Earth-Science Reviews 143 (2015) 147–160

Contents lists available at ScienceDirect

Earth-Science Reviews

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

Graph theory in the geosciences

Jonathan D. Phillips [a,*], Wolfgang Schwanghart [b], Tobias Heckmann [c]

[a] Department of Geography, University of Kentucky, Lexington, KY 40506-0027, USA
[b] Institute of Earth and Environmental Science, University of Potsdam, Germany
[c] Physical Geography, Catholic University of Eichstaett-Ingolstadt, Eichstaett, Germany

ARTICLE INFO

Article history:
Received 7 November 2014
Accepted 9 February 2015
Available online 18 February 2015

Keywords:
Graph theory
Geosciences
Networks
Spatially explicit models
Structural models
Complexity

ABSTRACT

Graph theory has long been used in quantitative geography and landscape ecology and has been applied in Earth and atmospheric sciences for several decades. Recently, however, there have been increased, and more sophisticated, applications of graph theory concepts and methods in geosciences, principally in three areas: spatially explicit modeling, small-world networks, and structural models of Earth surface systems. This paper reviews the contrasting goals and methods inherent in these approaches, but focuses on the common elements, to develop a synthetic view of graph theory in the geosciences. Techniques applied in geosciences are mainly of three types: connectivity measures of entire networks; metrics of various aspects of the importance or influence of particular nodes, links, or regions of the network; and indicators of system dynamics based on graph adjacency matrices. Geoscience applications of graph theory can be grouped in five general categories: (1) Quantification of complex network properties such as connectivity, centrality, and clustering; (2) Tests for evidence of particular types of structures that have implications for system behavior, such as small-world or scale-free networks; (3) Testing dynamical system properties, e.g., complexity, coherence, stability, synchronization, and vulnerability; (4) Identification of dynamics from historical records or time series; and (5) spatial analysis. Recent and future expansion of graph theory in geosciences is related to general growth of network-based approaches. However, several factors make graph theory especially well suited to the geosciences: Inherent complexity, exploration of very large data sets, focus on spatial fluxes and interactions, and increasing attention to state transitions are all amenable to analysis using graph theory approaches.

Contents

* Corresponding author. Tel.: +1 859 257 6950; fax: +1 859 323 1969.
E-mail address: jdp@uky.edu (J.D. Phillips).

http://dx.doi.org/10.1016/j.earscirev.2015.02.002

1. Introduction

Geoscientists are obliged to deal with complex, highly interconnected, spatially extensive Earth systems. The long tradition of addressing these by reduction and simplification has increasingly been complemented by approaches that attempt to deal with these complex systems more holistically, and to understand the synergistic interconnections in addition to the individual mechanistic and historical relationships. Many complex Earth systems can be represented as networks. Graph theory, a branch of mathematics well suited to network analysis, is thus emerging as a powerful tool in the Earth and environmental sciences. Network approaches in many disciplines (including Earth and environmental sciences) have highlighted a linkage between system properties and dynamics or behavior that can be addressed using graph theory. The purpose of this paper is to review graph theory applications, introduce geoscientists to some of the more promising techniques, and develop a synthesis of graph and network-based approaches in the geosciences.

An Earth system can be characterized as a set of interconnected components. These may be locations, sources, sinks, or nodes in flux networks; objects (e.g., landforms, mass or energy storage compartments); processes or process bundles or regimes (e.g. weathering, moisture advection; isostatic adjustments); or phenomena or events (e.g. overbank flows, tropical cyclones, earthquakes). These components are connected by fluxes of matter and energy, feedbacks, spatial or temporal sequencing or adjacency, statistical correlations, and process-response relationships. Table 1 summarizes some general examples. Thus Earth systems can often be represented as networks, whether in a literal sense (e.g., fluvial channel networks; rock fracture patterns) or as a conceptual and analytical tool. The components represent the nodes (vertices) of the network, and the relationships between them the edges (links). The box-and-arrow diagrams commonly used in geosciences may be treated as graphs, with the boxes as nodes and the arrows as edges or links.

Spatially explicit applications of graph theory in geosciences typically analyze large, complicated networks to identify critical nodes or regions, to measure connectivity properties, or to identify process coupling based on spatial patterns (e.g., Palus et al., 2011; Tsonis and Swanson, 2012; Heckmann and Schwanghart, 2013). Spatially explicit graphs are those where the nodes represent locations, regions, or points in space, such as sources and junctions in a stream channel network, geographical grid cells in climate models, or earthquake epicenters. The edges or links may be defined by flux patterns or other spatial interactions, or statistical relationships.

Structural graph models are network representations of the structure of geosystems. These may be empirically derived state-and-transition models or correlation structures (e.g. Donges et al., 2011; Phillips, 2011a), or theoretically-derived networks based on fundamental equations or system models (e.g., Tsonis and Roebber, 2004; Phillips, 2012). Applications of graph theory to structural models of geosystems typically seek to identify or measure the complexity or synchronization properties of environmental state transitions, or to deduce geosystem properties from network structures.

As an established branch of mathematics and systems theory, there exist numerous textbooks on graph theory. Harary (1994) has been especially influential among geographers and spatial analysts, and Arlinghaus et al. (2002) deal specifically with graph theory in geography. Biggs (1994) is a standard reference in algebraic graph theory, and Heckmann et al. (2014a) recently reviewed graph theory applications in geomorphology, and include a summary of graph theory definitions and principles. Thus only a very brief review is included here.

A graph consists of N nodes or vertices and m links or edges. A path exists between any pair of nodes if a sequence of edges can be followed between them. If links or edges are bidirectional (i.e., lines) the graph is undirected; otherwise (edges are arrows) it is directed. Directed graphs may be signed (positive or negative links) or unsigned. Both links and nodes may be weighted (assigned a value) or unweighted. A graph

Table 1
General examples of graph nodes and edges in geoscience applications of graph theory.

Examples of system components (graph nodes or vertices)
Locations (points, areas, regions)
Nodes of flux networks (sources, sinks, junctions)
Intersections of linear spatial patterns (e.g., rock fracture networks, cave passages, storm tracks)
Objects (e.g., landforms, soil types, geologic formations)
Mass/energy storage compartments
Process bundles or regimes (e.g., tectonic uplift, water erosion, evapotranspiration)
Events or phenomena (e.g. earthquakes, fires, storms)
Variables in process-response relationships (e.g., flow resistance equations)
Examples of system links (graph edges)
Transport pathways
Mass, energy fluxes & exchanges
Feedback relationships
Process-response linkages
Temporal sequences
Spatial adjacency
Statistical correlations
Theoretical or mathematical relationships
Similarity measures or indices

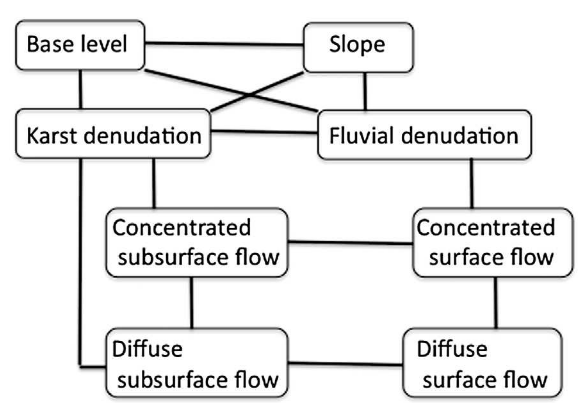

Fig. 1. An undirected graph showing connections among key components of fluviokarst landscapes at the landscape scale (top) and process scale (flow components, bottom). After Phillips (2012).

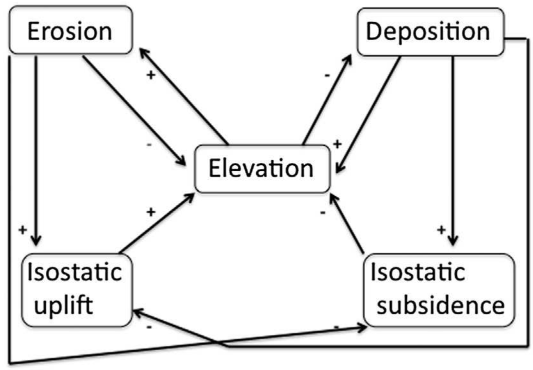

Fig. 2. Example of a directed graph, showing positive and negative relationships related to topographic evolution in the absence of tectonic forcing. After Phillips (2002).

where every node is linked to at least one other is connected; if every node is linked to every other node the graph is fully connected. Some examples are shown in Figs. 1 and 2.

A *N x N* adjacency matrix *A* is associated with any graph. Matrix entries are zero if the row and column nodes are unconnected. Nonzero elements a_{ij} in *A* denote an edge from node *i* to *j*. If *A* is the adjacency matrix of an undirected graph, that means that interactions between nodes are independent from direction, $a_{ij} = a_{ji}$ and *A* is symmetric. For the case of an unweighted, connected graph *A* is symmetric and all entries are 0 or 1. Weighted graphs may have values other than 0, 1. The incidence matrix of an undirected graph has a row for each node and a column for each link, with cell values of 0,1 according to whether the vertex is incident upon the edge. The incidence matrix of a directed graph has values of 0, −1, 1, the latter depending on whether the link originates or ends at the associated node.

A comprehensive discussion of software for graph and network analyses is beyond the scope of this paper, but it is worth noting that the available tools have been expanding both in number and sophistication, probably contributing to the increased use of graph theory across the sciences. A few illustrative examples are provided here. A versatile free and open source software tool for graph analysis is *igraph* (http://igraph.org/index.html). The *igraph* library can be embedded in Python, Perl, or R; it contains a large number of functions to create, manipulate, analyse and visualise graphs. Igraph can read and write common data formats and handles large graphs. Another related R package is *network* (http://statnet.org/); it forms part of *statnet* (http://statnetproject.org), a bundle that focuses on statistical modelling of network data. Implementation of packages such as *igraph* in R makes it possible to use all the power of R and its packages, including matrix calculations.

Several programming languages and software packages provide state-of-the-art functionality for analysing, modelling and visualizing graphs and network structures. Proprietary software include MATLAB and Mathematica. Owing to MATLAB's foundation in matrix algebra, various metrics related to spectral graph theory can be calculated using algorithms for calculating eigenvalues and eigenvectors. Graph structures are analysed with functions for reordering of matrices, and shortest paths or minimum spanning trees can be derived using linear programming. The MATLAB Bioinformatics Toolbox features several graph theory algorithms. Free MATLAB toolboxes include matlabBGL, which is very fast since it interfaces with the Boost Graph Library (Siek et al., 2002). Graph-tool and python-graph are among the most popular graph theory modules for the python programming environments.

