Supplementary Materials

Measurement report: A multi-year, high-resolution, county-scale atmospheric NH_3 dataset from the North China Plain

Jiyang Lyu[1#], Yu Niu[1#], Fanlei Meng[1], Wen Xu[1*]

[1] College of Resources and Environmental Sciences; National Observation and Research Station of Agriculture Green Development (Quzhou, Hebei), China Agricultural University, Beijing, 100193, China

*Correspondence to: W. Xu (wenxu@cau.edu.cn); Tel.: +86 18518151688

# Contributed equally to this work.

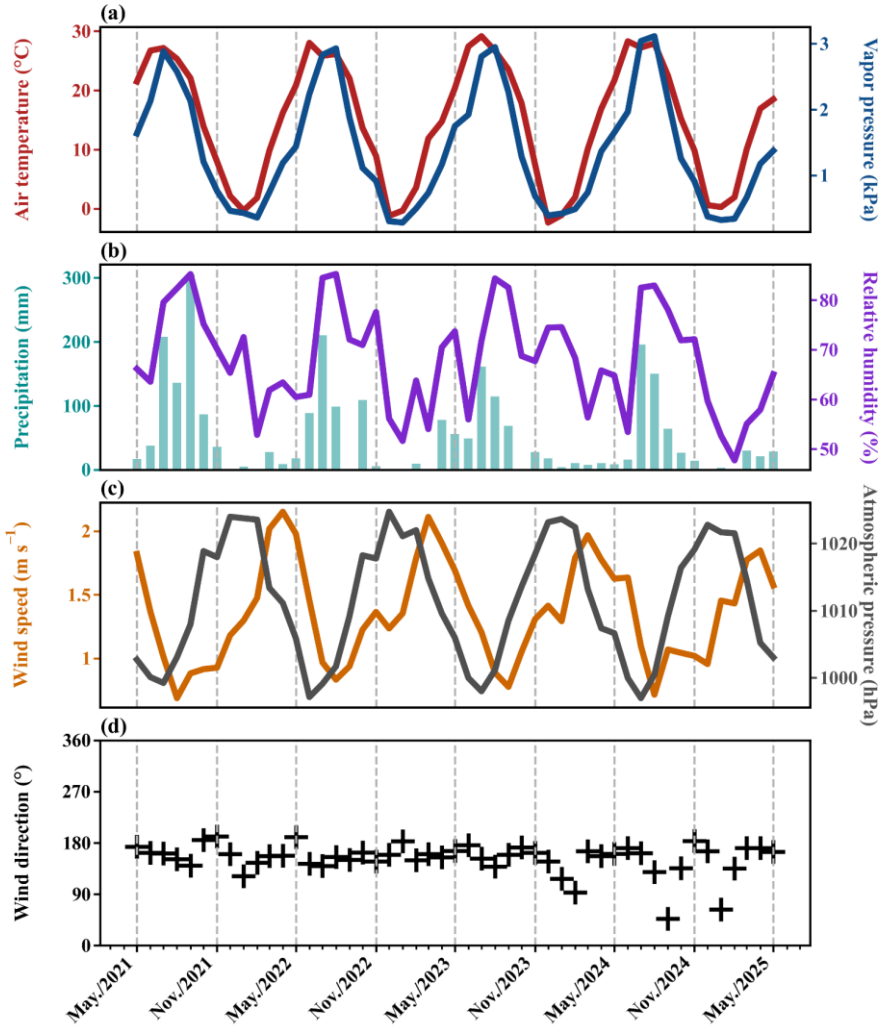

Figure S1. Monthly meteorological conditions in Quzhou County from May 2021 to May 2025. From top to bottom, the panels show air temperature and vapor pressure, precipitation and relative humidity, wind speed and atmospheric pressure, and wind direction. Precipitation is the monthly total averaged across the 10 stations. Dashed lines mark May and November.

