18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Irrigation Modelling Language for decision support

Car, N.J. [1,2,4], **Hornbuckle, J.W.**[1,2], **Christen, E.W.** [1,2,3,4] **and Moore G.A.**[1]

[1] University of Melbourne, Aust., [2]CSIRO Land & Water, Griffith, NSW, Aust. [3]University of New England, Armidale NSW, Aust., [4]CRC for Irrigation Futures, Aust.
Email: nicholas.car@csiro.au

Abstract: Irrigation Decision Support Systems (DSS) have seen poor uptake in Australia for a number of reasons, one of which is the lack of their flexibility in allowing users to choose data sources they perceive as most relevant to them to generate the decision support they receive. The number of local and remote data sources available to DSS users for use in irrigation scheduling is increasing due to technological progress in both the agricultural engineering and information technology sectors. A standardisation of the outputs of data sources used by irrigation DSS and their internal processes would help to solve this. While there are already large international projects, such as Sensor Web Enablement (SWE) and WaterML, aimed at standardising sensor communication, internet data transfer and natural resource data sources, irrigation-specific standardisation work has not been undertaken. This paper makes a start on this by providing some background information on irrigation data sources and other standardisation approaches affecting the irrigation sector, defining requirements for irrigation standardisation and providing details of tests of standardised data source output.

Surveys in the 2007/2008 and the 2008/2009 irrigation seasons listed data sources that irrigators currently use for their management decisions. A large range was found especially when non-biophysical factors, such as irrigators' personal calendars, were classified as data sources. Non-biophysical factors affecting irrigation management have long been known to cause a 'gap' between the support irrigation DSS can provide and industry practice, therefore this classification of non-biophysical factors as data sources is a first step towards codifying their effects so that future DSS may incorporate them and bridge this gap. Apart from the data source range, this survey work showed there is currently no consensus among irrigators as to the total set of useful data sources with irrigators changing data source use over time and accepting new data sources. This further highlights the need for data source flexibility if irrigation DSS are to remain relevant to irrigators.

The many data sources used, or potentially used in irrigation decision making, are often heterogeneous in form (units, structure, range timestep etc) and therefore Informatics – the science of information use – needs to be considered when working to combine them. Sub disciplines of informatics include the study of data formats, standards and semantics as well as information fusion, comparison and presentation. Informatics work within these disciplines aims to allow heterogeneous data sources to be meaningfully accessed and used through systems such as DSS. Much standards work, such as SWE and WaterML, results from informatics.

In order to allow irrigation DSS to offer users choice with respect to the data sources used through them and to allow that set of data sources to be expanded as technological progress generates new ones, not only must the technical standardisation of data source outputs occur but a conceptual informatics framework that describes the data required, and the techniques used, to generate irrigation decision support advice is also required. One way in which DSS may be able to help bridge the 'gap' between science and industry is to offer truly effective flexibility that will allow both biophysical and non-biophysical data sources to be used.

We present summarised results from our surveys that indicate the range and type of current data sources used for irrigation decision making and give examples that show where other standardisation projects have catered for them and where they have not. We determine the position and scope of an Irrigation Modelling Language (IML) that can be used to further add to the other projects' standardisation efforts and fill gaps between them. After detailing these general requirements of an IML, we present a start to the formalisation of the conceptual model that is needed to underpin an IML by defining decision support processes and their requisite data sources. We show such formalisation with data mark-up allows DSS more data source choice and conclude that the wide adoption of such an IML framework would lead to more flexible irrigation DSS.

