# Supplementary information to "Technical Note: Statistical generation of climate-perturbed flow duration curves"

Veysel Yildiz[1], Charles Rougé[1], Robert Milton[2], and Solomon Brown[2]

[1]Department of Civil and Structural Engineering, The University of Sheffield

[2] Department of Chemical and Biological Engineering, The University of Sheffield

**Correspondence:** Charles Rougé (c.rouge@sheffield.ac.uk)

**Abstract.** This supplementary information demonstrates that for triplets $(M,V,L)$ of streamflow statistics representing average behavior, variability, and low flows, there is unique parameterisation of the flow duration curve (FDC) according to the Kosugi model. We consider the "mean" case where $(M,V,L) = (\mu, \sigma, q_{low})$ where $\mu$ is the mean, $\sigma$ is the standard deviation and $q_{low}$ is the $1^{st}$ or $5^{th}$ percentile of flow, and the "median case" $(M,V,L) = (m, CV, q_{low})$ where $m$ is the median and $CV = \mu/\sigma$ is the coefficient of variation. It also provides conditions on $(M,V,L)$ for the existence of a parameterisation.

## 1 Kosugi function reminders

We model the flow duration curve (FDC) with the Kosugi equation, as proposed by Sadegh et al. (2016). The equation models streamflow $q$ as a function of the flow quantile $u \in [0,1]$:

$$q(u) = c + (a-c)\, z(u)^b \tag{1}$$

where $(a,b,c)$ are parameters, with $a$ and $c$ in the same units as $q$, and $b$ unitless. We need $a - c > 0$ and $b > 0$ for $z(u)$ is defined as follows, and represented in Figure S1:

$$z(u) = \exp\left[\sqrt{2}\,\mathrm{erfc}^{-1}(2u)\right] \tag{2}$$

This supplementary information will relate parameters $(a,b,c)$ to triplets of streamflow statistics $(M,V,L)$ representing average behavior, variability, and low flows. It will do so first in the "mean" case of $(M,V,L) = (\mu, \sigma, q_{low})$ in Section 2, where $\mu$ is the mean, $\sigma$ is the standard deviation and $q_{low}$ is the $1^{st}$ or $5^{th}$ percentile of flow. Then in Section 3 we will examine the "median case" of $(M,V,L) = (m, CV, q_{low})$, where $m$ is the median and $CV = \mu/\sigma$ is the coefficient of variation.

## 2 "Mean" case $(M,V,L) = (\mu, \sigma, q_{low})$

In this section we assume we know the streamflow mean $\mu$, standard deviation $\sigma$ and low flow percentile $q_{low}$. We assume we have $\mu > q_{low}$. We will prove the $(a,b,c)$ triplet of Kosugi parameter is unique, and give a sufficient condition on $(\mu, \sigma, q_{low})$ for its existence.

