Computers and Electronics in Agriculture 239 (2025) 110830

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Prediction of hop cone ripening through Internet of Things (IoT) and Machine Learning (ML) technologies

Giulia Oddi [a,1], Martina Galaverni [b,1], Laura Belli [a,*], Margherita Rodolfi [b,*], Luca Davoli [a], Gianluigi Ferrari [a,1], Tommaso Ganino [b,c,1]

[a] Internet of Things (IoT) Lab, Department of Engineering and Architecture, University of Parma, Parma, Italy
[b] Crop and Plant Science (Cro.P.S.) Lab, Department of Food and Drug, University of Parma, Parma, Italy
[c] National Research Council, Institute of BioEconomy (IBE), Sesto Fiorentino, Italy

ARTICLE INFO

Keywords:
Bitter acids
Internet of Things
Machine learning
Ripening prediction
Smart agriculture

ABSTRACT

Hop (*Humulus lupulus* L.) cones ripening is characterized by a gradual increase of the valuable brewing metabolites, namely bitter acids and essential oils (EOs), thus making the identification of the optimal harvest time pivotal to obtain high quality yields and avoid economical losses. Cone ripeness is currently evaluated visually: Smart Agriculture (SA) technologies, including the Internet of Things (IoT) paradigm and Machine Learning (ML) models, are expected to have a significant impact on it. In this work, IoT devices are employed to collect data in the time period 2021–2023 at the "Azienda Agricola Ludovico Lucchi" hop testbed located in Campogalliano, Modena, Italy. Two ML-based algorithms are proposed to forecast the optimal harvesting period: the *first* relies on Multiple Linear Regression (MLR), while the *second* exploits Principal Component Regression (PCR). Finally, both algorithms classify ripening stages (namely: *immature*, *mature*, *overripe*) through a soft voting classifier. To this end, the identification of the optimal ripening time required the hop cones to be morphologically and chemically characterized (approximately) weekly for three growing seasons. Our results indicated that during the first half of September, there was a contraction in cone width and an increase in the EOs content, representing the optimal harvest maturity. Finally, the proposed ML models forecasted the optimal harvesting period for the 2024 season in the same days and this was confirmed in the reality. The correspondence between predictions and analytical results highlights the potential of integrated IoT and ML techniques to provide decision support for farmers and to improve agricultural operations.

1. Introduction

Hop (*Humulus lupulus* L.) is a climbing, perennial and dioecious plant whose female inflorescences, usually called "hop cones", are widely used in the brewing industry because of the presence of a yellowish resinous substance, named *lupulin*, rich in secondary metabolites. Hop resins and essential oils (EOs), providing bitterness and aroma to the beer, respectively, are the most valuable brewing compounds, with the *first* ones representing the highest concentration of all active ingredients in the cones (Skomra and Koziara-Ciupa, 2020).

To this end, the bitter acids (5–20% of the cone dry weight) are the most important components of the hop resins and can be distinguished into α-acids and β-acids (Zanoli and Zavatti, 2008). The α-acids (*humulones*), consisting of a mixture of six *humulone* homologues (mainly Humulone, Cohumulone and Adhumulone), most affect the price and quality of hops because of their isomerization capacity during the wort-boiling due to high temperature and pH values (Zanoli and Zavatti, 2008; Rybka et al., 2017). In fact, the iso-α-acids are more water soluble and bitter compared to the not-isomerized form and, therefore, are the main responsible of the beer bitterness (Zanoli and Zavatti, 2008; Intelmann and Hofmann, 2010; Karabín et al., 2016). Instead, the β-acids (*lupulones*), unable to isomerize because of structural differences, contribute less to the bitter taste of beer, but exert a greater antimicrobial activity (Schönberger and Kostelecky, 2011; Almaguer et al., 2014; Krofta and Mikyska, 2014). The hop EOs (0.5–3% of the cone dry weight) are a complex mixture of over 1000 volatile compounds that, based on their concentration and ratios, can impart herbal, citrusy, fruity or the typical hoppy aromas (Schönberger and Kostelecky, 2011; Pearson et al., 2016; Raut et al., 2021).

The biosynthesis and the accumulation of bitter acids and EOs start with cone development and continue with significant variations

* Corresponding authors.
E-mail addresses: laura.belli@unipr.it (L. Belli), margherita.rodolfi@unipr.it (M. Rodolfi).
[1] The authors contributed equally to the study.

https://doi.org/10.1016/j.compag.2025.110830
Received 25 March 2025; Received in revised form 28 July 2025; Accepted 29 July 2025
Available online 21 August 2025

during cone maturation (Bailey et al., 2009). For instance, Skinner et al. (1974) observed a gradual increase in oil and bitter acids during the maturation process and similar findings have been confirmed by recent literature (Bailey et al., 2009, Sharp et al. 2014, Matsui et al. 2016, Lafontaine et al. 2019). Therefore, identifying the optimal ripening time is fundamental to achieve high-quality production. Avoiding both early and late harvesting is crucial: early harvesting prevents bitter acids and essential oils from reaching their ideal concentrations, while late harvesting may lead to the degradation of aroma compounds, discoloration, and an increased risk of exposure to pests and diseases, ultimately resulting in reduced storability (Darby et al., 2017, Skomra and Koziara-Ciupa 2020). Besides, although hop harvesting is now mostly performed using picking machines, it must be completed within a 2-week window during which quality parameters are stable: this process may be complicated by the simultaneous ripening of other varieties (Bailey et al., 2009, Sirrine et al. 2010).

In this context, the Smart Agriculture (SA) concept could offer valuable, precise, and efficient tools to assist farmers and could optimize processes, such as harvest, through the use, integration, and development of advanced technologies, such as Internet of Things (IoT) and Machine Learning (ML). In particular, the term IoT identifies a heterogeneous set of technologies enabling real-world objects to communicate with each other and with the internet, gathering data from the environment they are deployed in and processing them according to specific tasks and applications (Atzori et al., 2010). Therefore, IoT systems are beneficial for SA thanks to the integration of different types of smart devices, including environmental, soil and plant sensors. On the other side, ML algorithms include various methods to automatically identify patterns within large amounts of data, in order to reach different goals, including forecasting the future values of monitored data or performing decision-making tasks under uncertainty systems (Murphy, 2012). ML algorithms can then be used, alongside IoT technologies, in different application scenarios, including SA, thus allowing farmers to address challenges such as farm productivity, environmental impact and sustainability.

With particular focus on SA, it is well-known that its main objective is to improve and increase both quality and quantity of crop yields, as described by Kassim (2020). In fact, as highlighted by Ozdogan et al. (2017), SA facilitates to monitor and analyze a wide range of parameters, related to environmental factors, crop production status, cultivation health status, and soil conditions. This approach brings various advantages in agriculture, including increased crop yields, reduced costs, improved product quality, and optimized processes, thus helping farmers to face challenges of climate changes. Finally, the deployment of the IoT technologies in SA can help farmers to objectively and precisely identify the ripeness time of hop cones and, more generally, of other types of crops.

In SA scenarios, predicting fruits ripening is crucial to reach quality and sustainability goals, including waste reduction through supply chain and optimization, as highlighted by Li et al. (2018). In fact, in-field assessment of fruit ripeness, combined with accurate prediction of the optimal harvesting time using non-destructive technologies, can revolutionize traditional farming practices by optimizing harvesting operations and ensuring consumers to receive highest-quality products. In recent years, several studies have focused on the prediction of the maturation stages of various fruits, especially relying on ML and Deep Learning (DL) models. In the literature, these approaches, combined with image processing techniques, exploit the detection of the fruit color and shape to predict the maturation stages. Then, by integrating ML-based models, farmers can benefit from reliable and data-driven models also in the harvesting context. As an example, a specific Neural Network (NN) — namely, VGG-19 Convolutional Neural Network (CNN) — is proposed in Ramos et al. (2021) to allow a non-invasive grape classification based on ripening stages, in detail relying on images acquired from traditional optical cameras as input information. Moreover, Pereira et al. (2018) present a study about papaya maturation, exploiting fruit images as input for a Random Forest (RF) model and enabling (as a result) accurate classification of the three different maturation stages. Heterogeneous types of input could be used to predict maturation stages. As an example, dos Santos Costa et al. (2019) propose a PCA-based approach for the classification of reflectance spectroscopy information of wine grapes, in detail considering three different regression models — namely, Principal Component Regression (PCR) and MLR — and various classifiers, then enabling the discrimination of different maturation stages in wine grapes. Nevertheless, in the literature very limited applications of smart technologies on hop cultivation have been reported, basically for the evaluation of the crop status and growth through satellite images and Unmanned Aerial Vehicles (UAVs) (Kumhálová et al., 2021; Agehara et al., 2024; Řeřicha et al., 2025). Furthermore, recent evidence has shown that hop classification can be achieved using DL approaches (Castro et al., 2022). Additionally, plant disease assessment can be performed through (i) CNN architectures as well as (ii) Binary Particle Swarm Optimization (BPSO) feature selection combined with SVM (Degadwala et al., 2023; Farhanah and Al Maki, 2022). Although smart technologies have been employed to monitor hop growth throughout the growing season, the identification of the ideal hop harvest time, which is fundamental for achieving high yields and quality, still remains an open issue. To the best of our knowledge, no models — whether ML- or PCA-based — have been proposed for predicting hop ripeness.

In the context outlined above, the objective of this work is to develop a ML-based prediction model of hop cone ripening. Then, by integrating IoT technologies with the analysis of bitter acids and EOs in a real testbed, located in Campogalliano, Modena, Italy, a more comprehensive understanding of plants growth is achieved. This is crucial to ensure high-quality yields and to obtain the specific bitter acids and aromatic profiles that positively influence beer quality. The proposed approach aims to ease the harvesting operations, reducing hop cones' wasting and maximizing the final production quality and quantity. Hence, to support these goals, a cloud-based IoT data management platform, denoted as *Agriware* (Oddi et al., 2024; Galaverni et al., 2025), is employed to collect and monitor data from different devices deployed in the experimental Italian crop testbed, thus enabling to reach different goals: (i) gather different crop-related parameters; (ii) evaluate (in real-time) agronomic indicators, such as the Normal Heat Hours (NHH); (iii) apply ML models to implement algorithms able to analyze and generate data, as well as to forecast the optimal harvesting period.

More in detail, the main objectives of the proposed work can be summarized as follow.

1. Weekly morphological and chemical characterization of hop cones.
2. Integration and monitoring of crop-related data through IoT technologies and the *Agriware* platform.
3. Definition of algorithms for the optimal harvesting time prediction for a hop cultivation.
4. Harvesting support by reducing hop cones wasting and maximizing the final production quality and quantity.

The rest of the paper is organized as follows. Section 2 presents a description of materials and methods applied to the developed testbed, including a detailed description of the experimental crop setup, the hop sampling, the data analysis and the implemented algorithms. Section 3 discusses the obtained experimental results, providing a comprehensive analysis. Finally, in Section 4 we draw conclusions and outline potential future research directions.

2. Materials and methods

The proposed work is organized as a *chain* of consecutive experimentation steps (as shown in Fig. 1 for the sake of clarity), aiming at identifying the optimal hop cones harvesting period. Specifically,

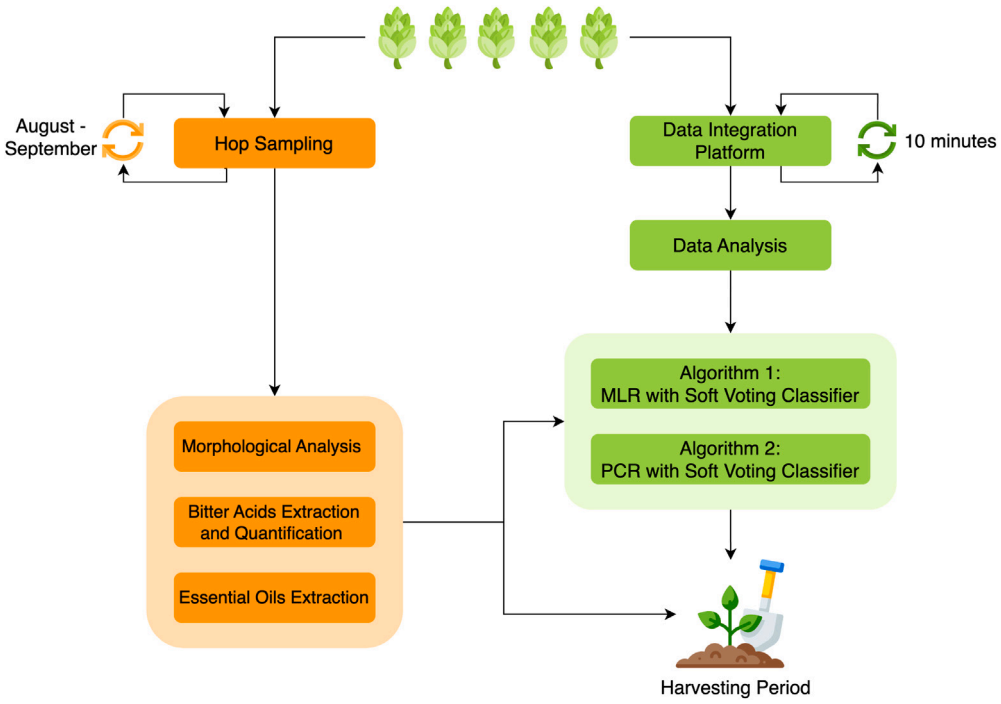


Fig. 1. Experimental steps representing the proposed hop harvesting period forecasting mechanism.

Table 1
Chemical and physical properties of the experimental field.

