IOP Conference Series: Earth and Environmental Science

OPEN ACCESS

Developing web-based data analysis tools for precision farming using R and Shiny

To cite this article: Ebrahim Jahanshiri and Abdul Rashid Mohd Shariff 2014 *IOP Conf. Ser.: Earth Environ. Sci.* **20** 012014

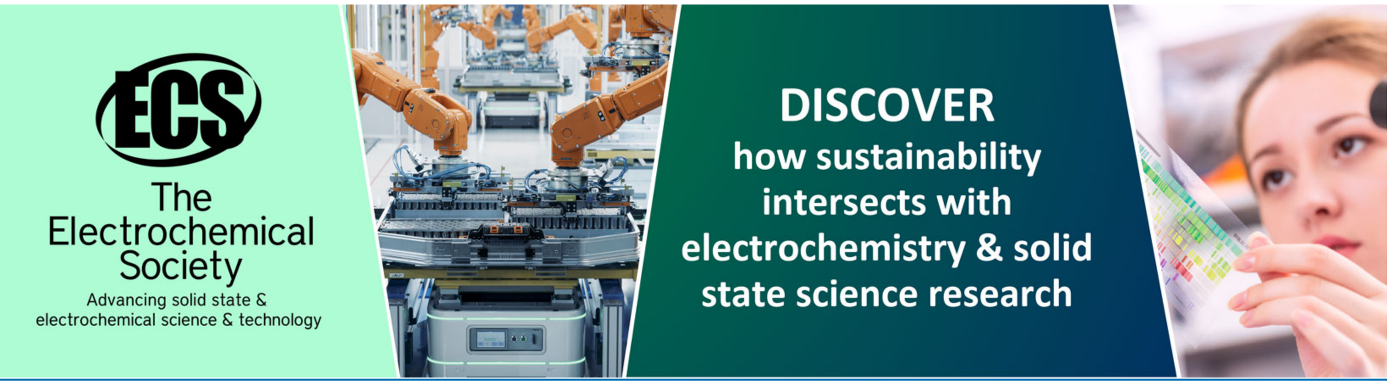
View the article online for updates and enhancements.

You may also like

- Development of an Irrigation System for Predicting Watering Time with ANFIS Method for Chili Plants
T George Michael, Mardi Turnip, Erni Muniarti et al.

- Spatial data gathering architecture for precision farming using web services
C Fibriani, A Ashari and M Riasetiawan

- Application of precision farming technologies for chemical protection of grain crops
F A Kipriyanov, D V Shemnyakov, P A Savinykh et al.



This content was downloaded from IP address 18.216.186.164 on 25/04/2024 at 06:59

IOP Conf. Series: Earth and Environmental Science **20** (2014) 012014 doi:10.1088/1755-1315/20/1/012014

Developing web-based data analysis tools for precision farming using R and Shiny

Ebrahim Jahanshiri, Abdul Rashid Mohd. Shariff

Geospatial Information Science Research Centre (GIS RC)
Faculty of Engineering
Universiti Putra Malaysia (UPM)
43400 Serdang, Selangor
Tel:603-89466424; Fax:603-89466425

e.jahanshiri@gmail.com; rashid@eng.upm.edu.my

Abstract. Technologies that are set to increase the productivity of agricultural practices require more and more data. Nevertheless, farming data is also being increasingly cheap to collect and maintain. Bulk of data that are collected by the sensors and samples need to be analysed in an efficient and transparent manner. Web technologies have long being used to develop applications that can assist the farmers and managers. However until recently, analysing the data in an online environment has not been an easy task especially in the eyes of data analysts. This barrier is now overcome by the availability of new application programming interfaces that can provide real-time web based data analysis. In this paper developing a prototype web based application for data analysis using new facilities in R statistical package and its web development facility, Shiny is explored. The pros and cons of this type of data analysis environment for precision farming are enumerated and future directions in web application development for agricultural data are discussed.

1. Introduction

World Wide Web has given new opportunities for making static and dynamic websites. With the advent of new programming languages like HTML5 [1] and JavaScript [2], making user-friendly web applications [3] [4] [5] is getting easier than ever. These advancements lead to the availability of many web applications that has truly influenced real time decision making. Data driven applications have also been recently used to facilitate data analysis online. Web Geo-porocessing applications [6] however have recently gained momentum in the business and research community. True geo-processing online which consists of accepting user data, processing it the way user requests and showing the results in a coherent and easy to understand manner has not been widely available to the developers due to complexity of data, process and visualization. Note that web mapping is different from web geo-processing that focuses on static online map delivery. In this case the maps pre-produced and the user can request them online. Most of the online mapping website including OpenGIS Web Map Service Interface Standard (WMS) is also based on web mapping concept [7] [8].

