Artificial intelligence reconstructs missing climate information

From missing measurements to higher resolution

Christopher Kadow[†a], Étienne Plésiat[a], Johannes Meuer[a], David Matthew Hall[b], Uwe Ulbrich[c], Hannes Thiemann[a], Thomas Ludwig[a]

a German Climate Computing Center (DKRZ), Hamburg, GERMANY, [†]kadow@dkrz.de
b NVIDIA, Santa Clara, CA, USA
c Institute of Meteorology, Freie Universität Berlin, Berlin, GERMANY

ABSTRACT

Historical temperature measurements are the basis of global climate datasets like HadCRUT4 and HadCRUT5. These datasets contain many missing values, particularly for periods before the mid-twentieth century, although recent years are also incomplete. Here we demonstrate that artificial intelligence can skillfully fill these observational gaps when combined with numerical climate model data. We show that recently developed image inpainting techniques perform accurate monthly reconstructions via transfer learning using either 20CR (Twentieth-Century Reanalysis) or the CMIP5 (Coupled Model Intercomparison Project Phase 5) experiments. In addition, high resolution in weather and climate was always a common and ongoing goal of the community. In this regard, machine learning techniques accompanied numerical and statistical methods in recent years. Here we demonstrate that artificial intelligence can skilfully downscale low-resolution climate model data when combined with numerical climate model data. We show that recently developed image inpainting technique perform accurate super-resolution via transfer learning using the HighResMIP of CMIP6 (Coupled Model Intercomparison Project Phase 6) experiments. Its huge database offers a unique training opportunity for machine learning approaches. The transfer learning purpose allows also to downscale other CMIP6 experiments and models, as well as observational data like HadCRUT5. Combined with the technology of Kadow et al. 2020 of infilling missing climate data, we gain a neural network which reconstructs and downscales the important observational data sets (IPCC AR6) at the same time.

KEYWORDS

Machine-Learning, Image-Inpainting, Super-Resolution, Convolutional-Neural-Network, Climate-Science, Historical Temperature Measurements, Transfer-Learning

Material from: Kadow, C., Hall, D.M. & Ulbrich, U. Artificial intelligence reconstructs missing climate information. *Nat. Geosci.* 13, 408–413 (2020). https://doi.org/10.1038/s41561-020-0582-5

1 Missing Measurements – Image Inpainting

Present day climate change research relies upon climate observations from the past. Historic temperature measurements form the basis for building global gridded datasets like HadCRUT4/5 [1], which are investigated for example in the IPCC reports. Artificial Intelligence (AI) provides a particularly powerful measure to reconstruct missing values. We found that image inpainting technology, using partial convolutions in deep neural networks [2], can be trained to produce accurate results using 20CR [3] reanalysis and CMIP5 [4] experiments. These climate data sets do not suffer from missing data as they are derived from climate models rather than observations. The trained 20crAI and cmipAI deep neural networks are able to independently and in a cross-manner reconstruct artificially trimmed versions of the 20CR and CMIP5 products in grid space for every given month using the HadCRUT4 missing value mask. The analysis for the global mean temperature exhibits high temporal correlations (r≥0.9940), low errors (rmse≤0.0547°C), and an additional value compared with a state-of-the-art kriging interpolation. Furthermore, we restore a spatial pattern of a hidden El Niño in July 1877 of HadCRUT4 (Fig. 1). As compared with the currently established overall picture, the AI-restored HadCRUT4 annual global mean reconstruction points to a cooler 19th century, a weaker hiatus in the 21st century, and a stronger global trend between 1850 and 2018. Our results demonstrate that the combination of long observational records, numerical Earth system modeling, and modern AI techniques can produce valuable hybrid datasets for investigating missing segments of the past [5].

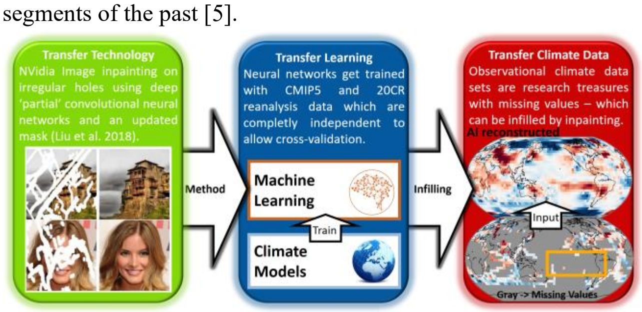

Figure 1: Transfer of the image technology towards a climate research tool to investigate climate change.

