Response to the comments of Reviewer 1

Thank you for reviewing our manuscript thoroughly and providing constructive comments and valuable suggestions. Please find our point-by-point responses listed below. The reviewer's comments are in black *Italics* followed by our responses (in blue). The red text within the quotation marks is the revisions in the manuscript, while the black text is the unmodified content.

This manuscript proposes to reduced storage size of weather and climate data without compromising scientific integrity, and investigates various precision truncation strategies (combined with lossless compression) with the data from the Weather Research and Forecasting (WRF) simulations. The authors choose 2016 for the WRF simulation period, with 4-D data assimilation. Results were compared with hourly 2m air temperature and humidity and 10m wind speed, and hourly precipitation.
Metrics of relative data compression are percentage of original data when further compressed using bzip2 or gzip.

Metrics on errors due to data compression consist of RMSE of the encoded values vs. reference values, Pearson Correlation R, and Normalized Mean Bias NMB. Additional metrics for assessing impacts on extreme precipitation include number of days exceeding the 95% or 99% percentile of wet days, the maximum 1-day or 5-day precipitation total, annual count of days with daily precipitation over 10mm, count and total precipitation in wet days over a year, simple daily intensity index derived from that.

The paper is generally well written. The results are encouraging but not new (see my point below on the literature review), and the authors are not providing final compression results for the optimal strategy; it is thus unclear why the reader should actually care about doing this extra work of data compression. The paper would strongly benefit from being improved for clarity.

Thank you for your constructive feedback. Following your suggestion, we carefully revised the manuscript to improve clarity and to better articulate the motivation, methodological framework, and practical implications of our work. Our detailed, point-by-point responses outlining the exact modifications made to address each of your specific comments and suggestions are provided below.

Overall, we substantially revised the manuscript in several aspects. First, the research background and literature context have been expanded to better position our work within existing studies on atmospheric data compression. Second, the scientific motivation and practical relevance of evaluating precision reduction within a WRF modeling workflow are now stated more explicitly in the Introduction section and Abstract. Third, the presentation of compression performance has been enhanced to more explicitly demonstrate the quantitative storage gains achieved across different precision-reduction configurations. Fourth, the methodological framework and evaluation analyses have been refined and expanded, including additional diagnostics of local deviation, structural fidelity and downstream scientific impacts. These revisions collectively improve the clarity of the manuscript.

Finally, to improve clarity and avoid potential confusion, we standardized the terminology throughout the manuscript. The term "precision truncation" used in the previous version has been replaced

with “precision reduction” or “decimal significant-digit rounding,” and expressions such as “truncation strategies” have been replaced by “precision reduction configurations.”

The fundamental limitation of the paper is that it is not properly situated in the comprehensive literature of data truncation and data compression, beyond three references: Baker et al (2016), Poppick et al (2020, lossy), Walters and Wong (2023).
The following work extensively investigated truncation strategies:
M Klöwer, M Razinger, JJ Dominguez, PD D uben, TN Palmer, Compressing atmospheric data into its real information content. Nat. Comput. Sci. 1, 713–724 (2021).
Moreover, several works have explored neural lossy compression:
L Huang, T Hoefler, Compressing multidimensional weather and climate data into neural networks. ICLR (2023).
T Han, S Guo, W Xu, L Bai, et al., Cra5: Extreme compression of era5 for portable global climate and weather research via an efficient variational transformer. arXiv preprint arXiv:2405.03376 (2024).
P Mirowski, D Warde-Farley, M Rosca, et al., Neural compression of atmospheric states. arXiv preprint arXiv:2407.11666 (2024).

Thank you for pointing out that the original manuscript did not sufficiently situate our study within the broader literature. Following this suggestion, we conducted a substantially expanded literature review and revised the Introduction (Lines 54–80) to more comprehensively position our work within the existing research landscape.

First, we incorporated the reviewer-recommended studies and clarified their relevance to our work. In particular, Klöwer et al. (2021) reframed atmospheric data compression from an information-theoretic perspective, demonstrating that the intrinsic precision requirements of atmospheric variables vary substantially and that a large portion of stored floating-point precision is redundant. This perspective strongly supports the motivation of our study and highlights the importance of variable-dependent precision reduction strategy. In addition, recent work on neural-network-based compression methods (Huang and Hoefler, 2023; Han et al., 2024; Mirowski et al., 2024) was added to place our approach within the broader context of emerging compression technologies.

Second, we expanded the discussion of classical precision reduction techniques and related evaluation frameworks. The revised Introduction now also references additional foundational studies, including Zender (2016) on bit grooming and unbiased quantization, Delaunay et al. (2019) on the Digit Rounding algorithm bridging decimal and binary representation, Silver and Zender (2017) on compression–error trade-offs in geophysical datasets, Baker et al. (2019) on structural fidelity diagnostics, and technical investigations of netCDF/HDF compression performance (Delaunay et al., 2019; Underwood et al., 2022). These revisions substantially broaden the literature context and more clearly distinguish our focus on analyzing the impacts of precision reduction across the entire WRF model workflow, including both input and output components.

Third, inspired by the evaluation framework used in Baker et al. (2019) and related studies, we strengthened the methodological analysis of structural fidelity in our manuscript. Specifically, we

introduced the Structural Similarity Index Measure (SSIM) as a complementary diagnostic to grid-scale numerical deviations (Section 2.4: Lines 256–273), which is presented in the newly added Section 3.3 titled "Maximum Deviations and Structural Fidelity". This section systematically analyzes the relationship between maximum absolute deviations and spatial structural similarity across precision-reduction configurations, providing additional insight into how precision reduction, especially for the input, affects the spatial integrity of simulated meteorological fields. The conclusion is as follows (Lines 623–640, Lines 726–730): Precision reduction applied to time-varying inputs introduces perturbations that interact with nonlinear model dynamics during integration. Although the large-scale atmospheric structures remain morphologically coherent, the resulting spatial phase shifts can substantially affect localized diagnostics (e.g., precipitation and 2-m temperature). Consequently, input precision reduction should generally be avoided when exact spatiotemporal correspondence is required, while it remains a viable approach for prioritizing the preservation of overall statistical distributions.

Lines 54–80:

"A variety of compression techniques applicable to atmospheric model archives have been developed to alleviate the rapidly growing storage demands of numerical simulations. Lossless algorithms alone preserve bitwise reproducibility but generally achieve only a modest compression ratio for floating-point geophysical fields because the high entropy of mantissa bits limits compressibility (e.g., Poppick et al., 2020). To address this limitation, combining precision reduction with lossless compression has emerged as a widely explored strategy for improving storage efficiency while attempting to preserve scientific fidelity. Within this paradigm, many approaches focus on manipulating the least significant bits of floating-point data. Early approaches such as bit shaving (Zender, 2016) reduce precision by zeroing trailing mantissa bits, which can introduce systematic bias. Zender (2016) subsequently introduced bit grooming, a statistically robust quantization approach that alternates bit shaving and bit setting of trailing IEEE mantissa bits, thereby preserving the mean in expectation while improving compressibility. Building on this line of work, Delaunay et al. (2019) introduced the Digit Rounding algorithm, which bridges decimal precision control and binary representation. Compared to bit grooming, Digit Rounding optimizes the allocation of mantissa bits required, thereby improving compression efficiency while retaining controlled numerical precision. In contrast to hybrid methods like Digit Rounding, alternative approaches regulate precision purely within the decimal domain. For example, decimal significant-digit rounding (Walters and Wong, 2023) directly constrains the number of significant digits, yielding a highly transparent and interpretable approach to precision reduction.

