A Physical Versus a Statistical Forecast Formulation for Clear Air Turbulence

Donald W. McCann
McCann Aviation Weather Research, Inc.
Overland Park, Kansas, USA

A report submitted to
DTN Weather
Burnsville, Minnesota, USA
16 November 2018

ABSTRACT

This report begins with a review of a physically-based clear air turbulence (CAT) conceptual model that dates its origins to the 1960s. The model's equations directly compute eddy dissipation rate (EDR), a measure of atmospheric turbulence. CAT results from a local gravity wave modification of the environmental wind shear and stability to lower the environmental Richardson number to less than 0.25, the threshold for Kelvin-Helmholtz instability, to initiate turbulence. To operationally implement this model, one needs to find locations and intensities of gravity waves of which several have been proposed. This model is in contrast to statistically-based CAT relationships which began as single formula indices and have progressed to multi-index formulations such as the Graphical Turbulence Guidance (GTG). These methods are statistically-based because they do not compute EDR but must statistically infer EDR from the index's value.

To test each formulation, we gathered more than two million archived aircraft eddy dissipation rate (EDR) observations for 100 days during the 2017-2018 winter. We compared these observations to one hour EDR forecasts computed on the Rapid Refresh (RAP) numerical forecast model. Six of the forecasts are based on various gravity wave indentifying equations, the Lighthill-Ford equations, divergence tendency, frontogenesis, the Plougonven-Zhang equations, stability advection, and acceleration divergence. Additionally, we tested the GTG Version 3. Overall, as measured by the True Skill Score (TSS) , all the gravity wave formulations significantly outperformed GTG3 with acceleration divergence leading the way at about ten times better than GTG3. This was because more than 99% of the EDR observations were less than 0.05 $m^{2/3}$ s^{-1} and GTG3 forecasts more than half the observations greater than 0.1 $m^{2/3}$ s^{-1}. However, it is far more important to forecast significant turbulence than smooth. To that end, we tested for skill at EDR > 0.25 $m^{2/3}$ s^{-1}. Here, the Lighthill-Ford equations captured 33% while all the others were less including GTG3 with only 5%. Combining the Probability of Detection (POD) of EDR > 0.25 $m^{2/3}$ s^{-1} with the POD(EDR < 0.05 $m^{2/3}$ s^{-1}) yields a positive TSS for all the gravity wave forecasts with the highest being 0.180 for Lighthill-Ford. Additionally, we gathered 41 major, mostly injurious, CAT reports within the RAP numerical forecast model domain between 2010-2018. Here, the Lighthill-Ford equations captured 83% with EDR > 0.4 $m^{2/3}$ s^{-1} while all the others were less than half including GTG3 with only 17% with the same threshold. Combining the Probability of Detection (POD) of major turbulence with the POD(EDR < 0.05 $m^{2/3}$ s^{-1}) from the 2017-2018 winter data yields a positive TSS for all the gravity wave forecasts with the highest being 0.678 for Lighthill-Ford. In contrast, since the GTG3 is weak in both POD measures, its TSS = -0.355. We conclude that the gravity wave physical model is much better than any statistical model such as GTG3.

1. INTRODUCTION

The DTN Flight Hazards turbulence product provides a complete turbulence forecast up to 18 hours over most of North America and up to 36 hours globally. By “complete,” we mean turbulence in all its known causes, boundary layer, mountain wave, clear air, and convective, and at multiple flight levels from the surface to FL530[1]. First released in 2012, the algorithms for computing the turbulence were developed many years before, boundary layer (McCann 1999a), mountain wave (McCann 2006), clear air (McCann 2001; Knox et al. 2008; McCann et al. 2012), and convective (McCann 1999b; Byers and Braham 1949). Over 9000 aircraft are benefiting from DTN’s Flight Hazards today from over 90 contracted airlines and business jet operations, and this list continues to grow.

Perhaps the most vexing to forecast is clear air turbulence (CAT), and, indeed, much of the recent literature has been devoted to that. CAT forecasting began in the mid twentieth century when aircraft started to fly at higher altitudes and encountered unexpected turbulence without significant cloudiness. The first CAT forecasts were mostly empirical rules and equations which were based relationships between observed atmospheric patterns and aircraft turbulence reports. See Chapters 1 (Sharman 2016) and 2 (Fahey et al. 2016) in the book *Aviation Turbulence* (Sharman and Lane, editors 2016), for a more complete early history. Researchers developed single equation indices based on processes thought to initiate CAT. Roach (1970) and Ellrod and Knapp (1992) are examples. Sharman et al. (2006) recognized that the single equation indices did not capture CAT in many situations and so combined many indices into one, called the Graphical Turbulence Guidance (GTG). This approach has dominated CAT forecasting in recent years.

McCann (2001) introduced an alternative CAT forecasting model which is physically based. Gravity waves initiate CAT given a favorable environment. Gravity waves modify the stability ($N^2 = [g/\Theta]/[\partial\Theta/\partial z]$; g is the gravity acceleration and Θ is the potential temperature) and wind shear ($S = \partial \mathbf{V}/\partial z$; $\mathbf{V}$ is the vector wind) within the wave which can locally reduce the Richardson number ($Ri \equiv N^2/S^2$) to less

---

[1] Flight levels (FL) are the altitudes that aircraft fly and by international agreement are designated in hundreds of feet above sea level. When aircraft fly above FL180, they set their altimeters to read the height with respect to the standard atmosphere.

than 0.25, the critical value below which the atmospheric layer becomes turbulent (Miles and Howard 1964; Thorpe 1969). The amount of stability and wind shear that gravity waves modify is primarily a function of the waves' amplitude.

McCann (2001) identified some methods to find gravity wave amplitudes but none were especially skillful. A few years later Knox et al. (2008) (hereafter KMW) noted that spontaneous imbalance can generate gravity waves, and they showed that many CAT reports occurred in areas of forecast spontaneous imbalance using the so called Lighthill-Ford radiation equation. Immediately after publishing Plougonven et al. (2009) asserted that this theory is not applicable to the flows related to CAT. While Knox et al. (2009) responded to many of the objections that Plougonven et al. raised, they did not answer one important Plougonven et al. concern, "Without a more systematic investigation of the variations of the excited waves relative to the forcing, it only provides an indication, not compelling evidence for a generation mechanism."

So how does the McCann (2001) gravity wave initiation forecast model for CAT compare with a combination of CAT forecast indices? McCann et al. (2012) found superior results over GTG Version 2.5 using their implementation of the Lighthill-Ford radiation equation in McCann's gravity wave model called ULTURB. In this report we further investigate Lighthill-Ford, five additional gravity wave diagnostics, and the GTG Version 3 forecast using automated *in situ* aircraft eddy dissipation rate (EDR, $m^{2/3}$ s^{-1}) observations. Our objective is to examine the robustness of the physical gravity wave model. If the model is robust, then an additional goal is to identify which gravity wave diagnostic best forecasts CAT.

First we review the physical gravity wave model. Next we describe the six gravity wave diagnostics that we use to compare EDR forecasts with the model and with the statistical GTG3. Then we describe the experimental comparison setup. We present the results and finish with conclusions.

2. A SIMPLE GRAVITY WAVE CONCEPTUAL MODEL FOR CAT

Endlich (1964) may have been the first to notice a connection of wave motions in the vicinity of stability discontinuities and wind shear and CAT. McCann (2001) then modeled the quantitative relationship between waves, stability, and wind shear which this section will review and amplify. The model was introduced in KMW and McCann et al. (2012) but was underemphasized. We presently give it its due importance. The model is analogous to the Doswell et al. (1996) conceptual model for forecasting thunderstorms that allows forecasters to assess the likelihood for thunderstorms in a given environment. In that model, there needs to be a favorable environment and a mechanism that lifts an air parcel to saturation and eventually to a level where it convects freely. Analogously, CAT may be initiated in a favorable environment when a gravity wave passes through.

There is a special relationship between turbulence and the Richardson number. Turbulence is suppressed whenever $Ri > 0.25$. Miles and Howard (1964) derived this critical Ri value, and Thorpe (1969) experimentally verified it.

