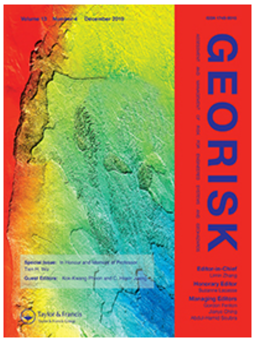

Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards

ISSN: 1749-9518 (Print) 1749-9526 (Online) Journal homepage: https://www.tandfonline.com/loi/ngrk20

A database of saturated hydraulic conductivity of fine-grained soils: probability density functions

Shuyin Feng & Paul J. Vardanega

To cite this article: Shuyin Feng & Paul J. Vardanega (2019) A database of saturated hydraulic conductivity of fine-grained soils: probability density functions, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 13:4, 255-261, DOI: 10.1080/17499518.2019.1652919

To link to this article: https://doi.org/10.1080/17499518.2019.1652919

View supplementary material

Published online: 15 Aug 2019.

Submit your article to this journal

Article views: 196

View related articles

View Crossmark data

Citing articles: 5 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ngrk20

GEORISK
2019, VOL. 13, NO. 4, 255–261
https://doi.org/10.1080/17499518.2019.1652919

Taylor & Francis
Taylor & Francis Group

Check for updates

A database of saturated hydraulic conductivity of fine-grained soils: probability density functions

Shuyin Feng and Paul J. Vardanega

Department of Civil Engineering, University of Bristol, Bristol, UK

ABSTRACT
Saturated hydraulic conductivity is a key soil mechanics parameter which has widespread use in many geotechnical applications. In order to set up stochastic analyses, geotechnical modellers require databases to calibrate the parameter ranges and distributions employed. This letter uses a recently compiled database of saturated hydraulic conductivity measurements called FG/KSAT-1358 and reports on the fitting of various probability density functions to the data of void ratio, liquid limit, water content ratio and the negative natural logarithm of k_{sat}. It is shown that the best fit distribution is the lognormal for void ratio, while the loglogistic distribution is most favoured for liquid limit and water content ratio, and the best fit distribution for -ln[k_{sat}(m/s)] is the logistic function. The data of -ln[k_{sat}(m/s)] is then subdivided according to liquid limit level, silt or clay classification, type of hydraulic conductivity test used and sample preparation/condition. When some subdivisions of the database are analysed, the best fit distribution is more variable with GEV and logistic being the most favoured for most of the studied subsets.

ARTICLE HISTORY
Received 22 February 2019
Accepted 4 August 2019

KEYWORDS
Saturated hydraulic conductivity; probability density functions; Akaike information criterion; corrected Akaike information criterion

1. Introduction

1.1. Saturated hydraulic conductivity

Saturated hydraulic conductivity (k_{sat}) is of key importance in many geotechnical designs (e.g. slopes, waste disposal facilities, road construction and foundation design). The ability to model potential variations in saturated hydraulic conductivity is important for those wishing to perform stochastic modelling of, e.g. slopes in the humid tropics (Almeida et al. 2017; Shepheard et al. 2018). To select parameter ranges and distributions for use in such modelling geo-databases are needed. Geo-databases are commonly employed in geotechnics to make a-priori assessments of more complex parameters from more readily obtainable ones (e.g. Kulhawy and Mayne 1990).

There are many empirical and semi-empirical approaches in the literature to model saturated hydraulic conductivity, e.g. "Hazen" (Hazen 1893, 1895, 1911); the "Kozeny-Carman" (Kozeny 1927; Carman 1937, 1939, 1956). The difference between "Hazen" style approaches and "Kozeny-Carman" style approaches is that the latter allows for variations of void ratio to be modelled (cf. Carrier 2003) while "Hazen" relies on selection of an effective particle size. Zhai, Rahardjo, and Satyanaga (2018) adopted the "pore-size distribution function" method which gave a good prediction for k_{sat} of sandy matierials. Feng, Vardanega, and Ibraim (2019) and Feng et al. (2019) show that the entire particle size distribution (PSD) curve can be used along with the Grading Entropy concept (e.g. Lőrincz et al. 2005) to compute the normalised entropy co-ordinates which have also been demonstrated to predict k_{sat} reasonably well for gravel and sands. Both the "Hazen" and "Kozeny-Carman" approaches require calibration and therefore "transformation models" (regression models) (cf. Phoon and Kulhawy 1999a, 1999b) are needed.

