Design Value Explorer: Methodology and Background Information
Version: February 15, 2022
Pacific Climate Impacts Consortium

General information about the climatic design values presented in the Design Value Explorer (DVE) and the methods used to process station data can be found in the *National Building Code of Canada, Appendix C: Climatic and Seismic Information for Building Design in Canada* (hereafter NBCC, 2015). The present document reviews the updates to the data presented there, including some methodological changes, and also describes the mapping method developed at PCIC that is an integral feature of DVE.

1. Station data processing and updates

PCIC obtained Canada-wide observational data from the Meteorological Service of Canada (MSC), comprising more than 150 variables, some measured since the early 20th century, and created a dedicated database for use in this project. Additional snow observations from provincial snow monitoring networks collected by the MSC in Quebec and British Columbia were also obtained, with permission. These data include snow depth, snow density and snow water equivalent observations that are essential to meeting the building code needs for both snow and rain-on-snow loading.

Processing procedures varied by the climatic design element, as described in NBCC (2015). Departures from these procedures are described below. For reference, **Table 1** provides a summary of all data used to calculate each design value (DV), including the variable name, sampling frequency, period of record, and number of stations used after screening and quality control. For most variables, the quality checks applied are those summarized in Tables 1, 2 and 4 of Durre et al. (2010).

1.1 Composite stations

Station records used for design variable calculations need to satisfy minimum length constraints. For example, at least 20 years of annual maxima of snow depth were required for the 50-year snow load DV, while a minimum of 8 years of complete hourly data (i.e., no missing hours) in January or July was required for the extreme lower and upper temperature percentiles. This disqualifies the majority of stations, which have shorter and/or incomplete records, and therefore constitutes an inefficient use of data. For several variables, composite stations were created by combining quality-controlled data from nearby short-record stations. The definition of 'nearby' varied by the DV: for example, snow depth annual maxima from stations with a horizontal separation of < 50 km and an elevation difference of < 100 m were combined into composites. For precipitation, a much smaller separation of 5 km was used, due to the finer scale structure of that field. While the chosen separations are somewhat arbitrary, the corresponding scales reflect the smoother nature of an accumulated variable— snow depth—versus a non-accumulated variable like precipitation. The addition of composites increased the number of analyzed stations modestly, by 5 to 9%, depending on the DV.

1.2 Precipitation

1.2.1 Climatological precipitation and rainfall

The previous analysis underlying NBCC (2015) used daily precipitation and rainfall data from stations over the interval 1961-1990. In addition to incorporating data from more recent decades, we introduced a method to make maximal use of station records with data in the longer interval (1961-2016) but that may have significant gaps. Referred to as the day-over-years (DoY) approach, a given station was required to have a minimum number of values occurring on each of January 1st, January 2nd, and so on, but contiguous values within any given calendar year were not required (and leap days were ignored). This procedure makes maximal use of the available data (e.g., 1386 stations using DoY compared to 965 stations for a 25-year climatology), under the assumption that climatological characterization is the principal interest.

Table 1. List of MSC variables analyzed, and properties of the subsetted data after completeness screening and quality control. The period of record and station totals include composite stations.

MSC variable	Sampling frequency	Period of record	No. stations (screened)
Snow depth	daily	1945-2017	520
Air temperature	daily	1982-2018	1322
Air temperature	hourly	1982-2018	531 (Jan) – 537 (Jul)
Dewpoint temperature	hourly	1961-2016	202
Wet bulb temperature	hourly	1982-2018	564
Relative humidity	hourly	1980-2018	554
Precipitation	daily (for climatology)	1961-2016	1390
Rainfall	daily (for climatology)	1961-2016	1386
Rainfall	daily (for extremes)	1947-2017	2116
Rainfall	5-min,10-min,15-min,30-min, 1-hr,2-hr,6-hr,12-hr,24-hr	1960-2017	547-640, depending on duration

1.2.2 Daily and sub-daily rainfall

An extensive analysis of one-day rainfall amounts at stations with 20 years or more of measurements was conducted. Since annual extrema of daily rainfall are required for the calculation of the 50-year return level of one-day rain, R1d50, 100% completeness of data within each year is desired. For this reason, the DoY approach used for climatologies is not appropriate. However, station records with gaps were still considered, by applying the method of Papalexiou & Koutsoyiannis (2013) which recognizes that high-ranking annual extremes are often preserved in incomplete records. Use of this procedure permitted an increase in the sample size by 15%. Annual maxima were fit using a Gumbel distribution and method of L-moments to estimate R1d50. The decision to use the Gumbel distribution fitted by the method of L-moments was based on extensive testing of several combinations of extreme value distributions and fitting methods. Daily maxima of interval rainfall from the MSC

DLY03 dataset (MSC, 2020) were analyzed for intervals ranging from 24 hours down to 5-min. The 10-year return level of 15-minute rainfall, R15m10, was computed in the same manner as for R1d50.

