



HLB Research and
Extension Outputs Network

Beyond the Psyllid: Are We Aligning HLB Research with Real-World Challenges?

Ulianova Vidal-Gomez¹, Kaity Coleman², Sandra Olkowski², Neil McRoberts², David McClure³, Uros Mojsilovic³, Mike Hill³, Dharmendra Kalauni⁴, Sebastian Galindo⁴, Lauren M. Diepenbrock¹, Tripti Vashisth¹, Yu Wang¹, and Megan M. Dewdney¹

¹Citrus Research and Education Center, University of Florida, Lake Alfred, FL. ²University of California, Davis, Davis, CA. ³Purdue University, West Lafayette, IN., ⁴University of Florida, Gainesville, FL.

INTRODUCTION

We analyzed USDA-funded* projects to identify research trends on huanglongbing (HLB). Text-mining 256 summaries and 2,000+ outputs reports reveals vector management as the dominant focus, followed by host resistance, pathogen detection, and disease surveillance. Citrus nutrition and antimicrobial studies are limited, while capacity building and training are common but rarely evaluated, highlighting thematic gaps, and guiding research toward growers' needs.

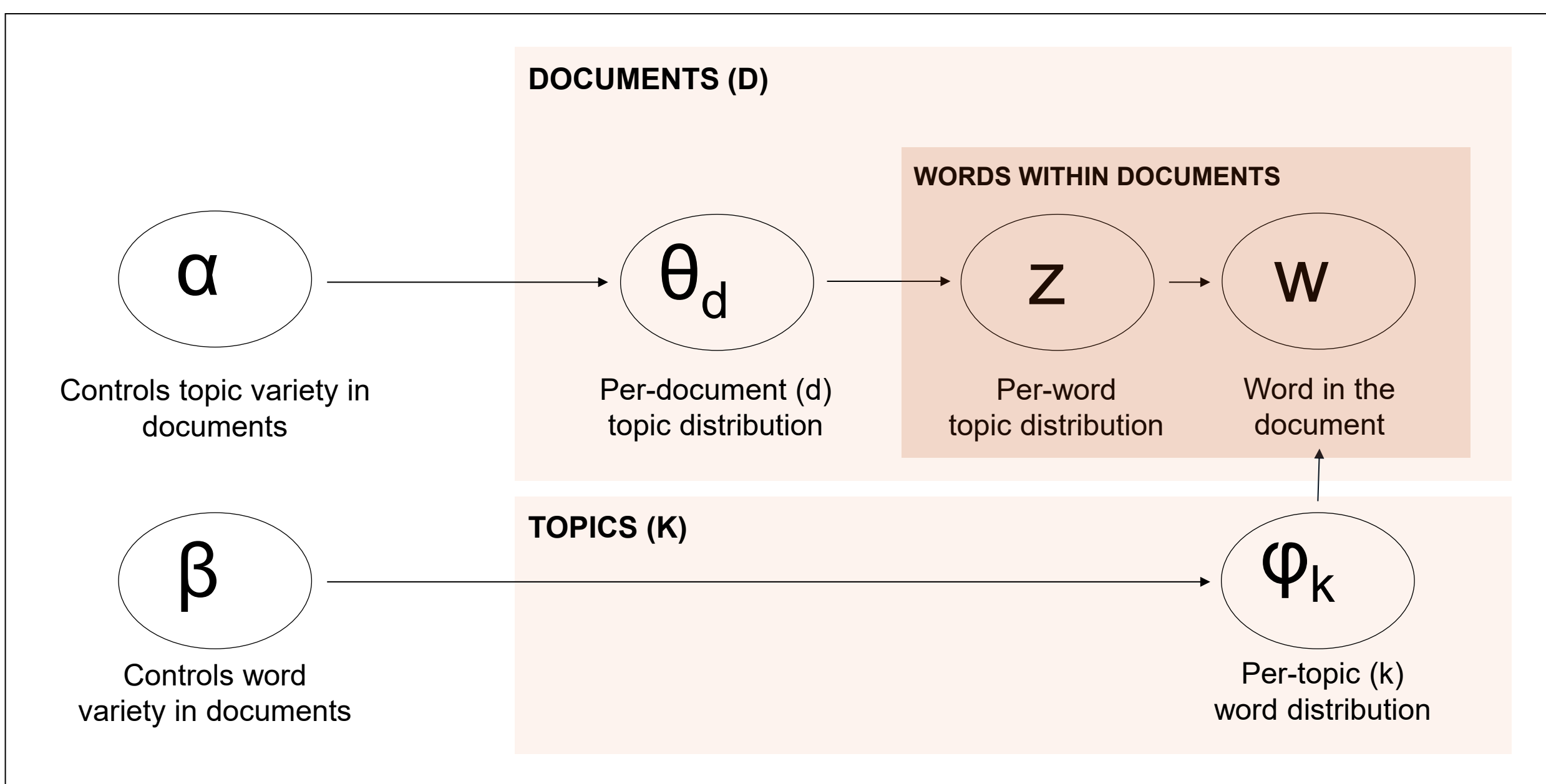
*United States Department of Agriculture (USDA)

