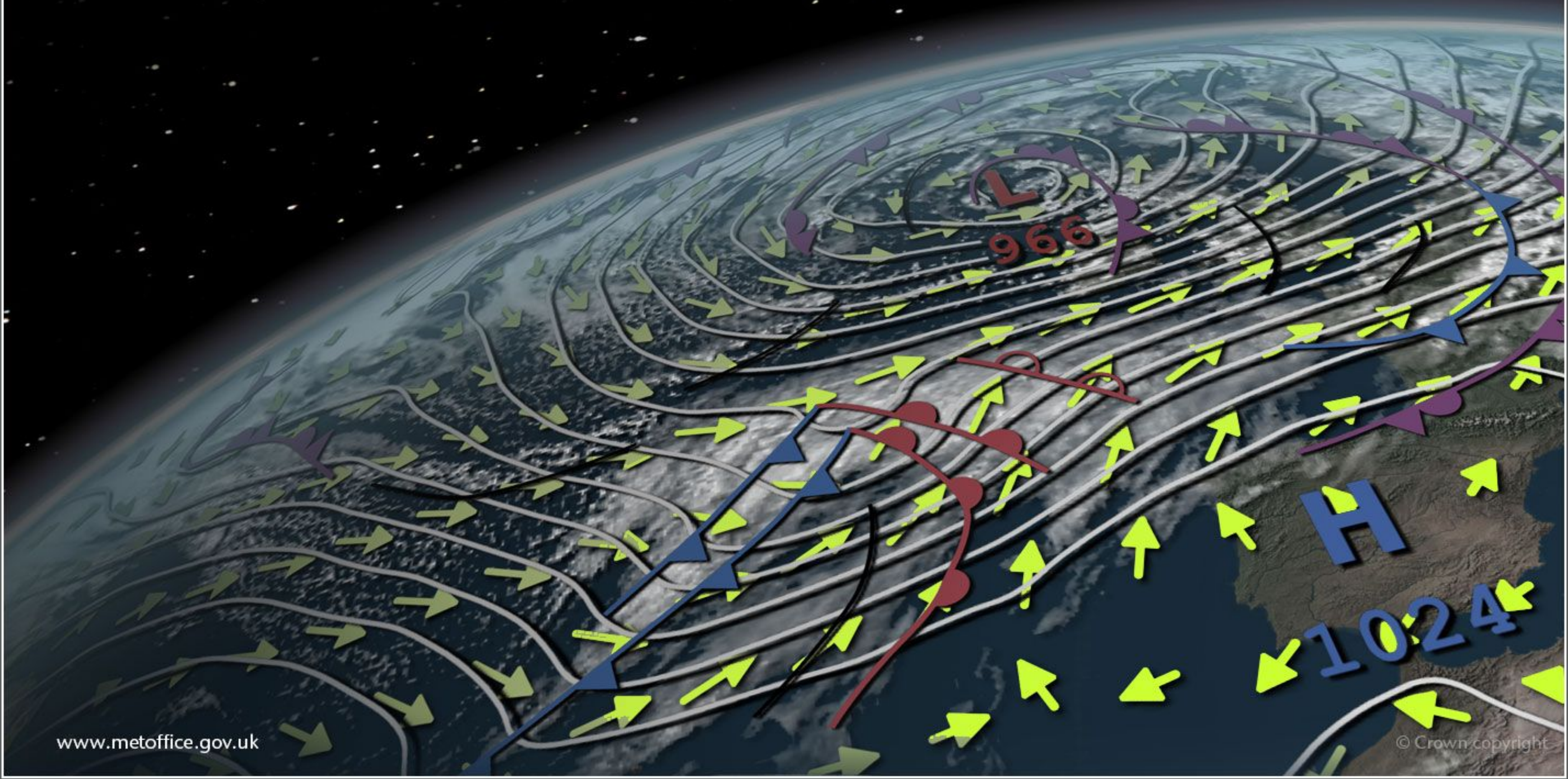
European Windstorm Event Response Service

From real time forecasts to historical analogues

Bernd Becker, Met Office, 1st September 2015



Collaborators

Applied climate science
Insurance and Capital Markets team: Paul Maisey, Claire Scannell, Hamish Steptoe, Lorna Mitchell
Building on the Unified Model
MOGREPS ensemble 7-day forecast (33km)
European 5-day forecast
Historical Windstorm Catalogue
At 4 km resolution

