Statistical Contrail Forecasting

ARTIE JACKSON, BRIAN NEWTON, DOUG HAHN, AND ALLAN BUSSEY

Air Force Research Laboratory, Hanscom Air Force Base, Bedford, Massachusetts

(Manuscript received 28 April 1999, in final form 7 May 2000)

ABSTRACT

The current operational Air Force Weather Agency condensation trail (contrail) forecast technique is based on the Appleman algorithm, developed in the 1950s, with minor modifications. The Appleman algorithm requires accurate measurements or forecasts of ambient flight-level temperature, relative humidity, and pressure, as well as the amount of heat and water vapor added by an aircraft to its exhaust to determine accurately the critical temperature for contrail formation. Several factors contribute to the relatively poor contrail forecasts produced by the Air Force Weather Agency contrail forecast technique, including insufficiently accurate atmospheric measurements and numerical weather prediction forecasts of temperature and humidity at flight level, as well as some of the procedures used in the Air Force Weather Agency's implementation of the Appleman algorithm. The Contrail Field Program was conducted in eastern Massachusetts during a 10-day period in September 1995. Radiosonde data and aircraft observations were collected from a five-station network. Radiosondes were launched every 3 h, and aircraft observations included aircraft type, aircraft speed, aircraft altitude, and whether the aircraft produced a contrail. This dataset of nearly coincident (in time and space) radiosonde and aircraft observations was used to develop a new statistical regression contrail prediction model and to compare the results of the new statistical model with nowcasts produced by the Schrader algorithm, which is very similar to the Air Force Weather Agency contrail prediction technique, known as ''JETRAX.'' The Statistical Contrail Forecast Model makes use of logistic regression techniques to relate the presence or absence of observed contrails with nearly coincident radiosonde measurements. The statistical model produced a correct diagnosis of contrail occurrence or nonoccurrence for 85% of the observations, as compared with 58% correct for the Schrader technique, for this particular dataset.

1. Introduction

Interest in condensation trail (contrail) formation has increased in recent years because of the deployment of expensive stealth military jet aircraft that are designed to avoid detection by sophisticated surveillance radar, because they are visible with the naked eye when the aircraft produce contrails. In response to this renewed interest and to verification studies that suggest there is room for improvement in contrail forecasts produced by the Air Force Weather Agency (AFWA) contrail prediction technique, a program was initiated to study the weather conditions conducive to the formation of contrails by jet aircraft. The objective of the program is to measure the spatial distribution of atmospheric variables, especially water vapor, with ground-, air-, and space-based sensors, to observe simultaneously aircraft in flight to determine whether they are producing contrails, and to develop a new statistical regression-based contrail forecast algorithm that can be compared with the Schrader (1997) contrail forecast algorithm, which is very similar to the AFWA contrail prediction technique, known as ''JETRAX.''

In pursuit of these goals, the Contrail Field Program was conducted in eastern Massachusetts during a 10-day period in September 1995, in which nearly coincident (in time and space) radiosonde measurements and aircraft observations regarding the presence (''contrail-yes'') or absence (''contrail-no'') of contrails were collected. The Statistical Contrail Forecast Model was then developed using logistic regression techniques. It relates contrail-yes/contrail-no observations with nearly coincident radiosonde measurements collected during the field program. The field program data were also used to evaluate nowcasts produced by the Schrader (1997) algorithm, which is the basis of the current JETRAX AFWA contrail prediction technique, though there are minor differences that are described later.

Corresponding author address: Artie Jackson, Air Force Research Laboratory/VSBE, Hanscom AFB, Bedford, MA 01731.
E-mail: jackson@plh.af.mil

2. The Contrail Field Program development dataset

The Contrail Field Program was conducted in eastern Massachusetts during the 10-day period of 18–22 and 25–29 September 1995. Radiosondes were launched

FIG. 1. Locations of Contrail Field Program radiosonde sites in eastern Massachusetts.

from five sites every 3 h from 1200 to 0000 UTC, local daylight hours, to measure the distribution of temperature, humidity, and winds within the project's domain. The five sites that are shown in Fig. 1, include the Air Force Research Laboratory (AFRL) at Hanscom Air Force Base in Bedford; the University of Massachusetts, Lowell; the Beverly Municipal Airport; the AFRL Weather Test Facility at Otis Air National Guard Base in Falmouth; and the National Weather Service radiosonde launch site in Chatham. Vaisala, Inc., RS80 loran radiosondes were used at three sites and VIZ, Inc., Microsonde loran radiosondes were used at two sites. Coincident space-based sounding and imagery data, and ''TPQ-11'' vertically pointing cloud radar (35 Ghz, located at AFRL at Hanscom Air Force Base) data were also archived.

In addition to collecting meteorological data, observations of aircraft by ground observers were documented at each radiosonde site. The vast majority of aircraft observed were commercial aircraft at cruising altitudes between 28 000 and 43 000 ft. Only aircraft flying under cruise conditions above 25 000 ft were documented. Aircraft were not included in this study if they were undergoing elevation changes, that is, ascending to cruise altitude after takeoff or descending in preparation for landing. Data recorded for each aircraft observation included reported aircraft altitude, aircraft type, aircraft speed, and whether the aircraft was producing a contrail.

An AFRL observer at the Federal Aviation Administration (FAA) Air Route Traffic Control Center (ARTCC) in Nashua, New Hampshire, communicated by cellular phone with a ground observer at a radiosonde site to verify each aircraft observation. When an aircraft was observed at a radiosonde site, the ground observer phoned the observer at the ARTCC. The ground observer described the observed aircraft's location and direction of flight to identify the particular aircraft on the Air Route Traffic Control Display (ARTCD). If the observer at the ARTCC did not receive a phone call from a ground observer within several minutes of an aircraft entering a site's celestial dome, which was plotted on the ARTCD, the observer at the ARTCC phoned the ground observer to inform him of an aircraft within his celestial dome. The observer at the ARTCC described the location of the aircraft and direction of flight. The ground observer sighted the aircraft with the naked eye or with binoculars to identify the particular aircraft seen on the ARTCD. The ground observers documented each aircraft observation, noting the time the aircraft was first observed, the aircraft position relative to the observation site when first observed, the aircraft heading, whether the aircraft produced a contrail, and the time the observation was verified with the observer at the ARTCC. The observer at the ARTCC documented all verified aircraft observations from the five ground observation sites. The observations documented by the observer at the ARTCC included the aircraft altitude, aircraft speed, aircraft heading, and aircraft type from information available on the ARTCD, as well as the time the observation was verified with the ground observer and whether the aircraft produced a contrail. The database documented by the ARTCC observer was used to develop the Statistical Contrail Forecast Model.

