The Pennsylvania State University

The Graduate School

Department of Geography

A COMPOSITE-BASED RETRO-PREDICTION METHOD FOR

JET CONTRAIL OUTBREAKS OVER THE UNITED STATES

A Thesis in

Geography

by

Armand D. Silva

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2009

The thesis of Armand Silva was reviewed and approved* by the following:

Andrew Carleton
Professor of Geography
Thesis Advisor

Brent Yarnal
Professor of Geography & Associate Department Head

Karl Zimmerer
Professor of Geography
Head of the Department of Geography

*Signatures are on file in the Graduate School

ABSTRACT

The cirrus-level contrails (condensation trails) produced by jet aircraft likely have

contributed to recent climate changes on regional and sub-regional scales in the U.S. and

Europe. Accordingly, there is growing concern about reducing the impacts of contrails

on climate, especially the surface temperature, through improved forecasting—in real

time—of when and where they are most likely to occur. This research develops a

climatology of upper troposphere (UT) meteorological conditions associated with

multiple occurrences of contrails, or outbreaks, for sub-regions of the U.S. in the mid-

season months of 2000-2002. The climatology consists of composites (i.e., multi-case

averages) of UT variables developed using the National Centers for Environmental

Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis data,

and is the first step in designing and verifying a retro-prediction method for contrail

outbreaks. For mapping the UT climatology of contrail outbreaks, the study develops an

objective (GIS-based) regionalization of the conterminous U.S. from the overlaps of

outbreak areas in the 2000-2002 period. The regionalization reveals that the high-

frequency areas of contrail outbreaks vary spatially according to mid-season month,

although the Midwest U.S. has the maximum frequencies on average for the year. The

UT composites of meteorological variables indicate that contrail outbreaks tend to occur

in advance (to the east) of baroclinic weather systems (troughs, fronts, jet streams), which

have associated upward vertical motion, moistening of the air, a thicker upper

troposphere (i.e., higher and colder tropopause), and horizontal wind shear. However,

statistical analyses (e.g., contingency, logit modeling) of UT conditions associated with

contrail outbreaks, reveal that the utility of particular meteorological variables and their

associated map characteristics (magnitude, pattern, gradient) in retro-predicting outbreaks

for the 2000-2002 study period differs by sub-region and mid-season month. Using the

statistical model results of which UT variables are the best retro-predictors for each sub-

region and mid-season month, the research conducts a verification study that involves

retro-predicting contrail outbreaks for July and October in 2008, and January and April

2009. The results of the verification study are mostly positive. They reveal that a

relatively simple map-based method of retro-predicting contrail outbreaks is successful in

certain U.S. sub-regions and mid-season months (e.g., the Central region in April).

Further verification studies are needed to refine the method and make it suitable for

forecasting contrail outbreaks.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Chapter 1

Introduction

The cirrus-level clouds produced by the engines of jet airplanes indicate a direct anthropogenic influence on the upper-troposphere (UT). These aviation-generated condensation trails (contrails) can occur singly but more often are evident as clusters – outbreaks – over regions characterized by a high density of commercial aircraft (e.g., Europe, the U.S.). Contrails and contrail outbreaks are becoming an important issue in climate change studies at regional and sub-regional-scales, and potentially also globally, because of their demonstrated role in reducing the diurnal temperature range (DTR) at Earth's surface. Obtaining improved knowledge of the synoptic climatology of contrail outbreaks (i.e., their regional and sub-regional scale occurrences on monthly to seasonal time scales) serves to both better understand this important anthropogenic forcing on climate, and provide a basis for improving the prediction of outbreaks in real time so that their strongest effects potentially may be reduced. The latter is becoming an important policy issue for some governments in Europe, because contrail outbreaks may intensify the effects of global warming (e.g., Williams et al. 2003). It is anticipated that the results of the present study will contribute to a similar debate in the U.S., involving the Federal Aviation Administration (FAA) and commercial airlines.

Contrails are ice crystal clouds at cirrus altitudes produced by the emission of water vapor and particulates from the engine exhaust of jet aircraft (Travis et al. 1997,

DeGrand et al. 1999), and comprise both the visible condensate of these aircraft

emissions and sublimated ambient moisture (DeGrand et al. 1999, Travis et al. 2006,

Carleton et al. 2008). Given their composition and altitude, contrails have characteristics

broadly similar to natural cirrus clouds, although they contain a higher density of smaller

ice crystals, especially soon after formation. Accordingly, contrail effects on the

shortwave and longwave radiation streams both are enhanced relative to clear-sky

conditions. The ability of contrails to reflect incoming solar radiation, which decreases

the surface insolation receipt, potentially cools the surface in the daytime (DeGrand et al.

2000, Travis et al. 2004). However, contrails may increase the surface temperature at

night relative to clear-sky conditions because of the cloud greenhouse effect (Travis et al.

2004). This combination of daytime cooling and nighttime warming, or contrail "cloud

forcing," decreases the DTR (Travis et al. 2002, 2004). The reduced DTR due to contrail

outbreaks is broadly similar to some other human impacts on near-surface climate at

regional and sub-regional scales (e.g., urbanization, irrigation).

Figure 1-1 shows a satellite IR image taken on a day in September 1995 over the

Midwest U.S., and shows a widespread contrail outbreak composed of numerous

individual contrails. The analysis of many such high-resolution satellite images

(DeGrand et al. 2000; Travis et al. 2007; Carleton et al. 2008) shows that contrail

incidence over the U.S. typically is highest during the transition seasons and lowest

during the summer months. Winter tends to have contrail frequencies in between these

maxima and minima. The seasonal variation in contrail occurrence coincides with the

latitudinal migration of the polar front jet stream. The transition seasons typically are

associated with the strongest horizontal temperature gradients, and therefore steepest

pressure/height gradients, over the U.S., and thereby produce the greatest baroclinicity.

Comparing the extreme seasons, winter has the higher incidence of contrail formation

because of the strong temperature/height gradients concentrated over the southern states

and the Gulf of Mexico, whereas in summer the baroclinicity associated with the jet

stream moves into southern Canada (DeGrand et al. 2000, Travis et al. 2006, Carleton et

al. 2008). At that time of year, contrails and contrail outbreaks over the U.S. mostly

occur in association with the cirrus anvils associated with deep convection.

Figure 1-1: AVHRR thermal IR image of contrail outbreak over Midwest U.S. on September 15, 1995. From Travis et al. (2004).

Although contrail outbreaks occur in every U.S. region, climatologically the Midwest has the highest frequency of occurrence (DeGrand et al. 2000; Carleton et al. 2008). This results from the combination of frequent favorable meteorological conditions associated with contrail formation in the UT, and the region's location beneath the dominant transcontinental flight corridor connecting major East Coast and West Coast cities (Changnon 1981; DeGrand et al. 2000; Carleton et al. 2008). The U.S. Northeast, Southeast and Northwest regions all have relatively high frequencies of contrail outbreaks, but the importance of favorable meteorological conditions in the UT seems to differ between regions; that is, there is an apparent geography to contrail outbreak occurrence (DeGrand et al. 2000, Carleton et al. 2008). This outbreak geography and its associated UT climatological conditions require clarification and characterization.

Contrail Formation

The pioneering study of contrail formation (Appleman 1953) utilized three assumptions for necessary conditions in the UT: (1) the saturation of water vapor with respect to ice; (2) the immediate freezing of water droplets after formation; and (3) an ice crystal content value of 0.004 g m^{-3}. These assumptions led Appleman to create the first conceptual forecast model of contrail formation using measurements of UT conditions as the predictive variables; specifically, the ambient pressure, temperature, and relative humidity (Hanson and Hanson 1995). Although this "model" has been refined subsequently (e.g., Jiusto and Pilie 1964; Schrader et al. 1997), Appleman's study was

important in that it provided the basis for most, if not all, current contrail prediction studies.

Clearly, the ability to predict when and where contrails and contrail outbreaks will occur requires an understanding of the UT meteorological conditions in conjunction with the presence of jet aircraft. As exhaust particles emitted by jet engines mix with the ambient air, the decrease in a wake parcel's absolute humidity is directly proportional to its decrease in temperature (Schrader 1997). In Figure 1-2, the straight line on the graph represents the jet engine emission and has a slope equal to the ratio of water vapor added by the engine exhaust to the increase in temperature from the heat added to the air parcel by the engine. This ratio is defined as the engine contrail factor, which typically has been found to range from 0.0295 g kg^{-1} K^{-1} to 0.049 g kg^{-1} K^{-1} (Schrader 1997, p.1725). Contrail formation in a sub-saturated environment is generally dependent upon the ambient temperature. Supersaturated environments are rare, but similarly depend upon ambient temperature for contrail formation (DeGrand et al. 2000). When the air is supersaturated with respect to ice, contrails can form if T < -39ºC. Although ice supersaturation typically requires RH > 150 percent for natural cirrus formation, contrail cirrus (i.e., cirrus and cirrostratus resulting from the merger of contrails in an outbreak) requires low humidity values because aircraft exhaust temporarily raises the associated local values of relative humidity with respect to liquid water to above 100 percent (Minnis et al. 2004). These UT conditions (high humidity, low temperatures) are most often encountered in the UT over Earth's middle latitude regions. Commonly associated with high humidity is upward vertical motion, due to the fact that rising (sinking) air typically results in moistening (drying) due to adiabatic expansion (compression).

Figure 1-2: Plot of temperature vs. vapor pressure. Saturation vapor pressure with respect to water represented by upper dashed line. Saturation vapor pressure with respect to ice represented by lower dashed line. Straight line represents path of aircraft wake parcel, emitted at highest values of temperature and vapor pressure and decreasing in both over time. From Schrader (1997).

Previous studies that have examined UT contributions to contrail formation have focused on the determination of the critical air temperature, or the temperature at which the existing ambient air mass is saturated for water vapor, thereby promoting contrail formation with any additional water emitted from the jet engines (Hanson and Hanson 1995). Derivation of the critical temperature is achieved by determining a line with a slope equal to the contrail factor, and the tangent point to the saturation vapor pressure curve (Figure 1-2). This derivation is based on the Goff-Gratch formulation, one of the primary saturation vapor pressure formulations used for computing water vapor pressure over liquid water and ice (Volmel 2006).

Contrail Persistence

Once a jet contrail forms, an additional set of favorable atmospheric conditions is required for persistence on time scales longer than a few minutes. Contrail persistence is important climatically, and also for their detection using satellite data. Although fresh contrails are detectable in satellite images (e.g., Figure 1-1), older contrails may become indistinguishable from natural cirrus clouds as the ice crystals comprising them disperse over wider areas. The latter process is accompanied by an increase in cloud particle sizes from their previously small (<10 μm) radius, and by a reduced particle density per unit volume (DeGrand et al. 2000) that may be aided by the influence of strong horizontal wind shear. The vertical thinning and horizontal spreading that accompanies persisting contrails tends to occur most frequently in association with baroclinic waves and the jet stream (DeGrand et al. 2000). As a result, contrails often also occur in conjunction with natural cirrus clouds (Carleton et al. 1986, Travis et al. 2006). Travis et al. (2006) found that the greatest contrail frequency increases coincided with the greatest decreases of tropopause temperature – and increases in tropopause height – in the eastern half of the U.S., occurring since about the mid-1970s. Coincident with the tropopause rising and cooling between the 1970s and today has been an increased average thickness associated with a warmer troposphere.

For the mid-season months (January, April, July, October) of 1977-79, DeGrand et al (2000) determined the associations of individual contrails over the conterminous U.S. with the synoptic-scale circulation using a classification scheme. This scheme was based upon one developed originally by Carleton and Lamb (1986), and distinguishes

between baroclinic and barotropic, or equivalent barotropic, circulation types. It was found that 75 percent of contrails occur in association with natural cirrus, 38 percent of which are associated with the jet stream. The present study expands upon such studies by examining contrail outbreak associations with UT variables. Because outbreaks occur on sub-regional scales (Carleton et al. 2008), a composite (i.e., map averaging technique) is suitable for determining outbreak associations with UT meteorological variables, as well as the relationships between these variables. From the foregoing discussion, the UT variables that can be used to characterize outbreaks are air temperature (T), relative humidity (RH), vertical motion, or omega (ω), and zonal (i.e., west-east component) of the total wind (U).

Study Objectives

There are two main objectives of this research:

1.   To provide reliable "climate diagnostics" (i.e., typically occurring UT meteorological conditions) of contrail outbreak atmospheric environments for those sub-regions of the conterminous U.S. having high frequencies of outbreaks in a given mid-season month (January, April, July, October);

2.   To develop and apply such a climatology-based method that can reliably predict contrail outbreak occurrences in those sub-regions of high frequency, and that can potentially be used in near-real time by the FAA and U.S. airlines.

To achieve Objective 1, I determine the geographical and seasonal dependencies of UT meteorological conditions associated with contrail outbreaks for the mid-season months of 2000-2002, using a composite (i.e., multi-case average) approach applied to reanalysis data on UT meteorological variables from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR). The contrail outbreak inventory for 2000-2002 was developed by Travis et al. (2007) and used by them to determine temporal changes in contrail occurrence between the 1970s and early 2000s. Although climate diagnostics of contrail outbreaks were determined for the 2000-2002 period by Carleton et al. (2008), these were for U.S. sub-regions identified subjectively on the basis of previous contrail studies (e.g., DeGrand et al. 2000), and for all mid-season months combined. In the present study, I use a GIS to objectively determine U.S. sub-regions having high frequencies of contrail outbreaks separately for each mid-season month. The resulting mid-season month and sub-region specific climate diagnostics (i.e., UT meteorological composites) accompanying outbreaks provide a basis for ultimately predicting contrail outbreaks on daily and sub-daily timescales (Objective 2).

To help realize Objective 2, I use the UT composite information derived in Objective 1 to develop a simple visual-rules based map method of identifying contrail outbreak-favorable areas on daily-averaged meteorological analyses. The UT variable(s) that are evaluated for a given region and mid-season month are determined from statistical analyses (e.g., logit modeling) for the 2000-2002 period, and the retro-

predictions[1] are verified by comparison with satellite-image analyses of outbreaks in the high-frequency regions for July and October 2008, and January and April 2009.

The remainder of this thesis is divided into four chapters. Chapter 2 discusses the data and methods used to develop the outbreak regionalization, UT composite analysis, and statistical analysis of the 2000-2002 high-frequency regions. Chapter 3 gives the results of the analyses performed using the data and methods described in Chapter 2. Discussion of the results and their significance in context of the study objectives are also included in Chapter 3. Chapter 4 discusses the retro-prediction technique applied to mid-season months for 2008-09, as well as the results and significance of this application. Finally, Chapter 5 summarizes the study results, and identifies future research that should be undertaken to help achieve reliable short-term (6-36 hr) forecasting of contrail outbreaks for the conterminous U.S.

---

[1] Backwards prediction of an event that has occurred in the past

Chapter 2

Data and Methods for Contrail Outbreak Climatology

This chapter examines the data and methods used in this study. It is divided into four sections. The first one presents the methodology used to develop the contrail outbreak regions. The next section describes the data and methods used to create composite maps of UT meteorological conditions associated with outbreaks. The third section discusses the four statistical measures used to determine the skill of retro-predicting the occurrence of contrail outbreaks. The concluding section covers the methods used to apply the retro-prediction techniques to an independent dataset and therefore to confirm the validity of the analysis.

Regionalization Methodology

Overview

Although previous studies have mapped frequency data on contrails at high spatial resolution (1°x1^{o}) according to mid-season months (e.g., DeGrand et al. 2000; Travis et al. 2004, 2007), this study is the first to develop an objective regionalization of contrail outbreaks for the conterminous U.S. in those months. This approach differs substantially from the regionalization method utilized in Carleton et al. (2008), for example, which developed a set of 10 regions for the U.S. based on those prior spatial

climatologies of contrails. In the Carleton et al. (2008) study, the regions are fixed,

meaning they do not change by mid-season month, and generally follow the boundaries

of surrounding groups of states. The current study maps the outbreak data for each mid-

season month (January, April, July, October), and creates regions based on where the

highest frequencies of outbreaks occur for similar months in 2000-2002. Regions are

determined by those grid cells having the greatest number of overlaps of coordinate

boxes (latitude/longitude) enclosing outbreaks. The result is a regionalization unique to

each mid-season month.

GIS analysis

Contrail outbreak box coordinates (upper left latitude and longitude, lower right

latitude and longitude), approximated to the nearest degree, were acquired for the 2000-

2002 mid-season months from Dr. David Travis at the University of Wisconsin –

Whitewater. These coordinates are for all outbreaks identified on the AVHRR and

MODIS IR imagery in those months. The latitude/longitude data were entered into a

Microsoft Excel spreadsheet, which facilitated the creation of polygons representing

outbreaks in ArcGIS (below). Figure 2-1 shows a sample of the arrangement of box

coordinate data in Excel. A total of four spreadsheets of box coordinate points were

created, one for each mid-season month (all 3 years included).

	1	2	3	4	5	6	7	8
1	Label	Lon	Longitude	Latitude				
2	1	76	-76	36	<--2000			
3	1	71	-71	34				
4	2	93	-93	41				
5	2	83	-83	39				
6	3	90	-90	40				
7	3	82	-82	38				
8	4	88	-88	32				
9	4	82	-82	29				
10	5	96	-96	33				
11	5	86	-86	31				
12	6	113	-113	38				
13	6	109	-109	34				
14	7	96	-96	33				
15	7	89	-89	29				
16	8	102	-102	32				
17	8	95	-95	30				
18	9	92	-92	33				
19	9	82	-82	30				
20	10	93	-93	33				
21	10	82	-82	30				
22	11	87	-87	30	<--2001			
23	11	84	-84	28				
24	12	101	-101	38				
25	12	96	-96	37				
26	13	90	-90	35				
27	13	83	-83	33				
28	14	83	-83	35				
29	14	81	-81	34				
30	15	83	-83	34				
31	15	75	-75	33				
32	16	130	-130	44				
33	16	117	-117	43				

Figure 2-1: Section of a sample Excel spreadsheet with box coordinate pairs. Each pair has a unique label number, the first of each pair being the upper left coordinate, and the second being the lower right. Longitudes are multiplied by -1 because they lie within the Western Hemisphere.

To spatially visualize outbreak occurrences in 2000-2002, a shapefile of the U.S. was acquired from the 2007 TIGER/Line Shapefiles download utility provided by the U.S. Census Bureau. After importing the shapefile into ArcMap, the 48 contiguous states and the District of Columbia were selected for the creation of a new data layer that would be utilized as the base map for the analysis. The next step required the overlay of a 1° by 1° grid onto the USA shapefile. The Hawth's Analysis Tools application was used to

produce this grid. Once the grid was created and displayed over the map, the contrail

outbreak box coordinate data were added, with the spatial reference set to North

American Datum of 1983 (NAD83). The result was a series of points at various grid

intersections, with every pair of points labeled with a common number for ease of

identification when creating outbreak extent polygons (below).

