

Basic Diagnosis and Prediction of Persistent Contrail Occurrence using High-resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part II: Evaluation of Sample Models

David P. Duda

National Institute of Aerospace, Hampton, Virginia

Patrick Minnis

Science Directorate, NASA Langley Research Center, Hampton, Virginia

Corresponding author address: David P. Duda, NASA Langley Research Center, Mail Stop 420, Hampton, VA 23681-2199.

E-mail: david.p.duda@nasa.gov, dduda@nianet.org

ABSTRACT

Previous studies have shown that probabilistic forecasting may be a useful method for predicting persistent contrail formation. A probabilistic forecast to accurately predict contrail formation over the contiguous United States (CONUS) is created by using meteorological data based on hourly meteorological analyses from the Advanced Regional Prediction System (ARPS) and from the Rapid Update Cycle (RUC) as well as GOES water vapor channel measurements, combined with surface and satellite observations of contrails. Two groups of logistic models were created. The first group of models (SURFACE models) is based on surface-based contrail observations supplemented with satellite observations of contrail occurrence. The second group of models (OUTBREAK models) is derived from a selected subgroup of satellite-based observations of widespread persistent contrails. The mean accuracies for both the SURFACE and OUTBREAK models typically exceeded 75 percent when based on the RUC or ARPS analysis data, but decreased when the logistic models were derived from ARPS forecast data.

1. Introduction

Current numerical weather analysis (NWA) systems are able to provide hourly meteorological data on horizontal scales as small as 10 km. In principle, these high-resolution NWAs including the 20-km Rapid Update Cycle (RUC; see Benjamin et al., 2004a,b) and the University of Oklahoma Center for Analysis and Prediction of Storms (CAPS) Advanced Regional Prediction System (ARPS; see Xue et al., 2003) can provide the meteorological information necessary to diagnose contrail formation. Unfortunately, the straightforward prediction of contrail-induced cloud cover from these analyses is hindered by systematic and random measurement errors. Duda and Minnis (2008), hereafter called Part I, show that logistic regression modeling can provide a method to deal with these errors and to diagnose contrail occurrence accurately based on NWA-derived atmospheric variables including temperature, relative humidity and vertical velocity.

Some probabilistic forecast models of contrail occurrence based on logistic regression have already been developed. Travis et al. (1997) used a combination of rawinsonde temperature and geostationary satellite water vapor absorption data to develop a logistic model of the occurrence of widespread persistent contrail coverage. Jackson et al. (2001) created a contrail prediction model using surface observations of contrails and rawinsonde measurements of temperature, humidity and winds. In this study, we use contrail observations from both the GLOBE (Global Learning and Observations to Benefit the Environment) program and geosynchronous satellite imagery, along with numerical weather analyses and forecasts to create forecast models for the prediction of persistent contrail formation. These models allow predictions of

widespread contrail occurrences on either a real-time basis or for long-term time scales. The real-time forecasts could be used in aviation for the prevention of persistent contrail production, while long-term studies could focus on estimating the radiative impact of contrails on regional or global climate.

Despite the success of the probabilistic forecast models described in Travis et al. (1997) and Jackson et al. (2001), several questions remain about the usefulness of these models. The former study used only a limited number of observations, while the latter only considered contrail observations within limited geographic (New England states) and temporal (two weeks in September) domains. Neither study attempted to use numerical weather forecast data to predict contrail occurrence. The use of prognostic meteorological data within the logistic models would allow for longer forecast lead-times than logistic models developed from observations only. Such longer lead-times would be helpful if contrail mitigation efforts are considered.

In this paper, we assess the ability of logistic models to provide a valuable and accurate diagnosis/prediction of persistent contrail occurrence via numerical weather models. Specifically, we evaluate a sample of logistic contrail forecasts based on RUC and ARPS data and observations of contrail occurrence. The value of the contrail prediction models is then discussed in the context of a forecast evaluation theory.

The next section describes the meteorological data and contrail occurrence observations used to develop the statistical contrail occurrence models, and section 3 presents and evaluates some examples of logistic models. The final two sections briefly summarize and discuss the overall value of the logistic forecasts.

2. Data and methodology

a. Meteorological Data

To provide atmospheric predictors for the logistic models, we use nearly 15 months (April 2004 – 27 June 2005) of meteorological data from two high-resolution, operational numerical weather analyses. Profiles of temperature, humidity, horizontal wind speed and direction, and vertical velocity were derived using hourly analyses from the 20-km resolution RUC model and from the 27-km resolution ARPS analyses in 25-hPa intervals from 400 hPa to 150 hPa. (After 12 UTC on 28 June 2005, the 13-km resolution version of the RUC model became operational, with significant differences in upper tropospheric humidity.) Due to limitations in computational resources, both the RUC and ARPS data were stored at approximately $1^\circ \times 1^\circ$ horizontal resolution. In addition to the RUC and ARPS analyses, ARPS 1-day, 2-day and 3-day forecasts were also used to build logistic models.

The meteorological data were downloaded each day to a local computer. The data are subject to interruptions including computer and power failures, full disks, operator errors, lack of data availability and other problems. Thus, approximately 77% of the hourly ARPS and 99.7% of the RUC data were collected during the time period. Two large gaps (between 20 August – 28 September 2004, and between 21 January – 21 February 2005) accounted for nearly 85% of the ARPS data loss. The ARPS forecasts had a slightly larger loss rate than the ARPS analyses as sometimes the forecasts were not available even though the analyses were available.

b. Satellite data

To supplement the meteorological data from the numerical weather analyses, radiance data from the 6.5- μm water vapor absorption channel on the 12th Geostationary Environmental Operational Satellite (*GOES-12*) were also used as atmospheric predictors of contrail formation. The 6.5- μm channel is sensitive to the top three millimeters of water vapor profile in the atmospheric column, and most of the detected emission is from the layer between 500 to 200 hPa, with a peak sensitivity near 400 hPa (Travis et al., 1997; Schmit et al., 2001). The water vapor channel on *GOES-12* is spectrally widened compared to the corresponding channel on *GOES-8*, and thus measures radiation slightly deeper in the atmosphere than previous sensors. The average 6.5- μm brightness temperatures on *GOES-12* are typically 2 to 3 K warmer than their counterparts on *GOES-8* or *GOES-10* (Schmit et al., 2001). The raw water vapor image counts and the calibrated 6.5- μm brightness temperatures and radiances were collected. Multi-spectral data from the GOES and the NOAA Advanced Very High Resolution Radiometer (AVHRR) imagers were also used to detect persistent contrail occurrence for some of the logistic models.

