https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Impact of Increased ROMEX GNSS Radio Occultation Data on Forecast Performance in the Korean Integrated Model

Eun-Hee Kim, Hyoung-Wook Chun, Yong Hee Lee

Korea Meteorological Administration, Daejeon, 35208, Republic of Korea

Correspondence to: Eun-Hee Kim (ehkim2010@korea.kr)

Abstract. Global Navigation Satellite System Radio Occultation (GNSS RO) data can be assimilated to improve global numerical weather prediction (NWP). While simulation studies and observing system experiments have indicated benefits from increasing RO data volume, the extent to which large-volume RO assimilation improves forecasts in an operational NWP configuration has been less thoroughly quantified. As part of the international Radio Occultation Modeling Experiment (ROMEX), which assesses the forecast impact of GNSS RO observations, this study evaluates the effect of increased GNSS RO data volume on the forecast performance of the Korean Integrated Model (KIM), the operational global NWP model of the Korea Meteorological Administration (KMA). Four baseline experiments were conducted: NoRO (without GNSS RO observations), CTL (operational-level GNSS RO volume), EXP_20k (≈20 000 profiles d^{-1}), and EXP_35k (≈35 000 profiles d^{-1}). To investigate the sensitivity of the EXP_35k results to the assimilation configuration, additional experiments were conducted by adjusting the observation-error inflation factor and the refractivity coefficient k_1 (the dry-term coefficient in the refractivity equation). A comparison of NoRO and CTL results showed that the assimilation of GNSS RO observations in KIM generally improved temperature and geopotential height forecasts in the lower stratosphere and upper troposphere. With larger GNSS RO datasets, EXP_20k showed relatively consistent improvements across most regions and levels, whereas EXP_35k maintained the global-mean improvement in 100 hPa temperature forecasts but showed weaker improvement in the tropics at that level. It also produced a pronounced increase in tropical 500 hPa geopotential height root-mean-square error, clearest through the first three forecast days and weakening only gradually through day 5. These results indicate that increasing the RO data volume does not translate linearly into forecast improvement and that the response of the current KIM assimilation configuration depends on the region, pressure level, and forecast variable. Increasing the observation-error inflation factor mitigated part of the mid-level geopotential height degradation but also indicated that this change weakens the upper-level temperature improvement attributable to RO data. In the k_1 adjustment experiments, increasing k_1 in KIM produced more favorable results than did decreasing it for certain variables and levels but did not directly resolve the main degradation observed in EXP_35k. The analysis-increment diagnostics showed similar tropical mid-level increment statistics across the three experiments, suggesting that the degradation is more likely related to accumulated differences in the analysis field or the forecast error growth than to excessive per-cycle analysis adjustment. These results show that the forecast impact of large-volume GNSS RO data is sensitive to the observation-error configuration and observation-operator coefficient. They further highlight that maximizing the value of future high-density GNSS RO observing systems depends not only on increased

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

observation availability but also on continued improvements in observation-error modeling, observation handling, and observation operators within data assimilation systems.

1 Introduction

Improving the accuracy of global numerical weather prediction (NWP) requires high-quality observational data that can be used to constrain a model's initial state to closely approximate the true state of the atmosphere. This is particularly true in regions where conventional surface-based observing networks are sparse, such as the oceans, polar regions, and the Southern Hemisphere, where satellite observations are key source of information for NWP. Among these observations, Global Navigation Satellite System Radio Occultation (GNSS RO) observations provide information on the temperature and density structure in the upper troposphere–lower stratosphere (UTLS) with high vertical resolution and globally homogeneous observational characteristics; therefore, they have played a key role in NWP data assimilation.

The contribution of GNSS RO data to NWP was systematically established through early impact-assessment studies using Challenging Minisatellite Payload (CHAMP) and Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) data. Healy and Thépaut (2006) showed that assimilating CHAMP RO observations could improve the global forecast performance, and Aparicio and Deblonde (2008) reported that RO refractivity data improved the representation of temperature structure in global forecasts. Subsequently, Cardinali (2009) and Bonavita (2014), among others, quantified the impact of RO data on short-range forecasts and analysis quality within observing systems and emphasized that in contrast to satellite radiances, GNSS RO requires minimal bias correction and has relatively stable error characteristics. These studies demonstrated that GNSS RO data can serve not merely as supplementary observations but also as an independent, stable anchor that constrains the atmospheric thermodynamic structure within global data assimilation systems. GNSS RO is also widely used as a key dataset for evaluating re-analyses and atmospheric observing systems more broadly, beyond NWP (Reiter et al., 2025).

However, the impact of GNSS RO data on NWP depends not only on the quality of the observational data but also on the data volume, observation-error configuration, observation operator, and characteristics of the assimilation system. The International Radio Occultation Working Group (IROWG) has recommended acquiring a baseline of at least 20 000 RO profiles per day with complete global and local-time coverage (Anthes et al., 2024). However, the data volume actually available to operational centers has often been lower than this baseline. In addition, the aging of key RO satellites and the uncertainty surrounding follow-on public missions have raised concerns about securing a stable supply of RO data in the future. Accordingly, active efforts are underway to expand RO data volume using commercial satellites and new satellite constellations.

The potential impact of large-volume GNSS RO data was first examined through observing system simulation experiments (OSSEs). Privé et al. (2022) evaluated the impact of increased RO data volume on NWP in an idealized experimental setting and found that the information from additional RO soundings did not saturate even at 100 000 soundings per day. However, OSSEs cannot fully represent the quality of real observational data, inter-satellite biases, quality-control procedures,

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

observation-operator errors, or the constraints of operational assimilation systems. Therefore, studies applying real RO data to an operational-level global assimilation system are needed to assess the impact of large-volume RO data. Motivated by this need, the international collaborative GNSS RO impact-assessment project, the Radio Occultation Modeling Experiment (ROMEX), was launched (Anthes et al., 2024; Shao et al., 2025). ROMEX aims to assess the effect of approximately 35 000 RO profiles per day, including commercial and new GNSS RO data not currently used by most operational centers, on operational NWP systems. The significance of ROMEX is not merely to reconfirm that GNSS RO observations improve forecasts. Rather, it is to determine the extent to which large-volume RO data beyond existing operational data volumes continue to contribute useful information in a real assimilation environment, and to identify how data assimilation components, including observation error, quality control, thinning, and observation-operator settings, should be adjusted as the data volume increases.

GNSS RO data are also an important observation type assimilated in the Korean Integrated Model (KIM), the operational global NWP model of the Korea Meteorological Administration (KMA). Previous KIM-based studies have adjusted the GNSS RO observation-error profile to improve the assimilation of COSMIC-2 data, thereby confirming the potential to enhance KIM forecast performance (Kim et al., 2025; Kim et al., 2023). However, these studies focused mainly on the use of operational or specific-constellation data and on observation-error optimization, whereas the manner in which large-volume RO data, including public and commercial missions across the US, Europe, and China, behave within KIM's 4D-EnVar assimilation system has not yet been thoroughly examined. In particular, when the data volume increases substantially, even if the quality of individual observations is good, several factors may act simultaneously: reduced effective independence among observations, increased representativeness error, limitations of the one-dimensional observation operator under horizontally inhomogeneous conditions, and excessive observation weighting.

As part of the ROMEX project, this study assesses the forecast impact of large-volume GNSS RO data in KIM. Beginning with an experiment without GNSS RO data, we examine in stages how forecast performance varies as the data volume increases from the operational level, through the IROWG-recommended level of approximately 20 000 profiles d^{-1}, to the ROMEX target level of approximately 35 000 profiles d^{-1}. Previous studies have shown that GNSS RO data contribute to improved global NWP performance. However, whether large-volume RO data, including commercial and new GNSS RO missions, translate linearly into forecast improvement within a real operational assimilation system requires further investigation. In particular, when the data volume increases substantially, observation density, observation-error configuration, thinning, quality control, and observation-operator characteristics can act in combination. Therefore, assessing the sensitivity of the assimilation system, in addition to the direct effect of increased data volume, is critical.

As the ROMEX project progressed, several operational centers and research groups reported results assessing the impact of large-volume GNSS RO data within their own assimilation systems. Zhang et al. (2025) used the JEDI assimilation system with the GFS model to increase RO data volume in stages from operational levels to approximately 20 000 and 35 000 profiles d^{-1}, and reported that forecast performance improved for most variables and levels at a progressively decreasing rate as data volume increased, without clear signs of saturation. Simultaneously, the European Centre for Medium-

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Range Weather Forecasts (ECMWF) (Lonitz, 2024) and the Met Office (Bowler and Lewis, 2026), among other ROMEX participants, reported increasing biases in lower-stratospheric temperature and geopotential height forecasts in 35 000 profiles d^{-1} experiments. In other works, although the use of large-volume RO data is generally positive, different assimilation systems have exhibited different sensitivities at specific levels and regions. Accordingly, this study aims to distinguish sensitivities common across the ROMEX community from responses specific to KIM by assessing the impact of large-volume RO data within KIM's assimilation configuration.

To this end, this study addresses three questions: First, what is the impact of GNSS RO data on KIM forecasts across regions and levels? Second, how does forecast performance change as RO data volume increases from operational levels to approximately 20 000 and 35 000 profiles d^{-1}? Third, how sensitive is the change in forecast performance observed with large-volume RO data to adjustments to the observation-error inflation factor and refractivity coefficient k_1? By informing how operational centers can assimilate large volumes of GNSS RO data more effectively, including appropriate observation-error settings, this study may contribute to more reliable NWP forecasts as future RO observing systems expand.

The remainder of this paper is organized as follows: Section 2 describes the ROMEX GNSS RO data, KIM assimilation system, GNSS RO observation-error configuration, and experimental design. Section 3 presents the impact of RO data, the change in forecast performance with increasing data volume, and the results of the observation-error-inflation and k_1 sensitivity experiments. Section 4 discusses the regional, vertical, and lead-time-dependent characteristics that emerge with large-volume RO data use, together with the analysis-increment diagnostics. Section 5 summarizes the main findings and future research directions.

2 Data and Methodology

2.1 ROMEX GNSS RO data

Both datasets used in this study, containing 20 000 and 35 000 profiles d^{-1}, were distributed to ROMEX participants by EUMETSAT's Radio Occultation Meteorology Satellite Application Facility (ROM SAF; Marquardt, 2024). The full ROMEX dataset at the 35 000 profiles d^{-1} level adds data from commercial and Chinese satellites to the operational constellation, while the 20 000 profiles d^{-1} dataset corresponds to the ROMEX20K subset, defined and distributed separately by ROM SAF to quantify the effect of increased data volume. Both datasets are cumulative: EXP_20k and EXP_35k successively add observations to the operational constellation used in the control experiment (CTL), forming a progressively expanding sequence of observation configurations. The differences in data composition at each stage are, therefore, not arbitrarily set but directly reflect the official data-volume tiers defined by the ROMEX project.

Figure 1 compares the GNSS RO observation distributions at 06 UTC on 1 September 2022 for CTL and EXP. CTL includes MetOp-B/C, TerraSAR-X, TanDEM-X, PAZ, Sentinel-6A, COSMIC-2, and KOMPSAT-5 data, which are used operationally in KIM. The EXP adds data from commercial and new missions (CICERO-OP1, GNOMES-A, SPIRE, FY-3C/D/E, Yunyao,

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

and Tianmu) to CTL. The global observation density was substantially higher in EXP than in CTL, with observation gaps notably reduced over the oceans, Southern Hemisphere, and high latitudes. This difference shows that ROMEX data not only increase the total number of observations but also supplement observational information in regions with relatively sparse observing networks.