Communication between the observer at the ARTCC and the ground observers at the radiosonde sites ensured that any aircraft, contrail-producing or noncontrail-producing, was likely to be observed and documented, reducing the common bias in many contrail observation studies of underobserving ''difficult to see'' noncontrail-producing aircraft. During the Contrail Field Program, more contrail-producing aircraft were observed than were noncontrail-producing aircraft, because atmospheric conditions at flight level favored contrail production, not because of observational deficiencies.

Thus, a comprehensive dataset was collected consisting of upper-air observations and aircraft contrail-yes/contrail-no observations. Each aircraft observation was paired with upper-air data (temperature, humidity, winds, and derived parameters) describing atmospheric conditions at flight level within 1.5 h of the aircraft observation. A daily summary of aircraft observations and daily mean flight-level conditions is presented in Table 1. Throughout the Contrail Field Program, 220 radiosondes were launched and 557 aircraft were observed and documented.

The temperature versus pressure and the relative humidity versus pressure distributions of the 557 aircraft observations are plotted in Figs. 2 and 3, respectively. Contrail-yes observations are indicated with a ''+,'' and contrail-no observations are indicated with a ''▲.'' Because overtyping, there are less than 557 points visible on the plots. There are 355 contrail-yes and 202 contrail-no data points plotted. Figure 2 shows the typical dis-

TABLE 1. Daily summary of aircraft observations and daily mean flight level conditions.

Date	Aircraft obs	Contrail yes	Contrail no	Wind speed (m s^{-1})	Wind direction	Temperature (°C)	Relative humidity
25 Aug	71	0	71	54.6	320.5	−43.0	5.6
18 Sep	51	35	16	24.9	236.2	−48.8	9.3
19 Sep	79	74	5	12.6	306.1	−48.7	9.5
20 Sep*	0						
21 Sep**	37	23	14	19.4	274.1	−44.3	24.1
22 Sep**	20	19	1	37.5	260.8	−43.8	38.9
25 Sep**	21	21	0	42.7	236.9	−52.6	38.5
26 Sep*	0						
27 Sep**	43	4	39	38.5	255.5	−46.5	7.9
28 Sep	109	78	31	18.9	277.1	−50.0	13.9
29 Sep	126	101	25	11.7	233.7	−50.1	22.8
Total	557	355	202				

* No observations on these dates because of rain.
** No observations for a portion of these dates because of overcast sky conditions.

tribution of contrail-yes/contrail-no observations, with the majority of contrail-yes observations occurring at lower pressures and colder temperatures; contrail-no observations are generally found at higher pressures and warmer temperatures. Figure 3 shows that there are only a few contrail-no observations at relative humidity greater than 20%; however, contrail-yes observations are present in significant numbers over the entire range of observed relative humidity. It is not uncommon to observe contrails at very low relative humidity because aircraft add significant quantities of water vapor to the atmosphere in their exhaust plume.

3. AFWA contrail forecast techniques

AFWA operational contrail forecast techniques during the past 40 years have been based on the Appleman (1953) algorithm. Successful application of the Appleman algorithm depends on accurate measurements or forecasts of ambient flight-level temperature, relative humidity, and pressure as well as accurate estimates of heat and water vapor added to the exhaust plume by an aircraft to determine the critical temperature for contrail formation. Several modifications have been incorporated over the years to account for the variations of water vapor and heat added to the exhaust plume by different engine types used in modern jet aircraft (Peters 1993; Schrader 1997). Bjornson (1992) developed an experimental contrail forecast technique using discriminant analysis methods to relate temperature, altitude, and vertical motion to contrail formation. The technique showed some improvement over the Appleman algorithm, but overall results were inconclusive.

The operational AFWA contrail forecast technique (JETRAX), based on the Appleman algorithm, uses the updated contrail factors suggested by Schrader (1997). The critical temperature is calculated iteratively using a form of the Goff–Gratch formula for saturation vapor pressure over a water surface (Shull 1998) and the definition of relative humidity in terms of vapor pressure

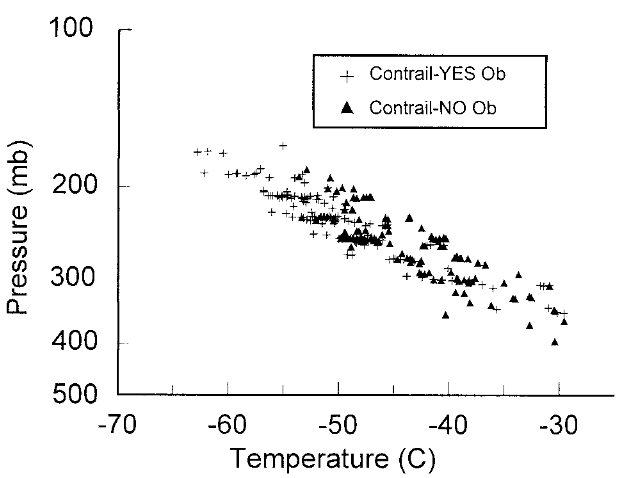

FIG. 2. Scatterplot of temperature (°C) vs ln pressure (hPa) for the 557 aircraft observations. There are 355 Contrail-yes and 202 Contrail-no observations plotted.

