5.7

NOAA OBSERVING SYSTEM INTEGRATED ANALYSIS - II: THE CHARACTERIZATION OF INTERCONNECTIVITY

Louis Cantrell * and David Helms
Technology, Planning and Integration for Observation (TPIO) Division
NOAA/NESDIS, Silver Spring, Maryland

1. INTRODUCTION

NOSIA-II traces hundreds of products and services paths down NOAA's value tree that quickly become complex.

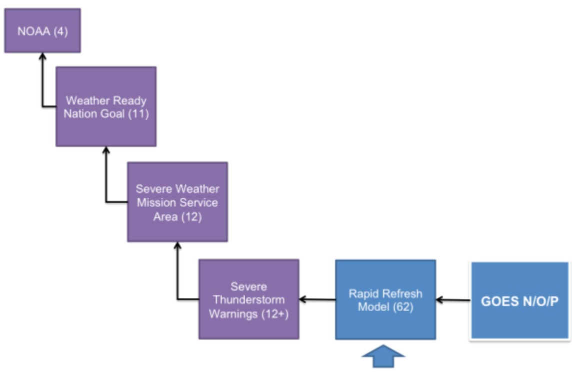


Figure 1

Fig. 1 above illustrates the top end of one of those paths up the value tree. GOES, at the right edge is one of 62 observing systems and other data sources feeding into the Rapid Refresh Model,

- which is used for local thunderstorm prediction,
- which is among more than a dozen data sources feeding into the National Weather Service's Severe Thunderstorm Warning product,
- which is among a dozen Key Surveyed Products feeding into the Severe Weather Mission Service Area,
- which is among 11 Mission Service Areas feeding into the NWS's Weather Ready Nation Goal,
- which is among 4 Goals feeding NOAA's overarching mission.

The mathematical relationships among the purple boxes are quite simple: they are linear, many to one, and they use simple weighting and averaging. The real complexity happens among the tiers in the model represented by these blue boxes.

* *Corresponding author address*: Louis Cantrell, PhD, CCM, Integrated Systems Solutions, Inc., Dunn Loring, VA 22030; email: Louis.Cantrell@ISSMgmt.com

Figure 2

A description of the molecular level of the model serves to illustrate how this complexity originates. First, PALMA™ functionally relates a data source (what we call and option) to a product using these three scores: a status quo score, a swing weight, and an overall satisfaction score. By aggregating data sources one at a time, you can see it grows geometrically.

PALMA™ computes the impact that an option or intermediate product has on the product by using a function called the Interval-Preserving Symmetric Extended Average Power Function (IPSEA-P). It's based on Russell Steinman's interpolation scheme. This function is able to not only determine the individual impact a data source has on the product but also the synergy multiple data sources working together have on the product. PALMA™ functionally relates every molecule of the Value Tree's full DNA. It then searches for the most "benefit-retaining" portfolio of observing systems under increasing budget constraints. PALMA™ then sorts these budget-constrained portfolios. This is illustrated below in what we call the Efficient Frontier.

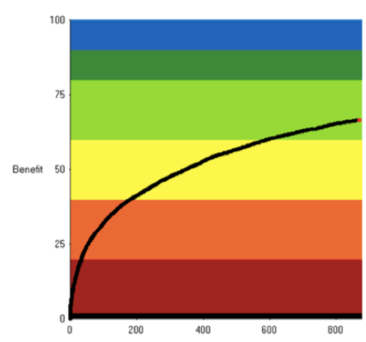


Figure 3

Each point on the line represents an entire portfolio of observing systems: the vertical axis indicates the NOAA-wide benefit of the portfolio, and the horizontal axis indicates the portfolio's cost. The red point on the end of the line represents NOAA's current portfolio, budget, and benefit score with respect to what we surveyed in NOSIA-II. This one red point, in fact every point on the line, compresses an incredible amount of information. We in TPIO are highly motivated to get a handle on this information (or this complexity). There is much to infer from this single red point.