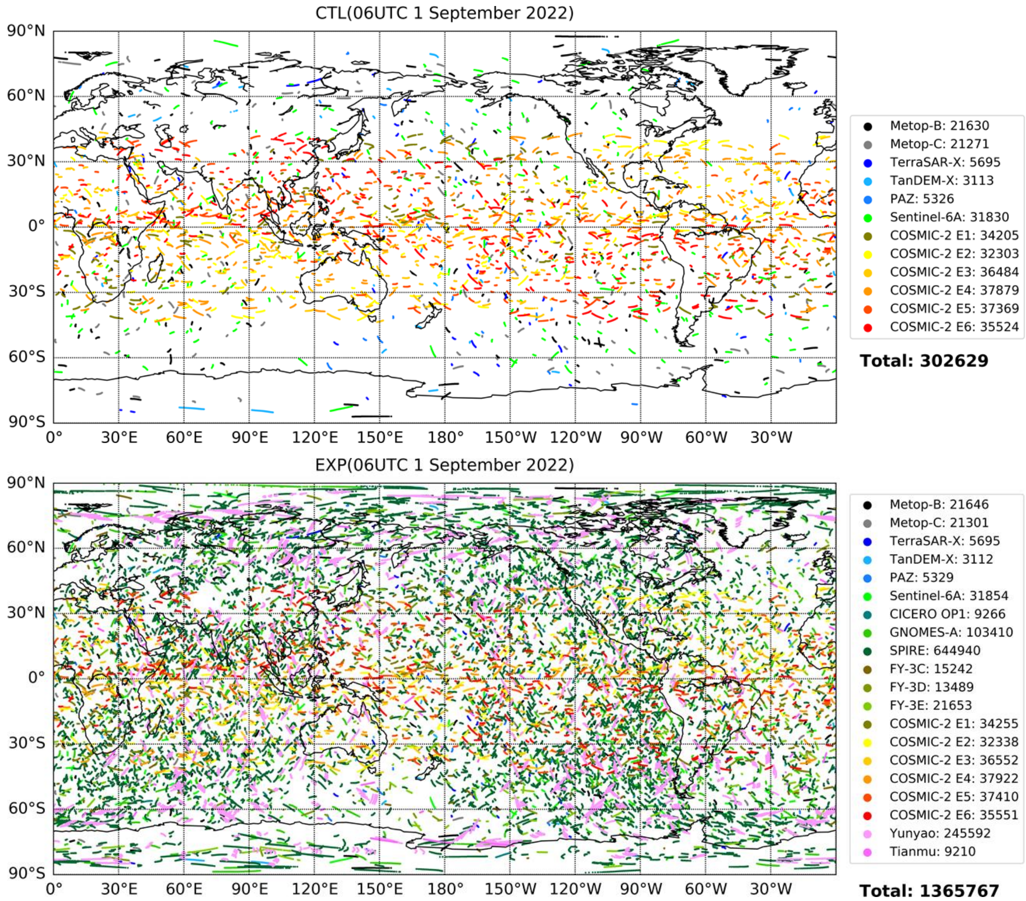


Figure 1. GNSS RO data distribution for CTL (top) and EXP (bottom) at 06 UTC on 1 September 2022. Legend values denote the number of bending-angle observation points per satellite, which is distinct from the number of profiles: a single RO profile consists of many vertical observation points, and because Fig. 1 shows the instantaneous distribution for a specific cycle (06 UTC), the legend totals appear large. The number of profiles distributed daily is approximately 8 000 profiles d^{-1} for CTL and 35 000 profiles d^{-1} for EXP.

Figure 2 shows the global statistics for the observation-minus-background (O−B) departures of GNSS RO bending-angle data collected during the ROMEX period as a function of impact height. The observation departure is defined as

$$O - B = y_o - H(x_b) \,, \tag{1}$$

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

where y_o denotes the GNSS RO observation, and $H(x_b)$ denotes the simulated observation computed by applying the observation operator H to the background field x_b. In this study, we computed the mean and standard deviation (STDV) of O−B in each altitude band to characterize the bias and scatter of the data.

In the lower troposphere, inter-satellite differences are relatively large, and a negative bias with large scatter is evident around 5–10 km, which is attributable to known sources of uncertainty in low-altitude GNSS RO observations, such as super-refraction, multipath propagation, and signal-tracking errors (Sokolovskiy, 2003; Xie et al., 2006). In the lower stratosphere, around 15–30 km, the O−B mean is close to zero and exhibits little scatter, indicating that ROMEX data are most stable in this altitude range. Above 30 km, the scatter increases again, which is interpreted as reflecting a reduced signal-to-noise ratio and residual ionospheric error. Apart from the lower troposphere, where inter-satellite differences are relatively large, the global O−B statistics in Fig. 2 show no clear quality differences among the satellite constellations. Overall, data from commercial and Chinese satellites exhibit vertical statistical characteristics similar to those of existing operational satellite data, indicating that they are of sufficient quality for assessing the impact of large-volume RO data.

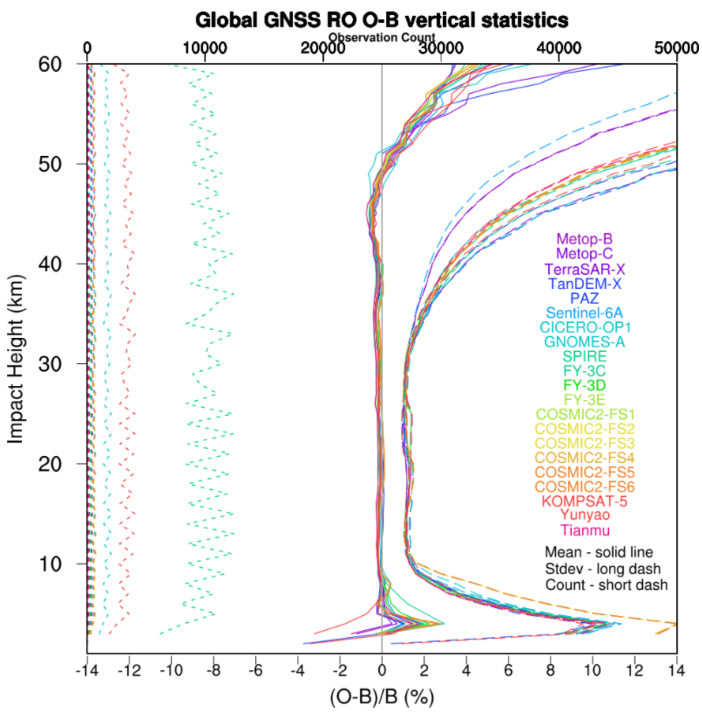


Figure 2. Vertical distributions of GNSS RO O−B statistics for September–November 2022.

2.2 KIM and GNSS RO assimilation

This study was conducted using the global data assimilation system of KIM. KIM is a global model based on a cubed-sphere grid (Choi, 2018; Hong and Jang, 2018); it uses a hybrid 4D-EnVar system for data assimilation (Kwon et al., 2018; Shin et al., 2018). The ROMEX experiments were run at a horizontal resolution of approximately 25 km with a 50-member ensemble

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

at a resolution of 50 km. In all the experiments, the GNSS RO data were assimilated in the bending-angle form. The observation operator used was ROPP's one-dimensional bending-angle operator; tangent point drift correction was applied, whereas bias correction was not. Vertical thinning at 400 m intervals was applied to accommodate large data volumes; no horizontal thinning was applied. Because the number of profiles increases with data volume while the horizontal distribution remains unthinned, the horizontal observation density in EXP_35k is correspondingly higher than that in CTL.

Figure 3 shows the rejection rate of the GNSS RO observations during quality control and thinning for EXP_35k. The quality control procedure consists of a gross check, duplicate check, background check, and vertical thinning. Each rejection rate is defined as the fraction of bending-angle observation points removed at that stage relative to the number entering it, so that the stages are sequential rather than additive; approximately 53 % of the input observation points remain after all stages. Of these, vertical thinning accounted for the largest removal fraction, reflecting the processing required to assimilate large-volume RO data within a limited computational budget. Excluding vertical thinning, the cumulative rejection rate of the three quality control steps (gross, duplicate, and background check) is about 15 %, which is relatively low and indicates that most of the ROMEX data pass standard quality control.

The GNSS RO observation error used in KIM has an altitude-dependent structure (Kwon et al., 2015). Larger observation errors are assigned in the lower troposphere to account for super-refraction and signal-tracking uncertainty. The assigned errors decrease with altitude, such that a relative error of approximately 1 % is applied at altitudes above approximately 10 km, except over parts of Antarctica. The O−B statistics presented in Fig. 2, large scatter in the lower and upper layers and small scatter in the lower stratosphere, are qualitatively consistent with this observation-error configuration.

To assess the sensitivity of large-volume RO data, an inflation factor α was applied to the baseline observation error in some experiments. The effective observation error is defined as

$$\sigma_o^{eff}(z) = \alpha\sigma_o^{ctl}(z) , \tag{2}$$

where, $\sigma_o^{ctl}(z)$ denotes KIM's baseline GNSS RO observation error, and α is the observation-error inflation factor. In the operational configuration, KIM applies an inflation factor of α = 2.5 to its baseline GNSS RO observation error, following the observation-error tuning study by Kim et al. (2025). In the sensitivity experiments conducted as part of this study, α = 4.3 and α = 5.3 were applied. These values were chosen to span a range of moderate and strong downweighting rather than being derived from an observation-error diagnostic. The goal of these experiments was not to permanently downweight RO data but to diagnose how sensitive the forecast-performance variation observed in EXP_35k is to the observation-error configuration.

In addition, to assess the sensitivity of the physical coefficient used in the refractivity calculation, a k_1 adjustment experiment was performed. The GNSS RO refractivity, comprising a dry term and wet term, is generally expressed as

$$N = k_1\frac{P}{T} + k_2\frac{e}{T^2} , \tag{3}$$

where, N denotes refractivity, P denotes pressure (hPa), T denotes temperature (K), and e denotes water-vapor pressure (hPa). In this study, the dry-term coefficient k_1 was adjusted by ±0.1 % relative to its reference value, to assess how variations in this

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

observation-operator physical coefficient influence forecast performance. As KIM does not apply bias correction to the GNSS RO data, a fine adjustment of k_1 shifts the global bias characteristics through the observation operator. Therefore, the k_1 adjustment experiment was designed not to identify an optimal value of this physical coefficient but to diagnose KIM's sensitivity to the interaction between background bias and the observation-operator coefficient.

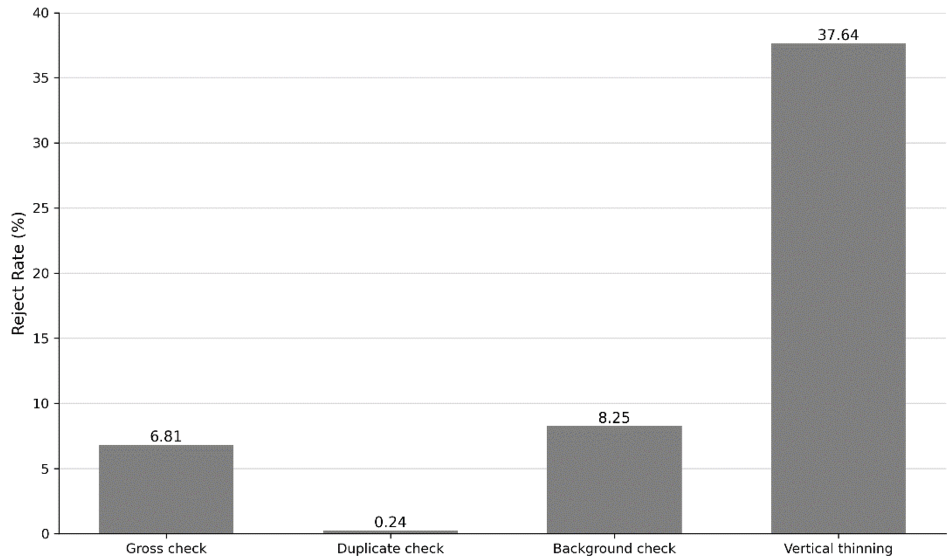


Figure 3. Rejection rate of GNSS RO bending-angle observation points at each stage of quality control and thinning in EXP_35k, September–November 2022. The stages are applied sequentially in the order shown, and each rate is calculated relative to the number of observation points entering that stage.

2.3 Experimental design and verification

Table 1 summarizes the experimental design. The experiments were organized into three groups. The first was the baseline group used to establish the impact of GNSS RO data itself: NoRO removed all GNSS RO data, CTL was the reference experiment using operational-level GNSS RO data, EXP_20k used approximately 20 000 profiles d^{-1} corresponding to the IROWG-recommended level, and EXP_35k used approximately 35 000 profiles d^{-1}, the ROMEX target level (hereafter referred to as the 20k and 35k levels, respectively). The second group adjusted the observation-error inflation factor based on EXP_35k: EXP_Errinfl43 and EXP_Errinfl53 applied $\alpha = 4.3$ and $\alpha = 5.3$, respectively, and were designed to assess how sensitive the forecast-performance change under 35k data use was to the observation-error configuration. The third group adjusted the refractivity coefficient k_1: EXP_k1red01 decreased k_1 by 0.1 % and EXP_k1add01 increased k_1 by 0.1 %. These sensitivity experiments assessed how a small change in this observation-operator physical coefficient affected forecast performance under large-volume RO data use. This experimental design was motivated by a prior report of observation-operator sensitivity in 35 000 profiles d^{-1} ROMEX assimilation (Lonitz, 2024) and aimed to independently diagnose KIM's observation-operator response under the same data-volume condition.

All experiments were run from 1 September to 30 November 2022. After excluding the spin-up period, the period from 8 September to 30 November 2022 was used for verification. Forecast performance was verified against the ECMWF Integrated Forecasting System (IFS) analysis and radiosonde observations. Because the IFS analysis itself is generated by assimilating a large volume of GNSS RO data, it cannot be regarded as a fully independent verification reference for experiments that add RO data. In particular, in the tropics and oceans where radiosonde observations are sparse, IFS analysis is more dependent on RO data. To mitigate this limitation, we combined IFS analysis verification with independent verification against radiosonde observations (Figs. 5 and 8).