To create rectangular polygons representing contrail outbreak locations and

extents, a new (empty) shapefile was made in ArcCatalog and added to ArcMap. The

result of this procedure was a shapefile with a series of polygons across various places on

the USA map for which outbreaks occurred, with multiple areas of overlap being

identified (Figure 2-2). In total, four shapefiles were created, one for each mid-season

month

Figure 2-2: Example of GIS shapefile consisting of multiple polygons, each representing the spatial extent of a contrail outbreak in the period 2000-2002. The polygon boundaries align with the 1º latitude by 1º longitude grid.

The grid covering the entire U.S. was set as the zonal polygon layer, and the shapefile of the outbreak extent rectangles was set as the summary polygon layer to be summed within the zonal polygon layer. This procedure allowed for a value to be given to each grid cell, based on the number of overlaps of contrail outbreak polygons in that cell. The count of outbreaks within each grid cell was created as a new attribute in the grid's attribute table. To enhance the visual display, a green-to-red color scheme was set to represent low-to-high values of outbreak frequencies: green represents low numbers of outbreaks, red represents high numbers. Uncolored (white) grid cells had no outbreaks observed for that mid-season month during the 2000-2002 study period.

Experimentation with coarser spatial resolution

Although the aforementioned coordinate box overlay analysis provides insights into the mid-season month spatial climatology of contrail outbreaks for the conterminous U.S., some subjectivity remains in terms of designating sub-regions visually. To test whether sub-regions could be more readily identified if the mapped overlay data were displayed at a coarser resolution, a filtering method in ArcGIS was applied to each of the mid-season summary maps to create larger cells from the original 1° by 1° data, (i.e., a spatial smoothing procedure comparable to the running mean in time series analysis). After conversion to raster format, the polygon of the 1° by 1° grid was altered using the Aggregate tool in Spatial Analyst, in which the mean of 9 grid cells was calculated and displayed accordingly to its value in a new grid cell covering a 3° latitude by 3° longitude area. An example of the resulting new map for the month of April is shown in Figure 2-3.

Figure 2-3: April 2000-2002 contrail outbreak occurrences showing filtered (spatial mean) values across a 3° latitude by 3° longitude grid

Comparing the original 1º by 1º resolution overlay maps to those at the 3º by 3º resolution reveals a broadly similar spatial pattern of outbreaks results. Furthermore, no clear pattern emerges as to where precisely the boundaries of the sub-regions having the most frequent outbreaks can be drawn. For this reason, the sub-region boundaries were determined using the original 1º by 1º resolution overlay maps. This has the advantage that inter-regional and inter-seasonal variances of outbreaks could differ by single degrees of latitude and longitude rather than by increments of three degrees.

Designation of sub-regions

For ease of analysis, and given that the NCEP-NCAR reanalysis data on UT variables are to be composited, sub-regions were designated as rectangular areas. Furthermore, sub-region boundaries are set at whole degrees of latitude and longitude. Attempts were made to include all grid cells that showed three or more outbreak occurrences for a particular 2000-2002 mid-season month. On each mid-season map, sub-region boundaries were drawn to include those multi-state areas comprising an outbreak overlap cluster. Accordingly, rectangular areas were large, each encompassing about 30 1º by 1º grid cells or more. It is worth noting that some high-frequency areas within a given sub-region may overlap. The emphasis on identifying those sub-regions having higher frequencies of outbreak overlaps in the 2000-2002 period means that the entire country has not been regionalized (cf. Carleton et al. 2008). This regionalization identified those U.S. sub-regions that are the most susceptible to contrail outbreaks, and for which prediction of outbreak occurrence would be most meaningful climatically.

Although the regionalization was a somewhat subjective process, and the exact boundaries of sub-regions could have been designated differently in size and number on each mid-season map, the overall subjectivity was reduced compared to previous studies by imposing sub-region boundaries after objectively determining a contrail outbreak climatology from overlapping coordinate boxes.

The above procedure to determine outbreak high-frequency sub-regions in the 2000-2002 study period, and for which the NCEP-NCAR reanalysis data are composited to generate UT climate diagnostics (below), resulted in the following sub-region locations by mid-season months. For January, three sub-regions are identified: Midwest (100ºW-80ºW, 37ºN-45ºN); Southeast (86ºW-74ºW, 32ºN-38ºN); and South (98ºW-78ºW, 25ºN-35ºN). April is the only month for which six sub-regions were designated, due to the high numbers of outbreaks covering much of the U.S.: Central (98ºW-84ºW, 35ºN-46ºN); Northeast (85ºW-71ºW, 37ºN-44ºN); South (91ºW-78ºW, 30ºN-36ºN); Upper South (90ºW-81ºW, 34ºN-39ºN); Pacific North (125ºW-116º, 40ºN-47ºN); and Pacific South (124ºW-113ºW, 32ºN-39ºN). In July, three high-frequency sub-regions were identified: Northeast (85ºW-69ºW, 40ºN-47ºN); East (88ºW-73ºW, 32ºN-42ºN); and West (125ºW-111ºW, 35ºN-45ºN). For October, three sub-regions were designated: Midwest/Upper South (96ºW-79ºW, 35ºN-46ºN); South (91ºW-81ºW, 28ºN-36ºN); and Northwest (121ºW-112ºW, 42ºN-48ºN). The spatial extents of these sub-regions are visualized in Figures 2-4a-d.

January

Figure 2-4a: Map showing high-frequency sub-regions of contrail outbreaks for January 2000-2002

April

Figure 2-4b: Map showing high-frequency sub-regions of contrail outbreaks for April 2000-2002

July

Figure 2-4c: Map showing high-frequency sub-regions of contrail outbreaks for July 2000-2002

October

Figure 2-4d: Map showing high-frequency sub-regions of contrail outbreaks for October 2000-2002

As indicated above, outbreak sub-regions are dynamic in that they change size and location between mid-season months. Even those regions with the same name in any two mid-season months (e.g., the Northeast in April and July) have slightly different spatial coverages. This strategy has implications for the analysis of NCEP-NCAR data: because each sub-region pertains to a particular mid-season month, daily-averaged data for that month can be composited rather than daily anomalies from the long-term monthly means.

Composite Methodology Applied to NCEP/NCAR Reanalysis Data

Overview

This section describes the data and methods used to create composite maps of UT meteorological conditions associated with outbreak, non-outbreak, and pre-outbreak days (defined below) for the 2000-2002 study period. These composites comprise the climate diagnostics of contrail outbreaks derived for U.S. sub-regions having high frequencies of outbreaks, and also form the basis for the retro-prediction method.

Data and composites

The meteorological UT composites were developed using the Daily Mean Composites webpage of the Physical Sciences Division, provided by NOAA's Earth System Research Laboratory (http://www.cdc.noaa.gov/data/composites/day). To determine the outbreak dates in each sub-region for each mid-season month, I identified

which polygons—representative of an outbreak on any given day in 2000-2002—fell

both entirely or partially within the sub-region in the 1° by 1° grid map of the

conterminous U.S. Those outbreak polygons lying mostly outside of a sub-region, but

that had a few grid cells located within the sub-region, were included in the composites.

Having compiled a list of outbreak dates in mid-season months of 2000-2002 for each

sub-region, composites for each of the four UT variables (temperature, RH, omega, and

zonal wind) were generated through the PSD website. Each region/mid-season month

had a unique set of individual outbreak days comprising its composites of the UT

variables. Each composite was downloaded in NetCDF format to permit further

manipulation using the Grid Analysis and Display System (GrADS) software. GrADS

can display the differences between two input NetCDF files, as well as display two or

more variables per map, which is useful for showing spatial relationships among

variables. To help consolidate the number of output maps, and to better display

relationships between UT variables, the composite temperature, RH, omega, and zonal

wind appear as a common plot for each sub-region. To facilitate interpretation, and to

permit inter-comparison of composite UT maps between regions, each variable was set to

a common interval for isoline plotting in the GrADS software: temperature at 0.5 °C;

relative humidity at 1.0 percent; omega at 0.01 Pa s^{-1}; and zonal wind at 1.0 m s^{-1}.

To determine the average meteorological conditions in the UT by sub-region in a

given mid-season month (section Designation of sub-regions, above), composite maps

were generated for: (1) outbreak days; (2) non-outbreak days; and (3) pre-outbreak days

(i.e., outbreak day minus 1). In addition, for a given variable, difference maps between

any two composites were generated and visually compared (e.g., outbreak minus non-

outbreak days for the Midwest in April; outbreak minus pre-outbreak for the South in January). For each sub-region and mid-season month, daily-averaged data at 300 hPa on the four variables previously identified as important for contrail outbreak occurrence and persistence (i.e., temperature, relative humidity, omega, and zonal wind) were composited. The 300 hPa level was chosen as the average lower-height representative of aircraft flight level, and was the highest level for which RH data were available. As indicated previously, for each variable and region, the composite mean rather than the composite anomaly was derived because data from different mid-season months were not being combined.

Of the fifteen outbreak high-frequency sub-regions identified earlier, a total of six were chosen for the comparison of composites between outbreak and non-outbreak days using the NCEP-NCAR reanalysis data. Each sub-region selected had the maximum number of outbreaks (2000-2002) in that mid-season month. A secondary consideration in this selection process was to compare several different sub-regions of the country (e.g., the Midwest, the South, the West) in terms of their respective UT meteorological conditions. Because the Midwest generally had the maximum number of outbreaks, this sub-region was analyzed for three of the four mid-season months. Accordingly, the sub-regions and months selected for the composite analysis were: (1) Midwest in January, (2) South in January, (3) Central in April, (4) Pacific South in April, (5) East in July, and (6) Midwest/Upper South in October.

Because of the need to determine whether the outbreak composites for a given region and mid-season month are sufficiently different from the atmospheric conditions on non-outbreak days—and for which UT variables—it is appropriate to develop

"undisturbed" composites based on days representing normal (i.e., non-outbreak)

conditions. For this purpose, non-outbreak dates were determined and compiled into a

separate database. I excluded the dates one day before outbreak dates (i.e., day-1) from

the undisturbed composite database as their associated atmospheric conditions were used

subsequently to evaluate whether there are consistent UT differences between outbreak

days and the days before.   Similarly, of the remaining non-outbreak dates, I stipulated

that no two consecutive days could be used in the composite because of the need to

obtain an independent (i.e., non-temporally correlated) sample database.

After I developed the UT mid-season month composites for non-outbreak days, I

downloaded the respective NetCDF files, opened them in GrADS, and displayed them in

similar map format to the composites for outbreak days. I did this to examine the UT

conditions for "normal" or undisturbed days.

To characterize and compare the UT conditions one day in advance of outbreaks

(i.e., pre-outbreak, or day-1) to the average outbreak and non-outbreak conditions, I

generated composites for the day immediately before each outbreak day (Appendix A).

Similar to the derivation of composites on outbreak and non-outbreak days, I created

these composites for the six regions of maximum frequency for their respective mid-

season month, using the same isoline intervals as those in the outbreak composites.

I also generated difference maps of the composites (i.e., outbreak minus non-

outbreak) using the GrADS software to obtain insights into how distinctive are the

outbreak composites compared to the non-outbreak composites. Just as was done

separately for the outbreak and non-outbreak composites, I created two maps for each of

the six high-frequency regions; one showing both temperature and RH, the other showing

both omega and zonal wind. In the difference maps, positive (negative) values indicate greater magnitudes of a given UT variable on outbreak (non-outbreak) days.

Map criteria for visually determining outbreak likelihood

To determine whether the differences in UT conditions depicted on the composites for outbreak days and non-outbreak days (by sub-region, mid-season month) permit skillful retro-prediction of contrail outbreak occurrences, the following three map criteria were considered: magnitude, pattern, and gradient. Magnitude differences refer to the change in a given UT variable's actual value (e.g., maximum, minimum) within the sub-region under consideration. Pattern differences are those in which the general arrangement of high-to-low values (e.g., location of maxima and minima, orientation of isolines) of a UT variable change across the region, but for which exact magnitudes are not explicitly considered because the magnitudes themselves do not change much between composites. Gradient differences, a form of magnitude difference, involve broadly similar patterns on outbreak and non-outbreak days, but the horizontal change in the values—as given by the isoline spacing—is either tightened or relaxed over the region. These three visual map criteria are not necessarily mutually exclusive; two or all three criteria may change between composites. In many cases, the outbreak minus non-outbreak differences were apparent by visually comparing each composite map side by side. To better distinguish the differences of outbreak days from non-outbreak days for more similar patterns (e.g., very small magnitude changes), the difference maps were utilized to interpret these changes. Although another investigator may not have

interpreted the map differences as I did, I attempted to highlight only the most notable

differences using the three map criteria. Such a simple visual map-based technique has

the advantage of being relatively easily and quickly used, greatly facilitating its potential

application as a contrail prediction method in real-time. There is confidence in this

method because of the statistical analyses that are performed and presented later.

Retro-prediction of contrail outbreaks using NCEP-NCAR daily-averaged reanalyses

To evaluate the visual map-based method of retro-predicting contrail outbreaks

using the magnitude, pattern and gradient characteristics evident on the regional

composites of outbreak versus non-outbreak days, I undertook a day-by-day inspection of

the daily-averaged (i.e., 4x6 hr) UT maps for each variable in all months of the 2000-

2002 study period. The occurrence of a contrail outbreak is a binary phenomenon; that

is, for each UT variable of interest (temperature, RH, omega, and zonal wind), the map

for that day was classified as either favorable or not favorable for a contrail outbreak to

occur according to its respective magnitude, pattern, or gradient criterion. The daily

mean plot of a given variable was deemed favorable for a contrail outbreak if it more

closely resembled the composite map on outbreak days for that region, contrasted with

the non-outbreak composite. For example, the visual assessment for the Midwest sub-

region in January was performed using each day's map of 300 hPa temperature for

January 1, 2000, then for January 2, 2000, and for each succeeding day through January

31, 2000; then for January 1, 2001 through January 31, 2001, and repeated for January 1,

2002 through January 31, 2002. This same process was repeated for RH300, for

omega(300), and for zonal wind at 300 hPa. In total, the January Midwest region had 372 individual interpretations, 93 for each of the four UT variables, for which each variable had three months of 31 days each. To reduce any bias on my part in determining outbreak-favorable conditions, I did not consult the dataset containing the actual outbreak dates, or those that went into the composites of non-outbreak days until after the analysis was completed.

To compare the retro-predicted and actual outbreak dates in the study period, I created an MS Excel spreadsheet with rows representing each day of the month for three years of the same mid-season month (e.g., January 2000, January 2001, January 2002), and columns representing each of the four UT variables. Whenever I deemed a daily-averaged map to be favorable—meaning that it met the criteria mentioned above either completely or very closely—I placed an "X" in the cell corresponding to that day and variable. I left the cell blank when I deemed a map to be unfavorable. Once this process was completed for all four variables in a given region, I highlighted the dates on which outbreaks had occurred, and compared statistically both sets of data (outbreak favorability, actual outbreak occurrence) to determine which UT variable(s) perform best for a given sub-region and mid-season month. The process was completed for all six regions.

Statistical Analysis of Contrail Outbreak Retro-Prediction

This section describes the statistical measures I used to determine which of the four UT variables considered (i.e., temperature, RH, omega, u) show skill in retro-

predicting the occurrence of contrail outbreaks for the 2000-2002 mid-season months and

sub-regions evaluated. For this purpose, I apply four statistical measures of skill that are

used to determine the success rate of meteorological forecasts compared to a verification

analysis (Wilks 1995). Each of these measures has its advantages and disadvantages

when applied to contrail outbreak retro-prediction (see below), so I use all four of them

and discuss their relative merits for determining skill by mid-season month and study

sub-region in the Results section of the thesis. To begin, I created contingency tables for

each UT variable within each region and mid-season month to determine individual

associations of favorability versus occurrence, and from which I computed the Chi-

squared statistics of distribution and success measures. Then, I applied a binary logistic

regression to the outbreak favorability and occurrence data to determine which

variable(s) are significant in terms of their main effects (single variable significance) and

also their two-way interactions of variables (multiplication of two variables). The results

of these statistical analyses for sub-regions and mid-season months guide the retro-

prediction study for recent months not used to develop the 2000-2002 UT composites

(July and October 2008, January and April 2009).

Background

The type of dichotomous event represented by a contrail outbreak lends itself to

verification of categorical forecasts of discrete predictands. A categorical forecast is one

in which only one set of possible events will occur. A discrete predictand is one that

takes just one finite set of values (Wilks 1995, p. 238). Accordingly, I apply binary

forecast—and retro-prediction—verifications to this study, using contingency analysis, four measures of forecast skill, and binary logistic regression (below).

Contingency analysis

First, I analyzed the success of each UT variable in determining outbreak occurrence or non-occurrence for the sub-regions and mid-season months of 2000-2002. Categorical data with discrete predictands can be verified using a contingency table, a plot of intersections of forecast categories and observation categories (Stanski et al. 1989). Here, the favorability of any one UT variable (temperature, RH, omega, and zonal wind) is the forecast category and the occurrence of a contrail outbreak is the observation category. Like outbreak occurrence, the favorability of a UT variable is also binary, meaning the variable is either favorable or not favorable. With two forecast categories (favorability: Yes, No) and two observation categories (occurrence: Yes, No), the contingency table contains four forecast combinations: Hit (both Yes), Miss (Forecast: No, Observation: Yes), False Alarm (Forecast: Yes, Observation: No), and Correct Negative (both No) (Table 2-1). A total of 24 such contingency tables were created, one for each UT variable (4), for each of the six regions of interest.

	Predicted	No
Occurrence	Hit	Miss
No	False Alarm	Correct Negative

Table 2-1: Contingency table with two binary axes and four resultant cells

A Chi-squared (χ^2) test is performed on each of the contingency tables to identify which UT variables for a given sub-region and mid-season month are most significant for predicting contrail outbreaks. The Chi-squared test can be used to determine the validity of categorical variables (here, forecast Yes or No, observation Yes or No) compared to an estimation of what their relationship would be as a function of chance (Stockburger 1996). I computed the Chi-squared statistic for each of the 24 contingency tables:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

where O is the observed value in a given cell in the contingency table, and E is the expected value for that cell.

The E value of a cell frequency is computed:

$$E = \left(\frac{RowTotal * ColumnTotal}{n} \right)$$

where n is the total number of observations; here the number of days (93 for January, July, October; 90 for April).

The degrees of freedom for a Chi-squared test of a contingency table (e.g., Hammond & McCullagh 1978) are calculated as follows:

$$df = (\#Rows - 1) \times (\#Columns - 1)$$

In every contingency table in this study, df = (2 – 1) * (2 – 1) = 1.

