Basic penultimate Weibull–XIMIS implementation

Alison Peard

2025-10-21

The sub-asymptotic Weibull–extended method of independent storms (XIMIS) method was found to produce the best fits for extreme wind speeds in a non-mixed mechanism climate. It was found to outperform both asymptotic approaches and epoch maxima approaches.

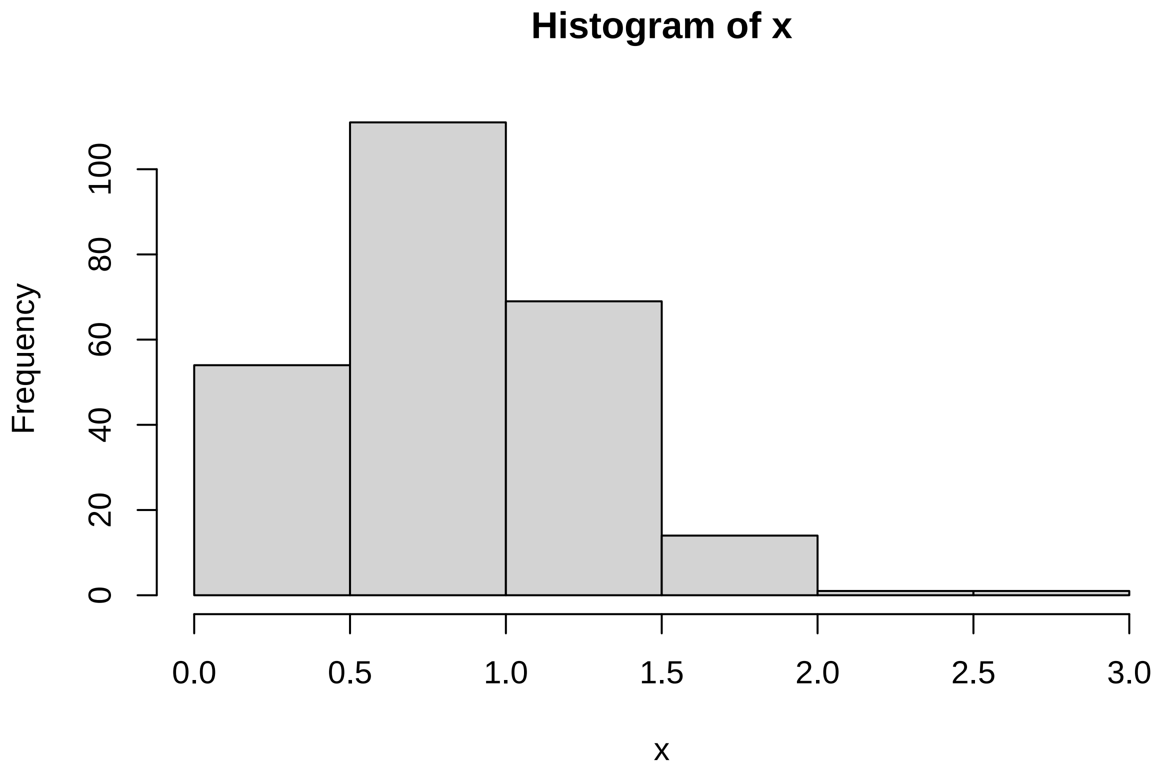
The improved performance over asymptotic methods is attributed to the slow convergence of extremes from a Weibull parent towards the Fisher–Tippett Type I (FT1) distribution. This means fitting GPD/GEVS to data from finite choices of threshold or block size assigns data to the wrong attractor. Indeed, when fitted to a finite sample, these models appear to give a better fit than the limiting distribution FT1 fit, but this is due to the third parameter in GEV/GPD models, which can handle the noise resulting from non-convergence.

