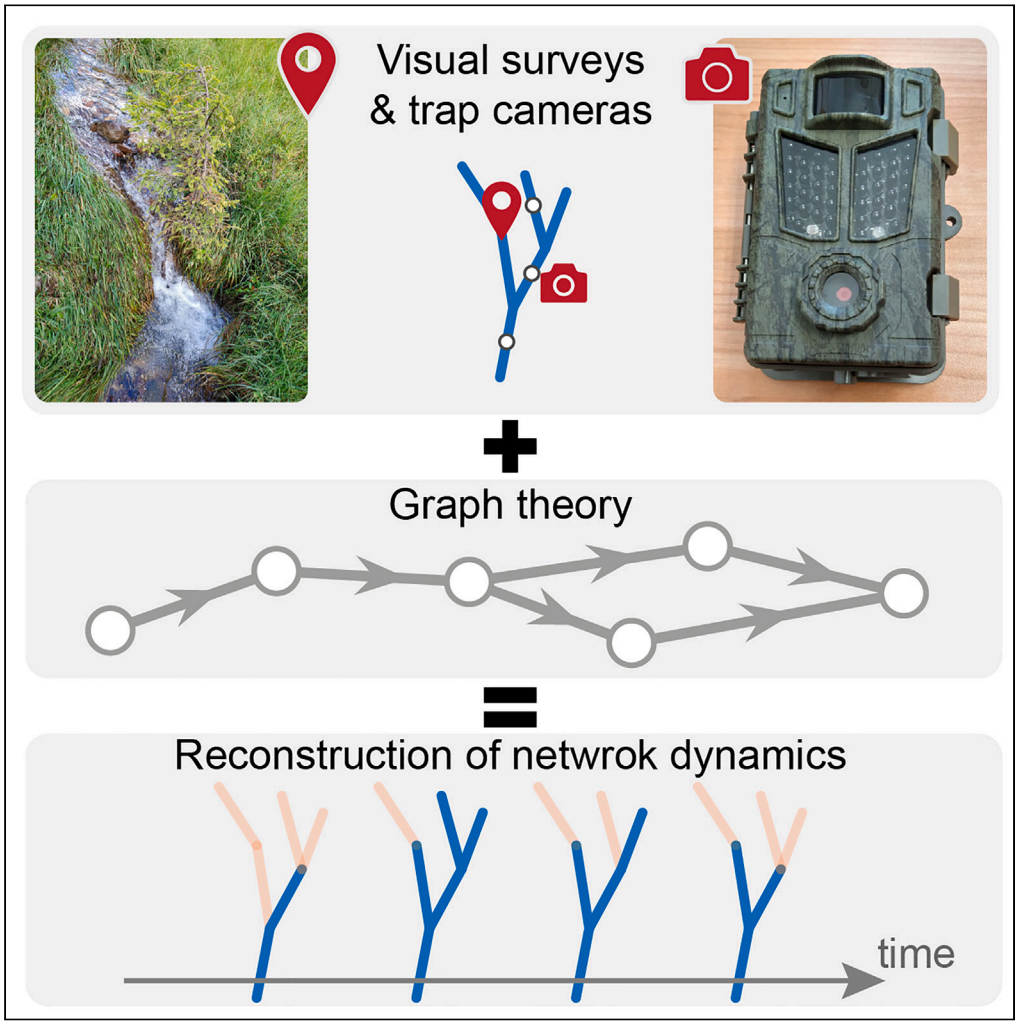
iScience

CellPress
OPEN ACCESS

Article

Integrating spatially-and temporally-heterogeneous data on river network dynamics using graph theory



Nicola Durighetto, Simone Noto, Flavia Tauro, Salvatore Grimaldi, Gianluca Botter

nicola.durighetto@unipd.it

Highlights

Non-perennial river networks expand and retract in a hierarchical manner

Graph theory can be used to reconstruct the activation hierarchy of stream reaches

The framework greatly reduces the data required to monitor the whole active network

Data with different spatial and temporal resolution can be easily integrated

Durighetto et al., iScience 26, 107417
August 18, 2023
https://doi.org/10.1016/j.isci.2023.107417

iScience

CellPress
OPEN ACCESS

Article

Integrating spatially-and temporally-heterogeneous data on river network dynamics using graph theory

Nicola Durighetto,[1,3,*] Simone Noto,[1] Flavia Tauro,[2] Salvatore Grimaldi,[2] and Gianluca Botter[1]

SUMMARY

The study of non-perennial streams requires extensive experimental data on the temporal evolution of surface flow presence across different nodes of channel networks. However, the consistency and homogeneity of available datasets is threatened by the empirical burden required to map stream network expansions and contractions. Here, we developed a data-driven, graph-theory framework aimed at representing the hierarchical structuring of channel network dynamics (i.e., the order of node activation/deactivation during network expansion/retraction) through a directed acyclic graph. The method enables the estimation of the configuration of the active portion of the network based on a limited number of observed nodes, and can be utilized to combine datasets with different temporal resolutions and spatial coverage. A proof-of-concept application to a seasonally-dry catchment in central Italy demonstrated the ability of the approach to reduce the empirical effort required for monitoring network dynamics and efficiently extrapolate experimental observations in space and time.

INTRODUCTION

Recent estimates suggest that over half of the global river network periodically ceases to flow,[1,2] owing to event-based and seasonal expansions and retractions of the wet channels. Non-perennial streams are ubiquitous features of landscapes, as they can be found not only in arid regions[3] but also in many humid headwaters.[4,5] The ceaseless switching between lentic, lotic, and terrestrial environmental conditions such as encountered in temporary streams affects many hydrological, biochemical, and ecological processes[6,7] that bear a profound influence on water quantity, quality, and ecosystem functioning of non-perennial river reaches and downstream water bodies.[8,9] Therefore, understanding the inner functioning of stream network dynamics proves to be crucial for describing the hydrological response of rivers and has important implications for many related research areas (e.g., stream ecology and biogeochemistry).

The availability of empirical data on surface flow presence along the network is a fundamental prerequisite for enhancing our understanding of the dynamics of temporary streams. However, channel network dynamics are usually very heterogeneous even at small spatial and temporal scales. Therefore, the empirical burden associated to the field mapping of the active portion of a branching river network remains one of the major bottlenecks in the study of non-perennial streams.[10]

In the literature of non-perennial streams, the most common monitoring technique consists in visual field surveys of the active network performed on different dates, so as to get a set of active network maps that refers to different hydrologic conditions.[11–24] Even though this method generates extremely accurate data (with associated errors usually smaller than a few meters), on-the-ground mapping proved to be highly time-consuming, owing to the difficulties in surveying the entire riparian corridor.[10] As such, visual inspection has usually been used to monitor monthly to seasonal dynamics in relatively small catchments–with a few notable exceptions in which the spatial scale and/or the temporal frequency were increased.[25–30]

Field mapping can be effectively complemented by hi-tech instruments. For instance, remote sensing (e.g., satellite and drone imagery) has been often used for the monitoring of temporary streams over larger areas.[31–35] While these applications look promising, they present significant limitations and generate several no-data. Field-deployed sensors have also been used several times to monitor the presence of surface water in a predetermined set of nodes in the network. The most common type of sensors consists in probes designed to measure either temperature[36–38] or electrical resistance,[39–50] although purpose-built

[1]Department of Civil, Environmental and Architectural Engineering, University of Padua, 35131 Padua (Padua), Italy

[2]Department of Innovation in Biological, Agro-food and Forest Systems, University of Tuscia, 01100 Viterbo (Viterbo), Italy

[3]Lead contact

*Correspondence: nicola.durighetto@unipd.it

https://doi.org/10.1016/j.isci.2023.107417

sensors[51] and trap cameras[52–55] have also been successfully employed to reconstruct the presence of surface flow at a pre-selected number of locations along the network. In the light of the fact that each monitoring system is characterized by its characteristic temporal resolution and spatial coverage, the creation of empirical datasets that are able to describe both seasonal and event-based network dynamics in catchments bigger than a few hectares might greatly benefit from the combined use of different techniques.[53] As an example, high-frequency data from a number of field-deployed sensors may be efficiently integrated with visual field surveys in order to better capture the full spatial complexity of network dynamics in specific areas of a catchment. While the integration of diverse techniques may allow both high spatial resolutions and fine temporal frequencies to be achieved, it also generates heterogeneous datasets in which different portions of the network may be observed with a different frequency and on different dates. This heterogeneity in the data complicates the direct use of empirical observations in many practical applications. As of now, in fact, a robust and organic procedure to merge different data about network dynamics into a single, coherent dataset is still missing.