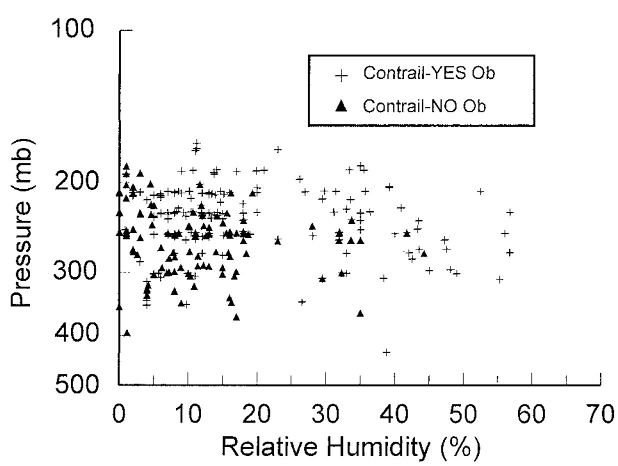

FIG. 3. Scatterplot of relative humidity (%) vs ln pressure (hPa) for the 557 aircraft observations. There are 355 Contrail-yes and 202 Contrail-no observations plotted.

and saturation vapor pressure as described by Schrader (1997) and Shull (1998). JETRAX also uses a blend of Numerical Weather Prediction (NWP) model forecast data and assumed relative humidity data as input (Shull 1998).

JETRAX is essentially the Schrader (1997) algorithm, modified to include a set of rules to convert critical temperatures to contrail layers. Beginning in the lower levels, the technique locates the first level where the atmospheric temperature is colder than the critical temperature for contrail formation. This level is a contrail layer base. The technique continues comparing the atmospheric temperature and the critical temperature at increasingly higher levels until the atmospheric temperature becomes greater than the critical temperature. This is the contrail layer top. Multiple contrail layers are possible. If a layer is less than 2000 ft thick, it is rejected. Layers less than 2000 ft apart are combined.

Because of the poor quality of relative humidity measurements and forecasts at typical flight altitudes, AFWA contrail forecast techniques have been tested experimentally using various relative humidity inputs during the past decade. In evaluating the Schrader (1997) algorithm, three relative humidity inputs that have been tested by AFWA were used in this study. The first relative humidity input tested, used in the current operational AFWA technique and referred to as Schrader 40/70/10, consists of Navy Operational Global Atmospheric Prediction System (NOGAPS) relative humidity data at flight levels below the 300-hPa level and a set of assumed relative humidity estimates for various layers above the 300-hPa level. For tests presented here, radiosonde relative humidity observations were substituted below the 300-hPa level. Relative humidity is assumed to be 40% from the 300-hPa level to 300 m below the tropopause, 70% from 300 m below the tropopause to 300 m above the tropopause, and 10% at altitudes greater than 300 m above the tropopause. The assumed 40/70/10 relative humidity values at a particular altitude above 300 hPa are constant geographically and seasonally. The second relative humidity input tested is the radiosonde-measured relative humidity observation at aircraft altitude (from the Contrail Field Program dataset), referred to as ''Schrader raob.'' The third relative humidity input tested for the Schrader algorithm is a climatological value derived from the Stratospheric Aerosol and Gas Experiment II (SAGE 2) (McCormick and Chiou 1994) database and is referred to as Schrader SAGE. SAGE climatological relative humidity values at a particular altitude vary geographically and seasonally.

The Schrader (1997) algorithm evaluated in this study is not exactly the same as the JETRAX. We applied the Schrader (1997) algorithm at the altitude of a particular aircraft observation, that is, the AFWA technique without the rules that convert critical temperatures to contrail layers. There are two other minor differences. The operational AFWA technique uses the tropopause height as determined by the NOGAPS model ouput. The tropopause height for the Contrail Field Program radiosonde data was determined using the World Meteorological Organization definition of the tropopause. The operational AFWA technique uses NOGAPS model forecast data as input (except for the assumed relative humidity values above 300 hPa as described above), whereas nearly coincident radiosonde measurements were used as input in this study. It is reasonable to assume that the Schrader (1997) algorithm, as applied in this study using near-real-time input data (as opposed to NWP model forecast data), and avoiding the intermediate step of converting critical temperature calculations to contrail layers, should perform equally well, or somewhat better than JETRAX.

To test the Schrader (1997) algorithm, pressure at the flight level of each aircraft observation from the Contrail Field Program dataset was input, along with the estimated contrail factor of the observed aircraft and the relative humidity appropriate for the flight altitude from the three relative humidity data sources described above. The contrail factor describes the ratio of water vapor produced by the combustion of jet fuel to heat added to the exhaust of a jet engine. The contrail factors suggested by Schrader (1997) are used in this study (nonbypass engine: 0.0300 g kg^{-1} $°C^{-1}$, low-bypass engine: 0.0340 g kg^{-1} $°C^{-1}$, high-bypass engine: 0.0390 g kg^{-1} $°C^{-1}$). A critical temperature was then calculated from the input parameters for each aircraft observation. If the observed flight level atmospheric temperature was less than or equal to the calculated critical temperature, contrails were forecast. If the observed atmospheric temperature was greater than the calculated critical temperature, contrails were not forecast.

4. Statistical logistic regression contrail forecast model

The Statistical Contrail Forecast Model using radiosonde data as input was developed as a proof of concept for the use of logistic regression techniques to forecast the occurrence of contrails produced by jet aircraft. NWP model output variables are currently used as predictors in regression techniques to forecast many meteorological parameters such as cloud amount, cloud height, and precipitation probability (Glahn and Lowry 1972). The ability of regression techniques to develop relationships between predictors and predictands and to account for biases in NWP model output variables has made the use of regression techniques a valuable weather forecasting tool. Travis et al. (1997) used satellite imagery data predictors as input to a logistic regression technique to predict successfully the occurrence of large-scale outbreaks of contrails in a study of the effects of long-lived contrails on climate.

Contrail forecasts produced by the operational AFWA contrail forecast technique based on the Appleman algorithm over the past decade have not been as accurate as desired because the technique requires accurate and precise input data. Errors in NWP model forecast temperature and relative humidity data, and assumed relative humidity data used as input to these techniques, often do not represent the true state of the atmosphere, thus contributing to less accurate contrail forecasts. The inability of radiosondes to provide accurate relative humidity measurements at high altitudes (cold temperatures) contributes to poor NWP model relative humidity analyses and forecasts and can be a significant factor in the accuracy of the resultant contrail forecasts. Although temperature is measured accurately by radiosondes, temperature errors in NWP forecasts can also contribute to significant errors in contrail forecasts, because the NWP forecast temperature is compared with the calculated critical temperature to make a contrail forecast. In applying regression techniques to contrail forecasting, the ability of regression techniques to develop relationships between the predictors (atmospheric measurements and derived parameters) and the predictand (contrail-yes/contrail-no observations) and to account for predictor biases provides a potential advantage over the operational contrail forecast technique. The Statistical Contrail Forecast Model was developed to take advantage of these beneficial qualities of regression techniques.

