1. Introduction

Water vapor in the atmosphere has high spatial and temporal variability and is interrelated with the thermodynamics of the atmosphere. Thus, water vapor is an important factor for predicting severe weather phenomena, such as localized heavy rainfall and heavy snowfall. Water vapor information from satellite observation data has often been utilized to fill the observation gap from radiosondes. In recent years, ground-based global navigation satellite system (GNSS) data, which have high spatial and temporal resolutions, have been used to rapidly provide vertically integrated atmospheric water vapor data that can help overcome the spatial and temporal limitations of radiosonde observations. The GNSS zenith total delay (ZTD) observation data can affect the lower and middle humidity fields and show a positive effect on the accuracy of predictions of localized heavy rainfall (Bennitt and Jupp 2012; de Haan 2013; Guerova et al. 2016; Sánchez Arriola et al. 2016).

The GNSS ZTD observation data are affected by various sources of uncertainty, leading to observation errors that must be estimated for data assimilation.

Characteristics and Effects of Ground-Based GNSS Zenith Total Delay Observation Errors in the Convective-Scale Model

Eun-Hee KIM, Eunhee LEE, Seung-Woo LEE, and Yong Hee LEE

Numerical Modeling Center, Korea Meteorological Administration, Republic of Korea

(Manuscript received 12 October 2018, in final form 11 July 2019)

Abstract

In this study, we evaluated the impacts of the revised observation error on ground-based global navigation satellite system (GNSS) zenith total delay (ZTD) data in the data-assimilation system of the Korea Meteorological Administration’s 1.5 km convective-scale model. Out of the 100 total stations on the Korean Peninsula, 40 ground-based GNSS data stations were assimilated using three-dimensional variational (3D-Var) data assimilation. The ZTD observation errors were diagnosed for each station using a posteriori methods, giving errors with a variety of spatial and temporal characteristics. These station-specific error data were then implemented using the data-assimilation system, and their impacts were evaluated for a one-month period in July 2016.

The root mean square error (RMSE) of the relative humidity in the lower troposphere was found to improve for the period from T+0 to T+36 hours, when using GNSS data. Replacing the errors used in the previous model with the average diagnosed errors also provided better results, but they were not as good as the results obtained using station-specific errors. The observation error is closely related to precipitable water vapor (PWV); therefore, correction values reflecting seasonal characteristics should be applied. In addition, the quantitative precipitation forecasts were improved in all experiments using GNSS data, but the effects were minimal.

Keywords
convective-scale model; ground-based GNSS; zenith total delay assimilation; observation error

Citation
Kim, E.-H., E. Lee, S.-W. Lee, and Y. H. Lee, 2019: Characteristics and effects of ground-based GNSS zenith total delay observation errors in the convective-scale model. J. Meteor. Soc. Japan, 97, 1009–1021, doi: 10.2151/jmsj.2019-056.
These errors include measurement, observation operator (forward model), representativeness, and quality control error (Bormann and Bauer 2010; Bormann et al. 2010; Waller et al. 2015). In addition to device-based errors, observation errors also occur during data processing because of the mapping functions that convert the slant path signals of multiple satellites into the zenith direction (Bennitt et al. 2017; Lindskog et al. 2017).

Observation errors can be estimated in several ways: Chun et al. (2015) combined all factors that lead to error or uncertainty, whereas Desroziers et al. (2005) and Hollingsworth and Lönnberg (1986) applied statistical methods, based on the results of data assimilation, to diagnose errors. Generally, errors in non-synoptic data use the standard deviation of the innovation, which is determined as the difference between the observation and model data (Boniface et al. 2009; Mahfouf et al. 2015). In the United Kingdom (UK) Met Office convective-scale forecasting model, almost no difference was found between the standard deviation of the innovations and the observation errors calculated by Desroziers et al. (2005), and these results were very similar for the summer and winter seasons (Bennitt et al. 2017). Furthermore, Poli et al. (2007) found that the overall ZTD observation errors for the entire European region in the winter and summer were 10 and 20 mm, respectively, and the ratios of the background error to the observation error were 1.0 and 0.8, respectively. Thus, the observation error has been observed to demonstrate different characteristics for each region and season.

A convective-scale, data-assimilation study over Korea improved the accuracy of precipitation pattern predictions and quantitative precipitation forecasts for heavy summer rains in August 2014 (Lee et al. 2015). Kim et al. (2017) used the method proposed by Desroziers et al. (2005) to calculate ZTD observation errors of 18–25 mm, at each of 15 stations in the same model, to analyze localized heavy rainfall cases from August 2014. The results showed that the location of precipitation was more accurately simulated when station-based specific errors were applied instead of the constant ZTD error of 6 mm, which is typically used for ZTD assimilation by the UK Met Office.

