# 2016 Oman Rainfall Enhancement

## Final Report

## February 2017

NATIONAL INSTITUTE FOR APPLIED
STATISTICS RESEARCH AUSTRALIA
University of Wollongong
NSW 2522
Australia
Tel: +61 2 4221 5435
Fax: +61 2 4221 4845

# Copyright, Disclaimer and Acknowledgements



**Prepared by:**

**Dr. Stephen Beare**
National Institute For Applied Statistics Research Australia
University of Wollongong
NSW 2522
AUSTRALIA

**Prof. Ray Chambers**
National Institute For Applied Statistics Research Australia
University of Wollongong
NSW 2522
AUSTRALIA

**Ms Mandy Moore**
National Institute For Applied Statistics Research Australia
University of Wollongong
NSW 2522
AUSTRALIA

**Mr. Scott Peak**
Australian Rain Technologies
Pier 8/9, 23 Hickson Rd, Millers Point
NSW 2000
AUSTRALIA

# Table of Contents

# 1 Executive Summary

The 2016 Atlant trial was the fourth consecutive year of an experimental rain enhancement program scheduled to run for six years, and conducted from June through October in the Hajar ranges of Oman. While the geographic scope of the project has expanded in each year, the experimental design, based on a randomised operating schedule with equal numbers of Atlants scheduled to operate / not operate each day, and the statistical methods of analysis have been maintained.

The year to year consistency in the trial methodology has allowed the focus of the statistical analysis of the trial data to be moved from consideration of data obtained in individual trials to data obtained across multiple trials. This along with the expanded number of Atlant sites and increased instrumentation covering a wider area has allowed for greater control for the high level of variability of natural rainfall, both over time and over the trial area, in the analysis.

The report on the 2015 trial included an analysis of data collected over the three years 2013-15. This indicated an estimated enhancement of 18 per cent over this three year period with a confidence level for a positive Atlant effect in excess of 99.9 per cent. This analysis has been repeated, with the inclusion of the 2016 trial data, to now cover the four years 2013-16. It can be seen in Table 1-1 that the additional 2016 trial data has further strengthened this result, with an estimated enhancement over the four years 2013-16 of 19.2 per cent and with a confidence level for a positive Atlant effect now more than 99.99 per cent. The added confidence derived from additional trials is clear even though 2016 was a low rainfall year across both target and control areas, which also limited the data available for estimation of a 2016-specific enhancement effect.

**Table 1-1 Summary results for the 2013-2016 Atlant trials**

| Year | Estimated Enhancement | Probability of a positive enhancement |
|---|---|---|
| **2013** | 18  per cent | 99 per cent |
| **2013-14** | 18.5 per cent | 99 per cent |
| **2013-15** | 18  per cent | 99.9 per cent |
| **2013-16** | 19.2 per cent | 99.99 per cent |

The consistency and level of precision of these results is unprecedented in the trialing of rainfall enhancement over 60 years. It  can be attributed to the extensive rainfall and meteorological data that have been collected, experimental design and the use of advanced fit for purpose statistical methodology for analysis of these data.

The four trials carried out to date support a convincing argument that the continued trialing of Atlant over the next two years is warranted. They also make a strong argument for the deployment of the Atlant technology as a long-term low-cost addition to Oman’s water supply.

# 2 Introduction

The 2016 Atlant trial was the fourth consecutive year of a program scheduled to run for six years, conducted from June through October in the Hajar ranges of Oman. The program is jointly run by the Australian Rain Technologies and the Trading and Investment Enterprise of Oman (TIE) with funding from the Omani government.

The analysis of the three previous trials all concluded that there was a significant rainfall enhancement signal. The percentage increase in rainfall in each year was more than 15 per cent. A comprehensive analysis of the combined data from the three trials 2013-15, gave an estimated enhancement effect of 18 per cent with a confidence level of more than 99 per cent.

While the geographic scope of the trial has expanded in each year, the experimental design, based on a randomised operating schedule with equal numbers of Atlants scheduled to operate / not operate each day, and the analytical methodology developed in the first year, which was a precursor to embarking on a five-year program, have been retained. In 2015 the primary focus of the statistical analysis moved from the analysis of individual trials to analysis of the combined data from all trials and a new robust permutation-based test statistic was developed. This approach, which still includes an analysis of the current year trial data has again been adopted in 2016. The report is structured as follows:

- An overview of the Atlant technology and experimental design
- A summary of changes in the physical design of the trial and the data collected;
- A comparison of meteorological conditions in 2016 and previous trial years;
- A discussion of the implications of and approaches adopted to deal with unanticipated issues arising in the 2016 trial;
- A summary of the statistical methodology and presentation of the 2016 trial results;
- A presentation of the combined four-year (2013 - 2016) trial results;
- Concluding comments; and
- Technical appendices covering the implementation of the experimental design and some of the statistical results.

# 3 Overview of the Atlant and experimental design

## 3.1 Atlant Technology

Atlant is a ground-based cloud ionisation system that is hypothesised to increase the amount of rainfall on the ground rather than initiate the formation of rain droplets in clouds. That is, Atlant is claimed to lead to rainfall enhancement when rain occurs. A comprehensive description of the Atlant system may be found on the Australian Rain Technologies website (http://www.australianrain.com.au). The aim of the six-year trial is to produce statistically reliable data that can be used to assess such an enhancement effect in the Hajar mountains of Oman.

## 3.2 Experimental Design

This statistical evaluation is based on an extensive set of surface level rainfall measurements covering target and control areas that vary from day and are defined by a randomised experimental design with strict operational controls, as opposed to opportunistic cloud measurements. The length of the trial and the randomised operating schedule used for the trial are specified well in advance of the start of the trial, as are rules governing adjustments for outages due to equipment issues or enforced stoppages due to weather conditions where there is risk of flooding. The statistical methodology used to analyse the rainfall data collected in the trial is set prior to the commencement of the trial, and is essentially the same methodology (extended to reflect the increased scope of the 2016 trial) that has been used in analysis of the trial data since 2013.

Data for the trial were collected from a combination of instruments placed and maintained by TIE and by the department of the Director General of Meteorology and Aviation (DGMAN). TIE maintains most of the rainfall gauge network and they operate and maintain the Atlant systems. DGMAN provide daily radiosonde data from Muscat airport on upper level wind directions, speeds and derived weather indices as well as data from weather stations located in and near the trial area. They also provided supplementary weather modelling and wind profile data on request.

Statistical modelling and analysis of rainfall data is notoriously difficult because of the huge variability associated with where and when rainfall occurs. However, such a trial is possible for a ground based system with a well-defined downwind target area, provided there are sufficient rainfall gauges spread over this area and adequate care is taken when analysing the data obtained from these gauges to account for the spatial and temporal variability in rainfall. Identifying whether variation in the operation of the Atlant systems leads to statistically significant differences in rainfall is the basic objective of this sequence of rainfall enhancement trials.

### 3.3 Changes in the 2016 Trial

As in previous years, the number of Atlant sites has been increased and the rainfall gauge network extended. Two new Atlant sites designated H7 and H8 were added. However, these sites only became fully operational mid-trial due to a combination of an anticipated delay in installation and unanticipated structural faults that needed repair following installation. Twenty gauges were added to improve the network coverage near these new Atlant sites. The instrumentation for the 2016 trial is shown in Figure 3-1 and provided in tabular format in Appendix A. New Atlant installations H7 and H8 were re-designed to enable flat pack delivery and facilitate ease of construction. H7 is depicted in Figure 3-2.

In addition, a new design of the Atlant system has been introduced at the two new Atlant sites to better cope with severe weather conditions in the Hajar ranges and to improve the efficiency of ion generation. As the physical process of generating free ions is well understood, the impact of these changes is not considered in the statistical analysis.

The most significant issues with respect to instrumentation was the change by DGMAN from a hand released radiosonde at Muscat International Airport to an automated system for this release. There were a high percentage of failures of the automated system in June and the first half of July. Data collected by this radiosonde are used to estimate steering wind directions, which in turn define target and control areas. The implications and the approach adopted to deal with these missing data are discussed in detail later in this report.

Two DGMAN stations in the trial area in 2016 also had a very high proportion of missing observations and so were not included in the 2016 analysis. One of these was the DGMAN station at Jabal Shams, which is the station with the highest elevation in the Hajar ranges and is in the middle of the trial area.

**Figure 3-1 The instrumentation for the 2016 trial in Oman.**

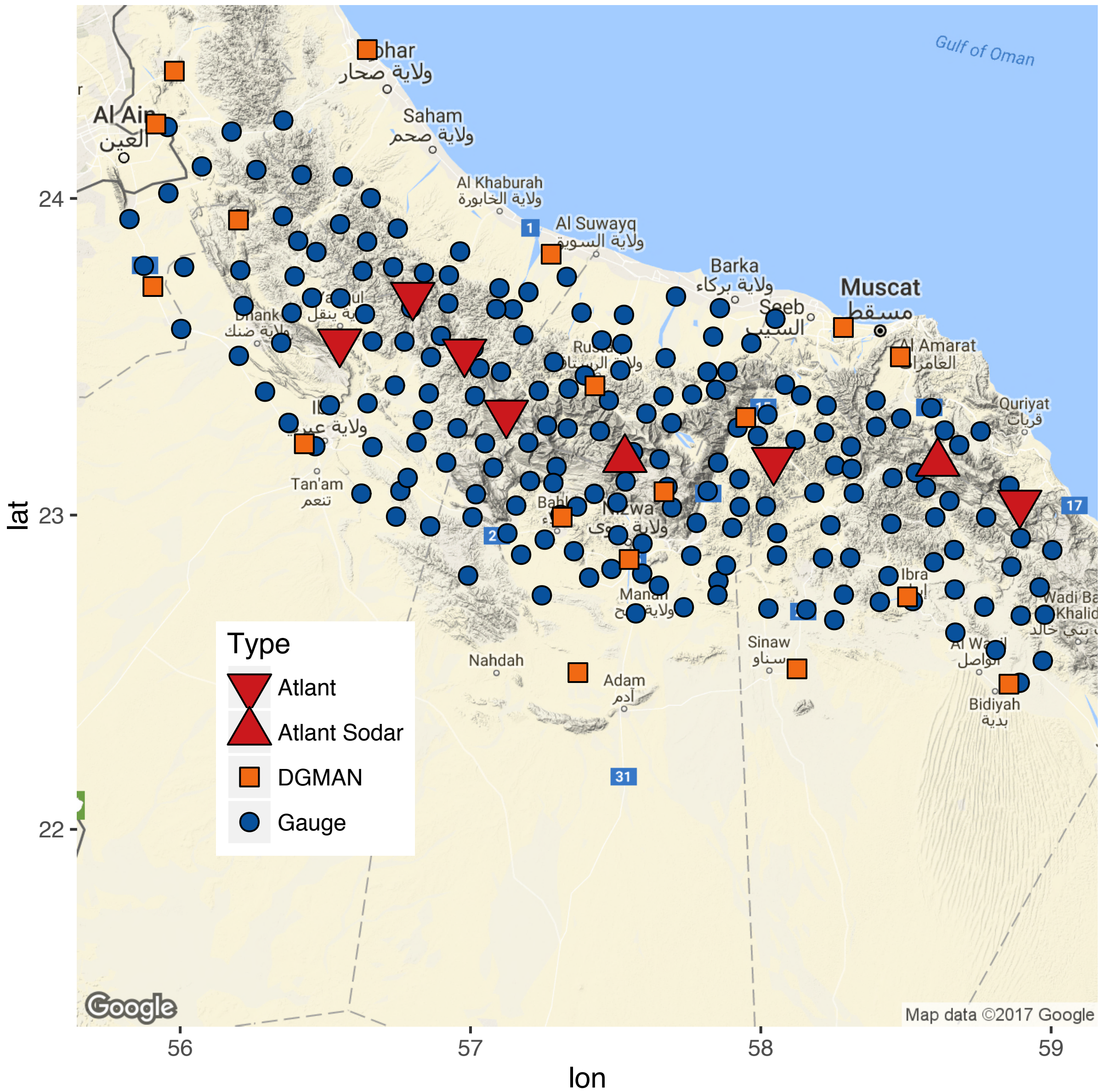


**Figure 3-2 New Atlant Installation designated H7**

# 4 Meteorological Conditions during the Trial

## 4.1 Rainfall

As the Atlant system is hypothesised to increase rather than create precipitation, general rainfall in the trial region is the most critical meteorological factor affecting a trial. Average daily gauge rainfall and the daily proportion of gauges recording a rainfall event (that is, with a rainfall reading of at least 0.2 mm) is shown in Table 4-1. The latter will be referred to as the propensity to rain or daily rainfall propensity.

Table 4-1 includes separate averages for all the trial gauges and for those that were in operation in all four years. It also includes all recorded rainfall readings as well as restricting readings to those that were greater than 0.2 mm. A reading of 0.2 mm is the minimum or trace gauge reading and gauge increment. It would take a percentage enhancement effect of over 100 percent to record a measurable increase in such a trace reading. Consequently, trace readings do not contribute significantly to the detection of an enhancement signal. It can be seen in the table that trace rainfall does not have much of an impact on average rainfall but does reduce the proportion of gauges recording rainfall on a given day and therefore the useful sample size.

The results for all operational gauges and those that were operational over the four years are similar. Average daily gauge rainfall in 2016 was the lowest observed in the trial to date, roughly half the previous minimum, which occurred in 2013 and less than one third of the maximum which occurred in 2015. The propensity to rain in 2016 was also the lowest of the four trials. The distribution of rainfall greater than 0.2 mm is shown as a box plot for each year in Figure 4-1.

The light blue box shown in Figure 4-1 extends from the 25 to the 75$^{th}$ percentile and the dark blue line is the median. The vertical blue line extends roughly to the 95$^{th}$ percentile. The red circles are outliers which would correspond to the heavy rainfall events in each year. Looking at the distribution of positive rainfall events gives a better picture of how rainfall affects the trial in each year. Average total rainfall in 2014 was greater than in 2013 but when it rained the falls were considerably smaller. The 2016 distribution is, apart from one extreme gauge level rainfall event, like 2014 though more compressed towards lower rainfall values.

The geographic distribution of rainfall and the propensity to rain is shown for 2016 in Figure 4-2 and for the three-year average from 2013 to 2015 in Figure 4-3. The colour scale of the circles corresponds to average daily rainfall and the size of the circle to the propensity to rain. The key differences are in the middle and leeward side of the centre of the trial area where the Hajar ranges are most elevated.

**Table 4-1 Average daily gauge rainfall and the average daily proportion of gauges recording rainfall and excluding trace rainfall over the four Atlant trials in Oman.**

| Trial Year | All Gauges | | | | Common Gauges | | | |
|---|---|---|---|---|---|---|---|---|
| | Average Rainfall (mm) | | Propensity | | Average Rainfall (mm) | | Propensity | |
| | All | No Trace | All | No Trace | All | No Trace | All | No Trace |
| **2013** | 0.41 | 6.26 | 0.08 | 0.07 | 0.45 | 6.30 | 0.09 | 0.07 |
| **2014** | 0.40 | 5.52 | 0.09 | 0.07 | 0.37 | 5.64 | 0.09 | 0.07 |
| **2015** | 0.65 | 7.10 | 0.11 | 0.09 | 0.76 | 7.53 | 0.12 | 0.10 |
| **2016** | 0.25 | 5.48 | 0.06 | 0.05 | 0.23 | 5.24 | 0.06 | 0.04 |

