



OPEN

# Mechanisms for log normal concentration distributions in the environment

August Andersson

Log normal-like concentration distributions are ubiquitously observed in the environment. However, the mechanistic origins are not well understood. In this paper, we show that first order exponential kinetics onsets log-normal concentration distributions, under certain assumptions. Given the ubiquity of exponential kinetics, e.g., source and sink processes, this model suggests an explanation for the frequent observation in the environment, and elsewhere. We compare this model to other mechanisms affecting concentration distributions, e.g., source mixing. Finally, we discuss possible implications for data analysis and modelling, e.g., log-normal rates and fluxes.

Log normal-like distributions are observed across scientific disciplines, including Earth and Environmental Sciences (Fig. 1A, Eq. 3)<sup>1–4</sup>. A fundamental question is how this dependency emerges. In the fields of Ecology and Economics, mathematical models for economic or population growth predicts log-normal distribution onset<sup>3,5</sup>. For environmental systems, a few studies discuss general mechanisms for the emergence of log-normal concentration distributions<sup>6–8</sup>. These studies are intuitively appealing, but with less theoretical rigor, while the underlying physico-chemical mechanisms remains largely unclear.

Ascribing empirical data to a given parametric distribution function, such as the log-normal distribution, is often advantageous for statistical analysis<sup>9–11</sup>. Much efforts have been dedicated towards such statistical analysis of environmental data. However, with a number of important exceptions, less interest has been directed towards understanding the underlying scientific question: what physico-chemical mechanisms drive the onset of different observational distributions, e.g., the log-normal distribution<sup>6–8,12,13</sup>?

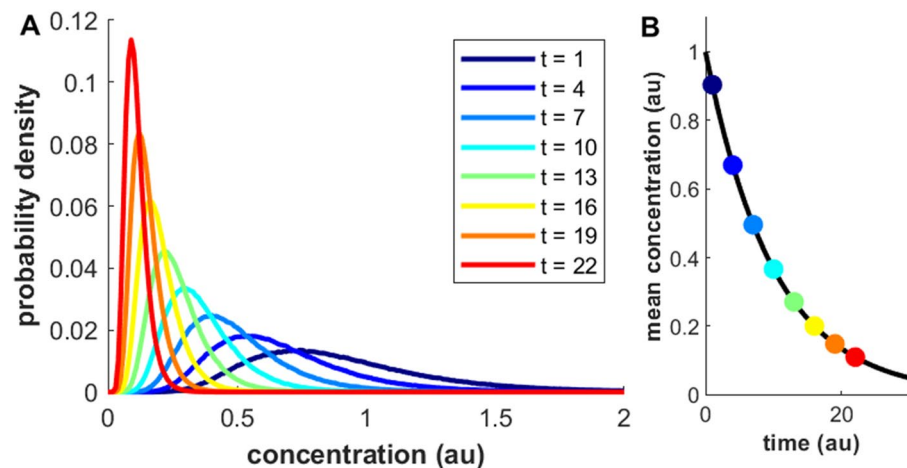
An example of a physical–chemical mechanism governing the onset of the normal concentration distribution is the mixing of multiple sources. For instance, consider emissions of carbon monoxide (CO) from many chimneys across a village. As the emission flux from each chimney vary randomly over time, the total concentration becomes the sum of many stochastic variables, and by the Central Limit Theorem the distribution tends to normality. Even though virtually any component in the environment is the sum of a number of different sources, much of environmental data display distinctively non-normal distributions, including examples with CO<sup>9–11,14</sup>. This suggests that there must be other mechanisms—mechanisms that to some extent override the source mixing mechanisms—pushing towards other concentration distributions. One such mechanism must then be the driver towards log normal-like distributions in the environment. What physical–chemical mechanism(s) may predict the onset of a log-normal dependence, as well as explain the comparable ubiquity in the environment?

In this paper, we investigate the relationship between first order kinetics and log-normal concentration distributions in environmental systems (Fig. 1A–B). A key concept in this regard is stochastic variability. To incorporate random fluctuations, we formulate a kinetic model as a stochastic differential equation (SDE). For clarity, we develop the model for the specific case of a kinetic sink process. However, we emphasize the general applicability of this mathematical framework to any process exhibiting first order kinetics, e.g., chemical formation or partitioning kinetics, also beyond environmental systems. While the mathematical treatment (see SI for details) provides rigor, we highlight physical interpretation: exponential kinetics is a possible mechanism that induces log-normal concentration distributions in the environment. This mechanism is compared with other processes that influence concentration distributions. Finally, we discuss a number of implications for data analysis and modelling.