The Korea Meteorological Administration operates its own convective-scale forecast model, and its prediction area was expanded from the Korean Peninsula to the entire East Asian region in June 2016. As such, re-evaluating the impacts of GNSS data assimilation on the quality of the initial conditions and the performance of the model is necessary. In this study, the ZTD observation errors were calculated using several different methods, the error characteristics were analyzed, and the effects of the models on the observation error were evaluated. The remainder of this paper is presented as follows: Section 2 describes the numerical model and experimental design used to evaluate the GNSS ZTD data and their effect on the predictions. Section 3 describes the observation errors for various calculation methods. Section 4 analyzes the effect of observation errors in the GNSS data on model predictions. Finally, Section 5 summarizes the results and related conclusions.

2. Data and numerical model

2.1 GNSS ZTD data

By the time GNSS radio signals are collected by ground-based receivers, they have been refracted by atmospheric vapor. A mapping function is usually used to convert the slant delay to ZTD. The ZTD observation data can show atmospheric refractivity (which is a function of atmospheric pressure, temperature, and water vapor pressure) as a vertically cumulative value. The signal delay that occurs between the satellite and receiver provides water vapor information that has been vertically accumulated in the troposphere (Bevis et al. 1992). The ZTD observation data are the sum of the zenith hydrostatic delay, caused by dry atmospheric gases and aerosols, and the zenith wet delay, caused by water vapor. The observation data are distributed within a range of approximately 2–3 m. The Korean ground-based GNSS observation network is comprised of 100 stations and managed by four organizations: the Korea Astronomy and Space Science Institute (KASI), National Geographic Information Institute (NGII), National Meteorological Satellite Center (NMSC), and National Maritime Positioning, Navigation, and Timing Office. In this study, we used hourly data from 40 stations produced using ground-based GNSS data processing software (Bernese v5.0) at the NMSC, and data from 671 automatic weather stations (AWSs) were used (precipitation and humidity) for verification (Fig. 1).

To investigate the accuracy of the ZTD observations, the data were converted to precipitable water and compared with the values calculated from the radiosonde data. Figure 2 shows the distribution of the precipitable water calculated at two GNSS stations, Suwon (SUWN) and Kwangju (KWNJ) (Fig. 1), and the precipitable water calculated from the radiosonde data collected nearest to each of these stations over a one-month period in July 2016. The distance between the radiosonde station and the GNSS station was
20 km for SUWN and 10 km for KWNJ. The GNSS and radiosonde data have correlations greater than 0.9 for SUWN and 0.8 for KWNJ, indicating that the data correspond well. For the SUWN GNSS data, the amount of precipitable water tended to be 4.6 mm less than the radiosonde value, and for the KWNJ GNSS data, it tended to be 0.5 mm greater. The root mean square errors (RMSEs) for the SUWN and KWNJ stations were nearly identical for the summer season, at 6.7 and 6.8 mm, respectively. These values are similar to the bias of −5.5 mm and the RMSE of 7.0 mm found at the SUWN GNSS station in the summer of 2014 by Kim et al. (2015).

2.2 Numerical weather prediction and experimental design

The convective-scale model at the KMA, named the Local Data Assimilation and Prediction System (LDAPS), is operated for short-range weather forecasts in East Asia, including the Korean Peninsula. From the Unified Model (UM) version 10.1, several of the physical dynamics were modified and we renamed the model as UM version 10.1k. The lateral boundary uses forecasts from the global KMA model, and the LDAPS has a variable grid system with an inner fixed grid of 1.5 km horizontal resolution and an outer grid of 4 km resolution (Fig. 1); the analysis cycle system uses a first guess at appropriate time (FGAT; Lorenc and Rawlins 2005) 3D-Var data assimilation at 3-h intervals in 70 vertical layers (Table 1), and a 36-h forecast is predicted in 6-h intervals. Observation data preprocessed by the observation processing system (OPS) are entered into the 3D-Var data-assimilation system with a 3-km resolution. To supplement the
3D-Var, the analysis increment is gradually entered as an input to mitigate the shock to the model, and an incremental analysis update (IAU) is used to increase the effect of each increment (Dixon et al. 2009). In addition, radar reflectivity data are assimilated into the model through the latent heat nudging method.

To explore the impacts of the GNSS data and observation errors on the performance of the LDAPS model forecasts, the four experiments described in Table 2 were conducted in July 2016. Conventional data describing the surface, radiosonde, aircraft, SCAT wind, and radar radial velocities were assimilated in all experiments. The control experiment (CNTL) did not use GNSS data. The other three experiments used the standard deviations of innovation from a previous model (9 mm at all stations; EXP1), the observation error for each station (calculated using the method of Desroziers et al. (2005); EXP2), and the monthly average observation error at each of the stations in EXP2 (EXP3).