METHODS

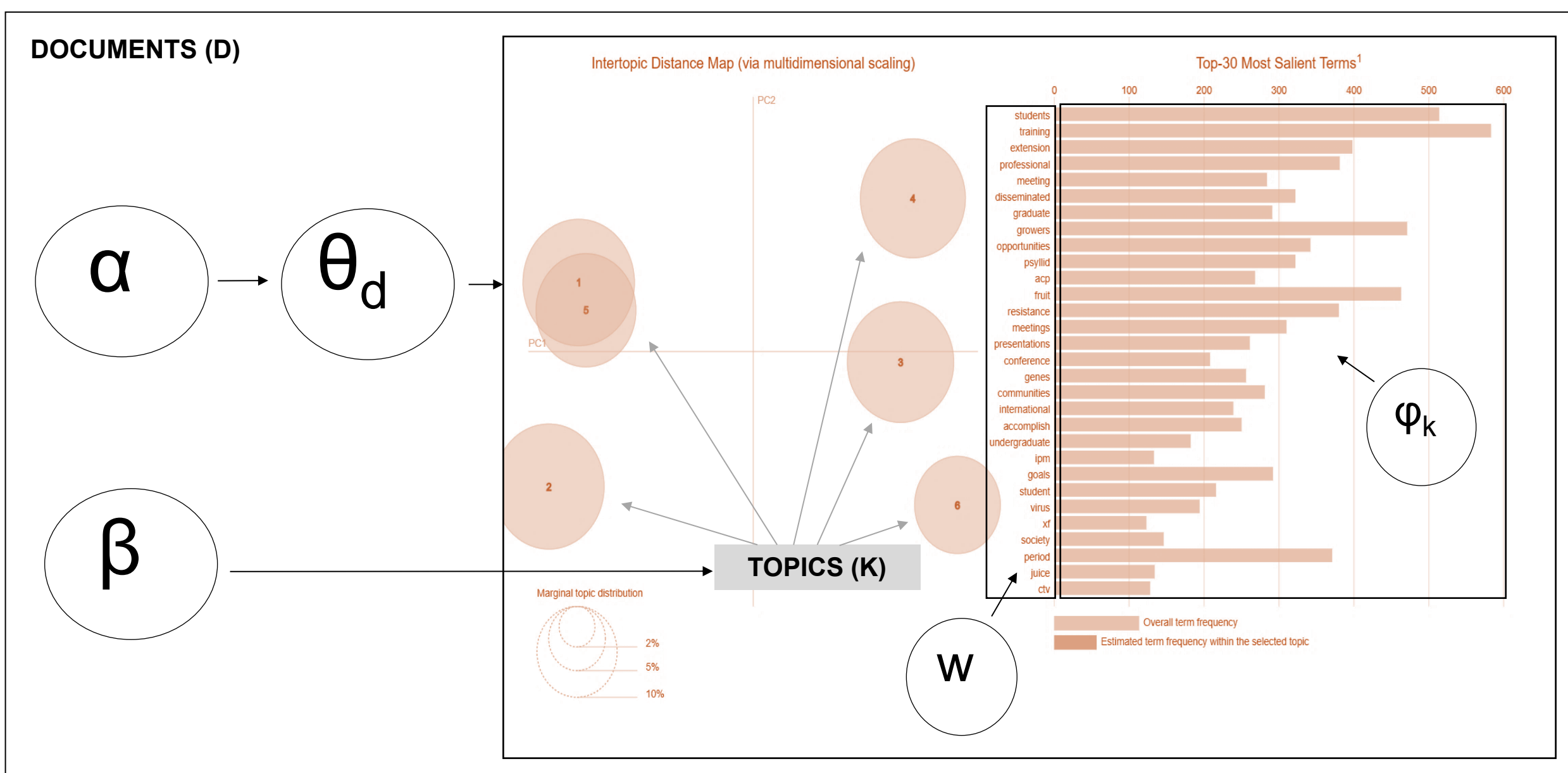
We analyzed **256 USDA-funded HLB projects** from the **USDA NIFA Project Reporting Portal (Gateway)**, including non-technical summaries and outputs (**>870 entries**). The workflow included:

- **Data Preparation** - Cleaned and standardized text; created a document-term matrix.
- **Vectorization** - Converted text into numerical form (**TF-IDF**), used **only for UMAP geometry** to calculate document distances.
- **Word Mapping** - Applied **UMAP** to visualize document relationships.
- **Topic Modeling** - Used **LDA** to identify six research themes from the top 300 words (β) and dominant topics (θ).

