# Bridging the gap between modelling advice and irrigator solutions through empirical reasoning techniques

**N. J. Car** [a] [b], **G. A. Moore** [a]

[a]*Department of Infrastructure Engineering, University of Melbourne, Australia*
[b]*CSIRO Land & Water, Griffith NSW, Australia*
Email: nicholas.car@csiro.au

**Abstract:** Science has led to increased on-farm irrigation productivity in the last decade in Australia through the widespread adoption of better crop varieties, management practices and mechanical tools. However, Decision Support Systems (DSS) for farmers, despite proved productivity benefits Inman-Bamber et al. (2005), have seen poor adoption. Inman-Bamber and Attard (2005) noted the vast majority of purpose-built irrigation DSS saw poor or no commercial adoption after development.

Australian authors, particularly McCown (2001, 2002b,a), have noted a science-industry 'gap' between science researchers and DSS designers and the agriculture industry resulting in systems not entirely fit for practical purpose. Current feeling in agricultural DSS development internationally is summed up by Matthews who states "there is a need to think beyond technocentric solutions" Matthews et al. (2008). They argue that DSS designers must engage with target users through collaborative approaches during development to prevent this 'gap' forming.

The issue central to the gap problem is DSS not catering for the myriad of hard to model, yet influential, factors that affect real-world decisions. This paper details an approach to decision support, implemented using 'technocentric' techniques that may do so. It allows open-ended problem parameterisation and user choice in determining advice relevant to industry that is derived from peers' empirical data.

The approach is implemented using a heavily modified Case-Based Reasoning (CBR) Artificial Intelligence (AI) technique with a novel interface. Traditional CBR systems look at the decision scenario at hand - the current case - and attempt to draw parallels to previous scenarios - stored cases - based on similarity measures. They then presents likely solutions to the current case from relevant stored cases.

Our CBR variant presents an enhanced version of emulated non-DSS farmer decision making. Added to human memory and reasoning is the power of DSS data storage and computer information manipulation. It allows users to select case parameters of interest to them, via an interface, to be used in case comparison in place of the standard, automated and thus predetermined, similarity measures used by most CBR systems. This approach does not restrict decision input parameters or pathways to a solution, unlike current model-based DSS. Advice delivered is still useful to users when the stored cases contain decisions which have been made using knowledge they do not possess.

Analysis of irrigation decision factors is given in Section 1 and the new theoretical approach is given in Section 2. Section 3 details CBR practices that relate to the theory, Section 4 describes a test implementation domain within Australian irrigation, Section 5 details a DSS prototype construction and Section 6 gives initial test results conducted with real irrigation decision data.

The aim of his paper is not to detail a complete, new, DSS using yet another AI technique. It is to highlight that at least one technical approach to DSS design exists which may sidestep current ongoing DSS problems. Ever more sophisticated, purely biophysical, model-based, DSS may never do this but Australian agricultural DSS designers need not give up all technical DSS design. Many more as-yet undiscovered approaches to decision support, including hybrid rule-based and CBR systems that require technical implementation may exist and may prove useful to farmers.

*Keywords:* Decision Support Systems, irrigation, Case-Base Reasoning, Empirical reasoning

## 1 INTRODUCTION

Much social science work has been done, for example Montagu et al. (2006), to understand the factors irrigation farmers do use but no computerised DSS has yet incorporated any of this work in its construction. The closest we have to this in Australia is an undertaking by irrigation DSS designers to use Participatory Action Research (PAR) to inform their design processes Inman-Bamber et al. (2006) and work that looks at social learning as an outcome of DSS use Jakku and Thorburn (2009).

From this and other work we know that farmers make decisions based on biophysical, economic, calendar, business and personal factors. For day-to-day irrigation decisions, they do not just focus on their crop's waterbalance and yet waterbalances, crop parameters and perhaps economic considerations are all that current irrigation DSS can help with. A great step forward in DSS design would therefore be to build a system that allows farmers to incorporate these other factors into the DSS' input parameters.

In interview work previously undertaken by the authors between September 2007 and May 2009, 72 winegrape farmers in Griffith, NSW, were asked what tool or method they used for day-to-day irrigation decisions. Their answers are given in Table 1. Multiple answers were allowed.