c. Surface data

Persistent contrail occurrence was also determined from a set of surface observations. The Global Learning and Observations to Benefit the Environment (GLOBE) program collects observations of contrail occurrence from primary and secondary schools across the contiguous United States (CONUS). (See www.globe.gov for more information about the GLOBE program.) In May 2003, GLOBE initiated a contrail observation protocol to gather and classify contrail observations. A primary goal

of the GLOBE program is to use detailed written protocols to enable students to provide scientifically valuable measurements of environmental parameters (Brooks and Mims, 2001). Over 18,500 observations of cloud coverage and contrail occurrence were reported over the CONUS between April 2004 and June 2005. The contrail observations are classified into three categories, short-lived (SHRT), non-spreading persistent contrails (NSPR), and spreading persistent contrails (SPRD). The contrail categories are defined as follows: short-lived contrails are contrails that dissipate as the aircraft moves across the sky. Persistent contrails are contrails that remain in the sky after the aircraft has flown out of view of the observer. Spreading contrails are defined as contrails wider than the width of a finger held at arm's length. This width corresponds to a contrail at least 350 m wide, based on a contrail altitude of 10 km (O'Shea, 1991), which is the minimum width expected to be detectable in high-resolution satellite imagery.

The GLOBE contrail dataset contains observations from 417 schools. The schools are mostly located in highly populated regions with substantial air traffic (Duda et al., 2008). Nearly all schools reported only one observation/day, but only 123 of the schools reported more than 30 observations during the 15-month period. Approximately 92% of all observations were between 1430 and 2030 UTC, and nearly 58% of the total were between 1630 and 1830 UTC.

d. Data processing

Before deriving the logistic models, the meteorological data were checked for missing data, and matched in time and location with the surface and satellite observations of contrails. No contrail observations with missing meteorological data were used in the statistical forecast models.

To match the RUC data with the contrail occurrence observations, meteorological variables from the RUC analyses closest in time with the contrail observations are linearly interpolated to the location of each contrail observation. An observation is not used if the time difference between the observation and the RUC analysis was greater than 2 hours (nearly all pairs were matched to within 1 hour). A similar procedure is used to match the ARPS analysis data with the contrail observations. For the ARPS forecast data, the meteorological data from the forecast time matching to within 1 hour of the observation were used. Because the ARPS forecasts begin at 00 UTC and all of the contrail observations used in this study occurred between 16 and 20 UTC, the 1-day forecasts refer to the 16 to 20 hour forecast model time, the 2-day forecasts refer to the 40 to 44 hour forecast model time, and the 3-day forecasts refer to the 64 to 68 hour forecast time.

For convenience, atmospheric humidity in both meteorological datasets was usually expressed in the form of the maximum relative humidity with respect to ice (RHI) between 150 and 400 hPa. For the ARPS data, the RHI was computed from the ARPS fields of potential temperature and specific humidity at the 25-hPa intervals to determine the level of maximum upper tropospheric humidity. Because it is expected that persistent contrails are most likely to form where relative humidity is greatest, for each contrail observation, the pressure level between 400 hPa and 150 hPa with the maximum RHI and having a temperature less than or equal to -40°C was identified. The temperature constraint was added to eliminate areas where the atmosphere is likely to be too warm to form contrails (Appleman, 1953).

e. Statistical technique

Logistic regression (Hosmer and Lemeshow, 1989) was used to create a probabilistic estimate of persistent contrail formation based on the meteorological variables from the RUC and ARPS models. The logistic model assumes the following fit:

$$P \approx \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)]}. \quad (1)$$

where P is the predictand (probability of persistent contrail formation) and β_i (for $i = 1, \dots, p$) are the set of coefficients used to fit the predictors (x_i) to the model. All predictors used in this study are based on meteorological quantities in the upper troposphere that are expected to be physically related to the formation of spreading, persistent contrails.

The maximum likelihood method was used to estimate the unknown coefficients β_i and to fit the logistic regression model to the data. The chi-square statistic (χ^2) was employed to assess the goodness of fit of each logistic model to the meteorological data. A stepwise regression technique was used to reduce the number of predictors to an optimal number of variables. In each step of the technique, a new predictor is added to the logistic model and the chi-square statistic is compared with the previous model. The new predictor that produces the largest improvement in model fit (that is, the largest increase in χ^2) is added to the model. To avoid overfitting the model, the stepwise regression technique is allowed to add predictors to the model until the test for statistical significance reaches a significance level (*i.e.*, p -value) is approximately 0.05.

f. Predictors

Table 1 includes all potential predictors considered for this study. A total of 82 potential predictors from the numerical weather models and satellite water vapor channel measurements were used to develop the logistic regression models. All of the variables are expected to influence the formation (or the spreading rate) of persistent contrails. From this set of potential predictors the stepwise regression method was used to reduce the number of predictors to approximately six. Several of the variables including the temperature, vertical velocity, wind speed and direction were computed at the level of maximum RHI, and the vertical shear of the horizontal wind and temperature lapse rate (a measure of atmospheric stability) are computed for the 25-hPa layer below the level of maximum RHI. The logistic models developed from the ARPS analyses and forecasts do not include precipitable water, tropopause temperature, tropopause pressure, or any other variable formed by the combination of those parameters because they are not included in the ARPS analyses.

Several other meteorological variables were also considered as possible predictors in the logistic models. In addition to determining the variables at the level of maximum RHI, the 200 – 300 hPa layer averages of several variables were computed, as well as regional mean variables, which are the mean (of the 200 – 300 hPa means) of all model grid points within 200 km of the contrail observation location. Most commercial air traffic over the CONUS cruises between 200 and 300 hPa (Garber et al., 2005). The regional mean was developed to account for some of the uncertainty in the meteorological fields forecasted by the ARPS model. Finally, “upstream” means of temperature and RHI were also computed. The upstream mean is defined as the 200 –

300 hPa layer mean average of the variable located 2 hours upstream from the contrail observation location. The upstream point is determined by computing a 2-hour backward trajectory using the 200 – 300 hPa mean wind from the original observation point. The upstream variables were included because most persistent contrails actually form 1 to 2 hours before they become visible within GOES infrared imagery (Duda et al., 2004).

