2023; Potts et al., 2010; Calovi et al., 2021). Indeed, climate change is anticipated to alter the current landscape of honey production (Holmes, 2002; Gordo and Sanz, 2006; Le Conte and Navajas, 2008; Switanek et al., 2017; Flores et al., 2019; Solovev, 2020), with notable impacts on countries like Italy, where beekeepers have already observed drastic fluctuations in honey yields, experiencing losses of up to 70% (Porrini et al., 2016; Gray et al., 2019).

To address these challenges, the insurance industry has introduced products that cover climate-related risks, but their diffusion remains limited by the difficulty of constructing reliable pricing models (Colivicchi et al., 2025). In this context, the availability of quantitative indicators of the risk of extreme honey-production losses, as well as reliable assessments of the beekeepers' exposure to this risk, becomes crucial not only for designing effective insurance products but also for informing policy interventions and public support measures aimed at mitigating the severe consequences of climate change on honey production.

The contribution of this paper is to move in this direction by introducing and carefully implementing a simple quantitative indicator, termed Honey-at-risk (HaR), that gauges the risk of extreme losses in honey production for clusters of hives exposed to homogeneous climatic conditions. Similarly to the Value-at-Risk metric employed in financial econometrics (see Ch. 2 of McNeil et al. (2015)), we define HaR as the α-quantile of the distribution of a cluster's aggregate honey production shock, measured in kilograms. In other words, HaR denotes the cluster-wise aggregate honey production shock that will not be exceeded with probability α and thus summarizes downside tail risk at the cluster level.

Several empirical studies have investigated the relationship between climatic conditions and honey production by focusing on the modeling of average yields, typically at seasonal or annual frequency. These include regression-based analyses linking honey yield to temperature and flowering conditions (Langowska et al., 2017), spatial regressions of honey productivity at the regional level (Tassinari et al., 2013), as well as more recent applications combining multiple meteorological variables to explain or predict average production levels (Sahin Demirel, 2025; Masaka, 2023). However, these methodological approaches do not provide information about the

[1] `https://agriculture.ec.europa.eu/farming/animal-products/honey_en`.

likelihood and magnitude of extreme losses, a piece of information which is particularly relevant in the presence of heavy-tailed production shocks such as those induced by climate change. By contrast, the HaR framework explicitly targets the lower tail of a cluster's production-shock distribution, thereby complementing standard analyses focused on average losses.

We implemented the HaR metric at the daily level using a unique dataset that allows extracting hive weight time series covering the period 2021–2024 for several locations within the Italian territory. In doing so, we followed a numerical procedure based on the Monte Carlo simulation of the dependence structure and the marginal distributions of the production shocks of clusters, which yields flexibility and improved accuracy in the multivariate modeling of tail events (see Ch. 7 of McNeil et al. (2015)). The results we obtained reveal a pronounced heterogeneity in HaR values across altitude categories, with mountain areas exhibiting not only a higher risk of extreme losses, compared to plains and hills, but also higher variability in HaR values. Moreover, in a logistic regression framework, we investigated the relationship between the likelihood of extreme daily production losses - defined as losses exceeding the estimated HaR metric at a given confidence level - and atmospheric temperature. Temperature represents a key driver of honey production, as it directly affects bees' foraging activity, colony thermal regulation, and nectar availability Burrill and Dietz (1981), Stabentheiner et al. (2010), Petanidou and Smets (1996), Vincze et al. (2025). Specifically, our analysis included not only the mean daily temperature but also extreme temperature realizations, modelled as a binary variable equal to one in the presence of a peak in the daily maximum temperatures exceeding a given threshold and to zero otherwise.

As for the role of the mean temperature, the results we obtained suggest that a statistically-significant negative link exists between the latter and the likelihood of extreme losses. This phenomenon can be explained by the sensitivity of honeybee colonies to low temperatures, which can impair colony development and reduce foraging activity, by slowing brood rearing, reducing worker emergence, and increasing physiological stress during the critical transition from winter to the foraging season, see, e.g., Vincze et al. (2025); Gordo and Sanz (2006). Consequently, our results suggest that periods characterized by lower average temperatures may increase the probability of severe production losses.

With regard to maxima peaks, we found that the latter were associated with an increase in the likelihood of extreme losses in all years analyzed. This is consistent with the literature showing that excessive heat, especially during flowering periods, reduces nectar secretion (Takkis et al. (2015)) and shortens bloom duration (Descamps et al. (2021)).

However, in our study, the link between extreme-loss probability and temperature maxima peaks was statistically significant in 2021, 2023, and 2024, but not in 2022. This might be justified by noting that cluster-level daily maximum temperatures in 2022 never exceeded 39.1°C in our sample, while critical thermal thresholds for domestic honeybee activity are estimated to be above 40°C (Vincze et al., 2025). By contrast, in the other years of the sample, temperature maxima surpassed 40°C, making peak realizations more likely to be associated with impactful heat stress.

The remainder of the paper is organized as follows. Section 2 details the dataset properties, the procedure followed to extract daily time series and the analysis of the properties of such series. Section 3 includes the illustration of the parametric model we adopted as the data-generating

process, the definition of the HaR metric and the description of the estimation procedure thereof. Section 4 illustrates empirical results related to HaR implementation and investigates the link between HaR exceedances and atmospheric temperature. Section 5 discusses empirical results and their policy implications. Finally, Section 6 concludes.

2 Data

In this section, we describe the high-frequency dataset of intraday hive weight recordings employed in our study, outline the procedure used to construct the daily hive weight time series that serve as the inputs to our analysis, and summarize their main statistical properties.

2.1 Dataset description

The dataset employed in our study was provided by the Italian technology company 3Bee S.R.L.[2]. It comprises hive weight observations collected from sensors installed on beehives, recorded approximately every three hours. The dataset covers hundreds of hives located across Italy over the period from 2021 to 2024. Each year, the monitoring network expands as additional hives are equipped with sensors, hence both the depth of available data and the geographical coverage across the Italian territory increase. However, the composition of the monitored sample is not strictly incremental: some hives are newly added, while others are removed from observation. Nonetheless, the dataset captures a wide range of environmental conditions, as the recorded hives span diverse landscapes and altitudes.

2.2 Daily series extraction

To overcome the presence of gaps at the intraday level, we extracted daily hive weight time series from the dataset.

As a preliminary step, for each year covered by the sample, we restricted our focus to the period between April 1 and September 30, which roughly corresponds to the period when honey is produced (Dolezal et al., 2019), and retained only the hives that were not subject to nomadism during that period, that is, hives whose location was stationary (Mezentsev, 2024).

Then, we resampled the intraday recordings relative to those hives to the daily frequency. Specifically, we retained the recording whose timestamp was the closest to 21:00 CET as the representative daily weight for a given hive, treating the daily value as missing when no observation was recorded between 20:00 CET and 23:59 CET. Hives with at least one missing daily observation were discarded.

Finally, we computed daily weight changes and corrected for outliers using a conservative procedure[3]. Specifically, a threshold proportional to the quantity of 1.2 kg per hour was adopted to flag outliers, in accordance with Arias-Calluari et al. (2023a). Larger Variations were replaced

[2] `https://www.3bee.com/en/`

[3] Outliers may be caused by beekeepers' actions, e.g., the removal of honey from the hive (negative outliers) or the accidental placement of additional equipment or weights on the hive (positive outliers).

with a three-day exponentially weighted moving average (EWMA), while smaller variations were set to zero. Eventually, we were able to extract daily hive weight change time series for 124 hives in 2021, 304 in 2022, 458 in 2023, and 457 in 2024.

To the best of our knowledge, no other study combines comparable spatial and temporal with daily high-frequency sampling, the typical sampling frequency employed being the monthly or the yearly one. In fact, high-frequency hive-weight series analyzed in the literature are typically narrower in both sample size and duration. For instance, Arias-Calluari et al. (2023b) report small-scale recordings from ten hives in Australia; Kulyukin et al. (2024) analyze a single-season monitoring of managed colonies in the United States; Senger et al. (2024) present a 78-hive dataset from Germany; Czekońska et al. (2023) provide a two-season 104-hive landscape dataset from Poland; and Zhu et al. (2024) introduce a 53-hive longitudinal multi-sensor dataset from Québec.

