Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

1 **Technical Note: Improved mathematical representation of**

2 **concentration-discharge relationships**

3

4 José Manuel Tunqui Neira[1, 2], Vazken Andréassian[1], Gaëlle Tallec[1] & Jean-Marie Mouchel [2]

5 [(1)] Irstea, HYCAR Research Unit, Antony, France

6 [(2)] Sorbonne University, UMR Metis 7619, Paris, France

7

8 **Abstract**

9 This Technical Note deals with the mathematical representation of concentration-discharge

10 relationships and with the identification of its parameters. We propose a two-sided power

11 transformation alternative to the classical log-log transformation, and a multicriterion identification

12 procedure allowing determining parameters that are efficient, both from the concentration and the

13 load points of view.

14 **Keywords**

15 Concentration-discharge relationships; log-log transformation; power transformation; multi-

16 objective calibration

17

18 **1.   Introduction**

19 The relationship between ion concentrations and river discharge is an age-old topic in hydrology (see

20 among others Bazerbachi and Probst, 1986;Durum, 1953;Foster, 1978;Gibbs, 1970;Gregory and

21 Walling, 1973;Hem, 1948;Hendrickson and Krieger, 1960;Johnson et al., 1969;Meybeck, 1976).

22 Several recent articles (e.g. Bieroza et al., 2018;Chanat et al., 2002;Godsey et al., 2009;Meybeck and

23 Moatar, 2012;Moatar et al., 2017;Rose et al., 2018;Kirchner, 2019) give an exhaustive view of the

24 current work on this relationship. Since at least seventy years, a one-sided power relationship has

25 been used to represent and model this relationship (Eq. (1)).

26

$$C = aQ^b \qquad \qquad \textbf{Eq. (1)}$$

$$\ln(C) = \ln(a) + b.\ln(Q) \qquad \qquad \textbf{Eq. (2)}$$

27

28 From a graphical point of view, the one-sided power relation presented in Eq. (1) is equivalent to

29 plotting concentration and discharge in a log-log space, where parameters *a* and *b* can be identified

30 very simply, either graphically or numerically (under the classical assumptions of linear regression).

31 Obviously, not all ion species present a clear relationship between concentration and discharge; but

32 a few exhibit a very clear one as shown by the Figure 1 which gives an example from Upper Hafren

33 catchment data (Neal et al., 2013a;Neal et al., 2013b). In the rest of this note, we will focus explicitly

34 on the ions which exhibit a clear dependency between discharge and concentration.
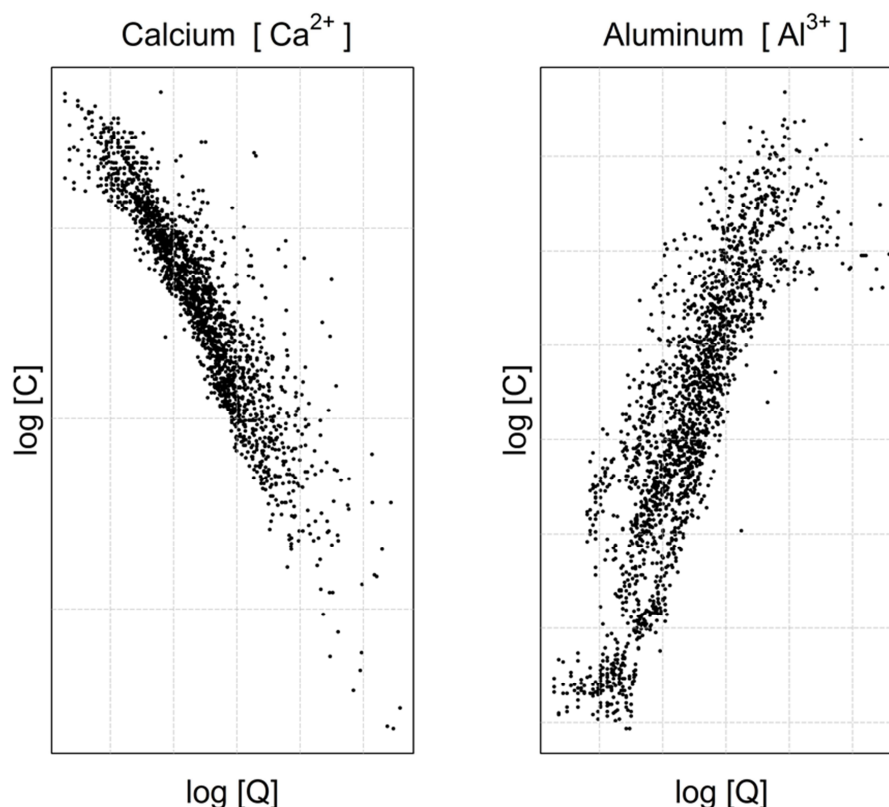
35