**Figure 4-1 The distribution of rainfall greater than 0.2mm over the four Atlant trials in Oman**

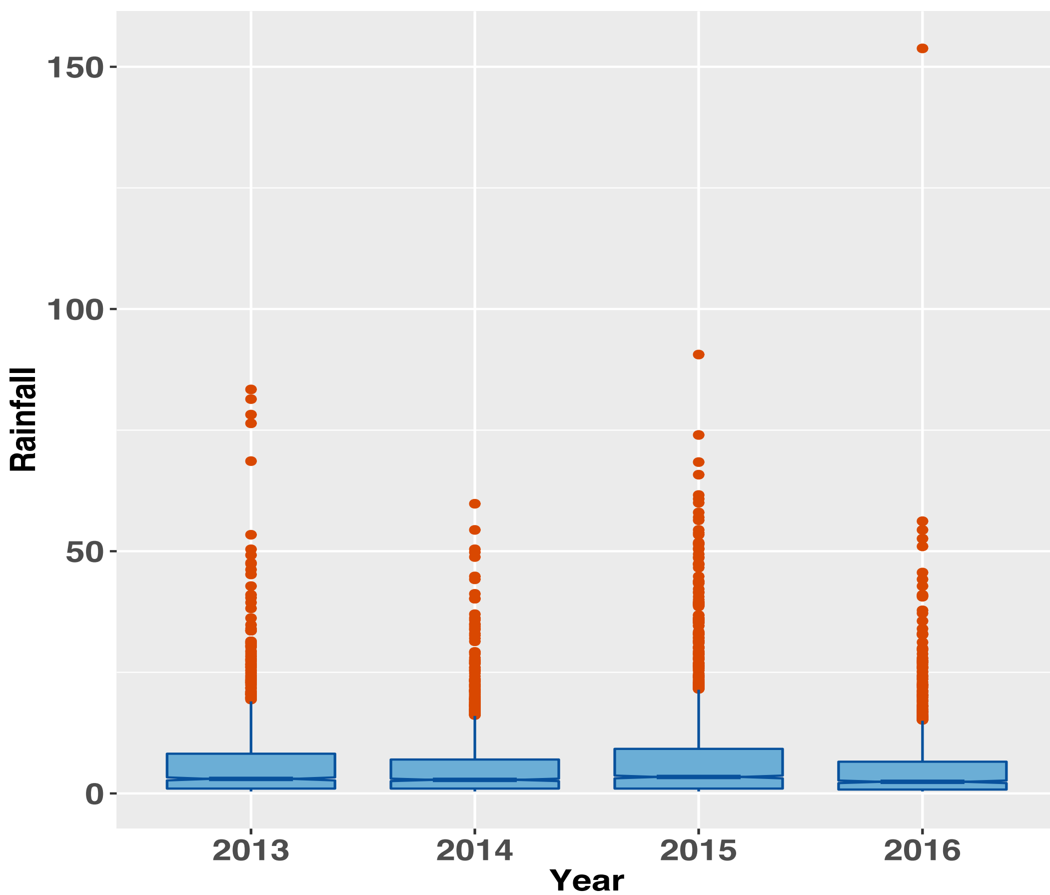


**Figure 4-2 The geographic distribution of rainfall over the trial area in 2016.**

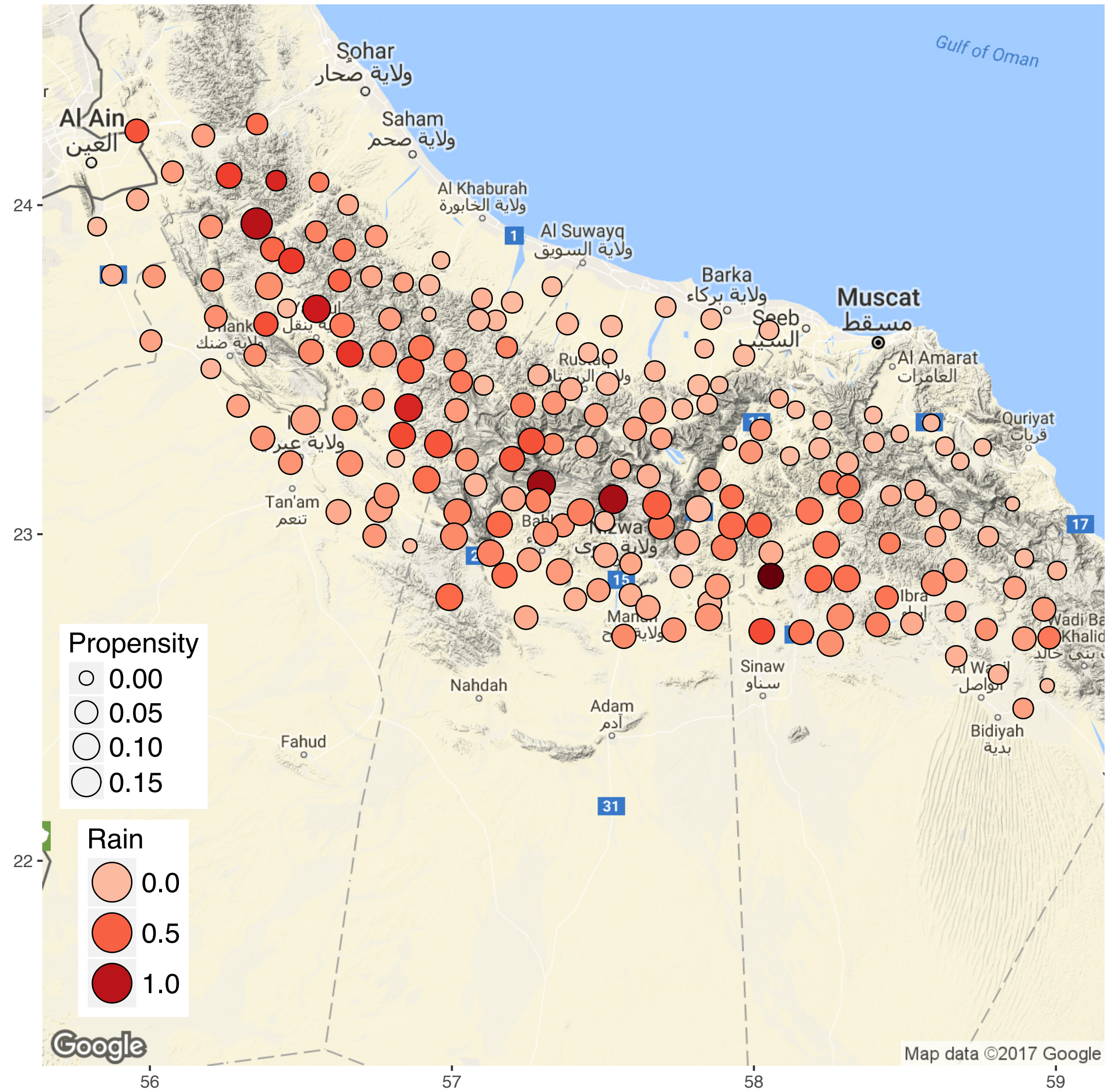
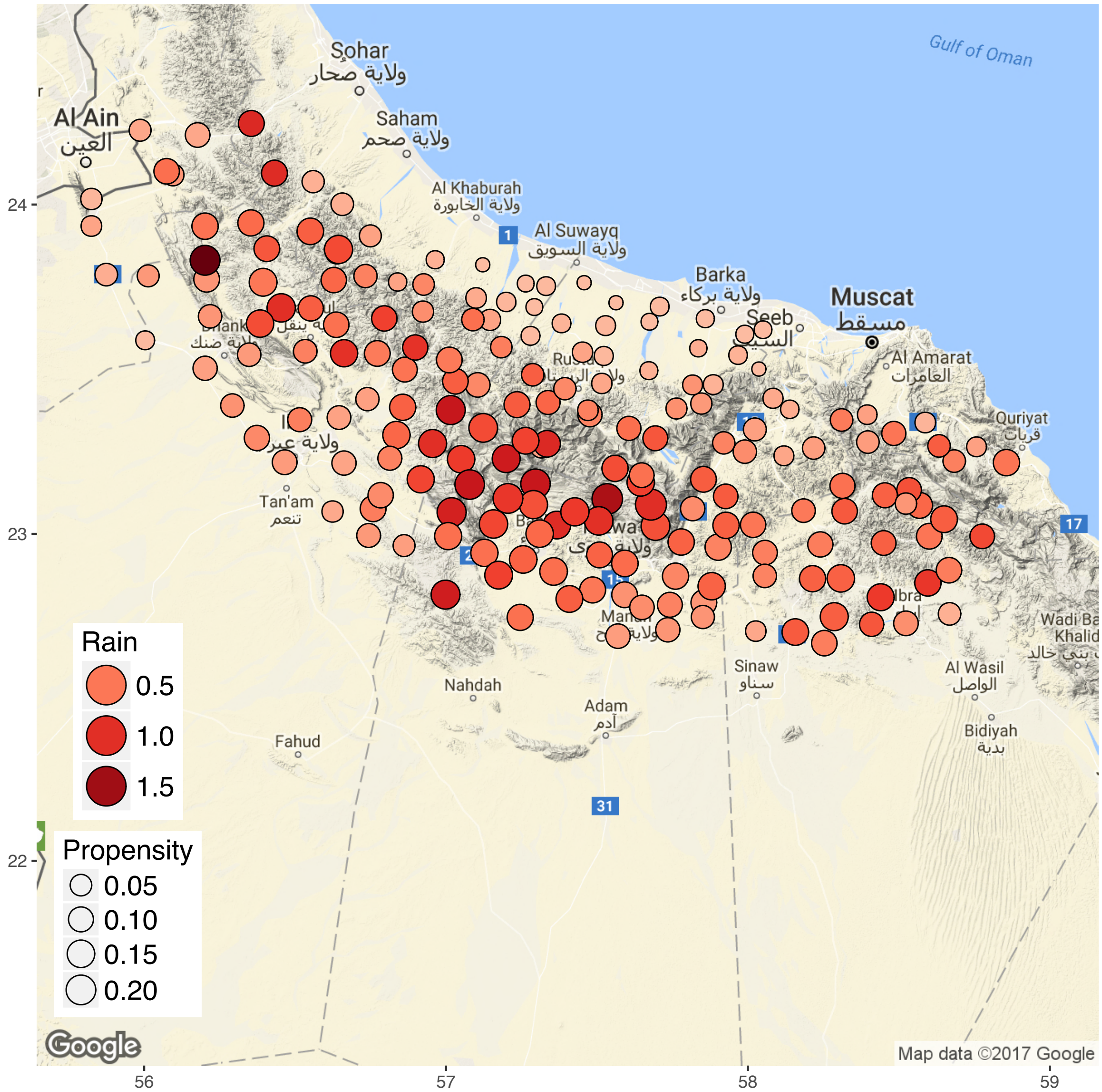


**Figure 4-3 The geographic distribution of rainfall over the trial area, three-year average based on data for 2013 to 2015.**



### 4.2 Wind Direction

From the perspective of analysis of the trial data, wind direction is the most important meteorological factor. Steering wind directions obtained from the 4am radiosonde released from Muscat International Airport are used to define the downwind target and control areas. Steering winds direct the flow of cloud moisture in the free atmosphere, calculated here as the speed weighted average of wind directions recorded by radiosonde sensors between 500 and 700 hPa pressure levels.

The distributions of steering wind directions for the four trial years are shown in Figure 4-4. The bars indicate the proportion of the time this steering wind was from the eight points of a compass rose. The pattern is consistent across all the trials. Steering winds are predominantly from the northeast and with about 70 per cent in the 90-degree range north to east.

**Figure 4-4 Distribution of steering wind directions during the four Atlant trials in Oman.**

### 4.3 Conditions Favouring Precipitation and Thunder Storms

The analysis of the trial data focuses on the prediction of rainfall. These predictions are used to try and control for the natural day to day variability in rainfall when comparing rainfall recorded by gauges located in the downwind areas, i.e. those areas defined by a 30 km wide corridor centred on and extending downwind from an Atlant, with Atlant status on the day determining whether a downwind gauge is a target (Atlant on) or a control (Atlant off). Determination of target or control status for a gauge is the basis for estimating how much rain would have been recorded at a downwind target in the counterfactual case that the Atlant system in question was off or at a downwind control for the counterfactual that the Atlant system in question was on. While the predictions make use of spatial characteristics such as the elevation of the gauges, the data collected by the 4am radiosondes and by the DGMAN weather stations are the two sources of meteorological information used.

There are several index measures derived from radiosonde measurements that indicate the amount of water in the atmosphere, and its stability, with the former being the water available to fall as rain and latter being the potential for storms, and hence rainfall. Four indices are used in the analysis:

- LCL Pressure
- Lift Index
- Precipitable Water
- Total Totals Index

Two of the indices are presented here. The Lift Index is shown **Figure 4-5** and the Precipitable Water Index is shown in **Figure 4-6**. The index value is on the horizontal axis and days for each of the trial years is shown on the vertical axis. The size of the circles indicates the amount average daily rainfall across the gauge network. Most of the larger rainfall events occur when the Lift Index is low and the Precipitable Water Index is high. Over the four trials the correlation between the daily values of these two indices is negative 66.8 per cent. This relationship is consistent over the four trials**.**

The DGMAN weather stations are spread over the trial area and represent a source of information on prevailing weather conditions that could potentially help differentiate between weather conditions at different downwind target and control gauges. This is a potential area for investigation. In previous trials these data were aggregated across DGMAN stations each day and this approach has been retained. The readings across stations, which are averaged between 10am and 8pm each day, are strongly correlated and so the different meteorological measurements are incorporated into the predictive models as principal components. The first component, which is essentially the common variation across all stations on a day, is shown for dew point temperature in Figure 4-7. There is clearly greater rainfall the higher the dew point temperature. The relationship is again consistent over the last three trials. The trial area in the first year was smaller, with coverage for only two Atlant sites, and there were substantially fewer number of DGMAN stations contributing. This may have contributed to the limited range of readings in 2013.

The natural level of variability of summer rainfall in the Hajar ranges is extremely high as the storm systems tend to quite turbulent. Meteorological data offers some degree of control over this natural variability in daily rainfall across the trial region when predictive models for rainfall are of interest. How well these models perform is taken up later in this report.

