



Technical note: A Flexible Framework for Precision Truncation and Lossless Compression in WRF Simulations with Application over the United States

Shang Wu¹, David C. Wong^{2,3}, Jiandong Wang¹, Yuzhi Jin¹, Junjun Li⁴, and Chunsong Lu¹

¹State Key Laboratory of Climate System Prediction and Risk Management, China Meteorological Administration Aerosol-Cloud and Precipitation Key Laboratory, and Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, China

²Atmospheric and Environmental Systems Modeling Division, US Environmental Protection Agency, Research Triangle Park, NC 27711, USA

³Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX 77204, USA

⁴National Institute of Education (NIE), Nanyang Technological University (NTU), Singapore 637616, Singapore.

Correspondence to: David C. Wong (wong.david-c@epa.gov) and Jiandong Wang (jiandong.wang@nuist.edu.cn)

Abstract. As climate simulations generate increasingly large datasets, reducing storage demands without compromising scientific integrity has become a critical challenge. This study evaluates the effectiveness of precision truncation, applied prior to lossless compression, in balancing storage efficiency and fidelity within regional Weather Research and Forecasting (WRF) simulations over the United States. We examine input-only, output-only, and combined input–output truncation strategies across both routine meteorological variables and extreme precipitation indices. Results show that conventional atmospheric fields remain robust when outputs are truncated to 5 or 4 significant digits, keeping biases within acceptable limits. Wind speed is largely insensitive to truncation, temperature and humidity are more vulnerable under aggressive output truncation (3 significant digits). Precipitation shows mixed responses, with deviations dominated by input perturbations. Extreme precipitation indices display more complex sensitivities: percentile- and maximum-based indices are highly susceptible to nonlinear, regionally heterogeneous biases under input truncation, whereas frequency- and intensity-based indices respond more systematically to output truncation, with substantial distortions emerging at 3 digits. These findings demonstrate that truncation strategies cannot be applied uniformly but must be tailored to variable type and diagnostic. Within this study, output-only truncation emerges as the most reliable strategy, with 4 significant digits identified as a safe lower bound and 5 digits preferable when fidelity of extreme-event is critical. To implement this in practice, we introduce a flexible error-tolerance framework that applies a predefined threshold across all indices and adapts truncation levels by region and season, enabling substantial storage savings while safeguarding the integrity of climate diagnostics.



1 Introduction

30 Climate variability and change, driven by both natural processes and human activities, have profound impacts on human societies and natural ecosystems. To better understand and project these processes, the scientific community increasingly demands climate simulations with higher spatial resolution, richer representations of physical processes, and large ensembles under multiple scenarios. This shift has greatly expanded the complexity and scale of numerical models, driving exponential growth in both input and output data volumes. Global-scale simulations now routinely generate tens to hundreds of petabytes of output (Overpeck et al., 2011), placing unprecedented pressure on storage and analysis capacities. The rapid rise in data 35 production poses significant challenges for existing storage infrastructures, constraining long-term archiving, data sharing, reproducibility of analyses, and the efficiency of postprocessing workflows. These challenges are further exacerbated by storage systems and network bandwidths that have not kept pace with the rapid gains in computational performance (Prein et al., 2015).

40 Importantly, the shift from global to regional modeling does not alleviate the challenges of data volume and management. The Weather Research and Forecasting (WRF) model (Skamarock et al., 2019), a widely used regional modeling system with a large international user base across academia, government agencies, private forecasting services, and independent researchers (Powers et al., 2017), exemplifies this issue. High-resolution WRF simulations generate several to hundreds of terabytes of output from a single project (Prein et al., 2015). Such datasets are not only central to advancing scientific understanding but 45 are also increasingly used to support infrastructure planning, disaster preparedness, and policy development (Akbar et al., 2013; Jam-Jalloh et al., 2024; Young et al., 2025; Zhang et al., 2025). At the same time, the rapid adoption of machine learning in atmospheric science has amplified data-handling demands. WRF outputs are now frequently harvested as training datasets, yet their size and temporal density often exceed the memory capacity of modern GPUs, creating bottlenecks in preprocessing and model optimization (Abdulla et al., 2022; Waqas et al., 2025). These combined pressures highlight the urgent need for effective 50 data compression techniques that can substantially reduce data storage requirements.

A variety of compression techniques, which are applicable for atmospheric model archives, have been developed to reduce the storage demands. Lossless algorithms such as gzip and bzip2 preserve bitwise reproducibility but generally achieve only modest compression ratio when applied to floating-point geophysical fields (e.g., Poppick et al., 2020). Error-bounded lossy approaches offer higher efficiency by discarding information beyond defined tolerances, expressed in terms of absolute error 55 or significant digits. This enables much greater compression while still controlling numerical accuracy (Baker et al., 2016; Walters and Wong, 2023). Among these methods, precision truncation is particularly appealing for both operational workflows because it is computationally lightweight, straightforward to implement, and easily combined with conventional lossless utilities (Walters and Wong, 2023). Early applications to global climate archives and regional air-quality simulations suggest that substantial reductions in file size can be achieved while retaining fidelity for climatological statistics.



60 Despite these advances, most implementations apply uniform truncation settings across all spatial domains, temporal periods, and variable types. Such a one-size-fits-all strategy may introduce unacceptable biases in regions or seasons where climate signals are highly sensitive to perturbations, while at the same time being unnecessarily conservative in situations where stronger compression could be applied without compromising data fidelity. This limitation underscores the need for adaptive frameworks that adjust precision thresholds according to region, season, and variable type, thereby maximizing storage efficiency while safeguarding the scientific integrity required for both routine meteorological analyses and the assessment of high-impact climate events. The latter are particularly challenging, as climate extremes exhibit pronounced regional and seasonal variability and often exert disproportionate societal and economic consequences (Seneviratne et al., 2021; Davenport et al., 2021). Many extreme-event diagnostics rely on threshold-based metrics, meaning that even subtle numerical perturbations introduced by lossy compression can alter exceedance frequencies, ultimately distorting long-term climate trends or biasing attribution studies. Robust evaluation of compression impacts on extreme climate indicators is therefore essential to ensure that truncation strategies preserve the fidelity of high-impact event characterization.

To address this gap, this study systematically evaluates the effects of precision truncation on both storage efficiency and scientific fidelity in WRF simulations over the contiguous United States (CONUS) domain. We analyze both fundamental meteorological variables and extreme precipitation indices to provide a comprehensive assessment of how truncation strategies influence core climate diagnostics. Beyond domain-averaged metrics, we examine regional and seasonal heterogeneity in error responses, highlighting cases where nonlinear sensitivity to truncation may compromise the robustness of scientific conclusions. The remainder of this paper is organized as follows. Section 2 describes the datasets, model configurations, truncation strategies, and analytical methods. Section 3 presents the results, quantifying storage efficiency and evaluating the performance of different truncation strategies for both meteorological fields and extreme precipitation indices, with particular emphasis on regional and seasonal variability. Based on an error-tolerance criterion, we then introduce region- and season-specific compression guidelines and discuss broader implications and limitations of the findings. Section 4 concludes the paper with a synthesis of key results and practical recommendations for data management in high-resolution climate modeling.

