2000

Statistical stratus forecasts for San Francisco International Airport

John Amirault Farley
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Farley, John Amirault, "Statistical stratus forecasts for San Francisco International Airport" (2000).
Master's Theses. 1989.
DOI: <https://doi.org/10.31979/etd.3hs7-vvkh>
https://scholarworks.sjsu.edu/etd_theses/1989

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

STATISTICAL STRATUS FORECASTS
FOR SAN FRANCISCO INTERNATIONAL AIRPORT

A Thesis

Presented to

The Faculty of the Department of Meteorology

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Masters of Science

by

John Amirault Farley

May 2000

UMI Number: 1399793

Copyright 2000 by
Farley, John Amirault

All rights reserved.

UMI[®]

UMI Microform 1399793

Copyright 2000 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2000

John Amirault Farley

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF METEOROLOGY

A handwritten signature in cursive script, appearing to read "Douglas M. Sinton", written above a horizontal line.

Dr. Douglas M. Sinton

A handwritten signature in cursive script, appearing to read "Jindra Goodman", written above a horizontal line.

Dr. Jindra Goodman

A handwritten signature in cursive script, appearing to read "F. Wesley Wilson", written above a horizontal line.

Dr. F. Wesley Wilson
(Massachusetts Institute of Technology, Lincoln Laboratory)

APPROVED FOR THE UNIVERSITY

A handwritten signature in cursive script, appearing to read "William Fish", written above a horizontal line.

ABSTRACT

STATISTICAL STRATUS FORECASTS FOR SAN FRANCISCO INTERNATIONAL AIRPORT

by John A. Farley

San Francisco International Airport (SFO) frequently operates below capacity for arriving aircraft. In the warmer months this is primarily due to stratus in the final approach to SFO. Using 1996 and 1997 (training) data from May to October, nine linear regression models are developed to forecast the time of stratus burn off: 3 each for initialization times of 1400 UTC, 1500 UTC, and 1600 UTC. These models are then tested on 1998 (independent) data. The models that don't use visible satellite data as input perform satisfactorily. The models that use visible satellite data perform best on the training data but poorest on the independent data. Two likely causes are discussed: 1) "drift" in the visible data on all GOES satellites, and 2) GOES-9 mechanical faults, which produced bad data during the summer of 1998. It is expected that the models that use visible data will once again perform well when calibrations have been made for the "drift."

ACKNOWLEDGMENTS

I begin by thanking the person that made this project possible, my wife Maura Farley. Without her encouragement it is likely that I never would have entered the field of meteorology. I can only begin to communicate the intense feeling of appreciation that I have to Maura for supporting, putting up with, encouraging, and inspiring me in this academic pursuit. A big thank you is also due to my children, John, Rosemary, and especially Grace, for tolerating a sometimes overworked and cranky dad.

My time at SJSU has exceeded any of my expectations, principally because of the great faculty and students in the Meteorology Department. A big thank you to Dr. Peter Lester for advising and then supporting me financially in the beginnings of this thesis project. I am significantly indebted to Dr. Douglas Sinton, my thesis advisor, for all of his work and support over the last 5 years. He has been a rock-solid "sounding board" on academic as well as other issues.

His selfless version of "trickle down economics" enabled my family and me to continue living here in the Bay Area.

I would like to thank Dr. Wes Wilson for his hard work and financial support in the Bay Area Stratus Initiative. He is the hardest working scientist that I have ever met. His drive and superior math mind are the significant contributors to the success of the forecast method presented in this thesis.

A special thank you goes to my peers in the department, most notably William Chan, Brian Kahn, and Jeff Van Dorn. Without their help it is likely that I would have needed to repeat a course or two, especially the programming ones. A final thank you is in order for Dr. Goodman, Dr. Bridger, and Dr. Steffens for continuing to encourage thinking about the atmosphere's many unknowns.

TABLE OF CONTENTS

A. Background

1. Aviation problem, ceiling and visibility requirements at San Francisco International Airport (SFO).....1
2. San Francisco Bay Area stratus cycle.....6
3. Climatology of approach zone reduced visibility..9

B. Stratus

1. Global climatology.....11
2. Synoptic influences on California coastal stratus.....14
3. How stratus forms - current theories.....17
4. Microstructure of California coastal stratus and fog.....22
 - a. Thickness of cloud layer.....24
 - b. Temperature profile.....24
 - c. Wind observations.....26
 - d. Relative humidity.....27
 - e. Droplet size.....27
 - f. Liquid water content.....33
 - g. Condensation and cloud condensation nuclei...33

C. Statistical Forecast Models

1. Introduction.....35
2. Methodology.....38

a. Overview.....	38
b. First attempts.....	40
c. Final models.....	42
3. Zi-Ceiling models.....	48
4. Cloud models.....	52
5. Combo models.....	57
6. Model verification.....	61
7. Zi-Ceiling model verification analysis.....	63
8. Cloud and Combo model verification analysis.....	63
9. Summary of results.....	76
10. Conclusions.....	76
APPENDIX A: List of candidate predictors.....	78
APPENDIX B: Cloud products.....	81
APPENDIX C: Model graphs.....	90
References.....	96

LIST OF TABLES

	Table Description	Page
Table 1	The ten stratus regions globally, their geographical extent, the season and amount of maximum stratiform cloudiness, the month and value of the maximum static stability, the season and amount of minimum stratiform cloudiness, and the month and value of minimum static stability.....	13
Table 2	Forecast model characteristics.....	61
Table 3	Model results on 1998 data.....	62
Table 4	Average Mean Brightness for 1500 UTC with side-bys beginning between 1924 and 2000 UTC.....	69
Table 5	Average Mean Brightness for 1500 UTC with side-bys beginning between 1728 and 1800 UTC.....	69

LIST OF FIGURES

	Figure Description	Page
Fig. 1	Map of SFO.....	2
Fig. 2	Map of San Francico Bay Area.....	3
Fig. 3	Normal arrival and departure patterns at SFO.....	3
Fig. 4	Typical weekly arrival demand at SFO.....	5
Fig. 5	Normal stratus cycle during the warmer months in the San Francico Bay Area.....	8
Fig. 6	Typical stratus burnoff during the warmer months in the San Francisco Bay Area.....	8
Fig. 7	Percent of side-by-side landings that that began after given time.....	10
Fig. 8	Average sea level pressure in July.....	14
Fig. 9	Mean surface winds for July.....	15
Fig. 10	Variable inversion base height.....	17
Fig. 11	Stratus lowering.....	20
Fig. 12	Temperature and relative humidity profile in marine boundary layer.....	25

Fig. 13	Vertical profiles of mean droplet radius in stratus 0-40 nmi offshore from Vanden- berg AFB.....	29
Fig. 14	Air parcels trajectories for 24, 48, and 72 hours.....	30
Fig. 15	Liquid water content, number of particles, and mean radius of stratus droplets.....	32
Fig. 16	Variation in condensation nuclei as a func- tion of time.....	34
Fig. 17	The San Francisco Bay Area broken up into sectors.....	56
Fig. 18	Visible imagery taken by GOES-8 and GOES-9.....	67
Fig. 19	Visible image take by GOES-9 on July 15, 1998 over Baja California.....	72
Fig. 20	Pre-processed visible image taken by GOES-9 on June 13, 1998 at 1315 UTC over the San Francisco Bay Area.....	73
Fig. 21	Pre-processed visible image taken by GOES-9 on June 13, 1998 at 1600 UTC.....	74
Fig. 22	Filtered visible image taken by GOES-9 on June 13, 1998 at 1600 UTC.....	75

Fig. 23	80x80 km satellite data grid centered over San Francisco Bay.....89
---------	--

A. Background

1. Aviation Problem, ceiling and visibility requirements at San Francisco International Airport

San Francisco International Airport (SFO) has two parallel runways, 28L and 28R, that are the primary sites for landing aircraft. These runways are about 250 meters apart and are used for simultaneous landings when visibility is good in the area just upstream of the runways. Arriving planes typically approach from the southeast through an area known as the approach zone. The approach zone ends where the runways begin and extends about 8 km to the southeast from the surface to roughly 900 m above sea level. (See Figure 1 for a map of SFO. Figure 2 shows SFO and the approach zone in relation to the San Francisco Bay Area.) Good visibility is defined by the side by side aircraft's ability to "see" each other throughout the approach zone. Excluding other factors, such as excessive wind, when visibility is excellent, SFO has a daytime acceptance rate of about 60, i.e., 60 aircraft can land per hour. When visibility is poor, i.e., the aircraft can't "see" each other, the acceptance

rate drops to about 30. Often visibility is good enough for parallel approaches, but not excellent, in which case the acceptance rate is about 45. Figure 3 illustrates the arrival and departure patterns at SFO.

Fig. 1. Map of SFO. The typical arrival runways are 28R and 28L, which run from the lower right (SE) to the upper left (NW). The planes arrive through the approach zone, which extends to the lower right (SE) over San Francisco Bay.

Fig. 2. Map of San Francisco Bay Area with approach zone highlighted in red.

Fig. 3. Typical arrival and departure patterns at SFO (Strach, W.J., 1991).

There are two daily peaks in airline arrivals at SFO: one between 0900 and 1300 PDT, and the other between 1900 and 2000 PDT (see Figure 4). During these peaks, or so-called rushes, well over 30 aircraft per hour are scheduled to arrive. If visibility is poor during either of these periods, air traffic may be significantly affected throughout the entire U.S. Air Traffic Control System (ATCS). If the forecast is for poor visibility during a rush, then a national ground delay program is initiated by ATCS, where aircraft bound for SFO are systematically delayed at their origins. Conversely, no national ground delay program is initiated if it is believed that the visibility will be good during a rush. Although delays are part of normal operations in the air traffic system, they are costly and should therefore be invoked only when conditions warrant.

SFO is the fifth busiest airport in the country rated by number of passengers. It is also consistently the airport with the highest incident of weather related delays that result in the implementation of ground delay programs (D. A. Clark, F. W. Wilson, 1996). The majority of these ground

delay decisions are centered around the short-term ceiling and visibility forecasts for the approach zone. Clearly, accurate visibility forecasts for the approach zone will minimize the adverse affects of a ground delay program. Conversely, inaccurate forecasts have costly repercussions for the U.S. ATCS, i.e., aircraft delayed at take off, or air-borne aircraft forced into holding patterns.

Fig. 4. Typical weekday arrival demand in one-hour intervals at SFO, (Evans, et. al, 1999).

Reduced visibility in the approach zone is typically caused by low clouds, although haze can also be a factor. During the colder months, there are two main sources for the low clouds in the approach zone: 1) transient cyclones, 2) radiation fog which either forms in the approach zone or advects in from the east (Livermore and Sacramento valleys).