Table 2. Observation errors for the sensitivity experiments of GNSS ZTD.

| Experiment | Observation Error (Method, Approximate Values) |
|------------|-----------------------------------------------|
| CNTL       | No GNSS                                       |
| EXP1       | Standard deviation of innovation, 9 mm        |
| EXP2       | Desroziers, station-dependent 18–30 mm        |
| EXP3       | Desroziers, 23.8 mm                           |

3. **GNSS observation error characteristics**

3.1 Quality control

To ensure the quality of the GNSS data, data points were removed when the ZTD innovation was greater than 55 mm, and the data were not used if the difference between the altitude of the model surface and receiver exceeded 300 m (Bennitt and Jupp 2012). To exclude the effect of time on the observation data, the data that were collected closest to the analysis time were used. As a result, innovations for the CNTL and EXP1 exhibited a normal distribution, except for certain stations, due to technical issues with their equipment. Figure 3 compares the time series for the observed ZTD and the model ZTD at Daejeon Station, which is an international standard observatory. The two values show similar distributions, and their differences are mostly distributed within ± 0.1 m, with the exception of two outliers. This finding indicates that the quality of the observed ZTD results is equivalent to that of international standard observations.

3.2 Bias correction

The 3D-Var method assumes that the observation data has no bias. Thus, the data from each station must be bias-corrected prior to data assimilation. Following Bennitt and Jupp (2012), the mean value of the innovation over 28 days, at each station, was used for bias correction. Whenever the model was upgraded or whenever seasonal changes occur, static bias correction methods were used to manually perform a renewal (Bennitt and Jupp 2012). Otherwise, an automatic renewal of the innovation, taking place 30 days before...
the data-assimilation time (de Haan 2013), was used. Outliers in the GNSS observation data were removed, and a static bias correction was performed for the one-month period under investigation. The difference in the mean ZTD bias of the 40 stations before and after the bias correction was 10.3 mm, which is similar to the 10 mm difference found by Yan et al. (2009).

### 3.3 Observation error characteristics

To determine the observation error of each station, the standard deviation of the innovation was calculated for July 2016 using the conventional methods for calculating the observation error in accordance with Hollingsworth and Lönnberg (1986) and Desroziers et al. (2005). The Hollingsworth and Lönnberg (1986) method is efficient for use with a sufficiently dense observation network that provides information in a variety of scales. With this method, the observation error is assumed to be separate and, thus, to exhibit no correlation between stations; however, the background error is assumed to have a spatial correlation. The covariance is calculated as the distance between the covariance of the innovation of the station for which the observation error is to be found and the remainder of the stations. With this method, the observation error is assumed to be separate and, thus, to exhibit no correlation between stations; however, the background error is assumed to have a spatial correlation. The covariance is calculated as the distance between the covariance of the innovation of the station for which the observation error is to be found and the remainder of the stations. Then, the value is extrapolated to 0 km to define the background error covariance, and the excluded value is the observation error covariance. The Desroziers et al. (2005) method calculates the observation error covariance by utilizing the statistical values of the innovation and analysis residual; the square root is taken to estimate the observation error through the standard deviation. This method also relies on an intrinsic assumption of the separation of scales and the correlation between the background and observation errors. Considering the existence of possible biases, the standard deviation of the observation error can be expressed by

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - A_i)(O_i - B_i)} - \left( \frac{1}{n} \sum_{i=1}^{n} (O_i - A_i) \right) \left( \frac{1}{n} \sum_{i=1}^{n} (O_i - B_i) \right),$$

where $\sigma_i$ is the standard deviation of the observation error of each station (m; hereafter named as the observation error); $O_i$ is the ZTD observation value (m); $A_i$ and $B_i$ are the ZTD values (m) calculated from the model variables of the analysis field and the background, respectively; and $n$ is the number of observations. This method is widely used, and unlike the Hollingsworth and Lönnberg (1986) method, it uses values from the same stations, providing an advantage by not needing to consider the entire observation network.

Figure 4 shows a time series comparing the ZTD observation errors of each station that were calculated using the three methods: the standard deviation of innovation, the Hollingsworth and Lönnberg (1986) method, and the Desroziers et al. (2005) method. The units were converted to millimeters for visibility. The observation errors found using the standard deviation of the innovation and Hollingsworth and Lönnberg (1986) methods exhibited similar means for 40 sta-
tions of 29.8 and 30.8 mm. Notably, using only 40 stations in this experiment does not comply with the assumption of the Hollingsworth and Lӧnnberg (1986) method; thus, the result may be inaccurate. The mean observation error calculated by the Desroziers et al. (2005) method was 23.8 mm, which was the lowest observation error from all of the methods, and the differences between the maximum and minimum values for each station were small. The observation errors were smaller in the case of using the Desroziers et al. (2005) method because it uses an analysis field that performs data assimilation. However, some stations of the NMSC showed much larger errors than others because their total number of observations was small in July 2016. The recommended approach for the data-assimilation system used in this experiment was a minimum of 168 data points for at least seven days, but more than 150 data points corresponding to 6.25 days were applied for many stations. Therefore, some points, which collected fewer than 150 data, were not used for data assimilation.