2 Methodology

2.1 Model Configuration

85 The year 2016 was chosen as the simulation period because it recorded the highest number of billion-dollar flood disasters and the second-largest economic loss in the United States since 2000 according to the NOAA National Centers for Environmental Information (NCEI) (NCEI, 2025). These factors make 2016 as a representative benchmark for assessing model performance under different truncation strategies, particularly with respect to extreme precipitation events. The CONUS domain was chosen because it encompasses diverse climatological regions and benefits from dense observational networks,



90 providing a robust basis for model assessment. The evaluation focused on a comprehensive suite of common atmospheric variables and extreme precipitation indices defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Frich et al., 2002; Alexander et al., 2006; Zhang et al., 2011).

Simulations were conducted using the WRF model version 4.4.1 over the CONUS domain at a 12-km horizontal spatial resolution with 35 vertical layers. To improve the realism of the simulated meteorological fields, we employed four-dimensional data assimilation (FDDA) within the WRF framework. The assimilated fields were nudged toward the North American Mesoscale (NAM) analysis, produced operationally by the National Centers for Environmental Prediction (NCEP). The nudging strength was set to minimal values for the upper atmosphere to constrain large-scale fields, and no nudging was applied at the surface to avoid suppressing near-surface variability. Table 1 summarizes the physical parameterization schemes used in the WRF simulations.

100

Table 1. Setup of physical parameterization schemes in WRF simulation.

Physical processes	Scheme	Reference
Microphysics	Morrison 2-moment	Morrison et al. (2009)
Radiation	RRTMG Shortwave and Longwave	Iacono et al. (2008)
Surface Layer	Pleim–Xiu	Pleim (2006)
Land Surface Model	Pleim–Xiu Land Surface Model	Pleim and Gilliam (2009)
Planetary Boundary Layer	Asymmetric Convection Model 2	Pleim (2007)

2.2 Precision Truncation Strategies

Precision truncation was applied at three levels, retaining 5, 4, or 3 significant digits, referring to Walters and Wong (2023). A Fortran-based routine was used to modify each floating-point variable except latitude and longitude fields. Specifically, the algorithm retained 5, 4, or 3 significant digits when values were expressed in Fortran scientific notation with a leading digit of zero. Two dataset categories were considered for the application of truncation: model input fields and model output fields. Truncation was applied in three configurations: (i) input-only, (ii) output-only, and (iii) both input and output simultaneously. After truncation, all datasets were compressed using two widely adopted lossless algorithms, gzip and bzip2, resulting in 15 distinct truncation strategies (Table 2).

110



Table 2. Description of 15 truncation strategies.

Case	Input Precision	Output Precision
WRF_fx5/fx4/fx3	Full precision	Keeping 5/4/3 significant digits
WRF_5	Keeping 5 significant digits	Full precision
WRF_5fx5/5fx4/5fx3	Keeping 5 significant digits	Keeping 5/4/3 significant digits
WRF_4	Keeping 4 significant digits	Full precision
WRF_4fx5/4fx4/4fx3	Keeping 4 significant digits	Keeping 5/4/3 significant digits
WRF_3	Keeping 3 significant digits	Full precision
WRF_3fx5/3fx4/3fx3	Keeping 3 significant digits	Keeping 5/4/3 significant digits

2.3 Observational Datasets

To quantitatively evaluate model performance, we employed observational datasets of surface meteorology and precipitation for the year 2016 within the CONUS domain. Hourly near-surface meteorological variables, including 2-m air temperature, 2-m relative humidity, and 10-m wind speed, were obtained from the National Climatic Data Center (NCDC). To ensure data reliability and spatial representativeness, a two-step quality control procedure was applied. First, only stations with valid records for at least 330 days in 2016 and with less than 5% missing data were retained. Second, to reduce representativeness errors caused by terrain mismatch, only stations with an elevation difference of less than 100 m from the corresponding WRF grid cell were selected. These criteria yielded a total of 1622 stations across CONUS domain (Fig. 1a), providing dense spatial coverage across diverse climatic and physiographic regions. This station network enabled robust validation of WRF-simulated near-surface meteorology at both regional and national scales.

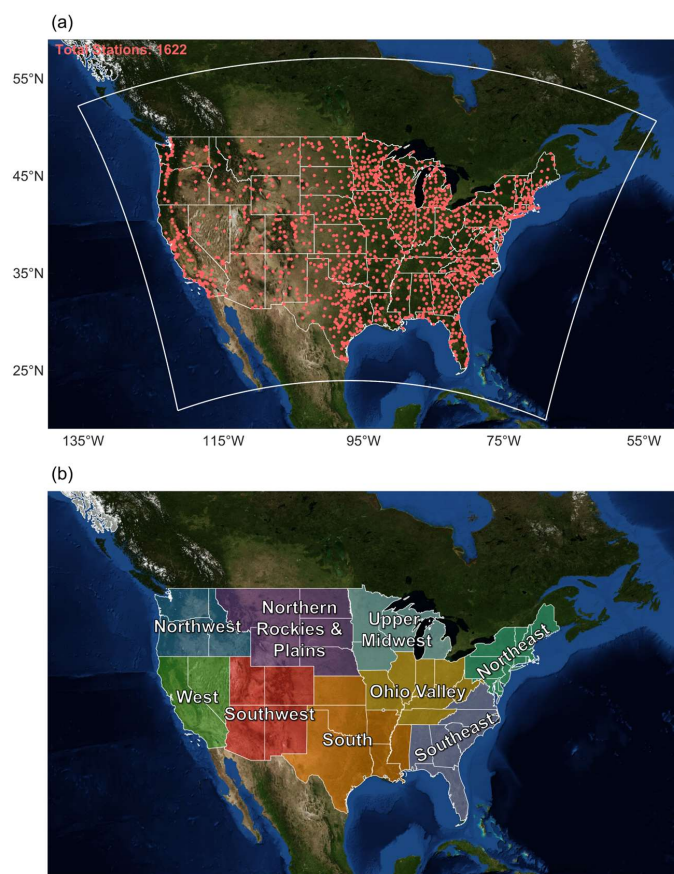


Figure 1: Spatial coverage of observational datasets, WRF model domain, and regional divisions used in this study: (a) locations of
 125 **NCDC surface meteorological stations (red dots) across CONUS domain; the white box outlines the WRF simulation domain; (b)**
delineation of 9 climatologically coherent subregions. Base map imagery from NASA Worldview, Earth Observing System Data and
Information System (EOSDIS).

For precipitation, we used the Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM
 130 (IMERG) Final Precipitation L3 product (Version 07) with a uniform spatial resolution of $0.1^\circ \times 0.1^\circ$. Two temporal resolutions
 were employed. First, daily IMERG data for the period 2001–2015 (Huffman et al., 2023a) were used to establish a
 climatological baseline. From these data, the 95th and 99th percentile precipitation thresholds were calculated and
 subsequently applied for computing extreme precipitation indices. Second, IMERG half-hourly data for 2016 (Huffman et al.,
 2023b) were aggregated to hourly resolution for comparison with observations and further aggregated to daily resolution for
 135 the computation of extreme precipitation indices. Bilinear interpolation was applied to all precipitation datasets to project them
 onto the WRF model grid, thereby ensuring consistency in spatial representation between observations and model results.



