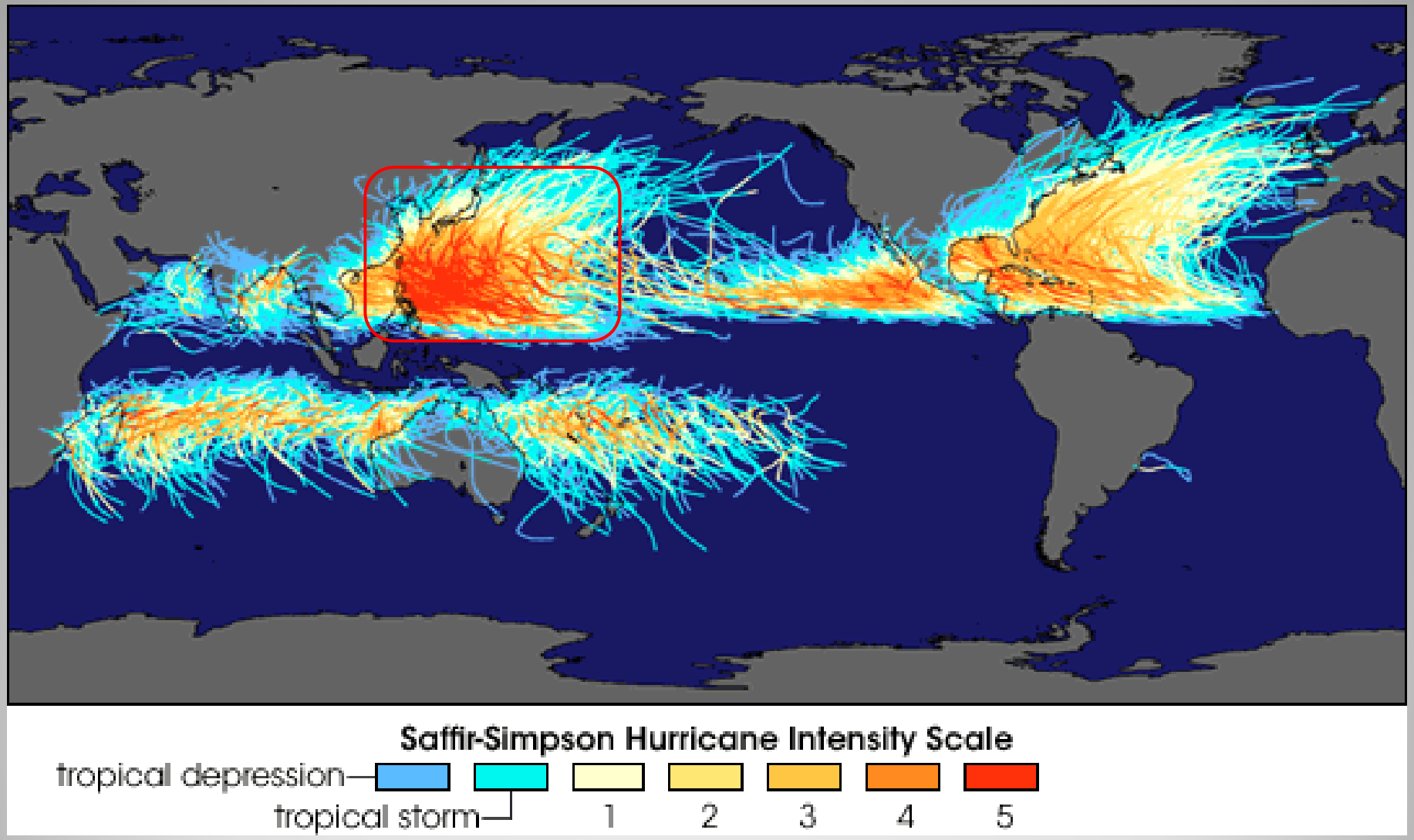
Statistical-Dynamical Seasonal Forecast of Western North Pacific and East Asia Landfalling Tropical Cyclones (TCs) using the High-Resolution GFDL FLOR Coupled Model

Wei Zhang [1,2], Gabriele Villarini[3] , Gabriel Vecchi[1,2], Hiroyuki Murakami[1,2]

[1]National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

[2]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA

[3] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, Iowa



http://earthobservatory.nasa.gov/IOTD/view.php?id=7079

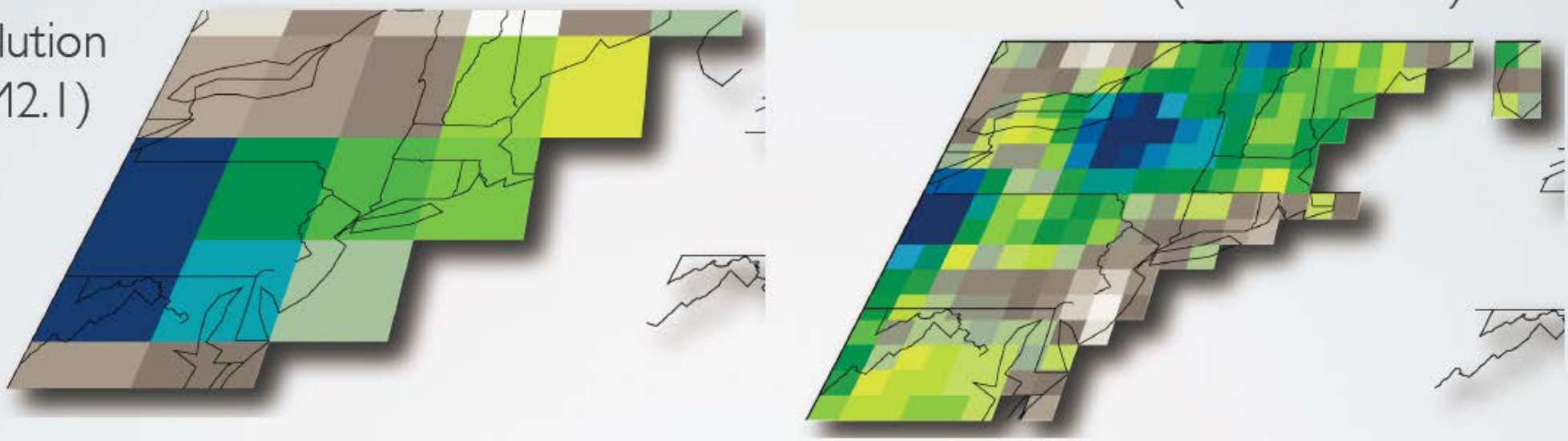
GFDL FLOR: Experimental high-resolution coupled seasonal to decadal prediction system

Goal: Build a seasonal to decadal forecasting system to:
Yield improved forecasts of large-scale climate
Enable forecasts of regional climate and extremes

Precipitation in Northeast USA

Medium resolution (CM2.1)

High resolution (CM2.5-FLOR)



Delworth et al. (2012), Vecchi et al. (2014)

Modified version of CM2.5 (Delworth et al. 2012):

- 50km cubed-sphere atmosphere
- 1° ocean/sea ice (low res enables prediction work)

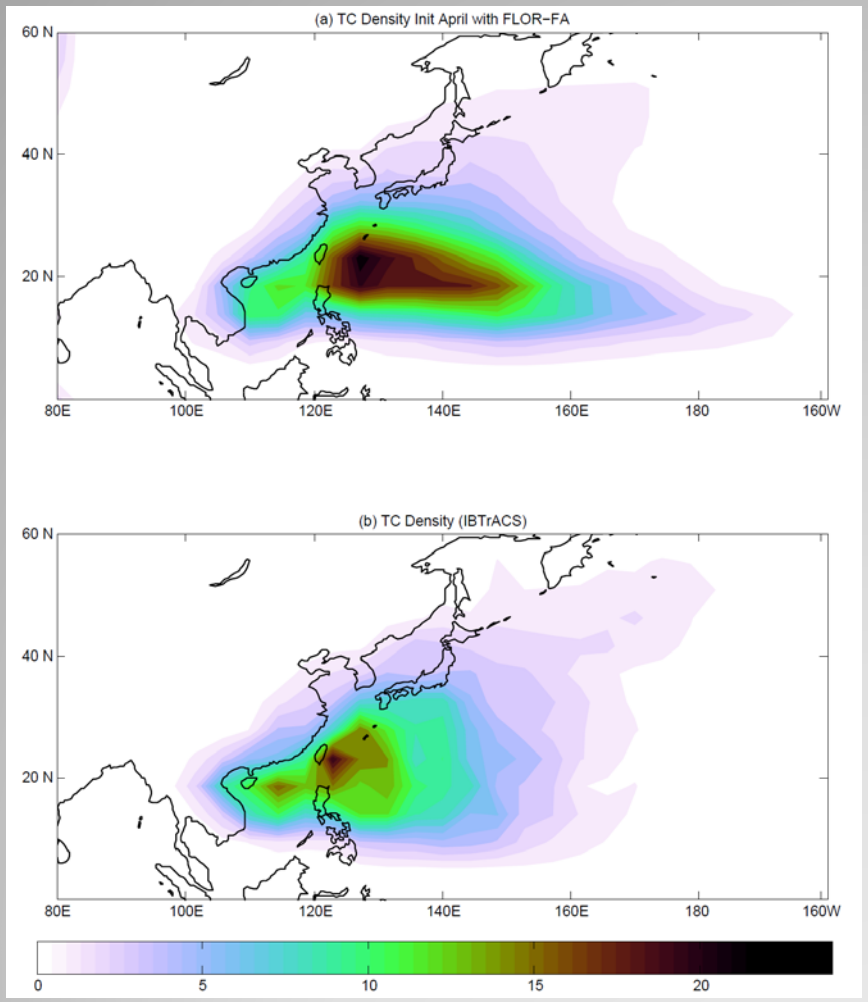
~15-18 years per day allows multi-century integrations.
Contributed to NMME from March 2014.

