

# Exploratory Data Analysis

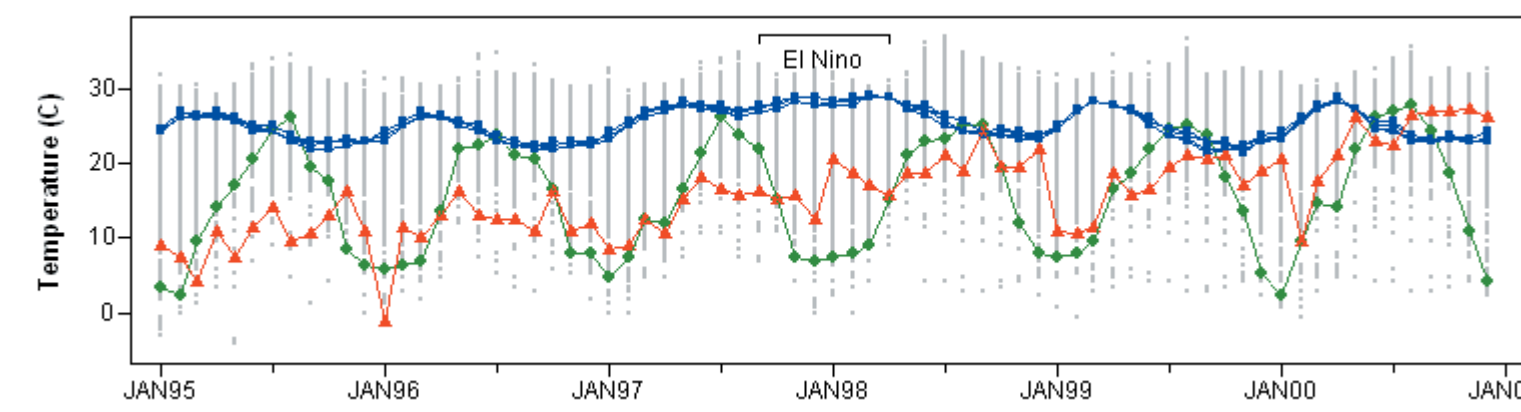
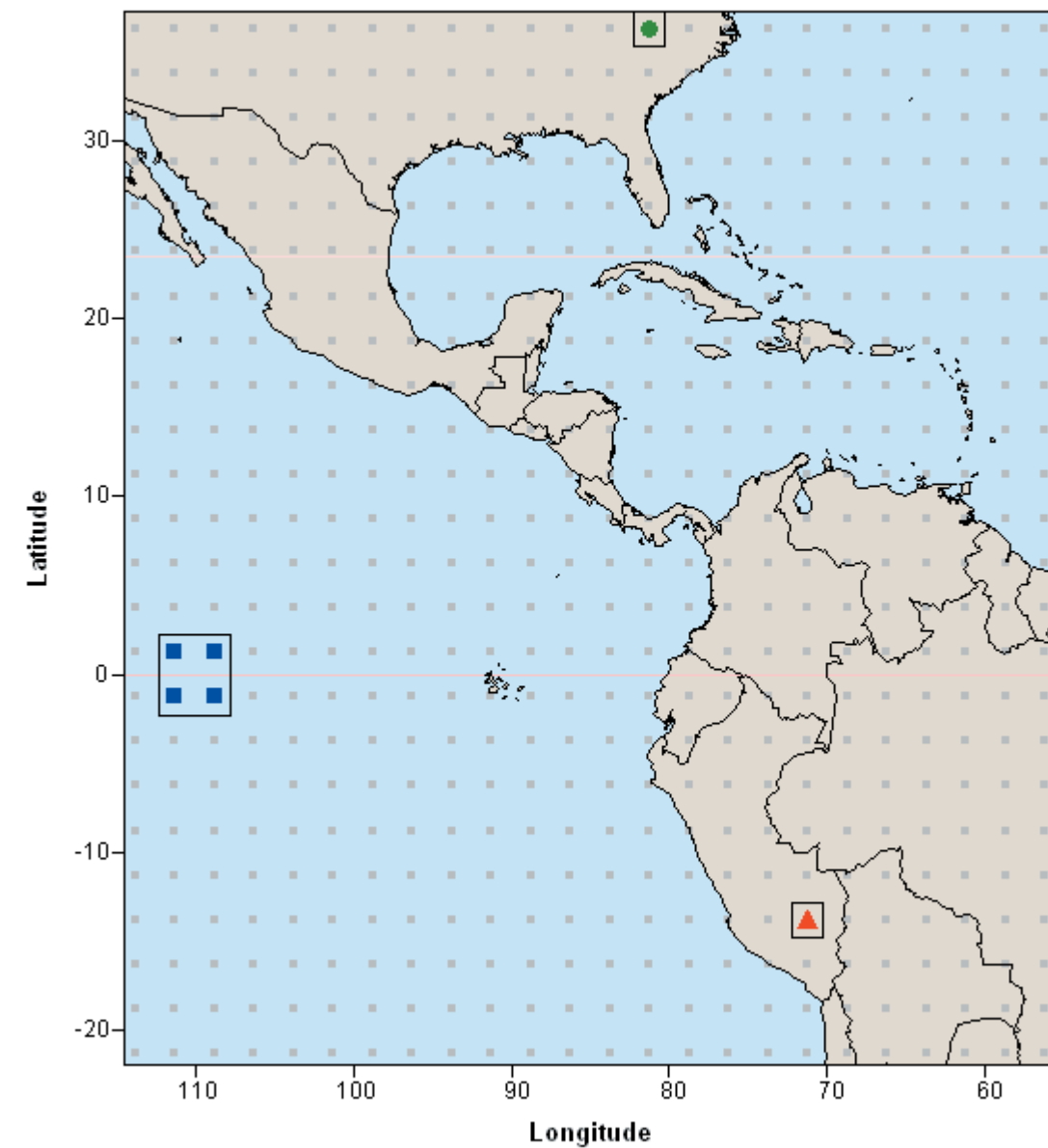
For each location or region, investigate corresponding time series for temperature, ozone, and pressure.