At the 0.05 significance level, the critical value of Chi-squared with one degree of freedom is 3.84. The larger a Chi-squared value, the less likely are the cell frequencies to be randomly distributed and, therefore, the less likely that the named categories (here, forecasts and observations, yes and no) are valid or reasonable (Hammond & McCullagh 1978). Thus, a Chi-squared statistic greater than 3.84 indicates a lesser probability of random distribution.

The binary categorical forecasts represented by the 2 x 2 contingency tables can be tested using several skill, or accuracy, measures (Wilks 1995, Chapter 7). These measures determine the success of using any one variable (here, UT map) to successfully retro-predict outbreaks, although each measures the accuracy slightly differently (e.g., the false alarm rate versus the hit rate). The first measure I discuss is the "hit rate" (H), which gives the accuracy of predicting an outbreak occurrence or a non-occurrence. The result is the fraction of correct forecasts. The H is calculated:

$$H = \frac{Hit + CorrectNegative}{n}$$

where n is the number of events; here, the number of days in the three study years (2000-2002) for any common mid-season month (e.g., January 2000, January 2001, January 2002). Thus, for each hit rate calculated for January, July, and October, n = (31 days)*(3 years) = 93. For each H calculated for April, n = (30 days)*(3 years) = 90. The best

possible H (i.e., maximum success) has a value of unity, and the worst possible H (no success) has a value of zero.

A broadly similar measure to H is the "threat score," or critical success index (CSI). The CSI measures the success of only the correctly predicted occurrences (Hit, or both "Yes"), not the correctly predicted non-occurrences (Correct Negative) (Wilks 1995, Chapter 7). The CSI is calculated:

$$CSI = \frac{Hit}{Hit + FalseAlarm + Miss}$$

The best possible CSI (maximum success) has a value of unity, and the worst possible CSI (no success) is zero.

The next accuracy measure I consider is the "probability of detection" (POD). The POD is a fraction representing the accuracy of a forecast, whether predicted (Yes) or not (No), for when the forecast event is observed (Wilks 1995, Chapter 7). When applied here, the POD considers only contrail outbreak days:

$$POD = \frac{Hit}{Hit + Miss}$$

The best possible POD has a value of unity, and the worst possible POD is zero.

The fourth measure of skill I use to determine a UT variable's ability to retro-predict an outbreak (i.e., favorability) is the false alarm rate (FAR). The FAR is the

fraction of forecasts predicting an event but for which the event does not occur (Wilks 1995); that is, FAR measures failure (or a lack of skill) in terms of the number of times a "Yes" prediction corresponds to a "No" observation, and is calculated:

$$FAR = \frac{FalseAlarm}{Hit + FalseAlarm}$$

In contrast to the other success measures just described, the best possible FAR is zero, and the worst possible FAR has a value of unity.

Another useful forecast verification tool measures the bias (B), or the correspondence between the frequency of "Yes" predictions and the frequency of "Yes" observations (Wilks 1995, Chapter 7). The B is a ratio calculated:

$$B = \frac{Hit + FalseAlarm}{Hit + Miss}$$

An unbiased dataset has a B value of unity, because the number of predictions and number of observations are the same. A value of B greater than unity, for which the "Yes" predictions outnumber the observations, is an over-prediction. Conversely, a value less than one, for which the forecasts are fewer than the observations, is an under-prediction.

Binary logistic regression

To verify the statistical significances of UT variables for each sub-region for predicting contrail outbreaks as determined by the contingency analysis, I performed a binary logistic regression on the "favorability" dataset (that which deems daily mean UT maps as being similar to the outbreak composites) to determine which combinations of UT variables (e.g., RH and temperature, RH and omega) were most successful in retro-predicting outbreaks in the high-frequency sub-regions by mid-season months of 2000-2002. Logistic regression estimates the probability of an event's occurrence, and is often used to determine the relationship between discrete binary responses (Yes, No) and a set of explanatory variables (SAS 2009). Forecasts of the probability of occurrence commonly transform the predictand into a binary with a value of zero (not occurring) or one (occurring) (Wilks 1995). Here, the predictand is the occurrence of a contrail outbreak. Binary logistic regression is used specifically when the dependent variable is dichotomous (Yes, No) and the independent variables are of any type (Garson 2009).

I used the SAS statistical software to fit the favorability and outbreak data to a logistic regression model. Because the predictand (outbreak occurrence) and predictor variables (UT temperature, RH, omega, and zonal wind) are binary, they each take on values of zero or one. Five categories were input to the model: Response, T, RH, Omega, and U Wind. The Response category represents contrail outbreak occurrences and is the dependent variable; a value of one represents an occurrence, a value of zero represents no occurrence. Categories T, RH, Omega, and U Wind are the independent variables, with a value of one representing favorable for contrail outbreak occurrences, and zero

representing unfavorable. Each data line, comprising the five categories, corresponded to one day in the dataset, and has five values (1 or 0) representative of the Response and UT variable categories.

I input the above transformed data into the PROC LOGISTIC procedure in SAS, which fits the logistic regression to the data and outputs maximum likelihood estimates for each of the independent (UT) variables. The table "Analysis of Maximum Likelihood Estimates" displays parameter estimates, standard errors, and the results of the Wald Chi-squared test and associated p-values. Each p-value is the probability that the test statistic will be equal to or greater than the actual observed value when the null hypothesis is true (Neter 1985). In this study, the test statistic is the value given by the Wald Chi-squared test. Because I want to reject the null hypothesis that the UT variables are related with outbreak occurrences by a random chance distribution, variables with lower p-values are more significant. A p-value less than the test level of 0.05 means that the null hypothesis can be rejected.

Two-way UT variable interactions

Given that many meteorological variables are inter-correlated (e.g., RH and omega; U and omega), it is possible that more than one UT variable may successfully predict contrail outbreaks in a given sub-region and mid-season month. I indicated earlier that the South sub-region in January had no single UT variable significantly associated with contrail outbreak occurrences. Thus, to confidently predict outbreaks for this region in January using UT variables (temperature, RH, omega, and zonal wind), it is

necessary to determine whether or not favorable conditions occur when considering two or more variables simultaneously. For this purpose, I fitted a binary logistic regression with two-way variable interaction to the favorability and outbreak data. A two-way interaction involves an added term in the logit regression model that is the cross product of the two independent variables (Neter et al. 1985). The general regression model for two separate variables has the form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

where β0, β1, β2 are the regression coefficients. The interaction model for X1 and X2 is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

I computed the interaction model separately for all six regions of maximum outbreak frequency using the PROC LOGISITIC procedure in the SAS software, but with cross product terms—signifying variable interactions—as follows: RH*Omega, RH*T, RH*UWind, Omega*T, Omega*UWind, and T*UWind. The maximum likelihood estimates were output for each of the main effects (i.e., individual) UT variables, as well as for the two-way interaction terms.

Retro-Prediction of Outbreaks for 2008-09

It is important to verify the skill of the UT variables in retro-predicting contrail outbreak occurrences for by testing them on an independent dataset not used to develop the UT composites. For this purpose, I performed a retro-prediction on the six high-frequency sub-regions for four recent mid-season months: the East in July 2008; the Midwest/Upper South in October 2008, the Midwest and the South in January 2009, and the Central and Pacific South in April 2009. The methods of this analysis are described below.

Retro-prediction method

The retro-prediction method was carried out similarly to the statistical analysis of UT variable favorability for 2000-2002. As in that analysis, the map data on UT conditions comprise daily NCEP/NCAR reanalyses for each UT variable (temperature, RH, omega, and zonal wind), and were downloaded from NOAA's CDC website (http://www.cdc.noaa.gov/data/gridded). I imported the NCEP-NCAR data into GrADS for visualization. The method of determining outbreak favorability was the same as I used for the 2000-2002 retro-prediction, except that I inspected only the maps of the UT variable(s) shown by the statistical analysis for that sub-region and mid-season month to be statistically significant.

Retro-prediction verification

To verify the outbreak retro-predictions for sub-regions in mid-season months of 2008 and 2009, I, along with the assistance of Matthew Aghazarian, interpreted visually the multi-band AVHRR images for the U.S. that are available approximately 5-8 times per day. These images were downloaded from the NOAA Satellite and Information Service (http://www.class.ngdc.noaa.gov), and imported into the ERDAS-Imagine software for visualization and manipulation (e.g., band subtraction, zoom, outbreak coordinate extraction). For each image, the presence or absence of a contrail outbreak was determined using the criteria of finding several visible contrails located in close proximity to one another, as described by Carleton and Lamb (1986). For each outbreak, the upper-left and lower-right coordinates (in degrees latitude and longitude) of a bounding box that completely encloses the contrails were determined (cf. Carleton et al. 2008), and recorded in a spreadsheet. These actual observations were then compared to the retro-predictions to determine the success of using significant UT variables from 2000-2002 for 2008-2009. Results of this comparison are described in Chapter 4.

Chapter 3

Results and Discussion of Contrail Outbreak Climatology

In this chapter, I present the results of the analyses described in Chapter 2 and discuss their implications for the real-time prediction of contrail outbreaks. First, I discuss the results of the objective regionalization of contrail outbreaks for mid-season months of 2000-2002. Second, I discuss the results of the composite analysis of UT variables for the six U.S. sub-regions that have high frequencies of outbreaks in the 2000-2002 study period. Third, I present the results of the statistical analyses of UT variables in context of their ability to retro-predict contrail outbreaks for sub-regions and mid-season months in the study period 2000-2002.

Regionalization of the conterminous U.S. by outbreak frequency

The objectively-derived GIS maps of contrail outbreak overlaps for mid-season months (Figures 3-1a-d) reveal the following characteristics. For January (2000-2002), there were relatively few outbreaks compared to the other mid-season months. However, in this month, outbreak occurrences were concentrated primarily over two locations: the Midwest (from the Mississippi River through central Ohio), and the South (from western Louisiana through north central Florida). Specifically in the South, there were two grid cells each having seven occurrences (total for the three January months), the maximum

frequency for this month. Both of these outbreak high-frequency regions display very strong patterns that span many more degrees of longitude than degrees of latitude. The most likely reason for this geographical pattern of contrail outbreaks involves the mean positions of both the subtropical and polar jet streams: in January, both jet streams typically are strongest and at their southernmost positions for the year. The polar front jet frequently is associated with a recurring series of synoptic-scale upper level troughs, and areas of horizontal wind shear and baroclinic instability occur around the mean center of the jet. These areas of baroclinicity increase the likelihood of upward vertical motion, which increases the relative humidity in the UT, in many instances, to that of ice supersaturation, which is are needed for contrail persistence. Reasonably large numbers of outbreak occurrences also occur over the Carolinas, northern Georgia, and eastern Tennessee, a sub-region that falls in between the two maxima latitudinally, and is centered more to the east longitudinally. This positioning could be the result of the southward (northward) oscillations of the polar (subtropical) jet stream. Most areas of the western U.S. and New England did not experience any outbreaks in the January months of 2000-2002, although cases of individual contrails likely occurred, as shown by Travis et al. 2007.

The outbreak summary map for April shows much higher frequencies than that for January, with the maximum of 13 outbreaks total for the three April months located over northwestern Indiana and northeastern Illinois. This concentration represents the highest number of contrail outbreak occurrences in all mid-season months for the 2000-2002 study period. Surrounding this absolute maximum is a larger-scale region comprising nine or more outbreaks in all grid cells of northern Illinois, northern Indiana,

southwestern Michigan, and extreme southern Wisconsin. This distinct Midwest

maximum results from the mean position of the polar front jet stream over the northern

U.S. in the spring (DeGrand et al. 2000). Large numbers of contrail outbreaks also occur

in a sub-region comprising eastern Ohio and western Pennsylvania, and in a separate sub-

region covering Kentucky, Tennessee, and western North Carolina. As was evident in

the outbreak map for January, both of these sub-regions exhibit greater longitudinal than

latitudinal extent. Also in April, the western U.S. sees the highest frequency of contrail

outbreaks of all mid-season months, at least for the 2000-2002 study period. Relative

maxima of contrail outbreaks occur in central California and western Oregon, with

several grid cells showing four to five occurrences total for the three April months. The

moderate numbers of outbreak occurrences in the Pacific coast sub-region is likely due to

increased activity of the jet streams, especially when highly meridional long-wave and

blocking patterns occur.

In July, there are fewer contrail outbreaks than in April, and the spatial coverage

is reduced across the conterminous U.S. However, outbreaks are highly concentrated in

the central Appalachians, including portions of North Carolina, Virginia, West Virginia,

Kentucky, and Tennessee. This pattern is most likely attributed to the onshore

(southeasterly) flow of moisture from the summertime Bermuda high (Mearns et al.

2003), aiding Appalachian-induced orographic convection. Indeed, the synoptic

climatology of individual contrails for 1977-79 presented by DeGrand et al. (2000)

showed that the majority of contrails in July occur in and near the cirriform anvil clouds

resulting from deep convection. Also in July, much of the Northeast sub-region and the

eastern Midwest, have relatively high outbreak frequencies, with multiple grid cells

showing six or more occurrences total, and a majority of grid cells with five or more occurrences. In contrast, most of the Great Plains extending westward to the Pacific Ocean did not have any contrail outbreaks during the July months of 2000-2002, although one sub-region comprising much of northern California and northern Nevada saw two total outbreaks. In general, it can be said of July that although the absolute maximum of contrail outbreaks occurs in the central Appalachians, the mean latitude of outbreak occurrences is at its most northerly in this month. This feature can be seen in the many grid cells of moderate frequencies in the Northeast and the relatively few in the Southeast, and is directly attributable to the northernmost mean position of the polar jet stream during the summer (DeGrand et al. 2000).

The contrail outbreak map for October (2000-2002) shows a pattern broadly similar to that of the other transition-season mid-month, April. However, there are differences of detail between these two months, the most notable being the relative absence of outbreaks in the western U.S. in October compared to April. The states of Illinois, Indiana, and Ohio again comprise the sub-region of highest outbreak frequencies, with the absolute maximum located in eastern Indiana. In this regard, a more minor difference between the distributions in October and April is that high outbreak frequencies do not extend as far east, with the region of relative maximum ending in eastern Ohio rather than in central Pennsylvania in October. Overall, the mean position of the subtropical jet stream (STJ) reflects similar spatial climatologies of outbeaks in October as it does in April.

January

Figure 3-1a: January contrail outbreak occurrences at 1º by 1º resolution for 2000-2002.

April

Figure 3-1b: April contrail outbreak occurrences at 1º by 1º resolution for 2000-2002.

July

Figure 3-1c: July contrail outbreak occurrences at 1º by 1º resolution for 2000-2002.

October

Figure 3-1d: October contrail outbreak occurrences at 1º by 1º resolution for 2000-2002

Composite UT Conditions for Outbreak and Non-Outbreak Days

This section presents the results of the composite analysis of NCEP-NCAR reanalysis UT data for the high-frequency sub-regions, by mid-season month. According to previous studies, a higher and, therefore, colder tropopause is a favorable environment for contrail formation (Travis et al. 2006). Also favorable are an increased thickness (i.e., warming) between geopotential levels in the troposphere (resulting from ridging and subsidence of air), stronger upward vertical motion, and stronger westerly winds (Carleton et al. 2008). Thus, synoptically, warm-cored high pressure and strong meridional gradients of vertical motion and UT humidity have been found to be associated with most contrail outbreaks, although these combinations show some dependence on sub-region within the U.S. (Carleton et al. 2008). The following results verify these previous findings.

For the Midwest sub-region in January (Figure 3-2a), there are small magnitude differences in 300 hPa temperature (T300) between the composite map of outbreak days and that of non-outbreak days. The largest temperature difference is 1ºC in the far northwest corner of the sub-region. In terms of the relative humidity (RH300), the magnitude differences are largest over Michigan and the Great Lakes region, with some areas having an average increase of 9 percent or more on composite outbreak days compared with non-outbreak days. For omega (Figure 3-2b), there is a distinct concentric pattern of ascending motion over much of the sub-region on outbreak days, with a maximum located over northern Missouri. The outbreak composite map of 300 mb zonal wind has a much more meridional pattern of isolines contrasted with non-

outbreak days, and these decrease in magnitude from east to west. Collectively, these composites for the Midwest in January suggest a mean troposopheric ridge in the eastern part of the sub-region on outbreak days, when zonal winds are also strongest. The west side of an upper-air ridge is a key area for warm air advection and positive vorticity advection, and therefore for upper-level divergence needed for the persistence of contrails in baroclinic conditions (e.g., DeGrand et al. 2000).

Also in January, only for the South sub-region (Figure 3-3a), the composite map of T300 on contrail outbreak days reveals relatively small temperature differences. These are largest (greater than 1ºC) over the area encompassing South Carolina, Georgia, and eastern Alabama. The composites signify cooler air in the UT on outbreak days in this area. For RH300, both the composite map on outbreak days and the difference map (outbreak composite minus non-outbreak composite), show a distinct pattern reversal such that outbreak days have moister air over the Gulf of Mexico (particularly in the western Gulf) and drier air inland, while non-outbreak days have the moisture maximum located well inland, with drier air over the Gulf. For UT omega on outbreak days (Figure 3-3b), a meridional gradient of decreasing positive values (i.e., reduced subsidence) from east to west is evident, which is distinctly different from the pattern on non-outbreak days. The map of zonal wind at 300 hPa shows stronger westerlies nearly everywhere in the sub-region on outbreak days, particularly over peninsular Florida. Considered together, the UT composites suggest that outbreaks for the South sub-region in January are associated with the west side of an upper-level ridge. The subsidence decreases and relative humidities increase (i.e., greater instability) on moving further westward. Moreover, contrail outbreaks show some association with the strong winds associated

with the STJ, especially over Florida. The latter is consistent with the average location of the STJ in January, and with the synoptic climatology results for this mid-season month described in DeGrand et al. (2000).

For the month of April in the Central sub-region (Figure 3-4a), large UT temperature differences between outbreak and non-outbreak days occur in the eastern and southeastern areas, amounting to 1.5ºC cooler on outbreak days. The composite plot of RH300 on outbreak days shows a clear pattern of drier air in the south and moister air to the north, as well as stronger gradients than those appearing on the composite map of non-outbreak days. For omega (Figure 3-4b), a strong difference is evident between outbreak and non-outbreak days: air is rising on non-outbreak days, and sinking on outbreak days. This result seems contradictory to what would be expected for contrail persistence. The composite maps of UT zonal wind depict substantially weaker westerlies across the region on outbreak days compared to non-outbreak days. This result suggests either a weak total wind (i.e., vector wind) or considerable meridional (north-south) flow in the UT when outbreaks are present, although the latter scenario is more likely due to its association with baroclinic activity. Given the large RH gradients on outbreak days, that UT humidity increases northward, and that subsidence also increases northward, it seems likely that the Central sub-region experiences increased instability when a more northerly flow is present. Such a pattern of UT moisture, vertical motion of air, and wind is indicative of the western side of a trough. Indeed, both Changnon et al. (1981) and DeGrand et al. (2000) show that individual contrails can occur in and west of a trough. The present results indicate that multiple occurrences of contrails (i.e., outbreaks) also occur in association with troughs, at least in the Central U.S. sub-region.