Chemical property	Value
Organic matter [%]	3.1
Total nitrogen [g kg^{-1}]	2.0
Assimilable Phosphorus (P_2O_5)	40.9
Exchangeable Calcium (Ca) [mg kg^{-1}]	5850
Exchangeable Potassium (K) [mg kg^{-1}]	382
Physical property	**Value**
Coarse sand [%]	1.1
Fine sand [%]	13.0
Coarse silt [%]	14.4
Fine silt [%]	35.0
Clay [%]	36.5

Table 2
IoT devices used at "Azienda Agricola Ludovico Lucchi" before 2023.

IoT device	Number of devices	Measured parameters	Dimension
Weather station	1	Air temperature	°C
		Air moisture	%RH
		Air pressure	kPa
		Dew point	°C
		Rain	mm
		Evapotranspiration	mm
		Solar Radiation	Wm^{-2}
		Wind speed	m/s
		Wind angle	°
Sensory unit	2	Soil moisture surface	%RH
		Soil moisture depth	%RH
		Soil temperature surface	°C
		Soil temperature depth	°C
		Foliar wetness	mV

from the experimental crop (described in Section 2.1), hop cones were collected with different sampling between August and September in a few consecutive years (2021–2024), as described in Section 2.2. Then, as detailed in Section 2.3, the cones were analyzed to characterize the behavior of bitter acids content and EO yields. At the same time, as described in Section 2.4, data coming from the IoT sensors deployed in the crop were integrated into a platform, with a 10 min sampling interval, and then proper agronomic indicators have been calculated (as detailed in Section 2.5). Finally, as described in Section 2.7, two different ML-based algorithms, based on MLR and PCR, and a soft voting classifier were implemented to predict the optimal hop harvest time, using as input the most correlated data.

2.1. Experimental crop setup

The experimental activities have been carried out in the time period 2021–2024 at the "Azienda Agricola Ludovico Lucchi" (with overall dimension of 4.89 ha) located in Campogalliano, Modena (Emilia-Romagna Region, Italy) (44° 42′ 19.9″ N, 10° 50′ 26.1″ E, 33.39 m a.s.l.) on hop plants cv. (*Cascade*). More in detail, the trial has been performed on a loamy clay soil whose chemical and physical properties are summarized in Table 1.

Then, the hop crop at the "Azienda Agricola Ludovico Lucchi" has been monitored during different time periods through several IoT devices. Specifically, before 2023, the testbed was managed using a Netsens AgriSense IoT weather station (Netsens, 2025) and two sensory units. In detail, the weather station provides measurements of different environmental parameters, including air temperature, air humidity and solar radiation, as described in Table 2, while the sensory units consist of Frequency Domain Reflectometry (FDR) probes, placed in representative areas of the hop crop to measure soil-related parameters, including soil temperature, soil moisture and foliar wetness (as detailed in Table 2).

Then, starting from 2023, a new set of IoT devices, managed by the commercial platform (XFarm, 2025), was introduced. In detail, *XFarm* is a platform (available also as a mobile App) that uses crop-related data to assist farmers to monitor the impact of their activities, improving practices and increasing positive contribution to the environment, through regenerative agriculture and sustainability. Consequently, the data in the time period 2023–2024 have been gathered and managed through this platform, exploiting an IoT network composed by a weather station and 14 IoT sensors, including tensiometers, volumetric

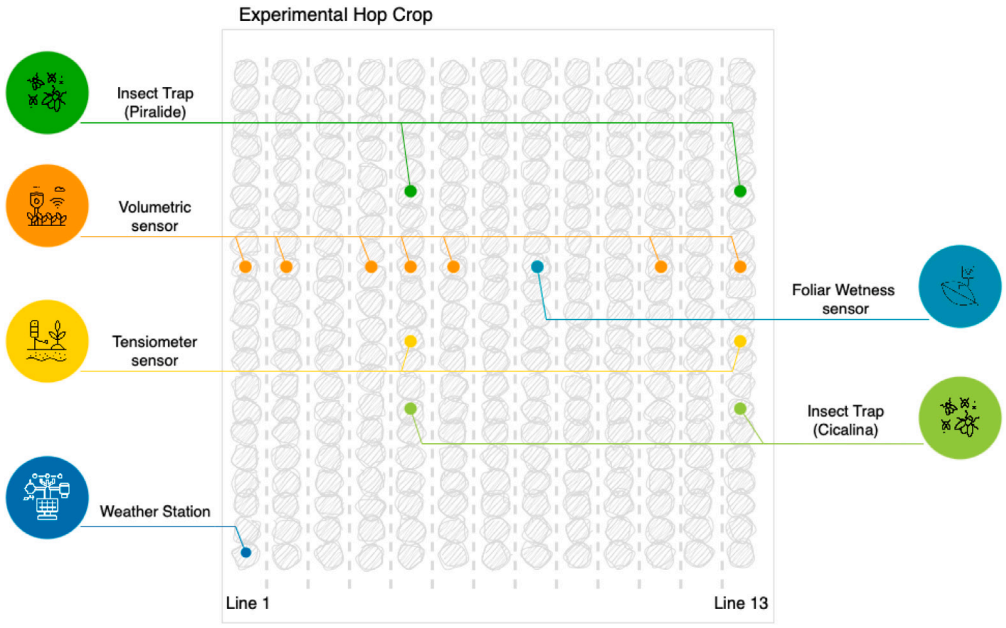


Fig. 2. Experimental *Cascade* hop crop setup with the 14 IoT devices position at "Azienda Agricola Ludovico Lucchi".

Table 3
IoT devices managed (starting from the year 2023) by the XFarm platform at "Azienda Agricola Ludovico Lucchi.".

IoT device	Number of devices	Measured parameters	Dimension
X Sense PRO weather station	1	Air moisture Air temperature Wind speed Wind angle	%RH °C m/s °
XNode Hydro tensiometer sensor	2	Soil temperature surface Soil temperature depth Water potential	°C °C kPa
XNode Soil Pro volumetric sensor	7	Soil moisture surface Soil moisture depth	%RH %RH
XLeaf foliar wetness sensor	1	Leaf wetness Hours leaf wetness	mV h
XTrap automatic insects trap	4	Image of trapped insects	–

sensors, foliar sensors and insect traps, as detailed in Table 3 and in Fig. 2. Within this setup, the weather station acts as a network gateway, aggregating data from other IoT devices and forwarding these data back to the *XFarm* cloud platform via the Long Range Wide Area Network (LoRaWAN) protocol (LoRa Alliance, 2025). In particular, LoRaWAN supports effective long-range communications and provides the ability to efficiently operate even in the presence of physical obstacles (e.g., underground), as shown in Aldhaheri et al. (2024). For this reason, LoRaWAN-enabled devices prove to be particularly beneficial for SA applications, allowing long-term monitoring and managing soil, crop, and air conditions.

2.2. Hop sampling

Hop cone samples were collected in the time period 2021–2024 at the "Azienda Agricola Ludovico Lucchi", Modena, Italy. More in detail, for each sampling, approximately 250 g of fresh hop cones have been collected for morphological and chemical analyses, then being sampled at different degrees of ripeness. For the sake of clarity, the number of samples and their dates are detailed in Table 4. Specifically, sampling in 2021 followed the BBCH scale for hops, beginning at growth stage 7 (BBCH code 71: "initial cone development") and ending st growth stage

Table 4
Time schedule of hop samplings in the time period 2021–2024.

Year	# samplings	August	September	October
2021	6	18 – 23 – 26	3 – 8 – 15	
2022	9	8 – 17 – 26	2 – 8 – 15 – 21 – 27	10
2023	3	24	8 – 15	
2024	1		15	

9 (BBCH code 92: "senescence, cone overripeness") (Rossbauer et al., 1995). The 2022 sampling campaign replicated this protocol, targeting the same developmental stages. Instead, in 2023 three samplings were performed in approximately BBCH codes 81, 89, and 92, corresponding to unripe, ripe and overripe cones (Fig. SM1), respectively. Finally, in 2024 hop cones were collected at optimal ripeness (BBCH code 89) (Rossbauer et al., 1995).

2.3. Hop cones analysis

All the samples collected in the four years of the experiment were characterized both morphologically — in terms of cone size — and chemically—in terms of bitter acids and EOs yield. After morphological analyses, the hop cones were dried until they reached a humidity of 12% and then stored at −20 °C for further analyses. In the following, each hop cone characterization will be detailed.

2.3.1. Hop cones morphological measurements

The hop cones morphological measurements (obtained from approximately 20 hop cones), at each sampling, were randomly chosen and manually measured directly in the field (on plants) in terms of width and length using a caliper.

2.3.2. Bitter acids extraction and quantification through HPLC-UV

The extraction of the bitter acids was conducted as described by Galaverni et al. (2024). The quantification of α- and β-acids was performed using a High Performance Liquid Chromatography (HPLC) system equipped with a pump, on-line vacuum degasser, auto-sampler, Peltier column oven, UV–Vis detector (PerkinElmer Series 200, Shelton, USA) and autosampler (PerkinElmer series 220, Shelton, USA). Then, chromatographic data were analyzed with a PerkinElmer Total Chrome workstation (version 6.3.1.). HPLC was equipped with Luna C18:2 column (5 µm, 100 A, 250 mm × 4.6 mm) (Phenomenex®, Castel Maggiore,

Bologna, Italy). For the mobile phase, solvent A (water (H_2O) + 0.1% orthophosphoric acid (H_3PO_4) and solvent B (methanol (CH_3OH) + 0.1% H_3PO_4) were employed. The chromatographic conditions used were the following: flow rate: 1.5 mL min^{-1} in isocratic, column temperature: 30 °C; injection volume: 10 μL; eluents: A (5%) and B (95%) for 15 min. Chromatograms were acquired at 314 nm. Three injections from three independent extractions were conducted for each sample. For the bitter acids quantification, a calibration curve was obtained from dilution of ICE-4 standard, in line with the official method (Analytica-EBC, method 7.7).

2.3.3. Essential oils extraction

EOs were extracted by steam distillation for 4 h with a Clevenger type apparatus to determine the oil content. Then, the EO yield was calculated as the ratio between the weight of the extracted oil and the dry weight of the used cones (w/w, dimension: [%]).

2.4. Data integration platform

In addition to the brief detail on the *XFarm* platform provided in Section 2.1, it should be highlighted that the platform also provides a Web dashboard for monitoring and controlling various crops, including the *Cascade* hop crop at the "Azienda Agricola Ludovico Lucchi", as well as providing different types of information, obtained directly from the sensors deployed in the crop itself. This information includes real-time IoT sensors data, insects' images, and crop conditions.

Moreover, the platform gives access to real-time data through specific HTTP APIs — available to all registered users — and enables the retrieval of historical data collected over a specified period. These APIs can be used to establish a connection with the platform and to retrieve information (in *JSON* format) related to the farm (in general) as well as to some specifically installed sensors. For completeness, the APIs operate through the following four distinct steps, each handled by a specific *HTTP request* and a dedicated *endpoint*.

1. **`POST` Login**: the *Login API endpoint* allows registered users or external softwares to sign in into the platform and obtain the *bearer token*, needed to authenticate the subsequent API calls. This call must be a `POST` HTTP request and requires valid *username–password* credentials as data payload. Then, the endpoint can return, as HTTP response, an *error* (`401 UNAUTHORIZED`, `403 FORBIDDEN`, `404 NOT FOUND`) or a *success* (`200 OK`).
2. **`GET` Own Farm**: the *Farms API endpoint*, that must be a `GET` HTTP request, allows to access the information on the farm associated with the authenticated user and to retrieve the corresponding `farms_id`.
3. **`GET` Device List**: using the `farms_id` obtained in the previous step, the *Equipment API endpoint* can be invoked to obtain a list of all the devices (denoted as "equipments") associated with the selected farm, where each device will be identified by a unique `device_id`. The remote call must be a `GET` HTTP request and the success response includes the list of available devices.
4. **`POST` Sensor Data**: once the list of the available devices is obtained, it is possible to call the *Telemetry API endpoint*, by specifying the `farm_id` and the `device_id`. This API endpoint allows users to specify a valid range of time and the HTTP response contains the list of all parameters monitored by the selected sensor.

Then, the heterogeneous data collected from IoT devices and from the *XFarm* platform have been integrated into a general data acquisition platform, denoted as *Agriware*. More specifically, *Agriware* targets: (i) the integration of information from heterogeneous SA data sources, in order to achieve the management and monitoring of different crops; (ii) the definition of custom software units for data processing, in order to handle data and generate new information streams; and (iii) the monitoring of the managed testbeds and parameters (as detailed in Tables 2 and 3, respectively) collected by IoT sensors. Finally, all this information is presented to external entities, commonly referred to as *Consumers*—for example, the Agriware Web application presents the results of interest to the farmer through the support of reports and charts.

The ingestion of data in the *Agriware* platform relies on a set of software modules, denoted as *Connectors*. The main purpose of a Connector is to allow the integration of a data stream from a specific IoT data source into the overall system. The specific structure and the implementation of a Connector depend on the technologies used to generate and transmit data.

Within the scope of this work, a specific Connector has been implemented to integrate the data from different data sources managed by the *XFarm* platform and, consequently, from the IoT devices deployed in the testbed in Campogalliano, Modena. More in detail, the XFarm Connector, which has been implemented in *Python* and is based on the HTTP communication protocol having to interact through the *XFarm* APIs, targets to (i) periodically invoke the provided APIs, (ii) handle the authentication process, as access is restricted to registered users only, and (iii) integrate IoT sensor-related data into the *Agriware* platform. To this end, the following two principal operations are periodically performed, through interactions with XFarm APIs and corresponding HTTP requests, by the XFarm Connector.