During the warmer months, from about late April to early October, the most frequent cause of reduced visibility in the approach zone is the presence of stratus. This research focuses on the issues surrounding forecasting of stratus burn-off in the approach zone during the morning rush from May to October.

2. San Francisco Bay Area Stratus Cycle

During the warmer months the stratus in the San Francisco Bay Area has a fairly regular daily cycle (although there can be significant deviations from this typical cycle). See Figure 5. The stratus resides at the top of the marine boundary layer while the bottom of the layer remains cloud free. The marine boundary layer is capped by a strong semi-permanent inversion (this is covered in more detail in chapter B.3.). The stratus advects in from the Pacific Ocean overnight and evaporates by afternoon because the morning solar radiation causes a positive net radiative flux convergence in the marine boundary layer. Although a small fraction of solar radiation is absorbed by the cloud, it isn't

enough to evaporate the stratus. The primary mechanisms responsible for stratus burn-off are long-wave radiation and entrainment. That is, in the morning the ground absorbs the solar radiation and in turn emits long-wave radiation into the boundary layer where it is absorbed by the cloud as well as the air below the cloud. As this process continues, the warmer marine boundary layer supports a thinner and thinner cloud until there are visible holes in the stratus layer. Typically the first areas to clear are over the land on either side of the San Francisco Bay. These clear spaces contain relatively dry and warmer air, which then gets entrained into the surrounding boundary layer and eventually the stratus evaporates everywhere inland from the coast. The typical occurrence from a humidity context is that the LCL rises above the base of the inversion making it impossible for a cloud layer to exist. Occasionally the base of the inversion drops below the LCL, but this is much less common. The typical evaporation process is illustrated in Fig. 6.

Fig. 5. Typical stratus cycle during the warmer months in the San Francisco Bay Area. The downward arrows represent the capping inversion. The left-to-right arrows are advection from the west. The upward arrows show the rising warm air in California's Central Valley. (Evans, et. al, 1999).

Fig. 6. Typical stratus burn-off during the warmer months in the San Francisco Bay Area. (Source: MIT Lincoln Labs, Lexington, MA).

3. Climatology of approach zone reduced visibility

There is no long-standing archive of surface observations in the approach zone. However, SFO is on the edge of the approach zone and SFO observes at least some stratiform clouds nearly 50% of the days from May to October (Western Regional Climate Center). From 1996 to 1998, roughly 40% of the time there was sufficient low cloudiness in the mornings that a burn-off forecast needed to be issued (archive kept by Center Weather Service Unit (CWSU) of the Air Route Traffic Control Center in Fremont, CA).

The reduced visibility in the approach zone is a significant problem for the U.S. ATCS. Data were compiled from roughly 1100 cases from May to October from 1987-1993 and are presented in Figure 7 as the frequency of commencement times for side-by-side landings (CWSU Archive). Most interesting to note is that nearly 40% of the time parallel landings began after 1700 UTC (60% began by 1700 UTC). By 1800 UTC, 2 hours into the rush, 25% of the time parallel landings had not yet begun. Obviously some of the mornings with reduced

capacity involve transient cyclones, haze, or excessive wind, but climatology indicates that the majority of these cases involve typical summertime stratus.

Fig. 7. Percent of side-by-side landings that began by the given time, e.g., roughly 25% of the time side-by-side landings began after 1800 UTC. Cases from May to October 1987-1993.

B. Stratus

1. Global Climatology

The low clouds discussed in this section are stratus, stratocumulus, fractocumulus, fractostratus, and sky obscuring fog. For simplicity, stratus is defined as including all of these clouds. On average 29% of the earth's surface is covered by stratus. Stratus is more extensive over the oceans than over the continents; on average 34% of the oceans are covered by stratus, compared with an average of 17.5% for the continents (Klein and Hartmann 1993).

Stratus is least abundant on the western side of the subtropical oceans. Stratus is most abundant on the eastern side of subtropical oceans, which coincides with the eastern side of subtropical anticyclones. These regions are characterized by relatively cold sea surface temperatures (SST), above which a well-mixed marine boundary layer persists with an average depth of about 1000 meters. The air temperature in the boundary layer deviates little from the underlying SST and the wind in the boundary layer blows toward the equator with speeds ranging about 5 to 10 m s⁻¹. Capping the boundary layer is a strong inversion that is maintained,

at least in part, by persistent subsidence from the large subtropical anticyclones. The largest and most persistent areas of stratus coincide with regions of high static stability in the lower troposphere (Klein and Hartmann 1993), where static stability or measure of the inversion strength is calculated by subtracting the average surface potential temperature from the average 700 mb potential temperature. Other areas of significant amounts of stratus are the mid-latitude and Arctic oceans.

During the warmer months the stratus coverage reaches its maximum. In June, July, and August (JJA) stratus covers more than 60% of the subtropical oceans off the coasts of California, Peru, and Namibia. More than 60% coverage is also observed in the North Atlantic and in the Arctic, while in the North Pacific more than 80% coverage is observed during JJA. In December, January, and February more than 60% of the circumpolar ocean south of about 60 degrees South Latitude is covered by stratus. Globally, there are 10 regions of extensive stratus coverage. Table 1 categorizes these regions according to location, season and amount of maximum and minimum stratus, and month and value of maximum and minimum static stability.

Stratus plays a significant role in the earth's radiation budget. Stratus, which tends to be optically thick, and have a high albedo, has a strong negative effect on the earth's radiation balance (Ramanathan et al. 1989; Klein and Hartmann 1993; Gruber et al. 1994).

Table 1. The ten stratus regions, their geographical extent, the season, and amount (%) of maximum stratiform cloudiness (DJF = December, January and February, etc.), the month and value (deg. C) of maximum static stability, the season and amount (%) of minimum stratiform cloudiness, and the month and value (deg. C) of minimum static stability. Note that in the Circumpolar Ocean there are few ship reports during austral winter; thus the annual cycle of stratus clouds is uncertain (Klein and Hartmann 1993).

Region	Location	Season and amount (%) of maximum stratus		Month and value (°C) of maximum stability		Season and amount (%) of minimum stratus		Month and value (°C) of minimum stability	
<i>Subtropical marine stratus</i>									
Peruvian	10°–20°S, 80°–90°W	SON	72	Oct	22	DJF	42	Feb	18
Namibian	10°–20°S, 0°–10°E	SON	75	Sep	22	MAM	48	Feb	17
Californian	20°–30°N, 120°–130°W	JJA	67	Jun	22	DJF	45	Feb	18
Australian	25°–35°S, 95°–105°E	DJF	45	Feb	18	JJA	41	Jun	15
Canarian	15°–25°N, 25°–35°W	JJA	35	Jun	16	SON	17	Oct	13
<i>Midlatitude marine stratus</i>									
North Pacific	40°–50°N, 170°–180°E	JJA	82	Jul	21	DJF	54	Jan	8
North Atlantic	50°–60°N, 35°–45°W	JJA	68	Jul	20	DJF	51	Dec	10
Circumpolar Ocean	50°–65°S	DJF?	62	Feb	15	?	?	Sep	13
Arctic	80°–90°N	JJA	62	Jan	28	DJF	18	Aug	20
China	20°–30°N, 105°–120°E	DJF	51	Feb	19	JJA	24	Aug	13

* In the computation of static stability (1), ECMWF climatological 1000-mb temperatures replace COADS surface air temperatures.

* In the Circumpolar Ocean, there are few ship reports during austral winter; thus the annual cycle of stratus clouds is uncertain.

2. Synoptic Influences on California Coastal Stratus

During the summer months the dominant synoptic feature is the semi-permanent Pacific anticyclone centered several hundred km west of the Central California coast (see Fig. 8). The location of the Pacific High leads to a persistent north to northwesterly flow along the California coast, with average surface wind speeds ranging from 7 to 10 ms⁻¹ (Simon 1977). See Figure 9 for illustration.

Fig. 8. Average sea level pressure (mb) in July (Mazur 1991).

Fig. 9. Mean surface winds for July (kts) (Simon 1977).

A typical vertical profile shows an elevated inversion over the ocean and extending inland. As seen in Figure 10, the inversion base height is at a minimum within about 10 km of the coast, and it slopes up gradually to the east and to the west (Bridger et. al 1993; Brick 1988). Average inversion base heights right along the coast have been measured at about 300 to 400 m above sea level (ASL) (Miller 1976; Goodman and Miller 1977; Mazur 1991). The inversion base height along the coast varies widely with extremes ranging from less than 10 meters to over 1000 meters. This variation is

largely a function of the location of the Pacific High and the strength of the subsidence from it.

The inversion is typically strong with temperature increases of about 10 °C per 100 to 300 m, and measurements by Goodman and Miller (1977) have shown temperature increases as high as 12 °C over 20 m. Typically, just below the inversion in the marine boundary layer, a stratus-type cloud layer persists, with stratocumulus west of about 125 degrees west longitude and stratus to the east. This is illustrated in Figure 10 where in zone 1 stratocumulus is observed and in zones 2 and 3 stratus is observed. The dividing line, located at the aforementioned longitude, is generally cloudless, which is due to strong surface divergence pushing the inversion base below the LCL (Simon 1977).

This brief description of a typical warm-month California coastal scenario is well documented. However, there remain many unanswered questions. Exactly how does fog form off the coast? What role does sea surface temperature play in fog formation? What exactly accounts for lifting the fog up such that it becomes a stratus layer? What mechanisms are responsible for maintaining the elevated inversion? What are the complex interactions that occur in the marine bound-

ary layer among convection, subsidence, radiation, turbulence, and advection?

Fig. 10. Variable inversion base height (Simon 1977).

3. How Stratus Forms - Current Theories

In this section, stratus and fog are described separately, where stratus includes all previously mentioned clouds with bases above the surface.

One frequently observed scenario is the so-called warm water case. In this case, fog and stratus develop in a wedge or

triangle-type pattern off the northern California coast. In this case the northern most edge is the point on the triangle where the stratus/fog is patchy and farther down wind the patches fill in to form an uninterrupted layer. A near surface based inversion is observed to the north and west of the stratus/fog, while an elevated inversion is observed within and at the edge of the stratus/fog.