36 **Figure 1. Illustration of the C-Q behavior of two ions (Calcium and Aluminum) showing a clear log-**

37 **log relationship. Data from Upper Hafren catchment, Wales, United Kingdom (Neal et al.,**

38 **2013a;Neal et al., 2013b)**

39

40 ## 2. About the "excess" of the log-log transformation

41 For many years, since the size of the C-Q datasets was limited by the cost of chemical analyses, it was

42 difficult to analyze in much detail the precise shape of the C-Q relationship. In many cases, the log-

43   log transformation appeared visually adequate (and conceptually simple), which explains its lasting

44   popularity. With the advent of high-frequency measuring devices in recent years, the size of the

45   datasets has exploded, and all the extremes of the relationship can now be included in the analysis.

46   Figure 2 shows an example from such a high-frequency dataset, collected from the Oracle-Orgeval

47   observatory (Tallec et al., 2015;Floury et al., 2017). The 17500 data points represent half-hourly

48   measurements of ions collected over a two-year period, during which the catchment was exposed to

49   a variety of high- and low-flow events, which provides an unprecedented opportunity for exploring

50   the shape of the C-Q relationship.

51
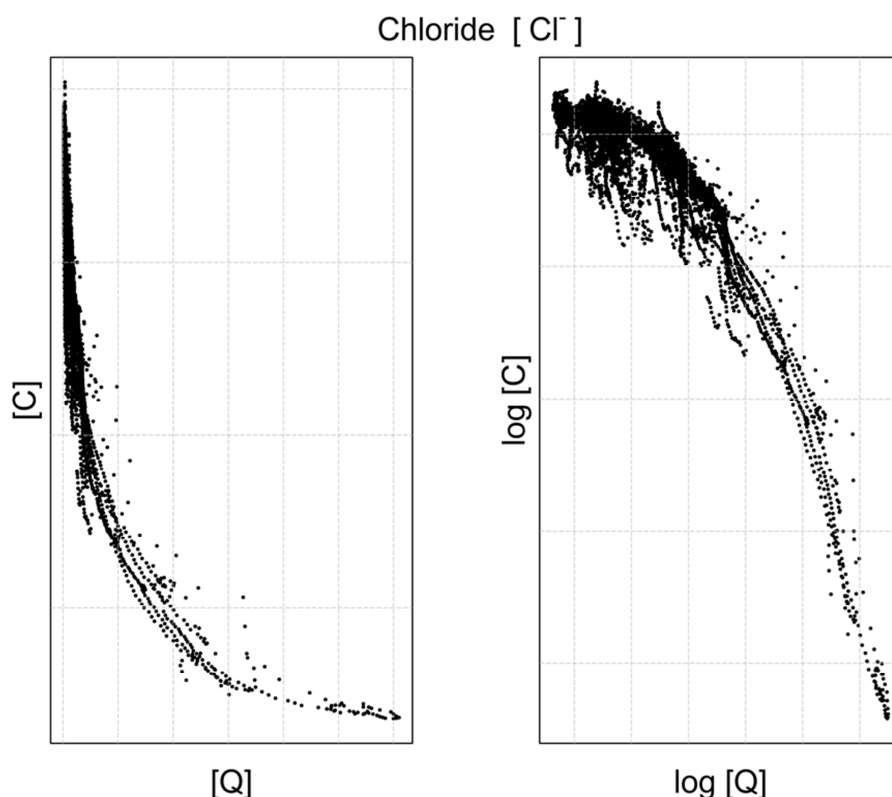
52   **Figure 2. Concentration-discharge relationship observed on the Oracle-Orgeval observatory**

53   **(measurements from the "RiverLab") for Chloride ions, with (right) and without (left) logarithm**

54   **transformation.**

55

56   Figure 2 illustrates what we call here the "excess" of the log-log transformation: the C-Q relationship

57   evolves from a clearly concave shape on the left to a slightly convex shape on the right. We have

58   gone beyond the straight shape that we aimed at. Note that this is not always the case, and the log-

59   log transformation can be well adapted in other situations (see Fig. 1 for the Upper Hafren data set

60   or further examples in the paper by Moatar et al., 2017). This slightly different shape may be due to

61   the high frequency of the time series (Moatar and Meybeck, 2007) or to catchment dynamics

62   (Kirchner, 2009), but in any case, it requires a different mathematical treatment than a logarithm

63   transformation.