CM2.5: 28 km in the tropics ;
8 km at high latitudes

Courtesy of Vecchi (2014).

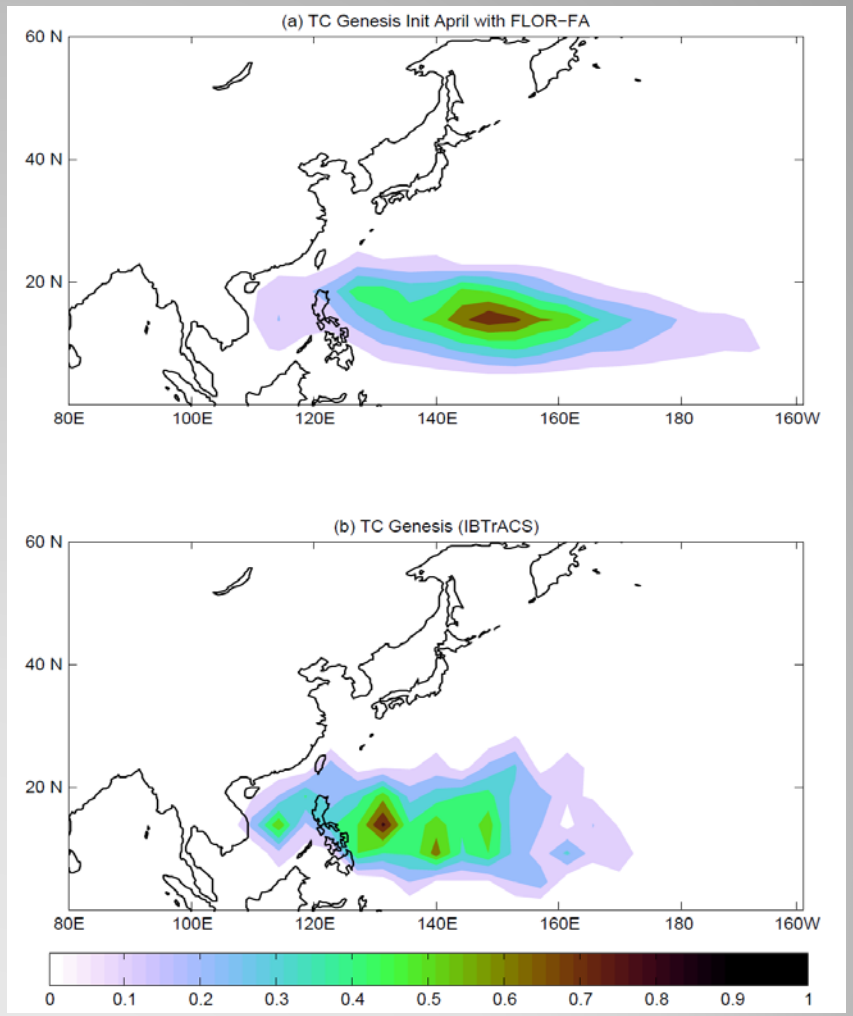
Performance of GFDL Forecast-oriented Low Ocean Resolution version of CM2.5 with Flux Adjustment (FLOR-FA) in Forecasting Tropical Cyclone in the Western North Pacific

5×5 degree grid boxes



Simulated and observed TC track density in the western North Pacific over 1980-2013.

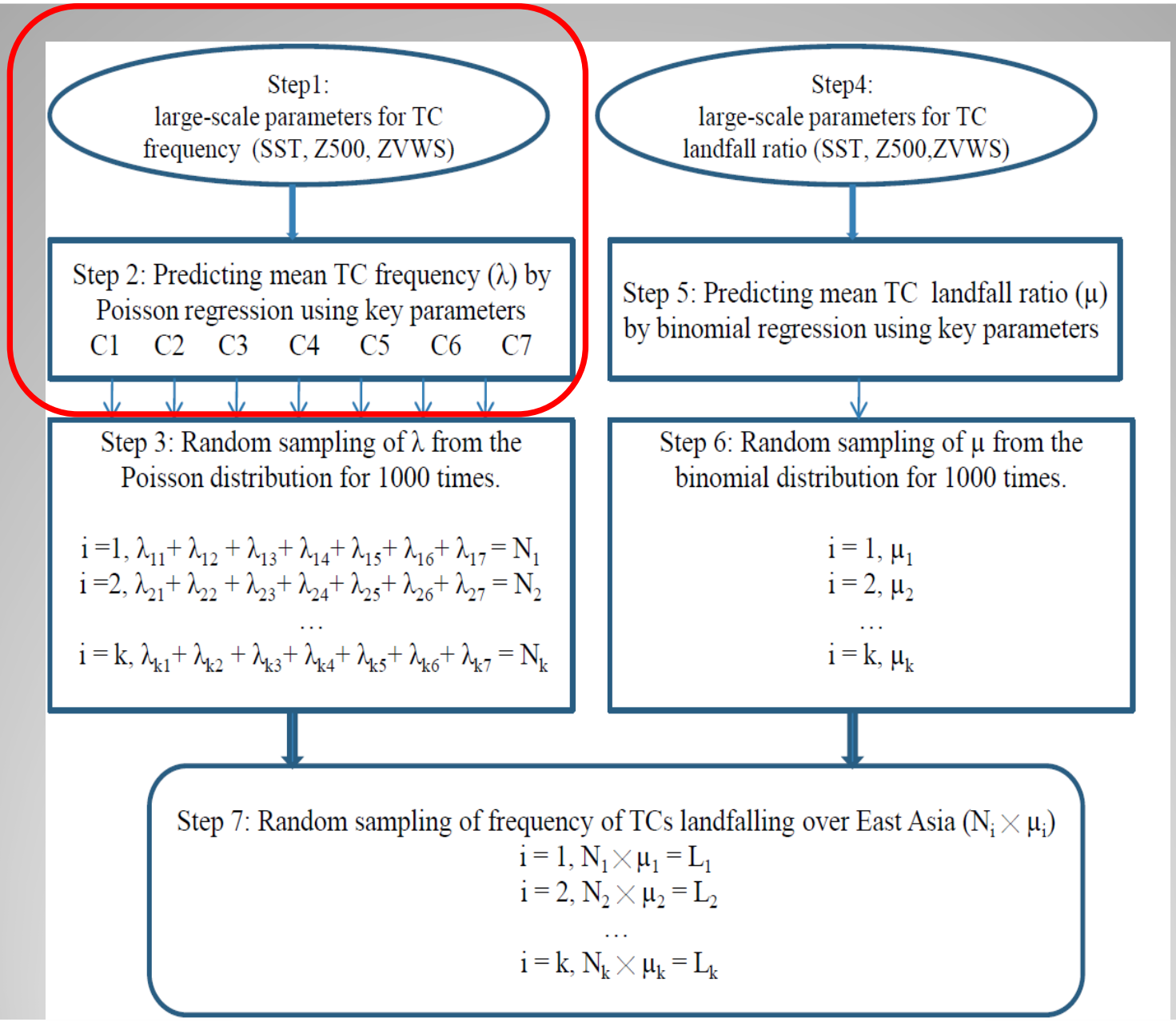
5×5 degree grid boxes



Simulated and observed TC genesis density in the western North Pacific over 1980-2013

Dynamical-statistical (Hybrid) Model for Forecasting TC

- The hybrid dynamical-statistical models are based on **statistical relationships** between TC frequency/landfall and physical processes tied to dynamical models.
- Hybrid models take advantage of both statistical associations derived from the observations and dynamic models, and environmental variables simulated by dynamic models to improve the skill of the prediction(e.g., Klotzbach, 2007; Elsner et al., 2008; Vecchi et al., 2010; LaRow, 2013;Li et al., 2013; Villarini and Vecchi, 2013; Vecchi et al., 2013, 2014; Kim et al., 2014).