In temporary streams, the wetting and drying of network nodes (and of the corresponding homogeneous reaches associated to these nodes, see Figure 1A) has been shown to follow a strict hierarchical structuring.[56–58] According to this principle, during network expansion, nodes are always activated following the same, fixed, sequence, and then they are deactivated in the reverse order when the flowing network retracts. This sequence defines a unique hierarchical ordering of all the network nodes: while the number of active nodes varies with time in response to climate variations, at any given time all the active nodes belong to the first part of the hierarchy, whereas all the dry nodes are located in the last part of the hierarchy. It must be noted that the order of the nodes in the hierarchy may not be related to their physical location along the network. In this way, even though nodes are always activated sequentially from the first to the last in the hierarchy, the corresponding active network may present a number of disconnections.[57,58]

This hierarchical mechanism is aligned with currently available conceptual models for network dynamics, which relate the generation of surface flow in a given node to the local imbalance between catchment-scale inflow and the longitudinal subsurface capacity.[5,16,58] In this context, knowing the hierarchy that links all the nodes of the river network can be very useful to link the total active length with the spatial configuration of the active network, thereby facilitating the modeling of stream network dynamics in many settings.[57,59,60] If the local persistency of each node in the network (i.e., the fraction of time for which each node experiences flowing water) is known, it follows naturally that the activation order indicated by the hierarchy goes from the most persistent node (i.e., the first to wet up and last to dry down) to the least persistent one (i.e., the last one to wet up and first to dry down).[56] However, when only few surveys are available, or if different subsets of nodes are observed with different frequencies and in different time periods, the relative number of observations in which a node was active might not be a good proxy of the local persistency. Consequently, evaluating the hierarchy among nodes only based on the "apparent" persistency as emerging from heterogeneous datasets might lead to strong biases, and a more robust procedure needs to be identified to define the hierarchy among the nodes in the network.

Graph theory has proven to be a valuable tool in many research fields, including water science, as it can be used to quantitatively analyze the topology and connectivity of fluvial systems. For example, graph theory has been applied to study morphodynamics,[61] sediment connectivity,[62] hillslope hydrology,[63] riverine deltas,[64] wet landscapes,[65] and surface[66,67] or subsurface[68] hydrological connectivity. While graph theory has also been employed for the description of non-perennial streams,[56] it has never been used to characterize the hierarchical behavior of non-perennial rivers.

On this basis, this paper addresses the following specific research questions:

(1) can graph theory be used to reconstruct and visualize the hierarchy of the nodes in a river network, in a way that allows heterogeneous and incomplete observations to be merged coherently?

(2) how can the hierarchical structuring of network dynamics be exploited to predict the spatial configuration of the active network on the basis of sporadic observations pertaining to limited subsets of nodes?

To answer these questions, in this paper we develop a robust mathematical framework based on the graph theory, which is designed to reconstruct the hierarchical structure of the nodes starting from spatially and

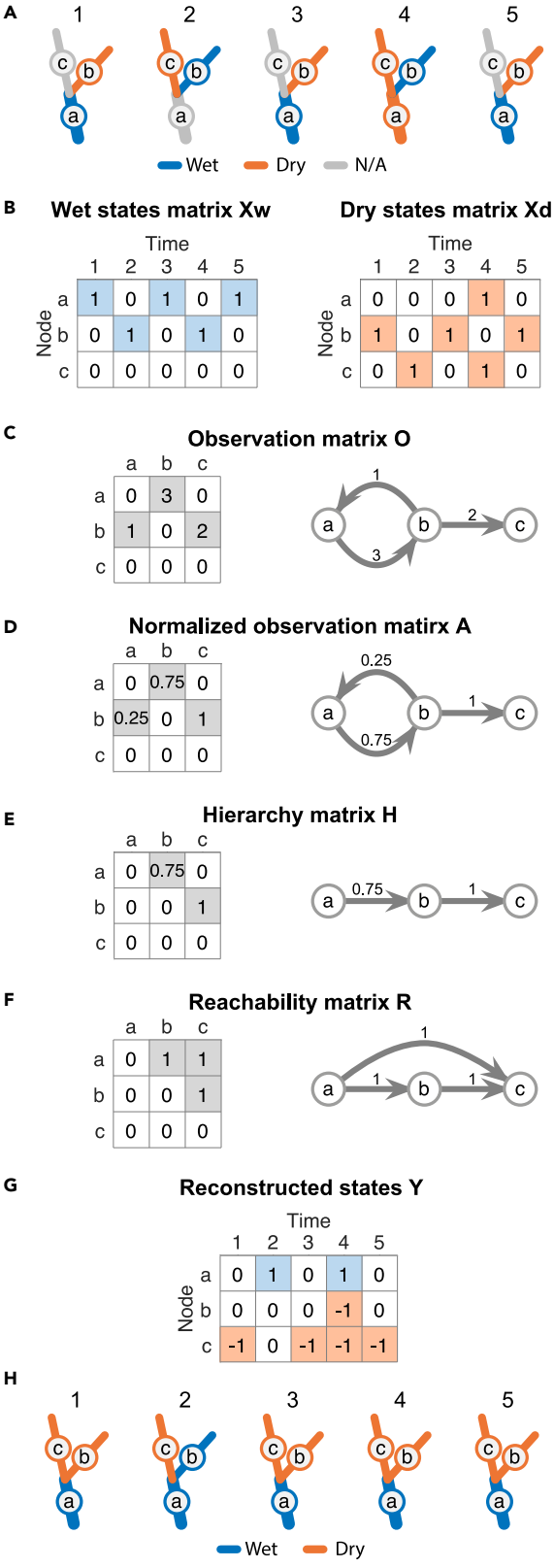


Figure 1. Schematic visualization of inner functioning of the proposed framework, as applied on an example three-nodes network ($n = 3$)

(A–H) Panel (A) reports maps of the available observations, relative to 5 distinct time instants ($n_T = 5$). Note that the observed network is not always hierarchical, as reach b in map A-4 is wet even though reach a is dry. The available data are summarized in matrix form in panel (B). Panels (C)–(E) display the steps for the construction of the hierarchy graph and reconstruction of node states, as specified by Equations 1, 2, 3, and 4. Note that loops are created in panels (C) and (D) due to the observed map A-4, which contradicts the reconstructed hierarchy. Panel (F) reports the reconstructed hierarchy graph, while (G) shows the matrix of reconstructed states. Panel (H) shows maps of the active network obtained combining the reconstructed states with the starting available observations. Note that, in all graphs, each node is associated to a uniform stream reach embedding the reference node (i.e., the status of the node represents the status of the associated reach). Likewise, each gray link describes statistical connections between nodes (rather than describing the stream reaches themselves).

temporally heterogeneous datasets. The framework is then applied to a study catchment in central Italy to demonstrate its potential.

RESULTS

A graph theory framework for representing and reconstructing hierarchical stream network dynamics

In this subsection, we present an analytical formulation to represent and reconstruct the hierarchical mechanism of activation/deactivation of the stream network. Though being to some extent methodological, we believe this part of the paper contains some key general results of our analysis, and thus well fits the results section of the manuscript.

Throughout the whole paper, we define a node as a unique entity that allows a discrete description of a river network. Specifically, a node represents the hydrological conditions observed in a given stream segment with uniform characteristics and behavior (Figure 1A). This definition enables us to decompose the network into a number of nodes and describe it via suitable graphs in which the edges represent statistical connections between pair of nodes (which could be far apart in a map). This implies that the statistical connections do not necessarily describe the physical connectivity of the flowing network in a map (e.g., neighboring nodes may not behave in a synchronous manner). Directed edges will in fact be used only to describe the order of activation of network nodes, following the hierarchical principle, as reconstructed from the available empirical data. The approach reported here was developed in order to deal with the case of a sparse dataset in which sporadic observations about the hydrological conditions across different nodes of the network are available. The empirical observations therefore consist in a non-homogeneous temporal sequence of states (wet/dry) for the node of the network, as in the example reported in Figure 1A. These data are summarized in two matrices, $\mathbf{X_w}$ and $\mathbf{X_d}$, which contain, respectively, the time-series of wet and dry states experienced by each node (Figure 1B). Both matrices have dimension $n \times n_T$, where n is the total number of nodes in the network and n_T is the number of times during which at least one node in the network was observed. As such, the element (i, t) of $\mathbf{X_w}$ is set to 1 if node i was observed as wet at time t, and 0 otherwise (i.e., if the node i was dry or was not observed at time t). Similarly, the element (i, t) of $\mathbf{X_d}$ is 1 if node i was observed as dry at time t, and 0 otherwise (i.e., if the node i was wet or not observed at time t). The overall observation matrix, $\mathbf{O}$, can be computed from $\mathbf{X_d}$ and $\mathbf{X_w}$ as:

$$\mathbf{O} = \mathbf{X_w} \cdot \mathbf{X_d}^{\mathsf{T}} \qquad \text{(Equation 1)}$$

where $\cdot$ denotes matrix multiplication, and $^{\mathsf{T}}$ denotes transposition. $\mathbf{O}$ is an $n \times n$ matrix, in which the element (i, j) contains the number of observations in which node i was wet and node j was dry. This matrix defines a directed graph, in which an edge from node i to j exists if and only if there is at least one date in which node i was wet and node j was dry (Figure 1C). The weight associated to the edge is the number of times nodes i and j were actually observed to be respectively wet and dry, as quantified by the value of $\mathbf{O}(i, j)$. In general, the graph associated to $\mathbf{O}$ may contain cycles if different observations (i.e., observations made in different dates) suggest the existence of contrasting hierarchies between at least two nodes. This can be due to observation errors, or to errors in the hierarchical model, in turn induced by changes in local morphology after intense events, heterogeneous rainfall fields, or local storage variations—see ref.[57]