Weibull–XIMIS is an alternative approach that instead fits the data to a penultimate FT1 distribution. This has been found to give better results. The main components of this approach are:

1. For an set of independent storm maxima from a simple climate x. (Obtaining this is discussed in [1].)
2. Fit a Weibull distribution to obtain the shape parameter ω.
3. Transform to be tail-equivalent to an exponential distribution using $z = x^{\omega}$. This accelerates convergence towards the FT1 asymptote.
4. Obtain Gringorten estimator in Eqs. (9)–(10) for the plotting positions y and standard errors σ of z.
5. Use the relationship from Eq. (2) to get estimators of z: $\hat{z} = yD + U$, where D is the dispersion and U the mode.
6. Optimise the weighted least squares $\frac{1}{\sigma^2}(z - \hat{z})$ for (U, D)

Let's start by simulating 250 i.i.d. Weibull variables as a training set. Later, we will simulate 10,000 random variables as a test set and see how well the data is estimating extremes. In practice, real data is unlikely to be i.i.d. and will exhibit serial correlation. Cook & Harris deal with this, see Cook (1982) to start.

```r
set.seed(123)

R <- 250
x <- rweibull(n = R, scale = 1, shape = 2)

hist(x)
```



Now, fit a Weibull distribution to these using maximum likelihood estimation.

```
par <-c(0.5, 0.5)

nll <- function(x, par) {
  shape = par[1]
  scale = par[2]
  -sum(dweibull(x, shape = shape, scale = scale, log = TRUE))
}

fit <- optim(par, nll, x = x, method = "L-BFGS-B", lower = c(0.01, 0.01))
```

The fitted shape parameter can be used to transform the Weibull to be tail-equivalent to Exponential using $Z = V^{\omega}$, leading to faster convergence to FT1.

```
omega <- fit$par[1]
z <- x^omega
```

For POT values, the estimators of the mean plotting position ($\bar{y}$) and statistical variance $\sigma^2(y_V)$ for the XIMIS model [2] were derived from asymptotic theory as

$$\begin{array}{cc} \hat{y}_1 = \gamma + \ln(R) & \hat{y}_{m+1} = \hat{y}_m - \frac{1}{m} \\ \sigma^2_{y_1} = \frac{\pi^2}{6} & \sigma^2_{y_{m+1}} = \sigma^2_{y_m} - \frac{1}{m^2} \end{array}$$

```
gamma <- 0.57721566490153286

z <- sort(z, decreasing = TRUE)
y <- vector(length=R)
sig2 <- vector(length=R);

y[1] <- gamma + log(R)
```

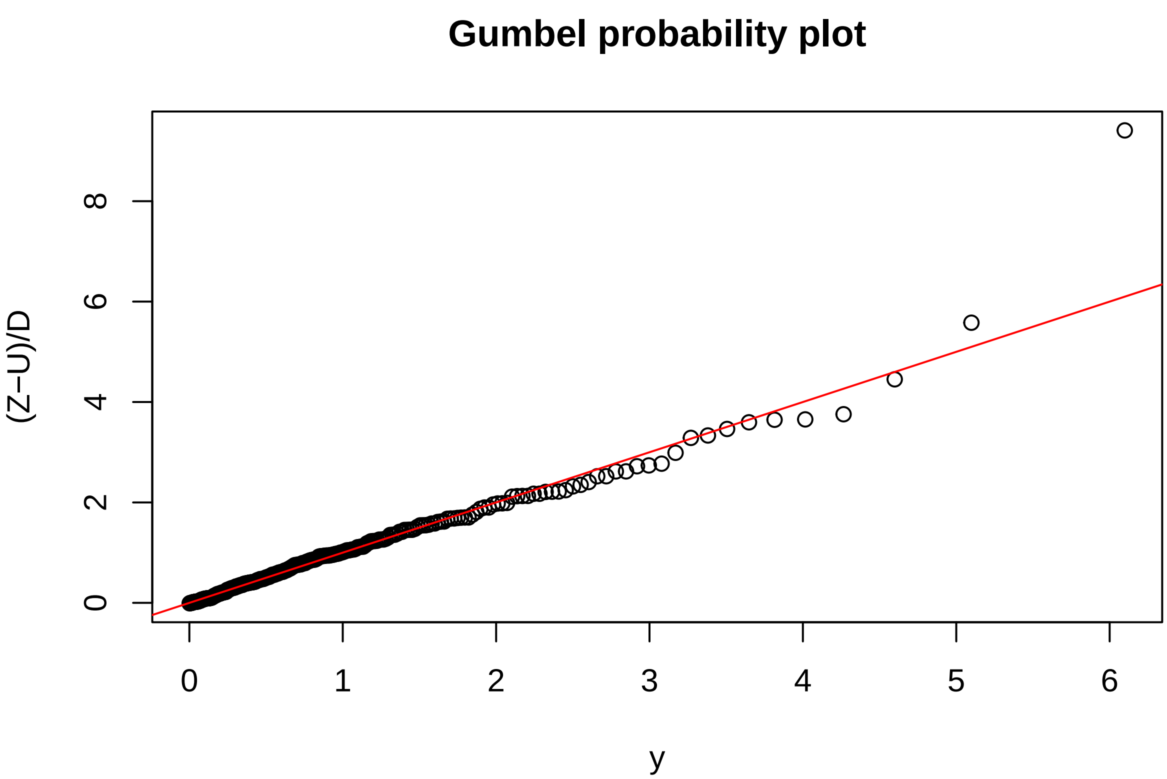
```
sig2[1] <- pi^2 / 6

for (m in seq(from = 2, to = R)) {
  y[m] <- y[m-1] - 1 / (m-1)
  sig2[m] <- sig2[m-1] - 1 / (m-1)^2
}
```