ESRI ArcGIS has been one of the pioneers in providing the online web mapping services based on Microsoft technologies [9]. Other companies have followed the lead and have provided facilities that are mainly tied to huge commercial data base systems and windows web application interfaces.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Google has recently provided its web mapping interfaces that will let the developers and users to map their own events and applications over the already existent Google maps and earth [10]. Most of these applications however lack the complexity that is needed by the geo-processing operations.

In the open sources world, there have been attempts to provide web geo-processing services using web clients and Python language (as web server) but that initiative did not become widespread mainly because Python does not have dedicated facilities for the level of statistical analysis of geographic data. With the advent of R open source statistical programming language [11] and gaining popularity among the spatial analysis community, and also the introduction of very useful "Shiny" package [12] that lets the programmers to create applications online, a new opportunity has been shown itself for creating web geo-processing services .

Precision farming or precision agriculture is highly reliant on the accurate and easy to access maps for management of variability in the field. Interpolation techniques used in the map making for precision farming are mostly using complex mythologies and therefore any application that attempts to provide the users such facility is limited to only a few options that are available within the software [13][14]. For example low-level ArcGIS interpolation elements are not available to the users to tweak. The rich data analysis facilities in R gives freedom to the user to access low level parameters that are necessary in order to create accurate maps.

This paper introduces an attempt to make an online web geo-processing service and mapping for precision farming using common interpolation and visualization techniques available in R.

2. Methodology

2.1. System Architecture

The interface is consisted of two parts; web interface and web server. Both of these components are controlled by the code that is written within the frame work of Shiny application in R. The building block of Shiny package is based on reactive programming [15]. Since the major task of the web application is to get the inputs and produce outputs, the whole programming language is designed so that a change in any input whether it is input data or method parameters will change the end result. The change will immediately be reflected into the form of texts, tables or figures. This is an advantage for the web applications that rely on user inputs. The shiny procedure can provide different outputs without the need to refresh the web page.

Within the shiny package, ordinary controllers or widgets are provided for ease of use of application programmers. Many of the procedures like uploading files and refreshing the page for drawing new plots and tables are provided automatically. These tasks are done based on the pre-built output widgets. The communication between the web interface and R is done with fast websockets package [16]. Websockets are exclusively important in situations where there are constant back and forth dialogue or data exchange between the clients and servers. Figure 1 provides general overview of the system components.

The communication between the client and server is done over the normal TCP connection. The bulk of live data traffic that is needed for many of web applications (i.e. online games) between the browser and the server is facilitated over the websockets protocol. This protocol operates separately and only handshake between the client and server is done over the HTTP protocol. The duplex connection is open all the time and therefore the authentication is not needed when exchange is done. The websocket is currently supported by many modern browsers.

2.2. Web Interface

The web interface provide means for the user to upload files, change the file reading format and provide necessary parameters to the R engine to produce the results and then reflecting the results in a consistent manner.

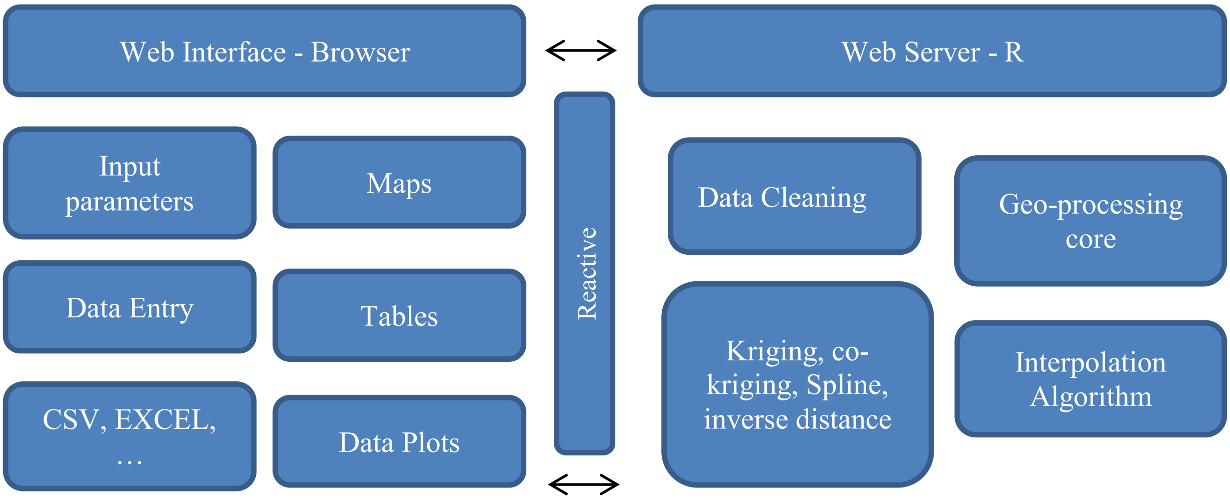


Figure 1: System architecture for the shiny application.

