Evaluation of the subseasonal forecast skill of floods associated with atmospheric rivers in coastal Western U.S. watersheds

Qian Cao [a], Shraddhanand Shukla [b], Michael J. DeFlorio [c], F. Martin Ralph [c], and Dennis P. Lettenmaier [a*]

[a] Department of Geography, University of California, Los Angeles, Los Angeles, CA

[b] University of California, Santa Barbara, Santa Barbara, CA

[c] Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego; La Jolla, CA

NOAA Climate Diagnostics and Prediction Workshop (CDPW)
Session 4: Hydroclimate Predictions
October 21, 2020

UCLA
DEPARTMENT OF
Geography

Background and Motivating Questions

Background

- Although the time scale of subseasonal forecasting is critical to proactive disaster mitigation efforts, such as reservoir operations for flood control, it has not received much attention until recently (Vitart et al., 2017).
- The NOAA/Climate Testbed Subseasonal Experiment (SubX) project (Pegion et al., 2019) consists of seven models and focuses on operational subseasonal forecasts with lead time of 32-45 days.
- Atmospheric rivers (ARs) are responsible for most of the storm events leading to extreme precipitation and runoff along the coastal Western U.S. (e.g., Ralph et al., 2006, 2019).
- The forecast skill of meteorological variables (particularly precipitation) is an important determinant of flood prediction skill; however, antecedent soil moisture (ASM) conditions play an important role as well.

Motivating questions:

- ➢ What is the subseasonal forecast skill (at 1-4 week lead times) of AR-related flooding driven by downscaled SubX reforecasts in coastal Western U.S. watersheds? Are SubX-based flood forecasts more skillful than traditional ensemble streamflow prediction (ESP)?
- ➢ What are the relative influences of ASM and SubX reforecast skill on subseasonal flood forecast skill?

Study domain

Three transect basins

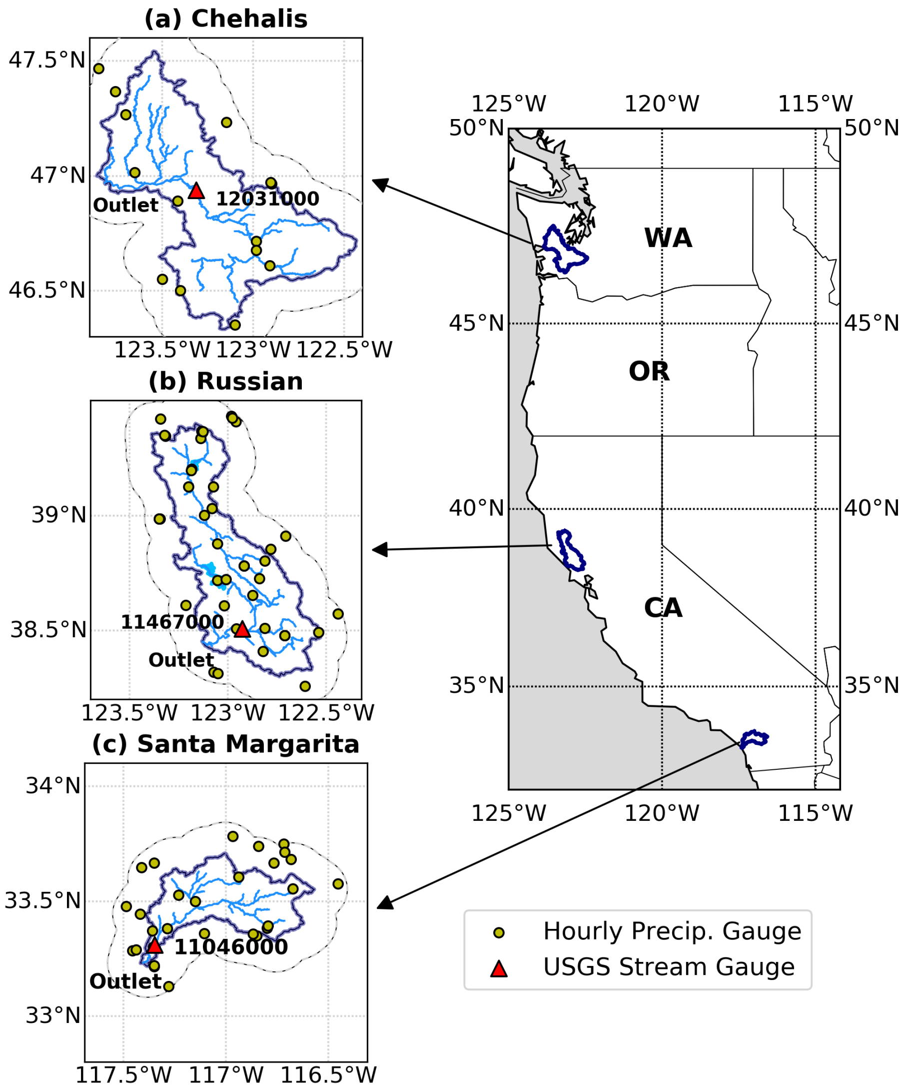


Table 1. Basin characteristics

Basin	Area [km^2]	Elevation range [m]	Annual precipitation [mm]	Precipitation falling between Oct-Mar
Chehalis	5400	0-1429	1560-2700	79%
Russian	3850	0-1324	320-1580	87%
Santa Margarita	1870	143-1736	160-750	83%

SubX models used

Model	Ens Members	Init Interval [days]	Forecast period [days]	Reference(s)
ECCC-GEPS5	4	7	32	Lin et al. (2016)
EMC-GEFS	11	7	35	Zhou et al. (2016, 2017); Zhu et al. (2018)
ESRL-FIMr1p1	4	7	32	Sun et al. (2018a,b)
GMAO-GEOS_V2p1	4	5	45	Koster et al. (2000); Molod et al. (2012); Reichle and Liu (2014); Rienecker et al. (2008)
RSMAS-CCSM4	3	7	45	Infanti and Kirtman (2016)
NCEP-CFSv2	4	1	44	Saha et al. (2014)

The multimodel ensemble (MME) mean is calculated as a lagged average (i.e. by averaging all forecasts from the same start date; in a similar manner to Pegion et al. (2019)).

