# Calculation of the water balance and analysis of agriculture drought data using a Business Intelligence (BI) system

Sašo Celarc, Mojca Gros

Bron d.o.o.
Družba za informacijsko tehnologijo
Litostrojska 44c
1000 Ljubljana, Slovenia
saso.celarc@bron.si
mojca.gros@bron.si

**Abstract:** One of the goals of the ongoing project BOBER (acronym for Better Observations for Better Environmental Solutions), founded by SEA (Slovenian Environment Agency) and co-funded by European Union through Cohesion Fund is the consolidation of data, knowledge, experience and tools for monitoring and analysis of data of agriculture drought. One of the challenges is computing of soil-crop water balance and irrigation needs based on meteorological, soil and crop data for different crops and locations. The process of water balance calculation is part of the ETL mechanism (E-Extract, T-Transformation, L-Load), results are stored in the data warehouse and can be visualised and further analysed on the dashboards in the form of numbers, graphs and geographical maps. From a technical point of view this BI solution is based on an Oracle set of tools. This set of tools is specific to a certain degree, because they have an Oracle Spatial database extension used for storing geolocated data in the database and MapViewer for visualising data on the geographical maps.

## 1 Introduction

Effective drought monitoring includes data analysis of measurements and soil/crop water modelling. Due to increased variability of water supply in agriculture in the last decade, it is necessary to analyze both extremes: water deficit and water surplus. Agricultural drought is mainly connected to soil water deficit. When soil water content approaches values when roots are not able to take water from soil (usually referred as wilting point), crops become drought-stressed. It is therefore important to monitor soil water values. In Slovene agrometeorological IT system we have access to soil water content in depths 10, 20 and 30 cm in six locations. Data is available on daily, hourly or 10-minutes basis. In addition to basic variables, evaluated variables like reference evapotranspiration and water balance are used. Water balance is defined as difference between measured precipitation amount and calculated reference evapotranspiration. It is correlated to soil water content and is therefore suitable for agricultural drought monitoring. The evapotranspiration rate is normally expressed in millimetres (mm) per unit time. The rate expresses the amount of water lost from a cropped surface in units of water depth. The time unit can be an hour, day, decade, month, or even an entire growing period or year.

In the past, the Department of Agricultural Meteorology had several IT solutions developed to cover the work process of monitoring weather conditions (drought) which stemmed from their current requirements. The scope of software solutions do not allow for the desired development within the capabilities of the available human resources and know-how. The need for a single platform, new content and transparency of the process generated the demand for a IS which will unify and extend the functionality of the existing system, with emphasis on the integration of new data (especially new internal data sources and extracting data from external partners databases), calculating individual quantities (load, interpolation) and a newly designed part with a GIS application. The new version of the system is expected to offer compatibility with existing processes and systems in ARSO and the possibility of its development in the future, especially in terms of adaptation and modularity of individual contents which will allow for future implementations of not yet determined needs (eg, communication with external data sources) by internal resources.

## 2 Application for water balance calculations

The basic concept of the project is agrometeorological irrigation scheduling model IRRFIB[SU05]. Model is used for the crops-soil water balance calculations. The model was developed by Slovene meteorological service in the frame of SEA. It is a model that simulates water use of plants during the growing season, taking into account the water content in the soil, plant phenological stages, rooting depth, and weather conditions. Input data are divided on three groups: meteorological data, soil data and crop data. The model also allows the calculation of the amount of irrigation for the normal development of plants. The application was developed as part of a system which makes it possible to carry out large-scale calculations of the water balance for a combination of different plants, different soil types, and different locations (monitoring stations) simultaneously. The results of the water balance based on these calculations are available for review and analysis in the application itself and they are transferred to the Business Intelligence system for analysis purposes. A subset of results are also published in different publications. The application enables the calculation of the water balance based on actual data gathered in the field. It also allows simulation of various scenarios of the water balance for selected crops, soil types and location. It offers possibilty to simulate water balance for hypothetical scenarios for planning agricultural prioduction. The user can in one step calculate all the combinations for a number of plants*location*soil*scenario. The water balance data is calculated for every day for each combination, which is visualized in graphs that show the soil moisture (%) and moisture deficit (mm) and the number of days the plant is under stress.

