University of Wollongong

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

TRADING & INVESTMENT ESTABLISHMENT

Australian Rain Technologies

Statistical Modelling for a Randomised Trial of a Rainfall Modification Technology

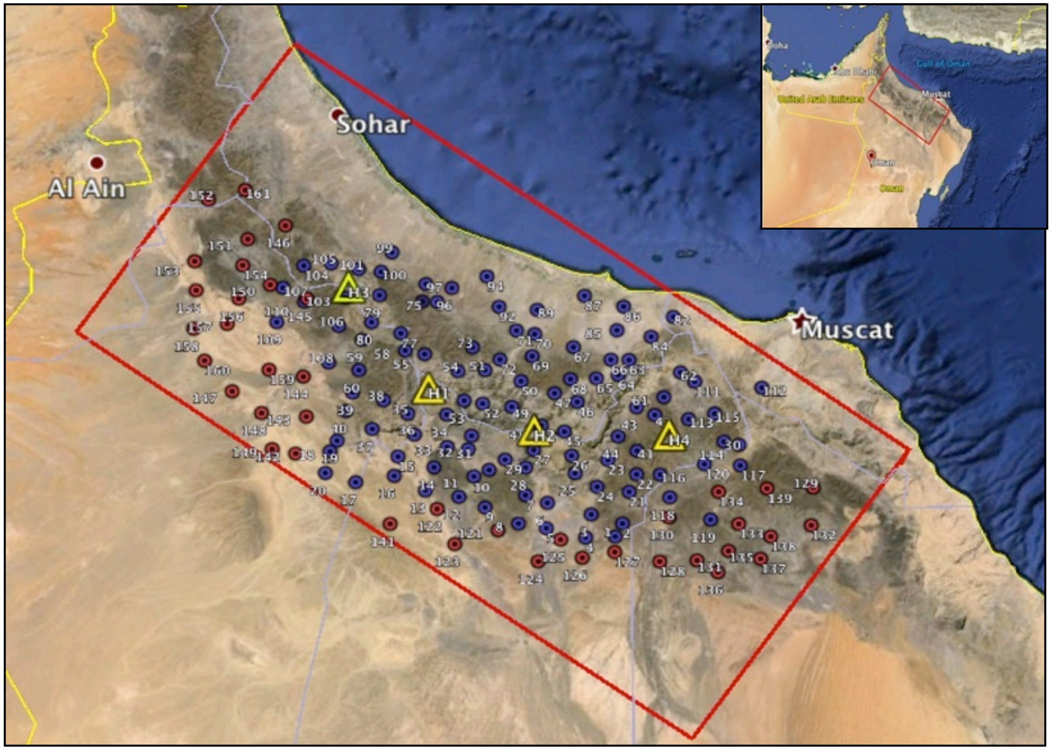
Ray Chambers
Stephen Beare
Scott Peak
Mohammed Al-Kabani

+ Introduction

- Spatio-temporal analysis of data from summer rainfall enhancement trials by TIE/ART in the Hajar mountains of Oman, currently entering its fourth year.
- Ground based ionisation system know as Atlant:
  - Operates in a similar way to ground based silver iodide seeding.
  - Hypothesised to facilitate raindrop coalescence as opposed to providing seed nuclei.
- Experimental design: randomised crossover experiment with dynamically (wind) defined target and control areas.
- Main statistical problem – maximising power.
  - Temporal and spatial aggregation reduces the high level of spatial and temporal heterogeneity but with a significant loss of power.
  - Using gauge level data gives greater power but need to account for the lack of independence in gauge observations due to spatial and temporal correlation.
  - Correlation is likely to be anisotropic with the direction of greatest correlation at any point in time being being downwind.
- Analysis - 2014 data from two sites with best instrumentation.

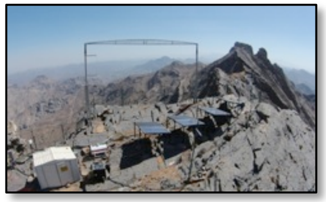
+ Scope of the 2014 Trial

- Four Atlant systems, H1 and H2 in the centre of the trial area, H3 in in the northwest and H4 in the southeast.
- 149 rain gauges installed over the trial area, made up of 119 gauges used in the 2013 trial (blue) + 30 new gauges (red).

