**The effect of ICAO type aerodrome weather forecasts on aircraft operations**

****Mr David Wilson, Senior Transport Safety Investigator, Australian Transport Safety Bureau (ATSB)**

*David has worked at the ATSB since early 2010. During this time David has conducted several systemic investigations. These include the analysis of Amateur-built aircraft accidents, investigating a series of near collisions, and the evaluation of reliability in the Australian aviation weather forecasting system. Prior to this David worked as a motor vehicle safety engineer, has degrees in Aeronautical Engineering and Physics, and is a private pilot who enjoys working on the art of aerobatics.*

# Introduction

The burgeoning of global data storage and categorization, coupled with the accessibility of modern computing power, presents unique opportunities for aviation safety investigations. The analysis of aviation data sets is increasingly being used to support and drive aviation safety investigations at the Australian Transport Safety Bureau (ATSB).

The likelihood and severity of specific safety related occurrences can be assessed with increased confidence as data analysis techniques are developed. These data analysis techniques complement expert knowledge and opinion, providing verification and objective assessment of investigation hypotheses, findings and recommendations.

Two case studies are presented to illustrate how the ATSB has used new data analysis techniques to complement weather related investigations. The discussion will focus on the investigative applications of a simulation developed that used historical aviation weather data to evaluate the likelihood and severity of flight crew retrieving ICAO type terminal area forecasts and trend forecasts (TAF and TTF), particularly where insufficient warning was provided when reported conditions were unsuitable for landing.

Selected results are presented for TAF and TTF reliability analyses performed to support two weather related investigations. This involved the analysis of Australia’s busiest capital city airports, four remote islands and Mildura Airport. Key results will focus on variations in reliability of forecasts for aspects such as time of day, weather phenomena and landing minima criterion.

These examples demonstrate how data modelling can be used to positively complement investigations, and how this opens the door for predictive, whole of system analysis.

# Context

Over the past 15 years, several unforecast weather episodes relating to flights into major Australian airports have led to unforeseen diversions, holding, and in some cases, landing below published safe limits.[[1]](#footnote-1) This research was initiated as a result of these occurrences, and because of the limited amount of Australian-based research in this area. The objective of this research was to understand how the reliability of weather forecasts affects the ability of flight crew to conduct safe landings at Australian airports.

To identify how weather forecasting at individual airports can potentially impact on safety, a comparative analysis was conducted between all aviation weather forecasts and coinciding reported observations. A further aim was to understand the effects that weather forecasting can have on aircraft engaged in air transport-related operations, specifically how weather forecasting may affect safe landing at airports.

While many studies have been performed on forecast verification, there was very limited research on how the weather forecasting system affected pilots. This study was performed by calculating the historical likelihood of the effect of forecasts if they were used according to pre-defined rule-sets, time of retrieval, the aircraft specifications, aerodrome equipment and flight scheduling. Note that the aim is to simulate the effect of pilots retrieving and using forecasts for operations.

The results will assist aircraft operators to focus on the highest risk seasons and times of day for weather reliability, facilitating better flight planning and support for pilots. They will also allow for more informed prioritization of investment decisions about aircraft and aerodrome navigational equipment.

# Methods

The results presented required the development of a computer program (the algorithm). Full details are available on the ATSB’s website ([www.atsb.gov.au](http://www.atsb.gov.au)) under investigation AR-2013-200.

The objective of the algorithm was to help calculate the probability of using the weather forecast system as it affects pilots, and number of expected aircraft arrivals during selected weather conditions. It uses actual historical weather forecasts (TAFs and TTFs) compared to reported observations (METARs and SPECIs) as they would be used by pilots. It also can examine what would have happened if a different type of forecast was used or the forecast was retrieved later in flight.

This is useful to help focus attention on higher risk areas, and to identify if the overall risk is too high. This places a tangible number on the risk and assists decision making when considering safety issues and recommendations.

The primary objective of the analyses presented was to calculate the probability of a pilot retrieving a TAF predicting conditions above the alternate minima and subsequently arriving during unsuitable landing conditions.

Two assumptions were applied regarding the operational effect of TAFs and reported observations:

* TAFs forecasting conditions below the alternate minima were assumed to warn a flight crew that a contingency plan was required.
* Reported observations below the landing minima were assumed to be unsuitable for landing.

Alternate and landing minima were obtained from the Airservices Australia Instrument Approach Procedures.

# Applications for investigations

The techniques and algorithm developed have been used to complement two recent investigations, providing tangible evidence for hypothesis testing for safety issues and supporting investigators when considering safety recommendations.

## Comparison between remote islands and mainland airports

This algorithm was used by the ATSB to support an investigation of a serious weather-related incident at a remote Australian island. Specific hypotheses were put forward regarding the reliability of remote island forecasting, and the analysis conducted allowed definitive conclusions to be drawn regarding the relative reliability of these forecasts from the perspective of the historical likelihood of unforecast weather conditions below the landing minima.

The ATSB compared reported weather observations (METARs and SPECIs) to aerodrome forecasts (TAFs) at Australia’s four remote island aerodromes (Norfolk, Cocos/Keeling, Christmas and Lord Howe Islands) and Australia’s five busiest airports (Brisbane, Sydney, Melbourne, Adelaide and Perth) between 2009 and 2014.

These were obtained for a Category C aircraft not equipped for RNAV approaches to match a specific aircraft type. As such RNAV based procedures were excluded from the analysis of all airports.

### Comparison of TAF effectiveness and weather for Australia’s remote islands and five busiest airports (conditions below landing minima)

To evaluate the relative likelihood of TAF effectiveness, TAFs for Australian remote islands were compared with TAFs for the top five Australian capital city airports by aircraft movements. The same type of simulations was used using a Category C aircraft without RNAV approaches.

Figure 1 shows observations below the landing minima in hours per year for each airport. This is divided for when the TAF predicted conditions above the alternate minima (yellow), between the alternate and landing minima (orange) and below the landing minima (blue). The average of years 2010 to 2013 is shown. These year ranges were selected due to common data availability across all the airports.

