Mobile Application Based Sustainable Irrigation Water Usage Decision Support System: An Intelligent Sensor CLOUD Approach

Cecil Li Ritaban Dutta Corné Kloppers
Claire D'Este Ahsan Morshed Auro Almeida

Commonwealth Scientific and Industrial Research Organisation (CSIRO), Hobart, Australia 7001

Aruneema Das Jagannath Aryal

University of Tasmania
Hobart, Australia 7001

Abstract— **In this paper a novel data integration approach based on three environmental Sensors – Model Networks (including the Bureau of Meteorology-SILO database, Australian Cosmic Ray Sensor Network database (CosmOz), and Australian Water Availability Project (AWAP) database) has been proposed to estimate ground water balance and average water availability. An unsupervised machine learning based clustering technique (Dynamic Linear Discriminant Analysis (D-LDA)) has been applied for extracting knowledge from the large integrated database. The Commonwealth Scientific and Industrial Research Organisation (CSIRO) Sensor CLOUD computing infrastructure has been used extensively to process big data integration and the machine learning based decision support system. An analytical outcome from the Sensor CLOUD is presented as dynamic web based knowledge recommendation service using JSON file format. An intelligent ANDROID based mobile application has been developed, capable of automatically communicating with the Sensor CLOUD to get the most recent daily irrigation, water requirement for a chosen location and display the status in a user friendly traffic light system. This recommendation could be used directly by the farmers to make the final decision whether to buy extra water for irrigation or not on a particular day.**

I. INTRODUCTION

Water shortage is a severe issue in many areas of Australia, particularly in the agricultural sector which is responsible for 65% of the total water usage. Agricultural water usage is a very expensive process in Australia due to the lack of water and highly priced electricity to run the irrigators. Thus, a short-term low cost solution is needed to provide reliable advices on irrigation planning for the farmers so that the water wastage is minimized. In Australia, much of the environmental degradation is associated with the changes in the near-surface water balance induced by massive clearing of native vegetation and deforestation. These artificial changes have led to significant increases in groundwater re-charge, which in turn have led to rising water tables and salinisation. Other important aspect of the hydrological system is vegetation. Vegetation influences the hydrological cycle through the exchange of energy, water, carbon and other substances and is therefore critical for many hydrological processes, in particular transpiration, infiltration and runoff. The movement of water through the hydrological cycle varies significantly in both time and space. Australia, the driest continent, has the highest variability in rainfall and runoff and is therefore a difficult system to model [1-2].

This paper presents a serious problem for sustainable agricultural development in the region because there is no reliable surface water for irrigation. Australian farmers must decide how much water to buy and use for their potential future irrigation and crop growth. Water usage for irrigation and associated electricity costs are extremely high in Australia. Efficient and timely decision support regarding the sustainable management of water usage is essential. One way to overcome the problem is to combine field experiments with the conventional water balance modelling. But field experiments are expensive and only a limited number of land-use options could be trialed. So generalisation of the water balance estimation is near impossible depending only on field experiments. There is a need for on demand complementary knowledge integration from multiple data sources to increase efficiency and automatic interpretation of the knowledge.

Even though regular weather data is available to the farmers ultimately farmers use their experience to make water management decisions, as it’s extremely hard to interpret multivariate data. There is genuine need for multi source data from sensors and models, as well as knowledge integration to

The authors wish to thank the Intelligent Sensing and Systems Laboratory and the Tasmanian node of the Australian Centre for Broadband Innovation. A grant from the Tasmanian Government which is administered by the Tasmanian Department of Economic Development, Tourism and the Arts helped to conduct this research.