**Figure 4-5 The 4am 'Lift Index' and average daily gauge rainfall over the four Atlant trials in Oman.**

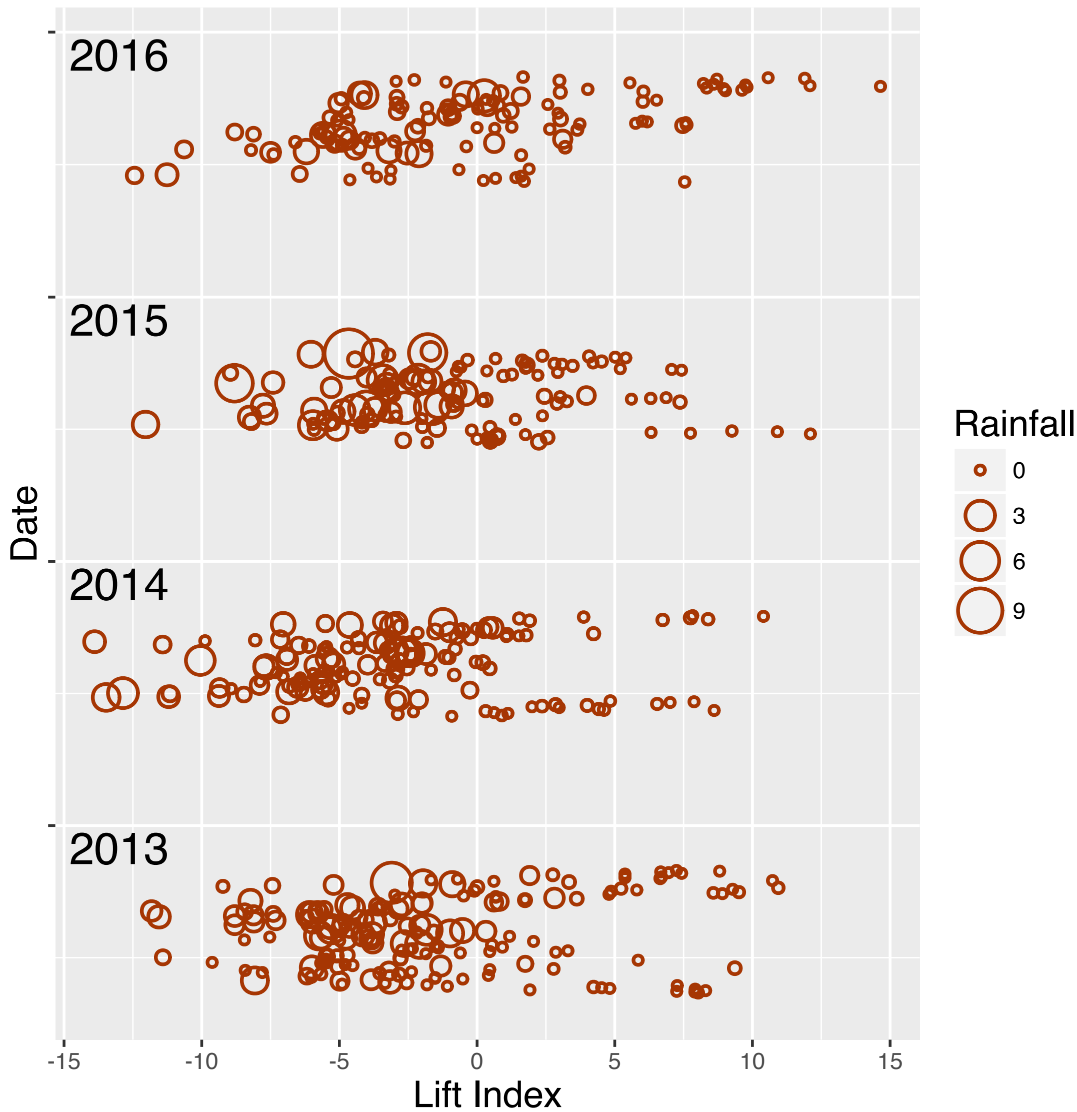


**Figure 4-6 The 4am 'Precipitable Water Index' and average daily gauge rainfall over the four Atlant trials in Oman.**

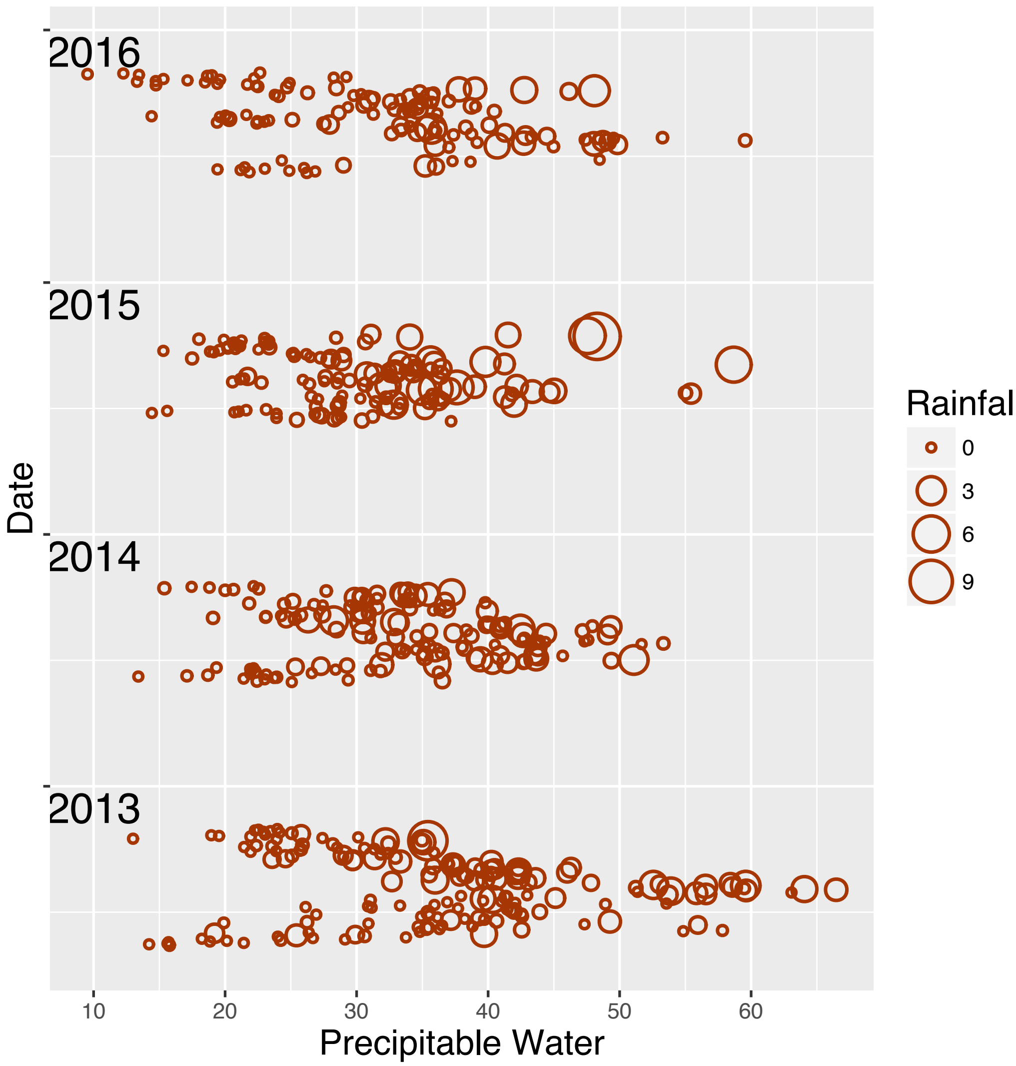
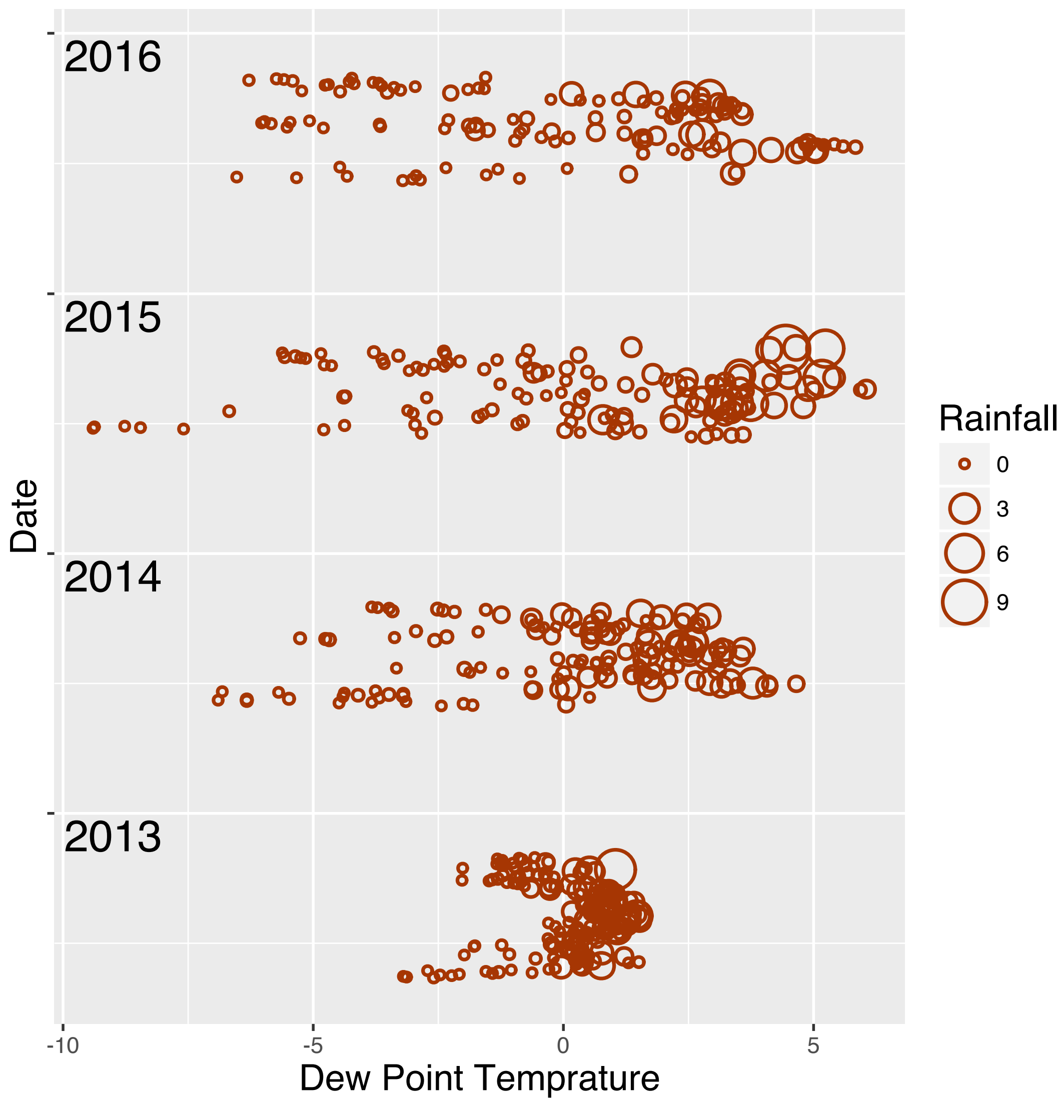


**Figure 4-7 The first principal component of DGMAN weather station readings of dew point temperature, averaged from 10am to 8pm daily over the four Atlant trials in Oman. Note that these principal component values are standardised to have a mean of zero.**



# 5 Unanticipated Issues Arising in 2016

## 5.1 Unanticipated Issues

There were two unanticipated issues in the 2016 trial:

- A delay in the planned start-up of the H7 and H8 Atlant systems due to technical faults; and
- A substantial number of days with missing radiosonde observations due to the change from hand held to automated release of the balloons.

The new systems at H7 and H8 were scheduled to come online on the 1st of July. The Atlant system at H8 came online on the 29th of July and the system at H7 came online on the 6th of August. Aside from the reduction in operating scope over essentially the first half of the 2016 trial, this slight delay lead to some imbalance in the experimental design, which called for an equal number of systems to be on and off each day. As this operating schedule was set prior to the start of the trial no adjustments were made, with the missing data treated as missing at random.

During the time H7 and H8 were offline the other six sites occasionally became imbalanced. On days without any recorded rainfall this was not a concern. On days with recorded rainfall there four systems were operating three days and two systems were operating two days. During the time H8 was online and H7 remained offline the design was again imbalanced. On days when rain was recorded four systems were operating on four days and three systems were operating on three days.

A much more important issue was the number of days when no radiosonde data were available. A missing radiosonde observation precludes the analysis of the rainfall data on that day using the established methodology, since steering wind direction data are then not available and so downwind areas cannot be determined. This left two choices if the basic structure of the methodology was to be retained in 2016:

- Analyse the data for the smaller number of days (primarily August - October) when radiosonde data were available; or
- Impute the missing radiosonde data to retain a larger sample.

A considerable amount of effort went into resolving this choice but the central issue was the trade-off between the reliability of the imputation and the impact of a reduced sample size. Several sources of information for the imputation were explored, including:

- Radiosondes from other sites in the region, notably Abu Dhabi;
- DGMAN wind profiler data at Muscat International Airport; and
- DGMAN modelled radiosonde data for Muscat International Airport.

The first two approaches did not predict the 4am steering wind directions adequately for those days when radiosonde data were available, while the reliability of the third approach could not be tested given that the model used by DGMAN is calibrated to the available radiosonde data. Consequently, a trend smoothing approach was utilised. This approach depends on having adequate support between missing observations, that is the missing data gaps cannot be too long. There was a missing sequence of 8 days at the start of June, a missing sequence of four days from the 20th to the 23rd of June and a missing sequence of 17 days from the 28th of June to the 14th of July. Radiosonde data were not imputed over these periods.

This left 27 days with missing radiosonde data, corresponding to gaps that were three days or less. The missing data for these days were imputed using a two-stage smoothing method based on the 'twicing' approach suggested by Tukey (1977) which exploits the nearby in time correlation of the data. The basic idea is to first calculate the local trend in the data using the observed radiosonde data and then use this trend as a first stage imputed value at the missing points in time. Second, the difference in the actual and trend values, again based on the observed data, is used to correct this first stage imputation. This is done using a second smooth of the deviations of the actual values from the first stage trend which is then used to impute for the first stage imputation error at the missing points. Both the first stage trend, and the second stage residual smooth, are computed using a non-linear local regression method called LOESS, and involve judicious choice of how much smoothing of the data is carried out, corresponding to more smoothing for the trend and much less smoothing for the residual.

The actual and imputed steering wind directions and speeds obtained after this process are shown for the eight quadrants of the compass rose in Figure 5-1. The figure also includes data from the previous trials. Observations derived from the 4am radiosonde are denoted by ‘Act’ and those corresponding to imputed values are denoted ‘Imp’ The height of the bars indicates the proportion of time the steering wind was from each quadrant while the colour of the bars shows the average steering wind speed. The actual and imputed wind speeds for 2016 show the same patterns observed in previous trials. All the imputed wind directions are from the northwest to the east.

The need to make use of these imputed wind directions is illustrated in Figure 5-2. This shows the distribution of gauge rainfall greater than 0.2 mm for the days with actual and with imputed wind directions. The vertical axis shows the frequency of gauge observations, with rainfall values on the horizontal axis. The shoulder of the imputed distribution is steeper and the tail is fatter, indicating that there is a higher proportion of larger rainfall readings on days when the steering wind was imputed.

The geographic distribution of average daily rainfall on days when actual steering wind information is available is shown in Figure 5-3 while Figure 5-4 shows the same distribution on days when only imputed steering wind information is available.

**Figure 5-1 Actual and imputed steering wind directions and speeds over the 2016 and previous trials.**

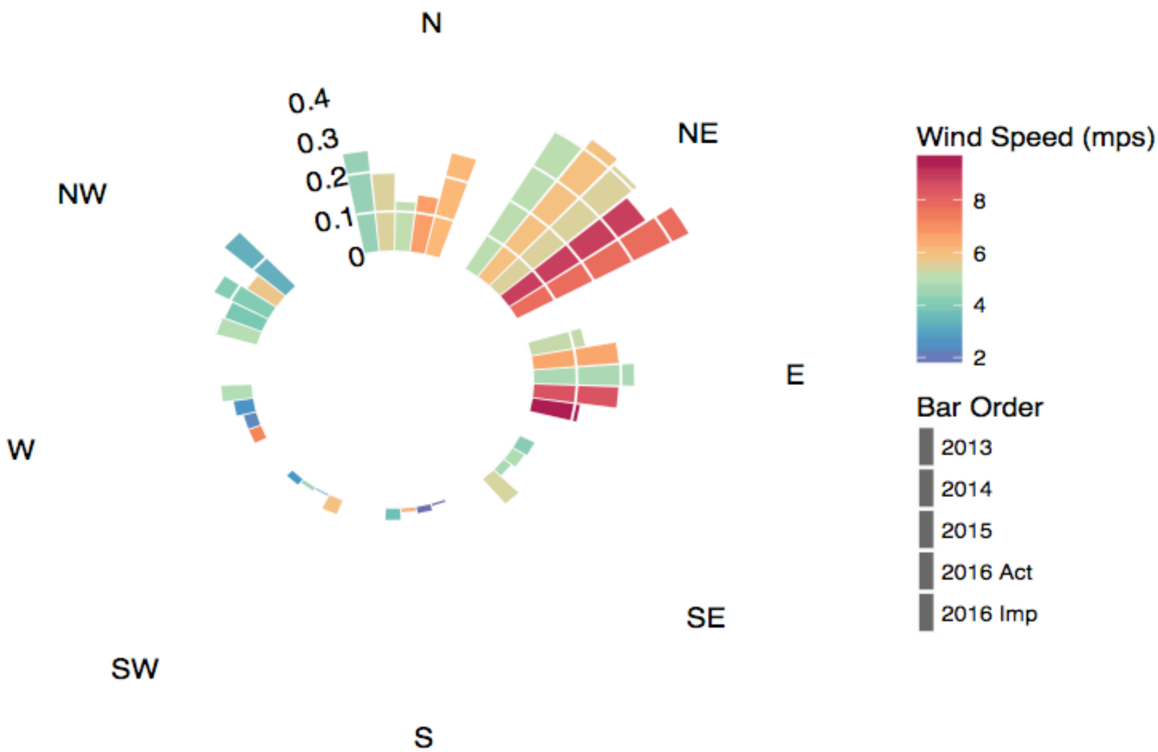


**Figure 5-2 The distribution of gauge rainfall greater than 0.2 mm on days with actual radiosonde data and on days with imputed steering winds in 2016.**

