Superstatistical approach to air pollution statistics

Griffin Williams,1 Benjamin Schäfer,1,∗ and Christian Beck1,†

1School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom

Air pollution by Nitrogen Oxides (NOx) is a major concern in large cities as it has severe adverse health effects. However, the statistical properties of air pollutants are not fully understood. Here, we use methods borrowed from non-equilibrium statistical mechanics to construct suitable superstatistical models for air pollution statistics. In particular, we analyze time series of Nitric Oxide (NO) and Nitrogen Dioxide (NO2) concentrations recorded at several locations throughout Greater London. We find that the probability distributions of concentrations have heavy tails and that the dynamics is well-described by χ²-superstatistics for NO and inverse-χ²-superstatistics for NO₂. Our results can be used to give precise risk estimates of high-pollution situations and pave the way to mitigation strategies.

Complex driven non-equilibrium systems with time scale separation are often well-described by superstatistical methods, i.e. by mixing several dynamics on distinct time scales and constructing effective statistical mechanics models out of this mixing [1, 2]. The intensive parameter β that fluctuates in a superstatistical way can be an inverse temperature of the system, a fluctuating diffusion constant, or simply a local variance parameter in a given time series generated by the complex system under consideration. This formalism is in particular relevant for heterogeneous spatio-temporally varying systems and has been successfully applied to many areas of physics and beyond, most notably Lagrangian turbulence [3], defect turbulence [4], wind velocity fluctuations [5, 6], share price dynamics [7], diffusion of complex biomolecules [8], frequency fluctuations in power grids [9], rainfall statistics [10, 11], and many more. Here, we apply a superstatistical analysis to an important topic of high relevance, namely the spatio-temporally dynamics of air pollution in big cities. Our main example of pollutants considered in the following are NO and NO₂, but the method can be similarly applied to other substances.

NO and NO₂ are examples of nitrogen oxides (NOₓ), which are gaseous air pollutants that are primarily discharged through combustion [12]. In urban areas, the main cause of this is through automobile emissions [13]. These chemicals have been shown to aggravate asthma and other respiratory symptoms and increase mortality. Further, short term exposure to NO₂ has been shown to correlate with ischaemic and hemorrhagic strokes [12, 14, 15]. While NO is a very important substance within Earth’s ecosystem, for example regulating the development in plants and acting as a signaling hormone in certain processes in animals, it can also cause severe environmental damage in high concentrations [16–18]. NO may form acids, e.g. by chemically mixing with water (H₂O) to form Nitric Acid (HNO₃), or reacting further to NO₂ [19] it can cause acid rain [15]. Though NOₓ are not thought to be carcinogenic, they indicate the presence of other harmful air pollutants [13].

The data analyzed here was taken from the publicly available London Air Quality Network (LAQN) website [20]. The site provides readings of pollutant levels at different locations throughout Greater London at different time intervals, with the 15-minute interval data analyzed here. Inspecting the measured concentration time series for both NO and NO₂ reveals that both pollutants’ concentrations vary on multiple time scales, see Fig. 1.

Concentrating on the case of NO for now, we notice that the probability distribution of the measured concentration exhibits power-law tails, see Fig. 2. The distribution is well-fitted by a q-exponential [21] of the form

\[ p(u) = (2 - q)\lambda_q [1 + (q - 1)\lambda_q u]^{\frac{1}{q-1}}, \]

where \( u \) are the concentration levels [\( \mu g/m^3 \)] of the pollutant, and the \( q \) and \( \lambda_q \)-values are parameters. \( q \) can be regarded as an entropic index [21–23] and it is related to \( \lambda_q \) and the mean \( \mu \) by

\[ \lambda_q = \frac{1}{\mu(3 - 2q)}, \]

see Appendices A and B for more details.