1.2. FG/KSAT-1358

Feng and Vardanega (2019) have recently assembled a large database of 1358 k_{sat} measurements on a variety of fine-grained soils, hereafter referred to as FG/KSAT-1358 (full details of the sources and composition of the database can be found in Feng and Vardanega (2019) and are not repeated here for brevity). Using this large data-set and following the previous works of a transformation model has been proposed linking water content ratio (which is defined as the water content of the soil (w) normalised by the water content at the liquid limit (w_L) i.e. w/w_L). The water content ratio is equivalent to (e/e_L) for the fully saturated case. The transformation model developed in Feng and Vardanega (2019)

CONTACT Paul J. Vardanega p.j.vardanega@bristol.ac.uk Department of Civil Engineering, University of Bristol, University Walk, Bristol BS8 1TR, UK
Supplemental data for this article can be accessed doi:10.1080/17499518.2019.1652919.
This article has been republished with minor changes. These changes do not impact the academic content of the article.
© 2019 Informa UK Limited, trading as Taylor & Francis Group

follows that from previous works of Nagaraj, Pandian, and Narashimha Raju (1993, 1994); Sivapullaiah, Sridharan, and Stalin (2000); Mbonimpa et al. (2002) who proposed similar models linking k_{sat} with w/w_L but with smaller data-sets. The new model is calibrated with $n = 1352$ measurements in Feng and Vardanega (2019), and is given here as Equation (1):

$$\ln[k_{sat}(m/s)] = 4.083\ln(w/w_L) - 20.074 \quad [R^2 = 0.62,\ n = 1352,\ SE = 1.58,\ p < 0.0001] \tag{1a}$$

which can be re-arranged to:

$$k_{sat}(m/s) = 1.91 \times 10^{-9}(w/w_L)^{4.083} \tag{1b}$$

In this letter, FG/KSAT-1358 is used for a different purpose: to determine the best fit probability density functions (PDFs) that describe the key parameters within the database with associated goodness-of-fit tests. It should be noted that as six datapoints from FG/KSAT-1358 were identified as potential outliers and thus excluded in the calibration of the transformation model presented in Feng and Vardanega (2019): this data is also not considered in the analysis presented in this letter giving a dataset of $n = 1352$.

2. Probability distributions in geotechnical engineering

Lumb (1966, 1970) examined the use of the normal, Gaussian and beta distributions for soil data-sets in Hong Kong. Rackwitz (2000) recognised that a lognormal distribution had a "strong precedent" (Shepheard et al. 2019) in soils engineering. Vardanega and Bolton (2016) cautioned against the use the use of PDFs of soil parameters to examine the ULS state (as opposed to the SLS state), in part due to the well-known problem of lack of data at the tails. Scott, Kim, and Salgado (2003) state "In geotechnical engineering, information about the mean and variance of a load or resistance is typically available, even though the exact distribution may not be known". To remedy this problem geo-databases are needed. Arguably, fitted distributions calibrated with databases of soil parameters should be used to set up stochastic and sensitivity analyses so that the range of potential outcomes for any particular geotechnical problem of interest can be better understood.

Recently Shepheard et al. (2019) have shown in that Weibull distribution may better describe peak effective friction angle (ϕ'_{peak}) (number of datapoints (n) = 85) and cohesion intercept (c') ($n = 86$) for a database of soils from the island of Saint Lucia. While a lognormal distribution may be commonly used in geotechnical engineering (e.g. Rackwitz 2000; Scott, Kim, and Salgado 2003), if a database is available then the engineer should investigate the applicability of a variety of statistical distributions (Shepheard et al. 2019).

