

Supplement of Hydrol. Earth Syst. Sci., 21, 3463–3481, 2017  
<https://doi.org/10.5194/hess-21-3463-2017-supplement>  
© Author(s) 2017. This work is distributed under  
the Creative Commons Attribution 3.0 License.



*Supplement of*

## **Variations in the correlation between teleconnections and Taiwan's streamflow**

**Chia-Jeng Chen and Tsung-Yu Lee**

*Correspondence to:* Tsung-Yu Lee (tylee@ntnu.edu.tw)

The copyright of individual parts of the supplement might differ from the CC BY 3.0 License.

### **Caveats of using the CRS detection method:**

As already indicated by the manuscript, there are several classes of approaches available for the detection of regime shifts, such as parametric methods (e.g., the classical  $t$ -test), non-parametric methods (e.g., the Mann-Whitney-Pettitt test), regression-based methods, cumulative sum methods, and sequential methods. Rodionov (2005) pointed out the pros and cons of some common approaches as well as his sequential method (Rodionov, 2004). The pros of his method are the automatic, early detection of a regime shift and the ability to monitor a possibility of a regime shift in real time. He has shown that his method can outperform Lanzante's method (another robust, non-parametric procedure developed by Lanzante, 1996). Therefore, since developed, Rodionov's method has been used in many studies in climate sciences. Nevertheless, the cons of his method are the requirement of some experimentation on the two parameters used (i.e., the cut-off length  $l$  and probability level  $p$ ) and inability to account for a gradual regime shift and data with obvious autocorrelation (or red noise, but this issue was later ameliorated by a prewhitening procedure introduced by him, Rodionov, 2006). According to Rodionov (2004), while the probability level  $p$  is known to determine the critical value of  $t$  (the Student's  $t$ -distribution), *"the cut-off length  $l$  determines the minimum length of the regimes, for which the magnitude of the shifts remains intact."* It has been tested that, the larger  $l$  is set, the fewer change points can be identified. By contrast, a smaller  $l$  does not necessarily lead to more change points since only those significant change points can be identified based on the  $t$ -test. In other words, if there is no strong regime shift in the data series, the method with some variations in  $p$  and  $l$  simply cannot identify any change point (Rodionov, 2015 and verified by us too). In any case, we should still use Rodionov's method with caution since the CRS detection method is purely statistical, and the CRS results could be meaningless without solid evidence from related research.

### **Hindcasting experiment using linear regression:**

Linear regression is widely used in climate forecasting studies (e.g., Hastenrath, 1995; Hastenrath et al., 2004; Chen and Georgakakos, 2015) to generate forecasts based on a calibrated, linear equation that depicts how a hydro-climatic predictand ( $\mathbf{Y}$ , the streamflow in our case) responds to selected predictors ( $\mathbf{X}$ , teleconnection indices in our case):

$$\mathbf{Y} = \mathbf{X}\beta,$$

where  $\beta$  represents coefficients estimated by ordinary least squares. In each catchment, we develop a linear regression equation as the prediction model. In terms of predictors (i.e., independent variables), we adopt 13 teleconnection indices described in the paper (with AAO excluded as relatively short in record) and perform stepwise model selection-based AIC (Akaike Information Criteria). Regarding model selection, we use both forward and backward directions to ensure a thorough search in the variable space. Afterwards, to avoid possible multicollinearity issues resulting from some highly correlated teleconnection indices, the variance inflation factor ( $VIF$ ) is assessed:

$$VIF_j = \frac{1}{1 - R_j^2},$$

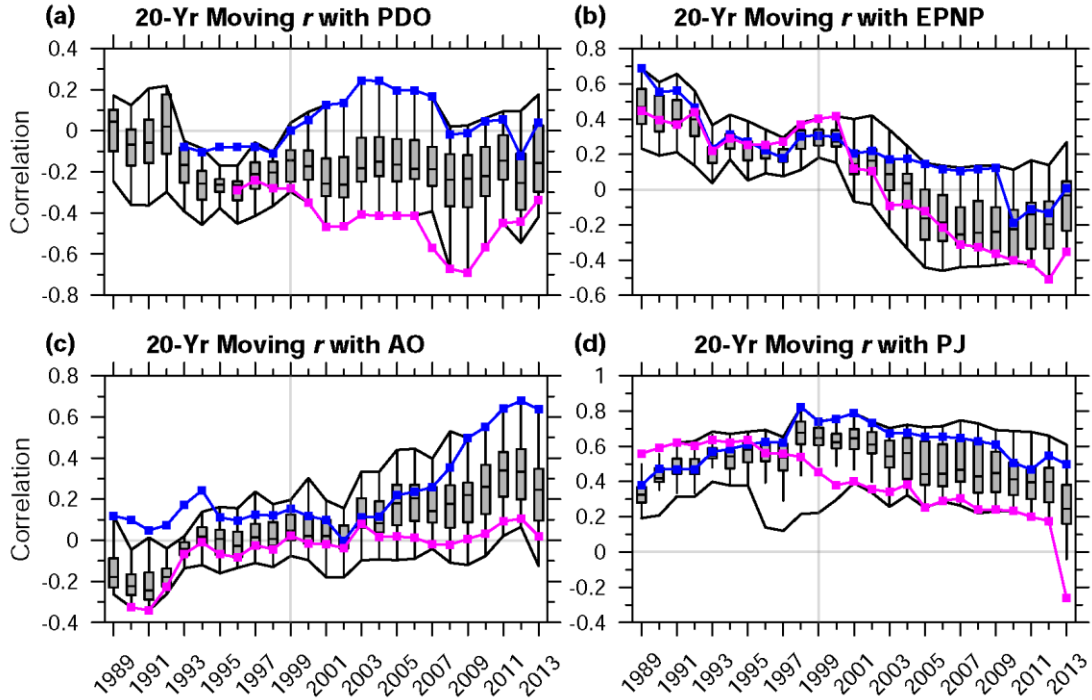
where  $R_j^2$  is the coefficient of determination from a regression of the  $j^{\text{th}}$  predictor on any other predictors. According to the literature (Chen and Georgakakos 2014; Hidalgo-Muñoz et al. 2015), the VIF tolerance threshold is set to be 4 for small samples. The final model is thus determined and used to generate hindcasts (i.e., retrospective forecasts) for that catchment. The generation of hindcasts is subject to the leave-one-out cross-validation (LOOCV) procedure to circumvent artificial skill. Lastly, the LOOCV correlation and Gerrity Skill Score (GSS, Gerrity, 1992) are calculated to assess the prediction skill in that catchment. The GSS, unlike other conventional skill scores (e.g., the Heidke skill score), is an “equitable skill score” with a specific scoring matrix known for rewarding correct forecasts in the less ordinary categories (e.g., high and low flows in our case). The GSS does not reward random or constant forecasts. The GSS has been adopted by a number of practitioners of climate forecasts (e.g., CPC and WMO) and many studies (Chen and Georgakakos 2014; Hidalgo-Muñoz et al. 2015). The above framework is repeated for all 41 catchments.

Figure S2 shows some hindcasting results for selected upstream and downstream catchments. LOOCV correlations vary from one catchment to another and can be as high as  $\sim 0.6$ . As a more stringent metric, cross-validated GSS values are generally lower but most of the time pass the significant threshold (e.g.,  $\sim 0.25$  for data size 30, determined by the bootstrap-based hypothesis testing, Chen and Georgakakos, 2014). Overall, using large-scale circulation indices can produce fair to good prediction skills in summer