Because of the dichotomous nature of the predictand (contrail-yes = 1, contrail-no = 0), logistical regression (Neter et al. 1989) was chosen to develop the Statistical Contrail Forecast Model. Logistical regression develops relationships between the predictand and a finite number of predictors. In this case, the predictor pool consists of measured and derived atmospheric parameters from the radiosonde measurements and aircraft observations collected during the Contrail Field Program.

The Statistical Contrail Forecast Model is a multivariate linear regression model. The relationship between the predictand (Contrail-yes/no observations) and the predictors is linear, and the effect of the predictors in the model is additive, that is, there is no interaction between predictors. The prediction output of a logistic regression model is a probability value between 0 and 1. For this study, contrails were predicted for probability values greater than or equal to 0.5, while contrails were not predicted for probabilities below 0.5.

Predictors for the model were chosen using a combination of a stepwise regression function and tests performed against a null model for which the coefficients of all predictors are assumed to be zero. The stepwise regression function evaluates the explanatory power of each predictor individually and in combination with the others, selecting only those predictors that explain a significant amount of the variance in the predictand. The model developed using only the predictors considered significant by the stepwise regression function did not perform well.

Tests were then performed with a null model. Predictors are added to the null model one at a time to determine the relationship between the predictand and each predictor. The C_p statistic is a useful quantity used in predictor selection (Afifi and Clark 1990). Predictors with a C_p value less than the C_p value for the null model are determined to be significant and may make a positive contribution to model development. The predictors determined to be significant in the null model tests, which included the predictors found to be significant in the stepwise regression model, were used to develop the logistic regression model.

The dataset used to develop the Statistical Contrail Forecast Model contains 557 contrail-yes/contrail-no aircraft observations listed chronologically in the order they were observed. Because of its limited size, the dataset was divided into four data "quadrants." The first data quadrant consisted of every fourth observation, starting with the first observation, the second data quadrant consisted of every fourth observation, starting with the second observation, and so on. Four logistic regression models were developed, each using three of the data quadrants for model development and the fourth data quadrant to test the model. For example, a logistical regression model, SM Quad 1, was developed using all the observations in data quadrants 1, 2, and 3, and the model was tested using the data from quadrant 4. Thus, each observation is tested only once on a statistical model developed on a dataset consisting of three-quarters (quadrants) of the entire observational dataset. The data quadrants used to develop and to test each logistic regression model are shown in Table 2. With a limited dataset, developing four logistic regression models with three-quarters of the dataset and testing each model on the remaining one-quarter of the dataset maximizes the use of the dataset and provides some confidence in the technique if results of the four quadrant models are of comparable quality.

Another logistical regression model, the Incestuous Logistic Regression Statistical Contrail Forecast Model (SM ILR) was developed using the entire dataset (all 557 contrail-yes/contrail-no observations). The model was then tested using the same entire development dataset. Developing and testing the SM ILR on the same dataset serves two purposes. The results produced by such a model may suggest that forecasting contrails with

TABLE 2. Data quadrants and number of observations used to develop and to verify the Statistical Contrail Forecast Models.

Logistic regression model	Development data quads	Verification data quad(s)	Development observations	Verification observations
SM ILR	1, 2, 3, 4	1, 2, 3, 4	557	557
SM Quad 1	2, 3, 4	1	417	140
SM Quad 2	1, 3, 4	2	418	139
SM Quad 3	1, 2, 4	3	418	139
SM Quad 4	1, 2, 3	4	418	139
SM Quad Avg	*	*	557	557

* See discussion in results section.

a logistic regression model is an idea worthy of further study. Also, the results of a regression model developed and tested on the same dataset will typically produce better results than a regression model developed on a portion of the dataset and tested on the remaining portion of the dataset. If the SM ILR produces favorable results, and the SM Quad models produce similar results to the SM ILR, it would provide some support for the validity of the SM Quad models. Thus, the SM ILR is a useful model to compare with the SM Quad models described above.

5. Results

Table 3 lists the regression coefficients and standardized coefficients for each predictor, for each statistical model. The regression coefficients are used directly in the regression equation to predict the occurrence or nonoccurrence of a contrail by a particular aircraft. The standardized coefficients can be compared to determine the relative contribution of each predictor (Afifi and Clark 1990). The larger the magnitude of the standardized coefficient, the more the predictor contributes to the prediction of the predictand, in this case the occurrence or nonoccurrence of a contrail by a particular aircraft. Twenty-one predictors were used to develop the logistic regression models. The predictors are listed in Table 3 in decreasing order of importance for SM Quad 1, that is, the predictors with the larger magnitude standard coefficients (the greatest contribution) at the top.

Temp, temp2, temprh, and rh (predictor abbreviations are described in Table 3) are the top four predictors in all five models. Temp and temp2 have significantly larger standardized coefficients (on the order of 80 to 132) than all other predictors do, indicating that temperature by far makes the most significant contribution to the prediction of contrail occurrence and is thus the most important parameter in predicting contrail formation in this model. Figure 2 suggests a strong relationship between the occurrence of contrails and flight-level temperature, with the majority of contrail-yes observations occurring at colder temperatures and lower pressures; contrail-no observations are generally found at higher pressures and warmer temperatures. Combined temperature and humidity predictors (temprh, temp2rh2) and relative humidity predictors (rh, rh2) also make a significant contribution, with standardized coefficients on the order of 14–31, though these contributions are only on the order of 15%–30% of the temperature predictors' contributions. Figure 3 shows that, although contrail-yes observations occur over the entire range of relative humidity, there is a strong tendency for contrail-no observations to occur at relative humidity less than 20%. The other predictors tend to contribute significantly less than the temperature and humidity predictors, with standardized coefficients ranging from 0.1 to 10.

The minor contributions by upper-air wind parameters (maxdir, maxspeed, wndspd) indicates a relationship between the occurrence of contrails and the flight-level synoptic situation. Southwesterly upper-air winds in New England typically imply the northward advection of high-level moisture, increasing the potential for aircraft to produce contrails. In diagnosing the presence of clouds using multiple linear regression techniques, Norquist et al. (1997) also found wind parameters to be important predictors. Detwiler and Pratt (1984) found that contrails often form under certain synoptic situations, such as ahead of cyclonic storms.

