nature physics

ARTICLES

https://doi.org/10.1038/s41567-019-0783-2

There are amendments to this paper

Universal gap scaling in percolation

Jingfang Fan[1*], Jun Meng[1], Yang Liu[1,2], Abbas Ali Saberi[3,4], Jürgen Kurths[1,5,6] and Jan Nagler[7*]

Universality is a principle that fundamentally underlies many critical phenomena, ranging from epidemic spreading to the emergence or breakdown of global connectivity in networks. Percolation, the transition to global connectedness on gradual addition of links, may exhibit substantial gaps in the size of the largest connected network component. We uncover that the largest gap statistics is governed by extreme-value theory. This allows us to unify continuous and discontinuous percolation by virtue of universal critical scaling functions, obtained from normal and extreme-value statistics. Specifically, we show that the universal scaling function of the size of the largest gap is given by the extreme-value Gumbel distribution. This links extreme-value statistics to universality and criticality in percolation.

Critical phenomena are associated with physical and socioeconomic systems that undergo a phase transition at a critical point, such as the critical temperature, pressure or population density[1]. A phase transition is characterized by scaling functions that govern the finite-size behaviours and a set of critical exponents that determine the system's universality class[2]. The concept of finite-size scaling provides a versatile tool to study critical phenomena from both theoretical[3] and experimental[4] perspectives. A prime example of critical dynamics is percolation—the transition from small-scale connectivity to extensive connectedness. Percolation fundamentally underlies the function, structure and robustness of networks[5]. Thus, the theory of percolation provides a theoretical foundation on which to assess the statistical characteristics of many complex systems[1,6–8]. For lattices, percolation denotes the spontaneous emergence of an interconnecting path from one side to the other. In networks, percolation is characterized by the emergence of a giant cluster of size $\mathcal{O}(N)$, announcing a long-range connectivity, with N being the total number of nodes in the network[9].

Universality in percolation has been studied extensively over the past decade[10–12]. However, previous studies could not utilize a universal framework of continuous and discontinuous percolation by virtue of finite-size scaling functions. Our theoretical framework developed here fills this gap.

Instead of performing a finite-size scaling analysis at the critical phase transition point, r_c, we base the finite-size scaling analysis on the statistics and critical scaling of the size of the largest gap in the order parameter. Since the size of the largest gap can be regarded as an extreme value of the percolation process, the corresponding scaling function is governed by extreme-value statistics. This trick allows us to develop a surprisingly simple but universal scaling approach.

Theory

In percolation, system properties are studied as a function of the control parameter. We denote the control parameter by r, the link density. This choice makes sense for bond and network percolation and is used throughout this paper, except for site percolation, where r denotes the site occupation density. Starting with a network with N isolated nodes, r is increased gradually by single links that are added randomly one by one[13]. The order parameter is the size of the largest cluster, given by the largest connected component in the entire system. During the evolution of the system, the order parameter is monitored, $S(T)$ at time step T, and its relative largest one-step gap Δ is calculated:

$$\Delta \equiv \frac{1}{N}\max_T(S(T+1)-S(T)) \tag{1}$$

Following ref. [14], we define the step with the largest jump as T_c and the reduced transition point as $r_c = T_c/N$, where N stands for the total number of nodes (in bond percolation on the lattice, $r_c = T_c/M$, with M denoting the maximal number of bonds). The percolation strength is defined as the size (number of sites or nodes) of the largest cluster at T_c, that is, $S_c = S(T_c)$. To study the statistics of growth and fluctuations, we run 10^4 realizations for a given system size. Each independent run results in one value of Δ, r_c and S_c. The finite-size scaling behaviours are studied for d-dimensional lattices as a function of the linear dimension L, where $N = L^d$, which is replaced by N in network percolation (Supplementary Information).

Gap exponents. Critical phenomena such as percolation exhibit scale-free behaviours that are quantified by scaling relations. Likewise, the averages $\overline{\Delta}$, $\overline{r}_c$ and $\overline{S}_c$ are anticipated to exhibit the following scaling relations, as a function of L

$$\overline{\Delta}(L) \sim L^{-\beta_1} \tag{2}$$

$$\overline{r}_c(L) - r_c(\infty) \sim L^{-1/\nu_1} \tag{3}$$

$$\overline{S}_c(L) \sim L^{d_{f_1}} \tag{4}$$

where β_1, ν_1 and d_{f_1} are three critical exponents characterizing the universality class of the percolation problem, and $r_c(\infty)$ is the percolation threshold in the thermodynamic limit. We also investigate the fluctuations $\delta\Delta = \Delta - \overline{\Delta}(L)$, $\delta r_c = r_c - \overline{r}_c(L)$ and $\delta S_c = S_c - \overline{S}_c(L)$, whose standard deviations are defined by

$$\chi_\Delta = \sqrt{\langle[\delta\Delta]^2\rangle} \tag{5}$$

[1]Potsdam Institute for Climate Impact Research, Potsdam, Germany. [2]Department of Computer Science, Technische Universität Berlin, Berlin, Germany. [3]Department of Physics, University of Tehran, Tehran, Iran. [4]Institut für Theoretische Physik, Universitat zu Köln, Köln, Germany. [5]Department of Physics, Humboldt University, Berlin, Germany. [6]Saratov State University, Saratov, Russia. [7]Deep Dynamics Group, Frankfurt School of Finance and Management, Frankfurt, Germany. *e-mail: jingfang@pik-potsdam.de; j.nagler@fs.de

$$\chi_{r_c} = \sqrt{\langle [\delta r_c]^2 \rangle} \tag{6}$$

$$\chi_{S_c} = \sqrt{\langle [\delta S_c]^2 \rangle} \tag{7}$$

with the following proposed scaling behaviours

$$\chi_\Delta \sim L^{-\beta_2} \tag{8}$$

$$\chi_{r_c} \sim L^{-1/\nu_2} \tag{9}$$

$$\chi_{S_c} \sim L^{d_{f_2}} \tag{10}$$

where β_2, ν_2 and d_{f_2} comprise another set of critical exponents. In general, different percolation models with the same critical exponents belong to the same universality class.

In the past decade, special attention has been paid to β_1 (ref. [14]). This is because $\beta_1 = 0$ immediately implies a discontinuous percolation transition, where models in the domain $0 < \beta_1 < 1$ can be continuous or discontinuous[15]. However, typically, $\beta_1 > 0$ implies continuity of the phase transition, although values close to zero generate 'explosive percolation', characterized by substantial gaps in the order parameter for large finite systems[10,16].

