Science Bulletin xxx (xxxx) xxx

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Short Communications

An artificial intelligence reconstruction of global gridded surface winds

Lihong Zhou [a], Haofeng Liu [b], Xin Jiang [a], Alan D. Ziegler [c], Cesar Azorin-Molina [d], Jiang Liu [b], Zhenzhong Zeng [a,*]

[a] School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
[b] Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
[c] Faculty of Fisheries and Aquatic Resources, Maejo University, Chiang Mai 50290, Thailand
[d] Centro de Investigaciones sobre Desertificación, Consejo Superior de Investigaciones Científicas (CIDE, CSIC-UV-Generalitat Valenciana), Climate, Atmosphere and Ocean Laboratory (Climatoc-Lab), Moncada, Valencia 46113, Spain

ARTICLE INFO

Article history:
Received 8 March 2022
Received in revised form 15 June 2022
Accepted 17 June 2022
Available online xxxx

Gridded wind speed data products with global coverage and continuous long-term time series are widely used in many applications, such as evaluating wind energy potential [1] and drought processes [2]. However, some available products do not accurately reproduce observed wind speed trends on land [3,4], leading to biased or inaccurate conclusions in studies on wind-related phenomena. *In-situ* weather stations involve direct measurements that accurately preserve wind speed trends. Still, the uneven distribution and incomplete time series have constrained their widespread applications in regional and global analyses. These limitations, which we have encountered firsthand in investigating the global wind stilling and reversal phenomena [4], have inspired us to create a new global gridded surface wind product that preserves observed wind patterns and trends.

To do so, we used the partial convolutional neural network (PCNN) to generate the wind product [5]. The PCNN is an artificial intelligence algorithm that maintains high performance even with large amounts of missing data [5–7]. In our case, we first trained our PCNN-based model with wind speed data (Table S1 online) from the Coupled Model Intercomparison Project Phase 6 (CMIP6, [8]). During this process, the PCNN-based model learned the wind pattern in CMIP6 by mastering the relationship between adjacent grid points linked by physical laws [7]. After getting a well-trained PCNN-based model, we used our model to restore the spatial–temporal gap of the gridded HadISD dataset (Fig. S1 online; [9]). The result was a globally complete gridded monthly wind speed dataset (1.25° × 2.5°), spanning from 1973 to 2020, which we call GGWS-PCNN (Global Gridded Monthly Wind Speed Dataset by the Partial Convolutional Neural Network). A diagram of our reconstruction process is represented in Fig. 1. Details about the data, reconstruction procedures, and our PCNN-based model can be found in the Supplementary materials (online).

We used three additional datasets to test the reconstruction performance of the well-trained model before creating the GGWS-PCNN (Fig. 1). Two were randomly selected models from the CMIP6 (i.e., CNRM-CM6-1, r1i1p1f2; EC-Earth3, and r1i1p1f1), and the other was the ERA5 reanalysis [10]. Neither CNRM-CM6-1 nor EC-Earth3 was among the CMIP6 models used for training (Table S1 online). The three datasets were fed into the well-trained model with the missing information of HadISD (i.e., masked datasets; see the Supplementary materials online for more details). The HadISD masks make the missing parts of the CNRM-CM6-1, EC-Earth3, and ERA5 the same as the HadISD. We then could compare their reconstructed version with the original data (examples are shown in Fig. S2 online).

Most grids (about 80%) are reasonably correlated ($r \geq |0.5|$, $P < 0.05$; Pearson correlation), indicating that the variability and trend of reconstructed data were reasonably close to that of the original data (Fig. S3a, c, e online). Interestingly, although our input data is very sparse for ocean areas (Fig. S1 online), the reconstructed wind speed in some locations had a strong relationship with the original data, for example, the Northeast Pacific Ocean, the North Atlantic Ocean, the Indian Ocean, and the Equator (Fig. S3a, c, e online). Also, some terrestrial areas with infrequent data were highly correlated, including India and West Siberia (Fig. S3a, c, e online).