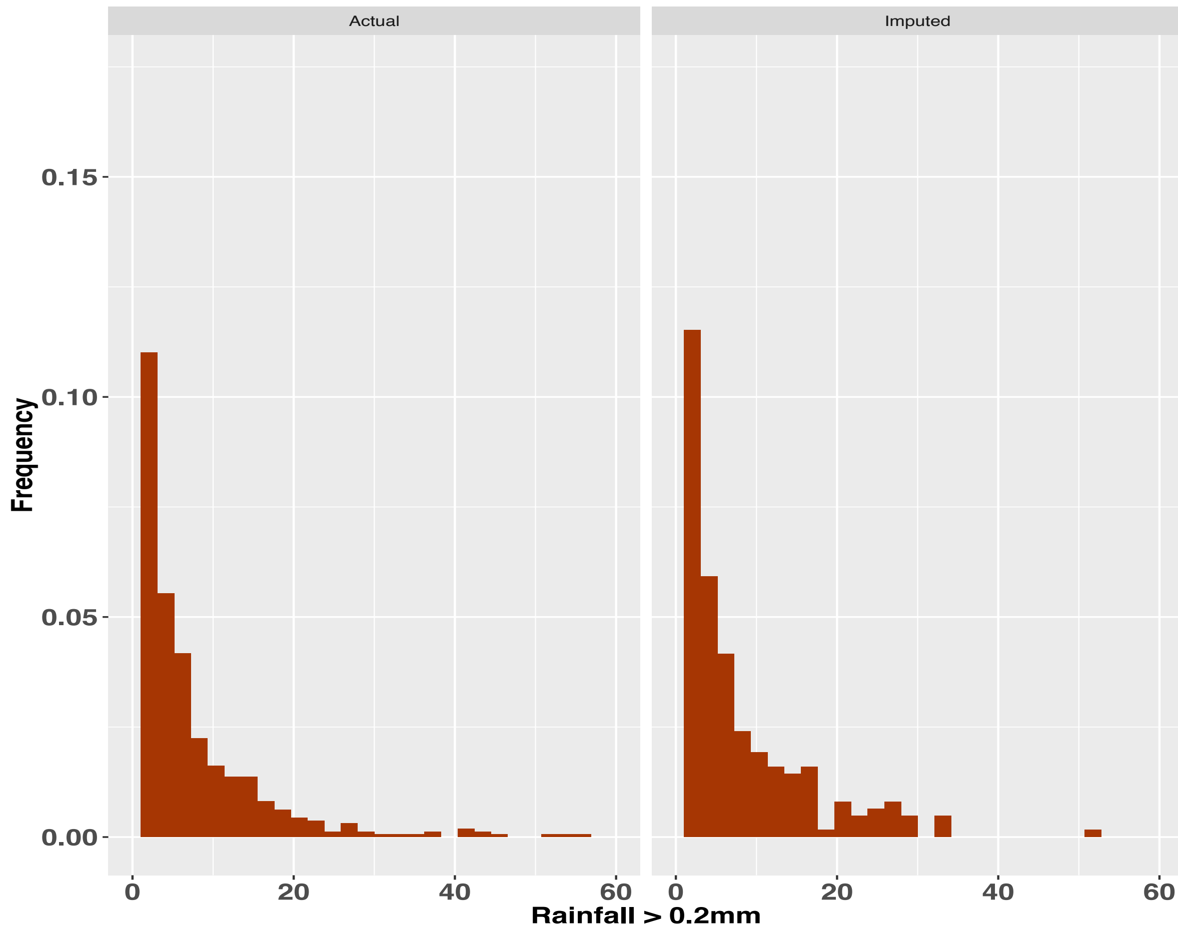


**Figure 5-3 The geographic distribution of average daily rainfall on days with actual steering winds in 2016.**

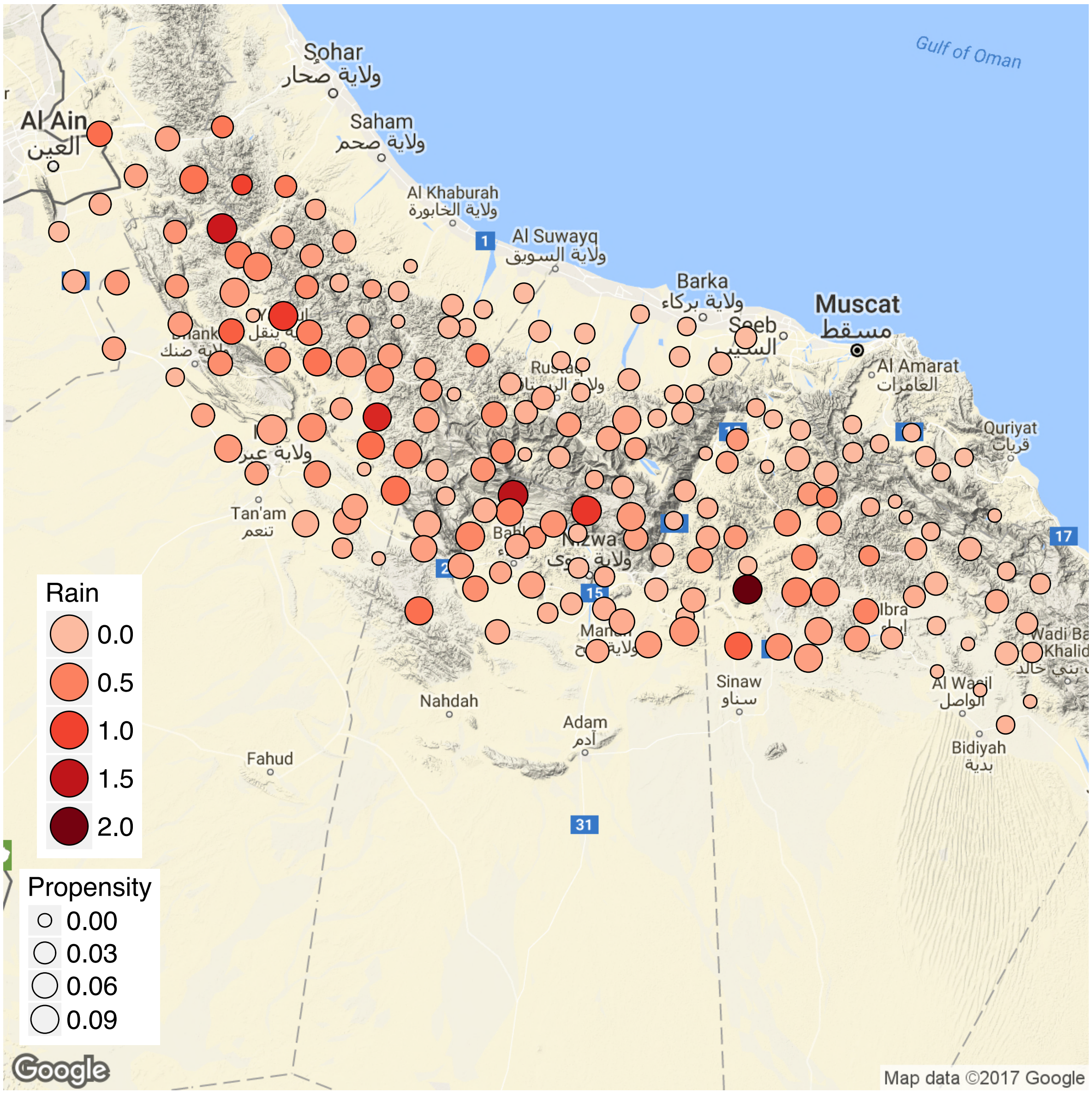
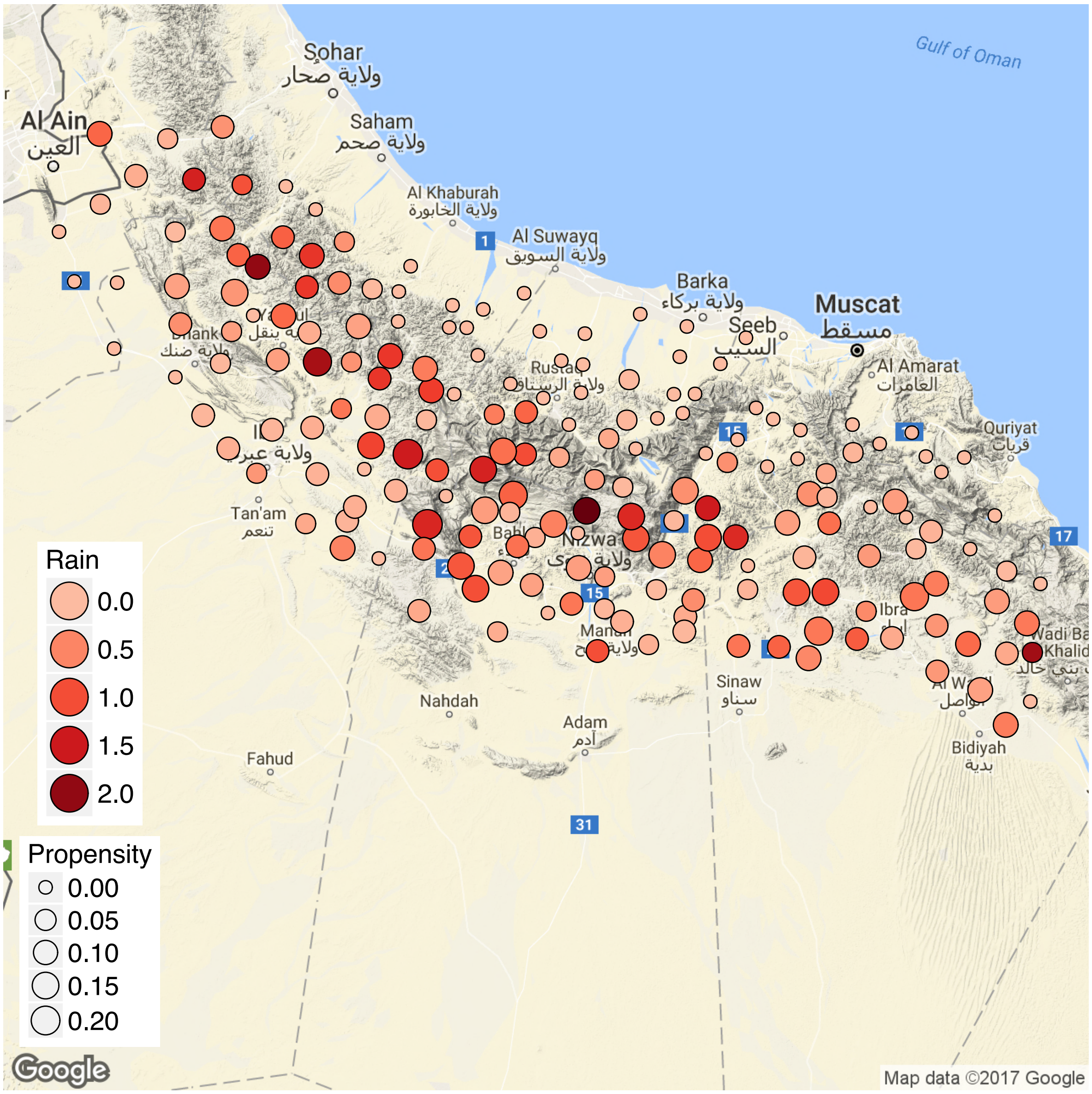


**Figure 5-4 The geographic distribution of average daily rainfall on days with imputed steering winds in 2016.**



Average rainfall was similar in the southwest of the trial area on days with actual as opposed to imputed steering wind directions. Average rainfall was higher in the northwest of the trial area on days with actual steering wind directions as opposed to those with imputed wind directions, but considerably lower than average rainfall for imputed days in the more elevated central trial area.

The propensity scales shown on Figure 5-3 and Figure 5-4 are quite different, making their comparison difficult. However, on days with actual steering wind directions the propensity for rain ranged up to nine per cent. This more than doubled to 20 per cent on imputed days. That is, the number of useful gauge level rainfall measurements on imputed days is higher than on days when actual steering wind directions were available.

The use of imputed wind directions adds to the uncertainty associated with steering wind directions and consequent designation of target and control gauges. However, it also adds substantially to the number of positive gauge readings available for analysis and so reduces the impact of the variability of rainfall observations in a year with substantially fewer rainfall events and less rainfall when rain did occur compared with previous trials.

An experiment was run to see how a similar imputation exercise would have impacted on the 2015 results. The experiment is described in detail in Appendix C but the general idea was to match days with missing radiosonde information (and hence with imputed wind direction in 2016) with similar days in 2015. Since matching by calendar day is likely to generate very different rainfall patterns in 2015 compared to what was observed on the days with missing observations in 2016, it was decided to match on actual rainfall. Ideally, one would seek to match on average daily rainfall. However, this was not possible because average rainfall and rainfall propensity were both much higher in 2015. Consequently, matching was based on rank order. Two rank orderings were matched:

- Rank order of average daily rainfall, as for example, if radiosonde data was missing on the 5th highest rainfall day in 2016, then wind direction was imputed for 5th highest rainfall day in 2015; and
- Rank order of daily rainfall propensity, as for example, if the 10th highest proportion of gauges reporting rainfall on a day when wind direction information was missing in 2016, then wind direction was imputed for 10th highest rainfall propensity day in 2015.

The results of this experiment are summarised in Table 5-1, and show that imputation for missing days based on rainfall rank has little impact on the 2015 trial results. Imputing for missing days in 2015 based on rainfall propensity rank does affect the results but not to the extent that it would change the conclusion that there was a highly significant and substantial attribution in 2015. While this reassurance cannot be taken for granted in 2016, it at least indicates that imputation of the missing radiosonde data in 2016 is a reasonable strategy.

## 5.2 The Sample Balance in 2016

The balanced randomised design used to control operation of the Atlant mechanisms over the course of a trial ensures the same number of operational and non-operational days for each Atlant, but cannot ensure a similar level of control or balance for target and control rainfall measurements because of day to day variability in wind direction. In order to assess the actual (i.e. achieved) balance in 2016, we examine the distribution of two variables that are not controlled by the trial design:

- The total number of daily downwind gauge observations (Gauge Days) when the site is active and when it is inactive for each site; and
- The trial wide rainfall propensity when a site is active compared with when it is inactive.

This achieved balance is summarised in Table 5-2. The H7 and H8 sites are included only when both systems were operational. On the first of these criteria, the balance of the sample with imputed steering wind directions does not appear to be a cause for concern. There is one exception. The trial area rainfall propensity when H6 was inactive was almost twice the same propensity when H6 was active.

**Table 5-1 The results of imputing for days with missing wind direction information in 2015 after imposing a pattern of missing radiosonde observations similar to that observed in the 2016 trial.**

| Value/Estimate | Bootstrap Distribution of Estimated 2015 Atlant Attribution (%) | | |
|---|---|---|---|
| | Actual 2015 Data | Missing Days Based on Rainfall Rank | Missing Days Based on Propensity Rank |
| **10th percentile** | 14.3 | 14.8 | 11.1 |
| **50th percentile** | 19.8 | 20.0 | 16.6 |
| **90th percentile** | 26.1 | 26.1 | 23.0 |
| **Mean** | 19.9 | 20.1 | 16.6 |
| **Standard Error** | 4.6 | 4.4 | 4.7 |

**Table 5-2 Summary of the sample design balance for 2016 based on downwind observations and rainfall propensity.**

| Atlant | Downwind Gauge Days | | Rainfall Propensity (%) | | Rainfall Propensity > 0.2mm (%) | |
|---|---|---|---|---|---|---|
| | Off | On | Off | On | Off | On |
| **H1** | 789 | 742 | 10 | 9 | 7 | 7 |
| **H2** | 849 | 874 | 11 | 9 | 9 | 6 |
| **H3** | 706 | 759 | 6 | 9 | 5 | 7 |
| **H4** | 831 | 828 | 7 | 7 | 5 | 6 |
| **H5** | 736 | 726 | 6 | 7 | 5 | 6 |
| **H6** | 413 | 417 | 13 | 7 | 11 | 6 |
| **H7** | 403 | 377 | 4 | 5 | 3 | 5 |
| **H8** | 597 | 565 | 10 | 10 | 8 | 7 |

# 6 Analysis of the 2016 Trial

## 6.1 Methodology

As in the previous three trials a multi-stage methodology was used for the gauge level analysis of the trial data. These stages are set out below:

1. Steering wind directions and speeds were calculated as speed weighted averages of 500 - 700 hPa wind directions and speeds recorded by the 4am radiosonde at Muscat International Airport. These steering wind directions were then used to define the 30 km wide and 75 km long upwind and downwind corridors centred at each Atlant site.
2. For each day of the trial, gauges located inside these corridors were then designated as;
    - A target when downwind of an active Atlant;
    - A control when not a target and downwind of an inactive Atlant site; and
    - Upwind and therefore eligible to be used to predict expected rainfall for the downwind target and control gauges.
3. A regression model with a random day effect was fitted to upwind gauges with recorded rain and used to estimate the conditional expectation of the logarithm of the amount of upwind rainfall given that a rainfall event occurred.
4. This upwind model was then used to estimate the expected value of the logarithm of the amount of rain for each target and control gauge in the downwind areas given that a rainfall event occurred at that gauge.
5. Another regression model with a random day effect was fitted to all downwind gauges recording rain. Like the upwind model, this also is a model for the logarithm of rainfall (i.e. it excludes zero values). However, it is expressed in terms of gauge elevation and the expected logarithm of rainfall derived in the previous step, together with indicators for the target status of the gauge rainfall reading on the day, allowing estimation of the expected level of rainfall at a gauge serving alternatively as a target as opposed to a control.
6. This downwind model was used to calculate gauge level attribution when rainfall was observed as the difference in rainfall under two counterfactual conditions:
    - The gauge was a target when in fact it was a control;
    - The gauge was a control when in fact it was a target.
7. The gauge level attribution was summed over all downwind gauges (i.e. both target and control) that recorded rain to obtain the overall attribution, which is then expressed as percentage of either:
    - Total observed rainfall for downwind gauges; or
    - Total predicted rainfall for downwind gauges that report rain, assuming that they are all controls. This is referred to as natural rainfall.
8. A bootstrap procedure that accounts for both spatial and temporal correlation in the model errors was then used to simulate an empirical distribution of estimated attribution values. Confidence intervals for the actual attribution were derived directly from this empirical distribution.

The rationale for the definitions of the downwind and upwind corridors, the methods used to specify both the upwind and downwind models and to compute the attribution have been documented in full in previous reports. We note that as in previous trials, rainfall modelling is carried out on the log scale, allowing the distribution of positive rainfall observations to be modelled on a linear scale. This is a standard approach when dealing with highly skewed data.

An additional test was introduced in 2015, and was repeated in 2016. This assesses whether there is a significant link between the observed attribution and the actual Atlant operating sequence. This is a robust non-parametric test used in large scale clinical trials for assessing whether there is significant evidence that application of a procedure (e.g. taking a drug) actually leads to the claimed benefit, rather than this benefit being due to a placebo effect. In the Atlant context, it is carried out by comparing the attribution obtained using the actual operating schedule for the trial with the attributions obtained from a large sample of random permutations of this schedule, all based on the same observed trial data.

Application of this test overcomes some if not most of the problem of unbalanced sample data due the random spatial and temporal variation in rainfall. If the rainfall variation between target and control gauges is largely a reflection of this variation, then any randomly chosen operating schedule that leads to a more favourable allocation of target gauges should generate a higher attribution. At the same time, the attribution generated by the actual operating schedule, given a large sample of randomised schedules, will tend to fall amongst the more ‘average’ outcomes.

On the other hand, if there is a significant attribution attributable to the operation of the Atlant systems, then the attribution generated by the actual operating schedule should be one of the highest ranking outcomes in this permutation distribution. In 2015 the attribution generated by the actual operating schedule was at the 99.5th percentile of the permutation distribution, indicating that there was less the 0.5 per cent chance of a randomly generated schedule leading to a higher attribution. For a single trial, this is very strong evidence of a significant positive rainfall enhancement effect.

### The Upwind Rainfall Model in 2016

Following the same approach as in previous trials, the structure of the upwind rainfall model for the logarithm of gauge level rainfall, denoted LogRain in what follows, was defined by selecting from the set of available meteorological covariates to give a balance between goodness of fit and the risk of overfitting. The latter is the result of including too many covariates in the model, and can lead to very unstable estimates of expected rainfall in the downwind areas.

An automated model selection process, known as stepwise regression, was used to choose the model covariates. Forward stepwise regression was carried out by maximising three different information criteria (BIC, AICc and Combined P-Value Threshold) which led to three candidate models. The day to day variation in upwind rainfall across all rainfall gauges was then swept out by adding a random day effect to each of these models. The final model selected combined significant effects from all three fits in order to define a compromise model with the highest proportion of coefficients with reliable (significant) estimates. These estimates are shown in Table 6-1. The abbreviation PC denotes a principal component.

The overall fit of the model was similar to that of upwind models in previous years, with just over 45% of the variation in LogRain explained by the covariates in the model plus the predicted values of the random day effects. Comparing this model specification with the corresponding 2015 upwind model, we see that the choice of covariates is different, reflecting the increase in the upwind areas due to the addition of H7 and H8 in 2016 and year on year changes in the relative importance of different factors known to be associated with rainfall. Notably, there were just 244 upwind observations of positive rainfall in 2016, compared with 542 in 2015, so the model in Table 6-1 characterises much drier conditions for upwind rainfall than were experienced in 2015.

### 6.2 The Downwind Rainfall Model in 2016

The downwind model for LogRain is displayed in Table 6-2 and includes two sets of covariates. The first set consists of gauge elevation and estimated downwind rainfall derived from the upwind rainfall model displayed in Table 6-1(the instrumental variable in the model). These covariates are included in the model in order to control for the natural variation in rainfall between days and between gauges. The second set is made up of eight target status indicators, one for each Atlant, showing which operating Atlants were upwind of the gauge on the day. These are used to calculate the attribution.

The fit of the downwind model specified in Table 6-2 is not as good as that of the upwind model but this is anticipated because downwind rainfall has much greater variability. In particular, the model explains a similar percentage of variability compared with the downwind model used in 2015 (24% in 2016 vs. 27% in 2015). However, there were significantly fewer rainfall events recorded downwind in 2016 compared with 2015 (772 over 8 downwind areas in 2016 compared with 1,118 over 6 downwind areas in 2015), so levels of significance for the target status indicators were much weaker in 2016 compared with 2015. Unlike 2015, gauge elevation was highly significant in 2016, perhaps reflecting higher variability in elevations due to the increase in the number, and spatial dispersion, of gauges deployed in 2016. The instrument (i.e. the estimated value generated by the upwind model) was highly significant, though with increased standard error compared with 2015. This general increase in standard errors flows through to lower significance levels for the coefficients of the target status indicators, where only the target indicator for H5 is significant at the 10 per cent level (it was the most significant indicator in 2015 as well). Overall, six out of the eight target indicators have positive coefficients in 2016, although these are very small for H4, H7 and H8. Interestingly, the H6 target indicator recorded the most negative coefficient in 2015 and in 2016, though in both years this value was not significant.