**Table 1**. Irrigation scheduling method and tool use by 72 irrigators

| Tool/Method | Irrigators |
|---|---|
| Experience | 66 |
| Moisture probes | 45 |
| Shovel | 28 |
| Visual plant inspection | 26 |
| Fixed schedule | 19 |
| Weather | 18 |
| $ET_0$ | 5 |
| $ET_c$ | 1 |
| Infra-red probe | 1 |

In addition to the tools and methods listed in Table 1, irrigators indicated through extended interviewing that other factors also affected scheduling volumes and timings. These included water availability, expected financial return, frost damage and so on. Some of items in Table 1 are able to be easily quantified, such as *IR Probe*: $-4°C$ *less than ambient* or *Fixed Schedule*: only irrigate on Weekends. Many of the others, such as visual plant inspection have been subjected to attempts at quantification[1] that relate to the FAO56 methodology Allen et al. (1998) but are not universally employed. Some factors, e.g. *experience*, being subjective, multi-factorial and ill-defined cannot be quantified as. There is thus a theoretical and undefinable upper limit to the ability of factor quantification to close the industry/science 'gap'.

*Data Fusion* is the area of work that investigates the combination of differing data source Hall and Llinas (1997) however work in this field is only of limited relevance in this irrigation context due to the extreme heterogeneity of the data sources already used by industry. Even data sources or factors numerically quantifiable may not easily be comparable for example, no mathematical data fusion methods exist to quantitatively compare soil moisture probe readings to millimetres of evapotranspiration. It is certainly not yet possible to compare a *shovel* indication of 'dry', which may mean "irrigate now", to an *IR probe* reading of "$-4°C$ *less than ambient*" which may mean "do not irrigate".

Assuming that irrigators considering multiple heterogeneous, quantifiable and non-quantifiable irrigation factors are not frivolous, some form of data quantification and then fusion is taking place in their head.

## 2 A NEW THEORETICAL DSS APPROACH

We suggest the implementation of a DSS that does not attempt to model a physical process involved with irrigation but one that assists current human decision making. Such a system would adapt to use the different decision inputs that farmers deemed relevant to certain decision scenarios and not be fixed to

---

[1]See Dept. of Primary Industry's webpage Determining Effective Area of Shade in Orchards and Vineyards to Estimate Crop Water Requirement that outlines an area of shade to crop coefficient conversion process.

accepting a limited, pre-defined set of inputs. Provided it could learn from such inputs, it would act as a form of automated PAR with the system recording farmers' input changes over time and changing the advice delivered. A successful combination of non-model-based decision support and pseudo-automated PAR research would supplement existing inter-grower decision support in the form of decision experience and benchmarking by using telecommunication technologies to facilitate communication beyond a farmer's immediate peers.

We should collect the factors affecting irrigation decisions that are relevant to farmers and record the decisions that they made based on those factors. Farmers faced with a decision could consult a reference collection and look for situations similar to their own and view the decisions that others made and possibly utility outcomes of those decisions as well. This would be similar to a benchmarking exercise but could present much more detail to the farmer about irrigation scenarios that others faced. A computerised collection of decisions would also allow a farmer to find situations similar to the one at hand that may not have been apparent in a normal benchmarking study in proximity in location or time. Over time, trends identifying correlations between certain factors and successful decisions would become apparent as would changing factor preference and solution reuse. No direct calculation would be necessary in the DSS, only adequate presentation of situational factors, decisions and decision outcomes.

## 3   CASE-BASED REASONING BACKGROUND

An artificial intelligence (AI) technique that attempts to use non-model-based approaches in generating decision support is Case-Based Reasoning (CBR). This approach looks at the decision scenario at hand - the current case - and attempts to draw parallels to previous scenarios - stored cases - based on similarity or utility measures. The CBR system then adapts successful solutions to previous cases to generate decision support. The basic logic flow of CBR systems was formalised over 15 years ago and involves a decision process lifecycle of *Retrieve*, *Reuse*, *Revise* and *Retain* Aamodt and Plaza (1994). Figure 1 shows the lifecycle with the access of stored cases (usually in a database) indicated.
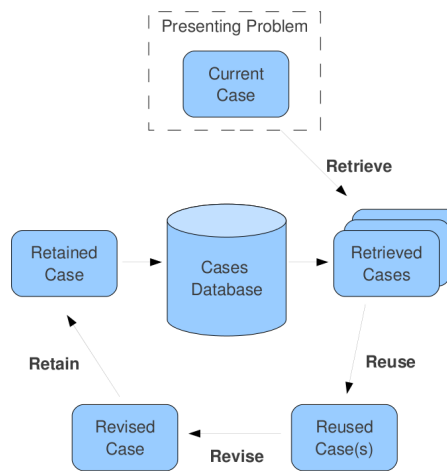