2.3 Daily series properties

Firstly, we tested for stationarity using the Augmented Dickey–Fuller test, see (Dickey and Fuller, 1979), at the 5% significance level, and obtained evidence that the null hypothesis of a unit root is rejected for all series across the different years of study.

Next, we examined the autocorrelation structure. Figure 1 presents the partial autocorrelation functions (PACF) for four representative hives that are present across all years of the dataset[4]. In these examples, the first few lags of the PACFs, particularly lags 1 through 3, are often significant. This evidence suggests that a low-order auto-regressive (AR) specification may represent a suitable modeling framework for daily hive weight changes. Moreover, to gain insight into the distributional modeling of the innovations in the AR framework, we analyzed the empirical distribution of daily hive weight changes. Figure 2 displays histograms for the representative hives, together with fitted Gaussian and Student-t distributions, obtained using maximum likelihood. The plots suggest that the Student-t distribution provides a better fit than the Gaussian, as it captures the heavy-tailed behavior present in the empirical data.

Based on these empirical observations, in Section 3 we model the dynamics of hive weights by fitting autoregressive models of order up to 3, with the innovation term's distribution specified either as a Student-t or a Gaussian[5].

[4]The representative hives correspond to different Italian regions: hive 1 was located in Piedmont at coordinates (45.686°N, 8.193°E); hive 2 in Piedmont at (44.722°N, 8.979°E); hive 3 in Lombardy at (45.399°N, 9.134°E); and hive 4 in Campania at (40.837°N, 15.330°E).

[5]We also allowed for the Gaussian specification since it represents an adequate modeling choice for a minority of hives, see the supplementary material.

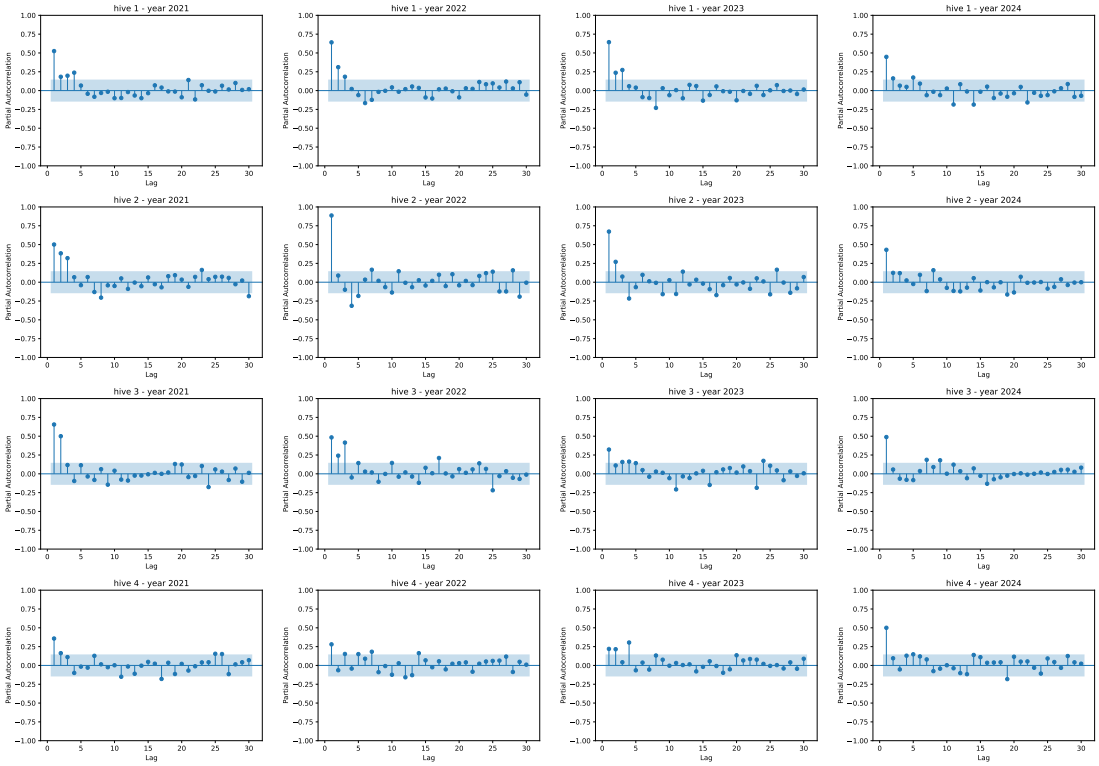

Figure 1: PACF of daily hive weight variation for hives 1–4 (top → bottom) across years 2021–2024 (left → right). The hives are located in the Italian regions of Piedmont (hives 1 and 2), Lombardy (hive 3), and Campania (hive 4).

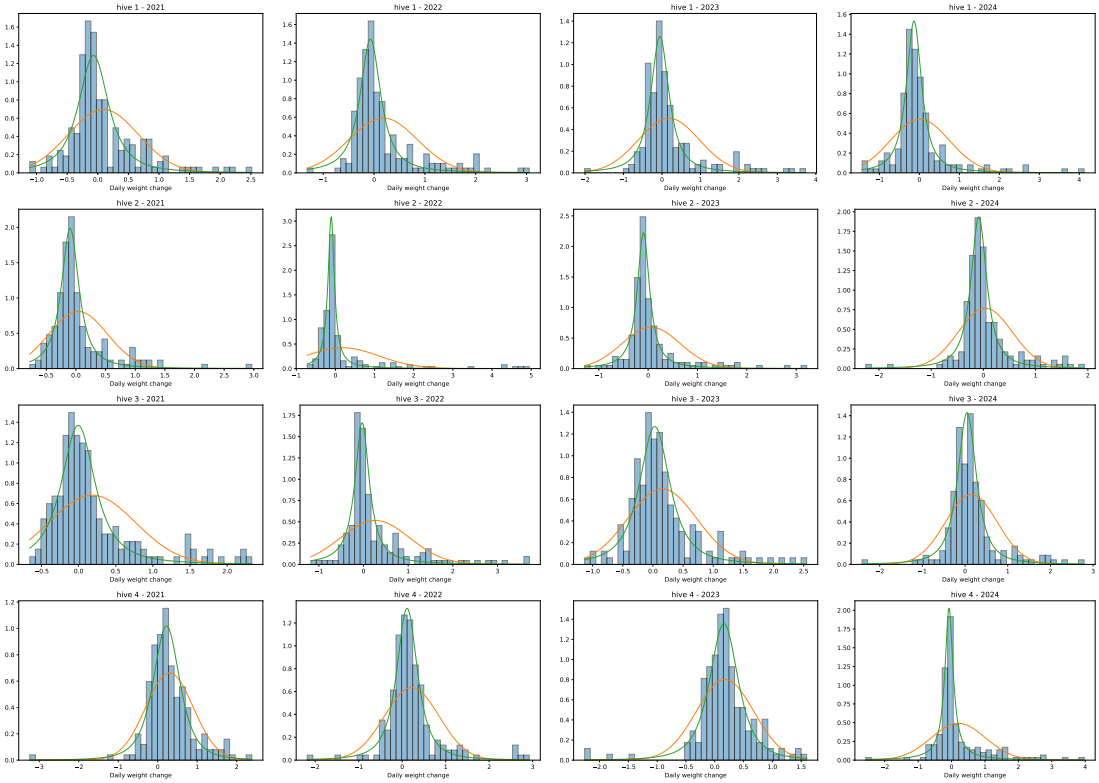

Figure 2: Empirical distributions of daily weight hive weight variation for hives 1–4 (top $\rightarrow$ bottom) across years 2021–2024 (left $\rightarrow$ right). The orange and green curves correspond to the fitted Normal and Student-t densities, respectively. The hives are located in the Italian regions of Piedmont (hives 1 and 2), Lombardy (hive 3), and Campania (hive 4).

3 Methodology

In this section, we describe the methodological aspects related to HaR estimation. Firstly, we illustrate the procedure that we used to group hives into clusters reflecting homogeneous climatic conditions. Then, we introduce the parametric data-generating process that we employed to model hive weight dynamics and isolate shocks cluster-wise. Finally, we define the HaR metric for the shocks of a given cluster and detail a numerical procedure to estimate it.

