Robust spatial frameworks for leveraging research on sustainable crop intensification

Patricio Grassini[a,*], Cameron M. Pittelkow[b], Kenneth G. Cassman[a], Haishun S. Yang[a], Sotirios Archontoulis[c], Mark Licht[c], Kendall R. Lamkey[c], Ignacio A. Ciampitti[d], Jeffrey A. Coulter[e], Sylvie M. Brouder[f], Jeffrey J. Volenec[f], Noemi Guindin-Garcia[g]

[a] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, P.O. Box 830915, Lincoln, NE 68583-0915, USA
[b] Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA
[c] Department of Agronomy, Iowa State University, Ames, IA 50011-1010, USA
[d] Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA
[e] Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108-6026, USA
[f] Department of Agronomy, Purdue University, West Lafayette, IN 47907-2054, USA
[g] USDA-NASS-Research and Development Division, Washington DC 20250, USA

ARTICLE INFO

ABSTRACT

Meeting demand for food, fiber, feed, and fuel in a world with 9.7 billion people by 2050 without negative environmental impact is the greatest scientific challenge facing humanity. We hypothesize that this challenge can only be met with current and emerging technologies if guided by proactive use of a broad array of relevant data and geospatial scaling approaches to ensure local to global relevance for setting research priorities and implementing agricultural systems responsive to real-time status of weather, soils, crops, and markets. Despite increasing availability of field-scale agricultural data, robust spatial frameworks are lacking to convert these data into actionable knowledge. This commentary article highlights this knowledge gap and calls attention to the need for developing robust spatial frameworks that allow appropriate scaling to larger spatial domains by discussing a recently developed example of a data-driven strategy for estimating yield gaps of agricultural systems. To fully leverage research on sustainable intensification of cropping systems and inform policy development at different scales, we call for new approaches combining the strengths of top-down and bottom-up approaches which will require coordinated efforts between field scientists, crop modelers, and geospatial researchers at an unprecedented level.

1. Text

A fundamental challenge facing agriculture is to address crop productivity gains and environmental quality concomitantly. Crop yield gains must accelerate to reduce pressure to convert natural ecosystems into farmland (Tilman et al., 2002, 2011; Cassman et al., 2003). Such conversion accounts for about 15% of anthropogenic greenhouse gas (GHG) emissions (Burney et al., 2010; Vermeulen et al., 2012) and much of the global biodiversity loss (IUCN, 2014; Laurance et al., 2014; Watson et al., 2016). However, the rate of crop yield increase is slowing or stagnating in many of the world's most productive regions which, in turn, has encouraged massive expansion of crop production area at the highest rate in all of human history (Lin and Huybers, 2012; Grassini et al., 2014). Rising demand for food, livestock feed, and biofuels coupled with global climate change are also putting increasing pressure on freshwater resources (Falkenmark et al., 1998; Rosegrant et al., 2009). Meanwhile, there is increasing concern about the impact of modern farming practices on natural resources including water quantity and quality, wildlife and biodiversity, greenhouse gas emissions, and soil and air quality (Vitousek et al., 1997; Linquist et al., 2012).

Given the diversity of environments where crop production takes place, we argue that it is inefficient to conduct research studies dealing with sustainable crop intensification without a robust framework to synthesize and upscale results to larger spatial scales while still ensuring local relevance. Here we provide an example of a strategy that not only advocates for 'boots on the ground' and 'white-peg' field experiments, but also addresses the urgent need for methods that allow appropriate scaling to larger spatial domains using frameworks specifically designed for their relevance and accuracy in predicting and

evaluating the performance of agricultural systems.

Most studies to date dealing with food security, agriculture's environmental footprint, and the impacts of climate change can be roughly grouped into two categories. One category includes an enormous and growing body of literature of studies focused on specific locations or small regions without means to identify the spatial "inference domain" for which the work is relevant. This type of field study represents the core of agronomic research, and while this approach has produced insights at a local level, limited efforts have been made to upscale these results to quantify their regional and global significance or to synthesize these results (e.g., with statistical meta-analysis) so that their collective inference space can be gleaned (Brouder and Gomez-Macpherson, 2014; Pittelkow et al., 2015).

A second category of studies focuses on regional to global scales using a top-down approach largely based on a gridded spatial framework for data on climate, soils, and crop production (Fig. 1, upper panels). Examples of such an approach are the platforms used to simulate production and environmental outputs from cropping systems models such as DSSAT and APSIM (Elliott et al., 2014; Gbegbelegbe et al., 2016). While useful to detect general global and regional trends, top-down approaches are less accurate at the spatial scale at which agricultural decisions are made (i.e., the field-scale) as a result of the coarse underpinning weather and soil data inputs and weak assumptions on cropping system context, and outcomes are difficult to validate (van Ittersum et al., 2013; Van Wart et al., 2013a; Grassini et al., 2015; Mourtzinis et al., 2016). In summary, existing spatial frameworks are inadequate because they were not designed to explicitly assess the performance of agricultural systems across different spatial scales while ensuring local to global relevance.