As indicated in Figure 1:

* Remote islands generally had more observed weather below the landing minima compared to the five busiest capital city airports. This was likely due a combination of factors such as more accurate instrument approaches available to Category C aircraft (excluding RNAV) at the capital cities, and more adverse weather conditions at the remote islands.
* Compared to the other remote islands, the Cocos Islands had the least weather below the landing minima (about 180 hours per year) and Christmas Island[[2]](#footnote-2) had the most weather below the landing minima (about 580 hours per year). Norfolk and Lord Howe Island had about the same amount of weather below the landing minima (just under 300 hours per year). However, Norfolk and Christmas Island had a similar amount of weather below the landing minima where the TAF was above the landing minima (orange and yellow combined).[[3]](#footnote-3)

Figure 1: Observations below landing minima (hours per year) by TAF above alternate minima, between alternate and landing minima and below landing minima averaged for simulated TAF retrieval 1 to 3 hours prior to arrival for a Category C aircraft without RNAV at selected airports, 2010 to 2013[[4]](#footnote-4)

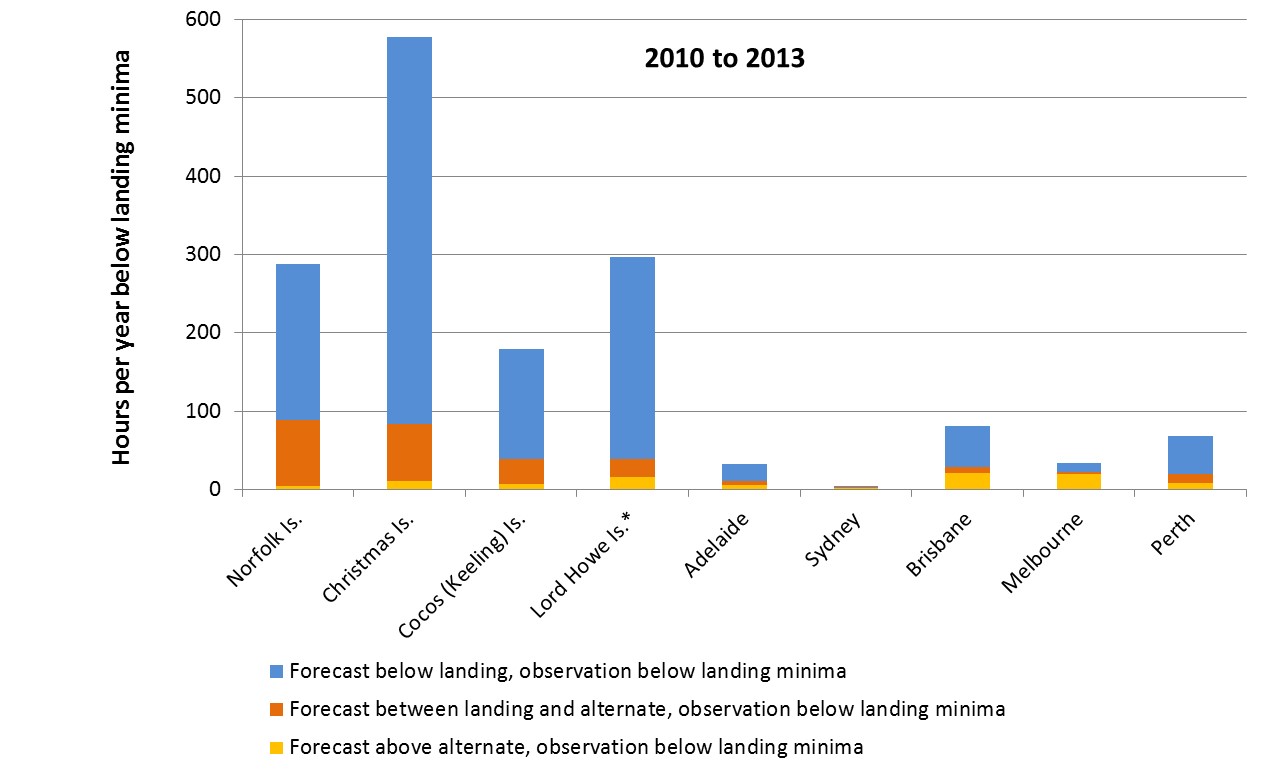
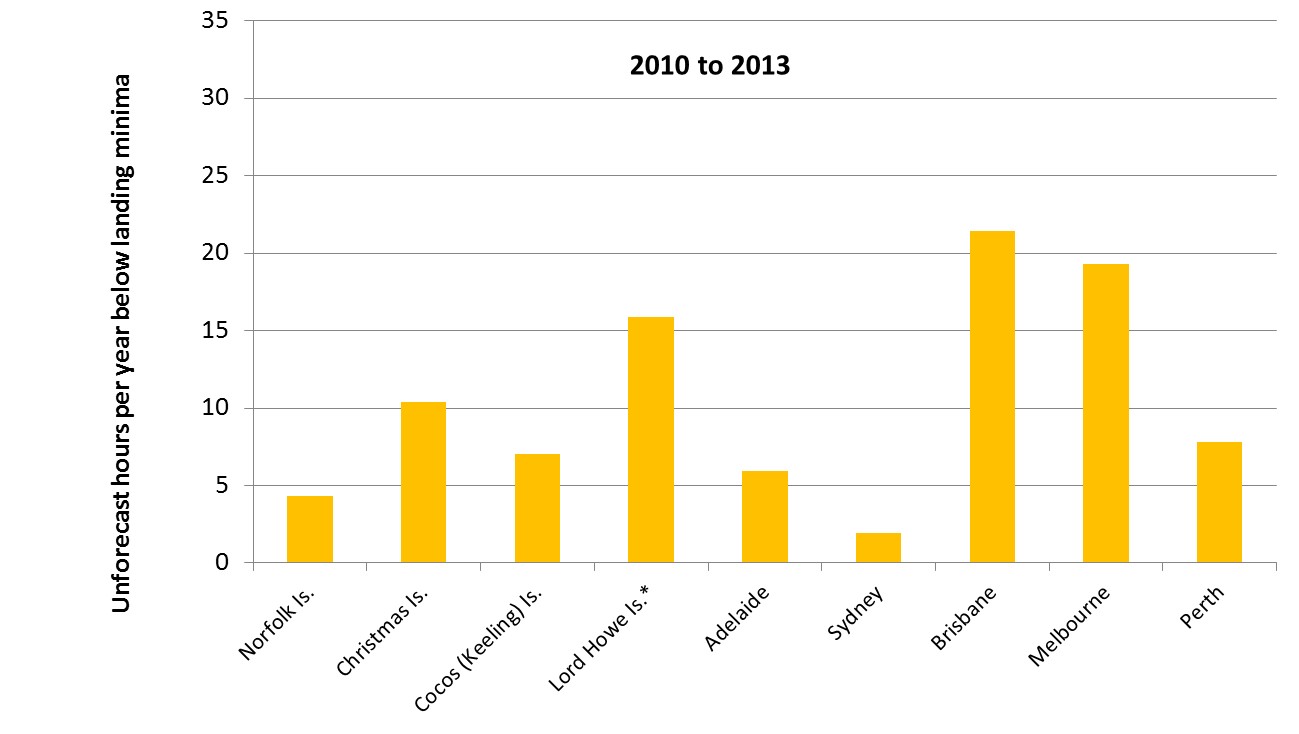


Figure 2 shows an expansion of the unforecast observations (or misses) below the landing minima depicted by yellow bars in Figure 1. This data indicates the likelihood (displayed in hours per year) of retrieving a TAF between 1 to 3 hours prior to arrival that predicts conditions above the alternate minima, and arriving during conditions below landing minima for the Category C aircraft.