3.1 Clustering

To aggregate hives into clusters reflecting homogeneous climatic conditions, we designed a procedure that jointly accounts for geographic location and altitude. For each year, we classified hives into altitude-based categories using elevation thresholds: plains (< 300 m), hills ($300 - 600$ m), mountains (> 600 m).

In the rest of the paper, we will refer to these three categories using P, H and M. We then assigned each hive to an Italian administrative region based on geographical coordinates. For each year and for each subset of hives falling in the same region and the same altitude category, we applied hierarchical agglomerative clustering using Ward's linkage method (Ward Jr, 1963) on hive coordinates (latitude and longitude). To enforce balanced group sizes, we implemented a recursive splitting algorithm: whenever a cluster exceeded a maximum size of 10 hives, we partitioned it into two subgroups and recursively applied the procedure until all clusters respected the size constraint. This approach allowed us to form clusters that are geographically compact[6]. Figure 3 shows the spatial distribution of hive clusters for each altitude category and year.

[6]If the regional constraint was not included, but only altitude was considered, the algorithm could lead to clusters spanning multiple and possibly very distant regions (although including hives with similar altitude), thereby not reflecting truly homogeneous climatic conditions.

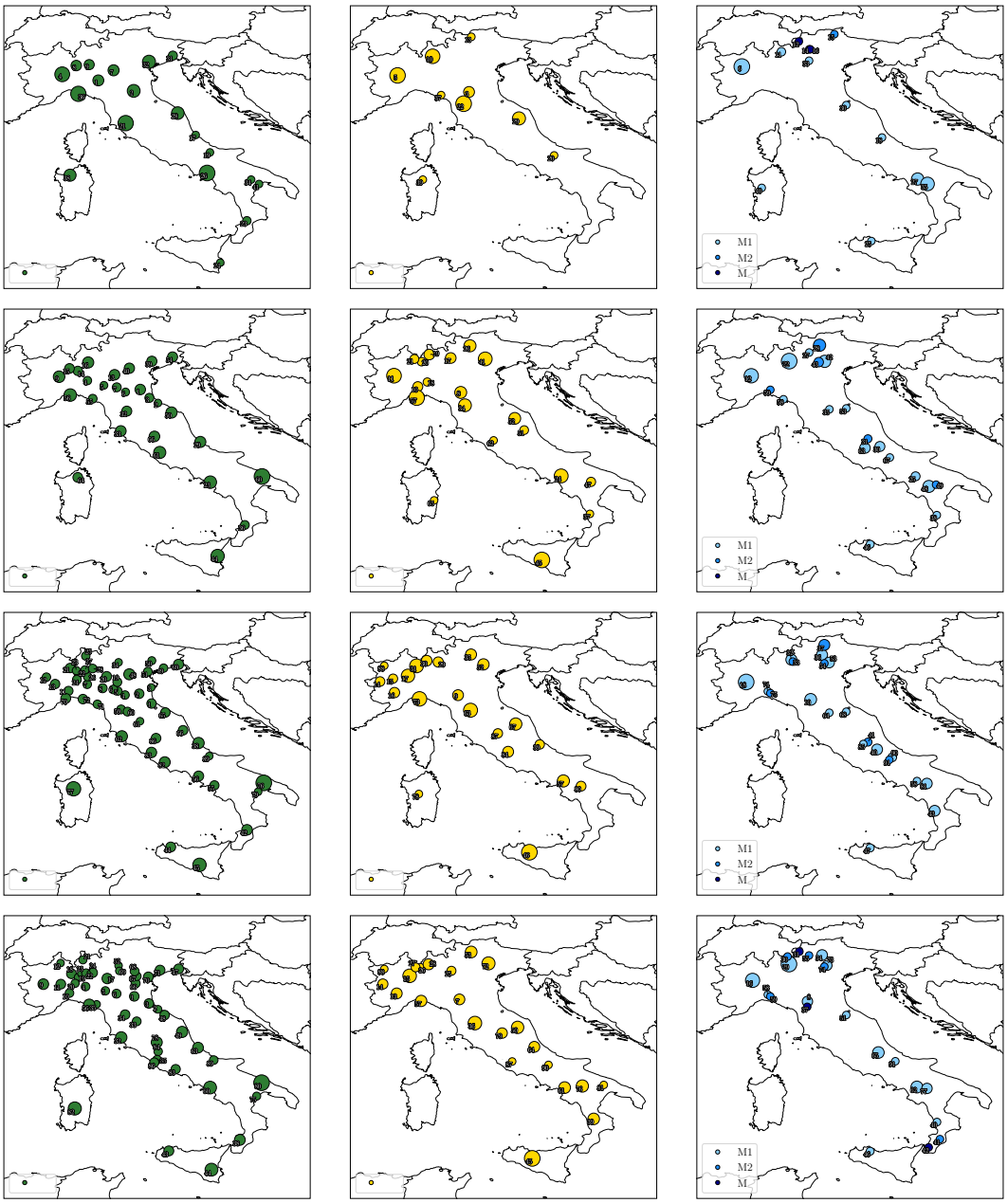

Figure 3: Maps by category P, H, and M (left→right) for the years 2021–2024 (top→bottom). Each marker is centered at the medoid, that is, the hive location that minimizes the sum of geodesic distances to all other hives in the cluster. Marker areas scale with a cluster's spread, defined as the maximum great-circle distance from the medoid to its member hives (haversine metric). The mountain category M is further divided into three subgroups, $M1$ (600–900 m), $M2$ (900–1200 m) and $M3$ ($\geq$1200 m), to better appreciate altitude-related differences.

3.2 Statistical modeling

Let J^y and $N^{j,y}$ denote, respectively, the number of clusters formed for year y, with $y \in \{2021, 2022, 2023, 2024\}$, and the number of hives belonging to the j-th cluster in year y. For brevity, in the rest of the Section, we drop the superscript y and consider a generic year. Since we focused on the period April 1 - September 30, in every year of study, we have $T = 183$ hive weight observations for a single hive.

For $t = 1, ..., T$, $j = 1, ..., J$, and $i = 1, ..., N^j$, let $w^j_{i,t}$ denote the weight of the i-th hive in the j-th cluster on day t. The variable of interest in our study is the daily variation in the hive weight, that is,

$$y^j_{i,t} := w^j_{i,t} - w^j_{i,t-1}, \quad t > 1.$$

Based on the empirical evidence from Section 2.3, for each j and i, we assume that $y^j_{i,t}$ is an AR process of order P^j_i, that is,

$$y^j_{i,t} = \bar{\phi}^j_i + \sum_{k=1}^{P^j_i} \phi^j_{i,k} y^j_{i,t-k} + \varepsilon^j_{i,t}, \tag{1}$$

where, for each j and t, $\varepsilon^j_t := (\varepsilon^j_{1,t}, \ldots, \varepsilon^j_{N_j,t})$ is a vector of random shocks following a meta-distribution, that is, a joint distribution with arbitrary copula and margins (see Sklar (1959))[7].

We assume that the joint distribution of ε^j_t is stationary, that is, its margins and copula are the same at every t. Furthermore, for each j, we assume that the N^j-dimensional sequence $\{\varepsilon^j_t\}_{t=1,\ldots,T}$ is independent through time. In our implementation, we considered two alternative parametric specifications for the copula of ε^j_t, namely the Gaussian copula $C^{\mathcal{G}}_{C^j}$ and the t copula $C^{\mathcal{T}}_{\nu^j, C^j}$, where C^j is a $N^j \times N^j$ correlation matrix and ν_j indicates the degrees of freedom. Furthermore, we considered two alternative parametric specifications for the marginal distributions, that is, a zero-mean Gaussian distribution $\mathcal{N}(0, \sigma^j_i)$, where σ^j_i indicates standard deviation of the i-th hive in the j-th cluster, and a zero-mean t distribution $t(0, \gamma^j_i, \nu^j_i)$, where γ^j_i and $\nu^j_i > 2$ represent, respectively, the scale parameter and the degrees of freedom of the i-th hive in the j-th cluster. Note that the t copula implies the existence of tail dependence between pairs of shocks, while t-distributed margins account for extreme events (see Ch. 7 of McNeil et al. (2015)).