Alongside algorithmic developments, some studies have expanded the evaluation of precision reduction beyond simple error statistics toward broader notions of statistical and structural fidelity. Baker et al. (2016) proposed an ensemble-based framework to assess whether compressed datasets remain statistically indistinguishable from internal climate variability, while Baker et al. (2019) emphasized the importance of evaluating structural integrity in compressed climate data. Silver and Zender (2017) quantified the compression–error trade-off for gridded datasets, demonstrating that carefully designed precision reduction can remove substantial false precision with limited impact on scientific conclusions. Complementary technical investigations have examined how data representation and codec behavior

influence achievable compression ratios within netCDF and HDF data formats (Delaunay et al., 2019; Underwood et al., 2022). Recently, Klöwer et al. (2021) reframed atmospheric data compression from an information-theoretic perspective, demonstrating that the intrinsic precision requirements of atmospheric variables vary substantially and that a large fraction of stored floating-point precision is redundant. In parallel, machine-learning-based compression methods have emerged that exploit nonlinear data manifolds to achieve very high compression ratios (Huang and Hoefler, 2023; Han et al., 2024; Mirowski et al., 2024)."

Lines 256–273:

"Second, to evaluate the impacts of precision reduction on both local numerical accuracy and spatial structures, we analyzed point-wise deviations together with spatial similarity metrics. At the grid scale, we computed the absolute grid-scale deviation (AD), defined as the absolute difference between the precision-reduced configurations and the full-precision baseline simulation WRF_bl at each grid point and hourly time step. Because point-wise metrics alone cannot capture changes in the spatial organization of meteorological fields, we additionally employed the Structural Similarity Index Measure (SSIM) (Wang et al., 2004; Baker et al., 2019; Klöwer et al., 2021). This metric is particularly important for evaluating simulations driven by precision-reduced inputs, where small numerical perturbations may propagate through nonlinear dynamics and alter evolving mesoscale structures. Unlike AD, which measures local differences, SSIM quantifies the preservation of large-scale spatial patterns and structural textures. For a full-precision reference spatial window x and the corresponding window y from precision-reduced configurations, the SSIM is calculated as:

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{3}$$

where μ_x and μ_y are the local means, σ_x^2 and σ_y^2 are the local variances, and σ_{xy} is the local covariance between the two fields. The stabilization constants C_1 and C_2 are scaled by the dynamic range of the specific meteorological variable being evaluated to accommodate diverse atmospheric fields. SSIM values range from 0 to 1, with unity indicating perfect structural similarity. Structural similarity was computed using an 11×11 Gaussian weighting window sliding across the domain. Given the 12 km horizontal grid spacing, this corresponds to a spatial footprint of approximately 130 km, enabling assessment of mesoscale structural consistency. To avoid artificial inflation of scores caused by large, structurally uniform dry regions, a standard wet-day mask with a threshold of 0.1 mm h^{-1} was applied prior to the calculation of precipitation SSIM."

Lines 623–640:

"**4.1 Implications of Input Precision Reduction**

Precision reduction applied to time-varying model inputs introduces perturbations that interact with nonlinear atmospheric dynamics during model integration. These perturbations propagate through the simulation and may alter the precise spatial and temporal positioning of weather systems. As demonstrated in Section 3.3, such perturbations can generate large grid-scale maximum AD. However, the simultaneously high SSIM values indicate that the overall morphology of the simulated atmospheric systems remains largely preserved. Consequently, the resulting errors primarily manifest as pronounced maximum AD driven by local spatial or temporal phase shifts, accompanied by a background of

widespread, disordered weak perturbations, rather than a fundamental structural collapse of the simulated fields. Therefore, input precision reduction may be acceptable for applications that prioritize large-scale statistics or aggregated diagnostics, where exact grid-scale correspondence is not required. However, this pathway introduces intrinsic nondeterministic perturbations into the simulation, meaning that the exact location and timing of individual weather systems may shift. For impact-oriented applications requiring strict spatiotemporal consistency, such as event attribution or local hydrological analyses, this loss of deterministic correspondence may be undesirable.

In addition, the chaotic amplification of these input perturbations exhibits strong seasonal dependence. Error propagation peaks during the summer when local land-atmosphere feedbacks and convective processes dominate. Hypothetically, aggressive precision reduction could be applied to input forcings during less sensitive, non-summer months, while retaining full precision during the dynamically volatile summer to help constrain the cascading growth of nonlinear errors. However, the extent to which this adaptive approach would restore summer fidelity remains an open question warranting further investigation."

Lines 726–730:

"Precision reduction applied to time-varying inputs introduces perturbations that interact with nonlinear model dynamics during integration Although the large-scale atmospheric structures remain morphologically coherent, the resulting spatial phase shifts can substantially affect localized diagnostics (e.g., precipitation). Consequently, input precision reduction should generally be avoided when exact spatiotemporal correspondence is required, while it remains a viable approach for preserving aggregate statistical distributions."

How does this work differ from the conclusions in all these previous works - is it by using the compressed data as inputs to WRF simulation? This should be made explicit.

Thank you for the suggestion. As you correctly noted, a key distinction of our work compared with previous studies is the explicit evaluation of precision reduction applied to time-varying input forcings used in WRF simulation. Most previous studies on atmospheric data compression primarily evaluated truncation strategies using static datasets, such as reanalysis products or model outputs, without propagating the compressed data through a dynamical model integration. In contrast, our study investigates precision reduction within the full WRF workflow, applying significant-digit rounding to both time-varying WRF input forcings and model outputs, and evaluating how these perturbations propagate through the nonlinear dynamics of the model.

In the revised manuscript, we have clarified this distinction more explicitly in the Abstract (Lines 20–24), the Introduction (Lines 81–94), and the Conclusions (Lines 726–730). We emphasize that precision reduction applied to time-varying inputs can manifest as spatial or temporal phase shifts in simulated weather systems. While the large-scale atmospheric structures generally remain morphologically coherent, the resulting phase shifts can substantially affect localized diagnostics. Consequently, our results show that input precision reduction may be acceptable for applications focusing on aggregate statistical properties but should generally be avoided when exact spatiotemporal correspondence is required.

In addition to evaluating input versus output precision reduction, our study also extends previous work by systematically examining how precision reduction affects downstream scientific diagnostics. The revised analysis framework now spans multiple levels of evaluation, including domain-scale statistical errors, regional extreme deviations, spatial structural fidelity, and downstream diagnostic metrics.

The prime focus of our study is precipitation, which represents one of the most challenging variables for compression because of its highly skewed distribution, intermittency, and cumulative representation in WRF (RAINNC + RAINC). To more comprehensively evaluate compression impacts across the full precipitation spectrum, we expanded our analysis beyond the extreme precipitation indices used in the original manuscript. In the revised Section 3.4, we also include diagnostics that explicitly assess the lower tail of the precipitation distribution, including precipitation occurrence metrics such as the Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), Frequency Bias (Bias), and a Zero Preservation Ratio (ZPR). These additions allow us to evaluate how precision reduction affects both light precipitation occurrence and extreme precipitation diagnostics. Together, these new analyses reveal that precision reduction applied to inputs and outputs produces fundamentally different impacts on downstream diagnostics, and they allow us to systematically identify safe precision thresholds for precipitation.

The methodological exposition and definition of the precipitation lower tail index have been incorporated into the revised manuscript (Section 2.4: Lines 274–283, Table 4). The extended analysis section is in Section 3.4 (Lines 511–555, Table 9).

Lines 20–24:

"Scientifically, precision-reduced inputs interact with nonlinear model dynamics and can induce spatial phase shifts in simulated meteorological systems. Although this process reduces deterministic grid-scale correspondence, the overall spatial morphology of the atmospheric structures remains largely preserved. Consequently, aggregate statistical distributions are weakly affected, especially during dynamically less active periods, rendering input precision reduction suitable for large-scale spatial aggregates and event-averaged statistical analyses."

Lines 81–94:

"Despite these advancements, most existing studies primarily consider precision reduction as a post-processing operation applied to static datasets or model outputs. Comparatively less attention has been paid to how precision reduction applied to time-varying model inputs propagates through nonlinear numerical integrations. A fundamental distinction exists between the precision reduction of input data and that of output data. Perturbations introduced into the input forcings may be amplified through nonlinear model physics, potentially influencing simulated trajectories and downstream diagnostics. Such responses are unlikely to be spatially or temporally uniform, as sensitivity to perturbations depends on the prevailing dynamical and thermodynamic regime.

