Modelling dendritic ecological networks in space: an integrated network perspective

Abstract

Dendritic ecological networks (DENs) are a unique form of ecological networks that exhibit a dendritic network topology (e.g. stream and cave networks or plant architecture). DENs have a dual spatial representation; as points within the network and as points in geographical space. Consequently, some analytical methods used to quantify relationships in other types of ecological networks, or in 2-D space, may be inadequate for studying the influence of structure and connectivity on ecological processes within DENs. We propose a conceptual taxonomy of network analysis methods that account for DEN characteristics to varying degrees and provide a synthesis of the different approaches within the context of stream ecology. Within this context, we summarise the key innovations of a new family of spatial statistical models that describe spatial relationships in DENs. Finally, we discuss how different network analyses may be combined to address more complex and novel research questions. While our main focus is streams, the taxonomy of network analyses is also relevant anywhere spatial patterns in both network and 2-D space can be used to explore the influence of multi-scale processes on biota and their habitat (e.g. plant morphology and pest infestation, or preferential migration along stream or road corridors).

Keywords
Connectivity, dendritic ecological network, directionality, network analysis, scales, spatial statistics, stream network, taxonomy of network analyses.

INTRODUCTION

The use of network frameworks and analyses to gain a better understanding of ecological structure and function has dramatically increased in recent years (Proulx et al. 2005; Dale & Fortin 2010). Network models, such as graph-theoretic-based approaches (Urban et al. 2009), are simplifications of reality used to conceptualise and describe relationships, either qualitatively or quantitatively, between a set of components interacting as an ecological system. Such topological structures have been used implicitly and explicitly in meta-population (Hanski 1998), metacommunity (Cadotte 2006), and metacoeosystem models (Massol et al. 2011). The appeal of a network-based approach across a suite of ecological and evolutionary systems stems from the explicit emphasis on the functional relationships (e.g. edges, links) between the entities of interest (e.g. nodes, points, patches). Hence, the same network framework can be used to investigate the effects of various processes, such as gene flow (Fortuna et al. 2009), predator–prey relationships (Bascompte et al. 2005) or energy fluxes (Proulx et al. 2005), between nodes (e.g. individuals, populations, communities). These network models are commonly applied to ecological phenomena represented non-spatially, such as multi-trophic interactions at a single location or area (e.g. Bascompte et al. 2005). In other cases, physical space is treated as a network (Gilarranz & Bascompte 2012), with nodes representing spatially explicit habitat patches, and edges denoting processes such as the rate of dispersal between habitat patches (Munepeerakul et al. 2008) or a georeferenced dispersal pathway (Fall et al. 2007).

Dendritic ecological networks (DENs; Grant et al. 2007) are used to describe spatial relationships in ecosystems that naturally exhibit a physical dendritic network topology (e.g. stream and cave ecosystems).
networks or plant architecture). A number of characteristics differentiate DENs from other types of ecological networks (Grant et al. 2007). First, movement of organisms, material or energy is primarily restricted to the physical network, which forms ecological corridors (Rodríguez-Iturbe et al. 2009). However, the permeability of these ecological corridors varies depending on the organism or process of interest. For example, troglobites may never move outside of a cave network (Barr & Holsinger 1985), while semi-aquatic organisms living in, or nearby streams may also move across the terrestrial landscape (Carranza et al. 2012). Second, DENs have fewer redundant pathways compared to other non-spatial or spatially structured ecological networks; though braiding may occur in some stream or cave networks. Third, directionality may also be important in some DENs, such as cave and stream networks, where flowing water strongly influences physicochemical and biological processes. Fourth, biological and physicochemical processes are not restricted to nodes, with relationships between nodes represented as edges; instead, processes occur on the network in DENs. For example, the availability and spatial arrangement of in-stream habitat may influence the potential distribution of a fish species, while the branching structure of the network affects in-stream dispersal to those habitats and resulting species interactions (e.g. competition and predation; Angermeier et al. 2002). Given the unique characteristics of DENs and the spatial complexity of processes on the physical network, many analytical methods used to quantify relationships in other types of ecological networks, or in 2-D space, are unsuitable for studying the influence of structure and connectivity on physicochemical and biological processes in these systems (Grant et al. 2007).

A variety of methods can be used to analyse DEN data, but they are scattered across the literature, ranging from graph-theoretic approaches (Dale & Fortin 2010), semivariogram analyses (Ganio et al. 2005), to metapopulation modelling (Fagan 2002). Thus, many researchers may be unfamiliar with these methods, as well as, the software needed to implement them. Instead, parametric statistical methods are commonly used to analyse DEN data, which either ignore spatial relationships altogether, or assume that proximity and connectivity are adequately described using Euclidean distance (e.g. ignore network topology). When these methods are used the implicit assumption is that topological relationships within the network are unimportant. Thus, there is a mismatch between conventional analytical approaches and the evolving ecological conceptualisation of DENs. This disparity limits our understanding of how DENs function, weakens our ability to make accurate and unbiased predictions of DEN attributes, and ultimately reduces the effectiveness of management actions in these unique ecosystems.

Our aim is to present a conceptual taxonomy of existing network analyses, which allows us to describe the characteristics of, and draw distinctions between, approaches used to describe network structure and connectivity within DENs. Within this context, we describe in more detail a new class of spatial statistical models for DENs that addresses a significant gap in previous approaches, with potential extensions to these methods. Finally, we discuss ways to combine different network analyses so that more complex and novel research questions may be explored. We mainly focus on freshwater stream networks because they are a common form of DEN (Box 1) and play a significant role in structuring spatio-temporal patterns and processes in both aquatic and terrestrial systems (Paola et al. 2006). In addition, human water security and threats to aquatic biodiversity are a major global concern (Vörösmarty et al. 2000).