In the warm water case, warm air, originating with subsidence from the Pacific High, advects over cool ocean water, at which time the heat flux is downward and the moisture flux is upward. This causes the lowest few meters of air to become moist and cool. The cool, moist air then advects over a patch of water that is warmer, at which time the heat flux is reversed. The upward heat flux causes the lowest few meters to warm and become superadiabatic. The now buoyant lowest layer of air rises, cools, and condenses into fog. Once the initial cloud layer is formed, strong radiative cooling from the top of the cloud drives further convection. The newly formed cloud droplets act as a vapor sink and enhance evaporation from the warm sea surface. The shallow convection, driven by cloud-top radiative cooling, increases downwind of the warm water and the fog develops into stratus. Further cloud-top radiative cooling strengthens and raises the inversion, and the result is an elevated

stratus layer capped by a strong inversion (Pilie et al. 1979).

Low level convergence has been observed as another mechanism for producing stratus (Pilie et al. 1979). In this process, air that has been modified by cool ocean water, flows into a convergence zone. In the convergence zone updrafts lift the air until condensation, forming fog and stratus. Cloud-top radiative cooling drives further instability resulting in more fog and stratus. In a 2400 km² convergence zone, located about 10 km off the northern California coast, Pilie et al. (1979) estimated an average updraft of 1 cm s⁻¹.

Another means by which stratus and fog form is adiabatic cooling. This occurs when an elevated inversion is in place with a well-established marine boundary layer below. When the inversion base rises above the lifted condensation level, a cloud can form. Once the cloud is established, the instability caused by radiative cooling can lead to a downward growth of the cloud. This so-called stratus lowering typically happens at night in the absence of solar radiation (Pilie et al. 1979; Noonkester and Logue 1976). This process, along with a vertical temperature profile, is depicted in Figure 11.

Fig. 11. Stratus lowering (Noonkester and Logue 1976).

Several authors (Lilly 1968; Pilie et al. 1979) have written about cloud-top radiative cooling as one essential element in establishing and maintaining fog and stratus. From this author's review of the literature, only two measurements of cloud-top cooling were actually recorded: Mack et. al (1974), and Goodman (1977). In both cases the cooling seemed to be caused by radiation loss from the top of the cloud.

In sharp contrast to the aforementioned authors, Telford and Chai (1984, 1993) conclude that radiative cooling is either non-existent or insignificant in the formation and maintenance of fog and stratus. This conclusion was reached after examining 100 or so marine soundings where two features were

consistently absent: 1) significant decrease in temperature near the cloud-top, and 2) significant low-level turbulence, both of which, they argue, would be present for strong radiative cooling destabilizing the layer.

Telford and Chai (1984, 1993) explain that the marine layer forms and increases in height as a result of the entrainment of air from above. The premise of this argument is that when the entrained air mixes with the nearly saturated air at the top of the surface layer, the result is a cooler less buoyant mixture. This is so because the stored up latent heat in the haze or mist particles, present at the top of the surface layer, provides sufficient cooling upon evaporation during this mixing. The entrained air continues to cool, by the evaporation of haze droplets, and sinks until the surface is reached. The marine layer grows vertically by this process until the wet-bulb potential temperature, which was initially increasing with height, becomes constant with height, thereby shutting off the evaporation.

Simultaneously occurring within the marine boundary layer is vapor driven convection. This process occurs because the layer of air closest to the surface is the recipient of the largest vapor flux. This nearly saturated or saturated air is less dense than the air just above it as it has a higher

virtual temperature. Any amount of turbulence sets off convection. Over time the convection produces fog and stratus.

4. Microstructure of California Coastal Stratus and Fog

Several reports documenting various aspects of the microstructure of California coastal fog and stratus have been published over the past half-century. These include, Neiburger and Wurtele (1949), Mack et. al (1974), Goodman (1977), Noonkester (1982), Telford and Chai (1993). In the forthcoming paragraphs there are frequent cross-referencing of the aforementioned authors. Henceforth, all references to the above publications include only names of the authors, and the dates stated above are assumed.

Mack et. al made extensive observations of the marine boundary layer using instruments aboard ship and aircraft. They documented several fog and stratus events over Central California coastal waters from Vandenberg Air Force Base (AFB) to San Francisco. Mack et. al also took measurements at Vandenberg AFB, about 1.6 km inland.

Goodman observed both fog and stratus, and in the majority of cases, collected data throughout the marine boundary layer, including above the cloud-top. Goodman's observa-

tions were made from a tower on the San Francisco Peninsula located roughly equidistant (about 5 km) from the Pacific Ocean (west), the Golden Gate Bridge (north), and the San Francisco Bay (east). Measurements were made from 250 m ASL (base of tower) to 470 m ASL (highest observation deck).

Noonkester exclusively documented stratus cases. The observations for his eight cases were taken throughout the marine boundary layer by a slow descent aircraft over the ocean southwest of San Diego. Observations made by Telford and Chai were primarily of fog. The data were collected by aircraft traveling from just south of San Francisco to over 300 kilometers due west. The aircraft stayed at one elevation (60 meters ASL) for the majority of the trip. The observations taken by Neiburger and Wurtele were of both fog and stratus, mostly in the Los Angeles Basin.

This paper will focus primarily on the work done by Mack et. al, Goodman, and Noonkester because of the comprehensive nature of these three studies.

a. Thickness of the cloud layer

The depth of the stratus layer averaged about 300 to 400 m over the ocean and ranged from a few meters to nearly 500 m. The average stratus cloud base was about 300 m ASL off the Central California coast and about 470 m ASL off the Southern California coast (Mack et. al; Noonkester).

Average thickness of sustained fog layers ranged from about 80 m up to a maximum of 400 m (Mack et. al; Goodman). Mack et. al observed that for fog to persist, an inversion not exceeding 400 m ASL was needed. Telford and Chai measured fog-tops at about 300 m ASL, at the base of the inversion.

b. Temperature profile

Temperature profiles within the marine boundary layer were similar for all cases and ranged from 8 to 18 °C. Noonkester observed an average lapse rate of about 9 °C km⁻¹ below the stratus and about 4.4 °C km⁻¹ within the stratus. In 5 of 8 cases he documented an unexplained superadiabatic lapse rate just below the cloud base (see Figure 12). It was not possible for Noonkester to observe cloud-top cooling, as he explains that his measurements were taken from 8:00 to 9:00 PST.

Fig. 12. Temperature and relative humidity profile in marine boundary layer (Noonkester, 1982).

Mack et. al observed a lapse rate through the entire marine layer of roughly $6\text{ }^{\circ}\text{C km}^{-1}$. They calculated nocturnal cloud-top cooling over 8 hours to be between 1.8 to $2.5\text{ }^{\circ}\text{C}$, although their measurements of cooling were between 0.0 and $0.7\text{ }^{\circ}\text{C}$.

Goodman measured nearly isothermal conditions in the cloud layer. Possible explanations for this are that the layers

observed by Goodman were considerably thinner than those observed by Mack et. al and Noonkester. Also, the up-sloping terrain just upwind of the observations may have produced enough updraft to result in a thermally well-mixed layer. Goodman measured nocturnal cooling at the top of the marine boundary layer from 1 to 3 °C, with the most cooling occurring just after the onset of fog.

c. Wind observations

Goodman (1977) measured the average horizontal wind speed in the marine layer between 3 and 10 m s^{-1} , and the average vertical wind speed was +20 cm s^{-1} . It is important to note that this updraft was probably influenced by orographic lifting and may not be an accurate measure of the general convection within a stratus layer.

Telford and Chai found no vertical wind velocity in the fog. As described earlier, Pilie et. al (1979) calculated an average updraft of 1 cm s^{-1} in a convergence zone with fog and stratus present.

d. Relative humidity

In Noonkester's study relative humidity averaged over 80 % at the ocean surface and increased with height. In the stratus layer, the relative humidity rarely reached 100 %. A relative humidity profile is contained in Figure 12. Neiburger and Wurtele observed droplets persisting with relative humidity values below 70 %.

e. Droplet size

In the stratus layer mean droplet size and width of the distribution increased with increasing height (Mack et. al; Goodman; Noonkester). However, in fog Mack et. al found the reverse; mean droplet size was negatively correlated with height. In stratus, Mack et. al observed an average droplet diameter between about 6 and 12 μm with a consistent 10 to 12 μm diameter observed at or near the top of the cloud layer (see Figure 13). Mack et. al found marine fog droplets to be larger with mean diameters from 11 to 19 μm .

Telford and Chai observed an average droplet diameter of about 11 μm in samples taken by aircraft off the central California coast at an elevation of about 60 m ASL. Goodman

and Noonkester measured slightly smaller average droplet diameters, averaging about 7 μm .

Drizzle size droplets were present that tended to precipitate downward in studies by Noonkester, Goodman, Mack et. al, and Telford and Chai. Noonkester (in stratus) and Mack et. al (in fog) observed some droplets with diameters larger than 100 μm .

In Goodman's observations, cloud droplet sizes and concentrations varied significantly from case to case. The mean diameter ranged from 4.5 to 10 μm and the droplet concentration ranged from 79 to 400 cm^{-3} . The explanation for this emerged after analysis of the air parcel trajectories over the 72-hour period prior to the sampling. The air masses were categorized according to their source regions: maritime, continental, and maritime-continental (see Figure 14). The average droplet concentration for the maritime cases was 89 cm^{-3} , for continental cases 266 cm^{-3} , and for maritime-continental cases 174 cm^{-3} . The higher concentrations tended to correlate positively with smaller mean diameters and generally narrower spectrums. Lower droplet concentrations tended to correlate with larger droplets and broader size distributions.

Vertical Profiles of Mean Drop Radius in Stratus
0-40 nmi Offshore Vandenberg, 15 July 1973

Fig. 13. Vertical profiles of mean droplet radius in stratus 0-40 nmi offshore from Vandenberg Air Force Base, July 15, 1973 (Mack et. al 1974).

time verses continental air masses. Noonkester's average maritime droplet concentration was about 270 cm^{-3} , with average droplet diameter of $8.0 \text{ }\mu\text{m}$. The average continental droplet concentration was 665 cm^{-3} , with average droplet diameter of $6.2 \text{ }\mu\text{m}$. Noonkester observed that droplet concentrations were fairly constant with height and then decreased sharply at the top of the cloud (see Figure 15). Noonkester's droplet concentrations are significantly higher than the others, indicating more cloud condensation nuclei. The reason for more abundant cloud condensation nuclei is not clear.

Telford and Chai found an average droplet concentration of somewhere between 60 to 185 cm^{-3} in fog. In this case, no attempt to document air parcel trajectories was made except to express the belief that the fog formed near the shore and drifted westward. Mack et. al observed droplet concentrations of 140 cm^{-3} in fog at an elevation of 52 m ASL about 1.6 km inland. At this same inland location the concentrations dropped to 11 cm^{-3} just above the surface at 11 m ASL.

