Designing What-if Analysis: Towards a Methodology

Matteo Golfarelli, Stefano Rizzi, Andrea Proli

DEIS – University of Bologna

Agenda:

1. Introduction
2. What-if enabled tools
3. Some lessons learnt
4. Methodological sketch
5. The case study
6. Conclusions and open issues

What-if analysis

- Decision makers need to evaluate beforehand the impact of a strategic or tactical move
  - *"How would my profits change if I ran a 3×2 promotion for one week on some product on sale?"*
    - Modeling the behavior of the customers
    - Modeling the side effects on similar product sales in the same week
    - Modeling the side effects on the product sales in the next weeks

What-if analysis can be described as a data-intensive simulation whose goal is to inspect the behavior of a complex system under some given hypotheses (called ***scenarios***)

- N.B. ***What-if analysis ≠ Forecasting***

What-if enabled tools

- A tool for what-if analysis should at least have the following features:
  - Allow interactive update of data.
  - Allow decision makers to hierarchically aggregate and disaggregate predictions and see the impact of modifications at every level.
  - Natively support a core set of techniques for expressing and building simulation models, plus a language for further extending the modeling capabilities.
  - Support decision makers in formulating hypothetical scenarios on the model.
  - Support statistical techniques for evaluating how reliable and accurate the predictions are.

What-if enabled tools

	Applix TM1	Powersim	QlikView	SAP BPS	SAS Forecast S.
Expression of simulation models	Limited to disaggregation rules	System dynamic	Scripting	✓	Based on function library and language
Support formulating hypothetical scenarios				✓	✓
Interactive update of data	✓	✓	✓	✓	✓
Multidimensional analysis of the results	✓	only if integrated with a DW	✓	✓	✓
Support analysis of the accuracy of the predictions		✓			✓
Data model	MOLAP (in-memory)	System dynamic	MOLAP (in-memory)	relational	relational

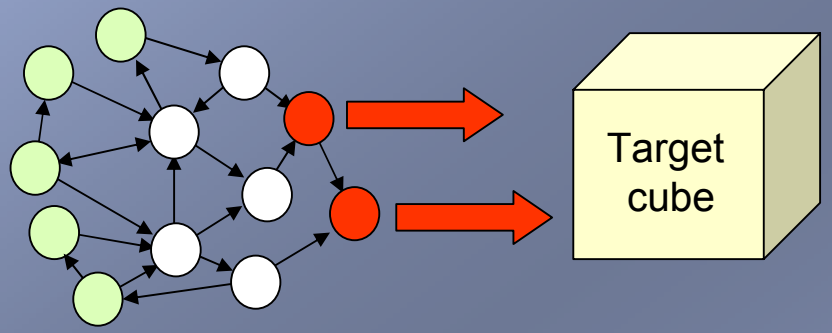
Some lessons learnt I

- In the context of BI, the multidimensional model should be taken as the reference
  - it is widely recognized to be the most suitable model for supporting information analysis;
  - it is inherently capable of representing historical trends;
  - it natively supports fruition of information at different abstraction levels;
  - what-if analysis is typically made on top of a DW system, where data are multidimensional.

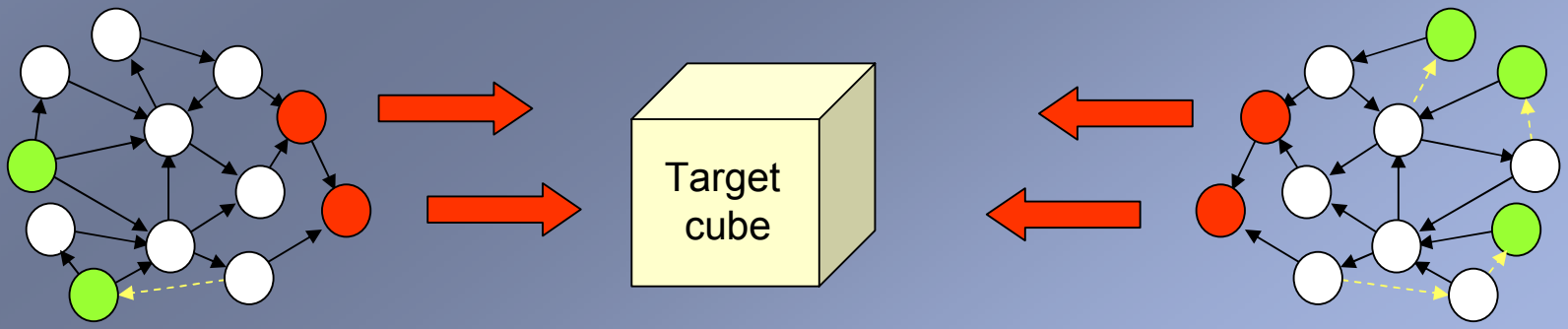
- We will call *target cube* the multidimensional schema that will host the *prediction*

Some lessons learnt II

- A what-if application is centered on a *simulation model*, that describes one or more alternative ways to populate the target cube with a prediction.



- Each alternative corresponds to a *class of scenarios* required by the users.