3.3 Honey-at-Risk

3.3.1 Definition

Consider the j-th cluster. For a given time t, we define the cluster-wise average shock in the hive weight as

$$\mathcal{L}^j_t := \frac{1}{N^j} \sum_{i=1}^{N^j} \varepsilon^j_{i,t},$$

[7] As discussed in Ch. 7 of McNeil et al. (2015), compared to a multivariate parametric model (e.g., a multivariate Normal or Student t distribution), the choice of a meta-distribution may increase flexibility and provide improved accuracy by allowing for the separate modeling of the marginal distributions and the copula of a multidimensional random vector. In our context, this aspect is particularly crucial for adequately modeling the univariate tails and the joint tails of the production shocks in cluster.

where $\varepsilon_{i,t}^j$ comes from model (1), and denote by $F_{\mathcal{L}^j}$ its cumulative distribution function (c.d.f.). Then, the Honey-at-Risk metric for the j-th cluster at the confidence level $\alpha \in (0,1)$ is defined as

$$\mathrm{HaR}_\alpha(\mathcal{L}^j) = \inf\{x \in \mathbb{R} : F_{\mathcal{L}^j}(x) \geq \alpha\}.$$

The estimator of $\mathcal{L}_t^j$ that we employ in our study is given by

$$L_t^j := \frac{1}{N^j} \sum_{i=1}^{N^j} e_{i,t}^j, \tag{2}$$

where $e_{i,t}^j$ is the residual from the estimation of model (1) for the i-th hive in j-th cluster at time t. The associated Honey-at-Risk estimator for the j-th cluster at the confidence level $\alpha \in (0,1)$ reads as

$$\mathrm{HaR}_\alpha(L^j) = \inf\{x \in \mathbb{R} : F_{L^j}(x) \geq \alpha\},$$

where F_{L^j} denotes the empirical c.d.f. of L_t^j.

3.3.2 Estimation procedure

For the estimation of the HaR metric, we followed four steps (see Ch. 7 of McNeil et al. (2015)).

Step 1: Univariate model estimation. The first step involves the estimation of the model in (1) for each hive in the cluster. For this task, we used OLS and estimated the model parameters for different lag orders. Eventually, we selected the combination of the lag parameter P_i^j and residual distribution that optimizes the Bayesian Information Criterion (BIC), see Schwarz (1978). As a result, for each i in cluster j, we obtained the AR parameter estimates $(\widehat{\phi}_{i,0}^j, \widehat{\phi}_{i,1}^j, ..., \widehat{\phi}_{i,\widehat{P}_i^j}^j)$ and the residual parameter estimates $(\widehat{\sigma}_i^j, \widehat{\nu}_i^j)$, in the case of a t-distributed shock, or $\widehat{\sigma}_i^j$ in the case of a Gaussian shock. Estimated parameter values are reported in the supplementary material, Tables 1–4.

Step 2: Copula estimation. Let $\{e_{i,t}^j\}_{t=1,...,T}$ denote the series of residuals from step 1 for hive i in cluster j. In the second step, we used the series of residuals to select a copula model for the dependence structure among the shocks. Firstly, using the model's residuals, for each i in cluster j and for each time t, we obtained the so-called copula pseudo-observations, defined as

$$u_{i,t}^j := \frac{1}{T+1} \sum_{s=1}^{T} \mathbb{1}_{\{e_{i,s}^j \leq e_{i,t}^j\}}.$$

Then, we employed the pseudo-observations to perform the maximum likelihood estimation of the copula models $C_{C^j}^{\mathcal{G}}$ and $C_{\nu^j,C^j}^{\mathcal{T}}$ and retained the one with the smallest BIC. Estimated parameter values are reported in the supplementary material, Tables 5–8.

Step 3: Meta-distribution simulation. In the third step, we simulated the meta-distribution of the shocks in cluster j. Specifically, we first simulated a number S of N^j-dimensional samples of length T of the selected copula model. Then, we applied the quantile

transform to turn such samples into samples from the estimated meta-distribution. Specifically, depending on the fitted marginal model, we performed either the transform

$$\widetilde{e}_{i,t}^{j,s} = \widehat{\sigma}_i^j t_{\widehat{\nu}_i^j}^{-1}\left(\widetilde{u}_{i,t}^{j,s}\right),$$

where $t_{\nu_i^j}^{-1}$ denotes the inverse of a zero-mean, unit-variance t distribution with ν_i^j degrees of freedom, or the transform

$$\widetilde{e}_{i,t}^{j,s} = \widehat{\sigma}_i \Phi^{-1}\left(\widetilde{u}_{i,t}^{j,s}\right),$$

where Φ^{-1} denotes the inverse of the standard normal c.d.f.

Step 4: HaR computation. In the last step, for every simulated sample $s \in \{1, ..., S\}$, we reconstructed the cluster-wise average hive residual weight changes, that is,

$$\widetilde{L}_t^{j,s} := \frac{1}{N^j} \sum_{i=1}^{N^j} \widetilde{e}_{i,t}^{j,s},$$

and obtained HaR values denoted by

$$\mathrm{HaR}_\alpha(\widetilde{L}^{j,s}) := \inf\{x \in \mathbb{R} : F_{\widetilde{L}^{j,s}}(x) \geq \alpha\},$$

where $F_{\widetilde{L}^{j,s}}$ denotes the empirical c.d.f. of $\widetilde{L}_t^{j,s}$ on the s-th simulated path.

The Monte Carlo HaR (MC-HaR) estimate for the j-th cluster was obtained as the average HaR over the S simulated samples, that is, as

$$\text{MC-HaR}_\alpha^j := \frac{1}{S} \sum_{s=1}^{S} \mathrm{HaR}_\alpha(\widetilde{L}^{j,s}). \tag{3}$$

4 Results

In this section, we illustrate empirical HaR estimates and investigate the link between extreme honey production losses, that is, losses that are smaller than HaR at a given confidence level, and atmospheric temperature. We first provide graphical evidence of the distribution of HaR estimates across the Italian territory and then summarize them through averages for clusters in the same region and altitude category, highlighting significant differences in magnitude and dispersion across altitude categories. Then, we implement a logistic regression to study the link between the likelihood of an extreme honey production loss in a given cluster and the average, minima and maxima temperatures associated with the cluster, at the daily level.

4.1 HaR estimates

We obtained cluster-wise MC-HaR values at the 10% confidence level, using $S = 10^5$ iterations. Figure 4 displays estimation results for the different altitude categories over the period 2021–2024.

To more effectively summarize estimation results, for each year under study, we also report average MC-HaR values for clusters in the same region and the same altitude category. Specifically, we define such an average in year y as

$$H^y_{r,c,\alpha} := \frac{1}{\#E^y_{r,c}} \sum_{j=1}^{J^y} \text{MC-HaR}^{j,y}_\alpha \mathbb{1}_{\{j \in E^y_{r,c}\}}$$

where $\text{MC-HaR}^{j,y}_\alpha$ is defined in (3)[8], $E^y_{r,c}$ denotes the subset of clusters belonging to the Italian administrative region r and the altitude category $c \in \{P, H, M\}$ in year y and $\#E^y_{r,c}$ indicates its cardinality.

Tables 1–3[9] detail, for each altitude category, the values of $H^y_{r,c,\alpha}$ for $\alpha = 10\%$.

We make the following remarks on the estimation results. Firstly, altitude appears to influence the magnitude of MC-HaR. For each year y and region r, to assess relative risk across altitude categories, we computed the differences

$$\Delta_{M,P} H^y_{r,\alpha} := H^y_{r,M,\alpha} - H^y_{r,P,\alpha}, \tag{4}$$

$$\Delta_{M,H} H^y_{r,\alpha} := H^y_{r,M,\alpha} - H^y_{r,H,\alpha} \tag{5}$$

and

$$\Delta_{H,P} H^y_{r,\alpha} := H^y_{r,H,\alpha} - H^y_{r,P,\alpha}. \tag{6}$$

If, e.g., $\Delta_{M,P} H^y_{r,\alpha} < 0$, the risk is higher in the M category, compared to the P category, in the region r during year y. For $\alpha = 10\%$, we found that $\Delta_{M,P} H^y_{r,\alpha}, \Delta_{M,H} H^y_{r,\alpha}$ and $\Delta_{H,P} H^y_{r,\alpha}$ are negative in, respectively, 71%, 70% and 58% of the instances across all years and regions, with average values of -0.30, -0.32 and -0.06. Such values suggest that a significant difference in MC-HaR magnitude is experienced, on average, in the M category compared to the other two categories, while no significant difference is experienced, on average, in the H category compared to the P category. To provide further support to this finding, we run two-sided t-tests on the sample values of (4), (5) and (6), under the null of a zero mean, and found p-values equal to 0.0143, 0.0028 and 0.558 for respectively (4), (5) and (6).