On sunny, windy days the Ri < 0.25 often in the planetary boundary layer. Above the boundary layer turbulent Richardson numbers are rare. Any atmospheric layers with Ri < 0.25 would quickly become turbulent, relieving the instability.

Gravity waves modify the temperature and wind, and thus the local stability and wind shear, as they pass through an environment. Using the dispersion relation for the vertically propagating wave component, Palmer et al. (1986) and Dunkerton (1997) derived a wave's impact on the local stability, $N = (\partial\Theta/\partial z)^{1/2}$ as

$$N_L^2 = N_E^{\,2}(1 + \hat{a}\cos\phi) \qquad (1)$$

and on the local vertical wind shear

$$\left(\frac{\partial \mathbf{V}}{\partial z}\right)_L = \left(\frac{\partial \mathbf{V}}{\partial z}\right)_E (1 + Ri_E^{\;1/2}\hat{a}\sin\phi) \qquad (2)$$

where the subscripts L and E are the local and environmental conditions, respectively, and ϕ is the angular position of the parcel within the gravity wave. The non-dimensional wave amplitude, $\hat{a} = Na/|\boldsymbol{V}\text{-}\boldsymbol{c}|$, where a is the actual wave amplitude and $\boldsymbol{c}$ is the wave's phase velocity; $|\boldsymbol{V}\text{-}\boldsymbol{c}|$ is the Doppler-adjusted wind speed. In a gravity wave both the stability and wind shear fluctuate around their environmental values. Equations (1) and (2) may be combined into a locally adjusted Richardson number.

$$Ri_L = Ri_E \frac{1 + \hat{a}\cos\phi}{(1 + Ri_E^{\;1/2}\hat{a}\sin\phi)^2} \qquad (3)$$

When $\phi = \pi$ and $\hat{a} > 1$, the numerator in (3) is negative, and the local Richardson number is less than zero thus causing turbulence. When $\phi = \pi/2$, wind shear is maximized. When $Ri_E^{1/2}$ $(2 - \hat{a}) < 1$, $Ri_L < 0.25$. Therefore, when $\hat{a} \geq 2$, Ri_L is always less than 0.25. When $\hat{a} < 2$, it is necessary for Ri_E to be below a particular value to lower Ri_L to below 0.25. That value is a function of $\hat{a}$ as depicted in Figure 1. As $\hat{a} \rightarrow 2$, the upper limit on Ri_E for turbulence approaches infinity. This implies that no matter the environmental Ri_E, any gravity wave with sufficient amplitude can initiate turbulence. The lower the turbulence-suppressing environmental Richardson number, the smaller the gravity wave needs to be to trigger turbulence. We point out that the environmental Richardson number is very analogous to the Doswell et al. (1996) thunderstorm ingredients-based model's convective inhibition (CIN) which measures the environmental suppression of thunderstorms. It takes less initiating lift when the CIN is low and more when the CIN is high.

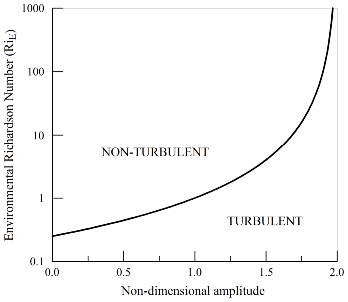


Figure 1. Curve of the bounding value of the environmental Richardson number as a function of the non-dimensional amplitude (â). When (â, $\mathbf{Ri_E}$) falls in the TURBULENT region, a gravity wave will locally increase the wind shear sufficiently to reduce the local Richardson number to less than 0.25. (From McCann 2001)

We can compute local Richardson numbers from observations or from numerical weather forecast models if we know gravity wave amplitudes. Theory is sufficient to estimate wave amplitudes generated by flow over mountains (McCann 2006). To estimate wave amplitudes in the free atmosphere requires two pieces of knowledge: 1) where is the flow generating these waves and 2) how strong are they. Free atmospheric gravity waves are theorized to form spontaneously as a result of unbalanced flow. In their chapter in *Aviation Turbulence* Plougonven and Zhang (2016) summarize much of the research. Unfortunately, they fall short of describing useful quantitative methods. In the next section we will describe six ways to diagnose gravity waves and their strength.

Before we do, there is one more aspect of this turbulence conceptual model that we will discuss. The model not only describes whether a gravity wave can initiate turbulence, but it also describes how to compute the turbulence intensity. A simple turbulent kinetic energy (TKE) equation describes the production and dissipation of TKE in terms of the stability (N) and wind shear (S) (Garratt 1992).

$$\varepsilon = K_m S^2 - K_h N^2 \tag{4}$$

where ε is the TKE dissipation rate due to molecular viscosity, K_m and K_h are eddy proportionality constants for momentum and heat, respectfully. Because the right hand side of (4) describes the TKE production rate, equation (4) simply states that over time, what TKE is produced is eventually dissipated. This production-to-dissipation process is described by L. F. Richardson's little poem:

> Big whirls have little whirls
> Which feed on their velocity,
> And little whirls have lesser
> whirls
> And so on to viscosity.

The TKE production often occurs in large eddies through which aircraft ride smoothly. Then the TKE cascades to smaller eddies that aircraft do feel. Eventually eddies become too small to affect aircraft.

Substituting (1) and (2) into (4) will compute TKE production/dissipation after a gravity wave modifies the environment.

$$\begin{aligned}\varepsilon = K_m \left(\frac{\partial \mathbf{V}}{\partial z}\right)_E^2 \left(1 + Ri_E^{1/2}\hat{a}\sin\phi\right)^2 \\ - K_h N_E^2 \left(1 + \hat{a}\cos\phi\right)\end{aligned} \tag{5}$$

The derivative of (5) with respect to ϕ will find the ϕ that maximizes ε for the given environment. Unfortunately, it is not a straightforward calculation, but it can be done numerically. To do this on numerical forecast model data takes substantial computations, so we simplify the process by assuming $\phi = \pi$ to calculate maximum ε due to buoyancy production,

$$\varepsilon_L = K_m S_E{}^2 - K_h N_E^2 (1 - \hat{a}) \tag{6}$$

and by assuming $\phi = \pi/2$ to calculate maximum ε due to wind shear production,

$$\varepsilon_L = K_m S_E^2 (1 + Ri_E^{1/2}\hat{a})^2 - K_h N_E{}^2 \tag{7}$$

We choose the maximum of the two production computations as our observation/forecast. Any negative TKE production/dissipation is set to zero. As a practical matter, wave breaking limits wave amplitudes, so we set the maximum $\hat{a} = 2.5$. We convert the resulting TKE dissipation to EDR, the standard for turbulence reporting, by taking the cube root of the TKE dissipation.

Depending on the location within the gravity wave (ϕ), (4) says that some portions of the gravity wave will not be turbulent and most of the remaining wave will be less turbulent than the maximum. We want to alert users of the strongest turbulence possible. Therefore, users should not expect maximum forecast turbulence intensity but should prepare for it. This is analogous to tornado watch forecasts which even the best watches grossly overforecast the actual threat.

We have introduced the necessary ingredients for this turbulence model. The environmental Richardson number/gravity wave non-dimensional amplitude amounts must be sufficient for the local Richardson number to be less than 0.25. If so, then we compute the local TKE production/dissipation maxima due to buoyancy and wind shear for an intensity diagnostic. Conceptually, this is fairly simple model to understand, but, as we will show in the next section, computing the non-dimensional gravity wave amplitude is not so simple.

3. DIAGNOSING GRAVITY WAVES

That gravity waves are observed in turbulence zones is undeniable. Intensive observations of mountain waves by Lilly and Kennedy (1973) and Lilly (1978) directly connected gravity waves and turbulence. Observations connecting free gravity waves with turbulence began to appear about the same time. In a CAT case over the eastern United States, Reed and Hardy (1972) observed waves aloft on a 10 cm radar in turbulent zones reported by research and commercial aircraft. Hooke and Hardy (1975) complemented that study with surface microbarograph data. Bedard et al. (1986) also connected microbarograph data with another CAT case. More recent observational campaigns, summarized in *Aviation Turbulence* by Plougenven and Zhang (2016), have further confirmed a gravity wave/turbulence connection.

