Proceedings of the 7th International Conference on Statistics: Theory and Applications (ICSTA'25)
Paris, France - August, 2025
Paper No. 122
DOI: 10.11159/icsta25.122

Bayesian Markov Switching Models for prediction in Precision Beekeeping

M. C. Robustillo[1], L. Naranjo[2], M. I. Parra[3], C. J. Pérez[3]
[1] Departamento de Análisis Matemático, Estadística e Investigación Operativa y Matemática Aplicada, Facultad de Ciencias, Universidad de Málaga, 29010 Málaga, Spain
mcrobustillo@uma.es
[2] Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México
Circuito Exterior s/n, Ciudad Universitaria, Coyoacán, 04510 Ciudad de México, Mexico
lizbethna@ciencias.unam.mx
[3]Departamento de Matemáticas, Facultad de Ciencias, Universidad de Extremadura
Avenida de Elvas s/n, 06006 Badajoz, Spain
mipa@unex.es; carper@unex.es

Extended Abstract

Precision beekeeping leverages sensor data to optimize hive management, monitor colony health, and address challenges such as declining bee populations and environmental stressors. However, the complexity of analysing sensor-generated time series data and the need for accurate predictions to support decision-making highlight significant methodological gaps. Predictive models capable of handling high-dimensional, nonlinear, and incomplete data are essential for unlocking the potential of these datasets to improve beekeeping practices and inform policy [1].

This study employs data from the *BeeObserver* database, a comprehensive repository of sensor readings from honeybee hives, to develop advanced statistical models [2]. Hidden Markov Models (HMMs) based on Bayesian methodology are applied to simultaneously manage multiple time series, capturing their dependencies and interactions [3]. Latent variables are incorporated to handle missing values, inherent in sensor data collection. To account for behavioural shifts in the data, a Markov-switching model framework [4] is integrated within the models, enabling the analysis of time series trends under varying conditions. Noninformative prior distributions are defined to ensure model flexibility. Parameter estimation and prediction are conducted using Hamiltonian Monte Carlo methods and Stan software [5, 6], addressing challenges posed by the high dimensionality, strong correlations, and nonlinearity of the posterior distribution.

Preliminary results demonstrate the efficacy of these methodologies in uncovering complex patterns within hive data. The models successfully identify temporal trends, manage data gaps, and adapt to regime changes, providing insights into hive dynamics that traditional methods fail to capture. Furthermore, the predictive capabilities of the developed framework show promise for early detection of anomalous hive conditions, such as health deterioration or environmental disruptions.

In conclusion, this study highlights the potential of Bayesian HMMs with Markov-switching model for precision beekeeping by offering robust tools for analysing and predicting sensor data. The integration of latent variables and Markov-switching models enables comprehensive management of multifaceted time series data, paving the way for data-driven decision-making in apiculture. These findings emphasize the role of advanced statistical techniques in addressing critical challenges in beekeeping and underscore the importance of predictive analytics in promoting sustainable hive management practices.

References

[1] Zhang, Y. and Thorburn, P. J. (2022). Handling missing data in near real-time environmental monitoring: A system and a review of selected methods. Future Generation Computer Systems. Elsevier. Volume 128, 63-72, doi: https://doi.org/10.1016/j.future.2021.09.033

[2] Senger, D., Gruber, C., Kluss, T., and Johannsen C. (2024). Weight, temperature and humidity sensor data of honey bee colonies in Germany, 2019-2022. Data in Brief. Volume 52, 110015, doi: https://doi.org/10.1016/j.dib.2023.110015.

[3] Zucchini, W., MacDonald, I. L., and Langrock, R. (2016). Hidden Markov models for time series: an introduction using R. Number 150 in Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, 2nd. edition.
[4] Frühwirth-Schnatter, S. (2006). Finite mixture and Markov switching models. Springer.
[5] Betancourt, M. (2018). A conceptual introduction to Hamiltonian Monte Carlo. arXiv: https://arxiv.org/abs/1701.02434
[6] Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.2.