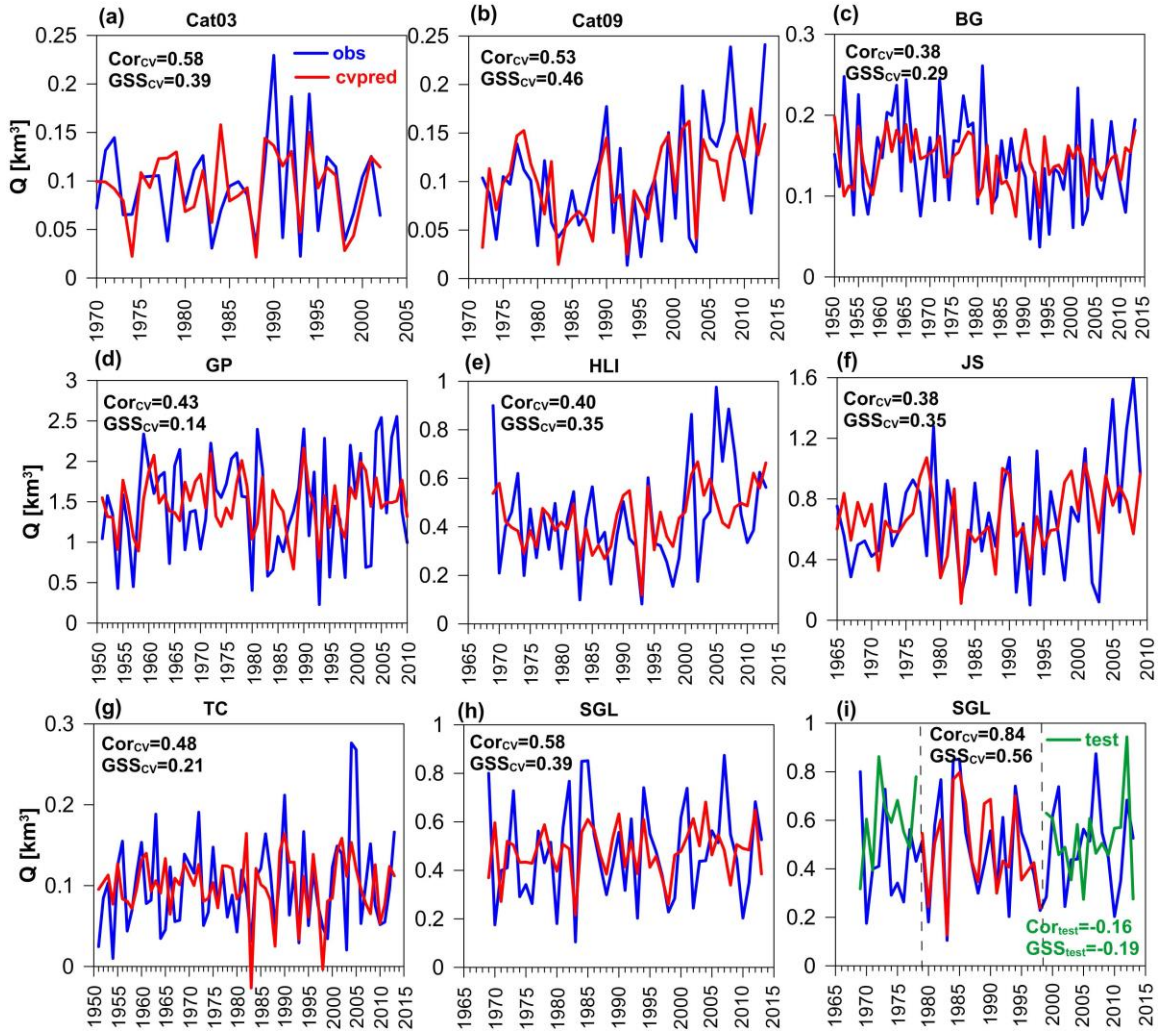
streamflow prediction in Taiwan. Among those many climate indices, the PDO and PJ indices are selected most frequently as the predictors for the catchments [while the former is selected seven times (except for SGL), the latter is selected five times (except for Catchments 3 and 9 and BG) for the results shown in Figure S2]. This result is, to a certain extent, consistent with our general correlation assessments (e.g., Table 3 in the main text) and indicates the general dominance of summer climate in Taiwan. Regarding the origin of the predictability, Chen and Chen (2011) indicated that the PDO coincides with the specific meridional SST contrast (i.e., warming in the tropical central and eastern Pacific and cooling in the extratropical North Pacific), which plays a dominant role in modulating summer rainfall in Taiwan. Choi et al. (2010), Kosaka et al. (2013), and Kubota et al. (2016) all provided sufficient evidence of the significant impact of the PJ on tropical cyclone activity and rainfall over the WNP during summer. Based on our findings, the predictability for summer rainfall can be extended to streamflow in Taiwan.