Data and Methods

1. **Downscaling of the SubX reforecast forcings** ($1^o x 1^o \rightarrow 1/16^o x 1/16^o$)
   - ➢ Bias correction and spatial downscaling (BCSD) (Wood et al., 2004) at a daily time scale
     Variables include precipitation, max. temperature (Tmax), min. temperature (Tmin), wind
     - ▪ We use SubX reforecasts during Oct-Mar months of 1999-2016
     - ▪ Training dataset: $1/16^o$ gridded observations (Livneh et al., 2013) (extended version)

2. Hydrological modeling

- ➢ Distributed Hydrology Soil Vegetation Model (DHSVM) (Wigmosta et al.,1994)
- ➢ Model is run at an hourly time step
  - ▪ Daily precipitation is disaggregated to hourly using a regionalized method of fragments (MoF) algorithm (Westra et al., 2012)
  - ▪ Other meteorological inputs are disaggregated using the Mountain Simulation Model (MTCLIM) algorithms following Bohn et al. (2013)
- ➢ Control run (1999-2016) driven by Livneh et al. (2013) forcings
  - ▪ Output model states (for 7^{th}, 14^{th}, 21^{st}, and 28^{th} of each month) from the control run to provide initial hydrological conditions (IHCs) for flood forecasts

3. Forecast evaluation

- ➢ Precipitation/temperature
- ➢ Flood★

4. ESP and reverse-ESP (revESP)

- ➢ To examine the relative influences of ASM and meteorological forcings

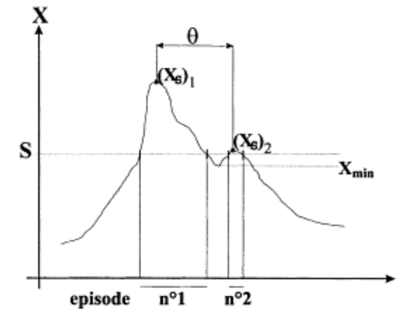
Identification of AR-related flood events

- Flood events → Peaks Over Threshold (**POT**) method
  - *Discharge events* are separated from each other using criteria on the interval between two peaks and a relative threshold on the intermediate flow from the U.S. Water Resources Council (USWRC, 1982).

POT$_{N1}$: one event per year on average
POT$_{N2}$: two events per year on average
POT$_{N3}$: three events per year on average

High
Low

POT threshold



Lang et al., *JH*, 1999

- Flood events associated with ARs
  - We examined AR contributions to extreme events by identifying the flood events that were coincident with AR events.
  - We used the AR date catalog based on the ECMWF Reanalysis-Interim (ERA-Interim) data set, from Guan and Waliser (2015)

Evaluation metrics of forecast skill

1) Precipitation and temperature skill

- Anomaly correlation coefficient (**ACC**; Wilks 2006)

2) Flood forecast skill

➢ Deterministic skill

We evaluated the deterministic skill of NCEP-CFSv2-based flood forecasts because they are initialized every day.

- Kling-Gupta efficiency (**KGE**) (Gupta et al., 2009)

➢ Probabilistic skill

We evaluated the probabilistic flood forecast skill of all six SubX models with 30 ensemble members in total.

- debiased Brier skill score (**BSS**) (Weigel et al., 2007)
- **Hit rate and false alarm rate**

$$\mathrm{BSS} = 1 - \frac{\mathrm{BS}}{\mathrm{BS}_{\mathrm{ref}} + \mathrm{D}} \qquad (1)$$