The pattern of the root mean square error (RMSE) was different from that of the correlation (Fig. S3b, d, f online). Overall, lower RMSE occurred in or near the grids with input data. In other words, the more original wind data available (Fig. S1 online), the smaller the RMSE for a given area. The RMSE ranged from 0 to 0.8 m s^{-1} for the terrestrial grids, and the values were greater than

* Corresponding author.
E-mail address: zengzz@sustech.edu.cn (Z. Zeng).

https://doi.org/10.1016/j.scib.2022.09.022

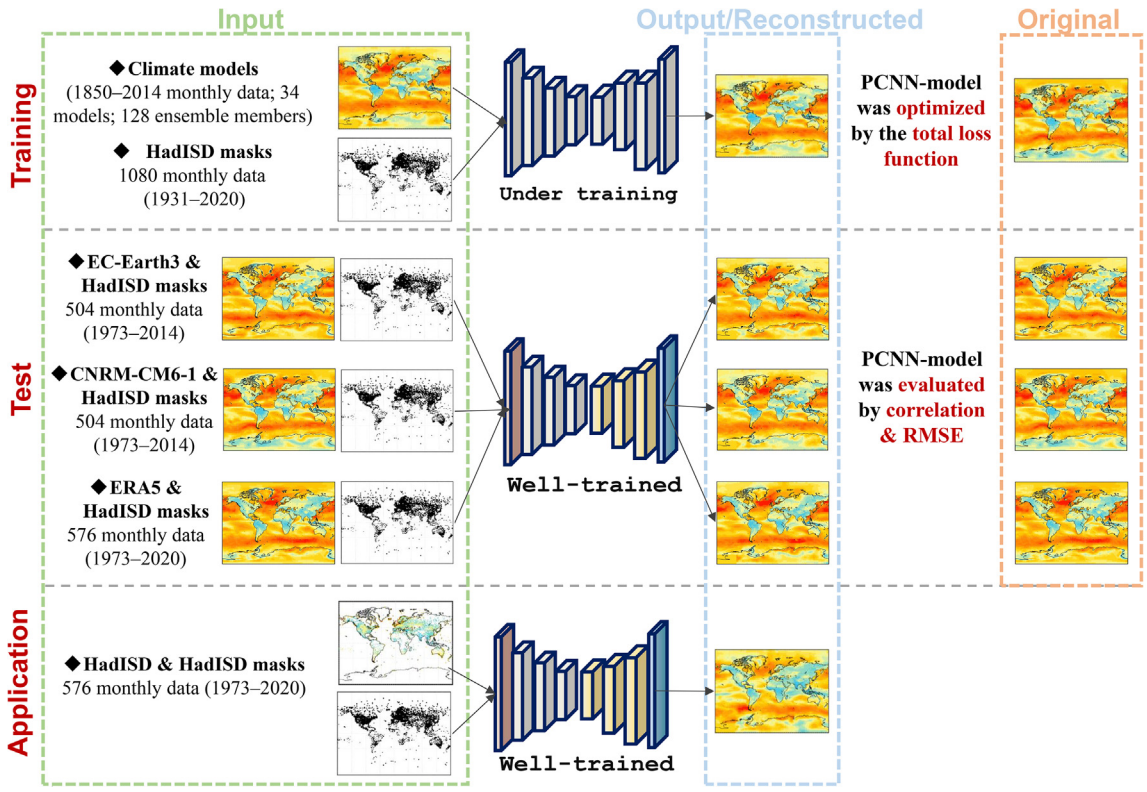

Fig. 1. Schematic of the PCNN-based model to reconstruct observed wind speed globally. HadISD masks enabled only the data in areas with observations to be used for reconstruction.

0.8 m s^{-1} for the marine grids. Our input data were mainly terrestrial wind speeds (Fig. S1 online), which are commonly weaker than ocean winds due to higher land friction caused by surface roughness. However, our reconstructed offshore wind speeds were not always lower than their original values (Fig. S4 online); only some areas that provided very little input data, such as the Antarctic and Greenland, were underestimated (Fig. S4 online). These comparisons demonstrated that the well-trained PCNN-based model could infer the wind fields of the whole area even when the number of observations was small.