In general, ZTD data are used for climate monitoring, as a task in which the long-term ZTD observation error is stable. To examine the seasonal characteristics of the ZTD observation error over a short period of time, the monthly averages of precipitable water, precipitation, and observation error for 2017 were calculated by utilizing the Desroziers et al. (2005) method with the LDAPS and continuous data from 40 GNSS observation stations (Fig. 5). The ZTD observation error was high in July and August, when there was a large amount of precipitation, because the model performance worsens during the humid summer season. As a 3D-Var with a constant B matrix was used, the variations in the model background error are presumably poorly represented, which increases the innovation. Thus, increasing the ZTD observation error to compensate for this effect may be beneficial. Observation errors in the spring and fall were similar and winter had the lowest error distribution. Indeed, in the summer (June–August) and winter (December–February), the mean observation errors were 30 and 14 mm, respectively. Thus, the mean observation error was approximately 2.1 times larger in summer, when a large volume of water vapor was present in the atmosphere, than that in winter, when little water vapor was present. This result agrees with the findings of other studies (Poli et al. 2007; Rohm et al. 2014) that show that ZTD observation data are closely related to atmospheric conditions. The ZTD depends on the refractivity, $N$, which is a function of atmospheric pressure, temperature, and humidity, all of which

---

**Fig. 4.** Estimation of observation errors (mm) for the ZTD data used in the LDAPS for the month of July 2016. The horizontal axis represents each station. The solid, dashed, and dotted line indicates Desroziers et al. (2005), Hollingsworth and Lӧnnberg (1986), and standard deviation of innovation, respectively.

**Fig. 5.** Monthly means of ZTD observation errors (solid line) and PWV values (dashed line) from 40 ground-based GNSS stations and precipitation (gray bars) from the AWS corresponding to each GNSS in 2017.
change seasonally in the mid-latitudes of the Northern Hemisphere. The observation error shows high seasonal variability depending on the precipitable water and precipitation, with a particularly high correlation (0.95) to precipitable water. This finding demonstrates that seasonal characteristics must be considered when the observation error is used to assimilate ground-based GNSS data.

4. Forecast effect in the local model

To examine the effects of the observation errors on the assimilation of ground-based GNSS data, this study examined the difference in specific humidity analysis increments between the CNTL and the other three experiments (EXP1–EXP3) at the initial time when the cycle forecast experiments started and at the 1-h accumulated precipitation times at 671 AWSs (Fig. 6). Figure 6b shows the difference between the analysis increments of the specific humidity of the CNTL and EXP1 at the lowest level. The area where the analysis increment is largest corresponds to the area around the GNSS observation station. Large positive increments in water vapor occurred around the GNSS stations near the southern coast, where more precipitation occurs. The analysis increment shows that water vapor information from the ground-based GNSS data was suitably reflected in the model. We examined the vertical distribution of the analysis increment between each experiment and the control experiment at Jinju Station (JINJ), which showed the largest positive increase (Fig. 6c). As expected, the increment was largest in EXP1 (black), which also showed the smallest observation error. Meanwhile, the increment was smallest in EXP2 (red), which used the observation error for each station calculated using the Desroziers et al. (2005) method. The increase in EXP3 (blue), which used the monthly mean observation error, showed a pattern similar to that of EXP2, but it was slightly larger at every level. These observations confirmed that the magnitude of the increment changes with variations in the observation error.
4.1 GNSS data-assimilation effects

The analysis of the residual histograms (Fig. 7) of the CNTL and EXP1 reveal that the bias and RMSE of EXP1 were lower than those of the CNTL. These findings show that the analysis field that assimilated the GNSS data matches the observation data better than the analysis field that did not perform ZTD assimilation. Thus, the GNSS data-assimilation module properly functioned. In the verification that used the ground-based humidity data of the 671 AWSs, the experiment that added the GNSS data also had a lower RMSE that was maintained from the initial forecast to the 30-h forecast (Fig. 8). The GNSS data showed a maximum improvement effect for a ground-based humidity of 1.3 % at 10 h and an overall improvement of 0.4 % over 36 h.