3. FG/KSAT-1358: probability distributions

3.1. Analysis

The key parameters that describe FG/KSAT-1358 are the void ratio (e), the liquid limit (w_L), the water content ratio (w/w_L) (or e/e_L) and the negative natural logarithm of saturated hydraulic conductivity -ln[k_{sat}(m/s)]. It should be noted that the data of k_{sat} varies over seven orders of magnitude (see Table 1). Various probability density functions were fitted to each of these parameters in turn. The PDFs trialled were: "Weibull" (W), "Normal" (N); "Lognormal" (LogN), "Exponential" (Exp), "Generalized extreme value (GEV)", "Logistic" (Logi), "LogLogistic" (LogLogi), "Gamma" (G) (all fitted with functions available in Matlab®). As the "Normal", "Generalized Extreme Value" and "Logistic" distributions can take negative values, in this work, they were not applied to the strictly positive parameters (e, w_L, w/w_L). However, the negative natural logarithm of saturated hydraulic conductivity -ln [k_{sat}(m/s)] in this database is always positive for the studied database (FG/KSAT-1358), but theoretically it can still be negative. Therefore, for completeness, all the aforementioned eight probability distributions functions were trialled for -ln[k_{sat}(m/s)]. Figure 1(a), Figure 2(a), Figure 3(a) and Figure 4(a) shows the aforementioned distributions fitted to the data of "e", "w_L", "w/w_L" and "-ln[k_{sat}(m/s)]" respectively, the fitted parameters for the trialled probability distribution functions along with their general form can be found in Table S1 of the online supplement.

3.2. Goodness of fit tests

Table 2 shows the different distributions' goodness of fit test results. In the analysis presented here both the and

Table 1. Summary statistics for FG/KSAT-1358 (*SD* = standard deviation; *COV* = coefficient of variation) ($n = 1352$).

Parameter	Min.	Max.	Mean	SD	COV
e	0.19	8.57	1.32	1.01	0.76
w_L (%)	22	675	84.84	83.09	0.98
w/w_L	0.058	2.99	0.69	0.35	0.51
k_{sat} (m/s)	1.44×10^{-13}	7.50×10^{-6}	2.21×10^{-8}	2.73×10^{-7}	12.4
-ln[k_{sat} (m/s)]	11.80	29.57	22.04	2.55	0.12

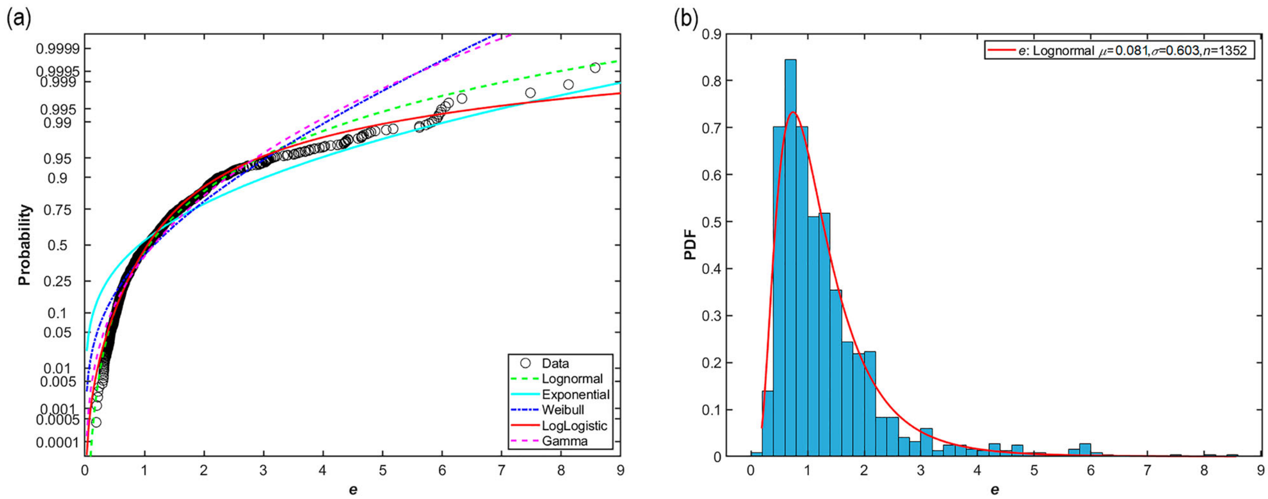

Figure 1. (a) Different PDFs fitted to the void ratio data from FG/KSAT-1358; (b) Best fit PDF (lognormal) shown ($n = 1352$).

