

Honeybees: leveraging Nature’s own environmental sensor for climate resilience

Ryan Coffee

August 8, 2022

Community Engagement Climate events are occurring with increasing rapidity. Our effort hopes to break ground with demonstration cases that spur partnerships across public and private sectors, span government agencies, and cross the divide between cultural and socioeconomic groups. Such partnerships are a requisite element [1] that underlies any national effort to build adaptive and resilient energy and climate strategies and policies that can leverage the latest emerging technologies in AI, machine learning, robotics, and applied mathematics. This initiative will lead the path for robust and adaptive precision agriculture and an intelligent equitable energy strategy that at once recognizes the individual interests of all active participants and shareholders, while predicting climate and humanitarian events, and optimizing dynamic adaptation to accommodate and even capitalize on them. Initiated under the auspices of an open-source community group called The Enterprise Neurosystem which includes members from SLAC/Stanford, multiple universities, and private and public companies, our community is coalescing around a very wide-reaching effort for an AI infrastructure that promotes climate resilience via adaptive agricultural practices, and provides course correction recommendations that mitigate the human and environmental impact of climate events. As an initial demonstration of the utility of broad information sharing for model training, the team is working in close collaboration with the Department of Energy’s Artificial Intelligence and Technology Office (AITO) to develop the secure integration of distributed datasets that intentionally span multiple scientific domains. This can be achieved with a federation model that preserves data privacy as needed, openly exposes publicly shared data where available, and creates mutually beneficial AI models that generalize across domains of human knowledge.

The Hive Mind Our pursuit of active adaptation to climate events will foster resilient agriculture and energy systems with the aim to bring artificial intelligence centrally into the task of integrating the full range of human environmental monitoring and response. We have an ultimate goal of overlaying climate-relevant satellite imagery with local agricultural health. We have begun with a project developed from a grass-roots partnership between the DOE’s AITO, SLAC’s cross-domain EdgeAI research, and the open-source community of industry partners—the Enterprise Neurosystem. Our shared vision is to merge together information from multiple modes of environmental monitoring and at vastly different geospatial scales.

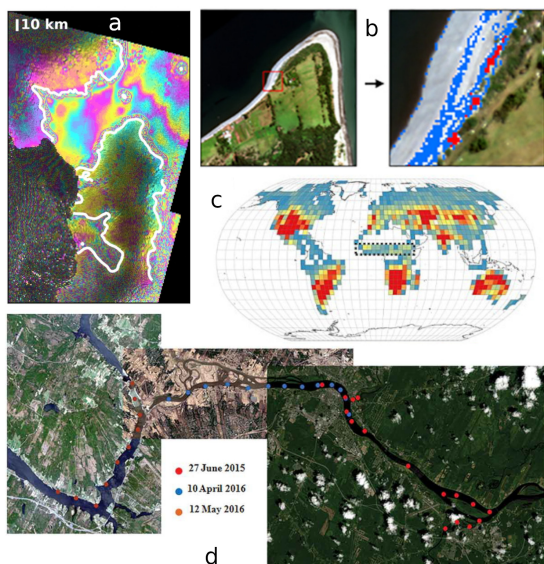


Figure 1: Composite of satellite imaging of various modalities, and regions, based on Refs. [2–5] for a-d respectively.

large-scale satellite imagery available, we have turned our attention to ground-level primary sources, thus we plan to ask the bees themselves.

Stanford collaborator Dr. Tadashi Fukami uses precise individual bee microbiome analysis [8] to reconstruct the nectar diversity for a hive. One can imagine point measurements of nectar diversity attributing spatial labels into geospatial images much like the water sample points (dots) in Fig. 1d. Our team is currently focused on continuous passive acoustic monitoring that would add point measurements of hive health and state. Starting small and local, we are quickly growing to partner with a much broader diversity of hive locations across complementary environments—industrial and local agricultural sites, native tribal lands, urban sites, and international sites where domestication of native honeybee species is a common practice.

Growth into diverse environments is precisely the benefit of engaging with the AITO. By working directly with typically under-engaged institutions that represent the agricultural and urban demographics that are key to project success, AITO is helping foster AI development in the communities that will likely have an outsized positive impact on the future of climate adaptability.

Figure 2 shows the first prototype testing of a tiny ultrasonic microphone, placed inside the beehive, that records the hive humming at well above human acoustic perception. The sensors are designed with easily attainable parts, require little advanced knowledge for construction, and cost less than \$3 US each when complete. Our team is also providing the AI software and design elements free of charge, to encourage adoption in broader agricultural and environmental-minded communities. We also encourage their direct engagement with the community to foster collective solution ownership, and a robust co-design paradigm.