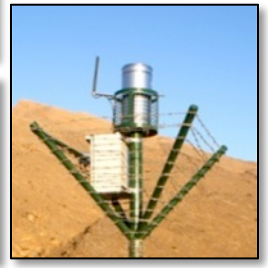


2 x SODAR

11 x AWS



4 x Atlants



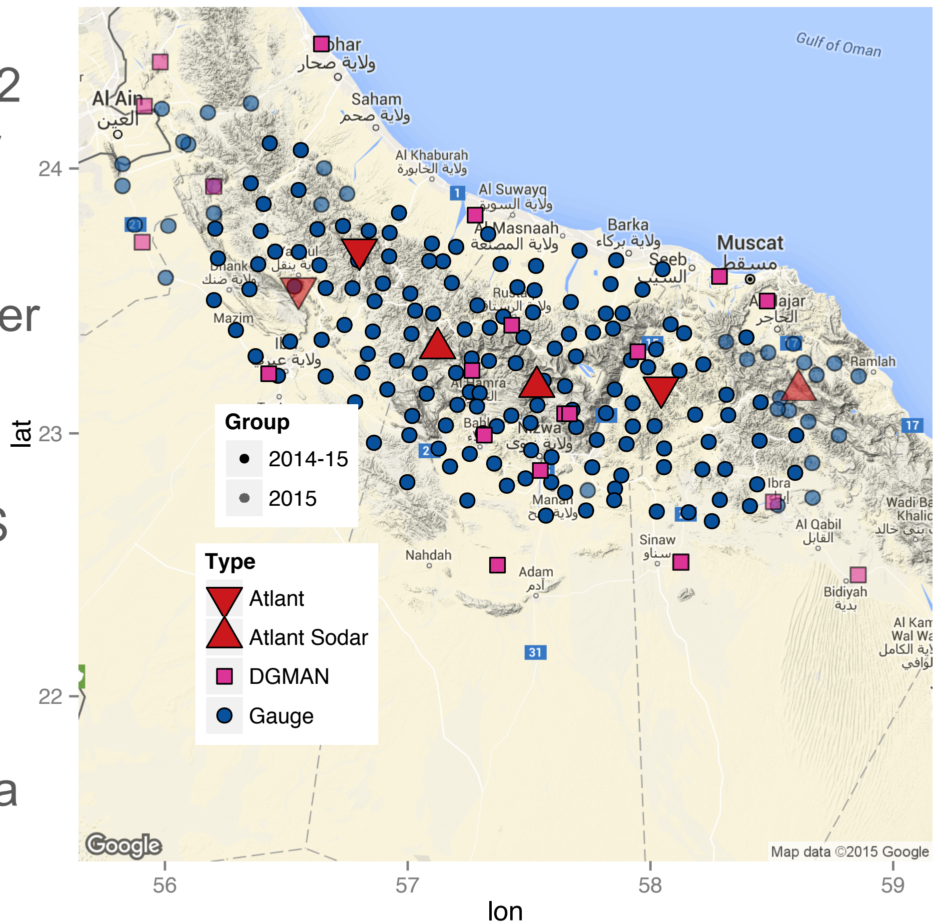
149 x Gauges

+ Trial Instrumentation

- Vertical sonic wind profilers (SODAR) installed at H1 and H2 in 2014 in order to gain a better understanding of the how the topography of the Hajar Mountains affects uplift and other atmospheric conditions.

- Data from TIE instruments are supplemented by DGMAN AWS data.

- Two more Atlants (H5 & H6) installed for 2015 trial, together with another SODAR (H5) and a further 32 gauges.



2014 Trial Design

- The trial was conducted over the 140 days 1 June 2014 - 18 October 2014

- The trial used a randomised cross-over design with:
  - The operational schedule set prior to the start of the trial;
  - Independent randomisation within 4 week blocks to balance against seasonal weather trends during the trial;
  - Basic statistical methodology prescribed prior to the start of the analysis.

- The four sites were operated in a pairwise on/off sequence to give maximum separation between active and control areas, with:
  - H1 paired with H3, and H2 paired with H4.

+ Analysis of the 2014 Trial Data

- Statistical models were used to control for the natural variability in rainfall over what can be achieved purely through the randomised experimental design.

- These models were also used to make a counterfactual prediction of the "natural" rainfall that would have been measured by a gauge located in the downwind footprint of an operational Atlant if it had ***not*** been operational.
  - The differences between observed rainfall and these predictions were used as estimates of the change in rainfall (+ve or -ve) caused by Atlant operation.