Box 1 Stream ecosystems as dendritic ecological networks

Most studies in lotic freshwater ecology (i.e. flowing, freshwater streams) have been undertaken at two disparate scales (Angermeier et al. 2002; Fausch et al. 2002); local studies of abundance or biotic interactions at discrete locations (≤ 200 m), and macroscale studies (> 100 km), which often use coarse, catchment or stream network averages to provide inference about species distributions, evolutionary process, and more recently, climate change impacts (e.g. Sanderson et al. 2009). However, key biological and physical processes, such as metapopulation dynamics and disturbance regimes, are thought to operate at intermediate scales (Schlosser & Angermeier 1995; Ward 1998; Fausch et al. 2002; Benda et al. 2004), where detailed information is often lacking (Falkse & Fausch 2010). This is also the scale at which conservation agencies and managers typically perceive the landscape and interact to prioritise conservation actions (Fausch et al. 2002). As a result, a spatially continuous view of streams and rivers over intermediate scales (1–100 km; Fausch et al. 2002) within the dendritic ecological network (DEN; Grant et al. 2007) is needed to better understand key physicochemical and biological processes (Schlosser & Angermeier 1995; Fisher 1997; Ward 1998; Fausch et al. 2002; Power & Dietrich 2002; Wiens 2002; Benda et al. 2004; Fisher et al. 2004). The terminology used to express these ideas is usually different than the terminology used in graph or metapopulation theory, but the goals are similar; to learn about ecosystem processes by investigating relationships among a set of components, or locations, rather than treating discrete locations independently (Proulx et al. 2005). Therefore, we refer to this intermediate scale as the ‘network scale’ or ‘network perspective’.

Fundamentally important characteristics of streams, such as their dendritic network structure, connectivity, stream-flow direction and spatio-temporal variability of in-stream habitat and flow, are particularly influential at the network scale. For example, the catchment (land area that a stream network drains) provides nutrient inputs to streams (i.e. lateral connectivity), where in-stream processes alter the form and concentration of those nutrients, which are then transported downstream (i.e. longitudinal connectivity; Finlay et al. 2011). In turn, mobile organisms such as fishes and amphibians respond to the spatio-temporal arrangement of conditions within or along the network (Fausch et al. 2002), which forms semi-restrictive corridors for the transport of water, materials and organisms (Grant et al. 2007). Barriers such as dams and water diversions may also cause longitudinal fragmentation in the network (Ward 1998), making it challenging for relatively mobile organisms to complete life histories across stream networks (Schlosser & Angermeier 1995). In addition, food-web structure and trophic dynamics may vary depending on network structure, position within the catchment and lateral connectivity (e.g. predation by terrestrial organisms; Power & Dietrich 2002). Despite the conceptualisation of stream networks as directed and highly connected DENs, there are few studies that have successfully incorporated all of these fundamentally important stream characteristics into network-scale analyses.
Nevertheless, the concepts are also applicable to other DENs, and in some cases, other spatially structured ecological networks.

CONCEPTS

A coordinate system for DENs

A key concept, when modelling DEN data, is that observations have a dual spatial representation; as points within the network (topology) and as points in 2-D space (geography). Figure 1 illustrates this concept using a minimum planar graph (Fall et al. 2007), but topology and geography could be represented in other ways. At a minimum, the network perspective requires a dual coordinate system (Fig. 1a), with the DEN represented as a network (Fig. 1b), embedded within the 2-D geographical environment (Fig. 1c). Although it may be simpler to explore network organisation in spatial ecosystems without explicitly representing geography, critical information about ecosystem function may be lost when models fit to DEN data do not adequately account for the dual coordinate system. This concept will be further explored in subsequent sections, but it is worth noting that the need for a dual spatial representation is not a new idea; a measurement always has 2-D coordinates because it is physically collected in geographical space. A variety of conceptual (Schlosser & Angermeier 1995; Fausch et al. 2002; Benda et al. 2004), metapopulation (Fisher 1997; Hanks 1998; Fagan 2002; Fisher et al. 2004; Muneepeerakul et al. 2008) and graph-theoretic-based models (Urban & Keitt 2001; Urban et al. 2009; Dale & Fortin 2010; Gilarranz & Bascompte 2012; Jabot & Bascompte 2012) have been used to account for the dual coordinate system. However, it is worth re-emphasising this concept because ecological data continue to be modelled solely within network space, in 2-D space, or independent of space altogether.

The dual spatial representation makes modelling DEN data more complex than data represented solely in 2-D or network space. For example, 2-D space simply forms the coordinate system for obtaining samples, which is consistent across study areas (i.e. the same coordinate system). In contrast, the branching structure and connectivity represented by network space is likely to differ and, as a subspace of 2-D, has interesting properties in its own right. When network properties are inadequately described, the analysis and results may be confounded. For example, data located in the same network space, but resulting from different processes are likely to produce different results (Peterson et al. 2006); yet data collected from different networks, but resulting from the same process may also provide different results (Fagan 2002). It is this complex interplay between 2-D and network space, as well as the need to separate and understand their influence on ecological processes that makes a taxonomy of network analyses necessary.

A taxonomy of network analyses

A wide range of data types has been used to describe the physical structure of DENs, as well as the structure and function of ecological processes. These data types fall into three general categories: (1) physical network structure, (2) physicochemical and biological processes and (3) an aggregation of structure or process (Table 1). Metrics describing structure can be further sub-divided into those describing the network as-a-whole (e.g. drainage density: the total length of the network divided by catchment area) or the sub-network (e.g. stream order: a measure of upstream branching complexity). Various methods have been used to analyse these data types, which we classify as non-network, about-network, over-network and across-network analysis methods (Table 1, Fig. 2). This taxonomy of network-analysis methods is not meant to drive or be organised by ecological or biological question. Instead, it acts as a pragmatic framework to help ecologists understand the similarities and differences between analytical methods commonly used to analyse ecological processes in stream networks and other DENs.