Given the Red Hat origins of the Enterprise Neurosystem community, this project sticks to a philosophy of interdisciplinary open-source sharing of ideas and methods, including

For instance, global scale environmental satellite imagery (Fig. 1), from meters to kilometers, can provide a family of diagnostic overlays for point samples on the ground, or in the water as shown in Fig. 1d adapted from Ref. [5]. Such overlays individually focus in on, e.g., crop diversity and regional pollution from refuse [3], agricultural runoff [6] as well as aquifer [7] and glacier health [2], and surface water concentrations [5]. And while a satellite modality might capture CO₂ concentration [4] at the kilometer meter scale, urban and agricultural beehive health could give a finer mesh of overall environmental health.

Correlating hive health with variations in numerous environmental factors is just the job of data-hungry machine learning algorithms. But this project starts with a few beehives at a time. Given the wealth of

hardware integration and software design. As a community endeavor, all members share a sense of ownership of the technology.

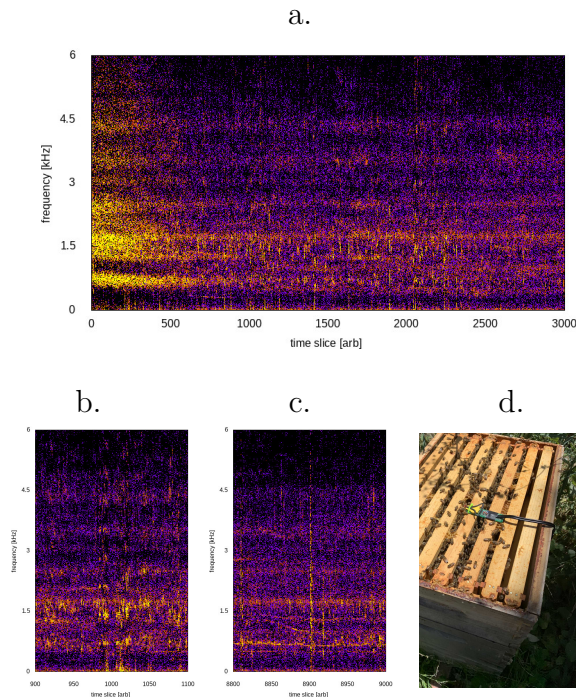


Figure 2: a. Agitated initial state and calming. b. Post-calming communication. c. Calm communication. d. Prototype acoustic sensor with frequency response from 20Hz to 85kHz.

our project is focused on acoustics for hive health, we plan to leverage compatible format data from fields as far-reaching as data-center acoustics and Tokamak magnetic fusion diagnostics (see Fig. 3).

Sharing is Daring The challenge of open data sharing is a topic that the research sector finds unnervingly challenging. The personal motivations that drive researchers to push the limits of possibility are often the same motivations that discourage the open sharing of uniquely produced datasets with understandably competitive community members. This personal motivation of Principal Investigators (PIs) is not unlike the intellectual property (IP) concerns of private sector entities, or the economic aspirations of both developed and developing nations. As noted in the UN Development Programme’s sustainable development goals [1], addressing global climate goals will require equitable partnerships that respect the economic aspirations of all participants.

The US research ecosystem, in particular across the Department of Energy, represents a functionally similar challenge for equitable data sharing, in support of a cohesive and sustainable future for the broad adoption of AI practices. We can only find true success in

Sharing is Caring As a seed idea to catalyze the direction for this multidisciplinary endeavor for climate and energy resilience, we take inspiration from the accomplishments that so-called Transformer [9] or Foundation [10] models have gleaned from tremendously diverse training corpora. The impressive results in language completion tasks for so-called transformer or foundation models like BERT [11] and GPT [12] have been achieved precisely owing to the diverse language corpora available as training sets. Unfortunately, the ability of tremendously large models to leverage vast datasets requires equally massive computational capability—a capability typically reserved for major industrial entities and the largest computational facilities in DOE.

Innovation, however, comes from non-traditional combinations of human ideas, and these combinations shine a spotlight into the gaps between research fields. Similarly, large machine learning (ML) models attain “generalizability” by training across increasingly diverse datasets. And although