Several variables based on the satellite water vapor channel radiance measurements were also regarded as potential predictors. The mean and standard deviation of the brightness temperature, the raw count, and the calibrated radiances were considered, as well as several multiplicative combinations of the three water vapor channel-based variables. Both the mean and the standard deviation of the variables were based on the 5×5 GOES pixel array centered on the observation location.

g. Equation development

Two groups of logistic models were created using actual meteorological data and satellite radiance measurements, and observations of contrail occurrence. The first group of models is based on surface-based contrail observations supplemented with satellite observations of contrail occurrence. These models were designed to relate the general occurrence of persistent contrails with the meteorological conditions, and are hereafter referred to as the SURFACE models. The second group of models is similar to the work of Travis et al. (1997) where a selected subgroup of observations within the presence of widespread persistent contrails (called both here and in Travis et al. (1997) as “outbreaks”) is used to build the logistic models. They are called the OUTBREAK models in this study. In addition to meteorological data from numerical weather analyses

and forecasts, satellite radiance data will also be used as potential predictors for both groups of logistic models.

The contrail observations were separated into a dependent (from which the statistical models were created) and an independent (on which the models were tested) dataset. Two-thirds of the data were randomly selected to build the dependent dataset, while the independent dataset comprises the remaining one-third of the data.

To determine the accuracy of the contrail models, two statistical measures used in Part I were employed. The contrail formation forecasts are separated into four categories based on the forecast and its outcome: a is the number of cases where persistent contrail formation is forecasted, and persistent contrails are observed (hits); b is the number of cases where contrails are predicted, but no contrails are observed (false alarms); c is the number of cases where contrails are not forecasted, but contrails are observed (misses), and d is the number of cases where contrails are not forecasted and no contrails are observed (correct rejections). The first measure is the percent correct (PC), and equals $(a + d)/(a + b + c + d)$. PC is defined as the ratio of the correct forecasts to the total number of forecasts. The second variable is known as the Hanssen-Kuipers discriminant or the true skill statistic (HKD) (Wilks, 1995). The HKD is calculated as $(ad - bc) / [(a + c)(b + d)]$. This measure of forecasting skill measures the skill of the “yes” and “no” forecasts of contrail occurrence equally, regardless of the relative numbers of each forecast. Gandin and Murphy (1992) show that HKD is the only equitable skill score for a two-event (*i.e.*, yes-or-no) forecast.

3. Logistic models based on numerical weather analyses

a. SURFACE models

A subset of 11 GLOBE reporting locations with at least 50 contrail observations under mostly clear skies (non-contrail cloud coverage less than 25%) were chosen for building the SURFACE logistic regression models. Because these schools provided multiple observations throughout the 15-month period, we expect that these locations would be more likely to provide high quality contrail observations among the GLOBE participants. Table 2 lists the location of each GLOBE school, and the locations of the GLOBE schools are also shown in Figure 2. All schools with the exception of Box Elder, MT are located in regions with substantial commercial air traffic during the observation period (Duda et al., 2008). From this group of locations, a set of 379 observations were selected that could be “verified” by visual inspection of time series of GOES imagery for contrail occurrence/non-occurrence. This verification is somewhat subjective as the surface observer under mostly clear skies can detect much narrower and thinner persistent contrails than the 4-km resolution satellite imagery. Of the 379 observations, the surface and satellite results matched nearly 75 percent of the time. In about 18 percent of the observations, the surface observer reported persistent contrails while none were apparent in the satellite imagery, and for the remaining seven percent of the observations contrail occurrence was detected by satellite but not reported by the GLOBE observer.

Two sets of probabilistic models were developed from the GLOBE surface observations and the numerical weather analysis data. The first set (called here Build 1) used the GLOBE contrail occurrence observations, while the second set (called here Build 2) used the GOES satellite observations of contrail occurrence. As discussed above,

the observations were randomly separated into dependent and independent datasets to create and to test the model, respectively. For convenience, the stepwise regression was stopped after 6 predictors were chosen, because additional predictors rarely improved the skill scores of the logistic models significantly. Due to the large number of potential predictors (some closely related to each other), many combinations of predictors produced chi-square statistics nearly equal to the best fitting model. Therefore, the logistic models were evaluated by averaging the skill scores from the five models with the highest chi-square statistic, thus producing a mean skill score from each group of meteorological data.

Table 3 presents the mean PC and HKD skill scores for both builds of the logistic model for each group of meteorological data. For simplicity, the critical threshold for determining contrail occurrence was 0.5 for all cases. The logistic models from the first four rows of Table 3 are built from analysis data (either with or without supplemental GOES water vapor channel data), and therefore diagnose contrail occurrence from the analysis data. The remaining models are true forecasts evaluated using forecast data, but are developed from either analysis or forecast data. In every case except one (the HKD score for the ARPS 3-day forecast), the skill scores of Build 2 were higher than the skill scores of Build 1. Also, the differences between the Build 2 and Build 1 scores were largest when analysis data were used, and smallest when 3-day forecast data were used. The skill scores also tended to improve when GOES water vapor channel data were used as potential predictors in the logistic models. The accuracy of the models generally decreased as the length of the forecast increased. When the independent datasets (the remaining third of the observations not used in the development of the models) were used

to evaluate the skill of the forecast models, the PC ranged from 0.69 to 0.86 for the Build 2 models, and the HKD varied from about 0.18 to 0.59. These scores are worse than the results from Jackson et al. (2001). They developed a regional contrail formation model (including non-persistent contrails) based on a network of surface observers across New England coordinated with air traffic control information such that the observers knew exactly when to expect flights. The observations were collected during a two-week period in September. Jackson et al. used non-synoptic radiosonde launches to gather humidity information, and report a PC around 0.85 and an HKD near 0.66.

The most common predictors for the Build 1 SURFACE models tend to be related to temperature when the models are derived from RUC/ARPS analysis data. No specific kind of variable is favored when the Build 1 models are derived from ARPS forecast data. The most common predictors for the Build 2 SURFACE models tend to be related to temperature, relative humidity and wind direction when the models are generated using RUC or ARPS analyses, and to vertical velocity and the product of temperature and relative humidity with respect to ice when the models are developed from ARPS forecasts.

b. OUTBREAK models

The logistic models created using GOES observations of contrail outbreaks are similar to the model created by Travis et al. (1997), who derived the meteorological data from a select set of atmospheric conditions. The water vapor channel data from non-contrail locations were taken from either completely clear or completely cloudy pixels close to the contrail outbreak regions, while contrail observations were taken at locations where contrails were wide enough to fill the entire satellite pixel. In this study, the

contrail observation locations were chosen from only two sets of conditions, either from clear skies near the contrail outbreak or from pixels within the contrail outbreak. This method allows for a sharp distinction in the meteorological and water vapor channel data between the contrail and non-contrail areas. As a result, logistic models of high accuracy can be produced.

