Environmental Modelling & Software 25 (2010) 770–781

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

TopoToolbox: A set of Matlab functions for topographic analysis

Wolfgang Schwanghart*, Nikolaus J. Kuhn

Physical Geography and Environmental Change, University of Basel, Klingelbergstrasse 27, 4056 Basel, Switzerland

ARTICLE INFO

Article history:
Received 7 May 2009
Received in revised form
23 November 2009
Accepted 6 December 2009
Available online 19 January 2010

Keywords:
Relief analysis
Flow networks
Digital elevation models
Linear algebra
Matlab
Graph Theory
Sparse matrices

ABSTRACT

TopoToolbox contains a set of Matlab functions that provide utilities for relief analysis in a non-Geographical Information System (GIS) environment. The tools have been developed to support the work flow in combined spatial and non-spatial numerical analysis. They offer flexible and user-friendly software for hydrological and geomorphological research that involves digital elevation model analysis and focuses on material fluxes and spatial variability of water, sediment, chemicals and nutrients. The objective of this paper is to give an introduction to the linear algebraic concept behind the software that employs sparse matrix computations for digital elevation model analysis. Moreover, we outline the functionality of the toolbox. The source codes are freely available in Matlab language on the authors' webpage (physiogeo.unibas.ch/topotoolbox).

Software availability

Program title: TopoToolbox
Developer: Wolfgang Schwanghart
First available: 2009
Source language: MATLAB
Requirements: MATLAB R2009a, Image Processing Toolbox
Availability: TopoToolbox is available free of charge and can be downloaded on http://physiogeo.unibas.ch/topotoolbox.

1. Introduction

The shape of the Earth's surface is of major concern to environmental scientists, and to geomorphologists and hydrologists, in particular. Relief is a product of endogenic and exogenic processes and human action that in combination create a variety of landforms. Landforms provide an access to the erosional and depositional past of the environment. As such, relief is an important archive of Earth's history. Relief, in addition, is the template for fluxes of water, sediment, chemicals and nutrients (Merritt et al., 2003; Callow and Smettem, 2009; Murray et al., 2009), and understanding the interactions between material fluxes and relief is a key to present environmental systems and their future evolution (Kuhn et al., 2009).

Digital elevation models (DEMs) form the basis of quantitative methods to analyze and model topography and its relation to geological, hydrological, biological and anthropogenic components of the landscape (Moore et al., 1993; Florinsky et al., 2002). As such, the analysis of DEMs is an essential method applied in morphometric analysis (Dikau, 1989; Klingseisen et al., 2008; Ames et al., 2009), prediction of soil properties (Florinsky et al., 2002) and ecological habitat characteristics (Beck and Kitching, 2009), distributed hydrologic modelling (Lacroix et al., 2002; Wu et al., 2007; Viviroli et al., 2009) and erosion and sediment transport modelling (Merritt et al., 2003).

A DEM usually is a finite set of georeferenced, irregularly or regularly spaced elevation points representing ground surface topography. Terrain is modelled with triangular irregular networks, contour vertices or hexagon tessellation. Most common, however, terrain is digitally represented as grids of squares because the regular data structure allows for simpler measurement algorithms (de By, 2001). Such algorithms are usually implemented in a Geographical Information System (GIS) software.

While GIS software provides powerful tools for spatial analysis, non-spatial analysis tools are usually deficient in GIS environments. Often, however, spatial and non-spatial analyses are combined when e.g. contaminant surveys, discharge measurements or soil carbon inventories meet relief analysis. In these cases flexible solutions are required to avoid large efforts in data exchange

* Corresponding author.
E-mail address: w.schwanghart@unibas.ch (W. Schwanghart).
URL: http://www.physiogeo.unibas.ch/

1364-8152/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2009.12.002

between different software packages. The solution presented here is TopoToolbox, a set of functions for topographic analysis implemented in Matlab programming language. TopoToolbox was designed to address the needs of researchers who analyze gridded DEM data to investigate material fluxes on Earth's surface. Matlab was chosen since it provides a versatile environment for mathematical and technical computing. It is based on a high-level programming language and as such, the functions provided here are easy to modify and open for amendments.

The major aim of this paper is to give an introduction to the linear algebraic concept for topographic analysis as implemented in TopoToolbox. The concept is explained using a miniature sample digital elevation model and various toolbox functions are illustrated. Finally, we present the functionality with a case study investigating stream discharge in the Upper Rhône Valley, Switzerland.

2. Digital elevation models, topographic attributes and their algebraic representation

2.1. Digital elevation models

DEMs are geographic field type of data that model topography and topographic change (Brooks, 2003). Most commonly DEMs are represented as rectangular grids where an elevation value is assigned to each cell. Mathematically this kind of data model is termed matrix. Matrices are rectangular tables of scalars termed elements. The elements of a matrix are referenced either by specifying rows and columns of each respective scalar in the matrix or by a linear, rowwise increasing index. Here, we adopt the latter referencing scheme. Consequently, elements in a DEM matrix **Z** are subscribed using linear indexing as shown in Fig. 1.