The basic idea of the superstatistical modelling approach is that a given time series, generated by a complex driven non-equilibrium system, obeys a stochastic differential equation (SDE) where the parameters of the SDE change randomly as well, but on a much longer time scale [24]. In our case, the concentration time series, whilst being q-exponentially distributed as a whole, appears to be non-stationary, in the sense that it can be divided into shorter time slices, of length \( T \), that each have locally an exponential density with different relaxation constant \( \lambda_q \) each, see Fig. 3. These changes in the pollutants statistics make sense due to the changing environment, e.g. because of changing weather conditions or traffic flow varying from week to week and also across seasons. Observing simple (exponential) statistics locally, while noting heavy tails in the aggregated statistics clearly suggests a superstatistical description. One of the simplest local models is based on an exponential probability distribution of the form

\[ p(u|\beta) = \beta \cdot \exp(-\beta u), \]

where
where $\beta = \lambda_e$ is a local relaxation parameter that fluctuates on the larger superstatistical time scale.

We determine this large time scale by computing the average local kurtosis $\kappa$ of a cell of length $\Delta t$, similarly as done in [2] for locally Gaussian distributions. We use a local average kurtosis defined as

$$\kappa (\Delta t) = \frac{1}{t_{\text{max}}-\Delta t} \int_0^{t_{\text{max}}-\Delta t} dt_0 \langle (u - \bar{u})^4 \rangle_{t_0,\Delta t} / \langle (u - \bar{u})^2 \rangle_{t_0,\Delta t}^2,$$  \hspace{1cm} (4)

where $t_{\text{max}}$ is the full length of the time series. The notation $\langle \ldots \rangle_{t_0,\Delta t}$ indicates the expectation for the time slice of length $\Delta t$ starting at $t_0$. Assuming local exponential distributions with kurtosis 9, we apply Equation (4) to determine the long time scale $T$ as the special time length $\Delta t$ such that

$$\kappa (\Delta t) = \kappa (T) = 9.$$  \hspace{1cm} (5)

We show an example of how $T$ is calculated in Appendix B, with a plot illustrating how the average local kurtosis depends on $\Delta t$. Once the time slice length $T$ is calculated, we define an intensive parameter $\beta$ for each local cell, by setting $\beta = \lambda_e = 1/\langle u \rangle_{t_0,T}$. The distribution of this parameter, $f(\beta)$, is then obtained from a histogram, as shown in Fig. 4a. In good approximation, we find $\beta$ to be $\chi^2$ distributed, i.e.,

$$f(\beta) = \frac{1}{\Gamma\left(\frac{q}{2}\right)} \left( \frac{n \beta_0}{2\beta_0} \right)^{\frac{q}{2} - 1} \exp \left( -\frac{n \beta}{2\beta_0} \right),$$  \hspace{1cm} (6)

where $n$ is the number degrees of freedom and $\beta_0$ is the mean of $\beta$.

Integrating out the $\beta$-parameter, the marginal distribution $p(u)$ is calculated as

$$p(u) = \int_0^{\infty} p(u|\beta)f(\beta)\, d\beta,$$  \hspace{1cm} (7)

which evaluates to

$$p(u) = (2-q)\lambda_q \left[ 1 + (q-1)\lambda_q u \right]^{-\frac{1}{1-q}},$$  \hspace{1cm} (8)

with

$$-\frac{n}{2} - 1 = \frac{1}{1-q},$$  \hspace{1cm} (9)

and

$$\frac{1}{2}(q-1)\lambda_q = \frac{\beta_0}{n}. $$  \hspace{1cm} (10)
Figure 3. NO concentrations locally follow exponential distributions. a: We plot the NO concentration for several months in 2010, with two periods of length $\Delta t = T \approx 6$ days highlighted, whose distributions are explored in panels b and c. The histograms of panels b and c are fitted with exponential distributions and we note their respective $\lambda_e$-values. Note that the exponent in panel b is only one fourth of the exponent recovered in panel c, i.e. high pollution levels are much more likely to be observed in panel b, a time period in January, than in panel c, a time period in June. These different exponents quantify the observation that pollutant levels tend to be higher during winter than during summer, see also Fig. 1.