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right)^2 \qquad (2)$$

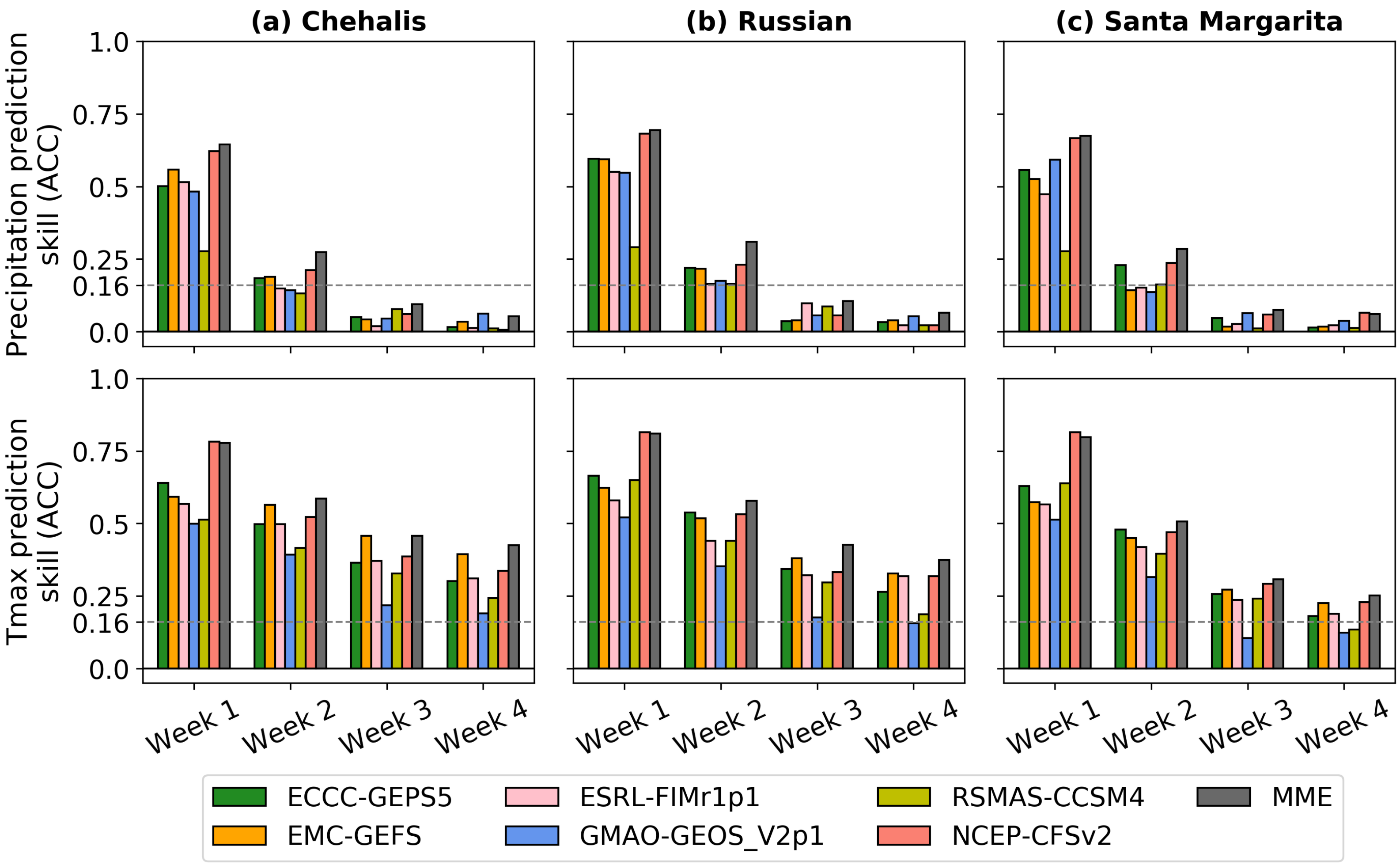
$$\mathrm{BS}_{\mathrm{ref}} = \frac{1}{N}\sum_{i=1}^{N}\left(P_{\mathrm{clim}} - O_i\right)^2 \qquad (3)$$

$$D = \frac{1}{M} P_{\mathrm{clim}}\left(1 - P_{\mathrm{clim}}\right) \qquad (4)$$

		Observed	
		Yes	No
Forecasted	Yes	a (Hit)	b (False alarm)
	No	c (Miss)	d (Correct rejection)

Hit rate = a / (a+c)
False alarm rate = b / (b+d)

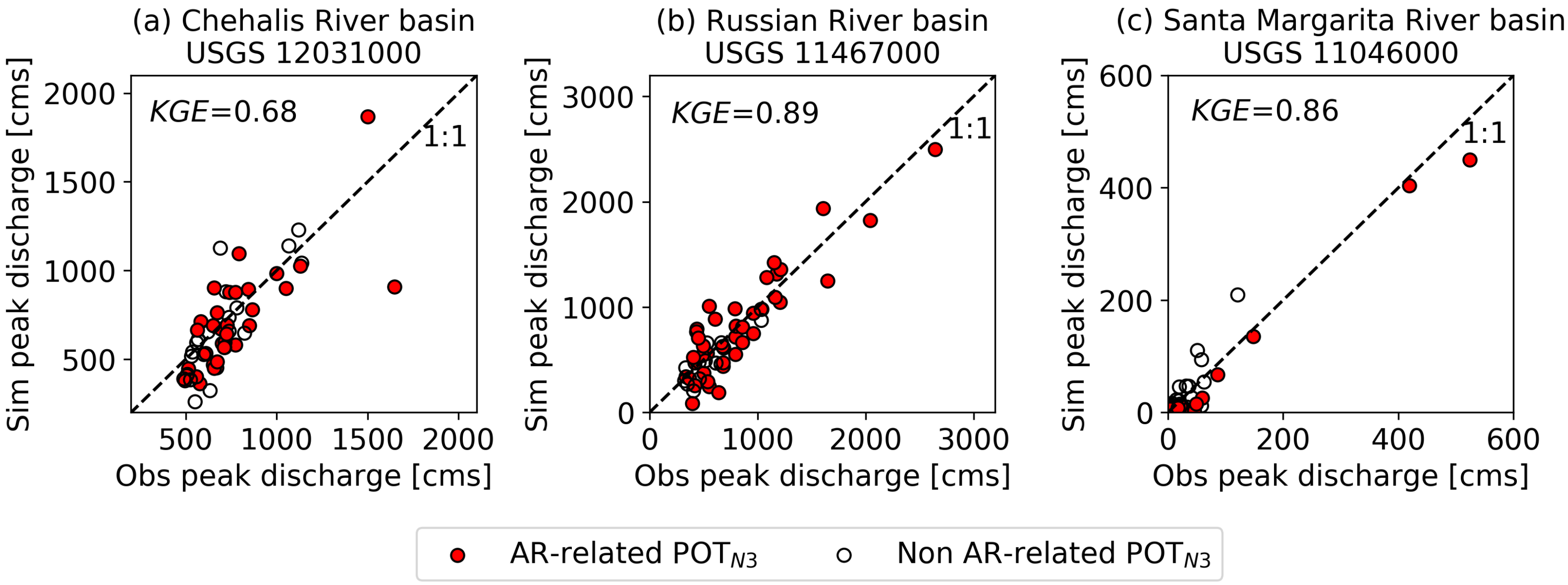
SubX precipitation and temperature skill