## 3 Dynamic data reviews

For the purpose of the application it is necessary to collect data from various sources and properly prepare them for the water balance. Since the data comes from sources that vary, both in structure and content, it first needs to be merged to a common denominator. The data can be located at different time intervals - eg. air temperature located on a daily level, data on the quantity of irrigation by different types of water (surface, underground) is sourced on an annual level. In addition, the data may contain inaccuracies or we need to adjust and refine the data for comparisons. The data transfer process from various sources offers regular daily calculations, such as the calculation of the number of days with precipitation in a given interval, the calculation of the spring and autumn temperature thresholds and other similar calculations. The data is collected in the so-called stage area, where it is processed and prepared for further processing. Thus, data produced from the stage area are transferred to the application for the water balance and the data warehouse. A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data and the support of management's decision making process [WI92]. Data warehouse and Business Intelligence tools are traditionally used for analysis and decision support on basic business data. On the other hand these tools are so general and rich in functionality that they can be used for the analysis of other data types-including spatial and environmental. In our case, the data warehouse is a uniform place where the data is collected and compiled in such a way that users can use Business Intelligence tools perform various analyzes and reviews. The calculated water balances are transferred back into the data warehouse at the end of the process.

The application for the water balance contains some preset data reviews and analysis. For users who want more detailed analysis and access to the data the analysis tool Oracle Business Intelligence can be used. As part of the project many dashboards have been prepared with predefined reports, which users can transform as they wish. In a very simple way, users can sort any data on the reports - eg. sorting by monitoring stations in relation to rainfall then in the next instance change the view and sort by air temperature. The user can practically with a single click completely change the content of the report. E.g. from a daily review we can get a cumulative overview of data by vegetation seasons, agro decades, months or years. In some reports, we want the system to alert us of specific information, so it is labeled differently than other information. E.g. a red dot indicates the day, when the plant is under stress. The user is thus at first glance aware of the deviation, which can then be further analyzed by examination of additional values.

In the pre-prepared reports various forms of digging through data are used. We have prepared a list of water balance calculations with basic data (plants, measuring station, period, number of stress intervals ...). Users and then click on the calculation that they are interested in which opens a new report with more detailed information. There are many levels of drilling. You can, for example, in the second step see data by plant stages and in the third step by day. The advantage of the system is the ability to build reports tailored to users and to their needs and content. Each user can also save a report in a folder where they can then transform it to their specifications (view changes, add new attributes, etc.). The tool is not limited only to tabular presentations of data, but the data can be displayed in various graphical presentations (eg temperature movements, precipitation within a certain period of time). Moreover, relatively simply, the data can appear on various maps (eg, the sum of precipitation in different geographical units).

The data warehouse stores data for a longer period of time (currently around 50 years). Users have the option to easily review and analyze their own data from a longer period of time and compare them with current data. For example, the long–term average air temperature can be compared to the data on the selected day in the same report. If the input parameters are changed the report can show the data for other locations or periods. As the data warehouse draws data from various sources, the reports can containing a variety of information - eg, air temperature, soil temperature and rainfall can be displayed together with the data from the water balance (eg. evapostranspiration). The data from the water balance simulations is also recorded in the data warehouse, so that can users analyze and examine different scenarios. Using these tools the time spent preparing reports and various analyzes is drastically reduced, in addition, they allow the user to view information from a different angle at any time. It is relatively simple to produce reports such as displaying the number of days in a given period for selected locations, where the temperature was greater than or equal to the selected. All reports created by users can also be publicized. For example, the content and format of reports can be exported to other formats (eg Excel, PDF, CSV, ...). It is also possible to prepare various publications containing any text (explanations of the content of the reports), facilities (eg logitipe, images, etc.) and prepare them for printing.

## 4 System development

The information system for agricultural drought monitoring presented in this article is currently in the final phase of project. The project is complex from both a technological aspect as a link between different technologies; data formats and displays, as well as in content (a broad range of data). That is why a clearly defined project objective is important; we followed the objective in all stages and good project management, which reminds us of deviations on the way to your objective. It is also very important, when constructing such a system, to always involve the people who will use the system. At the beginning of the construction of such a system it is also important to clearly define for whom and why the system is being designed. If the answers to these questions are not known, we should not expect that the system will itself give users answers. Although we used some of the tools that are at first glance more intended for business information, it became evident, that they are entirely equivalent and effective when using other data, despite having different characteristics.

## References

[AS05] Sušnik, A. 2005 IRRFIB model and its practical usage for drought estimation in Slovenia. In: Maracchi, G., L. Kajfez-Bogataj, S. Orlandini, F. Rossi & M. Barazutti (eds.): Irrigation and pest and disease models: Evaluation in different environments and web-based applications. European Commission, Brussels, 36-42.

[WI92] William H. Inmon: Building the Data Warehouse, 1992.