Latent Dirichlet Allocation (LDA) Model:



LDA Outputs:



Analyses were conducted in R 4.4.1, using “tidytext”, “topicmodels”, “tidyverse”, “textclean” (tm), “topicmodels”, “ldatuning”, “umap” (uwot), “LDAvis”, “FactoMineR”, “writexl”, “readr”, and “ggplot” libraries (R core team 2025).

RESULTS

SECTION 1: Comparative thematic clustering of HLB research grants

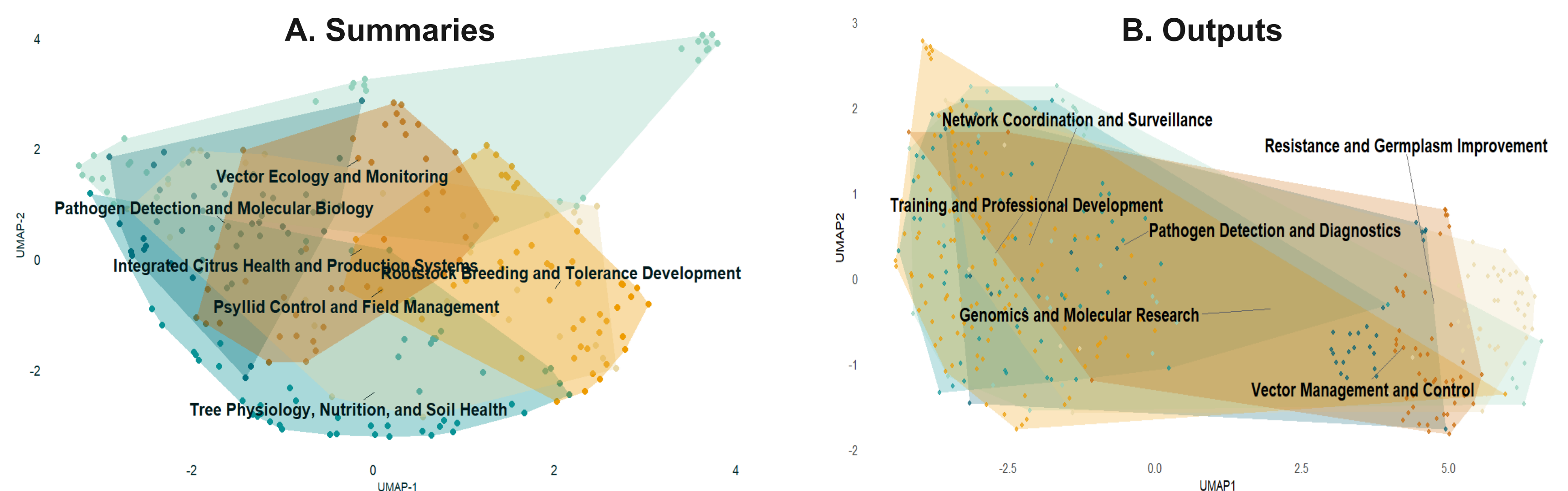


Fig. 1. UMAP projection showing the similarity among 256 NIFA-funded HLB research grants based on textual content from A. Summaries, and B. Outputs. Each point represents a grant. Grants with similar topic compositions appear closer together in the 2D space, revealing thematic clusters.

