

Agricultural Collaborative Research Outcomes System (AgCROS): A network of networks connecting food security, the environment, and human health

Jorge A. Delgado, Bruce Vandenberg, Nicole Kaplan, Donna Neer, Greg Wilson, Robert D'Adamo, Jennifer Carter, Laura O'Gan, Nadene Grow, Roger Marquez, Dan Arthur, Marlen Eve, Stephen J. Del Grosso, Jane M.F. Johnson, Douglas L. Karlen, Lisa Durso, John Finley, Veronica Acosta-Martinez, David B. Knaebel, Daren Harmel, and Justin D. Derner

National and international open-access agricultural research databases are needed to help solve problems at watershed, regional, and national scales, and to connect productivity, soil health, and environmental quality to food quantity and quality. There are some established, open-access agricultural research networks with extensive research data in the United States, but there is a major need to improve connections between those networks and the emerging data in order to address complex questions. Improving the connections and flow of information among agricultural research networks will enhance the scientific community's ability to simultaneously increase crop yield, sustainability of natural resources, and environmental quality, as well as food, feed, and forage quality, and thus human and animal health. Establishing a network of agricultural databases is crucial for facilitating information flow among different research disciplines. Doing so will also enhance multidisciplinary research opportunities and help build transdisciplinary teams that can provide answers to complex, whole-system research questions and thus solve some of the globe's greatest challenges.

GLOBAL CHALLENGES

The ever-growing human population is facing one of the greatest challenges of the twenty-first century: to ensure the sustainability of agricultural and natural systems. Both are under the pressure of a changing climate and the increase in extreme events that accompanies it. A statement by the United Nations (UN) Secretary General about a special climate change report recently released by the UN Intergovernmental Panel on Climate Change, warns the global community that climate change is occurring at a faster rate than humanity is addressing it and that environmental consequences, including more extreme weather and sea level rise, will continue impacting humanity in the years to come (United Nations 2018).

The occurrence of extreme weather events such as droughts, floods, and hurricanes will create havoc in some of the world's agricultural production systems with far-reaching impacts on food security and human health mediated by zoonotic and vector-borne diseases (Watts et al. 2015). Adaptation to a changing climate will be of the utmost importance. Conservation practices are key to adapting to a changing climate and will be essential to ensure that humanity maintains the productivity and the sustainability needed to confront those challenges (Delgado et al. 2011; Walthall et al. 2012; Spiegel et al. 2018).

Responding to great global challenges such as climate change will require breaking down silos and working in larger, more diverse teams to answer larger-scale research questions concerning the impacts of agricultural systems and how we could use management and conservation practices to adapt to a changing climate while providing ecosystem services. A key principle of adaptation to a changing climate is the need to "improve historical context" with long-term databases (Delgado et al. 2011). Well-described and organized open-access databases will increase the impact of research data by making it available to larger networks of scientists, extension personnel, consultants, private companies, professors, students, and members of the general public. This effort is not beginning from scratch; indeed, the agricultural scientific community has been developing database networks for the last two decades, but never have they been integrated into a linked system. The Agricultural Collaborative Research Outcomes System (AgCROS) brings together databases such as Greenhouse gas Reduction through Agricultural Carbon Enhancement Network/ Resilient Economic Agricultural Practices, formerly the Renewable Energy Assessment Project (GRACenet/REAP) (Del Grosso et al. 2014; Jawson et al. 2005; Leytem et al. 2014; Spiegel et al. 2018), Long-Term Agroecosystem Research Network

Jorge A. Delgado is a research soil scientist with the USDA ARS (Agricultural Research Service) Soil Management and Sugar Beet Research Unit, Fort Collins, Colorado. **Bruce Vandenberg** is an information technology specialist (data management) with the USDA ARS, Fort Collins, Colorado. **Nicole Kaplan** is a computational biologist with the USDA ARS Rangeland Resources and Systems Research Unit, Fort Collins, Colorado. **Donna Neer** is a biological science technician with the USDA ARS Soil Management and Sugar Beet Research Unit, Fort Collins, Colorado. **Greg Wilson** is a retired scientist, USDA ARS, Beltsville, Maryland. **Robert D'Adamo** is a physical science technician with the USDA ARS Soil Management and Sugar Beet Research Unit, Fort Collins, Colorado. **Jennifer Carter** is a computational biologist with the USDA ARS, Mandan, North Dakota. **Laura O'Gan** is an information technology specialist (data management) with the USDA Animal and Plant Health Inspection Service, Fort Collins, Colorado. **Nadene Grow** is an information technology specialist with the USDA ARS, Fort Collins, Colorado. **Roger Marquez** is computer assistant with the USDA ARS, Fort Collins, Colorado. **Dan Arthur** is an information technology specialist (data management) with the USDA ARS, University Park, Pennsylvania. **Marlen Eve** is deputy administrator with the USDA ARS Office of National Programs, Beltsville, Maryland. **Stephen J. Del Grosso** is a research soil scientist with the USDA ARS Soil Management and Sugar Beet Research Unit, Fort Collins, Colorado. **Jane M.F. Johnson** is a research soil scientist with the USDA ARS Soil Management Research Unit, Morris, Minnesota. **Douglas L. Karlen** is a soil scientist with the USDA ARS Soil, Water, and Air Resources Research Unit, Ames, Iowa. **Lisa Durso** is a research microbiologist with the USDA ARS Agroecosystem Management Research Unit, Lincoln, Nebraska. **John Finley** is national program leader (human nutrition) with the USDA ARS Office of National Programs, Beltsville, Maryland. **Veronica Acosta-Martinez** is a research soil scientist (microbiologist and biochemist) with the USDA ARS Wind Erosion and Water Conservation Unit, Lubbock, Texas. **David B. Knaebel** is national program leader (soil biology) with the USDA ARS Office of National Programs, Beltsville, Maryland. **Daren Harmel** is center director with the USDA ARS Center for Agricultural Resources Research, Fort Collins, Colorado. **Justin D. Derner** is research leader with the USDA ARS Rangeland Resources and Systems Research Unit, Cheyenne, Wyoming.