1.3 Relative humidity

Annual mean relative humidity (RH), an element of the Canadian Highway Bridge Design Code, was constructed from hourly RH measurements at stations with data from 1980 onward, to avoid a pervasive inhomogeneity issue prior to this time identified by Vincent et al. (2007). To maximize the number of stations (and so spatial coverage), climatologies were constructed from gap-filled 10-year records. Relatively short 10-year records were used to maximize station coverage after determining that the 10-year climatologies were statistically indistinguishable from 20-year climatologies at locations where both could be computed.

1.4 Snow and rain-on snow load

A systematic re-evaluation of ground snow loads over Canada was conducted using a methodology that improved upon NBCC (2015) in several respects. Two independent sources of historical ground snow data were employed: daily snow depth (SD) measurements from the MSC for 4412 stations, and the Manual Snow Survey data set, comprising snow water equivalent (SWE) data collected from 3320 locations, none exactly coincident with a MSC station. These extensive data along with a regional snow climate classification scheme (Sturm et al., 1995) were used to develop power-law relationships between annual maximum SD and SWE over Canada, allowing the derivation of a regional 'pseudo-density' relating the two variables and an annual maximum SWE time series at each MSC station (Wilks and McKay, 1996). After extensive quality control, an extreme value analysis was then applied to these time series to estimate the 50-year return level of annual maximum SWE and the corresponding snow load (SL50) at over 500 MSC stations across Canada. A subsequent analysis of temporally consistent rainfall data resulted in the estimation of the 50-year return level of rain-on-snow loads at the same locations.

1.5 Wind loads

Unlike other DVs, estimates of station-based wind speed extremes were obtained from the Engineering Climate Services Unit and those for station-based wind pressures conditional on the occurrence of rainfall above a specified threshold (1.8 mm h^{-1}) were obtained from the Climate Research Division of ECCC. PCIC used station-based estimates of the n-year return level of height- and exposure-adjusted wind speed from the former to derive the corresponding return level of wind pressure, WPn, at each station (where n = 10 or 50). Station-based estimates of 5-year driving rain wind pressure, DRWP5, provided by the Climate Research Division were used directly. Following the usual NBCC practice, "floor values" were applied to WP10 (0.23 kPa), WP50 (0.3 kPa) and DRWP5 (40 Pa) before the spatial interpolation.

2. Using regional climate model simulations to inform design value estimates

The relatively sparse distribution of meteorological stations over most of Canada, coupled with the fact that many of these stations cover periods that are too short to allow reliable estimates of DVs, creates challenges for

traditional interpolation methods. We therefore developed a mapping method that blends station based DV estimates with output from a regional climate model, here the Canadian Regional Climate Model, Version 4 (CanRCM4), to improve design value interpolation in data-sparse areas. A full description of the model, and its merits and limitations with respect to simulating DVs over Canada, is given in the comprehensive report of Cannon et al. (2020). CanRCM4's performance in simulating several key climate variables underlying the DVs over North America has been evaluated for temperature and precipitation extremes (Whan and Zwiers, 2016; Li et al., 2019), mean precipitation (Diaconescu et al., 2016), and surface wind speed (both annual values and extremes) and driving rain wind pressure (Jeong, Cannon and Morris, 2020).

Three key goals of the desired mapping method are: 1) it should be objective, in the sense that reconstructed DVs are produced without explicit ad-hoc or expert adjustments; 2) the final map should have a higher spatial resolution than the native RCM (45 km, true at 60 °N), in order to provide values at the community scale while allowing variation between communities in the same region (we set this target resolution to be ~4.5 km, or ten times finer than the native model grid), and; 3) it should also provide maps of future projected changes to DVs indexed to different levels of global warming (see ff., Section 4).

2.1 Model pre-processing

Two separate ensembles of initial condition simulations from CanRCM4 were used, a 35-member set with output provided at daily time resolution was used to calculate most DVs, while a 15-member set with hourly archived results was used for DVs that were based on hourly extremes. While any individual realization could be considered analogous to the observed climate system, we utilized the ensemble mean of the 35 (or 15) members to create a single smoother representation that is less affected by natural, unforced, climate variability.

