# How Theories of Induction Can Fruitfully Constrain Measurements of Scientific Performance


Slobodan Perović
Department of Philosophy
University of Begrade


- Bochum 2017 -

- Inductive analysis (IA): how to optimize scientific conclusion acquisition (hypothesis-driven)
- Operational analysis (OA): how to optimize scientific groups to make best conclusion acquisition (data-driven)

# Operational analysis of scientific networks

- Identifying optimal ways of organizing scientific networks of agents (e.g. types of connection, structure of networks, their size,…)
- Immediate applicability to science-policy
- Its scope is limited but precise

- Data-driven
  - Publication rates, citations, and its impact in various domains, size of teams…
- Simulations and decision theory
  - they require interpretation

Citation metrics:

- A pillar of science policy-making that affects organization of scientific pursuit and its outcomes across scientific fields
- A tool for measuring efficiency and optimizing
- Its downsides are rather typical of social science research, where research outcomes may have a negative impact (murky metrics and applications; loose analysis)

Example:

- Studies suggested publications in most reputable journals are the way for scientists and institutes to promote their career in an efficient way
- This led to implementation of the practice that in turn damaged the overall efficiency of research

# Inductive analysis as a constraint on operational analysis

- IA as a minimal assurance of methodological coherence of pursuit – justifying OA
- The citation metrics should effectively measure the efficiency of the inductive process.
- Seeking even basic coherence of a pursuit requires a model. Testing a sophisticated pursuit (e.g. HEP, phylogenetics) requires sophisticated inductive models
- Very selective choice of cases to illustrate or assess inductive models

## Optimal inductive procedures in science: Machine (Formal) Learning Theory

"[a]n important learning theory project is therefore to determine whether a proposed methodological norm prevents inquiry from being as reliable as it could have been." (Kelly, Schulte and Juhl, 247)

- *Computable agents* rather than ideal epistemological agents
- Generates *general principles and inference rules*
- Hypothetical, not categorical epistemic norms, which however, can be squared with the patterns of reasoning in concrete cases

# **Case 1:** Operational analysis of mega-experiments in high energy physics (HEP)

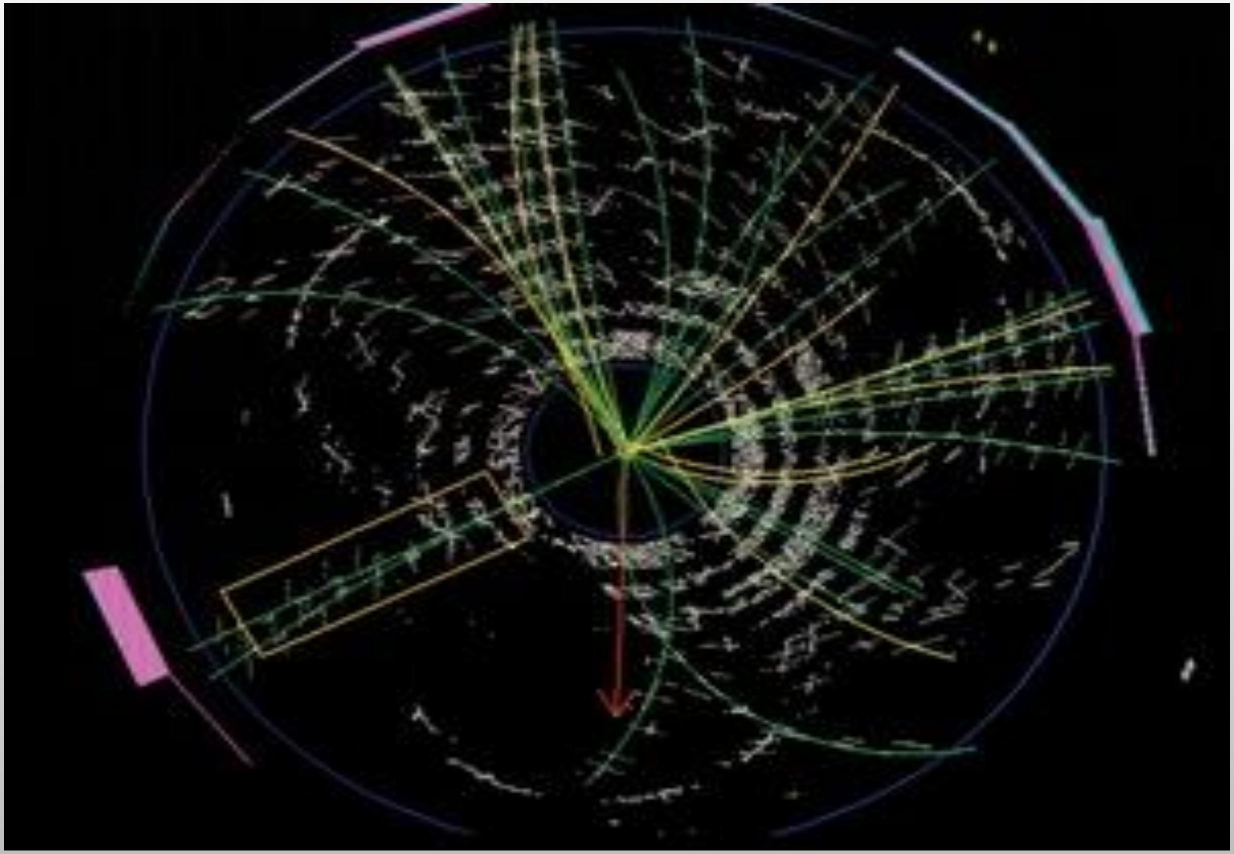
1. The convergence on the results is quick, stable and relevant (over long periods of time).
2. The experiments are either unique or almost unique (no oversight in tracking impact of results by publications/citations)
3. The bulk of citations are in journals within the specialized peer group (experts cite it)