Eq. (8) is the $q$-exponential distribution, which we derived from a superposition of local exponential distributions.

Fitting the $n$-parameter to find $f(\beta)$, as illustrated in Fig. 4a, we determine the $q$- and $\lambda_q$-parameters from Eqs. (9) and (10), respectively. Inserting the $\chi^2$-distribution $f(\beta)$, we use (7) and (8) to compute a probability density, which both approximates the fitted $q$-exponential and the data very well, successfully serving as a consistency check of the superstatistical approach, see Fig. 4b.

We already noticed different statistical behavior when inspecting the trajectory of the NO$_2$ data in Fig. 1, as compared to that of NO. Following a similar procedure as for the NO data, we find that for NO$_2$ the aggregated distribution is approximated by a $q$-Maxwell Boltzmann distribution, of the form

$$p(u) = \frac{1}{Z} u^2 \sigma_q^{3/2} \left[ 1 + (q - 1)\sigma_q u^2 \right]^{\frac{1}{1-q}},$$

where $Z$ is the normalization factor given by

$$Z = \frac{\sqrt{\pi} \Gamma \left( \frac{5-3q}{2(q-1)} \right)}{4(q-1)^{3/2} \Gamma \left( \frac{1}{q-1} \right)}$$

and where $q$ is related to the scale parameter $\sigma_q$ and the mean $\mu$ by

$$\sigma_q = \left( \frac{2(q-1)^{3/2} \Gamma \left( \frac{1}{q-1} \right)}{\mu \sqrt{\pi} (2-q) (3-2q) \Gamma \left( \frac{5-3q}{2(q-1)} \right)} \right)^2,$$

see Appendix C for details.

Analogously to the NO-case, we explain the aggregated distribution as a superposition of simple local distributions, here chosen as ordinary Maxwell-Boltzmann distributions. Each local Maxwell-Boltzmann distribution is defined as

$$p(u) = \sqrt{\frac{16}{\pi}} u^2 \sigma_{mb}^{3/2} \exp \left( -\sigma_{mb} u^2 \right),$$

where $\sigma_{mb}$ is the scale parameter of the local Maxwell-Boltzmann distribution.
where $\sigma_{mb} = \beta$ is a scale parameter.

We determine the long time scale $T$ by varying $\Delta t$ so that we locally obtain the kurtosis of a Maxwell-Boltzmann distribution, which is given as

$$\kappa(T) = \frac{15\pi^2 + 16\pi - 192}{(3\pi - 8)^2} \approx 3.1082,$$

(15)

see Appendix C for the calculation. If the $\beta$ values again follow a $\chi^2$-distribution, the superpositioned Maxwell-Boltzmann distributions would lead to an exact $q$-Maxwell-Boltzmann distribution, see Appendix C for a detailed calculation.

The results of our superstatistical analysis for the NO$_2$ data are displayed in Fig. 5. The $\chi^2$-distribution is only a rough fit and the histogram of $\beta$ is instead best fitted by an inverse-$\chi^2$-distribution

$$f(\beta) = \frac{\beta_0}{\Gamma\left(\frac{q}{2}\right)} \left(\frac{n\beta_0}{2}\right)^{\frac{q}{2}} \beta^{-\frac{q}{2}} \exp\left(-\frac{n\beta_0}{2\beta}\right).$$

(16)

This leads to a different superstatistics, compared to the case of NO. Integrating the conditional probability $p(u|\beta)$, as in (7) but with local Maxwell-Boltzmann distributions $p(u|\beta)$ and an inverse-$\chi^2$-distribution $f(\beta)$, gives the marginal distribution as

$$p(u) = \sqrt{\frac{64}{\pi}} \frac{\beta_0}{\Gamma\left(\frac{q}{2}\right)} \left(\frac{n\beta_0}{2}\right)^{\frac{n+1}{2}} u^{\frac{n+3}{2}} K_{1-n} \left(u\sqrt{2n\beta_0}\right),$$

(17)

where $K_v(z)$ is the modified Bessel function of the second order. This inverse-$\chi^2$-superstatistical model leads to a very good description of the data, see Fig. 5, and it approximates $q$-Maxwell-Boltzmann distributions for medium concentrations but it not exactly the same. In fact, the PDF of the observed statistics decays like $p(u) \sim \exp(-\text{const} \sqrt{u})$ for large $u$ values, see also Appendix C.