3.   Box-Cox transformations as a continuous alternative to the log-log transformations

66   As a progressive alternative to the log-log transformation, we propose to use a two-sided power

67   transformation as shown in Eq. (3 (Box and Cox, 1964;Howarth and Earle, 1979).

$$C^{\frac{1}{n}} = a + bQ^{\frac{1}{n}}$$ **Eq. (3)**

68

69   The advantage of this transformation was already underlined by Box and Cox (1964): when n takes

70   high values, Eq. (3) converges towards the logarithm transformation (Eq. (2)), thus offering a

71   progressive solution. The reason is simple:

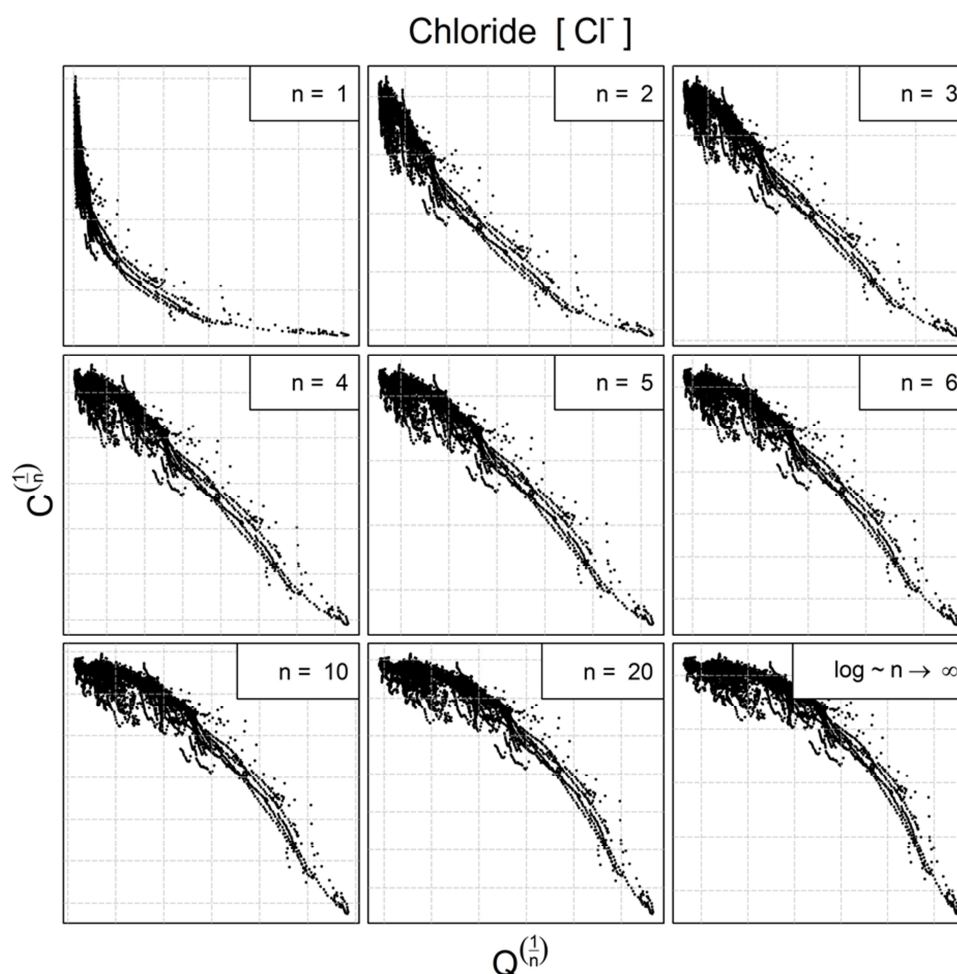$C^{\frac{1}{n}} = e^{\frac{1}{n} lnC} \approx 1 + \frac{1}{n} lnC$ when n is large.

72   Thus for large values of n, Eq. (3 can be written:

$$1 + \frac{1}{n} lnC \approx a + b + \frac{b}{n} lnQ$$

73   That is equivalent to

$\ln C \approx A + b . \ln Q$ (with $A = n(a + b - 1)$)

74   The progressive behavior and the convergence towards the log-log transformation are clearly

75   apparent in Figure 3.

76

Hydrology and
Earth System
Sciences

Discussions

Open Access

EGU

Figure 3. Evolution of the shape of the concentration-discharge scatterplot with an increasing value of *n*. Chloride ions concentrations measured on the Oracle-Orgeval observatory ("RiverLab").

4. Choosing an appropriate transformation for different ion species

To our knowledge, there is no physical or mathematical reason why all ionic species should have a C-Q relationship of the same shape. In Figure 4, we show the behavior of 3 ions and EC (Electrical conductivity) from the same catchment and the same dataset (all four from the Oracle-Orgeval observatory). The optimal shape could be chosen numerically (see Table 1), but we first followed the advice of Box et al. (2016, p. 331) and did it visually. Figure 4 shows the most adapted power transformation selected for each ion and EC.

