An update on linear regression and error correlation: Exploration of baseline climate change impacts on Arctic and North Atlantic fog

Richard E. Danielson[1,2] (rickedanielson@gmail.com), William A. Perrie[2], and Minghong Zhang[2]
[1]Danielson Associates Office, Inc., Halifax, Nova Scotia
[2]Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia

Introduction

The experience of fog in nature is as ethereal as it is challenging to capture in observations and numerical forecast models. Visibility is a measure of fog that is readily diagnosed from surface humidity, but seemingly not without bias. Gultepe and Milbrandt (2010) provide one parameterization that Danielson et al. (2020) use to explore 21st century regional trends in marine visibility. Following an approach that can be described as conventional, all linear adjustments developed in that study assume that the forecast model is further removed from reality than observations (i.e., in every way). By contrast, one may also question whether an analysis (specifically, parameterizations of visibility applied to an analysis that benefits from a forecast model and observations) can be considered in various ways better than, equivalent to, as well as worse than an observational proxy of reality. An emphasis on forecast model strengths that are *complementary* to the strengths of observations is an important proposition of Parker (2016). Oreskes et al. (1994) and Beven (2019) provide further motivation for the philosophical challenge of whether to accommodate an error of representation associated with the forecast model (or analysis) that is equivalent to an error of representation associated with the observations (cf. Daley 1991). It is important to acknowledge that representation error is a misnomer here, insofar as its inclusion provides a better representation of visibility.

An Updated Linear Regression

It is only by longstanding convention that historical marine visibility observations are recorded (i.e., as one of ten categories). Further standardization of these measures is not anticipated (cf. World Meteorological Organization 2017), but at least one may take them as a good indication of the presence or absence of fog. The same might be said of estimates of visibility derived from numerical forecast data. Although both estimates are exploratory, comparisons are also instructive. A nascent approach to comparing such measures (with no one dataset assumed to be uniformly better) is given by Danielson (2018). Our update focuses on a canonical (yet imperfect) measurement model, or linear regression framework (Danielson et al. 2020). If any dataset is a partial measure of truth with error, then a numerical translation from uncalibrated (U) to calibrated data (C) can be represented by an additive (α_U) and multiplicative (β_U) adjustment, where

$$\begin{aligned} C &= t + \epsilon + \epsilon_C \\ U &= \alpha_U + \beta_U t + \epsilon + \epsilon_U . \end{aligned} \tag{1}$$

We describe t as a *linear association* that is only partially shared by both datasets, and hence, linear calibration can be only partial as well. An interesting consequence of equation (1) is that the covariance between C and U (e.g., between in situ and gridded estimates of visibility) also involves *nonlinear association*. From a metrological point of view, equation (1) is a canonical expression of errors-in-variables linear regression (Fuller 2006; Dunn 2011), except that it allows Fuller's equation error (ϵ, the *nonlinear association* term) to be shared between two different datasets. The interpretive consequences of adding such a term are not yet fully understood, but with the benefit of well sampled data (Danielson 2018), relatively direct numerical solutions of (1) are available.

Conclusions

A linear calibration of forecast model output helps to reveal consistent decreasing trends in 21st century marine visibility (Danielson et al. 2020). Given the use of a conventional visibility parameterization, these trend estimates take in situ observations as a *perfect* reference. While this is consistent neither with equation (1) nor with the subjective nature of marine visibility observations, it provides a useful baseline and is easy to interpret. However, there seems to be a gap in our ability to interpret a seemingly simple linear calibration when this involves measurement error in both visibility datasets. Thus, we have begun to explore more than just linear calibration and to accommodate an interpretation of more than just measurement error among ϵ, ϵ_C, and ϵ_U. Idealized control experiments are also being explored.

Acknowledgements

This work has been supported by the Ocean Frontier Institute, the Belmont Forum, the Office of Energy Research and Development, and Fisheries and Oceans Canada.

References

Beven, K., 2019: Towards a methodology for testing models as hypotheses in the inexact sciences. *Proc. Roy. Soc. A*, **475**, 1–19, doi:10.1098/rspa.2018.0862.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, New York, New York, 457 pp.

Danielson, R. E., 2018: On retrieving parameters of a linear regression model that accommodates error correlation in well sampled data. Working Group on Numerical Experimentation Research Activities (Blue Book) accessed May 2020 at http://bluebook.meteoinfo.ru/uploads/2018/sections/BB_18_S10.pdf.

Danielson, R. E., M. Zhang, and W. A. Perrie, 2020: Possible impacts of climate change on fog in the Arctic and subpolar North Atlantic. *Adv. Statist. Clim. Meteor. Ocean.*, **475**, 1–19, doi:10.5194/ascmo–6–31–2020.

Dunn, G., 2011: Method Comparison Studies, *Int. Encyclopedia of Statistical Science*, M. Lovric, Ed., Springer, 815–816, doi:10.1007/978-3-642-04898-2_36.

Fuller, W. A., 2006: Errors in variables. *Encyclopedia of Statistical Sciences*, S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic and N. L. Johnson, Eds., doi:10.1002/0471667196.ess1036.pub2.

Gultepe, I., and J. A. Milbrandt, 2010: Probabilistic parameterizations of visibility using observations of rain precipitation rate, relative humidity, and visibility. *J. Appl. Meteor. Clim.*, **49**, 36–46, doi:10.1175/2009JAMC1927.1.

Oreskes, N., K. Shrader-Frechette, and K. Belitz, 1994: Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, **263**, 641–646, doi:10.1126/science.263.5147.641.

Parker, W. S., 2016: Reanalyses and observations: What's the difference? *Bull. Amer. Meteor. Soc.*, **97**, 1565–1572, doi:10.1175/BAMS–D–14–00226.1.

World Meteorological Organization, 2017: Guide to Meteorological Instruments and Methods of Observation, Chapter 9, Measurement of visibility, WMO-No. 8, Geneva.