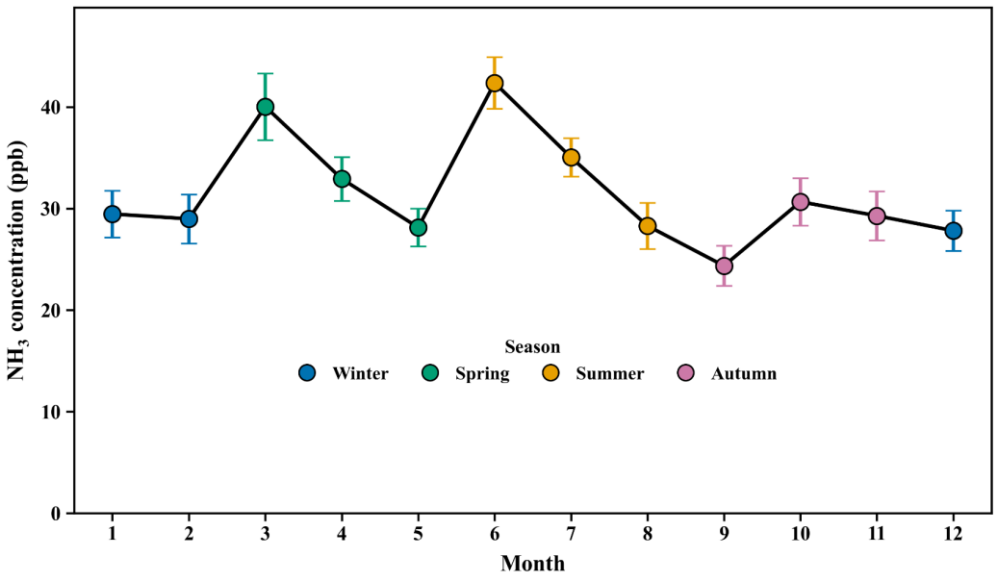

Figure S2. Monthly mean NH_3 concentrations in Quzhou County from May 2021 to May 2025. Each point is a monthly mean calculated from daily county means. Error bars show 95% confidence intervals, and colors indicate season.

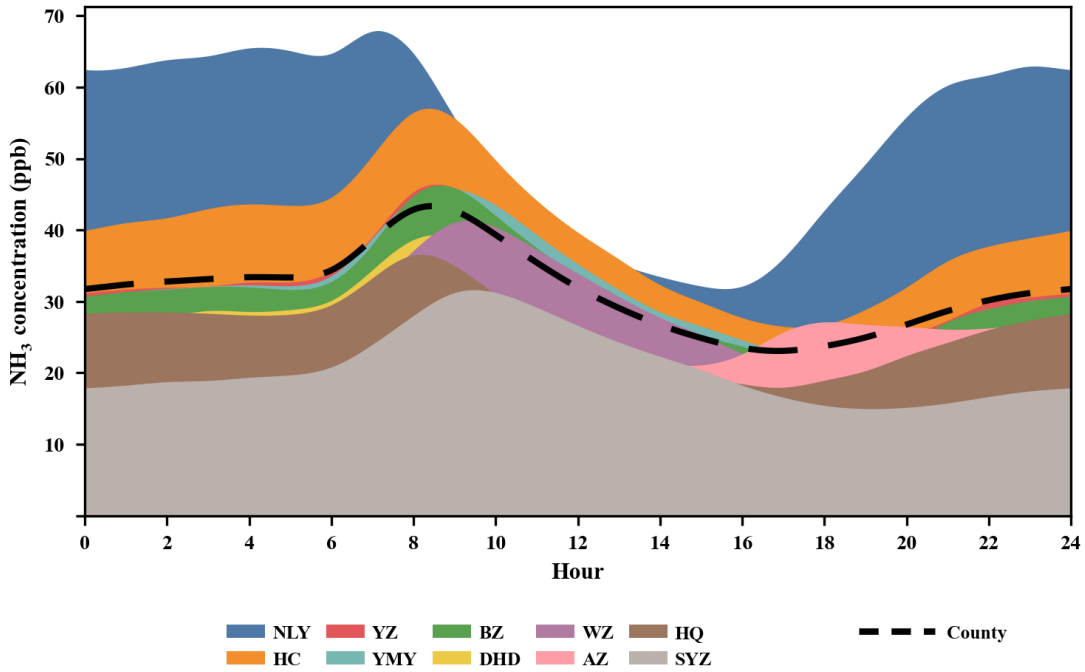

Figure S3. Mean diurnal cycles of NH_3 at the 10 monitoring sites and for Quzhou County. Colored areas show hourly means for individual sites connected with periodic cubic splines, and the black dashed line shows the county mean. Station abbreviations are defined in Fig. 1.

