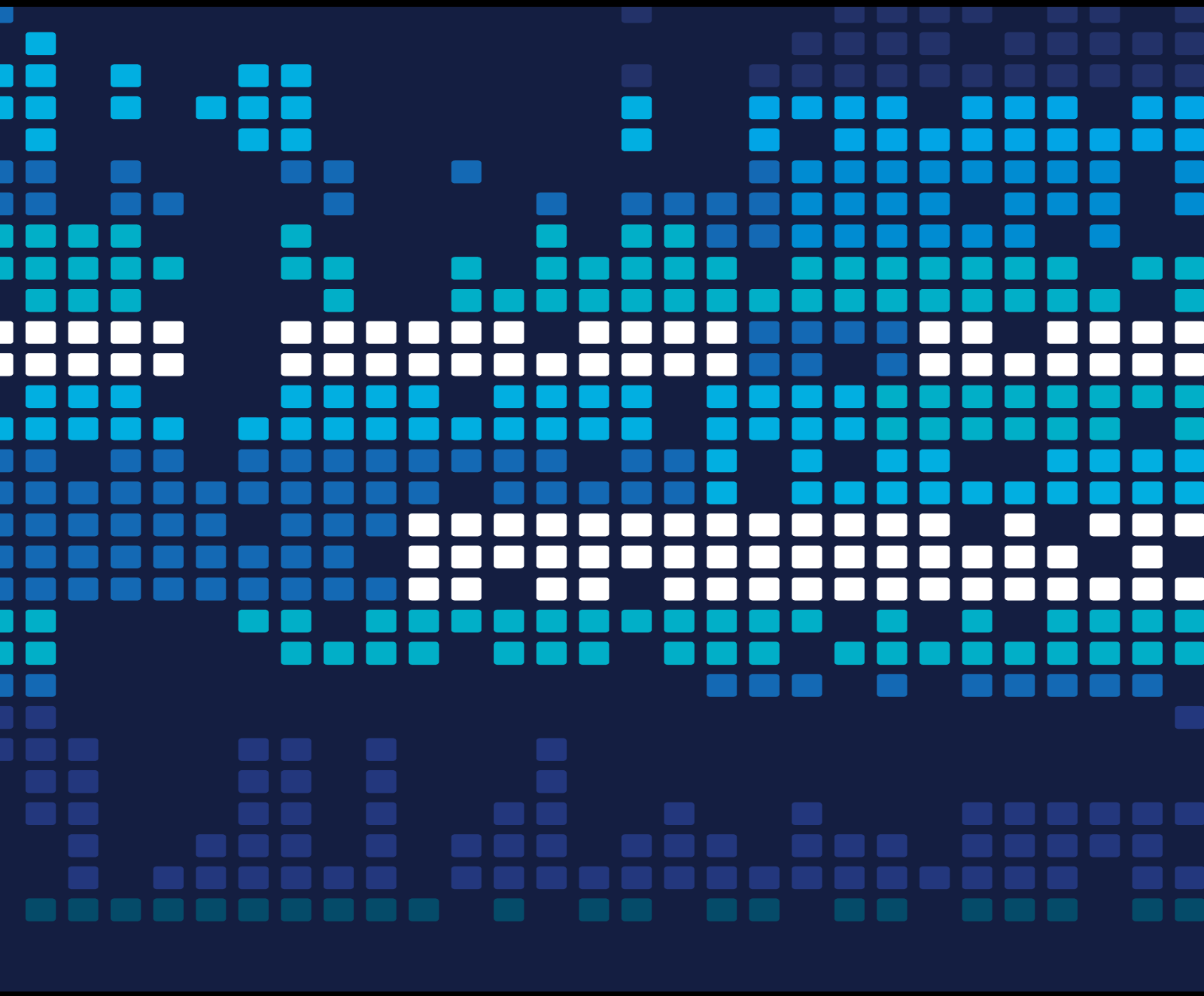
Scientific Programming

Scientific Programming Tools for Water Management

Lead Guest Editor: Francisco J. Alcalá
Guest Editors: Jaime Martínez-Valderrama, Francisco Gomariz, Carlos G. Hernandez, and José M. Cecilia



Scientific Programming Tools for Water Management

Scientific Programming

Scientific Programming Tools for Water Management

Lead Guest Editor: Francisco J. Alcalá
Guest Editors: Jaime Martínez-Valderrama, Francisco Gomariz, Carlos G. Hernandez, and José M. Cecilia

Contents

Hindawi
Scientific Programming
Volume 2021, Article ID 9828596, 3 pages
https://doi.org/10.1155/2021/9828596

Editorial

Scientific Programming Tools for Water Management

Francisco J. Alcalá,[1] **Jaime Martínez-Valderrama**,[2] **Francisco Gomáriz-Castillo**,[3,4] **Carlos G. Hernández**,[5] **and José M. Cecilia**[6]

[1]*Instituto Geológico y Minero de España (IGME), Madrid 28003, Spain*
[2]*Multidisciplinary Institute for Environment Studies 'Ramón Margalef', University of Alicante, San Vicente del Raspeig, Alicante 03690, Spain*
[3]*Department of Geography, University of Murcia, Murcia 30100, Spain*
[4]*Euro-Mediterranean Water Institute (IEA), Murcia 30100, Spain*
[5]*Departamento de Producción Agraria, Universidad Politécnica de Madrid, Madrid 28040, Spain*
[6]*Department of Computer Engineering (DISCA), Universitat Politècnica de València, Valencia 46022, Spain*

Correspondence should be addressed to Francisco J. Alcalá; fj.alcala@igme.es

Received 18 May 2021; Accepted 18 May 2021; Published 23 July 2021

This special issue delivers a platform in which researchers expose intersections between algorithm design, software platforms, and hardware architectures to deal with emerging challenges in the scientific field of management of water and water-dependent resources. Since the call for papers was announced in June 2019, this special issue has received 10 manuscripts. After a rigorous review process, 6 papers have been finally accepted for publication. Published papers deal with groundwater quality monitoring, coastal groundwater-dependent irrigation agriculture, desertification risk, water recovery from tailings, future scenarios of water resources, and vulnerability of coastal aquifers.

1. Introduction

One of the main humanity challenges is management of water and water-dependent resources. Global climatic driving forces determine the water availability and renewability rates of water-dependent resources such as soil, rangelands, crops, livestock, and similar others. The management of water and water-dependent resources is mandatory to cope with ecosystem preservation and human development. However, many of the terrestrial socio-ecological systems are typically sparse-data areas where the capability to adequately characterize the nonlinear processes that govern the dynamics of water and water-dependent resources is limited. Climate may determine the design of a computation platform. Many of the computational tools designed for humid and temperate regions often fail when trying to reproduce some complex human and environmental interactions that occur in drylands. In general, a lack of computational tools for modelling of water and water-dependent resources in the literature is observed.

The integration of the latest breakthroughs in geology, climatology, hydrology, agronomy, mining, and biotechnology from one side and high-performance computing, artificial intelligence, and computational modelling to the other has enabled remarkable advances in the field of water management. By merging these developments, scientists have started to create new strategies to cope with the consequences of global climatic forces and human action, such as water scarcity, ecosystem degradation, and decreasing renewability rates of water-dependent resources.

These scientific applications typically require big datasets and complex simulations to adequately characterize the nonlinear processes that govern the dynamics of water. Highly intensive computational strategies could greatly benefit from increased scientific computational resources to reproduce the complex environmental and human interactions that occur in water bodies and their associated ecosystems and dependent resources. All these scientific and programming advances are combined to provide innovative solutions to deep-seated problems for

society, especially with respect to forecasting and uncertainty analysis.

This special issue aims to deliver a platform. On the platform, researchers expose how the intersections between algorithm design, software platforms, and hardware architectures can be used to deal with emerging challenges in the scientific field of management of water and water-dependent resources. At the same time, this special issue showcases the main trends in scientific parallel processing, algorithm definition, and problem-domain requirements to anticipate future solutions which could be translated into real societal benefits. This special issue includes original research articles that describe specific computational tools related to natural and nonconventional water source management, optimization of water in agriculture, industrial and mining, and the dynamics of water-dependent resources.

Once the call for papers was announced in June 2019, this special issue has attracted attention. In total, 10 manuscripts were received. After a rigorous review process, 6 papers have been finally accepted for publication. To gain a better insight into the essence of the special issue, we offer brief highlights of the published papers in Section 2.

2. Published Papers

The paper "A New Clustering Algorithm and Its Application in Assessing the Quality of Underground Water" [1] introduces a new concept of "epsilon radius neighbours" in cluster analysis. Based on "epsilon radius neighbours," a new clustering algorithm in which the epsilon radius value is adapted to the characteristics of each cluster in the current partition is proposed for groundwater quality monitoring. Through this application, the authors found that the new algorithm might provide valuable reference information for groundwater quality monitoring management.

The paper "AQUACOAST: A Simulation Tool to Explore Coastal Groundwater and Irrigation Farming Interactions" [2] introduces a tool to analyse the impact of apparently unrelated factors such as the prices of crops, subsidies, or exploitation costs on the advance of saltwater intrusion, a great threat to coastal groundwater-dependent irrigation agriculture systems. The tool implements a system dynamics model in Vensim software that considers the integration of hydrological, agronomic, and economic drivers, and a Venapp push-button interface that allow users access to a Vensim model without going through the Vensim modelling environment.

The paper "Flow Enhancement of Mineral Pastes to Increase Water Recovery in Tailings: A Matlab-Based Imaging Processing Tool" [3] determines how water content affects the flow behaviour and depositional geometry of tailings and silica flour pastes in the mining copper industry. The modelling results indicate that the new technique can be incorporated to determine the proper solid content and modifiers for a given fluidity requirement. In general, the techniques are aimed at reducing current high water losses by evaporation and percolation during mineral recovering practices.

The paper "A Statistical Tool to Generate Potential Future Climate Scenarios for Hydrology Applications" [4] introduces GROUNDS to generate potential future climate scenarios in water resources systems of different typology and size (river basins, aquifers, mountain ranges, and countries) from historical and regional climate model information. The tool generates future series of precipitation, temperature, and potential evapotranspiration and uses different approaches and statistical correction techniques to generate individual local projections and ensembles of them. The tool is also useful for quantifying the uncertainties of future scenarios by combining them with stochastic weather generators.

The paper "SAT: A Software for Assessing the Risk of Desertification in Spain" [5] presents SAT, a decision support system for assessing the risk of desertification in Spain. SAT relies on three versions of a generic desertification model that integrates economics and ecology under the predator-prey paradigm. The models have been programmed using Vensim, a software to build and simulate system dynamics models. Through Visual Basic programming, these models are operated from the Excel environment. Other specially designed tools have been coupled to assess the risk of desertification and determine the ranking of the most influential factors.

The paper "GIS-SWIAS: Tool to Summarize Seawater Intrusion Status and Vulnerability at Aquifer Scale" [6] introduces GIS-SWIAS, an ArcGIS ArcToolbox to analyse seawater intrusion (SWI) status and vulnerability at aquifer scale. The tool uses georeferenced information and provides results in terms of reports and images of the SWI evolution over time. Depending on the data coverage, the tool can be applied to provide an initial picture of historical SWI or a rational quantification for detailed SWI management. The tool enables researchers and technicians to assess SWI dynamics and aquifer resilience under different management scenarios including adaptation strategies to potential future scenarios.

3. Conclusions

The guest editors envision that the published papers in this special issue would be of interest to researchers and practitioners and help identify further research routes. We also hope that the readers can find the material of this special issue both interesting and inspiring when exploring the field of management of water and water-dependent resources.

Conflicts of Interest

The editors declare that they have no conflicts of interest.

Acknowledgments

The editors also express gratitude to all authors for their submissions to this special issue. The editors would also wish to thank the editorial board for the opportunity to edit this special issue and for providing valuable suggestions to improve the selection of research manuscripts.

Francisco J. Alcalá
Jaime Martínez-Valderrama
Francisco Gomáriz-Castillo
Carlos G. Hernández
José M. Cecilia

References

[1] T. Vo-Van, A. Nguyen-Hai, M. V. Tat-Hong, and T. Nguyen-Trang, "A new clustering algorithm and its application in assessing the quality of Underground water," *Scientific Programming*, vol. 2020, Article ID 6458576, 12 pages, 2020.

[2] J. Martínez-Valderrama, J. Ibáñez, and F. J. Alcalá, "AQUA-COAST: a simulation tool to Explore coastal groundwater and irrigation farming interactions," *Scientific Programming*, vol. 2020, Article ID 9092829, 20 pages, 2020.

[3] S. L. Mondaca, C. A. Leiva, C. A. Acuña, and E. A. Serey, "Flow enhancement of mineral pastes to Increase water recovery in tailings: a matlab-based imaging processing tool," *Scientific Programming*, vol. 2020, Article ID 5607242, 9 pages, 2020.

[4] A. J. Collados-Lara, D. Pulido-Velazquez, and E. Pardo-Igúzquiza, "A statistical tool to generate potential future climate scenarios for hydrology applications," *Scientific Programming*, vol. 2020, Article ID 8847571, 11 pages, 2020.

[5] J. Martínez-Valderrama, J. Ibáñez, F. J. Alcalá, and S. Martínez, "SAT: a software for assessing the risk of desertification in Spain," *Scientific Programming*, vol. 2020, Article ID 7563928, 12 pages, 2020.

[6] L. Baena-Ruiz and D. Pulido-Velazquez, "GIS-SWIAS: tool to summarize seawater intrusion status and vulnerability at aquifer Scale," *Scientific Programming*, vol. 2021, Article ID 8818634, 15 pages, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 8818634, 15 pages
https://doi.org/10.1155/2021/8818634

Research Article

GIS-SWIAS: Tool to Summarize Seawater Intrusion Status and Vulnerability at Aquifer Scale

Leticia Baena-Ruiz and David Pulido-Velazquez

Department of Research on Geological Resources, Geological Survey of Spain (IGME), Urb. Alcázar del Genil, 4-Edif. Bajo, Granada 18006, Spain

Correspondence should be addressed to Leticia Baena-Ruiz; l.baena@igme.es

Received 13 March 2020; Accepted 26 April 2021; Published 12 May 2021

Academic Editor: Emiliano Tramontana

In this paper, we introduce GIS-SWIAS, a novel generalized ArcGIS ArcToolbox that helps to analyze seawater intrusion (SWI) status and vulnerability at aquifer scale (SWIAS). It is a user-friendly tool that can be applied to any aquifer and is fully integrated in the ArcGIS environment, which is a widely available software tool. It is the first ArcGIS tool with these characteristics focusing on SWI analyses that we can find in the literature. GIS-SWIAS is able to deal with georeferenced information; it is easy to introduce the required data (inputs) and to efficiently perform the demanding computational operations required. Its outputs are in the form of shapes, reports, and images (maps, conceptual cross sections, and time series of lumped indices) to summarize the magnitude, intensity, and temporal evolution of SWI within an aquifer for specific dates or by showing statistics for a chosen time period. It can be applied to assess historical SWI dynamic in cases where there is no groundwater flow model. In those cases, the spatial distribution is assessed by applying simple interpolation techniques. Nevertheless, if we want a rational quantitative analysis of the sustainability of alternative management scenarios to the SWI problem, the GIS-SWIAS tool requires that information on hydraulic head and chloride concentration distribution is generated from simulations of their impacts by a calibrated density-dependent flow model. In such cases, adaptation strategies to potential future scenarios—whose distributed impacts have to be propagated within the previously calibrated models—could usefully be analyzed and compared using this tool. Given all these ways that the GIS-SWIAS tool can be applied, it provides a valuable tool for both the researcher and technician to assess SWI dynamics and aquifer resilience under different scenarios. It can support the decision-making process by helping to make a rational selection of sustainable management strategies. Its performance for the analyses of historical and potential future scenarios has been tested and confirmed in two case studies described in previous research works.

1. Introduction

Seawater intrusion (SWI) in coastal aquifers is a worldwide problem affecting groundwater-dependent ecosystems and human health. In recent decades, society's awareness of this issue has grown and this has been reflected in the legal framework of many countries. For example, the European Union's Water Framework Directive (WFD) requires that river basin plans address the achievement of a good qualitative and quantitative status of groundwater bodies [1]. In coastal groundwater bodies, intrusion is one of the main issues that need to be considered to achieve or maintain good groundwater status.

The impacts of SWI on groundwater have a heterogeneous distribution. Analyses of spatiotemporal distribution of SWI require salinity or chloride concentration to be mapped for different dates. Depending on the issue to be addressed and the available information, SWI can be approximated using various models. SWI can be mapped either by applying simple interpolation models [2, 3] to existing point data or by simulating the physical processes using transient density-dependent groundwater flow models [4, 5]. The results obtained with these physical process models can be applied to assess sustainable management strategies, i.e., strategies that prevent deterioration of the aquifer resource due to SWI [6]. They can even be employed

to propagate impacts of potential local climate change (CC) [7] or global change (GC) scenarios and to identify adaptation strategies [8].

Based on the salinity or chloride concentration maps at different dates, some authors have defined indices to summarize SWI [9–12]. These indices provide an overview of the intensity and spatial distribution or percentage SWI at aquifer scale. Such indices need to offer descriptive and synthetic results so that the status of SWI in different aquifers and over different periods can be compared. These index-methods [9, 13] establish threshold values for chloride to define the basis of SWI that have been defined in various ways: by referencing natural background levels and/or by taking into account the concentrations required to protect dependent ecosystems or human health [14].

When simple interpolation is used to draw the maps used to define the indices, analyses must be limited to the historical period for which there are data. In contrast, physical process models can be applied to propagate various potential conditions and so maps can be obtained for different scenarios (e.g., alternative management scenarios or future potential CC scenarios); in this latter case, the output can be used to determine the optimum strategy and therefore support the decision-making process [15].

In addition to the SWI status and dynamics, another important issue to take into account in identifying sustainable management strategies for coastal aquifers is the aquifer's vulnerability to SWI. In the literature we found various methods for mapping SWI vulnerability, such as the GALDIT method [16]. Then, by applying a method to express vulnerability as an index, we can also get an overview of the intensity and spatial distribution or percentage of SWI at aquifer scale [13].

2. Related Works

In the literature we find many tools to assess and analyze water resource problems [17–19]. The success of these software tools lies in their usability. A user-friendly environment and the implementation in commonly used software are key factors for their success and popular use. For example, groundwater studies usually employ Geographic Information Systems (GIS) since they are powerful, widely available tools that can deal with large amounts of spatial georeferenced information [20] and make calculations in an efficient way to provide quick results [21]. The analysis and mapping of hydrogeological data provide useful spatio-temporal information to decision makers [22].

GIS tools have been widely used for different purposes related to groundwater issues [23, 24]. Several authors have developed different source codes within GIS environment (Alcaraz et al. [25]; Bhatt et al. [26]) for hydrogeological modelling. A free and open source module included in FREEWAT was developed by Criollo et al. [19] to analyze hydrochemical and hydrogeological data in order to simplify the characterization of the groundwater bodies at chemical risk. Almeida et al. [27] coupled a groundwater flow model into a GIS environment to simulate transient flow in a confined aquifer. Akbar et al. [28] and Ríos et al. [29] presented GIS-based models to simulate contaminants leaching into groundwater.

In coastal areas, a three-dimensional GIS-based groundwater flow model was developed [30] to simulate the aquifer's response to past climate changes. A new ArcGIS tool for groundwater flow simulation and visualization of results was also implemented by [19]. Other authors (De Filippis et al. [31]) applied a previously developed GIS-based tool (AkvaGIS), in addition to a groundwater flow model, to study the impacts of pumping on seawater intrusion in coastal aquifers in Malta and Italy. This tool was used in other studies (Perdikaki et al. [32]) to analyze hydrochemical parameters in a coastal aquifer that presented seawater intrusion problems.

Nevertheless, as far as we know, there is no ArcGIS tool focusing on analyzing seawater intrusion (SWI) status and vulnerability at aquifer scale.

In this paper, we describe the development of a new ArcGIS tool called GIS-SWIAS, which is the implementation of the index-based method for assessing aquifer status and vulnerability to SWI proposed by [13, 15]. It helps to analyze SWI status and/or vulnerability at aquifer scale using a mixed lumped-distributed analysis. It is a user-friendly ArcGIS® toolbox that performs all the required calculations for specific dates or temporal periods inside a GIS environment. The data inputs to the model are hydraulic head and chloride concentration maps. The tool provides two options to map these variables: the first is to use point data by applying interpolation techniques integrated within GIS-SWIAS, while the second is to take these data from existing external distributed models. The second approach allows different climatic and/or management scenarios to be assessed and compared. From those maps, extensive calculations have been fully automatized in GIS-SWIAS to display the results as distributed maps of affected and nonaffected volumes (at a specific moment or over a period of time), mean conceptual cross sections, and a lumped index (Ma and L_Vul) to analyze the global intensity and the dynamics of SWI.

Although there are many GIS-based tools in the literature that allow simulating groundwater flow and analyzing groundwater quality, none of them perform spatial and temporal analysis on groundwater quality and vulnerability issues. Moreover, this new tool provides simple pictures that summarize the proportional affected area within the aquifer according to a chloride threshold. For this purpose, GIS-SWIAS has been applied to analyze the seawater intrusion problem, but this tool could be applied to represent the global status of an aquifer to any contaminant. The main objective of GIS-SWIAS is to provide an easy-to-use tool through a user-friendly interface that can be used by users at different levels of expertise to summarize the SWI problem at aquifer scale. It allows analyzing long time periods with a low computing cost.

3. Description of GIS-SWIAS Tool: Models, Inputs, and Outputs

GIS-SWIAS is an ArcGIS ArcToolbox that contains the models required to analyze SWI status and vulnerability at aquifer scale, according to the methodology described in

previous works [13, 15]. Figure 1 shows the structure of the tool, which includes inputs, steps, and models, as well as the outputs generated.

To determine the overall status of the aquifer, the **inputs** to the tool include variables (to characterize the historical evolution of hydraulic head and chloride concentration) and parameters (to define aquifer geometry and hydrodynamic behaviour). Data describing the historical evolution can come from direct observations (monitoring network) or other techniques (geophysical applications, etc.). For the vulnerability assessment, an additional input is required: a distributed vulnerability index map, which comes from other intrinsic information (aquifer type, conductivity, distance from the shoreline, and bicarbonate concentration).

The results/**outputs** produced to summarize status and vulnerability to SWI through visual displays and time series are as follows: (1) maps of aquifer volumes affected by SWI; (2) 2D conceptual cross sections (with mean penetration inland and mean thickness on specific dates, or mean values over a period of time); (3) lumped index (mass of chloride that causes the concentration in some areas to exceed the SWI threshold and lumped vulnerability index) to summarize the global dynamic of SWI within the aquifer.

3.1. Description of the Outputs: Theoretical Background. In order to assess the maps of SWI-affected aquifer volumes for different dates, we need to compile (A) maps of chloride concentration or vulnerability to SWI; (B) maps of groundwater volumes for specific dates; (C) threshold of chloride concentration or vulnerability that will be used to tag which parts of the aquifer are impacted (areas with chloride concentration or vulnerability index exceeding the threshold). The tool provides two options with respect to the input maps (A) (chloride or vulnerability maps) and (B), either calculating the maps internally from point data by applying interpolation techniques integrated within GIS-SWIAS or taking the maps from existing external distributed models; the second option means that various potential climatic and/or management scenarios can be compared and assessed. Maps of groundwater volume are calculated by combining hydraulic head, geometry, and the storage coefficient. The maps of groundwater volume and chloride concentration are combined to assess the aquifer volume affected by using a chloride threshold (V_r). This threshold is assumed to be equal to the natural background level of the aquifer, or the reference quality standard determined by authorities in order to maintain a good groundwater status. The affected volume is defined as the groundwater volume of resource whose chloride concentration is above the established threshold.

2D conceptual cross section depicts the magnitude of the intrusion process in the aquifer at a specific moment, or the mean values in a period of time. The cross section is defined orthogonal to the shoreline. It summarizes the mean geometry of the affected volume, i.e., the mean thickness and inland penetration of the aquifer volume with chloride concentration above the threshold. The average affected thickness $T_{ha}(m)$ and inland penetration $P(m)$ of the intrusion can be calculated by summing the values in each cell i of the aquifer mesh where the chloride concentration exceeds the threshold:

$$T_{ha}(m) = \frac{\sum V_{i(>V_r)}}{\sum S_{i(>V_r)}},$$

$$P(m) = \frac{\sum V_{i(>V_r)}}{T_{ha} * L_{\text{coast}}}, \tag{1}$$

$$V_{i(>V_r)}\left(m^3\right) = S_i\left(m^2\right) * b_i(m) * \alpha_i,$$

where $V_{\text{i(>Vr)}}$ is the groundwater volume (m^3) in the cell i with a chloride concentration (or vulnerability) exceeding V_{r}; S_{i} is the surface area (m^2) of the cell i with chloride concentration (or vulnerability) exceeding V_{r}; b_{i} is the saturated thickness (m) within the cell i with Cl concentration (or vulnerability) above V_{r}; α_{i} is the storage coefficient in the cell i; L_{coast} is the length of coastline (m).

The mean chloride concentration (C) of the affected volume is

$$C\left(\frac{mg}{l}\right) = \frac{\sum\left(C_{i(>V_r)} * V_{i(>V_r)}\right)}{V_{(>V_r)}}, \tag{2}$$

where $C_{i(>V_r)}$ is the chloride concentration (mg/l) in cell i; $V_{i(>Vr)}$ is the groundwater volumes (m^3) in cell i with a concentration exceeding V_r; $V_{(>V_r)}$ is the total groundwater volume (m^3) of the affected area.

The increment of chloride concentration (IC) above the threshold (V_r) in the affected volume is

$$\text{IC}\left(\frac{mg}{l}\right) = C - V_r. \tag{3}$$

Both variables, the conceptual cross section and IC index, give an overview of the magnitude and intensity of the intrusion process per linear metre of coast at a specific moment in time.

The lumped index Ma (mass of chloride that causes the concentration in some areas to exceed the threshold) to summarize the global dynamic of SWI within the aquifer is obtained multiplying the increment of concentration (IC) by penetration (P) and affected thickness (T_{ha}) from (1) and (3).

$$Ma\left(\frac{kg}{m}\right) = P(m) * IC\left(\frac{mg}{l}\right) * 10^{-3} * T_{ha}(m). \tag{4}$$

The vulnerability to SWI (or vulnerability to contamination in general) is assessed and summarized following the same steps to assess the SWI status. In this case, instead of chloride concentration values, a distributed map of groundwater vulnerability has to be generated by applying any index-based method (e.g., GALDIT) and the threshold applied to identify the affected area is defined by a specific vulnerability class (e.g., high or very high vulnerability).

The affected volume corresponds to the groundwater that presents vulnerability values above the threshold (e.g., very high vulnerability). The average affected thickness

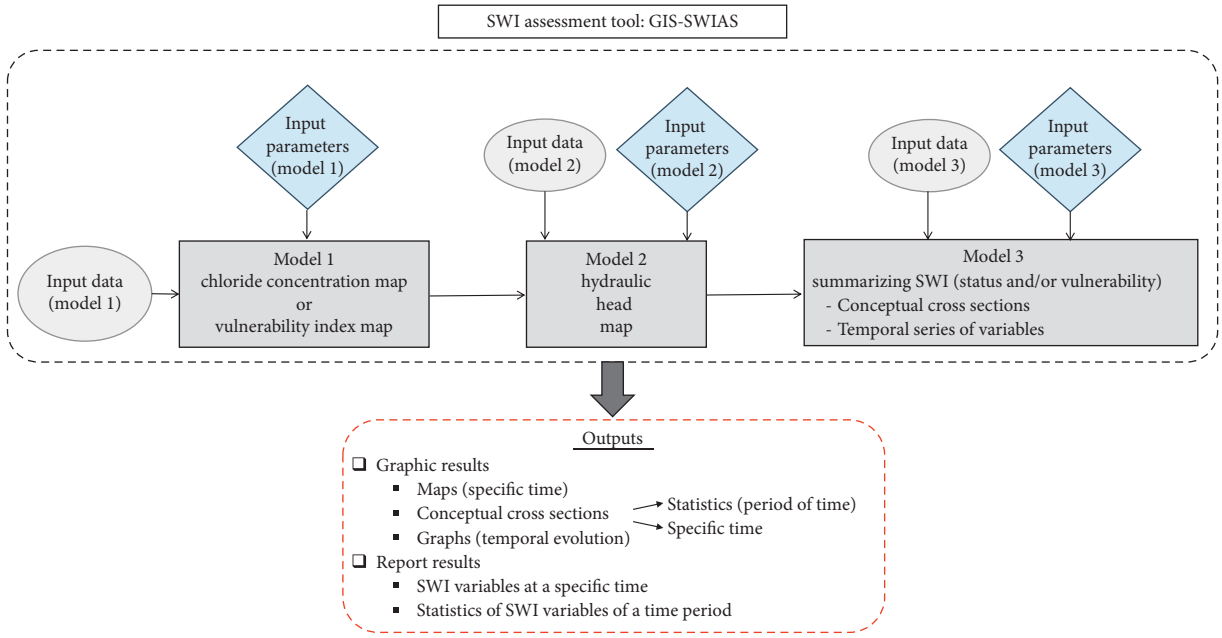


Figure 1: Workflow of the GIS-SWIAS ArcToolbox.

T_{ha}(m) and inland penetration P(m) are calculated by applying (1).

The lumped index to assess vulnerability is

$$L\_\mathrm{Vul}(-)\frac{\sum\left(\mathrm{Vul}_{i(>V_r)}\right) * V_{i(>V_r)}}{V_{(>V_r)}}, \quad (5)$$

where $\mathrm{Vul}_{i(>V_r)}$ is the value of the vulnerability index (-) in cell *i*.

The concept of Ma and L_Vul involves some simplifications, in accordance with the definition of the conceptual cross sections. While 2D maps and cross sections summarize the extent and magnitude of SWI and vulnerability in an aquifer at a specific time, Ma and L_Vul indices show the lumped intensity and temporal dynamic of the SWI and vulnerability to SWI at aquifer scale.

3.2. Tool Programming in ArcGIS. GIS-SWIAS is an ArcGIS ArcToolbox composed of a chain of models programmed in ModelBuilder. ModelBuilder is a visual programming language that allows chaining and sequencing several geoprocessing ArcGIS tools through a user-friendly interface. ModelBuilder is available within the tool bar in ArcGIS. It allows the addition of any geoprocessing tool of ArcGIS by linking and providing its output and transferring it to another tool as input.

Programming in ModelBuilder allows us to automate a workflow to create a model, which can be documented and shared as a ArcGIS tool. ModelBuilder contains a script tool to link with Python scripts and other models. It also allows iteration of a workflow, so it can be very useful to analyze the evolution of the historical hydrogeological processes.

The three models that compose GIS-SWIAS have been compiled by adding different tools from ArcToolbox to create shapes from point data, to interpolate, etc. Figure 2 shows the workflow of one of the three models.

Although ModelBuilder is an intuitive and easy-to-use tool, the integration of lots of geoprocesses in the same model may be difficult. Because several geoprocesses have dynamic parameters and the user interaction is necessary, GIS-SWIAS was divided into three steps (models) that have to be executed following the workflow shown in Figure 1.

3.3. Description of the Models within GIS-SWIAS. GIS-SWIAS contains three ModelBuilder models (Figure 2) that can be applied individually or used sequentially to produce a complete lumped assessment of the SWI at aquifer scale. GIS-SWIAS can be shared with other users and it can be added as a toolbox as shown in Figure 3. The run sequence follows the order shown in the workflow (Figure 1): "Chloride concentration map", "Hydraulic head map," and "Summarizing SWI". These models can be run within the ArcToolbox window providing a user-friendly graphical user interface.

3.3.1. "Chloride Concentration map" Model. The "Chloride concentration map" model generates a classified chloride concentration shapefile from a point feature table in text format by using Inverse Distance Weighting (IDW) interpolation technique (other interpolation techniques could be implemented in this tool). It can also import chloride concentration fields from Visual MODFLOW files. The dialog box shown in Figure 4 is opened by double-clicking the model tool on the ArcToolbox window.

The model requires a polygon shapefile of the aquifer and the workspace containing a text file for each date to be analyzed. The text files have to include *X* and Y UTM coordinates that define the locations of the point features

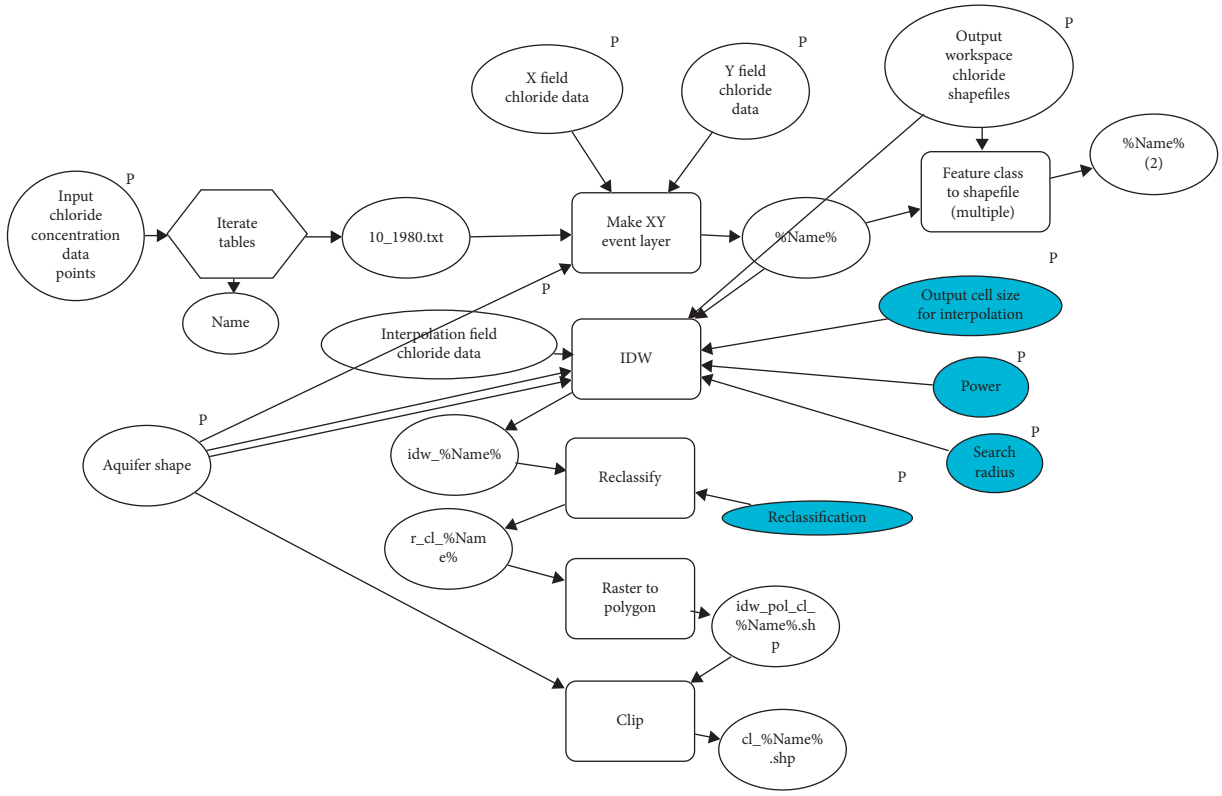


Figure 2: Workflow of "Chloride concentration map" model.

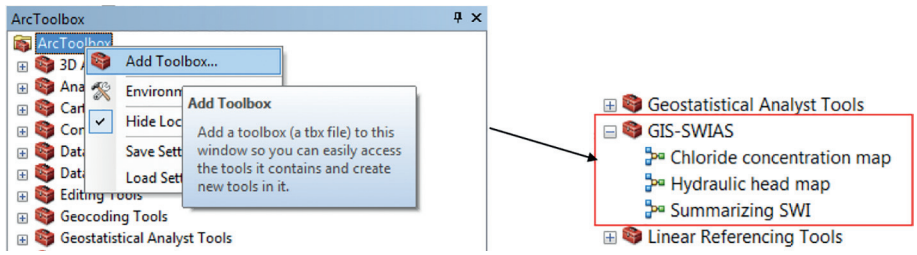


Figure 3: GIS-SWIAS in the ArcToolbox.

and chloride concentration values (mg/l) in the observation wells. Text files also have to include column header, as shown in Figure 4. The text filename cannot exceed eight characters nor include blank spaces or special characters (underscore can be used as a substitute). It is not necessary that all points contain data every date of the period to be analyzed.

The user has to indicate the fields (columns) in the input table that contain the *X* and *Y* coordinates and the chloride concentration value for each point (Figure 5). Optional settings concerning IDW interpolation techniques can be changed by the user. A reclassification of the values after interpolation is also required. Finally, the user has to choose a folder where the output shapefiles will be saved.

When all the required parameters are filled out, the model can be executed by clicking "OK" at the bottom of the dialog box. The execution screen (Figure 6) shows the running processes and it can be closed when the execution has successfully completed.

The "Chloride concentration map" model provides the following shapefiles for each date analyzed: (1) point shapefile of chloride concentration data; (2) raster from data interpolation; (3) polygon shapefile from interpolation covering the default extension; and (4) polygon shapefile from interpolation clipped to the shape of the aquifer.

3.3.2. "Hydraulic Head Map" Model. The "Hydraulic head map" model generates a classified hydraulic head shapefile from a point feature table in text format. It also generates a shapefile containing aquifer variables (chloride concentration and hydraulic head values) and aquifer parameters (storage coefficient and bottom of the aquifer). The dialog box is shown in Figure 7.

This model has the same input requirements as for the "Chloride concentration map" model but focuses on hydraulic head data (m.a.s.l.). It also allows hydraulic head fields to be imported from Visual MODFLOW. The name of

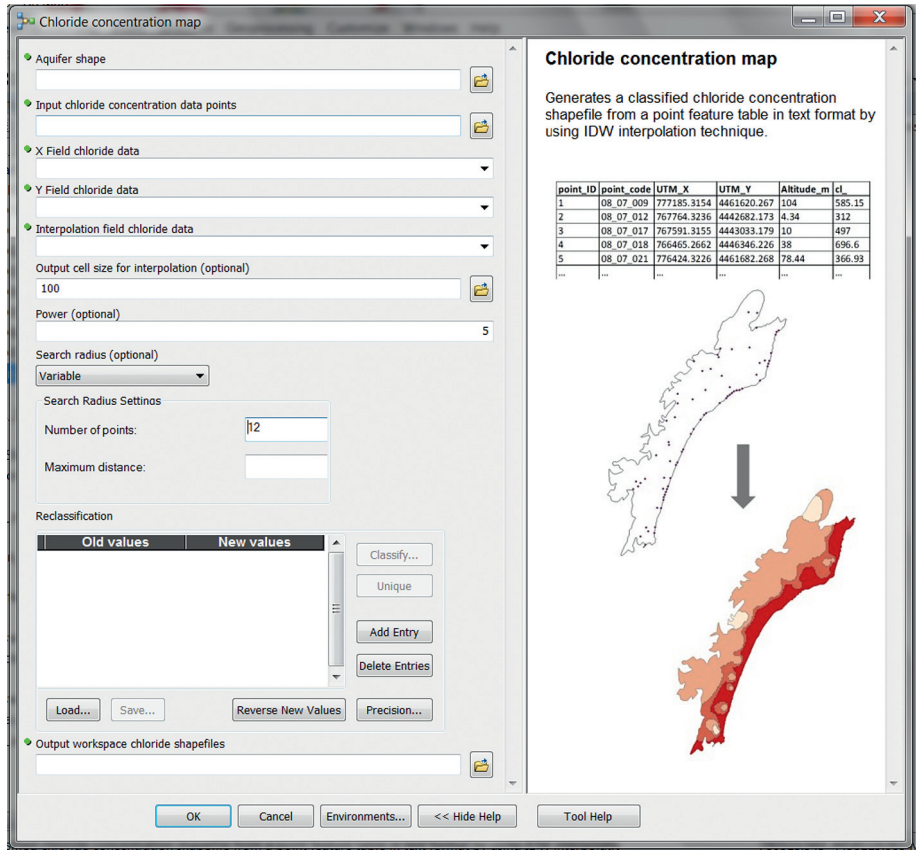


Figure 4: Dialog box of the "Chloride concentration map" model.

hydraulic head text files must be the same as the chloride concentration text files for each time period analyzed.

The user has to indicate the location of the chloride shapefiles generated in the previous model ("Chloride concentration map" model). It also requires polygon shapefiles of storage coefficient and bottom (m) of the aquifer as inputs.

The model generates shapefiles of hydraulic head data in an analogous way to the "Chloride concentration map" model. Moreover, it provides a polygon shapefile containing variables for each date analyzed (chloride concentration and hydraulic head values) and parameters (bottom and storage coefficient) of the aquifer. This shapefile is named "union_% name of the hydraulic head text file%_hh.shp", where "% name of the hydraulic head text file%" is variable if different dates are analyzed.

3.3.3. "Summarizing SWI" Model. For the "Summarizing SWI" model, the methodology proposed in Baena-Ruiz et al. [13,15] and described in Section 3.1 has been implemented in the ArcGis environment. This tool generates Excel® tables containing statistics that summarize SWI at aquifer scale. It also generates conceptual cross sections (.shp), where the mean affected and nonaffected volumes are drawn for the aquifer (average values over a time period or instantaneous values on a specific date). If different dates are analyzed, it shows graphs representing the temporal evolution of Pa and Ta variables, percentage of affected volume, chloride concentration within the aquifer, and Ma index (or lumped vulnerability index). The dialog box for global status assessment is shown in Figure 8.

The "Summarizing SWI" model requires the folder path where the results of the "Hydraulic head model" have been previously saved to be specified. In this folder, the shapefile named "union_%name of the hydraulic head text file% _hh.shp" contains chloride concentration, hydraulic head, bottom of the aquifer, and storage coefficient fields. The user has to select from the pull-down list the corresponding column in the input shapefile for each field, as shown in Figure 9.

The next required parameter in this tool is the "Chloride threshold." It is defined as the chloride concentration value above which the aquifer is considered to be affected by SWI. This threshold may be set as the natural background level of the aquifer or as the relevant environmental quality standards. Hinsby et al. [14] proposed a method to calculate groundwater thresholds values based on these criteria.

The shoreline length (m) is also required for subsequent calculations.

X and Y axes establish the coordinate system of the conceptual cross sections. The GIS-SWIAS tool provides polyline shapefiles for X and Y axes located at (0,0), but the user can translate them to another coordinate origin or create new ones.

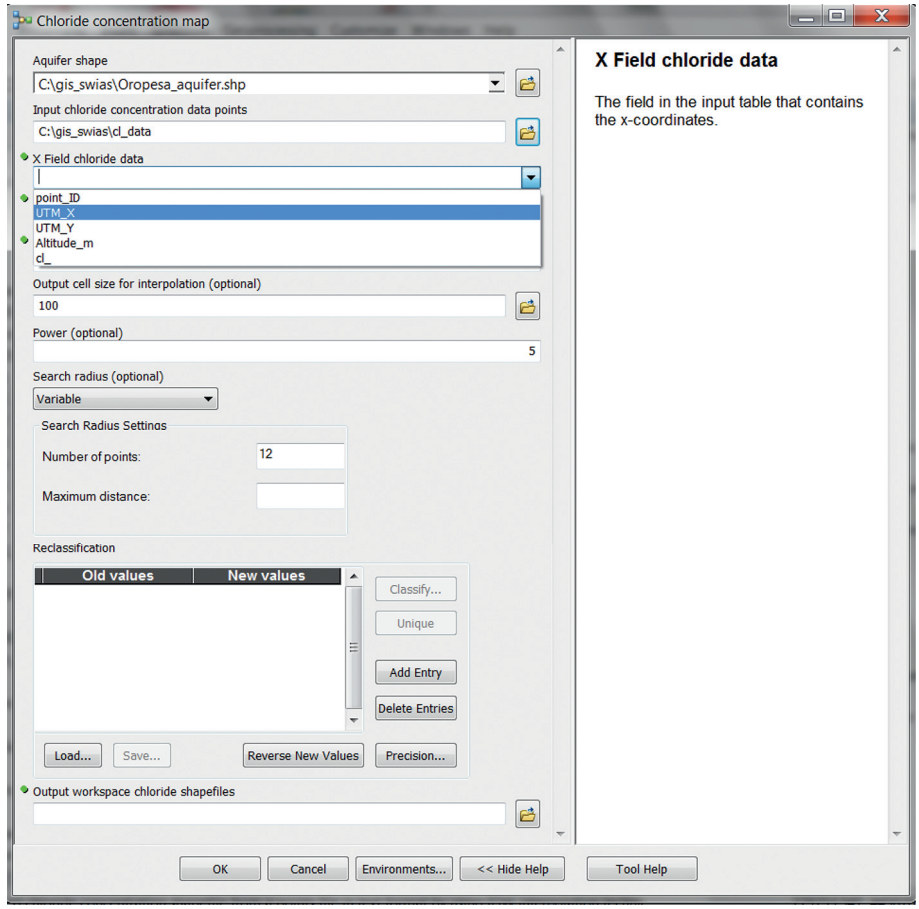


Figure 5: Setting options defined by the user.

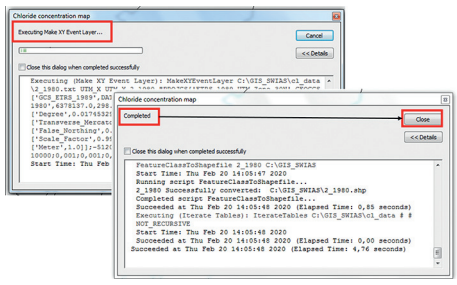


Figure 6: Execution screen of the model.

The vertical scale factor is used to rescale the vertical magnitude (Ta) of the conceptual cross section if the factor Ta/Pa is too small. If vertical scale factor = 1, the conceptual cross section will maintain the real size ratio.

Finally, two paths where the output results will be saved are required. "Output workspace statistic" will contain lumped variables reports in Excel table format for each date analyzed (Figure 10) and mean statistics for the entire period. Four graphs will be also saved in this path: (1) temporal evolution of Pa and Ta variables; (2) percentage volume affected; (3) Ma index; and (4) chloride concentration within the aquifer (mean chloride concentration in the aquifer, mean chloride concentration in the affected volume, and the increment of concentration within the affected volume above the threshold).

"Output workspace results" will contain the polygon shapefiles that allow the (1) mean affected and (2) non-affected conceptual cross section within the aquifer for each date analyzed to be drawn, (3) the mean affected and (4) nonaffected conceptual cross section within a time period, and the (5) maximum affected cross section for a time period. These two paths can be the same for all results, but they have to be different from the output paths of the previous models.

Figure 11 and Table 1 show the graphical and statistical summaries, respectively, from the GIS-SWIAS tool.

The GIS-SWIAS tool can provide results for each date where information is available; these are obtained by iterative application of the described method. GIS-SWIAS allows historical [13] and future periods [15] to be analyzed if the hydraulic head and chloride maps come from a density-dependent flow model. By this means, GIS-SWIAS can be used to analyze adaptation strategies [15] in terms of reducing SWI, taking into account future potential scenarios that might include CC and/or GC, also considering projected land use change scenarios (new urbanized areas, crop transformations) [15].

Moreover, this tool may be also used to summarize SWI vulnerability, for any index method applied to assess it. In

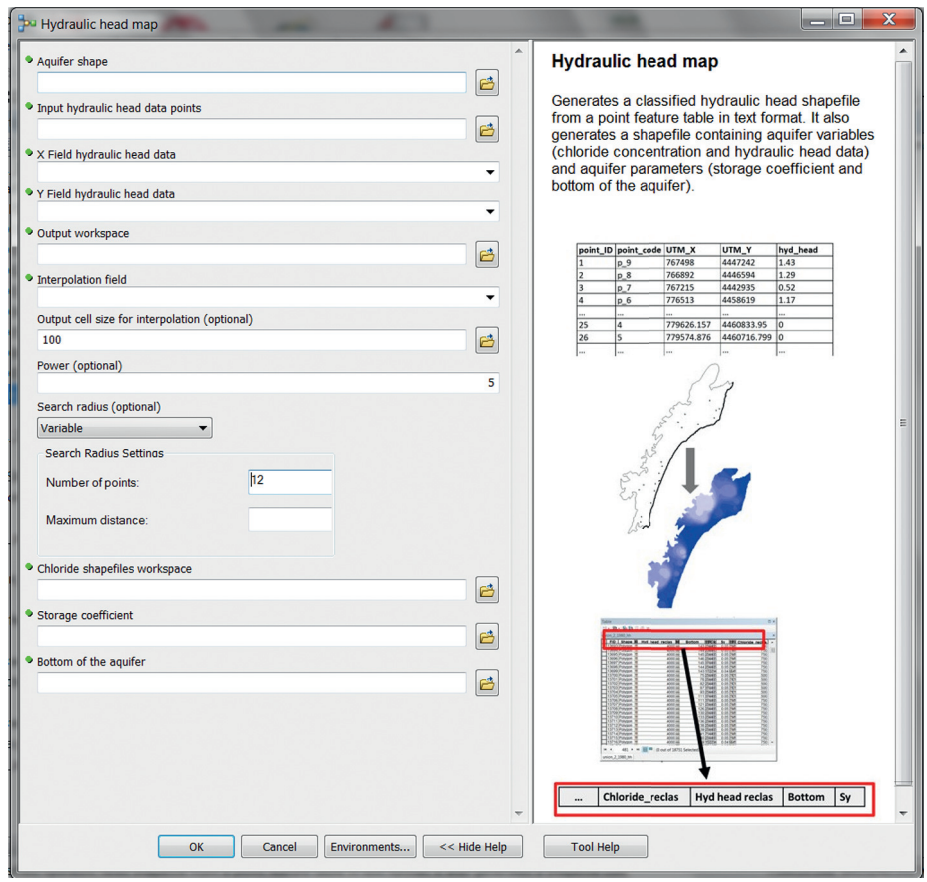


Figure 7: Dialog box of the "Hydraulic head map" model.