Keywords: *Irrigation, Decision Support Systems (DSS), semantic web, data fusion, Web Services*

INTRODUCTION

Decision Support Systems (DSS) are usually computer systems that aim to collect, process and present sensory, historical or calculated data to users to assist them in making decisions. The data they use comes from objects known as *data sources*, regardless of their physical make-up. Some typical DSS tasks are to; grant easier access to already accessible data, to fuse data from multiple sources into fewer, combined, sources of higher utility and to run calculations or models on data and present results. Displaying two datasets without data fusion, or combining data from some sources but not others used by irrigators, forces the decision maker to fuse that data in his own mind. All data is ultimately fused in a final decision, in our case scheduling of how much water to apply and when. Increasing the amount of data sources generally assists users in their own fusion efforts. It is important to know all the data sources, biophysical & non-biophysical, formal & informal, common & rare, that irrigators currently used to make irrigation decisions and how the data from these relate to each other to cater for DSS fusion.

The problem of irrigation DSS' inflexibility regarding data sources use is encountered by designers who wish to deliver useful information to irrigators but find the data sources they are able to provide access to are not those perceived by irrigators to be of the most utility. This mismatch results in a gap between DSS designers' and the target audience's data source use. Typically designers are able to work with biophysical data sources, such as automatic weather stations (AWS) & soil moisture (SM) probes but are unable to encapsulate less easily quantifiable biophysical data, such as plant health, or non-biophysical data, such as preferred personal working hours, as data sources and provide them through DSS even though irrigators can place much weight on them. Previous work on agricultural science/industry discrepancies such as McCown (2001), fail to recognise that non-biophysical influences on industry members' decisions can be thought of data sources and therefore do not consider that they may be encapsulated by future DSS.

Table 1: Summary of current data source use by Riverina viticulturalists

Data Source	Freq.
SM Probes	37
Shovel/auger	33
Visual check (wilting, colour, leaf angle)	24
Weather (forecast, past observations, numerical and non-numerical data)	17
Fixed schedule	17
Evapotranspiration	5
Other	2

Finally, even well known biophysical data sources producing quantifiable data in electronic form are often difficult to integrate into irrigation DSS due to the proprietary nature of their data formats.

To quantify the current data source usage of vine growers in the Riverina, we conducted detailed interviews with 68 of them about their scheduling techniques and data sources used over the 2007/2008 and 2008/2009 seasons. Table 1 summarises the data sources irrigators initially listed as being important for irrigation decisions. Of the 37 SM probes encountered, there were 7 different models used, all of which used proprietary data formats. Of the 17 irrigators listing weather sources, many used different types from a vast range of sources and many other irrigators, when questioned about weather, use indicated that they took it into account while not listing it initially as a data source.

The survey also revealed that the effort required by irrigators to access particular data sources, aside from the perceived value of their data, was a major factor for them in determining its overall utility.

After further questioning, it was apparent that irrigators could list other factors that much influenced their irrigation scheduling decision making but did not see those factors as data sources *per se*. These factors included personal calendars, system limitations on run times, water availability, manpower and off-peak electrical power times. Many treated these factors as inevitable and unalterable and as preconditions applied to their decisions and not variables to be accounted for in attempts to make better decisions.

As an experiment with the surveyed irrigators, we attempted to fuse some of the data sources they used with evapotranspiration (ET) data, which was new to many of them, to present some of their already used data as well as the new ones, in one place, namely a web page. This attempted to reduce the number of separate data source accesses an irrigator would have to make and therefore reduce their effort. Even for the common and quantified computer data sources this fusion process was fraught with difficulties due to their heterogeneity of form. Allowing choice of individual data source from one data source type, such as ET data, was also complex as different irrigator locations were served by different forms of the same type of source. In our trials, irrigators in two areas, Griffith and Hay, were able to receive ET data generated from an AWS using the Penman-Meyer (Meyer, 1999) modified equation but those from areas outside them, such as Hillston, could only receive ET data generated by an interpolation service (SILO[1]) using the Penman-Montieth equation (Allen, Pereira et al., 1998). This difference of equation can be catered for only with additional DSS design effort. Little progress was made in our attempts to fuse many data sources and no attempt was made to fuse any non-biophysical data sources due to the substantially more complex issues of integration.

