GroupMeeting

2019.11.04

张慕琪

Papers

1. A hierarchical analysis of the impact of methodological decisions on statistical downscaling of daily precipitation and air temperatures. Sara C. Pryor and Justin T. Schoof. International Journal of Climatology. 25 January 2019.

2. Deep Trouble for Deep Learning. Douglas Heaven. Nature. 10 October 2019

3. Statistical downscaling of precipitation using machine learning techniques. D.A. Sachindra, K. Ahmed, Md. Mamunur Rashid, S. Shahid, B.J.C. Perera. Atmospheric Research. 10 May 2018.

Received: 30 July 2018 | Revised: 2 January 2019 | Accepted: 4 January 2019 | Published on: 25 January 2019
DOI: 10.1002/joc.5990

International Journal of Climatology RMetS

RESEARCH ARTICLE

A hierarchical analysis of the impact of methodological decisions on statistical downscaling of daily precipitation and air temperatures

Sara C. Pryor[1] | Justin T. Schoof[2]

本文既有对于温度设计的模型又有针对降水设计的模型，在此更多的关注降水部分

Introduction

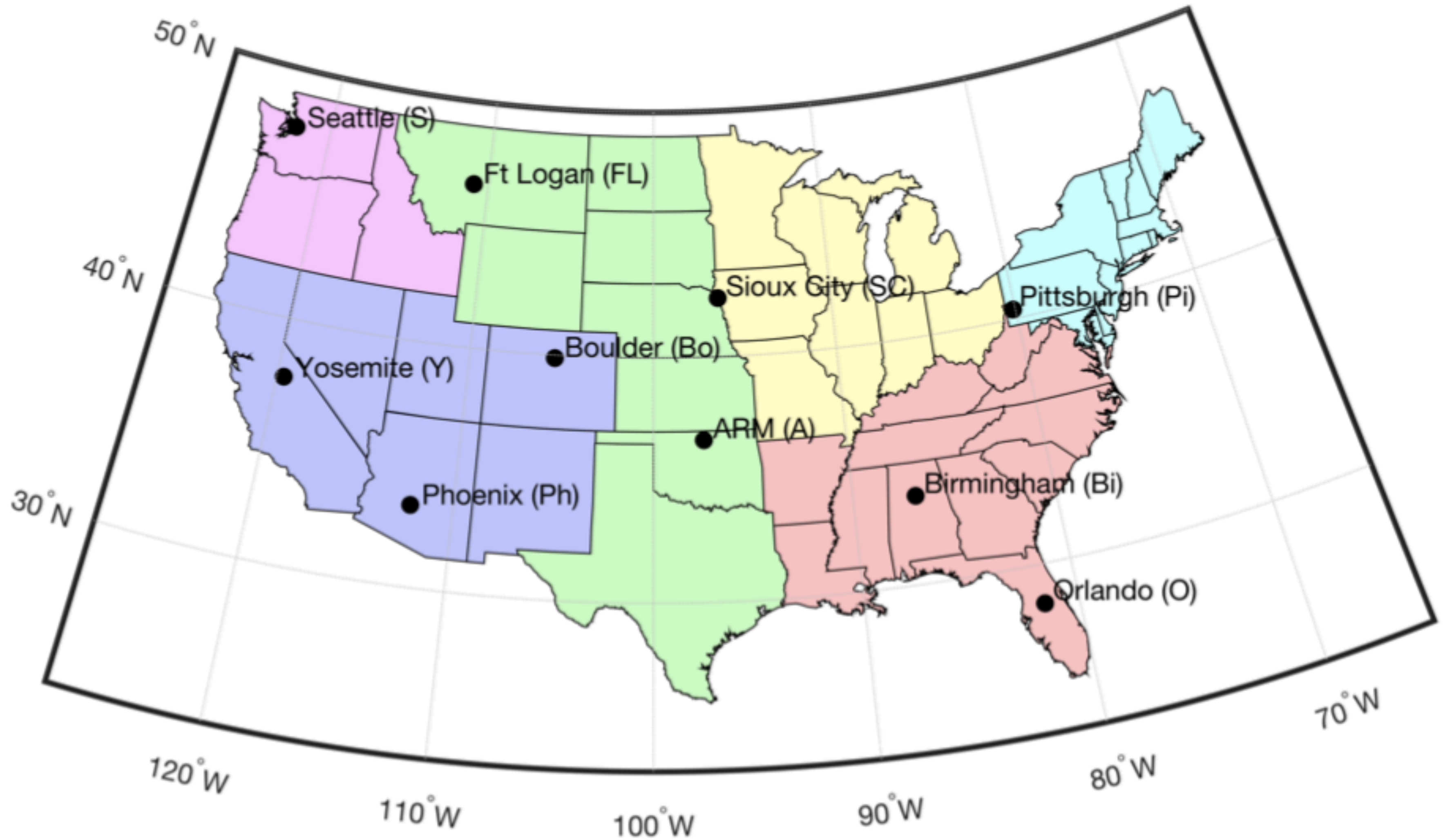
研究变量：Tmin, Tmax, PoP(Probability of precipitaion), Amount of precipitation on a wet day.

涉及模型：Generalized linear models[GLMs] and Artificial neural networks[ANNs].

Introduction

研究区域：涵盖美国不同气候区域的10个地区

数据来源：预测值(Tmin, Tmax, precipitation occurrence and amount on a wet day)来自Livneh data set(Livneh et al., 2013).



预测数据： Geopotential height at 500 hPa (Z500)， Air temperature at 700 hPa (T700)
Specific humidity at 700 hPa (Q700)， Air temperature at 500 hPa (T500)，
Specific humidity at 500 hPa (Q500)，
West–East (u-component) wind speed at 700 hPa (U700)，
South–North (v-component) wind speed at 700 hPa (V700)

Temperature Downscaling

Model 1: 使用GLM和三个输入（Z500 T700 Q700）;
Model 2: 使用ANN和三个输入;
Model 3: 使用GLM和七个输入;
Model 4: 使用ANN和七个输入;
Model 5 and 6: GLM和ANN使用PC scores。

Temperature Downscaling

Input网格并非越多越好

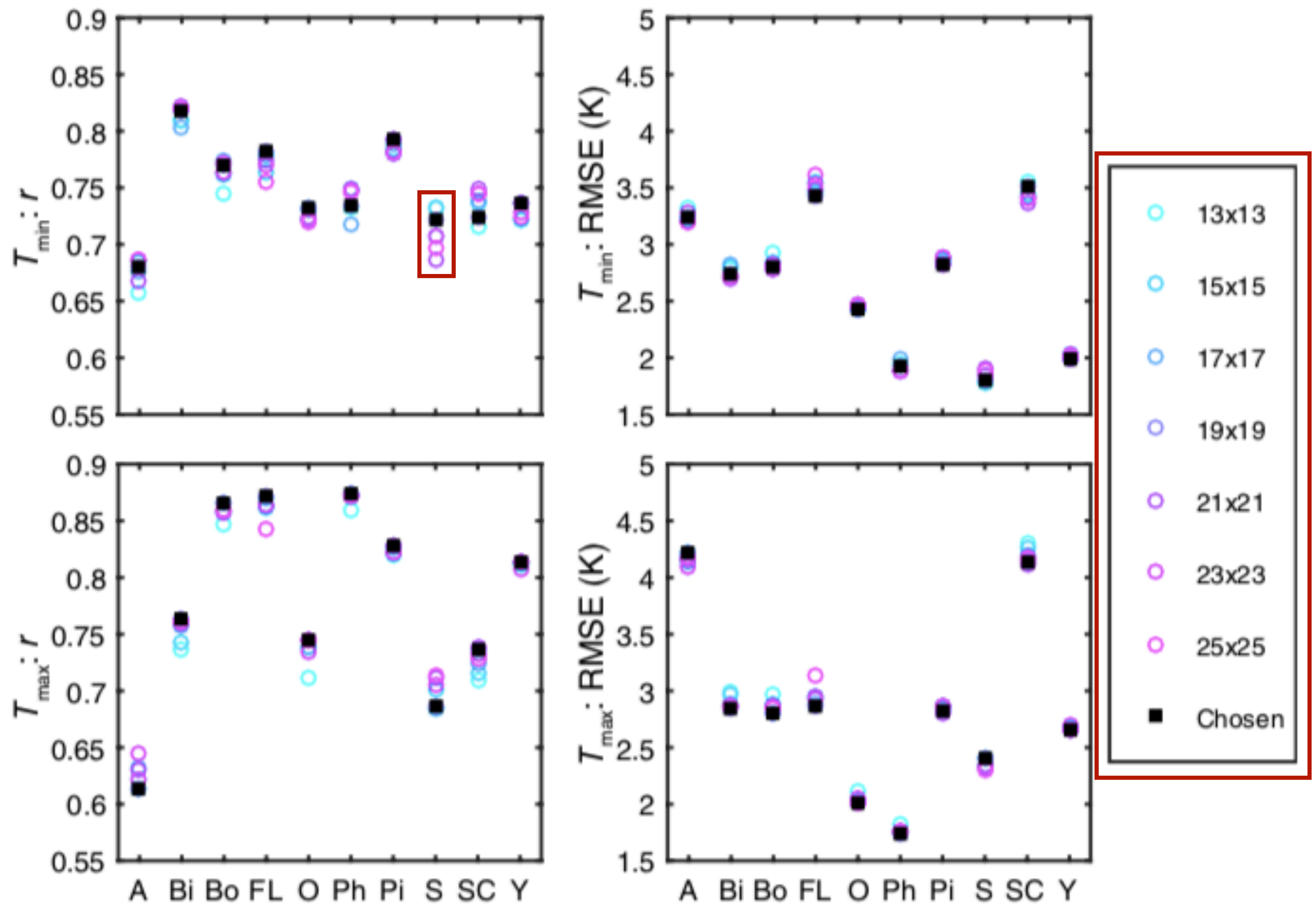


FIGURE 2 Pearson correlation coefficient (r) and RMSE between independent observed daily $T_{\min}$ and $T_{\max}$ anomalies and downscaled predictions for each of the 10 locations. The ESD transfer functions are built using ANN (using three hidden layers) where the predictors are PC scores from different domain sizes (where the number of grid cells used in the spatial domain presented to the PCA is shown in the legend). Results from a domain of 19 × 19 grid cells that is used within this manuscript are highlighted by the black squares. The locations are referred to using the abbreviations introduced in Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]