Step1:
large-scale parameters for TC frequency (SST, Z500, ZVWS)
Step4:
large-scale parameters for TC landfall ratio (SST, Z500,ZVWS)
Step 2: Predicting mean TC frequency (λ) by Poisson regression using key parameters
C1 C2 C3 C4 C5 C6 C7
Step 5: Predicting mean TC landfall ratio (μ) by binomial regression using key parameters
Step 3: Random sampling of λ from the Poisson distribution for 1000 times.
$i=1, \lambda_{11}+\lambda_{12}+\lambda_{13}+\lambda_{14}+\lambda_{15}+\lambda_{16}+\lambda_{17}=N_1$
$i=2, \lambda_{21}+\lambda_{22}+\lambda_{23}+\lambda_{24}+\lambda_{25}+\lambda_{26}+\lambda_{27}=N_2$
…
$i=k, \lambda_{k1}+\lambda_{k2}+\lambda_{k3}+\lambda_{k4}+\lambda_{k5}+\lambda_{k6}+\lambda_{k7}=N_k$
Step 6: Random sampling of μ from the binomial distribution for 1000 times.
$i=1, \mu_1$
$i=2, \mu_2$
…
$i=k, \mu_k$
Step 7: Random sampling of frequency of TCs landfalling over East Asia ($N_i \times \mu_i$)
$i=1, N_1 \times \mu_1 = L_1$
$i=2, N_2 \times \mu_2 = L_2$
…
$i=k, N_k \times \mu_k = L_k$

Cluster Analysis of Tropical Cyclone Tracks

- Finite mixture model (FMM)-based cluster analysis (Camargo et al., 2007a;b; Zhang et al., 2013)

In this study, the model is constructed following Camargo et al. (2007a, b)

$$y_i = T_i\beta + \varepsilon_i, \varepsilon_i \sim \mathcal{N}(0, \Sigma_k), \quad (1)$$

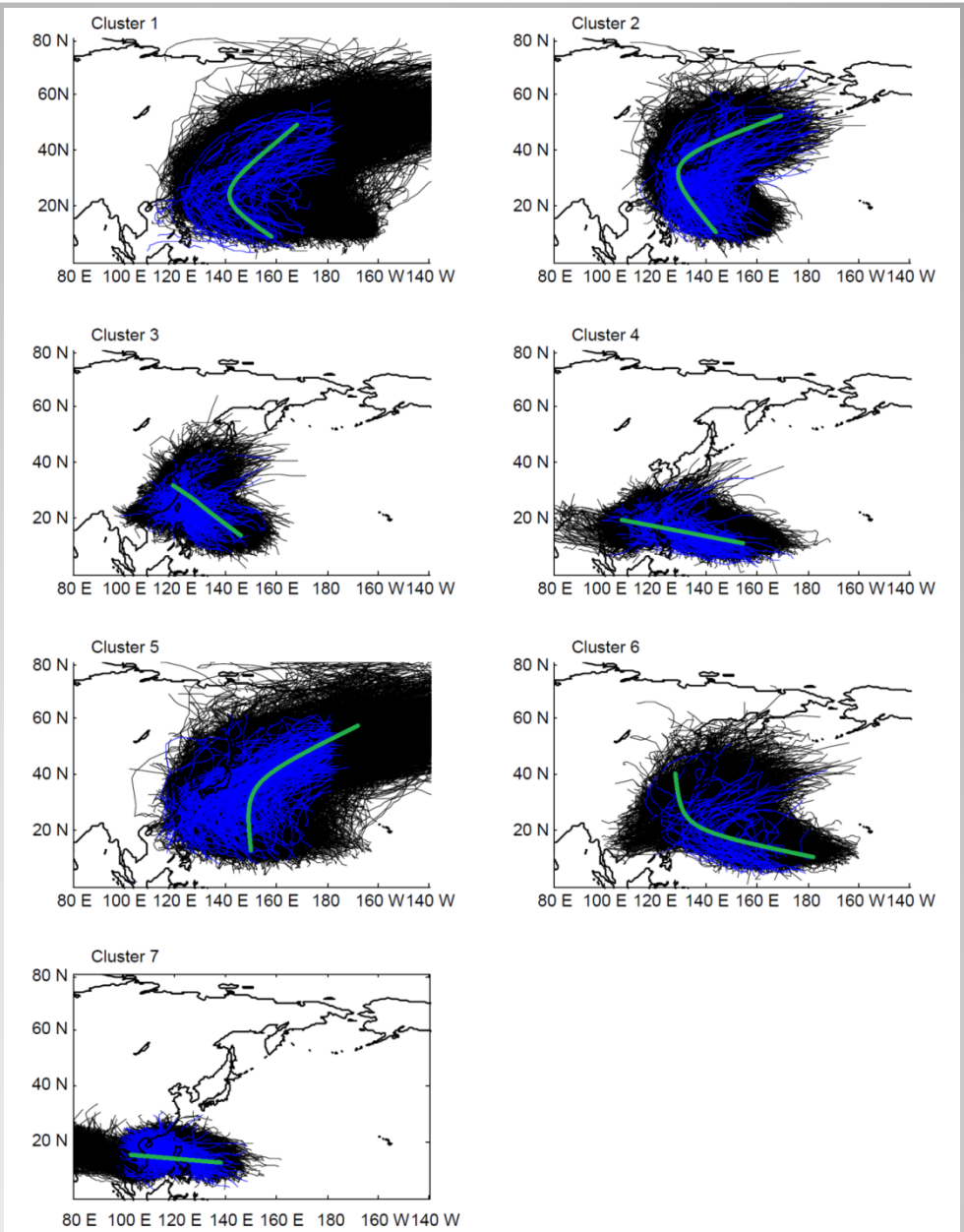
$$P(y_i \mid T_i) = \mathcal{N}(X_i\beta, \Sigma_k), \quad (2)$$

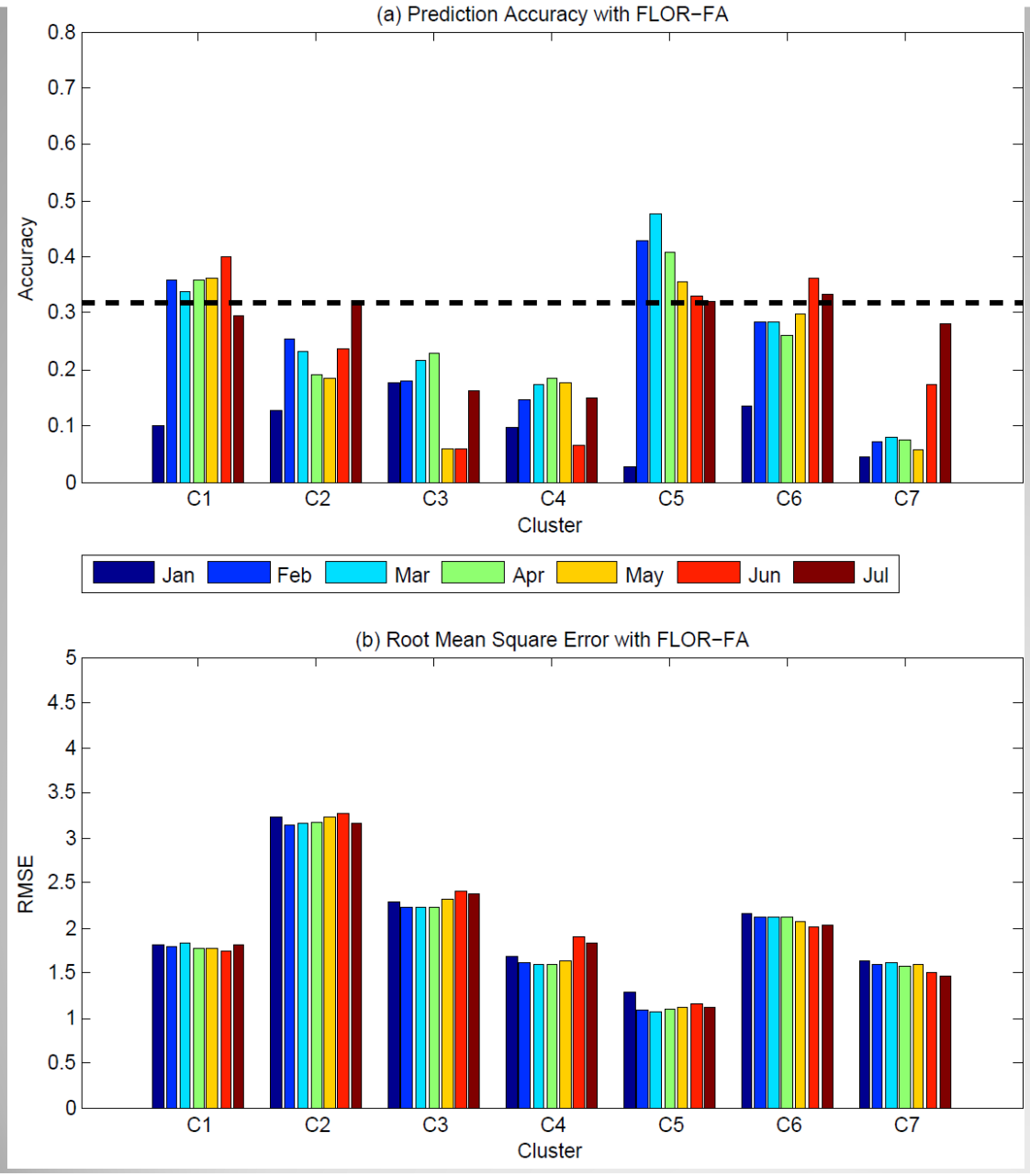
where

$$T_i = \begin{pmatrix} 1 & \cdots & x_{i1}^p \\ \vdots & \ddots & \vdots \\ 1 & \cdots & x_{in}^p \end{pmatrix}, \quad (3)$$