Crucially, the extent to which precision reduction applied to model inputs and outputs propagates into downstream diagnostic analyses remains largely unquantified. While the fidelity of fundamental thermodynamic and kinematic fields (e.g., temperature and wind) must be systematically verified, this uncertainty becomes particularly critical in downstream diagnostic studies of precipitation, which is

output as a cumulative quantity and exhibits intermittent and highly skewed characteristics. Given the disproportionate societal and economic impacts of extreme precipitation (Seneviratne et al., 2021; Davenport et al., 2021), it is therefore essential to assess whether precision reduction introduces artificial biases in extreme-event diagnostics, such as inflated dry-area coverage or altered peak precipitation intensities, that exceed expected analysis uncertainty."

Lines 274–283:

"Finally, the evaluation of precipitation for both the lower-tail thresholds that control the occurrence of light rainfall and the behavior of upper-tail extreme events is necessary given the highly skewed distribution of precipitation fields. Moreover, because total precipitation in WRF is represented as the cumulative sum of grid-resolved (RAINNC) and parameterized convective (RAINC) components, it exhibits a compounded sensitivity to numerical precision loss.

To assess structural fidelity at the lower bound of the precipitation spectrum, we first performed a dichotomous verification for hourly precipitation, using a wet–dry threshold of 0.1 mm h^{-1}. For each precision-reduced configuration, grid cells were classified relative to WRF_bl. Based on this comparison, several categorical verification metrics were calculated (Roebber, 2009), including the Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), and Frequency Bias (Bias). In addition, a Zero Preservation Ratio (ZPR) was introduced to quantify the fraction of dry grid cells that remain correctly classified after precision reduction."

Lines 511–555:

"3.4 Impacts on Precipitation Diagnostics: Zero-Value Preservation and Extreme Precipitation Indices

While the SSIM analysis reveals how these numerical artifacts alter the spatial structure of precipitation fields, an equally important question is how distortions introduced by precision reduction in cumulative precipitation fields propagate into downstream scientific diagnostics. Precipitation exhibits a highly skewed spatial and temporal distribution, characterized by two distinct and sensitive regimes: a widespread lower tail dominated by zero or near-zero values, and a rare but high-impact upper tail driven by extreme rainfall events. Both regimes are particularly vulnerable to numerical artifacts introduced by precision reduction. To quantify these impacts, the following section evaluates how precision-reduced configurations affect precipitation diagnostics, including zero-value preservation and extreme precipitation indices.

We first evaluate the preservation of the precipitation occurrence spectrum, specifically focusing on the zero-value boundary (using the 0.1 mm h^{-1} wet-day threshold). Standard contingency metrics, including ZPR, POD, FAR, CSI, and Bias, are employed to diagnose how different compression configurations alter the fundamental wet/dry morphology (Table 9).

For input-only precision reduction configurations, the metrics indicate that the overall precipitation spectrum is largely preserved. The Frequency Bias remains close to unity (e.g., 0.99942 for WRF_3), confirming that the total precipitation area is conserved. Furthermore, the nearly symmetric error rates, where the moderate FAR (7.84%–8.72%) closely balances the corresponding miss rates (1 - POD), together with a stable CSI (~0.84) indicate that input precision reduction induces spatial phase shifts in precipitation systems rather than systematically disrupting their structure. In contrast, aggressive output precision reduction introduces substantial numerical distortions that directly erode the bottom end of the precipitation spectrum, most prominently in the WRF_fx3 configuration. Under this configuration, the POD drops sharply to 71.94%, indicating that approximately 28% of valid light precipitation events are

artificially truncated to zero due to the loss of significance in the arithmetic of cumulative precipitation variables. As a result, a pronounced systematic dry bias emerges (Bias ≈ 0.77).

Interestingly, WRF_fx3 exhibits an apparently higher ZPR (99.12%) and a lower FAR (6.90%) compared with WRF_fx4 (98.5% and 8.98%, respectively). However, this does not indicate improved structural fidelity. Instead, it reveals that WRF_fx3 suppresses both numerical noise (reducing FAR) and genuine light stratiform precipitation (reducing POD), effectively converting them into widespread non-physical dry regions.

The structural degradation previously indicated by the late-stage decline in SSIM is further corroborated by the temporal evolution of precipitation occurrence metrics under the WRF_fx4 configuration (Fig. S4). At the domain scale, the ZPR remains high (> 97 % throughout the year), largely reflecting the overwhelming dominance of climatologically dry background grids. However, metrics that evaluate performance specifically within precipitation events (CSI, POD, and FAR) reveal a pronounced late-stage deterioration. Consistent with the SSIM decline observed in November, the CSI decreases from its summer peak (~0.89) to 0.769 by December. This degradation reflects a dual numerical distortion associated with the expanding accumulation baseline. As the quantization step increases during November and December, the FAR rises substantially (reaching 11.4%), indicating the growing occurrence of spurious drizzle signals introduced by rounding artifacts. At the same time, the POD declines to 85.4%, implying that approximately 15% of genuine winter stratiform precipitation events are artificially truncated to zero. The Frequency Bias further highlights this shift in behavior. While the system exhibits a slight wet bias in spring, it gradually transitions to a systematic dry bias by December (Bias = 0.964). The simultaneous increase in false alarms and the loss of weak precipitation events progressively disrupt the spatial consistency of precipitation structures, definitively explaining the late-stage SSIM collapse.

The analyses above focus on the lower boundary of the precipitation spectrum, where precision reduction primarily affects the occurrence of light rainfall and the preservation of dry conditions. However, precipitation diagnostics are also strongly influenced by the opposite end of the distribution."

Lines 726–730:

"Precision reduction applied to time-varying inputs introduces perturbations that interact with nonlinear model dynamics during integration. Although the large-scale atmospheric structures remain morphologically coherent, the resulting spatial phase shifts can substantially affect localized diagnostics (e.g., precipitation). Consequently, input precision reduction should generally be avoided when exact spatiotemporal correspondence is required, while it remains a viable approach for preserving aggregate statistical distributions."

Table 4. Definitions of categorical verification metrics for hourly precipitation and daily extreme precipitation indices.

Name	Definition	Units
ZPR	Percentage of dry grids (< 0.1 mm h^{-1}) in WRF_bl that remain below the threshold after precision reduction.	%
POD	Percentage of wet grids (≥ 0.1 mm h^{-1}) in WRF_bl that are correctly retained after precision reduction.	%
FAR	Percentage of wet grids after precision reduction that are false alarms relative to WRF_bl.	%

CSI	A comprehensive metric for the spatial fidelity of the precipitation field, penalizing both the artificial elimination (missed events) and generation (false events) of wet events.	-
Bias	Ratio of the total number of wet grids in the precision-reduced configurations to that in WRF_bl. Values > 1 indicate an artificial inflation of precipitation spatial extent.	-
R95p_days	Number of days per year with daily precipitation exceeding 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}

Table 9. Categorical verification metrics for hourly precipitation across different precision reduction configurations relative to the WRF_bl.

Configuration	ZPR (%)	POD (%)	FAR (%)	CSI	Bias
WRF_5	98.7086	92.1787	7.8366	0.8548	1.00017
WRF_4	98.7078	92.1526	7.8431	0.8545	0.99995
WRF_3	98.5641	91.2276	8.7198	0.8391	0.99942
WRF_fx5	99.8289	99.2410	1.0357	0.9822	1.00280
WRF_fx4	98.4663	94.3254	8.9820	0.8629	1.03634
WRF_fx3	99.1221	71.9411	6.8957	0.6830	0.77269
WRF_5fx5	98.6073	91.8428	8.4280	0.8468	1.00296
WRF_5fx4	97.4814	88.3594	14.7486	0.7664	1.03646
WRF_5fx3	98.5405	68.4151	11.4639	0.6285	0.77274
WRF_4fx5	98.6069	91.8190	8.4321	0.8466	1.00274
WRF_4fx4	97.4812	88.3404	14.7522	0.7663	1.03628
WRF_4fx3	98.5400	68.4076	11.4680	0.6284	0.77269
WRF_3fx5	98.4697	90.9292	9.2680	0.8320	1.00217
WRF_3fx4	97.3682	87.6111	15.4209	0.7554	1.03585
WRF_3fx3	98.4717	67.9713	12.0081	0.6220	0.77247

*Note: A threshold of 0.1 mm h^{-1} is applied to distinguish between dry and wet grids. ZPR is the Zero Preservation Ratio, POD denotes the Probability of Detection, FAR is the False Alarm Ratio, CSI is the Critical Success Index, and Bias represents the Frequency Bias.