With the increasing trend of big data analytics[1] and prescription agriculture services, the issue of scale is increasingly being addressed in the agricultural science community, yet many opportunities exist for improvement. While there are some important earlier examples of spatial frameworks developed for integrated assessment (e.g., Van Diepen et al., 1991, Rötter et al., 2005, Shirsat et al., 2016), these previous efforts were often different in scope looking at broader issues involving regional land use, climate change, crop production, and the environment. In contrast, the focus of this article is towards the development of a spatial framework that is specifically designed to make use of relevant, accurate local weather, soil, and cropping system data that can be used to help strategize agricultural research and development for sustainable intensification of crop production systems.

A bottom-up spatial framework has the inherent advantage of local to global relevance if the upscaling protocols are robust (Fig. 1, bottom panels). The costs of implementing a bottom-up approach, however, can be too expensive and time consuming if a large number of location-specific datasets are required to achieve adequate spatial coverage. Hence, an efficient method is needed to limit the number of location-specific datasets through use of an effective method of spatial upscaling. Here the scientific challenge is to develop a bottom-up framework that identifies the minimum number of location-specific datasets required to achieve robust prediction of cropping system performance at regional, national, and global scales.

To illustrate this concept, we discuss the bottom-up spatial framework developed for the Global Yield Gap Atlas (www.yieldgap.org), which offers a complementary approach to top-down studies for research on sustainable intensification (Fig. 1, bottom panels). The approach has at its core minimum data sets that include measured weather, soil, and cropping system data for representative locations to account for the greatest proportion of total regional or national production of the crop or cropping systems being evaluated (Van

Wart et al., 2013b, 2013c; Grassini et al., 2015; van Bussel et al., 2015). Results for these locations are subsequently upscaled to soil types and climate zones at national to regional and global spatial scales. This site selection and upscaling process helps to limit the number of locations for which site-specific data on weather, soils, and cropping systems are required, which in turn facilitates the focus on quality of the underpinning data and helps ensure local to global relevance of the analysis. By starting with the most relevant crop producing areas and scaling up, this approach allows for increased data quality and relevancy, an accurate understanding of local cropping system contexts and management approaches, and the ability to validate results at the field-scale, in contrast to top-down approaches which generally aim to achieve full terrestrial coverage and necessarily rely on coarse data.

The accuracy of this bottom-up approach has recently been validated for regions where high-quality data are available. Hochman et al. (2016) conducted a study on yield gaps of rainfed wheat in Australia following two approaches: (i) the bottom-up approach of the Global Yield Gap Atlas and (ii) a data rich analysis method using high-density data available in the Australian grain zone. These researchers reported that the two methods gave similar estimates of yield potential and yield gaps at climate zone and national levels.[2] Given the high spatial environmental variability within the Australian wheat zone, the remarkable level of agreement between results derived from these two methodologies provides evidence of the robust estimates provided by the spatial framework of the Global Yield Gap Atlas. Similarly, Aramburu-Merlos et al. (2015) and Morell et al. (2016) have shown that estimates of national average yields for Argentina and USA, calculated based on a limited number of selected locations which were upscaled to country level following the protocols of the Global Yield Gap Atlas, were remarkably similar to the reported national average yield based on data from hundreds of subnational-level administrative units covering the entire crop production area. Finally, Van Wart et al. (2013b) and van Bussel et al. (2015) showed that variability in weather and simulated yield potential was relatively low for sites located within same climate zones, which provides further support for a stratified (instead of random) selection of sites and use of the climate zone scheme as basis for upscaling results from location to region and country.

An inherent limitation of using a bottom-up approach is to leave out marginal or 'frontier' agricultural environments, which may not be relevant in terms of total food production but can be important relative to the environmental footprint of agriculture and climate change. To account for these regions, further development of a bottom-up approach is needed to capture both major and minor environments where crop production takes place so that productivity and environmental performance of agricultural systems can be evaluated more generally, informing strategic investments in agriculture and policy decisions. We believe that this can realistically be implemented by adding more sites, in addition to those selected based on their contribution to total national area. In other words, while the protocol designed by the Global Yield Gap Atlas sets a minimum threshold relative to the number of sites, the list can be extended to include other locations that are important for additional reasons besides crop production.

Another limitation of the bottom-up approach is the difficulty to be applied in crop production areas where high-quality data are not available, either because the required data do not exist or are not publicly available. This can be overcome by providing the most appropriate alternatives in a transparent manner. The methodology developed by the Global Yield Gap Atlas consists of a tiered approach

---

[1] Big data for agriculture includes geospatial data on soil properties, long-term weather data with a daily time step, short- and medium-term weather forecasts, and crop management practices over the recent past and in the current cropping season, all with fine spatial resolution required for decision making at local to global scales.

[2] Yield potential depends on solar radiation, temperature, and water supply during the crop growing season and can be calculated for both rainfed (water-limited yield potential) and irrigated conditions. The yield gap is defined as the difference between yield potential and farmer actual yield (van Ittersum et al., 2013).