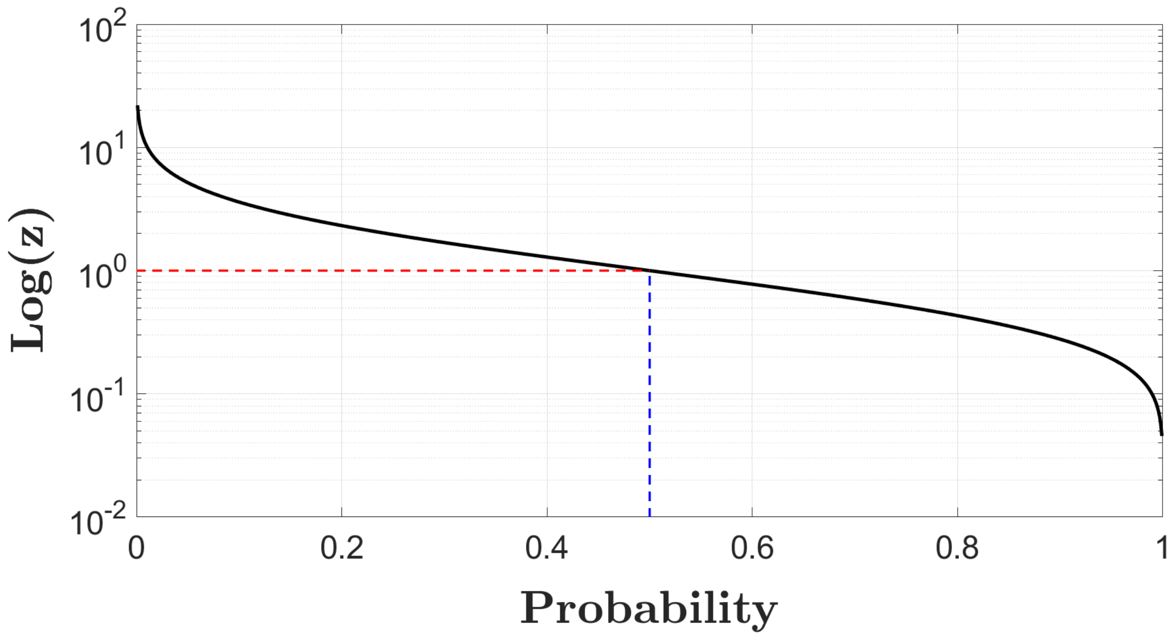


**Figure S1.** The $z(e)$ function.

### 2.1 Relating parameter triplets

Writing the definition of mean, standard deviation and low flow quantile for the Kosugi FDC yields three equations. For this, let us introduce the auxiliary function $f$:

$$f(b) = \int_0^1 z(u)^b \, du \tag{3}$$

We can then write the mean $\mu$ according to its definition as the integral of $q(u)$ for $u \in [0,1]$. Using the linearity properties of the integral yields:

$$\mu = c + (a-c) f(b) \tag{4}$$

Similarly, by definition of the variance, and using the definition of the mean above, we have:

$$\sigma^2 = \int_0^1 \left[c + (a-c) z(u)^b\right]^2 du - \left[c + (a-c) f(b)\right]^2 \tag{5}$$

Developing the squares and exploiting again the linearity of the integral enables us to simplify this into this definition of $V$:

$$\sigma = (a-c)\sqrt{f(2b) - f(b)^2} \tag{6}$$

Lastly, introducing $\varepsilon = z(q_{low})$ where $q_{low} = 0.99$ (respectively $0.95$) if we are interested in the first (resp. fifth) flow percentile, we have the following relationship for $q_{low}$:

$$q_{low} = c + (a-c)\,\varepsilon^b \tag{7}$$

### 2.2 Solution strategy

Clearly, for $b$ fixed, $(a,c)$ is the solution of a system of two linear equations. For instance, from equations (4) and (7), we get:

$$\begin{cases} \mu = c + (a-c) f(b) \\ q_{low} = c + (a-c)\,\varepsilon^b \end{cases} \tag{8}$$

which is equivalent to:

$$\begin{cases} a - c = \dfrac{\mu - q_{low}}{f(b) - \varepsilon^b} \\ a = \dfrac{q_{low}\,(f(b)-1) + \mu(1-\varepsilon^b)}{f(b)-\varepsilon^b} \\ c = \dfrac{q_{low}\, f(b) - \mu\,\varepsilon^b}{f(b)-\varepsilon^b} \end{cases} \tag{9}$$

We can then relate streamflow parameters $(\mu, \sigma, q_{low})$ to Kosugi parameter function of $b$ alone, by replacing $(a-c)$ into equation (6):

$$\frac{\sigma}{\mu - q_{low}} = \frac{\sqrt{f(2b) - f(b)^2}}{f(b) - \varepsilon^b} = \mathcal{F}(b) \tag{10}$$

Thus, whether we can find a unique triplet $(a,b,c)$ for $(\mu,\sigma,q_{low})$hinges on whether $\mathcal{F}(b)$ is monotonous for $b > 0$. Then existence will depend on (1) proving that $f(b) > \varepsilon^b$ for $b > 0$ (so $\mathcal{F}(b)$ is defined and positive), and (2) establishing the lower bond for $\mathcal{F}(b)$. For all of this, it would be easier to work with a simpler expression for $f(b)$. This is the topic of the next paragraph.