We optimise a least squares between the true z and its estimate from $\hat{z} = y \cdot D + U_T$, weighted by $1/\sigma^2$, as done in [2].

```
w <- 1 / sig2

residuals <- function(par, z, y, w) {
  u <- par[1]
  d <- par[2]
  z_hat <- y * d + u
  sum(w * (z - z_hat)^2)
}

fit2 <- optim(c(1,1), residuals, z=z, y=y, w=w)
u <- fit2$par[1]
d <- fit2$par[2]
```

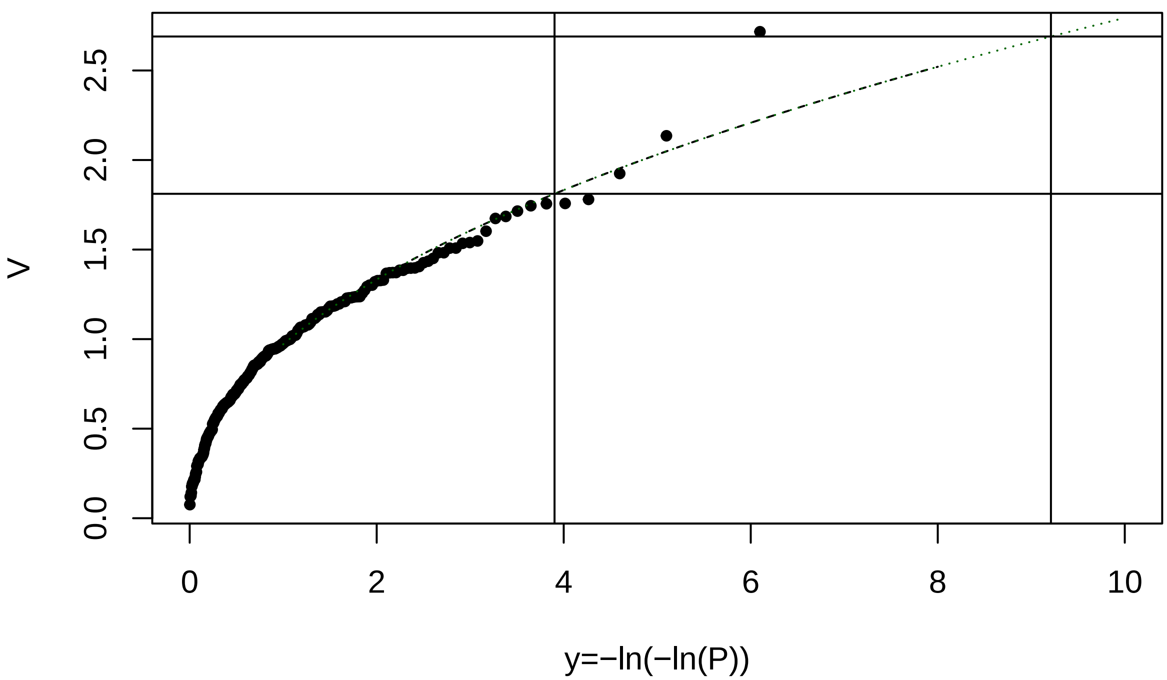
Now plot the estimates on a Type I (Gumbel) axis.

```
par(mfrow=c(1,1))
plot(y, (z - u) / d, main = "Gumbel probability plot",
     ylab = "(Z-U)/D", xlab = "y")
abline(0, 1, col = "red")
```