## Theory

Consider a component,  $X$ , in an environmental system, e.g., an organic pollutant, soot particles or CO. The sink or loss kinetics is often described by a first-order differential equation, with sink rate  $k$ :

Department of Environmental Science and the Bolin Centre for Climate Research, Stockholm University, 10691 Stockholm, Sweden. email: August.Andersson@aces.su.se



**Figure 1.** Illustration of the temporal evolution of a log-normal concentration distribution during continuous kinetic degradation. Panel (A): Concentration distributions depicted at different times ( $t = \text{time}$ ; Eqs. 3 and 4a–b). Panel (B): Temporal evolution of the corresponding mean concentrations, displaying exponential decay (Eq. 5).

$$\frac{d[X]}{dt} = -k[X] \quad (1)$$

This relation results in an exponentially decaying temporal dependence (Fig. 1B, Eq. 5). For a few systems, such as radioactive decay, this description captures the dynamics well, in the sense that  $k$  is a constant. However, for many environmental systems,  $k$  is dependent on, e.g., state variables such as temperature and pressure, but also on the concentrations of other components, since sinks often involve chemical reactions. Thus,  $k$  is commonly a function of time.

As an example, the chemical breakdown of CO in the atmosphere takes place through the reaction with the hydroxyl radical (OH), and the reaction rate thus varies in time with the OH concentration<sup>14</sup>. OH in the atmosphere is formed through photo-chemical reactions, and the abundance thus depends on, e.g., the diurnal variations in sunlight. In addition to such deterministic dependence, the OH concentration typically also displays a near-random variability over time, reflecting the fluctuations of many parameters (e.g., chemical reactants, humidity, turbulent diffusion, temperature or cloudiness). Consequently, the CO sink rate therefore also displays stochastic variability.

Overall, randomly varying rates are expected to be common in various environmental systems, including many different types of processes, e.g., chemical reactions (e.g., CO with OH), deposition, photo-degradation or kinetic partitioning. In general, we can write the rate  $k$  as the sum of a deterministic and a stochastic part. Here, we are mainly interested in understanding the impact of the stochastic part, and therefore for simplicity assume that the deterministic component is a constant ( $\mu_k$ ). We may re-write Eq. (1) in the form of a stochastic differential equation (SDE):

$$\frac{d[X]}{dt} = -(\mu_k + \sigma_k \eta(t))[X] \quad (2)$$

where  $\mu_k$  is the mean reaction rate and  $\sigma_k$  is the magnitude of the stochastic fluctuation. The function  $\eta(t)$  describes the time-dependency of the random fluctuations (with amplitude 1), which we here assume is independent and identically (i.i.d.) normal distributed.

Since  $\eta(t)$  fluctuates randomly at each time point, the solutions to Eq. (2) will also fluctuate randomly, and we obtain a different solution (or path) when solving it at different instances. However, our present goal is not to generate different realizations, but to evaluate the overall behavior: to obtain the concentration distribution, i.e., the probability density distribution (pdf) of the different realizations. The equation that describes the temporal evolution of the pdf is the Fokker–Planck equation (FPE), which can be derived from a corresponding SDE<sup>15</sup>. The FPE is a (non-stochastic) partial differential equation, and the solution is a pdf with two variables: the concentration of  $X$  and time (Fig. 1A). The different steps involved in obtaining the FPE that corresponds Eq. (2) and then solving it, are provided in the SI (S1 text).

The solution of the FPE corresponding to Eq. (2) is the log-normal distribution, with parameters  $\mu$  and  $\sigma$ , of the form (Fig. 1A):

$$P([X]) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln([X])-\mu)^2}{2\sigma^2}} \quad (3)$$

Thus, a generic kinetic sink model predicts the onset of log-normal variability, under the assumption of a randomly fluctuating sink rate.

