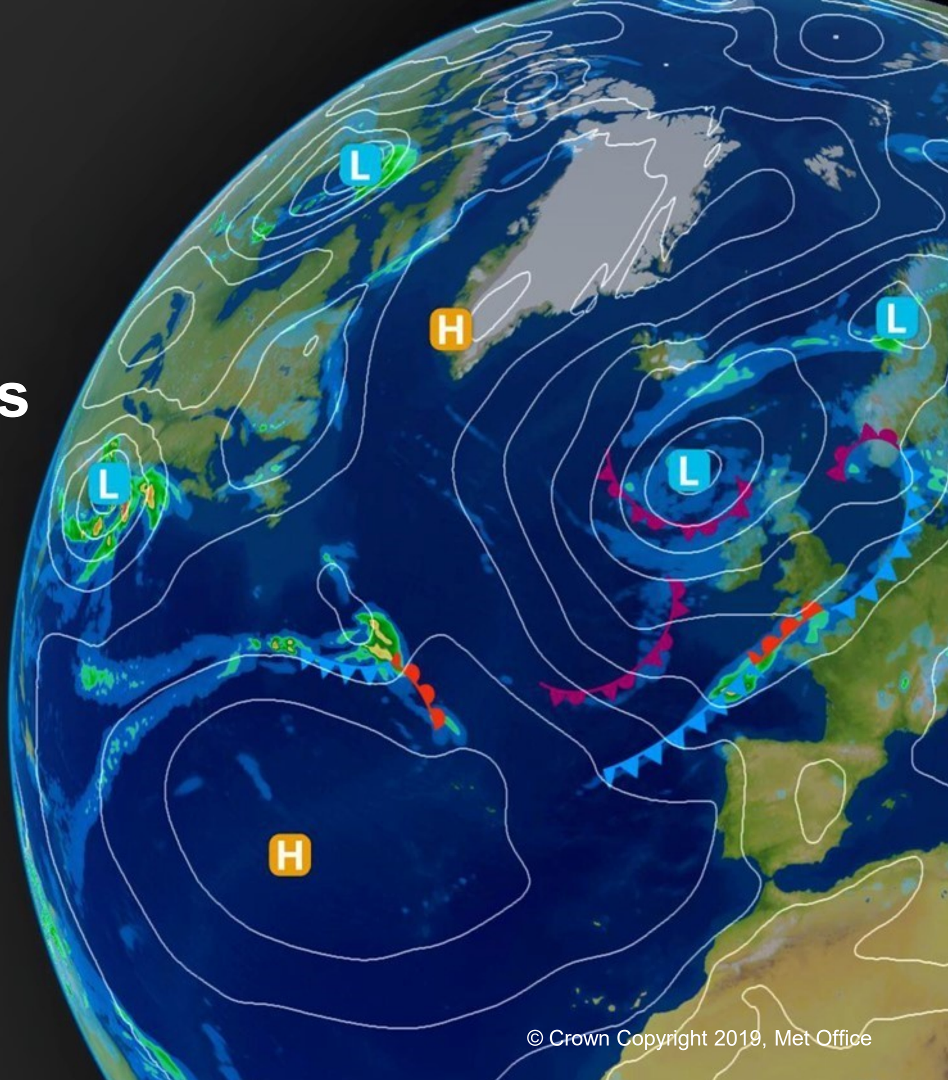
Representing Windstorm Footprints using Observations and Meteorological Models

Laura Dawkins, Met Office

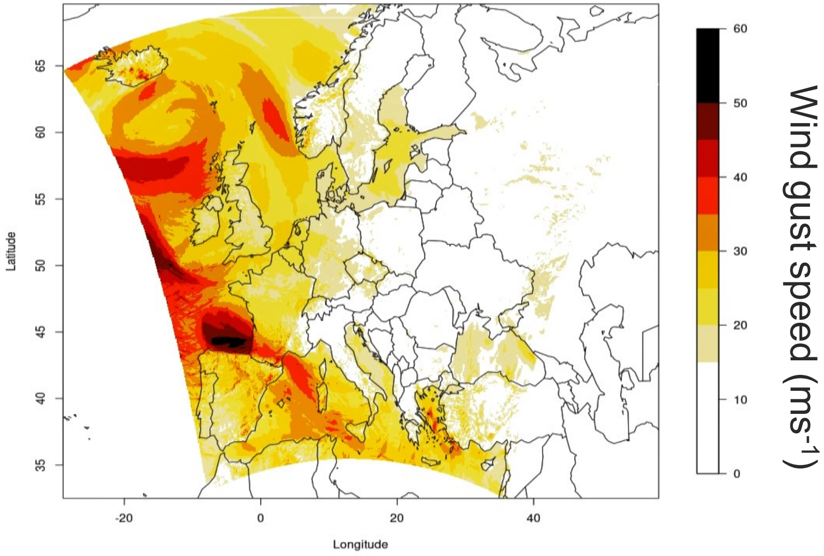
Tristan Perotin, AXA



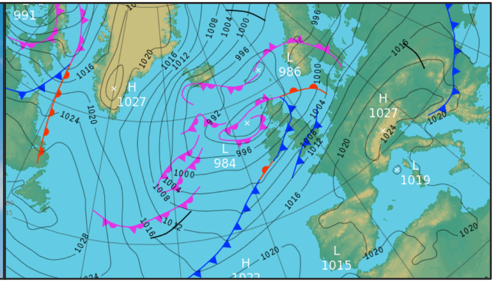
Motivation

- Insurance industry benefit from having the **most accurate** representation of the windstorm footprint at the **earliest opportunity**
- Prompt identification of the most **affected areas**
- Timely estimation of the **associated losses**
- Improve knowledge of **vulnerability** when combined with historical loss data

Klaus (23rd – 25th January 2009)
Met Office, North Atlantic European Model (EURO4): ~4.4km horizontal resolution



Windstorm Footprint: Maximum 3 second wind gust speed to occur in each location over the 72 hour lifespan of the storm



Aim

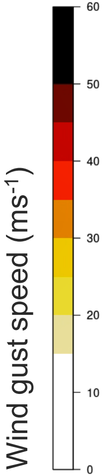
- Investigate **different methods** for estimating the windstorm footprint using **observations** and **numerical weather prediction (NWP) models**