Figure 2: Unforecast observations below landing minima (hours per year) averaged for simulated TAF retrieval 1 to 3 hours prior to arrival for a Category C aircraft without RNAV at selected airports, 2010 to 20135



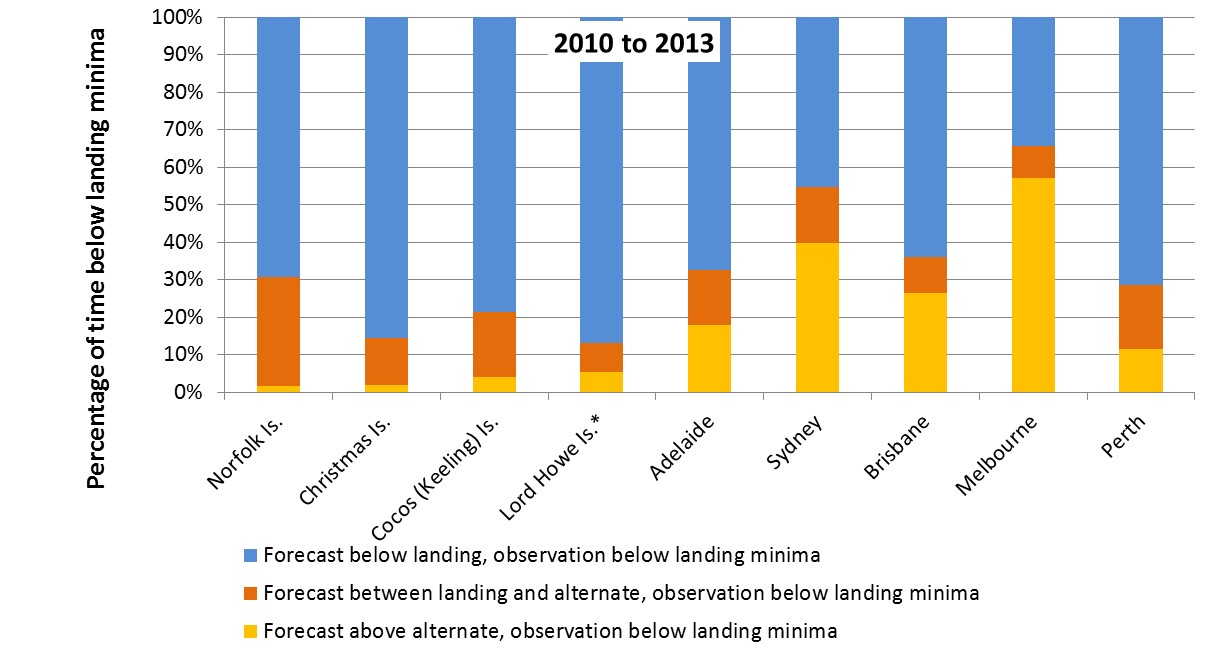
As indicated in Figure 2:

* There was significant variability across the airports in the amount of time when there was unforecast observations below the landing minima.
* Overall, remote island airports had a similar amount of time with unforecast conditions below the landing minima from 2010-2013 about 9.4 hours of unforecast weather below the landing minima to capital city airports with an average of about 11.3 hours.

Although remote islands had a similar amount of unforecast weather below the landing minima as the capital city airports, it should be noted that the potential consequences are more serious, given the locations of the airports.

Figure 2 provided data on the total time below the landing minima that was forecast and not forecast. To provide an indication of the relative accuracy of predictions between the airports, Figure 3 shows the same data as for Figure 1 calculated as a percentage of the total conditions below the landing minima at each airport for the Category C aircraft.

Figure 3: Percentage of time below landing minima by TAF above alternate minima, between alternate and landing minima and below landing minima averaged for simulated TAF retrievals 1 to 3 hours prior to arrival for a Category C aircraft without RNAV at selected airports, 2010 to 2013



As indicated in Figure 3:

* Remote islands had lower proportions of conditions below the landing minima that were unforecast (in yellow) compared to the selected capital city airports, with all the islands being less than half of the proportion of any selected capital city airport.
* Norfolk and Christmas Islands had the lowest proportions of conditions below the landing minima of the four remote islands.

Statistical tests were performed by grouping each the unforecast weather periods together. For example, in 2009 there were significantly less[[5]](#footnote-5) episodes with unforecast TAFs at Norfolk Island compared with all capital cities combined.

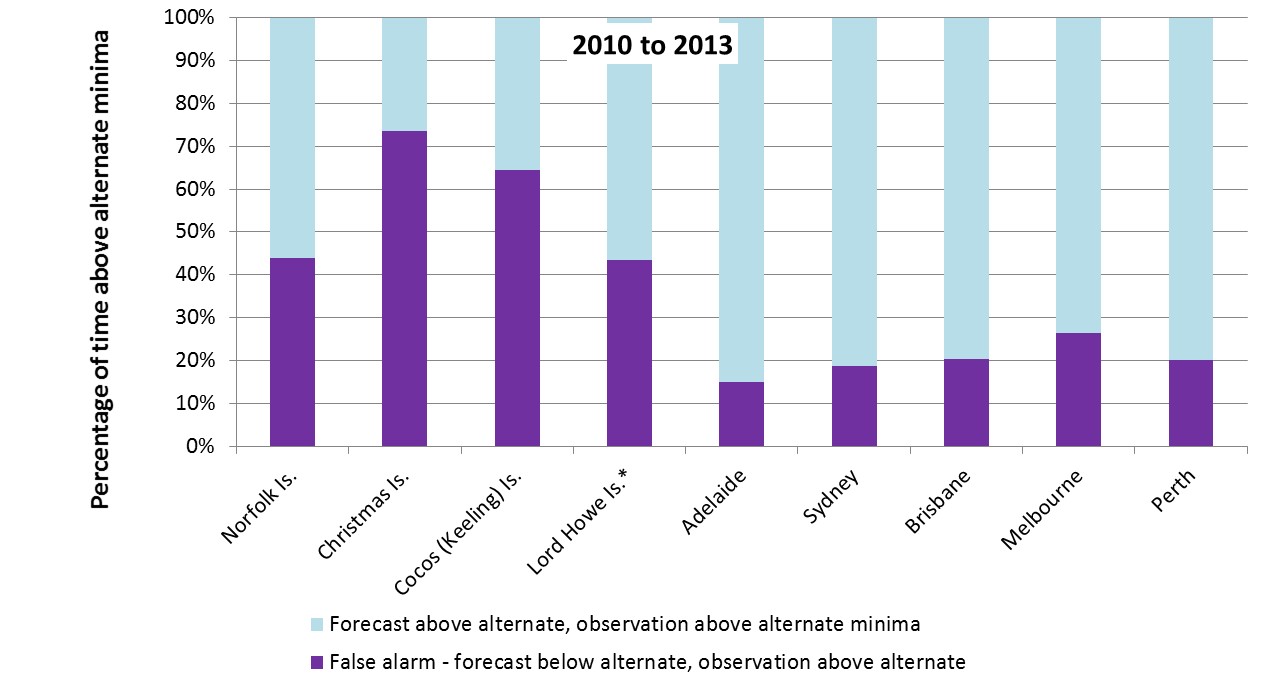
This data provided context for an ATSB investigation when considering the potential for a safety issue by providing the information required that was previously unattainable, dispelling commonly held beliefs about forecasting at islands.

