Assignments/Dates

- Chapter 3 notes due Feb 15
- http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-88-6-853 Cerveny Article
- Amazon Drought paper review Due Feb 17
- Odds Are It's Wrong. Due Feb 22
  - http://www.sciencenews.org/view/feature/id/57091/title/Odds_Are,_Its_Wrong
  - Read and summarize issues about significance testing in a few paragraphs
- Chapter 4 notes due Feb 24
- Exam March 1.

Extreme Weather Records
Compilation, Adjudication, and Publication

BY RANDALL S. CERVENY, JAY LAWRIMORE, ROGER EDWARDS, AND CHRISTOPHER LANDSEA

Table 1. U.S. weather elements and characteristics currently within the jurisdiction of the NCEC.

Weather element	Characteristic	Value	Date	Location
Temperature	Maximum	56.67°C (134°F)	10 Jul 1913	Greenland Ranch, California
	Minimum	–62.22°C (–80°F)	23 Jan 1971	Prospect Creek, Alaska
	Max 24-h change	39.44°C (103°F)	14–15 Jan 1972	Loma, Montana
Snow	Max 24-h snowfall	1.925 m (75.8 in.)	14–15 Apr 1921	Silver Lake, Colorado
	Max seasonal snowfall (July–June)	28.956 m (1140 in.)	1998–1999	Mt. Baker Ski Area, Washington
	Max snow depth	11.455 m (451 in.)	11 Mar 1911	Tamarack, California
Rain	Max 24 hr	1.092 m (43 in.)	25–26 Jul 1979	Alvin, Texas
	Least annual	0.0 m	1929	Death Valley, California
	Max annual	17.903 m (704.83 in.)	1982	Kukui, Hawaii
	Longest dry period	767 days	3 Oct 1912–8 Nov 1914	Bagdad, California
Wind	Max gust	103.3 m s^{-1} (231 mph)	12 Apr 1934	Mt. Washington, New Hampshire
Hail	Largest (diameter/circumference)	17.78 mm/47.625 mm (7 in./18.75 in.)	22 Jun 2003	Aurora, Nebraska
	Heaviest	0.7575 kg (1.67 lbs)	3 Sep 1970	Coffeyville, Kansas
Pressure	Lowest	892.3 mb (26.35 in. Hg)	2 Sep 1935	Matecumbe Key, Florida
	Highest	1078.6 (31.85 in. Hg)	31 Jan 1989	Northway, Alaska

US CLIVAR/NCAR ASP Researcher Colloquium (by Invitation/Application)
Statistical Assessment of Extreme Weather Phenomena under Climate Change
NCAR Foothills Lab, Boulder, Colorado, USA June 1317, 2011

The US CLIVAR/NCAR Research Colloquium will assemble climate researchers, statisticians, decision and policy makers to discuss the state-of-the-art in science of weather and climate extremes and its application to real-world decision-making.

Specific objectives:

- Determine climate and weather extremes that are crucial in resources management and policy making
- Identify the current state of the science of climate and weather extremes including uncertainties and information gaps in real-world applications
- Obtain insights into the capabilities of climate models in identifying and modeling such extreme events.
- Assess efficacy of statistical methods and tools to analyze and model extreme events under climate change
- Develop interdisciplinary research directions in modeling and application of climate extremes.

The Colloquium organizing committee invites the participation of researchers studying extreme events in observations and climate models; statistical modeling and identification of extremes; use of climate extremes in a suite of decision and policy making context.

If you are interested in participating, please submit an application with your CV and a brief statement of interest via an online application at http://www.regonline.com/Register/Checkin.aspx?EventID=931739. Travel support is available. You will be asked to identify any travel support needs in your application. **The deadline for applications is March 28, 2011.**