If we now, instead of a sink process, consider the formation of a reaction product. For example, the formation of CO<sub>2</sub> from the reaction of CO with OH: what concentration distribution do we expect for CO<sub>2</sub>? In the supplementary materials (S2 text) we show that a randomly fluctuating formation rate induces log-normal concentration distributions in reaction product as well. Taken together, we thus find that both source and sink kinetics may lead to log-normal concentration distributions. In fact, any first order kinetics process with a stochastic rate of the form of Eq. (2) will induce log-normal dependence.

Equation (3) is a generic expression for the log-normal distribution, and thus provides no specific information regarding the temporal kinetics associated with the system in Eq. (2). Following our treatment in the SI (S1 text), we can express the temporal evolution of the concentration distribution of X in Eq. (3), with the following parameters (Fig. 1A, Eq. S21):

$$\mu = \mu_0 - (\mu_k + 0.5\sigma_k^2)t \quad (4a)$$

$$\sigma^2 = \sigma_0^2 + \sigma_k^2 t \quad (4b)$$

The parameters  $\mu_0$  and  $\sigma_0$  represent the concentration distribution of X at  $t=0$ .

It is instructive to examine the temporal evolution of the mean ( $m$ ), which for a log-normal distribution is given by (here,  $k = \mu_k$ , since the mean of  $\mu(t)$  is zero):

$$m = \exp(\mu + 0.5\sigma^2) = [X(0)]e^{-kt} \quad (5)$$

The mean temporal dependence is thus recognized as the regular exponential solution to Eq. (1), where  $X(0)$  ( $= \exp(\mu_0 + 0.5\sigma_0^2)$ ) is the mean concentration of X at  $t=0$  (Fig. 1B).

## Results and discussion

**Log normal concentrations in the environment.** Empirical concentration distributions in the environment exhibit a wide variability, ranging from familiar parametric forms, e.g., normal and log-normal distributions, to more complex shapes<sup>9–12</sup>. Overall, the concentration distribution reflects the different processes that govern environmental fate, e.g., emissions/formation, transport processes, source mixing, sinks and dependences on other variables and processes. In principle, we may therefore extract mechanistic insights into the lifecycle of an environmental component from analysis of the concentration distribution. Although such deconvolution is intangible in the general case, there are specific cases where there are known mechanistic origins for certain distributions. For instance, source mixing in the environment will tend the distribution towards normal.