Evaluation of simulated peak discharge (control run)

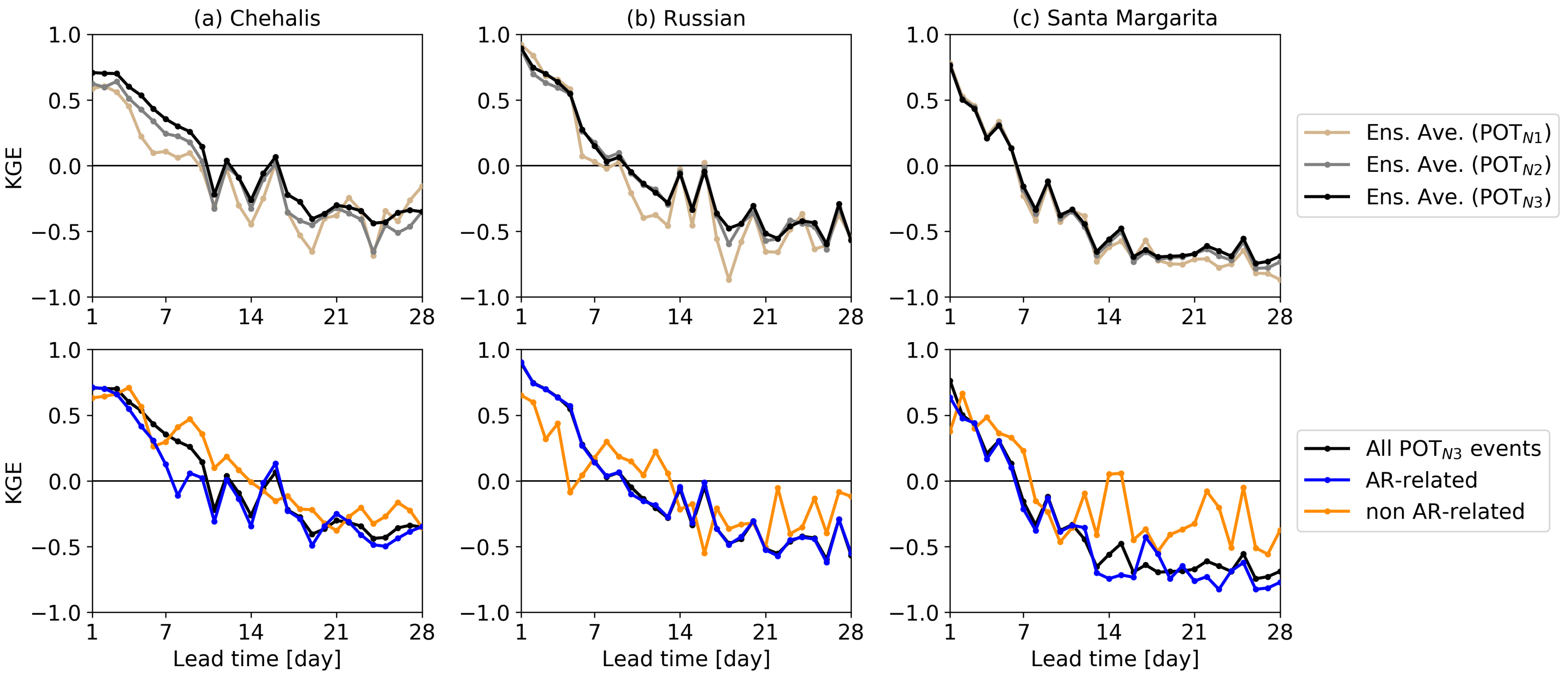
- **AR-related extreme discharge events**

  The percentages of POT_{N3} extreme discharge events that were coincident with ARs during 1999-2016 are 52%, 74% and 41% respectively in the Chehalis, Russian and Santa Margarita River basins.



