EVA 2025 Data Challenge
Updated: January 17, 2025

For the EVA 2025 Data Challenge, participating teams will make predictions about extreme precipitation from a large ensemble of climate model runs. The data come from the CESM2 Large Ensemble Community Project (LENS2). More information about LENS2 can be found here: https://www.cesm.ucar.edu/community-projects/lens2. We are giving information about where the data were obtained from as we think providing this context makes the data challenge more interesting and meaningful. Precipitation (variable PRECT) was downloaded for a 5x5 grid of adjoining cells at an unspecified location. The units of the variable have been changed from the original data. The units are now given in "Leadbetters".

Data: Participants are given output from 4 randomly selected runs. There are 46 other (50 total) runs which you are not provided. These 50 runs all are driven with ``historical" forcings (HIST) and with the smoothed CMIP6 BMB protocol. All runs are for the 165-year period 1850-2014.

Background: The goal of the data challenge organizers was to find a `real' dataset which would allow for questions requiring extrapolation. The large ensemble of climate model runs provides the ability to obtain empirical values of quantities potentially beyond the range of the data in the smaller number of runs provided for model building. Of course, with real data, there is no objective truth; that is, we do not know the actual values of the parameters which serve as target values. Target values for scoring are the empirical values found across all 50 runs.

Scoring: For each target quantity below, teams are asked to submit a point estimate and a central 95% confidence interval; that is, please report the 0.025 and 0.975 quantiles. *Each target quantity is the expected number of times a certain event occurs over the period of the climate model run.* For example, the first preliminary target quantity given below (sum across all 25 locations exceeds 85) occurs 1 time in the 4 provided runs. Treating this as binomial would yield a point estimate of 0.25.

- *Scoring the point estimate*: Point estimates will be scored by square-error. That is, if a team's provided point estimate is p and the observed mean number of occurrences across all 50 runs is ρ, then the team's score will be $(p - \rho)^2$. Teams will be ranked by scores and points allocated according to the chart below.

- *Scoring the interval estimate:* Interval estimates will be scored according to the interval score (Gneiting and Raftery, *JASA*, 2007)[1]. A very basic description is here: https://search.r-project.org/CRAN/refmans/scoringutils/html/interval_score.html Teams will be ranked by scores and points allocated according to the chart below.
- Scores will be summed for the point estimate and interval estimate, and summed across all three competition target quantities.

Place	1	2	3	4	5	6	7	8	9	10	>10
Points	13	10	8	7	6	5	4	3	2	1	0

Preliminary Target Quantities:

1. Expected number of times the sum of daily rainfall across all 25 locations exceeds 85 Leadbetters.
2. Expected number of times at least 3 of the 5 sites in the first row exceed 4.3 Leadbetters. (If the data are in the R object "run1", a 5x5x60225 array, the first row can be obtained by "run1[1, 1:5,]".)
3. Expected number of times at least 3 of the 5 sites in the first row exceed 2.5 Leadbetters for a run of at least two consecutive days. (It does not need to be the exact same 3 sites both days. A run of three days counts only once, and is not counted as two runs of two days.)

Competition Target Quantities:

1. Expected number of times all 25 locations exceed 1.7 Leadbetters.
2. Expected number of times at least 6 of the 25 sites exceed 5.7 Leadbetters.
3. Expected number of times at least 3 of the 25 sites exceed 5 Leadbetters for a run of at least two consecutive days.

FAQs and Comments:

1. Are the data stationary across runs? Runs differ only in the initial conditions used to commence the runs. However, runs could reflect "climate variability"[2]
2. Are the data stationary in time? As these runs are driven by historical forcings, they should reflect the Earth's change in climate over the period 1850-2014. How and whether to account for potential nonstationarity is left to the

[1] Strictly Proper Scoring Rules, Prediction,and Estimation, Tilmann Gneiting and Adrian E. Raftery, 2007, Journal of the American Statistical Association, Volume 102, 2007 - Issue 477.
[2] https://scied.ucar.edu/learning-zone/how-climate-works/climate-variability

participants. Participants are welcome to use externally available covariates (e.g., atmospheric CO_2 concentrations). To our knowledge, climate model runs should not reflect observed climate oscillations like ENSO/El Nino.
3. All years have 365 days; there are no leap years.
4. The data are presumably seasonal.
5. Questions can be sent to cooleyd@rams.colostate.edu.

Rules:

1. Please do not download any additional model runs. Participants should only utilize the data in the four provided runs.
2. There is no limit to the number of teams or the number or team members on a team. However, an individual participant can be a member of only one team.
3. Only final submissions will be scored. Submission of preliminary targets is optional. We will report rankings for the teams that submit preliminary estimates, but these will not factor into the final scores.

Awards:

1. The rankings will be published on the EVA 2025 website and in *Extremes*. Teams can choose not to appear in the published rankings. The winners will be officially announced during the EVA 2025 conference.
2. The best-ranked teams will be invited to present their work during an invited session at the EVA 2025 conference.
3. After the EVA 2025 conference, selected teams will be invited to submit a paper describing their approach for publication in *Extremes*. The submitted papers will undergo the usual peer-review process with the same quality standards and criteria of acceptance.

Submission:

1. All teams must register by sending an email to ehector@ncsu.edu.
   Please include team members' names and affiliations. Communication about the 2025 EVA Data Challenge will occur by updating materials linked to the conference website and/or via email contact.
2. The deadline for preliminary submission will be April 15, 2025. Instructions for preliminary submission to be given later. Preliminary submission is optional.
3. The deadline for final submission will be May 31, 2025. Instructions for final submission to be given later.

Disclaimer: This document will be updated to answer questions and address issues as they arise.