The bypass ratio of each aircraft was offered to the logistic regression model as a predictor, though the value was found to make no contribution to the model. The bypass ratio describes the ratio of the amount of air passed around an aircraft engine core to the amount of air passed through the engine core and ultimately determines the amount of heat and water vapor added to the exhaust of a particular aircraft. Modern jet engines are categorized into three bypass ratio types: high bypass, low bypass, and nonbypass. If a statistical contrail forecast model was developed utilizing a much larger dataset, a separate logistic regression model should be developed for each engine bypass type.

Tables 4 and 5 compare the distribution of correct and incorrect forecasts for the logistic regression Statistical Contrail Forecast Models and the Schrader (1997) algorithm using the three relative humidity inputs previously described. The tables show the number of correct forecasts and total number of forecasts for each technique. The number of correct forecasts is broken down into the number of observed contrail occurrences that were forecast as contrail occurrences (Yes/FYes) and the number of observed contrail nonoccurrences that were forecast as contrail nonoccurrences (No/FNo). The number of incorrect forecasts is broken down into the number of observed contrail occurrences that were forecast as contrail nonoccurrences (Yes/FNo) and the number of observed contrail nonoccurrences that were forecast as contrail occurrences (No/FYes). Table 6 compares performance statistics calculated for the logistic regression Statistical Contrail Forecast Models and the Schrader (1997) algorithm.

For the purpose of evaluation, and comparison of the statistical model with the Schrader (1997) algorithm, the results of the four logistic regression models (SM Quad 1–4) were combined so that statistics for the 557 Schrader (1997) algorithm contrail forecasts could be compared directly with the statistics of the combination of the 557 contrail forecasts produced by the four logistic regression models. Results of the four logistic regression models were totaled or averaged, as appropriate, and labeled SM Quad Avg in Table 2 and subsequent tables. In Table 4, the number of observations in each category, including the number of correct forecasts and the total number of forecasts, are totaled for the four regression models (SM Quad 1–4) and labeled SM Quad Avg. In Table 6, each of the statistical quantities presented, including percent correct forecasts and

TABLE 3. Regression coefficients and standardized coefficients (St coeff) for each statistical model. Standardized coefficient = coefficient$_{Xi}$ × (standard deviation$_{Xi}$/standard deviation$_Y$), X_i = predictors, Y = predictand. Predictors are given as follows. temp: temperature*, temp2: temp × temp*, rh: relative humidity*, temprh: temp × rh*, rh2: rh × rh*, temp2rh2: temprh × temprh*, ws: mixing ratio*, es: saturation vapor pressure*, troptemp: temperature at tropopause, rh300: rh at 300 hPa, rh200: rh at 200 hPa, windspd: wind speed*, wnddir: wind direction*, pre: pressure*, altd: altitude*, speed: aircraft speed, maxspeed: maximum wind speed**, maxdir: direction of maximum wind speed**, maxheight: height of maximum wind speed**, ra32avgdp: [(rh + rh300 + rh200)/3]/[max(pre, 300, 200) − min(pre, 300, 200)], drdpa2: (rh200 − rh)/(pre − 200). Here, * indicates at flight level, ** indicates in sounding.

	SM Quad 1		SM Quad 2		SM Quad 3		SM Quad 4		SM ILR	
	Coefficient	St coeff	Coefficient	St coeff	Coefficient	St coeff	Coefficient	St coeff	Coefficient	St coeff
Intercept	142.98		198.48		278.18		47.3497		184.75	
Temp	7.2573	87.24	8.6197	105.54	10.5851	132.15	8.3446	98.60	9.0047	109.33
Temp2	0.0723	79.79	0.0878	98.84	0.1021	117.69	0.0913	100.11	0.0913	102.25
Temprh	−0.0234	−29.27	−0.0210	−27.60	−0.0241	−31.53	−0.0213	−27.50	−0.0216	−27.92
Rh	−0.9680	−24.95	−0.8732	−23.39	−1.0232	−27.25	−1.0321	−27.25	−0.9130	−24.10
Rh2	0.0139	16.47	0.0111	14.16	0.0152	19.49	0.0145	18.21	0.0130	16.23
Temp2rh2	-5.90×10^{-6}	−16.45	-5.18×10^{-6}	−16.05	-6.02×10^{-6}	−18.66	-5.82×10^{-6}	−17.52	-5.45×10^{-6}	−16.35
Ws	−36 680.55	−8.57	−20 782.69	−4.98	−80 969.62	−19.34	37 040.43	8.30	−21 571.00	−5.05
Ra32avgdp	−25.7226	−6.26	−17.8347	−4.47	−30.7355	−7.66	3.3895	−0.84	−22.2190	−5.51
Es	−33.5965	−4.28	−69.0099	−9.00	3.4876	0.45	−153.73	−18.45	−74.5607	−9.43
Troptemp	−0.6338	−3.36	−0.6041	−3.20	−0.5708	−3.02	−0.3001	−1.58	−0.5380	−2.84
Rh300	0.0738	2.24	0.0782	2.43	0.0858	2.66	−0.0589	−1.81	0.0592	1.82
Wndspd	−0.0637	−2.03	−0.0971	−3.13	−0.0977	−3.13	−0.1132	−3.55	−0.0933	−2.98
Pre	0.0266	1.95	0.0010	0.07	−0.0512	−3.80	0.2505	17.39	0.0403	2.92
Maxspeed	0.0614	1.79	0.0822	2.38	0.0702	2.08	0.0900	2.59	0.0783	2.28
Maxdir	−0.0288	−1.72	−0.0332	−1.96	−0.0344	−2.07	−0.0308	−1.83	−0.0312	−1.86
Rh200	−0.0656	1.47	0.0447	1.03	0.0770	1.76	0.0319	0.72	0.0646	1.47
Altd	2.07×10^{-4}	1.38	-2.47×10^{-4}	−1.55	-2.28×10^{-4}	−1.54	0.0022	13.83	2.65×10^{-4}	1.74
Drdpa2	1.4761	0.58	0.7333	0.35	1.8172	0.86	0.4834	0.20	1.2611	0.56
Maxheight	-2.67×10^{-5}	−0.40	-3.76×10^{-5}	−0.58	-4.25×10^{-5}	−0.65	-3.81×10^{-5}	−0.58	-3.51×10^{-5}	−0.53
Speed	−0.0026	−0.27	−0.0051	−0.45	−0.0031	−0.32	−0.0097	−1.01	−0.0044	−0.44
Wnddir	−0.0017	−0.14	0.0013	0.10	0.0038	0.31	0.0058	0.44	0.0023	0.18