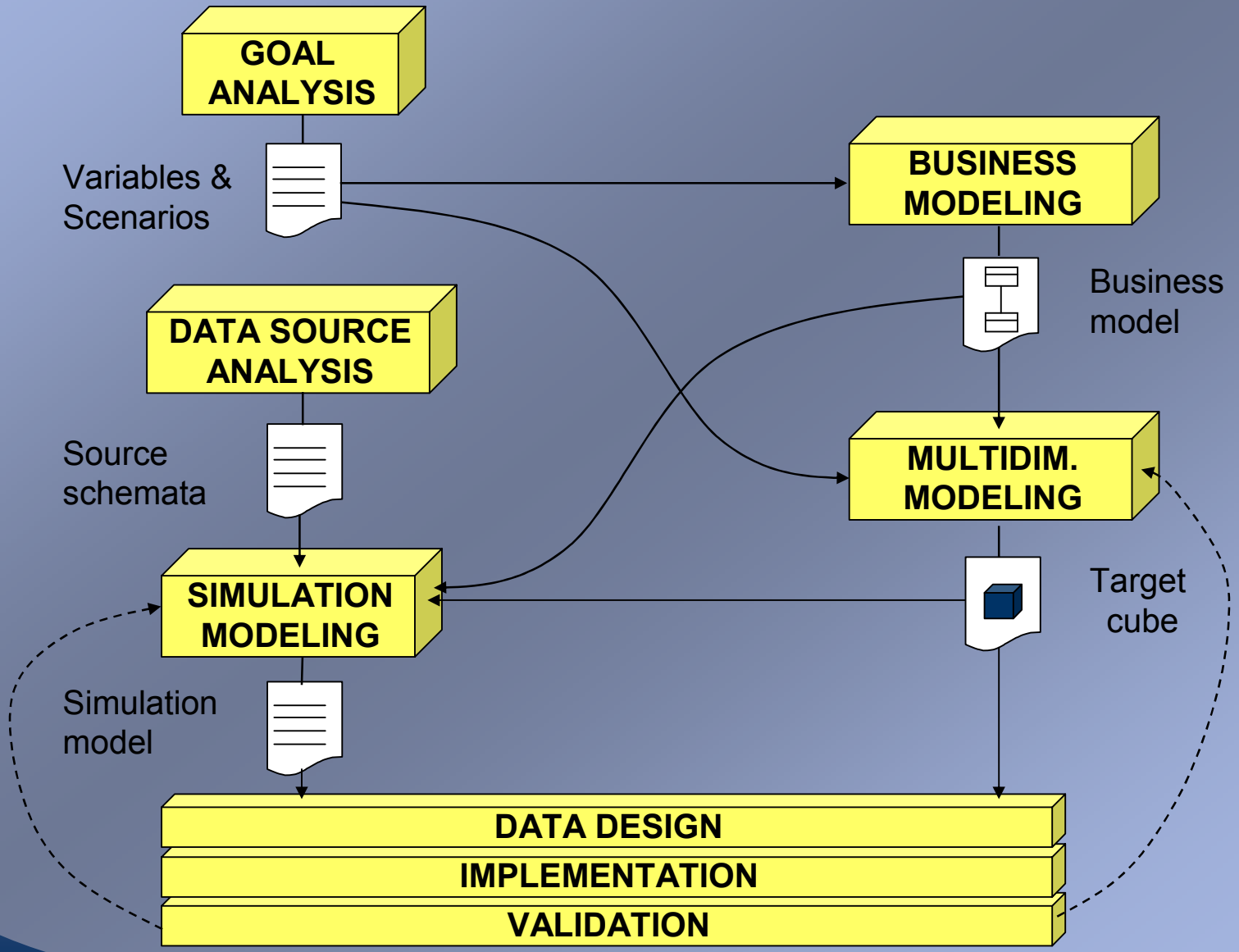


Some lessons learnt III

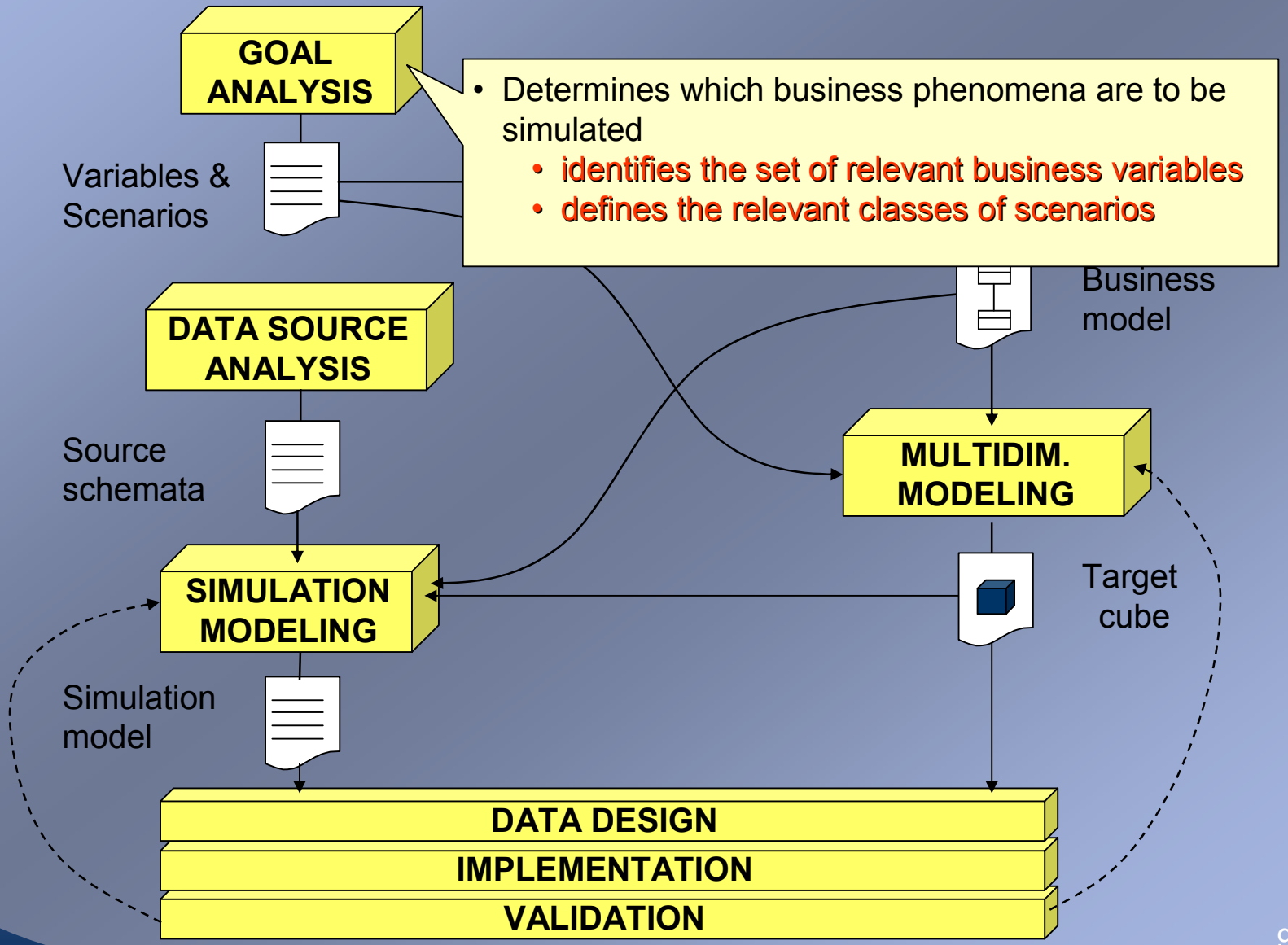
- Reliability of the simulation model strictly depends on the trade-off between precision and complexity.
  - Too precise model ⇒ high simulation costs
  - Rough simulation model ⇒ unreliable results