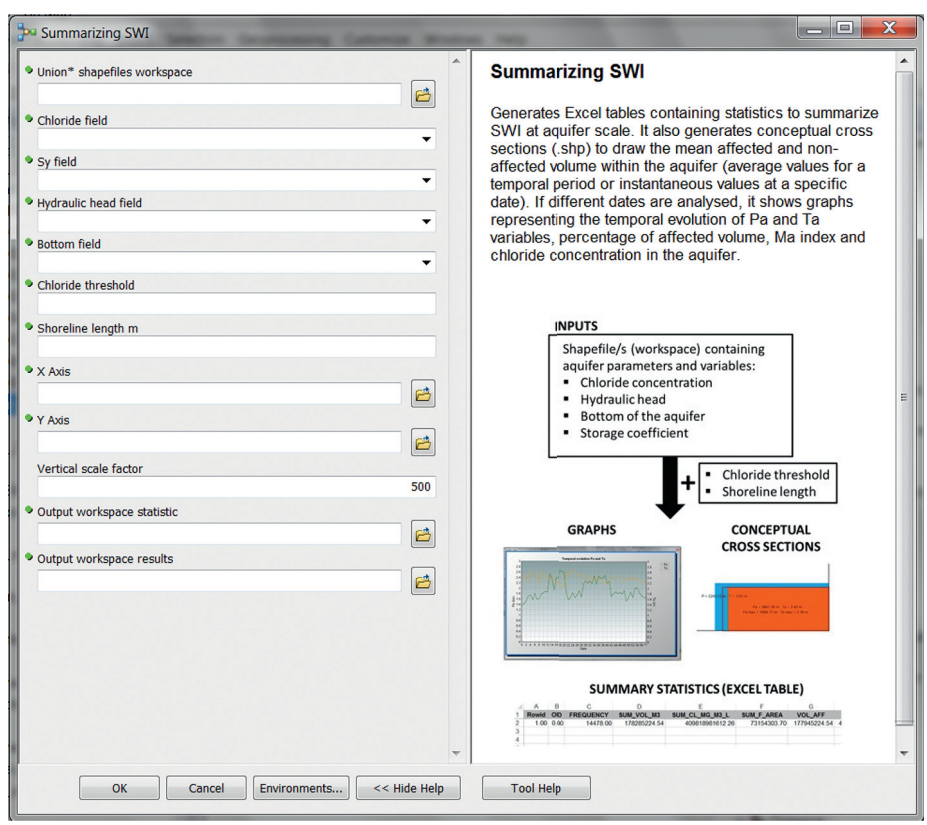


Figure 8: Dialog box of the "Summarizing SWI" model.

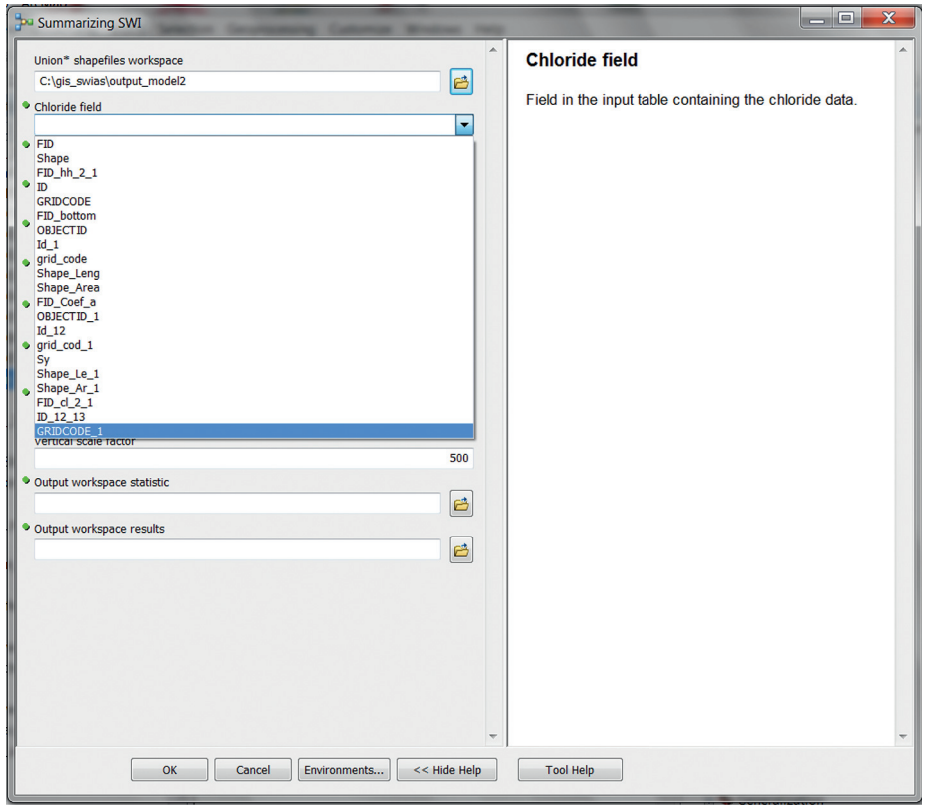


Figure 9: Selection of chloride field from the input table.

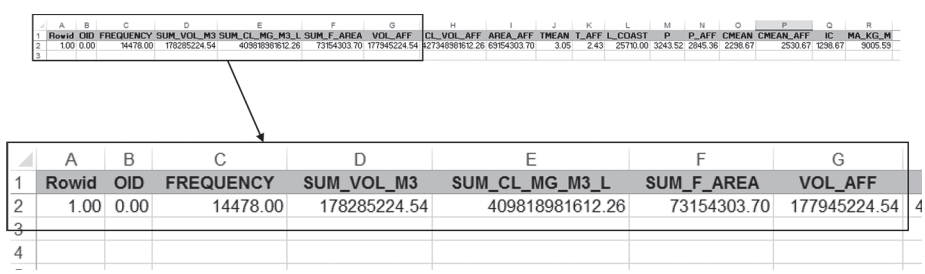


Figure 10: Summary statistics for a date.

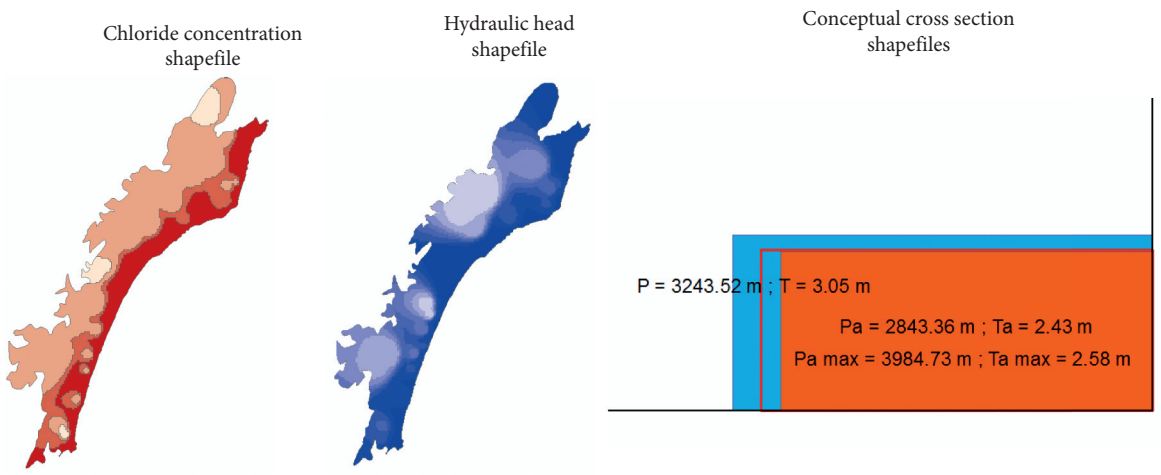
this case, instead of the chloride concentration maps, generated by executing the "Chloride concentration map" model, polygon shapefiles of the vulnerability index (previously prepared by the user) would be used as inputs of the model "Summarizing SWI vulnerability" (Figure 12), which also will require the "Hydraulic head map" shapefile generated by the tool, as previously described.

The vulnerability index maps must be included as numerical fields (values obtained before defining the vulnerability classes). In order to generate the conceptual cross sections that summarize the "affected" aquifer volume, i.e., where the vulnerability to SWI is identified, the tool requires a vulnerability threshold to be input that represents the reference value chosen to distinguish between affected and nonaffected volumes. This threshold will be also used to assess the lumped vulnerability index.

Just as in the definition of the Ma index, the lumped global value of vulnerability in the aquifer on a specific date is obtained by weighting the vulnerability score in each cell with its water storage. This lumped index also allows an analysis of the dynamic of SWI vulnerability of the system at aquifer scale to be performed. The lumped index can be also obtained using different threshold values [13, 15].

4. Discussion

GIS-SWIAS is a user-friendly polyvalent ArcGIS tool that provides a comprehensive overview of SWI status and vulnerability at aquifer scale. It integrates three models, which are documented in order to briefly explain the tool's description, its utility, and the data required for each item. This tool can be applied by scientists and decision makers, who may not be advanced users of GIS, to summarize SWI problems. Many GIS-based tools have demonstrated to be powerful and cost-effective to analyze groundwater issues (Criollo et al. and Perdikaki et al. [19, 32]). Moreover, GIS



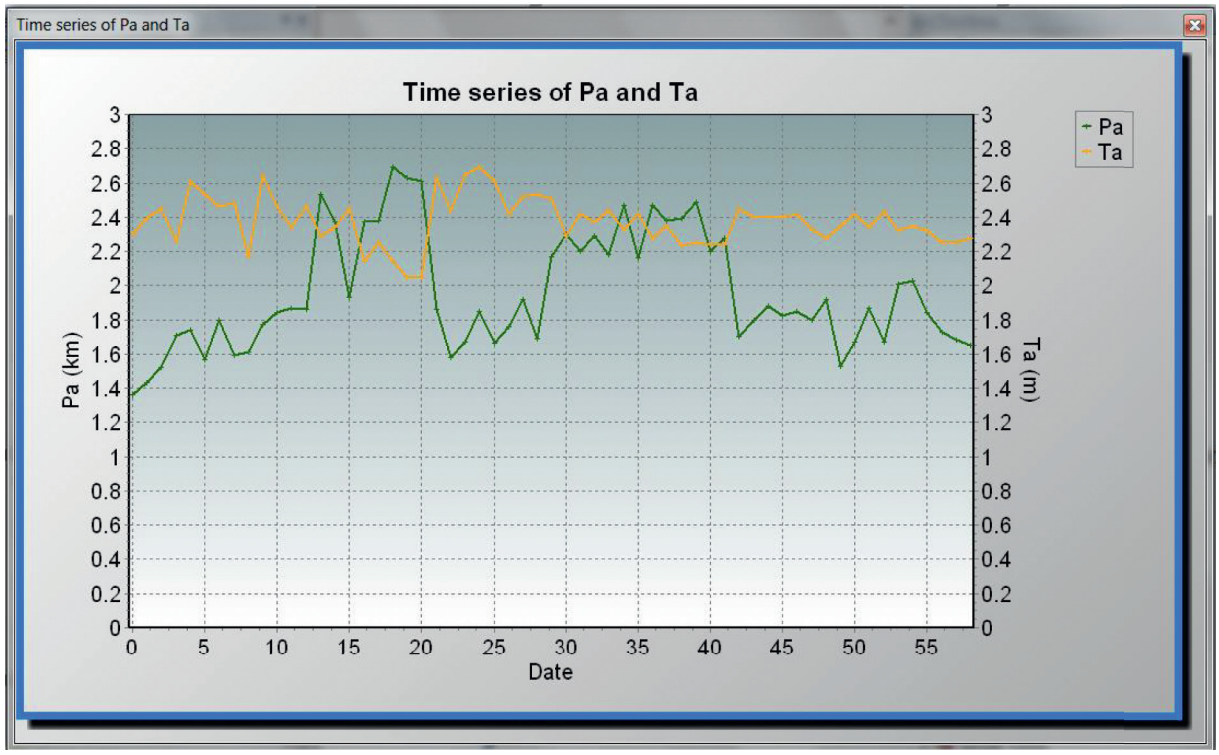


Figure 11: Graphical output from the GIS-SWIAS tool.

models as ModelBuilder models can be integrated into other platforms by using the Python script tool (Menezes and Inyang [33]).

Due to the heterogeneous distribution of seawater intrusion, **distributed information and assessments** are required to study its impacts [8, 30]. For this reason, the **methods** for modelling [34, 35] SWI impacts and the user-friendly tools developed based on them [36–38] also require distributed inputs and calculations. The GIS-SWIAS is a tool that could be classified as a postprocessing tool to summarize and help in the analysis of SWI impacts at aquifer scale. This tool produces both distributed and lumped results at aquifer scale, but, logically, it also requires distributed inputs and assessments, as described in the previous sections. In this group of postprocessing tools, we find in the literature, for example, [39]. GIS-SWIAS is a new tool, in which the method proposed by [13, 15] has been implemented. A significant novelty of this method with respect to other previously developed methods is that the proposed lumped index to summarize SWI status at aquifer scale is based on a variable with physical meaning (mass of chloride that causes the concentration in some areas to exceed the natural threshold). On the other hand, a novel aspect of this tool is that, from the distributed information and calculations, GIS-SWIAS allows easy computation of the affected volume containing a chloride concentration above a threshold. This

Table 1: Lumped variables output (Excel® table format) from the GIS-SWIAS tool.

Lumped variables (Excel table)	
At a specific moment in time	Statistics for a time period
Total aquifer volume	Average aquifer volume
Total aquifer affected volume	Average aquifer affected volume
Total chloride concentration ∗ aquifer volume	Average chloride concentration ∗ aquifer volume
Total chloride concentration ∗ aquifer volume in the affected volume	Average chloride concentration ∗ aquifer volume in the affected volume
Total aquifer area	Average aquifer area
Total aquifer affected area	Average aquifer affected area
Shoreline length	Shoreline length
Average aquifer thickness	Average aquifer thickness within the period
Average affected aquifer thickness	Average affected aquifer thickness within the period
Average orthogonal distance of the aquifer to the shoreline	Average orthogonal distance of the aquifer to the shoreline within the period
Average orthogonal distance of the affected aquifer to the shoreline	Average orthogonal distance of the affected aquifer to the shoreline within the period
Average chloride concentration within the aquifer	Average chloride concentration within the aquifer within the period
Average chloride concentration within the affected volume	Average chloride concentration within the affected volume within the period
Increment of chloride concentration above the threshold value	Increment of chloride concentration above the threshold value within the period
Ma index	Average Ma index within the period
Percentage of affected volume	Average percentage of affected volume within the period

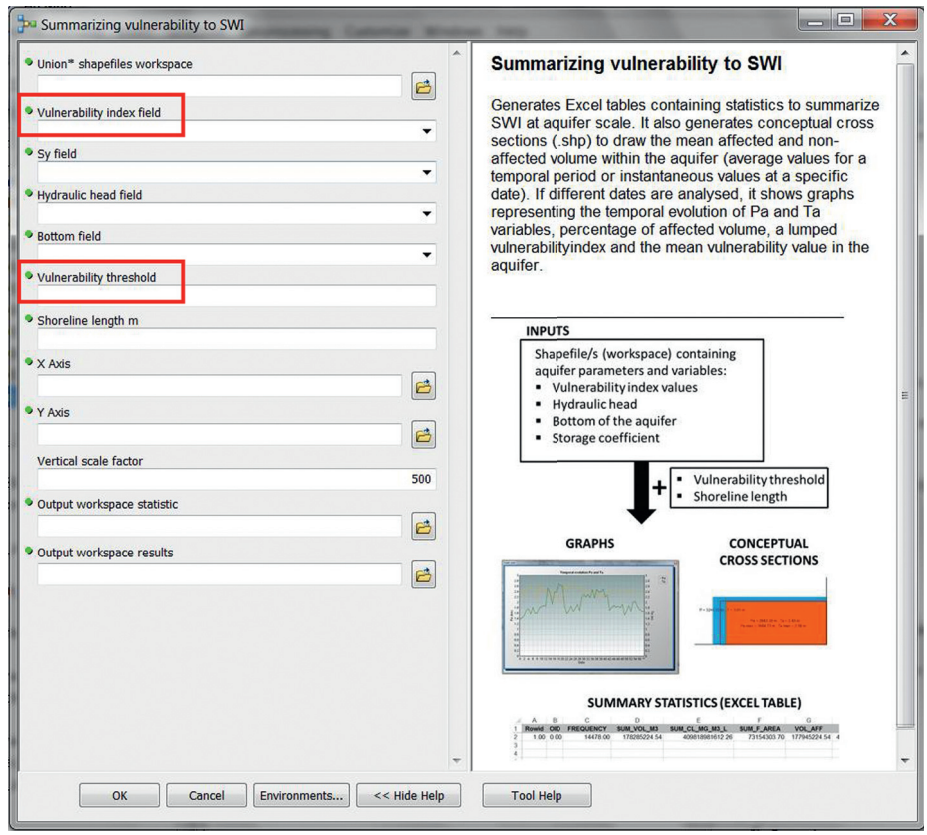


Figure 12: Dialog box of the "Summarizing SWI vulnerability" model.

tool also helps to produce lumped SWI outputs (indices) at aquifer scale. It produces valuable information that helps draw conclusions about the dynamic at aquifer scale, in terms of affected volume and global SWI intensity. Thus, it also provides insight into aquifer resilience and trends. Therefore, it will help to identify coastal groundwater bodies that require new management strategies to be implemented to achieve a good status.

The identification of SWI (the phenomenon that we want to analyze) requires a threshold value established that defines the inflection point beyond which the aquifer begins to register an impact. Previous research shows that the impact of SWI is significantly sensitive to the choice of the threshold value adopted [13]. The significant uncertainties around determining these threshold values [14] and the sensitivity of whether the aquifer is reported as being impacted by SWI or not increase the practical interest of the GIS-SWIAS tool: it is capable of performing the extensive calculations required to summarize SWI at aquifer scale, for the analyses of both historical and potential scenarios, considering different threshold values, which allow the comparison of the results.

With respect to the maps employed as inputs, the tool allows two options: to generate maps from available data using different interpolation techniques integrated in the tool and to take the maps directly from SEAWAT files. This functionality—which allows maps to be generated from point data or to be loaded from other commonly employed tools—has also been implemented in other SWI assessment tools [36, 37]. However, as far as we know, it is not available in postprocessing tools. In cases where map inputs are taken from density-dependent models, a comparative assessment of different scenarios (climatic conditions and/ or management strategies) could be performed. The physical-process approach can be applied to propagate and compare various potential conditions, and so in this case maps can be obtained and compared for different scenarios (e.g., management scenarios or future potential CC scenarios); this means that the output of the tool can support the decision-making process [15]. In contrast, when the maps employed to define the indices are obtained by applying simple interpolation approaches, analysis is limited to the historical period for which the data are available.

The tool also helps to analyze the vulnerability to seawater intrusion at aquifer scale. In the literature, we find different methods to assess groundwater vulnerability depending on the drivers of pollution (Aller et al., Vias et al., and Baena-Ruiz and Pulido-Velazquez [40–42]), pumping (Pulido-Velazquez et al. [43, 44]), and SWI [12, 16]. User-friendly tools have appeared to assist in this assessment; some of them were developed in a GIS environment [45]. Nevertheless, there are no tools that help in the assessment of SWI vulnerability with that focus on postprocessing. Therefore, this is the first postprocessing tool described that integrates SWI status and vulnerability assessment, which is very valuable information to help identify the significance of SWI problems in aquifers and potential sustainable solutions.

The GIS-SWIAS tool has been applied to two different case studies in the Mediterranean area of Spain (Plana de Oropesa Torreblanca and Plana de Vinaroz), obtaining the results described in previous papers [13, 15]. In [13], the process automation to generate the interpolated maps enabled authors to analyze SWI status and vulnerability over an extended time period (1977–2015) and to prove the sensitivity of results to the chloride threshold value (two threshold values were analyzed: 250 mg/l and 1100 mg/l) in Plana de Oropesa-Torreblanca and Plana de Vinaroz aquifers. In [15], the impacts of future GC scenarios were analyzed in Plana de Oropesa-Torreblanca aquifer. The methodology from [13] was adapted to compare six potential future scenarios including adaptation strategies. The historical period came from 1973 to 2010 and the six future scenarios covered the period 2011–2035.

The underlying methodology implemented in GIS-SWIAS was applied in [13] by interpolating chloride maps and hydraulic head from observation points, whereas the information to generate the field maps in [15] was loaded from SEAWAT model. The results of these studies for the Plana de Oropesa-Torreblanca aquifer show differences that reveal that the physical-process SWI approximations obtained using the density-dependent flow model give a more accurate representation. Despite these differences, the results are in the same order of magnitude. Other authors who have developed indices related to SWI [6, 9] have also proved that results do not differ considerably by including three-dimensional salinity data. Furthermore, the approximation obtained using interpolation will depend on the number of observations and the distribution of these points within the aquifer.

Although this tool has been developed to analyze SWI problem, it could be applied to study the lumped impact of any contaminant on groundwater and/or the groundwater vulnerability by applying any vulnerability index. In this case, instead of the "Shoreline length" parameter to generate the cross section, other equivalent lengths (e.g., the aquifer length orthogonal to the groundwater flow direction) should be considered. Therefore, GIS-SWIAS fulfils the requirements of flexibility, sturdiness, easy interaction, and user-friendliness, which make it a useful tool in the decision-making process. It will allow using them as "share vision models/tools" to help in the discussion of management alternatives between stakeholders and administration representatives [46]. Many Decision Support Systems tools were not successful because they were not user-friendly [47, 48].

4.1. Assumptions and Limitations. In this section we summarize the main assumptions/limitations of the GIS-SWIAS tool and in the implemented methodology.

4.1.1. Underlying Methodology.

(i) The methodology implemented [13] provides lumped results at aquifer scale. An analysis based exclusively on these lumped results would lose information about the heterogeneity of SWI in the system.

(ii) Although the method can be applied in cases where a distributed flow model is not available, a proper assessment for a specific date requires sufficient distributed data about hydraulic head, chloride concentration, and aquifer bottom. On the other hand, to generate the required fields from the data, it needs to apply an interpolation approach whose parameters influence on the results. It is recommended that the user tests different interpolation

approaches and parameters in order to obtain results in agreement with expert opinion.

(iii) In order to assess the dynamic of SWI and the resilience of the aquifer, there is need to have data for several dates over a long period.

(iv) If the aim is to assess sustainable management strategies, appropriate hydrological models to propagate the impacts are required, not only of the climatic conditions, but also of the management strategies.

4.1.2. GIS-SWIAS Tool.

(i) This tool has been implemented in ArcGIS and is licensed GIS software. Future implementations may be developed in an open environment.

(ii) The GIS-SWIAS tool had to be divided into three models due to some difficulties that arose to implement all the geoprocesses into one model. Nevertheless, in the future, we plan to combine these three steps into one, improving the usability of the model.

(iii) Graphic results are based on a predefined template. Although graphs in ArcGIS are customizable, there is not so much potential.

(iv) Output reports are generated automatically in Excel format and lack a refined style.

5. Conclusions

In this paper we describe a new general tool, GIS-SWIAS. It is an ArcGIS-based tool, designed to analyze SWI status and vulnerability at aquifer scale by applying the method presented by [13, 15]. It is a user-friendly tool that allows georeferenced information to be dealt with, and it is easy to introduce the required data (inputs) and to efficiently perform the demanding computational operations required. Its outputs are in the form of reports and images to summarize the magnitude, intensity, and temporal evolution of SWI within an aquifer.

The GIS-SWIAS tool can be applied to assess historical SWI dynamic in case studies where we do not have a previous model. Nevertheless, if we want to analyze in a rational quantitative analysis of the various alternative management scenarios to manage SWI in a sustainable manner, the GIS-SWIAS tool will need to take information on hydraulic head and chloride concentration distribution generated from simulations of their impacts by a calibrated density-dependent flow model. In such cases, adaptation strategies to potential future scenarios, whose distributed impacts have to be propagated within the previously calibrated models, could usefully be analyzed and compared using this tool. GIS-SWIAS can be applied to assess not only SWI status at aquifer scale, but also vulnerability to any contaminant.

Given all these ways that the GIS-SWIAS tool can be applied, it provides a valuable tool for both researcher and technician to assess SWI dynamics and aquifer resilience under different management scenarios. It can support the decision-making process in the rational selection of sustainable management strategies. The tool's performance has been tested and confirmed in two case studies described in previous research works.

It can be applied to any case study. The easy-to-use workflow and the few input data required facilitate its application to a large number of case studies in order to compare SWI.

Data Availability

The software developed in this study may be released upon application to the authors, who can be contacted at l.baena@igme.es or d.pulido@igme.es.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work has been supported by the GeoE.171.008-TACTIC and GeoE.171.008-HOVER projects from GeoERA Organization funded by European Union's Horizon 2020 Research and Innovation Program and SIGLO-AN (RTI2018-101397-B-I00) project from the Spanish Ministry of Science, Innovation and Universities (Programa Estatal de I+D+I orientada a los Retos de la Sociedad).

References

[1] Water Framework Directive (WFD), "Directiva 2000/60/CE del Parlamento Europeo y del Consejo de 23 de Octubre de 2000," *Diario Oficial de las Comunidades Europeas de 22/12/2000*, Water Framework Directive (WFD), L 327/ 1–327/32, 2000.

[2] N. Momejian, M. Abou Najm, I. Alameddine, and M. El-Fadel, "Groundwater vulnerability modeling to assess seawater intrusion: a methodological comparison with geospatial interpolation," *Water Resources Management*, vol. 33, no. 3, pp. 1039–1052, 2019.

[3] V. Elumalai, K. Brindha, B. Sithole, and E. Lakshmanan, "Spatial interpolation methods and geostatistics for mapping groundwater contamination in a coastal area," *Environmental Science and Pollution Research*, vol. 24, no. 12, pp. 11601–11617, 2017.

[4] C. Llopis-Albert and D. Pulido-Velazquez, "Discussion about the validity of sharp-interface models to deal with seawater intrusion in coastal aquifers," *Hydrological Processes*, vol. 28, no. 10, pp. 3642–3654, 2014.

[5] C. Llopis-Albert and D. Pulido-Velazquez, "Using MODFLOW code to approach transient hydraulic head with a sharp-interface solution," *Hydrological Processes*, vol. 29, no. 8, pp. 2052–2064, 2015.

[6] A. Renau-Pruñonosa, I. Morell, and D. Pulido-Velazquez, "A methodology to analyse and assess pumping management strategies in coastal aquifers to avoid degradation due to seawater intrusion problems," *Water Resources Management*, vol. 30, no. 13, pp. 4823–4837, 2016.

[7] A.-J. Collados-Lara, D. Pulido-Velazquez, and E. Pardo-Igúzquiza, "An integrated statistical method to generate potential future climate scenarios to analyse droughts," *Water*, vol. 10, no. 9, pp. 1224–1248, 2018.

[8] D. Pulido-Velazquez, A. Renau-Pruñonosa, C. Llopis-Albert et al., "Integrated assessment of future potential global change scenarios and their hydrological impacts in coastal aquifers - a new tool to analyse management alternatives in the Plana Oropesa-Torreblanca aquifer," *Hydrology and Earth System Sciences*, vol. 22, no. 5, pp. 3053–3074, 2018.

[9] B. J. Ballesteros, I. Morell, O. García-Menéndez, and A. Renau-Pruñonosa, "A standardized index for assessing seawater intrusion in coastal aquifers: the SITE index," *Water Resources Management*, vol. 30, no. 13, pp. 4513–4527, 2016.

[10] M. Tomaszkiewicz, M. Abou Najm, and M. El-Fadel, "Development of a groundwater quality index for seawater intrusion in coastal aquifers," *Environmental Modelling & Software*, vol. 57, pp. 13–26, 2014.

[11] I. S. Babiker, M. A. A. Mohamed, and T. Hiyama, "Assessing groundwater quality using GIS," *Water Resources Management*, vol. 21, no. 4, pp. 699–715, 2007.

[12] A. Zeynolabedin and R. Ghiassi, "The SIVI index: a comprehensive approach for investigating seawater intrusion vulnerability for island and coastal aquifers," *Environmental Earth Sciences*, vol. 78, 2019.

[13] L. Baena-Ruiz, D. Pulido-Velazquez, A.-J. Collados-Lara, A. Renau-Pruñonosa, and I. Morell, "Global assessment of seawater intrusion problems (status and vulnerability)," *Water Resources Management*, vol. 32, no. 8, pp. 2681–2700, 2018.

[14] K. Hinsby, M. T. Melo, and M. Dahl, "European case studies supporting the derivation of natural background levels and groundwater threshold values for the protection of dependent ecosystems and human health," *The Science of the Total Environment*, vol. 401, no. 1-3, pp. 1–20, 2008.

[15] L. Baena-Ruiz, D. Pulido-Velazquez, A.-J. Collados-Lara et al., "Summarizing the impacts of future potential global cdslfthange scenarios on seawater intrusion at the aquifer scale," *Environmental Earth Sciences*, vol. 79, no. 5, p. 99, 2020.

[16] A. G. Chachadi and J. P. Lobo Ferreira, "Assessing aquifer vulnerability to sea-water intrusion using GALDIT method: Part 2—GALDIT indicator descriptions," in *Water in Celtic Countries: Quantity, Quality and Climate Variability*, J. P. Ferreira, Ed., Proceedings of the fourth interceltic colloquium on hydrology and management of water resources IAHS Press Publications, Guimarães, Portugal, pp. 172–180, 2007.

[17] J. Andreu, J. Capilla, and E. Sanchís, "AQUATOOL, a generalized decision-support system for water-resources planning and operational management," *Journal of Hydrology*, vol. 177, no. 3-4, pp. 269–291, 1996.

[18] E. Vázquez-Suñé, E. Abarca, J. Carrera et al., "Groundwater modelling as a tool for the European water framework directive (WFD) application: the llobregat case," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 31, no. 17, pp. 1015–1029, 2006.

[19] R. Criollo, V. Velasco, A. Nardi et al., "AkvaGIS: an open source tool for water quantity and quality management," *Computers & Geosciences*, vol. 127, pp. 123–132, 2019.

[20] D. Machiwal, M. K. Jha, V. P. Singh, and C. Mohan, "Assessment and mapping of groundwater vulnerability to pollution: current status and challenges," *Earth-Science Reviews*, vol. 185, pp. 901–927, 2018.

[21] A. Rikalovic, I. Cosic, and D. Lazarevic, "GIS based multicriteria analysis for industrial site selection," *Procedia Engineering*, vol. 69, no. 12, pp. 1054–1063, 2014.

[22] D. Machiwal and M. K. Jha, "Role of geographical information system for water quality evaluation," in *Geographic Information Systems (GIS): Techniques, Applications and Technologies*, D. Nielson, Ed., pp. 217–278, Nova Science Publishers, New York, NY, USA, 2014.

[23] W. Gossel, A. M. Ebraheem, and P. Wycisk, "A very large scale GIS-based groundwater flow model for the Nubian sandstone aquifer in Eastern Sahara (Egypt, northern Sudan and eastern Libya)," *Hydrogeology Journal*, vol. 12, no. 6, pp. 698–713, 2004.

[24] L. Wang, C. R. Jackson, M. Pachocka, and A. Kingdon, "A seamlessly coupled GIS and distributed groundwater flow model," *Environmental Modelling & Software*, vol. 82, pp. 1–6, 2016.

[25] M. Alcaraz, E. Vázquez-Suñé, V. Velasco, and R. Criollo, "A loosely coupled GIS and hydrogeological modeling framework," *Environmental Earth Sciences*, vol. 76, no. 11, p. 382, 2017.

[26] G. Bhatt, M. Kumar, and C. J. Duffy, "A tightly coupled GIS and distributed hydrologic modeling framework," *Environmental Modelling & Software*, vol. 62, pp. 70–84, 2014.

[27] C. N. Almeida, J. Roehrig, and E. Wendland, "Development and integration of a groundwater simulation model to an open geographic information system," *JAWRA-Journal of the American Water Resources Association*, vol. 50, no. 1, pp. 101–110, 2014.

[28] T. A. Akbar, H. Lin, and J. DeGroote, "Development and evaluation of GIS-based ArcPRZM-3 system for spatial modeling of groundwater vulnerability to pesticide contamination," *Computers & Geosciences*, vol. 37, no. 7, pp. 822–830, 2011.

[29] J. F. Rios, M. Ye, L. Wang, P. Z. Lee, H. Davis, and R. Hicks, "ArcNLET: a GIS-based software to simulate groundwater nitrate load from septic systems to surface water bodies," *Computers & Geosciences*, vol. 52, pp. 108–116, 2013.

[30] R. Ali, D. McFarlane, S. Varma, W. Dawes, I. Emelyanova, and G. Hodgson, "Potential climate change impacts on the water balance of regional unconfined aquifer systems in south-western Australia," *Hydrology and Earth System Sciences*, vol. 16, no. 12, pp. 4581–4601, 2012.

[31] G. De Filippis, C. Pouliaris, D. Kahuda et al., "Spatial data management and numerical modelling: demonstrating the application of the QGIS-integrated FREEWAT platform at 13 case studies for tackling groundwater resource management," *Water*, vol. 12, no. 1, p. 41, 2020.

[32] M. Perdikaki, R. C. Manjarrez, C. Pouliaris, R. Rossetto, and A. Kallioras, "Free and open-source GIS-integrated hydrogeological analysis tool: an application for coastal aquifer systems," *Environmental Earth Sciences*, vol. 79, no. 14, p. 348, 2020.

[33] G. B. Menezes and H. I. Inyang, "GIS-based contaminant transport model for heterogeneous hydrogeological settings," *Journal of Environmental Informatics*, vol. 14, no. 1, 2009.

[34] M. G. McDonald and A. W. Harbaugh, "A modular three-dimensional finite-difference ground-water flow model," *U.S. Geological Survey Techniques of Water-Resources Investigations*, Washington, DC, USA, 1988, p. 586, book 6, chap. A1.

[35] M. Bakker, F. Schaars, J. D. Hughes, C. D. Langevin, and A. M. Dausman, "Documentation of the seawater intrusion (SWI2) package for MODFLOW," *U.S. Geological survey*

techniques and methods, Washington, DC, USA, 2013, p. 47, book 6, chap. A46, https://pubs.usgs.gov/tm/6a46/.

[36] Waterloo Hydrogeologic Inc., *Visual MODFLOW User's Manual*, Waterloo Hydrogeologic. Inc., Waterloo, Canada, 2005.

[37] W. H. Chiang and W. Kinzelbach, *3D-Groundwater Modeling with PMWIN: A Simulation System for Modeling Groundwater Flow and Transport Processes*, vol. 2, Springer, New York, NY, USA, 2001.

[38] R. B. Winston, *ModelMuse: A Graphical User Interface for MODFLOW-2005 and PHAST*, p. 52, US Geological Survey, Reston, VA, USA, 2009.

[39] R. Criollo, V. Velasco, E. Vázquez-Suñé, A. Serrano-Juan, M. Alcaraz, and A. García-Gil, "An integrated GIS-based tool for aquifer test analysis," *Environmental Earth Sciences*, vol. 75, no. 5, p. 391, 2016.

[40] L. Aller, T. Bennett, J. Lehr, R. Petty, and G. Hackett, *DRASTIC: A Standardized System for Evaluating Groundwater Pollution Potential Using Hydrogeologic Settings, U.S. Environmental Protection Agency,* p. 622, Washington, DC, USA, 1987.

[41] J. M. Vías, B. Andreo, M. J. Perles, F. Carrasco, I. Vadillo, and P. Jiménez, "Proposed method for groundwater vulnerability mapping in carbonate (karstic) aquifers: the COP method," *Hydrogeology Journal*, vol. 14, no. 6, pp. 912–925, 2006.

[42] L. Baena-Ruiz and D. Pulido-Velazquez, "A novel approach to harmonize vulnerability assessment in carbonate and detrital aquifers at basin scale," *Water*, vol. 12, no. 11, p. 2971, 2020.

[43] D. Pulido-Velazquez, A.-J. Collados-Lara, L. Baena-Ruiz, F. Fernández-Chacón, and F. J. Alcalá, *Assessment of Present and Future Vulnerability to Pumping in Spanish Groundwater Bodies Using a Natural Turnover Time Index*, IAH, Dubrovnik, Croatia, 2017.

[44] D. Pulido-Velazquez, J. Romero, A.-J. Collados-Lara, F. J. Alcalá, F. Fernández-Chacón, and L. Baena-Ruiz, "Using the turnover time index to identify potential strategic groundwater resources to manage droughts within continental Spain," *Water*, vol. 12, no. 11, p. 3281, 2020.

[45] L. Duarte, A. C. Teodoro, J. A. Gonçalves, A. J. Guerner Dias, and J. E. Marques, "Assessing groundwater vulnerability to pollution through the DRASTIC method," in *Proceedings of the International Conference on Computational Science and its Applications*, pp. 386–400, Minho, Guimaraes, Portugal, July 2014.

[46] D. P. Loucks and E. Van Beek, *Water Resource Systems Planning and Management: An Introduction to Methods, Models, and Applications*, Springer, Berlin, Germany, 2017.

[47] R. F. Reitsma, "Structure and support of water-resources management and decision-making," *Journal of Hydrology*, vol. 177, no. 3-4, pp. 253–268, 1996.

[48] M. T. Atta, "The effect of usability and information quality on decision support information system (DSS)," *Arts and Social Sciences Journal*, vol. 8, no. 2, p. 257, 2017.

Hindawi
Scientific Programming
Volume 2020, Article ID 8847571, 11 pages
https://doi.org/10.1155/2020/8847571

Research Article

A Statistical Tool to Generate Potential Future Climate Scenarios for Hydrology Applications

Antonio-Juan Collados-Lara,[1] **David Pulido-Velazquez**,[1] **and Eulogio Pardo-Igúzquiza**[2]

[1]*Spanish Geological Survey, Urb. Alcázar del Genil 4, Bajo, 18006 Granada, Spain*
[2]*Spanish Geological Survey, Ríos Rosas 23, 28003 Madrid, Spain*

Correspondence should be addressed to Antonio-Juan Collados-Lara; ajcollados@gmail.com

Received 13 March 2020; Revised 18 May 2020; Accepted 22 May 2020; Published 29 June 2020

Academic Editor: Francisco J. Alcalá

Global warming associated with greenhouse emissions will modify the availability of water resources in the future. Methodologies and tools to assess the impacts of climate change are useful for policy making. In this work, a new tool to generate potential future climate scenarios in a water resources system from historical and regional climate models' information has been developed. The GROUNDS tool allows generation of the future series of precipitation, temperature (minimum, mean, and maximum), and potential evapotranspiration. It is a valuable tool for assessing the impacts of climate change in hydrological applications since these variables play a significant role in the water cycle, and it can be applicable to any case study. The tool uses different approaches and statistical correction techniques to generate individual local projections and ensembles of them. The non-equifeasible ensembles are created by combining the individual projections whose control or corrected control simulation has a better fit to the historical series in terms of basic and droughts statistics. In this work, the tool is presented, and the methodology implemented is described. It is also applied to a case study to illustrate how the tool works. The tool was previously tested in different typologies of water resources systems that cover different spatial scales (river basin, aquifer, mountain range, and country), obtaining satisfactory results. The local future scenarios can be propagated through appropriate hydrological models to study the impacts on other variables (e.g., aquifer recharge, chloride concentration in coastal aquifers, streamflow, snow cover area, and snow depth). The tool is also useful in quantifying the uncertainties of the future scenarios by combining them with stochastic weather generators.

1. Introduction

Adaptation to climate change (CC) is one of the greatest challenges that humans will face in the coming decades [1, 2]. This is related to greenhouse gas emissions, mainly anthropogenic carbon dioxide emissions [3]. CC is accepted by the scientific community and most of the society [4]. In the last decades, unprecedented changes (increment in temperature (T) of the atmosphere and oceans, changes in the spatiotemporal variability of precipitation (P) (including more extreme events), reduction of ice, snow, and groundwater levels, and rises in sea level) have been observed [2, 5]. The Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) [3] includes four future trajectories of anthropogenic carbon dioxide emissions for the twenty-first century, called representative concentration pathways (RCPs). From these trajectories, different climate institutes and organizations develop regional climate models (RCMs) whose data are coordinated and collected by the World Climate Research Program (WCRP) through the CORDEX project [6].

Despite the multiple uncertainties related to RCMs, they provide valuable information that could be used for further evaluation and policy making [7, 8]. Water systems, which are essential for life, are and will be one of the most affected by CC. In general, CC will modify the spatiotemporal variability of climatic variables. Among them, P and T are the main variables that govern the hydrological cycle [9, 10].

Decision makers should have enough information about how these systems will change in the future to design appropriate adaption strategies. For this reason, local scenarios of CC have to be generated from the climatic model simulations to analyse impacts on water resources (WR) systems.

Potential local future scenarios of P and T can be used to propagate impacts of CC on other hydrological variables in different WR systems, such as river basin [11, 12], aquifer [13, 14], and mountain range [15, 16]. RCMs include control and future simulations of P and T among other climatic variables. However, the RCM simulations cannot be used directly due to the high amount of uncertainties [17, 18]. The statistics of the control series can differ a lot from the statistics provided by historical or observed time series for the considered baseline period. Therefore, the time series provided by the RCMs must be corrected [19, 20] before using them for other applications, generating local series for the specific WR system.

Different statistical correction techniques can be used, such as correction of first- and second-order moments or regression approaches. They have different degrees of complexity and can have different accuracy. These correction techniques can be applied under two conceptual approaches: bias correction and delta change [21, 22]. RCMs and/or correction techniques can be evaluated comparing the control or corrected control simulation series with the historical information. Equifeasible and non-equifeasible (defined by giving more weight to the more reliable projections) ensembles of individual predictions can be employed to obtain more robust predictions [23]. The generation of potential local scenarios, through the correction of the RCMs simulations, involves several nontrivial tasks (selection of the RCMs, RCPs, and reference and future periods and/or level of global warming and application of the correction approaches and techniques) which should be addressed before the propagation of the impacts to hydrological variables. Therefore, tools and methodologies that automate the process are valuable for hydrological applications. Some tools are available to generate local future climate scenarios. Most of them are focused on the bias correction approach using different correction techniques mainly based on the statistical distribution of data (e.g., qmap [24] and CMhyd [25]). However, as far as we know, the application of different correction techniques under the delta change approach requires generating scenarios for specific future horizons and/or level of global warming that have not been included in a tool. On the other hand, these tools do not include the option of creating non-equifeasible ensembles of projections with different weights depending on the reliability of RCMs and correction techniques attending to some basic and/or drought statistics. In this work, these possibilities are incorporated.

We present a new tool to generate potential future CC scenarios for hydrology applications. The GROUNDS (GeneRation of pOtential fUture CC sceNarios for hyDrology applicationS) tool generates distributed future monthly series of P, T (maximum, minimum, and mean) and potential evapotranspiration (ETP) using as inputs historical series and RCM simulations. This tool integrates some methodological aspects studied in previous works [12, 26] (generation of individual projections and creation of ensembles of them) and new potentialities (identification of the period of the RCM to be used for a specific local warming and generation of ETP projections) into an intuitive command-line interface applicable to any case study. The main advantages of the tool are (1) its applicability to any case study and series of data (RCMs, RCPs, simulation periods, etc.), (2) it only requires historical data and RCM simulations to generate the potential future scenarios, (3) it allows using different approaches and statistical correction techniques, and (4) it generates individual local projections and ensembles, including equifeasible and non-equifeasible approaches considering the approximation to historical basic and drought statistics.

The GROUNDS tool allows different possibilities of generating future scenarios and can be used for different purposes depending on the interest of the users:

- Specific level of local warming vs. a specific future period: users can be interested in studying the availability of WR under different RCMs or know how the system will change according to a defined warming.
- Individual projections or ensembles of projections: based on multicriteria analyses [8], they allow to select the individual projections that provide the best approximation to some selected basic statistics (first, second, and/or third moment) or drought statistics (frequency, intensity, magnitude, and duration), which are assessed by applying the Run Theory to the Standard Precipitation Index (SPI) series for the studied variable. They also allow to define ensembles scenarios by combining the selected individual projections. The first option is indicated to quantify uncertainty by using several RCMs and the second to obtain more robust prediction of mean values.
- Different variables according to the preferences of the user: P, maximum, mean, and minimum T and ETP.

These future scenarios are also useful to assess impacts of CC on the availability of WR in a system by propagating the climate inputs to other variables with appropriate hydrological models.

2. Description of the Tool: Data, Methods, and Example Results

Figure 1 shows a flow chart of the methodology implemented in the tool GROUNDS to generate potential climate scenarios to assess impacts on WR systems. The tool, which includes four subroutines, was written entirely in Fortran 90 and can be executed under a command-line interface. The description of the inputs, parameters, subroutines, and outputs of the tool are shown in Figure 2.

The target is to use historical long enough monthly series of P and T (maximum, minimum, and mean) to generate future potential local climatic scenarios of them. It also includes a subroutine to calculate ETP. Note that monthly

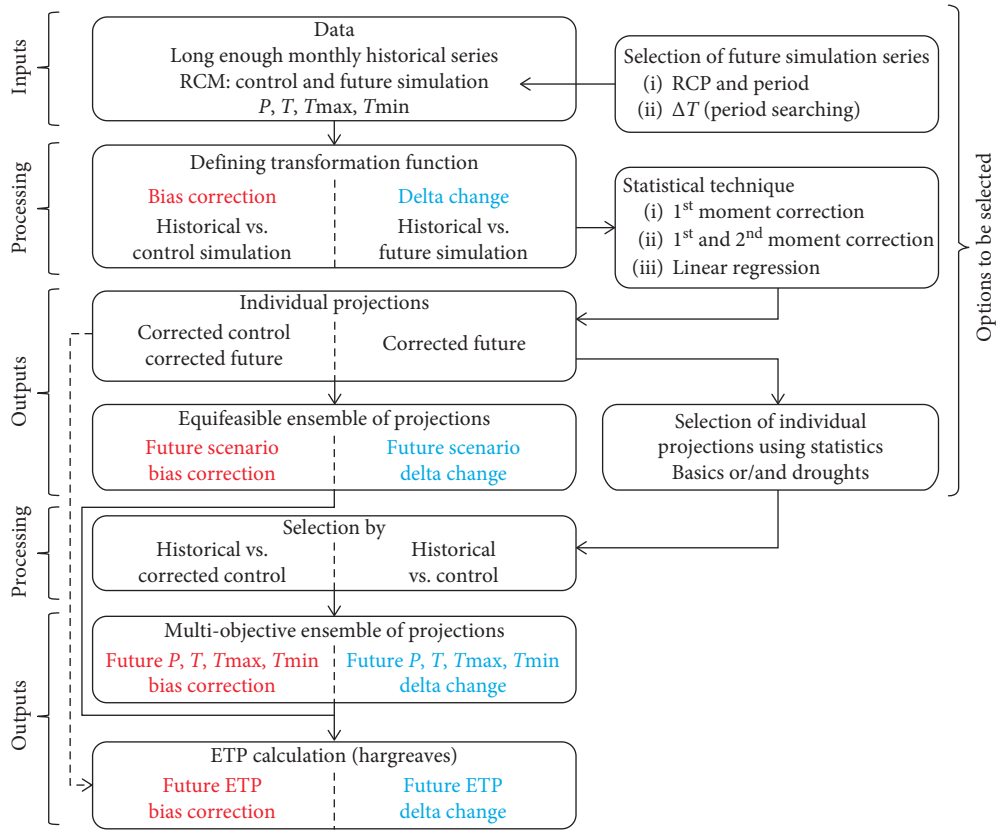


Figure 1: Flow chart of the assessment performed with the tool.

series should be complete (without gaps) and the quality of the data should be assessed by the user to ensure that they are representative of the climatology of the study area. It can be applied to point series that correspond to gauges measurements (in which gaps have to be previously completed) or to distributed fields, previously estimated on a grid with an appropriate method. Note that, in any case, those series have to cover the adopted historical reference period. The individual local series are generated using the subroutine *inpro.exe* that requires as inputs the monthly series of the climatic variables and the parameters L, N, and M which represent the length of the series (number of months), the numbers of points or stations considered in the case study, and the code (from 1 to the number of RCMs considered) that represents the RCM to be used, respectively. The historical series should be long enough in order to fulfil the assumption of stationarity of their statistics. If the series are not long enough, they might include drought or wet periods that modify the statistics of the series, making them not representative of the baseline period. A period of at least 30 years is recommended [12, 27]. With respect to the selection of the future projections, the tool allows two possibilities: (1) to fix a future horizon (specific period of time) or (2) to use an increment in mean T (e.g., local warming of 1°C) with respect to the value for the adopted historical reference period. The second procedure can be done by using the subroutine *search.exe*. It requires as input data the future simulation series of T for each RCM considered for the whole period available, the adopted longitude (e.g., 30 years) for the reference period, and the mean increment of T. This subroutine identifies a future period with a longitude equal to the historical (e.g., 30 years) in which the fixed local warming with respect to the historical is reached. The subroutine calculates a moving average considering an aggregation window equal to the considered length of the future simulation series and compares these values with the simulation series for the reference period. These increments of T represent different levels of local warming. The T and P series for the identified period will be used to perform the statistical correction of the series to generate the local scenarios that fulfil the adopted warming criteria.

In this work, we have applied it to a case study, used as an example to illustrate how the tool generates future series of climatic variables. The historical series employed correspond to a point (located in the province of Granada, south Spain) (longitude: −4.0877, latitude: 37.3789, and elevation: 1016 m a.s.l.) of the grid used by the Spain02 v04 project [28]. This project includes P and T interpolated for continental Spain. The CC information was obtained from RCMs from the CORDEX project [6]. We selected two RCMs (CCLM4-8-17 nested to the CNRM-CM5 and WRF331F nested to IPSL-CM5A-MR) and the most pessimist scenario (RCP8.5). Both project use the same grid for Spain (spatial resolution of 12.5 km).

