Solar Energy 187 (2019) 290–292

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Corrigendum

Corrigendum to “On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading” [Sol. Energy 173 (2018) 1087–1106]

Jamie M. Bright[a,*], Sven Killinger[a,b]

[a] Fenner School of Environment and Society, The Australian National University, 2601 Canberra, Australia
[b] Fraunhofer Institute for Solar Energy Systems ISE, 79100 Freiburg, Germany

This corrigendum specifically addresses mistakes or oversights into the original paper by Killinger et al. (2018). We explicitly state that none of the results or findings in the original paper were incorrect or indeed conflicted as a result of these mistakes. The nature of these changes merely impacts the ability of the reader to reproduce the statistics.

As a summary, the original paper (Killinger et al., 2018) collected metadata from 2.8 million photovoltaic (PV) systems and summarised the variables of tilt angle, azimuth angle, installed capacity and full load hours for each country of the study in terms of a fitted probability distribution. Many types of probability distribution were fitted and those that were most appropriate were assigned to each country and variable resulting in 96 parameterisations. The distributions were produced using Matlab but were stated in terms of $f(x|c_1, c_2, c_3, c_4) = \ldots$ where c_i is a parameter specified for the particular distribution type and $\cdots$ is the specific formulation.

This document aims to correct a statement about whether x was normalised prior to parameterisation (Section 1), correct typographic errors in the $f(x)$ formulations (Section 2) and to offer a Matlab® function and test script as supplementary material that provides simple random sampling of the distributions such that the reader may bypass the need for $f(x)$ formulations (Section 3).

1. Normalisation of x data

Each variable of metadata has specific x data that the distributions are fitted to (e.g. tilt is 0 to 90, azimuth is -180 to 180 etc.); it is essential that this x data is known prior to using it with the specified probability distribution. We stated in the original paper that x data was never normalised, however, this was incorrect. In reality, this occurred on 8 of the 96 parameterisations. We describe here the process of normalising of x data, which was absent in Killinger et al. (2018).

When fitting a distribution to the data, certain distributions were fit when the x data was normalised between 0 and 1; the resulting parameterisation is then only valid between 0 and 1 and must be subsequently converted back to the original x in order to randomly sample the distribution. For example, azimuth angle has x data between -180 and 180 and certain distributions could not be fit to negative data. To ensure the widest test of possible distributions, x was normalised between 0 and 1 to become $\hat{x}$ for these particular distributions.

We have reproduced the distribution coefficient table for all countries, variables and PV system size classification in Table 1. This time, however, we have added boldface indication and an asterisk specifying which of the distributions are parameterised using the normalised $\hat{x}$ data between 0 and 1.

For absolute clarity, $\hat{x}$ is calculated as:

$$\hat{x} = \frac{x_v - \textsc{min}(x_v)}{\textsc{range}(x_v)}; \quad (1)$$

where x_v indicates the x data for a the specific metadata variable: 0:90 for tilt in degrees, -180:180 for azimuth in degrees, 0:2 for full load hours in thousands (FLH), and 0:25 or 25:1000 for capacity in kWp.

2. Errors in $f(x)$ formulations

For readers to recreate the distributions presented in Killinger et al. (2018), expressions in terms of $f(x)$ are required should the reader not have access to Matlab. Unfortunately, there are typographic mistakes and an additional misinterpretation of the formulas in the original paper that are corrected here as updated in Table 2.

The extreme value distribution featured an additional minus sign. The Nakagai distribution had an order the wrong side of a bracket. The Logistic distribution was missing a bracket. The Generalized Extreme Value distribution had an oversight in that Matlab deploys different expressions of distribution depending on the sign of $F = 1 + k\frac{x-\mu}{\sigma}$. To accommodate this variation, we add an additional component $\lceil \max(0, \min(1, F)) \rceil$ where $\lceil \cdot \rceil$ represents the ceiling operation; the latter expression is equal to 0 if $F \leq 0$ and 1 if $F > 0$. The Loglogistic

DOI of original article: https://doi.org/10.1016/j.solener.2018.08.051
* Corresponding author.
E-mail addresses: jamiebright1@gmail.com, jamie.bright@anu.edu.au (J.M. Bright).

https://doi.org/10.1016/j.solener.2019.04.030
Received 4 April 2019; Accepted 8 April 2019
Available online 28 May 2019
0038-092X/

Table 1
The distribution coefficients C_i corresponding to the definitions in Table 2. The left side is for capacities $\leqslant$ 25 kWp and the right side is for >25 kWp. Those distributions in boldface followed by an asterisk (e.g. **Nakag.***) indicate that the x-data was normalised prior to parameterisation.

