

**EDITORIAL**

<https://doi.org/10.5232/ricyde2019.057ed>

**FIFA World Cup and climate change: correlation is not causation**  
**Copa del Mundo FIFA y cambio climático: correlación no implica causalidad**

**Aurelio Tobías<sup>1</sup>, Martí Casals<sup>2,3</sup>, Javier Peña<sup>2,3</sup>, Cristian Tebé<sup>4,5</sup>**

1. Institute of Environmental Assessment and Water Research (IDAEA), Spanish Council for Scientific Research (CSIC), Barcelona, Spain.
2. Sport and Physical Activity Studies Centre (CEEAF), University of Vic-Central University of Catalonia (UVic-UCC), Barcelona, Spain.
3. Sport Performance Analysis Research Group, University of Vic-Central University of Catalonia (UVic-UCC), Barcelona, Spain.
4. Biostatistics Unit at Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain.
5. Department of Basic Medical Sciences, Universitat Rovira i Virgili, Reus, Tarragona, Spain.

Sports analytics is defined as the process that identifies and acquires knowledge and insight about players and teams' performances. To do so, analysts use a wide variety of data sources coming from matches and individual players' performances (O'Donoghue & Holmes 2014; Jayal, McRobert, Oatley & O'Donoghue, 2018). Nowadays, detailed data from different nature including technical skills, individual physiological performances, team formations, or injuries are analysed on a daily basis by the analytics departments belonging to sports clubs and professional franchises. Even private companies like STATS or OPTA generate important revenues offering their movement tracking values and advanced metrics to media and fans. In the emerging field of Sports Analytics, as in many others, analysts must be aware of spurious correlations. These can come up due to the size (not nature) of data, a common-causal variable or just due to chance. For this reason, we always must keep in mind the lessons of the statistician Stephen John Senn and his sharp quote on the matter: "Statistics is not just about merely warning that correlation is not causation. Sometimes correlation isn't even correlation". Thus, we will explain an example of how climate change can be affecting, or not, on the FIFA World Cup performance statistics.

**The impact of the FIFA World Cup**

Every four years, since 1930, the world focuses its attention on the FIFA World Cup. A football competition in which men's national teams from around the world compete for the trophy in a tournament during a period of about a month. As a mega-sport event, the FIFA World Cup not only has a potential impact of diverse nature in the host country but also on the winner's team country. Different studies have been conducted to analyse the economic

impact of the FIFA World Cup in the past. Nicolau & Sharma (2018) explored the impact of tourism in the winning FIFA World Cup country between the years 1998 and 2010. Ramdas & van Gaalen (2015) analysed the impact of the hosting announcement from the FIFA on the host country stock markets between the period 1994 and 2010. Allmers & Maennig (2009) assessed the short- and long-term economic impact of the FIFA world cup editions of 1998 and 2006 on the hosting country.

Aside from the economic effects, to our knowledge, a wide assessment of the environmental impact on such mega-sport events has not been yet conducted. In this framework, the Intergovernmental Panel on Climate Change (IPCC), the leading international body for the assessment of climate change, concluded in its fifth assessment report (Pachauri et al., 2014) stating that human activities over the past 50 years have significantly warmed the planet. However, high-performance players and teams have been regularly studied without considering the potential impact of the environmental factors on their technical and tactical performance.

**The impact of climate change on the FIFA World Cup historical performances**

We correlated the FIFA World Cup performance statistics for the number of penalty shoot-outs at the round of 16 and the total number of hat-tricks (Wikipedia 2018, Jul. 18) with the temperature anomaly (change in global surface temperature relative to 1951-1980 average temperature) and minimum arctic sea ice (measured in million square km based on satellite observations) as climate change indicators (NASA 2018, Jul. 18) from 1986 up to 2018, considering that climate reports in FIFA World Cups are only available since the late '80s.

Figure 1 shows how the number of penalty shoot-outs at the round of 16 is strongly positively correlated with the land temperature anomaly ( $r = 0.82$ ) and negatively with the minimum Arctic sea ice ( $r = -0.74$ ). Thus, an increase in land temperature and a consequent decrease of the minimum Arctic sea ice lead to an increase in the number of penalty shoot-outs at the round of 16. Conversely, the total number of hat-tricks is negatively correlated with the land temperature anomaly ( $r = -0.55$ ) and positively with the minimum Arctic sea ice ( $r = 0.44$ ).

