



The i³S Document Series

Using ‘Case based reasoning’ for the solution of water stress problems

AquaStress

Mitigation of Water Stress through new Approaches to Integrating Management, Technical, Economic and Institutional Instruments

Contract FP6 511231

Title	The i ³ S Document Series: Using 'Case based reasoning' for the solution of water stress problems
Editor	Roger Moore and Isabella Tindall
Authors	Neno Kukuric, Frank Robijn, Jasper Griffioen
Current version	V1.0
Date	11/07/2008
Status	Final © Aquastress
Copyright	All methodologies, ideas and proposals in this document are the copyright of the Aquastress project participants. These methodologies, ideas and proposals may not be used to change or improve the specification of any project to which this document relates, to modify an existing project or to initiate a new project, without first obtaining written approval from those of the Aquastress participants who own the particular methodologies, ideas and proposals involved.

Preface

This document describes an application of Case Based Reasoning (CBR) for the solution of water stress problems. The (software) application is developed in the framework of Aquastress project that is focused on developing interdisciplinary methodologies and tools, enabling actors to mitigate water stress problems. In Europe, various regions may encounter similar kinds of water stress due to among others comparable physiographic conditions, land use functions and/or socio-economic conditions. Similar mitigating measures may therefore be used as solution to water stress.

DELTA RES developed a methodology for transferring the knowledge among various regions in Europe. A region related water situation can be considered as a “case”, being always “case-specific”. Resemblances among cases can be used to search for the regions with the most similar water-related characteristics. In this way successful water management measures may be applied elsewhere and failures can be minimized.

The drainage sub-catchment is chosen as a smallest spatial entity level where comparison should still be feasible. If water stress is a problem than mitigation measures could be a solution to the problem. In addition, a water situation is fully described only if natural conditions are included as well. Hence, a water-related case consists of indicators on natural conditions, stress(es) and mitigation measures.

An application is developed to provide comparison among the cases, using the set of indicators mentioned above. The core of the application is so-called Case-Based Reasoning technique that compares and ranks the cases. The Aquastress CBR software application is developed in cooperation with Good Heavens company and it uses a JColibri CBR Framework provided by Group of AI applications, University of Madrid.

The Aquastress CBR is a stand-alone application; nevertheless, it could be coupled with some components of the Aquastress i³S, as demonstrated in this document. Further coupling with the system is determined primarily by data and information that are -or could be- stored within the i³S. The Aquastress CBR is developed in a generic way, being able to handle various types of information and hence, various types of problems.

Further information

Further information on the Aquastress project can be found on the project website: <http://www.aquastress.net>.

Further Information on the Aquastress CBR software application and the access to the application can be found via Aquastress project website.

Contents

1	Introduction.....	7
2	Water stress and knowledge transfer	8
3	CBR software application.....	10
4	The main software features and the user interface	13
5	Concluding remarks.....	15
6	References	16

1 Introduction

This Many regions across Europe are exposed to water stress, triggering implementation of various mitigation measures. A water stress specifics and an environmental setup where it occurs can be comparable among certain regions. It is therefore useful to search for similar 'cases' and to (re)use the experience on stress behaviour and mitigation measures gained elsewhere.

In de framework of Aquastress, a long-term European research project, a knowledge transfer methodology has been developed for a comparison of water-stress related information in Europe. The environmental (or 'natural') conditions and the stress specifics are described by a set of attributes at a sub-catchment level. A significant effort has been made to collect basic information and define the attributes for each sub-catchment in Europe. At the same time, an off-the-shelf Case-Based Reasoning (CBR) tool has been employed as a search mechanism among the cases.

The CBR methodology is a very appropriate for comparison of complex cases where information is incomplete. Regrettably, the chosen CBR tool appeared to be error-prone, with limited input-output options and user interface. Therefore, a new CBR tool has been developed, specifically for the purpose of this project. The reasoner is based on jCOLIBRI (an application that generates CBR code) and consists of several components, namely: a builder, a publisher, a database, services and a web-based user interface. The newly developed software will be integrated in the AQUASTRESS solution support system. Nevertheless, it can be used as a stand-alone application as well.