Figure 1

The Agricultural Collaborative Research Outcomes System (AgCROS) is a growing “network of networks” that presently consists of multiple agricultural data networks: Nutrient Uptake and Outcome Network (NUOnet), the Greenhouse gas Reduction through Agricultural Carbon Enhancement Network (GRACEnet), Resilient Economic Agricultural Practices (REAP), Dairy Agriculture for People and the Planet (DAPP; Dairy Grand Challenge), Soil Health Assessment Network (SHAnet), Agricultural Antibiotic Resistance (AgAR), and the Long-Term Agroecosystem Research (LTAR) Network. By integrating these diverse database networks, AgCROS facilitates the flow of information and increases the cooperation among researchers participating in these networks.



(LTAR) (Spiegel et al. 2018), Nutrient Use and Outcome Network (NUOnet), Soil Health Assessment Network (SHAnet), and other databases, helping to integrate long-term agricultural research data and provide users data at the depth and breadth needed to develop viable solutions to whole-system challenges and assess conservation practices and their effectiveness in mitigating and adapting to climate change (figure 1).

CULTURAL AND TECHNICAL CHALLENGES TO OPEN ACCESS DATA

Cheruvilil and Soranno (2018) reported that in the field of ecology many of the challenges research is attempting to address will require larger sets of heterogeneous data, and they reported an open data system with a team approach will be more successful in answering big-picture research questions. We suggest that the same is true for the field of agriculture. Ecology and agriculture share many of the same cultural obstacles to developing large team

approaches, including the myth of the lone genius and scientists traditionally being trained in a single discipline and conducting research individually (Cheruvilil and Soranno 2018), with accomplishments predominantly being rewarded as individual efforts (Uriarte et al. 2007; Geman and Geman 2016). Agriculture is an applied research field that provides evidence-based outcomes for management decisions, and new grand challenges in agriculture will require data from various projects. However, data-intensive methods can be perceived as theoretical and not hypothesis-driven, and scientists may be hesitant to access and reuse data from other sources to complete a novel analysis (Cheruvilil and Soranno 2018; Elliott et al. 2016). Despite these cultural obstacles, “the powerful synergies realized by practicing team science and open science in conjunction with data-intensive science” (Cheruvilil and Soranno 2018) with data and information provided

by a comprehensive database could be realized in the agricultural research community.

Even with the advantages described above and new requirements to make data open (Obama 2013), many barriers remain that prevent researchers from embracing openness in their own work (Hampton et al. 2015) and sharing data (Couture et al. 2018). To overcome these barriers, there must be a paradigm shift in thinking from data ownership to data stewardship (Hampton et al. 2015; Easterday et al. 2018) and a collective approach to care of the data (Baker and Karasti 2018). However, some scientists may lack the resources or the particular expertise needed to transfer their data into repositories or database networks (Waide et al. 2017). Large volumes of data in file cabinets, floppy disks, and other media represent “dark data” (Heidorn 2008), and the absence of this data for others to use has hindered scientific advancement. Infrastructure for open-access data repositories would increase baseline information, thereby enabling assessments of changes with time and across space (Whitlock et al. 2010). Additionally, it would allow scientists to use these datasets to repeat studies at other sites and see if these studies could be replicated and validated. Finally, it would allow the preservation of datasets for future generations.

