QueryMSC

Nic Annau

March 2019

1 QueryMSC

QueryMSC is a Python tool to query design values from Meteorological Service of Canada (MSC) database hosted at the Pacific Climate Impacts Consortium. This project is currently in development, and be found at pacificclimate

2 Methods

Design values are physical and statistical derivations from samples of meteorological data that describe a given location's climatology and help inform the *National Building Code of Canada.* Canada has a large suite of historical meteorological data that are used to derive design values.

Most design values in this project are derived based on a description found in *National Building Code of Canada Volume 1 Appendix C*.

3 Gumbel Distribution

Some non-trivial statistical methods regarding the use of Gumbel extreme value distributions for rainfall amounts are described here.

The Gumbel distribution, also known as the Generalized Extreme Value distribution Type-I, models maximum and minimum values of extreme values.

The Cumulative Distribution Function (CDF) of a real-valued random variable X, evaluated at x, is the probability that X will take a value less than or equal to x. The general form of the CDF for the right-skewed Gumbel distribution is given by:

$$F(X) = e^{-e^{-(X-\xi)/\alpha}}$$

Where ξ and α are the first two moments of the Gumbel distribution, also known as the *location* and *scale* parameters respectively. The value of ξ and α, in practice, are estimated by moments derived from X, which will be the topic of this section.

Some design values, such as *15 Min Rain 1/10* and *One Day Rain 1/50* both require a fitting of the Gumbel distribution to annual maximum 15 minute rainfall and daily rainfall at a given station. In simpler terms, it provides a statistically robust and resistant way of determining the likelihood of extreme weather events to occur within a given time frame based on a historical record of a given weather station.

4 Extreme Value Analysis

For the purposes of Extreme Value Analysis (EVA) in climatology, extreme weather events are characterized by their *return period*.

Let t_r be the return period in years, and then let f_r be the expected frequency, and be defined as $f_r = \frac{1}{tr}$ with units of years^{-1}. The design value x_v is (the magnitude of) the extreme weather event that has probability f_r of being equalled or exceeded in any one year. This can be expressed using the CDFs of the Gumbel distribution evaluated at x_v, the design value.

$$P(x_v \leq X \leq 1) = F(1) - F(x_v) = f_r$$

This is equivalent to:

$$P(0 \leq X \leq x_v) = F(x_v) - F(0) = 1 - f_r$$

The latter form will be used to simplify the final expression.

5 Estimating Gumbel Distribution Moments

To estimate ξ and α, *L-moments* are used following the methods described in Hosking [1990]. The main motivators for using *L-moments*, as opposed to more conventional estimators, such as the *Method of Moments* found in Newark et al. [1989], is that *L-moments* are robust and resistant in the presence of highly variable data, and very large outliers, meaning that the moments are not heavily influenced by rogue data. Although *L-moments* are not completely resistant, they are more so than simply taking the *mean* or *standard deviation.*

L-moments must be estimated from samples drawn from an unknown distribution, and in practice, this is done using U-statistics introduced by Hoeffding [1948].

For a Gumbel distribution, only the first two *L-moments* (l_1, l_2) need to be calculated. Let N be the sample size, and X_i be the ordered sample.

$$l_1 = N^{-1} \sum_{i=1}^{N} X_i$$

$$l_2 = \frac{1}{2} \binom{N}{2}^{-1} \sum_{i>j}\sum (X_{i:N} - X_{j:N})$$

Then if $\hat{\xi}$ and $\hat{\alpha}$ estimate ξ and α respectively, then $\hat{\xi}$ and $\hat{\alpha}$ can be expressed by

$$\hat{\xi} = l_1 - \gamma\hat{\alpha}$$

where γ is the Euler–Mascheroni constant and,

$$\hat{\alpha} = \frac{l_2}{\log 2}$$

6 Estimating the Extreme Weather Event

Substituting $\hat{\xi}$ and $\hat{\alpha}$ into

$$P(0 \leq X \leq x_v) = F(x_v) - F(0) = 1 - f_r$$

gives us

$$1 - f_r = e^{-e^{-(x_v - \hat{\xi})/\hat{\alpha}}} - e^{-e^{\hat{\xi}/\hat{\alpha}}}$$

Solving for x_v, the magnitude of the extreme weather event with probability of occurring in any given year exceeding f_r gives

$$x_v = \hat{\xi} - \hat{\alpha} \log\left(-\log\left((1 - f_r) + e^{-e^{\hat{\xi}/\hat{\alpha}}}\right)\right)$$

For a given station, each parameter in the above equation are typically calculated from a distribution of annual values.

References

Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19(3):293–325, 09 1948. doi: 10.1214/aoms/1177730196. URL `https://doi.org/10.1214/aoms/1177730196`.

J. R. M. Hosking. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124, 1990. ISSN 00359246. URL `http://www.jstor.org/stable/2345653`.

M. J. Newark, L. E. Welsh, R. J. Morris, and W. V. Dnes. Revised ground snow loads for the 1990 national building code of canada. *Canadian Journal of Civil Engineering*, 16(3):267–278, 1989. doi: 10.1139/l89-052. URL `https://doi.org/10.1139/l89-052`.