Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/29/eaba1323/DC1

Supplementary Materials for

Future precipitation increase from very high resolution ensemble downscaling of extreme atmospheric river storms in California

Xingying Huang*, Daniel L. Swain, Alex D. Hall

*Corresponding author. Email: xingyhuang@gmail.com

Published 15 July 2020, *Sci. Adv.* **6**, eaba1323 (2020)
DOI: 10.1126/sciadv.aba1323

This PDF file includes:

Supplemental text

1) Computation and storage cost: The suite of WRF simulations required approximately 1,080,000 core-hours on the Cheyenne supercomputer (hosted by the National Center for Atmospheric Research at the NCAR-Wyoming Supercomputing Center). This computational cost was distributed over 120 discrete simulations in total, averaging 10 days per event, and requiring 1,800 core-hours per event-day. Test cases and troubleshooting required an additional ~60,000 core-hours. The WRF source code used in this study has previously been tested for scalability and optimized for parallel computing. Balancing both the computation cost and speed of the runs, we found the optimal number of processors to be 120-200 for simulating WRF's nested-domain configuration. Since the simulations are event-targeted and drawn from an existing dataset of boundary conditions, multiple events can be simulated in parallel (subject to CPU availability). From the initiation of test cases to completion of the final WRF integration, the overall simulation process spanned over 1.5 years. Due to the large spatial extent of our model domain and the high temporal resolution of model output (hourly), the daily file size can reach 70 GB. We note that the model output variable list has been trimmed to include only those that are essential to our AR analysis and hydroclimate research more broadly.

2) MLR analyses: Multiple individual variables were tested to determine which predictors account for the most fine-scale precipitation variance (R^2). Those variables which were tested but not included in the final model include: zonal wind (U), meridional wind (V), wind speed, wind direction, relative humidity (RH), specific humidity (Q), temperature (T), IWV and IVT. Variables that were tested and retained in the final model include: zonal IVT ($IVT_{U,}$), meridional IVT ($IVT_{V,}$), location parameter (Loc) and local vertical motion (W_{ij}). Due to the multiple vertical levels of U, V, and Q, a series of sensitivity tests were conducted to determine the most relevant vertical level from a range just above the planetary boundary layer (around 400-500 m from the surface) to the level above which atmospheric moisture content is negligible (~300 hPa). We found that the results were ultimately not sensitive to the vertical levels included for these variables. Considering the complexity and interdependence of the atmospheric variables in regulating the large-scale moisture transport and landfalling precipitation, a priori knowledge of potentially important physical processes is initial variable selection in the MLR model--requiring, in this case, a manual (vs. fully automated) statistical procedure. Still, the process of including or excluding predictors in MLR involves testing numerous possible combinations of potentially relevant individual variables. We generally follow the method of stepwise selection to successively add variables that contribute significantly to the explained variance. When two or more dependent variables are included at the same time, only the more/most significant one is used. We aim to build an MLR model with meaningful explanatory power, but also parsimonious and avoiding overfitting, and ultimately found that a 4-predictor, including $IVT_{U,}$ IVT_{V}, Loc, and W_{ij}, multiple linear model maximized the variance in precipitation explained.