1. **Login**: this operation is required to manage the authentication process and, subsequently, to login into the XFarm *Login* API to retrieve the *bearer token*. As mentioned before, the *bearer token* is required for the `GET` *Own Farm* and the `GET` *Device List* API invocations.
2. **Obtain sensor data**: once the list of devices has been obtained, this function is used to add or update the entities related to data sources. Consequently, the `GET` *Sensor Data* API is invoked to retrieve and store the obtained sensor data.

The *XFarm* Connector is designed to be scalable and effective, ensuring simplified management of IoT devices. This implies that, whenever a new sensor is installed, the *Connector* can integrate its data into the *Agriware* platform, without requiring internal modifications or updates.

Finally, as mentioned before, all the sensors communicate through the LoRaWAN protocol, transmitting different types of data (e.g., leaf wetness, air moisture, their DevEUI, etc.) in the message body. In order to manage these different parameters, a specific *descriptor* is assigned to each device type. The main objective of a descriptor is to establish a mapping between the sensor-measured parameters and the internal *ontology* of the *Agriware* middleware. All the values reported by IoT sensors are mapped into predefined ontology types, each associated with a data-related label, ensuring efficient data management.

2.5. Data analysis

As mentioned before, this work focuses on the analysis of the different testbed data conducted to understand the behavior of the *Cascade* hop crop. In detail, hop cone measurements (namely, length and width), bitter acids content (α-acids and β-acids), and EO yield have been analyzed using statistical methods. Then, proper agronomic indicators have been calculated, including the NHH curve, the Growing Degree Days (GDD), and the Heat Units (HUs). Finally, correlations between agronomic parameters, agronomic indicators, and data collected from IoT devices have been evaluated and analyzed: this is expedient to highlight how external environmental factors can affect crop growth, development and reproduction.

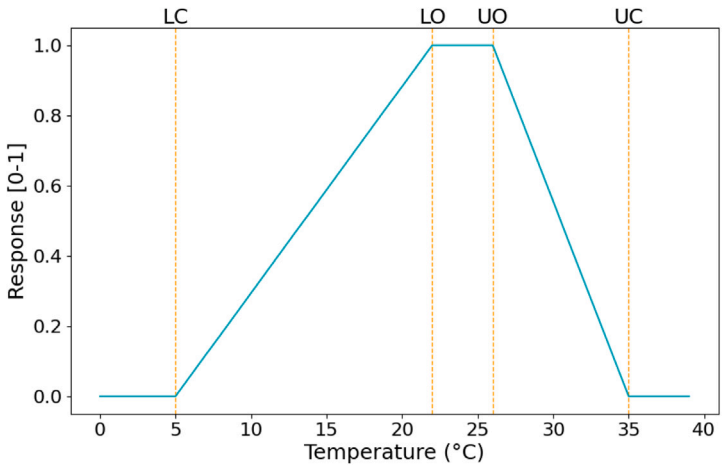


Fig. 3. NHH curve calculated for the *Cascade* hop crop. In particular, LC is set to 5 °C, LO to 22 °C, UO to 26 °C and UC to 35 °C.

2.5.1. Statistical analysis

Hop cone measurements, bitter acids content, and the EO yield have been evaluated through the XLSTAT software (Lumivero, 2025). All data have been analyzed in terms of mean ± standard error, while the normal distribution of the data and the variance homogeneity have been checked in accordance to Shaphiro and Wilk (1965) and Levene (Olkin, 1960).

2.5.2. NHH and GDD index

The NHH curve, calculated with the method described by Ferrante and Mariani (2018), is a crop-specific agronomic graphical indicator mapping average temperatures into a normalized value within the $0 \div 1$ range, and useful to understand the impact of temperatures above or below the optimum range for plant growth. With regard to the experimental evaluation, the calculated NHH curve for the *Cascade* hop cultivation is shown in Fig. 3.

In detail, the NHH curve returns that adequate growth and development only happen in a specific thermal range, delimited by the following two temperatures.

1. **Lower Cardinal** (LC) temperature: this temperature represents the *vegetation zero point*, also referred to as *base temperature* (denoted as T_{base} the following), which is specific to the cultivation being considered and represents the minimum biological temperature at which plants continue their vegetative activities. In particular, for the *Cascade* hop cultivation, T_{base} is set to 5 °C.
2. **Upper Cardinal** (UC) temperature: this temperature refers to the *cutoff temperature* (denoted as T_{cutoff}) and represents the temperature above which plants cannot grow. For the *Cascade* hop cultivation, T_{cutoff} is set to 35 °C.

Moreover, inside the described range, it is possible to identify a sub-optimal temperature range within which plants' growth and development occur without thermal limitation. Therefore, the response of the NHH curve in this range is set equal to 1, indicating that the plant can grow without thermal stress, making the conditions ideal for its development. This range is delimited by the following two temperatures.

1. **Lower Optimal** (LO) temperature: this temperature (denoted as T_{LO}) represents the minimum temperature for optimal growth of plants. Below this temperature, growth begins to slow down. For the *Cascade* hop cultivation, the lower optimal temperature is set to 22 °C.
2. **Upper Optimal** (UO) temperature: this temperature (denoted as T_{UO}) represents the optimal maximum temperature for plants' growth, while above this temperature the growth is reduced because of thermal stress. With regard to the *Cascade* hop cultivation, the UO temperature is set to 26 °C.

The values of the plant-related NHH curve have been calculated using the algorithm defined in Eq. (1), where T_{avg} represents the *average temperature*:

$$\text{NHH} = \begin{cases} 0 & \text{if} \quad T_{\text{avg}} \le T_{\text{base}} \quad \text{or} \quad T_{\text{avg}} \ge T_{\text{cutoff}} \\ 1 & \text{if} \quad T_{\text{avg}} \ge T_{\text{LO}} \quad \text{and} \quad T_{\text{avg}} \le T_{\text{UO}} \\ m \cdot T_{\text{avg}} + q & \text{if} \quad T_{\text{avg}} \ge T_{\text{base}} \quad \text{and} \quad T_{\text{avg}} \le T_{\text{LO}} \quad \text{where:} \\ & \quad m = \dfrac{1}{(T_{\text{LO}} - T_{\text{base}})} \quad \text{and} \\ & \quad q = 1 - \dfrac{T_{\text{LO}}}{(T_{\text{LO}} - T_{\text{base}})} \\ m \cdot T_{\text{avg}} + q & \text{if} \quad T_{\text{avg}} \ge T_{\text{UO}} \quad \text{and} \quad T_{\text{avg}} < T_{\text{base}} \quad \text{where:} \\ & \quad m = -\dfrac{1}{(T_{\text{cutoff}} - T_{\text{UO}})} \quad \text{and} \\ & \quad q = 1 + \dfrac{T_{\text{UO}}}{(T_{\text{cutoff}} - T_{\text{UO}})} . \end{cases} \tag{1}$$

The GDD, defined by Derscheid and Lytle (1981), is an agronomic indicator primarily used to describe the timing of biological processes. In detail, it corresponds to a measure of heat accumulation and is used in agriculture to predict plant development stages (Saadi et al., 2015). In particular, the GDD of day j, where j represents a generic day, so that $(j-1)$ represent its predecessor, is obtained by adding the GDD of the prior day $(j-1)$ according to the following equation:

$$\text{GDD}^{(j)} = \text{GDD}^{(j-1)} + \begin{cases} (T_{\text{avg}} - T_{\text{base}}) & \text{if} \quad T_{\text{avg}} > T_{\text{base}} \text{ and } T_{\text{avg}} < T_{\text{cutoff}} \\ (T_{\text{cutoff}} - T_{\text{base}}) & \text{if} \quad T_{\text{avg}} \ge T_{\text{cutoff}} \\ 0 & \text{if} \quad T_{\text{avg}} \le T_{\text{base}} . \end{cases} \tag{2}$$

where the daily average temperature (T_{avg}) and the *vegetation zero point* (T_{base}) are used, given the fact that the GDD is a daily agronomic indicator derived directly from the air temperature. In fact, the air temperature has a direct effect on the crop geographical distribution and growth.

HUs, defined in Machado et al. (2004), are valuable tools for predicting harvesting dates and determining the timing of successive plantings. More in detail, HUs are cumulative over time and are calculated daily using six different methods, introduced by Machado et al. (2004). These methods take into account, in addition to T_{cutoff} and T_{base}, two supplementary variables, such as (i) the daily maximum air temperature, denoted as T_{max}, and (ii) the daily minimum air temperature, denoted as T_{min}. The adopted methods, detailed in the

following, calculate the accumulation of six different types of HU, each of them directly derived from the air temperature.

1. Standard Day-Degrees: $(T_{max} + T_{min})/2 - T_{base}$.
2. Daily mean temperature: $(T_{max} + T_{min})/2$.
3. Daily maximum temperature above T_{base}: $T_{max} - T_{base}$.
4. Daily maximum temperature: T_{max}.
5. Daily maximum temperature above T_{base} with reduction of T_{cutoff} for plant development when T_{max} is higher then T_{cutoff}:

$$\text{if } T_x \leq T_{cutoff} : \quad T_x - T_{base} \tag{3}$$

$$\text{else if } T_{max} > T_{cutoff} : \quad T_{cutoff} - (T_{max} - T_{cutoff}) - T_{base} . \tag{4}$$

6. Ontario Units: $(T_a + T_b)/2$ where:

$$T_a = 3.33 \cdot (T_{max} - 10) - 0.084 \cdot (T_x - 10)^2$$
$$T_b = 1.8 \cdot (T_{min} - 4.4) .$$

2.6. Optimal harvesting time prediction

Building upon the previous analyses, the goal of the proposed work is the prediction of the maturation and, consequently, of the optimal harvesting period for the analyzed *Cascade* hop crop. To this end, two different ML models are considered: (i) the combined use of a MLR model with a soft voting classifier, and (ii) the use of a PCA method (Pearson, 1901) in the implementation of a PCR model. Both algorithms are implemented to (i) forecast parameters (or principal components) of the sensors and (ii) classify the predicted values into three predefined classes—namely, *immature*, *mature*, *overripe*. For the sake of clarity, it should be highlighted that MLR and PCR were preferred for this specific task, with respect to other algorithms—e.g., Artificial Neural Networks (ANN) or k-Nearest Neighbors (k-NN) thanks to their suitability to small datasets, such as the one available in the current work.

2.6.1. Multiple linear regression (MLR)

The first ML method employed is MLR (Mashaly and Alazba, 2016), a linear regression model using regression coefficients to determine the relationship between inputs and outputs, assuming the resistance of a linear relationship between them. This method, described by Eq. (5), optimizes regression coefficients by minimizing prediction error and is capable of simultaneously predicting multiple target variables (Ehteram and Banadkooki, 2023):

$$Y = \beta_0 + \beta_1 \cdot X_1 + \cdots + \beta_n \cdot X_n \tag{5}$$

where $X_1, \ldots, X_n$ are the input variables and $\beta_0, \ldots, \beta_n$ are regression coefficients.

MLR is used to forecast input features for the following day. Then, using the *iterated strategy* (Chevillon, 2007), the predictions are extended over multiple days, conducting a prediction model by minimizing the squares' residuals. Specifically, the iterated strategy generates a 1-step-ahead prediction, using the predicted values iteratively as input for the same model to subsequently forecast the following points. This process recursively continues until the desired prediction horizon is reached, as described by Bao et al. (2014). The features' prediction for the subsequent day is carried out considering a temporal *loopback* period comprising the previous 4 days. The window size, i.e., the *loopback* period, was chosen by identify the highest accuracy. Specifically, the accuracy of the model increases as the loopback period grows from 1 to 4 days, while it decreases with longer loopback periods. In order to achieve this, a *feature derivation* technique was applied: it calculates features for the 4-day loopback period, enabling the creation of an *ad-hoc* dataset. As a result, for each feature to be predicted, the corresponding features from the previous days were also included as additional inputs. New columns are created by shifting existing features over a 4-day period, allowing the model to use past feature values as input to predict future values.

2.6.2. Principal component regression (PCR)

The second regression model adopted in this work is PCR, corresponding to the combination of MLR with PCA. In this regressor, clustering and selection of input data for MLR have been performed via PCA, a statistical dimensionality reduction method introduced in 1901 by Karl Pearson (Pearson, 1901). In detail, the PCA allows to transform the original dataset into a lower-dimensional dataset composed of uncorrelated Principal Components (PCs), as detailed by Radzol et al. (2014). Determining the optimal number of PCs is a crucial task, as the PCs must effectively allow to characterize all the available information. In fact, several PCA stopping rules have been proposed to determine the number of PCs to retain, including: Kaiser's stopping rule; Cattell's scree test; and the Cumulative Percent of Variance (CPV) method.

1. Kaiser's stopping rule, also known as Eigenvalue-One-Criterion (EOC), was presented by Kaiser (1960) and states that only PCs with eigenvalues greater than one should be considered for further analysis. Essentially, this means that a component is considered significant if it accounts for at least as much variance as the equivalent of an original variable.
2. Cattell's scree test, proposed by Cattell (1966), determines the number of PCs to retain by identifying the point in the variance graph where a sharp change, also referred to as *elbow*, occurs. The PCs corresponding to the elbow are considered significant.
3. The CPV method is a criterion that retains components whose CPV meets a designated threshold, which can vary depending on the context and application.

Therefore, PCR is used to predict the PCs' values with a one-step-ahead approach. Similarly to MLR, PCR is employed in our work to forecast input features for the following day; subsequently, the predictions are extended to cover the desired prediction temporal horizon, using an iterated strategy.