Universal scaling function. The distributions of the fluctuations $\delta\Delta$, δr_c and δS_c are expected to take the following finite-size scaling forms[2,17]

$$P_\Delta(\delta\Delta, L) = L^{\beta_2} f_\Delta(\delta\Delta \times L^{\beta_2}) \tag{11}$$

$$P_r(\delta r_c, L) = L^{1/\nu_2} f_r(\delta r_c \times L^{1/\nu_2}) \tag{12}$$

$$P_S(\delta S_c, L) = L^{-d_{f_2}} f_S(\delta S_c \times L^{-d_{f_2}}) \tag{13}$$

where $f_\Delta(\cdot)$, $f_r(\cdot)$ and $f_S(\cdot)$ are three universal scaling functions. The values of the critical exponents β_2, ν_2 and d_{f_2} can be obtained from equations (8)–(10) numerically.

Thus far, we have introduced six critical (henceforth referred to as) gap exponents. It is easy to see and not surprising that they are neither independent nor totally distinct from the standard percolation exponents. In the Supplementary Information, we show that they partially coincide, namely, $\beta_1 = \beta_2$ and $d_{f_1} = d_{f_2}$. In addition, we derive relationships between the gap exponents and the standard percolation critical exponents as

$$\beta_1 = \beta_2 = \beta/\nu,\ \nu_2 = \nu, d_{f_1} = d_{f_2} = d_f \tag{14}$$

where β is the critical exponent of the order parameter $P_\infty \sim |r - r_c|^\beta$, $\nu = \nu_2$ describes the divergence of the correlation length $\xi \sim |r - r_c|^{-\nu}$ and d_f is the fractal dimensions in standard percolation[5], but in general, $\nu_1 \neq \nu_2$. In contrast to ν_2, the exponent ν_1 can be used to determine the percolation threshold in the thermodynamic limit. From the scaling law, $\beta/\nu = d - d_f$ of standard percolation theory[5], a similar law follows for the gap exponents, $\beta_1 = d - d_{f_1}$ (for more details, see Supplementary Information). Please note that the gap critical exponents proposed here are different from the so-called gap exponent Δ (ref. [18]).

In the following, we postulate the universal scaling functions, $f_\Delta(\cdot)$, $f_r(\cdot)$ and $f_S(\cdot)$. We regard the size of the largest gap of the order parameter as the maximum of random variables drawn from independent realizations of the entire percolation process. This stands in contrast to conventional critical finite-size scaling, where the critical percolation threshold, r_c, is obtained from independent realizations, and the scaling analysis is performed relative to this averaged observable[13,19–21].

For linking our framework to extreme-value theory, we make two observations. For a finite system of size N, close to r_c, the distribution of cluster sizes $n(s)$ typically exhibits a power law for cluster sizes up to a certain cut-off, related to the correlation length ξ. For some models of continuous percolation, the cluster size distribution is given by $n(s) \sim s^{-\tau} e^{-s/s_\xi}$ (refs. [5,9,22]; τ is the Fisher exponent), whereas, in general, the distribution may take a more complicated form[16]. In either case, for clusters much larger than this cut-off, that is for $s \gg s_\xi$, the distribution $n(s)$ decays rapidly. The size of the largest gap is exactly given by the size of the largest second largest cluster in the entire range of r (ref. [14]). Since the second largest cluster is, however, much larger than the cut-off size s_ξ, the largest gap size is necessarily characterized by the light-tailed part of the critical cluster size distribution. (1) This means that the distribution of largest gaps becomes necessarily sharp for large systems. (2) The largest gaps are maxima obtained from independent ensemble realizations. Thus, extreme-value theory[1,23] predicts equation (11) to follow the Gumbel distribution (even for non-convergent (non-self-averaging) percolation processes, where the order parameter does not converge to a non-random function, the gap distribution must take well-defined values with exponential-type tail behaviour on [0, 1], see refs. [15,24]) and the universal function in equation (11) can be expressed as

$$f_\Delta(\delta\Delta \times L^{\beta_2}) = f_0 + A e^{-e^{-z}-z},\ \text{with}\ z = (\delta\Delta \times L^{\beta_2} - B)/\omega \tag{15}$$

where f_0, A, B and ω are four constants and β_2 can be obtained numerically by fitting the data for χ_Δ according to equation (5).

In contrast, r_c is a non-extremal variable, drawn and averaged from independent realizations, and thus the central limit theorem suggests a Gaussian distribution. Hence, we expect that the universal function in equation (12) can be written as

$$f_r(\delta r_c \times L^{1/\nu_2}) = f_1 + A_1 e^{-\frac{(z-z_c)^2}{2\omega_1^2}},\ \text{with}\ z = \delta r_c \times L^{1/\nu_2} \tag{16}$$

where f_1 and A_1 are constants, z_c is the mean value and ω_1^2 is the variance. Note that ν_2 can also be obtained from equation (6).

The percolation strength and largest gap are related through

$$S_c = N\Delta + o(N) \tag{17}$$

and become equal in the thermodynamic limit[14]. Hence, for $N \gg 1$, we predict the distribution of δS_c to follow the Gumbel extreme-value distribution (equation (15)), although with different critical exponents

$$f_S(\delta S_c \times L^{-d_{f_2}}) = f_2 + A_2 e^{-e^{-z}-z},\ \text{with}\ z = (\delta S_c \times L^{-d_{f_2}} - B_2)/\omega_2 \tag{18}$$

where f_2, A_2, B_2 and ω_2 are constants.