Terrestrial wind speed changes have been a hot topic in the recent decade, e.g., regarding the stilling and reversal phenomena [11,12]. As another test, we investigated whether the well-trained PCNN-based model can reproduce global and regional wind speed trends. We divided the globe into six continents and compared the original and reconstructed data series to check their consistency. The reconstructed data in almost all regions captured the original data's variability and trend ($r \geq 0.78$, $P < 0.05$; RMSE ≤ 0.17 m s^{-1}; Fig. S5 online). This outcome reflects our model's ability to reproduce the characteristics of the original data. We also noticed that although input data showed an evident trend, the reconstructed data was not disturbed by this trend and remains close to the original data (Fig. S5a4, b3, b6, c3, c6 online).

The reconstructed dataset GGWS-PCNN has a variety of uses, including investigating long-term (1973–2020) terrestrial wind speed changes, an application that typically relies on reanalysis products or site-specific weather observations [12–14]. Wind speed trends in northern South America and central Africa varied among reanalysis products (ERA-I, MERRA-2, and JRA-55) [15]. Further, studies using observatories usually removed stations lacking continuous data during the study periods, which means trends in these studies may only reflect trends in areas with a high density of stations [12,13]. The GGWS-PCNN dataset allows for analyzing observed wind speed phenomena covering the globe and all continents (except the Antarctic).

The GGWS-PCNN dataset shows that the stilling phenomenon before the 2010s was mainly attributed to the decreasing wind speed in the Northern Hemisphere and Africa. Still, all six areas experienced a wind speed increase around 2010, contributing to the reversal of the global wind speed (red lines in Fig. 2; Table S2 online). The most rapid growth occurred in Asia. We further tested if Asia contributed disproportionally to the global increase by examining the time series of all the terrestrial areas except Asia. We found that the average wind speed of these areas increased after 2010 (Fig. S6 online). While Oceania is the only continent that recorded a continuous rise from 1973 to 2020, South America and Africa did not see a significant change (Fig. 2). Additionally, we noted that trends of the input/reconstructed/continuous series in the Northern Hemisphere and Oceania regions were similar (Fig. 2c, d, e). We suggest that this similarity indicates that the observational network in these areas is sufficient to show the regions' average wind speed state, as trends of the input series did not interfere the reconstructed results. This explanation is supported by the fact that the proportion of grids with observations is higher in these regions than in South America and Africa (Fig. 2a).

We also examined turning points and trends of time series in each terrestrial grid (Fig. S7 online). The turning points ($P < 0.05$) for 20% of grids occurred between 2008 and 2012 (subplots in Fig. S7 online); most of them were in northern Asia (Siberia). We further investigated the trends before and after the turning points, as well as trends across the series (Fig. S8 online). We used numbers one to three to denote trends of the former and D/I (i.e., decreasing/increasing) to denote those of the latter. Finally, there were six possible trend scenarios, which we named D1, D2, D3, I1, I2, and I3 (subplots in Fig. S8 online). For Europe and Asia, a large proportion of the land had an increasing trend after turning