This diversity of data sources and data formats that can be used and the requirements for DSS to allow users to choose different forms of the same types of data mean that a DSS that is to fuse or in some way use multiple sources successfully must choose either one of two approaches to deal with their heterogeneity. The first is to generate what, in software terms, is known as a 'wrapper' for each data source – that is a custom written piece of software code that converts the original data into a format that can be readily understood by the DSS. The second is to require that the data the DSS uses conforms to a standard. This first approach was the one taken by the authors and is historically the most common approach. An example of another, perhaps more successful, implementation of it is seen in Research Services New England’s *Probe for Windows*[2] which allows a single computer interface to display SM data from many different manufacturers' probes. Another implementation of the techniques is used for the *WaterSense* (Inman-Bamber, Attard et al., 2005) irrigation DSS which allows users to access on-line rain gauge and Bureau of Meteorology (BoM) AWS data. However, this approach will never be able to grant access to the multitude of data sources potentially for use due to the complexities of creating wrappers for data sources whose outputs are proprietary and thereby opaque to the DSS designers. This approach also requires new work for every DSS designed.

The second approach, that of requiring data sources to conform to a standard, shows promise in the light of recent developments concerning internet data standards, two of which, sensor Web Enablement (SWE) and WaterML, are described in relation to irrigation data use in the next section. Both SWE and WaterML, and indeed a great number of recent data standardisation efforts, rely on a internet data standards, at the heart of which is the eXtensible Mark-up Language[3] (XML). Standards that are XML-based are used in a great range of fields, including natural resource management (see next section), Graphical Information Systems[4] (GIS), web page display[5] and even chemistry[6] and therefore a large body of work is available to be drawn upon in order to determine how irrigation data sources may to conform to an XML-based standard. The diversity of XML-based standards indicate that it may be practical to implement it in yet another context, irrigation, and it’s extensibility suggests that it may also cater for future developments such as new data sources.

EXISTING DATA STANDARDS AND THE IRRIGATION CONTEXT

Several standardisation developments are of particular interest to the irrigation sector. Two of these are the establishment of the Sensor Web Enablement (SWE) Working Group who provide a large range of standards for various sensors to allow the standardisation of all of the processes required from the sensory measurement of data to the publishing of that data on the internet (Open Geospatial Consortium, 2007), and the specification of WaterML, a hydrological observations mark-up[7] language (Consortium of Universities for the Advancement of Hydrologic Science, 2009).

SWE aims to be applicable to all sensors and sensor networks and much of its early work has been influenced by the needs of environmental sensors, such as water gauges and weather stations that are commonly used in irrigation contexts. Work is currently being done in the area of automated, SWE-compliant, sensor-driven irrigation within the Australian context (McCulloch, McCarthy et al., 2008) but this work focuses on sensory

---

[1] SILO is a project that generates gridded daily ET information for all points on the Australian mainland at 5km intervals. See http://www.bom.gov.au/silo for information about the way they generate ET data.

[2] See the RSNE website for description of their “Probe for Windows” at http://www.rsne.com.au/

[3] XML is a general purpose, extensible, mark-up language, specific implementations of which can be made to mark up specific data. See http://www.w3schools.com/XML/xml_whatis.asp for more information.

[4] Google Earth uses the XML-based Keyhole Mark-up Language to display user-generated content. See http://code.google.com/apis/kml/documentation/.

[5] The Hypertext Mark-up Language or HTML is the best known of the XML-derived mark-up languages.

[6] See Chemical Mark-up Language at http://www.ch.ic.ac.uk/rzepa/cml/.

[7] For information to be ‘marked-up’ means it is annotated in order to allow users of it to structure, format and present it. Marked-up in the hydrological contexts means data is presented with metadata describing the units, timestamps and so on used and perhaps the names of their generating methods.

