An Analysis of Earth Science Data Analytics Use Cases

Goals of Earth Science Data Analytics

1. To calibrate data
2. To validate data (note it does not have to be via data intercomparison)
3. To derive new analytics tools
4. To perform coarse data preparation (e.g., subsetting, data mining, transformations, recover data)
5. To intercompare data (e.g., any data intercomparison; Could be used to better define validation/quality)
6. To tease out information from data
7. To glean knowledge from data and information
8. To forecast/predict phenomena (e.g., Special kind of conclusion)
9. To derive conclusions (e.g., that do not easily fall into another type)
10. To derive new analytics tools

Conclusions (thus far, with our limited number of use cases):

I. For Earth Science, defining results oriented Data Analytics types are more appropriate for categorizing Earth science data analytics…
   • They accommodate Earth science use cases which are typically results oriented
   • They invite better defined data analytics tools and techniques that address user goals

II. Most ESDA use cases tend to focus on data intercomparison, deriving new products, forecasting/predicting, and deriving conclusions

III. Most use cases were not gleaned from data/information. Perhaps some use cases were not recognized as such (more analysis needed)

IV. Distributed data sources, and data heterogeneity are persistent characteristics

V. Velocity issues are not significant (thus far)

VI. ESDA challenges provide interesting problems for data analytics tool/technique developers to ponder

VII. If any, use case 5 and 6 provides the true Big Data problem