The next step to making CAT forecasts using the gravity wave model is to identify areas and altitudes where the atmosphere is generating gravity waves. KMW implemented this model using the Lighthill-Ford radiation diagnostic. But that is not the only diagnostic that researchers have suggested. In this section we will describe six such diagnostics beginning with Lighthill-Ford.

a. Lighthill-Ford Radiation

James Lighthill (1952) described acoustic waves emitted from vortical flows. Then, Rupert Ford (1994) extended the problem to rotating shallow-water flow and gravity wave generation. Williams et al. (2005) set up a rotating annulus experiment and observed gravity waves. Rotating annulus experiments may be set up to be dynamically similar to atmospheric flow. Figure 2, from Williams et al. (2008), shows a snapshot of what they observed. They tested five possible generation mechanisms and found the Lighthill-

Ford theory best described the gravity waves.

KMW's contribution was to render the Lighthill-Ford radiation equation into terms which are computable on gridded meteorological data. The complete equation is the sum of three terms each of which is a sum of three subterms. Scaling each of the nine subterms reduced the equation to one leading order subterm and three second order subterms:

$$R = f\mathbf{V}\cdot\nabla\zeta + 2f\delta\zeta - f\mathbf{k}\cdot\mathbf{V}\times\nabla\delta - 2\frac{\partial}{\partial t}J(u,v) \tag{7}$$

where R is the Lighthill-Ford radiation, f is the Coriolis parameter, ζ is the relative vorticity, δ is the divergence, $\mathbf{k}$ is the unit vector pointing upward, and $J(u,v)$ is the Jacobian operator on the east and north wind components.

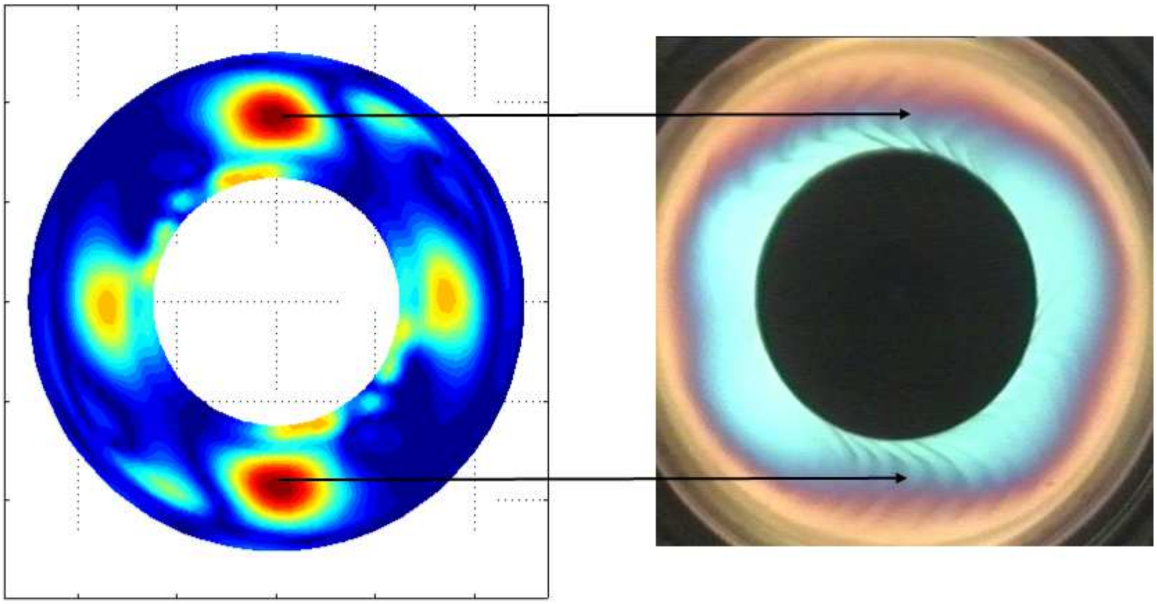


Figure 2. (Right) Waves observed in rotating annulus. (Left) Lighthill-Ford radiation computed from the parameters in the rotating annulus experiment. Arrows pair the computed Lighthill-Ford maxima to the locations in the experiment. Figure provided by Paul Williams (Williams, et al. 2003, 2005).

The first term on the right hand side of (7) is the Coriolis parameter times the relative vorticity advection and is the leading order subterm. Vorticity is often displayed on upper air charts, and the advection of it is easily visualized when it is overlaid on streamlines or even height contours. The other three second order subterms are not so easily visualized. Fortunately, all four subterms can be computed on gridded data such as numerical forecast models.

Plougonven et al. (2009) insist that Lighthill-Ford theory is inappropriate to the gravity waves observed in jets and fronts. They argue that Lighthill-Ford waves emanate from their source region and so are not located where they are generated. Furthermore, a portion of the Lighthill-Ford radiation may not be forcing gravity waves at all. Instead some of the forcing may cause adjustments to the balanced flow, for example, frontogenesis.

b. Divergence Tendency

Taking $(\nabla \cdot)$ the equation of motion and arranging terms results in the divergence tendency equation:

$$\frac{\partial \delta}{\partial t} = -\nabla^2 \Phi + 2J(u,v) + f\zeta - \beta u \quad (8)$$

where Φ is the geopotential height and $\beta = df/dy$. When $\partial\delta/\partial t = 0$, the equation is known as the nonlinear balance equation. When the divergence tendency is zero, the wind and the mass fields are in balance (Charney 1955). Growth in divergence ($\partial\delta/\partial t > 0$), leads to unbalanced flow (Uccellini and Koch 1987; Zack and Kaplan 1987). A nonzero divergence tendency is also known as the "nonlinear balance equation residual."

c. Frontogenesis

The processes by which fronts generate gravity waves are probably the most well- documented generation mechanisms. Fronts undergoing spatial contraction or frontogenesis exhibit gravity wave characteristics (Shapiro 1984), can excite gravity currents that propagate ahead of the front (Crook 1988), act as an obstacle to the flow (Ralph et al. 1993), or develop cross-front accelerations leading to gravity waves (Snyder et al. 1993). Unfortunately, these are studies that focus on surface fronts, not ones aloft where aircraft fly. Fronts develop aloft differently from those at the surface (Keyser and Shapiro 1986). Nevertheless, Koch et al. (2005) found extensive gravity wave activity in an upper front.

Both surface and upper frontogenesis describe the rate of temperature gradient increase (F, K m^{-1} s^{-1})

$$F = \frac{1}{2}\left|\nabla \Theta\right|\left(DEF \cos 2\vartheta - \delta\right) \quad (9)$$

where DEF is the total deformation

$$DEF = \left(SHR^2 + STR^2\right)^{1/2}$$

and ϑ is the angle between the axis of dilatation of the shear deformation (SHR) and the stretching deformation (STR) and the isolines of Θ.

d. Plougonven and Zhang (2007) Forcing

Plougonven and Zhang (2007) derived a gravity wave forcing equation written in dimensional variables as

$$\begin{aligned} G = & \frac{d}{dt}\frac{\partial}{\partial z}\left(\frac{\partial \delta}{\partial t}\right) \\ & + f\frac{\partial}{\partial z}\left(\frac{d\zeta}{dt} + f\delta + \beta v\right) \quad (10) \\ & - \frac{g}{\Theta}\nabla^2\left(\frac{d\Theta}{dt} + w\frac{\partial \Theta}{\partial z}\right) \end{aligned}$$

where G is the gravity wave forcing and the total derivative $d/dt = (\partial/\partial t + \mathbf{V}\cdot\nabla)$. Equation (10) is a quantitative diagnostic of the large-scale flow to be used as an indicator of gravity wave initiation. Each of the terms can be shown to be vertical acceleration terms through various vertical motion equations.