Met Office

Contents

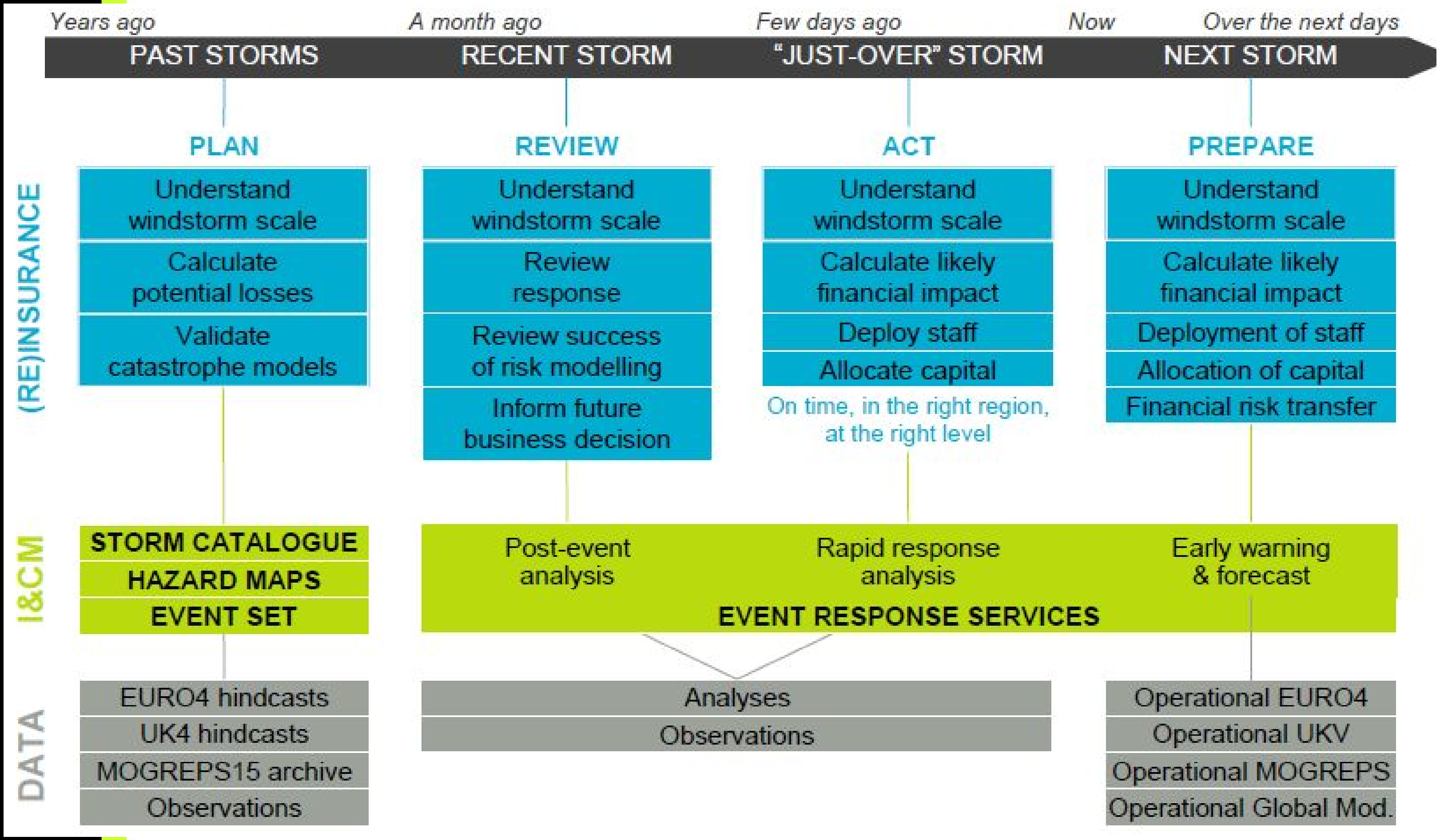
Met Office

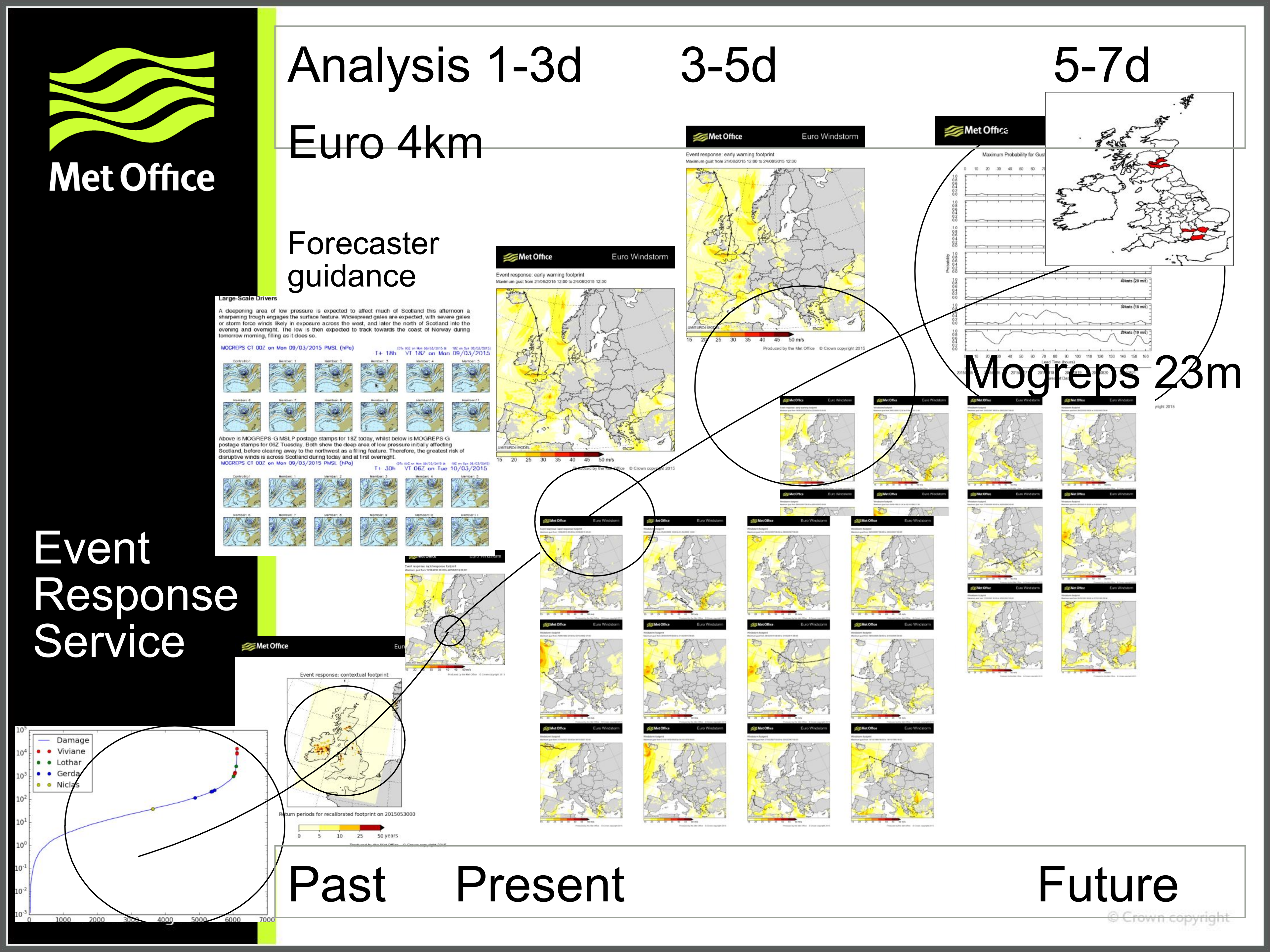
Introduction

What do we do?