Fig. 15. Liquid water content, number of particles, and mean radius of stratus droplets (Noonkester 1982).

f. Liquid water content

Liquid water contents were observed to increase with increasing heights in the stratus and fog layers (Noonkester; Goodman; Mack et. al). An example is depicted in Figure 15. Liquid water contents (g m^{-3}) were observed between 0.04 and 0.20, 0.01 to 0.08, and 0.009 to 0.16 for Mack et. al, Telford, and Goodman, respectively. Goodman estimates that her liquid water content values are an underestimation of actual values. This is attributed to instrument error.

Noonkester measured average liquid water contents (g m^{-3}) at 100 m above the cloud base. He documented values of 0.15 for maritime cases and 0.21 for continental cases. The range for Noonkester's study was 0.05 at the cloud base to over 0.60 near the cloud top (see Figure 15).

g. Condensation and cloud condensation nuclei

Condensation nuclei measured by Goodman peaked at 30,000 to 40,000 cm^{-3} during late afternoon and early morning rush hours. After scavenging processes in the fog worked through the night, minimum condensation nuclei concentrations as low as 1000 to 2000 cm^{-3} were reported. Figure 16 shows varia-

tion of condensation nuclei as a function of time for two 24-hour intervals.

Cloud condensation nuclei sampled from 25 to 100 km off shore by Mack et. al measured 50 cm^{-3} at 98 %percent relative humidity, 200 cm^{-3} at 0.5 % supersaturation and 240 cm^{-3} at 1.0 % supersaturation. Mack et. al strongly suggest that the cloud condensation nuclei are sea salt particles.

Fig. 16. Variation in condensation nuclei as a function of time (Goodman 1976).

C. Statistical Forecast Models

1. Introduction

Over the past 15 years San Francisco Bay Area forecasters have developed forecast aids to assist their predictions of when the morning marine stratus will burn-off sufficiently so that visual approaches are possible at San Francisco International Airport (SFO). Meteorologically, stratus burn-off is primarily a heat budget problem: heat added to the marine boundary layer evaporates the cloud and anything that impedes the heating will slow down or even reverse the evaporation process. The two major mechanisms for heat exchange in boundary layer are advection and terrestrial radiation. Terrestrial radiation is primarily a function the amount of incident solar radiation, i.e., the sun heats up the ground, which in turn heats up the marine boundary layer.

Given this understanding, forecasters have used predictors that in some way capture this heat budget as a function of time so that a forecast could be issued. The primary

predictors employed were cloud thickness, the height of the inversion, and either wind fields or pressure gradients. Cloud thickness was selected because a thicker cloud will take longer to burn off than a thinner one, all else being equal. The thicker cloud contains more liquid water and inhibits more solar radiation from reaching the ground than a thinner cloud, thus slowing the heating of the boundary layer and requiring more heating to complete the burn-off. The height of the inversion was chosen because it is the measure of the depth of the boundary layer, and is proportional to the volume of air that needs to be heated. A deeper marine boundary layer will require more time to heat up than a shallower one simply because there is more volume to heat. Therefore a cloud within a thicker boundary layer will take longer to evaporate than a similar cloud in a thinner boundary layer. The wind fields and pressure gradients were an attempt at capturing advection, e.g., a strong west wind tends to slow the burn-off as it is likely indicative of cool moist air filling into the approach zone.

Using the aforementioned predictors, the forecasters developed heuristic graphical forecast models. Though useful, these models have three significant shortcomings. First, the inputs to some of the models are not always objective, which means that the model could yield a different forecast depending on the forecaster. Second, the development approach is itself sometimes subjective. For example, trend lines on graphs are estimated and hand drawn. Third, the models can only issue one forecast at a given time of the day, i.e.; there is no opportunity for the model to give an updated forecast after 1400 UTC. Even so, the forecasters find their models to be generally helpful. The goal here is to mimic this general approach with an objective development and implementation process and to extend the approach to provide a sequence of models that can update the forecast until burn-off occurs.

2. Methodology

a. Overview

In the spirit of the stratus forecasters' models, statistical forecast models are developed using data collected from May to October of 1996 and 1997. These new models incorporate the handful of predictors used by the stratus forecasters and roughly 220 more potential predictors. A complete list of predictors, including the times of availability, is shown in Appendix A.

Although other model attempts are discussed, only 9 linear regression forecast models are presented: three models are developed for 1400 UTC, 1500 UTC, and 1600 UTC. The three models are named after the primary data inputs used.

The first model is the Zi-Ceiling model, where Zi refers to the height of the base of the inversion and Ceiling is the lowest at least broken layer reported from the hourly observation at SFO. The Zi is measured at two locations:

1) SFO and 2) San Carlos Airport which is about 17 km to the south east of SFO (roughly 6 km from the south west edge of the approach zone, Figure 2). The Z_i is measured using sodars. Sodars or acoustic radars indirectly measure the cloud-top by measuring the base of the inversion. The sodar sends sound waves propagating vertically from the ground, where some of the sound waves are back-scattered at the inversion base where the index of refraction (speed of sound) changes rapidly from the cool marine air below to warm dry air above. The back-scatter gets converted to a voltage, which is digitized and put in an array based on the time lag between the transmitted and received signal. From this the height of the Z_i is calculated using an algorithm developed at MIT/Lincoln Laboratory (MIT/LL).

The second model is the Cloud model where digitized visible satellite data are the only inputs into the models (this is discussed in more detail in chapter C.4). The third model is a combination of the first two types, called the Combo model.

For all the models, the data are from days where stratus was present in the morning such that a forecast of stratus burn-off needed to be issued by the CWSU forecasters - 103 days met this criterion. It is important to note that there were some days where stratus was present, but the data were missing. Days that reported rain were not included in the training set since the focus is on forecasting traditional stratus burn-off. Also excluded from the training data were days with a distinguishable at least broken layer above 900 meters, as these cases could have included more significant synoptic forcing that is not within the scope of this study.

b. First Attempts

The first model attempt is a mathematical imitation one of the CWSU forecaster models. It has the form

$$\Delta T = C_0 + C_1 Z_i + C_2 T_h$$

where ΔT is elapsed time after sunrise that burn-off is forecast to occur, C_0 , C_1 , and C_2 are the model coefficients, Z_i is the height of the base of the inversion, and T_h is the thickness of the cloud layer. The physical basis this model is that a cloud within a deeper boundary layer will require more time to evaporate than a similar cloud in a shallower boundary layer, and a thicker cloud will take longer to burn off than a thinner one. Although the physical reasoning for this model seems sound, when tested it produces unacceptably large time-errors.

After reviewing the physical properties associated with stratus, in particular that liquid water content increases quadratically with height (see Figure 15), several more attempts are made with this same basic model. The reasoning here is that possibly an exponent of the cloud thickness term, T_h^x , where x is allowed to be 2, 3, or 4, would better represent the burn-off process. However, this exponential model performs just as poorly as its linear predecessor.

c. Final Models

The most successful models attempted are of the form

$$\Delta T = C_0 + \sum C_i P_i$$

where ΔT is elapsed time after sunrise that burn-off is forecast to occur, and the model coefficients C_i are the weightings assigned to the various predictors P_i . The inclusion of the constant term C_0 allows the model to automatically adjust the origin for each set of predictors. Linear regression is used to determine the model coefficients, C_i .

Once the forecast models are created, one must attempt to quantify the ability of the models by measuring the model's error. The measures of error applied to the models are Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Standard Deviation (SD). MAE is simply the average number of minutes that the model is wrong. That is,

$$MAE = \sum/n (\text{Error}_i)$$

$$\text{Error}_i = \text{Forecast}_i - \text{Observed}_i$$

where the index i enumerates the training cases, and n is the number of cases. RMSE is

$$RMSE = SD^2 + BIAS^2$$

$$SD = (\sum(x_i - Xmean)^2 / n)^{.5}$$

$$Xmean = (\sum x_i) / n$$

$$Bias = (\sum \text{Error}_i) / n$$

The Principal of Least Squares asserts that the best model is the one with minimum residual.

$$\text{Residual} = \sum (\text{Error}_i)^2 / (n-p)$$

$$\text{Error}_i = \text{Forecast}_i - \text{Observed}_i$$

where p is the number of predictors. Therefore residual is used in developing the models, where minimizing the residual is the goal.

The optimal model has a relatively small number of predictors, while at the same time applying to a large number of the test cases. The reasoning behind a small number of predictors is that too many predictors used to develop the model on the training data may yield a model that "over-fits" the data. That is, the model produced may perform very well on the training data, but when applied to independent data the model loses substantial skill. To avoid "over-fitting" the data, the number of predictors is not allowed to be more than one-fifth of the number of development cases, and every effort is made to stay far below this limit.

The predictors are chosen by a step-wise selection process, but each model is constructed by the recomputation of all predictor weights, by using multiple regression. This

process is illustrated as follows: First, an initial model is developed using only a constant term,

$$\Delta T = C \quad (1)$$

Next, Model 1 error is computed for each case, where

$$\text{Error}_i = \text{Forecast}_i - \text{Observed}_i$$

The Model 1 error is then correlated with every potential predictor from Appendix A yielding a correlation coefficient, r , for each predictor. The correlation coefficient, r , is calculated as follows:

$$r = \Sigma(XY) / ((\Sigma X^2) (\Sigma Y^2))^{1/2}$$

$$X = x_i - X_{\text{mean}}$$

$$X_{\text{mean}} = (\Sigma x_i) / n$$

$$Y = y_i - Y_{\text{mean}}$$

$$Y_{\text{mean}} = (\Sigma y_i) / n$$

where x refers to the model error and y refers to the value of the predictor, the index i enumerates the cases, and n is the number of cases. The value of r is always between -1 and +1.

A high correlation (either negative or positive), between Model 1 error and a potential predictor, indicates a strong likelihood that the addition of this predictor will provide a substantial improvement to the model. That is, a predictor where the absolute value of the correlation coefficient is close to 1, will likely account for a significant part of the error of Model 1. Thus, the predictor with the highest positive or negative correlation coefficient, P_1 , is selected as the first input predictor, yielding Model 2,

$$\Delta T = C_0 + C_1 P_1 \quad (2)$$

If Model 2 has smaller residual than Model 1, then Model 2 replaces Model 1 as the "best" model. Then Model 2 error is correlated with all the potential predictors in Appendix A yielding new correlation coefficients for each predictor. P_1 now has a correlation coefficient of zero, and the candidate predictor, P_2 , with the highest positive or negative correlation coefficient is then selected, and a trial Model 3 is developed. This process is repeated for several iterations until there is no longer any meaningful contribution from additional predictors. Often in this process, several predictors yield similar correlation coefficients. In these cases each predictor is included and evaluated separately in an effort to produce the best model. In all, over 300 linear regression formulas were created and evaluated before the final 9 models were selected. The solver function in Microsoft Excel was used to produce the formulae.