- Iterative approach to reach the correct trade-off

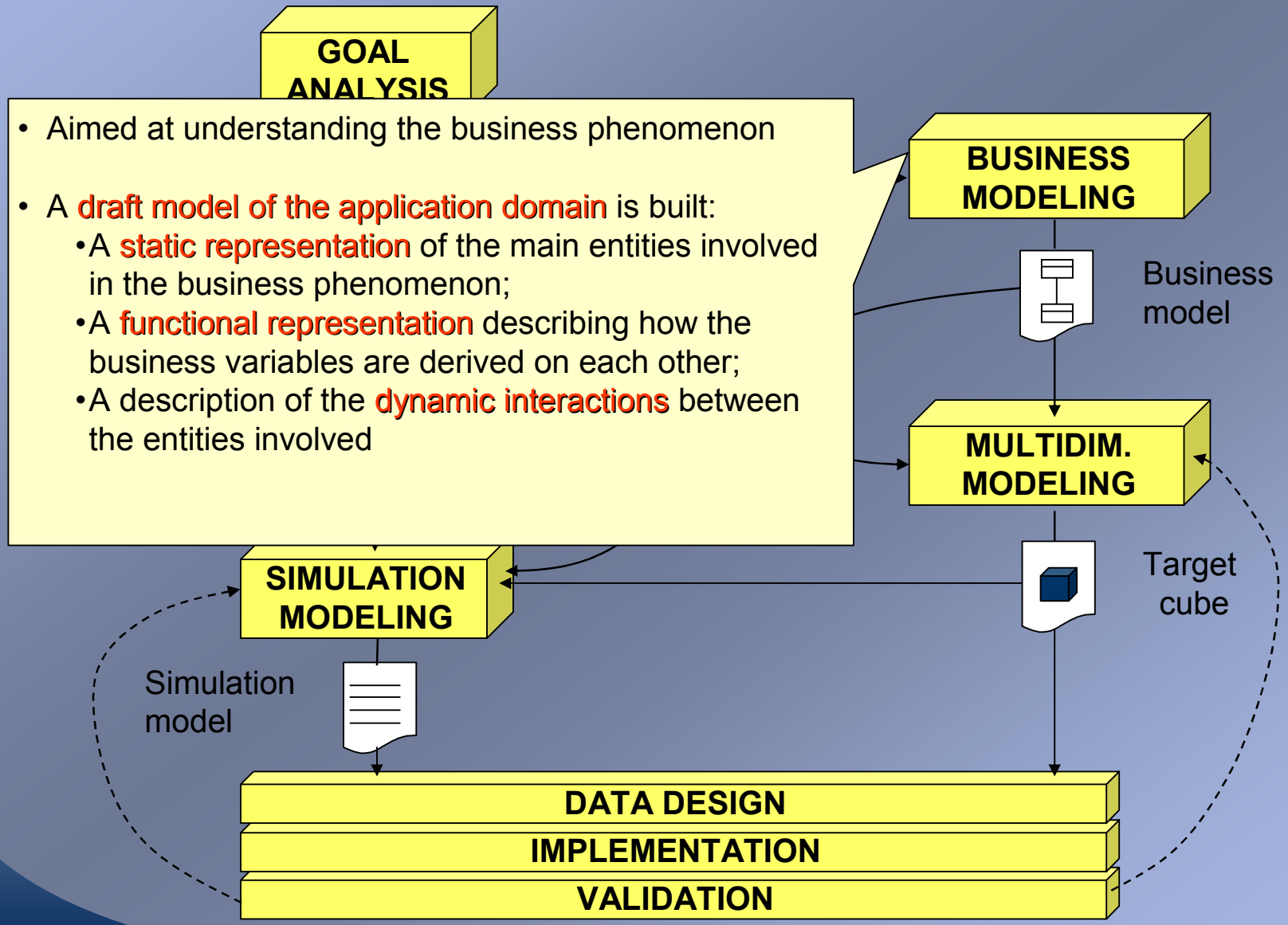
A methodological sketch



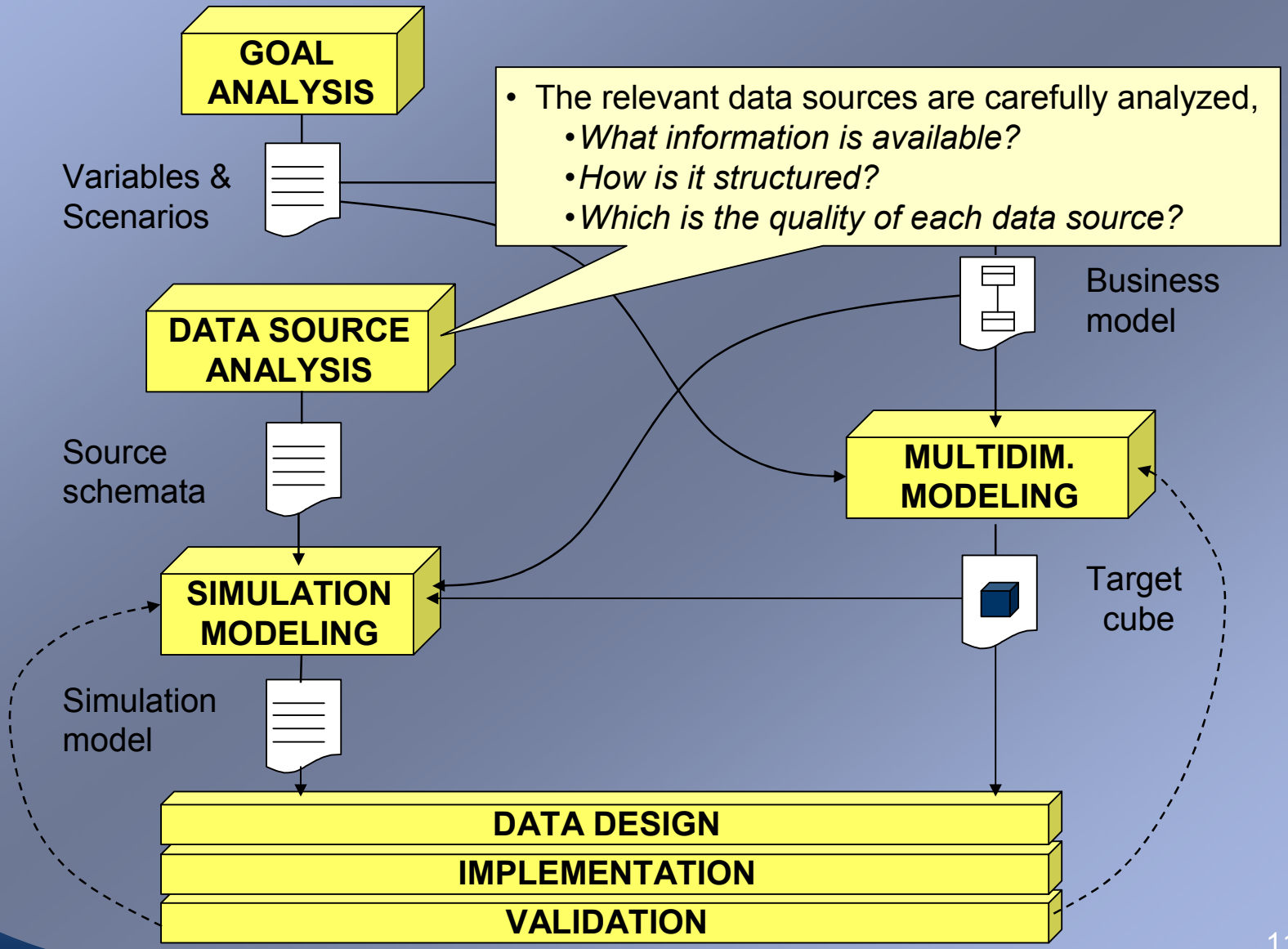
A methodological sketch



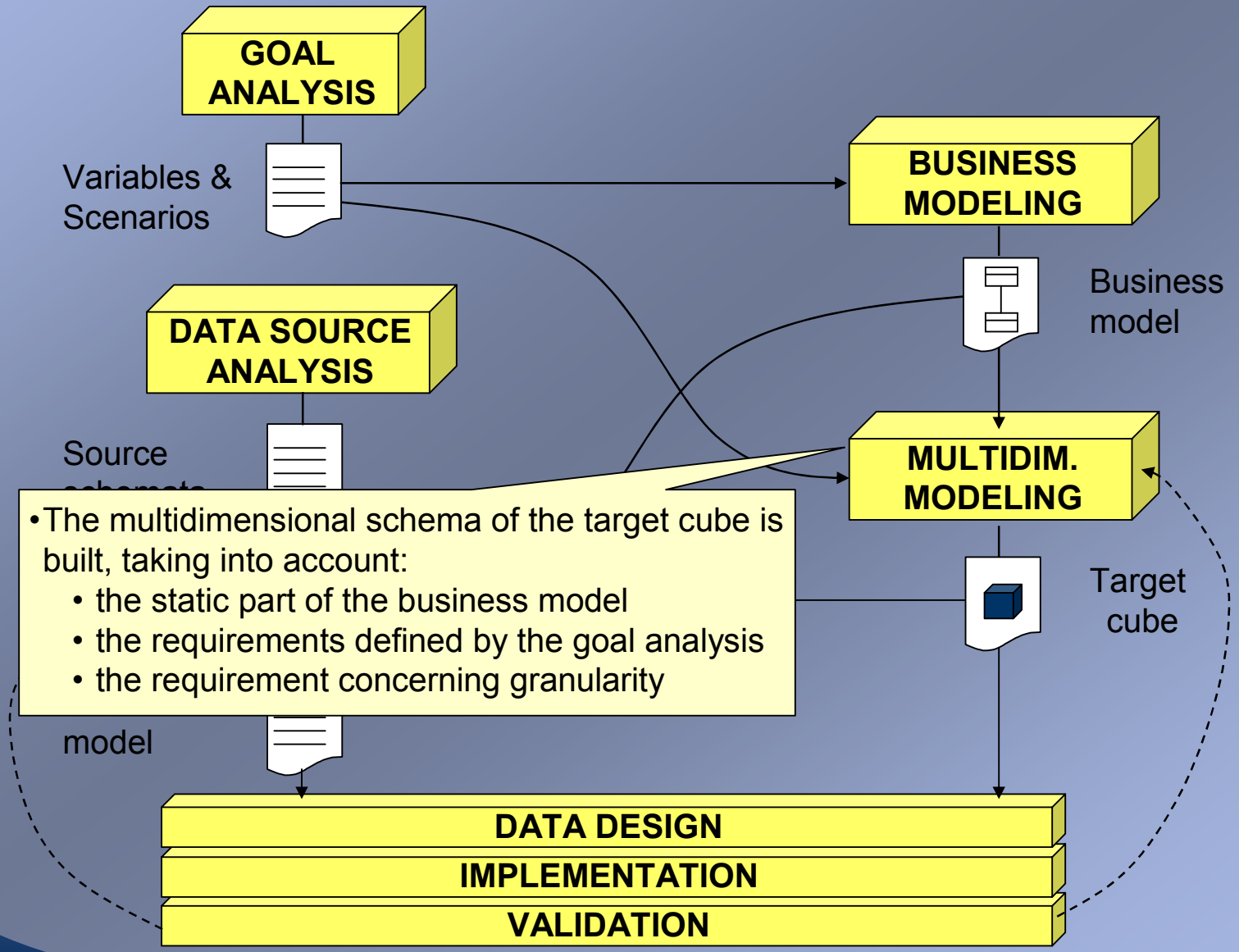
A methodological sketch



A methodological sketch

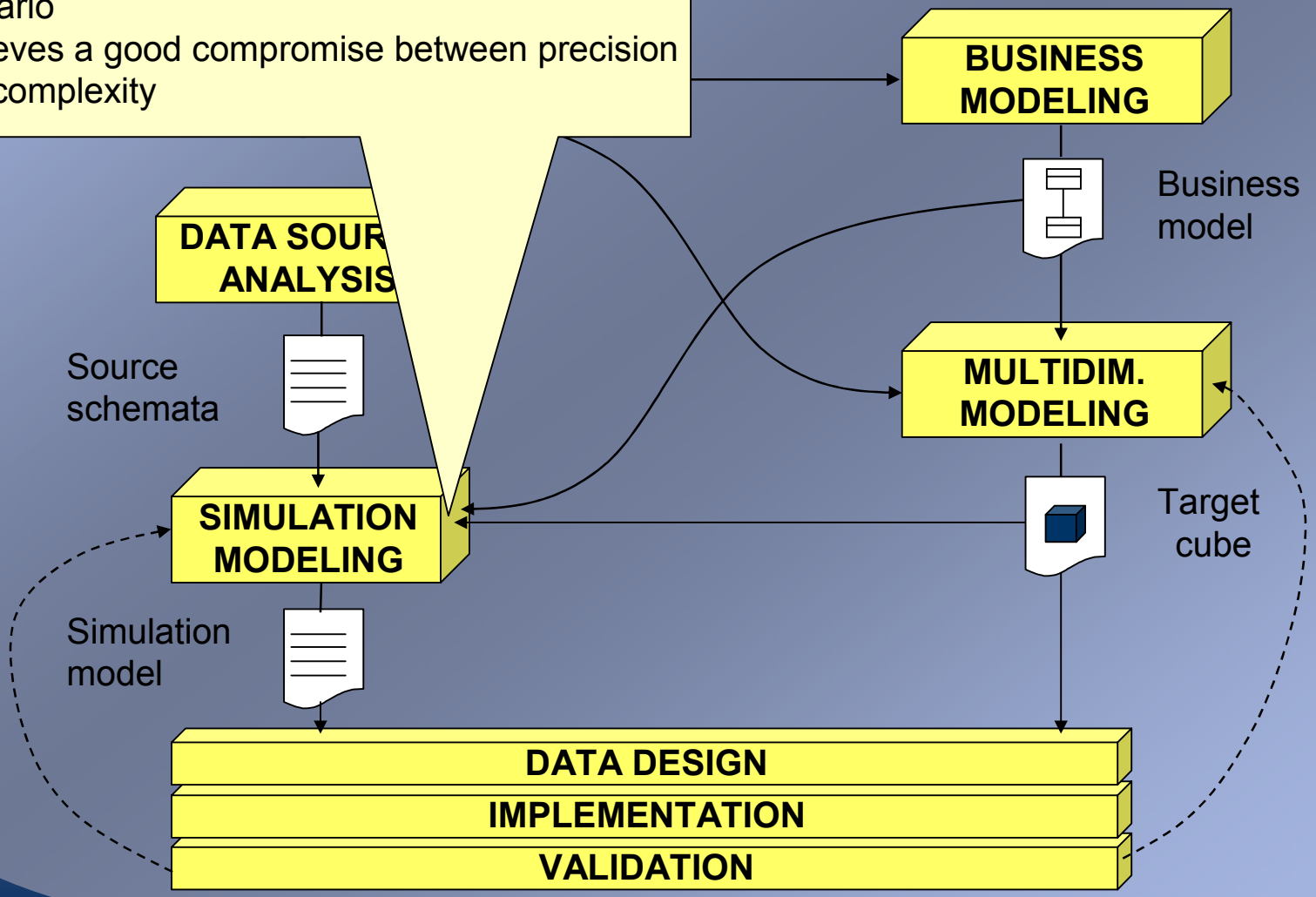