TABLE 4. Distribution of correct and incorrect contrail forecasts for the Statistical Contrail Forecast Models. Yes/FYes: observed contrail occurrences forecast as contrail occurrences; Yes/FNo: observed contrail occurrences forecast as contrail nonoccurrences; No/FYes: observed contrail nonoccurrences forecast as contrail occurrences; No/FNo: observed contrail nonoccurrences forecast as contrail nonoccurrences; #Correct: total number of correct forecasts; # Forecasts: total number of forecasts.

Models	Yes/FYes	Yes/FNo	No/FYes	No/Fno	#Correct	#Forecasts
SM ILR	330	25	44	158	488	557
SM Quad 1	84	6	11	38	122	139
SM Quad 2	81	7	12	40	121	140
SM Quad 3	80	4	15	40	120	139
SM Quad 4	77	16	12	34	111	139
SM Quad Avg	322	33	50	152	474	557

the discriminant ''*V*'' score (referenced below), for the four regression models (SM Quad 1–4) are averaged and labeled SM Quad Avg.

Results of the four quadrant tests of the statistical model (SM Quad 1–4), shown in Table 4, are consistent, though SM Quad 4 results are slightly poorer than the other three. The regression coefficients and standardized coefficients for the models, shown in Table 3, also are consistent, though SM Quad 4 again shows somewhat more variation. The consistency of the four quadrant tests provides confidence in the performance of the logistic regression technique. The SM ILR model produced 488 correct forecasts, only slightly outperforming the combined SM Quad Avg model, which produced 474 correct forecasts, thus increasing confidence in the statistical models. The SM ILR model should provide the best results possible, given that it is developed and tested on the same entire dataset of 557 contrail-yes/contrail-no observations.

The Schrader (1997) algorithm, shown in Table 5, produced very similar results for each of the three relative humidity inputs, though the Schrader 40/70/10 algorithm results were slightly better than the others. It is interesting to note that the Schrader 40/70/10 algorithm, which inputs assumed relative humidity values above the 300-hPa level, performed slightly better than the Schrader raob algorithm, which inputs near-real-time radiosonde relative humidity measurements. This apparent contradiction may be due to the poor quality of radiosonde-measured relative humidity at very cold temperatures (Wade 1994) or the relatively small size of the dataset. The Schrader 40/70/10 algorithm produced only 325 correct forecasts, significantly poorer than the combined SM Quad Avg and the SM ILR models. The combined SM Quad Avg and the SM ILR results will be compared with the Schrader 40/70/10 results in the following discussion.

The statistics calculated for each model, shown in Table 6, include the probability of detection (POD), false alarm rate (FAR), and critical success index (CSI) (Donaldson et al. 1975) for observed contrail occurrences (POD/Y, FAR/Y, CSI/Y) and for observed contrail nonoccurrences (POD/N, FAR/N, CSI/N), bias ratio (bias), percent correct (PC), and Hanssen and Kuipers' (Hanssen and Kuiper 1968) discriminant *V* score (VDS). Formulas for the statistical measures of accuracy are shown in Table 7, as calculated using a contingency table as shown in Table 8. The VDS, often used in contrail technique comparison studies, ranges from -1 (no skill) to $+1$ (total skill). The VDS accounts for biases in a test where events (contrail-yes observations) and nonevents (contrail-no observations) are not equally represented.

Pearson chi-square and *p* values were calculated for the three models presented in Table 6 to determine the statistical significance of each model's performance. The chi-square test for independence is used to demonstrate the existence of a relationship between the row and column classifications in a contingency table, that is, to show that the strong relationship shown in the contingency tables of observing contrails when they are forecast is not due to chance. The chi-square value is 292.43 for the SM ILR model, 249.48 for the SM Quad Avg model, and 83.72 for the Schrader 40/70/10 algorithm. Each model has a very small *p* value, and all results are statistically significant.

The SM ILR model only slightly outperforms the combined SM Quad Avg model for all statistics. The combined SM Quad Avg model significantly outperforms the Schrader 40/70/10 algorithm for all statistics except FAR/Y and POD/N. The VDS is 0.66 for the

TABLE 5. Distribution of correct and incorrect contrail forecasts for the Schrader contrail forecast models. Yes/FYes: observed contrail occurrences forecast as contrail occurrences; Yes/FNo: observed contrail occurrences forecast as contrail nonoccurrences; No/FYes: observed contrail nonoccurrences forecast as contrail occurrences; No/FNo: observed contrail nonoccurrences forecast as contrail nonoccurrences; #Correct: total number of correct forecasts; # Forecasts: total number of forecasts.

Models	Yes/FYes	Yes/FNo	No/FYes	No/Fno	#Correct	#Forecasts
Schrader 40/70/10	125	230	2	200	325	557
Schrader raob	115	240	1	201	316	557
Schrader SAGE	116	239	1	201	317	557

Table 6. Comparison of the ILR Statistical Contrail Forecast Model, the Quad Avg Statistical Contrail Forecast Model, and the Schrader 40/70/10 contrail forecast model statistics. All calculated p values were < 0.001. POD/Y, FAR/Y, CSI/Y, and POD/N, FAR/N, CSI/N: probability of detection, false alarm rate, and critical success index for observed contrail occurrences and observed contrail nonoccurrences, respectively; bias: bias ratio; PC: percent correct forecasts; VDS: discriminant V score.