2.4 Evaluation Metrics, Climate Regional Subdivision, and Extreme Precipitation Indices

Model fidelity under each truncation strategy was evaluated using three standard statistical metrics: the Root Mean Square Error (RMSE), the Pearson correlation coefficient (R), and the Normalized Mean Bias (NMB). These quantify, respectively, the overall magnitude of deviations, the strength of linear association, and the normalized systematic bias between compressed-simulation outputs and their reference counterparts. They are defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2} \quad (1)$$

$$R = \frac{\sum_{i=1}^N (M_i - \bar{M})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (M_i - \bar{M})^2 \sum_{i=1}^N (O_i - \bar{O})^2}} \quad (2)$$

$$\text{NMB} = \frac{\sum_{i=1}^N (M_i - O_i)}{\sum_{i=1}^N O_i} \times 100\% \quad (3)$$

Here M_i and O_i denote the evaluated values (precision-truncated cases) and the reference values (observations or the baseline WRF simulation without any data modification, WRF_bl), respectively, and N is the number of matched data pairs. To ensure spatial comparability, model outputs were matched to reference datasets using the nearest-neighbor method for station-based surface meteorological variables.

To capture spatial heterogeneity in truncation-induced errors, the CONUS domain was subdivided into 9 climatologically coherent regions following Karl and Koss (1984): Northwest, West, Northern Rockies & Plains, Southwest, Upper Midwest, Ohio Valley, South, Northeast, and Southeast (Fig. 1b).

In addition, to assess the impacts on extreme precipitation, 8 precipitation indices recommended by the ETCCDI (Frich et al., 2002; Nastos et al., 2013; Ozer and Mahamoud, 2013) were computed from both baseline and truncated-simulation precipitation fields. These indices encompass multiple dimensions of precipitation, including intensity, frequency, and persistence (Table 3). Percentile-based thresholds (e.g. R95p_days and R99p_days) were derived from the 2001–2015 daily GPM IMERG baseline dataset, ensuring consistency in threshold definitions across all evaluations. The truncated WRF precipitation fields for 2016 were then compared against this baseline to quantify truncation-induced biases in extreme-event diagnostics.



Table 3. Definitions of Extreme Precipitation Indices

Name	Definition	Units
R95p_days	Number of days per year with daily precipitation exceeds the 95th percentile of wet-day amounts (≥ 1 mm), thresholds derived from the 2001–2015 baseline period.	days
R99p_days	Same as R95p_days, but for the 99th percentile threshold.	days
Rx1_day	Maximum 1-day precipitation total in a year.	mm
Rx5_day	Maximum total precipitation accumulated over any consecutive 5-day period	mm
R10mm_days	Annual count of days with daily precipitation ≥ 10 mm.	days
PRCPTOT	Total annual precipitation from wet days.	mm
wet_days	Annual count of wet days (≥ 1 mm).	days
SDII	Simple Daily Intensity Index, calculated as PRCPTOT divided by wet_days.	mm day ⁻¹

160 3 Result

3.1. Compression Efficiency Analysis

Figure 2 presents the relative compression ratios achieved by applying gzip and bzip2 to the 15 truncation strategies, in which input and/or output data were truncated to 5, 4, or 3 significant digits. The baseline dataset totals 2991.3 GB, consisting of 837.0 GB of input data and 2154.3 GB of output.

165 For input data (Fig. 2a), compression efficiency improves progressively with stronger truncation. Under gzip, the relative compression ratio decreases from 69.0% at 5 significant digits to 34.9% at 3 significant digits, while bzip2 achieves a reduction from 52.4% to 18.5%. For output data (Fig. 2b), the pattern is similar but the relative gains are larger: under gzip, the ratio decreases from 84% to 44%, and under bzip2 from 64% to 25%. This indicates that while both input and output data benefit from truncation, output data exhibit a stronger relative gain when precision is reduced

170 from 5 to 3 significant digits. In combined input–output truncation strategies, the compression of output data is determined almost entirely by the precision applied to the output data, with input truncation exerting little additional influence. For the total dataset (Fig. 2c), gzip reduces the relative compression ratio from 91.8% in WRF_5 to 41.6% in WRF_3fx3, corresponding to a 58.4% reduction. Bzip2 achieves even stronger gains, decreasing from 86.5% to 23.4% (a 76.6% reduction). On average, bzip2 performs 15–30 percentage points better than gzip, with the advantage widening

175 under higher truncation.

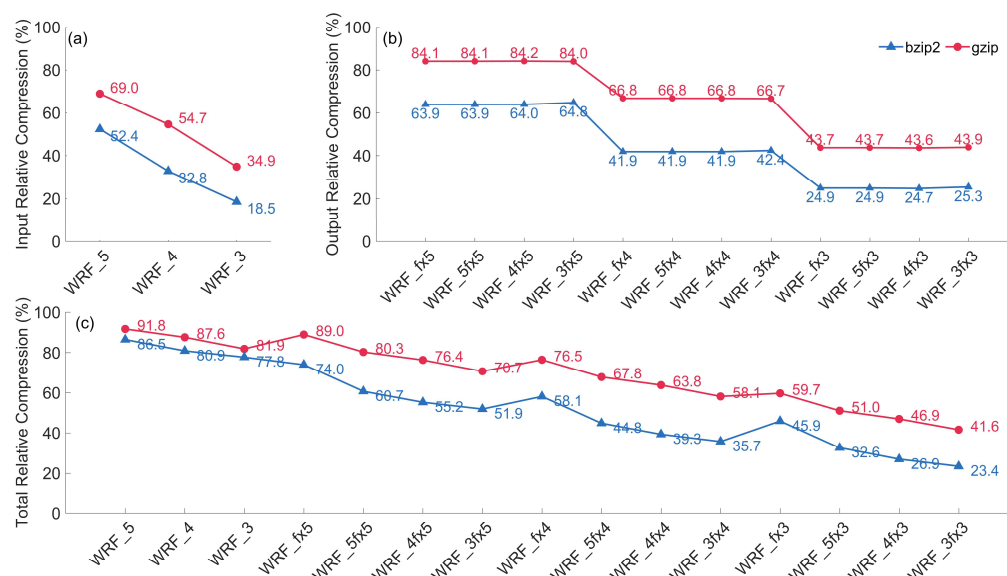


Figure 2: Relative compression performance under different truncation strategies on (a) input data, (b) output data, (c) the combined datasets. Lower percentages indicate greater compression efficiency.

180

These results highlight the significant potential of precision truncation as a preprocessing step for standard lossless compression. Bzip2 renders much better relative compression rate with respect to truncation amount, making it a favorable choice for high-resolution regional climate modeling frameworks such as WRF. However, while the storage benefits are substantial, the scientific consequences of aggressive truncation remain uncertain. In the following section, we evaluate how these truncation strategies affect the fidelity of key meteorological fields, with particular attention to variable-specific and nonlinear sensitivities.