A methodological sketch

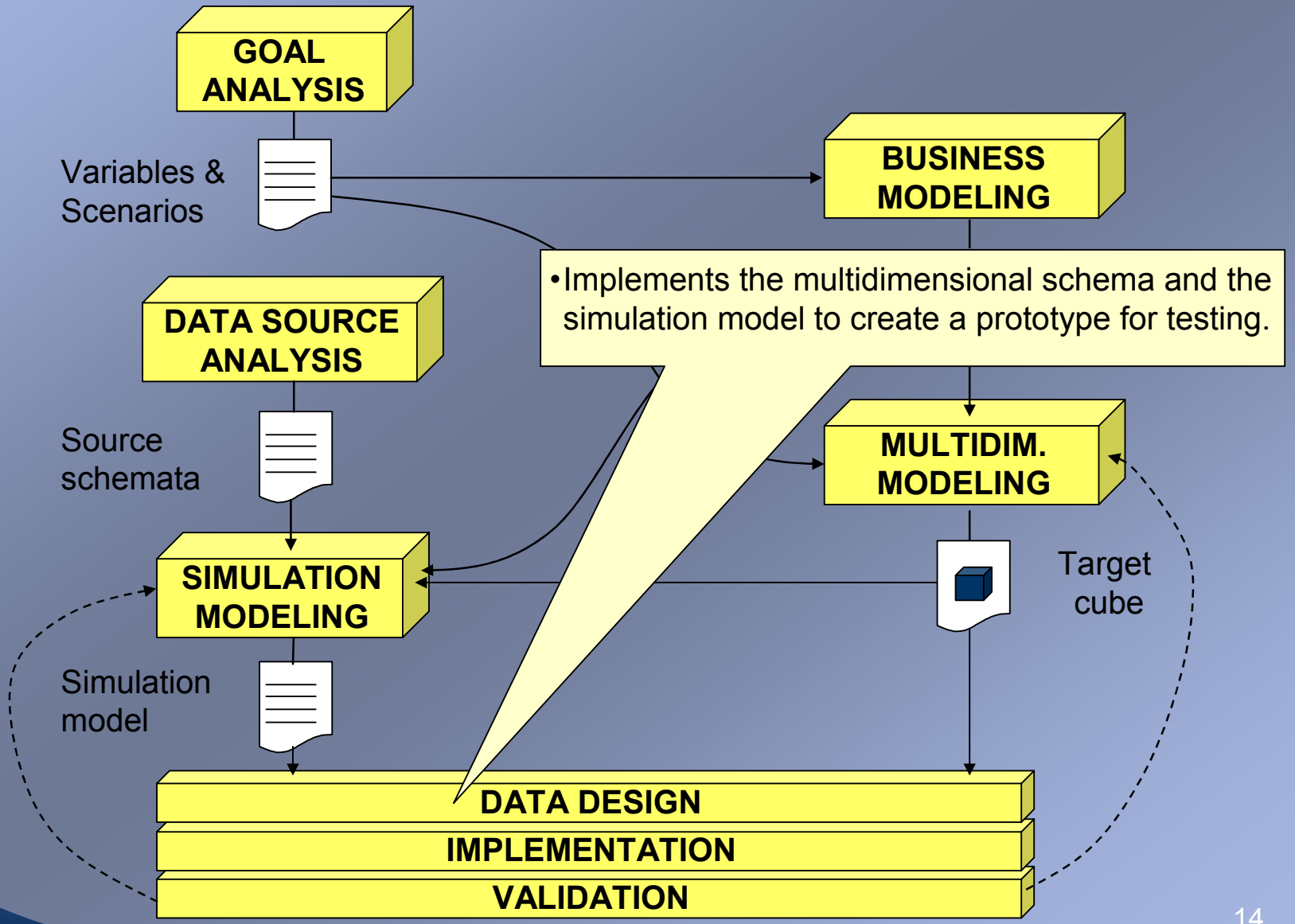


A methodological sketch

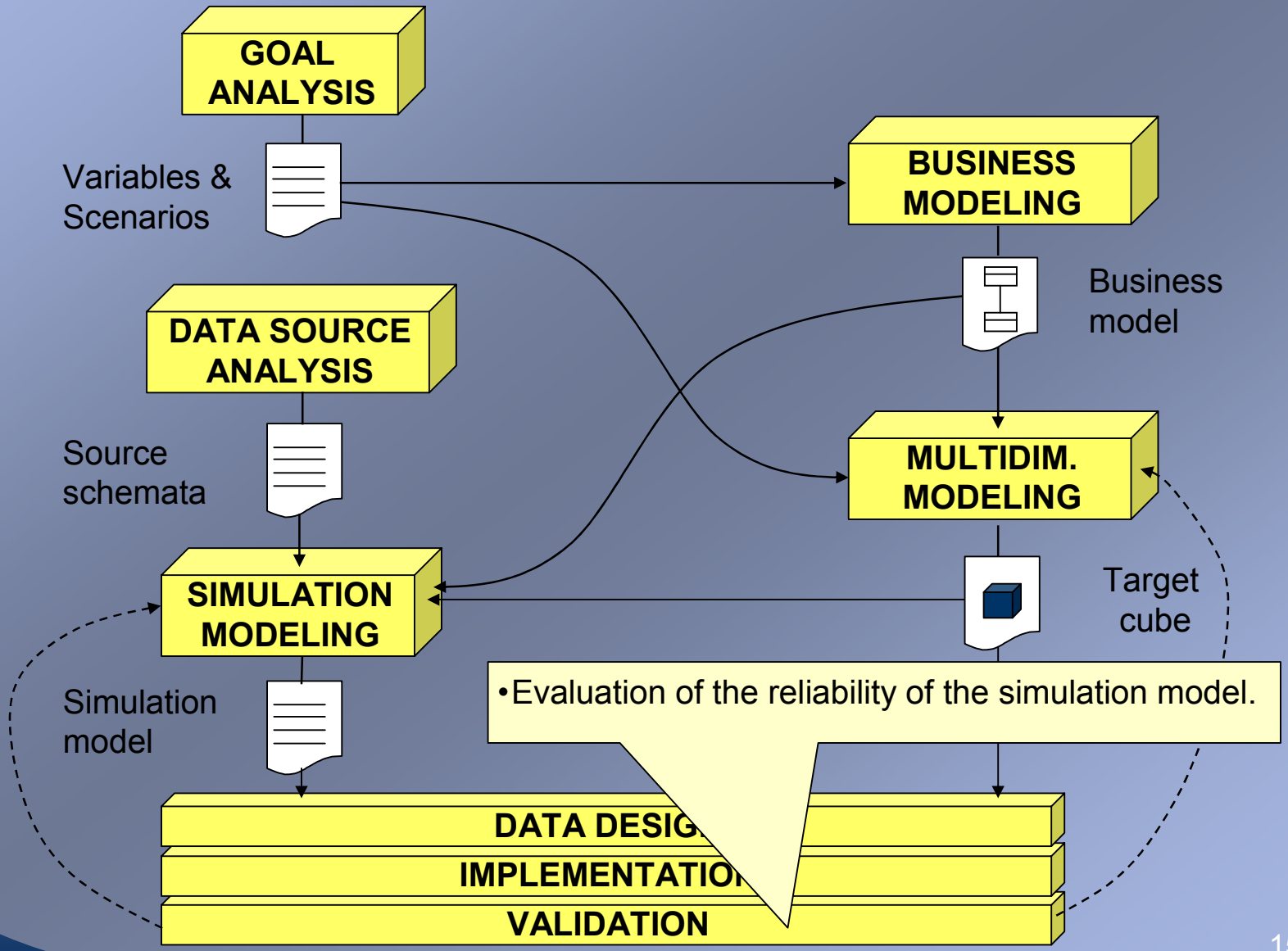
- Builds the functional/dynamic model allowing the prediction to be constructed for each given scenario
- Achieves a good compromise between precision and complexity



