Verification of Relative-Extreme Events

Michael Sharpe, Clare Bysouth and Becky Stretton

Operational Verification Systems and Products

7th June 2016

Met Office

© Crown copyright Met Office

Contents

- What is an extreme?

- How have we verified them?

- Are we any good at forecasting ...

  - Extreme heat

  - Extreme cold

  - Extreme rain

  - Extreme wind

- Summary

- Questions

Verification of Relative-Extreme Events

Days of Air Frost
Annual Average
1981-2010

Average Value (days)
> 100
80 to 100
60 to 80
40 to 60
20 to 40
10 to 20
< 10

absolute-extremes occur more frequently in some places...

...and less frequently in others

...whereas frosts on Scilly are big news...

- burst pipes

- dead plants

- ill pets

...are all likely
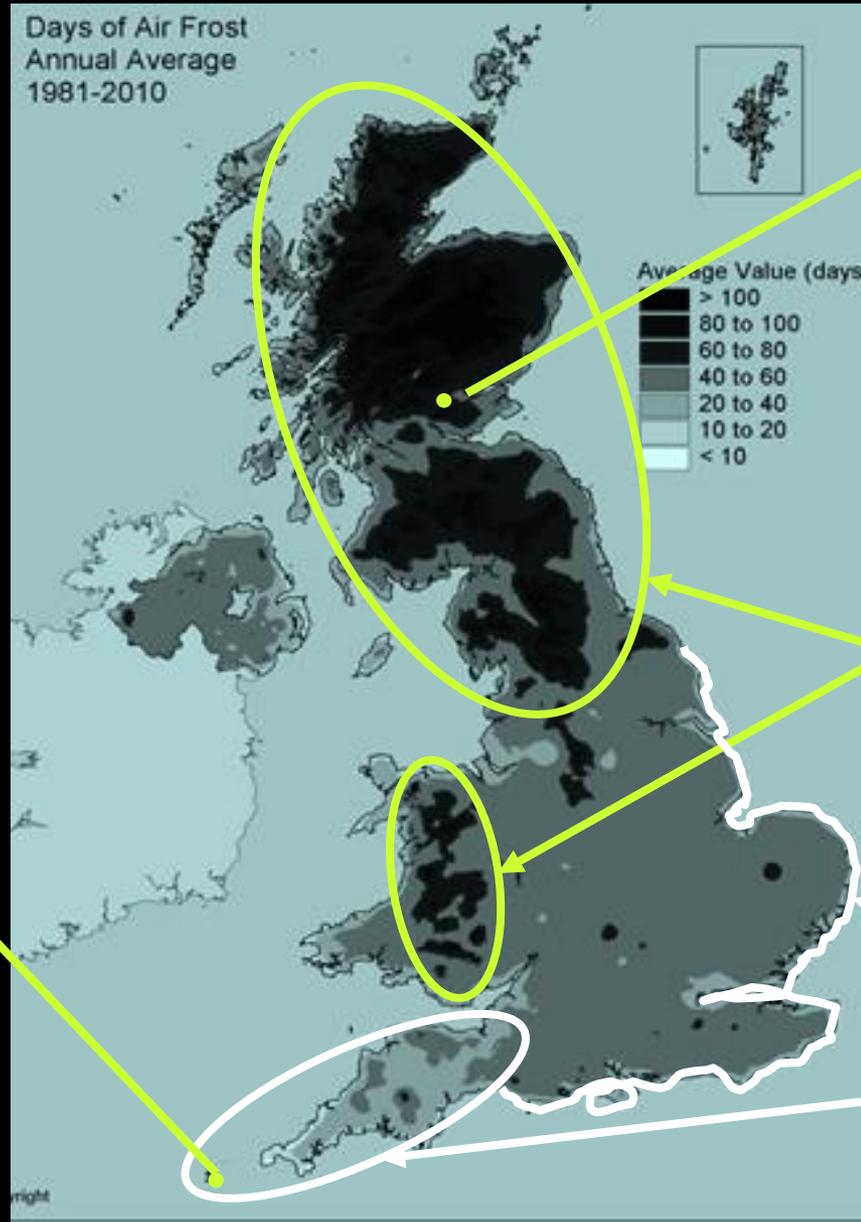
Frosts in Gleneagles are common...

...so they have little impact...

Performance measures tend to be dominated by cold areas...

...but they are often virtually unaffected by the skill in warmer areas

Verification of Relative-Extreme Events

So, what is an extreme weather event then?

Consider winter temperatures; is...