The article would benefit from an illustration of what are the input and output variables for the Weather Research Forecasting models, and a schematic of how the data interact.

Thank you for this helpful suggestion. The initial atmospheric state is provided via either an initial condition file, typically for the very first day of the study, or a restart file (wrfrst) which is generated after each day simulation, to serve as an initial condition for a subsequent day of simulation. Either initial condition file or wrfrst file, it provides full three-dimensional atmospheric and land-surface prognostic states required for seamless temporal continuity. For consistency purposes, the annual run started with a 2-day spin-up on 12/30/2015 so every day in 2016 the model started with a restart file and four prescribed time-varying external forcing files, which include atmospheric and surface analysis nudging fields (wrffdda, wrfsfdda), lateral boundary condition tendencies (wrfbdy), and lower boundary inputs such as sea surface temperature (wrflowinp). The outputs are written to wrfout files and contain near-surface and surface variables, three-dimensional atmospheric state fields, cloud microphysical quantities, radiation and energy fluxes, and other diagnostics. We have systematically detailed the exact data structures of the WRF model's inputs and outputs in the revised manuscript (Section 2.1: Lines 136–143, 145–147).

The entire precision reduction workflow is visually anchored by the newly introduced Figure 1. As detailed in the revised manuscript (Section 2.2: Lines 185–194), we delineate the workflow into preprocessing, model integration, and post-processing stages. Specifically, we clarify that input precision reduction exclusively targets time-varying forcing files (wrffdda, wrfsfdda, wrfbdy, wrflowinp) during the preprocessing stage. Evaluating each input precision level strictly necessitates a fully independent WRF simulation to physically capture the non-linear propagation of rounding errors. Conversely, output precision reduction is explicitly defined as a static, offline post-processing operation applied to the final wrfout files without the need to rerun the model.

Lines 136–143:

“To establish dynamical consistency prior to the evaluation period, the annual simulation for 2016 was preceded by a 2-day spin-up initialized on 30 December 2015. The annual integration was conducted in consecutive daily segments, with each day initialized from a model restart file (wrfrst) generated at the conclusion of the previous day. These restart files provide the three-dimensional atmospheric and land-surface state required for seamless temporal continuity. In addition to daily initialization, the model evolution was continuously constrained by prescribed time-varying external forcing fields, including atmospheric and surface analysis nudging (wrffdda, wrfsfdda), lateral boundary condition tendencies (wrfbdy), and lower-boundary updates such as sea surface temperature (wrflowinp).”

Lines 145–147:

“In contrast, model outputs (wrfout) contain near-surface variables, three-dimensional atmospheric states, cloud microphysical quantities, and energy fluxes, representing the prognostic results of the completed simulation.”

Lines 185–194:

“The systematic integration of this rounding framework into the overall experimental design is summarized in Fig. 1. The schematic depicts the three core workflow stages: preprocessing on input with the precision-reduction tool, model integration with full-precision input and precision-reduction input, and post-processing on model integration result, wrfout file, with the precision-reduction tool. Lossless

compression will be applied to all input and output. For input precision reduction configurations, the rounding procedure was applied during the preprocessing stage, exclusively targeting the time-varying forcing files (wrffdda, wrfsfdda, wrfbdy, and wrflowinp). Because these files define the dynamic external constraints on the model, evaluating each input precision level necessitated a fully independent WRF simulation to capture the non-linear propagation of rounding errors. Conversely, for output precision reduction, rounding was applied to the final wrfout files after simulation completion. As a purely static post-processing operation, output precision can be flexibly adjusted without requiring computationally expensive model re-runs."

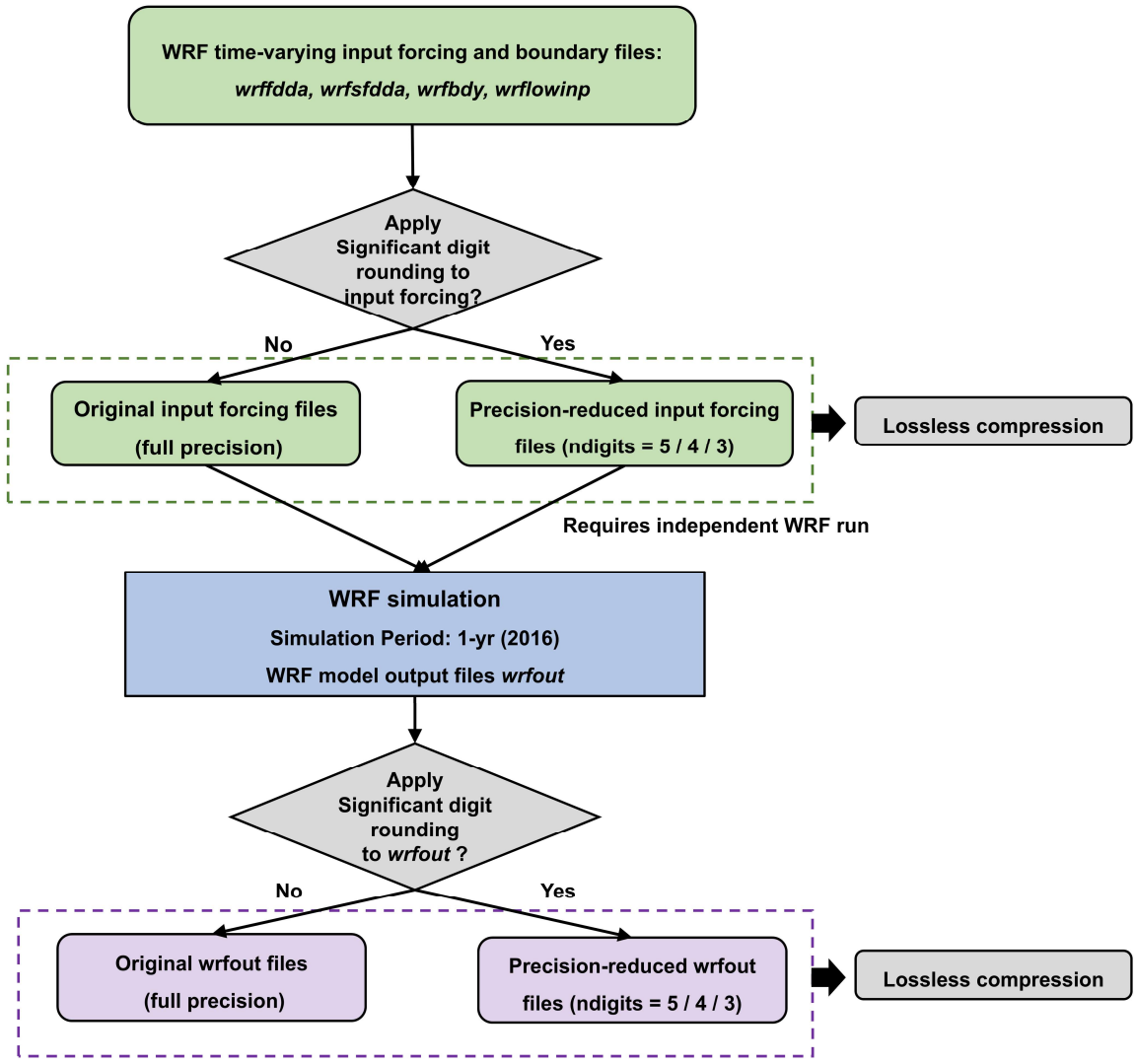

Figure 1: Schematic illustration of the experimental workflow for WRF data precision reduction and lossless compression.

