Modélisation d'incertitudes affectant les modèles numériques complexes

(work-in-progress)

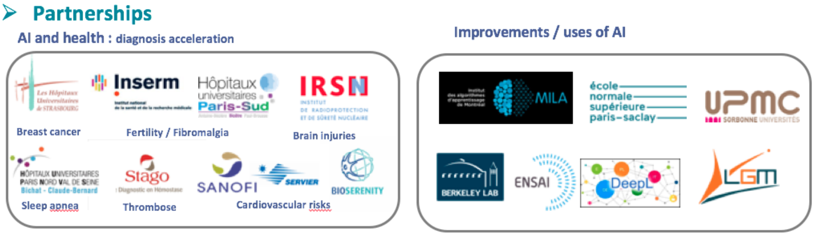
nicolas.bousquet@upmc.fr
nbousquet@quantmetry.com

with C. Denis (LIP6, EDF) for some parts

(lending agreement)

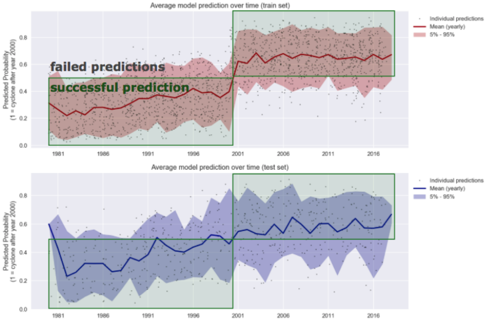
Leading the R&D programs on data science and AI tools

- Statistics, (deep) machine learning, decision theory, optimization, epistemology (for AI's interpretability)...
- Technological works on data architecture, engineering, IoT treatments



Joint work on climatic imputation of tropical storms with Berkeley Lab (NERSC) (Mr Prahbat, Michael Wehner) and Ouranos (Alexis Hannart)

Predicted Probability over Time



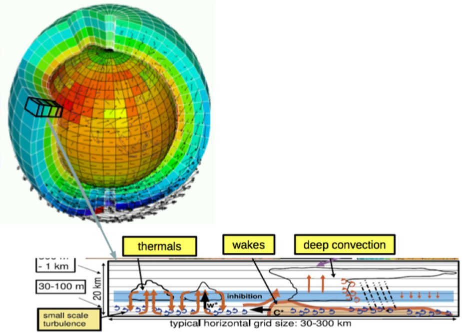
Observations

- proba > 0.5 : the classifier predicts that the storm happened after year 2000
- There is definitely some signal detected by the algorithm
- There is probably room for improvement using hyper optimisation

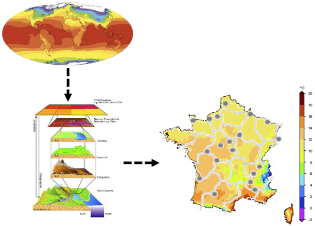
Based on documents from Jean-Louis Dufresne (LMD, IPSL)

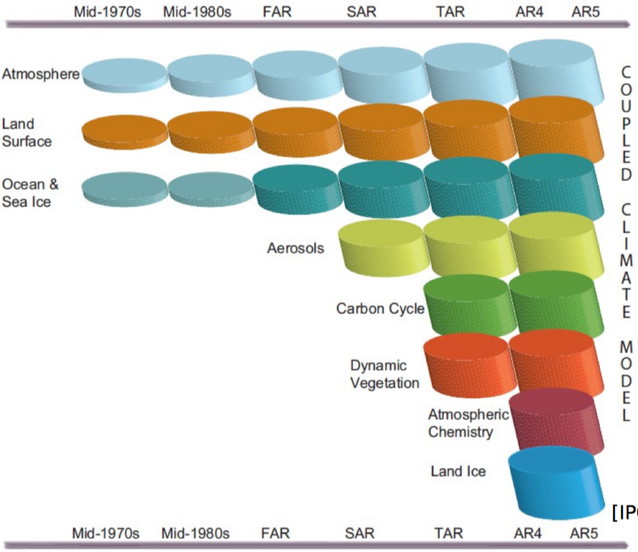
General circulation models (GCMs)

- **Dynamical core** : discretized version of the equations of fluid mechanics
- Involve terms other than fluid mechanics and unresolved scales ⇒ sub-grid models

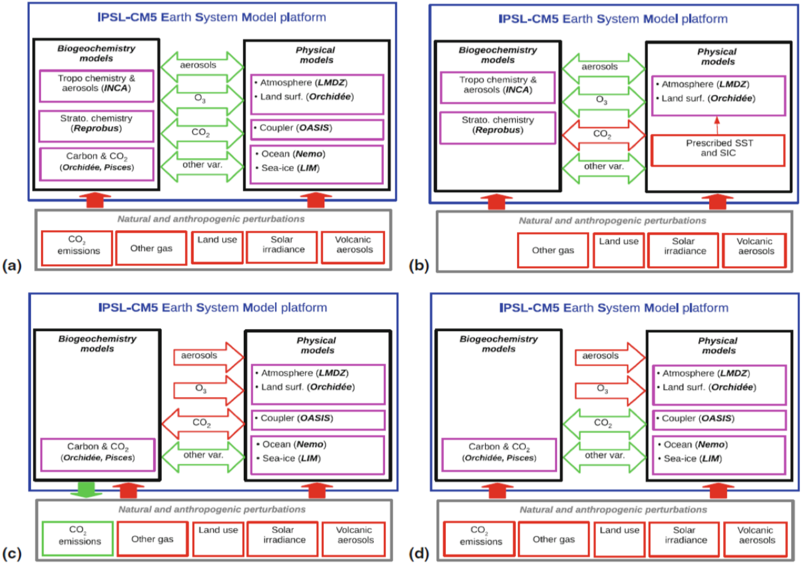


Downscaling (ex : in T°)