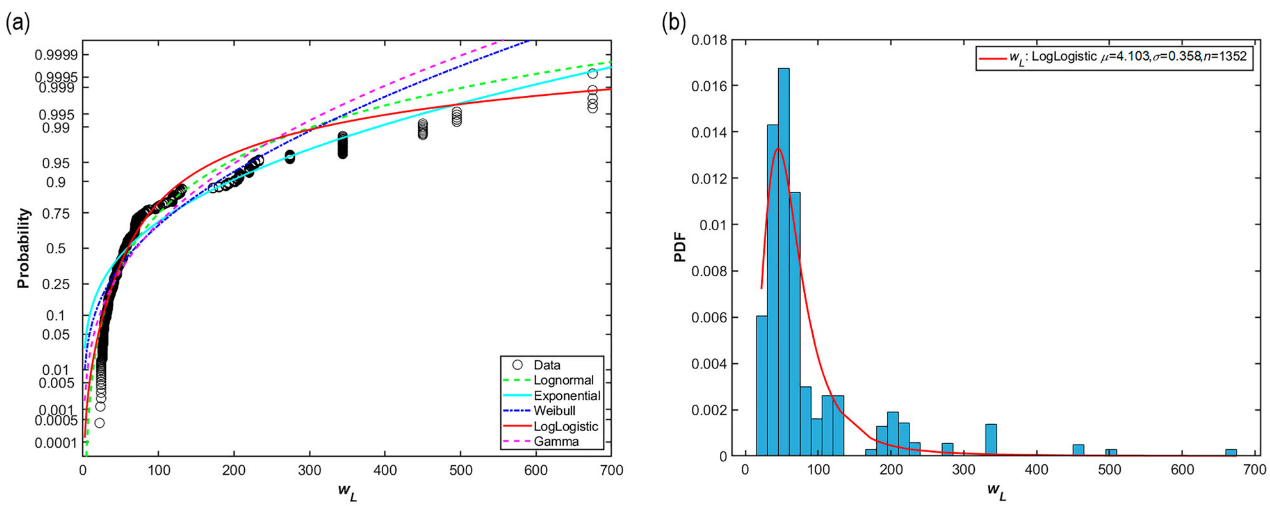

Figure 2. (a) Different PDFs fitted to the w_L data from FG/KSAT-1358; (b) Best fit PDF (loglogistic) shown ($n = 1352$).

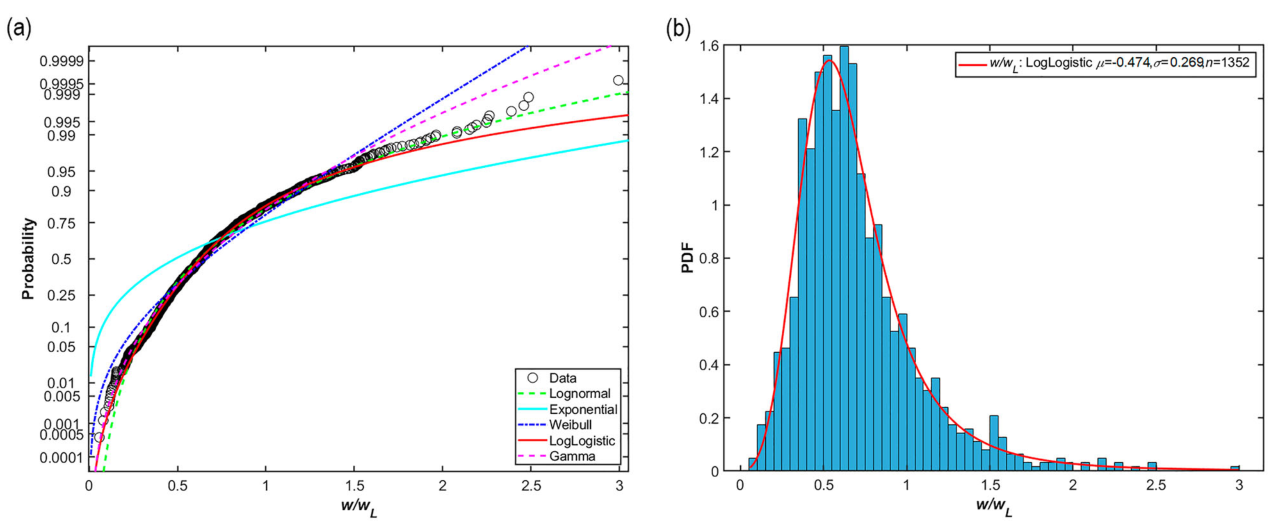

Figure 3. (a) Different PDFs fitted to the w/w_L data from FG/KSAT-1358; (b) Best fit PDF (loglogistic) shown ($n = 1352$).