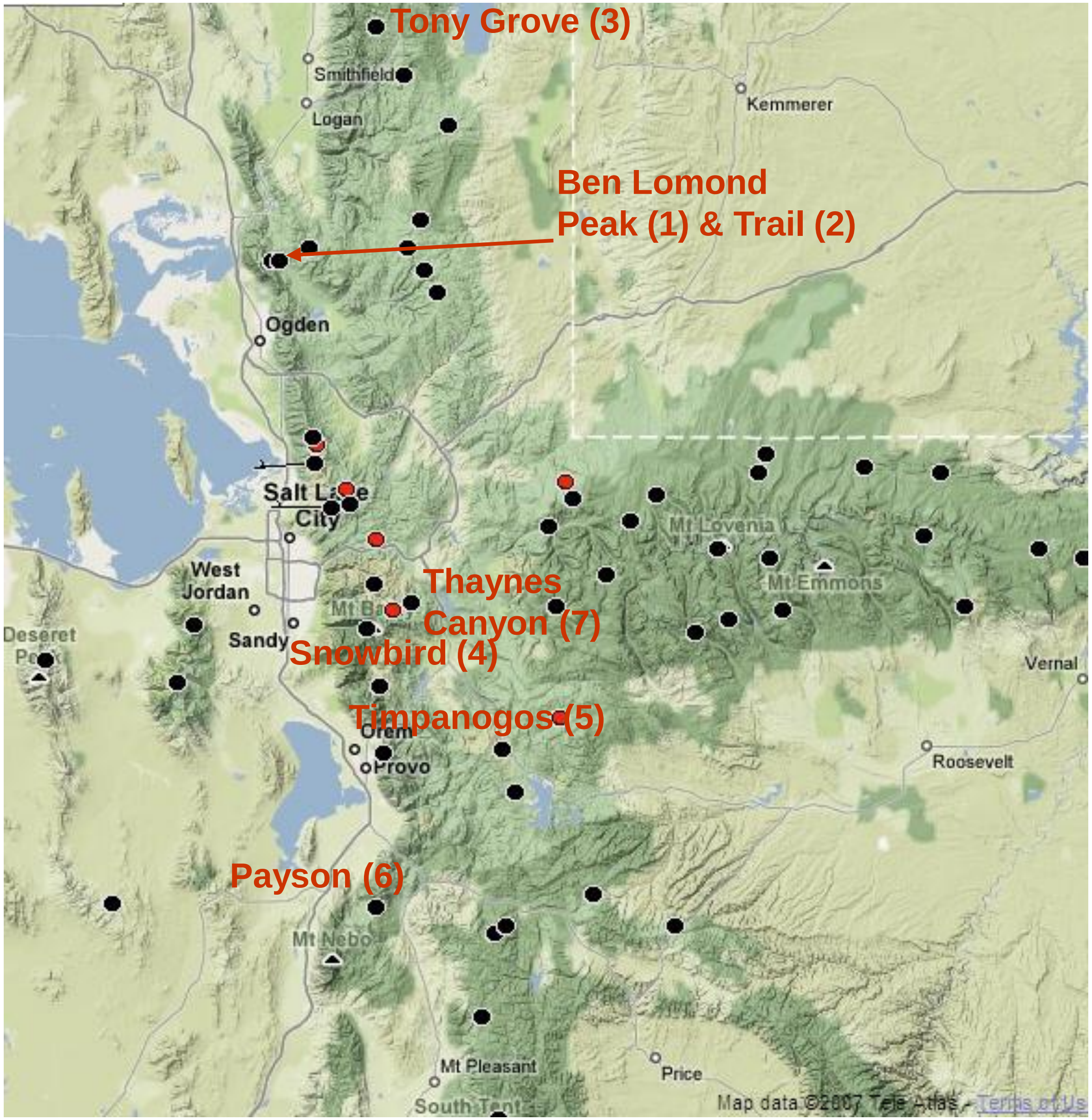
Science Magazine

- http://www.sciencemag.org/site/special/data/

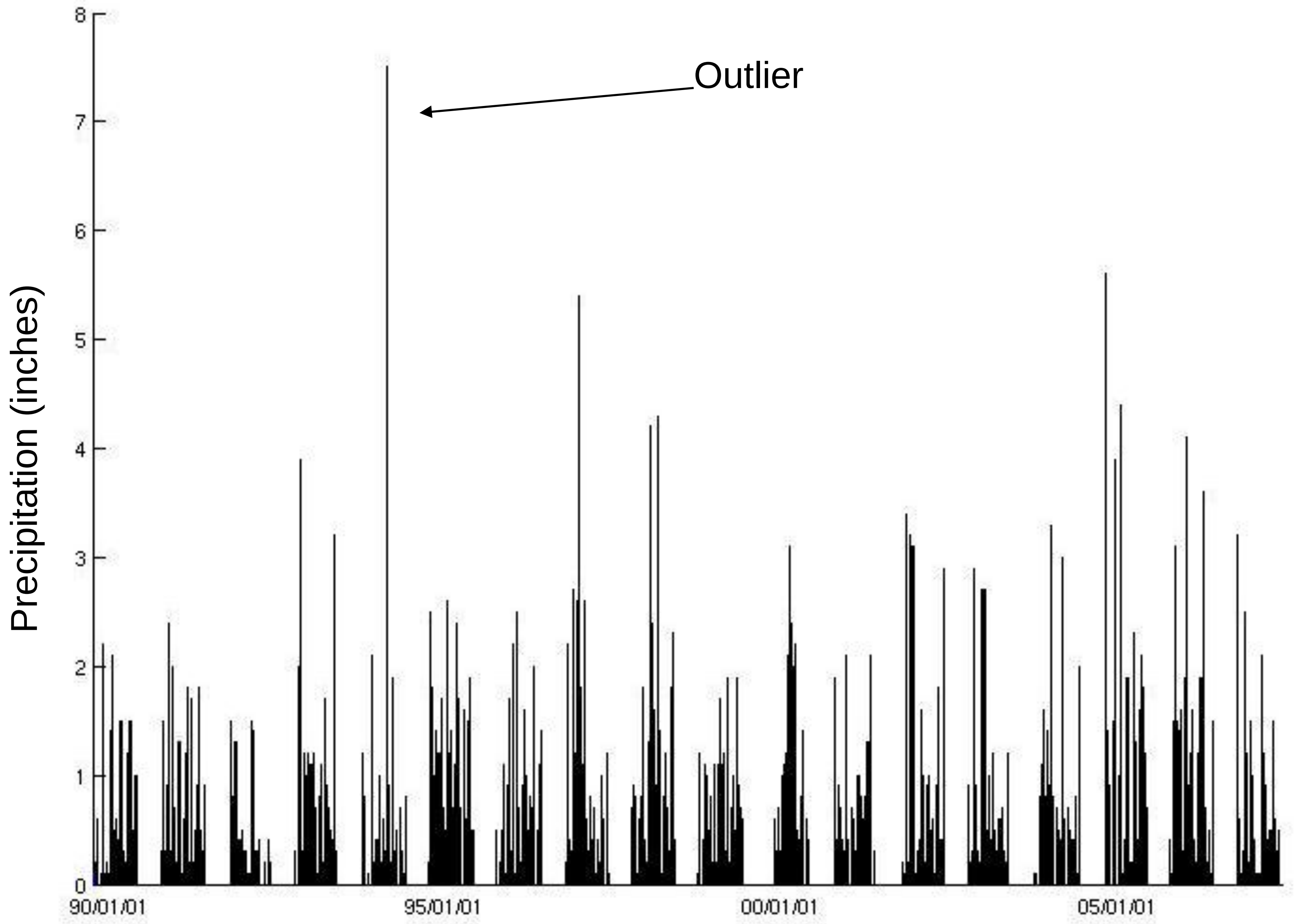
4. Exploratory Multivariate Data Analysis

