

Prediction of Energy Dissipation Rates for Aviation Turbulence. Part II: Nowcasting Convective and Nonconvective Turbulence

J. M. PEARSON AND R. D. SHARMAN

National Center for Atmospheric Research,^a Boulder, Colorado

(Manuscript received 14 September 2016, in final form 3 January 2017)

ABSTRACT

In addition to turbulence forecasts, which can be used for strategic planning for turbulence avoidance, short-term nowcasts can augment longer-term forecasts by providing much more timely and accurate turbulence locations for real-time tactical avoidance of turbulence hazards, especially those related to short-lived convection. This paper describes a turbulence-nowcasting algorithm that combines recent short-term turbulence forecasts with all currently available direct turbulence observations and inferences of turbulence from other sources. Building upon the need to provide forecasts that are aircraft independent, the nowcasts provide estimates of an atmospheric metric of turbulence, namely, the energy dissipation rate to the one-third power (EDR). Some observations directly provide EDR, such as in situ observations from select commercial aircraft and ground-based radar algorithm output, whereas others must be translated to EDR. A strategy is provided for mapping turbulence observations, such as pilot reports (PIREPs), and inferences from other relevant observational data sources, such as observed surface wind gusts, into EDR. These remapped observation values can then be combined with short-term turbulence forecasts and other convective diagnostics of turbulence to provide a turbulence nowcast of EDR in the national airspace. Case studies are provided to illustrate the algorithm procedure and benefits. The EDR nowcasts are compared with aircraft in situ EDR observations and PIREPs converted to EDR to obtain metrics of statistical performance. It is shown by one common performance metric, the area under the relative operating characteristic curve, that the turbulence nowcasts with assimilated observations considerably outperform the corresponding turbulence forecasts.

1. Introduction

This is the continuation of a two-part paper that describes a method for developing automated (strategic) forecasts and (tactical) nowcasts of energy dissipation rate to the one-third power (EDR) for aviation turbulence applications. Part I provided a description of the forecast method along with statistical-performance results from comparisons with observations (Sharman and Pearson 2017). The turbulence-forecasting technique is an extension of the Graphical Turbulence Guidance (GTG) product presented in Sharman et al. (2006), which uses post-processing algorithms applied to numerical weather prediction (NWP) model output to produce the turbulence forecasts. Part II (this paper) focuses on

aviation turbulence-nowcasting methods. Although the GTG product does provide short-term forecasts, its utility for tactical avoidance is limited by a 1-h update cycle, which, because of the latency in receiving the underlying NWP-model data and the time required to compute the turbulence forecasts, will usually not be available for at least 2 h beyond the valid time. Also, because of the highly transient and small spatial scales of turbulence associated with convective storms, neither the NWP model nor the turbulence-forecasting post-processing algorithms are particularly skillful at forecasting turbulence associated with convection [see, e.g., Pinto et al. (2015) and references therein and Sharman et al. (2016)]. Thus there is a need to provide more timely and accurate nowcasts for tactical turbulence avoidance that better reflect the current turbulence situation in the National Airspace System (NAS). To some extent, this information is provided by the current system of reporting by pilots and has been fairly successful in providing information for tactical turbulence avoidance, but these reports are point measurements and can have substantial errors in reported intensity,

^aThe National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author e-mail: Julia M. Pearson, jpearson@ucar.edu

position, and time (Schwartz 1996; Bass 2002; Sharman et al. 2006, 2014).

The Graphical Turbulence Guidance Nowcast (GTGN) system was developed to satisfy the need for better forecasts of turbulence that include the effects of convective sources. It is a rapidly updated short-term aviation turbulence-nowcasting product that aims to provide the best possible representation of 3D turbulence locations and intensities in the NAS for tactical turbulence-avoidance purposes. All turbulence sources are accounted for (e.g., low-level, clear-air, mountain-wave, and in- and near-cloud turbulence). Updates are provided every 15 min on a 3D grid that has the same structure as the operational GTG product (available for viewing on the Aviation Digital Data Service website: <http://www.aviationweather.gov/adds/turbulence/>), which in turn is based on NOAA's WRF Rapid Refresh (RAP) NWP-model horizontal grid structure (Benjamin et al. 2016). Like GTG, the output 3D gridded variable is an atmospheric metric of turbulence, namely, EDR ($\text{m}^{2/3}\text{s}^{-1}$), which is aircraft independent and is the International Civil Aviation Organization (ICAO) standard for turbulence reporting (ICAO 2010). The output grid has the same horizontal grid structure as the input NWP model and is interpolated vertically at 1000-ft (1 ft \approx 30.5 cm) intervals, from the surface to flight level (FL) 450.

To ensure that GTGN provides the best possible representation of the current turbulence situation in the NAS, recent turbulence observations from pilot reports (PIREPs), in situ EDR reports, and NEXRAD Turbulence Detection Algorithm (NTDA) output (Williams and Meymaris 2016), as well as turbulence inferences from the Diagnosis of Convectively Induced Turbulence (DCIT) product (Williams 2014) and EDR derived from meteorological terminal aviation routine weather reports (METAR) wind speed and gust observations, are used to nudge the GTG short-term forecast to agree better with the most recently available observations. Descriptions of PIREPs and their limitations are provided in Schwartz (1996) and Sharman et al. (2006), and the in situ EDR reporting system is described in Sharman et al. (2014). NTDA is a radar-based product that uses spectral width estimates from the U.S. network of WSR-88Ds (NEXRAD) to produce atmospheric turbulence intensity estimates (EDR) of in-cloud turbulence (Williams and Meymaris 2016). The DCIT algorithm fuses remote observations with NWP-model data to produce a 3D gridded diagnosis of convectively induced turbulence (CIT) over the continental United States and is more skilled at diagnosing CIT than GTG is (Williams 2014). DCIT was developed as an empirical algorithm that uses the random-forest machine-learning technique trained on

2D and 3D fields derived from the 13-km WRF-RAP NWP model associated with in situ EDR reports as truth. METARs, located primarily at airports, report on several meteorological phenomena, including temperature, dewpoint, sky condition, and visibility, but only wind speed and wind gust observations are used in GTGN. The METAR-provided wind speed is the mean speed in knots over the previous 2-min period. Wind gust is the maximum wind speed over the previous 10-min period when there is a variation of at least 10 kt ($1 \text{ kt} = 0.51 \text{ m s}^{-1}$) between peaks and lulls. METARs report every hour, with special reports that can occur when there are changing meteorological conditions. With over 2000 stations, there can be over 50 000 reports per day.

Figure 1 shows a flow diagram for the GTGN system. The current inputs for the first version of GTGN described here include GTG, DCIT, NTDA, in situ EDR, PIREPs, and METARs. These are shown along with potential inputs to future versions (within the orange box). As indicated in the diagram, the final 3D GTGN grids would be available from a designated processing center and could then be accessed by interested users for display. It is envisioned that these displays would ultimately be in the cockpit to have the most usefulness for tactical turbulence avoidance (e.g., Sharman et al. 2016).

