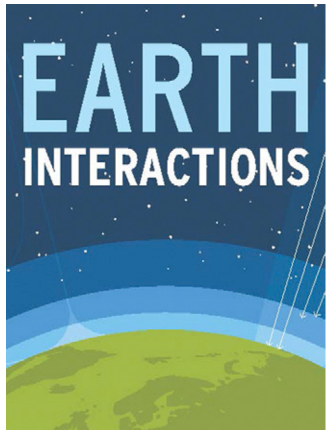

Comments on "Short-Term Precipitation and Temperature Trends along an Elevation Gradient in Northeastern Puerto Rico"

Angel R. Torres-Valcárcel[a]

COSUAM de Puerto Rico, and University of Puerto Rico, San Juan, Puerto Rico

Cesar Gonzalez-Avilés

COSUAM de Puerto Rico, San Juan, Puerto Rico

Received 19 July 2016; in final form 6 October 2016

ABSTRACT: The selection of statistical methods to evaluate data depends on study questions and characteristics of available data. In climate science, some methods are more popularly used than others; however, the use of applicable alternative methods does not invalidate study findings. Regardless of limitations, some methods like Pearson ordinary correlation are widely used in all sciences including climate and by scientists at government agencies like NOAA and the USGS. In addition, the use of the robust Student's *t* test is valid for near-Gaussian distributions with high sample numbers, since it is resistant to data distribution

Publisher's Note: This article was revised on 3 August 2017 to correct an editing error in the fourth paragraph.

[a] Corresponding author: Angel R. Torres-Valcárcel, atorresv09@gmail.com

DOI: 10.1175/EI-D-16-0018.1

inconsistencies. We wish to put in context the citation about our article and clarify the methods and justification for using them and to educate readers about the use of some conventional statistical tools and tests.

KEYWORDS: Statistical techniques; Statistics; Variational analysis; Climate variability; Interannual variability; Seasonal variability

1. Discussion

Our article Torres-Valcárcel et al. (2015) was cited in Van Beusekom et al. (2015, p. 18) as making improper use of statistical tools and tests. See below:

> A larger p value cutoff of $\alpha = 0.10$ was chosen for this entire study over the more traditional 0.01 or 0.05 because many of the climate trend studies covering PR. . . have used a t test and fit a least squares trend line, which exaggerates the significance of the trend, reporting a reduced p value, for non-Gaussian datasets.

The above citation is flawed, inaccurate, and misleading. The main focus of Torres-Valcárcel et al. (2015) was evaluating urban versus nonurban average temperature values, not about inferring about temperature trends. Hypothesis tests were done to detect and assess the magnitude of temperature signals for the entire period of study 1900 to 2007 from urban areas based on physical extension of intensively constructed area compared to rural areas (not intensively constructed areas). No hypothesis test regarding trends was done since this was not the main question or argument nor the main finding of our work, which would have required compliance with distribution assumptions. Consequently, inferential analysis of trends through a hypothesis test was not done, and so the use of Pearson ordinary correlation, also known as least squares linear correlation R^2, was for descriptive analysis of temperature trends based on their magnitude; this was explicitly noted and explained in Torres-Valcárcel et al. (2015, section 3.3.2). In other words, the use of linear correlation in our work was to compare actual R^2 values at different locations without making any inferences, just stating higher than or lower than, although no statistical tests were performed, as was transparently explained in the temperature analysis results and discussion section (Torres-Valcárcel et al. 2015, section 3.3.2).

Regardless of the known limitations of R^2, it is widely used in atmospheric sciences (Muhlbauer et al. 2009), including climatology such as in Nnamchi et al. (2015), Hartnett et al. (2014), Qi and Wang (2012), Fall et al. (2011), Anderson et al. (2010), and Angeles et al. (2007), to name several. In addition, the method is also applied by scientists at NOAA (Climate Prediction Center 2010) and the USGS, and so its use does not invalidate study findings in Torres-Valcárcel et al. (2015) or any other study. In any case, the use of R^2 may be uncertain when inferring future projections; in our study R^2 was only used as a descriptive way to compare the value of urban versus nonurban locations applying the same method to all. We did no test for any significance of temperature trends and so the above citation is inaccurate.

In our study, we made two separate analyses: one using the average value of maximum, minimum, and average temperature from surface stations, which were also analyzed independently from one another. The other analysis was done on inferred temperature data from interpolation methods for maximum, minimum, and average values for the period of study 1900 to 2007. In both, analysis data distribution parameters were measured and tested accordingly. Station data distribution met Gaussian distribution parameters and so analysis of variance (ANOVA) was done accordingly. Conversely, inferred interpolated data did not meet Gaussian parameters to proceed with ANOVA, but data characteristics did meet conditions for valid Student's t test application. Inferred data was nearly Gaussian, and the sample size was extraordinarily high to proceed with the Student's t test defined in Wilks (2006) as robust and therefore by definition resistant to data distribution inconsistencies. Our qualified and professionally certified statistician coauthor Cesar Gonzalez-Avilés, and professional engineer, appropriately recommended the valid use of Student's t test in substitution to ANOVA. Proceedings were clearly explained in the data and methods section (section 2) in Torres-Valcárcel et al. (2015), and a list of six references were provided (Student 1908; Hogg and Tanis 1996; Daniel 1999; Wigley 2006; Montgomery and Runger 2010; Laerd Statistics 2013) supporting the valid use of the Student's t test. Therefore, we found the above citation to be misleading.

