Verifying deterministic forecasts of extreme events

Chris Ferro

Walker Institute
Department of Meteorology
University of Reading

- Probability model
- Interpretation and estimation
- Application to precipitation forecasts

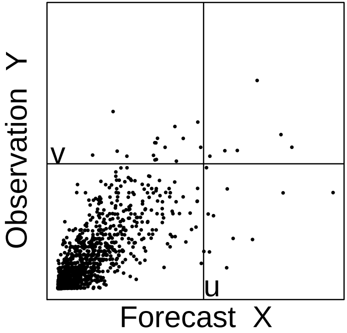
Direct approach

	Observed	Not Obs.	
Forecasted	a	b	$a+b$
Not Forecasted	c	d	$c+d$
	$a+c$	$b+d$	n

$$\text{Hit rate} = \frac{a}{a+c}$$

Forecast if $X > u$

Observe if $Y > v$



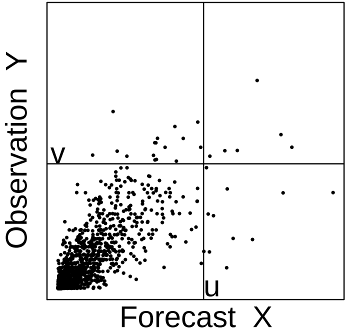
Probability approach

	Observed	Not Obs.	
Forecasted	$\Pr(X > u, Y > v)$	$*$	$\Pr(X > u)$
Not Forecasted	$*$	$*$	$*$
	$\Pr(Y > v)$	$*$	1

Hit rate $= \Pr(X > u \mid Y > v)$

Forecast if $X > u$

Observe if $Y > v$



Probability model

Imagine choosing u so that

$$\Pr(X > u) = \Pr(Y > v) =: p \qquad \text{(base rate)}$$

Extreme-value theory implies

$$\Pr(X > u, Y > v) = \kappa p^{1/\eta} \qquad \text{for small } p$$

under weak conditions.

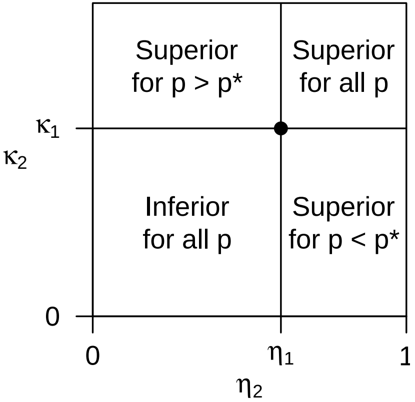
Ledford & Tawn (1996, Biometrika)

Interpretation

	Observed	Not Observed	
Forecasted	$\kappa p^{1/\eta}$	$*$	p
Not Forecasted	$*$	$1 - 2p + \kappa p^{1/\eta}$	$*$
	p	$*$	1

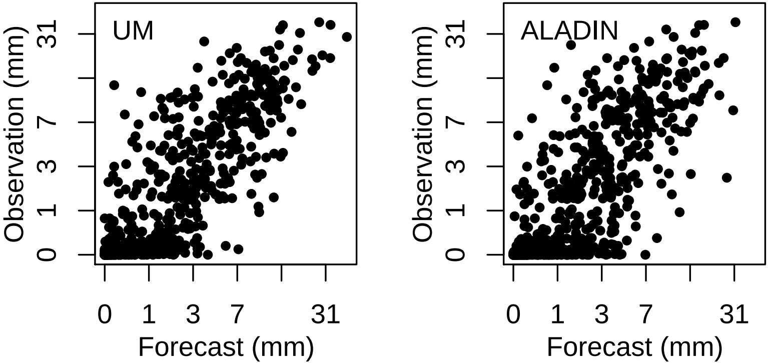
$$\text{Hit rate} = \kappa p^{1/\eta - 1}$$

$$\text{EDS} = 2\eta - 1$$



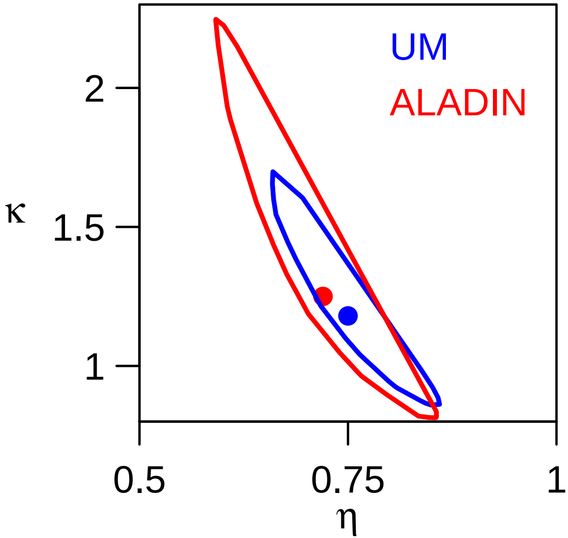
Daily precipitation: mid-Wales, 1 Jan 05 – 11 Nov 06

Thanks to Marion Mittermaier

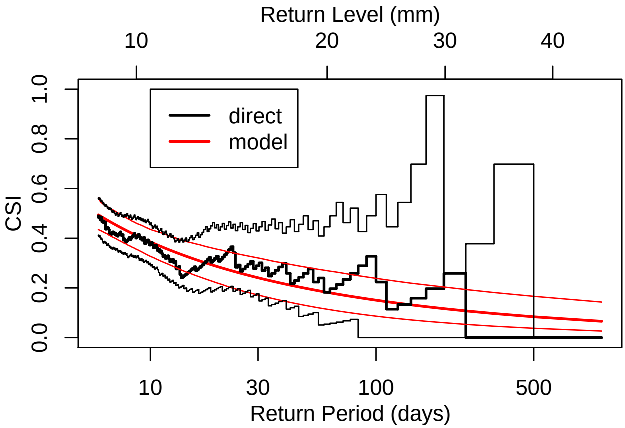


- Maximum-likelihood estimates of η and κ based on ranks
- Threshold choice and model assumptions

