ARTICLE IN PRESS
AGSY-02187; No of Pages 13
Agricultural Systems xxx (2016) xxx–xxx
Contents lists available at ScienceDirect
Agricultural Systems
journal homepage: www.elsevier.com/locate/agsy

Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology

Sander J.C. Janssen [a,*], Cheryl H. Porter [b], Andrew D. Moore [c], Ioannis N. Athanasiadis [a], Ian Foster [d], James W. Jones [b], John M. Antle [e]

[a] Wageningen University and Research Centre, Postbus 47, 6700AA Wageningen, Netherlands
[b] Agricultural & Biological Engineering, University of Florida, PO Box 110570, Gainesville, FL 32611, USA
[c] CSIRO Agriculture & Food, GPO Box 1600, Canberra, ACT, 2601, Australia
[d] Computation Institute, University of Chicago, IL, USA
[e] Oregon State University, Corvallis, OR, USA

ARTICLE INFO

Article history:
Received 20 November 2015
Received in revised form 14 September 2016
Accepted 21 September 2016
Available online xxxx

Keywords:
Agricultural models
ICT
Linked data
Big data
Open science
Sensing
Visualization

ABSTRACT

Agricultural modeling has long suffered from fragmentation in model implementation. Many models are developed, there is much redundancy, models are often poorly coupled, model component re-use is rare, and it is frequently difficult to apply models to generate real solutions for the agricultural sector. To improve this situation, we argue that an open, self-sustained, and committed community is required to co-develop agricultural models and associated data and tools as a common resource. Such a community can benefit from recent developments in information and communications technology (ICT). We examine how such developments can be leveraged to design and implement the next generation of data, models, and decision support tools for agricultural production systems. Our objective is to assess relevant technologies for their maturity, expected development, and potential to benefit the agricultural modeling community. The technologies considered encompass methods for collaborative development and for involving stakeholders and users in development in a transdisciplinary manner.
Our qualitative evaluation suggests that as an overall research challenge, the interoperability of data sources, modular granular open models, reference data sets for applications and specific user requirements analysis methodologies need to be addressed to allow agricultural modeling to enter in the big data era. This will enable much higher analytical capacities and the integrated use of new data sources. . Overall agricultural systems modeling needs to rapidly adopt and absorb state-of-the-art data and ICT technologies with a focus on the needs of beneficiaries and on facilitating those who develop applications of their models. This adoption requires the widespread uptake of a set of best practices as standard operating procedures.

1. Introduction

Information and computer technology (ICT) is changing at a rapid pace. Digital technologies allow people to connect across the globe at high speeds at any time (Gartner, 2016). Even those in remote, developing regions increasingly have the ability to connect online via telephone and Internet providers (Danes et al., 2014). Satellite and drone capabilities can provide remotely sensed data in real-time regarding in-season crop growth and development, soil moisture, and other dynamic variables (e.g. Capolupo et al., 2015). High performance computing can be used to process large amounts of data in a short time frame , to make sense of large quantities of structured and unstructured data (i.e. "big data"; NESSI, 2012) collected using new sensing technologies, and to scale and validate models in ways not previously possible. Web and cloud technologies permit these capabilities to be made available to large numbers of end users with a convenience and cost that was previously inconceivable (Foster, 2015). As a result of these and other developments, society expects more and higher-quality information to be available in support of daily decision-making.

Our enthusiasm for these new technologies in the agricultural sciences must be tempered by a realization that our modeling and decision support systems have not kept up with technology. Indeed, many frameworks used in these systems date back to the 1970s through the 1990s, prior to the availability of today's advanced data collection, computing, storage, access, processing technologies, software languages and coding standards. Thus, we see two distinct opportunities for applying modern ICT to agricultural systems modeling. First, advances such as big data, crowdsourcing (i.e. sourcing data and information through distributed networks of respondents), remote sensing, and high

* Corresponding author.
E-mail addresses: Sander.Janssen@wur.nl (S.J.C. Janssen), cporter@ufl.edu, jimj@ufl.edu (C.H. Porter), Andrew.Moore@csiro.au (A.D. Moore), ioannis.athanasiadis@wur.nl (I.N. Athanasiadis).

http://dx.doi.org/10.1016/j.agsy.2016.09.017
0308-521X/© 2016 Published by Elsevier Ltd.

performance computing can be used to advance the *science* of agricultural systems modeling. Second, new technologies can be used to transform the *practice and application* of agricultural systems modeling by making it far more collaborative, distributed, flexible, and accessible. As clearly shown in a recent thematic issue of *Environmental Modeling and Software* (Holzworth et al., 2014a; Athanasiadis et al., 2015), the *science* of agricultural systems modeling is progressing steadily and adopting various new ICT technologies to advance the science on a case-by-case basis. However, *the practice and application* of agricultural systems modeling is not progressing as fast, leading to lack of applications using agricultural systems models. Thus, an important feedback from *application* to *science* is absent and needs to be established, as also discussed by Holzworth et al. (2015) for cropping systems models. The focus of this review is thus not on relevant ICT technologies for the modeling scientist working at a university or research institute, but on ways to facilitate the involvement of actors beyond the academy. As discussed in the companion article by Antle et al. (this issue), the result of achieving such involvement will be a "next generation" modeling community (NextGen) that includes not only modelers and model developers working across disciplines, spatial scales, and temporal scales to exploit new data sources and to produce and apply new models, but also software developers to produce the NextGen modeling frameworks, data processing applications, and visualization tools.

