Authors' reply: Station data and modelled climate data in Africa

S Ng (sophian@princeton.edu)¹, B J Cowling²

1. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, United States

2. School of Public Health, The University of Hong Kong, Hong Kong, China

Citation style for this article:

Ng S, Cowling BJ. Authors' reply: Station data and modelled climate data in Africa. *Euro Surveill.* 2014;19(41):pii=20931. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20931>

Article submitted on 10 October 2014 / published on 16 October 2014

We thank Dr Shaman for his valuable comments [1] on our article [2]. We agree that it is challenging to study environmental determinants of diseases in locations where data are scarce. There are certainly weaknesses in the University of East Anglia Climate Research Unit (CRU) TS 3.21 dataset that could have affected our conclusions. This dataset was based on large climate datasets gathered by the World Meteorological Organization and the United States National Oceanographic and Atmospheric Administration. To maximise the use of available climate data, particularly in regions with imperfect station coverage, historical data, distant station data and station data of related climatic variables were processed and interpolated to provide modelled climatic data at a global scale. Technical details are available in an article written by Harris et al. [3].

The distribution of the 0.5° grid cells containing valid station data for mean temperature within the correlation decay distance in 1996 is shown in panel A of the Figure. Correlation decay distance was defined as the distance at which zonally averaged climatic conditions are no longer significantly correlated at the 95% confidence interval [3]. All of the Ebola virus disease outbreaks we studied occurred within the grid cells where at least one station data point was available. As reported by Harris et al., station data for vapour pressure were not widely available and these were inferred from station data on mean temperature and diurnal temperature range [3]. We show the distribution of grid cells containing valid station data points for these two predictor variables within correlation decay distance in 1980 in panel B of the Figure and in 1996 (panel C). Station data for mean temperature and

FIGURE

Grid cells containing one or more station data points within correlation decay distance

A. Area shaded in blue represents 0.5° grid cells containing station data for mean temperature within correlation decay distance in 1996.

B. Area shaded in green represents 0.5° grid cells containing station data for mean temperature and diurnal temperature range within correlation decay distance in 1980.

C. Area shaded in green represents 0.5° grid cells containing station data for mean temperature and diurnal temperature range within correlation decay distance in 1996.

Red circles represent the Ebola virus disease outbreak areas.

Source: [3].

diurnal temperature were not available in some previous Ebola virus disease outbreak areas and were less readily available for more recent years. Therefore the modelled climate data for these locations were mainly influenced by historical norms. New et al. [4] showed vapour pressure historical norms were widely available for all outbreak areas included in our analyses.

In our analyses, we standardised the climatic variables locally (standard deviation from the average climatic condition at the same location) to represent climatic condition as a variable varying seasonally above and below the average condition within the same outbreak location. Non-systematic discrepancies between the locally standardised modelled and the actual seasonal variation in climatic conditions added to the total noise presence in the data (random error). This noise was reflected in the confidence intervals of our estimates. Discrepancies that led to systematic bias might have some degree of influence on our main conclusion. Here we provide two examples of systematic discrepancies and their potential effects on our main conclusion.

Scenario 1

If the modelled seasonal variation in climate was consistently lagged behind or consistently phased ahead of the actual variation at a high number of outbreak locations, our analyses were vulnerable to systemic bias (e.g. humidity always peaked earlier at the predictor stations compared with the outbreak locations where climate data are interpolated). This type of systematic asynchrony influenced the best-fitting lag period of the model (maximum of three months allowed) and our main results may or may not be affected depending on the length of lag time between the modelled and actual climate variation.

Scenario 2

If the modelled climate data consistently inflated or consistently deflated the amplitude of seasonal variation in climatic conditions at many outbreak locations, our analyses were vulnerable to systemic bias (e.g. the locally standardised modelled data always showed larger peaks and/or troughs compared with the actual time series). Consistent deflation would have led to overestimation of the magnitude of odds ratios of zoonotic introduction associated with the standard deviation from mean climate conditions. Conversely, consistent inflation would have led to underestimation.

Conflict of interest

BJC reports receiving research funding from MedImmune Inc and consults for Crucell NV. The authors report no other potential conflicts of interest.

References

1. Shaman J. Letter to the Editor: Caution needed when using gridded meteorological data products for analyses in Africa. *Euro Surveill.* 2014;19(41):pii=20930. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20930>
2. Ng S, Basta, NE, Cowling, BJ. Association between temperature, humidity and ebolavirus disease outbreaks in Africa, 1976 to 2014. *Euro Surveill.* 2014;19(35):pii=20892.
3. Harris I, Jones PD, Osborn TJ, Lister DH. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int J Climatol.* 2014;34(3):623–42.
4. New M, Hulme M, Jones P. Representing twentieth-century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J Climate.* 1999;12(3):829–56.