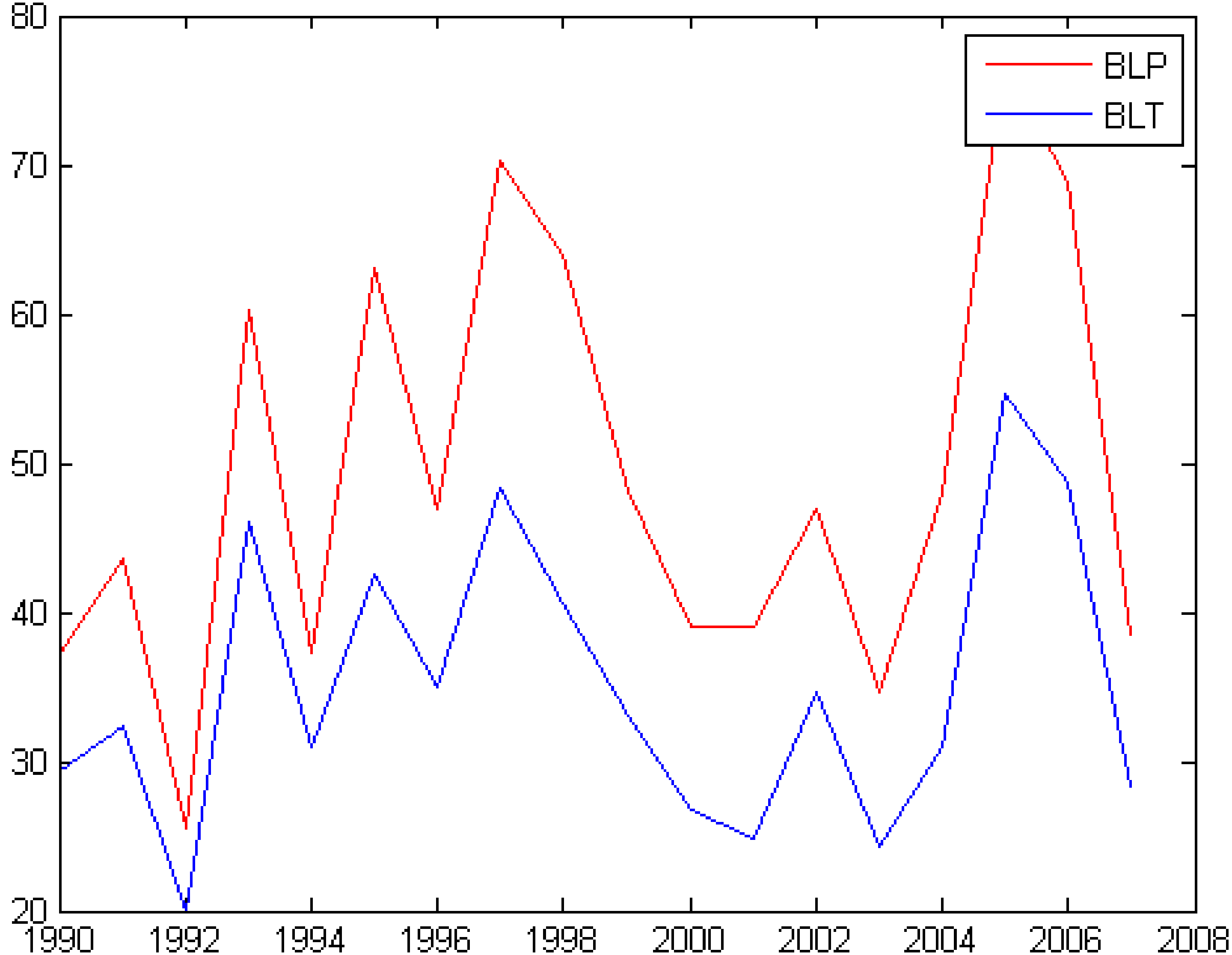
- want to relate how one or more phenomena are related to others
- Think of it as a multidimensional problem and the goal is to reduce the dimensionality through exploration of relationships
- Dealing with the dimensionality of environmental data sets in statistical analyses is of general concern (Murphy 1991; Mon. Wea. Rev., 1590-1601).
- Exploratory multivariate data analysis encompasses an array of tools to assess relationships between two or more samples