### 2.3 Simplifying $f(b)$

Let us operate a variable change $x = \text{erfc}^{-1}(2u)$ in the integral that defines $f(b)$. Then for $u = 0$, we have $x = +\infty$, for $u = 1$ we have $x = -\infty$. We also have $u = \text{erfc}(x)/2 = (1 - \text{erf}(x))/2$. Using the derivation of the error function to relate $du$ and $dx$ we can therefore write:

$$f(b) = \int_0^1 \exp\left[\sqrt{2}\,\text{erfc}^{-1}(2u)\right]^b du = \int_{+\infty}^{-\infty} e^{\sqrt{2}bx} \frac{-e^{-x^2}}{\sqrt{\pi}} dx \tag{11}$$

Which directly leads to:

$$f(b) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp\left(-(x - \frac{b}{\sqrt{2}})^2\right) e^{b^2/2} dx \tag{12}$$

Then a further change of variable $y = x - b/\sqrt{2}$ leads to:

$$f(b) = e^{b^2/2}\left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-y^2}\,dy\right) \tag{13}$$

and since the quantity between the parentheses is equal to $1$ this simplifies into:

$$f(b) = e^{b^2/2} \tag{14}$$

This remarkable equation simplifies the calculations going forward. It also demonstrates that $f(b) - \varepsilon^b = \exp(b^2) - \exp(b\ln(\varepsilon))$ is positive for $b > 0$, because $\ln(\varepsilon) < 0$.

### 2.4 Unicity of the parameterisation

Using the result from equation (14) into equation (10), we can write:

$$\mathcal{F}(b) = \frac{g(b)}{h(b)} \tag{15}$$

with:

$$\begin{cases} g(b) = \sqrt{e^{b^2} - 1} \\ h(b) = 1 - e^{-b^2/2}\varepsilon^b \end{cases} \tag{16}$$

Recall that to demonstrate unicity of the Kosugi parameterisation, it is enough to show that for $b > 0$, $\mathcal{F}(b)$ grows monotonically with $b$. Derivation with respect to $b$ yields:

$$\begin{cases} g'(b) = \dfrac{b\,e^{b^2}}{g(b)} \\ h'(b) = (b - \ln(\varepsilon))e^{-b^2/2}\varepsilon^b \end{cases} \tag{17}$$

where $\ln$ is the base $e$ logarithm, $(\ln(e) = 1)$. Since $g(b) > 0$, we can write for $b > 0$:

$$\mathcal{F}'(b) = \frac{1}{g(b)h^2(b)}\left(g(b)g'(b)h(b) - g^2(b)h'(b)\right) \tag{18}$$

Thanks to the above equation, $\mathcal{F}'(b)$ has the same sign as $\mathcal{U} = gg'h - g^2h'$. $\mathcal{U}(b)$ is given by:

$$\mathcal{U}(b) = be^{b^2} + \varepsilon^b\left[b(e^{-b^2/2} - 2e^{b^2/2}) + \ln(\varepsilon)\left(e^{b^2/2} - e^{-b^2/2}\right)\right] \tag{19}$$

Figure S2 graphically shows that $\ln(\mathcal{U}(b)) > 0$ for $b > 0$, in both cases where $\varepsilon = z(0.99)$ (if $q_{low}$ is the first percentile) or $\varepsilon = z(0.95)$ (if $q_{low}$ is the first percentile). We also represented $e^{b^2}$ on Figure S2, since it becomes the dominant term in $\mathcal{U}(b)$ as $b$ grows farther from 0. It is therefore clear that $\mathcal{F}(b)$ grows with $b$ when $b > 0$, and that therefore, there is at most a unique $b$ solution of equation (10). Equation (9) provides unique $a$ and $c$ for a value of $b$. This enables us to conclude on the uniqueness of the Kosugi parameterisation.