Secondly, dispersion seems to vary substantially across altitude categories. For $\alpha = 10\%$, in the P category, average MC-HaR values are relatively homogeneous across regions, typically concentrated between -0.7 and -1.5. By contrast, the M category displays much wider variation: in some northern regions (e.g., Lombardy, Veneto, Trentino-Alto Adige) average MC-HaR often falls below -2, while in central and southern regions, values remain closer to -1. This suggests that mountainous areas are not only exposed to higher risk on average but also subject to greater spatial heterogeneity.

[8] Here we add the superscript y to indicate the estimation year.

[9] For some regions, the number of clusters grows importantly from 2021 to 2023, reflecting the expansion of the dataset. For example, Emilia-Romagna and Lombardy show a noticeable increase in the number of clusters within the P category. Instead, the changes in the number of clusters between 2023 and 2024 are less pronounced, in line with the stabilization of the growth of the dataset.

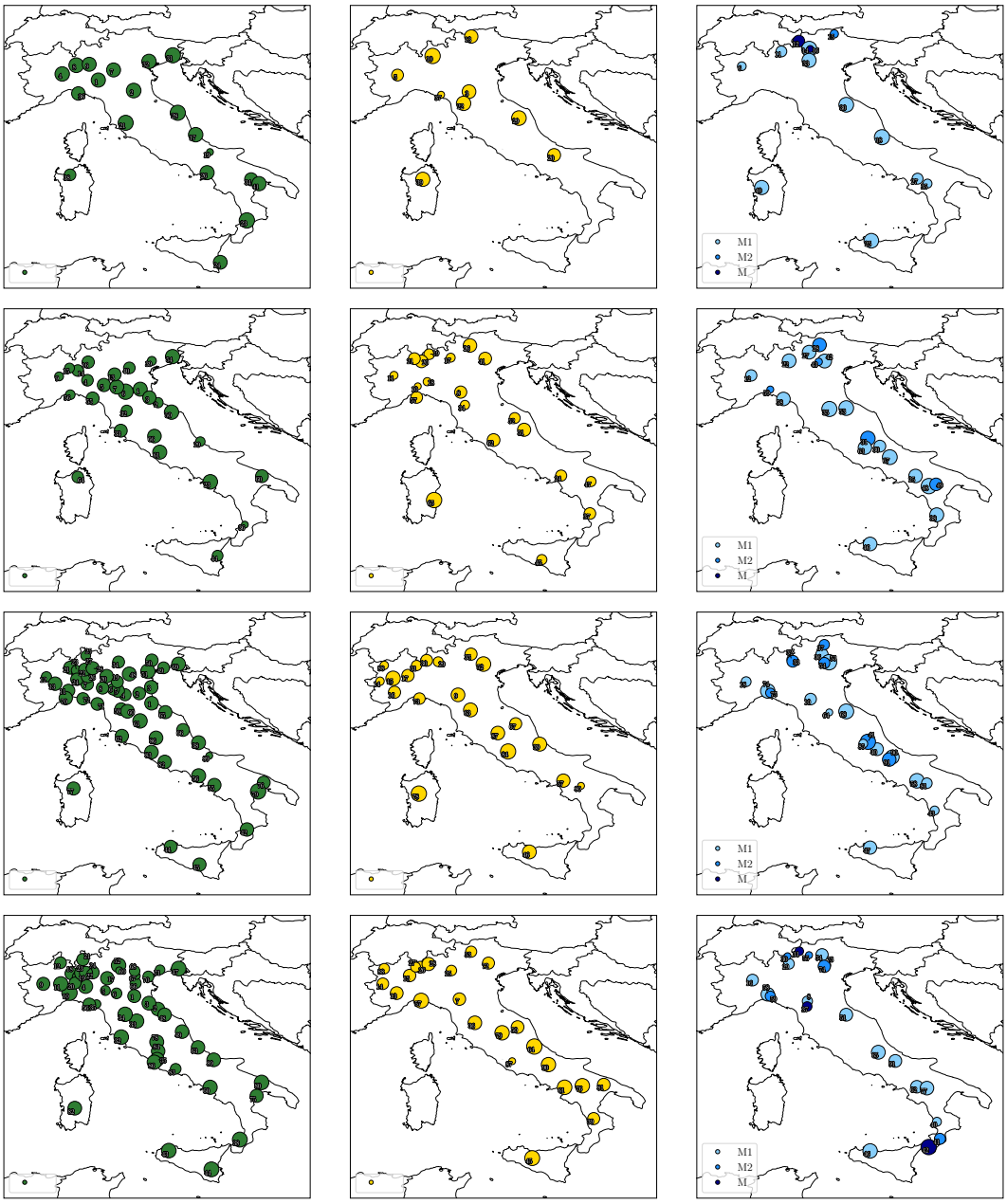

Figure 4: Maps for altitude categories P, H, and M (left→right) and years 2021–2024 (top→bottom), showing cluster-wise MC-HaR estimates at $\alpha = 10\%$. The area of each circle is proportional to the corresponding cluster's MC-HaR value, with larger markers indicating higher estimated risk.

	2021		2022		2023		2024	
	# Cl	Mean	# Cl	Mean	# Cl	Mean	# Cl	Mean
Abruzzo	1	-1.026	1	-1.847	1	-0.795	1	-1.172
Basilicata	1	-1.856	-	-	1	-0.007	1	-1.259
Calabria	1	-0.464	1	-2.254	1	-1.037	1	-0.656
Campania	1	-1.232	1	-0.714	2	-0.913	1	-0.908
Emilia-Romagna	2	-1.069	7	-1.063	8	-1.365	6	-1.421
Friuli-Venezia Giulia	1	-0.568	1	-0.803	1	-0.841	1	-0.749
Lazio	-	-	1	-0.851	2	-0.977	4	-1.249
Liguria	1	-1.594	2	-1.488	3	-1.351	1	-1.427
Lombardia	3	-1.149	4	-1.401	10	-1.558	8	-1.355
Marche	1	-0.571	1	-0.781	2	-1.273	2	-1.139
Molise	1	-4.435	-	-	1	-3.500	1	-1.066
Piemonte	1	-1.011	1	-1.960	3	-1.617	4	-1.083
Puglia	1	-1.148	1	-1.153	1	-1.307	1	-0.792
Sardegna	1	-2.559	1	-1.443	1	-1.182	1	-1.063
Sicilia	1	-1.401	1	-1.608	2	-1.222	2	-0.715
Toscana	1	-0.667	2	-1.322	4	-0.991	4	-1.242
Trentino-Alto Adige	-	-	-	-	1	-1.729	1	-0.813
Umbria	-	-	1	-0.967	1	-0.925	1	-0.982
Valle d'Aosta	-	-	-	-	-	-	-	-
Veneto	1	-1.070	2	-1.687	4	-1.130	6	-2.250

Table 1: Average MC-HaR by year and region for the plain category (P) at $\alpha = 10\%$. For each year and region, the table reports the number of clusters ($\#\ Cl$) and the cluster-wise mean ($Mean$).

