

Research on Wind Energy Data Quality Control Methods in Qinghai Province on the Qinghai-Tibet Plateau (2019–2021)

Bingyu Zhao¹, Yizhe Han^{2,*}, Xinhua Huang¹, Jing Wang¹, Xiaoyan Ma¹, Haowei Yang³, Xuanyu Zhang¹, Qiong Chao¹, Yan Guo¹

¹ Qinghai Meteorological Information Center, Xining 810001, Qinghai, China

² National Meteorological Information Center, Beijing 100081, China

³ Haidong Meteorological Bureau, Haidong 810600, Qinghai, China

* Corresponding author: Yizhe Han (Email: hanyizhe@cma.gov.cn)

Abstract: This paper presents the construction methodology and quality assessment results of the “Qinghai-Xizang Plateau Multi-source Integrated Wind Energy Basic Dataset (V1.0)”. Addressing the challenges in wind energy resource assessment over the complex terrain of the Qinghai-Xizang Plateau, this study establishes a comprehensive data quality control system. The dataset spans from 2019 to 2021 and integrates observations from 12 meteorological towers and 53 national-level weather stations across Qinghai Province, along with multi-source fused analysis products at a spatial resolution of 5 kilometers. A tiered quality control process is applied, including range checks, temporal consistency checks, and vertical consistency checks. Additionally, dynamic threshold adjustment rules are designed based on historical climate extremes, and a multi-factor correlation logic incorporating temperature-humidity coupling is introduced, leading to an innovative algorithm for wind tower data quality control that significantly improves data quality in high-altitude regions of Qinghai Province. Assessment results show that the overall accuracy of the data after quality control reaches 93.28%, substantially higher than the national average. This dataset provides solid foundational support for wind energy resource evaluation, climate research, and low-carbon transition efforts in northwestern China, and demonstrates the effectiveness of multi-source data fusion techniques in enhancing the accuracy of meteorological observations in high-altitude areas.

Keywords: Multi-source Fusion, quality control, wind energy resource, Qinghai Province.

1. Introduction

Qinghai Province, located in the northeastern part of the Qinghai-Tibet Plateau, possesses abundant wind energy resources and serves as a critical clean energy base in China. The unique geographical location and high-altitude topography endow the region with strong and stable wind conditions, offering superior natural advantages for large-scale wind power development. With the installed capacity of clean energy in Qinghai surpassing 53 million kW and renewable energy accounting for over 93% of the province's electricity consumption, accurate wind resource assessment is paramount for optimizing the energy structure and supporting the national "Dual Carbon" goals.

The development of a high-quality wind energy fundamental dataset is essential for precise wind power forecasting, resource evaluation, and disaster warning services. While international initiatives such as the New European Wind Atlas (NEWA) and the Wind Integration National Dataset (WIND Toolkit) in the United States have set precedents, the complex plateau environment in Qinghai demands localized solutions. Although the overall data quality of anemometer towers in Qinghai is high, specific challenges persist, particularly in the Chaidamu Basin's strong wind zones and the high-cold transition areas like Wudaoliang, where sensor adaptability and algorithm robustness require improvement.

This paper outlines the development process of the Qinghai Provincial Multi-Source Fused Wind Energy Fundamental

Dataset (V1.0). The primary objective is to establish a standardized, high-accuracy dataset by implementing a rigorous quality control system and advanced multi-source fusion algorithms. The subsequent sections describe the data sources and methodologies, present the evaluation results of the dataset, and discuss the specific characteristics of wind resources in the region.

2. Materials (Data) and Methods

2.1. Data Sources

The study utilized three primary categories of data to ensure comprehensive coverage and accuracy:

(1) Anemometer Tower Observations: Data from 12 anemometer towers across Qinghai Province, covering the period from April 2019 to December 2021. These towers are part of the Fourth National Wind Energy Resources Census.

(2) National Meteorological Station Data: Long-serial observational data from 53 national-level weather stations within Qinghai Province, including a 30-year climatological compilation.

(3) Multi-Source Fusion Data: A 5km resolution multi-source fusion dataset providing temperature, pressure, wind speed, and wind direction data for the region.

2.2. Quality Control Procedures

A hierarchical progressive quality control flow was adopted, consisting of the following stages:

2.2.1. Basic Checks

The process began with file format verification and metadata integrity checks. This was followed by the identification of missing data and range value validation. The allowable range checks for wind speed, wind direction, temperature, humidity, and pressure were defined based on physical and instrumental limits (e.g., wind speed: 0-75 m/s for heights below 200m).

2.2.2. Time Consistency Check

To detect sensor freezing or sudden jumps, a time consistency test was applied. The algorithm checked for variations between consecutive hours. For instance, the hourly variation threshold for wind speed was set to less than 20 m/s, and for temperature to less than 5°C. A novel multi-element coupling mechanism was introduced, specifically a temperature-humidity coupling validation, to distinguish between sensor faults and real persistent weather phenomena (e.g., advection fog).