After the "best" model is created, a two-step data randomization process is applied in an effort to avoid "over-fitting" the data. That is, the "best" model is reconstructed on half the data, i.e., half of the cases are

used to create a new model A, while the other half of the cases are used to produce a new model B. The coefficients from model A and B are then averaged together to produce the final model.

Developing the models requires a balance of several factors. Since the number of development cases can vary with the selection of predictors, the selection involves a balance between reducing the residual and maintaining as many cases as is practical. These subjective decisions depend on the experience and skill of the model developer.

3. Zi-Ceiling Models

The first set of models involves the height of the inversion, Z_i , and the height of the ceiling (lowest at least broken cloud layer), thus the name Zi-Ceiling Models. To have inputs for these models requires a measurable Z_i and Ceiling. This simple restriction eliminates several cases from the possible training set. Often the sodars are not able to detect a Z_i or the hourly observations only

report scattered, few, or no clouds at all. One shortfall of the Zi-Ceiling models is that sometimes the approach zone is filled with clouds, yet the nearby surface observations don't reflect that situation. Also, since arrivals at SFO are only impacted by clouds below 900 meters, all days with the lowest ceiling above 900 meters are excluded. These are not usually classic summer stratus days. Even with the aforementioned shortcoming and somewhat limited training set, the Zi-Ceiling models show promise. The 1400 UTC Zi-Ceiling model is

$$\Delta T = 2.8261 + 0.0021 \text{ Avg4Zi} - 0.0013 \text{ SFO4Cig} \quad (3)$$

where Avg4Zi is the average Zi at 1400 UTC measured by the SFO and SQL sodars (if either of the 2 is missing then the remaining one is taken as the average), and SFO4Cig is the height of the lowest ceiling at 1400 UTC at SFO. This model has 2 parameters, a RMSE of 1.07 hours, a MAE of 0.83 hours ($\cong 50$ min), and includes 67 cases. This model is conceptually similar to the heuristic forecaster model, which uses the 1400 UTC SFO ceiling and estimates Zi from

the 12Z Oakland sounding. A difference is that the forecasters choose to partition the data based on a ridge-trough rule, while the statistical approach finds no benefit in this partition.

The 1500 UTC Zi-Ceiling model is

$$\Delta T = 2.0155 + 0.0033 \text{ SFO5Zi} - 0.0022 \text{ SFO5Cig} - 0.0026 \text{ TR45SQL Zi} \quad (4)$$

where SFO5Zi is Zi at 1500 UTC measured by the SFO sodar, SFO5Cig is the height of the lowest ceiling at 1500 UTC at SFO. TR45SQLZi is the trend of the SQL Zi from 1400 UTC to 1500 UTC. That is,

$$\text{TR45SQLZi} = 14\text{ZSQLZi} - 15\text{ZSQLZi} \quad (5)$$

The 1500 UTC Zi-Ceiling model has 3 parameters, a RMSE of 0.91 hours, a MAE of 0.67 hours ($\cong 40$ min), and includes 46 cases.

The 1600 UTC Zi-Ceiling model is

$$\Delta T = 2.0644 - 0.0023 \text{ SFO6Cig} + 0.0009 \text{ TR46SQLZi} + 0.0034 \text{ Avg6Zi} \quad (6)$$

where SFO6Cig is the height of the lowest ceiling at 1600 UTC at SFO, TR46SQLZi is the trend of the San Carlos Zi from 1400 UTC to 1600 UTC and is calculated by taking the average slope of the best-fit line from 1400 UTC to 1600 UTC. Avg6Zi is the average Zi at 1600 UTC measured by the SFO and SQL sodars (if either of the 2 is missing then the remaining one is taken as the average). The 1600 UTC Zi-Ceiling model has 3 parameters, a RMSE of 0.90 hours, a MAE of 0.69 hours ($\cong 41$ min), and includes 46 cases.

4. Cloud Models

The second set of models exclusively uses satellite derived visible cloud data. Many would argue that the recent high-resolution visible satellite imagery is the most important type of data the forecaster has in issuing a short-term stratus forecast. If the forecaster can see where the stratus is and is not, the forecaster can extrapolate in the near-term where the stratus will or will not be. The shortcoming of this method is that little is quantified. That is, the forecaster can see where the stratus is, but has no further objective measure of the cloud.

MIT/LL has provided a Cloud Map product, which is obtained by applying a clustering algorithm to the GOES-9 visible data. This product was available only during the operational life of GOES-9, June 1996 until July 1998, after which GOES-10 became available. This product is based on a tiling of the greater San Francisco Bay Area by geographic sectors. Each sector is assigned three cloud attributes, for each 15-minute satellite scan:

1. mean brightness is a measure of the optical thickness of the cloud in that sector
2. standard deviation of brightness is a measure of cloud irregularity
3. coverage is a measure of the percentage of the sector that has visible cloud that is sufficient to impair visual approaches

These three data types - mean, standard deviation, and coverage - for each of 14 sectors, are used as the potential predictors for the cloud models. The sectors used are B1, B2, B3, B4, B5, B6, W1, W2, W3, W10, W11, W12, S5, and S6. See Figure 17 for a map of the cloud sectors. See Appendix B (Wilson, F.W. and D.A. Clark, 1999) for a complete description of the cloud data.

Since visible data are only available after sunrise, the 1400 UTC model has fewer cases from which to draw - from the beginning of May to the middle of August. There are also several days where the data are missing for a given time. The result is that the models are developed from 41 1400 UTC cases, 66 1500 UTC cases, and 55 1600 UTC cases.

The 1400 UTC Cloud model is

$$\Delta T = (-3.6923) + 5.0916 W1Mn + 7.0128 W11Mn + 1.0921 B4Cov \quad (7)$$

where W1Mn is the 1400 UTC mean value from sector W1, W11Mn is the 1400 UTC mean value from sector W11, and B4Cov is the 1400 UTC coverage value from sector B4. The 1400 UTC Cloud model has 3 parameters, a RMSE of 0.79 hours, a MAE of 0.57 hours ($\cong 34$ min), and includes 41 cases.

The 1500 UTC Cloud model is

$$\Delta T = (-1.3791) + 4.5337 W11Mn + 5.5988 W12Mn \quad (8)$$

where W11Mn is the 1500 UTC mean value from sector W11, W12Mn is the 1500 UTC mean value from sector W12. The 1500 UTC Cloud model has 2 parameters, a RMSE of 0.71 hours, a MAE of 0.54 hours ($\cong 32$ min), and includes 65 cases.

The 1600 UTC Cloud model is

$$\Delta T = 0.1843 + 7.6298 W1Mn + 4.7245 TR56B1Mn \quad (9)$$

where $W1Mn$ is the 1600 UTC mean value from sector W1, $TR56B1Mn$ is the trend of the mean value from 1500 UTC to 1600 UTC from sector B1. The trend is the 1500 UTC mean value - the 1600 UTC mean value. The 1600 UTC Cloud model has 2 parameters, a RMSE of 0.64 hours, a MAE of 0.48 hours ($\cong 29$ min), and includes 52 cases.

Fig. 17. The San Francisco Bay Area broken up into sectors. The approach zone runs through B1 and B2, with SFO as the circle in upper left of sector B1.

5. Combo Models

Since both the Zi-Ceiling models and Cloud models produce useful results, it seems plausible that a combination of the two could yield sharper results. The Combo models utilize all available data from Appendix A.

The 1400 UTC Combo model is

$$\Delta T = (-2.6635) + 0.0007 \text{ Avg4Zi} + 5.2795 \text{ W1Mn} + 3.2694 \text{ W11Mn} + 0.9892 \text{ B4Cov} \quad (10)$$

where Avg4Zi is the average Zi at 1400 UTC measured by the SFO and SQL sodars (if either of the 2 is missing then the remaining one is taken as the average), W1Mn is the 1400 UTC mean value from sector W1, W11Mn is the 1400 UTC mean value from sector W11, and B4Cov is the 1400 UTC coverage value from sector B4. The 1400 UTC Combo model has 4 parameters, a RMSE of 0.62 hours, a MAE of 0.47 hours ($\cong 28$ min), and includes 38 cases. This compares with the 1400

UTC Zi-Ceiling model, which has 2 parameters, a RMSE of 1.07 hours, a MAE of 0.83 hours ($\cong 50$ min), and includes 67 cases, and the 1400 UTC Cloud model, which has 3 parameters, a RMSE of 0.79 hours, a MAE of 0.57 hours ($\cong 34$ min), and includes 41 cases.

The 1500 UTC Combo model is

$$\Delta T = (-1.0375) + 4.5136 W11Mn + 4.9115 W12Mn \quad (11)$$

where W11Mn is the 1500 UTC mean value from sector W11, W12Mn is the 1500 UTC mean value from sector W12. The additional restriction on this model is that the SFO sodar detects a Zi at 1400 UTC, which restricts the cases for application. The 1500 UTC Combo model has 2 parameters and 1 data restriction, a RMSE of 0.60 hours, a MAE of 0.44 hours ($\cong 26$ min), and includes 56 cases. This compares with the 1500 UTC Zi-Ceiling model, which has 3 parameters, a RMSE of 0.91 hours, a MAE of 0.67 hours ($\cong 40$ min), and includes 46 cases, and the 1500 UTC Cloud model, which has

2 parameters, a RMSE of 0.71 hours, a MAE of 0.54 hours (≈ 32 min), and includes 65 cases.

The 1600 UTC Combo model is

$$\Delta T = 0.0032 + 7.5327 W1Mn + 3.6783 TR56B1Mn + 1.8785 TR56W12Mn + 0.5654 RT \quad (12)$$

where $W1Mn$ is the 1600 UTC mean value from sector $W1$, $TR56B1Mn$ is the trend of the mean value from 1500 UTC to 1600 UTC from sector $B1$. The trend is 1500 UTC mean value - the 1600 UTC mean value. $TR56W12Mn$ is the trend of the mean value from 1500 UTC to 1600 UTC from sector $W12$. Again the trend is 1500 UTC mean value - the 1600 UTC mean value. RT is the categorization of each day as either a ridge or trough case. This was done by the CWSU forecasters where the categorization ranged from strong ridge, ridge, weak ridge, neutral, weak trough, trough, to strong trough. The categories are given a value from 0 to 1.5, where 1.5 is assigned to the strong trough case, 1 to

trough case, 0.5 to the weak trough case, and 0 assigned to neutral or ridge cases. The additional restriction on this model is that one of the sodars detects a Zi at 1600 UTC. The 1600 UTC Combo model has 4 parameters and 1 restriction, a RMSE of 0.56 hours, a MAE of 0.42 hours ($\cong 25$ min), and includes 46 cases. This compares with the 1600 UTC Zi-Ceiling model, which has 3 parameters, a RMSE of 0.90 hours, a MAE of 0.69 hours ($\cong 41$ min), and includes 46 cases, and the 1600 UTC Cloud model, which has 2 parameters, a RMSE of 0.64 hours, a MAE of 0.48 hours ($\cong 29$ min), and includes 52 cases.