This document contains a detailed description of the developed software, presented in the context of its application for the water stress related problems. Hence, the following chapter describes a problem area, a knowledge transfer methodology and the role of the CBR application. The third chapter contains detailed description of the developed software and its components. Subsequently, the main software features and the user interface are addressed (Chapter 4). Some final comments on water stress mitigation and the role of CBR are presented in the Chapter 5, followed up by the most relevant references. The detailed description of the software components can be obtained from the authors.

2 Water stress and knowledge transfer

Water management is often confronted with a lack of data. This holds for both surface water and groundwater, and for the quantity and quality aspect. Therefore, a knowledge transfer from the data rich areas to those lacking data, should be maximised. A precondition for the knowledge transfer is a high degree of similarity among the areas. The similarity can be estimated by using a set of indicators that is well defined and consistently implemented throughout the region (in this case: Europe-wide).

A set of indicators required to describe a water-stress case and provide an input for the CBR application, consists of three subsets. The first one is about the natural conditions related to water, the second one specifies the water stress present due to use of water for a specific sector (agricultural, industrial, domestic, tourism, environment) and the last one describes measure(s) used to mitigate the stress.

Natural conditions in the present context refer to conditions that would exist if there were no (or only few) human-caused changes in the water system. The “natural” is a relative concept, since almost all water systems in Europe have been affected by men, centuries long. In general, the natural conditions exclude all fundamental anthropogenic changes to both water quantity (reservoirs, flow diversions, impervious areas) and quality (pollution from human activities). These conditions determine the initial volume and quality of water in any basin. The information on indicators for natural conditions has been collected and derived on the scale of drainage subbasins. The natural-conditions indicators used in this study are listed in the table 1 below.

Table 1. The natural-conditions indicators used in this study

Attribute	Description	Unit	Range
PREYEAR	Precipitation	mm/year	250 - 3500
PRESUM	Precipitation summer hydrological season (apr-sep)	mm/season	50 - 1500
PREWIN	Precipitation winter hydrological season (oct-mar)	mm/season	100 - 2000
PETYEAR	Evapotranspiration	mm/year	1-2500
NETPRE	Net precipitation	mm/year	-725 - 2636
SOILP	Soil permeability (soil infiltration potential)	% area with a certain permeability	high/medium/low
PERM	Hydrogeology permeability (aquifer storage potential)	% of area with a certain permeability	0-100
SLOPE	Mean slope in drainage basin	degree	0 - 25
SALINE	Salinity	% of area	0-100
PROXY	Groundwater recharge potential	symbol	high/medium/low
ARSENIC	Natural background quality (arsenic)	% of area	0-100
DUMP	Presence of extraction sites	% of area	0-100
VOLCAN	Presence of active volcanoes	frequency	0-4
RUNOFF	Runoff	mm/year	1-3500
STRLNG	River density	m/km ²	0 - 6500
WETLAND	Open water bodies	% of area	0-100

By definition, natural conditions alone cannot generate water stress, because in case of a change, the natural environment (i.e. vegetation, geomorphology) will simply move towards a new state. Water stress originates when societal requirements, i.e. the need for clean and sufficient water for consumption, agriculture industry, and safe environment, cannot be fulfilled. According to the AQUASTRESS definition “water stress occurs when the functions of water in the system do not reach the standards (of policies) or perceptions (of the population) on an appropriate quantity or quality, at an appropriate scale and the adaptability of reaching those is not given”. The information on indicators for water stress has been collected and derived on the scale of administrative units and, where possible drainage subbasins. The stress indicators used in this study are listed in the table 2 below.