CanRCM4 employs a horizontal grid of 130 by 155 points on a polar stereographic projection covering the entirety of North America except for the portion north of ~75 °N. The grid is irregularly spaced in latitude and longitude, but regularly spaced at $0.44° \times 0.44°$ in a rotated pole projection. After calculating the desired DV on the native CanRCM4 grid, the model land mask is applied. Each land grid cell in the rotated coordinate system was then divided into a 10 x 10 array, and the new grid cell centres used for bilinear interpolation. The target grid produced in this way has a nominal resolution of 4.5 km, true at 60 °N. However, the coarse representation of coastlines in CanRCM4 is problematic; specifically, the locations of many coastal meteorological stations and city/town centres listed in Table C-2 of the NBCC fall outside the land mask of the final, high-resolution target grid. To correct this, a second fine-scale land mask was created from a high-resolution (50-m) vector of the Canadian coastline, obtained from Natural Earth, https://www.naturalearthdata.com/. Any grid cell on the target grid that intersects the high-resolution coastline was assigned to the target grid land mask. To assign values of the DV field at the target grid cells land grid cells lying outside the CanRCM4 land mask (except in the high Arctic, which is treated separately; see below), a nearest-neighbour approach is used. The closest land value in the DV field at the target resolution, masked by the *original CanRCM4 land mask,* is assigned to these grid cells. CanRCM4 cells that lie outside the high-resolution coastline and the borders of Canada are excluded from the target map. The result of this procedure is illustrated in **Figure 1**, while the entire model preprocessing procedure is outlined in **Figure 2**.
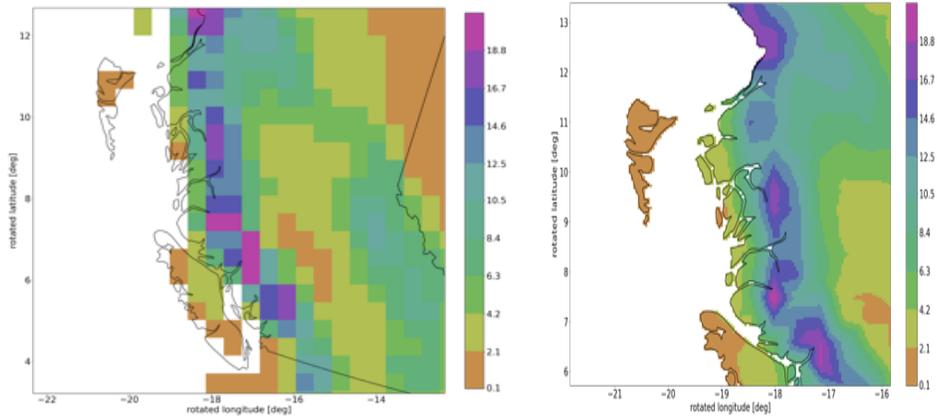
Figure 1. Result of the model preprocessing steps outlined in Figure 3. *Left:* Land-masked CanRCM4 DV field on the native model grid. *Right:* Close-up of the same DV field on the target grid, showing the refined land mask constructed from the 50-m resolution Canadian coastline from Natural Earth. The region shown is coastal British Columbia, with Haida Gwaii at top left, Vancouver Island at lower centre. The DV shown is the 50-year return level of annual maximum snow load (SL50, in kPa).
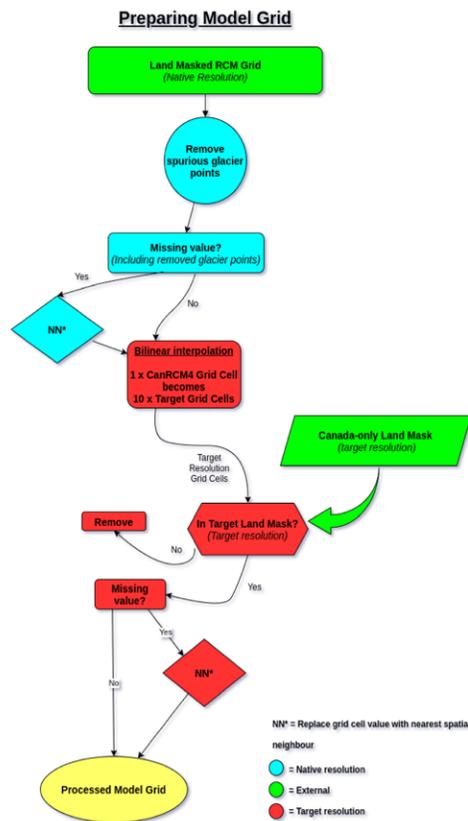
Figure 2. Flowchart illustrating the steps involved in preprocessing the CanRCM4 model output for use in the hybrid spatial mapping method.