A methodological sketch



A methodological sketch

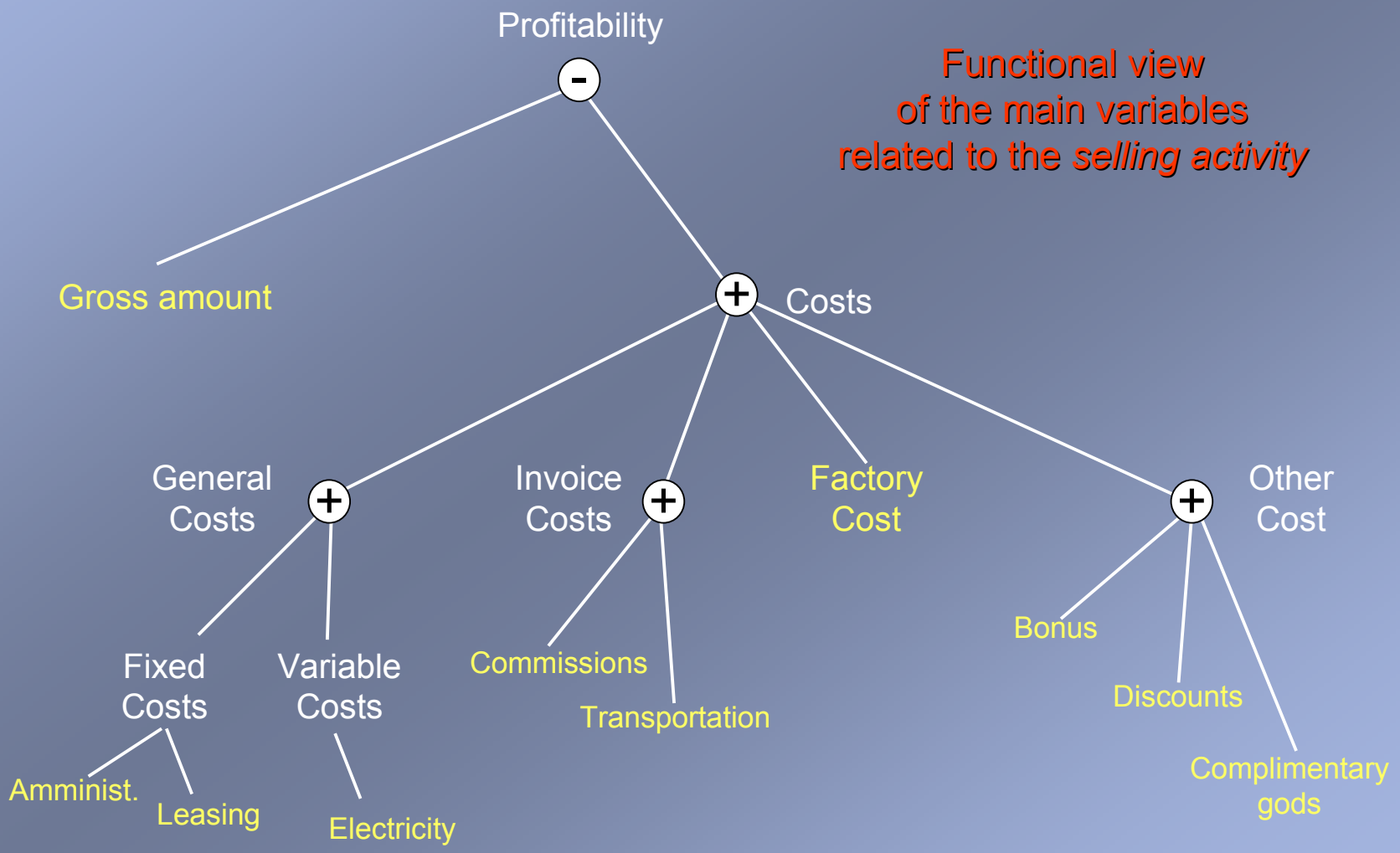


A case study: Orogel S.p.A.

- Orogel S.p.A. is a large Italian company in the area of deep-frozen food.
  - It sells and distributes its products in a set of branches scattered on the national territory
  - It is equipped with a DW covering most of its business area

- Goal analysis
  - Analyze the *profitability* of branches
  - Class of scenarios: analyze profitability during next *n* months if:
    - one or more new products were taken/dropped by a branch
    - one or more new customers were taken/dropped by a branch.