Visual inspection of AVHRR brightness temperature difference (BTD) imagery between the 10.8 and 12.0 μm channels, and of loops of GOES infrared and water vapor imagery was used to identify approximately 50 examples of contrail outbreaks – areas of distinct, line-shaped contrails covering at least 100,000 km^2 at various locations around the CONUS between August 2004 and June 2005. (See Figure 1 for an example of a typical contrail outbreak.) Because the horizontal resolution of the GOES infrared imagery is 4 km, the contrail outbreaks are composed of extremely wide, well-developed spreading persistent contrails. Figure 2 shows the locations of the contrail outbreaks.

A total of 104 satellite measurements in and around large contrail outbreaks were used to make the independent and dependent datasets. The stepwise regression was stopped after 4 predictors were chosen, and the skill scores from the five models with the highest chi-square statistic were averaged to produce a mean skill score for each group of meteorological data. The mean PC and HKD skill scores for the top five models produced from each group of meteorological data are presented in Table 7. (Once again, the critical threshold for determining contrail occurrence was 0.5 for all cases.) The skill scores in Table 7 are similar to the results from Travis et al. (1997) who used raw GOES water vapor channel count data to obtain upper tropospheric humidity information and

radiosondes to get temperature information. They reported PCs near 0.90 and an HKD around 0.85.

Similar to the skill scores for the SURFACE models, the skill scores of the logistic models developed from the OUTBREAK data tended to improve when GOES water vapor channel data were used as potential predictors in the logistic models. The accuracy of the OUTBREAK models also generally decreased as the length of the forecast increased. The differences between the skill scores of the logistic models created from the ARPS analysis with the scores from the models developed from the ARPS forecasts are larger in Table 4 than in Table 3. The OUTBREAK models are developed from a smaller set of observations than the SURFACE models, and because the OUTBREAK models are tuned to sharply defined meteorological conditions, they are more sensitive to the errors in the meteorological fields that are present in forecast models. Thus, this group of logistic models highlights the effects of forecast errors in the logistic model results. The most common predictors for the OUTBREAK models tend to be wind direction, atmospheric lapse rate (dT/dz), temperature, RHI, and the product of temperature and RHI.

4. Discussion

A comparison of results between all of the models presented here gives some insight into the overall quality of the logistic models, where they perform well, and where further improvement is necessary. As shown in Section 3, the Build 2 SURFACE models are consistently better than the Build 1 models. While the difference in model performance could be easily explained by postulating the superior quality of satellite-based contrail observations compared to the observations of primary and secondary

school students with little training, it is important to note some differences between surface-based and satellite-based observations. Surface observers often miss contrails forming above lower cloudiness (although by choosing only mostly clear observations, this type of error should be minimal here), or record the observation incorrectly in another category (some GLOBE observations are suspected to suffer from such a clerical error). Both the manual and the automated detection of persistent contrails in satellite imagery is also hampered by cloud cover and the misidentification of cloud streets as contrails, or contrails as cloud streets (Mannstein et al., 1999). Surface observers, however, can detect much narrower and probably optically thinner contrails than those seen in the 4-km resolution satellite imagery of this study. If the students are detecting relatively thin, but persistent contrails within thin layers of supersaturation in the upper troposphere, then a weaker correlation between the numerical weather model variables and the occurrence of persistent contrails would be expected.

The skill scores of the logistic models using RUC/ARPS analyses show improvement when water vapor data are added. Many of the logistic models built with water vapor data selected the standard deviation of the water vapor brightness temperature rather than the brightness temperature itself as the best predictor of contrail occurrence. Since contrails tend to reduce the brightness temperature compared to the contrail-free surroundings, the standard deviation of the water vapor brightness temperatures should increase in the presence of contrails. Thus, the improvements in the logistic models may result from the sharp distinction between contrail areas and non-contrail areas in the water vapor images. This is akin to using the contrail images themselves to predict contrails. Thus, the use of the water vapor imagery in a logistic

contrail prediction model is only applicable in a diagnosis mode, or perhaps for short-term (3 to 6 h) forecasts.

The accuracy of the SURFACE and OUTBREAK models is less than the accuracy of the test case models in Part I created from synthetic observations. Part of the reason for the lower accuracy is that factors other than temperature, relative humidity and vertical velocity affect the development of spreading persistent contrails. The results from Part I show that the addition of vertical velocity to the determination of contrail formation resulted in slightly less accurate models, even when all factors were known and accounted for in the logistic model. The most common predictors chosen in the SURFACE and OUTBREAK models tended to be related to temperature and humidity, but other variables including vertical velocity, wind direction and speed, and atmospheric lapse rate were frequently chosen as predictors. Previous studies of contrail occurrence suggest that high contrail incidence is associated with areas of baroclinity, and thus with areas where wind speed, vertical velocity and lapse rate may have significant departures from mean conditions (DeGrand et al., 2000). The results from Carleton et al. (2008) suggest that atmospheric variables lower in the atmosphere that were not included in this study may also be valuable predictors. The list of meteorological variables in Table 4 is not exhaustive, and other combinations of variables not presented here may be better predictors of contrail occurrence.

The differences between the accuracy of the dependent and independent results in Tables 3 and 4 indicate whether an adequate number of observations have been used to build the logistic models. The dependent results are often better than the independent results suggesting that the logistic models sometimes are too finely tuned to the

dependent sets and that more observations are needed to build the logistic models. The results from the OUTBREAK models, which were developed using only a subsample of possible atmospheric conditions highlight this problem. The results for the OUTBREAK models show differences between the dependent and independent results that are 3 to 4 times larger than those for the SURFACE models. The OUTBREAK models are designed to have high accuracy when using the dependent data. However, the models do much worse when evaluated with the independent data, which are not used in the construction of the logistic models. The large differences between the dependent data and the independent data skill scores are especially apparent in the OUTBREAK models developed from ARPS forecast data.

Tables 3 and 4 present the skill scores of several prognostic models developed from both ARPS analysis data and from ARPS forecast data. One advantage of developing logistic models for forecasts using the analysis data is that they have the most accurate meteorological data, and should allow for more accurate short-term forecasts than models built with forecast data. An examination of Tables 3 and 4 confirms this point. When evaluating the accuracy of the SURFACE and OUTBREAKS models with the independent data, the models developed from the analysis data tend to do better than the forecast data-developed models for the 1-day forecasts, and to do worse for the 3-day forecasts. The models developed from the analysis data are always less accurate, however, when evaluating the accuracy of the models with the dependent data. This is expected because the forecast data-developed models are developed directly from the dependent data, and should have the best fit. The results support the viewpoint that logistic models developed from analysis data are at least comparable in accuracy to the

models developed from forecast data. Any gains from the forecast data-developed models are lost when applied to independent data. The meteorological data in the forecasts have so much error that the logistic models respond to that error and tend to choose predictors that fortuitously correlate with contrail occurrence within that particular dependent dataset.