Results

1. Phoenix和Yosemite的偏差在±2C内；
2. ARM的偏差在±5C内；
3. 无论是Tmin还是Tmax变量，GLM有更多的输入（L7和LP）会比仅有三个输入（L3）时表现更佳；
4. 使用ANN比GLM的偏差更小。

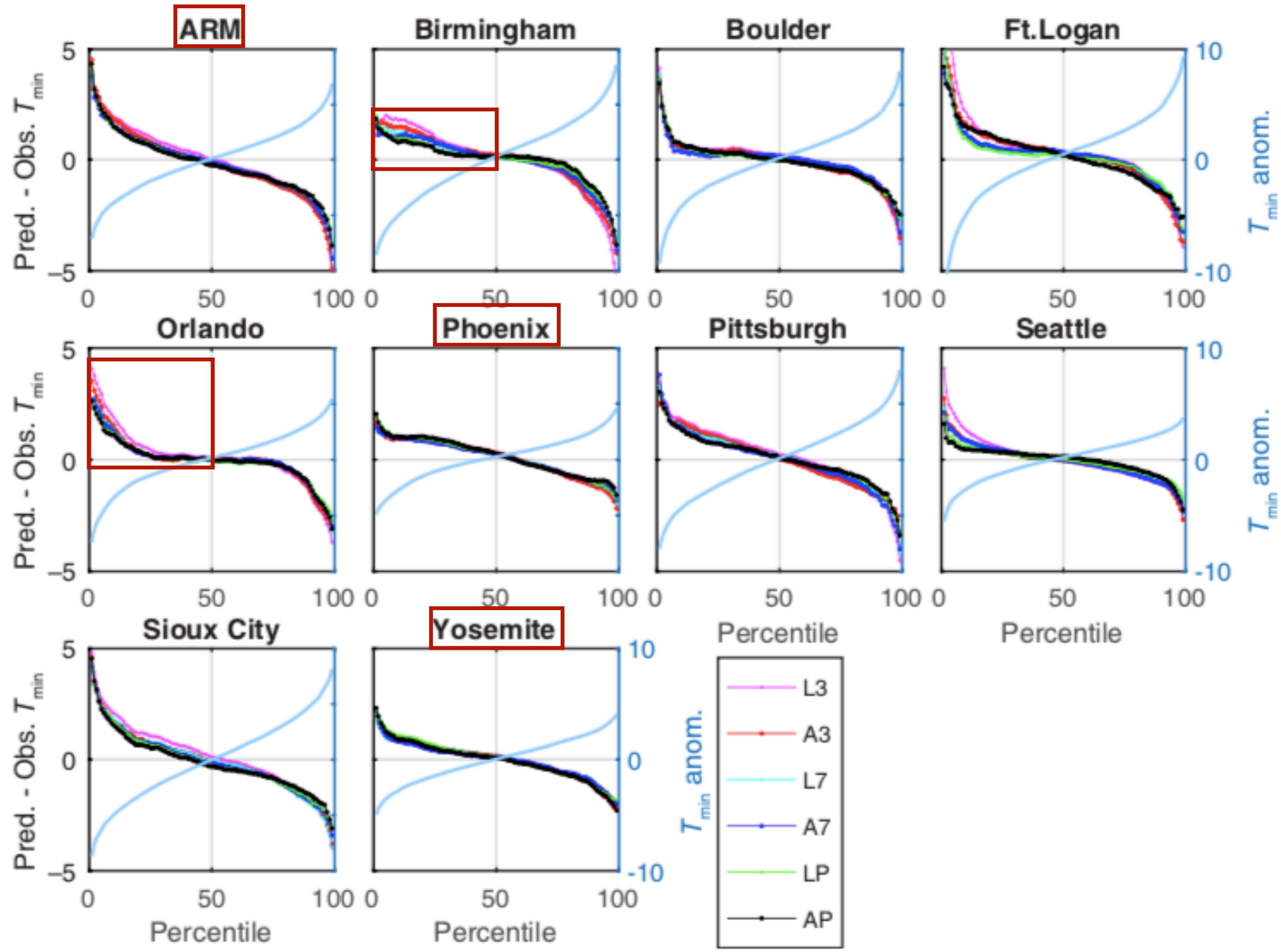


FIGURE 3 Difference in the 1st to 99th percentile of daily anomalies of minimum temperature (T_{min}) from climatology in observations from the independent data set (even years) and six different transfer functions. The solid light-blue lines denote the 1st to 99th percentile observed daily T_{min} anomalies (the scale is shown on the right-hand axes). The abbreviations used to identify the transfer functions are as follows; L3 denotes a transfer function built using multiple linear regression with three grid cell predictors (Z500, T700 and Q700), A3 indicates transfer functions built using ANNs with three grid cell predictors (Z500, T700 and Q700), L7 is for a transfer function built using the seven grid cell predictors (Z500, T700, Q700, T500, Q500, U700 and V700) but allowing for first-order interactions of the predictors and stepwise procedure for selecting the predictors (based on BIC) and the regularized multiple linear regression. A7 is for a transfer function built using ANNs with seven grid cell predictors (Z500, T700, Q700, T500, Q500, U700 and V700). LP is for a transfer function built using the PC scores, allowing for first-order interactions of the predictors and stepwise procedure for selecting the predictors (based on BIC) and the regularized multiple linear regression. AP is for a transfer function built using ANNs and PC scores [Colour figure can be viewed at wileyonlinelibrary.com]

Results

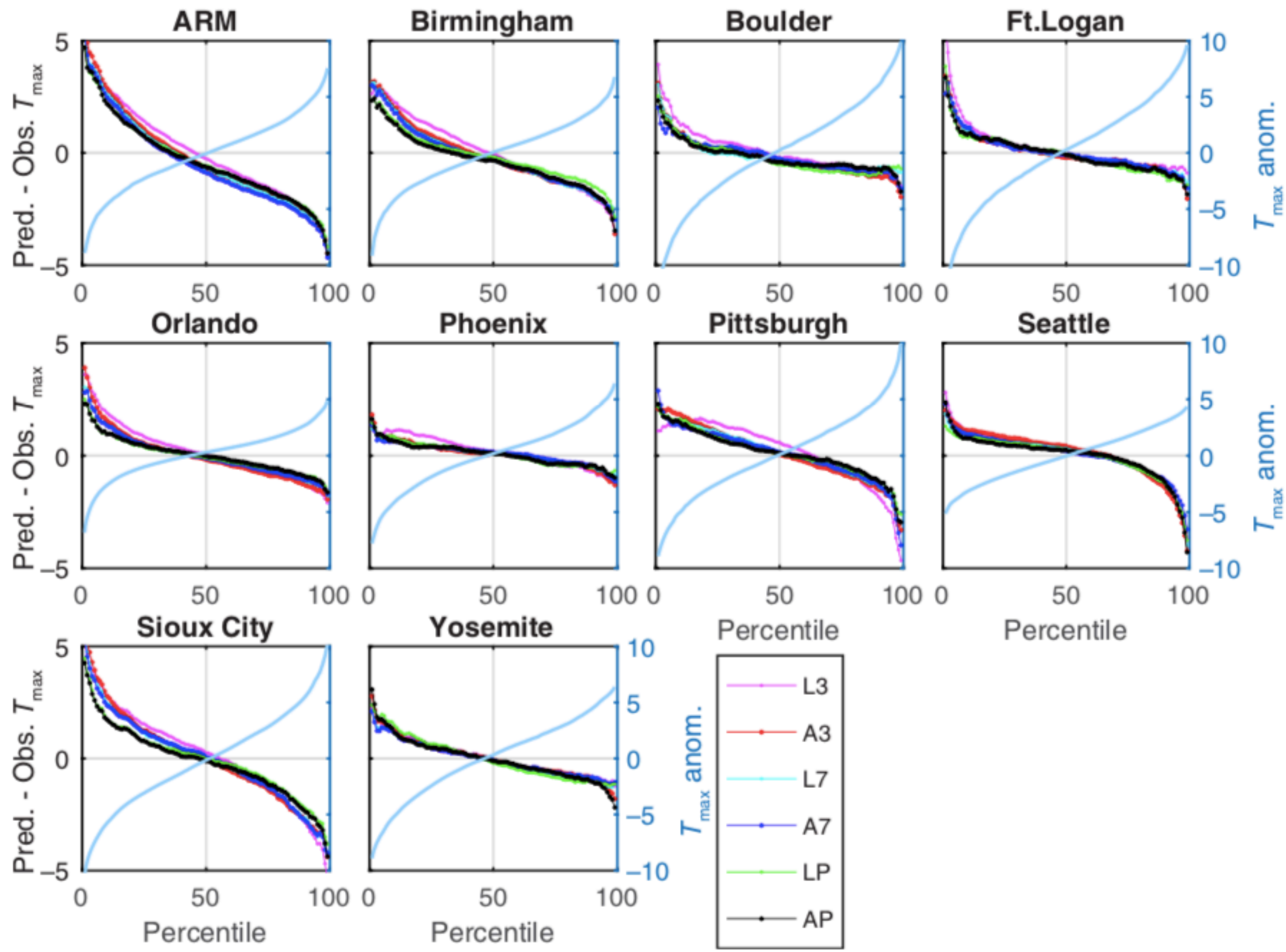


FIGURE 4 As Figure 3 but for daily anomalies of maximum temperature (T_{max}) from climatology in observations from the independent data set (even years) and six different transfer functions. The solid light-blue lines denote the 1st to 99th percentile observed daily T_{max} anomalies (shown on the right-hand axes) [Colour figure can be viewed at wileyonlinelibrary.com]