In the second step of the proposed procedure, the observation matrix is normalized, so that the element (i, j) of $\mathbf{A}$ contains the fraction of observations indicating that node i is more persistent than node j (i.e., number of observations in which node i was wet and node j was dry divided by all observations in which nodes i and j had different states, as in the example of Figure 1D):

$$\mathbf{A} = \frac{\mathbf{O}}{\mathbf{O}+\mathbf{O}^{\mathsf{T}}} \qquad \text{(Equation 2)}$$

Third, the graph is transformed to a directed acyclic graph (DAG) by breaking all the cycles (Figure 1E). A new matrix representing the DAG of the hierarchy, $\mathbf{H}$, is therefore created as follows:

$$\mathbf{H} = f(\mathbf{A}), \qquad \text{(Equation 3)}$$

where $f(\mathbf{A})$ is a function used to break the cycles of matrix $\mathbf{A}$ while preserving its logical structure.[69] This is done by removing a specific subset of edges from $\mathbf{A}$. In particular, edge removal is performed in a way that minimizes the sum of the weights associated to the removed edges (i.e., the hierarchy is defined in a way that minimizes the relative number of observations that are not in line with the inferred hierarchy, which

correspond to the removed edges). The element (i, j) of $\mathbf{H}$ is thus the fraction of available observations indicating that node i is more persistent than node j. This matrix defines the DAG describing the hierarchy among all the nodes in the network. In this DAG, a directed edge between nodes i and j exists if most of the observations indicates that node i is located before j in the hierarchy. Therefore, the DAG can be used to visualize and describe the order of node activation (deactivation) during network expansion (contraction). It should be noted that H always includes all the nodes in the network. However, if not enough information is available through empirical observations, one or more nodes may be completely isolated from the rest of the DAG, suggesting that the ranking of the isolated node within the activation order cannot be identified with the available data.

Crucially, matrix $\mathbf{H}$ and the corresponding DAG can be used to estimate the status of non-observed nodes starting from a limited number of observations (as in the case of a survey in which not all the network is monitored): if a given node is observed as wet, all the previous nodes in the hierarchy must be wet too, and if a given node is dry, all the subsequent nodes in the hierarchy must also be dry. Since node ranking within the hierarchy is only related to the wetting/drying behavior and might not be related to the physical location of the nodes along the network, the presented framework can be applied to heterogeneous streams characterized by complex spatial patterns of wetting and drying, including dynamically fragmented stream networks. Mathematically, to estimate the status of non-observed nodes we first need to define the reachability matrix $\mathbf{R}$, which represents the transitive closure of $\mathbf{H}$. Therefore, $\mathbf{R}$ represents a new DAG, which is based on $\mathbf{H}$, in which an edge between nodes i and j is present if an indirect path between i and j exists in $\mathbf{H}$ (Figure 1F). Note that the weight of the edges in $\mathbf{R}$ is always unitary.

Then, given the wet- and dry-node observations represented in the matrices $\mathbf{X_w}$ and $\mathbf{X_d}$, the status of the nodes that were not observed in any date in which at least one node was monitored can be estimated as:

$$\mathbf{Y} = \mathbf{R} \cdot \mathbf{X_w} - \mathbf{R}^T \cdot \mathbf{X_d}, \quad \text{(Equation 4)}$$

where $\mathbf{Y}$ is a $n \times n_T$ matrix in which the element (i, t) expresses the net number of nodes observed at time t that are being used through the hierarchy to predict the status of node i (i.e., the number of observed wet nodes that are located after i in the hierarchy, minus the number of observed dry nodes that precede i in the hierarchy). Therefore, the element (i, t) is positive if node i is estimated to be wet at time t, negative if it is estimated to be dry, or null if the available information is not sufficient to reconstruct the status of such node (Figure 1G). The latter condition applies to cases in which the reconstructed hierarchy is not able to connect node i to a wet observed node in the subsequent part of the hierarchical DAG, or to a dry observed node in the previous part of the DAG. This definition of $\mathbf{Y}$ allows us to summarize all the estimated node states in a single matrix, which could be easily compared to the analogous compound data matrix $\mathbf{X} = \mathbf{X_w} - \mathbf{X_d}$ that summarizes the observed states of all the nodes in the network.

A proof-of-concept application to the Montecalvello catchment

To elucidate the potential of the theory developed above, the framework was applied to a representative case study located in central Italy, the "Fosso di Montecalvello" creek.[54,55] The flowing portion of the stream network in the Montecalvello catchment was monitored from October 2019 to August 2022. A total of 40 visual surveys were carried out on a biweekly to monthly base to monitor $n = 58$ nodes (or a subset of them). On top of that, a total of 21 trap cameras were installed in selected locations along the network, to enhance the temporal resolution of the observation of a subset of nodes to a daily timescale (Figure 2). Combining the datasets obtained from visual surveys and trap cameras, network dynamics on the Montecalvello catchment were monitored for a total of $n_T = 466$ days. These included 16 days in which only visual surveys were carried out, 426 days in which only trap cameras were used, and 25 days in which both methodologies were jointly used, as reported in Figure 3. The stream network was found to be highly dynamic, with a flowing length ranging from 1.11 km to 5.28 km. The temporal dynamics of the active network followed a clear seasonal pattern, with the shortest-flowing network that was observed during the summer.

The hierarchy among the nodes of the network was reconstructed by means of a DAG, using the procedure detailed in the previous section. The idea underlying the procedure is that a new edge is added to the graph whenever a pair of nodes with different states is observed. The new link is directed from the node observed as wet to the one observed as dry. Such a link indicates that, in the hierarchy, the wet node must precede the dry node. The hierarchy graph for the Montecalvello network is reported in Figure 4, while the steps required to build such graph are reported in Figure S1, using a subset of 6 nodes for the

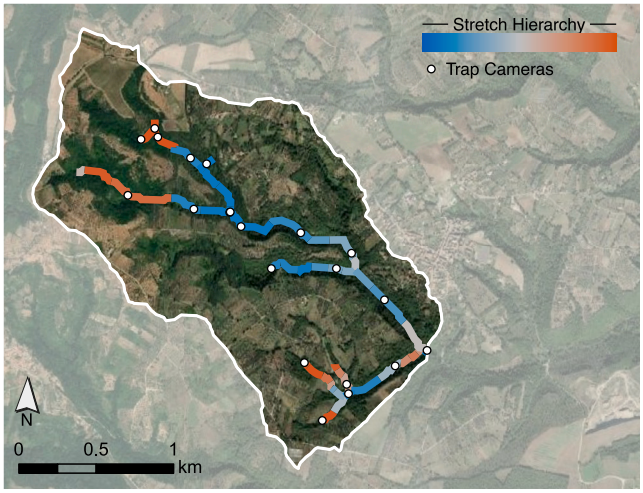


Figure 2. Map of the Montecalvello catchment

The blue-orange color scheme of the network (see figure legend) reflects the color gradient along the hierarchy, as reported in Figure 4. Map generated with ESRI ArcGIS Pro.

sake of clarity. Nonetheless, the procedure can be applied regardless of the number of nodes involved. Given that the edges among nodes in the graph mean to describe statistical connections rather than the hydraulic connectivity of nodes along the network, the hierarchy graph H cannot be reconstructed from a single map, but requires multiple empirical observations. As the amount of information on the hydrological conditions in the nodes of the network increased significantly in 2021 and 2022 (Figure 3), the structure of the hierarchy became clearer as new empirical data were available. At the end of the summer of 2022, most of the hierarchical structure was unveiled, particularly in its last part (which comprises the less persistent nodes, color-coded in light orange in Figure 4). The order of activation of the most persistent nodes (color-coded in dark blue), instead, has not been fully observed, thus generating a graph with a more complicated structure in the first part of the hierarchy, which in fact does not resemble a chain. Nonetheless, the available data provided a good description of the dynamics of the stream network, and the accuracy of the reconstructed hierarchy was about 99.6%, suggesting that less than 0.4% of all the empirical observations were not in accordance with the identified hierarchy. Such observations are attributed to the

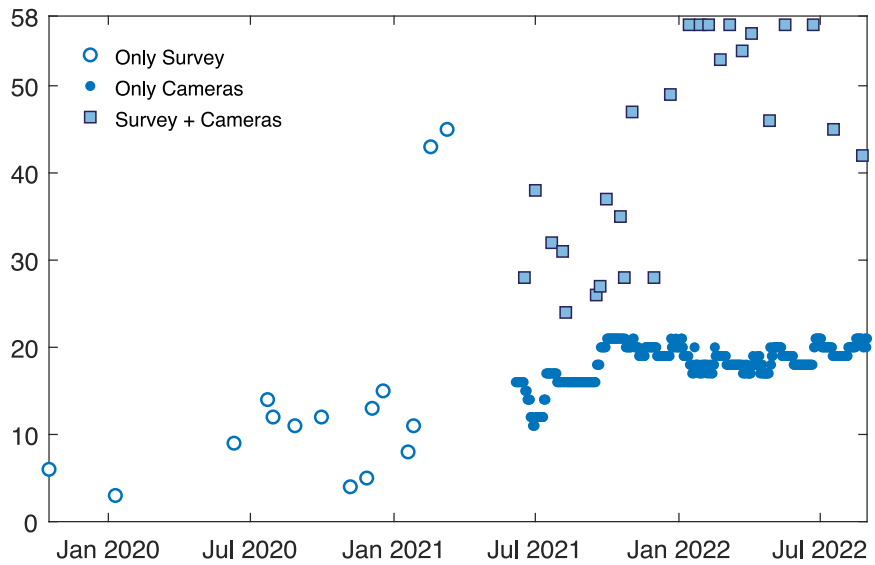


Figure 3. Number of observed nodes for each date, as a function of the monitoring method employed