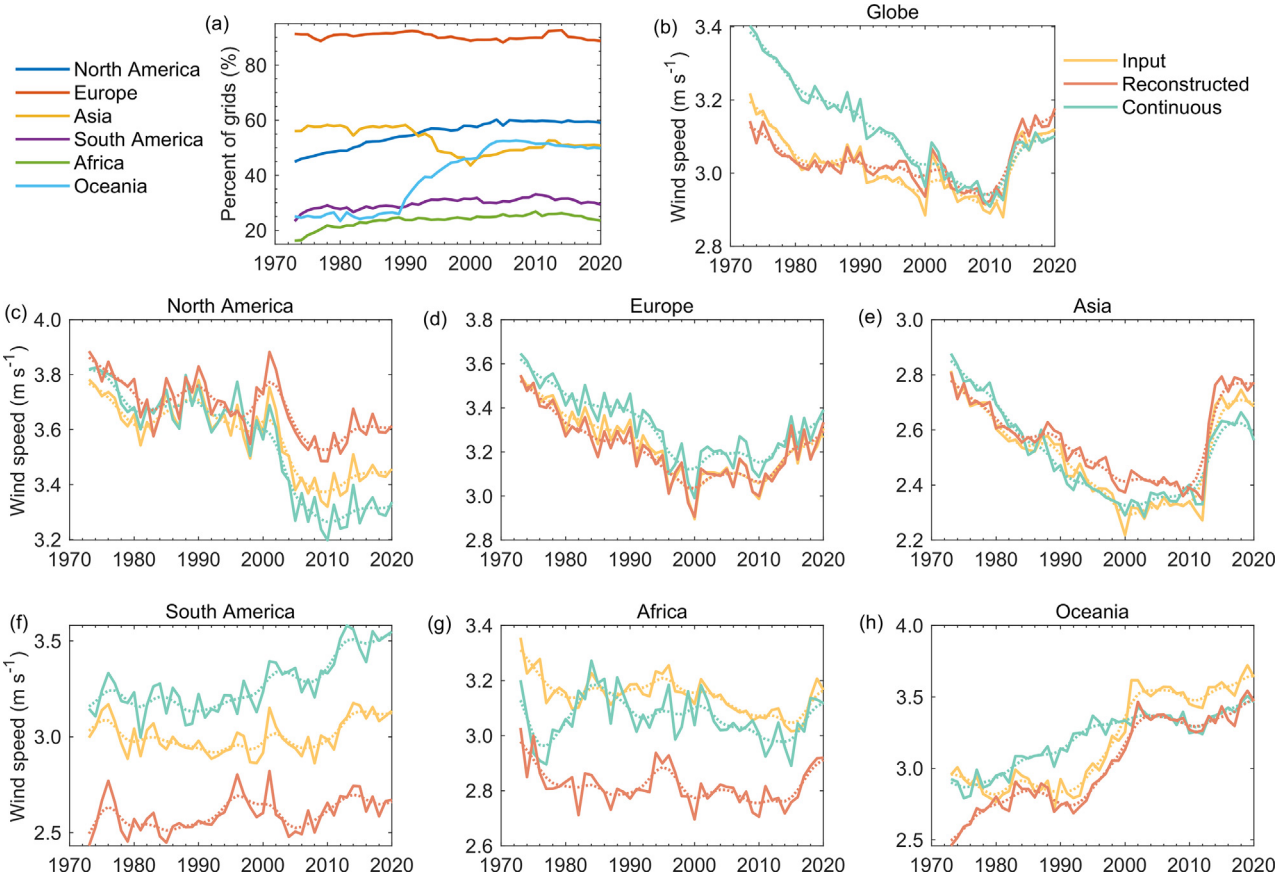

Fig. 2. Annual percentage of land grids sampled for wind speed (a) and time series of HadISD-based reconstructed terrestrial wind speed in the globe and six continents (b–h). The dash lines in (b–h) are the multidecadal variability of the corresponding time series (10-a Gaussian low-pass filter). Note that the input data is from limited grids, whose number varies from yearly; their trends are thus meaningless. Reconstructed data is from all grids in the corresponding area. Continuous data represents the average wind speed of grids with continuous 576 monthly observations (1973–2020).

points despite the whole period's declining tendency (Scenario D1; Fig. S8 online). Almost the entire Australian mainland experienced a rise in wind speed between 1973 and 2020; a similar trend occurred in large parts of Africa and South America (Fig. S8 online). North America seems to be the only continent where most grids showed decreasing trends (Fig. S8 online). However, our method of detecting turning points may mask North America's recent wind speed reversal (subplots in Fig. S8 online).

We intended to reconstruct all available monthly data in HadISD (i.e., 1931–2020). However, some invalid reconstructed data appeared in the pre-1973 images (see areas with the same colour in pre-1973 subplots in Fig. S9 online) because of large gaps in input data (white areas in pre-1973 subplots in Fig. S10 online). Thus, while the PCNN algorithm is robust for inpainting holes in images, it can fail when the missing portion of the images is greater than 88% (Fig. S11 online). The lack of input data partially explains why the reconstructed wind speed for 1973–2020 in the southeastern Pacific Ocean has a low correlation with the original, as well as a large RMSE (Fig. S3 online).

