Robust nonlinear machine learning methods applied to climate and weather

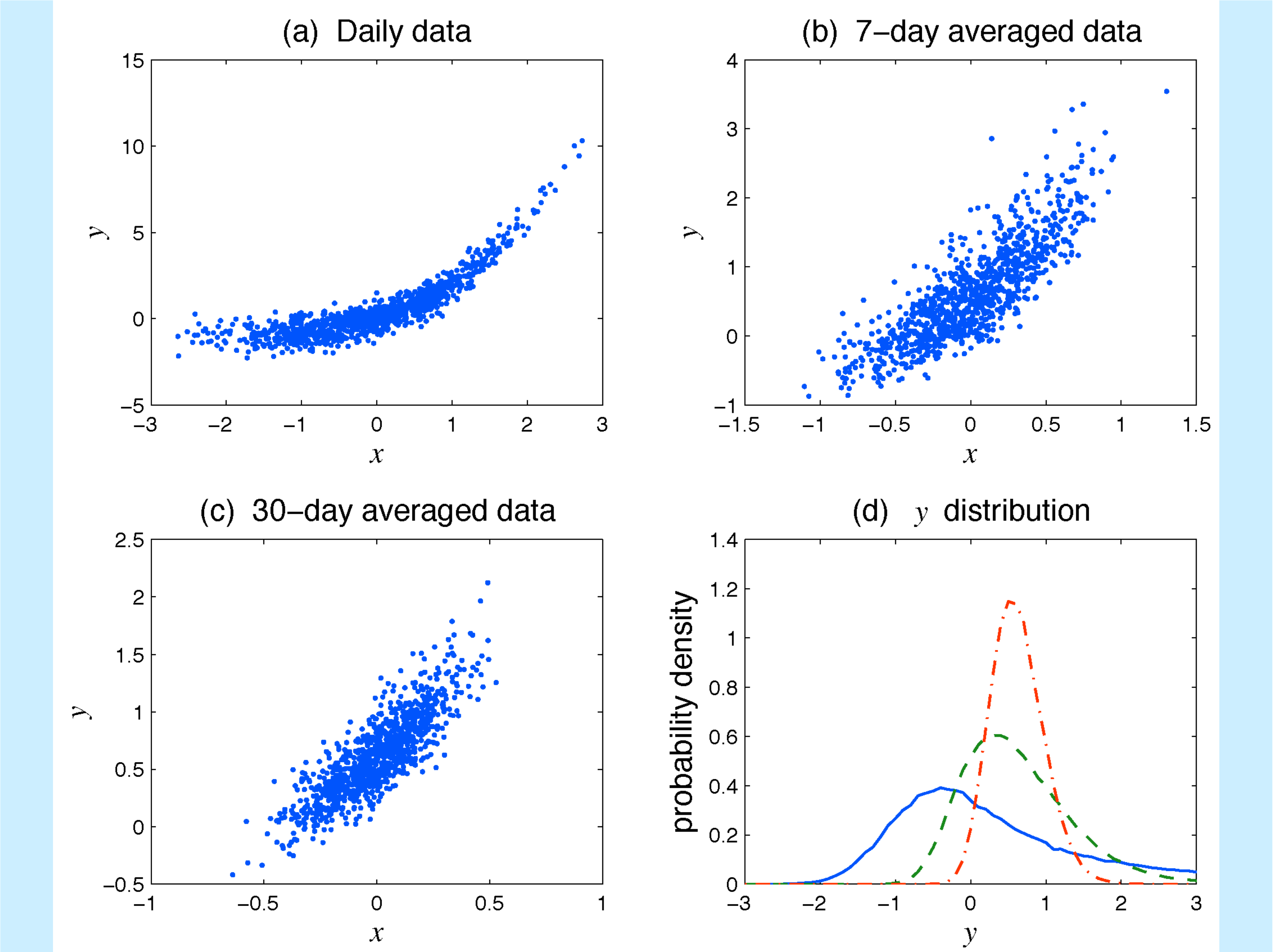
William W. Hsieh

University of British Columbia