There are also technical and logistical challenges to sharing and reusing data, and ecology presents some lessons learned. Comparing data collected under different objectives is a difficult endeavor, particularly if the same methods were not applied to all the sites and various collection efforts were made at different sampling intensities or frequencies, even when sites or ecosystems don’t differ dramatically (Cushing et al. 2008). Working with data from organized networks and within scientific collaborations can help mitigate the challenges to comparing data across various sites and projects, and produce data of higher quality with more impactful research results with far broader implications (Moran 2016). Peters et al. (2014) reported that there is a need for coordination among ecological networks to provide data that can be scaled to address challenges such as a changing climate, and

a need to use similar methods to measure attributes, using a controlled nomenclature when joining different datasets into a network to reduce bias. For recommendations to alleviate several of these challenges (e.g., using common metadata protocols for consistent units and terminology and establishing requirements for data curation and repositories), see Eagle et al. (2017).

LESSONS LEARNED TO ENSURE LONG-TERM AVAILABILITY OF DATA

There are large global efforts for open data and there are open-data policy initiatives taken on by governments and research sponsors. However, long-term data availability tends to remain limited without a reliable and robust data infrastructure or funding to support development of systems to contain and provide open-access data (Peters et al. 2014). There are many examples in the ecological research community (DataONE.org and Knowledge Network for Biocomplexity) and the agricultural research community (the Global Open Data for Agriculture and Nutrition Initiative [GODAN]). New technologies, including increased storage capacity in “the cloud,” a network of servers accessed through the Internet; emerging data documentation standards (e.g., metadata); and more rapid development via Web applications, set the stage for a future of long-term access to data and information. Along with the opportunities presented by these emergent technologies across multiple scientific fields today, and just as envisioned by Peters et al. (2014) and Waide et al. (2017) for ecology, we see it as key to develop a strategic vision for infrastructure to link data repositories at a national scale in order to help simplify data discovery and integrate the available data from across domains, such as ecology and agriculture. Coordinating the cyberinfrastructure needed to manage a “network of networks” would be resource challenging (Waide et al. 2017), but data availability in the future is dependent on coordinated efforts for availability and accessibility now.

A linked data landscape includes connections between data repositories, so that one may discover data and information from various starting points, among a web-of-repositories (Baker 2009).

However, as a first step, data must be shared to these repositories, which in agricultural research has not been the tradition. Included in that infrastructure should be services for data providers and data users to ensure data are prepared to be findable via online portals to data repositories, accessible for download, interoperable between machines, and understood by humans to be reusable in new analyses, according to FAIR principles (Wilkinson 2016). Such services should begin with inclusion of a data management plan within any research plan and approaches to data collection that allow for more efficient quality assurance, analysis, documentation, and sharing of data (Michener and Jones 2012). Preparing data for long-term availability can be challenging, but data should be treated as valuable assets because it is difficult to judge the future value of any data in helping to accelerate the speed of science, forecast future conditions, catalyze a collaboration, or amplify the impact of research.

ECOLOGICAL NETWORKS

Ecological studies have moved from fine-scale and plot-level emphases to larger scale and long-term temporal studies that are oriented toward examining regional- to continental- and global-scale phenomena (Peters et al. 2014). Ecological data in the public domain, however, are disseminated through several different data repositories in various academic, government, and other institutions (Waide et al. 2017). For example, US governmental ecological databases are available from the Environmental Protection Agency, US Geological Survey, National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, Department of Energy, USDA Forest Service, National Science Foundation Geosciences, National Science Foundation Biological Sciences, and National Institutes of Health (Waide et al. 2107). An additional ecological nutrient network is NutNet, which has over 40 worldwide sites and addresses the effects of nutrient inputs and herbivores on ecosystem composition and function (Adler et al. 2011; Borer et al. 2014).

EXISTING AND EMERGING AGRICULTURAL DATA NETWORKS

While the ecological community has faced the need to address larger-scale problems similar to the ones faced by modern agriculture, the agricultural community has developed a number of key, focused databases that have advanced agricultural research and outcomes. Some of these can be integrated with the AgCROS network as needed to further advance solutions to changing climate challenges. Research networks can provide a venue for virtual research communities that rely on sharing data among collaborators and inform development of data infrastructure for their work. They can be at an advantage by having the provision of support for participants and data managers to collaborate on specifications for organizing, documenting, and sharing data. Many of our lessons learned come from the Long-Term Ecological Research (LTER) Network, and in agriculture we have the new LTAR Network, which is coordinating data efforts via a working group consisting of data managers from each LTAR site to address how to share a large variety of data types and sizes, much like the LTER network did when designing their first network information system (Baker 2000).

Some examples of successful projects that have contributed to the generation of large agricultural databases include the USDA Agricultural Research Service (ARS) GRACEnet, REAP, and Conservation Effects Assessment Project (CEAP). These projects, with locations collecting data across different regions of the United States, generated the GRACEnet, REAP, and Sustaining the Earth's Watersheds, Agricultural Research Data System (STEWARDS) networks, with millions of publicly available records with supportive data on site descriptors, experimental design, and methods; as well as management and measured data. The data, methods, and protocols used in these networks have been used by other US and international scientists.