Renormalization group theory for critical phenomena[25] dictates that universal scaling functions must be homogeneous functions. Therefore, we assume the additive constants in equations (15), (16) and (18) to be zero, $f_0 = f_1 = f_2 = 0$. In addition, since the parameters B (equation (15)) and B_2 (equation (18)) are the modes of unimodal distributions, we can determine them without fitting. This is relevant for determining the amplitude A and the scale parameter ω in equation (15). In particular, if the Gumbel distribution is normalized, that is, $\int f_\Delta(\delta\Delta \times L^{\beta_2}) d(\delta\Delta \times L^{\beta_2}) = 1$, then we have $A = 1/\omega$ and

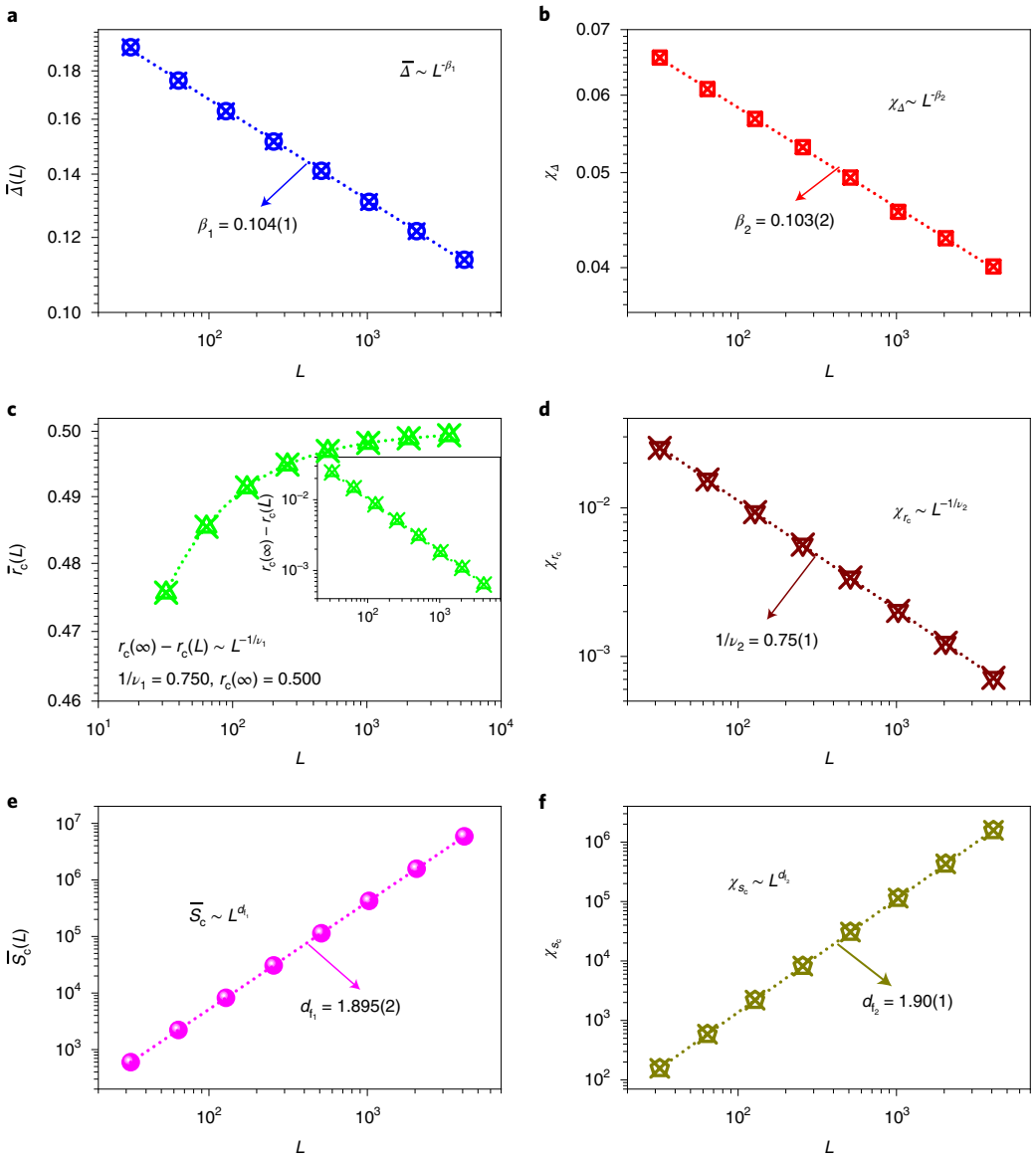

Fig. 1 | Critical scaling for 2D percolation. Fits of critical exponents for bond percolation on a 2D square lattice. **a**,**b**, The average size of the largest gap $\overline{\Delta}(L)$ (**a**) and its fluctuation χ_Δ (**b**) as functions of L. **c**,**d**, Log–log plot of the percolation threshold $\overline{r}_c(L)$ (**c**) and its fluctuation χ_{r_c} (**d**) versus L. **e**,**f**, The power-law relations of $\overline{S}_c(L)$ (**e**) and χ_{S_c} (**f**) with L. The inset in **c** shows the log–log plot of $r_c(\infty) - \overline{r}_c(L)$ as a function of L. The dashed lines in each figure are the best-fit lines for the data with $R^2 > 0.99$ and error bars refer to the standard error. Numbers in parentheses indicate the uncertainty in the final digits. Results obtained from 10,000 independent realizations.

$A_2 = 1/\omega_2$, which requires only a one-parameter fit of the scales (see Supplementary Information for details).

To verify the proposed theoretical framework, we apply it to a number of continuous percolation processes and underlying structure, that is, two- and three-dimensional lattices, Erdős–Rényi (ER) networks, random regular (RR) graphs and explosive percolation (EP)[10]. In addition, we demonstrate its applicability for models with a wide range of critical behaviours, namely, discontinuous cluster aggregation (DCA)[26,27], the restricted Erdős–Rényi (r-ER) model[28] and fractional Brownian motion (fBm) surfaces with a long-range correlated topography[29,30].

Numerical results

Continuous percolation. We first analyse prototypical bond percolation on an $L \times L$ lattice with periodic boundary conditions. Bonds are added one by one randomly, until all initially open bonds have been closed. In Fig. 1a,b, we study $\overline{\Delta}$ and χ_Δ as a function of L, and find that $\beta_1 = \beta_2 \approx 0.104$. The exponents are in agreement with equation (14) and agree well with the ratio β/ν known from percolation theory, predicting $\beta = 5/36$ and $\nu = 4/3$ (ref. [31]). The percolation threshold $\overline{p}_c$ and its fluctuation χ_p are shown in Fig. 1c,d. We obtain $1/\nu_1 = 1/\nu_2 \approx 1/\nu$, and $r_c(\infty) \approx 0.5$, which are in agreement with the known values[32]. In addition, we find that $d_{f_1} = d_{f_2} \approx 1.895$, which is close to the fractal dimension $d_f = 91/48$ of standard percolation[31] (see $\overline{S}_c$ and its fluctuation χ_{S_c} plotted as a function of L in Fig. 1e,f).