Last time the temperature
fell to -5C was 1991

- *-1C an extreme?*

- *-5C an extreme?*

- *-10C an extreme?*

Last time the temperature
fell below -1C was 1987

Last time the temperature
didn't fall below -5C was 1990

St Jame's Park	Isles of Scilly	GlenEagles

The answer depends on where you live!
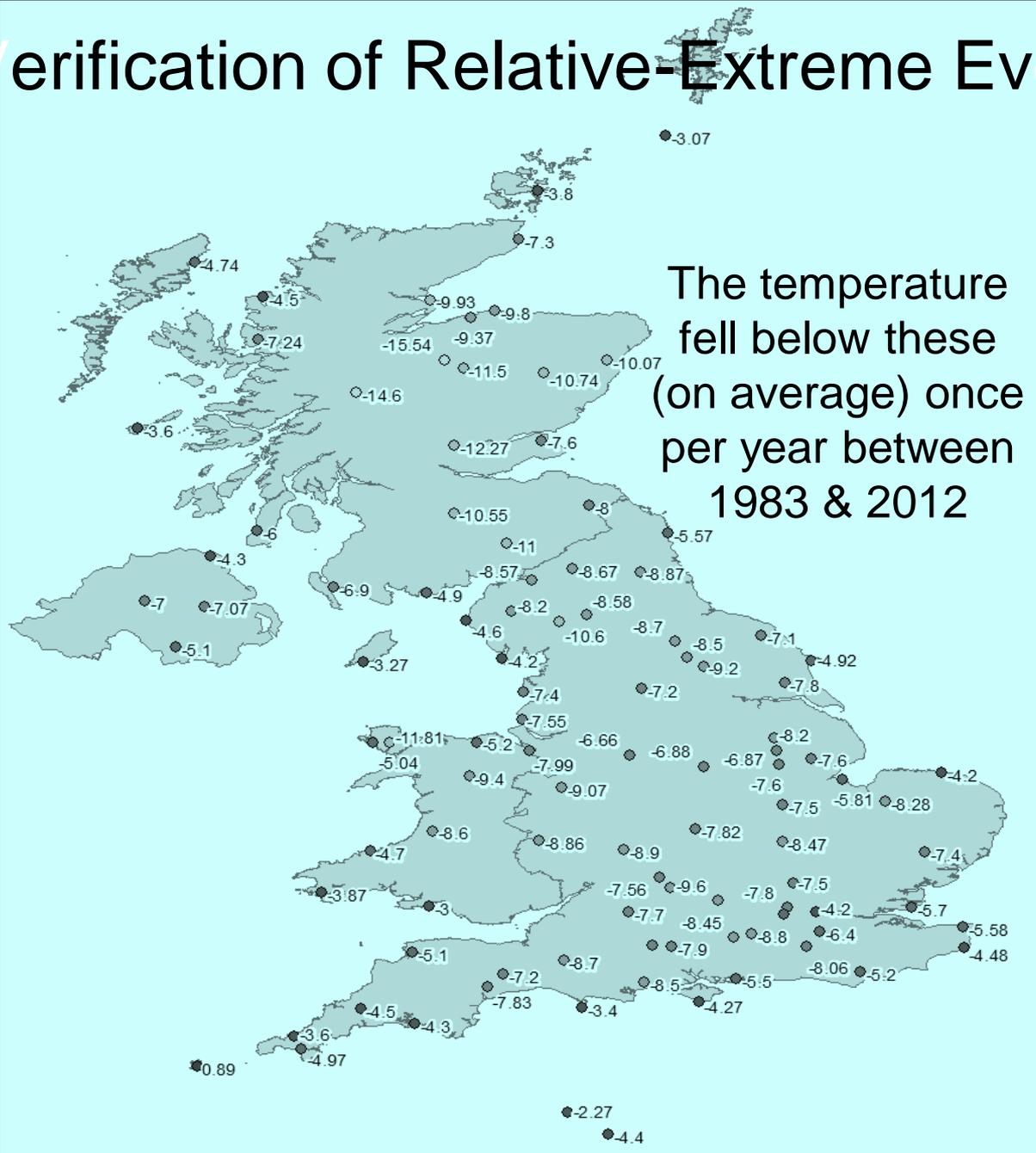
Verification of Relative-Extreme Events

Advantages of using local climatology to define extreme events...

a) *Base rate is the same everywhere*

b) *Simple aggregation gives all-site performance*

c) *Similar impact everywhere*

d) *Event defined in terms of a return period*

Disadvantages...

a) *Can't verify where climatology is unavailable*

b) *Assimilation affects model climatology (use observed)*

c) *multiple thresholds → technically more challenging*

d) *Not as simple to explain to your boss!*

Verification of Relative-Extreme Events

The forecast data which populates the Met Office website...

- *looks deterministic...*

  ....but looks can be deceiving

- *these values are the 'most likely'...*

  ...chosen from a distribution

20%

5%

Probability

11 11.5 12 12.5 13 13.5 14 14.5 15 15.5
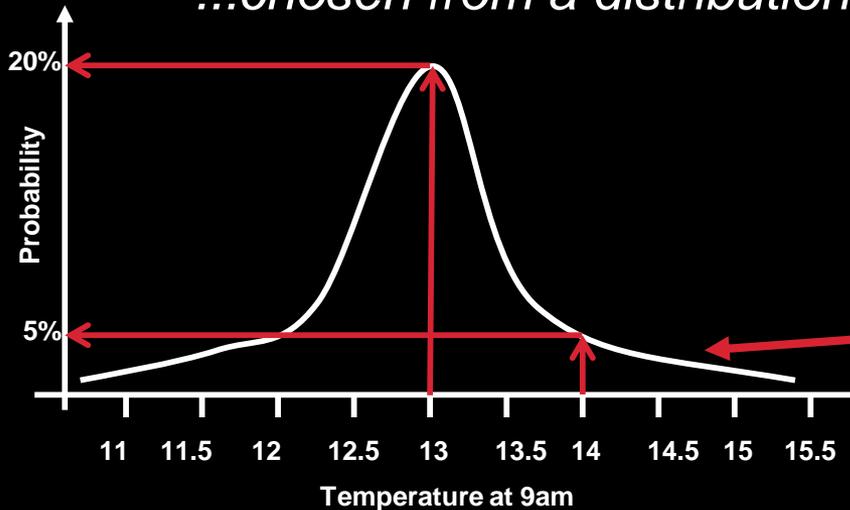
Temperature at 9am

... each temperature has a different likelihood

Weather | Climate | Learning | Research | Products | News | Holiday weather | Get ready for winter

| UK forecast | Severe weather warnings | Specialist forecasts | Last 24 hours |
| UK climate | World weather | Weather guide | Weather services |

Plymouth

Enter town, city or postcode to find a location.