Supplemental figures

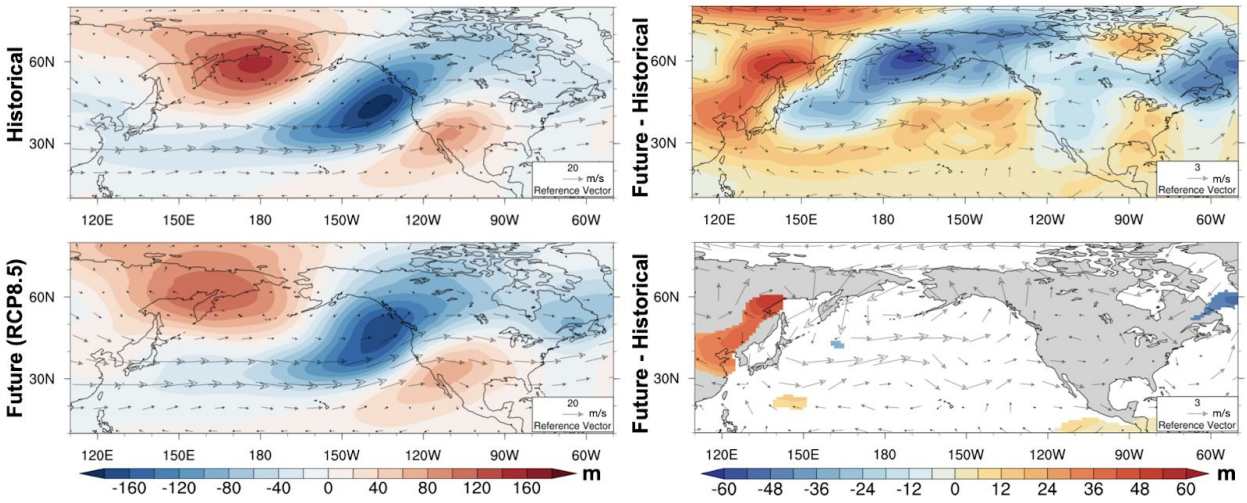

Figure S1: Geopotential height anomaly at 500 hPa (Z500) and the difference between future and historical periods. Left: Z500 anomaly (red/blue contours, m) relative to the monthly mean averaged across 60 AR events when maximum daily precipitation occurs in the historical and future periods. The 500 hPa circulation pattern (UV500) (m/s) during the AR events is overlaid as black vectors. **Right:** Difference between future and historical Z500 anomaly and wind changes (Z500 differences not significant, $p>0.05$, are masked out).

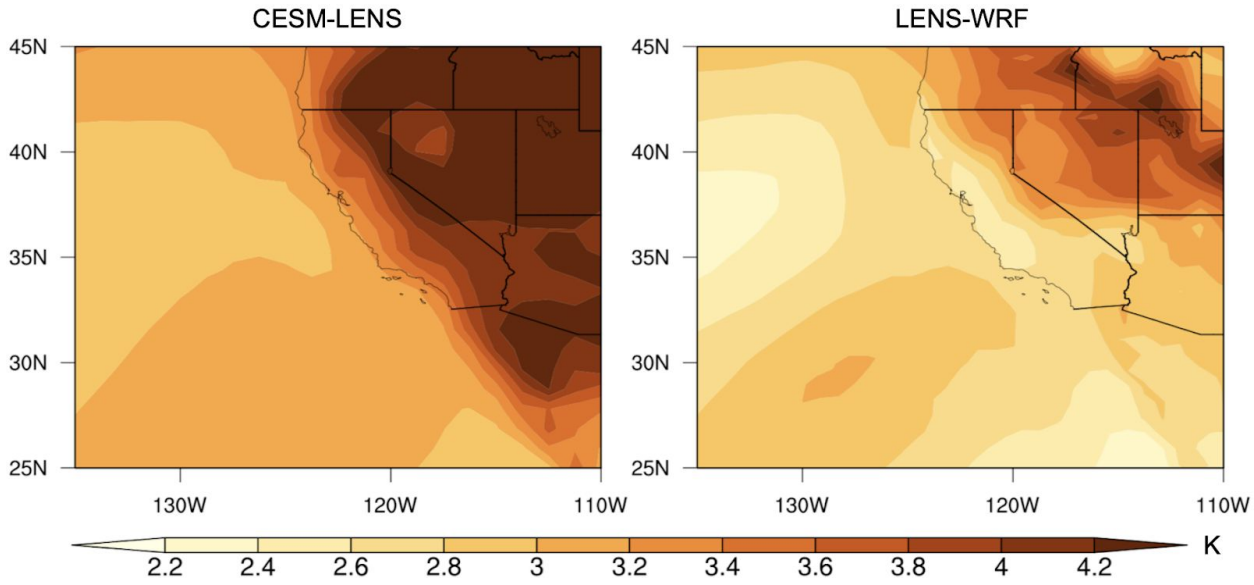

Figure S2: Warming from historical to future period under the RCP8.5 emissions scenario. Left: CESM-LENS climatological warming during Nov-Mar between 2071-2080 and 1996-2005 (RCP8.5 minus Historical simulations in CESM-LENS). **Right**: Similar as left but for warming during the AR events only (for LENS-WRF simulations at coarsest WRF resolution, 81km).