The distributions of the fluctuations $\delta\Delta$, δr_c and δS_c as a function of L are studied in Fig. 2. We find a pronounced finite-size dependence. Combining the proposed finite-size scaling forms, equations (11)–(13), with the gap scaling exponents as shown in Fig. 1, yields collapses onto the respective universal scaling function (Fig. 2). In particular, we find that $f_\Delta(\cdot)$ is well described by the Gumbel function, equation (15) (see the solid green line in Fig. 2d, with

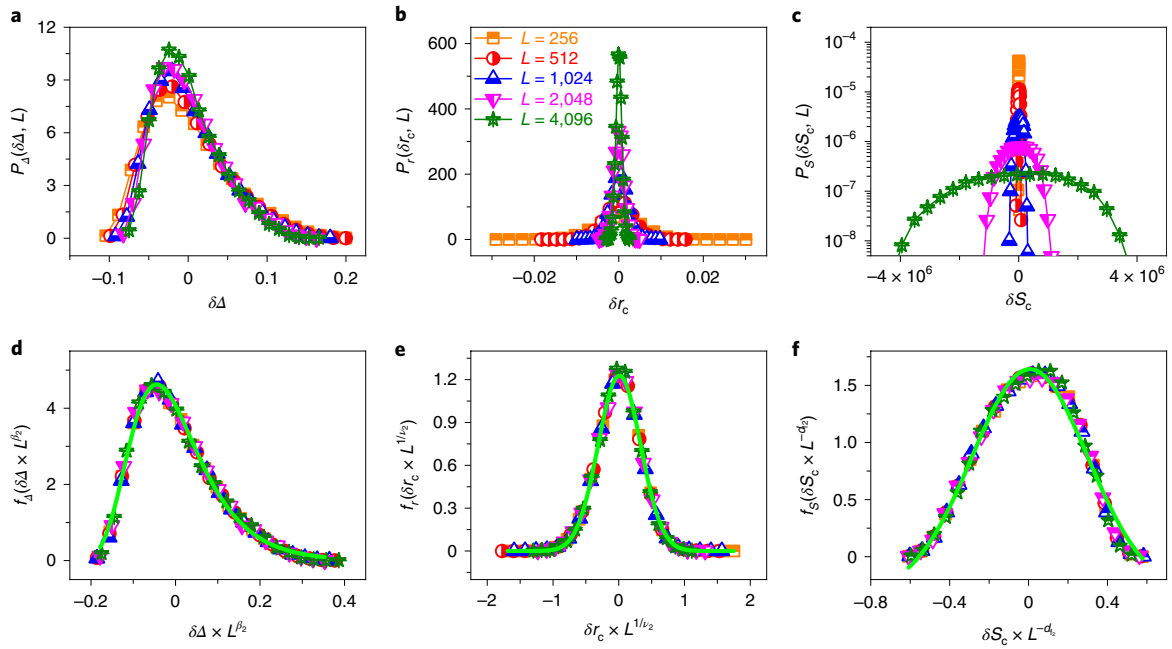

Fig. 2 | Universal gap scaling functions for 2D percolation. Universal scaling functions for bond percolation on a 2D square lattice. **a**–**c**, The distributions of the fluctuations $\delta\Delta$ (**a**), δr_c (**b**) and δS_c (**c**) with different L, respectively. **d**–**f**, The corresponding universal scaling functions for $\delta\Delta$ (**d**), δr_c (**e**) and δS_c (**f**). The green solid line in **d** is obtained from the Gumbel distribution, equation (15) with $R^2 > 0.99$; in **e** is obtained from the Gaussian function, equation (16) with $R^2 > 0.99$; in **f** is obtained from equation (18) with $R^2 > 0.90$. The values of critical exponents, β_2, ν_2 and d_{f_2} are from Fig. 1. Results obtained from 10,000 independent realizations.

Table 1 | Table of critical exponents

Model	β_1	$1/\nu_1$	d_{f_1}	β_2	$1/\nu_2$	d_{f_2}	β	ν	d_f	Type of PT
2D lattice	0.104(1)	0.75(1)	1.895(2)	0.103(2)	0.75(1)	1.90(1)	5/36	4/3	91/48	Continuous
3D lattice	0.47(1)	1.14(1)	2.53(2)	0.47(1)	1.12(4)	2.53(1)	0.417(3)	0.875(8)	2.52(2)	Continuous
ER	0.333(5)	0.33(1)	0.67(1)	0.33(1)	0.33(1)	0.67(1)	1	3	2/3	Continuous
RR	0.333(5)	0.33(1)	0.67(1)	0.33(1)	0.33(1)	0.67(1)	1	3	2/3	Continuous
EP	0.065(1)	0.75(1)	0.935(1)	0.064(1)	0.500(1)	0.935(1)	0.086(1)	2	–	Continuous
DCA	0.0	0.75(1)	1.0(1)	0.0	0.50(1)	1.0(1)	0	2	–	Discontinuous
r-ER	0.0	0.75(1)	1.0(1)	0.0	0.50(1)	1.0(1)	0	2	–	Discontinuous
fBm	0.0(1)	0.0(2)	2.0(1)	0.0(1)	0.0(2)	2.0(1)	0	–	–	Discontinuous

The numerical values of the standard percolation critical exponents (or exact ones, if known) for β, ν and d_f for the 2D lattice model are taken from ref. [31]; for the 3D lattice from ref. [42]; for the ER network from ref. [43]; for the RR graph from ref. [44]; for EP from ref. [12]; for the DCA model from refs. [26,27]; for the r-ER model from ref. [28]; and for the 2D fBm surfaces with Hurst exponent H from refs. [29,30]. Type of percolation transition (PT) is specified for each case. Note that the value of ν and d_f in network systems contains the upper critical dimension d_u, that is, $\nu = \nu' d_u$ and $d_f = d_f^* / d_u$, where ν' is the critical exponent for the correlation length and d_f^* is the fractal dimension. For instance, the ER (RR) process is a mean-field case of percolation with $\beta = 1$, $\nu' = 1/2$, $d_f^* = 4$ and $d_u = 6$ (refs. [45,33]). We thus obtain $\nu = 3$ and $d_f = 2/3$. Numbers in parentheses indicate the uncertainty in the final digits.