The root-mean-square error (RMSE) was calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2}, \tag{4}$$

The forecast performance change was defined as the RMSE improvement rate of an experiment relative to the reference experiment, as follows:

$$\text{Improvement rate (\%)} = \frac{RMSE_{Ref} - RMSE_{Exps}}{RMSE_{Ref}} \times 100, \tag{5}$$

Here, $RMSE_{Ref}$ is the RMSE of the comparison reference experiment, and $RMSE_{Exps}$ is the RMSE of the experiment being evaluated. A positive value indicates that the RMSE of the evaluated experiment is smaller than that of the reference experiment (i.e., forecast performance improved), whereas a negative value indicates that the RMSE of the evaluated experiment increased (i.e., forecast performance degraded). Accordingly, in the scorecards of this study, positive improvement rates are highlighted in green, and negative improvement rates are highlighted in red.

In this study, regional verification distinguished the Northern Hemisphere (NH, 20°N–90°N), tropics (20°S–20°N), and Southern Hemisphere (SH, 20°S–90°S). This classification was applied consistently in all scorecards (Figs. 4, 6, 9–11; Table 2) and the analysis increment cross-sections (Figs. 12 and A3). The reference experiment in each figure was set according to the purpose of each comparison: NoRO was used as the reference when evaluating the effect of CTL relative to NoRO, and CTL was used as the reference when computing the improvement rates of EXP_20k and EXP_35k in the data-volume-increase experiments. In the additional sensitivity experiments, each experiment's performance was compared against CTL, and as a supplementary analysis, the relative change in the observation-error-inflation and k_1 adjustment experiments was also assessed using EXP_35k as the reference.

Table 1. Summary of the experimental design.

Experiment group	**Experiment name**	**Daily mean profile count**	**Observation-error-inflation factor α**	**k_1 (K hPa^{-1})**	**Purpose**

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Baseline	NoRO	0	–	–	Assess overall contribution of RO data
Baseline	CTL	~8 000	2.5	77.6	Operational baseline
Baseline	EXP_20k	~20 000	2.5	77.6	IROWG-recommended level
Baseline	EXP_35k	~35 000	2.5	77.6	ROMEX target level
Obs. error	EXP_Errinfl43	~35 000	4.3	77.6	Moderate error inflation
Obs. error	EXP_Errinfl53	~35 000	5.3	77.6	High error inflation
Physical parameter	EXP_k1red01	~35 000	2.5	77.522	k_1 −0.1 % sensitivity
Physical parameter	EXP_k1add01	~35 000	2.5	77.678	k_1 +0.1 % sensitivity

3 Results

3.1 Baseline impact of GNSS RO data

We first compared NoRO and CTL to assess the impact of GNSS RO data on KIM forecast performance; this comparison served as the interpretive baseline for subsequent data-volume-increase experiments, that is, the difference between CTL and NoRO represents the basic forecast contribution currently provided by the operational-level GNSS RO data in KIM. Figure 4 shows the RMSE improvement rate of CTL relative to NoRO, verified against the IFS analysis, by variable, level, region, and lead time. In the scorecard, green indicates that the CTL RMSE is smaller than that of NoRO (i.e., improved forecast performance), and red indicates that the CTL RMSE is larger than that of NoRO.

CTL exhibited positive RMSE improvement rates relative to NoRO for most variables and levels. In particular, improvements in temperature and geopotential height forecasts were pronounced in the lower stratosphere and upper troposphere. This finding shows that GNSS RO data play a significant role in constraining KIM's upper-level thermodynamic structure. However, the improvement was not consistent across all regions; degradation appeared at certain lead times in lower- to mid-level geopotential height forecasts in parts of the Northern Hemisphere and in lower-level temperature and mid-level geopotential height forecasts in the tropics. Thus, although the baseline effect of the GNSS RO data was generally positive, its impact varied across region and altitude.

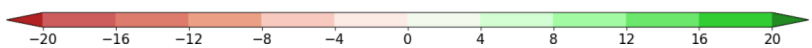
The lower part of Fig. 4 shows the statistical reliability of the RMSE improvement rates. To test whether the mean improvement rate at each lead time was significantly different from zero, we used the following Z-score:

$$Z = \frac{\bar{X}-\mu}{SE}, \ \mathrm{SE} = \frac{\sigma}{\sqrt{n}} \quad (6)$$

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Here, $\bar{X}$ denotes the mean RMSE improvement rate, and μ denotes the assumed population mean of the RMSE improvement rate; in this study, μ was set to zero to test whether each experiment's mean improvement is statistically significant relative to the no-change baseline. We tested whether the mean improvement in each experiment was significantly different from the reference point of no performance change (zero) relative to the reference experiment. In the formulation of the standard error SE, σ is the STDV and n is the sample size. In the scorecards, the size of the triangle reflects the magnitude of the standard error (1SE, 2SE, 3SE), corresponding to confidence levels of 68.3 %, 95.4 %, and 99.7 %, respectively. A filled triangle (▲) indicates a statistically significant improvement (green) or degradation (red) of at least 3SE (≥99.7 % confidence), and an open triangle (△, ▽) denotes a significant change between 2SE and 3SE.

		Global						Northern Hemisphere						Tropics						Southern Hemisphere					
		0	24	48	72	96	120	0	24	48	72	96	120	0	24	48	72	96	120	0	24	48	72	96	120
Q	700hPa	3.88	1.82	1.16	0.69	0.51	0.55	1.75	1.14	0.37	0.22	0.81	0.61	5.45	2.54	1.59	0.79	0.16	0.32	2.40	1.35	1.79	1.64	1.33	1.23
WS	50hPa	21.23	16.04	12.40	10.52	8.70	7.49	3.74	2.33	1.21	0.83	0.73	0.96	29.86	23.43	18.83	16.62	14.86	13.77	7.76	4.46	3.03	2.88	2.29	2.84
WS	100hPa	11.20	6.17	3.79	2.37	1.59	1.53	5.33	3.10	2.17	1.25	1.17	1.20	15.40	8.10	4.68	2.66	1.46	1.15	8.15	5.51	4.32	3.71	2.66	2.39
WS	250hPa	5.30	3.26	2.25	1.50	1.41	1.21	1.89	1.40	1.11	0.55	1.33	1.09	7.01	4.36	2.81	1.69	1.21	0.40	6.28	4.01	2.83	2.14	1.33	1.25
WS	500hPa	2.44	1.98	1.55	1.48	1.44	1.41	1.21	1.03	0.41	0.21	0.90	1.02	3.58	2.88	2.11	2.00	1.81	1.08	2.46	2.08	2.07	2.15	1.58	1.56
WS	850hPa	0.51	0.55	0.55	0.55	0.92	0.97	0.45	0.21	-0.05	-0.24	0.58	0.76	0.70	0.62	0.22	0.08	-0.05	0.26	0.36	0.75	1.15	1.25	1.37	1.20
GPH	50hPa	15.26	11.89	9.31	7.16	5.80	5.41	9.51	7.66	7.39	6.30	5.75	5.42	16.57	13.01	11.15	10.13	9.13	8.66	18.53	14.02	9.02	6.00	4.77	4.84
GPH	100hPa	20.59	15.02	11.11	7.38	5.65	4.53	21.33	15.24	11.32	6.74	4.92	3.53	18.87	13.61	11.44	9.45	8.27	7.34	24.86	18.67	10.76	6.64	5.18	4.47
GPH	250hPa	16.35	6.59	3.40	2.25	2.69	2.17	9.83	3.40	2.22	0.57	2.00	1.09	23.03	12.53	5.40	2.98	2.08	1.40	13.99	6.24	3.76	3.34	2.96	2.74
GPH	500hPa	0.01	0.40	1.30	1.68	2.67	2.14	-1.72	-0.38	0.75	0.34	1.76	1.01	2.72	-0.90	-1.33	-1.88	-0.95	-0.35	-0.05	1.07	1.83	2.68	3.18	2.77
GPH	850hPa	1.73	0.92	1.01	1.05	1.90	1.44	2.26	0.68	-0.07	-0.64	0.89	1.06	2.30	0.97	1.18	-0.29	-0.66	-0.50	1.16	0.91	1.51	2.16	2.45	1.58
T	50hPa	15.36	6.38	3.11	1.77	1.28	1.15	6.21	3.43	2.42	1.21	1.15	1.42	21.91	8.54	3.81	2.24	1.27	0.36	8.67	4.34	2.85	2.08	2.19	2.75
T	100hPa	17.62	9.07	5.85	3.62	2.42	1.89	7.23	4.39	3.34	2.02	2.17	1.65	23.15	10.74	6.77	4.18	2.68	1.91	14.77	9.76	6.95	4.97	2.95	2.55
T	250hPa	14.57	5.90	3.12	1.78	1.62	1.26	6.56	2.70	1.51	0.73	0.86	0.67	14.75	4.50	2.32	1.31	1.13	1.08	20.37	9.79	5.06	2.89	2.31	1.64
T	500hPa	9.45	4.60	2.25	1.44	1.58	1.54	6.35	2.53	0.72	0.02	1.16	1.16	9.05	4.75	2.31	1.65	1.02	0.62	12.24	6.23	3.41	2.45	1.85	1.65
T	850hPa	0.39	0.65	0.66	0.62	0.84	1.17	0.46	0.37	0.21	-0.07	0.60	0.95	-1.30	-0.53	-0.21	-0.00	-0.37	-0.56	1.79	1.82	1.51	1.33	1.46	1.86



https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

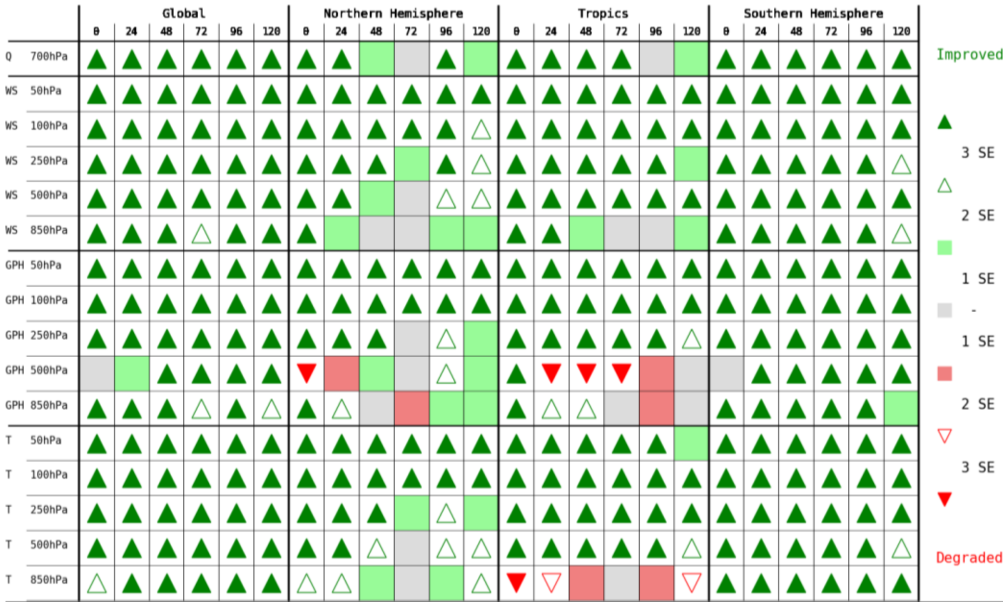


Figure 4. Scorecard of the RMSE improvement rate of CTL relative to NoRO, verified against the IFS analysis, by variable, level, region, and lead time: (upper) improvement rate and (bottom) statistical significance. Green indicates improvement; red indicates degradation.

3.2 Impact of increased RO data volume

Next, we assessed the impact of adding the supplementary RO data provided by ROMEX. The purpose of this section is to determine how forecast performance changes as the GNSS RO data volume increases from the operational level to approximately 20 000 and 35 000 profiles d^{-1}. Figures 5–7 show radiosonde-based background field verification, IFS analysis-based forecast scorecards, and the relationship between data volume and global mean performance, respectively.

Figure 5 shows temperature O−B statistics of the 6-hour background field against radiosonde observations. The mean error (ME) represents systematic bias in the background field, and the STDV represents the scatter in the differences between the radiosonde observations and background field. ME and STDV are calculated using Eq. (7):

$$ME = \frac{1}{N}\sum_{i=1}^{N}(O_i - B_i)\ , \mathrm{STDV} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(O_i - B_i) - ME]^2} \tag{7}$$

Here, B_i is the background field, and O_i is the radiosonde observation.