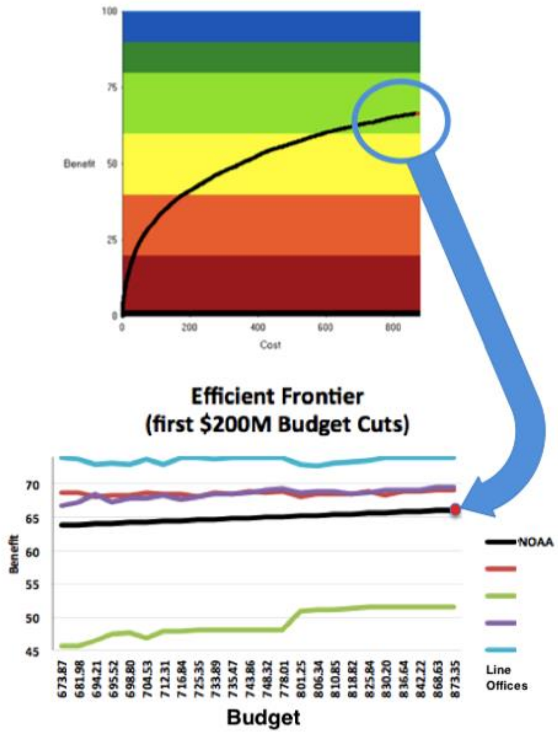


Figure 4

2. GOING DEEPER

The chart in Fig. 4 zooms into the first $200M of observing portfolio budget cuts NOAA might consider and then decomposes the budget into individual Goal benefit scores for each budget profile. This kind of view tells us how much benefit each Goal looses with increasingly budget-constrained portfolios.

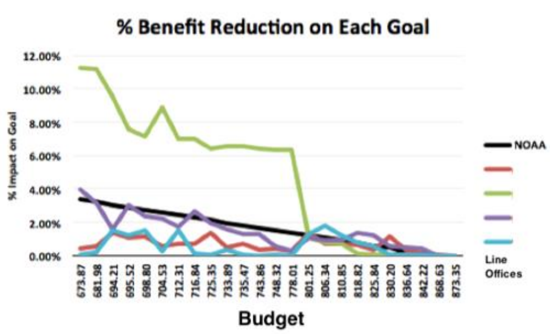


Figure 5

The chart in Fig. 5 examines how much of that budget constraint is a burden to each Goal. In the current model configuration for example we see that there is some give and take managed by each Goal for the first $75M but then bifurcates and one of the Goals begins taking the brunt of the budget cuts. Indicators like this inform us of where the budget may become more sensitive to the decision making process.

Looking backward, the Efficient Frontier has compressed the complexity of the Goal impacts, which is what NOAA's decision maker's need from the TPIO office. They need the problem reduced. There is a great deal more functionality we can derive from the Efficient Frontier in addition to examining budget restrictions. With NOSIA-II, we can also examine

- Budget restrictions,
- Investment redistributions for things like trades among observing systems as a data source for a particular product,
- Investment in new observing systems, or
- Investments in performance enhancements in current observing systems.

3. NOSIA-II'S METRIC SPACE

PALMA's Efficient Frontier serves to compress this complexity into a more manageable decision space but we can't lose sight of the complexity. NOSIA's metric space can provide insight into more than the two dimensions of Budget and Benefit (Value Tree Impact) imply. Other dimensions include Policy, Performance, Inter-Dependence, Economic Impact, as well as multi period (time) and integration with the larger earth observing enterprise.

We've begun examining optimization among budget and benefit, but what about optimization in these other dimensions? Each dimension represents a decision-relevant degree of freedom in NOSIA's information space.

Changes in **Policy** can dramatically affect the relevance of an observing system because policy controls which products and services have the highest priority, which in turn influences the relevance of the observing systems those products and services rely on.

Changes in **Performance** can affect the relevance of an observing system and the budget. NOSIA may be able to tell us if investment dollars for continued maintenance or enhanced capability would have a greater return on investment for one observing system over another.

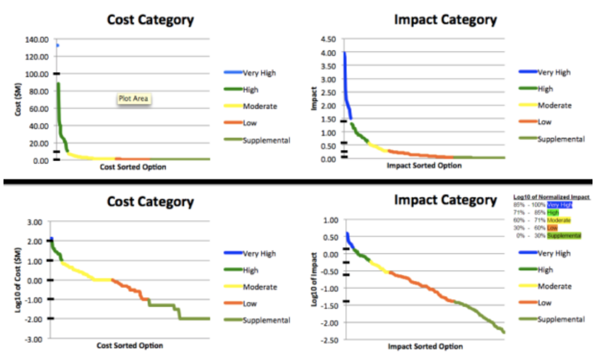
Changes in **Inter-Dependence** can dramatically affect the relevance of an observing system. Those observing systems that support more than one of NOAA's products and services synergistically cumulate impact. On the other hand, isolating the relevance of an observing system to only a few products may give resilience to NOAA as a whole if the removal of an observing system from the portfolio affects only a few minor products.