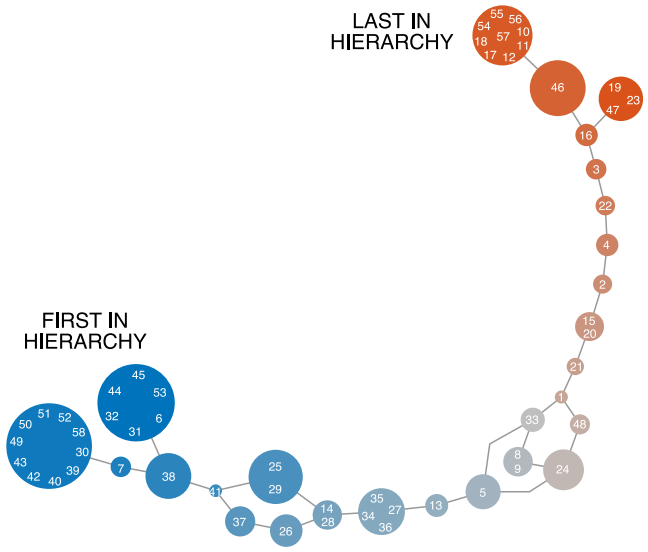


Figure 4. Graph representing the activation hierarchy of the nodes within the stream network

Network nodes with similar properties have been merged together to simplify visualization. Node size is proportional to the associated length in the river network. The color gradient from blue to orange indicates the direction of the hierarchy. The numbers identify each single node. Edges are used to characterize the statistical properties of node activation, and may not reflect the connectivity between nodes along the river network, which are reported in Figure 2.

effect of local hydrological processes observed in certain parts of the network that possibly generate a non-univocal activation/deactivation order of the nodes. This includes any unpredictable change in the pattern of local water input in some nodes during a few rain events and temporal changes in the river bed permeability. Non-hierarchical behaviors could be also generated by the unavoidable uncertainty in the collected data (e.g., image interpretation and processing).

The hierarchy shown in Figure 4 was to estimate the status of the nodes whenever node observations were not available (no-data). These no-data here refer mainly to nodes that were not observed continuously by the trap cameras, or periods during which a trap camera was not active for maintenance purposes. The proposed application represents a good example of how different monitoring techniques can be integrated in a robust manner to generate a comprehensive dataset that minimizes the amount of missing information. The estimation of the nodes' states was done by applying Equation 4, which is based on the hierarchical structuring of network dynamics: if a given node in the hierarchy is observed as wet, all the previous DAG nodes must be wet too, and if a given node is dry, all the subsequent nodes in the DAG must be dry. The hierarchical method employed for estimating the status of unmonitored network nodes is shown in detail in Figure 5, in which 4 different configurations of the active network are reported (from very dry to almost completely wet). For each date, the hierarchical graph and the observations were used to determine the position of the transition between the wet and dry nodes in the DAG. Afterward, the status of the nodes which were not observed was reconstructed based on their position in the chain, as per Equation 4. The spatial distribution of the active and dry nodes was also reported in the right panels of Figure 5, which indicate that the observed flowing networks emerging in hierarchical rivers can be highly fragmented. In some cases, the status of some nodes could not be reconstructed (gray nodes in Figure 5), because the empirical information available on a given date was insufficient (the precise position of the transition between wet and dry nodes in the DAG was not known).

The same procedure was repeated for all the dates in which at least one node was monitored, and the results are summarized in Figure 6, in which the observed (or estimated) status of each node is shown as a function of time using a suitable color-coding. In line with Figure 5, the applied procedure did not allow the estimation of the status of the nodes that, in the hierarchy, fell between the last observed wet node and the first observed dry node. Therefore, a limited amount of no-data values was indicated in gray in Figure 6. In spite of this limitation, in many cases the observation of a very limited number of nodes was sufficient to estimate the status of the whole network based on the hierarchical DAG. On average, the empirical monitoring of a mean active network length

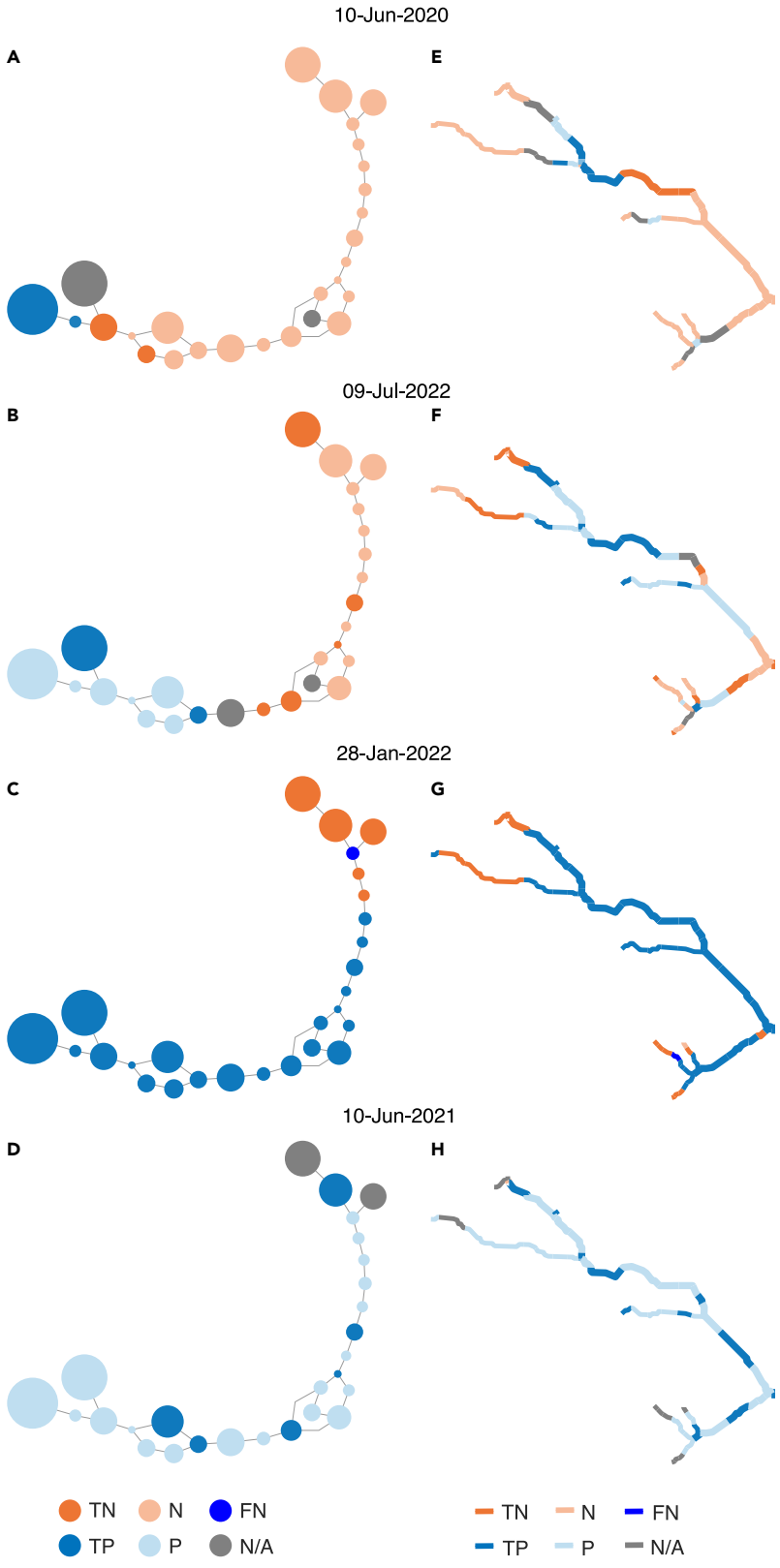

10-Jun-2020
A
E
09-Jul-2022
B
F
28-Jan-2022
C
G
10-Jun-2021
D
H
TN
N
FN
TP
P
N/A
TN
N
FN
TP
P
N/A

Figure 5. Application of the hierarchical framework on the Montecalvello catchment

Hierarchy graphs of the stream for 4 selected dates (panels A, B, C, and D) and the corresponding active networks (panels E, F, G, and H). True positives (TP) and true negatives (TN) refer to nodes correctly identified by the model as wet or dry, respectively. False negatives (FN), instead, highlight model errors. Positives (P) and negatives (N) show node states estimated by the model, for which there are no corresponding empirical observations. N/A identifies no-data. False positives (FP) are not present in the selected dates.

of 2.24 km allowed the estimation of the status of 5.56 km of river network during each survey. Figure 6 also demonstrates the high accuracy of the proposed method: in our application only 0.37% of the observed network length was misclassified as wet (False Positives = 0.24%) or dry (False Negatives = 0.13%), with a total accuracy of the hierarchical model of about 99.6% (see STAR Methods).