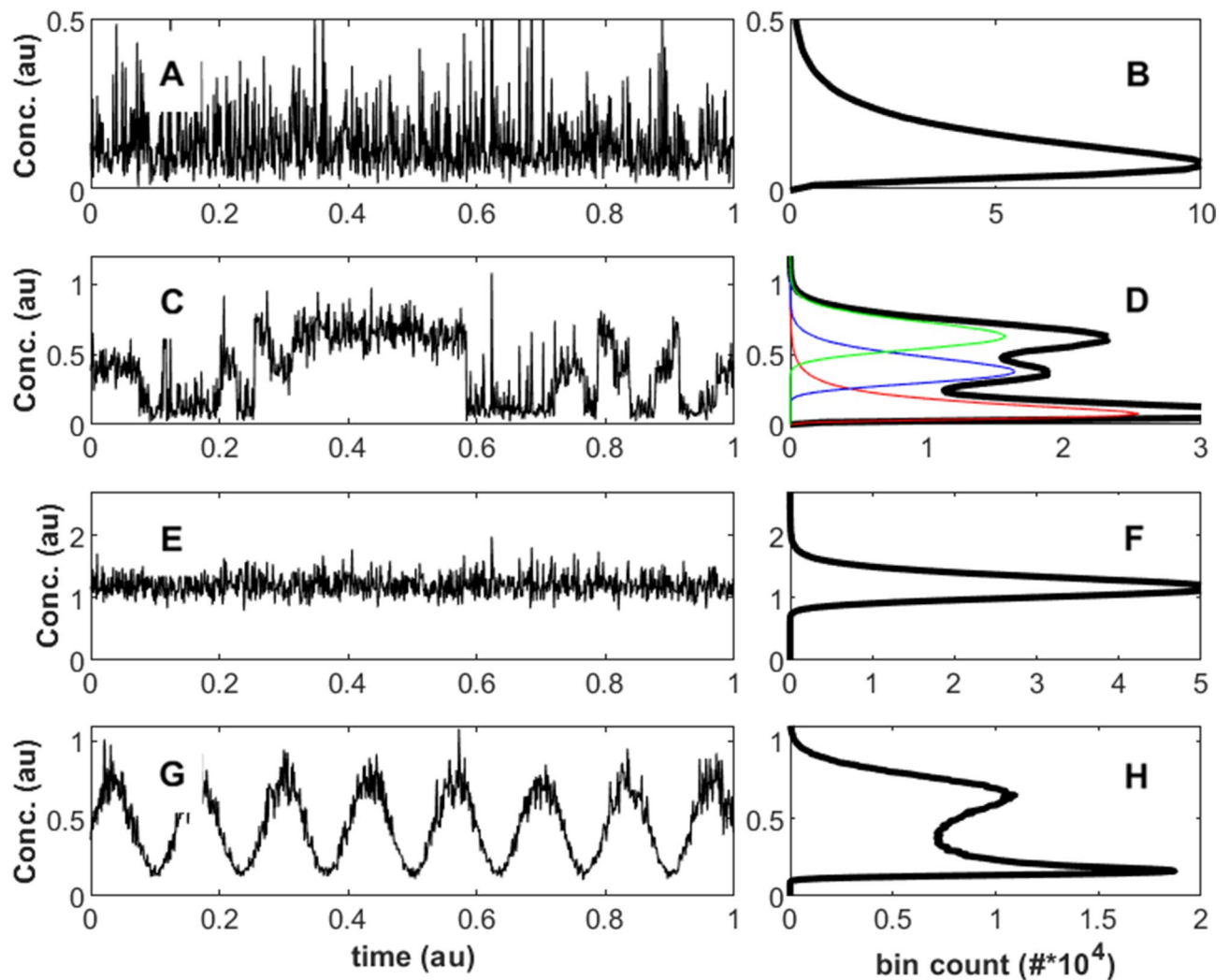
In this paper we present a model for how log-normal distributions may emerge in the environment (Fig. 1A). This topic has been addressed in a few earlier studies<sup>6–8</sup>. A joint argument in these is the implication of the ‘Multiplicative Central Limit Theorem’ (or ‘Gibrat’s law’), which applies to processes that include the product of many random variables. If we take the logarithm of the product of many random variables, we have the sum of many random variables, by which the Central Limit Theorem applies, which predict a normal distribution. Back-transforming the logarithm, yields a log-normal distribution. Although this heuristic argumentation does not provide specific physico/chemical mechanisms as to how the multiplication of multiple random variables may commonly appear in the environment, it does provide an intuitive framework.

The present contribution is based on a physico/chemical first order exponential kinetics model (Fig. 1B). Starting with an ordinary differential equation (Eq. 1), we introduce a stochastically variable rate, resulting in a stochastic differential equation (Eq. 2). By solving the corresponding Fokker–Planck equation (FPE) (see Theory section and SI for details) we derive the log-normal distribution (Eq. 3), under certain assumptions about stochastic variability. Since exponential kinetics are commonly observed for many processes in the environment, this model provides a potential explanation for the relative ubiquity of lognormal distributions in the environment.

Observations of lognormal-like concentration distributions in the environment include a wide array of components (e.g., aerosols; bulk organic carbon; heavy metals; organic pollutants; minerals; trace gases; inorganic ions; humic matter; biomarkers; radionuclides; microplastics; pharmaceuticals and pesticides) in various environmental sub-compartments (e.g., groundwater; watersheds; urban air; precipitation; rocks; stormwater; indoor air; soils; wastewater; marine sediments; landfills; lake sediments; peat; glaciers; background air; biota and humans)<sup>2,3,12–14,16–38</sup>. However, even though log normality is relatively common in the environment, we emphasize that it by no means is generally observed<sup>10–12</sup>; see discussion in the next section.

Ascribing a data set to a specific parametric form, e.g., log-normal, may provide mechanistic insights, but also typically simplifies statistical treatment<sup>11</sup>. Among multiple approaches, a straight-forward approach to test for log-normality is by log-transforming the data and then use one of the many well-established tests for normality, e.g., the Shapiro–Wilks test<sup>9,39,40</sup>. Specific aspects of analysis of log-normal data sets, e.g., the impact of non-detects or data averaging, or data sampling strategies, have been discussed in some detail in the literature<sup>12,41–43</sup>. However, even though we may attribute a certain data set to a parametric distribution, there are many empirical situations where the data appears to follow a complex/un-known distribution. Can we identify some general principles that may help disentangle observed variability, even beyond common parametric forms?

