Ecological Informatics 26 (2015) 192–202

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Increasing dependence on foreign water resources? An assessment of trends in global virtual water flows using a self-organizing time map

Stephanie Clark [a,*], Peter Sarlin [b,d], Ashish Sharma [c], Scott A. Sisson [a]

[a] School of Mathematics and Statistics, University of New South Wales, Sydney, Australia
[b] Center of Excellence SAFE, Goethe University, Frankfurt, Germany
[c] School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia
[d] RiskLab Finland at IAMSR, Abo Akademi University and Arcada University of Applied Sciences, Finland

ARTICLE INFO

Article history:
Received 30 November 2013
Received in revised form 1 May 2014
Accepted 28 May 2014
Available online 7 June 2014

Keywords:
Virtual water
Renewable water resources
Self-organizing map
Self-organizing time map
Temporal dynamics
Cluster visualization

ABSTRACT

Water resources are continually redistributed across international borders as a result of virtual water flows associated with global trade, where 'virtual water' is the term describing water used in the production of commodities. This transfer of virtual water allows some countries to rely heavily on the water resources of other countries without having to transport the water itself. This paper conducts an investigation into the relationship between international virtual water flows and domestically available renewable water resources for a number of countries, to determine trends in national dependencies on foreign water resources over time. Countries with similar states of dependence are clustered, and changes in these clusters are tracked from 1965 to 2010 to determine country-specific and global trends. We make use of a temporal version of the self-organizing map (SOM), the self-organizing time map (SOTM), which provides the means for visualizing structural changes in spatiotemporal data. The SOTM is investigated through a second-level clustering to visualize emerging, changing and disappearing clusters in the data. A post-processing technique is introduced to facilitate interpretation of individual country trends on the SOTM. This study reveals a global trend towards an increased dependence on foreign water resources between 1965 and 2010. The method presented in this study is a workflow tool that results in a visualization of countries with similar and diverging trends of water resource dependencies. This tool can be used to inform national trade, water resources, and environmental management decisions which must take international hydrologic connectivity into account. The sustainability of current virtual water trade and water use trends can be examined with respect to the level of water scarcity experienced by individual and groups of countries.

1. Introduction

The water resources of a country can be significantly impacted by cross-border virtual water flows as a result of international trade (Hoekstra & Mekonnen, 2012). Virtual water is defined as the water used in the production of commodities, such as the quantity of water required to produce a tonne of apples or cereal (Allan, 1998). The transport of trade items across borders can convey large quantities of water virtually 'embedded' in the traded items (without requiring transport of the water itself). This connects the water resources of separate countries, leading to an effective redistribution of water resources between countries and significantly affecting the dispersal of global water resources (Hoekstra et al., 2011; Tamea et al., 2013).

In general, agricultural production is the largest contributor to global water use and pollution (92%) (followed by industrial production (4.4%) and domestic water supply (3.6%)) (Hoekstra & Mekonnen, 2012). Consequently, it is not surprising that the international food trade results in large fluxes of virtual water across international borders (Hoekstra & Mekonnen, 2012). It is less complicated to import crops than to import the water required to grow them, therefore domestic water resources can be supplemented by importing water-intensive food from more hydrologically advantaged regions (Allan, 1998). In this way, countries have the ability to make use of more water than they have available domestically due to the influx of virtual water through imported agricultural products (Suweis et al., 2013).

When combined with an investigation of the naturally occurring available water resources within a country, the consideration of virtual water flows allows for an appraisal of a nation's actual water scarcity (Ercin et al., 2013). In this paper, we will investigate the relationship between virtual water imports through agricultural trade and the available domestic renewable water resources for a set of countries, in

* Corresponding author at: School of Mathematics and Statistics, UNSW, Sydney, NSW 2052, Australia.
E-mail address: stephanie.clark@student.unsw.edu.au (S. Clark).

http://dx.doi.org/10.1016/j.ecoinf.2014.05.012
1574-9541/

order to explore trends in dependencies on foreign water resources. Countries with high virtual water imports and low internal water resources will be considered relatively dependent on foreign water resources. These dependencies will naturally evolve over time as a result of changing national circumstances, policies, consumption patterns, and environmental factors. Considering the dynamic patterns of water dependencies in a global context will allow for the exploration and comparison of the trends of individual countries.

1.1. Related work

The literature provides several applications of statistical methods to the virtual water trade network. Konar et al. (2011) used complex network theory to investigate virtual water trade connections as a framework for network optimization, creating a model of nodes (countries) and links (virtual water flows). This study highlighted how individual countries fit into the global structure of the virtual water trade at a certain point in time. Carr et al. (2012) investigated the connections of the virtual water network using trade matrices, to describe changes in the flows to and from specific countries over time. Tamea et al. (2013) showed trends in the virtual water balance on a national basis for a selection of countries. Suweis et al. (2013) calculated the 'carrying capacity of nations' based on the domestic water currently used in food production compared with the virtual water imports of each country. Whilst current research is focused on quantifying national water footprints (that is the total volume of water used to produce goods and services consumed by a country's population) (Hoekstra & Mekonnen, 2012; Mekonnen & Hoekstra, 2011), and investigating flows between countries (Carr et al., 2012; Konar et al., 2011; Tamea et al., 2013), there is a gap in research relating virtual water flows to available domestic water resources. Investigating this relationship will enable an assessment of the actual state of a country's reliance on external water resources at specific points in time.