### Comparison of TAF effectiveness (conditions above alternate minima)

The columns in Figure 4 show percentage of the total conditions above the alternate minima at each airport for a Category C aircraft without RNAV. The lower purple bars show the total percentage of time of false alarms each year (where forecast conditions were below the alternate minima, and conditions on arrival 1 to 3 hours after were observed above the alternate minima).

As indicated in Figure 4, the false alarm rates for the capital city airports averaged about 20 per cent during 2010-2013, being considerably lower than the remote island airports.

Figure 4: Percentage of time above alternate minima by TAF above alternate minima, between alternate and landing minima and below landing minima for averaged for simulated TAF retrievals 1 to 3 hours prior to arrival for a Category C aircraft without RNAV at selected airports, 2010 to 2013



### The effect of time of retrieval prior to arrival on TAF reliability at Norfolk Island Airport, 2009 to 2014

Figure 5 shows the time of unforecast observations below the landing minima per year as a percentage of all time (any type of observed weather) for each quadrant of the day (starting from 0000 local time). The horizontal axis represents the time before arrival when pilots retrieved TAFs up to 900 minutes (15 hours) prior to arrival.

This figure shows that likelihood of arriving during unforecast observations below the landing minima increased as the duration between TAF retrieval and arrival increased. This was particularly important for arrival in the evening (between 1800 and 0000 local time) at Norfolk Island. For example, if a TAF was retrieved 6 hours prior to arrival, the chances of arrival during unforecast conditions below the landing minima are more than two times greater than if a TAF was retrieved 2 hours prior to arrival.

From the perspective of operational decision making, this shows that forecasts retrieved at the latest possible time (before the point of no return) provided a better warning of potential conditions below the landing minima at Norfolk Island Airport during 2009-2014.

Figure 5: Likelihood of arriving during unforecast observations below the landing minima by quadrant of day for simulated TAF retrieval times 0 to 900 minutes prior to arrival, Norfolk Island Airport, 2009 to 2014 for a Category C aircraft without RNAV