	2021		2022		2023		2024	
	# Cl	Mean	# Cl	Mean	# Cl	Mean	# Cl	Mean
Abruzzo	-	-	-	-	1	-1.007	1	-0.618
Basilicata	-	-	1	-1.599	1	-2.772	1	-0.967
Calabria	-	-	1	-1.231	-	-	1	-1.767
Campania	-	-	1	-1.352	1	-1.225	1	-0.912
Emilia-Romagna	1	-1.256	1	-1.090	1	-1.094	1	-1.646
Friuli-Venezia Giulia	-	-	-	-	-	-	-	-
Lazio	-	-	1	-0.950	1	-0.576	1	-3.038
Liguria	1	-3.084	1	-1.209	1	-1.777	1	-0.771
Lombardia	1	-0.644	5	-1.486	3	-1.962	4	-1.970
Marche	1	-0.924	2	-1.127	1	-1.607	1	-1.568
Molise	1	-1.520	-	-	-	-	1	-0.996
Piemonte	1	-1.885	2	-2.098	4	-1.609	3	-1.697
Puglia	-	-	-	-	-	-	1	-1.576
Sardegna	1	-1.071	1	-0.159	1	-0.525	-	-
Sicilia	-	-	1	-1.579	1	-1.052	1	-0.776
Toscana	1	-1.014	1	-1.776	1	-1.010	1	-1.254
Trentino-Alto Adige	1	-1.378	1	-0.725	1	-1.545	1	-2.047
Umbria	-	-	-	-	1	-1.117	1	-0.977
Valle d'Aosta	-	-	-	-	1	-2.489	1	-2.236
Veneto	-	-	1	-0.952	1	-0.796	1	-1.786

Table 2: Average MC-HaR by year and region for the hill category (H) at $\alpha = 10\%$. For each year and region, the table reports the number of clusters retained ($\#\ Cl$) and the cluster-wise mean ($Mean$).

	2021		2022		2023		2024	
	# Cl	Mean	# Cl	Mean	# Cl	Mean	# Cl	Mean
Abruzzo	1	-0.488	2	-1.651	2	-0.977	1	-1.084
Basilicata	1	-2.000	2	-1.224	1	-1.997	1	-1.175
Calabria	-	-	1	-1.259	1	-2.400	3	-1.777
Campania	1	-1.600	1	-1.462	1	-0.803	1	-1.611
Emilia-Romagna	-	-	-	-	1	-1.755	1	-2.414
Friuli-Venezia Giulia	-	-	-	-	-	-	-	-
Lazio	-	-	1	-1.519	1	-2.400	-	-
Liguria	-	-	2	-2.749	2	-1.576	2	-1.848
Lombardia	2	-1.522	1	-1.335	2	-2.055	3	-2.796
Marche	1	-0.604	1	-0.857	1	-0.678	1	-1.482
Molise	-	-	1	-0.781	2	-1.090	1	-1.520
Piemonte	1	-2.044	1	-2.514	1	-2.187	1	-1.915
Puglia	-	-	-	-	-	-	-	-
Sardegna	1	-0.763	-	-	-	-	-	-
Sicilia	1	-0.631	1	-1.413	1	-1.264	1	-0.815
Toscana	-	-	1	-0.628	1	-2.785	1	-2.600
Trentino-Alto Adige	3	-1.631	2	-1.406	2	-1.802	2	-2.475
Umbria	-	-	-	-	-	-	-	-
Valle d'Aosta	-	-	-	-	-	-	-	-
Veneto	1	-0.788	2	-2.610	2	-1.195	2	-2.432

Table 3: Average MC-HaR by year and region for the aggregated mountain category (M) at $\alpha = 10\%$. For each year and region, the table reports the number of clusters retained ($\#\ Cl$) and the cluster-wise mean ($Mean$).

4.2 HaR exceedances and atmospheric temperature

In this section, we study the link between atmospheric temperature and the occurrence of extreme daily honey production losses in our dataset.

For each cluster j and day t in year y, we defined the binary variable

$$Z_{t,\alpha}^{j,y} := \mathbb{1}_{\{L_t^{j,y} < \text{MC-HaR}_\alpha^{j,y}\}},$$

where $L_t^{j,y}$ and $\text{MC-HaR}_\alpha^{j,y}$ are given, resp., in (2) and (3)[10]. If $Z_{t,\alpha}^{j,y} = 1$, we are in the presence of an extreme loss for cluster j on day t of year y, in the sense that the loss exceeds, in absolute value, the risk measure computed at the confidence level α.

Moreover, letting $A_t^{j,y}$ and $M_t^{j,y}$ denote, resp., the daily average and maximum atmospheric

[10]The superscript y is added to indicate the estimation year.

temperature recording associated with cluster j on day t of year y, we defined the binary variable

$$U_t^{j,y} := \mathbb{1}_{\{M_t^{j,y} \geq \vartheta_U^{j,y}\}},$$

where $\vartheta_U^{j,y}$ denotes a suitable real-valued threshold associated with cluster j and year y[11]. If $U_t^{j,y} = 1$, the daily maximum temperature on day t of year y associated with cluster j violates the threshold, indicating an extreme realization of the temperature maximum[12].

For each year y considered in our study, we estimated the following logistic regression model:

$$\ln\left(\frac{P(Z_{t,\alpha}^{j,y} = 1 | \, U_t^{j,y}, A_t^{j,y})}{1 - P(Z_{t,\alpha}^{j,y} = 1 | \, U_t^{j,y}, A_t^{j,y})}\right) = \beta_0 + \beta_1 A_t^{j,y} + \beta_2 U_t^{j,y}, \tag{7}$$

where $j \in \{1, \ldots, J^y\}$, $t \in \{1, \ldots, T\}$ and β_0, β_1 and β_2 are real-valued coefficients.

In implementing model (7), we employed atmospheric temperature data from the Copernicus ERA 5 database[13]. Specifically, we downloaded hourly observations for the 2-meter atmospheric temperature and computed daily averages and maxima. For each cluster, we used temperatures recorded in correspondence with the medoid of the cluster, that is, the hive location that minimizes the sum of geodesic distances to all other hives in the cluster. The confidence level α was set at 0.25, a value which reflects a compromise between targeting relatively extreme realizations and retaining a sufficiently large number of exceedances for statistical analysis. The threshold $\vartheta_U^{j,y}$ was chosen as the 0.95-quantile of the daily maxima associated with cluster j on year y.

Table 4 illustrates the results of the estimation of the model in (7) for 2021, 2022, 2023 and 2024. The intercept is negative and statistically significant for all the years considered. The coefficient β_1 shows negative and statistically significant values in all years of study, indicating that days characterized by warmer average temperatures were associated with a lower probability of extreme losses. The values of β_2 are positive throughout the period of study and suggest that the occurrence of hot-temperature peaks was associated with an increase in the odds of extreme losses. If we exclude 2022, the estimates of β_2 can be deemed statistically significant at the usual 5% significance level, the largest p-value being 0.033 in 2021. The lack of statistical significance in 2022 can be justified by the fact that the highest value of the daily maximum temperature records is only 39.1°C in 2022, well below the corresponding quantity for the other years of study (41.5°C in 2021, 42.3°C in 2023, and 40.4°C in 2024). As the presence of maxima peaks is detected relative to each year's maxima distribution (specifically, relative to the year-specific 95-th percentile), the peaks of 2022 may not correspond to values typically associated with extreme-heat stress for honeybees, in that the literature places such values above 40°C (see, e.g., Vincze et al. (2025)).

[11] A cluster- and year-specific threshold reflects local climatic conditions, allowing temperature maxima peaks to be carefully identified across heterogeneous environments, unlike in the case of a fixed absolute threshold independent of the cluster and/or the year.

[12] We recall that the study employs time series sampled over the period April 1 - September 30, during which peaks of excessive heat may play a crucial role.

[13] https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels

Year	coefficient	estimate	std error	p-value
2021	β_0	-2.069	0.120	0
	β_1	-0.034	0.007	0
	β_2	0.294	0.138	0.033
2022	β_0	-2.476	0.152	0
	β_1	-0.030	0.008	0
	β_2	0.077	0.125	0.537
2023	β_0	-2.037	0.132	0
	β_1	-0.068	0.007	0
	β_2	0.367	0.118	0.002
2024	β_0	-2.726	0.136	0
	β_1	-0.039	0.007	0
	β_2	0.276	0.117	0.018

Table 4: Estimation results for the model in (7). P-values smaller than 10^{-3} are reported as 0.

5 Discussion

5.1 Interpretation

The geographical pattern of HaR estimates indicates that downside tail risk related to honey production was unevenly distributed across Italy during the period of study. In particular, clusters located in mountainous areas exhibited not only higher exposure to such risks but also greater dispersion of HaR values, compared to plains and hills. This evidence suggests that higher-altitude increases the risk of extreme adverse realizations. This finding can be explained not only by the more erratic weather and shorter flowering periods typical of high-altitude environments, but also by the longer duration of foraging trips and the reduced rate of flower visits per unit of time. Together, these factors may result in a shorter foraging window and a higher energy expenditure for bees at higher elevations. In addition, the prolonged winters typical of high-altitude environments require colonies to rely more heavily on their stored reserves, which can leave insufficient surplus honey for harvest. By contrast, lowland areas are typically associated with longer foraging seasons and more stable climatic regimes, which tend to mitigate the occurrence of extreme negative production shocks.