Sample regions:

Eastern United States (green), elevation 306 m

Equatorial Pacific Ocean (blue), elevation 0 m

Peruvian Andes (red), elevation 4270 m

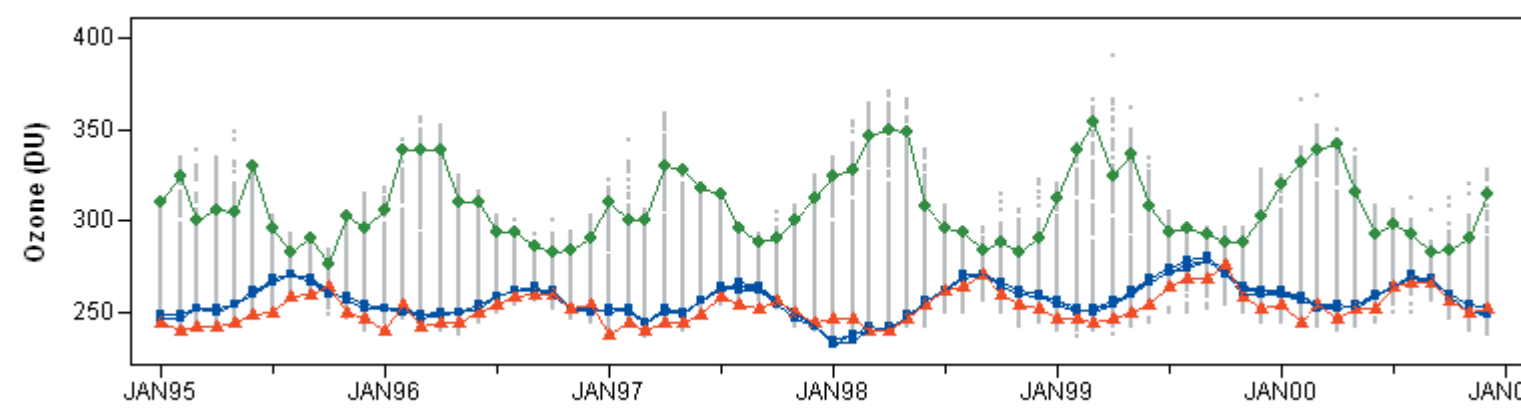


## Near-Surface Air Temperature

**Eastern US:** hot in August; cold in January; large seasonal variation

**Pacific:** hot in April; cold in October; small seasonal variation  
- Homogeneous region: high spatial correlation  
- 1997–98 El Niño caused warm temperatures.