Nearby locations

Recent locations   Plymouth ×   Burntisland ×   Beer (Beach) ×   Fernwood Ottery St. Mary ×   Honiton ×

Share

‹ Plymouth last 24 hours

Fri 30 Oct		Sat 31 Oct		Sun 1 Nov		Mon 2 Nov		Tue 3 Nov	
16°C 13°C		16°C 11°C		16°C 11°C		16°C 11°C		14°C 13°C	

0000	0300	0600	0900	1200	1500	1800	2100
12 °	13 °	12 °	13 °	14 °	14 °	13 °	13 °

Feels like temperature (°C)

| 10 ° | 10 ° | 10 ° | 11 ° | 12 ° | 12 ° | 12 ° | 12 ° |

Precipitation probability

| <5% | 10% | 10% | 40% | 40% | 50% | 50% | 40% |

Wind direction, speed & gust (mph)

ESE	E	ESE	ESE	ESE	ESE	ESE	SE
11	11	11	12	13	11	11	10
19	20	19	20	20	18	18	16

Visibility (E = Excellent, VG = Very Good, G = Good, M = Moderate, P = Poor, VP = Very Poor)

| G | G | M | G | G | M | M | M |

Humidity

| 92% | 93% | 94% | 92% | 89% | 91% | 94% | 95% |

UV index

| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |

Sunrise: 07:10   Sunset: 16:49   Moon phase: Third quarter   Daily air quality index: 2

Verification of Relative-Extreme Events

The forecast data which populates the Met Office website...

- *looks deterministic...*

  ....but looks can be deceiving

- *these values are the 'most likely'...*

  ...chosen from a distribution

- Extremes are <u>*very unlikely*</u> to be chosen as the *'most likely'* value

  - **shouldn't** just verify the *'most likely'*

  - **should** verify the distribution

- **BUT** *how.......?*
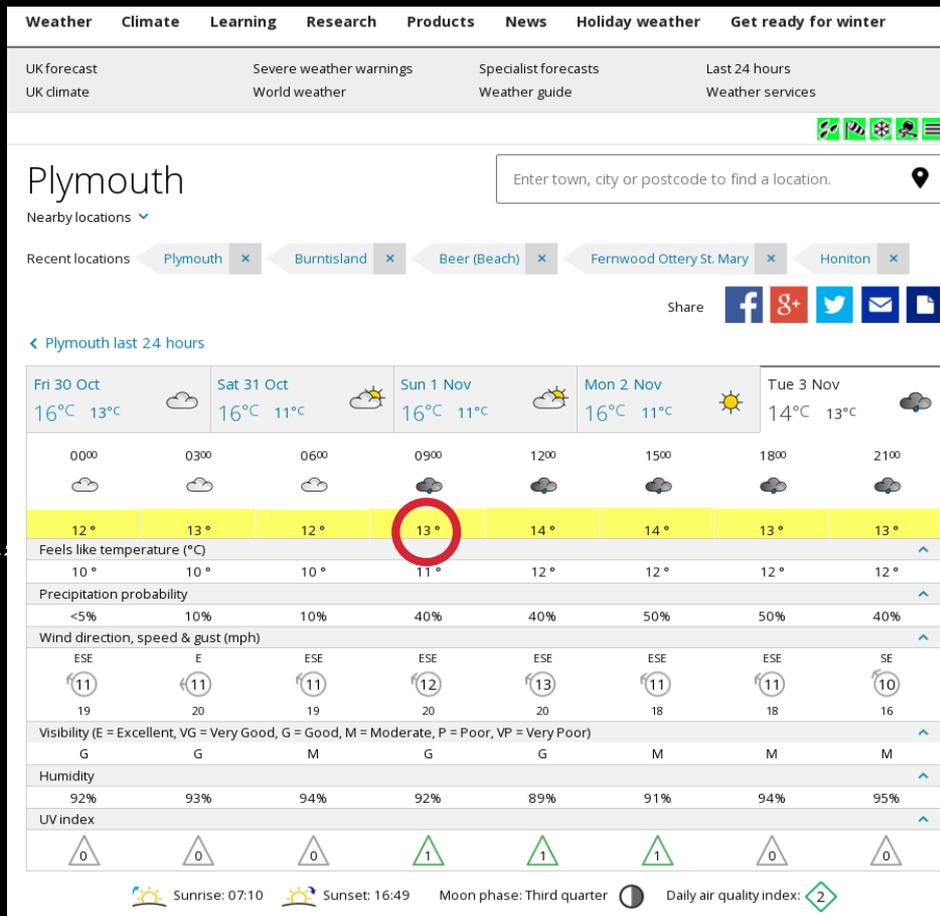
Verification of Relative-Extreme Events

We've tried ...

1. *SEDI (deterministic method)*

    a) *Consider each forecast percentile separately*

    b) *Choose one based on its frequency bias*

    c) *& use it as the 'deterministic' forecast*

2. *CRPS (probabilistic method)*

    a) *Compare the CDF of the forecast with the CDF of the observation (a Heaviside function)*

    b) *Restrict CRPS to only examine extreme events by partitioning or thresholding*

    c) *Integrate numerically over forecast percentiles*

Verification of Relative-Extreme Events

developed specifically to verify rare events

1. SEDI

Symmetric Extremal Dependency Index - Ferro & Stephenson (2011)

$$SEDI = \frac{\log F - \log H - \log(1-F) + \log(1-H)}{\log F + \log H + \log(1-F) + \log(1-H)}$$

Hit rate $H = \dfrac{a}{a+c}$ False alarm rate $F = \dfrac{b}{b+d}$

Condition: frequency bias $FB = \dfrac{a+b}{a+c}.$ must be ≈ 1

We've used the forecast percentiles as a means of calibration....

.... at each site we verify the percentile that minimises |FB - 1|

Verification of Relative-Extreme Events

2. CRPS

Continuous ranked probability score

Calculates the difference between

- forecast CDF
- observed CDF

We are only interested in extremes...

... but the CRPS evaluates everything

So we've tried...