2.2 Spatial mapping method

2.2.1 Incorporation of station values

The main goal is to use the spatially complete, two-dimensional DV fields from CanRCM4 to guide interpolation between observationally-derived DVs at stations. Hence the principal utility of the model is in the spatial pattern, and only secondarily in the magnitude, of the DV field it simulates. The key steps of the method, which we refer to as hybrid spatial mapping (HSM), are outlined in **Figure 3**. RCM values M_i are generally biased relative to observations, and thus station values S_i are often used to bias-adjust the models in some fashion. After model pre-processing, the mean model bias over a region of interest (comprising $i = 1,…, N_S$ stations) is removed by the simple rescaling

$$M'_j = (\langle S_i \rangle / \langle M_i \rangle)M_j, \qquad j = 1, …, N_m \qquad (1)$$

where $\langle \cdot \rangle$ denotes the arithmetic average over the station locations i and the index j ranges over all model grid cells N_m (at the target resolution) in the region of interest.

Next, we define the bias ratio $B_i = S_i/M'_i$ and aim to find a method that, first, interpolates this quantity to all j locations on the target grid, B_j, and second, brings B_j as close to unity as possible at each point in the region (since $B_j = 1$, $j = i$, indicates no bias). Given such a method, the final step is to estimate the local DV field through multiplication of the interpolated bias B_j' by the rescaled model field, M_j', the adjusted model DV field at the target resolution. That is, we "reconstruct" the DV field as:

$$R_j = M_j' B_j'. \qquad (2)$$

If, for example, the chosen interpolation method happens to be exact at the station grid cells, then $B_i' = B_i$ and $R_i = M_j' (S_i/ M_j') = S_i$; i.e., the reconstruction matches the station value exactly at these grid locations. At grid cells away from stations, i.e. over the majority of the domain, R_j incorporates both station and model information in a manner that automatically applies a weighting between the two according to station proximity.

For the interpolation we use ordinary kriging (ORKG), which is designed to construct the best linear unbiased estimator of predicted field values (Isaacs and Srivastava, 1989). The spatial structure of the DV field surrounding each station is approximated by an isotropic, exponential covariance function with three parameters (nugget, sill and range). We allow for a positive, non-zero nugget, indicating a discontinuous variogram at zero separation. Hence, while the interpolation does not exactly match the station values, it does account for unknown measurement errors, resulting in a smoother field. ORKG is applied at the regional scale to the discrete bias field B_i, with parameters estimated by maximum likelihood estimation (MLE). The ORKG-MLE method converges to parameter estimates that minimize the error variance of the set of j estimated values B_j' across the entire domain of interest. This makes the method preferable to purely mathematical interpolation or curve-fitting techniques that do not explicitly account for the local spatial covariance structure that is a feature of most physical fields.

In order to apply HSM over the entirety of Canada, we implemented ORKG-MLE in a moving polygon mode, with each polygon defined by the k nearest stations to a target station (we chose $k = 30$; see, e.g., Haas [1990]). This allows for regionally varying covariance parameters since, although the form of the covariance function is kept fixed in each polygon, the anisotropic station distribution is reflected in the varying size and location of the polygons. Significant overlap was allowed to avoid edge discontinuities. Once every station has acted as a target,