For the Pacific South sub-region in April (Figure 3-5a), UT temperatures are higher on outbreak days than non-outbreak days, particularly over the inland desert areas, with the northwestern portion of the sub-region showing an average increase of 1.8°C and higher. The outbreak composite of RH300 shows a greatly strengthened west-east gradient relative to that on non-outbreak days, comprising an average minimum of 26 percent located in the east to a maximum of 53 percent in the west. The composite patterns of omega (Figure 3-5b) also are strongly different between outbreak and non-outbreak days for this sub-region: on outbreak days there is strong subsidence of air in eastern areas but strong ascendance in the northwest. The composites of UT zonal wind for outbreak and non-outbreak days show broadly similar magnitudes and gradients; however, there is a clear pattern reversal on outbreak days, when magnitudes generally decrease to the southward. A distinctive feature of the outbreak composite of UT zonal wind for the Pacific South sub-region in April is the "col" or "saddle point" pattern occurring over the western portion. All four of the UT composites on outbreak days seem to imply weak stability, particularly as given by the dramatic increase in RH and upward vertical motion. Furthermore, the reversal in the pattern of zonal wind strength between north and south on outbreak versus non-outbreak days suggests an upper-level high located to the east and an upper-level low to the westward, yielding a region of more meridional flow and enhanced upper-level divergence on outbreak days.

In July, the U.S. sub-region of maximum contrail outbreak frequencies is the East. In that sub-region, the composite map of UT temperature (Figure 3-6a) shows generally small differences between outbreak and non-outbreak days, although the extreme northeastern portion sees an average decrease of 1.5°C on outbreak days. The composite

map of RH300 on outbreak days shows a distinct maximum of moisture located over eastern Tennessee, but drier air elsewhere, particularly off the coast of South Carolina and over Lake Erie. The UT composite omega map for outbreak days (Figure 3-6b) shows a greater range of values across the region, and therefore a stronger gradient, contrasted with non-outbreak days (Figure 3-12b). In general, positive values of UT omega (subsidence) occur in the northwest, with negative values (ascendance) in the southeast of the sub-region. The composite map of UT zonal winds for outbreak days shows weaker westerlies in the far northwestern corner of the region, but stronger westerlies over the Atlantic, east of North Carolina's Outer Banks. These patterns may suggest that the mean center of the STJ lies to the north of the region during this month.

In October, for the Midwest/Upper South sub-region (Figure 3-7a), the horizontal gradients of UT temperature are highly different between outbreak and non-outbreak days: weaker (stronger) on outbreak (non-outbreak) days, and with a north-south difference of about 5ºC (7.5ºC). The RH300 on both the outbreak and non-outbreak composite plots have similar magnitudes but with different patterns; on outbreak days, moister air in the west and north and drier air elsewhere. The composite map of UT omega on contrail outbreak days (Figure 3-7b) shows a much weaker spatial gradient than on non-outbreak days. Moreover, the composite difference map (omega on outbreak days minus omega on non-outbreak days) indicates the greatest values over the Mississippi River region, comprising positive omega (subsiding motion) on outbreak days, but negative values (ascending motion) on non-outbreak days. The composite map of zonal winds for the Midwest/Upper South in October shows substantially weaker westerlies on outbreak days contrasted with non-outbreak days, particularly over

Missouri, Illinois, Indiana, and western Kentucky. Similar to what was observed in this sub-region in April, the greater subsidence rather than ascendance on outbreak composite days is contradictory to what is expected for contrail formation and persistence. However, as in April, the UT zonal winds are weaker on outbreak days than on non-outbreak days. This likely means that there is an upper-tropospheric high or ridge giving enhanced subsidence, which also explains the reduced westerlies. This finding is consistent with what Carleton et al. (2008) found for the Midwest (on average for the 4 mid-season months combined).

GrADS: COLA/IGES

Figure 3-2a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-2b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-3a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-3b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-4a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-4b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-5a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

April Pacific South 300mb Omega & U Wind

Figure 3-5b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

July East 300mb Temp & RH

Figure 3-6a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-6b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-7a: 2000-2002 composites of temperature (colored, C°) and RH (black, %).

Figure 3-7b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-8a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

January Midwest 300mb Omega & U Wind, Non-Outbreak

GrADS: COLA/IGES

Figure 3-8b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

January South 300mb Temp & RH, Non-Outbreak

GrADS: COLA/IGES

Figure 3-9a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-9b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-10a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

April Central 300mb Omega & U Wind, Non-Outbreak

Figure 3-10b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

April Pacific South 300mb Temp & RH, Non-Outbreak

Figure 3-11a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure 3-11b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-12a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

July East 300mb Omega & U Wind, Non-Outbreak

Figure 3-12b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

October Midwest/UpperSouth 300mb Temp & RH, Non-Outbreak

Figure 3-13a: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

ctober Midwest/UpperSouth 300mb Omega & U Wind, Non-Outbrea

GrADS: COLA/IGES

Figure 3-13b: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

January Midwest 300mb Temp & RH, Outbreak-Non

GrADS: COLA/IGES

Figure 3-14a: 2000-2002 difference maps of temperature (colored, C°) and RH (black,

January Midwest 300mb Omega & U Wind, Outbreak−Non

Figure 3-14b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

January South 300mb Temp & RH, Outbreak − Non

Figure 3-15a: 2000-2002 difference maps of temperature (colored, Cº) and RH (black, %).

January South 300mb Omega & U Wind, Outbreak–Non

Figure 3-15b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

April Central 300mb Temp & RH, Outbreak–Non

Figure 3-16a: 2000-2002 difference maps of temperature (colored, Cº) and RH (black, %)

April Central 300mb Omega & U Wind, Outbreak−Non

Figure 3-16b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

April Pacific South 300mb Temp & RH, Outbreak−Non

Figure 3-17a: 2000-2002 difference maps of temperature (colored, Cº) and RH (black, %)

Figure 3-17b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-18a: 2000-2002 difference maps of temperature (colored, Cº) and RH (black, %)

Figure 3-18b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

Figure 3-19a: 2000-2002 difference maps of temperature (colored, Cº) and RH (black, %).

ctober Midwest/UpperSouth 300mb Omega & U Wind, Outbreak–No

GrADS: COLA/IGES

Figure 3-19b: 2000-2002 difference maps omega (colored, Pa/s) and zonal wind (black, m/s).

Table 3-1 summarizes the differences in UT variables in terms of map magnitude, pattern, or gradient, between composites of outbreak days (Figures 3-2 through 3-7) and non-outbreak days (Figures 3-8 through 3-13) in each of the six high-frequency regions. By visual comparison with each day's average UT analysis, these criteria are used to determine the favorability of a sub-region in a given mid-season month to contrail outbreak occurrence.

Table 3-1: Primary UT variable differences of outbreak day composites with respect to non-outbreak day composites

Midwest in January

RH300	Omega(300)	T300	U(300)
Magnitude: -40ºC+ over the Great Lakes states, -45ºC+ over Michigan	Pattern: negative values (ascendance) centered over northern Missouri, Iowa, western IL	Magnitude: -49ºC or greater over Minnesota and South Dakota	Pattern: increasing eastward, gradients meridional

South in January

RH300	Omega(300)	T300	U(300)
Pattern: moister over Gulf of Mexico, drier to the north over land	Pattern: decreasing positive values (subsidence) from east to west	Magnitude: cooler in the northeast (-43.5ºC or lower)	Magnitude: stronger over peninsular Florida

Central in April

RH300	Omega(300)	T300	U(300)
Pattern: drier in the south, moister in the north	Pattern: stronger positive values (subsidence) everywhere	Magnitude: -41.5ºC or greater everywhere	Magnitude: less than 29 m/s everywhere

Pacific South in April

RH300	Omega(300)	T300	U(300)
Gradient: strong across Nevada and California, from low in the east to high in the west	Pattern: very strong subsidence in the east, very strong ascendance in the northwest	Magnitude: warmer over land (-42.5ºC or greater)	Pattern: decreasing southward

East in July

RH300	Omega(300)	T300	U(300)
Pattern: moister over Tennessee, Kentucky, western North Carolina, western and central Virginia; drier elsewhere	Gradient & Pattern: moderately strong gradient (0.06 Pa/s+); more positive values in the northwest, more negative values in the southeast	Magnitude: less than -36.5ºC in northeast from Maryland northward	Magnitude: less than 13 m/s in northwestern Indiana, greater than 11 m/s off the North Carolina Outer Banks

Midwest/Upper South in October

RH300	Omega(300)	T300	U(300)
Pattern: moister in the west and north, drier elsewhere	Magnitude: positive values (subsidence) over AR, MOi, IL, Lake Michigan	Gradient: weaker (difference of 5ºC or less) from north to south	Magnitude: 25 m/s or less over most of the region

Statistical Analysis of Contrail Outbreak Retro-Prediction

This section shows and explains the results of the statistical analyses performed to determine which of the four UT variables of interest (temperature, RH, omega, U wind) are most skillful at retro-predicting contrail outbreaks. Statistical analyses were based on the resulting tables (Appendix B, Tables B1 through B6) that were created to compare the retro-predicted and actual outbreak dates in the 2000-2002 study period based on individual UT variable favorability.

Contingency tables

The 24 contingency tables created for each UT variable (4), for each of the six regions of interest, are shown in Tables 3-2a-f. Also shown for each table are the marginal totals of "Yes" and "No" for both forecasts and observations. The lower right corner of each table gives the sum of either the row or column marginal totals, which is the number of days in the dataset. Tables 3-3a-f show the values of each of the four accuracy measures, as well as the B values (bias), for each of the four UT variables, within each of the six sub-regions of interest. Please note that for variable T300 for the Pacific South sub-region in April, I present both a full (four-cell) contingency table, as well as a two-cell table, which combines the success categories (Hit, Correct Negative) into one cell, and the failure categories (Miss, False Alarm) into the other. I created the latter table because many of the values in each of the four cells are fewer than 5 occurrences, which violates the assumptions of the standard Chi-squared test.

January Midwest

RH	Predicted	No	Σ
Occurrence	7	9	16
No	27	50	77
Σ	34	59	93

Omega	Predicted	No	Σ
Occurrence	9	7	16
No	21	56	77
Σ	30	63	93

T	Predicted	No	Σ
Occurrence	9	7	16
No	24	53	77
Σ	33	60	93

U Wind	Predicted	No	Σ
Occurrence	1	15	16
No	16	61	77
Σ	17	76	93

Table 3-2a: Contingency table for T, RH, Omega, U Wind for January Midwest

January South

RH	Predicted	No	Σ
Occurrence	12	13	25
No	25	43	68
Σ	37	56	93

Omega	Predicted	No	Σ
Occurrence	11	14	25
No	22	46	68
Σ	33	60	93

T	Predicted	No	Σ
Occurrence	12	13	25
No	26	42	68
Σ	38	55	93

U Wind	Predicted	No	Σ
Occurrence	14	11	25
No	28	40	68
Σ	42	51	93

Table 3-2b: Contingency table for T, RH, Omega, U Wind for January South

April Central

RH	Predicted	No	Σ	
Occurrence	15	14		29
No	15	46		61
Σ	30	60		90

Omega	Predicted	No	Σ	
Occurrence	13	16		29
No	19	42		61
Σ	32	58		90

T	Predicted	No	Σ	
Occurrence	20	9		29
No	32	29		61
Σ	52	38		90

U Wind	Predicted	No	Σ	
Occurrence	22	7		29
No	36	25		61
Σ	58	32		90

Table 3-2c: Contingency table for T, RH, Omega, U Wind for April Central

April Pacific South

RH	Predicted	No	Σ	
Occurrence	3	5		8
No	21	61		82
Σ	24	66		90

Omega	Predicted	No	Σ	
Occurrence	5	3		8
No	10	72		82
Σ	15	75		90

T	Predicted	No	Σ	
Occurrence	4	4		8
No	18	64		82
Σ	22	68		90

U Wind	Predicted	No	Σ	
Occurrence	3	5		8
No	25	57		82
Σ	28	62		90

Table 3-2d: Contingency table for T, RH, Omega, U Wind for April Pacific South

July East

RH	Predicted	No	Σ	
Occurrence	11	13		24
No	9	60		69
Σ	20	73		93

Omega	Predicted	No	Σ	
Occurrence	8	16		24
No	12	57		69
Σ	20	73		93

T	Predicted	No	Σ	
Occurrence	19	5		24
No	66	3		69
Σ	85	8		93

U Wind	Predicted	No	Σ	
Occurrence	3	21		24
No	10	59		69
Σ	13	80		93

Table 3-2e: Contingency table for T, RH, Omega, U Wind for July East

October
Midwest/Upper South

RH	Predicted	No	Σ	
Occurrence	13	14		27
No	11	55		66
Σ	24	69		93

Omega	Predicted	No	Σ	
Occurrence	10	17		27
No	22	44		66
Σ	32	61		93

T	Predicted	No	Σ	
Occurrence	11	16		27
No	17	49		66
Σ	28	65		93

U Wind	Predicted	No	Σ	
Occurrence	11	16		27
No	21	45		66
Σ	32	61		93

Table 3-2f: Contingency table for T, RH, Omega, U Wind for October Midwest/U South

January Midwest	χ2
RH	0.430833
Omega	5.090341
T	3.64019
U Wind	1.872043

Table 3-3a: Chi-squared scores for January Midwest

January South	χ2
RH	0.963188
Omega	0.850884
T	0.640871
U Wind	1.621869

Table 3-3b: Chi-squared scores for January South

April Central	χ2
RH	6.512154
Omega	1.745907
T	1.452569
U Wind	2.434273

Table 3-3c: Chi-squared scores for April Central

April PacSo	χ2
RH	0.526954
Omega	13.28049
T	17
U Wind	0.167226

Table 3-3d: Chi-squared scores for April Pacific South

July East	χ2
RH	11.34147
Omega	2.680881
T	6.155155
U Wind	0.058805

Table 3-3e: Chi-squared scores for July East

October	
MUS	χ2
RH	9.918378
Omega	0.116462
T	2.044214
U Wind	0.675911

Table 3-3f: Chi-squared scores for October Midwest/Upper South

Chi-squared statistics and accuracy measures

The Chi-squared statistics for each UT variable by sub-region are shown in Tables 3-4a-f. For the Midwest in January, the only map variable that is statistically significant is omega300 ($\chi^2 = 5.09$), indicating an association with the favorability of occurrence of contrail outbreaks in the 2000-2002 period. Also in January, except for the South sub-region, no variables individually are significantly associated with outbreak events. For the Central sub-region in April, RH300 is the significant variable associated with contrail outbreaks ($\chi^2 = 6.51$). Also in April, except for the Pacific South sub-region, omega300 and T300 are both significantly associated with contrail outbreak occurrences (χ^2 values = 13.28 and 17.0, respectively). For the East sub-region in July, RH300 is significantly associated with outbreaks ($\chi^2 = 11.34$), while temperature ($\chi^2 = 6.15$) is less so. For the Midwest/Upper South sub-region in October, RH300 is the only UT variable strongly associated with outbreak occurrences ($\chi^2 = 9.91$).

January Midwest

	RH	Omega	T	U Wind
Predicted/Occurred	7	9	9	1
Predicted/Not	27	21	24	16
Not Predicted/Occurred	9	7	7	15
Not/Not	50	56	53	61
Hit Rate	0.612903	0.698925	0.666667	0.666667
TS or CSI	0.162791	0.243243	0.225	0.03125
POD	0.4375	0.5625	0.5625	0.0625
FAR	0.794118	0.7	0.727273	0.941176
Bias	2.125	1.875	2.0625	1.0625

Table 3-4a: Accuracy measures and biases for January Midwest

January South

	RH	Omega	T	U Wind
Predicted/Occurred	12	11	12	14
Predicted/Not	25	22	26	28
Not Predicted/Occurred	13	14	13	11
Not/Not	43	46	42	40
Hit Rate	0.5914	0.6129	0.580645	0.580645
TS or CSI	0.24	0.234043	0.235294	0.264151
POD	0.48	0.44	0.48	0.56
FAR	0.675676	0.666667	0.684211	0.666667
Bias	1.48	1.32	1.52	1.68

Table 3-4b: Accuracy measures and biases for January South

April Central

	RH	Omega	T	U Wind
Predicted/Occurred	15	13	20	22
Predicted/Not	15	19	32	36
Not Predicted/Occurred	14	16	9	7
Not/Not	46	42	29	25
Hit Rate	0.677778	0.6111	0.5444	0.5222
TS or CSI	0.340909	0.270833	0.327869	0.338462
POD	0.517241	0.448276	0.689655	0.758621
FAR	0.5	0.59375	0.615385	0.62069
Bias	1.034483	1.103448	1.793103	2

Table 3-4c: Accuracy measures and biases for April Central

April Pacific South

	RH	Omega	T	U Wind
Predicted/Occurred	3	5	4	3
Predicted/Not	21	10	18	25
Not Predicted/Occurred	5	3	4	5
Not/Not	61	72	64	57
Hit Rate	0.711111	0.866667	0.755556	0.666667
TS or CSI	0.103448	0.277778	0.153846	0.090909
POD	0.375	0.625	0.5	0.375
FAR	0.875	0.666667	0.818182	0.892857
Bias	3	1.875	2.75	3.5

Table 3-4d: Accuracy measures and biases for April Pacific South

July East

	RH	Omega	T	U Wind
Predicted/Occurred	11	8	19	3
Predicted/Not	9	12	66	10
Not Predicted/Occurred	13	16	5	21
Not/Not	60	57	3	59
Hit Rate	0.763441	0.698925	0.236559	0.666667
TS or CSI	0.333333	0.222222	0.211111	0.088235
POD	0.458333	0.333333	0.791667	0.125
FAR	0.45	0.6	0.776471	0.769231
Bias	0.833333	0.833333	3.541667	0.541667

Table 3-4e: Accuracy measures and biases for July East

October Midwest/Upper South

	RH	Omega	T	U Wind
Predicted/Occurred	13	10	11	11
Predicted/Not	11	22	17	21
Not Predicted/Occurred	14	17	16	16
Not/Not	55	44	49	45
Hit Rate	0.731183	0.580645	0.645161	0.602151
TS or CSI	0.342105	0.204082	0.25	0.229167
POD	0.481481	0.37037	0.407407	0.407407
FAR	0.458333	0.6875	0.607143	0.65625
Bias	0.888889	1.185185	1.037037	1.185185

Table 3-4f: Accuracy measures and biases for October Midwest/Upper South

To evaluate the success of which UT variables were deemed favorable (unfavorable) in relation to contrail outbreak occurrences (non-occurrences), the accuracy measures of the UT variables found to be significant by the Chi-squared test are examined. This involves comparing the statistical values of the H, CSI, POD, and FAR measures of skill for each of the six sub-regions of interest, and interpreting them for the significant UT variables associated with contrail outbreaks in the 2000-2002 study period. This allows for us to see the specific criteria which contributed to a UT variable's overall success, in terms of how it performed within each of the four contingency categories (Hit, Miss, False Alarm, Correct Negative).