In addition, the citation above is completely flawed, suggesting that the selected sigma α somehow affects the processing computation of the p value. Sigma α in hypothesis testing is just the criteria level for acceptance or rejection of the null hypothesis and is not used in p value computations in any way; this value does not change in response to the selected level of α, so it will be the same regardless of the selected sigma level. In our study, we selected the traditional conservative 0.05 cutoff value which represent the halfway point of the sigma α selection level interval between 0.01 to 0.1, reducing the risks of making the known type I error (reject a true null hypothesis) and type II error (fail to reject or accept a false null hypothesis). In our case, we got p values so low ($p = 0.00$) that they would have been found significant even at the most restrictive sigma selection level of $\alpha = 0.01$. These basic statistical concepts are widely discussed and explained in any basic statistical textbook that covers hypothesis testing methods. It is therefore unfortunate that Van Beusekom et al. (2015) did not make proper citation of Torres-Valcárcel et al. (2015). A list of web links are here provided with detailed explanation of hypothesis testing proceedings (accessed 2 October2016: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/hypothesis-tests/basics/type-i-and-type-ii-error/; http://statistics.about.com/od/Inferential-Statistics/a/Type-I-And-Type-II-Errors.htm; https://www.ma.utexas.edu/users/mks/statmistakes/errortypes.html; http://onlinestatbook.com/2/logic_of_hypothesis_testing/errors.html; http://www.stat.berkeley.edu/~hhuang/STAT141/Lecture-FDR.pdf).

Acknowledgments. I want to thank Earth Interactions's editor Dr. Rezaul Mahmood and everyone who works on the journal for taking the time and effort to proceed with this comment. Also, thanks to the American Meteorological Society for allowing us to clarify this issue and continue to contribute to science.

References

Anderson, B. T., J. Wang, G. Salvucci, S. Gopal, and S. Islam, 2010: Observed trends in summertime precipitation over the southwestern United States. *J. Climate*, **23**, 1937–1944, doi:10.1175/2009JCLI3317.1.

Angeles, M. E., J. E. Gonzalez, D. J. Erickson III, and J. L. Hernández, 2007: Predictions of future climate change in the Caribbean region using global general circulation models. *Int. J. Climatol.*, **27**, 555–569, doi:10.1002/joc.1416.

Climate Prediction Center, 2010: Global ENSO temperature and precipitation linear regressions. National Weather Service Climate Prediction Center, accessed 2 October 2016. [Available online at http://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/regressions/readme.shtml.]

Daniel, W. W., 1999: *Biostatistics: A Foundation for Analysis in the Health Sciences.* 7th ed. John Wiley and Sons, 907 pp.

Fall, S., A. Watts, J. Nielsen-Gammon, E. Jones, D. Niyogi, J. R. Christy, and R. A. Pielke Sr., 2011: Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends. *J. Geophys. Res.*, **116**, D14120, doi:10.1029/2010JD015146.

Hartnett, J. J., J. M. Collins, M. A. Baxter, and D. P. Chambers, 2014: Spatiotemporal snowfall trends in central New York. *J. Appl. Meteor. Climatol.*, **53**, 2685–2697, doi:10.1175/JAMC-D-14-0084.1.

Hogg, R. V., and E. Tanis, 1996: *Probability and Statistical Inference.* 5th ed. Prentice-Hall, 826 pp.

Laerd Statistics, 2013: Independent t-test using SPSS statistics. Accessed 10 July 2013. [Available online at https://statistics.laerd.com/spss-tutorials/independent-t-test-using-spss-statistics.php.]

Montgomery, D. C., and G. C. Runger, 2010: *Applied Statistics and Probability for Engineers.* John Wiley and Sons, 784 pp.

Muhlbauer, A., P. Spichtinger, and U. Lohmann, 2009: Application and comparison of robust linear regression methods for trend estimation. *J. Appl. Meteor. Climatol.*, **48**, 1961–1970, doi:10.1175/2009JAMC1851.1.

Nnamchi, H. C., J. Li, F. Kucharski, I. S. Kang, N. S. Keenlyside, P. Chang, and R. Farneti, 2015: Thermodynamic controls of the Atlantic Niño. *Nat. Commun.*, **6**, 8895, doi:10.1038/ncomms9895.

Qi, L., and Y. Wang, 2012: Changes in the observed trends in extreme temperatures over China around 1990. *J. Climate*, **25**, 5208–5222, doi:10.1175/JCLI-D-11-00437.1.

Student, 1908: The probable error of a mean. *Biometrika*, **6**, 1–25, doi:10.1093/biomet/6.1.1.

Torres-Valcárcel, Á. R., J. Harbor, A. L. Torres-Valcárcel, and C. J. González-Avilés, 2015: Historical differences in temperature between urban and non-urban areas in Puerto Rico. *Int. J. Climatol.*, **35**, 1648–1661, doi:10.1002/joc.4083.

Van Beusekom, A. E., G. González, and M. M. Rivera, 2015: Short-term precipitation and temperature trends along an elevation gradient in northeastern Puerto Rico. *Earth Interact.*, **19**, doi:10.1175/EI-D-14-0023.1.

Wigley, T. M., 2006: Statistical issues regarding trends. *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*, T. R. Karl et al., Eds., U.S. Climate Change Science Program, 129–139.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.