Big data networks are now emerging in agricultural research through advances in molecular genetics networks that contain high-throughput sequencing data and provide important biological information about potential gene func-

tions. An example of such a network is MaizeGDB. These networks will have long-term research data focused on maize (*Zea mays* L.). Tian et al. (2018) reported on MCENet, which is a database for a maize conditional co-expression network oriented toward helping researchers identify maize functional genes or modules that regulate important agronomic traits. The USDA ARS Germplasm Resources Information Network (GRIN) provides germplasm information about plants, animals, microbes, and invertebrates.

Agricultural databases focused on food have also been developed. An example is the International Network of Food Data Systems (INFOODS), which was founded at the Food and Agriculture Organization and which has contributed to the harmonization and compilation of food data (Murphy et al. 2016). The USDA ARS Food Data System (FoodS) is a unique, national data system that includes the Food and Nutrient Database for Dietary Studies and the Food Patterns Equivalents Database, which have important information about specifics of food composition.

Other agricultural data networks and/or repositories are the Ag Data Commons of the National Agricultural Library, and Measured Annual Nutrient loads from Agricultural

Environments (MANAGE) (Harmel et al. 2016). Additionally, there are soil database networks, such as the USDA Natural Resources Conservation Service (NRCS) Geospatial Data Gateway, which has soil survey data. The USDA NRCS also has the PLANTS national resources inventory, as well as the National Soil Survey Center. There are also drought networks, such as the North American Drought Monitor (NADM), which is a network to monitor drought in Canada, Mexico, and the United States.

The Australian Soil Resource Information System (ASRIS) is a national data repository that includes almost 5,000 nutrient site-years of experimental data across different regions of Australia with different climates and soil types. The data was collected using a standard template, and the process allows for ongoing collection of data. The Farm Nutrient Loss Index (FNLI) decision support tool can also be downloaded from the ASRIS site to assess the risk of nutrient loss from runoff, gases, or leaching. The European Soil Bureau and the European Soil Data Center provide datasets, services/applications, maps, documents, events, projects, and external links related to soil data. These are just a few of the growing number of agricultural data networks in existence.

AGCROS: A NEW APPROACH TO ENHANCE AGRICULTURAL DATA DISCOVERY

AgCROS is a “network of networks” that links multiple agricultural databases using consistent variable names, units, and methods (figure 2). The AgCROS system is set up so that data contributors can provide their research data using commercial data management systems and/or the Data Entry Template (DET) originally designed for the GRACenet and REAP efforts (Del Grosso et al. 2014). The DET is a collaborative tool for bringing data together. Some of the significant impacts of both AgCROS and the DET approach are that they have not only helped bring many agricultural data networks together but also have facilitated communication and cooperation between teams of researchers working in different areas of research.

AgCROS and the DET approach have enabled cooperation among data teams that compile data, database managers that execute scripts and merge data from multiple locations, and a system administrator that publishes the data on the cloud. Metadata for agricultural datasets about management practices; animal management practices; and weather, physical, chemical, and biological data for soil, water, air, and crop

Copyright © 2018 Soil and Water Conservation Society. All rights reserved. Journal of Soil and Water Conservation 73(6): 158A-164A www.swcs.org

Figure 2

Website for the Agricultural Collaborative Research Outcomes System (AgCROS) network, with links to participating networks.



resources are aggregated for the National Agricultural Library's Ag Data Commons. To the best of our knowledge, AgCROS is the first "network of networks" for agricultural data. AgCROS has linked research from the GRACEnet, REAP, NUOnet, SHAnet, Dairy Agriculture for People and the Planet (DAPP), and LTAR networks (figure 3). It is also inherently designed to allow addition of future network efforts.

The most recent example of AgCROS connecting different agricultural database networks is the new NUOnet (figure 4). NUOnet is a multidisciplinary, open-access database that stores data about nutrient inputs, cycling, and sources (Delgado et al. 2016; Eve et al. 2016). Organic and inorganic nutrient inputs, crops, forage, and livestock can all be managed to increase productivity and sustainability. One of the goals of NUOnet was the connection of NUOnet to other database networks. NUOnet is strengthening existing and established USDA ARS networks by contributing additional types of data that were not previously available in the already established GRACEnet and REAP databases. NUOnet is helping to provide more complete information that could be used to fine-tune models used for greenhouse gas (GHG) inventories since it adds data about nutrient management and the effects of best management practices on the potential to reduce nitrate-nitrogen ($\text{NO}_3\text{-N}$) leaching, ammonia (NH_3) volatilization, and nitrogen (N) runoff.

NUOnet has been developed using the established framework of the USDA ARS GRACEnet and REAP networks. Using the DET approach from GRACEnet, the GRACEnet/REAP DET was expanded with newly added tables for entering additional information about nutrient management research and pathways for nutrient losses (e.g., $\text{NO}_3\text{-N}$ leaching, NH_3 volatilization, and surface carbon [C] and N losses). The connection of networks such as the NUOnet, GRACEnet, and REAP can help researchers answer larger-scale questions.