2. Algorithm overview

The basis for the GTGN product is the most recently issued GTG forecast whose valid time is closest to the target GTGN valid time. As an example, the GTGN updates at 1730, 1745, 1800, and 1815 UTC would use either the 2-h GTG forecast issued at 1600 UTC or the 1-h GTG forecast issued at 1700 UTC as its base. Because the lead time for the GTG forecast is short, the weights used to combine the different turbulence diagnostics can be chosen dynamically on the basis of each diagnostic's agreement with recent turbulence observations from PIREPs and in situ EDR reports. The GTG combination of diagnostics is (Sharman et al. 2006; Sharman and Pearson 2017)

$$\text{GTG} = \frac{1}{N} \sum_{i=1}^N W_i D_i, \quad (1)$$

where N is the number of diagnostics used in the combination, D is a particular automated turbulence diagnostic remapped to an EDR scale, and W is the weight of the diagnostic. For short-term GTG forecasts, the weights are chosen to maximize the computed probability of detection yes (PODy) and probability of detection no (PODn) and to minimize the volume of moderate or

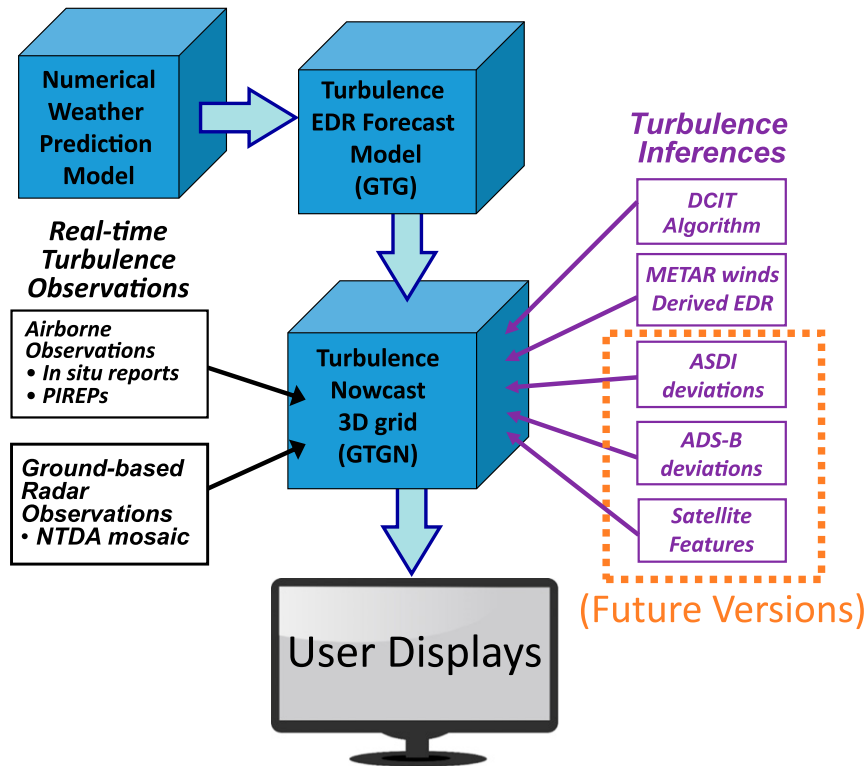


FIG. 1. Flow diagram showing various inputs to the GTGN system. Inputs for the version of GTGN described in this paper are shown along with potential inputs [aircraft situation display to industry (ASDI) deviations, Automatic Dependent Surveillance-Broadcast (ADS-B) deviations, and satellite features, listed within the orange box] for future versions.

greater (MOG) turbulence that the diagnostic forecasts over the domain calculated for the most recent forecast [see Sharman et al. 2006, their Eq. (1)].

This short-term GTG forecast is then blended with DCIT in a first pass, followed by a blending with recent observations in a second pass to produce the final nowcast. The underlying GTG grid and all observations used must be in EDR units ($\text{m}^{2/3} \text{s}^{-1}$) so that the final nowcast is also in EDR units. During the blending process, some observations are not in EDR units and must be empirically converted to EDR. The first blending combines the GTG with DCIT (Williams 2014) 3D forecast grids. DCIT uses artificial intelligence to blend radar data, convective diagnoses, and other information to provide EDR on a 3D grid covering the same domain as GTG, and hence GTGN, and updates on a 15-min cycle. Along with EDR, DCIT provides a convective-interest map that is 1 in areas of probable convection and that decreases to 0 radially away from convection. To guarantee that the most up-to-date convective diagnosis information is included, the most recent DCIT update within the last 20 min is used. The blending uses a linear combination of the GTG and DCIT EDR fields,

weighted by the convective-interest map to provide the first-pass GTGN⁽¹⁾:

$$\text{GTGN}^{(1)}(x, y, z) = W_1 \text{DCIT}(x, y, z) + (1 - W_1) \text{GTG}(x, y, z), \quad (2)$$

where W_1 is convective interest ($0 \leq W_1 \leq 1$). It thus favors DCIT in areas of probable convection and then decreases linearly to zero away from convective activity. This is done because it has been shown that DCIT is more skillful than GTG is at identifying CIT, even in areas where the underlying NWP model does capture the convection (Williams 2014).

Next, all point turbulence observations from PIREPs, in situ EDR reports, and NTDA output, along with the EDR values estimated from METAR wind speed and gust observations, are included. All PIREPs, in situ EDR reports, and METAR EDR values from the previous hour are used as well as all NTDA output within the previous 15 min. NTDA data are produced at 2-km resolution on 3000-ft vertical levels and are available every 5 min (Williams and Meymaris 2016), and therefore three NTDA updates are usually used. The NTDA

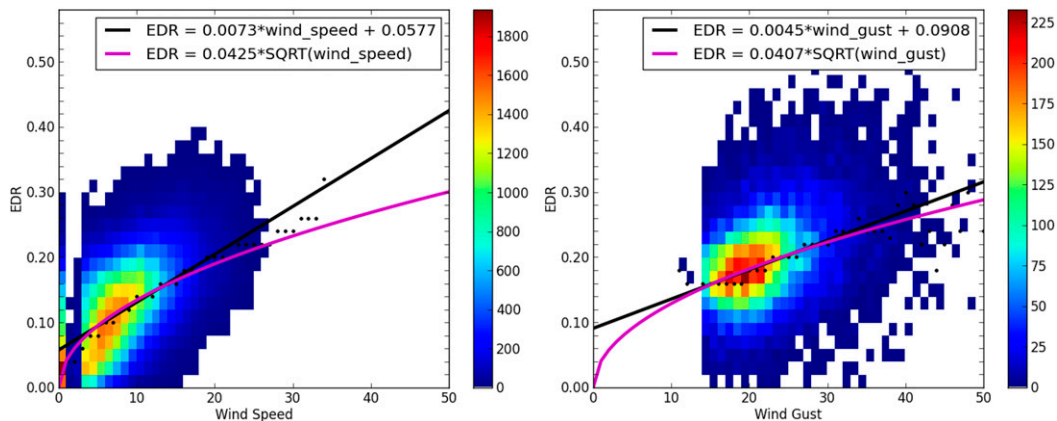


FIG. 2. Plots of (left) 117 069 wind speed and (right) 15 382 wind gust observations (kt) matched with in situ EDR reports within 5 km, 5 min, and 1000 ft of each other. Median speed and gust values within each 1-kt bin are shown as black dots. Fit lines to the median values are shown in black (linear) and purple (square root).

mosaic of EDR values from 140 NEXRADs over the continental United States is projected onto the GTG domain and, using NTDA values provided every 3000 ft starting at 3000 ft and going up to 45 000 ft, is linearly interpolated vertically to 1000-ft levels. Because NTDA values are limited to in-cloud regions, the output grids can be sparse, and so gridcell EDR values are treated as point observations. In situ EDR reports have the desired units, but each PIREP turbulence intensity value (which is typically described subjectively as “smooth,” “light,” “moderate,” or “severe”) must be converted to EDR on the basis of intensity and aircraft type as described in Sharman et al. (2014).