In order to solve this diagnostic on gridded data, we assume that the local derivatives ($\partial/\partial t$) are negligible compared with the advection **($\mathbf{V}\cdot\nabla$).** This is a reasonable assumption for the divergence tendency [$\partial(\partial\delta/\partial t)/\partial t$], the vorticity ($\partial\zeta/\partial t$), and the potential temperature ($\partial\Theta/\partial t$), all of which are usually small outside of convective situations. We also assume that the vertical advection of potential temperature [$w(\partial\Theta/\partial z)$] is also small in non-convection.

Plougonven and Zhang (2007) left the relevance of their diagnostic to future research. Wang and Zhang (2010) assessed the contributions of each of the forcing terms on an idealized vortex-dipole jet system. They found that the second (vorticity) term and particularly the vorticity advection subterm dominated the gravity wave forcing. As far as we know, we are the first to compute these forcing terms on real atmospheric data. Since no one else has shown similar results, we compute all three terms with the assumptions noted previously.

e. Stability Advection

Yasuda et al. (2015a) derive a theory of gravity wave generation coming from the divergence and ageostrophic vorticity components of the balanced flow. They derive a complete description of gravity wave generation and radiation using the renormalization group (RG) method (Chen et al. 1994, 1996). The RG equations are very complex and are difficult to compute. However, in the second of their papers, Yasuda et al. (2015b) describe how the RG equations are consistent with both a mountain wave-like generation mechanism (McIntyre 2009) and a velocity-variation generation mechanism (Viúdez 2007). We will discuss the second mechanism in the next subsection.

Figure 3 shows how gravity waves develop when flow moves through deformed potential temperature surfaces. Yasuda et al. (2015b) show that these deformed surfaces are depicted well in the stability field with stability defined as in Section 2 as the square root of the vertical potential temperature (θ) gradient. In Fig. 3 the highest stability is where the θ lines are the closest together vertically. The stability is analogous to terrain elevation. Just as waves form as air flows down a mountain slope, waves also form as air flows to lesser stability. We capture this process as stability advection.

$$G_\theta = \mathbf{V}\cdot\nabla N \qquad (11)$$

f. Acceleration Divergence

The Yasuda et al. RG equations also explain the Viúdez (2007) velocity-variation gravity wave generation mechanism. As shown in Figure 4, whenever the flow exhibits acceleration convergence, gravity waves will be emitted. We estimate flow acceleration with the Uccellini et al. (1984) inertial advective wind. The inertial advective wind is defined as $\mathbf{V}\cdot\nabla\mathbf{V}$ and is the total acceleration vector minus the local change in the wind velocity. As with the Plougonven-Zhang equation earlier, we assume the local change is minimal except in convective areas.

$$G_\delta = \nabla\cdot(\mathbf{V}\cdot\nabla\mathbf{V}) \qquad (12)$$

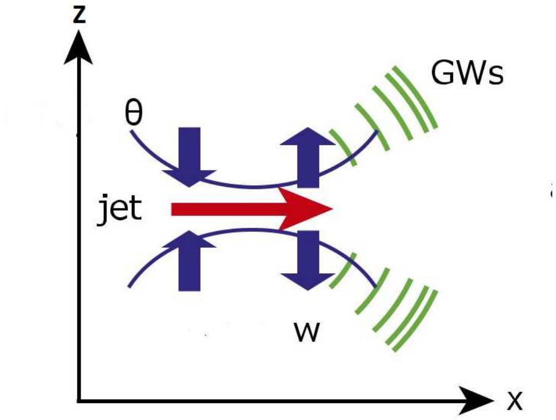


Figure 3. Schematic drawing of gravity wave radiation from a jet stream moving over deformed potential temperature (θ) surfaces. The blue arrows indicate gravity wave (GW) generating vertical motions. This mechanism is like flow over a mountain. Taken from Yasuda et al. (2015a)

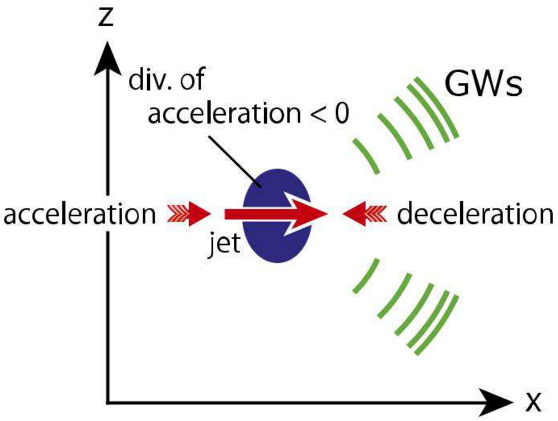


Figure 4. Schematic drawing of gravity wave radiation from a jet stream maximum. The jet is located by the long red arrow. The two short red arrows indicate flow acceleration and deceleration. The blue ellipse is the region where the acceleration has horizontal convergence. Gravity waves (GW) are generated as a result of this acceleration convergence. Taken from Yasuda et al. (2015a).

4. TESTING THE GRAVITY WAVE DIAGNOSTIC AND GTG3

When KMW computed the Lighthill-Ford diagnostic, it had to be converted into a non-dimensional gravity wave amplitude so it could be used as input into the turbulence model equations. How to make such a conversion has not been theoretically derived or observationally observed. However, Williams et al. (2008) noted that the gravity wave amplitudes in laboratory experiments vary linearly with Rossby number, and the Rossby number varies with the square root of the Lighthill-Ford radiation term. Therefore KMW assumed the same relationship:

$$\hat{a} \propto \sqrt{R}$$

KMW empirically found a proportionality constant by examining the range of R in several CAT outbreaks, then, because the maximum $\hat{a}$ is 2.5, the constant was (max R)/2.5^2.

For this experiment we assume that the other gravity wave diagnostics have a similar relationship. McCann (2001) gave value ranges for the divergence tendency. For frontogenesis, the Plougonven-Zhang diagnostic, stability advection, and acceleration divergence we found maximum values similarly as KMW found for Lighthill-Ford. As it turns out, computed EDRs were only somewhat sensitive to the constants we found as long as they were reasonably representative. Using the ULTURB software (McCann et al. 2012) as a template, we wrote similar programs for each of the other five gravity wave diagnostics, substituting the appropriate diagnostic for the Lighthill-Ford one.

We created a combined clear air and mountain wave GTG3 above FL200 (Sharman and Pearson 2017) replica forecast following their method. Because we only had archived verification data (see below), we had to be able to create forecasts from past model data. Our GTG3 version is not exact because the official GTG3 dynamically varies its weights for each input diagnostic. In our facsimile we weighted all input diagnostics equally as Sharman and Pearson suggest. While there are some differences, our version using the Rapid Refresh (RAP) forecast model looks very similar to official GTG3 forecasts as seen in Figure 5 and in other comparisons.

We verified the physical gravity wave model using each of the six gravity wave diagnostics and our GTG3 forecast against automated *in situ* aircraft turbulence observations (Cornman et al. 1996). Onboard aircraft software analyzes aircraft movement to determine an aircraft-independent turbulence measurement. Observations can be transmitted as frequently as one per minute and include the average eddy dissipation rate (EDR) and the maximum EDR for the previous minute. Observations are archived on the National Centers for Environmental Prediction (NCEP) Meteorological Assimilation Data Ingest System (MADIS) website (www.madis.ncep.noaa.gov). We retrieved maximum EDR observations at or above FL200 plus or minus ten minutes of the top of every hour every day from 25 November 2017 to 4 March 2018. Our goal was to compare CAT gravity wave model forecasts and GTG3 with CAT observations. The chosen 100 day period climatologically provided the most CAT observations relative to convective induced turbulence observations. The maximum EDR observations were rounded to the nearest 0.1 $m^{2/3}$ s^{-1} and put into bins of 0.0, 0.1, 0.2, 0.3, 0.4, and 0.5 or greater $m^{2/3}$ s^{-1}. We averaged more than 21 000 observations per day distributed as in Table 1. This distribution is very similar to other larger studies, i.e. Sharman et al. (2006).

EDR	Percentage
0.0	99.265
0.1	0.655
0.2	0.071
0.3	0.007
0.4	0.0009
>0.5	0.0002

Table 1. Percentage distribution of EDR observations in each 0.1 $m^{2/3}$ s^{-1} EDR bin.