Indirect emissions of nitrous oxide (N_2O) in the United States account for about 22% of the GHG losses from agriculture when expressed as carbon dioxide (CO_2) equivalent emissions (USEPA 2010). With the connection of networks

Figure 3

The Agricultural Collaborative Research Outcomes System (AgCROS) consists of multiple agricultural database networks. This map shows Long-Term Agroecosystem Research (LTAR) locations and collaborators; the locations of Greenhouse gas Reduction through Agricultural Carbon Enhancement Network (GRACEnet), Nutrient Uptake and Outcome Network (NUOnet), Resilient Economic Agricultural Practices (REAP), and Agricultural Antibiotic Resistance (AgAR); and Natural Resources Conservation Service Land Resource Regions in the United States.

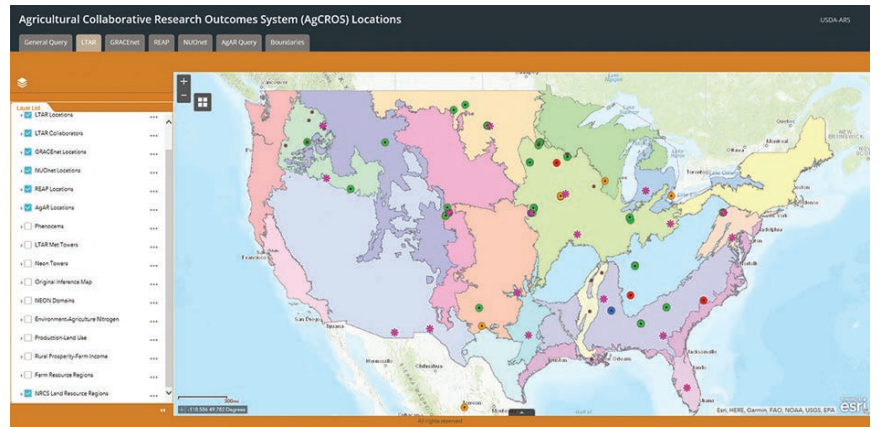
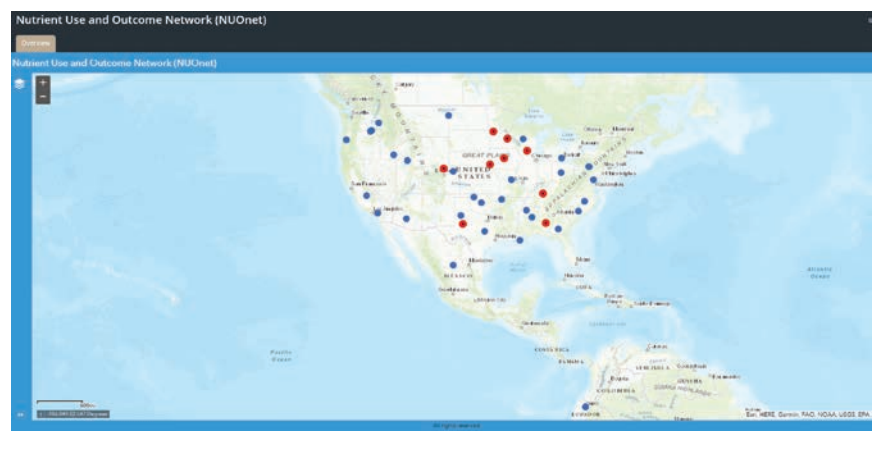


Figure 4

Locations of sites contributing data and/or metadata for the first Nutrient Uptake and Outcome Network (NUOnet) data collection cycle (red points) and sites that have expressed interest in contributing data to NUOnet in future data collection cycles (blue points).



in AgCROS, it is possible to download datasets that provide more complete information about C sequestration, GHG emissions, and nutrient losses in order to answer how best management practices can reduce direct and indirect GHG emissions, increase C sequestration, and reduce nutrients losses to the environment. Additionally, NUOnet is connected to the USDA ARS Dairy Grand Challenge and aims to connect datasets on forage management, nutrient use efficiencies, effects

on forage quality, and potential effects on milk quality and animal and human health. These are just a few examples of how AgCROS is helping to connect information and make it publicly available to enable discovery of viable management practices in a whole-system approach.

Similarly, work is ongoing to expand the wealth of soil biological data through SHAnet. Significant advances have been made in how we envision the soil as a living component, and there is now

improved understanding of soil biology through new methodology that explores the changes in size, diversity, and activity of soil organisms at different levels, including the microbiome. The soil biological data being collected across regions will enhance our understanding of soil health as affected by the complex interactions of management and climate variability.

AgCROS is connecting agricultural research databases, an effort that we believe is needed to help answer research questions at the local, regional, national, continental (e.g., North America), and even global scale for agricultural, natural, and urban landscapes.