$R^2 > 0.99$), $f_r(\cdot)$ is well fitted by the Gaussian function, equation (16) (see the solid green line in Fig. 2e, with $R^2 > 0.99$] and $f_S(\cdot)$ is also described by the Gumbel function, equation (18) (see the solid green line in Fig. 2f, with $R^2 > 0.90$), yet with larger deviations compared with Fig. 2d,e. The deviations may result from finite-size effects from the relation equation (17). In particular, we expect S_c to become an independent extremal variable only in the thermodynamic limit.

Overall, the universality class of two-dimensional (2D) percolation is well characterized by the gap exponents and the proposed universal scaling functions.

In the Supplementary Information, we apply our framework to a 3D lattice, random network ER percolation, RR graphs and EP model. The critical exponents and fitting values for each model are summarized in Table 1 and Supplementary Table 1, whereas universal scaling is studied as shown in Supplementary Figs. 1–8. The performed scaling analysis and the goodness-of-fit results as shown in Table 2 are an excellent support of our framework.

In short, the simulation results suggest that the proposed scaling theory based on extreme-value statistics provides a firm theoretical foundation for universal scaling in continuous percolation.

Discontinuous percolation. Traditionally, a network or lattice is expected to undergo a continuous percolation phase transition during a gradual random occupation process[33]. Explosive (yet continuous) and genuinely discontinuous percolation for not entirely random but competitive link-addition processes has attracted much attention in recent years[10,11,14,15,26,34,35].

Can our framework be successfully applied to percolation models showing an unusual finite-size scaling and to models of

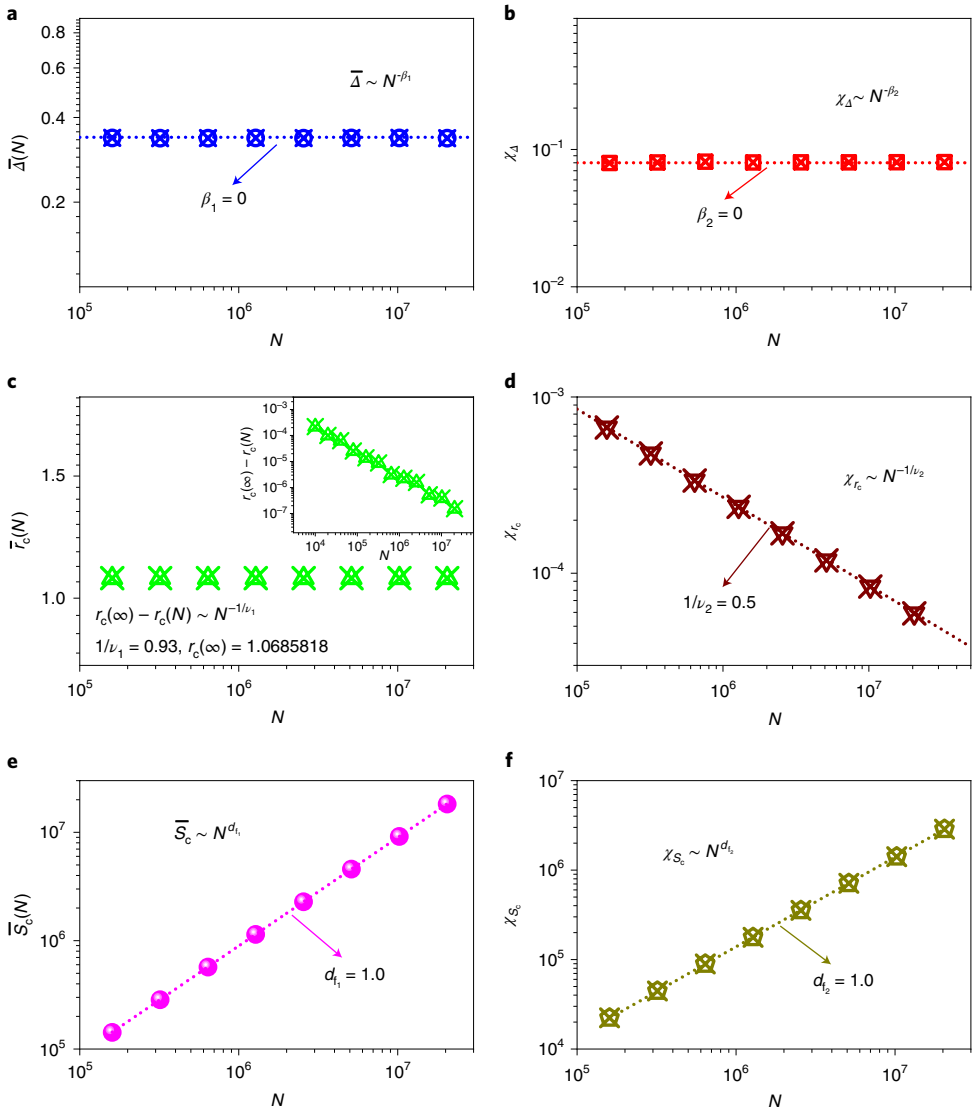

Fig. 3 | Critical scaling for DCA. a–f, Analogous to Fig. 1a–f, critical exponents for DCA, for which $\nu_1 \neq \nu_2$. The dashed lines in each figure are the best-fitting lines for the data with $R^2 > 0.99$.

genuinely discontinuous percolation? To test this, first we perform our scaling analysis for a model of DCA[26].

The evolution of the cluster size distribution is traditionally described by the Smoluchowski coagulation equation[36]

$$\frac{\mathrm{d}n_k}{\mathrm{d}t} = \sum_{i+j=k} K_{ij} n_i n_j - 2n_k \sum_j K_{kj} n_j \qquad (19)$$

where n_k is the density of k-size clusters and K_{ij} is called the collision kernel that accounts for the adhesion of two clusters[37]. Here we study the size-independent kernel $K_{ij} = 1$, which coincides with Smoluchowski's original choice but also with a particular case of the network model proposed and studied in refs. [26,27], except that we allow for links in the same cluster. This model leads to a maximal delay of the discontinuous transition to global connectivity, yet featuring non-trivial finite-size behaviours[26,27]. For more details of the algorithm, see Supplementary Information.