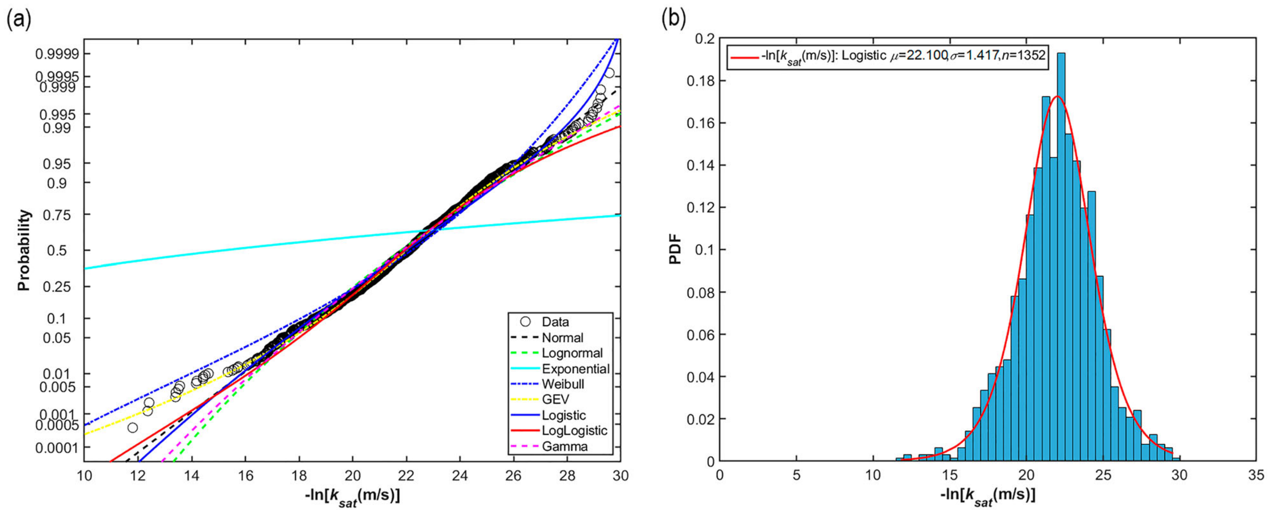

Figure 4. (a) Different PDFs fitted to the -ln[k_{sat}(m/s)] data from FG/KSAT-1358; (b) Best fit PDF (logistic) shown ($n = 1352$).

Table 2. AIC and AICc for the fitted distributions shown in Figures 1–4 (strongest fits shown in bold type).

	AIC or AIC$_C$							
	N	Exp	LogN.	W	GEV	Logi.	LogLogi.	G
e	–	3464	**2694**	3091	–	–	2702	2901
w_L (%)	–	14714	13992	14564	–	–	**13945**	14391
w/w_L	–	1720	621	776	–	–	**568**	623
-ln[k_{sat} (m/s)]	6375	11069	6465	6422	6408	**6350**	6395	6426

Akaike Information Criterion (AIC) (Equation 2) (Akaike 1974) and the Corrected Akaike Information Criterion (AIC$_C$) (Equation 3) which takes into account the size of the sample thus eliminate the risk of "over-fitting" the data (e.g. Sugiura 1978; Hurvich and Tsai 1989; Hurvich and Tsai 1995; Burnham and Anderson 2004) are given. Based on Burnham and Anderson (2004), the AIC and AIC$_C$ can be expressed as

$$\mathrm{AIC} = -2\log\left(L(\hat{\theta})\right) + 2K \tag{2}$$

$$\mathrm{AIC}_C = -2\log\left(L(\hat{\theta})\right) + 2K + \frac{2K(K+1)}{n-K-1} \tag{3}$$

where $L(\hat{\theta})$ is the likelihood function, K is the number of estimable parameters in the approximating model, while n is the sample size.

