Author Accepted Manuscript

Publisher Version:
Imre, E., Illés, Z., Casini, F., Guida, G., Feng, S., Datcheva, M., Baille, W., Bálint, Á., Mwinken, D.K., Lorincz, J. and Leak, J., 2024. Grading curve relations for saturated hydraulic conductivity of granular materials. *Environmental Geotechnics*, pp.1-85. https://doi.org/10.1680/jenge.23.00131

- Article type: Paper
- Date text revised: 20/06/2024
- Number of words in main text and tables: 5000 app: 733
- Number of figures: 10 + 2 in appendices

Grading curve relations for saturated hydraulic conductivity of granular materials

Author 1

- Emoke Imre, Ph.D., Habil
- Bánki Donát Faculty of Mechanical and Safety Engineering and EKIK HBM Research Centers, Óbuda University, Budapest Hungary
- orcid.org/0000-0001-6746-027X

Author 2

- Zsombor Illés, M.Sc.
- Department of Engineering Geology and Geotechnics, Budapest University of Technology and Economics, Budapest, Hungary
- orcid.org/0000-0001-9351-1763

Author 3

- Francesca Casini, Ph.D.
- Department of Civil Engineering and Informatics, Università degli studi di Roma Tor 33Vergata, Rome, Italy
- orcid.org/0000-0001-7933-9055

Author 4

- Giulia Guida, Ph.D.
- Department of Civil Engineering and Informatics, Università degli studi di Roma Tor Vergata, Rome, Italy
- orcid.org/0000-0003-1129-7906

Author 5

- Shuyin Feng, Ph.D.
- Department of Civil Engineering, Birmingham City University, UK
- orcid.org/0000-0002-3837-6762

Author 6

- Maria Datcheva, Ph.D.
- Institute of Mechanics and Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia, Bulgaria
- orcid.org/0000-0002-0801-4795

Author 7

- Wiebke Baille, Ph.D.
- Ruhr-Universität Bochum, Bochum, Germany
- orcid.org/0000-0001-6852-9427

Author 8

- Ágnes Bálint, Ph.D., Habil
- Óbuda University, Institute of Environmental Engineering and Natural Sciences, Budapest, Hungary
- orcid.org/0000-0003-3527-6835

Author 9

- Delphin Kabey Mwinken
- AIAM Doctoral School, Óbuda University, Hungary
- orcid.org/0000-0002-2540-9027

Author 10

- János Lorincz
- AIAM Doctoral School, Óbuda University, Hungary
- orcid.org/0000-0002-2540-9027

Author 11

- James Leak, Ph.D.
- School of Computing, Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, United Kingdom

Author 12

- Daniel Barreto, M.Sc., Ph.D, DIC.
- School of Computing, Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, United Kingdom
- orcid.org/0000-0003-4790-3250

Full contact details of the corresponding author.

- Daniel Barreto, M.Sc., Ph.D, DIC.
- School of Computing, Engineering and the Built Environment, Edinburgh Napier University, Merchiston Campus, Edinburgh, United Kingdom, EH105DT
- orcid.org/0000-0003-4790-3250
- E-mail: d.barreto@napier.ac.uk

Abstract

Estimation of hydraulic conductivity in soils is challenging. The primary aim of this study is to demonstrate that such predictions may be improved if grading curves are appropriately quantified and described, as well as by including density-related values in such relationships. Various saturated hydraulic conductivity models were tested with the assumption that predictions would improve if different grading curve statistics are used. A unimodal database was elaborated using old and new data. Three types of permeability models were examined. One using the traditional

variables consisting of the product of harmonic mean d_h or d_{10} and void ratio, the hydraulic radius; as well as additional density information. The second using the grading entropy coordinate pair S_0, ΔS or the similar pair d_{10}, C_U, expressing the mean grain size on logarithmic scale along with the spread of the grain size distribution and containing similar information on pore size distribution (POSD) by duality. When these were combined in the third type, including also relative density for coarse materials, the fit was the best, verifying the hypothesis that the full pore size range may be the missing pore geometry information of the Taylor's equation (hence predictions are better if grading curve parameters consider the entire distribution of particle sizes). The parameters identified for the various data series were dependent on the data themselves as found from early times in literature. The similarity of grading entropy coordinate pairs and the pair d_{10}, C_U, as well as d_h and d_{10}, was analysed by simulations and by using the same measured data.

Keywords chosen from the ICE Publishing list
Granular materials; Permeability & pore-related properties; Statistical analysis

List of notations

A	is the relative base entropy
B	is the normalised entropy increments
C_U	is the coefficient of uniformity (= d_{60}/d_{10})
C_i	are model parameters
C_S	is the skew
C_K	is the kurtosis
cv	coefficient of variation (Standard deviation/expected value)
d_h	is the harmonic mean diameter
d_{10}	is the diameter of which 10% of the particles are finer
d_{50}	is the diameter of which 50% of the particles are finer
d_{60}	is the diameter of which 60% of the particles are finer
e	is the void ratio
e_{max}	is the maximum void ratio
e_{min}	is the minimum void ratio
k	is the saturated hydraulic conductivity in [cm/s]
v, N	is the fraction number
R_D	is the relative density
r_m	is the hydraulic radius
ρ_v	is the pore volume on unit pore surface
ρ_s	is the grain density

S_s is the surface area of voids

S_0 is the base entropy

ΔS is the entropy increment

S_{sA} is the specific surface area per volume [1/m]

S_{sm} is the specific surface area per mass [m^2/g]

SD standard deviation

PSD is the particle size distribution

POSD is the pore size distribution

GSD is the grain size distribution by dry mass

V is the total volume

V_v is the volume of voids

V_s is the volume of solid

x_i is the relative frequency of fraction i

1. Introduction

The estimation of hydraulic permeability in coarse-grained is challenging. Consequently, several relationships have been proposed (e.g. Hasen, 1893; Kozeny, 1927; Taylor, 1948; Carrier, 2003; Ren & Santamarina, 2018). It is also relatively well accepted that permeability is affected by particle morphology and mineralogy (e.g. Li, et al 2023). Chen, et al (2019) have further considered the effect of grading on permeability at the pore scale. However the focus of the present study is that some existing studies have considered the effect of void ratio and grain size distribution on the estimation of hydraulic conductivity.

With regards to grain size distributions, it is common for existing relationship to use parameters such as d_{10}, d_{50} and c_u (=d_{60}/d_{10}) which do not fully quantify/describe the entire grading curve. It is hypothesized that better estimations of hydraulic conductivity can be made if parameters that characterise the entire grading curve and alternative measures of density are also used for such estimations. To demonstrate such hypotheses, we perform statistical analyses on a combined granular database of hydraulic conductivity experiments, and we have re-evaluated results obtained by other researchers. The combined database has been complemented by new and extensive experiments by the authors covering a very wide range of grading curves for coarse-grained soils. Using the full database, the variables of both the classical theory and the grading entropy theory were applied to develop empirical relationships between grading curves and hydraulic conductivity.

The structure of the paper is as follows. Hereafter, the grain size distribution (GSD), pore size distribution (POSD) and density variables for saturated hydraulic conductivity models of granular materials are summarized. Then the Taylor's permeability model—the foundation of all subsequent models—is presented, highlighting its open question regarding the description of pore geometry and the assumption of this research. Next, we present the methods of data processing, parametric model definition, parameter identification, and model discrimination. Moreover, we detail the results concerning the established database, elaborated models, along with the results of the model discrimination. Finally, we discuss the elaborated model equations and analyse the similarities among various grading curve variables (S_0 - ΔS and d_{10} - C_U , moreover, d_h- d_{10}).

1.1 Grain size distribution curve and statistics

The measured grading curve represents a finite, discrete distribution with *N* uniform statistical cells, based on *N* sieve data (Figure 1). Using a logarithmically uniform cell system representing the size fractions (Table 1, Appendix 1), some additional statistical variables can be defined beyond the traditional quantiles like d_{60}, d_{50}, d_{10} and other derived quantities such as the coefficient of uniformity C_U as follows.

Table 1. Fraction *i* in terms of diameter *d* and *D* (dimensionless diameter variable)

Fraction number *i*	**1**	**23**	**24**
Limits in terms of *d*	1 d_0 to 2 d_0	2^{22} d_0 to 2^{23} d_0	2^{23} d_0 to 2^{24} d_0
D or S_{0i} [-]	1	23	24

1.1.1. Harmonic mean diameter and related variables

The harmonic mean diameter (d_h) from all measured GSD data is computed as follows:

$$d_h = \frac{1}{\sum_{i=1}^{N} \frac{x_i}{d_i}}$$

1.

where d_i is an arbitrary diameter value selected from fraction ("sieve") *i.* This value can be chosen in various ways, however the choice of diameter has a negligible effect on the results.

The mean pore volume is defined as (Imre *et al.*, 2014):

$$\rho_v = \frac{V_v}{S_s}$$

2.

where V_v is the volume of voids and S_s is the specific surface. ρ_v may also be expressed using the harmonic mean diameter, assuming spherical grains:

$$\rho_v = \frac{V - V_s}{6V_s \sum_{i=1}^{N} \frac{x_i}{d_i}} = \frac{1}{6}\frac{e}{\sum_{i=1}^{N} \frac{x_i}{d_i}} = \frac{e}{6} d_h,$$

3.

where e is the void ratio. It can be noted that the value of the mean pore volume is equal to the hydraulic radius r_m (Taylor, 1948), containing the product of the void ratio and the harmonic mean diameter.

The specific surface area per volume of the soil is defined as:

$$S_{sA} = \frac{6}{(1+e)} \sum_{i=1}^{N} \frac{x_i}{d_i} = \frac{6}{(1+e)d_h},$$

4.

The specific surface area per mass of the soil is defined as:

$$S_{sm} = \frac{6}{\rho_s} \sum_{i=1}^{N} \frac{x_i}{d_i} = \frac{6}{\rho_s d_h}$$

5.

1.1.2. The grading entropy coordinates

In Figure 1, the GSD is represented, where the sieve fractions, with sieve hole diameters doubling at each step, create a uniform cell system. The four grading entropy coordinates, derived from all measured GSD data, are calculated as follows (Lorincz, 1986, Singh, 2014).

$$S_0 = \sum x_i S_{0i}$$

6.

$$A = \frac{S_0 - S_{0min}}{S_{0max} - S_{0min}}$$

7.

$$\Delta S = \frac{-1}{\ln(2)} \sum_{i=1}^{N} x_i \ln x_i$$

8.

$$B = \frac{\Delta S}{\ln N}$$

9.

where $S_{0i} = i$ is the i-th fraction entropy (see Table 1), N is the number of fractions including the smallest and largest diameter non-zero fractions.

The d_0 in Table 1 is limited by the smallest diameter which may approximately be equal to the diameter of the SiO_4 tetrahedron (~2.68E-8 m). In this work, d_0=3.05175E-08 m is used. It can be noted that the relation between diameter limits and the S_{0i} is not unique.

By specifying the arbitrary smallest (i-th) and the arbitrary largest ((i+N-1)-th) non-zero fractions, infinite many grading curves can be defined. It can be shown that for every fixed value of A, the subgraph area of the related GSD-s is the same, and there is a unique, optimal grading curve with maximum B and finite fractal distribution. Since this optimal grading curve has no inflexion point, it is a kind of mean grading curve. It follows that the fractal grading curve series depending on A can be used to elaborate “mean” relationships of the various grading curve statistics (Imre *et al.*, 2022).

1.2 Density type permeability model variables

The density variables employed in the saturated hydraulic conductivity models of granular materials are summarized hereafter. The most popular density variables are the void ratio (e), the porosity (n), dry density (ρ_d), the solid volume ratio (s) or its inverse, the specific volume (v). Their basic relations are given below:

$$n = 1 - \frac{1}{1+e} = \frac{e}{1+e}$$

10.

$$e = \frac{1-s}{s}$$

11.

$$s = \frac{1}{1+e} = \frac{1}{v}$$

12.

$$\rho_d = s\rho_s$$

13.

The most informative parameter is the relative density (I_D or R_d), which is dependent on three variables: the void ratio *e* and the minimum and maximum dry densities in terms of $e_{\max}$ and $e_{\min}$:

$$I_D = \frac{e_{\max} - e}{e_{\max} - e_{\min}}$$

14.

Notably, (Kabai, 1974) observed that the ratio e_{min} / e_{max} remains approximately constant for most sands but begins to decrease as the soil contains more silt, see some values in (Imre *et al.*, 2011). Furthermore, the e_{max} of fractal grain size distributions has a minimum at *A*=2/3 by observation which is also a boundary defining stable and instable packings, (e.g. Lorincz, 1986; Imre et al., 2019*).* In practical terms this highlights that grading entropy parameters may be as or more useful than common parameters such as C_u and d_{10} to define the suitability of granular filters and the stability of fills and embankments.

1.3 The Taylor's equation; aim and structure of the paper

In Taylor's derivation, the saturated permeability relation is derived from Poiseuille's law of hydraulics, considering soil pores as a group of tubes. The Taylor permeability equation (Taylor, 1947), reads:

$$k = \left(\frac{V_v}{S_s}\right)^2 \frac{\gamma_w}{\mu} \frac{e}{(1+e)} C = r_m^2 \frac{\gamma_w}{\mu} \frac{e}{(1+e)} C,$$

15.

where γ_w is permeant's unit weight, μ is the dynamic viscosity of the permeant. The $r_m = ed_h/6$ is the hydraulic radius or the mean pore size in the case of spherical grains, *e/(1+e)* is the porosity, and the free parameter *C* depends on additional pore geometry characteristics. It follows that a good permeability model may contain the product of the third power of void ratio and the second power of the harmonic mean diameter.

In the present study it was assumed that the pore geometry can be characterized by the four grading entropy coordinates (i.e. Eqs 5 – 8). These are precise grading curve statistics, based on all data measured during sieving. Hence they can enhance the accuracy of the soil permeability models.

It is well-known that the S_0, ΔS and their normalised forms *A* and *B*, are related to a kind of mean logarithmic grain diameter, to the spread of the distribution (similar to C_U) and to the internal structure and stability information.

In the present study, the pairs S_0 - ΔS and d_{10} - C_U were incorporated into various permeability models. The parameters of the so defined models depending on the underlying data, (Taylor, 1948), were identified using a new database containing relatively unimodal grain size distributions, ranging from silt to gravel sizes.

2. Methods

2.1 Databases

The databases used in the present study correspond to both existing and new hydraulic conductivity experiments. Existing databases were chosen with the aim of considering grading effects including particle sizes ranging from silts to gravels. Furthermore, it is considered important to make comparisons that involve identical testing conditions and repeatability. In that regard all tests and series included consider constant head tests with identical testing conditions and many of them with repeated measurements that add reliability to our databases.

Series 1 to 4 are based on the research of the Central Organisation for Flood Protection, Hungary, collecting different soils from 10 pits along various dikes. The series were created by mixing soils, ranging from silt to gravel, with four fixed, different d_{10} ranges in the silt fraction between 0.006 to 0.016mm. As increasing amounts of coarse materials were added, the mixtures became progressively bimodal, with C_U ranging from 2 to 530. Falling head tests were repeated three times on 74 soil mixtures (Nagy, 2011, 2012); the saturated hydraulic conductivity ranged from 6E-6 to 5E-3 cm/s.

Series 5 (coarse sand and fine gravel) was based on the database by Feng, et al (2019). The soils had a d_{10} range of 0.72 to 5.82 mm, C_U ranged from 1.9 to 6.9, saturated hydraulic conductivity (constant head test) ranged from 0.378 to 50.107 cm/s. At least half of the samples were significantly bimodal.

Series 6 to 8 are from Pap and Mahler (2018) and Nagy 92010, 2011). Series 6 comprises various soils and measuring techniques extending for finer soils. The d_{10} ranged of 0.004 to 0.01 mm, C_U ranged from 3 to 84, saturated hydraulic conductivity (constant head test) ranged from 1E-7 to 1E-6 cm/s. Series 7 and 8 are part of Nagy's data (Series 1 to 4). Hence partly overlapping Series 1 to 4.

In the present research, some new measurements were conducted by the authors on 2-fraction soils with four fractions: 0.25-0.5 mm (medium sand), 0.5-1 mm and 1-2 mm (coarse sand), and 2-4 mm (fine gravel). These comprised Series 9 and 10 featuring 15 (one repeat) and 45 (3 repeats) results on 15 identical compositions, differing only in density. The d_{10} range was of 0.28 to 1.4 mm, C_U ranged from 1.6 to 2.2, saturated hydraulic conductivity (constant head test) ranged from 0.079 to 2.2 cm/s.

In the data processing phase, the Weibull fitting (Guida *et al.,* 2016; Casini *et al.,* 2017) was applied, to provide the ordinates of the GSD in the cell system (Table 1) and some completion of the measured data for fines, if it was needed. Then – besides the traditional quantile parameters and derived grading curve parameters, including d_{10}, d_{30}, d_{50}, d_{60} and the uniformity coefficient C_U - the four grading entropy coordinates and central moments (based on parameter *D* in Table 1) were computed. Bimodal grading curves were excluded based on the normalised grading entropy coordinate values (see Appendix 1). Where the fine content was not precisely measured in series 1 to 4, the PSD was extrapolated below d=2E-3 mm down to d_{min}=6.1E-05 mm (which generally did not influence the value of d_{10}).

2.2 Permeability modelling

2.2.1 Some existing models

The simplest, single-linear models contain only d_{10}, like the model in (Hasen, 1893):

$$k = 1.3 C_H d_{10}^2$$

16

where the parameter C_H is Hazen's empirical coefficient and d_{10} is the characteristic particle diameter.

Lumped/composite parameters are commonly used because (i) the hydraulic radius component of Taylor's model (Equation 15) is the product of a diameter value and the void ratio, (ii) density is an important additional variable. An example is the Chapuis's equation (Chapuis, 2004):

$$k = 2.4622 \left[d_{10}^2 \frac{e^3}{1+e} \right]^{0.7825}$$

17.

In the context of multi-variable models with parameters identifiable through multiple linear regression, certain variable pairs can be highlighted (Carman, 1937, 1939). For example, (Kozeny, 1927) gives the following formula by using a value for d_{10} less than 1.0 mm:

$$k = 1.2\ C_U^{0.735}\ d_{10}^{0.89} \frac{e^3}{1+e}$$

18.

Carrier (2003) proposes a similar equation using d_h. Instead of using d_{10} and C_U the grading entropy coordinates pairs (*A, B*) and (S_0, ΔS) along with void ratio were used by Feng *et al.* (2017) and Imre *et al.* (2021).

2.2.2 The parametric models used in the model discrimination study

The following parametric models were included in the model discrimination study.

A single-variable model with two unknown parameters given by the expression:

$$k = C_1\, p^{C_2}$$

19.

where parameters C_1 and C_2 depend on the unit of the variables *p* and *k*. The *p* is either a diameter value (e.g., d_{10}, d_{30}, d_{50}, d_{60} and d_h) or a lumped/composite variable. The latter is the product of some diameter value or a harmonic mean - based variable, the void ratio and the porosity, expressed, for example, as $d_{10}^3\, e^3/(1+e)$.

The parametric form of the model by Ren & Santamarina (2018):

$$k = C \cdot S_{sA}^{-2} e^{C_2}$$

20.

Note that for comparison with Equation 19, $C_1 = C\, S_{sA}{}^2$.

Multivariable models with three or four unknown parameters were also used:

$$k = \exp C_3 \Delta S^{C_1} S_0^{C_2}$$

21.

$$k = \exp C_4 \Delta S^{C_1} S_0{}^{C_2} p^{C_3}$$

22.

The base entropy S_0 and the entropy increment ΔS were exchanged with d_{10} and C_U in some cases (and with *A* and *B*, see section 4).

2.2.3 Model fitting, discrimination and validation methods

The inverse problem was linear for most of the considered conductivity models when using the logarithmic form of Equation 21 (or 22), e.g.:

$$\ln k = C_3 + C_1 \ln \Delta S + C_2 \ln S_0 \quad \text{and}$$

$$\ln k = C_4 + C_1 \ln \Delta S + C_2 \ln S_0 + C_3 \ln p$$

23.

The (multi-)linear model fitting was based on the weak solution of the Gauss Normal Equations of the formulated inverse problem. Subsequently, the standard deviation and coefficient of variation were estimated (Press *et al*., 2007). The model discrimination was based on either the minima of the normalised merit function called fitting error *F* or on the R^2 value (defined as one minus the ratio of the residual variance to the total variance of the dependent variable, quantifying the fraction of data variance explained by the model). The results of the model discrimination study are further analysed in Appendix 2, which explores the dependence of model parameters on the data used for identification and evaluates model accuracy both on the training data and withheld data.

3. Results

3.1 The database

3.1.1 Grading curve statistics

The results are shown in Figures 2 to 4, and Tables 1 and 2, as well as Appendix 1. The Weibull fitting provided the ordinates of the GSD in the cell system (Table 1) to compute the various GSD statistics. Highly bimodal samples were left out on the basis of the GSD statistics (see Appendix 1). According to the results (Imre *et al.*, 2021), the entropy coordinates changed significantly if the fines were considered by extending the grading curves up to the possible smallest grain sizes which were not measured. The precise value of the fines in the grading curve measurement was essential for the normalized entropy coordinates.

Table 2 contains the range and mean values of the d_{10} and C_U for the selected samples, Table 3 contains the range and mean values of the non-normalised grading entropy coordinates for the selected samples. It can be seen that the mean of S_0 increases for Series 1 to 5, the mean of ΔS decreases for Series 1 to 5 with series number and there is a gap between Series 4 and 5.

The mean of the selected grading curves of the various series is shown in Figure 2(b). The grading curves of the selected and all samples of Series 1 to 4 are shown in Figure 3. The grading entropy coordinates of the selected samples of the old series, the planned samples (to address the gap) and of the new series are shown in Figure 4(a), in the non-normalised grading entropy diagram. The three groups of grading curves of the new Series 9 and 10 (with identical composition) are shown in Figure 4(b).

Table 2. The statistical features of selected data.

d_{10} [mm]	C_U [-]

Series	mean	min	max	mean	min	max
1	0.0055	0.0042	0.0065	27	19	36
2	0.0117	0.0076	0.0154	30	22	38
3	0.0195	0.0089	0.0805	14	2	14
4	0.0391	0.0101	0.0990	13	4	36
5	2.4196	0.7200	5.8200	4.08	2	7
9 – 10*	0.7000	0.2800	1.5900	1.9313	1.59	2.15

*new data, entire series

Table 3. Some statistics of the non-normalised entropy parameters of selected mixtures

Series	mean S_0	min S_0	max S_0	mean ΔS	min ΔS	max ΔS
1	12.8	11.0	14.4	3.2	1.9	3.8
2	14.2	12.0	16.1	3.2	2.4	3.7
3	14.6	11.7	19.1	2.9	1.1	3.5
4	15.2	11.4	18.2	2.6	2.2	3.5
5	17.5	16.1	18.6	1.6	0.6	2.2
9 – 10*	16.5	15.3	17.8	0.9	0.8	1.00

*new data, entire series

3.1.2 Saturated water hydraulic conductivity

In Figure 5, the k is shown in terms of the single (diameter or lumped) curve variables. Each d_i correlated positively; the best was the d_{10}. However, the lumped variables like $d_{10}d_{10}e^3/(1+e)$ with extra porosity term showed correlation improvement.