Using the subroutine *search.exe*, we calculated a 30-year moving average that allows us to identify when a specific warming is reached. Figure 3 shows the increment of temperature with respect to the reference period

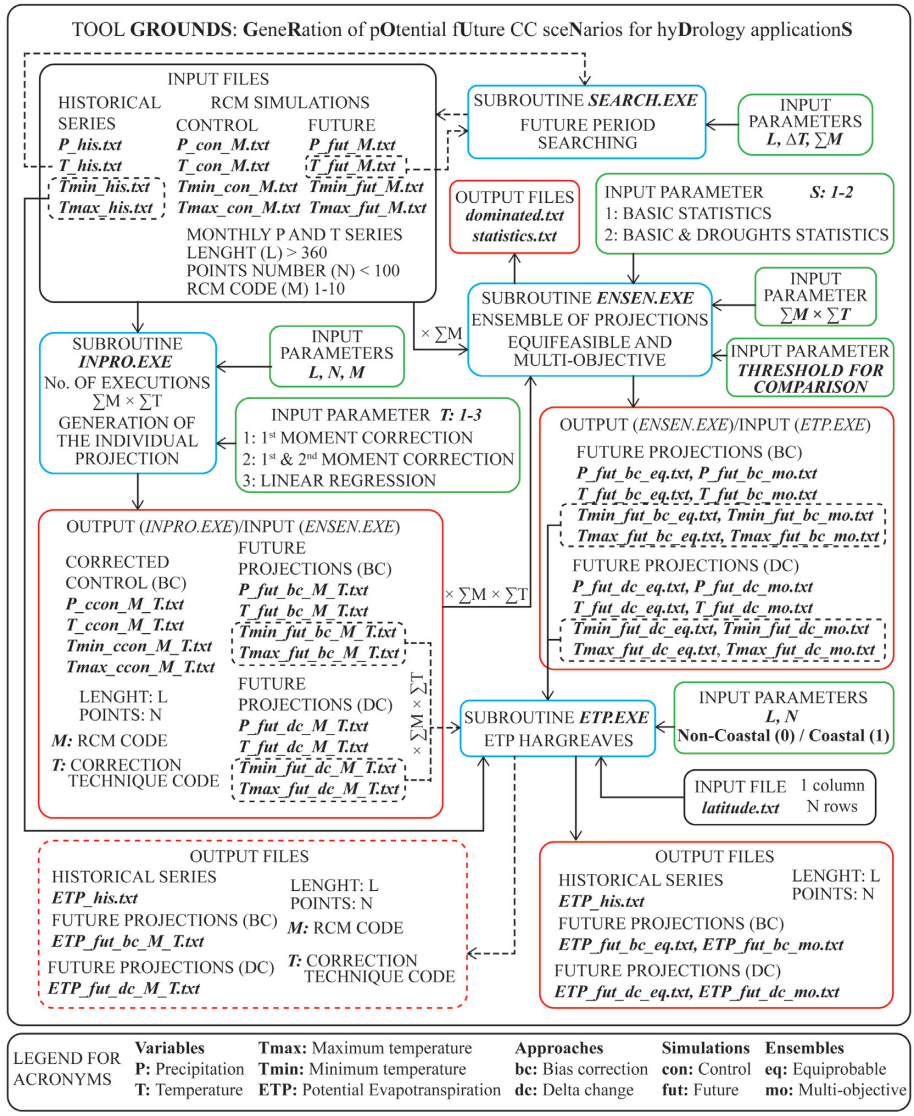


Figure 2: Flow chart of the GROUNDS tool subroutines.

(1971–2000) for the two RCMs considered and different future periods.

In this study, we selected the same future period for the generation of scenarios (2071–2100). Note that if a specific local warming is selected, different periods (periods in which the constraint is fulfilled) should be considered for each RCM. For example, in our case, a local warming of 4°C is reached for 2071–2100 in the case of CCLM4-8-17 nested to CNRM-CM5 and for 2064–2093 in the case of WRF331F nested to IPSL-CM5A-MR (see Figure 3).

The subroutine *inpro.exe* uses two conceptual approaches (bias correction and delta change) and three correction techniques, defined by a transformation function, to generate future potential scenarios (see Figure 1). The bias correction approach defines a transformation function to modify the control simulation series to another one whose statistics are more similar to the historical ones. The same function is applied to the future simulation series assuming that the bias between the historical data and the RCM simulations will be maintained invariant in the future [29, 30]. On the other hand, the delta change approach assumes that the RCMs provide good approximation of the relative changes between future and control simulations, and those changes are applied to the historical series to generate future projections [22, 31]. Note that delta change approach does not correct the RCM simulations; it only modifies the historical series to obtain the future series. From the point of view of the tool, the main difference between bias correction and delta change is that the first generates corrected control series for each combination of correction technique and RCM. The tool allows using three different correction techniques: first moment correction (code 1), first and second moment correction (code 2), and linear regression (code 3) (see Figure 2). These techniques, unlike quantile mapping techniques, do not assume a specific statistical distribution of data. The different correction techniques are applied at monthly scale. In the first moment correction, the transformation function is performed considering only the mean values, while the first and second moment correction technique considers mean and standard deviation. The

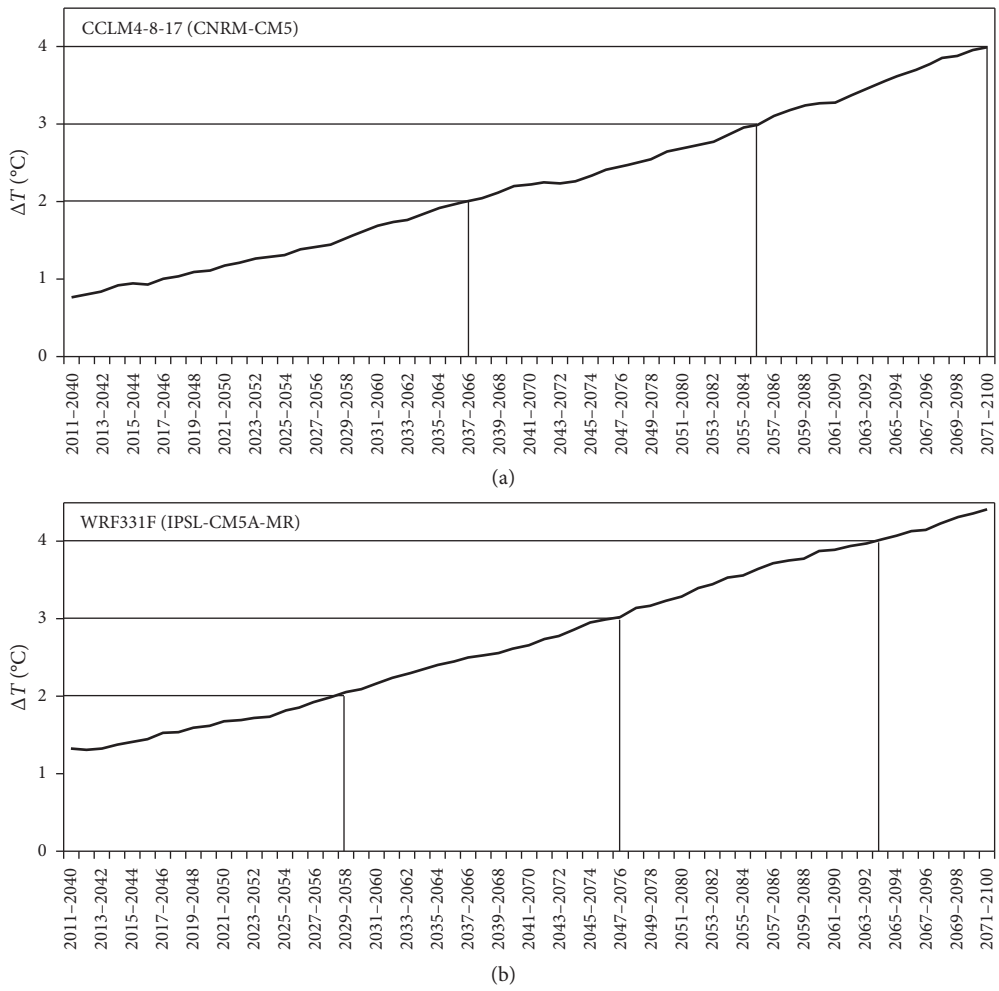


Figure 3: Local warming reached with respect to the reference period (1971–2000) for the (a) RCM CCLM4-8-17 nested to GCM CNRM-CM5 and (b) RCM WRF331F nested to GCM IPSL-CM5A-MR.

transformation function in the linear regression is elaborated through fitting a regression model between historical data and control simulation in the case of the bias correction approach and between the control simulation and future simulation in the case of the delta change approach. Note that the tool employs the best linear fit and provides the correlation coefficient (R^2) of it. The user should decide if its accuracy is enough. For the case study presented in this work, R^2 is above 0.85 for RCM CCLM4-8-17 nested to CNRM-CM5 and 0.67 for WRF331F nested to IPSL-CM5A-MR in the case of P. In the case of T, these values are 0.70 and 0.75, respectively. Figure 4 shows the historical data and the individual projections obtained for P and T for the considered case study. The average monthly results for the mean year have been represented in Figure 5.

The generated scenarios can be used as individual local projections or create ensembles of them that could produce more robust climate scenarios than those based on a single projection [23]. The subroutine *ensen.exe* allows to generate four ensembles of individual projections. Two of them are defined by combining, as equifeasible members, all the individual projections generated by applying delta change or bias correction approaches (see Figure 2). Two other options are defined by combining only the noneliminated projections in the multiobjective analysis for the two approaches (delta change and bias correction). The tool allows performance of the multiobjective analyses by considering only basic statistics (mean, standard deviation, and asymmetry coefficient) or including also drought statistics (duration, magnitude, and intensity). The drought statistics are calculated by applying the Run Theory to the corresponding SPI monthly series [12, 32]. The drought statistics series are calculated considering different thresholds of SPI values (it covers the whole range of negative values of SPI obtained) to define drought periods. In the calculation of the SPI, the probability of occurrence of P in the RCM simulations was obtained by using the parameters calibrated from the historical series to perform an appropriate comparison [12]. In this analysis, an RCM or a combination of RCM and correction technique is eliminated (dominated solution in multiobjective terminology), if another projection has a better fit to the historical data for all the statistics. The index employed to measure the goodness of fit is the mean of the quadratic difference between simulations and historical

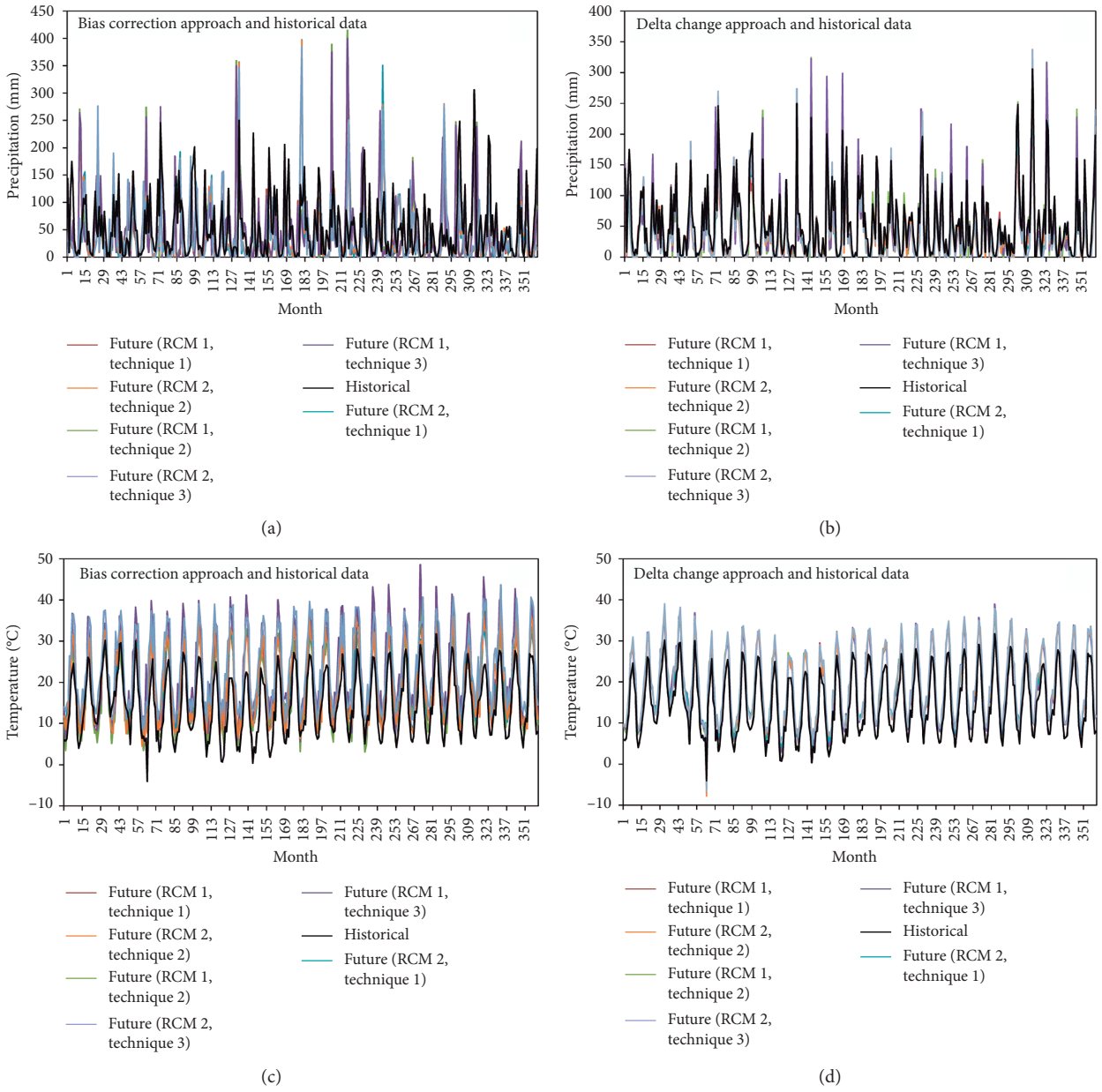


Figure 4: Historical data and individual projections of P using the (a) bias correction approach and (b) delta change approach and of T using the (c) bias correction approach and (d) delta change approach generated by using the GROUNDS tool.

series divided by the square of the mean historical value for each statistic. This index is applied to each considered statistic. The tool also includes a threshold (it can be modified by the user) that allows us to define when an approach is considered significantly better than others with respect to a specific statistic. This threshold, measured in percentage of relative difference, allows us to assume that certain differences in some statistics are not significant. For the considered example, this threshold has been considered 3%. The subroutine *ensen.exe* also generates the files *dominated.txt* and *statistics.txt*. The first indicates the combinations of RCMs and techniques (in the case of bias correction approach) and RCMs (in the case of delta change approach) that are dominated solutions. These solutions are eliminated to obtain the non-equifeasible ensembles. The second file includes the values of the indices used for the multiobjective analysis. For the case study, the combination of the second RCM (WRF331F nested to IPSL-CM5A-MR) and the first correction technique (first moment correction) was identified as dominated solution. The mean values of P and T obtained with the ensembles and their standard deviations for the case study are represented in Figure 6. In the case of delta change, the equifeasible and the multiobjective ensembles are equivalent because there is no dominated solution attending to the two RCMs considered. For the case study, an important reduction in P (mean of the four ensembles of 23.7%) and increase in T (mean of the four ensembles of 32.6%) are obtained. With respect to the

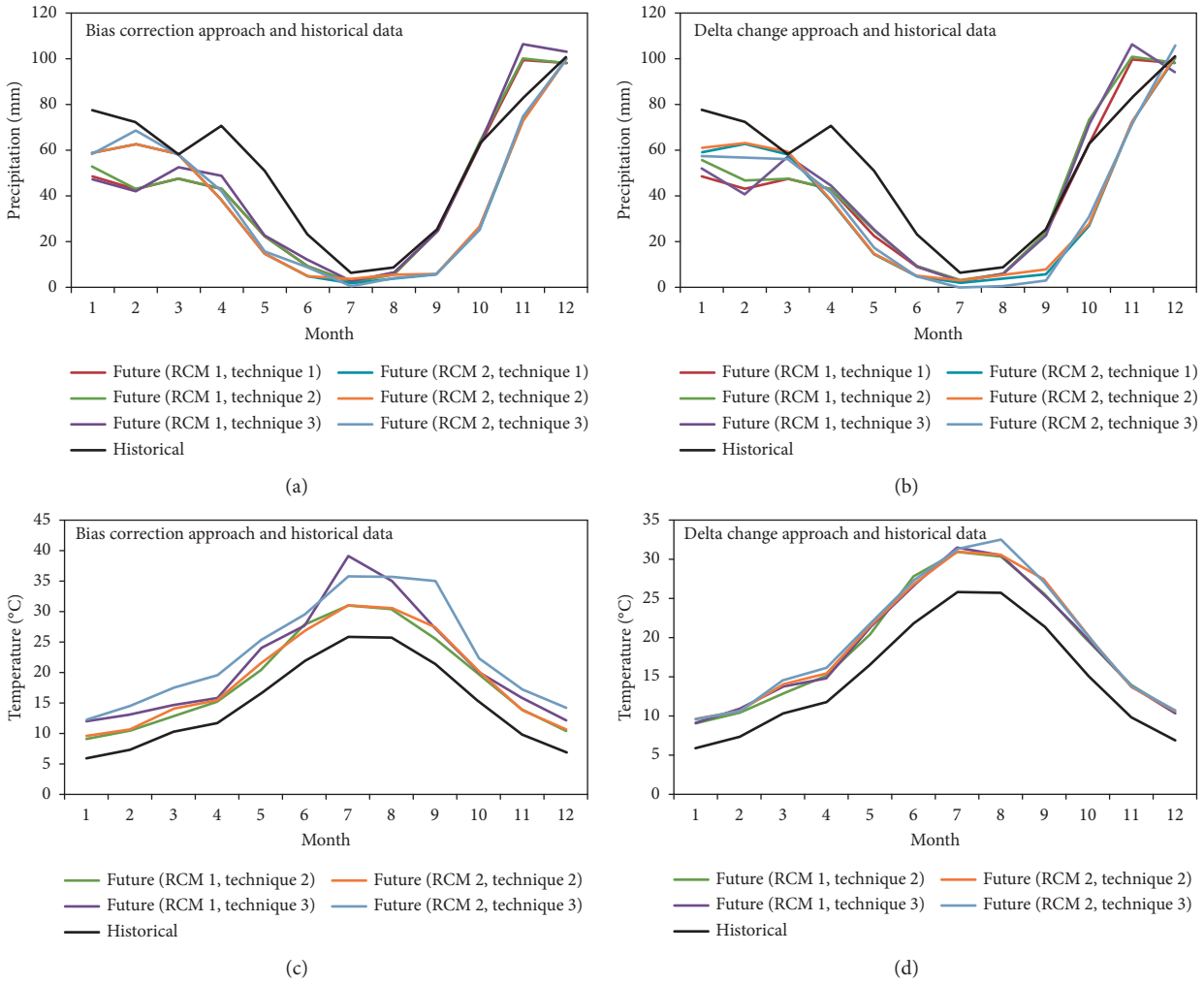


Figure 5: Mean year at monthly scale for the historical data and individual projections of P using the (a) bias correction approach and (b) delta change approach and of T using the (c) bias correction approach and (d) delta change approach generated by using the GROUNDS tool.

standard deviation, in the case of P, the ensembles show a lower value with the exception of the autumn where some ensembles predict an increment of the variability. In the case of the temperature, the bias correction approaches predict a lower variability and the delta change approaches a higher variability.

Finally, the tool also allows calculating historical and future ETP series by means of the subroutine *ETP.exe*. This subroutine employs the Hargreaves method [33, 34] and requires as inputs the historical and the future ensembles (or individual projections) of maximum and minimum T and the latitude of each location or station considered. The tool employs the modified method proposed by Samani [35]. This method uses tabulated values (depending on the latitude) to calculate solar radiation. It is also needed to indicate if the study area is located in a coastal or noncoastal zone to determine an empirical coefficient, which should be different for interior regions and for coastal regions [36]. The mean historical and future ETP for the case study have been represented in Figure 7. The mean increment of the four future ensembles of ETP with respect to the historical data is 31.8%.

3. Utility and Discussion

The GROUNDS tool allows correcting the future climatic simulations using the observed series to reduce the uncertainties of the RCMs. The tool requires complete monthly series (without gaps) and an adequate assessment of the quality in order to ensure that they represent the climatology of the case study. These corrected projections are plausible scenarios that intend to be representative of a period or a mean increment of T (local or global warming). The IPCC in some of its latest reports opts for using a mean increment of global warming to better reach users and decision makers (e.g., [37]). Some research works included this approach too [38, 39]. Multiple future series could be generated for the same climatic conditions. The stochastic weather generators (SWG) provide multiple synthetic time series of weather data for a location or multiple locations based on the statistical characteristics of the observed climate at these locations [40, 41]. A large amount of works employed SWG to assess the impact of CC [42, 43]. The tool presented in this work could be combined with a SWG to generate multiple

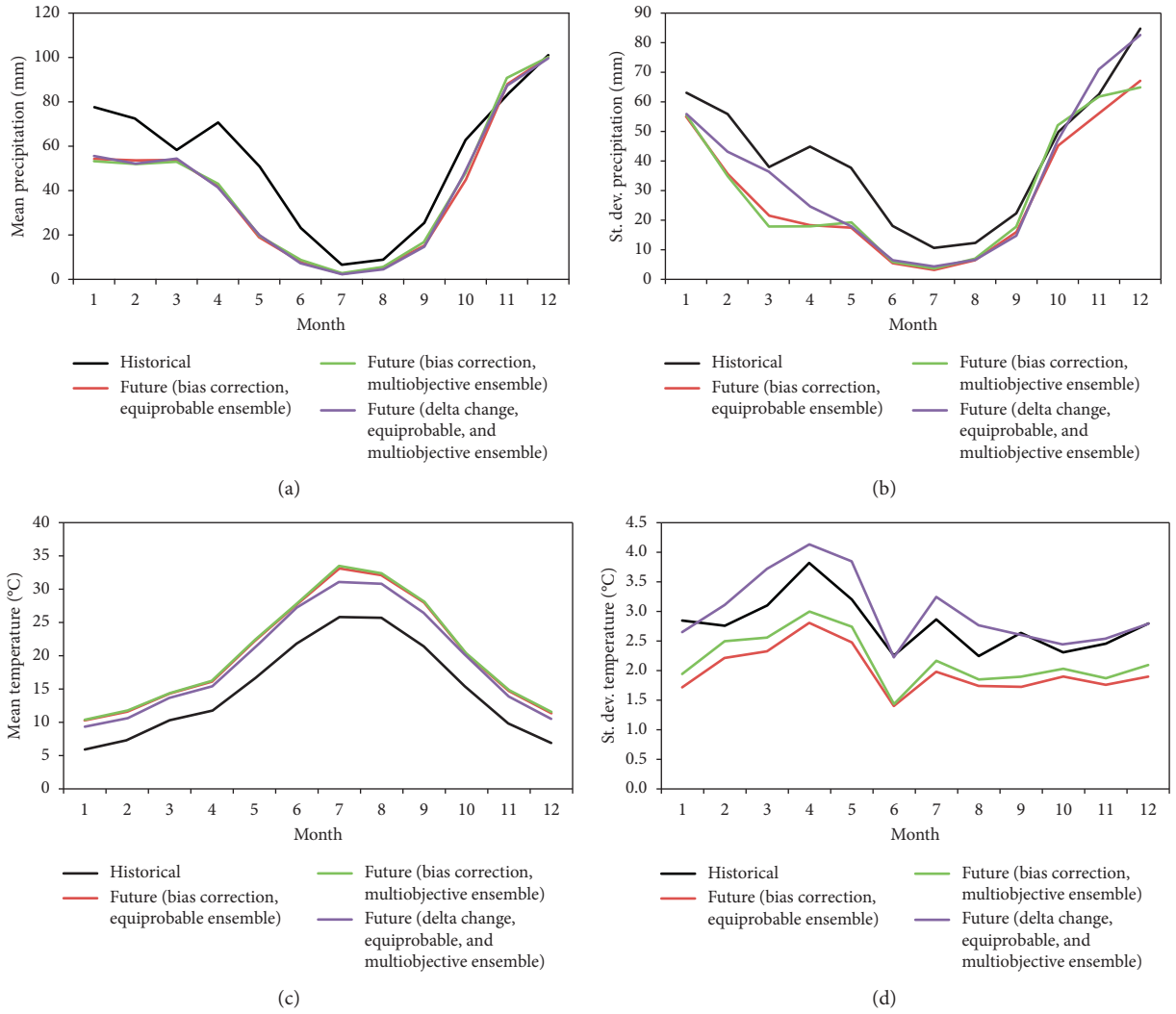


Figure 6: Mean year at monthly scale for the historical data and future ensembles of (a), (b) P and its standard deviation and (c), (d) T and its standard deviation generated by using the GROUNDS tool for the case study. Note that for the example considered in the case of the delta change approach, the multiobjective ensembles are equivalent to the equifeasible ensembles because there is not any dominated solution.

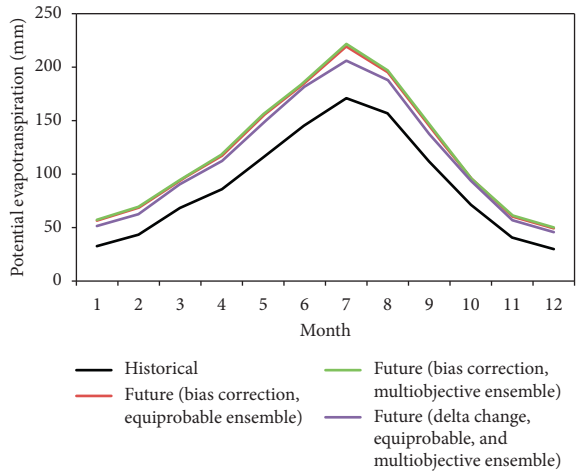


Figure 7: Mean year at monthly scale for the historical and future ensembles of ETP generated by using the GROUNDS tool for the case study. Note that for the example considered in the case of the delta change approach, the multiobjective ensembles are equivalent to the equifeasible ensembles because there is not any dominated solution.

series of potential future local scenarios of CC which could feed hydrological models quantifying the uncertainty of the projections through a Monte Carlo analysis [44, 45].

As pointed in the Introduction section, one of the main advantages of the GROUNDS tool is its applicability to different cases studies. In recent years, two subroutines of the tool (*inpro.exe* and *ensen.exe*) have been employed to generate potential scenarios for assessing CC impacts in different typologies of WR systems. It has been tested in cases covering different spatial scales: country [26], basin [12], aquifer [14, 46, 47], and mountain range [16, 48]. Furthermore, the local scenarios of CC generated by using the GROUNDS tool have been propagated to different variables in different case studies: recharge [16, 26], ground subsidence [47], and snow cover area [48]. The tool has been proven to be effective for hydrological applications and could contribute to the analyses to support decision-making processes by experts or public participation [49, 50].

All the cited applications correspond to case studies located in Spain, where important reduction in P and increment in T are expected due to CC. The tool also allows quantifying the change in variability of the climatic variables at seasonal and monthly scale. This variability in general will be higher due to CC [51, 52]. For the case showed in this study, the tool predicts a higher variability of P in autumn. In the case of T, the delta change ensembles show higher variability than the historical data (see Figure 6).

We have developed a general tool that uses several correction techniques (first moment correction, first and second moment correction, and linear regression). It does not include correction techniques based on the cumulative distribution function, which have been employed by other authors with good results [53, 54]. However, the differences between these techniques and more simple techniques, as first and second moment corrections, are not significant for the hydrological applications where we have tested them [12]. On the other hand, the tool employs and generates monthly series. This could be an inconvenience for some hydrological applications that use daily data [55–57]. The possibility of generating daily data was tested, but if the second moment correction is applied, it might generate negative values of P. If these values are moved to zero, the statistics of the series can suffer modifications and the information from RCMs can be altered [31]. However, the monthly series can be transformed to daily series using the daily pattern of distribution of the monthly historical values to distribute the generated future monthly values [48].

On the other hand, the local future scenarios can be propagated through appropriate hydrological models to study the impacts on other variables (e.g., aquifer recharge or discharge [26], chloride concentration in coastal aquifers [46], streamflow, snow cover area, and snow depth [48]). They can be even propagated with other models to assess nonhydrological physical variables (e.g., subsidence [47]).

4. Conclusions

In this work, the new tool GROUNDS, which generates potential future local scenarios of CC, has been presented. The tool is applicable to any case study and requires historical and RCM simulation information to generate individual local projections. The generated future scenarios can be used as individual local projections or create ensembles of them that could produce more robust climate scenarios taking into account basic and droughts statistics. The tool also allows determination of the future simulation period from a specific increment in mean T (local warming scenario) and propagation the impacts on T to ETP.

For the case study (located in south Spain) used to illustrate the functionalities of the tool, an important reduction in P (in mean 23.7%) and an increment in T and ETP (in mean 32.6% and 31.8%, respectively) have been obtained.

Two subroutines of the tool have proven to be effective for different scales of application and spatiotemporal resolutions in previous works (see the previous section). The tool can be also combined with appropriate hydrological models to propagate impacts of CC to other hydrological variables or SWG to assess the uncertainty of the generated future scenarios in these models.

Data Availability

The historical data and the RCM simulations used to support the findings of this study come from previously reported studies and datasets, which have been cited. The software developed in this study may be available for research purposes upon request from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was partially supported by the research projects SIGLO-AN (RTI2018-101397-B-I00) from the Spanish Ministry of Science, Innovation and Universities (Programa Estatal de I+D+I Orientada a los Retos de la Sociedad) and the GeoE.171.008-TACTIC Project from the GeoERA Organization funded by the European Union's Horizon 2020 Research and Innovation Program.

References

[1] A. K. Misra, "Climate change and challenges of water and food security," *International Journal of Sustainable Built Environment*, vol. 3, no. 1, pp. 153–165, 2014.

[2] W. Cramer, J. Guiot, M. Fader et al., "Climate change and interconnected risks to sustainable development in the Mediterranean," *Nature Climate Change*, vol. 8, no. 11, pp. 972–980, 2018.

[3] IPCC, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, R. K. Pachauri and L. A. Meyer, Eds., p. 151, IPCC, Geneva, Switzerland, 2014.

[4] J. S. Dryzek, R. B. Norgaard, and D. Schlosberg, "Climate change and society: approaches and responses," in *The Oxford Handbook of Climate Change and Society*, S. Dryzek John and

S. David, Eds., pp. 490–503, Oxford University Press, Oxford, UK, 2011.

[5] IPCC, *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, C. B. Field, V. Barros, T. F. Stocker et al., Eds., p. 582, Cambridge University Press, Cambridge, UK, 2012.

[6] CORDEX Project, "The coordinated regional climate downscaling experiment CORDEX. Program sponsored by World climate research program (WCRP)," 2013, http://www.cordex.org/.

[7] P. Jacques, "Downscaling climate models and environmental policy: from global to regional politics," *Journal of Environmental Planning and Management*, vol. 49, no. 2, pp. 301–307, 2006.

[8] I. M. Picketts, S. J. Déry, and J. A. Curry, "Incorporating climate change adaptation into local plans," *Journal of Environmental Planning and Management*, vol. 57, no. 7, pp. 984–1002, 2014.

[9] I. M. Held and B. J. Soden, "Robust responses of the hydrological cycle to global warming," *Journal of Climate*, vol. 19, no. 21, pp. 5686–5699, 2006.

[10] P. Wu, N. Christidis, and P. Stott, "Anthropogenic impact on Earth's hydrological cycle," *Nature Climate Change*, vol. 3, no. 9, pp. 807–810, 2013.

[11] M. Safeeq and A. Fares, "Hydrologic response of a Hawaiian watershed to future climate change scenarios," *Hydrological Processes*, vol. 26, no. 18, pp. 2745–2764, 2012.

[12] A.-J. Collados-Lara, D. Pulido-Velazquez, and E. Pardo-Igúzquiza, "An integrated statistical method to generate potential future climate scenarios to analyse droughts," *Water*, vol. 10, no. 9, p. 1224, 2018.

[13] T. Y. Stigter, J. P. Nunes, B. Pisani et al., "Comparative assessment of climate change and its impacts on three coastal aquifers in the Mediterranean," *Regional Environmental Change*, vol. 14, no. S1, pp. 41–56, 2014.

[14] L. Baena-Ruiz, D. Pulido-Velazquez, A.-J. Collados-Lara et al., "Summarizing the impacts of future potential global change scenarios on seawater intrusion at the aquifer scale," *Environmental Earth Sciences*, vol. 79, no. 5, p. 99, 2020.

[15] N. Salzmann, J. Nötzli, C. Hauck, S. Gruber, M. Hoelzle, and W. Haeberli, "Ground surface temperature scenarios in complex high-mountain topography based on regional climate model results," *Journal of Geophysical Research*, vol. 112, no. F2, 2007.

[16] E. Pardo-Igúzquiza, A. J. Collados-Lara, and D. Pulido-Velazquez, "Potential future impact of climate change on recharge in the Sierra de las Nieves (southern Spain) high-relief karst aquifer using regional climate models and statistical corrections," *Environmental Earth Sciences*, vol. 78, no. 20, p. 598, 2019.

[17] M. Déqué, D. P. Rowell, D. Lüthi et al., "An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections," *Climatic Change*, vol. 81, no. S1, pp. 53–70, 2007.

[18] C. Dobler, S. Hagemann, R. L. Wilby, and J. Stötter, "Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed," *Hydrology and Earth System Sciences*, vol. 16, no. 11, pp. 4343–4360, 2012.

[19] J. Chen, F. P. Brissette, D. Chaumont, and M. Braun, "Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America," *Water Resources Research*, vol. 49, no. 7, pp. 4187–4205, 2013.

[20] M. Luo, T. Liu, F. Meng et al., "Comparing bias correction methods used in downscaling precipitation and temperature from regional climate models: a case study from the kaidu river basin in western China," *Water*, vol. 10, no. 8, p. 1046, 2018.

[21] S. Watanabe, S. Kanae, S. Seto, P. J.-F. Yeh, Y. Hirabayashi, and T. Oki, "Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models," *Journal of Geophysical Research: Atmospheres*, vol. 117, no. D23, 2012.

[22] J. Räisänen and O. Räty, "Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations," *Climate Dynamics*, vol. 41, no. 5-6, pp. 1553–1568, 2013.

[23] AEMET (Spanish Meteorologial Agency), *Generación de Escenarios Regionalizados de Cambio Climático Para España*, Agencia Estatal de Meteorología (Ministerio de Medio Ambiente y Medio Rural y Marino), Madrid, Spain, 2009.

[24] L. Gudmundsson, "Statistical transformations for post-processing climate model output," 2016, https://cran.r-project.org/web/packages/qmap/qmap.pdf.

[25] H. Rathjens, B. Bieger, S. Srinivasan, I. Chaubey, and J. G. Arnold, "CMhyd user manual," 2016, http://swat.tamu.edu/software/cmhyd/.

[26] D. Pulido-Velazquez, A.-J. Collados-Lara, and F. J. Alcalá, "Assessing impacts of future potential climate change scenarios on aquifer recharge in continental Spain," *Journal of Hydrology*, vol. 567, pp. 803–819, 2018.

[27] D. Pulido-Velazquez, L. Garrote, J. Andreu, F.-J. Martin-Carrasco, and A. Iglesias, "A methodology to diagnose the effect of climate change and to identify adaptive strategies to reduce its impacts in conjunctive-use systems at basin scale," *Journal of Hydrology*, vol. 405, no. 1-2, pp. 110–122, 2011.

[28] S. Herrera, J. Fernández, and J. M. Gutiérrez, "Update of the Spain02 gridded observational dataset for euro-CORDEX evaluation: assessing the effect of the interpolation methodology," *International Journal of Climatology*, vol. 36, no. 2, pp. 900–908, 2016.

[29] C. Piani, J. O. Haerter, and E. Coppola, "Statistical bias correction for daily precipitation in regional climate models over Europe," *Theoretical and Applied Climatology*, vol. 99, no. 1-2, pp. 187–192, 2010.

[30] L. P. Seaby, J. C. Refsgaard, T. O. Sonnenborg, and A. L. Højberg, "Spatial uncertainty in bias corrected climate change projections and hydrogeological impacts," *Hydrological Processes*, vol. 29, no. 20, pp. 4514–4532, 2015.

[31] D. Pulido-Velazquez, J. L. García-Aróstegui, J.-L. Molina, and M. Pulido-Velázquez, "Assessment of future groundwater recharge in semi-arid regions under climate change scenarios (Serral-Salinas aquifer, SE Spain). Could increased rainfall variability increase the recharge rate?" *Hydrological Processes*, vol. 29, no. 6, pp. 828–844, 2015.

[32] S. Khan, H. F. Gabriel, and T. Rana, "Standard precipitation index to track drought and assess impact of rainfall on watertables in irrigation areas," *Irrigation and Drainage Systems*, vol. 22, no. 2, pp. 159–177, 2008.

[33] G. H. Hargreaves and Z. A. Samani, "Reference crop evapotranspiration from temperature," *Applied Engineering in Agriculture*, vol. 1, no. 2, pp. 96–99, 1985.

[34] R. G. Allen, L. S. Pereira, and D. Raes y Smith, "Crop evapotranspiration-guidelines for computing crop water requirements-FAO irrigation and drainage paper 56," 1998, http://www.fao.org/docrep/X0490E/X0490E00.htm#Contents.

[35] Z. Samani, "Estimating solar radiation and evapotranspiration using minimum climatological data," *Journal of Irrigation and Drainage Engineering*, vol. 126, no. 4, pp. 265–267, 2000.

[36] G. H. Hargreaves, *Simplified Coefficients for Estimating Monthly Solar Radiation in North America and Europe*, Utah State University, Logan, UT, USA, 1994.

[37] IPCC, "Summary for policymakers. In: global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways," in *The Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*, P. Zhai, H.-O. Pörtner, D. Roberts et al., Eds., p. 32, World Meteorological Organization, Geneva, Switzerland, 2018.

[38] A. D. King, D. J. Karoly, and B. J. Henley, "Australian climate extremes at 1.5°C and 2°C of global warming," *Nature Climate Change*, vol. 7, no. 6, pp. 412–416, 2017.

[39] J. Wu, Z. Han, Y. Xu, B. Zhou, and X. Gao, "Changes in extreme climate events in China under 1.5°C–4°C global warming targets: projections using an ensemble of regional climate model simulations," *Journal of Geophysical Research: Atmospheres*, vol. 125, Article ID e2019JD031057, 2020.

[40] D. S. Wilks and R. L. Wilby, "The weather generation game: a review of stochastic weather models," *Progress in Physical Geography: Earth and Environment*, vol. 23, no. 3, pp. 329–357, 1999.

[41] N. Peleg, S. Fatichi, A. Paschalis, P. Molnar, and P. Burlando, "An advanced stochastic weather generator for simulating 2-D high-resolution climate variables," *Journal of Advances in Modeling Earth Systems*, vol. 9, no. 3, pp. 1595–1627, 2017.

[42] M. A. Semenov and E. M. Barrow, "Use of a stochastic weather generator in the development of climate change scenarios," *Climatic Change*, vol. 35, no. 4, pp. 397–414, 1997.

[43] J. Chen, H. Chen, and S. Guo, "Multi-site precipitation downscaling using a stochastic weather generator," *Climate Dynamics*, vol. 50, no. 5-6, pp. 1975–1992, 2018.

[44] R.-S. Blasone, H. Madsen, and D. Rosbjerg, "Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling," *Journal of Hydrology*, vol. 353, no. 1-2, pp. 18–32, 2008.

[45] H. K. McMillan, I. K. Westerberg, and T. Krueger, "Hydrological data uncertainty and its implications," *WIREs Water*, vol. 5, no. 6, Article ID e1319, 2018.

[46] D. Pulido-Velazquez, A. Renau-Pruñonosa, C. Llopis-Albert et al., "Integrated assessment of future potential global change scenarios and their hydrological impacts in coastal aquifers-a new tool to analyse management alternatives in the Plana Oropesa-Torreblanca aquifer," *Hydrology and Earth System Sciences*, vol. 22, no. 5, pp. 3053–3074, 2018.

[47] A.-J. Collados-Lara, D. Pulido-Velazquez, R. M. Mateos, and P. Ezquerro, "Potential impacts of future climate change scenarios on ground subsidence," *Water*, vol. 12, no. 1, p. 219, 2020.

[48] A.-J. Collados-Lara, E. Pardo-Igúzquiza, and D. Pulido-Velazquez, "A distributed cellular automata model to simulate potential future impacts of climate change on snow cover area," *Advances in Water Resources*, vol. 124, pp. 106–119, 2019.

[49] H. Macian-Sorribes and M. Pulido-Velazquez, "Integrating historical operating decisions and expert criteria into a DSS for the management of a multireservoir system," *Journal of Water Resources Planning and Management*, vol. 143, no. 1, 2017.

[50] C. Llopis-Albert, J. M. Merigó, H. Liao, Y. Xu, J. Grima-Olmedo, and C. Grima-Olmedo, "Water policies and conflict resolution of public participation decision-making processes using prioritized ordered weighted averaging (OWA) operators," *Water Resources Management*, vol. 32, no. 2, pp. 497–510, 2018.

[51] A. G. Pendergrass, R. Knutti, F. Lehner, C. Deser, and B. M. Sanderson, "Precipitation variability increases in a warmer climate," *Scientific Reports*, vol. 7, no. 1, p. 17966, 2017.

[52] F. J. M. M. Nijsse, P. M. Cox, C. Huntingford, and M. S. Williamson, "Decadal global temperature variability increases strongly with climate sensitivity," *Nature Climate Change*, vol. 9, no. 8, pp. 598–601, 2019.

[53] L. Gudmundsson, J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, "Technical Note: downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods," *Hydrology and Earth System Sciences*, vol. 16, no. 9, pp. 3383–3390, 2012.

[54] L. Sangelantoni, A. Russo, and F. Gennaretti, "Impact of bias correction and downscaling through quantile mapping on simulated climate change signal: a case study over Central Italy," *Theoretical and Applied Climatology*, vol. 135, no. 1-2, pp. 725–740, 2019.

[55] Y. Huang, A. Bárdossy, and K. Zhang, "Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data," *Hydrology and Earth System Sciences*, vol. 23, no. 6, pp. 2647–2663, 2019.

[56] J. Senent-Aparicio, S. Liu, J. Pérez-Sánchez, A. López-Ballesteros, and P. Jimeno-Sáez, "Assessing impacts of climate variability and reforestation activities on water resources in the headwaters of the segura river basin (SE Spain)," *Sustainability*, vol. 10, no. 9, p. 3277, 2018.

[57] H. Moutahir, M. Fernández-Mejuto, J. M. Andreu et al., "Observed and projected changes on aquifer recharge in a Mediterranean semi-arid area, SE Spain," *Environmental Earth Sciences*, vol. 78, p. 671, 2019.

Hindawi
Scientific Programming
Volume 2020, Article ID 7563928, 12 pages
https://doi.org/10.1155/2020/7563928

Research Article

SAT: A Software for Assessing the Risk of Desertification in Spain

Jaime Martínez-Valderrama,[1] **Javier Ibáñez**,[2] **Francisco J. Alcalá**,[3,4] **and Silvio Martínez**[5]

[1]*Instituto Multidisciplinar para el Estudio del Medio "Ramón Margalef", Universidad de Alicante, Carretera de San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Alicante, Spain*
[2]*Departamento de Economía Agraria, Estadística y Gestión de Empresas, Universidad Politécnica de Madrid, 28040 Madrid, Spain*
[3]*Instituto Geológico y Minero de España (IGME), 28003 Madrid, Spain*
[4]*Instituto de Ciencias Químicas Aplicadas, Facultad de Ingeniería, Universidad Autónoma de Chile, 7500138 Santiago, Chile*
[5]*Centro de Ciencias Humanas y Sociales, Consejo Superior de Investigaciones Científicas, 28037 Madrid, Spain*

Correspondence should be addressed to Jaime Martínez-Valderrama; jaime.mv@ua.es

Received 12 December 2019; Revised 8 May 2020; Accepted 25 May 2020; Published 29 June 2020

Academic Editor: Autilia Vitiello

Desertification is a major global environmental issue exacerbated by climate change. Strategies to combat desertification include prevention which seeks to reverse the process before the system reaches the stable desertified state. One of these initiatives is to implement early warning tools. This paper presents SAT (the Spanish acronym for Early Warning System), a decision support system (DSS), for assessing the risk of desertification in Spain, where 20% of the land has already been desertified and 1% is in active degradation. SAT relies on three versions of a Generic Desertification Model (GDM) that integrates economics and ecology under the predator-prey paradigm. The models have been programmed using Vensim, a type of software used to build and simulate System Dynamics (SD) models. Through Visual Basic programming, these models are operated from the Excel environment. In addition to the basic simulation exercises, specially designed tools have been coupled to assess the risk of desertification and determine the ranking of the most influential factors of the process. The users targeted by SAT are government land-use planners as well as desertification experts. SAT tool is implemented for five case studies, each one of them representing a desertification syndrome identified in Spain. Given the general nature of the tool and the fact that all United Nations Convention to Combat Desertification (UNCCD) signatory countries are committed to developing their National Plans to Combat Desertification (NPCD), SAT could be exported to regions threatened by desertification and expanded to cover more case studies.

1. Introduction

Desertification is defined as land degradation (i.e., reduction or loss of the biological or economic productivity of the land) in drylands (i.e., arid, semiarid, and dry subhumid areas) resulting from various factors, "including climatic variations and human activities" [1]. The magnitude of desertification, which was the first major environmental issue to be recognized as occurring on a global scale [2], was significant enough to set the United Nations Convention to Combat Desertification (UNCCD), a degree of importance only achieved by the Conventions on Biodiversity and Climate Change.

Drylands occupy 47% of the terrestrial surface [3] and are home to over 38% of the world human population [4]. Furthermore, 90% of the human settlements in drylands are located in developing countries [5]. It is valued that severe ecosystem degradation is present in 10–20% of drylands, and its consequences are estimated to affect around 250 million people in the developing world [6]. Progression of desertification worldwide is expected, as deduced from the increased rainfall variability, frequency of droughts, and persistent dry conditions foreseen by the future climate change scenarios [7].

Desertification has been recognized as one of the biggest problems facing the European Mediterranean countries [8, 9]. UNCCD reserves Annex IV to cope with the special

features of the region. Two differing mechanisms can trigger desertification, both associated with the decoupling of natural resource provision and use [10]. The most common has to do with the overexploitation of natural resources, while the second, limited to the European case, is related to rural abandonment [11]. These opposing forces operate simultaneously and allow us to see that the issue of desertification requires a very fine adjustment in the intensity of land use. The challenge is further complicated by climate change. Climate projections predict a 5–30% drop in precipitation in large portions of the drylands in the Mediterranean Basin [12, 13] and an increase in aridity [14].

Strategies to combat desertification include mitigation and prevention. The former are applied when desertification has already deteriorated the territory. The main flagship of these actions is revegetation programmes [15]. They seek to control soil erosion, carbon sequestration, and even the reversal of water imbalances, emphasizing that increasing forest cover raises water yield [16]. The second group of initiatives includes sustainable land management practices [17] or the development of early warning indicators. This last line of action is that addressed in this work, specifically, through the development of computer tools to assess desertification, as the United Nations promote [1].

Early warning indicators seek to detect signals that warn the system is heading towards degradation before it reaches the stable state of "desertification" [18]. Many studies rely on studying the statistical properties of time series in order to detect the proximity of a tipping point or a catastrophic bifurcation point [19, 20]. Different statistical analyses, such as autocorrelation, have emerged in different fields and have been applied to different types of transitions [19]. In addition to studying trends over time, changes in spatial characteristics of the system can provide early warnings of approaching transitions as well [21]. In essence, these indicators seek to detect warn signals that a transition from some initial desired state to a final degraded state is prone to happen. However, our indicators (described in Section 2.2) are of a completely different nature because they are conceived for socioecological systems that are currently undergoing transitions, whose initial states could already be undesired in some senses, and where the question is whether the final equilibrium states will exceed some critical thresholds or not.

Spain is one of the countries most affected by desertification in the Mediterranean region. Drylands occupy 73% of the country. Overall, 20% of the territory is degraded and an additional 1% is actively degrading [22, 23]. Following UNCCD recommendations, the Ministry of Agriculture, Fisheries and Food of Spain (MAFF) implemented the Spanish NPCD in 2008 [24], relying on different methodological approaches for the monitoring and evaluation of desertification.

In order to improve its NPCD, MAFF decided to create an early warning system based on dynamic simulation models. The result is SAT, a computational tool to assess desertification risk in five hot spots. SAT, the Spanish acronym for Early Warning System, is a DSS implemented in Excel that handles SD models performed with Vensim (Vensim® DSS Version 5.8b) [25]. SAT yields three kinds of outputs: (i) temporal trends of all the variables contained in models, (ii) desertification risk assessment as the percentage of 1,000 simulations in which the values of key variables are above (or below) certain thresholds, and (iii) a ranking of desertification factors according to their impact on key variables.

This paper is structured as follows: Section 2 provides an outline of the simulation models that serve as the basis for the DSS and gives details about the analysis methods implemented to assess desertification. SAT tool is presented in Section 3. Finally, discussion and main conclusions drawn from the study are summarized in Section 4.

2. Materials and Methods

2.1. A Generic Desertification Model. SD modelling [26, 27] is a methodology that embodies system thinking through the implementation of differential equation systems. SD is interdisciplinary and is grounded in the theory of nonlinear dynamics and feedback control developed in mathematics, physics, and engineering. An SD model consists in a set of ordinary differential equations that makes a stock-and-flow representation of the studied system. SD models are conceived as a structure made up of causal feedback loops including nonlinear relationships and delays.

Agricultural systems and their environmental effects make up complex systems. Their study encompasses socioeconomic variables (demand of raw materials, employment, and land-use allocation), geochemical and ecological processes (aquifer exploitation, soil erosion, and biomass regeneration), and agricultural production techniques (tillage, supplementary feed, and types of crop). To understand the interaction of all these elements, a systemic approach capable of explicitly managing the temporal dimension, sustainability conditions, uncertainty, and externalities is mandatory [28].

The basis of SAT is a GDM [29] that represents the use of resources (soil, water, and pasture) by economic agents (shepherds and farmers) under the predator-prey paradigm [30, 31]. The consideration of human-nature interaction on the basis of this pattern of biological interaction has been successfully adapted to study grazing systems [32, 33]. GDM makes use of this idea and extends it to other land uses. The model highlights the dealings between environmental and socioeconomic variables, clarifies the processes and drivers behind land use and desertification, and copes with the critical gaps of tracking the origins of desertification [34].