2.6.3. Soft voting classifier

The soft voting classifier is a *parallel ensemble method* operating on the principle that the predictions of multiple basic classifiers can be combined to achieve more reliable and accurate classification results. In particular, this type of voting classifier employs the *soft voting* technique, defined as (Raschka, 2023):

$$\mathcal{L} = \arg \max_{i=1,\ldots,C} \sum_{k=1}^{M} w_k p_{ik} \tag{6}$$

where: $\mathcal{L}$ corresponds to the predicted class label; M is the number of classifiers; C is the number of classes; w_k is the weight that can be assigned to the kth classifier; and p_{ik} corresponds to the probability of the ith class predicted by the kth classifier. Therefore, the soft voting classifier averages the class probabilities predicted by the considered classifiers and selects the class with the highest average probability (Manconi et al., 2022).

2.7. Optimal harvesting time prediction algorithms

Two optimal harvesting time prediction algorithms, based on the ML methods described in Section 2.6, were implemented. Each algorithm combines two main types of ML models: (i) a *regressor*, capable of predicting the input features; and (ii) a *supervised multi-class classifier*, able to decide the correct class (*immature*, *mature*, *overripe*) of the predicted features. For the *regression task*, data were divided in order to retrieve different datasets: data collected from 2019 to 2022 were used as the training set, and data from 2023 served as the test set. Finally, the trained algorithms were used to predict the optimal harvest period for the 2024 season. Instead, for the *classification task*, only the data collected during the hop-growing season (July–October period) were used, with data from 2020 to 2023 serving as the training set, while data from 2024 were used as the test set.

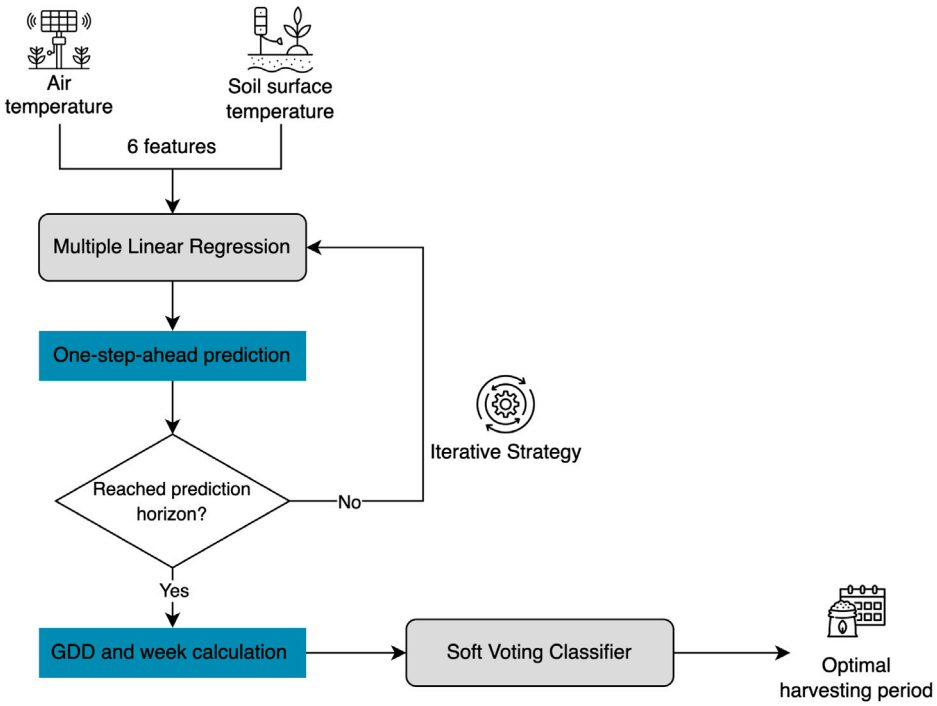


Fig. 4. Flowchart of the first proposed algorithm for the forecasting and the classification of the *Cascade* hop optimal harvesting period.

2.7.1. First prediction algorithm

The *first prediction algorithm* combines a MLR with a soft voting classifier. The workflow performed by this algorithm is shown in Fig. 4.

Initially, a MLR is implemented using, as input, only sensor data highly correlated to agronomic parameters and classes. These parameters consist of 6 different features, including air temperatures (namely: daily average, maximum, minimum) and soil surface temperatures (namely: daily average, maximum, minimum). Subsequently, the soft voting classifier uses, as input, the output of the iterated strategy applied to the predictions made by MLR. This input includes (i) data highly correlated to agronomic parameters and classes (such as air temperature, soil surface temperature, GDD values) and (ii) the week number (namely, week no.1, …, week no.52) of the year, as it resulted in the data analysis phase. To achieve this, it is necessary to calculate the GDD values and the corresponding week number for each predicted day. Finally, the output provides a prediction of the optimal harvesting period, corresponding to the days whose data were labeled as *mature*. More in detail, the MLR model was defined and implemented using the `LinearRegression` model provided by the `scikit-learn` Python package. Moreover, this model was trained using the features derived from the training set (2019–2022).

2.7.2. Second prediction algorithm

The second algorithm combines PCR with a soft voting classifier. The workflow performed by this algorithm is shown in Fig. 5.

In this approach, the initial input consists of all 42 features, including collected sensors data and calculated agronomic indicators. Consequently, a dimensionality reduction through PCA is required before running PCR: only the 4 most relevant PCs feed the subsequent PCR followed by a soft voting classifier. In detail, the most relevant PCs are detected by implementing the PCA `decomposition` offered by the `scikit-learn` Python package. The PCR model is implemented using the `LinearRegression` model provided by the `scikit-learn` Python package. The PCR operates using the iterated strategy to cover the entire prediction horizon. As with the first algorithm, the iterative forecasting strategy stops when the prediction horizon is reached. Then, the predicted PCs, together with the calculated GDD values and the week of the year, are used as input for the soft voting classifier. The final output provides a prediction of the optimal harvesting period, corresponding to the days whose data are labeled as *mature* by the soft voting classifier.

3. Results

3.1. Data correlations

The data collected from the IoT devices have been analyzed in order to highlight correlations among the monitored parameters. Table 5 depicts the values of the Pearson's correlation coefficient (denoted as ρ) between pairs of the parameters measured by IoT devices, while Fig. 6 shows the trend of correlations between pairs of parameters with a high ρ ($\rho = 0.89$) — namely, the soil surface temperature and the air temperature — measured throughout the 2023 hop season. Then, even if 2023 season is illustrative to highlight this correlation, the same data have been analyzed using the daily average data: the data have been interpolated, evaluating the corresponding angular coefficients and dispersions, as shown in Fig. 7. In Figs. 6 and 7, the colors indicate the elapsed time, with blue dots representing the earliest available data (from July 5, 2023) and red dots referring to the last available data (until October 31, 2023).

Subsequently, an additional data analysis has been performed to determine the strongest correlations between sensor data and agronomic parameters, as well as between calculated agronomic indicators and agronomic parameters. In fact, only the most correlated sensor data and agronomic indicators are considered as input for the forecasting algorithms. Tables 6 and 7 present the values of ρ for all pairs of data, respectively. The following agronomic parameters of hop cones, essential to understand the harvesting time, are considered in this analysis:

1. percentage of oil in the cone;
2. α-acids content (*Cohumulone* and *Adhumulone*);
3. β-acids content (*Colupulone* and *Adlupulone*);
4. cone humidity;

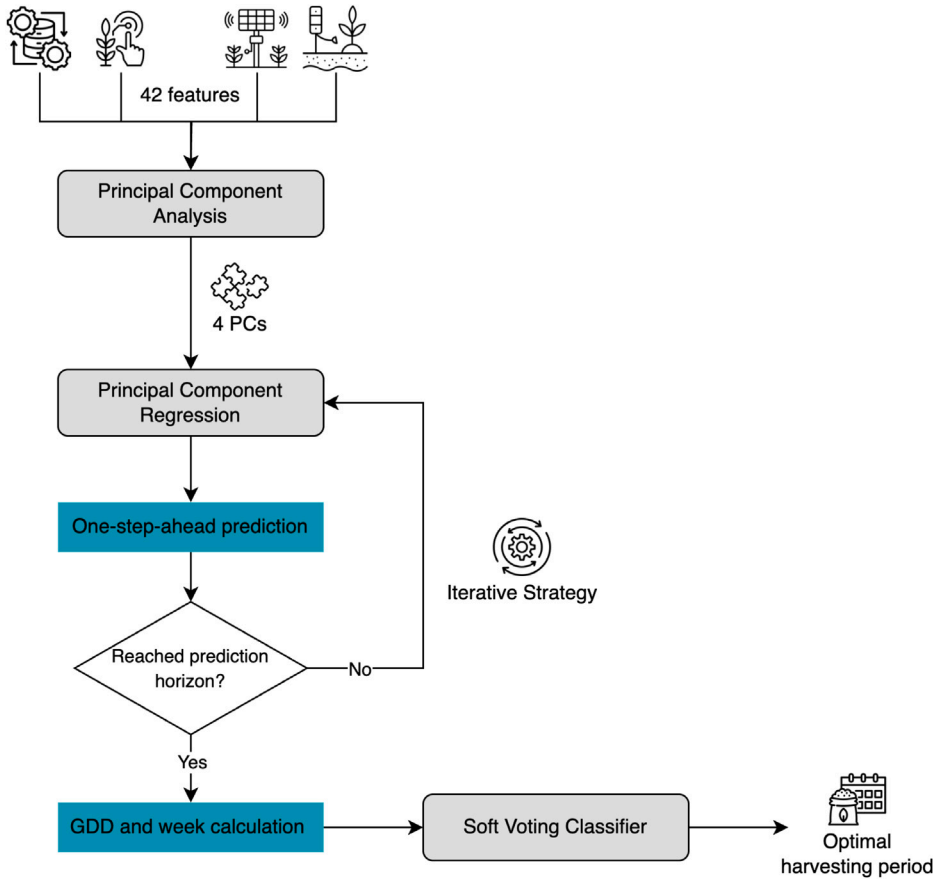


Fig. 5. Flowchart of the second proposed algorithm for the forecasting and the classification of the *Cascade* hop optimal harvesting period.

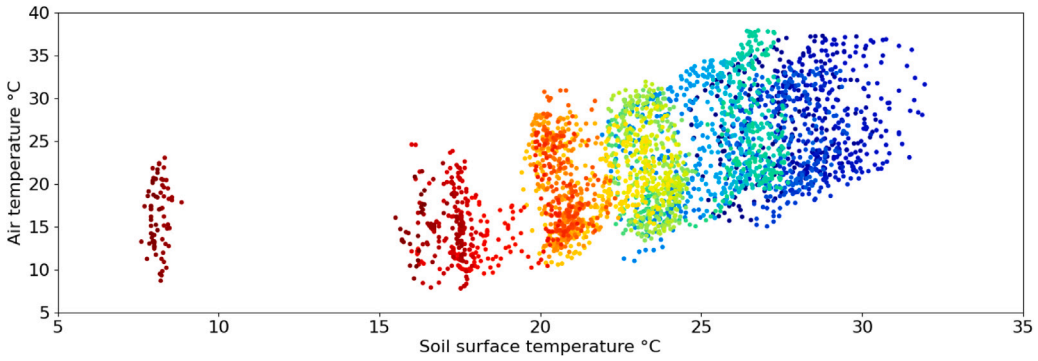


Fig. 6. Correlations representation during time between soil surface temperature (°C) and air temperature (°C) 10–minutes values for 2023 data ($\rho = 0.89$).

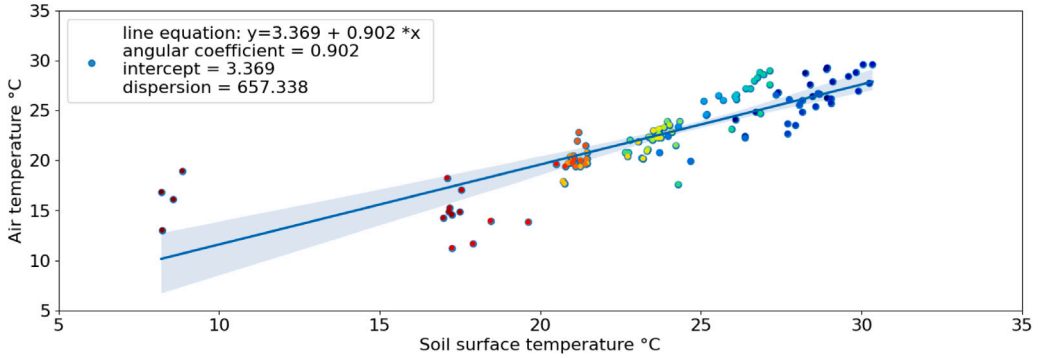


Fig. 7. Interpolation, with angular coefficient and dispersion parameters, between soil surface temperature (°C) and air temperature (°C) daily values for 2023 data ($\rho = 0.89$).

Table 5
Values of ρ between pairs of collected IoT sensors parameters.