In Figure 6, the k is shown in terms of C_U and in terms of S_0 and ΔS. Nagy's research gave separate equation of type $A/(C_U+B)+C\ d_{10}^2$ for series 1 to 4, with parameters A, B and C, predicting decrease with C_U and increases with d_{10}.

In Figures 6(a), (b), semi-linear correlation trends of hydraulic conductivity in terms of C_U are shown at the various fixed d_{10} range for the unimodal samples of Series 1 to 4, and an evidence of suffosion. There is a basically increasing trend with the series number (related to increasing d_{10} ranges) and decreasing with increasing C_U, in accordance to (Nagy, 2011). A similar trend is shown in Figure 6(c) and (d) in terms of the entropy coordinates: *k* decreases with increasing ΔS, like with C_U, and increases with increasing S_0, similarly to d_{10}, but with a less significant correlation. This will be discussed in section 4.3.

In Figure 7, the *k* is shown in terms of simple, single variable S_{sA}, the relative density and the lumped Santamarina's variable, moreover the e is shown in terms of sample number 1 to 15, for the new 2-fraction coarse mixtures.

The regression is not acceptable in terms of single variables without density term only. The *k* – relative density graphs exhibit significant separation between loose and dense samples and among grain sizes as follows. In coarsest group 1 *k* ranges from 0.05 to 0.09 cm/s for dense and from 0.22 to 0.48 cm/s for loose samples. In group 2 *k* ranges from 0.03 to 0.06 cm/s for samples and from 0.13 to 0.27 cm/s for loose samples. In finest group 3 *k* ranges from 0.007 to 0.012 cm/s for dense and from 0.05 to 0.14 cm/s for loose samples. The great difference can tentatively be explained by a different – possibly honeycomb – structure for the loose samples.

The Taylor equation simplifies assuming that parameter *C* is constant as follows:

$$\frac{k_1}{k_2} = \frac{e_1^2}{e_2^2}$$

24.

Since identical samples were tested at different densities, the fit to Equation 24 was used to indicate the different structure for the denser samples and the looser samples.

In Figure 7(d), *e* values are depicted, with e_{min} and e_{max} derived from the literature. The e_{max} was measured by (Lorincz, 1986) on samples with identical composition, while e_{min} was estimated based on data from (Kabai, 1974, Imre *et al.,* 2011).

3.2 Model discrimination results

The results are shown in Tables 4 to 6, and are presented separately for the old part, the new part and the entire unimodal database (Figure 2(a)).

3.2.1.1 Old part of database (joint series 1 to 5)

The results are shown in Table 4. The following is the fitting quality in accuracy increasing order:

1. Single variable using diameter - type variables (e.g., d_{10}).
2. Grading entropy parameter pair only.
3. Lumped single variable with porosity.
4. Grading entropy parameter pair combined with void ratio (density information).
5. Grading entropy parameter pair combined with lumped variables, which provided the best accuracy.

The pair S_0 - ΔS was found interchangeable with the pair d_{10} - C_U. The latter gave slightly worse values for R^2 as for the grading entropy parameters only and a similar value as when using d_{10} alone. This indicates that the pair S_0 - ΔS was better for representing the data than the pair d_{10} - C_U.

Table 4. Model discrimination based on the old database, using Series 1-5 jointly..

variable *p*	R^2 (multiple linear model, entropy parameters and *p*)	R^2 (single-linear model with *p*)
$d_{10} \cdot d_{30} \cdot d_{60} \cdot e^3/(1+e)$	0.963	0.949
$d_{10}^2 \cdot e^3/(1+e)$	0.968	0.946
1	0.934*	
d_{10}		0.9117
d_{30}		0.8231
d_{60}		0.716
e	0.953	0.131

* C_U, d_{10} gave R^2=0.9074

3.2.1.2 Elaborated single-variable models

The elaborated equations for the single variables are shown in Equations 25 to 28. For the predictor variable d_{10}, the following equation resulted (R^2=0.9117):

$$k = 0.878\, d_{10}^{\;2.0213}$$

25.

This is the Hazen equation (Hazen, 1893) with fitted model parameters for this database. The equation obtained with R^2=0.823 employing as a predictor variable d_{30} reads

$$k = 2.265\, d_{30}^{\;2.5571}$$

26.

The equation with R^2=0.716 using as a predictor variable d_{60} is:

$$k = 6.047\, d_{60}^{\;2.6393}$$

27.

In terms of $d_{10}^2e^3/(1+e)$, the regression analysis gives for k with R^2 = 0.946 the following:

$$k = 5.868\left[d_{10}^2 \frac{e^3}{1+e}\right]^{1.0322}$$

28.

This is the Chapuis's equation (Chapuis, 2004) adapted to the new soil data set (the original equation fits the data with R^2=0.2025).

3.2.1.3 Elaborated multi-variable models

Concerning the multi-variable Equations 21, 22; results are shown in Tables 5 and 6 and in Equations 29 to 32. The parameters were determined in the equivalent, natural logarithm form Equation 22. The exponents of the entropy increment ΔS and the base entropy S_0 were the identified parameters C_1, C_2, and the coefficient in Equation 21 was equal to $\exp(C_3)$.

Table 5 shows the parameters of the 3-parameter entropy variable equation fitted on individual Series 1 to 5. The parameters depended on the position of series in the entropy diagram (see Appendix 2), the difference was more significant than the linear error of the parameter identification. The absolute value of the exponent of ΔS was between 0.45 and 6.41, and the value of the exponent of S_0 *was* between 2.8 and 32.9. The exponent of ΔS decreased, while the exponent of S_0 increased as soil became coarser.

Table 5 Results of fit of the 3-parameter Equation 21, using data from Series 1 to 5 and joint Series 1 to 4, parameter estimates and fitting errors (selection of S_0, as shown in Appendix 1).

Series	5	1	2	3	4	1..4
C_1	-0.45	-3.80	-1.63	-1.23	-0.93	-6.41
C_2	32.87	7.11	-4.58	7.77	2.76	4.69
C_3	-92.46	-23.68	3.88	-27.24	-13.82	-14.34
SD(C_1)	0.59	1.15	0.44	3.44	1.13	0.92
SD(C_2)	4.96	1.81	0.68	8.82	2.09	1.58
SD(C_3)	14.39	4.94	1.40	18.80	4.45	3.97
CV(C_1)	-1.32	-0.30	-0.27	-2.80	-1.21	-0.14
CV(C_2)	0.15	0.25	-0.15	1.13	0.76	0.34
CV(C_3)	-0.16	-0.21	0.36	-0.69	-0.32	-0.28
Fitting Error [-]	0.07	2.7E-4	2E-2	2.6E-2	4.8E-3	2E-02

Table 6 shows the estimated parameters of the various 4-parameter entropy variable equations identified using joint Series 1 to 5. The absolute value of the exponent of ΔS was between 2.4 and 6.2, and the value of the exponent of S_0 *was* between 2.6 and 22.2. The value of the coefficient *C* varied between exp(–59.4) to exp(–5.7), small numbers occurred being in the same interval as for the 3-parameter case.

The sign of the exponent of S_0 was generally positive, and the sign of the exponent of ΔS was generally negative. The coefficient of variation was smaller for the "global" equation (using joint Series 1-5) than for the series separately.

Table 6 Results of the fit of the 4-parameter equations, estimated parameters, coefficients of variation and R^2

p	$d_{10}d_{10}e^3/(1+e)$	$d_{10}^3e^3/(1+e)$	$d_{10}d_{30}d_{60}e^3/(1+e)$	*e*	entropy parameters only
C_1	-2.4	-3.0	-3.3	-3.2	-6.2
C_2	8.8	8.3	2.6	22.3	17.5
C_3	0.6	0.3	0.4	5.1	-46.6
C_4	-22.3	-21.1	-5.7	-59.4	
CV(C_1)	-0.3	-0.2	-0.2	-0.2	-0.1
CV(C_2)	0.2	0.2	1.0	0.1	0.1
CV(C_3)	0.1	0.1	0.1	0.2	0.1
CV(C_4)	-0.2	-0.2	-1.2	-0.1	2
R^2	0.968	0.965	0.963	0.953	0.934

The fitting error was smaller if individual Series 1 and 4 were considered. If Series 2, 3 and 5, or Series 1-4 or 1-5 were used jointly in derivations, the magnitude of the fitting eeror was up to two orders of magnitude larger.

Some equations obtained using Series 1 to 5 jointly are:

$k = 7.67\text{E-}21\, \Delta S^{-6.21}\, S_0^{17.51}$

29.

$k = 2.38\text{E-}26\, \Delta S^{-3.24}\, S_0^{22.28}\, e^{5.15}$

30.

$k = 0.003375\, \Delta S^{-3.32}\, S_0^{2.64}\, [d_{10} d_{30} d_{60}\, e^3/(1+e)]^{0.43}$

31.

$k = 2.40\text{E-}10\, \Delta S^{-2.4}\, S_0^{8.8}\, [d_{10}^2\, e^3/(1+e)]^{0.6}$

32.

3.2.2. New data

Using new data for two-fraction mixtures consisting of medium- to coarse-grained sand and fine gravel, with identical composition but different densities, the model fitting yielded an R^2 value less than 0.2 when density was not considered as an extra variable and was greater than 0.8 incorporating the pair of grading entropy coordinates and the relative density parameter (Table 7).

Table 7, New data (small C_U), model discrimination,

Independent variables (predictors)	R^2
Entropy parameters and relative density	0.8746
Entropy parameters and void ratio	0.7811
Entropy parameters	0.1700

3.2.3 New and old data together

Being included old and new data (Series 1-5, 9-10), using the Ren and Santamaria model on 150 samples, the identified with R^2 = 0.846 exponent of the void ratio was 8.6, i.e.:

$$k \sim \frac{e^{8.6}}{S_{sA}^2}$$

33.

The fitting of the multiple linear equation using the grading entropy parameters and the Ren-Santamarina's variable was successful, achieving R^2 of 0.924. The equation with the estimated parameters reads:

$$k = 1.729 \cdot 10^{-19} \Delta S^{1.005} \cdot S_0{}^{17.5} \cdot [e^{8.6}/S_{sA}^2]^{0.53}$$

34.

4. Discussion

4.1 The analysis of the results

Various analyses were performed on the results. First, the effect of the training data on the identified parameters and on the model accuracy were considered. The case in which training data were selected is considered in detail in Appendix 2 (Figures A2-1 and 2, Table A2-1). The identified parameters of Equation 21 – depending on the two grading entropy variables - were represented in 1-dimensional form by fixing the one variable. The functions did not intersect each-other if the "hulls" of the entropy coordinates of the training data series were disjunct. The multivariable Equation 21 was more precise if the training data and the tested data were similar. The single-variable Equation 19 (the simple d_{10} - model) was more precise if the training data set was larger than the tested data set. The model discrimination result of the original Series 1 to 4 (with bimodal samples) aligned with the foregoing result, however, there was a notable disparity in the R^2 values where gap-graded soil with possible suffusion were included.

4.2 Some results with normalised grading entropy coordinates and with level lines

The definition of *N* - the number of fractions including the smallest and largest non-zero fractions – is not the same in the literature, for example in (Feng *et al.,* 2019), the arbitrary smallest and largest fractions can be zero fractions. The so computed coordinates are not differing from the non-normalised coordinates due to the functional relationships defined by Equations 6 to 9. The two approaches are equivalent; only numerical differences may occur.

This similarity is illustrated in Figure 8 where the permeability zones and k-level lines are presented in the non-normalised and normalised entropy diagrams. The results are also similar to the earlier data of (Feng et al., 2019).

Feng *et al*., 2019 used normalized entropy coordinates *A* and *B* for Equations 21 and 22, in combination with the void ratio. For Series 5, models employing the independent variable combinations *A* - *B*, and *A* – *B* – e showed R^2 values of 0.90 and 0.96, respectively. In contrast, for the joint unselected Series 1-2, the R^2 values were 0.23 and 0.27 (Imre *et al*., 2021). These results support the present model discrimination study.

However, in the internal stability rule of the grading entropy concept, based on the entropy coordinate *A,* the sharp definition of *N* is needed. In future research, it would be interesting to combine the non-normalised grading entropy coordinates with the normalised entropy coordinate *A,* computed using the sharp definition of *N.* Some early results are presented on the effect of A in (Imre *et al.*, 2020).

4.3 The grading curve statistics

The pair S_0 - ΔS was found interchangeable with pairs d_{10} – C_U (or *A* – *B*, using the wider definition of *N*). To explain this, simulations of mean relations using fractal or mean grading curves (see section 1) and experimental data were considered in Figures 9 and 10, respectively.

According to Figure 9, the theoretical mean relations for ΔS - C_U determined using fractal grading curves with *N*=5, 7 and 20 are non-unique (different branch is related to *A*< or *A* > 2/3) while the theoretical mean relations for *A* - d_{10} determined using fractal grading curves with *N*= 7 is unique.

In Figure 9(a), the measured, unselected data are within, the gap-graded data are outside the band bounded by the theoretical mean ΔS - C_U relation. Similarly, the experimental relation of selected data for ΔS - C_U seems to have a regression along the area bounded by the theoretical band of the mean relation (Figure 10(a)). The regression is stronger for S_0 – d_{10} along the theoretical, unique, mean relation (Figure 10(b)). The theoretical, mean relations may explain that the regression C_U - ΔS gives slightly smaller R^2 than the regression d_{10} - S_0.

Concerning other regressions to measured S_0 - d_i data (Table 8, Figure 10(c)), R^2 is larger for S_0 - d_{60} than for S_0 - d_{10} since d_{60} is closer to the mean abstract diameter value, which is the meaning of S_0. Concerning the experimental relation for Series 2 and 5, the regressions d_{10}-d_h and d_{10}- r_m are linear is semi-log plot. The d_h was slightly larger than d_{10} for gravel and that d_h was smaller than d_{10} for fine sand. Further research is suggested on this matter.

Table 8. Multiple linear regression results for the independent variable combinations of S_0 and selected d_i.

Variable	R^2	Equations
d_{10}	0.7672	$\ln S_0 = 0.0587 \ln d_{10} + 2.8628$
d_{30}	0.8494	$\ln S_0 = 0.0777 \ln d_{30} + 2.8138$
d_{60}	0.8363	$\ln S_0 = 0.0886 \ln d_{60} + 2.7553$

5. Summary and conclusions

5.1 Model discrimination

The parametric saturated hydraulic conductivity models examined here were three types. The first model set was based on single, classical variables like the harmonic mean (d_h) or d_{10}, the void ratio e, porosity $e/(1+e)$ or a lumped variable of these. The second model set was based on variable pairs (the non-normalised grading entropy coordinate pair S_0 - ΔS or the pair d_{10} - C_U). The third model set was based on the combination of variables of the first model set and on a variable pair of the second model sets.

A unimodal database was started to be built for this purpose re-evaluating some previous databases and providing some new data for granular materials ranging from silt to gravel. The bimodal samples of the old databases were left out since for these mixtures the permeability test was not precise due to suffusion. The identification of the parameters of the suggested models was made by multiple linear regression based on various subsets of the new database. The R^2 was generally significant, but the result was the best when thye model sets 1 and 2 were combined. The difference among the model variants was small for Series 5 with small C_U.

All identified parameters depended on the range of grading entropy coordinates of the data used for parameter identification, as well as implicitly on grain shapes and other factors. The successfully tested models were as follows:

- single-variable models, using one lumped variable consisting of some squared diameter variables and at least the second power of void ratio, or more complicated forms, like $d_{10} \cdot d_{30} \cdot d_{60} \cdot e^3/(1+e)$ being related to the hydraulic radius and the porosity terms of Taylor's equation;
- .multi-variable models, with a variable pair $S_0 - \Delta S$ or pair $d_{10} - C_U$ (alone or being completed by either a density terms or one of the previous lumped variables), expressing the missing pore geometry information of Taylor's equation.

5.2 Taylor equation

The Taylor equation contains the porosity, squared hydraulic radius (mean pore volume), and the constant that expresses the pore geometry. This fact may explain why most *k* - models contain the third product of the void ratio and the second pore a diameter value (the best is the harmonic mean diameter based on the fraction cell system and all measured data).

The grading entropy theory (Lorincz, 1986) offers a coplementary statistical system of quantifying the GSD, using all data measured in the sieving test. The model discrimination results supported the hypothesis that the non-normalised grading entropy parameters may give information on the geometry of GSD (mean abstract diameter and spread of the distribution) and by duality on the geometry of POSD, which is needed for the Taylor equation.

5.3 Future research

In future research, in the multivariable models, the non-normalised grading entropy coordinates are planned to be completed by the normalised entropy coordinate *A,* computed using the sharp definition of *N* which may link some additional packing information.

The dependencies of the parameters on the data used for model fitting are planned to be quantitatively determined in future research. To achieve this, the database will be completed and parallel tests with varying grain shapes, e.g., laboratory experiments and numerical simulations using the discrete element method, are planned to be conducted. More research is needed on the present database and on the suggested models families.

The relative density -- giving the best result for sands as density variable, depending on three parameters (the void ratio e, the e_{max}, e_{min} being the void ratio at minimum and maximum dry densities) -- can be a relevant parameter of the future saturated hydraulic conductivity of sands. Note that our data analysis has considered the use of void ratio both as an individual and lumped/composite grain-size permeability relationships. It can be determined from two tests only, based on the research of (Kabai, 1974,) using that the ratio e_{max}/ e_{min} is about constant.

Acknowledgements

The discussions with Professor Nagy and Professor Paul Vardanega are greatly acknowledged. The help of former students Levente Répássy, Dang Thi Quynh Huong, and Lizeth Lamas Lopez is acknowledged. M. Datcheva acknowledges the partial support by Grant No BG05M2OP001-1.001-0003-C01, financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European Structural and Investment funds. D. Barreto acknowledges the support provided by the Royal Academy of Engineering Leverhulme Trust Fellowship 2023 (LTRF2223-19-134) that enabled him to focus on research work.

Contributions

Emoke Imre. Conceptualisation. Data curation. Formal analysis. Funding acquisition. Methodology. Writing. Review and editing.

Zsombor Illés. Conceptualisation. Data curation. Investigation. Writing
Francesca Casini. Data curation. Investigation. Validation
Giulia Guida. Data curation. Investigation. Validation
Shuyin Feng. Data curation. Validation. Writing
Maria Datcheva. Conceptualisation. Data curation. Funding acquisition. Methodology. Writing. Review and editing.
Wiebke Baille. Data curation. Investigation. Validation
Ágnes Bálint. Conceptualisation. Methodology. Writing. Review and editing.
Delphin Kabey Mwinken. Investigation. Validation
János Lorincz. Conceptualisation. Methodology. Writing
James Leak. Data curation. Investigation. Writing. Review and editing.
Daniel Barreto. Conceptualisation. Data curation. Funding acquisition. Writing. Review and editing.

References

Casini F, Guida G, Bartoli M, Viggiani GMB (2017) Grading evolution of an artificial granular material from medium to high stress under one-dimensional compression. Rivista Italiana di Geotecnica 51(4): 69–80, https://doi.org/DOI:10.19199/2017.4.0557-1405.69

Carman PC (1937) Fluid flow through granular beds. Chemical Engineering Research and Design 15: S32–S48, https://doi.org/10.1016/S0263-8762(97)80003-2

Carman PC (1939) Permeability of saturated sands, soils and clays. The Journal of Agricultural Science 29(2): 262–273, https://doi.org/10.1017/S0021859600051789

Carrier WD (2003) Goodbye, Hazen; Hello, Kozeny-Carman. Journal of Geotechnical and Geoenvironmental Engineering 129(11): 1054–1056, https://doi.org/10.1061/(ASCE)1090-0241(2003)129:11(1054)

Chapuis RP (2004) Predicting the saturated hydraulic conductivity of sand and gravel using effective diameter and void ratio. Can Geotech J 41(5): 787–795, https://doi.org/10.1139/t04-022

Chen, R.-P, Liu, P., Liu, X.-M, Wang, P.-F. & Kang, X. (2019) Pore-scale model for estimating the bimodal soil–water characteristic curve and hydraulic conductivity of compacted soils with different initial densities. Engineering Geology, 260, 105199,

Feng S (2017) Assessing the permeability of pavement construction materials by using grading entropy theory. MSc thesis, University of Bristol, Bristol, UK

Feng S, Vardanega PJ, Ibraim E, Widyatmoko I, Ojum C (2019) Permeability assessment of some granular mixtures. Géotechnique 69(7): 646–654, https://doi.org/10.1680/jgeot.17.T.039

Guida G, Bartoli M, Casini F, Viggiani GMB (2016) Weibull Distribution to Describe Grading Evolution of Materials with Crushable Grains. Procedia Engineering 158:75–80, https://doi.org/10.1016/j.proeng.2016.08.408

Hazen A (1893) Some physical properties of sand and gravels. Massachusetts State Board of Health, Wright & Potter Printing, Boston, MA, USA, pp.

Imre E, Fityus S, Keszeyné E, Schanz T (2011) A Comment on the Ratio of the Maximum and Minimum Dry Density for Sands. Geotechnical Engineering Journal of the SEAGS & AGSSEA 42(4):77–82.

Imre E, Lorincz J, Hazay M, Juhász M, Rajkai K, Schanz T, Lins Y, Singh VP, Hortobágyi Zs (2014) Sand mixture density in: of UNSAT2014. pp. 691–697.

Imre E, Lorincz J, Barreto D, Trang PQ, Csonka I, Kaczvinszki-Szabó V, Telekes G. (2019) Preliminary study on the relationship between dry density of sands and the grading entropy parameters. XVII ECSMGE. Reykjavík, Iceland : 1–6 of September 2019 pp. 996-1002. Paper: 167797.

Imre E, Balint A, Barreto D, Khaliunaa A, Boldbaatar T, Dang T, Guadalupe L, Illés Z (2020) The Saturated Hydraulic Conductivity of 2-Fraction Granular Soils and the Internal Stability, In: Hosam, E.A.F. Bayoumi Hamuda.11th International Annual Conference on Sustainable Environmental Protection & Waste Management Responsibility, ICEEE 2020: Proceedings Book. Budapest, Magyarország: Óbudai Egyetem, pp. 267-274.

Imre E, Bálint Á, Nagy L, Lorincz J and Illés Zs (2021) Examination of saturated hydraulic conductivity using grading curve functions, in Proceedings of 6th International Conference on Geotechnical and Geophysical Site Characterisation, Budapest, Hungary, pp. 1–12.

Imre E, Hortobágyi Zs, Illés Zs, Nagy L, Talata I, Barreto D, Fityus S, Singh VP (2022) Statistical parameters and grading curves. In: Rahman R, Jaksa J (eds) Proceedings of the 20th International Conference on Soil Mechanics and Geotechnical Engineering. Australian Geomechanics Society, Sydney Australia, pp 713–718

Imre E, Firgi T, Baille W, Datcheva M, Barreto D, Feng S, Singh V (2023) Soil parameters in terms of entropy coordinates. E3S Web of Conf. (Vol. 382, p. 25003). UNSAT 2023, https://doi.org/10.1051/e3sconf/202338225003

Kabai I (1974) The effect of grading on the compatibility of coarse-grained soils. Periodica Polytechnica Civil Engineering 18(4):255–275.