Changes in **Economic Impact** can dramatically drive Policy or focus NOAA's budget for improvements or maintenance. We hope to soon have the ability to link NOAA's mission service areas, which are modeled in NOSIA, to components of the national economy, and when we do, policy, budget, and performance decisions regarding the portfolio of observing systems will be made to enhance the national economy.

And… these dimensions don't even include variation over time, or synergies among the global network of observing capability portfolios including those funded by other federal agencies such as DoD and DoI, state and local, international, and commercial providers.

3.1 NOSIA-II's Metric Space: Cost and Impact

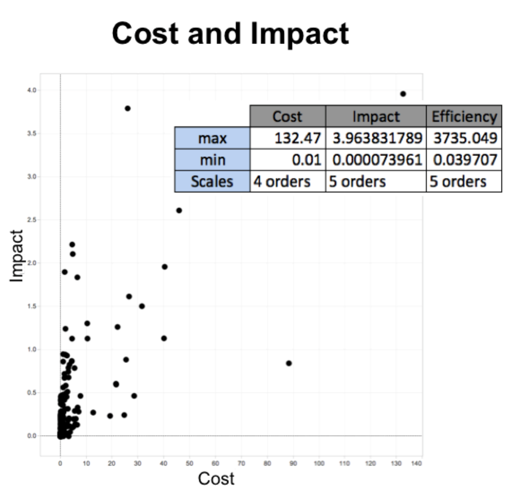
A closer look at the Budget (cost) and Benefit (impact) metric space starts with examining the distributions of cost and impact.



Log Transformed

Figure 6

The charts in Fig. 6 show the distribution of cost and impact for the untransformed data and the log transformed data. We developed a method of categorizing cost and impact categories that would compress some of this information for decision makers. We use the categories you see above. The distributions of the data are logarithmic so we transformed them for correlation analysis.



	Cost	Impact	Efficiency
max	132.47	3.963831789	3735.049
min	0.01	0.000073961	0.039707
Scales	4 orders	5 orders	5 orders

Figure 7

Before beginning a correlation analysis, it's a good idea to look at how the data behave together. Fig. 7 illustrates un-transformed data which bunches up in the lower left of the graph. There isn't much information here but there must be more to see because the cost, impact, and efficiency scores vary by 4 or 5 orders of magnitude so it's necessary to see the log-transform to the data.

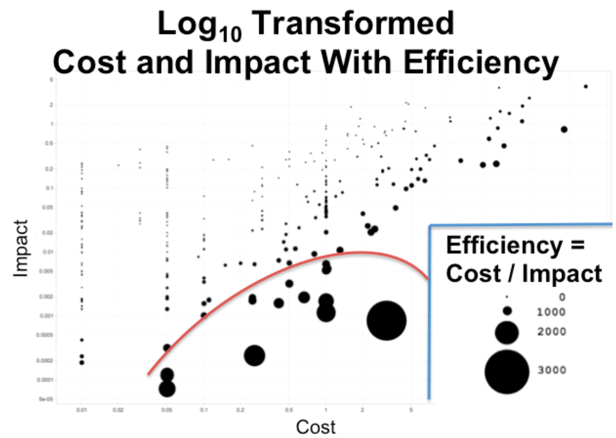


Figure 8

This is how the log-transformed data appear with an addition layer of information. Circle sizes are scaled by an index called "Efficiency", which we compute as the ration of cost to impact. This index represents investment efficiency with larger circles less efficient than smaller circles. It makes it easier to find observing systems that we should give a closer look in a budget constraint situation (under the red curve). There is a visible bias in these data: generally, as cost goes up, so does impact but cost is not strictly related to impact because you can obviously have high impact systems at low cost.

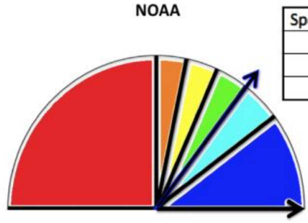
Looking at a scatterplot is not enough, so we need to do some correlation estimates. One method of estimating correlation for a sample is to assign each member of a sample its own dimension. This results in a multi-dimensional vector that represents the sample. I can create such a vector for the cost data and one for the similarly ordered impact data. Correlation then becomes the angle between the two vectors.

One feature of the sorted log-transformed impacts is that it produces a straight line. This implies that the distribution of impacts is not normal. In other words, equal bin sizes for impact ranges would produce a flat histogram. Because of this, we used an approach called Spearman's Rank Correlation, which is a non-parametric (rank-based) method that doesn't depend on the normal distribution assumption or on statistical parameters to find correlation between cost and impact.