NoRO showed a relatively large negative temperature bias in the middle-to-upper troposphere, but this bias approached zero in CTL, indicating that operational-level GNSS RO data alone improved the quality of KIM's temperature background field. EXP_20k maintained an ME structure similar to that of CTL, while reducing the scatter at some levels. EXP_35k showed further improvements in parts of the troposphere but exhibited a different bias structure from CTL and EXP_20k in the lower stratosphere. This result indicates that while increased data volume can have a positive effect overall, it can also induce different responses at different altitudes.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Figure 6 shows changes in the forecast performance of EXP_20k and EXP_35k relative to CTL as lead-time–altitude cross-sections, using the IFS analysis as the reference. EXP_20k showed neutral-to-improved performance relative to CTL across most regions, with a particularly positive impact in the upper troposphere and lower stratosphere; however, a moderate degradation was already present in terms of tropical 500 hPa geopotential height (Table 3). EXP_35k, by contrast, showed further improvement for some variables and regions but a different response in the tropics. For lower-stratospheric temperature forecasts, the global-mean improvement was maintained but, in the tropics, it weakened or reversed (Table 2), and a pronounced degradation occurred in tropical 500 hPa geopotential height forecasts. This geopotential height degradation was not confined to the earliest lead times; it was clearest through the 3-day forecast and weakened only gradually through day 5, suggesting a structural response related to forecast error growth rather than a simple initial-condition fluctuation.

Figure 7 and Table 2 summarize the effects of increased data volume in terms of global and regional means. Figure 7 shows how the RMSE and STDV reduction rates for the 100 hPa temperature, relative to NoRO, changed with the data volume. For the global mean, the 100 hPa temperature forecast performance improved as the GNSS RO data volume increased. In particular, the RMSE reduction rate increased markedly from NoRO to CTL and then to EXP_20k, which is consistent with the earlier result that GNSS RO data consistently improved lower-stratospheric temperature forecasts. However, further increasing the data volume from EXP_20k to EXP_35k did not produce a proportionally greater improvement, and for some verification metrics, the improvement rate of EXP_35k was similar to or smaller than that of EXP_20k. This result shows that the relationship between RO data volume and forecast-performance improvement is not linear; as the data volume increases, more observational information becomes available, but the extent to which that information translates into forecast improvement can depend on the assimilation system, observation-error configuration, thinning, the observation operator, and regional background-error characteristics. However, the STDV reduction rate showed no clear improvement with increasing data volume when verified against the IFS analysis, with only a modest increase observed against radiosondes. This result suggests that compared with the IFS analysis, the 100 hPa temperature improvement from the increased RO data volume primarily reflects a reduction in systematic bias.

The regional results in Table 2 demonstrate this nonlinearity. In the Southern Hemisphere, both EXP_20k and EXP_35k showed RMSE reductions relative to CTL, indicating that the effect of adding RO data is relatively stable in regions where observational data are scarce. Improvements in the Northern Hemisphere were smaller but remained positive in both experiments. However, in the tropics, the change observed in EXP_20k was small, and in EXP_35k, both the 100 hPa temperature RMSE and the 500 hPa geopotential height RMSE increased (Table 2, Fig. 11). In particular, the 500 hPa geopotential height RMSE, averaged across the 12–120 h lead times, increased by approximately 8.8 % relative to CTL (Table 3). For the tropical 100 hPa temperature, RMSE increased, while the STDV decreased slightly (Table 2), suggesting that this degradation may be attributable primarily to a change in mean bias rather than to increased scatter. In addition, because the geostrophic balance is weak in the tropics, the absolute forecast error of mid-level geopotential height is smaller than that at mid-latitudes, and the uncertainty in the analysis field used as the verification reference may be relatively larger; therefore, the

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

percentage degradation should be interpreted with caution. These results indicate that while a 35k-level increase in data volume may appear neutral or weakly positive in the global mean, it can produce negative regional effects.

Accordingly, the key finding of the data-volume-increase experiments is that increases up to the 20k level produced a relatively consistent improvement, whereas at the 35k level the marginal improvement diminished in the global mean and regional degradation was observed in the tropics. This finding does not mean that EXP_35k is a failed experiment; rather, it shows that KIM's current assimilation configuration can respond sensitively to ROMEX-level RO data volumes for specific regions and variables. To further investigate this sensitivity, the next section analyzes the observation-error inflation and refractivity-coefficient (k_1) adjustment experiments based on EXP_35k.

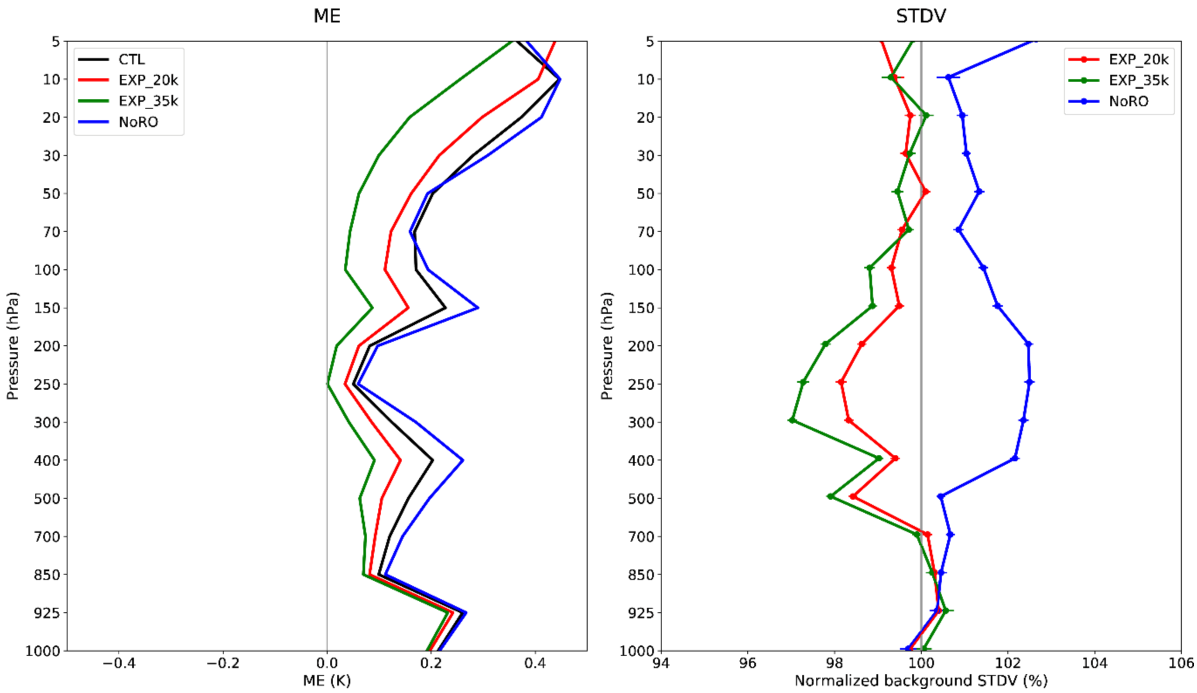


Figure 5. Mean error (left) and STDV (right) of temperature O−B for the 6-hour background field, verified against radiosonde observations, for NoRO, CTL, EXP_20k, and EXP_35k. Averaged over September–November 2022.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

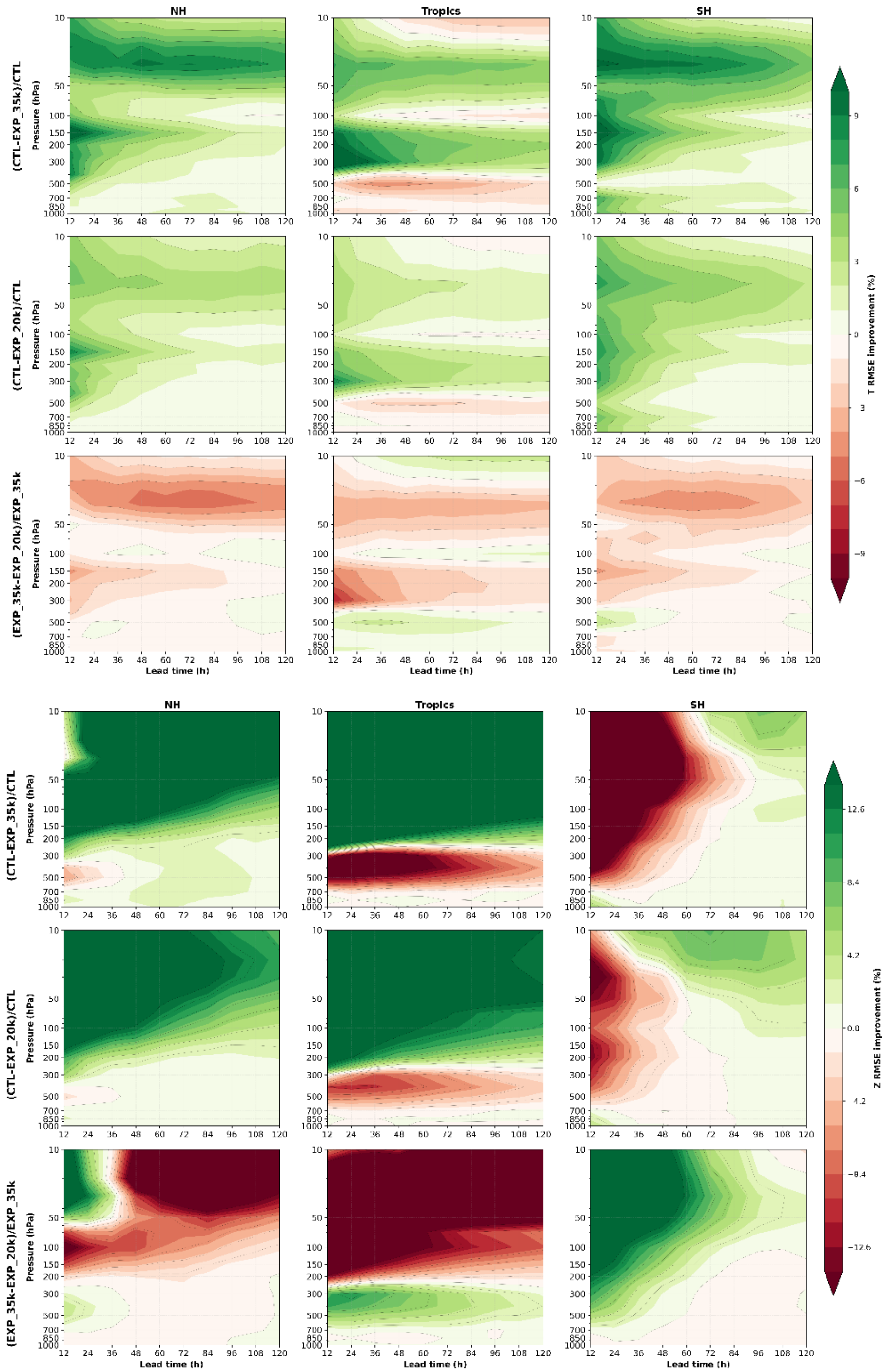
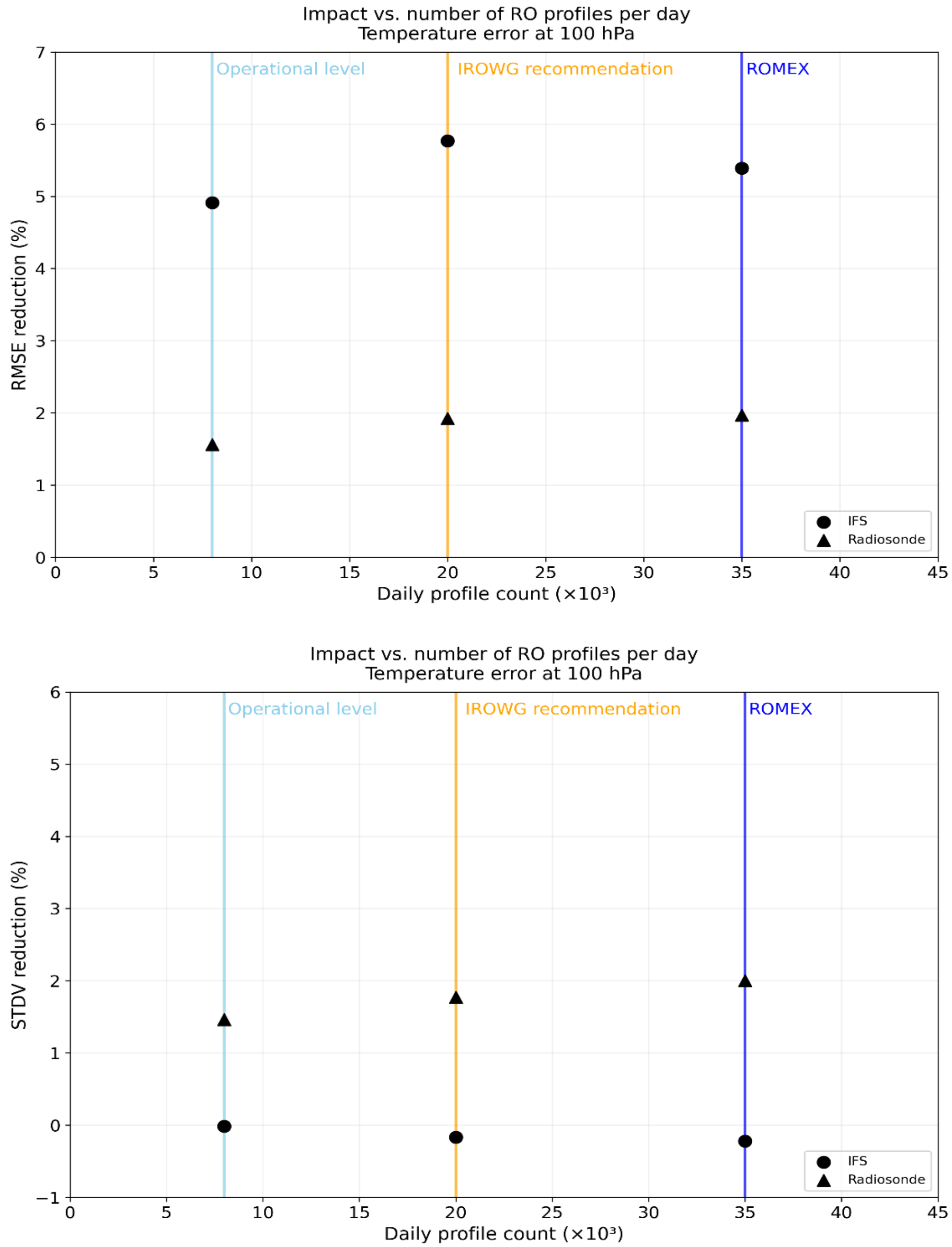