Cluster	Param.	Dist. Type	C_1	C_2	C_3	C_4	Dist. Type	C_1	C_2	C_3	C_4
Australia	Azimuth	Logistic	8.3472	34.2603	–	–	**Nakag.***	28.0028	0.2537	–	-
	Tilt	Logistic	15.0891	6.0136	–	–	Normal	14.6875	8.4183	–	–
	Capacity	tLoc.Scale	4.9355	1.4855	3.2434	–	Lognorm.	4.3491	0.8328	–	–
Austria	Azimuth	Logistic	0.4566	17.6159	–	–	Logistic	0.6107	19.9366	–	–
	Tilt	Weibull	34.6425	3.6381	–	–	Weibull	26.2683	2.7083	–	–
	Capacity	G.Ext.Val.	0.1633	2.9111	4.8713	–	G.Ext.Val.	0.4337	8.5691	33.5998	–
	FLH	tLoc.Scale	1.0668	0.0971	2.0439	–	Logistic	1.0144	0.1090	–	–
Belgium	Azimuth	Logistic	−1.2365	24.3969	–	–	Logistic	−3.5828	19.3034	–	–
	Tilt	Ext.Val.	39.8100	7.8703	–	–	Logistic	29.6932	5.1091	–	–
	Capacity	Stable	1.4734	0.9578	1.3688	4.8772	Loglog.	3.3900	0.0905	–	–
	FLH	Burr	0.5814	9.2775	5.8277	–	tLoc.Scale	0.8570	0.1623	3.8904	–
Denmark	Azimuth	Logistic	1.5526	20.6455	–	–	G.Ext.Val.	−0.3684	49.6733	−13.0930	–
	Tilt	**Loglog.***	−1.1441	0.2256	–	–	Stable	0.9410	−1.0000	6.5572	41.6962
	Capacity	tLoc.Scale	5.7414	1.5014	2.3622	–	Stable	0.9767	1.0000	2.8140	28.6026
	FLH	Stable	1.2818	−0.8035	0.0948	0.8767	Logistic	0.9500	0.0778	–	–
France	Azimuth	tLoc.Scale	0.3346	34.5161	5.4215	–	Logistic	−2.2831	19.1336	–	–
	Tilt	**Lognorm.***	−1.2139	0.4144	–	–	Burr	16.5673	5.0314	0.6793	–
	Capacity	Stable	0.4000	0.5250	0.0345	3.0118	Burr	27.9585	19.1798	0.0934	–
	FLH	tLoc.Scale	1.1011	0.1416	7.6129	–	tLoc.Scale	1.1194	0.1111	2.9295	–
Germany	Azimuth	Logistic	−0.1366	23.0048	–	–	Stable	1.0309	−0.0002	9.9430	−0.2506
	Tilt	Ext.Val.	38.0648	9.6547	–	–	**Nakag.***	1.1890	0.0887	–	–
	Capacity	G.Ext.Val.	0.1143	3.5745	6.2413	–	Stable	0.7373	1.0000	5.2331	30.8089
	FLH	Stable	1.6611	−0.9423	0.1152	0.9086	Stable	1.5803	−0.8322	0.0902	0.9530
Italy	Azimuth	tLoc.Scale	−5.2371	32.9151	2.1254	–	Ext.Val.	4.4388	46.3326	–	–
	Tilt	Stable	1.8917	1.0000	5.5359	19.0462	**Nakag.***	0.6147	0.0448	–	–
	Capacity	Stable	0.9088	0.9411	1.5334	4.4130	Burr	37.2635	8.0048	0.1480	–
	FLH	tLoc.Scale	1.1774	0.1415	2.5624	–	Logistic	1.1096	0.1246	–	–
Japan	Azimuth	Logistic	−0.7254	15.9481	–	–	Stable	0.4045	0.0010	0.2121	0.0007
	Tilt	Ext.Val.	27.6791	6.4206	–	–	Weibull	18.9055	2.6245	–	–
	Capacity	Stable	1.3094	0.9774	1.0384	4.4144	tLoc.Scale	48.0538	7.9913	1.4353	–
	FLH	Burr	1.3730	11.1707	2.7755	–	Burr	1.4833	13.7688	5.1025	–
Netherl.	Azimuth	Stable	1.0309	0.0001	11.2285	0.0294	Logistic	−0.3912	15.7031	–	–
	Tilt	Ext.Val.	38.8465	11.6153	–	–	Loglog.	3.0459	0.2311	–	–
	Capacity	tLoc.Scale	3.2381	1.3462	1.6898	–	Stable	0.6078	1.0000	3.7179	28.2348
	FLH	Stable	1.3553	−0.8354	0.0893	0.9334	Stable	1.4819	−0.9477	0.0760	0.9577
UK	Azimuth	Logistic	0.2461	26.0886	–	–	Logistic	−0.8287	16.4894	–	–
	Tilt	tLoc.Scale	31.5952	4.6591	3.3818	–	Gamma	3.9484	4.5939	–	–
	Capacity	Stable	1.8937	−0.0544	0.5898	2.9504	G.Ext.Val.	0.5210	10.8742	35.6490	–
	FLH	Stable	1.6851	0.0000	0.0776	0.9035	Stable	1.3744	−0.8038	0.0638	0.9395
USA Nor.	Azimuth	Logistic	0.4600	28.1468	–	–	tLoc.Scale	0.1686	12.6681	1.2295	–
	Tilt	**Loglog.***	−1.2944	0.2061	–	–	Stable	1.3001	1.0000	4.8192	10.6480
	Capacity	G.Ext.Val.	0.0868	2.5493	4.6450	–	G.Ext.Val.	1.1806	31.0335	45.3525	–
	FLH	Stable	1.3201	−0.7411	0.0666	1.0628	Logistic	1.0679	0.0575	–	–
USA Sou.	Azimuth	Logistic	8.5676	29.3711	–	–	Stable	0.4000	0.4410	0.2829	0.0822
	Tilt	tLoc.Scale	20.4150	3.8106	3.0306	–	**Weibull***	0.1647	1.3876	–	–
	Capacity	**Burr***	0.3092	2.5724	2.2109	–	G.Ext.Val.	1.2834	38.6717	49.6038	–
	FLH	Ext.Val.	1.4898	0.0963	–	–	Logistic	1.4035	0.0579	–	–