**Andes:** strong upward trend; little seasonality (!)  
Lows in January, during southern summer (!)

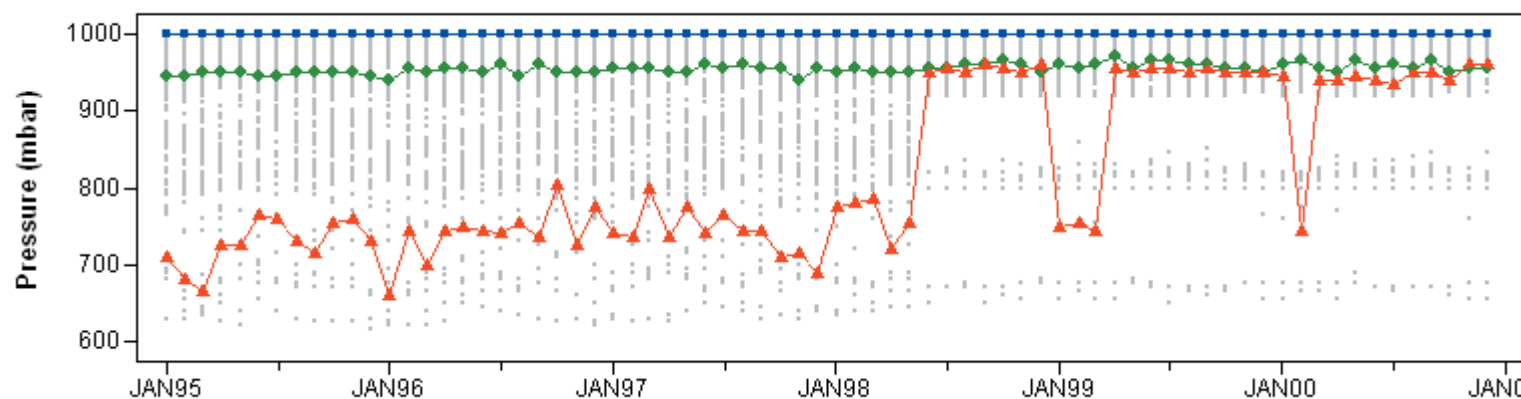


## Ozone

**Eastern US:** high values due to northerly location; peaks in April

**Pacific:** low values; peaks in August

**Andes:** similar to the equatorial