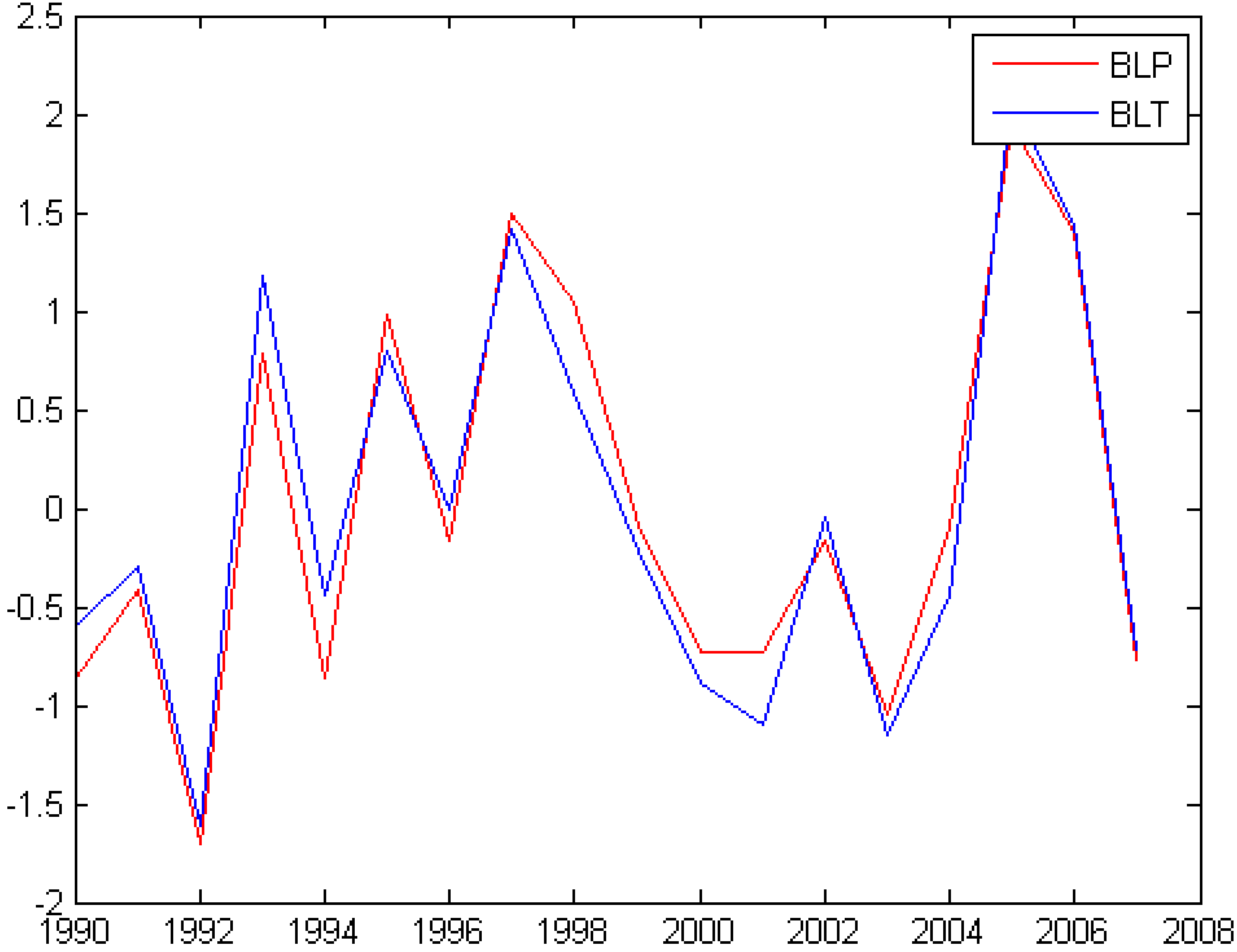
SNOTEL Sites

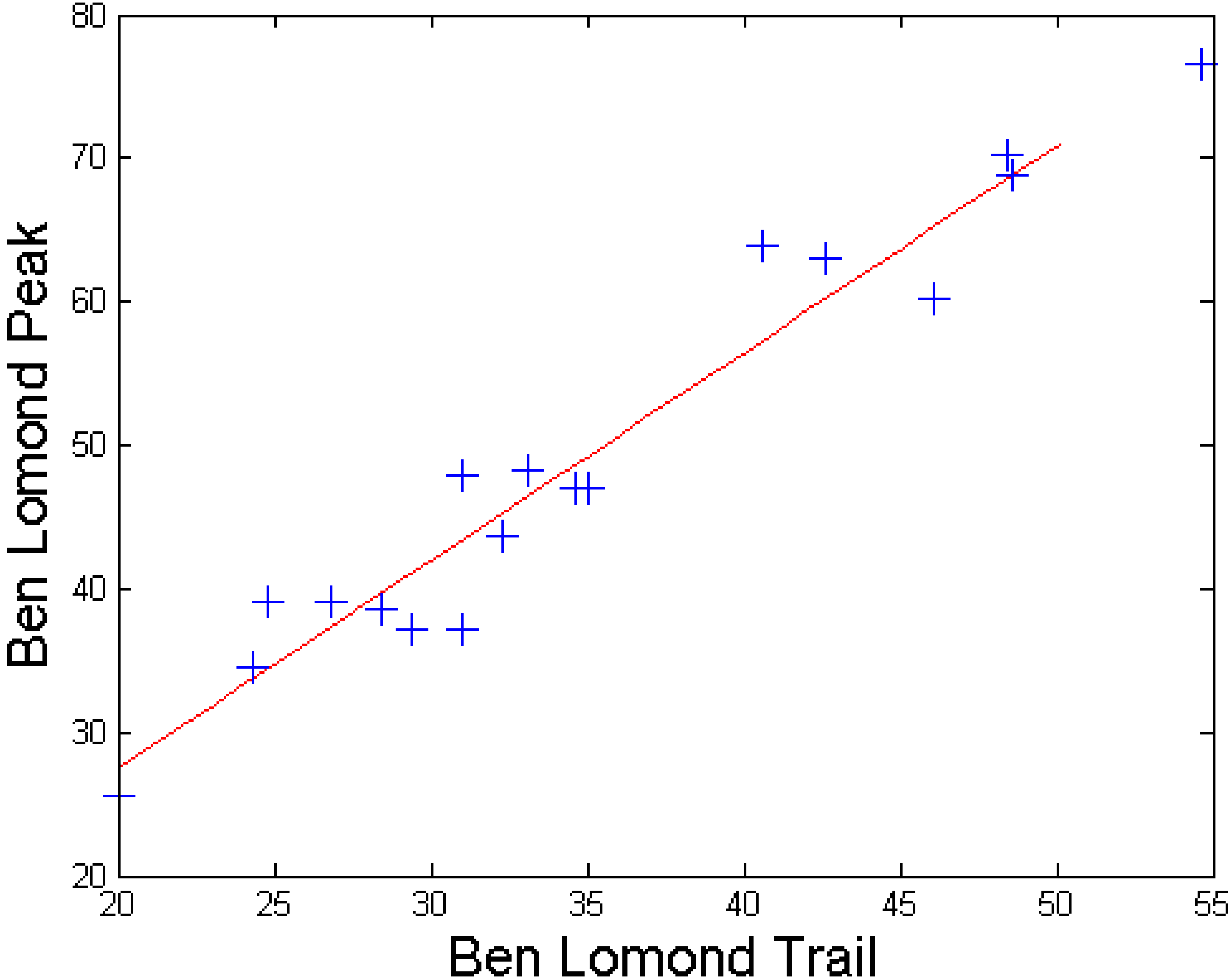


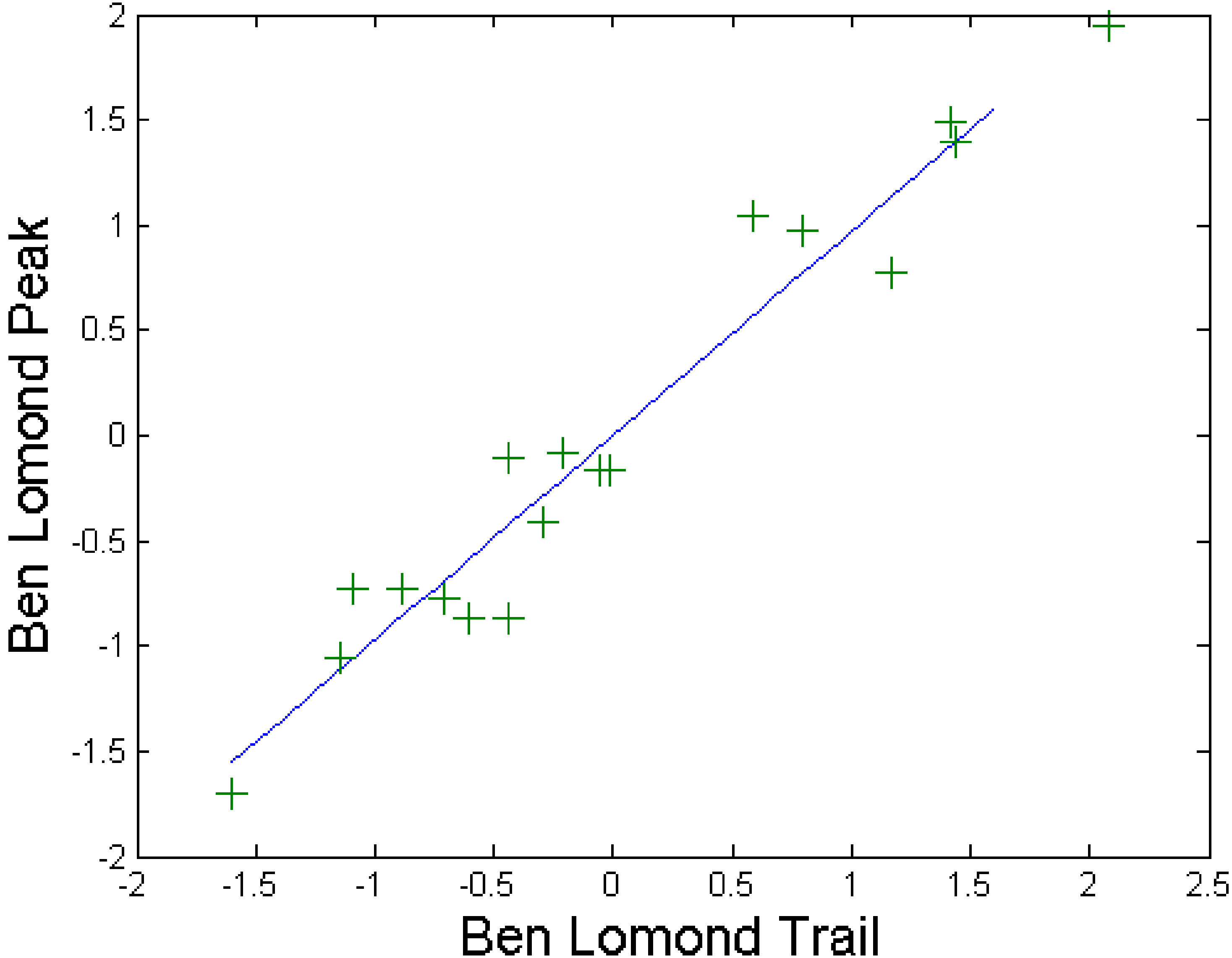
Ben Lomond Daily Precipitation




80
70
60
50
40
30
20
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
BLP
BLT


