Supplementary material

Supplemental text:

Sample size (ablation study): When using different sizes of samples for testing (e.g. 1 year, 2 years, or 5 years), the results still look good overall. In general, 5 years or more can show better accuracy, although it also depends on the dataset availability and the requirement of different purposes.

Supporting dataset: For precipitation, more supporting datasets are used than temperature as described in the main text. It is found that using more relevant supporting datasets instead of elevation data can help improve the prediction to a certain extent, although the detailed contribution from each support variable is not decided. For further applications, the author suggests using more supporting dataset given the availability, although topographic is still the most important factor here.

Residual network: Residual blocks have also been tested during the training processes, and it is found that using this scheme will not improve the prediction accuracy but instead produce noisy values occasionally.

Other loss functions: More complicated loss functions including perceptual loss and style loss have also been tested. It is found that perceptual loss and style loss combined with L1 loss leads to more details and better captured spatial patterns in the daily output. However, the overall accuracy is improved in a pretty minor way compared to the simpler loss functions used here (i.e. L1 and L2 loss). Therefore, to balance both simplicity and efficiency, this feature-based loss function is not discussed in this study.

Supplemental figures

Figure S1: Training loss curve

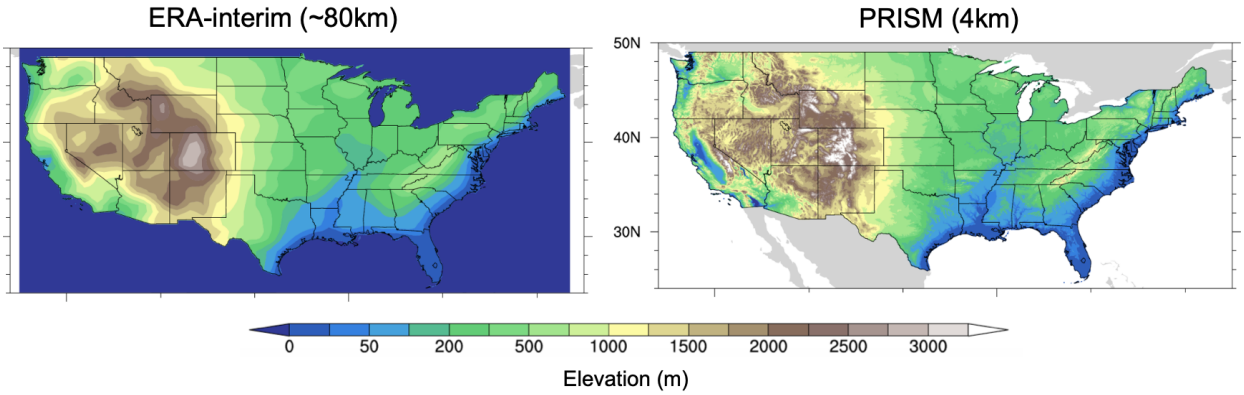

Figure S2: Topographic details from input dataset (i.e. ERA-interim) and target dataset (i.e. PRISM) over the U.S.

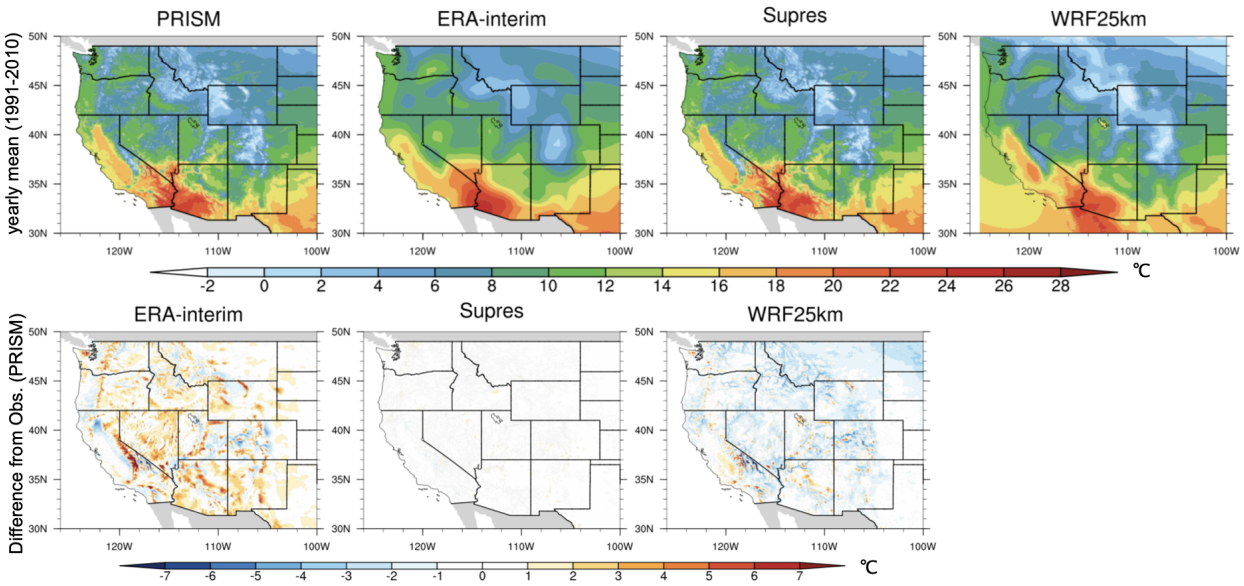

Figure S3: Yearly average temperature (T2) over 1991-2010 from PRISM, ERA-interim, predict (referring as Supres, hereafter), and WRF 25km results (first row); and absolute differences from PRISM (second row).