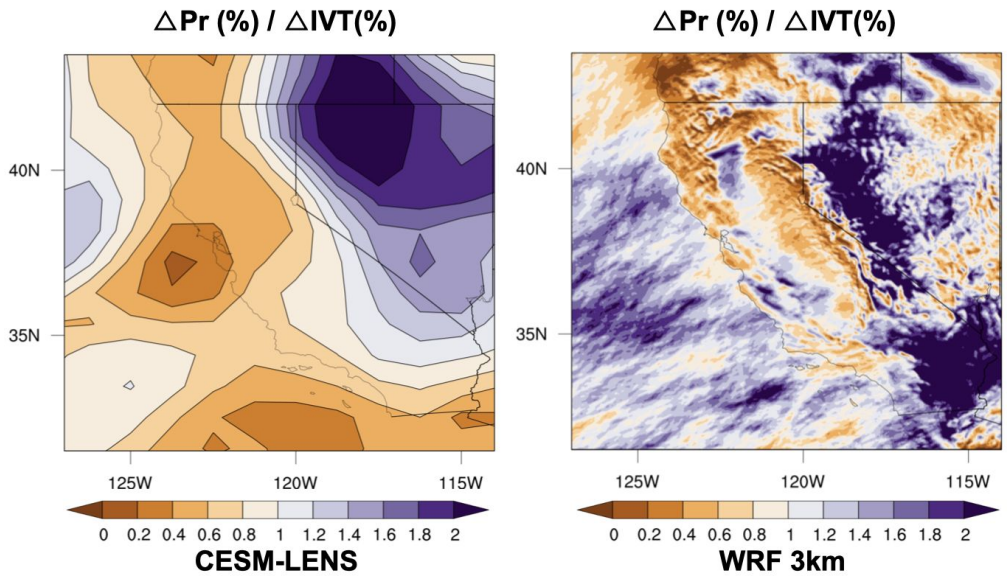

Figure S3: Change in precipitation efficiency during extreme AR events. Left: Ratio of change (%) in precipitation to change (%) in integrated vapor transport (IVT), RCP8.5 minus Historical simulations, in CESM-LENS. **Right**: Same as left, except for WRF 3km simulations.