GDM is lumped spatially [35] as it does not yield georeferenced results, and its temporal basis is the year. The analysis is focused on desertification syndromes. Their outputs are not designed to fill pixels' data but to spot threatens posed by some land uses. The core of the GDM, outlined in Figure 1, can be described succinctly as follows: the level of resources depends on their renewal and consumption rates. The former relies on the current stock of natural resources, the physical framework (climate and soil type), and a limiting factor, the dynamics of which can be affected by the stock of resources. In fact, soil erosion, soil

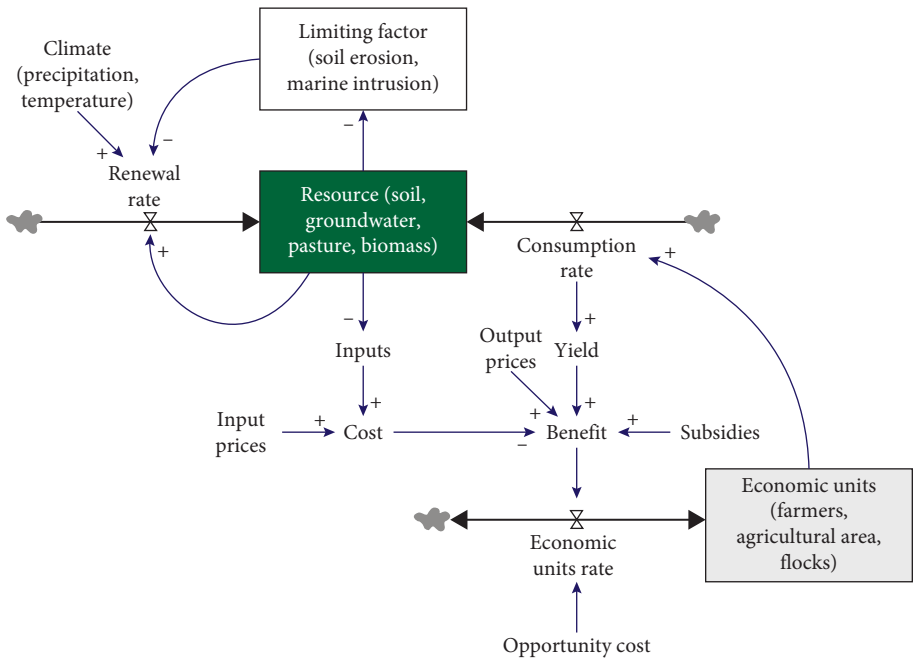


Figure 1: General overview of GDM.

salinization, saltwater intrusion, and others have an inhibitory effect on resource regeneration that can be irreversible once critical thresholds are exceeded [36].

Consumption is linked to the number of economic units in the area. They are understood as the number of farmers who make decisions on the size of herds or areas under cultivation. Decisions depend on their benefits compared to the alternative profit, i.e., the opportunity cost of farmers [37, 38]. When, on average, the ratio is above the unit, economic agents intensify their activity by increasing the number of livestock or the agricultural area; the opposite is also true.

Profit increases with income, which depends on yield, production prices, and subsidies. The costs, which penalize the benefit, respond to the use of additional inputs needed for production (supplementary feed, fertilizers, pesticides, fuel, etc.) and their price. Note that the amount of inputs required is conditioned by the stock of natural resources. For example, water table plummeting increases groundwater pumping costs, while the loss of soil and/or biomass drives up the costs of supplementary livestock feed.

Therefore, the deterioration of natural resources has a negative impact on economic activity that should slow down the overexploitation of resources. However, the signs of scarcity fade in an economic context in which certain circumstances (favourable prices, subsidies, technological improvements, etc.) generate the certainty—through indicators based on the profitability of the activity—that those good decisions are being made. For example, if the demand of some crop is very high, its market price can be high enough for revenues to offset or even exceed the rising of pumping costs.

The main output of an SD model is the time trajectories produced from numerical scenarios that capture the initial conditions of the system, i.e., the state in which the model is started, the values of the form parameters, and the hypotheses established by the user. The simulation period for GDM is set at 200 years, an appropriate time horizon when studying desertification. Indeed, environmental degradation only emerges when we observe the evolution of "slow variables" [6].

Time trends, despite their quantitative accuracy, should not be considered as forecasts. In this sense, it is important to take the nature of the laws underlying the model equations into account. Since they belong to the socio-ecological sphere, i.e., they are not strictly based on physical laws, it is more rigorous to speak of exploration than of prediction [39]. Hence, GDM is intended to be a "means of exploration" [40] for a better understanding of how systems may behave. In this context, the importance of simulation relies in the relative results yielded under different scenarios.

2.2. Assessing Desertification Risk. Ultimately, GDM is a state and transition model [41] that uses box-and-arrow diagrams accompanied by data-driven narratives to describe the states of key variables, i.e., the stock of resources and economic units. These models address both reversible and irreversible changes, the details of which are fundamental to understanding desertification and communicating it effectively. This approach requires the establishment of a state called "desertified" with respect to another "non-desertified" or historical/referential state [42].

The desertification risk assessment is carried out within the framework of this dichotomy. The analysis calculates the long-term equilibrium of key variables given current conditions. In this way, it attempts to know the effects of current land-use policies on the system, acting as an early warning

system. Therefore, there are two types of time trends. Those corresponding to stable behaviours are labelled as sustainable and those showing extinction of the resource or some other socioeconomic magnitude are considered desertification (Figure 2).

In particular, the risk of desertification is associated with the probability of losing a critical amount of a key resource within a time frame set by the user. To this end, a thousand simulations are carried out using randomly generated scenarios. Specifically, the RANDOM_NORMAL command from Vensim has been implemented. This returns a random number that is different on each successive time step. This algorithm considers a normal distribution of the parameter; in this analysis, some parameters of the scenario are converted into random variables with a mean value and a standard deviation that allows associating a normal distribution to it. For example, when random values of precipitation are used to assess risk, 1,000 precipitation scenarios are generated where values close to the average value of precipitation are more likely to appear than extreme values. After implementing scenarios of different random variables such as subsidies, prices, or the mentioned precipitation, the response of the system is analysed. The final value of the key variables is used to assess the risk of desertification. This indicator is defined as the percentage of simulations in which key resources are exhausted or critical thresholds are surpassed. In addition, it records the number of years required to lose that critical amount of resources.

2.3. Ranking of Desertification Factors. Sensitivity Analysis (SA) transcends the validation/calibration stage of a model. It is also a robust technique for obtaining information about a system. The objective is to classify factors (i.e., parameters) and distinguish orders of magnitude from their influences on system behaviour. Specifically, the Plackett–Burman Sensitivity Analysis [43] was performed. This is a sound statistical procedure that measures the effects of each parameter on the target variables in an efficient way in terms of the number of necessary scenarios. An important feature is that the effects of every parameter are not measured with the all-other-things-being-equal assumption but are averaged over variations made in all other parameters. PBSA also enables measuring two-way interactions of pairs of parameters, although this option has not been used in this case.

To apply PBSA, upper and lower values (10% for this evaluation) must first be assigned to each parameter in the model. The next step is to design $2d$ scenarios, where d is any multiple of 4 greater than the number of n parameters. Each scenario is a set of n parameter values that are sampled from the previously assigned upper and lower values. The design of these scenarios follows the patterns originally proposed by Plackett and Burman [43]. The effects of each parameter are obtained by adding the $2d$ outputs of the target variable and then dividing it by d.

The evaluation procedure focuses on tailor-made variables that assess the time needed to lose a key resource. These are "slow" variables, which are appropriate indicators for assessing the system change in the long term [6]. The positive sign indicates the fact that incrementing 10%, that parameter value delays degradation while a 10% decrease accelerates degradation. A negative impact should be interpreted in the opposite way.

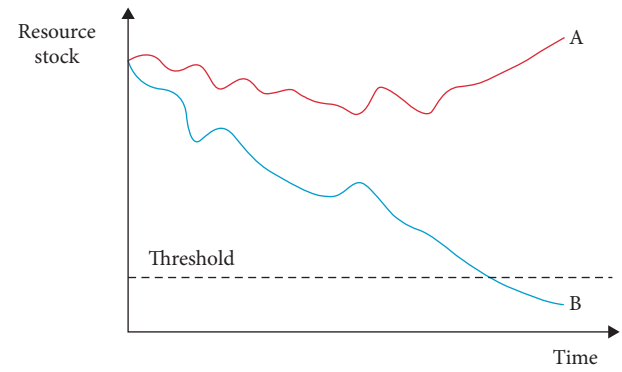


Figure 2: Basis for SAT early warning indicators. Some simulations do not exceed the degradation thresholds considered (A), while others do (B). The latter determine the risk of desertification.

3. Results

SAT was implemented for five case studies (Figure 3), each one of them representing a DL identified in the Spanish NPCD [24]: (DL1) "Woody crops affected by erosion"; (DL2) "Rainfed herbaceous crops with erosion risk"; (DL3) "Overgrazed agroforestry-pastoral systems"; (DL4) "Irrigated areas with desertification risk"; and (DL5) "Degraded shrublands and wastelands."

The five cases were covered with three versions of GDM previously developed by the authors:

(i) GDM1 (rangelands): water erosion in rangelands and shrublands [28, 37, 44, 45] for Dehesas in Cáceres province (DL3) and Sierra de Los Filabres in Almería province (DL5)

(ii) GDM2 (croplands): water erosion in extensive croplands for olive orchards (DL1) and wheat/sunflower crops rotation (DL2) in Córdoba province [46]

(iii) GDM3 (irrigation agriculture): hydrological models linked to groundwater-based irrigation agriculture [36, 47, 48] for Eastern La Mancha aquifer in Ciudad Real and Albacete provinces apply to DL4

Interaction with simulation models has improved considerably in recent years. However, it can still be difficult for a non-modelling user to deal directly with programs such as Vensim. User-friendly interfaces are essential to involve a broader audience in exploiting and improving models. The Vensim DSS® software [25] includes the possibility of developing simple Venapps applications to visualize the model structure, examine its causality, simulate the model behaviour, and visualize its results.

Alternatively, it is possible to control Vensim through other applications. This was the option chosen, as one of the main requirements in the SAT design was to use a platform to which most users were familiar. SAT is a Vensim

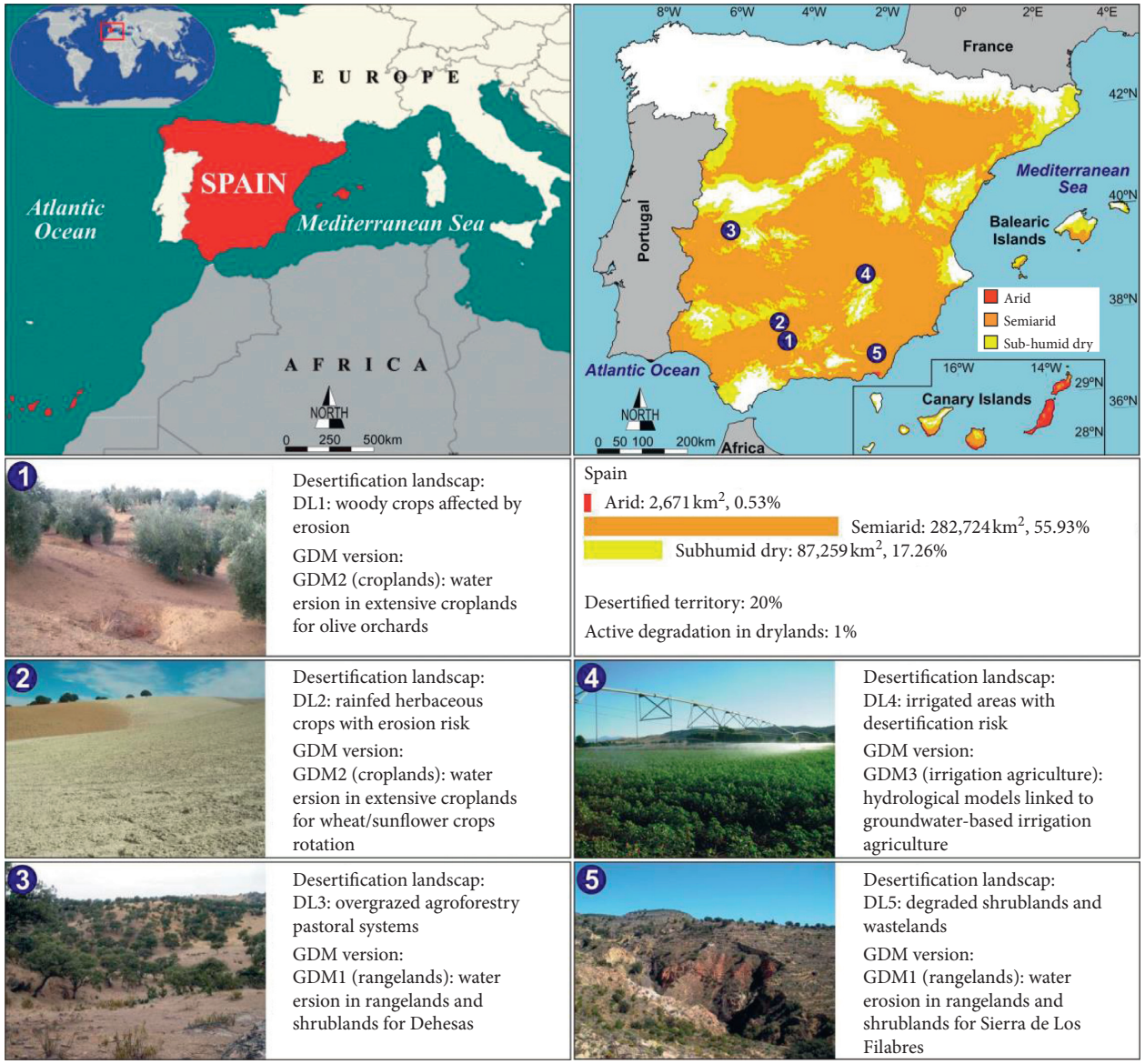


Figure 3: Drylands (arid, semiarid, and subhumid dry areas) in Spain and location of the five case studies.

Published Application (vpa) implemented in Excel that is ready to work with Vensim Model Reader (https://vensim.com/free-download/). In our case, the vpa consists of a SD model (SAT.vmf) and a caller (SAT.xls), the true user interface. Around these main files, other ones help to manage the model outside the Vensim environment (Figure 4). Table 1 lists all of them, and their interaction is illustrated in Figure 5.

SAT is organized into three blocks (Figure 6 (top)), each of which corresponds to a version of GDM. However, to simplify programming, the three GDM adaptations have been merged into a single Vensim file (SAT.vmf). The Main Menu (Figure 4) is used to go from one case to another and provides general information about the project. The other parts of the Excel book with which the user interacts are the scenarios and results. In each of these spreadsheets, the scenarios are established, the different indicators are calculated, and the results are presented (Figures 7 and 8).

Beyond the interface elements, there are hidden structures where all calculations and operations are executed (Figure 6 (bottom)). They consist of sheets inside the main workbook (SAT.xls) and external files that are necessary for the operation of the Vensim models. Thus, each version of GDM includes four additional worksheets to organize intermediate calculations and store the results and data used. These are as follows: (i) "s1" and "s2" contain the numerical values of the time trajectories that feed the graph, the "s1" being also used to store the base scenario values; (ii) the "risk" sheet stores and processes the results of the 1,000 simulations needed to assess desertification risk; and (iii) in "pbsa" the scenarios required by PBSA are implemented. In addition, here, the matrix calculations are executed to establish the ranking of the factors of desertification.

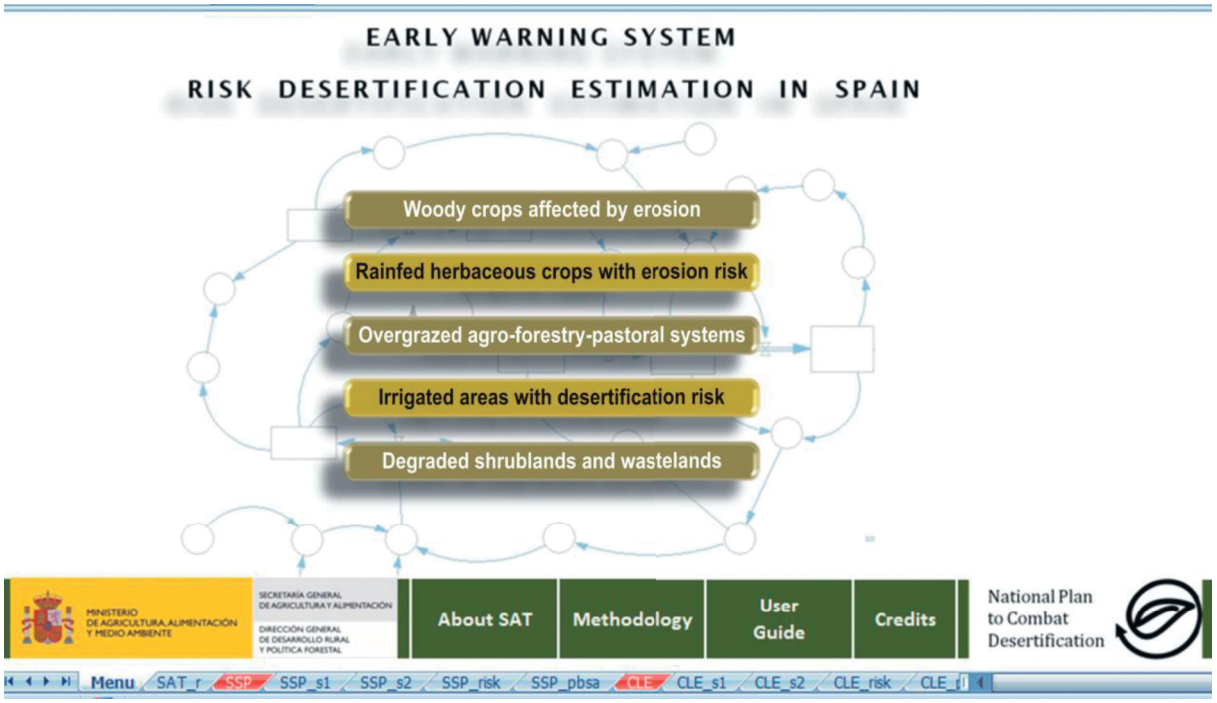


Figure 4: SAT Main Menu.

Table 1: Files compiled in SAT DSS.

File	Type	Number	Description
SAT.vmf	Vensim simulation model	1	The usual extension of a Vensim model is *.mdl. However, to be compiled, it must be saved in binary format (*.vmf).
SAT.xls	Excel file	1	Contains the user interface and works as a caller for the rest of files.
Results.xls	Excel file, v 4.0	2	Provides the link between Vensim and Excel.
Parameters.xls	Excel file, v 4.0	1	Provides the link between Excel and Vensim.
Name.vdf	Vensim Data file	*n*	Contains scenario results from a simulation; the name of the exercise is provided by the user.
Parameters.vdf	Vensim Data file	1	Contains data from the scenario of simulation provided through SAT.xls.
List.lst	Vensim list	9	List of model variables to export to Excel.
Read.frm	Vensim input form	1	Specifies how Excel data are read by Vensim.

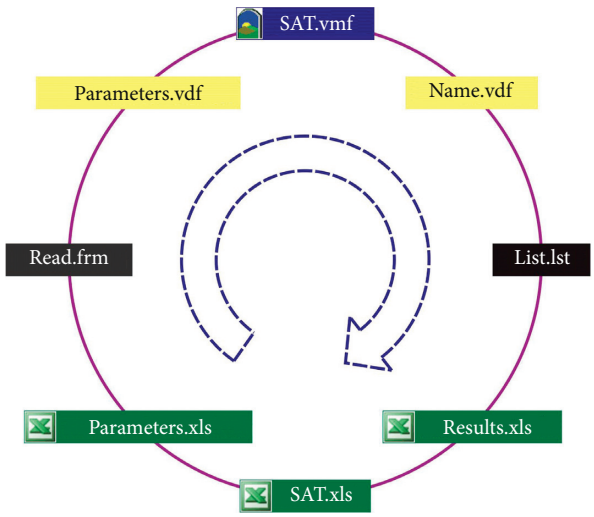
The additional files to connect the SAT interface with the model are Excel 4.0 files, the format required by Vensim. Finally, at the last level is the simulation model, the real core of the DSS.

The SAT tool is programmed with Visual Basic. A part of this code automates tasks in Excel and another one interacts with Vensim DSS through DLL commands. A dynamic link library (DLL) is a collection of small programs that can be loaded when larger ones need them. Each of these programs is made up of commands, i.e., actions performed by Vensim. They are introduced as a class, followed by the specific command and the options that allow its customization in the form class > command|option1|option2....

The command Vensim_command is the DLL used in SAT. Actually, a small portion of the available commands is enough to take control of Vensim from Excel. They belong to three classes (Table 2): (1) MENU commands have essentially the same effect as selecting a command from the Vensim workbench menu; (2) SIMULATE class allows controlling simulation input control parameters and also reads results from other runs to start another simulation; and (3) SPECIAL commands allow manipulating the Vensim environment in similar ways to what can be done from the Control Panel.

Figure 5 illustrates the operation of SAT. In a simple simulation exercise (which generates time trajectories), the loop is started after clicking on the "Trends" button (Figure 7(c)). Then, the *Parameters.xls* file is updated with the parametric values that the user has entered in the scenario. They are automatically converted to *Parameters.vdf*, the Vensim format (vdf denotes Vensim data file). The command used for this operation (MENU > XLS2VDF) involves the file *Read.frm*. It specifies the import options for Vensim to read the Excel file correctly.

Once data files are set to be used in the simulation (SIMULATE > ADDDATA; SIMULATE > DATA), the model is run (MENU > RUN) under a name given by the user (*Name.vdf*) (SIMULATE > RUNNAME). The



Type of exercise	Loop iterations	Time frame, years
Trends	1	200
Risk of desertification	1,000	100/2,000
Rankings	2,500/4, 100/4, 800	2,000

Figure 5: SAT basic operation scheme. The table shows the number of iterations and time frame for each of the simulation options.

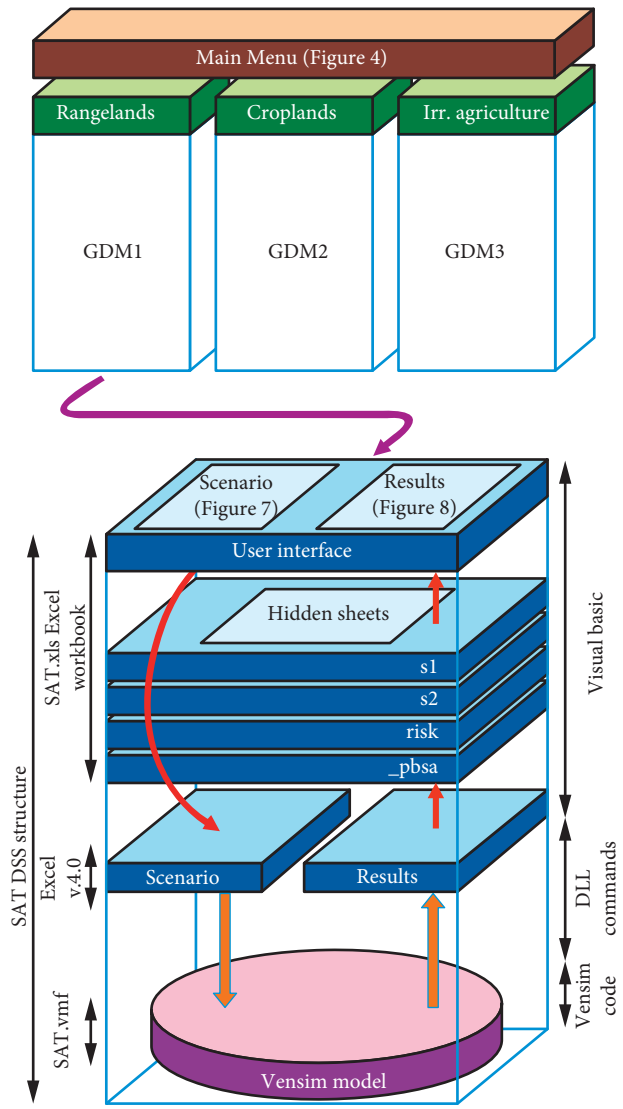


Figure 6: Overview of SAT DSS structure (top portion). The five case studies are implemented using only three simulation models. Each of them forms an independent block (accessible from the Main Menu) in which the relevant SAT.xls sheets and diverse Excel and Vensim files specific to the version of the GDM used are interconnected (see Figure 5). Detailed arrangement of the files for the Rangeland GDM version is shown (bottom portion).

information generated in each simulation includes values for each variable and for each time step (0.0625 years) for a long simulation period (200 years). However, only a small part of this huge amount of data is used in SAT. The Vensim data file (*Name.vdf*) is converted back to Excel (MENU > VDF2XLS), generating *Results.xls*. The file *List.lst* is used to determine which variables are exported. Finally, the information in this file is updated in *SAT.xls*.

For desertification risk assessment, this loop is repeated 1,000 times. To determine the ranking of factors, the number of simulations ranges from 2,400 to 4,800, since the amount of PBSA scenarios is subjected to the number of parameters of each GDM version. At each turn of the loop, the results of a simulation are saved in *SAT.xls*. Once the exercise is completed, a set of calculations is executed to calculate the indicators that the user finally sees. In addition, they can be saved in ∗.txt format for consultation.

When the user opens the application, the first screen that appears is Main Menu (Figure 4). Here the user can find general background information: a user guide, an overview of the implemented methodology, and the software credits. The Main Menu gives access to the five DLs included in SAT. After selection, the worksheet is updated with the corresponding baseline scenarios. The available options include (Figure 7): (a) links to the Main Menu and to the supporting documents on the analysed case; (b) implement new simulation scenarios; (c) possibility of launching three types of simulation exercises; (d) files management; and (e) summary report of the results of the exercises.

The basic two steps for working with SAT are to fill the data in the simulation scenario and then to run any of the three exercises. The scenario can be created by changing the baseline scenario loaded by default. It is also possible to make use of other scenarios that have been saved previously or create a completely new one by clicking on the "Clear scenario" option.

The "Trends" exercise immediately generates a graph representing the trajectory (red line) of the variable selected in the upper combo box. For comparison, an alternative simulation (blue line) can be loaded.

The other two exercises take longer (see table in Figure 5), as thousands of simulations to calculate desertification indicators and rankings are needed. After a few minutes, the results can be consulted by scrolling down (Figure 8). Together with the simulation scenario, the user has to set the time horizon of the exercise "Risk" in 100 or 2,000 years.

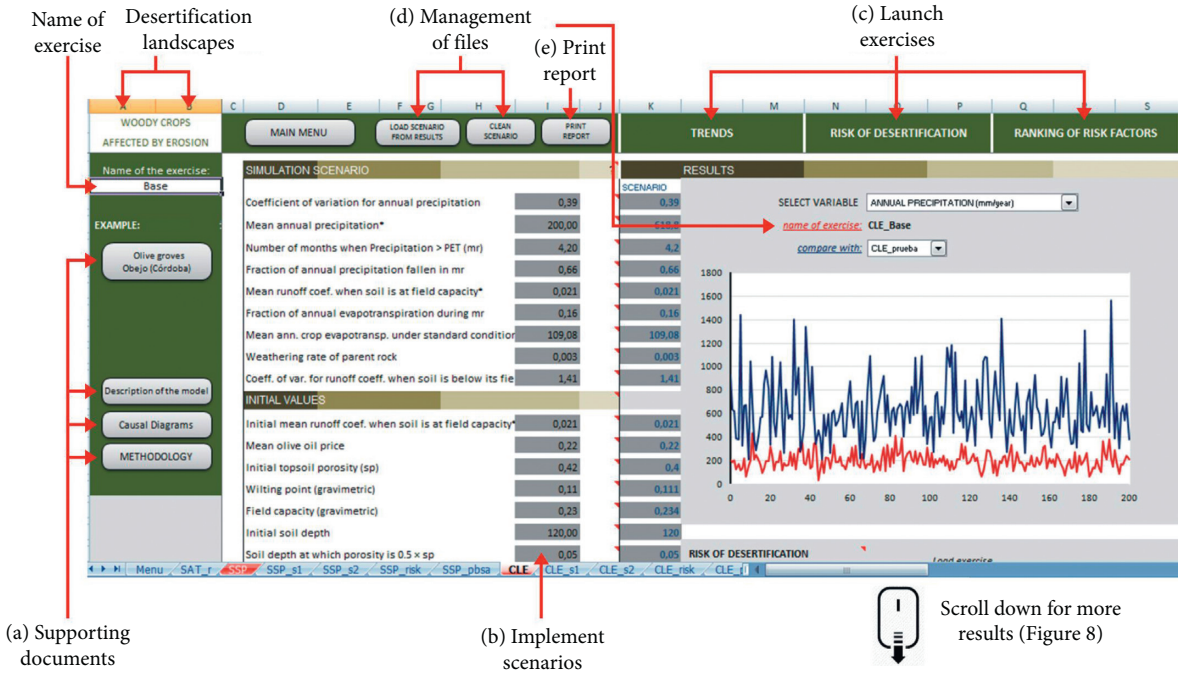
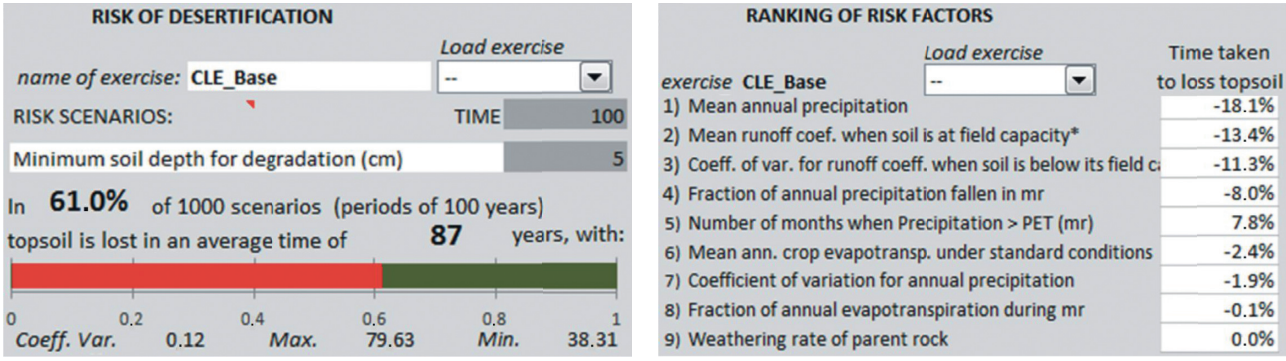


Figure 7: Screen for one of the cases implemented in SAT DSS.



(a) (b)

Figure 8: Outputs for risk (a) and ranking (b) exercises. Example for landscape DL1 "Woody crops affected by erosion."

Table 2: Vensim commands implemented in SAT.

Command	Function
MENU > XLS2VDF	Converts the Excel format file to the ∗.vdf format file
MENU > RUN	Runs the model using the current simulation setup
MENU > VDF2XLS	Converts the Vensim dataset to an Excel v. 4.0 format file
SIMULATE > ADDDATA	Adds a ∗.vdf file to the list of data files to be used in the simulation
SIMULATE > DATA	Sets the data file(s) to be used in the simulation
SIMULATE > RUNNAME	Sets the name of the simulation run
SPECIAL > LOADMODEL	Loads the model

By way of illustration, the results for DL1 are shown in Figure 8(a). They can be interpreted as follows: the risk of soil thickness falling below a critical threshold for productivity (set at 5 cm for the user in this example) is estimated at 61%, and this can occur within 87 years. Together with the main output, a statistic summary (mean, coefficient of variation, max, and min) is included. Finally, Figure 8(b) shows the classification of the parameters. The results are interpreted as follows: a 10% increase in "mean annual precipitation" means that the "time needed to lose the topsoil layer," the target variable for this case, increases by 18.1%.

All simulation exercises are saved by default for later use. Trends are saved in ∗.vdf format generated by Vensim. From there, both simulation scenarios and results can be retrieved.

Table 3: Time for soil collapse under two assessment methodologies.

Case study	"Simple" soil erosion indicator, years	SAT estimation, years	Percentage of variation (%)
DL2	200	26	−87
DL1	657	169	−74
DL5	1,045	186	−82
DL3	1,739	349	−80

From the other two exercises, only the final output, i.e., the desertification risk and the ranking of risk factors, is saved in ∗.txt format, but not the thousands of simulations necessary for their calculation.

4. Discussion and Conclusions

Prevention and mitigation of desertification includes understanding its causes, monitoring and evaluating its progression, and designing and implementing site-specific management strategies [18]. Given the irreversible nature of desertification, it is essential to anticipate the problem [22, 49, 50]. SAT is a software aimed at contributing to this problematic of the fight against desertification. The mere fact of constructing a SD model that allows the understanding of the physical and socioeconomic processes involved is a great advance to propose sustainable solutions. Furthermore, by coupling numerical methods such as SA or probability calculations, we achieve an objective assessment of the risk of desertification and of the most important factors.

Our early warning system complements the indicators that the Spanish NPCD [24] already had. These were based on the methodology of MEDALUS [51], a widely known technique for assessing desertification in the Mediterranean area. This provided a qualitative evaluation of the risk of desertification of the territory by analysing each pixel from basic information that was readily available (estimated erosion, aridity index, aquifer overexploitation, and recurrence of forest fires). Its weakness has been highlighted in the latest World Atlas of Desertification [52]. The main criticism of this approach is that land degradation cannot be mapped by a single indicator or by any arithmetic or modelled combination of variables [53]. In addition, the dynamics of the processes are ignored, the importance of which can be illustrated by the following example. Table 3 shows the time for soil loss in the case studies included in SAT under two methodological approaches: (i) a "simple" indicator (second column) estimates soil loss by dividing the soil thickness by the average annual rate of erosion (previously converted from tons/ha to mm through soil bulk density) and (ii) SAT procedure (third column). The percentage of variation relative to the "simple" indicator appears in the last column.

It is noted that the degradation times for the "simple" indicator are much longer than the SAT results. The reason is due to the static nature of the above indicators. In fact, the assumption of a constant erosion rate, which implicitly incorporates the "simple" indicator, is not supported from a medium-to-long-term perspective. This consideration is essential for desertification, as this process is governed by "slow variables" [6]. By way of illustration, erosion rates may change due to fluctuations in vegetation cover—the literature describes the exponential nature of this relationship [54]— driven by livestock density which, in turn, is sensitive to price changes in livestock markets [55].

In fact, the differences observed between the two types of indicators are the result of considering positive feedback in the soil loss process. Soil erosion exposes layers of higher apparent density, i.e., deep layers having lower porosity and lower infiltration rates. Soil crusting prevents seeds from sprouting and reduces water entry into the soil. This, in turn, increases runoff rates and soil erosion. All these mechanisms are ignored by the "simple" indicators, which consider soil loss as a linear process.

There are also clear limitations in the case of groundwater-based irrigation agriculture. It could be said that the sustainable groundwater use implies that the aquifer piezometric level is steady. However, the shift from a steady state (aquifer under natural regime) to the other state (aquifer under sustainable use) implies that the recharge-to-discharge ratio turns less than one for the period in which the disturbance (the development of irrigation agriculture) remains. Hence, just like the use of average constant erosion rates, the implementation of constant recharge-to-discharge indicators ignores the dynamics (feedbacks and delays) of the groundwater resource. Actually, a less-than-one ratio is a necessary, though not sufficient, condition to declare groundwater overexploitation [56, 57]. Some key questions arise as follows: What is the expected trend for irrigation areas? Is it possible that the recharge-to-discharge ratio returns to one? How long can it take? What would be the stock of groundwater at the time? It is impossible to get an idea of the answers to these questions without taking into account the set of dynamic processes—at least the most important ones—involved in the use of the groundwater resource.

Monitoring indicators over time helps to overcome their static disadvantage. The inherent problem with indicators based on "flash" observations or remote sensing techniques is that they show the current, actual state of the system. This merit becomes a disadvantage when it comes to inertial systems. When irreversible thresholds and delays are part of the system [19, 27, 58], it is imperative to have an idea of the general trend of the system before the symptoms of desertification become so evident. Otherwise, monitoring results in a kind of necropsy that records the path to collapse.

The users targeted by SAT are government land-use planners as well as desertification experts. The former are in charge of updating the NCPD, and SAT will provide them with land use trends under different scenarios and their associated desertification risk. The latter will contribute their knowledge and data to improve process modelling and to refine thresholds and parametric values.

The growing perception of the complex interaction between the sociocultural, economic, and biophysical components of systems is the cause of the growing interest in environmental DSSs [59, 60]. Such platforms play a key role in the expansion of research results to society. They allow, in an easy way, to explore modelling and its results and introduce their users to the decision-making process [61]. In addition, they diminish the resilience of society in dealing with the problems of desertification [18].

DSS should work in two ways: first, by bringing scientific research closer to social agents. Thus, when the decision-making process and the premises on which it is based are opened to public debate, there is a gain in transparency [62]. Second, this approach can be useful for researchers. In fact, the participation of non-scentific stakeholders and experts raises interesting questions as they reflect society's main environmental concerns.

On SAT, one of the main tasks of stakeholders and experts was to agree on degradation thresholds. There were some difficulties in interpreting the results and defining the scenarios, as some of the parameters were difficult to understand. This is part of the communication barriers between scientific research and the real world, and as in [59], "there is a trade-off between the attempt to simplify the intrinsic complexity and the need for scientifically robust approaches and detailed high-quality data."

An important conclusion after SAT is the need to model the land-use change (LUC) [42], as these are the main drivers of global environmental change [63, 64]. Model development for particular land uses provides a limited perspective on the causes and consequences of desertification [6, 65]. We agree with the demand to develop models that cover various land uses in an increasingly interconnected world (e.g., landless livestock trigger soybean cultivation). However, GDM and SAT were conceived under the umbrella of DLs linked to specific land-uses, nourished by the *hotspot* concept, i.e., a sudden increase of economic productivity driven by the intensification and expansion of a certain land-use. What is more, GDM implicitly considers the most important form of land conversion, which is the encroachment of arable and grazing lands on natural ecosystems [63].

This reason encourages us to include more case studies in SAT. In addition, specific land-use modelling is a preliminary step for its integration into the LUC framework. Increasing the collection of models will give us a precise idea of the desertification risk state in Spain. Even more, the computational system is ready to be exported to other countries. GDM has been already adapted to study groundwater-dependent oases in Morocco [56]; Alcalá et al. [66, 48], Greek rangelands [37], and Algerian steppe rangelands [50]. Given that all the UNCCD signatory countries are committed to develop their NPCDs, this kind of tools can help threatened regions to spot desertification processes in their initial stages.

Under the increasing climatic and social uncertainties enveloping the entire planet, SAT pursues—to quote Abraham Lincoln—not only to know where we are but also to sketch where we are trending. In that case, we could better judge what to do and how to do it.

Abbreviations

DL: Desertification landscape
DLL: Dynamic link library
DSS: Decision support system
GDM: Generic desertification model
LUC: Land-use change
MAFF: Ministry of Agriculture, Fisheries, and Food of Spain
NPCD: National Plan to Combat Desertification
PBSA: Plackett–Burman sensitivity analysis
SA: Sensitivity analysis
SAT: Sistema de Alerta Temprana (early warning system)
SD: System dynamics
UNCCD: United Nations Convention to Combat Desertification.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was funded by TRAGSATEC Public Enterprise on behalf of the Spanish Ministry of Agriculture, Food, and Environment under Contract 23.674. The generic desertification model was developed under the umbrella of DeSurvey IP (European Commission FP6 Contract no. 003950). The paper was written within the framework of the project BIODESERT, funded by the European Research Council (ERC Grant agreement no. 647038).

References

[1] UNCCD, "United Nations convention to combat desertification in countries experiencing serious drought and/or desertification, particularly in Africa," *International Legal Materials*, vol. 33, no. 5, pp. 1328–1382, 1994.

[2] N. Middleton, *Deserts: A Very Short Introduction*, Oxford University Press, Oxford, UK, 2009.

[3] R. Lal, "Potential of desertification control to sequester carbon and mitigate the greenhouse effect," *Storing Carbon in Agricultural Soils: A Multi-Purpose Environmental Strategy*, vol. 51, pp. 35–72, 2001.

[4] Millenium Ecosystem Assessment (MEA), *Ecosystems and Human Well-Being: Desertification Synthesis*, World Resources Institute, Washington, DC, USA, 2005.

[5] U. Safriel and Z. Adeel, "Dryland systems," in *Ecosystems and Human Well-Being: Current State and Trends*, R. Hassan, Ed., vol. 1, pp. 623–662, Island Press, Whasington DC, USA, 2005.

[6] J. F. Reynolds, D. M. S. Smith, E. F. Lambin et al., "Global desertification: building a science for dryland development," *Science*, vol. 316, no. 5826, pp. 847–851, 2007.

[7] IPCC, Intergovernmental Panel on Climate Change, "Summary for policymakers," in *Climate Change 2007: The Physical*

Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC, pp. 1–23, Cambridge University Press, New York, NY, USA, 2007.

[8] N. A. Geeson, C. J. Brandt, and J. B. Thornes, *Mediterranean Desertification: A Mosaic of Processes and Responses*, John Wiley & Sons, Ltd., Hoboken, NJ, USA, 2002.

[9] J. B. Thornes, "Environmental crisis in the mediterranean," in *Geography, Environment and Development in the Mediterranean*, R. King, P. De Mas, and J. Mansvelt Beck, Eds., Sussex Academic, Brighton, UK, 2001.

[10] J. Puigdefabregas, "Desertification: stress beyond resilience, exploring a unifying process structure," *Ambio*, vol. 24, pp. 311–313, 1995.

[11] J. Hill, M. Stellmes, T. Udelhoven, A. Röder, and S. Sommer, "Mediterranean desertification and land degradation," *Global and Planetary Change*, vol. 64, no. 3-4, pp. 146–157, 2008.

[12] B. C. Bates, Z. W. Kundzewicz, S. Wu, and J. P. Palutikof, *Climate Change and Water*, American Society of Civil Engineer, Reston, VA, USA, 2008.

[13] L. Samaniego, S. Thober, R. Kumar et al., "Anthropogenic warming exacerbates European soil moisture droughts," *Nature Climate Change*, vol. 8, no. 5, pp. 421–426, 2018.

[14] A. Mariotti, N. Zeng, J. Yoon, V. Artale, A. Navarra, and P. Alpert, "Mediterranean water cycle changes : transition to drier 21st century conditions in observations and CMIP3 simulations," *Environmental Research Letters*, vol. 3, no. 4, 2008.

[15] J.-F. Bastin, Y. Finegold, C. Garcia et al., "The global tree restoration potential," *Science*, vol. 365, no. 6448, pp. 76–79, 2019.

[16] O. Branch and V. Wulfmeyer, "Deliberate enhancement of rainfall using desert plantations," *Proceedings of the National Academy of Sciences*, vol. 116, no. 38, pp. 18841–18847, 2019.

[17] WOCAT, "Where the land is greener—case studies and analysis of soil and water conservation initiatives worldwide," in *WOCAT—World Overview of Conservation Approaches and Technologies*, H. Liniger and W. Critchley, Eds., WOCAT, Switzerland, 2007.

[18] P. D'Odorico, A. Bhattachan, K. Davis, S. Ravi, and C. Runyan, "Global desertification: drivers and feedbacks," *Advances in Water Resources*, vol. 51, pp. 326–344, 2013.

[19] V. Dakos, S. R. Carpenter, W. A. Brock et al., "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data," *PLoS One*, vol. 7, no. 7, Article ID e41010, 2012.

[20] M. Scheffer, J. Bascompte, W. A. Brock et al., "Early-warning signals for critical transitions," *Nature*, vol. 461, no. 7260, pp. 53–59, 2009.

[21] S. Kéfi, V. Guttal, W. A. Brock et al., "Early warning signals of ecological transitions: methods for spatial patterns," *PLoS One*, vol. 9, no. 3, Article ID e92097, 2014.

[22] J. Martínez-Valderrama, J. Ibáñez, G. Del Barrio et al., "Present and future of desertification in Spain: implementation of a surveillance system to prevent land degradation," *Science of The Total Environment*, vol. 563-564, pp. 169–178, 2016.

[23] M. E. Sanjuán, G. Del Barrio, A. Ruiz, L. Rojo, J. Puigdefábregas, and A. Martínez, *Evaluación y seguimiento de la desertificación en España: Mapa de la Condición de la Tierra 2000–2010*, Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid, Spain, 2014.

[24] MAGRAMA, *Programa de Acción Nacional contra la Desertificación*, MAGRAMA, Madrid, Spain, 2008.

[25] Ventana Systems Inc., *VensimDSS Software*, Ventana Systems, Inc., Harvard, MA, USA, 2006.

[26] J. W. Forrester, *Industrial Dynamics*, The MIT Press, Cambridge, MA, USA, 1961.

[27] J. D. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, Mc Graw Hill, New York, NY, USA, 2000.

[28] J. Martínez-Valderrama and J. Ibañez, "Foundations for a dynamic model to analyze stability in commercial grazing systems," in *Sustainability of Agrosilvopastoral Systems–Dehesas, Montados*, S. Schnabel and A. Ferreira, Eds., pp. 173–182, 2004.

[29] J. Ibáñez, J. M. Valderrama, and J. Puigdefábregas, "Assessing desertification risk using system stability condition analysis," *Ecological Modelling*, vol. 213, no. 2, pp. 180–190, 2008.

[30] A. J. Lotka, *Elements of Mathematical Biology*, Dover Publications, New York, NY, USA, 1956.

[31] V. Volterra, "Variations and fluctuations of the number of individuals in animal species living together," in *Animal Ecology*, B. M. Chapman, Ed., pp. 409–448, McGraw-Hill, New York, NY, USA, 1931.

[32] I. Noy-Meir, "Stability of grazing systems: an application of predator-prey graphs," *The Journal of Ecology*, vol. 63, no. 2, pp. 459–481, 1975.

[33] J. B. Thornes, "Erosional equilibria under grazing," in *Conceptual Issues in Enviromental Archaelology*, J. L. Bintliff, D. A. Davidson, and E. G. Grant, Eds., Edimburgh University Press, Edimburgh, UK, 1988.

[34] E. Diez and B. S. McIntosh, "Organisational drivers for, constraints on and impacts of decision and information support tool use in desertification policy and management," *Environmental Modelling & Software*, vol. 26, no. 3, pp. 317–327, 2011.

[35] R. A. Kelly, A. J. Jakeman, O. Barreteau et al., "Selecting among five common modelling approaches for integrated environmental assessment and management," *Environmental Modelling and Software*, vol. 47, pp. 159–181, 2013.

[36] J. Martínez-Valderrama, J. Ibáñez, and F. J. Alcalá, "AQUACOAST: a simulation tool to explore coastal groundwater and irrigation farming interactions," *Scientific Programming*, vol. 2020, Article ID 9092829, 20 pages, 2020.

[37] J. Ibañez, J. M. Valderrama, V. Papanastasis, C. Evangelou, and J. Puigdefabregas, "A multidisciplinary model for assessing degradation in Mediterranean rangelands," *Land Degradation & Development*, vol. 25, no. 5, pp. 468–482, 2014.

[38] M. F. Quaas, S. Baumgärtner, C. Becker, K. Frank, and B. Müller, "Uncertainty and sustainability in the management of rangelands," *Ecological Economics*, vol. 62, no. 2, pp. 251–266, 2007.

[39] G. L. W. Perry and J. D. A.c Millington, "Spatial modelling of succession-disturbance dynamics in forest ecosystems: concepts and examples," *Perspectives in Plant Ecology, Evolution and Systematics*, vol. 9, pp. 191–210, 2008.

[40] T. Oxley, B. McIntosh, N. Winder, M. Mulligan, and G. Engelen, "Integrated modelling and decision-support tools: a Mediterranean example," *Environmental Modelling & Software*, vol. 19, no. 11, pp. 999–1010, 2004.

[41] M. Westoby, B. Walker, and I. Noy-Meir, "Opportunistic management for rangelands not at equilibrium," *Journal of Range Management*, vol. 42, no. 4, pp. 266–274, 1989.

[42] B. T. Bestelmeyer, G. S. Okin, M. C. Duniway et al., "Desertification, land use, and the transformation of global drylands," *Frontiers in Ecology and the Environment*, vol. 13, no. 1, pp. 28–36, 2015.

[43] R. L. Plackett and J. P. Burman, "The design of optimum multifactorial experiments," *Biometrika*, vol. 33, no. 4, pp. 305–325, 1946.

[44] J. Ibáñez, J. F. L. Contador, S. Schnabel, M. P. Fernández, and J. M. Valderrama, "A model-based integrated assessment of land degradation by water erosion in a valuable Spanish rangeland," *Environmental Modelling & Software*, vol. 55, pp. 201–213, 2014.

[45] J. Ibáñez, J. Martínez, and S. Schnabel, "Desertification due to overgrazing in a dynamic commercial livestock-grass-soil system," *Ecological Modelling*, vol. 205, no. 3-4, pp. 277–288, 2007.

[46] J. Ibáñez, J. Martínez-Valderrama, E. V. Taguas, and J. A. Gómez, "Long-term implications of water erosion in olive-growing areas in southern Spain arising from a model-based integrated assessment at hillside scale," *Agricultural Systems*, vol. 127, pp. 70–80, 2014c.

[47] J. Ibáñez, J. M. Valderrama, and J. Puigdefábregas, "Assessing overexploitation in Mediterranean aquifers using system stability condition analysis," *Ecological Modelling*, vol. 218, no. 3-4, pp. 260–266, 2008.

[48] J. M. Valderrama, J. Ibáñez, F. J. Alcalá, A. Dominguez, M. Yassin, and J. Puigdefábregas, "The use of a hydrological-economic model to assess sustainability in groundwater-dependent agriculture in drylands," *Journal of Hydrology*, vol. 402, no. 1-2, pp. 80–91, 2011.

[49] X. Feng, B. Fu, S. Piao et al., "Revegetation in China's Loess Plateau is approaching sustainable water resource limits," *Nature Climate Change*, vol. 6, no. 11, pp. 1019–1022, 2016.

[50] J. Martínez-Valderrama, J. Ibáñez, G. Del Barrio et al., "Doomed to collapse: why Algerian steppe rangelands are overgrazed and some lessons to help land-use transitions," *Science of the Total Environment*, vol. 613-614, pp. 1489–1497, 2018.