Table 2. The stress indicators used in this study

Attribute	Description	Unit	Range
ENV-C	Number of dams / length of water course (1/1000 km)	1/1000km	0-1000
DOM-B	Per capita investment in water treatment (EUR)		0-100
DOM-C	Losses in Infrastructure	%	0-100
AGR-A1	Ratio of irrigated to total agricultural area	%	0-100
AGR-B	Percentage of irrigated area equipped with micro/drip-irrigation	%	0-100
AGR-C	Ratio of irrigated to potentially irrigated area		0-100
AGR-D	Agricultural water use per ULA (m3)	m3	0-10000
IND-B	Percentage companies with ISO-14001	%	0-100
IND-D	Added value by industry	EUR per m3 water used	0-10000
TOU-A1	Percent change in population in tourist peak season (August)	%	0-100
ENV-A2	Percentage Volume of Waste Water treated	%	0-100
ENV-A2	Fertilizer (N+P) distributed, kg per ha,		0-10000
ENV-B	Percentage protected area	%	0-100
ENV-D	Percentage of humid areas under high anthropogenic stress	%	0-100
W4	Water use by tourists as fraction of the total	%	0-100
ENV-1A	Number of excess months with river runoff lower than 30Q90		-14,3-144
DOM-A1	Per capita consumption (l/d)	l/d	0-1000
DOM-C1	Percentage families reporting irregularities in water supply (2007)	%	0-100
DOM-D	No. of reported incidents of diarrhoea per 100000 inhabitants (2004)		0-10000
IND-A1	Potential for alternative water use by industry		0-100
TOU-C	capacity of accomodation with 4 or 5 stars as percentage of total	%	0-100
W1	A certain water use as fraction of the total	%	dom/arg/ind/tou/ot

Mitigating measures are needed to combat water stress. The related set of indicators has not been defined yet. At this stage of the project, a substantial general knowledge on mitigation methods has been collected, together with some examples from the project test sites. This knowledge is encapsulated in a knowledge base of AQUASTRESS i3S and should be easily accessible to the user of the CBR application. A possibility to link these applications is discussed further in the paper (in a chapter describing the main software features).

A proper selection of indicators is crucial for the success of the knowledge transfer. Indicators are widely used in practice to report, summarise, simplify and/or clarify state of water resources. Unfortunately, most of these indicators show aggregated values at country level, being of limited use at sub-catchment scale (as chosen for this study). Besides, country level indicators are usually yearly values, whereas the seasonal phenomena such as droughts define substantially water stress situation. Finally, information relevant for the water management is dispersed over various application fields and often not easy to trace.

Having the abovementioned in mind, not only data relevancy, but also their availability was used as a criterion for selection of the indicators. A preference was given to widely available data with some relevancy (proxy data), over highly relevant but sparsely monitored and/or difficult to obtain data. The study showed that a proper combination and processing of the proxy data can provide surprisingly vivid new pieces of information. For more information on AQUASTRESS indicators and their derivation see Vermooten et al (2007) and Griffioen et al (2007). The next chapter describes characteristics of a CBR application that is developed to accommodate selected indicator sets and make the transfer of water-stress related knowledge possible.