the mean value of B_j' over all overlapping polygons is computed at each target grid cell j. In station-rich areas, more windows are averaged than in station-poor areas, reducing measurement error in the former regions.
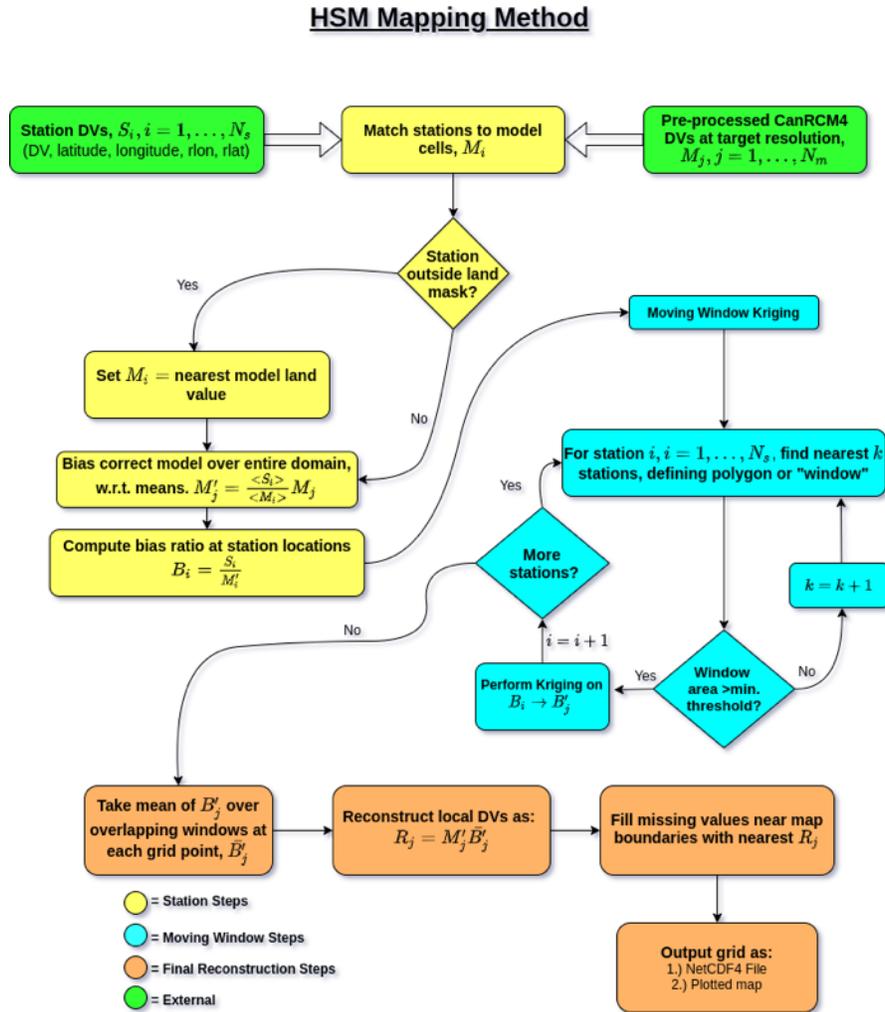
HSM Mapping Method

Figure 3. Flowchart illustrating various steps in the hybrid spatial mapping (HSM) method. The steps leading to the pre-processed model inputs at upper right are shown in Figure 2.

2.2.2 Regions outside the model domain: Border areas and the Upper Arctic Archipelago

After moving window averaging, the resulting bias field $\overline{B}_j$ covers most of Canada. However, since the moving window method is station-based, certain coastal portions and station-poor areas near the borders of the map, e.g. in Northern Canada, are not covered by a window. We therefore filled each of these missing grid cells with a $\overline{B}_j$ value copied from the nearest grid cell (Fig. 2). The reconstruction then proceeded according to eq. (2), with the grid-wise multiplication of $\overline{B}_j$ by the model DV field, M_j, to produce the reconstructed field R_j.

The domain of CanRCM4 extends no further north than 75 $^{\circ}$N, presenting the challenge of reconstructing values in a region (hereafter referred to as the Upper Arctic Archipelago, or UAA) where the bias ratio, B_i, cannot be computed. We take the UAA to comprise the islands north of the continuous ocean passage stretching from M'Clure Strait in the west to Barrow Strait in the east. Station density is also very sparse in this area (typically 0-4 stations, depending on the DV), meaning that direct interpolation of station DVs is not a viable option. We use two sources of information for the UAA reconstructed values. First, we take the arithmetic mean of all available stations DVs in the UAA, $\bar{S}_{UAA}$. If no UAA stations are available, we assign the DV of the northernmost station on the map to $\bar{S}_{UAA}$. Second, we make use of reconstructed values $R_{LAA,j}$ from the nearby islands to the south of this channel (where LAA denotes the Lower Arctic Archipelago). Specifically, we take the arithmetic mean of all land-only target grid values in a thin strip between 72 $^{\circ}$N and 73 $^{\circ}$N, denoting this quantity as $\bar{R}_{LAA}$. The final UAA value is then defined as the mean of the averaged UAA station and LAA reconstruction means:

$$\bar{R}_{UAA} = \tfrac{1}{2} \left(\bar{R}_{LAA} + \bar{S}_{UAA} \right). \qquad (3)$$

Note that over the UAA, $R_j = \bar{R}_{UAA}$ is taken as a single, spatially constant DV. In our view, the paucity of stations and complete absence of model values does not justify a spatially varying reconstruction in this region. With the R_j values in the UAA now fixed, this completes the Canada-wide reconstruction of the DV field.

3. Example and comparison with NBCC (2015)

The final result of the HSM procedure for SL50 is shown in **Figure 4**, which also shows the pre-processed CanRCM4 and station DVs used as input to the HSM method. We draw attention here to some key characteristics of the reconstructions.