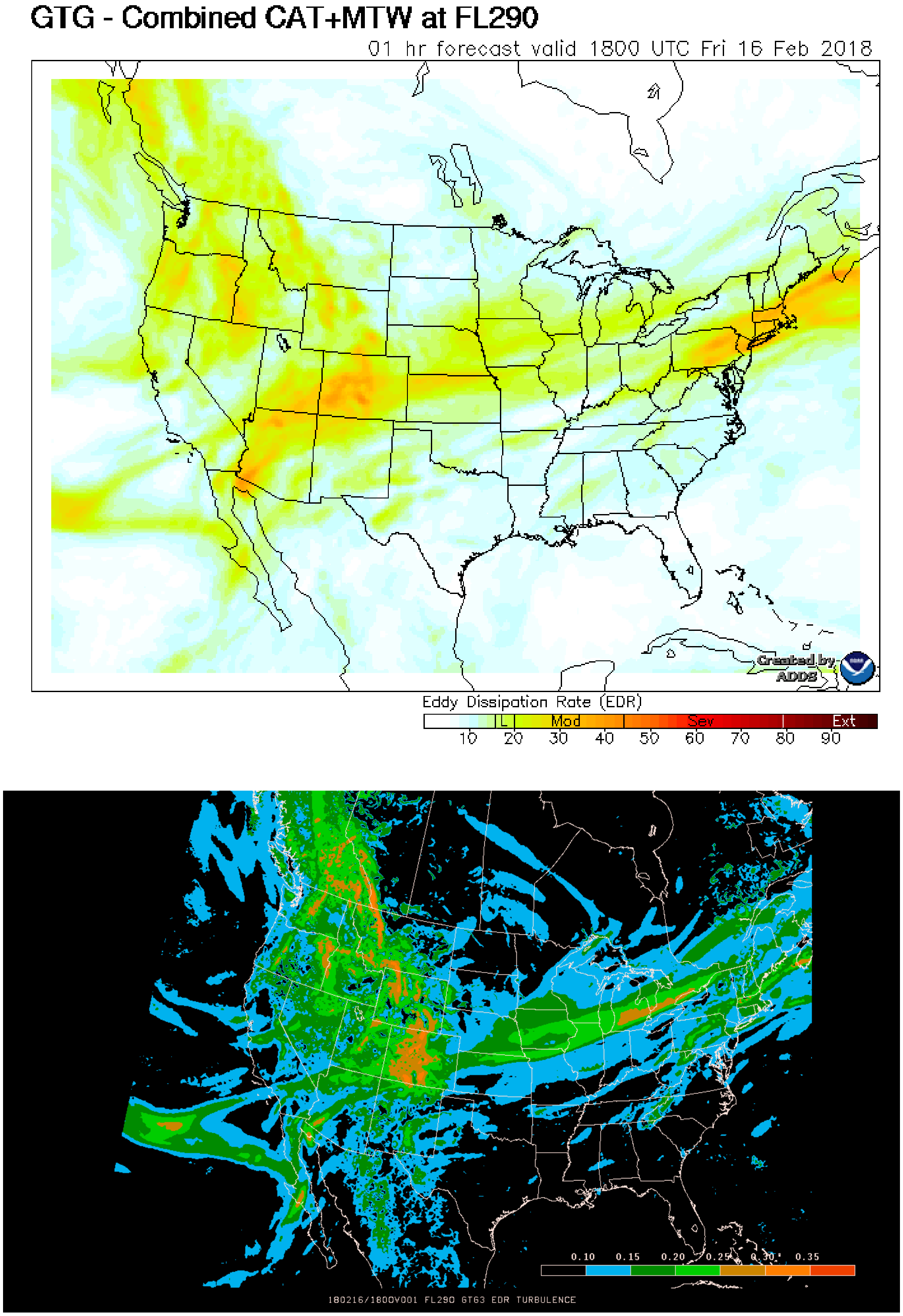
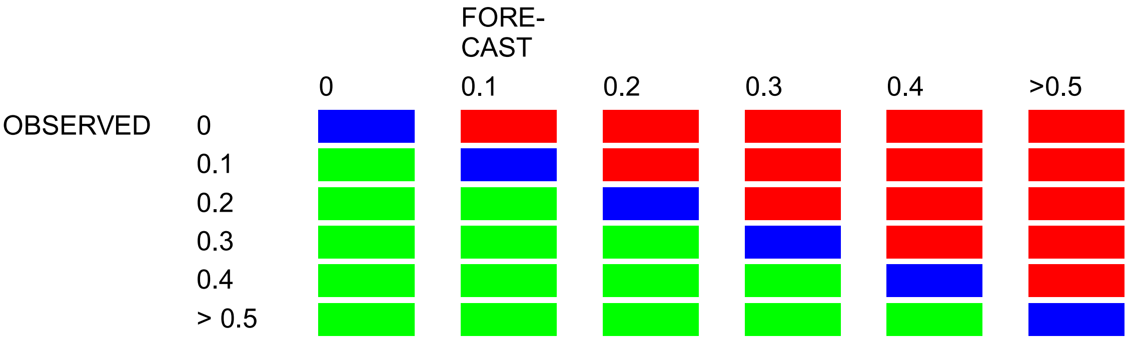


Figure 5. (Top) One hour combined GTG3 forecast at fL290 valid 1800UTC 16 February 2018 taken from the Aviation Weather Center website (www.aviationweather.gov). (Bottom) Our GTG3 replica for the same forecast time and altitude.

We matched the EDR observations with archived Rapid Refresh (RAP) numerical forecast model data retrieved from the National Centers for Environmental Information (www.ncdc.noaa.gov/nomads/data-products). The archived RAP model forecasts cover the contiguous United States, southern Canada, northern Mexico, and adjacent waters. The model runs every hour with archived forecasts available hourly on top of each hour out to 18 hours. We excluded any EDR observation outside the RAP model domain. For each first hour RAP model forecast we ran each of our gravity wave turbulence algorithms and our GTG3 version, all of which produced EDR forecasts. For each EDR observation ±10 minutes that forecast time we interpolated the forecast EDR to the exact location of the observation and rounded the forecast to the nearest 0.1 $m^{2/3}$ s^{-1}. We then binned them the same as the observations. We could pair each observation/forecast in a six by six contingency table as in Figure 6. Because we did not distinguish turbulence observations caused by CAT mechanisms, mountain waves, or convection, our analyses of the 6 X 6 contingency tables do not absolutely compute skill for CAT-only forecasts. But since our goal was to establish superiority of one method over another, our analyses are relative.



		FORE-CAST					
		0	0.1	0.2	0.3	0.4	>0.5
OBSERVED	0	🟦	🟥	🟥	🟥	🟥	🟥
	0.1	🟩	🟦	🟥	🟥	🟥	🟥
	0.2	🟩	🟩	🟦	🟥	🟥	🟥
	0.3	🟩	🟩	🟩	🟦	🟥	🟥
	0.4	🟩	🟩	🟩	🟩	🟦	🟥
	> 0.5	🟩	🟩	🟩	🟩	🟩	🟦

Figure 6. Observed and forecast EDRs are paired and put into the appropriate place in the 6 X 6 contingency table. Blue represents accurate forecasts, red represents overforecasts, and green represents underforecasts. The overall 6 X 6 contingency table may be easily reduced to smaller degrees such as 2 X 2 depending what statistics are desired.

First, we show how well each turbulence forecast fared overall. We calculated a 6 X 6 True Skill Score (Doswell et al. 1990)

$$TSS = \frac{tr(\mathbf{R})}{tr(\mathbf{R}^*)} \tag{13}$$

where **R** is a matrix of the contingency table values minus their expected value if the values were random and **R*** is a similar matrix of perfect forecast values. The *tr* function is the diagonal of the matrix (blue values in Fig. 6). For these contingency tables the TSS measures how well the forecasts are within 0.1

		FORECAST						
		0	0.1	0.2	0.3	0.4	>0.5	
OBSERVED	0	2078177	34525	23322	12586	4210	918	2153738
	0.1	11619	974	851	419	160	99	14122
	0.2	1180	100	102	83	44	33	1542
	0.3	115	13	9	6	8	3	154
	0.4	13	1	3	0	1	1	19
	>0.5	2	2	0	1	0	0	5
		2091106	35615	24287	13095	4423	1054	2169580

Table 2. The 6 X 6 verification contingency table for turbulence forecasts using frontogenesis.