METAR EDR values are calculated on the basis of wind gust observations if they exist, or from wind speed observations otherwise, as an empirically derived function:

$$\begin{aligned} EDR_{\text{METAR}} &= W_1 \sqrt{G_{\text{obs}}} \quad \text{if } G_{\text{obs}} \text{ exists} \\ &= W_2 \sqrt{S_{\text{obs}}} \quad \text{otherwise,} \end{aligned} \quad (3)$$

where $W_1 = 0.0407$, $W_2 = 0.0425$, G_{obs} is the wind gust, and S_{obs} is the wind speed, with the latter two both being in knots.

These equations were derived by linear regression of matched in situ EDR reports with METAR observations. Figure 2 shows 117 069 wind speed (left panel) and 15 382 wind gust (right panel) observations matched with in situ EDR within 5 km, 5 min, and 1000 ft of each other over the 6-yr period from October 2008 to October 2014. Linear and square root functions [Eq. (3)] are shown that fit the median EDR values (black dots in Fig. 2) within each speed and gust bin of 1 kt. The equation for the square root fit was chosen for converting wind observations to EDR because of its goodness of fit for the matched

observations that occur with the most frequency as well as its natural behavior of converting zero wind to zero EDR.

The blending of the turbulence observations and inferences into the GTGN grid is a complicated procedure that depends on observation type, EDR intensity, and what the underlying combined GTG–DCIT [GTGN⁽¹⁾] grid is predicting. The following discussion summarizes the process.

The order of the blending of the turbulence observations and inferences depends on type, age, and intensity. First, the point EDR observations are sorted separately by observation type on the basis of time and intensity, with the order going from oldest to newest and, within the same time ranges, lowest to highest EDR intensity. The observations are then combined into one ordered list for blending by including all METAR values first and then all PIREPs, followed by in situ and NTDA EDR data based on time. This ordering was chosen so that observations with more confidence and higher EDR values take precedence in the final blended GTGN output.

Each point observation has a default vertical and horizontal influence region about its reported location that is based on observation type and EDR intensity. PIREPs, in situ EDR reports, and NTDA cells that do not have other NTDA values above and/or below them all have a vertical influence extent [H_i in Eqs. (4) and (5) below] of 2000 ft. All other NTDA EDRs have no additional vertical influence past their observed altitude [i.e., H_i in Eqs. (4) and (5) is set to zero]. The vertical extent of METAR EDR values depends on the height of the planetary boundary layer (PBL) at the observation station’s location: they extend from the surface to 2000 ft above the top of the PBL (as provided in the NWP model output) at each location. This strategy is based on observations of the magnitude of the individual terms in the turbulent kinetic energy equation in the unstable

TABLE 1. Radius of the horizontal influence region (km) by observation type and EDR value.

EDR value	Observation type	
	PIREPs, in situ EDR reports, and METAR EDRs	NTDA cell
0.0–0.15	100 km	30 km
0.15–0.22	100 km	30 km
0.22–0.38	64 km	30 km
>0.38	40 km	30 km

PBL that show a fairly constant dissipation above the surface layer (e.g., [Lenschow 1974](#)). In the future, this will be refined to better fit available observations in both the stable and unstable PBLs. For the purpose of determining the horizontal influence, EDR intensity (provided directly from in situ estimates and NTDA or derived from empirical mappings for PIREPs and METARs) is binned into severity categories on the basis of the guidance provided in [Sharman et al. \(2014\)](#) for medium-sized aircraft. These categories are null ($0 \leq \text{EDR} < 0.15$), light ($0.15 \leq \text{EDR} < 0.22$), moderate ($0.22 \leq \text{EDR} < 0.38$), and severe ($\text{EDR} \geq 0.38$). The default horizontal influence region of each of the observations is shown in [Table 1](#). The horizontal influence region of NTDA does not change with regard to EDR intensity, but for in situ EDR, PIREPs, and METAR EDRs the radius changes on the basis of the EDR value, with lower EDR values having a larger horizontal extent and higher EDR values having a smaller one. In situ reports, PIREPs, and METAR EDRs that have an EDR intensity of null turbulence ($\text{EDR} < 0.15$) or MOG turbulence ($\text{EDR} \geq 0.22$), however, can have a larger horizontal influence region on the basis of the predicted values of the underlying GTG–DCIT [$\text{GTGN}^{(1)}$] grid. In areas where the underlying grid is also predicting either null or MOG turbulence, the horizontal influence of the observation is 2 times that shown in [Table 1](#). The influence-region extent was established on the basis of sensitivity studies.

Each observation is blended in the assigned order into the combined GTG–DCIT [$\text{GTGN}^{(1)}$] grid using a weighted combination of the observation EDR and the grid-based EDR. For NTDA observations, as well as for PIREPs, in situ EDR reports, and METAR EDRs, using the default influence region, a simple linearly weighted combination that is based on distance within the influence region is used, with the weighting function for the observation EDR linearly decaying from 1 to 0 according to

$$\text{GTGN}^{(2)}(x, y, z) = W_1 \text{EDR}_{\text{obs}} + W_2 \text{GTGN}^{(1)}(x, y, z), \quad (4)$$

where W_1 is equal to $(1 - d_{\text{obs}}/R)F_V$ for (x, y, z) within the influence region and is 0 otherwise, W_2 is equal to

$1 - W_1$, and F_V is 1 for $H_i = 0$ and is equal to $(H_i - |h_{\text{obs}} - h|)/H_i$ otherwise, and where EDR_{obs} is the EDR from the current observation being incorporated (from in situ, PIREP, METAR, or NTDA), R is given in [Table 1](#), h_{obs} is the height of the observation, h is the height of the current grid cell, and H_i is described above. Note that the observation EDR takes precedence at the observation location and that the underlying grid takes precedence at the edge of the influence-region boundary. The observations are always blended with the original unmodified $\text{GTGN}^{(1)}$ grid EDR value, and in areas where multiple observations may impact the blended grid the closest observation takes precedence, unless a more distant observation has an elevated EDR value whose influence would increase the resultant EDR. In areas of the grid outside the influence regions of all observations, the final GTGN value [$\text{GTGN}^{(2)}$] is the value of $\text{GTGN}^{(1)}$ at the corresponding grid point. The linear-interpolation weighting function is simple and is in the spirit of [Cressman's \(1959\)](#) first-pass objective-analysis scheme. More-sophisticated schemes that produce smoother results could be used, but it is unclear whether they could substantially benefit the final results, since dynamic consistency cannot be enforced here.

Equation (4) specifies the default weighting function. For in situ EDR reports, PIREPs, and METAR EDR values that are either null ($\text{EDR} < 0.15$) or MOG ($\text{EDR} \geq 0.22$) in areas where the underlying grid is also predicting similar EDR intensities, the horizontal influence radius is 2 times that shown in [Table 1](#) and a slightly different weighted combination is applied to blend in the observed EDR values:

$$\text{GTGN}^{(2)}(x, y, z) = W_1 \text{EDR}_{\text{obs}} + W_2 \text{GTGN}^{(1)}(x, y, z), \quad (5)$$

[i.e., the same form as Eq. (4)] but where now W_1 is equal to 0 for (x, y, z) outside the influence region, is equal to $[4(F_d - 0.5)^3 + 0.5]F_V$ for $F_d > 0.5$, and is equal to $F_d F_V$ otherwise, F_d is equal to $[1 - d_{\text{obs}}/(2R)]$, and the values for W_2 and F_V are calculated as in Eq. (4).

Once all of the turbulence observations and inferences are blended with the combined GTG–DCIT [$\text{GTGN}^{(1)}$] grid, the output is smoothed using a 2D 3×3 gridcell average in areas where observations were incorporated and for which the EDR intensity is below the designated MOG threshold of $0.22 \text{ m}^{2/3} \text{ s}^{-1}$.