[51] EC-European Commission, "The Medalus Project: Mediterranean desertification and land use: manual on key indicators of desertification and mapping environmentally sensitive areas to desertification. European Commission Publication EUR 18882," in *Directorate-General Science, Research and Development, Office for Official Publications of the European Communities, Luxembourg*, C., Kosmas, M. J., Kirkby, and N., Geeson, Eds., p. 87, 1999.

[52] M. Cherlet, C. Hutchinson, J. Reynolds, J. Hill, S. Sommer, and G. von Maltitz, *World Atlas of Desertification*, Publication Office of the European Union, Brussels, Belgium, 2018.

[53] S. D. Prince, "Where does desertification occur? Mapping dryland degradation at regional to global scales," in *The End of Desertification? Disputing Environmental Change in the Drylands*, R. Behnke and M. Mortimore, Eds., pp. 225–263, Springer, Berlin, Germany, 2016.

[54] H. A. Elwell and M. A. Stocking, "Vegetal cover to estimate soil erosion hazard in Rhodesia," *Geoderma*, vol. 15, no. 1, pp. 61–70, 1976.

[55] J. Ibáñez, J. Martínez-Valderrama, J. F. L. Contador, and M. P. Fernández, "Exploring the economic, social and environmental prospects for commercial natural annual grasslands by performing a sensitivity analysis on a multidisciplinary integrated model," *Science of the Total Environment*, vol. 705, Article ID 135860, 2020.

[56] F. J. Alcalá, J. Martínez-Valderrama, P. Robles-Marín et al., "A hydrological-economic model for sustainable groundwater use in sparse-data drylands: application to the Amtoudi Oasis in southern Morocco, northern Sahara," *Science of the Total Environment*, vol. 537, pp. 309–322, 2015.

[57] E. Custodio, "Aquifer overexploitation: what does it mean?" *Hydrogeology Journal*, vol. 10, no. 2, pp. 254–277, 2002.

[58] M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, and B. Walker, "Catastrophic shifts in ecosystems," *Nature*, vol. 413, no. 6856, pp. 591–596, 2001.

[59] M. Matthies, C. Giupponi, and B. Ostendorf, "Environmental decision support systems: current issues, methods and tools," *Environmental Modelling & Software*, vol. 22, no. 2, pp. 123–127, 2007.

[60] A. E. Rizzoli and W. J. Young, "Delivering environmental decision support systems: software tools and techniques," *Environmental Modelling & Software*, vol. 12, no. 2-3, pp. 237–249, 1997.

[61] H. van Delden, R. Seppelt, R. White, and A. J. Jakeman, "A methodology for the design and development of integrated models for policy support," *Environmental Modelling & Software*, vol. 26, no. 3, pp. 266–279, 2011.

[62] K. A. Stave, "Using system dynamics to improve public participation in environmental decisions," *System Dynamics Review*, vol. 18, no. 2, pp. 139–167, 2002.

[63] E. F. Lambin and P. Meyfroidt, "Global land use change, economic globalization, and the looming land scarcity," *Proceedings of the National Academy of Sciences*, vol. 108, no. 9, pp. 3465–3472, 2011.

[64] B. L. Turner II, E. F. Lambin, and A. Reenberg, "The emergence of land change science for global environmental change and sustainability," *PNAS*, vol. 103, no. 128, pp. 13070–13075, 2009.

[65] H. J. Geist and E. F. Lambin, "Dynamic causal patterns of desertification," *BioScience*, vol. 54, no. 9, pp. 817–829, 2004.

[66] F. J., Alcalá, M., Martín-Martín, F., Guerrera, J. Martínez-Valderrama, and P. A Robles-Marín, "Feasible methodology for groundwater resource modelling for sustainable use in sparse-data drylands: application to the Amtoudi Oasis in the northern Sahara," *Science Total Environment*, vol. 630, pp. 1246–1257, 2018.

Hindawi
Scientific Programming
Volume 2020, Article ID 5607242, 9 pages
https://doi.org/10.1155/2020/5607242

Research Article

Flow Enhancement of Mineral Pastes to Increase Water Recovery in Tailings: A Matlab-Based Imaging Processing Tool

S. L. Mondaca,[1] C. A. Leiva ,[1] C. A. Acuña,[2] and E. A. Serey[1]

[1]*Department of Chemical Engineering, Universidad Católica del Norte, 1270709 Antofagasta, Chile*
[2]*Department of Chemical and Environmental Engineering, Universidad Técnica Federico Santa María, 2390123 Valparaíso, Chile*

Correspondence should be addressed to C. A. Leiva; cleiva01@ucn.cl

Received 8 January 2020; Revised 29 April 2020; Accepted 11 May 2020; Published 26 May 2020

Academic Editor: Francisco J. Alcalá

The rate of growth of mining copper industry in Chile requires higher consumption of water, which is a resource limited in quality and quantity and a major point of concern in present times. In addition, the efficient use of water is restricted due to high levels of evaporation (10 to 15 (l/m^2) per day), in particular at the north highland mining sites (Chile). On the contrary, the final disposal of tailings is mainly on pond, which loses water by evaporation and in some cases by percolation. An alternative are the paste thickeners, which generate stable paste (70% solids), reducing evaporation and percolation and therefore reducing water make up. Water is a resource with more demand as the industries are expanding, making the water recovery processes more of a necessity than a simple upgrade in efficiency. This technology was developed in Canada (early 80s) and it has widely been used in Australia (arid zones with similar weather conditions to Chile), although few plants are using this technology. The tendency in the near future is to move from open ponds to paste thickeners. One of the examples of this is Minera El Tesoro. This scenario requires developing technical capacity in both paste flow characterization and rheology modifiers (fluidity enhancer) in order to make possible the final disposal of this paste. In this context, a new technique is introduced and experimental results of fluidity modifiers are discussed. This study describes how water content affects the flow behavior and depositional geometry of tailings and silica flour pastes. The depositional angle determined from the flume tests, and the yield stresses is determined from slump test and a rheological model. Both techniques incorporate digital video and image analysis. The results indicate that the new technique can be incorporated in order to determine the proper solid content and modifiers to a given fluidity requirement. In addition, the experimental results showed that the pH controls strongly the fluid paste behavior.

1. Introduction

The growth rate of the copper mining industry in Chile requires a greater water consumption, a resource with limited availability whose industrial use is a major concern nowadays [1, 2]. The recovery of water from thickened tailings in mining results in an important benefit economically speaking, although it gets more difficult to dewater tailings if they have fine particles [3].

In addition, the efficient use of water is restricted due to the high levels of evaporation (10 to 15 (l/m^2) per day) [4], particularly on the sites of the mines located in the northern highlands. The use of paste thickeners for water recovery proves to be a very efficient method to prevent water loses in tailings depositions, taking into account the large volume of water lost to evaporation [5].

In the case of waste materials from the flotation process (tailings), they range between 95–98% [4] by weight of the treated material, which are disposed in tailings dams in pulp form. Its transport is carried out by pipeline with an approximate solids consistency from 40% to 60% [4], to prevent the line from getting stuck and to reduce energy consumption during pumping [6]. Given the large volumes of water handled, part of the water is recovered and recirculated to the flotation stage after the solids are settled in the tailings dam. Alternatively, another option to increase the

water recovery is to thicken the pulp up to 75%, which is then available as a paste, using new technologies, such as thickened tailings disposal (TTD) [7].

The use of new technologies or "Paste Technology" [8] was firstly utilized at the end of the year 1980 [9] and since then it is used to describe the disposal of paste with a high content of fines (silt or clay) and low contents of water in the filling material that can be transported through distribution pipes (10 to 20 cm in diameter). This "Paste," unlike the filling sludge, does not separate in different phases (solid and liquid) at low speeds in transport pipes. The transport of pastes by pipes requires that enough fine material (20 [μm] or less) is contained within the mixture so that the lubrication of the pipe walls creates a state of piston flow. Tailings can be transported for filling uses or, given their high surface, in a tailings management facility.

This natural abundance of fine-grained materials is not exclusive in tailings from mines, and it is also common in waste streams from other industries, which are transported as pastes by pipes and are located in places without producing particle segregation. The potential use of paste technology has expanded in recent years [3, 5] because it is recognized as a low-cost, environmentally superior technology with technologically controllable methods for transporting and allows commercializing certain waste streams from materials with a low water content and high density. An example of the economic impact of implementing this technology is the comparison made by JRI Engineering, for the Codelco Norte Division, where it is concluded that paste technology yields up to 15% savings, compared to the conventional alternative [10]. Rheology control in pastes has various possible effects or consequences. The nature of the paste facilitates the mixing of one or more secondary waste materials, which allows all waste materials to be transported by pipes. The increase in density decreases the ratio of supernatant/treatment, reduces mobilization of contaminants, hydraulic conductivity, and oxygen/reactivity of water through encapsulation within the paste matrix [8]. Due to these effects, among others [11–15], it is proposed to set up a laboratory with characterization tests, which determines yield stress and angle of repose in pastes. Additionally, taking into account that the use of a more automatic control system allows increasing the efficiency of the procedures in different areas [16–19], it is proposed to develop a technique based on digital video and image processing using Toolbox Image Processing-Matlab, system widely utilized to facilitate the analysis of data for different types of processes [20–23].

The results allow to quantitatively evaluate the effect of pH on the fluidility and angle of repose in the test pastes. With respect to the additive used, sulfuric acid acts as a dispersant increasing the fluidity and decreasing the angle of repose in the pastes; in the case of the technique implemented, it is concluded that this allows to analyze the dynamic behavior and the profile of the paste throughout time, making the rheological study in pastes possible.

2. Materials and Methodology

2.1. Designs of Slump Test Equipment. Slump and flume test equipment are designed and constructed based on the criteria and methodologies developed in previous investigations [24, 25]. In the slump test, the height of the paste is measured over time, to determine the corresponding yield stress. Analysis models have been developed to relate the slump value to the value of the corresponding yield stress and to predict the slump behavior of the material. Slump models are derived from the first principles with the model variables expressed in dimensionless form. Thus, the slump is not an empirical model and provides the independent matter with a unique relationship between the yield stress and the slump height.

The slump test, which is usually performed on conical geometries, was adapted to a cylindrical geometry for the alumina industry. This is due to the fact that there is a relationship between the slump height and the flow behavior in the bauxite waste, correlation which is adaptable to the case of copper mining pastes due to similar physical characteristics.

Figure 1 shows the comparison between the cone and cylinder models, both the theoretical results and those obtained in the laboratory. It can be clearly seen that the experimental results when using the cylinder model are quite close to the theoretical model, unlike the cone model that shows a significant deviation compared to the empirical results. Taking this information into account, it is decided to use the cylindrical model, as it predicts more precisely the yield stress in pastes.

The results are expressed in terms of dimensionless variables to allow the generalization of the slump test model and are defined as follows:

$$\text{Cylinders}' = 1 - 2\tau_y' \left[1 - \ln\left(2\tau_y'\right)\right], \quad (1)$$

where $\tau_y' = \tau_y/\rho g H$ dimensionless creep limit. $s' = s/H$ dimensionless drop height. $h_0' = h_0/H$ dimensionless paste height without deformation. $h_1' = h_1/H$ dimensionless deformed paste height.

Figure 2 determines that, for the cylinder test, the diameter that obtained a lower value of deviation (<10% between the theoretical-experimental curves) was the geometry with diameter of 200 mm, with a value in the diameter-height ratio of 1. From this it is concluded that the diameter of the cylinder to be used in the experimentation will be 200 mm.

To determine the yield stress in the slump test, a series of assumptions are used, which are analyzed in more detail in reference [24]:

(a) It is assumed that no deformation of the paste occurs in the removal of the cylinder in the drop test.

(b) Only the effort acts on the test paste and is assumed as a vertical effort associated with the weight of the paste itself.

(c) The test paste is supposed to behave like an elastic solid so that the maximum shear stress that can be

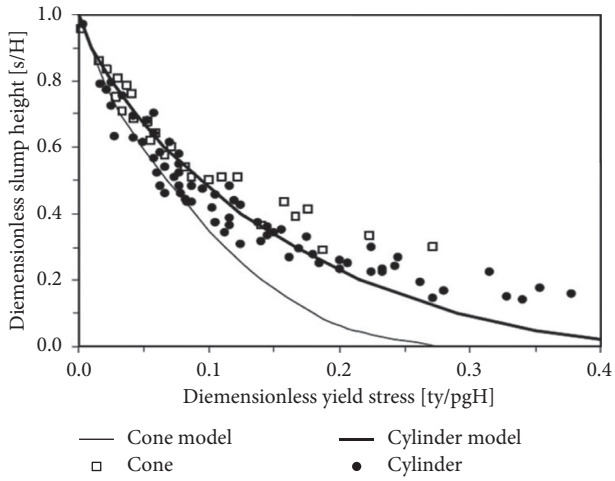


Figure 1: Theoretical-experimental comparison of cylinder and cone models, figure adapted from reference [24].

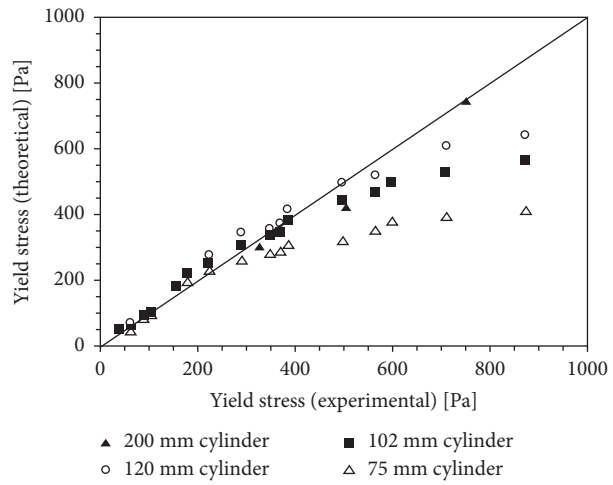


Figure 2: Relationship between the diameter of the cylinder and the theoretical prediction of the yield stress, figure adapted from reference [24].

applied to a body when a "P" pressure is applied in a normal direction which is equal to half the pressure.

(d) At some point along the height of the cylinder, the test paste experiences a tension that is greater than the creep limit and the paste flows until the tension is again reduced to the creep limit and remains motionless.

(e) In the process of falling, it is assumed that the interface layer between the paste that exhibits an elastic and inelastic behavior is a horizontal surface that moves down as the test paste flows below it.

(f) The flow occurs until the cross section increases the area so that the effort required to support the weight of the paste above any given plane is reduced to the creep limit. Thus, the product of the tension and the cross section is proportional to the weight of the paste on the plane.

The manufacturing material is acrylic to ensure a homogeneous and complete filling of the geometry. To achieve good video quality, the geometry must have a contrasting base and background surface, in addition to good lighting.

2.2. Designs of Flume Test Equipment. The flume test, used to determine the angle of repose in mineral pastes [25], was developed considering the characteristics of the paste under study. To determine the angle of repose of the paste, equation (2) is used. A schematic diagram is presented in Figure 3.

$$\theta_1 = \tan^{-1}\left(\frac{H_1 - H_2}{L}\right). \quad (2)$$

The equipment is built in acrylic to observe the profile of the paste over time, and the geometry used is rectangular and similar to a flume with a feeding chamber with a sliding door. The representation of the equipment is shown in Figure 4. To achieve good video quality, the geometry must have a contrasting base surface and background lighting.

2.3. Designs of Technique for Video Monitoring of Pastes. For analyzing results and image processing and recording tools, Matlab® Toolbox Image Processing and Free Video to Jpg. Converter software were used. For moving pastes, a digital video camera with 720 × 480 video resolution and color depth of 8 bits per pixel (256 colors) is used. The videos are recorded in.wmv format. The videos captured in the camera are edited with the software Camtasia Studio 6 to remove the time slots that are not considered in the image processing.

The original set of captured images (video) is separated into 30 frames per second with Free Video to JPG Converter software. These individual images or frames correspond to a series of points (pixels) capable of storing a color intensity ranging from 0 to 255, using the RGB color model (red, green, and blue); thus, the extensive combination of intensities that each pixel can store is the product of the mixture of primary colors at different intensities, with white being a mixture of maximum intensity (256, 256, 256) and black a combination of minimum intensity (0, 0, 0).

From the original image that can be seen in Figure 5, the black and white image shown in Figure 6 is extracted, which allows defining the contour surface and making measurements in the tests. In addition, Toolbox Image Processing allows defining the contour surface of the test in white or black so that from now on the combination of black test surface will be used for the results images.

When working with the RGB color model and using the image processing software tools contained in Matlab's Toolbox Image Processing, it is possible to separate only one color intensity in the frame (red, green, or blue), depending on the contrast surface used in the experience and that better reflects the contour surface for a treatment of the image in binary colors (white or black).

2.4. Description of Operational Parameters and Procedures. Powdered quartz is used, with an approximate granulometry of 100% under 400 Tyler meshes (<33 mm). The selection of

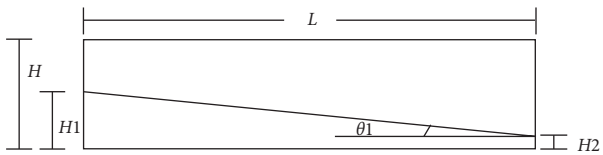


Figure 3: Relation schematic diagram of the flume test, where θ_1 = angle of repose, H = Initial height (cm), H_1 = highest height of the paste (cm), H_2 = lowest height of the paste (cm), and L = travel length (cm).

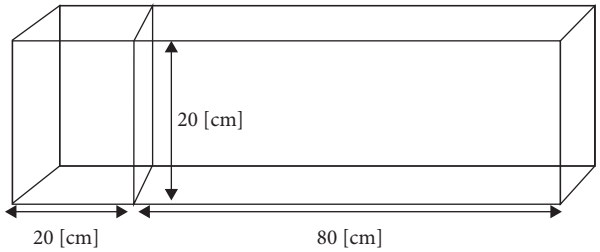


Figure 4: Geometrical characteristics for the flume test.

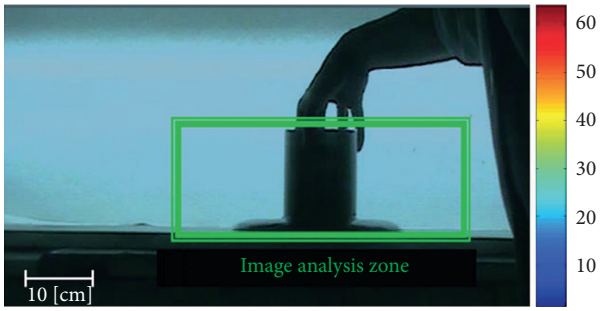


Figure 5: Example of the original image of pastes in movement.

the reagents used in the experimentation is based on the effect produced by a polymer and a pH modifier on the paste, considering quantifying the effect that these additives produce on the pastes at different concentrations of solids, the polymer in 75% solids and the pH modifier in 72% solids. In the case of the polymer, 1 ppm of a polyacrylamide was used with a cationic derivative of acrylic acid, Clarisol, and in the case of the pH modifier, 1 ppm of sulfuric acid was used.

For the slump test of 70% solids by weight, the test paste is prepared without adding sulfuric acid, the video camera is installed and configured, and the pH is measured. Then, the recording begins and the paste is emptied into the geometry of the cylinder. If the paste contains a large quantity of solids, the filling is done in two steps with compaction of the paste to eliminate the dead zones; otherwise, the filling is performed in only one step. Once the entire surface of the geometry is completely filled, the cylinder is lifted, the paste is allowed to flow and the recording stopped in the stationary state of the paste. The paste is then collected and concentrated sulfuric acid (98%) is added in a ratio of 0.0018 g of sulfuric acid/ton of dry solid, the pH is measured, and the same specifications described in the first paragraph of Section2.3 are followed from the paragraph where the recording begins. The experiences are repeated twice and one video is selected for image processing.

In the case of the slump test for the paste of 75% solids by weight, the paste is prepared without adding cationic polymer, Clarisol; then, the same procedure for the paste with 70% solids is followed and only the additive used as a rheology modifier changes, using a cationic polyacrylamide concentrated at 10% w/w at a ratio of 0.001 g of Clarisol/ton dry solid.

In the case of the flume test for 75% solids by weight, the paste is prepared without adding flocculant, Clarisol; then, the same procedure for the paste with 70% solids is followed and only the additive used as a rheology modifier is changed, this time using a concentrated cationic polyacrylamide at 10% w/w in a ratio of 0.001 g of Clarisol/ton dry solid.

The characteristics that the video camera must fulfill are to remain in a fixed position, parallel to the test plane, and at a distance (L) that covers the entire surface of the geometry. In addition, it must have manual brightness adjustment so that there are no automatic adjustments during video recording and image recording quality must be 720×480 pixels.

3. Results and Discussion

3.1. Effect of Sulfuric Acid with 70% Solid in Slump Test. The image capture was made in time intervals of 1/30 s without sulfuric acid, with a total duration of 1.97 s. The final ratio between the width of the cylinder and the spread of the paste was 2.76. With the use of 0.0018 g of sulfuric acid/ton dry solid and with a total duration of 1.97 s, the final ratio between the width of the cylinder and the spread of the sludge was 3.26.

When measuring the spread length of the paste with and without the addition of sulfuric acid, it can be observed that a change in the spread velocity of the paste occurs in both tests, being higher in the case with sulfuric acid as presented in Tables 1 and 2 and Figure 7. The graph shows that the addition of acid achieves higher spread velocities at the beginning of the test and that at the end of both tests the velocities obtained were similar.

3.2. Effect of Sulfuric Acid with 70% Solid in Flume Test. For the first case without the use of sulfuric acid, the image capture was made at 1/30 s time intervals with a total duration of 5.37 s, for which the angle of repose of the paste decreases in 2.3°. For the test with 0.0018 g of sulfuric acid/ton dry solid, the image capture was performed in time intervals 1/30 s, with a total duration of 1.97 s the angle of repose of the paste decreases 3.47°.

When measuring the length and height of the spread of the paste with and without the addition of sulfuric acid, the angle of repose is calculated to quantify the variation of the angle of repose in time, showing lower values in the paste sequence 70% solid with acid sulfuric. This parameter on an industrial scale means that the solid with these

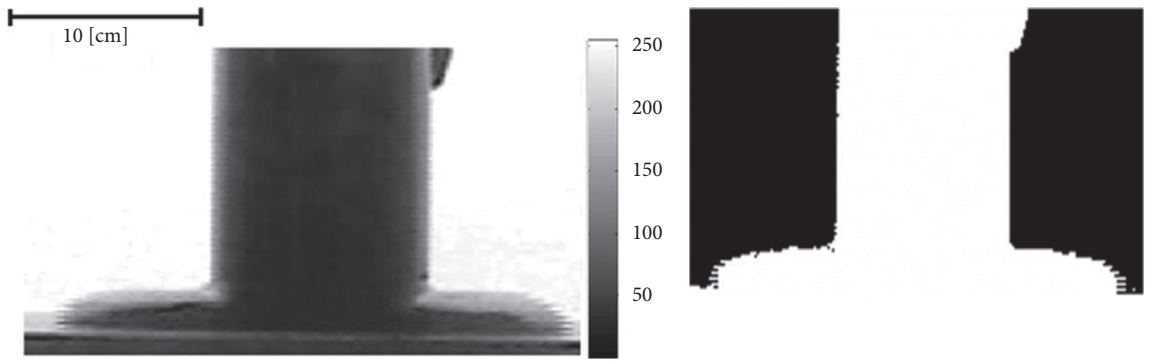


Figure 6: Example of binary image: (a) original image and (b) binary image.

Table 1: Slump test without sulfuric acid, cylinder.

Length (cm)	Real length (cm)	Ratio L/C	Time (s)	Real length increase	Spread velocity (cm/s)
4.74	20.00	1.00	0	0	0
5.50	23.21	1.16	0.13	3.21	24.05
7.30	30.80	1.54	0.30	10.80	36.01
8.15	34.39	1.72	0.47	14.39	30.83
9.40	39.66	1.98	0.63	19.66	31.05
10.10	42.62	2.13	0.80	22.62	28.27
10.98	46.33	2.32	0.97	26.33	27.24
11.36	47.93	2.40	1.13	27.93	24.65
11.80	49.79	2.49	1.30	29.79	22.91
12.40	52.32	2.62	1.47	32.32	22.04
12.62	53.25	2.66	1.63	33.25	20.36
12.80	54.01	2.70	1.80	34.01	18.89
13.10	55.27	2.76	1.97	35.27	17.94

Table 2: Slump test with sulfuric acid, cylinder.

Length (cm)	Real length (cm)	Ratio L/C	Time (s)	Real length increase	Spread velocity (cm/s)
4.53	20.00	1.00	0	0	0
7.28	32.14	1.61	0.13	12.14	91.06
11.35	50.11	2.51	0.30	30.11	100.37
12.90	56.95	2.85	0.47	36.95	79.19
14.15	62.47	3.12	0.63	42.47	67.06
14.24	62.87	3.14	0.80	42.87	53.59
14.38	63.49	3.17	0.97	43.49	44.99
14.62	64.55	3.23	1.13	44.55	39.31
14.65	64.68	3.23	1.30	44.68	34.37
14.76	65.17	3.26	1.47	45.17	30.79
14.76	65.17	3.26	1.63	45.17	27.65
14.76	65.17	3.26	1.80	45.17	25.09
14.76	65.17	3.26	1.97	45.17	22.97

characteristics needs less energy to be transported. Graphically the results are reflected in Figure 8 and the data is in Tables 3 and 4. From Figure 8, it can be seen that the flume test of 70% solids with the addition of sulfuric acid causes a decrease in the angle of repose of the paste and shorter spread time.

3.3. Effect of Cationic Polymer Clarisol with 75% Solids in Slump Test. The slump test with 75% solids does not present variations in height, as shown in Figure 9, which represents the initial, intermediate, and final images of the test at times 0, 1.07 and 1.63 s, respectively. It us worth mentioning that without the addition of cationic polymer; the high

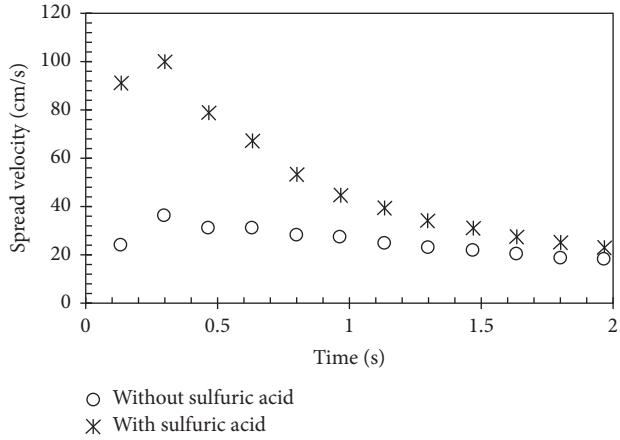


Figure 7: Graph of spread velocity vs. time, slump test.

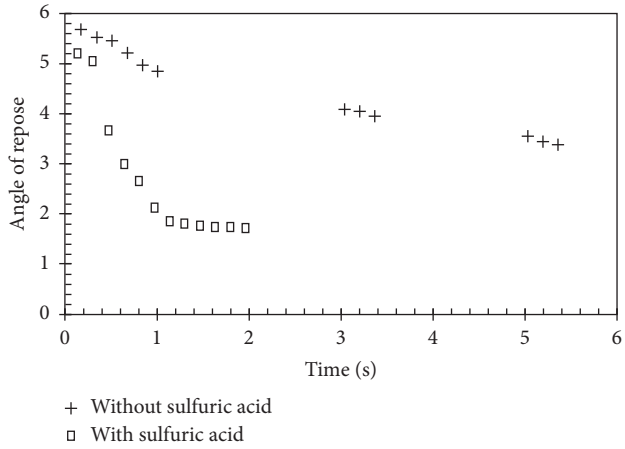


Figure 8: Graph of angle of repose vs. time, flume test.

Table 3: Flume test without sulfuric acid, cylinder.

Height (cm)	Length (cm)	(H_1/L)	Degrees	Time (s)
—	—	—	—	0.00
0.91	9.10	0.10	5.71	0.17
0.88	9.10	0.10	5.52	0.33
0.88	9.20	0.10	5.46	0.50
0.86	9.40	0.09	5.23	0.67
0.83	9.50	0.09	4.99	0.83
0.81	9.50	0.09	4.87	1.00
0.81	11.30	0.07	4.10	3.03
0.81	11.45	0.07	4.05	3.20
0.79	11.45	0.07	3.95	3.37
0.75	12.00	0.06	3.58	5.03
0.73	12.05	0.06	3.47	5.20
0.72	12.10	0.06	3.41	5.37

concentration of solids would prevent the movement of the test paste. The sequence of images is performed at 1/30 time intervals s with a total duration of 1.63 s.

When adding Clarisol cationic polymer (0.001 g Clarisol/ton dry solid), agglomeration of particles is observed. This way, an increase in the apparent viscosity of the paste is

Table 4: Flume test with sulfuric acid, cylinder.

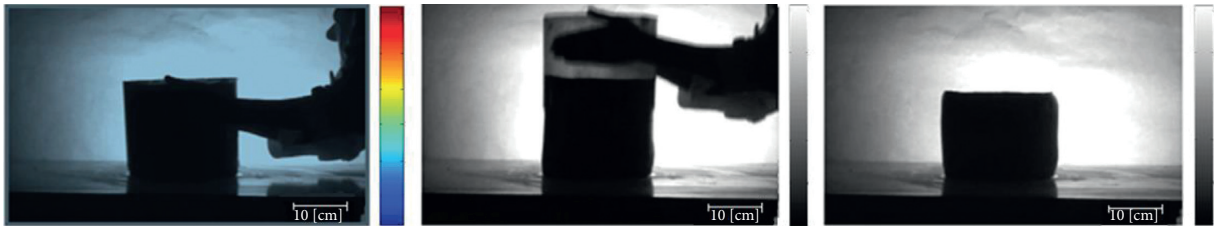
Height (cm)	Length (cm)	(H_1/L)	Degrees	Time (s)
—	—	—	—	0.00
0.53	5.83	0.09	5.19	0.13
0.53	6.00	0.09	5.05	0.30
0.55	8.57	0.06	3.67	0.47
0.53	10.10	0.05	3.00	0.63
0.52	11.18	0.05	2.66	0.80
0.46	12.40	0.04	2.12	0.97
0.43	13.30	0.03	1.85	1.13
0.43	13.67	0.03	1.80	1.30
0.42	13.67	0.03	1.76	1.47
0.42	13.67	0.03	1.76	1.63
0.42	13.67	0.03	1.76	1.80
0.41	13.67	0.03	1.72	1.97

inferred, as shown in Figure 10, which represents the initial, intermediate, and final image of the test, at time 0, 1.43 and 2.87 s, respectively. The image sequence is performed at 1/30 time intervals (s) with a total duration of 2.87 s.

For both tests, with and without addition of polymer, no changes in the fluidity of the paste occur. Another observation made when comparing the tests is the increase in the time it takes to remove the cylinder during the experiment, from 1.63 s without additive to 2.87 s with additive, so the effect that the polymer produces on the test paste would not be favorable for the objectives set out in this investigation.

3.4. Effect of Clarisol Cationic Polymer with 75% Solids in Flume Test. The flume test with 75% solids and without added Clarisol does not present variations in height and length. This is reflected in Figure 11, which represents the initial, intermediate, and final image of the test at times 0, 1.90, and 3.83 s, respectively. The sequence of images is performed at 1/30 time intervals (s) with a total duration of 3.83 s.

In the case of the flume test with 75% solids and with 0.001 g Clarisol/ton dry solid, there are no variations in height and length, which is reflected in Figure 12, which represents the initial, intermediate, and final image of the

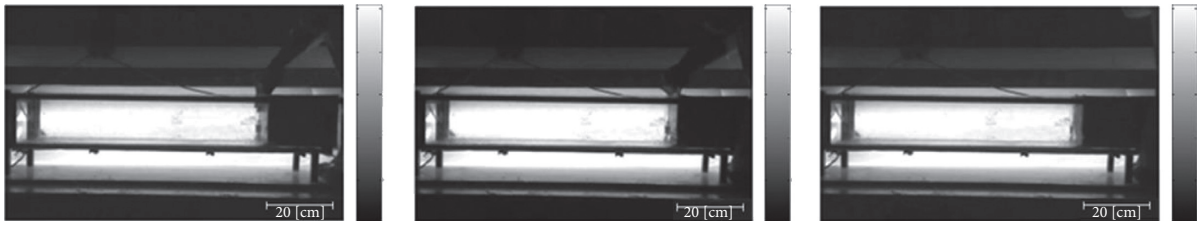


(a) (b) (c)

Figure 9: Dynamic monitoring with 75% solids, slump test without cationic polymer. (a) $t = 0$ s. (b) $t = 1.07$ s. (c) $t = 1.63$ s.

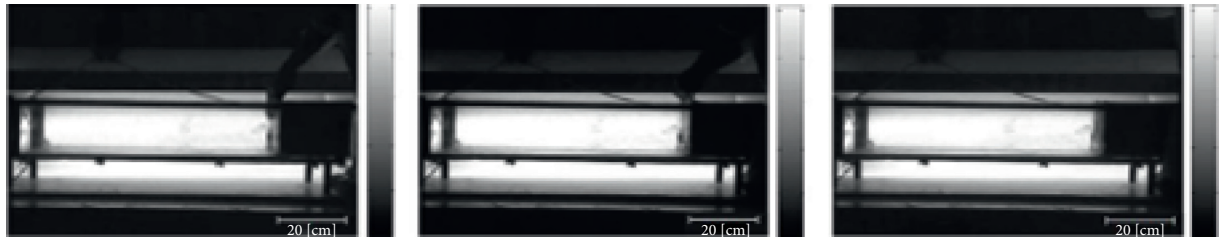
(a) (b) (c)

Figure 10: Dynamic monitoring with 75% solids, slump test with cationic polymer. (a) $t = 0$ s. (b) $t = 1.43$ s. (c) $t = 2.87$ s.



(a) (b) (c)

Figure 11: Dynamic monitoring with 75% solids, flume test without cationic polymer. (a) $t = 0$ s. (b) $t = 1.9$ s. (c) $t = 3.83$ s.



(a) (b) (c)

Figure 12: Dynamic monitoring with 75% solids, flume test with cationic polymer. (a) $t = 0$ s. (b) $t = 1.9$ s. (c) $t = 3.83$ s.

test at times 0, 2.43 and 4.87 s, respectively. The sequence of images is performed at 1/30 time intervals (s) with a total duration of 4.87 s. In summary, for the flume tests with 75% solids, with or without addition of cationic polymer Clarisol, no changes in the fluidity of the paste are produced. The time it takes to remove the sliding gate from the flume increases from 3.83 s without additive to 4.87 s with additive.

4. Conclusions

4.1. Software Utilized. For analyzing results and image processing and recording tools, Matlab® Toolbox Image Processing and Free Video to Jpg.Converter software were used. The videos captured in the camera were edited with the software Camtasia Studio 6 to remove the time slots that are

not considered in the image processing. In addition, Toolbox Image Processing allows defining the contour surface of the test in white or black so that the combination of black test surface is used for the results images. It us recommended to use the tools and techniques of Matlab® Toolbox Image Processing to analyze, quantify, and compare the effect of the additives over time.

4.2. Novelty and Advantages of the Technique. A technique based on digital video and image processing was developed through the implementation of a laboratory for the characterization of pastes including slump and flume tests. The relevance of implementing image analysis by video should be noted, since it allows to quantify the dynamic behavior in pastes. Regarding the type of rheology modifier used, it is concluded that sulfuric acid meets the objectives, allowing to recover water by increasing fluidity. It is worth mentioning that the use of this technology generates significant savings in water resources compared to current processes, which not only means an economic benefit, but an important step in favor of the sustainable development that is so much needed in these times.

4.3. Limitations. On the contrary, it was not possible to analyze the effect of the polymeric and pH modifier in tailings pastes, and it was possible to analyze and quantify the effect of a chemical additive as a rheology modifier in pastes of fine granulometry (<38 μm) and basic pH.

4.4. Future Work Ahead. It is recommended to modify the dimensions of the flume, specifically the length through which the test paste flows, since the dimensions considered do not allow to evaluate the effect that the addition of more acid produces on the paste. To observe and quantify the behavior of the test pastes throughout time and using image processing, it is recommended to work with percentages of solids comprised between 70% and 74% in the cylinder and flume test.

Data Availability

All data supporting this study are provided as supplementary information accompanying this paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Claudio Acuña conceptualized the study; Sofía Mondaca carried out data curation; Claudio Leiva, Claudio Acuña, and Sofía Mondaca carried out formal analysis; Sofía Mondaca carried out investigation; Claudio Acuña and Sofía Mondaca provided the methodology; Claudio Acuña carried out validation; Eduardo Serey wrote the original draft; Claudio Leiva and Eduardo Serey wrote the review and edited the article.

Acknowledgments

The authors wish to acknowledge the material support provided by the Gabriela Mistral Division, Codelco, financial support provided by Universidad Católica del Norte, and to the student Sofía Mondaca for his contribution with his chemical engineer thesis.

References

[1] S. Herrera-León, F. A. Lucay, L. A. Cisternas, and A. Kraslawski, "Applying a multi-objective optimization approach in designing water supply systems for mining industries. The case of Chile," *Journal of Cleaner Production*, vol. 210, pp. 994–1004, 2019.

[2] G. M. Mudd, *Mining and Water Resources, Reference Module in Earth Systems and Environmental Sciences*, Elsevier, Amsterdam, Netherlands, 2020.

[3] J. Huazhe, W. Shufei, Y. Yixuan, and C. Xinming, "Water recovery improvement by shearing of gravity-thickened tailings for cemented paste backfill," *Journal of Cleaner Production*, vol. 245, 2020.

[4] COCHILCO, *Good Practices and Efficient Use of Water in the Mining Industry*, INEH publications, Santiago, Chile, 2008.

[5] J. Johnson and W. Xie, "Paste and thickened tailings water benefits–case studies," in *Proceedings of the 20th International Seminar on Paste and Thickened Tailings*, A. Wu and R. Jewell, Eds., University of Science and Technology Beijing, Beijing, China, pp. 387–392, 2017.

[6] P. Slatter, "The role of rheology in the pipelining of mineral slurries," *Mineral Processing and Extractive Metallurgy Review*, vol. 20, pp. 281–300, 1999.

[7] E. I. Robinsky, "Tailings dam failures need not be disasters-the Thickened Tailings Disposal (TTD) system," *CIM Bulletin*, vol. 92, pp. 140–142, 1999.

[8] D. Landriault, J. M. Johnson, and F. Palkovits, "Thickened tailings and paste technology: the future of industrial waste disposal," *SME Annual Meeting: Got Mining-Preprints*, vol. 34, pp. 899–908, 2005.

[9] W. Cincilla, D. Landriault, and R. Verburg, "Application of paste technology to surface disposal of mineral wastes," in *Proceedings of the Fourth International Conference on Tailings and Mine Waste*, Fort Collins, CO, USA, 1997.

[10] J. Rayo, R. Fuentes, and R. Orellana, "Large tailings disposal-conventional versus paste," in *Proceedings of the Twelfth International Seminar on Paste and Thickened Tailings*, R. Jewell, A. B. Fourie, S. Barrera, and J. Wiertz, Eds., Australian Centre for Geomechanics, Perth, Australia, pp. 271–278, 2009.

[11] C. Bazin and C. B-Chapleau, "The difficulty associated with measuring slurry rheological properties and linking them to grinding mill performance," *International Journal of Mineral Processing*, vol. 76, no. 1-2, pp. 93–99, 2005.

[12] S. B. Johnson, G. V. Franks, P. J. Scales, D. V. Boger, and T. W. Healy, "Surface chemistry–rheology relationships in concentrated mineral suspensions," *International Journal of Mineral Processing*, vol. 58, no. 1-4, pp. 267–304, 2000.

[13] Q. D. Nguyen and D. V. Boger, "Application of rheology to solving tailings disposal problems," *International Journal of Mineral Processing*, vol. 54, no. 3-4, pp. 217–233, 1998.

[14] L. Spinosa and V. Lotito, "A simple method for evaluating sludge yield stress," *Advances in Environmental Research*, vol. 7, no. 3, pp. 655–659, 2003.

[15] D. Wasan, A. Nikolov, and B. Moudgil, "Colloidal dispersions: structure, stability and geometric confinement," *Powder Technology*, vol. 153, no. 3, pp. 135–141, 2005.
[16] V. Flores, B. Keith, and C. A. Leiva, "Using artificial intelligence techniques to improve the prediction of copper recovery by leaching," *Journal of Sensors*, vol. 34, 2020.
[17] L. G. Bergh, J. B. Yianatos, and C. A. Leiva, "Fuzzy supervisory control of flotation columns," *Minerals Engineering*, vol. 11, no. 8, pp. 739–748, 1998.
[18] C. A. Leiva, V. M. Flores, F. Salgado, D. Poblete, and C. Acuña, "Applying softcomputing for copper recovery in leaching process," *Science Program*, vol. 2017, p. 6, 2017.
[19] C. Leiva, K. Arcos, D. Poblete, E. Serey, C. Torres, and Y. Ghorbani, "Design and evaluation of an expert system in a crushing plant," *Minerals*, vol. 8, no. 10, p. 469, 2018.
[20] M. A. Molina, C. A. Acuña, and C. A. Leiva, "Characterization of filamentous flocs to predict sedimentation parameters using image analysis," *Journal of Sensors*, vol. 34, 2020.
[21] B. Klein and M. Pawlik, "Rheology modifiers for mineral suspensions," *Mining, Metallurgy & Exploration*, vol. 22, no. 2, pp. 83–88, 2005.
[22] M. Katzer, R. Klimpel, and J. Sewell, "Example of the laboratory characterization of grinding aids in the wet grinding of ores," *Minning Engineering*, vol. 33, pp. 1471–1476, 1982.
[23] D. B. Braun and M. R. Rosen, "Part 1-practical rheology," in *Rheology Modifiers Handbook*, pp. 1–69, William Andrew Publishing, Norwich, NY, USA, 1999.
[24] S. Clayton, T. G. Grice, and D. V. Boger, "Analysis of the slump test for on-site yield stress measurement of mineral suspensions," *International Journal of Mineral Processing*, vol. 70, no. 1-4, pp. 3–21, 2003.
[25] M. Kwak, D. F. James, and K. A. Klein, "Flow behaviour of tailings paste for surface disposal," *International Journal of Mineral Processing*, vol. 77, no. 3, pp. 139–153, 2005.

Hindawi
Scientific Programming
Volume 2020, Article ID 9092829, 20 pages
https://doi.org/10.1155/2020/9092829

Research Article

AQUACOAST: A Simulation Tool to Explore Coastal Groundwater and Irrigation Farming Interactions

Jaime Martínez-Valderrama,[1] **Javier Ibáñez**,[2] **and Francisco J. Alcalá**[3,4]

[1]*Department of Ecology & Multidisciplinary Institute for Environment Studies "Ramón Margalef", University of Alicante, 03690 San Vicente del Raspeig, Alicante, Spain*
[2]*Departamento de Economía Agraria, Estadística y Gestión de Empresas, Universidad Politécnica de Madrid, 28040 Madrid, Spain*
[3]*Instituto Geológico y Minero de España (IGME), 28003 Madrid, Spain*
[4]*Instituto de Ciencias Químicas Aplicadas, Facultad de Ingeniería, Universidad Autónoma de Chile, 7500138 Santiago, Chile*

Correspondence should be addressed to Jaime Martínez-Valderrama; jaime.mv@ua.es

Received 10 January 2020; Accepted 28 April 2020; Published 14 May 2020

Academic Editor: Cristian Mateos

In the framework of coastal groundwater-dependent irrigation agriculture, modelling becomes indispensable to know how this renewable resource responds to complex (usually not conceptualized nor monitored) biophysical, social, and economic interactions. Friendly user interfaces are essential to involve nonmodeling experts in exploiting and improving models. Decision support systems (DSS) are software systems that integrate models, databases, or other decision aids and package them in a way that decision makers can use. This paper addresses these two issues: firstly with the implementation of a System Dynamics (SD) model in Vensim software that considers the integration of hydrological, agronomic, and economic drivers and secondly with the design of a Venapp, push-button interfaces that allow users access to a Vensim model without going through the Vensim modelling environment. The prototype designed, the AQUACOAST tool, gives an idea of the possibilities of this type of models to identify and analyze the impact of apparently unrelated factors such as the prices of cultivated products, subsidies or exploitation costs on the advance of saltwater intrusion, and the great threat to coastal groundwater-dependent irrigation agriculture systems.

1. Introduction

A large proportion of the world's population depends on groundwater resources [1]. Groundwater accounts for as much as 33% of the global freshwater demand and is increasing [2]. This resource is especially crucial in regions where the surface water network does not cover a fraction of the freshwater demand at least, such as in many coastal drylands [3]. In these areas, aquifers play a critical role in sustaining the economy and the environment [4], but water scarcity has led to high groundwater abstraction rates to supply the increasing urban, tourism, industrial, and agriculture demands, thus adding stress to groundwater bodies [5, 6]. The cases with alarming signs of groundwater quantity depletion and quality degradation do not stop increasing [7].

Groundwater degradation in coastal drylands is often the result of complex interactions between global driving forces and human-induced actions [8, 9]. Global climate scenarios forecast high evaporation rates and lower high-salinity aquifer recharge rates [10]. Some human-induced actions may contribute to this initial "climatic" groundwater salinity, such as pumping inducing saltwater intrusion and mobilisation of brines [4, 11], high-efficiency irrigation systems producing high-salinity, low irrigation return rates [12], and misuse of fertilizers and agrochemicals, among others. For instance, the resulting groundwater salinity in many southern Europe coastal drylands exceed the standards of quality that the European Drinking Water Directive [13] recommends for human consumption and health, crop production, and survival of the ecosystems biodiversity [12].

In this context, agriculture has invariably had the highest freshwater demand [14, 15], contributing at growing of local societies, but accounting for 70% of withdrawals [16, 17]. The exposition of coastal aquifers to saltwater intrusion

(SWI) after disruption of natural conditions [18, 19] had reduced crop surface and therefore profit in many regions, with dramatic consequences for food production and social inequality. Users often forget that seawater is another water balance component. An unplanned exploitation is the cause behind the alteration of the fragile freshwater-saltwater equilibrium with typical consequences such as freshwater-saltwater mixing zone displacement inland and reduction of crop production. These anticipated negative indicators may end with abandonment of land devoted to irrigated crops.

Understanding coastal groundwater and irrigation farming interactions is a challenge because the equilibrium between sustainable pumping and agricultural profit is subjected to a fragile threshold determined by the crop-limiting salinity. This paper aims to contribute to this topic by means of a user-friendly programming tool that combines hydrological, agronomic, and economic components of the system [20, 21]. This multidisciplinary tool must deal with the complex network of variables that conform feedbacks loops subjected to delays and nonlinear relationships. For this, SD modelling is a very suitable approach. When the scope of the study goes beyond the academic field and aims to involve stakeholders and land-use planners, it is essential to link these models to DSS. These are software systems that integrate models, databases, or other decision aids and package them in a way that decision makers can use [22].

The perception of the complex interaction between the sociocultural, economic, and biophysical components of the systems is the cause of the growing interest in environmental DSSs [22–24]. This paper presents a Vensim application (Venapp) for the AQUACOAST simulation model, which aims to evaluate the sustainability of coastal groundwater-dependent agriculture. As described, the great threat to coastal aquifers is saltwater intrusion triggered by excessive freshwater pumping. Analyzing under what scenarios such a disaster may occur is the purpose of this simulation tool.

2. Materials and Methods

2.1. System Dynamics Modelling. SD modelling [25] is an ideal framework to integrate variables coming from different disciplines because it is based on a holistic view of the system under study and considers the model structure as a whole. SD modelling embodies system thinking through the implementation of differential equation systems. It is grounded in the theory of nonlinear dynamics and feedback control developed in mathematics, physics, and engineering.

A SD model is a stock-and-flow structure of first-order differential equations, which is commonly deterministic and used to generate the time trajectories of the model variables under different simulation scenarios. The model structure is made up of causal feedback loops, which include nonlinear relationships and delays, and constitute a holistic and easily overlooked because of its behavior [26].

Strong points of SD may be summarized as follows [27]: (i) SD models are useful learning tools that help improve system understanding and foster system thinking skills and knowledge integration for modelers and end users. This is true even for their first stage of development, i.e., as conceptual models. (ii) Advances in the development of high-level dynamic modelling software platforms have made computational SD modelling more accessible. (iii) SD literature has made rich contributions to approaches informing on the modelling process, including data collection methods, knowledge elicitation/mapping techniques, and policy analysis.

The integrated models that allow the implementation of SD are essential to address environmental problems [28–30]. The degradation of the environment is forged by the imbalance created between the natural rate of regeneration of resources and their rate of exploitation, which is due to socioeconomic factors. In order to understand how economic systems are decoupled from natural ones and to seek sustainable solutions, it is essential to analyze the dynamic interactions between economic and natural biophysical systems [21, 31].

Advances in the development of high level dynamic modelling software platforms, such as iThink (isee systems, http://www.iseesystems.com), Vensim (Ventana Systems, http://www.vensim.com), and Powersim Studio (Powersim Software AS, http://www.powersim.com), have rendered computer SD models very accessible (even with minimal technical background). These applications are often designed as communication layers: user interface, stock-flow, mathematical equations, and programming code [27].

The important feature of these modelling tools is that they allow the user to write a model, even a complex one, without writing a single equation. The advantage is that the model diagram is focused on the relationships among the different variables and processes in the system and can be understood by a person with a good background in the problem domain of the model, but with little knowledge of mathematics. Models are designed "visually" by connecting blocks and aggregating them at different levels of resolution [22] what makes these tools very useful in the model prototyping phase and for educational purposes.

From all available options, we chose Vensim [32]. Although the software lacks certain possibilities, for example, does not support spatial treatment and its graphical interface is improvable [33], it is very suitable for the purpose we aim, to understand the typical functioning complexity of a coastal groundwater-dependent agricultural system subjected to different water-competing sectors and land-use policies. Complexity departs from the lack of data in many groundwater systems [34], thus complicating numerical modelling. There may be unknown essential information, such as the actual groundwater stored in an aquifer. In addition, cause and effect are separated in time and space which adds difficulty in conceptualizing basic hydrological processes. For instance, aquifer recharge may take place tens to hundreds of kilometers from the discharge area and may occur months to years before this water is pumped [4, 8].