In this paper, we approach the envisioned NextGen of agricultural models and the supporting modeling community from the ICT perspective. Our objective is to assess relevant technologies for their maturity, expected development, and potential to benefit the agricultural modeling community. The technologies considered encompass methods for collaborative development and for involving stakeholders and users in development in a transdisciplinary manner.

We assess recent ICT developments through five use cases that have been formulated to support the vision for a NextGen modeling community (see also the introductory overview by Antle et al. (submitted, a) and accompanying papers by Jones et al. (submitted) and Antle et al. (submitted, b)).

A NextGen of applications based on agricultural systems modeling can help companies, governments, and farmers in the food chain to make informed decisions. The concepts of knowledge chain (Fig. 1) and application chains (Fig. 2) provide complementary perspectives on the value and positioning of modeling in the broader context of decision making and ICT, and are used to loosely organize the content of this paper. A **knowledge chain** is a set of linked steps by which data are processed into information, knowledge and finally wisdom as used in decision making. This perspective postulates that data comprise a raw material that, when combined with description and quality attributes, leads to information. Information can be linked to other information sources and placed in causal chains to produce knowledge. Ultimately, knowledge serves as an input for decisions based on wisdom, which cannot be digitized and which exists in the mind of a decision-maker. A second perspective focuses on **application chains**. Agricultural models must be engaged in an infrastructure consisting of both software (e.g. in layers of data access, processing, analysis and visualization) and hardware (i.e., servers, computing capacity, and storage) as depicted in Fig. 2. Based on the data in the infrastructure, applications targeted at end-users serve information and knowledge, e.g. a yield forecast to a supply chain manager; effects on farm income of a policy change; estimates of disease related crop damage to a farmer. Application chains may be simple or complex, and may include, for example, data access, extraction, transformation (e.g. summarization or interpolation), and integration operations; one or multiple models; integration of output from different models; and model output transformation, analysis, and visualization steps. Design of the application chains must consider not only the end-users, but the full spectrum of users of NextGen ICT infrastructure including primary data collectors, database professionals, software developers, modelers and the end-users of knowledge and information.

This paper is organized according to the different layers of the knowledge chain and the covering actors and elements of the application chains, focused on the NextGen of agricultural systems model applications. Section 2 introduces the users and the use cases where NextGen models could potentially play a large role. Section 3 describes the actors active along the application chain. Sections 4 and 5 cover developments in data as the raw material of any modeling effort and new technologies that can assist with the presentation of modeling analyses to users. In Section 6, we discuss how agricultural models (in the strict sense) will need to be developed and implemented to enhance their usefulness in concert with parallel ICT advances, and in Section 7 we discuss the importance of new interfaces from software to users and

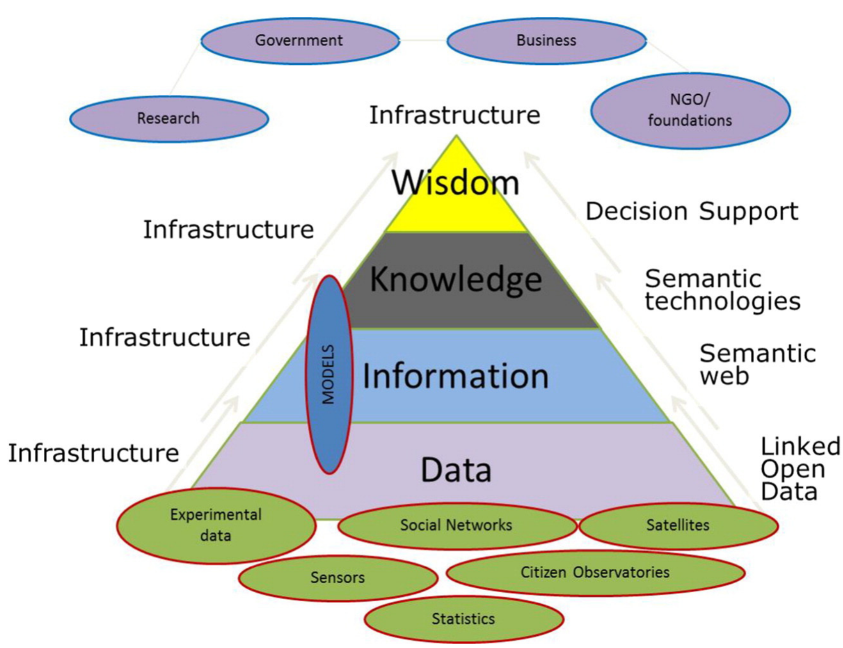

Fig. 1. Knowledge Pyramid linking data to information to knowledge and wisdom, in which data is the raw material for the development of applications addressing decision making through wisdom in research, government, business and ngo/foundations (adapted from Lokers et al., 2016).