Diagnostic	PODsmooth	TSS
Lighthill-Ford	.849	.022
Divergence Tendency	.946	.016
Frontogenesis	.965	.034
Plougonven-Zhang	.928	.036
Stability Advection	.989	.072
Acceleration Divergence	.994	.078
GTG3	.474	.0076

Table 3. Probability of Detection of EDR< 0.05 $m^{2/3}$ s^{-1} (smooth) and overall True Skill Scores (TSS) for turbulence forecasts each of the six gravity wave diagnostics and for GTG3.

$m^{2/3}$ s^{-1} of the observations. These are stringent criteria, so because the EDR observations may not all be CAT-related, we are not looking for absolute values but are looking for relative skill of one method over another. The TSS measures how well the observations fall into the contingency table's diagonal (blue in Fig. 6). Table 2 is the 6 X 6 contingency table for frontogenesis.

Table 3 shows the overall verification statistics for turbulence forecasts using each of the six gravity wave diagnostics and for GTG3. The highest TSS for acceleration divergence is about ten times the lowest TSS for GTG3. Even the low TSS for divergence tendency is more than twice GTG3's. There is a rough correlation between the Probability of Detection of EDR < 0.05 $m^{2/3}$ s^{-1} (hereafter "smooth") and the TSS. The fraction of the grid volume of forecast EDR to the total grid volume helps explain each forecast's TSS. This fraction is the number of grid points between and including FL200 and FL400 forecast above a certain threshold to the total number of grid points. Table 4 shows these fractions. Given that more than 99% of the observations were smooth, it is not surprising that those methods that forecast small fractions of grid volume have the highest overall TSSs. That nearly half of GTG3's grid point volume has EDR > 0.05 $m^{2/3}$ s^{-1} explains why its overall TSS is so low.

We generated bias statistics of our observations/forecasts. The bias is simply the number of forecasts divided by the number of observations above a certain threshold. Table 5 shows these bias statistics. What immediately jumps out is the grossly overforecast of the turbulence by almost all methods, especially when forecasting EDR > 0.25 $m^{2/3}$ s^{-1}. The highest biases are for Lighthill-Ford.

	EDR Forecast greater than				
Forecast	0.05	0.15	0.25	0.35	0.45
Lighthill-Ford	.130	.075	.030	.008	.0015
Divergence Tendency	.054	.039	.019	.003	.0003
Frontogenesis	.024	.017	.008	.002	.0005
Plougonven-Zhang	.068	.043	.013	.003	.0007
Stability Advection	.011	.005	.002	.0002	.00006
Acceleration Divergence	.007	.002	.0004	.00004	.000003
GTG3	.445	.053	.003	.00003	.000006

Table 4. The forecast fraction of grid points above various thresholds. This fraction is the number of grid points between and including FL200 and FL400 forecast above a certain threshold to the total number of grid points.

	Bias for EDR Forecast greater than				
Forecast	0.05	0.15	0.25	0.35	0.45
Lighthill-Ford	10	79	318	716	518
Divergence Tendency	1.8	33	237	294	60
Frontogenesis	2.5	16	85	233	211
Plougonven-Zhang	8.0	19	106	171	187
Stability Advection	1.2	4.5	11	14	1.2
Acceleration Divergence	0.8	2.0	4.1	2.1	0.4
GTG3	71	90	64	12	0.2

Table 5. Bias statistics for the various gravity wave diagnostics and GTG3. The bias is the number of forecasts divided by the number of observations above a certain threshold.

Forecast	POD(EDR>0.25)	POD(EDR<0.05)	TSS
Lighthill-Ford	.331	.849	.180
Divergence Tendency	.095	.946	.041
Frontogenesis	.112	.965	.077
Plougonven-Zhang	.179	.928	.107
Stability Advection	.051	.989	.040
Acceleration Divergence	.033	.994	.027
GTG3	.051	.474	-.475

Table 6. Probabilities of Detection of EDR > 0.25 $m^{2/3}$ s^{-1} and EDR < 0.05 $m^{2/3}$ s^{-1} and the resulting True Skill Score for each turbulence forecast method using automated aircraft EDR observations.

While the overall TSSs mostly indicate how well the methods forecast smooth turbulence, users want to avoid the strong turbulence. To measure that skill, we reduced the 6 X 6 contingency tables to 2 X 2 tables setting an EDR > 0.25 $m^{2/3}$ s^{-1} threshold for PODyes and EDR < 0.05 $m^{2/3}$ s^{-1} for PODno. For a 2 X 2 table, the TSS is simply (PODyes + PODno – 1). Table 6 shows the results. Here it is the methods with large volumes of high EDR that excel with the Lighthill-Ford method leading the way. Not only does GTG3 poorly forecast smooth turbulence, but it also doesn't forecast strong turbulence well thus yielding a very low negative skill.

Additionally, we have been saving clear air turbulence case studies on which we have been testing ULTURB (McCann et al. 2012) which uses McCann's (2001) physical gravity wave model with the Lighthill-Ford gravity wave diagnostic. We have 41 cases within the RAP operational numerical forecast model domain above FL200 between 2012-2018 and within the Rapid Update Cycle model between 2010-2012. We gathered the cases from the archives of the *Aviation Herald* (www.aviationherald.com). The *Aviation Herald* reports on commercial aircraft mishaps worldwide, and the turbulence reports were primarily when the turbulence caused injuries. The *Aviation Herald* does not report on smaller general aviation aircraft, nor does it report every incident or accident. The *Aviation Herald's* editor verifies each report from two independent sources or from a government aviation safety agency. Of course, there were many more than 41 turbulence reports in the archive, so we confirmed that each report was clear-air by eliminating reports associated with convection by examining satellite imagery. Because it can be difficult to distinguish mountain wave turbulence from clear air turbulence, we did not filter our data for the former.

The *Aviation Herald* reports typically gave the event's time, location, and altitude. Even with that information, the incident was often reported when the aircrew took action rather than when the turbulence occurred. Whenever the report was inadequate in some way, we obtain the flight's track from FlightAware (www.flightaware.com).

We assumed the incident could have occurred as much as 10 minutes prior to the reported time. We discarded any report that we could not resolve to this accuracy.

We verified each of the 41 cases similarly as we did for the automated EDR observations. If the EDR > 0.4 $m^{2/3}$ s^{-1}, we considered it a "hit" for the tested diagnostic. EDR = 0.4 $m^{2/3}$ s^{-1} is about moderate to severe turbulence for most commercial aircraft (Sharman et al. 2014). We assume that the PODno statistic computed from our automated EDR observations is representative of each diagnostic's smooth turbulence forecasting skill in general.

Forecast	POD(EDR>0.40)	POD(EDR<0.05)	TSS
Lighthill-Ford	.829	.849	.678
Divergence Tendency	.366	.946	.312
Frontogenesis	.341	.965	.306
Plougonven-Zhang	.463	.928	.391
Stability Advection	.195	.989	.184
Acceleration Divergence	.122	.994	.116
GTG3	.171	.474	-.355

Table 7. Probabilities of Detection of EDR > 0.40 $m^{2/3}$ s^{-1} and EDR < 0.05 $m^{2/3}$ s^{-1} and the resulting True Skill Score for each turbulence forecast method using 41 cases of significant turbulence found in the *Aviation Herald* archives 2010-2018.

	Lighthill-Ford CAT forecast	
Statistic	Original	Reduced
TSS overall	.022	.052
POD(EDR<0.05 $m^{2/3}$ s^{-1})	.849	.952
TSS (EDR>0.25 $m^{2/3}$ s^{-1})	.180	.115
TSS (EDR>0.40 $m^{2/3}$ s^{-1})	.678	.464
Forecast Fraction (smooth)	.130	.039
Forecast Fraction (EDR>0.25 $m^{2/3}$ s^{-1})	.030	.0007

Table 8. Statistics comparing CAT forecasts from the original Lighthill-Ford diagnostic to one reduced by half.