The critical exponents, $\beta_1 = \beta_2 = 0$, and $d_{f_1} = d_{f_2} = 1$, as shown in Fig. 3, confirm the genuinely discontinuous percolation featured by the DCA model. We find that the critical exponent $\nu_1 \neq \nu_2$; in particular, $1/\nu_2 = 0.5$, which, according to ref. [12], may describe the divergence of the correlation length for a large class, including EP.

Table 2 | Goodness of fit and significance

Model	$f_\Delta(\cdot)$	$f_r(\cdot)$	$f_S(\cdot)$
2D lattice	$R^2 > 0.99$	$R^2 > 0.99$	$R^2 > 0.90$
3D lattice	$R^2 > 0.99$	$R^2 > 0.99$	$R^2 > 0.97$
ER	$R^2 > 0.99$	$R^2 > 0.99$	$R^2 > 0.99$
RR	$R^2 > 0.99$	$R^2 > 0.99$	$R^2 > 0.99$
EP	$R^2 > 0.98$	$R^2 > 0.99$	$R^2 > 0.88$
DCA	$R^2 > 0.98$	$R^2 > 0.99$	$R^2 > 0.98$
r-ER	$R^2 > 0.95$	$R^2 > 0.99$	$R^2 > 0.88$
fBm	$R^2 > 0.96$	$R^2 > 0.95$	$R^2 > 0.96$

R^2 values for the best-fitting curves of the universal scaling functions. $f_\Delta(\cdot)$ is the Gumbel function, with parameters of equation (15), $f_r(\cdot)$ is a Gaussian, equation (16), and $f_S(\cdot)$ is a Gumbel function, with parameters of equation (18).

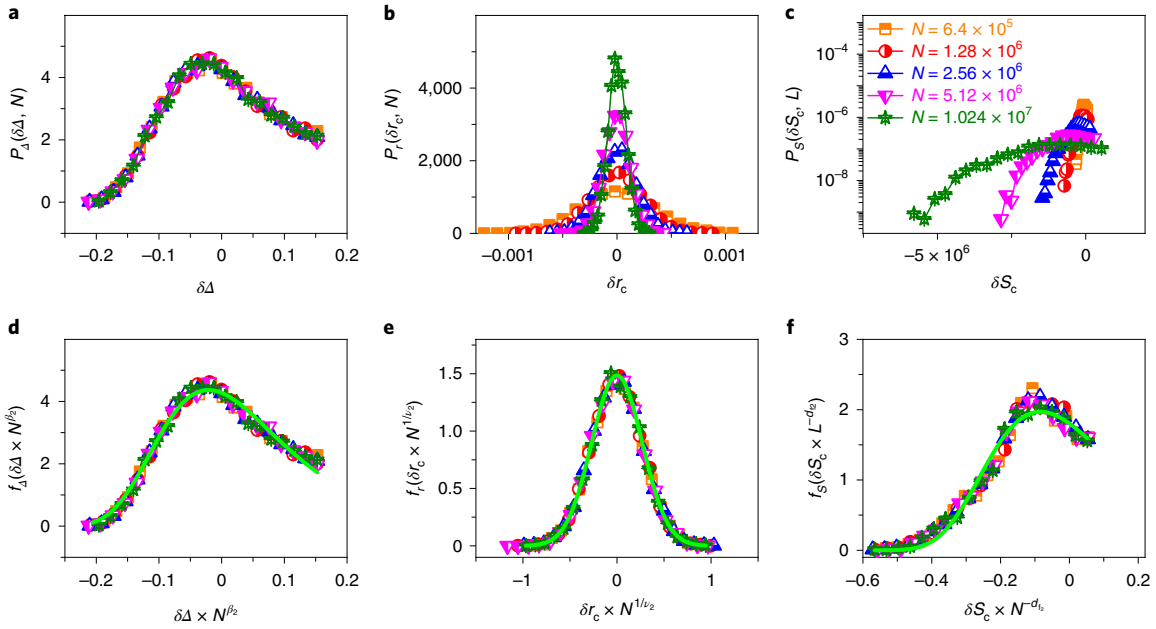

Fig. 4 | Universal gap scaling functions for DCA. a–f, Analogous to Fig. 2a–f, universal scaling functions for DCA. Note that **a** and **d** are the same, since $\beta_2 = 0$. R^2 values for the best-fitting curves of the universal scaling functions are summarized in Table 2.

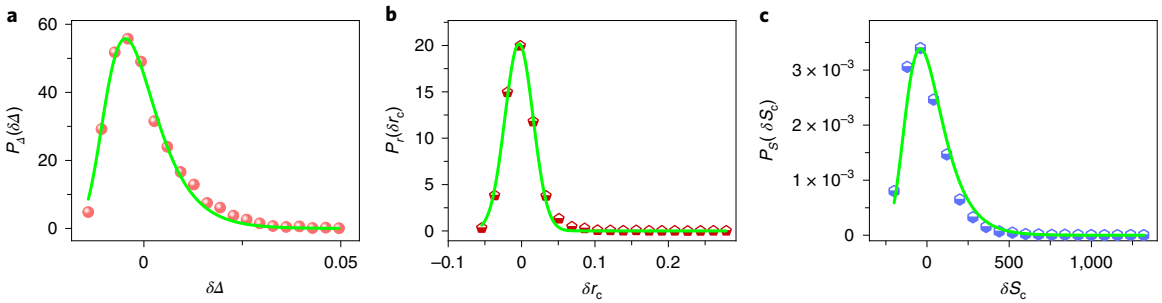

Fig. 5 | Gap scaling functions for human protein–protein interactions network. a–c, Analogous to Fig. 2a–c, the distributions of the fluctuation $\delta\Delta$, δr_c and δS_c. The green solid line in **a** is a Gumbel distribution fit with $R^2 > 0.98$; in **b** fit of Gaussian with $R^2 > 0.99$; in **c** Gumbel distribution fit with $R^2 > 0.96$. Results obtained for 10,000 independent realizations of a random link removal process.

For the DCA model, as shown in Fig. 4, we find that the distributions of $\delta\Delta$, δr_c, δS_c also collapse onto universal scaling functions of Gumbel and Gaussian type as for continuous percolation. In particular, Fig. 4 suggests that the distributions of $\delta\Delta$, for both Fig. 4a,d, are size independent, which results from $\beta_2 = 0$.

