Supplementary Information
Reconstructing Flow Dynamics of Temporary Streams Using Citizen Science Data and a Hierarchical Network Structure

Mirjam Scheller[1,2], Nicola Durighetto[3,4], Anja Fischer[1], Reinhard Furrer[5], Marc Vis[1], Jan Seibert[1], Gianluca Botter[2], and Ilja van Meerveld[1]

[1]University of Zurich, Department of Geography, Zurich, Switzerland
[2] now at Deutsche Rückversicherung AG, Düsseldorf, Germany
[3]University of Padua, Department of Civil, Environmental and Architectural Engineering, Padova, Italy
[4] now at University of Oulu, Faculty of Technology, Water, Energy and Environmental Engineering Unit, Oulu, Finland
[5]University of Zurich, Department of Mathematical Modeling and Machine Learning, Zurich, Switzerland

Correspondence to: Mirjam Scheller (mirjam.scheller@uzh.ch)

Contents of this file

Supplementary Materials S1 to S4
Figures S1 to S12
Tables S1 to S2

Introduction

In this supplementary information, we present additional information and figures for the study area and stream sites, the method used for the calculation of the nodewise accuracies for a random model, and the methods and results of the correlation between the cumulative net precipitation and the fraction of flowing nodes based on the skew-normal distribution.

Supplementary material S1: Additional information on the study area and data sets

S1.1 Photos of the stream sites

Figure S1. Pictures of temporary streams on Zuriberg. The ID of each site (as used in this study) is printed above each picture (see Figure S2 for the location), with the ID used for each site in the CrowdWater (CW) in parentheses (CW App-ID). The CrowdWater ID can be used to download the data and see all photos for a site: https://www.spotteron.com/crowdwater/spots/CW App-ID).

S1.2 Maps of the study area

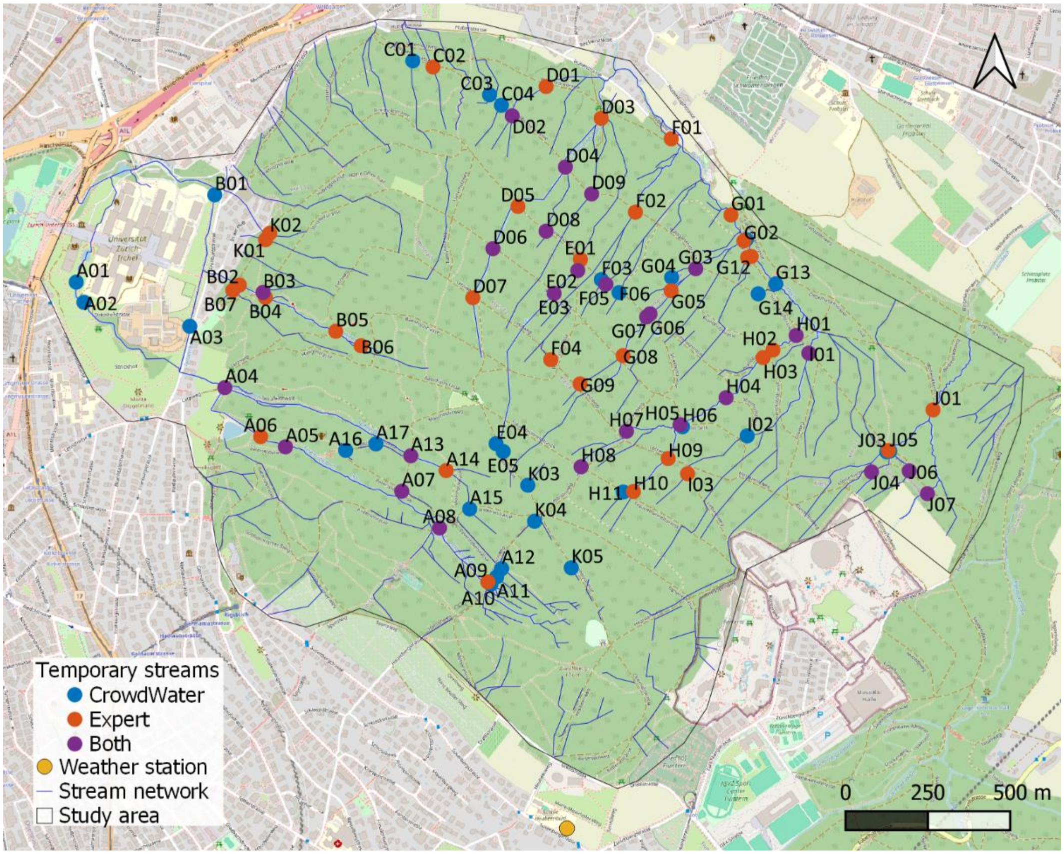


Figure S2. Map of the study area (Zuriberg) with the stream sites and their ID, and the mapped geomorphic stream network. Background image: Open Street Map (OpenStreetMap contributors 2017).