Property	Air temperature	Air humidity	Air pressure	Dew point	Rain	Solar radiation	Wind	ETP	Soil humidity 1	Soil humidity 2	Soil temperature 1	Soil temperature 2
Air temperature	1	−0.191	−0.174	0.799	0.036	0.345	−0.637	0.468	0.122	0.032	0.891	0.877
Air humidity	−0.191	1	−0.158	0.431	0.596	−0.150	0.364	−0.209	−0.325	−0.392	−0.088	−0.050
Air pressure	−0.174	−0.158	1	−0.299	0.048	−0.077	0.074	−0.078	−0.013	0.081	−0.094	−0.142
Dew point	0.799	0.431	−0.299	1	0.384	0.193	−0.351	0.271	−0.090	−0.222	0.760	0.780
Rain	0.036	0.596	0.048	0.384	1	−0.200	0.428	−0.201	−0.213	−0.209	0.286	0.247
Solar radiation	0.345	−0.150	−0.077	0.193	−0.200	1	−0.230	0.988	−0.221	−0.238	0.359	0.418
Wind	−0.637	0.364	0.074	−0.351	0.428	−0.230	1	−0.298	−0.458	−0.456	−0.333	−0.269
ETP	0.468	−0.209	−0.078	0.271	−0.201	0.988	−0.298	1	−0.195	−0.221	0.463	0.515
Soil humidity 1	0.122	−0.325	−0.013	−0.090	−0.213	−0.221	−0.458	−0.195	1	0.943	0.005	−0.124
Soil humidity 2	0.032	−0.392	0.081	−0.222	−0.209	−0.238	−0.456	−0.221	0.943	1	−0.046	−0.214
Soil temperature 1	0.891	−0.088	−0.094	0.760	0.286	0.359	−0.333	0.463	0.005	−0.046	1	0.974
Soil temperature 2	0.877	−0.050	−0.142	0.780	0.247	0.418	−0.269	0.515	−0.124	−0.214	0.974	1

Table 6
Values of ρ for pairs of sensor (rows) and agronomic (columns) parameters.

Property	Oil%	Cohumulone	Adhumulone	α-acids	Colupulone	Adlupulone	β-acids	Cohumulone%	Humidity%	Length	Width
Air temperature	−0.6808	−0.0236	0.0873	0.0574	−0.2580	−0.0867	−0.1575	−0.5513	0.5365	−0.3714	−0.2663
Air humidity	0.0405	−0.4917	−0.5296	−0.5198	−0.0050	0.0583	0.0753	−0.0064	−0.2662	0.0715	0.0171
Air pressure	0.1800	0.0726	0.0651	0.0683	0.0040	−0.0582	−0.0977	0.0509	−0.2410	−0.0163	−0.1191
Dew point	−0.5480	−0.3295	−0.2547	−0.2753	−0.2387	−0.0452	−0.0938	−0.4962	0.3404	−0.2968	−0.2479
Rain	−0.4789	−0.3254	−0.3141	−0.3133	−0.1866	−0.0465	−0.1007	−0.2059	−0.0917	0.0089	−0.1505
Solar radiation	−0.3228	−0.0868	−0.0538	−0.0636	−0.3062	−0.2993	−0.2953	−0.1920	0.0615	−0.2014	−0.0140
Wind	0.5583	−0.1529	−0.2342	−0.2099	0.0891	0.0345	0.0721	0.3649	−0.6287	0.3595	0.1337
ETP	−0.4019	−0.0679	−0.0198	−0.0337	−0.3207	−0.2901	−0.2995	−0.2587	0.1312	−0.2304	−0.0376
Soil humidity 1	−0.2017	0.5418	0.6092	0.5929	0.4730	0.5003	0.4776	−0.0992	0.5174	−0.1490	−0.1415
Soil humidity 2	−0.2580	0.4866	0.5371	0.5252	0.3091	0.2863	0.2680	−0.0400	0.5528	−0.2855	−0.2408
Soil temperature 1	−0.8179	−0.0826	0.0099	−0.0138	−0.3718	−0.2135	−0.2682	−0.4820	0.4845	−0.4386	−0.4017
Soil temperature 2	−0.5588	−0.2033	−0.1159	−0.1390	−0.4220	−0.2527	−0.3040	−0.4822	0.4187	−0.4125	−0.3982

Table 7
Values of ρ between pairs of calculated agronomic indicators (rows) and agronomic parameters (columns).

Property	Oil%	Cohumulone	Adhumulone	α-acids	Colupulone	Adlupulone	β-acids	Cohumulone%	Humidity%	Length	Width
GDD	0.8507	−0.0401	−0.1319	−0.1069	0.3859	0.2927	0.3564	0.4635	−0.8576	0.6804	0.5803
Standard day degree	0.3747	−0.1543	−0.1915	−0.1796	−0.1507	−0.1560	−0.1234	0.1200	−0.5814	0.3166	0.2765
Daily mean temperature	0.3025	−0.1881	−0.2202	−0.2100	−0.1848	−0.1821	−0.1525	0.0853	−0.5481	0.2779	0.2474
Daily max above T_{base}	0.3195	−0.1892	−0.2228	−0.2122	−0.1760	−0.1747	−0.1443	0.0928	−0.5597	0.2892	0.2567
Daily max	0.2941	−0.1970	−0.2287	−0.2186	−0.1885	−0.1846	−0.1552	0.0807	−0.5458	0.2742	0.2450
Daily max reduction	0.3282	−0.1831	−0.2172	−0.2064	−0.1746	−0.1746	−0.1440	0.0964	−0.5630	0.2920	0.2584
Ontario units	0.3377	−0.1636	−0.1978	−0.1868	−0.1719	−0.1740	−0.1430	0.1020	−0.5611	0.2933	0.2585

5. cone growth, measured in terms of *length* and *width*.

The correlation analysis reveals that the most strongly correlated parameters, among those monitored by IoT sensors, are the air temperature (ρ = 0.68 with oil cone percentage) and the soil surface temperature (ρ = 0.82 with oil cone percentage), while the GDD is the most correlated indicator (ρ = 0.85 with both the oil percentage and cones humidity). Notably, air and soil surface temperature correlate well with the oil percentage, which is an indicator of hop aroma potential in beer and is known to be an increasing function of harvest date (Lafontaine et al., 2019). Additionally, a strong correlation is found between GDD and cone humidity, which, together with phytochemical compounds, is considered by the brewing industry when assessing hop maturation (Kavalier et al., 2011).

3.2. Ripening classes characterization

Following the data correlation analysis, a new dataset has been created to address the classification problem. Specifically, the information about the optimal harvesting period for the *Cascade* hop crop testbed, spanning from 2021 to 2023, has been exploited. Consequently, only the data collected during the hop growth period (July–October for each of the three years) were labeled. The labeling process enabled the separation of the data rows into three distinct classes, representing different stages of crop development, on the basis of the data generation timestamp and the date intervals shown in Table 8.

Table 8
Data split into three classes, according to the generation date (indicated as d).

Year	Immature	Mature	Overripe
2021	$d \leq$ August 23	August 24 $\leq d \leq$ September 5	$d \geq$ September 6
2022	$d \leq$ August 29	August 30 $\leq d \leq$ September 18	$d \geq$ September 19
2023	$d \leq$ September 4	September 5 $\leq d \leq$ September 12	$d \geq$ September 13

- **[0] Immature**: this class identifies the daily sensor information corresponding to the period when the hop crop was not yet ready for harvest. This period corresponds to the phenological stages related to the early development of the cones.
- **[1] Mature**: this class represents the daily sensor information generated and collected in the optimal harvesting time period, ensuring the production of the highest quality hop cones.
- **[2] Overripe**: this class captures the daily sensor data corresponding to the time period when it was too late to harvest the cones, resulting in a crop with lower quality compared to the previous stage.

The labeled dataset has been analyzed to identify correlations with each class, and to determine the most significant features associated

Table 9
Values of ρ between classes and atmospheric features.

Air temperature	Air humidity	Air pressure	Dew point	Rain	Solar radiation	Wind	ETP
−0.74139	0.32925	0.19493	−0.60605	−0.01827	−0.56524	0.34136	−0.57790

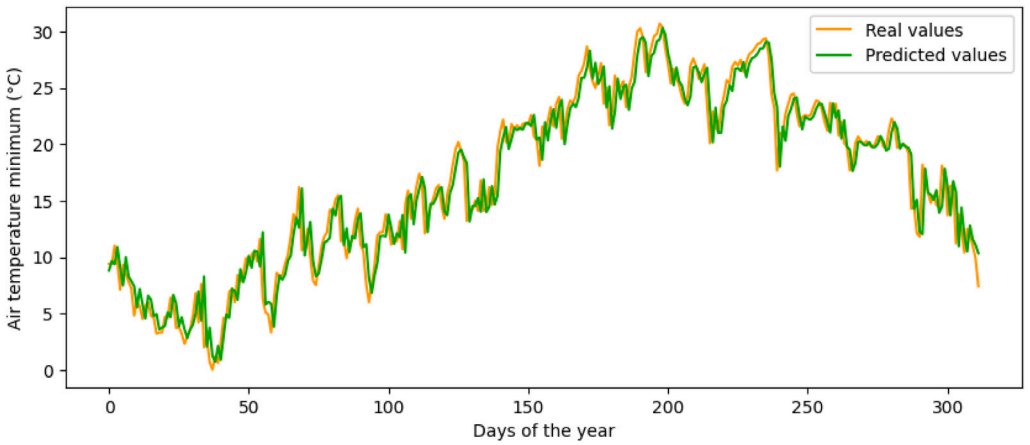


Fig. 8. Air average temperature values (in 2023) predicted by the MLR algorithm trained on most correlated features.

Table 10
Values of ρ between class division and soil features.

Soil surface humidity	Soil depth humidity	Soil surface temperature	Soil depth temperature
−0.08421	−0.22191	−0.70375	−0.62758

with the *mature* class. In fact, Table 9, Tables 10 and 11 present the values of ρ calculated between class labels and atmospheric related characteristics, soil-related characteristics and calculated agronomic indicators, respectively. These correlations highlight that the most significant features, with regard to class division, are: (i) air temperature ($\rho = -0.74$); (ii) soil surface temperature ($\rho = -0.70$); and (iii) GDD ($\rho = 0.72$). In addition, the relationship between the class division and the temporal indicator, represented by the *week* of the year, has been studied. To this end, by analyzing the available data related to the hop crop phenological stages, it is evident that hop plants, in recent years, have never been ready for harvesting before week no.34. In particular, the week becomes highly significant in identifying the optimal harvest time ($\rho = 0.82$ with class division).

3.3. Algorithms outputs

The performance of the MLR regressor, evaluated on the test set (2023), is summarized in Table 12. The evaluation metrics are Mean Absolute Error (MAE), Mean Squared Error (MSE), and Coefficient of Determination (R^2). The selected indicators are expedient to assess the model's accuracy and reliability in forecasting the features on the basis of the input data. In Fig. 8, the predicted values of the air average temperature feature are shown, comparing them with the true air temperature values: an excellent accuracy can be observed.

Moreover, for the first presented algorithm, the adopted soft voting classifier combines 5 different supervised ML classifiers, i.e., the ones that achieved an accuracy greater than 0,9: RF (0.929), Decision Tree (DT, 0.905), k-NN (0.929), Extra Trees (0.952), and Gradient Boosting (0.952). The soft voting classifier achieves an F1-score of 0.929 and an accuracy of 0.929 on the test set (20% of all dataset). The confusion matrix corresponding to this classifier is shown in Fig. 9(a).

For the second algorithm presented in this work, the PCs were investigated before implementing the PCR model to find out how many PCs need to be considered for the implementation of the PCR model. In particular, from the PCA, the Kaiser's stopping rule indicates that 25 PCs can be retained. However, Cattell's scree test was applied to further refine the selection and identify the most significant PCs. As shown in Fig. 10, the scree plot test highlights that 4 PCs can be considered as essential. Additionally, the CPV applied to our data, as shown in Fig. 11, reveals that these 4 PCs account for over 75% of the total variance in the dataset, while 95% of the variance is explained using 11 PCs. However, our experimental tests show that using 11 PCs degrades the accuracy—in general, increasing the number of PCs beyond 4 lowers the accuracy.

On the basis of these results, four PCs were considered for all the algorithms, as these components effectively summarize most of the information contained in all the features of the original dataset. The PCR model was trained on the training dataset associated with the years 2019–2022 and, then, tested on the test dataset associated with the year 2023. The performance metrics (in terms of MAE, MSE, and R^2) are shown in Table 13.

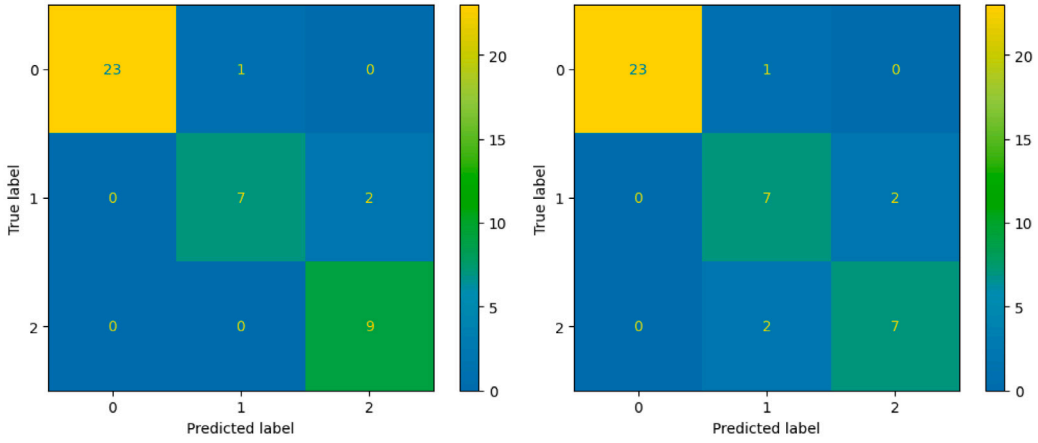
Moreover, in Fig. 12 the predicted values of the first PC in 2023 are shown and compared with the actual values. The soft voting classifier combines different supervised ML classifiers, specifically those achieving an accuracy higher than 0.9: RF (0.952), DT (0.905), Support Vector Classificator (0.905), Extra Trees (0.929), AdaBoost (0.929). The classifier achieves a final F1-score of 0.884 and an accuracy of 0.881 on the test set (20% of the dataset). The confusion matrix for this classifier is shown in Fig. 9(b).