**Figure 1**. The CBR *cycle* after Aamodt and Plaza (1994)

*Similarity* measures are usually determined by a multi-parameter (one for each case *feature*) mathematical analysis of the current case and how it compares to stored cases. The measure may use a series of nearest-neighbour or other equations and requires manual selection and pre-tested algorithms. An example CBR implementation, Ros et al. (2006), computes similarity between the current and stored cases for each case feature using geometric similarity for spatial features, linear similarity for a time-domain feature and a custom 'strategic' function for a situational, time-based feature relevant to their cases. The individual similarity functions are: assumed to be independent, heterogeneous, implemented against a threshold and normalised giving an aggregate similarity of $S = [0, 1]$. The overall sequence of similarity calculation is determined by the authors based on their particular environment needs. Non-numerical case features used in similarity are often ascribed binary values and similarity calculated as a Hamming distance. In almost

all examples of CBR, the set $A$ of features $a$ for a case definition $c$ is fixed although, for a particular current case $c_{cur}$, not all features need be present.

*Utility* is a more recent method used to determine cases relevant to $c_{cur}$ Bergmann et al. (2001). It can use knowledge of previous solution choice, stored as an additional metric in the Case Base (*CB*), to perform sensitivity analysis between input features and particular cases that make good solutions.

## 4  IMPLEMENTATION DOMAIN

An implementation domain for testing - and the ultimate target of the future DSS - is that of daily irrigation scheduling for grape vines, in South Eastern Australia, under drip systems. This is intentionally narrow in order to keep the focus on DSS mechanics.

Much theoretical CBR testing is conducted against a reference Case Base[2] however testing this is not appropriate for this paper due to the fixed nature of case features in it: it is unlike the irrigation decision domain we wish to study.

Current cases, $c_{cur}$, are presented in the form of a problem-solution pair with the problem, $p_{cur}$ being the question "How much should I irrigate today?" to which a solution, $s_{cur}$, consists of a single answer in millimetres, $s \in \mathbb{R}, s \geq 0$, representing an irrigation depth per unit area which may be expressed in drip system run times, $DRT, DRT \in \mathbb{R}, DRT \geq 0$ for ease of use.

Current cases have an open-ended list of possible features $c_{cur} = (a_0, a_1, ...a_n), a \in A, n \in \mathbb{N}, n>0$. Case features, $a$, are key-value pairs consisting of a free text key, $k$ and a numerical or a non-numerical value $v$ thus $a(k, v), k = $ string value, $v \in \mathbb{R}$ OR $v = $ string value. Though not mandated, it is expected that $v$ values will be in the form of fuzzy logic linguistic variables {'few','some','lots'} or {'very dry','dry','damp','very damp','wet'} etc.

There are deliberately no restrictions on the individual items in the $c_{cur}$ feature or the total number of them, other than there being at least one. Not even a broad time domain feature (such as the month in which the decision was made) is mandated. Selecting features from those defined previously for cases in the *CB* is facilitated through the interface, as is the creation of new feature definitions where one does not already exists for a parameter of interest to the user.

The *CB* is populated with cases $c_i$ of problem-solution pairs thus $c_i(p_i, s_i), i \in \mathbb{N}$. Cases are collected from irrigators who have faced and answered the problem of "How much should I irrigate today?" with recorded features and provided a solution.

## 5  PROTOTYPE CONSTRUCTION

### 5.1  Similarity Measure

The similarity measure for this CBR case selection is an interactive process containing the following repeatable steps:

1. *Preference Order*: User specifies (choosing existing/creating new) features up to any number $n$;
2. *Feature Count*: Order cases from the *CB* according to the number of features they have in common with $c_{cur}$. A maximal match would yield a $c_i$ with $n$ features in common with $c_{cur}$;
3. *Weighted Sum*: Order cases based on a weighted sum of the first $n$ features with the sum progressing $\sum_{m=1}^{n}((n+1)-1)a_m$;
4. Allow the user to change either the number and order of features used.