As the sample size here in this analysis is considerable ($n = 1352$), the computed results for AIC and AIC$_C$ only differ in the decimal places (not shown here for brevity), the values quoted in the subsequent analysis can be taken indicating that the AIC and AICc are essentially identical for the calculations presented in this letter. For the data of "e" the best fit function is the lognormal, for "w_L", "w/w_L" the best fit distribution is the loglogistic, and for the data of "-ln[k_{sat}(m/s)]" the best fit function is the logistic function, which is similar to a normal distribution but with larger tails (Birnbaum and Dudman1963; Mudholkar and George 1978) therefore k_{sat} is essentially modelled with a log-normal distribution. Figures 1(b), 2(b), 3(b) and 4(b) show the best fit distributions plotted for each of the four studied parameters.

4. PDFs for -ln[k_{sat}(m/s)] database sub-sets

A PDF fitted to a parameter across the entire database may not be the best choice when the database is subdivided in certain ways. In Feng and Vardanega (2019) the effects on the transformation function were studied when FG/KSAT-1358 was split into four subcategories: liquid limit range (i.e. w_L greater than or less than 50%); position on the Casagrande chart (e.g. ASTM 2017) (i.e. whether the material would classify as a clay or silt); permeability test method (i.e. falling head, constant head; flow pump or consolidation) and sample state (i.e. "disturbed" or "undisturbed") would change the best fit PDF. For the sake of brevity only the effects of these sub-divisions on -ln[k_{sat}(m/s)] are shown in detail in this letter. It should be noted that for the two sub-categories: above and below the A-line $n = 1277$ points as 75 samples are without sufficient soil classification information (e.g. plasticity index). Also the sample states of 122 samples in the database were not clearly specified and therefore for the sub-categories: "disturbed" and

Table 3. AIC and AICc for the fitted distributions to the -ln[k_{sat}(m/s)] data when split by liquid limit level, silt or clay classification, type of hydraulic conductivity test used and sample state (strongest fits shown in bold type).

-ln[k_{sat} (m/s)]	n	AIC or AIC$_C$ N	Exp	LogN	W	GEV	Logi	LogLogi	G	Figure in supplement
Falling head	580	2692	4706	2768	2620	**2603**	2685	2727	2740	S1
Consolidation	512	2299	4238	2284	2383	**2283**	2302	2291	2288	S2
Flow pump	91	474	749	478	471	**470**	481	484	476	S3
Constant head	169	748	1379	754	745	**741**	757	762	752	S4
$w_L \geq$ 50%	854	**3919**	7044	3932	3991	3925	3929	3935	3924	S5
w_L<50%	498	2288	4025	2355	**2217**	2220	2282	2320	2330	S6
Above A-line	934	4289	7680	4356	4346	4324	**4245**	4269	4327	S7
Below A-line	343	1614	2771	1611	1648	1612	1628	1626	**1610**	S8
Disturbed	1103	5235	9044	5309	5275	5260	**5214**	5251	5277	S9
Undisturbed	127	509	1038	503	535	**493**	514	509	505	S10

"undisturbed" combined n = 1230. All the calculated AIC (AIC$_C$) of different PDF models for each sub-dataset are summarised in Table 3 with their fitted results presented in Figures S1-S10 in the online supplement.

4.1. Test method

Table 3 shows that for the "Falling head"; "Consolidation"; "Flow pump" and "Constant Head" categories all suggest that the GEV is the best fit PDF.

4.2. Liquid limit level

For materials with w_L ≥50 the normal distribution is favoured with the Gamma ranking as second and the GEV ranking a close third, while for the w_L < 50 the Weibull distribution is favoured with the GEV ranking second.

4.3. Location with respect to the A-line

For materials that would plot above the A-line the logistic distribution is favoured with loglogistic distribution ranking second and normal ranking third. For materials that would plot below the A-line the Gamma distribution is favoured with lognormal closely ranking second.

4.4. Sample type

For those samples classed as "disturbed" or remoulded, which comprise most of the database, the Logistic function is favoured with normal distribution ranking second and loglogistic ranking third. For those samples classed as "undisturbed" albeit potentially subjected to varying stress levels in the laboratory work the GEV distribution is favoured with the lognormal ranking second.

