Development of a high-resolution gridded climate dataset for Vietnam by blending ERA5-Land and in-situ data in the period 1981—2019

Hong Nguyen Ngoc Kim[1], Thanh Ngo Duc[2], Dairaku Koji[1]

Department of Engineering Mechanics and Energy, University of Tsukuba, Japan [1]
Department of Space and Applications, University of Science and Technology of Hanoi, Vietnam[2]

筑波大学 University of Tsukuba

25-29 SEPTEMBER 2023 TRIESTE, ITALY
ICRC-CORDEX 2023
International Conference On Regional Climate

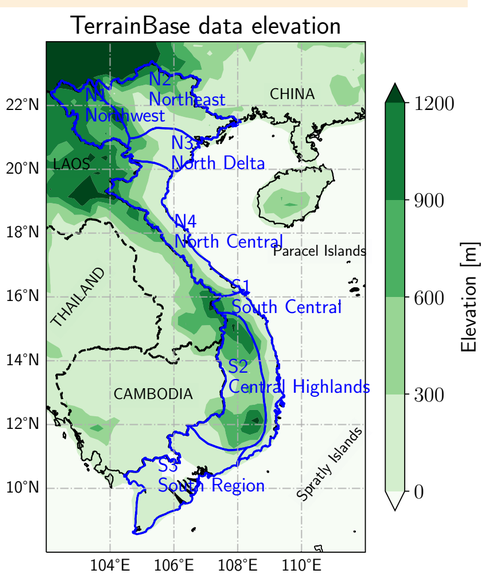
Abstract

By applying the **Cressman interpolation method**, a new precipitation and 2m-temperature dataset was created from 128 and 63 meteorological stations, respectively, and the **ERA5-Land** reanalysis for Vietnam. The dataset includes the period 1981—2019 with a native resolution of 9 km x 9 km monthly. To evaluate the performance of the newly created dataset, an effective statistical method, the Taylor diagram is used. Besides, comparisons between the newly created one, other regional datasets, and GCMs are performed. The results show that the new **high-resolution gridded climate dataset**has significant potential and better improvement.

Introduction

- Vietnam is characterized by diverse topography and extensive mountainous regions, giving rise to a complex climate pattern that spans the entire country.
- While it grapples with limited meteorological data sources, exacerbating the vulnerability of the country to the impacts of climate change (MONRE, 2009).

$\Longrightarrow$ It is essential to develop new datasets, that serve as valuable supplements by combining ground-based stations with derived data from models and advanced techniques. This integrated approach will significantly enhance the accuracy and reality.



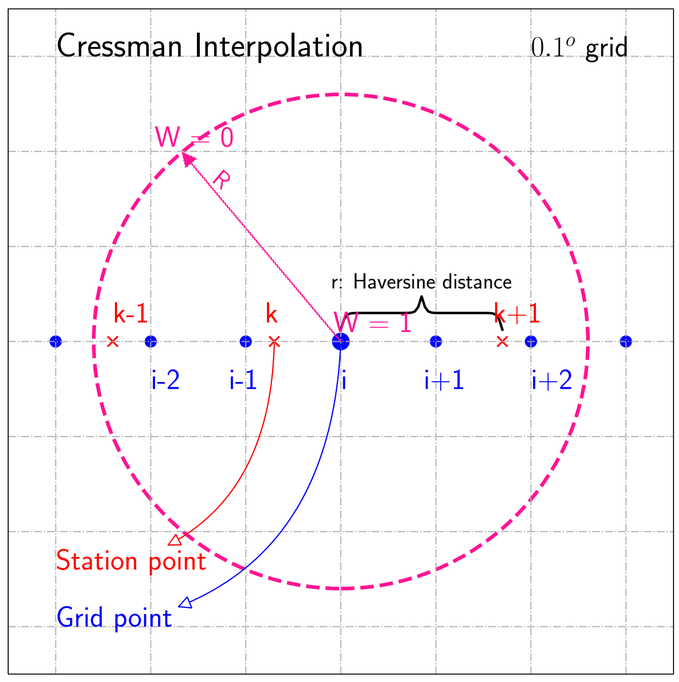
Used data and Methodology

ERA5-Land reanalysis data
a global land-surface by ECMWF
$0.1^o \times 0.1^o$; hourly; 1981—2019