Met Office

Products & Activities




Met Office
Event Response Service
Analysis 1-3d
3-5d
5-7d
Euro 4km
Forecaster guidance
Mogreps 23m
Past
Present
Future
Large-Scale Drivers
A deepening area of low pressure is expected to affect much of Scotland this afternoon a sharpening trough engages the surface feature. Widespread gales are expected, with severe gales or storm force winds likely in exposure across the west, and later the north of Scotland into the evening and overnight. The low is then expected to track towards the coast of Norway during tomorrow morning, filling as it does so.
MOGREPS CT 00Z on Mon 09/03/2015 PMSL (hPa)
Above is MOGREPS-G MSLP postage stamps for 18Z today, whilst below is MOGREPS-G postage stamps for 06Z Tuesday. Both show the deep area of low pressure initially affecting Scotland, before clearing away to the northwest as a filling feature. Therefore, the greatest risk of disruptive winds is across Scotland during today and at first overnight.
MOGREPS CT 00Z on Mon 09/03/2015 PMSL (hPa)
Euro Windstorm
Event response: early warning footprint
Maximum gust from 21/08/2015 12:00 to 24/08/2015 12:00
UM/EURO4 MODEL
15 20 25 30 35 40 45 50 m/s
Produced by the Met Office © Crown copyright 2015
Event response: contextual footprint
Return periods for recalibrated footprint on 2015053000
0 5 10 25 50 years
Damage
Viviane
Lothar
Gerda
Niclas
0 1000 2000 3000 4000 5000 6000 7000
Maximum Probability for Gust
40knts (20 m/s)
30knts (15 m/s)
20knts (10 m/s)
Lead Time (hours)
Probability
© Crown copyright

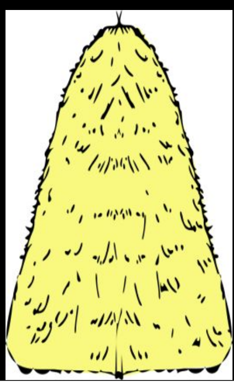
Historical Windstorm Catalogue (72 hours at 4 km)

Commercially available
- for both commercial and academic use

Contains information on storm tracks, surface winds and windstorm footprints for storms over the past 35 years (6110 storms)

Facilitates research into storm characteristics and the influence of large-scale atmospheric variability on European windstorms

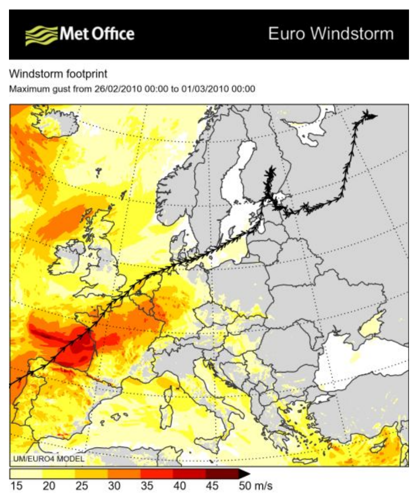
Provides a common performance benchmark for both catastrophe models and climate models

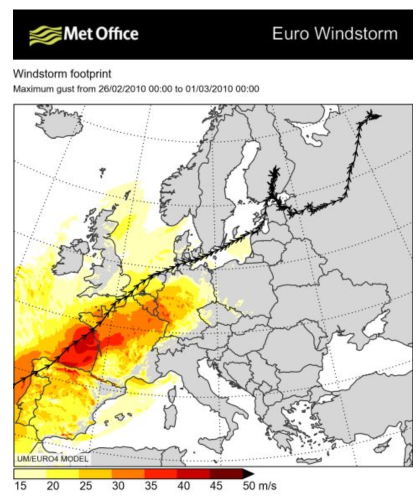


Haystack

Full field maximum wind gust:

Xynthia
27 Feb 2010



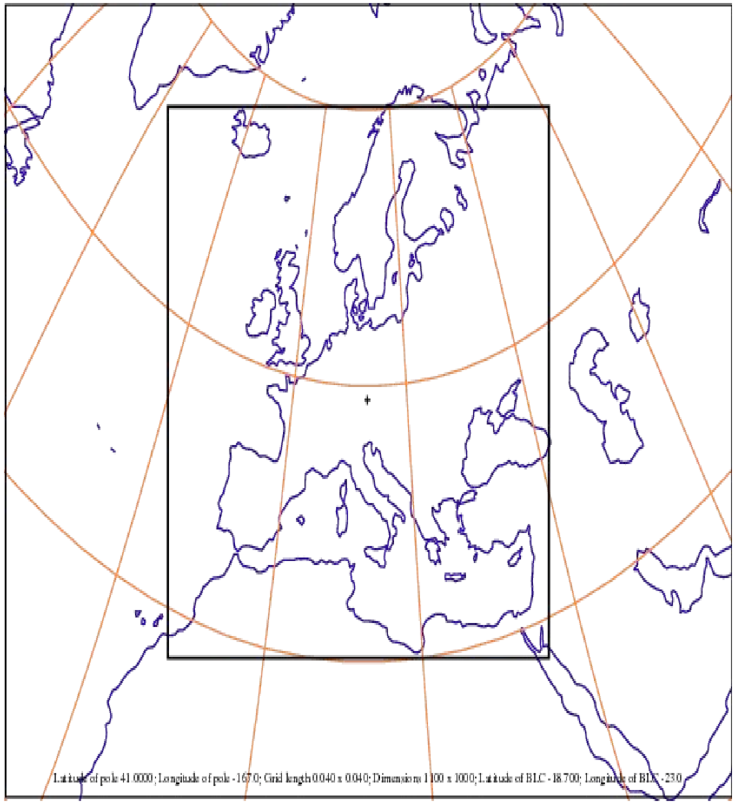


Foot print:
Decontaminated
1000 km
radius
around track

The real time forecast system

- State of the art, world leading, operational, high resolution
- EURO4: European 4.4km Global down-scaling model

  enlarged domain to
  cover most of Europe
  driven by 3-hourly
  Global boundary conditions
  starting from the
  Global T+0 analysis.



The operational storm footprint (SFP) monitor

Post-processing the latest EURO4 run every 6 hours

Concatenate short range forecast period t+3 to t+9 from consecutive forecasts to describe a 72 hour time series in the recent past