Similarly, the relatively lower accuracy of the forecast models compared to the diagnostic models built from analysis data is mostly the result of larger temperature and humidity errors in the forecasts. This is evident because the accuracy of the models tended to decrease as the length of the forecast increased. The atmospheric conditions for the formation of persistent contrails in the absence of natural cirrus tend to occur at the edges of areas of high humidity and cooler temperatures in the upper troposphere, and the exact location and timing of these regions are not always represented well in numerical weather models. Most of the variables chosen as possible predictors in the logistic models are based on meteorological quantities at the point of interest. Even if the general synoptic features of the forecast are accurate, relatively small errors in the motion or size of high humidity areas could reduce the accuracy of the contrail prediction models substantially. Model errors may be mitigated if more regionally or temporally averaged variables were used in the creation of the logistic models.

This study attempted to build a universal contrail model suitable for all times across the CONUS, whereas both seasonal and regional differences in contrail occurrence are common (DeGrand et al., 2000). It appears that a universal model for the entire CONUS may not allow for the highest accuracy, and as in probabilistic precipitation forecasts, local forecasts for a specific location or region, or a specific season may allow

for more accurate models. The superior results from the spatially and temporally limited study of Jackson et al. (2001) support this idea to some extent, but they rely on having enhanced direct observations of the meteorological fields, not degraded NWA fields.

An important quality of the contrail forecast model is its overall value. The value of a forecast is defined here following Murphy and Ehrendorfer (1987), such that forecasts are of positive value only if they can lead to different actions than those that the decision maker would have taken in the absence of the forecasts. If persistent contrail forecasts are to be of any value, for example, in the case of diverting flights in order to reduce persistent contrail cloud cover, then the cost (C) of diverting the flights must be weighed against the losses (L) that may result as a consequence of the additional cloud cover produced by the contrails. In the absence of a contrail forecast, the decision maker would have to compare this cost/loss ratio (C/L) to the climatological occurrence of persistent contrails (p_c). If $C/L > p_c$, then no flights would ever be diverted, while all flights would be diverted if $C/L < p_c$. In general, it is expected that any valuable forecast must be accurate enough such that the percentage of forecast misses (p_m , defined here as $c/(c + d)$) must be less than or equal to the climatological frequency (and the cost/loss ratio), and the percentage of forecast hits (p_h , defined here as $a/(a + b)$) must be greater than or equal to p_c (and C/L). Murphy and Ehrendorfer (1987) show that if $0 \leq C/L \leq p_m \leq p_c \leq p_h \leq 1$ or $0 \leq p_m \leq p_c \leq p_h \leq C/L \leq 1$, then the problem of diverting or not diverting flights becomes trivial. In the former case, the cost of diverting flights is so inexpensive that all flights should be diverted to avoid making contrails, while in the latter case the cost of diverting flights is so expensive that no flights should be diverted despite the loss incurred from the production of contrails. The potentially wide range between p_m and p_h

in the test case models from Part I suggests that logistic models would be able to produce valuable persistent contrail occurrence forecasts for a variety of cost/loss situations. The results from Part I derived from synthetic observations show that $p_m = 0.028$ when the climatological frequency is used as the probability threshold, and $p_h = 0.506$ even in scenario B1m where the random error is maximized. If 0.5 is used as the probability threshold, then $p_m = 0.077$ and $p_h = 0.726$ in scenario B2m. For comparison, the Build 2 SURFACE models built from (the dependent) and evaluated with (the independent) ARPS analysis data have a $p_m = 0.169$ and a $p_h = 0.573$. Considering that the p_c measured from surface observers was 0.170 in Duda et al. (2008), and 0.152 in Minnis et al. (1993), the models presented here have marginal value as p_m approximately equals p_c . Logistic models built from a larger number of observations, however, may have positive value because the p_m and p_h for the Build 2 SURFACE models built from and evaluated with the same ARPS analysis data (dependent data) are 0.094 and 0.810, respectively.

It is important to note that the conclusions of Murphy and Ehrendorfer (1987) apply to a simple two-parameter [occurrence versus non-occurrence] system. Much more complicated cost/loss relationships could be possible if the full capability of a probabilistic forecasting system were used. For example, given a forecast probability p in a forecast region, a fraction p_d of all flights within that region could be diverted. Also, reliable probabilistic forecasts inherently have extra value to users compared to categorical (simple yes or no occurrence) forecasts because users can take advantage of cost/loss analyses better with probabilistic forecasts (Keith, 2003).

5. Summary and concluding remarks

Probabilistic models of persistent contrail occurrence within the CONUS were developed from high-resolution numerical weather analyses and forecasts. Meteorological data from the 20-km Rapid Update Cycle (RUC) and the Advanced Regional Prediction System (ARPS), as well as GOES water vapor channel measurements were combined with observations of persistent contrail occurrence from surface reports and visual inspection of satellite imagery. Two groups of logistic models were created. The first group of models (SURFACE models) is based on surface-based contrail observations supplemented with satellite observations of contrail occurrence. The second group of models (OUTBREAK models) is derived from a selected subgroup of satellite-based observations of widespread persistent contrails. The mean accuracies for both the SURFACE and OUTBREAK models typically exceeded 75 percent when based on the RUC or ARPS analysis data, but decreased when the logistic models were derived from ARPS forecast data.

Some unanswered issues about the effectiveness of the logistic model are not addressed here, and require future study. Aircraft may not fly at all times through some regions where persistent contrails are possible, although this is not expected to be a major problem for this study as much of the CONUS is nearly continually traveled by jet aircraft throughout the day. Also, persistent contrails are unlikely in regions where adverse weather conditions (such as convection, turbulence, and icing) are expected to occur and aircraft are likely to avoid. These errors in the accurate determination of contrail occurrence should be quantified and their impact on the logistic model should be addressed.