We consider the self-organizing map (SOM) and its temporal extension, the self-organizing time map (SOTM), to be useful exploratory approaches for this investigation, due to the difficulty in quantifying links between hydrological and other (in this case, trade) data or comparing data collected using different methods in different countries (UN Water, 2009). The SOM enables dimension and data reduction of a complex dataset through projection and clustering, and has previously been used to illustrate refined relationships between countries (Kaski & Kohonen, 1996). Countries can be grouped, and inferences drawn on their relative attributes with respect to other countries internationally, leading to the use of SOMs as a decision support tool (Kaski & Kohonen, 1996; Shanmuganathan et al., 2006). The SOM has been increasingly used in water resource applications over the past decade (Kalteh et al., 2008), but although water resource data often contain a temporal component, investigations frequently focus on either spatial structure or temporal structure, not allowing for an assessment of the changes in spatial structure over time. As the virtual water network is extremely dynamic (Carr et al., 2012), it is important to understand not only spatial, but also temporal, trends in the data. The literature has provided a number of approaches for incorporating time in SOMs (Chappell and Taylor, 1993; Guo et al., 2006; Kohonen, 1988; Kohonen, 1991). But these SOM-based approaches are not aimed at visualizing temporal changes in cluster structures, which is the key aim of the SOTM (Sarlin, 2013). The SOTM includes time as a dimension of the map, thereby providing insight into the trends of the data over time. In this study, the use of the SOTM will enable assessment of which groups of countries have experienced similar transformations in their dependence on foreign water resources over the timeline of the study.

1.2. This paper

The key focus of this paper is to develop a decision support tool to study the changing dependencies of countries on foreign water resources over time, as decisions regarding national water resources must account for international hydrologic connectivity (UN Water, 2009). Recently, interest in applying the concept of virtual water fluxes to government policy has grown, recognizing the need to understand the effects of trade on water resources. Increased understanding will produce better-informed management decisions, which have conventionally relied only on domestic water use statistics (Hoekstra & Mekonnen, 2012).

This tool is exploratory in nature, as it aims to provide visual insights into data that are evolving over time. It seeks to cluster countries based on their state of dependence on foreign water resources and to investigate how this cluster structure has progressed. In order to achieve this, we develop extensions to the existing SOTM framework. In particular the response to missing values in the data (which are common in hydrological datasets) is modified, and a post-processing technique is developed to disentangle the trends of individual countries on the SOTM.

The paper is structured as follows: Section 2 provides a description of the SOM and the SOTM algorithm, along with accompanying clustering and visualization tools; Section 3 presents the data and implementation used in this study; a discussion of the results follows in Section 4; and a conclusion in Section 5.

2. Methods

2.1. The self-organizing map

2.1.1. Overview

The SOM is an unsupervised learning algorithm from the family of artificial neural networks, used for defining and visualizing non-linear relationships for high-dimensional, multivariate systems. Through training of the SOM with a data set, a topology-preserving mapping of the data is produced from a high-dimensional input space to a low-dimensional output grid (Kohonen, 1998). A key benefit of the SOM is the ability to extract unseen patterns from large quantities of data without requiring an explicit understanding of the underlying relationships.

The SOM grid is first initialized based on the overall structure of the input data set (the training data), and then trained based on the individual input data items. The initialized map units assume linearly spaced values along a set of axes aligned with the eigenvectors corresponding to the principal components of the input data space (Kohonen, 1998). As a primary purpose of the SOM is to provide visualization of a data set (Kohonen, 1998), 1D or 2D output grids are usually used.

The training of the SOM consists of applying two iterative processes: selection of the best map unit to match each item of input data, and updating of the map to better represent the input data. These processes seek the optimal map structure to represent the form of the input data (Kohonen, 1998). During the selection step of training, the map node that best matches each item of input data is selected (the best matching unit, or BMU) based on minimum Euclidean distance. In the updating stage, the BMU and its neighbouring map units (within a specified neighbourhood radius) move to become closer to the input. The neighbourhood radius decreases with each iteration of selection and updating, producing a smoothed final map. For a 2D SOM, Hastie et al. (2009) encourages the reader to consider the map units as buttons that have been sewn in a regular pattern onto the 2D principal component plane of the input data (the input data may be in two or more dimensions), and the training process of the SOM bends and stretches the plane until the buttons best approximate the distribution of the data.

Fig. 1 depicts the SOM input and output using a sample dataset (source: Bache & Lichman, 2013). The input data (black) are compared with the initialized, but untrained, map (light grid) as well as with the final, trained, map (dark grid). The orientation of the initial map is along the first two principal components of the data set. It can be seen