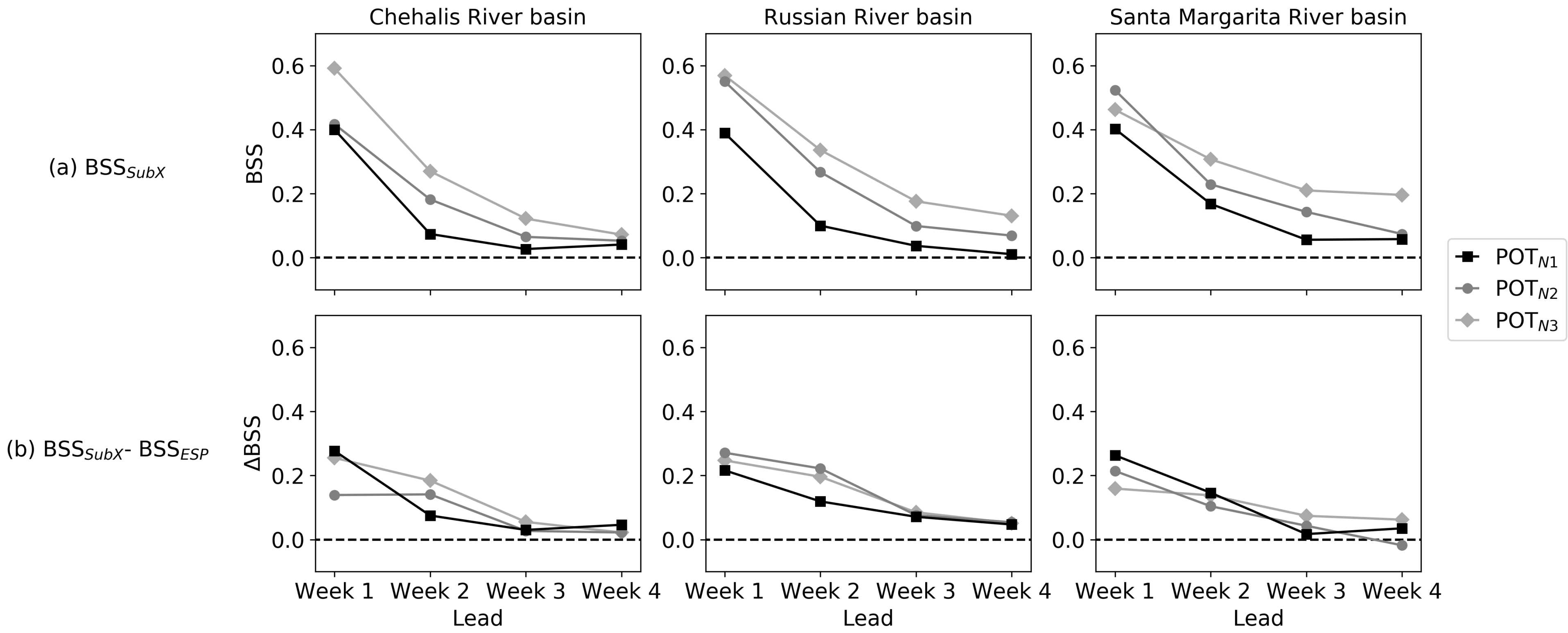
Assessment of flood forecast skill

➢ **Deterministic skill of NCEP-CFSv2-based flood forecasts**



Assessment of flood forecast skill

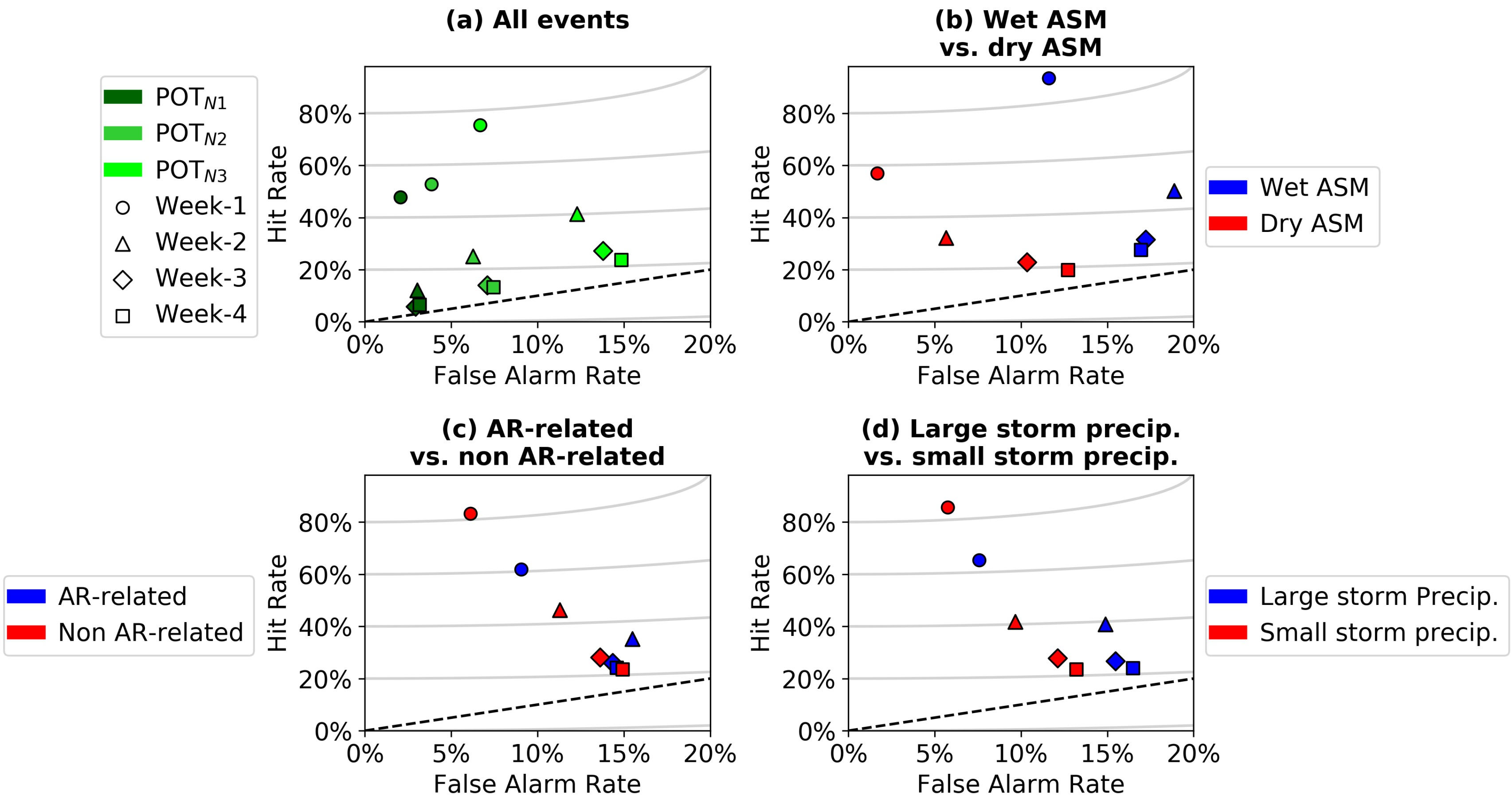
➢ **Probabilistic flood forecast skill: BSS values**