3.4. Morphological analysis of Hop cones

Changes in the morphology of hop cones are shown in Table 14, which clearly shows that, during the period 2021–2022, there was a significant growth of cones, in terms of both length and width, between the first and third samplings. From that point onward, the size remained almost constant. Each hop variety is characterized by its specific morphological features, including cone size and shape (Roberts and Wilson, 2006). Since there is a direct relationship between cone development and the accumulation of secondary metabolites (Kavalier et al., 2011), cone length is an important feature to be considered when evaluating the degree of ripeness, as cone elongation ceases only at physiological maturity (Čeh et al., 2012). For instance, Kavalier et al. (2011) morphologically and chemically characterized hop cones of the Willamette and Zeus cultivars across five developmental stages, reporting a significant increase in cone volume and mass alongside a

Table 11
Values of ρ between classes and agronomic indicators.

GDD	Standard degree day	Daily mean	Daily max above T_{base}	Daily max	Daily max with reduction	Ontario units
0.72397	0.14851	0.10408	0.117012	0.09903	0.122426	0.12367



(a) Confusion matrix first soft voting classifier. (b) Confusion matrix second soft voting classifier.

Fig. 9. Confusion matrix of classification results on the test set (20% of the training set 2020–2023) made by the two proposed soft voting classifiers.

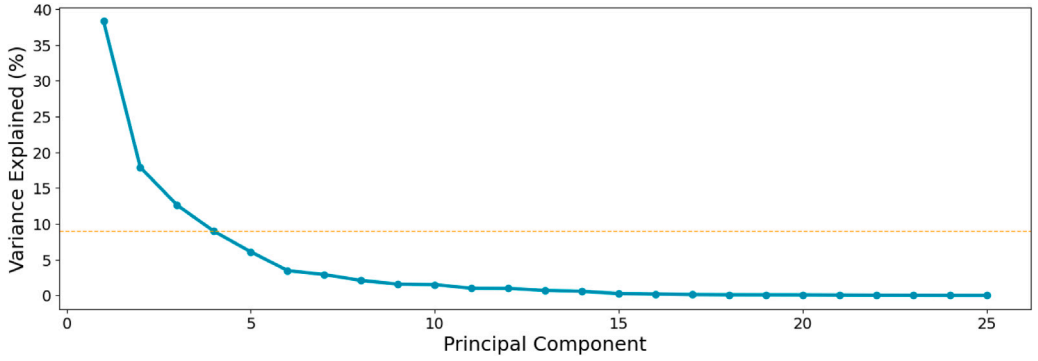


Fig. 10. Cattell's scree plot method applied with 25 PCs.

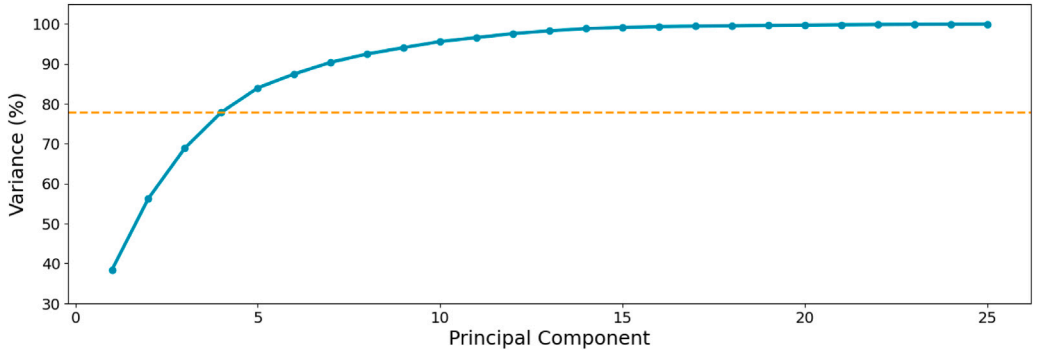


Fig. 11. Cumulative Percent of Variance (CPV) method applied with 25 PCs.

rise in the bitter acids content. Therefore, it is reasonable to assume that the first three samplings in 2021 and 2022 were carried out at the following three stages of hops' growth: during developing (BBCH 78–79); at the very beginning of the ripening process (BBCH 81–89); and, finally, at the optimal ripeness of the cone, once elongation had stopped (BBCH 89) (Rossbauer et al., 1995). In contrast, in 2023 no significant variations in cone length were observed across the three samplings, but for a decrease in width, generally occurs during the final ripening stages when cones become increasingly compact until they close and are ready for harvest (BBCH 89) (Rossbauer et al., 1995).

3.5. Chemical analyses of Hop cones

Besides morphological characterization, hop cones were analyzed also in term of bitter acid content, with values from each sampling time reported in Table 15. In 2021, no significant variations in the α-

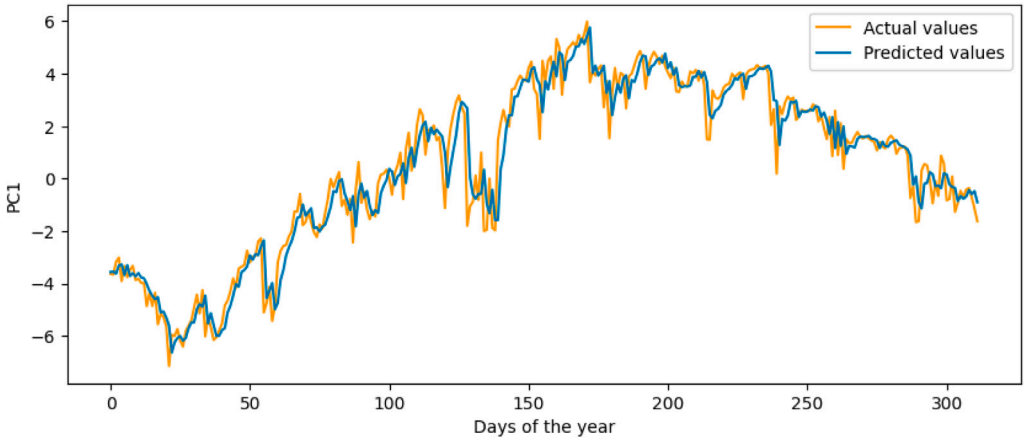


Fig. 12. Principal component (PC1) values prediction for test set (2023) by PCR.

Table 12
Performance (in terms of MAE, MSE and R^2) of MLR on the 6 parameters of the test set data (2023).

Parameter	MAE	MSE	R^2
Air average temperature	1.3498	3.1288	0.9477
Air maximum temperature	2.1733	8.0614	0.8980
Air minimum temperature	1.5197	3.6929	0.9251
Soil surface average temperature	0.2831	0.3653	0.9814
Soil surface maximum temperature	0.3618	0.4455	0.9778
Soil surface minimum temperature	0.2980	0.3829	0.9803

Table 13
Performance (in terms of MAE, MSE and R^2) of PCR applied on the 4 PCs identified in the test set data (2023).

Principal Component	MAE	MSE	R^2
PC1	0.5933	0.6965	0.9316
PC2	0.2524	0.1158	0.8166
PC3	0.7288	0.9731	0.5650
PC4	0.7592	1.2031	0.6954

Table 14
Morphological measurements of cones at different sampling times.

Year	Function	Day	Length [cm]	Width [cm]
2021	Training	08/18	0.76 ± 0.06^{c}	0.56 ± 0.05^{c}
		08/23	1.81 ± 0.10^{b}	1.02 ± 0.03^{b}
		08/26	2.34 ± 0.07^{a}	1.20 ± 0.03^{a}
		09/03	2.61 ± 0.17^{a}	1.19 ± 0.04^{a}
		09/08	2.57 ± 0.11^{a}	1.20 ± 0.04^{a}
		09/15	2.75 ± 0.10^{a}	1.16 ± 0.03^{a}
2022	Training	08/08	0.76 ± 0.06^{c}	0.76 ± 0.05^{d}
		08/17	1.85 ± 0.09^{b}	1.02 ± 0.03^{c}
		08/26	2.34 ± 0.07^{a}	1.21 ± 0.03^{b}
		09/02	2.61 ± 0.17^{a}	1.20 ± 0.04^{b}
		09/08	2.56 ± 0.10^{a}	1.20 ± 0.03^{b}
		09/15	2.75 ± 0.10^{a}	1.17 ± 0.03^{b}
		09/21	2.53 ± 0.09^{a}	1.15 ± 0.02^{bc}
		09/27	2.78 ± 0.08^{a}	1.43 ± 0.02^{a}
		10/05	2.45 ± 0.07^{a}	1.27 ± 0.03^{b}
2023	Test	08/24	2.35 ± 0.06^{a}	1.36 ± 0.02^{a}
		09/08	2.46 ± 0.09^{a}	1.12 ± 0.02^{b}
		09/15	2.63 ± 0.13^{a}	1.04 ± 0.02^{c}

The data are reported in terms of mean ± standard error. Different letters indicate statistically significant differences at $p < 0.05$ by Tukey's test.

and β-acids contents were observed throughout the trial, with values ranging from 6.25 to 6.62 ppm and 5.93 to 6.64 ppm, respectively. In 2022, a significant rise in bitter acids content was found during the first three samplings, especially between the first two, when growths of more than 250% and 296% for the α- and β-acids content, respectively, were observed. Such increases in both α- and β-acids are consistent with the cone size growth discussed above during the same period (Table 14), with similar findings reported by Kavalier et al. (2011). At the opposite, in 2023 a significant decline ($p < 0.0001$) in the α-acids content across the samplings was observed, unlike in the β-acids. The biosynthesis of α- and β-acids is affected by several external factors, with the environment playing a pivotal role in determining their concentration (Marceddu et al., 2022). Hot summer temperatures and low precipitations, as in summer 2023, have been ascribed as causes of reduction in the α-acids content in several studies (Mozny et al., 2009; Srečec et al., 2008; MacKinnon et al., 2020; Donner et al., 2020; Marceddu et al., 2022). Srečec et al. (2008) found negative ($r = -0.39$, $p < 0.05$) and positive ($r = 0.46$, $p < 0.05$) correlations between the average daily temperature and total rainfall with the production of α-acids, respectively. Similar negative ($r = -0.78$, $p < 0.01$) and positive ($r = 0.72$, $p < 0.01$) impacts of temperature and rainfall on the α-acids content were reported by MacKinnon et al. (2020); likewise, positive ($r = 0.59$ and 0.61) and negative (from −0.56 to −0.83) correlations between seasonal rainfalls and temperatures with the α-acids content were reported by Donner et al. (2020). However, despite the unfavorable weather conditions in summer 2023, the hops exhibited α-acid levels consistent with the average values for Italian hops in the same year (Barth Haas Hops Companion, 2024) and those typical of the *Cascade* variety (Haas, Inc., 2016). Typical α- and β-acid levels of the *Cascade* variety were also attained in 2021 and 2022, with slightly lower values in 2022, consistent with global trend (Barth Haas Hops Companion, 2023). Differences in the bitter acids content across different growing years in the same location are already reported by Matsui et al. (2016), who attribute these variations to weather conditions (temperature and rainfall).

As summarized in Table 16, the EO yields in all the experimental years (2021, 2022, and 2023) gradually increased during cone maturation and reached, every year, their maximum values at the last samplings. Based on the hops growth patterns and accumulation of secondary metabolites, the optimal harvest window of *Cascade* cones grown in Northern Italy is in the first half of September. Our concentrations of humulones and lupulones plateaued at the end of August and remained stable during the harvest window in agreement with Sharp et al. (2014). At the opposite, the oil content increased as a function of the harvest date as previously reported by several authors (Bailey et al., 2009; Sharp et al., 2014; Matsui et al., 2016; Lafontaine et al., 2019). In particular, Bailey et al. (2009) study the influence of hop harvest date on hop aroma and observe, on average, a 30% rise in the oil content in 24 days. Similarly, Matsui et al. (2016) observe the impact of the harvest timing on the amount of essential oil and its composition in different years and locations. Likewise, Lafontaine et al. (2019) verify the impact of harvesting time on the essential oil content,

Table 15
Bitter acids content at different sampling times.

Year	Function	Day	α-acids	β-acids
2021	Training	08/18	6.38 ± 0.49^{a}	5.93 ± 0.44^{a}
		08/23	6.33 ± 0.13^{a}	6.60 ± 0.47^{a}
		08/26	6.62 ± 0.13^{a}	6.62 ± 0.12^{a}
		09/03	6.40 ± 0.42^{a}	6.35 ± 0.48^{a}
		09/08	6.62 ± 0.04^{a}	6.41 ± 0.06^{a}
		09/15	6.25 ± 0.12^{a}	6.64 ± 0.13^{a}
2022	Training	08/08	0.97 ± 0.00^{c}	1.35 ± 0.06^{c}
		08/17	3.41 ± 0.35^{b}	5.35 ± 0.62^{b}
		08/26	5.55 ± 0.11^{a}	7.41 ± 0.17^{ab}
		09/02	5.72 ± 0.14^{a}	7.48 ± 0.22^{ab}
		09/08	5.18 ± 0.13^{a}	7.58 ± 0.20^{ab}
		09/15	5.30 ± 0.75^{a}	7.82 ± 1.13^{a}
		09/21	4.30 ± 0.35^{ab}	6.19 ± 0.50^{ab}
		09/27	5.06 ± 0.47^{a}	5.52 ± 0.47^{b}
		10/05	5.06 ± 0.86^{a}	6.54 ± 0.86^{ab}
2023	Test	08/24	7.22 ± 0.42^{a}	5.50 ± 0.42^{a}
		09/08	6.98 ± 0.14^{b}	5.62 ± 0.14^{a}
		09/15	3.79 ± 0.11^{c}	5.63 ± 0.11^{a}

α-acids and β-acids contents (*%w/w* on dry weight) are reported in terms of mean $\pm$ standard deviation. Different letters indicate statistically significant differences at $p < 0.05$ by Tukey's test.