+ The Statistical Modelling Process

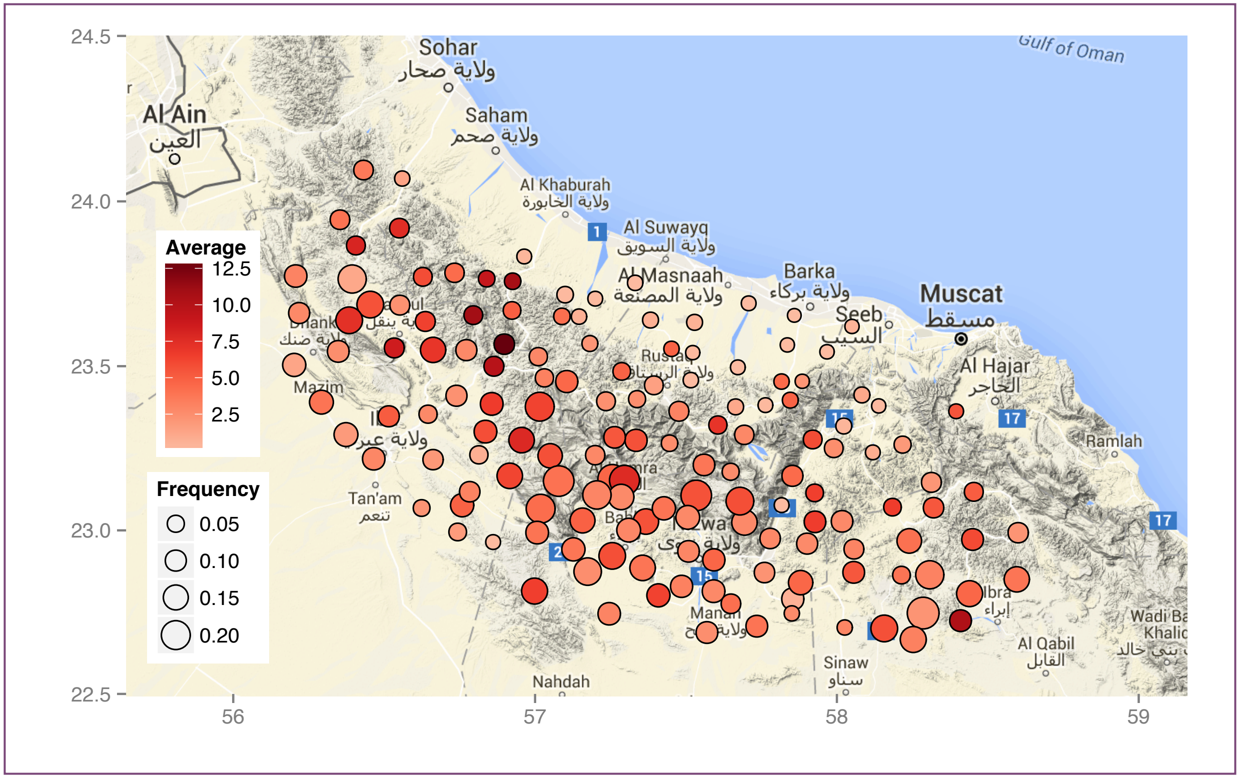
- The focus is on whether there is increased rain when rain occurs, so positive rainfall values are modelled. These are highly right-skewed, so model is for LogRain = log(rainfall).
  - **Stage 1:** Construction of an instrumental variable that correlates with the expected amount of natural rainfall at a gauge on a day.
    - Defined by fitted values for LogRain generated by data from gauges that are upwind of the target and control areas on a day.
    - Covariates used in this model include gauge elevation, steering wind direction and speed (based on 04:00 radiosonde measurements made at Seeb), data collected from DGMAN AWS sites and indexes of rainfall propensity defined by Seeb radiosonde data.
  - **Stage 2:** Modelling LogRain using the rainfall data in the downwind (i.e. target and control) areas.
    - This model includes the instrumental variable values generated in Stage 1, gauge elevation, meteorological conditions at the Atlant sites (based on SODAR data) and binary indicators for target status.
    - It also included a random day effect for potential omitted variable misspecification bias (there was no evidence of a significant random gauge effect).