4.2 GNSS observation error sensitivity

The data from 58 radiosondes included in the LDAPS domain were used to compare the analysis field to the 36-h forecast, at 6-h intervals, in the middle to lower troposphere below 500 hPa, where water vapor is primarily distributed. Figure 9 shows the humidity RMSE improvement rates of EXP1–EXP3 compared with the CNTL, in which a positive value indicates that the experimental RMSEs were lower than those of the CNTL. The improvement rates were calculated using Eq. (2).

\[
\text{Improvement rate} = \left(1 - \frac{\text{RMSE}_{\text{EXPs}}}{\text{RMSE}_{\text{CNTL}}}\right) \times 100. \tag{2}
\]

The lowest humidity RMSE occurred in EXP2, in the analysis field at the lower troposphere below 700

Fig. 7. Histograms of the differences between the observed ZTD and analytical ZTD results for the 40 GNSS stations in six hourly interval analyses, from 0000 UTC on July 1 to 1800 UTC on July 31, 2016 for (a) the CNTL and (b) EXP1.

Fig. 8. Verification using relative humidity from the AWSs in July 2016. Lines indicate the RMSE (%) of the relative humidity for the CNTL and EXP1; bars indicate the improvement rate (%) of EXP1 compared with the CNTL.
hPa; EXP2 also showed the highest improvement rate (1.8 %). For the 6-h forecast time, the improvement rates of EXP1 and EXP2 were the same (3.7 %), and for the 12-h forecast time, the improvement rates showed a trend similar to that of the analysis field. The overall improvement rate of the 18-h forecast’s humidity field was better than that of the 6-h forecast, with improvements in EXP2 and EXP3 of 4.2 % and 4.8 %, respectively. The improvement rates from the analysis time to the 36-h forecast were 0.5 %, 2.0 %, and 1.3 % for EXP1, EXP2, and EXP3, respectively.

The lower level humidity field improvement was the largest in the results that used the newly calculated observation error for each station, reflecting the current model’s characteristics. The use of the monthly average at all stations, which made the observation error easier to apply, did not have a substantial effect. These results reveal that using a small observation error can produce results that are worse than the background error. These findings are similar to those of Benáček et al. (2016), who empirically increased or decreased the observation error to decrease the data’s spatial error correlation.

The Critical Success Index (CSI) and bias in the 1-h accumulated precipitation data from 671 AWSs were examined for the experiments that added GNSS data and for the CNTL, which did not add GNSS data (Fig. 10). Compared with the CNTL, the CSI of EXP1–EXP3 increased at all threshold values, showing that adding GNSS data alone increases precipitation forecast accuracy. No substantial difference was observed in the precipitation forecast performance resulting from the observation error during the overall forecast time. However, EXP2 had the highest mean CSI (by a small amount) for the threshold value of 3–15 mm during the 36-h forecast. In addition, in EXP1–EXP3, the underestimation of precipitation values under 1 mm was improved, compared with that in the CNTL, and the tendency of the CNTL to overestimate precipitation greater than 10 mm was clearly improved. Furthermore, the bias in EXP2 was improved at all threshold values.

5. Summary and conclusions

In this study, South Korea’s ground-based GNSS data were used in a LDAPS via 3D-Var to expand and optimize the use of local observation data. Quality control was performed on the data collected from the 40 stations, and the results confirmed that, with the exception of certain stations, most of the errors show an almost normal distribution and that the quality of the results was as good as that of international standard observations. To make optimal use of the GNSS observation data in the convective-scale forecast model, the observation error characteristics were examined and their effects on the model were

Fig. 9. Experimental RMSE improvement rate (%) for 925–700 hPa relative humidity based on the 36-h forecast for CNTL and radiosonde observations in the July 2016 assimilation cycles.

Fig. 10. (a) CSI and (b) bias score against rain-gage observations for 1-h, accumulated precipitation for EXP1–EXP3 and CNTL forecast runs, in July 2016.
evaluated. The observation error was calculated using three different methods: one method that utilized the standard deviation of the innovation from the existing convective-scale forecast model and two methods (e.g., Hollingsworth and Lonnberg 1986; Desroziers et al. 2005) that utilized the statistical values of the data-assimilation results. The comparison results of these methods showed that the values calculated using the Hollingsworth and Lonnberg (1986) method had no substantial difference from the values derived from the standard deviation of innovation, but the observation errors calculated using the Desroziers et al. (2005) method were on average 6 mm smaller at all stations than those of the other two methods. The observation error and precipitable water showed a similar seasonal variation, and the observation error in summer was approximately 2.1 times larger than that in winter. In contrast with a study using the same observation error for each season and station (Bennitt and Jupp 2017), utilizing different observation errors for different seasons or months was considered to be more appropriate for the Korean climate.

