

Prediction of the Flowering of Chilean Native Species and Honey Yield Using Machine Learning An Approach to Climate Change Adaptation

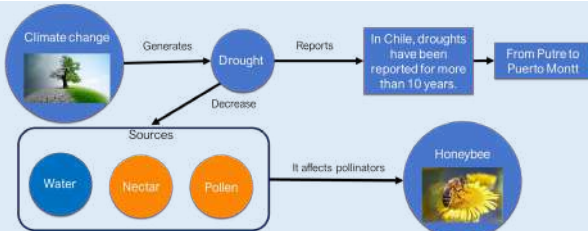
Roberto Ahumada-García¹, David Zabala-Blanco¹, Víctor Hugo Monzón²

¹ Faculty of Engineering Sciences, Universidad Católica del Maule, Talca, Chile.

² Bee Ecology Laboratory, Department of Biology and Chemistry, Faculty of Basic Sciences, Universidad Católica del Maule, Talca, Chile.

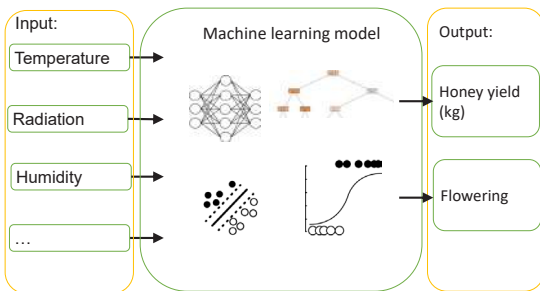
robertoahumadagarcia@gmail.com, rahumada@ucm.cl, vmonzon@ucm.cl

Introduction



Studies have established a relationship between climatic variables such as temperature and rainfall, flowering, and honey production. Currently, predictive models are being proposed to support decision-making and adaptation to climate change in beekeeping.

General Model



Machine learning models: neural networks, random forests, support vector machines (SVM), logistic regression.

Honey Yield Model

This study used historical harvest records and their corresponding climate data. The classification considered three categories: poor year (below 20 kg), moderate year (20 - 40 kg), and good year (over 40 kg). Machine learning algorithms called Bagging (Bootstrap Aggregating) were used, a technique that combines multiple models trained on random subsets of the original data.

Results Classification			
Model	Normalization	Accuracy	Time (seconds)
1 Random Forest	Without normalization	0.7347	13
2 Random Forest	Normalization with norm	0.7347	9
3 Random Forest	Normalization min-max	0.7347	12
4 Random Forest	Normalization z-score	0.7347	18
5 Random Forest	Normalization div-max	0.7347	6
6 Logistic Regression with Bagging	Without normalization	0.7347	360
7 Logistic Regression with Bagging	Normalization with norm	0.7351	43
8 Logistic Regression with Bagging	Normalization min-max	0.7355	60
9 Logistic Regression with Bagging	Normalization z-score	0.7355	60
10 Logistic Regression with Bagging	Normalization div-max	0.7347	60
11 SVM with Bagging	Without normalization	0.7347	60
12 SVM with Bagging	Normalization with norm	0.6939	29
13 SVM with Bagging	Normalization min-max	0.7359	24
14 SVM with Bagging	Normalization z-score	0.6863	30
15 SVM with Bagging	Normalization div-max	0.7347	60
16 Gradient Boosting	Without normalization	0.6735	25
17 Gradient Boosting	Normalization with norm	0.6735	28
18 Gradient Boosting	Normalization min-max	0.6735	33
19 Gradient Boosting	Normalization z-score	0.6735	33
20 Gradient Boosting	Normalization div-max	0.6735	33

Data from Australia were used (Campbell et al., 2020). The model was validated with 5-fold cross-validation using statistical techniques.

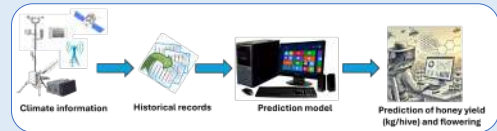
Results: The models show an accuracy of over 80% (correctly classified yields) up to three months before the main flowering period.

Materials and Methods

This study evaluates machine learning models (a type of artificial intelligence) to predict honey yield per hive and the flowering of native species.

The data supported by the literature for this application are:

- Meteorological data.
- Historical records of honey production.
- Historical records of plant phenological stages (budburst, flower bud, flowering).
- It may also include beekeeper management, internal hive conditions, and botanical information.
- In this first approach, only variables dependent on temperature, radiation, and humidity were considered, calculating monthly mean, standard deviation, minimum, maximum, and total values.



Flowering Model

This study focused on native species of high importance for beekeeping in Chile: Avellano, Corontillo, Litre, Peumo, Quillay, Tevo, and Tralhuén.