In this prototype, the simplest numerical or proximity matching is used. For numerical features (evapo-transpiration, soil moisture probe readings etc.) similarity between values is determined by the following normalised triangular function:

$$sim(a_{m_{cur}}, a_{m_i})_{\text{numerical}} = \frac{-1}{range(a)} |a_{m_i} - a_{m_{cur}}| + 1$$

---

[2] http://ai-cbr.cs.auckland.ac.nz/cases.html

Where $a_{m_{cur}}$ is the feature $m$ of the current case, $a_{m_i}$ is the feature $m$ of the $i$th case in the Case Base, if it exists, and $range(a)$ is the absolute value of the numerical range of possible values for the feature $m$.

For non-numeric values, the range of the feature is taken to be the range of ordinal states of that value so, for example, a feature of *moisture by shovel* may have the range set of {'very dry','dry','damp','very damp','wet'} giving an ordinal set of {0,1,2,3,4,5} for a total range of 6 thus:

$$sim(a_{m_{cur}}, a_{m_i})_{\text{non numerical}} = \frac{-1}{range(a)}\,|a_{m_i} - a_{m_{cur}}| + 1$$

When the current case feature being assessed does not exist in the target case, the similarity is zero which then gives the full feature similarity equation:

$$sim(a_{m_{cur}}, a_{m_i}) = \begin{cases} sim(a_{m_{cur}}, a_{m_i})_{\text{numerical}} & \text{if } a_m \text{ is numerical} \\ sim(a_{m_{cur}}, a_{m_i})_{\text{non numerical}} & \text{if } a_m \text{ is not numerical} \end{cases} \tag{1}$$

Table 2 gives possible similarity inputs and results for likely features. A typical case in the inital dataset loaded into the case base contained date, crop type, crop canopy size, soil type and waterbalance/$ET_c$ as features, but there are less frequent and more surprising features represented.

**Table 2**. Likely similarity features, inputs and similarity results

| Feature Key $(a_m(k))$ | Feature range set | Feature range value $(a_m(v))$ | Current case feature value $(a_{m_{cur}})$ | $i$th case feature value $(a_{m_i})$ | similarity |
|---|---|---|---|---|---|
| *moisture by shovel* | 'very dry', 'dry', 'damp', 'very damp', 'wet' | 6 | 'very damp' | 'dry' | $\frac{1}{5}$/0.2 |
| *Total $ET_c$* | $ET_c = [0,15]$ | 16 | 3.8 | 5 | 0.925 |
| *Day of Year* | $DoY = [0,365]$ | 366 | 54 Feb 23 | 49 Feb 18 | $\frac{361}{365}$ |
| *water availability* | 'little', 'some', 'plenty', 'unlimited' | 4 | 'some' | 'some' | 1 |
| *water availability* | 'little', 'some', 'plenty', 'unlimited' | 4 | 'some' | 'unlimited' | 0.25 |

## 5.2 Similarity measure refinement

The previous section gives a first attempt similarity measure only. Sensitivity testing, either through direct study or through statistical 'learning' in a growing *CB* will reveal much better individual feature similarity measures over time. The second method is preferred.

## 5.3 Utility measure

Utility metrics being, in this initial system, expressions of the frequency of case selection per feature frequency and features' relative weightings, were added to the database and allowed to form as simulated irrigation decisions were made as detailed in Section 6.

## 5.4 Interface

Critical to the operation of the DSS is interactivity in feature number and preference. Interactivity enables the iterative similarity measure and allows users to head towards potentially useful cases. Through interface visual controls, the user is kept away from mathematical relations entirely.

The test interface for this system is a web page with access to the Case Base and is live, online, at http://111.118.171.88/cbr. A screenshot of it is also given in Figure 2.



**Figure 2**. A screenshot of the test interface. This instance shows a current case with 'month', 'temperature' and 'waterbalance' defined.

The interactive, repeatable steps of the similarity measure are facilitated at different points in the interface. The selection of pre-defined or the creation of new features of $c_{cur}$, up to the open-ended number number $n$ takes place in the left column the the body of the page. Feature weighting is done by the vertically re-ordering them through mouse clicks. Cases are returned from the *CB* using the similarity measure run appear in the centre column with most similar appearing at the top. The utility metric re-orders cases in the center column and displays in a new order in the right column. Main actions, such as starting a new current case, are chosen from links in the left sidebar.

## 5.5 Case Base

The initial *CB* was constructed by the authors from irrigator decisions recorded between 2007 and 2009 in Griffith NSW. Case definitions were constructed artificially, by researchers, not DSS users, but using real decision data. The results of this construction can be seen through the interface as detailed above.