978-1-4673-4642-9/13/$31.00 ©2013 Crown

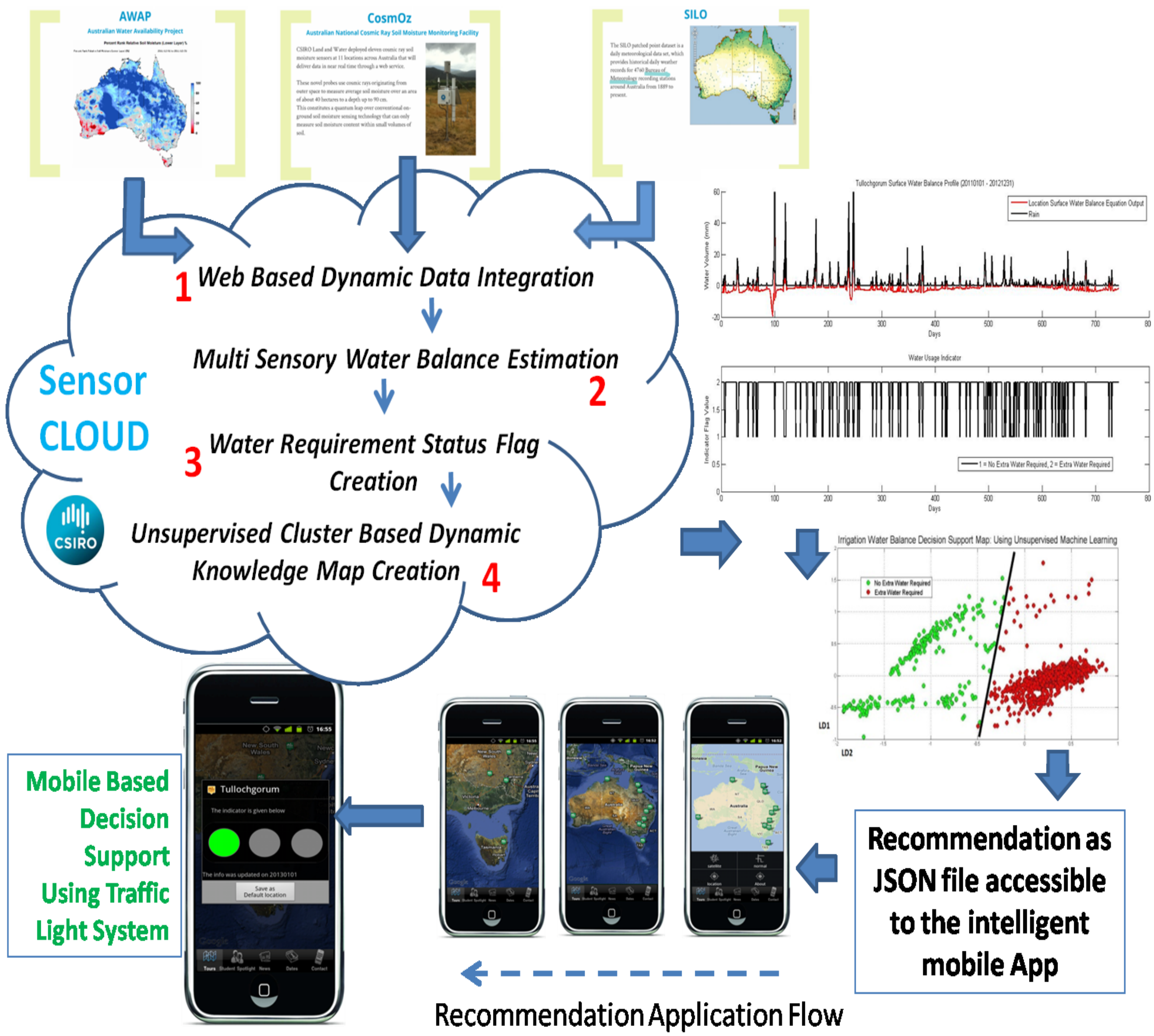

Fig. 1: CSIRO Sensor CLOUD based multi sensory knowledge integration architecture and ANDROID mobile based recommendation framework for the irrigation water balance decision support system. Unsupervised machine learning clustering technique LDA has been used for the traffic light indicator system.

provide better decision support for the farmers based on strong scientific foundation rather than intuition.