Table 7 shows the results for each diagnostic. Four of the diagnostics captured more than 30% of the *Aviation Herald* reports indicating that the physical gravity wave model describes well the clear air turbulence production process. The Lighthill-Ford diagnostic is by far the best of the bunch, finding 83% of the reports. In fact, none of the other diagnostics captured any of the seven reports that Lighthill-Ford missed. Williams et al. (2013) suggested that some reported turbulence is wake turbulence caused by other nearby aircraft. Thus, some of the "misses" may be explained in this fashion.

As an additional experiment, we halved the proportionality constant that converts the Lighthill-Ford diagnostic to non-dimensional gravity wave amplitude. Because of the assumed square root relationship between any diagnostic and gravity wave amplitude, this reduced the effect to a quarter. Table 8 compares the original and the reduced Lighthill-Ford diagnostics. While the reduced Lighthill-Ford diagnostic forecast was better overall with a smaller forecast fractions and smaller biases, it suffered when forecasting stronger turbulence.

The results from this additional experiment and the overall statistics illustrate the ever apparent tradeoff between smaller forecasts and forecasts of significant events. While the ideal forecast will be small and highlight the strong CAT, it is obvious that the research has not advanced far enough to date to do so.

Furthermore, we would like to comment on the inadequacy of the GTG3 as a CAT forecast. From the forecast fractions in Table 4, the GTG3 forecasts much more positive turbulence, nearly half the forecast volume, compared with any diagnostic used in the gravity wave initiating conceptual model presented in section 2. Moreover, GTG3's forecasts of significant turbulence, i.e. EDR > 0.35 $m^{2/3}$ s^{-1} are nearly non-existent. We conclude that GTG3 forecasts both ends of the turbulence spectrum poorly.

We entered into this experiment with the idea of augmenting or changing our Lighthill-Ford diagnostic with another one if it were to improve our turbulence forecasts. As it turns out, we believe that we already have the best diagnostic and are distributing the most valuable clear air turbulence forecasts to our customers. These forecasts are available from DTN (https://www.dtn.com/industries/weather/aviation/).

5. CONCLUSIONS

We presented a simple ingredients-based conceptual model of CAT. Gravity waves locally alter the environmental stability and wind shear when passing through. If the modification can lower the Richardson number to less than 0.25, then CAT will develop. The turbulence maximum potential intensity is computed from the modified stability and wind shear. Thus, we have three ingredients, environmental stability and environmental wind shear from which we compute an environmental Richardson number and gravity wave non-dimensional amplitudes.

We compared six gravity wave indicators' turbulence forecast skill on automated aircraft EDR observations using the physical gravity wave conceptual CAT model. Because of the overwhelming number of smooth observations, overall, the acceleration divergence indicator, which forecast the smallest airspace volume, fared the best. In contrast, the Lighthill-Ford indicator, which forecast the largest airspace volume, did the best on stronger turbulence, EDR > 0.25 $m^{2/3}$ s^{-1}. When we tested the gravity wave indicators on 41 significant turbulence events occurring between 2010-2018, the Lighthill-Ford indicator excelled over all the others with an 83% POD at EDR > 0.4 $m^{2/3}$ s^{-1}. Because the Lighthill-Ford indicator was not too far behind the other gravity wave indicators in forecasting smooth, its True Skill Score, 0.678, was much higher than the others, all < 0.4. The statistically-based GTG3 forecast paled in comparison to the physical gravity wave model forecasts at both ends of the turbulence spectrum. Its overall skill score was two to ten times lower than the gravity waves', depending on which method was the GTG3 compared, and its 17% POD for significant turbulence was near the bottom, resulting in a -0.355 TSS.

Because of the high bias for forecasting positive CAT for the physical gravity wave conceptual model, especially the Lighthill-Ford version, one should not take its forecasts too literally. Users should interpret them as forecasting CAT potential, not actual CAT; an aircraft entering into a high EDR volume is not likely to experience the maximum EDR forecast, but it could. This is analogous to tornado watches which cover vast areas compared to the tornadoes' actual areal extent.

We can not overemphasize how important the physical gravity wave conceptual CAT model is to diagnosing and forecasting CAT. Without its framework, gravity wave indicators are not successful (e.g. Wilson 2012). Within its framework all of the gravity wave indicators that we tried work with a varying degree of success. We wouldn't be surprised if other more traditional CAT indices that may be related to gravity wave generation were applied in this simple model's context, they would have some success also.

DTN recognizes that clear air, mountain wave, and convective turbulence may interact with each other. Therefore they now integrate forecast algorithms of each (Lennartson and McCann 2014). Airlines and other users are discovering that they can proactively avoid costly turbulence events both from a financial as well as a brand erosion aspect. Significant turbulence events often make headlines, and when they do, the traveling public's fear level increases. The already serious turbulence forecast problem is expected to worsen as the climate warms (*Aviation Turbulence*, Williams and Joshi 2016). The industry needs good turbulence forecasts, better than those supplied to it today. A physically-caused turbulence conceptual model such as the one presented in this report stands to be more successful than a statistically-based one.

ACKNOWLEDGEMENTS

This report was commissioned by DTN Weather, Burnsville, MN. We thank Wes Hyduke for his comments.

REFERENCES

Bedard, A.J. Jr., F. Canavero, and F. Einaudi, 1986: Atmospheric gravity waves and aircraft turbulence encounters. *J. Atmos. Sci.*, **43**, 2838-2844.

Charney, J., 1955: The use of the primitive equations of motion in numerical prediction. *Tellus*, **7**, 22-26.

Chen, L.-Y., N. Goldenfeld, and Y. Oono, 1994: Renormalization group theory for global asymptotic analysis. *Phys. Rev. Lett.*, **73**, 1311–1315.

Chen, L.-Y., N. Goldenfeld, and Y. Oono, 1996: Renormalization group and singular perturbations: Multiple scales, boundary layers, and reductive perturbation theory. *Phys. Rev.*, **54E**, 376–394.

Cornman, L.B., C.S. Morse, and G. Cunning, 1995: Real-time estimation of atmospheric turbulence severity from in-situ aircraft measurements. *J. Aircraft*, **32**, 171-177.

Crook, N.A., 1988: Trapping of low level internal gravity waves. *J. Atmos. Sci.*, **45**, 1533-1541.

Doswell, C.A. III, H.E. Brooks, and R.A. Maddox, 1996: Flash flood forecasting: An ingredients based methodology. *Wea. Forecasting*, **11**, 560-581.

Doswell, C.A. III, R. Davies-Jones, and D.L. Keller: On summary measures of skill in rare event forecasting based on contingency tables, *Wea. Forecasting*, **5**, 576-585.

Dunkerton, T.J., 1997: Shear instability of inertia-gravity waves. *J. Atmos. Sci.*, **54**, 1628-1641.

Ellrod, G.P. and D.I. Knapp, 1992: An objective clear-air turbulence forecasting technique: Verification and operational use. *Wea. Forecasting*, **7**, 150-165.

Endlich, R.M., 1964: The mesoscale structure of some regions of clear-air turbulence. *J. Appl. Meteor.*, **3**, 261-276.

Fahey. T., E.N. Wilson, R. O'Loughlin, M. Thomas, and S. Klipfel: A history of weather reporting from aircraft and turbulence forecasting for commercial aviation. *Aviation Turbulence*, Springer International Publishing, Switzerland, ISBN 978-3-319-23629-2, Chapter 2, 31-58.

Ford, R., 1994: Gravity wave radiation from vortex trains in rotating shallow water. *J. Fluid Mech.*, **281**, 81-118.

Hooke, W.H. and K.R. Hardy, 1975: Further study of the atmospheric gravity waves over the Eastern Seaboard on 18 March 1969. *J. Appl. Meteor.*, **14**, 31-38.

Garratt, J.R., 1992: *The Atmospheric Boundary Layer*, Cambridge University Press, Cambridge, UK. 316 pp.