BLP
BLT
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008


80
70
60
50
40
30
20
Ben Lomond Peak
20
25
30
35
40
45
50
55
Ben Lomond Trail



2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
Ben Lomond Peak
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Ben Lomond Trail

Estimating Values of One Variable From Another

- X- Ben Lomond Trail
- Y- Ben Lomond Peak
- Want to estimate Peak from Trail
- Use pairs of observations from sample
- Need to determine coefficient b or r
- b- slope of linear estimate
- r- linear correlation

$$\hat{y}_i = \bar{y} + b(\hat{x}_i - \bar{x})$$

$$\hat{y}^*{}_i = r\hat{x}^*{}_i$$

Definitions

- Estimate $\hat{y}_i = \bar{y} + b(\hat{x}_i - \bar{x})$
- Error of estimate $e_i = y_i' - \hat{y}_i'$
- Want $\sum_{i=1}^{n} e_i^2$ to be a minimum
- Need to find the value of b that minimizes that sum

$$\frac{\partial}{\partial b}\sum_{i=1}^{n} e_i^2 = 0$$

- The value of b that minimizes the total error in the sample $b = \overline{x_i' y_i'} / \overline{(x_i')^2} = \overline{x_i' y_i'} / s_x^2$

Covariance

- Relates how departures of x and y from respective means are related
- Units are the product of the units of the two variables x and y
- Large and positive if sample tendency for:
  - large + anomalies of x occurring when large + anomalies of y AND
  - large - anomalies of x occurring when large - anomalies of y
- Large and negative if sample tendency for:
  - large + anomalies of x occurring when large - anomalies of y AND
  - large - anomalies of x occurring when large + anomalies of y
- Near zero when tendency for cancellation
  - large + anomalies of x occurring when both large – and + anomalies of y AND
  - large - anomalies of x occurring when both large – and + anomalies of y

Linear Correlation

$$r^2 = b^2 s_x{}^2 / s_y{}^2 = (\overline{x_i' y_i'})^2 / (s_x{}^2 s_y{}^2) \qquad r = (\overline{x_i' y_i'}) / \sqrt{\overline{x_i'^2}\,\overline{y_i'^2}}$$

$$x_i^* = x_i' / s_x, y_i^* = y_i' / s_y, r = (\overline{x_i^* y_i^*})$$

$$1 = r^2 + \frac{\overline{e_i^2}}{s_y{}^2}$$

y’s total sample variance = fraction of variance estimated by x + fraction of variance NOT explained by x

- Dimensionless number relates how departures of x and y from respective means are related taking into account variance of x and y
- r = 1. Linear fits estimates ALL of the variability of the y anomalies and x and y vary identically
- r = -1 perfect linear estimation but when x is positive, y is negative and vice versa
- r = 0. linear fit explains none of the variability of the y anomalies in the sample. Best estimate of y is the mean value

Linear Algebra is your friend

$$\vec{X}' = \begin{bmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{bmatrix} \quad \vec{Y}' = \begin{bmatrix} y'_1 \\ y'_2 \\ \dots \\ y'_n \end{bmatrix}$$

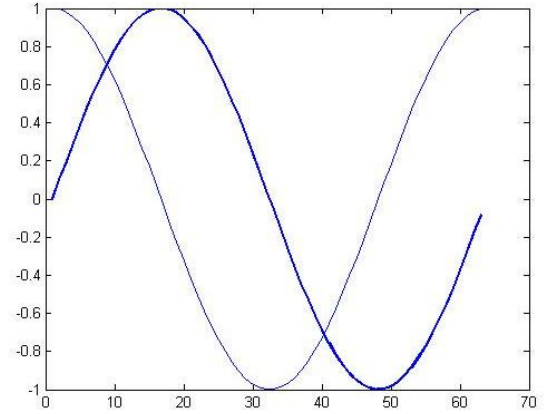
$$\overline{x'_i y'_i} = \vec{X}'^T \vec{Y}' / n$$

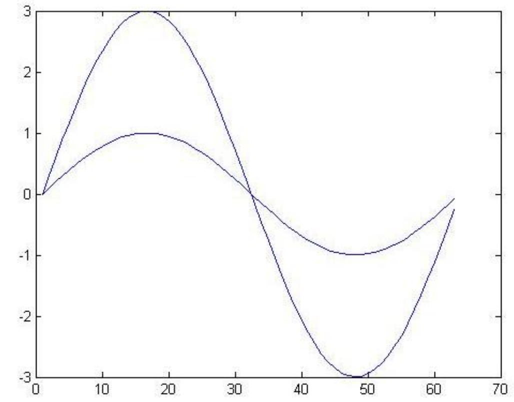
Variations on a theme

- Pearson correlation vs. Spearman rank correlation
- Spearman- correlating ranks not values
  - Good for skewed distributions

Stop and think before blindly computing correlations

- tendency to use correlation coefficients of 0.5-0.6 to indicate “useful” association.
  - 75%-64% of the total variance is NOT explained by a linear relationship if the correlation is in that range
- linear correlations can be made large by leaving in signals that may be irrelevant to the analysis. Annual and diurnal cyles may need to be removed
- large linear correlations may occur simply at random, especially if we try to correlate one variate with many, many others
- relationships in the data that are inherently nonlinear will not be handled well
- when two time series are in quadrature with one another then the linear correlation is 0
- Linear correlation provides no information on the relative amplitudes of two time series