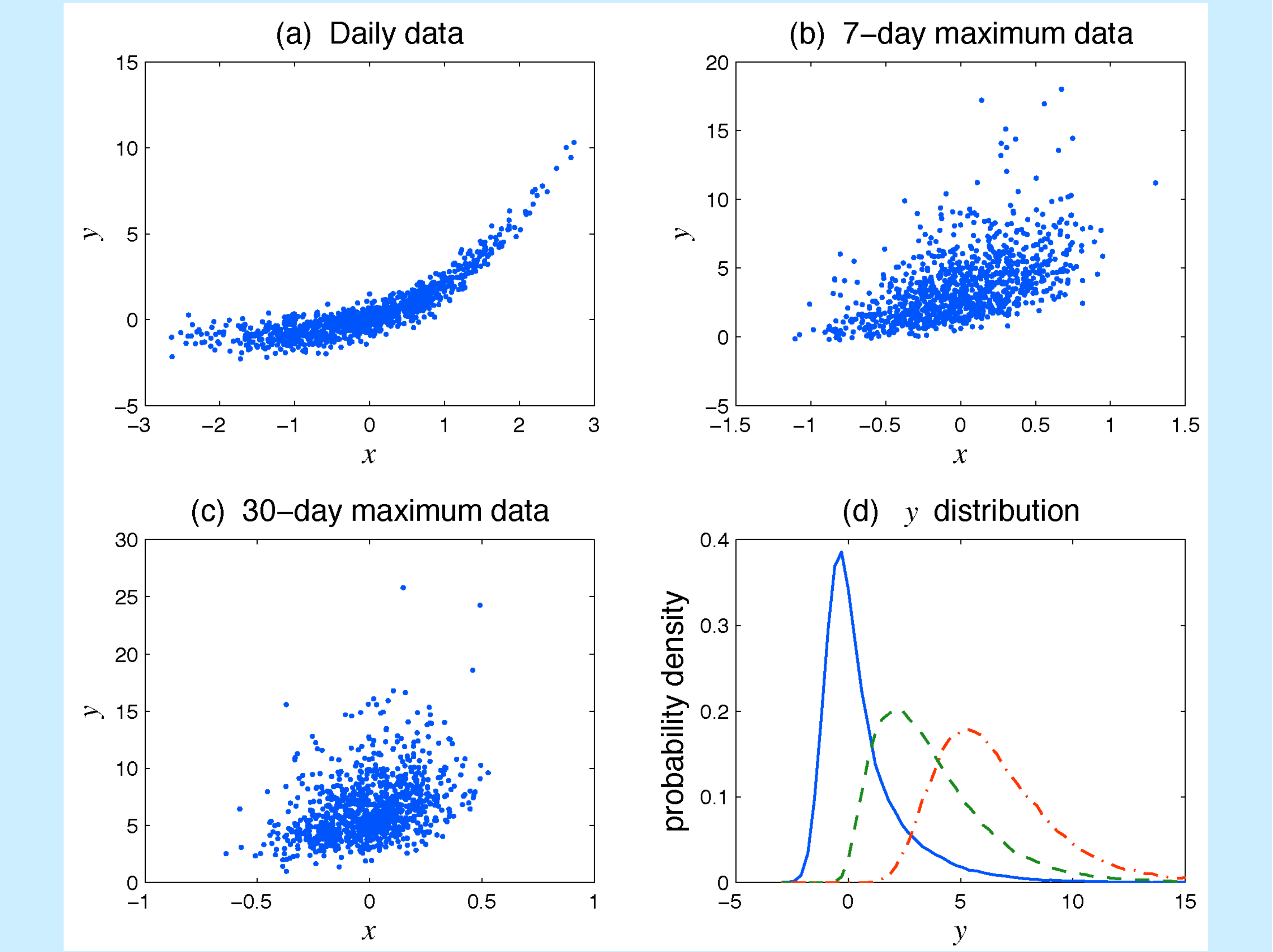
Collaborators:

UBC: Zhen Zeng, Joel Finnis,

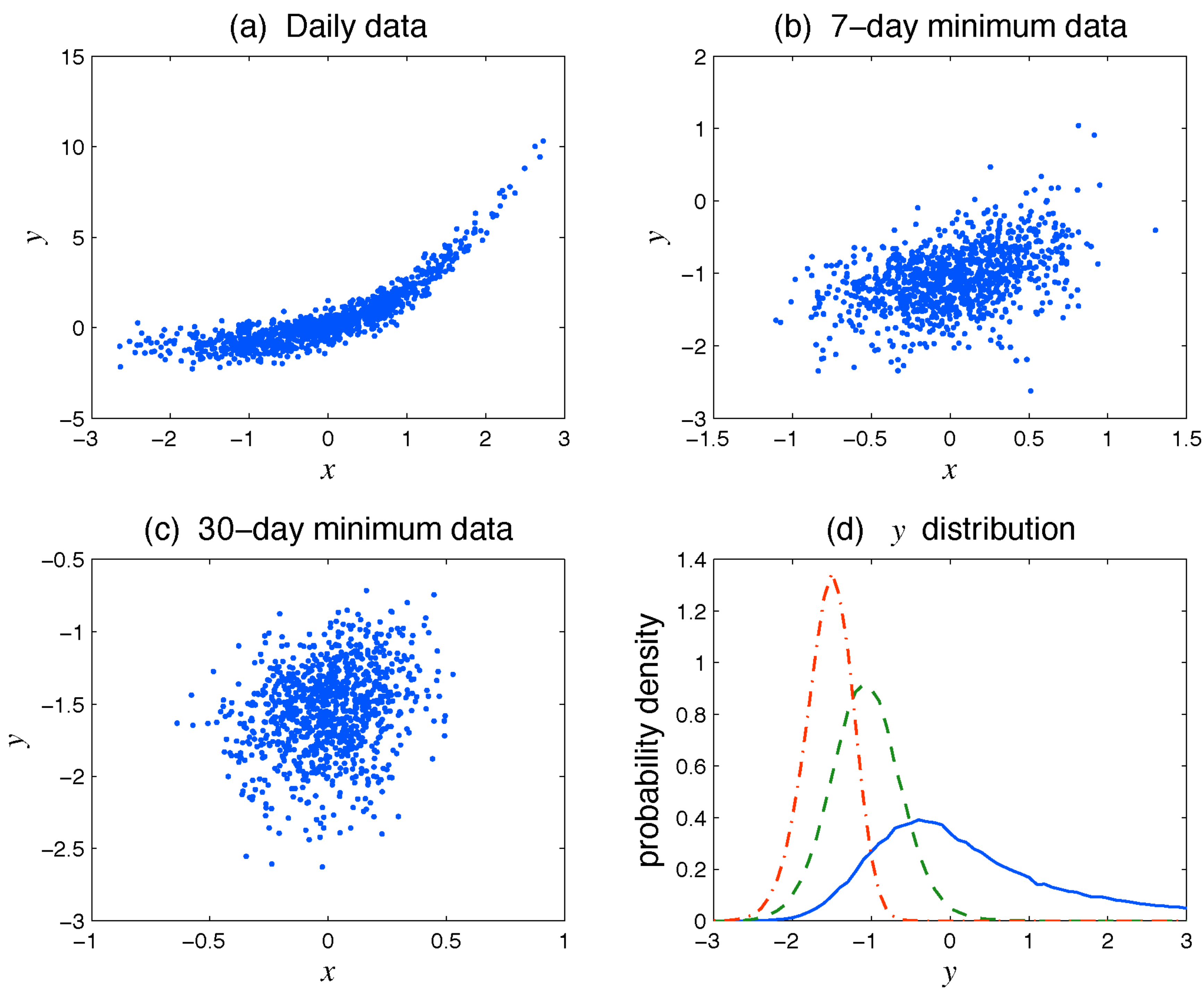
EC: Alex Cannon, Amir Shabbar, Bill Burrows, Bill Merryfield, Hai Lin