Three one-month experiments were performed on the LDAPS using data from July 2016 for stable quality. The largest (2 %) relative-humidity improvement rate of the experiment, with individual ZTD error characteristics for each station, was exhibited from the analysis time to the 36-h forecast time in the lower atmosphere (925–700 hPa) in summer. For precipitation forecasts, each experiment showed better quantitative precipitation forecast performance than the control experiment. The effect of observation error was not large, but the precipitation forecast performance of the experiment that used the observation error of each station was high for precipitation threshold values below 15 mm, whereas the bias was improved for all precipitation threshold values. By providing the water vapor information at the initial time, via ground-based GNSS data assimilation, the analysis field was made to be more humid, mitigating the precipitation-production delay phenomenon. Thus, the quantitative precipitation forecast performance can be improved by providing realistic water vapor information at the initial time. Indeed, the best forecast performance was obtained when using observation errors that reflected the model’s characteristics and observations for each station, when considering seasonal variation.

In future works, additional research on GNSS data bias corrections in convective-scale forecast models will be conducted and the use of GNSS data to forecast severe weather and improve nowcasting will be examined.

**Appendix A: ZTD observation operator**

The ZTD operator uses model variables from the forecast time of each GNSS station (i.e., the model background) to calculate the ZTD. This function vertically integrates the refractivity, \(N\), which depends on the atmospheric pressure, temperature, and humidity. The refractivity has a large value at the surface and decays exponentially with increasing altitude as follows (Poli et al. 2007):

\[
ZTD = 10^{-6} \int_{z=0}^{z=\infty} N\,dz, \tag{A1}
\]

where \(z\) is the altitude (m) above the surface and \(N\) is the refractivity based on the refractive index \(n\), which is expressed by:

\[
N = k_1 \frac{P_d}{T} + k_2 \frac{P_v}{T} + k_3 \frac{P_d}{T^2}, \tag{A2}
\]

where \(P_d\) is the dry air partial pressure, \(P_v\) is the water vapor partial pressure, \(T\) is the temperature, and \(k_1, k_2,\) and \(k_3\) are the refractivity constants (Bevis et al. 1994). By substituting \(p_d = p - p_v\), into Eq. (A2):

\[
N = k_1 \frac{P}{T} + (k_2 - k_3) \frac{p_v}{T} + k_3 \frac{P_d}{T^2}. \tag{A3}
\]

According to Smith and Weintraub (1953), a negligible error occurs when the second and third terms of Eq. (A3) are lumped within a limited temperature range (−50–40°C). The two terms are divided by \(p_v T^{-1}\), and if \(T\) is 273 K, then the composite coefficient \((k_3)\) is \(3.73 \times 10^4\). Thus, Eq. (A4) can be approximated for use in the LDAPS as an observation operator, as follows:

\[
N = k_1 \frac{P}{T} + k_4 \frac{P_v}{T^2}. \tag{A4}
\]

Figure A1 shows the model levels, variables, and GNSS station for the ZTD calculation. The observation operator starts at the bottom of the model and, between each model level, assumes that the potential temperature and specific humidity are constant to calculate the refractivity at each theta (\(\theta\)) level until the top of the model is reached (Bennitt and Jupp 2012). The Exner pressure on the rho (\(\rho\)) level is given by

\[
\Pi_{\rho_i} = \left( \frac{P_{\rho_i}}{P_{ref}} \right)^{R/c_p}, \tag{A5}
\]

which can be linearly interpolated to the height of the \(i^{th}\) \(\theta\) level as follows:
\[ \Pi_i = \alpha \Pi_i + (1 - \alpha) \Pi_{i+1}, \]  
\[ \Pi_i = \alpha \Pi_i + (1 - \alpha) \Pi_{i+1}, \]  
\[ \alpha = \frac{z_{\rho,i+1} - z_{\theta,i}}{z_{\rho,i+1} - z_{\theta,i}}, \]  
\[ \alpha = \frac{z_{\rho,i+1} - z_{\theta,i}}{z_{\rho,i+1} - z_{\theta,i}}, \]  
where

\[ P_{\theta} = P_{\text{ref}} (\Pi_{\rho \theta})^{c_i/R}. \]  
\[ P_{\theta} = P_{\text{ref}} (\Pi_{\rho \theta})^{c_i/R}. \]

The pressure on the \( i \)th \( \theta \) level is given by

\[ T = \frac{T_c}{1.0 + q_i C_{\text{virtual}}} . \]  
\[ T = \frac{T_c}{1.0 + q_i C_{\text{virtual}}} . \]