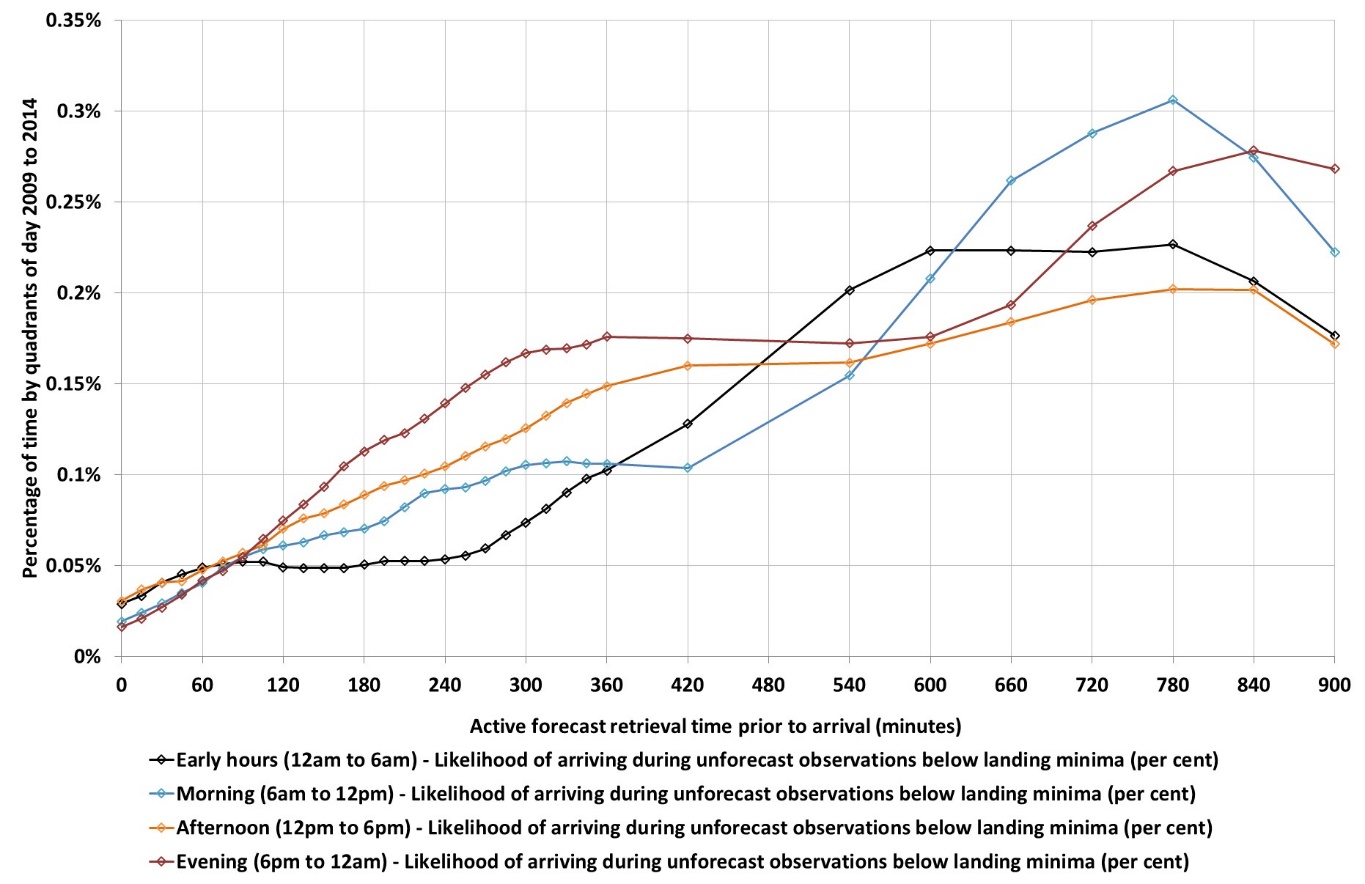
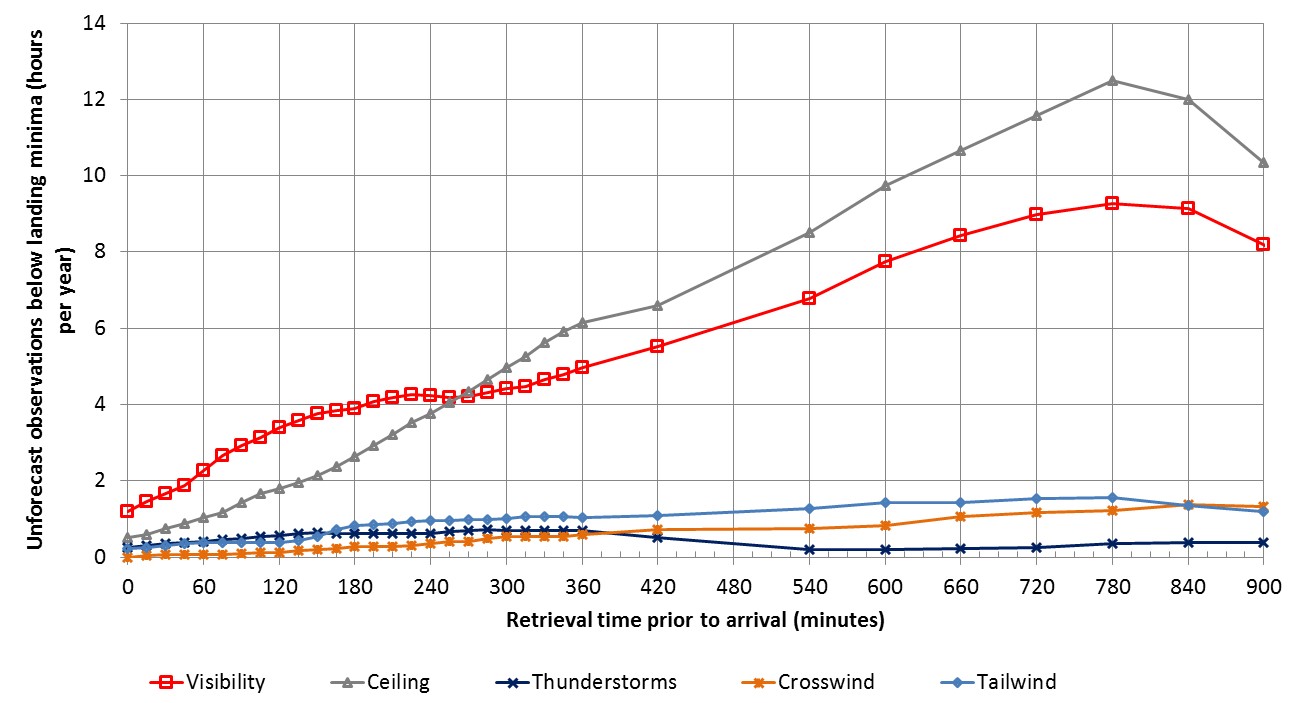


Figure 6 shows the hours each year of unforecast observations below the landing minima for each of the minima criteria for TAF retrievals up to 15 hours prior to arrival. Increasing the TAF retrieval time prior to arrival had the greatest effect for reported visibility and ceiling below the landing minima.

Although low ceiling was more prevalent overall, it was more common to arrive during unforecast visibility below the landing minima when pilots retrieved TAFs up to 4.5 hours prior to arrival.

Figure 6: Likelihood of arriving during unforecast observations below the landing minima by landing minima parameter for simulated TAF retrieval times 0 to 900 minutes prior to arrival, Norfolk Island Airport, 2009 to 2014 for a Category C aircraft without RNAV



### Conclusions from analysis

The ATSB’s analysis of weather forecasting was conducted using a Category C aircraft without capacity to conduct RNAV approaches. Different results would be expected if a different category aircraft or other assumptions were made regarding the types of approaches that were available.

Key results from the ATSB’s analysis included:

* The remote islands had considerably more weather below the landing minima compared to the five busiest capital city airports in Australia.
* Overall, false alarms were more prevalent at remote islands than at capital city airports. This, combined with the lower proportion of unforecast observations below the landing minima indicates that forecasting for remote islands was more conservative than forecasting for capital city airports.
* From the perspective of operational decision making, forecasts retrieved at the latest possible time (before the point of no return) provided a better warning of potential conditions below the landing minima at Norfolk Island during 2009-2014. In other words, as the time between retrieving a TAF and arriving at the airport increased, the likelihood of encountering unforecast weather increased. This was particularly applicable for aircraft arrivals in the evening between 1800 and 0000 local time for a TAF retrieved up to 6 hours prior to arrival.

## Selected results from research investigation AR-2013-200

This section shows selected results from the ATSB research investigation AR-2013-200 to demonstrate how weather forecasting was likely to affect aircraft arrivals between 2009 and 2013 at Adelaide and Mildura Airports. These airports related to a significant ATSB investigation. On 18 June 2013, involved two Boeing 737-800 aircraft that encountered unforecast weather en route to Adelaide leading to a diversion to Mildura Airport. Upon arrival, both aircraft encountered weather unsuitable for landing. Both aircraft landed in conditions below the published safe limits, with one aircraft also landing in a fuel critical state (ATSB investigation [AO-2013-100](http://www.atsb.gov.au/publications/investigation_reports/2013/aair/ao-2013-100/)).

An analysis of the systemic reaction time to unexpected poor weather is presented. The ATSB report AR-2013-200 has been released covering the full analysis and techniques used for Mildura and Adelaide Airports.