Figure 6. Forecast performance by lead time and region, verified against the IFS analysis, for temperature (upper panels) and geopotential height (lower panels). Rows show (top) EXP_35k relative to CTL, (middle) EXP_20k relative to CTL, and (bottom) EXP_20k relative to EXP_35k. Green indicates improvement and red degradation relative to the reference experiment; in the bottom row, green indicates better performance in EXP_20k than in EXP_35k.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

(a)



(b)

345 **Figure 7. Global-mean reductions (%) in 100 hPa temperature (a) RMSE and (b) STDV as a function of GNSS RO data volume, verified against the IFS analysis and radiosonde observations. Reductions are calculated relative to NoRO. Regional differences are presented in Table 2.**

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Table 2. Regional reductions (%) in 100 hPa temperature RMSE and STDV for EXP_20k and EXP_35k relative to CTL, verified against the IFS analysis and averaged over the 12–120 h forecast lead times for 8 September–30 November 2022. Positive values indicate a reduction in RMSE or STDV relative to CTL (i.e., improved performance), and negative values indicate an increase (i.e., degradation).

Area / Exps	NH		Tropics		SH	
	RMSE	STDV	RMSE	STDV	RMSE	STDV
EXP_20k	1.3	-0.2	0.0	0.0	1.9	-0.1
EXP_35k	1.0	-0.3	-0.8	0.2	2.0	-0.2

Table 3. Tropical 500 hPa geopotential height RMSE (m), verified against the IFS analysis and averaged over the 12–120 h forecast lead times for 8 September–30 November 2022, together with the difference and the percentage increase in RMSE relative to CTL.

Experiment	RMSE (m)	Difference from CTL (m)	Increase relative to CTL (%)
CTL	6.425	–	–
EXP_20k	6.731	+0.306	4.8
EXP_35k	6.989	+0.564	8.8

3.3 Sensitivity to observation-error inflation and refractivity coefficient k_1

To explore probable causes of the regional and vertical performance degradation seen in EXP_35k, two types of sensitivity experiments were carried out: one adjusting the observation-error inflation factor α, and the other adjusting the refractivity coefficient k_1. These experiments were not intended to identify a final optimal configuration for large-volume RO data use but rather to assess the sensitivity of the forecast performance changes seen in the 35k-level assimilation to the observation-error configuration and observation-operator settings.

Figure 8 shows the temperature verification results for the 6-hour background field against radiosonde observations. EXP_Errinfl43 and EXP_Errinfl53 increased the observation-error inflation factor, thereby downweighting the RO observations. In these experiments, the ME increased at some levels, and in EXP_Errinfl53, the STDV increased in parts of the upper troposphere. This trend indicates that while substantially inflating the observation error can mitigate some of the negative effects seen in EXP_35k, it can also weaken the beneficial impact of the RO data.

In the k_1 adjustment experiments, EXP_k1red01 (k_1 decreased by 0.1 %) and EXP_k1add01 (k_1 increased by 0.1 %) exhibited different responses. In terms of ME, the k_1-decrease experiment produced values closer to zero at some levels; however, when the STDV and forecast performance were considered together, the k_1-increase experiment showed more stable results in parts of the middle to upper levels. This result indicates that changing the refractivity coefficient can affect forecast performance, but the effect is more limited than that of observation-error inflation and varies by variable and level.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Figure 9 shows the forecast performance of the observation-error-inflation experiments verified against the IFS analysis. EXP_Errinfl43 and EXP_Errinfl53 tended to mitigate some of the lower-stratospheric and mid-level geopotential height degradation observed in EXP_35k. In particular, RMSE reductions appeared in some regions as the lead time increased, indicating that the impact of the 35k-level assimilation was sensitive to the observation-error configuration. However, observation-error inflation does not always improve forecast performance: at $\alpha = 5.3$, the degradation was further mitigated in some cases, but the improvement in temperature forecasts also tended to weaken. Observation-error inflation is therefore best interpreted not as a solution for large-volume RO data use but as a sensitivity experiment to determine whether the influence of the observations in the current configuration was excessive.

Figure 10 shows the forecast performance of the k_1 adjustment experiments. Both the k_1-decrease and k_1-increase experiments showed a structure broadly similar to EXP_35k overall, but differences appeared in some regions and lead times. In particular, the k_1-increase experiment showed relatively more favorable results than the k_1-decrease experiment for mid-level geopotential height and lower- to mid-level temperature forecasts. This trend indicates that KIM is sensitive to the refractivity coefficient, but adjusting this coefficient had limited ability to resolve the tropical mid-level geopotential height degradation observed in the 35k experiment.

Figure 11 summarizes the experiments by regional average (a) and by an overall value obtained by averaging the RMSE equally across the three regions before computing the reduction relative to CTL (b). The evaluated variables were 100 hPa temperature, representative of the upper-level thermodynamic impact of GNSS RO, and 500 hPa geopotential height, where degradation was pronounced in EXP_35k. Each value denotes the RMSE reduction rate averaged across lead times from +12 to +120 h and calculated relative to CTL. 100 hPa temperature showed improvement relative to CTL for most experiments and regions, which is consistent with the earlier result that increased GNSS RO data volume generally has a positive effect on lower-stratospheric temperature forecasts. By contrast, the RMSE reduction for 500 hPa geopotential height varied substantially among regions. In particular, in the tropics, negative RMSE reduction rates were observed for both EXP_35k and the k_1 adjustment experiments, with EXP_k1red01 exhibiting the largest degradation (Fig. 11). This result shows that mid-level geopotential height forecasts respond sensitively to 35k-level data use, and that decreasing k_1 does not mitigate this sensitivity but rather amplifies the tropical degradation. Because the overall value in Fig. 11(b) weights the three regions equally, the pronounced tropical degradation in 500 hPa geopotential height is partly offset by the extratropical regions; thus, the regional breakdown in Fig. 11(a) should be considered together with the overall value.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

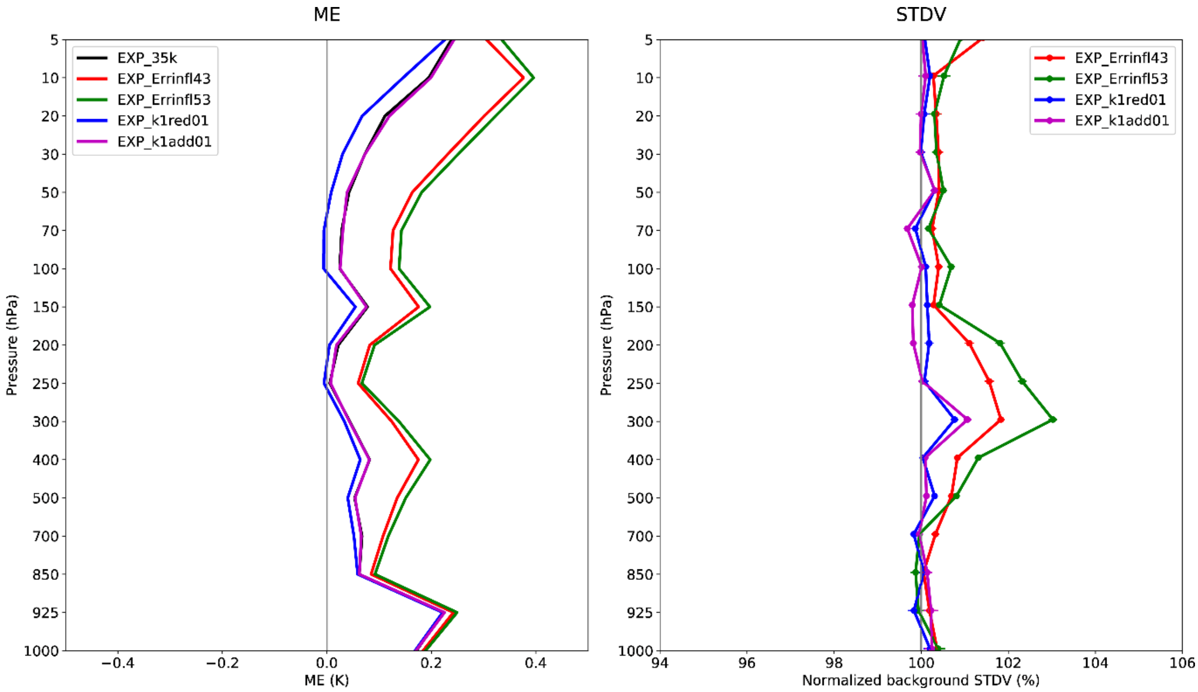
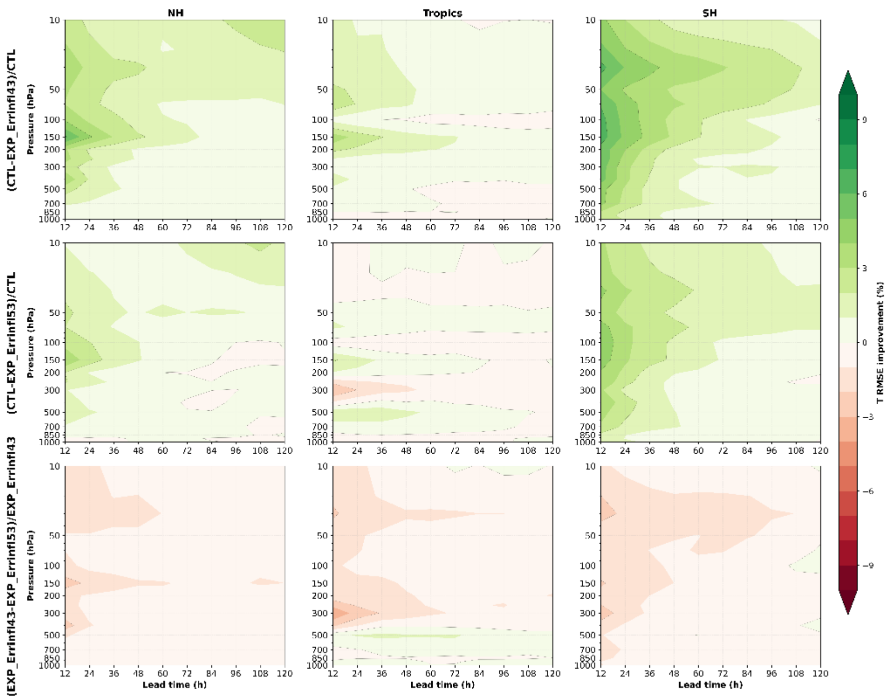


Figure 8. Mean error (left) and STDV (right) of temperature O−B for the 6-hour background field, verified against radiosonde observations, for EXP_35k and the additional observation-error-inflation (EXP_Errinfl43, EXP_Errinfl53) and k_1 adjustment (EXP_k1red01, EXP_k1add01) experiments. Averaged over September–November 2022.