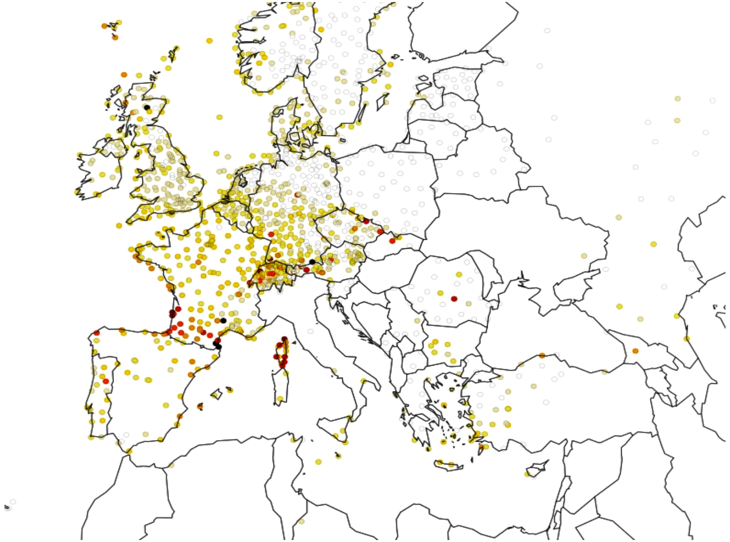
- Observations: relatively accurate but **spatially heterogeneous**
- Meteorological NWP models: spatially complete but **biased**

- **How can we effectively combine these two sources of information?**

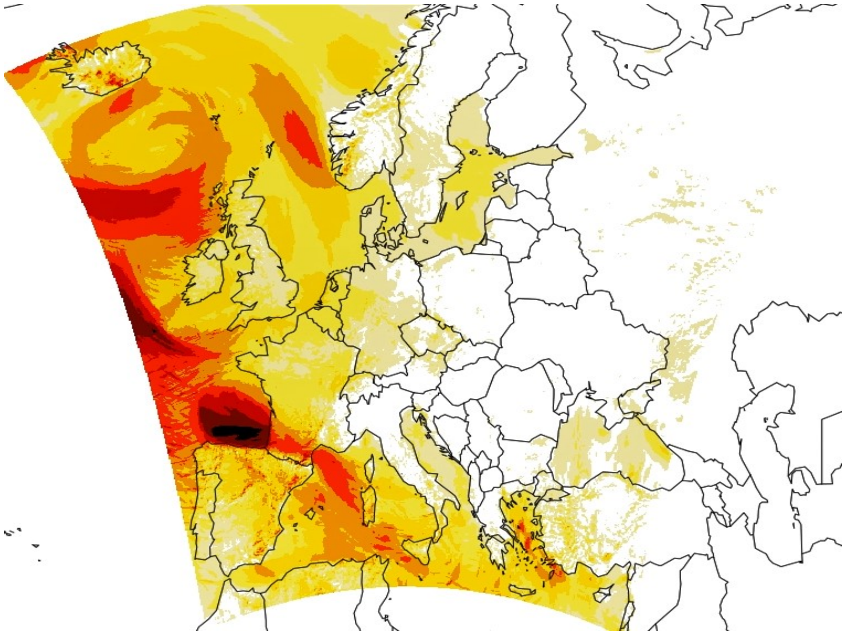
Observed (left) and modelled (right) footprints for windstorm Klaus (23rd – 25th January 2009)

Data



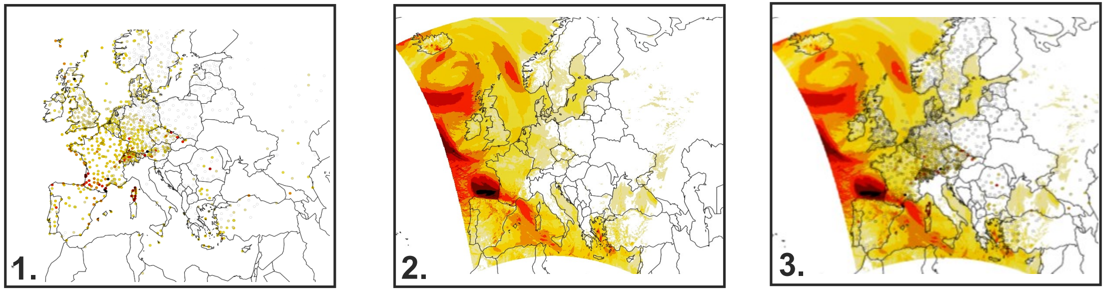


Observations: Wind gust speeds taken from a **station network** of ~1500 stations across Europe



NWP Model: Met Office, North Atlantic European Model (EURO4), ~4.4km horizontal resolution

Method

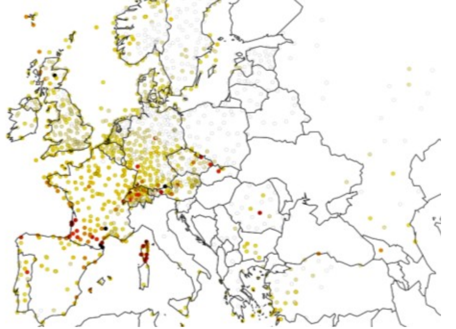


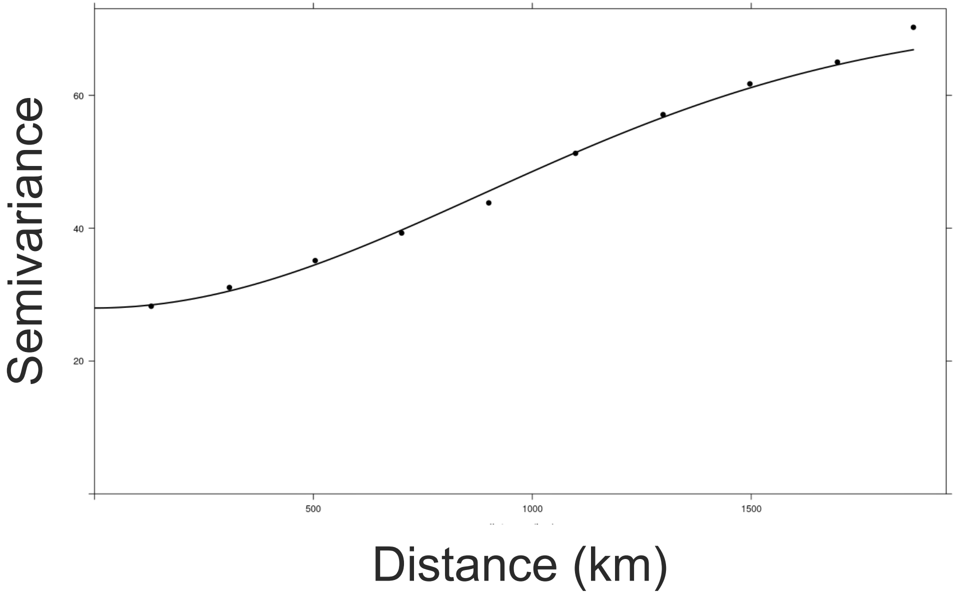
- Compare how well **three different approaches** for representing a single windstorm footprint are able to **predict** observations at locations not included in model fitting

1. Using **observations** only: a spatial geostatistical model, kriged predictions
2. Using **meteorological NWP model** only: interpolate to the prediction location
3. **Combined approach:** using the statistical recalibration approach of Youngman and Stephenson (2019)

Youngman, B. D. and Stephenson, D. B. (2019). Spatial inference for hazard event intensities using imperfect observation and simulation data. Preprint available from http://empslocal.ex.ac.uk/people/staff/by223/youngman-stephenson_ recalibration.pdf