More work is needed to realize the potential of logistic contrail forecasts. The most direct way to make the logistic models better is to reduce the errors within the meteorological data used to build the models. Reductions in the uncertainties of meteorological variables to a point where acceptable contrail forecasts are produced would be a good goal for NWA modelers. As mentioned earlier, meteorological errors directly affect the regressions developed in the logistic model, and if the errors are large enough, may cause the model to choose less pertinent predictors, further reducing model accuracy. Meteorological analyses could be improved by using the Atmospheric Infrared Sounder (AIRS) onboard the Aqua satellite to supplement the temperature and relative humidity data in numerical weather models. Methods to reduce errors in the determination of contrail occurrence could also be pursued. Additional studies are needed to determine if other regionally or temporally averaged variables would increase the accuracy of logistic models based on numerical weather forecasts, and if other atmospheric variables may be relevant. Regional and seasonal models of contrail occurrence may help improve the overall performance of this type of persistent contrail prediction model. Finally, logistic models of contrail occurrence provide an additional advantage that has not been used here. Because logistic models compute a probability of occurrence, they could be useful in global circulation model (GCM) simulations of contrail coverage (Ponater et al., 2002; Marquart et al., 2003) to determine the impact of contrail radiative forcing on global climate. Such models use a simple analytical formula based on relative humidity and cirrus cloud coverage to determine contrail coverage. The logistic models could be easily used within the GCM to determine an appropriate contrail coverage fraction for a region based upon the product of the air traffic and the computed

probability. Because the logistic model can be developed by comparing GCM model simulations to actual contrail observations, it may provide more accurate simulations of contrail coverage than current methods.

Acknowledgments.

This material is based upon work supported by the NASA Earth Science Enterprise Radiation Sciences Division, the NASA Modeling, Analysis, and Prediction Program, NASA contracts NAG1-02044 and NCCI-02043 NIA-2579, and by the National Science Foundation under Grant No. 0222623.

References

- Appleman, H., 1953: The formation of exhaust condensation trails by jet aircraft. *Bull. Amer. Meteorol. Soc.*, **34**, 14–20.
- Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004a: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494.
- Benjamin, S. G., D. Dévényi, S. S. Weygandt, K. J. Brundage, J. M. Brown, G. A. Grell, D. Kim, B. E. Schwartz, T. G. Smirnova, T. L. Smith, and G. S. Manikin, 2004b: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.
- Brooks, D. R., and F. M. Mims III, 2001: Development of an inexpensive handheld LED-based Sun photometer for the GLOBE program, *J. Geophys. Res.*, **106(D5)**, 4733–4740.
- Carleton, A. M., D. J. Travis, K. Master, and S. Vezhapparambu, 2008: Composite atmospheric environments of jet contrail outbreaks for the United States. *J. Appl. Meteorol. Climatol.*, **47**, 641–667.
- DeGrand, J. Q., A. M. Carleton, D. J. Travis, and P. J. Lamb, 2000: A satellite-based climatic description of jet aircraft contrails and associations with atmospheric conditions, 1977–79. *J. Appl. Meteorol.*, **39**, 1434–1159.
- Duda, D. P., P. Minnis, L. Nguyen, R. Palikonda, 2004: A case study of the development of contrail clusters over the Great Lakes. *J. Atmos. Sci.*, **61**, 1132–1146.
- Duda, D. P., and P. Minnis, 2008: Basic diagnosis and prediction of persistent contrail occurrence using high-resolution numerical weather analyses/forecasts and logistic regression. Part I: Effects of random error. *Submitted to J. Appl. Meteorol. Climatol.*.

- Duda, D. P., R. Palikonda, and P. Minnis, 2008: Relating satellite and surface observations of persistent contrail occurrence to numerical weather analyses and forecasts. *Submitted to Atmos. Chem. Phys.*
- Gandin, L. S., and A. H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon. Wea. Rev.*, **120**, 361–370.
- Garber, D. P., P. Minnis, and P. K. Costulis, 2005: A commercial flight track database for upper tropospheric aircraft emission studies over the USA and southern Canada. *Meteorol. Z.*, **14**, 445–452.
- Hosmer, D. W., and S. Lemeshow, 1989: *Applied Logistic Regression*. John Wiley & Sons, New York, 307 pp.
- Jackson, A., B. Newton, D. Hahn, A. Bussey, 2001: Statistical contrail forecasting. *J. Appl. Meteorol.*, **40**, 269–279.
- Keith, R., 2003: Optimization of value of aerodrome forecasts. *Wea. Forecasting*, **18**, 808–824.
- Mannstein, H., R. Meyer, P. Wendling, 1999: Operational detection of contrails from NOAA-AVHRR data. *Int. J. Remote Sensing*, **20**, 1641–1660.
- Marquart, S., M. Ponater, F. Mager, and R. Sausen, 2003: Future development of contrail cover, optical depth, and radiative forcing: Impacts of increasing air traffic and climate change. *J. Climate*, **16**, 2890–2904.
- Minnis, P., J. K. Ayers, M. L. Nordeen, and S. P. Weaver, 2003: Contrail frequency over the United States from surface observations. *J. Climate*, **16**, 3447–3462.
- Murphy, A. H., and M. Ehrendorfer, 1987: On the relationship between the accuracy and value of forecasts in the cost-loss ratio situation. *Wea. Forecasting*, **2**, 243–251.

- O'Shea, R. P., 1991: Thumb's rule tested: visual angle of thumb's width is about 2 deg, *Perception*, **20(3)**, 415–418, doi:10.1068/p200415.
- Ponater, M., S. Marquart, and R. Sausen, 2002: Contrails in a comprehensive global climate model: Parameterisation and radiative forcing results. *J. Geophys. Res.*, **107(D13)**, 4164, doi:10.1029/2001JD000429.
- Schmit, T. J., E. M. Prins, A. J. Schreiner, and J. J. Gurka, 2001: Introducing the GOES-M imager. *Natl. Wea. Assoc. Digest*, **25(3,4)**, 28–37.
- Travis, D. J., A. M. Carleton, S. A. Changnon, 1997: An empirical model to predict widespread occurrences of contrails. *J. Appl. Meteor.*, **36**, 1211–1220.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.
- Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation, *Meteor. Atmos. Physics*, **82**, 139–170.

List of Figures

FIG. 1. Persistent contrails are highlighted in a Channel 4 ($10.8\text{ }\mu\text{m}$) minus channel 5 ($12.0\text{ }\mu\text{m}$) brightness temperature difference image from the *NOAA-17* overpass at 1531 UTC on 18 April 2005 over the northeastern United States.

FIG. 2. Locations of persistent contrails identified from *GOES* imagery for OUTBREAK models (crosses), and GLOBE schools reporting persistent contrail coverage for SURFACE models (squares).

TABLE 1. Potential parameters used in logistic regression models.