(a) Daily data
(b) 7–day averaged data
(c) 30–day averaged data
(d) y distribution
x
y
probability density



(a) Daily data
(b) 7–day maximum data
(c) 30–day maximum data
(d) y distribution
x
y
probability density


(a) Daily data
(b) 7–day minimum data
(c) 30–day minimum data
(d) y distribution
probability density
x
y

- Nonlinear methods should beat linear methods when working with:
  - weather variables,
  - some seasonal extreme variables,
  - but doubtful with seasonal mean variables
- But seasonal extreme variables are noisier than seasonal mean variables,
  - need robust methods!

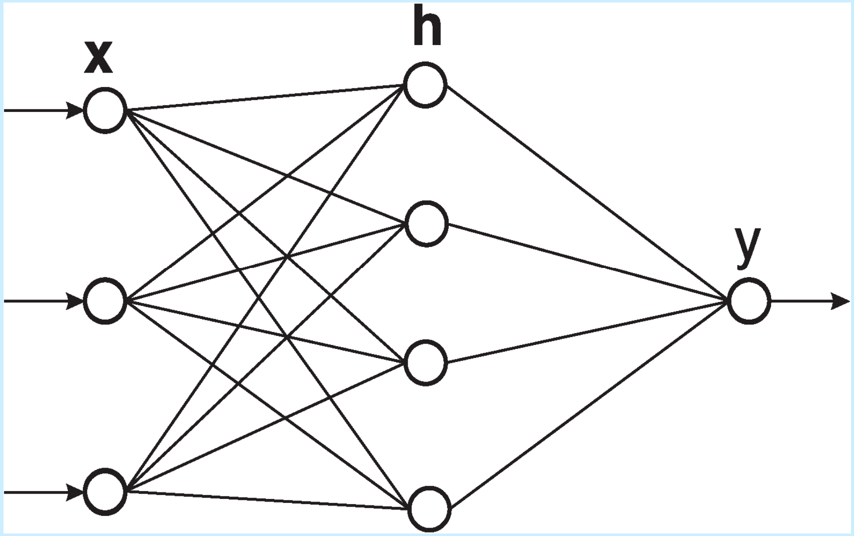
Regression methods

Linear regression (LR):

$$y = \sum_i a_i\, x_i + a_0$$

Neural networks (NN):

Adaptive basis fns h_j



$$y = \sum_j a_j \; h_j(\mathbf{x};\mathbf{w}) + a_0$$

$$= \sum_j a_j \overbrace{\tanh\left(\sum_i w_{ij} x_i + w_{0j}\right)} + a_0$$

Kernel methods

Non-adaptive basis fns.: $$y = \sum_j a_j\, \phi_j(\mathbf{x}) + a_0$$

Adv.: linear optimization, no local minima.
Disadv.: Many (infinite?) no. of basis fns.

If optimization problem can involve only dot products like $\phi^{\top}(\mathbf{x}')\phi(\mathbf{x})$

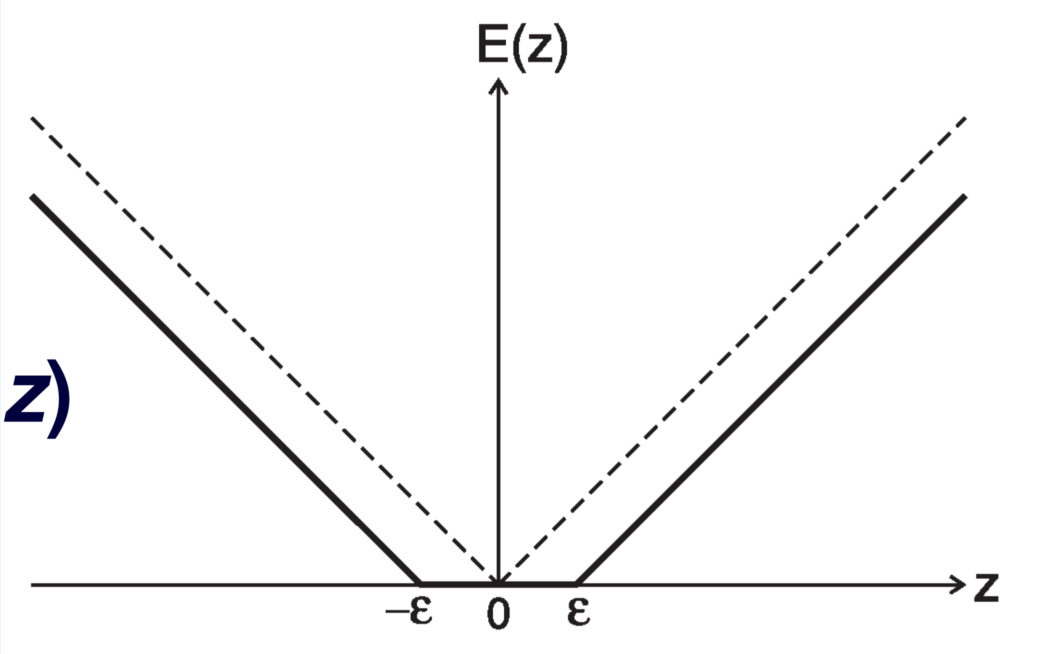
and the dot product given by a kernel function *K*: $$\phi^{\top}(\mathbf{x}')\phi(\mathbf{x}) = K(\mathbf{x}', \mathbf{x})$$