1. Geostatistical model for observations





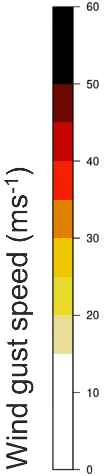
- For each pair of locations, empirically calculate the **semivariance** (a measure of dissimilarity), plot against separation distance

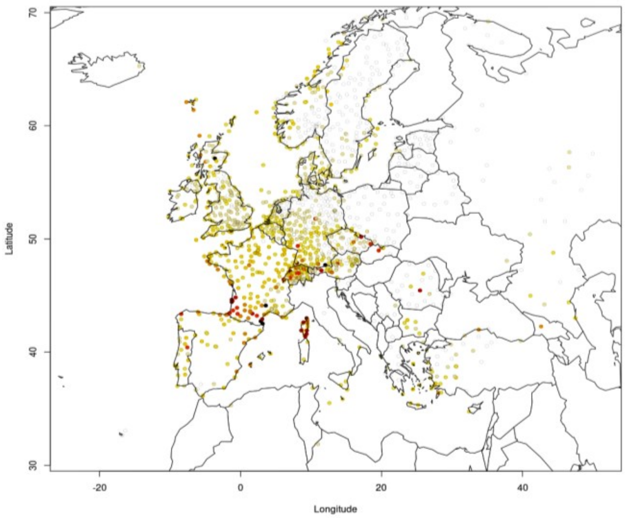
- Calculate the average semivariance for **separation distance bins** (here every 200km)

- Fit a **parametric covariance function** to these points – here the Gaussian model

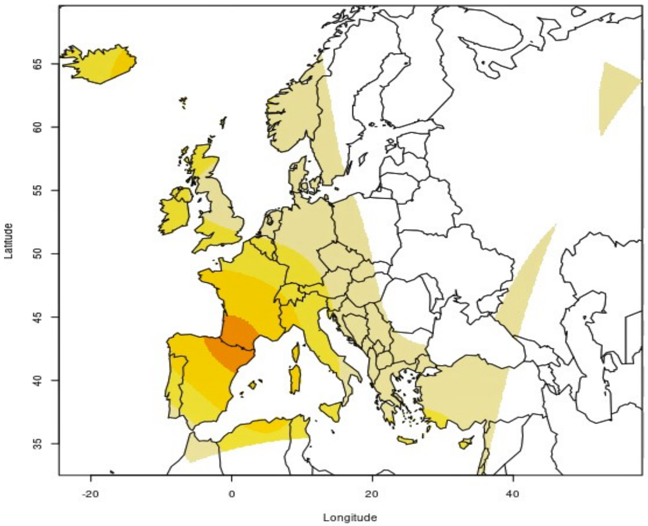
- This model can be used to **predict** at unobserved locations, based on a **weighted average** of neighbouring locations (**ordinary kriging**)

1. Geostatistical model for observations



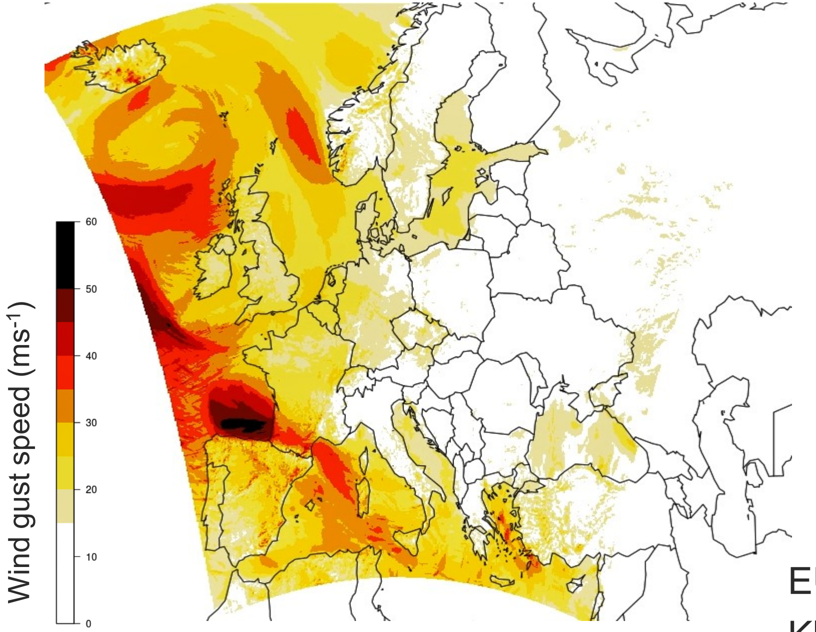


Observed footprint for windstorm Klaus



Kriged observation footprint (4km resolution) for windstorm Klaus

2. Interpolating NWP model



- Use **bilinear interpolation** to **non-parametrically** estimate wind gust speed at any desired location, based on wind gust speeds at **surrounding locations**

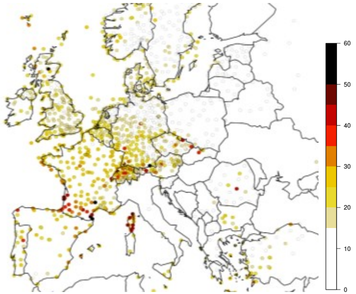
- Use the `interp.surface()` function in R

EURO4 footprint for windstorm Klaus (4km resolution)

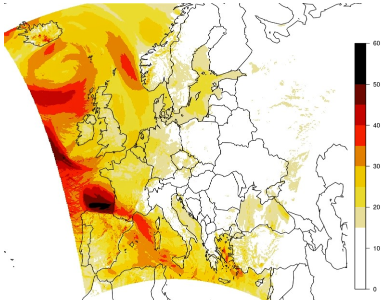
Observation = **Unobservable truth** + Measurement error **(1)**
Unobservable truth = **NWP model** + Model discrepancy **(2)**

3. Recalibration

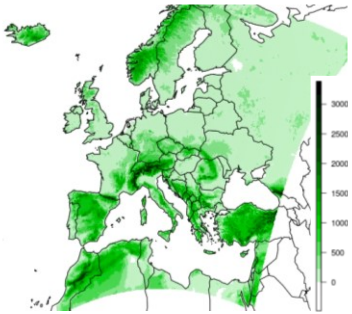
- Model **Observations** such that the spatial mean process is a function of **NWP model**, and known model parameterisations (e.g. orography)
- Quantifying the difference in spatial structures and measurement error