**Three mechanisms influencing concentration distributions.** Any environmental system is governed by a multitude of different processes, occurring simultaneously. Fortunately, some mechanisms appears to be more prevalent than others, and we can make simplifications, allowing modeling, e.g., the currently presented model for log-normal dependence. To attempt a more general description for concentration distributions, it may prove useful to consider a few different principal mechanisms. Here we explore three such mechanisms:



**Figure 2.** Numerical simulations (random sampling) of time-series concentration data and corresponding distributions, illustrating how different processes may influence concentration distributions. Panels (A)–(B): Log normally-distributed data (Eq. 3), e.g., influenced by exponential kinetics. Panels (C)–(D): Data reflecting the random fluctuation between three different log-normally distributed states (green, blue and red), yielding a multi-modal (mixture) distribution. Panels (E)–(F): Random mixing of the three distributions from panels (C)–(D), yielding a convolution distribution, tending towards a normal distribution. Panels (G)–(H): The log-normal distribution of panels (A)–(B), modulated by an oscillatory function (here sine function), illustrating the impact of, e.g. diurnal or seasonal cycles on observed distributions. The simulations were conducted using Matlab (ver. R2019b).

**Kinetics.** The model presented here is a common first order approximation of kinetics, e.g., of source and sink processes (Fig. 2A–B). In reality, dynamical systems may be highly complex, beyond the present formulation with a constant deterministic part ( $\mu_k$ , Eq. 2), with multiple coupled components, non-linearities, feedbacks and heterogeneous interactions. How such more complex kinetics influence concentration distributions is a topic for further scientific discovery.

**Mixing.** The concentration of a component in the environment is typically the sum of emissions from many sources. If these combine randomly, the concentration distribution approach normality (Fig. 2E–F). However, this type represents one limit of mixing—where the resulting distribution at each time point is the sum distribution (convolution) of many variables. Another limit is where the contributions from different sources are separated, such that the concentration at any given measurement point (e.g., time point) reflects one individual source (Fig. 2C–D). An example could be an atmospheric measurement site located between two cities: by wind direction each time point is mainly dominated by either city. In this limit the concentration data will tend to a multi-modal (or ‘mixture’) distribution.

**External variables.** In the model described in the Theory section the non-stochastic part of Eq. (2) is a constant ( $\mu_k$ ). But this parameter may also vary with external parameters, e.g., state variables. For instance, diurnal

or seasonal variations of temperature, pressure or sunlight, including phase transitions, may strongly influence concentration distributions (Fig. 2G–H).

These three classes of mechanisms influence the lifecycle of a component in the environment, but their relative importance regarding how concentration distributions are affected may be highly variable. For certain instances it may be possible to ascribe a certain parametric function to the data, e.g., a normal distribution for well-mixed sources or a log-normal suggesting a kinetic domain. However, for situations with more complex, multi-modal or broad distributions both statistical treatment and interpretations are more challenging. For instance, quite similarly looking distributions may emerge from rather different mechanisms, e.g., dependence on an externally oscillating parameter or a stochastically jumping mixture distribution (e.g., compare Fig. 2C–D with Fig. 2G–H).

**Outlook.** The shape of empirical concentration distributions is determined by the lifecycle dynamics and thereby contain information of environmental fate. In this paper we show that log-normality suggests influence by first order exponential kinetics (Eqs. 1 and 5; Fig. 1A–B). Reflecting the overall non-equilibrium state of environmental systems, there are a number of different processes/mechanisms that may exhibit exponential kinetics and thereby drive concentrations towards log-normality, e.g., emissions/chemical formation (e.g., of primary or secondary pollutants), degradation/decomposition (e.g., chemical reactions or deposition), kinetic transfer between different pools/layers/reservoirs (e.g., between the troposphere and the stratosphere) or kinetic partitioning (e.g., between gas and liquid phases). Some of these processes, e.g., sink kinetics, are active throughout the lifecycle of a component, and thereby continuously push concentrations towards log-normality.