+ The Statistical Modelling Process

- **Stage 3:** Estimation of the amount of increase (or decrease) in downwind rainfall given Atlant operation.
  - A logistic regression model is first fitted to the binary variable indicating whether rainfall occurs at a gauge or not on a day. This model is used to simulate rainfall incidence for all downwind gauges on the day.
  - The LogRain model fitted at Stage 2 is used to simulate the amount of rainfall recorded for the gauge-days that are simulated as recording rainfall in previous step.
  - To preserve the original correlation between gauges in the downwind areas each day, simulated LogRain (and rainfall) values are randomly drawn only from gauges that are also downwind that day. That is, observations that are nearby in space and time are grouped into blocks.
  - The LogRain model is refitted to each bootstrap rainfall simulation, and log scale fitted values calculated. These are back-transformed (with bias correction) to generate a population of alternative realistic rainfall values which are then used to estimate the amount of natural rainfall, and, by comparison with the simulated actual rainfall, the estimated amount of increase (or decrease) in rainfall due to operation of the Atlant.

+ Rainfall Distribution

- Distribution of rainfall was relatively even through most the trial area. The exception was in the northeast where there were a number of large rainfall events that occurred on isolated days.

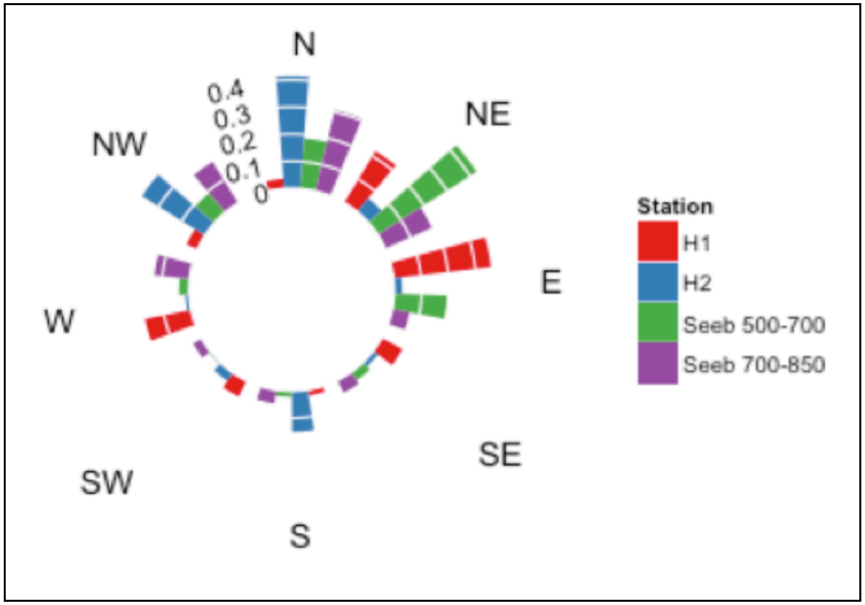


Bubble plot of the locations of the TIE rain gauges, with gauge rainfall frequency proportional to bubble size and gauge average actual rainfall (i.e. excluding zeros) denoted by the colour scale

+ 2014 Wind Direction Profile at H1, H2

Muscat International Airport

- The prevailing winds are from the northwest to the east, with the majority from the north to the northeast.
- The orographic effects of the Hajar ranges are clear at H1 and H2 as well as at the lower elevation (700hPa to 850hPa) wind directions above Seeb (now Muscat International Airport).
- When compared to the upper elevation wind directions at Seeb (500hPa to 700hPa), winds at H1 have clocked to the east, while winds at H2 and at lower elevations above Seeb have clocked to the west.
- Graphic shows the distribution of daily upper level wind directions at Seeb, H1 and H2. Each colour sums to one around the eight points of the compass rose.



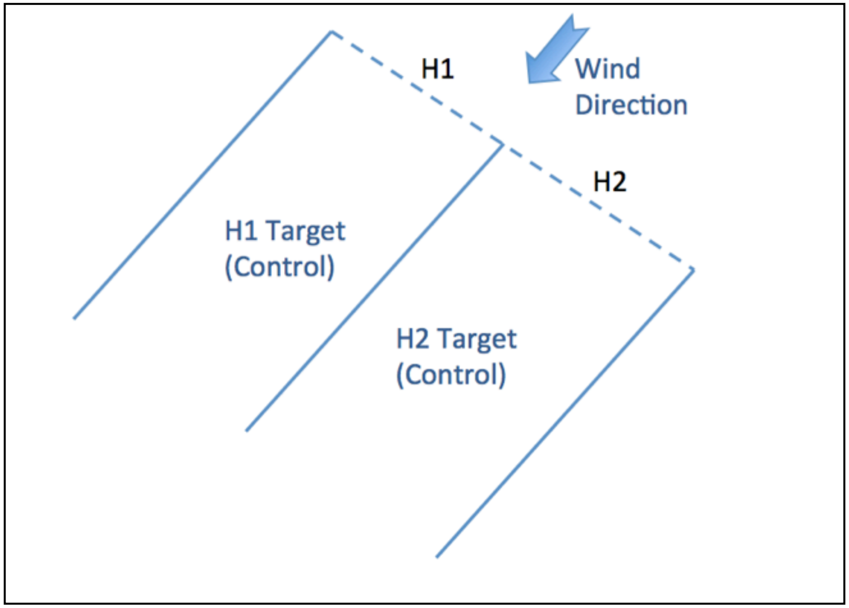
The distribution of daily upper level wind directions at Seeb airport, H1 and H2: speed weighted averages.