NON-network analysis

A non-network analysis ignores the structure, connectivity and directionality of the network (Fig. 2b). Although not technically a network analysis, non-network warrants mention because many studies conducted at landscape to regional scales ignore spatial relationships between locations altogether (e.g. regression; Pandey et al. 2012). In fewer cases, spatial statistical models based on a 2-D coordinate system (Fig. 1c; Box 2) are used to account for spatial dependence between observations (Yuan 2004). Results and conclusions from many of these studies may be adversely affected by ignoring the properties of stream networks, as we demonstrate in the Spatial Statistical Methods for Network Analysis section.

Many researchers attempt to overcome the limitations of non-network analyses by including covariates that represent sub-network structure (Hitt & Angermeier 2008), direction and connectivity (Dunham & Rieman 1999; Isaak et al. 2007; Fitzcort et al. 2012; Table 1). For example, Hitt & Angermeier (2008) quantified the structural position of each survey site relative to the main stem based on stream network topology and then used Mann–Whitney U-tests to determine whether fish metrics in headwater tributaries (small segment at the periphery of the network) differed from main tributaries (larger segment draining to the main stem). Other metrics representing habitat quality, proximity, connectivity and arrangement are typically measured using least-cost path analyses and moving-window approaches (Le Pichon et al. 2006), or patch size, composition and distance measures (Dunham & Rieman 1999; Isaak et al. 2007). Incorporating measures of sub-network structure as covariates in non-network analyses allows researchers to explore specific questions relating to the influence of physical network structure and in-stream habitat on physicochemical and biological stream pro-

![Figure 1](image-url)

Figure 1 Locations within a stream network can be characterised using a (a) dual, (b) network or (c) 2-D spatial representation. A 2-D coordinate system lacks information about network connectivity. A network coordinate system only uses relative position within the network, not the 2-D coordinates; thus, distances between points within the network are equal in (b) and in that sense they are equivalent network structures. A dual representation combines the 2-D coordinate system with the network coordinate system. Various statistical models use none, one or both coordinate systems.
Spatial autocorrelation, or autocovariance, is the degree to which measurements are similar as a function of the distance separating them (i.e. separation distance). It is inherent in geographical and environmental data sets and occurs in both aquatic and terrestrial systems at multiple scales (Peterson et al. 2006; Peterson & Ver Hoef 2010). Spatial statistical modelling is a well-established branch of statistics that provides a convenient way to model these spatial dependencies (Cressie 1993; Diggle et al. 1998). As an example, we present a spatial linear-mixed model in the usual vector/matrix form,

\[ y = X\beta + z + \epsilon \]  

(1)

where \( y \) is a vector of random variables (i.e. the response variable) measured at multiple locations on the stream network(s), \( X \) is a design matrix for fixed effects, which contains the covariates (i.e. explanatory variables), \( \beta \) is a vector of parameters for the fixed effects (i.e. regression coefficients), \( z \) is a vector of random variables that are spatially correlated and \( \epsilon \) is a vector of independent random errors. The linear model is convenient because it decomposes data into three components: (1) covariates that are measured in the field or remotely, which may be spatially patterned themselves in \( X\beta \) (e.g. percent shade at a location, land use or climate); (2) unmeasured spatially patterned covariates as random variation in \( z \); these include factors that are known to be influential, but were not measured (e.g. land use or biotic interactions), as well as unknown factors resulting from a lack of understanding about the process; and (3) independent errors, including measurement errors (e.g. calibration error), in \( \epsilon \).

An autocovariance function is simply the covariance between any two values from \( z \) as a function of separation distance, controlled by the covariance parameters. Three parameters are commonly used to describe the variance structure of the spatial linear model (1): the nugget effect, the partial sill and the range. The nugget effect is the variance of \( \epsilon \) and describes the variation between sites as the separation distance approaches zero. This may be due to variation at a scale finer than the shortest separation distance or measurement error. The variance of \( z \) is called the partial sill and it is the spatially structured component of the random variation that is modelled. Note that, together the partial sill and nugget make up the sill, which represents the overall variance. Finally, the range parameter describes how fast autocorrelation decays to zero between any two values from \( z \) (e.g. the distance within which spatial autocorrelation is expected to occur).

When ecological processes are autocorrelated, a spatial statistical approach provides parameter estimates with the proper amount of uncertainty, whereas wrongly assuming independence often means that significant relationships may be identified that do not exist (Cressie 1993). In addition, spatial statistical models use autocorrelation to make better local predictions, with estimates of uncertainty at unobserved locations. An in-depth discussion of spatial statistical models can be found in Cressie (1993).
These studies account for the dual spatial representation of streams data and as a result have significantly improved our theoretical understanding of the relationship between network structure, connectivity and function. From a practical standpoint, these findings are now being used to assess management-related questions specifically focused on physical changes to connectivity, such as interbasin transfers (Grant et al. 2012). However, to our knowledge they have not been used to address issues such as the influence of land-management practices on in-stream habitat and organismal distributions, or to provide spatially explicit estimates for broad-scale monitoring (Cressie et al. 2006; Garreta et al. 2010; Isaak et al. 2010). The model predictions have also been used for a variety of purposes including the assessment of in-stream thermal suitability (Ruesch et al. 2012) and to provide spatially explicit estimates for broad-scale monitoring (Garetta et al. 2009; Money et al. 2009a). These ON-network methods are similar to traditional linear regression techniques commonly applied to point measurements in stream ecology; except that the assumption of independent errors is replaced with the notion that random errors co-vary in both 2-D and network space. These concepts will be further explored in the Spatial Statistical Methods for Network Analysis section.