https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

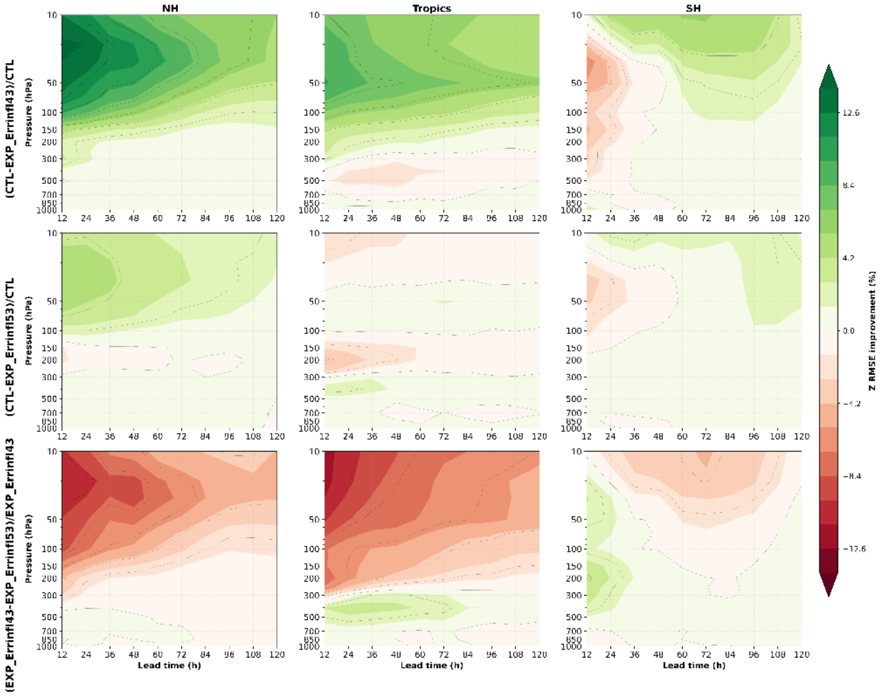
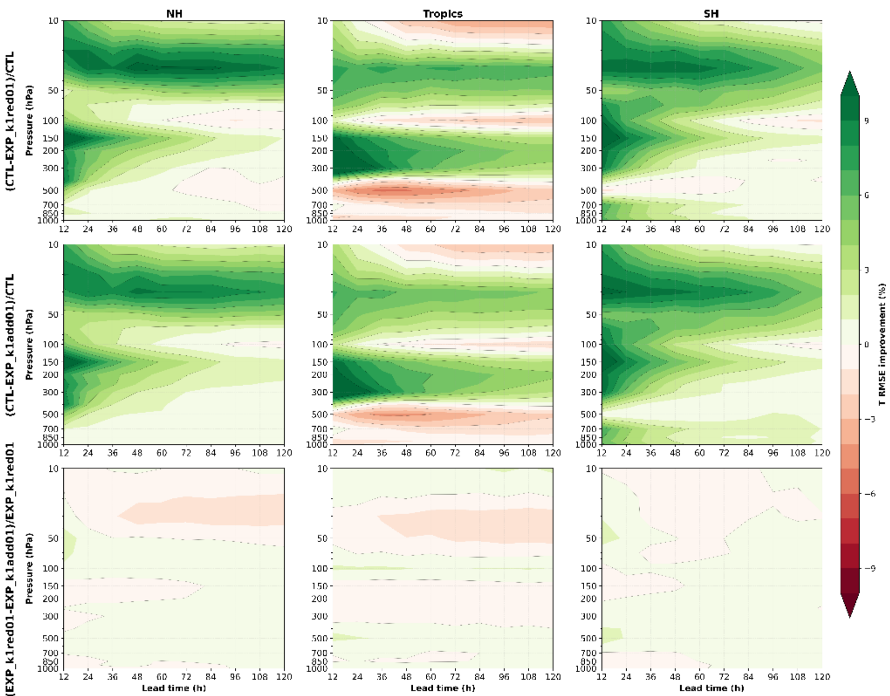


Figure 9. Forecast performance by lead time and region, verified against the IFS analysis, for temperature (upper panels) and geopotential height (lower panels). Rows show (top) EXP_Errinfl43 relative to CTL, (middle) EXP_Errinfl53 relative to CTL, and (bottom) EXP_Errinfl53 relative to EXP_Errinfl43. Green indicates improvement and red degradation relative to the reference experiment; in the bottom row, green indicates better performance in EXP_Errinfl53 than in EXP_Errinfl43.



https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

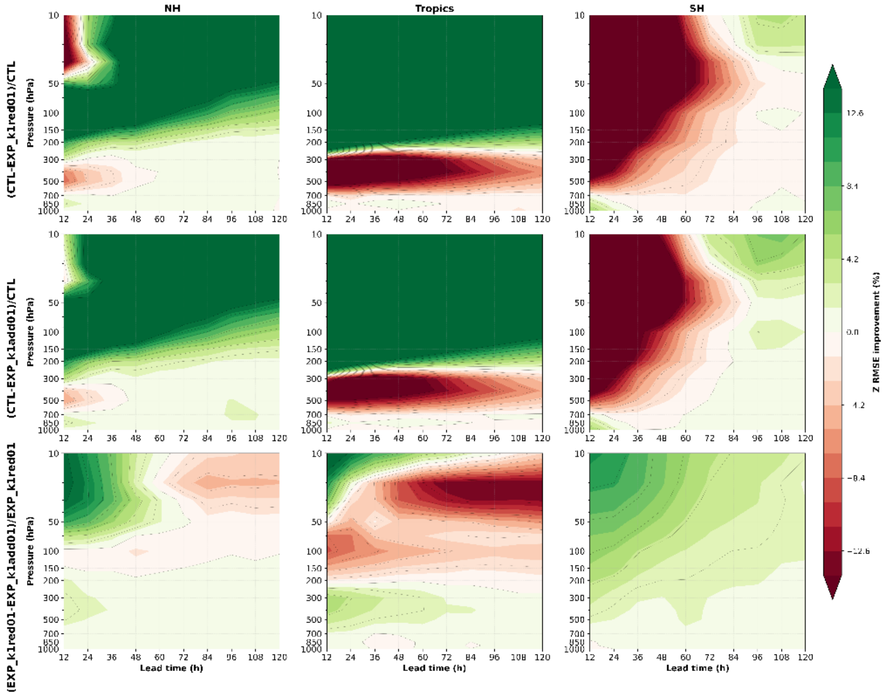
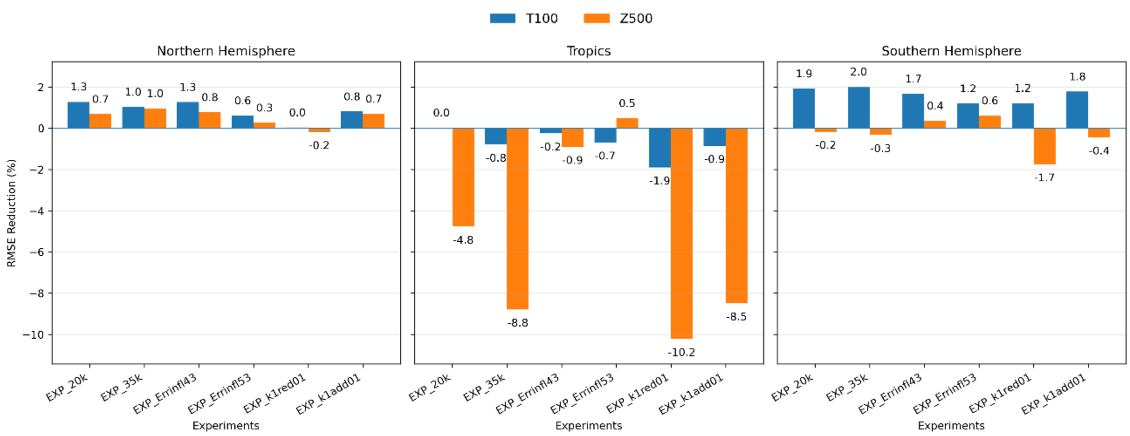


Figure 10. Forecast performance by lead time and region, verified against the IFS analysis, for temperature (upper panels) and geopotential height (lower panels). Rows show (top) EXP_k1red01 relative to CTL, (middle) EXP_k1add01 relative to CTL, and (bottom) EXP_k1add01 relative to EXP_k1red01. Green indicates improvement and red degradation relative to the reference experiment; in the bottom row, green indicates better performance in EXP_k1add01 than in EXP_k1red01.

(a)



https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

(b)

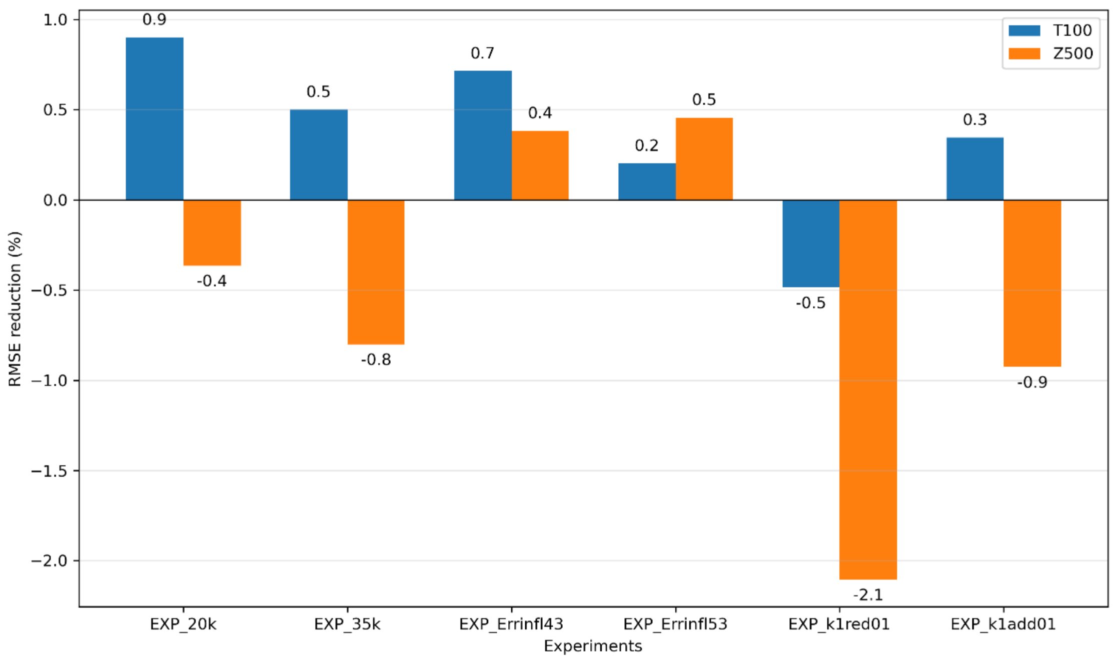


Figure 11. RMSE reductions in 100 hPa temperature (T100) and 500 hPa geopotential height (Z500) relative to CTL for each experiment. (a) Regional mean RMSE reductions for the Northern Hemisphere (20°N–90°N), Tropics (20°S–20°N), and Southern Hemisphere (20°S–90°S). (b) Mean RMSE reduction relative to CTL, computed from RMSE values averaged equally across the three regions (not the average of the regional reductions in (a)). The reductions are averaged over forecast lead times from 12 to 120 h. Positive values indicate improved forecast performance relative to CTL.

4 Discussion

This study examined the baseline contribution of GNSS RO data within the KIM assimilation system, the change in forecast performance with increasing data volume, and the sensitivity that emerged when assimilating ROMEX-level, large-volume RO data. GNSS RO data had an overall positive effect on KIM forecast performance, but the improvement did not increase linearly with increasing data volume. In particular, at approximately 20 000 profiles d^{-1}, the improvement was relatively stable, whereas at 35 000 profiles d^{-1}, the global-mean lower-stratospheric temperature improvement was maintained, but the tropics showed a weakening of the 100 hPa temperature improvement and a pronounced degradation in mid-level geopotential height. This trend indicates that when using large-volume RO data, both data volume and factors in the assimilation system, such as the observation-error configuration, thinning, the observation operator, and the background error covariance, must be considered.

4.1 Baseline impact and nonlinear response to RO data volume

A comparison between NoRO and CTL results shows the basic contribution of GNSS RO data to KIM forecasts. CTL exhibited RMSE reductions relative to NoRO for most variables and levels, with pronounced improvements in temperature and

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

geopotential height forecasts in the lower stratosphere and upper troposphere. This observation implies that GNSS RO data act as a key observation type that constrains KIM's upper-level thermodynamic structure, consistent with the result that the O−B statistics of RO bending-angle observations are most stable around 15–30 km.

In the data-volume increase experiments, this positive effect remained relatively stable up to EXP_20k. EXP_20k showed a neutral-to-improved performance relative to CTL across most regions and levels, with a stable performance gain, particularly for 100 hPa temperature forecasts. This result suggests that RO data volume close to the IROWG-recommended level can be effectively accommodated in the current KIM configuration, although the tropical 500 hPa geopotential height already showed a moderate degradation at this data volume (Table 3). In EXP_35k, however, the further increase in data volume did not translate into further improvement for all variables: the global-mean positive effect on 100 hPa temperature forecasts was maintained, but in the tropics, the 100 hPa temperature improvement weakened (Table 2), and the 500 hPa geopotential height showed a pronounced RMSE increase, a degradation that was clearest from the early lead times through the 3-day forecast and weakened only gradually through day 5. Therefore, it is more appropriate to interpret the EXP_35k result not as invalidating the effectiveness of large-volume RO data, but as showing that the current KIM assimilation configuration can respond sensitively to 35k-level high-density RO data in specific regions and for specific variables.