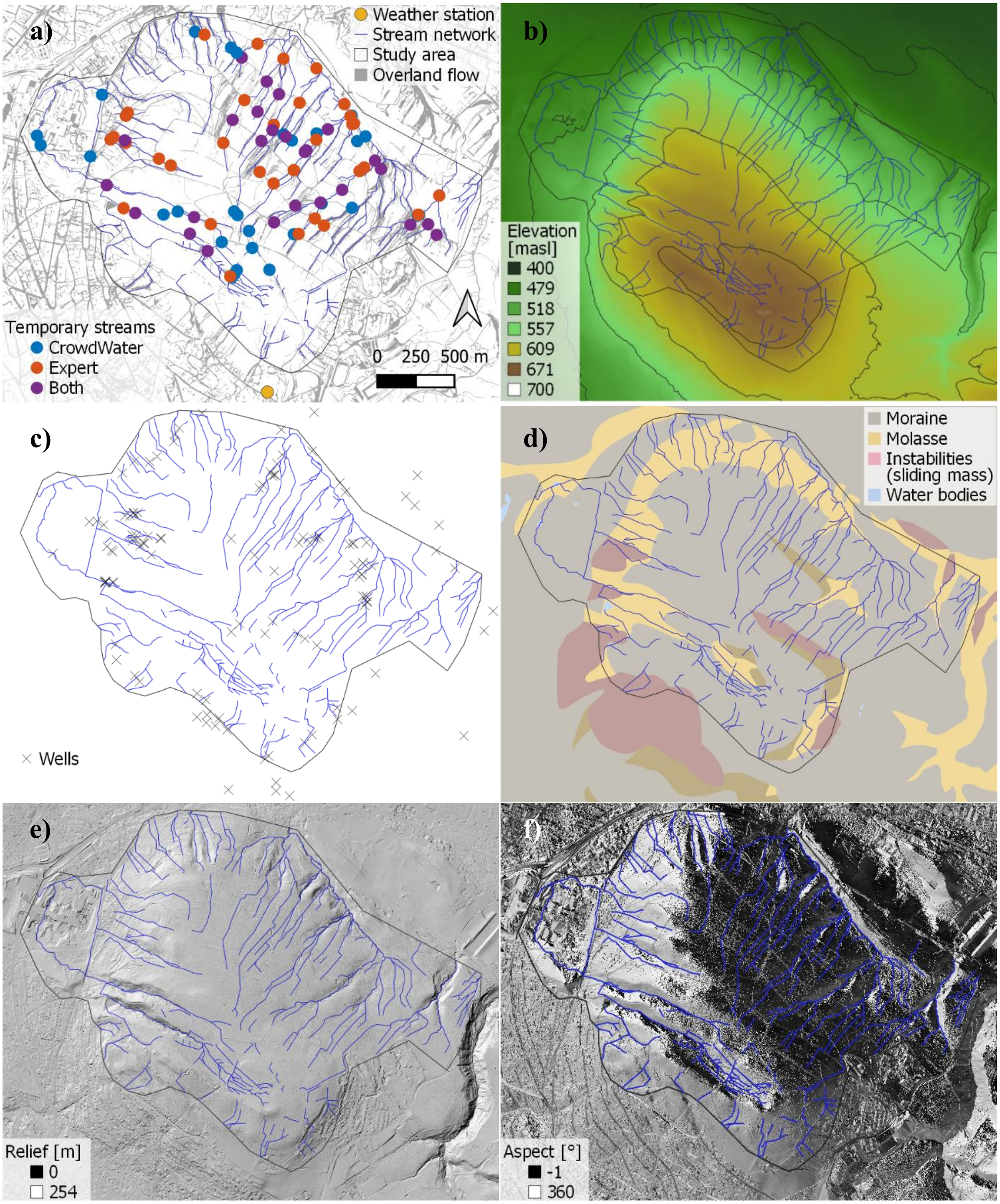


Figure S3. Maps of the study area showing a) the location of the study sites (color-coded by dataset) and likely locations of overland flow based on the Oberflächenabfluss map of the Swiss Federal Office for the Environment, b) the 2 m Digital Elevation Model (swissALTI3D from the Swiss Federal Office of Topography, SwissTopo), c) the location of wells (based on the Geographisches Informationssystem des Kantons Zürich (GIS-ZH)), d) the surficial Geology (based on the GeoCover map from SwissTopo), e) the relief (swissALTI3D from SwissTopo), and f) the aspect (swissALTI3D from SwissTopo).

S1.3 Derivation of geomorphic stream network

We mapped the geomorphic stream network for about one third of the study area during three days in August and September 2023. We included dry streambeds as part of the stream network and mapped all visible traces of surface flow. To derive the stream network for the entire study area, we used the simulated overland flow pathways from the Oberflächenabfluss map of the Swiss Federal Office for the Environment (FOEN) and the roads and paths indicated in the Open Street Map (see Figure S3a). The mapped stream network aligned well with the information provided by the available geodata (based on a visual comparison). To validate the stream network derived from the overland flow map, we mapped the streams in an area that had thus far not been mapped. Again, the derived stream network aligned well with the field observations (based on a visual comparison).