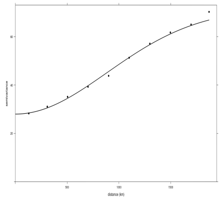
Observations ~



NWP model +



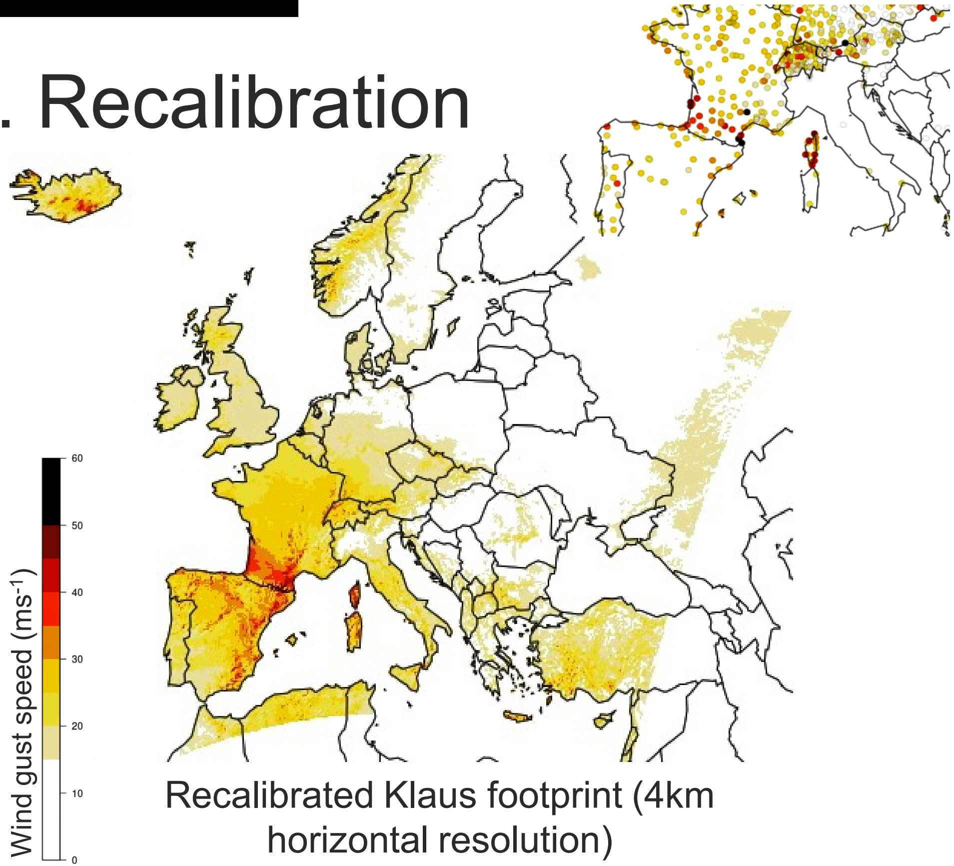
Orography ,



Spatial structure adjustment and measurement error

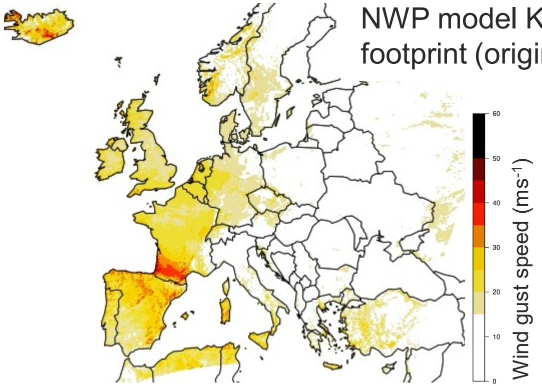
- Use this model to predict **Unobservable truth** at a given location using its joint distribution with **Observations** derived from equation **(1)**

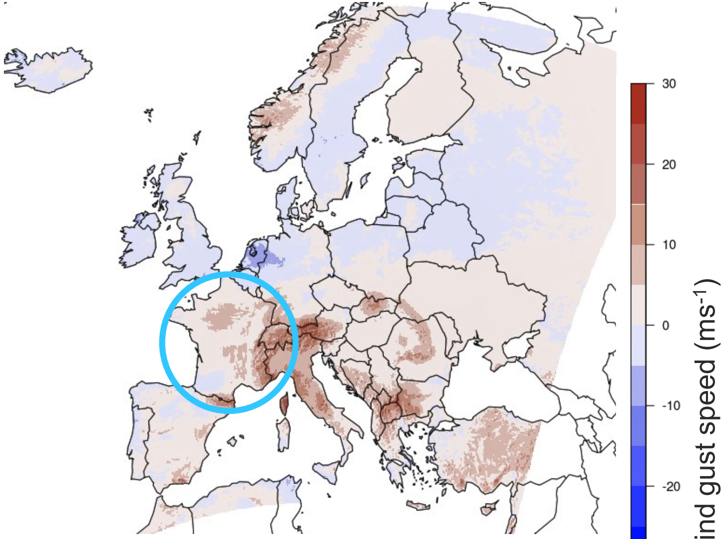
3. Recalibration



Recalibrated Klaus footprint (4km horizontal resolution)

NWP model Klaus footprint (original)





Difference (recalibrated – original)

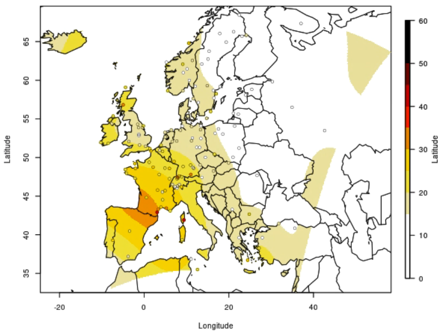
*Which approach gives **best predictions** of observations not included in model fitting?*

Results: Klaus

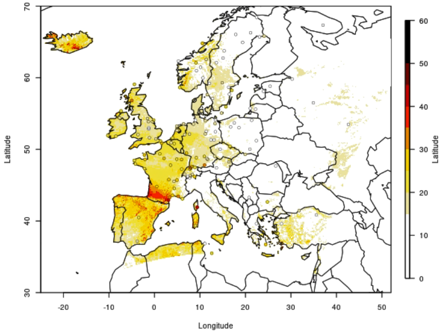
- **10-fold cross validation** (~140 observations per validation sample)
- For each of the 10 cross validations, calculate the **Root Mean Squared Error (RMSE)**

1. For **all** wind gust speeds
2. For observed wind gust speed **> 25ms^{-1}**, most relevant for insured loss estimation