To conclude, we have illustrated that superstatistical methods, originally introduced in non-equilibrium statis-
We start by providing an overview of the different local distributions (exponential or Maxwell-Boltzmann) and the \( \beta \)-distributions of the individual scale parameters:

$$\mu = \langle u \rangle = (2 - q) \lambda_q \int_0^\infty u [1 + (q-1) \lambda_q u]^{\frac{1}{q-1}} du$$

$$= \frac{(2 - q) \Gamma(2) \Gamma \left( \frac{3 - 2q}{q-1} \right)}{\lambda_q (q-1)^2 \Gamma \left( \frac{1}{q-1} \right)}$$

$$= \frac{1}{\lambda_q (3 - 2q)}$$

which gives

$$\lambda_q = \frac{1}{\mu (3 - 2q)}.$$  \hspace{1cm} (B5)

\section{Calculation of mean \( \mu \) of Exponential Distribution}

We briefly recall the relationship between the exponential scale parameter \( \beta = \lambda_e \) and the mean \( \mu \) as

$$\mu = \langle u \rangle = \beta \int_0^\infty u \exp(-\beta u) \, du$$

$$= \frac{1}{\beta}.$$  \hspace{1cm} (B6)
3. Calculation of kurtosis $\kappa$ of Exponential Distribution

In the main text, we determined the time scale $T$ as the time scale where the local kurtosis of the local exponential distribution takes on the value $\kappa = 9$. Here, we provide the corresponding calculation

$$\kappa = \frac{\langle (u - \mu)^4 \rangle}{\langle (u - \mu)^2 \rangle^2} = \frac{\int_0^\infty (u - \mu)^4 \beta \exp(-\beta u) \, du}{\left( \int_0^\infty (u - \mu)^2 \beta \exp(-\beta u) \, du \right)^2} = \frac{\int_0^\infty (u^4 - 4u^3\beta + 6u^2\beta^2 - 4u\beta^3 + \beta^4) \exp(-\beta u) \, du}{\left( \int_0^\infty (u^2 - 2u\beta + \beta^2) \exp(-\beta u) \, du \right)^2} = \frac{24 - 24\mu\beta + 12\mu^2\beta^2 - 4\mu^3\beta^3 + \mu^4\beta^4}{4 - 8\mu\beta + 8\mu^2\beta^2 - 4\mu^3\beta^3 + \mu^4\beta^4}$$

$$= 9,$$

(B7)

using $\mu = \frac{1}{\beta}$.

4. Superstatistical calculation of $q$-Exponential Distribution

We show how integrating several local exponential distributions, whose exponents $\beta$ follow a $\chi^2$-distribution, leads to an overall $q$-exponential distribution: Each local distribution is given as

$$p(u|\beta) = \beta \exp(-\beta u).$$

(B8)

Integrating over all of these distributions can be expressed as

$$p(u) = \int_0^\infty p(u|\beta) f(\beta) \, d\beta = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left( \frac{n}{2\beta_0} \right)^{\frac{n}{2}} \int_0^\infty \beta^{\frac{n}{2}} \exp\left(-\beta \left( \frac{n}{2\beta_0} + u \right) \right) \, d\beta,$$

(B9)

which we evaluate to

$$p(u) = (2 - q)\lambda_q \left[ 1 + (q - 1)\lambda_q u \right]^{-\frac{1}{q}},$$

(B10)

if we identify

$$-\left( \frac{n + 2}{2} \right) = \frac{1}{1 - q},$$

(B11)

$$\frac{1}{2}(q - 1)\lambda_q = \frac{\beta_0}{n},$$

(B12)

where $\beta_0$ is the mean of $\beta$.