185

3.2. Effects on Meteorological Fields

While gains in compression efficiency are important, they must be weighed against potential degradation in scientific fidelity. Figure 3 presents the relative changes in RMSE and correlation coefficient R for wind speed, temperature, relative humidity, and precipitation. For precipitation, values are derived from spatially averaged time series before calculating temporal statistics; whereas for meteorological variables, statistics are computed at each station-grid pair and then averaged spatially.

190

Wind speed exhibits the strongest response to input truncation, with larger and nonlinear deviations compared to output truncation, suggesting that wind dynamics are sensitive to perturbations in forcing and initial conditions. However, the relative magnitude of truncation-induced errors in wind speed is the smallest among all variables, indicating that despite its sensitivity

195



to input perturbations, the overall bias introduced is limited in scale. Temperature exhibits different behavior: truncating output to 3 significant digits produces substantially larger changes than other strategies, with RMSE increases of about 1.1%. This indicates that thermal variability is more vulnerable to precision loss when applied directly to prognostic outputs. Relative humidity shows a similar response, with marked increases in RMSE and decreases in R under 3 significant digits output truncation, underscoring the high sensitivity of moisture fields to aggressive precision reduction in outputs. Precipitation demonstrates a distinct pattern compared with the other variables. Input truncation at 3 significant digits leads to the largest deviations, while output-only truncation results in the smallest changes. This reflects the diagnostic nature of precipitation, which is strongly modulated by upstream dynamical and thermodynamical input fields but remains relatively robust to precision loss on the output side.

Overall, these results demonstrate that truncation impacts are variable-dependent. Dynamical fields such as wind speed respond more sensitive to input truncation, thermodynamical fields such as temperature and humidity respond more strongly to output truncation, and precipitation shows mixed sensitivity dominated by perturbations to input data. These contrasts provide motivation for a more detailed analysis of regional and seasonal heterogeneity, as shown in Fig.

4.

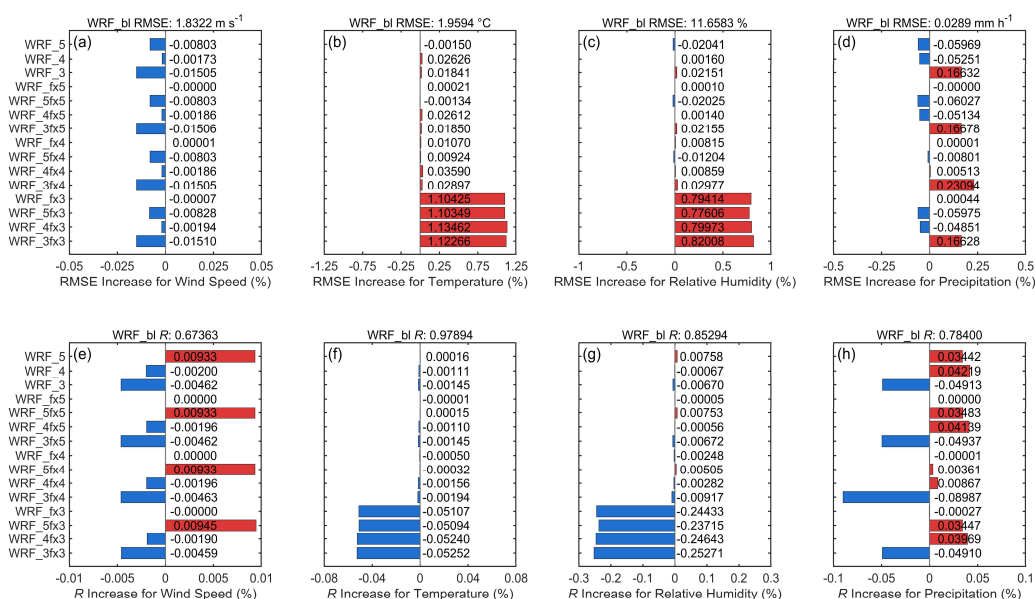


Figure 3: Relative changes with regard to WRF_bi, in RMSE (a–d) and correlation coefficient R (e–h) for wind speed, temperature, relative humidity, and precipitation across all cases.



Figure 4 presents the absolute RMSE increases across nine climate regions and four seasons. When only the output is truncated to four or five significant digits, with inputs left unaltered, RMSE changes are negligible for all variables across regions and seasons. By contrast, truncating the output to three digits introduces systematic effects: wind speed remains largely stable, but temperature and relative humidity show consistent error increases across all regions, and precipitation displays noticeable degradations, particularly during winter and fall.

Wind speed remains the most robust variable, with RMSE increments below 0.1% almost everywhere, except for minor increases in the Southwest and Ohio Valley during summer. Temperature is more vulnerable, with errors reaching about 1.2% under 3-digit truncation, a result partly attributable to the Kelvin unit in WRF outputs, where three digits remove all decimal precision. Humidity follows a similar but less pronounced pattern, with elevated errors that peak near 1.6% in the Upper Midwest during winter and fall. Precipitation emerges as the most heterogeneous field: RMSE increases are largest in fall, followed by summer, and smaller in winter and spring, with regional contributions shifting seasonally from the Upper Midwest, West, Southwest, and Northern Rockies & Plains in summer to the West, Southeast, Upper Midwest, and South in fall. Compared with these controlled and predictable output responses, input truncation exerts a less systematic influence, producing non-monotonic and regionally heterogeneous biases, particularly evident in wind speed and precipitation. Together, these results emphasize that moderate output truncation (5 and 4 significant digits) preserves fidelity effectively, whereas aggressive output truncation and input truncation both carry risks of introducing systematic or spatially variable biases that can compromise model reliability.