Parameter estimates



Verification measures



- Direct estimates degenerate for rare events
- Model estimates change smoothly and are more precise

Conclusion

- Deterministic forecasts of rare, extreme events
- Only two parameters are needed to describe how the quality of calibrated forecasts changes with base rate
- The model gives more precise estimates of forecast quality

Paper and R code available at

www.met.rdg.ac.uk/∼sws02caf

c.a.t.ferro@reading.ac.uk

Appendix

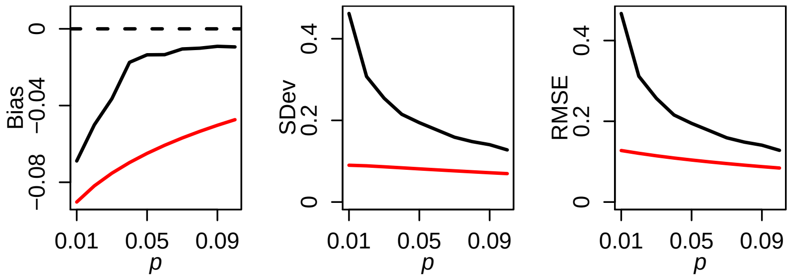
Simulation study

Model theory

Limiting behaviour of verification measures

Simulation study

- Bivariate Normal data: correlation 0.8
- Direct and model estimates of hit rate



Model theory – 1

Imagine choosing u so that

$$\Pr(X > u) = \Pr(Y > v) =: p \qquad \text{(base rate)}.$$

Define

$$\tilde{X} = -\log[1 - F(X)] \quad \text{where} \quad F(x) = \Pr(X \leq x)$$
$$\tilde{Y} = -\log[1 - G(Y)] \qquad\qquad G(y) = \Pr(Y \leq y)$$

Then $\tilde{X}$ and $\tilde{Y}$ are Exponential with unit means and

$$\begin{aligned} \Pr(X > u, Y > v) &= \Pr(\tilde{X} > -\log p, \tilde{Y} > -\log p) \\ &= \Pr(Z > -\log p) \end{aligned}$$

where $Z = \min\{\tilde{X}, \tilde{Y}\}$.

Model theory – 2

For $\tilde{X}$ and $\tilde{Y}$ Exponential with unit means and $Z = \min\{\tilde{X}, \tilde{Y}\}$,

$$\Pr(Z > z) = \begin{cases} \exp(-z) & \text{if } \tilde{X} \equiv \tilde{Y} \\ \exp(-2z) & \text{if } \tilde{X} \perp\!\!\!\perp \tilde{Y} \end{cases}$$

Assume

$$\Pr(Z > z) \sim \mathcal{L}(e^z)\exp(-z/\eta) \qquad \text{as } z \to \infty,$$

where $0 < \eta \leq 1$ and $\mathcal{L}(rt)/\mathcal{L}(r) \to 1$ as $r \to \infty$ for all $t > 0$.

e.g. $(X, Y) \sim$ Normal has $\eta = [1 + \text{cor}(X, Y)]/2$.

Ledford & Tawn (1996, Biometrika)

Model theory – 3

$\Pr(Z > z) \sim \mathcal{L}(e^z)\exp(-z/\eta)$ where $\mathcal{L}(rt)/\mathcal{L}(r) \to 1$ as $r \to \infty$.

For a high threshold w_0,

$$\begin{aligned}\Pr(Z > w_0 + z) &\approx \mathcal{L}(e^{w_0+z})\exp[-(w_0+z)/\eta] \\ &\approx \mathcal{L}(e^{w_0})\exp[-(w_0+z)/\eta]\end{aligned}$$

so model

$$\Pr(Z > z) = \kappa \exp(-z/\eta) \qquad \text{for all } z > w_0$$

i.e.

$$\Pr(Z > -\log p) = \kappa\, p^{1/\eta} \qquad \text{for all } p < \exp(-w_0).$$

Limiting behaviour of measures

$$\text{Hit rate} = \frac{a}{a+c} \sim \kappa p^{1/\eta - 1} \to \begin{cases} 0 & \text{if } \eta < 1 \\ \kappa & \text{if } \eta = 1 \end{cases}$$

$$\text{PC} = \frac{a+d}{n}, \qquad \text{PSS} = \frac{ad - bc}{(a+c)(b+d)}, \qquad \text{OR} = \frac{ad}{bc}$$

	$\eta < \frac{1}{2}$	$\eta = \frac{1}{2}$	$\eta > \frac{1}{2}$	$\eta = 1$
PC	$1 - 2p \uparrow 1$	$1 - 2p \uparrow 1$	$1 - 2p \uparrow 1$	$1 - 2\bar{\kappa}p \uparrow 1$
PSS	$-p \uparrow 0$	$-\bar{\kappa}p \updownarrow 0$	$\kappa p^{\delta-1} \downarrow 0$	$\kappa - \bar{\kappa}p \uparrow \kappa$
OR	$\kappa p^{\delta-2} \downarrow 0$	$\kappa - 2\kappa\bar{\kappa}p \updownarrow \kappa$	$\kappa p^{\delta-2} \uparrow \infty$	$\kappa/(\bar{\kappa}^2 p) \uparrow \infty$

where $\delta = 1/\eta$ and $\bar{\kappa} = 1 - \kappa$

Contradictory skill scores?

ERA-40 daily rainfall forecasts: $\eta = 0.81$, $\kappa = 1.16$