distribution contained an additional component that should never have featured; furthermore, the common logarithm was stated but should in fact be the natural logarithm.

Lastly, the approach for utilising the stable distribution was unclear due to its lack of description in the form $f(x|\alpha, \beta, \gamma, \delta) = ...$ Simply expressing the stable distribution in this format is not possible in a simple basis due to the numerous logical queries that change its formulation. Ideally, the user shall use the inbuilt Matlab® formulations with the following call:

```
makedist('Stable','alpha',c1,'beta',c2,'gam',c3,'delta',c4)
```

where c_x parameters as stated in Table 1; this approach would guarantee reproduction. However, we appreciate that not all readers work in Matlab; an excellent alternative exists in R (Wuertz and Maechler, 2016), and additional packages in alternative coding languages are available from (Nolan, 2018, 2019).

3. New Matlab function and test script

As supplementary material, a Matlab function called `GetPVStats.m` and a testing script called `testGetPVStats.m` are provided to the reader. `GetPVStats.m` is a function that randomly samples the appropriate distribution depending on the requirements of the user. The usage is simple with the call:

```
samples = GetPVStats(system_size, variable, country, n_samples),
```

where the `system_size` is either the string `'small'` for capacities between 0 and 25 kWp, or `'large'` between 25 and 1000 kWp. The `variable` can be any of the strings `'azimuth'`, `'tilt'`, `'capacity'` or `'flh'`; `country` can be the string name of any of the countries within the study, e.g. `'germany'`. Lastly, `n_samples` is the number random samples to be returned by the function e.g. `500` would return as many from the distribution. The $f(x)$ formulations are

Table 2
Probability distribution formulations used in the research. The coefficients correspond to those presented in Table 1.