The complete set of reconstructions for all DVs, along with the corresponding input station and CanRCM4 maps, reveals that the overall goals of the method have been met. Specifically, the HSM method is objective and optimal, in the sense that it evidently produces a reconstruction whose resemblance to station DVs is high where station density is high, and whose resemblance to the spatial pattern of model DVs is high where stations are sparse or absent. The upper left panel of Fig. 4 illustrates the effect of HSM on the empirical distribution of the input data for the SL50 DV. The CanRCM4 distribution is characterized by values that are quite centrally-concentrated with a low median, few large values, and positive skewness, while the station observations exhibit a broader spatial distribution with a higher median, but again with positive skewness and a long upper tail. The distribution of the HSM reconstruction resembles that of the stations much more than that of the model, with a median DV close to the station median. In particular, the reconstruction adds more values between 5 and 10 kPa than are present in the model, bringing it more into line with the station distribution. However, the reconstruction also has a lower maximum DV compared to both the model and stations, at ~15 kPa compared to > 20 kPa. The lower maximum is an expected consequence of using a spatial interpolation technique that accounts for a local "nugget" effect that is designed to account for uncertainty in local design value estimates. This differs from an interpolating technique that requires smooth interpolating surfaces to pass through all stations (i.e., a zero nugget).

We caution that despite the evident success of HSM in bringing the reconstructed DVs into line with station values, comparison with the NBCC (2015) Table C-2 values may reveal both random and systematic differences. First, the station-based DVs used as input to the HSM differ from those underlying NBCC (2015), both because the input data have been extended to 2018 (in most cases) and because analysis methods used to derive many

of the DVs have been modified (see Section 1). Second, the use of a new, objective interpolation methodology would be expected to lead to some differences.. Although we feel that that the objective, RCM-based method is preferable, biases may nevertheless result where the station density is low and/or where local topographic influences are unresolved by the model. Finally, while the reconstructions are rendered at fairly high resolution, they may inherit model bias from larger scales, particularly in areas where station data are sparse.
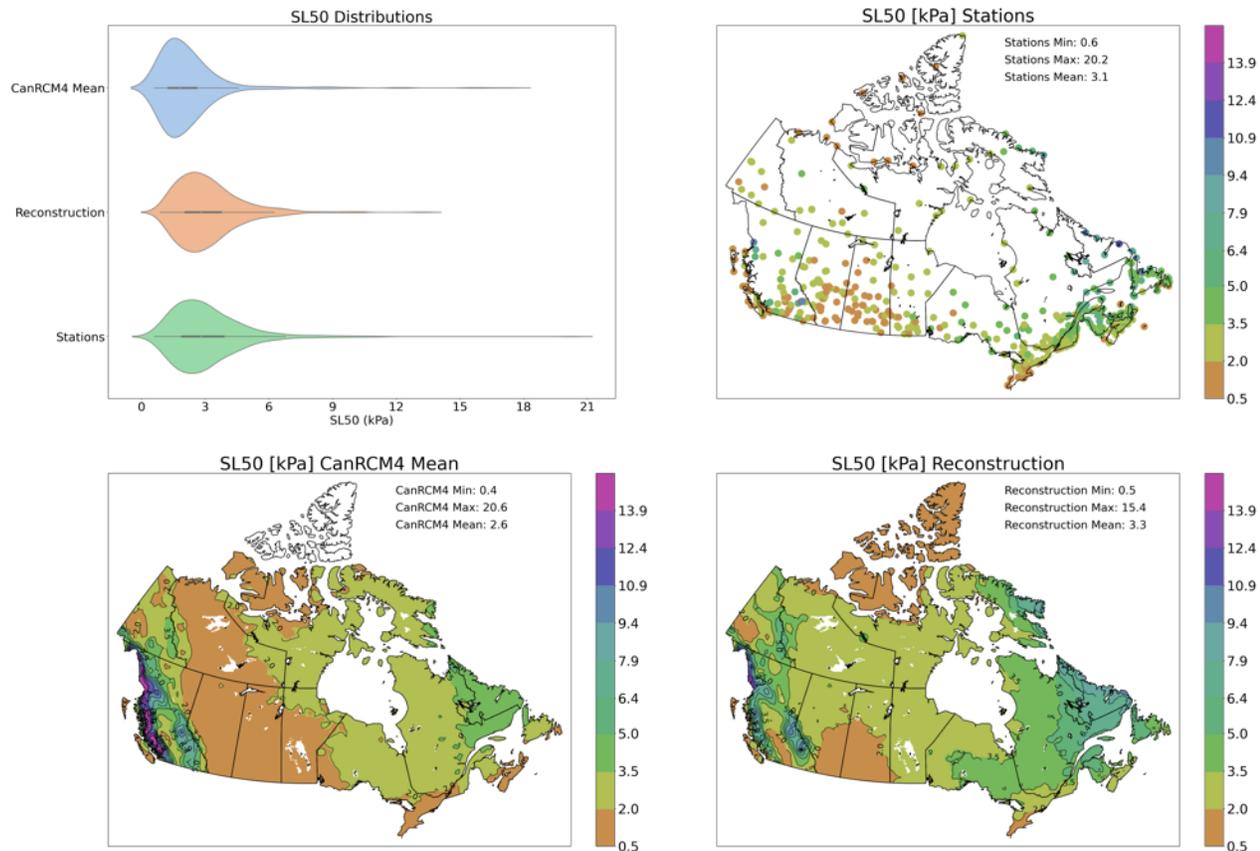
Figure 4. *Upper left:* Empirical spatial distribution function ("violin plot") of the SL50 DV. Blue shows the distribution of model design values amongst grid boxes containing stations, green shows the distribution for design values at stations, and salmon shows the distribution of reconstructed design values amongst grid boxes containing stations. The horizontal axis indicates design value magnitudes, and the width of the "violin" describes the relative frequency of occurrence of those design values. *Upper right:* Map of SL50 at stations, *Lower left:* CanRCM4 ensemble mean map of SL50. *Lower:* Final reconstruction, R_j, of the SL50 DV.