The estimated 2016 attribution, i.e. the relative increase in the total amount of downwind rain estimated to be due to the presence of the target indicators in the downwind model specified in Table 6-2, was then calculated as the ratio of the total contribution of rainfall from the target indicators in this downwind model to the total rainfall contributions from all other sources in this model, including both the random day effect and the residual effect. This showed that there was an estimated 3,014 mm of natural rainfall in downwind gauges in 2016 out of a total of 3,521 mm of observed rainfall. Attributed downwind rainfall in 2016 therefore totalled 507 mm, or 16.8% as a percentage of estimated natural rainfall.

While the estimated attribution can be calculated directly from the model specified in Table 6-2, its precision cannot be similarly calculated. Consequently, a numerically intensive resampling strategy known as a bootstrap is used for this purpose. This strategy is described in previous reports and was repeated in 2016. It involves creating a large number of independent simulations of the downwind gauge readings, and for each repeating the model fitting process leading to Table 6-2 and the subsequent attribution estimation step. This results in a distribution of estimated attribution values from which confidence intervals and associated significance levels can be extracted.

A key property of this bootstrap is that it allows for rainfall propensity as well as actual rainfall by using a hurdle model to generate potential downwind gauge readings, i.e. zero readings are first simulated (using the downwind model specification but applied to presence / absence of rainfall) and then rainfall values are simulated for gauges where rainfall is predicted to occur (using the downwind model set out in Table 6-2). This strategy controls for unexplained spatio-temporal variability in observed rainfall by using random sampling of average daily residuals and block sampling of within day residuals to recreate day and residual random effects for the downwind rainfall model. This is necessary since downwind gauges are likely to be exposed to similar meteorological and orographic conditions on any given day. As these conditions are only partially controlled for by gauge elevation and the instrumental variable in the downwind model, there is the possibility of within day correlation in observed downwind rainfall values. This correlation effectively reduces the sample size and, if ignored, leads to the true precision of the estimated attribution being overstated. By restricting resampling of residuals to the same day, this correlation structure, should its exist, is retained in the bootstrap values and hence reflected in the calculated precision of the estimated attribution.

A total of 10,000 bootstrap samples were generated via the above procedure. The average attribution over this bootstrap distribution, measured as the increase over estimated natural rainfall, was 17.2 per cent with a standard deviation of 5.1 per cent. This is a highly significant result, with less than 0.5% of the bootstrap distribution reflecting attribution estimates less than 6%. In particular, all bootstrapped attribution estimates were positive. This is reflected in the bootstrap distribution of attribution estimates shown in the first panel of Figure 6-1.

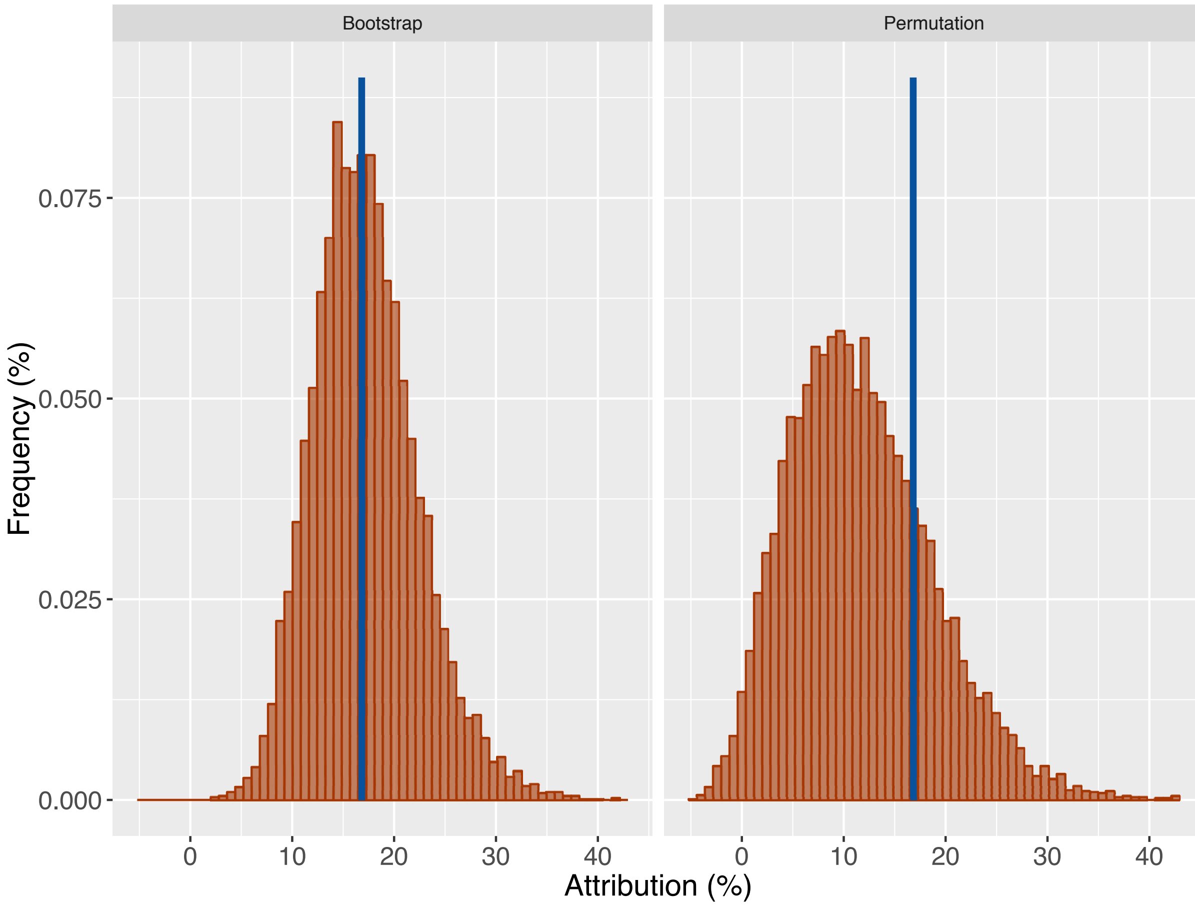
**Table 6-1 The upwind model estimates for the 2016 trial**

| **Summary** | | | | |
|---|---|---|---|---|
| **R Squared** | 0.453 | | | |
| **Observations** | 244 | | | |
| **Parameters** | **Estimate** | **Standard Error** | **T-Value** | **Significance** |
| **Intercept** | -0.695 | 0.196 | -3.555 | 99.9 |
| **Dry Bulb Temp PC2** | 0.642 | 0.159 | 4.047 | 100.0 |
| **Dew Point Temp PC1** | 0.499 | 0.137 | 3.656 | 99.9 |
| **Humidity PC1** | -0.242 | 0.096 | -2.533 | 98.4 |
| **Random Effects** | **Variance** | **Standard Error** | **Percent of Total** | |
| **Day** | 0.413 | 0.190 | 21.9% | |
| **Residual** | 1.468 | 0.144 | 78.1% | |

**Table 6-2 The downwind model estimates for the 2016 trial.**

| **Summary** | | | | |
|---|---|---|---|---|
| **R Squared** | 0.243 | | | |
| **Observations** | 772 | | | |
| **Parameters** | **Estimate** | **Standard Error** | **T-Value** | **Significance** |
| **Intercept** | -1.074 | 0.238 | -4.515 | 100.0 |
| **Gauge Elevation** | 0.000 | 0.000 | 3.019 | 99.7 |
| **Instrument** | 1.345 | 0.319 | 4.213 | 100.0 |
| **H1 Target** | 0.307 | 0.197 | 1.556 | 88.0 |
| **H2 Target** | -0.022 | 0.195 | -0.113 | 9.0 |
| **H3 Target** | 0.284 | 0.195 | 1.458 | 85.5 |
| **H4 Target** | 0.054 | 0.202 | 0.265 | 20.9 |
| **H5 Target** | 0.425 | 0.229 | 1.856 | 93.6 |
| **H6 Target** | -0.097 | 0.287 | -0.337 | 26.4 |
| **H7 Target** | 0.048 | 0.342 | 0.141 | 11.2 |
| **H8 Target** | 0.003 | 0.210 | 0.014 | 1.1 |
| **Random Effects** | **Variance** | **Standard Error** | **Percent of Total** | |
| **Day** | 0.350 | 0.101 | 15.3% | |
| **Residual** | 1.933 | 0.103 | 84.7% | |

**Figure 6-1 The attribution distributions for the 10,000 bootstrap and permutation samples, with the estimated attribution denoted by the line, for the 2016 trial.**




As noted earlier, an alternative test of the significance of the estimated attribution is to check whether this value is clearly a consequence of the actual operation of the Atlant systems. If there was no true enhancement effect, then the estimated attribution calculated using the actual operating schedule for the Atlant systems over the trial would not be discernably different from the estimated attribution derived from a random permutation of this operating schedule. Conversely, if there is a true positive enhancement effect, then the estimated attribution based on the actual schedule should rank very highly when compared to estimated attributions derived from randomly permuted schedules.

The details of this permutation test were provided in the 2015 report and the procedure is briefly summarised here. The actual schedule was randomly permutated 10,000 times, with attribution estimated as described previously using the permuted schedules, but otherwise making no change to the 2016 trial data. The resulting permutation distribution is shown in the second panel of Figure 6-1. In this case approximately 22% of the permuted attribution estimates were as large, or larger, than the estimate based on the actual 2016 operating schedule. As a consequence, this result would not be regarded as significant.

The lower significance of the 2016 permutation test result may be due to a combination of low frequency and intensity of rainfall events in 2016 relative to previous years and a substantial number of missing days in the 2016 trial due to missing information on steering wind direction. More importantly, however, we can see from this result that the permutation test provides an independent perspective on the significance of the observed attribution value when all non-Atlant related conditions in the 2016 trial, including the presence of missing wind direction data, are held fixed. In contrast, the bootstrap analysis reported earlier keeps all Atlant related conditions (i.e. the operating schedule) fixed, and reports significance relative to variation of the non-Atlant related conditions. Ideally, both tests would concur, but there is no particular reason why this should be the case. In any case, using both tests then provides a much more robust assessment of real significance, adding confidence to the application of both tests in our analysis below of the combined data over the four trial years.

# 7 Analysis of Combined Data from the 2013 to 2016 Trials

## 7.1 2013-2016 Combined Dataset

Prior to presenting this combined analysis it is worth noting that with a sufficiently large sample, and with stable meteorological conditions both across years and across sites, it would not have been necessary to control for the spatial and temporal variability in rainfall in the Oman Rainfall Enhancement Trial. This is because the randomised experimental design used in the Trial would have averaged out any contributions to either target or control rainfall arising from differences in meteorological conditions between target and control measurements. However, this variability is quite large, and so a large sample extending over many years would have been required to allow a simple comparison of average downwind target and control rainfall. In this context, it is interesting to look at rainfall averages over the full four-year trial period, as shown in Table 7-1.

Overall the number of observations is well balanced except for H7 and H8 where there are a disproportionate number of control observations. At H1 and H2 there are over 12,000 downwind rainfall observations and positive rainfall at the target gauges is, on a weighted average, 10.7 per cent higher. There are over 8,000 downwind rainfall observations at H3 and H4 and while control gauges downwind of H4 record a higher average positive rainfall than target gauges, the weighted average positive rainfall across these two target areas is 7 per cent higher. Of the remaining sites, average positive target rainfall is higher only at H5. However, the number of downwind rainfall observations at H6 is much less than at H5, while there are relatively few downwind rainfall observations at H7 and H8, and these, because of the delay in operating these sites, are quite skewed towards control observations.

**Table 7-1 Rainfall statistics over the four trial years**

| Site | Number of Downwind Rainfall Observations | | Average Positive Rainfall (mm) for Downwind Gauge-Days | |
|---|---|---|---|---|
| | Control | Target | Control | Target |
| **H1** | 3,598 | 3,244 | 4.86 | 5.56 |
| **H2** | 3,708 | 3,666 | 4.84 | 5.20 |
| **H3** | 2,108 | 2,259 | 4.66 | 6.46 |
| **H4** | 2,595 | 2,390 | 4.75 | 3.71 |
| **H5** | 1,491 | 1,406 | 5.28 | 5.45 |
| **H6** | 817 | 830 | 6.77 | 4.29 |
| **H7** | 777 | 377 | 3.94 | 2.00 |
| **H8** | 962 | 585 | 5.22 | 3.07 |

## 7.2 Analysis using 2015 Methodology

The analysis of the combined trial data set was conducted using the same methodology as in 2015. In addition, a variation to the analysis methodology, that restricts the set of downwind gauge-day rainfall values in order to ensure a more appropriate balance of target and control observations in each year, was also conducted. In each year, two Atlant sites have been added. The trial area and gauge network was also expanded. There are two ways in which controls can be incorporated in the combined sample given these changes:

- All observations in a downwind corridor corresponding to a location that has not yet had an Atlant installed are treated as controls, provided that they are not also downwind of an operational Atlant in that year. This means that all gauges that are downwind relative to an Atlant site essentially only generate control values in any year prior to its installation; and
- Downwind observations only generate control values in years when the corresponding Atlant is operational. That is, there are no control values from downwind areas in years prior to installation of the Atlant.

The first option above describes the methodology used in 2105. The second is a more balanced approach. For example, under the first (unbalanced) approach, there were only control observations added downwind of H7 and H8 in the first three trials unless the downwind areas for these sites intersected with the existing gauge network at that time. These sites became operational in 2016 and the gauge network was also expanded at the same time to increase the proximity of coverage near H7 and H8. So the imbalance generated using the 2015 methodology may have two dimensions, one with respect to having more distant gauge locations, and the other due to the number of control versus target observations downwind of these sites. In what follows we generate four-year attribution estimates under the unbalanced and the balanced approaches to defining control gauges.

The upwind model was specified using the same comparison of different stepwise model selection procedures described earlier. The larger sample allowed for the inclusion of more covariates including:

- Indicators for the trial years;
- Gauge elevation (in units of 1000 m);
- Meteorological variables derived from the 4am radiosonde readings at Muscat International Airport; and
- Principal components for weather related variables derived from the DGMAN automatic weather station network.

The upwind model fit based on the unbalanced control gauges design (i.e. data from gauges upwind of non-installed Atlants were included) is summarised in Table 7-2. There were 1,385 rain events upwind of the H1 - H8 locations over the four years, and the model fit explains just under 46% of the corresponding variation in upwind LogRain over this period.

The specification for the four year downwind model set out in Table 7-2 is an updated version of the three year downwind model specification used in 2015. That is, it includes main effects and second order interactions for the three year indicators, gauge elevation and the instrument defined by the upwind model fit set out in Table 7-2. Target indicators for H1 - H8 are included, as well as the interactions of gauge elevation with the target indicators for H1 and H2. The fit of the downwind model for the combined trial data is presented in Table 8. Note that this fit is for the unbalanced controls design.

Four of the target status indicators, H1, H2, H3 and H5, are positive and significant. Two of the target status indicators, H4 and H6, are negative and marginally significant. The remaining two target status indicators (H7 and H8) are not significant.

Both H1 and H2 target status by elevation interaction terms are negative and significant. H1 and H2 are at the highest elevations of all the sites, with both above 2,000 metres. A few of the nearby gauges are between 1,000 and 2,000 metres. These interactions were noted in the 2013 Atlant trial and have been included in all subsequent analyses that include data from 2013. Particularly at H2, they serve to moderate the impact of a gauge being a H1 or H2 target with increasing elevation.