Overall, the implications of log-normal concentration distributions span a broad spectrum of potential applications, ranging from data analysis methodologies, sampling strategies, emissions estimation, source apportionment calculations, modelling of chemical fate, estimation of toxic exposure to future climate scenarios<sup>9,25,29,41–52</sup>. Given the general mathematical formulation of the model (Eq. 2), we note that the applicability extends to log-normal concentration distributions also beyond Environmental systems<sup>53,54</sup>.

A specific example, where mechanistic insights may be derived from analysis of concentration distributions is the estimation of lifetimes ( $\tau = 1/k$ , Eq. 1)<sup>55,56</sup>. Our present findings suggest that observation of log-normally distributed concentration data, perhaps especially at remote or receptor sites, may indicate sink kinetics, and may therefore provide insights into the sink rate (Eqs. 4a–b and 5). Another specific situation where concentration distribution analysis is of importance is the common challenge of trend detection in environmental data<sup>11</sup>. For certain time-series data, log-transformation is commonly used prior to statistical analysis<sup>51,57</sup>. The present analysis may provide a physico-chemical motivation for such transformations, potentially contributing to interpretations. A third specific example regards the analysis of ratios of diagnostic markers, which are common tools to assess sources and processes across Earth and Environmental sciences. Examples include ratios of chemical markers and isotope signatures, where the latter commonly are reported as concentration ratios. The ratio of two log-normal random variables is another log-normal variable<sup>58</sup>. We may then predict that if the overall concentrations are log-normal, then, e.g., isotope signatures should also be log-normal, with potential implications for data analysis methodology and interpretation.

Finally, we note that rates in environmental systems often are proportional to concentrations, e.g., reactants. As an example, the sink rate for CO is proportional to the OH concentration (see Theory). Consequently, if the concentration is log-normal, so is the rate. Furthermore, fluxes (e.g., emission or sink fluxes) are often defined as the product of a rate and a concentration (or at least the amount, by some proportional measure). Taken together, this suggests that log-normal concentration distributions may imply log-normal distributions also for rates and/or fluxes for certain systems, with implications for, e.g., emission estimation or box models<sup>47,49,59</sup>.

All-in-all, this paper presents a mechanistic model for how log-normal concentrations may emerge in the environment, which in turn suggests an explanation for the relative abundance.

Received: 17 March 2021; Accepted: 23 July 2021

Published online: 12 August 2021

## References

- Gaddum, J. H. Lognormal distributions. *Nature* **156**, 463–466 (1945).
- Ahrens, L. H. The lognormal distribution of the elements (A fundamental law of geochemistry and its subsidiary). *Geochim. Cosmochim. Acta.* **5**, 49–73 (1954).
- Crow, E. L. & Shimizu, K. Lognormal distributions: Theory and Applications. ISBN: 0367580276. (CRC Press, 1988).
- Limpert, E., Stahel, W. A. & Abbt, M. Log-normal distributions across the sciences: Keys and clues. *Bioscience* **51**, 341–352 (2001).
- Black, F. & Scholes, M. The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**, 637–645 (1973).
- Kahn, H. D. Note on the distribution of air pollutants. *J. Air Poll. Contr. Assoc.* **23**, 973 (1973).
- Esmen, N. A. & Hamad, Y. Y. (1977) Log-normality of environmental sampling data. *J. Environ. Sci. Health Part A.* **12**, 29–41 (1977).
- Ott, W. R. A. physical explanation of the lognormality of pollutant concentrations. *J. Air Waste Manage. Assoc.* **40**, 1378–1383 (1990).
- Singh, A. K., Singh, A., Engelhardt, M. The lognormal distribution in environmental applications. EPA/600/R-97/006 (US EPA, 1997).
- Reimann, C. & Filzmoser, P. Normal and lognormal data distribution in geochemistry: Death of a myth. Consequences for the statistical treatment of geochemical and environmental data. *Environ. Geol.* **39**, 1001–1014 (1999).
- Chandler, R. & Scott, M. Statistical Methods for Trend Detection and Analysis in the Environmental Sciences. ISBN: 9780470015438 (Wiley, 2011).
- Bencala, K. E. & Seinfeld, J. H. On frequency distributions of air pollutant concentrations. *Atmos. Environ.* **10**(10), 941–950 (1976).
- Georgopoulos, P. G. & Seinfeld, J. H. Statistical distributions of air pollutant concentrations. *Environ. Sci. Technol.* **16**, 401–416 (1982).