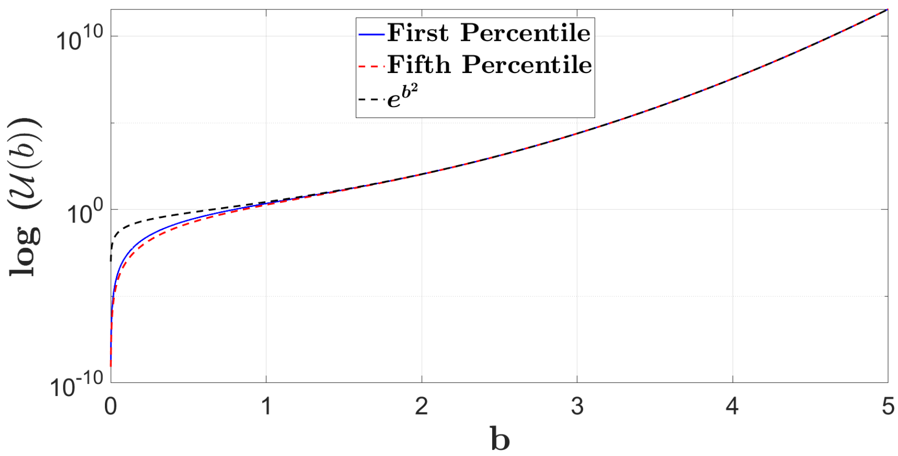


**Figure S2.** Representation of $\log(\mathcal{U}(b))$ to establish that $\mathcal{F}'(b) > 0$. Blue coloured line and dashed red line represent the derivation based on first percentile and fifth percentile of flow respectively. The dashed black line signifies $e^{b^2}$.

### 2.5 Condition for existence

$\mathcal{F}(b) = g(b)/h(b)$ grows monotonically with $b$ when $b > 0$, and goes to $+\infty$ as $b \to +\infty$. Therefore, a solution exists if:

$$\frac{\sigma}{\mu - q_{low}} \lim_{b \to 0^+} h(b) > \lim_{b \to 0} g(b) \tag{20}$$

and these limits are defined because both $g$ and $h$ are continuously differentiable over $]0, +\infty)$. The respective first-order Taylor expansions of $g$ and $h$ at $0^+$ yield:

$$\begin{cases} g(b) = b + o(b) \\ h(b) = -\ln(\varepsilon)\, b + o(b) \end{cases} \tag{21}$$

Since Taylor expansions are unique, the results from equation (21) into equation (20) yields the existence condition:

$$\frac{\sigma}{\mu - q_{low}} > \frac{-1}{\ln(\varepsilon)} \tag{22}$$

where $\varepsilon < 1$ so $\ln(\varepsilon) < 0$ and $-1/\ln(\varepsilon) \approx 0.43$ if $q_{low}$ is the first percentile; $0.61$ if $q_{low}$ is the fifth percentile. Note this condition is sufficient: if it is met, one can find the unique $b$ with equation (10), then $a$ and $c$ with equation (9).

## 3 Median, coefficient of variation and low flow quantile

In this section we assume we know the streamflow median $m$, coefficient of variation $CV = \sigma/\mu$, and low flow percentile $q_{low}$. We assume flow is not constant for large-stretches of the FDC (true for natural flows in perennial rivers) so we have $m > q_{low}$ and $CV > 0$. We will prove the $(a, b, c)$ triplet of Kosugi parameter is unique, and exists given a condition on $(\mu, \sigma, q_{low})$ that is often met in practice.

### 3.1 Relating parameter triplets

The median $m$ corresponds to $q(0.5)$ in equation (1). Since for $u = 0.5$ we have $z(u) = 0$, we have:

$$M = a \tag{23}$$

$CV$ is the ratio of standard deviation and mean. These two quantities are given by equations (6) and (4), respectively, so:

$$CV = \frac{(a-c)\sqrt{f(2b) - f(b)^2}}{c + (a-c)f(b)} \tag{24}$$

Finally, $q_{low}$ still verifies equation (7).