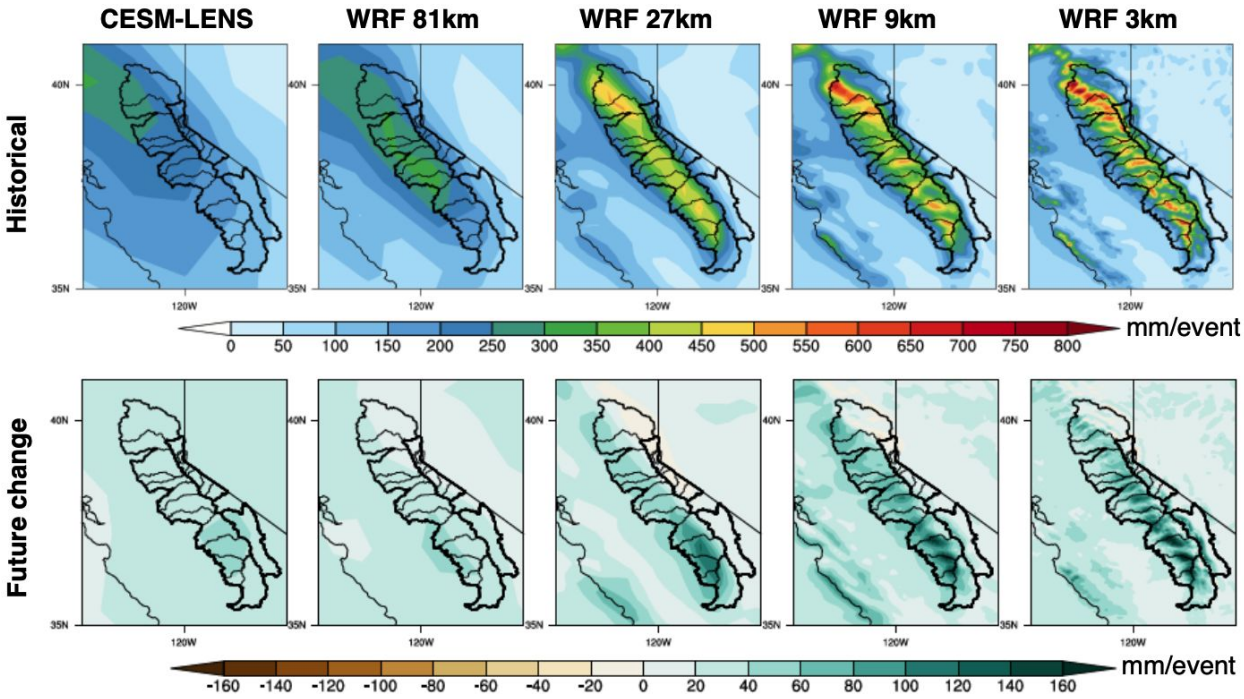

Figure S4: Average event-total precipitation from CESM-LENS and WRF output. First row: Average event-total precipitation from CESM-LENS and WRF output for the historical period; **Second row:** Future changes. All averaged across each 60 AR events and zoomed-in Sierra Nevada region with watershed boundaries overlaid.

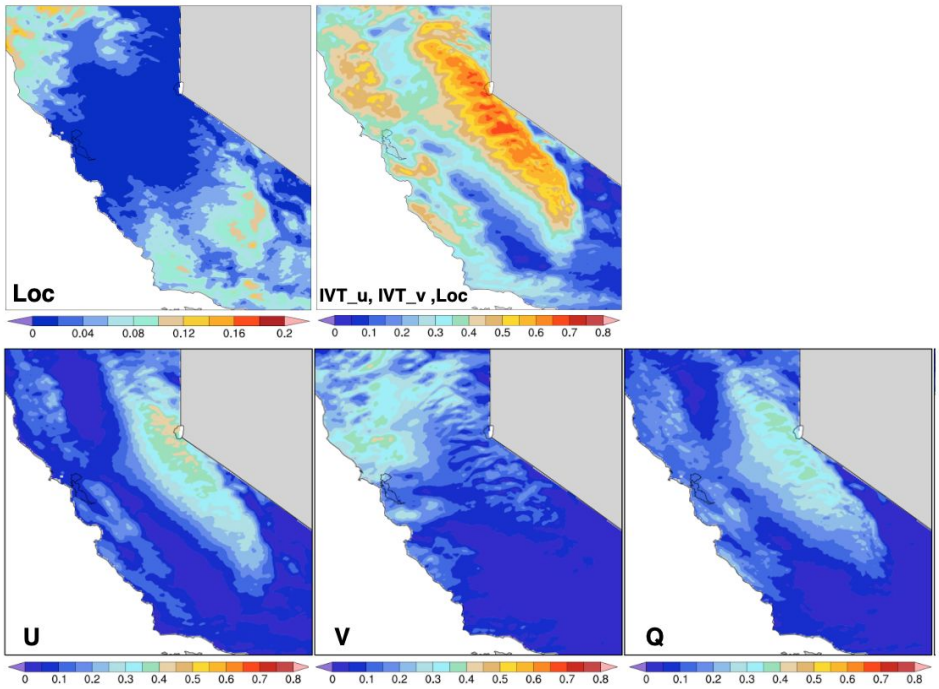

Figure S5: Fraction of explained variance (R^2) in precipitation using single predictor. First row: Variance (R^2) in precipitation explained using Loc single predictor and the MLR model with the three parameters ($IVT_{U,}$ IVT_V, Loc). Here, the Loc factor refers to the grid box location with maximum IVT among the near-coast locations at each regression time-step. **Second row:** similar as first row, but for U, V, and Q single predictors.