DISCUSSION

In this study, we developed a new graph theory framework that enables the identification of the hierarchical structuring of the nodes in a river network using empirical data on surface water presence characterized by heterogeneous temporal frequencies and spatial resolution. As a proof-of-concept, the graph theory framework presented in this paper has been successfully applied to the Montecalvello catchment, where it was used to integrate sporadic and incomplete visual surveys with data gathered through 21 camera traps located at a subset of nodes of the network.

The main advantage of the proposed mathematical setup lies in its flexibility, as the method can exploit information derived from non-homogeneous surveys and monitoring techniques, in which different subsets of nodes are observed with diverse temporal frequencies. In particular, the framework provides an objective procedure to integrate the high temporal frequency typical of remote sensing tools with the higher spatial coverage typical of traditional on-the-ground monitoring techniques.[18,30,48] The approach is a purely data-driven method, which is particularly useful in settings where more advanced physically-based models (that are typically data-demanding) cannot be applied.[24,70,71]

The hierarchical chain (or graph) can be updated incrementally, by adding new edges between pairs of disconnected nodes as long as new empirical data becomes available, thereby facilitating the planning of forthcoming surveys. As an example, given the hierarchy depicted in Figure 4, one could foresee an additional field survey during a dry period, in order to observe the network in a relatively contracted state and gather additional information about the structure of the first part of the DAG, which is not still resolved. In fact, while the nodes in the second part of the hierarchy are arranged in a nearly-continuous chain, the first part of the DAG is characterized by a more complicated structure, suggesting that the order of node activation/deactivation of these nodes (and therefore the spatial configurations of the active network when the flowing length is short) has not been fully identified yet. Owing to the transitive property of DAGs, the structure of $\mathbf{H}$ also allows the hierarchy between pairs of nodes that have never been simultaneously observed to be inferred. This could facilitate the monitoring of network dynamics in relatively large catchments (with a catchment area of, say, some tens of km^2), which can hardly be monitored as a whole with appropriate spatial and temporal resolution using currently available technologies.[58] To reconstruct the hierarchy, in fact, different (yet overlapping) subsets of nodes could be surveyed in different dates, and these observations could be subsequently combined together in the hierarchy DAG using the proposed theory.

Thanks to the hierarchical principle, if the river network is properly fragmented into segments/nodes such as the internal heterogeneity of surface flow patterns is properly captured, and the hierarchy among the nodes has been reconstructed with a sufficient amount of surveys (say, 8–10), the observation of a limited number of nodes (again, 8– 10 nodes) could be sufficient to reconstruct the whole spatial distribution of flowing and dry channels in the network (spatial extrapolation). In the case study presented in this paper, the application of the method made it possible to reconstruct, on average, the temporal dynamics of 3.3 km of non-observed streams (corresponding to about 56% of the geomorphic network). The field mapping of the status of the nodes which were not observed in this campaign would otherwise have required significant efforts (several tens of hours of additional fieldwork), demonstrating how the proposed framework can greatly facilitate monitoring activities. Notably, the more the hierarchy graph is unveiled by empirical data, the more the operational advantages during the monitoring of network dynamics. As an example, given the Montecalvello hierarchy depicted in Figure 4, if node 1 is observed as dry we can expect all the subsequent nodes in the hierarchy to be dry too. This would save us from surveying the part of network that contains such nodes. An analogous reasoning can be done when a given node is observed as wet (i.e., all the preceding nodes in the DAG/chain could be predicted as wet). Therefore, in the

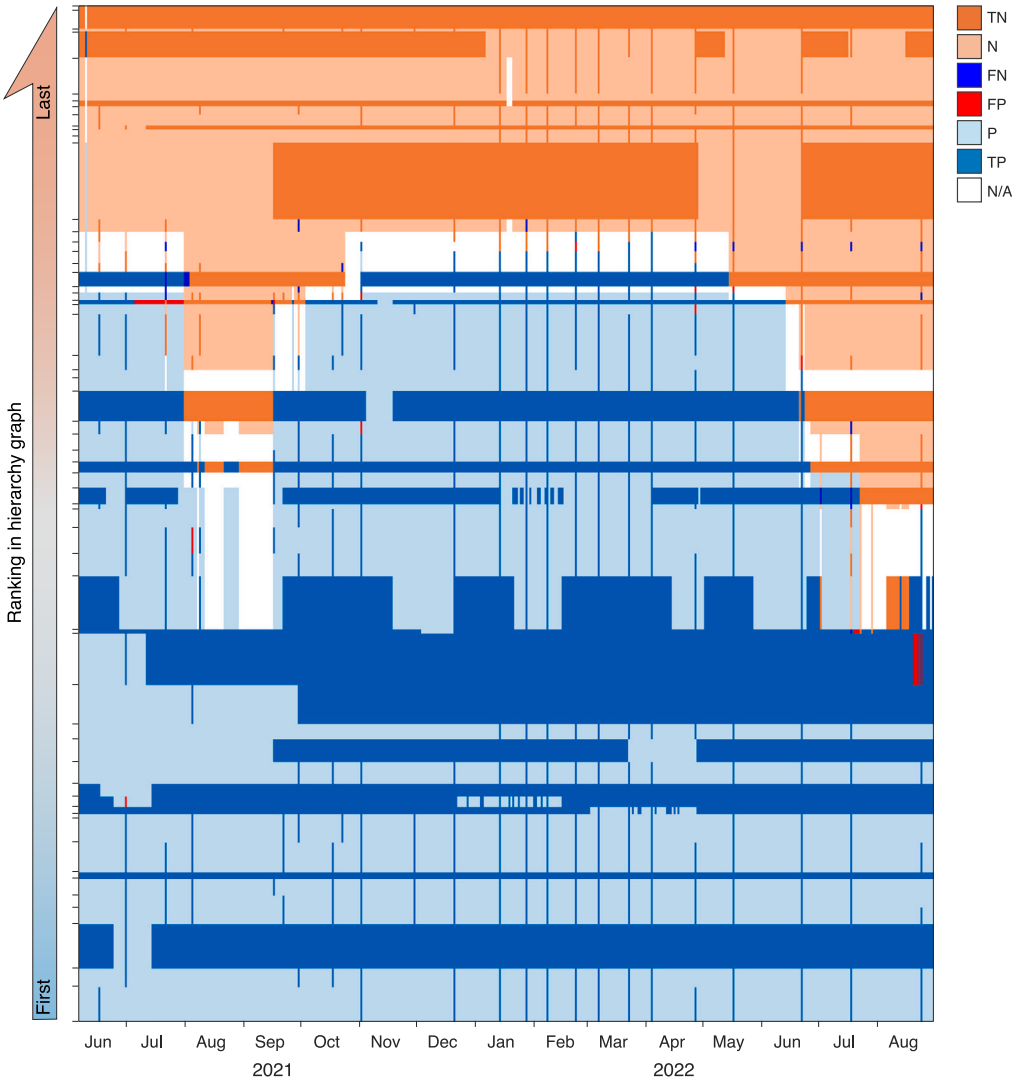


Figure 6. Observed and estimated states of network nodes as a function of time in the period from June 6th, 2021 to August 30th, 2022

True positives (TP) and true negatives (TN) refer to nodes correctly identified by the model as wet or dry, respectively. False positives (FP) and false negatives (FN), instead, highlight model errors. Positives (P) and negatives (N) show node states estimated by the model, for which there are no corresponding empirical observations. N/A identifies no-data.

most favorable scenario only two nodes need to be observed: the two nodes where the transition from wet to dry takes place along the hierarchy.

In conclusion, we propose that reconstructing the hierarchy among the network nodes using a DAG is a fundamental step in the analysis of channel network dynamics, as it allows for an optimized use of empirical data on network expansion and contraction—a feature that distinguishes the method presented in this paper from other commonly used techniques.

Limitations of the study

The mathematical framework proposed here is quite general and applies to any kind of river network that experiences expansion/retraction cycles in response to climate variations. A crucial assumption of the method is the hierarchical structuring of channel network dynamics, which is functional to transform the original graph of the nodes in a DAG (or a chain) and enables one to extrapolate in space the empirical information on the status (wet/dry) of some nodes. The hierarchical behavior of network dynamics, however, has been observed

ubiquitously across a broad range of climatic and geomorphic conditions.[56,57] Therefore, even though the single benchmark application presented here does not enable a comprehensive evaluation of the proposed mathematical framework under a broad range of settings, we believe that the method is well suited to be applied to all the case studies where the hierarchical structuring of stream dynamics has been proved valid. Another limitation of the study is related to the data-driven nature of the method, which cannot be effectively used for predictive purposes in catchments where empirical data about the dynamics of the river network is unavailable. In fact, when too few observations on the status of the network nodes are available (i.e., when most of the elements of the observation matrix are null), many nodes will be either lumped together or disconnected in the hierarchical graph, meaning that their relative position in the DAG will not be defined. This hinders the ability to estimate the activation order of all the nodes in the network and reconstruct the status of one node on the basis of the observation of other reaches. For this reason, we suggest that about 8–10 empirical observations of the whole active network (or any subset of it) is the minimum requirement to get a basic version of the DAG and enable the application of the proposed method for reconstructing the temporal dynamics of the state of unmonitored nodes. These empirical observations should be best carried out under different climatic conditions (i.e., for different configurations of the active network), since new edges are generated in the DAG only between pair of nodes that are observed in a different state (one wet and the other dry). Therefore, the hierarchy between a couple of nodes that has always been observed as wet (or dry) cannot be identified, thereby limiting the capabilities of extrapolating empirical information in space and time. Furthermore, even though currently available conceptual models for network dynamics suggest that the spatial variability of local persistency could be related to measurable physiographic properties such as contributing area and slope,[5,16,58] incorporating this type of morphometric predictors in our data-driven framework may not necessarily improve the underlying model accuracy, due to the pronounced spatial heterogeneity of local persistency in non-perennial stream networks,[57] and the observed limited correlation between the physical position of a node along the network and its ranking in the hierarchy. In spite of these limitations, the mathematical framework outlined here is quite general, and widely applicable in hydrological studies that analyze the intermittent nature of rivers and streams.