In addition, a number of social network analysis packages (many free and open-source) are available with capabilities for graph theory analysis and graph visualization, and are often readily applied to adaptable for geosciences. One example is SNAP (Stanford Network Analysis Program; http://snap.stanford.edu/).

2. Networks and graphs in geosciences: overview

2.1. Spatial networks and connectivity

Spatial graphs are characterized by their nodes being located in geographical space (e.g., Fig. 3). Depending on the research problem and spatial scale, nodes represent sampling locations or other point-like objects, event locations (e.g., earthquake epicenters or tropical cyclone landfalls), habitat or resource patches (ecology), river junctions or reaches (hydrology), landforms (geomorphology), or any other spatial unit. The edges linking these nodes then have a length, though their exact location and geographic course can be explicit or not (Dale and Fortin, 2010). Similarly, the length of an edge can correspond to distance as the crow flies, or any other distance such as the "effective" length (least-cost path, e.g. Adriaensen et al., 2003). Depending on the research problem, edges represent the directed transfer of matter or energy, or are undirected (indicating neighborhood). A set of spatial nodes can be linked by a number of different algorithms, resulting in a hierarchy of spatial graphs (Dale and Fortin, 2010); examples include the linkage of (mutually) nearest neighbors, the minimum spanning tree, Delaunay triangulation and the complete graph. The latter links all *n* nodes to each other using *n-1* edges. Both nodes and edges can have additional attributes that may relate to spatial properties (size, width, etc.).

A key property of ecological, hydrological and geomorphic systems is connectivity; it governs how organisms, matter and/or energy are exchanged between system components. Graph models of such systems were introduced above, and graph theory connectivity measures are discussed below. Landscape connectivity was first defined in ecology as "the degree to which the landscape facilitates or impedes movement [of organisms] among resource patches" (Taylor et al., 1993). It depends on the interaction of landscape structure and properties of particular species and influences–among others, the sensitivity of populations to disturbance; connectivity is therefore relevant for conservation issues. The idea that habitat and resource patches in a landscape, separated by a more or less permeable matrix, form a network led to the generation of spatial graph models. Habitat or resource patches form the nodes, and (usually undirected) edges are formed based on a distance measure between nodes and relevant properties (e.g., its dispersal range) of the particular species in question. This graph-theoretic perspective on landscape connectivity has been developed and applied in a large number of studies over the last 20 years (e.g. Cantwell and Forman, 1993; Bunn et al., 2000; Urban and Keitt, 2001; Pascual-Horta and Saura, 2006; Minor and Urban, 2008; Kong et al., 2010; Zetterberg et al., 2010; Erős et al., 2012; Segurado et al., 2013). Cantwell and Forman (1993) and Galpern et al. (2011) review and discuss methods to construct and analyse such graphs.

The quantification of landscape connectivity has been an important field of research (e.g. Tischendorf and Fahrig, 2000), especially related to conservation issues, and graph-theoretic measures form part of a large number of metrics proposed so far (Calabrese and Fagan, 2004). They may address single nodes or edges, or properties of the whole network; for example, measures address the connectivity of a network with respect to the number and properties of clusters, i.e. subgraphs that contain well connected nodes but are poorly connected to other subgraphs in the network. Such measures include the number of graph components, and the diameter of the largest graph component (Calabrese and Fagan, 2004, and references therein). Other indices have been proposed by van Langevelde et al. (1998; with comments on the connectivity of directed graphs) and Pascual-Horta and Saura (2006). Rayfield et al. (2011) reviewed and classified network measures; they found >60 of them for ecologists to choose from. The relative contribution of nodes or edges to overall connectivity can be assessed by removing the respective component from the network and re-calculating connectivity (e.g. Urban and Keitt, 2001). Tables 2 and 3 are summaries of some common graph connectivity measures.

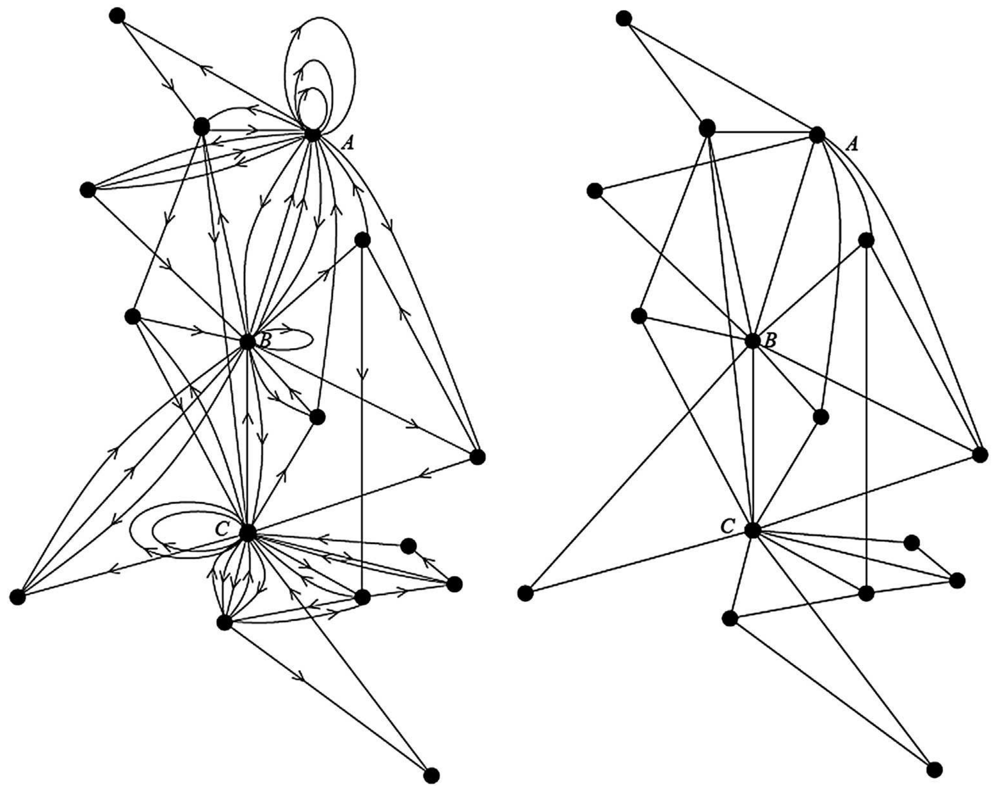

Fig. 3. Left is a schematic graph of an earthquake network from Abe and Suzuki (2006), with nodes indicating spatial cells where earthquakes occurred and edges representing event correlations. At right the network is reduced to a simple undirected graph. Nodes A, B, C are locations of main shocks and are graph hubs with large connectivity values (Abe and Suzuki, 2006).

Similar to Calabrese and Fagan (2004), who proposed a "comparison-shoppers guide to connectivity metrics" in ecology, Ali and Roy (2010) provided an overview of connectivity metrics in hydrology. Hydrological connectivity addresses the spatiotemporal variability of runoff and runon, i.e. the (directed) linkage of runoff generation and the channel network, and has received considerable attention in the last decade (Bracken and Croke, 2007; Ali and Roy, 2009; Bracken et al., 2013). Graph theory applications to hydrological research problems have increased in recent years. Poulter et al. (2008), set up and analysed graph models of artificial drainage systems; and Aurousseau et al. (2009) and Gascuel-Odoux et al. (2011) analysed networks of surface flows between agricultural plots. Runoff dynamics and hydrological connectivity were measured using graph theory by Phillips et al. (2011), and Cui et al. (2011) employed graph theory for the description of drainage networks in operational river forecasting. Spence and Phillips (2014) linked catchment-scale connectivity measures to temporal variations in runoff-generating processes. Channel-floodplain connectivity of alluvial wetlands was found to be non-linear and initiated at subbanktop flows via a graph analysis by Kupfer et al. (2014).

Closely linked to hydrological connectivity is geomorphic, or sediment, connectivity (Brierley et al., 2006; Fryirs, 2013; Bracken et al., 2014). Here, components of a landscape, e.g. landforms, are coupled to each other by geomorphic processes effecting sediment transfer, and connectivity is a system property related to the continuum of the cascading system (Bracken et al., 2014). The analysis of sediment connectivity using directed graphs represents a new approach that still is in its infancy, but early results are promising and have yielded geomorphic insights unlikely to have obtained otherwise. Heckmann and Schwanghart (2013) identified and analysed sediment cascades in a high mountain catchment as sequences of edges connecting the raster cells of a digital elevation model; the edges are generated using numerical models simulating the sediment transport by rockfall, slope-type debris flows, slope wash, and within the channel network. Recently, Heckmann et al. (2014b) presented first results of a graph-theoretic analysis of sediment connectivity that is based on coupling relationships of adjacent landforms depicted on a geomorphological map

While the analysis of ecological and hydrogeomorphic spatial networks is conceptually similar, the implications of connectivity for the respective systems are different. Well-connected habitat networks, for example, are thought to be resilient to disturbances (Minor and Urban, 2008; Urban et al., 2009). By contrast, well-connected geomorphic systems are highly sensitive to change (e.g. Phillips, 2011a; Kuo and Brierley, 2014). A lack of intermittent storage landforms such as aggradational fans or intramontane basins entails that sediment is readily transferred along the river network, and facilitates the propagation of changes in both up- and downstream direction. Blöthe and Korup (2013) showed that storage landforms such as fluvial terraces, deposits accumulated upstream to landslide dams and large-scale mass wasting may introduce millennial lag times in fluvial sediment routing systems, buffer seismic and climatic disturbances, and complicate their direct correlation with downstream sedimentary archives.

2.2. Complexity

For a graph with a given number of nodes, complexity is a function of the number of edges, and the "wiring" of those connections. In general, complexity increases with the number of potential paths between any two nodes, and the number of loops (paths within the graph of any length that start and end at the same node). The highest complexity is associated with a fully connected graph (every node is connected to every other); the lowest with a pattern where one key node is connected to all others; which in turn are connected only to the key node.

The adjacency matrix A has N eigenvalues λ_i, such that $\lambda_1 > \lambda_2 \ldots > \lambda_{N-1} > \lambda_N$. The largest eigenvalue is an important determinant of a number of system properties (Biggs, 1994; Restrepo et al., 2007). λ_1, called the spectral radius, is the best and most

Table 2
Connectivity measures for graphs.