Regarding the link between atmospheric temperature and losses exceeding HaR, the existing literature has extensively documented that temperature matters for honeybee activity, colony performance, and ultimately honey production (González-Varo et al., 2013; Abou-Shaara et al., 2017; Meikle et al., 2018; Norrström et al., 2021; Zhao et al., 2021; Gounari et al., 2022; Porras et al., 2023; O'Connell et al., 2024). However, much of this evidence is typically framed in terms of correlations between average (yearly or seasonal) production responses and climatic variables, and therefore does not explicitly isolate how temperature relates to the lower tail of the production-shock distribution. Our findings contribute to this strand of work by showing, in a tail-risk analysis framework, that average temperature is primarily informative about exposure

to colder conditions, which are associated with a higher likelihood of extreme losses, whereas extreme realizations of daily maximum temperature capture the role of excessive heat.

5.2 Limitations

Our study is subject to some limitations, which mainly relate to data availability constraints and a natural trade-off between spatial coverage and local detail.

Firstly, while the dataset we employed is uniquely detailed in spatial coverage and hive-level resolution, it spans only four years and thus offers limited temporal depth. This limitation is mitigated by the extraction and use of daily production series, which provide a higher temporal resolution than the monthly or annual data typically used in the literature and support statistically meaningful inference even over shorter horizons. Moreover, it is worth noting that in the presence of rapidly evolving climate conditions, extending the window far into the past may incorporate historical regimes that are less representative of the current risk profile across clusters.

Secondly, we acknowledge that the clustering procedure, which relies on geographic coordinates and altitude, does not incorporate local factors, such as micro-climatic or land-use heterogeneities, that can influence foraging conditions and, consequently, colony productivity. Nonetheless, such fine-scale analyses are generally more feasible for a limited number of hives within narrowly defined regions, while our study encompasses the entire national territory and includes hundreds of hives. Moreover, since the dataset does not cover the same hive across multiple years, clusters differ from one year to another.

Finally, although average temperatures and temperature peaks emerge as statistically significant explanatory factors for extreme production losses, other local environmental stressors, such as droughts, pesticide use, or diseases, remain outside the current scope of the analysis. Besides, we note that honey production is affected by climate change not only through direct channels but also through indirect ones, which affect bee physiology, colony health, and foraging behavior. Thus, honey production can be deemed as inherently more complex than most climate-sensitive agricultural activities for which quantile-based risk measures are used to deal with rare but impactful shocks.

5.3 Implications

The spatial heterogeneity of HaR values allows the identification of vulnerable production systems, such as northern and mountainous zones, and highlights regions where extreme climatic events disproportionately affect honey yields (Flores et al., 2019). By capturing tail risks that commonly used elasticity measures cannot capture[14], HaR provides a basis for targeted adaptation and risk management, including evidence-based interventions such as smart hive monitoring, beekeeping training, and forage diversification in high-risk areas.

[14]Climate–yield elasticity quantifies the mean change in honey yield per unit change in climate variables (e.g., Celsius degrees), see Holmes (2002).

At the policy and financial level, HaR can inform the design of weather-indexed insurance and compensation schemes (by defining payout triggers for extremely low-yield years and guiding premium calculations), and help identify apiaries most exposed to catastrophic losses. In combination with elasticity, it offers a more complete picture: while elasticity quantifies sensitivity to average climatic variation, HaR addresses rare but severe events that often drive the most significant losses.

From a methodological perspective, this shift—from modeling mean outcomes to focusing on the tails—is consistent with the broader literature on environmental extremes, where rare events require dedicated inferential tools rather than extrapolation from average responses (Davison and Gholamrezaee, 2012; Coles, 2001). Abnormal negative outcomes in honey production frequently result from nonlinear extremes, such as late frosts, heatwaves, droughts, or prolonged bad weather that limit foraging (Vincze et al., 2025).

Overall, HaR should be seen as a complement to traditional elasticity measures, aligning statistical assessment with the practical interest in extreme production losses and enabling more informed decisions in policy, insurance, and operational management of honey production.

6 Conclusions

This paper studies honey production losses through a tail-oriented framework applied to a unique dataset of hive-weight data from Italy over the period 2021–2024. The main contribution of this work is measurement-oriented. By constructing HaR as a cluster-level quantile of daily production shocks, the statistical target is shifted from the center of the distribution to its lower tail, to capture risk exposure more accurately in the presence of heavy-tailed shocks.

The paper also provides an empirical contribution that reveals pronounced spatial heterogeneity in downside honey-production tail risk. Mountain clusters exhibit higher HaR values and greater tail risk dispersion than plains and hills, suggesting that exposure to extreme losses may vary across altitude levels. As these differences may only be weakly reflected in average production levels, these findings highlight the importance of tail-based risk measures for characterizing vulnerability and inform risk assessment, insurance design, and policy interventions aimed at effectively mitigating the effects of increasing climatic variability.

The analysis further indicates that atmospheric temperature is associated with extreme honey production losses through two distinct dimensions. Indeed, average temperature levels and extreme realizations of temperature maxima exhibit different associations with the occurrence of extreme losses. This highlights the fact that focusing exclusively on average weather conditions may obscure mechanisms relevant to explaining extreme honey production outcomes.

7 Acknowledgements

The authors acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European

Union – NextGenerationEU– Project Title *Honey BEE VOLatility: An environmental index for assessing climatic risk impact on ecosystems service provision*, CUP B53D23029520001 - Grant Assignment Decree No. P2022AX5TH by the Italian Ministry of University and Research (MUR). The authors thank 3BEE S.R.L. for providing the data on bee-hives.

References

Abou-Shaara, H., Owayss, A., Ibrahim, Y., and Basuny, N. (2017). A review of impacts of temperature and relative humidity on various activities of honey bees. *Insectes Sociaux*, 64(4):455–463.

Arias-Calluari, K., Colin, T., Latty, T., Myerscough, M., and Altmann, E. G. (2023a). Modelling daily weight variation in honey bee hives. *PLoS Comput Biol*, 19(3).

Arias-Calluari, K., Colin, T., Latty, T., Myerscough, M., and Altmann, E. G. (2023b). Modelling daily weight variation in honey bee hives. *PLOS Computational Biology*, 19(3):e1010880.

Burrill, R. M. and Dietz, A. (1981). The response of honey bees to variations in solar radiation and temperature. *Apidologie*, 12(4):319–328.

Calovi, M., Grozinger, C.-M., Miller, D. A., and Goslee, S. C. (2021). Summer weather conditions influence winter survival of honey bees (apis mellifera) in the Northeastern United States. *Scientific reports*, 11 (1):1553.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer London, London.

Colivicchi, I., Dell'Acqua, S., and Russo, V. (2025). A parametric insurance policy for beekeepers and honey production: random forest regressions and real-world pricing. *Decisions in Economics and Finance*.

Czekońska, K., Łopuch, S., Miścicki, S., Bańkowski, J., and Szabla, K. (2023). Monitoring of hive weight changes in various landscapes. *Apidologie*, 54(3):30.

Davison, A. C. and Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences*, 468:581–608.

Delaplane, K., Mayer, D., et al. (2000). *Crop pollination by bees*. CABI publishing.

Descamps, C., Boubnan, N., Jacquemart, A. L., and Quinet, M. (2021). Growing and flowering in a changing climate: Effects of higher temperatures and drought stress on the bee-pollinated species impatiens glandulifera royle. *Plants (Basel)*, 10(5):988.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

Dolezal, A. G., Clair, A. L. S., Zhang, G., Toth, A. L., and O'Neal, M. E. (2019). Native habitat mitigates feast–famine conditions faced by honey bees in an agricultural landscape. *Proceedings of the National Academy of Sciences*, 116(50):25147–25155.