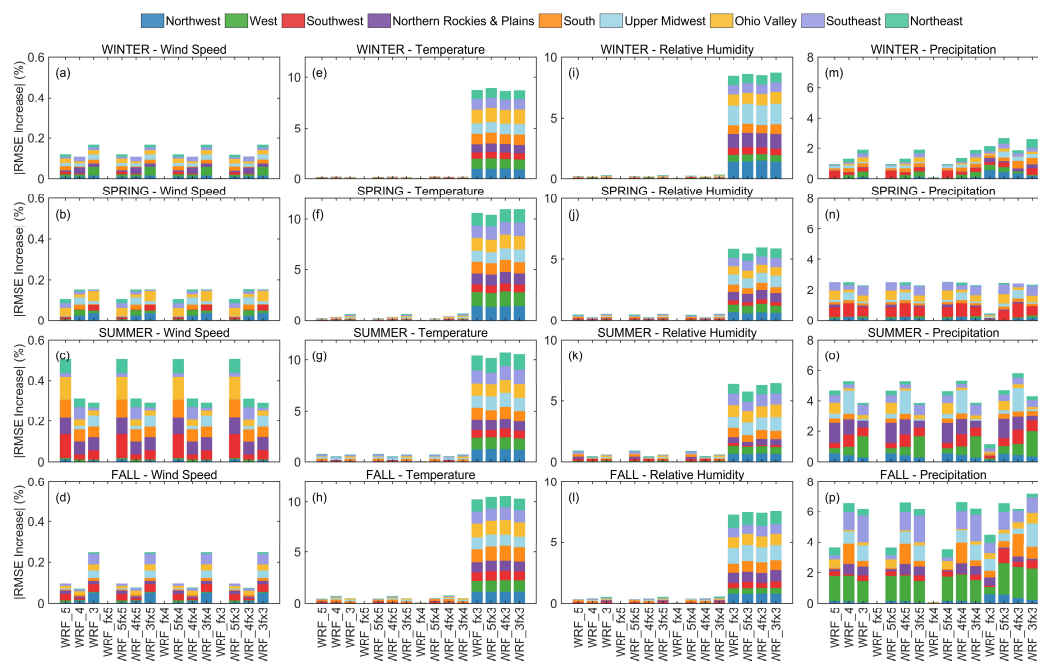




Figure 4: Seasonal and regional absolute RMSE increases for meteorological variables across 15 WRF truncation strategies. Stacked bars show RMSE differences relative to the WRF_bl, aggregated over nine climate regions. The total bar heights do not represent aggregated RMSE across regions but only serve to display regional contributions. Results are presented separately for (a–d) wind speed, (e–h) temperature, (i–l) relative humidity, and (m–p) precipitation.

3.3. Sensitivity of Extreme Precipitation Indices to Precision Truncation

In this section, we evaluate the sensitivity of ETCCDI extreme precipitation indices, which are critical for understanding climate variability and societal risks. The analysis focus on how these indices respond to different precision truncation strategies and whether truncation compromises their reliability. For clarity, the indices are grouped into two categories. The first category consists of percentile-based and maximum-based extreme precipitation indices (R95p_days, R99p_days, Rx1_day, Rx5_day), which emphasize the behavior of the upper tail of the daily precipitation distribution and the magnitude of rare extreme events. The second category includes precipitation frequency and accumulation indices (R10mm_days, PRCPTOT, wet_days, SDII), which describe total precipitation amounts, mean intensity, and the occurrence of moderate wet events. In the main text, we focus on two representative diagnostics, R99p_days and SDII. Results for the other six indices are provided in the Supplementary Information.

For the percentile-threshold index R99p_days, biases are generally modest in winter, spring, and fall (typically within $\pm 3\%$ across regions and strategies), but increase sharply in summer (Fig. 5a–d). The largest positive deviations arise when input precision is truncated to 3 significant digits, either alone or in combination with output truncation, with NMB reaching 7–8% in the Northeast. Output-only truncation produces a relatively small and near-linear response from fx5 to fx3, with domain-wide shifts remaining below 2%. Once input truncation is introduced, however, the response becomes distinctly nonlinear, demonstrating that small distributional perturbations near the 99th-percentile threshold can be disproportionately amplified. These results underscore the pronounced sensitivity of percentile-based exceedance metrics to input-related perturbations.

By contrast, SDII shows a different pattern. Negative biases dominate across most regions and seasons, and these biases are primarily governed by output precision (Fig. 5e–h). When only output is truncated, fx3 produces widespread and substantial negative deviations, while fx4 introduces consistent negative biases, ranging from -1% to -3%. In comparison, fx5 maintains an NMB close to zero, with minimal overall impact. Input-only truncation generates relatively small, mixed positive and negative shifts, while combined input–output truncation reveals that output precision plays a dominant role in determining the final bias. Overall, more aggressive truncation strategies lead to larger SDII biases, following a relatively linear response pattern.

The contrasting behaviors of these two indices reveals two distinct sensitivity regimes. Percentile-threshold



265 exceedance metrics such as R99p_days are dominated by input precision and can deviate in either direction, whereas
 mean-intensity metrics like SDII respond more systematically and predominantly negatively to output truncation.

This distinction becomes clearer when the full set of 8 extreme precipitation indices is examined through regional
 box plots (Fig. 6). For the percentile- and maximum-based group (R95p_days, R99p_days, Rx1day, Rx5day), output-
 only truncation (fx5–fx3) keeps tight interquartile ranges with medians close to zero. However, once input truncation is
 270 introduced, the distributions become more dispersed and a large number of outliers appear, indicating pronounced
 nonlinear behavior driven by threshold crossing. In contrast, the frequency and accumulation group (R10mm_days,
 PRCPTOT, wet_days, SDII) is governed primarily by output precision. When outputs are truncated to 3 significant digits,
 R10mm_days and wet_days shift strongly positive, PRCPTOT shows modest positive shifts, and SDII exhibits consistent
 negative deviations. Input truncation, by comparison, has relatively little influence on these indices.