SECTION 2. Topic modeling of USDA-NIFA HLB grants

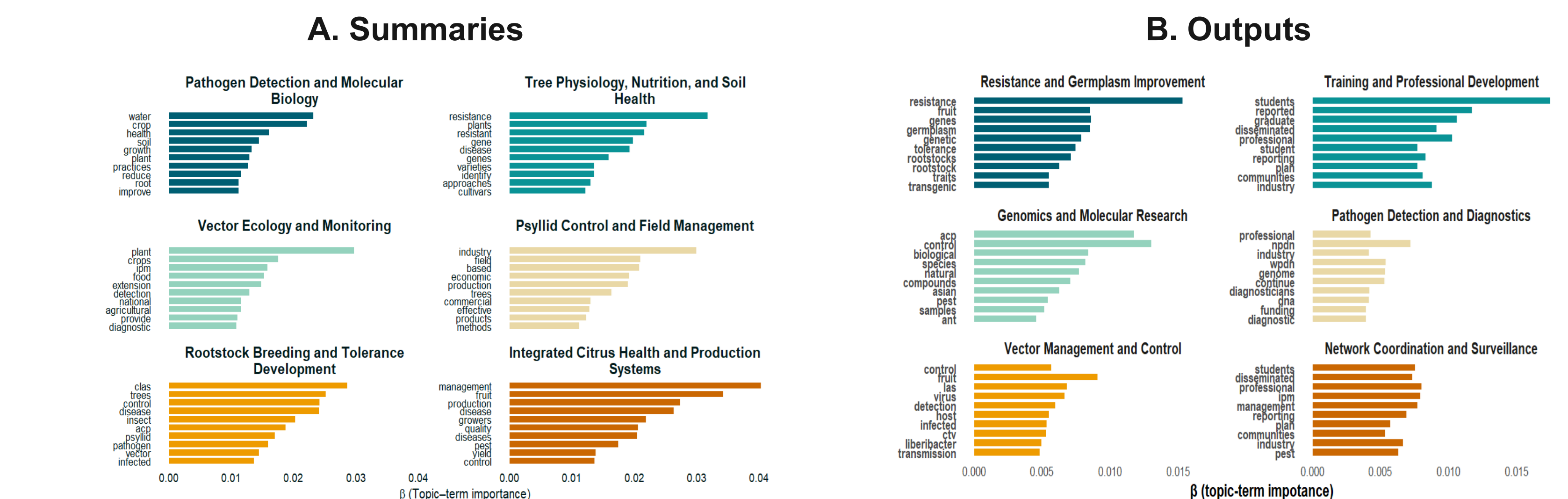


Fig. 2. Top 10 words with the highest β values (topic-term importance) for each topic identified by the LDA model ($k = 6$). Word frequency patterns reveal the key terms defining each research area in USDA-NIFA HLB project summaries and outputs.

SECTION 3. Number of grants per research theme

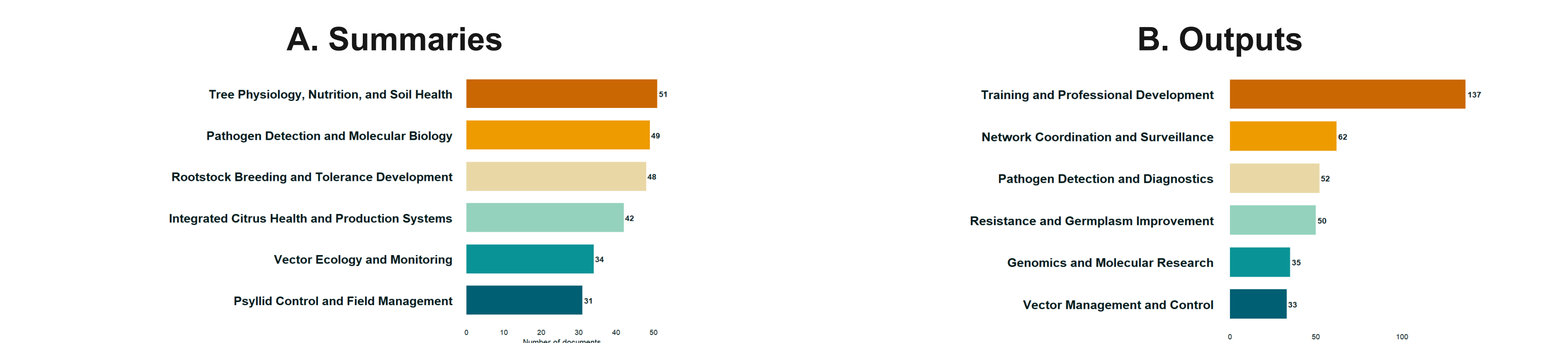


Fig. 3. Figure: Number of documents assigned to each topic generated by the LDA model ($k = 6$), based on USDA-NIFA HLB project summaries and outputs.

DISCUSSION

- Vector management and control (Asian Citrus Psyllid - ACP) dominate the HLB research narrative and is closely linked with breeding, resistance, molecular biology and gene expression.
- Despite significant progress in vector control, areas such as citrus nutrition or antimicrobial strategies remain underexplored yet essential for rapid, field-level solutions.
- Training and workforce development appear across grants, but their impacts remain unclear and unquantified.
- To advance HLB management, it is crucial to continue diversifying research beyond vector control and other key areas, as this may lead to more diverse, integrated, and applicable outcomes.

CONCLUSION

Understanding how research priorities shift from proposals to outputs reveals critical gaps and opportunities. While vector control remains central, advancing progress requires diversification into citrus nutrition, antimicrobial strategies, and other emerging areas.

REFERENCES

- USDA (n.d.). *NIFA Reporting Portal*. U.S. Department of Agriculture: <https://portal.nifa.usda.gov>
- We used Latent Dirichlet Allocation (LDA; Blei et al., 2003) for topic modeling, visualized topics with LDAvis (Sievert & Shirley, 2015) following principles by Chuang et al. (2012), and applied UMAP (McInnes et al., 2018) for 2D projection.