In-situ data 128 precipitation and 63 2m-temperature stations; daily; 1981—2019

Cressman interpolation method (Cressman, 1959)

$$fo_i = fb_i + \frac{\sum_{k=1}^{K} W_{ik}\,(o_k - fb_k)}{\sum_{k=1}^{K} W_{ik}}$$

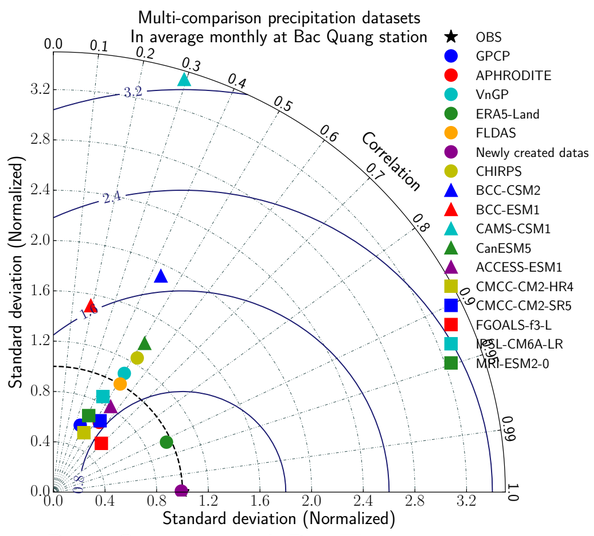


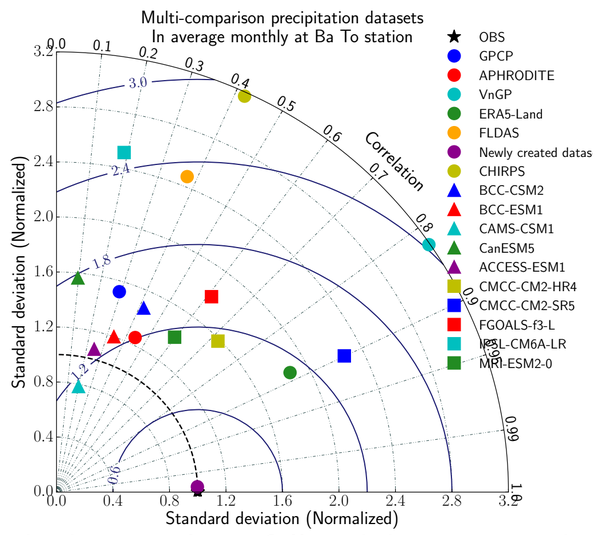
- fo, fb, o_k, and W_{ik}: output, background observed data, and weighted coefficients, respectively; i, k: indices for locating grid cells.

$$W_{ik} = \frac{R^2 - r_{ik}^2}{R^2 + r_{ik}^2},\ r_{ik}^2 \le R^2$$
$$W_{ik} = 0,\ r_{ik}^2 > R^2$$

- R: the influent radius; r_{ik}: the distance of the grid point to the located station.
- Stations located closer to the grid point have the greater weight and conversely, the greater the distance, the less weight the stations carry.