14. Parrish, D. D. *et al.* Carbon Monoxide concentrations and their relation to concentrations of total reactive oxidized nitrogen at two rural US sites. *J. Geophys. Res.* **96**, 9309–9320 (1991).
15. Zwanzig, R. Non-equilibrium statistical mechanics. ISBN: 978-0195140187 (Oxford University Press, 2001).
16. Vistelius, A. The skew frequency distributions and the fundamental law of the geochemical processes. *J. Geol.* **68**, 1–22 (1960).
17. Noshkin, V. E., Eagle, R. J. & Wong, K. M. Plutonium levels in Kwajalein Lagoon. *Nature* **262**, 745–748 (1976).
18. Giesy, J. P. & Wiener, J. G. Frequency distributions of trace metal concentrations in five freshwater fishes. *Transac. Am. Fish. Soc.* **106**, 393–403 (1977).
19. Skeffington, R. A. The chemistry of bulk precipitation at a site in southeast England—1. Small-scale spatial variation, frequency distributions and variation with time. *Atmos. Environ.* **18**, 1683–1693 (1984).
20. Nero, A. V. *et al.* Distributions of airborne Radon-222 concentrations in US homes. *Science* **234**, 992–997 (1986).
21. Rice, J. A. & MacCarthy, P. Statistical evaluation of the elemental composition of humic substances. *Org. Geochem.* **17**, 635–648 (1991).
22. Van Buren, M. A., Watt, W. E. & Marsalek, J. Application of the log-normal and normal distributions to stormwater quality parameters. *Water Res.* **31**, 95–104 (1997).
23. Wilcke, W. *et al.* Urban soil contamination in Bangkok: heavy metal and aluminium partitioning in topsoils. *Geoderma* **86**, 211–228 (1998).
24. O'Neill, N. T. *et al.* The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data. *Geophys. Res. Lett.* <https://doi.org/10.1029/2000GL011581> (2000).
25. Saito, H. & Goovaerts, P. Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. *Environ. Sci. Technol.* **34**, 4228–4235 (2000).
26. Efrogmson, R. A., Sample, B. E. & Suter, G. W. II. Uptake of inorganic chemicals from soil by plant leaves: Regressions of field data. *Environ. Toxicol. Chem.* **20**, 2561–2571 (2001).
27. Jönsson, A. *et al.* Global accounting of PCBs in the continental shelf sediments. *Environ. Sci. Technol.* **37**, 245–255 (2003).
28. Maltby, L. *et al.* Insecticide species sensitivity distributions: importance of test species selection and relevance to aquatic ecosystems. *Environ. Toxicol. Chem.* **24**, 379–388 (2005).
29. Chen, S.-C. & Liao, C.-M. Health risk assessment on humans exposed to environmental polycyclic aromatic hydrocarbons pollution sources. *Sci. Tot. Environ.* **366**, 112–123 (2006).
30. Lissemore, L. *et al.* An exposure assessment for selected pharmaceuticals within a watershed in Southern Ontario. *Chemosphere* **64**, 717–729 (2006).
31. Liu, D. *et al.* Spatial distribution of soil organic carbon and analysis of related factors in croplands of the black soil region Northeast China. *Agricult. Ecosyst. Environ.* **113**, 73–81 (2006).
32. Döscher, A. *et al.* A 130 years deposition record of sulfate, nitrate and chloride from a high-alpine glacier. *Water Air Soil Pollut.* **85**, 603–609 (1995).
33. Dodds, W. K. *et al.* Eutrophication of US freshwaters: Analysis of potential economic damages. *Environ. Sci. Technol.* **43**, 12–19 (2009).
34. Bossew, P. Radon: Exploring the log-normal mystery. *J. Environ. Radioact.* **101**, 826–834 (2010).
35. Pulles, T. *et al.* Emission factors for heavy metals from diesel and petrol used in European vehicles. *Atmos. Environ.* **61**, 641–651 (2012).
36. McClintock, N. Assessing soil lead contamination at multiple scales in Oakland, California: Implications for urban agriculture and environmental justice. *Appl. Geography.* **35**, 460–473 (2012).
37. Pleil, J. D. *et al.* Estimating common parameters of lognormally distributed environmental and biomonitoring data: Harmonizing disparate statistics from publications. *J. Toxicol. Environ. Health B. Crit. Rev.* **17**, 341–368 (2014).
38. Chin, K. F. *et al.* Statistical analysis of trace contaminants measured in biogas. *Sci. Tot. Environ.* **729**, 138702 (2020).