1. excluding non-extremes inside the integrand (Gneiting & Ranjan, 2011)

2. Partitioning using the CRPS$^{o|f}$

...using each site-specific climatology to calculate skill scores

e.g. for maximum temperature

$$CDF\ O(t) = \begin{cases} 0 \text{ for all } t < ob \\ \\ 1 \text{ for all } t \geq ob \end{cases}$$

Heaviside function using the relative-extreme thresholds

$$CRPS = \int \left(F(t) - O(t)\right)^2 w(t)$$

CRPS measures the difference

100%

31 31.5 32 32.5 33 33.5 34 34.5 35 35.5
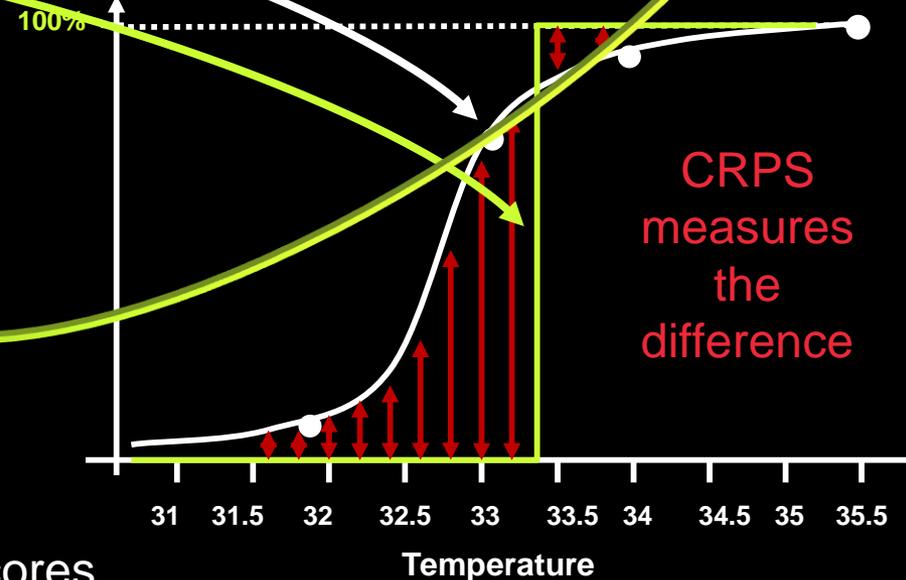
Temperature

Verification of Relative-Extreme Events

Minimum:

Minimum:

Winter night-time minimum temperatures

Summer day-time maximum temperatures

4.4th percentile → 4 nights / yr

95.6th percentile → 4 days / yr
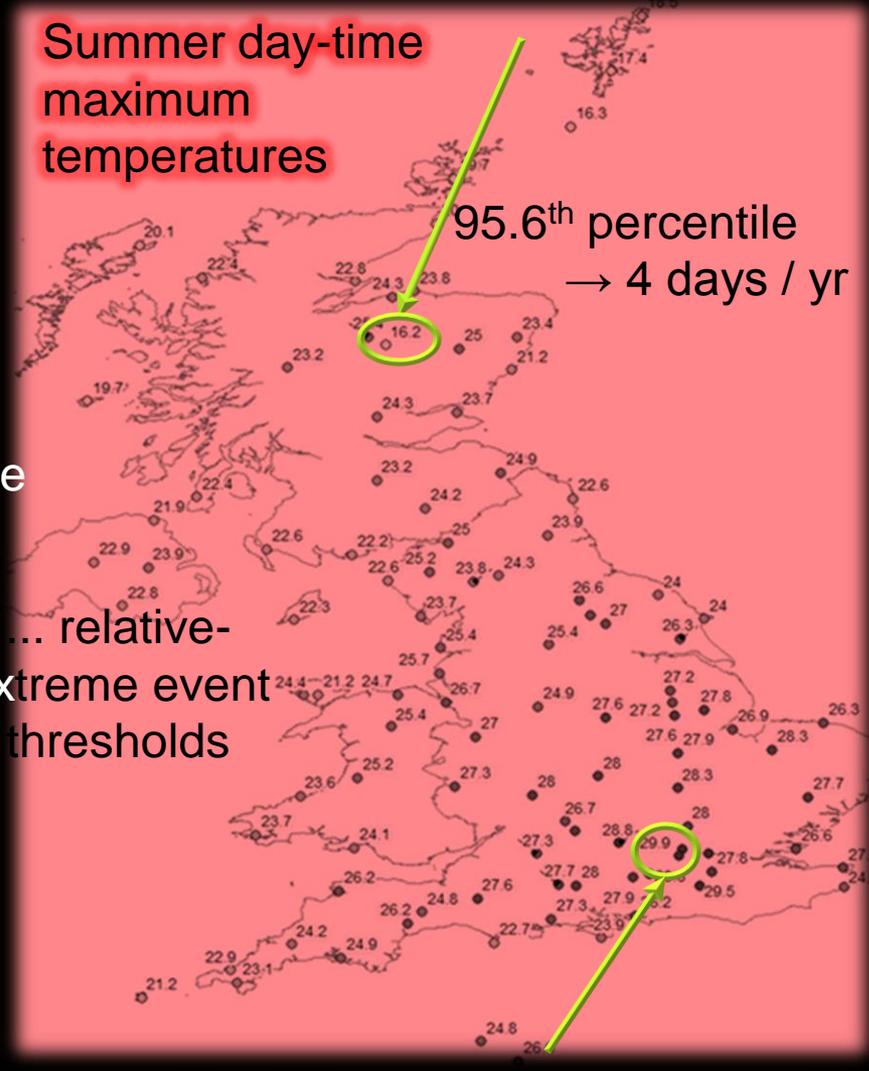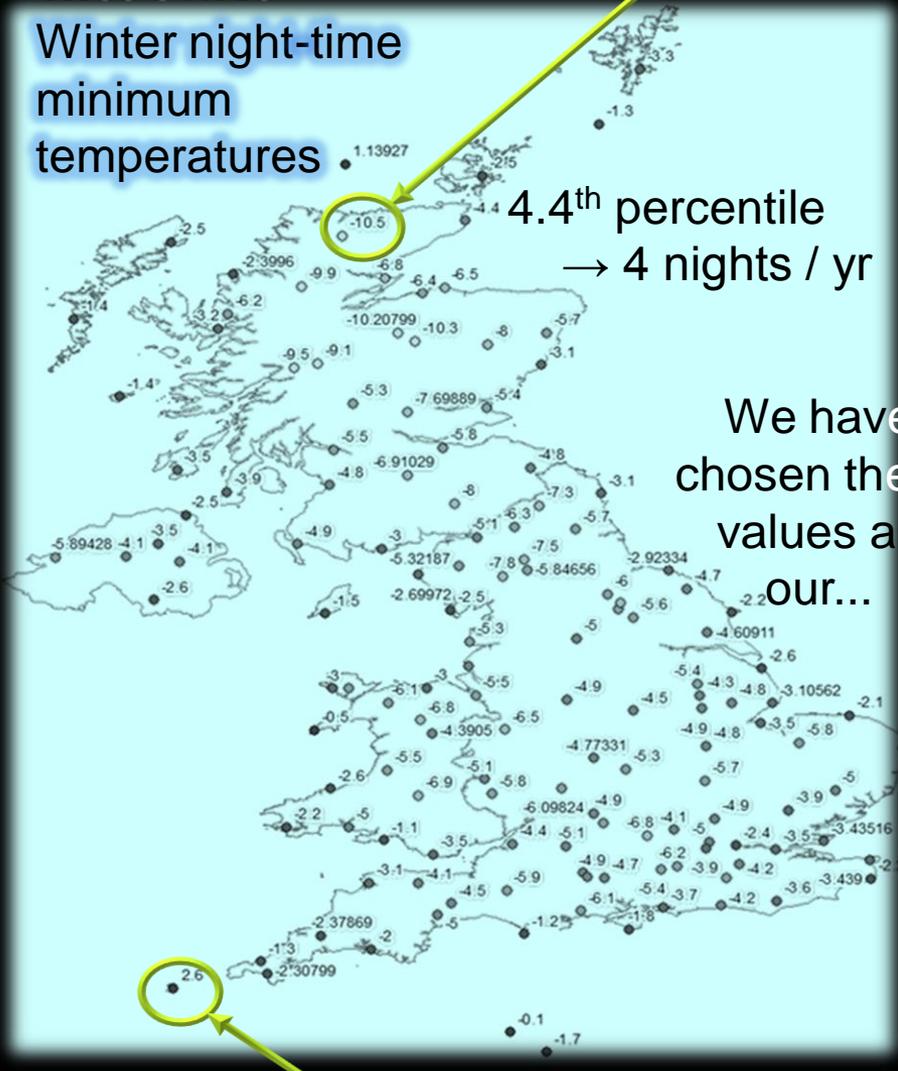
We have chosen these values as our...