4. Future-projected design value estimates

4.1 Presentation of climate projections as a function of global temperature change

As described in the comprehensive report of Cannon et al. (2020), ECCC and PCIC used CanRCM4 to derive future DV fields under a high-emissions greenhouse gas scenario (RCP8.5). Consistent with the approach taken in international climate policy, which is focused on specific global warming levels, like 2.0°C above pre-industrial

levels, the DVE presents model projections of DV changes at specified levels of annual mean global surface air temperature change (ΔT) rather than during fixed time periods. Framing climate projections in this way shifts the impact of scenario uncertainty from the magnitude of change to the timing of ΔT increases: the relationship between the two can be seen in Table 2. The timing of warming was assessed at fixed levels of global mean ΔT = 0.5°C, …, 3.5°C, in 0.5°C increments, from a historical baseline of 1986-2016. Specifically, we determined the year by which global warming permanently surpasses these levels averaged across an ensemble of climate change simulations consisting of one run from each of the Coupled Model Intercomparison Project Phase 5 (CMIP5) models. As shown in Table 2, the closer the future time horizon, the less sensitivity there is to different emissions pathways (RCPs+.

Table 2. Timing of projected global warming. The year at which the indicated global mean warming ΔT relative to 1986-2016 is irrevocably exceeded by the CMIP5 single-member per model ensemble mean, under different Representative Concentration Scenarios (RCPs). Model results were obtained from KNMI Climate Explorer https://climexp.knmi.nl/start.cgi. A dash ("−") indicates that the corresponding ΔT is not reached before 2100 for that RCP. Results from different RCPs are averaged if the years of exceedence are within 6 years of each other.

Global ΔT	RCP8.5	RCP6.0	RCP4.5	RCP2.6
0.5°C			2023	
1.0°C	2035		2046	−
1.5°C	2047		2070	−
2.0°C	2059	2087	−	−
2.5°C	2069	−	−	−
3.0°C	2080	−	−	−
3.5°C	2090	−	−	−

4.2 Future climate projections as change factors

The DVE provides future-projected changes relative to baseline values either as increments (for temperature-related DVs) or multiplicative factors (for all other DVs). So, for example, if at a given location the change factor (CF) for JulT2.5 at a global ΔT of 3.0°C is provided as 2.2, then 2.2°C should be added to the 1986-2016 reference value for JulT2.5. Alternatively, if the CF for R1d50 is provided as 0.94 for a given level of global ΔT, then the reference value for R1d50 at that location should be multiplied by 0.94. Change factors for future IDF curves (the design variable IDFCF in the corresponding DVE menu) are to be interpreted similarly. As described more fully in the Cannon et al. (2020) report, the ability of CanRCM4 to accurately simulate short duration rainfall at the local scale is very limited. Therefore, as recommended in the report, we assume that the rainfall intensity at all durations increases according to the multiplicative CF

$$CF = (1 + \alpha)^{\Delta T_{loc}} \qquad (4)$$

where α = 0.07 corresponds to the Clausius-Clapeyron rate of ~7% per °C for the increase of the water vapour content of air with increasing temperature and ΔT_{loc} is the local air temperature change as simulated by the CanRCM4 model ensemble mean (or more precisely, the closest grid cell value in the model to a geographic location of interest) that corresponds to one of the selected changes in global mean temperature, ΔT. Future

periods for calculating ΔT_{loc} are defined according to the global mean temperature change, ΔT in Table 2. Thus, since ΔT_{loc} is a function of global ΔT and geographic location, so is the CF. We therefore constructed maps, first at the model native resolution of 45 km (true at 60°N) and then at 10 times higher resolution, of CFs for different levels of global warming.