### Calculating the impact of forecasting on aircraft operations – comparing the expected overall effects of using TAFs instead of TTFs – Adelaide Airport

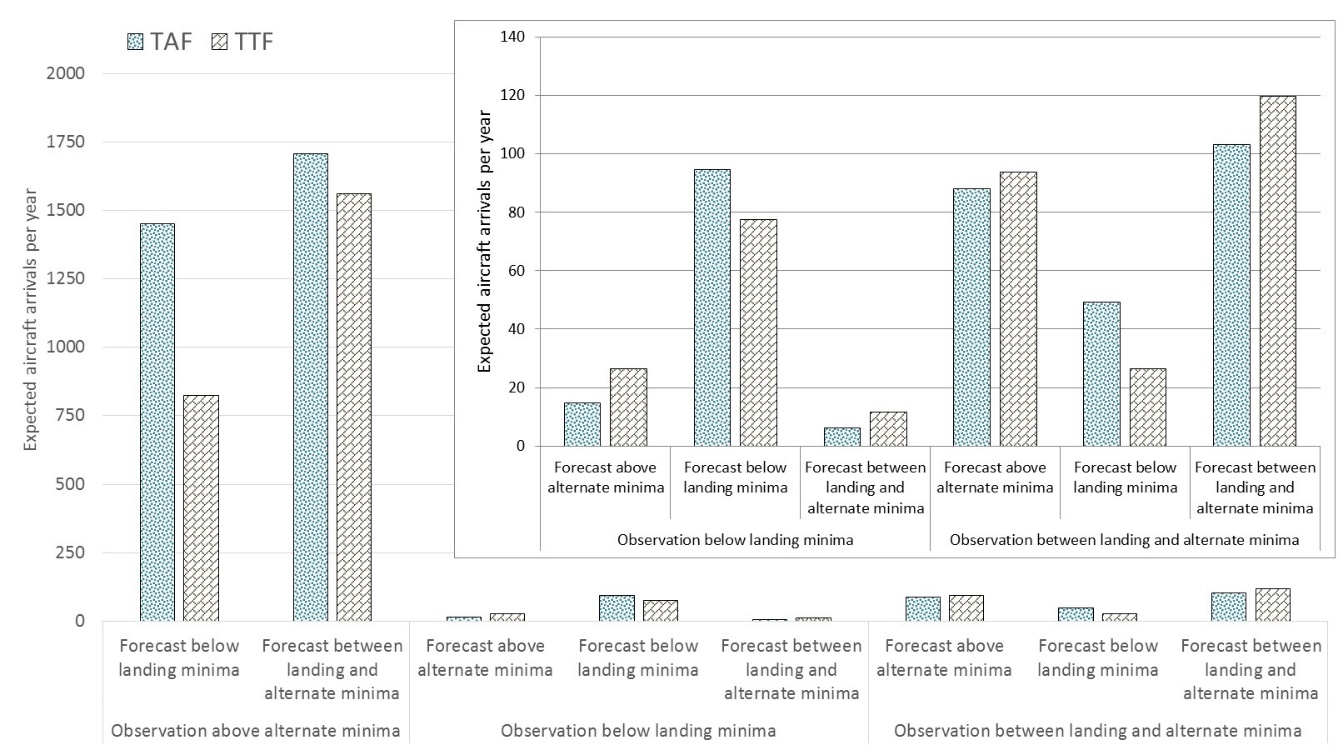
The following section shows an example of comparing the effectiveness of two types of forecasts on historical aircraft operations. In this case, this is showing the expected aircraft arrivals if solely using a TAF or solely using a TTF, retrieved 1 hour prior to arrival (on average).

Although this was aimed at identifying the number of aircraft likely to arrive in these scenarios, the same techniques could be used to test different scenarios, such as measuring the effect of the installation of new prescribed instrument approach procedures, changing flight schedules or ICAO guidelines. This is an example indicating the predictive nature of these techniques.

Comparisons were performed between reported observations (METARs and SPECIs) and forecast predictions (TAFs and TTFs). Simulated forecast retrievals between 0 and 2 hours prior to arrival (in 15 minute intervals) were used to estimate the likelihood of flight crews retrieving an unprimed forecast. An aircraft arrivals model was then used to estimate the number of aircraft expected to arrive during these times. This time frame was expected to provide the best-case scenario for forecast retrieval.

The number of expected aircraft arrivals each year for all combinations of forecasts and observations are shown in Figure 7. That is, above the alternate minima, between alternate and landing minima, and below landing minima (except when both forecast and observations were above the alternate minima). This assumes the crew are using forecasts from either the TAF or TTF, but not both. The inset graph on the right of Figure 7 shows an expanded scale of observed conditions below the alternate minima. The left side of the graph shows when the forecasts were false alarms as observations were above the alternate minima.

Figure 7: Expected distribution of aircraft arrivals per year for sole use of TAF or TTF forecasts, simulations averaged for retrieval 0 to 2 hours prior to arrival (excluding correct rejections), Adelaide Airport 2009 to 2013



Pilots retrieving TTFs for Adelaide were expected to arrive the most often during unforecast observations below the landing minima. However, the number of expected arrivals with false alarms was relatively lower for TTFs compared to TAFs.

The following are selected findings comparing TAFs and TTFs at Adelaide Airport.

* + Fifteen aircraft per year, if using a TAF, would be expected to arrive at Adelaide during an unforecast observation below the landing minima.
  + Twenty-seven aircraft per year, if using a TTF, would be expected to arrive at Adelaide during an unprimed observation below the landing minima.
  + The largest difference in unforecast TAFs and TTFs below the landing minima were found for observations of fog/mist/low cloud.

False alarms (including fine weather periods surrounded by conditions below the alternate minima) created the largest operational differences between forecast types (TAFs or TTFs):

* + A difference of around 773 more aircraft per year were expected to arrive at Adelaide with false alarms if only using TAFs instead of TTFs (following the stipulated rule sets for each forecast).[[6]](#footnote-6)

About 90 aircraft per year would be expected to arrive during conditions between the landing and alternate minima without warning from a forecast (88 for a TAF, 94 for a TTF). This indicates where the safety buffer between the alternate and landing minima values of ceiling or visibility has positive safety benefits.

### Expected aircraft arrivals during unforecast conditions by time of day – Mildura Airport

This section shows the effect of combining an expected aircraft arrivals model with unforecast conditions below the landing minima by each hour of day at Mildura Airport. This high-fidelity data allows the identification of the historical risk to flight operations.

