

Outlier Detection for SGPMET Data

YUPING LU

University of Tennessee

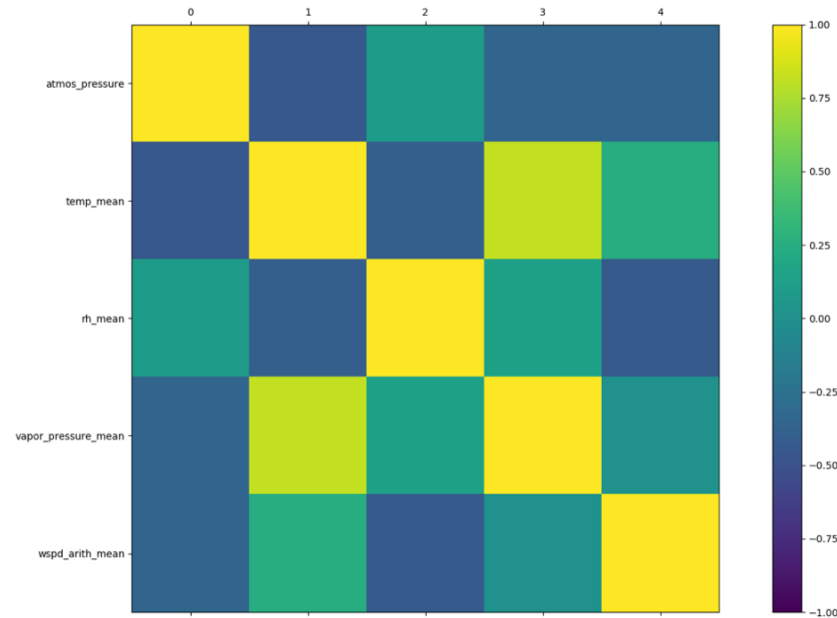
ARM Data Center

Analysis of co-located co-varying data

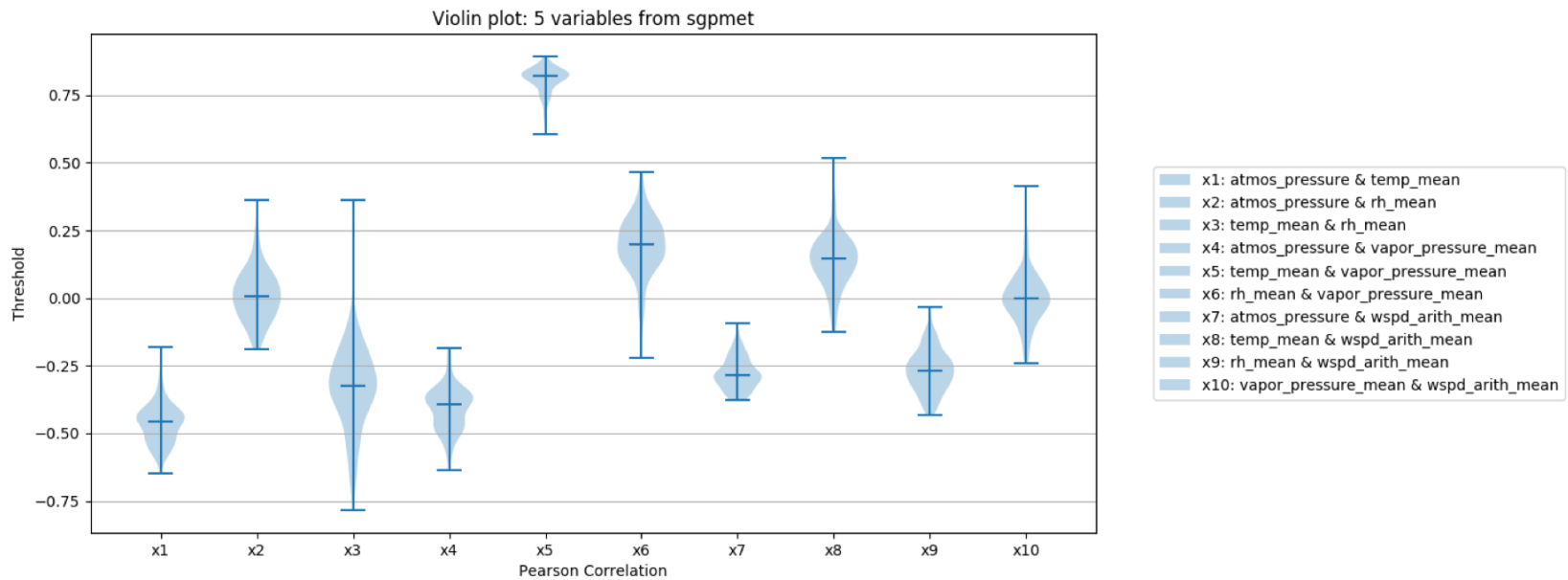
- Multiple instruments across multiple facilities at SGP collect co-located observations for a number of meteorological variables.
- Meteorological variables are expected to show auto- and cross-correlations, strength and patterns of which may exhibit inter and intra-annual variability within and across facilities.
- Analysis of historical data can help mine known and expected multivariate patterns, deviations from which can be used to determine an “outlier”.
- An “outlier” may result due to either an extreme weather event, or an instrument issue – both of which are of interest to us.