Am I right that the truncation of input data to WRF has an impact on the output results coming from the WRF, and that this is the reason why, given the same output truncation, different input truncations can reduce the relative compression size of the outputs?

Thank you for this question. You are correct from a modeling perspective: input precision reduction effectively acts as a small perturbation. Due to the nonlinear dynamics of the WRF model, these

perturbations definitely propagate and produce quantitatively different final output fields (as we evaluate extensively in our scientific impact analysis). However, regarding the compressibility (storage footprint) of these outputs, our results indicate that the impact of these input perturbations is actually negligible. Modifying the input precision does not meaningfully change the information entropy or the overall numerical distribution of the final outputs in a way that affects the performance of the backend lossless codecs.

We realize that our original visualization in Figure 2 caused this exact misunderstanding. In the original manuscript, Figure 2 included three subplots: (a) input relative compression ratio, (b) output relative compression ratio, and (c) total (combined) relative compression ratio. You reasonably observed in subplot (c) that under the same output precision reduction, applying different input precision reduction improved the combined compression ratio. However, because the calculation for subplot (c) pooled both datasets (inputs + outputs) together, the apparent improvement in the total relative compression ratio was driven entirely by the size reduction of the input dataset itself, not by any increased compressibility of the output dataset.

Recognizing that the combined metric in the original subplot (c) was misleading, we have removed it. We believe that presenting the input and output compression ratios separately is much clearer and accurately conveys our results. This updated visualization is now presented as Figure 3 in the revised manuscript. We have also added an explicit statement in the revised manuscript (Section 3.1: Lines 309–311) to clarify that while input precision reduction perturbs the physical output values, its impact on the ultimate compressibility of the model outputs is negligible.

Lines 309–311:

"As physically expected, the final compressed volume of the output is predominantly dictated by the precision applied during the output post-processing stage; varying the input precision exerts no immediate influence on the final compressed wrfout file size."

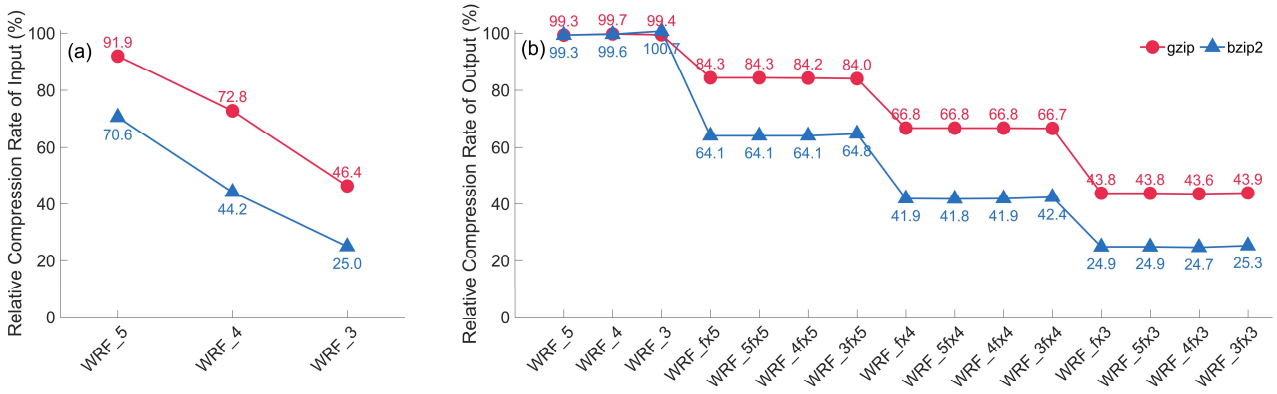

Figure 3: Relative compression ratios of (a) input data and (b) output data across diverse precision-reduction configurations.

Are input variables forcing variables, or are they also weather data? It is only on line 325 that we can infer that output variables are not recursively fed back into the model, since output-only truncation can happen after the model is run.

Thank you for the question. In this study, the WRF input data encompasses both internal weather states and external forcing variables. In our workflow, precision reduction was strictly applied only to these time-varying external forcing fields, leaving the internal weather states within the restart files intact. We have explicitly defined the targeted time-varying input forcings in the revised Methodology (Section 2.1: Lines 129–131; Section 2.2: Lines 188–190).

Regarding the model outputs, you are correct that the output variables are not recursively fed back into the model integration. The output precision reduction evaluated in this study functions strictly as an offline, post-processing filter. We have explicitly stated the non-recursive nature of the output in the revised manuscript (Section 2.2: Lines 192–194). The newly added Figure 1 in the revised manuscript also illustrates this information.

Lines 129–131:

"The sheer volume of these high-resolution input arrays significantly exacerbates the overall storage and data management burden. Consequently, these massive time-varying input files represent a critical target for our proposed precision reduction."

Lines 188–190:

"For input precision reduction configurations, the rounding procedure was applied during the preprocessing stage, exclusively targeting the time-varying forcing files (wrffdda, wrfsfdda, wrfbdy, and wrflowinp)."

Lines 192–194:

"Conversely, for output precision reduction, rounding was applied to the final wrfout files after simulation completion. As a purely static post-processing operation, output precision can be flexibly adjusted without requiring computationally expensive model re-runs."

The relationship between the 1622 stations and the surface data is unclear. Do the authors have access to simulated and data-assimilated dense surface data?

Thank you for the question. The 1622 stations represent the independent, point-based observational dataset used for evaluation. To establish a clear spatial relationship between these two distinct data collections, we applied a nearest-neighbor interpolation method to align the point-based observations with the simulated model grid. Specifically, each station was spatially matched to the WRF grid point with the smallest geographical distance. We have explicitly detailed this spatial matching procedure in the revised manuscript (Section 2.3: Lines 219–223).

Considering that you might still be concerned about whether the data used for evaluation and the nudging method we employed in the simulation process are in conflict, we emphasize that we used weak nudging coefficients for the free troposphere, which acts as a large-scale constraint to anchor the synoptic circulation. Atmospheric nudging was explicitly disabled within the planetary boundary layer and at the surface to prevent the suppression of near-surface thermodynamic variability. We have supplemented the details regarding the application of the nudging method in our simulation (Section 2.1: Lines 116–129).

Lines 116–129:

"To prevent systematic model drift during the long-term integration, the four-dimensional data assimilation (FDDA) via grid-nudging was employed within the WRF framework, utilizing the North American Mesoscale (NAM) analysis, produced operationally by the National Centers for Environmental Prediction (NCEP). In model settings, weak nudging coefficients (1.0×10^{-4} s^{-1} for wind and temperature, and 1.0×10^{-5} s^{-1} for moisture) were deliberately selected for the free troposphere. This configuration acts as a large-scale constraint to anchor the synoptic circulation while granting the model sufficient degrees of freedom to generate free-evolving, high-resolution mesoscale features. Furthermore, atmospheric nudging was explicitly disabled within the planetary boundary layer and at the surface to avoid suppressing near-surface thermodynamic variability. Because the boundary-layer and surface atmospheric fields evolved entirely according to model physics, subsequent evaluations of the WRF outputs against independent observational station data remain scientifically valid and uncontaminated by the direct assimilation process.

To mitigate long-term model drift in deep soil temperature and moisture, we employed the Pleim-Xiu land surface model combined with indirect soil nudging, which specifically requires the continuous ingestion of two-dimensional surface analysis fields (wrfsfdda). Coupled with the three-dimensional atmospheric fields (wrffdda) utilized for upper-air nudging, the model is forced to ingest massive datasets every 3 hours."

Lines 219–223:

"To spatially align the point-based measurements with the gridded model output, a nearest-neighbor interpolation scheme was applied, assigning each station to the WRF grid point with the smallest Euclidean distance in the latitude–longitude space. Despite the inherent representativeness limitations typical of point-to-grid regional climate validation, this extensive 1622-station network provides a robust benchmark for assessing the model's meteorological fidelity at both regional and CONUS-wide scales."