Notably, in this experimental design, the increase in data volume and change in satellite composition occurred simultaneously. As a substantial portion of the data added in EXP_35k originates from specific satellite constellations such as SPIRE and Yunyao, it is difficult, from the present results alone, to distinguish whether the tropical mid-level geopotential height degradation reflects the assimilation system's response to increased observation density or characteristics specific to particular satellite constellations. Indeed, differences in processing specifications, such as the GNSS clock data rate, have been reported to affect satellite-specific bending-angle characteristics (Padovan et al., 2025). The fact that the new satellite data exhibited vertical statistical characteristics similar to those of existing operational data in the O−B statistics of Fig. 2 (apart from the lower troposphere) makes the latter possibility less likely; nonetheless, potential bias arising from the spatial distribution of the observations is a separate issue that should be confirmed through constellation-denial experiments.

This nonlinear response may be attributed to several factors. Even as the data volume increases, the effective independence of observations may not increase proportionally. When a diagonal observation-error covariance matrix is used, the information content of densely packed observations may be overestimated. In addition, because the one-dimensional bending-angle observation operator does not explicitly account for horizontal inhomogeneity, representativeness errors or observation operator errors may become more important in regions with large horizontal and vertical variabilities, such as the tropics. To use 35k-level RO data reliably, the assimilation configuration must be reviewed by considering not only the data volume itself but also the data density, constellation-specific characteristics, and spatial distribution of observation influence.

4.2 Sensitivity to observation-error inflation

Observation-error inflation experiments were performed to determine the sensitivity of the performance changes observed in EXP_35k to the relative weighting of RO observations. Increasing the observation-error inflation factor reduces the

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

assimilation weight of individual RO observations; therefore, EXP_Errinfl43 and EXP_Errinfl53 can be regarded as sensitivity experiments testing whether the influence of the observations was excessive in EXP_35k.

The experiments revealed that observation-error inflation mitigated some of the 500 hPa geopotential height degradation observed in EXP_35k. In particular, EXP_Errinfl43 showed an intermediate response, largely maintaining a 100 hPa temperature improvement while partially mitigating mid-level geopotential height degradation. EXP_Errinfl53, in some cases, mitigated the geopotential height degradation more clearly but also showed the potential to weaken the upper-level temperature improvement provided by the RO data (Fig. S1).

This result does not imply that uniformly downweighting RO observations is optimal for large-volume RO data use. Rather, it should be interpreted as a diagnostic result showing that the 35k assimilation outcome is sensitive to observation-error configuration and observation weighting. In future work, selective adjustments such as level- and latitude-dependent observation-error tuning, adaptive thinning based on observation density, incorporation of satellite-specific quality characteristics, and consideration of the observation-error correlation structure are likely to be more appropriate than simple uniform inflation. However, because the inflation factor in this study was applied uniformly across all levels, the level- and region-dependent effects implied by the opposing responses observed here (mitigation of mid-level geopotential height degradation versus weakening of upper-level temperature improvement) could not be directly assessed using the present experimental design. This limitation needs to be addressed in future studies. In addition, this study does not quantify the extent to which the inflation adjustment reduced the effective observational information (e.g., degrees of freedom for signal, DFS). It is therefore difficult, based on this study's diagnostics alone, to determine the extent to which performance changes resulting from observation-error inflation are attributable to reduced effective information content. This issue warrants quantitative assessment in future work.

4.3 Diagnosis of the tropical degradation: analysis increments

To diagnose whether the source of the performance degradation seen in the data-volume-increase experiments lies in the assimilation step or in the forecast process, we compared the analysis increments (A−B) of geopotential height for the three experiments (CTL, EXP_20k, EXP_35k) as latitude–pressure cross-sections (Fig. 12a). Averaged over the full verification period, no clear large-scale differences appeared in the incremental structure between the three experiments at the tropical mid-levels. In the vertical profile averaged over the tropical band (20°S–20°N), the time-mean increment for EXP_35k near 500 hPa was somewhat more negative than that for CTL, but the difference was only approximately 0.1 m (Fig. 12c). By contrast, the mean Z500 RMSE over 12–120 h in the same region increased from CTL to EXP_20k and again to EXP_35k, corresponding to degradations of approximately 0.31 m and 0.56 m relative to CTL (Table 3). This monotonic increase with data volume is substantially larger than the difference seen in the initial increments. This small systematic difference in the initial analysis increment cannot by itself explain the magnitude of the degradation seen in the forecast.

The cycle-to-cycle STDVs of the increments decreased with increasing data volume, in the order of CTL, EXP_20k, and EXP_35k. However, near 500 hPa, EXP_20k and EXP_35k were nearly identical, and the differences between the three

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

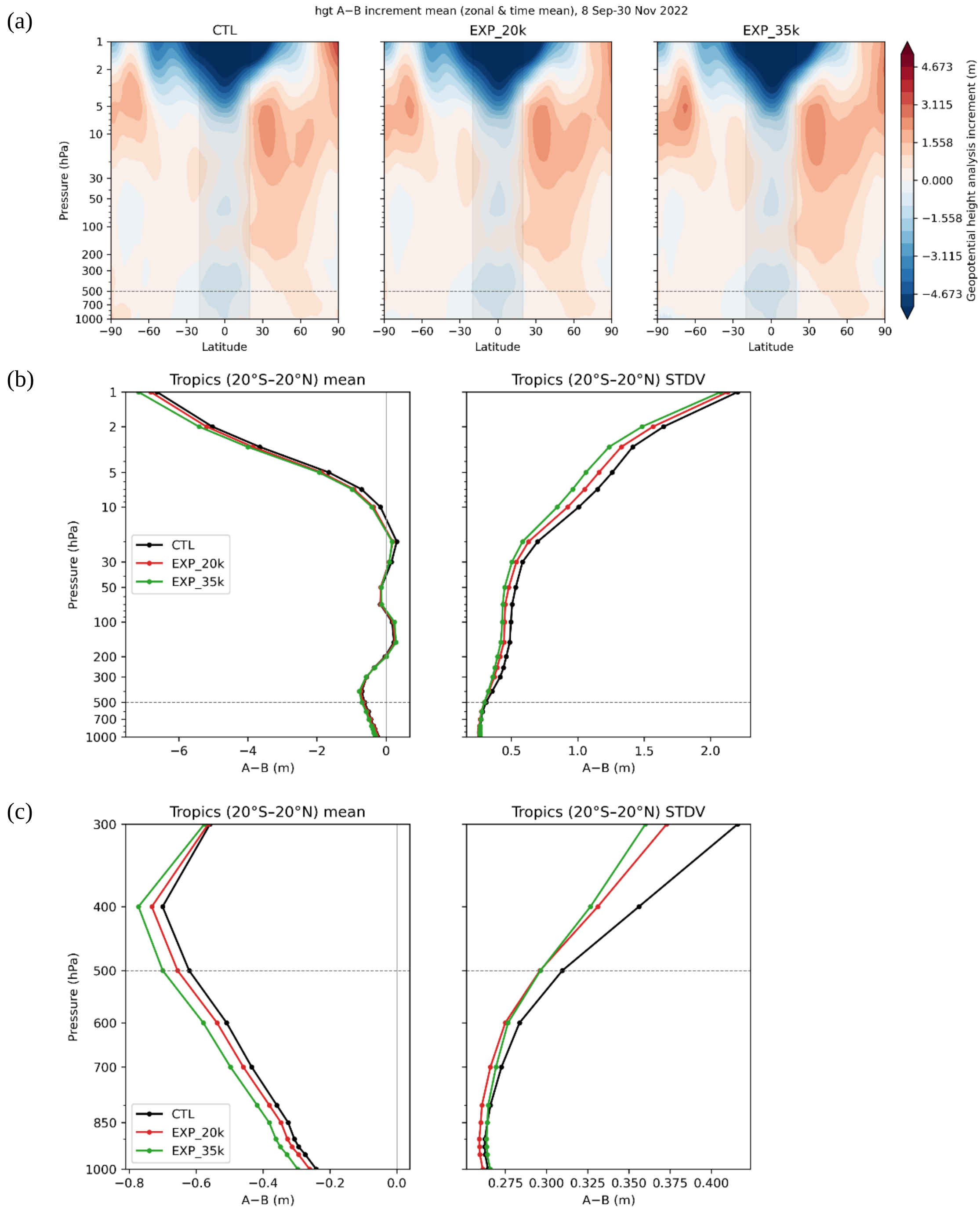
experiments progressively narrowed below 500 hPa and became indistinguishable below 850 hPa (Fig. 12c). The clearest difference in the STDV among the three experiments appeared in the upper stratosphere above 10 hPa (Fig. 12b). In other words, near 500 hPa in the tropics, EXP_20k and EXP_35k showed no clear difference either in the mean of the initial increments or in their cycle-to-cycle variability.

Similarly, the tropical-band-averaged vertical profiles of temperature and zonal-wind increments showed no clear difference among the three experiments in either the increment mean or the cycle-to-cycle STDV near 500 hPa (Fig. A3). In particular, the small difference near 500 hPa in the zonal-wind increment, which is closely coupled to the mass field in the tropics, suggests that EXP_35k's initial analysis field is not in a state of pronounced mass–wind imbalance. The decrease in cycle-to-cycle increment variability with increasing data volume indicates that the observations provide stronger constraints on the background field. Therefore, the performance degradation in EXP_35k is more likely attributable to accumulated differences in the analysis field or to error growth during the forecast process than to an excessively large initial increment resulting from overweighting dense RO observations during assimilation.

The temporal evolution of the degradation differed by variable and level. The degradation in the tropical 500 hPa geopotential height was clearest from the earliest lead times through the 3-day forecast and weakened only gradually through day 5, whereas the tropical 100 hPa temperature improved up to 24 h before gradually worsening at longer lead times (Fig. 6). As noted above, the initial analysis increments of the three experiments showed no substantial differences in the tropical mid-levels, which likely reflects how the initial field constrained by large-volume RO data evolves differently by variable and level during the forecast, rather than the amplitude of the initial increment. The specific mechanism underlying the opposing time dependence observed in these two variables cannot be identified from this study's analysis alone and warrants further diagnosis of early forecast error evolution.

It should also be noted that this diagnosis is based on per-cycle analysis increments (A−B) and therefore does not directly assess differences among experiments in the cycling-accumulated mean analysis-field state. Given that the tropical 500 hPa geopotential height degradation is already apparent from the +12 h forecast, it is possible that even if the per-cycle increments are similar, a systematic difference exists in the accumulated analysis field itself. Future work should, therefore, directly compare the mean analysis-field differences between experiments (e.g., a cross-section of the EXP_35k−CTL analysis-field difference) to confirm this possibility.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026



530 **Figure 12. (a) Latitude–pressure cross-sections (zonal- and time-mean) of the geopotential-height analysis increment (A−B) for CTL (left), EXP_20k (middle), and EXP_35k (right). Shading marks the tropical band; the dashed line marks 500 hPa. (b) Vertical profiles of the tropical-band (20°S–20°N) averaged increment: (left) mean and (right) STDV. (c) Close-up of (b) over 300–1000 hPa. Dashed lines in (b) and (c) mark 500 hPa. Averaged over 8 September–30 November 2022.**

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

4.4 Sensitivity to refractivity coefficient k_1

k_1 adjustment experiments were conducted to assess how a small change in the physical coefficient in the observation operator affected the outcome of the 35k RO data assimilation. Some ROMEX participants reported a positive effect of decreasing k_1 (Lonitz, 2024), but in this study's KIM experiments, increasing k_1 produced relatively more favorable results than decreasing it for some metrics of mid-level geopotential height and lower- to mid-level temperature.

This result suggests that the direction of the effect of k_1 adjustment is not universal but instead depends on each model's background-field bias, assimilation system, observation-operator configuration, and level-dependent weighting structure with which RO observations are assimilated. For KIM in particular, increasing k_1 may have been more consistent with the current background field and observation-operator configurations. However, the effect of k_1 adjustment was more limited than that of observation-error inflation, and it did not directly resolve the tropical 500 hPa geopotential height degradation seen in EXP_35k. In fact, the experiment with decreased k_1 (EXP_k1red01) showed a larger tropical 500 hPa geopotential height degradation than EXP_35k (Fig. 11), indicating that decreasing k_1 worsened, rather than mitigated, the degradation in this region. This trend further supports the view that the optimal direction of k_1 adjustment depends on the background field-bias characteristics of each assimilation system. The k_1 adjustment is therefore best interpreted as a supplementary result illustrating the observation-operator sensitivity of KIM (Fig. S2).