2 Downscaling – Super Resolution

Global climate change acts as a regional phenomena on a local scale. However, ground measurements are sparse and even more sparse, the further we go into the past. Therefore gridded datasets like HadCRUT4 and HandCRUT5, which combine station or ship or boey measurements into one larger grid cell, reach 5x5 degree grids – more than 500km per cell. On the other side to potentially counterplay this with high resolved numerical modeling, weather and climate models need extensive compute resources on high performance computers (HPC). The climate community is aiming for 1km climate simulations on exascale computers. In the meantime, the Coupled Intercomparison Project 6 (CMIP6) [6] set a High Resolution Model Intercomparison Project (HighResMIP) to investigate global models with much higher resolution. Multiple institutes and models participate and offer a database of climate experiments of higher model outputs than 1x1 degree grids. All available models were conservatively interpolated and remapped to a common 1x1 degree grid for near-surface air temperature focusing in the historical experiment *hist-1950* [7]. Furthermore, this set got conservatively interpolated to the original 5x5 degree grid of HadCRUT4/5. This enables us to set the training, validation and test dataset from 5 to 1 degree. Already the original study from Liu et al. 2018 [2] investigated partly the ability of its technology for super-resolution of images. Therefore, they basically introduced missing values around available pixels to upscale the information – and then infill using again image inpainting. We transformed this approach into the climate model setup comparable to the climate reconstruction of missing values. We obtained a valid strategy for super-resolution of climate data, which is basically downscaling in climate research (Fig. 2). We reached a high agreement with the original test datasets (ground truth). As the scientific transfer learning idea is also valid here, we are able to downscale also the observational data sets HadCRUT4/5.

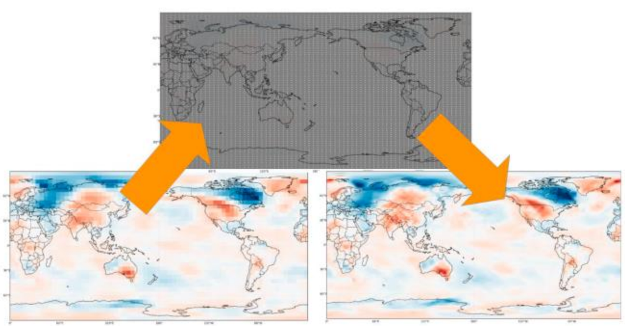

Figure 2: Super-Resolution or downscaling of climate data from a 5x5° grid to a 1x1° grid with the example of near-surface air temperature for an application observational data HadCRUT4/5.

3 Conclusion

Within this study, we are able to combine two very important tasks for climate research to investigate climate change:

1. Missing measurements
2. Downscaling

These two were performed by a transformed image technology originally developed to repair broken photographs and to upscale them. While the use of that technology for photographs was still possible on strong but regular PCs, the application around training with climate data needed the use of super-computing. Loading and balancing Tera-Bytes of NetCDF/HDF data needed huge resources for computing and storage.

The developed technology is now able to perform both tasks for the climate community. This successful application shows, that it is important to connect the communities for further research to investigate the important challenges humanity is facing in the upcoming decades.

ACKNOWLEDGMENTS

We thank the HPC-Service of ZEDAT, Freie Universität Berlin and the German Climate Computing Center (DKRZ) for the computation resources; the Climatic Research Unit (CRU) of the University East Anglia (UEA) and the MetOffice UK for providing the HadCRUT4 and HadSST4 datasets; the Earth System Grid Federation (ESGF) for providing the CMIP5 and CMIP6 experiments; Nature Geoscience and its reviewers.

REFERENCES

[1] Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., et al. (2021). An updated assessment of near-surface temperature change from 1850: the HadCRUT5 data set. Journal of Geophysical Research: Atmospheres, 126, e2019JD032361. https://doi.org/10.1029/2019JD032361

[2] Liu, G. et al. in *Computer Vision—ECCV 2018* Lecture Notes in Computer Science, Vol. 11215 (eds Ferrari, V. et al.) 19–35 (Springer, 2018).

[3] Compo, G. P. et al. The Twentieth Century Reanalysis project. *Q. J. R. Meteorol. Soc.* **137**, 1–28 (2011).

[4] Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteor. Soc.* **93**, 485–498 (2012).

[5] Kadow, C., Hall, D.M. & Ulbrich, U. Artificial intelligence reconstructs missing climate information. Nat. Geosci. 13, 408–413 (2020). https://doi.org/10.1038/s41561-020-0582-5

[6] Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

[7] Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185-4208, doi:10.5194/gmd-9-4185-2016, 2016.