Flores, J., Gil-Lebrero, S., Gámiz, V., M.I., R., Ortiz, M., and Quiles, F. (2019). Effect of the climate change on honey bee colonies in a temperate mediterranean zone assessed through remote hive weight monitoring system in conjunction with exhaustive colonies assessment. *Science of The Total Environment*, 653:1111–1119.

Garibaldi, L. A., Carvalheiro, L. G., Leonhardt, S. D., Aizen, M. A., Blaauw, B. R., Isaacs, R., Kuhlmann, M., Kleijn, D., Klein, A. M., Kremen, C., et al. (2014). From research to action: enhancing crop yield through wild pollinators. *Frontiers in Ecology and the Environment*, 12(8):439–447.

González-Varo, J., Biesmeijer, J., Bommarco, R., Potts, S., Schweiger, O., Smith, H., Steffan-Dewenter, I., Szentgyörgyi, H., Woyciechowski, M., and Vilà, M. (2013). Combined effects of global change pressures on animal-mediated pollination. *Trends in Ecology & Evolution*, 28:524–530.

Gordo, O. and Sanz, J. (2006). Temporal trends in phenology of the honey bee Apis mellifera (L.) and the small white Pieris rapae (L.) in the Iberian Peninsula (1952–2004). *Ecological Entomology*, 31(3):261–268.

Gounari, S., Proutsos, N., and Goras, G. (2022). How does weather impact on beehive productivity in a mediterranean island? *Italian Journal of Agrometeorology*, 1:65–81.

Gray, A., Brodschneider, R., Adjlane, N., Ballis, A., Brusbardis, V., Charrière, J.-D., Chlebo, R., Coffey, M. F., Cornelissen, B., da Costa, C. A., Csáki, T., Dahle, B., Danihlík, J., Dražić, M. M., Evans, G., Fedoriak, M., Forsythe, I., de Graaf, D., Gregorc, A., Johannesen, J., Kauko, L., Kristiansen, P., Martikkala, M., Martín-Hernández, R., Medina-Flores, C. A., Mutinelli, F., Patalano, S., Petrov, P., Raudmets, A., Ryzhikov, V. A., Simon-Delso, N., Stevanovic, J., Topolska, G., Uzunov, A., Vejsnaes, F., Williams, A., Zammit-Mangion, M., and Soroker, V. (2019). Loss rates of honey bee colonies during winter 2017/18 in 36 countries participating in the coloss survey, including effects of forage sources. *Journal of Apicultural Research*, 58(4):479–485.

Holmes, W. (2002). The influence of weather on annual yields of honey. *The Journal of Agricultural Science*, 139(1):95–102.

Klein, A. M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., and Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings. Biological sciences*, 274(1608):303–313.

Kulyukin, V. A., Coster, D., Kulyukin, A. V., Meikle, W., and Weiss, M. (2024). Discrete time series forecasting of hive weight, in-hive temperature, and hive entrance traffic in non-invasive monitoring of managed honey bee colonies: part i. *Sensors*, 24(19):6433.

Langowska, A., Zawilak, M., Sparks, T. H., Tomkins, P. W., and Tryjanowski, P. (2017). Long-term effect of temperature on honey yield and honeybee phenology. *Int J Biometeorol*, 61:1125–1132.

Le Conte, Y. and Navajas, M. (2008). Climate change: impact on honey bee populations and diseases. *Revue Scientifique et Technique-Office International des Epizooties*, 27(2):499–510.

Masaka, K. (2023). Yearly fluctuations in honey production in Hokkaido, Northern Japan, with special reference to weather conditions and masting behavior. *International Journal on Food, Agriculture, and Natural Resources*, 4(3):62–68.

McNeil, A., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools*. Princeton University Press.

Meikle, W. G., Holst, N., Colin, T., Weiss, M., Carroll, M. J., McFrederick, Q. S., and Barron, A. B. (2018). Using within-day hive weight changes to measure environmental effects on honey bee colonies. *PLOS ONE*, 13(5).

Mezentsev, S. (2024). Nomadic beekeeping as an effective way to physically protect and preserve the lives of honey bees from pesticides used in agriculture. In *AIP Conference Proceedings*. AIP Publishing LLC.

Norrström, N., Niklasson, M., and Leidenberger, S. (2021). Winter weight loss of different subspecies of honey bee *Apis mellifera* colonies (Linnaeus, 1758) in southwestern Sweden. *PLOS ONE*, 16(10).

O'Connell, D. P., Baker, B. M., Atauri, D., and Jones, J. C. (2024). Increasing temperature and time in glasshouses increases honey bee activity and affects internal brood conditions. *Journal of Insect Physiology*, 155:104635.

Petanidou, T. and Smets, E. (1996). Does temperature stress induce nectar secretion in mediterranean plants? *New Phytologist*, 133:513–518.

Porras, M. F., Navas, C. A., Agudelo-Cantero, G. A., Santiago-Martínez, M. G., Loeschcke, V., Sørensen, J. G., Crandall, S. G., Biddinger, D., and Rajotte, E. G. (2023). Extreme heat alters the performance of hosts and pathogen. *Frontiers in Ecology and Evolution*, 11.

Porrini, C., Mutinelli, F., Bortolotti, L., Granato, A., Laurenson, L., Roberts, K., Gallina, A., Silvester, N., Medrzycki, P., Renzi, T., Sgolastra, F., and Lodesani, M. (2016). The status of honey bee health in Italy: Results from the nationwide bee monitoring network. *PloS one*, 11(5):e0155411.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., and Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution*, 25(6):345–353.

Sahin Demirel, A. N. (2025). Investigating the impact of climate variables on the organic honey yield in Turkey using xgboost machine learning. *Journal of the Science of Food and Agriculture*, 105(1):84–92.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Senger, D., Gruber, C., Kluss, T., and Johannsen, C. (2024). Weight, temperature and humidity sensor data of honey bee colonies in Germany, 2019–2022. *Data in Brief*, 52:110015.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris*, 8:229–231.

Solovev, V. V. (2020). Influence of weather conditions on the honey productivity of bee colonies in the Valdai district of the Novgorod region. *IOP Conference Series: Earth and Environmental Science*, 613(1):12142.

Stabentheiner, A., Kovac, H., and Brodschneider, R. (2010). Honeybee colony thermoregulation - regulatory mechanisms and contribution of individuals. *PlosOne*, 5(1).

Switanek, M., Crailsheim, K., Truhetz, H., and Brodschneider, R. (2017). Modelling seasonal effects of temperature and precipitation on honey bee winter mortality in a temperate climate. *Science of the Total Environment*, 579:1581–1587.

Takkis, K., Tscheulin, T., Tsalkatis, P., and Petanidou, T. (2015). Climate change reduces nectar secretion in two common Mediterranean plants. *AoB Plants*, 7.

Tassinari, W., Lorenzon, M., and Peixoto, E. (2013). Spatial regression methods to evaluate beekeeping production in the state of Rio de Janeiro. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 65:553–558.

Tscharntke, T., Clough, Y., Wanger, T. C., Jackson, L., Motzke, I., Perfecto, I., Vandermeer, J., and Whitbread, A. (2012). Global food security, biodiversity conservation and the future of agricultural intensification. *Biological Conservation*, 151(1):53–59.

van Engelsdorp, D. and Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *Journal of Invertebrate Pathology*, 103:S80–S95.

Van Espen, M., Williams, J. H., Alves, F., Hung, Y., de Graaf, D. C., and Verbeke, W. (2023). Beekeeping in europe facing climate change: A mixed methods study on perceived impacts and the need to adapt according to stakeholders and beekeepers. *Science of the Total Environment*, 888:164255.

Vincze, C., Leelőssy, Á., Zajácz, E., and Mészáros, R. (2025). A review of short-term weather impacts on honey production. *International Journal of Biometeorology*, 69(2):303–317.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244.

Zhao, H., Li, G., Guo, D., Li, H., Liu, Q., Xu, B., and Guo, X. (2021). Response mechanisms to heat stress in bees. *Apidologie*, 52(2):388–399.

Zhu, Y., Abdollahi, M., Maucourt, S., Coallier, N., Guimarães, H. R., Giovenazzo, P., and Falk, T. H. (2024). Mspb: a longitudinal multi-sensor dataset with phenotypic trait measurements from honey bees. *Scientific Data*, 11(1):860.