Correlations Between Variables

- First, we want to see how well each pair of variables are related. Pearson Correlation are used to measure it.
- Five variables are selected.

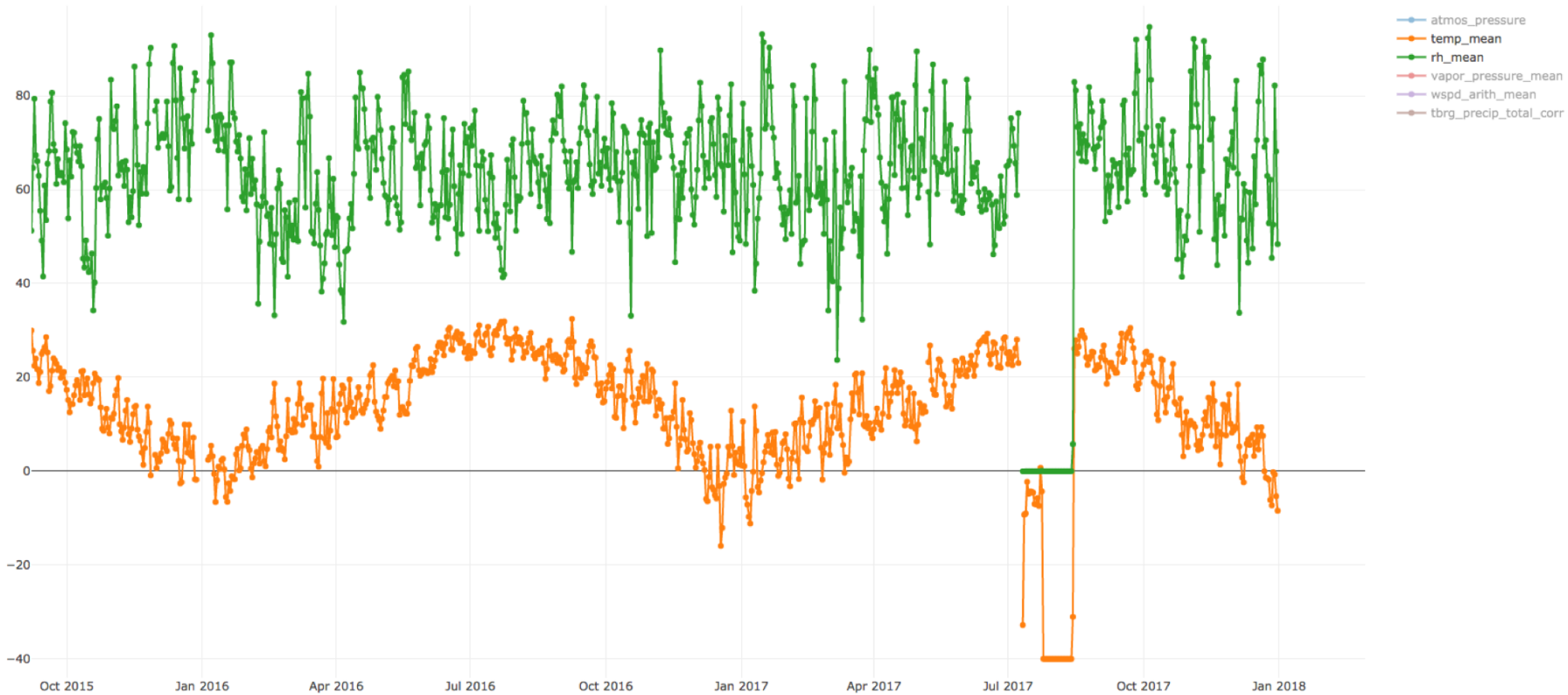


Correlations Between Variables

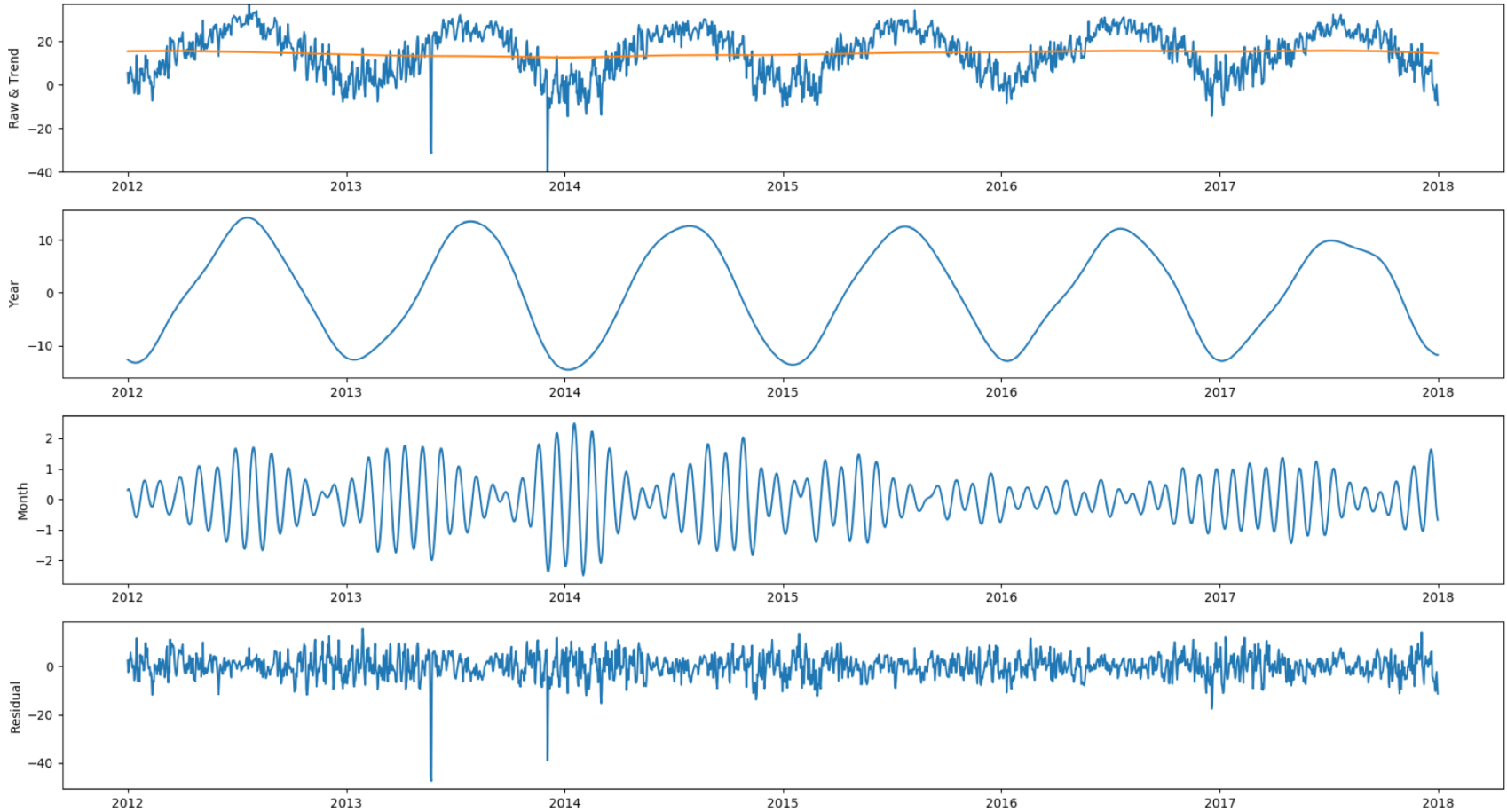


Drawback of Pearson Correlation

E11 1993-2017



Singular Spectrum Analysis (SSA)



E33 2012-2017 temp-mean

Singular Spectrum Analysis (SSA)

Outlier 1
May 2013

Outlier 2
Dec 2013

DQRID : D130528.2

Start Date	Start Time	End Date	End Time	Data Quality Metric
05/21/2013	1704	05/24/2013	2045	Incorrect

Subject:	SGP/MET/E33 - Incorrect temp and wind data
DataStreams:	sgpmetE33.b1
Description:	Cattle chewed temperature and wind sensor instrument cables causing incorrect readings during this time period.
Suggestions:	
Measurements:	sgpmetE33.b1: <ul style="list-style-type: none"> wdir_vec_std wspd_vec_mean wsod arith mean more

DQRID : D131205.1

Start Date	Start Time	End Date	End Time	Data Quality Metric
12/02/2013	1000	12/04/2013	1715	Incorrect

Subject:	SGP/MET/E33 - Temp/RH sensor failure
DataStreams:	sgpmetE33.b1
Description:	Temp/RH sensor failed, causing problems with wind and logger data until 12/3/13 at 21:36 (sensor was disconnected until a replacement was installed).
Suggestions:	
Measurements:	sgpmetE33.b1: <ul style="list-style-type: none"> wdir_vec_std wspd_vec_mean wsod arith mean more

<https://github.com/YupingLu/arm-pearson>
<https://github.com/YupingLu/arm-ssa>