Table 16
EOs yield at different samplings.

Year	Function	Day	EO yield [%]
2021	Training	08/18	0.38 ± 0.00^{e}
		08/23	0.82 ± 0.00^{d}
		08/26	1.39 ± 0.01^{c}
		09/03	1.89 ± 0.00^{b}
		09/08	1.95 ± 0.00^{ab}
		09/15	2.05 ± 0.08^{a}
2022	Training	08/08	0.21 ± 0.02^{d}
		08/17	0.34 ± 0.01^{d}
		08/26	1.20 ± 0.03^{c}
		09/02	1.73 ± 0.39^{bc}
		09/08	2.10 ± 0.00^{b}
		09/15	2.33 ± 0.11^{ab}
		09/21	2.48 ± 0.04^{ab}
		09/27	2.20 ± 0.28^{ab}
		10/05	3.00 ± 0.35^{a}
2023	Test	08/24	1.11 ± 0.03^{c}
		09/08	1.86 ± 0.05^{b}
		09/15	1.99 ± 0.04^{a}

The data are reported in terms of mean $\pm$ standard deviation. Different letters indicate statistically significant differences at $p < 0.05$ by Tukey's test.

reporting higher values at later harvest dates. However, despite the progressive increase of oils, it is crucial to avoid late harvests that might lead to decrease in cone yield and lupulin loss (Sharp et al., 2014), as well as to the formation of sulfur analytes (dimethyl disulfide, S-methylthioisovalerate, and S-methylthiohexanoate), responsible for negative onion-garlic notes (Kammhuber et al., 2017). Furthermore, late harvest timings seem to be related to reductions in the concentrations of monoterpene diglycosides, which are important non-volatile aroma precursors (Lafontaine et al., 2021).

3.6. Algorithms' harvesting period predictions

In order to test the ripeness prediction accuracy, the algorithms described in Section 2.7 were tested on the data of the 2024 season. In detail, the first algorithm was applied to predict the optimal harvesting period: following the 60-day-ahead predictions of air temperature (average, maximum, minimum) and soil surface temperature (average, maximum, minimum) generated by MLR, the corresponding GDD values and the relative weeks, for all the predicted data, were calculated. Then, all these data were used as input for the soft voting classifier. The predictions (from August 10, 2024) of the optimal harvest period in 2024 are shown in Fig. 13: the obtained results show that the period predicted by the first algorithm falls between September 2 and September 16, indicated in the plot by the green color.

Then, we also applied the second proposed algorithm, based on the PCA, to the 2024 data to find the optimal harvesting period. More in detail, following the 60-day-ahead predictions of the four PCs generated by the PCR, the predicted data have then been used as input for the soft voting classifier. The predicted optimal harvest period is shown in Fig. 14: it can be seen that the predicted optimal period falls between September 10 and September 25. In fact, during the 2024 season, the hop cones have been collected at their optimal maturity on September 15 when α- and β-acids contents were equal to $3.96 \pm 0.13\%$ and $3.87 \pm 0.16\%$, respectively, while the EO yield was equal to $1.20 \pm 0.04\%$. Hence, as expected, the results for the EO yield agrees with the typical values for the *Cascade* variety at ripeness (Haas, Inc., 2016), unlike the bitter acids content that is slightly lower.

Finally, the performance of the considered algorithms, in term of accuracy, will have to consider climate change effects, especially in the case of extreme weather events — e.g., heat and drought as well as torrential rain — which pose a serious challenge to the hop industry, as they negatively affect both yield and quality (Barth Haas Hops Companion, 2024). In fact, according to Mozny et al. (2023), hop yield and α-acids content are predicted to decline by 4–18% and 20–31%, respectively, by 2050, highlighting the urgent need for adaptive measures to safeguard the industry. Therefore, the assessment of the optimal cone ripeness becomes crucial to avoid further yield and quality losses. This further motivates the need to conduct analyses on the same variety under different climatic conditions and in different growing regions, which may affect crop production and quality (Acosta-Rangel et al., 2021; Rodolfi et al., 2019). Furthermore, the determination of optimal harvest time should be extended to other hop varieties, as they exhibit different maturation kinetics and, consequently, distinct harvest windows (Marceddu et al., 2020; Lafontaine et al., 2021).

4. Conclusions and future works

The approach proposed in this study aims at assessing the optimal hop harvesting period on the basis of real experimental data collected during the time period 2021–2023 at the "Azienda Agricola Ludovico Lucchi" in Campogalliano, Modena, Italy, mainly targeting to support farmers in achieving high yields of good quality.

More in detail, the proposed work considers two ML-based prediction algorithms: (i) a first algorithm, based on the MLR features (namely, air temperature and soil surface temperature) strongly correlated to the measured agronomic parameters; (ii) a second algorithm, based on PCA, which highlights that the input features' space can be reduced to only 4 PCs, still accounting for over 75% of the total variance. Then, both algorithms have been integrated with a (subsequent) soft voting classifier trained over three ripening classes (*immature*, *mature*, *overripe*) of hop growth period data. To this end, the cones were collected and, then, morphologically and chemically characterized approximately on a weekly basis for three growing seasons. Our results show a gradual increase in the EOs' content and a reduction in the cone width during the first half of September, which can thus be considered the best harvest window for the *Cascade* variety under our experimental conditions. Both ML-based prediction algorithms presented in this study forecasted the optimal harvesting period in the 2024 season in the same days and were confirmed by the bitter acids' content and EO yield. These results suggest the potential of AI-based prediction models in the identification of the optimal hop cone ripening period, providing a tool that can be used to support farmers in optimizing agricultural operations.

However, the proposed method is affected by some limitations. First, this analysis considers a single cultivar (*Cascade*), in a single field, with the same environmental conditions during the years. Moreover,

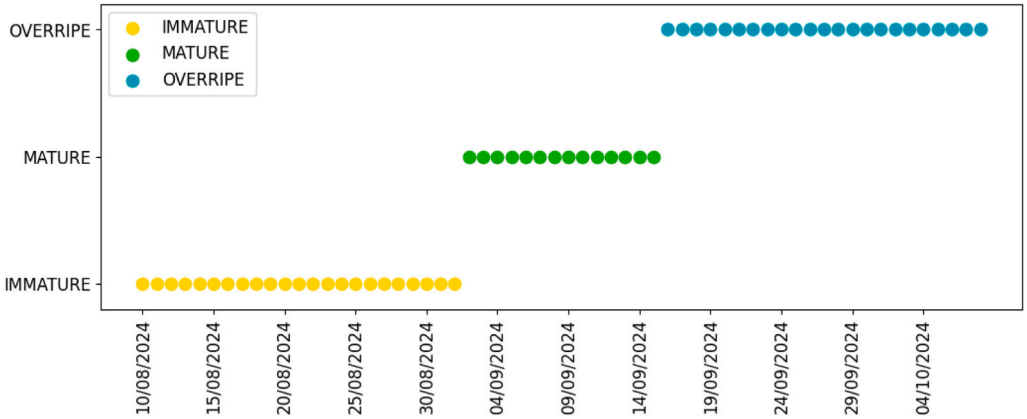


Fig. 13. Optimal harvesting period prediction by first proposed algorithm (MLR and soft voting classifier) from August 10, 2024.

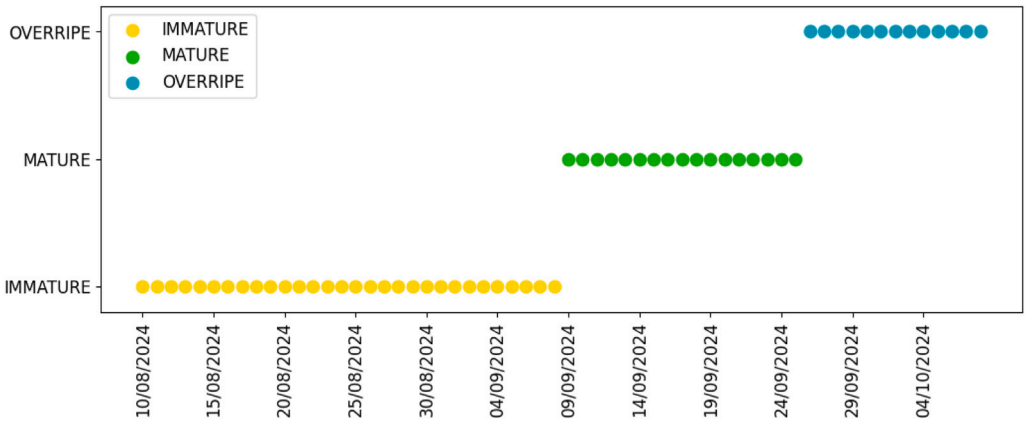


Fig. 14. 2024 optimal harvesting period prediction by second proposed algorithm (PCR and soft voting classifier) from August 10, 2024.

Fig. SM1. Hop cones in different maturity stages (*immature, mature, overripe*).

the proposed approach is only based on ML-based algorithms, such as MLR and PCR: this choice was motivated by the fact that the limited dataset collected and used was not suitable to train more complex DL models, such as ANN or *k*-NN. In the presence of richer datasets, the proposed approach could exploit DL models applied to different hop varieties (e.g., *Columbus, Comet, Chinook, Lotus*) and environmental conditions. As an example, finer and more frequent data sampling — e.g., taking measurements every hour, or even every 10 min, thanks to the employed IoT devices — could be used to implement new DL models.

Future research activities may include the collection and integration of new types of information to improve both prediction and classification tasks. Moreover, a larger number of classes — e.g., adding the *pre-mature* and *early post-mature* classes to the already considered classes — may be expedient to refine the classifiers and narrow the optimal harvesting period. Finally, the integration of images captured

by smart cameras may generate new information strictly correlated with plant growth and maturity, useful to estimate morphological parameters such as cone height and length. Another relevant activity that can be evaluated as a future research direction would be the definition of a novel Hop Harvesting Index (HHI), which may provide a comprehensive metric for the calculation of the harvesting period and hop cone maturation based on available IoT sensory information. The HHI could guide the farmers during harvesting operations.

CRediT authorship contribution statement

Giulia Oddi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Martina Galaverni:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laura Belli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Margherita Rodolfi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Luca Davoli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gianluigi Ferrari:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Tommaso Ganino:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This work is partially supported by the Agritech project - “National Research Centre for Agricultural Technologies”, project code CN00000022, funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender no. 3138 of 16/12/2021 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU, Concession Decree no. 1032 of 17/06/2022 adopted by the Italian Ministry of University and Research (MUR). This work is partially supported by the National Project “Luppolo, Orzo, Birra: biodiversità ITaliana da valorizzare (Acronim: LOB.IT)” and received funding from Ministry of Agriculture, Food Sovereignty and Forestry (MASAF) (DD n. 667550 - 30.12.2022). The work reflects only the authors’ views and opinions; neither the European Union nor the European Commission can be considered responsible for any use that may be made of the information it contains.

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.

Appendix. Supplementary materials

This section contains supplementary material that expands on the data and analyses discussed in the main text. These include an additional figure that visually illustrate the three utilized maturity stages of a hop cone, namely, immature, mature, overripe.

Data availability

Data will be made available on request.

References

Acosta-Rangel, A., Rechcigl, J., Bollin, S., Deng, Z., Agehara, S., 2021. Hop (’humulus lupulus’ l.) phenology, growth, and yield under subtropical climatic conditions: Effects of cultivars and crop management. Aust. J. Crop. Sci. 15 (5), 764–772.

Agehara, S., et al., 2024. Phenological assessment of hops (humulus lupulus l.) grown in semi-arid and subtropical climates through BBCH scale and a thermal-based growth model. Agronomy 14 (12), http://dx.doi.org/10.3390/agronomy14123045.

Aldhaheri, L., et al., 2024. Lora communication for agriculture 4.0: Opportunities, challenges, and future directions. http://dx.doi.org/10.48550/arXiv.2409.11200, arXiv:2409.11200. URL: https://arxiv.org/abs/2409.11200.

Almaguer, C., et al., 2014. Humulus lupulus–a story that begs to be told. a review. J. Inst. Brew. 120 (4), 289–314. http://dx.doi.org/10.1002/jib.160.

Atzori, L., Iera, A., Morabito, G., 2010. The internet of things: A survey. Comput. Netw. 54 (15), 2787–2805. http://dx.doi.org/10.1016/j.comnet.2010.05.010.

Bailey, B., et al., 2009. The influence of hop harvest date on hop aroma in dry-hopped beers. Tech. Q. Master Brew. Assoc. Am 46 (2), 1–7.

Bao, Y., Xiong, T., Hu, Z., 2014. Multi-step-ahead time series prediction using multiple-output support vector regression. Neurocomputing 129, 482–493. http://dx.doi.org/10.1016/j.neucom.2013.09.010.

Barth Haas Hops Companion, 2023. BarthHaas Report 2022/2023. Technical Report.

Barth Haas Hops Companion, 2024. BarthHaas Report 2023/2024. Technical Report.

Castro, P.H.N., Moreira, G.J.P., Luz, E.J.d.S., 2022. An end-to-end deep learning system for hop classification. IEEE Lat. Am. Trans. 20 (3), 430–442. http://dx.doi.org/10.1109/TLA.2022.9667141.

Cattell, R.B., 1966. The scree test for the number of factors. Multivar. Behav. Res. 1 (2), 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10.