For the Midwest sub-region in January, when the omega300 map pattern was found to be the significant variable, the Hits (H) value is 0.70, meaning that 70 percent of the days in the dataset consist of either Hits (i.e., Yes for retro-prediction, Yes for observation) or Correct Negatives (both No). However, the critical success index (CSI) is considerably lower (0.24), indicating that most of the success in H comes from correctly not predicting an outbreak, rather than in correctly predicting one. The probability of detection (POD) value (0.56) indicates that just over half the time that outbreaks occurred in January, a "Yes" prediction was correctly made on the basis of the omega300 pattern.. However, the False Alarm Rate (FAR) of 0.7 means that 70 percent of the time that an outbreak was predicted using omega300, it did not occur. Overall however, the ability of omega300 to retro-predict contrail outbreaks for this sub-region in this month seems to be reasonably good, due to an overall high H value and a POD above 0.5.

The South sub-region in January, for which the Chi-squared test determined that no single UT variable is significantly associated with contrail outbreaks, is not evaluated using the 4 skill measures. Instead, this relationship of UT variables to outbreaks in this sub-region/month is examined further using single-variable binary logistic regression and logit modeling for multi-variable interaction statistics (next section).

For the Central sub-region in April, when the RH300 map pattern is the significant UT variable associated with contrail outbreak occurrences, the corresponding H value is 0.67, indicating that an outbreak occurrence or non-occurrence was correctly predicted two-thirds of the time. However, similar to the Midwest in January, the lower CSI (= 0.34) means that most of these correct retro-predictions were for outbreak non-occurrences. The POD value of 0.51 indicates that RH300 predicted "Yes" for an outbreak on only about half of the actual outbreak days. However, the FAR (0.5) means that only half the time an outbreak was predicted it did not occur. The ability of RH300 to retro-predict contrail outbreaks for this sub-region in this month seems to be fair, given a relatively good POD and Hit Rate, and an average FAR.

For the Pacific South region in April, both omega300 and T300 were determined to be significantly associated with contrail outbreak occurrences in the 2000-2002 study period. However, because I determined that T300 has a very high bias (B) value (= 2.75), I do not be consider it to be a suitable UT variable for outbreak retro-prediction; at least, on its own. For the omega300 map pattern, the H value of 0.87 is the highest of the four UT variables across all sub-regions studied by mid-season months. Thus, for a significant majority of days in April, an outbreak occurrence or non-occurrence in the Pacific South was correctly retro-predicted on the basis of the omega300 map pattern.

However, this impressive result is tempered by the corresponding CSI value (0.27), which indicates that most of these correct retro-predictions were for non-occurrences rather than occurrences of outbreaks. The POD value (= 0.625) is moderately high, and means that over 60 percent of the time when an outbreak occurred, it was also retro-predicted on the basis of omega300. However, the FAR value of 0.67 indicates that two-thirds of the "Yes" retro-predictions of outbreaks yielded a non-occurrence, meaning that omega300 was over-predicted. Omega300 seems to be a good UT variable for retro-predicting contrail outbreaks for this sub-region in this month, despite a high FAR.

For the East sub-region in July, the Chi-squared test indicated that RH300 and T300 are significantly associated with contrail outbreaks in the 2000-2002 study period. The H value for the RH300 map pattern is 0.76, meaning that over three-quarters of the retro-predictions (either Yes or No) were successful. However, the CSI value of 0.33 means that RH300 more correctly retro-predicted outbreak non-occurrences than occurrences. The POD (0.46) is moderately low, indicating that fewer than half of the outbreak occurrences were predicted as "Yes." However, the FAR value of 0.45 means that RH300 performed quite well in not falsely retro-predicting a contrail outbreak. RH300 should ideally be a successful variable for retro-prediction for this month and sub-region, given the high Hit Rate and low FAR.

Although the T300 magnitude was significantly associated with contrail outbreaks in the East during July 2000-2002, the H value (0.24) is very low, indicating that three-quarters of the time neither outbreaks nor non-outbreaks were correctly retro-predicted. Also, the B value of 3.54 means that contrail outbreaks were over-predicted on the basis of the T300 magnitude. Given this poor performance, UT temperature is not a useful

individual retro-predictor of contrail outbreaks for the East region in July, and I do not consider it further.

For the Midwest/Upper South region in October, when the RH300 map is significantly associated with contrail outbreak occurrences in 2000-2002, the corresponding H value (0.73) indicates a 73 percent success rate for both outbreak and non-outbreak retro-predictions. However, the CSI value (0.34) means that most of these correct retro-predictions were for outbreak non-occurrences rather than for occurrences. The POD value (0.48) indicates that slightly fewer than half of the outbreak occurrences actually were associated with a "Yes" retro-prediction. The FAR measure performed reasonably well (= 0.45), meaning that less than half the time an outbreak was retro-predicted when did one not occur. With a high H value and relatively low FAR, RH300 should be a good retro-predictor for outbreaks during this month for this sub-region.

Binary logistic regression

The binary logistic regression analysis was performed to provide another set of measures (in addition to those from the contingency analyses) for retro-predicting outbreaks. Similarities in the results of this analysis (Tables 3-5a-f) to contingency help to solidify which UT variables should be used for forecasting outbreaks within each of the six high-frequency sub-regions. For the Midwest sub-region in January, the table indicates that only the omega300 map pattern has a p-value lower than the 5 percent test level (p = 0.0076). This finding indicates that the relation between prediction of this variable and actual outbreak occurrence is substantially better than random chance. For

the South sub-region in January, there are no significant p-values for the single variable

logistic regression, indicating that no UT variables, when used for prediction, perform

better than chance distribution. For the Central sub-region in April, the p-value for the

RH300 map pattern is the significant UT variable ($p = 0.049$). This indicates that the use

of RH in prediction is slightly better than random chance, as it is lower than 0.05. For

the Pacific South sub-region in April, the omega300 map pattern is the significant

variable ($p = 0.0070$). As this is significantly lower than the test level of 0.05, omega

appears to be a good retro-predictor for outbreaks. The East sub-region in July has two

significant UT variables according to their p-values: RH300 ($p = 0.0044$), and T300 ($p = 0.0334$). These results indicate that they are both ideal for retro-prediction of contrail

outbreaks. Finally, for the Midwest/Upper South sub-region in October, RH300 is the

significant UT variable (p-value $= 0.0050$). This is a value much lower than the test

level, and is therefore a good variable for retro-prediction of outbreaks, at least as

determined from the events in 2000-2002.

January Midwest

```
                           The SAS System          13:22 Tuesday,

                        The LOGISTIC Procedure

                  Analysis of Maximum Likelihood Estimates

                                    Standard        Wald
   Parameter    DF    Estimate        Error    Chi-Square      Pr > ChiSq

   Intercept     1     -2.4504       0.5540       19.5665         <.0001
   RH            1     -0.5914       0.7112        0.6915         0.4056
   Omega         1      1.8973       0.7111        7.1196         0.0076
   T             1      1.2045       0.6398        3.5446         0.0597
   UWind         1     -1.4722       1.1365        1.6780         0.1952


          Association of Predicted Probabilities and Observed Responses

             Percent Concordant    69.1    Somers' D    0.497
             Percent Discordant    19.4    Gamma        0.561
             Percent Tied          11.5    Tau-a        0.143
             Pairs                 1232    c            0.748
```

Table 3-5a: January Midwest SAS LOGISTIC output, single-variable

January South

```
                           The SAS System          13:23 Tuesday,

                        The LOGISTIC Procedure

                  Analysis of Maximum Likelihood Estimates

                                    Standard        Wald
   Parameter    DF    Estimate        Error    Chi-Square      Pr > ChiSq

   Intercept     1     -1.8520       0.5039       13.5090         0.0002
   RH            1      0.3364       0.5075        0.4393         0.5075
   Omega         1      0.8759       0.5203        2.8347         0.0922
   T             1      0.1573       0.5472        0.0826         0.7738
   UWind         1      0.6577       0.5769        1.2996         0.2543


          Association of Predicted Probabilities and Observed Responses

             Percent Concordant    61.5    Somers' D    0.314
             Percent Discordant    30.1    Gamma        0.343
             Percent Tied           8.4    Tau-a        0.125
             Pairs                 1700    c            0.657
```

Table 3-5b: January South SAS LOGISTIC output, single-variable

April Central

```
                              The SAS System            13:17 Tuesday, June

                         The LOGISTIC Procedure

                    Analysis of Maximum Likelihood Estimates

                                    Standard          Wald
    Parameter     DF    Estimate      Error       Chi-Square      Pr > ChiSq

    Intercept      1     -1.8563      0.5494        11.4163          0.0007
    RH             1      1.0174      0.5163         3.8822          0.0488
    Omega          1      0.1598      0.5134         0.0969          0.7556
    T              1      0.3309      0.5220         0.4020          0.5261
    UWind          1      0.6934      0.5295         1.7149          0.1904



          Association of Predicted Probabilities and Observed Responses

            Percent Concordant     64.3     Somers' D     0.362
            Percent Discordant     28.1     Gamma         0.392
            Percent Tied            7.6     Tau-a         0.160
            Pairs                  1769     c             0.681
```

Table 3-5c: April Central SAS LOGISTIC output, single-variable

April Pacific South

```
                              The SAS System            13:20 Tuesday, Ju

                         The LOGISTIC Procedure

                    Analysis of Maximum Likelihood Estimates

                                    Standard          Wald
    Parameter     DF    Estimate      Error       Chi-Square      Pr > ChiSq

    Intercept      1     -3.4601      0.7160        23.3511          <.0001
    RH             1      0.4656      0.8728         0.2846          0.5937
    Omega          1      2.2663      0.8409         7.2626          0.0070
    T              1      0.8603      0.8962         0.9213          0.3371
    UWind          1     -0.2331      0.9174         0.0646          0.7994



          Association of Predicted Probabilities and Observed Responses

            Percent Concordant     84.8     Somers' D     0.735
            Percent Discordant     11.3     Gamma         0.765
            Percent Tied            4.0     Tau-a         0.120
            Pairs                   656     c             0.867
```

Table 3-5d: April Pacific South SAS LOGISTIC output, single-variable

July East

The SAS System   13:25 Tuesday, Jun

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1324	0.7942	0.0278	0.8676
RH	1	1.6231	0.5693	8.1277	0.0044
Omega	1	0.9026	0.5882	2.3547	0.1249
T	1	-1.7878	0.8403	4.5268	0.0334
UWind	1	-0.0267	0.7746	0.0012	0.9725

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.0	Somers' D	0.481
Percent Discordant	17.9	Gamma	0.573
Percent Tied	16.1	Tau-a	0.186
Pairs	1656	c	0.740

Table 3-5e: July East SAS LOGISTIC output, single-variable

October Midwest/Upper South

The SAS System   13:26 Tuesday, Ju

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5928	0.4154	14.7029	0.0001
RH	1	1.4467	0.5159	7.8647	0.0050
Omega	1	0.2746	0.5258	0.2728	0.6015
T	1	0.5920	0.5897	1.0078	0.3154
UWind	1	-0.0717	0.5676	0.0160	0.8995

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.0	Somers' D	0.418
Percent Discordant	24.2	Gamma	0.463
Percent Tied	9.7	Tau-a	0.175
Pairs	1755	c	0.709

Table 3-5f: October Midwest/U South SAS LOGISTIC output, single variable

The Chi-squared statistics calculated from the contingency tables (previous section) and the p-values generated by running the favorability data in binary form through the LOGISTIC procedure give identical results as to the significant UT variables, associating contrail outbreak favorability with actual occurrences in the study period. These results confirm that the composite maps of those UT variables found to be significant by region and mid-season month can reasonably be compared to a daily mean UT map when retro-predicting contrail outbreaks using the simple visual map pattern technique that I propose. Once again, these variables and map criteria are: omega300 map pattern for the Midwest sub-region in January; RH300 map pattern for the Central sub-region in April; omega300 map pattern for the Pacific South sub-region in April; the RH300 map pattern and T300 gradient for the East sub-region in July; and RH300 map pattern for the Midwest/Upper South sub-region in October. The fact that RH300 and omega300 are the dominant UT variables associated with contrails in the six sub-regions of highest outbreak frequency indicates the dependence of moisture on the vertical motion of air (e.g., Carleton et al. 2008), with moistening in the UT typically accompanying air that is rising.

Two-way UT variable interactions

A binary logistic regression with two-way interactions of UT variables was performed in order to see if combinations of two variables, as opposed to single UT variables alone, were successful in predicting outbreak occurrences. The results of this analysis are shown in Tables 3-6a-f. For the South sub-region in January (Table 3-6b),

the most noteworthy interaction term is UT Omega*T (p-value = 0.0552). This value is considered significant, even though slightly above 0.05, because test levels can be expected to increase when additional terms are added to the model. Accordingly, UT omega and temperature will be used in combination for the verification study for mid-season months of 2008 and 2009 (Chapter 4). Among the remaining five regions and mid-season months, two more interaction terms are determined to be significant. For the East sub-region in July (Table 3-6e), RH*T is significant (p-value = 0.0484). However, recall from the analysis of single variable interactions that RH and temperature are significant individually, and therefore, the favorability of either mapped variable (RH300 or T300) should adequately retro-predict a contrail outbreak for a given day. The other significant interaction term (Table 3-6f) occurs for T*UWind in the Midwest/Upper South sub-region in October (p < 0.0174). Because the only UT significant variable in the single-variable analysis for this region and mid-season month was RH, I will use either (1) RH only or (2) UT temperature and zonal wind in combination for the outbreak verification study.

January Midwest

The SAS System   21:24 Monday, M

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.2877	0.6225	13.5032	0.0002
RH	1	0.2737	1.8303	0.0224	0.8811
Omega	1	1.1735	1.1509	1.0396	0.3079
T	1	0.6270	0.9335	0.4512	0.5018
UWind	1	0.6782	1.2600	0.2897	0.5904
RH*Omega	1	-0.4220	1.9085	0.0489	0.8250
RH*T	1	-0.9901	1.9097	0.2688	0.6041
RH*UWind	1	-10.2485	327.7	0.0010	0.9751
Omega*T	1	2.5011	1.9059	1.7221	0.1894
Omega*UWind	1	-11.5568	284.2	0.0017	0.9676
T*UWind	1	-0.4623	593.7	0.0000	0.9994

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.3	Somers' D	0.542	
Percent Discordant	17.1	Gamma	0.613	
Percent Tied	11.5	Tau-a	0.156	
Pairs	1232	c	0.771	

Table 3-6a: January Midwest SAS LOGISTIC output, two-way interaction

January South

The SAS System   21:55 Monday, M

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1533	0.6108	3.5658	0.0590
RH	1	-2.0306	1.4588	1.9374	0.1640
Omega	1	0.1621	0.8352	0.0377	0.8461
T	1	-1.9601	1.4495	1.8286	0.1763
UWind	1	0.6231	1.0787	0.3336	0.5635
RH*Omega	1	2.2806	1.5834	2.0746	0.1498
RH*T	1	2.1844	1.4563	2.2497	0.1336
RH*UWind	1	1.5068	1.6003	0.8865	0.3464
Omega*T	1	3.0809	1.6067	3.6768	0.0552
Omega*UWind	1	-2.7812	1.6605	2.8052	0.0940
T*UWind	1	0.6705	1.5938	0.1770	0.6740

Association of Predicted Probabilities and Observed Responses

Percent Concordant	69.5	Somers' D	0.474	
Percent Discordant	22.1	Gamma	0.517	
Percent Tied	8.4	Tau-a	0.188	
Pairs	1700	c	0.737	

Table 3-6b: January South SAS LOGISTIC output, two-way interaction

April Central

```
                          The SAS System              21:46 Monday, M

                       The LOGISTIC Procedure

                 Analysis of Maximum Likelihood Estimates

                                  Standard        Wald
   Parameter     DF    Estimate     Error     Chi-Square     Pr > ChiSq

   Intercept      1     -1.2582     0.7033       3.2000         0.0736
   RH             1     -0.5149     1.3898       0.1373         0.7110
   Omega          1     -0.3591     1.3234       0.0736         0.7861
   T              1      0.1653     0.9628       0.0295         0.8637
   UWind          1     -0.2931     0.8999       0.1061         0.7447
   RH*Omega       1     -1.2896     1.2018       1.1515         0.2832
   RH*T           1      1.0841     1.2217       0.7874         0.3749
   RH*UWind       1      1.8110     1.3300       1.8542         0.1733
   Omega*T        1     -0.3565     1.1828       0.0908         0.7631
   Omega*UWind    1      1.5739     1.3535       1.3523         0.2449
   T*UWind        1     -0.0263     1.1994       0.0005         0.9825



         Association of Predicted Probabilities and Observed Responses

           Percent Concordant     70.7    Somers' D    0.490
           Percent Discordant     21.7    Gamma        0.530
           Percent Tied            7.6    Tau-a        0.216
           Pairs                  1769    c            0.745
```

Table 3-6c: April Central SAS LOGISTIC output, two-way interaction

April Pacific South

```
                          The SAS System              21:50 Monday, M

                       The LOGISTIC Procedure

                 Analysis of Maximum Likelihood Estimates

                                  Standard        Wald
   Parameter     DF    Estimate     Error     Chi-Square     Pr > ChiSq

   Intercept      1    -25.5190     181.3        0.0198         0.8881
   RH             1     22.1097     181.3        0.0149         0.9030
   Omega          1     25.1135     181.3        0.0192         0.8898
   T              1     22.9962     181.3        0.0161         0.8991
   UWind          1     -7.8667     170.6        0.0021         0.9632
   RH*Omega       1    -21.6245     181.3        0.0142         0.9051
   RH*T           1    -19.3155     205.0        0.0089         0.9249
   RH*UWind       1      9.7136     170.6        0.0032         0.9546
   Omega*T        1    -21.5446     181.3        0.0141         0.9054
   Omega*UWind    1     -2.0065     2.3238       0.7455         0.3879
   T*UWind        1      7.7808     170.6        0.0021         0.9636



         Association of Predicted Probabilities and Observed Responses

           Percent Concordant     89.8    Somers' D    0.835
           Percent Discordant      6.3    Gamma        0.870
           Percent Tied            4.0    Tau-a        0.137
           Pairs                   656    c            0.918
```

