Environmental Modelling & Software 52 (2014) 264–272

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Development of an intelligent environmental knowledge system for sustainable agricultural decision support

Ritaban Dutta [a,*], Ahsan Morshed [a], Jagannath Aryal [b], Claire D'Este [a], Aruneema Das [c]

[a] Intelligent Sensing and Systems Laboratory, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Hobart 7001, Australia
[b] School of Geography and Environmental Studies, University of Tasmania, Hobart 7001, Australia
[c] School of Engineering, University of Tasmania, Hobart 7001, Australia

ARTICLE INFO

Article history:
Received 9 May 2013
Received in revised form
1 October 2013
Accepted 3 October 2013
Available online 15 November 2013

Keywords:
Feature
Semantic matching
Knowledge integration
i-EKbase
Linked Open Data cloud

ABSTRACT

The purpose of this research was to develop a knowledge recommendation architecture based on unsupervised machine learning and unified resource description framework (RDF) for integrated environmental sensory data sources. In developing this architecture, which is very useful for agricultural decision support systems, we considered web based large-scale dynamic data mining, contextual knowledge extraction, and integrated knowledge representation methods. Five different environmental data sources were considered to develop and test the proposed knowledge recommendation framework called Intelligent Environmental Knowledgebase (i-EKbase); including Bureau of Meteorology SILO, Australian Water Availability Project, Australian Soil Resource Information System, Australian National Cosmic Ray Soil Moisture Monitoring Facility, and NASA's Moderate Resolution Imaging Spectroradiometer. Unsupervised clustering techniques based on Principal Component Analysis (PCA), Fuzzy-C-Means (FCM) and Self-organizing map (SOM) were used to create a 2D colour knowledge map representing the dynamics of the i-EKbase to provide "prior knowledge" about the integrated knowledgebase. Prior availability of recommendations from the knowledge base could potentially optimize the accessibility and usability issues related to big data sets and minimize the overall application costs. RDF representation has made i-EKbase flexible enough to publish and integrate on the Linked Open Data cloud. This newly developed system was evaluated as an expert agricultural decision support for sustainable water resource management case study in Australia at Tasmania with promising results.

Software and/or data availability

NASA MODIS data is publicly available from NASA
ASRIS is publicly available data base from CSIRO, Australia
Jena framework for data translation http://jena.apache.org/
Sesame TripleStore http://www.openrdf.org/-
Pubby tool is Linked Open Data front end tool
http://wifo5-03.informatik.uni-mannheim.de/pubby/

1. Problem space and motivation

The ultimate challenge of an environmental and agricultural decision support system is to overcome uncertainty associated with the data quality, to cross validate the knowledge automatically, and to improve the efficiency of the decision making process. This research study proposed a knowledge integration platform (Knapen et al., 2013; Ren et al., 1988; Sheth, 2012; Henson et al., 2012) and machine learning analysis based recommendation architecture (Intelligent Environmental Knowledgebase (i-EKbase)) (Morshed et al., 2013a,b; Gilbert et al., 2010) to represent the knowledge in a more meaningful way. Knowledge integration and expert knowledge representation for sustainable agricultural decision support systems (SADSS) is an ever changing domain problem for environmental modelling and associated software system development. Often sensory system and model integrations do not reflect the natural fluidity (Voinov and Shugart, 2013) of the environment hence create significant performance limitations. In many cases SADSS also does not reflect or satisfy the end user requirements due to lack of domain knowledge capturing (McIntosh et al., 2011). These early works have motivated this research project to consider heterogeneous data sources and domain information systems as dynamic software modules before the knowledge integration to capture the natural dynamics of the environment in a more realistic

* Corresponding author. Tel.: +61 3 6232 5423; fax: +61 2 6232 5229.
E-mail addresses: ritaban_dut@yahoo.com, ritaban.dutta@csiro.au (R. Dutta).

1364-8152/$ – see front matter
http://dx.doi.org/10.1016/j.envsoft.2013.10.004

way possible. Automated multi-scale knowledge integration in a software system for emergency management has been reported (Akbar et al., 2013). It is also evident from literature that awareness and handling of uncertainty in SADSS could only be addressed by using multi-source multi-criteria knowledge integration and recommendation approach (Nino-Ruiz et al., 2013; Arnold, 2013). These previous reporting motivated the main focus of this paper to make the environmental knowledge more robust and trustworthy in a decision support context applicable to agricultural design. The application of i-EKbase to sustainable water resource management was studied in this paper as a proof of concept. In Australia, water usage for irrigation and associated electricity costs are extremely high and provides a compelling case study for the i-EKbase. Selection of this generic irrigation water management case study was inspired by the fact that there is a strong requirement for demonstrating the effectiveness of multi-source multi-scale spatio-temporal environmental knowledge integration and usefulness of such approach for much wider performance improvement in various environmental decision support systems.

Five different environmental data sources were considered for the development of i-EKbase, namely, SILO (SILO Website), AWAP (AWAP Website), CosmOz (COSMOZ Website), ASRIS (ASRIS Website), and MODIS (MODIS Website) (Fig. 1).