88
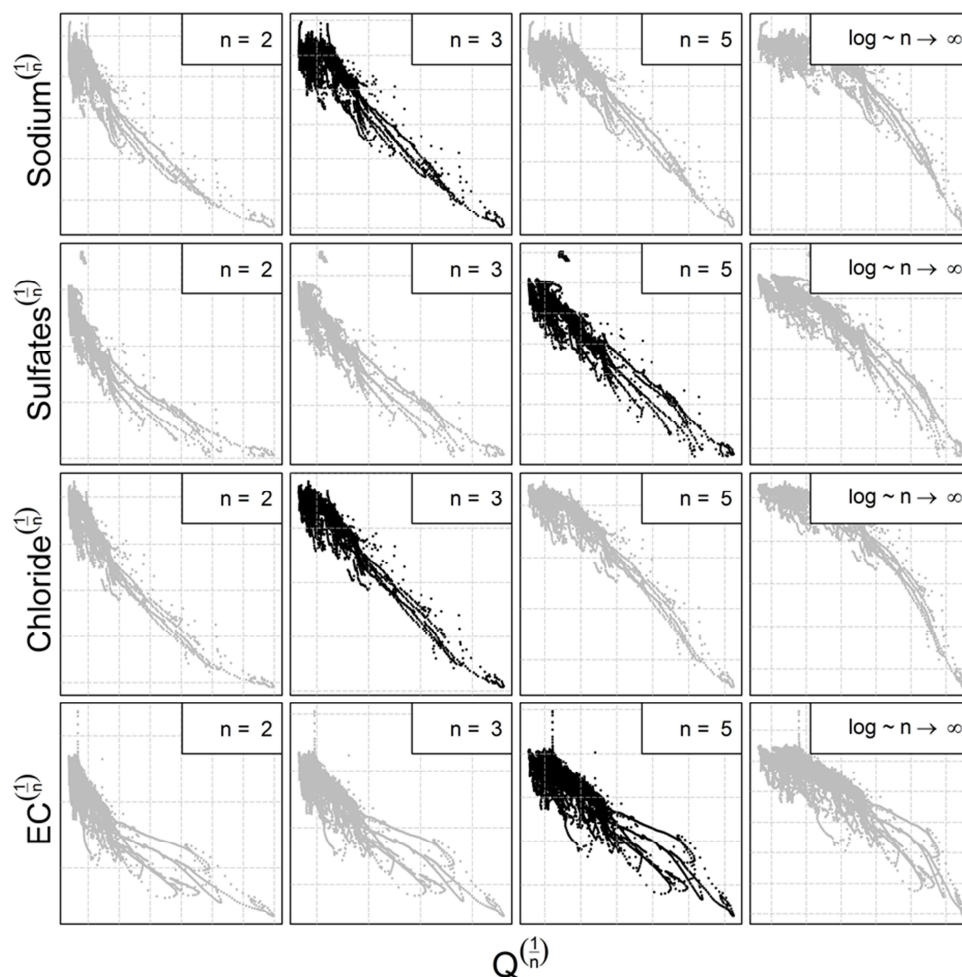
89

90 **Figure 4: C-Q behavior of three different chemical species and the conductivity with different**

91 **transformations (*n* =2, 3, 5 and log). The optimal value of the power transformation (black dots)**

92 **has been chosen visually.**

93

94

95

96

97 **Table 1. Coefficient of determination (R^2) calculated for n =1 (no transformation), n = optimal value**

98 **for two-sided power transformation (Figure 4) and n $\rightarrow \infty$ (log-log transformation) for each ion and**

99 **for EC**

ion	n	R^2
Sodium	$n = 1$ (no transformation)	0.53
	$n = 3$ (optimal)	0.73
	$n \rightarrow \infty$ (log-log)	0.53
Sulfate	$n = 1$ (no transformation)	0.32
	$n = 5$ (optimal)	0.81
	$n \rightarrow \infty$ (log-log)	0.77
Chloride	$n = 1$ (no transformation)	0.52
	$n = 3$ (optimal)	0.88
	$n \rightarrow \infty$ (log-log)	0.69
EC	$n = 1$ (no transformation)	0.38
	$n = 5$ (optimal)	0.79
	$n \rightarrow \infty$ (log-log)	0.74

100

101 Although we indicated above that the value of n could be chosen visually, we have also calculated

102 the coefficient of determination (R^2 see Table 1) to confirm our choice numerically. For each ion and

103 EC, the n considered optimal has the highest R^2 value.

5. Multi-objective identification of the parameters of the C-Q relationship

105 Once the most appropriate value of the power transformation has been determined, the numerical

106 identification of the optimal values of parameters a and b (see. Eq. (3)) should be easy. However, the

107 extremely large number of values in the high-frequency dataset can prevent a robust identification

108 over the full range of discharges, because the largest discharge values are in small numbers (in our

109 dataset only 1% of discharges are in the range [2.6 m^3s^{-1}, 12.2 m^3s^{-1}], and they correspond to the

110 lowest concentrations (see Figure 2).