2.2.3. Vertical Consistency Check

Given the vertical structure of the atmosphere, a vertical consistency check was crucial for anemometer towers with multiple layers. The study optimized the threshold parameters for the northwestern plateau environment. The wind speed difference threshold for layers 20m apart was adjusted from 2 m/s to 4 m/s, and for 30m apart from 3 m/s to 6 m/s. Additionally, a wind shear coefficient coupling validation was introduced. If the logarithmic wind shear exponent (α) exceeded 0.2, the data was flagged for further analysis to differentiate between real atmospheric instability and sensor errors.

2.3. Multi-Source Fusion Algorithms

To address the limitations of single-source data, two key fusion algorithms were developed:

Dynamic Threshold Adjustment for Temperature: Instead of using fixed climatological limits, a dynamic threshold was calculated by adjusting historical extreme values based on the elevation of each tower (using a lapse rate of 0.06°C/100m) with a tolerance of $\pm 5^\circ\text{C}$.

Collaborative Judgment of Temperature and Pressure: The observed data was cross-validated with the 5km multi-source fusion analysis data. If the deviation from the fusion data exceeded 3%, the observation was flagged as erroneous, effectively filtering out data affected by local microclimates or instrument drift.

3. Results and Analysis

3.1. Data Quality Evaluation

The quality control assessment revealed that the overall correctness rate of the 12 anemometer towers was 93.28%, with a missing rate of 4.59% and a zero error rate. This performance is significantly higher than the national average anemometer tower correctness rate of 87.87%.

As shown in the evaluation results, 11 out of the 12 stations achieved a correctness rate above 91%, and 8 stations exceeded 95%. Notably, the Mangya Town station (29001) and Kuaierma station (29009) demonstrated exceptional quality, with correctness rates surpassing 99.95%. However, the Wudaoliang station (29012) presented a challenge, with a correctness rate of only 51.02%, primarily due to a high missing rate of 37.69% caused by the extreme high-altitude environment.

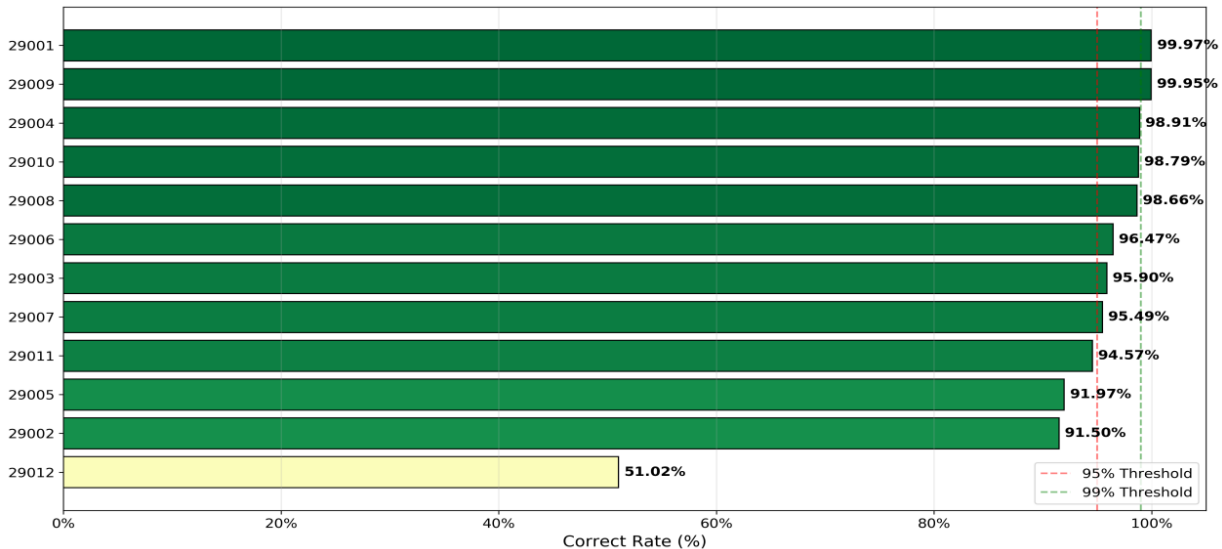


Figure 1. Correctness Rate Statistics of Qinghai Anemometer Tower Sites

3.2. Wind Resource Characteristics

The dataset reveals distinct wind patterns across Qinghai. Taking the Mangya Town station (29001) as a representative example, the wind rose diagrams indicate a dominant northwest wind direction. The frequency of the northwest wind (292.5° and 315.0°) is overwhelming, and this dominance intensifies with height. At 70m, the frequency of the northwest wind reached 3579 occurrences, indicating a

highly stable wind field structure. This suggests that wind turbines in this region should be primarily oriented to capture the prevailing northwest winds.

Furthermore, the analysis indicates that the wind data in the western Chaidamu Basin, while generally reliable, showed slightly lower correctness during extreme wind events compared to temperature, pressure, and humidity data. This highlights the need for continued refinement of quality control logic for wind sensors under gale conditions.