... relative-extreme event thresholds

Maximum:

Maximum:

Verification of Relative-Extreme Events
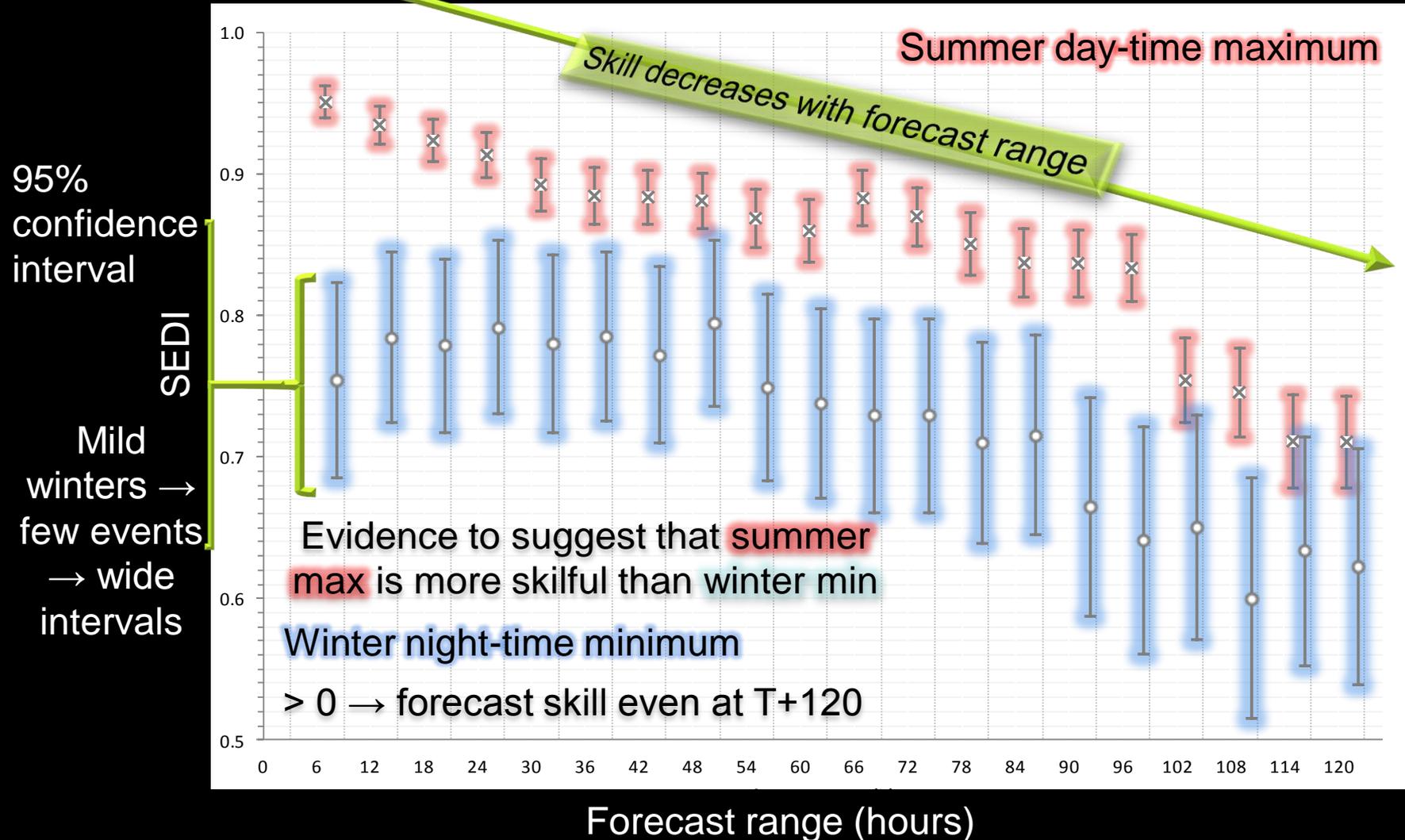
On verifying the last two winters & summers....