2.2.1 Data

The mapping procedure will start by providing the input data. Since the data is geographic by nature, the coordinates (x and y or Northing and Easting) should be part of the data. Also at least one target variable should be present in the data so that the interpolation can be done for that variable. Currently the input format is comma separated values (CSV). This is a very common format for storing the tabular data and is recognized by R in addition to the commercial data analysis software like Microsoft Excel. The support for other types of data inputs can be added in the future given the flexibility of R to accept many types of data.

Some elementary data screening like checking variable names, checking for duplicates in the coordinate points are done in this stage. In this stage the data also will be converted to specific spatial point data or SpatialPointsDataframe object [17], [18] so that spatial methods in R can take advantage of these formats. Most of these activities are done on the server side of the application and therefore hidden from the user.

2.2.2 Exploratory Analysis.

In this section normal exploratory spatial data analysis is done on the target of interest. As can be seen in Figure 2, the user must choose a target variable and the histogram, boxplot and distribution of variable (bubble plot) is drawn in the web page. The plot visualizations are done using elegant methods in the ggplot2 package [19]. Using these plots the user can decide whether to use any specific measure in the interpolation like transformation of target variable.

2.2.3 Mapping. In this section the result of interpolation using the standardized kriging for the selected variable is shown in the form of a map with a legend showing different levels of concentrations. The interpolation algorithm is currently limited to inverse distance weighting and geostatistical kriging (ordinary kriging). Other interpolation algorithms like co-kriging, universal and regression kriging will be implemented in the future. Due to extensive methods and procedures that are available in R, the program can become a very sophisticated interpolation web app for everyone who is willing to make a map for agricultural or geological variables whenever the data is available.