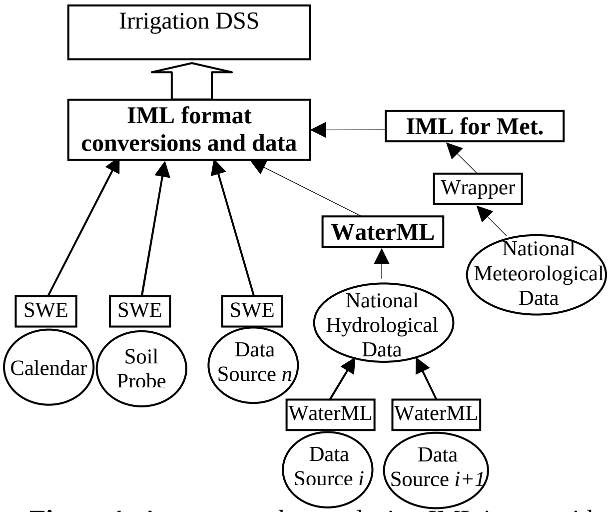

Figure 1: A concept schema placing IML in use with mark-up languages. *IML for Met.* is hypothetical IML output from a meteorological data source wrapper. Square boxes are standards, round ones, data sources.

biophysical data sources only and does not include non-sensor data sources such as personal timetables that our work has shown affects irrigation management decisions (see previous section).

This notwithstanding, future sensors are likely to adhere to the SWE standards given its acceptance by a large number of institutions world wide and the general acceptance that internet-connected sensors will need standardisation as their numbers increase exponentially. Later in this paper we explore how non-sensory non-biophysical data may be adapted for use in a way similar to adapting sensors for SWE-compliant use.

The WaterML project is large, well supported by a range of hydrological researchers group and provides a mark-up language to describe hydrological observations. As with SWE it is internet focused and could easily be used as a starting point for irrigation-specific hydrological processes mark-up. With both SWE and WaterML, there is nothing in their respective specifications that directly prevent their use by new data sources in the irrigation context. Further to this, excessive development would not be required to adapt some existing data sources to their standards so that even though their wide-spread acceptance in the irrigation sector is currently many years away, catering for them is possible and advisable.

An example of WaterML use is the reissuing of data from the United States Geological Surveys' (USGS) National Water Information System (NWIS) website[8] and the US' Environmental Protection Agency's (EPA) Storage and Retrieval System (STORET)[9]. These data sources provide historical and current water observations. The data is not marked-up such that 3rd party data providers, such as *Hydroseek* can collect and reissue it without specific software wrappers (Beran and Piasecki, 2008). Such wrappers have been provided by organisations such as the San Diego Supercomputer Center which scrape the websites' data, save them to databases and then reissue them using WaterML. *Hydroseek* or other 3rd parties may then use that data without reference to the original data sources. *Hydroseek* implements its own map-based interface to the NWIS and STORET data which is unlike the original standard webpage interfaces of the originals. It could implement more interfaces for example for mobile devices all from the same marked-up data.

The above are taken form a recent paper (in print at the time of writing) and represent current work in the international hydrological data standardisation field. Such mark-up retrofitting of hydrological data sources, is also presently taking place in the Australian context with possible implications for irrigation due to projects such as the Water Resources Observation Network (WRON)(CSIRO Land and Water, 2007) which can be expected to present a vast amount of hydrological data in a marked-up format[10] similar to WaterML.

As with the automated irrigation work previously mentioned, both WaterML and the WRON's efforts falls short of what would be required of a mark-up language to allow irrigation DSS data source choice. Irrigation DSS need to do more than simply present data and need to access more than simply sensor-collected biophysical data sources. Any mark-up language to be of use to them then needs to be wider in scope than WaterML/WRON and SWE and needs to have additional information allowing it to determine if it may fuse that data with other data or calculate values from it in order to deliver decision support. Thus an irrigation standard or mark-up/modelling[11] language would act as an additional layer, on top of standardised, marked-up hydrological data, SWE standardised non-hydrological sensor data and, of course, hither to non-marked-up sensory data also used for irrigation decision support. It would not clash with either WaterML or SWE

---

[8] USGS National Water Information System http://waterdata.usgs.gov/.