$$y = \sum_{k=1}^{n} \alpha_k\, K(\mathbf{x}_k, \mathbf{x}) + \alpha_0$$

Gaussian/RBF kernel $$K(\mathbf{x}_k, \mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_k)\|^2}{2\sigma^2}\right)$$

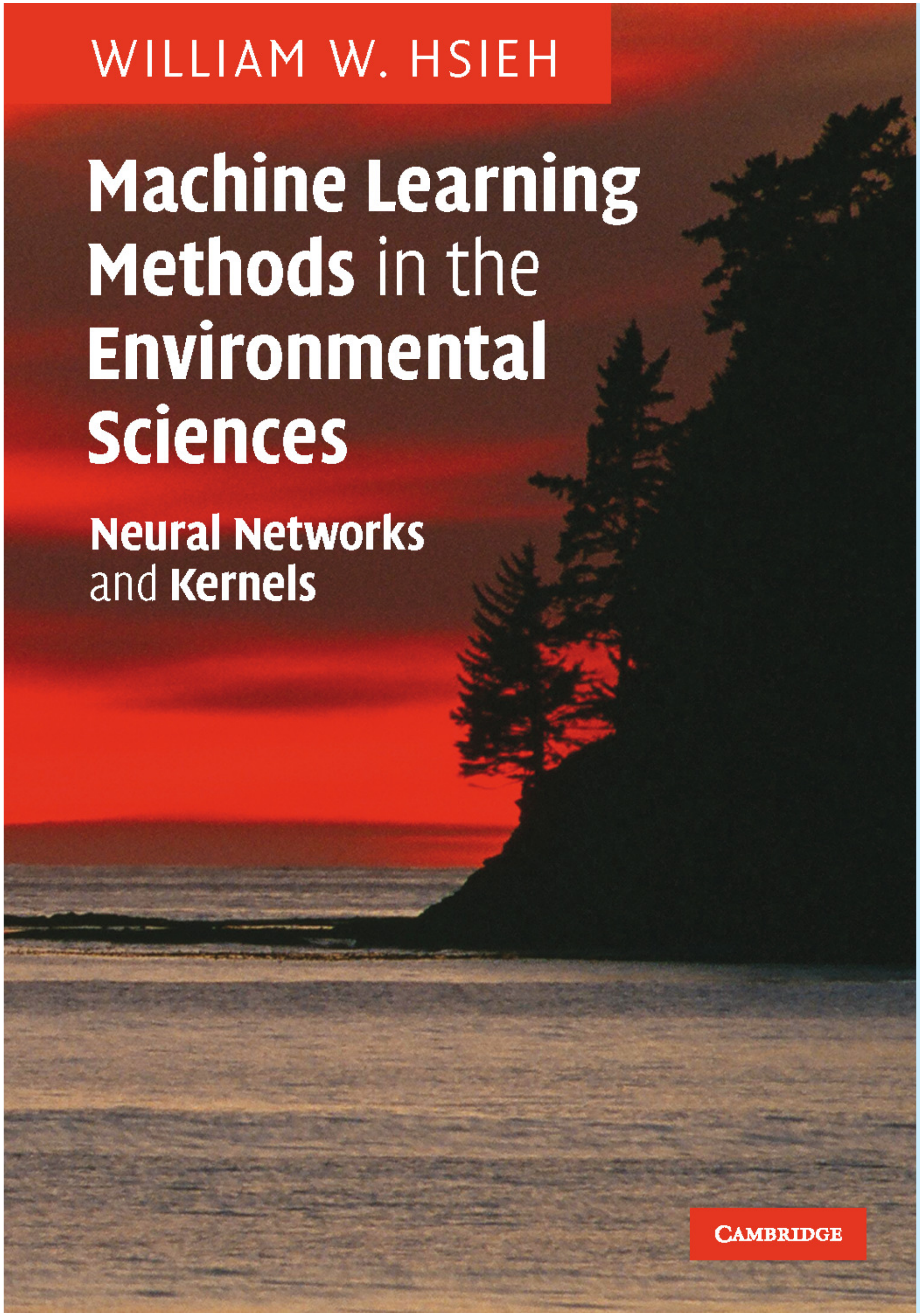
Common kernel method:

Support vector machines (SVM) for regression (SVR)