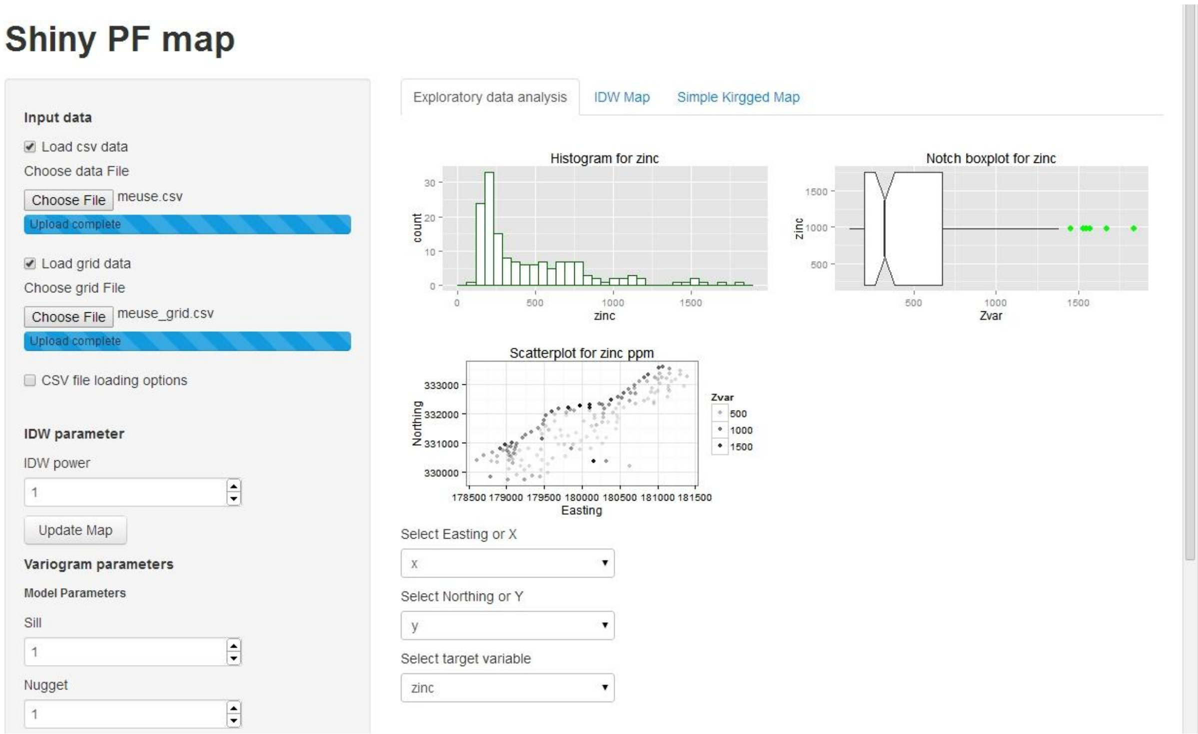


Figure 2: Exploratory spatial data analysis for the given target variable.

Figure 3 shows the mapping panel for kriging method. The variogram of the target variable in conjunction with the thematic map is shown in this panel. The advantage of this system is that the user can easily change the variogram and kriging variable and almost instantly see the results on both variogram fit and the resulting map. This will give an invaluable tool to analysts to fine tune the variogram and therefore the map. Given the vast potential of R, extending this application to encompass even minute parameters in modelling and mapping is also possible.

3. Future Direction

R as an extensive statistical programming language has provided suitable means for the spatial analysis. These methods are created and enhanced by the researchers and are subject to improvement and enhancement by a global network of researchers who test these methods. Many of the methods that are available in R for analysis are not implemented in the proprietary GIS software world. The implementation of shiny package unleashed this potential for web geo-processing applications. This web mapping application will provide basis to further enhance the analysis methods. The interpolation and mapping will be extended to encompass many new and older methods for spatial analysis of point data. New data screening facilities will be added in the future so that using different data methods (from SPSS files to ESRI shapefiles and geodatabases) will be possible with the Shiny PF app. Also providing on-the-fly projection for the data and then projecting the data and the resulting map on the Google maps or Open Street maps will be considered in the future so that the user will have a better understanding of the data and map. Further visualization plots for the data can also be added to this web app to enrich its utility everywhere.

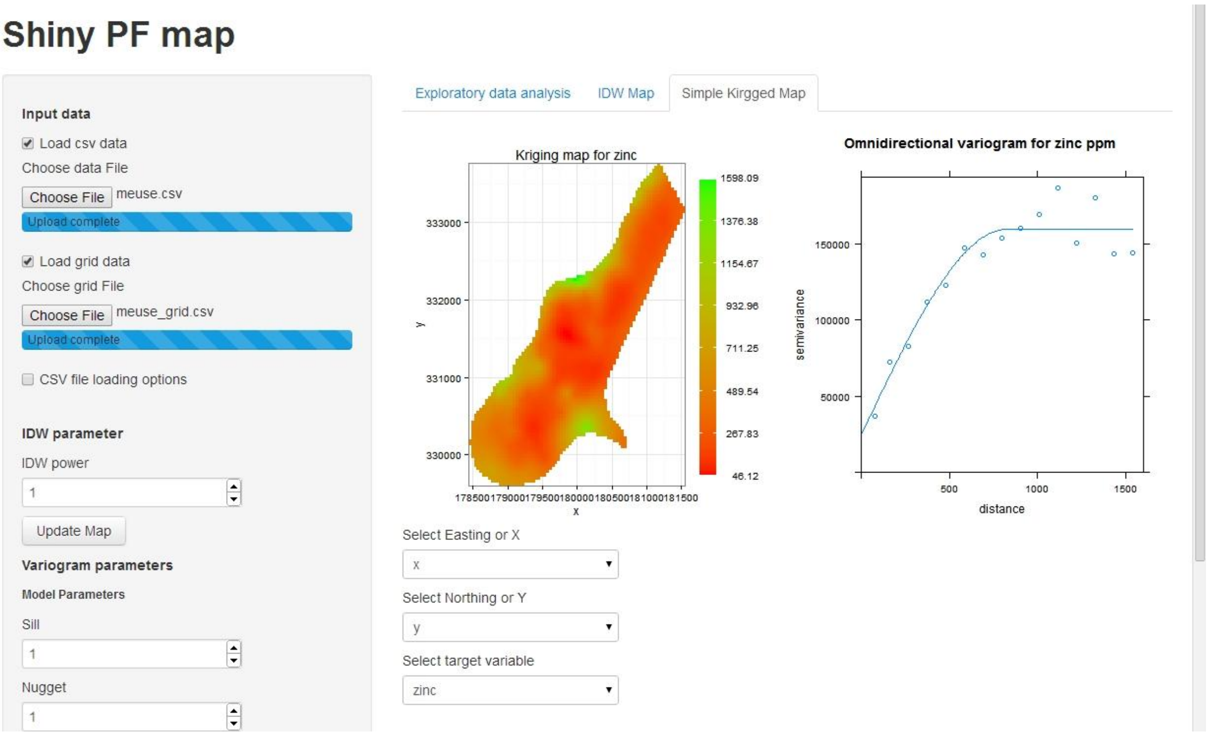


Figure 3: Mapping panel for ordinary kriging method.