Kozeny J (1927) Über kapillare Leitung des Wassers im Boden. Akademie der Wissenschaften, Wien, pp 271–306

Li, S.-L., He, R.-J and Kang, X. (2023). Permeability of polymer-modified kaolinite and fly ash–kaolinite mixtures. Géotechnique Letters, 13:4, 211-217

Lorincz J (1986) Grading entropy of soils Doctoral Thesis. Technical Sciences, TU of Budapest, Hungary

Lorincz J, Imre E, Fityus S, Trang P, Tarnai T, Talata I, Singh V (2015) The Grading Entropy-based Criteria for Structural Stability of Granular Materials and Filters. Entropy 17(5): 2781–2811, https://doi.org/10.3390/e17052781

Nagy L (2011) Permeability of well graded soils. Per Pol Civil Eng 55(2): 199–204, https://doi.org/10.3311/pp.ci.2011-2.12

Nagy L (2012) Characterization of piping with grading entropy on the Surány example. Per Pol Civil Eng 56(1):107-113, https://doi.org/10.3311/pp.ci.2012-1.12

Pap M, Mahler A (2018) Comparison of Different Empirical Correlations to Estimate Permeability Coefficient of Quaternary Danube Soils. Period Polytech Civil Eng 63(1): 25–2, https://doi.org/10.3311/PPci.13108

Press WH, Teukolsky SA, Wetterling WT, Flannery BP (2007) Numerical recipes in C: the art of scientific computing. Second Edition, Cambridge University Press, Cambridge, UK; New York, pp. 1–994.

Ren XW, Santamarina JC (2018) The hydraulic conductivity of sediments: A pore size perspective, Engineering Geology, 233: 48–54, https://doi.org/10.1016/j.enggeo.2017.11.022

Singh VP. Entropy theory in hydraulic engineering: an introduction (2014). American Society of Civil Engineers, p.656.

Taylor DW (1948) Fundamentals of soil mechanics. John Wiley & Sons, Inc., New York, USA, pp. 1–700.

Figure captions

Figure 1. (a, b) Grain size distribution and density functions of a sand. Legend: in terms of diameter *d* (dashed line); in terms of abstract diameter *D* (solid line), see Table 1.

Figure 2. Data processing of old and new samples (a). Weibull fit, with completion for fines (series 2, sample 13, bimodal). (b) Mean, selected, unimodal-grading curves for series 1 to 9.

Figure 3. Sample selection for series 1-4. (a) Selected unimodal samples. (b) All samples.

Figure 4. Data processing results of old and new samples. (a) Selected, new and planned samples in the non-normalised grading entropy diagram. (b) Series 9 (= 10), grading curves of new, 2-fraction mixtures, three groups, serial numbers (sample id-s) 1 to 15.

Figure 5. The saturated hydraulic conductivity in terms of a single variable. (a) to (c) *k* in terms of d_{10}, d_{30}, d_{60}, respectively (each d_i correlates positively, and the best is d_{10}). (d) *k* in terms of lumped-variable of the Chapuis model (significant improvement in correlation).

Figure 6. The saturated hydraulic conductivity in terms of elements of pairs d_{10} - C_U or ΔS - S_0. (a) The *k* in terms of C_U, Series 1 to 4, selected samples, and (b) Series 2, all samples (gap-graded, non-selected soils showed suffusion). (c) and (d) The *k* in terms of ΔS and S_0.

Figure 7. Newly measured 2-fraction mixtures data. (a) to (c): Saturated hydraulic conductivity in terms of S_{sA}; relative density; Ren and Santamarina variable; resp. (d) the void ratio with approximate values at the maximum and minimum dry density. Legend: circles – loose samples, squires – dense samples.

Figure 8 The zones of saturated permeability *k* (a) Approximate level lines in non-normalised entropy diagram (Imre *et al.,* 2021). (b) Series 5 in the normalised entropy diagram (Feng *et al.* 2019) (c) The same as (a) in the normalised diagram.

Figure 9. Mean relations of fractal soils, various *N* values. (a) C_U - ΔS with measured, unselected data for Series 3, the unimodal samples are within the band of mean relation, (b) d_{10}- *A.*

Figure 10. (a) and (b): The experimental, relations C_U - ΔS and d_{10}- S_0 on selected data. (c) The experimental relations of d_{10}-d_h and d_{10}- r_m, Series 2 and 5 data.

Figure A2-1. The simplified *k* – models related to Equation 21 with parameters identified from various training series. (a) with respect to ΔS and (b) with respect to S_0

Figure A2-2. Goodness-of-fit for models using various datasets. (a) results considering Eq. 21. (b) results considering Eq. 19. (c) results considering Eq. 21 and a larger dataset. (d) Model performance in terms of d_{10}

Appendix 1 Some notes to the fraction system, and to the use of the abstract diameter

An abstract fraction system is given in terms of d_0 being limited by the smallest diameter, some related information is given in Table A1-1 The soil types related to the suggested d_0 and base entropy values are shown in Table A1-2 in two different ways.

Table A1 -1. Diameter sizes

order of magnitude	size in m	2^n, exponent *n*	fraction serial number
	1.53E-08	-26	0
SiO_4 tetrahedron	3.05E-08	-25	1
a few microns	3.91E-06	-18	8*
1 mm sieve	1.00E-03	-10	16
gravel	6.40E-02	-4	22
km (~4 km)	4.19E+03	12	38

*Range of comminuting limit of small particles by compression (Kendall, 1978)

Table A1-2. The soil fractions in terms of abstract diameter *D*

name	gravel	sand	silt
d [mm]	32-2	2- 0.0625	0.065 - 0.001955
S_{0i} or *D* [-]	20 - 17	16– 12	11 - 7
S_{0i} or *D* [-]*	21 - 18	17 - 13	12 - 8

*alternative, upper values (Tables 4 to 6 are based on the values in row 2).

Using the fact, that the normalized grading entropy parameters *A* and *B* have unique, monotonic mean relationship with the skewness and kurtosis of (of abstract diameter) variable *D*, the various soil series were tabulated in terms of these. Extreme values indicated bi-unimodal, near gap-graded soils which were left out from the unimodal database (Table A1 - 3).

Table A1-3. Selecting unimodal grading curves on the basis of skew C_S and kurtosis C_K of *D* and the normalised grading entropy coordinates, on the example of Series 1. Extreme values (deviating most from the mean) were bi-modal, near gap-graded soils (indicated in bold).

	CS	A	C_K+3	B
1*	-1.08	0.68	**5.05**	**1.00**
2	-0.58	0.56	3.36	1.16
3	-0.45	0.56	3.05	1.21
4	-0.41	0.58	2.74	*1.28
5	-0.20	0.59	2.30	1.28
6	-0.11	0.61	1.99	1.26
7	0.16	0.64	0.83	1.32
8	-1.29	0.60	4.67	1.03
9	-1.46	0.66	4.99	0.86
10*	**0.07**	0.64	1.08	1.29
11*	**0.06**	0.67	0.89	1.25
12*	-1.86	0.57	**6.57**	**0.90**
13*	-1.52	0.57	**5.53**	**0.97**
mean	-0.67	0.61	3.31	1.14

Appendix 2: Discussion of the models

The following questions were investigated to prove the reliability of the new modelsl, using various training and testing datasets, in a prelimary manner.

Question 1. How does the simulated *k* with Eq 21 at various training datasets compre in terms of the entropy variables?

In Eq 21, the *k* is a function of ΔS and S_0. The parameters identified by the various training data subsets are different, as shown in Tables 5, 6. The question how the identified coefficients of Equation 21 may vary in terms of various training data subsets was examined such that $\Delta S - k$ functions and $S_0 - k$ functions were defined by using fixed S_0 and ΔS, resp., (being equal to the mean value given in Table 3). In Figure A2-1(a) and (b), generally the *k* decreases with ΔS and increases with S_0 but differently for the various subsets. The Series 1 appears disjoint in (also in the entropy diagram, Figure 2(a)).

Question 2. How does the goodness of *k* simulated with Eq 21 and Eq 19 compare at various training set on general data of Series 2?

The prediction accuracy of Equation 21 and 19 was tested using the whole, unselected Series 2. The results are shown in Figure A2-2 (a) and (b). For Eq 21 (with only entropy variables), the prediction was better for smaller than from larger training data set. For Eq 19 (with only d_{10}), the reverse was true. In other words, the finding was different for Equation 21 and for 19.

Question 3. How does the goodness of *k* simulated with Eq 21 compare at various training set on selected data of Series 1 to 4? The accuracy of Equation 21 (prediction with only on entropy parameters) on selected series 1 to 4 was tested. The prediction was better for smaller than from larger training data set set (Figure A2-2 (c)), similarly to the previous case.

Question 4. How does the *k* simulated with Eq 21 and Eq 19 compare in terms of d_{10}?

The S_0-d_{10} relation determined here (see Table 8) was used to change variable S_0 into d_{10}, and the value of DS was fixed at various values. It was found in Figure A2- 2(d) that the tested models were basically similar in terms of $d_{10.}$ It is important to note that the entropy variable based equation does not work for DS =0, but is working for DS >0.5.

Question 5. How does the model rank compare on general data (interpolation case)?

Four models were used, *p*, S_0 - ΔS, S_0-ΔS-*e*, and S_0-ΔS-*p* where *p* was the Ren-Santamarina's variable with exponent 8. The results showed the same model discrimination results as for the selected series, according to result in Table A2-1, for various models, considering the complete data, the R^2 variation was similar. Series 2 showed the worse result, possibly due to the inclusion of the gap-graded mixtures showing suffusion during the permeability test.

Question 6. How is the goodness of fit of various models for extrapolation?

Seven experiments of Series 9 were used to validate the elaborated relationships presented in this study. The results in Table A2-2 were characterized by the mean squared difference of the measured and computed values, called variance here. The most sophisticated model-versions Eq 22 were not working (with entropy parameters and ***p***), the single variable models Eq 19 and the only entropy parameter models Eq 21 were better. The d_{10} equation without any other variables also provided good results.

Table A2-1. The R^2 in case of various model variables, non-selected samples

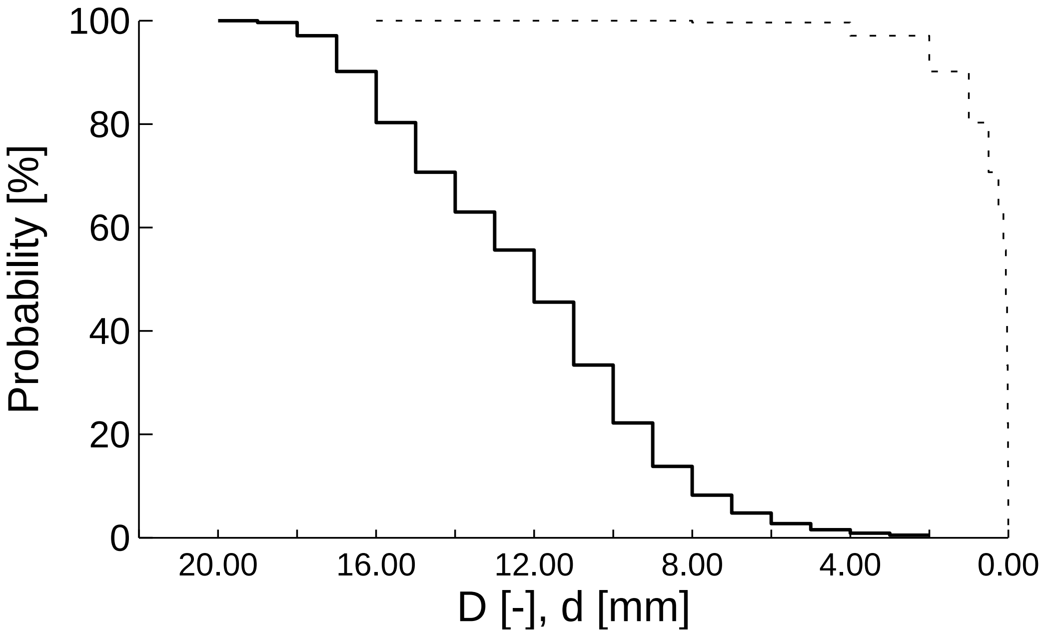
series	test #	p	S_0 ,ΔS	S_0 ,ΔS, e	S_0 ,ΔS, p
1	13	0.0613	0.7077	0.7093	0.7119
2	18	0.1837	0.2424	0.2519	0.3088*
3	30	0.0107	0.6117	0.595	0.5963
4	12	0.4362	0.622	0.7086	0.7593
5	30	0.8905	0.83	0.8917	0.9362

Note: The variable *p* was Ren & Santamarina's (2018) variable with exponent 8

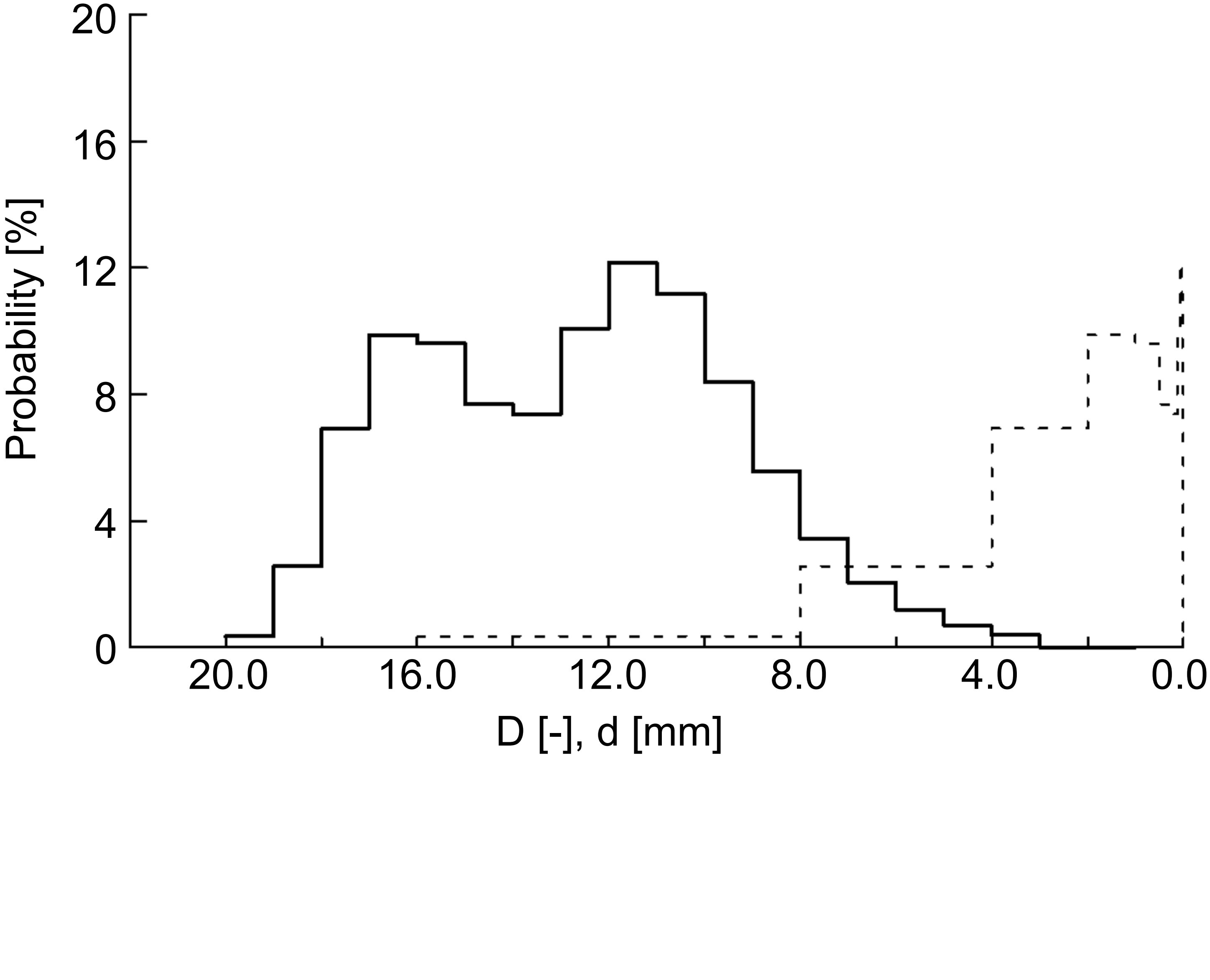
Table A2-2. The variance of extrapolation error, training data set 1..5, tested data from series 9.

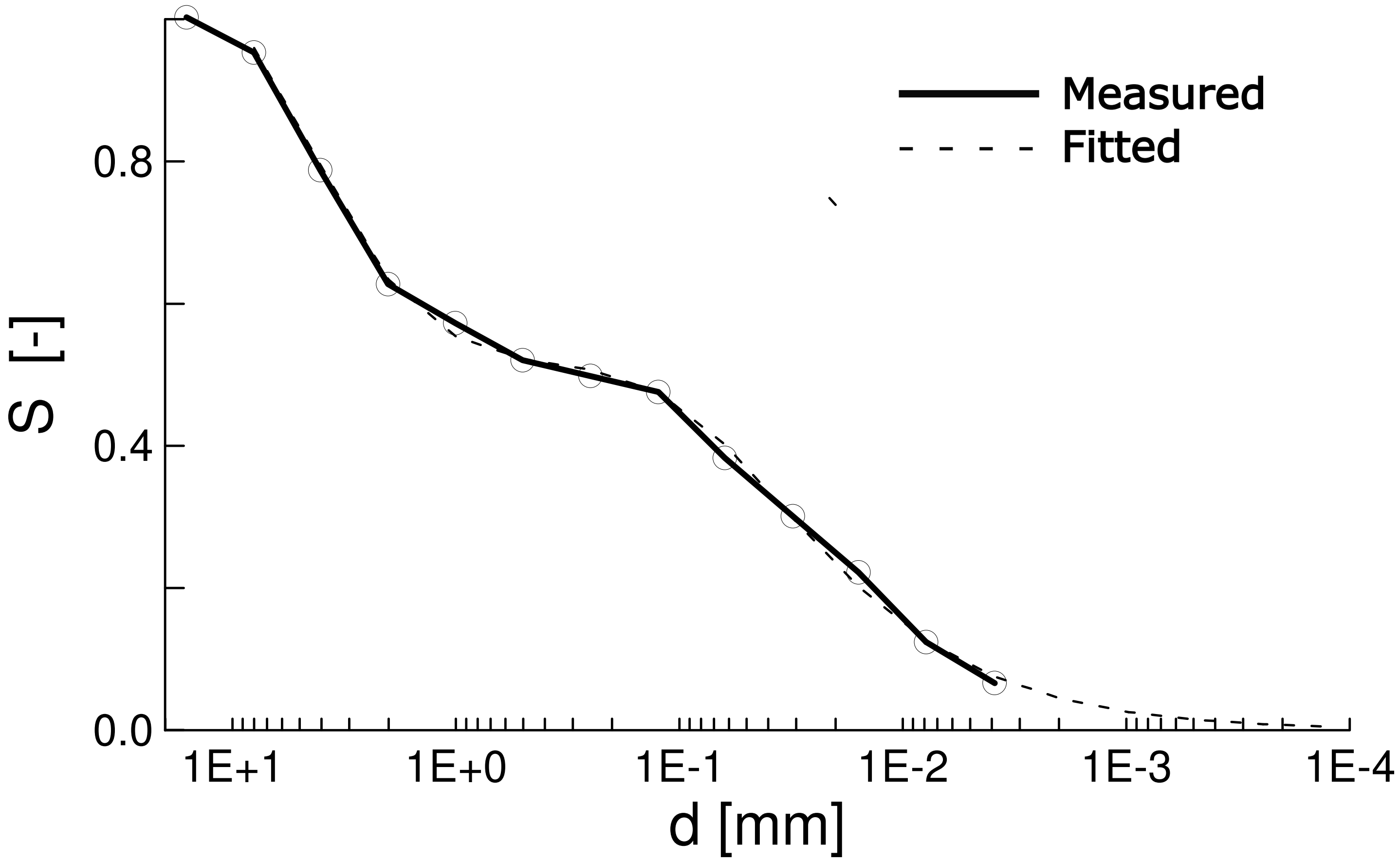
	Eq 21 ^{0}	Eq 22 $^+$	Eq 22 $^+$	Eq 19 *
p	-	d_{10}^2 e^3/ (1+e)	d_{10}* d_{30}* d_{60}* e^3/(1+e)	d_{10}^2
error	1.1	57.7	28.9	0.4

^{0} only entropy coordinates, * classical variables only, $^+$ with entropy coordinates

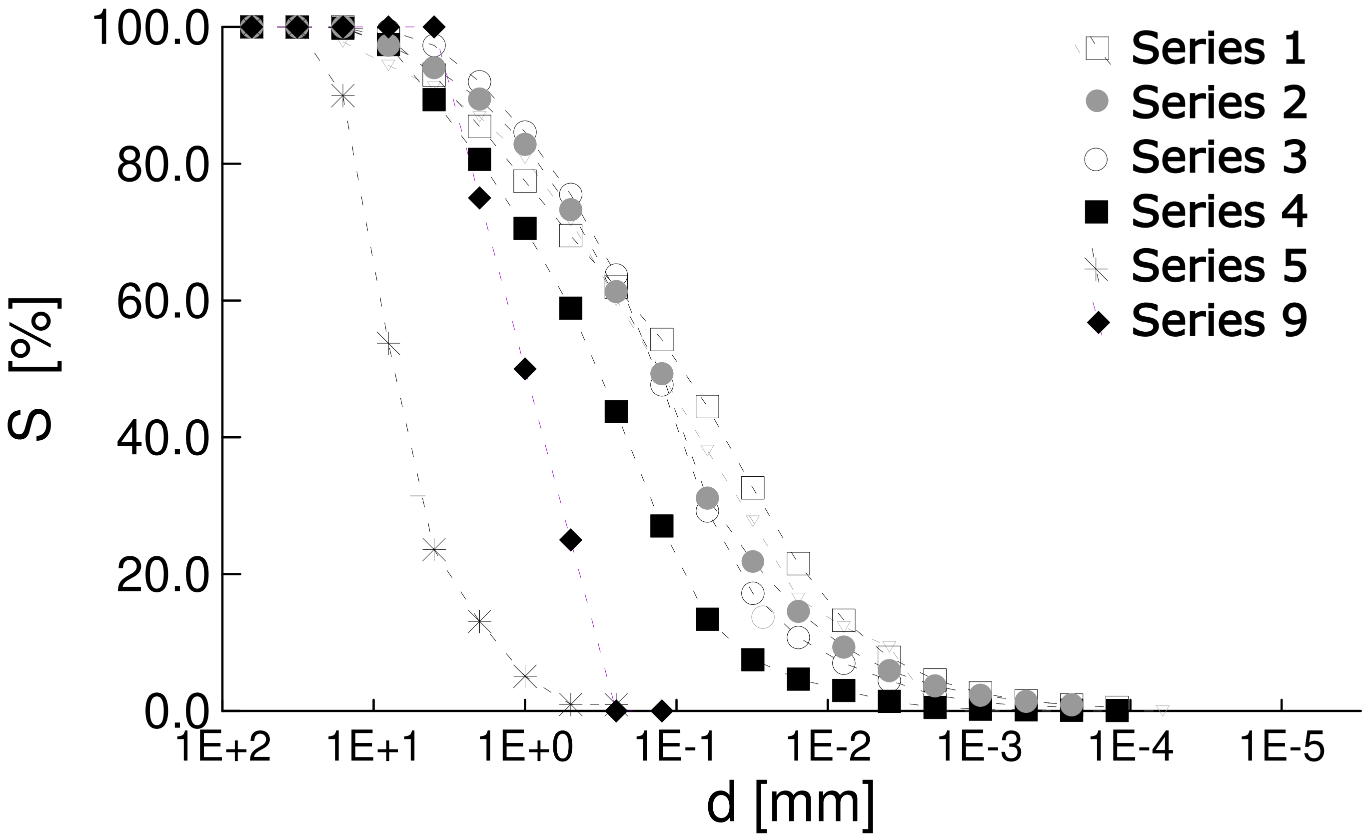


Probability [%]
100
80
60
40
20
0
20.00
16.00
12.00
8.00
4.00
0.00
D [-], d [mm]

