**Canadian Large Ensembles Adjusted Dataset version 1 (CanLEADv1)**

The dataset contains large ensembles of bias adjusted daily climate model outputs of minimum temperature, maximum temperature, precipitation, relative humidity, surface pressure, wind speed, incoming shortwave radiation, and incoming longwave radiation on a 0.5-degree grid over North America. Intended uses include hydrological/land surface impact modelling and related event attribution studies.

The CanLEADv1 dataset is based on archived climate model simulations in the Canadian Regional Climate Model Large Ensemble (CanRCM4 LE) <https://open.canada.ca/data/en/dataset/83aa1b18-6616-405e-9bce-af7ef8c2031c> and Canadian Earth System Model Large Ensembles (CanESM2 LE) <https://open.canada.ca/data/en/dataset/aa7b6823-fd1e-49ff-a6fb-68076a4a477c> datasets.

Specifically, CanLEADv1 provides bias adjusted daily climate variables over North America derived from 50 member initial condition ensembles of CanESM2 (ALL and NAT radiative forcings) and CanESM2-driven CanRCM4 (ALL radiative forcings) simulations (Scinocca et al., 2016; Fyfe et al., 2017). Raw CanESM2 LE and CanRCM4 LE outputs are bias adjusted (Cannon, 2018; Cannon et al., 2015) so that they are statistically consistent with two observationally-constrained historical meteorological forcing datasets (S14FD, Iizumi et al., 2017; EWEMBI, Lange, 2018).

File names, formats, and metadata headers follow the recommended Data Reference Syntax for bias-adjusted Coordinated Regional Downscaling Experiment (CORDEX) simulations (Nikulin and Legutke, 2016).

Multiple initial condition simulations can be used to investigate the externally forced response, internal variability, and the relative role of external forcing and internal variability on the climate system (e.g., Fyfe et al., 2017). Large ensembles of ALL and NAT simulations can be compared in event attribution studies (e.g., Kirchmeier-Young et al., 2017). Availability of bias adjusted outputs from the CanESM2-CanRCM4 modelling system can be used to investigate the added value of dynamical downscaling (Scinocca et al., 2016). Multiple observational datasets are used for bias adjustment to partly account for observational uncertainty (Iizumi et al., 2017).

For CanESM2 LE, there are two sets of radiative forcing scenarios (ALL, which consists of historical and RCP8.5 forcings for the periods 1950-2005 and 2006-2100, respectively, and NAT, which consists of historicalNat forcings for the period 1950-2020), two observationally-constrained target datasets for bias adjustment (S14FD and EWEMBI), and 50 ensemble members, which gives a total of 2 × 2 × 50 = 200 sets of outputs. For CanRCM4 LE, historicalNat simulations were not run; hence, there are 2 × 50 = 100 sets of outputs. In both cases, CanLEADv1 provides variables on the CORDEX NAM-44i 0.5-degree grid. CanESM2 outputs (~2.8-degree grid) and CanRCM4 outputs (0.44-degree grid), are bilinearly interpolated onto the NAM-44i grid before bias adjustment.

A multivariate version of quantile mapping (Cannon, 2018) is used to adjust the distribution of each simulated variable, as well as the statistical dependence between variables, so that these properties match those of the target observational dataset. Bias adjustment is performed on a grid cell by grid cell basis. Outside of the historical calibration period, the climate change signal simulated by the climate model is preserved (Cannon et al., 2015).

**Key words:** bias adjustment/correction; temperature; precipitation; humidity; pressure; wind; radiation; climate model; land surface model; hydrological model; event attribution; dynamical downscaling; internal variability; climate change

**Citation:**

Cannon, A.J., H. Alford, R.R. Shrestha, M.C. Kirchmeier-Young, and M.R. Najafi, 2021. Canadian Large Ensembles Adjusted Dataset version 1 (CanLEADv1): Multivariate bias-corrected climate model outputs for terrestrial modelling and attribution studies in North America. Geoscience Data Journal. doi:10.1002/gdj3.142

**References:**

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Climate Dynamics, 50(1-2), 31-49.

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? Journal of Climate, 28(17), 6938-6959.

Fyfe, J. C., Derksen, C., Mudryk, L., Flato, G. M., Santer, B. D., Swart, N. C., Molotch, N. P., Zhang, X., Wan, H., Arora, V. K., Scinocca, J., & Jiao, Y. (2017). Large near-term projected snowpack loss over the western United States. Nature Communications, 8, 14996.

Iizumi, T., Takikawa, H., Hirabayashi, Y., Hanasaki, N., & Nishimori, M. (2017). Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. Journal of Geophysical Research: Atmospheres, 122(15), 7800-7819.

Kirchmeier-Young, M. C., Zwiers, F. W., Gillett, N. P., & Cannon, A. J. (2017). Attributing extreme fire risk in Western Canada to human emissions. Climatic Change, 144(2), 365-379.

Lange, S. (2018). Bias correction of surface downwelling longwave and shortwave radiation for the EWEMBI dataset. Earth System Dynamics, 9(2), 627-645.

Nikulin, G., & Legutke, S. (2016). Data Reference Syntax (DRS) for bias-adjusted CORDEX simulations. <http://is-enes-data.github.io/CORDEX_adjust_drs.pdf>

Scinocca, J. F., Kharin, V. V., Jiao, Y., Qian, M. W., Lazare, M., Solheim, L., Flato, G. M., Biner, S., Desgagne, & Dugas, B. (2016). Coordinated global and regional climate modeling. Journal of Climate, 29(1), 17-35.