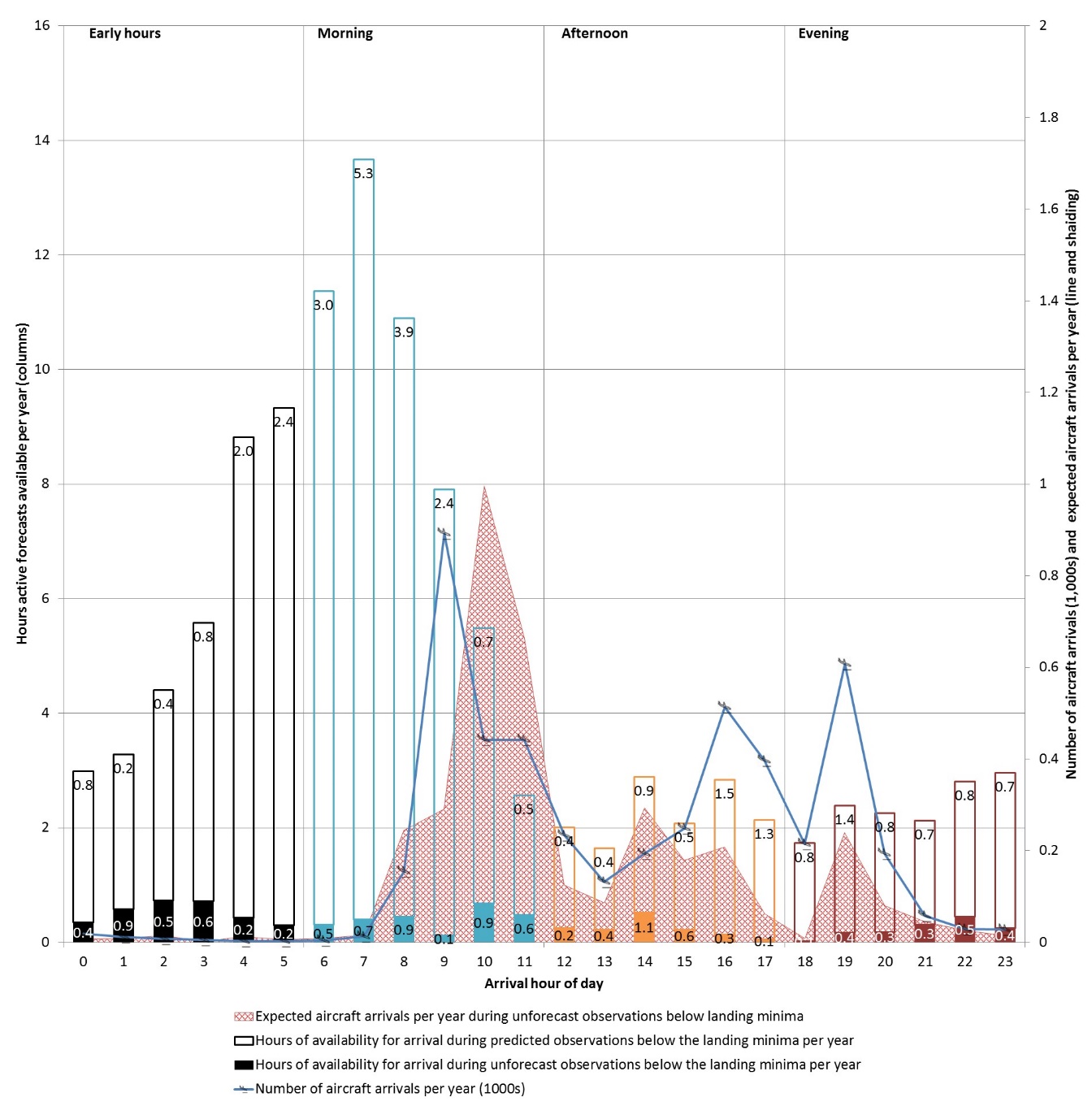
The frequency of all observations below the landing minima varied significantly at Mildura across the day. This can be seen by the combined solid and transparent columns in Figure 8. Variation also differs between forecasts above the alternate minima (solid columns) and forecasts below the alternate minima (transparent columns) across the 24 hours. Many of these observations below the landing minima coincided with times outside of the common arrival times for aircraft landing at Mildura (blue line).

Figure 8 also shows the estimated number of aircraft affected each year by unforecast observations below the landing minima by hour of day (red shading). The most common time that aircraft would have been affected by unforecast conditions was between 1000 and 1100 (and surrounding hours from 0800 to 1200). Between 1000 and 1100, it was estimated that 1.0 aircraft were affected per year on average.

The hour from 0900 to 1000 represents the hour of day where the most aircraft arrived, and coincided with a high number of total hours below the landing minima (either primed or unprimed). Between 1000 and 1100, however, more aircraft were affected by unprimed observations due to the higher hours of availability of unprimed observations below the landing minima even though there were a lower number of arrivals.

While there are relatively low numbers of affected aircraft in the early morning, when there was a higher availability of forecasts with unprimed conditions (from midnight and 0600), there was a higher likelihood per arrival for any aircraft electing to arrive during this time.

Figure 8: Hours active forecasts were available for observations below the landing minima (columns), thousands of aircraft arrivals (blue line), and expected aircraft arrivals during unforecast observations below the landing minima (red shading) per year, by hour of day, averaged for forecast retrieval times between zero to 2 hours prior to arrival 2009 to 2013. Column labels indicate the number of weather episodes per year for each category and are additive.