© Crown copyright   Met Office

Verification of Relative-Extreme Events

On verifying the last two winters & summers....

Summer day-time maximum

FromT+24: little evidence that summer max is more skilful than winter min

Winter night-time minimum

> 0 → forecast more skilful than site-specific climatology

Reference = site-specific climate
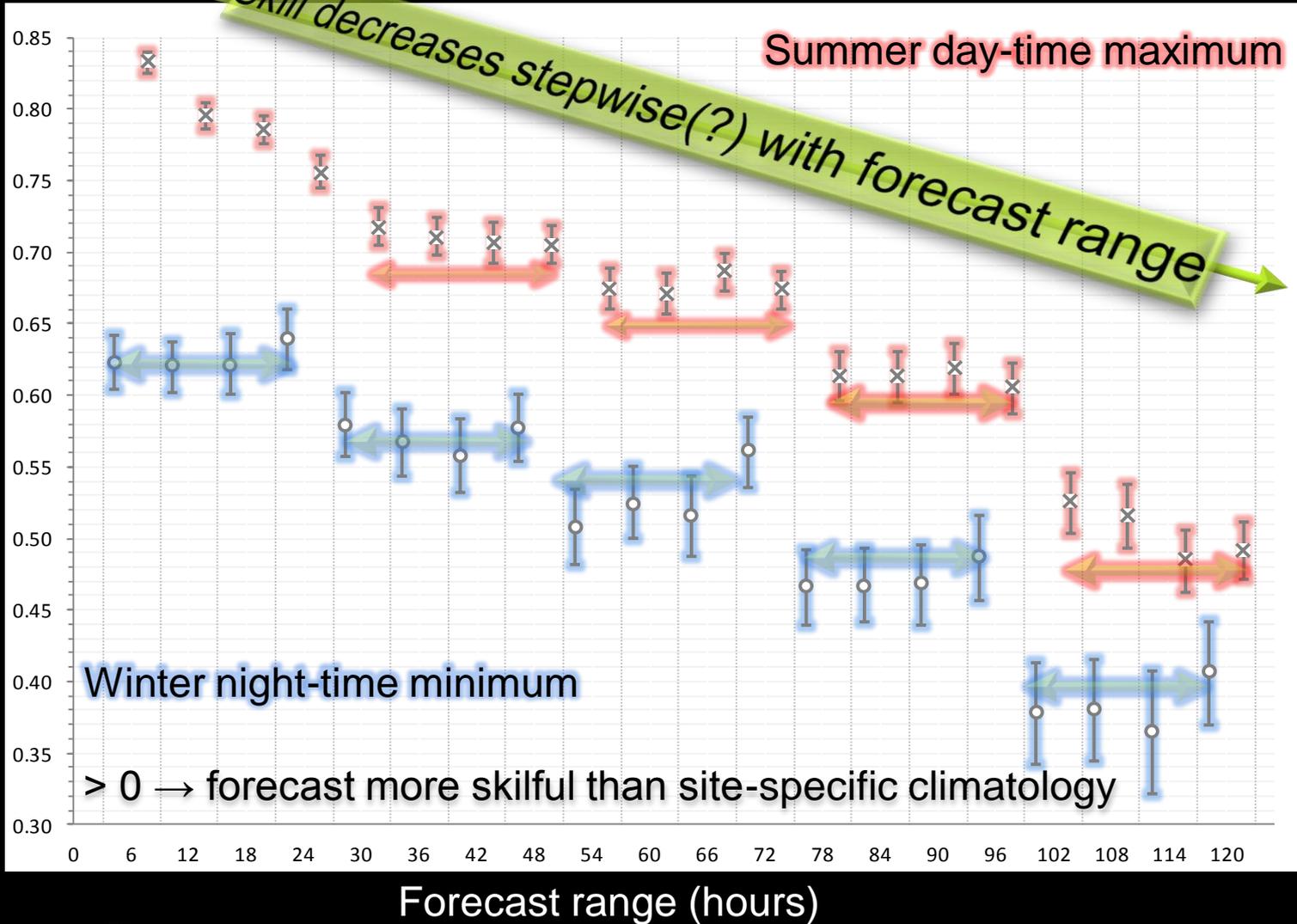
twCRPSS

Forecast range (hours)

© Crown copyright   Met Office

Verification of Relative-Extreme Events

On verifying the last two winters & summers....

Verification of Relative-Extreme Events

On verifying the last two winters & summers....