Keyser, D. and M.A. Shapiro, 1986: A review of the structure and dynamics of upper-level frontal zones. *Mon. Wea. Rev.*, **114**, 452-499.

Knox, J.A., D.W. McCann, and P.D. Williams, 2008: Application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting. *J. Atmos. Sci.*, **65**, 3392-3304.

Knox, J.A., D.W. McCann, and P.D. Williams, 2009: Reply to Plougonven et al. (2009), *J. Atmos. Sci.*, **66**, 2511-2516.

Koch, S.E., B.D. Jamison, C. Lu, T.L. Smith, E.I. Tollerude, C. Girz, N. Wang, T.P. Lane, M.A. Shapiro, D.D. Parish, and O.R. Cooper, 2005: Turbulence and gravity waves within an upper-level front. *J. Atmos. Sci.*, **62**, 3885-3908.

Lennartson, D. W. and D.W. McCann, 2014: Integrated turbulence forecasts*, Fourth Aviation, Range, and Aerospace Meteorology Special Symposium*, https://ams.confex.com/ams/94Annual/webprogram/Paper230413.html

Lighthill, M.J., 1952: On sound generated aerodynamically, I. General theory. *Proc. Roy. Soc. London*, **211A**, 564-587.

Lilly, D.K., 1978: A severe downslope windstorm and aircraft turbulence event induced by a mountain wave. *J. Atmos. Sci.*, **35**, 59-77.

Lilly, D.K. and P.J. Kennedy, 1973: Observations of a stationary mountain wave pattern and its associated momentum flux and energy dissipation. *J. Atmos. Sci.*, **30**, 1135-1152.

McCann, D.W., 1999a: A simple turbulence kinetic energy equation and aircraft boundary layer turbulence. *Natl. Wea. Digest*, **23**, 13-19.

McCann, D.W. 1999b: VVSTORM – Convection diagnosed from numerical models. *Proc. Eighth Conf. on Aviation, Range, and Aerospace Meteorology*, Dallas TX, Amer. Meteor. Soc., 240-243.

McCann, D.W., 2001: Gravity waves, unbalanced flow, and aircraft clear air turbulence. *Natl. Wea. Digest*, **25**, 3-14.

McCann, D.W., 2006: Diagnosing and forecasting aircraft turbulence with steepening mountain waves. *Natl. Wea. Digest*, **30**, 77-92.

McCann, D.W., J.A. Knox, and P. D. Williams, 2012: An improvement in clear-air turbulence forecasting based on spontaneous imbalance theory: the ULTURB algorithm. *Meteorol. Appl.,* **19**, 71–78.

McIntyre, M.E., 2009: Spontaneous imbalance and hybrid vortex-gravity structures. *J. Atmos. Sci.*, **66**, 1315-1326.

Miles, J.W. and L.N. Howard, 1964: Note on a heterogeneous flow. *J. Fluid Mech*, **20**, 331-336.

Palmer, T.N., G.J. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave parameterization. *Quart. J. Roy. Met. Soc.*, **112**, 1001-1039.

Plougonven, R., C. Snyder, and F. Zhang, 2009: Comments on "Application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting." *J. Atmos. Sci.*, **66**, 2506-2510.

Plougonven, R. and F. Zhang, 2007: On the forcing of inertia-gravity waves by synoptic flows. *J. Atmos. Sci.*, **64**, 1737-1742.

Plougonven, R. and F. Zhang, 2016: Gravity waves generated by jets and fronts and their relevance for clear-air turbulence. *Aviation Turbulence*, Springer International Publishing, Switzerland, ISBN 978-3-319-23629-2, Chapter 19, 385-406.

Ralph, F.M., C. Mazaudier, M. Crochet, and S. V. Venkateswaran, 1993: Doppler sodar and radar wind profiler observations of gravity wave activity associated with a gravity current. *Mon. Wea. Rev.*, **121**, 444–463.

Reed, R.J. and K.R. Hardy, 1972: A case study of persistent, intense clear air turbulence in an upper level frontal zone. *J. Appl. Meteor.*, **11**, 541-549.

Roach, W.T., 1970: On the influence of synoptic development on the production of high level turbulence. *Quart. J. Roy. Met. Soc.,* **96**, 413-429.

Shapiro, M.A., 1984: Meteorological tower measurements of a surface cold front. *Mon. Wea. Rev.*, **112**, 1634-1639.

Sharman, R., 2016: Nature of aviation turbulence. *Aviation Turbulence*, Springer International Publishing, Switzerland, ISBN 978-3-319-23629-2, Chapter 1, 3-30.

Sharman, R.D., 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.

Sharman, R. and T. Lane (editors), 2016: *Aviation Turbulence*, Springer International Publishing, Switzerland, ISBN 978-3-319-23629-2, 523 pp.

Sharman, R., 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.

Sharman, R., C. Tebaldi, G. Weiner, J. Wolff, 2006: An integrated approach to mid- and upper-level turbulence forecasting. *Wea. Forecasting*, **21**, 268-287.

Snyder, C., W.C. Shamarock, and R. Rotunno, 1993: Frontal dynamics near and following frontal collapse. *J. Atmos. Sci.*, **50**, 3194-3211.

Thorpe, S.A., 1969: Experiments on the stability of stratified shear flows. *Radio Sci.* **4**, 1327-1331.

Uccellini, L.W., P.J. Kocin, R.A. Petersen, C.H. Wash, and K.F. Brill, 1984: The President’s Day cyclone of 18-19 February 1979: Synoptic overview and analysis of the subtropical jet streak influencing the pre-cyclogenetic period. *Mon. Wea. Rev.*, **112**, 31-55.

Uccellini, L.W. and S.E. Koch, 1987: The synoptic setting and possible energy sources for mesoscale wave disturbances. *Mon. Wea. Rev.*, **115**, 721-729.

Viúdez, Á., 2007: The origin of the stationary frontal wave packet spontaneously generated in rotating stratified vortex dipoles. *J. Fluid Mech***.**, **593**, 359–383.

Williams, J.K., G. E. Blackburn, J. A. Craig, R. K. Goodrich, J. Johnson, F. McDonough, G. Meymaris, J. M. Pearson, and R. D. Sharman, 2013: Identifying upper-level wake vortex encounters using routine turbulence reports. *Proc. 16th Conf. on Aviation, Range, and Aerospace Meteorology*, Austin TX, Amer. Meteor. Soc., Available at https://ams.confex.com/ams/93Annual/webprogram/Paper219378.html.

Williams, P.D., T.W.N. Haine, and P.L. Read, 2005: On the generation mechanisms of short-scale unbalanced modes in rotating two-layer flows with vertical shear. *J. Fluid Mech*. **528**, 1–22.

Williams, P.D., T.W.N. Haine, and P.L. Read, 2008: Inertia-gravity waves emitted from balanced flow: Observations, properties, and consequences. *J. Atmos. Sci.*, **65**, 3543-3556.

Williams. P.D. and M.M. Joshi, 2016: Clear-air turbulence in a changing climate. *Aviation Turbulence*, Springer International Publishing, Switzerland, ISBN 978-3-319-23629-2, Chapter 23, 465-480.

Wilson, E.N., 2012: *Case studies of clear-air turbulence: evaluation and verification of new forecasting techniques*. M.S. thesis, University of Georgia, 127 pp.

Yasuda, Y., K. Sato, and N. Sugimoto, 2015a: A theoretical study on the spontaneous radiation of inertia–gravity waves using the renormalization group method. Part I: Derivation of the renormalization group equations. *J. Atmos. Sci.*, **72**, 957-983.

Yasuda, Y., K. Sato, and N. Sugimoto, 2015b: A theoretical study on the spontaneous radiation of inertia–gravity waves using the renormalization group method. Part II: Verification of the theoretical equations by numerical simulation. *J. Atmos. Sci.*, **72**, 984-1009.

Zack, J.W. and M.L. Kaplan, 1987: Numerical simulations of the subsynoptic features associated with the AVE-SESAME case. Part 1: The preconvective environment. *Mon. Wea. Rev.*, **115**, 2367-2394.