S1.4 Additional information on the datasets

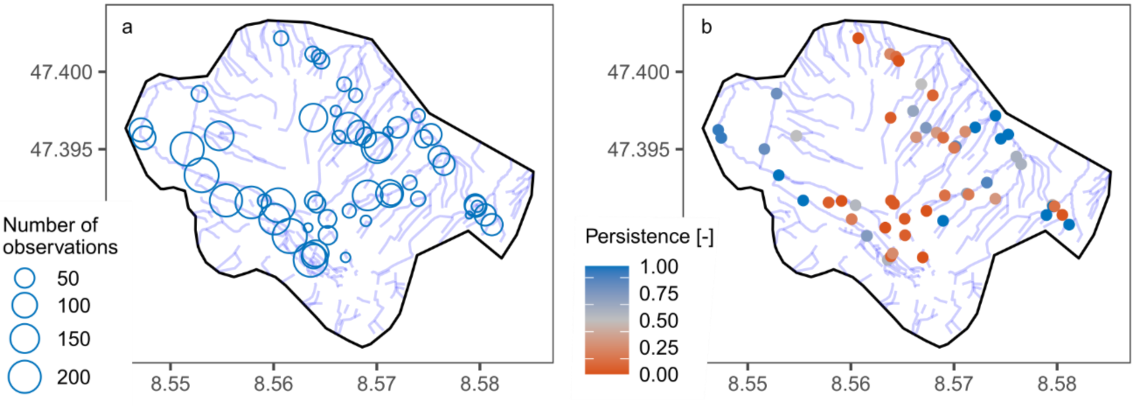


Figure S4. Number of observations per study site (a) and fraction of the observations for which the stream was observed to be flowing (i.e., persistence; b) for the CrowdWater dataset.

Table S1. Overview of the 88 temporary stream sites observed by citizen scientists (CW: CrowdWater) and an expert. (Coordinate system: EPSG:2056 - CH1903+ / LV95 - Projected). For the calculation of the area and mean slope, see section S1.5. The persistence is shown for the observations (i.e., fraction of the observations for which the stream was observed to be flowing) and the application of the hierarchical structure (fraction of days for which there was at least one observation for which the stream was inferred to be flowing based on the application of the hierarchical model). Sites for which no persistence is shown (-) were either not observed by the citizen scientists or the expert.

ID	Lon.	Lat.	Area [m²]	Mean Slope [°]	Number of observations		Persistence observations		Persistence predicted	
					CW	Expert	CW	Expert	CW	Expert
A01	2683681	1250084	547792	12	85	0	1.00	-	0.95	-
A02	2683704	1250023	544120	12	81	0	0.99	-	0.95	-
A03	2684026	1249950	507681	13	220	0	0.92	-	0.83	-
A04	2684132	1249764	156170	12	232	24	1.00	1.00	1.00	1.00
A05	2684317	1249584	216881	8	201	24	1.00	1.00	1.00	1.00
A06	2684241	1249615	226940	11	0	24	-	1.00	-	1.00
A07	2684670	1249449	4660	5	177	24	0.19	0.17	0.18	0.13
A08	2684785	1249337	103572	12	235	24	0.76	1.00	0.74	1.00
A09	2684932	1249174	74418	6	0	24	-	0.88	-	0.88
A10	2684941	1249167	3423	5	230	0	0.54	-	0.58	-
A11	2684959	1249189	5891	5	92	0	0.15	-	0.04	-
A12	2684974	1249214	23013	6	117	0	0.31	-	0.25	-
A13	2684698	1249557	50760	5	226	24	0.54	0.92	0.55	1.00
A14	2684805	1249512	31998	5	0	24	-	0.50	-	0.50
A15	2684877	1249395	5273	5	3	0	0.00	-	0.00	-
A16	2684500	1249573	1478	5	194	0	0.03	-	0.04	-
A17	2684592	1249593	5661	5	14	0	0.00	-	0.00	-
B01	2684102	1250349	697915	12	26	0	0.66	-	0.81	-
B02	2684177	1250075	110398	10	0	24	-	0.58	-	0.57
B03	2684248	1250053	8403	15	151	24	0.50	0.50	0.50	0.55
B04	2684256	1250038	98679	10	0	24	-	0.92	-	0.92
B05	2684470	1249935	74302	7	0	24	-	0.79	-	0.86
B06	2684547	1249891	5926	7	0	24	-	0.17	-	0.17
B07	2684157	1250059	30925	13	0	24	-	0.00	-	0.00
C01	2684704	1250756	26334	17	18	0	0.00	-	0.00	-
C02	2684765	1250737	90485	10	0	24	-	0.08	-	0.05
C03	2684937	1250653	5276	16	22	0	0.58	-	0.23	-
C04	2684973	1250621	56177	9	18	0	0.00	-	0.11	-
D01	2685109	1250678	43213	13	0	24	-	0.54	-	0.57