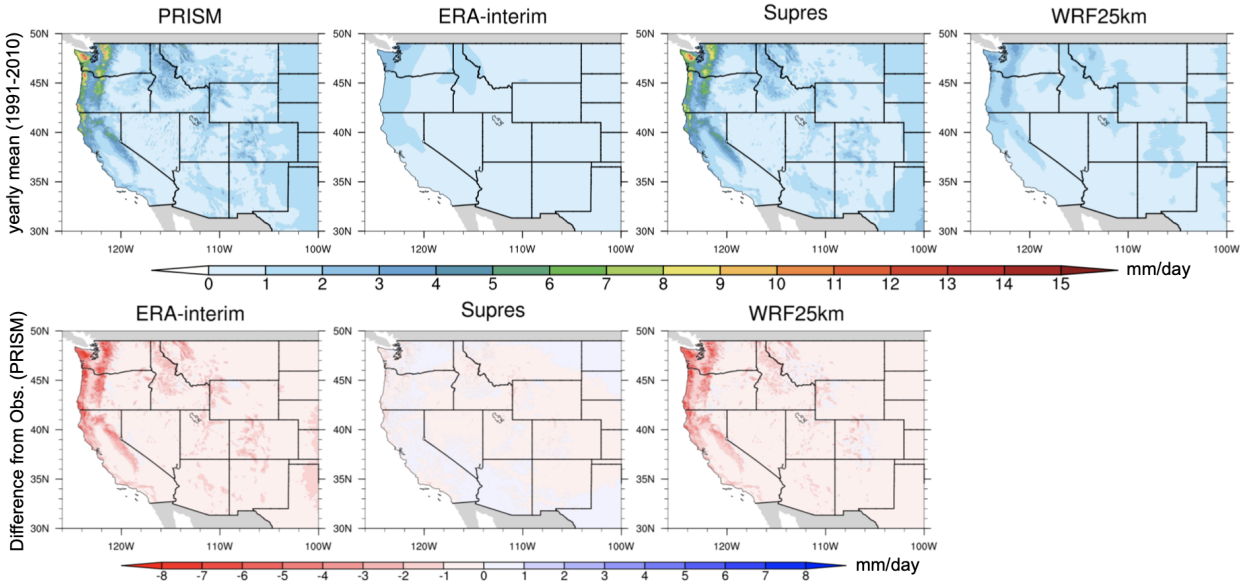

Figure S4: Yearly average precipitation (Pr) over 1991-2010**. Upper row:** Daily precipitation mean from PRISM, ERA-interim, predict, and WRF 25km; **Bottom row:** Differences from PRISM for the input, prediction, and WRF dynamical downscaling results.

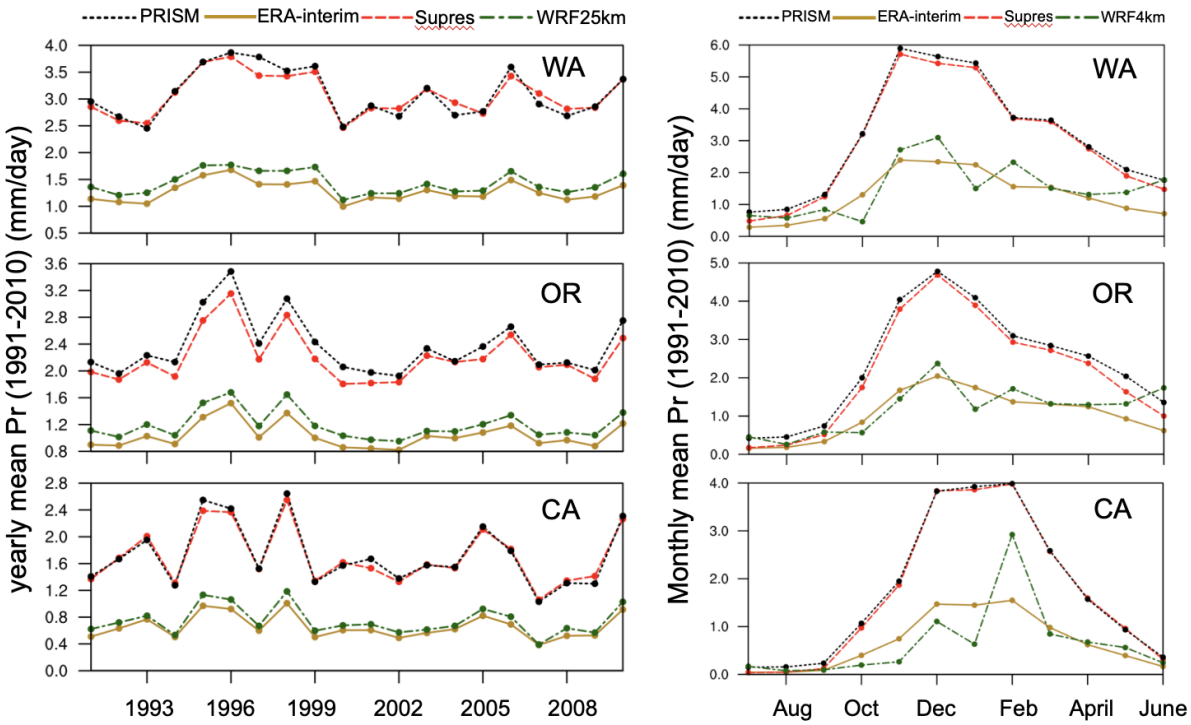

Figure S5: Time series features for Pr from 1991 to 2010 over the western US states including WA, OR, and CA. Left panel: Yearly averaged for PRISM, ERA-interim, Supres, and WRF 25km results; **Right panel:** Similar as left panel, but for the seasonal cycle (i.e. monthly average).