We also perform the scaling analysis for two other percolation models, the r-ER model that features a hybrid continuous–discontinuous transition[28,38] and 2D fractional Brownian motion surfaces with long-range correlated topography[29,30,39]. Details are presented in Supplementary Figs. 9–12.

In addition, the robustness of the numerical results regarding binning, fitting and collapse of cumulative distribution functions is demonstrated in the Supplementary Information.

Taken together, the results consistently demonstrate that our proposed scaling theory based on extreme-value statistics provides a unifying framework for continuous and discontinuous percolation.

Application to real networks. Many real networks are only available in one size. However, for a given stochastic network percolation process, our framework predicts the largest gap statistics to be of Gumbel type. Thus, we investigate the gap distribution of the size of the largest connected component of network instances induced by repeated random-link pruning of the original network. This also allows us to study the percolation threshold r_c and the largest cluster size S_c of the percolation process. Specifically, we study (1) the human protein–protein interactions network, (2) the DBLP collaboration network and (3) the Wikipedia Talk network. As shown in Figs. 5 and Supplementary Figs. 19 and 20, we find the respective distributions in excellent agreement with our theory.

Discussion

Conventional percolation theory could not provide a consistent framework that unifies continuous, discontinuous, explosive and hybrid percolation. Our theoretical framework unifies these domains by means of universal scaling functions based on finite-size scaling and extreme-value statistics regarding the size of the largest one-step gap in the order parameter.

The connection to traditional finite-size scaling is given by the relation of the proposed six critical exponents to the traditional ones: whereas $\beta_1 = \beta_2$, ν_2 and $d_{f_1} = d_{f_2}$ serve as the analogues of β/ν

(ratio of percolation strength and correlation length exponent), ν (correlation exponent) and d_f (fractal dimension), respectively, we also find that, in general, $\nu_1 \neq \nu_2$, except for random network percolation.

Phase transitions underlie critical phenomena in general. The universality of extreme values may be successfully applied in other critical phenomena as well, which may lead to a deeper understanding. This perspective may be helpful, in particular, for critical systems that lack a theoretical foundation, yet show some degree of universality. Our framework may help develop powerful tools for a universal formulation of various relevant problems in related areas ranging from synchronization[40] and epidemic spreading[41] to other critical phenomena[1,16].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-019-0783-2.

Received: 1 July 2019; Accepted: 17 December 2019;
Published online: 10 February 2020

References

1. Sornette, D. *Critical Phenomena in Natural Sciences* 2nd edn (Springer-Verlag, 2006).
2. Privman, V. & Fisher, M. E. Universal critical amplitudes in finite-size scaling. *Phys. Rev. B* **30**, 322–327 (1984).
3. Privman, V. *Finite Size Scaling and Numerical Simulation of Statistical Systems* (World Scientific Singapore, 1990).
4. Rhee, I., Gasparini, F. M. & Bishop, D. J. Finite-size scaling of the superfluid density of ^{4}He confined between silicon wafers. *Phys. Rev. Lett.* **63**, 410–413 (1989).
5. Stauffer, D. & Aharony, A. *Introduction to Percolation Theory* (Taylor and Francis, 2003).
6. Saberi, A. A. Recent advances in percolation theory and its applications. *Phys. Rep.* **578**, 1–32 (2015).
7. Meng, J., Fan, J., Ashkenazy, Y. & Havlin, S. Percolation framework to describe El Nino conditions. *Chaos* **27**, 035807 (2017).
8. Schröder, M., Nagler, J., Timme, M. & Witthaut, D. Hysteretic percolation from locally optimal individual decisions. *Phys. Rev. Lett.* **120**, 248302 (2018).
9. Cohen, R. & Havlin, S. *Complex Networks: Structure, Robustness and Function* (Cambridge Univ. Press, 2010).
10. Achlioptas, D., D'souza, R. M. & Spencer, J. Explosive percolation in random networks. *Science* **323**, 1453–1455 (2009).
11. da Costa, R. A., Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Explosive percolation transition is actually continuous. *Phys. Rev. Lett.* **105**, 255701 (2010).
12. Grassberger, P., Christensen, C., Bizhani, G., Son, S.-W. & Paczuski, M. Explosive percolation is continuous, but with unusual finite size behavior. *Phys. Rev. Lett.* **106**, 225701 (2011).
13. Newman, M. E. J. & Ziff, R. M. Fast Monte Carlo algorithm for site or bond percolation. *Phys. Rev. E* **64**, 016706 (2001).
14. Nagler, J., Levina, A. & Timme, M. Impact of single links in competitive percolation. *Nat. Phys.* **7**, 265–270 (2011).
15. D'Souza, R. M. & Nagler, J. Anomalous critical and supercritical phenomena in explosive percolation. *Nat. Phys.* **11**, 531–538 (2015).
16. D'Souza, R. M., Gómez-Gardeñes, J., Nagler, J. & Arenas, A. Explosive phenomena in complex networks. *Adv. Phys.* **68**, 123–223 (2019).
17. Fan, J. & Chen, X. General clique percolation in random networks. *Europhys. Lett.* **107**, 28005 (2014).
18. Nguyen, B. G. Gap exponents for percolation processes with triangle condition. *J. Stat. Phys.* **49**, 235–243 (1987).
19. Sornette, D. Discrete-scale invariance and complex dimensions. *Phys. Rep.* **297**, 239–270 (1998).
20. Chen, W., Schröder, M., D'Souza, R. M., Sornette, D. & Nagler, J. Microtransition cascades to percolation. *Phys. Rev. Lett.* **112**, 155701 (2014).
21. Schröder, M., Chen, W. & Nagler, J. Discrete scale invariance in supercritical percolation. *N. J. Phys.* **18**, 013042 (2016).
22. Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001).
23. Gumbel, E. J. *Statistics of Extremes* (Courier Corporation, 2012).
24. Schröder, M., Rahbari, S. H. E. & Nagler, J. Crackling noise in fractional percolation. *Nat. Commun.* **4**, 2222 (2013).
25. Fisher, M. E. Renormalization group theory: its basis and formulation in statistical physics. *Rev. Mod. Phys.* **70**, 653–681 (1998).
26. Cho, Y. S., Kahng, B. & Kim, D. Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E* **81**, 030103 (2010).
27. Cho, Y. S., Mazza, M. G., Kahng, B. & Nagler, J. Genuine non-self-averaging and ultraslow convergence in gelation. *Phys. Rev. E* **94**, 022602 (2016).
28. Cho, Y., Lee, J., Herrmann, H. & Kahng, B. Hybrid percolation transition in cluster merging processes: continuously varying exponents. *Phys. Rev. Lett.* **116**, 025701 (2016).
29. Du, C., Satik, C. & Yortsos, Y. C. Percolation in a fractional Brownian motion lattice. *AIChE J.* **42**, 2392–2395 (1996).
30. Isichenko, M. B. Percolation, statistical topography, and transport in random media. *Rev. Mod. Phys.* **64**, 961–1043 (1992).
31. Nijs, M. P. Md A relation between the temperature exponents of the eight-vertex and q-state Potts model. *J. Phys. A* **12**, 1857–1868 (1979).
32. Essam, I. W., Gaunt, D. S. & Guttmann, A. J. Percolation theory at the critical dimension. *J. Phys. A* **11**, 1983–1990 (1978).
33. Bollobás, B. *Random Graphs* (Cambridge Univ. Press, 2001).
34. Riordan, O. & Warnke, L. Explosive percolation is continuous. *Science* **333**, 322–324 (2011).
35. Cho, Y. S., Hwang, S., Herrmann, H. J. & Kahng, B. Avoiding a spanning cluster in percolation models. *Science* **339**, 1185–1187 (2013).
36. Smoluchowski, M. Drei vorträge über diffusion, brownsche bewegung und koagulation von kolloidteilchen. *Z. Phys.* **17**, 557–585 (1916).
37. Family, F. & Landau, D. P. *Kinetics of Aggregation and Gelation* (Elsevier, 2012).
38. Lee, D., Choi, W., Kertész, J. & Kahng, B. Universal mechanism for hybrid percolation transitions. *Sci. Rep.* **7**, 5723 (2017).
39. Fan, J., Meng, J. & Saberi, A. A. Percolation framework of the Earth's topography. *Phys. Rev. E* **99**, 022304 (2019).
40. Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: A Universal Concept in Nonlinear Sciences* 1st edn (Cambridge Univ. Press, 2003).
41. Romualdo, P.-S., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
42. Strenski, P. N., Bradley, R. M. & Debierre, J.-M. Scaling behavior of percolation surfaces in three dimensions. *Phys. Rev. Lett.* **66**, 1330–1333 (1991).
43. Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* **5**, 17–60 (1960).
44. Nachmias, A. & Peres, Y. Critical percolation on random regular graphs. *Random Struct. Algorithms* **36**, 111–148 (2010).
45. Essam, J. W. Percolation theory. *Rep. Prog. Phys.* **43**, 833–912 (1980).