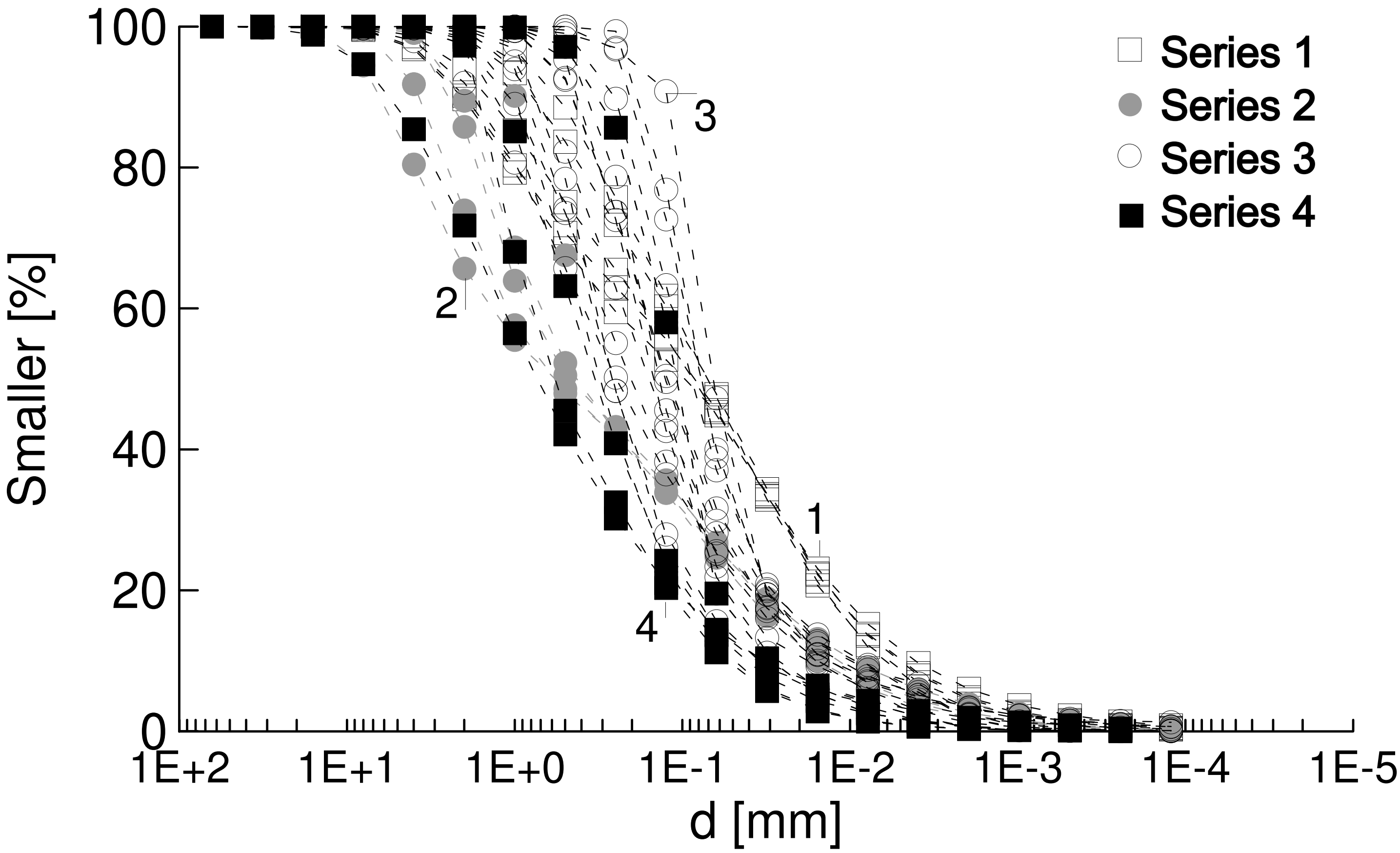

Probability [%]
20
16
12
8
4
0
20.0
16.0
12.0
8.0
4.0
0.0
D [-], d [mm]



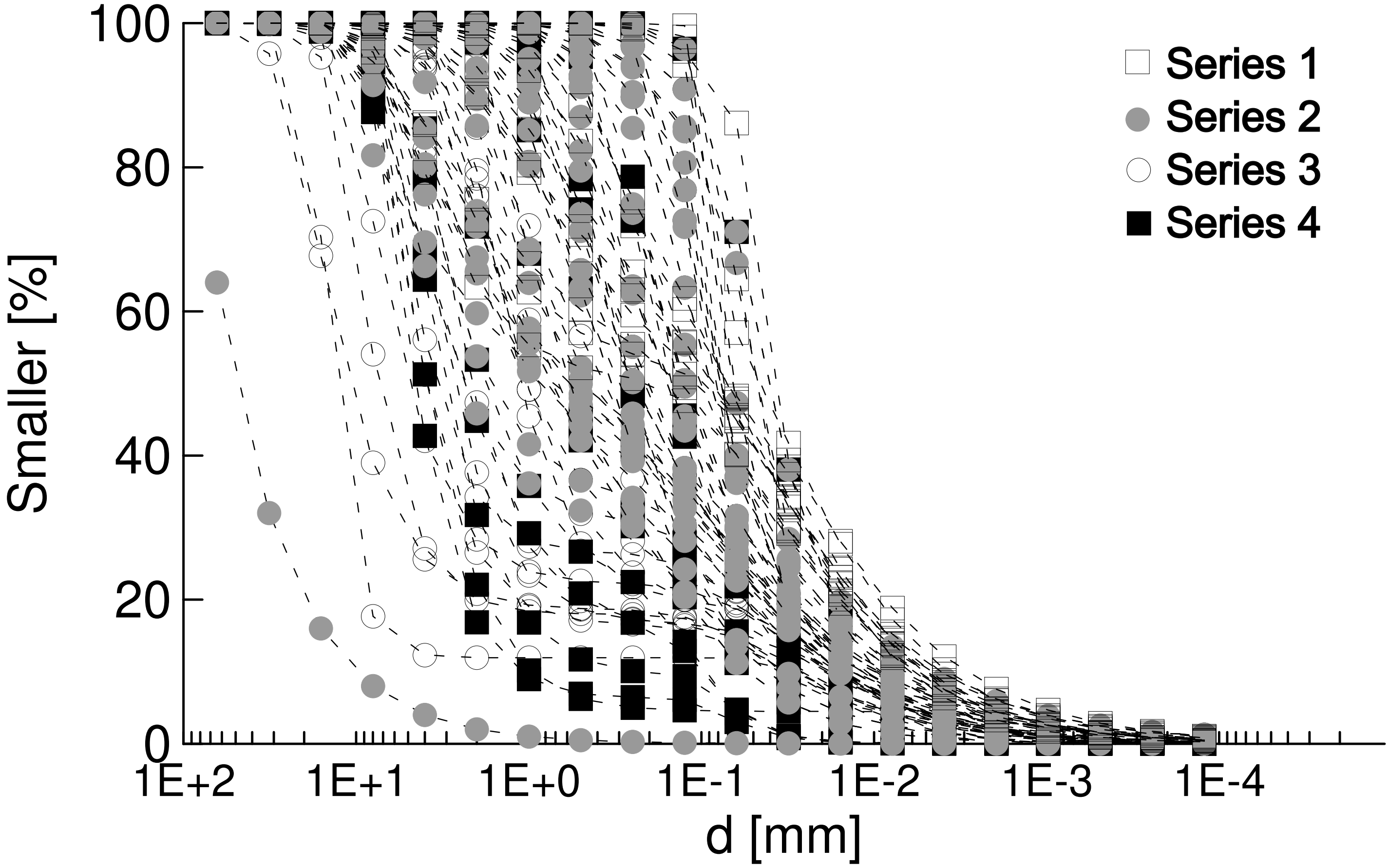
Measured
Fitted
S [-]
0.0
0.4
0.8
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
d [mm]



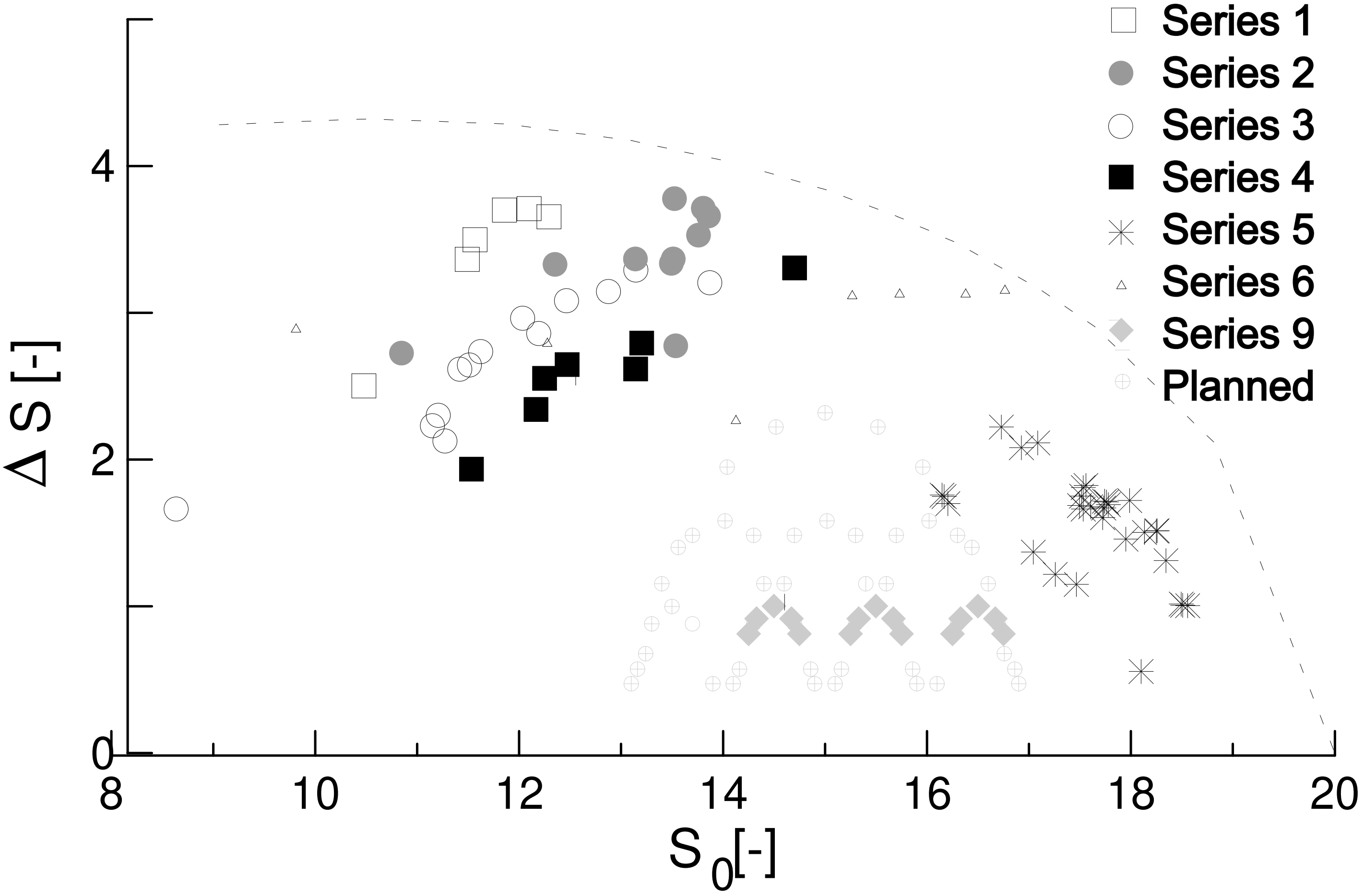
Series 1
Series 2
Series 3
Series 4
Series 5
Series 9
100.0
80.0
60.0
40.0
20.0
0.0
S [%]
1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
d [mm]



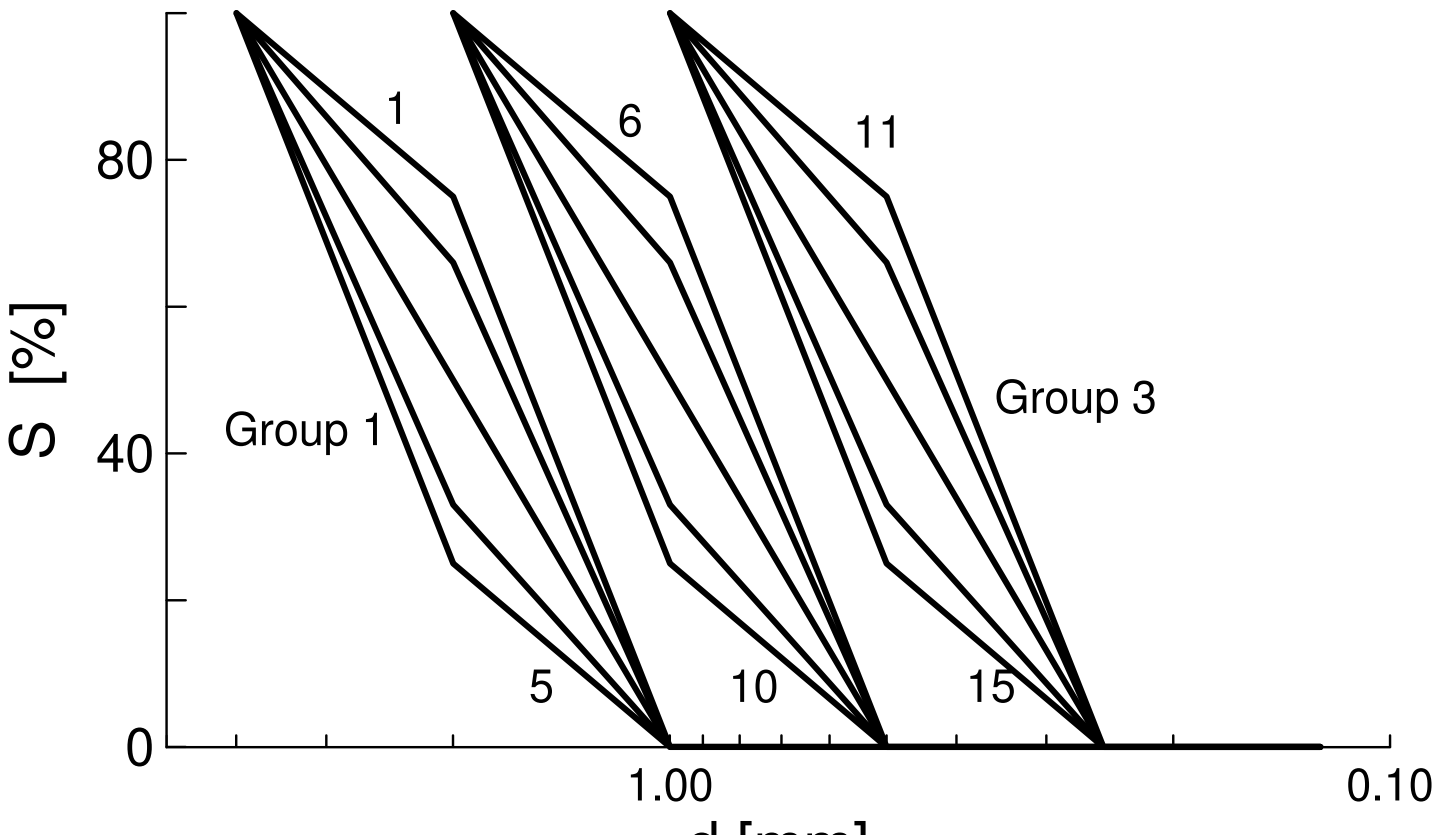
100
80
60
40
20
0
Smaller [%]
1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
d [mm]
Series 1
Series 2
Series 3
Series 4
1
2
3
4



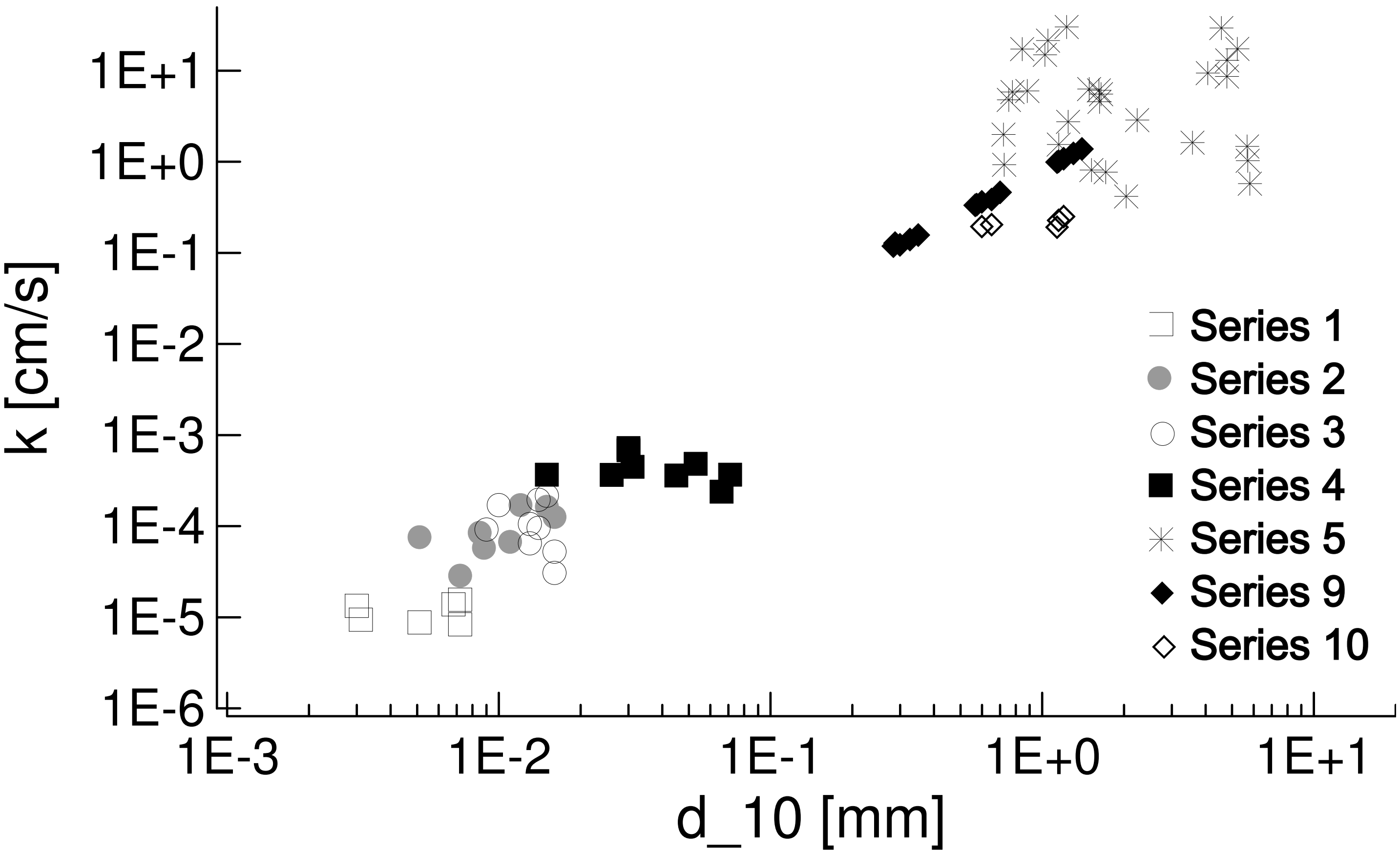
100
80
60
40
20
0
Smaller [%]
1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
d [mm]
Series 1
Series 2
Series 3
Series 4

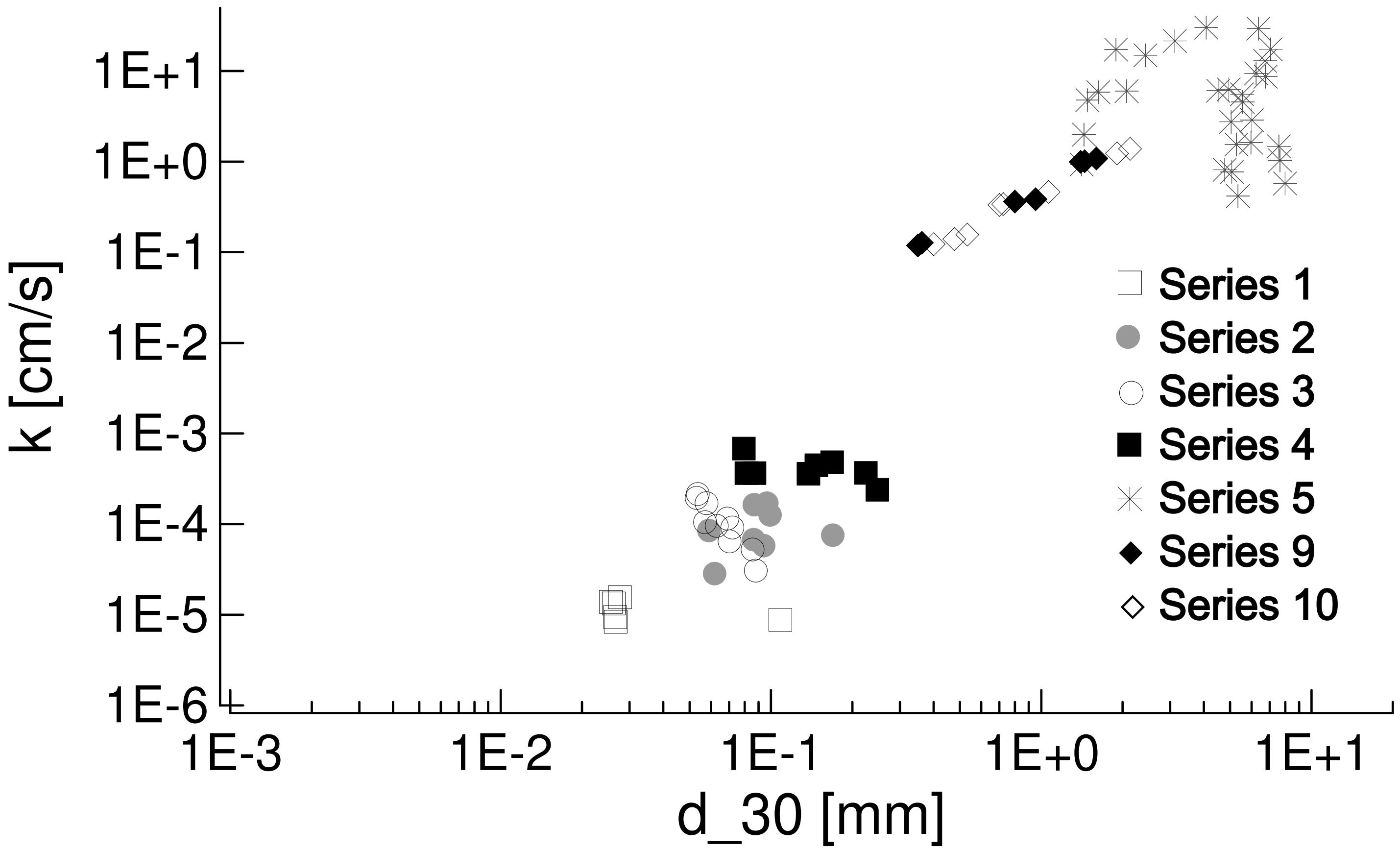


Series 1
Series 2
Series 3
Series 4
Series 5
Series 6
Series 9
Planned
Δ S [-]
S 0 [-]
0
2
4
8
10
12
14
16
18
20