3. GTGN examples

To illustrate the GTGN procedure, three example cases are provided in [Figs. 3–5](#). The examples show the

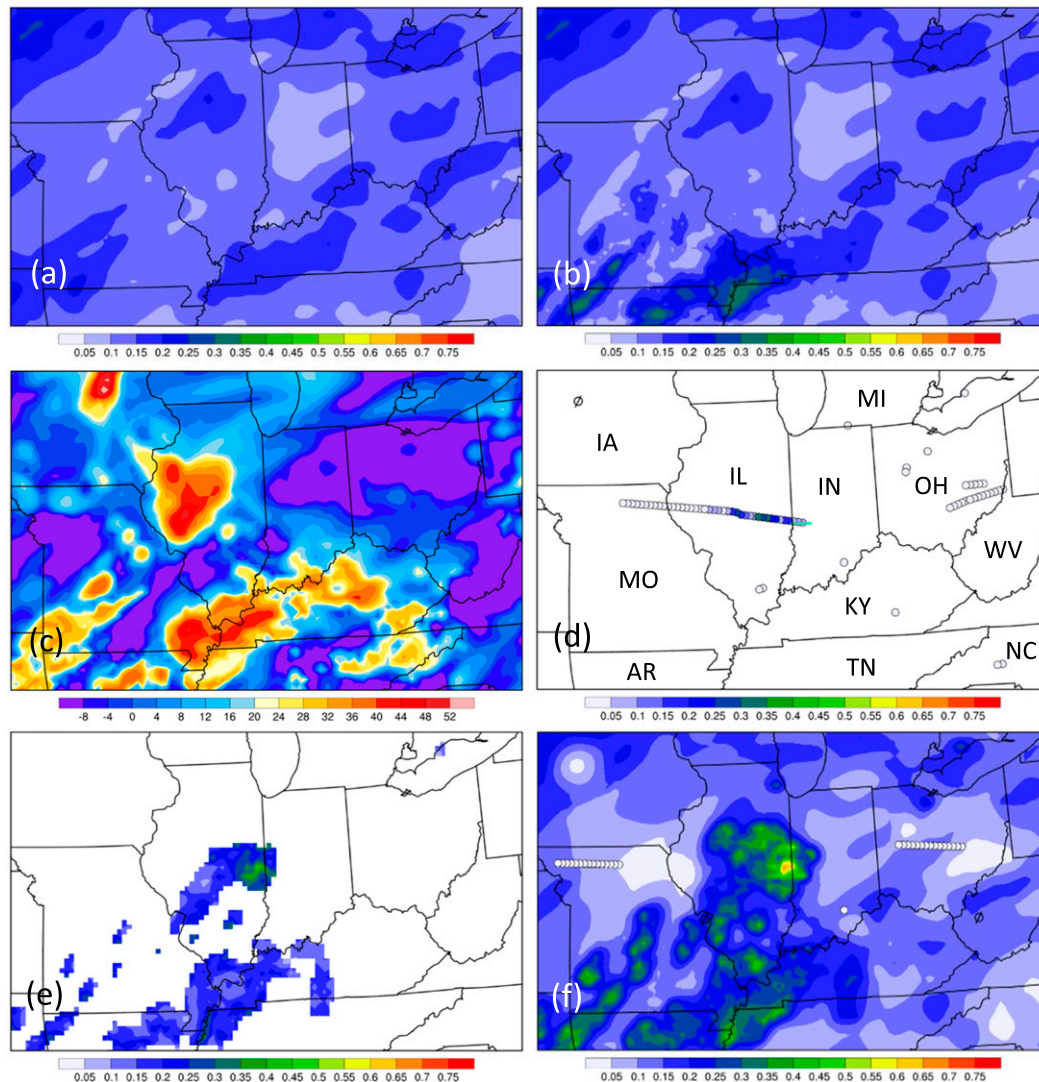


FIG. 3. (a) GTG 2-h forecast valid at 1500 UTC, (b) combined GTG-DCIT [GTGN⁽¹⁾] grid using DCIT valid at 1500 UTC, (c) composite reflectivity from the 2-h forecast WRF-RAP valid at 1500 UTC, (d) PIREPs and in situ EDR reports within the previous hour to 1500 UTC, (e) NTDA valid at 1455 UTC, and (f) GTGN valid at 1500 UTC with verifying observations from 1500 to 1515 UTC overlaid—all from 26 Dec 2015 at FL300. State labels are shown in (d).

GTGN components of and final output obtained within a zoomed-in region for the 1) 1500 UTC update on 26 December 2015 at FL300, the 2) 1800 UTC update on 7 March 2016 at FL360, and the 3) 1700 UTC update on 15 April 2016. The first two examples show the GTG forecast, combined GTG and DCIT [GTGN⁽¹⁾] grids, corresponding WRF-RAP computed composite reflectivity, one of the three NTDA input grids (they vary little), in situ EDR reports, and PIREPs that are combined as described above to form the GTGN grid along with the final GTGN output. The last example illustrates the influence of the METAR EDR values from wind speed

and gust observations on the GTG forecast at 2000 ft MSL.

a. Example 1: 26 December 2015 valid at 1500 UTC, FL300

For this example, the GTG 2-h forecast from 1300 UTC provides the basis for GTGN. Figure 3a shows a zoomed-in portion of the GTG forecast at FL300. Most of the EDR forecast values in this region range from 0.1 to $0.2 \text{ m}^{2/3} \text{ s}^{-1}$, which would correspond to null or light turbulence for a medium-sized aircraft. There are also a few patches of forecast EDR values ranging from 0.2 to $0.25 \text{ m}^{2/3} \text{ s}^{-1}$, which correspond to moderate turbulence