**OVER-network analysis**

Data describing the physical network structure, physicochemical and biological processes or an aggregation of those characteristics within an area or feature may be summarised OVER a network or multiple networks (Table 1; Fig. 2e). The complexity of the OVER-network analysis depends on data type and spatial representation (Fig. 1). For example, measurements describing physical structure at single time points, such as confluence angles (the angle of two stream segments converging) calculated from a static geographic information system (GIS) data set, do not have a variance; therefore, a simple OVER-network summary (e.g. mean) may be sufficient. However, biological or physicochemical measurements (e.g. stream temperature) are temporally dynamic and each has their own variance, which can be summarised OVER the network(s). For instance, empirical semivariogram analysis has been used to explore OVER-network spatial

---

**Table 1** A description of network data types and the potential network analysis methods that can be used to analyse them

<table>
<thead>
<tr>
<th>Data type</th>
<th>Examples</th>
<th>Analysis method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical network structure (whole network)</td>
<td>Drainage density, fractal dimension or metric of network connectivity</td>
<td>NON  ABOUT  ON  OVER  ACROSS</td>
</tr>
<tr>
<td>Physical network structure (sub-network)</td>
<td>Confluence angle, stream order, node degree, edge weight, patch connectivity or composition</td>
<td>✔✔✔✔ ✔✔✔ ✔✔✔ ✔</td>
</tr>
<tr>
<td>Physicochemical and biological processes or attributes</td>
<td>Pool depth, pH or fish counts</td>
<td>✔✔✔</td>
</tr>
<tr>
<td>Structure and processes aggregation</td>
<td>A special case where measurements are aggregated over an area (e.g. a hydrologic unit), a network, or multiple networks</td>
<td>✔✔✔</td>
</tr>
<tr>
<td></td>
<td>Mean fish count or confluence angle</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2** (a) A real stream is represented as a network embedded in 2-D space. (b) Non-network analyses use explicitly sampled values in 2-D space, but they make no use of the network structure. (c) Analyses about a network look at characteristics of the network structure such as segment lengths, branching and connectivity. (d) Analyses on a network use explicitly sampled values at points in network space. (e) Analyses over a network average or aggregate the spatially explicit point values, whereas (f) analyses across networks compare network values, either about or over the individual networks.
dependency of Oncorhynchus clarki clarki (coastal cutthroat trout) counts in western Oregon as a function of hydrologic (i.e. in-stream) distance (Ganio et al. 2005). Multiple over-network patterns of spatial dependency may also be evaluated separately for network and 2-D space, allowing spatial patterns across the dual coordinate system to be explored simultaneously (Isak et al. 2010). Another example is block kriging, which can be used to scale up an on-network model to an over-network analysis (Ver Hoef et al. 2006).

Across-network analysis
An across-network analysis is used to compare or contrast characteristics between whole networks or sets of networks (Fig. 2f). The analysis generally takes two forms, depending on whether it is based on data describing the whole-of-network structure (a single measurement without an variance estimate) or data that have been aggregated previously using an over-network analysis (estimates with a standard error) (Table 1). Across-network analyses are relatively simple for whole-of-network structure data measured at a single time point; a t-test could be used to compare two types of networks, while an ANOVA may be appropriate when whole networks can be categorised (e.g. catchment size or climatic region), assuming proper transformations. When an across-network analysis is based on aggregated measurements of sub-network structure or physicochemical or biological processes, these data will have an associated variance measure. In that case, statistical models that incorporate measurement error (i.e. uncertainty or variance in the data value), such as a Bayesian hierarchical model (Cressie et al. 2009), should be employed. For example, the mean heavy metal concentration in networks draining mined and unmined catchments could be obtained by block kriging point samples, after which a hierarchical model that uses the means and variances from the block kriging model could be used in the across-network analysis.

SPATIAL STATISTICAL METHODS FOR NETWORK ANALYSIS
As we examined the literature and developed the taxonomy described above, it was clear that most research that explicitly acknowledges fundamental stream characteristics has been based on about- or on-network models that use sub-network data structures. Data describing stream-network structure are readily available at broad spatial scales via remote sensing or GIS data sets and may be used with all of the network-analysis methods (Table 1). The primary focus of these studies has been to investigate the influence of physical network structure on biological processes, such as dispersal (Fagan 2002; Schick & Lindley 2007). In contrast, on- or over-network models fitted to measurements of point data representing physicochemical and/or biological processes are needed to study the effects of stream processes on another physicochemical or biological response; for example, the influence of heterogeneous in-stream water quality on organismal distributions. These studies are less common because near-continuous, network-wide data sets describing in-stream processes are rare (Falke & Fausch 2010). Spatial statistical methods fill a number of needs for two types of spatial relationships based on hydrologic distance: flow-connected and flow-unconnected (Fig. 3). Two locations are considered flow-connected if water flows from the upstream location to the downstream location. Flow-unconnected locations reside on the same stream network (e.g. share a common confluence somewhere downstream), but do not share flow. Although the tail-up (Fig. 4) and tail-down (Fig. 5) models are both adapted for branching in streams and account for directionality, there are significant differences in the way that spatial relationships are represented in the two models. In the case of the ‘tail-up’ model, the tail of the moving-average function points in the upstream direction (Fig. 4).

As a result, the function must be split at confluences to allow for

over-network analyses using spatial statistical methods, which were briefly described in the On Network Analysis and Over-Network Analysis sections.