What are the inputs and outputs to the WRF? Do the authors re-run the WRF at different input data truncation strategies?

Thank you for the question. Regarding the specific definitions of WRF inputs and outputs within our workflow, we have described in our previous response (please refer to that reply).

Concerning the experimental execution, we reiterate that a completely separate, independent WRF simulation was conducted with distinct input (full precision input and precision-reduced input). This independent integration strategy is strictly necessary to physically capture how errors non-linearly propagate through the model dynamics. To make this parallel experimental design unambiguously clear,

we have added an explicit declaration in the revised manuscript (Section 2.2: Lines 190–192). The newly added Figure 1 in the revised manuscript also illustrates this scheme.

Lines 190–192:

"Because these files define the dynamic external constraints on the model, evaluating each input precision level necessitated a fully independent WRF simulation to capture the non-linear propagation of rounding errors."

What is the N in the error formulas: is it the number of input/output datapoints in 2016? Or is it the number of discrete observation station measurements? How is the distribution of observation points compared to that of the data used in the WRF?

Thank you for the question. The term *N* dynamically represents the total number of valid, spatiotemporally matched data pairs used in each specific comparison. Specifically, for validation against discrete ground-based observations, *N* corresponds to the total number of valid station-hour pairs across the NCDC network (where each station measurement is spatially matched to its nearest WRF grid point). Conversely, for comparisons against gridded satellite products, which were spatially interpolated directly onto the WRF model grid, *N* refers to the total number of grid-time pairs within the WRF domain. We have incorporated this explicit definition of *N* in the revised manuscript (Section 2.4: Lines 253–255).

Regarding the spatial relationship between the datasets, the WRF model generates a continuous, high-resolution gridded surface over the contiguous United States (CONUS). In contrast, the 1622 NCDC surface stations provide a discrete, albeit dense, spatial sampling network. The geographical distribution of these observation stations explicitly visualized in the manuscript (presented as Figure 1a in the original submission, and now as Figure 2a in the revised manuscript).

Lines 253–255:

"For station-based validation against NCDC observations, *N* corresponds to the number of valid station–hour pairs across all stations. For comparisons against gridded satellite products, *N* corresponds to the total number of grid–time pairs within the evaluated domain."

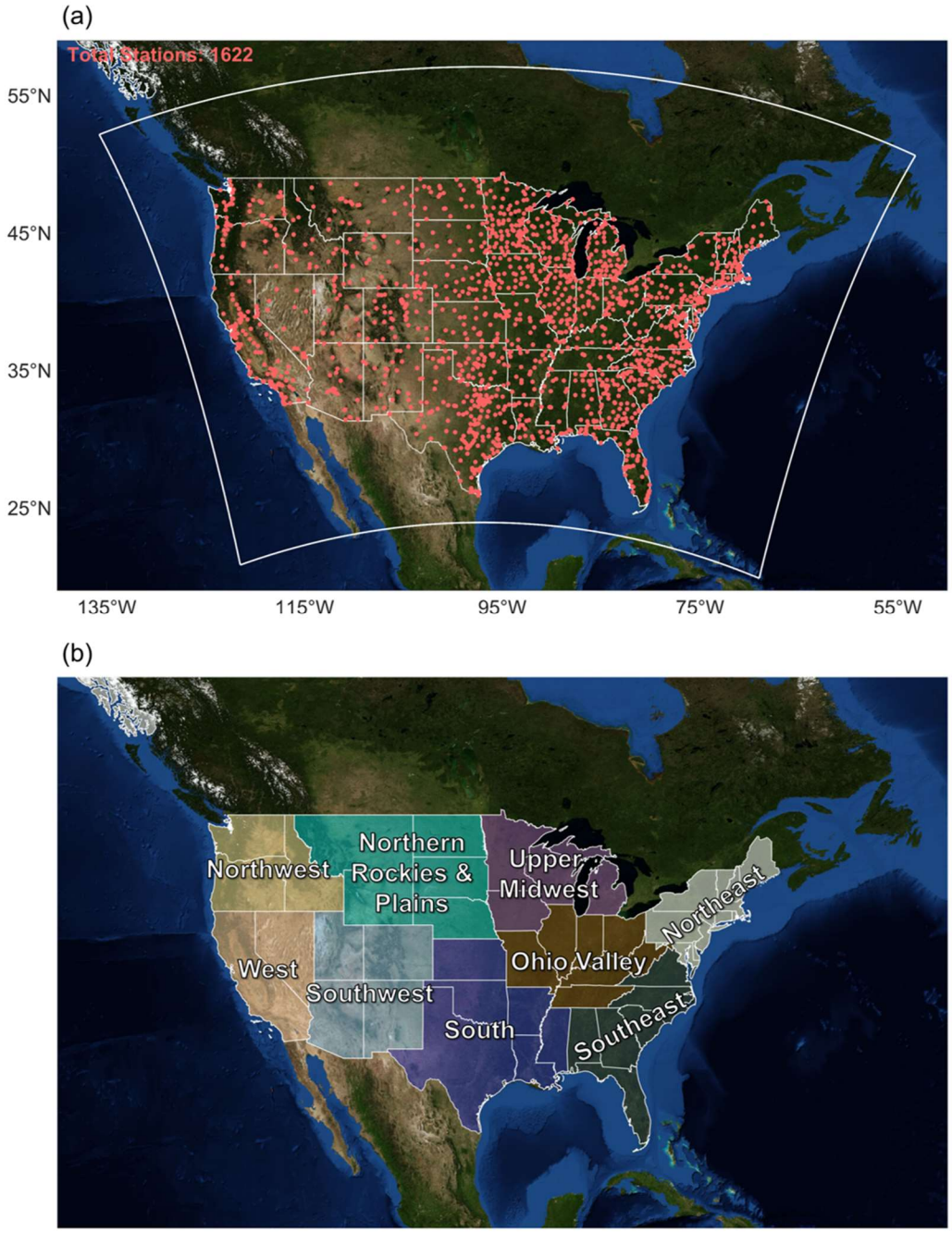

Figure 2: Spatial coverage of surface observational datasets, WRF model domain, and regional divisions: (a) locations of NCDC surface meteorological stations (red dots) across CONUS domain; the white box outlines the WRF simulation domain; (b) delineation of nine climatologically coherent subregions. Base map imagery from NASA Worldview, Earth Observing System Data and Information System (EOSDIS).

It was unclear if 69% at 5 significant digits in the input data (WRF_5) meant that:

1) original data at full precision are further compressed using lossless gzip

2) 5-digit truncated input data are further compressed using lossless gzip

3) the ratio of the storage size 2) over 1) is computed.

Thank you for the comment. Your interpretation is entirely correct. The “relative compression ratio”

is explicitly designed to isolate the additional space savings attributable strictly to decimal significant-digit rounding. It is formally calculated as the storage size of the precision-reduced data (subsequently compressed via a lossless codec) divided by the storage size of the full-precision data (compressed using the identical lossless codec). To avoid any ambiguity, we have provided an explicit and formal definition of the relative compression ratio in the revised manuscript (Section 3.1: Lines 304–306).

Lines 304–306:

"Here, the relative compression ratio is defined as the compressed size of the precision-reduced configurations divided by the compressed size of the full-precision baseline. This metric isolates the additional space savings attributable specifically to decimal significant-digit rounding."

When the authors write that the baseline dataset has 837GB of input data, are these full precision data or data compressed using gzip / bzip2?

Thank you for the question. The reported 837 GB baseline explicitly refers to the original, completely uncompressed WRF input forcings stored in the standard 32-bit single-precision floating-point format (with the output also adhering to the same standard 32-bit single-precision floating-point format). To permanently eliminate any ambiguity, we have provided a clearer expression in the revised manuscript (Section 3.1: Lines 300–302).

During this revision process, we proactively identified and rectified two reporting discrepancies in the original manuscript. First, the uncompressed baseline output volume was previously misstated as 2154.3 GB; the correct physical full-precision size is 7395.8 GB. Second, in the original Figure 2a, the input relative compression ratios were inadvertently calculated against the uncompressed raw data, rather than the lossless compressed baseline. This visualization error has been fully rectified in the updated Figure 3a to strictly align with the formal definition of relative compression. We emphasize that these corrections do not alter any scientific conclusions.