Measure	Definition	Comments
Number of nodes (N)	Number of nodes (vertices)	
Number of links (m)	Number of links (edges)	Maximum possible $m = (N^2 - N)/2$ Minimum possible $m = N - 1$ (connected graphs)
Link density	m/N	
Mean degree (d)	$(\Sigma\, d)/N$	$(\Sigma\, d) = 2\, m$ in simple undirected graph
Degree distribution	$P(k) = N_k/N$ N_k = number of nodes of degree k	Scale-free networks characterized by $P(k) \sim k^{-a}$, $a > 1$
Graph diameter	Maximum geodesic distance between any two nodes	Geodesic distance = shortest path
Characteristic path length	Characteristic (e.g., mean) number of edges between any two nodes	Other measures of central tendency can be used
Assortativity coefficient	Pearson correlation coefficient between nodes of same degree	Extent to which nodes link to other nodes of the same degree
Rich-club coefficient	$\phi(k) = (2m_k)/N_k(N_k\text{-}1)$ N_k, m_k = number of nodes, edges of degree k	Indicates extent to which graph's most highly-connected nodes are connected to each other
Average node distance	Mean geodesic distance between pairs of nodes	
Average node eccentricity	Mean maximum geodesic distance between pairs of nodes	
Harary Index	$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1,\, j\neq i}^{n}\frac{1}{nl_{ij}}$ nl_{ij} = shortest path between nodes i and j (topological distance = number of links)	$nl_{ij} = \infty$ if i and j are unconnected A normalized variant exists for comparison between systems with different n
Integral index of connectivity	$\sum_{i=1}^{n}\frac{\sum_{j=1,\, j\neq i}^{n}\frac{a_i a_j}{1+nl_{ij}}}{A_L{}^2}$ A_L = total landscape area a_i = area of node i	

widely used general indicator of graph complexity (Biggs, 1994; Restrepo et al., 2007). λ_1 is sensitive to the number of cycles, and reflects the extent to which externally imposed changes are likely to be amplified, critical coupling strength for coherent behavior, and dynamical stability (Tinkler, 1972; Schreiber and Hastings, 1995; Restrepo et al., 2006, 2007; Yuan et al., 2008; Phillips, 2011a,b; 2013a,b,c; Logofet, 2013).

The maximum spectral radius for a graph of given number of nodes and edges is

$$\lambda_{1,max} = [2m(N\text{-}1)/N]^{0.5} \tag{1}$$

This is useful because it allows observed spectral radius values to be compared to reference maximum values for specific graph structures (e.g., Phillips, 2012; 2014; Logofet, 2013). Other metrics are listed in Table 4.

2.3. Synchronization

Studies of synchrony in geosciences have been almost exclusively concerned with the concurrence of changes or events in different locations (for instance, geographical similarities or differences in timing of glacial advances or retreats). These studies have examined the contemporaneity of features (for instance, whether planation surfaces were formed at the same time); and the problems of correlating time series and data sets (for instance, reconciling climate, sea level, and stratigraphic chronologies).

Synchronization in general refers to the property or situation whereby things happen or operate simultaneously or in unity, or in rapid and predictable sequence. Graph theory is widely applied in systems theory, computer science, and engineering studies of synchronization to identify system properties that influence synchronization, and to facilitate control and management of networks to achieve synchronization. Network synchronization studies (as opposed to synchronization as described above) are rare in geosciences, but examples exist. These applications have the potential to address chronic issues such as quantifying effects of scale linkage and historical contingency (Phillips, 2012, 2013a; Beven, 2014; Cohen et al., 2014).

In algebraic graph theory the most common measure of synchronization is algebraic connectivity, defined as the second-smallest eigenvalue of the Laplacian matrix $L(A)$ of the adjacency matrix, $\lambda(A)_{N\text{-}1}$ (Fiedler, 1973; Biggs, 1994; Duan et al., 2009).

2.4. Small-world networks

Initially inspired by studies of social networks and biological phenomena, in the late 1990s the phenomenon of "small world networks" (SWN) gained wide attention. The most common examples of, and

Table 3
Graph metrics that apply to individual nodes or links; mean values apply to entire graph.

Measure	Definition	Comments
Betweenness centrality or node betweenness	Proportion of paths between other nodes of the graph that pass through a given node	Originally developed to quantify individual's influence within a social network
Link betweenness	Proportion of paths in graph that include a given link	
Mean neighbor degree	Mean degree of neighbors of a given node	
Clustering coefficient	$C_u = (2e)/k(k\text{-}1)$ e = number of edges between k neighbors	Ratio of neighbors of node u to total possible.
Degree centrality	Node degree normalized to interval 0,1	Important nodes are involved in a large number of interactions
Closeness centrality	Geodesic distance between nodes	Indicates how quickly a node can communicate with other nodes
Node coreness	Maximum k such that node is still present in k-core but absent in $(k + 1)$ core	k-core is a subgraph obtained from original by recursive removal of all nodes of $d \le k$

Table 4
Algebraic graph metrics applied to graph adjacency matrix (***A***).

Measure	Definition	Comments
Spectral radius (λ_1)	Largest eigenvalue of ***A***	Measures graph complexity
Algebraic connectivity ($\lambda(A)_{N-1}$)	Second smallest (largest nonzero) eigenvalue of Laplacian of ***A***	Measures graph synchronizability
S-metric ($S(g)$)	$S(g) = \sum_{i-1}^{N} d_i d_{i+1}$	Indicates scale-free properties; d_i = degree of node i
Laplacian spectral radius ($\lambda_1(A)$)	Largest eigenvalue of Laplacian of ***A***	Indicator of network stability
Critical coupling strength (c)	Critical value at which coupled subgraphs transition from incoherent to coherent behavior	$c = \mu/\lambda_1$ μ = parameter depending on dynamics of subsystems

metaphors for, SWNs are based on social networks and the famous "six degrees of separation" phenomenon. This is the idea that, on average, any two humans are separated by only six levels of kinship or acquaintance – a Chinese engineer and a farmer in Argentina who have never had any contact could, for example, be connected by a handful of intermediary individuals. SWNs are characterized by relatively sparse overall networks – to stick with the social metaphor, even the most outgoing and cosmopolitan individuals, for example, will have met only a miniscule fraction of other humans. The network of connections between local social, economic, and kinship networks, however, allows numerous "shortcuts" so that only six links (on average) are necessary to link any two of us. Early work on SWNs is reviewed by one of the principal discoverers, Duncan Watts (1999).

Watts and Strogatz (1998) introduced SWNs into the scientific literature, noting that small-world phenomena are not limited to social networks. Subsequent work has borne that out, including geosciences. Yang (2001) pointed out potential applications of SWN theory to geophysics, and Valentini et al. (2007a) analyzed networks of rock fractures as graphs, and found that these have SWN characteristics. The SWN framework was also applied to magma transport beneath mid-ocean ridges (Valentini et al., 2007b). Seismic data from various locations were found to have small-world features by Abe and Suzuki (2006) and Jiménez et al. (2008). Tsonis et al. (2008) identified SWN properties as a potential explanation of long-distance climate teleconnections. Phillips (2012) used SWN structures in geomorphic phenomena to address scale linkage issues.

SWNs are characterized by power-law degree distributions, large clustering coefficient values, and small average path lengths. These metrics have been applied in geosciences by, e.g., Abe and Suzuki (2006). Phillips (2012, 2013a) has made use of the *S*-metric (Li et al., 2005; Table 4), developed to measure the extent to which a graph has scale-free properties. It applies to simple, undirected, connected graphs. $S(g)$ is maximized when high-degree nodes are connected each other. Graph self-similarity increases $S(g)$, increasing the likelihood that inferences about system evolution are independent of specific time periods. Scale-free graphs have a power-law degree distribution and feature a small number of highly connected nodes, such as the hubs of major airline networks, and are efficient for information transfer within the network. This may explain, for example, scale linkage phenomena in hierarchical Earth systems (e.g., Phillips, 2012), and the workings of teleconnections in ocean–atmosphere interactions (e.g., Tsonis and Roebber, 2004; Donges et al., 2009; Palus et al., 2011). Scale-free networks are considered more stable than other relatively sparse graphs, in that random removal of a node is unlikely to disconnect the graph, because of the low probability that a randomly removed node will be a key hub (see e.g., Tsonis et al., 2008).

2.5. Geoscientific insights from graph theory

Before discussing the goals and categorizing the applications of graph theory, in this section we show that these methods can lead to novel insights in Earth and atmospheric sciences. Some geoscience applications of graph theory represent improved (e.g., faster or more efficient) techniques for certain types of analysis or applications (e.g., Bouille, 1976; Pacheco, 1998; Arge et al., 2003; Steinhaeuser et al., 2011, 2012; Grace et al., 2012; Braun and Willett, 2013; Schwanghart and Scherler, 2014). Other applications are geared toward contextualizing geoscience networks in the broader context of graph and network theory. For instance, SWNs are known to have interesting and important properties potentially relevant to Earth systems, so some studies have been geared toward determining whether, for instance, climate, seismic, or rock fracture networks have small-world properties (e.g., Yang, 2001; Abe and Suzuki, 2006; Valentini et al., 2007a; Baek et al., 2012). These types of application are legitimate and important in their own right, and may lead indirectly to novel understandings of geosciences. In this section, however, we focus on a selective sampling (i.e., illustrative rather than comprehensive) of new findings about geophysical phenomena arising directly from application of graph theory.

The application of graph theory to climate networks has led to the development of a new subfield of *climate networks*, Peron et al. (2014) argued. The climate networks paradigm is based on spatial grid points in climate data as nodes, with edges established on the basis of statistical correlations. In their work, Peron et al. (2014) discovered climate-topography interactions at continental to global scales that are independent of the known effects of orography. These interactions had not been detected by non-graph theory methods.

Network analysis – particularly SWN – has also provided key insights into climate teleconnection patterns. The SWN properties of some climate networks are in fact a plausible explanation for the existence of some teleconnection patterns that (due to complex dynamics and vast distances involved) are difficult to explain otherwise. By measuring degree centrality, Tsonis et al. (2006; Tsonis and Swanson, 2008) identified key nodes and linked them to the North Atlantic Oscillation. This line of inquiry has also uncovered long-distance connections associated with ENSO (El Nino-Southern Oscillation) phenomena. Temperature predictability decreases in El Nino episodes, due to impacts of ENSO on climate system stability (Gozolchiani et al., 2008; Tsonis and Swanson, 2008).

Donges et al. (2009) used graph theory methods to find previously unknown wave-like structures linked to ocean currents. These represent a key pathway of increased energy and mass flux in the global air temperature field. Donges et al. (2009) also showed that the patterns uncovered using their methods are not detectible using traditional analyses.