## 6 INITIAL TESTS

Case base functionality, similarity/utility algorithm and interface functionality test only had been performed. First results indicate the 'one-size fits all' similarity function does generate somewhat relevant cases and the interface is capable of listing perhaps more cases than could be achieved by human memory recall alone. A test comparing decision support generated with synthetic case data - case features and results determined by a simple waterbalance model that formed the basis of the IrriSatSms DSS (see Hornbuckle et al. (2009)) - was performed and results form this test will be presented at MODSIM11. This test shows the use of model-based DSS with this CBR system.

## CONCLUSION

CBR is a technical response to the non-technical issue of the science/industry 'gap' irrigation practice. This paper has quantified one aspect of that gap, determined a possible pathway to address it, introduced

the CBR concept to the irrigation research domain, defined a particular decision scenario for study thus linking abstract CBR concepts to real-world situations. It has also provided specific parameter definition, selection and quantification mechanics, demonstrated construction of an irrigation and provided a start for further CB & irrigation theoretical investigation. Importantly, it has addressed the usage of non-biophysical parameters in irrigation decision making through a DSS implementation which, in future DSS iterations, may be successful at delivering decision support on a par with, or above, that of of model-based systems.Construction also yielded unexpected results regarding the use of hybrid model-based/CBR techniques which indicates a possible joint future for them in irrigation decision support.

## REFERENCES

Aamodt, A. and E. Plaza (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AICom - Artificial Intelligence Communications 7*(1), 39–59.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998). Crop evapotranspiration - guidelines for computing crop water requirements. Irrigation and Drainage paper 56, FAO, FAO, Rome.

Bergmann, R., M. M. Richter, S. Schmitt, A. Stahl, and I. Vollrath (2001). Utility-oriented matching: A new research direction for case-based reasoning. In *In Professionelles Wissenmanagement: Erfahrungern und Visionen. Proc. 1st Conf. on Prof. Knowledge Mgt. Shaker*, pp. 264–274.

Hall, D. L. and J. Llinas (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE 85*(1), 6–23. 0018-9219.

Hornbuckle, J. W., N. J. Car, E. W. Christen, T. M. Stein, and B. Williamson (2009, April). Irrisatsms - irrigation water management by satellite and sms - a utilisation framework. Science Report 04/09, CSIRO Land & Water, CSIRO Land and Water, Griffith, NSW, Australia.

Inman-Bamber, G. and S. Attard (2005). Inventory of australian software tools for on farm water management. Technical report, CRC-IF.

Inman-Bamber, N. G., S. A. Attard, C. Baillie, D. Lawson, and L. Simpson (2005, May). A web-based system for planning use of limited irrigation water in sugarcane. In *Proc. 2005 Conf. of the Aust. Soc. of Sugar Cane Technologists*, Bundaberg, Queensland, Australia, pp. 170–181.

Inman-Bamber, N. G., W. A. Webb, and S. A. Verrall (2006). Participatory irrigation research and scheduling in the ord: R and d. *Proc Aust. Soc. Sugar Cane Tech. Year 2006*, 155–163.

Jakku, E. and P. Thorburn (2009, October). A conceptual framework for guiding the participatory developement of agricultural decision support systems. Working Paper 2009-12, CSIRO, CSIRO Sustainable Ecosystems GPO Box 284, Canberra ACT 2601 Australia.

Matthews, K., G. Schwarz, K. Buchan, M. Rivington, and D. Miller (2008). Wither agricultural dss? *Computers and Electronics in Agriculture 61*, 149–159.

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.

McCown, R. L. (2002a). Changing systems for supporting farmers' decisions: problems, paradigms, and prospects. *Agricultural Systems 74*(1), 179–220.

McCown, R. L. (2002b). Locating agricultural decision support systems in the troubled past and socio-technical complexity of 'models for management'. *Agricultural Systems 74*(1), 11–25.

Montagu, K., B. Thomas, G. Thomas, E. Christen, J. Hornbuckle, C. Baillie, C. Linehan, P. Smith, F. Gallo, S. North, W. Meyer, R. Stirzacker, and P. Cornish (2006). *Understanding Irrigation Decisions: From Enterprise Planning to the Paddock*, Volume 6 of *Irrigation Insights*. Land and water Australia.

Ros, R., M. Veloso, R. L. D. Mntaras, C. Sierra, and J. L. Arcos (2006). Retrieving and reusing game plays for robot soccer. In *In Advances in Case-Based Reasoning*, pp. 2006. Springer.