The geographic location of each DEM cell is determined by a coordinate system which assigns a projected location in the *xy*-plane to each element in the DEM matrix. Thus, the 2D-Euclidean distance *d* between two cells with the linear index *i* and *j* is calculated by

$$d_{ij} = \sqrt{(\mathbf{X}_i - \mathbf{X}_j)^2 + (\mathbf{Y}_i - \mathbf{Y}_j)^2} \tag{1}$$

with **X** and **Y** being matrices with the same size as **Z** where each element refers to the *x*- and *y*-coordinate, respectively.

Processes represented in hydrological and geomorphological cellular models are governed by local values like roughness coefficients, grain size or vegetation density contained in further matrices. When taking elevations into account, however, altitude in each cell is usually less important for process representations than the relation of a cell to each of its eight neighboring cells (Moore neighborhood). This relation is what we refer to as relief and which determines the type and intensity of lateral material fluxes such as gravitational processes.

Mathematically, neighborhoods can be represented in so called *adjacency matrices*. In *Graph Theory* these matrices are used to define edges between nodes of graphs (Bondy and Murty, 1976). An adjacency matrix $\mathbf{A(Z)} = [a_{ij}]$ of a DEM matrix **Z** with n elements is a symmetric $n \times n$ matrix, where the elements a_{ij} equal one when a cell with the linear index *i* is a neighbor of cell *j*. Otherwise, a_{ij} is zero. **A** is populated primarily with zeros ($>n^2$–$8n$) and can thus be classified as sparse. The sparse data structure allows for efficient storage organization and operations compute sparse results in time proportional to the number of arithmetic operations on nonzeros. Fig. 2a visualizes the sparsity pattern of **A** for the DEM **Z** displayed in Fig. 1.

Adjacency matrices and some modifications of them have a great value for DEM analysis. Some examples are explained in greater detail in the following sections.

2.2. Gradient

One of the most important factors of surface processes is gradient since it provides a measure for the potential energy in a specific location towards a local basis. The gradient of a DEM cell may be differently defined (Warren et al., 2004) but in many cases the trigonometrical maximum downward gradient is taken. Finding the maximum gradient requires the calculation of all slopes between all neighboring cells. The slope S_{ij} between two neighbor cells *i* and *j* is calculated by

$$S_{ij} = \frac{Z_i - Z_j}{d_{ij}} \tag{2}$$

where d_{ij} is the distance between the cells *i* and *j* as calculated in equation (1). Applied to all cells and their respective neighbor cells the results are transferred to a matrix **S** with *i* as row indices and *j* as column indices of the slope values. The resulting matrix **S** has the same sparsity pattern as the adjacency matrix **A**. This time, however, the nonzero elements contain the tangent of the slope that are – according to equation (2) – positive when indicating

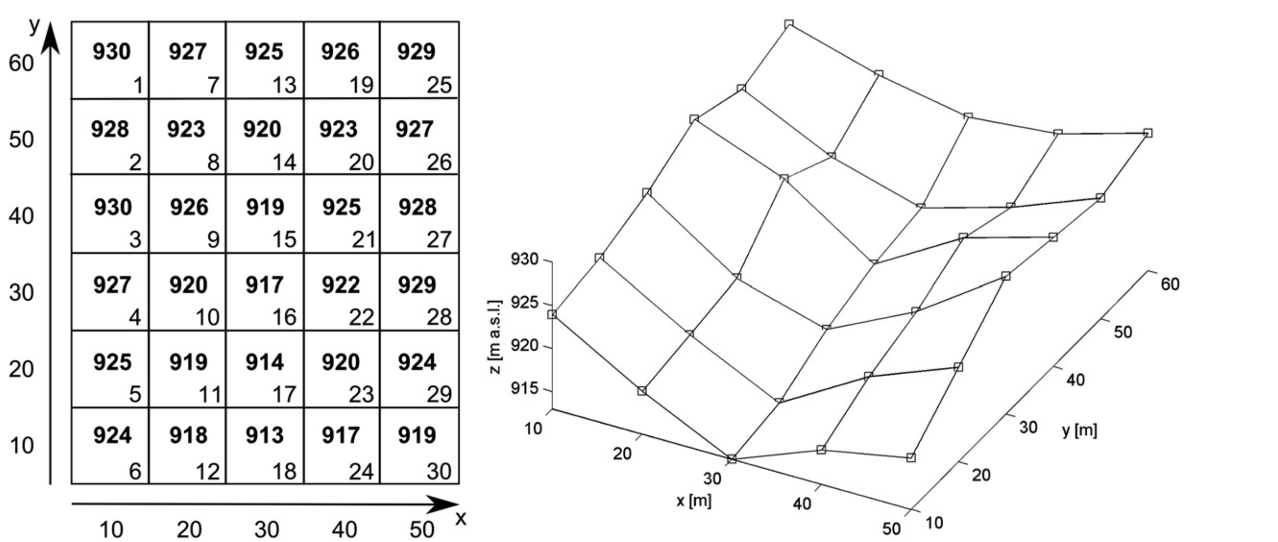

Fig. 1. Sample digital elevation model in matrix notation and surface illustration. Centered values in the matrix refer to cell elevation and lower right values refer to a linear index of the cells.