Parameter	Name
Pressure at level of maximum RHI	prs
Gradient Richardson number at level of maximum RHI	grad_ri
Vertical wind shear at level of maximum RHI	shr
Mean vertical wind shear (200 – 300 hPa)	mnsr
Lapse rate in 25-hPa layer above level of maximum RHI	dtdz
Mean lapse rate (200 – 300 hPa)	mnstdz
North-south wind speed at level of maximum RHI	uwnd
East-west wind speed at level of maximum RHI	vwnd
Mean north-south wind speed (200 – 300 hPa)	mnuwnd
Mean east-west wind speed (200 – 300 hPa)	mnvwnd
Vertical velocity at level of maximum RHI	vv
Mean vertical velocity (200 – 300 hPa)	mnvv
Regional mean vertical velocity	regvv
Temperature at level of maximum RHI	tmp
Mean upstream temperature	upt
Mean temperature (200 – 300 hPa)	mnt
Regional mean temperature	regt
Maximum upper tropospheric RHI	rhi
Mean upstream RHI	upr
Mean RHI (200 – 300 hPa)	mnr
Regional mean RHI	regr
Tropopause pressure *	trp
Tropopause temperature *	trt
Precipitable water *	pwat
Raw water vapor count	raw
Standard deviation of raw water vapor count	raw_s
Water vapor channel radiance	rad
Standard deviation of water vapor channel radiance	rad_s
Water vapor channel brightness temperature	brt
Standard deviation of water vapor channel brightness temp.	brt_s
$uwnd \times uwnd$	uwnd2
$vwnd \times vwnd$	vwnd2
$mnuwnd \times mnuwnd$	mnuwnd2
$mnvwnd \times mnvwnd$	mnvwnd2
$uwnd \times vwnd$	uv
$mnuwnd \times mnvwnd$	mnuv
wind speed (based on uwnd and vwnd)	windspd
mean wind speed (based on mnuwnd and mnvwnd)	mnwindspd
wind direction	winddir
mean wind direction	mnwinddir
$vv \times vv$	vv2

TABLE 1 (CONTINUED). Potential parameters used in logistic regression models.

Parameter	Name
$mnvv \times mnvv$	$mnvv2$
$regvv \times regvv$	$regvv2$
$tmp \times tmp$	$tmp2$
$upt \times upt$	$upt2$
$mnt \times mnt$	$mnt2$
$regt \times regt$	$regt2$
$rhi \times rhi$	$rhi2$
$upr \times upr$	$upr2$
$mnr \times mnr$	$mnr2$
$regr \times regr$	$regr2$
$rhi \times tmp$	$rhitmp$
$upt \times upr$	$uptupr$
$mnt \times mnr$	$mntmnr$
$regt \times regr$	$regtregr$
$mnt \times rhi$	$mnrtrhi$
$(mnt \times rhi) \times (mnt \times rhi)$	$mnt2rhi2$
$tmp \times regt$	$tmpregt$
$mnt \times upt$	$mntupt$
$mnt \times regt$	$mntregt$
$rhi \times regr$	$rhiregr$
$rhi \times upr$	$rhiupr$
$rhi \times regvv$	$rhiregvv$
$mnt \times regvv$	$mntregvv$
$pwat \times pwat^*$	$pwat2$
$mnt \times trt^*$	$mnttrt$
$mnt \times pwat^*$	$mntpwat$
$raw \times raw$	$raw2$
$raw_s \times raw_s$	raw_s2
$rad \times rad$	$rad2$
$rad_s \times rad_s$	rad_s2
$brt \times brt$	$brt2$
$brt_s \times brt_s$	brt_s2
$raw \times mnt$	$rawmnt$
$rad \times mnt$	$radmnt$
$brt \times mnt$	$brtmnt$
$raw_s \times mnt$	raw_smnt
$rad_s \times mnt$	rad_smnt
$brt_s \times mnt$	brt_smnt
$(raw \times mnt) \times (raw \times mnt)$	$raw2mnt2$
$(rad \times mnt) \times (rad \times mnt)$	$rad2mnt2$
$(brt \times mnt) \times (brt \times mnt)$	$brt2mnt2$

TABLE 2. Locations of the GLOBE schools used in the development of the SURFACE models.

GLOBE school code	School Name	Location	Latitude	Longitude
LJhOS6Y	Most Pure Heart of Mary	Mobile, AL	30.70 N	88.05 W
c8t2giz	Ponderosa Elem. School	Fayetteville, NC	35.05 N	78.59 W
pWouwAn	Norfolk Elem. School	Norfolk, AR	36.20 N	92.27 W
YP8wiev	Norfolk Rebels 4-H Club	Mountain Home, AR	36.24 N	92.32 W
hzJ5KKx	Hartland Consolidated School	Hartland, ME	44.88 N	69.45 W
ZZSo0PT	Gold Trail School	Placerville, CA	38.78 N	120.89 W
ztYjGF9	Agua Caliente Park	Tucson, AZ	32.17 N	110.44 W
usozUPL	Waynesboro Sr. High School	Waynesboro, PA	39.75 N	77.57 W
bxU7W5h	Stone Child College	Box Elder, MT	48.29 N	109.87 W
mA5dQYm	Whitehall High School	Whitehall, MI	43.38 N	86.32 W
xXVJ4PP	Park View Elem. School	Washington, DC	38.56 N	77.01 W

TABLE 3. Percentage correct/skill scores for several versions of RUC and ARPS analyses and forecasts for the SURFACE models.

Model	Build 1 (dep.)	Build 1 (indep.)	Build 2 (dep.)	Build 2 (indep.)
RUC analysis + wv	0.704/0.291	0.644/0.186	0.873/0.627	0.853/0.534
RUC analysis	0.697/0.275	0.660/0.247	0.845/0.525	0.811/0.379
ARPS analysis + wv	0.794/0.522	0.714/0.347	0.914/0.751	0.841/0.507
ARPS analysis	0.721/0.370	0.672/0.206	0.884/0.661	0.778/0.351
ARPS 1-day forecast	0.730/0.305	0.667/0.214	0.872/0.661	0.856/0.597
Analysis eval. w/ 1-day	0.688/0.230	0.681/0.265	0.820/0.424	0.807/0.398
ARPS 2-day forecast	0.710/0.346	0.734/0.388	0.832/0.554	0.831/0.553
Analysis eval. w/ 2-day	0.674/0.262	0.682/0.188	0.798/0.431	0.802/0.450
ARPS 3-day forecast	0.707/0.202	0.679/0.187	0.808/0.351	0.688/0.179
Analysis eval. w/ 3-day	0.616/0.084	0.603/0.104	0.712/0.241	0.689/0.236

TABLE 4. Percent correct (PC) and Hanssen and Kuiper's determinant skill score for several versions of RUC and ARPS analyses and forecasts for the OUTBREAK models.