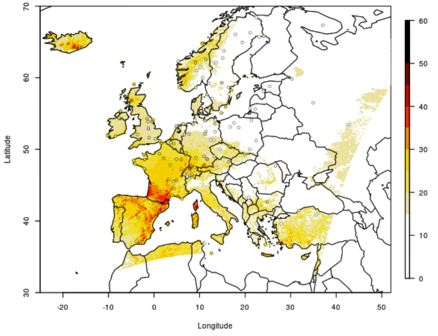
1. Kriged Observations

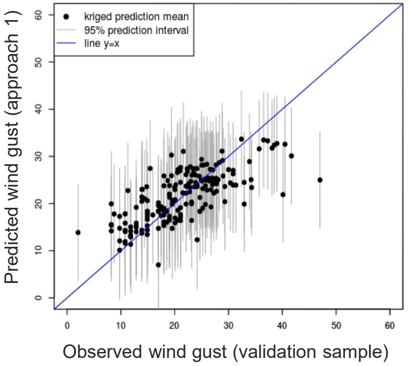


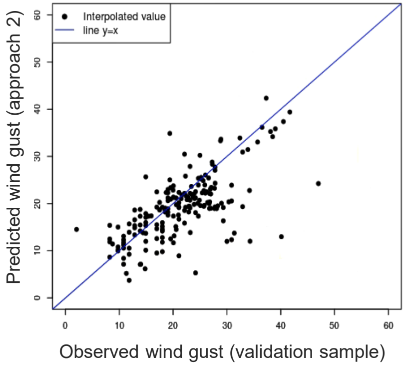
2. Interpolated NWP model

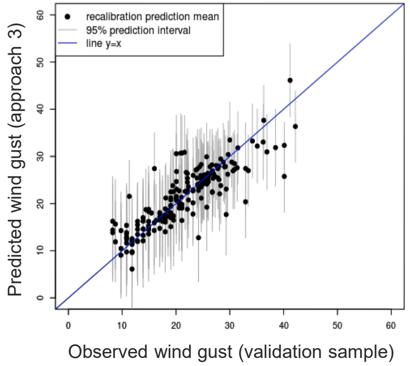


3. Recalibrated footprint





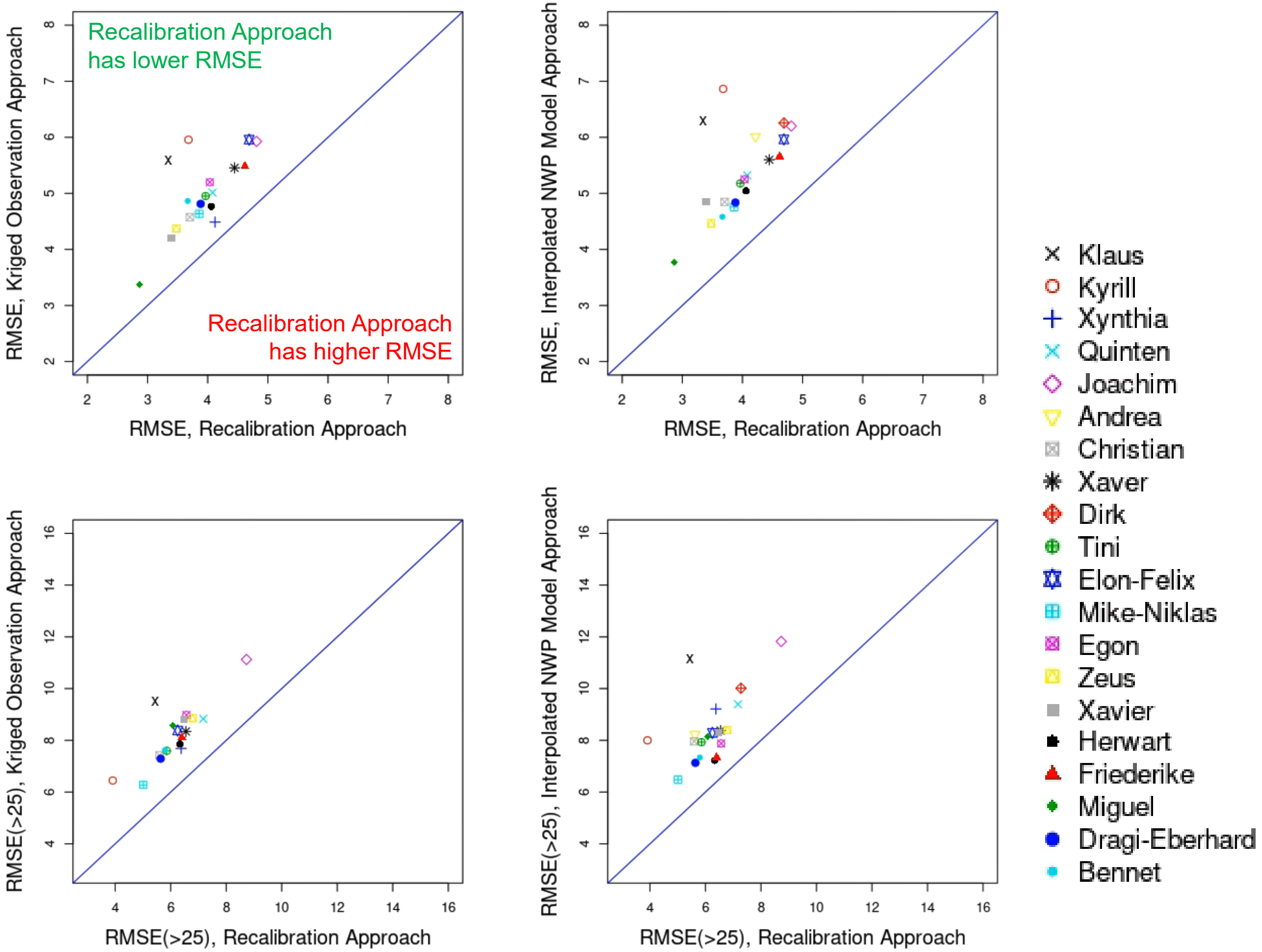




Results

1. RMSE for all 10-fold cross validations for Klaus
2. Mean RMSE for Klaus
3. Mean RMSE for 20 storms (2007-2019)

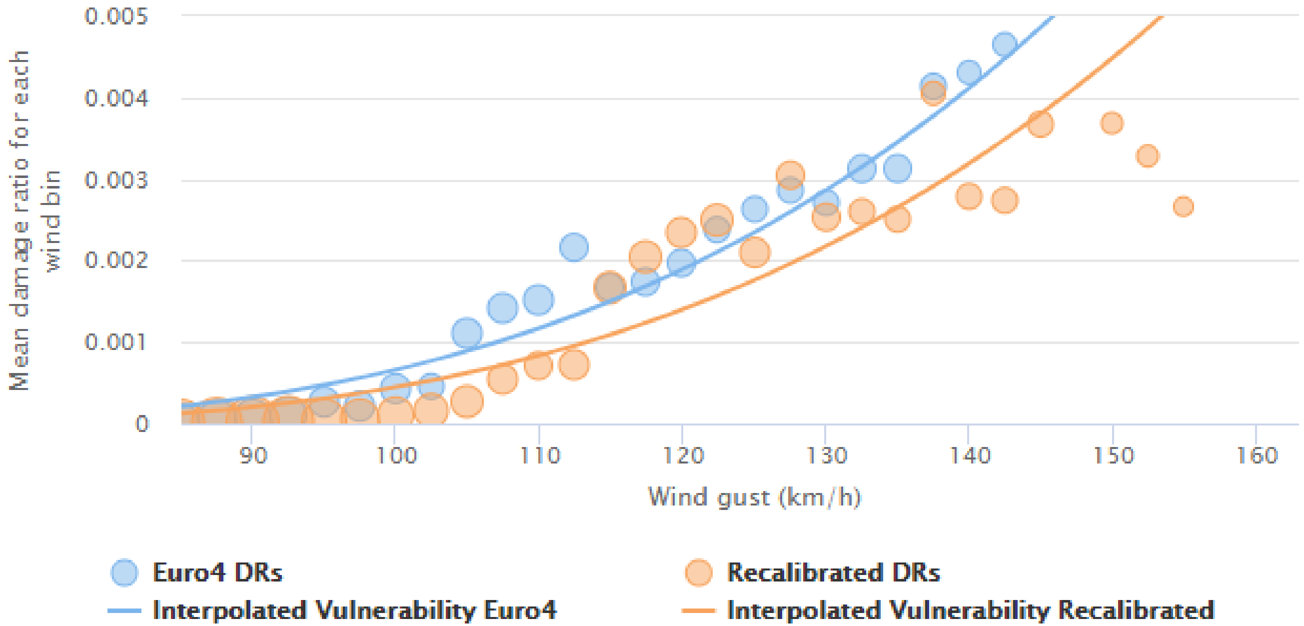
- For all 20 storms, the **recalibration approach** gives **more accurate** predictions
- Both for all wind gust magnitudes and **extreme wind gusts**