- Judgements of the peers are reliably tracked by the publication rates and citation rates (convergence time indicator)

- The goal: provide an independent argument that the actual quick convergence on the results – the actual pursuit - is not spurious, merely an artefact of a peculiar traits of the scientific network in HEP

# **Case1**: Inductive analysis of HEP experiments

- Principle: the experimental particle physicists constrain their derivations from data (hypotheses) with the conservation principle (conservation of the momentum, energy, charge…)
- Rules: Real-world practical derivation procedures, as well as IA computations make recommendations based on inferences bounded by the principle as the base-line
- Method: physicists opt for the closest fit with the data



Machine learning theory (MLT) models of the pursuit:

- "an exhaustive search in the space of quark models for baryons followed by the mesons reveals the standard quark model stands out nearly uniquely as the simplest, when the constraints of complementary pairs is imposed." (Perez and Zytkov, 2109)

- The proof of restrictive selection of the rules:

"Under pure induction (i.e. without additional assumptions), more than one selection rule and quantum property are never needed to distinguish any set of allowed reactions from any set of prohibited ones." (Perez and Erdmann, 172)

1. <u>The physicists project the theory</u>: they expect that the theory will be valid for some future expected evidence – and the methods of selection based on conservation principles warrant this expectation -> fast and reliable convergence

2. <u>The constraint on the selection rules is strong</u>: assuming conservation laws the number of selection laws that are not redundant turns out to be small

- The inductive method based on learning theory produces the same procedure that is in use by physicists for particular particles
- A quick and stable convergence on the experimental results based on warranted (strict selection rules) projecting of the theory

**IA (FLT/MLT) test** (the conditions for evaluating whether a pursuit is inductively coherent):

1. Computable models matching the actual pursuit (over a set of data) developed
2. A core of various models: a base-line inductive principle and restrictive inference rules
3. The models warrant/imply fast and stable convergence by successfully computing to data set with restrictive rules based on the postulated principle

**OA test**:

1. Fast and stable convergence on the results
2. Reflected in publication and citation rates
3. Passes MLT test

- The citation metrics effectively measures the efficiency of the inductive process.