Table 3-6d: April Pacific South SAS LOGISTIC output, two-way interaction

July East

```
                                    The SAS System              22:00 Monday, Ma

                               The LOGISTIC Procedure

                     Analysis of Maximum Likelihood Estimates

                                        Standard         Wald
    Parameter       DF     Estimate       Error     Chi-Square    Pr > ChiSq

    Intercept        1       1.0986      1.1547        0.9052        0.3414
    RH               1      -1.7918      1.6833        1.1331        0.2871
    Omega            1       9.8124       234.1        0.0018        0.9666
    T                1      -3.2268      1.2479        6.6869        0.0097
    UWind            1       1.0296      0.9437        1.1905        0.2752
    RH*Omega         1       9.8816     76.4549        0.0167        0.8972
    RH*T             1       3.6968      1.8727        3.8968        0.0484
    RH*UWind         1      -0.8065      1.8277        0.1947        0.6590
    Omega*T          1      -9.1883       234.1        0.0015        0.9687
    Omega*UWind      1     -19.4561       116.4        0.0279        0.8673
    T*UWind          0            0           0             .             .



            Association of Predicted Probabilities and Observed Responses

                Percent Concordant      72.8    Somers' D     0.617
                Percent Discordant      11.1    Gamma         0.735
                Percent Tied            16.1    Tau-a         0.239
                Pairs                   1656    c             0.809
```

Table 3-6e: July East SAS LOGISTIC output, two-way interaction

October Midwest/Upper South

```
                                    The SAS System              22:05 Monday, M

                               The LOGISTIC Procedure

                     Analysis of Maximum Likelihood Estimates

                                        Standard         Wald
    Parameter       DF     Estimate       Error     Chi-Square    Pr > ChiSq

    Intercept        1      -1.8597      0.5392       11.8953        0.0006
    RH               1       1.9162      0.9903        3.7439        0.0530
    Omega            1       0.6733      0.8325        0.6540        0.4187
    T                1       2.3087      1.0248        5.0748        0.0243
    UWind            1       0.5036      1.3260        0.1442        0.7041
    RH*Omega         1      -2.7189      1.4992        3.2889        0.0697
    RH*T             1      -1.3448      1.4649        0.8428        0.3586
    RH*UWind         1       2.4652      1.5432        2.5519        0.1102
    Omega*T          1       2.0661      1.6990        1.4787        0.2240
    Omega*UWind      1      -0.2248      1.5395        0.0213        0.8839
    T*UWind          1      -3.9544      1.6633        5.6524        0.0174



            Association of Predicted Probabilities and Observed Responses

                Percent Concordant      74.0    Somers' D     0.578
                Percent Discordant      16.2    Gamma         0.640
                Percent Tied             9.7    Tau-a         0.242
                Pairs                   1755    c             0.789
```

Table 3-6f: October Midwest/U South SAS LOGISTIC output, two-way interaction

Summary

Using contingency Chi-squared tests, four measures of statistical skill, and binary logistic regression for both single variable and two-way variable interactions, I determined the mapped UT variables (out of temperature, RH, omega, and zonal wind) and their combinations that are most closely associated with retro-predicted contrail outbreaks in the high-frequency sub-regions for mid-season months of 2000-2002. These variables and their combinations are as follows: for the Midwest sub-region in January, omega300 map pattern; for the South sub-region in January, both omega300 map pattern and T300 magnitude; for the Central sub-region in April, RH300 map pattern; for the Pacific South sub-region in April, omega300 map pattern; for the East sub-region in July, either RH300 map pattern or T300 magnitude; and for the Midwest/Upper South sub-region in October, only RH300 map pattern or the combination of T300 magnitude and zonal wind magnitude. Accordingly, these variables and variable combinations are used in the retro-prediction verification study for mid-season months of 2008 and 2009 undertaken in the next chapter.

In general, the results of this chapter yielded the following:

1. RH300 and omega300 are most frequently associated with contrail outbreaks in the high-frequency sub-regions because they are inter-related variables. Greater UT moisture—as given by the higher values of relative humidity—is required for persisting contrails and outbreaks

(e.g., Carleton et al. 2008), and this enhanced moisture content is most likely to occur in areas where air is rising (negative omega).

2.   Chi-squared tests and the binary logit model give similar results as do those UT variables most closely associated with contrail outbreaks (by sub-region and mid-season month), indicating a robustness to the results and increasing the confidence with which they can be used in the verification study for 2008-09 (Chapter 4).

3.   The skill score results suggest that the simple map-based interpretation method of retro-predicting contrail outbreaks—indicated by the criteria of variable magnitude, pattern, or horizontal gradient on the composite analyses—is moderately successful. The April regions (Central and Pacific South) seemed to have the most success when retro-predicting outbreaks in terms of composite similarity, while the South region in January performed the least well with this method. A simpler method that is relatively quick to use is preferable to one that is complex or time-consuming to use and, therefore, is the one more likely to be adopted for use in forward prediction of contrail outbreaks in real time.

Chapter 4

Retro-Prediction of Contrail Outbreaks for 2008 and 2009 Mid-Season Months

In this chapter, I report and discuss the results of the retro-prediction procedure

for 2008 and 2009. This retro-prediction used the UT variables, individually or in

combination, that were found to be statistically significant when retro-predicting contrail

outbreaks for the 2000-2002 mid-season months, applied to the same high-frequency

regions for July and October 2008 and January and April 2009. I also visualize and

discuss some of the relations between contrail outbreak frequency, duration, and size, as

derived from the analysis of AVHRR images for these months.

Retro-prediction results

As shown in Chapter 3, the significant variables by region resulting from the

statistical analysis of the 2000-2002 datasets are: either RH or temperature for the East in

July; only RH or the combination of temperature and zonal wind for the Midwest/Upper

South in October; omega300 for the Midwest in January; both omega and temperature for

the South in January; and RH for the Central U.S. in April and omega300 for the Pacific

South in April. As described in Chapter 2, the retro-prediction process was completed for

each of the six aforementioned regions in April and January 2009 and October and July

2008. Retro-predictions for contrail outbreaks were made based on combinations of UT

variables found to be statistically significant. Tables 4-1a-f show the resulting

comparisons of retro-predictions to actual outbreak occurrences. Cells in yellow

represent days for which one or more outbreaks were predicted. Cells in blue represent

days on which outbreaks actually occurred. Green cells represent days for which there

was a retro-prediction and an observation. Un-colored cells represent days for which

there was neither an outbreak retro-prediction nor an observation. Therefore, as in the

analysis for 2000-2002, the latter two categories indicate successful retro-prediction in

that a contrail was successfully predicted or successfully not predicted.

2008	RH	T
1		X
2		X
3		X
4		X
5	X	X
6	X	
7		
8		
9	X	
10		
11	X	
12	X	
13	X	
14		
15		X
16		
17		
18		
19		
20		
21	X	
22		
23	X	
24		
25	X	
26		
27		
28		X
29		X
30		
31		

Table 4-1a: Retro-prediction table for East sub-region in July 2008

2008	RH	T	UWind
1			X
2	X	X	X
3	X		
4			
5	X		
6	X	X	
7		X	X
8		X	X
9		X	
10	X	X	
11			X
12	X		X
13			X
14		X	
15		X	
16			
17			
18			
19	X		X
20			
21			
22	X		
23			
24			
25			
26		X	
27	X		
28	X		X
29			
30		X	X
31	X	X	X

Table 4-1b: Retro-prediction table for Midwest/Upper South sub-regions in October 2008

2009	RH
1	
2	
3	
4	X
5	X
6	X
7	
8	
9	
10	
11	X
12	
13	X
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	X
26	X
27	
28	
29	
30	

Table 4-1c: Retro-prediction table for Central sub-region in April 2009

2009	Omega
1	
2	
3	
4	
5	
6	
7	X
8	
9	X
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	X
25	
26	
27	
28	
29	
30	

Table 4-1d: Retro-prediction table for Pacific South sub-region in April 2009

2009	Omega
1	X
2	
3	X
4	X
5	
6	
7	
8	
9	
10	X
11	
12	
13	
14	
15	
16	
17	X
18	
19	
20	
21	
22	
23	
24	
25	X
26	X
27	X
28	
29	
30	
31	

Table 4-1e: Retro-prediction table for Midwest sub-region in January 2009

2009	Omega	T
1		X
2		X
3		
4		
5	X	
6		
7		
8		
9	X	
10		
11		
12		
13		
14		X
15		X
16		X
17	X	X
18		X
19		X
20		X
21	X	X
22		X
23	X	X
24		
25		
26		
27		
28		
29		
30		
31	X	X

Table 4-1f: Retro-prediction table for South sub-region in January 2009

The verification study to retro-predict contrail outbreaks for the East sub-region in July 2008 (Table 4-2a) showed modest success. Recall that the UT map criteria I used for this month and sub-region were either a favorable pattern of RH300, or favorable magnitudes of T300. The combined Hit Rate (Hit + Correct Negative) for the East sub-region in July 2008 was 0.51, meaning that I only correctly retro-predicted half of the days having outbreaks. The number of Hits was reasonably good, with eight days having

"Yes" retro-prediction and "Yes" occurrences of outbreaks. However, the CSI of 0.34 indicates that only about one of every three "Yes" occurrences corresponded to a "Yes" retro-prediction on those days. The POD measure performed reasonably well (= 0.57), indicating that slightly over half of the days on which outbreaks occurred I correctly retro-predicted. In contrast, there were a fairly high number of days (9) on which False Alarms occurred (FAR = 0.52). Like the Midwest/Upper South region in October 2008, the verification for this month was fair, with no very good accuracy measures.

The retro-prediction of contrail outbreaks for the Midwest/Upper South sub-region in October 2008 (Table 4-2b) had limited success on the basis of determining favorability using RH300 map pattern or the combined favorability of T300 gradient and U(300) magnitude. I correctly retro-predicted outbreaks for six days of the month; however, there were just as many Misses as Hits, with 6 days having a "No" retro-prediction but a "Yes" occurrence. In addition, there were more days (8) having False Alarms—a "Yes" retro-prediction of an outbreak but a "No" occurrence—giving this month and sub-region a FAR = 0.57. The CSI of 0.3 indicates that only 30 percent of the "Yes" retro-predictions correctly correspond to the "Yes" observations. The POD of 0.5 means that I correctly retro-predicted exactly half of the days on which outbreaks occurred. Combining the correct forecasts (Hit + Correct Negative), the Hit rate for this sub-region in October 2008 was 0.54, meaning that retro-predictions were successful on just over half the days in the month. The overall verification for this month was only fair, as none of the accuracy measure scores ranged above good.

The retro-prediction of contrail outbreaks undertaken for the Midwest sub-region in January 2009 (Table 4-2c), which used the map pattern of omega300, resulted in

moderate success. Three days were verified as Hits, meaning they had both "Yes" retro-predictions and "Yes" observations. The calculated CSI was 0.30, meaning that only 30 percent of the total "Yes" retro-predictions and "Yes" observations occurred on the same day—a poor result. However, the POD value was reasonably good, at 0.6, meaning that more often than not, contrail outbreaks were retro-predicted correctly for the days on which they occurred. Because a greater number of "Yes" retro-predictions yielded "No" observations than "Yes" observations, the FAR for this sub-region in January 2009 was somewhat high at 0.625. However, the relatively high number of days with Correct Negatives ("No" retro-prediction of contrail outbreak verified as "No" observation of an outbreak) helped to bring the Hit (success) rate to a respectable 0.77. Overall, the verification for this month is good, given a high Hit Rate and relatively high POD.

The retro-prediction of contrail outbreaks for the South sub-region in January 2009 (Table 4-2d) was unsuccessful in comparison to those undertaken for the July 2008, October 2008, and April 2009 mid-season months and associated sub-regions of high outbreak frequency. Recall that the UT map-based criterion for outbreak retro-prediction in this sub-region is the omega300 pattern in combination with the magnitude of T300 (both favorable). Using this criterion, I deemed only four days to be favorable for contrail outbreaks in the South during January 2009; however, the satellite image verification showed 15 days with outbreaks. Furthermore, only one day (January 23) had both a "Yes" retro-prediction and "Yes" observation. As a result, the CSI—indicating how well the "Yes" retro-predictions correspond to the "Yes" observations—is extremely poor, at 0.055. The POD, which measures the accuracy of retro-prediction for the day when the event is observed, was also very low, at 0.067. Interestingly, because three out

of four days on which "Yes" retro-predictions were made actually saw outbreaks, the FAR was high at 0.75. In contrast, because 13 days in the month had successful retro-predictions of non-occurrences of contrail outbreaks (Correct Negative), the Hit rate was somewhat poor, at 0.45. The verification for this month was poor overall.

The retro-prediction of contrail outbreaks for the Central sub-region in April 2009 (Table 4-2e) using RH300 map pattern showed reasonable success. The Hit Rate (H) value of 0.7 means that 70 percent of the time the retro-prediction for the presence or absence of contrail outbreaks was correct. Five of the seven days for which outbreaks were retro-predicted had a verifying contrail outbreak on the satellite imagery; therefore, the FAR score was very good, at 0.28. Of the days for which outbreaks occurred but were not predicted, seven days observed outbreaks in the sub-region while the remaining 16 did not. The CSI for this sub- region in April 2009 was 0.36, meaning that over one of every three "Yes" retro-predictions corresponded to a "Yes" observation. In addition, the POD—the prediction success measured only in terms of days for which the forecast event was observed—was 0.41. Overall, the verification for this month and sub-region is good.

The retro-prediction undertaken for the Pacific South region in April 2009 using omega300 map pattern (Table 4-2f) also showed reasonable success according to the skill score statistics. The hit rate value of 0.8 means that 80 percent of the time the retro-prediction of contrail outbreak presence or absence was correct. Of the three days for which "Yes" retro-predictions were made, two outbreaks occurred (April 7, April 9), giving a fairly good FAR (= 0.33). However, there were five days for which contrail outbreaks were observed but not retro-predicted. As a result, the CSI, which measures the correspondence between "Yes" retro-predictions and "Yes" observations, was low

(0.25), as was the POD, which measures the fraction of successful retro-predictions for "Yes" observation days (0.28). Conversely, there were 22 days for which outbreaks were correctly not retro-predicted (i.e., Correct Negative). Thus, the verification for this month and sub-region is mixed, but reasonably good overall given the high Hit Rate and FAR.

	Forecast	No
Occurrence	8	6
No	9	8

CSI	0.347826
POD	0.571429
FAR	0.529412
Hit Rate	0.516129

Table 4-2a: Contingency table and accuracy measures for the East sub-region in July 2008

	Forecast	No
Occurrence	6	6
No	8	11

CSI	0.3
POD	0.5
FAR	0.571429
Hit Rate	0.548387

Table 4-2b: Contingency table and accuracy measures for Midwest/Upper South region in October 2008

	Forecast	No
Occurrence	3	2
No	5	21

CSI	0.3
POD	0.6
FAR	0.625
Hit Rate	0.774194

Table 4-2c: Contingency table and accuracy measures for Midwest sub-region in January 2009

	Forecast	No
Occurrence	1	14
No	3	13

CSI	0.055556
POD	0.066667
FAR	0.75

Hit Rate	0.451613

Table 4-2d: Contingency table and accuracy measures for South sub-region in January 2009

	Forecast	No
Occurrence	5	7
No	2	16

CSI	0.3571429
POD	0.4166667
FAR	0.2857143

Hit Rate	0.7

Table 4-2e: Contingency table and accuracy measures for Central sub-region in April 2009

	Forecast	No
Occurrence	2	5
No	1	22

CSI	0.25
POD	0.285714
FAR	0.333333

Hit Rate	0.8

Table 4-2f: Contingency table and accuracy measures for Pacific South sub-region in April 2009

Addition of non-significant UT variables

Given the above mixed retro-prediction results of the 2008 and 2009 verification study, I investigated which additional (not statistically significant) UT variables would have successfully detected contrail outbreaks on days for which they occurred but were not retro-predicted. For these additional analyses, I downloaded the NCEP/NCAR reanalysis data from the CDC website, imported them into GrADS, and viewed each on a day-to-day basis within the month of concern, similar to the way I interpreted the maps of statistically significant variables. The results are shown in Appendix C.

The East sub-region in July 2008 does not detect any additional outbreak occurrences using U300. However, the addition of omega300 includes four more outbreaks that were not accounted for by the RH300-or-T300 criteria. Nonetheless, because omega300 favorability adds more to the number of False Alarms than it does to the number of Hits, the retro-prediction skill is significantly decreased for this region and month if omega is incorporated into the retro-prediction criteria. "Yes" retro-predictions would be made for 23 of the 31 days in the month.

For the Midwest/Upper South sub-region in October 2008, three of the four UT variables analyzed in this study are used, as the criteria for the retro-prediction of outbreaks on any given day were either only the favorability of RH300 or the combination of T300 and U300 favorability. Neither changing the criteria to RH, T, or U Wind, nor to all three simultaneously, improves the retro-prediction. By adding omega300, only two more of the missed outbreaks ("no" retro-prediction and "yes"

occurrence) were detected. This finding does not merit the addition of omega to increase retro-prediction success.

For the Midwest sub-region in January 2009, T300, U300, and RH300 all showed little success in detecting outbreaks that were not associated with omega favorability. Worth noting, however, is that omega300 and RH300 had similar patterns of favorability throughout the month, meaning that, on most days, either both would be present simultaneously or neither would be present. We can recall that the criteria for the UT variables' favorability, based on the composites of outbreak days for 2000-2002, were relatively high values of RH and negative values of omega (ascendance) over the region. Therefore, the occurrence of these two phenomena together agree with the properties of synoptic-scale dynamics—that rising air is associated with upper-level divergence and condensation, and therefore a moistening of the UT.

Retro-prediction of contrail outbreaks for the South sub-region in January 2009 can be improved not by using the same two variables of significance used in the retro-prediction simultaneously, but by using them interchangeably (i.e., either/or). A contingency table was created that accounted for using either omega300 or T300 in outbreak retro-prediction (Table 4-3a), and success measures were calculated based on this table. Compared to the contingency table requiring both variables to be simultaneously favorable for retro-prediction, although still not as successful as in other regions, the CSI and POD improved drastically, at 0.25 and 0.4 respectively. This finding means that there is a one in four correspondence rate between "yes" retro-predictions and "yes" occurrences, and a 40 percent retro-prediction success rate among the days for which outbreaks did occur. Although the FAR was improved as well, to 0.6,

the increase in the number of False Alarm days took away from the number of Correct Negatives, thereby leading to a reduced Hit Rate (Hit + Correct Negative) of 0.42. The addition of RH300 and U300 do not aid significantly in the detection of more outbreak occurrences.