Results

Tmin

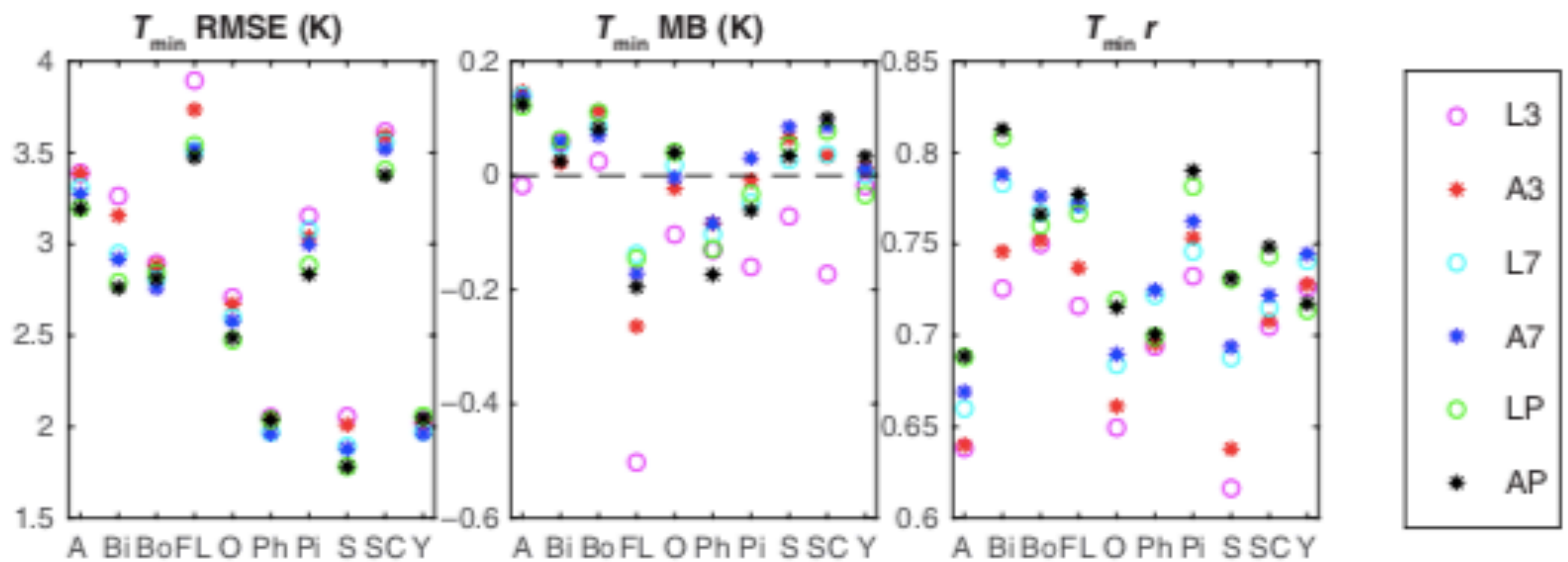


FIGURE 5 Skill assessment of the different ESD models as applied to T_{min}. RMSE, r and MB are computed for the daily deviations from climatology. Bias in the number of frost days and tropical nights and *OR* for $T_{min} < 0°C$ is computed once the climatology has been added to the predicted anomalies. A value of −0.1 for tropical nights indicates the downscaled values underestimate the number of tropical nights by 10% relative to observations. *OR* are computed from the hit rate and false alarm rate and thus test time synchronicity. The abbreviations used to identify the ESD models are as in Figures 3 and 4 [Colour figure can be viewed at wileyonlinelibrary.com]

Tmax

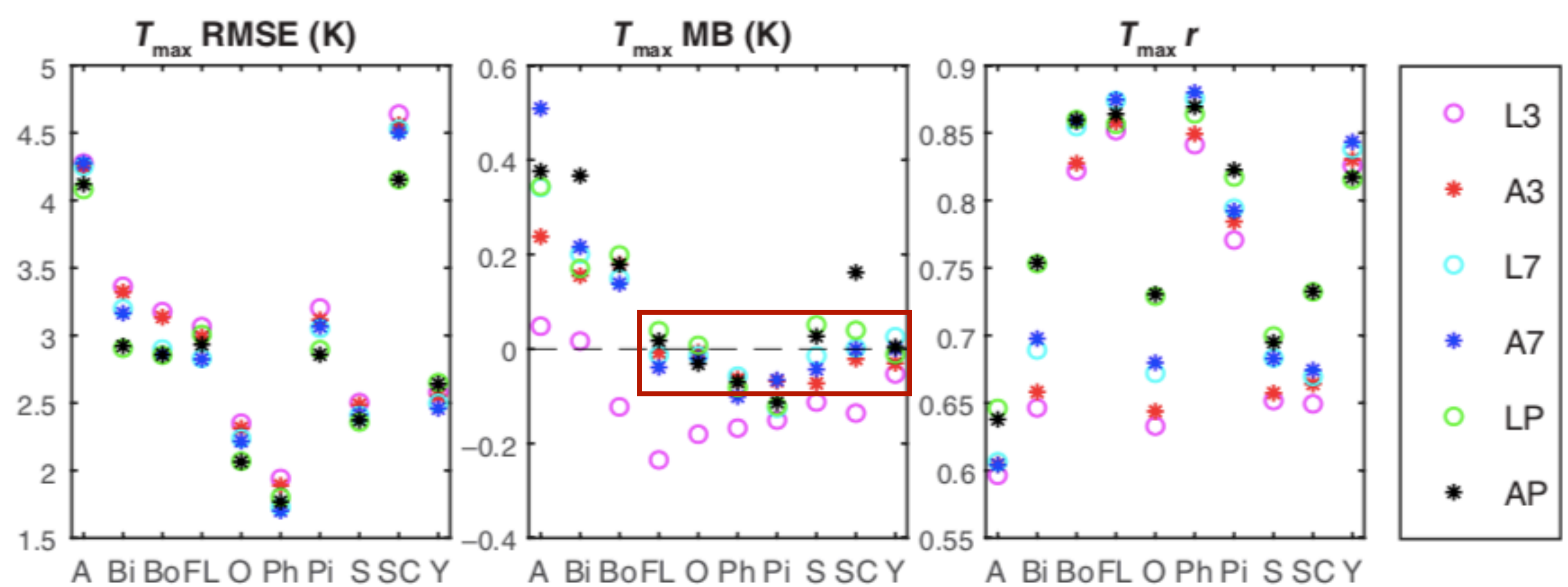


FIGURE 6 As in Figure 5. Skill assessment of the different ESD methods as applied to T_{max}. RMSE, r and MB are computed for the daily deviations from climatology. The bias in the number of icing days and summer days in the 13 years and the *OR* for the occurrence of $T_{max} > 32.2°C$ are computed once the climatology has been added [Colour figure can be viewed at wileyonlinelibrary.com]

Precipitation Downscaling

RCMs的数据总是出现“经常下雨”的情况，且由于方法的系统偏差RCMs会低估降水的强度；

研究发现ANN对于预报PoP表现较差；

研究发现，方法的偏差是由潮湿天气定义的阈值决定的，因此本文设定三种阈值：>0 mm/day, >0.1 mm/day and >1 mm/day;

模型设计：三种阈值

输入变量3个/7个/使用PCA方法确定具体输入几个

GLM使用Poisson或gamma分布 VS. ANN

数据集划分采用奇偶年的方法

P.S. ANN有三层隐藏层

Results

OR = H/(1-H)
H = a/(a+c)
a: 事件发生且预报正确的个数；
c: 事件发生但未预报
H: 成功率

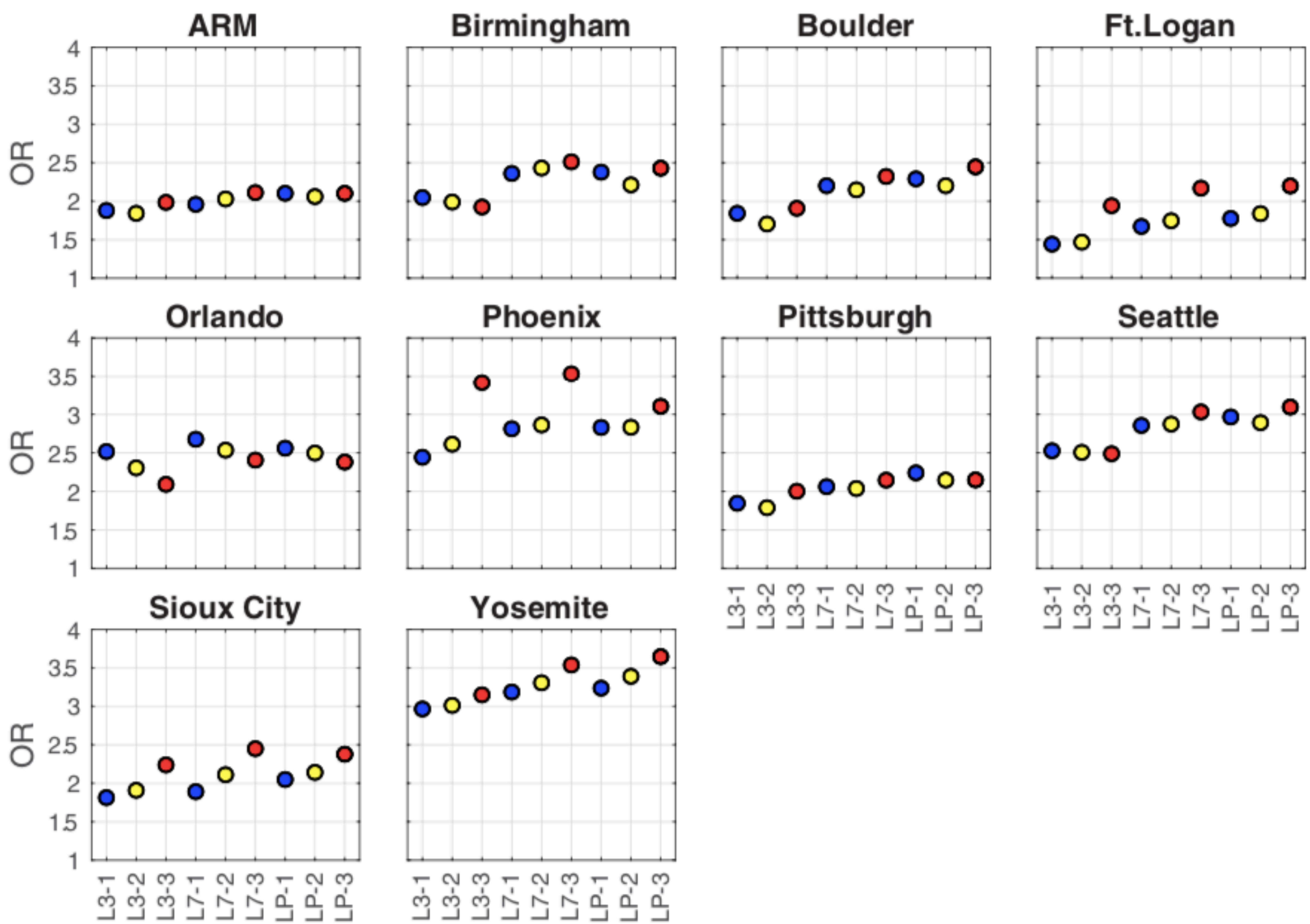


FIGURE 8 The ORs for the PoP at the 10 stations for the nine ESD models applied to the independent test data. Model abbreviations are as follows: L3 denotes logistic regression using grid cell values of Z500, T700 and Q700. L7 denotes stepwise logistic regression with L1 regularization wherein the predictors are Z500, T700, Q700, T500, Q500, U700, V700. LP indicates stepwise logistic regression with L1 regularization with PC scores as predictors. −1 indicates a wet-day threshold of >0 mm/day (red), −2 indicates a threshold of 0.1 mm/day (yellow) and 3 is used for a threshold of >1 mm/day (red) [Colour figure can be viewed at wileyonlinelibrary.com]