Conclusion

In this paper, we propose a graph-theory approach to describe hierarchical channel network dynamics in temporary streams. The hierarchical structure of the nodes in the network, which drives the activation (deactivation) order of the nodes when the wet channel length increases (decreases), was represented here by a DAG. Direct knowledge on the hierarchy of the nodes provides important information on the spatial and temporal dynamics of the active portion of the network, and it can be used to reconstruct the spatial configuration of the entire active length on the basis of the observation of a limited subset of nodes. The approach, therefore, provides a mathematical framework to merge heterogeneous datasets characterized by different temporal frequency and spatial coverage, an instance that applies to all the cases in which different techniques are combined (e.g., visual surveys, on-site sensors, and remote sensing). The theory was successfully applied to predict the stream network dynamics in the Montecalvello catchment (Central Italy). The results demonstrated the ability of the method to integrate data collected via visual surveys and trap cameras, and predict the status of unmonitored nodes. The proposed approach enables an optimized use of available empirical data on stream network dynamics, and thus facilitates the planning and execution of monitoring campaigns in non-perennial rivers.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Study catchment and data collection
  - Numerical implementation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Accuracy of the hierarchy

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107417.

ACKNOWLEDGMENTS

This research was supported by the European Community's Horizon 2020 Excellent Science Programme (grant no. H2020-EU.1.1.-770999). A special thanks goes also to all the personnel that provided help with the empirical activities.

AUTHOR CONTRIBUTIONS

Conceptualization, N.D. and G.B.; Methodology, N.D.; Software, N.D.; Formal Analysis, N.D.; Investigation, S.N.; Writing – Original Draft, N.D. and G.B.; Writing –– Review & Editing, S.N., F.T., and S.G. Visualization, N.D.; Supervision, G.B.; Project Administration, G.B.; Funding Acquisition, G.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 27, 2023
Revised: June 6, 2023
Accepted: July 14, 2023
Published: July 23, 2023

REFERENCES

1. Goodrich, D., et al. (2018). Southwestern Intermittent and Ephemeral Stream Connectivity. JAWRA Journal of the American Water Resources Association *2*, 400–422. https://doi.org/10.1111/1752-1688.12636.

2. Messager, M.L., et al. (2021). Global Prevalence of Non-Perennial Rivers and Streams. Nature *594*, 391–397. https://doi.org/10.1038/s41586-021-03565-5.

3. Tooth, S. (2000). Process, Form and Change in Dryland Rivers: A Review of Recent Research. Earth Sci. Rev. *51*, 67–107. https://doi.org/10.1016/S0012-8252(00)00014-3.

4. Larned, S.T., Datry, T., Arscott, D.B., and Tockner, K. (2010). Emerging Concepts in Temporary-River Ecology. Freshw. Biol. *4*, 717–738. https://doi.org/10.1111/j.1365-2427.2009.02322.x.

5. Lapides, D.A., Leclerc, C.D., Moidu, H., Dralle, D.N., and Hahm, W.J. (2021). Variability of Stream Extents Controlled by Flow Regime and Network Hydraulic Scaling. Hydrol. Process. *3*. https://doi.org/10.1002/hyp.14079.

6. Skoulikidis, N.T., Sabater, S., Datry, T., Morais, M.M., Buffagni, A., Dörflinger, G., Zogaris, S., Del Mar Sánchez-Montoya, M., Bonada, N., Kalogianni, E., et al. (2017). Non-Perennial Mediterranean Rivers in Europe: Status, Pressures, and Challenges for Research and Management. Sci. Total Environ. *577*, 1–18. https://doi.org/10.1016/j.scitotenv.2016.10.147.

7. Giezendanner, J., et al. (2021). A Note on the Role of Seasonal Expansions and Contractions of the Flowing Fluvial Network on Metapopulation Persistence. Water Resour. Res. *11*, 1944–7973. https://doi.org/10.1029/2021WR029813.

8. Datry, T., Larned, S.T., and Tockner, K. (2014). Intermittent Rivers: A Challenge for Freshwater Ecology. Bioscience *64*, 229–235. https://doi.org/10.1093/biosci/bit027. http://academic.oup.com/bioscience/article/64/3/229/224292/Intermittent-Rivers-A-Challenge-for-Freshwater.

9. Bertassello, L.E., Durighetto, N., and Botter, G. (2022). Eco-Hydrological Modelling of Channel Network Dynamics—Part 2: Application to Metapopulation Dynamics. Royal Society Open Science *9*, 220945. https://doi.org/10.1098/rsos.220945.

10. Borg Galea, A., Sadler, J.P., Hannah, D.M., Datry, T., and Dugdale, S.J. (2019). Mediterranean Intermittent Rivers and Ephemeral Streams: Challenges in Monitoring Complexity. Ecohydrology *8*, 1–11. https://doi.org/10.1002/eco.2149.

11. Gregory, K.J., and Walling, D.E. (1968). The Variation of Drainage Density within a Catchment. Int. Assoc. Sci. Hydrol. Bull. *2*, 61–68. https://doi.org/10.1080/02626666809493583.

12. Morgan, R.P.C. (1972). Observations on Factors Affecting the Behaviour of a First-Order Stream. Trans. Inst. Br. Geogr. *56*, 171. https://doi.org/10.2307/621547.

13. Blyth, K., and Rodda, J.C. (1973). A Stream Length Study. Water Resour. Res. *5*, 1454–1461. https://doi.org/10.1029/WR009i005p01454.

14. Day, D.G. (1978). Drainage Density Changes during Rainfall. Earth Surf. Process. *3*, 319–326. https://doi.org/10.1002/esp.3290030310.

15. Jaeger, K.L., Montgomery, D.R., and Bolton, S.M. (2007). Channel and Perennial Flow Initiation in Headwater Streams: Management Implications of Variability in Source-Area Size. Environ. Manag. *5*, 775. https://doi.org/10.1007/s00267-005-0311-2.

16. Godsey, S., and Kirchner, J. (2014). Dynamic, Discontinuous Stream Networks: Hydrologically Driven Variations in Active Drainage Density, Flowing Channels and Stream Order. Hydrol. Process. *28*, 5791–5803. https://doi.org/10.1002/hyp.10310.

17. Ågren, A., Lidberg, W., and Ring, E. (2015). Mapping Temporal Dynamics in a Forest Stream Network—Implications for Riparian Forest Management. Forests *6*, 2982–3001. https://doi.org/10.3390/f6092982.

18. Whiting, J.A., and Godsey, S.E. (2016). Discontinuous Headwater Stream Networks with Stable Flowheads, Salmon River Basin, Idaho. Hydrol. Process. *13*, 2305–2316. https://doi.org/10.1002/hyp.10790.

19. Datry, T., Pella, H., Leigh, C., Bonada, N., and Hugueny, B. (2016). A Landscape Approach to Advance Intermittent River Ecology. Freshw. Biol. *8*, 1200–1213. https://doi.org/10.1111/fwb.12645.

20. Jensen, C.K., McGuire, K.J., and Prince, P.S. (2017). Headwater Stream Length Dynamics across Four Physiographic Provinces of the Appalachian Highlands. Hydrol. Process. *19*, 3350–3363. https://doi.org/10.1002/hyp.11259.

21. Shaw, S.B., Bonville, D.B., and Chandler, D.G. (2017). Combining Observations of Channel Network Contraction and Spatial Discharge Variation to Inform Spatial Controls on Baseflow in Birch Creek, Catskill Mountains,

USA. J. Hydrol. Reg. Stud. *12*, 1–12. https://doi.org/10.1016/j.ejrh.2017.03.003.

22. Floriancic, M.G., Meerveld, I., Smoorenburg, M., Margreth, M., Naef, F., Kirchner, J.W., and Molnar, P. (2018). Spatio-temporal Variability in Contributions to Low Flows in the High Alpine Poschiavino Catchment. Hydrol. Process. *26*, 3938–3953. https://doi.org/10.1002/hyp.13302.

23. van Meerveld, H.J.I., Kirchner, J.W., Vis, M.J.P., Assendelft, R.S., and Seibert, J. (2019). Expansion and Contraction of the Flowing Stream Network Alter Hillslope Flowpath Lengths and the Shape of the Travel Time Distribution. Hydrol. Earth Syst. Sci. *11*, 4825–4834. https://doi.org/10.5194/hess-23-4825-2019. https://hess.copernicus.org/articles/23/4825/2019/.