111 To avoid the difficulties linked with the overrepresentation of low-discharge / high-concentration

112 data points, we tested successfully a multi-objective criterion for identifying the optimal couple (a,b).

113 We used an optimizing simultaneously on the quality of reproduction of the concentrations and the

114 load (i.e. the discharge-weighted concentrations); otherwise, the large discharge-low concentration

115 data points do not have enough weight to influence the selection of the parameter set.

116 The numerical criterion used for the concentration and the load is a bounded version of the Nash and

117 Sutcliffe criterion (Mathevet et al., 2006). The Nash and Sutcliffe (1970) criterion (see Eq. (4) and Eq.

118 (6), in Table 2) is well-known and widely used in the field of hydrology. The rescaling proposed by

119 Mathevet et al. (2006) transforms NSE in NSEB, which varies between -1 and 1 (see Eq. (5) and Eq. (7)

120   in Table 2). The advantage of this rescaled version is to avoid the occurrence of large negative values

121   (the original NSE criterion varies in the range]-∞, 1]).

122   Last, we also used a combined criterion for both concentration and load, by averaging $NSEB_{conc}$ and

123     $NSEB_{load}$ (see Eq. (8) in Table 2).

124

125   **Table 2. Numerical criteria used for optimization (C_{obs} – observed concentration, C_{cal} – computed**

126   **concentration, Q – observed discharge)**

|	
$$NSE_{conc} = 1 - \frac{\sum_t (C_{obs}^t - C_{cal}^t)^2}{\sum_t (C_{obs}^t - \overline{C_{obs}})^2}$$	Eq. (4)
$$NSEB_{conc} = \frac{NSE_{conc}}{2 - NSE_{conc}}$$	Eq. (5)
$$NSE_{load} = 1 - \frac{\sum_t (Q^t C_{obs}^t - Q^t C_{cal}^t)^2}{\sum_t (Q^t C_{obs}^t - \overline{QC_{obs}})^2}$$	Eq. (6)
$$NSEB_{load} = \frac{NSE_{load}}{2 - NSE_{load}}$$	Eq. (7)
$$NSEB_{comb} = \frac{1}{2}(NSEB_{conc} + NSEB_{load})$$	Eq. (8)

127

128   For the three ions and EC of the Oracle-Orgeval observatory, Figure 5 plots the performances of 6725

129   random pairs of parameters a and b (see Eq. (3)). The Pareto plot allows us to visualize the best

130   compromise between the criterion focusing on concentration ($NSEB_{conc}$) and the criterion focusing on

131   load ($NSEB_{load}$). The black point in Figure 5 corresponds to the optimum identified by the combined

132   criterion $NSEB_{comb}$.