5. Determining the long time scale $T$

We determine the long time scale $T$ from the time series using Eq. (4) from the main text given as

$$\kappa(\Delta t) = \frac{1}{t_{max} - \Delta t} \int_0^{t_{max} - \Delta t} du \left( \frac{\langle (u - \bar{u})^4 \rangle}{\langle (u - \bar{u})^2 \rangle^2} \right)_{\bar{u}, \Delta t}.$$

(B13)

The long time scale $T$ is defined as $T := \Delta t$ such that $\kappa(\Delta t) = 9$. Here, we compute the average local kurtosis $\kappa$ as a function of the time window $\Delta t$ and thereby determine $T$.

Figure 6. The average kurtosis $\kappa$ as given by Eq.(4) is plotted as a function of the time window $\Delta t$ (blue). The crossing between the horizontal line at $\kappa = 9$ (the kurtosis of an exponential distribution) and the $\kappa$ vs. $\Delta t$ curve gives the value for $\Delta t = T \approx 6$.

Appendix C: NO$_2$

In the case of NO$_2$ concentrations, we approximate local concentrations as Maxwell-Boltzmann distributions

$$p(u) = \frac{1}{\pi} \frac{16}{u^2} \sigma_{mb}^{3/2} \exp\left(-\sigma_{mb} u^2\right)$$

(C1)

and the aggregated statistics is a $q$-Maxwell-Boltzmann distribution

$$p(u) = \frac{1}{\pi} \frac{16}{u^2} \sigma_{qmb}^{3/2} \left[ 1 + (q - 1)\sigma_{qmb} u^2 \right]^{-\frac{1}{q}}$$

(C2)

if $\sigma_{mb}$ is $\chi^2$ distributed. These $q$-Maxwell-Boltzmann distributions would strictly arise if the scale parameters $\beta = \sigma_{mb}$ were following a $\chi^2$-distribution. Empirically, we observe that for our data, the $\beta$ values follow an inverse-$\chi^2$-distribution instead:

$$f(\beta) = \frac{\beta_0}{\Gamma\left(\frac{n}{2}\right)} \left( \frac{n\beta_0}{2} \right)^{\frac{n}{2}} \beta^{-\frac{n}{2} - 2} \exp\left(-\frac{n\beta_0}{2\beta}\right).$$

(C3)
We obtain this inverse-$\chi^2$-distribution from a simple transformation of random variables when $\beta^{-1}$, rather than $\beta$, is $\chi^2$-distributed.

1. Calculation of $\sigma_q$

Here, we show how to express the scale parameter of $q$-Maxwell-Boltzmann distributions $\sigma_q$ as a function of the mean of the distribution $\mu = \langle u \rangle$ and the $q$-parameter. We start by computing the mean:

$$\mu = \langle u \rangle = \frac{\sigma_q^{3/2}}{Z} \int_0^\infty u^3 \left[1 + (q-1)\sigma_q u^2\right]^{-1/q} \, du$$

$$= \frac{\Gamma(2)\Gamma\left(\frac{3-2q}{q-1}\right)}{2Z\sigma_q^{1/2}(q-1)^2\Gamma\left(\frac{1}{q-1}\right)}$$

$$= \frac{2(2-q)(3-2q)\sigma_q^{1/2}}{Z(2-q)(3-2q)\Gamma\left(\frac{5-3q}{q-1}\right)}$$

which gives

$$\sigma_q = \left(\frac{2(2-q)(3-2q)\Gamma\left(\frac{1}{q-1}\right)}{\mu \sqrt{\pi(2-q)(3-2q)\Gamma\left(\frac{5-3q}{q-1}\right)}}\right)^2.$$  

(C4)