In this paper, a novel knowledge integration and machine learning analysis based water usage recommendation system has been investigated and proposed based on the CSIRO Sensor CLOUD. The ultimate challenge in environmental forecasting and decision support systems is to overcome the uncertainty associated with the data quality, to cross validate the knowledge automatically, and to make the decision making process more efficient. Here, we have also proposed and developed an ANDROID mobile phone based recommendation framework (see Fig. 1) to represent the dynamic water availability for any farming area in Australia.

II. KNOWLEDGE INTEGRATION ON CLOUD

Based on any given location information (latitude and longitude) the nearest available Bureau of Meteorology (BOM) weather station was selected based on geographical distance. A SILO data file was downloaded and processed for the chosen station. The nearest available CosmOz (Cosmic ray based soil moisture measuring sensor network) data was also downloaded for the selected station. The AWAP database was connected through a secured FTP server and grid files were downloaded locally. For the same location a pixel position was derived on the daily continental AWAP gridded data and time series were extracted for individual variable for a given time frame.

A unified knowledge integration and representation model was developed using the unified Resource Description Framework (RDF). Unified knowledge RDFs were created for all of the three data sources based on pre-processed data, available Meta data, and original provenance information. All RDFs were integrated in the next stage. The main purpose of this RDF based approach was to store the processed knowledge on a triplestore framework. A triplestore is a framework used for storing and querying RDF data. It provides a mechanism for persistent storage and access of RDF graphs. Recently, there has been a major development initiative in query processing, access protocols and triple-store technologies. The integrated knowledge representation system was developed using a triple called "Sesame triplestore". Sesame is an open source framework for storage inference and querying of RDF data. This integrated system was incorporated on the CSIRO CLOUD infrastructure.

A. *Experimental Study*

Data from Daly River (131.4°E, 14.2°S) was integrated and processed for this study. This location was selected to induce significant data variance in the generalization experiments as geographically they are quite different in nature. Daly River is a tropical savannah. The time period of the experiment was 01/01/2011 – 31/12/2012. The study was conducted on a daily basis.

III. CLOUD BASED WATER BALANCE ESTIMATION

A system can be one of several hydrological domains, such as a column of soil, or a drainage basin. Water balance can also refer to the ways in which an organism maintains water in dry or hot conditions. It is often discussed in reference to plants or arthropods, which have a variety of water retention mechanisms, including a lipid waxy coating that has limited permeability (Fig. 2) [2].

Water balance is based on the law of conservation of mass: any change in the water content of a given soil volume during a specified period must equal the difference between the amount of water added to the soil volume and the amount of water withdrawn from it. The root zone water balance is usually expressed as:

$$\Delta S = P - E - T - RO - DD \tag{1}$$

Where ΔS is the change in root zone soil water storage over the time period of interest, P is precipitation that get to the soil, E is direct evaporation from the soil and water body surface, T is transpiration by plants and grass, RO is surface runoff or overland flow, and DD is deep drainage out of the root zone. All attributes are expressed in terms of volume of water per unit land area or equivalent depth of water over the period considered. The recharge to the groundwater system can be calculated as:

$$RO = DD - SSF \tag{2}$$

Where SSF is the lateral subsurface flow calculated according to Darcy's law. When the control volume is the entire catchment represented by given latitude and longitude information, the surface water balance equation can be expressed as:

$$\Delta{<}S{>} = {<}P{>} - {<}ET{>} - {<}Q{>} - {<}R{>} \tag{3}$$

Where $\Delta{<}S{>}$ the change in spatially averaged catchment water storage is, ${<}P{>}$ is spatially averaged precipitation, ${<}ET{>}$ is the spatially averaged catchment evapotranspiration, ${<}Q{>}$ is the spatially averaged catchment surface runoff, and ${<}R{>}$ is the spatially averaged catchment recharge. Equation 3 was used for the historic water balance calculation. A multi data source based "Water Balance Ontology" (Fig.2) was created represent this model in a machine readable structure.