The study of earthquake networks in seismology was pioneered by Abe and Suzuki (2004, 2006), who argued that the graph/network approach produces unique descriptions and insights. Using a variety of network analysis methods, they have produced a number of papers regarding the hierarchical structures, scaling, connectivity, and complexity of earthquake networks. This has led to some new geophysical insights, such as the fact that the clustering coefficient peaks during main shocks, indicating that the network is dynamic, with new shocks organizing communities (highly connected subgraphs within the network; Abe and Suzuki, 2007). Later, Abe and Suzuki (2012) showed that the main shock absorbs and combines existing (graph) communities to create new ones.

Graph community identification was also key in Halverson and Fleming's (2014) analysis of an array of stream gauges in Canada.

Beyond showing that graph theory can be usefully applied to the assessment and design of hydrometric networks, they identified 10 separate communities. These communities are associated with the processes dominating the flow regime and exposure (or not) to similar climate and meteorological forcings. Graph properties of gauge networks were also linked to runoff-producing process regimes by Spence and Phillips (2014). Vevatne et al. (2014) analyzed a graph based on fractures in sea ice. The graph properties suggested preferential growth phenomena, indicating that long, dominant fractures appear first. This enabled them to develop a new model of fracture network growth of sea ice.

Graph theory has allowed significant progress to be made on some chronic problems in geomorphology. Heckmann and Schwanghart (2013) identified modes of geomorphic coupling among different mass movement processes, and between hillslopes and channel systems. Given the massive size of the dataset of flux source/origination and termination zones, this would not have been possible without use of graph theory methods. A chronic problem in soil landscape analysis is the possible dependence of results on the vagaries of soil taxonomy. Phillips (2013c) used algebraic graph theory to develop a method for measuring the contribution of taxonomy to soil map uncertainty.

Historical contingency is ubiquitous in Earth sciences, but except where long, high-resolution time series are available it has been generally unmeasurable (c.f. Beven, 2014; Cohen et al., 2014). Phillips (2013a) used algebraic graph theory to develop methods to quantify contribution of historical contingency to the complexity of Earth surface systems. Graph theory methods have also provided new measures of connectivity in hydrological and geomorphic systems by linking network connectivity properties with flux connectivity in flow systems (e.g., Santiago et al., 2008; Zaliapin et al., 2010; Pardo-Iguzquiza et al., 2011; Marra et al., 2014).

Even this brief review makes it clear that graph theory has enabled new insights into geoscience problems, as well as streamlining or simplifying certain types of analysis, and connecting geoscience with the broader scientific studies of networks and complexity. In the next section we categorize applications of graph theory according to the broad goals of the analysis.

3. Goals of graph analysis

Applications of graph theory in Earth, atmospheric, and ocean sciences can be tentatively classified according to the primary goals of the analysis. Like nearly any classification in Earth sciences, there is some overlap (i.e., a given application may have multiple goals), and not every case will fit neatly into the categories. Goals of graph theory applications in geosciences include identifying and quantifying properties of complex networks, and searching for evidence of particular types of graph structure. These are mainly spatially explicit applications, and are broadly analogous to statistical analyses of large, complicated data sets to identify key structures and relationships. Graph theory is also employed in spatial analysis of morphometric or network patterns, much in the tradition of graph theory applications in quantitative

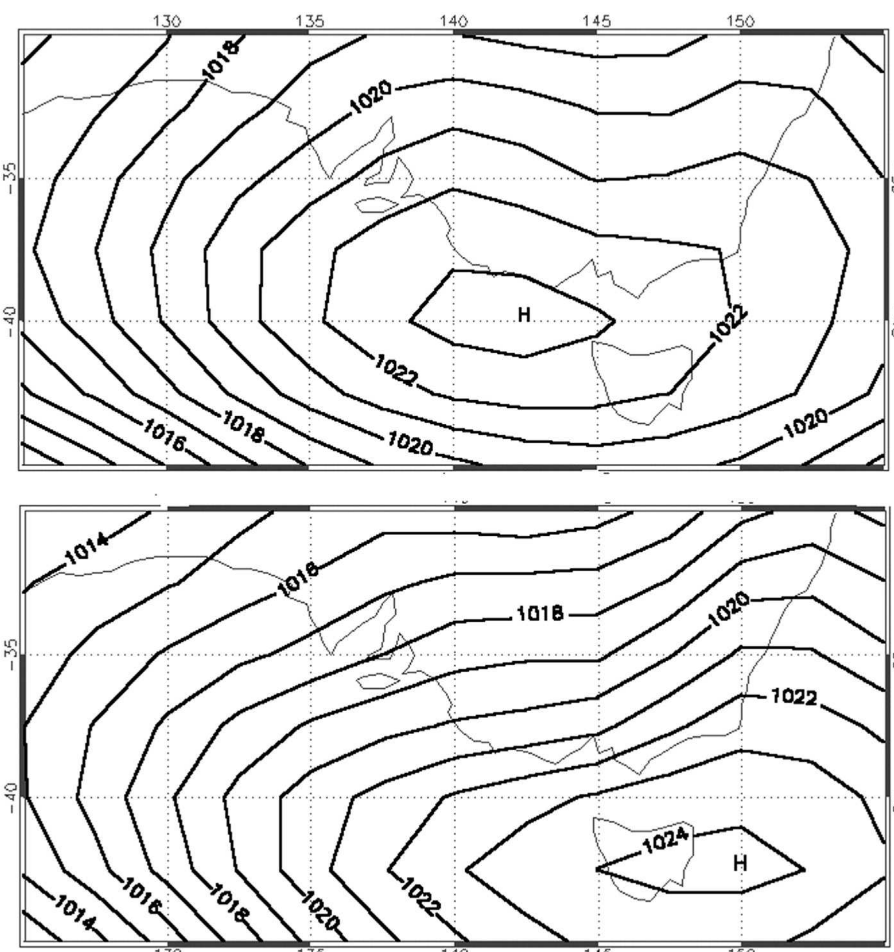

Fig. 4. Mean sea level pressure maps for synoptic types 14 (top), and 32 (bottom), in the vicinity of the state of Victoria, Australia. These are associated with, respectively, weak ESE and moderate ENE anticyclonic flow (Stern, 2003). By treating synoptic types as graph nodes and identifying links or edges based on temporal sequences, connectivity, complexity, and other properties of climate may be explored via graph theory. This is the approach used by, e.g., Zagouras et al. (2012).

geography. In the temporal domain, graph theory is sometimes employed to identify patterns from historical records or time series – again, in very broad terms of the major goals, analogous to statistical time series analysis. Structural graph theory approaches often have goals of testing, evaluating, and measuring dynamical system properties such as complexity, coherence, dynamical stability, and synchronization. Finally, the applications of graph theory to design or optimize networks also finds occasional use in the geosciences. Each of these broad categories of goals is briefly reviewed below.

3.1. Complex network properties

The main goal of these analyses is to identify and quantify properties of spatial networks, such as connectivity properties, critical nodes and regions, clustering, and centrality (relative importance of graph nodes). Several examples come from climatology (Fig. 4). Tsonis et al. (2006; Tsonis and Swanson, 2008) analyzed data on 500 hPa levels, defining grid cells for data collection as nodes, and determining edges based on the strength of correlations between 500 hPa heights. They identified "supernodes" in the resulting graph and showed that these correspond with major climate teleconnections as described above. They also showed that these supernodes and associated teleconnections make the network more stable and efficient in information transfer. Donges et al. (2009) used similar methods applied to global surface air temperatures. Classification of atmospheric circulation types (based on 850 hPa levels in the eastern Mediterranean region) was accomplished by applying graph theory to synoptic maps, thereby providing an additional objective, quantitative tool to synoptic climatology (Zagouras et al., 2012).

In seismology, Baek et al. (2012) used graph theory to identify topological properties of earthquake data, linking phenomena of shock clustering on the Korean peninsula to similar findings in other locations . Structural equation models have long been used to identify properties of ecological and hydrological networks, and Grace et al. (2012) developed a graph theory-based method for implementing this approach in wetland ecohydrology.

3.2. Graph structures

The primary purpose of these analyses is to search or test for particular types of graph or network structures, most commonly SWNs or scale-free networks(e.g., Baiesi and Paczuski, 2004). The latter are defined as networks whose degree distribution follows a power law.

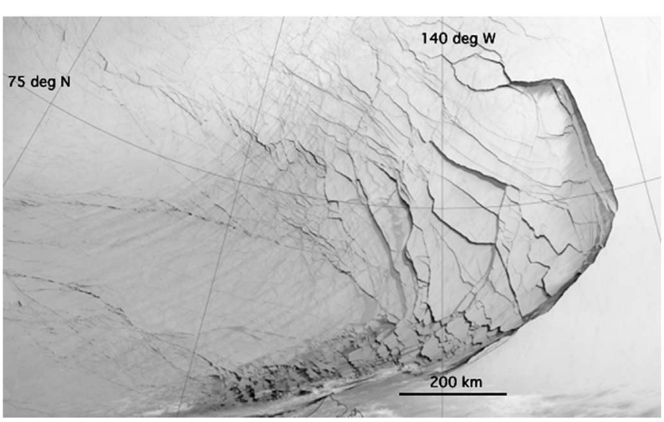

Fig. 5. Satellite image from 23 February, 2013 showing cracks in sea ice in the Beaufort Sea (original image from U.S. NASA Earth Observatory, http://earthobservatory.nasa.gov/IOTD/view.php?id=80752). Fracture intersections can be defined as nodes, and the connecting fractures as links in a graph. Vevatne et al. (2014) used this approach in the Kara Sea to develop a model of growing fracture networks in ice.

Scale-free properties are related to the robustness or vulnerability of a network to disruption and properties of preferential attachment (tendency of some highly-connected nodes to become more so over time). SWNs are one type of scale-free network. A large body of work in geomorphology, hydrology, and geophysics concerns itself with scale-free properties such as fractality, self-similarity, and power-law scaling. However, most of this work does not use graph theory methods. Example applications of SWNs were given in Section 2.4.

Prairie wetland networks in North America are scale-free, Wright (2010) found, but "meso-world" rather than small-world, and these dynamical scaling properties have implications for management and conservation. Vevatne et al.'s (2014) study of sea ice fracture networks (Fig. 5) is another example of a study based on identifying graph structures.

3.3. Spatial analysis

The very roots of graph theory are closely tied to analysis of spatial patterns and problems of cartographic representation. These are reviewed and linked to applications primarily in transportation, urban, and economic geography, cartography and quantitative spatial analysis by Arlinghaus et al. (2002).