Distribution Name	C_1	C_2	C_3	C_4	Probability Density Function, $f(x\|\ldots)$
Burr type XII	α Scale	c Shape 1	k Shape 2	–	$f(x\|\alpha, c, k) = \dfrac{\frac{kc}{\alpha}\left(\frac{x}{\alpha}\right)^{c-1}}{\left(1+\left(\frac{x}{\alpha}\right)^{c}\right)^{k+1}}$
Extreme Value	μ Location	σ Scale	–	–	$f(x\|\mu, \sigma) = \sigma^{-1}\exp\left(\frac{x-\mu}{\sigma}\right)\exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right)$
Gamma	a Shape	b Scale	–	–	$f(x\|a, b) = \frac{1}{b^{a}\Gamma(a)}x^{a-1}\exp\left(\frac{-x}{b}\right)$
Generalized Extreme Value	k Shape	σ Scale	μ Location	–	$f(x\|k, \mu, \sigma) = \left(\frac{1}{\sigma}\right)\exp\left(-\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right)\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}} \times \lceil \max(0, \min(1, F)) \rceil$
Inverse Gaussian	μ Scale	λ Shape	–	–	$f(x\|\mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^{3}}}\exp\left(-\frac{\lambda}{2\mu^{2}x}(x-\mu)^{2}\right)$
Logistic	μ Location	σ Scale	–	–	$f(x\|\mu, \sigma) = \dfrac{\exp\left(\frac{x-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(\frac{x-\mu}{\sigma}\right)\right)^{2}}$
Loglogistic	μ Log Loc.	σ Log Scale	–	–	$f(x\|\mu, \sigma) = \frac{1}{\sigma}\frac{1}{x}\dfrac{\exp\left(\frac{\ln(x)-\mu}{\sigma}\right)}{\left(1+\exp\left(\frac{\ln(x)-\mu}{\sigma}\right)\right)^{2}}$
Lognormal	μ Log Loc.	σ Log. Scale	–	–	$f(x\|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(\frac{-(\ln x-\mu)^{2}}{2\sigma^{2}}\right)$
Nakagami	μ Shape	ω Scale	–	–	$f(x\|\mu, \omega) = 2\left(\frac{\mu^{\mu}}{\omega}\right)\frac{1}{\Gamma(\mu)}x^{(2\mu-1)}\exp\left(\frac{-\mu}{\omega}x^{2}\right)$
Normal	μ Location	σ Scale	–	–	$f(x\|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(x-\mu)^{2}}{2\sigma^{2}}\right)$
Stable	α Shape 1	β Shape 2	γ Scale	δ Location	Use inbuilt Matlab functions, `stabledist` R package or alternatives from Nolan (2018, 2019).
t Location-Scale	μLocation	σ Scale	ν Deg. of Freedom	–	$f(x\|\mu, \sigma, \nu) = \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(\dfrac{\nu+\left(\frac{x-\mu}{\sigma}\right)^{2}}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$
Weibull	a Scale	b Shape	–	–	$f(x\|a, b) = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1}\exp\left(-\left(\frac{x}{a}\right)^{b}\right)$

included yet commented out within the function; the code itself is well documented and has examples provided. The test script `testGetPVStats.m` iterates through every variable, country and system size and produces a figure that is very similar to those that featured in Killinger et al. (2018); however, no data on PV systems is stored. This script demonstrates how sample metadata can be extracted from the distributions.

The reader is encouraged to join or follow the ResearchGate project for this research at https://www.researchgate.net/project/Characteristics-and-statistics-of-PV-system-metadata. Questions and discussion posted there is the best place for collaboration on this topic.

Acknowledgement

Dr. Jamie M. Bright is funded by the Australian Renewable Energy Agency (ARENA) under Research and Development Programme Funding G00854.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.solener.2019.04.030.

References

Killinger, S., Lingfors, D., Saint-Drenan, Y.-M., Moraitis, P., Van Sark, W., Taylor, J., Engerer, N.A., Bright, J.M., 2018. On the search for representative characteristics of PV systems: data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. Sol. Energy 173, 1087–1106.

Nolan, J.P., 2018. Stable Distributions - Models for Heavy Tailed Data. Birkhauser, Boston, in progress, Chapter 1 online at <http://fs2.american.edu/jpnolan/www/stable/stable.html>.

Nolan, J.P., 2019. John Nolan's stable distributions page: information on stable distributions. <http://fs2.american.edu/jpnolan/www/stable/stable.html>.

Wuertz, D., Maechler, M., 2016. Package 'stabledist'. <https://cran.r-project.org/web/packages/stabledist/stabledist.pdf>.