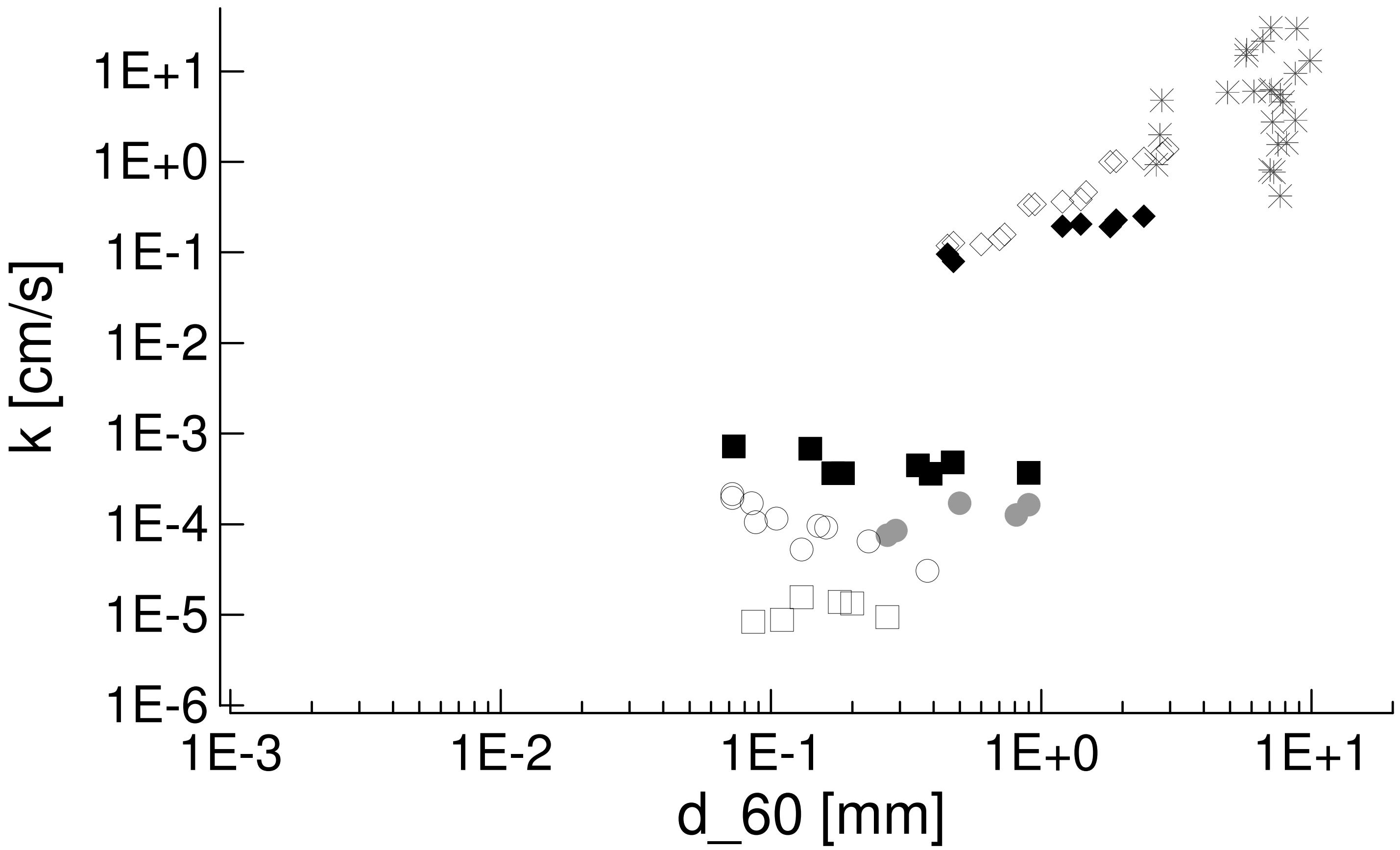


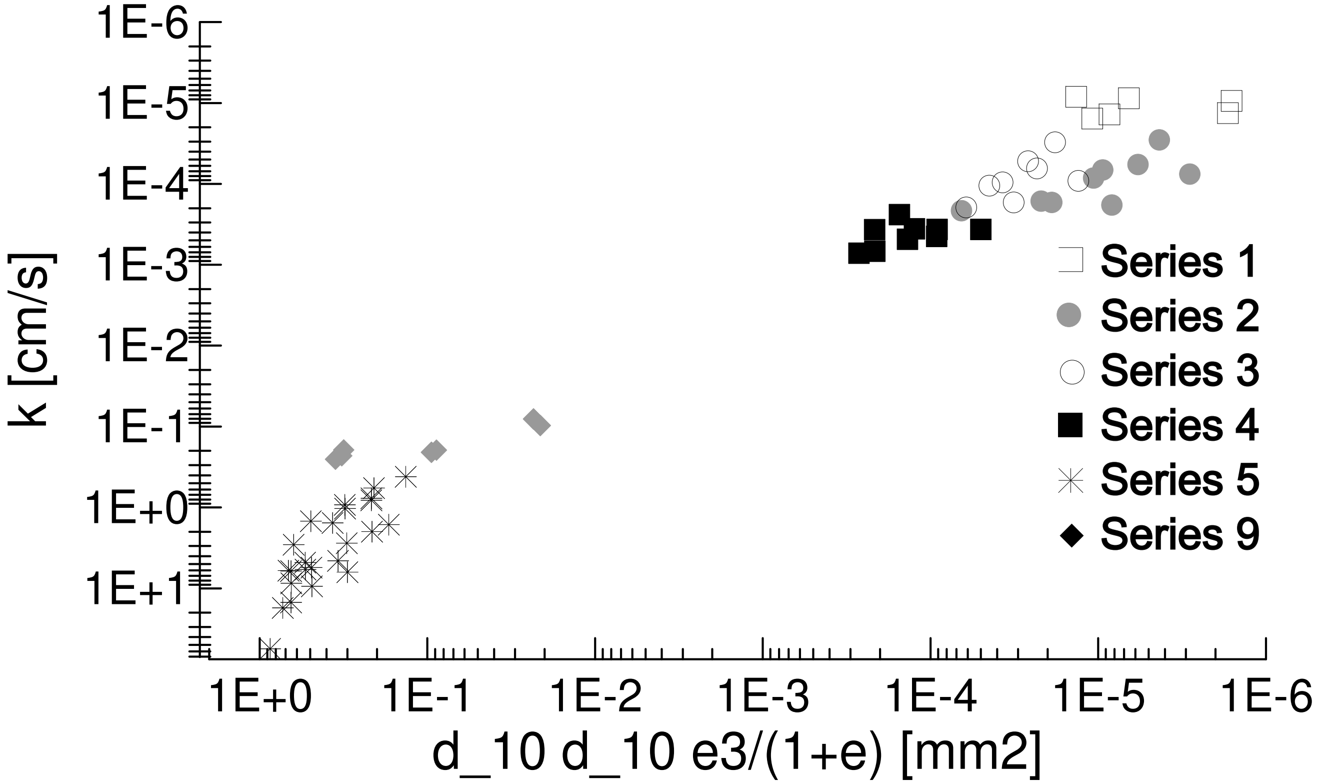
S [%]
80
40
0
1
6
11
Group 1
Group 3
5
10
15
1.00
0.10
d [mm]

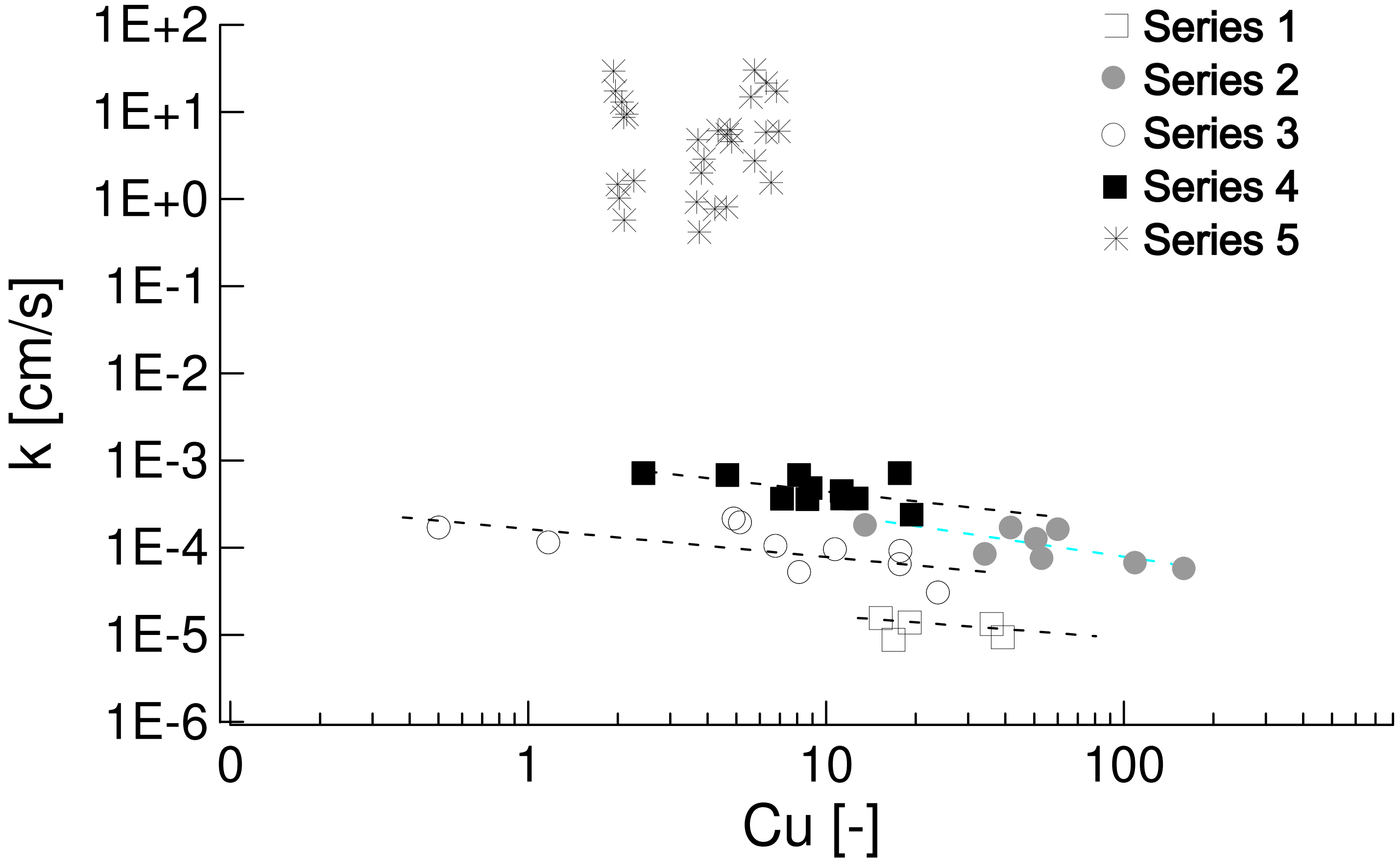

k [cm/s]
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-3
1E-2
1E-1
1E+0
1E+1
d_10 [mm]
Series 1
Series 2
Series 3
Series 4
Series 5
Series 9
Series 10



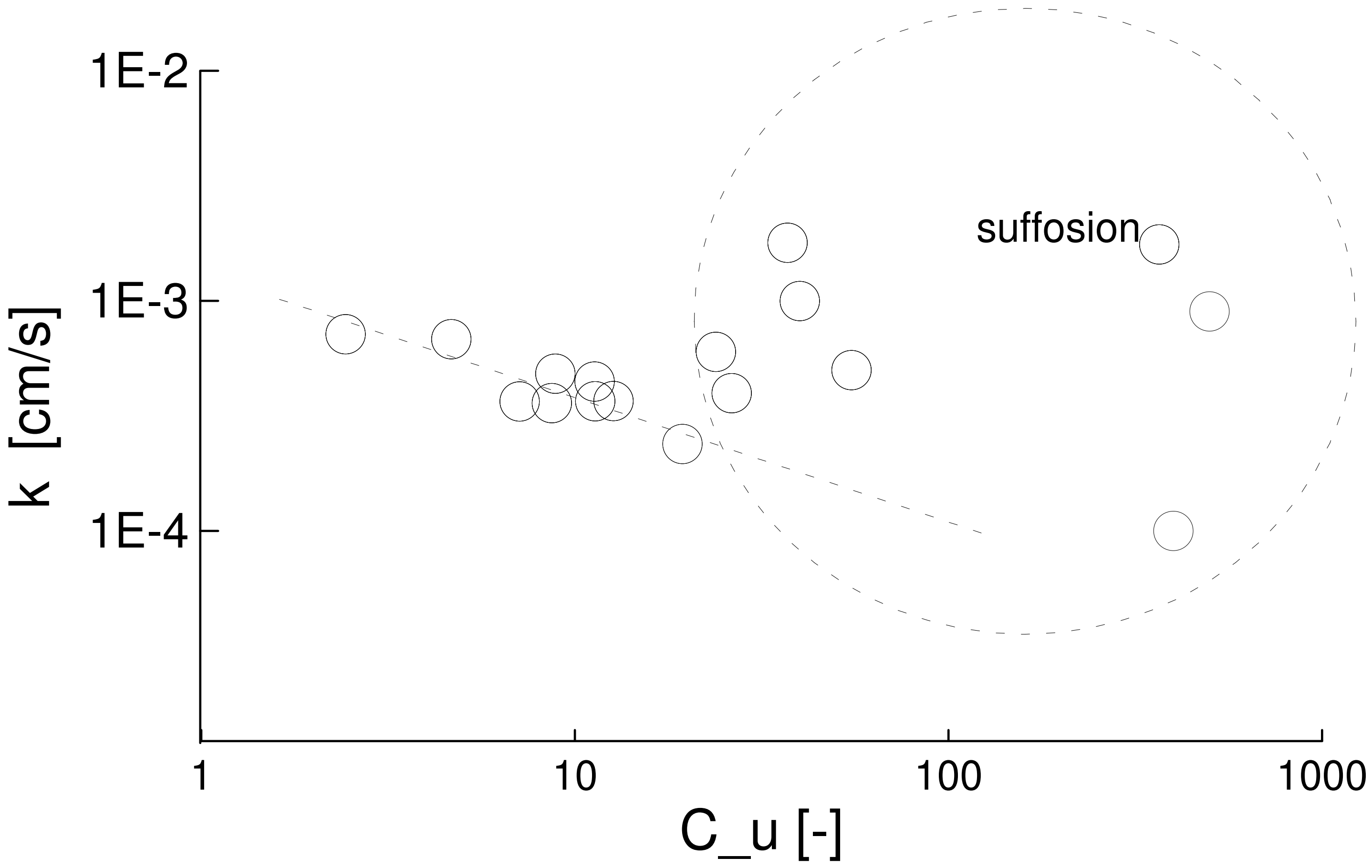
k [cm/s]
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-3
1E-2
1E-1
1E+0
1E+1
d_30 [mm]
Series 1
Series 2
Series 3
Series 4
Series 5
Series 9
Series 10


k [cm/s]
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-3
1E-2
1E-1
1E+0
1E+1
d_60 [mm]

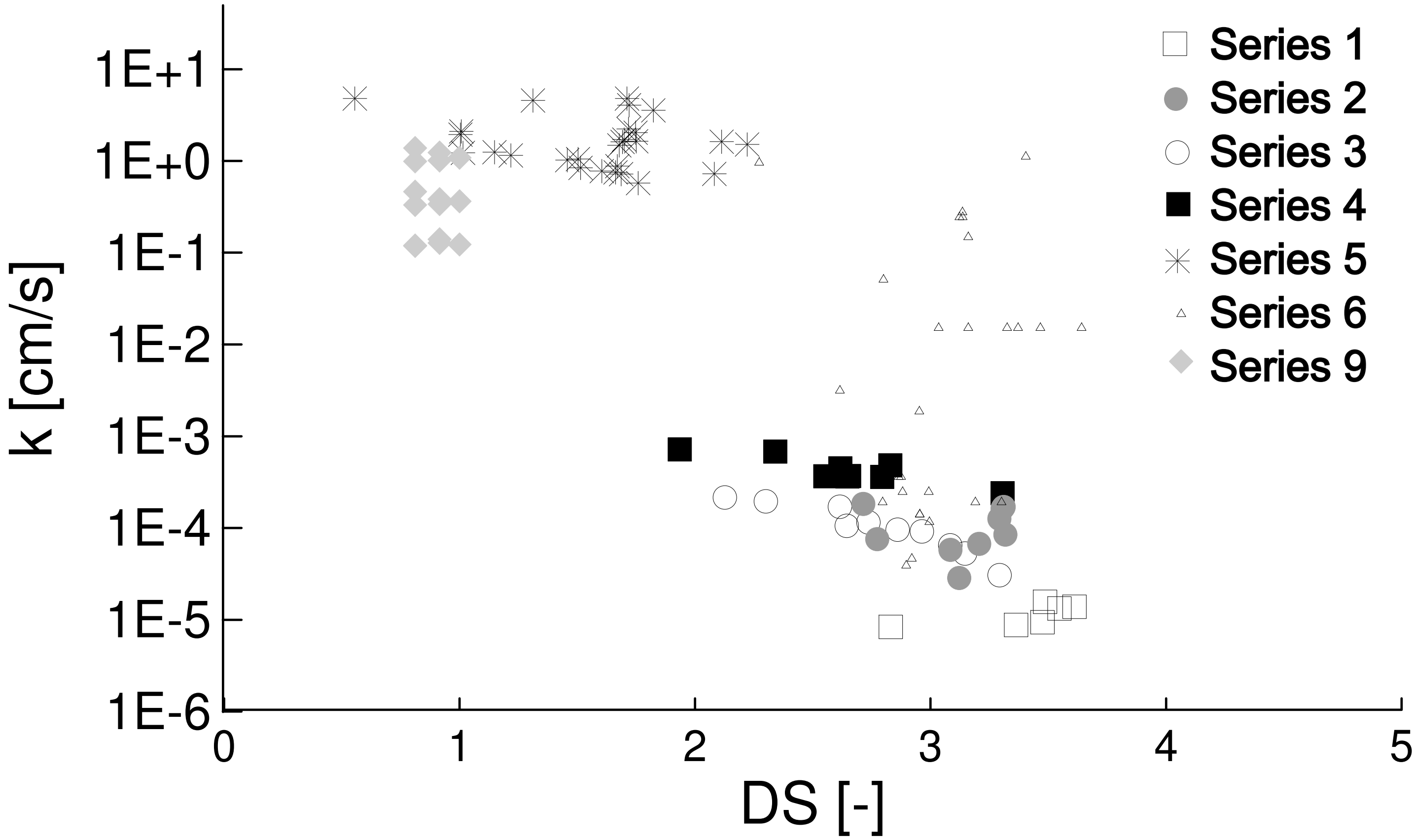

1E-6
1E-5
1E-4
1E-3
1E-2
1E-1
1E+0
1E+1
k [cm/s]
Series 1
Series 2
Series 3
Series 4
Series 5
Series 9
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
d_10 d_10 e3/(1+e) [mm2]