133
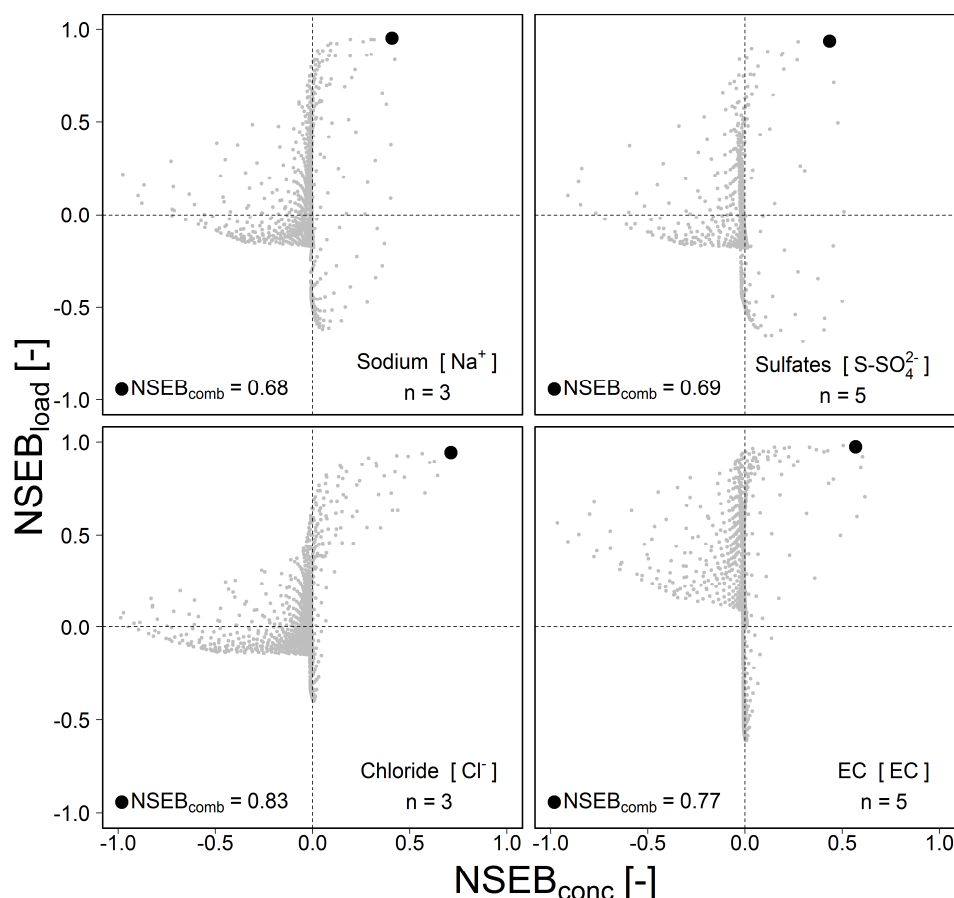
Figure 5. Multi-objective criterion *NSEB_{conc}* vs *NSEB_{load}* (Pareto plot) obtained from 3 chemical species (Sodium, Sulfates, Chloride) and Electrical conductivity (EC) from the Oracle-Orgeval observatory. The black points represent the performance of the pair having the best-combined criterion *NSEB_{comb}*. Note that like NSE, the optimal value of NSEB is 1, and that a value of 0 shows a very poor fit.

The interest of the Pareto plot is clearly seen in Figure 5. Not all ion species and EC have the same behavior: Chloride has a common optimum for the concentration and the load, while the others do not: they benefit clearly from a compromise that can be either identified visually on the Pareto plot or identified by the combined criterion *NSEB_{comb}*. The optimal *NSEBc_{omb}* criterion values for the three

145 ions and EC range from 0.68 to 0.83 (these values of NSEB correspond to "unbounded" NSE values

146 ranging from 0.77 to 0.91, which most hydrologists would consider very satisfying).

6. Results

148 The optimal values of *a* and *b* corresponding to the one shown on the Pareto plot (see Figure 5) and

149 the *n* value identified on Figure 4 are presented in Table 3, and Figure 6 illustrates the quality of the

150 fit over the entire calibration dataset (17500 points). Overall, the two-sided power transformation

151 model and the multicriterion identification procedure fit very well the concentrations. On can only

152 mention the difficulty of the model to reproduce the high extreme concentrations and conductivity

153 that would perhaps require a more elaborated model, which is out of the scope of this Technical

154 Note.

155

156 **Table 3. Summary of values *a, b* and *n* used to obtain the optimal $NSEB_{comb}$ criterion**

Ion	n	a	b	$NSEB_{comb}$
Sodium	3	-0.60	2.70	0.68
Sulfate	5	-0.55	2.20	0.69
Chloride	3	-1.00	3.70	0.83
EC	5	-0.70	4.20	0.77