FUTURE OPPORTUNITIES

AgCROS will continue to expand with additional networks and will serve as a platform to bring scientists together to connect networks with their respective databases in a way that enables information to be used to answer complex research questions and provide solutions to customers. As it grows, AgCROS will be able to provide larger datasets to the public to enhance the discovery of information, the development of viable products, and model calibration and validation, among other advantages.

We propose that AgCROS can provide data that could eventually be used in new technologies such as machine learning, robotics, and artificial intelligence. Robotics applications are rapidly advancing and will increasingly be incorporated into agricultural applications such as pesticide sprayers, automatic harvesting devices, soil systems, automatic grafting processes, and other applications. Together with drones and unmanned aerial vehicles, some of these future systems will use networks to complete their applications (Belforte et al. 2018).

While we cannot know what new developments may emerge in the agricultural sector over the next two to three decades, we suggest that the AgCROS approach of team cooperation and a “network of networks” for agricultural data will be used by new technologies in the future in ways that cannot be imagined today. We agree with the Cheruvilil and Soranno (2018) proposal for ecological studies, and propose

a similar approach for agricultural studies. AgCROS is designed to be a unique “network of networks” for agricultural data that will allow researchers test new empirical patterns and help provide new insights of large processes at greater agricultural scales to resolve grand-scale challenges that cannot be adequately addressed or understood with finer-scale studies. We hope that AgCROS will assist management practice and system research and new technological innovations for maintaining or increasing yields and increasing profitability while reducing negative environmental impacts and increasing soil, crop, animal, and human health.

DISCLAIMER

The USDA prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotope, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.

Trade and manufacturer's names are necessary to report factually on available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by USDA implies no approval of the product to the exclusion of others that may also be suitable.

REFERENCES

Adler, P.B., E.W. Seabloom, E.T. Borer, H. Hillebrand, Y. Hautier, A. Hector, W.S. Harpole, L.R. O'Halloran, J.B. Grace, T.M. Anderson, J.D. Bakker, L.A. Biederman, C.S. Brown, Y.M. Buckley, L.B. Calabrese, C.-J. Chu, E.E. Cleland, S.L. Collins, K.L. Cottingham, M.J. Crawley, E.I. Damschen, K.F. Davies, N.M. DeCrappeo, P.A. Fay, J. Firn, P. Frater, E.I. Gasarch, D.S. Gruner, N. Hagenah, J. Hille Ris Lambers, H. Humphries, V.L. Jin, A.D. Kay, K.P. Kirkman, J.A. Klein, J.M.H. Knops, K.J. La Pierre, J.G. Lambrinos, W. Li, A.S. MacDougall, R.L. McCulley, B.A. Melbourne, C.E. Mitchell, J.L. Moore, J.W. Morgan, B. Mortensen, J.L.

Orrock, S.M. Prober, D.A. Pyke, A.C. Risch, M. Schuetz, C.J. Stevens, L.L. Sullivan, G. Wang, P.D. Wragg, J.P. Wright, and L.H. Yang. 2011. Productivity is a poor predictor of plant species richness. *Science* 333(6050):1750-1753.

Baker, K.S. 2009. Data stewardship: Environmental data curation and a web of repositories. *Digital Discourse: The E-volution of Scholarly Communication* 1(1).

Baker, K.S., B.J. Benson, D.L. Henshaw, D. Blodgett, J.H. Porter, and S.G. Stafford. 2000. Evolution of a multisite network information system: The LTER information management paradigm. *BioScience* 50(11):963-978.

Baker, K.S., and H. Karasti. 2018. Data care and its politics: Designing for local collective data management as a neglected thing. *In Proceedings of the 15th Participatory Design Conference: Full Papers-Volume 1:10*. New York: Association for Computing Machinery.

Belforte, G., G. Eula, T. Raparelli, S. Sirolli, P. Piccarolo, P. Gay, D. Riccauda Aimonino, and L. Comba. 2018. Preliminary design of an electro-pneumatic automatic machine for herbaceous grafting. *In Advances in Service and Industrial Robotics, Mechanisms and Machine Science* 49, eds. C. Ferraresi and G. Quaglia, pp. 255-263. Springer International Publishing AG 2018, doi:10.1007/978-3-319-61276-8_28.

Borer, E.T., W.S. Harpole, P.B. Adler, E.M. Lind, J.L. Orrock, E.W. Seabloom, and M.D. Smith. 2014. Finding generality in ecology: A model for globally distributed experiments. *Methods in Ecology and Evolution* 5:65-73.

Cheruvilil, K.S., and P.A. Soranno. 2018. Data-intensive ecological research is catalyzed by open science and team science. *BioScience* 68(10):813-822.