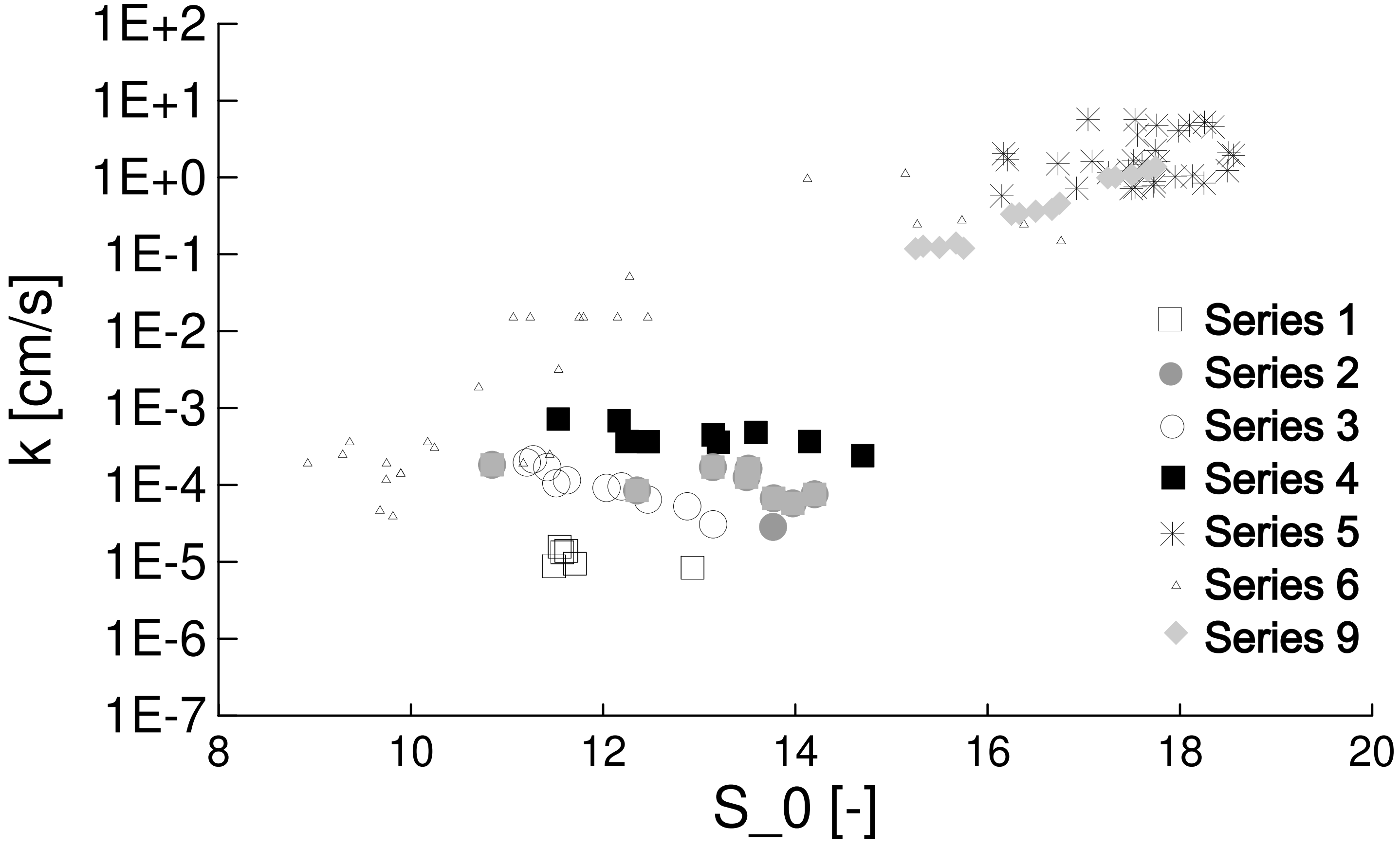


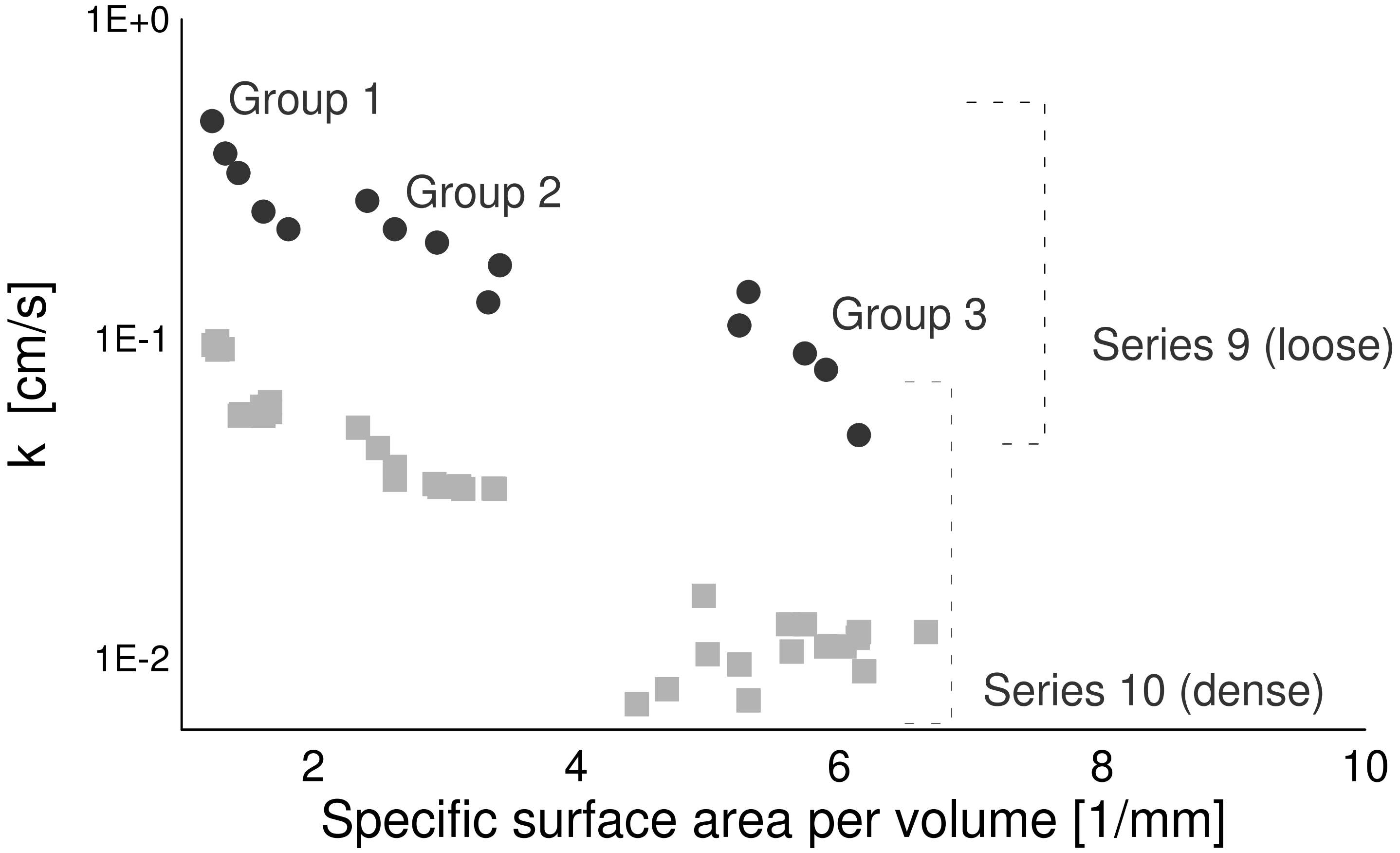
Series 1
Series 2
Series 3
Series 4
Series 5
k [cm/s]
1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
0
1
10
100
Cu [-]

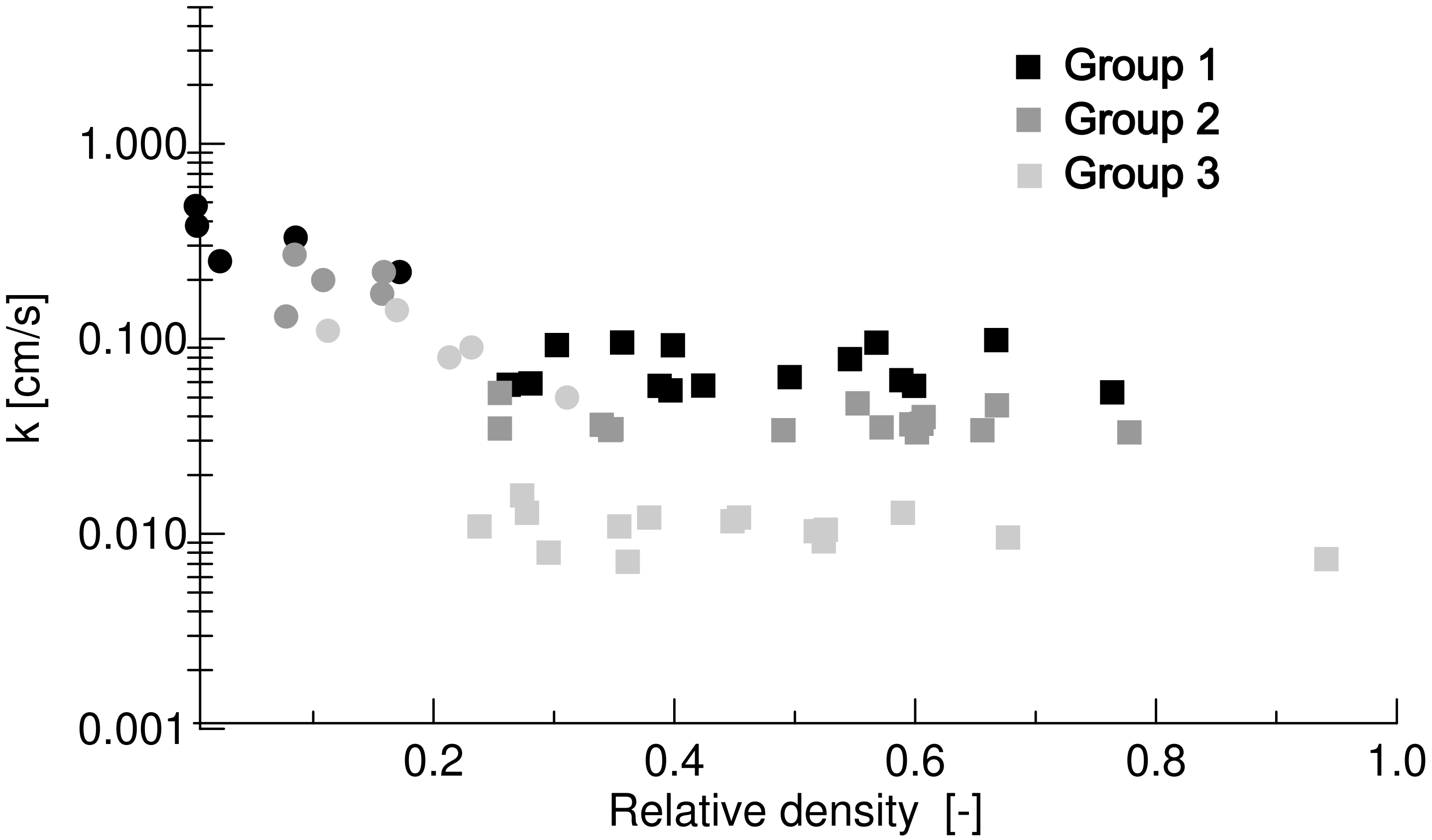


1E-2
1E-3
1E-4
k [cm/s]
1
10
100
1000
C_u [-]
suffosion

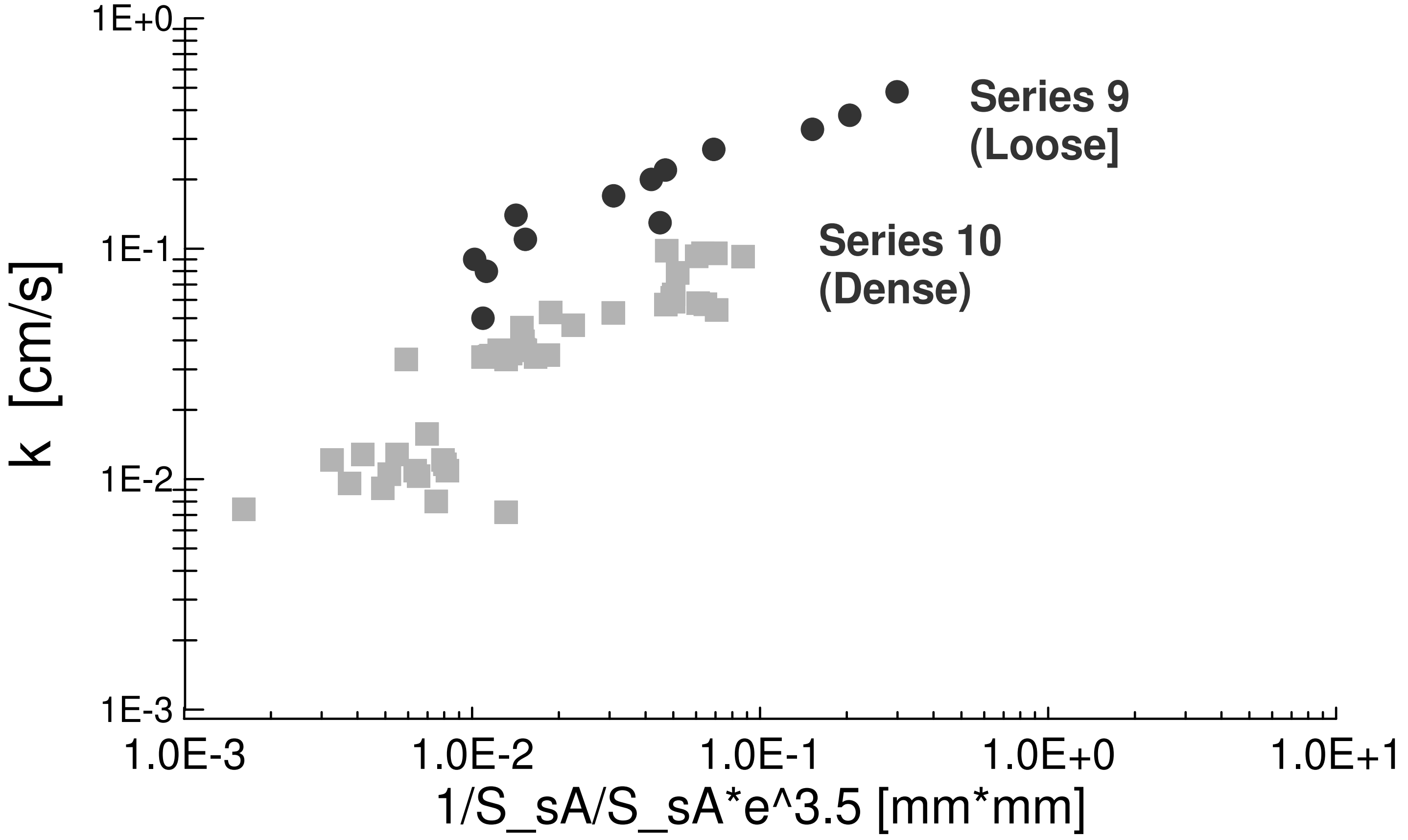

Series 1
Series 2
Series 3
Series 4
Series 5
Series 6
Series 9
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
k [cm/s]
0
1
2
3
4
5
DS [-]

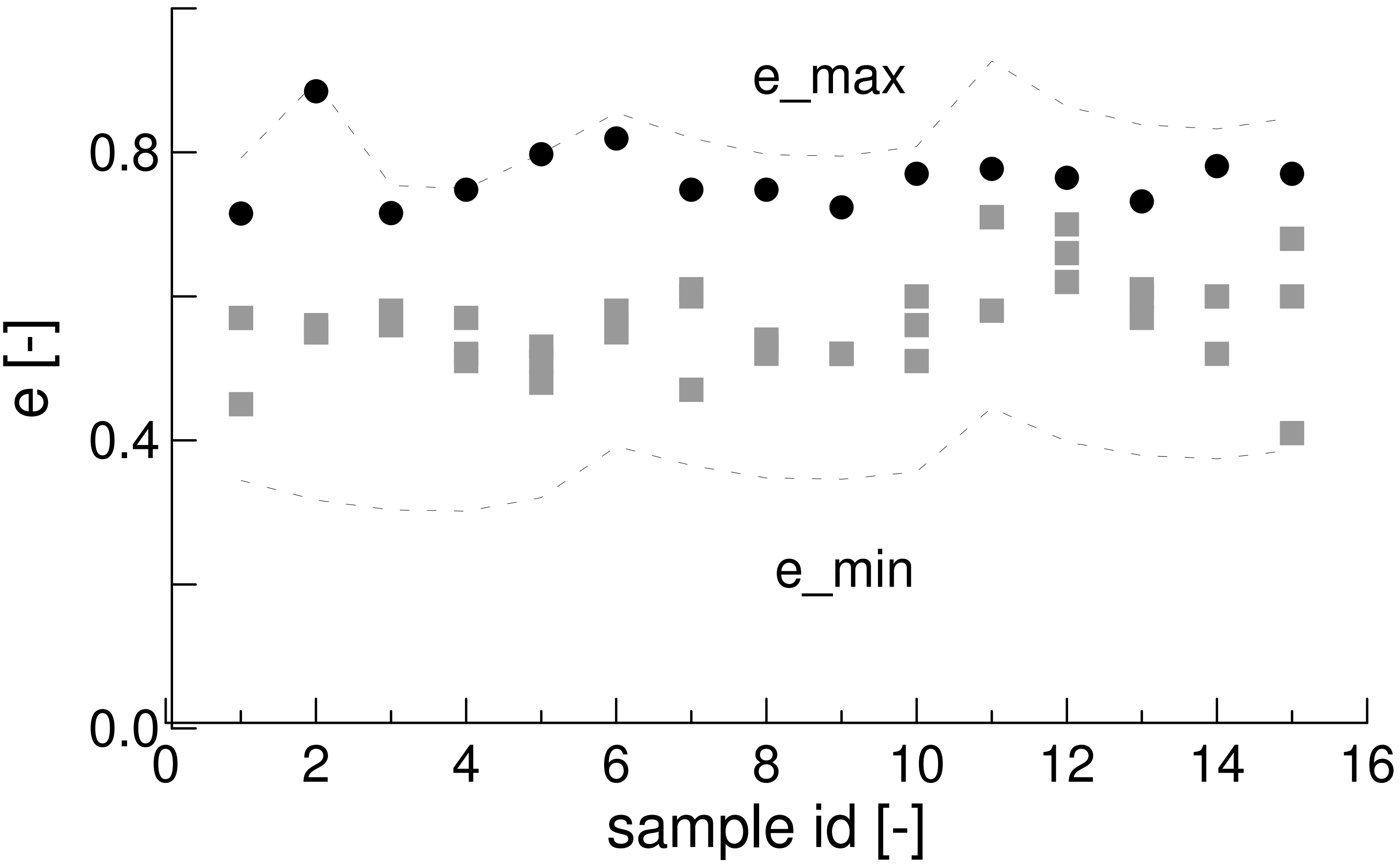

k [cm/s]
1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-7
8
10
12
14
16
18
20
S_0 [-]
Series 1
Series 2
Series 3
Series 4
Series 5
Series 6
Series 9


1E+0
1E-1
1E-2
k [cm/s]
Group 1
Group 2
Group 3
Series 9 (loose)
Series 10 (dense)
2
4
6
8
10
Specific surface area per volume [1/mm]

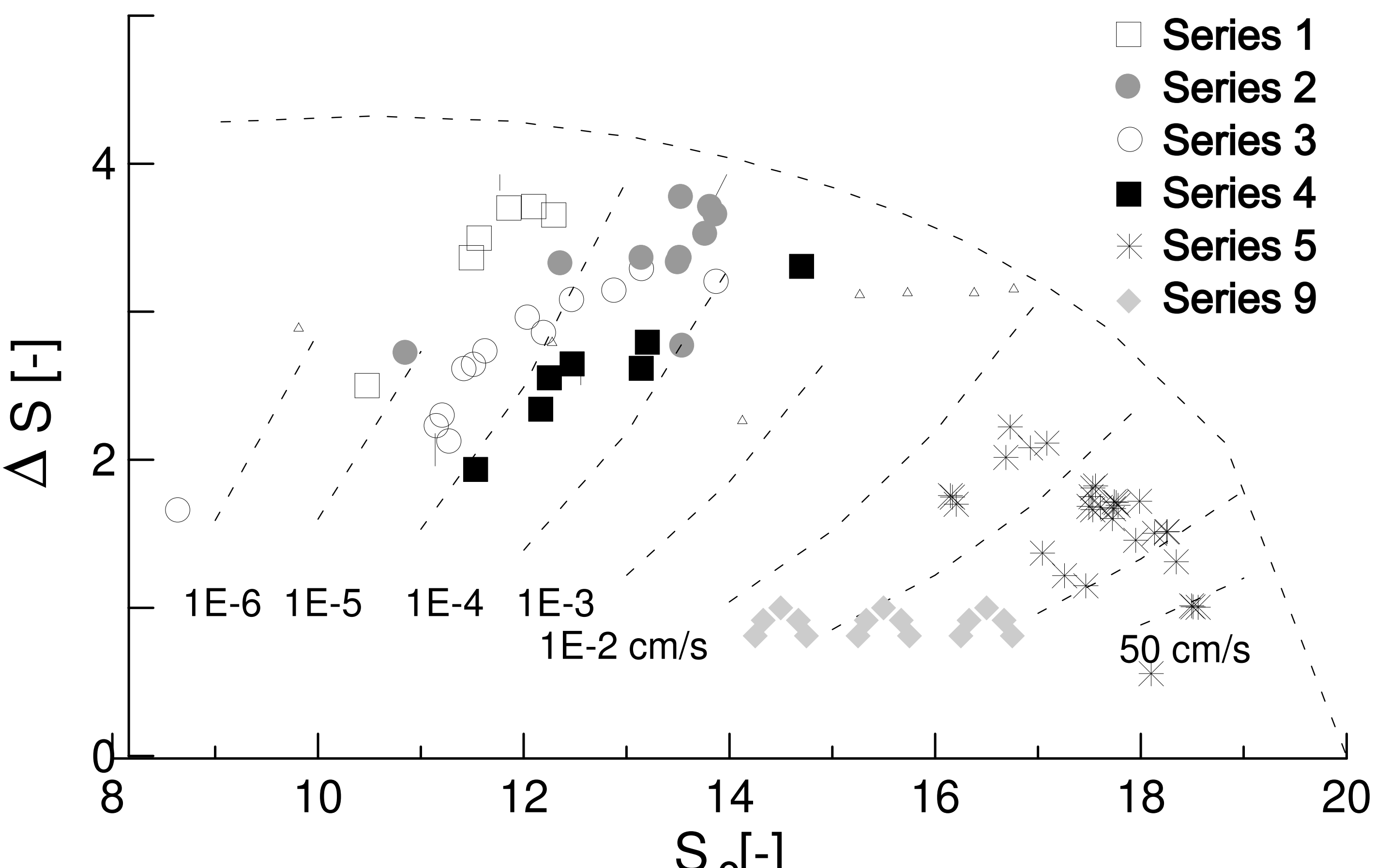


Group 1
Group 2
Group 3
1.000
0.100
0.010
0.001
k [cm/s]
0.2
0.4
0.6
0.8
1.0
Relative density [-]

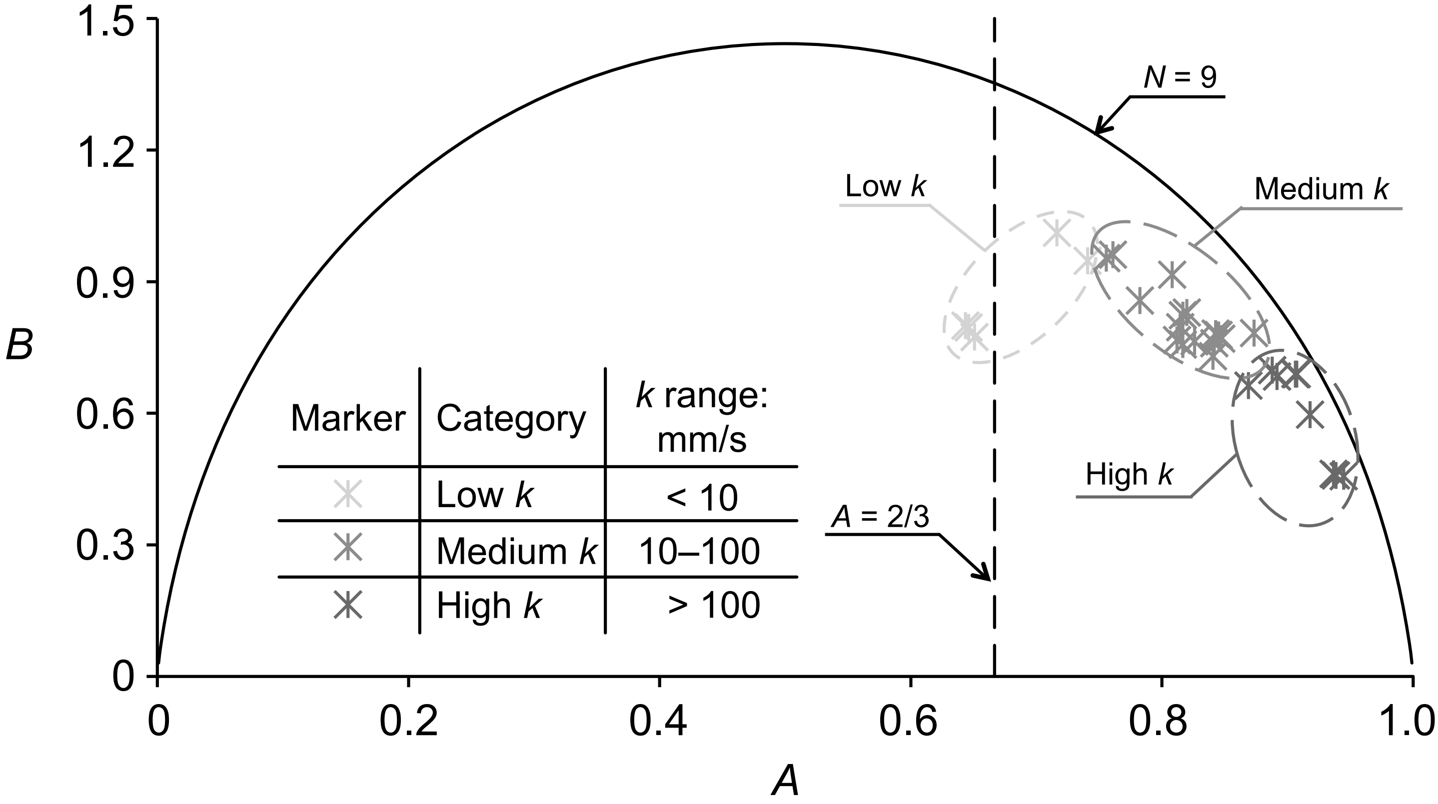

1E+0
1E-1
1E-2
1E-3
k [cm/s]
1.0E-3
1.0E-2
1.0E-1
1.0E+0
1.0E+1
1/S_sA/S_sA*e^3.5 [mm*mm]
Series 9
(Loose]
Series 10
(Dense)