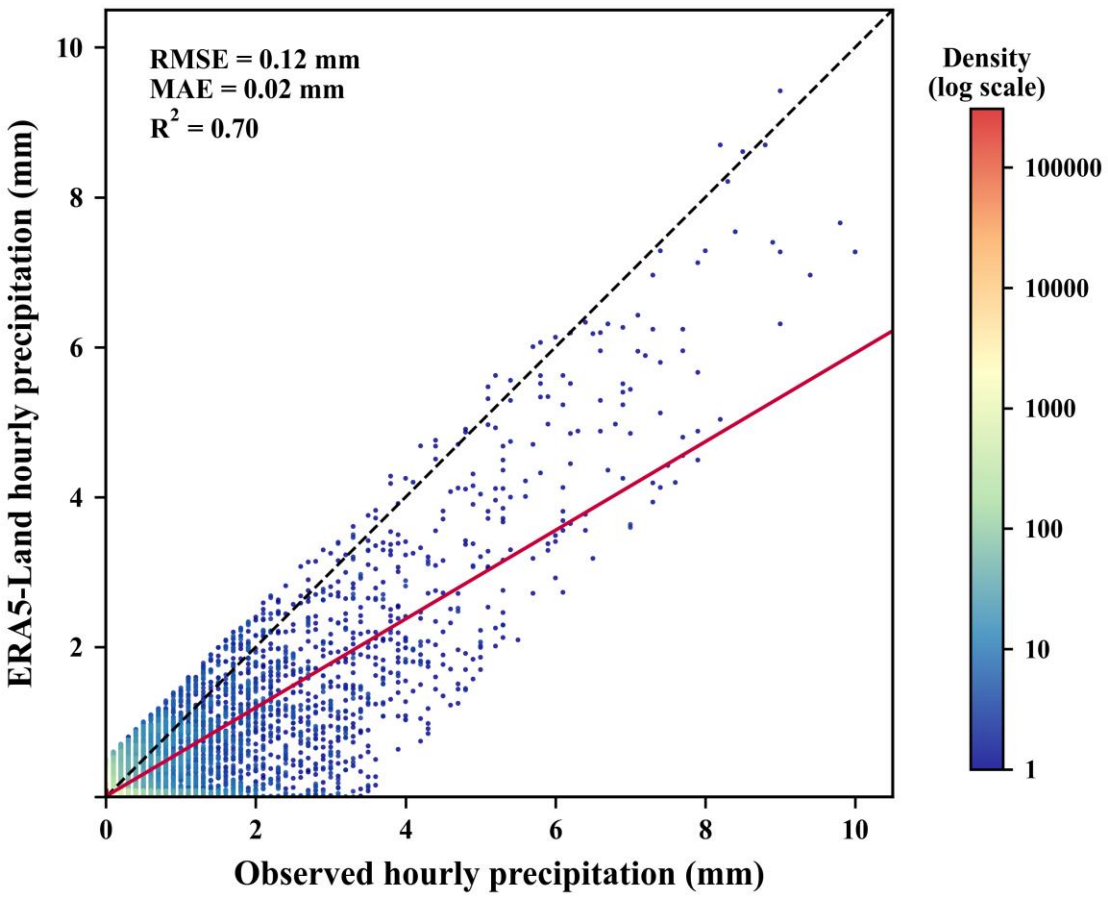

Figure S4. Comparison of hourly precipitation from ERA5-Land with rain gauge observations at the 10 monitoring stations. Point color indicates logarithmic density. The black dashed line is the 1:1 line, and the red line is the least squares regression.

Table S1. Information on the 10 NH_3 monitoring stations in Quzhou County.