275

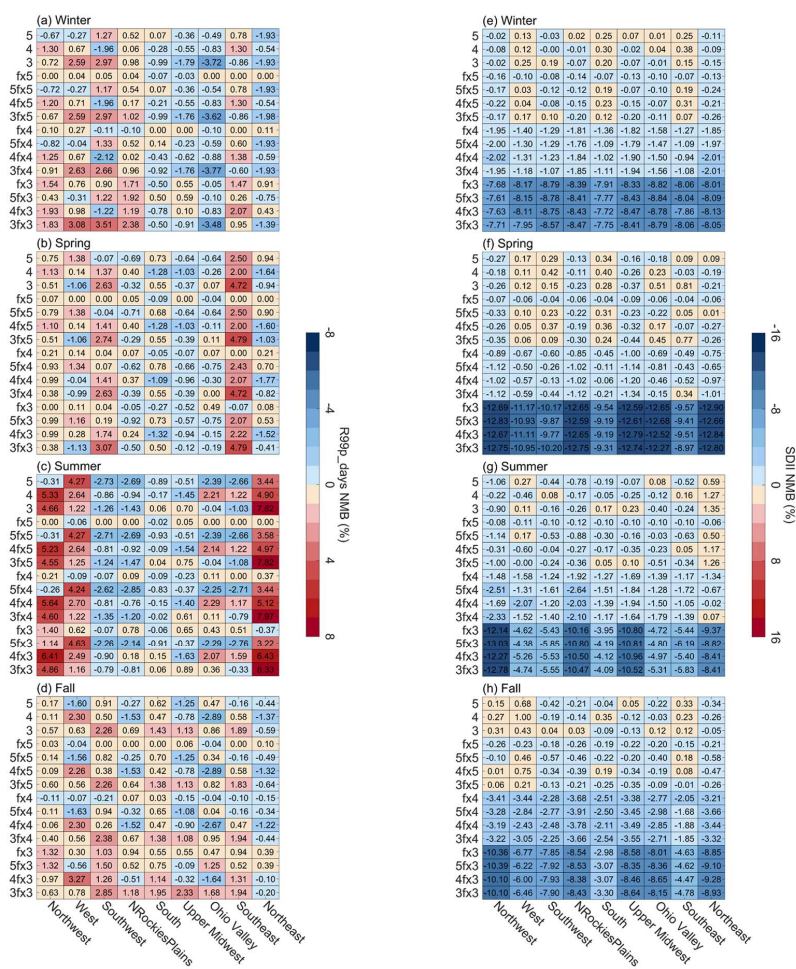


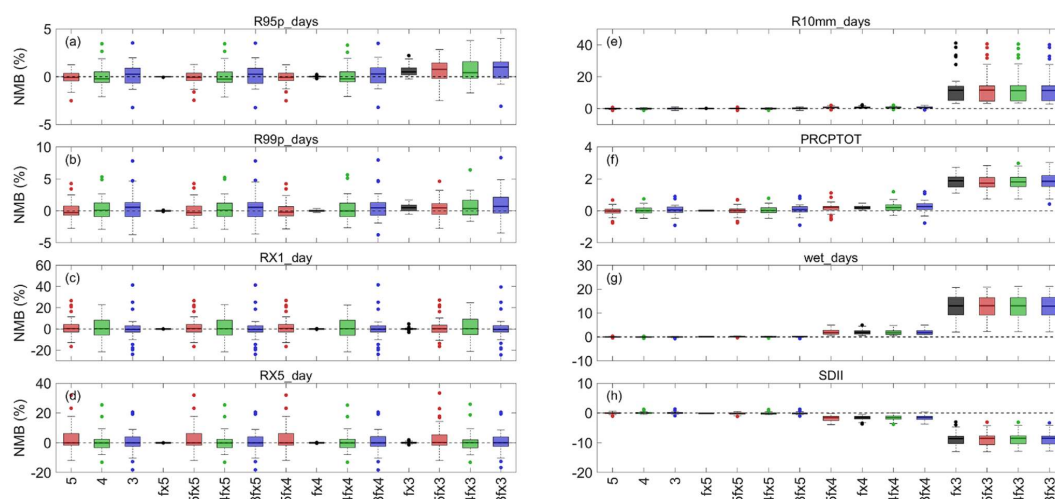
Figure 5: Seasonal and regional NMB of R99p_days (a–d) and SDII (e–h) relative to the WRF_bl, shown for different WRF



truncation strategies across 9 climate regions and 4 seasons.

280 A synthesis of central tendencies using the regional median $|NMB|$ (Fig. 7) reinforces the earlier findings. Percentile-
 and maximum-based indices (Fig. 7a) maintain low medians under output-only truncation but rise sharply once input
 precision is reduced to three digits, indicating that departures are primarily triggered by input perturbations. In contrast,
 frequency, accumulation, and mean-intensity indices (Fig. 7b) exhibit a systematic monotonic response governed by
 output precision, with marked deterioration at the 3 significant digits.

285 Together with Figs. 5–6, these results highlight two distinct regimes: nonlinear, threshold-driven variability in
 percentile-based indices versus systematic, output-driven biases in frequency and intensity indices. Recognizing this
 difference is essential for designing truncation strategies that safeguard both storage efficiency and scientific fidelity.



290 **Figure 6: Box plots showing the distribution of NMB values for 8 extreme precipitation indices across 15 truncation strategies and 9 climate regions. Each box spans the interquartile range (IQR; 25th–75th percentiles), with the median shown as a horizontal line. Whiskers extend to $1.5 \times IQR$, and outliers are plotted as colored circles corresponding to the truncation groups based on input data. Dashed horizontal lines indicate zero bias.**

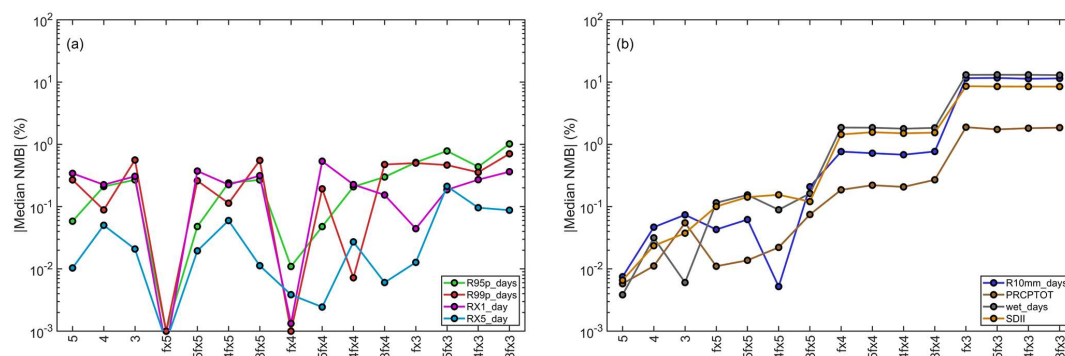


Figure 7: Median NMB of extreme precipitation indices across 15 truncation strategies and 9 climate regions relative to the WRF_bl. (a) shows the indices for extreme precipitation events (R95p_days, R99p_days, Rx1_day, Rx5_day), while panel (b) displays the indices for precipitation frequency and intensity (R10mm_days, PRCPTOT, wet_days, SDII).

3.4. Optimal Truncation Strategy

Building on the sensitivity analyses presented in Sections 3.2 and 3.3, we now translate these findings into practical truncation strategies that balance storage optimization with the preservation of scientific fidelity. The objective is not simply to minimize file size, but to do so in a way that safeguards critical features of both routine meteorological variables and extreme precipitation indices. As shown earlier, the impacts of precision truncation vary substantially across variable types, regions, and seasons, with the clearest contrasts emerging for extreme precipitation metrics that play a central role in assessing climate extremes and their associated risks.

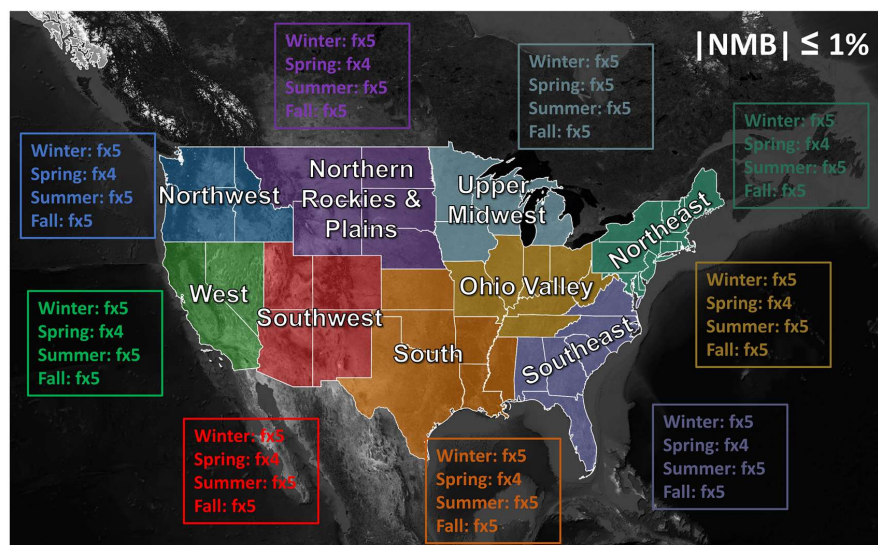
Although input truncation offers attractive potential for stronger compression, our analysis shows that it consistently produces nonlinear and regionally heterogeneous biases, especially for percentile-based indices. Reducing input precision to 3 significant digits amplifies small perturbations near extreme thresholds, causing substantial distortions and undermining the reliability of extreme-event characterization. For instance, even minor shifts in daily precipitation totals near the 99th percentile can substantially alter how extremes are quantified, eroding confidence in the diagnostic signals. Based on these results, we focus on output-only truncation as the more reliable strategy, since it preserves the fidelity of all analyzed variables more effectively.