The exponential refractivity observation operator \( \text{ZenithDelay}_i \) is given by the following equations (Mark and Gemma 2014):

\[ \text{ZenithDelay}_i = 10^{-6} \int_{z_{\theta,i}}^{z_{\rho,i}} N_i \exp(-c_i(z - z_{\theta,i})) dz \]

\[ \text{ZenithDelay}_i = 10^{-6} \int_{z_{\theta,i}}^{z_{\rho,i}} N_i \exp(-c_i(z - z_{\theta,i})) dz \]

\[ = -10^{-6} \frac{N_i}{c_i} \exp(c_i z_{\theta,i}) \]

\[ = -10^{-6} \frac{N_i}{c_i} \exp(c_i z_{\theta,i}) \]

\[ \times (\exp(-c_i z_{\theta,i}) - \exp(-c_i z_{\theta,i})), \]  
\[ \times (\exp(-c_i z_{\theta,i}) - \exp(-c_i z_{\theta,i})), \]  
where \( \text{ZenithDelay}_i \) is the delay for each layer and is calculated from difference between the heights of the \( \rho \) level,

\[ c_i = \frac{\ln \left( \frac{z_{\rho,i+1}}{z_{\theta,i+1}} \right)}{z_{\theta,i+1} - z_{\theta,i}}. \]  
\[ c_i = \frac{\ln \left( \frac{z_{\rho,i+1}}{z_{\theta,i+1}} \right)}{z_{\theta,i+1} - z_{\theta,i}}. \]  

When the model reaches the final \( \theta \) level, the refractivity is integrated to the top \( \rho \) level, such that:

\[ \text{ZenithDelay}_{\text{final}} = -10^{-6} \frac{N_i}{c_i} \exp(c_i z_{\theta}) \]

\[ \times (\exp(-c_i z_{\theta,i+1}) - \exp(-c_i z_{\theta})). \]  
\[ \times (\exp(-c_i z_{\theta,i+1}) - \exp(-c_i z_{\theta})). \]

After calculating the refractivity on every level, the operator starts at the bottom level and checks if the station height lies below each level. The GNSS station can lie between model levels (Fig. A1b) or below the model surface (Fig. A1c). The method of accounting for these situations is the same. The operator iterates up the levels until reaching

\[ z_{\rho,i+1} > \text{GNSS station height}. \]  
\[ z_{\rho,i+1} > \text{GNSS station height}. \]

At this level, the zenith delay for the partial layer is calculated by

\[ \text{ZenithDelay}_i = -10^{-6} \frac{N_i}{c_i} \exp(c_i z_{\theta}) \]

\[ \times (\exp(-c_i z_{\theta,i+1}) - \exp(-c_i z_{\text{station}})). \]  
\[ \times (\exp(-c_i z_{\theta,i+1}) - \exp(-c_i z_{\text{station}})). \]

Acknowledgments

This research was supported by the Numerical Modeling Center of the Korea Meteorological Administration under grant number (NTIS-1365002970), “Development of Numerical Weather Prediction and Data Application Techniques”. The authors would like to thank the National Meteorological Satellite Center for providing GNSS ZTD observations in near-real time. We also thank Hyeyoung Kim for helpful discussion and advice. The authors are grateful to two
anonymous reviewers for their suggestions and comments, which have helped to improve this manuscript.

References
Abel, S. J., and I. A. Boutle, 2012: An improved representation of the raindrop size distribution for single-moment microphysics schemes. Quart. J. Roy. Meteor. Soc., 138, 2151–2162.

Benáček, P., R. Brožková, A. Bučánek, and A. Trojáková, 2016: Tuning of 3D-Var ALADIN-CZ system for aircraft data assimilation. Advanced Meteorological Information for Aviation, TH01010503, 21 pp. [Available at http://www.rclace.eu/File/Data_Assimilation/2016/rep.pdf.]

Bennitt, G. V., and A. Jupp, 2012: Operational assimilation of GPS zenith total delay observations into the Met Office numerical weather prediction models. Mon. Wea. Rev., 140, 2706–2719.

Bennitt, G. V., H. R. Johnson, P. P. Weston, J. Jones, and E. Pottiaux, 2017: An assessment of ground-based GNSS Zenith Total Delay observation errors and their correlations using the Met Office UKV model. Quart. J. Roy. Meteor. Soc., 143, 2436–2447.

Bevis, M., S. Businger, S. Chiswell, T. A. Herring, R. A. Anthes, and R. H. Ware, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. J. Geophys. Res., 97, 15787–15801.

Bevis, M., S. Businger, S. Chiswell, T. A. Herring, R. A. Anthes, C. Rocken, and R. H. Ware, 1994: GPS meteorology: Mapping zenith wet delays onto precipitable water. J. Appl. Meteor., 33, 379–386.

Boniface, K., V. Ducrocq, G. Jaubert, X. Yan, P. Brousseau, F. Masson, C. Champollion, J. Chéry, and E. Doerflinger, 2009: Impact of high-resolution data assimilation of GPS zenith delay assimilation on Mediterranean heavy rainfall forecasting. Ann. Geophys., 27, 2739–2753.

Bormann, N., and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. Quart. J. Roy. Meteor. Soc., 136, 1036–1050.

Bormann, N., A. Collard, and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data. Quart. J. Roy. Meteor. Soc., 136, 1051–1063.