Appendix: Software implementation

The software implementation of the method in Python and R, *climpyrical*, was developed over the course of the project and makes use of a number of publicly available routines. The ORKG-MLE method was implemented using the R Statistical Computing package (R Core Team, 2020), specifically the routine spatialProcess included in the library fields (CRAN, 2019; Wiens and Krock, 2019). Since the rest of the analysis was implemented in Python (Python Software Foundation, 2020), the spatialProcess routine was called from the Python package rpy2 (Gauthier, 2020). Other Python routines utilized include: NumPy (Oliphant, 2006) for main matrix operations; Pandas (McKinney et al., 2010) for station data processing; GeoPandas (Jordahl, 2014) for mask generation and coordinate transformations; and Scikit-Learn (Pedregosa, 2011) and SciPy (Virtanen, 2020) for nearest neighbour and interpolation modules.

References

Cannon et al., 2020. *Climate-Resilient Buildings and Core Public Infrastructure: Guidance on Updating Climate Design Values for Enhanced Resilience to Future Climate Conditions*, available on-line from ECCC at https://climate-scenarios.canada.ca/?page=buildings-report or http://publications.gc.ca/site/eng/9.893021/publication.html.

CRAN, 2019. https://cran.r-project.org/web/packages/fields/index.html. Last accessed on December 8, 2020.

Diaconescu, E. P., et al., 2016. Evaluation of precipitation indices over North America from various configurations of regional climate models. *Atmosphere-Ocean*, 54(4), 418-439.

Gauthier, L. 2020. rpy2, R wrapper for Python, version 3.3.6. Available at https://pypi.org/project/rpy2/

Haas, T. C., 1990. Kriging and automated variogram modeling within a moving window. *Atmospheric Environment. Part A. General Topics*, *24*(7), 1759-1769.

Isaacs, E.H. and R. M. Srivastava, 1989. *Applied Geostatistics* (New York: Oxford University Press), Ch. 12.

Jeong, D. I., A. J. Cannon, & R. J. Morris, 2020. Projected changes to wind loads coinciding with rainfall for building design in Canada based on an ensemble of Canadian regional climate model simulations. *Climatic Change*, 162, 821–835.

Jordahl, K., 2014. GeoPandas: Python tools for geographic data. https://Github. Com/Geopandas/Geopandas.

Li, C., F.W. Zwiers, X. Zhang, and G. Li, 2019. How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, 7. https://doi. org/10.1029/2018EF001001

McKinney, W., and coauthors, 2010. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (Vol. 445, pp. 51–56).

MSC, 2020. Technical Documentation Digital Archive of Canadian Climatological Data. https://climate.weather.gc.ca/doc/Technical_Documentation.pdf. Last accessed on December 2, 2020.

NBCC, 2015. *National Building Code of Canada, Appendix C: Climatic and Seismic Information for Building Design in Canada*. https://nrc.canada.ca/en/certifications-evaluations-standards/codes-canada/codes-canada-publications/national-building-code-canada-2015.

Papalexiou, S.M. and Koutsoyiannis, D., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49, 187-201.

Oliphant, T. E., 2006. A guide to NumPy (Vol. 1). Trelgol Publishing USA.

Pedregosa, F., Varoquaux, Ga"el, Gramfort, A., Michel, V., Thirion, B., Grisel, O., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Python, 2020. Python Software Foundation. Python Language Reference, version 3.6. Available at http://www.python.org

R Core Team, 2020. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, available at: https://www.R-project.org/. Last accessed on February 15, 2022.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & van der Walt, S. J., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, *17*(3), 261-272.

Whan, K., Zwiers, F., 2016. Evaluation of extreme rainfall and temperature over North America in CanRCM4 and CRCM5. *Climate Dynamics*, 46**,** 3821–3843. https://doi.org/10.1007/s00382-015-2807-7.

Wiens, A. and M. Krock, 2017. `Fields` vignette. https://github.com/NCAR/fields/blob/master/fieldsVignette.pdf. Last accessed on December 8, 2020.

Wilks, D. S., and McKay, M., 1996. Extreme-value statistics for snowpack water equivalent in the northeastern United States using the cooperative observer network. *Journal of Applied Meteorology*, 35, 706-713.