

Extreme Value Analysis

For attribution, please cite this work as:

Cal-Adapt. (2026). *Extreme Value Analysis*. Cal-Adapt.

<https://analytics.cal-adapt.org/scientific-guidance/extreme-value-analysis.html>

Overview

The [climakitae](#) Python package available on the Analytics Engine provides numerous tools to assist with [Extreme Value Analysis](#). This includes the calculation of [return probabilities](#), [return periods](#), and [return values](#). These terms are defined below within the context of meteorology and hydrology.

Return probability: The probability that a climate or weather event will exceed a specified magnitude in any given year¹.

- For example, if the maximum temperature at a location has a 10% chance of exceeding 105°F in any year, the annual **return probability** of that temperature is 0.10.

Return period: The reciprocal of the return probability. This is commonly interpreted as the average expected wait time between threshold exceedances¹⁻⁴.

- If the return probability of exceeding 105°F is 0.10 per year, the **return period** of that event is $1 \div 0.10$ or 10 years.

Return value: The magnitude of a climate or weather event associated with a given return probability. Also called a return level^{2,5}.

- The 105°F threshold that has a 10% chance of being exceeded in a given year would be the 1-in-10 year **return value** for maximum temperature.

Extreme value analysis can be applied within the [Global Warming Levels](#) approach to compare return variables (return probability, period, or value) across warming levels^{6,7}. A time-based approach can also be used.

Methodology

Block Maxima/Minima Series

The threshold tools on the Analytics Engine have implemented the Block Maxima/Minima Series (BMS) method for the stationary, single-variable case. In this method, a time series is divided into a series of evenly sized blocks. The collection of maximum or minimum values from these blocks is called the BMS. A statistical distribution is fitted to the BMS and then used to obtain return probabilities, values, or periods^{2,5}. A typical block size for this method is one year, giving an [Annual Maximum Series \(AMS\)](#). The AMS or BMS can be obtained for data with any sub-annual time frequency including hourly, daily, and monthly frequencies. This method cannot be used to study extremes that are expected to occur more frequently than once per block.



Effective Sample Size

Meteorological data is often temporally autocorrelated, meaning that the value at a given time depends to some degree on the value(s) at past times. When the data within a block of the BMS has too much autocorrelation it can bias the result of the extreme value analysis. One way to manage temporal autocorrelation in the blocks is to compute the [effective sample size \(ESS\)](#) within each block and increase the block size as needed to obtain a sufficiently large ESS. The ESS is a scaled-down sample size that reflects the loss of information in a sample due to autocorrelation⁸.

The *climakitae* package offers the option to check the effective sample size of blocks and generate a warning when the effective sample size is below a minimum threshold value (by default this threshold is 25). To calculate the ESS, a variance inflation factor V is first estimated for each block^{1,8} with:

$$V = 1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) r_k$$

Where n is the sample size and r is the autocorrelation at lag k . The ESS is then:

$$\text{ESS} = \frac{n}{V}$$

For large time series with more than 500 time points per block, a logarithmic lag sampling scheme is used to speed up calculation time.

Distribution Fit

Once the BMS has been constructed, a known statistical distribution can be fitted to model the distribution of extreme values. This distribution is then used to obtain the return values, periods, and probabilities.

The *climakitae* package offers the following probability distributions for modeling extreme values. While all of these distributions are frequently used in climate and hydrologic applications, the choice of distribution will depend on the specific application and underlying data^{2,9–11}. When in doubt, the [Generalized Extreme Value distribution](#) is a good starting point as it is commonly used with climate data and the BMS method^{2,5}.