Assessment of flood forecast skill

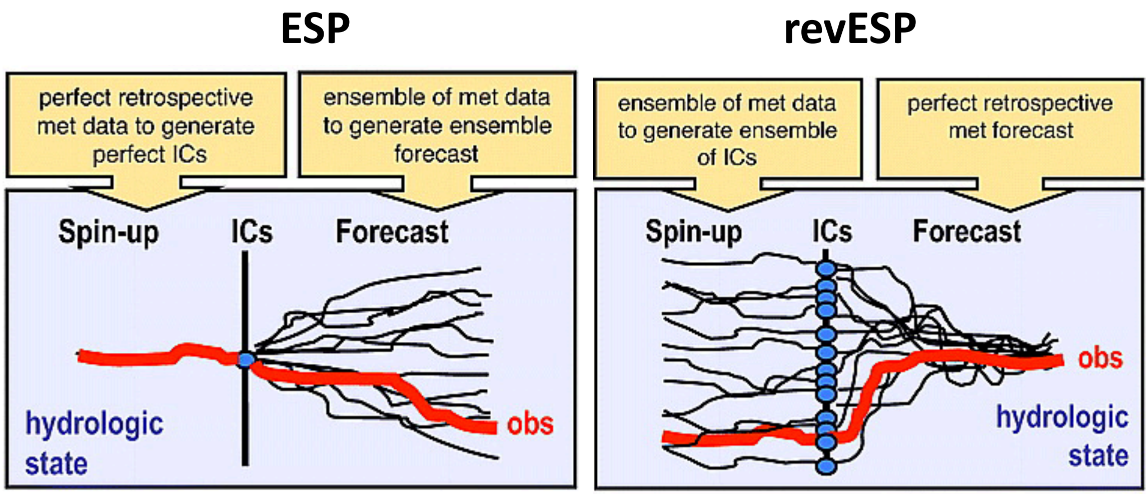
➢ **Probabilistic flood forecast skill: Hit rate vs. False alarm rate**

Chehalis River basin

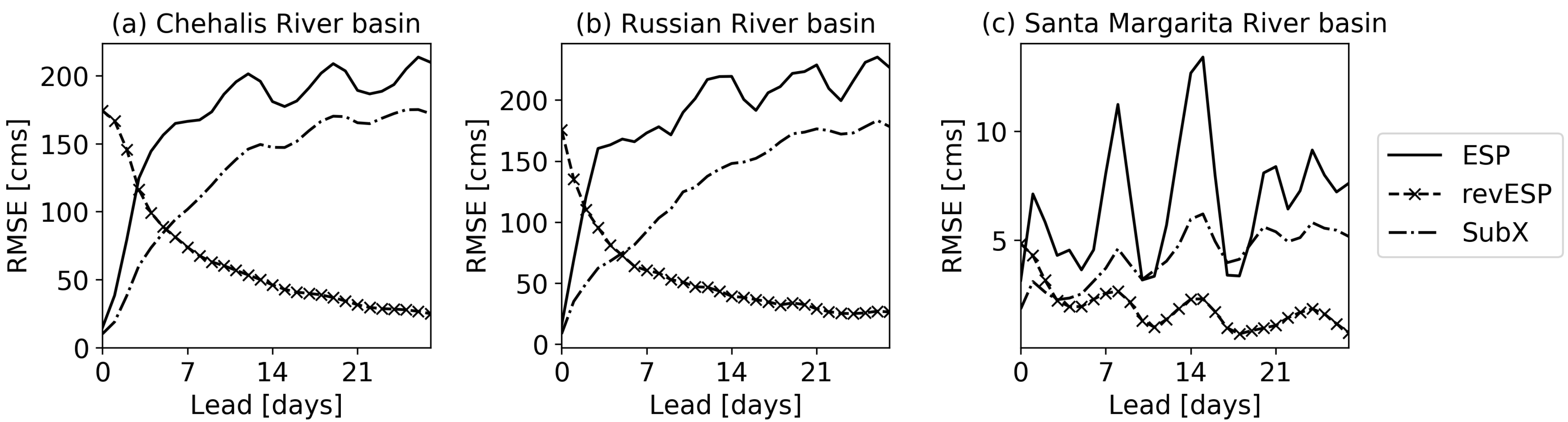


Relative influences of ASM and SubX reforecast skill on subseasonal flood forecast skill

The ESP/revESP method is used to partition the relative contributions of IHCs and meteorological forecast skill to errors in streamflow forecasts