157
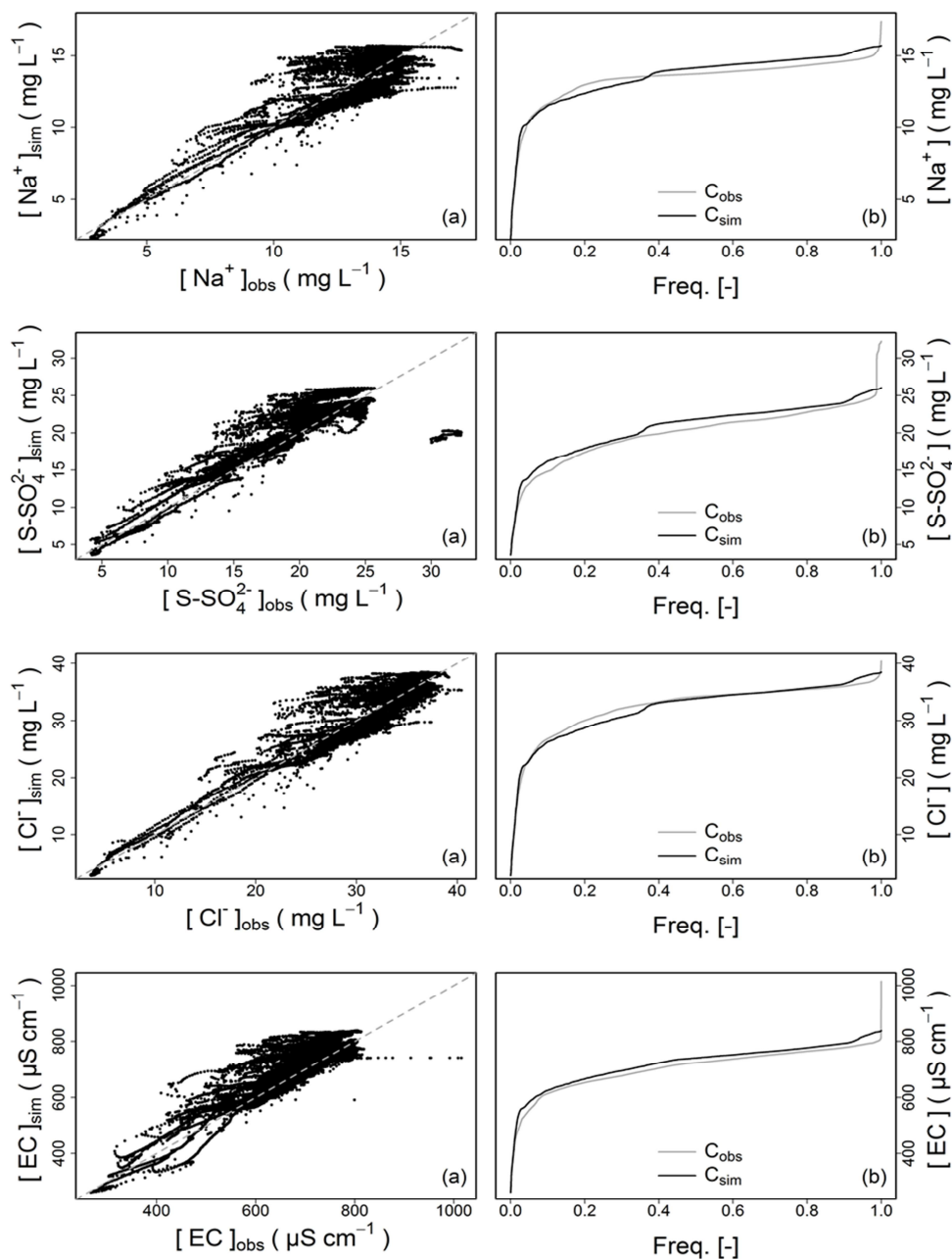
Figure 6. Comparison of observed concentrations with simulated concentrations by a two-sided power transformation model: (a) Scatterplot between the observed and simulated concentration; (b) Comparison of the cumulative frequency of the observed and simulated concentration.

161

Hydrology and
Earth System
Sciences
Discussions

7. Conclusion

In this technical note, we discussed the log-log transformation, widely used by hydrologists to represent concentration-discharge relationships, and showed that it is sometimes inadequate. The two-sided power transformation we proposed is a valid and progressive alternative. We also showed how the identification of the parameters of this relation can benefit of a multicriterion identification procedure, combining efficiency in concentration and load representation. The simulated concentrations for the 3 ions and the EC show a good performance.

Data availability. Data will be available in a dedicated database website after a contract accepted on behalf of all institutes.

Competing interests. The authors declare that they have no conflict of interest.

8. References

Bazerbachi, A., and Probst, J.-L.: Transports en solution et en suspension par la Garonne supérieure. *Solute and particulate transports by the upstream part of the Garonne river*, Sciences Géologiques, bulletins et mémoires, 79-98, 1986.

Bieroza, M. Z., Heathwaite, A. L., Bechmann, M., Kyllmar, K., and Jordan, P.: The concentration-discharge slope as a tool for water quality management, Science of The Total Environment, 630, 738-749, https://doi.org/10.1016/j.scitotenv.2018.02.256, 2018.

Box, G. E., and Cox, D. R.: An analysis of transformations, Journal of the Royal Statistical Society: Series B (Methodological), 26, 211-243, 1964.

Chanat, J. G., Rice, K. C., and Hornberger, G. M.: Consistency of patterns in concentration-discharge plots, Water Resources Research, 38, 22-21-22-10, 10.1029/2001wr000971, 2002.

Durum, W. H.: Relationship of the mineral constituents in solution to stream flow, Saline River near Russell, Kansas, Eos, Transactions American Geophysical Union, 34, 435-442, 10.1029/TR034i003p00435, 1953.

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., and Roubaty, J. L.: The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data, Hydrol. Earth Syst. Sci., 21, 6153-6165, 2017.