补充

ROC curve

Receiver operating characteristic curve

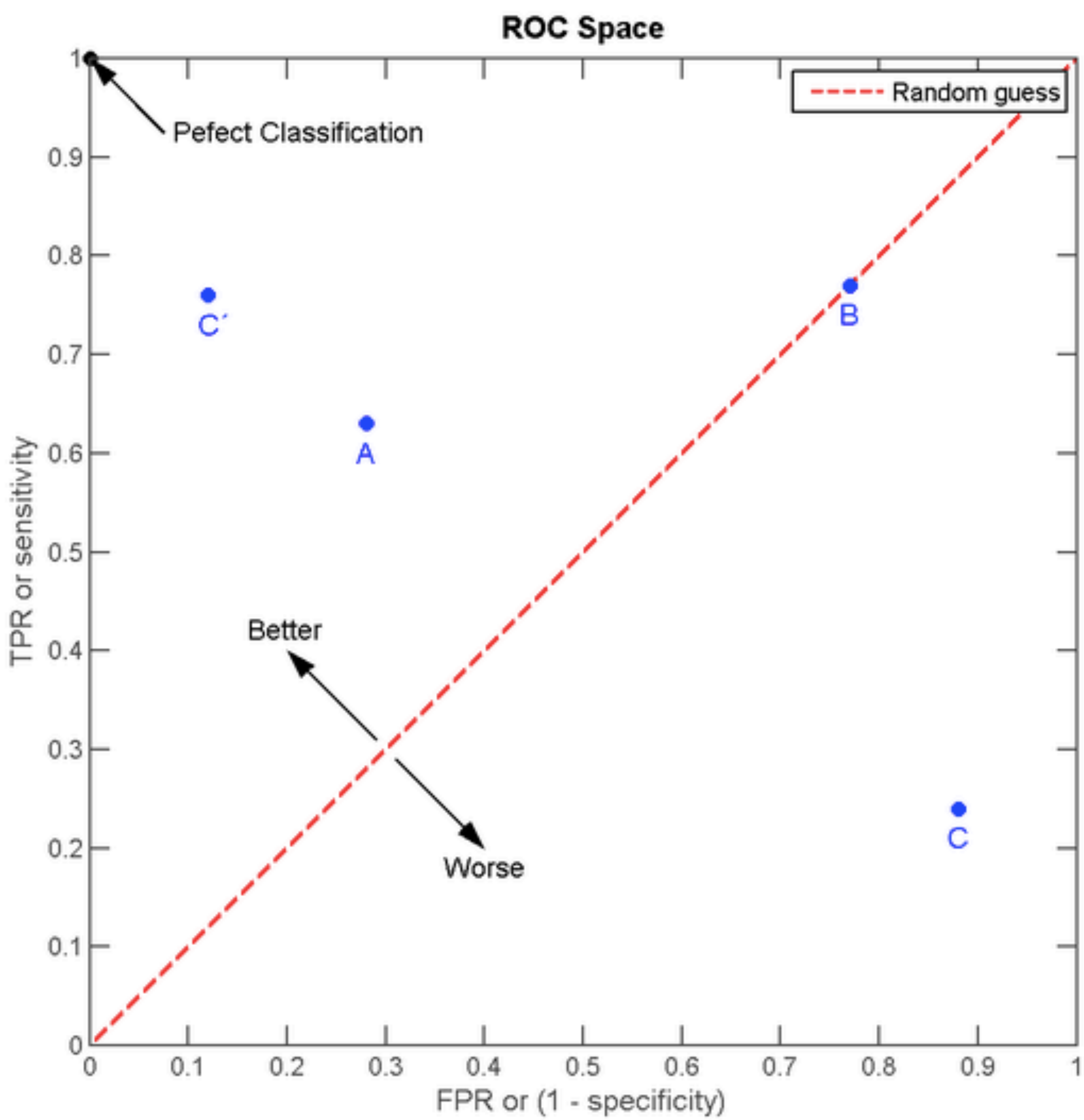
Let us look into four prediction results from 100 positive and 100 negative instances:

A			B			C			C′		
TP=63	FP=28	91	TP=77	FP=77	154	TP=24	FP=88	112	TP=76	FP=12	88
FN=37	TN=72	109	FN=23	TN=23	46	FN=76	TN=12	88	FN=24	TN=88	112
100	100	200	100	100	200	100	100	200	100	100	200
TPR = 0.63			TPR = 0.77			TPR = 0.24			TPR = 0.76		
FPR = 0.28			FPR = 0.77			FPR = 0.88			FPR = 0.12		
PPV = 0.69			PPV = 0.50			PPV = 0.21			PPV = 0.86		
F1 = 0.66			F1 = 0.61			F1 = 0.23			F1 = 0.81		
ACC = 0.68			ACC = 0.50			ACC = 0.18			ACC = 0.82		

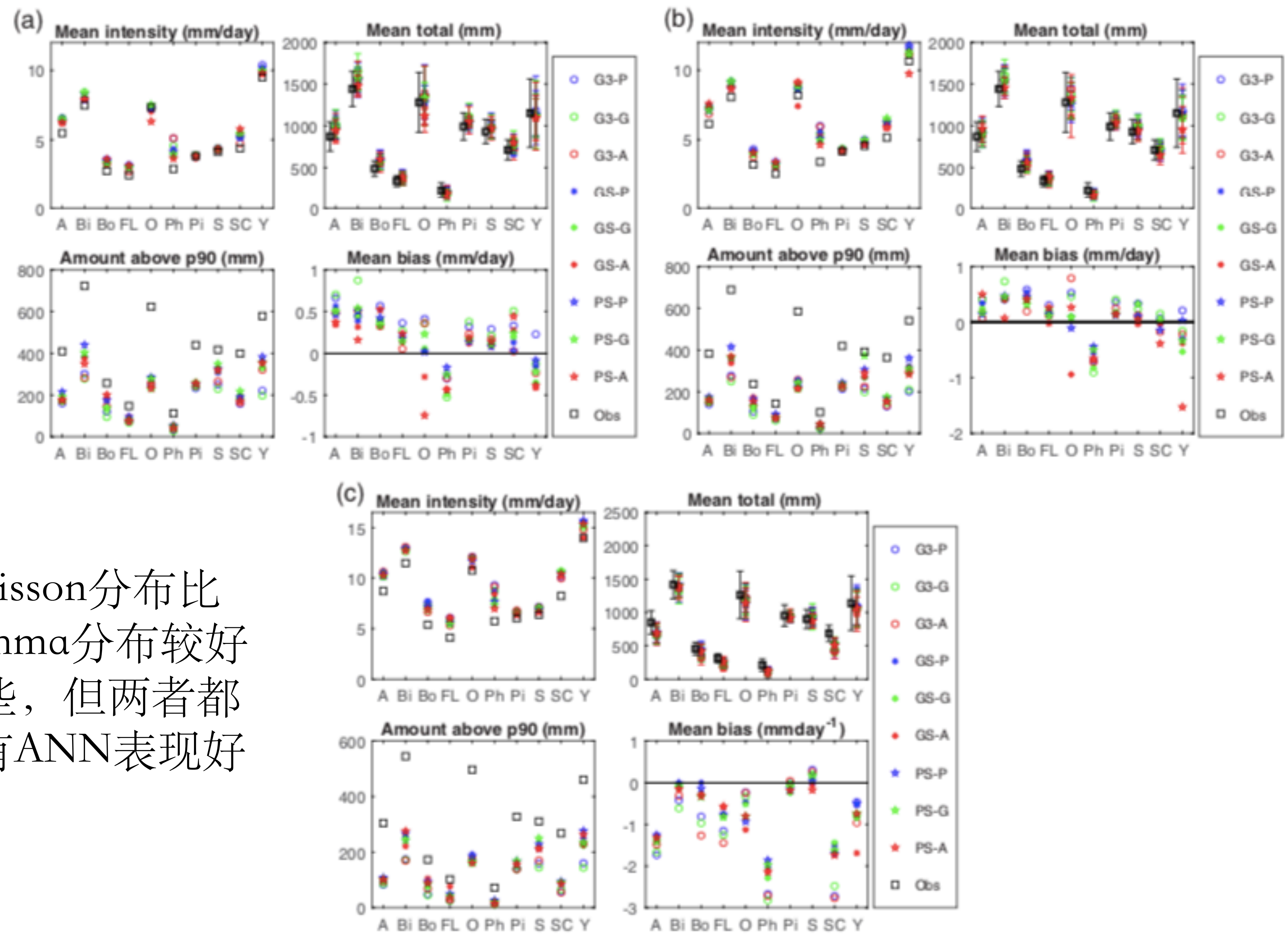
补充

ROC curve

Receiver operating characteristic curve



Results



Poisson分布比Gamma分布较好一些，但两者都没有ANN表现好

FIGURE 11 Results for the 10 stations of four aspects of the precipitation climate; mean intensity on a wet day, MB on a wet day, mean annual total (and interannual variability) and the amount (in mm) of precipitation above the 90th percentile value. Note in the panel showing annual total precipitation and interannual variability therein, the observed values are slightly displaced on the horizontal axis to aid legibility. Frame (a) shows results for a precipitation threshold of 0 mm/day, (b) for 0.1 mm/day and (c) for 1 mm/day for a wet day. Nine transfer functions (ESD models) are shown the first two letters denote the predictors; G3 = 3 predictors are used (grid cell values of Z500, T700 and Q700), GS = 7 gird cell predictors, and PS denotes models built using the PC scores as predictors. The final letter denotes the transfer function form; P is Poisson for the link function in the logistic regression, G is for gamma for the link function in the logistic regression, and A is ANN [Colour figure can be viewed at wileyonlinelibrary.com]

Conclusions

1. 无论哪种回归模型对于PoP的预测总是不尽人意;
2. 本文使用的数据划分是奇偶年特征，有人曾提出如果使用干旱年进行训练，非干旱年进行测试的新方法，但这种方法更多的是侧重于捕捉气候的内部变化而非辐射平衡导致的气候变化;
3. 非线性模型（ANN）要比GLM更具有技巧性，表现更好。

ILLUSTRATION BY EDGAR BĄK

DEEP TROUBLE FOR DEEP LEARNING

BY DOUGLAS HEAVEN

Introduction

DNNs do not actually understand the world.