[9] EPA Storage and Retrieval System http://www.epa.gov/storet/.

[10] See http://www.csiro.au/science/WRONoverview.html for details of how the WRON plans to "enable water information interoperability" through standards development.

[11] A modelling language can be used to model concepts and may include specifications for marking-up data.

mark-up languages and will indeed make use of them. This then is how Irrigation Modelling Language (IML) is differentiated from other standardisation projects related to the irrigation sector. A concept schema showing the relations of SWE, WaterML and IML implementations is given in Figure 1.

IML REQUIREMENTS

The **first** requirement for IML is that it be compatible with SWE, WaterML and other similar standards. Open Geospatial Consortium [12] (OGC) standards are wide ranging, well supported and no reasons exists to prevent their use. Additionally their standards already cover geographic information and this can be used by an IML when geographic details are needed. The OGC aims for full compatibility with the International Organization for Standardization (ISO) therefore this adoption will place an IML in line with much international standards work.

Table 2: Requirements of an IML

	Requirement Description
1	Comply with SWE, WaterML and future WRON mark-up languages
2	Accept wrapper-generated mark-up
3	Provide a conceptual model of irrigation DSS data & process requirements
4	Provide data conversion procedures

As many of the current data sources used in irrigation scheduling are not marked-up in any way, the **second** requirement then is that an IML is able to accept software wrapper-generated marked-up output from non-marked-up data sources commonly used in irrigation. This requirement can be satisfied by simply publishing all IML specifications freely. In this way, should people want to design software wrappers that output IML or internet web services that generate IML, they will be able to follow the standard easily.

As mentioned above, irrigation DSS aim to do more than just display data: they aim to fuse it and calculate values from it. The **third** requirement for an IML is that it specifies a conceptual model of common irrigation decision support data, and processes that defines data requirements for using various data sources in ways such as fusion. It needs to be open-ended allowing new data sources and processes to be added over time.

Satisfying this requirement is a large and ongoing task so that in this short paper we only provide brief examples of a conceptual model used for irrigation scheduling. We also explain how this conceptual model may be extended to other irrigation decision support processes and what conceptual linkages may be defined between industry-used biophysical and non-biophysical data sources in order to prepare for their fusion in future DSS.

Data source units and timesteps need to be standardised for those sources to be used effectively. Rather than forcing adherence to a particular set of units we can define a further, **fourth**, requirement then for an IML is that it specifies procedures to be used convert the units, timesteps and other characteristics of data from similar sources. This paper will demonstrate test implementations of unit conversion in following section.

EXAMPLES OF IMPLEMENTATION

If we use OGC standards with provision for data source wrappers we move to address the third requirement of an IML by making a conceptual model of a part of the irrigation scheduling decision processes and demonstrate some test implementations of conversion to satisfy the fourth requirement.

Irrigation scheduling conceptual model

A layered conceptual model of irrigation scheduling processes, starting with the simplest and becoming more complex, can show how data sources may fit together. A full IML specification would define conceptual models for all of the process that DSS developers use. Figure 2 shows a series of conceptual models for irrigation scheduling, starting with the most basic *water in/water out* model and then moving to ET-based scheduling, SM-based scheduling and finally an ET & SM combined model including a calendar non-biophysical data source that affects the time an irrigator's availability.

Such a conceptual model, if formalised, would specify what data sources could be used together or interchangeably and what sources are mandatory for certain types of decision support. For example, data sources for irrigations, rainfall, ET and drainage would be mandatory for ET-based scheduling but for ET there would be a choice between SILO and an AWS (and potentially other data sources not shown here).

[12] "[The] OGC is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial & location based services." – http://www.opengeospatial.org.