At present, a validated predictive model is available for different flowering stages of Quillay: initial flowering (10%), peak flowering (10 - 50%), and terminal flowering (10% in decline).



The model predicts the different flowering stages starting from the moment the beekeeper reports the flower bud stage. The machine learning algorithm used is a multilayer neural network.

Results: The flowering models show an error of about one week, allowing the prediction of different flowering stages approximately one month in advance. These predictive model results will provide beekeepers with a reliable tool to improve their hive management and production planning.

Conclusions

- The main studies suggest a relationship between temperature and rainfall with flowering and honey yield per hive.
- This work supports the possibility of predicting phenological flowering stages of native Chilean trees using climatic variables such as temperature, radiation, and humidity.
- These results indicate that using machine learning algorithms together with meteorological data makes it possible to predict honey yield.
- These results can be applied in decision-making and in adapting beekeeping to climate change.

Future Work

- The authors' goal is to strengthen collaboration between universities and beekeepers in order to gather information for the development of predictive models, with a focus on Chilean, Latin American, and global beekeeping.
- This study will promote the evaluation of different machine learning methods and datasets on flowering and hive yields.
- Similarly, the use of deep learning models (the foundation of ChatGPT) is being explored for these applications.

Acknowledgments

The author Roberto Ahumada-García acknowledges the financial support of ANID-Subdirección de Capital Humano/Doctorado Nacional/2024-21241043. The authors acknowledge Richard Manríquez Ramírez, founder and CEO of Bee Fractal (Chile), for providing flowering data. Richard Manríquez Ramírez has also launched the B-Mai AI application. QR code for download:



References

- M. Gajardo-Rojas et al., "Declining honey production and beekeeper adaptation to climate change in Chile," Prog. Phys. Geogr., vol. 46, no. 5, pp. 737–756, Oct. 2022.
- T. Campbell, K. W. Dixon, K. Dods, P. Fearn, and R. Handcock, "Machine Learning Regression Model for Predicting Honey Harvests," Agric. 2020, Vol. 10, Page 118, vol. 10, no. 4, p. 118, Apr. 2020.
- N. Grandon, R. Ahumada-García, J. Reyes-Suárez, and F. Besoain, "Information system for improving local productivity and decision making in organic beekeeping," 2016 IEEE Int. Conf. Autom. ICA-ACCA 2016, Dec. 2016.
- R. Ahumada and J. Reyes-Suarez, "A first approximation for prediction of bloom phenology of *Quillaja saponaria* (quillay), a Chilean native species of beekeeping interest," in BMEICB-02, Nov. 2016, p. 17009.
- Meteorological Department of Chile, "Monitoring of Meteorological Drought." Drought Bulletin. 2023.

Prediction of the Flowering of Chilean Native Species and Honey Yield Using Machine Learning – An Approach to Climate Change Adaptation

Roberto Ahumada-García¹, David Zabala-Blanco¹, Víctor Hugo Monzon²

¹Faculty of Engineering Sciences, Universidad Católica del Maule, Talca, Chile

²Ecology bees Laboratory. Faculty of Basic Sciences, Universidad Católica del Maule, Talca, Chile

Honey production per hive in countries such as Chile has been irregular and has decreased to values below 20 kg per hive, mainly attributable to climate change. These conditions are largely associated with changes in the regularity and timing of temperatures, precipitation, and the phenological stages of native species. Therefore, this study develops predictive models for the flowering of Chilean native species of apicultural interest (Avellano, Corontillo, Litre, Peumo, Quillay, Tevo, and Tralhuén) and estimates hive honey yield weeks or even months before the main flowering events. To achieve this, climatic data such as temperature, precipitation, and humidity from the nearest station are used to predict flowering stages of native Chilean trees and honey yield per hive. This research employs machine learning algorithms including Random Forest, SVM, Logistic Regression, together with Bagging techniques and Neural Networks. Current flowering models have shown an error of about one week, with predictions made roughly one month in advance. This has been corroborated with phenological state models for the Quillaja Saponaria species (Quillay), predicting the timing of flowering transitions from the floral bud stage to the onset of flowering, peak flowering, and final flowering. Following the same approach, the method has been used to predict hive honey yield, with prediction errors of less than 10 kg and accuracy exceeding 80%. These findings demonstrate the feasibility of using climatic variables to forecast the flowering of Chilean native trees, as well as to predict apicultural yield. This technology is essential for developing information systems that help reduce uncertainty in beekeeping operations. The above information is now being used to build an expert system that beekeepers can utilize to optimize production decisions, such as controlling hive development and diseases during specific periods.