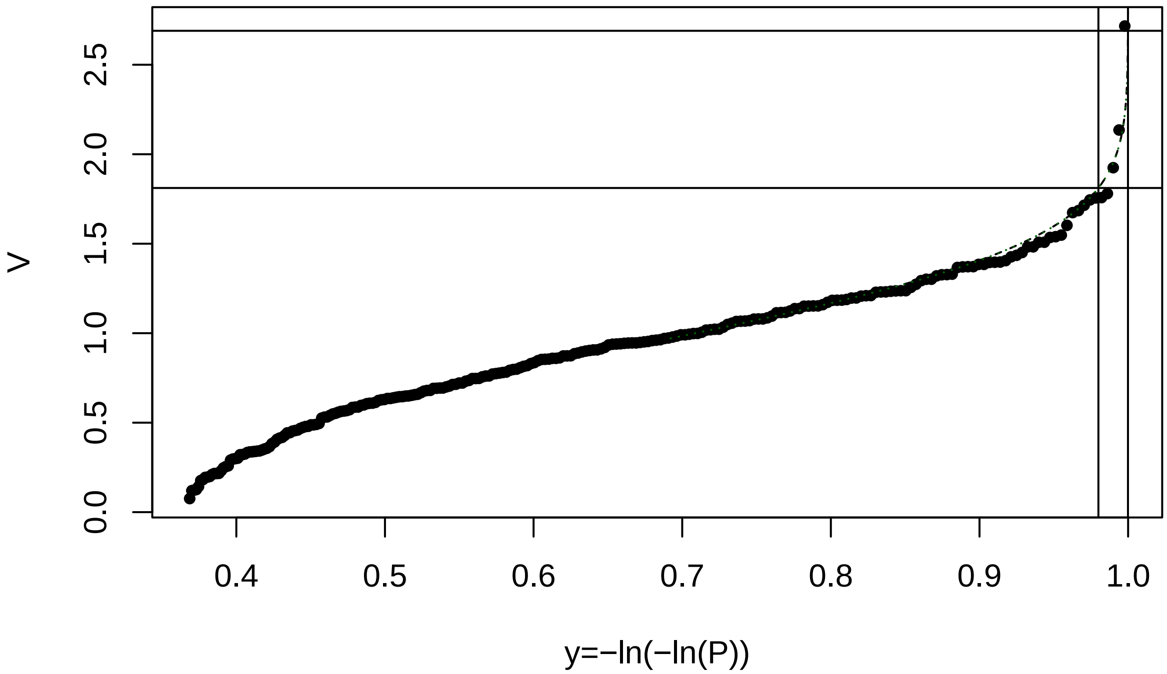
Comparing to Fig. 3 of Cook (2023).

```r
par(mfrow=c(1,1))

y_grid <- seq(0, 8, length = 1000)

#omega_true <- 2

plot(y, z^(1/omega), col="black", xlab = "y=-ln(-ln(P))", ylab = "V", xlim = c(0, 10), pch = 20) # data
lines(y_grid, (y_grid*d + u)^(1/omega), col="black", lty = 2) # fit

y_grid <- seq(1, 10, length = 1000)
x_grid <- (y_grid * d + u)^(1/omega)
lines(y_grid, x_grid, col="darkgreen", pch = 20, lty = 3)

# get y values for return periods
y50 <- -log(-log(1 - 1/50))
y10k <- -log(-log(1 - 1/10000))
abline(v=y50, col="black")
abline(v=y10k, col="black")

# calculate corresponding return levels
z50 <- (y50 * d + u)^(1/omega)
z10k <- (y10k * d + u)^(1/omega)
abline(h=z50, col="black")
abline(h=z10k, col="black")
```




```r
par(mfrow=c(1,1))

plot(exp(-exp(-y)), z^(1/omega), col="black", xlab = "y=-ln(-ln(P))", ylab = "V", pch = 20) # data
lines(exp(-exp(-y_grid)), (y_grid*d + u)^(1/omega), col="black", lty = 2) # fit

y_grid <- seq(1, 10, length = 1000)
x_grid <- (y_grid * d + u)^(1/omega)
```

```r
lines(exp(-exp(-y_grid)), x_grid, col="darkgreen", pch = 20, lty = 3)

# get y values for return periods
y50 <- -log(-log(1 - 1/50))
y10k <- -log(-log(1 - 1/10000))
abline(v=exp(-exp(-y50)), col="black")
abline(v=exp(-exp(-y10k)), col="black")

# calculate corresponding return levels
z50 <- (y50 * d + u)^(1/omega)
z10k <- (y10k * d + u)^(1/omega)
abline(h=z50, col="black")
abline(h=z10k, col="black")
```