D02	2685006	1250589	33099	12	25	24	0.00	0.00	0.00	0.00
D03	2685276	1250581	156844	10	0	24	-	1.00	-	1.00
D04	2685167	1250433	93154	9	18	24	0.96	1.00	0.50	1.00
D05	2685023	1250314	65542	8	0	24	-	0.25	-	0.25
D06	2684947	1250186	52458	8	138	24	0.01	0.08	0.04	0.05
D07	2684886	1250037	27928	8	0	24	-	0.17	-	0.13
D08	2685109	1250239	7776	11	6	24	0.84	0.79	0.67	0.87
D09	2685247	1250352	43032	10	15	24	0.00	0.33	0.07	0.43
E01	2685213	1250153	63358	7	0	24	-	0.96	-	1.00
E02	2685205	1250119	61974	7	170	24	0.86	0.50	0.75	0.50
E03	2685133	1250050	56425	7	14	24	0.29	0.58	0.29	0.57
E04	2684957	1249593	3772	6	39	0	0.03	-	0.03	-
E05	2684978	1249570	2810	5	36	0	0.03	-	0.03	-
F01	2685489	1250519	93721	12	0	24	-	1.00	-	1.00
F02	2685380	1250296	62945	9	0	24	-	0.58	-	0.57
F03	2685277	1250092	1790	12	94	0	0.33	-	0.31	-
F04	2685123	1249849	11490	10	0	24	-	0.17	-	0.13
F05	2685291	1250077	16390	9	66	24	0.60	0.50	0.45	0.45
F06	2685331	1250052	5398	10	55	0	0.17	-	0.11	-
G01	2685671	1250288	819847	10	0	24	-	1.00	-	1.00
G02	2685710	1250209	178258	10	0	24	-	1.00	-	1.00
G03	2685564	1250125	134907	9	61	24	1.00	1.00	1.00	1.00
G04	2685490	1250098	17071	11	4	0	0.00	-	0.25	-
G05	2685489	1250058	72978	8	0	24	-	0.71	-	0.83
G06	2685425	1249988	70469	8	142	24	0.97	1.00	0.87	1.00
G07	2685414	1249978	68380	8	174	24	0.13	0.04	0.16	0.04
G08	2685343	1249861	59694	8	0	24	-	0.08	-	0.09
G09	2685212	1249775	12352	8	0	24	-	0.04	-	0.00
G10	2685717	1250213	625719	10	15	0	0.99	-	1.00	-
G11	2685721	1250159	40918	11	0	24	-	1.00	-	1.00
G12	2685732	1250162	552792	9	0	24	-	1.00	-	1.00
G13	2685807	1250079	549480	9	64	0	1.00	-	1.00	-
G14	2685753	1250048	1914	14	38	0	1.00	-	1.00	-
H01	2685869	1249922	231961	9	66	24	0.71	0.92	0.59	0.95
H02	2685798	1249878	224864	9	0	24	-	1.00	-	1.00
H03	2685768	1249855	219503	9	0	24	-	1.00	-	1.00

H04	2685656	1249733	192537	9	18	24	0.88	0.96	0.89	1.00
H05	2685516	1249651	167928	9	126	24	0.81	1.00	0.67	1.00
H06	2685523	1249645	23	10	65	0	0.24	-	0.17	-
H07	2685353	1249631	68208	8	167	24	0.08	0.13	0.13	0.13
H08	2685215	1249523	57758	7	15	24	0.01	0.00	0.00	0.00
H09	2685480	1249548	55662	9	0	24	-	0.21	-	0.17
H10	2685374	1249448	41529	8	0	24	-	0.67	-	0.57
H11	2685343	1249447	4434	10	6	0	0.97	-	1.00	-
I01	2685907	1249868	192903	10	68	24	0.72	0.96	0.59	1.00
I02	2685720	1249617	49934	10	20	0	0.15	-	0.30	-
I03	2685539	1249502	26470	11	0	24	-	0.79	-	0.83
J01	2686285	1249696	223695	10	0	24	-	1.00	-	1.00
J02	2686149	1249573	32558	8	0	24	-	1.00	-	1.00
J03	2686144	1249569	11908	10	66	0	1.00	-	1.00	
J04	2686098	1249508	7811	10	2	24	1.00	1.00	1.00	1.00
J05	2686153	1249571	20469	7	8	0	0.03	-	0.13	-
J06	2686212	1249511	16651	6	66	24	0.03	0.04	0.02	0.00
J07	2686268	1249442	6708	7	66	24	1.00	1.00	1.00	1.00
K01	2684256	1250213	37362	12	0	24	-	0.00	-	0.00
K02	2684269	1250232	32335	13	0	24	-	0.08	-	0.05
K03	2685053	1249467	36632	6	38	0	0.03	-	0.03	-
K04	2685074	1249358	5078	6	43	0	0.05	-	0.05	-
K05	2685185	1249217	3048	5	4	0	0.00	-	0.00	-

S1.5 Catchment characteristics

Due to anthropogenic alterations of the study area (e.g., forest roads, culverts at forest roads, etc.), the drainage areas could not be derived directly from the 2 m digital elevation model (DEM) provided by the Federal Office of Topography SwissTopo (Fig. S3b). We, therefore, burned the geomorphic stream network 2 m deep into the DEM and applied the following formula to the raster calculator in ArcMap:

$$DEM_{burned} = DEM - \left(\frac{2}{2+Euclidean\ Distance(Stream\ network)}\right)^{k} * H \tag{S1}$$

where H describes the depth and k the shape factor. The calculated drainage areas best represented what we observed in the field when setting k to 0.5 and H to 2 m. We then used the functions of the Spatial Analyst Tools in ArcMap to first fill the sinks in the burned DEM and determined the flow directions with the D8 method. We had to manually adjust the position of four points in the CrowdWater dataset and one point in the Expert dataset because they were not located in a stream. Afterwards, the flow accumulation (i.e., watershed areas) was calculated for each point with the Watershed function and the D8 method. Note that these watershed areas are uncertain due to uncertainties in e.g., the determination of the geomorphic stream network, anthropogenic alterations of the network, microtopography, etc. In addition to the area of the watershed draining to each point, we also calculated the mean slope for each watershed area in R with the raster package and the terrain function. Finally, we derived the Topographic Wetness Index (TWI; Beven and Kirkby, 1979) for each point in R by dividing the watershed area (A) by the slope (B) for that location:

$$TWI = \frac{\ln(A)}{\tan(B)} \tag{S2}$$

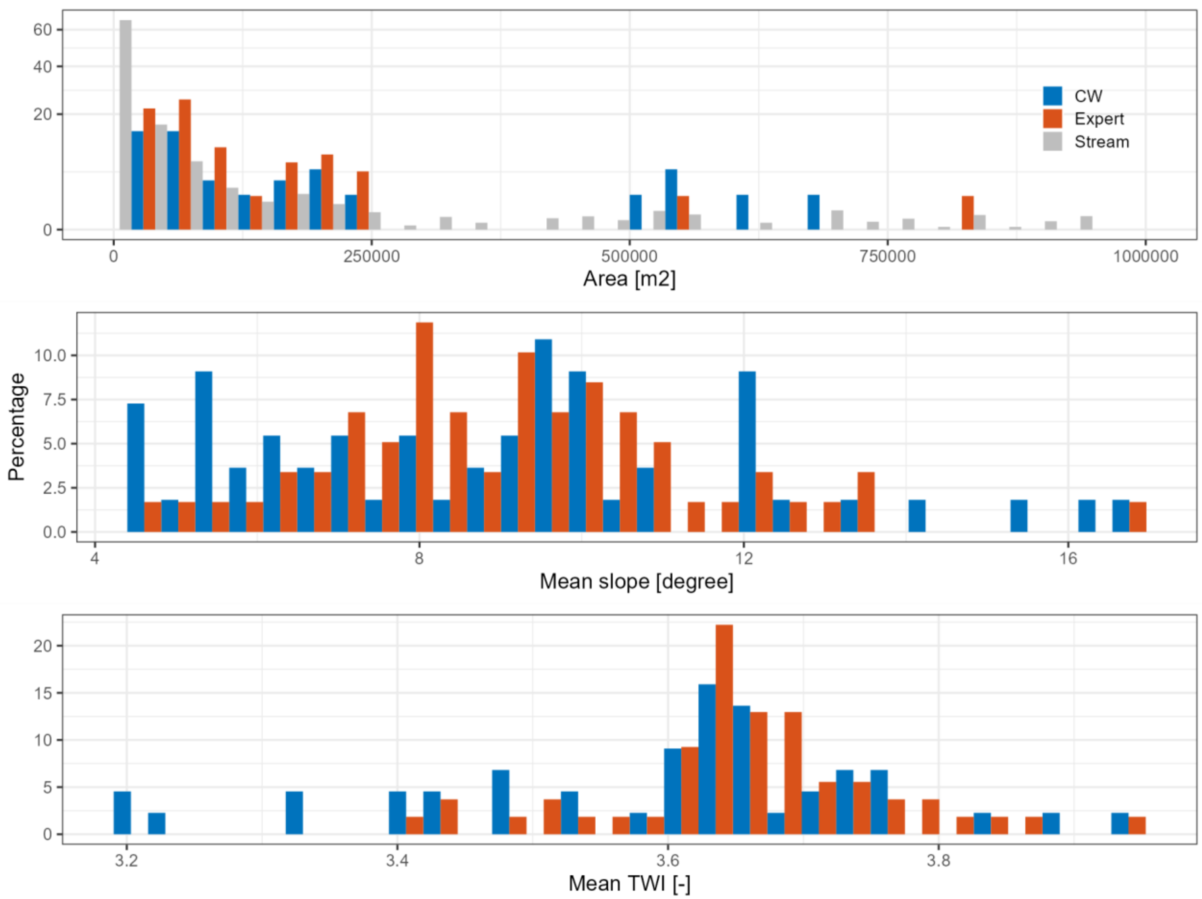


Figure S5. Frequency distributions of the drainage area (A in m^2, top), mean slope (middle) and mean value of the topographic wetness index (TWI) (bottom) for the area draining to each temporary stream sites for the CrowdWater dataset (blue) and the Expert dataset (red). The drainage areas for all points along the full stream network are shown for completeness in the top figure as well (grey). Note that the y-axis scale for the drainage area is squared to better visualize the variation for the small streams.

Supplementary material S2: Nodewise accuracies for random model

For the random model, we assume that the nodewise persistence (i.e., the fraction of time that a node is flowing) derived from the hierarchical model is correct but the timing of flowing and not flowing states varies randomly. The accuracy of a random model for node i ($A_{random,i}$) with persistence P_i can then be described by the following equation:

$$A_{random,i} = P_i^2 + (1 - P_i)^2 \tag{S3}$$

The accuracy of the hierarchical model was compared to the accuracy of the random model to interpret the nodewise accuracies. Note that the minimum accuracy for the random model is 0.5.

Supplementary material S3: Application of the skew-normal distribution

S3.1 Method

The relation between the fraction of flowing notes and the cumulative effective precipitation $P_{eff,T}$ was derived using the following equation:

$$F_{active} = \frac{psn(P_{eff,T}, x_i, \ \omega, \ \alpha) + a}{a+1} * b \quad \text{(S4)}$$

Where x_i is the location parameter, ω the scale parameter and α the slant parameter. The a and b parameters are used to adjust the range and the shape of the s-shaped function: parameter a is the minimum offset, and parameter b is the range of the offset. The five parameters were calibrated using the nls() function (stats package in R; Azzalini 2000). The parameter ranges were set to -50 to 250 for x_i, 50 to 500 for ω, -3 to 3 for α, 0.01 to 0.5 for a, and 0.75 to 1 for b.