Werner (1993, 1994) used graph theory methods in the quantitative analysis of topography, based on networks of fluvial channels and ridgelines. As the first quantitative analyses of the integrated network of ridge and drainage lines, Werner (1993, 1994) turned to graph theory as the most intuitive and efficient method, though at the time these methods were not well known in geomorphology. Larsen et al. (2012) used graph connectivity analyses to study spatial patterns of wetland interconnections, showing the utility of the methods for quantifying directional hydrological and ecological connectivity. The often complex networks of cave passages would seem to lend themselves to graph theory analysis, but there have been thus far only a few applications. Briesch (2011) used graph theory both to analyze networks of cave passages, and to optimize cave mapping and search-and-rescue strategies. Pardo-Iguzquiza et al. (2011, 2012) applied graph theory to morphometric analysis and to stochastic simulation of karst conduit networks in three dimensions.

In geomorphology, explicit, spatial applications of graph theory are scarce, yet we detect an upsurge in interest recently (e.g., Carling et al., 2014; Kleinhans et al., 2013). Although a plethora of studies on river networks exist, relatively few of these make use of graph theory and its measures (Heckmann et al., 2014a). A reason for the general lack may be that many spatially explicit networks in geomorphology such as river networks or sediment cascades are modelled as directed graphs that are representative for the flux of sediment and water, and for which fewer methods and measures exist compared to the toolbox available for undirected graphs. Rethinking graph theoretical applications in geomorphology will likely be spurred by the recent interest in sediment connectivity (Fryirs, 2013; Bracken et al., 2014; Heckmann and Schwanghart, 2013). Further discussion on graph theory applications in geomorphology is given by Heckmann et al. (2014a).

The examples above are based on morphological networks. Spatial analysis using graph theory has also been applied to networks based on starting and end points of sediment pathways in an alpine watershed (Heckmann and Schwanghart, 2013). Many graph theory applications in landscape ecology are conceptually similar, in that nodes and edges are defined on the basis of patches and corridors connecting them (e.g., Urban et al., 2009; Galpern et al., 2011).

Another approach to spatial analysis is based on patterns of spatial adjacency (Fig. 6). In this case the graph nodes are features such as soil types, landforms, or habitat units. These features are considered connected if they occur contiguously to each other. Phillips (2013b) applied this method to analyze spatial complexity of soil landscapes, and showed that the approach could separate uncertainty associated with

Fig. 6. Soil map of a 4568 ha area near Carrollton, KY, USA. By treating soil map units as nodes and defining links as spatial adjacency, the spatial structure and complexity of the soil cover can be determined using algebraic graph theory (Phillips, 2013b). Map taken from the U.S. Department of Agriculture *Web Soil Survey* (http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm).

variability in soil-forming factors from that associated with localized instabilities or contingent factors.

3.4. Temporal dynamics

Spatial applications of graph theory have a long tradition in geography and cartography, and are well established in landscape ecology. However, applications in the temporal domain are more recent. All the examples below seek to identify system dynamics or characteristics from some type of historical record or sequence; thus the chief aims of these applications can be broadly categorized as time series analysis, identification of state transitions, and investigation of historical contingency and system "memory."

In time series analysis, Marwan et al. (2009) used recurrence networks to analyze marine paleoclimate records (Fig. 7). These graphs are constructed by identifying nodes based on vectors within phase space. The nodes are considered connected if the associated states are very similar. The graph thus represents the recurrence of specific phenomena, system states, or indicators over time. Visibility graphs are conceptually similar; in this case connections are based on geometric relationships in the time series. In terrain analysis, the nodes of visibility graphs are locations on or above the topographic surface, and edges indicate mutual visibility (Puppo and Marzano, 1997; O'Sullivan and Turner, 2001). Visibility graphs for other phenomena are constructed by treating data sets as a surface. Donner and Donges (2012) showed how visibility graphs can be applied to a variety of geophysical time series. Radebach et al. (2013) applied the method to discriminate among different types of ENSO episodes and Telesca and Lovallo (2009) analyzed seismic data using visibility graphs.

Recurrence networks and visibility graphs are sometimes employed to detect state transitions, such as climate regime shifts (e.g., Elsner et al., 2009; Donges et al., 2011; Radebach et al., 2013). Tsonis and Swanson (2012) took a more direct approach by linking graph coupling strength and synchronization to state changes in climate. Spence and Phillips (2014) also addressed state changes by identifying graph properties of pedological, ecological, and hydrological chronosequences.

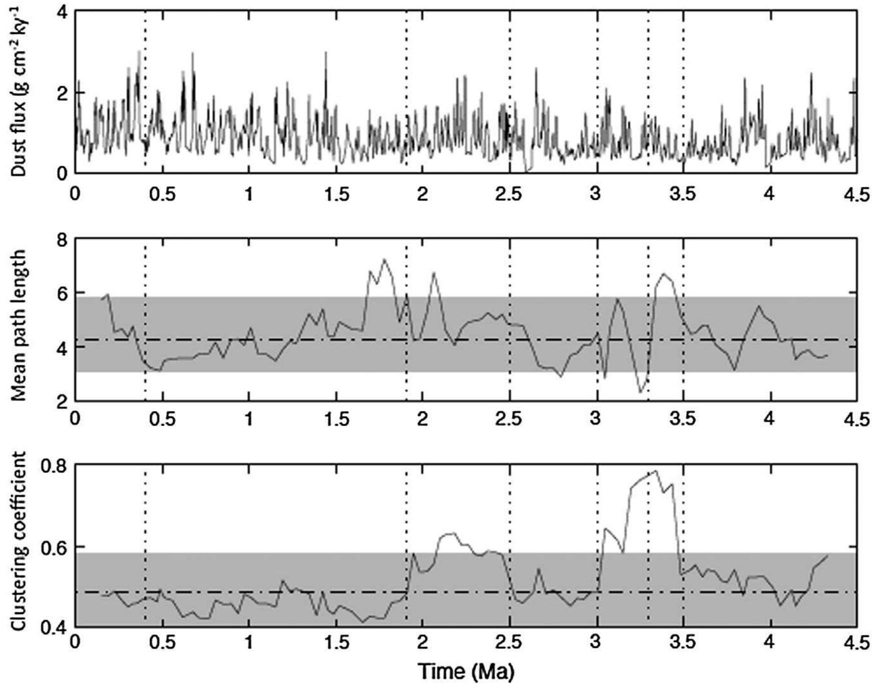

Fig. 7. Example of graph theory application to time series analysis. Using inferred terrestrial dust flux (top) from Ocean Drilling Program site 659, Marwan et al. (2009) used recurrence analysis procedures to identify nodes and links, and computed graph mean path length (middle) and connectivity (bottom). These parameters successfully identified regime shifts in African climate (slightly modified from Fig. 5 in Marwan et al., 2009).

Historical contingency and memory are related to the extent that Earth system phenomena are characterized by path dependent evolution and persistence or inheritance. Hendrix et al. (2011) and Palus et al. (2011) used graph-based clustering techniques to identify long term dependencies in temperature records. Phillips (2013a) developed methods for quantitatively estimating historical contingency in landscape evolution, applying them to several fluvial landscapes.

3.5. Dynamical system properties

The properties of Earth systems can sometimes be determined from interaction graphs composed of system properties or components as nodes, and their functional or feedback relationships comprising the links (e.g., Fig. 1). These analyses typically focus on complexity, coherence, stability, and synchronization. This approach is relatively new in the geosciences, though at least one example is much earlier: rock engineering systems concepts were used to define relevant state variables and relationships for rock mechanics by Jiao and Hudson (1995), who analyzed the resulting graph to determine its stability properties.

Huang et al. (2009) used a graph-based approach to analyze nonlinear dynamics of a landslide. While acknowledging that the approach is not yet capable of precisely predicting the time of failure, their analysis revealed evolutionary dynamics not evident from monitoring movement. Patterns of environmental change in a river delta were analyzed by Phillips (2011a), who identified environmental states, and transitions between them driven by environmental forcings. The complexity and synchronization properties of the resulting graph were assessed using algebraic graph theory to address the issue of the likelihood of asynchronous, spatially complex responses vs. advancing fronts from, e.g., rising sea-level or reduced fluvial inputs. Similar tools were used to examine scale linkage in geomorphic systems and geomorphic state transitions (Phillips, 2013a, 2014).

3.6. Network optimization and design

Engineering, computer science, and operations research studies commonly use graph theory methods to support design, assessment, and optimization of human-constructed networks. There exist a few examples of such approaches in the geosciences. A graph-based method to streamline digitization of geological maps was developed by Bouille (1976). More recently, Poulter et al. (2008) applied graph analysis to assess the vulnerability of artificial coastal drainage networks to sea level rise. Cui et al. (2011) employed graph theory to optimize drainage network representations for operational river forecasting, and Segurado et al. (2013) used measures of structural connectivity in rivers to prioritize stream restoration. By discretizing the surface ocean to define nodes, and identifying edges based on mass transport using Lagrangian modeling, Rossi et al. (2014) defined an ocean connectivity network. They applied graph theory methods to the problem of defining marine protection reserves based on larval transport.

4. Discussion

Why the upsurge in graph theory applications in the geosciences? Or, why would or should a geoscientist consider applying these methods? One answer is that graph theory provides an additional set of analytical tools for scientific problem solving, which happens to have experienced considerable development and elaboration in the past 15 years. It is to be expected that any set of tools with general applicability in science and engineering will have some utility in geosciences. Thus it is not surprising that a general exponential increase in publications on, e.g., small-world or scale-free networks would include some examples from Earth, ocean, and atmospheric sciences. However, there are reasons to believe graph theory applications will become increasingly important, beyond the effects of a general scientific trend that includes geosciences.

4.1. Complex systems

Graph theory by definition addresses system components and the interconnections and interrelationships among them. It is thus well suited to systems analyses, and particularly analysis of complex systems characterized by numerous components, complicated patterns of interaction, and nonlinearity. As stated in the introduction, these traits are typical of geoscience problems.

The global carbon cycle, for instance, is characterized by a complex network of interactions between multiple components. Logofet (1997) used graph theory methods to test the stability of several representations of the C cycle, as a step toward understanding the extent to which apparent resilience is affected by how the cycle is represented in models. Other examples of geoscience applications of graph theory to complex systems problems (such as measuring complexity, detecting chaotic dynamics, and determining dynamical stability) were given in Section 3.5.