Summer day-time maximum temperatures
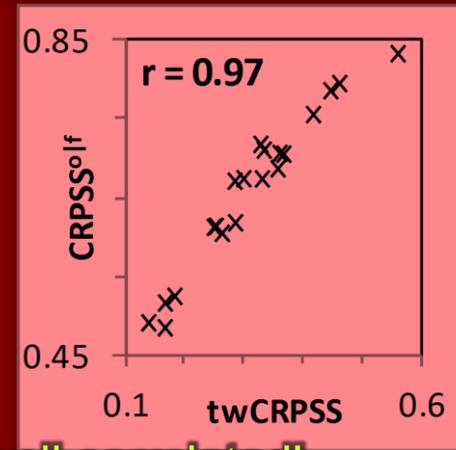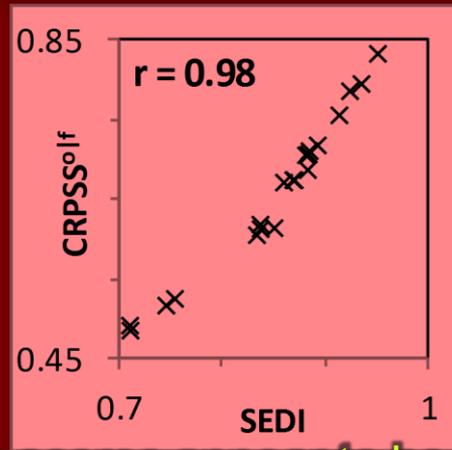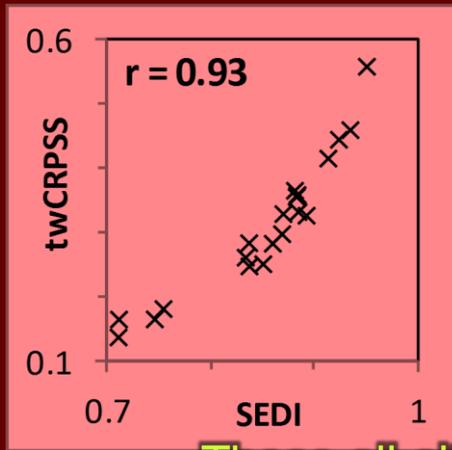
These all-site scores appear to be well correlated!
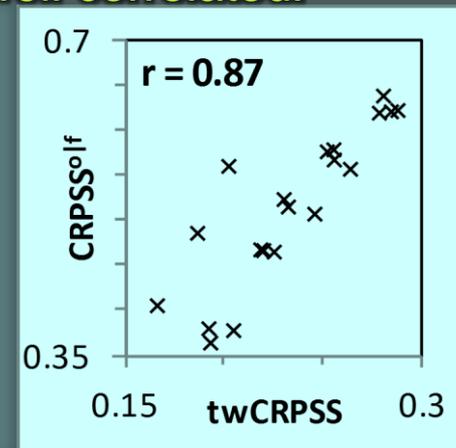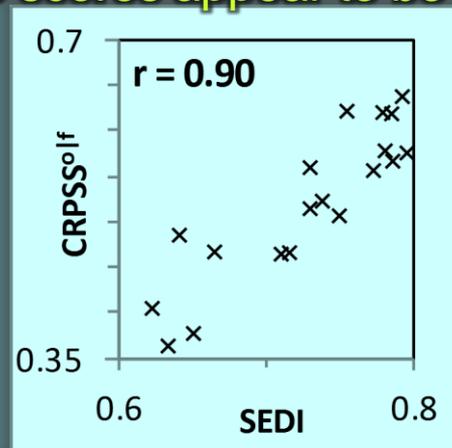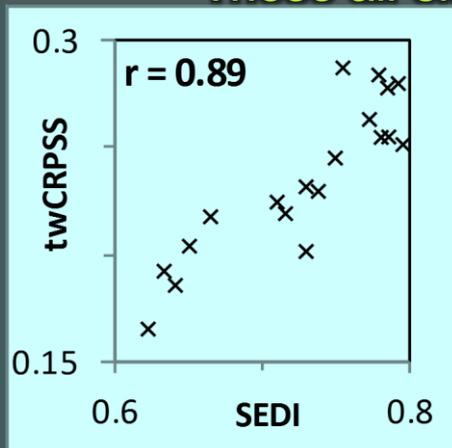
Winter night-time minimum temperatures

Verification of Relative-Extreme Events

Met Office

On verifying the last two winters & summers....

Summer day-time maximum temperatures

Above diagonal: observed more than forecast

When event was forecast to occur 3/10 times it actually occurred ~5/10 times

#events when forecast probability = 80%

More events when T+120 forecast probability is low

No events when T+120 forecast probability is high
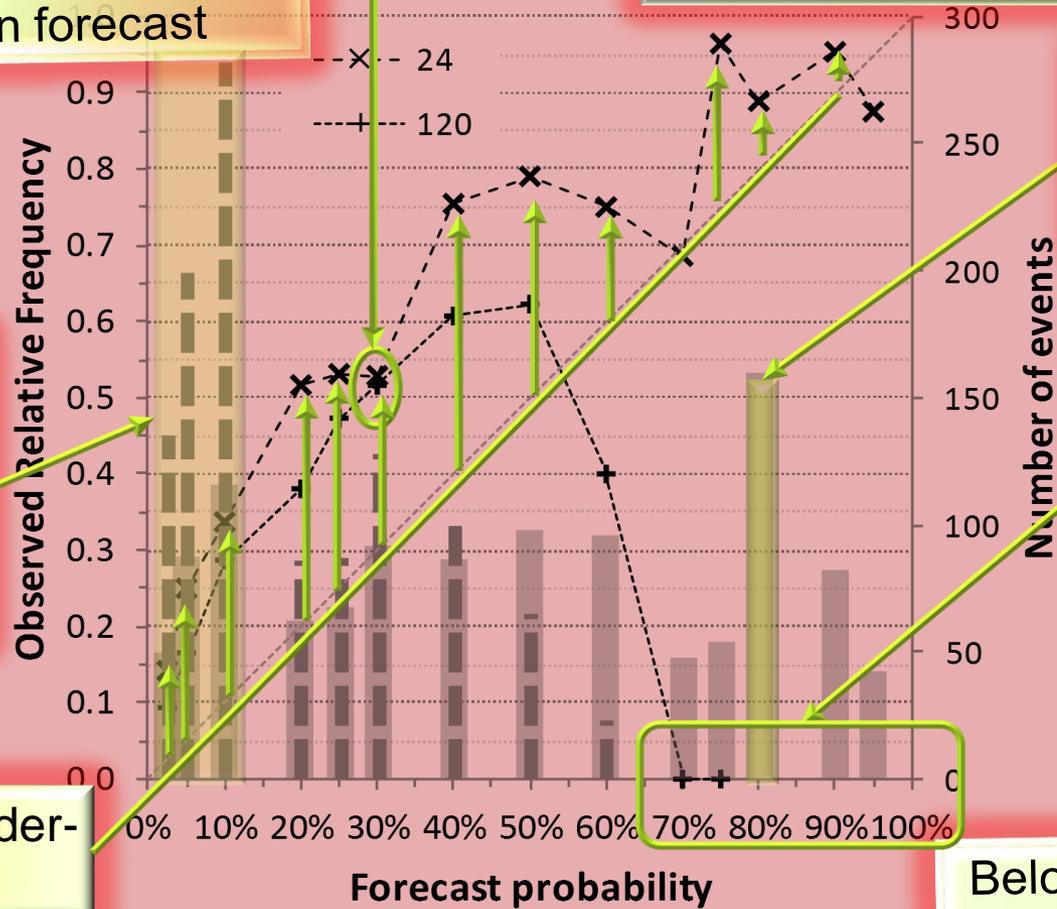
Tends to under-forecast