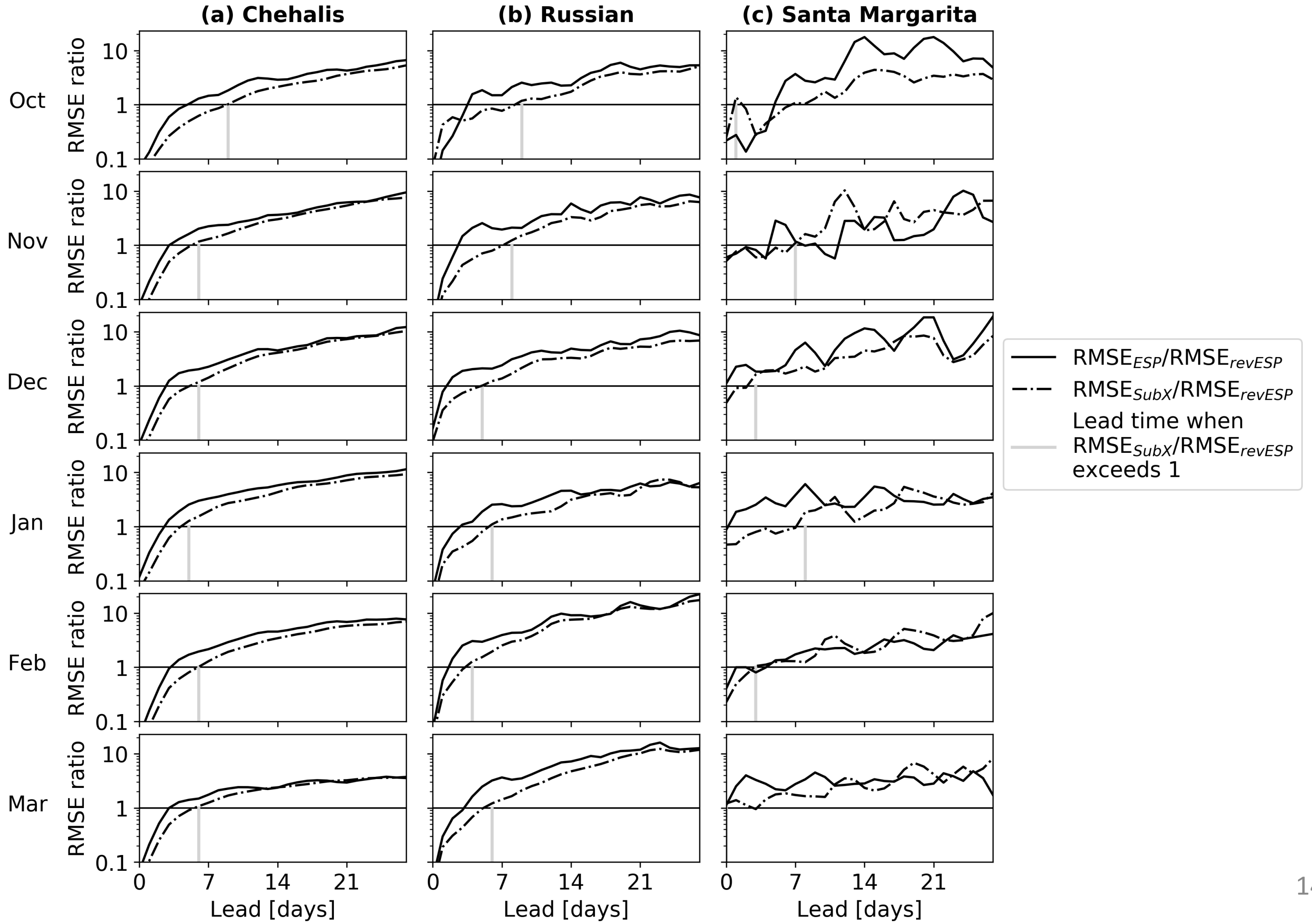
Wood and Lettenmaier, *GRL*, 2008



Relative influences of ASM and SubX reforecast skill on subseasonal flood forecast skill

Role of ASM in streamflow forecast

- ASM dominates streamflow deterministic forecast skill at leads up to 9 days with the maximum lead length occurring in Oct (following generally dry summers).



Relative influences of ASM and SubX reforecast skill on subseasonal flood forecast skill

Role of ASM in flood forecast

- ASM dominates flood probabilistic forecast skill only for small flood events in the three basins at week 1. For most large flood events (i.e. POT_{N1}) in the three basins, the SubX reforecast skill dominates the flood probabilistic forecast skill at all weeks.