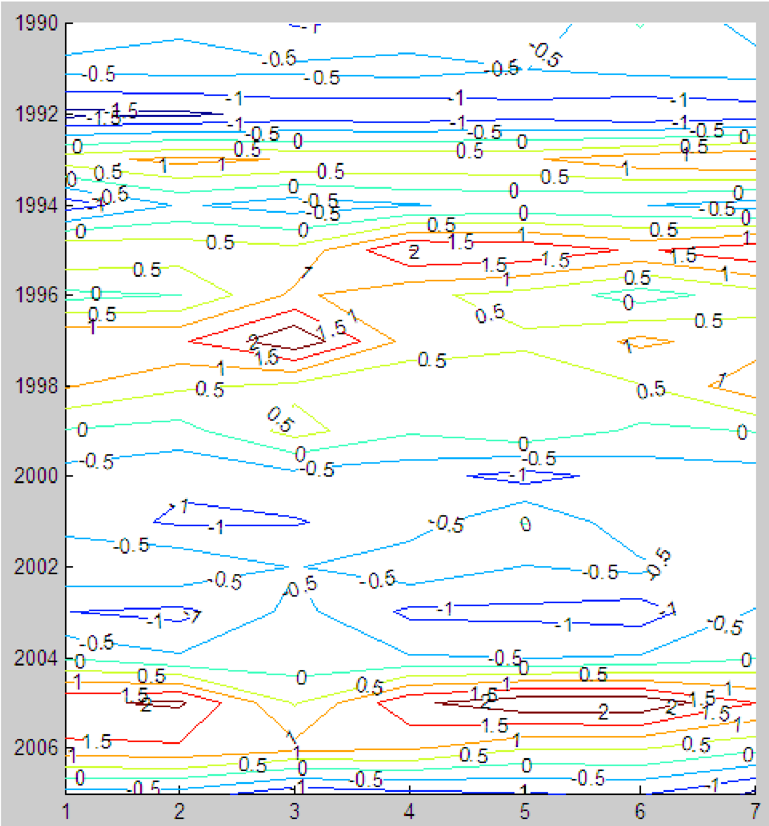




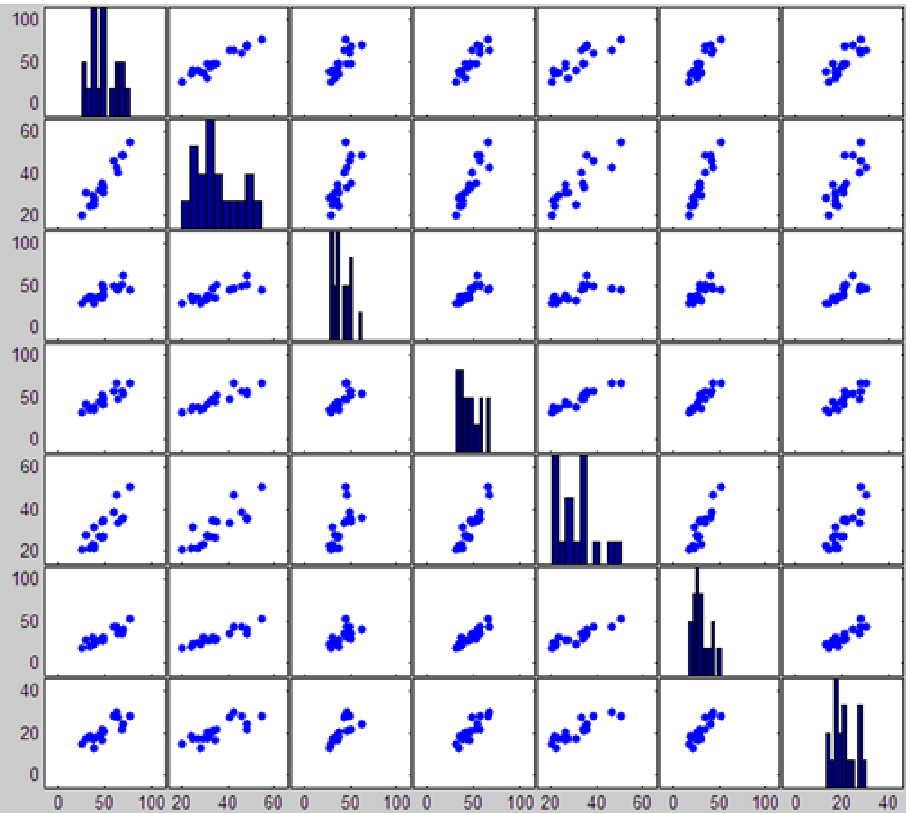
Multivariate Linear Correlations

$$\vec{X}* = \begin{bmatrix} x*_{11} & x*_{12} & \dots & x*_{17} \\ x*_{21} & x*_{22} & \dots & x*_{27} \\ \dots & \dots & \dots & \dots \\ x*_{n1} & x*_{n2} & \dots & x*_{n7} \end{bmatrix}$$

- 7 stations and n years
- Standardized anomalies

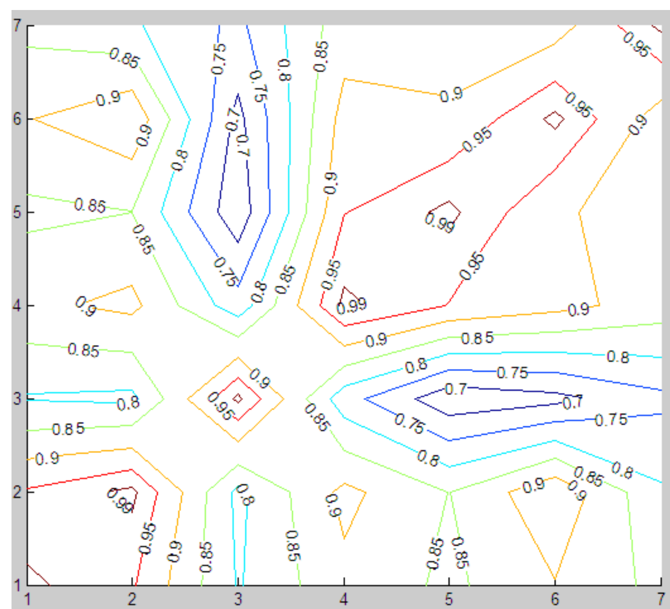


Multivariate Linear Correlations



$$\vec{X}^* = \begin{bmatrix} x^*_{11} & x^*_{12} & \dots & x^*_{17} \\ x^*_{21} & x^*_{22} & \dots & x^*_{27} \\ \dots & \dots & \dots & \dots \\ x^*_{n1} & x^*_{n2} & \dots & x^*_{n7} \end{bmatrix}$$

$$\vec{R} = \vec{X}^{*T} \, \vec{X}^* / n$$



Linear Correlation Coefficients

TABLE 1. Matrix of contemporaneous correlation coefficients (×100) between the time series shown in Figs. 2, 3, 4, 7 and 10. The number of winter seasons used in each correlation, if less than the complete 28 seasons (1951–78), is indicated in parentheses.