4.2. "Big data"

Technological developments both in geosciences and in data storage and transfer have led to both opportunities to examine very large, dense data sets, and problems in making sense of very large data sets. Many graph theory methods – for example, some of the measures listed in Tables 2 and 3 – lend themselves to both characterization of large, complex networks and the identification of critical or disproportionately influential members of the network.

For instance, large climate datasets have been analyzed to identify critical nodes (locations and regions) in the climate network, and to identify patterns of connectivity. The latter are basically invisible without network analysis (e.g., Tsonis and Roebber, 2004; Tsonis et al., 2008; Palus et al., 2011; Zagouras et al., 2012; Hlinka et al., 2013). Steinhaeuser et al. (2011, 2012) applied graph theory-based methods of data mining to climate data.

A suite of techniques to segment remotely sensed imagery into meaningful objects is related to spectral graph theory (Shi and Malik, 2000). Finding patterns in unstructured data such as massive, geographical data such as point clouds derived from airborne or ground based laser scanning has been supported by graph theory. de Almeida et al. (2013) use graph-traversal techniques to detect higher-level spatial structures such as buildings and their spatial arrangement from LiDAR point cloud data in urban areas. Nelson et al. (2014) employ a spectral clustering approach to automatically delineate river bed-surface patches from high-resolution laser scanned DEMs of a flume experiment.

Graph theoretic algorithms are increasingly applied in representing and analyzing flow networks in digital elevation models (DEMs). In terms of graph theory, vertices in flow networks are grid cells between which directed links exist if there is an exchange of water from one cell to a neighboring cell. Links may be weighted if multiple flow directions exist such that flow from one cell is distributed among several neighbors (Schwanghart and Kuhn, 2010). Flow networks are thus represented by directed, acyclic networks (DAG), a property which ensures that flow networks can be topologically sorted. A topological sorting refers to a ordering of vertices that ensures that vertices appear in a vertex list before their downstream neighbors. Topological sorting has recently been exploited by Braun and Willett (2013), Huang and Lee (2013) and Schwanghart and Scherler (2014) to speed up flow-related algorithms such as flow accumulation, drainage basin delineation or flow distance calculations, outperforming conventional approaches that utilize recursive calculations on grids. Arge et al. (2003) introduced grid graphs, watershed graphs and flow graphs to facilitate memory efficient input/output algorithms on massive DEMs.

4.3. Spatial fluxes and interactions

Geosciences are inherently geographical; they are often concerned with spatial interactions and flows and exchanges of mass and energy. Such fluxes through networks are increasingly being addressed using the concept of connectivity (Larsen et al., 2012; Bracken et al., 2013, 2014). The roots of graph theory are largely concerned with spatial and cartographic problems. There exists an extensive tradition of graph theory analyses in spatial contexts such as landscape ecology and transportation geography. Many recent and developing applications of graph theory can be viewed as the convergence of this spatial analysis tradition with recent and emerging complex systems applications. Examples include Pardo-Igúzquiza et al.'s (2012) simulations of karst conduit networks, Larsen et al.'s (2012) studies of wetland hydrology, Heckmann and Schwanghart's (2013) work on sediment cascades, Marra et al.'s (2014) analysis of multichannel fluvial systems, and Phillips (2013a) analysis of spatial structures of soil landscapes.

4.4. State transitions

In ecology, conceptual models of change have expanded from classical succession theories postulating a monotonic, predictable sequence of stages toward a single stable state. Multiple possible successional pathways, alternative stable states, state-and-transition models and other frameworks recognize the multiple possible states (e.g., vegetation communities) and possible transitions between them.

Analogous developments have occurred in the geosciences. Huggett (1995, 1997), for example, articulated an expansion from developmental views of a single pathway towards a predetermined end-state to evolutionary perspectives that recognize the possibility of path-dependent development. Alternative stable states are now recognized, for instance, in global climate (e.g., Overpeck and Cole, 2006; Holmes et al., 2011), and some traditional developmental frameworks such as channel evolution models have recently been expanded to include multiple states and pathways in a state-and-transition style framework (Van Dyke, 2013). These more complex patterns of change (as well as traditional linear and cyclical patterns) are readily represented, implicitly or explicitly, in a state-and-transition model (STM) framework. STMs are a type of network, and are thus amenable to analysis with graph theory methods (Phillips, 2011a,b).

5. Conclusions

5.1. Summary

More, and increasingly sophisticated, applications of graph theory concepts and methods in geosciences have appeared in recent years, principally in spatially explicit modeling, studies of small-world networks, and structural models of Earth surface systems. Metrics and methods applied in Earth and environmental sciences fall generally into three categories: connectivity measures of entire networks; measures assessing various aspects of the importance or influence of particular nodes, links, or regions of the network; and indicators of system dynamics based on graph adjacency matrices.

Geoscience applications of graph theory can be grouped, based on the primary goal or purpose, in five general categories: (1) Identification and quantification of complex network properties such as connectivity, centrality, and clustering; (2) Tests for evidence of particular types of network or system structures, such as small-world or scale-free networks; (3) Testing and evaluation of dynamical system properties, e.g., complexity, coherence, stability, and synchronization; (4) Identification of dynamics from historical records or time series; and (5) spatial analysis, continuing the long tradition of graph theory applications to spatial problems. A sixth application area, design, testing, and optimizations of networks constructed or manipulated by humans, has also seen a few geoscience applications.

Recent and future expansion of graph theory applications in geosciences is in part tied in with a general growth of network-based approaches in science. However, several factors make graph theory especially well suited in Earth, ocean, atmospheric, and geographical sciences. Complexity science, exploration of very large data sets, focus on spatial fluxes and interactions, and increasing attention to state transitions are all amenable to analysis using graph theory approaches.

5.2. Future prospects

Graph theory clearly provides some very useful tools for geoscientists, and there exist several examples of novel, fundamental insights revealed by graph theory analyses. No scientist can become proficient with every potentially useful mathematical and statistical tool, but graph theory should certainly be included on the standard menu of relevant methods for Earth and atmospheric sciences. Viewing and treating systems (Earth and otherwise) as complex networks has been touted as an emerging paradigm across the sciences (e.g., Chung and Lu, 2006). Earthquake and climate networks have emerged as distinct subfields of seismology and climatology (Abe and Suzuki, 2004, 2012; Peron et al., 2014), and extension of the state-and-transition model framework to geomorphology as a guiding conceptual framework has been proposed (Phillips, 2011a,b; Van Dyke, 2013). Connectivity, for the quantification and analysis of which graph theory methods represent a very promising approach, is now a critical issue in hydrology, geomorphology, and ecology. Thus we believe that graph theory tools and network-based conceptual models more generally will become increasingly important in the geosciences.

Research to this point also suggests some intriguing possibilities. For instance, both Abe and Suzuki's (2012) earthquake and Vevatne et al.'s (2014) ice fracture studies indicate some type of preferential attachment phenomena at work. This may in turn be related to more general geophysical principles regarding behaviors of brittle solids. Similarly, in the study of both long-distance climate teleconnections (e.g., Gozolchiani et al., 2008; Tsonis et al., 2008) and scale linkage (e.g., Phillips, 2012, 2013a) it appears that SWN network structures may explain the functional connections between distant phenomena where the causal chain is difficult or impossible to observe directly. It is early days yet in graph theory applications, and other convergent phenomena are likely to be identified.

There is also great promise for network-based approaches in identifying archetypes or prototypes that may further facilitate linking of Earth system phenomena. Previously, these have been based on properties of individual objects or areas rather than properties of interconnections and linkages. Graph theory has been used thus far to identify and categorize synoptic atmospheric circulation patterns, ENSO phenomena, hydrological response regions, climate zones, and chronosequence archetypes (Zagouras et al., 2012; Radebach et al., 2013; Halverson and Fleming, 2014; Peron et al., 2014; Phillips, 2014; Spence and Phillips, 2014). The recognition of common phenomenologies (or not) with respect to network structures has the potential to shed considerable light on the ultimate complex network—the Earth system.

Acknowledgements

Comments by two anonymous reviewers on a previous draft helped us improve this paper considerably, and we appreciate it.

References

Abe, S., Suzuki, N., 2004. Scale-free network of earthquakes. EPL Europhys. Lett. 65, 581–586.

Abe, S., Suzuki, N., 2006. Complex-network description of seismicity. Nonlinear Process. Geophys. 13, 145–150.

Abe, S., Suzuki, N., 2007. Dynamical evolution of clustering in complex network of earthquakes. Eur. Phys. J. B 59, 93–97.

Abe, S., Suzuki, N., 2012. Dynamical evolution of the community structure of complex earthquake network. EPL Europhys. Lett. 99, 39001. http://dx.doi.org/10.1209/0295-5075/99/39001.

Adriaensen, F., Chardon, J., de Blust, G., Swinnen, E., Villalba, S., Gulinck, H., Matthysen, E., 2003. The application of 'least-cost' modelling as a functional landscape model. Landsc. Urban Plan. 64, 233–247.

Ali, G.A., Roy, A.G., 2009. Revisiting hydrologic sampling strategies for an accurate assessment of hydrologic connectivity in humid temperate Systems. Geogr. Compass 3, 350–374.

Ali, G.A., Roy, A.G., 2010. Shopping for hydrologically representative connectivity metrics in a humid temperate forested catchment. Water Resour. Res. 46. http://dx.doi.org/10.1029/2010WR009442.

Arge, L., Chase, J.S., Halpin, P., Toma, L., Vitter, J.S., Urban, D., Wickremesinghe, R., 2003. Efficient flow computation on massive grid terrain datasets. GeoInformatica 7, 283–313.

Arlinghaus, S.L., Arlinghaus, W.C., Harary, F., 2002. Graph Theory and Geography. An Interactive View. John Wiley, New York.

Aurousseau, P., Gascuel-Odoux, C., Squividant, H., Trepos, R., Tortrat, F., Cordier, M.O., 2009. A plot drainage network as a conceptual tool for the spatial representation of surface flow pathways in agricultural catchments. Comput. Geosci. 35, 276–288.

Baek, W.H., Lim, G., Kim, K., Choi, Y.-J., Chang, K.-H., Jung, J.-W., Yi, M., Lee, D.-I., Ha, D.-H., 2012. Analysis of topological properties in a seismic network. Phys. A 391, 2279–2285.

Baiesi, M., Paczuski, M., 2004. Scale-free networks of earthquakes and aftershocks. Phys. Rev. E 69, 066106. http://dx.doi.org/10.1103/PhysRevE.69.066106.