Spatial statistical modelling on stream networks
Spatial statistical modelling is a well-established branch of statistics, which provides a convenient way to model spatially dependent data (Box 2). However, standard spatial statistical models may not adequately represent the unique spatial relationships found in stream networks and other DENs. For example, lattice models, which are used to model spatial dependency in aerial units, model autocovariance (i.e. autocorrelation) based on neighbouring areas, whereas Euclidean distance has typically been used to build autocovariance models in geostatistics (i.e. kriging; Cressie 1993). These metrics of distance and proximity do not reflect the influence of dendritic structure, connectivity and directionality within a network. In addition, a model is not guaranteed to be statistically valid when hydrologic distance is used in a geostatistical model developed for Euclidean distance in 2-D space (Ver Hoef et al. 2006).

Ver Hoef & Peterson (2010) summarised the development of the tail-up and tail-down autocovariance models for stream networks (Cressie et al. 2006; Ver Hoef et al. 2006; Money et al. 2009a,b; Garreta et al. 2009), which are based on a branching, continuous spatial analogue to moving averages in time series. These models account for two types of spatial relationships based on hydrologic distance: flow-connected and flow-unconnected (Fig. 3). Two locations are considered flow-connected if water flows from the upstream location to the downstream location. Flow-unconnected locations reside on the same stream network (e.g. share a common confluence somewhere downstream), but do not share flow. Although the tail-up (Fig. 4) and tail-down (Fig. 5) models are both adapted for branching in streams and account for directionality, there are significant differences in the way that spatial relationships are represented in the two models. In the case of the ‘tail-up’ model, the tail of the moving-average function points in the upstream direction (Fig. 4).

As a result, the function must be split at confluences to allow for

Figure 3 There are two types of spatial relationships in a stream network: flow-connected and flow-unconnected. For two locations to be flow-connected, water must flow from an upstream location to a downstream location (S3 and S2, and S1). Flow-unconnected locations share a common confluence (e.g. junction) somewhere downstream, but do not share flow (S2 and S3).
The need to quantify patterns of spatial autocorrelation that are best described in network space is intuitive in a DEN where the network structure is obvious. However, DENs are also embedded within 2-D space and complex, multi-scale processes and interactions occur across the dual coordinate system. This is especially true in stream ecosystems where topographic and climatic gradients (e.g., elevation and air temperature), as well as, land-management or disturbances within the catchment and riparian zone (e.g. tree cover or wildfires) have direct and indirect effects on physicochemical and biological in-stream processes (Isaak et al. 2010). Thus, it is not uncommon for stream data to show evidence of multiple Euclidean and/or hydrologic patterns of spatial autocorrelation (Peterson et al. 2006; Garreta et al. 2009). To address this issue, autocovariance models developed for Euclidean distance may be combined with stream-network models to produce mixed models based on variance components (Ver Hoef & Peterson 2010), through an extension of the spatial linear model:

\[ y = X\beta + z_{TU} + z_{TD} + z_E + e \]  

where \( y \) is a vector of response variables, \( X \) is a matrix of covariates, \( \beta \) is a parameter vector, \( z_{TU} \) and \( z_{TD} \) are vectors of zero-mean random variables with a correlation structure based on the tail-up and tail-down stream-network models, respectively, \( z_E \) is a vector of zero-mean random variables with a correlation structure based on Euclidean distance, and \( e \) is a vector of independent random errors (see Box 2 for an overview of the spatial linear-mixed model). When spatial random effects are added to form a mixed-covariance structure and then combined with covariates within a single model, a flexible modelling framework is formed that can be used to account for measured and unmeasured variables at multiple scales (Peterson & Ver Hoef 2010).

Is it worth it?

Creating a stream-network model involves more effort than employing standard geostatistical methods. Calculating hydrologic distances and spatial weights requires advanced GIS expertise, whereas the Euclidean distances can be calculated easily using site coordinates, with or without a GIS. So, how much is really gained by using spatially explicit on-network models? We explore this with a simple example from the Middle Fork of the Lower Snake River, Idaho (Fig. 6a). Daily stream temperatures were recorded in the summer of 2004 and summarised to produce a summer mean temperature for each location (Isaak, D.J., unpublished data). We fit two models to these data: (1) a standard geostatistical model using Euclidean distance with a constant mean (no covariates) and a spherical autocovariance model (i.e. ordinary kriging; Cressie 1993), and (2) a stream-network model, with a constant mean and a tail-up spherical autocovariance model (Ver Hoef et al. 2006). A small portion of the stream network (black square, Fig. 6a) is shown in Fig. 6b, which contains three locations labelled with the observed temperature values. First, we used all of the data locations shown in Fig. 6a to estimate the covariance parameters (nugget, partial sill, and range; Box 2) for both the standard geostatistical model and the tail-up
Furthermore, the prediction intervals provided by the estimate from the stream-network model, 12.94 °C. Thus, the kriging estimate of 11.98 °C should be some weighted average of temperatures from the two upstream segments; one has a temperature of 11.16 °C, and the other temperature is unknown, but is surely greater than 12.42 °C. Thus, the kriging estimate of 11.98 °C is not sensible, whereas the estimate from the stream-network model, 12.94 °C, is much more reasonable. Furthermore, the prediction intervals provided by the tail-up model are wider, which better reflects the uncertainty coming from the physical structure of the network, rather than the interpolation based on Euclidean distance.