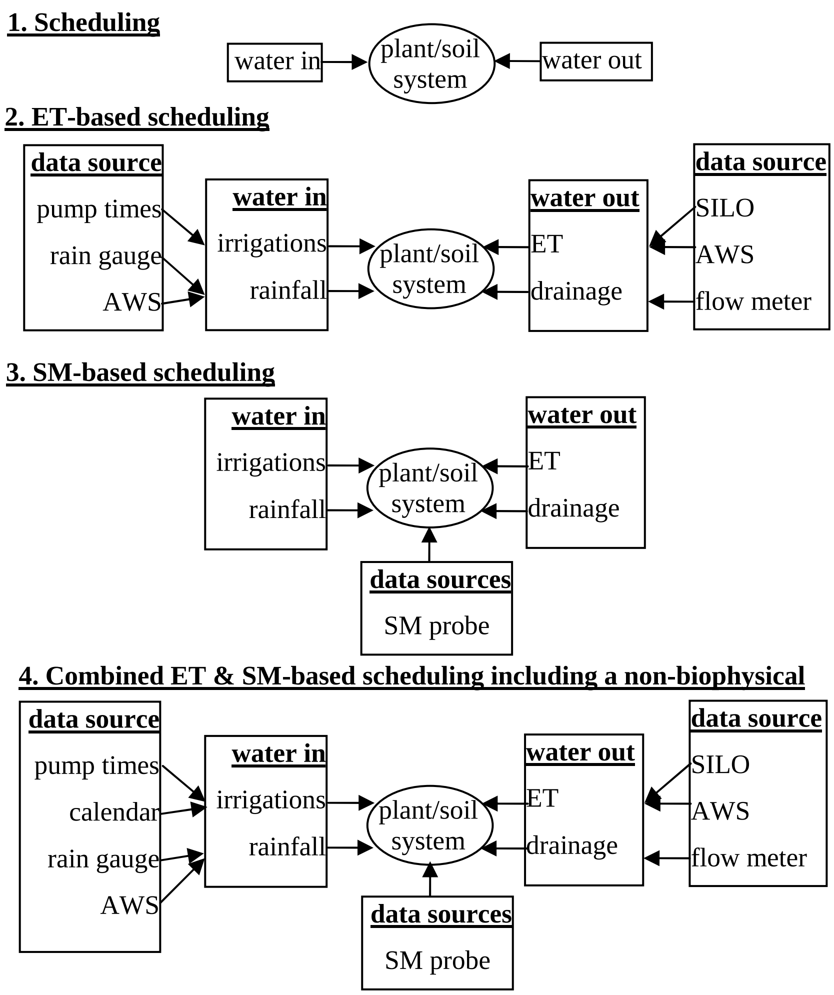

Figure 2: Conceptual model of irrigation scheduling

Calendar information would not be mandatory for ET-based scheduling as the simpler ET-based process does not require it. This is also the case for SM probes for use with ET-based scheduling. This layered approach could be expanded indefinitely to include further process and data sources.

Formal implementations of such a conceptual model could use the Universal Modelling Language (UML) (Fowler, 2004) to formalise compulsory and optional data sources and their relations. UML, like XML, is extendable and used for many purposes and should be suitable for irrigation.

Using SILO or AWS data interchangeably requires knowledge of the choice of equation employed by them. The conceptual model would define XML namespaces[13] that list the various ET equations and their properties. Examples of XML code snippets for marked-up AWS and SILO ET data are show in Listing 1. Working examples of an IML-producing wrapper for live Griffith AWS ET data and a data wrapper for historical SILO data for a location close to the Griffith AWS are supplied at http://irrigateway.net/IML/v0.1/datasources/scheduling/.

Non-biophysical data sources that affect irrigation can be encapsulated and wrapped in IML if they can be expressed in units similar to other data sources.

