Challenges and Opportunities of Data Science for Climate Research

Reto Knutti, David Bresch, Gertrude Hirsch Hadorn, Christoph Baumberger, Benedikt Knüsel, Marius Zumwald

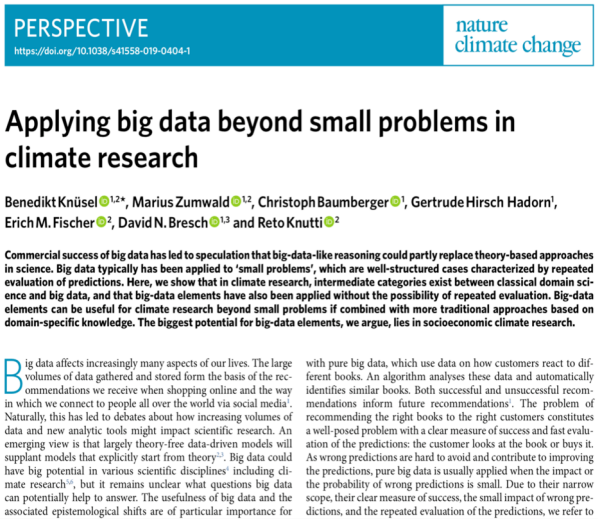
Conceptual work

Investigate epistemological d ifferences between process-based and data-driven models. Characterise uncertainties of inferences for decision-making.

Interdisciplinary Exchange

Applied case studies

Modelling with machine learning and non-conventional data sources such as crowd-sensed citizen weather stations.



PERSPECTIVE
https://doi.org/10.1038/s41558-019-0404-1
nature climate change

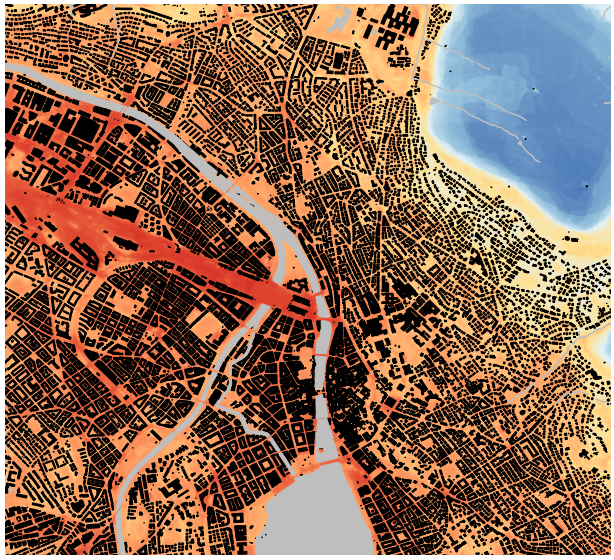
Applying big data beyond small problems in climate research

Benedikt Knüsel[1,2*], Marius Zumwald[1,2], Christoph Baumberger[1], Gertrude Hirsch Hadorn[1], Erich M. Fischer[2], David N. Bresch[1,3] and Reto Knutti[2]

Commercial success of big data has led to speculation that big-data-like reasoning could partly replace theory-based approaches in science. Big data typically has been applied to 'small problems', which are well-structured cases characterized by repeated evaluation of predictions. Here, we show that in climate research, intermediate categories exist between classical domain science and big data, and that big-data elements have also been applied without the possibility of repeated evaluation. Big-data elements can be useful for climate research beyond small problems if combined with more traditional approaches based on domain-specific knowledge. The biggest potential for big-data elements, we argue, lies in socioeconomic climate research.

Big data affects increasingly many aspects of our lives. The large volumes of data gathered and stored form the basis of the recommendations we receive when shopping online and the way in which we connect to people all over the world via social media[1]. Naturally, this has led to debates about how increasing volumes of data and new analytic tools might impact scientific research. An emerging view is that largely theory-free data-driven models will supplant models that explicitly start from theory[2,3]. Big data could have big potential in various scientific disciplines[4] including climate research[5,6], but it remains unclear what questions big data can potentially help to answer. The usefulness of big data and the associated epistemological shifts are of particular importance for with pure big data, which use data on how customers react to different books. An algorithm analyses these data and automatically identifies similar books. Both successful and unsuccessful recommendations inform future recommendations[7]. The problem of recommending the right books to the right customers constitutes a well-posed problem with a clear measure of success and fast evaluation of the predictions: the customer looks at the book or buys it. As wrong predictions are hard to avoid and contribute to improving the predictions, pure big data is usually applied when the impact or the probability of wrong predictions is small. Due to their narrow scope, their clear measure of success, the small impact of wrong predictions, and the repeated evaluation of the predictions, we refer to

Knüsel, B., Zumwald, M., Baumberger, C., Hirsch Hadorn, G., Fischer, E. M., Bresch, D. N., & Knutti, R. (2019). Applying big data beyond small problems in climate research. Nature Climate Change, 9(3), 196.



Zumwald et al., Mapping urban temperature using crowd-sensing data and machine learning, in preparation