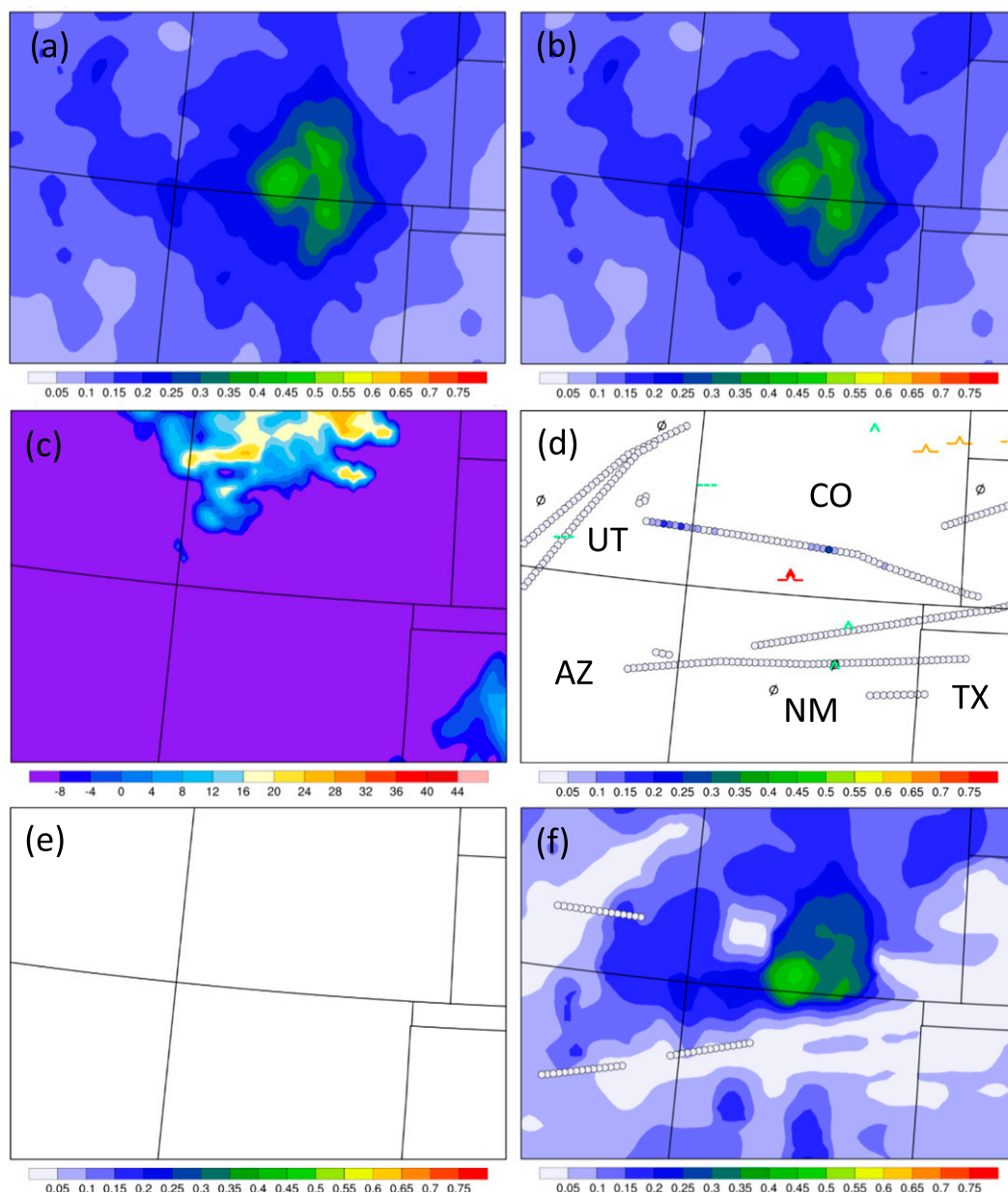


FIG. 4. (a) GTG 2-h forecast valid at 1800 UTC, (b) combined GTG–DCIT [GTGN⁽¹⁾] grid using DCIT valid at 1800 UTC, (c) composite reflectivity from the 2-h WRF–RAP valid at 1800 UTC, (d) PIREPs and in situ EDR reports within the previous hour to 1800 UTC, (e) NTDA valid at 1755 UTC, and (f) GTGN valid at 1800 UTC with verifying observations from 1800 to 1815 UTC overlaid—all from 7 Mar 2016 at FL360. State labels are shown in (d).

for a medium-sized aircraft (Sharman et al. 2014). The combined GTG–DCIT [GTGN⁽¹⁾] grid is shown in Fig. 3b. The DCIT output indicates some areas of CIT in the southern portion of the image, resulting in regions of combined EDRs between 0.2 and 0.35 m^{2/3} s^{−1}. Some dampening of the EDR forecast values can also be seen in the region over central Missouri. For reference, the composite reflectivity from the WRF–RAP model is shown in Fig. 3c. The NWP model does show higher

reflectivity values in the region, but not in the exact same places as indicated by the DCIT forecast or the 1455 UTC NTDA update shown in Fig. 3e, implying the NWP model misplaced the convection. This NTDA update along with the updates at 1450 and 1445 UTC was included in the GTGN forecast at 1500 UTC. NTDA identifies regions of moderate and severe turbulence in cloud in this area. Areas shown in white denote no value for NTDA, which could be due to the reflectivity being too low for reliable

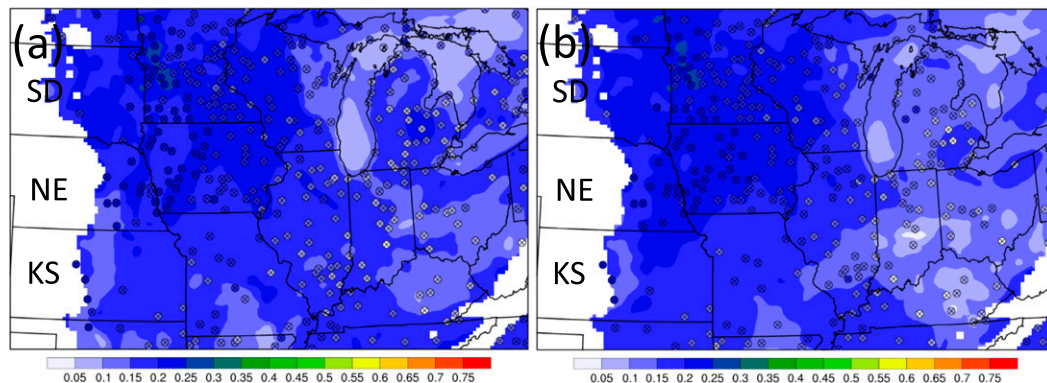


FIG. 5. (a) GTG 2-h forecast valid at 1700 UTC with METAR EDRs from 1600 to 1700 UTC overlaid and (b) GTGN using only GTG and METAR EDR valid at 1700 UTC with METAR EDRs from 1700 to 1715 UTC overlaid—both for 15 Apr 2016 at 2000 ft MSL. State labels are shown for South Dakota, Nebraska, and Kansas.

detection or to missing data in the region. This is an example in which GTG was unable to capture the observed CIT in the area. Even if GTG were able to capture the convective turbulence, its location would be misplaced because of the spatial error in the NWP-model output. Both PIREPs and in situ EDR reports that occurred within the region at FL300 are shown in Fig. 3d. The two PIREPs shown are for null (central Iowa) and smooth-to-light turbulence (western Indiana border). The EDR reports shown as circles correspond to EDR values ranging from 0 to $0.35 \text{ m}^{2/3} \text{ s}^{-1}$. These figures show that the NTDAs EDR values and the in situ EDR reports agree very well in the areas where they overlap. Figure 3f shows GTGN at FL300 valid at 1500 UTC derived from the combined inputs. Observations reported within the 15 min beginning at 1500 UTC are overlaid for verification. This example shows a substantial difference between the original GTG forecast and the final GTGN output that are both valid at the same time. Areas where the forecast was dampened by null in situ EDR and the null PIREP can be seen over Ohio, southern Michigan, and the northern corner of North Carolina, as well as the region near the Illinois, Iowa, and Missouri borders. Also, there are now significant areas of moderate and severe turbulence because of the inclusion of NTDA and the elevated in situ EDR reports throughout southern and central Illinois and Missouri, as well as over northern Arkansas and Tennessee and western Kentucky. The verifying observations show null in situ EDR reports occurring over Ohio and northern Missouri where the GTG forecast EDRs were dampened to null and null-to-light EDR nowcast values in GTGN. This case illustrates that the incorporation of observations into GTGN creates a nowcast that is more consistent with the current state of the atmosphere

when compared with the NWP-model-based GTG forecast.