Beven, K., 2014. What we see now: even-persistence and the predictability of hydro-eco-geomorphological systems. Ecol. Model. http://dx.doi.org/10.1016/j.ecolmodel.2014.07.019.

Biggs, N., 1994. Algebraic Graph Theory. 2nd ed. Cambridge University Press.

Blöthe, J.H., Korup, O., 2013. Millennial lag times in the Himalayan sediment routing system. Earth Planet. Sci. Lett. 382, 38–46.

Bouille, F., 1976. Graph theory and digitization of geological maps. J. Int. Assoc. Math. Geol. 8, 375–393.

Bracken, L.J., Croke, J., 2007. The concept of hydrological connectivity and its contribution to understanding runoff-dominated geomorphic systems. Hydrol. Process. 21, 1749–1763.

Bracken, L., Wainwright, J., Ali, G., Tetzlaff, D., Smith, M., Reaney, S., Roy, A., 2013. Concepts of hydrological connectivity: Research approaches, pathways and future agendas. Earth Sci. Rev. 119, 17–34.

Bracken, L.J., Turnbull, L., Wainwright, J., Bogaart, P., 2014. Sediment connectivity: a framework for understanding sediment transfer at multiple scales. Earth Surf. Process. Landf. http://dx.doi.org/10.1002/esp.3635.

Braun, J., Willett, S.D., 2013. A very efficient O(n), implicit and parallel method to solve the stream power equation governing fluvial incision and landscape evolution. Geomorphology 180–181, 170–179.

Brierley, G., Fryirs, K., Jain, V., 2006. Landscape connectivity. The geographic basis of geomorphic applications. Area 38, 165–174.

Briesch, R.L., 2011. Lost in a Cave: Applying Graph Theory to Cave Exploration. National Speleological Society, Huntsville, AL.

Bunn, A., Urban, D., Keitt, T., 2000. Landscape connectivity: A conservation application of graph theory. J. Environ. Manag. 59, 265–278.

Calabrese, J.M., Fagan, W.F., 2004. A comparison-shopper's guide to connectivity metrics. Front. Ecol. Environ. 2, 529–536.

Cantwell, M.D., Forman, R.T., 1993. Landscape graphs: Ecological modeling with graph theory to detect configurations common to diverse landscapes. Landsc. Ecol. 8, 239–255.

Carling, P., Jansen, J., Meshkova, L., 2014. Multichannel rivers: their definition and classification. Earth Surf. Process. Landf. 39, 26–37.

Chung, F., Lu, L., 2006. Complex Graphs and Networks (Cbms Regional Conference Series in Mathematics). American Mathematical Society, Boston.

Cohen, B.L., Kaulfuss, A., Lüter, C., 2014. Craniid brachiopods: aspects of clade structure and distribution reflect continental drift. Zool. J. Linnean Soc. 171, 133–150.

Cui, Z.T., Koren, V., Cajina, N., Voellmy, A., Moreda, F., 2011. Hydroinformatics advances for operational river forecasting: using graphs for drainage network descriptions. J. Hydroinf. 13, 181–197.

Dale, M.R.T., Fortin, M.-J., 2010. From graphs to spatial graphs. Annu. Rev. Ecol. Evol. Syst. 41, 21–38.

de Almeida, J.P., Morley, J.G., Dowman, I.J., 2013. A graph-based algorithm to define urban topology from unstructured geospatial data. Int. J. Geogr. Inf. Sci. 27, 1514–1529.

Donges, J.F., Zou, Y., Marwan, N., Kurths, J., 2009. The backbone of the climate network. EPL Europhys. Lett. 87, 48007. http://dx.doi.org/10.1209/0295-5075/87/48007.

Donges, J.F., Donner, R.V., Trauth, M.H., Marwan, N., Schellnhuber, H.-J., Kurths, J., 2011. Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. Proc. Natl. Acad. Sci. U. S. A. 108. http://dx.doi.org/10.1073/pnas.1117052108.

Donner, R.V., Donges, J.F., 2012. Visibility graph analysis of geophysical time series: potentials and possible pitfalls. Acta Geophys. 60, 589–623.

Duan, Z.-S., Wang, W.-X., Chao, L., Chen, G.-R., 2009. Are networks with more edges easier to synchronize, or not? Chin. Phys. 18, 3122–3129.

Elsner, J.B., Jagger, T.H., Fogarty, E.A., 2009. Visibility network of United States hurricanes. Geophys. Res. Lett. 36, L16702. http://dx.doi.org/10.1029/2009GL039129.

Erős, T., Olden, J.D., Schick, R.S., Schmera, D., Fortin, M.-J., 2012. Characterizing connectivity relationships in freshwaters using patch-based graphs. Landsc. Ecol. 27, 303–317.

Fiedler, M., 1973. Algebraic connectivity of graphs. Czechoslov. Math. J. 23, 298–305.

Fryirs, K., 2013. (Dis)Connectivity in catchment sediment cascades: a fresh look at the sediment delivery problem. Earth Surf. Process. Landf. 38, 30–46.

Galpern, P., Manseau, M., Fall, A., 2011. Patch-based graphs of landscape connectivity: a guide to construction, analysis and application for conservation. Biol. Conserv. 144, 44–55.

Gascuel-Odoux, C., Aurousseau, P., Doray, T., Squividant, H., Macary, F., Uny, D., Grimaldi, C., 2011. Incorporating landscape features to obtain an object-oriented landscape drainage network representing the connectivity of surface flow pathways over rural catchments. Hydrol. Process. 25, 3625–3636.

Gozolchiani, K., Yamaski, K., Gazit, O., Havlin, S., 2008. Pattern of climate network blinking links follows El Nino events. EPL Europhys. Lett. 83, 28005. http://dx.doi.org/10.1209/0295-5075/83/28005.