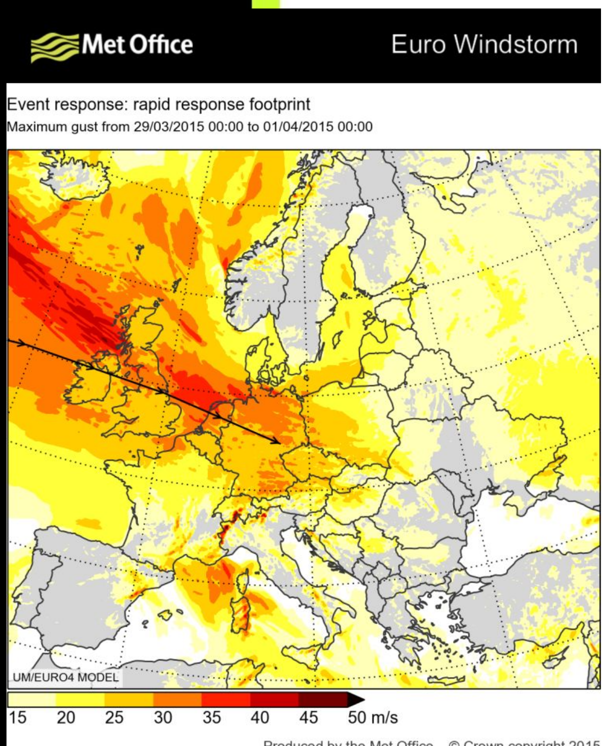
Event Response Rapid Response: past 72 hours

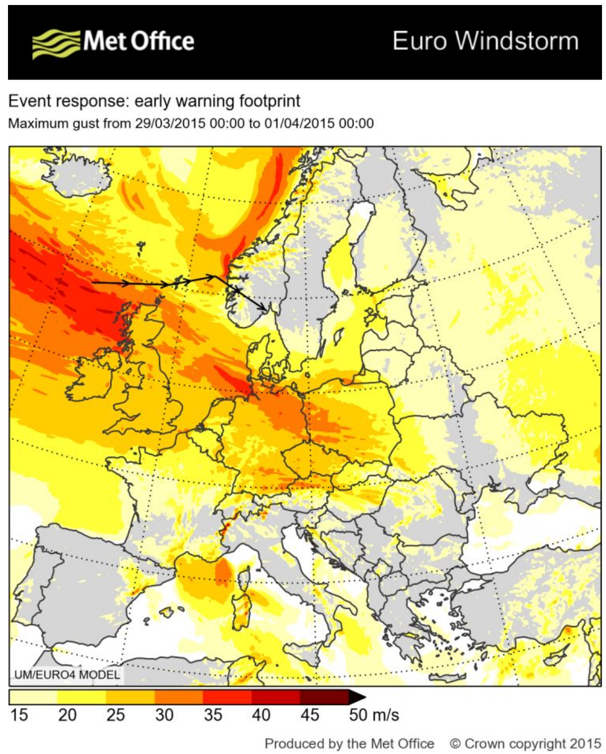
Complement with:

Event Response Early Warning: 1-3-5 day forecast

The operational storm footprint (SFP) monitor provides:

Early Warning





Rapid response

A succinct description of the latest SFP close to real time.

May provide near real time forcing data for CAT models.

May be extended into longer lead times

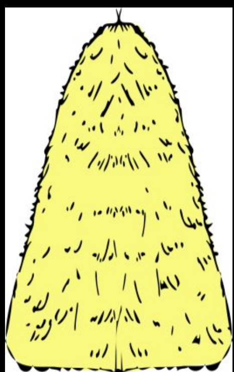
(images, gridded data, geotiffs, csv)

(Niklas)

Needle

Met Office

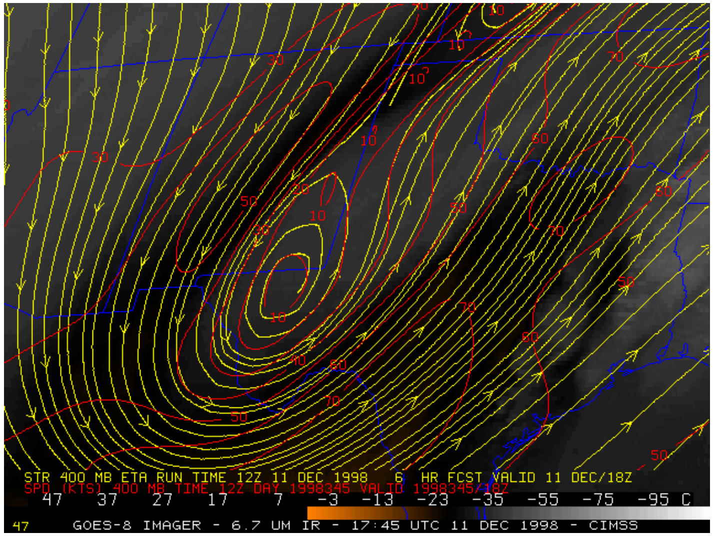
“Matching” challenge



To put current extreme wind events into perspective:
Find real time storm footprint (SFP) data (the needle)
In a catalogue of past events (in the haystack)
Provide re/insurers with historical wind storm data similar to the current/most recent event
Known impacts of historical storms in (user) terms of damage, claims, loss, costs
Guide customers during the planning stages in mitigating the impact of the most recent event.
May provide forcing data for CATastrophe models

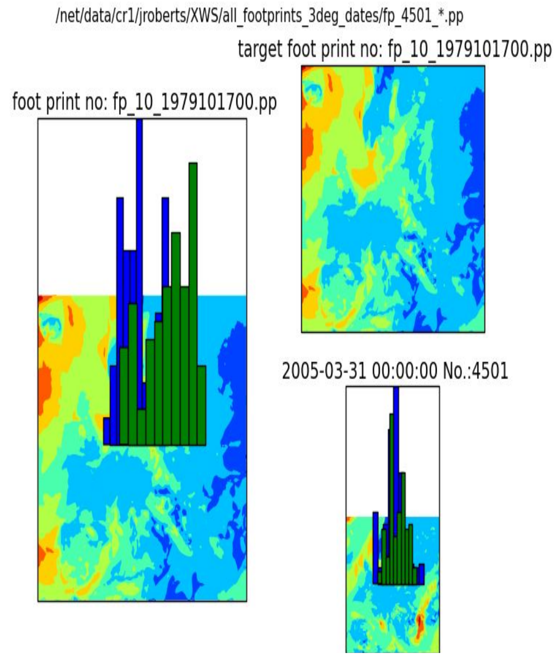
Met Office

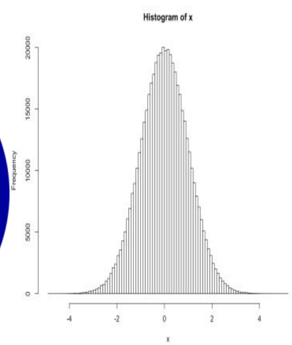
Idea: compare fingerprint and streamlines



Fingerprint reading software now commonplace, on every phone.
What software is available for computer vision and image processing, searching for similarities and feature tracking?

Matching storm footprints: blind alleys...