2.2. Vensim Applications: Venapps. Interaction with simulation models has improved considerably in recent years. However, it can still be difficult for a nonexpert user to deal directly with programs such as Vensim. Friendly user interfaces are essential to involve nonmodeling experts in exploiting and improving models. The VensimDSS® software

[35] includes the possibility of developing simple Venapps to visualize the model structure, examine its causality, simulate the model behavior, and visualize its results [36–38].

Venapps are simplified, push-button interfaces that allow users access to a Vensim model without going through the Vensim modelling environment. A Venapp uses a model and a set of rules for interacting with the model to give users simplified access to that model. To the user, the Venapp appears as a series of buttons, menus, or a sequence of screens allowing him or her to use and analyze the model in straightforward and meaningful way.

The goal of these interactive simulators is to provide the user with the following tasks: examine the model structure, establish simulation scenarios, simulate the model, and analyze the impact of decisions on the behavior of the system under study. The VensimDSS environment includes an extension called Vensim application (Venapp^{TM}) which allows the creation of graphical user interfaces in the file format "vpa" (Vensim packaged application). Both model and Venapp can be run with the software "Vensim Model Reader" which can be downloaded from the Vensim® homepage for free (http://www.vensim.com/freedownload.html).

3. AQUACOAST Model

3.1. Overall Description. The AQUACOAST model implements the two-dimensional saltwater-freshwater interface (SFI) analytical solution proposed in [39] for an idealized unconfined (i.e., its upper boundary is the fluctuating water table) coastal aquifer, which is horizontal, enough thick, and overlying an impervious bedrock, isotropic, and hydraulically homogeneous with steady inflow (aquifer recharge) and outflow (aquifer discharge) rates (Figure 1). The variables and parameters of the AQUACOAST model concerning irrigation practices, groundwater hydrology, and groundwater allotment for agriculture have been added to the schematization of Verruijt [39]. In the steady schema of Verruijt [39], the SFI position is reached automatically depending on the fresh groundwater flow. AQUACOAST modifies this hypothesis by introducing delays in order to reach the equilibrium described by Verruijt, which magnitude differs, which differ for the intrusion phase (minor delay) and the freshening phase (longer delay) [40]. Groundwater allotment for agricultural use is the main flow component. The model focuses on determining critical distances of agronomic interest from the shoreline, such as the current and equilibrium freshwater and salinization boundaries or the irrigation farming abandonment boundary (Figure 1).

AQUACOAST is a lumped spatial model [27], and thus, its variables represent either totals or averages over the entire area modelled. For the sake of illustration, the simulation period is 50 years, an appropriate time horizon when studying environmental degradation that can only be detected when observed the evolution of "slow variables" [41]. Variables and parameters use the SI units, capital letters being variables, whereas small letters denote parameters. Tables 1 and 2 include the notation, units, and adopted values for variables and parameters, respectively, after data from [4, 8, 20, 21, 40, 42, 43] and references therein. The AQUACOAST model uses 26 straightforward formulations divided into seven sections as follows.

3.2. Net Aquifer Discharge. One of the main hypotheses of the AQUACOAST model is that fresh groundwater flow moves towards the sea and discharges at the shoreline in natural regime. Net aquifer discharge (TDFW) is expressed as difference in inflow and outflow water balance components and is measured per unit thickness of the aquifer. In a disturbed regime, the model assumes that TDFW decreases due to groundwater pumping and the SFI displaces inland. The depth and extent of the SFI determine the fraction of usable fresh groundwater. The time-zero TDFW responds to initial pumping. This initial TDFW can be used to set other initial states in the AQUACOAST model. TDFW is expressed as

$$\mathrm{TDFW}_t = \frac{[\mathrm{appt}\cdot(1-\mathrm{cfro})-\mathrm{aaet}-\mathrm{PUMP}_t\cdot(1-\mathrm{cfrf})]\cdot\mathrm{lgth}}{365}+\frac{(\mathrm{lttf}\cdot 10^3)}{(\mathrm{wdth}\cdot 365)}, \tag{1a}$$

$$\mathrm{TDFW}_0 = \frac{[\mathrm{appt}\cdot(1-\mathrm{cfro})-\mathrm{aaet}-\mathrm{PUMP}_0\cdot(1-\mathrm{cfrf})]\cdot\mathrm{lgth}}{365}+\frac{(\mathrm{lttf}\cdot 10^3)}{(\mathrm{wdth}\cdot 365)}, \tag{1b}$$

where TDFW = net aquifer discharge, PUMP = groundwater pumping, appt = annual precipitation, cfro = runoff coefficient, aaet = annual actual evapotranspiration, cfrf = return flow coefficient, lgth = coastal aquifer length, lttf = lateral groundwater transference from others aquifers, wdth = coastal aquifer width, t = state of the system at a given time, and 0 = time zero.

3.3. Saltwater and Freshwater Boundaries Dynamics. Verruijt [39] arrives at the formulae giving the depth of the SFI below sea level (h_S) and the height of the free groundwater surface (FGS: water table) (h_F) as functions of the shoreline distance along x-axis, TDFW, and three constants: hydraulic conductivity (cfpm), freshwater density (dtfw), and saltwater density (dtsw). Note that x, h_S, and h_F,

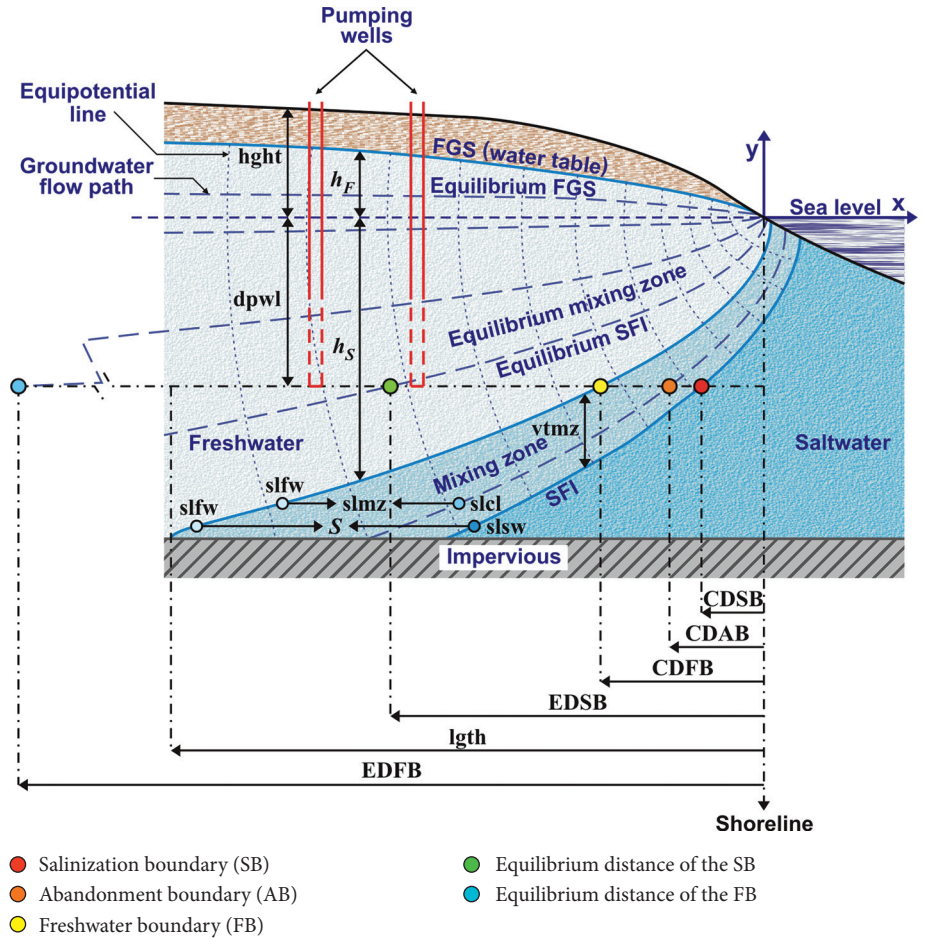


Figure 1: The AQUACOAST model added to Verruijt's schematization for an idealized unconfined coastal aquifer. h_F is the height of the free groundwater surface (FGS: water table) above sea level (it is not a model parameter), h_S is the depth of the saltwater-freshwater interface (SFI) below sea level (it is not a model parameter), hght is the average height of land surface above sea level, vtmz is the vertical thickness of the mixing zone, dpwl is the depth of pumping wells below sea level, CDSB is the current distance of the SB from the shoreline, CDAB is the current distance of the AB from the shoreline, CDFB is the current distance of the FB from the shoreline, EDSB is the equilibrium distance of the SB from the shoreline, EDFB is the equilibrium distance of the FB from the shoreline, lgth is the coastal aquifer length, slfw is the freshwater salinity, S is the mixing water salinity along the x-axis distance from the shoreline (CDSB $\leq x \leq$ CDFB), slmz is the actual average salinity across a line segment connecting FB and AB (CDAB $\leq x \leq$ CDFB), slcl is the crop-limiting salinity, and slsw is the saltwater salinity.

Table 1: Notation for variables used.

Variable	Equation	Definition	Units[1]
ARIR	15, 26	Current irrigated area	$m^{-2}\cdot day^{-1}$
ARPI	11, 13	Potential irrigated area	$m^3\cdot ha^{-1}\cdot year^{-1}$
ARTI	13, 15	Target for the irrigated area	$€\cdot year^{-1}$
CDAB	8, 11, 14, 19	Current distance of the AB from the shoreline	$m^3\cdot ha^{-1}\cdot year^{-1}$
CDFB	5, 7, 8, 19	Current distance of the FB from the shoreline	$€\cdot ha^{-1}\cdot year^{-1}$
CDSB	3, 6, 7, 8, 14	Current distance of the SB from the shoreline	$€\cdot m^{-3}$
CHFW	10, 20	Current average height of the FGS above sea level	m
CSTW	20, 21, 24	Water cost per cubic meter	m
EDFB	4, 5	Equilibrium distance of the FB from the shoreline	$m^3\cdot ha^{-1}\cdot year^{-1}$
EDSB	2, 3, 6, 14	Equilibrium distance of the SB from the shoreline	$m^3\cdot ha^{-1}\cdot year^{-1}$
EHFW	9, 10	Equilibrium average height of the FGS above sea level	km
EXPF	13, 22	Expected profit per irrigated farm	km
EXWX	23, 25	Expected water requirement for profit maximization	km
IRDS	17, 21, 23, 26	Irrigation dose from groundwater	km
PFTH	21, 22	Profit per hectare	km
PPIR	16, 17, 23	Contribution of direct precipitation to irrigation	$year^{-1}$

Table 1: Continued.

Variable	Equation	Definition	Units[1]
PUMP	1, 26	Groundwater pumping	dmnl
RAIA	14, 15	Rate of abandonment of irrigated area	ha
S(x)	7, 18	Salinity of mixing water along the x-axis distance from the shoreline, CDSB $\leq x \leq$ CDFB	kg·m^{-3}
SLFT	17, 18, 19, 24	Salinity factor	ha
TDFW	1, 2, 4, 9	Net aquifer discharge	ha
TMBE	3, 5, 6, 14	Time for the SB and FB to achieve equilibrium	year
WRXP	24, 25	Water requirement for profit maximization	mm·year^{-1}
YLDH	17, 21	Crop yield per surface unit	kg·ha^{-1}·year^{-1}

1 – dmnl = dimensionless; ha = hectare; € = EU Euros.

Table 2: Notation for parameters used.

Parameter	Equation	Definition	Units[1]	Value[2]
aaet	1	Annual actual evapotranspiration	mm·year^{-1}	200
appt	1, 16	Annual precipitation	mm·year^{-1}	300
arfr	15, 22	Area of one farm	ha	2.84
cdap	7, 8, 19	Calibration parameter to represent nonlinear variations in salinity	dmnl	3
cfpm	2, 4, 9	Hydraulic conductivity	m·day^{-1}	10
cfrf	1	Return flow coefficient	dmnl	0.2
cfro	1, 16	Runoff coefficient	dmnl	0.2
cwot	20	Costs of water other than energy	€·m^{-3}	1.4
dpwl	2, 4	Depth of pumping wells below sea level	m	100
dtcp	2, 4, 9	Density contrast parameter	dmnl	0.025
dtfw	2, 4, 9	Freshwater density	kg·m^{-3}	1.000
dtsw	2, 4, 9	Saltwater density	kg·m^{-3}	1.025
efir	21, 24, 26	Irrigation system efficiency	dmnl	0.8
egcm	20	Energy required to pump one cubic meter one meter	kwh·m^{-4}	0.017
expf	22	Initial expected profit per irrigated farm	€·year^{-1}	23,000
hght	20	Average height of land surface above sea level	m	100
idqt	23	Irrigation quota	m^3·ha^{-1}·year^{-1}	20,000
lgth	1, 2, 4, 9, 11, 14, 19, 26	Coastal aquifer length	km	10
lttf	1	Lateral transference from other aquifers	hm^3·year^{-1}	7.9
oaag	16	If 1, then open-air agriculture else under plastic agriculture	dmnl	0
ocfm	13	Average opportunity cost for a farmer	€·yr^{-1}	23,000
ocfv	13	Coefficient of variation of the opportunity cost for a farmer	dmnl	0.29
octh	21	Other costs per hectare	€·ha^{-1}·year^{-1}	40,053
prcp	21, 24	Crop price	€·kg^{-1}	0.76
preg	20	Price of energy	€·kwh^{-1}	0.05
pxys	23	If 1, then farmers seek economic optimum else they seek agronomic optimum	dmnl	1
pycm	17, 24	Potential marginal yield per cubic meter of water	kg·m^{-3}	30
sfmz	19	Average salinity factor across a line segment connecting the FB and AB	kg·m^{-3}	imp
slcl	8, 18, 19	Crop-limiting salinity	kg·m^{-3}	2
slfp	18, 19	Crop-specific calibration parameter	kg·m^{-3}	0.5
slfw	7, 8, 19	Freshwater salinity	kg·m^{-3}	0.15
slmz	19	Actual average salinity across a line segment connecting FB and AB, CDAB $\leq x \leq$ CDFB	kg·m^{-3}	imp
slsw	7, 8	Saltwater salinity	kg·m^{-3}	35
subh	21	Subsidies per hectare	€·ha^{-1}·year^{-1}	0
tfex	15, 22, 25	Time for farmers to adjust expectations	year	3
tmbf	6	Time for current SB and FB to achieve equilibrium, freshening phase	year	10
tmbi	6	Time for current SB and FB to achieve equilibrium, intrusion phase	year	5
tmhe	10	Time for the FGS to achieve equilibrium	year	5
vtmz	4	Vertical thickness of the MZ	m	20
wdth	1, 11, 26	Coastal aquifer width	km	4
wrao	17, 23, 24, 25	Water requirement for agronomic optimum yield	m^3·ha^{-1}·year^{-1}	7500
subscript 0		Time zero, initial state of the system	year	unt
subscript s		Time step	year	0.0078125
subscript t		State of the system at a given time	year	unt

1 – dmnl = dimensionless; ha = hectare, 104 m2; € = EU Euros; 2 – imp = intermediate modelling parameter; unt = undefined time start or span.

as such, are not variables of the AQUACOAST model despite that they are included in Figure 1 for illustrative purposes.

As a convenient simplification, the model assumes that the depth of all the pumping wells below sea level (dpwl) is the same (Figure 1). In real aquifers, pumping wells may show substantial differences in depth. In this case, dpwl will adopt the depth of the deepest one or the depth of other shallowest ones as the SFI rises due to progressive pumping.

Let us call salinization boundary (SB) the line, in theory parallel to the shoreline for an isotropic and hydraulically homogeneous aquifer (whose projection in Figure 1 is a point) where the SFI intersects the horizontal plane that connects the bottom of pumping wells (whose projection in Figure 1 is a line). It may be thought that the distance of the SB from the shoreline could be obtained by equaling Verruijt's h_S to dpwl and solving for x. However, this would entail assuming that the SFI moves instantaneously, which is not the case in real aquifers. Therefore, it is assumed that the expression obtained as mentioned gives the equilibrium distance from the shoreline that the SB would reach given the TDFW (EDSB) as in Figure 1. This expression includes a MIN function that avoids EDSB to be greater than the coastal aquifer length, as follows:

$$\mathrm{EDSB_t} = \mathrm{MIN}\left\{\mathrm{lgth}, \left(\left[\frac{\mathrm{dpwl}^2 \cdot \mathrm{cfpm} \cdot \mathrm{dtcp} \cdot (1+\mathrm{dtcp})}{(2 \cdot \mathrm{TDFW_t})}\right] - \left[\frac{\mathrm{TDFW_t} \cdot (1-\mathrm{dtcp})}{(2 \cdot \mathrm{dtcp} \cdot \mathrm{cfpm})}\right]\right) \cdot 10^{-3}\right\}, \tag{2a}$$

$$\mathrm{EDSB_0} = \mathrm{MIN}\left\{\mathrm{lgth}, \left(\left[\frac{\mathrm{dpwl}^2 \cdot \mathrm{cfpm} \cdot \mathrm{dtcp} \cdot (1+\mathrm{dtcp})}{(2 \cdot \mathrm{TDFW_t})}\right] - \left[\frac{\mathrm{TDFW_0} \cdot (1-\mathrm{dtcp})}{(2 \cdot \mathrm{dtcp} \cdot \mathrm{cfpm})}\right]\right) \cdot 10^{-3}\right\}, \tag{2b}$$

$$\mathrm{dtcp} = \frac{(\mathrm{dtsw}-\mathrm{dtfw})}{\mathrm{dtfw}}, \tag{2c}$$

where EDSB = equilibrium distance of the SB from the shoreline, TDFW = net aquifer discharge, lgth = coastal aquifer length, dpwl = depth of pumping wells below sea level, cfpm = hydraulic conductivity, dtcp = density contrast parameter, dtsw = saltwater density, dtfw = freshwater density, t = state of the system at a given time, and 0 = time zero.

The AQUACOAST model assumes that the SB moves towards equilibrium following an exponential fitting. Thus, the current distance of the SB from the shoreline (CDSB) is given by a first-order linear negative feedback. As deduced, no pumping wells whose distance from the shoreline is less than CDSB take saltwater (Figure 1). It is assumed that, initially, the system is in the equilibrium associated with initial TDFW as follows:

$$\mathrm{CDSB_{t+s}} = \mathrm{CDSB_t} + \mathrm{s} \cdot \left[\frac{(\mathrm{EDSB_t} - \mathrm{CDSB_t})}{\mathrm{TMBE_t}}\right], \tag{3a}$$

$$\mathrm{CDSB_0} = \mathrm{EDSB_0}, \tag{3b}$$

where CDSB = current distance of the SB from the shoreline, EDSB = equilibrium distance of the SB from the shoreline, TMBE = time for the SB and FB to achieve equilibrium, t = state of the system at a given time, 0 = time zero, and s = time step.

The AQUACOAST model represents the mixing zone (MZ) between the interface and the freshwater body (at left of the interface in Figure 1). The MZ has a constant vertical thickness (vtmz) (Figure 1). Let us call freshwater boundary (FB) the line, in theory parallel to the shoreline for an isotropic and hydraulically homogeneous aquifer, where the upper-inland surface of the MZ intersects the horizontal plane connecting the bottom of pumping wells (Figure 1). The distance of the FB from the shoreline in the equilibrium associated with the current TDFW (EDFB) results from substituting dpwl by dpwl + vtmz in equations (2a)–(2c) as follows:

$$\mathrm{EDFB_t} = \mathrm{MIN}\left\{\mathrm{lgth}, \left(\left[\frac{(\mathrm{dpwl}+\mathrm{vtmz})^2 \cdot \mathrm{cfpm} \cdot \mathrm{dtcp} \cdot (1+\mathrm{dtcp})}{(2 \cdot \mathrm{TDFW_t})}\right] - \left[\frac{\mathrm{TDFW_t} \cdot (1-\mathrm{dtcp})}{(2 \cdot \mathrm{dtcp} \cdot \mathrm{cfpm})}\right]\right) \cdot 10^{-3}\right\}, \tag{4a}$$

$$\mathrm{EDFB_0} = \mathrm{MIN}\left\{\mathrm{lgth}, \left(\left[\frac{(\mathrm{dpwl}+\mathrm{vtmz})^2 \cdot \mathrm{cfpm} \cdot \mathrm{dtcp} \cdot (1+\mathrm{dtcp})}{(2 \cdot \mathrm{TDFW_0})}\right] - \left[\frac{\mathrm{TDFW_0} \cdot (1-\mathrm{dtcp})}{(2 \cdot \mathrm{dtcp} \cdot \mathrm{cfpm})}\right]\right) \cdot 10^{-3}\right\}, \tag{4b}$$

where EDFB = equilibrium distance of the FB from the shoreline, TDFW = net aquifer discharge, lgth = coastal aquifer length, dpwl = depth of pumping wells below sea level, vtmz = vertical thickness of the MZ, cfpm = hydraulic conductivity, dtcp = density contrast parameter, t = state of the system at a given time, and 0 = time zero.

Again, the AQUACOAST model assumes that the FB moves towards equilibrium following an exponential fitting.

Thus, the current distance of the FB from the shoreline (CDFB) is given by a first-order linear negative feedback. The two surfaces limiting the MZ would move at the same velocity, so equations (3a) and (3b) for CDSB and equations (5a) and (5b) for CDFB use the same average adjustment time. As deduced, no pumping wells whose distance from the shoreline is greater than CDFB take freshwater (Figure 1). It is assumed that, initially, the system is in the equilibrium associated with the current TDFW as follows:

$$\mathrm{CDFB}_{t+s} = \mathrm{CDFB}_t + s \cdot \left[\frac{(\mathrm{EDFB}_t - \mathrm{CDFB}_t)}{\mathrm{TMBE}_t}\right], \quad (5a)$$

$$\mathrm{CDFB}_0 = \mathrm{EDFB}_0, \quad (5b)$$

where CDFB = current distance of the FB from the shoreline, EDFB = equilibrium distance of the FB from the shoreline, TMBE = time for the SF and SB to achieve equilibrium, t = state of the system at a given time, 0 = time zero, and s = time step.

The time needed for SB and FB to achieve equilibrium (TMBE) differs when saltwater is encroaching (intrusion phase) and when freshwater is rinsing the salinized aquifer (freshening phase) [40]. In real aquifers, the freshening phase is quite longer than the intrusion phase. This is the reason why saltwater encroached during the last global sea-level maximum may still remain in some coastal aquifers [4, 44]. The TMBE expression includes a conditional function to reflect this difference as

$$\mathrm{TMBE}_t = \{\mathrm{IF}\ \mathrm{EDSB}_t > \mathrm{CDSB}_t\ \mathrm{THEN\ tmbi\ ELSE\ tmbf}\}, \quad (6)$$

where TMBE = time for the SB and FB to achieve equilibrium, EDSB = equilibrium distance of the SB from the shoreline, CDSB = current distance of the SB from the shoreline, tmbi = time for current boundaries to achieve equilibrium, intrusion phase, tmbf = time for current boundaries to achieve equilibrium, freshening phase, and t = state of the system at a given time.

3.4. Water Salinity. Verruijt's analytical solution assumes that water salinity in the FB (at distance CDFB from the shoreline) equals freshwater salinity (slfw), whereas water salinity in the SB (at distance CDSB from the shoreline) equals saltwater salinity (slsw). The AQUACOAST model defines the water salinity as the mass fraction of total dissolved solids relative to the sample weight: pristine coastal fresh groundwater is around $0.2\,\mathrm{kg\,m^{-3}}$ [10, 45], fresh groundwater in yielded coastal aquifers is typically in the $0.2–0.5\,\mathrm{kg\,m^{-3}}$ range [46], saltwater is $35\,\mathrm{kg\,m^{-3}}$ [47], and crop-limiting irrigation water is $2\,\mathrm{kg\,m^{-3}}$ [43]. Let us assume that mixing water salinity (S) along the x-axis distance from the shoreline ($\mathrm{CDSB} \le x \le \mathrm{CDFB}$) connecting SB and FB (Figure 1) is given by slsw and slfw as

$$S(x) = \mathrm{slfw} + (\mathrm{slsw} - \mathrm{slfw}) \cdot \left[\frac{(\mathrm{CDFB} - x)}{(\mathrm{CDFB} - \mathrm{CDSB})}\right]^{\mathrm{cdap}}, \quad (7)$$

where S = mixing water salinity along the x-axis distance from the shoreline ($\mathrm{CDSB} \le x \le \mathrm{CDFB}$), slfw = freshwater salinity, slsw = saltwater salinity, CDFB = current distance of the FB from the shoreline, CDSB = current distance of the SB from the shoreline, and cdap = calibration parameter to represent nonlinear variations in salinity.

Once defined S and slfw, the distance of the abandonment boundary (AB) (CDAB) from the shoreline (Figure 1) is obtained by equaling equation 7 to the crop-limiting salinity (slcl), i.e., the maximum water salinity for crop production, and then solving for x. Note that no pumping well whose distance from the shoreline is less than CDAB (i.e., slfw > slcl) can use groundwater for irrigation (Figure 1). CDAB is expressed as

$$\mathrm{CDAB}_t = \mathrm{CDFB}_t - (\mathrm{CDFB}_t - \mathrm{CDSB}_t) \cdot \left[\frac{(\mathrm{slcl} - \mathrm{slfw})}{(\mathrm{slsw} - \mathrm{slfw})}\right]^{(1/\mathrm{cdap})}, \quad (8)$$

where CDAB = current distance of the AB from the shoreline, CDFB = current distance of the SB from the shoreline, CDSB = current distance of the SB from the shoreline, slcl = crop-limiting salinity, slfw = freshwater salinity, slsw = saltwater salinity, cdap = calibration parameter to represent nonlinear variations in salinity, and t = state of the system at a given time.

3.5. Free Groundwater Surface. The FGS (water table) is not a horizontal plane in Verruijt's conceptualization (Figure 1). However, the AQUACOAST model only makes use of its average height above sea level. The expression giving such an average FGS in the equilibrium associated with the current TDFW (EHFW) results from integrating the water table height above sea level between 0 and the coastal aquifer length (lgth) and then dividing the result by lgth as

$$\mathrm{EHFW}_t = \left[\frac{8 \cdot \mathrm{dtcp} \cdot \mathrm{TDFW}_t \cdot \mathrm{lgth} \cdot 10^3}{(9 \cdot \mathrm{cfpm} \cdot (1 + \mathrm{dtcp}))}\right]^{(1/2)}, \quad (9a)$$

$$\mathrm{EHFW}_0 = \left[\frac{8 \cdot \mathrm{dtcp} \cdot \mathrm{TDFW}_0 \cdot \mathrm{lgth} \cdot 10^3}{(9 \cdot \mathrm{cfpm} \cdot (1 + \mathrm{dtcp}))}\right]^{(1/2)}, \quad (9b)$$

where EHFW = equilibrium average height of the FGS above sea level, TDFW = net aquifer discharge, lgth = coastal aquifer length, cfpm = hydraulic conductivity, t = state of the system at a given time, and 0 = time zero.

Once again, the current average height of the FGS (CHFW) is assumed to move towards equilibrium following an exponential fitting as

$$\mathrm{CHFW}_{t+s} = \mathrm{CHFW}_t + s\left[\frac{\mathrm{EHFW}_t - \mathrm{CHFW}_t}{\mathrm{tmhe}}\right], \quad (10a)$$

$$\mathrm{CHFW}_0 = \mathrm{EHFW}_0, \quad (10b)$$

where CHFW = current average height of the FGS above sea level, EHFW = equilibrium average height of the FGS above sea level, tmhe = time for the FGS to achieve equilibrium,

t = state of the system at a given time, 0 = time zero, and s = time step.

3.6. Irrigated Area. The potential irrigated area (ARPI) is the land area over the coastal aquifer where a pumping well with depth dpwl below sea level would take fresh groundwater whose salinity does not exceed the crop-limiting salinity, i.e., the area covered by distance lgth-CDAB (Figure 1). ARPI is expressed as

$$\mathrm{ARPI}_t = \mathrm{wdth} \cdot \left(\mathrm{lgth} - \mathrm{CDAB}_t\right) \cdot 10^2, \tag{11}$$

where ARPI = potential irrigated area, CDAB = current distance of the AB from the shoreline, wdth = coastal aquifer width, lgth = coastal aquifer length, hekm = hectare-to-square-kilometer conversion parameter, and t = state of the system at a given time.

Let us assume that a number of people own equally sized pieces of agricultural land over the coastal aquifer. These persons consider entering or leaving irrigated farming by comparing the average expected profit per irrigated farm (EXPF) with the returns from other alternative economic activities, i.e., with their respective opportunity costs. However, it takes some time for them to decide and carry out their plans. Additionally, suppose that the opportunity cost follows a generalized Rayleigh distribution across farmers, which is commonly used to analyze skewed data [48]. The cumulative distribution function of this distribution is generically expressed as

$$\mathrm{P}\left(\mathrm{X} \le \mathrm{x}\right) = 1 - \mathrm{EXP}\left\{\frac{-\mathrm{x}^2}{\left[\mathrm{A}^2\left(\mathrm{C}^2 + 1\right)\right]}\right\}. \tag{12}$$

In equation (12), X is the opportunity cost of a farmer and A and C are the mean and the coefficient of variation of the opportunity cost across farmers, respectively. Note that P, X, x, A, and C, such as, are not variables of the AQUACOAST model. In this way, $\mathrm{P}(X \le \mathrm{EXPF})$ is the fraction of owners whose opportunity cost is less than the average expected profit per irrigated farm, i.e., the fraction of owners that regard irrigation as profitable. But, given that farms are equally sized, $\mathrm{P}(X \le \mathrm{EXPF})$ is also the fraction of the potential irrigated area where irrigation is seen as profitable. Multiplying this fraction by the potential irrigated area, the target for the irrigated area (ARTI) is obtained as

$$\mathrm{ARTI}_t = \mathrm{ARPI}_t \cdot \left[1 - \mathrm{EXP}\left\{\frac{-\mathrm{EXPF}_t^2}{\left[\mathrm{ocfm}^2 \cdot \left(\mathrm{ocfv}^2 + 1\right)\right]}\right\}\right], \tag{13}$$

where ARTI = target for the irrigated area, ARPI = potential irrigated area, EXPF = expected profit per irrigated farm, ocfm = average opportunity cost for a farmer, ocfv = coefficient of variation of the opportunity cost for a farmer, and t = state of the system at a given time.

If irrigated farms are uniformly distributed over the coastal aquifer (excluding the coastal fringe CDAB where slfw > slcl), the rate of abandonment of irrigated area (RAIA), because pumping wells take groundwater whose salinity exceeds the crop-limiting salinity, equals the rate of saltwater encroachment. Thus, the expression of RAIA results from dividing the velocity of the SB displacement inland by the length of the potential irrigated area. RAIA will equal zero in the freshening phase, i.e., when EDSB < CDSB. The RAIA expression is

$$\mathrm{RAIA}_t = \mathrm{MAX}\left\{0, \frac{\left[\mathrm{EDSB}_t - \mathrm{CDSB}_t\right]}{\left[\mathrm{TMBE}_t \cdot \left(\mathrm{lgth} - \mathrm{CDAB}_t\right)\right]}\right\}, \tag{14}$$

where RAIA = rate of abandonment of irrigated area, EDSB = equilibrium distance of the SB from the shoreline, CDSB = current distance of the SB from the shoreline, TMBE = time for the SB and FB to achieve equilibrium, CDAB = current distance of the AB from the shoreline, lgth = coastal aquifer length, and t = state of the system at a given time.

The current irrigated area (ARIR) approaches towards the target for the irrigated area (ARTI) following an exponential fitting though some irrigated area is simultaneously abandoned when saltwater is encroaching, i.e., when RAIA > 0. The AQUACOAST model assumes that there is only one irrigated farm pumping from the aquifer at time zero. ARIR is expressed as

$$\mathrm{ARIR}_{t+s} = \mathrm{ARIR}_t + \mathrm{s} \cdot \left[\frac{\left(\mathrm{ARTI}_t - \mathrm{ARIR}_t\right)}{\mathrm{tfex}} - \mathrm{RAIA}_t \cdot \mathrm{ARIR}_t\right], \tag{15a}$$

$$\mathrm{ARIR}_0 = \mathrm{arfr}, \tag{15b}$$

where ARIR = current irrigated area, ARTI = target for the irrigated area, RAIA = rate of abandonment of irrigated area, tefx = time for farmers to adjust expectations, arfr = area of one farm, t = state of the system at a given time, and 0 = time zero.

3.7. Profit per Irrigated Farm. The parameter oaag allows the user to define whether the AQUACOAST model represents an open-air (oaag = 1) or an under plastic agricultural system (oaag ≠ 1). In the former case, precipitation contributes to irrigation, whereas, in the latter, it does not. The contribution of precipitation to irrigation (PPIR) is expressed as

$$\mathrm{PPIR}_t = \left\{\mathrm{IF\ oaag} = 1\ \mathrm{THEN\ appt} \cdot \left(1 - \mathrm{cfro}\right) \cdot 10\ \mathrm{ELSE}\ 0\right\}, \tag{16}$$

where PPIR = contribution of precipitation to irrigation; oaag = if 1, then open-air agriculture else under plastic agriculture; appt = annual precipitation; cfro = runoff coefficient; mmhe = cubic-meter-to-millimeter-per-hectare conversion parameter; and t = state of the system at a given time.

The crop yield per hectare (YLDH) is related to the sum of the irrigation dose from groundwater (IRDS) and the contribution of precipitation to irrigation (PPIR) by means of a logistic or parabolic function. The function involves a salinity factor that measures the negative effect of water salinity on crop yield and ranges between zero and one. It can be proved that, for any nonnull value of the salinity

Table 3: Files compiled in the AQUACOAST Venapp.

File	Type	Number	Description
AQUACOAST.vmf	Simulation model	1	The usual extension of a Vensim model is ∗.mdl. However, for Venapp programming, it must be saved in binary format (∗.vmf).
AQUACOAST.vcd	Venapp	1	Vensim application.
Intro.vmf	Simulation model	1	Simulation model to explore Vensim options.
Intro.vcd	Venapp	1	Vensim application.
Intro.vgd	Vensim graph data		Customized graphs that allow displaying more than one variable, with more than one scale, on a single graph.
∗.vdf	Vensim data file	n	Contains results from a simulation, including scenario values; the name of the exercise (∗) is provided by the user.
Default.vdf	Vensim data file	1	Contains results from a default simulation, including scenario values.

factor, crop yield is maximized when IRDS + PPIR equals the water requirement for agronomic optimum yield (wrao) as

$$\mathrm{YLDH_t} = \mathrm{pycm} \cdot \mathrm{SLFT_t} \cdot \left(\mathrm{IRDS_t} + \mathrm{PPIR_t}\right) \cdot \mathrm{MAX}\left\{0, 1 - \left[\frac{\left(\mathrm{IRDS_t} + \mathrm{PPIR_t}\right)}{(2 \cdot \mathrm{wrao})}\right]\right\}, \tag{17}$$

where YLDH = crop yield per surface hectare, SLFT = salinity factor, IRDS = irrigation dose from groundwater, PPIR = contribution of precipitation to irrigation, pycm = potential marginal yield per cubic meter of water, wrao = water requirement for agronomic optimum yield, and t = state of the system at a given time.

The expression of the salinity factor (SLFT) combines two values of another salinity factor, the one given by [49]. This factor is generically expressed as

$$\mathrm{SLFT} = \mathrm{MAX}\left\{0, 1 - \mathrm{MAX}\frac{\{0, \mathrm{S} - \mathrm{slfp}\}}{(\mathrm{slcl} - \mathrm{slfp})}\right\}, \tag{18}$$

where SLFT = salinity factor generically expressed, S = mixing water salinity along the x-axis distance from the shoreline (CDSB ≤ x ≤ CDFB) as in equation 7, slcl = crop-limiting salinity, slfp = another crop-specific calibration parameter, and t = state of the system at a given time.

The first value of SLFT is one, its length is lgth–CDFB (Figure 1), and it corresponds to the freshwater zone where pumping wells take fresh groundwater. The second value of SLFT corresponds to the inland part of the MZ where pumping wells take salty groundwater still useful for crop production (slcl ≥ slfw), and its length is CDFB–CDAB (Figure 1). Integrating equation 7 between CDAB and CDFB and then dividing the result by CDFB–CDAB gives the actual average salinity across the line segment connecting FB and AB (slmz). Thus, the value of SLFT in the second zone (sfmz) results from using slmz instead of S in equation (19). Finally, SLFT is a weighted average of the two mentioned values of SLFT, where the weights are the lengths of the respective zones. Note that, in doing so, (1) the crop yield per hectare (YLDH) is also a weighted average of the crop yields corresponding to both zones, and (2) if salinity increased linearly along the line segment between CDAB and CDFB, i.e., if cdap = 1, slmz would be the simple average of the crop limiting (slcl) and the freshwater (slfw) salinities. SLFT is expressed as

$$\mathrm{SLFT_t} = \frac{\left[\left(\mathrm{lgth} - \mathrm{CDFB_t}\right) + \mathrm{sfmz} \cdot \left(\mathrm{CDFB_t} - \mathrm{CDAB_t}\right)\right]}{\left(\mathrm{lgth} - \mathrm{CDAB_t}\right)}, \tag{19a}$$

$$\mathrm{sfmz} = \mathrm{MAX}\left\{0, 1 - \mathrm{MAX}\frac{\{0, \mathrm{slmz} - \mathrm{slfp}\}}{(\mathrm{slcl} - \mathrm{slfp})}\right\}, \tag{19b}$$

$$\mathrm{slmz} = \mathrm{slfw} + \frac{(\mathrm{slcl} - \mathrm{slfw})}{(\mathrm{cdap} + 1)}, \tag{19c}$$

where YLDH = salinity factor, CDFB = current distance of the FB from the shoreline, CDAB = current distance of the AB from the shoreline, lgth = coastal aquifer length, sfmz = average salinity factor across a line segment connecting the FB and AB, slmz = actual water salinity across a line segment connecting the FB and AB (CDAB ≤ x ≤ CDFB), slfp = calibration parameter to represent nonlinear variations in salinity, slcl = crop-limiting salinity, slfw = freshwater salinity, and t = state of the system at a given time.

The cost of water per cubic meter (CSTW) combines the cost of the energy needed to pump groundwater, which depends on the FGS (water table) depth and the price of energy and other costs as

$$\mathrm{CSTW_t} = \mathrm{egcm} \cdot \left(\mathrm{hght} - \mathrm{CHFW_t}\right) \cdot \mathrm{preg} + \mathrm{cwot}, \tag{20}$$

where CSTW = water cost per cubic meter, CHFW = current average height of the FGS above sea level, egcm = energy required to pump one cubic meter one meter, hght = average height of land surface above sea level, preg = price of energy, cwot = costs of water other than energy, and t = state of the system at a given time.

The profit per hectare (PFTH) is the difference in incomes and costs per hectare. Incomes come from the selling of agricultural production and from subsidies. The costs per hectare are the cost of water and other costs. The former results from multiplying the water cost per cubic meter (CSTW) and the actual irrigation dose from groundwater (IRDS). This product tries to reach the desired irrigation dose over the efficiency of the irrigation system. Other costs

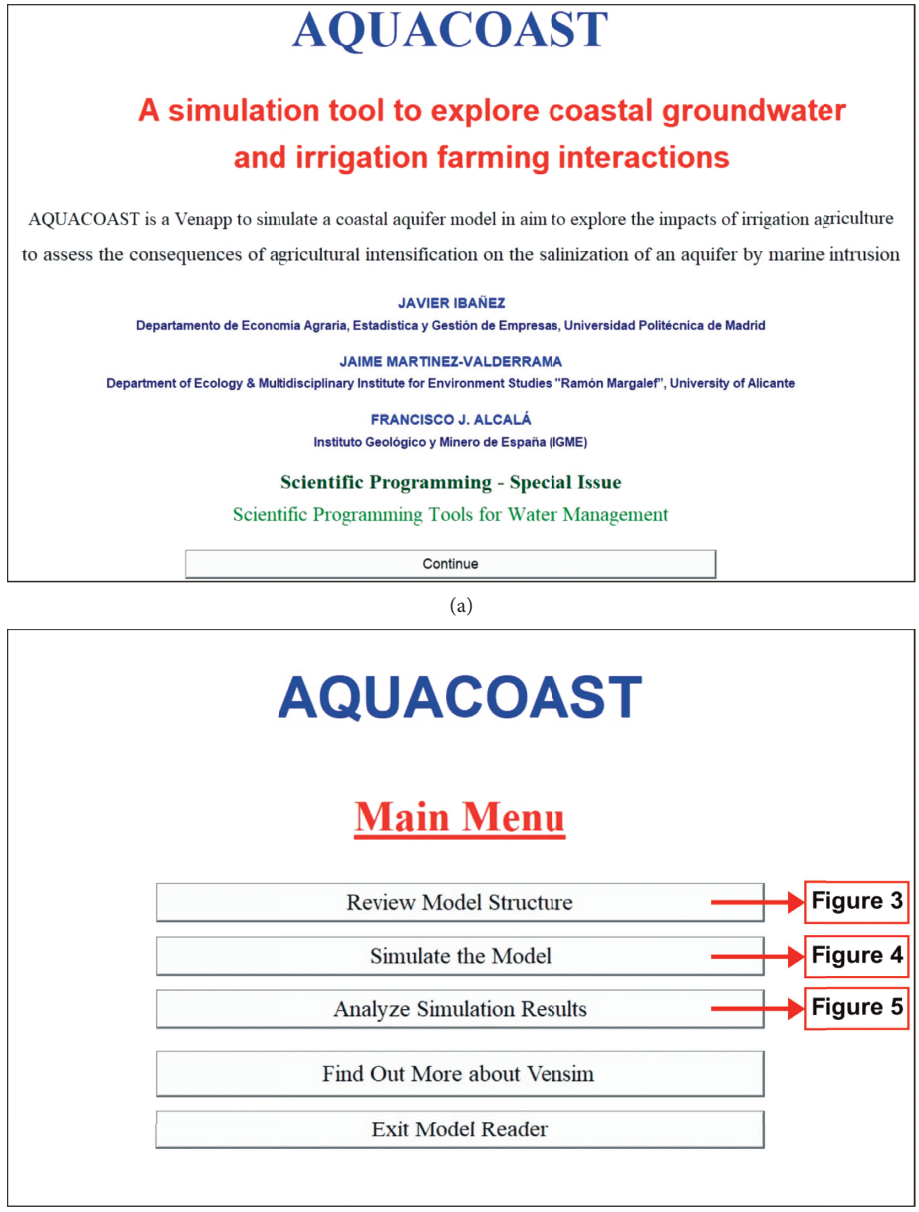


Figure 2: (a) Welcome screen of the AQUACOAST Venapp; (b) main menu screen.

per hectare are aggregated into a single model parameter (octh). PFTH is expressed as

$$\mathrm{PFTH}_t = \mathrm{prcp} \cdot \mathrm{YLDH}_t + \mathrm{subh} - \frac{(\mathrm{CSTW}_t \cdot \mathrm{IRDS}_t)}{\mathrm{efir} - \mathrm{octh}}, \quad (21)$$

where PFTH = profit per hectare, YLDH = crop yield per hectare, CSTW = water cost per cubic meter, IRDS = irrigation dose from groundwater, prcp = crop price, subh = subsidies per hectare, efir = irrigation system efficiency, octh = other costs per hectare, and t = state of the system at a given time.

The expected profit per irrigated farm (EXPF) results from multiplying the profit per hectare (PFTH) and the area of one farm (arfr). The AQUACOAST model implements the expectation about such a profit is a moving average of observed past profits where weights exponentially decrease over time. This is the well-known exponential smoothing, which can be expressed as a first-order linear negative feedback [26] as

$$\mathrm{EXPF}_{t+s} = \mathrm{EXPF}_t + s \cdot \left[\frac{\mathrm{arfr} \cdot \mathrm{PFTH}_t - \mathrm{EXPF}}{\mathrm{tfex}}\right], \quad (22a)$$

$$\mathrm{EXPF}_0 = \mathrm{expf}, \quad (22b)$$

where EXPF = expected profit per irrigated farm, PFTH = profit per hectare, arfr = area of one farm, tfex = time for farmers to adjust expectations, expf = initial expected profit per irrigated farm, t = state of the system at a given time, 0 = time zero, and s = time step.

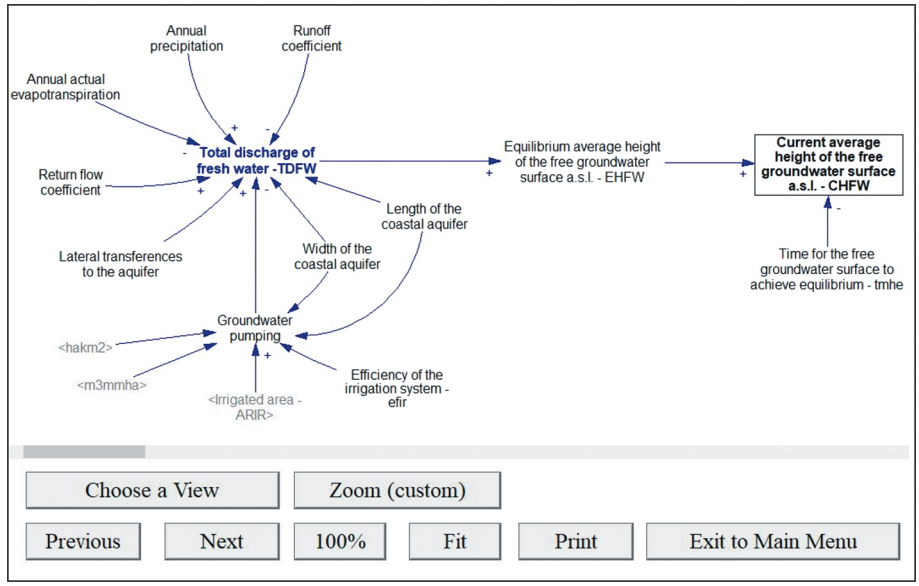


Figure 3: "Coastal aquifer dynamics" view of the AQUACOAST model.[[parms resize(1),pos(50,50),size(200,200),bgcol(156)]]

3.8. Groundwater Pumping. The AQUACOAST model includes the parameter pxys that allows the user to define whether farmers seek to maximize the profit per irrigated farm (pxys = 1) or crop yield (pxys ≠ 1). Note that the water requirement for each of these goals is different. In the latter case, the water requirement is given by the parameter wrao as in equation (17). The expression of the expected water requirement for profit maximization (EXWX) will be given later. Whatever the case, the irrigation dose from groundwater (IRDS) is the difference in water requirement and contribution of precipitation to irrigation. The resultant irrigation dose cannot be greater than the irrigation quota, if any. The expression of the initial value of IRDS simply considers EXWX = wrao. IRDS is expressed as

$$\mathrm{IRDS}_t = \mathrm{MIN}\{\mathrm{idqt}, \mathrm{MAX}\{0, \{\mathrm{IF\ pxys} = 1\ \mathrm{THEN\ EXWX}_t \ \mathrm{ELSE\ wrao}\} - \mathrm{PPIR}_t\}\}, \tag{23a}$$

$$\mathrm{IRDS}_0 = \mathrm{MIN}\{\mathrm{idqt}, \mathrm{MAX}\{0, \mathrm{wrao} - \mathrm{PPIR}_t\}\}, \tag{23b}$$

where IRDS = irrigation dose from groundwater; PPIR = contribution of precipitation to irrigation; EXWX = expected water requirement for profit maximization; idqt = irrigation quota; pxys = if 1, then farmers seek economic optimum else they seek agronomic optimum; wrao = water requirement for agronomic optimum yield; t = state of the system at a given time; and 0 = time zero.

The water requirement for profit maximization (WRXP) results from calculating the partial derivative of equation (21) to deduce profit per hectare with respect to the irrigation dose, equaling it to zero, and then solving for the irrigation dose as

$$\mathrm{WRXP}_t = \mathrm{wrao} \cdot \left[\frac{1 - \mathrm{CSTW}_t}{(\mathrm{efir} \cdot \mathrm{prcp} \cdot \mathrm{pycm} \cdot \mathrm{SLFT}_t)}\right], \tag{24}$$

where WRXP = water requirement for profit maximization, CSTW = water cost per cubic meter, SLFT = salinity factor, wrao = water requirement for agronomic optimum yield, efir = irrigation system efficiency, prcp = crop price, pycm = potential marginal yield per cubic meter of water, and t = state of the system at a given time.

For the current year, farmers cannot know the water requirement for profit maximization because many of the factors involved are unknown a priori, e.g., crop and energy prices, or the incidence of salinization. However, farmers can form an approximation of such a water requirement for the current year (EXWX) by learning from past experiences. This learning process would follow an exponential fitting to the true water requirement for profit maximization. In the AQUACOAST model, the initial value of EXWX is taken to be the water requirement for agronomic optimum yield, a technical parameter which is easier to grasp by farmers. EXWX is expressed as

$$\mathrm{EXWX}_{t+s} = \mathrm{EXWX}_t + s \cdot \left[\frac{(\mathrm{WRXP}_t - \mathrm{EXWX}_t)}{\mathrm{tfex}}\right], \tag{25a}$$

$$\mathrm{EXWX}_0 = \mathrm{wrao}, \tag{25b}$$

where EXWX = expected water requirement for profit maximization, WRXP = water requirement for profit maximization, tfex = time for farmers to adjust expectations, wrao = water requirement for agronomic optimum yield, t = state of the system at a given time, 0 = time zero, and s = time step.