IV. IRRIGATION WATER USAGE FLAG

The irrigation water usage indicator was calculated for the whole duration of the research experiment. Historically there

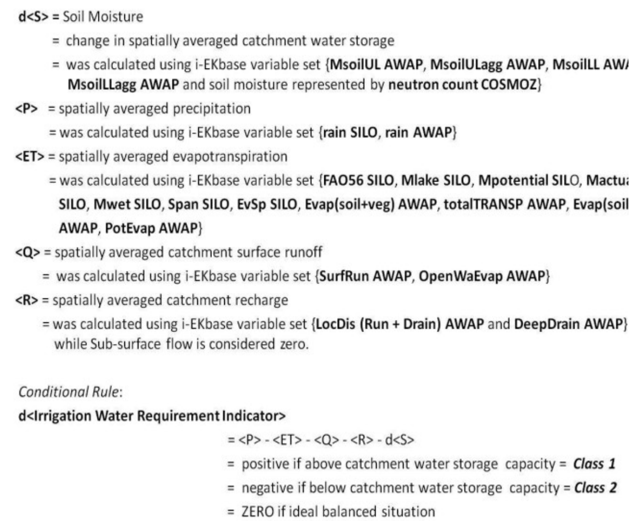

Fig. 2: Flow of environmental attributes integration and irrigation water usage flag generation on Sensor CLOUD. The CLOUD based Intelligent Environmental Knowledgebase (i-EKbase) system has continuously been used to estimate daily water balance dynamically.

were only two possible water management decisions that one farmer could take on a given day - buy extra water for crop irrigation, or there was enough water in soil that they did not need to. Positive results from the CLOUD based water balance calculation system indicated enough water in the soil (represented by Class 1 or '1' values) whereas negative results indicated the necessity of buying water (represented by Class 2 or '2' values). A new time series was created to represent the variance of irrigation water requirement indicator over two complete years. Fig. 3 shows the irrigation indicator profile for the Daly River.

V. LINEAR DISCRIMINANT CLUSTERING

Linear discriminant analysis (LDA) and the related Fisher's linear discriminant are methods used in statistics, pattern recognition and machine learning to find a linear combination of features, which characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier or, more commonly, for dimensionality reduction before later classification [3].

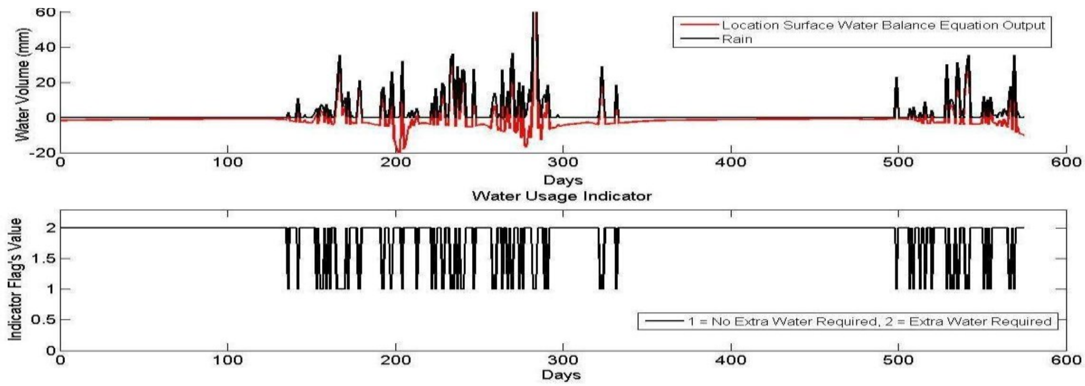

Fig. 3: Irrigation water usage flag generation on Sensor CLOUD. The CLOUD based Intelligent Environmental Knowledgebase (i-EKbase) system.