	Contrail-yes obs			Contrail-no obs			All obs		
Models	POD/Y	FAR/Y	CSIY	POD/N	FAR/N	CSIN	Bias	PC	VDS
SM ILR	0.93	0.12	0.83	0.78	0.14	0.70	1.05	0.88	0.71
SM Quad Avg	0.91	0.13	0.80	0.75	0.17	0.65	1.05	0.85	0.66
Schrader 40/70/10	0.35	0.02	0.35	0.99	0.53	0.46	0.36	0.58	0.34

combined SM Quad Avg model, but the VDS is 0.34 for the Schrader 40/70/10 algorithm, indicating that the combined SM Quad Avg model provides a significant improvement in contrail forecasting skill. POD/Y for the combined SM Quad Avg model is 0.91 and is 0.35 for the Schrader 40/70/10 algorithm, pointing out that the Schrader 40/70/10 algorithm substantially underpredicts the occurrence of contrails. Percent Correct (PC) for the combined SM Quad Avg model is 0.85; it is 0.58 for the Schrader 40/70/10 algorithm. In all aspects except FAR/Y and POD/N, the combined SM Quad Avg model performance is superior to the Schrader 40/70/10 algorithm.

The smaller FAR/Y and POD/N for the Schrader 40/70/10 algorithm are deceiving, because it characteristically underpredicts contrail occurrences and overpredicts contrail nonoccurrences for the observations tested in this study. JETRAX, based on the Schrader (1997) algorithm, also underpredicts contrail occurrences and overpredicts contrail nonoccurrences (Shull 1998). Note, in Tables 4 and 5, that the combined SM Quad Avg model correctly predicted 322 of 355 observed contrail occurrences, but the Schrader 40/70/10 algorithm correctly predicted only 125 of 355 observed contrail occurrences. The bias ratio for the Schrader 40/70/10 algorithm is 0.36, indicating a strong tendency to underforecast contrail occurrences. The bias ratio for the combined SM Quad Avg model is 1.05, indicating a slight tendency to overforecast contrail occurrences. A perfectly unbiased forecast technique would produce a bias ratio of 1.00. The combined SM Quad Avg model correctly predicted 152 of 202 observed contrail nonoccurrences. The Schrader 40/70/10 algorithm correctly predicted 200 of 202 contrail nonoccurrences, but in total predicted 430 contrail nonoccurrences. The tendency of the Schrader (1997) algorithm to underpredict the occurrence of contrails and to overpredict the nonoccurrence of contrails is typical of past and current operational AFWA contrail forecast techniques, is well documented (Speltz 1995; Peters 1993), and results in a deceivingly small FAR/Y and large POD/N. However, the overprediction of contrail nonoccurrences results in a large FAR/N of 0.53 for the Schrader 40/70/10 algorithm; the combined SM Quad Avg model has a more reasonable FAR/N of 0.17. The Schrader 40/70/10 algorithm statistics are significantly poorer than the combined SM Quad Avg model, though they are consistent with recent Air Force contrail studies reported by Speltz (1995) and Peters (1993). In a more recent verification study, JETRAX produced somewhat better results (e.g., higher hit rates), which were comparable to nowcasts produced using the Schrader (1997) algorithm, indicating that JETRAX may perform somewhat better under some synoptic conditions (Walters and Shull 1999).

Table 7. Formula table. POD/Y, FAR/Y, CSI/Y, and POD/N, FAR/N, CSI/N: probability of detection, false alarm rate, and critical success index for observed contrail occurrences and observed contrail nonoccurrences, respectively; bias ratio: bias ratio; PC: percent correct forecasts; VDS: discriminant V score. Here, A, B, C, and D are elements in a 2×2 contingency table as shown in Table 8, and N is the total number of observations.

POD/Y	A/(A + C)
POD/N	D/(B + D)
FAR/Y	B/(A + B)
FAR/N	C/(C + D)
CSIY	A/(A + B + C)
CSIN	D/(B + C + D)
Bias ratio	(A + B)/(A + C)
PC	(A + D)/N
VDS	(AD − BC)/[(A + C)(B + D)]

6. Summary

The Contrail Field Program was conducted in eastern Massachusetts during a 10-day period in September 1995. Radiosonde data and aircraft contrail observations were collected. This dataset of nearly coincident radiosonde observations and aircraft observations was used to test the Schrader (1997) contrail forecast algorithm, which is very similar to the operational AFWA contrail forecast technique (JETRAX), and to develop a new statistical regression-based contrail forecast algorithm. The Statistical Contrail Forecast Model was developed, which makes use of logistic regression techniques to relate contrail-yes/contrail-no observations with nearly coincident radiosonde measurements. In this study,

Table 8. Contingency Table. Fcst yes: forecast yes; Fcst no: forecast no; Obs yes: observed yes; Obs no: observed no.

	Obs yes	Obs no
Fcst yes	A	B
Fcst no	C	D

flight-level temperature and relative humidity parameters were found to be the most important predictors. The Statistical Contrail Forecast Model provides results that are superior to the Schrader (1997) contrail prediction algorithm. The statistical model produced a correct diagnosis of contrail occurrence or nonoccurrence for 85% of the observations, as compared with 58% correct for the Schrader 40/70/10 algorithm. The probability of detection of contrail occurrences (POD/Y) was 0.91 for the statistical model and 0.35 for the Schrader 40/70/10 algorithm. The results for the Schrader (1997) algorithm presented here are consistent with past Air Force contrail studies that document the tendency of past and present AFWA techniques to underpredict the occurrence of contrails and overpredict the nonoccurrence of contrails. For this dataset, the statistical model provides a significant improvement over the Schrader (1997) algorithm tested concerning percent correct, POD and FAR statistics. For reasons stated earlier, the results obtained using the Schrader (1997) algorithm should be comparable to or better than results obtained using the JETRAX. Thus it can be inferred that, for this dataset, the Statistical Contrail Forecast Model would provide a significant improvement over JETRAX.

Current operational contrail forecast techniques often produce poor forecasts caused by errors in input variables such as temperature, humidity, and contrail factor. NWP model forecast data and climatological or assumed relative humidity data used as input to these techniques often do not represent the true state of the atmosphere, thus contributing to less accurate contrail forecasts. Two important potential sources of error in applying the JETRAX technique are the procedure that eliminates layers that are less than 2000 ft thick and combines layers that are less than 2000 ft apart, and the use of estimated contrail factors. In applying regression techniques to contrail forecasting, the ability of regression techniques to develop relationships between the predictors and the predictand and to account for predictor biases provides a distinct advantage in this study over nowcasts made with the Schrader (1997) algorithm, which would likely be more accurate than similar JETRAX forecasts. The Statistical Contrail Forecast Model has been developed to take advantage of these beneficial qualities of regression techniques.