4.5. Comparison to the n = 1352 dataset

As already mentioned, for the entire database for -ln [k_{sat}(m/s)] the logistic function is favoured with the normal ranking second and the loglogistic ranking third. From the above discussion it can be seen that the GEV and logistic features are either at the top or near the top of most of the rankings of the PDFs for each database subset.

5. Conclusions

The database FG/KSAT-1358 (Feng and Vardanega 2019) has been analysed using probability density functions fitted to data or four key parameters: e, w_L, w/w_L and -ln[k_{sat}(m/s)]. For e the lognormal distribution is top ranked PDF, while for w_L and w/w_L the loglogistic distribution was calculated to be the best fit based on examination of both the AIC and AICc. For the "-ln[k_{sat}(m/s)]" data the logistic function is the best fit. For the various subsets of "-ln[k_{sat}(m/s)]" examined: test method; liquid limit level; location above or below the A-line and sample state generally the logistic and GEV distributions are the most favoured or ranked in the top three distributions of those studied in this letter. The results may be useful for those wishing to stochastically model variations of saturated hydraulic conductivity for various geotechnical applications.

Acknowledgments

The authors thank Dr Flavia De Luca for her helpful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The first author acknowledges the support from the China Scholarship Council (CSC) under the [grant number CSC No. 201708060067].

Data availability statement

This research has not generated new experimental data.

ORCID

Shuyin Feng http://orcid.org/0000-0002-3837-6762
Paul J. Vardanega http://orcid.org/0000-0001-7177-7851

References

Akaike, H. 1974. "A new Look at the Statistical Model Identification." *IEEE Transactions Automatic Control* 19 (6): 716–723.

Almeida, S., E. A. Holcombe, F. Pianosi, and T. Wagener. 2017. "Dealing with Deep Uncertainties in Landslide Modelling for Disaster Risk Reduction Under Climate Change." *Natural Hazards and Earth System Sciences* 17: 225–241.

ASTM. 2017. *Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System)*. West Conshohocken, USA: ASTM International.

Birnbaum, A., and J. Dudman. 1963. "Logistic Order Statistics." *The Annals of Mathematical Statistics* 34: 658–663.

Burnham, K. P., and D. P. Anderson. 2004. "Multimodal Inference: Understanding AIC and BIC in Model Selection." *In: Sociological Methods and Research* 33 (2): 261–304.

Carman, P. C. 1937. "Fluid Flow Through Granular Beds." *Transactions-Institution of Chemical Engineers* 15: 150–166.

Carman, P. C. 1939. "Permeability of Saturated Sands, Soils and Clays." *The Journal of Agricultural Science* 29 (2): 262–273.

Carman, P. C. 1956. *Flow of Gases through Porous Media*. London: Butterworths Scientific Publications.

Carrier III, W. D. 2003. "Goodbye, Hazen; Hello, Kozeny-Carman." *Journal of Geotechnical and Geoenvironmental Engineering* 129(11): 1054–1056.

Feng, S., and P. J. Vardanega. 2019. "Correlation of the Hydraulic Conductivity of Fine-Grained Soils with Water Content Ratio Using a Database." *Environmental Geotechnics* 1–16. doi:10.1680/jenge.18.00166.

Feng, S., P. J. Vardanega, and E. Ibraim. 2019. "Comparison of Prediction Models for the Permeability of Granular Materials Using a Database." In *Contemporary Issues in Soil Mechanics. GeoMEast 2018. Sustainable Civil Infrastructures*, edited by S. Hemeda, and M. Bouassida, 1–13. Cham: Springer.

Feng, S., P. J. Vardanega, E. Ibraim, I. Widyatmoko, and C. Ojum. 2019. "Permeability Assessment of Some Granular Mixtures." *Géotechnique* 69 (7): 646–654.

Hazen, A. 1893. "Some physical properties of sand and gravels. Massachusetts State Board of Health." 24th Annual Report. Wright & Potter Printing, Boston, USA.

Hazen, A. 1895. *The Filtration of Public Water-Supplies*. New York, USA: John Wiley & Son.

Hazen, A. 1911. "Discussion of 'Dam on Sand Foundation'by A. C. Koenig." *Transactions of American Society of Civil Engineers* 73: 199–203.

Hurvich, C. M., and C.-L. Tsai. 1989. "Regression and Time Series Model Selection in Small Samples." *Biometrika* 76 (2): 297–307.