is the regression matrix with an n×p dimension; β denotes the regression coefficients, y_i is the ith TC track consisting of n latitude×longitude points (i.e., observations), ε_i is an n_i×2 matrix of multivariate Gaussian noise with zero mean and 2×2 covariance matrix Σ_k containing the diagonal elements σ_{k1}^2 and σ_{k2}^2 (noise variances of each longitude and latitude observation).

Seven clusters of TCs from both observations and 12-member FLOR-FA predictions.





$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(P_i - O_i\right)^2}{n}}$$

P_i: predicted TC frequency;
O_i: observed TC frequency.

Prediction skills measured by correlation and root mean square error of each cluster at different initial months with FLOR-FA.

Prediction Skills for Potential Predictors of TCs in FLOR-FA

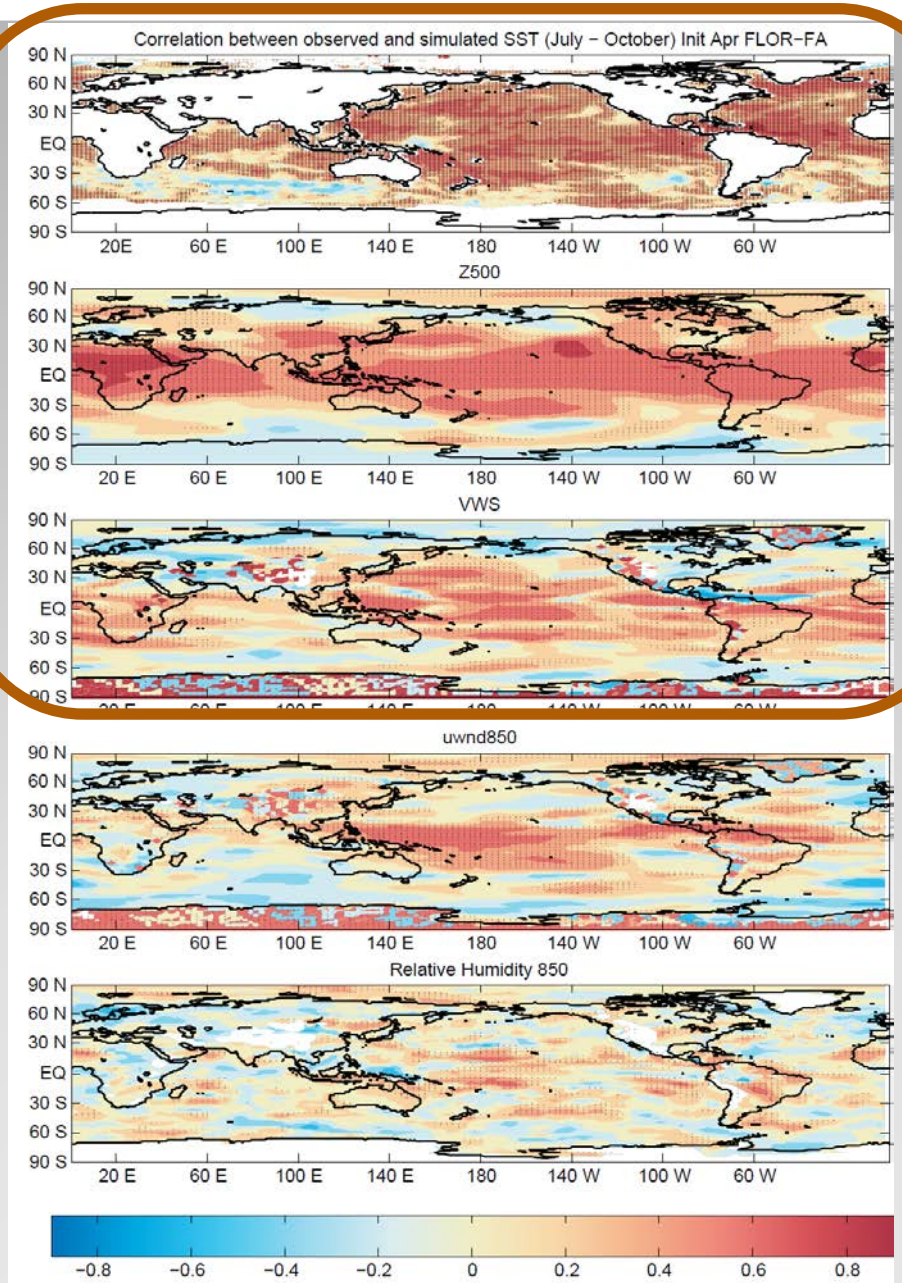
- SST
- 500-hPa geopotential height
- Vertical wind shear (200hPa – 850hPa)
- 850-hPa zonal wind(u)
- 700-hPa relative humidity

(e.g., Chan et al., 1998; 2001; Li et al., 2013)

SST

Z500

VWS

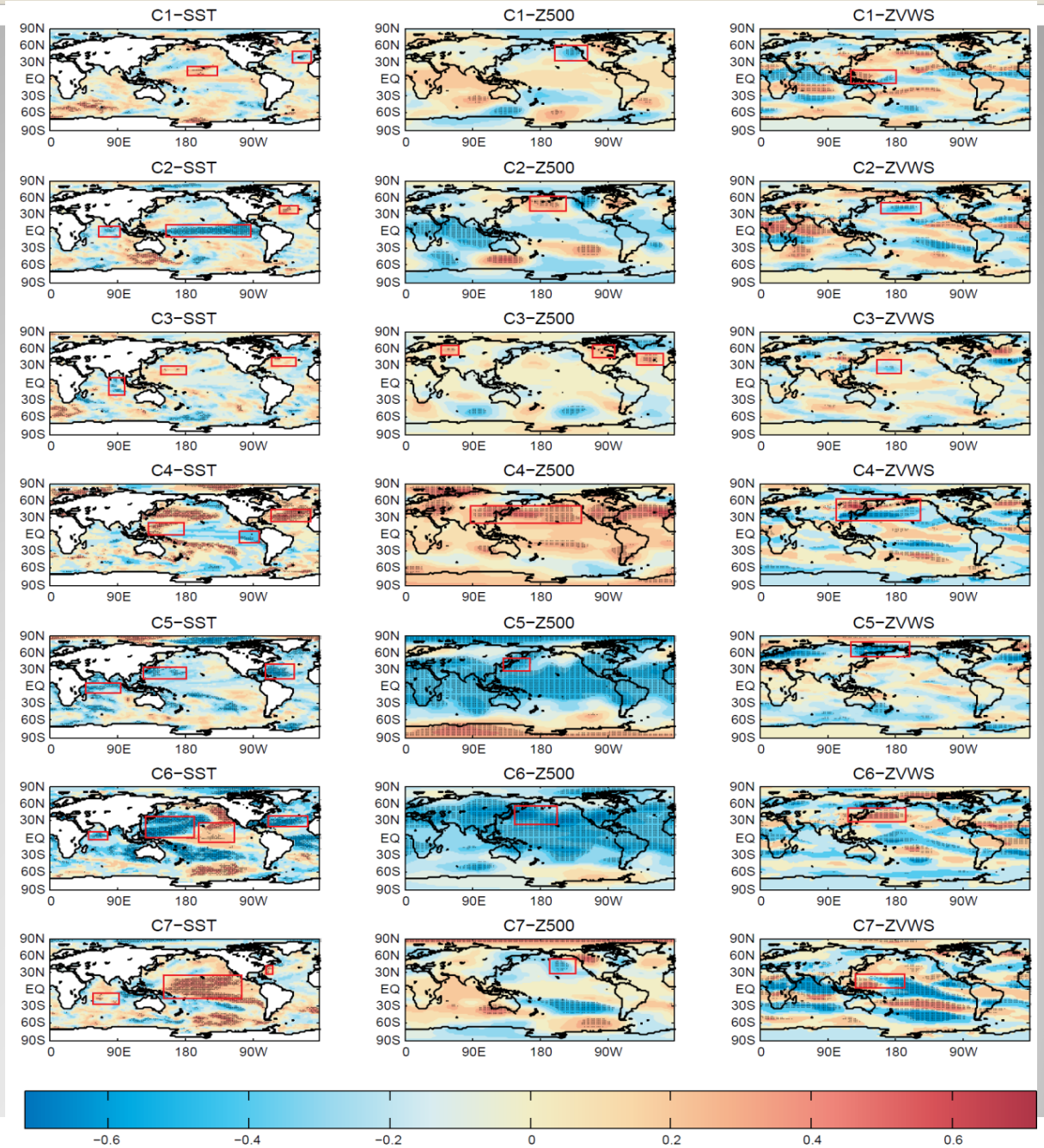


Initialized in April using FLOR-FA

How to find spatial domains for the predictors?