In sum, the PCNN algorithm was capable of overcoming a large number of missing values (at most 88%). Our application of the PCNN represents an encouraging first attempt to create a wind product based on artificial intelligence, observations, and climate models. Once the Met Office updates the HadISD, our GGWS-PCNN dataset will be available to the global community. In future work, we will consider developing a finer resolution product by adopting other state-of-the-art algorithms, assimilating other wind products (such as buoy data or ship observations), or applying simulated wind speed datasets with higher resolution. The data and codes in this study will be open-source at https://www.zhenzhongzeng.com/resources/.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (42071022), the start-up fund provided by Southern University of Science and Technology (29/Y01296122), and Highlight Project on Water Security and Global Change of Southern University of Science and Technology (G02296302). Cesar Azorin-Molina was supported by Evaluación y atribución de la variabilidad de la velocidad media y las rachas máximas de viento: causas del fenómeno stilling (RTI2018-095749-A-100), Cambios observados, proyecciones futuras e índices de la velocidad del viento y sus extremos en la Comunidad Valenciana (AICO/2021/023), and the Spanish National Research Centre Interdisciplinary Thematic Platform PTI-CLIMA. We thank Met Office (HadISD), ECMWF (ERA5), and CMIP6 for providing wind speed data used in this work.

Author contributions

Zhenzhong Zeng designed the research; Lihong Zhou performed the analysis and wrote the draft; Haofeng Liu ran the AI model

together with Lihong Zhou. All authors contributed to the interpretation of the results and the writing of the paper.

Appendix A. Supplementary materials

Supplementary materials to this short communication can be found online at https://doi.org/10.1016/j.scib.2022.09.022.

References

[1] Huang JL, McElroy MB. A 32-year perspective on the origin of wind energy in a warming climate. Renew Energy 2015;77:482–92.
[2] Sheffield J, Wood EF, Roderick ML. Little change in global drought over the past 60 years. Nature 2012;491:435.
[3] Miao HZ, Dong DH, Huang G, et al. Evaluation of Northern Hemisphere surface wind speed and wind power density in multiple reanalysis datasets. Energy 2020;200:117382.
[4] Zhang ZT, Wang KC. Stilling and recovery of the surface wind speed based on observation, reanalysis, and geostrophic wind theory over China from 1960 to 2017. J Clim 2020;33:3989–4008.
[5] Liu GL, Reda FA, Shih KJ, et al. Image inpainting for irregular holes using partial convolutions. In: Proceedings of the European Conference on Computer Vision (ECCV). p. 85–100.
[6] Lops Y, Pouyaei A, Choi Y, et al. Application of a partial convolutional neural network for estimating geostationary aerosol optical depth data. Geophys Res Lett 2021;48. e2021GL093096.
[7] Kadow C, Hall DM, Ulbrich U. Artificial intelligence reconstructs missing climate information. Nat Geosci 2020;13:408–13.
[8] Eyring V, Bony S, Meehl GA, et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 2016;9:1937–58.
[9] Dunn RJ, Willett KM, Parker DE, et al. Expanding HadISD: quality-controlled, sub-daily station data from 1931. Geosci Instrum Meth 2016;5:473–91.
[10] Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc 2020;146:1999–2049.
[11] Vautard R, Cattiaux J, Yiou P, et al. Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness. Nat Geosci 2010;3:756–61.
[12] Zeng ZZ, Ziegler AD, Searchinger T, et al. A reversal in global terrestrial stilling and its implications for wind energy production. Nat Clim Chang 2019;9:979–85.
[13] Yang Q, Li MX, Zu ZQ, et al. Has the stilling of the surface wind speed ended in China? Sci China Earth Sci 2021;64:1036–49.
[14] Deng KQ, Azorin-Molina C, Minola L, et al. Global near-surface wind speed changes over the last decades revealed by reanalyses and CMIP6 model simulations. J Clim 2021;34:2219–34.
[15] Torralba V, Doblas-Reyes FJ, Gonzalez-Reviriego N. Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison. Environ Res Lett 2017;12:114019.

Lihong Zhou received her B.S. degree from Tongji University in 2019, and M.S. degree from Southern University of Science and Technology in 2022. Her research interest is wind speed changes and their causes under global warming.

Zhenzhong Zeng received his B.S. degree from Sun Yat-sen University in 2011, and Ph.D. degree from Peking University in 2016. He led a tropical deforestation project at Princeton University during 2016–2019. Since 2019, he led a lab on Earth system and global change, with a focus on global climate change and its natural-based solutions.