### 3.2 Solution strategy

Finding $a$ is immediate thanks to equation (23), and combined with equation (7), this directly leads to the following expression for $c$:

$$c = \frac{q_{low} - m\varepsilon^b}{1 - \varepsilon^b} \tag{25}$$

which means that $c$ can be easily and uniquely computed once $b$ is known. To find $b$, we use equation (24) and replace $f(b)$ with $e^{b^2/2}$ thanks to equation (14). This leads to:

$$CV = \frac{\sqrt{e^{b^2} - 1}}{e^{-b^2/2}\frac{c}{a-c} + 1} \tag{26}$$

Let us introduce $R$ as the ratio of $L$ by $M$:

$$R = \frac{L}{M} \tag{27}$$

Clearly, we have $0 < R < 1$. Equations (23) and (25) then become:

$$\frac{c}{a-c} = \frac{R - \varepsilon^b}{1 - R} \tag{28}$$

And finally:

$$CV = (1-R)\frac{\sqrt{e^{b^2} - 1}}{1 - R + (R - \varepsilon^b)e^{-b^2/2}} = \mathcal{G}(b) \tag{29}$$

where $\mathcal{G}$ only depends on $b$ because $R$ is a known parameter. As was the case in Section 2, we need to establish that there is (at most) a single $b > 0$ for a given value of $V$, and find the condition for existence. Then we can then deduce $c$. Yet, before establishing unicity and condition for existence, it is important to clarify on which range for $b > 0$ we can say that $\mathcal{G}(b)$ is defined.

### 3.3 Range of $b$ for which the equation for $CV$ is defined

Similar to equation (15), we can write:

$$\mathcal{G}(b) = (1-R)\frac{g(b)}{k(b)} \tag{30}$$

with $g(b)$ (and $g'(b)$) defined as in equations (16) and (17), and $k(b)$ defined as:

$$k(b) = 1 - R + (R - \varepsilon^b)e^{-b^2/2} \tag{31}$$

$\mathcal{G}(b)$ is defined in the range for $b > 0$ in which $k(b) \neq 0$. Since $g(b) > 0$, it corresponds to a coefficient of variation in the range in which $k(b) > 0$. We have $k(0) = 0$ and $\lim_{b\to\infty} k(b) = 1 - R > 0$, and need to understand variations to know what happens in between. $k(b)$ is derivated as follows:

$$k'(b) = \left[-bR + (b - \ln(\varepsilon))\varepsilon^b\right] e^{-b^2/2} \tag{32}$$

so that $k'(b)$ has the sign of $v(b) = \left[-bR + (b - \ln(\varepsilon))\varepsilon^b\right]$. In turn we have:

$$v'(b) = -R + \left[1 + (b - \ln(\varepsilon))\ln(\varepsilon)\right]\varepsilon^b \tag{33}$$

Since $\ln(\varepsilon) < -1$, for any positive value of $b$, $[1 + (b - \ln(\varepsilon))\ln(\varepsilon)] < 0$. This means that $v$ is monotonously decreasing for $b \geq 0$. We have $v(0) = -\ln(\varepsilon) > 1$, and $\lim_{b\to\infty} k(b) = -\infty$ because $-bR$ is the dominant term. Therefore, there is a $b_{lim}$ such that $v(b_{lim}) = 0$. For $b < b_{lim}$, $k(b)$ grows strictly and monotonously to a global maximum $k(b_{lim})$, then it degrows for $b > b_{lim}$ towards its limit value $1 - R > 0$. This means that $k(b_{lim}) > 0$, and $k(b) > 0$ for $b > 0$.

### 3.4 Unicity of the parameterisation

For $b > b_{lim}$, We know that the numerator $g(b)$ always grows with $b > 0$, and for $b > b_{lim}$, the numerator decreases strictly, so $\mathcal{G}(b)$ is strictly and monotonously growing. To demonstrate this is the case for any $b > 0$ we need to prove that $\mathcal{G}'(b)$ never reaches 0. This will complete the proof that the Kosugi parameterisation is unique if it exists. The following equivalence is true:

$$\mathcal{G}'(b) = 0 \Longleftrightarrow gg'k - g^2k' = 0 \tag{34}$$

This is equivalent to this linear equation in $R$:

$$R = \varepsilon^b + \frac{\ln(\varepsilon)\,\varepsilon^b\,(e^{b^2} - 1) + b\,e^{3b^2/2}\,(1 - \varepsilon^b)}{b\,(1 - 2\,e^{b^2} + e^{3b^2/2})} \tag{35}$$

The last expression is plotted in Figure S3 for both $L =$ first percentile (blue line) and fifth percentile (dashed red line). Clearly, stationarity requires $R < 0$ or $R > \lim_{b\to\infty} R = 1^+$. Both of these are impossible, because $0 < R < 1$ by definition of $R$ as the ratio of $q_{low}$ by $m$. Therefore $\mathcal{G}(b)$ is monotonic and strictly growing with $b$. This means that there is at most one $b$ for a given $CV$.