Couture, J.L., R.E. Blake, G. McDonald, and C.L. Ward. 2018. A funder-imposed data publication requirement seldom inspired data sharing. *PLOS One* 13(7):e0199789, doi:10.1371/journal.pone.0199789

Cushing, J.B., N.E. Kaplan, C. Laney, J. Mallett, K. Ramsey, K. Vanderbilt, and D. Milchunas. 2008. Integrating ecological data: Notes from the Grasslands ANPP Data Integration Project. *In Proceedings of the Environmental Information Management Conference, Albuquerque, New Mexico, September 10-11, 2008*.

Del Grosso, S.J., J.W. White, G. Wilson, B. Vandenberg, D.L. Karlen, R.F. Follett, J.M.F. Johnson, A.J. Franzluebbers, D.W. Archer, H.T. Gollany, M.A. Liebig, J. Ascough, M. Reyes-Fox, L. Pellack, J. Starr, N. Barbour, R.W. Polunsky, M. Gutwein,

- and D. James. 2014. Introducing the GRACEnet/REAP data contribution, discovery, and retrieval system. *Journal of Environmental Quality* 42:1274–1280.
- Delgado, J.A., P.M. Groffman, M.A. Nearing, T. Goddard, D. Reicosky, R. Lal, N.R. Kitchen, C.W. Rice, D. Towery, and P. Salon. 2011. Conservation practices to mitigate and adapt to climate change. *Journal of Soil and Water Conservation* 66(4):118A–129A, doi:10.2489/jswc.66.4.118A.
- Delgado, J.A., S. Weyers, C. Dell, D. Harmel, P. Kleinman, K. Sistani, A. Leytem, D. Huggins, T. Strickland, N. Kitchen, J. Meisinger, S. Del Grosso, J. Johnson, K. Balkcom, J. Finley, N. Fukagawa, M. Powell, and S. Van Pelt. 2016. USDA Agricultural Research Service creates Nutrient Uptake and Outcome Network (NUOnet). *Journal of Soil and Water Conservation* 71(6):147A–148A. doi:10.2489/jswc.71.6.147A.
- Eagle, A., L.E. Christianson, R.L. Cook, R.D. Harmel, F. Miguez, S.S. Qian, and D.R. Diaz. 2017. Meta-analysis constrained by data: Recommendations to improve relevance of nutrient management research. *Agronomy Journal* 109:1–9.
- Easterday, K., T. Paulson, P. DasMohapatra, P. Alagona, S. Feirer, and M. Kelly. 2018. From the field to the cloud: A review of three approaches to sharing historical data from field stations using principles from data science. *Frontiers in Environmental Science* 6:88.
- Elliott, K.C., K.S. Cheruvilil, G.M. Montgomery, and P.A. Soranno. 2016. Conceptions of good science in our data-rich world. *BioScience* 66:880–889.
- Eve, M., J.A. Delgado, S.L. Weyers, C. Dell, and D. Harmel. 2016. NUOnet: A network to help gather, use big data. *CSA News* 61:26–26, doi:10.2134/csa2016-61-10-11.
- Geman, D., and S. Geman. 2016. Science in the age of selfies. *Proceedings of the National Academy of Sciences* 113:9384–9387.
- Hampton, S.E., S.S. Anderson, S.C. Bagby, C. Gries, X. Han, E.M. Hart, M.B. Jones, W.C. Lenhardt, A. Macdonald, W.K. Michener, J. Mudge, A. Pourmohktarian, M.P. Schildhauer, K.H. Woo, and N. Zimmerman. 2015. The Tao of open science for ecology. *Ecosphere* 6(7):120.
- Harmel, R.D., L.E. Christianson, D.R. Smith, M.W. McBroom, and K.D. Higgs. 2016. Expansion of the MANAGE database with forest and drainage studies. *Journal of the American Water Resources Association* 52:1275–1279, doi:10.1111/1752-1688.12438.
- Heidorn. 2008. Shedding light on the dark data in the long tail of science. *Library Trends* 57:280–299.
- Jawson, M.D., S.R. Shafer, A.J. Franzluebbers, T.B. Parkin, and R.F. Follett. 2005. GRACEnet: Greenhouse gas Reduction through Agricultural Carbon Enhancement network. *Soil and Tillage Research* 83:167–172.
- Leytem, A.B., and R.S. Dungan. 2014. Livestock GRACEnet: A workgroup dedicated to evaluating and mitigating emissions from livestock production. *Journal of Environmental Quality* 43:1101–1110.
- Michener, W.K., and M.B. Jones. 2012. Ecoinformatics: Supporting ecology as a data-intensive science. *Trends in Ecology and Evolution* 27(2):85–93.
- Moran, M.S., P. Heilman, D.P. Peters, C. Holifield Collins. 2016. Agroecosystem research with big data and a modified scientific method using machine learning concepts. *Ecosphere* 7(10).