Introduction

2013年，谷歌研究员Christian Szegedy首次提出一个新概念——“对抗样本”。在他本人的印本“神经网络的有趣的可能”中，他提出他们小组发现DNN可以成功认出狮子的图像，但通过更改个别像素数据，DNN则认为自己在观看完全不同的一幅图像，例如一个图书馆。Clune也通过实验得到了类似的结论，他认为这种错误在人类大脑中是完全不可能想象的。

这也意味着黑客们有各种不同的方法来攻击一个系统，而当攻击开始后工作者往往很难解决这类问题。

Why Great Power Comes Great Fragility?

目前，大家公认的解决办法是向AI输入更多的数据，然而这样的训练往往是一个漫长的过程；训练仅仅一个模型所需要的数据有可能需要花费几年的时间来完成，并且数据并不完全是可靠的，而硬件中传感器的校正可能会随着时间变化，硬件设施的性能也会随着时间降低。

针对使用较少数据进行学习的方法，人们提出了“转移学习”对训练方法。即使用几个甚至一个例子即可训练出一个新的网络。这个想法建立在已有一个提前训练好的DNN的前提之上，例如，有一个DNN已经见过犯罪数据中几百万的面部图片并习得了一些有用的信息，现在向其提供一张新的图片，DNN可以快速找到数据集中与之最相近的一个图像。

Learning From Less Data

然而，即使是目前最成功的AI系统例如AlphaZero也只能在很狭窄的领域中获得成功，AlphaZero的算法仅针对国际象棋和GO，但两种竞技游戏并不是同时训练的，同时训练将由于各自的干扰而降低胜率。然而从人类的角度出发，会觉得这是件很荒诞的事情，因为人类是不会轻易忘记曾经学过的知识并且人类是可以学以致用。

DNNs don’t have a good model of how to pick out what matters.

科学家希望理想的DNNs输出应该是不会收到图像细节改变的干扰的， 而这件事目前没有人可以优化。

Atmospheric Research 212 (2018) 240–258

ELSEVIER

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Statistical downscaling of precipitation using machine learning techniques

D.A. Sachindra[a,*], K. Ahmed[b], Md. Mamunur Rashid[c], S. Shahid[d], B.J.C. Perera[a]

[a] Institute for Sustainability and Innovation, College of Engineering and Science, Victoria University, P.O. Box 14428, Melbourne, Victoria 8001, Australia
[b] Faculty of Water Resources Management, Water and Marine Sciences, Lasbela University of Agriculture, Uthal, Balochistan, Pakistan
[c] Civil, Environmental, and Construction Engineering Department, University of Central Florida, Orlando, Florida 32816-2450, USA
[d] Faculty of Civil Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia

Introduction

机器学习方法

MLR、GLMs、

ANNs、SVM、RVM、GP、GEP……

计算成本低、简便

降尺度

统计降尺度

建立GCM和所需尺度内水文气候变量间的经验统计关系

基于回归的方法

基于天气发生器的方法

基于天气分类器的方法

动力降尺度　基于物理方程以达到同样目的

·相较于MLR，基于GP的降尺度模型可以更好的模拟日最高温和日最低温；

·相较于ANN和MLR，基于SVM的降尺度模型可以更好的模拟月最高温和最低温；

·相对于其他所有方法，基于SVM的降尺度模型对降水的月平均模拟最好；

——根据上述研究，在众多传统的统计回归方法中，依据机器学习方法建立的降尺度模型比其他方法更加有优势。

Introduction

1. GP（Genetic programming）

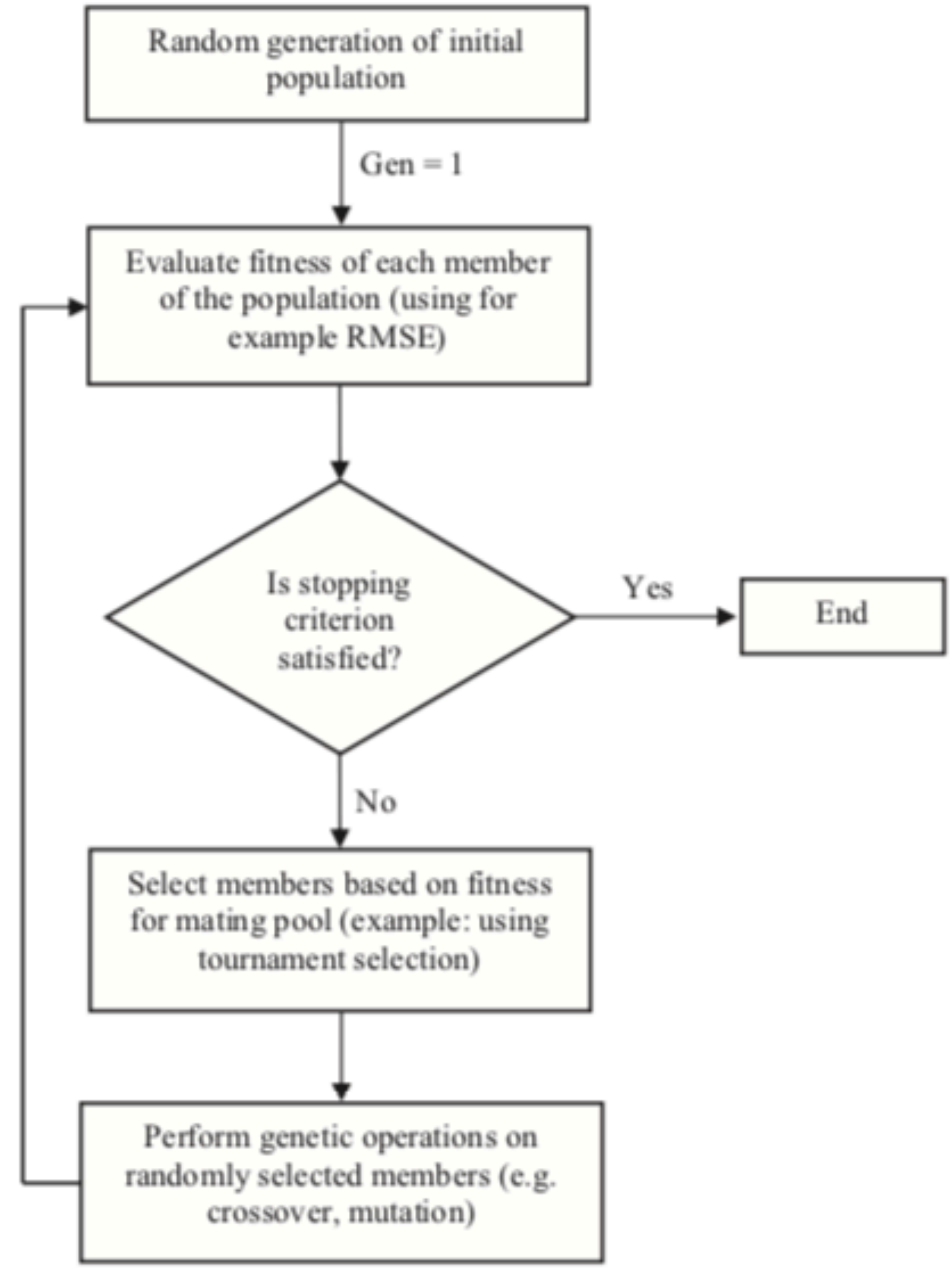


Fig. 3. Genetic programming algorithm.

2. Artificial neural networks（ANNs）

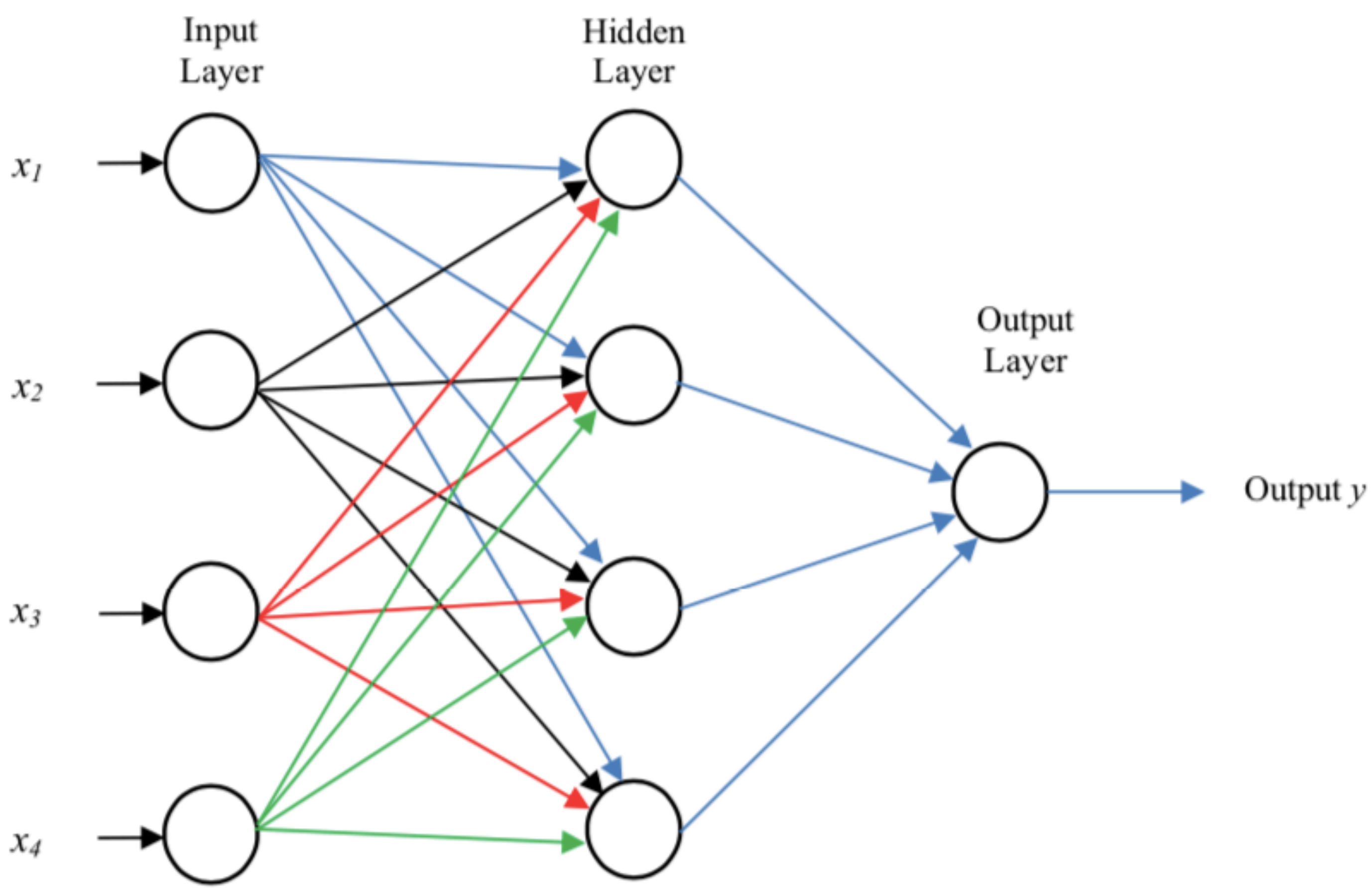


Fig. 4. A typical artificial neural network.