	Correlation	Degrees
Negative		
Very Weak	0.0 - 0.2	90.00 - 78.46
Weak	0.2 - 0.4	78.46 - 66.42
Moderate	0.4 - 0.6	66.42 - 53.13
Strong	0.6 - 0.8	53.13 - 36.87
Very Strong	0.8 - 1.0	36.87 - 00.00

Table 1

Spearman's correlations range between +1 to -1 and can be interpreted to be equivalent to the cosine of the angle between the two vectors (high correlation has a small angle between them and low correlation approaches a right angle or worse). To simplify an interpretation of the correlation analysis, we created arbitrary categories for correlation by simply dividing the relevant part of the correlation range into five equal segments.



Spearman Coefficient	0.61
Angle In Degrees	52.52
Moderate	
Strong	

Figure 9

This is the correlation between cost and impact for NOAA. For NOAA, correlation is **"moderate to strong".** Fortunately, investment is not negatively correlated or uncorrelated. There is a positive correlation between impact and cost, but we need to keep in mind that there is uncertainty in our cost data and there is uncertainty in impact scores. We are working now to characterize uncertainties on these correlations.

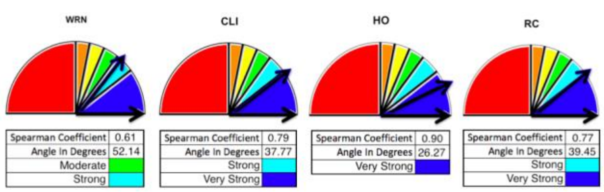


Figure 10

We also extended this analysis to the Goals by extracting the Goal-relevant components of the correlation. Here, correlation among the Goals weights an individual option's cost according to its impact on the goal. This was to eliminate unfairness in carrying the full cost of a system to each Goal whether they depended on it or not. Because investments were done in the past without the benefit of objective impact measurements, the result implies NOAA indeed considered impact in their observing system investment decision-making. The Goal subsets extracted from the NOAA data all have higher correlations than the whole sample. This implies that the Line Offices are individually driving NOAA-wide purchases to benefit their Goals. Moving forward by actually measuring impact gives us a measurement for improvement and success. We believe that a truly optimized portfolio would have a clear "strong" or "very strong" correlation.

Another way of looking at NOSIA's cost and impact relationship is to examine how much information in the relationship is available to a decision maker. The tool we've selected follows from Information Theory, which provides a metric called "Entropy". We define Entropy in the context of decision-making: Entropy is the amount of information the decision maker lacks prior to learning the outcome of a decision; i.e., "I'm going to invest in an observing system, but I don't know how much impact it will have." Entropy is a measure of uncertainty or unpredictability. It is also a measure of information content. Claude Shannon developed a mathematical approach to measuring it with the unit of measure called the "bit" (binary digit), which is in wide use today.

NOSIA Objective:
Make Investments Equilibrated with Impact

ENTROPY (RANDOM)	ENTROPY (NOSIA-II)	ENTROPY (PERFECT)
4.91	**3.72**	**2.32**
Upper Bound	NOSIA-II	Lower Bound

Entropy is the amount of information the decision maker lacks prior to learning the outcome of a decision.

Without the aid of NOSIA-II,
NOAA has successfully managed
46 % of the available information space.

NOSIA-II will help lower the uncertainty of portfolio decisions.

Figure 11

It is our goal to inform NOAA's leadership to help them make the most beneficial decision. We can use entropy to describe the breadth of the decision context if we can adequately state what the context is. In this case, I describe the context of the decision as "How can I make investments in my observing systems equilibrated with its impact?" In other words, if you were a decision maker, you want to invest most energetically in observing systems that have high impacts on NOAA and you want to invest less if the impacts are not very substantial. Entropy can serve as a NOAA-wide objective measurement to characterize our observing system acquisition decision space.

In this figure, we've computed Shannon's Entropy for NOAA's distribution of cost and impact. This metric helps me define the range of entropy that can potentially exist in our portfolio of observing systems. It has a maximum value (based on an assumption of random correlation) and a minimum value (based on an assumption of all impact and cost ranking categories highly correlated).

From this measurement, we've found that each impact and cost category relationship has a range of 4.91 bits to 2.32 bits. Our measurement of the amount of entropy existing in the NOSIA-II data implies that even without the benefit of an impact analysis, NOAA's leadership has successfully managed about half of the uncertainty. We think that with NOSIA's information now available, we can close the gap even further.