24. Senatore, A., et al. (2021). Monitoring and Modeling Drainage Network Contraction and Dry Down in Mediterranean Headwater Catchments. Water Resour. Res. *6*, 1944–7973. https://doi.org/10.1029/2020WR028741.

25. Shaw, S.B. (2016). Investigating the Linkage between Streamflow Recession Rates and Channel Network Contraction in a Mesoscale Catchment in New York State. Hydrol. Process. *3*, 479–492. https://doi.org/10.1002/hyp.10626.

26. Zimmer, M.A., and McGlynn, B.L. (2017). Ephemeral and Intermittent Runoff Generation Processes in a Low Relief, Highly Weathered Catchment. Water Resour. Res. *8*, 7055–7077. https://doi.org/10.1002/2016WR019742.

27. Lovill, S.M., Hahm, W.J., and Dietrich, W. (2018). Drainage from the Critical Zone: Lithologic Controls on the Persistence and Spatial Extent of Wetted Channels during the Summer Dry Season. Water Resour. Res. *8*, 5702–5726. https://doi.org/10.1029/2017WR021903.

28. Zimmer, M.A., and McGlynn, B.L. (2018). Lateral, Vertical, and Longitudinal Source Area Connectivity Drive Runoff and Carbon Export Across Watershed Scales. Water Resour. Res. *3*, 1576–1598. https://doi.org/10.1002/2017WR021718.

29. Jaeger, K., Sando, R., McShane, R., Dunham, J., Hockman-Wert, D., Kaiser, K., Hafen, K., Risley, J., and Blasch, K. (2019). Probability of Streamflow Permanence Model (PROSPER): A Spatially Continuous Model of Annual Streamflow Permanence throughout the Pacific Northwest. J. Hydrol. X *2*, 100005. https://doi.org/10.1016/j.hydroa.2018.100005.

30. Durighetto, N., et al. (2020). Intraseasonal Drainage Network Dynamics in a Headwater Catchment of the Italian Alps. Water Resour. Res. *4*. https://doi.org/10.1029/2019WR025563.

31. Wigington, P.J., Moser, T.J., and Lindeman, D.R. (2005). Stream Network Expansion: A Riparian Water Quality Factor. Hydrol. Process. *8*, 1715–1721. https://doi.org/10.1002/hyp.5866.

32. Malard, F., et al. (2006). Flood-Pulse and Riverscape Dynamics in a Braided Glacial River. Ecology *3*, 704–716. https://doi.org/10.1890/04-0889.

33. Spence, C., and Mengistu, S. (2016). Deployment of an Unmanned Aerial System to Assist in Mapping an Intermittent Stream. Hydrol. Process. *3*, 493–500. https://doi.org/10.1002/hyp.10597.

34. Roelens, J., Rosier, I., Dondeyne, S., Van Orshoven, J., and Diels, J. (2018). Extracting Drainage Networks and Their Connectivity Using LiDAR Data. Hydrol. Process. *8*, 1026–1037. https://doi.org/10.1002/hyp.11472.

35. Micieli, M., Botter, G., Mendicino, G., and Senatore, A. (2021). UAV Thermal Images for Water Presence Detection in a Mediterranean Headwater Catchment. Rem. Sens. *1*, 108. https://doi.org/10.3390/rs14010108.

36. Constantz, J., Stonestorm, D., Stewart, A.E., Niswonger, R., and Smith, T.R. (2001). Analysis of Streambed Temperatures in Ephemeral Channels to Determine Streamflow Frequency and Duration. Water Resour. Res. *37*, 317–328. https://doi.org/10.1029/2000WR900271.

37. Blasch, K.W., Ferré, T.P.A., Christensen, A.H., and Hoffmann, J.P. (2002). New Field Method to Determine Streamflow Timing Using Electrical Resistance Sensors. Vadose Zone J. *1*, 289–299. https://doi.org/10.2136/vzj2002.2890.

38. Buttle, J., et al. (2012). An Overview of Temporary Stream Hydrology in Canada. Can. Water Resour. J. *37*, 279–310. https://doi.org/10.4296/cwrj2011-903.

39. Blasch, K.W., Ferré, T.P.A., and Hoffmann, J.P. (2004). A Statistical Technique for Interpreting Streamflow Timing Using Streambed Sediment Thermographs. Vadose Zone J. *3*, 936–946. https://doi.org/10.2136/vzj2004.0936.

40. Adams, E.A., et al. (2006). Electrical Resistance Sensors Record Spring Flow Timing, Grand Canyon, Arizona. Ground Water *44*, 5. https://doi.org/10.1111/j.1745-6584.2006.00223.x.

41. Goulsbra, C.S., Lindsay, J.B., and Evans, M.G. (2009). A New Approach to the Application of Electrical Resistance Sensors to Measuring the Onset of Ephemeral Streamflow in Wetland Environments. Water Resour. Res. *45*, 1–7. https://doi.org/10.1029/2009WR007789.

42. Bhamjee, R., and Lindsay, J.B. (2011). Ephemeral Stream Sensor Design Using State Loggers. Hydrol. Earth Syst. Sci. *3*, 1009–1021. https://doi.org/10.5194/hess-15-1009-2011.

43. Jaeger, K.L., and Olden, J.D. (2012). Electrical Resistance Sensor Arrays as a Means to Quantify Longitudinal Connectivity of Rivers. River Res. Appl. *10*, 1843–1852. https://doi.org/10.1002/rra.1554.

44. Chapin, T.P., Todd, A.S., and Zeigler, M.P. (2014). Robust, Low-Cost Data Loggers for Stream Temperature, Flow Intermittency, and Relative Conductivity Monitoring. Water Resour. Res. *8*, 6542–6548. https://doi.org/10.1002/2013WR015158.

45. Goulsbra, C., Evans, M., and Lindsay, J. (2014). Temporary Streams in a Peatland Catchment: Pattern, Timing, and Controls on Stream Network Expansion and Contraction. Earth Surf. Process. Landforms *39*, 790–803. https://doi.org/10.1002/esp.3533.

46. Peirce, S.E., and Lindsay, J.B. (2015). Characterizing Ephemeral Streams in a Southern Ontario Watershed Using Electrical Resistance Sensors. Hydrol. Process. *29*, 103–111. https://doi.org/10.1002/hyp.10136.

47. Bhamjee, R., Lindsay, J.B., and Cockburn, J. (2016). Monitoring Ephemeral Headwater Streams: A Paired-Sensor Approach. Hydrol. Process. *30*, 888–898. https://doi.org/10.1002/hyp.10677.

48. Jensen, C.K., McGuire, K.J., McLaughlin, D.L., and Scott, D.T. (2019). Quantifying Spatiotemporal Variation in Headwater Stream Length Using Flow Intermittency Sensors. Environ. Monit. Assess. *191*, 226. https://doi.org/10.1007/s10661-019-7373-8.

49. Paillex, A., Siebers, A.R., Ebi, C., Mesman, J., and Robinson, C.T. (2020). High Stream Intermittency in an Alpine Fluvial Network: Val Roseg, Switzerland. Limnol. Oceanogr. *65*, 557–568. https://doi.org/10.1002/lno.11324.

50. Zanetti, F., Durighetto, N., Vingiani, F., and Botter, G. (2022). Technical Note: Analyzing River Network Dynamics and the Active Length–Discharge Relationship Using Water Presence Sensors. Hydrol. Earth Syst. Sci. *26*, 3497–3516. https://doi.org/10.5194/hess-26-3497-2022.

51. Assendelft, R.S., and van Meerveld, H.J.I. (2019). A Low-Cost, Multi-Sensor System to Monitor Temporary Stream Dynamics in Mountainous Headwater Catchments. Sensors *19*, 4645. https://doi.org/10.3390/s19214645.

52. Schoener, G. (2018). Time-Lapse Photography: Low-Cost, Low-Tech Alternative for Monitoring Flow Depth. J. Hydrol. Eng. *23*, 06017007. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001616.

53. Kaplan, N.H., Sohrt, E., Blume, T., and Weiler, M. (2019). Monitoring Ephemeral, Intermittent and Perennial Streamflow: A Dataset from 182 Sites in the Attert Catchment, Luxembourg. Earth Syst. Sci. Data *11*, 1363–1374. https://doi.org/10.5194/essd-11-1363-2019.

54. Noto, S., Tauro, F., Petroselli, A., Apollonio, C., Botter, G., and Grimaldi, S. (2022). Low-Cost Stage-Camera System for Continuous Water-Level Monitoring in Ephemeral Streams. Hydrol. Sci. J. *67*, 1439–1448. https://doi.org/10.1080/02626667.2022.2079415.

55. Tauro, F., Noto, S., Botter, G., and Grimaldi, S. (2022). Assessing the Optimal Stage-Cam Target for Continuous Water Level Monitoring in Ephemeral Streams: Experimental

Evidence. Rem. Sens. *14*, 6064. https://doi.org/10.3390/rs14236064.