Hurvich, C. M., and C.-L. Tsai. 1995. "Model Selection for Extended Quasi-Likelihood Models in Small Samples." *Biometrics* 51 (3): 1077–1084.

Kozeny, J. 1927. "Über Kapillare Leitung des Wassers im Boden (Aufstieg, Versickerung und Anwendung auf die Bewässerung)." [in German.] *Sitzungsber. Akad. Wiss., Wien.* 136(a): 271–306.

Kulhawy, F. H., and P. W. Mayne. 1990. Manual on estimating soil properties for foundation design, *Report. No. EL*-6800, Electric Power Research Institute, Palo Alto, California, USA.

Lőrincz, J., E. Imre, M. Gálos, Q. P. Trang, K. Rajkai, S. Fityus, and G. Telekes. 2005. "Grading Entropy Variation due to Soil Crushing." *International Journal of Geomechanics* 5 (4): 311–319.

Lumb, P. 1966. "The Variability of Natural Soils." *Canadian Geotechnical Journal* 3 (2): 74–97.

Lumb, P. 1970. "Safety Factors and the Probability Distribution of Soil Strength." *Canadian Geotechnical Journal* 7 (3): 225–242.

Mbonimpa, M., M. Aubertin, R. Chapuis, and B. Busssière. 2002. "Practical Pedotransfer Functions for Estimating the Saturated Hydraulic Conductivity." *Geotechnical and Geological Engineering* 20 (3): 235–259.

Mudholkar, G. S., and E. O. George. 1978. "A Remark on the Shape of the Logistic Distribution." *Biometrika* 65 (3): 667–668.

Nagaraj, T. S., N. S. Pandian, and P. S. R. Narashimha Raju. 1993. "Stress State-Permeability Relationships for Fine-Grained Soils." *Géotechnique* 43 (2): 333–336.

Nagaraj, T. S., N. S. Pandian, and P. S. R. Narashimha Raju. 1994. "Stress State-Permeability Relations for Overconsolidated Clays." *Géotechnique* 44 (2): 349–352.

Phoon, K.-K., and F. H. Kulhawy. 1999a. "Characterization of Geotechnical Variability." *Canadian Geotechnical Journal* 36 (4): 612–624.

Phoon, K.-K., and F. H. Kulhawy. 1999b. "Evaluation of Geotechnical Property Variability." *Canadian Geotechnical Journal* 36 (4): 625–639.

Rackwitz, R. 2000. "Reviewing Probabilistic Soil Modelling." *Computers and Geotechnics* 26 (3-4): 199–223.

Scott, B., B. J. Kim, and R. Salgado. 2003. "Assessment of Current Load Factors for Use in Geotechnical Load and Resistance Factor Design." *Journal of Geotechnical and Geoenvironmental Engineering* 129 (4): 287–295.

Shepheard, C. J., P. J. Vardanega, E. A. Holcombe, R. Hen-Jones, and F. De Luca. 2019. "Minding the Geotechnical Data gap: Appraisal of the Variability of key Soil Parameters for Slope Stability Modelling in Saint Lucia." *Bulletin of Engineering Geology and the Environment*. doi:10.1007/s10064-018-01451-5. ahead of print.

Shepheard, C. J., P. J. Vardanega, E. A. Holcombe, and K. Michaelides. 2018. "Analysis of Design Choices for a Slope Stability Scenario in the Humid Tropics." *Proceedings of the Institution of Civil Engineers – Engineering Sustainability* 171 (1): 37–52.

Sivapullaiah, P. V., A. Sridharan, and V. K. Stalin. 2000. "Hydraulic Conductivity of Bentonite-Sand Mixtures." *Canadian Geotechnical Journal* 37 (2): 406–413.

Sugiura, N. 1978. "Further Analysts of the Data by Akaike' s Information Criterion and the Finite Corrections." *Communications in Statistics - Theory and Methods* 7 (1): 13–26.

Vardanega, P. J., and M. D. Bolton. 2016. "Design of Geostructural Systems." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 2 (1): 04015017.

Zhai, Q., H. Rahardjo, and A. Satyanaga. 2018. "A Pore-Size Distribution Function Based Method for Estimation of Hydraulic Properties of Sandy Soils." *Engineering Geology* 246: 288–292.