Grace, J.B., Schoolmaster Jr., D.R., Guntenspergen, G.R., Little, A.M., Mitchell, B.R., Miller, K.M., Schweiger, E.W., 2012. Guidelines for a graph-theoretic implementation of structural equation modeling. Ecosphere 3 (article 73).
Halverson, M., Fleming, S., 2014. Complex networks, streamflow, and hydrometric monitoring design. Hydrol. Earth Syst. Sci. Discuss. 11, 13663–13710.
Harary, F., 1994. Graph Theory. Perseus.
Heckmann, T., Schwanghart, W., 2013. Geomorphic coupling and sediment connectivity in an alpine catchment - exploring sediment cascades using graph theory. Geomorphology 182, 89–103.
Heckmann, T., Schwanghart, W., Phillips, J.D., 2014a. Graph theory—recent developments and its application in geomorphology. Geomorphology http://dx.doi.org/10.1016/j.geomorph.2014.12.024.
Heckmann, T., Hilger, L., Mebenzehl, K., Hoffman, T., Schwangart, W., Götz, J., Buckel, J., 2014b. Network analysis of sediment cascades derived from digital geomorphological maps—a comparative study of three catchments in the Austrian and Swiss Alps. Geophys. Res. Abstr. 15, 2014–11452 (EGU).
Hendrix, W., Tetteh, I.K., Agrawal, A., Semazzi, F., Liao, W.-K., Choudhary, A., 2011. Community dynamics and analysis of decadal trends in climate data. IEEE International Conference on Data Mining Workshops 11, pp. 9–14 http://dx.doi.org/10.1109/ICDMW/2011.51.
Hlinka, J., Hartman, D., Vejmelka, M., Runge, J., Marwan, N., Kurths, J., Palus, M., 2013. Reliability of inference of directed climate networks using conditional mutual information. Entropy 15, 2023–2045.
Holmes, J., Lowe, J., Wolff, E., Srokosz, M., 2011. Rapid climate change: lessons from the recent geological past. Glob. Planet. Chang. 79, 157–162.
Huang, P.-C., Lee, K.T., 2013. An efficient method for DEM-based overland flow routing. J. Hydrol. 489, 238–245.
Huang, Z., Law, K.T., Liu, H., Jiang, T., 2009. The chaotic characteristics of landslide evolution: a case study of Xintan landslide. Environ. Geol. 56, 1585–1591.
Huggett, R.J., 1995. Geoecology. An Evolutionary Approach. Routledge, London.
Huggett, R.J., 1997. Environmental Change: The Evolving Ecosphere. Routledge, London.
Jiao, Y., Hudson, J.A., 1995. The fully-coupled model for rock engineering systems. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 32, 491–512.
Jiménez, A., Tiampo, K.F., Posadas, A.M., 2008. Small world in a seismic network: the California case. Nonlinear Process. Geophys. 15, 389–395.
Kleinhans, M.G., Ferguson, R.I., Lane, S.N., Hardy, R.J., 2013. Splitting rivers at their seams: bifurcations and avulsion. Earth Surf. Process. Landf. 38, 47–61.
Kong, F., Yin, H., Nakagoshi, N., Zong, Y., 2010. Urban green space network development for biodiversity conservation: Identification based on graph theory and gravity modeling. Landsc. Urban Plan. 95, 16–27.
Kuo, C.-W., Brierley, G., 2014. The influence of landscape connectivity and landslide dynamics upon channel adjustments and sediment flux in the Liwu Basin, Taiwan. Earth Surf. Process. Landf. http://dx.doi.org/10.1002/esp.3598.
Kupfer, J.A., Meitzen, K.M., Gao, P., 2014. Flooding and surface connectivity of *Taxodium-Nyssa* stands in a southern floodplain forest ecosystem. River Res. Appl. http://dx.doi.org/10.1002/rra.2828.
Larsen, L.G., Choi, J., Nungesser, M.K., Harvey, J.W., 2012. Directional connectivity in hydrology and ecology. Ecol. Appl. 22, 2204–2220.
Li, L., Alderson, D., Doyle, J.C., Willinger, W., 2005. Towards a theory of scale-free graphs: definition, properties, and implications. Internet Math. 2, 431–523.
Logofet, D.O., 1997. Svicobians of the compartment models and DaD-stability of the Svicobians: aggregating '0-dimensional' models of global biogeochemical cycles. Ecol. Model. 104, 39–49.
Logofet, D.O., 2013. Projection matrices in variable environments: λ_1 in theory and practice. Ecol. Model. 251, 307–311.
Marra, W.A., Kleinhans, M.G., Addink, E.A., 2014. Network concepts to describe channel importance and change in multichannel systems: test results for the Jamuna River, Bangladesh. Earth Surf. Process. Landf. 39, 766–778.
Marwan, N., Donges, J.F., Zou, Y., Donner, R.V., Kurths, J., 2009. Complex network approach for recurrence analysis of time series. Phys. Lett. A 373, 4246–4254.
Minor, E.S., Urban, D.L., 2008. A graph-theory framework for evaluating landscape connectivity and conservation planning. Conserv. Biol. 22, 297–307.
Nelson, P.A., Bellugi, D., Dietrich, W.E., 2014. Delineation of river bed-surface patches by clustering high-resolution spatial grain size data. Geomorphology 205, 102–119.
O'Sullivan, D., Turner, A., 2001. Visibility graphs and landscape visibility analysis. Int. J. Geogr. Inf. Sci. 15, 221–237.
Overpeck, J.T., Cole, J.E., 2006. Abrupt change in Earth's climate system. Annu. Rev. Environ. Resour. 31, 1–31.
Pacheco, F., 1998. Finding the number of natural clusters in groundwater data sets using the concept of equivalency class. Comput. Geosci. 24, 7–15.
Palus, M., Hartman, D., Vejmelka, 2011. Discerning connectivity from dynamics in climate networks. Nonlinear Process. Geophys. 18, 751–763.
Pardo-Iguzquiza, E., Durán-Valsero, J.J., Rodríguez-Galiano, V., 2011. Morphometric analysis of three-dimensional networks of karst conduits. Geomorphology 132, 17–28.
Pardo-Igúzquiza, E., Dowd, P.A., Xu, C., Durán-Valsero, J.J., 2012. Stochastic simulation of karst conduit networks. Adv. Water Resour. 35, 141–150.
Pascual-Horta, L., Saura, S., 2006. Comparison and development of new graph-based landscape connectivity indices: towards the prioritization of habitat patches and corridors for conservation. Landsc. Ecol. 21, 959–967.
Peron, T.K.D., Comin, C.H., Amancio, D.R., Costa, L. da F., Rodrigues, F.A., Kurths, J., 2014. Correlations between climate network and relief data. Nonlinear Process. Geophys. 21, 1127–1132.
Phillips, J.D., 2002. Erosion, isostasy, and the missing peneplains. Geomorphology 45, 225–241.
Phillips, J.D., 2011a. Predicting modes of spatial change from state-and-transition models. Ecol. Model. 222, 475–484.
Phillips, J.D., 2011b. The structure of ecological state transitions: amplification, synchronization, and constraints. Ecol. Complex. 8, 336–346.
Phillips, J.D., 2012. Synchronization and scale in geomorphic systems. Geomorphology 137, 150–158.
Phillips, J.D., 2013a. Networks of historical contingency in Earth surface systems. J. Geol. 121, 1–16.
Phillips, J.D., 2013b. Sources of spatial complexity in two coastal plain soil landscapes. Catena 111, 98–103.
Phillips, J.D., 2013c. Evaluating taxonomic adjacency as a source of soil map uncertainty. Eur. J. Soil Sci. 64, 391–400.
Phillips, J.D., 2014. State transitions in geomorphic responses to environmental change. Geomorphology 204, 208–216.
Phillips, R.W., Spence, C., Pomeroy, J.W., 2011. Connectivity and runoff dynamics in heterogeneous basins. Hydrol. Process. 25, 3061–3075.
Poulter, B., Goodall, J.L., Halpin, P.N., 2008. Applications of network analysis for adaptive management of artificial drainage systems in landscapes vulnerable to sea level rise. J. Hydrol. 357, 207–217.
Puppo, E., Marzano, P., 1997. Discrete visibility problems and graph algorithms. Int. J. Geogr. Inf. Sci. 11, 139–161.
Radebach, A., Donner, R.V., Runge, J., Donges, J.F., Kurths, J., 2013. Disentangling different types of El Nino episodes by evolving climate network analysis. Phys. Rev. E 88, 052807. http://dx.doi.org/10.1103/PhysRevE.88.052807.
Rayfield, B., Fortin, M.-J., Fall, A., 2011. Connectivity for conservation: a framework to classify network measures. Ecology 92, 847–858.
Restrepo, J.G., Ott, E., Hunt, B.R., 2006. Emergence of synchronization in complex networks of interacting dynamical systems. Phys. D 224, 114–122.
Restrepo, J.G., Ott, E., Hunt, B.R., 2007. Approximating the largest eigenvalue of network adjacency matrices. Phys. Rev. E 76. http://dx.doi.org/10.1103/PhysRevE.76.056119.
Rossi, V., Ser-Giacomi, E., López, C., Hernández-Garcia, E., 2014. Hydrodynamic provinces and ocean connectivity from a transport network help designing marine reserves. Geophys. Res. Lett. 41, 2883–2891.
Santiago, A., Cardenas, J.P., Losada, J.C., Benito, R.M., Tarquis, A.M., Borondo, F., 2008. Multiscaling of porous soils as heterogeneous complex networks. Nonlinear Process. Geophys. 15, 893–902.
Schreiber, M.A., Hastings, H.M., 1995. Stability of random matrix models. Rocky Mt. J. Math. 25, 471–478.
Schwanghart, W., Kuhn, N.J., 2010. TopoToolbox: A set of Matlab functions for topographic analysis. Environ. Model Softw. 25, 770–781.
Schwanghart, W., Scherler, D., 2014. TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. Earth Surf. Dyn. 2, 1–7.
Segurado, P., Branco, P., Ferreira, M.T., 2013. Prioritizing restoration of structural connectivity in rivers: a graph based approach. Landsc. Ecol. 28, 1231–1238.
Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22, 888–905.
Siek, J.G., Lee, L.-Q., Lumsdaine, A., 2002. The Boost Graph Library: User Guide and Reference Manual (C++ in depth). Addison-Wesley Longman, Amsterdam.
Spence, C., Phillips, R.W., 2014. Refining understanding of hydrological connectivity in a boreal catchment. Hydrol. Process. http://dx.doi.org/10.1002/hyp.10270.
Steinhaeuser, K., Chawla, N.V., Ganguly, A.R., 2011. Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science. Stat. Anal. Data Min. 4, 497–511.
Steinhaeuser, K., Ganguly, A.R., Chawla, N.V., 2012. Multivariate and multiscale dependence in the global climate system revealed through complex networks. Clim. Dyn. 39, 889–895.
Stern, H., 2003. Progress on a knowledge-based internet weather forecasting system. Commonwealth Bureau of Meteorology, Australia (URL: http://www.weather-climate.com/internetforecasts.html (last accessed 9 January, 2015).).
Taylor, P.D., Fahrig, L., Henein, K., Merriam, G., 1993. Connectivity is a vital element of landscape structure. Oikos 68, 571–573.
Telesca, L., Lovallo, M., 2012. Analysis of seismic sequences by using the method of visibility graph. Europhys. Lett. 97, 5002. http://dx.doi.org/10.1209/0295-5075/97/5002.
Tinkler, K.J., 1972. The physical interpretation of eigenfunctions of dichotomous matrices. Trans. Inst. Br. Geogr. 55, 17–46.
Tischendorf, L., Fahrig, L., 2000. How should we measure landscape connectivity? Landsc. Ecol. 15, 633–641.
Tsonis, A.A., Roebber, P.J., 2004. The architecture of the climate network. Phys. A 333, 497–504.
Tsonis, A.A., Swanson, K., 2008. Topology and predictability of El Nino and La Nina networks. Phys. Rev. Lett. 100, 228502. http://dx.doi.org/10.1103/PhysRevLett.100.228502.
Tsonis, A.A., Swanson, K.L., 2012. On the origins of decadal climate variability: a network perspective. Nonlinear Process. Geophys. 19, 559–568.
Tsonis, A.A., Swanson, K.L., Roebber, P.J., 2006. What do networks have to do with climate? Bull. Amer. Meteor. Soc. 87, 585–595.
Tsonis, A.A., Swanson, K.L., Wang, G., 2008. On the role of atmospheric teleconnections in climate. J. Clim. 21, 2990–3001.
Urban, D., Keitt, T., 2001. Landscape connectivity: A graph-theoretic perspective. Ecology 82, 1205–1218.
Urban, D.L., Minor, E.S., Tremi, E.A., Schick, R.S., 2009. Graph models of habitat mosaics. Ecol. Lett. 12, 260–273.
Valentini, L., Perugini, D., Poli, G., 2007a. The "small-world" topology of rock fracture networks. Phys. A 377, 323–328.
Valentini, L., Perugini, D., Poli, G., 2007b. The "small-world" nature of fracture/conduit networks: possible implications for disequilibrium transport of magmas beneath mid-ocean ridges. J. Volcanol. Geotherm. Res. 159, 355–365.

Van Dyke, C., 2013. Channels in the making—an appraisal of channel evolution models. Geogr. Compass 7, 759–777.
van Langevelde, F., van der Knapp, W.G.M., Claasen, G.D.H., 1998. Comparing connectivity in landscape networks. Environ. Plan. B Plan. Des. 25, 849–863.
Vevatne, J.N., Rimstad, E., Hope, S.M., Korsnes, R., Hansen, A., 2014. Fracture networks in sea ice. Front. Phys. http://dx.doi.org/10.3389/fphy.2014.00021.
Watts, D., 1999. Small Worlds. Princeton University Press.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of "small-world" networks. Nature 393, 440–442.
Werner, C., 1993. Ridge-channel path interdependence in drainage basins. Geogr. Ann. Ser. A Prog. Phys. Geogr. 75, 73–81.
Werner, C., 1994. Explorations into the formal structure of drainage basins. Earth Surf. Process. Landf. 19, 747–762.
Wright, C.K., 2010. Spatiotemporal dynamics of prairie wetland networks: power-law scaling and implications for conservation planning. Ecology 91, 1924–1930.
Yang, X.S., 2001. Small-world networks in geophysics. Geophys. Res. Lett. 28, 2549–2552.
Yuan, W.-J., Luo, X.-S., Jiang, P.-Q., Wang, B.-H., Fang, J.-Q., 2008. Transition to chaos in small-world dynamical network. Chaos, Solitons Fractals 37, 799–806.
Zagouras, A., Argiriou, A.A., Flocas, H.A., Economou, G., Fotopoulos, S., 2012. An advanced method for classifying atmospheric circulation types based on prototypes connectivity graph. Atmos. Res. 118, 180–192.
Zaliapin, I., Foufoula-Georgiou, E., Ghil, M., 2010. Transport on river networks: a dynamic tree approach. J. Geophys. Res. Earth Surf. 115, F00A15. http://dx.doi.org/10.1029/2009JF001281.
Zetterberg, A., Mörtberg, U.M., Balfors, B., 2010. Making graph theory operational for landscape ecological assessments, planning, and design. Landsc. Urban Plan. 95, 181–191.