For the Central sub-region in April 2009, when omega300 is added to RH300—the only other UT variable found to be statistically significant in the 2000-2002 base period—the result was only slightly improved compared to RH300 alone. Changing the criteria for retro-prediction (to increase the chance of detection) to either favorability of RH300 or favorability of omega300 added two more days of Hits ("Yes" retro-prediction and "Yes" observation) to the month, but also added two more False Alarm days ("Yes" retro-prediction and "No" observation, FAR = 0.36). The resulting contingency table (Table 4-3b) shows that the CSI, which measures the correspondence between "Yes" retro-predictions and "Yes" observations, increased to 0.44 (0.36 for RH only). The POD, which measures successful retro-predictions only on those days on which outbreaks occurred, also improved to 0.58 (0.42 for RH only). The Hit Rate, or ratio of correct retro-predictions, remained the same, at 0.7.

For the Pacific South sub-region in April 2009, the addition of RH300 to omega300 aids significantly in the improvement of outbreak retro-prediction in this region. After having added the RH favorability column in the table, a contingency table (Table 4-3c) was created for which the criteria for retro-predicting "yes" for an outbreak on any particular day were either the favorability of omega or the favorability of RH. A series of success measures were computed based on this contingency table. All seven of the days on which outbreaks occurred had "yes" predictions, yielding a POD perfect

score of 1. The calculated CSI was 0.5, meaning that there was a 50 percent

correspondence rate of "yes" retro-predictions to "yes" observations, reflecting the fact

that there were twice as many "yes" retro-predictions as there were occurrences, given

the perfect POD. Thus, the FAR was also 0.5, an improvement from the previous (omega

only) FAR of 0.33. Although the hit rate was slightly reduced, to 0.77, the addition of

RH to omega vastly improved the detection of outbreaks when they occurred.

	Forecast	No
Occurrence	6	9
No	9	7

CSI	0.25
POD	0.4
FAR	0.6

Hit Rate	0.419355

Table 4-3a: Contingency table and success measures for South sub-region in January 2009, using *either/or* instead of both omega and T together

	Forecast	No
Occurrence	7	5
No	4	14

CSI	0.4375
POD	0.5833333
FAR	0.3636364

Hit Rate	0.7

Table 4-3b: Contingency table and accuracy measures for Central sub-region in April 2009, with non-significant UT variables added

	Forecast	No
Occurrence	7	0
No	7	16

CSI	0.5
POD	1
FAR	0.5

Hit Rate	0.766667

Table 4-3c: Contingency table and accuracy measures for Pacific South sub-region in April 2009, with non-significant UT variables added

Contrail Duration, Size, and Frequency Statistics for 2008 and 2009 Mid-season Months

In addition to verifying the contrail outbreak retro-prediction analysis for mid-season months of 2008 and 2009, satellite-based image interpretation permits me to comment upon some basic characteristics of contrail outbreaks in the six high-frequency sub-regions of the U.S. Determining outbreak attributes such as duration and size requires a high temporal resolution of images, which was not available in the 2000-2002 base dataset of contrail outbreaks. Moreover, these outbreak attributes have not previously been presented for the U.S., although such statistics for contrails (i.e., whether single or comprising larger outbreaks) are available for the trans-Atlantic flight corridors (Bakan et al. 1994).

Figures 4-1 through 4-6 show the relations between outbreak duration and outbreak frequency. Four duration classes are indicated according to the temporal sampling of the satellite imagery: 0 to 3 hours, 3 to 6 hours, 6 to 9 hours, and 9 to 12 hours. It can be seen that the great majority (87.5 percent) of outbreaks last between 0 and 3 hours. More specifically, all outbreaks over the Midwest in January 2009, the Central sub-region in April 2009, the Pacific South sub-region in April 2009, and the East sub-region in July 2008, lasted between 0 and 3 hours. For the South sub-region in January 2009, where most outbreaks (77.78 percent) lasted between 0 and 3 hours, there were two outbreaks (11.1 percent) that lasted between 3 and 6 hours, and two outbreaks that lasted between 6 and 9 hours. Similarly, in the Midwest/Upper South sub-region in October 2008, most outbreaks (65.2 percent) also lasted between 0 and 3 hours; however,

five outbreaks (21.7 percent) lasted 3 to 6 hours, two lasted 6 to 9 hours (8.7 percent),

and one lasted 9 to 12 hours (4.3 percent). Thus, most outbreaks have durations of

between zero and three hours, and the frequency of outbreaks decreases sharply with

increased duration. This result most likely reflects the fact that the UT supersaturation of

air supporting jet contrail persistence (as described in Chapter 1), only lasts for a few

hours or less. These time scales represent the larger meso-scales, and these scales are

embedded within the larger synoptic (several day, approximately 1000 km size) scales

represented by frontal systems, jet maxima, and medium-scale cyclone waves.

Figure 4-1: July 2008 East histogram showing duration vs. number of occurrences

October 2009 Midwest/Upper South

Figure 4-2: October 2008 Midwest/U South histogram showing duration vs. number of occurrences

January 2009 South

Figure 4-3: January 2009 South histogram showing duration vs. number of occurrences

January 2009 Midwest

Figure 4-4: January 2009 Midwest histogram showing duration vs. number of occurrences

April 2009 Central

Figure 4-5: April 2009 Central histogram showing duration vs. number of occurrences

April 2009 Pacific South

Figure 4-6: April 2009 Pacific South histogram showing duration vs. number of occurrences

The next set of histograms (Figures 4-7 through 4-12) show the UTC time ranges within the day during which contrail outbreaks occurred in each of the analyzed sub-regions for mid-season months of 2008-2009. This analysis demonstrates when the highest (and lowest) concentrations of outbreaks occur during any given day. The 24-hour (diurnal) time scale is divided into four 6-hr time windows: 0000 to 0600, 0600 to 1200, 1200 to 1800, and 1800 to 0000 UTC. All sub-regions, with the exception of the Pacific South (April 2009), span the Eastern and Central time zones of the U.S. This convention means that the diurnal time windows correspond, respectively, to the overnight hours, early morning, the midday hours, and afternoon/evening. The Pacific

South sub-region comprises the U.S. Pacific Time Zone, but also extends slightly into the Mountain Time Zone to the east. Therefore, the four diurnal time ranges for the Pacific South correspond, respectively, to the evening/nighttime hours, the early morning hours, late morning, and the afternoon.

Across the six sub-regions of high contrail outbreak frequencies, the 0600 to 1200 UTC time range has the least number of outbreaks. This time window corresponds roughly to early morning local time. In the South sub-region in January 2009 and the Midwest/Upper South sub-region in October 2008, this time range had the fewest number of outbreaks of the four ranges of the day. However, in the Central and Pacific South sub-regions in April 2009, the 06-12 UTC time window was tied with the 00 to 06 UTC window for the least number of outbreaks on average diurnally. Overall, the time window having the highest frequency of outbreaks was that of 18 to 00 UTC, which corresponds to around 1-7 pm local time. In all but the East sub-region in July 2008, this diurnal window had a higher number of outbreaks than each of the other three during the day. For the East in July 2008, the highest number of outbreaks occurred in the 1200 to 1800 time range, or around 8 am-2 pm local time. These results indicate that, generally, contrail outbreaks occur most during the late morning and midday hours and least during the early morning hours. This diurnal variation of contrail outbreak maxima and minima broadly agrees with previous observations that contrail occurrence peaks in the mid- to late morning hours (Minnis et al. 1997, Palikonda et al. 2004). This maximum is likely due to the high frequencies of jet aircraft flights during the morning and early afternoon hours (Minnis et al. 1997). This finding suggests that contrail observations by time of day are not only a function of UT conditions, but of commercial aircraft flight frequency.

This diurnal pattern of contrail outbreaks is slightly different from that shown for the

northern North Atlantic (Bakan et al. 1994), where most contrails occur in the afternoon

hours—at least, for the summer season studied by these authors.
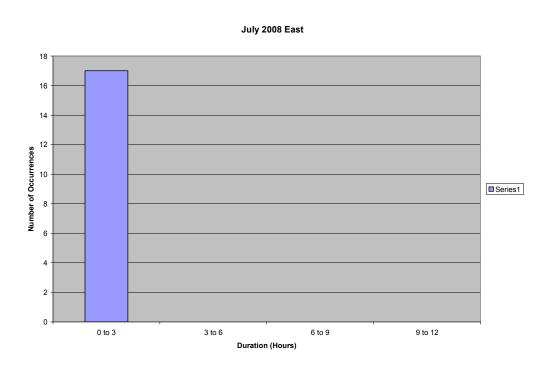
July 2008 East

Figure 4-7: July 2008 East histogram showing 6-hour time range vs. number of occurrences

October 2008 Midwest/Upper South

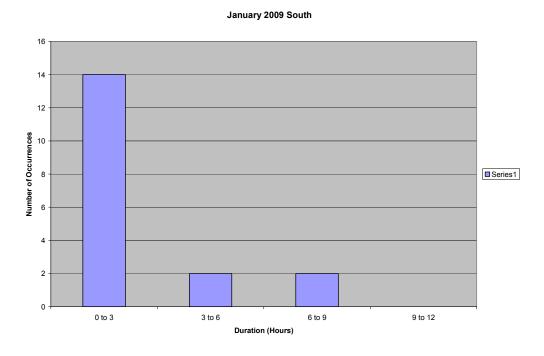Figure 4-8: October 2008 Midwest/U South histogram showing 6-hour time range vs. number of occurrences

January 2009 South

Figure 4-9: January 2009 South histogram showing 6-hour time range vs. number of occurrences

January 2009 Midwest

Figure 4-10: January 2009 Midwest histogram showing 6-hour time range vs. number of occurrences

April 2009 Central

Figure 4-11: April 2009 Central histogram showing 6-hour time range vs. number of occurrences

April 2009 Pacific South

Figure 4-12: April 2009 Pac South histogram showing 6-hour time range vs. number of occurrences

Finally, outbreak sizes are compared to their time (hour range) durations. This is to test the hypothesis that larger outbreaks are of longer duration, as increases in spatial scales of physical phenomena are directly related to increases in their temporal scales. These data are represented by categorical scatterplots (Figures 4-13 through 4-18). The height of each dot along the Y-axis represents an outbreak's size (in square km), and is plotted continuously against the discrete time ranges (0 to 3 hours, 3 to 6 hours, 6 to 9 hours, 9 to 12 hours) located along the X-axis. The South sub-region in January 2009 shows that, on average, the spatial extent of outbreaks remains the same, and possibly decreases, as time durations increase. In the Midwest/Upper South sub-region in October 2008, a similar pattern is apparent, as the average size of outbreak appears to remain

somewhat the same as temporal durations increase. The Midwest sub-region in January 2009, the Central and Pacific South sub-regions in April 2009, and the East sub-region in July 2008 have no outbreaks greater in temporal extent than the 0 to 3 hour time range. I can conclude from this analysis that contrail outbreaks that last for longer periods are not necessarily larger in spatial extent than those that last for shorter periods. This is an unexpected result due to the fact that, generally, spatial scales of natural phenomena increase with increasing temporal scales. However, it is worth considering that contrails are anthropogenic, and therefore may not possess the same physical attributes and processes that natural entities have.

Figure 4-13: July 2008 East scatterplot showing outbreak duration vs. size

October 2008 Midwest/Upper South

Figure 4-14: October 2009 Midwest/U South scatterplot showing outbreak duration vs. size

January 2009 Midwest

Figure 4-15: January 2009 Midwest scatterplot showing outbreak duration vs. size

January 2009 South

Figure 4-16: January 2009 South scatterplot showing outbreak duration vs. size

April 2009 Central

Figure 4-17: April 2009 Central scatterplot showing outbreak duration vs. size

April 2009 Pacific South

Figure 4-18: April 2009 Pacific South scatterplot showing outbreak duration vs. size

Summary

The retro-prediction for 2008 and 2009 had moderate success in the use of the statistically significant UT variables in that over 50 percent of the time—in other words, more times than not—retro-predictions for either outbreaks or no outbreaks for that day were successful when verified with the observations. The exception to this was the South region in January 2009, which had an overall success rate less than 50 percent. Conversely, the two April 2009 regions had a high rate of success as outbreaks were retro-predicted correctly over 70 percent of the time. In several cases, however, success was improved with the addition of one more non-statistically significant UT variable.

Analysis also demonstrated that most contrail outbreaks tend to last between zero to three hours, and decrease in number with increasing time; i.e., there are fewer outbreaks that last between three to six hours, and still fewer that last six to nine hours, and so on. Diurnally, outbreaks occur most frequently in late morning and afternoon. Finally, the average sizes of outbreaks remain fairly constant as time increases.

Chapter 5

Summary and Future Work

This study has addressed the primary and secondary study objectives stated in the thesis introduction: (1) to create climate diagnostics of outbreaks for six high-frequency contrail outbreak regions, and (2) to develop a prediction method for contrail outbreaks based on climatological composites.

For Objective 1, I developed climate diagnostics of contrail outbreaks for the high outbreak frequency sub-regions of the U.S. for each of the mid-season months (January, April, July, October) in 2000-2002. I achieved this by performing an objective regionalization of contrail outbreak locations using a GIS technique of overlapping 1º by 1º grid cells.

For Objective 2, I selected the six highest frequency sub-regions for climatological analysis comprising composites of UT variables (temperature, RH, omega, and zonal wind) associated with contrail outbreaks from the 15 high-frequency sub-regions identified across the four mid-season months. These composites are the average of the daily mean values for each variable on days for which outbreaks occurred within that region. I generated similar composites for non-outbreak days and also pre-outbreak days (i.e., the outbreak day minus 1) for each sub-region and mid-season month in the 2000-2002 base period. I compared outbreak to non-outbreak composite patterns for each variable using difference mapping to determine the UT map criteria (magnitude,

pattern, and gradient) for retro-predicting contrail outbreaks in the six high-frequency sub-regions.

For the retro-prediction of contrail outbreaks for the 2000-2002 study period, I compared each daily mean map (average of the 4x6 hr maps) of each UT variable with the outbreak composite. Based on the closeness of the daily map to its respective outbreak composite map, given by the magnitude, pattern, or gradient criteria, I deemed that day either favorable or unfavorable for a contrail outbreak. From these results, I computed chi-squared statistics (for statistical significance) and a number of skill score measures (hit rate, critical success index, probability of detection, false alarm rate) relating the retro-predicted outbreak favorability to the satellite-observed outbreak occurrences. I also computed one-way and two-way binary logistic regressions from the favorability versus observation results both to verify the statistical significances of each UT variable, and to yield the "best" variable(s) for each sub-region and mid-season month (which I used for an independent retro-prediction verification analysis of outbreaks in mid-season months of 2008 and 2009).

The retro-predictions I undertook for the four most recent mid-season months—July and October 2008, and January and April 2009—were for the same six high-frequency sub-regions identified for the 2000-2002 study period. I verified the retro-predictions by interpreting AVHRR multi-band satellite imagery for outbreaks, and developed skill score statistics by sub-region and mid-season month. Because the AVHRR imagery was available at fairly high temporal resolutions, I determined outbreak duration and aerial extent statistics by sub-region for the 2008 and 2009 mid-season months studied.

The regionalization analysis shows that, in general, the geographic distribution and frequency of contrail outbreaks is broadly similar to that shown in previous studies that relied upon a subjective regionalization (e.g., DeGrand et al. 2000; Carleton et al. 2008). However, I showed that there is a variation in location of the high-frequency outbreak regions by mid-season month (2000-2002). This variation indicates a dependence on the seasonal movement of the subtropical and polar front jet streams. The maps of contrail outbreaks by mid-season month for 2000-2002 compare very closely with those for individual contrails presented by Travis et al. (2007). Overall, the eastern half of the conterminous U.S. experiences many more outbreak events than does the western half of the country (e.g., Carleton et al. 2008), as more warm air advection associated with a moister UT generally is available (Travis et al. 2006). In the West, April is the most significant mid-season month for outbreaks in the Pacific states. As found in previous studies (Changnon 1981; DeGrand et al. 2000; Carleton et al. 2008), the Midwest has the highest frequency of outbreaks and is the region of absolute maximum occurrence in three out of four mid-season months.

The two UT variables most frequently associated with contrail outbreaks in the high-frequency sub-regions are RH300 and omega300, probably due to the fact that they are inter-related variables. Greater UT moisture—as given by the higher values of relative humidity—is required for persisting contrails and outbreaks (e.g., Carleton et al. 2008), and this requirement is most likely to be met in areas where air is rising (negative omega). Relative humidity patterns were significant in most of the mid-season months and corresponding regions, with increased moisture taking place on outbreak days. Likewise, omega (vertical velocity) in general behaved as expected across mid-season

months with changes from subsidence to ascendance, or a decrease in values of subsidence from east to west. Overall, the composites agree with the previous findings that contrail outbreaks accompany environments that are increasing in moisture, zonal wind speed, and horizontal wind shear, and changing vertical motion from subsidence to ascendance, as mentioned in the thesis introduction.

The Chi-squared tests and the binary logit model give similar results as to those UT variables most closely associated with contrail outbreaks (by sub-region and mid-season month). This indicates a robustness to the results and increases the confidence with which they could be used in the retro-prediction verification study for 2008-09 (Chapter 4). This robustness occurred despite the fact that the individual UT variables and variable combinations found to be statistically significant appeared to be somewhat random, as the comparison of similar regions (e.g., Midwest in January, Central in April) and similar seasons (e.g., the transition seasons of April and October) had very different criteria as to which variables were most accurate for use in predicting outbreaks. RH and omega, however, were most frequent among the significant UT variables in terms of their favorability, or, in other words, their similarities in daily mean conditions to those of the outbreak composites for 2000-2002.

In terms of the retro-prediction applied to the 2008 and 2009 mid-season months using the criteria of variable magnitude, pattern, or horizontal gradient on the composite analyses, the measures suggest moderate success. The April regions (Central and Pacific South) had the most success when predicting outbreaks in terms of composite similarity, while the South region in January performed the least success with this method.

The retro-prediction for 2008 and 2009 had moderate success in the use of the statistically significant UT variables in that over 50 percent of the time; in other words, more times than not, retro-predictions for either outbreaks or no outbreaks for that day were successful when verified with the observations. The exception to this rule was the South region in January 2009, which had an overall success rate less than 50 percent. Conversely, the two April 2009 regions had a high rate of success as outbreaks were retro-predicted correctly over 70 percent of the time. In several cases however, success was improved with the addition of one more non-statistically significant UT variable.