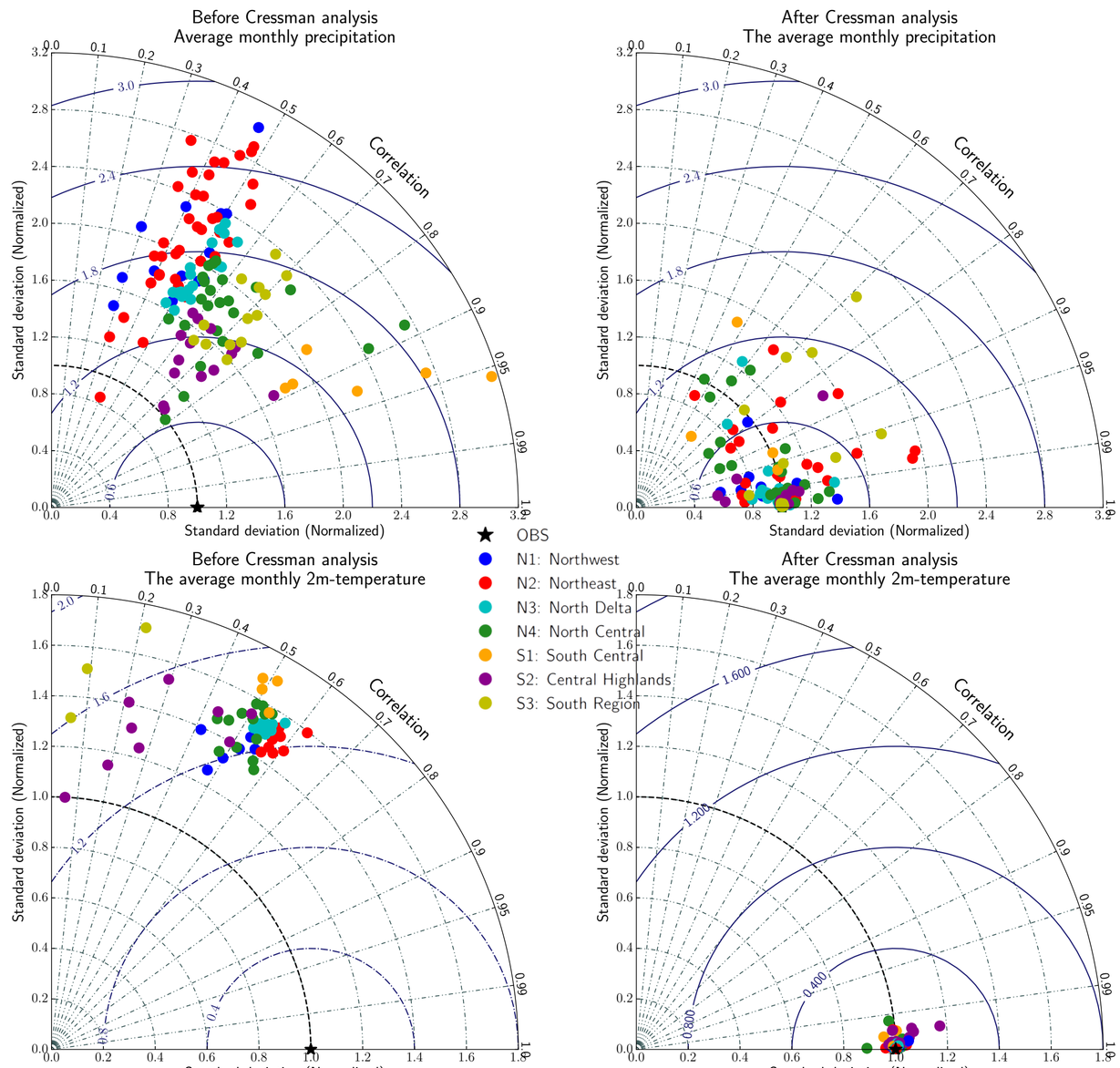
Multi-comparison datasets/ GCMs





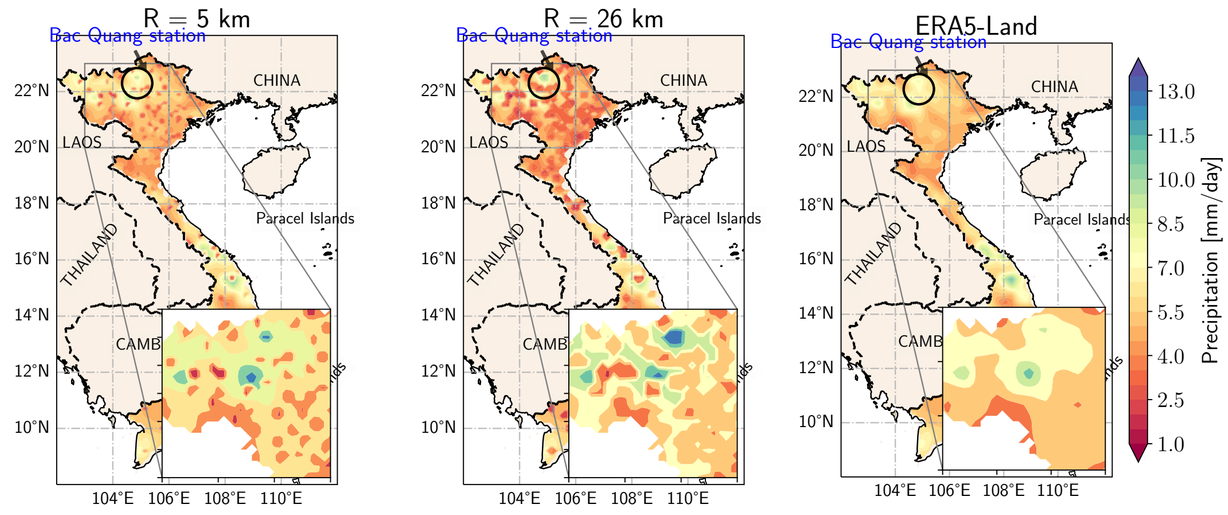
Bac Quang and Ba To, two stations with high annual rainfall, exhibit improved correlation in the new dataset, stemming from its finer resolution compared to most models. Some GCMs excel at one station but struggle at the other.

Taylor diagram: 7 climatic sub-regions



Correlation values of the Northern regions typically fall on 0.3—0.6. Toward the South, these correlation values increase slightly, around 0.5—0.9, except for the S1 sub-region. After Cressman in precipitation, the valued points oscillate around the reference, which is due to the cumulative and complex nature of rain.

The effect of influenced radius



In the dataset, opting for a radius of 26 km stands out as the optimal choice for enhancing data quality in the Northern region, characterized by the densest station allocations. The utilization of the Taylor diagram for comparison purposes reveals significant improvements compared to a radius of 11.5 and 5km, which proves insufficient to influence neighboring locations.

Conclusions

The new dataset, spanning from 1981 to 2019, boasts a high-resolution grid of $\mathbf{0.1^o \times 0.1^o}$ achieved via the Cressman method. It seamlessly is blended from ERA5-Land and 128 ground