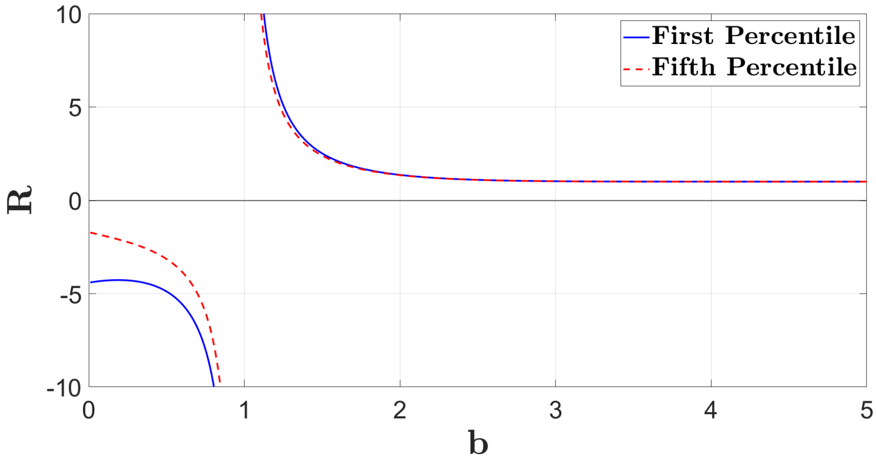


**Figure S3.** Plot of the stationary point locus in $b - R$ space on which $\mathcal{G}'(b) = 0$, as given by Equation 35. Blue coloured line and dashed red line represent the derivation based on first percentile and fifth percentile of flow respectively

### 3.5 Condition for existence

$\mathcal{G}(b) = (1-R)\, g(b)/k(b)$ is monotonous and grows with $b$ when $b > 0$, and goes to $+\infty$ as $b \to +\infty$. Therefore, a solution exists if:

$$\frac{CV}{1-R} \lim_{b\to 0^+} k(b) > \lim_{b\to 0} g(b) \tag{36}$$

and these limits are defined because both $g$ and $k$ are continuously differentiable over $]0, +\infty)$. The respective first-order Taylor expansions of $g$ at $0^+$ is given in equation (21) and for $k$ we have:

$$k(b) = -\ln(\varepsilon)\, b + o(b) \tag{37}$$

Since Taylor expansions are unique, the results from equation (37) into equation (36) yields the existence condition:

$$\frac{CV}{1-R} > \frac{-1}{\ln(\varepsilon)} \tag{38}$$

where $\varepsilon < 1$ so $\ln(\varepsilon) < 0$ and $-1/\ln(\varepsilon) \approx 0.43$ if $q_{low}$ is the first percentile; $0.61$ if $q_{low}$ is the fifth percentile.

Note this condition is sufficient: if it is fulfilled, one can find the unique $b$ with equation (29) then, 23 and 25 equations above directly lead to obtaining unique values of $a$ and $c$.

**Acknowledgments**

Prof. Solomon Brown and Dr Robert Milton are supported by the UK Engineering and Physical Sciences Research Council (EPSRC) through the 'Table Top Manufacturing of Tailored Silica for Personalised Medicine [SiPM]' project (Ref: EP/V051458/1).

*Code and data availability.* All data needed and the Python scripts to reproduce the analysis in this manuscript are are available in Yildiz et al. (2022) at https://doi.org/10.5281/zenodo.7423056

## References

Sadegh, M., Vrugt, J., Gupta, H. V., and Xu, C.: The soil water characteristic as new class of closed-form parametric expressions for the flow duration curve, Journal of Hydrology, 535, 438–456, 2016.

Yildiz, V., Rougé, C., Milton, R., and Brown, S.: Veysel-Yildiz/ClimatePerturbed_FDCs: v1.0.0, https://doi.org/10.5281/zenodo.7423056, 2022.