# **Case 1**: optimal organization via citation metrics in HEP experiments

- A three-part assessment (CERN) of the performance of CERN with respect to other HEP laboratories
- Performance of individual accelerators of the laboratory
- Offered various quantified results with the ambitious normative intentions of improving performance of existing HEP laboratories as a whole

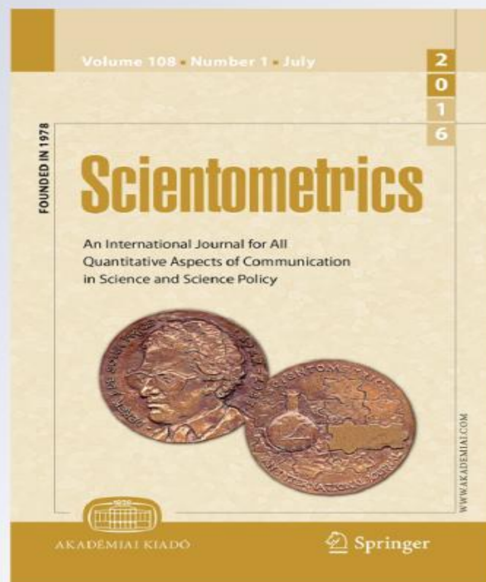
*Optimal research team composition: data envelopment analysis of Fermilab experiments*

**Slobodan Perović, Sandro Radovanović, Vlasta Sikimić & Andrea Berber**

**Scientometrics**
An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science Policy

ISSN 0138-9130
Volume 108
Number 1

Scientometrics (2016) 108:83-111
DOI 10.1007/s11192-016-1947-9




Springer

- A study conducted on the data from the Fermi National Laboratory (Perović, Sikimić, Radovanović, Berber 2016)
- Based on the actual data from 27 experiments with the goal of computing their efficiencies in relation to the team sizes (DEA)
- The inefficient experiments in the quantitative study among the largest teams in the group (stalled at the level of realization or the data analysis could not be completed). The efficient among smallest.

- *The most efficient teams were small compared to the rest:* ***they are better at the inductive process (knowledge production)***
  - *It is the bulk of what the experimenters do/an organizing structure*