A summary of all the model characteristics is presented in Table 2.

Table 2. Forecast model characteristics. RMSE is Root Means Square Error. MAE is Mean Absolute Error. SD is Standard Deviation. Units are decimal hours.

Model	Bias	RMSE	MAE	SD	Cases
Zi-Ceiling 1400 UTC	.02	1.07	.83	1.03	67
Cloud 1400 UTC	-.04	.79	.57	.89	41
Combo 1400 UTC	-.03	.62	.47	.78	38
Zi-Ceiling 1500 UTC	.02	.91	.67	.95	46
Cloud 1500 UTC	-.01	.71	.54	.85	65
Combo 1500 UTC	.05	.60	.44	.77	56
Zi-Ceiling 1600 UTC	.02	.90	.69	.95	46
Cloud 1600 UTC	-.04	.64	.48	.80	52
Combo 1600 UTC	-.02	.56	.42	.75	46

6. Model Verification

The models are run on 1998 data and the results are presented in Table 3.

The number of cases in 1998 available for analysis is small and this is especially apparent with the models that use the cloud data as input. This occurs for several reasons. First there are not that many days where stratus is present as compared with 1996 and 1997. Second the cloud data are limited by the number of days that GOES-9 was operational.

GOES-9 was replaced by GOES-10 in late July 1998 and the GOES-10 visible data are significantly different from the GOES-9 data. Since the models were developed on data from GOES-9, it seemed logical to only test on GOES-9 data. Third, there were intermittent data dropouts.

Table 3. Model results on 1998 data. RMSE is Root Means Square Error. MAE is Mean Absolute Error. SD is Standard Deviation. Units are decimal hours.

Model	Bias	RMSE	MAE	SD	Cases
Zi-Ceiling 1400 UTC	.12	.93	.79	.96	32
Cloud 1400 UTC	-.68	1.39	.91	.96	15
Combo 1400 UTC	-.83	1.45	.98	.87	14
Zi-Ceiling 1500 UTC	.17	1.14	.92	1.05	32
Cloud 1500 UTC	-.53	1.09	.85	.90	16
Combo 1500 UTC	-.47	.90	.77	.82	15
Zi-Ceiling 1600 UTC	.18	.93	.70	.95	30
Cloud 1600 UTC	-.71	1.03	.83	.72	14
Combo 1600 UTC	-.79	1.10	.88	.69	14

Overall the 1998 Zi-Ceiling models perform reasonably well with RMSEs and MAEs comparable to those observed on the training data. However, the models that include the cloud data perform poorly overall on the 1998 data. On average

the models forecast for good visibility has a 40 minute early bias. The bias gets worse for later burn-offs. The RMSE and MAE are on the order of double those for the training data. The Combo models perform the worst on the 1998 data, whereas the Combo models perform the best on the training data.

7. Zi- Ceiling Model Verification Analysis

As stated earlier, the forecasts issued on the 1998 data by the Zi-Ceiling models perform well and therefore there is confidence in their robustness. All the model results are presented in graphical format in Appendix C. Note that the 1998 Zi-Ceiling model graphs fairly closely resemble the training data graphs.

8. Cloud and Combo Model Verification Analysis

The forecasts issued on the 1998 data by the Cloud and Combo models perform poorly. There are several possible explanations for the weak results. The most obvious

candidate is that the models are over fitting the data. This however is very unlikely as the model with the lowest parameter to case ratio has 4 parameters applied to 46 cases and 2 of the models have only 2 parameters applied to over 50 cases. Furthermore, the aforementioned 2-step data randomization process was carefully employed on all the models.

Another possible explanation is the small number of 1998 cases. These cases are bunched in June and July and could be biased by the time of year alone. Again this is not very likely. The 1400 UTC models were developed on cases that ranged from May to July, 83% of which were June and July cases. The 1400 UTC models that use cloud data perform just as poorly as the 1500 UTC and 1600 UTC models that use cloud data.

Another possible explanation is that there could also be a bias for any given year that relates more to larger scale issues that are beyond the scope of this investigation. For example, the summer of 1998 followed a very active El

Nino rainy season (in fact it was among the top 4 rainiest on record in the San Francisco Bay Area). Some Bay Area meteorologists have argued that summers following El Nino winters are warmer than normal in the Bay Area which could have an impact on stratus. However, the robustness of the Zi-Ceiling models seems to cast doubt on this.

Another possibility is that the 1998 visible cloud data are contaminated. Although the visible data yield the most promising results on the training data, the cloud data remain more or less a "black box" to the user. That is, all of the other data inputs are very familiar to any meteorologist, e.g., Zi, cloud base, etc. One can readily determine if the data from any of the familiar inputs are suspect. For example, a Zi report of 2000 meters would clearly be out of bounds in forecasting approach-zone stratus. However, it is nearly impossible to tell if the digitized satellite data are erroneous. Recall that the digitized visible data used measure 3 things: 1) cloud brightness or optical thickness, 2) standard deviation of brightness, and 3) coverage or the percentage of the sector that has visible cloud that is sufficient to impair visual

approaches. In viewing a visible image with the naked eye it is easy to tell where clouds are or aren't, but nearly impossible to determine a cloud's optical thickness, standard deviation of that optical thickness, or coverage with any degree of accuracy. When dealing with cloud data, there are two main possibilities for error: 1) the so-called "drift" in the visible channels that occurs over time with the GOES satellites, 2) the satellite itself malfunctioning.

In an effort to find more specifics on this "drift" or darkening of the visible channels, experts were contacted. It was learned that the lens used to take visible images on all GOES satellites darkens exponentially over time. This is due to the exposure to solar radiation. Weinreb, et. al, (1997) suggest that the lens darkens roughly 5% per year. See Figure 18 for visible evidence of the darkening. Currently there is no calibration process in place to compensate for this darkening. So as time goes on the satellite reports back darker and darker clouds, which the model interprets as thinner and thinner clouds. The result of this is that as time goes on, the model would continue

Fig. 18. Visible image taken by GOES-8 (top) and GOES-9 (bottom) on the afternoon of September 13, 1995. The differences in perspective for 75W and 90W are obvious, and you can also see that GOES-9's visible image is brighter than GOES-8's. GOES-8 had been operational for roughly 16 months longer than GOES-9 when this image was taken. Image provided by NASA Goddard Space Flight Center.

to forecast earlier and earlier burn-offs. This is consistent with the early bias found for the 1998 cloud data.

In an effort to quantify this drift, two comparisons are made between similar cases on the training data and the 1998 data. Both comparisons are made using 1500 UTC data from June and July cases. In the first comparison, cases are selected with burn off times from 1924Z to 2000Z. There are 3 cases in the training data where the average mean brightness is 0.724 for sector W11 (See Figure 17 for sector location), and the average mean brightness is 0.729 for sector W12. There are 4 cases from 1998 where the average sector W11 mean brightness is 0.678, and the average sector W12 mean brightness is 0.649. In the second comparison, cases are selected with burn off times from 1728Z to 1800Z. There are 4 cases in the training data where the average sector W11 mean brightness is 0.646, and the average sector W12 mean brightness is 0.647. There are 8 cases in the 1998 data where the average sector W11 mean brightness is 0.583, and the average sector W12 mean brightness is 0.616. This averages out to roughly an 8%

decrease in brightness from the training data to the 1998 data. This is summarized in tables 4 and 5. It should be pointed out here that the assumption in this analysis is that the optical thickness of the cloud in sectors W11 and W12 is a good indication of when the stratus will burn off. This seems reasonable since the statistics bear this out. The second caveat is that this is a small data set, which unfortunately is all the data that are available for this analysis. Even so, the results support this darkening of the lens.

Table 4. Average Mean Brightness for 1500 UTC with side-bys beginning between 1924 UTC and 2000 UTC.

Sector	Training Data Average Mean Brightness	1998 Data Average Mean Brightness
W11	0.724	0.678
W12	0.729	0.649

Table 5. Average Mean Brightness for 1500 UTC with side-bys beginning between 1728 UTC and 1800 UTC.

Sector	Training Data Average Mean Brightness	1998 Data Average Mean Brightness
W11	0.646	0.583
W12	0.647	0.616

The second issue associated with the quality of the 1998 visible data concerns the GOES-9 satellite itself. Investigations into GOES-9 reveal that the satellite started experiencing trouble with its momentum wheels in early 1998. The momentum wheels stabilize the satellite so that it is able to take clean images. By June of 1998 the problem was very apparent in the imagery. Figure 19 illustrates the significant adverse impact on the visible imagery in 1998. Figure 20 shows another problem case focused on the San Francisco Bay Area. Investigations into more visible images taken at this time show that this problem was intermittent in intensity, but appears to have impacted nearly all of the images. By the end of July 1998, NASA removed GOES-9 from service for this reason.

Clearly, the visible data collected from GOES-9 during June and July of 1998 are contaminated. The algorithm that processes the visible data employs a median filter, which is an attempt to filter out irregularities. The median filter is not capable of filtering out all of the "dark spots" on the images. Figure 21 shows a visible image of

the San Francisco Bay area before the filter is applied, while Figure 22 shows a post-processed image from the same time. Recall that minor variations in the mean cloud brightness values can yield moderate differences in forecasts. It is therefore imperative to account for every small difference in cloud brightness. Although one can't distinguish any obvious contamination in Figure 22, the image is slightly darker than it would have been without the contamination. As stated previously, the result of a darker image is an early bias in the forecasts.

The two problems with the 1998 visible data, drift and contamination, are likely causes for the early bias and overall poor results in the 1998 forecasts of the Cloud and Combo models.

Fig. 19. Visible image taken by GOES-9 on July 15, 1998 over Baja California. Notice the "lines" streaking from left to right across the image. Image provided by NASA Goddard Space Flight Center.

Fig. 20. Visible image taken by GOES-9 on June 13, 1998 over the San Francisco Bay Area. Notice the "lines" streaking from left to right across the image. Image provided by the Naval Research Lab, Monterey, CA.