**Table 7-2 The upwind model estimates for the combined trials: unbalanced control gauges (2015 methodology).**

| **Summary** | | | | |
|---|---|---|---|---|
| **R Squared** | 0.458 | | | |
| **Observations** | 1385 | | | |
| **Parameters** | **Estimate** | **Standard Error** | **T-Value** | **Significance** |
| **Intercept** | -1.460 | 0.553 | -2.642 | 99.1 |
| **Year 2013** | 0.377 | 0.209 | 1.805 | 92.7 |
| **Year 2014** | 0.139 | 0.219 | 0.637 | 47.4 |
| **Year 2015** | 0.230 | 0.203 | 1.129 | 73.9 |
| **Gauge Elevation (1000m)** | 0.551 | 0.087 | 6.343 | 100.0 |
| **Steering Wind Speed** | -0.063 | 0.028 | -2.233 | 97.3 |
| **Totals Total Index** | 0.024 | 0.011 | 2.195 | 97.0 |
| **Dry Bulb Temp PC2** | 0.198 | 0.070 | 2.820 | 99.5 |
| **Humidity PC1** | 0.199 | 0.028 | 7.025 | 100.0 |
| **Pressure PC1** | 0.078 | 0.028 | -2.839 | 99.5 |
| **Random Effects** | **Variance** | **Standard Error** | **Percent of Total** | |
| **Day** | 0.494 | 0.088 | 22.9% | |
| **Residual** | 1.669 | 0.068 | 77.1% | |

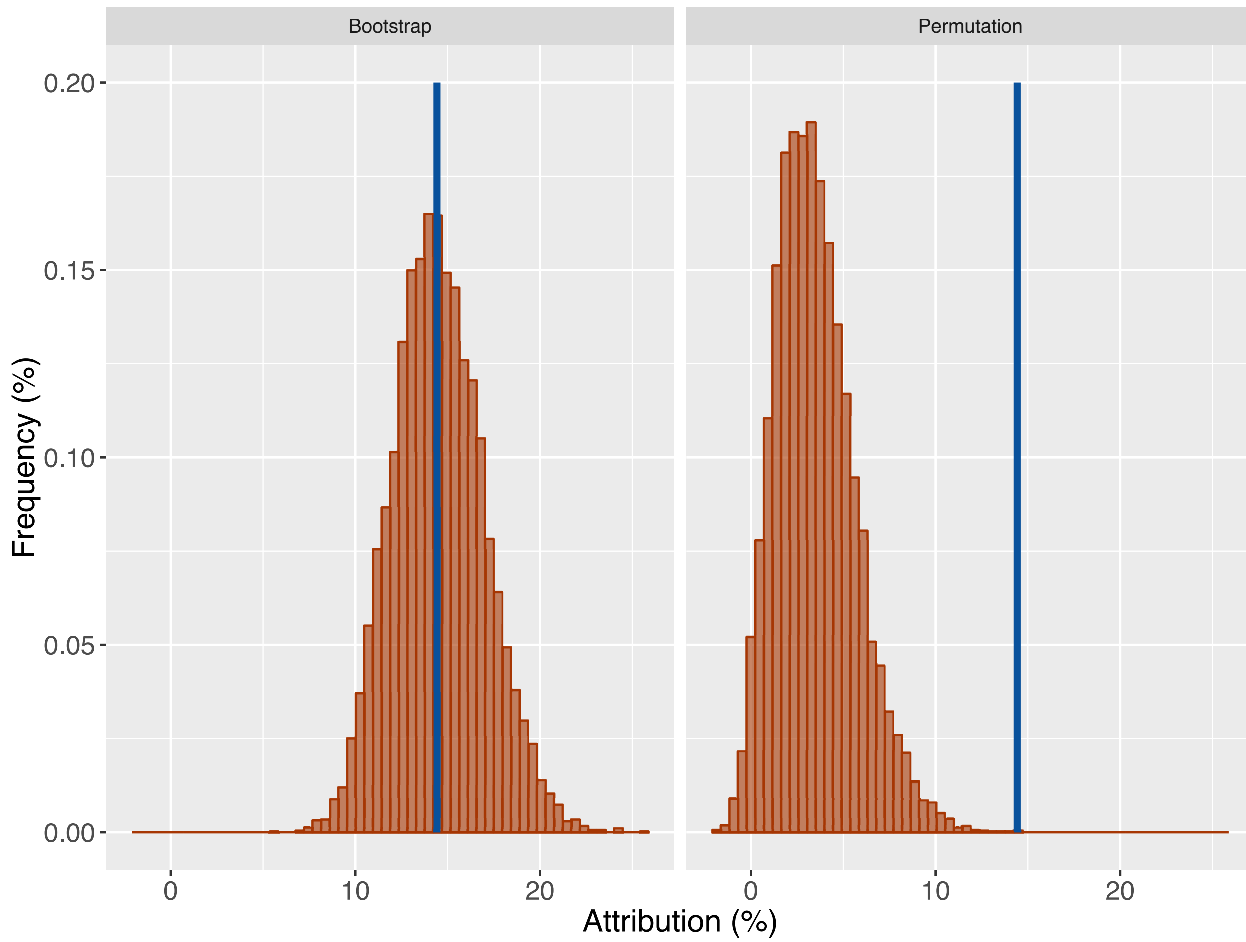
**Table 7-3 The downwind model estimates for the combined trials: unbalanced control gauges (2015 methodology).**

| Summary | | | | |
|---|---|---|---|---|
| R Squared | 0.247 | | | |
| Observations | 4107 | | | |
| Parameters | Estimate | Standard Error | T-Value | Significance |
| Intercept | -1.247 | 0.229 | -5.451 | 100.0 |
| Year 2013 | 0.557 | 0.337 | 1.656 | 95.1 |
| Year 2014 | 0.080 | 0.265 | 0.301 | 61.8 |
| Year 2015 | -0.047 | 0.268 | -0.175 | 57.0 |
| Gauge Elevation (1000m) | 0.284 | 0.220 | 1.295 | 90.2 |
| Gauge Elevation*2013 | -0.076 | 0.245 | -0.308 | 62.1 |
| Gauge Elevation*2014 | -0.437 | 0.220 | -1.981 | 97.6 |
| Gauge Elevation*2015 | -0.152 | 0.218 | -0.698 | 75.7 |
| Instrument | 1.939 | 0.357 | 5.438 | 100.0 |
| Instrument*2013 | -0.538 | 0.555 | -0.971 | 83.4 |
| Instrument*2014 | 0.453 | 0.430 | 1.051 | 85.3 |
| Instrument*2015 | -0.076 | 0.388 | -0.195 | 57.7 |
| Instrument*Elevation | -0.301 | 0.154 | -1.962 | 97.5 |
| H1 Target | 0.311 | 0.150 | 2.075 | 98.1 |
| H2 Target | 0.335 | 0.133 | 2.525 | 99.4 |
| H3 Target | 0.326 | 0.104 | 3.145 | 99.9 |
| H4 Target | -0.155 | 0.102 | -1.517 | 93.5 |
| H5 Target | 0.528 | 0.160 | 3.296 | 100.0 |
| H6 Target | -0.280 | 0.176 | -1.597 | 94.5 |
| H7 Target | 0.061 | 0.334 | 0.183 | 57.3 |
| H8 Target | -0.013 | 0.203 | -0.066 | 52.6 |
| Elevation*H1 Target | -0.251 | 0.161 | -1.555 | 94.0 |
| Elevation*H2 Target | -0.330 | 0.134 | -2.464 | 99.3 |
| Random Effects | Variance | Standard Error | Percent of Total | |
| Day | 0.302 | 0.041 | 13.8% | |
| Residual | 1.883 | 0.043 | 86.2% | |

The estimated attribution across all four years for all downwind gauges for H1 - H8 was a 14.4 per cent increase over estimated natural rainfall over the same period. A bootstrap analysis of this estimated attribution lead to a bootstrap average attribution value that was almost identical (14.5 per cent) and a bootstrap standard deviation of 2.5 per cent. This result is highly significant, at a level of significance greater than 99.9 per cent. This is clear from the bootstrap distribution of the estimated attribution shown in the left panel of Figure 7-1, which is entirely above zero.

The permutation distribution for the estimated attribution defined by the four-year operating sequence is shown in the right panel of Figure 7-2. The estimated attribution can be seen to sit at the positive extremum of this distribution providing robust confirmation that, given the range of conditions in the four trials, the four-year attribution is highly significant.

**Figure 7-3 The attribution distributions for the 10,000 bootstrap and permutation samples, with the estimated attribution denoted by the line, for the 2013 - 2016 combined trials: unbalanced control gauges (2015 methodology).**



### 7.3 Analysis of the Balanced Control Gauge Design

The unbalanced control gauges design leads to many more controls than targets, and consequently a reduction in estimated attribution (since it is spread over many more gauges that are not exposed to any Atlant effect). Consequently, we now show model fits and estimated attribution when the design of the combined data is restricted to downwind areas that are defined by installed (and operated) Atlants. The upwind model is presented in Table 7-4. The results are similar to those obtained with the unbalanced sample, despite a seven per cent reduction in sample size. Most of the coefficients and the significance levels are of the same order of magnitude. The variance components associated with the random day effects are virtually identical.

The corresponding downwind model is shown in Table 7-5. Removing control values associated with non-installed Atlants reduces the sample size by about 8 per cent. The target status indicators at H1, H2, H3 and H5 are now all highly significant. The target status indicator for H6 is the only one that is negative, and significant at the ninety per cent level. The target status indicators for the remaining three sites are not significant. The elevation interaction term with H2 target status remains significant.

The balanced downwind gauges design leads to an estimated attribution of 19.0 per cent over the four years. A bootstrap analysis of this estimated attribution, using the four year balanced downwind gauges data set, leads to an average attribution of 19.2 per cent with a standard deviation of 3.3 per cent. This is significant at the 99.9 per cent level, and is again clear from the empirical bootstrap distribution of the estimated attribution shown in the left panel of Figure 7-4, which is entirely above zero.

The permutation distribution of the estimated attribution distribution, based on random permutations of the operating schedule across all four years is shown in the right panel of Figure 7-5. The estimated attribution based on the actual schedule is at the positive extremum of this distribution, confirming that the four year attribution is positive and highly significant.

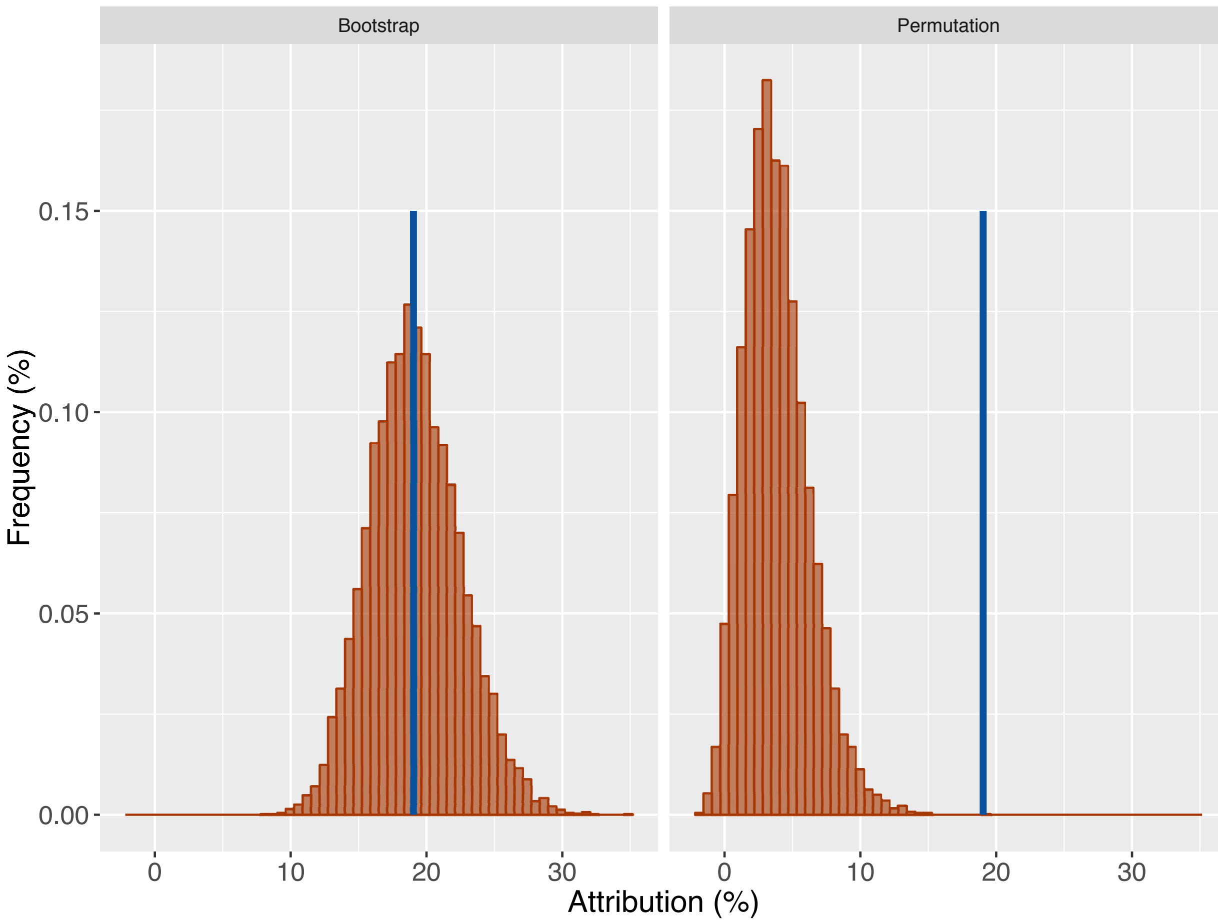
**Table 7-4 The upwind model estimates for the combined trials: balanced control gauges.**

| Summary | | | | |
|---|---|---|---|---|
| R Squared | 0.467 | | | |
| Observations | 1292 | | | |
| **Parameters** | **Estimate** | **Standard Error** | **T-Value** | **Significance** |
| Intercept | -1.463 | 0.544 | -2.690 | 99.6 |
| Year 2013 | 0.201 | 0.214 | 0.942 | 82.7 |
| Year 2014 | 0.184 | 0.225 | 0.817 | 79.3 |
| Year 2015 | 0.234 | 0.203 | 1.153 | 87.6 |
| Gauge Elevation (1000m) | 0.500 | 0.083 | 6.015 | 100.0 |
| Steering Wind Speed | -0.068 | 0.028 | -2.461 | 99.3 |
| Totals Total Index | 0.025 | 0.011 | 2.342 | 99.0 |
| Dry Bulb Temp PC2 | 0.193 | 0.070 | 2.770 | 99.7 |
| Humidity PC1 | 0.202 | 0.029 | 7.034 | 100.0 |
| Pressure PC1 | -0.075 | 0.028 | -2.711 | 99.7 |
| **Random Effects** | **Variance** | **Standard Error** | **Percent of Total** | |
| Day | 0.498 | 0.090 | 23.6% | |
| Residual | 1.609 | 0.068 | 76.4% | |

**Table 7-5 The downwind model estimates for the combined trials: balanced control gauges.**

| Summary | | | | |
|---|---|---|---|---|
| R Squared | 0.262 | | | |
| Observations | 3224 | | | |
| **Parameters** | **Estimate** | **Standard Error** | **T-Value** | **Significance** |
| Intercept | -1.241 | 0.232 | -5.341 | 100.0 |
| Year 2013 | 0.566 | 0.357 | 1.586 | 88.6 |
| Year 2014 | 0.032 | 0.277 | 0.114 | 9.1 |
| Year 2015 | -0.025 | 0.273 | -0.091 | 7.2 |
| Gauge Elevation (1000m) | 0.293 | 0.220 | 1.331 | 81.6 |
| Gauge Elevation*2013 | -0.086 | 0.254 | -0.339 | 26.5 |
| Gauge Elevation*2014 | -0.415 | 0.215 | -1.925 | 94.5 |
| Gauge Elevation*2015 | -0.125 | 0.212 | -0.588 | 44.3 |
| Instrument | 1.859 | 0.350 | 5.317 | 100.0 |
| Instrument*2013 | -0.360 | 0.606 | -0.594 | 44.7 |
| Instrument*2014 | 0.408 | 0.425 | 0.961 | 66.2 |
| Instrument*2015 | -0.145 | 0.377 | -0.386 | 29.9 |
| Instrument*Elevation | -0.243 | 0.166 | -1.462 | 85.6 |
| H1 Target | 0.368 | 0.160 | 2.292 | 97.8 |
| H2 Target | 0.386 | 0.138 | 2.797 | 99.5 |
| H3 Target | 0.374 | 0.106 | 3.540 | 100.0 |
| H4 Target | -0.096 | 0.104 | -0.926 | 64.5 |
| H5 Target | 0.556 | 0.160 | 3.467 | 99.9 |
| H6 Target | -0.247 | 0.176 | -1.408 | 84.0 |
| H7 Target | 0.076 | 0.333 | 0.228 | 18.0 |
| H8 Target | -0.006 | 0.202 | -0.029 | 2.3 |
| Elevation*H1 Target | -0.282 | 0.174 | -1.625 | 89.6 |
| Elevation*H2 Target | -0.361 | 0.138 | -2.620 | 99.1 |
| **Random Effects** | **Variance** | **Standard Error** | **Percent of Total** | |
| Day | 0.291 | 0.044 | 13.5% | |
| Residual | 1.868 | 0.049 | 86.5% | |

**Figure 7-6 The attribution distributions for the 10,000 bootstrap and permutation samples, with the estimated attribution denoted by the line, for the 2013 - 2016 combined trials: balanced control gauges.**



# 8 Concluding Comments

The 2016 trial was conducted in a year which had a considerably lower frequency of gauge level rainfall events that any of the previous three trials. This resulted in a smaller downwind sample of rain events that could be used to model positive rainfall. The sample size was further reduced because missing radiosonde observations meant that wind directions were not available for extended periods at the start of the trial, and so downwind direction could not be reliably ascertained. To some extent, this was mostly recovered by using trend based imputation of steering wind directions and other derived information from the missing radiosondes. These imputations were only carried out where there were three or fewer consecutive days missing in order to ensure that the imputation had adequate support. The installation of the two new Atlant sites (H7 and H8) was also delayed due to a technical fault in the new system design.

Gauge level rainfall events, when they did occur, tended to be lower than in previous trials as well. This implied that rainfall enhancement, if it did occur, would be smaller making the signal more difficult to detect. Nevertheless, the estimated enhancement effect for 2016 was within the range of previous trials at about 17 per cent. The bootstrap distribution of attribution indicated that estimate was significant. However, the permutation test of the operating schedule was not.

The 2016 trial did contribute a substantial number of observations to the data from the previous three trials. The methodology developed in 2015 to analyse data from combined trials was therefore used to analyse an expanded four year dataset, combining the 2013 - 2016 trials. A slight modification to the combined trial methodology used in 2015 was adopted in order to ensure target and control gauges were better balanced in the combined downwind data set. This was necessary because the 2015 assumption that gauges downwind of non-installed Atlants could be treated as controls meant that by 2016 the ratio of the number of controls to number of target gauge observations had become quite large. A alternative balanced design in which only gauges downwind of operational sites were designated as either controls or targets was therefore also implemented.

Under both designs the bootstrap attribution was highly significant. In the permutation test the estimated attribution generated under the actual four year operating schedule was the maximum of the randomly permutated schedules in each case. However, the estimated attributions were different, reflecting the fact that the unbalanced design calculated attribution relative to many more control gauges than target gauges, while in the balanced design these two counts were much closer. The estimated attributions were:

- 14.4 per cent for the unbalanced controls design; and
- 19.0 per cent for the balanced controls design.