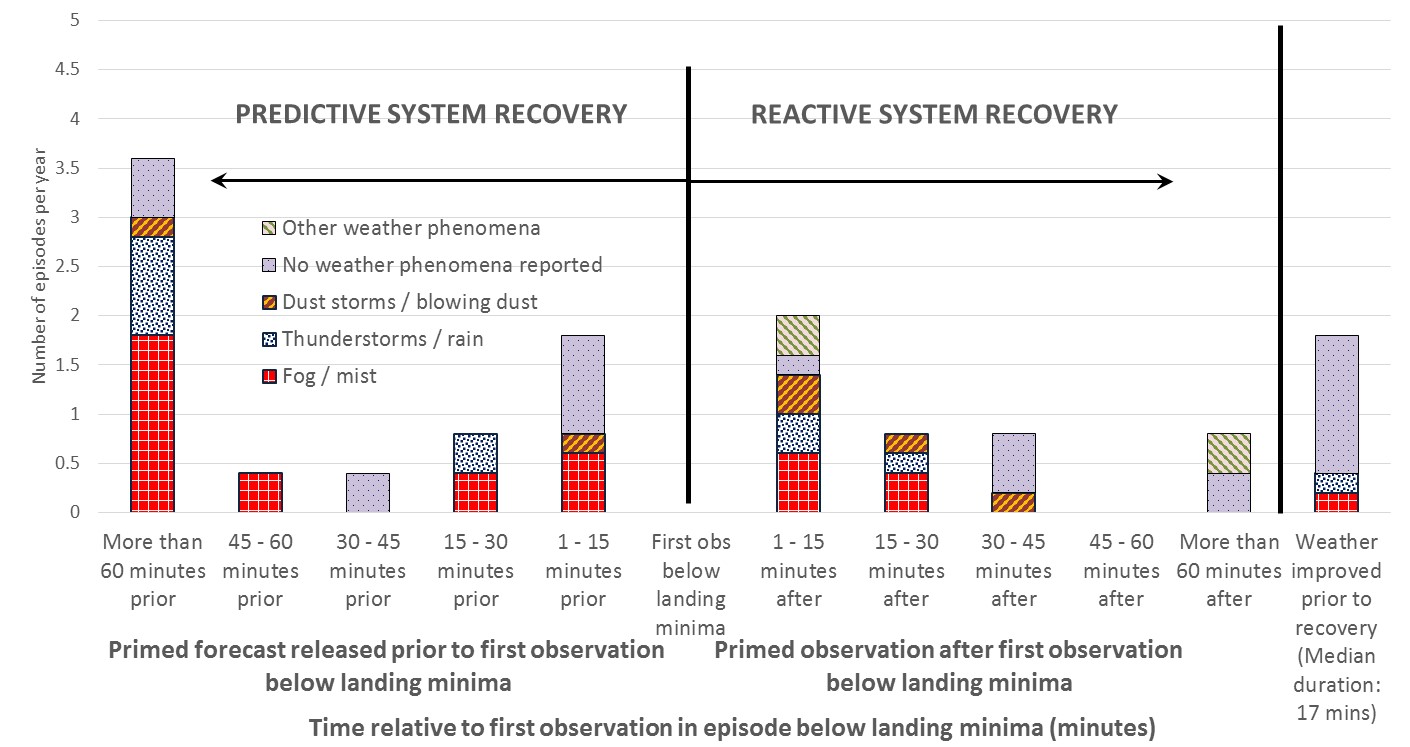
### Investigating overall systemic effects – forecast release reaction time to the onset of poor weather – Mildura Airport

A question of interest for investigation AO-2013-100 was how long it took for a forecast to be released following unexpected weather observation below the landing minima. The following shows an analysis used to calculate this.

Figure 9 shows episodes per year where TAFs that were not forecasting weather conditions below the alternate minima were active prior to or following an observation below the landing minima, however, a subsequent forecast was released predicting conditions below the alternate minima. The labels ‘predictive’ and ‘reactive’ system recovery refer to relative timing of the forecast correction.

Episodes relating to reactive scenarios are shown to the right of the first observations below landing minima in Figure 9, and indicate situations where weather below the landing minima has occurred, followed by the system recovering with a correctly forecasted observation. This can be considered as the reaction time to unforecast observations below the landing minima. This shows that for about five episodes every 2 years, the system takes longer than 15 minutes to predict conditions below the alternate minima following unforecast observations below the landing minima. This is in addition to around least two episodes every year on average that no prediction exists.

Figure 9: Below alternate minima forecast release times relative to first observation below landing minima in each weather episode (Predictive system recovery and reactive system recovery) and unforecast episodes with no recovery by observed weather phenomena, Mildura Airport 2009 to 2013



# Conclusions

The expanding availability of safety related data combined with modern computer power enables new techniques to be developed to measure transport safety risk and support the identification of safety issues and recommendations in aviation safety investigations.

This has been demonstrated by the ATSB’s predictive weather algorithm, which has been used to support the findings of two major investigations. This was achieved by comparing forecasts and observations in a way that affects pilots and their passengers, instead of for forecast verification purposes. Using the maximum fidelity of the data provided the best opportunity to identify high risk areas.

Combining historical likelihood of weather forecasting with expected aircraft arrivals allowed the historical impact on aircraft operations to be estimated.

As demonstrated, this simulation has applications ranging from investigation of weather reliability at any aerodrome with a TAF or TTF. The example shown demonstrates the relative impacts to actual aircraft operations of using a TAF in lieu of TTF forecasts. In the same fashion, the algorithm can be used to determine the impact of guideline changes to any ICAO type forecast. Furthermore, the prediction of the safety impact of changes to weather forecasting systems before they are implemented is possible. For example:

* the installation of advanced airport equipment and aircraft instrumentation
* changes to pilot strategies when using and interpreting forecasts
* additional contingency planning
* using multiple or different forecasts before and during a flight
* changing of flight schedules
* retrieving a forecast closer to arrival.

This may help investigations providing tangible, repeatable evidence when testing hypotheses for safety issues and recommendations.

Further research

Further data driven research is planned utilizing this algorithm. This is the first report in a series covering Australian airports supporting regular public transport operations.

# Contact

Email: david.wilson@atsb.gov.au

1. ATSB investigations of incidents related to unforecast weather include: [199904029](http://www.atsb.gov.au/publications/investigation_reports/1999/aair/aair199904029/), [200401270](http://www.atsb.gov.au/publications/investigation_reports/2004/aair/aair200401270/), [200605473](http://www.atsb.gov.au/publications/investigation_reports/2006/aair/aair200605473/), [AO-2012-073](http://www.atsb.gov.au/publications/investigation_reports/2012/aair/ao-2012-073/), [AO-2015-067](http://www.atsb.gov.au/publications/investigation_reports/2015/aair/ao-2015-067/), and [AO-2013-100](http://www.atsb.gov.au/publications/investigation_reports/2013/aair/ao-2013-100/), available on the ATSB’s website at www.atsb.gov.au. [↑](#footnote-ref-1)
2. Observations below the landing minima at Christmas Island had a similar median release frequency to Norfolk Island Airport (about 23 minutes), despite routine observations (METARs) at Christmas Island being released every hour (instead of every 30 minutes). Based on this it was assumed that the SPECI trigger mechanisms at Norfolk and Christmas Islands were similar, reducing the risk of over counting time below the landing minima at Christmas Island. [↑](#footnote-ref-2)
3. Although TAFs at Christmas Island were quoted in ERSA as being released twice daily, the vast majority of TAFs (98 per cent) were released within every 7 hours of one another. This was very similar to Norfolk Island TAF and amended TAF release frequency. [↑](#footnote-ref-3)
4. Approximately 28 per cent of Lord Howe Island data was not decoded - numbers are expected to be larger than presented. [↑](#footnote-ref-4)
5. Χ2, P < 0.05, 1df. Norfolk Island 2009 vs Capital cities 2009, TAF above alternate minima during conditions below the landing minima as a proportion of all conditions below the landing minima. [↑](#footnote-ref-5)
6. TAFs have associated 30 minute timing buffers added to any prediction below the alternate minima. [↑](#footnote-ref-6)