Conclusion

- Explored three approaches for using observations and meteorological NWP model for **estimating the windstorm footprint** - separately and in combination
- The combined approach followed the hazard footprint recalibration approach of **Youngman and Stephenson (2019)**
- For all 20 storms we have explored, the **recalibration approach** gives **more accurate** predictions of ‘new’ observed wind gusts speeds
- This is true for wind gust speed of all magnitudes and **extreme wind gusts** (>25ms^{-1})
- **AXA should employ the recalibration method** to achieve more accurate representations of both historical and future windstorm footprints

Applications: Vulnerability Modelling (Klaus Storm)



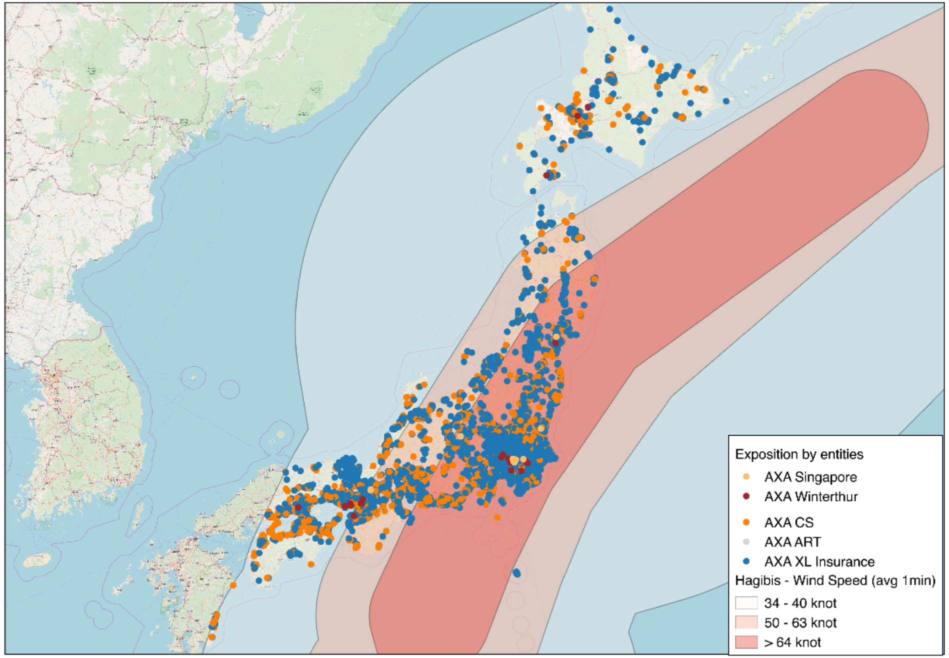
- **30% difference in implied vulnerability**
- Consistency with gust observations is necessary for **model modularity**
- It also improves reliability of **comparisons between events**

Applications: Event Response (Hagibis Typhoon)

- Improved **alert systems** and **loss prevention measures** thanks to finer vulnerability knowledge

- Improved **early estimations of number of claims** and **total event losses**

- Improved **claim handling** and **urgent assistance services** thanks to better identification of clients at risks



Snapshot from the internal Hagibis typhoon early event response report

Extra slides

ID	Storm Name	Start Date	Impacted Countries
1	Klaus	23/01/2009	FRA
2	Kyrill	10/01/2007	BEL, CHE, DEU, FRA, GBR, IRL, LUX, NLD
3	Xynthia	28/02/2010	BEL, CHE, DEU, FRA
4	Quinten	10/02/2009	FRA, BEL, NLD, DEU
5	Joachim	16/12/2011	CHE, DEU, FRA
6	Andrea	04/01/2012	BEL, CHE, DEU, FRA, GBR, NLD
7	Christian	27/10/2013	BEL, DEU, DNK, GBR, NLD, SWE
8	Xaver	04/12/2013	DEU, DNK, GBR, NLD, NOR, SWE
9	Dirk	22/12/2013	FRA, GBR
10	Tini	12/02/2014	GBR, IRL
11	Elon-Felix	08/01/2015	DEU, DNK, GBR, NOR, SWE
12	Mike-Niklas	30/03/2015	AUT, BEL, CHE, DEU, GBR, NLD
13	Egon	12/01/2017	DEU, FRA
14	Zeus	05/03/2017	FRA
15	Xavier	04/10/2017	DEU
16	Herwart	28/10/2017	AUT, DEU
17	Friederike	16/01/2018	BEL, DEU, GBR, NLD
18	Miguel*	07/06/2019	POR, SPA, FRA, UK, NLD
19	Dragi-Eberhard*	08/03/2019	BEL, CHE, DEU, FRA, GBR, LUX, NLD
20	Bennet*	02/03/2019	BEL, CHE, DEU, FRA, LUX

interp.surface

From fields v9.8-6
by Douglas Nychka

Percentile

Fast Bilinear Interpolator From A Grid.

Uses bilinear weights to interpolate values on a rectangular grid to arbitrary locations or to another grid.

Keywords spatial

Usage

```
interp.surface(obj, loc)
interp.surface.grid(obj, grid.list)
```

Arguments

obj A list with components x,y, and z in the same style as used by contour, persp, image etc. x and y are the X and Y grid values and z is a matrix with the corresponding values of the surface

loc A matrix of (irregular) locations to interpolate. First column of loc isthe X coordinates and second is the Y's.

grid.list A list with components x and y describing the grid to interpolate. The grids do not need to be equally spaced.

Details

Here is a brief explanation of the interpolation: Suppose that the location, (locx, locy) lies in between the first two grid points in both x an y. That is locx is between x1 and x2 and locy is between y1 and y2. Let ex= (l1-x1)/(x2-x1) ey= (l2-y1)/(y2-y1). The interpolant is

(1-ex)(1-ey)*z11 + (1- ex)(ey)*z12 + (ex)(1-ey)*z21 + (ex)(ey)*z22

Where the z's are the corresponding elements of the Z matrix.