Height of the Boundary Layer

- Airflows above the boundary layer are generally non-turbulent and consistent over a region the size of the trial area. The cloud layer tends to form just above the boundary layer in the free atmosphere.
- These different wind flows affect an ion plume and therefore the location of the "footprint" of the effect
- The height of the boundary layer in the afternoon can be measured by the height at which radio waves are no longer reflected back to the SODAR.
- This varies from day to day but the average height of the boundary layer above:
  - H1 is 340 metres, and
  - H2 is 510 metres.
- Carried by the average afternoon vertical wind speeds an ion plume would reach the top of the boundary layer in:
  - Less than 5 minutes at H1, and
  - About 20 minutes at H2.

+ Defining the Downwind Footprint

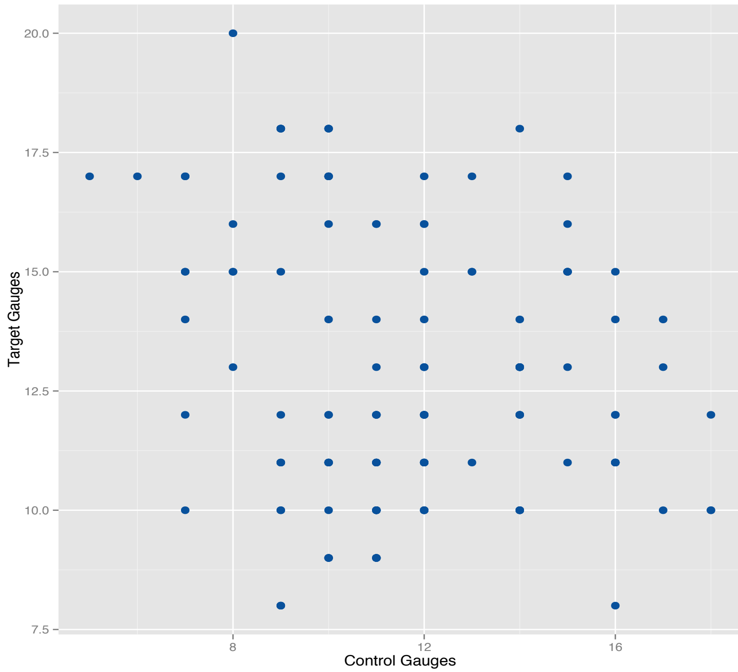
- Distribution of the ion plume should be determined by the upper level wind directions in the free atmosphere.
  - These are measured on a daily basis at Seeb.
- In the average time it would take an ion plume to reach the free atmosphere, given the average horizontal wind speeds above the site, the plume would drift laterally by:
  - 1.2 km at H1 , and
  - 4.8 km at H2.
- Footprint model assumes that the plume stays within a 30 km downwind corridor.