3. Support vector machine（SVM）
4. Relevance vector machine（RVM）

Method

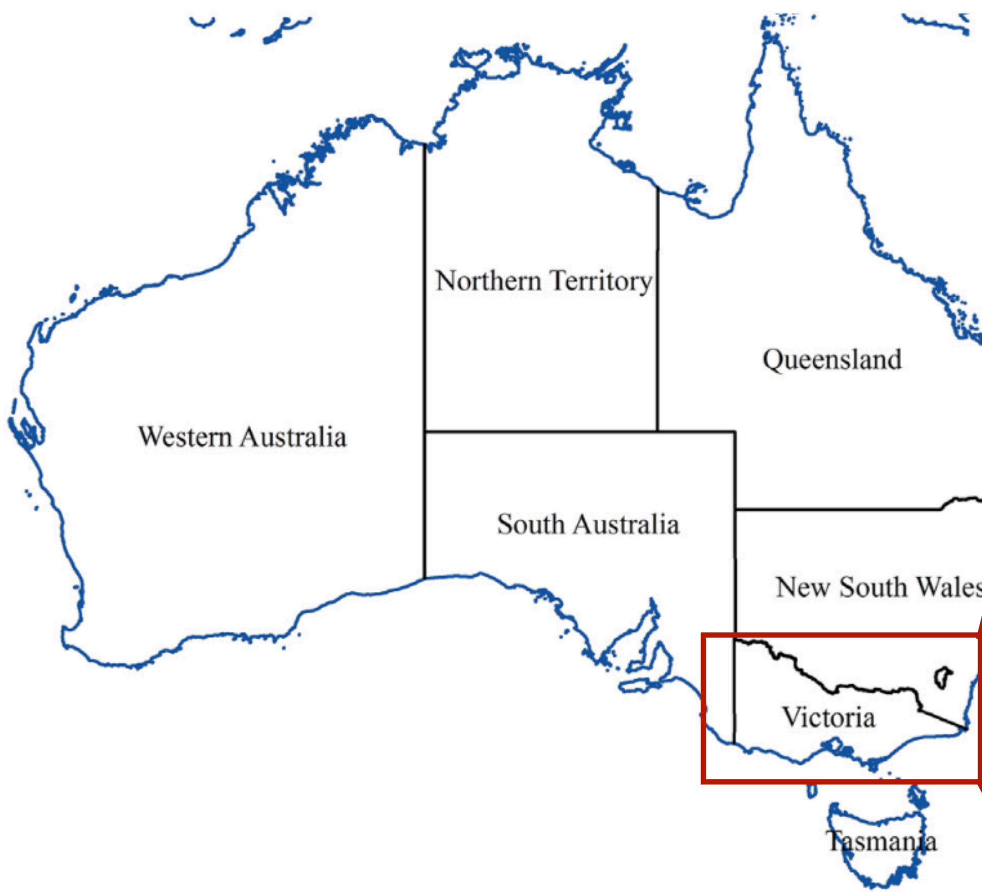


Fig. 1. Location of Victoria in Australia.

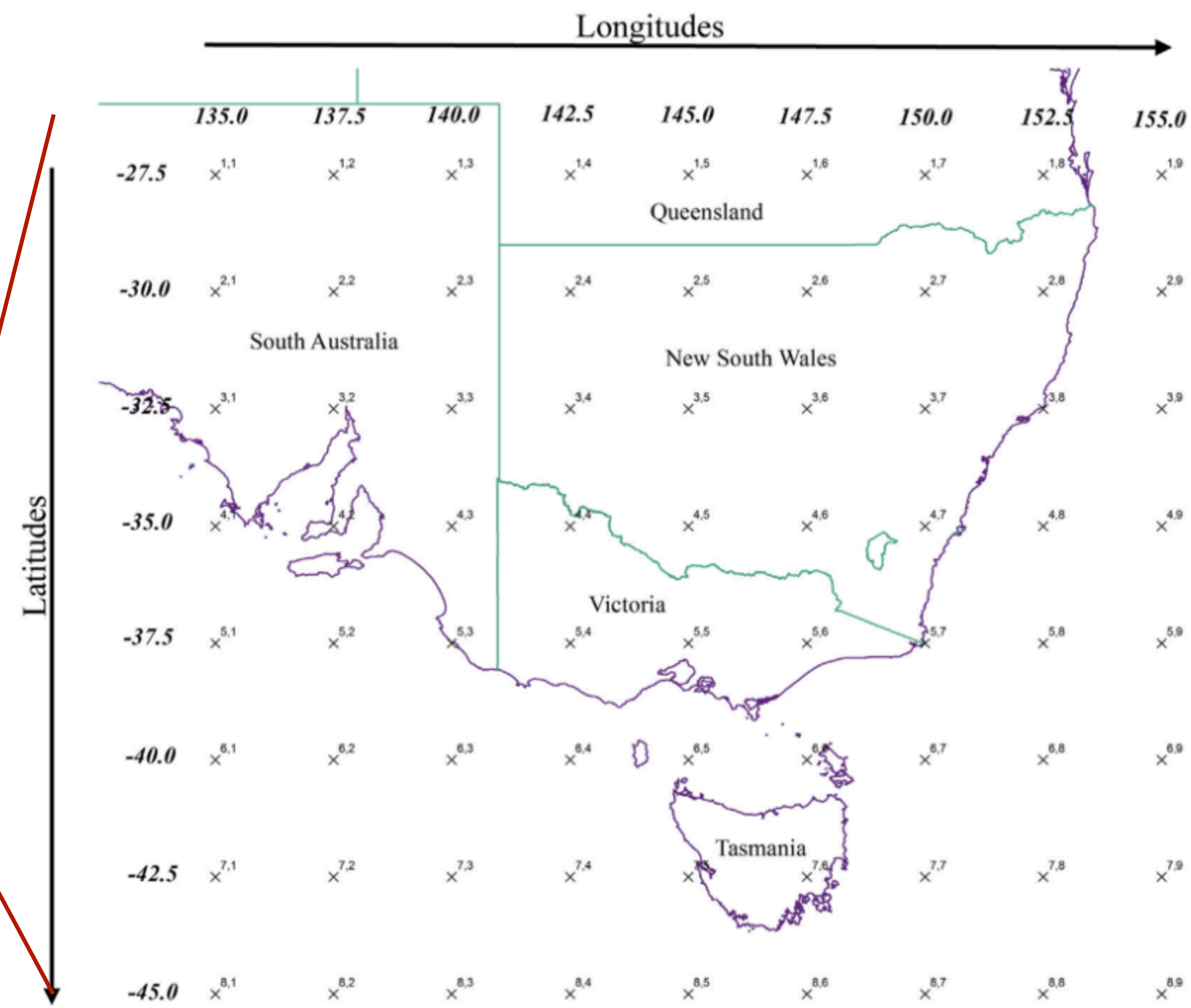


Fig. 5. Atmospheric domain used in this study.

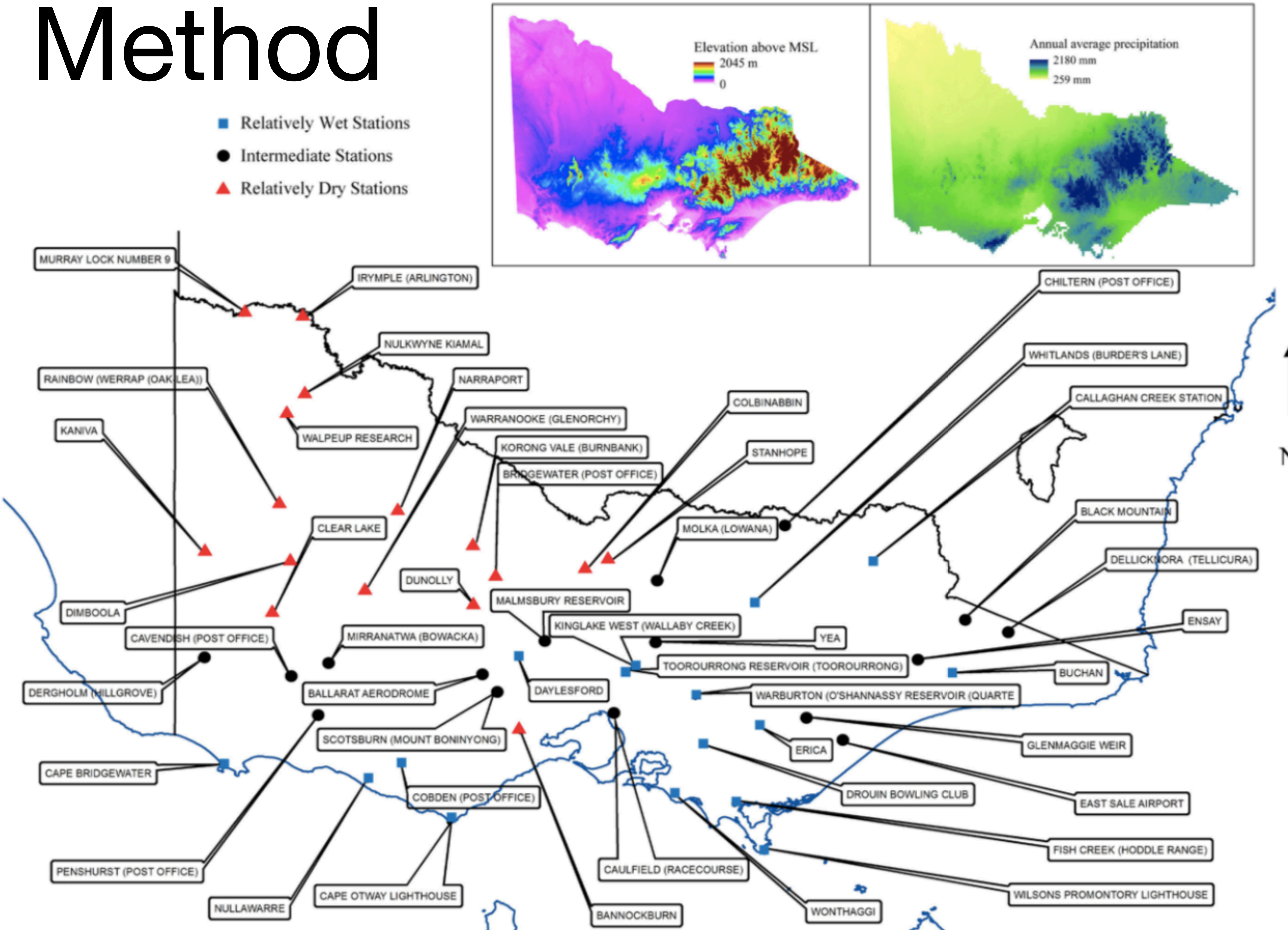
NCEP/NCAR reanalysis data set；

48 precipitation observation stations located across Victoria (237,000km^2), Australia were selected；