2. Calculation of mean $\mu$ of Maxwell-Boltzmann Distribution

We briefly recall the relationship between the scale parameter $\beta = \sigma_{mb}$ of the Maxwell-Boltzmann distribution and the mean $\mu$ as

$$\mu = \langle u \rangle = \sqrt{\frac{16}{\pi}} \beta^{3/2} \int_0^\infty u^3 \exp(-\beta u^2) \, du$$

$$= \frac{16}{\pi} \beta^{3/2} \left(\frac{\Gamma(2)}{2\beta^2}\right)$$

$$= \frac{2}{\sqrt{\pi} \beta}$$

(C6)

3. Calculation of kurtosis $\kappa$ of Maxwell-Boltzmann Distribution

In the main text, we determined the time scale $T$ as the time scale where the local kurtosis of the local Maxwell-Boltzmann distribution takes on the value $\kappa \approx 3.1082$.

Here, we provide the corresponding calculation

$$\kappa = \frac{(u - \mu)^4}{(u - \mu)^2/2}$$

$$= \frac{16}{\pi} \beta^{3/2} \int_0^\infty (u - \mu)^4 u^2 \exp(-\beta u^2) \, du$$

$$= \frac{16}{\pi} \beta^{3/2} \left(\frac{1}{(2\beta_0)^2} \left(\frac{n}{2\beta_0}\right)^2 u^2 \right. \times \int_0^\infty \beta^{n+1} \exp\left(\beta \left(\frac{n}{2\beta_0} + u^2\right)\right) \, d\beta,$$

which can be evaluated to

$$p(u) \sim u^2 \sigma_q^{3/2} \left[1 + (q-1)\sigma_q u^2\right]^{-1/q},$$

(C10)

if we identify

$$-\left(\frac{n + 3}{2}\right) = \frac{1}{1 - q},$$

(C11)

$$\frac{1}{2}(q - 1)\sigma_q = \frac{\beta_0}{n},$$

(C12)

where $\beta_0$ is the mean of $\beta$.

4. Superstatistical calculation of $q$-Maxwell-Boltzmann Distribution

We show how integrating several local Maxwell-Boltzmann distributions, whose exponents $\beta$ follow a $\chi^2$-distribution, leads to an overall $q$-Maxwell-Boltzmann distribution: Each local distribution is given as

$$p(u|\beta) = \frac{16}{\pi} u^2 \sigma_{mb}^{1/2} \exp\left(-\sigma_{mb} u^2\right).$$

(C8)

Integrating over all of these distributions can be expressed as

$$p(u) = \int_0^\infty p(u|\beta) f(\beta) \, d\beta$$

$$= \frac{16}{\pi} \beta^{3/2} \left(\frac{1}{(2\beta_0)^2} \left(\frac{n}{2\beta_0}\right)^2 u^2 \right. \times \int_0^\infty \beta^{n+1} \exp\left(\beta \left(\frac{n}{2\beta_0} + u^2\right)\right) \, d\beta,$$

which can be evaluated to

$$p(u) \sim u^2 \sigma_q^{3/2} \left[1 + (q-1)\sigma_q u^2\right]^{-1/q},$$

(C10)

5. Calculation of inverse-$\chi^2$ superstatistical distribution based on superimposed Maxwell-Boltzmann distributions

Instead of following a $\chi^2$-distribution, our data indicates that the scale parameters $\beta$ of the local Maxwell-Boltzmann distribution rather follow an inverse-$\chi^2$-distribution. Here, we derive the probability density...
function for such a superposition. Taking the conditional density $p(u|\beta)$ as given in (C8) and the inverse-$\chi^2$-distribution for the distribution $f(\beta)$, the probability density $p(u)$ is given as

$$
p(u) = \sqrt{\frac{16}{\pi}} \frac{\beta_0}{\Gamma(\frac{n+1}{2})} \left( \frac{n\beta_0}{2} \right)^{\frac{n+1}{2}} u^n \exp\left( -\frac{n\beta_0 u^2}{2\beta} - \beta u^2 \right) \, d\beta
$$

$$
= \sqrt{\frac{64}{\pi}} \frac{\beta_0}{\Gamma(\frac{n}{2})} \left( \frac{n\beta_0}{2} \right)^{\frac{n+1}{2}} u^{\frac{n+3}{2}} K_{\frac{n}{2}} \left( u \sqrt{2n\beta_0} \right).
$$

This PDF has still heavy tails, which decay as $p(u) \sim \exp(-\text{const} \sqrt{u})$ for large $u$ values.