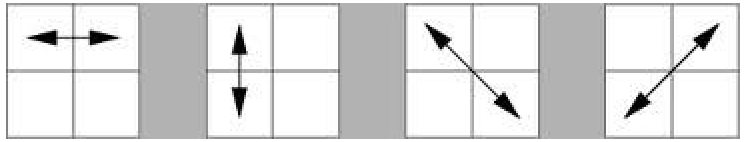
Using a single feature to describe a SFP:

- mean variance (longitude) * mean variance (latitude)
- Draw one histogram from a north-south oriented strip of data and another from an east-west oriented strip of data. Calculate Wilcoxon rank sum statistic to describe the departure from “blobbiness” (here we are dipping our toe into feature analysis!)

Insufficient discrimination: similar numbers describe dissimilar SFP

Try computer vision for indexing database of SFP

Haralick texture features:

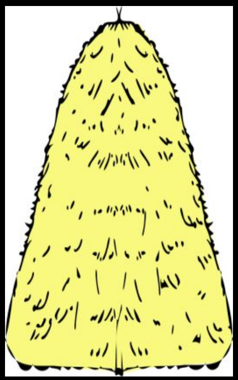


- Derived from the Gray-Level-Co-occurrence matrix
- Defined over an image, to be the distribution of co-occurring values at a given offset which
- Contains information about how image intensities in pixels with a certain position in relation to each other occur together.
- Mahotas computer image processing library for Python calculates thirteen features for each SFP

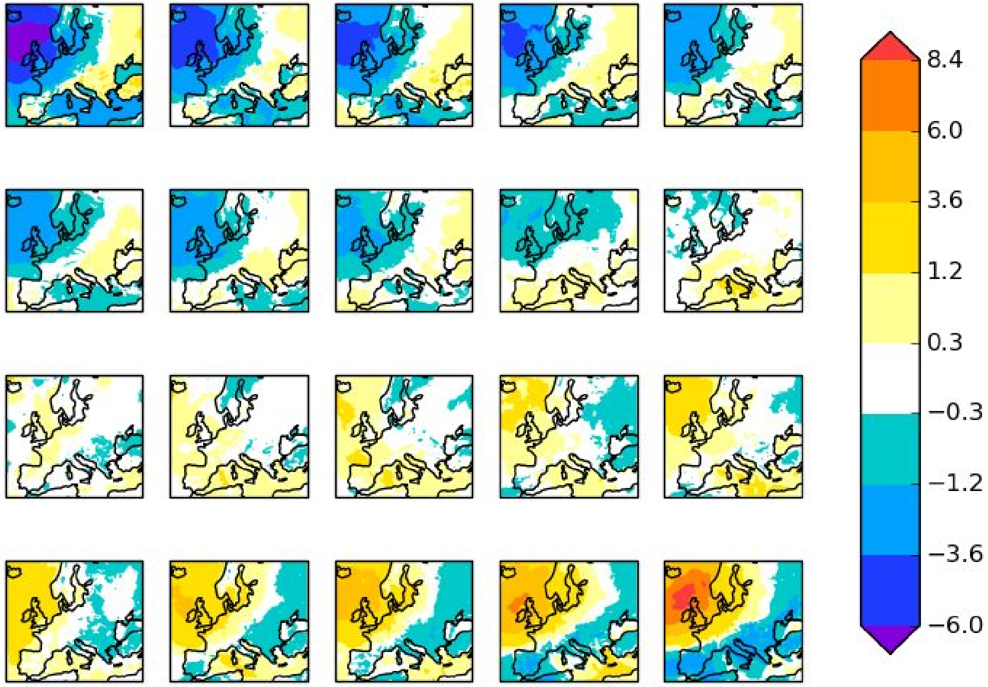
H1: Angular Second Moment
H2: Contrast
H3: Correlation
H4: Sum of Squares: Variation
H5: Inverse Difference Moment
H6: Sum Average
H7: Sum Variance
H8: Sum Entropy
H9: Entropy
H10: Difference Variance
H11: Difference Entropy
H12: Information Measure of Correlation 1
H13: Information Measure of Correlation 2

Robert M Haralick, K Shanmugam, Its'hak Dinstein (1973). "Textural Features for Image Classification". IEEE Transactions on Systems, Man, and Cybernetics. SMC-3 (6): 610–621.

Met Office



Sufficient discrimination



Ranked mean Haralick image texture features for each of the 6110 storm foot prints:

Plotted averages of 20 chunks at 305 storms blurs the characteristics of individual events but the overall change in the pattern is promising

Sign of the potential of the method to finding similar features between recently observed and historically reported European windstorms. (or helps us to find the needle in the haystack)

Match decision tree, machine learning

Learn closest similarity by ranking a list of similarity measures and by choice of closest members.
For each storm footprint (SFP), save:
SFP number
Date
Longitude (average from U > 25 m/s mask)
Latitude (average from U > 25 m/s mask)
Storm intensity over SFP area (3)
13 Haralick features (*4)
Expandable to many more storm features!
Add current storm to catalogue and find best possible match (twin)
Note nearest neighbours (rank retrieval mode)
Keep a number of similar candidates to evaluate and plot
Calculate (costly!) distance measures to confirm closest candidate

Met Office

www.metoffice.gov.uk

Evaluation of close matches

In each similarity measure of Storm footprints (SFP) we:

- Note the distance in rank for each measure
- Pick the top N best matches
- Calculate the distance ||F|| between candidate and group members
- Pick a winner

1. Rank proximity measure:
   Count positional distance in rank from perfect match in each category, integrate over all categories and declare the winner with the smallest rank proximity measure

Met Office

Met Office

Best Match RPM : 1751, 427, 86, 2586, 404, 1954, 4203, 2777, 5105, 3174,

6 8 13 14 15 21 24 24 25 27

Feature:	1:		2:		3:		4:	4203	5:
								1954	
		2777						3174	
2586	-5	3174	-5		-5		-5	404	-5
86	-4	5105	-4		-4		-4	5105	-4
404	-3	1954	-3		-3		-3	2586	-3
1954	-2	4203	-2	1751	-2	86	-2	2777	-2
1751	-1	1751	-1	86	-1	427	-1	1751	-1
Candi	0	1634	0	1634	0	1634	0	1634	0
4203	1	404	1	427	1	1751	1	427	1
427	2	2586	2	2586	2	2586	2	86	2
3174	3	427	3	404	3	404	3		3
5105	4	86	4	1954	4	2777	4		4
2777	5		5	4203	5	1954	5		5
				3174		5105			
				5105		3174			
				2777		4203			