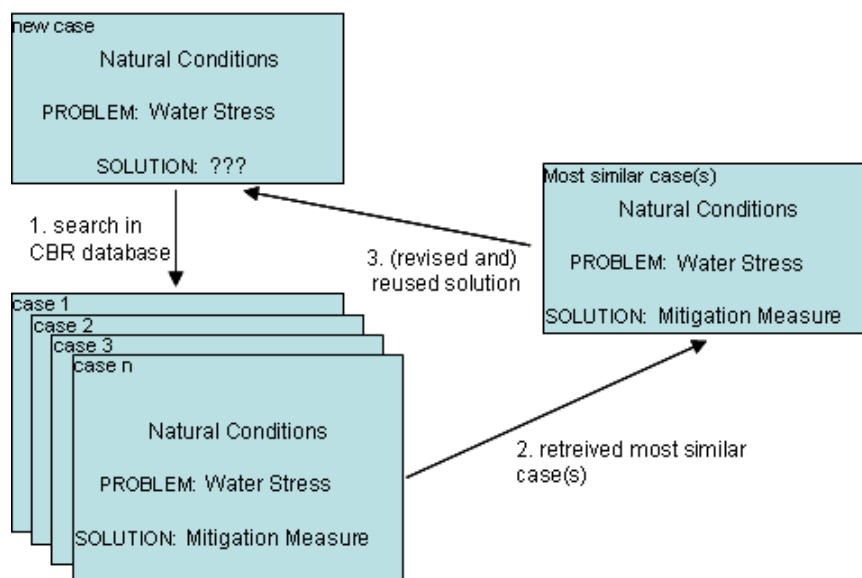
3 CBR software application

The purpose of a Case-Based Reasoner (CBR) application is to retrieve information on existing case(s) that have highest similarity with a new defined case (a new problem), in attempt to (adapt and) reuse already available solution(s) for the new problem (the new case). A water related case consists of indicators on natural conditions, stresses and mitigation measures (Figure 1).

The idea of Case-Based Reasoning emerged in the late seventies as an antithesis of the traditional view of human reasoning. CBR is a direction in Artificial Intelligence and a methodology for the knowledge taxonomy, retrieval and reuse. Since very well described in the literature (e.g. Aarmodt and Plaza, 1994, Kukuric and Vasak, 2000, Kolodner 1993), the postulates of the CBR will not be elaborated here. The attention will rather be paid to a concrete application of the CBR idea.

Originally, DELTARES developed a CBR prototype for the AQUASTRESS project by using commercial software CBR-Works 4.0. This tool provides a shell for development of dedicated CBR applications and it was used by DELTARES in the past to develop research prototypes for lithostratigraphy and in-situ soil remediation. In meantime, many code-errors were discovered and the tool needed urgent updates. Further on, a possible coupling with other AQUASTRESS i3S components became questionable due to quite limited input/output functionality of the tool. Since 2006 CBR-Works is not supported by developers any more and the software code is not available from the software owner. Due to all these reasons, a decision was taken to develop a new, dedicated CBR software for the AQUASTRESS project.

Figure 1. Illustration of the CBR concept



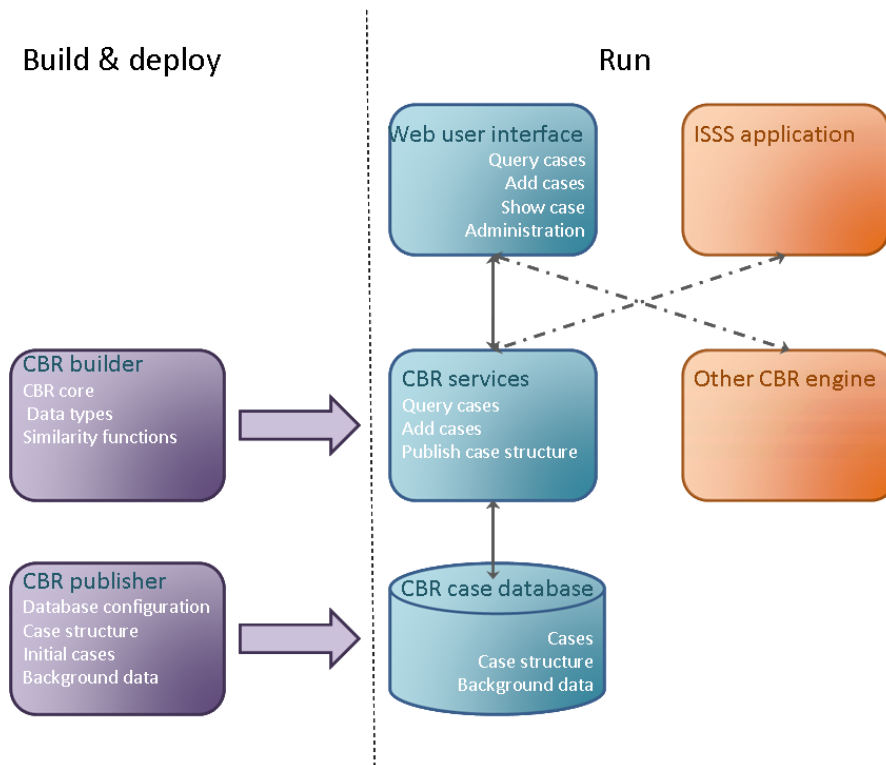
Requirements for the new software were specified according to the objectives of the knowledge transfer activity and the i3S development within the AQUASTRESS project. Some of the major requirements are listed below:

- The software developed in the CBR project should be able to run both as a stand-alone (web) application, and as part of the AQUASTRESS i3S.

- The CBR database should be easily extendible with custom data types and similarity functions. The restructuring of the database should not affect the web user interface, including the graphics.
- The software should produce overall similarity not only for each case, but also for each indicator. The results of query should be presented both numerically and graphically.
- The software should be deployable to a standard hosting service and not rely on the presence of services that are not widely available.

The CBR software is split in three groups of software applications (Figure 2): the CBR engine is implemented as part of a set of web services, collectively known as CBR Services, which also include some basic mapping capability. The web user interface is a generic web application that queries the CBR Services to discover its capabilities and case structure. Finally, (desktop) tools are provided for software developers to build and configure the CBR Services (CBR Builder) and for domain experts to modify the CBR Case database (CBR Publisher).

Figure 2. F The architecture of the CBR Software



Separation of the CBR engine and web user interface via a web service interface makes it possible to reuse either component in distributed i3S applications. The CBR Services can be called by an i3S application that wants to use its query engine or that wants to add new cases. The web user interface can communicate with every engine as long as it implements the same web service interface as the CBR Services.

The CBR Services consists of software that implements the web service interface, a configuration that details the case structure and the similarity rules for the indicators, and a data store. The CBR engine proper is built using jCOLIBRI, an application that generates

code for case based reasoning applications. jCOLIBRI provides out-of-the-box support for various data types and similarity functions, and offers a framework to embed Case Based Reasoning into a web service. jCOLIBRI also has an extension mechanism to add custom components for the requirements that are not fulfilled by the standard components. The CBR Builder tool is a combination of jCOLIBRI and Eclipse, an environment that allows software developers to create custom extensions, generate the CBR engine and use that to implement the web service methods.

The software built in the CBR Builder environment uses configuration files to specify the location of the case data and the case structure (indicators and similarity functions). Within the capabilities provided by the software, the case structure and case data can be modified by experts (i.e. non-developers) in CBR Publisher to set up and fine tune the CBR system. CBR Publisher is a flexible and relatively simple to use spreadsheet based tool; nevertheless, instructions are provided to make the configuration task easy for experts.

Both the web user interface and the CBR Services are developed as Java applications that can be deployed to an Apache/Tomcat application server. At this moment the case data are stored in a flat file but support for a SQL database (e.g. MySql) is partly implemented. All other (configuration and background) data are stored as XML or binary files.

The CBR Services include web services to add, remove and modify cases. These services can be used to import or synchronise cases from external data sources, such as the other AQUASTRESS i3S systems, or any other system for that matter. This would require an i3S application that can obtain data from the other i3S system and call the CBR service with the data.