- [Generalized extreme value \(GEV\) distribution](#)
- [Gumbel distribution](#)
- [Weibull distribution](#)
- [Generalized Pareto distribution](#)
- [Gamma distribution](#)
- [Pearson Type III distribution](#)

The [Kolmogorov-Smirnov \(K-S\) Goodness of Fit test](#) can be used to evaluate how well the selected probability distribution fits the data with the following null and alternative hypotheses^{1,12}:

- H_0 = input data is distributed according to the selected distribution
- H_1 = input data is *not* distributed according to the selected distribution



The corresponding p-value shows the significance of the fit:

- If p-value < α , reject H_0 . If p-value $\geq \alpha$, fail to reject H_0
- Where α is the significance level (often set to 0.05 or 5%)

The *climakitae* package can be used to obtain these p-values. Users should be aware that the K-S test is sensitive to sample size and p-values will change as sample size increases.

Once a distribution is selected, the [Maximum Likelihood Estimation](#) method is used to estimate the distribution parameters⁵. As currently implemented in *climakitae*, all parameters are constants.

The fitted probability distribution can then be used to calculate the return probability, return period, and return value as described below.

Return Probability

The [cumulative distribution function \(CDF\)](#) is calculated for the fitted probability distribution. Using the CDF, the exceedance probability is read for a desired return value¹. In the maximum case, the return probability is one minus the exceedance probability. In the minimum case, the return probability is equal to the exceedance probability.

- Maximum case: `return_probability = (1 - fitted_distribution.cdf(return_value))`
- Minimum case: `return_probability = fitted_distribution.cdf(return_value)`

Figure 1 provides a demonstration of this procedure for the Maximum case.

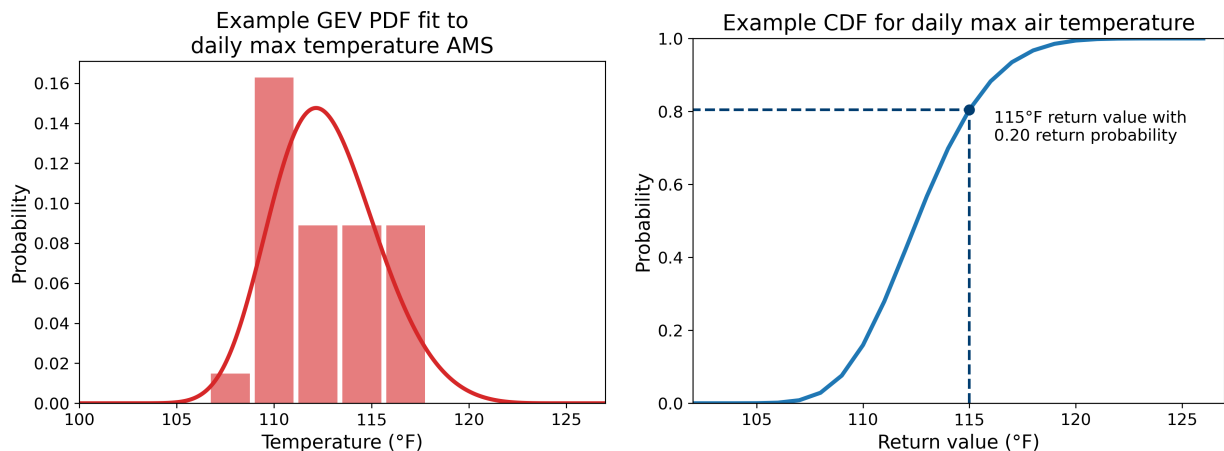


Figure 1: This figure shows the CDF (right) for a GEV distribution (left) fitted to an AMS of daily maximum air temperature from a simulation. The return probability of 0.2 is read for a return value of 115°F ($1 - 0.8 = 0.2$), meaning that there is a 20% chance of temperatures exceeding 115°F in any given year at this location.

Return Period

The return probability is obtained as described above, then inverted to obtain the return period. Using the example in Figure 1, the return period would be $1 / 0.2$ or 5 years.

`return_period = 1 / return_probability`



Return Value

A [percentage point function](#) (PPF, also known as an inverse CDF)⁵ for the fitted probability distribution is used to obtain the return value for a given event probability. In the maximum case, the event probability is one minus the return probability. In the minimum case, the event probability is equal to the return probability:

- Maximum case: `return_value = fitted_distribution.ppf(1 - return_probability)`
- Minimum case: `return_value = fitted_distribution.ppf(return_probability)`

Figure 2 provides a demonstration of this procedure for the Maximum case.