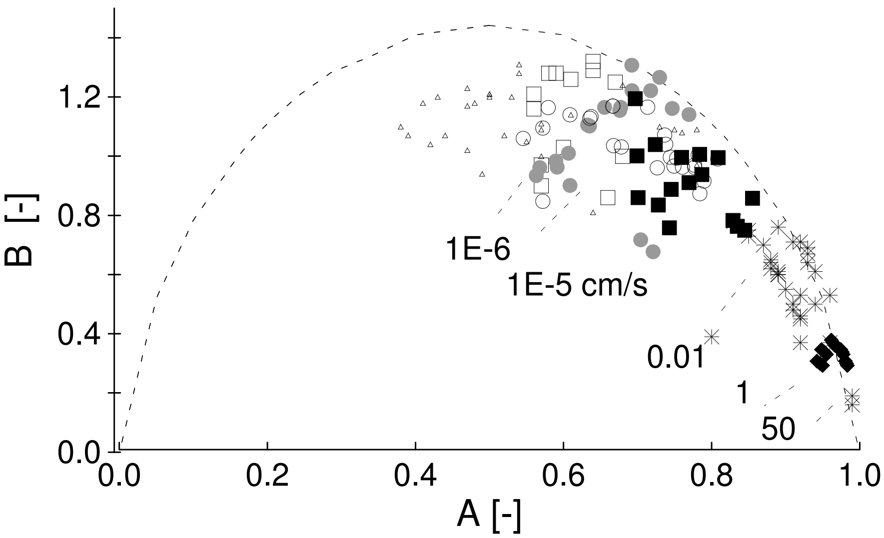


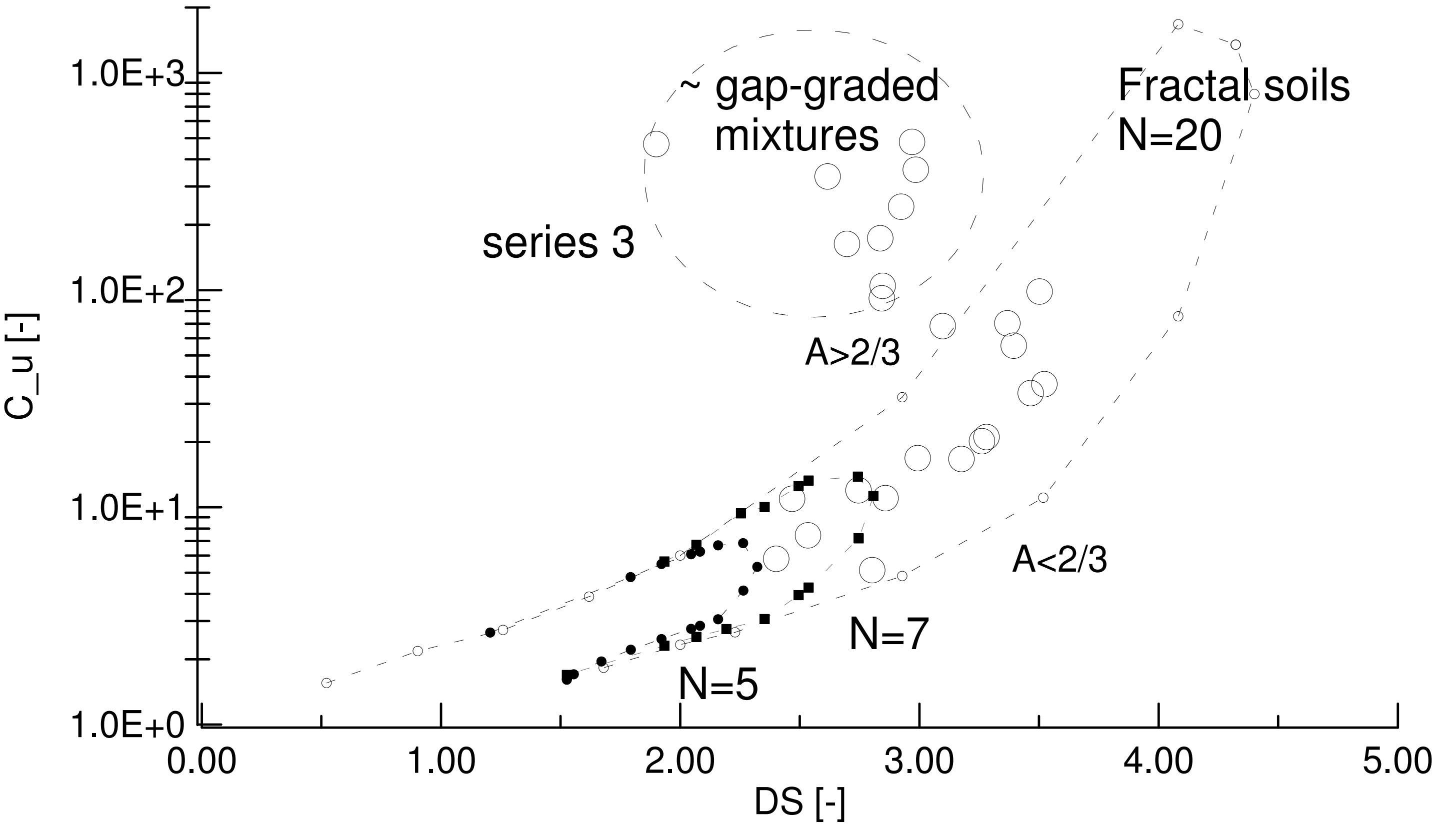
e_max
e_min
0.8
0.4
0.0
e [-]
0
2
4
6
8
10
12
14
16
sample id [-]

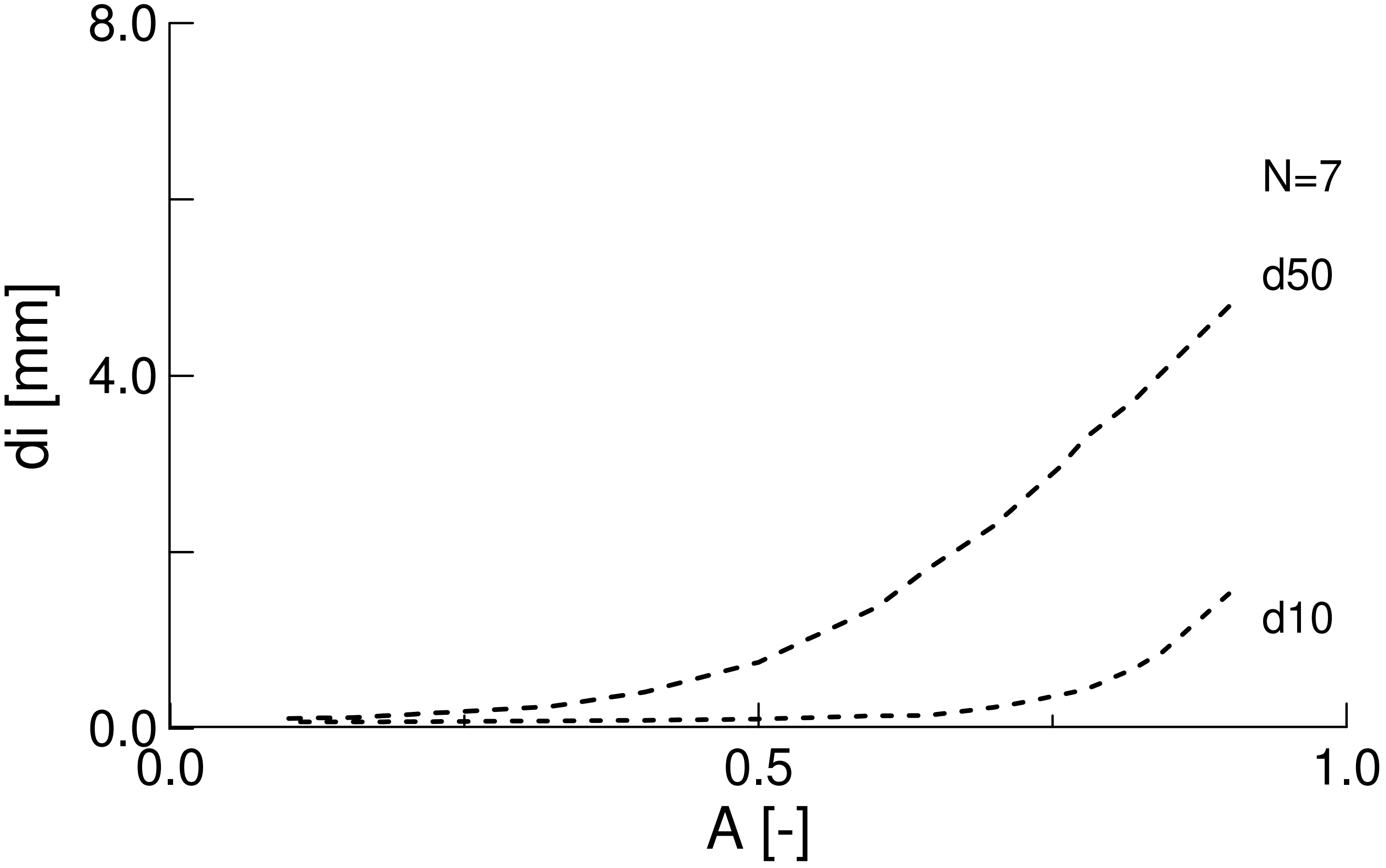


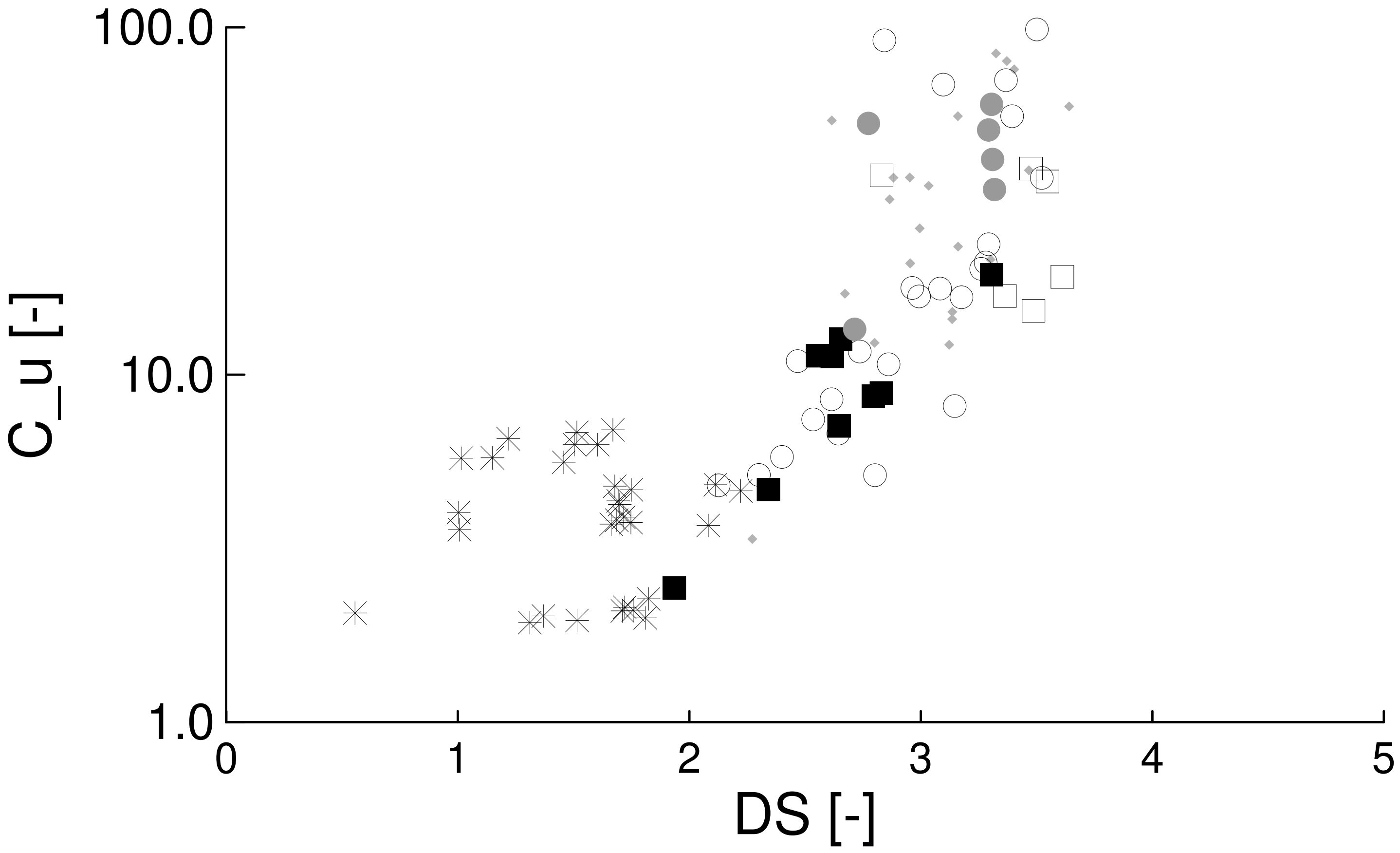
Series 1
Series 2
Series 3
Series 4
Series 5
Series 9
ΔS [-]
S0[-]
1E-6
1E-5
1E-4
1E-3
1E-2 cm/s
50 cm/s
0
2
4
8
10
12
14
16
18
20


1.5
1.2
0.9
0.6
0.3
0
B
0
0.2
0.4
0.6
0.8
1.0
A
N = 9
Low k
Medium k
High k
A = 2/3
Marker
Category
k range: mm/s
Low k
< 10
Medium k
10–100
High k
> 100

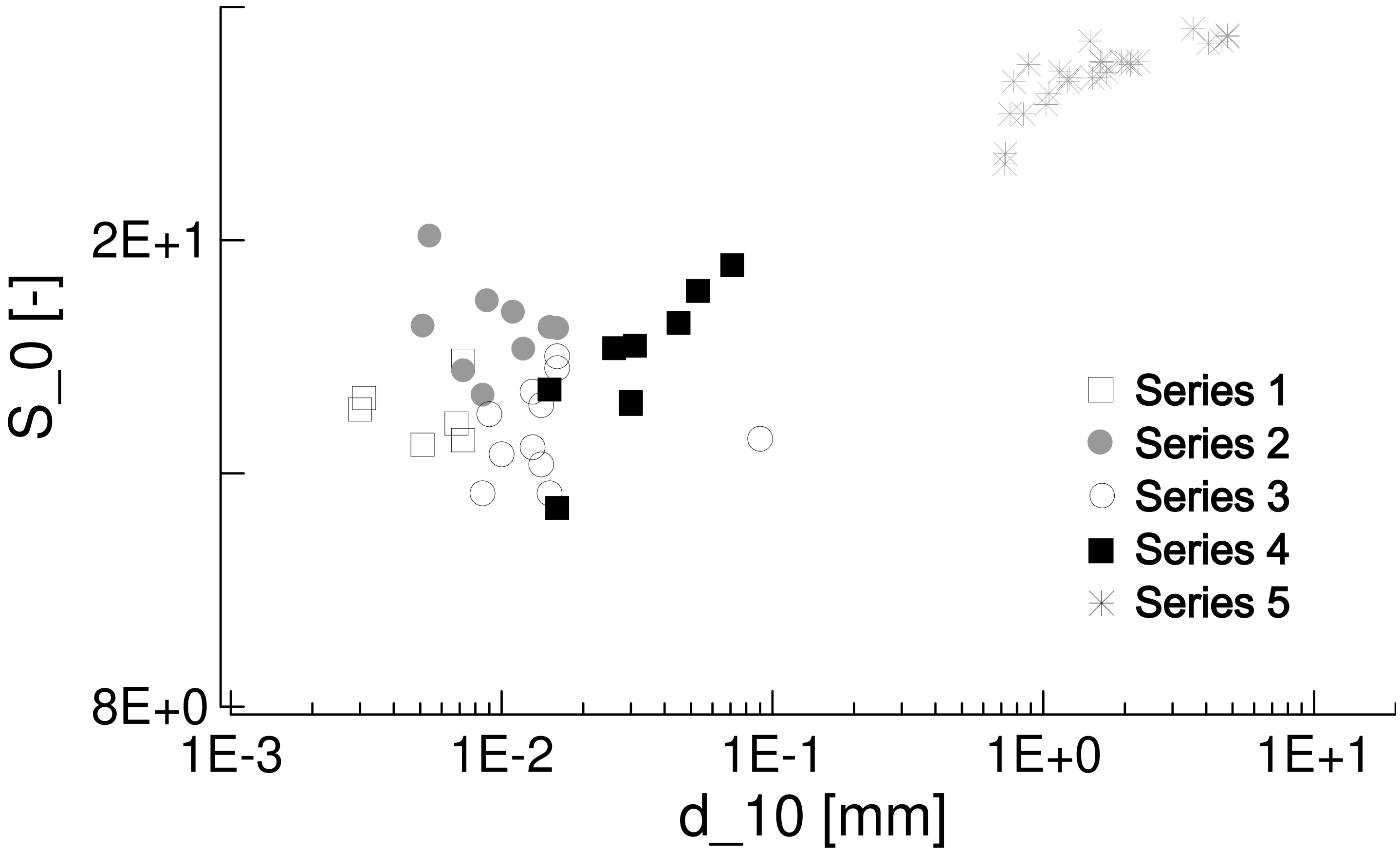

1.2
0.8
0.4
0.0
B [-]
0.0
0.2
0.4
0.6
0.8
1.0
A [-]
1E-6
1E-5 cm/s
0.01
1
50

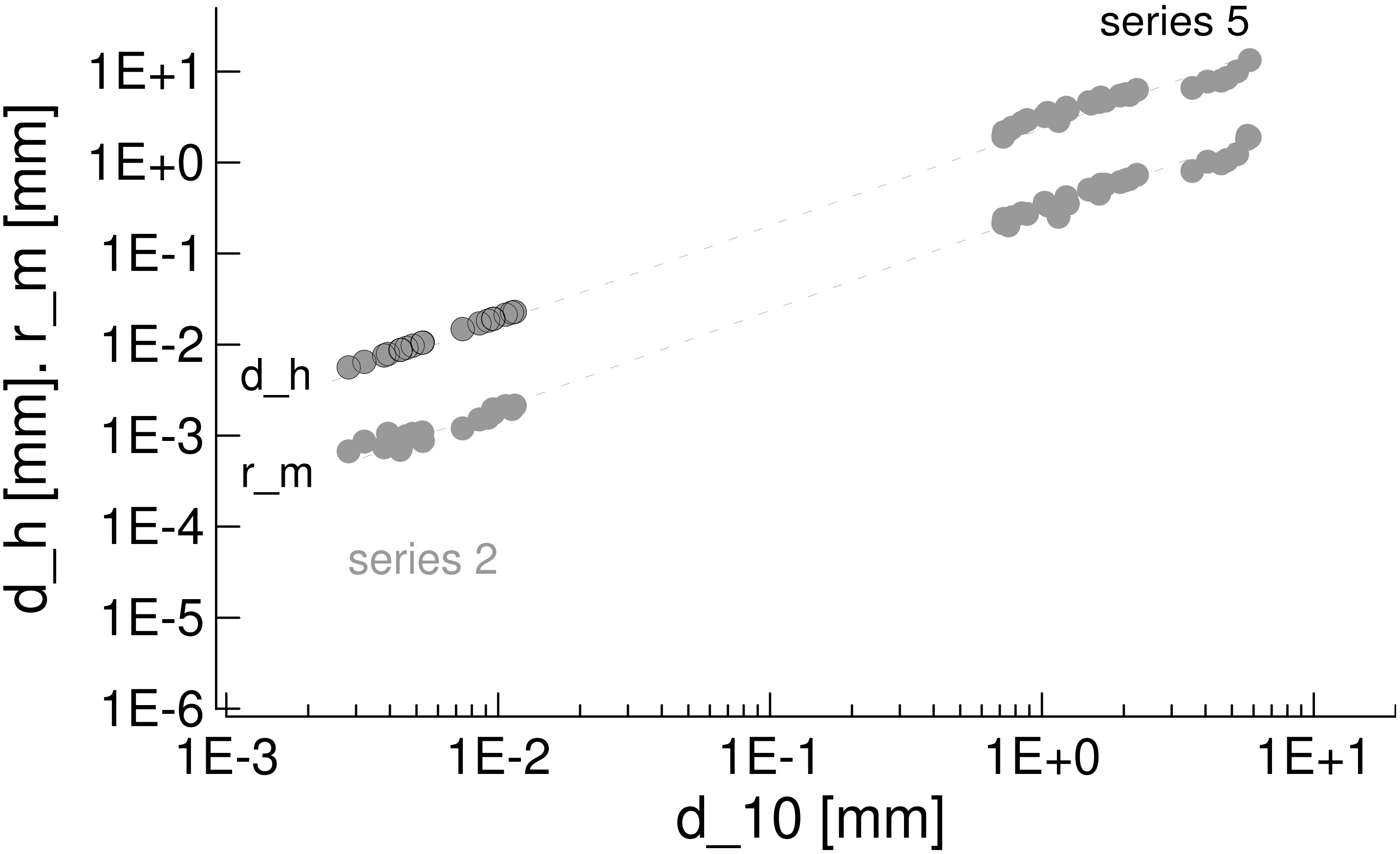

~ gap-graded
mixtures
series 3
Fractal soils
N=20
A>2/3
A<2/3
N=7
N=5
C_u [-]
DS [-]
1.0E+3
1.0E+2
1.0E+1
1.0E+0
0.00
1.00
2.00
3.00
4.00
5.00