时间：1950-1991(calibration)、1992-2014(validation)

Train, Validation, Test

Method



Relatively Wet Stations
Intermediate Stations
Relatively Dry Stations
Elevation above MSL
2045 m
0
Annual average precipitation
2180 mm
259 mm
N
MURRAY LOCK NUMBER 9
IRYMPLE (ARLINGTON)
NULKWYNE KIAMAL
RAINBOW (WERRAP (OAKLEA))
NARRAPORT
KANIVA
WALPEUP RESEARCH
WARRANOOKE (GLENORCHY)
KORONG VALE (BURNBANK)
BRIDGEWATER (POST OFFICE)
COLBINABBIN
STANHOPE
CHILTERN (POST OFFICE)
WHITLANDS (BURDER'S LANE)
CALLAGHAN CREEK STATION
CLEAR LAKE
MOLKA (LOWANA)
BLACK MOUNTAIN
DELLICKNORA (TELLICURA)
DUNOLLY
DIMBOOLA
MALMSBURY RESERVOIR
KINGLAKE WEST (WALLABY CREEK)
ENSAY
CAVENDISH (POST OFFICE)
MIRRANATWA (BOWACKA)
YEA
TOOROURRONG RESERVOIR (TOOROURRONG)
BUCHAN
DERGHOLM (HILLGROVE)
BALLARAT AERODROME
DAYLESFORD
WARBURTON (O'SHANNASSY RESERVOIR (QUARTE
SCOTSBURN (MOUNT BONINYONG)
GLENMAGGIE WEIR
ERICA
CAPE BRIDGEWATER
COBDEN (POST OFFICE)
DROUIN BOWLING CLUB
EAST SALE AIRPORT
FISH CREEK (HODDLE RANGE)
PENSHURST (POST OFFICE)
CAULFIELD (RACECOURSE)
WILSONS PROMONTORY LIGHTHOUSE
CAPE OTWAY LIGHTHOUSE
NULLAWARRE
BANNOCKBURN
WONTHAGGI

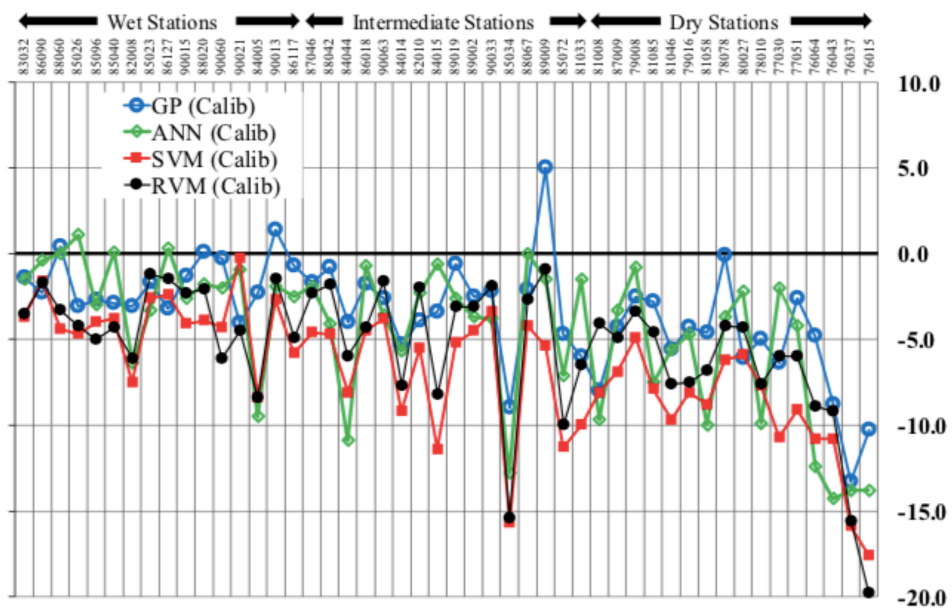
Results

Table 4
Percentage of selection of a kernel as the best for a given climate regime.

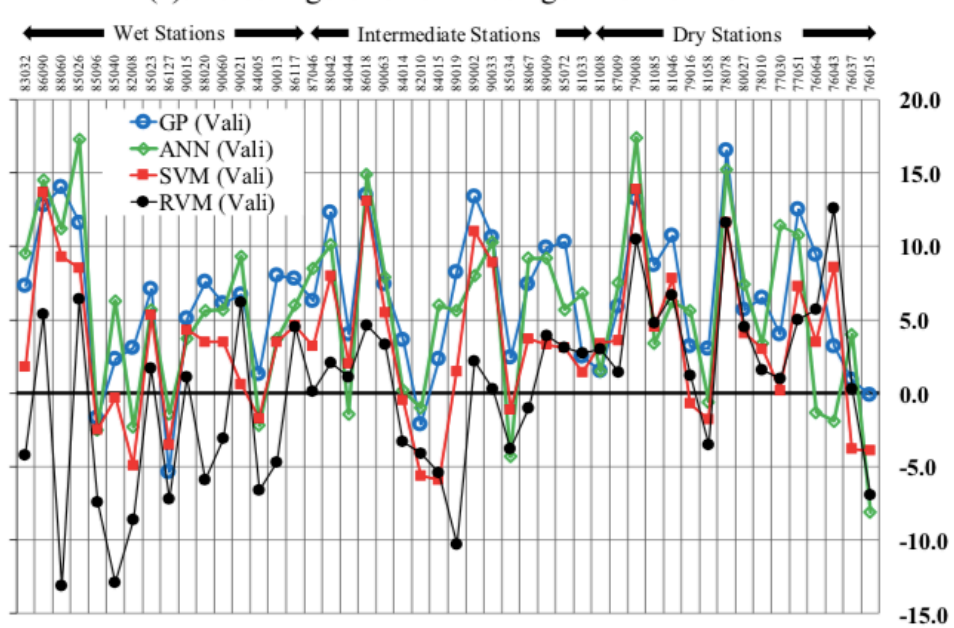
Machine learning technique	Climate regime	Kernel								
		Hyperbolic tangent	Polynomial	RBF	Laplacian	Bessel	ANOVA	Spline	String	Linear
SVM	Wet	0.0	34.4	13.0	25.0	18.8	8.9	0.0	0.0	0.0
	Intermediate	0.0	42.2	13.0	19.3	16.7	8.9	0.0	0.0	0.0
	Dry	0.0	39.1	10.4	25.0	14.6	10.9	0.0	0.0	0.0
RVM	Wet	1.0	33.9	13.0	25.0	18.2	8.9	0.0	0.0	0.0
	Intermediate	1.6	39.6	13.0	19.8	16.1	9.9	0.0	0.0	0.0
	Dry	0.5	38.0	10.4	25.5	15.1	9.9	0.5	0.0	0.0

无论是那种气候区域，SVM和RVM使用Polynomial kernel，Laplacian，Bessel，RBF和ANOVA都可以很好的模拟降水变量；同时，Spline，String，Linear和Hyperbolic tangent kernals无法得到降尺度模型所需的信息。

Results



(a). Percentage of bias in average in Calibration

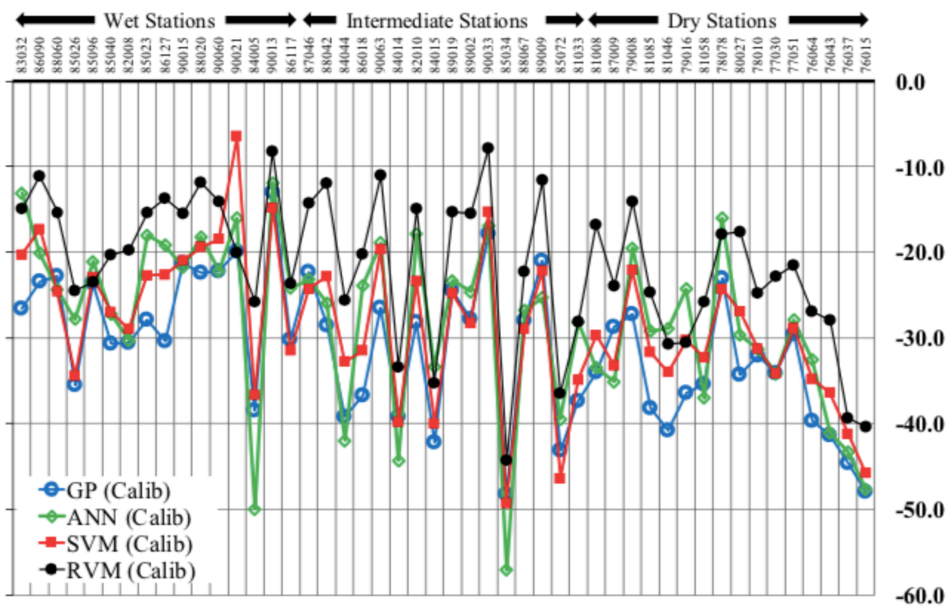


(b). Percentage of bias in average in Validation

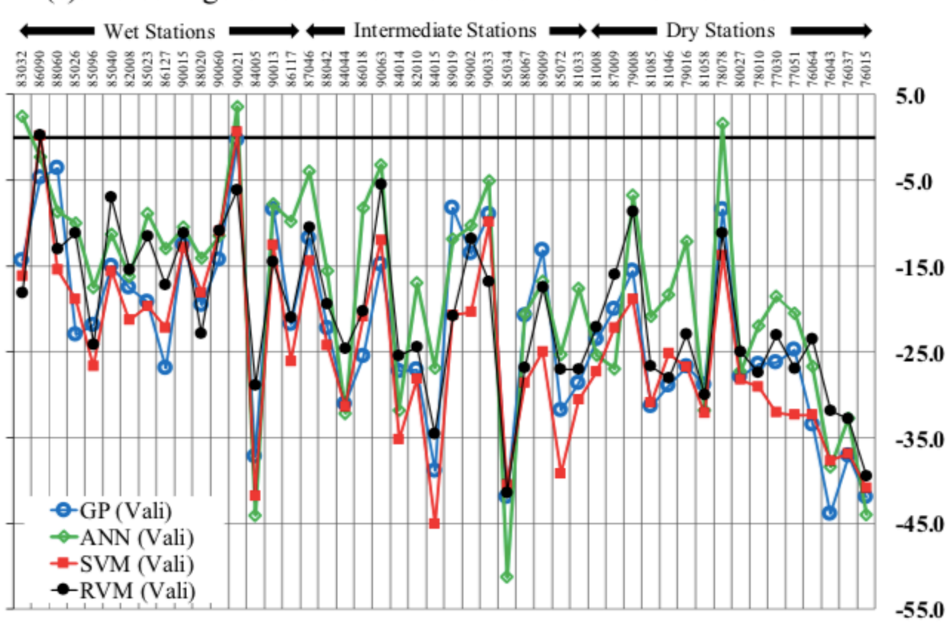
在训练集中，模型对于降水的模拟在干燥地区偏差较大，在湿润地区偏差较小（图6a）；同样的，基于GP的方法偏差最小，而基于SVM方法的偏差最大。