This simple example clearly demonstrates the potential benefits of implementing a spatial stream-network model. However, these benefits only materialise when (1) the data are spatially correlated, (2) spatial autocorrelation is best described using a hydrologic (e.g. flow-connected or flow-unconnected) rather than Euclidean relationship and (3) data are distributed across a branching network rather than a single, non-branching stream channel. In addition, the spatial distribution of survey sites has important implications on the number of neighbouring pairs used to fit the autocovariance function (Peterson et al. 2006; Box 2). If there are too few flow-connected or flow-unconnected locations, there is little to gain from fitting a spatial stream-network model.

**Generalised linear models and other extensions**

Non-Gaussian data, such as counts of organisms or species presence-absence, are commonly collected for monitoring programs and ecological studies. Spatial linear models may be applied to non-Gaussian data if transformations are used to normalise the response and homogenise the variance. However, another approach is the generalised linear model (GLM), which uses Poisson or binomial distributions directly, and this approach has already been adapted for spatial statistical models based on Euclidean distance (Diggle et al. 1998). To our knowledge, spatial GLMs using stream-network models as reviewed by Ver Hoef & Peterson (2010) have not been described in the literature, but in principle stream-network covariances can be used in GLMs in exactly the same way as Euclidean distance covariances are used; consequently, no new methodological developments are needed to fit a spatial GLM for stream networks.

**Empirical semivariogram analysis and the Torgegram**

Empirical semivariogram analysis is used to explore how the spatial dependence between observations changes as a function of distance. These patterns may be particularly interesting in stream networks, where the dendritic structure, as well as longitudinal and lateral connectivity can produce multiple patterns of spatial autocorrelation (Peterson & Ver Hoef 2010). An empirical semivariogram estimates the semivariance \( \gamma(h) \) for all \( h \neq 0 \) plotted as a function of increasing distance among observed locations, where pairwise distances (i.e. separation distances) are aggregated into bins. Empirical semivariograms that display a patterned increase in semivariance with increasing distance indicate that the data, or model residuals, exhibit positive spatial autocorrelation.

In the case of stream-network processes, a Torgegram (Ver Hoef et al. in review) is used to display semivariance as a function of hydrologic distance separately for flow-connected and flow-unconnected relationships (Fig. 3); making them useful exploratory tools for visualising different network-based patterns of spatial autocorrelation in raw data or model residuals, which we illustrate next using two examples.

**The Torgegram: Visualising network-based patterns of spatial autocorrelation**

We constructed a Torgegram using 178 mean summer stream temperature observations (Fig. 7a) collected in the Bear Valley Creek catchment (13,000 km²), upper Middle Fork of the Salmon River, Idaho, USA (Isaak, D.J., unpublished data), assuming a constant...
mean among all locations. The second Torgegram (Fig. 7b) relies on 386 juvenile Oncorhynchus mykiss (rainbow trout) abundance data collected in the Elk River catchment (238 km²), located in southwestern Oregon, USA (Burnett 2001). The Torgegram in Fig. 7b was constructed using residuals from a spatial linear model where abundance values were first log-transformed, and a regression coefficient for trend in upstream distance was estimated using model covariates. Three parameters are used to describe the shape of the semivariogram: the nugget effect, the sill, and the range (Box 2). For stream temperature, there can be only one overall sill, which appears to be around 5. The nugget effect for flow-connected sites is near zero and the semivariance increases more slowly towards the sill (Fig. 7a), which suggests that the range of spatial correlation for flow-connected stream temperature sites may be near 15,000 m, or greater. In contrast, the flow-unconnected pairs exhibit a larger nugget effect and the semivariance increases more rapidly to a range of approximately 10,000 m. These characteristics suggest that the data exhibit both flow-connected and flow-unconnected patterns of spatial autocorrelation and that fitting a mixed-covariance structure that includes both tail-up and tail-down autocorrelation models may be appropriate. A relatively strong pattern of spatial autocorrelation between flow-connected pairs is also evident in the Torgegram for the abundance residuals (Fig. 7b); the nugget effect is approximately 0.8 for flow-connected sites, with a sill near 1.6, and a range approaching 4,000 m. Yet, the semivariance for the flow-unconnected pairs does not appear to change as a function of hydrologic distance, which suggests that flow-unconnected locations may not be spatially correlated after accounting for the upstream trend. As such, adding a single tail-up autocorrelation model might be sufficient.

Abundance estimation and block kriging

Estimates of averages or totals, along with their estimated precision, over stream networks, sub-networks or stream segments are particularly important for managing populations of aquatic organisms or monitoring pollution impacts. For example, Poos et al. (2012) derived sub-catchment-scale population estimates of Clinostomus elongatus (redside dace), an endangered minnow, by extrapolating pool-scale density estimates based on a combination of quantitative and qualitative rules. These pool and sub-catchment-scale estimates were then used to better understand the relationship between the distribution of redside dace and impervious land use. Classical approaches to abundance estimation use exhaustive surveys to minimise bias and achieve reasonable precision, which increases the cost of sampling and limits the survey area (e.g. Hankin & Reeves 1988). Classical random-sampling techniques can also be used, such as simple or stratified-random sampling, but lack predictive ability and precision for small areas. In addition, it may not be feasible to truly randomise sample placement due to a lack of access (e.g. no roads, steep canyons or uncooperative land owners). Even if a randomised design can be used, there may be better estimators, such as block prediction, or universal block kriging on stream networks (Ver Hoef et al. 2006), to scale up from an on-to an over-network analysis.