As noted above, the impacts of truncation differ not only with specific variable type but also by season and region, making adaptation essential. To address this, we adopt a flexible error-tolerance framework that applies a threshold of $|NMB| \leq 1\%$ as the benchmark for acceptable fidelity. Any truncation strategy must ensure that this criterion is simultaneously satisfied across all indices, regions, and seasons. Within this framework, the truncation level selected is the one that maximizes compression efficiency while keeping all metrics within the tolerance bounds. In practice, this



design allows the threshold to be adjusted if needed, enabling different tolerance levels for specific variables, regions,
 320 and/or seasons. We implemented this framework in a Fortran-based routine that systematically applies truncation
 strategies once error analyses are completed. Rather than prescribing a universal standard, we recommend that users
 conduct dataset-specific assessments and then select truncation levels that best meet their scientific objectives. It should
 be emphasized that, for the sake of scientific consistency, simulation results generated under different input conditions
 cannot be merged directly. Input truncation alters the initial and boundary conditions of the model, and thus must remain
 325 fixed across all regions and seasons within a single experiment. By contrast, output truncation is applied after the
 simulation has been completed, which allows truncation levels to vary flexibly across regions and seasons without
 affecting the underlying physical consistency of the run. This distinction ensures that the compression strategy preserves
 methodological rigor while offering practical adaptability for large-scale climate datasets.

Figure 8 summarizes the resulting optimal truncation strategies by region and season, providing a clear visualization
 330 of how the error-tolerance criterion translates into tailored recommendations. Output-only truncation is consistently the
 most robust choice. Among these, fx5 offers the most conservative option, ensuring the highest fidelity though with
 smaller compression gains. In many regions and seasons, however, fx4 can be applied to achieve stronger storage
 reductions while still preserving scientific reliability. These results underscore the value of adaptive strategies that adjust
 truncation according to regional and seasonal sensitivities, offering a balanced pathway to maximize storage efficiency
 335 without compromising the robustness of both meteorological and extreme-event diagnostics. Looking forward, we
 encourage further error analyses of precision truncation for additional variables, which will be essential for strengthening
 data management strategies in high-resolution climate modeling.





340 **Figure 8: Optimal truncation strategies across 9 climate regions and 4 seasons for at a tolerance of $|NMB| \leq 1\%$. Base map imagery from NASA Worldview, Earth Observing System Data and Information System (EOSDIS).**

4. Conclusions

This study provides a comprehensive evaluation of how significant-digit truncation influences both storage efficiency and scientific fidelity in regional atmospheric modeling, with a focus on routine meteorological variables and
 345 ETCCDI extreme precipitation indices. The results demonstrate that different diagnostics respond in fundamentally different ways: conventional atmospheric fields can be well preserved when the output is truncated to 5 or 4 significant digits, whereas percentile- and maximum-based precipitation indices exhibit nonlinear sensitivities to input truncation, and frequency- and intensity-based indices follow systematic, monotonic responses to output precision.

These contrasting behaviors argue strongly against uniform, one-size-fits-all solutions. Instead, the findings support
 350 the development of an adaptive truncation framework that explicitly integrates error-tolerance thresholds with regional and seasonal sensitivities. Within the scope of this work, output-only truncation emerged as the most robust option, with 4 significant digits serving as a safe lower bound and 5 significant digits advisable in regions or seasons dominated by extremes. However, the core contribution of this study lies not in prescribing a single “best” truncation level, but in advancing a methodological framework that can be flexibly applied across variables, domains, and research objectives.

355 By requiring all indices to remain within predefined error tolerances, this framework provides a precise way to evaluate trade-offs between compression efficiency and scientific robustness. This approach is inherently extensible: thresholds can be adapted to different scientific applications, and truncation levels can be tuned by variable, region, and/or season. Overall, this framework supports context-aware decisions that preserve the integrity of both routine meteorological diagnostics and extreme-event analyses while delivering substantial storage savings.

360 Looking ahead, the methodology presented here offers a foundation for future work on climate data storage reduction. Incorporating advances in lossless compression algorithms, and applying it across diverse modeling systems will further enhance its utility. Ultimately, this study demonstrates how precision truncation can be integrated into a systematic, error-aware workflow for climate data management, enabling more efficient use of storage resources without compromising the fidelity of climate diagnostics.

365 Code availability

A tool for formatting data to a user-specified number of significant digits is available for download on Zenodo at <https://doi.org/10.5281/zenodo.17156737> (Wong and Wu, 2025). The download includes the full source code, a run script, and detailed usage instructions.



Data availability

370 The processed analytic datasets that support the main findings of this study are openly available on Zenodo at
<https://doi.org/10.5281/zenodo.17139028> (Wu and Wong, 2025).

Author contributions

Conceptualization: DCW and SW; Data curation: SW and DCW; Formal analysis: SW, DCW, JW, and CL; Funding acquisition:
CL; Investigation: SW, YJ, and JL; Methodology: SW, DCW, JW, and CL; Visualization: SW and JL; Writing – original draft:
375 SW; Writing – review & editing: SW, DCW, YJ, JW, CL, and JL.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the views or the policies of the
380 U.S. Environmental Protection Agency (EPA).

Acknowledgements

We acknowledge the High-Performance Computing Center of Nanjing University of Information Science & Technology for
their support of this work. Surface meteorological data (wind speed, temperature, and humidity) were sourced from the NOAA
National Centers for Environmental Information (NCEI) Global Hourly dataset, accessible
385 at: <https://www.ncei.noaa.gov/data/global-hourly/archive/csv/> (last access: 28 September 2025). Precipitation data were
obtained from the GPM IMERG Final Precipitation L3 product (Version 07), with daily and hourly data available
via <https://doi.org/10.5067/GPM/IMERGDF/DAY/07> and <https://doi.org/10.5067/GPM/IMERG/3B-HH/07>, respectively.

The authors thank these agencies for providing these publicly available datasets. We acknowledge the use of imagery from
NASA Worldview, operated by NASA's Earth Observing System Data and Information System (EOSDIS). The maps were
390 obtained from <https://worldview.earthdata.nasa.gov/> (last access: 28 September 2025).



Financial support:

This work has been supported by the National Natural Science Foundation of China (Grant no. 42422505).