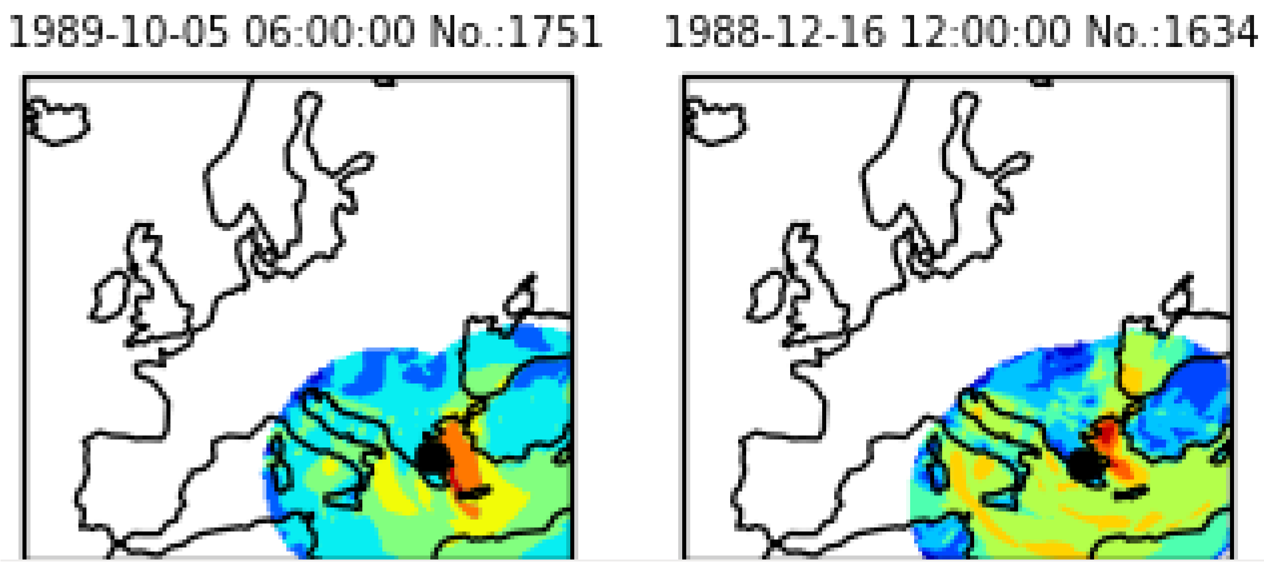
No. 1634 Best match **F**: 1751, 427, 86, 2586, 2777, 3174, 404, 5105, 4203, 1954.

Met Office

Match for Candidate 1634 is 1751!

The storm from 5 October 1989 in the Aegean Sea (Diagnostics of Cyclogenesis Over the Aegean Sea Using Potential Vorticity Inversion, H. A. Flocas) is very similar to

The storm on December 16th the year before. 20th December 1988 saw a record low temperature in central Greece (http://weatherspark.com/history/32209/1988/Dervenochoria-Central-Greece)



Met Office

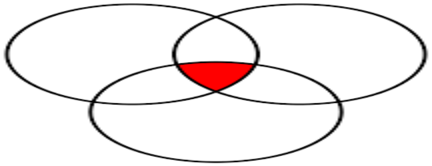
Evaluate close matches

Frobenius norm F of the distance between candidate and close members

$$||A||_F = [\sum_{i,j} abs(a_{i,j})^2]^{1/2}$$

Dot norm: subtract the vectors, make inner product and average (mean of dot product), np.dot(A.T*A).mean()
Here Dot offers an alternative set of close matches

Take top M candidates for each of the 3 distance measures and calculate intersection!

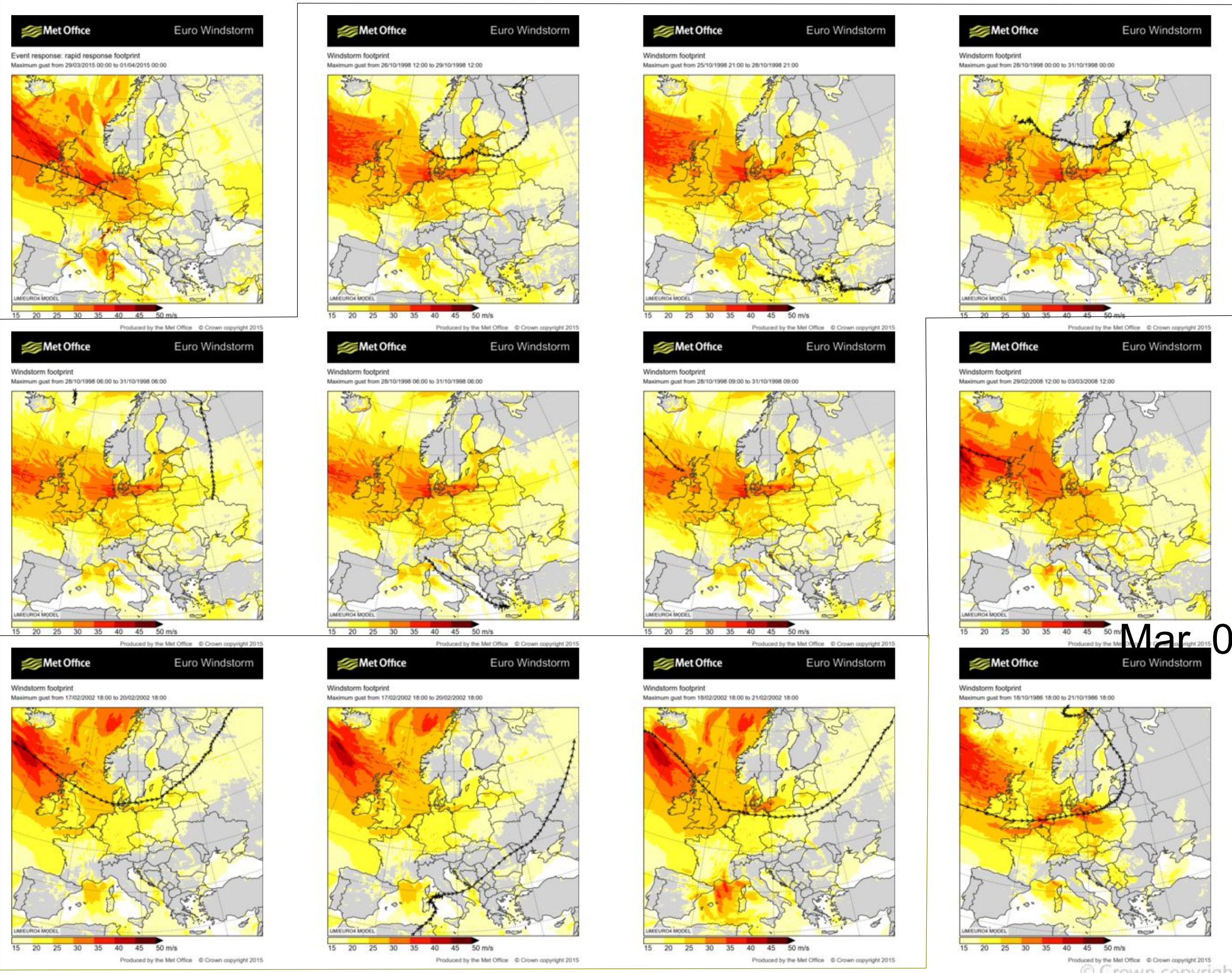
Provide an envelope of similar h  ch real time SFP