Here, an increase of land temperature and a consequent decrease of the minimum Arctic sea ice lead to a decrease in the total number hat-tricks scored in the World Cup.

However, similar correlations have been observed between other World Cup performances and climatic indicators. For example, the total number of own goals is also positively correlated with the sea level ( $r = 0.75$ ), contributing to increasing evidence that climate change is correlated with the World Cup competitive performance.

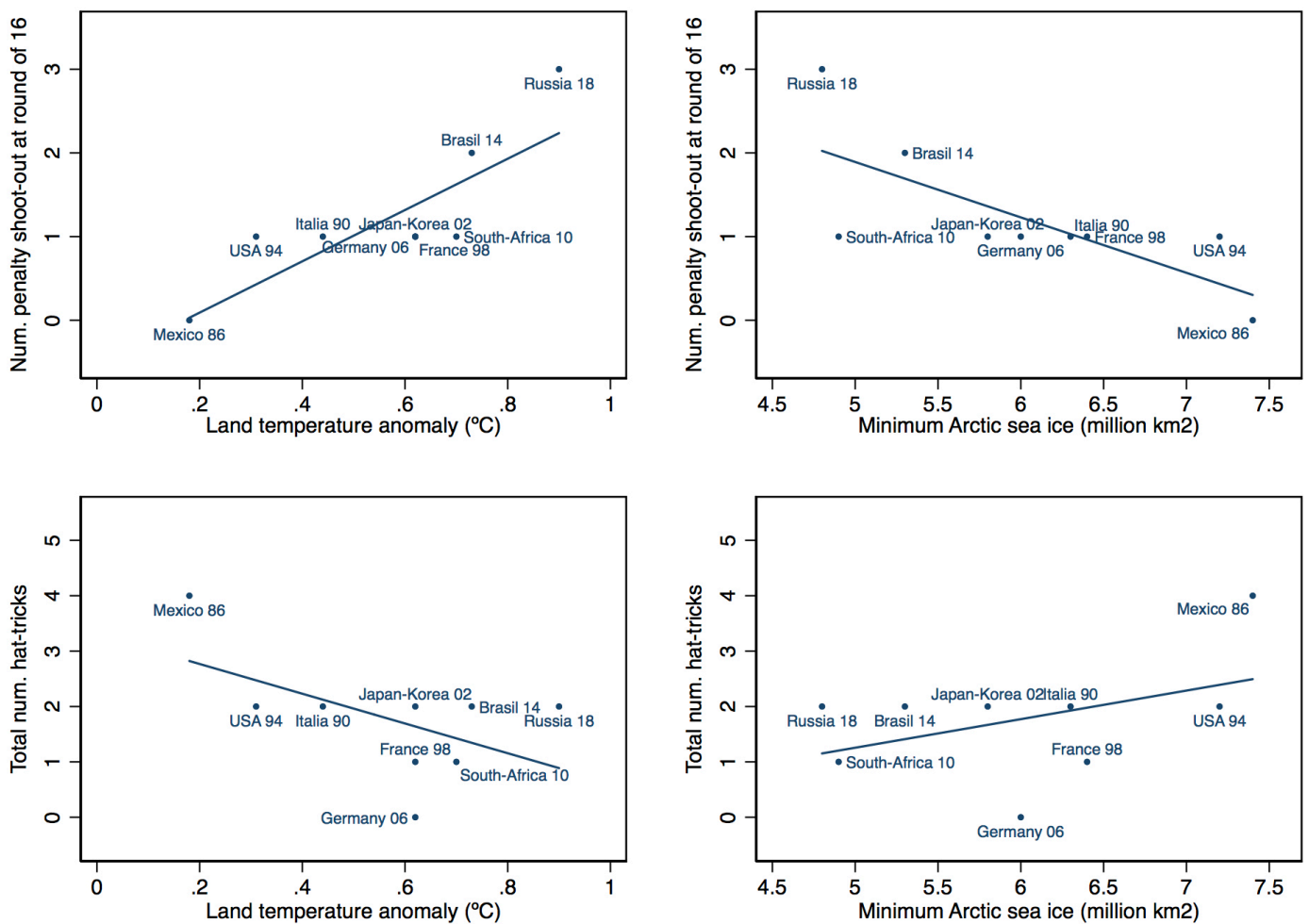


Figure 1. Scatter plots between selected FIFA World Cup performance statistics and climate change indicators.