Let's see how this new sampled extremes compare to the true distribution. *Need better tests, review Cook and Harris more.*

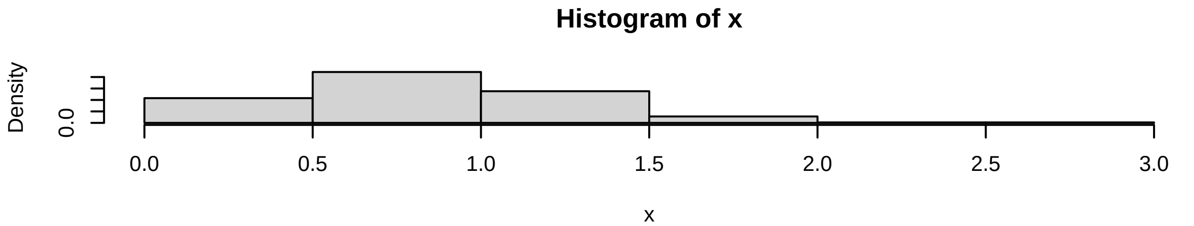
```r
range(y); range(omega)
```

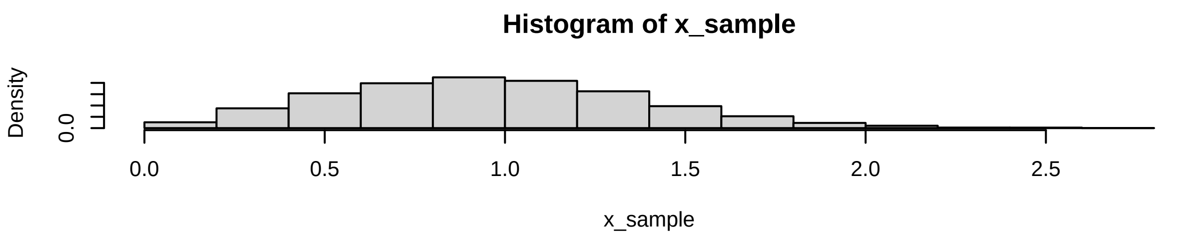
```
## [1] 0.002001333 6.098676583
```

```
## [1] 2.167627 2.167627
```

```r
# simulate new samples
n_test <- 10000
u_sample <- runif(n_test)#, min=0.0001, max=0.9999)
y_sample <- -log(-log(u_sample))
z_sample <- d * y_sample + u
x_sample <- z_sample^(1/omega)

x_test <- rweibull(n_test, scale = 1, shape = 2)

par(mfcol=c(3, 1))
hist(x, freq = FALSE)
```

```r
hist(x_test, freq = FALSE)
hist(x_sample, freq = FALSE)
```





References

1. Cook, N. J. (2023). Reliability of Extreme Wind Speeds Predicted by Extreme-Value Analysis. *Meteorology*, 2(3), 344–367.
2. Harris, R. I. (2009). XIMIS, a penultimate extreme value method suitable for all types of wind climate. *Journal of Wind Engineering and Industrial Aerodynamics*, 97(5–6), 271–286.