39. Shapiro, S. S. & Wilks, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
40. Razali, N. M. & Wah, Y. B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Mod. Anal.* **2**, 21–33 (2011).
41. Leidel, N. H., Busch, K. A., Lynch, J. R. Occupational exposure sampling manual. DHHS (NIOSH) Publication Number 77-173 (1977).
42. Rappaport, S. M. Assessment of long-term exposures to toxic substances in air. *Ann. Occup. Hyg.* **35**, 61–121 (1991).
43. Antweiler, R. C. & Taylor, H. E. Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: I summary statistics. *Environ. Sci. Technol.* **42**, 3732–3738 (2008).
44. Pedit, J. & Miller, C. T. Heterogeneous sorption processes in subsurface systems. 1. Model formulations and applications. *Environ. Sci. Technol.* **28**, 2094–2104 (1994).
45. Di Toro, D. M. *et al.* Biotic ligand model of the acute toxicity of metals. 1. Technical basis. *Environ. Tox. Chem.* **20**, 2383–2396 (2001).
46. MacLeod, M., Fraser, A. J. & Mackay, D. Evaluating and expressing the propagation of uncertainty in chemical fate and bioaccumulation models. *Environ. Tox. Chem.* **21**, 700–709 (2002).
47. Zhao, Y. *et al.* Quantifying the uncertainties of a bottom-up emission inventory of anthropogenic atmospheric pollutants in China. *Atmos. Chem. Phys.* **11**, 2295–2308 (2011).
48. Massoudieh, A. & Kayhanian, M. Bayesian chemical mass balance method for surface water contaminant source apportionment. *J. Environ. Eng.* [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000645](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000645) (2013).
49. Frankenberg, C. *et al.* Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. *Proc. Nat. Acad. Sci.* **113**, 9734–9739 (2016).
50. Rogelj, J. *et al.* Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development in *Global warming of 1.5°C*. ISBN: 978-92-9169-151-7 (IPCC, 2018).
51. Liao, K. *et al.* A multiple linear regression model with multiplicative log-normal error term for atmospheric concentration data. *Sci. Tot. Environ.* **767**, 144282 (2021).
52. Wolf, R. & Tollefsen, K. E. A Bayesian approach to incorporating spatiotemporal variation and uncertainty limits into modeling of predicted environmental concentrations from chemical monitoring campaigns. *Environ. Sci. Technol.* **55**, 1699–1709 (2021).
53. Bengtsson, M. *et al.* Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res.* **15**, 1388–1392 (2005).
54. Lu, P. *et al.* Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotech.* **1**, 117–124 (2008).
55. Junge, C. E. Residence time and variability of tropospheric trace gases. *Tellus* **26**, 477–488 (1974).
56. Jobson, B. T. *et al.* Trace gas mixing ratio variability versus lifetime in the troposphere and stratosphere: Observations. *J. Geophys. Res.* <https://doi.org/10.1029/1999JD900126> (1999).
57. Collaud Cohen, M. *et al.* Aerosol decadal trends—Part 1: In-situ optical measurements at GAW and IMPROVE stations. *Atmos. Chem. Phys.* **13**, 869–894 (2013).
58. Springer, M. D. The algebra of random variables. ISBN: 0471014060 (Wiley, 1979).
59. Forney, D. C. & Rothman, D. H. Common structure in the heterogeneity of plant-matter decay. *J. R. Soc. Interface.* **9**, 2255–2267 (2012).

## Acknowledgements

This work was supported by research grants from the Swedish Research council (VR Contract No. 2017-05687 and 2020-05384) and the Swedish Research Council for Sustainable Development (FORMAS Contract No. 2020-01951). I thank Jens Danielsson for valuable discussions.

## Author contributions

A.A. came up with the idea, carried out research and wrote the paper.

## Funding

Open access funding provided by Stockholm University.

## Competing interests

The author declares no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96010-6>.

**Correspondence** and requests for materials should be addressed to A.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021