Storm Footprint matches

Met Office

Niklas

Xylia 96



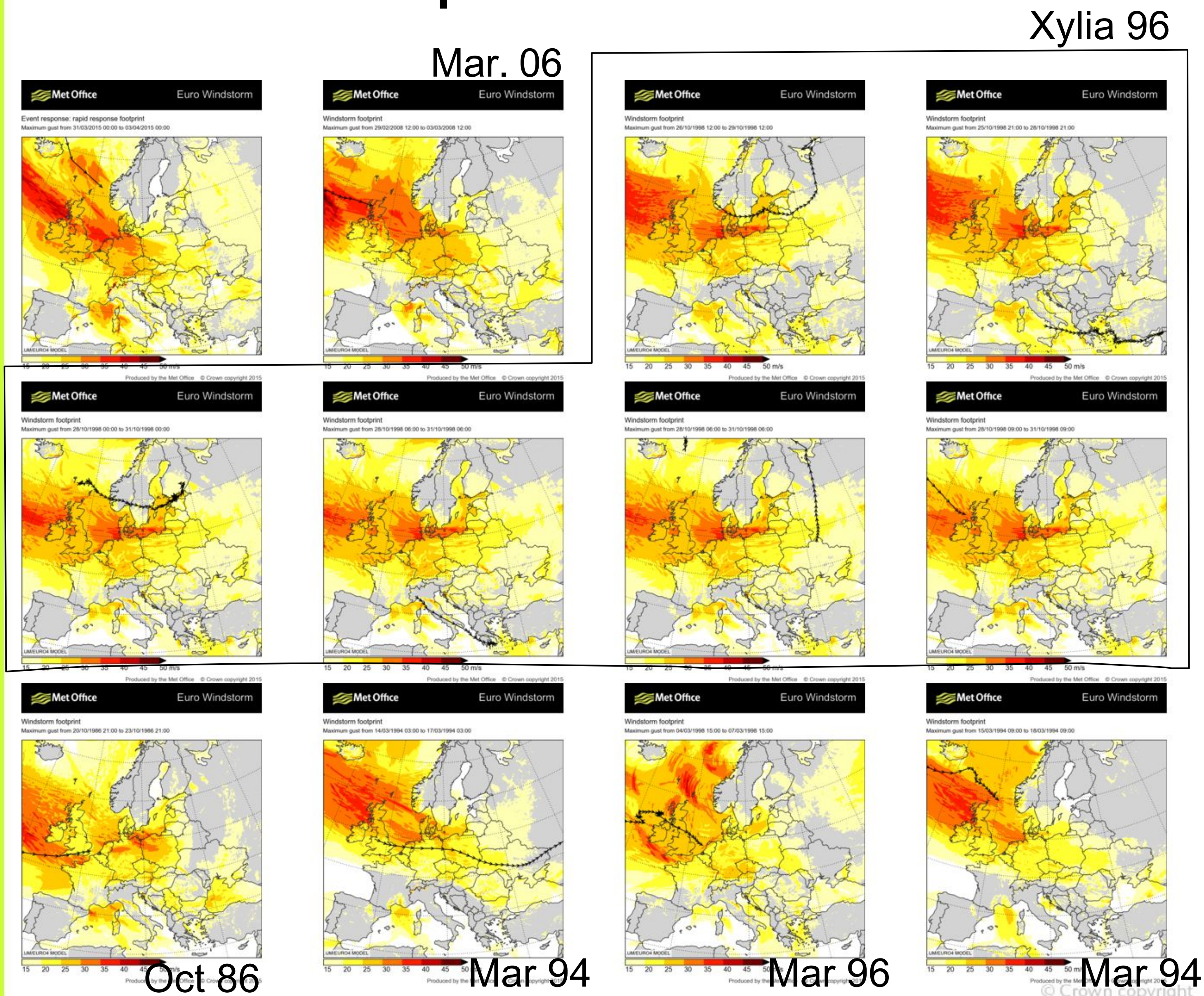
Mar. 06

Anna 02

Oct. 86

Storm Footprint matches

Mike



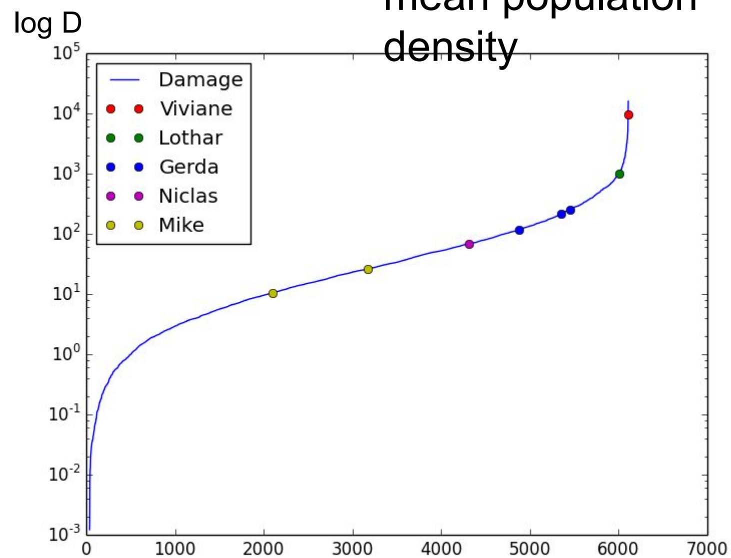
European Windstorm damage curve

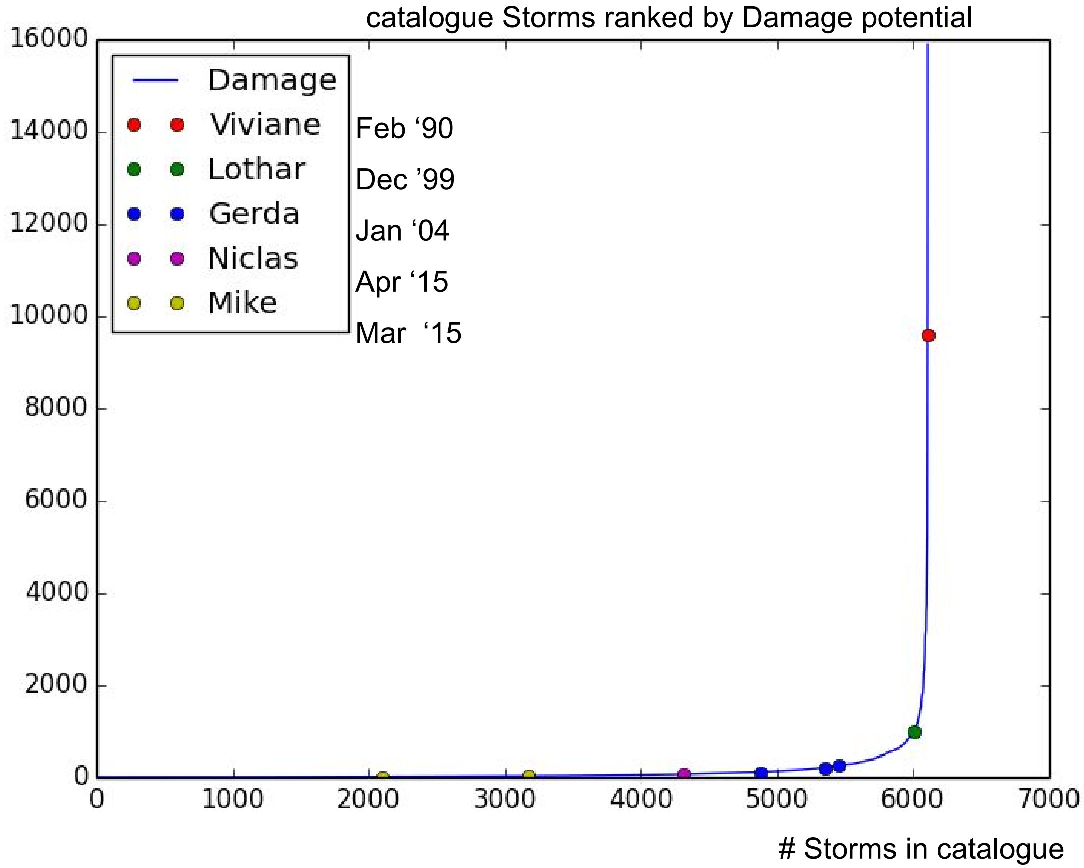
Damage potential: function (Gust speed, Size of storm and Value at risk)
Land area mean where U>25:

Gust speed:
Sum (U>25 -25)^3

Size of storm:
Fraction of land grid points

Value at risk:
mean population density





Storm Footprint matching

Match a recent or predicted SFP with
The Historical Windstorm Catalogue
By generating a list of features for all
catalogue members (one off!)
Rank features and pick closest members
Evaluate final group of candidates
For each real time SFP, provide an
envelope of similar historic events.

Apply pattern matching to:

Find similar storms in catalogue

Unpick the strongest/most damaging events (rather than relying on other naming agents)

Prove similarity between catalogue (ERA) and event response (Euro4) for overlapping period (late 2013)

Match event response SFP with appropriate cluster from Catalogue

To gauge impact of predicted event and prepare!

Met Office

Questions & Answers

www.metoffice.gov.uk/insurance
or contact us on peril@metoffice.gov.uk

Met Office

Used in machine learning context, to analyse medical images, remote sensing, crowd behaviour from surveillance videos, search engines….

- p(i,j): (i,j)th entry in a normalized gray-tone spatial dependence matrix, p(i,j) = P(i,j) / R * P(i,j) is the co-occurrence matrix and R is the sum of values in it, thus P(i,j) can be considered as the joint distribution of i and j, which are gray levels of the original image. The value of entry p(i,j) is supposed to be very small due to the large size of the co-occurrence matrix.