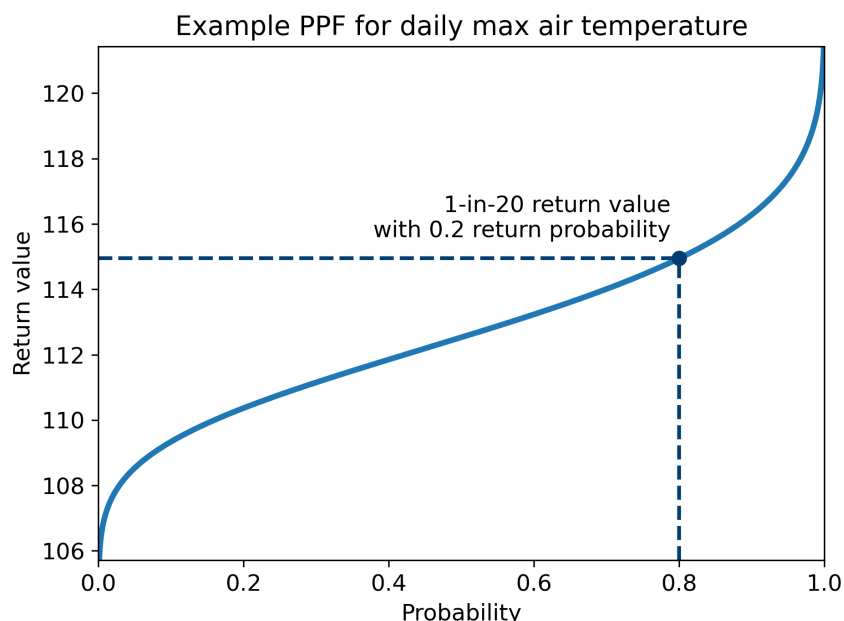


Figure 2: This figure shows the PPF for a GEV distribution fitted to an AMS of daily maximum air temperature from a simulation. The return value of 115°F is read for the return probability of 0.2 ($1 - 0.8 = 0.2$), meaning that there is a 20% chance of temperatures exceeding 115°F in any given year at this location.

Confidence Intervals

The bootstrap method is used to obtain confidence intervals on these return variables^{13,14}. This involves pulling random samples (with replacement) of the same sample size from the original BMS series and obtaining return variable estimates from those samples. *Climakitae* allows users to adjust the number of bootstrap samples, with 50–200 often being sufficient to obtain confidence limits¹³. Percentiles are used to select the confidence limits from the bootstrapped estimates — for example, the 95% confidence intervals are obtained from the 2.5 and 97.5 percentile values of these estimates. These confidence intervals represent uncertainty due to internal variability and sampling error, but there are other sources of uncertainty in the return variable results such as model physics and scenario uncertainty that are not captured by these confidence intervals¹⁵.



Aggregation

This section describes some general ways in which extreme value datasets can be aggregated. Users should always consider which methods are most appropriate for their specific applications.

Time Aggregation

Time aggregation may be used on input data as long as the final version of the input data is a time series. For example, an hourly precipitation time series could be resampled to a daily sum that is then used in the extreme value analysis.

Spatial Aggregation

Spatial aggregation can be applied prior to creating the BMS or after calculating extreme value results. For gridded data, *climakitae* threshold tools include an option to stack time series from multiple grid cells to increase sample size^{9,10}. Alternatively, summary statistics such as spatial mean and median can be applied to a region to create a BMS for that entire region. Such statistics can also be applied to return variables, for example to obtain an average regional return value for a 1-in-10 event that has been calculated on a gridcell basis.

The stage at which spatial aggregation should be applied will depend on the variable used and research question being asked. For example, one investigation of basin-wide snow volume may require a single, basin-wide time series per simulation. A different investigation of extreme precipitation could focus on the most extreme return values in the region and choose to look at a spatial maximum of the return value. Users do not need to use any spatial aggregation and could instead look at single stations or maps of values.