b. Example 2: 7 March 2016 valid at 1800 UTC

For the second example, the 2-h GTG forecast issued at 1600 UTC 7 March 2016 was used as the basis for the GTGN update valid at 1800 UTC. Figure 4a shows a zoomed-in region of the GTG forecast at FL360. This region contains most of Colorado along with large areas of Utah, Arizona, New Mexico, and parts of other adjacent states. Note that GTG is forecasting a large area of moderate and severe turbulence in southern Colorado, with EDR values ranging from 0.20 to $0.50 \text{ m}^{2/3} \text{ s}^{-1}$, and widespread areas of null-to-light turbulence, shown in light blue and corresponding to EDR values between 0.1 and $0.15 \text{ m}^{2/3} \text{ s}^{-1}$. This region is apparently not impacted by convectively induced turbulence since the combined GTG–DCIT [GTGN⁽¹⁾] grid (Fig. 4b) shows no difference from the GTG forecast in Fig. 4a. The WRF-RAP NWP model does predict some reflectivity (Fig. 4c) in central and northwestern Colorado, but this is not apparent in the GTG forecast. In addition, this example region does not show in-cloud turbulence data from NTDA (Fig. 4e). There were, however, in situ EDR and PIREP observations as shown in Fig. 4d. The majority of the observations are null or light, but there is one in situ EDR report of $0.26 \text{ m}^{2/3} \text{ s}^{-1}$ shown as a dark blue circle to the northeast of a severe PIREP (red inverted triangles) located in the region where GTG forecasts moderate and severe turbulence, as well as three EDR reports ranging from 0.1 up to 0.25 EDR over southeastern Utah. Thus, although GTG does capture the elevated turbulence in the area, the total area of forecast elevated EDRs seems too large. Figure 4f shows the final GTGN output with verifying observations from 1800 to 1815 UTC overlaid. In

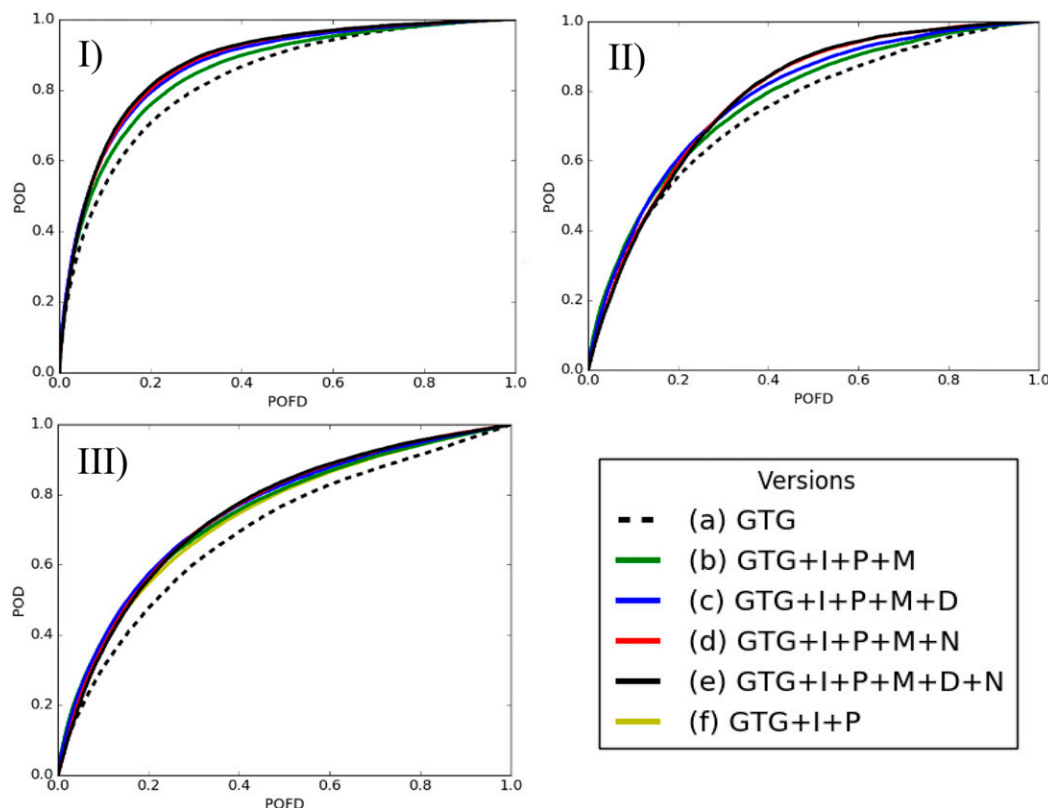


FIG. 6. ROC curves for (I) high, (II) mid-, and (III) low levels showing GTG only (version a); GTG with PIREPs (P), in situ EDR reports (I), and METARs only (M) (version b); GTG with PIREPs, in situ EDR, METARs, and DCIT (D) (version c); GTG with PIREPs, in situ EDR, METARs, and NTDA (N) (version d); GTG with PIREPs, in situ EDR, METARs, DCIT, and NTDA (version e); and, for low levels only, GTG with PIREPs and in situ EDR reports only (version f), for the combined winter and summer seasons.

general, the GTGN output EDR values are less than the corresponding GTG base grid because of the inclusion of the many null turbulence observations. The area of forecast moderate and severe turbulence relative to GTG is also less widespread, with EDR values being more consistent with the input observations. The verifying observations shown in Fig. 4f show null in situ EDR reports in areas nowcast as null turbulence in New Mexico and Arizona, as well as in regions of null and light-to-moderate nowcast turbulence in southeastern Utah. This case illustrates the benefit of GTGN to tighten up areas of forecast MOG turbulence and to greatly reduce the volume of MOG forecast by GTG.

c. Example 3: 15 April 2016 valid at 1700 UTC

This example illustrates the effect of the inclusion of ground-based wind and gust observations from METARs converted to EDR on the low-level turbulence predicted by the GTGN product. Only the GTG forecast, the combined GTG forecast with the METAR EDR, and the associated METAR EDR observation values

are shown. Figure 5a shows the GTG forecast valid at 1700 UTC at 2000 ft MSL over parts of the Midwest with the METAR wind speed and gust observations [converted to EDR using Eq. (3)] from 1600 to 1700 UTC overlaid. Figure 5b shows the adjusted forecast that is based on incorporating these METAR-derived EDR values with the METAR EDRs from 1700 to 1715 UTC overlaid for verification. The impact of the inclusion of the METAR EDR values is to increase the EDR intensity of the GTG forecast, as shown on the western side of Iowa, as well as to decrease the EDR intensity of the forecast in parts of Kentucky, Ohio, Indiana, central Illinois, and Michigan. The verifying METAR EDRs show good agreement with GTGN in these modified regions.

4. GTGN verification statistics

Standard verification statistics were derived on the basis of calculated 2×2 contingency tables (e.g., Gill 2016) for GTG and various versions of GTGN for

TABLE 2. Statistics for GTG and various versions of GTGN; winter and summer seasons.

Version	AUC	TSS (0.22)	PODy (0.22)	PODn (0.22)	RMSE	RMSE > 0.22
Altitude region: high (20 000 ft and above); $N = 2\,483\,757$						
GTG	0.83	0.34	0.39	0.95	0.10	0.10
GTG + in situ + PIREPs + METAR	0.86	0.32	0.36	0.97	0.08	0.11
GTG + in situ + PIREPs + METAR + DCIT	0.87	0.37	0.41	0.96	0.08	0.10
GTG + in situ + PIREPs + METAR + NTDA	0.88	0.42	0.47	0.95	0.08	0.10
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.88	0.44	0.50	0.94	0.09	0.10
Altitude region: middle (10 000–20 000 ft); $N = 244\,481$						
GTG	0.75	0.26	0.34	0.92	0.10	0.11
GTG + in situ + PIREPs + METAR	0.77	0.26	0.34	0.93	0.09	0.11
GTG + in situ + PIREPs + METAR + DCIT	0.78	0.31	0.41	0.90	0.09	0.11
GTG + in situ + PIREPs + METAR + NTDA	0.79	0.39	0.58	0.81	0.12	0.11
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.79	0.40	0.62	0.79	0.12	0.11
Altitude region: low (surface–10 000 ft); $N = 681\,979$						
GTG	0.70	0.21	0.32	0.89	0.09	0.10
GTG + in situ + PIREPs	0.74	0.28	0.39	0.89	0.08	0.10
GTG + in situ + PIREPs + METAR	0.74	0.29	0.39	0.90	0.08	0.10
GTG + in situ + PIREPs + METAR + DCIT	0.75	0.32	0.45	0.88	0.09	0.09
GTG + in situ + PIREPs + METAR + NTDA	0.75	0.34	0.48	0.86	0.09	0.09
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.75	0.35	0.52	0.84	0.10	0.09