Lines 300–302:

"All calculations are benchmarked against a full-precision baseline (stored in the standard 32-bit single-precision floating-point format) comprising 837.0 GB of uncompressed input forcings and 7395.8 GB of uncompressed output data."

A table summarizing the effective storage space after various truncation strategies, as well as compression, would be very useful.

Thank you for the suggestion. As recommended, we have added a Table 5 (see below) and related description in the revised manuscript (Abstract: Lines 18–20; Section 3.1: Lines 303–304, 311–312; Conclusion: Lines 710–716).

Lines 18–20:

“From a storage perspective, combining precision reduction (retaining 5–3 digits) with bzip2 compression reduces model outputs to 19.2%–7.5% of their original uncompressed sizes and model inputs to 52.4%–18.5%.”

Lines 303–304:

“As summarized in Table 5 and illustrated in Fig. 3, data compressibility improves progressively with more aggressive precision reduction.”

Lines 311–312:

“On average, bzip2 outperforms gzip by 15–30 percentage points across all precision-reduced configurations (Table 5).”

Table 5. Storage and compression performance of input and output data across diverse precision-reduced configurations.

Data Type	Configurations	gzip			bzip2		
		Compressed size (GB)	% of Orig. Size	Relative compression ratio (%)	Compressed size (GB)	% of Orig. Size	Relative compression ratio (%)
INPUT (Original size: 837 GB)	WRF_bl	628.4	75.1	-	621.5	74.3	-
	WRF_5	577.4	69.0	91.9	438.5	52.4	70.6
	WRF_4	457.5	54.7	72.8	274.5	32.8	44.2
	WRF_3	291.7	34.9	46.4	155.1	18.5	25.0
OUTPUT (Original size: 7395.8 GB)	WRF_bl	2229.5	30.1	-	2212.6	29.9	-
	WRF_5	2214.0	29.9	99.3	2196.6	29.7	99.3
	WRF_4	2223.3	30.0	99.7	2203.8	29.8	99.6
	WRF_3	2217.0	30.0	99.4	2228.8	30.1	100.7
	WRF_fx5	1879.4	25.4	84.3	1418.5	19.2	64.1
	WRF_fx4	1489.8	20.2	66.8	926.0	12.5	41.9
	WRF_fx3	977.2	13.2	43.8	551.6	7.5	24.9
	WRF_5fx5	1879.6	25.4	84.3	1418.8	19.2	64.1
	WRF_5fx4	1489.9	20.2	66.8	926.1	12.5	41.9
	WRF_5fx3	977.3	13.2	43.8	551.4	7.5	24.9
	WRF_4fx5	1876.1	25.4	84.2	1414.8	19.2	64.1
	WRF_4fx4	1489.5	20.1	66.8	925.8	12.5	41.8
	WRF_4fx3	972.1	13.1	43.6	545.8	7.4	24.7
	WRF_3fx5	1872.2	25.3	84.0	1433.2	19.4	64.8
	WRF_3fx4	1486.5	20.1	66.7	937.6	12.7	42.4
	WRF_3fx3	978.2	13.2	43.9	560.5	7.6	25.3

*Notes: Relative compression ratio = (compressed size of precision-reduced configurations/compressed size of WRF_bl).

Lines 710–716:

"From a storage perspective, removing redundant numerical precision prior to compression yields substantial additional storage savings beyond conventional lossless compression. For model outputs, retaining 5 to 3 significant digits prior to bzip2 compression reduces data volumes to 19.2%–7.5% of their original uncompressed sizes (corresponding to 64.1%–24.9% relative to lossless compression alone). Comparable reductions relative to the original uncompressed dataset are achieved with Zstd (21.8%–10.2%), zlib (25.6%–13.0%), and gzip (25.4%–13.2%). Similarly, precision-reduced model inputs paired with bzip2 compress files to 52.4%–18.5% of their original volumes (corresponding to 70.6%–25.0% relative to lossless compression alone), while Zstd, zlib and gzip achieve 58.3%–25.5%, 69.1%–33.6% and 69.0%–34.9%, respectively."

The color scheme of Fig. 6 is confusing and does not correspond to previous figures.

Thank you for this comment. We fully agree that the categorical color mapping in the original Figure 6 lacked visual consistency with the preceding analyses. To resolve this, we have completely redesigned this visualization (now presented as Figure 9 in the revised manuscript). The updated figure strictly employs a standardized matrix of colors and marker shapes that systematically maps to specific climate regions and seasons, ensuring seamless visual continuity across the entire paper. Furthermore, during this graphical revision, we recognized that our original palette was not scientifically optimized for red-green color vision deficiencies. Consequently, we have systematically overhauled the color schemes across multiple visualizations in the manuscript (e.g., the newly revised Figure 5).

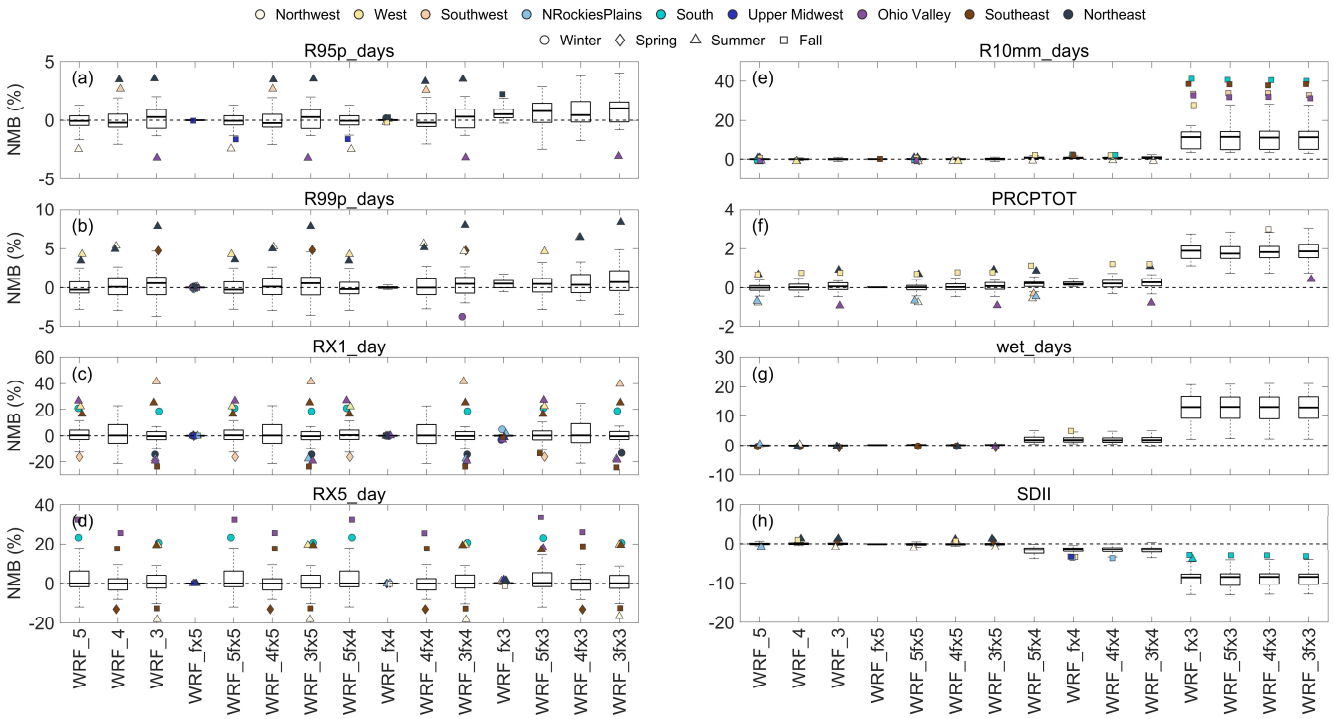

Figure 9: Box plots showing the distribution of NMB values for extreme precipitation indices across climate regions and seasons. Each box spans the interquartile range (IQR; 25th–75th percentiles), with

the median shown as a horizontal line. Whiskers extend to 1.5 × IQR, and outliers are plotted as colored shapes corresponding to different climate regions and seasons. Dashed horizontal lines indicate zero bias.