3.2 NOSIA-II's Metric Space: Policy, Impact, and Inter-dependence

What about other dimension's in NOSIA's metric space such as Policy and Inter-Dependence (inter-connectivity)? This is a 3-D scatterplot of observing system rank (Policy), impact (Benefit), and interconnectivity (Inter-Dependence). Scatterplots in 3-D are a little more difficult to visually inspect, but we can infer some important features in this data set.

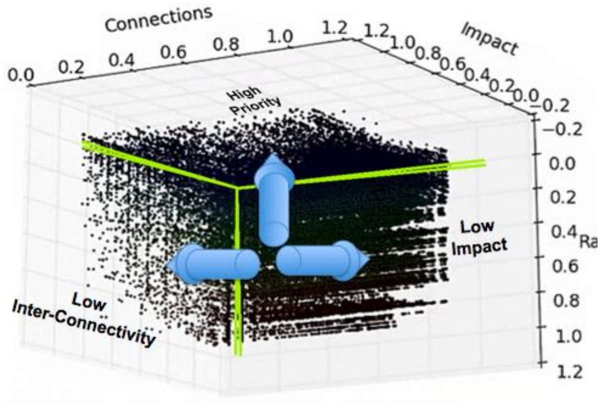


Figure 12

Here is a view from inside the box. The top face of the box (the high ranking observing systems) shows a great accumulation of high impact and high interconnectivity. The data are attracted to the low inter-connectivity face and to the low impact face. These are biases in the data. However, as ranking goes up toward the top face of the box (the high ranking observing systems), the data show a bias toward both high impact and high interconnectivity.

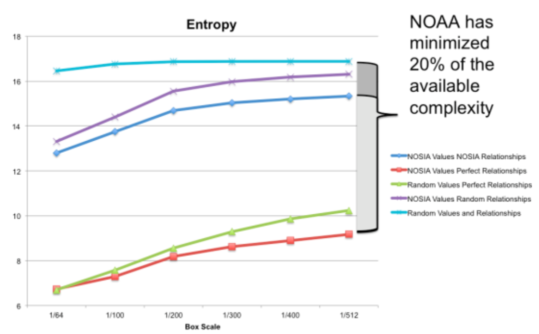


Figure 13

This is a graph of NOSIA's entropy measurements for these dimensions (Policy, Benefit, and Inter-Dependence) as it varies with sampling scale. In these dimensions, NOAA has successfully navigated only 20% of the complexity of the value tree.

The interconnectivity of NOAA's products and services provides more avenues for transmitting the benefit of observations throughout the value tree than a "stovepipe" observing system/product branch of the tree would. A stovepipe observing system/product branch can have very high impact in our value tree (and appears budgetarily secure) as long as the product that depends on it has high policy ranking. If for instance a specific climate or air quality product relied on only a few niche observing systems, and if climate and air quality are politically sensitive, then those observing systems would be politically vulnerable.

Increasing inter-dependence throughout the value tree has the effect of increasing an observing system's impact when you remove it from NOAA's observing system portfolio. Inter-dependence is a double-edged sword: it synergizes products and services, but it also makes them more vulnerable to impact from a single observing system failure. Less inter-dependence and more duplication is unacceptable in a budget-constrained environment. If NOAA's decision makers articulate this objective, then we can measure the entropy or the complexity of the model to answer that objective.

References:

10th Symposium on Societal Applications: Policy, Research and Practice, AMS

5.1. Evolution of NOAA's Observing System Investment Assessment Process. Eric Miller, NOAA/NESDIS, Silver Spring, MD; and H. W. Tileston III.

5.2. Overview of the NOAA Observing System Integrated Analysis (NOSIA) II Assessment Capability. David Helms, NOAA/NESDIS, Silver Spring, MD; and M. Yapur

5.3. Constructing NOAA's Value based Tree: Representing Relationships between NOAA's Organization, Services, and Observing Systems. Martin Yapur, NOAA/NESDIS, Silver Spring, MD; and D. Helms and E. Miller.

5.4. NOAA Observing System Integrated Analysis (NOSIA) Methodology, Martin Yapur, NOAA/NESDIS, Silver Spring, MD; and D. Helms and E. Miller.

5.5. Improving Understanding and Use of NOAA Observing Systems on Societal Benefit. Louis Cantrell, Integrated Systems Solutions, Inc.; and M. Austin.

5.6. OSEs, OSSEs and the NOSIA Capability Aaron Pratt, NOAA/NESDIS, Silver Spring, MD; and D. Helms, L. Cantrell, R. C. Reining, and V. Ries.