Because KIM does not apply bias correction to the GNSS RO data, the fine adjustment of k_1 acts through the observation operator as an implicit bias correction that shifts the bias structure of the simulated bending angle. In the O−B statistics of Fig. 2, the bending angle shows a positive bias in the lower troposphere (below approximately 5 km) and a weak negative bias above the lower stratosphere; increasing k_1 may have compensated for this residual structure more effectively given KIM's background-field bias. Therefore, the optimal direction for k_1 adjustment may depend not on a universal value of the coefficient itself but on the background field bias characteristics of each assimilation system and whether bias correction is applied. However, because this study did not directly diagnose the level-dependent change in O−B resulting from k_1 adjustment, this interpretation remains at a qualitative level.

4.5 Regional, vertical, and temporal dependence of RO impacts

The impact of the GNSS RO data in this study varied by region, altitude, and lead time. Vertically, the most stable positive effect appeared in the lower stratosphere, consistent with the altitude range where the O−B statistics of RO observations are most stable. Mid-level geopotential height forecasts, particularly those at 500 hPa in the tropics, responded most sensitively to the 35k-level data use. The degradation in this region was partially mitigated by observation-error-inflation experiments but was not clearly mitigated by k_1 adjustment experiments. This finding suggests that tropical mid-level geopotential height degradation may be more closely related to observation weighting, background error covariance, horizontal inhomogeneity, and observation density than to a single physical coefficient of the observation operator. As described in Sect. 4.3, the initial increments of the three experiments showed no clear differences in the tropical mid-levels, indicating that this degradation was unlikely to result from an excessive per-cycle analysis adjustment. However, because similar increments do not necessarily imply similar analysis fields, the present diagnostics cannot determine whether the dominant factor is the accumulation of mean-field differences during cycling or error growth during the forecast.

Regionally, the positive effect of the increased RO data volume appeared relatively stable in the Southern Hemisphere. Therefore, in the Southern Hemisphere and oceans, where the existing observing network is relatively sparse, additional RO data may contribute more directly to improving the analysis field. By contrast, in the tropics, the impact of RO observations appeared more complex, which we interpret as resulting from active convection, a weak geostrophic balance, and large horizontal and vertical variability. Some effects were also more pronounced in the medium range than in the short range, indicating that small differences in the initial analysis field could be amplified during forecast error growth.

Taken together, the effects of the GNSS RO data cannot be adequately assessed using a single global-mean metric. To judge the potential of large-volume RO data, it is necessary to consider not only variables that show consistent improvement, such as the 100 hPa temperature, but also variables that respond sensitively, such as the 500 hPa geopotential height. Through this multivariate, multiregional, and multi-lead-time analysis, this study has highlighted both the potential of large-volume RO data use in KIM and the areas requiring further optimization.

5 Conclusions and Future Work

As part of the ROMEX project, this study assessed the impact of large-volume GNSS RO data on the forecast performance of KIM. We compared NoRO (without GNSS RO data), operational-level CTL, EXP_20k (using approximately 20 000 profiles d^{-1}), and EXP_35k (using approximately 35 000 profiles d^{-1}). We also conducted observation-error inflation and refractivity-coefficient (k_1) adjustment experiments to diagnose the regional and vertical sensitivities that emerged in EXP_35k.

Overall, this study showed that expanding GNSS RO data use in KIM has the potential to improve forecast performance. In particular, increasing the data volume to approximately 20 000 profiles d^{-1} produced relatively consistent improvements under the current configuration. However, at the 35 000 profiles d^{-1} level, lower-stratospheric temperature improvement and tropical mid-level geopotential height degradation occurred simultaneously, indicating that further optimization of the assimilation configuration is needed for the reliable use of large-volume RO data. These results indicate that the impact of large-volume GNSS RO data should be assessed by considering variables, altitude, region, and lead time together, rather than using a single global-mean metric alone. For other operational centers preparing to assimilate ROMEX-level RO volumes, these results suggest that the benefit of large-volume RO data cannot be assumed a priori and should be verified within each system, with particular attention to the observation-error configuration and observation-operator settings rather than data volume alone.

Future work should proceed in three directions. First, level- and latitude-dependent observation-error tuning, incorporation of satellite-specific O−B characteristics, and observation-density-based adaptive thinning should be examined to mitigate negative impacts in specific regions and levels while preserving the information content of RO data. Second, the treatment of

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

observation-error correlations and improvements to the observation operator that account for horizontal inhomogeneity should be examined to represent the effective information content of dense RO observations more realistically. Third, because this study was based on experiments at approximately 25 km resolution, the same experiments should be repeated at operational resolution. Assessing the applicability of bias correction and comparing the results with those of other ROMEX participants should help distinguish KIM-specific responses from sensitivities common to large-volume RO data assimilation more broadly. Such follow-up studies should support the development of assimilation strategies that enable more stable and expanded use of commercial and new GNSS RO data in operational global NWP.

Code and data availability

The ROMEX GNSS RO bending-angle data used in this study were processed and distributed to ROMEX participants by EUMETSAT's ROM SAF under the ROMEX terms and conditions; further information is available at https://irowg.org/ro-modeling-experiment-romex/ (last access: 7 August 2026). The Korean Integrated Model (KIM) source code is operational software developed and maintained by the Korea Meteorological Administration (KMA) and is not publicly available.

Author contributions

E.-H. Kim designed and conducted the experiments, performed the analysis and verification, and prepared the manuscript. H.-W. Chun and Yong Hee Lee contributed to the interpretation of the results and reviewed the manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Special issue statement

This article is part of the special issue "The Radio Occultation Modeling EXperiment (ROMEX): observational quality, processing, and numerical weather prediction (NWP) applications". It is not associated with a conference.

Acknowledgements

We thank the ROMEX steering committee for coordinating the ROMEX project and for their guidance throughout this study. We also thank EUMETSAT's ROM SAF and all ROMEX data providers for the processing and distribution of the GNSS RO observations used in this study. The authors also thank the National Meteorological Satellite Center for making the ROMEX data available in Korea.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Financial support

This research was supported by "Development of Numerical Weather Prediction and Data Application Techniques (KMA2018-00721)" of the Numerical Weather Prediction Center of the Korea Meteorological Administration.

References

Anthes, R. A., Marquardt, C., Ruston, B. and Shao, H.: Radio Occultation Modeling Experiment (ROMEX): Determining the impact of radio occultation observations on numerical weather prediction, Bull. Amer. Meteor. Soc., 105, E1552–E1568, https://doi.org/10.1175/BAMS-D-23-0326.1, 2024.

Aparicio, J. M. and Deblonde, G.: Impact of the assimilation of CHAMP refractivity profiles on Environment Canada Global Forecasts, Mon. Wea. Rev., 136, 257–275, https://doi.org/10.1175/2007MWR1951.1, 2008.

Bonavita, M.: On some aspects of the impact of GPSRO observations in global numerical weather prediction, Q. J. Roy. Meteor. Soc., 140, 2546–2562, https://doi.org/10.1002/qj.2320, 2014.

Bowler, N. E. and Lewis, O.: Experiments with a large number of GNSS-RO observations through the ROMEX collaboration in the Met Office NWP system, Atmos. Meas. Tech., 19, 2479–2505, https://doi.org/10.5194/amt-19-2479-2026, 2026.

Cardinali, C.: Monitoring the observation impact on the short-range forecast, Q. J. Roy. Meteor. Soc., 135, 239–250, https://doi.org/10.1002/qj.366, 2009.

Choi, S.-J.: Structure of eigenvalues in the advection-diffusion equation by the spectral element method on a cubed-sphere grid, Asia-Pac. J. Atmos. Sci., 54, 293–301, https://doi.org/10.1007/s13143-018-0020-4, 2018.

Healy, S. B. and Thépaut, J.-N.: Assimilation experiments with CHAMP GPS radio occultation measurements, Q. J. Roy. Meteor. Soc., 132, 605–623, https://doi.org/10.1256/qj.04.182, 2006.

Hong, S.-Y. and Jang, J.: Impacts of shallow convection processes on a simulated boreal summer climatology in a global atmospheric model, Asia-Pac. J. Atmos. Sci., 54, 361–370, https://doi.org/10.1007/s13143-018-0013-3, 2018.

Kim, E.-H., Chun, H.-W., Son, J. and Lee, Y. H.: A Study on the Adjustment of Global Navigation Satellite System Radio Occultation Observation Errors to Improve KIM Forecast Performance, Asia-Pac. J. Atmos. Sci., 61, 33, https://doi.org/10.1007/s13143-025-00415-6, 2025.

Kim, E.-H., Jo, Y., Chun, H.-W., Ha, J.-H. and Kim, S.: A Study on improvement of the observation error for optimal utilization of COSMIC-2 GNSS RO data, Atmosphere (Korean Meteorological Society), 33, 33–47, https://doi.org/10.14191/Atmos.2023.33.1.033, 2023 (in Korean with English abstract).

Kwon, H., Kang, J.-S., Jo, Y. and Kang, J. H.: Implementation of a GPS-RO data processing system for the KIAPS-LETKF data assimilation system, Atmos. Meas. Tech., 8, 1259–1273, https://doi.org/10.5194/amt-8-1259-2015, 2015.

Kwon, I.-H., Song, H.-J., Ha, J.-H., Chun, H.-W., Kang, J.-H., Lee, S., Lim, S., Jo, Y., Han, H.-J., Jeong, H., Kwon, H.-N., Shin, S. and Kim, T.-H.: Development of an operational hybrid data assimilation system at KIAPS, Asia-Pac. J. Atmos. Sci., 54, 319–335, https://doi.org/10.1007/s13143-018-0029-8, 2018.

https://doi.org/10.5194/egusphere-2026-4767
Preprint. Discussion started: 13 August 2026

Lonitz, K.: Sensitivity to the RO observation operator when assimilating 35,000 radio occultations per day during ROMEX, IROWG-10, UCAR, Colorado, U.S., 12–18 September 2024, https://www.cosmic.ucar.edu/sites/default/files/2024-09/1%20-%20Lonitz_Katrin_2024.09.16.pptx.pdf, 2024.

Marquardt, C.: ROMEX data processing, The First ROMEX workshop, EUMETSAT headquarter, Darmstadt, Germany, 17–19 April 2024, https://cdn.eventsforce.net/files/ef-xnn67yq56ylu/website/61/565d7153-abac-414f-92dc-5466867616fc/20240417_13_marquardt_et_al_eumetsat_romex.pdf, 2024.

Padovan, S., von Engeln, A., Paolella, S., Andres, Y., Galley, C. R., Notarpietro, R., Rivas Boscan, V., Sancho, F., Martin Alemany, F., Morew, N. and Marquardt, C.: Observed impact of the GNSS clock data rate on radio occultation bending angles for Sentinel-6A and COSMIC-2, Atmos. Meas. Tech., 18, 3217–3228, https://doi.org/10.5194/amt-18-3217-2025, 2025.

Privé, N. C., Errico, R. M. and Akkraoui, A. E.: Investigation of the potential saturation of information from global navigation satellite system radio occultation observations with an observing system simulation experiment, Mon. Wea. Rev., 150, 1293–1316, https://doi.org/10.1175/MWR-D-21-0230.1, 2022.

Reiter, A., Danzer, J. and Steiner, A. K.: The potential of GNSS radio occultation data for the analysis of the tropical width: a comparison with reanalyses. Weather Clim. Dynam., 6, 1683–1698, https://doi.org/10.5194/wcd-6-1683-2025, 2025.

Shao, H., Foelsche, U., Mannucci, A. J., Azeem, I., Bowler, N. E., Braun, J. J., Lonitz, K., Marquardt, C., Steiner, A. K., Ruston, B. and Vergados, P.: Advances in GNSS-Based Remote Sensing for Weather, Climate, and Space Weather: Missions, Applications, and Emerging Techniques, Bull. Amer. Meteor. Soc., 106, E1471–E1477, https://doi.org/10.1175/BAMS-D-25-0138.1, 2025.

Shin, S., Kang, J.-H., Chun, H.-W., Lee, S., Sung, K., Cho, K., Jo, Y., Kim, J.-E., Kwon, I.-H., Lim, S. and Kang, J.-S.: Real data assimilation using the local ensemble transform kalman filter (LETKF) system for a global non-hydrostatic NWP model on the cubed-sphere, Asia-Pac. J. Atmos. Sci., 54, 351–360, https://doi.org/10.1007/s13143-018-0022-2, 2018.

Sokolovskiy, S.: Effect of superrefraction on inversions of radio occultation signals in the lower troposphere, Radio Sci., 38, 1058, https://doi.org/10.1029/2002RS002728, 2003.

Xie, F., Syndergaard, S., Kursinski, E. R. and Herman, B. M.: An approach for retrieving marine boundary layer refractivity from GPS occultation data in the presence of superrefraction, J. Atmos. Ocean. Technol., 23, 1629–1644, https://doi.org/10.1175/JTECH1996.1, 2006.

Zhang, H., Shao, H., Ruston, B. and Braun, J.: Impact study of increased radio occultation observations during the ROMEX period using JEDI and the GFS atmospheric model, Atmos. Meas. Tech., 18, 6167–6184, https://doi.org/10.5194/amt-18-6167-2025, 2025.