- Correlation between the predictors (SST, VWS and GPH-500) and the observed frequency of TCs in each cluster in the WNP during peak season (July to October).
- Find domains similar for each cluster in all initial months (i.e., January to July).

Spatial domains based on correlation and physical interpretations



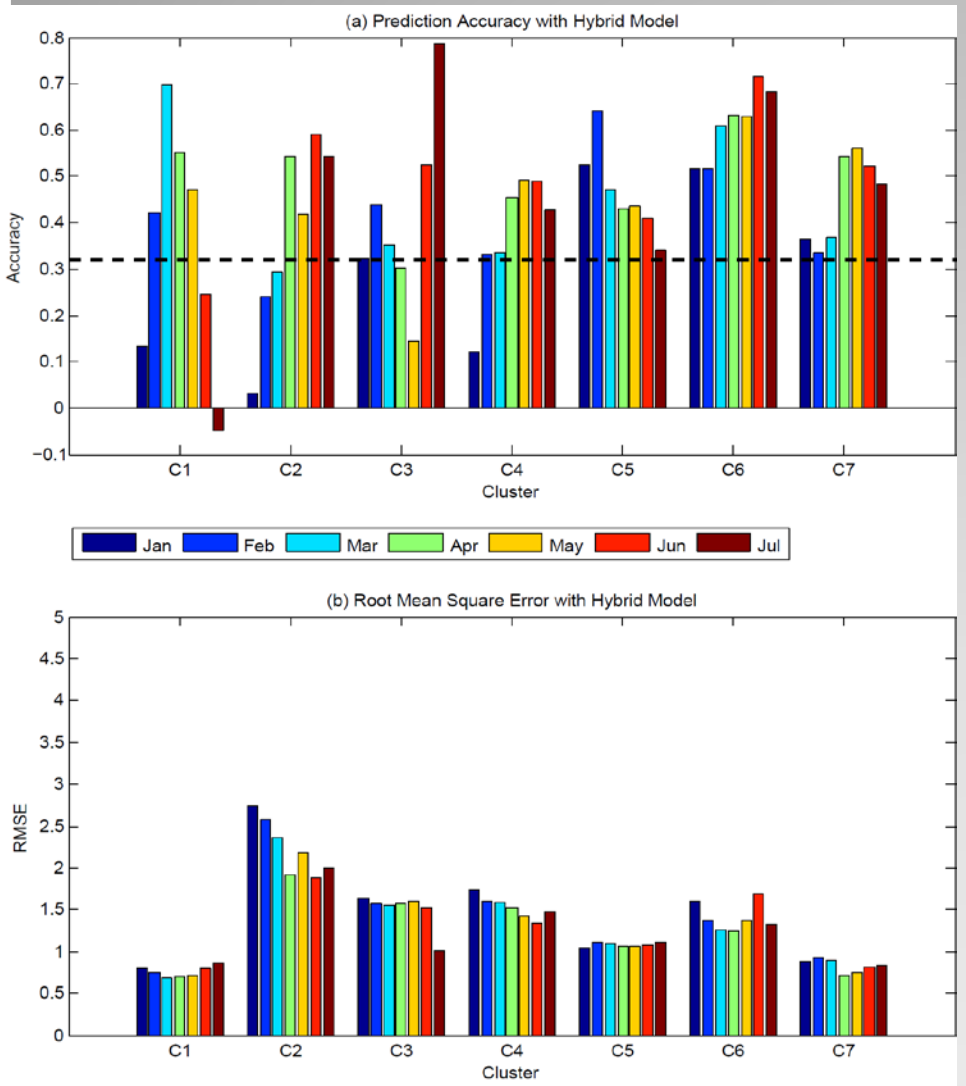
Poisson regression model has been widely used to analyze the probability for the occurrence of rare events (e.g., hurricanes, tornados and epidemics; e.g., Elsner and Jagger, 2006; Villarini et al., 2010; Wilks, 2011).

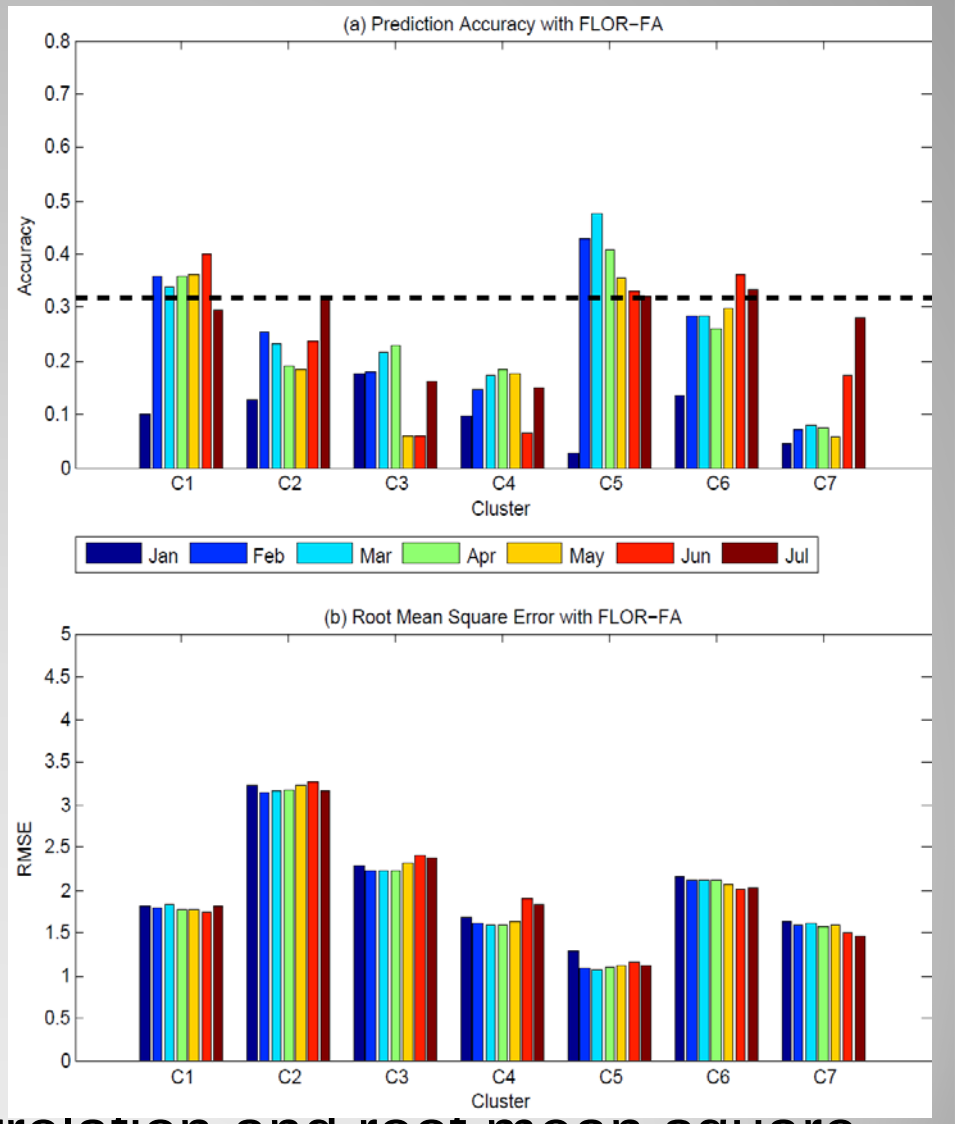
$$P(Y_i = y) = \frac{\mu_i^y e^{-\mu_i}}{y!}, y = 0, 1, 2 \ \ldots \infty \qquad (11)$$

where $\mu_i = \exp(\beta_0 + \sum_j(\beta_j x_{ij})$

If Y follows a Poisson distribution, the logarithm of the expected number of landfalling TC occurrence, μ, can be modeled as a linear combination of the predictors x_{ij}, with j being the specified predictor during the season i; β_j is the corresponding Poisson regression coefficient, β_0 the intercept, and y the observed count of landfalling TCs. In a Poisson model, the variance is equal to the mean (λ).