4. Conclusion

In this paper we set out to present the feasibility of creating a web mapping application for precision farming data. Because of the variability in the field, it is important for the farmers and managers to be able to view the data visually. Because of the nature of agricultural data, that is geographical, any mapping and statistical analysis and has an added complexity. Given the vastness of methods that are available in R statistical package for analysis of spatial data and freedom of the analyst to analyse and map theses data, developing based upon R has tremendous advantage. The newly available shiny package from RStudio that lets the developer inter-connect R statistical system to the web interface has opened new horizons. Developing a complete data analysis and mapping solution however does not come without challenges. Coordinate systems transformations, variety of data formats, inaccurate and missing data and complexity of some of the interpolation algorithms are among some of them. These challenges can be overcome as data collections formats are becoming standardized and as new upgrades on the statistical methods and web interface are becoming available to the mainstream data analysts.

References

[1] B. Lawson and R. Sharp, *Introducing HTML5 (Google eBook)*. New Riders, 2011, p. 312.
[2] D. Flanagan, *JavaScript Pocket Reference (Google eBook)*. “O’Reilly Media, Inc.,” 1998, p. 89.
[3] D. Nations, “Web Applications,” *About.com*.
[4] M. Calore, “How Do Native Apps and Web Apps Compare?,” *Wired.com*.
[5] P. Fraternali, “Tools and approaches for developing data-intensive Web applications: a survey,” *ACM Comput. Surv.*, vol. 31, no. 3, pp. 227–263, Sep. 1999.

[6] P. Zhao, F. Lu, T. Foerster, and P. Yue, "The Geo-processingWeb," *Comput. Geosci.*, vol. 47, pp. 3–12, 2012.
[7] "Web mapping - Wikipedia, the free encyclopedia." [Online]. Available: https://en.wikipedia.org/wiki/Web_mapping.
[8] "Web Processing Service | OGC®." [Online]. Available: http://www.opengeospatial.org/standards/wps. [Accessed: 05-Oct-2009].
[9] ESRI, "ArcGIS Help 10.1 - What is a geo-processingservice?" ESRI, 2014.
[10] "Google Maps API Web Services - Google Maps API Web Services — Google Developers." [Online]. Available: https://developers.google.com/maps/documentation/webservices/.
[11] R Development Core Team, "R: A Language and Environment for Statistical Computing." Vienna, Austria, 2010.
[12] RStudio and Inc., "shiny: Web Application Framework for R." 2013.
[13] A. Fauzul, A. Aziz, A. Rashid, M. Shariff, M. Amin, A. A. Rahim, E. Jahanshiri, N. Norasma, C. Ya, and F. Azhan, "GIS based System for Paddy Precision Farming," in *World Conference on Agricultural Information and IT*, 2008, no. Did, pp. 417–422.
[14] C. Y. N. Norasma, A. R. M. Shariff, E. Jahanshiri, M. S. M. Amin, S. Khairunniza-Bejo, and A. R. Mahmud, "Web-Based Decision Support System for Paddy Planting Management," *Pertanika J. Sci. Technol.*, vol. 21, no. 2, pp. 343–364, 2013.
[15] Z. Wan and P. Hudak, "Functional reactive programming from first principles," *ACM SIGPLAN Not.*, vol. 35, no. 5, pp. 242–252, May 2000.
[16] "Introducing Shiny: Easy web applications in R | RStudio Blog." [Online]. Available: http://blog.rstudio.org/2012/11/08/introducing-shiny/.
[17] R. S. Bivand, E. J. Pebesma, and V. Gómez-Rubio, *Applied Spatial Data Analysis with R*, 1st ed. Springer, 2008, p. 390.
[18] E. Pebesma and R. Bivand, "Classes and methods for spatial data in R," *R News*, vol. 5, no. 2, pp. 9 – 13, 2005.
[19] H. Wickham, *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.