S3.2 Additional results

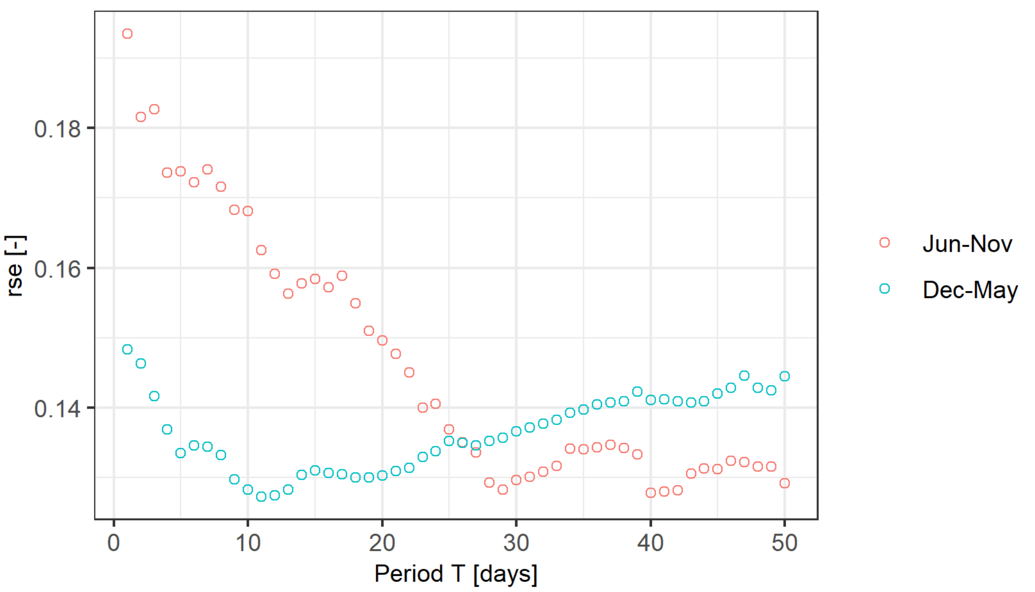


Figure S6. Residual standard errors (rse) for for the regression between the cumulative effective precipitation and fraction of flowing nodes (based on the skew-normal distribution) as a function of the period T over which the cumulative effective precipitation is calculated for the months Jun-Nov (red) and the months Dec-May (blue). The residual standard error was smallest when T is 40 days for the summer and fall and 11 days for the winter and spring.

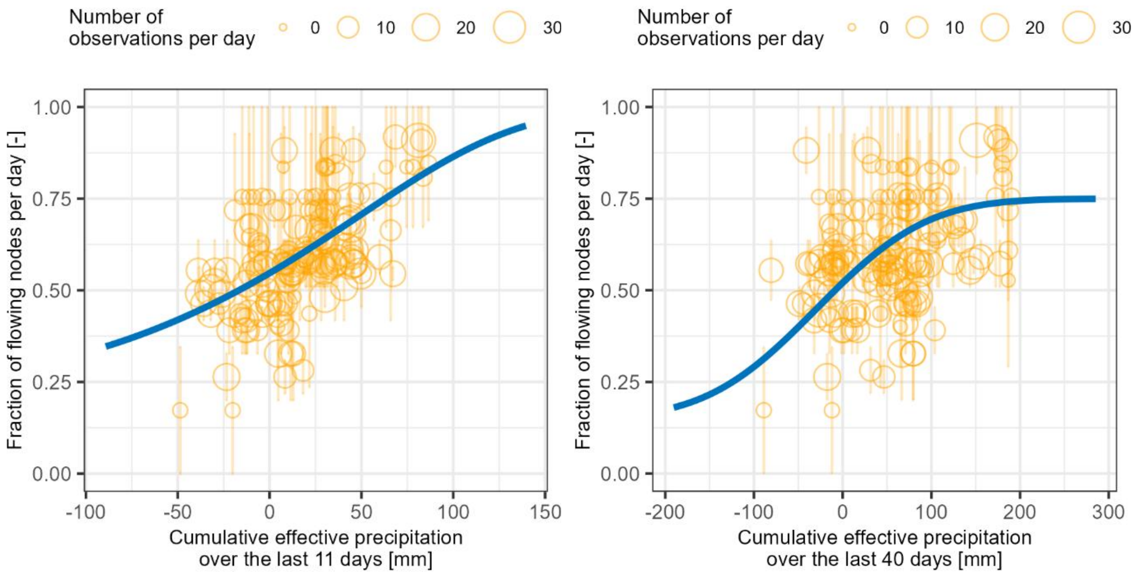
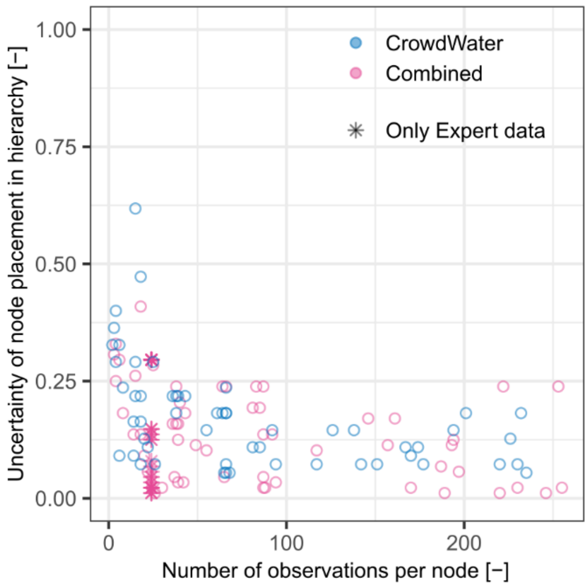