Simulation Aggregation

When working with the [CMIP6](#) GCM data and tools provided on AE, extreme value analysis should be conducted on a simulation-by-simulation basis. Time series should not be aggregated over simulations, for example by taking a time mean by simulation, prior to creating the BMS. Each simulation should have its own time series and corresponding return variables.

Once the return variables are generated, it may be appropriate to aggregate by simulation. For example, a user interested in finding the return period of 105°F temperatures at a location could generate the return period by simulation and take the median over all the simulations. A subset of simulations could also be selected, such as simulations that have statistically significant p-values for the extreme value distribution fit.

There are many examples of multi-model assessments of return periods and return values in the peer-reviewed literature including^{12,15,16}.

References

1. Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Academic Press.



2. Cooley, D. (2013). Return Periods and Return Levels Under Climate Change. in AghaKouchak, A., D. Easterling, K. Hsu, S. Schubert, & S. Sorooshian (Eds.) *Extremes in a Changing Climate* (pp. 97–114) Springer Netherlands. https://doi.org/10.1007/978-94-007-4479-0_4.
3. Douglas, E. M., Vogel, R. M., & Kroll, C. N. (2001). Impact of Streamflow Persistence on Hydrologic Design. in *Bridging the Gap* (pp. 1–10) American Society of Civil Engineers. [https://doi.org/10.1061/40569\(2001\)75](https://doi.org/10.1061/40569(2001)75).
4. Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., & Koutsoyiannis, D. (2015). One hundred years of return period: Strengths and limitations. *Water Resources Research*, 51(10), 8570–8585 <https://doi.org/10.1002/2015WR017820>.
5. Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer London. <https://doi.org/10.1007/978-1-4471-3675-0>.
6. Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., & Wehner, M. (2021). Changes in Annual Extremes of Daily Temperature and Precipitation in CMIP6 Models. *Journal of Climate*, 34(9), 3441–3460 <https://doi.org/10.1175/JCLI-D-19-1013.1>.
7. Slater, R., Freychet, N., & Hegerl, G. (2021). Substantial changes in the probability of future annual temperature extremes. *Atmospheric Science Letters*, 22(11), e1061 <https://doi.org/10.1002/asl.1061>.
8. Zwiers, F. W., & Von Storch, H. (1995). Taking Serial Correlation into Account in Tests of the Mean. *Journal of Climate*, 8(2), 336–351 [https://doi.org/10.1175/1520-0442\(1995\)008%3C0336:TSCIAI%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008%3C0336:TSCIAI%3E2.0.CO;2).
9. England, J., Cohn, T., Faber, B., Stedinger, J., Thomas, W., Veilleux, A., Kiang, J., & Mason, R. (2018). *Guidelines for determining flood flow frequency*. USGS.
10. Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, 25(8-12), 1287–1304 [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8).
11. Furrer, E., Katz, R., Walter, M., & Furrer, R. (2010). Statistical modeling of hot spells and heat waves. *Climate Research*, 43(3), 191–205 <https://doi.org/10.3354/cr00924>.
12. Kharin, V. V., Zwiers, F. W., Zhang, X., & Hegerl, G. C. (2007). Changes in Temperature and Precipitation Extremes in the IPCC Ensemble of Global Coupled Model Simulations. *Journal of Climate*, 20(8), 1419–1444 <https://doi.org/10.1175/JCLI4066.1>.
13. Efron, B., & Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1(1) <https://doi.org/10.1214/ss/1177013815>.
14. Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, 82(397), 171–185 <https://doi.org/10.1080/01621459.1987.10478410>.
15. Kharin, V. V., & Zwiers, F. W. (2005). Estimating Extremes in Transient Climate Change Simulations. *Journal of Climate*, 18(8), 1156–1173 <https://doi.org/10.1175/JCLI3320.1>.
16. Wehner, M., Gleckler, P., & Lee, J. (2020). Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 1, model evaluation. *Weather and Climate Extremes*, 30, 100283 <https://doi.org/10.1016/j.wace.2020.100283>.