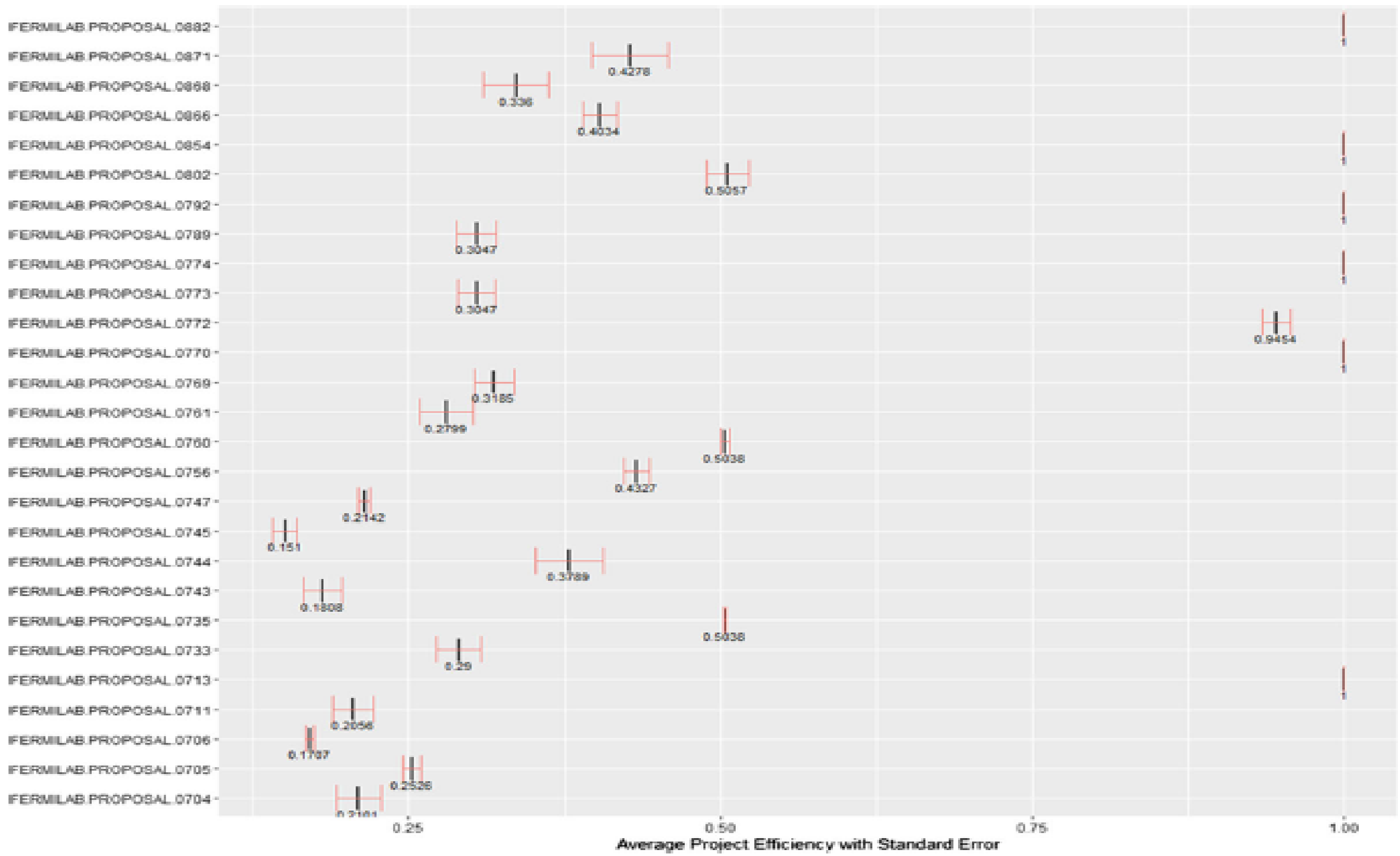



**Fig. 1** Jackknife DEA results

# **Case 2:** Phylogeny research

- In biology, the convergence on results is usually not fast and reliable
- Find a pursuit that passes the test

Phylogenetics:

- Principle: the parsimony principle - all other things being equal, the best hypothesis on evolutionary relationship is the one that requires the fewest evolutionary changes
- Restrictive rules
- Method: sequence similarities between genes

- Biologists themselves are using machine learning as a tool of their analysis

- Calculating the similarities (MLT):
    1. The numerical scores are assigned to differences in the nucleotide, or amino acid sequence and changes in the gene structure.
    2. The scores are determined based on the frequencies of these differences.
    3. The greater the frequency (in large data sets) the smaller the number assigned.

- The calculation gives an optimal tree which has the smallest number of differences between the branches.

# AAA-AAB:1, AAA-BBA:2, and AAB-BBA:3

| | AAA | AAB | BBA |
|---|---|---|---|
| AAA | | 1 | 2 |
| AAB | 1 | | 3 |
| BBA | 2 | 3 | |

- OA in phylogenetics: one can test efficiency of various distributions of tasks across laboratories via citation metrics
- Does it pass OA test? Extract suitable data.

- Contrast with the problematic case:
  - Why was OA applied to that particular set of data and why was citation metrics used?
  - Was it applied to a coherent pursuit, as it was identified by an IA?
  - Was the pursuit examined, or intertwined domains of various pursuits identified at all?
- If not, then it should not have been applied to start with. The base-line constrain on the OA prevents spurious analysis and unwanted side-effects

- Cross-pursuits do not pass MLT test
- Exploratory science does not pass MLT test
- Mature and streamlined pursuits suitable

# A wider test convergence?

- Convergence of different inductive analyses on reliability of a specific research pursuit would even more strongly argue in favour of a justified application of the operational analysis of the scientific network in which the pursuit is embedded

- Statistical models across disciplines; DEA and IA tested results within “individual” pursuits