6. Extra Figures

We repeat Fig. 3 from the main text for local $NO_2$ concentrations, i.e. analyzing brief periods of the $NO_2$ trajectory. Fig. 7 illustrates that in good approximation the local behavior follows Maxwell-Boltzmann distributions with fluctuating variance. Further, we also determine the long time scale $T$ for the local Maxwell-Boltzmann distributions by setting $T := \Delta t$ for the local kurtosis $\kappa(\Delta t) \approx 3.1082$, following Eq. (4) from the main text again. See Figs. 7 and 8.

Appendix D: Thresholds and policies

Policies to tackle pollution in cities often focus on thresholds and compliance with these thresholds. This then often leads to pollution just below the threshold in many regions. However, in many cases it may be more reasonable to reduce the overall exposure to pollutants [26]. Weather is responsible for many fast variations of the air pollution [26]. The focus on threshold is especially dangerous as there is strong evidence that even small concentrations of pollutants, below national standards, increase the death rate [27]. The UK government (Dept. of Health & Social Care) calls air pollution a "health emergency" and the World Health Organization (WHO) judges the situation similarly on a global scale [28]. Note that about 9500 premature deaths are attributed to air pollution in London alone every year [29]. On a global scale this number rises to about 4.9 million premature deaths, making air pollution the fifth leading cause of mortality worldwide [30].

So let us finally sketch how the knowledge of the probability density function (PDF) allows estimates of threshold crossings and total exposure. Setting thresholds on pollutant concentrations is a popular tool when setting pollution policies. Simultaneously, data coverage might not be very good in all locations. Especially if data is not available for the full time period of interest but only for a few months, estimates on how frequently thresholds are violated are difficult to obtain. As soon as we are able to determine the PDF $p(u)$ for the concentration levels $u$, we can easily compute the number of days where thresholds are violated as

$$
N_{u>u\text{threshold}} = 365 \int_{u\text{threshold}}^{\infty} p(u)du.
$$

This integral can be easily evaluated numerically, using $q$-exponential distributions for $NO$, see Eq. (8) from the main text, and Eq. (17) from the main text for $NO_2$, consistent with experimental observations.

The number of threshold violations explicitly depends on the parameters $\lambda_q$ and $q$, which have to be determined from the data sets but can then be compared between different locations and time periods. Having these two parameters explicitly disentangles two aspects: The rate of (rare) extreme events, encoded in $q$ and the overall variability of the pollutant concentration, encoded in $\lambda_q$ (or $\lambda_{mb}$ for $NO_2$).

Complementary to thresholds, policies could instead focus on the total pollution exposure (TPE) describing the average amount citizens are exposed to. Again, this is easily computed from the PDF as