30km dynamic downwind corridor footprint model

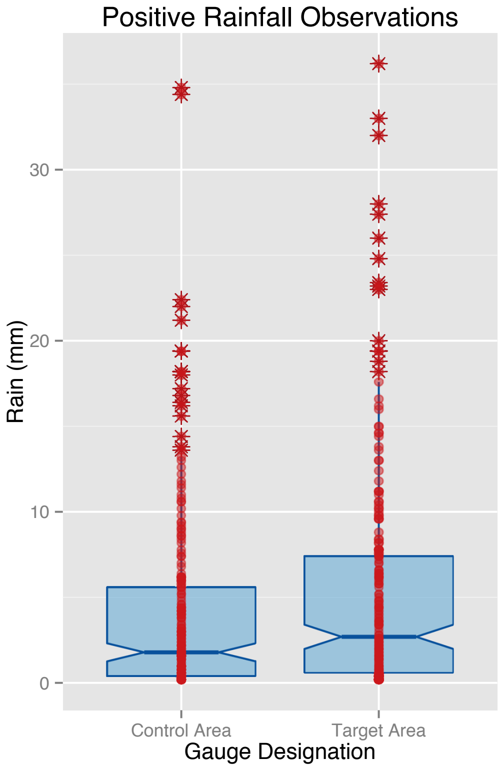
+ Rainfall Downwind of H1, H2 - 2014

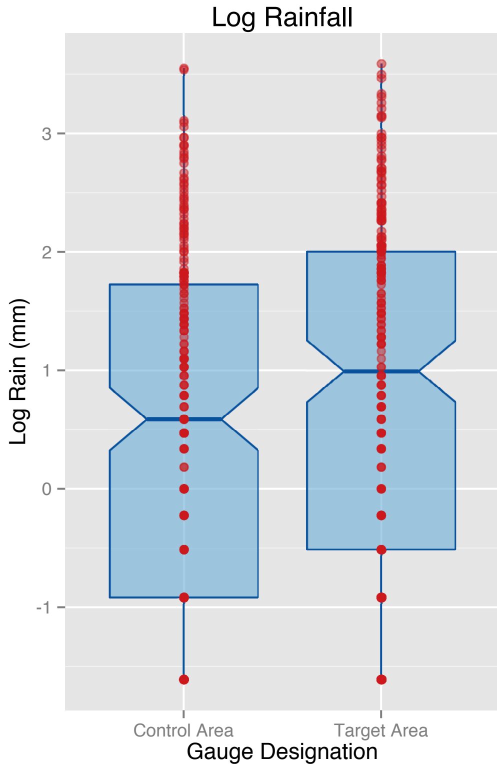
- Daily gauge numbers small and varied significantly between target and control areas.
  - Target and control gauges also not balanced on a day to day basis.
- Target areas recorded non-zero rainfall 50% of the time, compared with 59% of control areas (232 vs. 247 positive gauge readings).
  - But average gauge-level daily rainfall for target vs. control gauges was 0.71 mm vs. 0.62 mm (5.41 vs. 4.08 if only positive readings averaged).



Log Scale Modelling of Positive Downwind Rainfall

- Boxplot on left shows positive rain values highly right-skewed, so modelled on a log scale (LogRain).
- Box plot on right shows control and target area values for LogRain.
- Comparing LogRain values, average target (control) area value = 0.81 (0.52), 2-sided p-value for raw difference = 0.0308.





Fitted Model for LogRain

Term	Estimate	Std Error	2-sided p
Intercept	0.7952	0.2093	0.0002
Target status	0.2676	0.1268	0.0354
Expected Upwind Rainfall	0.7597	0.1474	<.0001
H1 SODAR U	-0.1724	0.0577	0.0040
H2 SODAR V	0.1854	0.0688	0.0090
Gauge Elevation	-0.0007	0.0002	0.0006

- Target status = binary indicator (1 = target reading, 0 = control reading).
- Expected Upwind Rainfall = instrumental variable characterising expected daily rainfall levels upwind of H1 and H2.
  - Defined by fitted values for gauge-level upwind LogRain, based on meteorological indicators on the day + average readings from DGMAN weather stations.
- H1 SODAR U = 50m-400m average U-transform of 10:00-20:00 wind directions recorded by H1 SODAR.
- H2 SODAR V = 50m-400m average V-transform of 10:00-20:00 wind directions recorded by H2 SODAR.
- Random day effect explains just over 11% of total residual variation in fitted model and just under 21% of total variation in corresponding null model (no covariates).

+ How Much Extra Rain?

- 2263 mm of rain recorded downwind of H1 and H2 during the 2014 trial.
- Hurdle-type Random Effect block bootstrap estimates (std errors) of "natural" and "extra" rainfall (mm) components of this total are 1866 (282) and 397 (122).
- Corresponds to estimated 21.7% extra rainfall relative to natural downwind rainfall.
- Graphic shows bootstrap distribution of estimated extra rainfall:
  - Est Pr(+ve extra rainfall) = 98%
  - Est Pr(at least 3% extra rainfall) = 98%
  - Est Pr(at least 10% extra rainfall) = 80%
- Similar estimates obtained when all 4 (H1 - H4) downwind areas are analysed, and also when rainfall data from H1 and H2 downwind areas for 2013 and 2014 are combined.