Methods

Details on numerics and fitting. In the first step, we use an efficient Monte Carlo algorithm, that is, the Newman–Ziff algorithm[13], to calculate Δ, r_c and S_c. By subtracting the respective mean values, we obtain the fluctuations that are studied in the next step. For a given L, we compute the distributions by using a fixed number of bins (if not stated otherwise, we chose $n=20$). In the third step, we rescale the data with the factors L^{β_2}, L^{1/ν_2} and $L^{-d_{f_2}}$, respectively. We then compute the distributions of the scaled data by using a fixed number of bins (also $n=20$, if not stated otherwise). Please note that we do not use exponential binning. In the last step, we use the Levenberg–Marquardt iteration algorithm to fit our numerical data.

Fitting with Levenberg–Marquardt. Since the mode of a unimodal continuous probability distribution is the value at which the probability density function attains its maximum value, the parameter B in equation (15) can be identified as the mode of data. The parameter B is thus determined from the empirical data. As explained in the main text, from theory, we have $f_0=f_1=f_2=0$. This leaves a simple parameter space spanned by A and ω for fitting. However, the parameter A is an amplitude—a pre-factor of the homogenous (scaling) function—while ω also has a simple physical interpretation as well, that is, it is a simple scaling parameter. This allows a simple numerical method to be used for fitting, which is explained in the Supplementary Information and is used also for the parameter fitting of scaling functions defined in equations (16) and (18) in the main text. Specifically, we use the Levenberg–Marquardt[46,47] algorithm as a standard tool for nonlinear fitting. It combines two minimization methods, namely gradient-descent and the Gauss–Newton method. In the gradient-descent method, the sum of the squared errors is reduced by updating the parameters in the steepest-descent direction. In the Gauss–Newton method, the sum of the squared errors is reduced by assuming the least squares function behaves locally quadratic. The Levenberg–Marquardt algorithm acts more like a gradient-descent method when the parameters are far from their optimal value, and acts more like the Gauss–Newton method when the parameters are close to their optimal value. The iterative process completes when the difference between reduced chi-squared values of two successive iterations is less than a certain tolerance value (throughout this paper, we set this threshold to 10^{-9}). In Supplementary Table 2, following reduced chi-squared statistics, standard errors were scaled with the square root of the reduced chi-squared.

Data availability

The data represented in Figs. 1–5 are available as Source Data. All other data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Code availability

The C++ and Python codes used for the analysis is available on GitHub (https://github.com/fanjingfang/Universal-gap-scaling-in-percolation). All figures are plotted by Origin 2018.

References

46. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **2**, 164–168 (1944).
47. Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**, 431–441 (1963).

Acknowledgements

We acknowledge the 'East Africa Peru India Climate Capacities — EPICC' project, which is part of the International Climate Initiative (IKI). The Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) supports this initiative on the basis of a decision adopted by the German Bundestag. The Potsdam Institute for Climate Impact Research (PIK) is leading the execution of the project together with its project partners The Energy and Resources Institute (TERI) and the Deutscher Wetterdienst (DWD). A.A.S. acknowledges support from the Alexander von Humboldt Foundation and partial financial support from the research council of the University of Tehran.

Author contributions

J.F., J.M., A.A.S., J.K. and J.N. designed the research, conceived the study, carried out the analysis and prepared the manuscript. J.F., J.M. and J.N. analysed data. J.F., J.M., Y.L., A.A.S. and J.N. discussed results and contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41567-019-0783-2.

Correspondence and requests for materials should be addressed to J.F. or J.N.

Peer review information *Nature Physics* thanks Filippo Radicchiand the other, anonymous, reviewer(s) for their contribution to the peer review of this work.