Station	**Lon** (°E)	**Lat** (°N)	Period	Crops	Missing **(%)**
AZ	114.98	36.67	May 2021 to May 2025	Grapes	14.94%
BZ	114.95	36.71	May 2021 to May 2025	Wheat, maize	23.04%
DHD	114.92	36.68	May 2021 to May 2025	Wheat, maize	22.66%
WZ	115.03	36.87	May 2021 to May 2025	Wheat, maize	15.84%
HC	115.12	36.64	May 2021 to May 2025	Wheat, maize	25.91%
HQ	115.06	36.77	May 2021 to May 2025	Wheat, maize	20.62%
NLY	115.02	36.72	May 2021 to May 2025	Wheat, maize	23.39%
YZ	115.16	36.63	May 2021 to May 2025	Wheat, maize	18.25%
YMY	114.94	36.76	May 2021 to May 2025	Vegetables	23.38%
SYZ	115.01	36.86	May 2021 to May 2025	Wheat, maize	7.68%

Note: Missing rate is the percentage of the 35064 expected hourly NH_3 observations that were unavailable after quality control at each station.

Table S2. Quality control limits for meteorological variables. The precipitation limits were applied to both rain gauge observations and the ERA5-Land series for each station.

Meteorological parameter	Unit	Physical lower limit	Physical upper limit
air temperature	°C	−30	50
relative humidity	%	0	100
atmospheric pressure	hPa	800	1100
wind speed	$m\ s^{-1}$	0	60
wind direction	°	0	360
precipitation	$mm\ h^{-1}$	0	120
vapor pressure	kPa	0	

Note: Before model fitting, meteorological observations were screened using physical-range, persistence, and internal-consistency checks. General limits followed China Meteorological Administration (2020). The upper wind-speed and hourly precipitation screening limits were based on Park and Baek (2023) and Longman et al. (2018), respectively. Wind direction was invalid outside 0° to 360° or when the logger recorded 0° while wind speed was nonzero. Persistence screening was used to identify constant-value episodes (Kaffashzadeh, 2023); six identical consecutive hourly values were used as the study-specific criterion for the ammonia analyzer and meteorological sensors. Hourly precipitation data missing because of rain gauge malfunctions were filled with ERA5-Land total precipitation and replaced with the corresponding station's hourly ERA5-Land precipitation (Muñoz-Sabater et al., 2021; see Sect. 2.2 of the main text).

Table S3. Performance of the three models on the test set.

Model	R^2	RMSE	Fitting equation	MAE (> 60 ppb)	Maximum absolute residual	R^2 (> 60 ppb)
Random Forest	0.73	16.85	$y = 0.56x + 13.93$	28.52	389.12	0.52
LightGBM	0.74	15.46	$y = 0.7x + 9.39$	23.92	406.26	0.58
XGBoost	0.78	14.43	$y = 0.72x + 8.79$	22.16	387.74	0.61

Note: R^2 is the squared Pearson correlation between observed and predicted NH_3 concentrations. RMSE, the MAE for observations above 60 ppb, and the maximum absolute residual are reported in ppb. In each fitting equation, x is the observed concentration and y is the predicted concentration.

References:

China Meteorological Administration: Quality control of meteorological observation data - Surface (Standard No. QX/T 118-2020), Meteorological Press, Beijing, 2020.

Kaffashzadeh, N.: A data-driven persistence test for robust (probabilistic) quality control of measured environmental time series: constant value episodes, Atmos. Meas. Tech., 16, 3085-3100, https://doi.org/10.5194/amt-16-3085-2023, 2023.

Longman, R. J., Giambelluca, T. W., Nullet, M. A., Frazier, A. G., Kodama, K., Crausbay, S. D., Krushelnycky, P. D., Cordell, S., Clark, M. P., Newman, A. J., and Arnold, J. R.: Compilation of climate data from heterogeneous networks across the Hawaiian Islands, Sci. Data, 5, 180012, https://doi.org/10.1038/sdata.2018.12, 2018.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349-4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Park, M.-S. and Baek, K.: Quality management system for an IoT meteorological sensor network - Application to Smart Seoul Data of Things (S-DoT), Sensors, 23, 2384, https://doi.org/10.3390/s23052384, 2023.