It must be noted that the Statistical Contrail Forecast Model, using near-real-time radiosonde data as input, was developed with a relatively small dataset (557 observations) from a two-week period in September 1995 in eastern New England. Thus, application of the model may be limited geographically and seasonally. However, the result presented in this study is a very successful proof of concept for regression-based contrail forecasting techniques.

Application of a statistical contrail forecast model operationally would likely use input from a NWP model to produce 6- to 48-h contrail forecasts, or potentially would use input from real-time satellite sounding and imagery data for 0- to 6-h contrail forecasts. Future work will concentrate on the use of these data sources, which are available in all areas, including those void of conventional upper-air data, as input to regression-based contrail forecast techniques.

The accuracy of contrail forecast techniques based on the Appleman algorithm is limited by the accuracy of the measurements or forecasts used as input and the accuracy of the estimated contrail factor. Improvements in the accuracy of atmospheric measurements and NWP forecasts are difficult. The statistical approach to contrail forecasting can account for the biases in these measurements and forecasts and offers the possibility of improving the accuracy of contrail forecasts, though attaining the database necessary for development of statistical contrail forecast models is also difficult.

Building the dataset necessary to develop an operational statistical contrail forecast model for a particular region is not a trivial task. The Air Force operates in a large number of theaters around the globe. Many are of strategic interest for extended periods of time ranging from several years to decades. All of the theaters have high-quality NWP model output and satellite data available. Likely, high-quality aircraft/contrail observations do not exist for most of the theaters. A high-quality database consisting of aircraft position, aircraft altitude, and whether the aircraft produced a contrail would need to be developed for each theater, for each season. For regions where strategic planning suggests that operations may occur over an extended period, development of a high-quality aircraft/contrail database could be beneficial both to develop a new statistical contrail forecast model and to evaluate the existing contrail forecast techniques based on the Appleman algorithm. An alternative to developing several statistical contrail forecast models for several strategic regions is to develop a statistical contrail forecast model that is applicable globally through the rigorous choice of NWP and satellite-derived predictors that are less dependent on geographical influences. Developing the necessary aircraft/contrail database would be a significant, though manageable, task. The successful results obtained using regression techniques in the development of the Statistical Contrail Forecast Model using radiosonde data as input provide encouragement for future model development using NWP output and satellite data as input to regression-based contrail forecast techniques.

Acknowledgments. We thank Mr. Scott Lambert from the FAA Air Route Traffic Control Center in Nashua, New Hampshire, for providing access to the FAA facilities, allowing us to document aircraft observations in real time. The aircraft data obtained were invaluable to this study. We also thank Dr. Andrew Detwiler for contributing his vast knowledge of contrails, including both the production of contrails by jet aircraft and the cloud physics of contrails, to this study. The comments of Dr. Frank Ruggiero in reviewing the original manuscript are also appreciated.

REFERENCES

Afifi, A. A., and V. Clark, 1990: *Computer-Aided Multivariate Analysis.* Van Nostrand Reinhold, 505 pp.

Appleman, H. S., 1953: The formation of exhaust condensation trails by jet aircraft. *Bull. Amer. Meteor. Soc.,* **34,** 14–20.

Bjornson, B. M., 1992: SAC contrail formation study. USAFETAC/PR-92/003, 40 pp. [Available from USAF Environmental Technical Applications Center, Scott AFB, IL 62225.]

Detwiler, A., and R. Pratt, 1984: Clear-air seeding: Opportunities and strategies. *J. Wea. Modif.,* **16,** 46–60.

Donaldson, R. J., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluation of techniques for predicting severe weather events. Preprints, *Ninth Conf. on Severe Local Storms,* Norman, OK, Amer. Meteor. Soc., 321–326.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

Hanssen, A. W., and W. J. A. Kuipers, 1968: On the relationship between the frequency of rain and various meteorology parameters. *Koninklijk Nederlands Meteorologisch Institute,* Med Ed. Verhand., **81,** 2–15.

McCormick, M. P., and E. W. Chou, 1994: Climatology of water vapor in the upper troposphere and lower stratosphere determined from SAGE II observations. Preprints, *Fifth Conf. on Global Change Studies,* Nashville, TN, Amer. Meteor. Soc., 213–215.

Neter, J., W. Wasserman, and M. H. Kutner, 1989: *Applied Linear Regression Models.* Irwin, 661 pp.

Norquist, D. C., H. S. Muench, D. Aiken, and D. C. Hahn, 1997: Cloud cover predictions diagnosed from global numerical weather prediction model forecasts. PL-TR-97-2015, ERP 1203, 138 pp. [Available from Air Force Research Laboratory, Hanscom AFB, MA 01731.]

Peters, J. L., 1993: New techniques for contrail forecasting. AWS/TR-93/001, 26 pp. [Available from HQ Air Weather Service, Scott AFB, IL 62225.]

Schrader, M. L., 1997: Calculations of aircraft contrail formation critical temperatures. *J. Appl. Meteor.,* **36,** 1725–1729.

Shull, J. D., 1998: A validation study of the Air Force Weather Agency (AFWA) JETRAX contrail forecast algorithm. M.S. thesis, Dept. of Meteorology, Air Force Institute of Technology, 120 pp. [Available from Air Force Institute of Technology; Wright Patterson AFB, OH 45433.]

Speltz, D. J., 1995: Validation of the Appleman contrail forecasting scheme using engine-specific aircraft data. Preprints, *Cloud Impacts on DOD Operations and Systems Conf. 95,* Hanscom AFB, MA, Air Force Research Laboratory, 169–171.

Travis, D. J., A. M. Carleton, and S. A. Changnon, 1997: An empirical model to predict widespread occurrences of contrails. *J. Appl. Meteor.,* **36,** 1211–1220.

Wade, C. G., 1994: An evaluation of problems affecting the measurement of low relative humidity on the United States radiosonde. *J. Atmos. Oceanic Technol.,* **11,** 687–700.

Walters, M. K., and J. D. Shull, 1999: A validation study of the Air Force Weather Agency (AFWA) JETRAX contrail forecast algorithm. Preprints, *Eighth Conf. on Aviation, Range, and Aerospace Meteorology,* Dallas, TX, Amer. Meteor. Soc., 591–594.