Čeh, B., et al., 2012. Hmelj od sadike do storžkov. In: Žalec, vol. 15, Inštitut za hmeljarstvo in pivovarstvo Slovenije, pp. 65–69.

Chevillon, G., 2007. Direct multi-step estimation and forecasting. J. Econ. Surv. 21 (4), 746–785. http://dx.doi.org/10.2139/ssrn.914618.

Darby, H., Bruce, J., Lewins, S., 2017. Hop harvest timing. 58, URL: https://scholarworks.uvm.edu/nwcsp/58.

Degadwala, S., Vyas, D., Panesar, S., Ebenezer, D., Pandya, D.D., Shah, V.D., 2023. Revolutionizing hops plant disease classification: Harnessing the power of transfer learning. In: 2023 International Conference on Sustainable Communication Networks and Application. ICSCNA, pp. 1706–1711. http://dx.doi.org/10.1109/ICSCNA58489.2023.10370692.

Derscheid, L.A., Lytle, W.F., 1981. Growing degree days (GDD). SDSU Extension Fact Sheets.

Donner, P., et al., 2020. Influence of weather conditions, irrigation and plant age on yield and alpha-acids content of Czech hop (humulus lupulus l.) cultivars. Plant Soil Env. 66 (1), 41–46. http://dx.doi.org/10.17221/627/2019-PSE.

Ehteram, M., Banadkooki, F.B., 2023. A developed multiple linear regression (MLR) model for monthly groundwater level prediction. Water 15 (22), 3940. http://dx.doi.org/10.3390/w15223940.

Farhanah, A., Al Maki, W.F., 2022. Hops plants disease detection using feature selection based BPSO-SVM. In: 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics. EECSI, pp. 389–393. http://dx.doi.org/10.23919/EECSI56542.2022.9946620.

Ferrante, A., Mariani, L., 2018. Agronomic management for enhancing plant tolerance to abiotic stresses: High and low values of temperature, light intensity, and relative humidity. Horticulturae 4 (3), 21. http://dx.doi.org/10.3390/horticulturae4030021.

Galaverni, M., Ganino, T., Margherita, R., 2024. Evaluation of the stability of hop pellets subjected to different storage conditions during three years. J. Stored Prod. Res. 107, 102353. http://dx.doi.org/10.1016/j.jspr.2024.102353.

Galaverni, M., et al., 2025. An IoT-based data analysis system: A case study on tomato cultivation under different irrigation regimes. Comput. Electron. Agric. 229, 109660. http://dx.doi.org/10.1016/j.compag.2024.109660.

Haas, Inc., 2016. Barth-Haas Hops Companion: A Guide to the Varieties of Hops and Hop Products. Independent Publisher, URL: https://books.google.it/books?id=QvBfvgAACAAJ.

Intelmann, D., Hofmann, T., 2010. On the autoxidation of bitter-tasting iso-α-acids in beer. J. Agricult. Food Chem. 58 (8), 5059–5067. http://dx.doi.org/10.1021/jf100083e.

Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20 (1), 141–151. http://dx.doi.org/10.1177/001316446002000116.

Kammhuber, K., Hundhammer, M., Weihrauch, S., 2017. Influence of the date of harvest on the sulphur compounds of the" special flavour hop" varieties cascade, mandarina bavaria, hallertau blanc, huell melon and polaris. Brew. Sci. 70 (9–10), 124–130.

Karabín, M., et al., 2016. Biologically active compounds from hops and prospects for their use. Compr. Rev. Food Sci. Food Saf. 15 (3), 542–567. http://dx.doi.org/10.1111/1541-4337.12201.

Kassim, M.R.M., 2020. IoT applications in smart agriculture: Issues and challenges. In: 2020 IEEE Conference on Open Systems. ICOS, IEEE, pp. 19–24. http://dx.doi.org/10.1109/ICOS50156.2020.9293672.

Kavalier, A.R., et al., 2011. Phytochemical and morphological characterization of hop (humulus lupulus l.) cones over five developmental stages using high performance liquid chromatography coupled to time-of-flight mass spectrometry, ultrahigh performance liquid chromatography photodiode array detection, and light microscopy techniques. J. Agricult. Food Chem. 59 (9), 4783–4793. http://dx.doi.org/10.1021/jf1049084.

Krofta, K., Mikyska, A., 2014. Hop beta acids: Properties, significance and utilization. Kvas. Prumysl 60 (4), 96–105. http://dx.doi.org/10.18832/kp2014010.

Kumhálová, J., et al., 2021. Evaluation of UAV and sentinel 2 images to estimate condition of hop (humulus lupulus l.) plants. Acta Hortic. (1328), 95–102. http://dx.doi.org/10.17660/ActaHortic.2021.1328.13.

Lafontaine, S., et al., 2019. Impact of harvest maturity on the aroma characteristics and chemistry of cascade hops used for dry-hopping. Food Chem. 278, 228–239. http://dx.doi.org/10.1016/j.foodchem.2018.10.148.

Lafontaine, S., et al., 2021. Evaluation of variety, maturity, and farm on the concentrations of monoterpene diglycosides and hop volatile/nonvolatile composition in five humulus lupulus cultivars. J. Agricult. Food Chem. 69 (15), 4356–4370.

Li, B., Lecourt, J., Bishop, G., 2018. Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction—A review. Plants 7 (1), 3. http://dx.doi.org/10.3390/plants7010003.

LoRa Alliance, 2025. Lorawan specification v1.1. https://resources.lora-alliance.org/technical-specifications/lorawan-specification-v1-1. (Accessed 24 July 2025).

Lumivero, 2025. XLSTAT – statistical software for excel. https://www.xlstat.com/. (Accessed 24 July 2025).

Machado, R., et al., 2004. Prediction of optimal harvest date for processing tomato based on the accumulation of daily heat units over the fruit ripening period. J. Hortic. Sci. Biotechnol. 79 (3), 452–457. http://dx.doi.org/10.1080/14620316.2004.11511789.

MacKinnon, D., et al., 2020. The impact of weather conditions on alpha-acid content in hop (humulus lupulus l.) cv. Aurora. Plant Soil Environ. 66 (10), http://dx.doi.org/10.17221/344/2020-PSE.

Manconi, A., et al., 2022. A soft-voting ensemble classifier for detecting patients affected by COVID-19. Appl. Sci. 12 (15), 7554. http://dx.doi.org/10.3390/app12157554.

Marceddu, R., Carrubba, A., Sarno, M., 2020. Cultivation trials of hop (humulus lupulus l.) in semi-arid environments. Heliyon 6 (10).

Marceddu, R., Carrubba, A., Sarno, M., 2022. Resilience of hop (humulus lupulus l.) to salinity, heat and drought stresses: A mini-review. Front. Plant Sci. 13, 1064922. http://dx.doi.org/10.3389/fpls.2022.1064922.

Mashaly, A.F., Alazba, A., 2016. MLP and MLR models for instantaneous thermal efficiency prediction of solar still under hyper-arid environment. Comput. Electron. Agric. 122, 146–155. http://dx.doi.org/10.1016/j.compag.2016.01.030.

Matsui, H., et al., 2016. The influence of pruning and harvest timing on hop aroma, cone appearance, and yield. Food Chem. 202, 15–22. http://dx.doi.org/10.1016/j.foodchem.2016.01.058.

Mozny, M., et al., 2009. The impact of climate change on the yield and quality of saaz hops in the Czech Republic. Agricult. Forest. Meterol. 149 (6–7), 913–919. http://dx.doi.org/10.1016/j.agrformet.2009.02.006.

Mozny, M., et al., 2023. Climate-induced decline in the quality and quantity of European hops calls for immediate adaptation measures. Nat. Commun. 14 (1), 6028.

Murphy, K.P., 2012. Machine Learning: a Probabilistic Perspective. MIT Press.

Netsens, 2025. Netsens. https://www.netsens.it. (Accessed 24 July 2025).

Oddi, G., et al., 2024. Optimizing tomato production through IoT-based smart data collection and analysis. In: 2024 IEEE 20th International Conference on Automation Science and Engineering. CASE, pp. 45–50. http://dx.doi.org/10.1109/CASE59546.2024.10711738.

Olkin, I., 1960. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press.

Ozdogan, B., Gacar, A., Aktas, H., 2017. Digital agriculture practices in the context of agriculture 4.0. J. Econ. Financ. Account. 4 (2), 186–193. http://dx.doi.org/10.17261/Pressacademia.2017.448.

Pearson, K., 1901. LIII. On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin Philos. Mag. J. Sci. 2 (11), 559–572. http://dx.doi.org/10.1080/14786440109462720.

Pearson, B.J., Smith, R.M., Chen, J., 2016. Growth, strobile yield, and quality of four humulus lupulus varieties cultivated in a protected open-sided greenhouse structure. HortScience 51 (7), 838–842. http://dx.doi.org/10.21273/HORTSCI.51.7.838.

Pereira, L.F.S., Barbon, Jr., S., Valous, N.A., Barbin, D.F., 2018. Predicting the ripening of papaya fruit with digital imaging and random forests. Comput. Electron. Agric. 145, 76–82. http://dx.doi.org/10.1016/j.compag.2017.12.029.

Radzol, A., et al., 2014. Principal component analysis for detection of melamine adulterated milk products. In: 2014 IEEE Conference on Biomedical Engineering and Sciences. IECBES, IEEE, pp. 850–854. http://dx.doi.org/10.1109/IECBES.2014.7047631.

Ramos, R.P., et al., 2021. Non-invasive setup for grape maturation classification using deep learning. J. Sci. Food Agric. 101 (5), 2042–2051. http://dx.doi.org/10.1002/jsfa.10824.

Raschka, S., 2023. Ensemblevoteclassifier: A majority voting classifier. https://rasbt.github.io/mlxtend/user_guide/classifier/EnsembleVoteClassifier/. (Accessed 24 July 2025).

Raut, S., et al., 2021. Influence of pre-drying storage time on essential oil components in dried hops (humulus lupulus l.). J. Sci. Food Agric. 101 (6), 2247–2255. http://dx.doi.org/10.1002/jsfa.10844.

Řeřicha, J., Kohútek, M., Vandírková, V., Krofta, K., Kumhála, F., Kumhálová, J., 2025. Assessment of UAV imageries for estimating growth vitality, yield and quality of hop (humulus lupulus l.) crops. Remote. Sens. 17 (6), 970.

Roberts, T., Wilson, R., 2006. Hop analysis. Handbook Brewing, 2º ed. CRC Press, Florida, pp. 181–278.

Rodolfi, M., Chiancone, B., Liberatore, C.M., Fabbri, A., Cirlini, M., Ganino, T., 2019. Changes in chemical profile of cascade hop cones according to the growing area. J. Sci. Food Agric. 99 (13), 6011–6019.

Rossbauer, G., et al., 1995. Phenological growth stages of hop (humulus lupulus l.). Coding and description according to the extended BBCH scale with illustrations. https://www.openagrar.de/servlets/MCRFileNodeServlet/openagrar_derivate_00010428/BBCH-Skala_en.pdf. (Accessed 24 July 2025).

Rybka, A., et al., 2017. Effect of storage duration and atmosphere on the content and price of hop alpha bitter acids. Sci. Agric. Bohem. 48 (4), 245–251. http://dx.doi.org/10.1515/sab-2017-0032.

Saadi, S., et al., 2015. Climate change and mediterranean agriculture: Impacts on winter wheat and tomato crop evapotranspiration, irrigation requirements and yield. Agricult. Water. Manag. 147, 103–115. http://dx.doi.org/10.1016/j.agwat.2014.05.008.

dos Santos Costa, D., et al., 2019. Development of predictive models for quality and maturation stage attributes of wine grapes using vis-nir reflectance spectroscopy. Postharvest Biol. Technol. 150, 166–178. http://dx.doi.org/10.1016/j.postharvbio.2018.12.010.

Schönberger, C., Kostelecky, T., 2011. 125Th anniversary review: The role of hops in brewing. J. Inst. Brew. 117 (3), 259–267. http://dx.doi.org/10.1002/j.2050-0416.2011.tb00471.x.

Shaphiro, S., Wilk, M., 1965. An analysis of variance test for normality. Biometrika 52 (3), 591–611. http://dx.doi.org/10.2307/2333709.

Sharp, D.C., et al., 2014. Effect of harvest maturity on the chemical composition of cascade and Willamette hops. J. Am. Soc. Brew. Chem. 72 (4), 231–238. http://dx.doi.org/10.1094/ASBCJ-2014-1002-01.

Sirrine, J.R., Rothwell, N.L., Goldy, R.G., Marquie, S., Brown-Rytlewski, D., 2010. Sustainable Hop Production in the Great Lakes Region. Michigan State University Extension Publisher.

Skinner, R.N., et al., 1974. Hop maturation and resin formation: Field trials. J. Sci. Food Agric. 25 (9), 1121–1133. http://dx.doi.org/10.1002/jsfa.2740250908.

Skomra, U., Koziara-Ciupa, M., 2020. Stability of the hop bitter acids during long-term storage of cones with different maturity degree. Pol. J. Agron. (40), 16–24. http://dx.doi.org/10.26114/pja.iung.406.2020.40.03.

Srečec, S., et al., 2008. Influence of climatic conditions on accumulation of α-acids in hop clones. Agric. Conspec. Sci. 73 (3), 161–166.

XFarm, 2025. Xfarm platform. https://xfarm.ag. (Accessed 24 July 2025).

Zanoli, P., Zavatti, M., 2008. Pharmacognostic and pharmacological profile of humulus lupulus l.. J. Ethnopharmacol. 116 (3), 383–396. http://dx.doi.org/10.1016/j.jep.2008.01.011.