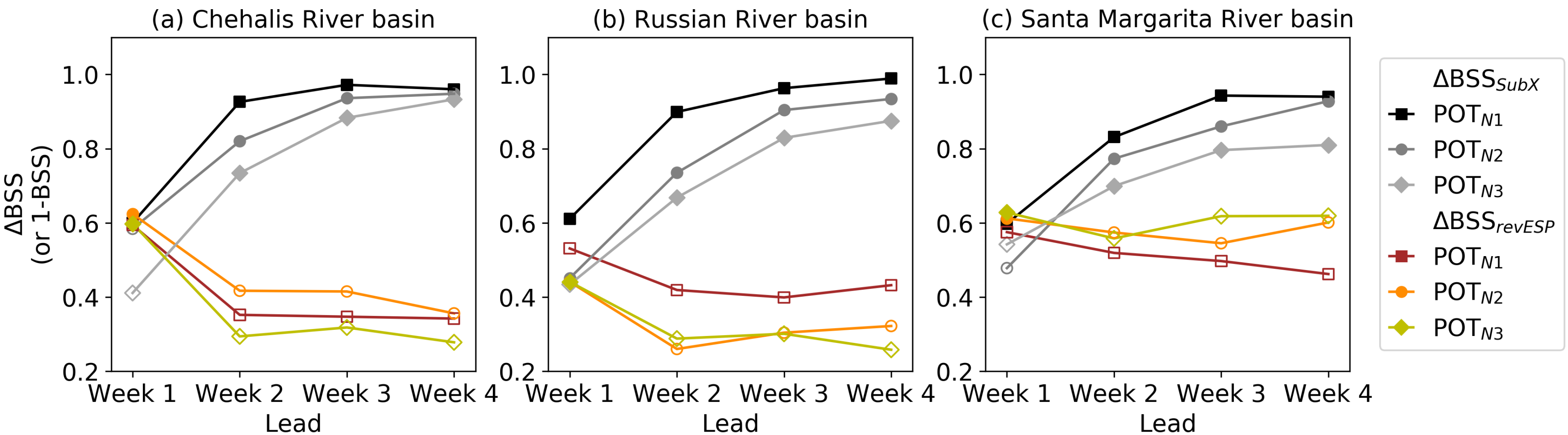


Figure. BSS difference (denoted as "ΔBSS") between perfect skill (i.e. BSS=1) and SubX-based BSS (denoted as "ΔBSS_{SubX}") as well as revESP-based BSS (denoted as "ΔBSS_{revESP}"). If $\Delta BSS_{SubX} \geq \Delta BSS_{revESP}$, the marker of ΔBSS_{revESP} is shown as a hollow symbol and vice versa.

Conclusions

1) Precipitation and temperature skill

- SubX precipitation forecast skill drops quickly after week 1 lead, but still has usable skill at week 2, while at week 3-4, models show minimal skill.
- There is higher skill in temperature than precipitation forecasts, with all models showing usable skill through lead 4 weeks.
- Across models, NCEP-CFSv2 performs best in weeks 1-2, with performance that is comparable with MME, while in weeks 3-4 GMAO-GEOS_V2p1 generally performs best for precipitation and EMC-GEFS performs best for temperature across the three basins.

2) Flood forecast skill

- The deterministic forecast skill of NCEP-CFSv2 drops quickly with lead time, with little skill by lead days 9, 7, and 6 in the Chehalis, Russian and Santa Margarita River basins respectively for the largest (POT_{N1}) events.
- SubX-based probabilistic skill drops quickly after week 1, with minimal forecast skill by week 3. Forecast skill is slightly higher for small events (lower POT thresholds).

3) Role of ASM in flood forecast

- ASM dominates flood probabilistic forecast skill only for small flood events in the three basins at week 1. For most large flood events in the three basins, the SubX reforecast skill dominates the flood probabilistic forecast skill at all weeks.

Acknowledgement

The research support to co-authors of this paper was provided by the Center for Western Weather and Water Extremes (CW3E) at the Scripps Institution of Oceanography UC San Diego via AR Program Phase II, sponsored by the California Department of Water Resources, and NOAA Regional Integrated Sciences and Assessments (RISA)'s support through the California–Nevada Applications Program.

This presentation:

Cao, Q., S. Shukla, M.J. DeFlorio, F.M. Ralph, and D.P. Lettenmaier, 2020: Evaluation of the subseasonal forecast skill of floods associated with atmospheric rivers in coastal Western U.S. watersheds, *Journal of Hydrometeorology*, (under review).

Land Surface Hydrology Research Group

University of California, Los Angeles

Climate Hazards Center

UC SANTA BARBARA

Center for Western Weather and Water Extremes

SCRIPPS INSTITUTION OF OCEANOGRAPHY

References

Bohn, T. J., B. Livneh, J. W. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. Agricultural and Forest Meteorology, *Agr. Forest Meteorol.*, **176**, 38-49, https://doi.org/10.1016/j.agrformet.2013.03.003.

Guan, B., and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies, *J. Geophys. Res. Atmos.*, **120**, 12514-12535, https://doi.org/10.1002/2015JD024257

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *J. Hydrol.*, **377**, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Lang, M., T. Ouarda, and B. Bobee, 1999: Towards operational guidelines for over-threshold modeling. *J. Hydrol.*, **225**, 103-117, https://doi.org/10.1016/s0022-1694(99)00167-5.

Livneh B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer, and D.P. Lettenmaier, 2013: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Clim.*, **26**, 9384–9392, https://doi.org/10.1175/JCLI-D-12-00508.1

Pegion, K., and coauthors, 2019: The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1

Ralph, F.M., P.J. Neiman, G. Wick, S. Gutman, M. Dettinger, D. Cayan, and A. White, 2006: Flooding on California's Russian River: Role of atmospheric rivers. *Geophys. Res. Lett.*, **33**, L13801, https://doi.org/10.1029/2006GL026689.

——, J.J. Rutz, J.M. Cordeira, M.D. Dettinger, M.L. Anderson, D. Reynolds, L.J. Schick, and C. Smallcomb, 2019: A Scale to Characterize the Strength and Impacts of Atmospheric Rivers. *Bull. Amer. Meteor. Soc.,* **100**, 269-289, https://doi.org/10.1175/BAMS-D-18-0023.1.

USWRC, 1982: Guidelines for determining flood flow frequency. Bulletin 17B of the Hydrology Subcommittee, 183 pp., https://water.usgs.gov/osw/bulletin17b/dl_flow.pdf.

Vitart, F., and Coauthors, 2017: The Subseasonal to Seasonal (S2S) Prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1.

Weigel, A.P., M.A. Liniger, and C. Appenzeller, 2007: The Discrete Brier and Ranked Probability Skill Scores. *Mon. Wea. Rev.*, **135**, 118–124, https://doi.org/10.1175/MWR3280.1

Westra, S., R. Mehrotra, A. Sharma, and R. Srikanthan, 2012: Continuous Rainfall Simulation: 1—A regionalised sub-daily disaggregation approach, *Water Resour. Res.*, **48**, W01535, https://doi.org/10.1029/2011WR010489.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679, https://doi.org/10.1029/94WR00436.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 648 pp.