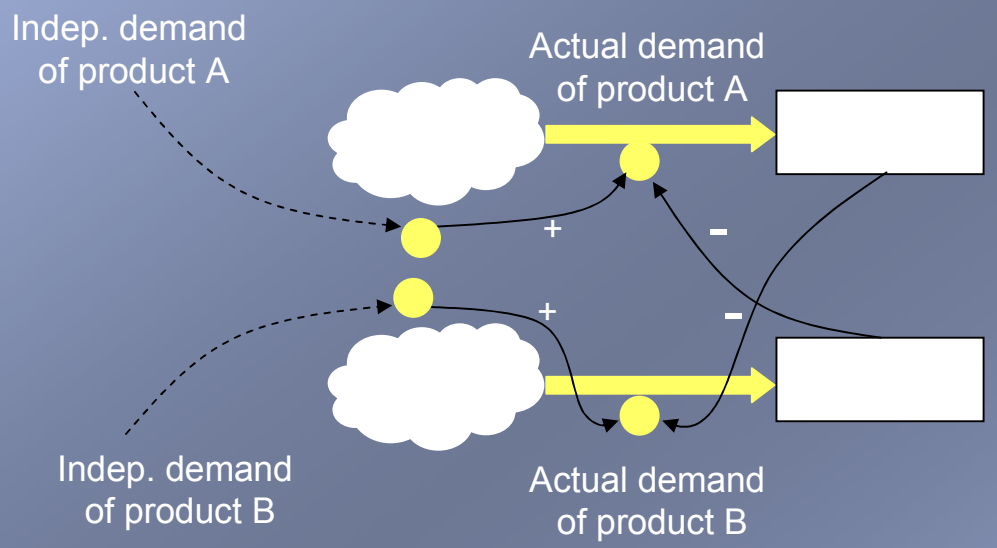
Business Modelling



Functional view
of the main variables
related to the *selling activity*

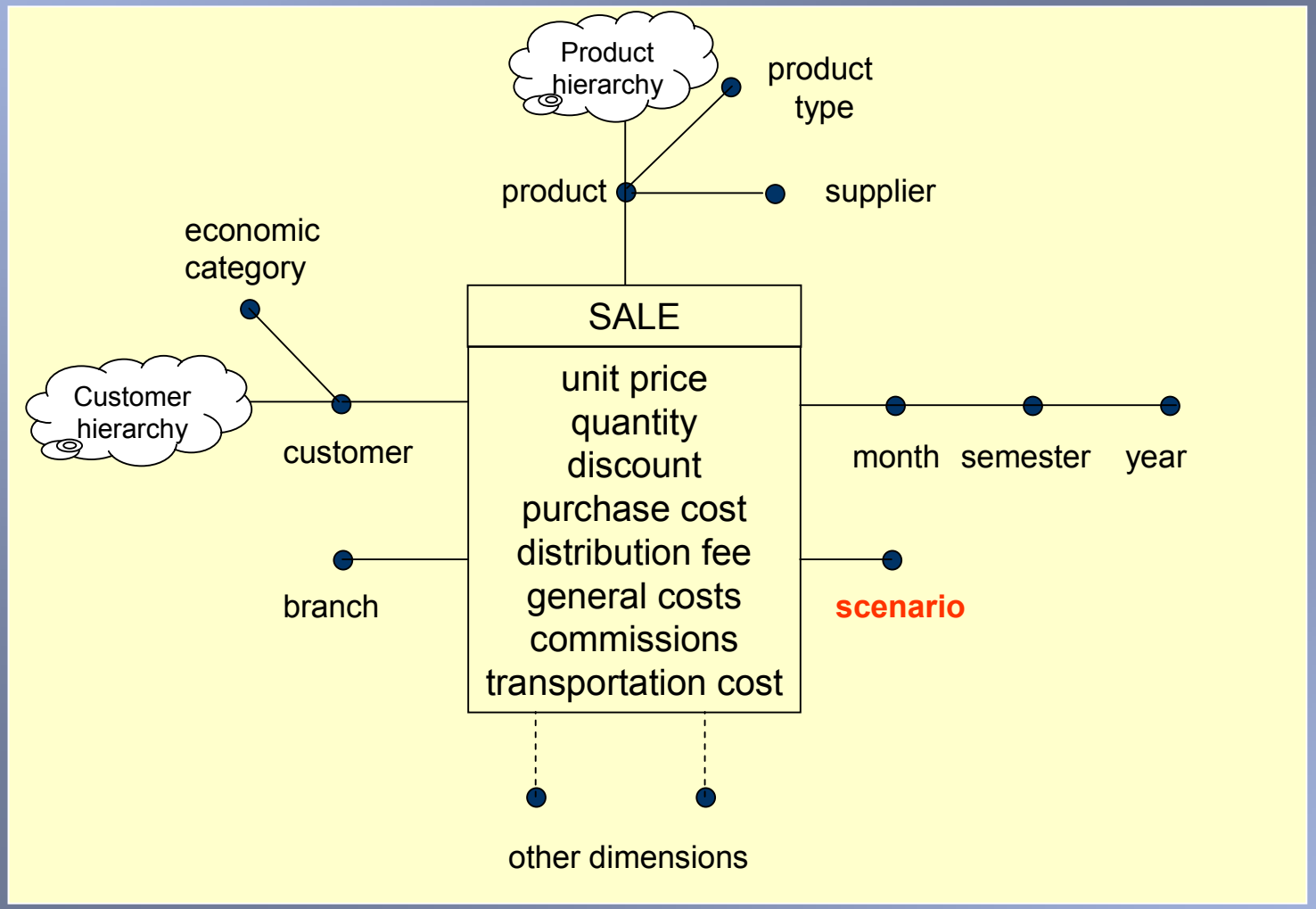
Business Modelling

Cannibalization: the process by which a new product gains sales by diverting sales from existing products