Fig. 21. Visible image taken by GOES-9 on June 13, 1998 at 1600 UTC, over the San Francisco Bay Area. The "lines" streaking from left to right across the image, though still distinguishable are not as apparent as in Fig. 21. Image provided by the Naval Research Lab, Monterey, CA.

Fig. 22. Image from June 13, 1998 at 1600 UTC, over the San Francisco Bay Area. This is the same image as in Fig. 22, except that the data have been filtered. The "lines" streaking from left to right across the image are not apparent on this image. Image provided by the MIT/LL.

9. Summary of Results

A total of nine forecast models have been presented in this report. They are partitioned by the forecast time: 1400 UTC, 1500 UTC, 1600 UTC, and by the data type: Inversion, Clouds, or Combination. The predictor weightings for each model are determined by linear regression, selected to minimize the residual. MAE, RMSE, and SD have been provided for each case as measures of model skill. Each model was tested on the independent 1998 data and the results presented.

10. Conclusions

Statistical analysis of regional data has led to the development of several forecast models for marine stratus burn-off at SFO. Based on the measures of MAE, RMSE, and SD the models all show an improving skill as the forecast hour advances on the training data. The 14 UTC surface model is a direct extension of the graphical cloud analysis technique, which is employed by the operational

forecasters. The 1500 UTC and 1600 UTC training data analogues show skill improvement. The training data Zi-Ceiling models lag the cloud models. The training data combination models provide a substantial additional level of improvement, with indicated model errors below one-half hour for the 1600 UTC model.

When the models are tested on the 1998 independent data, the Zi-Ceiling models perform reasonably well, but the Cloud and Combo models perform poorly. Two probable causes are examined as the most likely reason for the poorly performing models: 1) drift in the visible imagery on GOES satellites, and 2) contamination of GOES-9 data. There is no calibration made for the drift in the GOES imagery for any of the data. Even so, the training data appear to produce impressive results. Once the calibration is made for the drift in the GOES imagery, it seems likely that the models that employ cloud data will yield results that are at least as good as the training data models presented here.

Appendix A Listing of the candidate predictors

Surface Predictors:

Time	Label	Description
12Z	CWSU Zi	CW SU estimate of Zi (Oakland Sounding)
12Z	OAK Zi	Automated estimate of Zi (Oakland Sounding)
12Z	R/T score	CWSU Ridge/Trough score
14Z	SQL Zi	Inversion base - San Carlos Sodar
14Z	SFO Zi	Inversion base - SFO Sodar
14Z	SFO Cig	SFO Ceiling
14Z	SQL Cig	San Carlos Ceiling
14Z	SF Dir	SFO Wind direction
14Z	SF Spd	SFO Wind speed
14Z	SF West	SFO West Wind
14Z	SQ Dir	SQL Wind direction
14Z	SQ Spd	SQL Wind speed
14Z	SQ West	SQL West Wind
14Z	SODAR AVG	SQL Zi & SFO Zi average
15Z	SQL Zi	Inversion base - San Carlos Sodar
15Z	SFO Zi	Inversion base - SFO Sodar
15Z	SFO Cig	SFO Ceiling
15Z	SQL Cig	San Carlos Ceiling
15Z	SF Dir	SFO Wind direction
15Z	SF Spd	SFO Wind speed
15Z	SF West	SFO West Wind
15Z	SQ Dir	SQL Wind direction
15Z	SQ Spd-kt	SQL Wind speed
15Z	SQ West	SQL West Wind
15Z	SODAR AVG	SQL Zi & SFO Zi average
16Z	SQL Zi	Inversion base - San Carlos Sodar
16Z	SFO Zi	Inversion base - SFO Sodar
16Z	SFO Cig	SFO Ceiling
16Z	SQL Cig	San Carlos Ceiling
16Z	SF Dir	SFO Wind direction
16Z	SF Spd	SFO Wind speed
16Z	SF West	SFO West Wind
16Z	SQ Dir	SQL Wind direction
16Z	SQ Spd-kt	SQL Wind speed
16Z	SQ West	SQL West Wind
16Z	SODAR AVG	SQL Zi & SFO Zi average
15Z	TR45 SQL Zi	Trend: Inversion base - San Carlos Sodar
15Z	TR45 SFO Zi	Trend: Inversion base - SFO Sodar
15Z	TR45 SFO Cig	Trend: SFO Ceiling
15Z	TR45 SQL Cig	Trend: San Carlos Ceiling
15Z	TR45 SF Dir	Trend: SFO Wind direction
15Z	TR45 SF Spd	Trend: SFO Wind speed
15Z	TR45 SF West	Trend: SFO West Wind
15Z	TR45 SQ Dir	Trend: SQL Wind direction
15Z	TR45 SQ Spd-kt	Trend: SQL Wind speed
15Z	TR45 SQ W	Trend: SQL West Wind
15Z	TR45 SODAR	Trend: SQL Zi & SFO Zi average
16Z	TR56 SQL Zi	Trend: Inversion base - San Carlos Sodar
16Z	TR56 SFO Zi	Trend: Inversion base - SFO Sodar
16Z	TR56 SFO Cig	Trend: SFO Ceiling
16Z	TR56 SQL Cig	Trend: San Carlos Ceiling
16Z	TR56 SF Dir	Trend: SFO Wind direction
16Z	TR56 SF Spd	Trend: SFO Wind speed
16Z	TR56 SF W	Trend: SFO West Wind
16Z	TR56 SQ Dir	Trend: SQL Wind direction
16Z	TR56 SQ Spd-kt	Trend: SQL Wind speed
16Z	TR56 SQ W	Trend: SQL West Wind
16Z	TR56 SODAR	Trend: SQL Zi & SFO Zi average

Cloud Predictors:

Time	Sector	Param.	Time	Sector	Param.	Time	Sector	Param.
14z	B1	M	15z	B1	M	16z	B1	M
14z	B1	SD	15z	B1	SD	16z	B1	SD
14z	B1	C	15z	B1	C	16z	B1	C
14z	B2	M	15z	B2	M	16z	B2	M
14z	B2	SD	15z	B2	SD	16z	B2	SD
14z	B2	C	15z	B2	C	16z	B2	C
14z	B3	M	15z	B3	M	16z	B3	M
14z	B3	SD	15z	B3	SD	16z	B3	SD
14z	B3	C	15z	B3	C	16z	B3	C
14z	B4	M	15z	B4	M	16z	B4	M
14z	B4	SD	15z	B4	SD	16z	B4	SD
14z	B4	C	15z	B4	C	16z	B4	C
14z	B5	M	15z	B5	M	16z	B5	M
14z	B5	SD	15z	B5	SD	16z	B5	SD
14z	B5	C	15z	B5	C	16z	B5	C
14z	B6	M	15z	B6	M	16z	B6	M
14z	B6	SD	15z	B6	SD	16z	B6	SD
14z	B6	C	15z	B6	C	16z	B6	C
14z	W1	M	15z	W1	M	16z	W1	M
14z	W1	SD	15z	W1	SD	16z	W1	SD
14z	W1	C	15z	W1	C	16z	W1	C
14z	W2	M	15z	W2	M	16z	W2	M
14z	W2	SD	15z	W2	SD	16z	W2	SD
14z	W2	C	15z	W2	C	16z	W2	C
14z	W3	M	15z	W3	M	16z	W3	M
14z	W3	SD	15z	W3	SD	16z	W3	SD
14z	W3	C	15z	W3	C	16z	W3	C
14z	W10	M	15z	W10	M	16z	W10	M
14z	W10	SD	15z	W10	SD	16z	W10	SD
14z	W10	C	15z	W10	C	16z	W10	C
14z	W11	M	15z	W11	M	16z	W11	M
14z	W11	SD	15z	W11	SD	16z	W11	SD
14z	W11	C	15z	W11	C	16z	W11	C
14z	W12	M	15z	W12	M	16z	W12	M
14z	W12	SD	15z	W12	SD	16z	W12	SD
14z	W12	C	15z	W12	C	16z	W12	C
14z	S5	M	15z	S5	M	16z	S5	M
14z	14z S5	SD	15z	14z S5	SD	16z	14z S5	SD
14z	S5	C	15z	S5	C	16z	S5	C
14z	S6	M	15z	S6	M	16z	S6	M
14z	S6	SD	15z	S6	SD	16z	S6	SD
14z	S6	C	15z	S6	C	16z	S6	C

Cloud Mean Trend Predictors for the indicated sectors

TR14-15	B1	TR15-16	B1	TR14-16	B1
TR14-15	B2	TR15-16	B2	TR14-16	B2
TR14-15	B3	TR15-16	B3	TR14-16	B3
TR14-15	B4	TR15-16	B4	TR14-16	B4
TR14-15	B5	TR15-16	B5	TR14-16	B5
TR14-15	B6	TR15-16	B6	TR14-16	B6
TR14-15	W1	TR15-16	W1	TR14-16	W1
TR14-15	W2	TR15-16	W2	TR14-16	W2
TR14-15	W3	TR15-16	W3	TR14-16	W3
TR14-15	W10	TR15-16	W10	TR14-16	W10
TR14-15	W11	TR15-16	W11	TR14-16	W11
TR14-15	W12	TR15-16	W12	TR14-16	W12
TR14-15	S5	TR15-16	S5	TR14-16	S5
TR14-15	S6	TR15-16	S6	TR14-16	S6

Appendix B: Cloud Products

1.0 Satellite Derived Products

There are currently two products under development that are derived from GOES West (GOES-9 and GOES-10) satellite data. They are the Satellite Gridded Product, and the Satellite Sector Product. Both of these are derived data products, as defined in the Product Requirements Document, as they are derived from a single sensor, and are intended for use as input for generating other products rather than as stand alone products. In addition to these two products, there is under consideration the development of a Satellite Cloud Forecast Product which would serve as a stand alone Individual Forecast Product.

Archived satellite data are being sent via tape to Lincoln Laboratory from the Naval Postgraduate School in Monterey. Data are available from approximately 400 summer days from 1996-1998. The Visible channel data are available for all three years, while the IR and Water Vapor channels are available for only a subset of the days. The initial focus

for product development is on the Visible data, which most directly relate to the stratus burnoff problem. These data are available at 15 minute intervals throughout the day. The Visible satellite data consist of brightness values at a nominal 1 km resolution. (For practical purposes, the actual data resolution over the SF Bay area is closer to 1.5 km due to the oblique viewing angle of the satellite.) The reported brightness values are a measure of albedo or reflectivity; in the case of stratus cloud, this is proportional to liquid water content of the cloud, which is in turn proportional to cloud thickness. The returned brightness values also are impacted by the daily and seasonal variability of the sun angle.