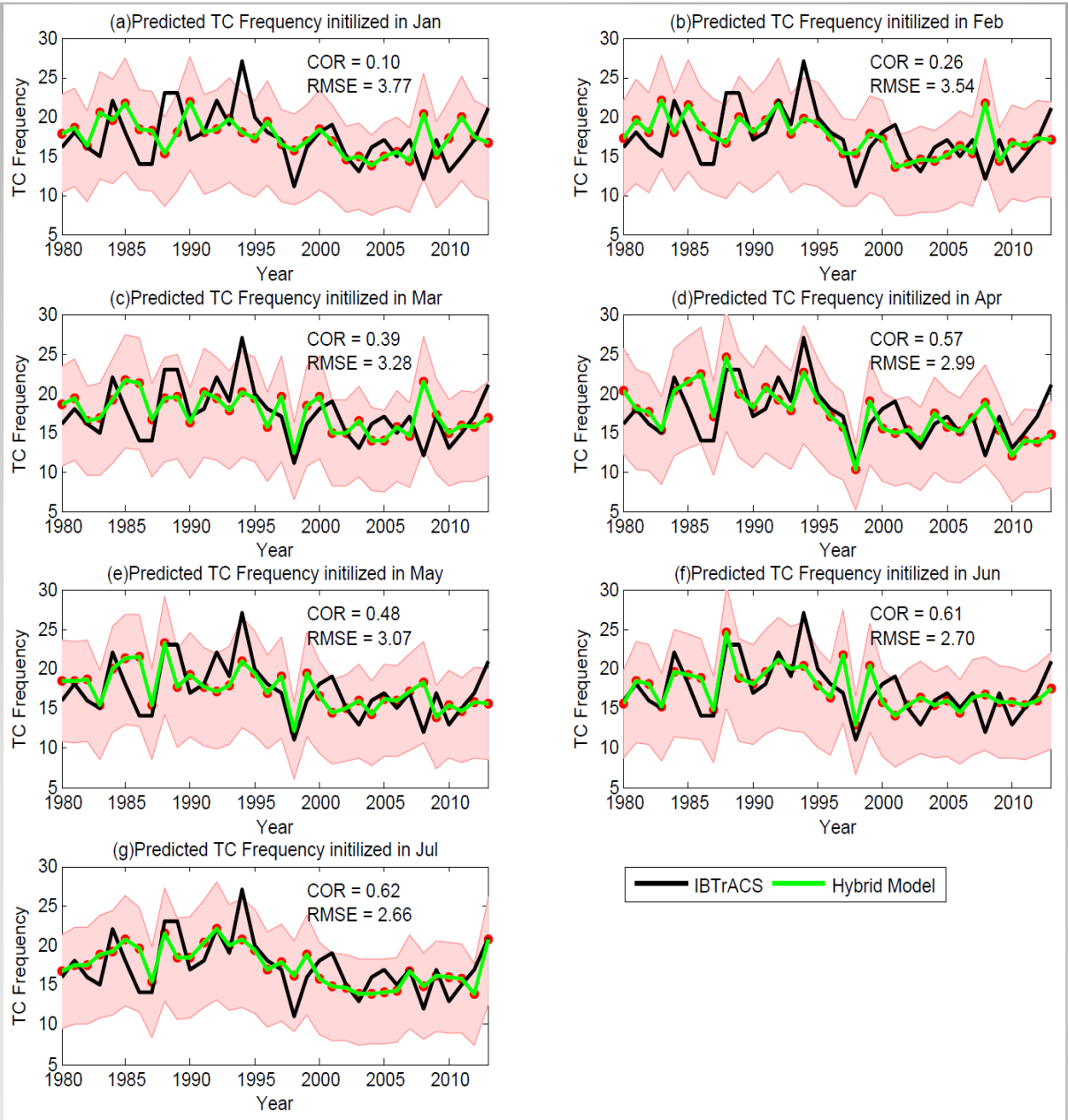
(Zhang et al., 2015)





Prediction skills measured by correlation and root mean square error of each cluster at different initial months with hybrid model after 10-fold cross validation.

Skill lowest in initial month **January** and highest in initial month **July (cor = 0.62)**



Prediction of TC Landfall over East Asia

Step1:
large-scale parameters for TC frequency (SST, Z500, ZVWS)

Step 2: Predicting mean TC frequency (λ) by Poisson regression using key parameters
C1 C2 C3 C4 C5 C6 C7

Step 3: Random sampling of λ from the Poisson distribution for 1000 times.

$i = 1,\ \lambda_{11} + \lambda_{12} + \lambda_{13} + \lambda_{14} + \lambda_{15} + \lambda_{16} + \lambda_{17} = N_1$
$i = 2,\ \lambda_{21} + \lambda_{22} + \lambda_{23} + \lambda_{24} + \lambda_{25} + \lambda_{26} + \lambda_{27} = N_2$
…
$i = k,\ \lambda_{k1} + \lambda_{k2} + \lambda_{k3} + \lambda_{k4} + \lambda_{k5} + \lambda_{k6} + \lambda_{k7} = N_k$

Step4:
large-scale parameters for TC landfall ratio (SST, Z500,ZVWS)

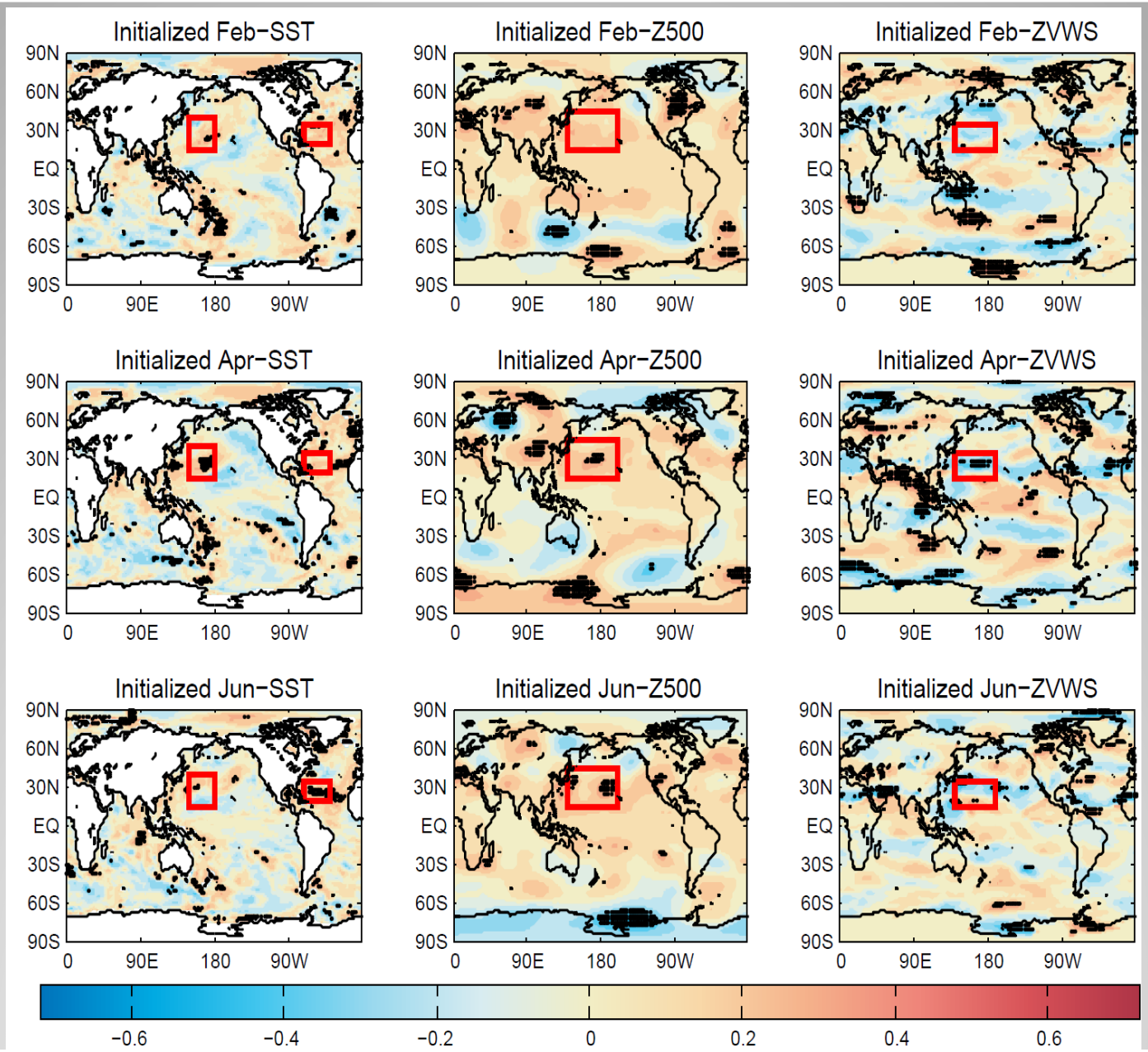
Step 5: Predicting mean TC landfall ratio (μ) by binomial regression using key parameters

Step 6: Random sampling of μ from the binomial distribution for 1000 times.