While there are good arguments for adopting the balanced controls design in 2016 and into the future, the take home message is that although attribution results may be highly significant, their absolute size is conditional on the target vs. control balance of the sample achieved under trial conditions. In this case the imbalance was due to the retrospective inclusion of gauge observations that fell downwind of a future site in the data set used to compute attribution. This led to an increased estimate in the total amount of natural rainfall with no corresponding increase in attributed rainfall (since there was no operating Atlant that could have been responsible). Given the spatial variability of summer rainfall in the Hajar ranges the potential for obtaining a similar imbalance between target and control gauges is not inconsequential. Similar issues arise when data become missing, for example because of non-operation of one or more Atlants. Both the bootstrap and permutation analyses reported here are most efficient when this balance exists, and combining data from several years can help smooth temporal and spatial imbalance, as well as protect against inefficiencies due to missing data.

# Appendix A – 2016 Instrumentation

## Atlant Installations

| Station | Latiutude (Decimal) | Longitude (Decimal) | Elevation (m) |
|---|---|---|---|
| **H1** | 23.32246667 | 57.12330000 | 2670 |
| **H2** | 23.17633333 | 57.53166667 | 2157 |
| **H3** | 23.69440000 | 56.80110000 | 1621 |
| **H4** | 23.17320000 | 58.04420000 | 1395 |
| **H5** | 23.16423600 | 58.60931000 | 1876 |
| **H6** | 23.54725500 | 56.54952000 | 1466 |
| **H7** | 23.03750000 | 58.89194400 | 1574 |
| **H8** | 23.51541700 | 56.97934200 | 1501 |

## DGMAN Weather Stations

| Station | Latiutude (Decimal) | Longitude (Decimal) | Elevation (m) |
|---|---|---|---|
| **Adam Airpoirt** | 22.49982400 | 57.37006600 | 328 |
| **Al-Amirat** | 23.50008100 | 58.48083500 | 105 |
| **Bahla** | 22.99322200 | 57.31592200 | 592 |
| **Bidiyah** | 22.46282800 | 58.85467500 | 316 |
| **Buraimi** | 24.23391400 | 55.91617400 | 372 |
| **Ibra** | 22.74044400 | 58.50525000 | 469 |
| **Ibri** | 23.19541700 | 56.42944400 | 323 |
| **Mahadha** | 24.39990000 | 55.97980000 | 438 |
| **Majis** | 24.46744400 | 56.64442800 | 2 |
| **Muscat Airport** | 23.59534500 | 58.29846100 | 8 |
| **Nizwa** | 22.85907500 | 57.54629400 | 462 |
| **Qumaira** | 23.93202800 | 56.20072200 | 633 |
| **Rushtaq** | 23.40250000 | 57.42861100 | 322 |
| **Saiq** | 23.07397200 | 57.66652800 | 1995 |
| **Smail** | 23.30900000 | 57.94805600 | 417 |
| **Sunaynah** | 23.72258300 | 55.90597200 | 257 |
| **Suwaiq** | 23.82651900 | 57.28013100 | 39 |

## TIE Rain Gauges

| Gauge ID | Latitude | Longitude | Elevation |
|---|---|---|---|
| **1** | 22.7905 | 57.85393333 | 479 |
| **2** | 22.84041667 | 57.8799 | 540 |
| **3** | 22.87118333 | 57.76013333 | 495 |
| **5** | 22.8143 | 57.5914 | 445 |
| **6** | 22.90993333 | 57.59236667 | 508 |
| **7** | 22.93595 | 57.50846667 | 547 |
| **8** | 22.82906667 | 57.48546667 | 483 |
| **9** | 22.88541667 | 57.35626667 | 555 |
| **10** | 23.0269 | 57.3671 | 631 |
| **11** | 23.00066667 | 57.31008333 | 591 |
| **12** | 22.92211667 | 57.25565 | 540 |
| **13** | 22.94205 | 57.12701667 | 519 |
| **14** | 23.03008333 | 57.15721667 | 557 |
| **15** | 23.06598333 | 57.01848333 | 886 |
| **16** | 22.99316667 | 57.00695 | 607 |
| **17** | 22.9643 | 56.86095 | 559 |
| **18** | 23.0767 | 56.7584 | 542 |
| **19** | 23.11783333 | 56.7835 | 464 |
| **20** | 22.99511667 | 56.74363333 | 462 |
| **21** | 22.9595 | 57.9015 | 253 |
| **22** | 23.02603333 | 57.92741667 | 747 |
| **23** | 23.07695 | 57.81686667 | 677 |
| **24** | 22.97521667 | 57.77885 | 588 |
| **25** | 23.0241 | 57.69378333 | 2015 |
| **26** | 23.0892 | 57.67875 | 2023 |
| **27** | 23.10643333 | 57.53411667 | 2000 |
| **28** | 23.03963333 | 57.50585 | 603 |
| **29** | 23.0672 | 57.42678333 | 771 |
| **30** | 23.15676667 | 58.25585 | 1958 |
| **31** | 23.15261667 | 57.2963 | 1263 |
| **32** | 23.10181667 | 57.28525 | 655 |
| **33** | 23.108 | 57.20495 | 762 |
| **34** | 23.2292 | 57.19921667 | 1968 |
| **35** | 23.22735 | 57.04995 | 729 |
| **36** | 23.15026667 | 57.07833333 | 964 |
| **37** | 23.16625 | 56.91615 | 590 |
| **38** | 23.2751 | 56.95531667 | 464 |
| **39** | 23.30091667 | 56.83566667 | 587 |
| **40** | 23.23026667 | 56.81385 | 547 |
| **41** | 23.11428333 | 57.9264 | 897 |

| | | | |
|---|---|---|---|
| **42** | 23.25038333 | 57.98971667 | 448 |
| **43** | 23.27705 | 57.91983333 | 426 |
| **44** | 23.16615 | 57.85278333 | 564 |
| **45** | 23.17681667 | 57.65133333 | 915 |
| **46** | 23.29108333 | 57.69321667 | 517 |
| **47** | 23.32113333 | 57.60571667 | 405 |
| **48** | 23.19951667 | 57.55973333 | 729 |
| **49** | 23.26558333 | 57.4456 | 587 |
| **50** | 23.36283333 | 57.47665 | 398 |
| **51** | 23.39985 | 57.33798333 | 442 |
| **52** | 23.27431667 | 57.33466667 | 715 |
| **53** | 23.28351667 | 57.2635 | 1677 |
| **54** | 23.39326667 | 57.23501667 | 535 |
| **55** | 23.45328333 | 57.105 | 677 |
| **57** | 23.37656667 | 57.0154 | 961 |
| **58** | 23.4644 | 57.0309 | 897 |
| **59** | 23.49973333 | 56.86356667 | 708 |
| **60** | 23.38526667 | 56.8564 | 693 |
| **61** | 23.31795 | 58.02283333 | 352 |
| **62** | 23.41255 | 58.083 | 263 |
| **63** | 23.4541 | 57.88523333 | 236 |
| **64** | 23.3966 | 57.84565 | 304 |
| **65** | 23.3815 | 57.76328333 | 321 |
| **66** | 23.45325 | 57.81663333 | 228 |
| **67** | 23.49655 | 57.67183333 | 186 |
| **68** | 23.37623333 | 57.66453333 | 337 |
| **69** | 23.45768333 | 57.51601667 | 275 |
| **70** | 23.54116667 | 57.5224 | 174 |
| **71** | 23.55263333 | 57.45201667 | 195 |
| **72** | 23.44176667 | 57.39368333 | 340 |
| **73** | 23.48413333 | 57.28715 | 377 |
| **75** | 23.65056667 | 57.14568333 | 288 |
| **76** | 23.5688 | 57.18168333 | 405 |
| **77** | 23.529 | 57.0109 | 764 |
| **78** | 23.65141667 | 57.08898333 | 413 |
| **79** | 23.66926667 | 56.9236 | 413 |
| **80** | 23.5669 | 56.89798333 | 808 |
| **82** | 23.62076667 | 58.05003333 | 35 |
| **84** | 23.5437 | 57.9673 | 116 |
| **85** | 23.56455 | 57.83578333 | 118 |
| **86** | 23.65433333 | 57.85886667 | 31 |
| **87** | 23.69096667 | 57.7077 | 29 |
| **89** | 23.63318333 | 57.52853333 | 107 |
| **92** | 23.64005 | 57.38251667 | 145 |

| | | | |
|---|---|---|---|
| **94** | 23.753 | 57.33158333 | 71 |
| **96** | 23.70503333 | 57.19936667 | 167 |
| **97** | 23.71715 | 57.0996 | 243 |
| **99** | 23.83296667 | 56.96405 | 104 |
| **100** | 23.75793333 | 56.92568333 | 310 |
| **101** | 23.76551667 | 56.83961667 | 364 |
| **103** | 23.65495 | 56.79543333 | 850 |
| **104** | 23.77128333 | 56.62815 | 750 |
| **105** | 23.78348333 | 56.73318333 | 501 |
| **106** | 23.5488 | 56.77233333 | 691 |
| **107** | 23.68541667 | 56.55198333 | 727 |
| **108** | 23.4101 | 56.7396 | 525 |
| **109** | 23.55543333 | 56.53355 | 532 |
| **110** | 23.63585 | 56.63636667 | 714 |
| **111** | 23.37913333 | 58.13855 | 257 |
| **112** | 23.36283333 | 58.39515 | 360 |
| **113** | 23.23778333 | 58.1191 | 394 |
| **114** | 23.14635 | 58.31315 | 739 |
| **115** | 23.2613 | 58.21771667 | 486 |
| **116** | 23.02756667 | 58.0171 | 683 |
| **117** | 23.06911667 | 58.32026667 | 711 |
| **118** | 22.94276667 | 58.05643333 | 587 |
| **119** | 22.86338333 | 58.21401667 | 650 |
| **120** | 23.07078333 | 58.18416667 | 522 |
| **121** | 22.80153333 | 57.40861667 | 672 |
| **122** | 22.8742 | 57.17423333 | 698 |
| **123** | 22.74526667 | 57.24565 | 560 |
| **124** | 22.68801667 | 57.5695 | 508 |
| **125** | 22.77585 | 57.64893333 | 554 |
| **126** | 22.70733333 | 57.73465 | 543 |
| **127** | 22.74645 | 57.85066667 | 695 |
| **128** | 22.7033 | 58.02601667 | 518 |
| **129** | 22.99278333 | 58.6007 | 624 |
| **130** | 22.87183333 | 58.05576667 | 357 |
| **131** | 22.70041667 | 58.15693333 | 406 |
| **132** | 22.85043333 | 58.59606667 | 471 |
| **133** | 22.86436667 | 58.30776667 | 636 |
| **134** | 22.96781667 | 58.2398 | 534 |
| **135** | 22.74743333 | 58.2858 | 322 |
| **136** | 22.6666 | 58.253 | 333 |
| **137** | 22.72413333 | 58.40973333 | 316 |
| **138** | 22.80675 | 58.43965 | 671 |
| **139** | 22.97213333 | 58.44955 | 573 |
| **140** | 23.11751667 | 58.45363333 | 518 |

| | | | |
|---|---|---|---|
| **141** | 22.80798333 | 56.99153333 | 481 |
| **142** | 23.06801667 | 56.6241 | 753 |
| **143** | 23.21561667 | 56.66201667 | 400 |
| **144** | 23.35405 | 56.64546667 | 589 |
| **145** | 23.54943333 | 56.6622 | 396 |
| **146** | 23.91891667 | 56.55045 | 314 |
| **147** | 23.29193333 | 56.37348333 | 398 |
| **148** | 23.21816667 | 56.46545 | 296 |
| **149** | 24.06913333 | 56.55985 | 710 |
| **150** | 23.68706667 | 56.45405 | 671 |
| **151** | 23.86631667 | 56.40633333 | 573 |
| **152** | 23.9444 | 56.35355 | 320 |
| **153** | 23.77371667 | 56.207 | 481 |
| **154** | 23.75468333 | 56.39401667 | 753 |
| **155** | 23.66138333 | 56.21836667 | 400 |
| **156** | 23.63938333 | 56.38401667 | 589 |
| **157** | 23.54493333 | 56.3479 | 396 |
| **158** | 23.50378333 | 56.20198333 | 314 |
| **159** | 23.3476 | 56.51526667 | 398 |
| **160** | 23.39058333 | 56.29218333 | 296 |
| **161** | 24.07411667 | 56.41856667 | 710 |
| **162** | 24.08966667 | 56.2624 | 856 |
| **163** | 23.58866667 | 56.00343333 | 267 |
| **164** | 24.10058333 | 56.07456667 | 516 |
| **165** | 24.21011667 | 56.17711667 | 594 |
| **166** | 24.24501667 | 56.35495 | 352 |
| **167** | 23.78411667 | 56.0134 | 320 |
| **168** | 24.2245 | 55.95666667 | 411 |
| **169** | 23.78786667 | 55.87471667 | 261 |
| **170** | 23.935 | 55.82521667 | 285 |
| **171** | 24.0167 | 55.9591 | 373 |
| **172** | 23.8313 | 56.4688 | 905 |
| **173** | 23.93441667 | 56.20175 | 642 |
| **174** | 23.86398333 | 56.64383333 | 612 |
| **175** | 24.00091667 | 56.65635 | 333 |
| **176** | 23.30573333 | 58.48318333 | 330 |
| **177** | 23.27953333 | 58.39725 | 512 |
| **178** | 23.22173333 | 58.68331667 | 277 |
| **179** | 23.2679 | 58.633 | 298 |
| **180** | 23.26515 | 58.75718333 | 163 |
| **181** | 23.33895 | 58.58741667 | 211 |
| **182** | 23.08681667 | 58.56865 | 412 |
| **183** | 23.04473333 | 58.64965 | 337 |
| **184** | 22.99258333 | 58.77608333 | 325 |

| | | | |
|---|---|---|---|
| **185** | 23.13403333 | 58.53388333 | 451 |
| **186** | 23.90515 | 56.74936667 | 209 |
| **187** | 22.88926667 | 58.6659 | 558 |
| **188** | 22.76376667 | 58.6678 | 429 |
| **189** | 22.72708333 | 58.52333333 | 467 |
| **190** | 23.09233333 | 58.85635 | 139 |
| **191** | 23.217 | 58.3101 | 687 |
| **192** | 23.34671667 | 58.2266 | 352 |
| **193** | 22.92661667 | 58.89511667 | 695 |
| **194** | 22.70915 | 58.76951667 | 452 |
| **195** | 22.83636667 | 58.86326667 | 700 |
| **196** | 22.62678333 | 58.67003333 | 397 |
| **197** | 22.88875 | 59.00393333 | 1277 |
| **198** | 22.77068333 | 58.96036667 | 886 |
| **199** | 22.68076667 | 58.89528333 | 518 |
| **200** | 22.57131667 | 58.80908333 | 382 |
| **201** | 22.68376667 | 58.97805 | 585 |
| **202** | 22.4673 | 58.8917 | 305 |
| **203** | 22.53715 | 58.97118333 | 404 |