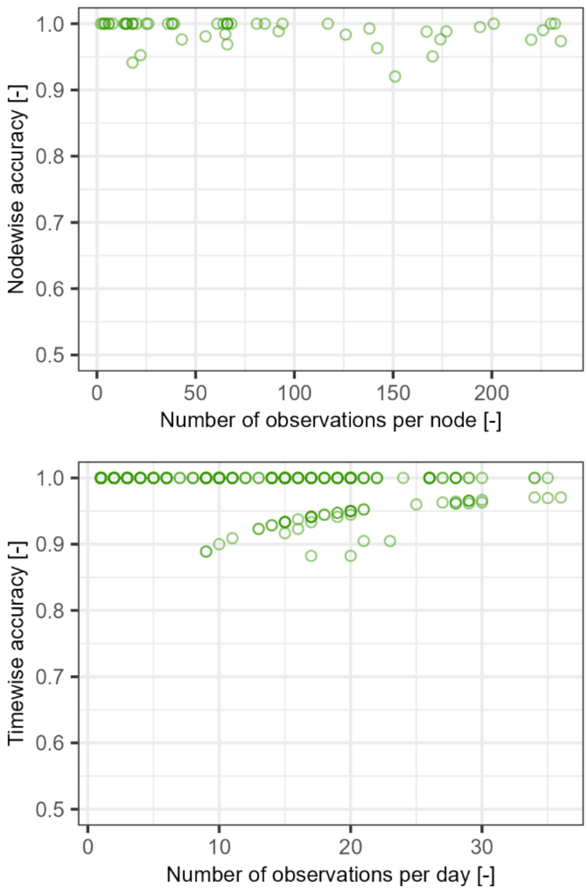
Figure S7. Fraction of flowing nodes per observation day between December and May plotted as a function of the cumulative effective precipitation calculated over a 11-day period (left) and 40-day period (right) for the CrowdWater dataset (yellow) and the best fitted model (skew-normal distribution; blue). The uncertainty of the fraction of flowing nodes is represented by the vertical lines. The size of the symbol represents the number of observations for that day.

Supplementary material S4: Additional results



125

Figure S8. Uncertainty of a node's placement in the hierarchical structure as a function of the number of observations for that node for the different datasets (CrowdWater, Expert and Combined). For the Combined dataset, the nodes that were only observed by the expert are indicated with a star.



130 **Figure S9. Nodewise (top) and timewise accuracy (bottom) for the hierarchical structure based on the CrowdWater dataset as a function of the number of observations for the node (top) or day (bottom). Note that many points plot on top of each other (represented by a darker green colour). The y-scale ranges from 0.5 to 1 to enable better visualization.**

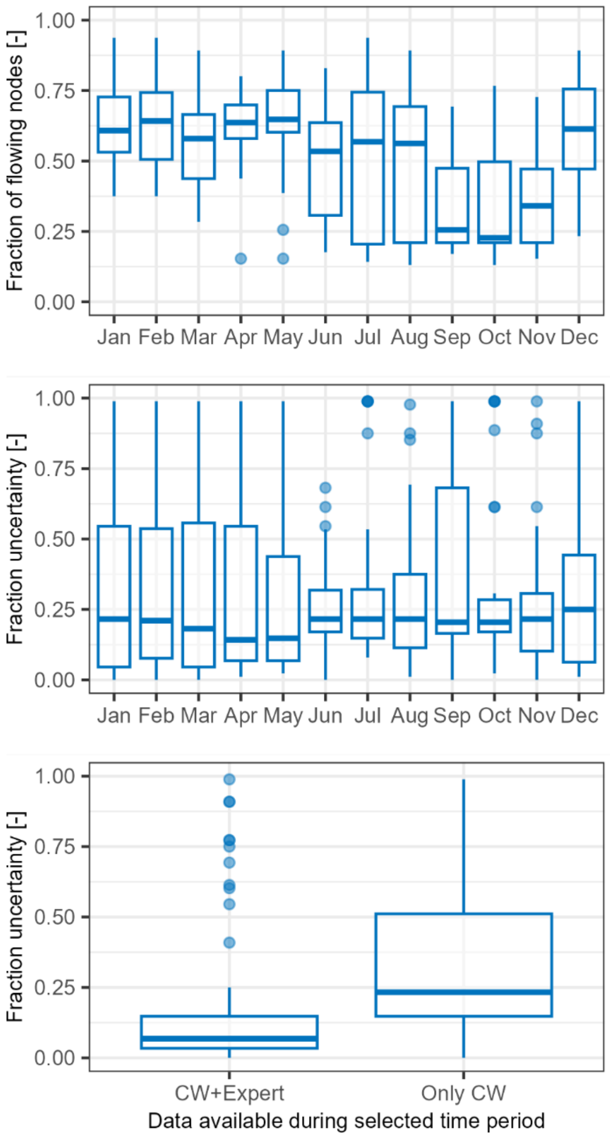


Figure S10. Boxplots of the estimated fraction of flowing nodes for each observation day (median values) (top) and uncertainty of this fraction per observation day (middle) for each month of the year for the Combined dataset for the entire observation period (Sep 2020 - Nov 2023), and, the uncertainty in the fraction of flowing nodes for the Combined dataset for the time periods for which expert data are available (Nov 2022-May2023) and the period for which only Citizen science data were available (bottom). The box represents the interquartile range, and the thick line the median. The whiskers extend to 1.5 times of the interquartile range.

