Predict extreme influenza epidemics

Maud Thomas[1], Magali Lemaitre[2,3], Mark L. Wilson[4], Cécile Viboud[3], Youri Yordanov[5], Hans Wackernagel[6] and Fabrice Carrat[2,5]

[1]LSTA-LPMA, UPMC [2] iPLESP, UPMC, Inserm [3] FIC, NIH [4] University of Michigan [5] APHP [6] MINES ParisTech

Influenza

- High morbidity
- High mortality: $\geq$ 500,000 deaths/year worldwide
- 3 different types of viruses $\Rightarrow$ yearly epidemics
- Central issue for resource planning in public health

Objective:
Predict severe or exceptional influenza epidemics using EVT

Extreme Value Theory (EVT) [2]

Main goals of EVT:

- Estimate the probability to exceed a value larger than the observed maximum
- Estimate a very high quantile

Fundamental result:
The limit distribution of the maximum of n i.i.d. random variables belongs to a three-parameter Generalized Extreme Value (GEV) family

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

where μ, σ and ξ are localization scale and shape parameters, respectively.
Three classes of GEV:

- Fréchet domain ($\xi > 0$) heavy-tailed distributions
- Gumbel domain ($\xi = 0$) light-tailed distributions
- Weibull domain ($\xi < 0$) finite-tailed distributions

Block maxima method:

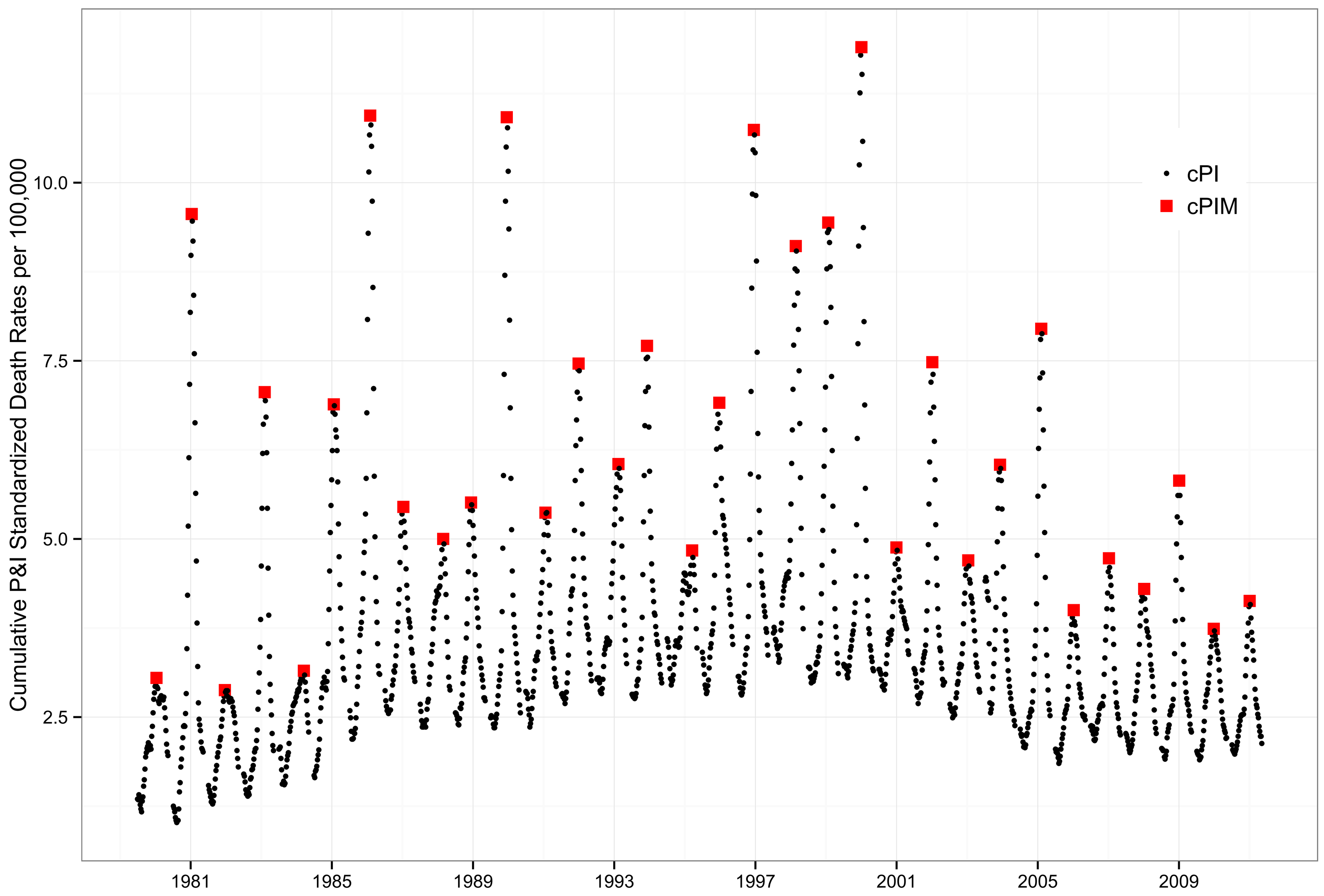
- $X_1, \ldots, X_{mn}$ i.i.d.