From Figure S2, we can note that the relatively better performance of each LOOCV time series seems to occur during the period from the late 1970s to the late 1990s, coinciding with the CRS epoch discussed earlier. Hindcasts during the pre-regime shift epoch seem to still be able to capture the general variability of the observed runoff (with relatively poorer performance), whereas hindcasts during the post-regime shift epoch appear to present more opposite signals and apparent departures from the observed runoff. It is worth noting that some of the departures occur in years when JAS typhoon activity is abnormally high. For example, in 2007, typhoons Pabuk, Sepat, and Wipha together generated the highest amount of cumulative rainfall for some watersheds over the past decade, and in 2008, typhoons Kalmaegi, Fung-Wong, Sinlaku, Hagupit, and Jangmi made a record of continuous invasions of intense typhoons (all Category 2 and above) in JAS. To further illustrate the effect of CRSs on streamflow prediction, we fitted a new regression model using the data from 1979 to 1998 and then evaluated how the fitted model performed in the remaining years. Using the SGL watershed as an example, Figure S2-i shows the new hindcasting result. In comparison with Figure S2-h, the new fitted model exhibits some definite improvement during the period from 1979 to 1998, showing the outstanding CV correlation and GSS values as 0.84 and 0.56, respectively. However, the fitted model can generate nothing but extremely poor hindcasts for the remainders. In fact, both skill metrics show a reverse sign, clearly illustrating distinct climate regimes over the temporal horizon. In contrast with the above experiment, if we fit another regression model using the data outside the time frame of 1979–98, the stepwise model selection scheme discloses no climate indices that are qualified to be a predictor. To summarize, the linear regression experiment indicates that the assumption of a stable predictor-predictand relationship could be quite problematic for hydro-climatic forecasting due to the observation of CRS.