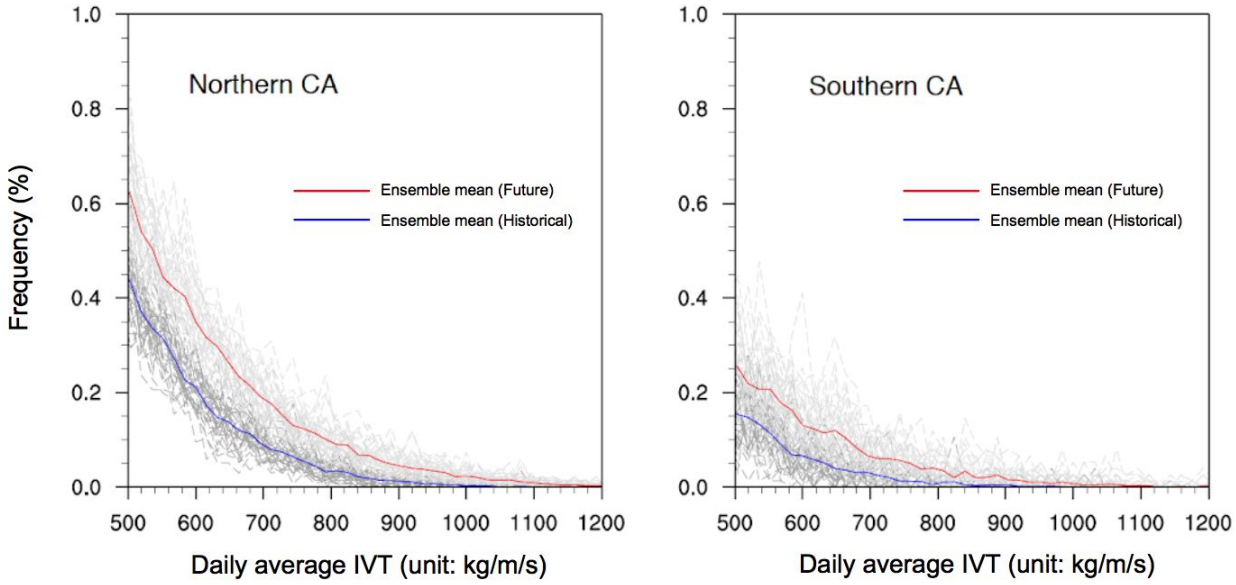

Figure S6: Daily average IVT frequency distributions during the wet season (November 1st to March 31st) from historical (1996-2005) and future (2071-2080) periods in CESM-LENS (40 members). Left: IVT values over northern California (poleward of 35N) coastal grid boxes. **Right**: IVT values over Southern California coastal grid boxes.

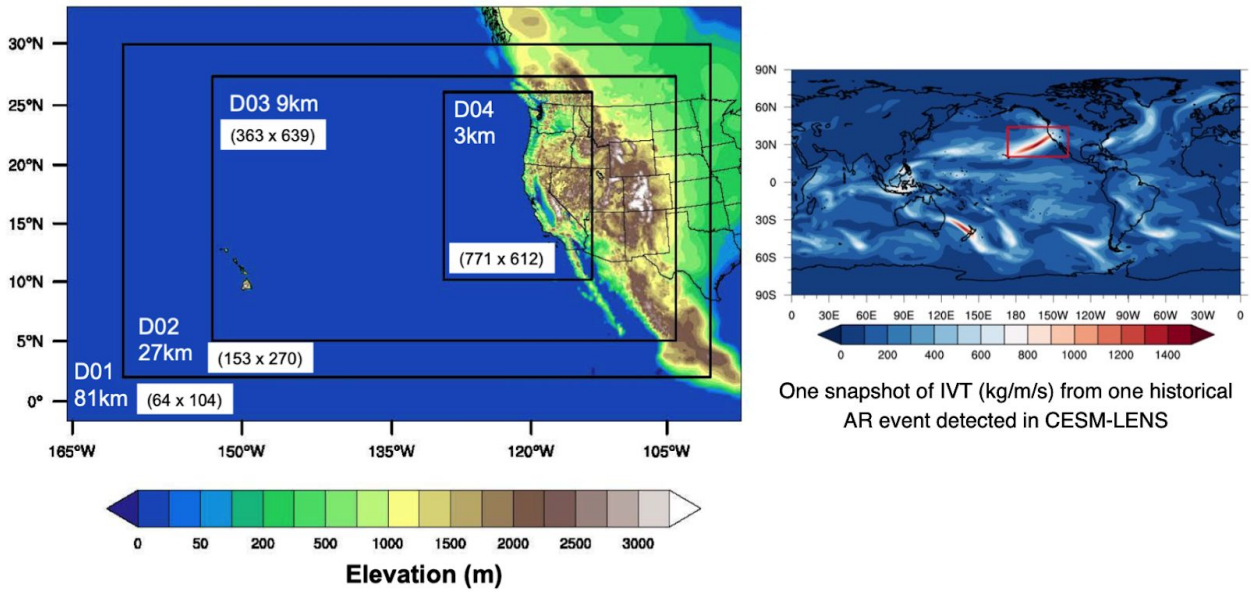

Figure S7: CESM-LENS forced WRF downscaling configuration. Left: Domain settings from D01 (81 km) to D04 (3 km) Domain numbers refer to horizontal resolution, with topography overlaid. **Right**: Example of historical moisture transport (IVT) pattern during an event simulated by CESM-LENS.