

MAPPING RESEARCH TRENDS AND KNOWLEDGE GAPS IN USDA-NIFA-FUNDED HLB RESEARCH USING TOPIC MODELING

Ulianova Vidal-Gomez¹, Neil McRoberts², David McClure³, Mike Hill³, Sebastian Galindo⁴, Lauren M. Diepenbrock¹, Luca Mantegazza⁴, and MEGAN M. DEWDNEY¹

¹Citrus Research and Education Center, University of Florida.

700 Experiment Station Rd. Lake Alfred, FL. 33850. FL. USA.

mmdewdney@ufl.edu

²University of California, Davis, Davis, CA

³Purdue University, West Lafayette, IN

⁴University of Florida, Gainesville, FL

This project analyzes and translates research outputs from United States Department of Agriculture (USDA)-funded efforts addressing Huanglongbing (HLB), one of the greatest threats to the U.S. citrus industry. The initiative supports the development of a publicly accessible web-based database designed to systematize and share more than 15 years of fragmented research with growers, stakeholders, and researchers through accessible, actionable information. Mapping research progress is a critical step toward better knowledge transfer, data-driven decisions, and identifying where more work is still needed.

To achieve this goal, text mining and unsupervised topic modeling approaches are being applied to 342 HLB-related projects funded through the USDA National Institute of Food and Agriculture (NIFA). The study includes original project statements, progress reports, and associated research outputs (e.g., peer-reviewed papers, conferences' abstracts, pre-prints). Previous evaluations, using the Latent Dirichlet Allocation (LDA) method, identified major research efforts focused on Asian citrus psyllid management, disease surveillance, and monitoring, while areas such as citrus nutrition and antimicrobial strategies appeared less frequently, suggesting underexplored opportunities with potential for rapid field-level impact. Preliminary results also revealed shifts between original research objectives and eventual publications, reflecting the dynamic evolution of research priorities during project implementation.

Current work uses an embedding-based topic modeling approach that provides additional benefits: it automatically identifies the number of topics and creates semantically similar clusters of documents and words. Applying this method to

our dataset will generate a topic map of projects statements, impacts reports, and linked publications that will be used to identify semantic relationships among research areas and analyze differences between the topics of proposed research goals and published outputs. To our knowledge, it is the first effort to systematically connect research with an accessible knowledge framework intended not only for scientists, but also for growers and stakeholders seeking practical information for HLB-management and decision-making.

Conference:

Vidal Gomez, U., McRoberts, N., M., McClure, D., Hill, M., Galindo S., Diepenbrock, L., Mantegazza, L. and Dewdney, M. (2026, July 12-17). Mapping research trends and knowledge gaps in USDA-NIFA-funded HLB research using topic modeling. [Poster presentation]. 58th Annual Meeting Caribbean Food Crops Society, 1st Annual Virgin Islands Caribbean Agricultural Conference with the participation of the American Phytopathological Society - Caribbean Division. "Increasing Food Security in the Caribbean Through Innovative Technology". University of Virgin Islands, Albert A. Sheen Campus, St. Croix, United States Virgin Islands.