	SST Index	SLP Index	200 mb Index	PNA Index	WP Index	Tarawa rainfall	Canton rainfall	Christmas rainfall	Fanning rainfall
SST Index									
SLP Index	−83								
200 mb Index	80	−68							
PNA Index	46	−31	57						
WP Index	67	−57	44	−00					
Tarawa rainfall	78	−71	63	32	65				
Canton rainfall	82 (23)	−65 (23)	61 (23)	40 (23)	65 (23)	60 (23)	(23)		
Christmas rainfall	64 (24)	−49 (24)	57 (24)	40 (24)	48 (24)	55 (24)	82 (19)	(24)	
Fanning rainfall	79 (23)	−60 (23)	69 (23)	38 (23)	70 (23)	72 (23)	85 (18)	84 (23)	(23)

Teleconnections

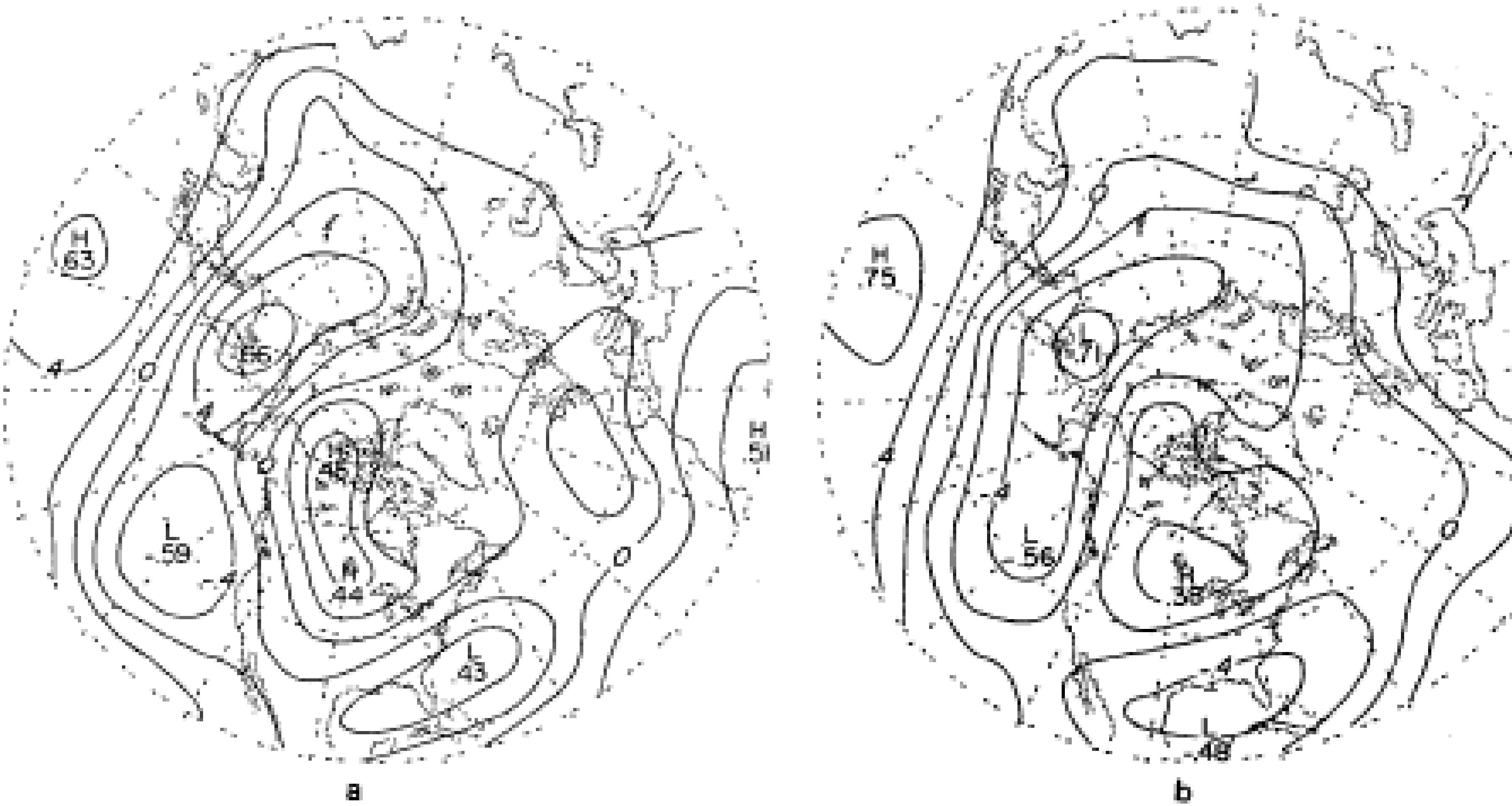


FIG. 9. Correlation coefficients between 700 mb geopotential height at gridpoints poleward of 20°N and (a) the Sea Surface Temperature Index, (b) December–February rainfall at Fanning, (c) our Southern Oscillation Index, and (d) the tropical 200 mb Index. Contour interval 0.2. The locations of the centers of action of the Pacific/North American and West Pacific patterns are denoted, respectively, by dots and open circles in Fig. 9d.