Future Work

As a first step towards the eventual successful forward prediction (i.e., forecasting) of contrail outbreaks in real-time, this study demonstrates the validity of a simple visual (i.e., subjective) technique of comparing composite UT maps—a synoptic climatology—with daily-averaged UT maps for retro-prediction purposes. Although certain of the retro-prediction verification results (by sub-regions in 2008-09 mid-season months) were mixed, others were more successful. Further verification studies are needed to refine the method developed here, and to make it suitable for forward prediction of contrail outbreaks in near-real time. This future work might involve adding contrail outbreak data for the mid-season months of 2003-2006, along with its associated UT conditions, to comprise a considerably longer-term (7-year) synoptic climatology. Additional information about contrail outbreaks that may help improve the ability to retro-predict and ultimately forecast their occurrence concerns their vertical extent. The

associated UT conditions could be assessed by analyzing temperature and/or dew point lapse rates for atmospheric layers on thermodynamic diagrams for available sounding data. Finally, there is a continuing need to determine the climatic impacts of contrail outbreaks, particularly on surface diurnal temperatures ranges and their interactions with other anthropogenic forcings of climate change.

References

Bakan, S., M. Betancor, V. Gayler, and H. Grassl, 1994: Contrail frequency over Europe from NOAA-satellite images. *Ann. Geophys*., 12, 962-968

Carleton, A.M. and P.J. Lamb, 1986: Jet Contrails and Cirrus Cloud: A Feasibility Study Employing High-Resolution Satellite Imagery. *Bulletin of the American Meteorological Society*, 67, 301-309.

Carleton, A.M., D.J. Travis, K. Master, and S. Vezhapparambu: 2008. Composite Atmospheric Environments of Jet Contrail Outbreaks for the United States. *Journal of Applied Meteorology and Climatology*, 47, 641-667.

Changnon, S.A., 1981: Midwestern cloud, sunshine and temperature trends since 1901: Possible evidence of jet contrail effects. *Journal of Applied Meteorology*, 20, 496-508.

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. *Journal of Applied Meteorology*, 39, 1434-1459.

Duric, D. *Weather Analysis*. Upper Saddle River, NJ: Prentice Hall, Inc., 1994.

Garson, D. G., 2009: "Logistic Regression." *Quantitative Research in Public Administration*. <http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm>

Haby, J. "The 300 / 200 mb chart." *The Weather Prediction* <http://www.theweatherprediction.com>

Hammond, R. and McCullagh, P. *Quantitative Techniques in Geography, An Introduction*. Oxford, UK: Clarendon Press, 1978.

Hanson, H.M. and D.M. Hanson, 1995: A Reexamination of the Formation of Exhaust Condensation Trails by Jet Aircraft. *Journal of Applied Meteorology*, 34, 2400-2405.

Kastner, M., R. Meyer, and P. Wendling, 1999: Influence of weather conditions on the distribution of persistent contrails. Meteo. Appl., 6, 261-271.

Mearns, L.O., F. Giorgi, L. McDaniel, and C. Shields, 2003: Climate scenarios for the southeastern U.S. based on GCM and regional model simulations. *Climatic Change*, 60, 7-35.

Minnis, P., J.K. Ayers, R. Palikonda, and D. Phan, 2004: Contrails, Cirrus Trends, and Climate. *Journal of Climate*, 17, 1671-1685.

Minnis, P., J.K. Ayers, and S.P. Weaver, 1997: Surface-based observations of contrail occurrence frequency over the US, April 1993 – April 1994. NASA Reference Publication 1404.

Neter, J., W. Wasserman, and M.H. Kutner. *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*. Homewood, IL: Richard D. Irwin, Inc., 1985.

Palikonda, R., P. Minnis, D.P. Duda, and H. Mannstein, 2005: Contrail coverage derived from 2001 AVHRR data over the continental United States of America and surrounding areas. *Meteorologische Zeitschrift*, 14, 525-536.

Schrader, M.L., 1997: Calculations of Aircraft Contrail Formation Critical Temperatures. Notes and Correspondence, *Journal of Applied Meteorology*, 36, 1725-1729.

Stanski, H.R., L.J. Wilson, and W.R. Burrows, 1989: *Survey of common verification methods in meteorology*. World Weather Watch Tech. Rept. No.8, WMO/TD No.358, WMO, Geneva, 114 pp.

Stockburger, D. W, 1996: "Chi-square and tests of contingency tables." *Introductory Statistics: Concepts, Models, and Applications*.
<http://www.psychstat.missouristate.edu/introbook/sbk28.htm>

Travis, D.J., A.M. Carleton, and S.A. Changnon, 1997: An Empirical Model to Predict Widespread Occurrences of Contrails. Journal of Applied Meteorology, 36, 1211-1220.

Travis, D.J., A.M. Carleton, and R.G. Lauritsen, 2004: Regional Variations in U.S. Diurnal Temperature Range for the 11-14 September 2001 Aircraft Groundings: Evidence of Jet Contrail Influence on Climate. Journal of Climate, 17, 1123-1134.

Travis, D.J., A.M. Carleton, J.S. Johnson, and J.Q. DeGrand, 2006: U.S. Jet contrail frequency changes: influences of jet aircraft flight activity and atmospheric conditions. International Journal of Climatology, 27, 621-632.

Volmel, H. "Saturation vapor pressure formulations." CIRES, University of Colorado, Boulder, CO. 27 November 2006. <http://cires.colorado.edu/~voemel/vp.html>

Wallace, J.M. and P.V. Hobbs. *Atmospheric Science: An Introductory Survey*. San Diego: Academic Press, 1977.

Wilks, D. S. *Statistical Methods in the Atmospheric Sciences*. San Diego: Academic Press, 1995.

Williams, V. R.B. Noland, and R. Toumi, 2003: Air transport cruise altitude restrictions to minimize contrail formation. Climate Policy, 3, 207-219.

Appendix A

Composites of 2000-2002 Pre-Outbreak Days

Figure A1: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

January South 300mb Omega & U Wind, Pre-Outbreak

GrADS: COLA/IGES

Figure A2: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s)

April Central 300mb Temp & RH, Pre-Outbreak

GrADS: COLA/IGES

Figure A3: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

April Central 300mb Omega & U Wind, Pre-Outbreak

Figure A4: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

April Pacific South 300mb Temp & RH, Pre-Outbreak

Figure A5: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

April Pacific South 300mb Omega & U Wind, Pre-Outbreak

GrADS: COLA/IGES

Figure A6: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s)

July East 300mb Temp & RH, Pre-Outbreak

GrADS: COLA/IGES

Figure A7: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

July East 300mb Omega & U Wind, Pre-Outbreak

Figure A8: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s)

October Midwest/UpperSouth 300mb Temp & RH, Pre-Outbreak

Figure A9: 2000-2002 composites of temperature (colored, Cº) and RH (black, %).

Figure A10: 2000-2002 composites of omega (colored, Pa/s) and zonal wind (black, m/s).

Appendix B

2000-2002 UT Favorability vs. Outbreak Occurrence Tables

2000	RH	Omega	T	U Wind
1	X		X	
2	X	X		
3	X	X		X
4				X
5			X	
6		X		
7				
8		X		X
9		X		
10				
11				
12	X			
13			X	
14			X	
15	X		X	
16	X		X	
17	X	X	X	
18			X	
19		X		
20				
21				
22	X	X		X
23				
24				
25			X	
26				
27				
28	X	X		
29	X	X		
30				
31				
2001				
1				
2				
3		X	X	
4			X	
5	X		X	
6			X	

7				
8				
9			X	
10		X	X	
11	X	X		X
12	X			
13				
14	X	X		
15				
16				
17	X	X		
18		X		
19	X			X
20				
21				
22			X	
23	X	X		
24				X
25				
26	X	X		
27				
28	X	X		X
29	X	X	X	X
30			X	
31				
2002				
1				
2				
3			X	
4	X		X	
5	X	X		X
6	X		X	X
7			X	
8	X		X	
9	X		X	
10	X	X	X	
11				
12			X	
13				
14				
15				
16	X	X		
17		X		
18				
19		X		X
20	X		X	
21				X
22				X
23				X

	RH	Omega	T	U Wind
24	X	X		X
25				X
26			X	
27	X		X	
28	X	X	X	
29	X	X	X	
30	X	X	X	
31	X	X	X	

Table B1: Favorability table (X = favorable; highlight = outbreak) for January Midwest sub-region

2000	RH	Omega	T	U Wind
1				
2		X		
3		X		
4	X			
5				X
6	X	X		
7				
8				
9				
10	X			X
11	X			X
12	X			X
13	X			
14	X			
15		X		
16		X	X	
17	X		X	
18		X	X	
19			X	
20		X	X	X
21		X	X	X
22	X		X	
23	X			
24	X		X	
25			X	X
26				X
27	X	X		X
28	X		X	X
29			X	X
30			X	X
31			X	X
2001				
1	X		X	X
2	X	X	X	X
3	X		X	X
4			X	X

Day				
5			X	
6		X		
7			X	
8	X		X	X
9	X		X	X
10	X	X	X	X
11			X	X
12				X
13				
14				
15	X	X		
16		X		
17	X	X		X
18		X		
19		X		
20				
21			X	
22			X	X
23			X	X
24			X	X
25	X			X
26	X	X	X	X
27	X		X	X
28	X		X	X
29		X		
30	X	X		
31				X
2002				
1	X		X	X
2			X	X
3	X		X	X
4	X		X	X
5		X	X	X
6				X
7	X			X
8			X	X
9		X	X	
10		X		
11	X	X		
12	X	X		X
13	X		X	X
14	X			X
15				
16		X		
17		X		
18				
19				
20				
21				

	RH	Omega	T	U Wind
22	X	X		
23		X		
24		X		
25				
26	X			
27	X			
28				
29		X		
30		X		
31		X		

Table B2: Favorability table (X = favorable; highlight = outbreak) for January South sub-region

2000	RH	Omega	T	U Wind
1		X	X	
2				
3				
4				X
5	X	X	X	X
6	X	X	X	
7	X		X	X
8			X	X
9		X	X	X
10	X	X	X	
11			X	
12			X	
13	X		X	X
14	X		X	X
15	X	X	X	X
16	X	X	X	X
17			X	X
18	X			X
19	X	X		X
20	X			X
21	X		X	X
22				X
23		X	X	X
24			X	X
25	X	X	X	X
26			X	
27	X	X	X	X
28			X	X
29		X	X	X
30		X	X	X
2001				
1			X	X
2		X	X	
3			X	X

Day	1	2	3	4
4	X		X	X
5	X	X		X
6				
7				X
8				X
9		X	X	X
10				X
11		X		X
12	X			
13	X		X	
14	X		X	
15	X		X	
16			X	
17			X	X
18	X		X	X
19		X		X
20				X
21				X
22				X
23				X
24			X	
25			X	X
26	X	X	X	X
27	X	X	X	X
28		X	X	X
29				X
30		X		X
2002				
1	X	X	X	
2	X		X	
3			X	
4			X	
5			X	
6	X	X	X	X
7		X	X	X
8				X
9				X
10	X	X	X	X
11	X	X	X	X
12			X	X
13			X	X
14			X	X
15	X	X	X	X
16		X		X
17				X
18				X
19				
20				
21				

	RH	Omega	T	U Wind
22				
23		X		
24				X
25				
26				
27				
28	X			
29		X		
30		X		

Table B3: Favorability table (X = favorable; highlight = outbreak) for April Central sub-region

2000	RH	Omega	T	U Wind
1				
2				
3			X	X
4		X	X	X
5	X		X	X
6				
7	X			
8		X		
9	X			X
10				
11	X			
12	X	X		
13		X		
14				
15	X			
16			X	
17			X	
18			X	
19	X			
20		X		
21				
22	X			
23				X
24				X
25				X
26			X	X
27		X	X	X
28	X			
29				
30	X			
2001				
1	X	X	X	X
2				X
3				
4				

Day	A	B	C	D
5				
6	X			
7				
8				
9				
10				
11				
12	X			
13	X			X
14				X
15	X			X
16	X	X	X	X
17			X	
18		X	X	
19				
20				
21				
22				
23				
24				X
25				X
26	X	X	X	
27			X	
28	X			
29	X			
30			X	X
2002				
1			X	
2				
3				
4				
5	X			
6				
7				
8		X		
9	X	X		X
10			X	X
11			X	X
12				X
13	X			
14			X	X
15			X	
16				X
17				X
18				X
19				
20	X			
21	X			
22				

	RH	Omega	T	U Wind
23		X		
24		X	X	
25				
26				
27				
28		X		X
29			X	X
30				

Table B4: Favorability table (X = favorable; highlight = outbreak) for April Pacific South sub-region

2000	RH	Omega	T	U Wind
1			X	
2			X	
3			X	
4	X		X	
5			X	
6			X	
7			X	
8		X	X	
9			X	X
10			X	X
11	X	X	X	X
12	X		X	
13			X	
14			X	
15			X	
16			X	
17			X	
18			X	
19		X	X	X
20		X	X	X
21			X	
22			X	
23	X		X	
24			X	
25			X	
26			X	
27	X		X	
28			X	
29	X		X	X
30			X	X
31			X	X
2001				
1	X		X	X
2			X	
3			X	
4			X	

Day				
5			X	
6			X	
7		X	X	
8		X	X	
9			X	X
10			X	
11			X	
12				
13			X	
14				
15		X	X	
16	X	X	X	
17			X	
18			X	
19			X	
20			X	
21		X	X	
22			X	
23			X	
24			X	
25			X	
26		X	X	
27			X	X
28			X	X
29			X	
30			X	
31			X	X
2002				
1			X	
2	X			
3	X			
4	X			
5				
6		X		
7				
8			X	
9			X	
10			X	
11			X	
12	X		X	
13	X		X	
14	X	X	X	
15	X	X	X	
16	X		X	
17	X	X	X	
18			X	
19			X	
20	X		X	
21		X	X	

	RH	Omega	T	U Wind
22			X	X
23				X
24			X	X
25				X
26				X
27			X	X
28			X	X
29				X
30	X			X
31	X	X		X

Table B5: Favorability table (X = favorable; highlight = outbreak) for July East sub-region

2000	RH	Omega	T	U Wind
1	X	X		X
2	X	X		
3	X			
4				
5				
6				
7		X		
8		X		
9	X	X		X
10		X		X
11		X		X
12	X	X		X
13	X		X	X
14			X	
15				
16				
17				
18		X	X	X
19			X	X
20			X	X
21			X	X
22		X	X	X
23	X		X	X
24			X	X
25			X	X
26	X		X	X
27			X	X
28		X		X
29	X			X
30				X
31	X			X
2001				
1			X	X
2		X	X	X

Day				
3			X	
4	X		X	
5			X	
6				
7		X		
8	X			
9	X		X	X
10			X	X
11				
12				X
13				X
14			X	X
15				
16				
17	X	X		
18	X	X	X	
19				
20		X		
21			X	
22	X			
23				
24				
25		X		
26		X		
27		X		
28		X	X	X
29	X		X	X
30	X		X	X
31		X		X
2002				
1				
2	X		X	
3			X	
4				
5	X	X		
6		X		
7		X		
8				
9				
10				
11		X		X
12	X	X	X	
13		X		
14		X		
15				
16				
17		X		
18	X			
19				

153

20		
21	X	
22		
23		X
24		
25		
26		
27		
28		
29		
30	X	X
31		X

Table B6: Favorability table (X = favorable; highlight = outbreak) for October Midwest/Upper South sub-region

Appendix C

2008 and 2009 Retro-Prediction Tables with Non-Significant UT Variable Favorability

2009	Omega	T	U	RH
1	X			X
2				
3	X	X		
4	X	X		
5				
6				
7		X	X	
8		X		
9		X		
10	X			X
11				
12				
13			X	
14				
15				
16				
17	X			X
18		X		
19		X		
20		X		
21		X		X

2009	Omega	T	RH	U
22		X		X
23		X		X
24				
25	X			X
26	X			
27	X			X
28				X
29				
30				
31		X		

Table C1: Retro-prediction table for January 2009 Midwest sub-region showing favorability of non-statistically significant UT variables

2009	Omega	T	RH	U
1		X	X	X
2		X		
3				
4				X
5	X			
6				
7				
8				
9	X			
10				
11				
12				
13			X	X
14		X		
15		X	X	X
16		X		
17	X	X		
18		X		
19		X		X
20		X		
21	X	X		
22		X		
23	X	X		
24				
25				
26				
27				
28				

2009	RH	U	T	Omega
29				
30				
31	X	X		

Table C2: Retro-prediction table for January 2009 South sub-region showing favorability of non-statistically significant UT variables

2009	RH	U	T	Omega
1			X	
2			X	
3			X	
4	X		X	
5	X			
6	X		X	
7		X	X	X
8			X	X
9			X	
10		X	X	
11	X		X	X
12		X	X	
13	X	X	X	
14		X	X	
15				X
16		X	X	
17		X	X	
18		X	X	
19		X		
20		X	X	
21		X	X	
22		X	X	X
23		X	X	
24		X	X	
25	X			
26	X	X		
27		X		
28				

	Omega	RH
29		X
30		X

Table C3: Retro-prediction table for April 2009 Central sub-region showing favorability of non-statistically significant UT variables

2009	Omega	RH
1		
2		
3		
4		X
5		
6		
7	X	
8		
9	X	X
10		
11		X
12		X
13		X
14		
15		
16		X
17		
18		
19		X
20		X
21		X
22		
23		X
24	X	
25		
26		
27		X
28		

29
30   X

Table C4: Retro-prediction table for April 2009 Pacific South sub-region showing favorability of non-statistically significant UT variables

2008	RH	T	Omega	U
1		X		X
2		X		X
3		X		
4		X		
5	X	X		X
6	X		X	X
7			X	
8				
9	X			
10			X	
11	X		X	
12	X		X	
13	X			
14				
15		X	X	
16			X	
17			X	
18			X	
19				
20				
21	X			
22				
23	X			
24				
25	X			
26			X	
27			X	
28		X	X	

29	X		X	X
30				
31				

Table C5: Retro-prediction table for October 2008 Midwest/Upper South sub-region showing favorability of non-statistically significant UT variables

2008	RH	T	UWind	Omega
1			X	
2	X	X	X	
3	X			
4				
5	X			
6	X	X		
7		X	X	
8		X	X	
9		X		
10	X	X		
11			X	
12	X		X	
13			X	
14		X		
15		X		
16				
17				
18				
19	X		X	
20				
21				X
22	X			
23				
24				
25				
26		X		
27	X			

28	X			X	
29					X
30			X	X	
31	X		X	X	

Table C6: Retro-prediction table for July 2008 East sub-region showing favorability of non-statistically significant UT variables