Future research

The full suite of spatial statistical models described thus far, including mixed models, spatial GLMs and block kriging on stream networks, may be fit using the SSN package (Ver Hoef et al. in review) for R statistical software (R Development Core Team 2010). However, other methods would also be useful and more research is needed to make them spatially explicit on stream networks. For example, incomplete-detection occupancy models (MacKenzie et al. 2002) provide a way to estimate occupancy rates from binary data, while also accounting for the probability of detection. Analysis of extremes for water quality often involves converting continuous data to a binary response at the threshold level (Clement & Thas 2009) or using generalised extreme-value models that depend on a distribution (Towler et al. 2010). Analyses at broad scales may lead to computational issues with large sample sizes (e.g. in-stream sensor networks, Porter et al. 2012); therefore, spatial statistical methods for large data sets, such as fixed-rank kriging (Cressie & Johannesson 2008), could be adapted for stream networks. Current methods only account for spatial dependency in stream data, but there is clearly a temporal-dynamic structure that should be incorporated simultaneously using spatio-temporal analytic methods (Cressie & Wikle 2011). Finally, inferences for spatial data are substantially affected by the spatial configuration of survey sites on the network (Zimmerman 2006). Many survey designs seek spatial balance over the geographical range (Stevens & Jensen 2007), but proximity and connectivity in stream networks are functionally different than in terrestrial systems, requiring research on stream-specific survey designs to optimise objectives.

INTEGRATING NETWORK ANALYSES

Previous network analyses in stream ecosystems have focused on the influence of either network-explicit variables (e.g. physical network structure and flow direction) or network-implicit variables (e.g. continuous, hierarchical and spatio-temporally heterogeneous in-stream habitat quality). However, the tendency to focus solely on either spatial structure or function is not unique to stream ecology and there has been an effort to integrate disparate perspectives to gain a more holistic understanding of the study system (Paola et al. 2006; Rodríguez-Iturbe et al. 2009). For example, Massol et al. (2011) noted that much of the research on spatial food webs in ecology has been developed independently from research on ecosystem dynamics, and proposed a ‘metasystem’ framework for bringing concepts from landscape ecosystem and food-web metacommunity ecology together. In another study, Jabot & Bascompte

© 2013 Blackwell Publishing Ltd/CNRS and Commonwealth of Australia
(2012) demonstrated how a metacommunity model, which focuses on spatial processes in a single trophic group, could be integrated with a network model that considers multiple trophic groups at a single location, to obtain a better understanding of how network structure affects metacommunity dynamics. The combined influence of spatial structure and ecological interactions on in- or near-stream processes is also of interest in stream ecology (Angermeier et al. 2002; Power & Dietrich 2002; Fisher et al. 2004; Falke & Fausch 2010), but the dual coordinate system makes explicitly accounting for their combined influence more complicated than in other spatially structured ecological networks.

The taxonomy of network analyses provides a framework to aid in the integration of different models preferred by ecological subdisciplines, such as community, landscape or population ecology. As a first step, studies that include both on- and about-network analyses, potentially within a single statistical framework, are relatively straightforward to implement and will provide a means of investigating the influence of in-stream habitat availability and physical network structure on the distribution of organisms. For example, a spatial statistical model could be used to generate semi-continuous predictions of in-stream habitat based on physical sub-network structure or other remotely derived covariates, such as climate, topography or land cover. Then, about-network metrics could be used to relate the configuration and connectivity of predicted habitat patches to species distributions using a graph-theoretic-based model. Different types of on-network analyses may also be combined to address more complex questions. Goldberg et al. (2010) developed a matrix population model to investigate the effects of dendritic network structure and within-network habitat patches on species distributions. Habitat patch characteristics were assigned based on distance upstream from the stream outlet; however, a spatial statistical model could be used to estimate more realistic segment-scale habitat characteristics (e.g. temperature or substrate type) thought to influence organismal dispersal, survival or reproduction, based on catchment land-management practices. These examples demonstrate how spatial statistical methods can be used to predict biologically relevant information at an intermediate scale (1–100 km), which can then be combined with about- or on-network analyses, thus, accounting for the interplay between network structure, within-network habitat characteristics or processes, and/or the characteristics of the 2-D landscape the network resides within.

The ability to integrate various network analyses in DENs also opens the door to a suite of previously intractable research questions. For example, cutthroat trout is a species of concern in the northwestern United States, where their distribution is relatively limited compared to historical distributions (Young 1995). Evidence suggests that these declines may be due to habitat degradation (Harig et al. 2000), isolation of populations (Haak et al. 2010) and competition with invasive species such as Salvelinus fontinalis (brook trout; Fausch 2008). It remains unclear, however, which factors are responsible for most of the decline, and whether their influence is spatially heterogeneous or varies depending on scale. An integrated approach using both on- and about-network methods provides a way to investigate each component’s respective contribution, as well as, its influence on other ecological dynamics. This information could then be used to develop a network-explicit reserve design (Urban et al. 2009) or to undertake a risk or cost–benefit analysis to identify areas with the greatest conservation or restoration potential (Urban & Keitt 2001). Other taxa of concern also share habitat with cutthroat trout (e.g. Dicamptodon sp.) and a spatially structured network model (Jabot & Bascompte 2012) would provide a way to move from single species conservation to a multispecies approach (Urban et al. 2009). In addition, air and stream water temperatures are expected to increase in the future, causing shifts in fish distributions (Hari et al. 2006), which adds to the challenge of spatially explicit, conservation prioritisation efforts. One solution would be to study potential metacommunity dynamics using time-ordered networks (Blonder et al. 2012) under a series of future climate and thermal habitat scenarios to account for new species interactions, as well as, changes in habitat quality and network structure resulting from lower stream flows. Finally, ecological processes (e.g. movement and dispersal) are often facilitated or impeded by non-natural transport mechanisms such as human reintroductions and the intercatchment transfer of water, nutrients or organisms (Fullerton et al. 2011; Grant et al. 2012). Thus, there is a need to understand the effects of these human-imposed networks on physically constrained networks, such as DENs. The taxonomy of network analyses provides a conceptual framework to select and combine complimentary analytical methods to understand complex ecological systems composed of both physicochemical and biological processes that operate across multiple scales in the dual coordinate system.