$z = y - y_{obs}$

Robust error norm $E(z)$



To appear, Sept.2009

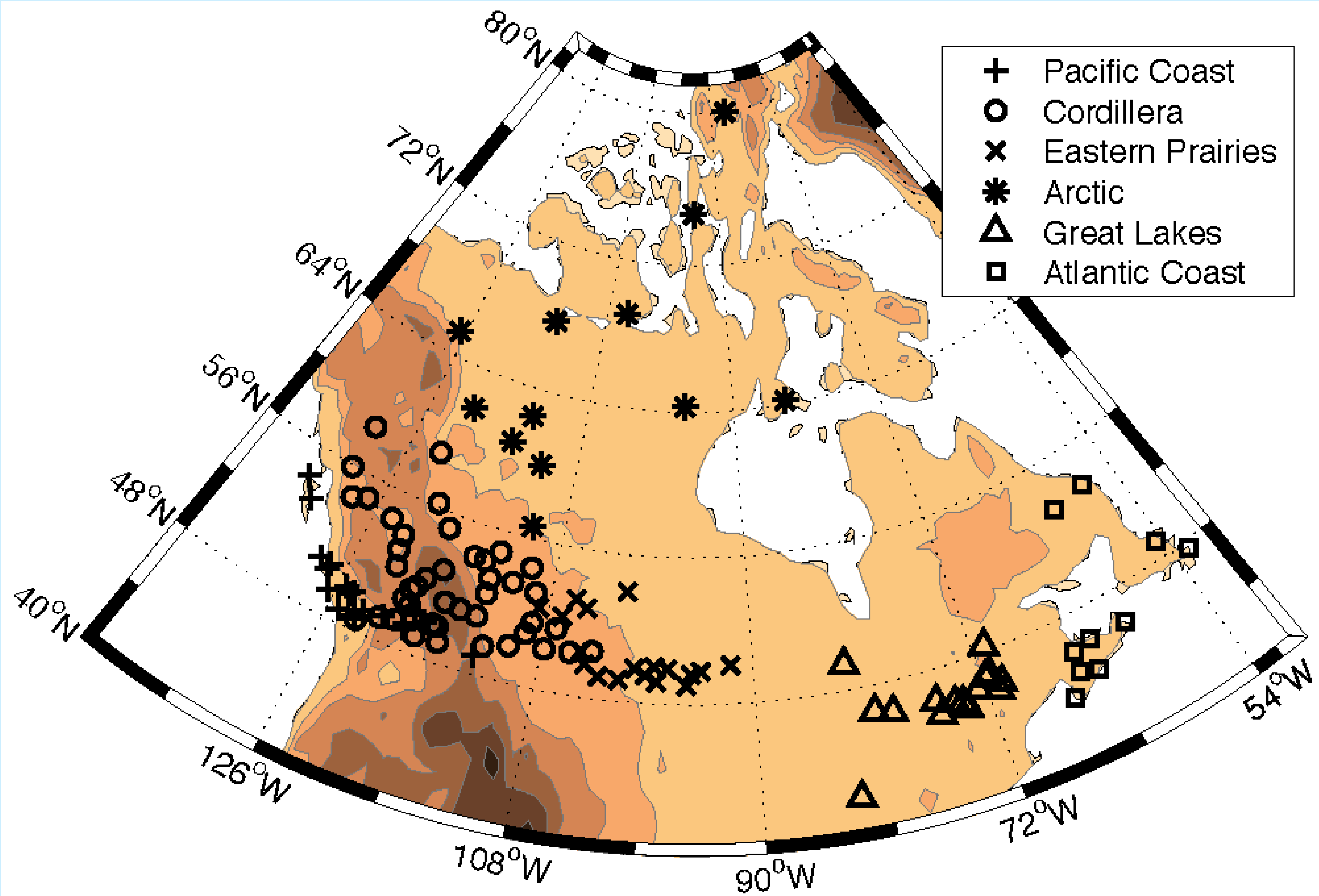
Applications in:

- Remote sensing
- GCM
- post-processing & downscaling GCM output
- data analysis
- forecasting

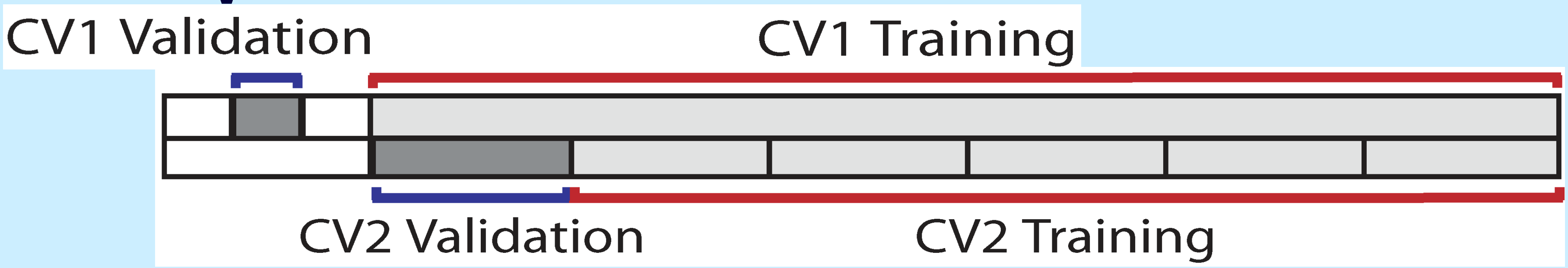
etc.