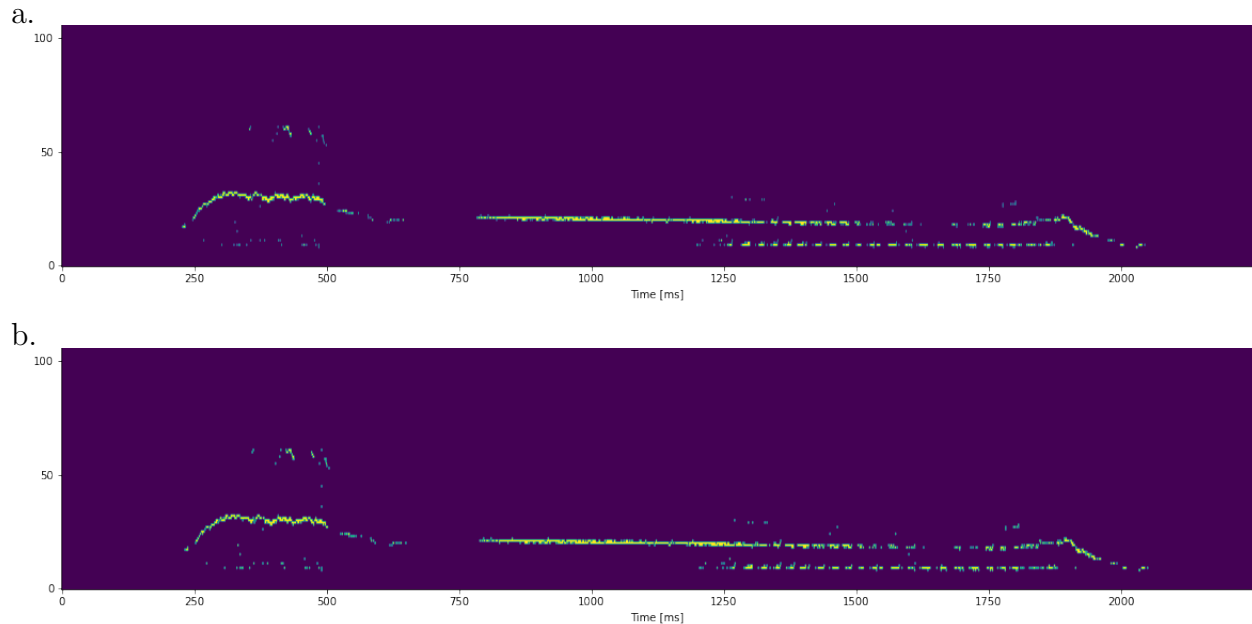


Figure 3: a. Ground truth spectrogram of ECE signal. b. Forecast of signal by 5ms.

a multi-modal, data-rich, and geographically and culturally distributed national computing ecosystem.

Solving a grand challenge and building a far reaching vision, however, always begins with a small exemplar—one hive at a time.

References

- [1] 'United Nations Development Programme'. Undp: Sustainable development goals, 2015. url: <https://www.undp.org/sustainable-development-goals>.
- [2] Yara Mohajerani, Seongsu Jeong, Bernd Scheuchl, Isabella Velicogna, Eric Rignot, and Pietro Milillo. Automatic delineation of glacier grounding lines in differential interferometric synthetic-aperture radar data using deep learning. *Scientific Reports*, 11(1):4992, Mar 2021.
- [3] Tomás Acuña-Ruz, Diego Uribe, Richard Taylor, Lucas Amézquita, María Cristina Guzmán, Javier Merrill, Paula Martínez, Leandro Voisin, and Cristian Mattar B. Anthropogenic marine debris over beaches: Spectral characterization for remote sensing applications. *Remote Sensing of Environment*, 217:309–322, 2018.
- [4] Zhao-Cheng Zeng, Liping Lei, Kimberly Strong, Dylan B. A. Jones, Lijie Guo, Min Liu, Feng Deng, Nicholas M. Deutscher, Manvendra K. Dubey, David W. T. Griffith, Frank Hase, Bradley Henderson, Rigel Kivi, Rodica Lindenmaier, Isamu Morino, Justus Notholt, Hirofumi Ohyama, Christof Petri, Ralf Sussmann, Voltaire A. Velazco, Paul O. Wennberg, and Hui Lin. Global land mapping of satellite-observed co2 total columns

- using spatio-temporal geostatistics. *International Journal of Digital Earth*, 10(4):426–456, 2017.
- [5] Essam Sharaf El Din, Yun Zhang, and Alaeldin Suliman. Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. *International Journal of Remote Sensing*, 38(4):1023–1042, 2017.
 - [6] Muhammad Shoaib Saleem, Zachary Y. Huang, and Meghan O. Milbrath. Neonicotinoid pesticides are more toxic to honey bees at lower temperatures: Implications for overwintering bees. *Frontiers in Ecology and Evolution*, 8, 2020.
 - [7] John W. Crowley, Jerry X. Mitrovica, Richard C. Bailey, Mark E. Tamisiea, and James L. Davis. Land water storage within the congo basin inferred from grace satellite gravity data. *Geophysical Research Letters*, 33(19), 2006.
 - [8] Leslie E. Decker, Priscilla A. San Juan, Magdalena L. Warren, Cory E. Duckworth, Cheng Gao, and Tadashi Fukami. Higher variability in fungi compared to bacteria in the foraging honey bee gut. *Microbial Ecology*, Jan 2022.
 - [9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.
 - [10] Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri S. Chatterji, Annie S. Chen, Kathleen Creel, Jared Quincy Davis, Dorottya Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah D. Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, Omar Khattab, Pang Wei Koh, Mark S. Krass, Ranjay Krishna, Rohith Kuditipudi, and et al. On the opportunities and risks of foundation models. *CoRR*, abs/2108.07258, 2021.
 - [11] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805, 2018.
 - [12] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2018.