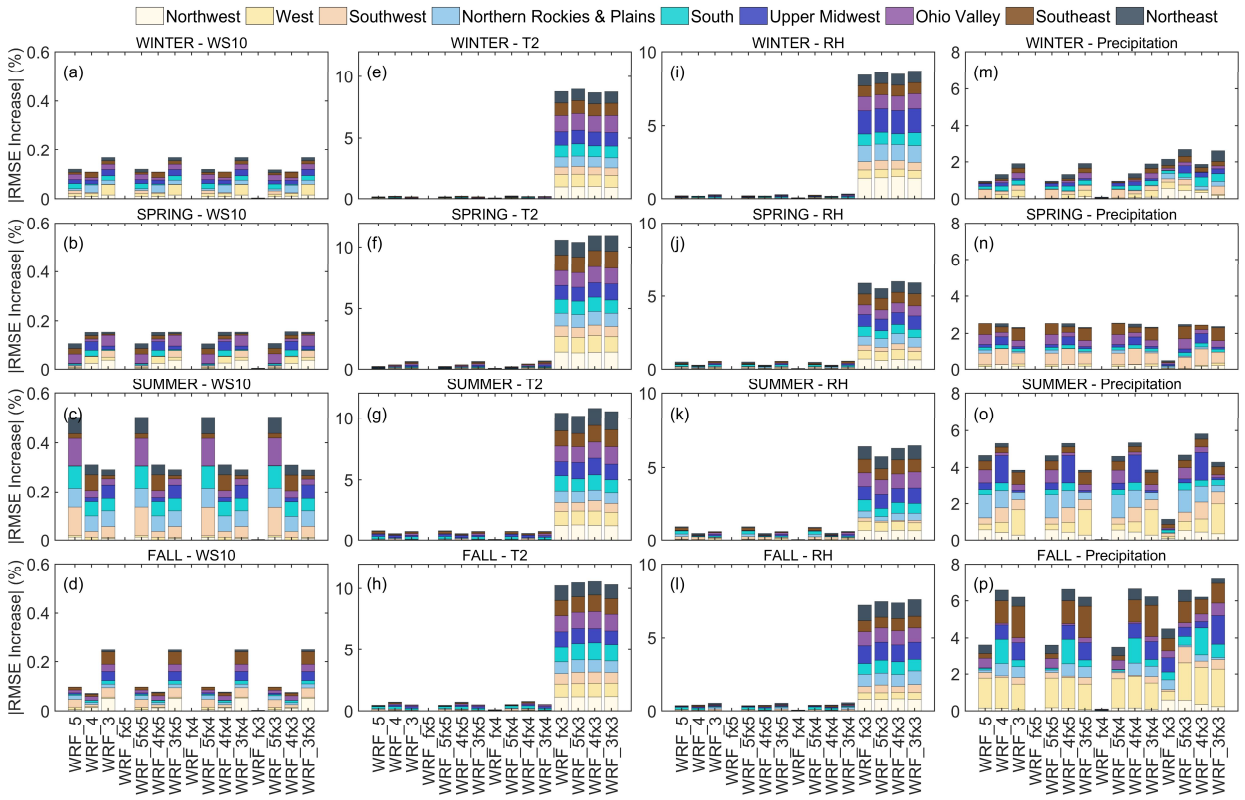

Figure 5: Magnitude of seasonal and regional relative changes in RMSE for meteorological variables across 15 precision-reduced configurations. Stacked bars illustrate the regional breakdown of |RMSE increase| relative to the WRF_bl across nine climate regions. Note that the total bar heights serve solely to visualize relative regional contributions and do not represent a mathematically aggregated domain-wide RMSE percentage. Results are presented separately for (a–d) wind speed, (e–h) temperature, (i–l) relative humidity, and (m–p) precipitation.

What is also missing is a visual showing one measurement (e.g., temperature at 2m) with the FORTRAN representation and corresponding truncations.

Thank you for this suggestion. In the revised manuscript, we have added a table (Table 2) and corresponding description in Section 2.2 (Lines 182–184). The table presents variables (T2, Q2, U10, and V10) at a single grid point and time step, showing the original values together with the corresponding values after retaining 5, 4, and 3 significant digits.

Table 2. An example of retaining different significant digits for WRF output variables.

Variables	Full precision	5-digit	4-digit	3-digit
T2 (K)	290.3516	290.35	290.4	290
Q2 (kg kg^{-1})	0.01852282	0.018523	0.01852	0.0185
U10 (m s^{-1})	0.5504633	0.55046	0.5505	0.551
V10 (m s^{-1})	-1.483089	-1.4831	-1.483	-1.48

Lines 182–184:

"To provide an intuitive visualization of the rounding results, Table 2 contrasts the original full-precision values against those retained at 5, 4, and 3 significant digits for key near-surface variables at a single grid point and time step: 2-m temperature (T2), 2-m specific humidity (Q2), and 10-m wind components (U10 and V10)."

References

Baker, A. H., Hammerling, D. M., and Turton, T. L.: Evaluating image quality measures to assess the impact of lossy data compression applied to climate simulation data, Computer Graphics Forum, 38, 517–528, https://doi.org/10.1111/cgf.13707, 2019.

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.

Delaunay, X., Courtois, A., and Gouillon, F.: Evaluation of lossless and lossy algorithms for the compression of scientific datasets in netCDF-4 or HDF5 files, Geoscientific Model Development, 12, 4099-4113, https://doi.org/10.5194/gmd-12-4099-2019, 2019.

Han, T., Chen, Z., Guo, S., Xu, W., and Bai, L.: CRA5: Extreme Compression of ERA5 for Portable Global Climate and Weather Research via an Efficient Variational Transformer, arXiv:2405.03376, https://doi.org/10.48550/arXiv.2405.03376, 2024.

Huang, L. and Hoefler, T.: Compressing multidimensional weather and climate data into neural networks, arXiv:2210.12538, https://doi.org/10.48550/arXiv.2210.12538, 2023.

Klöwer, M., Razinger, M., Dominguez, J. J., Düben, P. D., and Palmer, T. N.: Compressing atmospheric data into its real information content, Nature Computational Science, 1, 713-724, https://doi.org/10.1038/s43588-021-00156-2, 2021.

Mirowski, P., Warde-Farley, D., Rosca, M., Grimes, M. K., Hasson, Y., Kim, H., Rey, M., Osindero, S., Ravuri, S., and Mohamed, S.: Neural Compression of Atmospheric States, arXiv:2407.11666, https://doi.org/10.48550/arXiv.2407.11666, 2024.

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.

Roebber, P. J.: Visualizing multiple measures of forecast quality, Weather and Forecasting, 24, 601-608, https://doi.org/10.1175/2008WAF2222159.1, 2009.

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.

Silver, J. D. and Zender, C. S.: The compression–error trade-off for large gridded data sets, Geoscientific Model Development, 10, 413–423, https://doi.org/10.5194/gmd-10-413-2017, 2017.

Underwood, R., Bessac, J., Di, S., and Cappello, F.: Understanding the effects of modern compressors on the community earth science model, 2022 IEEE/ACM 8th International Workshop on Data Analysis and Reduction for Big Scientific Data (DRBSD), https://doi.org/10.1109/DRBSD56682.2022.00006, 2022.

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.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P.: Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13, 600–612, https://doi.org/10.1109/TIP.2003.819861, 2004.

Zender, C. S.: Bit Grooming: statistically accurate precision-preserving quantization with compression, evaluated in the netCDF Operators (NCO, v4.4.8+), Geoscientific Model Development, 9, 3199–3211, https://doi.org/10.5194/gmd-9-3199-2016, 2016.