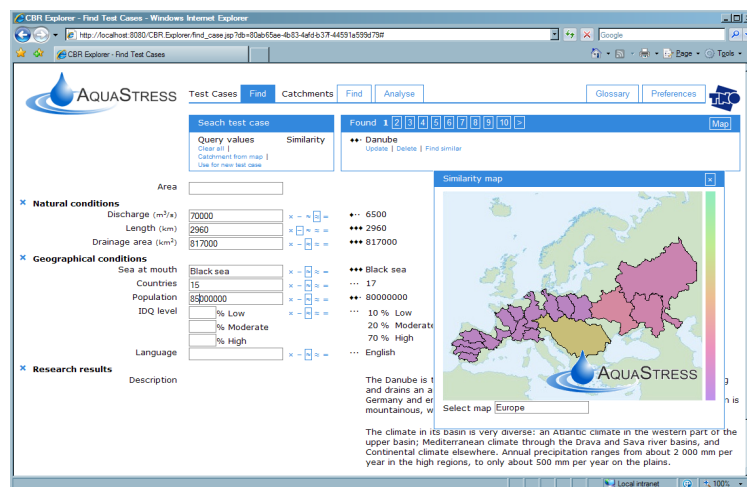
The main features of the CBR software will be outlined in the next chapter, in the way in which the user experiences them: through the CBR web user interface.

4 The main software features and the user interface

The main feature of the software is the CBR-based search of possible solutions (mitigation measures) for (water stress related) problems: the first two steps of Figure 1.. The search page is part of the CBR Explorer website (see Figure 3). The user enters all or some of the characteristics of the water situation. The CBR Explorer immediately returns the best matching cases, and shows a three-star rating of the similarity of the case to the characteristics the user searched for. The same rating is displayed for the individual characteristics. The user can refine the search by changing the importance of a characteristic and hence its weight in the overall similarity.

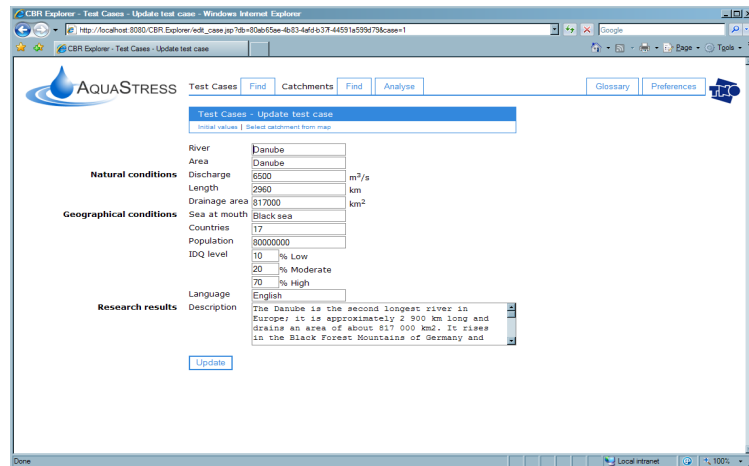
The CBR-based search is text-based, but the CBR Explorer also offers map-based search support. If the area the user is investigating (or a similar area) is part of the CBR Services database, the user can select the area from a map and include the natural conditions of that area in the search. The CBR Explorer can also display a map of the similarity of the best matching case for each area, and the user can click on an area to display the full case data.

Figure 3: The main search page of the CBR Explorer where the user enters the relevant characteristics of the region, or selects an area with similar problems. The CBR Explorer immediately returns the best matching cases. A map of the similarity of the best matching cases is also available.