Pacific, but phase-shifted



## Pressure

**Eastern US:** fairly constant,  $\mu = 954$  mbar,  $\sigma = 6$  mbar

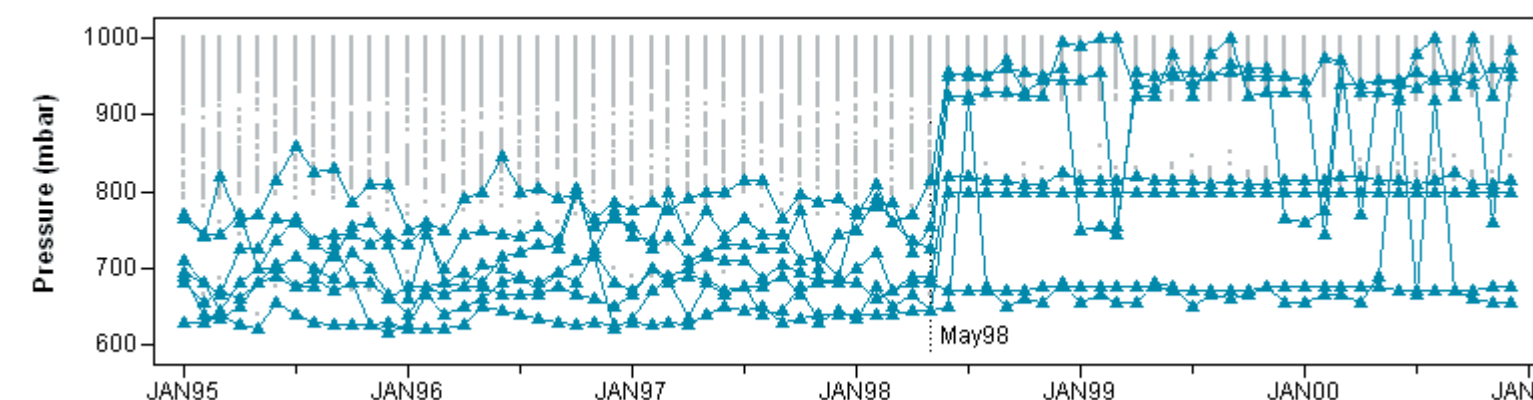
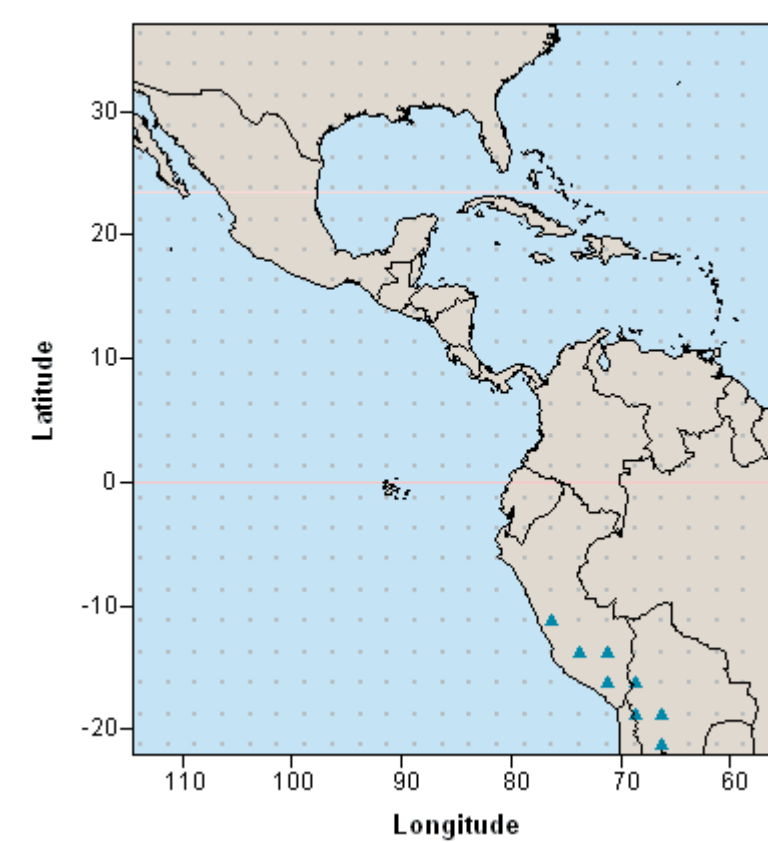
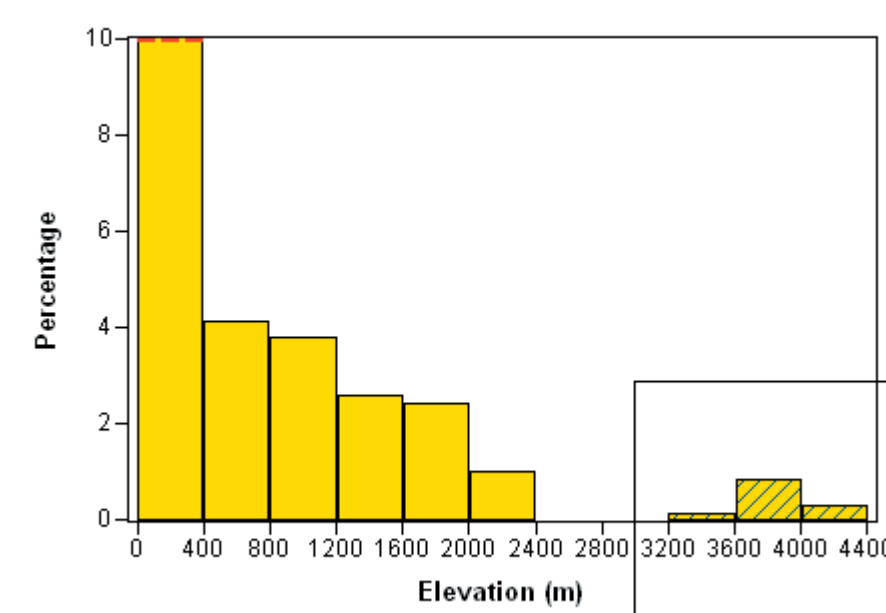
**Pacific:** constant sea-level pressure of 1000 mbar

