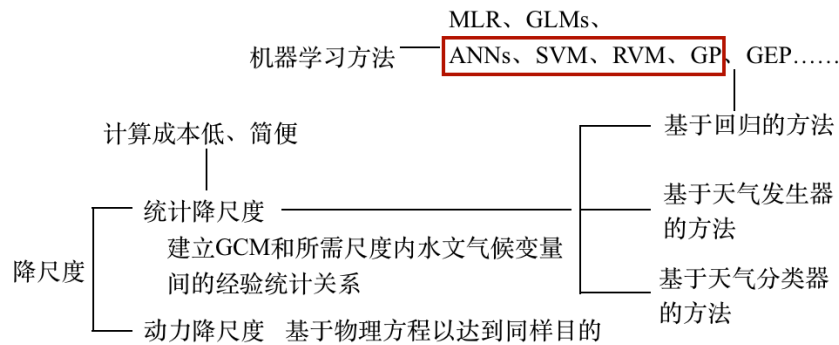


2019年10月20日

Statistical downscaling of precipitation using machine learning techniques

Introduction:



- 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。

Methods:

NCEP/NCAR reanalysis data set;

48 precipitation observation stations located across Victoria (237,000km²), Australia were selected;

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

Results:

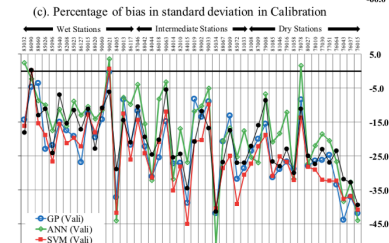
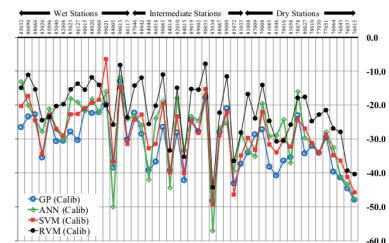
Table 4
Percentage of selection of a kernel as the best for a given climate regime.

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	0.0
	Intermediate	0.0	42.2	13.0	19.3	16.7	8.9	0.0	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	0.0
RVM	Wet	1.0	33.9	13.0	25.0	18.2	8.9	0.0	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	0.0
	Dry	0.5	38.0	10.4	25.5	15.1	9.9	0.5	0.0	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 kernels无法得到降尺度模型所需的信息。

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 kernels无法得到降尺度模型所需的信息。



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。