These examples show the potential impact of climate change on the players' performance in a mega-sport event. Initially, climate change seems to worsen players' performance and consequently, that fact is reflected in the match-related statistics of the FIFA World Cup. Climate change can directly affect human health by varying exposure to non-optimal outdoor temperature. In the case of footballers, several studies have demonstrated impaired high-intensity running (Mohr, Nybo, Grantham & Racinais, 2012) or a reduced number of sprints (Nassis, Brito, Dvorak, Chalabi & Racinais, 2015), to set a few examples, when professional players suffer from heat stress. From the initial edition of the World Cup in 1930, an increased number of national teams have accessed the tournament, with more matches played, more stadiums built and more people than ever attending the championship. More teams are also expected for the 2022 edition. All this unstoppable growth implies not only more games and players participating, but also more visitors attending the forthcoming championships who could also be affected by the increase in temperature, with the consequent impact on the public health system of the organiser country. Therefore, it would be appropriate for the organizing committee to consider the adoption of appropriate preventive measures for the effects of high temperatures (Matzarakis & Fröhlich, 2015). A similar recommendation has already been issued by Kamaku Wada, Smith, Endo & Fukushima (2017) to prevent heat effects for the foreseeable hot climate in the Tokyo 2020 Summer Olympic Games.

### Conclusion

It is well known that performance indicators such as total shots, shots on target or ball possession have great discriminatory power in favour of successful teams in the World Cup (Castellano, Casamichana & Lago, 2012; Liu, Gomez, Lago-Peñas & Sampaio, 2015; Rumpfet, Silva, Hertzog, Farooq & Nassis, 2017). However, actual performance-related indicators often are in contradiction with non-related variables leading to spurious correlations and misleading interpretations. Correlation simply describes the strength of a linear relationship between two variables. The increase in the value of one variable, such as land temperature anomaly, may be followed by the increase in the value of a second one, such as the number of penalty shoot-outs at the round of 16. The simplicity of a correlation coefficient hides the considerable complexity in interpreting its causal meaning.

The most common error is to fall into an ecological fallacy when a conclusion about individuals is reached based on group-level data (Robinson 1950). In our case, we have calculated the correlation coefficient at the aggregate level (each World Cup edition since 1986) and then mistakenly used that value to reach a conclusion about the individual performance-level, but data at the individual level was unknown. Ecological correlations have been displayed in a big number of examples of non-sense quantitative studies (Vigen 2015), like per capita consumption of mozzarella cheese and civil engineering

doctorates awarded in the US ( $r=0.96$ ). To avoid misinterpretations, researchers must carefully check the causal framework of their research, using measures such as formalized counterfactual reasoning aided by directed acyclic graphs (Vandenbroucke, Broadbent & Pearce, 2016).

### Disclaimer

This editorial is not intended to disprove the value of sports science or environmental epidemiology research. It pretends to be only an example of how studies based on unsubstantiated theory, jointly with arbitrarily collected data and references, can generate and support coincidental statistical associations leading to apparent scientific endorsements. Associations and spurious correlations between phenomena do not mean they are causally related. This scientific commentary only tries to remind about the importance of research methodology in education and statistical thinking to maintain rigour in sports sciences and performance analysis. In the end, it is just a humorous case study for education in sports analytics.

### References

- Allmers, S., & Maennig, W. (2009). Economic impacts of the FIFA Soccer World Cups in France 1998, Germany 2006, and outlook for South Africa 2010. *Eastern Economic Journal*, 35(4), 500-519.  
<https://doi.org/10.1057/ej.2009.30>
- Castellano, J.; Casamichana, D., & Lago, C. (2012). The use of match statistics that discriminate between successful and unsuccessful soccer teams. *Journal of Human Kinetics*, 31(1), 137-147.  
<https://doi.org/10.2478/v10078-012-0015-7>
- Jayal, A., McRobert, A., Oatley, G., & O'Donoghue, P. (2018). *Sports Analytics: Analysis, Visualisation and Decision Making in Sports Performance*. Routledge.  
<https://doi.org/10.4324/9781315222783>
- Kakamu, T.; Wada, K.; Smith, D. R.; Endo, S., & Fukushima T. (2017). Preventing heat illness in the anticipated hot climate of the Tokyo 2020 Summer Olympic Games. *Environmental Health and Preventive Medicine* 22(68), 1-6.  
<https://doi.org/10.1186/s12199-017-0675-y>
- Matzarakis, A., & Fröhlich, D. (2015). Sport events and climate for visitors-the case of FIFA World Cup in Qatar 2022. *International Journal of Biometeorology*, 59(4), 481-586.  
<https://doi.org/10.1007/s00484-014-0886-5>
- Moh, r, M.; Nybo, L.; Grantham, J., & Racinais, S. (2012). Physiological Responses and Physical Performance during Football in the Heat. *PLoS ONE*, 7(6), e39202.  
<https://doi.org/10.1371/journal.pone.0039202>