Teleconnections

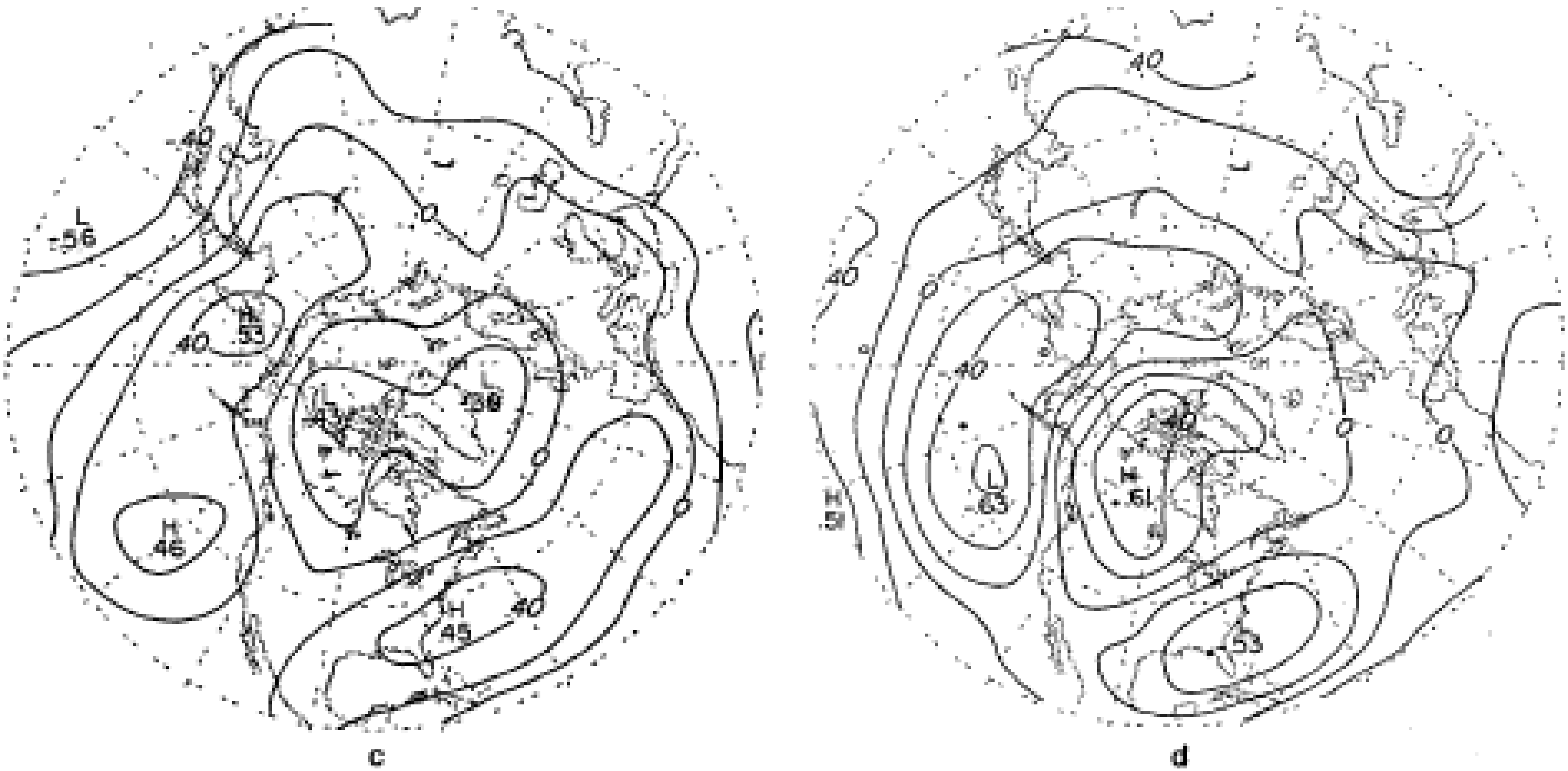
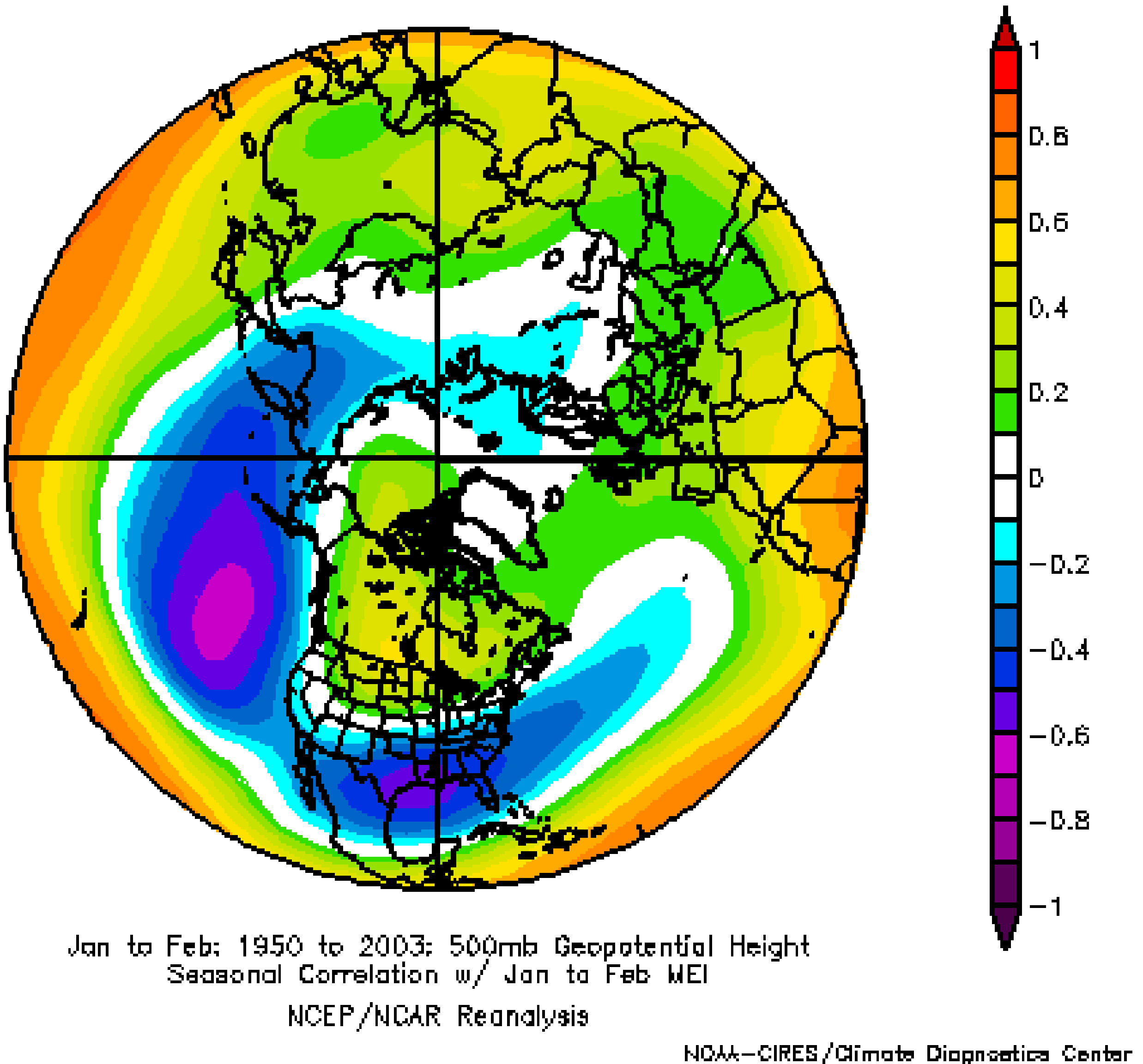


FIG. 9. Correlation coefficients between 700 mb geopotential height at gridpoints poleward of 20°N and (a) the Sea Surface Temperature Index, (b) December–February rainfall at Fanning, (c) our Southern Oscillation Index, and (d) the tropical 200 mb Index. Contour interval 0.2. The locations of the centers of action of the Pacific/North American and West Pacific patterns are denoted, respectively, by dots and open circles in Fig. 9d.

http://www.cdc.noaa.gov/Correlation



Correlating Maps Rather Than Time Series

$$\vec{\hat{X}} = \begin{bmatrix} \hat{x}_{1,1} & \hat{x}_{1,2} & \dots & \hat{x}_{1,18} \\ \hat{x}_{2,1} & \hat{x}_{2,2} & \dots & \hat{x}_{2,18} \\ \dots & \dots & \dots & \dots \\ \hat{x}_{7,1} & \hat{x}_{7,2} & \dots & \hat{x}_{7,18} \end{bmatrix}$$

- Comparing variability in one year over 7 locations to the variability in all of the other n= 18 years

$$\vec{S} = \vec{\hat{X}}*^{T}\ \vec{\hat{X}}*\ \ /7$$

Comparing Forecast Anomaly Maps to Analyses