comparison. The GTG 2-h forecast was considered for all GTGN versions that were scored. To assess the importance of various components of GTGN, different configurations were tested and scored: GTG only (version a in Fig. 6); GTG modified with PIREPs, in situ EDR reports, and METARs only (version b); GTG modified with PIREPs, in situ EDR, METARs, and DCIT (version c); GTG modified with PIREPs, in situ EDR, METARs, and NTDA (version d); GTG modified with PIREPs, in situ EDR, METARs, DCIT, and NTDA (version e); and, for low levels, GTG modified with PIREPs and in situ EDR reports only (version f). Tables 2–4 show the total observations used in the evaluations N , area under the relative operating characteristic curve (ROC) (AUC), true skill score (TSS), PODy, PODn, root-mean-square error (RMSE), and the RMSE for observations that have EDR of greater than 0.22. The threshold used to distinguish MOG forecasts for TSS, PODy, and PODn is an EDR value of 0.22. The statistics were based on two time periods, July–September 2015 (summer) and January–March 2016 (winter), for GTGN times valid for each 15-min time update between 1130 and 0215 UTC. Table 2 contains various performance statistics for the combined winter and summer performance. Tables 3 and 4 contain

performance statistics for winter and summer, respectively. Recall that a 2-h GTG forecast contributes to four different GTGN updates, as described above, such that the valid times of the GTGN nowcasts span the 1-h period centered about the GTG forecast valid time. Each GTGN update was scored using the PIREPs and in situ EDR observations that were reported during the duration of each update valid time, that is, the 15-min period following the GTGN valid time. Using the observations within ± 30 min of the GTG forecast valid time to score GTG results in exactly the same set of observations being used to verify both products. The statistics in Table 2 for both seasons show that GTGN, in all variants, statistically outperforms GTG when considering near-term forecast performance. For the AUC, it can be seen that all variants of GTGN are superior to GTG. At upper levels, the improvement in AUC varies between 0.03 (for GTG + in situ + PIREPs + METAR) to 0.05 (GTG + in situ + PIREPs + METAR + DCIT + NTDA), which is an improvement of 3.6%–6%. For midlevels, the AUC increase is not quite as high but is still substantial, with improvement between 2.7% and 5.3%. Low levels had the largest overall increase in the AUC, with improvements between 5.7% and 7.1%. Sample ROC curves are shown in Fig. 6 for high, mid-,

TABLE 3. Statistics for GTG and various versions of GTGN; winter season.

Version	AUC	TSS (0.22)	POD _y (0.22)	POD _n (0.22)	RMSE	RMSE > 0.22
Altitude region: high (20 000 ft and above); $N = 796\,366$						
GTG	0.83	0.38	0.48	0.90	0.11	0.09
GTG + in situ + PIREPs + METAR	0.85	0.37	0.44	0.93	0.09	0.10
GTG + in situ + PIREPs + METAR + DCIT	0.85	0.39	0.47	0.93	0.10	0.10
GTG + in situ + PIREPs + METAR + NTDA	0.86	0.43	0.51	0.92	0.10	0.10
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.86	0.44	0.53	0.91	0.10	0.10
Altitude region: middle (10 000–20 000 ft); $N = 92\,172$						
GTG	0.78	0.35	0.50	0.85	0.12	0.09
GTG + in situ + PIREPs + METAR	0.80	0.35	0.48	0.87	0.10	0.09
GTG + in situ + PIREPs + METAR + DCIT	0.80	0.37	0.51	0.87	0.10	0.09
GTG + in situ + PIREPs + METAR + NTDA	0.78	0.41	0.69	0.72	0.14	0.11
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.79	0.42	0.70	0.72	0.14	0.11
Altitude region: low (surface–10 000 ft); $N = 254\,643$						
GTG	0.74	0.26	0.38	0.88	0.09	0.09
GTG + in situ + PIREPs	0.77	0.33	0.45	0.88	0.08	0.09
GTG + in situ + PIREPs + METAR	0.78	0.35	0.47	0.88	0.09	0.09
GTG + in situ + PIREPs + METAR + DCIT	0.78	0.36	0.49	0.87	0.09	0.09
GTG + in situ + PIREPs + METAR + NTDA	0.77	0.37	0.56	0.82	0.10	0.09
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.77	0.38	0.57	0.81	0.10	0.09

and low levels for the combined winter and summer seasons. These curves show the fraction of correctly forecast events, POD_y or simply POD, versus the false-alarm rate (POFD = 1 – POD_n) for various forecast EDR thresholds. Again, GTGN with all inputs (version e described above) has the largest area under the ROC curve, indicating greatest discrimination skill. For this GTGN version, the POD_y, based on the MOG threshold of 0.22, has an even greater increase at 28%, 82%, and 63% for upper, mid-, and lower levels, respectively, with the trade-off of small decreases in POD_n (1%, 14%, and 6%). Overall, using DCIT with NTDA (black curve in Fig. 6) generally provides the best performance, but the combination including NTDA but without DCIT is very close (red curve in Fig. 6). Consistent with GTG performance, the statistical performance of GTGN with respect to AUC, TSS, POD_y, and POD_n degrades from upper to low levels. Last, the overall performance of GTGN when considering RMSE does not change much relative to GTG.

Similar trends of better performance of GTGN relative to GTG are seen in the winter and summer statistics in Tables 3 and 4. Performance for all products degrades in the summer months when compared with winter. The performance of GTG drops considerably between winter and summer, whereas the performance of GTGN

with all inputs included has a smaller degradation. This is presumably related to GTGN's ability to better capture convective-turbulence sources. When comparing GTG at upper levels, the decrease in AUC, TSS, and POD_y between winter and summer periods is 8%, 55%, and 58%, respectively, while the same decreases for GTGN are only 1%, 11%, and 17%.

5. Summary and conclusions

A method has been developed to combine turbulence observations and turbulence inferences with forecasts to create a nowcast of the current turbulence conditions in the NAS, from the surface to FL450. The method starts with a GTG forecast of EDR derived from a 1- or 2-h NWP model forecast. In principle, other turbulence-forecasting models could be used as long as they provide forecasts of EDR. Here the WRF-RAP 2-h forecast is used as input to GTG. Next, in a first blending pass, the convectively induced turbulence diagnostic DCIT is combined with the GTG forecast in areas of convection. Observations and inferences of turbulence from NTDA, PIREPs, in situ EDR, and METAR wind observations are then blended into the grid as point observations in a second pass. The method assumes that all inputs are

TABLE 4. Statistics for GTG and various versions of GTGN; summer season.