- px(i) / py(i): ith entry in the marginal-probability distribution matrix obtained by summing the rows/columns of p(i,j).
- Ng: Number of distinct gray levels in the image.

Angular Second Moment	$\sum_i \sum_j p(i,j)^2$
Contrast	$\sum_{n=0}^{N_g-1} n^2 \{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)\}, \lvert i-j \rvert = n$
Correlation	$\frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ where μ_x , μ_y , σ_x , and σ_y are the means and std. deviations of p_x and p_y , the partial probability density functions
Sum of Squares: Variance	$\sum_i \sum_j (i-\mu)^2 p(i,j)$
Inverse Difference Moment	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Sum Average	$\sum_{i=2}^{2N_g} i p_{x+y}(i)$ where x and y are the coordinates (row and column) of an entry in the co-occurrence matrix, and $p_{x+y}(i)$ is the probability of co-occurrence matrix coordinates summing to $x+y$
Sum Variance	$\sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i)$
Sum Entropy	$-\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} = f_8$
Entropy	$-\sum_i \sum_j p(i,j) log(p(i,j))$
Difference Variance	$\sum_{i=0}^{N_g-1} i^2 p_{x-y}(i)$
Difference Entropy	$-\sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}$
Info. Measure of Correlation 1	$\frac{HXY-HXY1}{\max\{HX,HY\}}$
Info. Measure of Correlation 2	$(1-\exp[-2(HXY2-HXY)])^{\frac{1}{2}}$ where $HXY = -\sum_i \sum_j p(i,j) \log(p(i,j))$, HX , HY are the entropies of p_x and p_y , $HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i)p_y(j)\} HXY2 = -\sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$
Max. Correlation Coeff.	Square root of the second largest eigenvalue of $\mathbf{Q}$ where $\mathbf{Q}(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$