Foster, I. D. L.: Seasonal solute behaviour of stormflow in a small agricultural catchment, CATENA, 5, 151-163, https://doi.org/10.1016/0341-8162(78)90006-1, 1978.

Hydrology and
Earth System
Sciences
Discussions

197 Gibbs, R. J.: Mechanisms Controlling World Water Chemistry, Science, 170, 1088-1090,
198 10.1126/science.170.3962.1088, 1970.

199 Godsey, S. E., Kirchner, J. W., and Clow, D. W.: Concentration-discharge relationships reflect
200 chemostatic characteristics of US catchments, Hydrological Processes, 23, 1844-1864,
201 10.1002/hyp.7315, 2009.

202 Gregory, K. J., and Walling, D. E.: Quantitative evaluation of drainage basin process, in: Drainage
203 basin form and process, Fletcher & Son Ltd, United Kingdom, 184-225, 1973.

204 Hem, J. D.: Fluctuations in concentration of dissolved solids of some southwestern streams, Eos,
205 Transactions American Geophysical Union, 29, 80-84, 10.1029/TR029i001p00080, 1948.

206 Hendrickson, G., and Krieger, R.: Relationship of chemical quality of water to stream discharge in
207 Kentucky, Proceedings of 21st International Geological Congress, Norden, Denmark, 66-75, 1960.

208 Howarth, R., and Earle, S.: Application of a generalized power transformation to geochemical data,
209 Journal of the International Association for Mathematical Geology, 11, 45-62, 1979.

210 Johnson, N. M., Likens, G. E., Bormann, F. H., Fisher, D. W., and Pierce, R. S.: A working model for
211 variation in stream water chemistry at Hubbard-Brook-experimental-forest, new-hampshire, Water
212 Resources Research, 5, 1353-&, 10.1029/WR005i006p01353, 1969.

213 Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff
214 modeling, and doing hydrology backward, Water Resources Research, 45, 10.1029/2008wr006912,
215 2009.

216 Kirchner, J. W.: Quantifying new water fractions and transit time distributions using ensemble
217 hydrograph separation: theory and benchmark tests, Hydrology & Earth System Sciences, 23, 2019.

218 Mathevet, T., Michel, C., Andreassian, V., and Perrin, C.: A bounded version of the Nash-Sutcliffe
219 criterion for better model assessment on large sets of basins, IAHS PUBLICATION, 307, 211, 2006.

220 Meybeck, M.: Total mineral dissolved transport by world major rivers Hydrological Sciences Bulletin,
221 21, 265-284, 10.1080/02626667609491631, 1976.

222 Meybeck, M., and Moatar, F.: Daily variability of river concentrations and fluxes: indicators based on
223 the segmentation of the rating curve, Hydrological Processes, 26, 1188-1207, 2012.

224 Moatar, F., and Meybeck, M.: Riverine fluxes of pollutants: Towards predictions of uncertainties by
225 flux duration indicators, Comptes Rendus Geoscience, 339, 367-382,
226 https://doi.org/10.1016/j.crte.2007.05.001, 2007.

227 Moatar, F., Abbott, B., Minaudo, C., Curie, F., and Pinay, G.: Elemental properties, hydrology, and
228 biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major
229 ions, Water Resources Research, 53, 1270-1287, 2017.

230 Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion
231 of principles, Journal of hydrology, 10, 282-290, 1970.

232 Neal, C., Kirchner, J., and Reynolds, B.: Plynlimon research catchment high-frequency hydrochemistry
233 data. NERC Environmental Information Data Centre, 2013a.

234   Neal, C., Reynolds, B., Kirchner, J. W., Rowland, P., Norris, D., Sleep, D., Lawlor, A., Woods, C.,
235   Thacker, S., Guyatt, H., Vincent, C., Lehto, K., Grant, S., Williams, J., Neal, M., Wickham, H., Harman,
236   S., and Armstrong, L.: High-frequency precipitation and stream water quality time series from
237   Plynlimon, Wales: an openly accessible data resource spanning the periodic table, Hydrological
238   Processes, 27, 2531-2539, 2013b.

239   Rose, L. A., Karwan, D. L., and Godsey, S. E.: Concentration–discharge relationships describe solute
240   and sediment mobilization, reaction, and transport at event and longer timescales, Hydrological
241   Processes, 32, 2829-2844, 10.1002/hyp.13235, 2018.

242   Tallec, G., Ansard, P., Guérin, A., Delaigue, O., and Blanchoin, A.: Observatoire Oracle [Data set],
243   edited by: Irstea, https://bdoh.irstea.fr/ORACLE/, 2015.

244

245