Version	AUC	TSS (0.22)	PODy (0.22)	PODn (0.22)	RMSE	RMSE > 0.22
Altitude region: high (20 000 ft and above); $N = 1\,687\,391$						
GTG	0.76	0.17	0.20	0.99	0.09	0.12
GTG + in situ + PIREPs + METAR	0.80	0.15	0.17	0.98	0.07	0.13
GTG + in situ + PIREPs + METAR + DCIT	0.85	0.25	0.28	0.97	0.07	0.12
GTG + in situ + PIREPs + METAR + NTDA	0.87	0.34	0.38	0.96	0.08	0.11
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.87	0.39	0.44	0.95	0.08	0.11
Altitude region: middle (10 000–20 000 ft); $N = 152\,309$						
GTG	0.69	0.08	0.12	0.96	0.08	0.13
GTG + in situ + PIREPs + METAR	0.72	0.09	0.14	0.96	0.08	0.13
GTG + in situ + PIREPs + METAR + DCIT	0.75	0.18	0.27	0.92	0.09	0.12
GTG + in situ + PIREPs + METAR + NTDA	0.76	0.29	0.42	0.86	0.11	0.10
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.76	0.32	0.49	0.83	0.11	0.10
Altitude region: low (surface–10 000 ft); $N = 427\,336$						
GTG	0.65	0.15	0.25	0.90	0.08	0.11
GTG + in situ + PIREPs	0.70	0.21	0.31	0.91	0.08	0.11
GTG + in situ + PIREPs + METAR	0.70	0.21	0.30	0.91	0.08	0.11
GTG + in situ + PIREPs + METAR + DCIT	0.72	0.28	0.39	0.88	0.08	0.10
GTG + in situ + PIREPs + METAR + NTDA	0.72	0.27	0.38	0.88	0.09	0.10
GTG + in situ + PIREPs + METAR + DCIT + NTDA	0.72	0.31	0.45	0.85	0.09	0.09

converted to EDR for blending and nudges forecast values toward observations within specified areas of influence. The final nowcast is EDR updated every 15 min.

One major benefit of the GTGN algorithm over the current GTG forecast is the inclusion of convective-turbulence sources from the inclusion of DCIT, in-cloud observations from NTDA, as well as from in-cloud PIREPs and in situ EDR reports. This was exemplified in Fig. 3, which shows a region of convectively induced turbulence that was missed by GTG but identified by DCIT and observed by NTDA and in situ EDR reports. With NTDA updating every 5 min, DCIT available every 15 min, and PIREPs and in situ EDR reports available shortly after encountered elevated events, GTGN is able to capture convectively induced turbulence in a way that is not feasible using current NWP models alone and hence GTG. In addition, GTGN is able to enhance areas of elevated turbulence forecasts but, as important, reduce the volume of moderate or greater turbulence forecast by GTG by dampening the forecast in areas of null or light turbulence observations. Figure 4 illustrates both the enhancement of a region of moderate or greater turbulence as well as the reduction of forecast elevated turbulence areas.

Both GTG and GTGN, in various forms, were evaluated using standard statistical verification metrics,

including the AUC, TSS, PODy, PODn, and RMSE. These metrics showed that the inclusion of all proposed inputs improved the AUC of GTGN over GTG by 6%, 5%, and 7% at upper, mid-, and low levels, respectively. The TSS also improved, with increases in PODy that were at the expense of minor decreases in PODn. The comparisons showed that the inclusion of only PIREPs and in situ EDR improve upon GTG performance, and the inclusion of NTDA provides still further improvement. As expected, the improvement in GTGN performance relative to that of GTG during the summer period, when CIT is more prevalent, is the most significant.

The incorporation and blending of observations into the GTG grid are done through simple linear influence regions. More-sophisticated influence functions and relevant data-assimilation techniques could be tried in future versions, but, unlike data-assimilation techniques used in NWP model initializations, there are no obvious dynamic constraints that should be enforced. In a similar way, blending the EDR observations directly into NWP model initialization procedures to modify the modeled dissipation and primitive variable fields would require a substantial research effort. To move toward operational implementation, GTGN will include the turbulence observations, in situ EDR reports, PIREPs, and NTDA.

DCIT and METAR EDR inferences are scheduled to be included in a future version.

Acknowledgments. The authors are grateful to Domingo Muñoz-Esparza and Frank McDonough for their constructive comments that led to clarifications in the manuscript. This research is in response to requirements and funding by the Federal Aviation Administration (FAA; Grant DTFWA-15-D-00036). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

REFERENCES

- Bass, E. J., 2002: Turbulence assessment and decision-making in commercial aviation: Investigating the current state of practice and the efforts of technology interventions. *Int. J. Appl. Aviat. Stud.*, **2**, 11–22.
- Benjamin, S., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, doi:[10.1175/MWR-D-15-0242.1](https://doi.org/10.1175/MWR-D-15-0242.1).
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, doi:[10.1175/1520-0493\(1959\)087<0367:AOOAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2).
- Gill, P. G., 2016: Aviation turbulence forecast verification. *Aviation Turbulence: Processes, Detection, Prediction*, R. Sharman and T. Lane, Eds., Springer, 261–283.
- ICAO, 2010: Meteorological service for international air navigation. Annex 3, Convention on International Civil Aviation, 17th ed. 206 pp. [Available online at <http://store1.icao.int/index.php/publications/annexes/3-meteorological-service-for-international-air-navigation.html>.]
- Lenschow, D. H., 1974: Model of the height variation on the turbulence kinetic energy budget in the unstable planetary boundary layer. *J. Atmos. Sci.*, **31**, 465–474, doi:[10.1175/1520-0469\(1974\)031<0465:MOTHVO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0465:MOTHVO>2.0.CO;2).
- Pinto, J. O., J. A. Grim, and M. Steiner, 2015: Assessment of the High-Resolution Rapid Refresh model's ability to predict mesoscale convective systems using object-based evaluation. *Wea. Forecasting*, **30**, 892–913, doi:[10.1175/WAF-D-14-00118.1](https://doi.org/10.1175/WAF-D-14-00118.1).
- Schwartz, B., 1996: The quantitative use of PIREPs in developing aviation weather guidance products. *Wea. Forecasting*, **11**, 372–384, doi:[10.1175/1520-0434\(1996\)011<0372:TQOPI>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0372:TQOPI>2.0.CO;2).
- Sharman, R. D., and J. Pearson, 2017: Prediction of energy dissipation rates for aviation turbulence. Part I: Forecasting non-convective turbulence. *J. Appl. Meteor. Climatol.*, **56**, 317–337, doi:[10.1175/JAMC-D-16-0205.1](https://doi.org/10.1175/JAMC-D-16-0205.1).
- , C. Tebaldi, G. Wiener, and J. Wolff, 2006: An integrated approach to mid- and upper-level turbulence forecasting. *Wea. Forecasting*, **21**, 268–287, doi:[10.1175/WAF924.1](https://doi.org/10.1175/WAF924.1).
- , L. B. Cornman, G. Meymaris, J. Pearson, and T. Farrar, 2014: Description and derived climatologies of automated in situ eddy dissipation rate reports of atmospheric turbulence. *J. Appl. Meteor. Climatol.*, **53**, 1416–1432, doi:[10.1175/JAMC-D-13-0329.1](https://doi.org/10.1175/JAMC-D-13-0329.1).
- , T. Lane, and U. Schumann, 2016: Research needs. *Aviation Turbulence: Processes, Detection, Prediction*, R. Sharman and T. Lane, Eds., Springer, 501–518.
- Williams, J. K., 2014: Using random forests to diagnose aviation turbulence. *Mach. Learn.*, **95**, 51–70, doi:[10.1007/s10994-013-5346-7](https://doi.org/10.1007/s10994-013-5346-7).
- , and G. Meymaris, 2016: Remote turbulence detection using ground-based Doppler weather radar. *Aviation Turbulence: Processes, Detection, Prediction*, R. Sharman and T. Lane, Eds., Springer, 149–177.