$$\underbrace{X_1, \ldots, X_n}_{M_1}, \quad \underbrace{X_{n+1}, \ldots, X_{2n}}_{M_2}, \quad \cdots \quad \underbrace{X_{n(m-1)+1}, \ldots, X_{mn}}_{M_m}$$

- Fit a GEV to the series $M_1, \ldots, M_m$

Pneumonia and influenza (P&I) mortality data

- Specific indicator of influenza mortality
- Weekly number of P&I deaths in France from July 1979 to June 2011 $\Rightarrow$ weekly standardized rates (Cépidc)
- Cumulative rates of P&I mortality (cPI) = sum of weekly P&I mortality rates over eight consecutive weeks using a moving window through the entire time series
- $cPIM$ = annual maxima of cPI

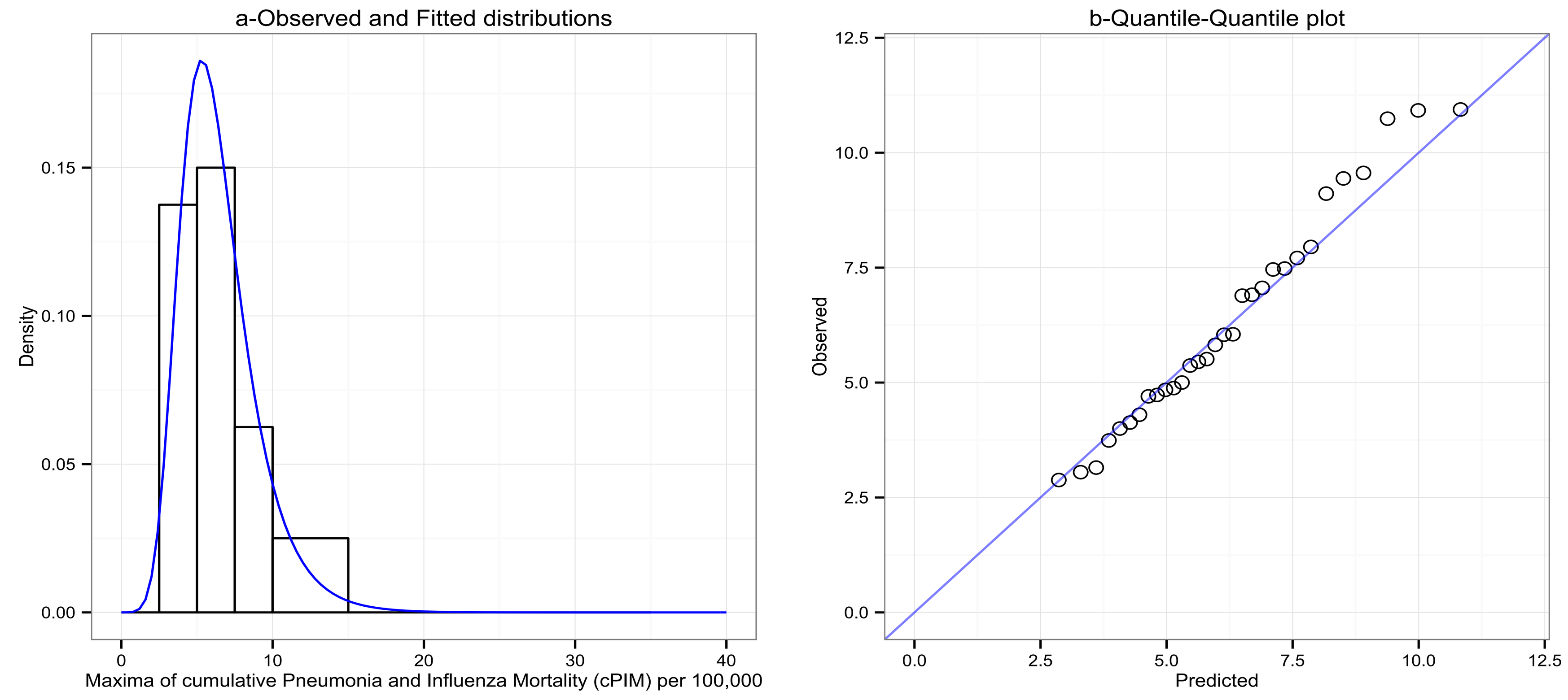


Fitting a GEV distribution to $cPIM$

Parameter estimates:

	μ	σ	ξ
Estimates	5.33	1.97	0.004
95% Confidence interval	$[4.51, 6.14]$	$[1.35, 2.59]$	$[-0.36, 0.37]$