（图6b）在验证集中，湿润地区的观测站中，基于RVM方法的偏差最大，然而在大部分干燥或是温和区域中，RVM方法却反而有较小的偏差；同样，对于大部分观测站，基于SVM的方法具有较小的偏差，特别是和基于GP的方法比较时。

Results



(c). Percentage of bias in standard deviation in Calibration



(d). Percentage of bias in standard deviation in Validation

无论是在训练集还是在验证集，对于所有机器学习的方法，在大部分台站上，均低估了降水的标准差；而这一点在相对干燥的台站上表现更为明显。

机器学习在降尺度过程中对于变量标准差的低估已经在许多文献中提及过，这是由于大尺度气候信息所包含的方差为含有可解释所需水文气候变量方差的信息。

在验证集中，对于大多数的台站，ANN方法的标准差偏差比是最小的。

Results

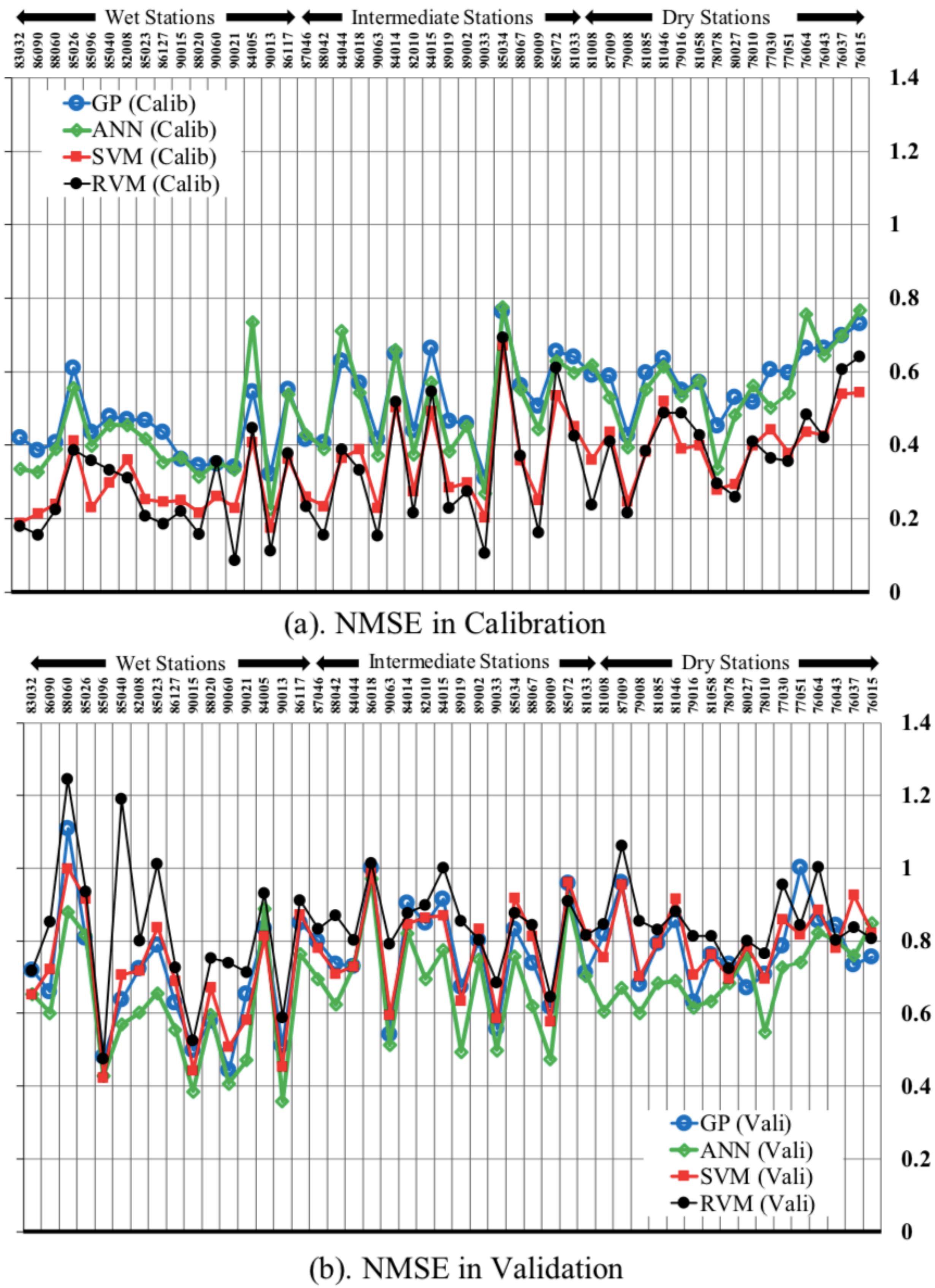


Fig. 9. Normalised mean square error of precipitation simulated by downscaling models.

在训练集中，无论是哪个地区，基于RVM方法的降尺度模型的NMSE最小，而GP和ANN的NMSE较大。

在验证集中，基于RVM方法的降尺度模型在大部分潮湿地区的NMSE较大。

Results

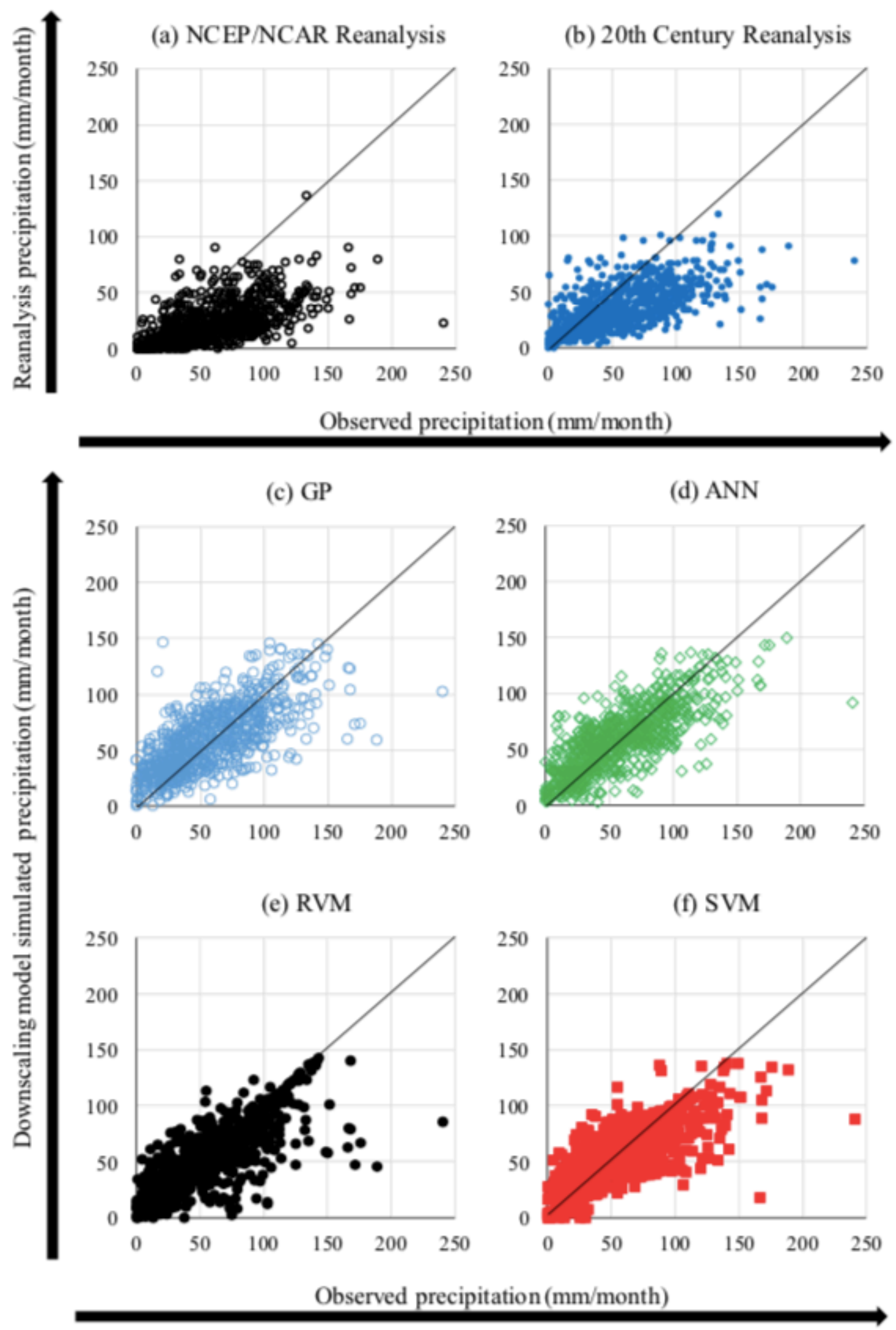


Fig. 10. Agreement between observed precipitation, reanalysis precipitation, and downscaled precipitation.

a significant underestimation

Far more accurate than the raw precipitation

Conclusions

1. 基于RVM和ANN方法的降尺度模型在潮湿地区训练集和验证集上都具有较小的偏差，因此对于洪水的预报上更加推荐使用RVM和ANN方法；
2. 对于干燥地区的降水，RVM方法具有较小的偏差，因此在做干旱地区的分析时更加推荐使用RVM方法；
3. 无论是那种气候区域，SVM和RVM使用Polynomial kernel，Laplacian，Bessel，RBF和ANOVA都可以很好的模拟降水变量；同时，Spline，String，Linear和Hyperbolic tangent kernals无法得到降尺度模型所需的信息。

谢谢