OTHER DENs AND SPATIALLY STRUCTURED ECOLOGICAL NETWORKS

Although our primary focus has been on stream ecology (Box 1), the conceptual taxonomy of network analyses is relevant for any dendritic ecological network, which exists within a dual coordinate system. The same concepts and models could be used to select and combine analyses of preferential, but not exclusive, use and migration of terrestrial or semi-aquatic species along riparian corridors (Naiman & Decamps 1997; Carranza et al. 2012); the effects of both pollution and the distribution of refugia on fauna in cave networks (Wood et al. 2008); or studies investigating the effects of plant architecture on foraging intensity or pest infestation patterns in dendrically structured plants (Legrand & Barbosa 2003; Sylvaine et al. 2012). For example, spatial statistical methods have previously been used to explore pest and nutrient distributions in trees (Habib et al. 1991; Audergon et al. 1993) and the tail-down covariance model would be particularly suited for these problems. In addition, a model based on a covariance mixture (i.e. tail-down, tail-up and Euclidean) would allow complex patterns of spatial autocorrelation associated with proximity to the plant’s main stem or differences in light exposure to be accounted for, in addition to those associated with network structure.

The taxonomy of network analyses would also be relevant in other ecological settings where processes are not limited to Euclidean space, but rather follow pathways that are constrained by the physical environment. For example, the effects of ocean currents on larval dispersal (Hidalgo et al. 2011); patterns of dissolved oxygen in estuaries (Rathbun 1998); animal movement along habitat corridors (Castellon & Sieving 2006); and plant (Spooner et al. 2004) and animal dispersal along road networks (Brock & Kelt 2004). Note that, the specific models and examples provided here may not be suitable in every situation (e.g. a spatial statistical model for stream networks cannot be applied to a non-dendritic road network). Nevertheless, other models found within the same families of models, such as
Spatial statistical methods, graph-theoretic approaches and metapopulation models, can be combined in a myriad of ways to systematically account for the interplay between network structure, within-network characteristics or processes, and the characteristics of the 2-D landscape the network resides within.

**SUMMARY**

Dendritic ecological networks, such as streams ecosystems, are a unique form of spatially structured ecological network that play a vital role in ecology (Paola et al. 2006). Analytical methods that explicitly account for fundamental characteristics, such as network structure and connectivity within the dual coordinate system, are needed to better understand the processes governing physicochemical and biological properties within DENs and the surrounding environment. This is especially true in streams, where longitudinal connectivity strongly influences in-stream processes and lateral connectivity blurs the boundary between the aquatic and terrestrial environment (Fisher et al. 2004). If the fundamental characteristics of DENs are not accounted for in the analyses, it can lead to poor scientific inference, and in turn, poor management decisions. Therefore, the ability to account for these fundamental characteristics within an analytical framework is especially important for bridging the gap between research pertaining to fine-scale processes and broad-scale management decisions (Fausch et al. 2002).

We proposed a unifying taxonomy of analyses non-, about-, on-, over- and across- networks to help researchers (1) understand the differences between the processes of interest and the analytical methods available, (2) select the most suitable method for their study and (3) integrate network analyses to acquire a more coherent system-wide understanding. We then considered on-network analysis in more detail because it has received the least attention and there have been recent novel developments, while the taxonomy of network analyses provided the context for such development. There are undoubtedly other analytical methods that were not discussed, which account for network characteristics to varying degrees, such as process-based models used to predict sediment movement in streams (Gassman et al. 2007) or network-explicit spatial optimisation methods used to prioritise conservation efforts (Hermoso et al. 2011). Nevertheless, the taxonomic framework allows ecologists to place other methods, including those yet to be developed, within the broader context of potential network analyses.

Our hope is that this taxonomic framework will help ecologists quantitatively embrace the spatial complexity of DENs, as well as, explore and test the evolving ecological conceptualisation of DENs as a unique form of spatially structured ecological network. Interdisciplinary collaboration between ecologists and statisticians made it possible to develop this framework. Further cross-disciplinary collaboration is needed to ensure that new statistical methods represent the fundamental characteristics of spatially structured ecological networks, so that ecologists can push the boundaries of their science, while also providing managers with tools for solving real-world problems.

**ACKNOWLEDGEMENTS**

This study was conducted as part of the Spatial Statistics for Streams Working Group supported by the National Center for Ecological Analysis and Synthesis, a Center funded by the National Science Foundation (Grant #EF-0553768), the University of California, Santa Barbara and the State of California. We thank G.H. Reeves and K.M. Burnett at the US Forest Service Pacific Northwest Research Station for providing the rainbow trout abundance data used to develop the Torgegram in Fig. 7b. We also thank Frederike Kroon, Nick Bond, Dan Pagendam and three anonymous reviewers for constructive comments on previous versions of this manuscript. Note that, any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

**AUTHORSHIP**

This manuscript was developed as part of the National Center for Ecological Analysis and Synthesis, Spatial Statistics for Streams working group and all authors provided major contributions to its conceptual development. The first three authors contributed portions of text, synthesised contributions from co-authors, developed examples and provided general editing. The remaining authors wrote portions of text, developed examples, contributed to editing and/or provided feedback on previous versions of the manuscript; these authors are listed in alphabetical order.

**REFERENCES**


Editor, Bernd Blasius

Manuscript received 19 September 2012
First decision made 31 October 2012
Manuscript accepted 14 January 2013

© 2013 Blackwell Publishing Ltd/CNRS and Commonwealth of Australia