Note that bilinear interpolation can produce some artifacts related to the grid and not reproduce higher behavior in the surface. For, example the extrema of the interpolated surface will always be at the parent grid locations. There is nothing special about about interpolating to another grid, this function just includes a `for` loop over one dimension and a call to the function for irregular locations. It was included in fields for convenience. since the grid format is so common.

See also the akima package for fast interpolation from irrgeular locations. Many thanks to Jean-Olivier Irisson for making this code more efficient and concise.

Value

An vector of interpolated values. NA are returned for regions of the obj$z that are NA and also for locations outside of the range of the parent grid.

Met Office

2.2 Inference

2.2.1 Parameter estimation

Relations (1) and (3) in Section 2.1 imply the marginal model

$$Y(s) \,|\, x(s), \sigma^2, \boldsymbol{\beta}, \boldsymbol{\theta} \sim GP\big(m(x(s)), \sigma^2 c(\,,\,)\big), \tag{4}$$

where $m(\,)$ is as in Relation (3) and $\sigma^2 c(\,,\,) = \sigma_Y^2 c_Y(\,,\,) + \sigma_X^2 c_X(\,,\,)$, as Relation (1) may be written as a GP with covariance function $\sigma_Y^2 c_Y(\,,\,)$. For tractability, suppose that $m(x) = \mathbf{h}^T(x)\boldsymbol{\beta}$, where $\mathbf{h}(\,)$ comprises q basis functions (e.g. $\mathbf{h}(x) = (1, x)^T$) and $\boldsymbol{\beta}$ comprises q regression coefficients. Depending on the forms chosen for the correlation functions, not all their parameters, collectively denoted $\boldsymbol{\theta}$, may be identifiable without prior knowledge, in particular if both $c_X(\,,\,)$ and $c_Y(\,,\,)$ contain nugget terms. We address this in Section 3 by specifying the measurement error. Relation (4) allows us to directly establish the relationship between the observations and simulator output and in turn perform inference. A possible drawback to this tractability is that only observation locations are used to infer $Z(s)$, whereas in Fuentes et al. (2003) all simulator output locations are used. Careful consideration must be given as to whether observation locations are sufficient for inferring $Z(s)$ for any s of interest.

Let $\mathbf{y} = \big(y(s_1), \ldots, y(s_n)\big)'$ denote observations on a hazard event at locations $s_1, \ldots, s_n$, and let $\mathbf{x} = \big(x(s_1), \ldots, x(s_n)\big)'$ denote corresponding simulator output. Construct $n \times q$ matrix $\mathbf{H}$ with ith row $\mathbf{h}^T\big(x(s_i)\big)$ for $i = 1, \ldots, n$, and the $n \times n$ matrix $\mathbf{A}(\boldsymbol{\theta})$ with (i,j)th element $c(s_i, s_j)$. The restricted log-likelihood, obtained by integrating over a uniform prior for $\boldsymbol{\beta}$ (Harville, 1974), is given by

$$\ell_R(\boldsymbol{\theta}) = -\frac{n-q}{2}\big(\log(2\pi) + \log\hat{\sigma}^2\big) - \frac{1}{2}|\mathbf{A}(\boldsymbol{\theta})| - \frac{1}{2}|\mathbf{H}^T\{\mathbf{A}(\boldsymbol{\theta})\}^{-1}\mathbf{H}|, \tag{5}$$

where

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}^T\{\mathbf{A}(\boldsymbol{\theta})\}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\{\mathbf{A}(\boldsymbol{\theta})\}^{-1}\mathbf{y}, \tag{6}$$

$$\hat{\sigma}^2 = \frac{1}{n-q}(\mathbf{y} - \mathbf{H}\hat{\boldsymbol{\beta}})^T\{\mathbf{A}(\boldsymbol{\theta})\}^{-1}(\mathbf{y} - \mathbf{H}\hat{\boldsymbol{\beta}}). \tag{7}$$

We choose $\boldsymbol{\theta}$ to maximise Equation (5).

2.2.2 Actual process estimation

We can use an estimate of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$, and the assumptions made in Section 2.1 to infer $Z(s)$ for the hazard event for any $s \in R$. Noting that

$$\begin{pmatrix} \mathbf{Y} \\ Z(s) \end{pmatrix} \sim MVN\left(\begin{pmatrix} \mathbf{H}\hat{\boldsymbol{\beta}} \\ \mathbf{h}^T(x(s))\hat{\boldsymbol{\beta}} \end{pmatrix}, \begin{pmatrix} \hat{\sigma}^2\mathbf{A}(\hat{\boldsymbol{\theta}}) & \hat{\sigma}_X^2\mathbf{t}(s) \\ \hat{\sigma}_X^2\mathbf{t}^T(s) & \hat{\sigma}_X^2 c_X(s,s) \end{pmatrix}\right), \tag{8}$$

where $\mathbf{t}^T(s) = \big(c_X(s_1, s), \ldots, c_X(s_n, s)\big)$, it follows that

$$Z(s) \,|\, \mathbf{Y} = \mathbf{y} \sim GP\big(m^*(x(s)), c^*(\,,\,)\big), \tag{9}$$

where

$$m^*(x(s)) = \mathbf{h}^T(x(s))\hat{\boldsymbol{\beta}} + \hat{\sigma}_X^2\mathbf{t}^T(s)\{\hat{\sigma}^2\mathbf{A}(\hat{\boldsymbol{\theta}})\}^{-1}(\mathbf{y} - \mathbf{H}\hat{\boldsymbol{\beta}}), \tag{10}$$

$$c^*(s, s') = \hat{\sigma}_X^2\big[c_X(s,s') - \hat{\sigma}_X^2\mathbf{t}^T(s)\{\hat{\sigma}^2\mathbf{A}(\hat{\boldsymbol{\theta}})\}^{-1}\mathbf{t}(s')\big]. \tag{11}$$