- Murphy, S.P., U.R. Charrondiere, and B. Burlingame. 2016. Thirty years of progress in harmonizing and compiling food data as a result of the establishment of INFOODS. *Food Chemistry* 193:2–5.
- Obama, B. 2013. Executive Order—Making Open and Machine Readable the New Default for Government Information. May 9, 2013. <https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government->
- Peters, D.P.C., H.W. Loescher, M.D. Sanclements, and K.M. Havstad. 2014. Taking the pulse of a continent: Expanding site-based research infrastructure for regional- to continental-scale ecology. *Ecosphere* 5(3):Article 29.
- Spiegel S., B.T. Bestelmeyer, D.W. Archer, D.J. Augustine, E.H. Boughton, R.K. Boughton, M.A. Cavigelli, P.E. Clark, J.D. Derner, E.W. Duncan, C.J. Hapeman, R.D. Harmel, P. Heilman, M.A. Holly, R. Huggins, K. King, P.J.A. Kleinman, M.A. Liebig, M.A. Locke, G.W. McCarty, N. Millar, S.B. Mirsky, T.B. Moorman, E.B. Pierson, J.R. Rigby, G.P. Robertson, J.L. Steiner, T.C. Strickland, H.M. Swain, B.J. Wienhold, J.D. Wulffhorst, M.A. Yost, and C.L. Walthall. 2018. Evaluating strategies for sustainable intensification of US agriculture through the Long-Term Agroecosystem Research network. *Environmental Research Letters* 13:034031
- Tian, T., Q. You, H. Yan, W. Xu, and Z. Su. 2018. MCENet: A database for maize conditional co-expression network and network characterization collaborated with multi-dimensional omics levels. *Journal of Genetics and Genomics* 45(7):351–360.
- United Nations. 2018. Statement by the Secretary-General on the IPCC Special Report Global Warming of 1.5°C. October 8, 2018. <https://www.un.org/sg/en/content/sg/statement/2018-10-08/statement-secretary-general-ipcc-special-report-global-warming-15-%C2%BAc>.
- Uriarte, M., H.A. Ewing, V.T. Eviner, and K.C. Weathers. 2007. Constructing a broader and more inclusive value system in science. *BioScience* 57:71–78.
- USEPA (US Environmental Protection Agency). 2010. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2008. Washington, DC: US Environmental Protection Agency. http://www.epa.gov/climatechange/emissions/downloads/10/508_Complete_GHG_1990_2008.pdf.
- Waide, R.B., J.W. Brunt, and M.S. Servilla. 2017. Demystifying the landscape of ecological data repositories in the United States. *BioScience* 67(12):1044–1051.
- Walthall, C.L., J. Hatfield, P. Backlund, L. Lengnick, E. Marshall, M. Walsh, S. Adkins, M. Aillery, E.A. Ainsworth, C. Ammann, C.J. Anderson, I. Bartomeus, L.H. Baumgard, F. Booker, B. Bradley, D.M. Blumenthal, J. Bunce, K. Burke, S.M. Dabney, J.A. Delgado, J. Dukes, A. Funk, K. Garrett, M. Glenn, D.A. Grantz, D. Goodrich, S. Hu, R.C. Izaurralde, R.A.C. Jones, S.-H. Kim, A.D.B. Leaky, K. Lewers, T.L. Mader, A. McClung, J. Morgan, D.J. Muth, M. Nearing, D.M. Oosterhuis, D. Ort, D., C. Parmesan, W.T. Pettigrew, W. Polley, R. Rader, C. Rice, M. Rivington, E. Rosskopf, W.A. Salas, L.E. Sollenberger, R. Srygley, C. Stöckle, E.S. Takle, D. Timlin, J.W. White, R. Winfree, L.W. Morton, and L.H. Ziska. 2012. Climate Change and Agriculture in the United States: Effects and Adaptation. USDA Technical Bulletin 1935. Washington, DC: USDA.
- Watts N., W.N. Adger, P. Agnolucci, J. Blackstock, P. Byass, W. Cai, S. Chaytor, T. Colbourn, M. Collins, A. Cooper, P.M. Cox, J. Depledge, P. Drummond, P. Ekins, V. Galaz, D. Grace, H. Graham, M. Grubb, A. Haines, I. Hamilton, A. Hunter, X. Jiang, M. Li, I. Kelman, L. Liang, M. Lott, R. Lowe, Y. Luo, G. Mace, M. Maslin, M. Nilsson, T. Oreszczyn, S. Pye, T. Quinn, M. Svendsdotter, S. Venevsky, K. Warner, B. Xu, J. Yang, Y. Yin, C. Yu, Q. Zhang, P. Gong, H. Montgomery, and A. Costello. 2015. Health and climate change: Policy responses to protect public health. *Lancet* 386(10006):1861–1914 (published online June 22, 2015).
- Whitlock, M.C., M.A. McPeck, M.D. Rausher, L. Rieseberg, and A.J. Moore. 2010. Data archiving. *American Naturalist* 175:145–146.
- Wilkinson, M.D., M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, and J. Bouwman. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3(1):160018.