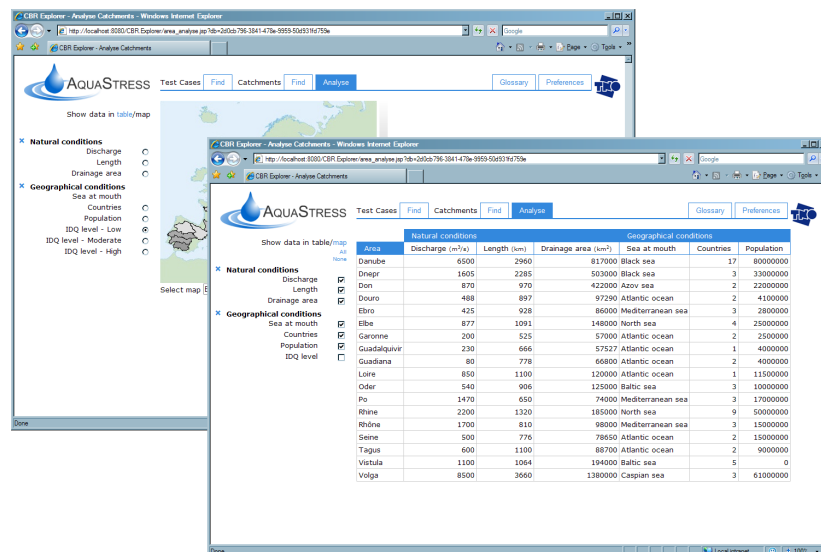
The third step in the CBR cycle (see Figure 1) occurs if a user has updated information on an existing case, or has learned of a new solution. The CBR Explorer provides links on the search page to update an or delete existing case, or to add a new case based on the characteristics the user searched for (see Figure 4). The user has to be authorised for case managements tasks.

Figure 4: The CBR Explorer supports case management tasks like adding and updating cases.



The database with the CBR cases can also be used to study the natural conditions in a wide area. The CBR Explorer allows users to search (using the CBR technique) for areas that best match the specified natural conditions, and to explore the data on the areas, both via a map and tabular data (see Figure 5).

Figure 5: The case database holds a wealth of information on natural conditions; the CBR Explorer allows the user to browse through and visualise the data.



Finally, the CBR Explorer includes a glossary that elucidates the characteristics of the cases (natural conditions, water stress and mitigation measures) and key notions of the CBR methodology.

Another program, the CBR Publisher, offers administrators a tool to define the CBR case database structure: what characteristics are present, how similarities are defined, what maps are available, etc. The CBR Publisher is an Excel front-end to the CBR Services (the actual CBR engine). The administrator modifies the CBR case structure and load an initial set of cases in Excel. The user then uses the CBR Publisher to check the data for consistency, and to upload the structure and data to the CBR Services.

5 Concluding remarks

Developed Case-based Reasoner is a flexible and robust tool for comparison of water-stress related information in Europe. In general, the benefit of CBR software depends primarily on selection of indicators and their mutual relationships (defined through similarity functions and weights). Therefore Case-Based Reasoning needs to be understood and accepted as a methodology for knowledge structuring and encapsulation.

Selection of AQUASTRESS indicators has been a long and difficult process, that extends far beyond development of this software tool. The indicator list is certainly neither definitive nor exclusive. A lack of indicators for mitigation measures prevents the CBR tool be employed fully by providing solutions of resembling cases (at this moment only comparison of the cases is possible). Nevertheless, a methodology for knowledge transfer and a related CBR tool are developed within this project. Since a process of data collection and processing is still ongoing, CBR software could soon be employed by project partners in daily practice throughout Europe.

Finally, developed tool is not restricted only for the water stress problems; it can be used in any problem area showing case-specific but still comparable behaviour.

6 References

- [1] Aamodt, A. and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches", AICOM, Vol. 7, No. 1, (1994) pp 39-59.
- [2] Griffioen J., S. Vermooten, N. Kukuric and L. Vasak, "A European case-based reasoning tool to intercompare hydrological information on water stress among drainage" in *Climate Variability and Change: Hydrological Impacts*, Siegfried Demuth and Arthur Askew Eds, IAHS Press, (2006).
- [3] Kukuric, N. and L. Vasak, "Knowledge Encapsulation: a Case from Groundwater Remediation Practice" Proc. Int. Conf. Hydroinformatics, Iowa, USA, (2000).
- [4] Kolodner, J.L. "Case-Based Reasoning" Morgan Kaufmann , San Francisco, California USA, (1993).
- [5] Vermooten, S., J. Buma, J. Griffioen, N. Kukuric, J. Reckman and L. Vasak, "A method to intercompare water stress-related information among drainage subbasins", AQUASTRESS report, Deliverable 3.8-3, TNO, Utrecht, The Netherlands, (2007).