Based on the hydrological water balance calculation, a daily water requirement status indicator was generated and used as class boundaries between two types of irrigation scenarios, namely, "Extra water required" and "No Extra water required". D-LDA was applied to cluster the historical

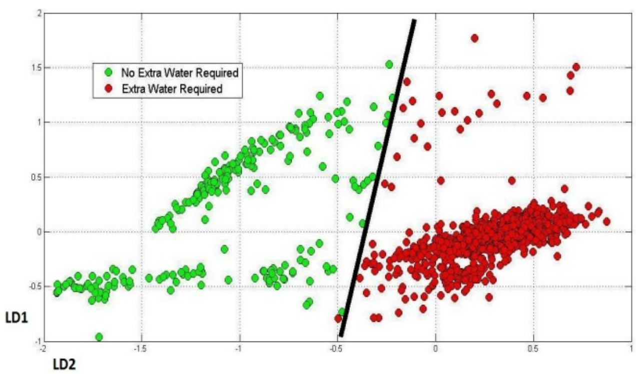

Fig. 4: D-LDA based unsupervised cluster knowledge map for irrigation water balance decision support system on the sensor CLOUD.

sensory data according to these two categories to create a water availability decision support knowledge map (see Fig. 4). Overall accuracy of this newly proposed framework was measured using a two year long historical data set. 50% data the historical data set was used for creating the two clusters based unsupervised dynamic knowledge map representing two scenarios. The remaining 50% was used to test the correctness

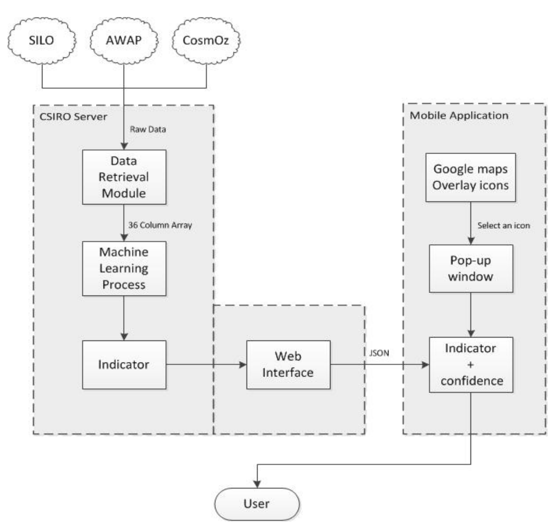

Fig. 5: The sensor CLOUD guided ANDROID based mobile application system for dynamic water requirement decision support system.

of the map. Daily input variables were represented as a point on the two clusters based map. Clusters were coloured according to the water balance calculation, where green cluster represented "No Extra water required" and the scenario for "Extra water required" was represented by the red cluster. Distances from those two clusters' center was calculated for the data points representing testing set. The minimum distance guided the most probable scenario classification. A knowledge map was updated daily depending on the new data added to the historical data (two years time window was used at any given time). Classification performance was encouraging with 92.3% overall accuracy with 93% sensitivity and 94% specificity [1-4].

VI. ANDROID MOBILE WATER APP

The LDA based classification of the newly integrated multi source based daily stream data into 'Class 1' or 'Class 2' is main building block of the decision support system. Classification results are then passed to a JSON file for the ANDROID app to access directly from the CLOUD and visualize on the mobile as a traffic light system (Fig.5). This process is available on demand. As the unsupervised data analytics and decision support is conducted on the CLOUD, mobile application is used for display purposes only.

VII. CONCLUSION

This study demonstrates how intelligent data analytics on CLOUD computing infrastructure and an associated mobile application can be applied in the irrigation domain; however the same approach has the potential for wide use in agricultural decision support and natural resource management.

REFERENCES

[1] L. Zhang , "Estimating episodic recharge under different crop/pasture rotations in the Mallee region. 1. Experiments and model calibration", Agricultural Water Management, 42, pages 219–235, 1999.

[2] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, "Cloud computing—The business perspective", Decision Support Systems, 51(1), 176-189.

[3] Linear discriminant analysis (August 2013). https://en.wikipedia.org/wiki/Linear_discriminant_analysis

[4] R Dutta, A Morshed, "Performance Evaluation of South Esk Hydrological Sensor Web: Using Machine Learning and Semantic Linked Data Approach", Accepted in Special Issue of IEEE Sensor Journal on IoT, May 2013.