56. Botter, G., and Durighetto, N. (2020). The Stream Length Duration Curve: A Tool for Characterizing the Time Variability of the Flowing Stream Length. Water Resour. Res. *56*. https://doi.org/10.1029/2020WR027282.

57. Botter, G., Vingiani, F., Senatore, A., Jensen, C., Weiler, M., McGuire, K., Mendicino, G., and Durighetto, N. (2021). Hierarchical Climate-Driven Dynamics of the Active Channel Length in Temporary Streams. Sci. Rep. *11*, 21503. https://doi.org/10.1038/s41598-021-00922-2.

58. Durighetto, N., and Botter, G. (2022). On the Relation Between Active Network Length and Catchment Discharge. Geophys. Res. Lett. *49*, 1944–8007. https://doi.org/10.1029/2022GL099500.

59. Durighetto, N., and Botter, G. (2021). Time-lapse Visualization of Spatial and Temporal Patterns of Stream Network Dynamics. Hydrol. Process. *2*. https://doi.org/10.1002/hyp.14053.

60. Durighetto, N., Bertassello, L.E., and Botter, G. (2022). Eco-Hydrological Modelling of Channel Network Dynamics—Part 1: Stochastic Simulation of Active Stream Expansion and Retraction. R. Soc. Open Sci. *9*, 220944. https://doi.org/10.1098/rsos.220944.

61. Connor-Streich, G., Henshaw, A.J., Brasington, J., Bertoldi, W., and Harvey, G.L. (2018). Let's Get Connected: A New Graph Theory-based Approach and Toolbox for Understanding Braided River Morphodynamics. WIREs Water *5*. https://doi.org/10.1002/wat2.1296.

62. Heckmann, T., and Schwanghart, W. (2013). Geomorphic Coupling and Sediment Connectivity in an Alpine Catchment — Exploring Sediment Cascades Using Graph Theory. Geomorphology *182*, 89–103. https://doi.org/10.1016/j.geomorph.2012.10.033.

63. Masselink, R.J.H., Heckmann, T., Temme, A.J.A.M., Anders, N.S., Gooren, H.P.A., and Keesstra, S.D. (2017). A Network Theory Approach for a Better Understanding of Overland Flow Connectivity: Networks for a Better Understanding of Overland Flow Connectivity. Hydrol. Process. *31*, 207–220. https://doi.org/10.1002/hyp.10993.

64. Tejedor, A., Longjas, A., Zaliapin, I., and Foufoula-Georgiou, E. (2015). Delta Channel Networks: 2. Metrics of Topologic and Dynamic Complexity for Delta Comparison, Physical Inference, and Vulnerability Assessment: METRICS OF TOPOLOGIC AND DYNAMIC COMPLEXITY FOR DELTAS. Water Resour. Res. *51*, 4019–4045. https://doi.org/10.1002/2014WR016604.

65. Bertassello, L.E., Aubeneau, A.F., Botter, G., Jawitz, J.W., and Rao, P.S.C. (2020). Emergent Dispersal Networks in Dynamic Wetlandscapes. Sci. Rep. *10*, 14696. https://doi.org/10.1038/s41598-020-71739-8.

66. Phillips, R.W., Spence, C., and Pomeroy, J.W. (2011). Connectivity and Runoff Dynamics in Heterogeneous Basins. In Hydrological Processes. https://doi.org/10.1002/hyp.8123.

67. Liao, C., et al. (2023). Topological relationship-based flow direction modeling: Mesh-independent river networks representation. J. Adv. Model. Earth Syst. *15*, e2022MS003089. https://doi.org/10.1029/2022MS003089.

68. Zuecco, G., Rinderer, M., Penna, D., Borga, M., and van Meerveld, H.J. (2019). Quantification of Subsurface Hydrologic Connectivity in Four Headwater Catchments Using Graph Theory. Sci. Total Environ. *646*, 1265–1280. https://doi.org/10.1016/j.scitotenv.2018.07.269.

69. Sun, J., et al. (2017). Breaking Cycles in Noisy Hierarchies. WebSci 2017 - Proc. 2017 ACM Web Sci. Conf. 151–160. https://doi.org/10.1145/3091478.3091495.

70. Ward, A.S., Schmadel, N.M., and Wondzell, S.M. (2018). Simulation of Dynamic Expansion, Contraction, and Connectivity in a Mountain Stream Network. Adv. Water Resour. *114*, 64–82. https://doi.org/10.1016/j.advwatres.2018.01.018.

71. Jensen, C.K., McGuire, K.J., Shao, Y., and Andrew Dolloff, C. (2018). Modeling Wet Headwater Stream Networks across Multiple Flow Conditions in the Appalachian Highlands. Earth Surf. Process. Landforms *13*, 2762–2778. https://doi.org/10.1002/esp.4431.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Monitored network dynamics of the Montecalvello catchment	Authors	Research Data Unipd: https://doi.org/10.25430/researchdata.cab.unipd.it.00000831
Software and algorithms		
Code to reproduce the results of this study	Authors	Zenodo: https://doi.org/10.5281/zenodo.7956623

RESOURCE AVAILABILITY

Lead contact

Further information and results for resources should be directed to and will be fulfilled by the lead contact, Nicola Durighetto (nicola.durighetto@unipd.it).

Materials availability

This study did not generate new datasets.

Data and code availability

This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.

All original code has been deposited on Zenodo and is publicly available as of the date of publication. Links are listed in the key resources table.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Study catchment and data collection

The Montecalvello catchment (Figure 2) is located in central Italy, about 75 km north of Rome, and covers an area of 3.73 km^2 with elevation ranging from 130 to 340 m a.s.l.. In the hillslopes, the land cover is dominated by olive groves, vineyards and other crops, while a broad-leaved forest is mainly found in the riparian areas. The site has a typical Mediterranean climate, characterized by arid summers with sparse rainfall events and mild, rainy winters. The potential stream network, defined by the presence of permanent channelization signs and/or flowing water, is about 5.9 km long. Channel widths range from about 10 cm to more than a meter. The riverbed is mainly composed of silt and clay, but gravel and conglomerates can also be found in the riparian zone.

The presence/absence of flowing water along the network was monitored by combining observations from field surveys with a number of trap cameras. Field surveys were carried out on a biweekly to monthly time frequency, from October 2019 to August 2022. Each survey consisted in walking along the whole network and collecting the GPS coordinates of the points were surface flow starts or stops, thus allowing the monitoring of the active portion of the stream network regardless of its connectivity to the outlet. At total of 40 visual surveys of the active network was carried out. Based on the observations, the network was described by a total of $n = 58$ nodes with an average length of about 100 m. The actual network length associated to each node varied from 18 to 450 m, and was defined by ensuring that each node described a uniform reach. Therefore, shorter nodes were placed in locations where network dynamics were more heterogeneous, in order to provide a more precise description of the active network.

To enhance the temporal resolution of the surveys, 21 trap cameras were installed along the network, with a mean distance of 280 m. The specific location of each camera was chosen on the basis of the first few surveys, in order to deploy only one camera per each uniform stretch of network, therefore limiting the total

number of installed instruments. The specific locations of the cameras also corresponded to nodes identified by the visual surveys, in order to ease the merging of the two datasets. Each trap camera collected one image every 20 minutes; the pictures were then manually classified as wet (i.e. flowing water) or dry. The data was then aggregated at the daily time scale: a node was considered as active if at least one picture suggested flowing water during the day. This allowed us to properly take sub-daily dynamics into account. The cameras were active for a total of 451 days, in the period from June 6th, 2021, to August 30th, 2022. However, some cameras were operational for slightly shorter time periods due to malfunctions.

Numerical implementation

Starting from the collected empirical data, ESRI ArcGIS was used to create a shapefile describing the geometry of the geomorphic network. The attribute table associated with the shapefile was composed of two main sections: the first section reported the main morphological properties of the nodes (length, average slope, strahler order), while the second section stored the empirical data collected from visual surveys (one column per date, reporting the wet/dry/NA status of each node). The same structure was also employed to build an excel table reporting the status data derived from trap cameras. The theoretical framework proposed in this paper was then implemented via a MATLAB script. The numerical computations strictly followed the equations reported in Section "a graph theory framework for representing and reconstructing hierarchical stream network dynamics". The Graph and Network Algorithms built into MATLAB was used to generate the hierarchy DAG starting from matrix $\mathbf{H}$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Accuracy of the hierarchy

The need for cycles to be broken by removing edges suggests that the inferred hierarchy may not agree with all the available empirical observations. The degree of agreement between the hierarchy and the empirical data can be quantified through the accuracy. Without loss of generality, let's assume that nodes are reordered in order of hierarchy (this can be achieved with a topological sort starting from $\mathbf{H}$). In such case, $\mathbf{H}$ and $\mathbf{R}$ are upper triangular. More interestingly, the upper triangle of $\mathbf{O}$ contains the observations favorable to the estimated hierarchy, while the lower triangle contains the observations against. As such, the accuracy of the hierarchy can be calculated as

$$acc_h = \sum \mathbf{O}^u \Big/ \sum \mathbf{O} \qquad \text{(Equation 5)}$$

where $\mathbf{O}^u$ is the upper component of the LDU decomposition of $\mathbf{O}$ (i.e. the upper triangle), and $\sum$ denotes the summation of all the elements in the matrix.