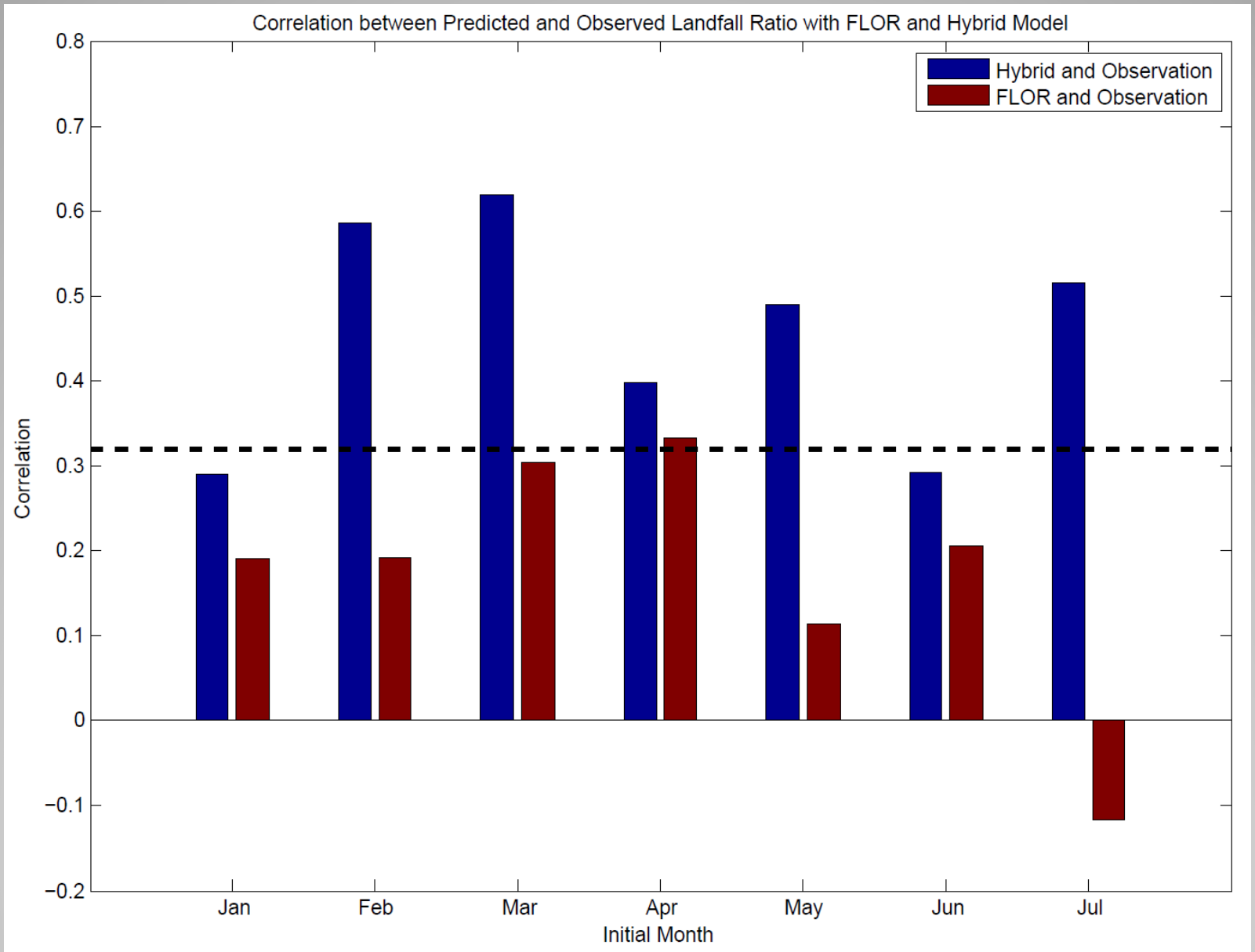
$i = 1,\ \mu_1$
$i = 2,\ \mu_2$
…
$i = k,\ \mu_k$

Step 7: Random sampling of frequency of TCs landfalling over East Asia ($N_i \times \mu_i$)

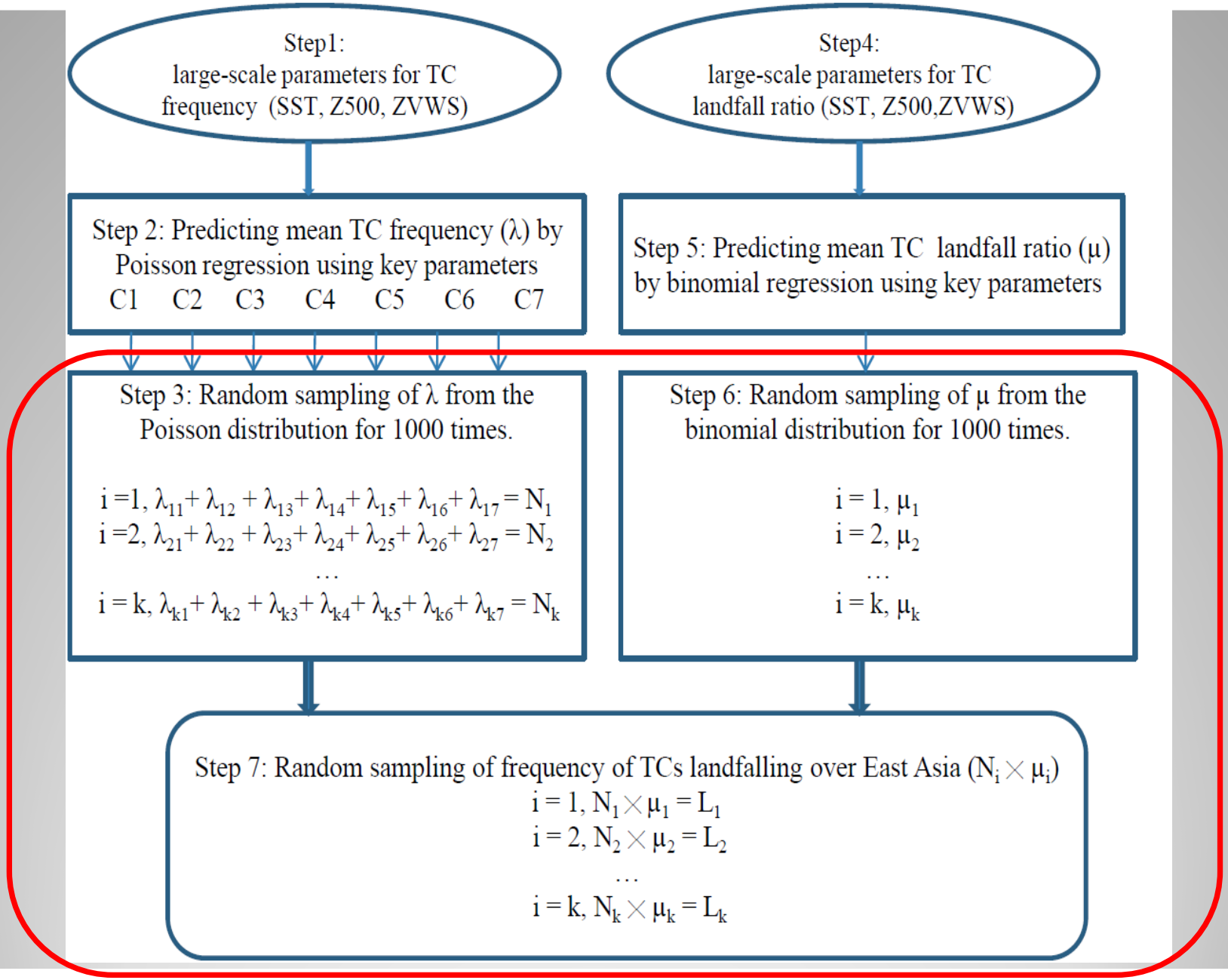
$i = 1,\ N_1 \times \mu_1 = L_1$
$i = 2,\ N_2 \times \mu_2 = L_2$
…
$i = k,\ N_k \times \mu_k = L_k$