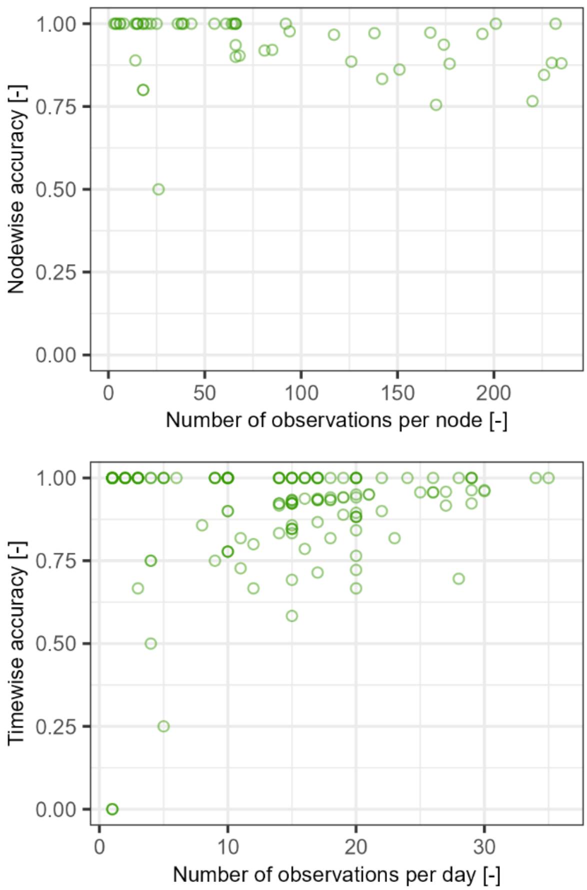
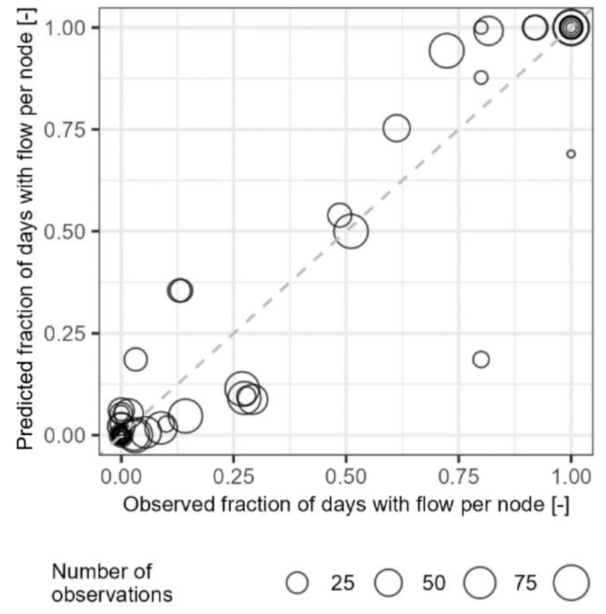


Figure S11. Nodewise (top) and timewise accuracy (bottom) for the predictions based on the regression between the cumulative net precipitation for a 40 day period and the fraction of nodes with flowing water, and the application of the hierarchical structure for the CrowdWater data as a function of the number of observations for each node (top) or day (bottom).



145 **Figure S12. Predicted fraction of days with flowing water for each node compared to the fraction of observations for which the site was observed to be flowing between June and November. The size of the symbol represents the number of observations for each site. The dashed line represents the 1 to 1 line. The Pearson correlation is 0.95.**

Table S2. Order of the nodes in each hierarchy according to the topological order of the directed graph for the respective dataset. Nodes at the top of the table correspond to those at the top of the hierarchy in Fig. 4.

CrowdWater	Expert	Combined
A04	A06	A04
A05	A05	A05
J07	A08	J07
J03	A04	G03
G13	D03	J03
G03	D04	G13
G14	F01	G14
A01	G02	A01
A02	G03	A02
G10	G06	G10
G06	G01	G06
A03	G12	J04
H04	H02	D04
E02	J01	A06
H11	J07	D03
J04	J02	F01
H05	J04	G02
D08	H03	G01
A08	H05	G12
D04	G11	H02
I01	H04	J01
H01	I01	J02
B01	E01	H03
F05	A13	G11
A13	B04	H04
A10	H01	H11
B03	A09	A08
C03	B05	H05
E03	D08	I01
F03	I03	E01
A12	G05	H01
H06	H10	B04
A07	B02	H10

I02	F02	A03
A11	E03	A13
F06	D01	A09
G07	A14	B05
H07	B03	I03
K04	F05	G05
K03	E02	E02
E04	D09	B01
E05	D05	D08
J06	H09	C03
J05	A07	E03
A16	B06	B02
D06	D07	F02
H08	F04	F05
C04	H07	D01
D02	D06	A10
G04	G08	B03
A15	C02	F03
K05	K02	A14
C01	G07	D09
D09	G09	A12
A17	J06	D05
	B07	H09
	D02	D07
	K01	H06
	H08	A07
		I02
		F06
		B06
		F04
		H07
		A11
		G08
		C02
		K02
		G07

	D06
	J06
	K04
	K03
	E04
	E05
	J05
	G09
	H08
	C04
	D02
	A16
	G04
	A15
	K05
	C01
	A17
	B07
	K01

References

Azzalini, Adelchi (2000): CRAN: Contributed Packages.

Beven, K. J.; Kirkby, M. J. (1979): A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d’appel variable de l’hydrologie du bassin versant. In *Hydrological Sciences Bulletin* 24 (1), pp. 43–69. DOI: 10.1080/02626667909491834.

OpenStreetMap contributors (2017): Planet dump retrieved from https://planet.osm.org.