Below diagonal: forecast more than observed

Chart axes: Observed Relative Frequency (left, 0.0–1.0), Number of events (right, 0–300), Forecast probability (bottom, 0%–100%). Legend: ×- 24, +- 120

© Crown copyright   Met Office

Verification of Relative-Extreme Events

On verifying the last two winters & summers....

Winter night-time minimum temperatures

Few events occur when forecasts probability > 30%

Above diagonal: observed more than forecast
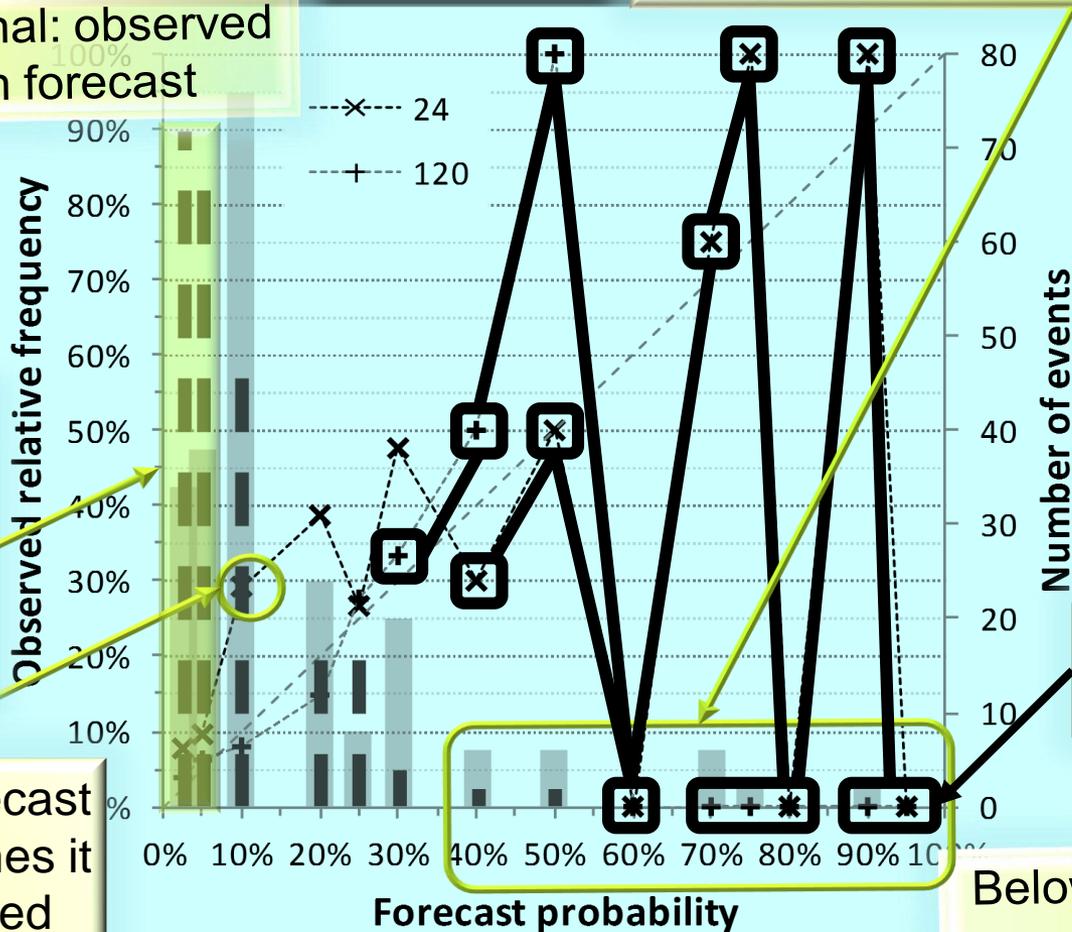
More events when T+120 forecast probability is low

When event forecast to occur 1/10 times it actually occurred 3/10 times

Far fewer cold events

few events: values very inaccurate

Below diagonal: forecast more than observed

Observed relative frequency (y-axis left): 100%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 0%

Number of events (y-axis right): 80, 70, 60, 50, 40, 30, 20, 10, 0

Legend: ×--- 24   +--- 120

Forecast probability (x-axis): 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Met Office

Summary...

- SEDI – verifies just 1 forecast percentile (as deterministic solution)

- twCRPSS – verifies only the percentiles that exceed the relative-
extreme event thresholds

- CRPSS$^{o|f}$ – verifies all percentiles **BUT** only when a relative-extreme
was forecast or observed

BUT they all tell a similar story...

Not significant for twCRPSS

- Skill decreases with forecast range

- More skill at forecasting summer maximums than winter minimums

- Evidence (95% level) that forecast is better than site-specific
climatology at predicting relative-extremes

Reliability

- Few high probability extreme event forecasts

- Tendency to under-forecast the probability

© Copyright David Martin and licensed for reuse under Creative Commons Licence 2.0 Generic

Thank you for listening
Questions....