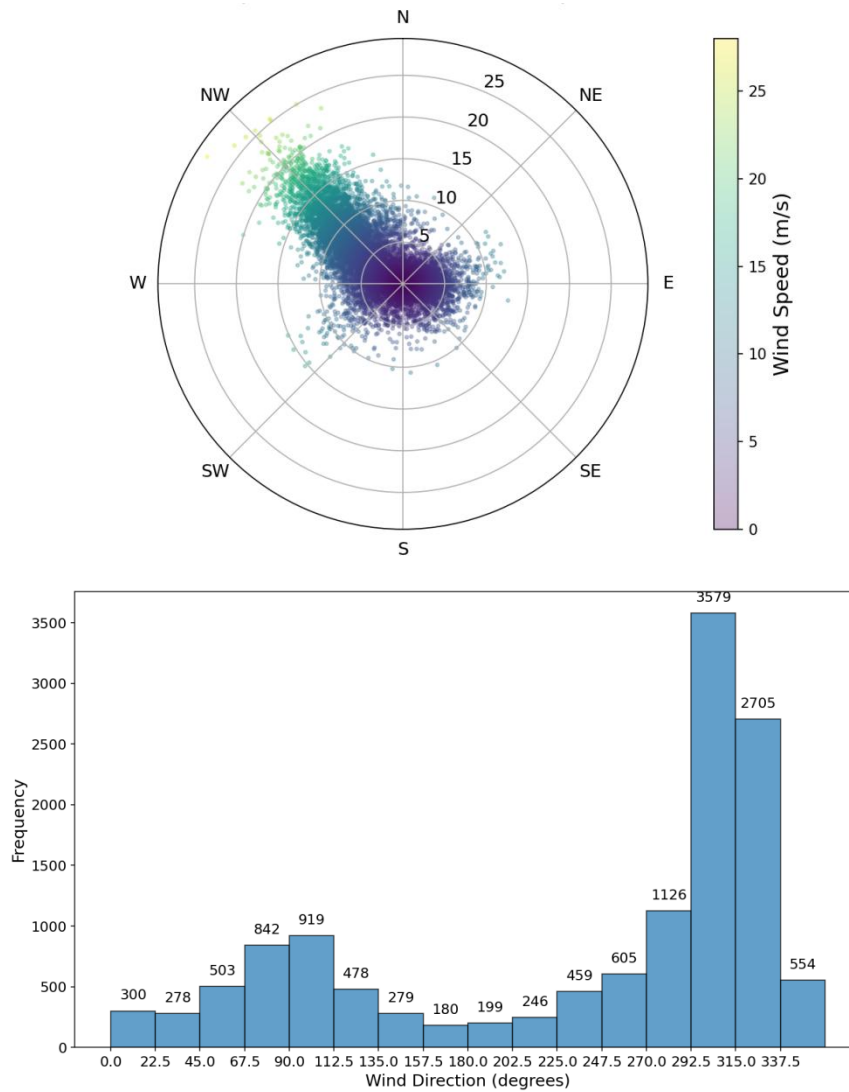


Figure 2. Wind Rose Diagram (a) and Wind Frequency Distribution Chart (b) at 70 m Height for Station 29001

4. Conclusion

This study successfully developed the Qinghai Provincial Multi-Source Fused Wind Energy Fundamental Dataset (V1.0). By integrating observational data from 12 anemometer towers, 53 national stations, and 5km multi-source fusion products, and by applying a rigorous quality control scheme with innovative fusion algorithms, a high-accuracy dataset covering the period 2019–2021 was established.

The evaluation demonstrates that the dataset has a high level of integrity and correctness (93.28%). The research overcame the limitations of single-source data by implementing dynamic threshold adjustments and multi-element coupling logic, significantly improving the reliability of wind observations in complex plateau terrain. This dataset provides a vital scientific foundation for wind energy resource assessment, renewable energy planning, and meteorological services in Qinghai Province. Future work will focus on continuously optimizing the correction algorithms based on user feedback and expanding the dataset's temporal coverage.

Acknowledgment

We thank the Qinghai Meteorological Information Center for providing the raw observational data and technical support. This work was supported by the Qinghai Meteorological

Bureau.

References

- [1] Dörenkämper, M., Olsen, B. T., Witha, B., et al. (2020). The making of the New European Wind Atlas—part 2: Production and evaluation. *Geoscientific Model Development Discussions*, 1–37.
- [2] Hahmann, A. N., Sile, T., Witha, B., et al. (2020). The making of the new European Wind Atlas, Part 1: model sensitivity. *Geoscientific Model Development Discussions*, 1–33.
- [3] Draxl, C., Clifton, A., Hodge, B. M., et al. (2015). The Wind Integration National Dataset (WIND) Toolkit. *Applied Energy*, 151, 355–366. <https://doi.org/10.1016/j.apenergy.2015.04.072>
- [4] Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- [5] Badger, J., Hahmann, A. N., Witha, B., et al. (2020). The New European Wind Atlas: Part 1—Model Sensitivity and Year Selection. *Wind Energy Science*, 5, 1251–1270. <https://doi.org/10.5194/wes-5-1251-2020>
- [6] National Energy Administration of China. (2018). NB/T 31147-2018 Technical code for wind energy resources measurement and assessment of wind farm engineering. China Water & Power Press.

- [7] National Bureau of Statistics of China. (2023). China Energy Statistics Yearbook 2023. China Statistics Press.