The Long Paddock SILO database is operated by the Queensland Climate Change Centre of Excellence (QCCCE) from the Australian Department of Science, Information Technology, Innovation and the Arts (DSITIA). The Australian Water Availability Project (AWAP) database is developed to monitor the state and trend of the terrestrial water balance of the Australian continent, using model-data fusion methods to combine both measurements and modelling. The Australian Soil Resource Information System (ASRIS) database provides online access to the best publicly available information on soil and land resources in a consistent format across Australia. The Australian Cosmic Ray Sensor Network (CosmOz)

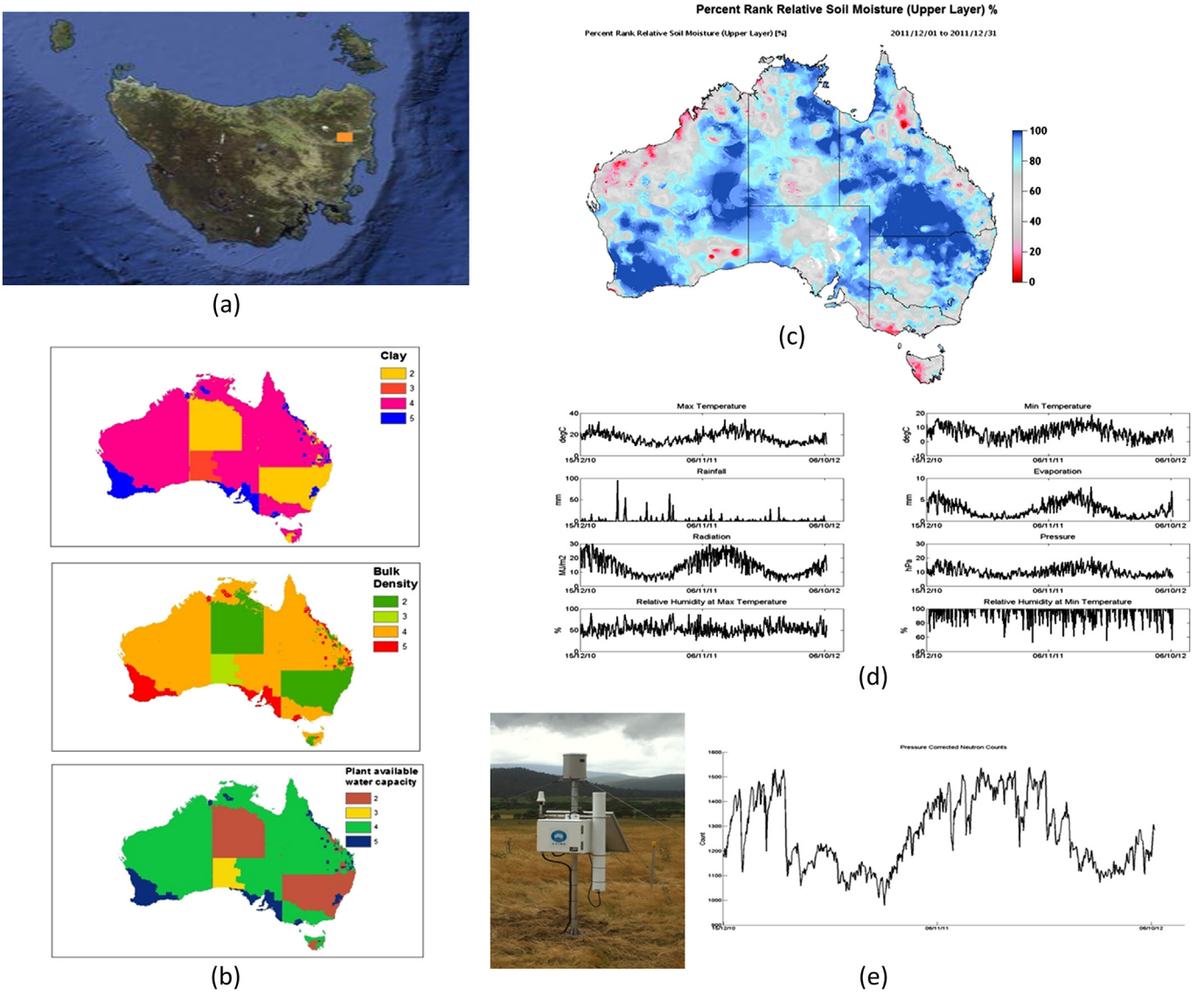

Fig. 1. (a) Tullochgorum, Tasmania is represented on a MODIS image as the selected study location for this study, (b) ASRIS maps showing clay content, bulk density, and the plant available water capacity for Australia. A web adaptor was used to extract static data for a particular location, (c) Examples of AWAP gridded map data, (d) Example of SILO data time series extracted using dynamic web adaptor, (e) CosmOz Hydroinnova CRS-1000 cosmic ray soil moisture probe based neutron count data relating to area-average soil moisture over its horizontal footprint.