Forecast max. 5-day PRCP in DJF

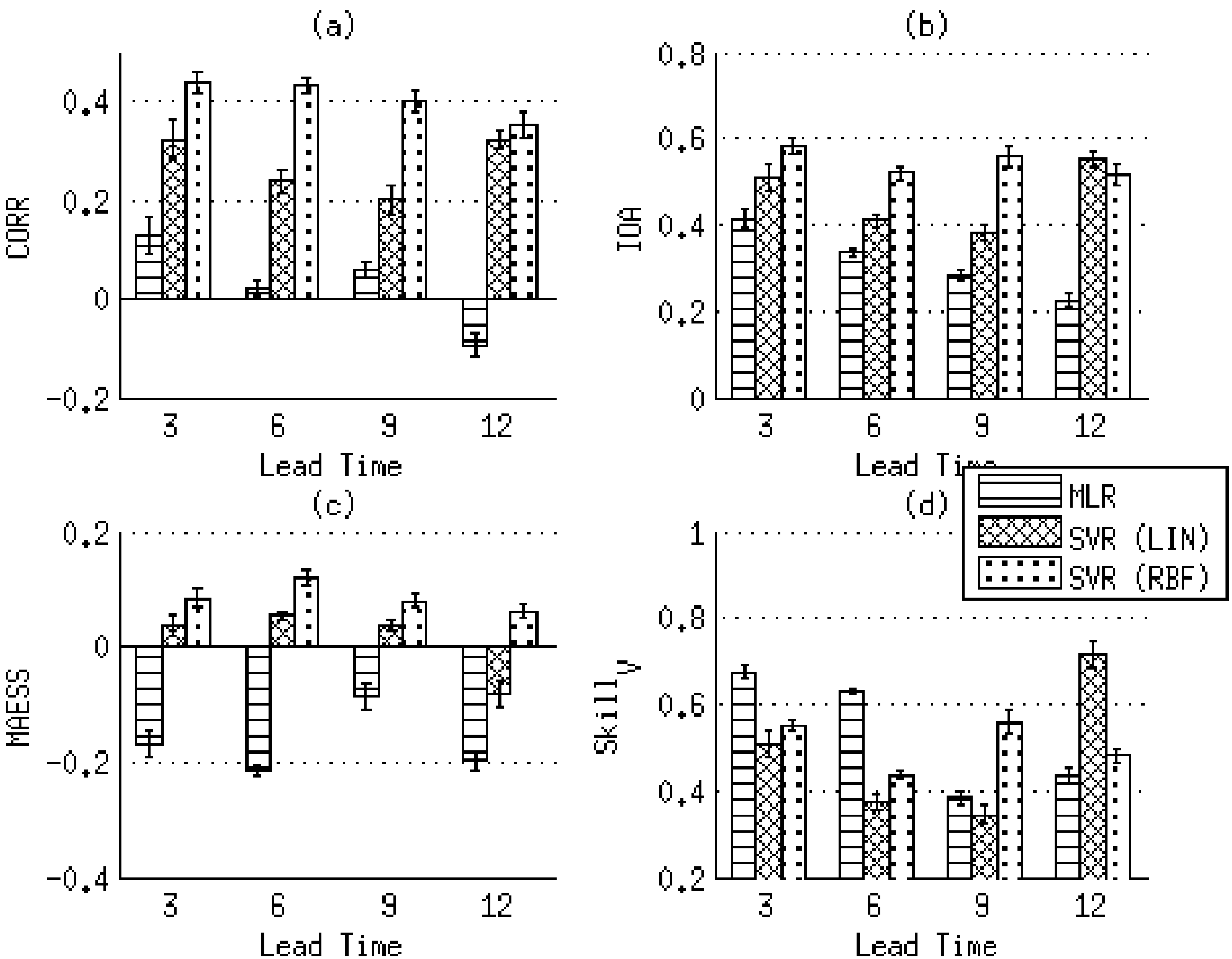
- Cluster analysis gave 6 regions



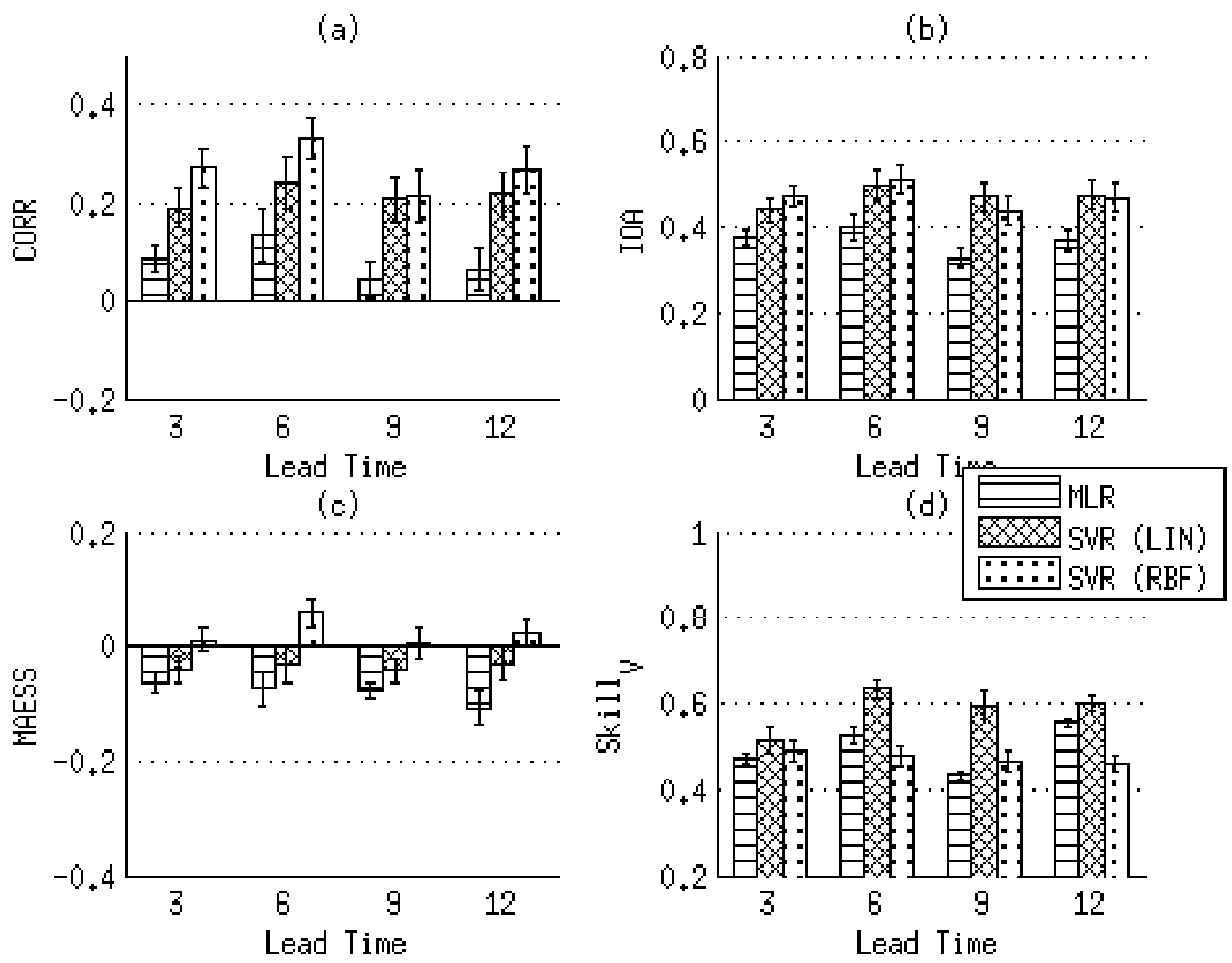
- **Predictors:**
  - **Quasi-global SST**
  - **N. Hem. Z500**
  - **6 climate indices (Nino3.4, PNA, PDO, NAO, SC, EA)**
- **Forecast scores**
  - **Correlation**
  - **Willmott's Index of Agreement**
  - **MAE skill score**
  - **Skill_V = S.D. of forecasts / S.D. of observ.**