meteorological stations for precipitation, as well as 63 ones for 2m-temperature. The Taylor diagram reveals marked improvements in precipitation, validated against multiple datasets and GCMs. Through the independent approach, a 26 km radius emerges as the most suitable choice. It's essential to recognize both the strengths and limitations of the parent dataset concerning specific regions or phenomena. Future developments may explore advanced techniques such as Optimal Interpolation, 4D-Var, and Kalman Filter.

References

[1] Thanh Nguyen-Xuan, Thanh Ngo-Duc, Hideyuki Kamimera, Long Trinh-Tuan, Jun Matsumoto, Tomoshige Inoue, Tan Phan-Van, The Vietnam Gridded Precipitation (VnGP) Dataset: Construction and Validation, SOLA, 2016, Volume 12, Pages 291-296, Released on J-STAGE November 14, 2016, Online ISSN 1349-6476, https://doi.org/10.2151/sola.2016-057
[2] MONRE 2009 Climate Change, Sea Level Rise Scenarios for Vietnam. Vietnam Natural Resources, Environment and Mapping Publishing House, Hanoi, Vietnam, p. 34.
[3] Cressman G. P. 1959 An operational objective analysis system. Mon. Weather Rev. 87, 367–374. http://dx.doi.org/10.1175/1520-0493
[4] Nguyen D. N., Nguyen T. H. 2004 Vietnamese Climate and Climatic Resources. Science and Technology Publishing House, Vietnam, p. 230, (in Vietnamese).

hongnguyen.ngock2@gmail.com