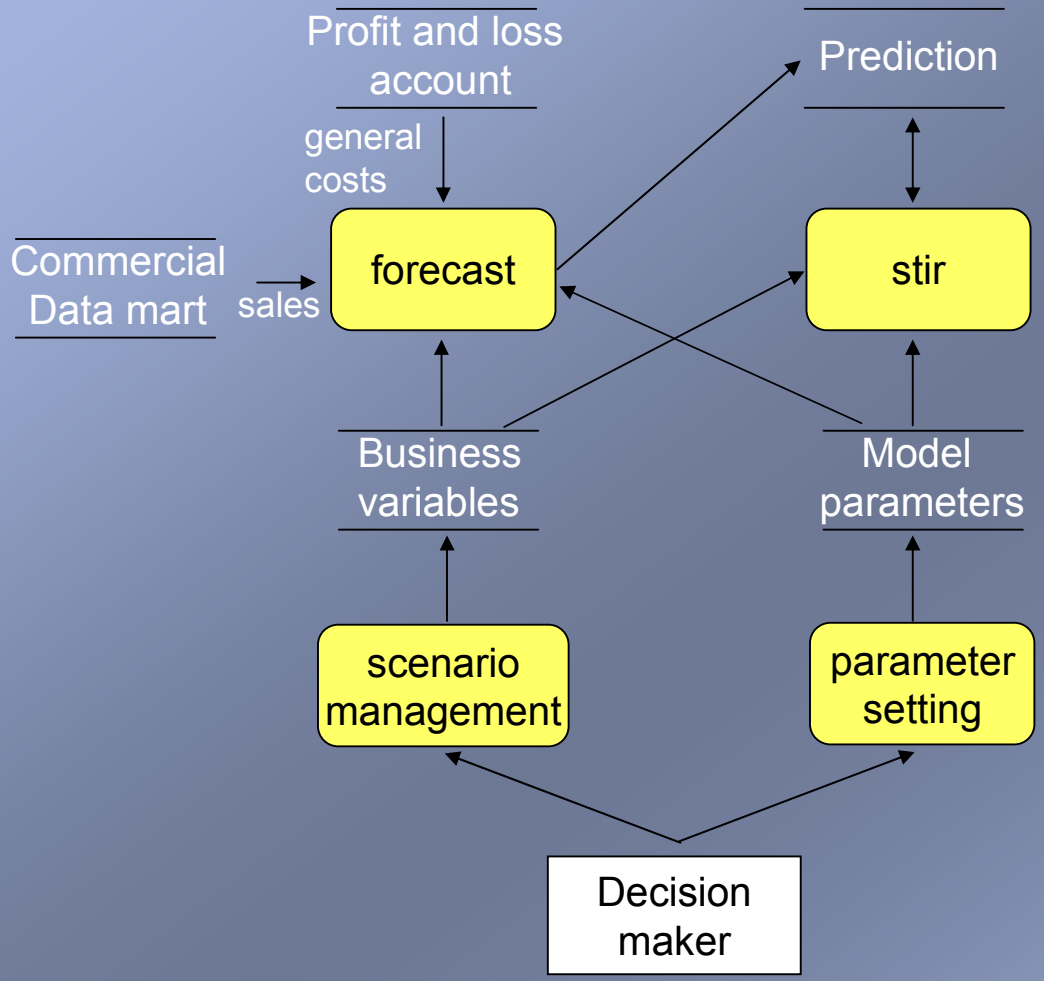


Dynamic model for cannibalization
(system dynamic graphical formalism)

Multidimensional Modelling



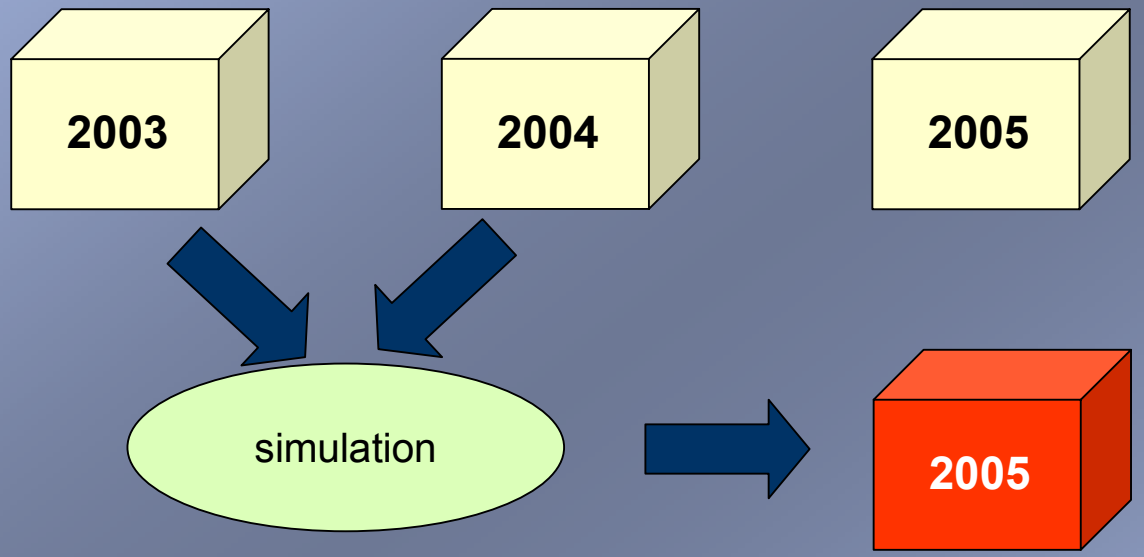
Simulation Modelling



- Functional view of the simulation model
  - statistical techniques have been adopted for both forecasting and stirring
  - Simplifications have been adopted in computing discounts and cannibalization
  - Events related to new product/customer have been simulated by reproducing the sales of a representative product/customer

Validation

- Test carried out on 2003-2004 data and compared with 2005 ones



- Average error of 18% on the total profitability of the single branches
- Global error on profitability was about 7% due to a compensation

Conclusions and …

- The spread of what-if analysis projects is surprisingly low
- Several factors contribute to this situation:
  - Immature technology:
    - The new generation of analytic tools are now compensating the technological gap
  - Complexity of design
    - Complexity can be overcome by relying on pre-configured models (e.g., SAP-BPS is based on the business models captured by its ERP)
  - Lack of a design methodology:
    - Development of a well-structured design methodology

Conclusions and (*some*) open issues

- Extend OLAP with new operators specifically devised for what-if analysis
  - E.g. *apportion*: disaggregates a quantitative information down a hierarchy according to some given criterion (*driver*);

- Find an adequate formalism to express the simulation model, so that it can be discussed and agreed upon with the users.

- Enforce consistency of multiple previsions of the same phenomenon taken at more than one abstraction level
  - E.g. *Sales prevision for Europe for next year +10%*
    - Sales prevision for Italy + 20%
    - Sales prevision for Greece + 15%
    - Sales prevision for Germany ???

Conclusions and (*some*) open issues

- Extend OLAP with new operators specifically devised for what-if analysis
  - E.g. *apportion*: disaggregates a quantitative information down a hierarchy according to some given criterion (*driver*);

- Find an adequate formalism to express the simulation model, so that it can be discussed and agreed upon with the users.

- Enforce consistency of multiple previsions of the same phenomenon taken at more than one abstraction level
  - E.g. *Sales prevision for Europe for next year +10%*
    - Sales prevision for Italy + 20%
    - Sales prevision for Greece + 15%
    - Sales prevision for Germany ???

Questions?