Groundwater pumping (PUMP) results from multiplying the current irrigated area (ARIR) and the actual irrigation dose (IRDS) and then dividing by the coastal aquifer area as

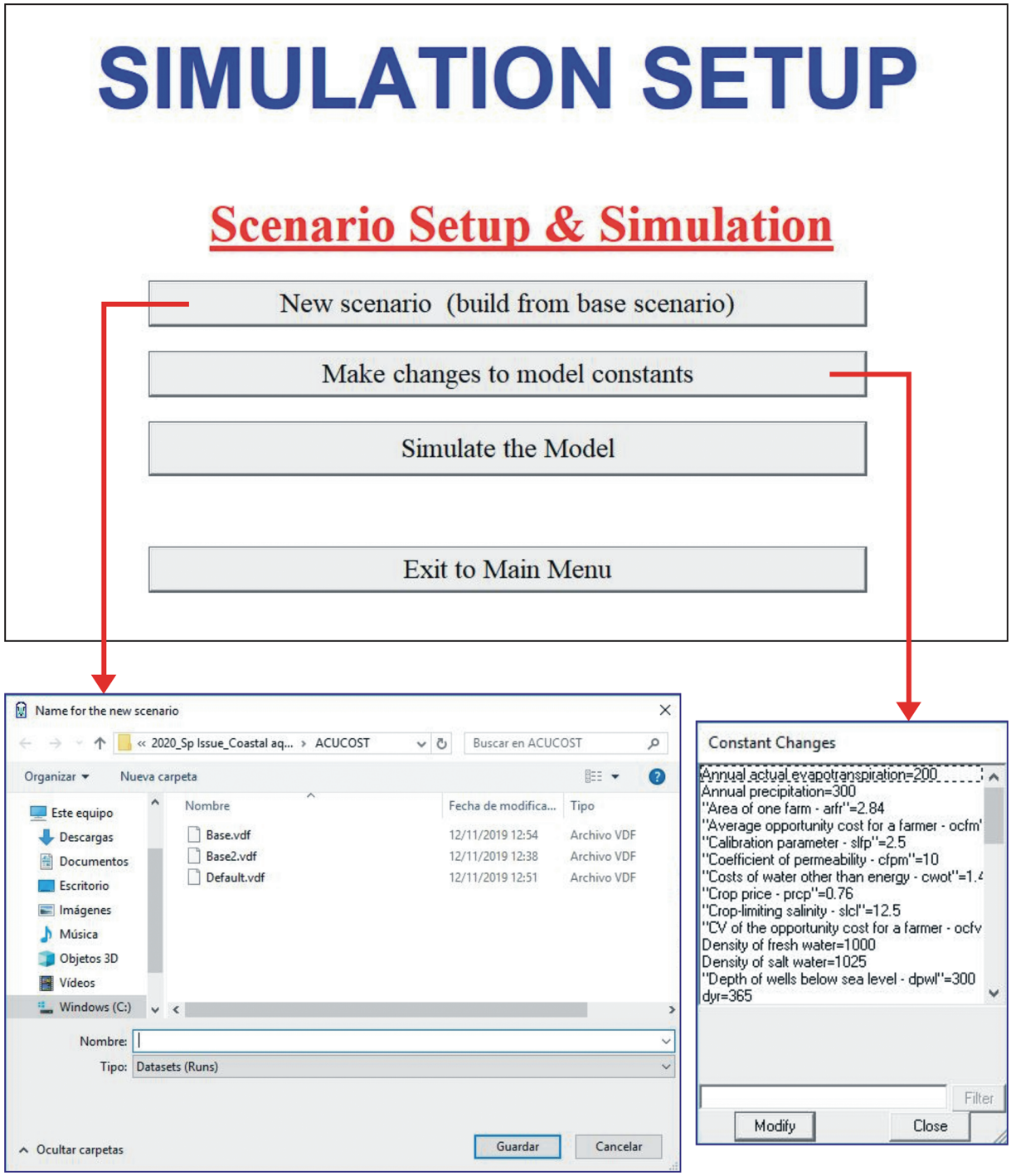


Figure 4: Simulation setup screen, including setting up and simulation scenarios.

$$\mathrm{PUMP}_t = \frac{\mathrm{ARIR}_t \cdot \mathrm{IRDS}_t}{(\mathrm{wdth} \cdot \mathrm{lgth} \cdot \mathrm{efir} \cdot 10^3)}, \tag{26a}$$

$$\mathrm{PUMP}_0 = \frac{\mathrm{ARIR}_0 \cdot \mathrm{IRDS}_0}{(\mathrm{wdth} \cdot \mathrm{lgth} \cdot \mathrm{efir} \cdot 10^3)}, \tag{26b}$$

where PUMP = groundwater allotment for irrigation, ARIR = current irrigated area, IRDS = irrigation dose from groundwater, wdth = coastal aquifer width, lgth = coastal aquifer length, efir = irrigation system efficiency, t = state of the system at a given time, and 0 = time zero.

4. AQUACOAST Venapp

4.1. Overall Description. The Vensim® DSS extension Venapps™ was utilized to program the user-friendly interface of the above described formulations. To use AQUACOAST, it is necessary to install Vensim Reader and

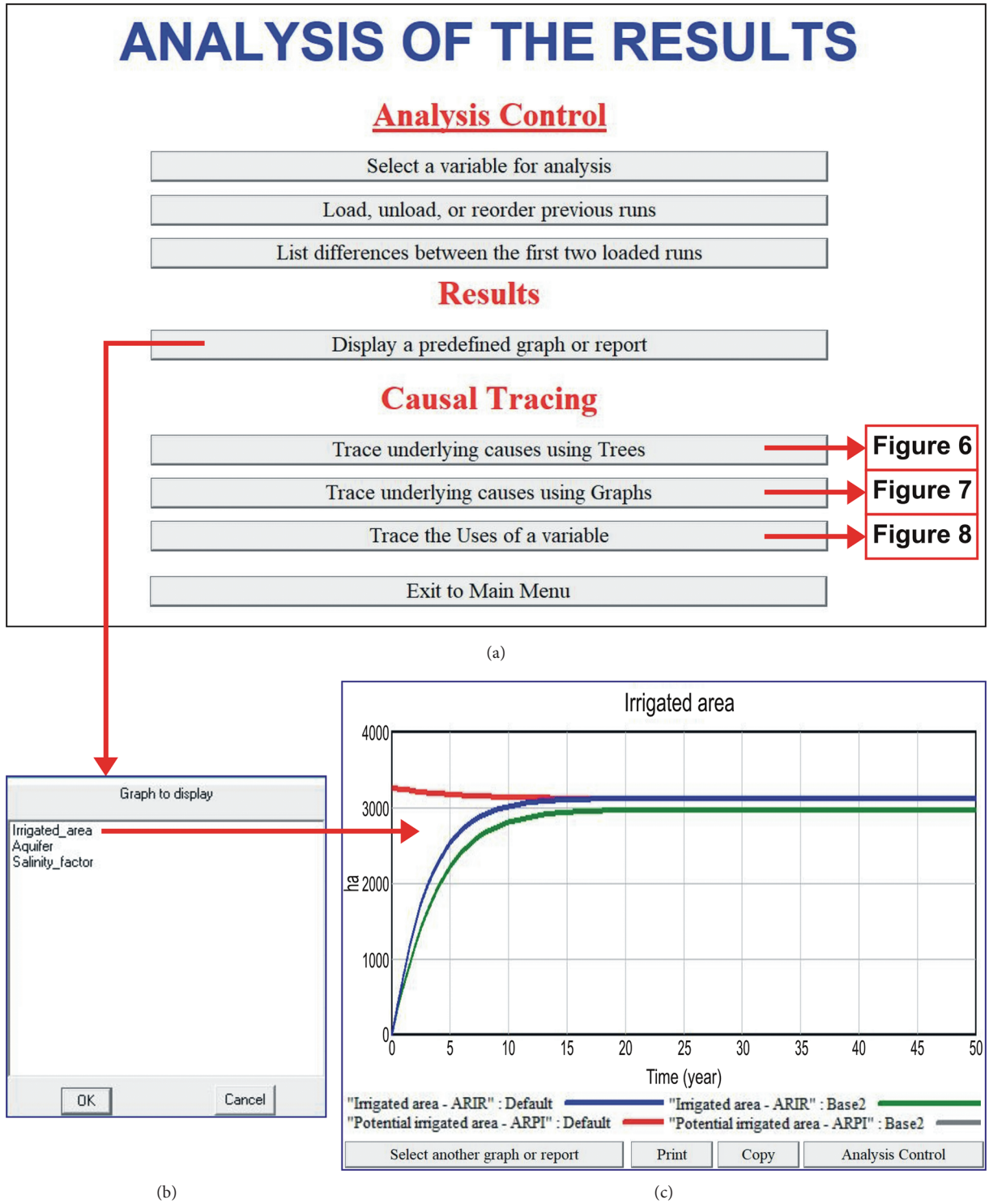


Figure 5: (a) "Analysis of the results" screen including three blocks: analysis control, results, and causal tracing; (b) for the "result" block, menu of available graphs; (c) temporal trends of the main variables for the default optional scenarios, in this example, "irrigated area."

the folder with all the necessary files (Table 3) on the computer.

After opening the AQUACOAST file, the welcome screen will appear (Figure 2(a)). Here is brief information about the model and its authors. Clicking on the "Continue" button takes the user to the main menu (Figure 2(b)). From this point, it is possible to (i) review the model structure, (ii) access the screen from which the simulation scenarios configured the model is executed, and (iii) analyze the results using different tools of Vensim. The bottom part of the screen gives the option to explore the possibilities offered by Vensim (button "Find Out More about Vensim"), which is

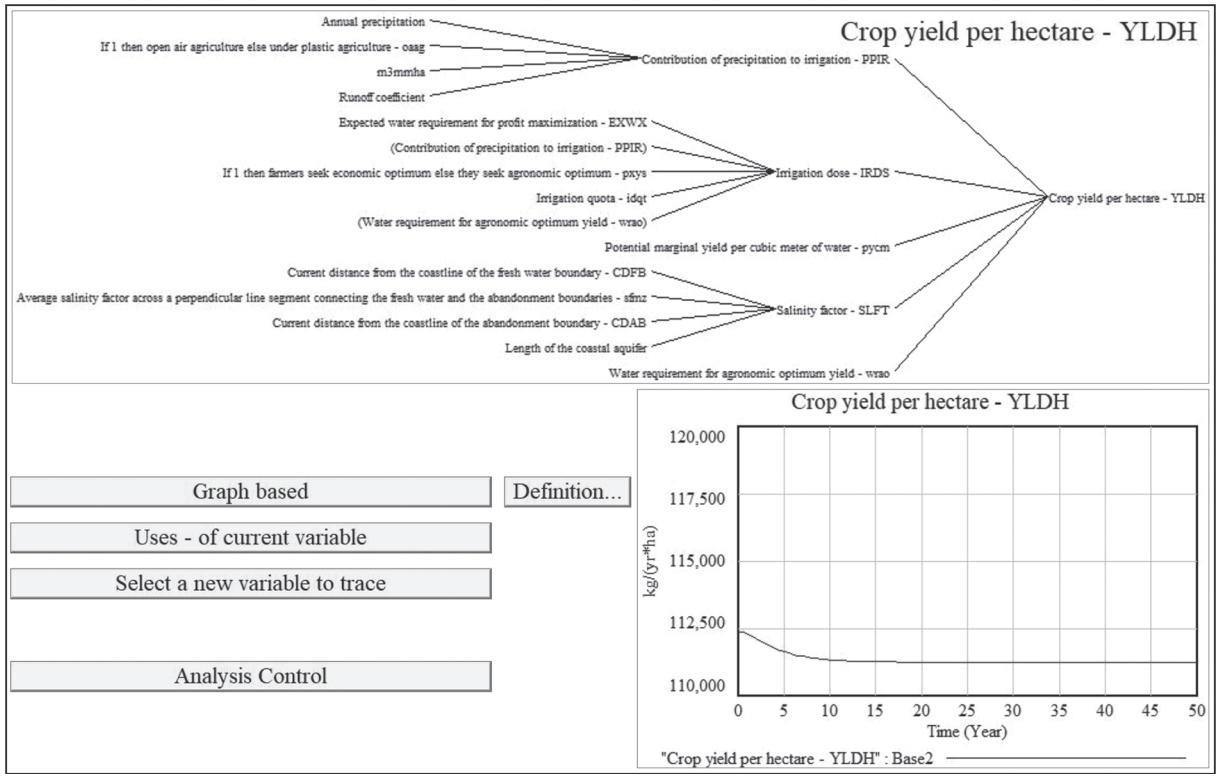


Figure 6: "Trace underlying causes with Trees" screen. Example of the underlying causes of "Crop yield per hectare" variable using trees.

supported by the Intro.vmf model file (Table 3) and is part of the Vensim examples, or leave the simulator (button "Exit Model Reader").

4.2. Exploring the Structure of the Model. One of the main purposes of SD is to understand how a system works. To do this, Vensim offers multiple tools. The first allows studying the model structure. For this to be practical, the model must be split into different screens or "Views" by clicking the button "Review Model Structure" as in Figure 2(b) so that the user can gradually assimilate the relationship between the variables involved. Figure 3 shows one of these screens as an example. The button captions themselves explain the available utilities.

4.3. Running the Model. The button "Simulate the Model" as in Figure 2(b) enables the model to be executed. In this part of the Venapp (Figure 4), the simulation scenarios are set. For the former, the "New scenario. . ." button opens a dialog box to select an existing simulation or create a new one. After that, the "Make changes. . ." button will show all the model parameters and their default values; by modifying them, alternative scenarios will be created.

These changes are not permanent, for it would be necessary to access the model through a commercial license of Vensim® DSS and change them. Therefore, if the model is simulated without any modification, the user will obtain the results contained in the Default.vdf simulation file (Table 3). After setting the scenario, the model is ready to be simulated. After clicking on the button to execute Vensim "Simulate the Model," a blank screen will appear on which the user has to click to move on to the next stage, which consists of visualizing and analyzing the results.

4.4. Analyzing Results. After clicking button "Analyze Simulation Results" as in Figure 2(b), this stage offers three blocks of possibilities: analysis control, results, and causal tracing (Figure 5(a)). The "Analysis Control" block allows the user to manage simulations, select variables, and compare scenarios. This is an useful option when starting to accumulate simulations. In that case, it is hard to remember what changes from one scenario to another. Using the "Load, unload. . ." button, this option details which parameters change (and how much) between the two previously selected simulations.

The "Results" block displays predesigned graphs showing the temporal evolution of the most important variables of the model (Figure 5(a)). Clicking on the button "Display a predefined...," a list will appear with the available graphics (Figure 5(b)). In all of them, the results of the default scenario will be included, as well as the results of the simulation that appear in the first position of the list of loaded simulations, which are managed from the "Load, unload, or reorder. . ." button, as in Figure 5(c).

The "Causal Tracing" block (Figure 5(a)) enables the use of Vensim's own analysis tools, as shown in Figures 6–8. The tree diagram tool creates a graphical representation of the

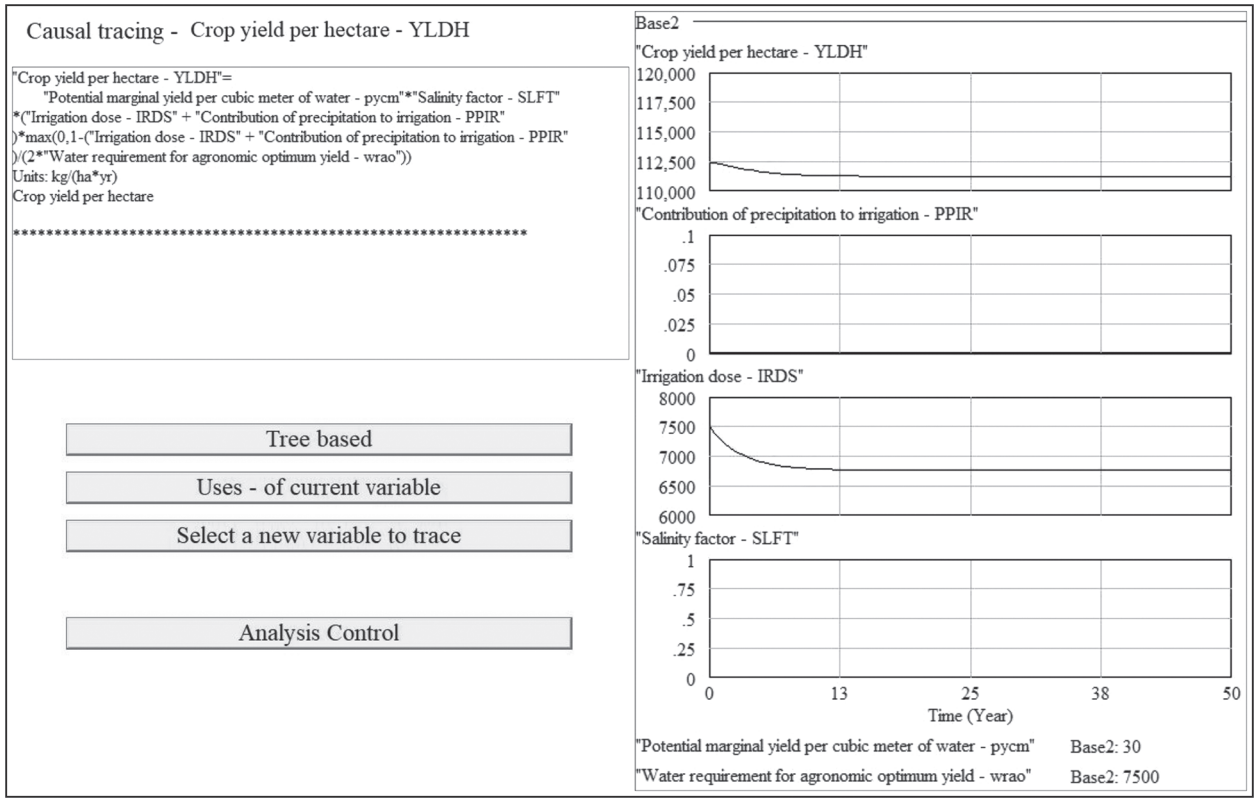


Figure 7: "Trace underlying causes with Graphs" screen. Example of the underlying causes of "Crop yield per hectare" variable using graphs.

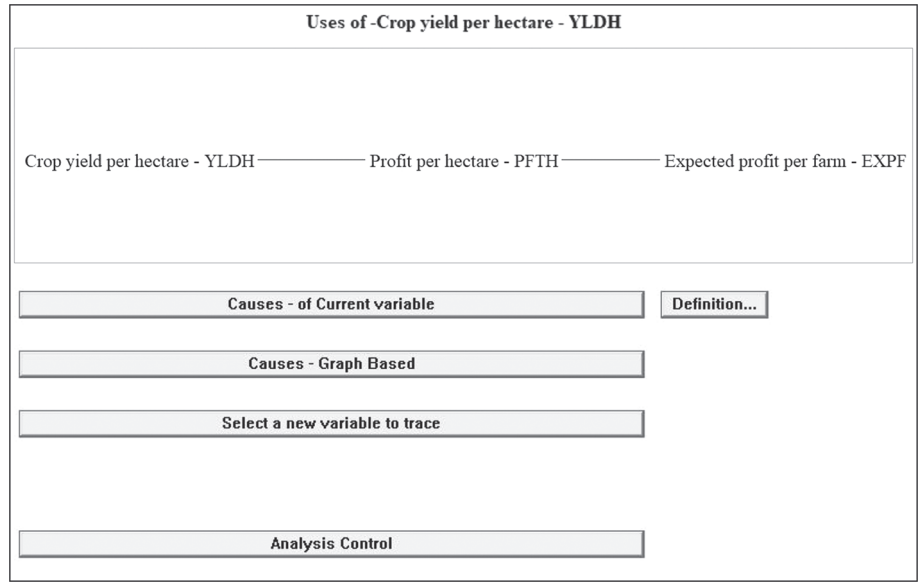


Figure 8: "Trace the Uses of a variable" screen. Example of the underlying causes of "Crop yield per hectare" variable.

structure of a model associated with the selected workbench variable; the currently active variable is displayed at the top left of the screen. The tree diagram tool provides an easy-to-read, graphical representation of the causes and uses of variables. This is a convenient way for displaying structural relationships and therefore information about a complex structure in a relatively small space.

The causal structure of the model can be explored using trees or charts, whereas the uses of each variable can only be analyzed using trees by displaying the tree of variables affected by the selected one. Figure 6 shows the screen that appears after clicking the button "Trace underlying causes using Trees." At the top is the causal tree of the currently active variable. It is possible to select

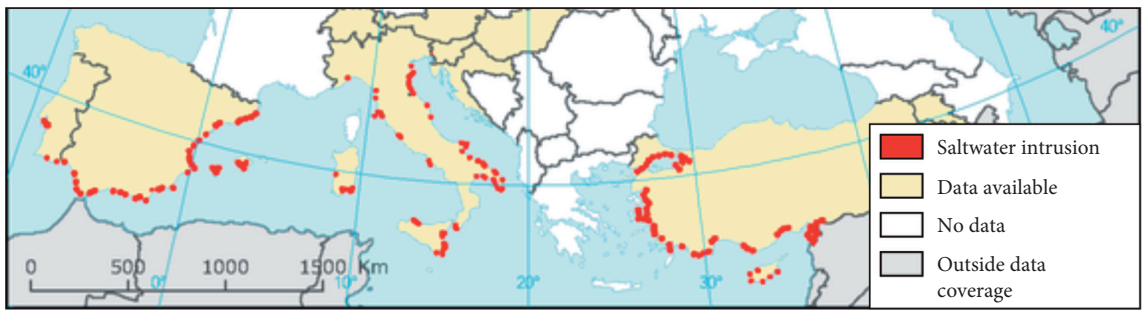


Figure 9: Southern Mediterranean European groundwater bodies affected by saltwater intrusion due to overexploitation (modified from [55]).

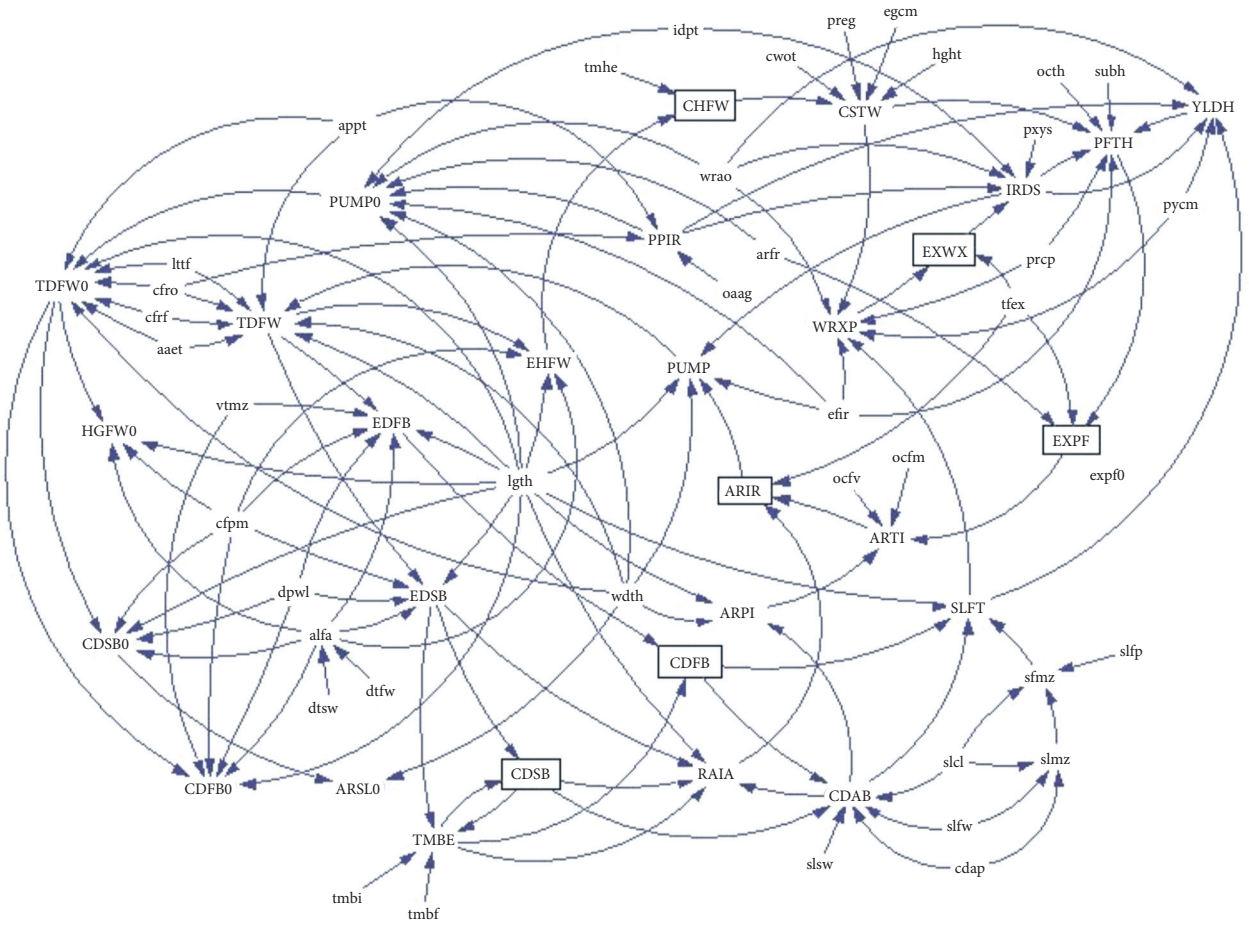


Figure 10: Original sketch of the AQUACOAST model. Acronyms for variables (capital letters) and parameters (small letters) are as in Tables 1 and 2, respectively.

other variables using the "Select a new variable..." button or, interactively, by clicking on any variable in the same causal tree. In addition, the user can (i) access to the model equations through the button "Definition..." and (ii) see its temporal trajectory according to the last simulated or loaded scenario. The rest of the buttons allow the user to scroll to other screens, as their own caption indicates. In this way, all the options of using the tree diagram tool are interconnected.

Clicking the button "Trace underlying causes using Graphs" (Figure 5(a)), the graphical option to study the causal tree of a variable can be used (Figure 7). In this case, only first-order variables, i.e., those directly affected by the selected variable, are shown. On the right side of the screen, the temporal trajectories of the variables included in the performance equation can be observed, as well as the value of the parameters involved. The definition of the variable is also presented in the box on the left. The rest of the options allow

the user to navigate between the screens related to the diagram tool.

The third analysis option offered by this Venapp (Figure 5(a)) is to "Trace the Uses of a variable" (Figure 8). Again, it is possible to see its definition and scroll to other analysis screens. Finally, it is important to highlight the interactivity of the tree diagram tool. Any of the outputs it generates allows the user to select a new variable by clicking on it. This option is designed to make it easier to explore and study the model, one of Vensim's hallmarks.

5. Discussion and Conclusions

5.1. Global Relevance of Coastal Groundwater Resources. Groundwater is one of the more strategic renewable resources, especially in coastal drylands [4]. Over two billion people rely on groundwater as their primary potable freshwater source [50], whereas half or more of the irrigation water used for food production has also this source [17]. Groundwater also acts as the key strategic reserve in times of drought [51] and to set population in drylands with sparse economic options except agriculture [8, 20].

Soaring water withdrawal worsens water scarcity already prevalent in those coastal drylands having limited surface water resources. A consequence of the aquifer overexploitation to cope with water requirements is saltwater intrusion, which threats the sustainable food production and economic development in an increasing number of countries [15, 52, 53].

The use of SD modelling in water management has a long tradition due to the complex nature of the problems addressed [54]. However, few models consider the interaction between terrestrial water fluxes and human activities and associated uses, and even fewer are the models intended to coastal groundwater-dependent irrigation agriculture [15]. The AQUACOAST tool is a contribution to fill this gap of knowledge.

The AQUACOAST is a global application tool despite that the Mediterranean region is postulated as a main scenario in which to apply. Coastal aquifers in this area support strong groundwater pumping rates for crop production, and saltwater intrusion has extended in most of them (Figure 9). For instance, in Spain, this problem affects most of the Mediterranean coastal aquifers, so they were considered one of the "desertification landscapes" in the National Program to Combat Desertification [56, 57]. Our immediate purpose is to apply AQUACOAST to coastal aquifers of the Mediterranean arc where tourist and agricultural pressure has triggered problems of marine intrusion such as Campo de Dalías, Campo de Níjar, Vélez River and Verde River in Spain, Martíl-Alila and Smir in Morocco, and Nador and Tipaza in Argelia.

5.2. Using SD for Modelling Coastal Groundwater-Dependent Irrigation Systems. SD models express in mathematical language (the most precise way that exists) the dynamic processes that take place in a given field of study. The academic and scientific value of understanding how many variables relate, determining what is cause and what is effect, making explicit feedback loops of a system, and, finally, translating them into a set of operative differential equations is very high.

However, it is important to make an additional effort so that this information reaches the less specialized public involved in decision-making processes. In order to do this, the use of DSSs is essential. An example taken from AQUACOAST can help us understand the value of programming an application aimed at a wider audience. Figure 10 shows the causal diagram drawn up for purely academic use as an indecipherable amalgam in the eyes of someone other than the modeler. Variables are represented by acronyms, and it is difficult to follow the path of causality in the complex web of causal relationships. However, a careful edition of the model, expliciting the full name of each variable and splitting the diagram into five screens (Figure 3 shows one of them), allows the user to create a much more user-friendly version.

The precision of mathematical language should not confuse us about the nature of the SD models. In reality, when confronted with socioecological systems, we lack universal physical or chemical laws that allow us to talk about predictions. Socioecological systems require input from a wide range of sources and types of knowledge [58]. This includes qualitative and quantitative pieces on information from various stakeholder groups such as scientists, policy makers, and community members. To collect, synthesize, and use these datasets in useful ways, integrated assessment and modelling needs to utilize and combine different methods (i.e., conceptual, numerical, and participatory) in appropriate methodological designs that best fit the project's context, objective, and constraints, in the latter case including resource availability [27, 33].

The AQUACOAST model is not intended for prediction or forecasting, even though they provide outputs over variable time periods. This is because there are not enough data to validate the models for such purposes; remember, for example, the lack of some essential information such as the initial groundwater storage or similar others like past weather data or former land uses. Thus, the aim is to get qualitative rather than quantitative outputs, answering the "what if" questions, which helps to compare simulation scenarios that address different hypotheses.

System understanding and experimentation also constitute important goals of the models, as well as social learning, though this is limited to those cases where the assessment shows a special risk of degradation. Since degradation usually proceeds slowly, stakeholders are not sufficiently aware of its consequences, no matter how they value the water and land use being degraded [59]. Thus, the AQUACOAST model is intended as a "means of exploration" [60, 61] for helping them to better understand how a coastal groundwater-dependent agricultural system may behave.

5.3. The AQUACOAST Venapp. Just as there are several SD modelling platforms, DSS can be built using a wide range of possibilities. Venapps offer clear advantages when the software chosen is VensimDSS. The construction is very fast since there are templates that with few modifications allow

the user to implement the main options of a DSS: configuration of scenarios, simulation, and visualization of the results and their analysis. The use of the tool is simple since it only requires to install the free version of VensimDSS (Vensim Reader) and to copy the Vensim files provided by the modeler.

The main problem is the graphical limitation of Vensim, which prevents the use of images or figures or the implementation of options such as global sensitivity analysis or risk estimation. More complex DSS based on Vensim can be built using various programming languages (e.g., Visual Basic) and platforms (e.g., Excel), such as the application designed by authors to study desertification risk in Spain, which combined some hotspots such as soil erosion, grassland overgrazing, and aquifer overexploitation [57]. However, the choice of the type of DSS depends, as does the construction of a model, on its purpose.

Since the AQUACOAST tool is aimed at (i) facilitating the study and understanding of coastal groundwater-dependent irrigated agriculture systems and (ii) creating an application whose installation and use are very simple, Venapp is a perfect ally since it makes available to the user, in a very simple way, all the analysis tools of Vensim.

Since Meadows et al. [62] presented the revolutionary best-seller "Limits to growth," for which they used systems thinking and SD concepts in order to explain how short-term development policies can lead to "overshoot and collapse" behavior of socioecological systems, SD has shown its potential as a tool to help understand complex socio-ecological systems and is still regarded as a valuable resource for thinking about sustainable futures [33, 63]. The AQUACOAST tool is in line with these principles and aims to highlight the fragility of such a strategic renewable resource as groundwater. As long as the model describes reality with a certain accuracy, the modelling process and its outcomes can be used to improve our understanding of the problem of excessive pumping under the increasing food production needs a necessary step towards affecting sustainable and effective change [64].

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was funded by the European Research Council grant agreement no. 647038 (BIODESERT).

References

[1] C. Small and R. J. Nicholls, "A global analysis of human settlement in coastal zones," *Journal of Coastal Research*, vol. 19, pp. 584–599, 2003.

[2] J. S. Famiglietti, "The global groundwater crisis," *Nature Climate Change*, vol. 4, no. 11, pp. 945–948, 2014.

[3] R. J. Naiman, J. J. Magnuson, D. M. McKnight, J. A. Stanford, and J. R. Karr, "Freshwater ecosystems and their management: a national initiative," *Science*, vol. 270, no. 5236, pp. 584-585, 1995.

[4] E. Custodio, "Coastal aquifers of Europe: an overview," *Hydrogeology Journal*, vol. 18, no. 1, pp. 269–280, 2010.

[5] E. Custodio, J. M. Andreu-Rodes, R. Aragón et al., "Groundwater intensive use and mining in south-eastern peninsular Spain: hydrogeological, economic and social aspects," *Science of The Total Environment*, vol. 559, pp. 302–316, 2016.

[6] S. Sabater and D. Barceló, *Water Scarcity in the Mediterranean. Perspectives under Global Change*, Springer-Verlag, Berlin, Germany, 2010.

[7] C. Dalin, Y. Wada, T. Kastner, and M. J. Puma, "Groundwater depletion embedded in international food trade," *Nature*, vol. 543, no. 7647, pp. 700–704, 2017.

[8] F. J. Alcalá, M. Martín-Martín, F. Guerrera, J. Martínez-Valderrama, and P. Robles-Marín, "A feasible methodology for groundwater resource modelling for sustainable use in sparse-data drylands: application to the Amtoudi Oasis in the northern Sahara," *Science of The Total Environment*, vol. 630, pp. 1246–1257, 2018.

[9] P. Baudron, F. Barbecot, J. L. G. Aróstegui, C. Leduc, Y. Travi, and D. Martinez-Vicente, "Impacts of human activities on recharge in a multilayered semiarid aquifer (Campo de Cartagena, SE Spain)," *Hydrological Processes*, vol. 28, no. 4, pp. 2223–2236, 2014.

[10] F. J. Alcalá and E. Custodio, "Spatial average aquifer recharge through atmospheric chloride mass balance and its uncertainty in continental Spain," *Hydrological Processes*, vol. 28, no. 2, pp. 218–236, 2014.

[11] E. Custodio, W. M. Edmunds, and Y. Travi, "Management of coastal palaeowaters," in *Paleowaters in Coastal Europe: Evolution of Groundwater since the Late Pleistocene*, W. M. Edmunds and C. J. Milne, Eds., vol. 189, pp. 313–327, Geological Society. Special Publications, London, UK, 2001.

[12] M. Rodríguez-Rodríguez, J. Benavente, F. J. Alcalá, and M. Paracuellos, "Long-term water monitoring in two Mediterranean lagoons as an indicator of land-use changes and intense precipitation events (Adra, Southeastern Spain)," *Estuarine, Coastal and Shelf Science*, vol. 91, no. 3, pp. 400–410, 2011.

[13] European Union, "Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption," *Official Journal of the European Union*, vol. 330, pp. 32–54, 1998.

[14] B. Scanlon, C. Faunt, L. Longuevergne et al., "Groundwater depletion and sustainability of irrigation in the US high plains and central valley," *Proceedings of the National Academy of Sciences*, vol. 109, no. 24, pp. 9320–9325, 2012.

[15] Y. Wada, D. Wisser, and M. F. P. Bierkens, "Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources," *Earth System Dynamics*, vol. 5, no. 1, pp. 15–40, 2014.

[16] P. Döll and S. Siebert, "Global modeling of irrigation water requirements," *Water Resources Research*, vol. 38, no. 4, 2002.

[17] S. Siebert, J. Burke, J. M. Faures et al., "Groundwater use for irrigation—a global inventory," *Hydrology and Earth System Sciences*, vol. 14, no. 10, pp. 1863–1880, 2010.

[18] D. M. FitzGerald, M. S. Fenster, B. A. Argow, and I. V. Buynevich, "Coastal impacts due to sea-level rise,"

Annual Review of Earth and Planetary Sciences, vol. 36, no. 1, pp. 601–647, 2008.

[19] V. E. A. Post, "Fresh and saline groundwater interaction in coastal aquifers: is our technology ready for the problems ahead?" *Hydrogeology Journal*, vol. 13, no. 1, pp. 120–123, 2005.

[20] F. J. Alcalá, J. Martínez-Valderrama, P. Robles-Marín et al., "A hydrological-economic model for sustainable groundwater use in sparse-data drylands: application to the Amtoudi Oasis in southern Morocco, northern Sahara," *Science of The Total Environment*, vol. 537, pp. 309–322, 2015.

[21] J. Martínez-Valderrama, J. Ibáñez, F. J. Alcalá, A. Dominguez, M. Yassin, and J. Puigdefábregas, "The use of a hydrological-economic model to assess sustainability in groundwater-dependent agriculture in drylands," *Journal of Hydrology*, vol. 402, no. 1-2, pp. 80–91, 2011.

[22] A. E. Rizzoli and W. J. Young, "Delivering environmental decision support systems: software tools and techniques," *Environmental Modelling & Software*, vol. 12, no. 2-3, pp. 237–249, 1997.

[23] M. Matthies, C. Giupponi, and B. Ostendorf, "Environmental decision support systems: current issues, methods and tools," *Environmental Modelling & Software*, vol. 22, no. 2, pp. 123–127, 2007.

[24] J. D. Sterman and L. B. Sweeney, "Cloudy skies: assessing public understanding of global warming," *System Dynamics Review*, vol. 18, no. 2, pp. 207–240, 2002.

[25] J. W. Forrester, *Industrial Dynamics*, The MIT Press, Cambridge, UK, 1961.

[26] J. D. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill, New York, NY, USA, 2000.

[27] R. A. Kelly, A. J. Jakeman, O. Barreteau et al., "Selecting among five common modelling approaches for integrated environmental assessment and management," *Environ. Model. Softw.*vol. 47, pp. 159–181, 2013.

[28] J. Ibáñez, J. Martínez-Valderrama, and J. Puigdefábregas, "Assessing overexploitation in Mediterranean aquifers using system stability condition analysis," *Ecological Modelling*, vol. 218, no. 3-4, pp. 260–266, 2008.

[29] J. Martínez-Valderrama, J. Ibáñez, G. Del Barrio et al., "Doomed to collapse: why Algerian steppe rangelands are overgrazed and some lessons to help land-use transitions," *Science of The Total Environment*, vol. 613-614, pp. 1489–1497, 2018.

[30] J. M. Fernández and M. A. E. Selma, "The dynamics of water scarcity on irrigated landscapes: mazarrón and Aguilas in south-eastern Spain," *System Dynamics Review*, vol. 20, no. 2, pp. 117–137, 2004.

[31] T. Uehara, Y. Nagase, and W. Wakeland, "Integrating economics and system dynamics approaches for modelling an ecological-economic system," *Systems Research and Behavioral Science*, vol. 33, no. 4, pp. 515–531, 2015.

[32] R. L. Eberlein, D. W. Peterson, V. Systems, and J. G. Road, "Understanding models with Vensim," *European Journal of Operational Research*, vol. 59, no. 1, pp. 216–219, 1992.

[33] S. Elsawah, S. A. Pierce, S. H. Hamilton et al., "An overview of the system dynamics process for integrated modelling of socio-ecological systems: lessons on good modelling practice from five case studies," *Environmental Modelling & Software*, vol. 93, pp. 127–145, 2017.

[34] M. Moench, J. Burke, and Y. Moench, "Rethinking the approach to groundwater and food security," vol. 24, FAO, Rome, Italy, 2003, Water Reports.

[35] Ventana Systems Inc., *VensimDSS*, Ventana Systems Inc., Harvard, MA, USA, 2019.

[36] Y. Barlas and V. Dicker, "Decision support for strategic university management: a dynamic interactive game," in *Proceedings of the International System Dynamics Conference*, Istanbul, Turkey, August 1997.

[37] P. Pfaffenbichler, "Modelling with systems dynamics as a method to bridge the gap between politics, planning and science? Lessons learnt from the development of the land use and transport model MARS," *Transport Reviews*, vol. 31, no. 2, pp. 267–289, 2011.

[38] P. A. Walker, R. Greiner, D. McDonald, and V. Lyne, "The Tourism Futures Simulator: a systems thinking approach," *Environmental Modelling & Software*, vol. 14, no. 1, pp. 59–67, 1998.

[39] A. Verruijt, "A note on the ghyben-herzberg formula," *International Association of Scientific Hydrology. Bulletin*, vol. 13, no. 4, pp. 43–46, 1968.

[40] E. Giménez-Forcada, "Space/time development of seawater intrusion: a study case in Vinaroz coastal plain (Eastern Spain) using HFE-Diagram, and spatial distribution of hydrochemical facies," *Journal of Hydrology*, vol. 517, pp. 617–627, 2014.

[41] J. F. Reynolds, D. M. S. Smith, E. F. Lambin et al., "Global desertification: building a science for dryland development," *Science*, vol. 316, no. 5826, pp. 847–851, 2007.

[42] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Irrigation and Drainage*, Food and Agriculture Organization, Rome, Italy, 1998.

[43] R. S. Ayers and D. Westcot, *Water Quality for Agriculture. Irrigation and drainage paper 29*, Food and Agriculture Organization, Rome, Italy, 1985.

[44] H. Kooi, J. Groen, and A. Leijnse, "Modes of seawater intrusion during transgressions," *Water Resources Research*, vol. 36, no. 12, pp. 3581–3589, 2000.

[45] E. Dahl, "Ecological salinity boundaries in poikilohaline waters," *Oikos*, vol. 7, no. 1, pp. 1–21, 1956.

[46] H. A. Swenson and H. L. Baldwin, *A Primer on Water Quality*, USGS. Department of Interior, Washington, DC, USA, 1965.

[47] W. Stumm and J. J. Morgan, *Aquatic Chemistry, Chemical Equilibria and Rates in Natural Waters*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1996.

[48] M. Z. Raqab and M. T. Madi, "Generalized Rayleigh distribution," in *International Encyclopedia of Statistical Science*, M. Lovric, Ed., Springer Berlin Heidelberg, Berlin, Germany, pp. 599–603, 2011.

[49] R. Aragües, *Manejo de los suelos. Interpretación. Caso particular de salinización y sodificación. Correcciones., Jornada Técnica del Colegio Oficial de Ingenieros Técnicos Agrícolas de Lérida "Fertilización y Riego: Aplicaciones en Fruticultura*, Colegio Oficial de Ingenieros Técnicos Agrícolas de Lérida, Lleida, Spain, 2011.

[50] W. M. Alley, R. W. Healy, J. W. LaBaugh, and T. E. Reilly, "Flow and storage in groundwater systems," *Science*, vol. 296, pp. 1985–1990, 2002.

[51] J. S. Famiglietti, M. Lo, S. L. Ho et al., "Satellites measure recent rates of groundwater depletion in California's Central Valley," *Geophysical Research Letters*, vol. 38, pp. 2–5, 2011.

[52] P. Döll, K. Fiedler, and J. Zhang, "Global-scale analysis of river flow alterations due to water withdrawals and reservoirs," *Hydrology and Earth System Sciences*, vol. 13, no. 12, pp. 2413–2432, 2009.

[53] M. Kummu, P. J. Ward, H. De Moel, and O. Varis, "Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia," *Environmental Research Letters*, vol. 5, pp. 1–10, 2010.

[54] P. P. Rogers and M. B. Fiering, "Use of systems analysis in water management," *Water Resources Research*, vol. 22, no. 9S, pp. 146S–158S, 1986.

[55] European Environment Agency (EEA), "The European environment: state and outlook 2005," Technical Report No 1/2005, European Environment Agency (EEA), Copenhagen, Denmark, 2005.

[56] MAGRAMA, *Programa de Acción Nacional contra la Desertificación*, MAGRAMA, Madrid, Spain, 2008.

[57] J. Martínez-Valderrama, J. Ibáñez, G. Del Barrio et al., "Present and future of desertification in Spain: implementation of a surveillance system to prevent land degradation," *Science of The Total Environment*, vol. 563–564, pp. 169–178, 2016.

[58] A. J. Jakeman, R. A. Letcher, and J. P. Norton, "Ten iterative steps in development and evaluation of environmental models," *Environmental Modelling & Software*, vol. 21, no. 5, pp. 602–614, 2006.

[59] J. Ibáñez, J. F. L. Contador, S. Schnabel, M. P. Fernández, and J. Martínez-Valderrama, "A model-based integrated assessment of land degradation by water erosion in a valuable Spanish rangeland," *Environmental Modelling & Software*, vol. 55, pp. 201–213, 2014.

[60] T. Oxley, B. McIntosh, N. Winder, M. Mulligan, and G. Engelen, "Integrated modelling and decision-support tools: a Mediterranean example," *Environmental Modelling & Software*, vol. 19, no. 11, pp. 999–1010, 2004.

[61] G. L. W. Perry and J. D. A. Millington, "Spatial modelling of succession-disturbance dynamics in forest ecosystems: concepts and examples," *Perspectives in Plant Ecology, Evolution and Systematics*, vol. 9, no. 3-4, pp. 191–210, 2008.

[62] D. Meadows, D. Meadows, J. Randers, and W. W. Behrens III, *The Limits to Growth*, Universe Books, New York, NY, USA, 1972.

[63] G. M. Turner, "On the cusp of global collapse? Updated comparison of the limits to growth with historical data," *GAIA—Ecological Perspectives for Science and Society*, vol. 21, no. 2, pp. 116–124, 2012.

[64] I. Winz, G. Brierley, and S. Trowsdale, "The use of system dynamics simulation in water resources management," *Water Resources Management*, vol. 23, no. 7, pp. 1301–1323, 2009.

Hindawi
Scientific Programming
Volume 2020, Article ID 6458576, 12 pages
https://doi.org/10.1155/2020/6458576

Research Article

A New Clustering Algorithm and Its Application in Assessing the Quality of Underground Water

T. Vo-Van,[1] A. Nguyen-Hai,[2] M. V. Tat-Hong,[2] and T. Nguyen-Trang[3,4]

[1]*College of Natural Science, University of Can Tho, Can Tho City, Vietnam*
[2]*The Institute for Environment and Resources, Vietnam National University, Ho Chi Minh City, Vietnam*
[3]*Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[4]*Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

Correspondence should be addressed to T. Nguyen-Trang; nguyentrangthao@tdtu.edu.vn

Received 9 October 2019; Revised 7 January 2020; Accepted 4 February 2020; Published 7 March 2020

Guest Editor: Francisco Gomariz

Cluster analysis, which is to partition a dataset into groups so that similar elements are assigned to the same group and dissimilar elements are assigned to different ones, has been widely studied and applied in various fields. The two challenging tasks in clustering are determining the suitable number of clusters and generating clusters of arbitrary shapes. This paper proposes a new concept of "epsilon radius neighbors" which plays an essential role in the cluster-forming process, thereby determining both the number of clusters and the shape of clusters, automatically. Based on "epsilon radius neighbors," a new clustering algorithm in which the epsilon radius value is adapted to the characteristics of each cluster in the current partition is proposed. Recently, clustering has been widely applied in environmental applications, including underground water quality monitoring. However, the existing studies have simply applied conventional clustering techniques, in which the abovementioned two challenging tasks have not been solved already. Therefore, in this paper, the proposed clustering algorithm is applied in assessing the underground water quality in Phu My Town, Ba Ria-Vung Tau Province, Vietnam. The experimental results on benchmark datasets demonstrate the effectiveness of the proposed algorithm. For the quality of underground water, the new algorithm results in four clusters with different characteristics. Through this application, we found that the new algorithm might provide valuable reference information for underground water management.

1. Introduction

Cluster analysis is to discover the underlying structure of a dataset by partitioning the data into groups so that similar elements are assigned to the same group and dissimilar elements are assigned to different ones [1–5]. Recently, along with the development of big data, cluster analysis has been extensively studied and widely applied in various fields, such as physics, biology, economics, engineering, sociology, and data mining. [6]. For solving the problem of clustering, several approaches have been proposed in the literature, which includes: nonhierarchical clustering (k-means, k-means ++, etc. [7, 8] and other variances), hierarchical clustering [9], clustering for probability functions [1], or fuzzy clustering [10]. Among the abovementioned approaches, k-means clustering is the most well known and widely applied in various fields. However, the k-means algorithm and its extensions usually require a user-defined number of clusters that is often unknown in practice. (i) Furthermore, the k-means algorithm constructs spherical clusters, which is unsuitable for arbitrary-shaped clusters. (ii) The above two problems have been the major drawbacks of clustering so far, which lead to many difficulties and challenges in solving this problem [6].

For (i), to determine the suitable number of clusters, the most commonly used approach is running the clustering algorithm several times with different number of clusters each time, and evaluating them based on a number of

internal validity measures, such as S-index, F-index, Dunn index, and Xie-Beni index [11–14]. This approach can investigate the suitable number of clusters, but it repeats the clustering process many times to find the best number of clusters, thereby increasing the amount of time and space required, according to [6]. Moreover, the abovementioned evaluation indices are distance-based measures; therefore, they can only evaluate the qualities of spherical clusters and cannot be used for arbitrary-shaped clusters. In [15], Mavridis et al. proposed the algorithm PFClust (Parameter Free Clustering). The term "parameter free" means that the algorithm can automatically determine the number of clusters without requiring any user-defined parameters. For this purpose, PFClust performs an agglomerative algorithm on many subdatasets that are randomly sampled several times. Given an internal validity measure and a set of threshold corresponding to the number of clusters, the suitable threshold is then chosen based on the distribution of the given internal measure for all possible clustering results. In comparison to other conventional clustering algorithms, PFClust can result in a little better performance; however, it repeats the process of sampling and evaluating internal measures of the given thresholds in several times. Consequently, PFClust tends to be more time-consuming and expensive than other clustering methods. References [16–18] found the optimal partition by combining the metaheuristic optimization method and the clustering. These studies used the abovementioned internal validity measures as objective functions that need to be optimized to find the best clustering solution. It is well known that the metaheuristic optimization method, e.g., the genetic algorithm, results in an extreme computational cost, which reduces the efficiency of the algorithm. Furthermore, in spite of outputting the number of clusters and partitioning automatically, the metaheuristic optimization method requires a few of its own user-defined parameters that have effects on the optimal solution. As a result, avoiding the challenge of specifying the number of clusters, k leads to the challenge of specifying many other parameters. In [19], an automatic clustering algorithm was conducted using a function of force that can control the movements of the objects. The farther the distance, the weaker the force between two objects. In the end, each object converges to the center of the cluster it belongs to. Since the computing of force also requires a user-defined parameter denoted λ and the value of λ also has effects on the number of clusters, the attempt to overcome the problem of [16–18] of this algorithm is not too significant.