Listing 1: Marked-up ET data sources

1. AWS using Penman-Meyer equation

```
<datasource type="irrigateway.net/IML/datasources/scheduling/aws">
  <name>Griffith AWS</name>
  <location coordSys="WGS84">146.069321,-34.321571</location>
  <data name="et" timestep="day" units="mm" >
      <datum datetime="2009-02-15 09:10" value="9.13" equation="penman-meyer" />
  </data>
</datasource>
```

2. SILO using Penman-Montieth equation

```
<datasource type="irrigateway.net/IML/datasources/scheduling/SILO">
  <name>SILO</name>
  <location coordSys="WGS84">146.0685,-34.3170</location>
  <data name="et" timestep="day" units="mm" >
    <datum datetime="2009-02-15 09:00" value="7.5" equation="penman-montieth" />
  </data>
</datasource>
```

[13] Namespaces are XML constructs that define canonical names & descriptions for conceptual or real objects.

For the optional calendar data source shown in Figure 2, the common unit linking it to other biophysical data sources is time. This common unit allows calendar-specific data, for example *availability*, to be fused with other data such as pump run times. The web location given above also provides an example calendar data source that outputs data in an IML. The data source is canonically located in the same conceptual region as other biophysical data sources.

Data conversion procedures

A DSS may be able to convert information between datasources if it has the relevant knowledge, such at the timestep, units and equation details. If the 3rd party to the DSS designers or the data source suppliers were able to supply such a converter, it could define a conversion 'data source' and publish that on the internet under the relevant canonical namespace. An example to convert between SILO to AWS ET readings is given at the URL supplied above as is a conversion between average wind speed measured in km/day or m/s. This uses the namespace-defined unit codes to determine the required conversion.

The functions given at the above URL supply their definitions and outputs through the internet data exchange XML-based format known as Web Services allowing any internet-connected computer system to use them. An IML would adopt such a data exchange format and thus allow 3rd party data source developers to use any underlying computer system. Web Services are a fundamental part of SWE.

CONCLUSION

With such a range of data sources and such a range of forms, little progress can be made without a standardisation framework, perhaps in the form of a modelling language. Such a modelling language may be able to supply a data source structure and syntax and, though a series of namespaces, a common definition set that allows data sources to be more easily used, fused and converted. If existing and new data sources supply their data in an IML form, DSS may achieve data source flexibility and thus come closer to bridging the gap between science and practice.

REFERENCES

Allen, R. G., L. S. Pereira, et al. (1998), Crop evapotranspiration - Guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. Rome, FAO.

Beran, B. and M. Piasecki (2008), "Engineering new paths to water data." Computers & Geosciences **In Press, Corrected Proof**.

Consortium of Universities for the Advancement of Hydrologic Science, I. (2009). "WaterML Project Website." Retrieved March 2009, 2009, from http://river.sdsc.edu/wiki/MainPage.ashx.

CSIRO Land and Water. (2007), "WRON - Water Resources Observation Network." http://wron.net.au/ Retrieved March 2009, 2009, from http://wron.net.au/.

Fowler, M. (2004), UML distilled : a brief guide to the standard object modeling language: A Brief Guide to the Standard Object Modeling Language, Addison-Wesley.

Inman-Bamber, N. G., S. A. Attard, et al. (2005), A web-based system for planning use of limited irrigation water in sugarcane. 2005 Conference of the Australian Society of Sugar Cane Technologists. Bundaberg, Queensland, Australia**:** 170-181.

McCown, R. L. (2001), "Learning to bridge the gap between science-based decision support and the practice of farming: Evolution in paradigms of model-based research and intervention from design to dialogue." Australian Journal of Agricultural Research **52**(5): 549-572.

McCulloch, J., P. McCarthy, et al. (2008), Wireless sensor network deployment for water use efficiency in irrigation. Proceedings of the workshop on Real-world wireless sensor networks. Glasgow, Scotland, ACM.

Meyer, W. (1999), Standard reference evaporation calculation for inland, south eastern Australia. CSIRO Land & Water Technical Report. Adelaide, SA, CSIRO Land & Water.

Open Geospatial Consortium. (2007), "Sensor Web Enablement." Retrieved March 2009, 2007, from http://www.opengeospatial.org/projects/groups/sensorweb.