David LeBauer, Ph.D.

2911 E Canyon Bend St, Tucson, AZ, 27408; 760-468-8621; dlebauer@gmail.com

Education

2008 **Ph.D.** Earth System Science, University of California at Irvine
2001 **M.S.** Ecology, Agriculture, University of California at Davis
1998 **B.S.** Biology with honors, Ecology concentration, Duke University

Selected Work Experience

Feb 2023–Feb 2024. **Staff Scientist**, Indigo Ag

- Technical lead for team that calibrated, validated, and wrote report following Climate Action Reserve Soil Enrichment Protocol to support generation of verified carbon credits.
- Harmonized and migrated an extensive dataset of observed soil carbon, greenhouse gas fluxes, and yields; migrated dataset to a Postgres relational geospatial database.
- Wrote science requirements for automated model calibration and validation.

Jul 2019–Jan 2023. **Director of Data Science**, University of Arizona Division of Agriculture, Life and Veterinary Sciences, and Cooperative Extension

- Founded and built the data science support team comprising 4-6 data scientists and software engineers to provide research support for the Division of Agriculture, Life Science, and Cooperative Extension at the University of Arizona. datascience.cct.arizona.edu
- Delivered and coordinated workshops to enable data intensive and reproducible research. CCT Data Science YouTube & datascience.cct.arizona.edu/workshops
- Developed and launched a data science incubator program, leveraging institutional funds to provide researchers with technical support for data intensive projects with outcomes that include: producing preliminary results; upgrading scientific software; publishing papers, data, and code; creating data visualization dashboards and analysis pipelines: datascience.cct.arizona.edu/incubator
- Facilitated cross-disciplinary, collaborative research projects through data curation, statistical guidance, and translation of research priorities across domains.
- Acquired over $2m in extramural funding to support modeling, computing, and informatics projects from ARPA-E, DARPA, USDA, and NSF.

Jun 2015– Jul 2018 (part time). **Senior Scientific Advisor**, Agrible, Inc

- Designed and implemented analyses of yield data used to inform seed choice.
- Consulted with software engineers and scientists to help in the development and analysis of a crop model that was used to inform and optimize farms and supply chains.
- Taught quantification and propagation of uncertainty in crop models, including soil physical properties and advised on implementation of these methods.

Jul 2012– Jul 2018. **Research Scientist**, University of Illinois

2015–2019 **Data and Computing Pipeline Lead**, *TERRA-REF*

- Awarded $1.7m from the DOE plus over $1m in computing and storage from NSF-funded computing resources to lead a team of three full time and ten part time programmers in the design and implementation of a pipeline of a cutting edge sensor based plant observation data stream.
- Coordinated development of software and data products across six collaborating research groups representing diverse domains across genomics, physiology, breeding, remote sensing, computer vision, and robotics. Obtained and integrated feedback from end users, an advisory committee, and industry partners into data products and metadata standards.
- Published the world's largest public domain agricultural data set including over 1PB of sensor, trait, and genomics data.

2012-2015 **Scientific Manager**, EBI Ecosystem Modeling Program

- Managed scientific research program that predicted yield potential, yield stability, and environmental impacts of bioenergy feedstock production.

- Responsible for $600k annual budget that supported research across four faculty, three postdocs, one research programmer, a project manager, multiple contractors, and dozens of part-time student employees.
- Led development and application of simulation, statistical, and informatics software.

Aug 2009-Jul 2012. **Postdoctoral Researcher**, University of Illinois

- Created, implemented (as PEcAn and BETYdb), and distributed a new software platform and database to support model-data synthesis.

Data & Software Projects

PEcAn: The Predictive Ecosystem Analyzer
Creator, developer, & co-PI 2009-present
Ecoinformatics toolbox for model-data synthesis, analysis, and prediction. Integrates observations with scientific understanding of crop and ecosystem functioning as an iterative, directed learning process.
web:pecanproject.org; code:github.com/PecanProject/pecan

BETYdb: Biofuel Ecophysiological Traits and Yields database
Creator & development lead 2009-present
Database network and web interface to harmonize heterogeneous plant, ecosystem, and agronomic data. Supports meta-analysis, simulation modeling, and synchronization across research teams; currently used for data management by six research teams.
web: betydb.org; code: github.com/PecanProject/bety

TERRA REF: TERRA Reference Phenotyping Platform
Lead, data and computing pipeline, 2015-2019
Computing pipeline, reference data products, and cloud environments advancing the use of technology to improve crops and agriculture. Data and metadata products that adopt existing conventions and implement new ones that promote interoperability.
web: terraref.org; code: github.com/terraref

BioCro: Bioenergy Crop simulator
Development lead 2013-2015
Dynamical simulation model combining physics, chemistry and physiological processes to predict crop growth and water use. Development lead.
code: github.com/ebimodeling/biocro

GHGVC: Ecosystem Climate Regulation Services Calculator.
Development lead 2013-2017
Web interface and R package to compute biophysical and biogeochemical impacts of land use change.
web: www.ecosystemservicescalc.org; code: github.com/ebimodeling/ghgvc

Notable Publications

2020 What does TERRA-REF's high resolution, multi sensor plant sensing public domain data offer the Computer Vision Community? Proc. IEEE/CVF International Conference on Computer Vision doi: 10.1109/iccvw54120.2021.00162
2017 BETYdb: a yield, trait, and ecosystem service database applied to second-generation bioenergy feedstock production; Global Change Biology-Bioenergy doi: 10.1111/gcbb.12420
2013 Facilitating feedbacks between field measurements and ecosystem models, Ecological Monographs doi: 10.1890/12-0137.1
2008 Nitrogen limitation of net primary productivity is globally distributed, Ecology doi: 10.1890/06-2057.1

Honors and Awards

2023 Equity, Inclusion, and Diversity Award. North American Plant Phenotyping Network
2021 Best Paper, 7th workshop on Computer Vision in Plant Phenotyping and Agriculture, ICCV 2021
2017 Outstanding Mentor, Students Pushing Innovation program at NCSA
2014 Faculty Fellow, National Center for Supercomputing Applications
2007 Graduate Student Representative, Department of Earth System Science
2005 Graduate Fellow, Kearney Soil Science Foundation
1998 Benenson Award in the Arts: Built a mushroom garden in the Duke Gardens