For (ii), DBSCAN [20], a density-based algorithm, is the most well-known method to construct arbitrary-shaped clusters. The algorithm utilizes two connectivity functions termed as density-reachable and density-connected, and each data instance is indicated as either a core point or a border point. The algorithm works to expand core points to form a cluster around itself. A drawback of DBSCAN is that when clusters of different densities exist, only particular kinds of noise points are captured [21]. Besides, two user-defined parameters regarding the minimum size of clusters and the radius need to be carefully turned. The other approaches, such as kernel k-means [22] and spectral clustering [23] can construct arbitrary-shaped clusters; these methods, however, also require a predefined number of clusters.

Because of the abovementioned drawbacks, an investigation of a new clustering method which can automatically determine the number of clusters and the clusters' shape is necessary. This paper proposes a new clustering method based on a new definition called "ε-radius neighbors" of a given point $\mathbf{x}_0$. ε-radius neighbors play a key role in constructing clusters with arbitrary shapes. When any new ε-radius neighbor is not found, the algorithm stops processing the current cluster and thereby the number of clusters is automatically determined. Furthermore, the radius ε can be adapted to specific cluster density, which is an advantage of the proposed methods in comparison with DBSCAN.

The quality of underground water depends on various factors, such as climate, characteristics of aquifers, pH, alkalinity, redox potential of the geological environment, initial sources, contamination due to human activities, and biological processes. The conventional methods of assessing the quality of groundwater are usually based on comparing the parameters representing water quality, which are collected by sensors, with the permitted standards. Clustering can help explain complex data matrix, analyze the similarities in water quality characteristics, and group them into clusters, thereby showing their general characteristics, as well as the causes that affect water quality. Therefore, clustering has been widely applied in environmental applications, including underground water quality monitoring. Some studies, for example, [24–28], have applied clustering in order to classify the water qualities in the whole region and design a future spatial sampling strategy in an optimal manner, which can reduce the number of sampling stations and associated costs. However, the abovementioned studies simply applied conventional clustering methods, such as hierarchical clustering with Ward distance, and k-means clustering. These methods, in general, have encountered the disadvantages, as mentioned in the previous parts. Therefore, in this paper, the proposed clustering algorithm is applied in assessing the underground water quality in Phu My Town, Ba Ria-Vung Tau Province, Vietnam. This application is expected to produce more reliable and valuable information so that the administrators can monitor underground water behavior.

The remainder of this paper is organized as follows. Section 2 presents the study area, the data collection, and the proposed method. The results and discussion are presented in Section 3 in which Section 3.1 is the validation of the proposed algorithms for different datasets and Section 3.2 is the application in assessing the underground water quality in Phu My Town, Ba Ria-Vung Tau Province, Vietnam. Finally, Section 4 is the conclusion.

2. Materials and Methods

2.1. The Proposed Clustering Method. Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, $\mathbf{x}_i \in R^d$ be a set of n points, $\mathbf{x}_0 \in R^d$ be a given point, and ε be an arbitrarily positive integer. A set $S \subseteq X$ is called as ε-radius neighbors of $\mathbf{x}_0$ if

$$S = \{\mathbf{x}_i \in X: d(\mathbf{x}_i, \mathbf{x}_0) \le \varepsilon\}, \quad (1)$$

where $d(\mathbf{x}_i, \mathbf{x}_0)$ is the Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_0$.

Obviously, ε-radius neighbors of $\mathbf{x}_0$ are located in a hypersphere of radius ε around $\mathbf{x}_0$. As a result, a cluster can be extended by searching on the dataset and adding new objects pertaining to any hypersphere of radius ε around the current objects. This process still depends on the value of ε. This parameter plays a role which is the same as the parameter ε in the well-known DBSCAN algorithm. The choice of this parameter has effects on the clustering result. A fixed value of ε has low generalization ability because different datasets and clusters with different densities in a dataset could require different values of ε. A natural strategy is simply to adapt ε using the current cluster density. For the sake of presentation, the set of pairwise distances in the current cluster is called the set of "historical extending." Based on the set of "historical extending" in the current cluster (samples), we can estimate the maximum "extending" of the entire cluster (population). In this case, two basic principles are as follows:

(1) We know that if data has the normal distribution with mean μ and standard deviation σ, then 95% of the data values belong to the interval $\mu \pm 2\sigma$. If the two abovementioned parameters are unknown and data is enough large, we can estimate them from the sample data. For example, the mean and the adjusted standard deviation of the sample can be selected as alternatives for μ and σ, respectively. Therefore, to estimate the maximum extending of the cluster (population), we can use the following formula:

$$\varepsilon = \max D = \overline{d} + 2s_d, \quad (2)$$

where $\overline{d}$ and s_d are the mean and adjusted standard deviation of "historical velocities" in the current-processing cluster (sample). Obviously, about 97.5% of the extending pertaining to the true cluster (population) must be less than the extending estimated by formula (2), and thus, this formula can be used to approximate the maximum extending of the true cluster (population).

(2) Let n be the sample size or the number of objects in the current-processing cluster and $\overline{d}$ and s_d be the sample mean and adjusted standard deviation, respectively. Assuming that the value of n is large enough or d has the normal distribution with the mean $\mu(d)$ and the variance s_d^2/n. Consequently, with a significant level of 0.05, the mean of d belongs to the interval $\overline{d} \pm (1.96s_d/\sqrt{n})$. As a result, the maximum of the mean extending can be directly estimated using the following formula:

$$\varepsilon = \overline{d} + \frac{1.96s_d}{\sqrt{n}}. \quad (3)$$

The maximum value of the confidence interval is then used as the representative extending of the cluster.

It can be observed from formulas (2) and (3) that, in the earlier processing stage, when the sample size is too small, the standard deviation and the adaptive extending must be large. Therefore, we can avoid unreasonable extending in the earlier processing stage when the current sample is not a good representation of the population. Meanwhile, in the later stage, the number of objects in the current-processing cluster or the sample size is large enough for maintaining a stable adaptive extending.

Based on formulas (2) and (3), we propose a new clustering method called adaptive radius clustering for automatically determining the number of clusters and clusters shapes. Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, $\mathbf{x}_i \in R^d$ be an original dataset of N objects. The new clustering algorithm is presented as the following pseudocode and in Figure 1.

Initialize $C_1^{\text{old}} = \varnothing$, $C_1^{\text{new}} = \varnothing$, and $\text{centroid}_{\text{new}} = \varnothing$, where C_1^{old} and C_1^{new} are current-processing cluster obtained before and after an update, respectively.

Step 1. Get the first three objects of the cluster using the formulas below:

$$\mathbf{v}_1 = \arg\min_{x_i \in X} \sum_{j=1}^{n} d(\mathbf{x}_i, \mathbf{x}_j), \quad (4)$$

$$\mathbf{v}_2 = \arg\min_{x_i \in X/\{v_1\}} d(\mathbf{x}_i, \mathbf{v}_1), \quad (5)$$

$$\mathbf{v}_3 = \arg\min_{x_i \in X/\{v_1, v_2\}} d(\mathbf{x}_i, \mathbf{v}_1), \quad (6)$$

which subject to

$$d(\mathbf{v}_2, \mathbf{v}_1) \le d(\mathbf{v}_3, \mathbf{v}_1) < \binom{2}{n}^{-1} \sum_{i \ne j} d(\mathbf{x}_i, \mathbf{x}_j). \quad (7)$$

Update

$$C_1^{\text{old}} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}; C_1^{\text{new}} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}. \quad (8)$$

In formulas (4), (5), and (6), argument "arg" of a function is the value that must be provided to obtain the function's result; hence $\mathbf{v}_1 = \text{argmin}_{x_i \in X} \sum_{j=1}^{n} d(\mathbf{x}_i, \mathbf{x}_j)$ is a point $\mathbf{x}_i$ in X such that the sum of distances between it and other points is the minimum. In other words, $\mathbf{v}_1$ is the centroid of the current dataset. Similarly, $\mathbf{v}_2$ is the nearest point of $\mathbf{v}_1$ and $\mathbf{v}_3$ is the nearest point of $\mathbf{v}_1$ when excluding $\mathbf{v}_2$. Formula (7) is defined to overcome the problem of bad initialization. For example, if $\mathbf{v}_2$ and $\mathbf{v}_3$ are two nearest neighbors of $\mathbf{v}_1$, but the corresponding distances are larger than the average of pairwise distances between points in the current dataset, then $\mathbf{v}_1$ will be considered as a single cluster and the current extending process will be stopped.

In the abovementioned formulas, d is the Euclidean distance between any two d-dimensional points. In some illustration below, for the sake of visualization, $\mathbf{x}$ will be chosen as a 2-dimensional point ($\mathbf{x}_1$ and $\mathbf{x}_2$) so that we can draw the scatter plot of data. In fact, $\mathbf{x}_1$ and $\mathbf{x}_2$ not only can be the coordinates but also can be other informations such as height, weight, Ca^{2+}, Mg^{2+}, and Na^{+}. Furthermore, $\mathbf{x}$ can be a

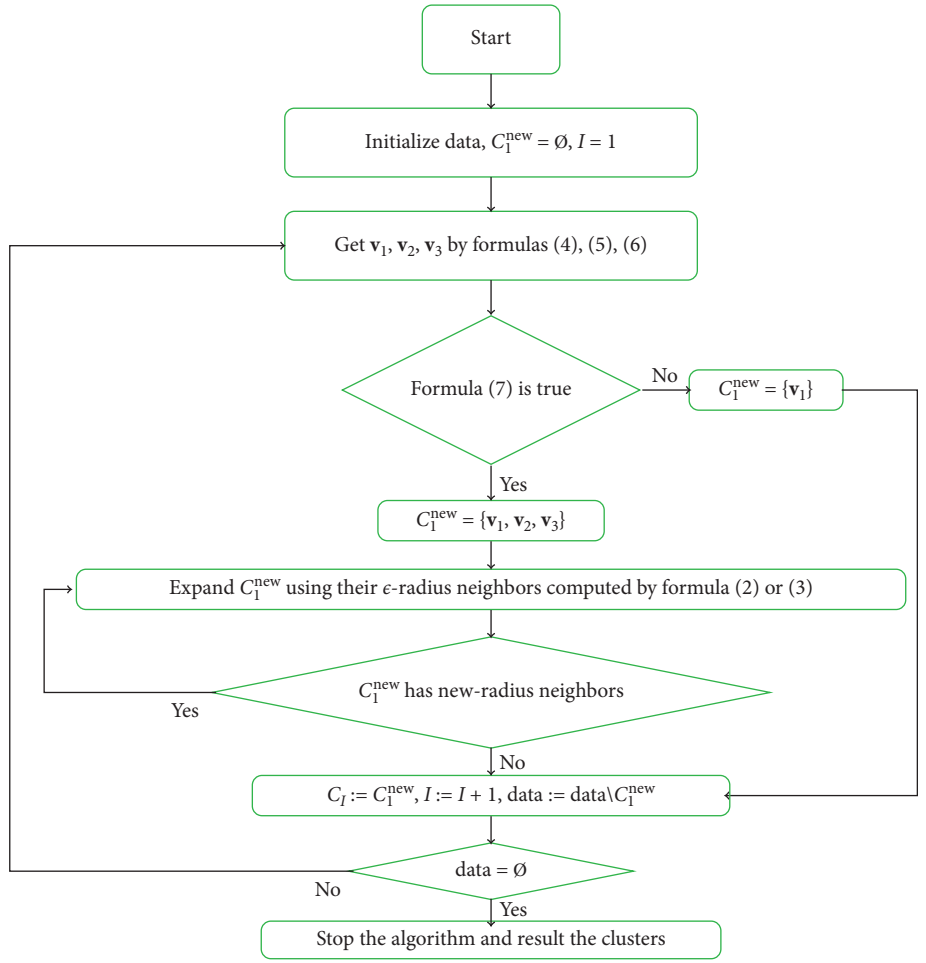


Figure 1: The flowchart of proposed algorithm, where I is the number of iterations.

d-dimensional vector, in general. Certainly, we can calculate Euclidean distance between two d-dimensional points $\mathbf{x}$ and $\mathbf{y}$ using the following formula:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{d} (x_i - y_i)^2}. \tag{9}$$

In addition, because variables measured at different scales do not contribute equally when calculating the distance, the data are normalized into [0, 1] interval using the following formula:

$$z_{ij} = \frac{x_{ij} - \min_i(x_{ij})}{\min_i(x_{ij}) - \min_i(x_{ij})}, \tag{10}$$

where x_{ij} is the value of variable j ($j = \overline{1, d}$) at the point i ($i = \overline{1, n}$), z_{ij} is the normalized value of variable j at point i and $\min_i(x_{ij})$ and $\max_i(x_{ij})$ are the minimum and maximum value of variable j, respectively.

Step 2. For each $\mathbf{v}_i \in C_1^{new}$, compute the adaptive ε-radius and the corresponding ε-radius neighbors S_i using Definition 1 and either formula (2) or formula (3); update C_1^{new} and centroid$_{new}$ by the following formulas:

$$C_1^{new} := C_1^{new} \cup S_i,$$
$$\text{centroid}_{new} := \frac{\text{centroid}_{new}}{v_i}. \tag{11}$$

In this step, formulas (2) and (3) are utilized to compute the adaptive ε-radius and the corresponding ε-radius neighbors S_i. Note that, the two abovementioned formulas are now just some options that need to be tested. In the numerical results, after applying both, the best option will be selected in the application.

Step 3. If $C_1^{new}/C_1^{old} \neq \emptyset$, then $C_1^{old} := C_1^{new}$ and centroid$_{new}$: = centroid$_{new} \cup C_1^{new}/C_1^{old}$. Repeat Step 2 and Step 3 until centroid$_{new} = \emptyset$, then stop the current-processing cluster.

Step 4. Repeat the three steps above until all objects are assigned to their clusters.

The main idea of the proposed algorithm is that from a number of points initialized using formulas (4), (5), and (6)

subject to (7), the cluster can automatically expand based on formulas (2) or (3). When the cluster does not extend more, the abovementioned process will repeat over the rest of the data until all points in the data are assigned to a specific cluster. With formulas (2) or (3), the ε-radius neighbor can adapt to different cluster densities; hence, the proposed algorithm can determine the number of clusters and find clusters of arbitrary shapes in cases of both balanced and imbalanced cluster densities. This is an advantage of the proposed algorithm in comparison to conventional methods, such as k-means, k-medoids, and DBSCAN.

2.2. Study Area and Data Used. The clustering method proposed above will be applied in assessing the underground water quality in Phu My Town, Ba Ria-Vung Tau Province, Vietnam. The study area and data used are described as follows.

2.2.1. Study Area. Phu My town has a natural area of 33,825 hectares and a population of 137,334 people. To the east, it borders Chau Duc district, Ba Ria-Vung Tau province. To the West, it borders Can Gio district, Ho Chi Minh City, and Vung Tau City, Ba Ria-Vung Tau province. To the South, it borders Ba Ria City, Ba Ria-Vung Tau province, and to the North is the Long Thanh district, Dong Nai province. Phu My town is located in the climate region of the Southern Delta, Vietnam, with a tropical climate and is influenced mainly by the northeast and southwest monsoon. There are two distinct seasons in a year, dry season and rainy season. The first lasts from December to April with an average annual temperature of 26.3 Celsius, and the second is between May and November with an average annual rainfall of 1356.5 mm.

Phu My town is the most concentrated industrial area and is one of the most developed areas in Ba Ria-Vung Tau province, Vietnam. To serve economic development, the demand for water in this area is quite high, but the sources of surface water from rivers and lakes do not meet the demand. According to the 2012 survey data of the Department of Natural Resources and Environment of Ba Ria-Vung Tau province, the total volume of underground water exploitation in this town had accounted for 18,608,430 m^3/year (mainly from Phu My-My Xuan water station and Toc Tien Water Plant). Groundwater exploitation has been reported to be mainly in the Pleistocene aquifer, which is composed of coarse-grained soil of Cu Chi Formation, Thu Duc Formation, and Trang Bom Formation with the main minerals: fluorite-apatite, feldspar, gypsum, tourmaline, montmorillonite, ilmenite, and some other impurities.

2.2.2. Data Used. The dataset has been provided by the Department of Natural Resources and Environment of Ba Ria-Vung Tau Province. The groundwater samples in the Middle-Upper Pleistocene (qp_{2-3}) aquifer and Upper Pleistocene (qp_3) aquifer, which consist of 11 variables, have been collected from 17 monitoring wells. The locations of 17 monitoring wells are shown in Figure 2, and the detailed dataset is presented in Table 1.

In this study, the contribution of variables is the same when calculating distance, that is, the proposed method considers the equal importance for each chemical parameter. In case in which some chemical parameters are more important than the others, the proposed method can be performed by using the weighted Euclidean distance instead of using the standard Euclidean distance. Also, note that, in this application, well's location is not considered as a variable, that is, the wells will only be grouped by their chemical parameters. The algorithm thereby will not be too focused on location, but more on chemical properties. Naturally, if wells in the same region have the same chemical properties, they will be assigned to the same cluster. As a result, we have wells sorted by locations. In contrast, through the clustering results, we can still identify wells that are in the same region, but have different chemical properties, or wells that are in different regions, but have similar chemical properties. In such cases, the corresponding explanation will also be provided.

3. Results and Discussion

3.1. Numerical Example. In this section, a simple dataset is used in order to illustrate the proposed algorithm in detail. The dataset consists of 20 bivariate points presented in Table 2; the normalized data points are presented in Figure 3.

Using formulas (4), (5), and (6), we found the three initial points $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ of the first cluster, which are represented by red in Figure 4. It can be seen from Figure 4 that the distance between these three points is really small in comparison with the distance between all points; therefore, condition (7) is satisfied and we can use these three points for extending the cluster.

Now, we use the points in the processing cluster to build up the cluster itself. For example, in Figure 5, starting from the green point, $\mathbf{v}_2$, using formula (3), we calculate the adaptive radius and determine the three new ε-neighbors, based on the circle formed. After that, the processing cluster will be extended by adding these three new points, and the point $\mathbf{v}_2$ will no longer be used to extend the cluster in the next steps. Using another point in the processing cluster, for example, the green point in Figure 6, we also calculate the adaptive radius and determine the new ε-neighbors, based on the circle formed.

Repeat the abovementioned process until the processing cluster cannot be extended more, that is, all points in the processing cluster have been used for the extending process and we cannot find any new points linked to them, as shown in Figure 7.

Figure 7 completely determines the first cluster; we can repeat the abovementioned process for the remainder of the dataset and obtain the final partition, as shown in Figure 8.

3.2. Experiments in Benchmark Datasets. Section 3.1 step-by-step illustrated the proposed algorithm. In this section, to test the partitioning performance of the proposed algorithm

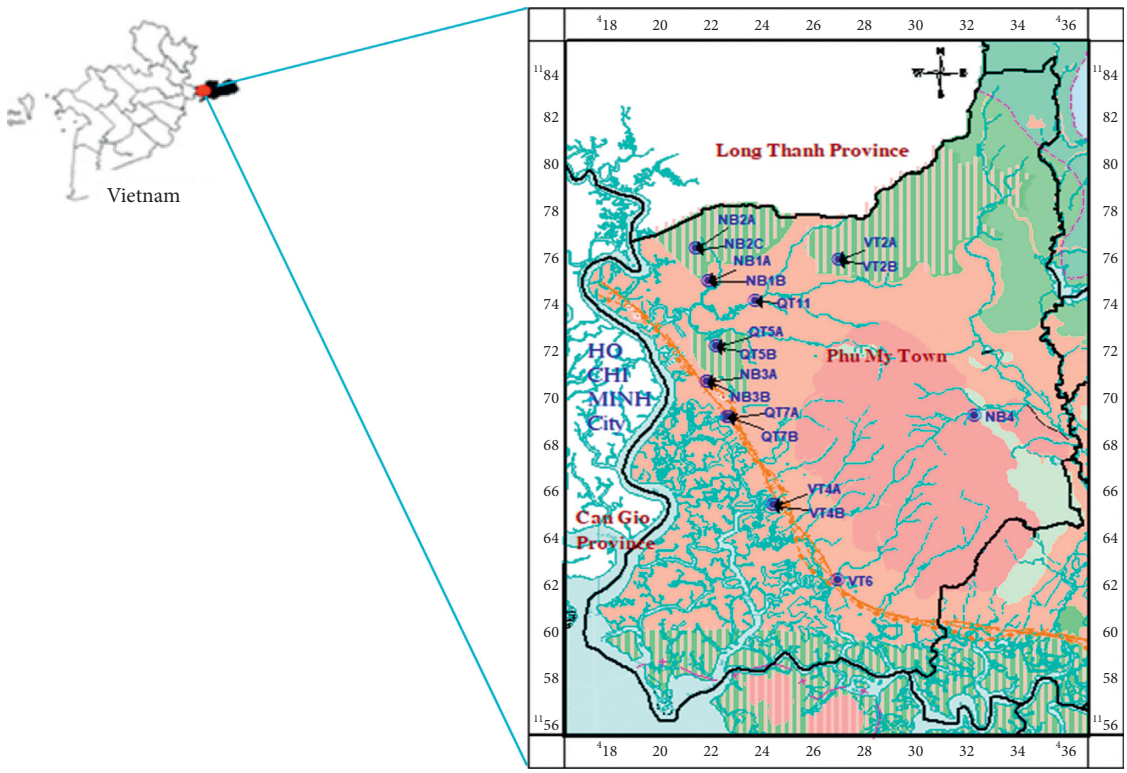


Figure 2: The position of wells.

Table 1: Concentration of chemical parameters (mg/l) collected at wells.

ID	Na^+	K^+	Ca^{2+}	Mg^{2+}	NH_4^+	Al^{3+}	HCO_3^-	Cl^-	SO_4^{2-}	NO_3^-	NO_2^-
NB3A	4.19	1.47	17.03	0.61	0.00	0.00	54.92	8.15	2.40	1.20	0.00
NB3B	6.56	3.85	2.00	0.49	0.00	0.00	6.10	17.73	2.40	0.64	0.00
QT5B	6.57	4.07	17.03	0.61	0.00	0.08	54.92	10.64	9.61	1.49	0.00
QT7B	192.73	9.00	22.04	19.46	2.31	8.51	0.00	375.77	81.65	1.15	0.00
NB2C	8.29	2.90	35.07	1.82	0.00	0.00	103.73	17.73	7.20	11.24	0.00
NB1B	4.14	2.32	1.60	0.24	0.00	0.00	12.20	7.80	2.40	0.41	0.00
VT4B	277.65	17.60	26.05	31.62	2.24	1.67	0.00	514.03	115.27	1.44	0.00
VT6	33.79	7.25	10.02	0.61	0.00	0.00	24.41	49.63	19.21	2.88	7.77
NB4	11.00	1.44	21.04	0.61	0.24	0.00	67.12	14.18	9.61	0.91	0.01
QT5A	10.43	1.22	1.40	0.36	0.00	0.00	6.10	16.66	2.40	7.91	0.00
QT7A	644.44	57.90	100.20	118.56	36.10	0.00	494.26	946.52	528.33	6.23	7.28
NB1A	5.00	5.38	2.00	0.49	0.00	0.00	18.31	8.86	2.40	0.59	0.00
NB2A	3.86	3.61	11.02	0.97	3.72	0.00	48.82	7.09	3.84	0.85	0.32
VT4A	82.73	5.76	54.11	21.89	3.65	0.00	85.43	223.34	31.22	0.81	0.00
QT11	4.89	1.65	1.00	0.24	0.00	0.00	12.20	7.80	0.96	1.32	0.00
VT2B	6.88	1.94	9.02	0.61	0.11	0.00	24.41	12.41	2.40	15.14	0.01
VT2A	4.33	2.14	5.61	1.09	0.04	0.00	18.31	8.15	3.36	11.34	0.01

and compare it with other methods, and the proposed algorithm is implemented on different datasets with different characteristics.

The tested datasets can be downloaded from (https://cs.joensuu.fi/sipu/datasets/), which include

(i) Spiral: a dataset with spiral-shaped clusters

(ii) Aggregation: a dataset with different cluster shapes

(iii) Compound: a compound dataset with different cluster shapes and densities

(iv) Gauss: a dataset simulated in [6] with three Gaussian clusters

The tested algorithms include

(i) ARC1: the proposed method with the adaptive radius defined according to formula (3).

(ii) ARC2: the proposed method with the adaptive radius defined according to formula (4).

(iii) k-mean, DBSCAN: two popular clustering algorithms. The k-means requires an initial number of

Table 2: Illustrated data.

Data	X_1	X_2	Data	X_1	X_2
1	42	72	11	41	58
2	44	71	12	41.5	59
3	46	73	13	42.5	59
4	47	72	14	43	60
5	49	71	15	45	61
6	51	71	16	45.5	61
7	52	70	17	47	61
8	54	69	18	48	61
9	55	68	19	49	61
10	57	67	20	50	60

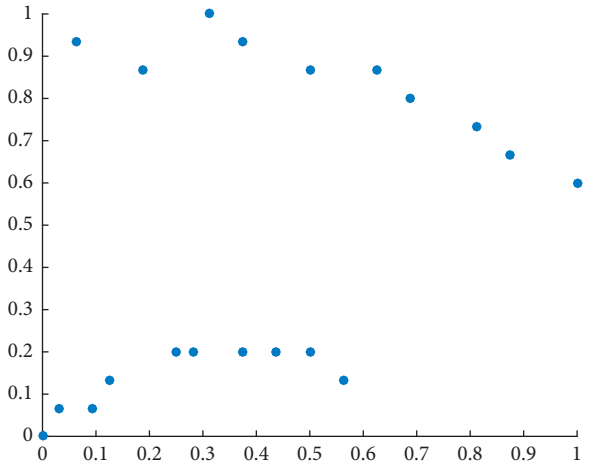


Figure 3: Illustrated data points.

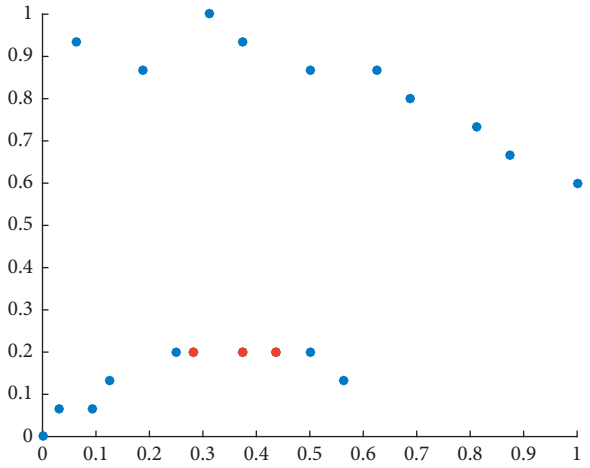


Figure 4: The initial prototypes.

clusters and results in the spherical clusters, while the DBSCAN is a density-based clustering algorithm that is suitable for clusters of arbitrary shapes.

(iv) SU: an automatic clustering algorithm recently presented by [19] for determining the number of clusters, automatically.

In this paper, the Adjusted Rand Index, ARI [29, 30], is employed to evaluate the performance of the five compared

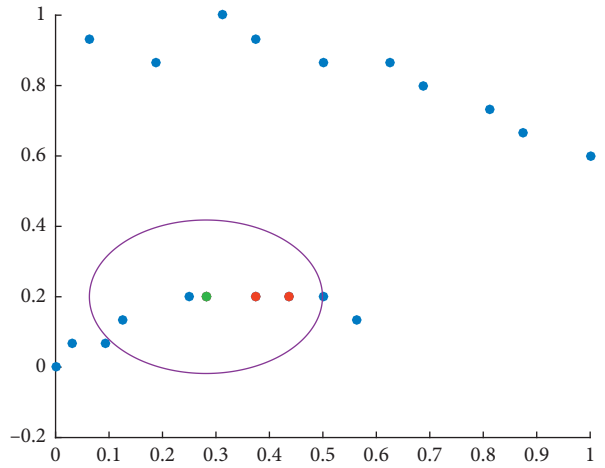


Figure 5: Determining the adaptive radius and new neighbors.

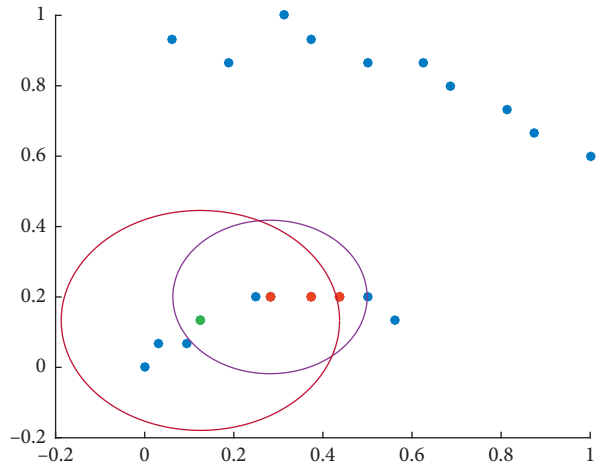


Figure 6: The processing cluster is extended at the green point.

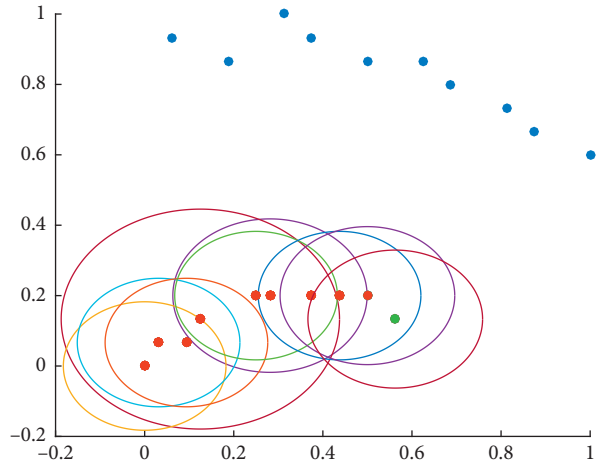


Figure 7: Determining the first clustering.

methods. ARI is an external measure that can make the comparison between the partition produced by a clustering algorithm (P) and the actual partition (Q), where "ground-truth" labeling is known. Particularly, given P and Q, the formulation of ARI is defined as follows:

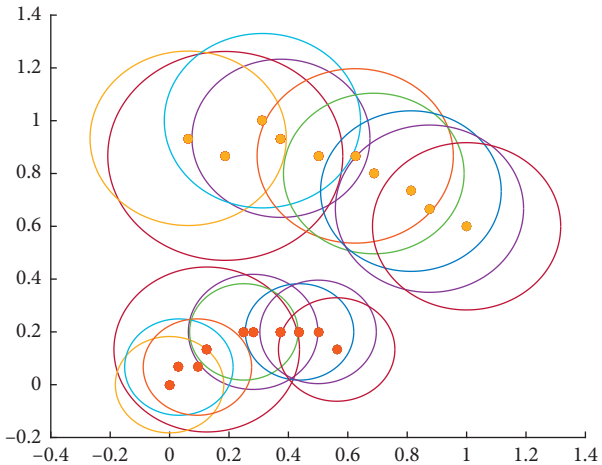
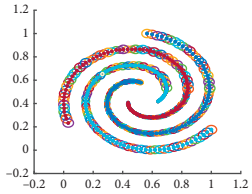
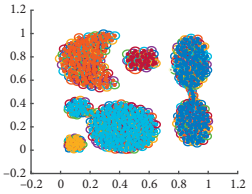
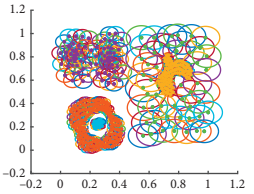
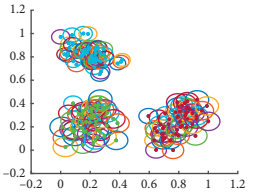
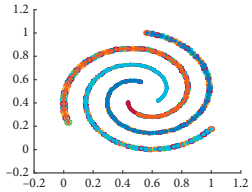
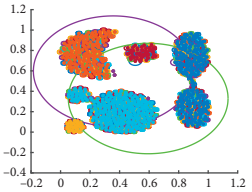
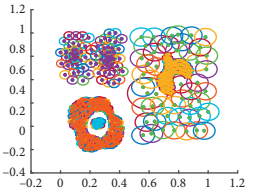
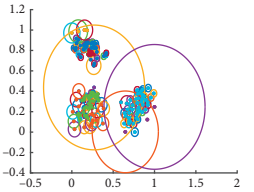
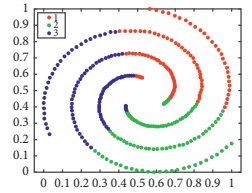
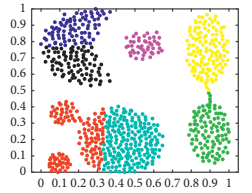
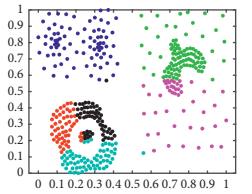
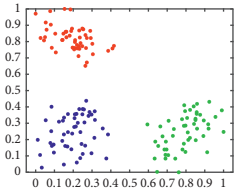
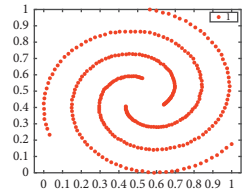
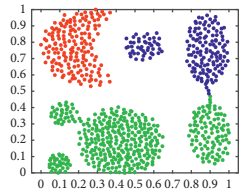
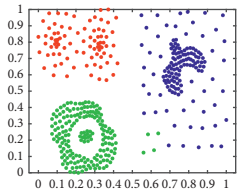
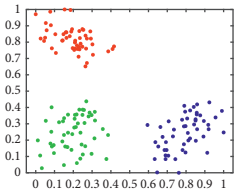
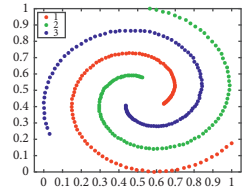
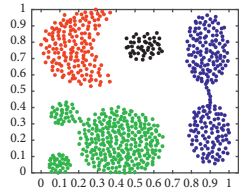
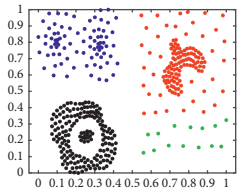
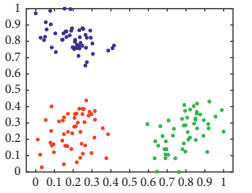


Figure 8: Determining the two clusters.

Table 3: The clusters formed the by the tested algorithms.

	Spiral	Aggregation	Compound	Gauss
ARC1				
ARC2				
k-means				
SU				
DBSCAN				

$$\mathrm{ARI} = \frac{a - (a+c)(a+b)/(a+b+c+d)}{((a+c)+(a+b))/2 - (a+c)(a+b)/(a+b+c+d)}, \quad (12)$$

where a is the number of pairs of elements in the same cluster in P and Q, b is the number of pairs of elements in the same cluster in P, but in different clusters in Q, c is the number of pairs of elements in a different cluster in P, but in the same cluster in Q, and d is the number of pairs of elements in a different cluster in both P and Q. The closer the ARI is to 1, the better the clustering result is (it can be seen from formula (12) that when P and Q are the same, $b = c = 0$ and ARI = 1).

Table 3 intuitively presents the clustering results of the five tested algorithms on the four used datasets.

Remarks:

(i) For the nonspherical clusters, the performance of the DBSCAN is better than that of SU and k-means algorithms. This result is reasonable because DBSCAN can easily group the data points into arbitrary shape clusters, based on the density and the connection rather than the distance between them. ARC2 algorithm, in general, is quite efficient in terms of ARI and outperforms the DBSCAN on two of the three datasets. Meanwhile, the ARC1 achieves the largest ARI values, which indicates the best performance in terms of clustering accuracy.

(ii) For the spherical or Gaussian clusters, most of the methods render good performance, in which ARC1, SU, and DBSCAN are the proper methods. The k-means algorithm also provides the best result, for $k = 3$; however, when k is randomly changed and does not satisfy $k = 3$, this method shows poor performance. Tables 3 and 4 also show that the ARC2 performs better than the k-means; however, it is not good enough for the Gaussian clusters.

(iii) In summary, it can be claimed that ARC1 is an effective algorithm. Specifically, the ARC1 can automatically determine the number of clusters and has notably larger ARI values or notably better clustering results for any given dataset.

3.3. Application for Underground Water Quality Assessment. In this section, we cluster the samples of groundwater quality parameters provided by the Department of Natural Resources and Environment of Ba Ria-Vung Tau Province. The study area and data used have been presented in Section 2. The clustering results in Figure 9 showed that the 17 monitoring wells are classified into 4 groups based on the water quality characteristics:

(i) Cluster 1: NB3A, QT5B, NB4

(ii) Cluster 2: NB3B, NB1B, NB1A, QT11

(iii) Cluster 3: QT7B, NB2C, VT4B, VT6, QT5A, NB2A, VT4A, VT2B, VT2A

(iv) Cluster 4: QT7A

Table 4: The ARI of the comparative methods on the four datasets.

	Spiral	Aggregation	Compound	Gauss
ARC1	**1.0000**	**0.8089**	**0.9438**	**1.0000**
ARC2	0.9253	0.8035	0.9438	0.7034
k-means	0.0924	0.5906	0.5890	0.6968
SU	0.0000	0.5638	0.7257	**1.0000**
DBSCAN	**1.0000**	0.7338	0.7568	**1.0000**

A comparison of some parameters among clusters is shown in Figure 10. We have the following remarks:

(i) Cluster 4 consists of only 1 well, QT7A, with very high parameter values. This result demonstrates that the water quality in this well is really bad compared to the remaining clusters. In addition, it can be seen from Table 1 and Figure 10(a) that QT7A has more salt ions (Mg^{2+}, Na^{+}, K^{+}, Ca^{2+}, HCO_3^{-}, NH_4^{+}, Cl^{-}, SO_4^{2-}, and Nitrite) compared to the remaining clusters. According to National Technical Regulation on Groundwater Quality of Vietnam, the permitted standard for Cl^{-} is 250 mg/l and for SO_4^{2-} is 400 mg/l. Therefore, the Cl^{-} and SO_4^{2-} values of QT7A exceed the permitted standards 3.78 and 1.3 times, respectively. This demonstrates that QT7A may be overaffected by saline intrusion because this well is located near the saline boundary. Additionally, it can be seen in Figure 9 that two wells QT7A and QT7B are located in the same region, but they belong to different clusters. Actually, they are both contaminated wells, but they have different depths, representing separate aquifers. As a result, QT7A exhibits a higher level of contamination than QT7B.

(ii) For the three remaining clusters, it can be seen from Figures 9 and 10(b) that Cluster 1 consists of three wells, with high HCO_3^{-} values. To our knowledge, the two wells, NB3A and QT5B, are located near My Xuan B1 industrial zone, and the well NB4 is located near Toc Tien landfill. As a result, those wells may be contaminated by the waste discharge process of the abovementioned industrial zone and landfill.

(iii) Cluster 2 consists of four wells with relatively good quality. In this cluster, most of the parameter values are lower than those of other clusters and are within safe ranges. It can be concluded that the wells of Cluster 2 are not affected by agricultural activities as well as saline intrusion.

(iv) Cluster 3 consists of eight wells with higher values of Mg^{2+}, Na^{+}, K^{+}, Ca^{2+}, Cl^{-}, and SO_4^{2-} compared to those of Cluster 1 and Cluster 2. Especially, Cl^{-} value exceeds the permitted standard at 2/8 wells. This indicates a number of wells in Cluster 3, which are located near the coast as well as salinity boundaries, are capable of being affected by salinity intrusion. In addition, as shown in Figure 10(b), in Cluster 3, the average value of NO_3^{-} is higher than that of Cluster 1 and Cluster 2. This demonstrates

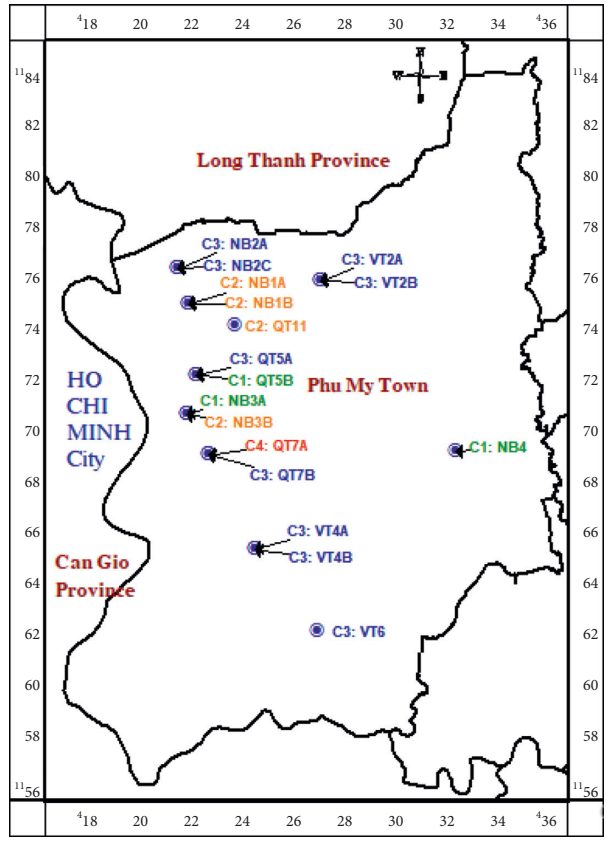


Figure 9: The clustering result for 17 wells.

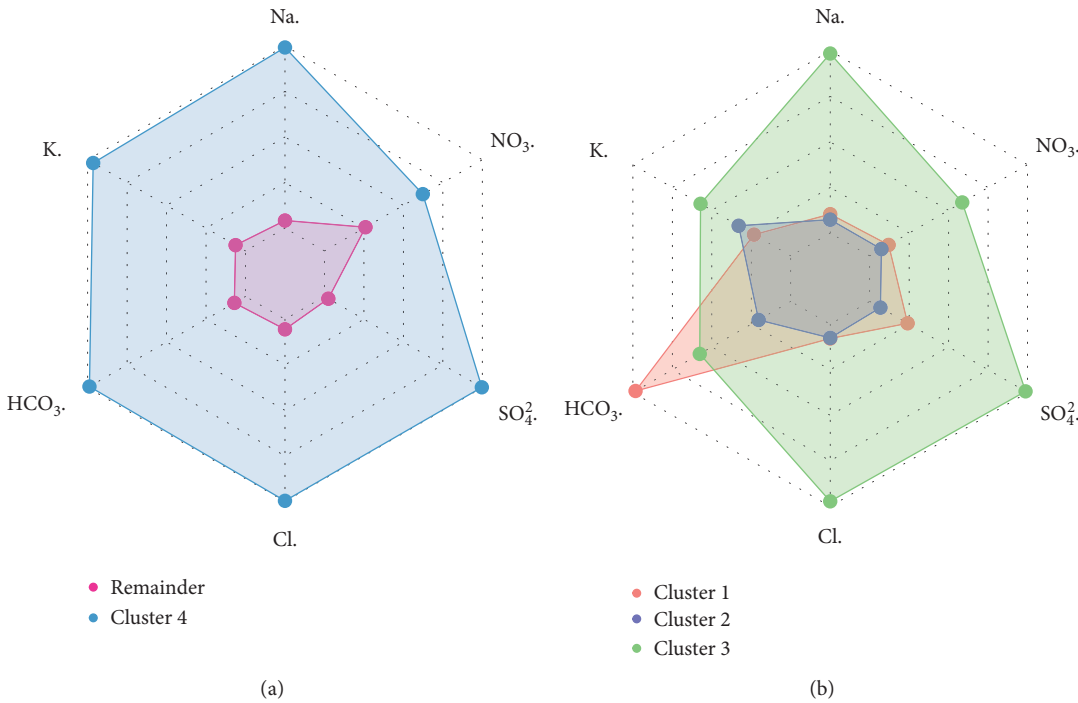


Figure 10: Comparing the parameters among clusters. (a) Cluster 4 and the remaining clusters. (b) Cluster 1, Cluster 2, and Cluster 3.

that agricultural activities taking place around the monitoring area served as large contributors to the underground water quality of this cluster. In particular, well NB2C, VT2B, and VT2A are located near the industrial planting area. Meanwhile, well VT6, which is located near the aquaculture area,

may be seriously affected by organic matter from the residual feed; therefore, the NO_3^- value reaches 7.77 times higher than the permitted standard.

4. Conclusion

Based on the definition of epsilon radius neighbors, this paper has proposed a new clustering algorithm that can automatically determine the number of clusters and can find clusters with different sizes, shapes, and densities. The radius or extending is adapted to the current-processing cluster and has good generalization ability. The proposed algorithm is tested on benchmark datasets and is then applied to underground water quality assessment in Phu My Town, Ba Ria-Vung Tau province, Vietnam. For the experiments with many datasets, the ARC1 algorithm exhibits a better performance than the other tested algorithms in terms of the Adjusted Rand index. The ARC2 algorithm performs better than the conventional clustering algorithms in the case of nonspherical clusters but worse in the case of spherical clusters. For the underground water quality assessment in Phu My Town, Ba Ria-Vung Tau province, Vietnam, the proposed algorithm indicated that there are four clusters of water quality that represent different source contributions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Vietnam National University Ho Chi Minh City (VNU-HCM) under Grant no. C2018-24-01.

References

[1] T. Vo Van and T. Pham-Gia, "Clustering probability distributions," *Journal of Applied Statistics*, vol. 37, no. 11, pp. 1891–1910, 2010.

[2] T. VoVan and T. NguyenTrang, "Similar coefficient for cluster of probability density functions," *Communications in Statistics—Theory and Methods*, vol. 47, no. 8, pp. 1792–1811, 2018.

[3] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Computing Surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[4] A. K. Jain, "Data clustering: 50 years beyond *K*-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.

[5] Z. Xie, R. Dong, Z. Deng et al., "A probabilistic approach to latent cluster analysis," in *Proceedings of the Twenty Third International Joint Conferences on Artificial Intelligence (IJCAI)*, Beijing, China, August 2013.

[6] T. VoVan and T. Nguyen Trang, "Similar coefficient of cluster for discrete elements," *Sankhya B*, vol. 80, no. 1, pp. 19–36, 2018.

[7] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, Oakland, CA, USA, June 1967.

[8] D. Arthur and S. Vassilvitskii, "*k*-means++: the advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1027–1035, Society for Industrial and Applied Mathematics, New Orleans, LA, USA, January 2007.

[9] G. Karypis, E.-H. Eui-Hong Han, and V. Kumar, "Chameleon: hierarchical clustering using dynamic modeling," *Computer*, vol. 32, no. 8, pp. 68–75, 1999.

[10] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: the fuzzy *c*-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[11] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *Journal of Cybernetics*, vol. 4, no. 1, pp. 95–104, 1974.

[12] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 841–847, 1991.

[13] J. C. Bezdek and N. R. Pal, "Cluster validation with generalized Dunn's indices," in *Proceedings of the 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems*, pp. 190–193, Dunedin, New Zealand, November 1995.

[14] R. Babuška, *Fuzzy Modeling for Control*, Springer Science & Business Media, Berlin, Germany, 2012.

[15] L. Mavridis, N. Nath, and J. B. Mitchell, "PFClust: a novel parameter free clustering algorithm," *BMC Bioinformatics*, vol. 14, no. 1, p. 213, 2013.

[16] L. E. Agustı, S. Salcedo-Sanz, S. Jiménez-Fernández, L. Carro-Calvo, J. Del Ser, and J. A. Portilla-Figueras, "A new grouping genetic algorithm for clustering problems," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9695–9703, 2012.

[17] S. Das, A. Abraham, and A. Konar, "Automatic clustering using an improved differential evolution algorithm," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, vol. 38, no. 1, pp. 218–237, 2008.

[18] T. Vo-Van, T. Nguyen-Thoi, T. Vo-Duy, V. Ho-Huu, and T. Nguyen-Trang, "Modified genetic algorithm-based clustering for probability density functions," *Journal of Statistical Computation and Simulation*, vol. 87, no. 10, pp. 1964–1979, 2017.

[19] W.-L. Hung and J.-H. Yang, "Automatic clustering algorithm for fuzzy data," *Journal of Applied Statistics*, vol. 42, no. 7, pp. 1503–1518, 2015.

[20] M. Ester, H.-P. Kriegel, J. Sander et al., "A density-based algorithm for discovering clusters in large spatial databases with noise," *KDD*, vol. 96, no. 34, pp. 226–231, 1996.

[21] C. Dobbins and R. Rawassizadeh, "Towards clustering of mobile and smartwatch accelerometer data for physical activity recognition," *Informatics*, vol. 5, no. 2, p. 29, 2018.

[22] G. F. Tzortzis and A. C. Likas, "The global kernel *k*-means algorithm for clustering in feature space," *IEEE Transactions on Neural Networks*, vol. 20, no. 7, pp. 1181–1194, 2009.

[23] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel *k*-means: spectral clustering and normalized cuts," in *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD'04*, pp. 551–556, ACM, New York, NY, USA, August 2004.

[24] M. O. Mavukkandy, S. Karmakar, and P. S. Harikumar, "Assessment and rationalization of water quality monitoring network: a multivariate statistical approach to the Kabbini

River (India)," *Environmental Science and Pollution Research*, vol. 21, no. 17, pp. 10045–10066, 2014.

[25] U. N. Kura, F. M. Ramli, N. W. Sulaiman, S. Ibrahim, A. Aris, and A. Mustapha, "Evaluation of factors influencing the groundwater chemistry in a small tropical island of Malaysia," *International Journal of Environmental Research and Public Health*, vol. 10, no. 5, pp. 1861–1881, 2013.

[26] S. Shrestha and F. Kazama, "Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan," *Environmental Modelling & Software*, vol. 22, no. 4, pp. 464–475, 2007.

[27] M. Varol and B. Ş r, "Assessment of nutrient and heavy metal contamination in surface water and sediments of the upper Tigris River, Turkey," *Catena*, vol. 92, pp. 1–10, 2012.

[28] Q. Yang, J. Zhang, Y. Wang, Y. Fang, and J. Martin, "Multivariate statistical analysis of hydrochemical data for shallow ground water quality factor identification in a coastal aquifer," *Polish Journal of Environmental Studies*, vol. 24, 2015.

[29] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 846–850, 1971.

[30] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985.