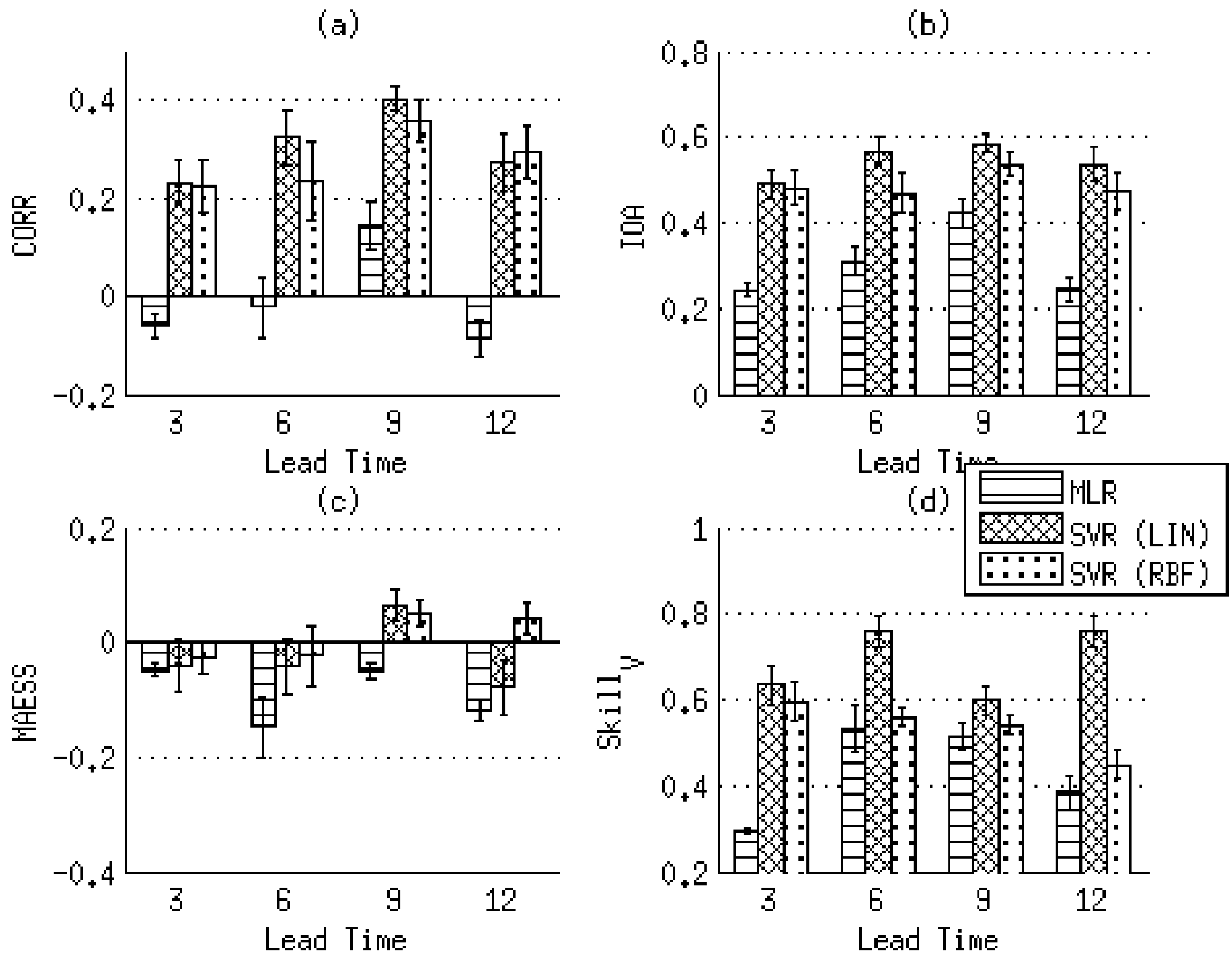
Eastern Prairies



Pacific coast

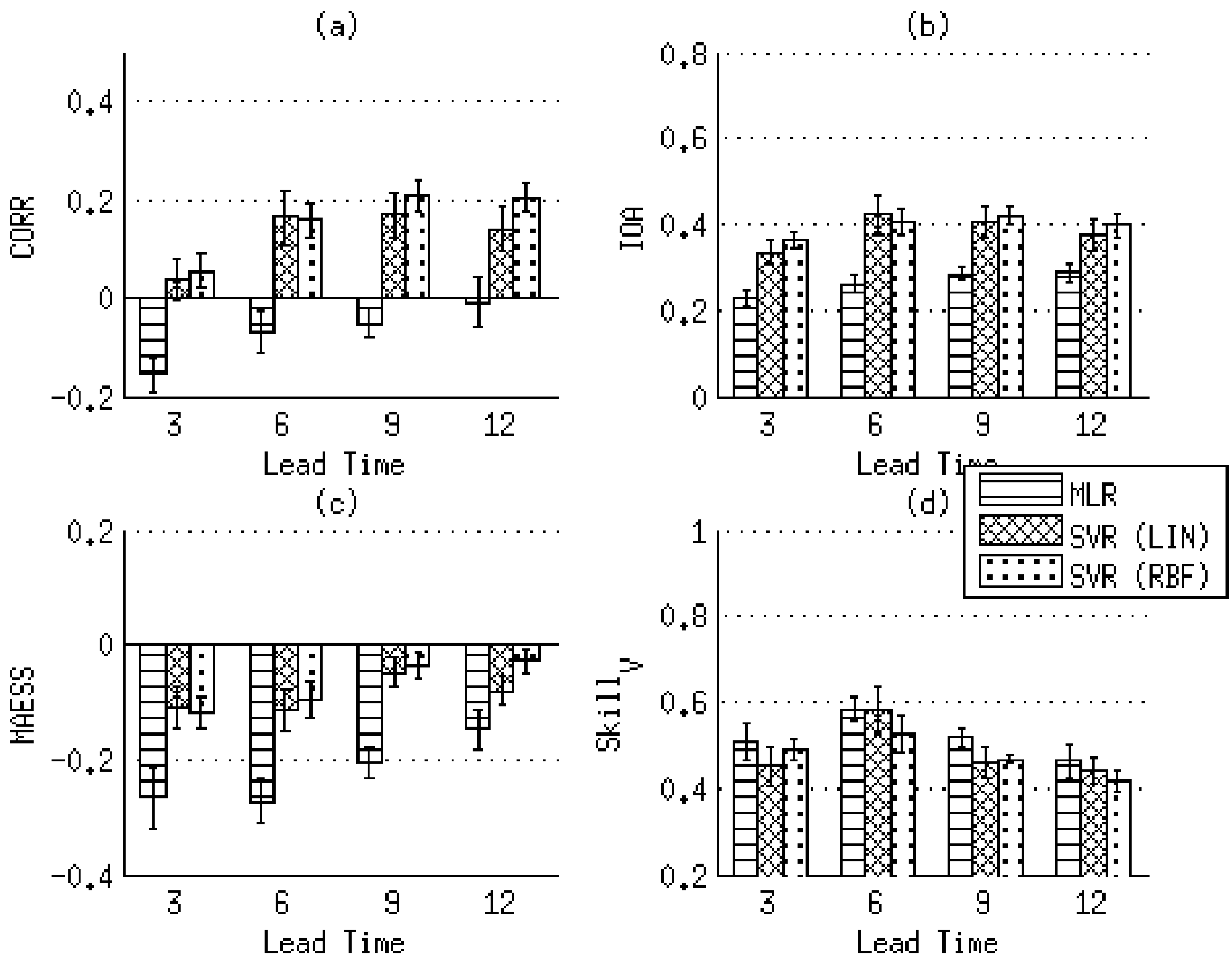


Arctic



(a)
CORR
0.4
0.2
0
-0.2
3
6
9
12
Lead Time
(b)
IOA
0.8
0.6
0.4
0.2
0
3
6
9
12
Lead Time
(c)
MAESS
0.2
0
-0.2
-0.4
3
6
9
12
Lead Time
(d)
Skill_V
1
0.8
0.6
0.4
0.2
3
6
9
12
Lead Time
MLR
SVR (LIN)
SVR (RBF)

Atlantic coast



Conclusion

- Nonlinear methods most suited for weather and some seasonal extreme variables, but not for seasonal mean variables.
- For seasonal extreme PRCP forecasting, SVR (nonlin.) and SVR (lin.) can beat MLR, indicating importance of robust error norm.
- SVR (nonlin.) beats SVR (lin.) in eastern Prairies, but not in Arctic.