1.1 Satellite Gridded Product

The intent of the Satellite Gridded Product is to have a stable, numerical representation of satellite data on a fixed cartesian grid for use in generation of satellite images and as possible input to other forecast algorithms. This product is initially being developed using the Visible

satellite data. It has an 80x80 km domain centered over the SFO approach zone, consisting of scaled satellite brightness values on a 2 km grid. The primary value of this product, compared to simply using the source data, is the data pre-processing which provides a numerically reliable output that is more appropriate for trend analysis. This is an important consideration for the cloud burnoff forecast mission. Specifically it provides:

- a) Improved data registration
- b) Data mapped to a fixed cartesian grid
- c) Modest filtering for numerical stability (spatially and temporally)
- d) Normalization for sun angle

One of the initial pre-processing steps is a simple test for data registration - locating the data geographically. cursory checks of registration indicated that errors of up to 4 km are possible, with errors of 1-2 km more typical. Although the overall registration reliability is superior to that of the previous generations of GOES-West,

improvements to data registration have value, since small horizontal variations in cloud cover are important operationally with regard to the stratus forecast mission. As such, a simple check examines brightness variance in regions of known geophysical contrast, to shift the raw data field when necessary.

The source data field is then projected onto a grid. These data are mapped to a 2-km cartesian grid by interpolating from the field of raw brightness values. The standardized Universal Transverse Mercator (UTM) grid, which is shown on all quadrangle maps provided by the U.S. Geological Survey, is used.

The 2-km gridded data field is then passed through a simple 9-point median filter to provide increased numerical stability to the data. The objective is to smooth out some of the small time- and space-scale fluctuations in the cloud field that would impact trend analysis. Experiments with a few different filters were made before choosing the 9-point filter which removed most of the very fine-scale

short-lived features without over-smoothing the sharpness of the cloud edges.

The final major pre-processing step is a normalization of data to account for the variation of brightness due to sun angle. This is perhaps the most important value added component to the gridded product, since it provides a useful quantitative cloud assessment input for other algorithms, and allows for numerical trend analysis. The normalization method is adapted from an existing technique by Cherna et al. (1985).

The end-product is a filtered and normalized 2-km resolution 80x80 km grid of brightness values scaled to a range of 0 to 1, centered over the SFO final approach zone. The gridded product is used to create and archive GIF satellite imagery, and to generate the Satellite Sector Product.

1.2 Satellite Sector Product

A number of applications for satellite data information have been identified during cloud burnoff forecast product development. For many of these applications, the desirable input is a more simplified mapping of cloud information which can be used in algorithm logic.

In order to meet this objective, development is continuing on a Satellite Sector Product. This product provides satellite information in a format that is similar to that of individual observations (e.g., cloud reports from surface observations), which are more compatible with the framework of the existing algorithm development.

The Satellite Sector Product is generated from the Satellite Gridded Product. It is created by dividing the 80x80 km gridded domain (6400 sq. km areal coverage) into approximately 50 sectors, whereby each sector has some homogeneity with respect to the physical geography that influences the cloud formation and dissipation processes. (See Figure 23.) The division of sectors is designed to isolate specific areas that are the typical focus of

interest for the stratus forecast, e.g., the San Bruno Pass to the northwest and the Crystal Gap to the southwest of the airport, the coastal range, the inland valley areas, etc. Particular attention is paid to making sector boundaries that separate land surfaces from water surfaces. In particular there are groups of sectors that define the Pacific Ocean (both along the coast and away from the coast), the SFO Peninsula (coastal hills and interior valley), the San Francisco Bay, the Golden Gate area north of the Bay, and the interior ranges to the east of the Bay.) In addition, sectors are also defined to identify the approach corridor out to 15 miles (25 km), as well as the corridors to the west of the approach zone (land surface), east of the approach zone (water surface), and the merge zones at the south end of the Bay.

Each sector includes a number of data points from the gridded product. Thus, a statistical representation of each sector can be created from the gridded data contained within, thereby allowing each to act as an individual "observation." Depending on its size, a single sector may include as few as 8 to as many as 90 grid points. Most of

the smaller sectors are located in the vicinity of the approach zone where details of the extent of cloud coverage and thickness are most important.

For each sector, a time-series of statistical observations are generated from the 15-minute satellite data. The observations include the mean and median brightness value within the sector, the variance of values within the sector, and the percent of data points within the sector that exceed one or more significant brightness thresholds. These thresholds are determined empirically as they relate to visible cloud observations. Thus the mean/median value provide a weighted observation of total cloud amount (areal coverage and thickness) within the sector. The variance gives an indication of the cloud texture or fragmentation (often an indicator of the physical "forcing" associated with a stratus event). The percent coverage gives an indication similar to the areal cloud coverage reported in surface observations.

In addition to using the sector statistics as input to other forecast techniques under development, the availability of sectorized and normalized values allows for development of a stand alone forecast algorithm/product based on time and space correlation of cloud coverage.

Figure 23. 80x80 km satellite data grid centered over SF Bay, divided into sectors. Grid of dots identifies satellite data points at 2 km resolution. Dark lines indicate sector boundaries. Shading indicates topography (elevation in meters) as identified in legend at top of figure.

Appendix C

Zi-Ceiling Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

Appendix C

Zi-Ceiling Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

Appendix C

Cloud Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

Appendix C

Cloud Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

Appendix C

Combo Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

Appendix C

Combo Models

Plots of 1996-1997 (training data)

Plots of 1998 (independent data)

REFERENCES

- Brick, W. C., 1988: Aircraft observations of marine inversion during MABLES WC: data processing and preliminary analysis. M.S. Thesis, Dept. of Meteorology, San Jose State University, San Jose, CA, 108 pp.
- Bridger, A. F. C., W. C. Brick, and P. F. Lester, 1993: The structure of the marine inversion layer off the central California coast: Mesoscale conditions. *Mon. Wea. Rev.*, **121**, 335-351.
- Cherna, E., A. Bellon, B. L. Austin, and A. Kilambi, 1985: An objective technique for the delineation and extrapolation of thunderstorms from GOES satellite data. *J. of Geophysical Research*, **90**, 6203-6210.
- Clark, D. A., and F. W. Wilson, 1996: The marine stratus initiative at San Francisco International Airport. Project Report ATC-252, MIT Lincoln Laboratory, Lexington, MA.
- Evans, J. E., T. J. Dasey, D. A. Rhoda, R. E. Cole, F. W. n Wilson, and E. R. Williams, 1999: Weather sensing and data fusion to improve safety and reduce delays at major west coast airports. Project Report ATC-290, MIT Lincoln Laboratory, Lexington, MA.
- Goodman, J. K., 1976: The microstructure of California coastal fog and stratus. Rep. No. 76-09 Dept. of Meteorology, San Jose State University, San Jose, CA, 61 pp.
- _____, 1977: The microstructure of California coastal fog and stratus. *J. Appl. Meteor.*, **16**, 1057-1067.
- Goodman, J. K., and A. Miller, 1977: Mass transport across a temperature inversion. *J. Geophys. Res.*, **82**, 3436-3471.

- Gruber, A., R. Ellingson, P. Ardanuy, M. Weiss, S. K. Yang, and S. N. Oh, 1994: A comparison of ERBE and AVHRR longwave flux estimates. *Bull. Amer. Meteor. Soc.*, **75**, 2115-2130.
- Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. of Climate*, **6**, 1587-1606.
- Lee, S. C., C. L. Lin, and S. W. Pepper, 1974: A theoretical study of steam fog. Paper presented at 1974 fall annual meeting of the American Geophysical Union, San Francisco, CA, Dec. 12-16, 1974.
- Lilly, D. K., 1968: Models of cloud-topped mixed layers under a strong inversion. *Quart. J. Roy. Meteor. Soc.*, **94**, 292-309.
- Mack, E. J., U. Katz, C. Rogers, and R. J. Pilie, 1974: The Microstructure of California coastal stratus and fog at sea. Calspan Rep. No. CJ-5404-M-1. 67 pp.
- Mazur, A., 1991: Wind and temperature profiles through the marine inversion: Mt. Sutro tower July 1979. M.S. Thesis, Dept. of Meteorology, San Jose State University, San Jose, CA, 171 pp.
- Miller, A., 1976: Wave properties in the west coast temperature inversion. Rep. No. 76-01, Dept. of Meteorology, San Jose State University, San Jose, CA, 95 pp.
- Neiburger N., and M. G. Wurtele, 1949: On the nature and size of particles in haze, fog, and stratus of the Los Angeles region. *Chem Rev.*, **44**, 322-335.
- Noonkester, V. R., 1979: Coastal marine fog in southern California. *Mon. Wea. Rev.*, **107**, 830-851.
- _____, 1982: Aerosol and humidity structure beneath maritime stratus clouds: 1981 data. Naval Ocean Systems Center, Technical Report 783, 128 pp.

- Noonkester, V. R., and L. E. Logue, 1976: Fog related to stratus clouds in southern California. Naval Electronics Laboratory Center, Technical Report 1989, 59 pp.
- Pillie R. J., E. J. Mack, C. W. Rogers, U. Katz, and W. C. Kocmond, 1979: The formation of marine fog and the development of fog-stratus systems along the California coast. *J. Appl. Meteor.*, **18**, 1275-1286.
- Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. L. Hartmann, 1989: Cloud-radiative forcing and climate: results from the earth radiation budget experiment. *Science*, **243**, 57-63.
- Simon, R. L., 1977: The summertime stratus over the offshore waters of California. *Mon. Wea. Rev.*, **105**, 1310-1314.
- Strach, W. J., 1991: The effects of summer time stratus at San Francisco International Airport on the nationwide flow of commercial airline traffic. First Aviation Weather Workshop, Kansas City, MO.
- Telford, J. W. and S. K. Chai, 1984: Inversions, fog, stratus, and cumulus formation in warm air over cooler water. *Bound. Layer Meteor.*, **29**, 109-137.
- _____, 1993: Marine fog and its dissipation over warm water. *J. Atmos. Sci.*, **50**, 3336-3349.
- Weinreb, M. P., M. Jamison, N. Fulton, Y. Chen, J. X. Johnson, J. Bremer, C. Smith, and J. Baucom, 1997: Operational calibration of Geostationary Operational Environmental Satellite-8 and -9 imagers and sounders. *Applied Optics*, **36**, 6895-6904.
- Wilson, F. W. and D. A. Clark, 1999: Interim Report: Products of the SFO marine stratus initiative. Weather Project Memorandum No. 43PM Wx-0058, MIT Lincoln Laboratory, Lexington, MA.