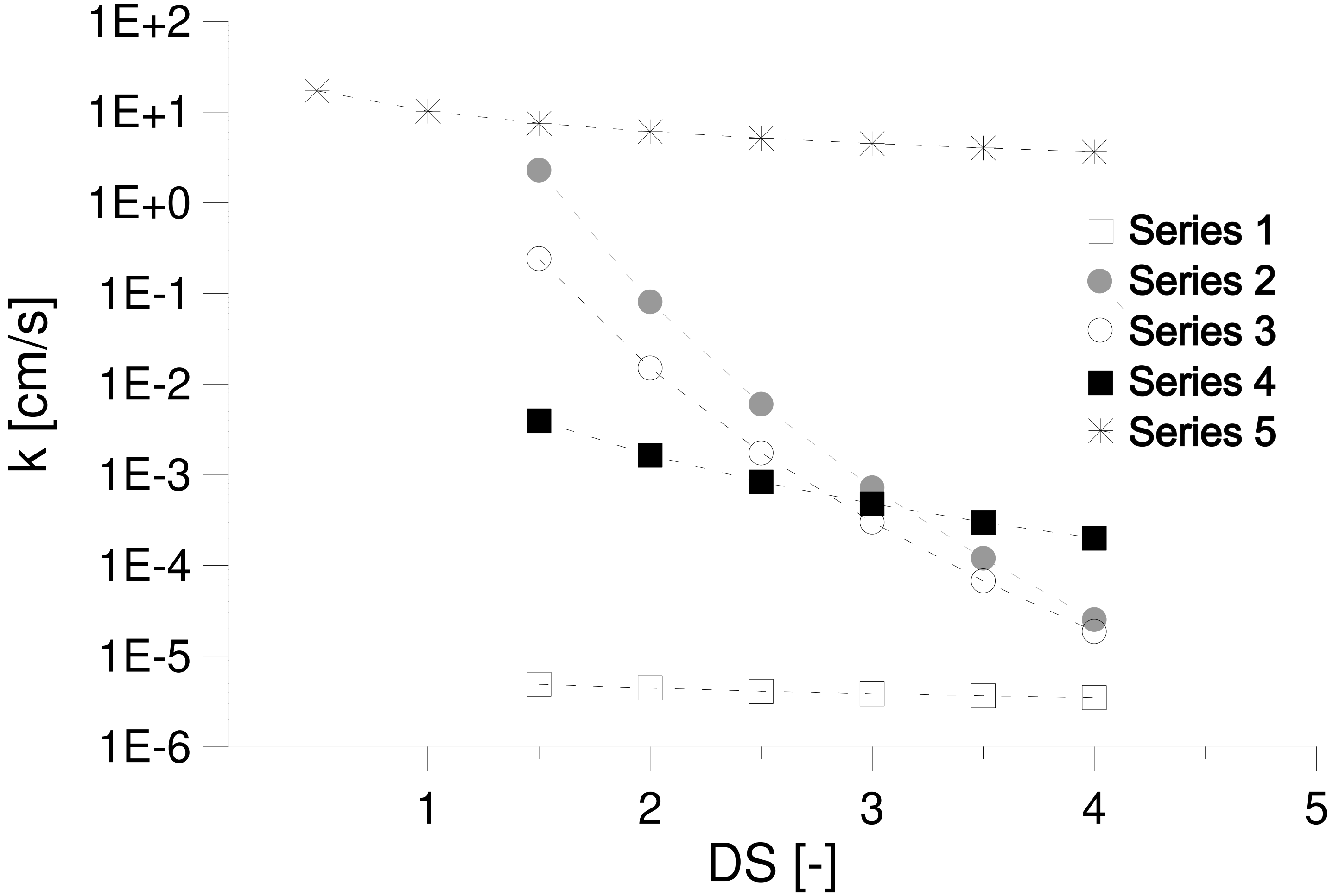

8.0
4.0
0.0
di [mm]
N=7
d50
d10
0.0
0.5
1.0
A [-]

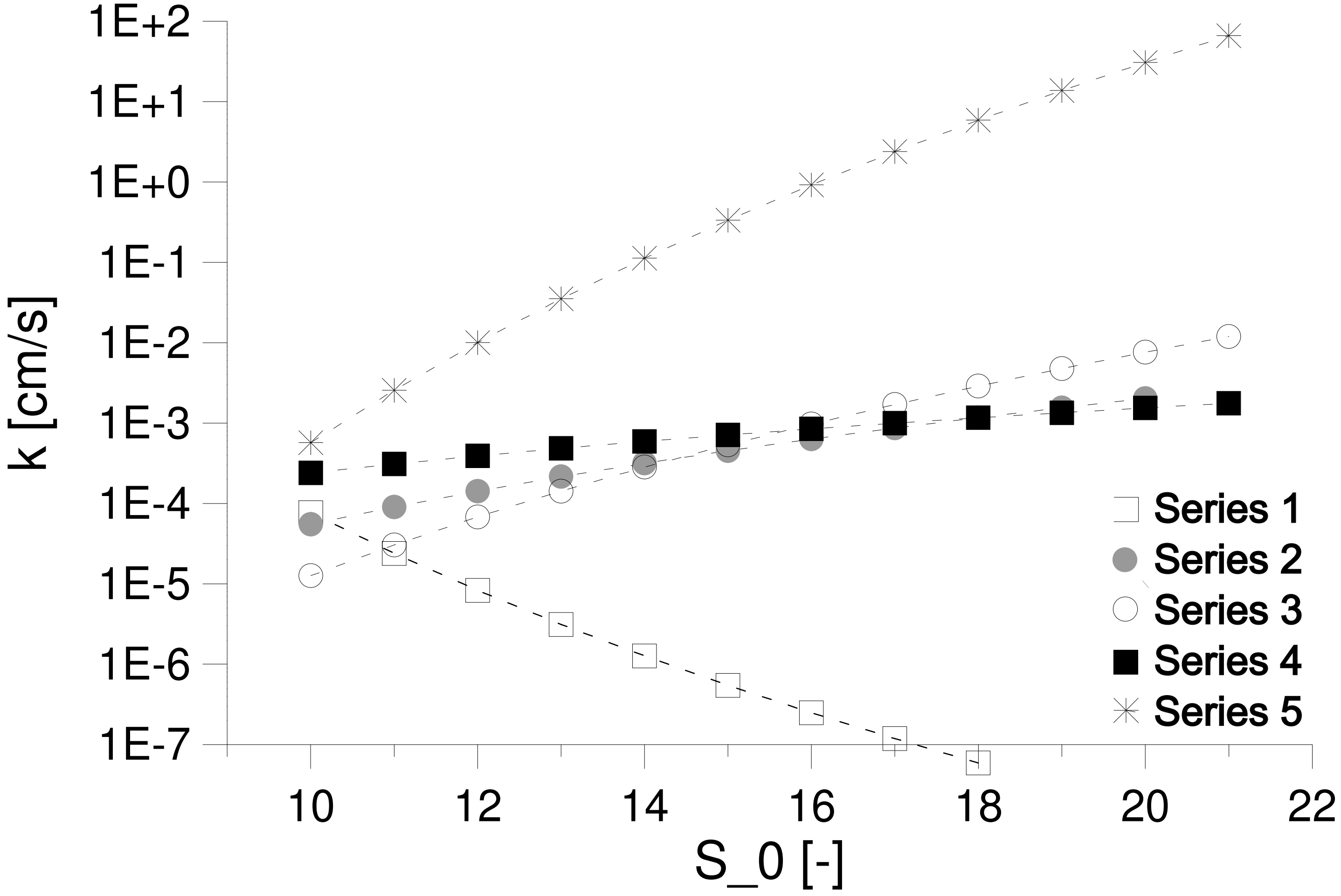


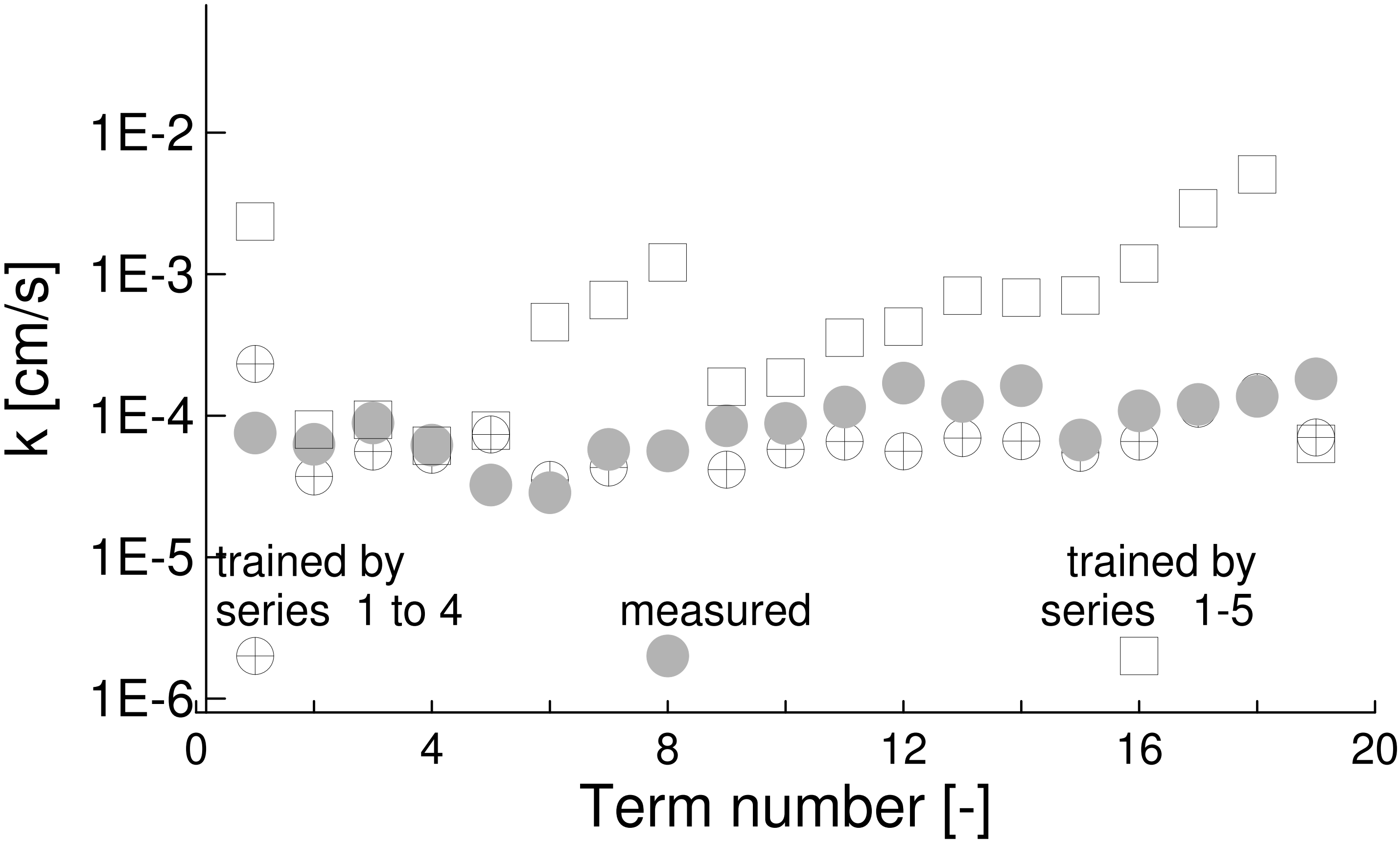
100.0
10.0
1.0
C_u [-]
0
1
2
3
4
5
DS [-]


S_0 [-]
2E+1
8E+0
1E-3
1E-2
1E-1
1E+0
1E+1
d_10 [mm]
Series 1
Series 2
Series 3
Series 4
Series 5

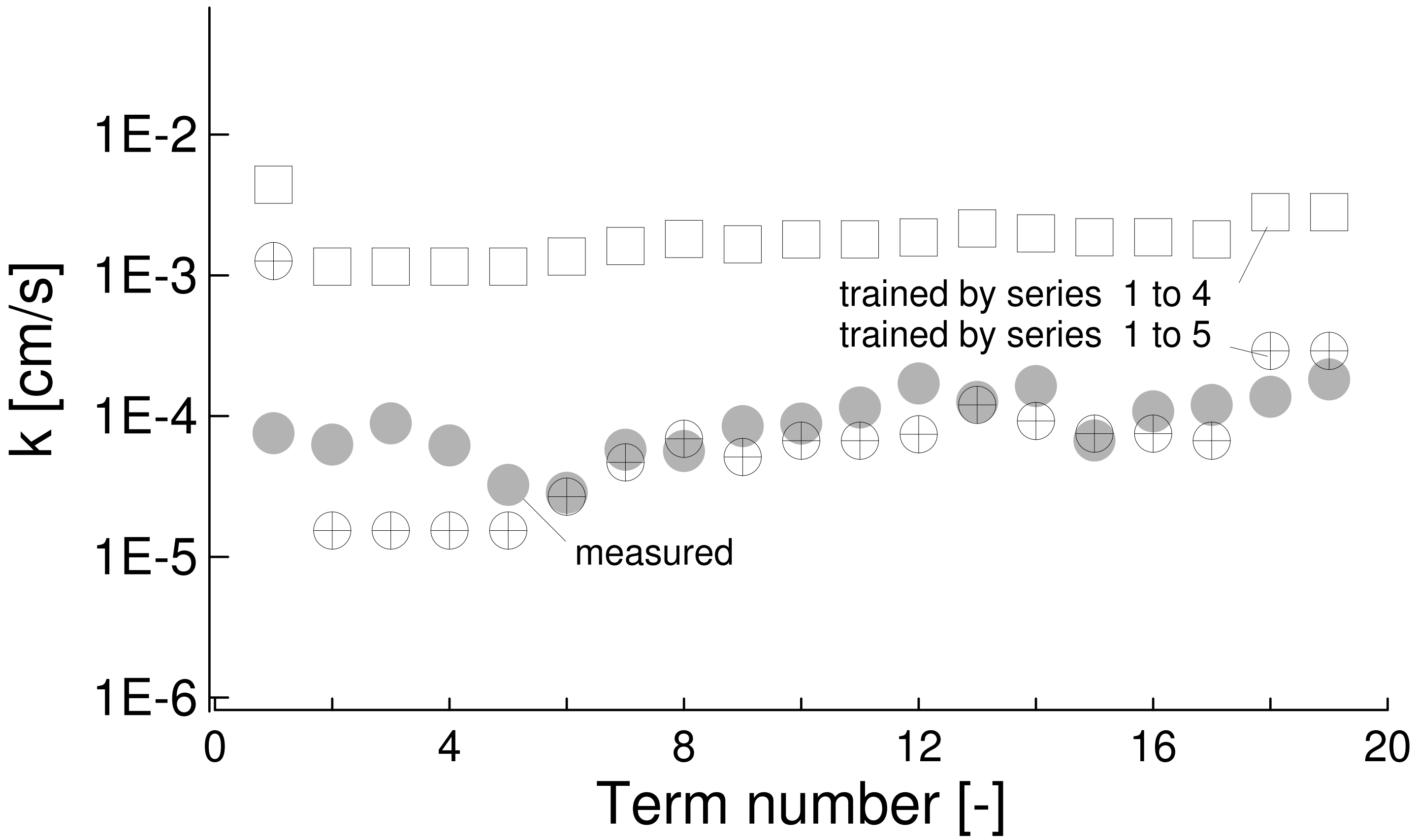

series 5
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
d_h
r_m
series 2
d_h [mm]. r_m [mm]
1E-3
1E-2
1E-1
1E+0
1E+1
d_10 [mm]

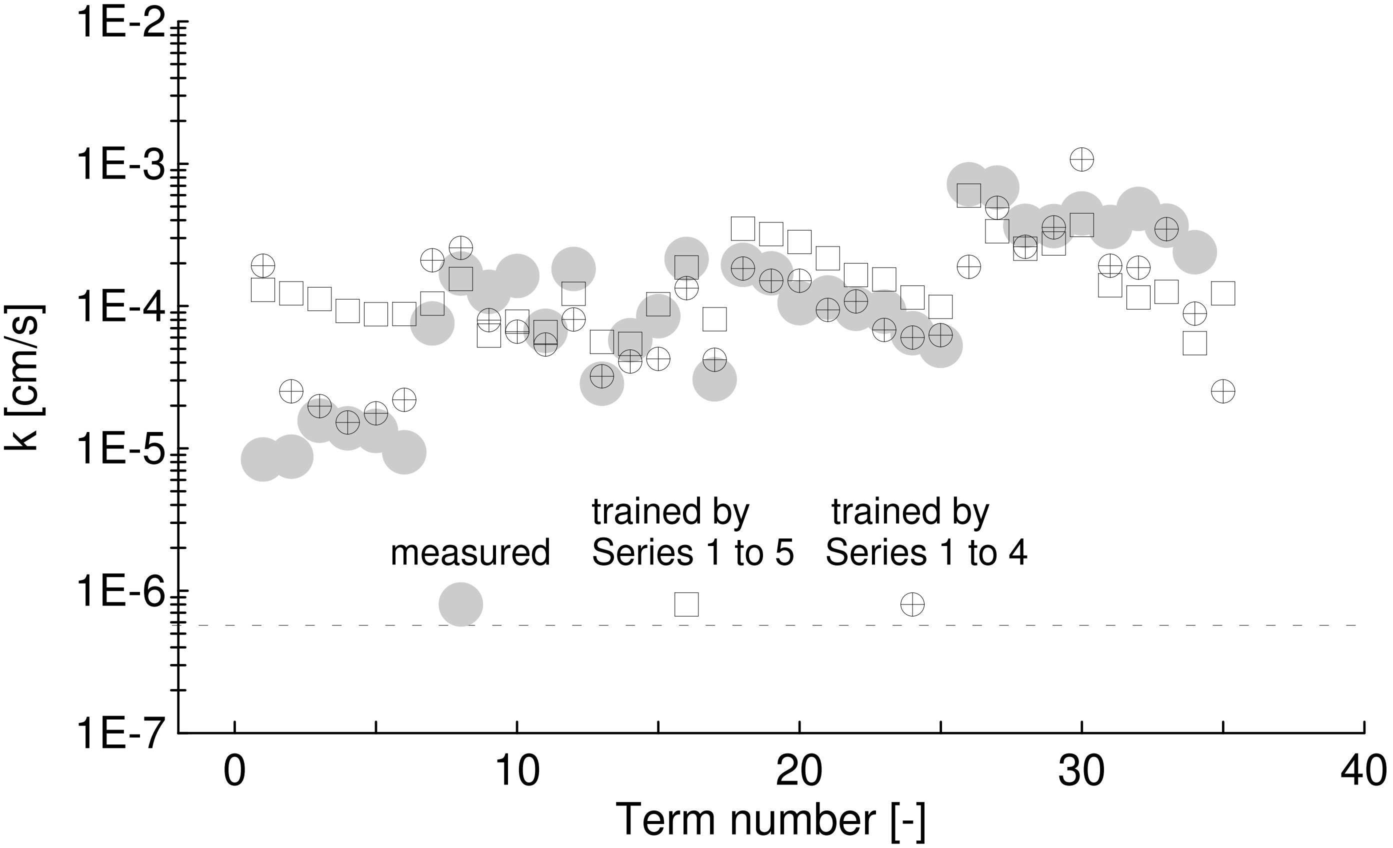

1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
k [cm/s]
1
2
3
4
5
DS [-]
Series 1
Series 2
Series 3
Series 4
Series 5


1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-7
k [cm/s]
10
12
14
16
18
20
22
S_0 [-]
Series 1
Series 2
Series 3
Series 4
Series 5

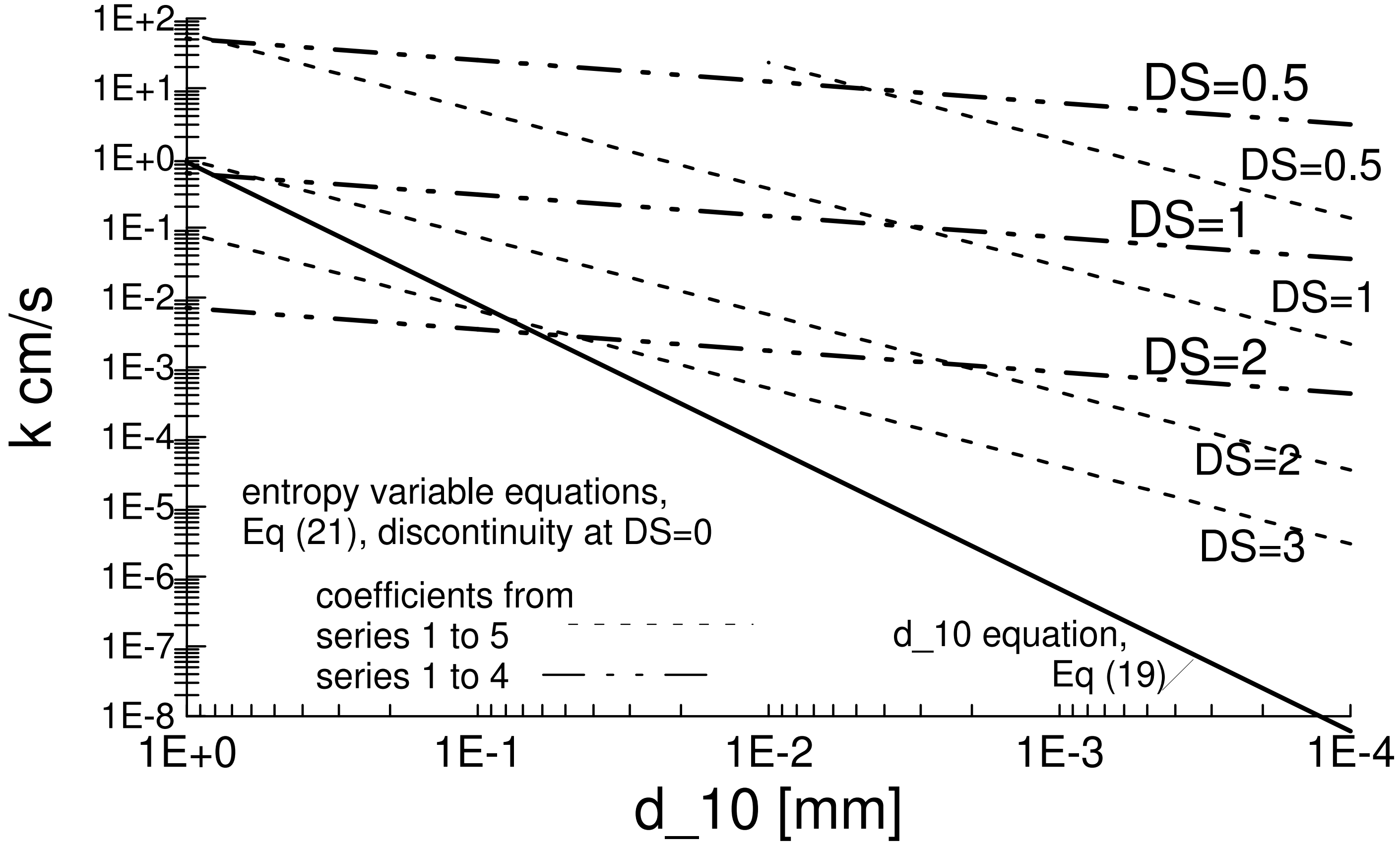


1E-2
1E-3
1E-4
1E-5
1E-6
k [cm/s]
trained by
series 1 to 4
measured
trained by
series 1-5
0
4
8
12
16
20
Term number [-]


1E-2
1E-3
1E-4
1E-5
1E-6
k [cm/s]
0
4
8
12
16
20
Term number [-]
trained by series 1 to 4
trained by series 1 to 5
measured



1E-2
1E-3
1E-4
1E-5
1E-6
1E-7
k [cm/s]
0
10
20
30
40
Term number [-]
measured
trained by
Series 1 to 5
trained by
Series 1 to 4



1E+2
1E+1
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
1E-6
1E-7
1E-8
k cm/s
DS=0.5
DS=0.5
DS=1
DS=1
DS=2
DS=2
DS=3
entropy variable equations,
Eq (21), discontinuity at DS=0
coefficients from
series 1 to 5
series 1 to 4
d_10 equation,
Eq (19)
1E+0
1E-1
1E-2
1E-3
1E-4
d_10 [mm]