- NASA (2018). *Global Climate Change, Vital Signs of the Planet*. [Internet] <http://climate.nasa.gov/> (last accessed, Jul. 18)
- Nassis, G. P.; Brito, J.; Dvorak, J.; Chalabi, H.; & Racinais, S. (2015). The association of environmental heat stress with performance: analysis of the 2014 FIFA World Cup Brazil. *British Journal of Sports Medicine*, 49(9), 609-613. <https://doi.org/10.1136/bjsports-2014-094449>
- Nicolau, J. L., & Sharma, A. (2018). A generalization of the FIFA World Cup effect. *Tourism Management*, 66(June), 315-317. <https://doi.org/10.1016/j.tourman.2017.12.014>
- Liu, H.; Gomez, M. Á.; Lago-Peñas, C., & Sampaio, J. (2015). Match statistics related to winning in the group stage of 2014 Brazil FIFA World Cup. *Journal of Sports Sciences*, 33(12), 1205-1213. <https://doi.org/10.1080/02640414.2015.1022578>
- O'Donoghue, P., & Holmes, L. (2014). *Data analysis in sport*. Routledge. <https://doi.org/10.4324/9781315816357>
- Pachuari R. K.; Allen, M. R.; Barros, V. R.; Broome, J.; Cramer, W.; Christ, R.; et al. (2014). *Climate change 2014: synthesis report. Contribution of working groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change, Geneva. Retrieved from: [https://www.ipcc.ch/site/assets/uploads/2018/05/SYR\\_AR5\\_FINAL\\_full\\_wcover.pdf](https://www.ipcc.ch/site/assets/uploads/2018/05/SYR_AR5_FINAL_full_wcover.pdf)
- Ramdas, B.; van Gaalen, R., & Bolton, J. (2015). The Announcement Impact of Hosting the FIFA World Cup on Host Country Stock Markets. *Procedia Economics Finance*, 30, 226-238. [https://doi.org/10.1016/S2212-5671\(15\)01290-3](https://doi.org/10.1016/S2212-5671(15)01290-3)
- Robinson, W. S. (1950). Ecological correlations and the behaviour of individuals. *American Sociological Review*, 15(3), 351-357. <https://doi.org/10.2307/2087176>
- Rumpf, M. C.; Silva, J. R.; Hertzog, M.; Farooq, A., & Nassis, G. (2017). Technical and physical analysis of the 2014 FIFA World Cup Brazil: winners vs. losers. *The Journal of sports medicine and physical fitness*, 57(10), 1338-1343.
- Vandenbroucke, J. P.; Broadbent, A.; Pearce, N. (2016). Causality and causal inference in epidemiology: the need for a pluralistic approach. *International Journal of Epidemiology*, 45(6), 1776-1786. <https://doi.org/10.2307/2087176>
- Vigen, T. (2015). *Spurious Correlations*. Hachette Books.
- Wikipedia (2018). *FIFA World Cup-related lists*. [Internet] (last accessed, Jul. 18). Retrieved from: [https://en.wikipedia.org/wiki/Category:FIFA\\_World\\_Cup-related\\_lists](https://en.wikipedia.org/wiki/Category:FIFA_World_Cup-related_lists)

Tobías, A.; Casals, M.; Peña, J., & Tebé, C. (2019). FIFA World Cup and climate change: correlation is not causation. *RICYDE. Revista Internacional de Ciencias del Deporte*, 57(15), 280-283. <https://doi.org/10.5232/ricyde/2019.057ed>

Correspondence/correspondencia: Martí Casals  
Sport Performance Analysis Research Group, University of Vic-Central University of Catalonia (UVic-UCC), Barcelona, Spain.  
Email: marticasals@gmail.com