Implementation of equations (9)–(11) may be simplified by noting that

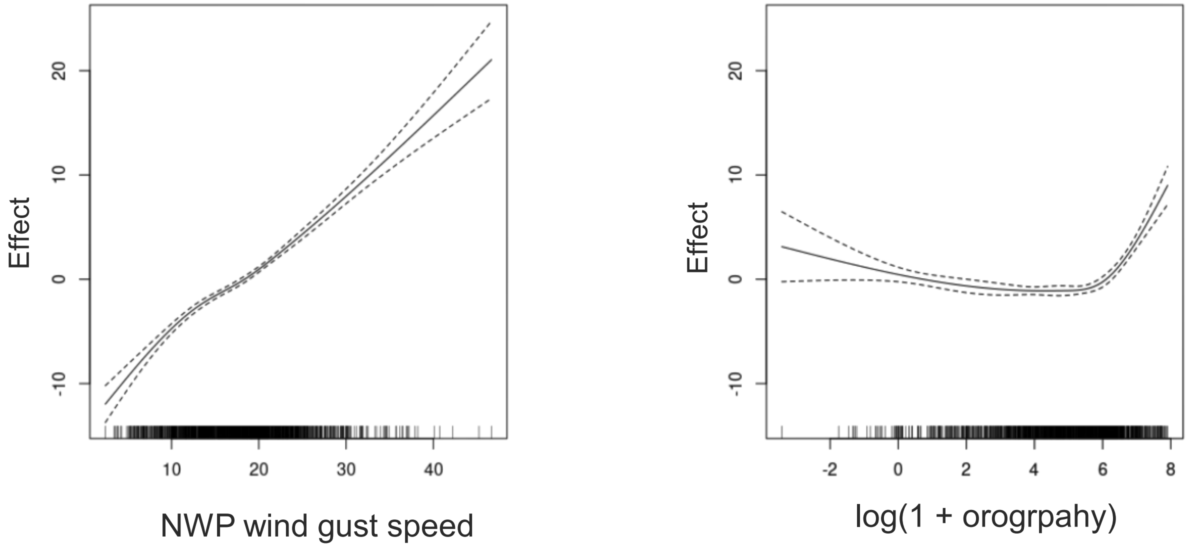
$$\{\hat{\sigma}^2\mathbf{A}(\hat{\boldsymbol{\theta}})\}^{-1} = \{\hat{\Sigma}_X(\hat{\boldsymbol{\theta}}) + \hat{\sigma}_Y^2\mathbf{I}_n\}^{-1} \tag{12}$$

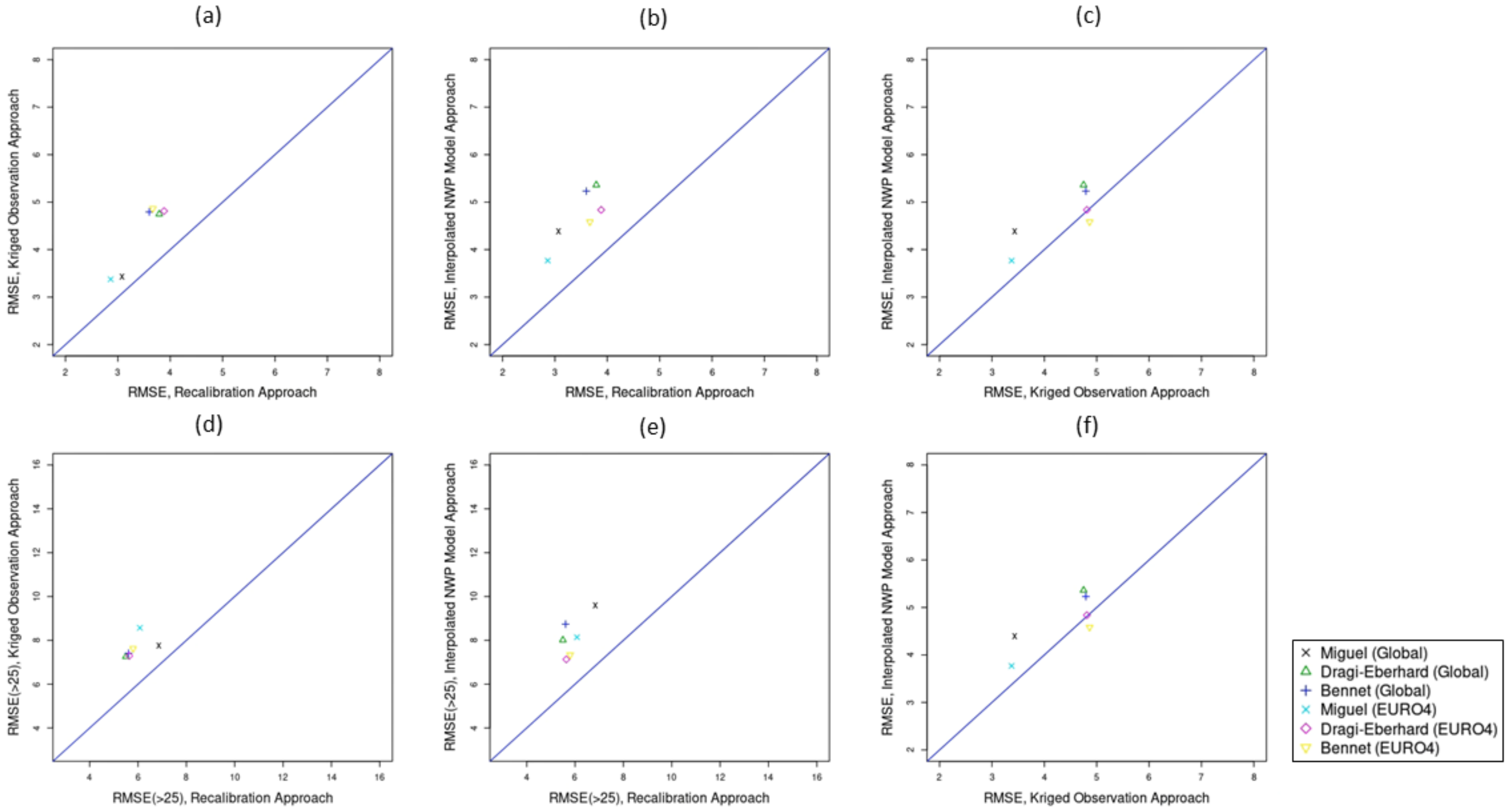
$$= \hat{\sigma}_Y^{-2}[\hat{\Sigma}_X(\hat{\boldsymbol{\theta}})]^{-1}\{[\hat{\Sigma}_X(\hat{\boldsymbol{\theta}})]^{-1} + \hat{\sigma}_Y^{-2}\mathbf{I}_n\}^{-1}, \tag{13}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix and the (i,j)th elements of $\hat{\Sigma}_X(\hat{\boldsymbol{\theta}})$ are given by $\hat{\sigma}_X c_X(s_i, s_j)$.

3. Recalibration

- In Youngman and Stephenson (2019) and here, we use **cubic regression splines** to relate the **NWP** modelled wind gusts and **orography** with **observed wind gusts**




(a)
RMSE, Kriged Observation Approach
RMSE, Recalibration Approach
(b)
RMSE, Interpolated NWP Model Approach
RMSE, Recalibration Approach
(c)
RMSE, Interpolated NWP Model Approach
RMSE, Kriged Observation Approach
(d)
RMSE(>25), Kriged Observation Approach
RMSE(>25), Recalibration Approach
(e)
RMSE(>25), Interpolated NWP Model Approach
RMSE(>25), Recalibration Approach
(f)
RMSE, Interpolated NWP Model Approach
RMSE, Kriged Observation Approach
Miguel (Global)
Dragi-Eberhard (Global)
Bennet (Global)
Miguel (EURO4)
Dragi-Eberhard (EURO4)
Bennet (EURO4)