$$
TPE = 365 \int_{0}^{\infty} p(u)duu.
$$

---

[1] C. Beck and E. G. D. Cohen, Physica A: Statistical mechanics and its applications 322, 267 (2003).
[2] C. Beck, E. G. D. Cohen, and H. L. Swinney, Physical Review E 72, 056133 (2005).
[3] C. Beck, Physical Review Letters 98, 064502 (2007).
[4] K. E. Daniels, C. Beck, and E. Bodenschatz, Physica D: Nonlinear Phenomena 193, 208 (2004).
[5] S. Rizzo and A. Rapisarda, in AIP Conference Proceedings (AIP, 2004), vol. 742, pp. 176–181.
[6] J. Weber, M. Reyers, C. Beck, M. Timme, J. G. Pinto, D. Witthaut, and B. Schäfer, arXiv preprint arXiv:1810.06391 (2018).
[7] P. Jizba and H. Kleinert, Physical Review E 78, 031122 (2008).
[8] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, Physical Review X 7, 021002 (2017).
[9] B. Schäfer, C. Beck, K. Aihara, D. Witthaut, and M. Timme, Nature Energy 3, 119 (2018).
[10] G. C. Yalcin, P. Rabassa, and C. Beck, Journal of Physics A: Mathematical and Theoretical 49, 154001 (2016).
[11] C. De Michele and F. Avanzu, Scientific Reports 8, 14204 (2018).
[12] E. D. Amster, M. Haim, J. Dubnov, and D. M. Broday, Environmental Pollution 186, 20 (2014).
Figure 7. NO$_2$ concentrations locally follow Maxwell-Boltzmann distributions. a: We plot the NO$_2$ concentration for several months in 2010, with two periods of length $\Delta t = T \approx 2.6$ days highlighted, whose distributions are explored in panels b and c. The histograms of panels b and c are fitted with Maxwell-Boltzmann distributions and we note their respective $\lambda_{mb}$-values, which are strongly varying in time.

Figure 8. The average kurtosis $\kappa$ is plotted as a function of the time window $\Delta t$. The crossing between the horizontal line at $\kappa = 3.1082$ (the kurtosis of a Maxwell-Boltzmann distribution) and the $\kappa$ vs. $\Delta t$ curve gives the value $\Delta t = T = 2.61$.

[13] G. B. Hamra, F. Laden, A. J. Cohen, O. Raaschou-Nielsen, M. Brauer, and D. Loomis, Environmental Health Perspectives 123, 1107 (2015).
[14] A. S. Shah, K. K. Lee, D. A. McAllister, A. Hunter, H. Nair, W. Whiteley, J. P. Langrish, D. E. Newby, and N. L. Mills, BMJ 350, h1295 (2015).
[15] United States Environmental Protection Agency, Nitrogen Dioxide (NO$_2$) Pollution, http://www.epa.gov/no2-pollution/basic-information-about-no2/
[16] P. Domingos, A. M. Prado, A. Wong, C. Gehring, and J. A. Feijo, Molecular Plant 8, 506 (2015).
[17] J. Astier, I. Gross, and J. Durner, Journal of Experimental Botany 69, 3401 (2017).
[18] M. M. Musameh, C. J. Dunn, M. H. Uddin, T. D. Sutherland, and T. D. Rapson, Biosensors and Bioelectronics 103, 26 (2018).
[19] M. E. Barsan (2007).
[20] Environmental Research Group (ERG) at King’s College London, London air, URL https://www.londonair.org.uk.
[21] C. Tsallis, Introduction to nonextensive statistical mechanics: approaching a complex world (Springer, 2009).
[22] R. Hanel, S. Thurner, and M. Gell-Mann, Proceedings of the National Academy of Sciences 108, 6390 (2011).
[23] P. Jizba and J. Korbel, Physical Review Letters 122, 120601 (2019).
[24] C. Beck, Physical Review Letters 87, 180601 (2001).
[25] L. L. Chen and C. Beck, Physica A: Statistical Mechanics and its Applications 387, 3162 (2008).
[26] G. W. Fuller and A. Font, Science 365, 322 (2019).
[27] Q. Di, Y. Wang, A. Zanobetti, Y. Wang, P. Koutrakis, C. Choirat, F. Dominici, and J. D. Schwartz, New England Journal of Medicine 376, 2513 (2017).
[28] World Health Organization (WHO), How air pollution is destroying our health (2019), URL https://www.who.int/air-pollution/news-and-events/how-air-pollution-is-destroying-our-health.
[29] H. Walton, D. Dajnak, S. Beevers, M. Williams, P. Watkiss, and A. Hunt, London: Kings College London, Transport for London and the Greater London Authority (2015).
[30] Health Effects Institute, Special Report (2018).