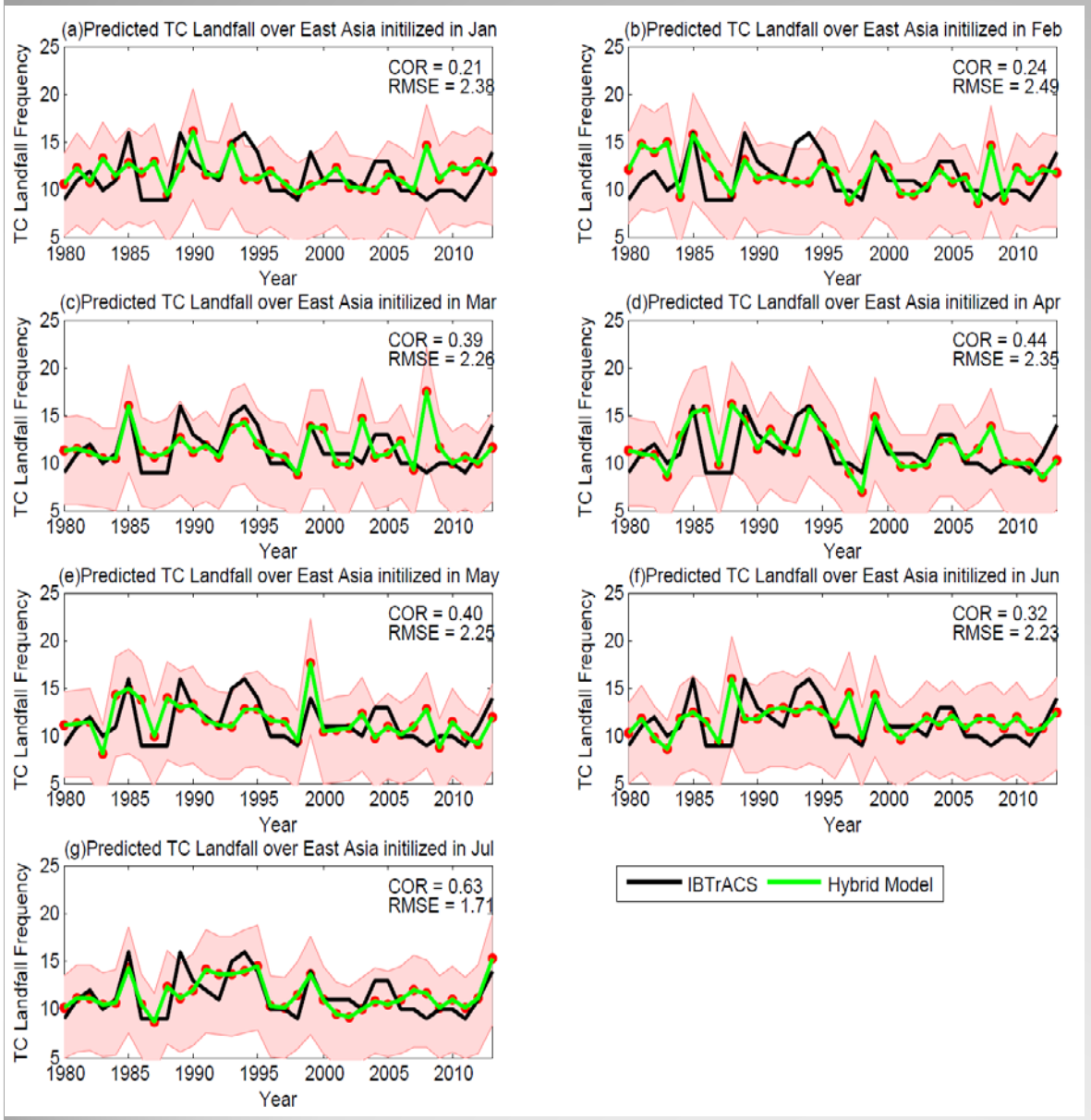


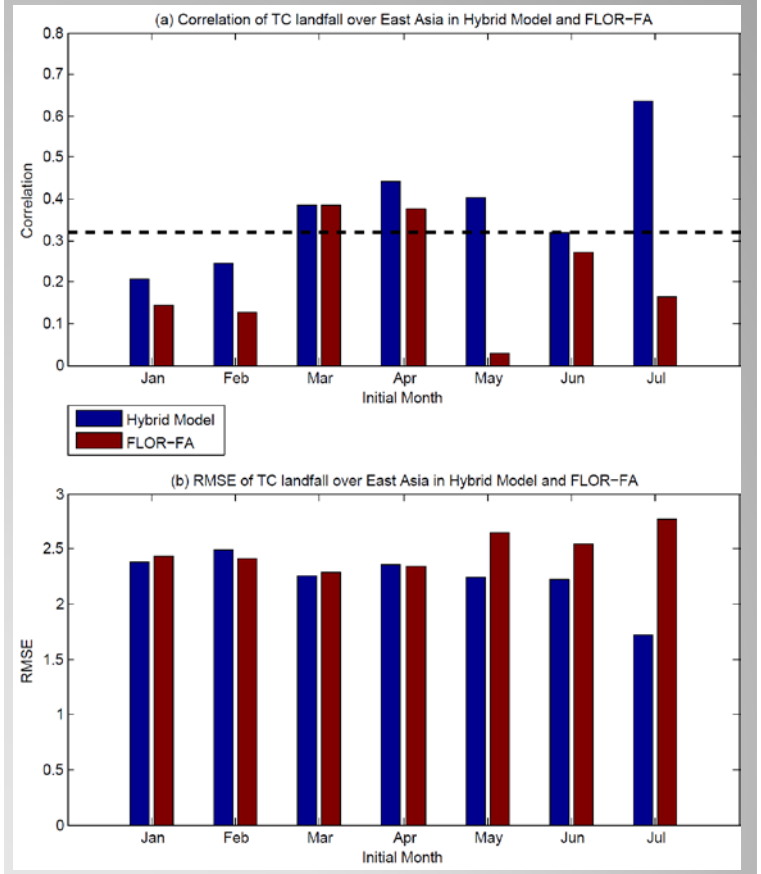
Spatial domains for the prediction of landfall ratio over East Asia



The prediction skill of landfall ratio over East Asia using FLOR-FA and hybrid model


Step1:
large-scale parameters for TC frequency (SST, Z500, ZVWS)
Step4:
large-scale parameters for TC landfall ratio (SST, Z500,ZVWS)
Step 2: Predicting mean TC frequency (λ) by Poisson regression using key parameters
C1 C2 C3 C4 C5 C6 C7
Step 5: Predicting mean TC landfall ratio (μ) by binomial regression using key parameters
Step 3: Random sampling of λ from the Poisson distribution for 1000 times.
$i=1, \lambda_{11}+\lambda_{12}+\lambda_{13}+\lambda_{14}+\lambda_{15}+\lambda_{16}+\lambda_{17}=N_1$
$i=2, \lambda_{21}+\lambda_{22}+\lambda_{23}+\lambda_{24}+\lambda_{25}+\lambda_{26}+\lambda_{27}=N_2$
…
$i=k, \lambda_{k1}+\lambda_{k2}+\lambda_{k3}+\lambda_{k4}+\lambda_{k5}+\lambda_{k6}+\lambda_{k7}=N_k$
Step 6: Random sampling of μ from the binomial distribution for 1000 times.
$i=1, \mu_1$
$i=2, \mu_2$
…
$i=k, \mu_k$
Step 7: Random sampling of frequency of TCs landfalling over East Asia ($N_i \times \mu_i$)
$i=1, N_1 \times \mu_1 = L_1$
$i=2, N_2 \times \mu_2 = L_2$
…
$i=k, N_k \times \mu_k = L_k$





The prediction skill for landfall frequency over East Asia using FLOR-FA and hybrid model

Conclusion

- Using the FMM-based cluster analysis, seven clusters are identified from the historical and FLOR-FA-predicted TC tracks for the period 1980-2013. By building Poisson regression models for each cluster using key predictors, the predictive skill for almost all the clusters at all initial months improves with respect to the dynamic prediction.
- The prediction of total WNP TC frequency made by combining hybrid predictions for each of the seven clusters in the hybrid model shows skill higher than what achieved using the TC frequency directly from FLOR-FA initialized from April to July. However, the hybrid predictions initialized from January to March exhibit lower skill than FLOR-FA.
- The prediction of TC landfall over East Asia made by combining the two hybrid models for total WNP TC frequency and TC landfall ratio also outperforms FLOR-FA for all start months January through July.