**Andes:** extremely surprising! Some event happened about May 1998?

# Detecting Outliers: Unusual Pressures

Exploratory data analysis reveals an unusual pressure series in the Peruvian Andes. Is this location unique, or an indicator of a widespread phenomenon? Investigate locations with high elevations.

Select high-elevation positions (> 3000 m)

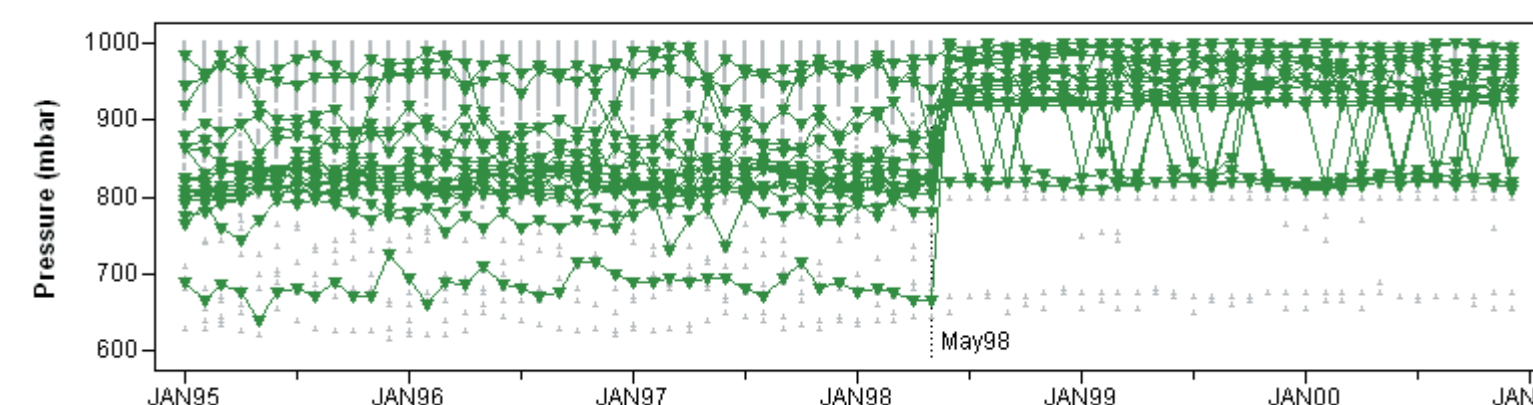
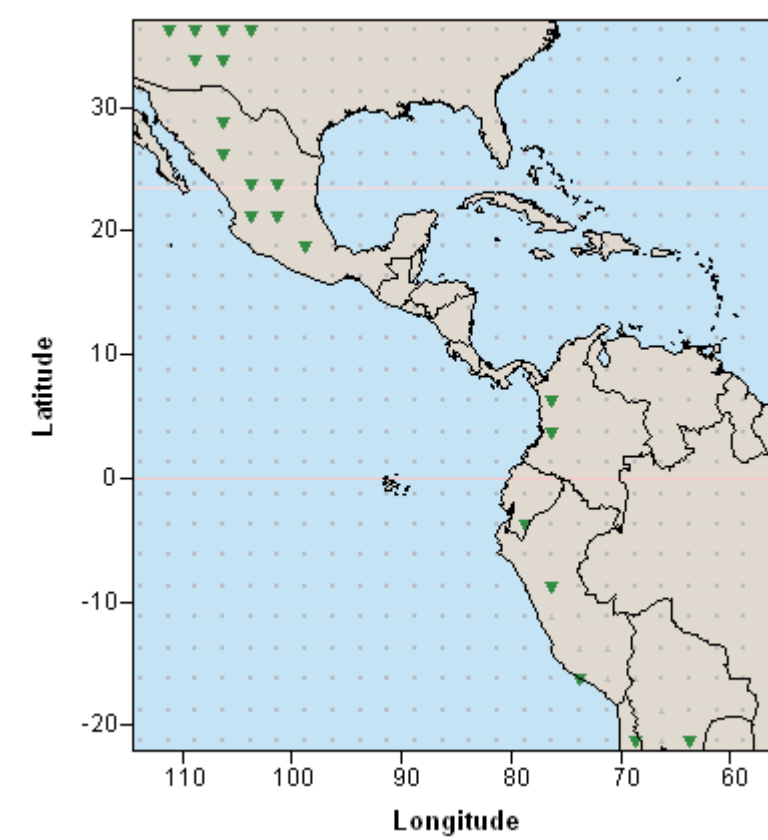
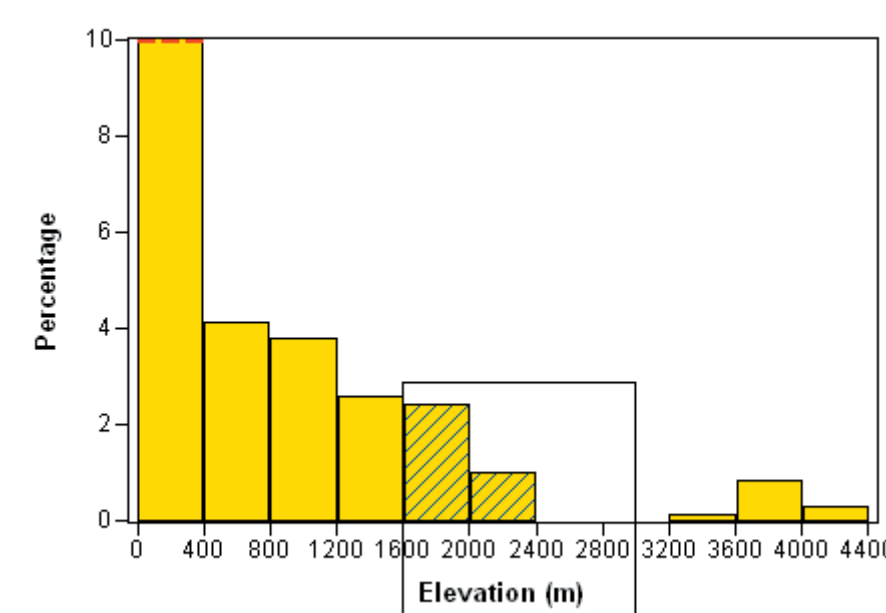


Qualitative change in pressures after May 1998

## Analysis

- **Prior to May 1998**
  - Data record **station pressure?** (Higher elevations have lower station pressure.)
  - **Large seasonal variations** (> 100 mbar) too large to be from natural phenomena
- **After May 1998**
  - Most pressures > **920 mbar**
  - Most positions have **small variations**

Select mid-elevation positions (1600-3000 m)



Anomalies are spatially widespread

## Conjectures

- **Change** in meaning or method of recording pressure after May 1998?
- **Low pressures** (> 920 mbar) after May 1998 due to:
  - Stations not conforming to change?
  - Interpolating onto a 24x24 spatial grid to estimate data?
    - In mountains, nearby positions are at vastly different elevations
    - Averaging pressures between mountains and sea leads to errors