Reference

- Abdulla, N., Demirci, M., and Ozdemir, S.: Design and evaluation of adaptive deep learning models for weather forecasting, *Engineering Applications of Artificial Intelligence*, 116, 105440, <https://doi.org/10.1016/j.engappai.2022.105440>, 2022.
- Akbar, M., Aliabadi, S., Patel, R., and Watts, M.: A fully automated and integrated multi-scale forecasting scheme for emergency preparedness, *Environ. Model. Software*, 39, 24-38, <https://doi.org/10.1016/j.envsoft.2011.12.006>, 2013.
- Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., and Rahimzadeh, F.: Global observed changes in daily climate extremes of temperature and precipitation, *J Geophys Res-Atmos*, 111, <https://doi.org/10.1029/2005JD006290>, 2006.
- Baker, A. H., Hammerling, D. M., Mickelson, S. A., Xu, H., Stolpe, M. B., Naveau, P., Sanderson, B., Ebert-Uphoff, I., Samarasinghe, S., and De Simone, F.: Evaluating lossy data compression on climate simulation data within a large ensemble, *Geoscientific Model Development*, 9, 4381-4403, <https://doi.org/10.5194/gmd-9-4381-2016>, 2016.
- Davenport, F. V., Burke, M., and Diffenbaugh, N. S.: Contribution of historical precipitation change to US flood damages, *Proceedings of the National Academy of Sciences*, 118, e2017524118, <https://doi.org/10.1073/pnas.2017524118>, 2021.
- Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Tank, A. K., and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, *Clim. Res.*, 19, 193-212, <https://doi.org/10.3354/cr019193>, 2002.
- Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Tan, J.: GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V07 [dataset], <https://doi.org/10.5067/GPM/IMERGDF/DAY/07>, 2023a.
- Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Tan, J.: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V07 [dataset], <https://doi.org/10.5067/GPM/IMERG/3B-HH/07>, 2023b.
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.-Atmos.*, 113, 8, <https://doi.org/10.1029/2008jd009944>, 2008.
- Jam-Jalloh, Umar, S., Liu, J., Wang, Y., and Liu, Y.: Coupling WRF with HEC-HMS and WRF-Hydro for flood forecasting in typical mountainous catchments of northern China, *Natural Hazards and Earth System Sciences*, 24, 3155-3172, <https://doi.org/10.5194/nhess-24-3155-2024>, 2024.
- Karl, T. R. and Koss, W. J.: Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983, *Historical Climatology Series 4-3*, National Climatic Data Center, Asheville, NC, 38, 1984.



- Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes, *Mon. Weather Rev.*, 137, 991-1007, <https://doi.org/10.1175/2008MWR2556.1>, 2009.
- Nastos, P., Kapsomenakis, J., and Douvis, K.: Analysis of precipitation extremes based on satellite and high-resolution gridded data set over Mediterranean basin, *Atmospheric Research*, 131, 46-59, <https://doi.org/10.1016/j.atmosres.2013.04.009>, 2013.
- NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2025): <https://www.ncei.noaa.gov/access/billions/>, <https://doi.org/10.25921/stkw-7w73>, last access: 2025-09-06.
- Overpeck, J. T., Meehl, G. A., Bony, S., and Easterling, D. R.: Climate data challenges in the 21st century, *Science*, 331, 700-702, <https://doi.org/10.1126/science.1197869>, 2011.
- Ozer, P. and Mahamoud, A.: Recent extreme precipitation and temperature changes in Djibouti City (1966–2011), *Journal of Climatology*, 2013, 928501, <https://doi.org/10.1155/2013/928501>, 2013.
- Pleim, J. E.: A simple, efficient solution of flux–profile relationships in the atmospheric surface layer, *J. Appl. Meteorol. Climatol.*, 45, 341-347, <https://doi.org/10.1175/JAM2339.1>, 2006.
- Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing, *J. Appl. Meteorol. Climatol.*, 46, 1383-1395, <https://doi.org/10.1175/JAM2539.1>, 2007.
- Pleim, J. E. and Gilliam, R.: An indirect data assimilation scheme for deep soil temperature in the Pleim–Xiu land surface model, *J. Appl. Meteorol. Climatol.*, 48, 1362-1376, <https://doi.org/10.1175/2009JAMC2053.1>, 2009.
- Poppick, A., Nardi, J., Feldman, N., Baker, A. H., Pinard, A., and Hammerling, D. M.: A statistical analysis of lossily compressed climate model data, *Computers & Geosciences*, 145, 104599, <https://doi.org/10.1016/j.cageo.2020.104599>, 2020.
- Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., and Peckham, S. E.: The weather research and forecasting model: Overview, system efforts, and future directions, *Bull. Am. Meteorol. Soc.*, 98, 1717-1737, <https://doi.org/10.1175/BAMS-D-15-00308.1>, 2017.
- Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., and Feser, F.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, *Reviews of geophysics*, 53, 323-361, <https://doi.org/10.1002/2014RG000475>, 2015.
- Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and Climate Extreme Events in a Changing Climate, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy,



- E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1513–1766, <https://doi.org/10.1017/9781009157896.013>, 2021.
- 455 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., and Barker, D. M.: A description of the advanced research WRF version 4, National Center for Atmospheric Research, Boulder, CO, USANCAR tech. note ncar/tn-556+ str, <https://doi.org/10.5065/1dfh-6p97>, 2019.
- Walters, M. S. and Wong, D. C.: The impact of altering emission data precision on compression efficiency and accuracy of simulations of the community multiscale air quality model, *Geoscientific model development*, 16, 1179–1190, <https://doi.org/10.5194/gmd-16-1179-2023>, 2023.
- 460 Waqas, M., Humphries, U. W., Chueasa, B., and Wangwongchai, A.: Artificial intelligence and numerical weather prediction models: A technical survey, *Natural Hazards Research*, 5, 306–320, <https://doi.org/10.1016/j.nhres.2024.11.004>, 2025.
- Wong, D. C. and Wu, S.: Truncation tool for the paper 'A Flexible Framework for Precision Truncation and Lossless Compression in WRF Simulations: Method and Application over the United States' [dataset], <https://doi.org/10.5281/zenodo.17156737>, 2025.
- 465 Wu, S. and Wong, D. C.: Dataset for the paper 'A Flexible Framework for Precision Truncation and Lossless Compression in WRF Simulations: Method and Application over the United States' [dataset], <https://doi.org/10.5281/zenodo.17139028>, 2025.
- Young, A., Bhattacharya, B., Daniëls, E., and Zevenbergen, C.: Integrating WRF forecasts at different scales for pluvial flood forecasting using a rainfall threshold approach and a real-time flood model, *Journal of Hydrology*, 656, 132891, <https://doi.org/10.1016/j.jhydrol.2025.132891>, 2025.
- 470 Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., and Zwiers, F. W.: Indices for monitoring changes in extremes based on daily temperature and precipitation data, *Wiley Interdisciplinary Reviews: Climate Change*, 2, 851–870, <https://doi.org/10.1002/wcc.147>, 2011.
- 475 Zhang, Y., Deng, C., Xu, W., Zhuang, Y., Jiang, L., Jiang, C., Guan, X., Wei, J., Ma, M., and Chen, Y.: Long-term variability of extreme precipitation with WRF model at a complex terrain River Basin, *Scientific Reports*, 15, 156, <https://doi.org/10.1038/s41598-024-84076-x>, 2025.