Model	(dep.) PC/HKD	(indep.) PC/HKD
RUC analysis + wv	0.962/0.923	0.917/0.838
RUC analysis	0.827/0.652	0.833/0.681
ARPS analysis + wv	0.971/0.942	0.861/0.713
ARPS analysis	0.959/0.919	0.872/0.731
ARPS 1-day forecast	0.866/0.722	0.675/0.323
Analysis eval. w/ 1-day	0.784/0.560	0.812/0.628
ARPS 2-day forecast	0.769/0.494	0.500/0.108
Analysis eval. w/ 2-day	0.691/0.375	0.600/0.287
ARPS 3-day forecast	0.779/0.555	0.467/-0.067
Analysis eval. w/ 3-day	0.672/0.344	0.347/-0.307

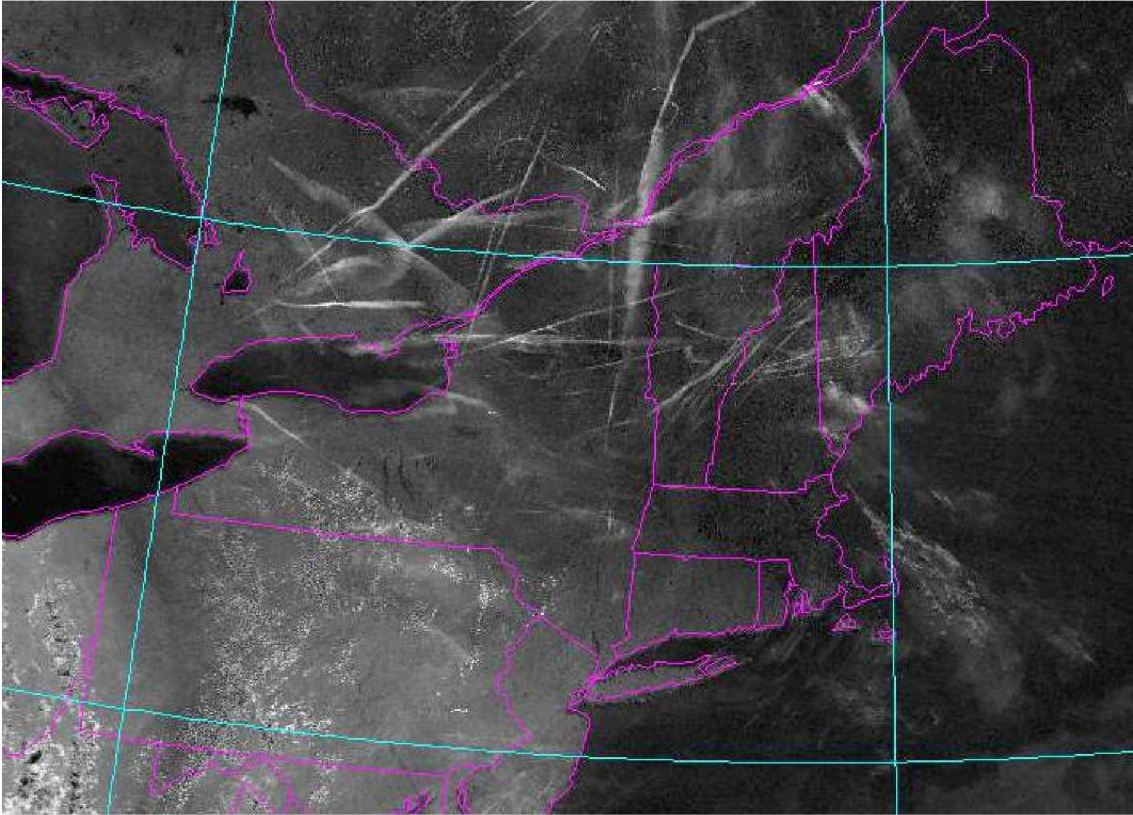


FIG. 1. Persistent contrails are highlighted in a Channel 4 ($10.8\text{ }\mu\text{m}$) minus channel 5 ($12.0\text{ }\mu\text{m}$) brightness temperature difference image from the *NOAA-17* overpass at 1531 UTC on 18 April 2005 over the northeastern United States.

Aug 2004--Aug 2005 CT outbreak obs. (over CONUS)

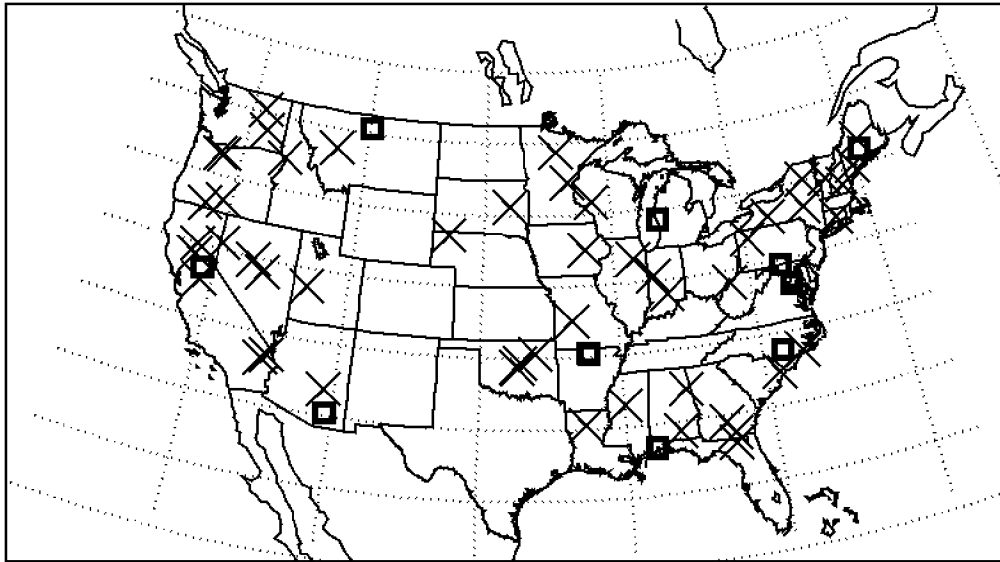


FIG. 2. Locations of persistent contrails identified from *GOES* imagery for OUTBREAK models (crosses), and GLOBE schools reporting persistent contrail coverage for the SUels (squares).