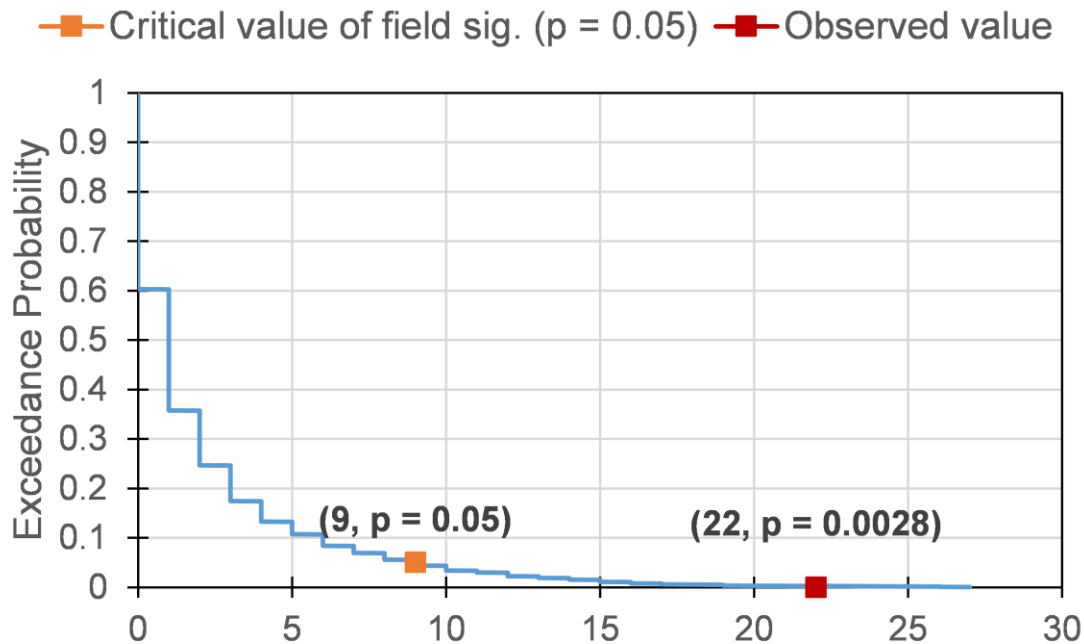
**Supplementary Figures:**



**Figure S1.** Selected moving-window correlation results. Each boxplot encapsulates correlation values derived from 28 upstream catchments (JAS runoff vs. specific teleconnection index). The year label indicates the end of a moving window (e.g., 1989 indicates the moving window is taken from 1970 to 1989). Blue (magenta) time series denotes the highest (lowest) moving-window correlations over the temporal horizon, and can go outside the box plot range (i.e., deemed an outlier) since the IQR with a factor of 1.5 is used to determine a rather diminished range.



**Figure S2.** (a) to (h) are selected hindcasting results for upstream and downstream catchments in Taiwan using linear regression. Time series in red are model estimates based on the leave-one-out cross-validation (LOOCV) procedure. Cross-validated (CV) correlation and GSS values are also denoted in each plot. (i) is similar to (h), but the linear regression model is trained/fitted with the data from 1979 to 1998, and then the fitted model is tested with the rest of the data points (i.e., 1969–78 and 1999–2013)



**Figure S3.** Empirical null distribution of field significance test. The abscissa represents the number of catchments with correlations between streamflow and lagged QBO significant at the 95% level ( $p = 0.05$ ) in 2000 Monte Carlo trials.

## References

- Chen, C.-J. and Georgakakos, A. P.: Hydro-climatic forecasting using sea surface temperatures: methodology and application for the southeast US, *Climate Dyn.*, 42, 2955–2982, 2014.
- Chen, C.-J. and Georgakakos, A. P.: Seasonal prediction of East African rainfall, *Int. J. Climatol.*, 35(10), 2698–2723, 2015.
- Chen, J. M. and Chen, H. S.: Interdecadal variability of summer rainfall in Taiwan associated with tropical cyclones and monsoon, *J. Climate*, 24(22), 5786–5798, 2011.
- Choi, K. S., Wu, C. C., and Cha, E. J.: Change of tropical cyclone activity by Pacific-Japan teleconnection pattern in the western North Pacific, *J. Geophys. Res.*, 115, doi:10.1029/2010JD013866, 2010.
- Hastenrath, S.: Recent advances in tropical climate prediction, *J. Climate*, 8(6), 1519–1532, 1995.

Hastenrath, S., Polzin, D., and Camberlin, P.: Exploring the predictability of the 'short rains' at the coast of East Africa, *Int. J. Climatol.*, 24(11), 1333–1343, 2004.

Hidalgo-Muñoz, J. M., Gámiz-Fortis, S. R., Castro-Díez, Y., Argüeso, D., and Esteban-Parra, M. J.: Long-range seasonal streamflow forecasting over the Iberian Peninsula using large-scale atmospheric and oceanic information, *Water Resour. Res.*, 51(5), 3543–3567, 2015.

Kosaka, Y., Xie, S. P., Lau, N. C., and Vecchi, G. A.: Origin of seasonal predictability for summer climate over the Northwestern Pacific, *Proc. Natl. Acad. Sci. (USA)*, 110, 7574–7579, 2013.

Kubota, H., Kosaka, Y., and Xie, S. P.: A 117-year long index of the Pacific Japan pattern with application to interdecadal variability, *Int. Climatol.*, 36, 1575–1589, 2016.

Lanzante, J. R.: Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. Climatol.*, 16(11), 1197–1226, 1996.

Rodionov, S. N.: A brief overview of the regime shift detection methods. Large-Scale Disturbances (Regime Shifts) and Recovery in Aquatic Ecosystems: Challenges for Management Toward Sustainability, pp.17–24, 2005.

Rodionov, S. N.: Use of prewhitening in climate regime shift detection. *Geophys. Res. Lett.*, 33(12), 2006.