# Appendix B – 2016 Operating Schedule

| Month | Day | H6 | H3 | H8 | H1 | H2 | H4 | H5 | H7 |
|---|---|---|---|---|---|---|---|---|---|
| June | 1 | 0 | 1 | | 0 | 0 | 1 | 1 | |
| June | 2 | 1 | 1 | | 0 | 1 | 0 | 0 | |
| June | 3 | 0 | 0 | | 1 | 0 | 1 | 1 | |
| June | 4 | 0 | 1 | | 0 | 1 | 1 | 0 | |
| June | 5 | 1 | 0 | | 1 | 0 | 0 | 1 | |
| June | 6 | 0 | 1 | | 0 | 0 | 1 | 1 | |
| June | 7 | 1 | 0 | | 1 | 1 | 0 | 0 | |
| June | 8 | 0 | 1 | | 1 | 0 | 1 | 0 | |
| June | 9 | 1 | 0 | | 0 | 1 | 0 | 1 | |
| June | 10 | 1 | 1 | | 0 | 0 | 1 | 0 | |
| June | 11 | 0 | 0 | | 1 | 1 | 0 | 1 | |
| June | 12 | 1 | 0 | | 1 | 0 | 1 | 0 | |
| June | 13 | 0 | 1 | | 0 | 1 | 0 | 1 | |
| June | 14 | 1 | 0 | | 0 | 1 | 1 | 0 | |
| June | 15 | 0 | 1 | | 1 | 0 | 0 | 1 | |
| June | 16 | 1 | 0 | | 0 | 1 | 0 | 1 | |
| June | 17 | 0 | 1 | | 1 | 0 | 1 | 0 | |
| June | 18 | 1 | 0 | | 1 | 1 | 0 | 0 | |
| June | 19 | 0 | 1 | | 0 | 0 | 1 | 1 | |
| June | 20 | 1 | 0 | | 1 | 0 | 0 | 1 | |
| June | 21 | 0 | 1 | | 0 | 1 | 1 | 0 | |
| June | 22 | 0 | 1 | | 1 | 0 | 0 | 1 | |
| June | 23 | 1 | 0 | | 0 | 1 | 1 | 0 | |
| June | 24 | 0 | 1 | | 0 | 1 | 0 | 1 | |
| June | 25 | 1 | 0 | | 1 | 0 | 1 | 0 | |
| June | 26 | 0 | 0 | | 1 | 1 | 0 | 1 | |
| June | 27 | 1 | 1 | | 0 | 0 | 1 | 0 | |
| June | 28 | 0 | 0 | | 1 | 0 | 1 | 1 | |
| June | 29 | 1 | 1 | | 0 | 1 | 0 | 0 | |
| June | 30 | 1 | 0 | | 1 | 1 | 0 | 0 | |
| July | 1 | 1 | 1 | x | 1 | 0 | 1 | 0 | x |
| July | 2 | 0 | 0 | x | 0 | 1 | 0 | 1 | x |
| July | 3 | 0 | 1 | x | 1 | 0 | 1 | 0 | x |
| July | 4 | 1 | 0 | x | 0 | 1 | 0 | 1 | x |
| July | 5 | 0 | 1 | x | 1 | 0 | 1 | 1 | x |
| July | 6 | 1 | 0 | x | 0 | 1 | 0 | 0 | x |
| July | 7 | 0 | 1 | x | 0 | 0 | 1 | 1 | x |
| July | 8 | 1 | 0 | x | 1 | 1 | 0 | 0 | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **July** | **9** | 0 | 1 | x | 0 | 1 | 0 | 1 | x |
| **July** | **10** | 1 | 0 | x | 1 | 0 | 1 | 0 | x |
| **July** | **11** | 0 | 1 | x | 0 | 1 | 1 | 0 | x |
| **July** | **12** | 1 | 0 | x | 1 | 0 | 0 | 1 | x |
| **July** | **13** | 0 | 1 | x | 0 | 1 | 0 | 1 | x |
| **July** | **14** | 1 | 0 | x | 1 | 0 | 1 | 0 | x |
| **July** | **15** | 0 | 1 | x | 1 | 1 | 0 | 0 | x |
| **July** | **16** | 1 | 0 | x | 0 | 0 | 1 | 1 | x |
| **July** | **17** | 1 | 0 | x | 0 | 0 | 1 | 0 | x |
| **July** | **18** | 0 | 1 | x | 1 | 1 | 0 | 1 | x |
| **July** | **19** | 1 | 0 | x | 0 | 1 | 0 | 0 | x |
| **July** | **20** | 0 | 1 | x | 1 | 0 | 1 | 1 | x |
| **July** | **21** | 1 | 0 | x | 1 | 1 | 0 | 0 | x |
| **July** | **22** | 0 | 1 | x | 0 | 0 | 1 | 1 | x |
| **July** | **23** | 1 | 1 | x | 0 | 1 | 0 | 0 | x |
| **July** | **24** | 0 | 0 | x | 1 | 0 | 1 | 1 | x |
| **July** | **25** | 0 | 1 | x | 1 | 1 | 0 | 1 | x |
| **July** | **26** | 1 | 0 | x | 0 | 0 | 1 | 0 | x |
| **July** | **27** | 0 | 0 | x | 1 | 0 | 1 | 1 | x |
| **July** | **28** | 1 | 1 | x | 0 | 1 | 0 | 0 | x |
| **July** | **29** | 0 | 0 | 1 | 1 | 0 | 1 | 0 | x |
| **July** | **30** | 1 | 1 | 0 | 0 | 1 | 0 | 1 | x |
| **July** | **31** | 0 | 1 | 1 | 0 | 0 | 1 | 0 | x |
| **August** | **1** | 1 | 0 | 0 | 1 | 1 | 0 | 1 | x |
| **August** | **2** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | x |
| **August** | **3** | 1 | 0 | 0 | 1 | 0 | 1 | 1 | x |
| **August** | **4** | 1 | 0 | 1 | 0 | 1 | 1 | 0 | x |
| **August** | **5** | 0 | 1 | 0 | 1 | 0 | 0 | 1 | x |
| **August** | **6** | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| **August** | **7** | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| **August** | **8** | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| **August** | **9** | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| **August** | **10** | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **August** | **11** | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| **August** | **12** | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| **August** | **13** | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| **August** | **14** | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **August** | **15** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| **August** | **16** | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| **August** | **17** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| **August** | **18** | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **August** | **19** | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **August** | **20** | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| **August** | **21** | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| **August** | **22** | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| **August** | **23** | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| **August** | **24** | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| **August** | **25** | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| **August** | **26** | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| **August** | **27** | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| **August** | **28** | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| **August** | **29** | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| **August** | **30** | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| **August** | **31** | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| **September** | **1** | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| **September** | **2** | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| **September** | **3** | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| **September** | **4** | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| **September** | **5** | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| **September** | **6** | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| **September** | **7** | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| **September** | **8** | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| **September** | **9** | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **September** | **10** | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| **September** | **11** | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| **September** | **12** | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| **September** | **13** | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| **September** | **14** | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |
| **September** | **15** | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| **September** | **16** | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| **September** | **17** | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| **September** | **18** | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| **September** | **19** | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| **September** | **20** | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| **September** | **21** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| **September** | **22** | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| **September** | **23** | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| **September** | **24** | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| **September** | **25** | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| **September** | **26** | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| **September** | **27** | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |
| **September** | **28** | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **September** | **29** | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| **September** | **30** | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| **October** | **1** | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| **October** | **2** | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| **October** | **3** | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| **October** | **4** | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| **October** | **5** | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| **October** | **6** | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| **October** | **7** | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| **October** | **8** | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **October** | **9** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| **October** | **10** | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **October** | **11** | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| **October** | **12** | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| **October** | **13** | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| **October** | **14** | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| **October** | **15** | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| **October** | **16** | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| **October** | **17** | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **October** | **18** | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| **October** | **19** | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| **October** | **20** | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| **October** | **21** | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **October** | **22** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| **October** | **23** | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| **October** | **24** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| **October** | **25** | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| **October** | **26** | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **October** | **27** | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| **October** | **28** | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| **October** | **29** | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| **October** | **30** | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| **October** | **31** | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| **Total Operating Days** | | **77** | **77** | **47** | **76** | **77** | **77** | **76** | **43** |

# Appendix C – Imputation of Steering Winds

### Trend Smoothing Approach

In order to impute the missing radiosonde values for wind direction and speed a trend-based imputation process was developed. This approach requires sufficient data on either side of days with missing observations in order to provide a reasonable basis for trend interpolation. The largest data gap allowed is three consecutive days, with any gaps larger than this identified as non-imputable. Steering wind data for these days were regarded as missing values in the final analysis. There were three sequences of days with missing radiosonde data that were regarded as non-imputable in the 2016 trial.

Missing steering wind direction values for imputable days were imputed using a two-stage smoothing method based on the 'twicing' approach suggested by Tukey (1977) which exploits the nearby in time correlation of the data. Similar technology is used in the function “smooth” in R (Svetunkov, 2017). The twiced version of a smoother F(y) is defined as F(y) + F(y - F(y)), i.e., smoothed residuals from the original smooth are added to the original smoothed values. This decreases imputation bias but also increases imputation variability.

The procedure used was to first calculate the local linear trend in the steering wind direction data from the observed radiosonde values. This trend then defined first stage imputed wind direction values for those days with missing radiosonde values. Second, local linear deviations from this local trend, defined as the differences between actual radiosonde values and trend values, i.e. based on the observed data, were used to create a second local trend by a further local linear smoothing of these residuals. This second stage smooth was next used to predict the errors in the first stage smooth for the days with missing radiosonde data. This second stage residual smooth was finally added to the first stage imputed smooth to produce a “twiced” smooth, which defined imputed values for the missing radiosonde wind direction data.

Both the first stage trend and the second stage residual smooth are computed using a local linear regression method called LOESS (Cleveland and Devlin,1988). The method involves calculating a fitted value for any point of interest in the support of the sample data based on a low-degree polynomial fit to a subset of the sample data that is "local" to the point of interest. The polynomial fit is carried out using weighted least squares, giving more weight to sample points near the point of interest and less weight to sample points that are further away. The fitted value at the point of interest is then obtained by evaluating the local polynomial at the explanatory variable values for that point. The subset of data used for each weighted least squares fit in LOESS is determined by a nearest neighbors algorithm. A user-specified input to the procedure called the "smoothing parameter" determines how much of the data is used to fit each local polynomial. Judicious choice of this smoothing parameter ensured that the fitted trend to the original data was smoothed more than the fitted trend to the residuals. Also, the smoothing procedure was applied to the u and v transforms of the observed radiosonde data, with the final imputed values for these transforms then converted back to wind directions. The twiced smoothing process for the u transform of 2016 steering wind direction is shown in Figure C - 1 and Figure C - 2.

**Figure C-1 The two LOESS smooths of the u transform of steering wind direction used in the imputation of missing values for these wind directions. The first stage smooth is of the actual values for u, while the second stage smooth is of the residuals from the first stage smooth.**

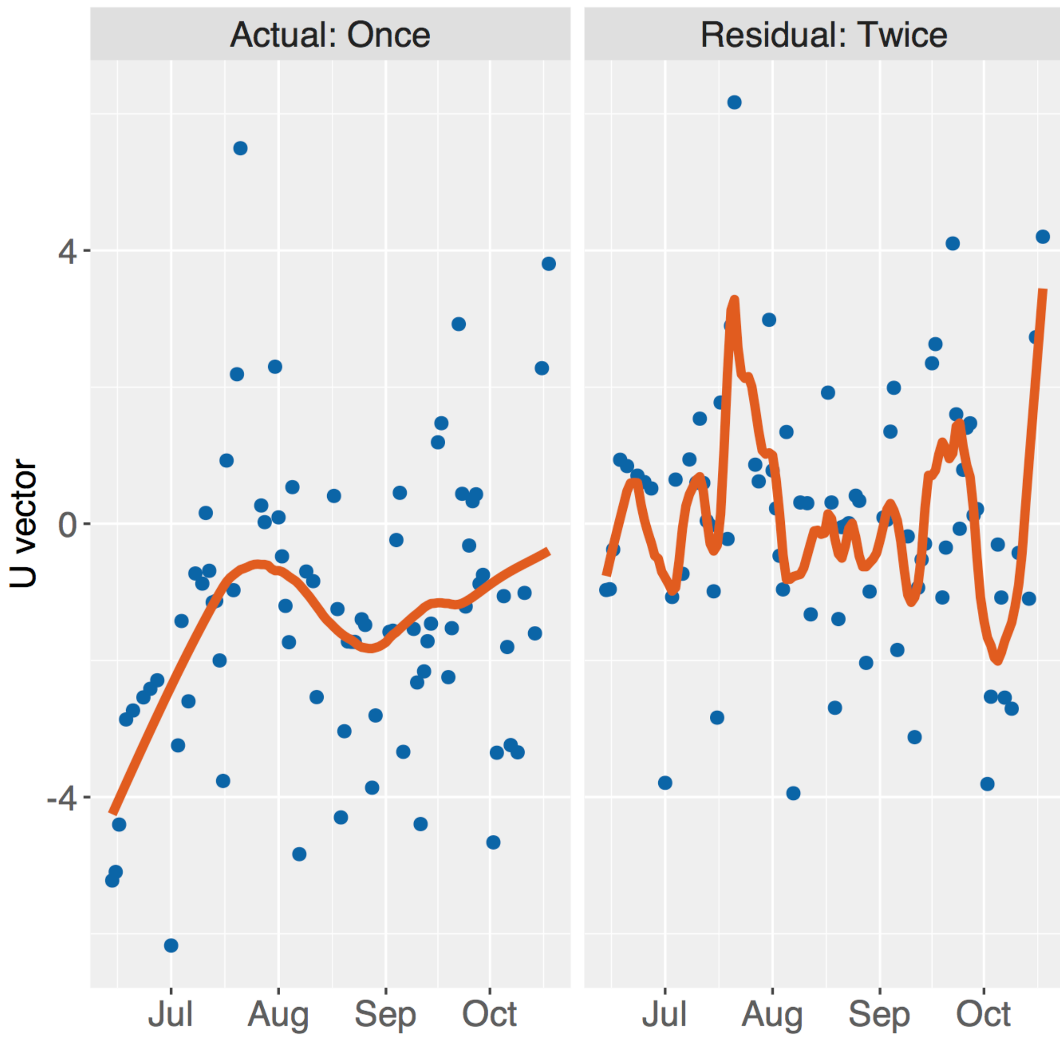
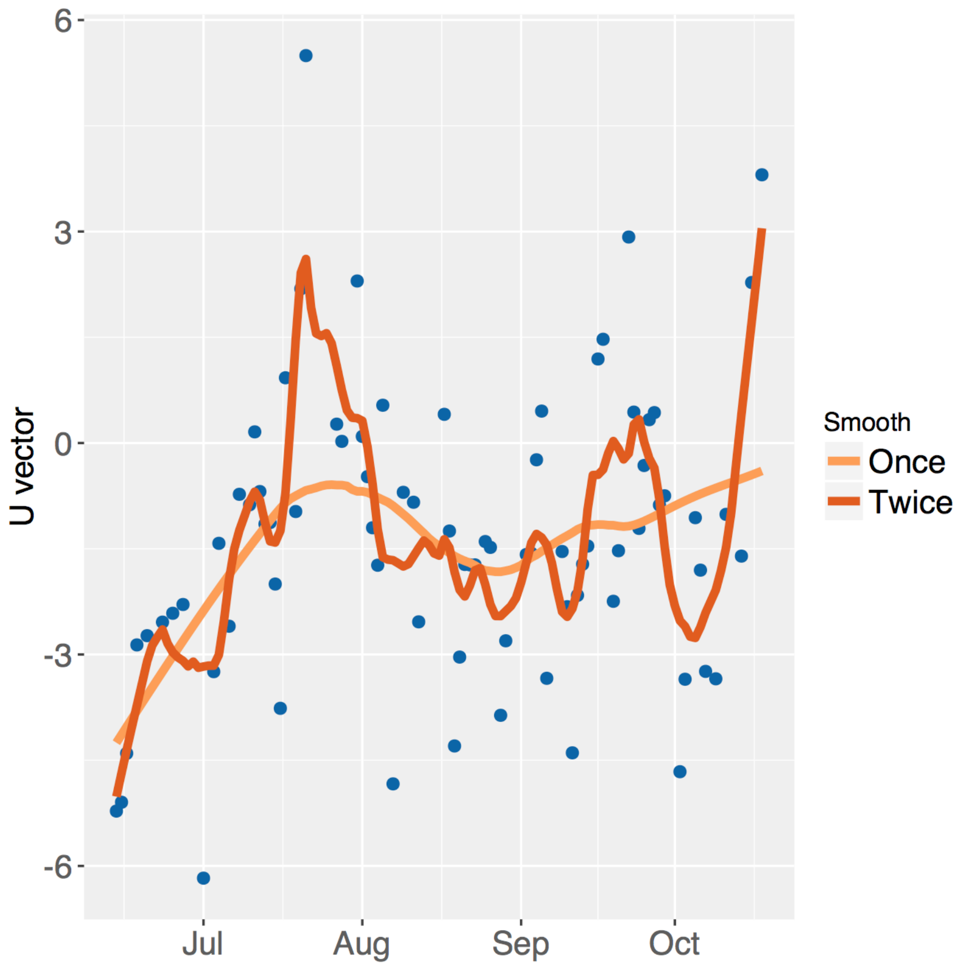


**Figure C-2 Once and twiced LOESS smooths of the 2016 values of the u transform of steering wind direction. The twiced smooth is used in the imputation of missing values for these wind directions.**



**A Sensitivity Analysis of the Imputation Method Using 2015 Trial Data**

In order to assess the appropriateness of using imputed values for missing wind direction in the analysis of the 2016 trial data a sensitivity analysis was carried out to see how a similar imputation exercise would have impacted on the 2015 trial results.

The data available for 2015 ranged from the 14th of June to the 18th of October, corresponding to a total of 127 days (referred to as in-scope days below). Data from the 2016 trial for these in-scope days was then used to ensure that the sensitivity analysis was based on the same days in each year.

In-scope days from the 2016 and 2015 data sets were ranked and ordered based on the average daily rainfall and also on daily rainfall propensity. An indicator variable, which specified whether radiosonde data was available for in-scope days in 2016, was also created, and ordered similarly to the 2016 rank ordering of interest. A corresponding missing data indicator variable for 2015 was then defined by matching the daily data from both years on the basis of their respective rank orders as defined by either average daily rainfall or daily rainfall propensity. The ordered indicator variable for 2016 missingness was then copied to the rank ordered 2015 data set, and wind directions for those days that were indicated as missing according to this newly created 2015 indicator variable were imputed. For example, if radiosonde data was missing on the 5th highest rainfall day in 2016, then it was assumed missing for the 5th highest rainfall day in 2015 and wind direction was imputed for that day. In the case of the rainfall propensity rank ordering, if the day with the 10th highest proportion of gauges reporting rainfall was a day when wind direction information was missing in 2016, then the day with the 10th highest proportion of gauges reporting rainfall in 2015 was assumed to have missing radiosonde data in 2015 and wind direction was imputed for that day.

We note that this somewhat complicated way of generating missing radiosonde data days in 2015 was necessary because simple matching of days between 2016 and 2015 would have led to missing radiosonde data on days in 2015 that were not comparable (in terms of rainfall) with days with missing radiosonde data in 2016.

Imputed wind direction and wind speed for these artificially generated missing radiosonde data days were finally calculated via the two-stage smoothing method described above, using the data from the remaining (non-missing data) days. As outlined previously those days with artificially missing data that were part of a sequence of four or more days with missing data were considered not imputable in this exercise. For the 2015 data with missingness based on average daily rainfall, four days out of 127 were not included, while with missingness based on the daily rainfall propensity, nine days out of 127 were not included. Table 5-1 compares the final bootstrap distributions of estimated attribution in 2015 based on the actual 2015 data set with the same distributions where attribution is estimated using the 2015 data set with imputed radiosonde wind directions based on the pattern of missingness in 2016 ordered by average rainfall and daily rainfall propensity. It can be seen that the impact of imputation in both these cases is not excessive.

## References

Cleveland, W.S. and Susan J. D. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the Am. Statist. Assoc.* 83, 596-610.

Svetunkov, I. (2017). smooth: Forecasting Using Smoothing Functions. R package version 1.7.0. https://CRAN.R-project.org/package=smooth

Tukey, John W. (1977). *Exploratory Data Analysis*. Pearson. ISBN 978-0201076165.