LESS IS MORE: RUMINATIONS ON DATA ANALYSIS AND MINING

Andrés Muñoz-Jaramillo

www.solardynamo.org

Georgia State University
University of California - Berkeley
Stanford University
• Less is oftentimes more. Sophisticated algorithms are very valuable but should be used after careful consideration.

• Humans are optimum pattern recognizing machines, don’t forget they are a valid resource too.

• The more you understand your data and goals, the more effective your mining is going to be.

• Use colors and devote time an energy to defining your color schemes.
ACTIVE REGIONS AND THE SOLAR CYCLE
Sunspots are the visible signature of regions with very strong magnetic field (Hale 1908)
Sunspots are the visible signature of regions with very strong magnetic field

(Hale 1908)

Fields generally appear at the surface in the form of bipolar structures called active regions
Sunspots are the visible signature of regions with very strong magnetic field

(Hale 1908)

Active regions present systematic orientation and inclination

SDO/HMI
They originate from large-scale magnetic belts located inside the Sun

Yeates & Muñoz-Jaramillo (2013)
ARs give shape to the large-scale field
BUILDING AN AR DATABASE
The Bipolar Active Region Detection (BARD) code

1. Detection of positive and negative regions separately.
2. Aggregation into BMRs based on size, flux and distance.
3. Tracking of BMRs across different magnetograms.
Human Supervision Module

A human observer corrects pairing and labeling errors. Only about 5% of active regions needs significant intervention.
Current Status

Two instruments finished (about 8,000 magnetograms processed).
Current Status

Two instruments finished (about 8,000 magnetograms processed).
UNDERSTANDING YOUR DATA AND GOALS
See the panda?
Knowing what you are looking for helps you see!
CONVEYING INFORMATION THROUGH COLOR
Munsell color system
Don’t default to a rainbow color map
Hue best conveys types of data
Lightness best conveys quantity
Magic happens when you use both
Magic happens when you use both
Colorbrewer2.org is your friend
TAKE HOME MESSAGE
• Less is oftentimes more. Sophisticated algorithms are very valuable but should be used after careful consideration.

• Humans are optimum pattern recognizing machines, don’t forget they are a valid resource too.

• The more you understand your data and goals, the more effective your mining is going to be.

• Use colors and devote time an energy to defining your color schemes.