Chun, H.-W., R. Eresmaa, A. P. McNally, N. Bormann, and M. Matricardi, 2015: A physically-based observation-error covariance matrix for IASI. The 20th International TOVS Study Conference (ITSC-20), 12 pp.

de Haan, S., 2013: Assimilation of GNSS ZTD and radar radial velocity for the benefit of very-short-range regional weather forecasts. Quart. J. Roy. Meteor. Soc., 139, 2097–2107.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. Quart. J. Roy. Meteor. Soc., 131, 3385–3396.

Dixon, M., Z. Li, H. Lean, N. Roberts, and S. Ballard, 2009: Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model. Mon. Wea. Rev., 137, 1562–1584.

Edwards, J. M., and A. Slingo, 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. Quart. J. Roy. Meteor. Soc., 122, 689–719.

Guerova, G., J. Jones, J. Dousa, G. Dick, S. de Haan, E. Pottiaux, O. Bock, R. Pacione, G. Elgered, H. Vedel, and M. Bender, 2016: Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. Atmos. Meas. Tech., 9, 5385–5406.

Hollingsworth, A., and P. Lönneberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A, 38, 111–136.

Kim, E.-H., K.-D. Ahn, H.-C. Lee, J.-C. Ha, and E. Lim, 2015: A study on the effect of ground-based GPS data assimilation into very-short-range prediction model. Atmosphere, 25, 623–637 (in Korean with English abstract).

Kim, E.-H., E. Lee, Y. H. Lee, and S. Joo, 2017: Assimilation of GNSS data in KMA NWP models. Proceedings of IROWG-6, 20 pp. [Available at https://cpaess.ucar.edu/meetings/cosmic-2017/abstracts/assimilation-gnss-data-kma-nwp-models.]

Lee, E., Y. No, Y. Kim, J.-R. Lee, S. Joo, H.-C. Shin, and B. J. Sohn, 2015: Status and plans for satellite data assimilation at the Korea Meteorological Administration. The 20th of International TOVS Study Conference. [Available at https://cimss.ssec.wisc.edu/itwg/itsc/itsc20/program/PDFs/29Oct/session6/posters/6p_03_lee.pdf.]

Lindskog, M., M. Ridal, S. Thorsteinsson, and T. Ning, 2017: Data assimilation of GNSS zenith total delays from a Nordic processing centre. Atmos. Chem. Phys., 17, 13983–13998.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Mon. Wea. Rev., 128, 3187–3199.

Lorenc, A. C., and F. Rawlins, 2005: Why does 4D-Var beat 3D-Var? Quart. J. Roy. Meteor. Soc., 131, 3247–3257.

Mahfouf, J.-F., A. Ahmed, P. Moll, and F. N. Teferle, 2015: Assimilation of zenith total delays in the AROME France convective scale model: A recent assessment. Tellus A, 67, 26106, doi:10.3402/tellusa.v67.26106.

Mark, H., and H. Gemma, 2014: OSDP 13—Processing of zenith total delay measurements from ground-based
GNSS receivers. Met Office OPS Scientific Documentation Papers, 13, 13 pp. [Available at ftp://ftp.ncmrf.gov.in/pub/outgoing/indira/Rani/Ops_docs/doc/scientific/OSDP13/OSDP13.docx.]

Poli, P., P. Moll, F. Rabier, G. Desroziers, B. Chapnik, L. Berre, S. B. Healy, E. Andersson, and F.-Z. El Guelai, 2007: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. J. Geophys. Res., 112, D06114, doi:10.1029/2006JD007430.

Rohm, W., Y. Yuan, B. Biadeqigne, K. Zhang, and J. L. Marshall, 2014: Ground-based GNSS ZTD/IWV estimation system for numerical weather prediction in challenging weather conditions. Atmos. Res., 138, 414–426.

Sánchez Arriola, J., M. Lindskog, S. Thorsteinsson, and J. Bojarova, 2016: Variational bias correction of GNSS ZTD in the HARMONIE modeling system. J. Appl. Meteor., 55, 1259–1276.

Smith, E. K., and S. Weintraub, 1953: The constants in the equation for atmospheric refractive index at radio frequencies. Proc. IRE, 41, 1035–1037.

Tang, Y., H. W. Lean, and J. Bornemann, 2013: The benefits of the Met Office variable resolution NWP model for forecasting convection. Meteor. Appl., 20, 417–426.

van Genuchten, M. T., 1980: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J., 44, 892–898.

Waller, J. A., S. L. Dance, and N. K. Nichols, 2015: Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. Quart. J. Roy. Meteor. Soc., 142, 418–431.

Wood, N., A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat, and J. Thuburn, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. Quart. J. Roy. Meteor. Soc., 140, 1505–1520.

Yan, X., V. Ducrocq, P. Poli, M. Hakam, G. Jaubert, and A. Walpersdorf, 2009: Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall. J. Geophys. Res., 114, D03104, doi:10.1029/2008JD011036.