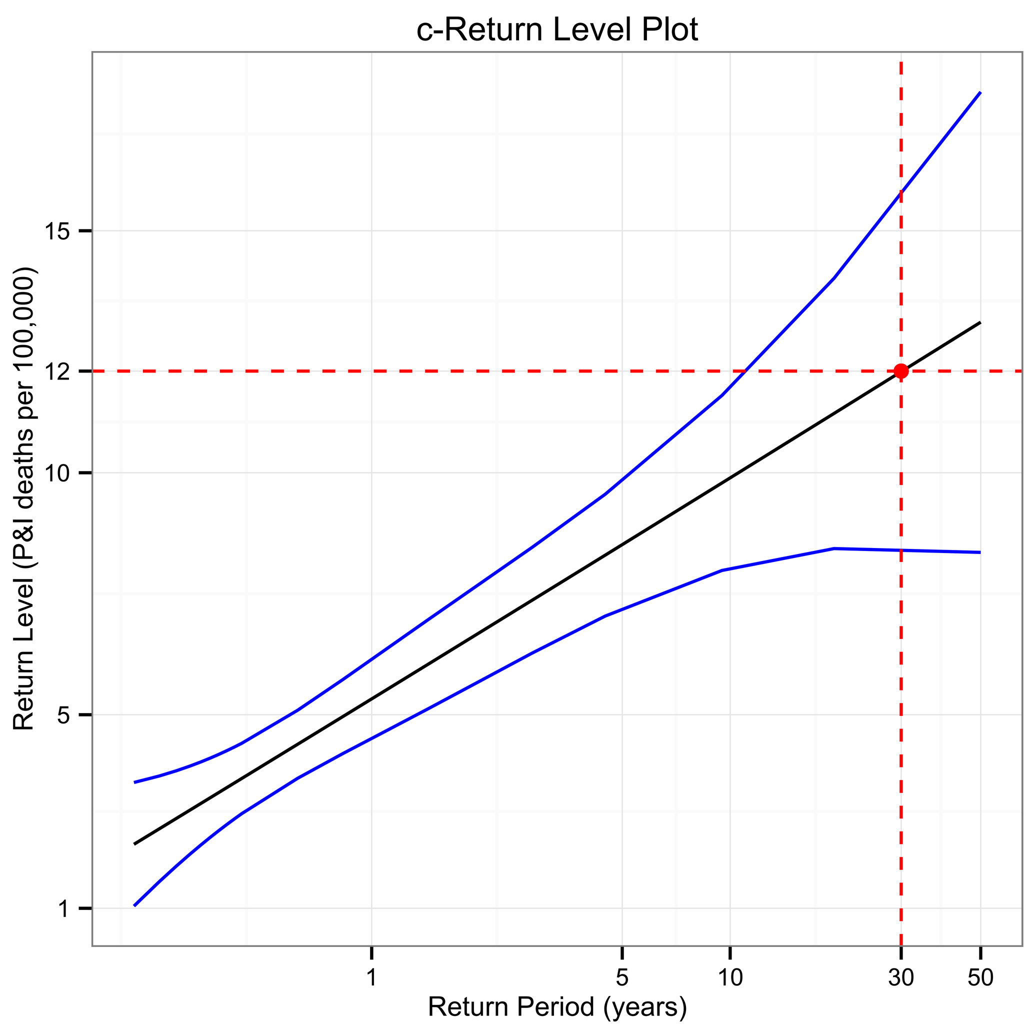
- $\gamma \approx 0 \Rightarrow$ Gumbel domain
- Fit seems correct: empirical and fitted cPIM distributions and QQ-plots.



Return levels

- $1/p$-return level = $(1-p)$-quantile z_p of a GEV
- **Interpretation**: z_p is expected to be exceeded on average once every $1/p$ years
- z_p has a closed form expression

$$z_p = \begin{cases} \mu - -\frac{\sigma}{\xi}\left[1-\{-\log(1-p)\}^{-\xi}\right] & \text{for } \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\} & \text{for } \xi = 0 \end{cases}$$



- Linear aspect confirms $\gamma \approx 0$
- **Interpretation**: over the next 30 years, cPIM should exceed 12 deaths per 100,000

Results

- Risk to exceed the observed highest $cPIM$
  $\hookrightarrow$ 1.11% chance that the $cPIM$ will exceed 14 deaths per 100,000 next year
- Value exceeded in average once every n year
  $\hookrightarrow$ over the next 30 years, $cPIM$ should exceed 12 deaths per 100,000 once

References

[1] Thomas M., Lemaitre M., Wilson M.L., Viboud C., Yordanov Y., Wackernagel H. & Carrat F. Applications of Extreme Value Theory in Public Health. *Accepted in PlosOne*

[2] Coles S. an introduction to statistical modeling of extreme values. 2001