Mid-1970s
Mid-1980s
FAR
SAR
TAR
AR4
AR5
Atmosphere
Land Surface
Ocean & Sea Ice
Aerosols
Carbon Cycle
Dynamic Vegetation
Atmospheric Chemistry
Land Ice
COUPLED CLIMATE MODEL
Mid-1970s
Mid-1980s
FAR
SAR
TAR
AR4
AR5

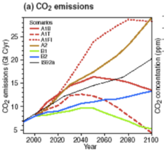
[IPCC, AR5]

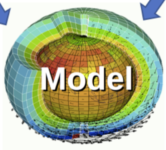

IPSL-CM5 Earth System Model platform
Biogeochemistry models
Tropo chemistry & aerosols (INCA)
Strato. chemistry (Reprobus)
Carbon & CO_2 (Orchidée, Pisces)
aerosols
O_3
CO_2
other var.
Physical models
• Atmosphere (LMDZ)
• Land surf. (Orchidée)
• Coupler (OASIS)
• Ocean (Nemo)
• Sea-ice (LIM)
Prescribed SST and SIC
Natural and anthropogenic perturbations
CO_2 emissions
Other gas
Land use
Solar irradiance
Volcanic aerosols
(a)
(b)
(c)
(d)

Initial conditions
T°, water vapour, salinity all over the globe

Boundary conditions (forcings)
Ex: greenhouse gases, insolation

RCP scenarios





Simulation results
T°, water vapour, salinity

Statistical analysis

Simulations are based on some variations of initial conditions (sensitive dependence)

The range of simulations is amplified by the chaotic nature of evolution equations of the atmosphere, and errors introduced because of model imperfections

Simulation produce members

Of which nature ? How dealing with them ?

Common use of ensemble forecasting approaches, assimilated to probabilistic forecasts (Monte Carlo-like)

Common use of averages, “standard deviations" as measures of spread, dispersion...

Complex simulation models in general

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Complex simulation models in general

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

⇒ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

Models can be complex because :

- they represent complex phenomena at various scales
- they are hard to understand (black or grey-box syndrome)
- their input are hard to calibrate
- they are very time-consuming to explore by simulation means

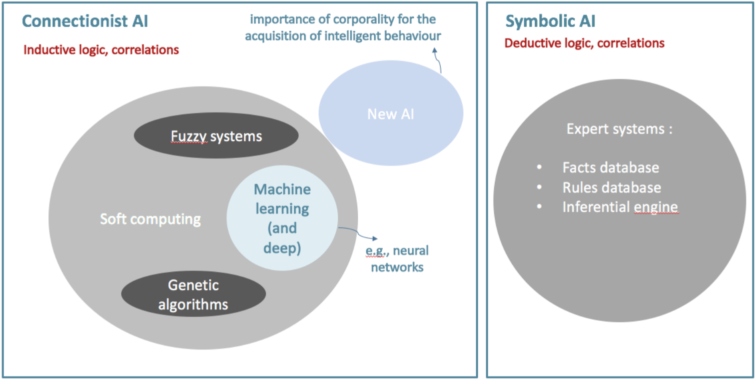
Examples :

- Energy systems as nuclear reactors (CATHARE-type models)
- Ecosystemic models (incorporating trophic relations, human forcings, etc.)
- Socio-economic models

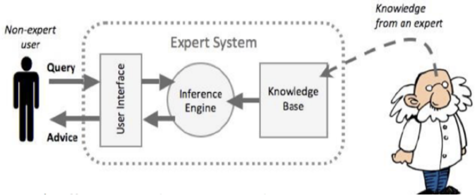
$\Rightarrow$ they are often characterized by a lack of understanding and justification about :

1. the nature and features of the uncertainties that affect them
2. the most suitable formal tools (e.g., random sequences) for representing such uncertainty
3. the most suitable computing tools (e.g., quasi-random generators) in the present state of techniques

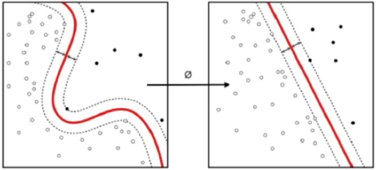
What about Artificial Intelligence (AI) ?



Symbolic AI



Machine-learning based (statistical) AI



I do not know where this image come from but I find it nice ! (sorry for non-citing the anonymous author)

Reducing connectionist AI tools to statistical approaches to mimic physics (*computer analytics*)

- Based on physically-based simulation results, producing simpler meta-models
- Relevance mostly based for neural networks on the universal approximation property
- Correlation structure constrained to respect physical properties
- Same ideas (basically) than regression, interpolation...

Making new knowledge emerging at unresolved scales for physics

- E.g., merging real data at lowest scale and simulated outputs for downscaling ?

Reducing connectionist AI tools to statistical approaches to mimic physics (*computer analytics*)

- Based on physically-based simulation results, producing simpler meta-models
- Relevance mostly based for neural networks on the universal approximation property
- Correlation structure constrained to respect physical properties
- Same ideas (basically) than regression, interpolation...

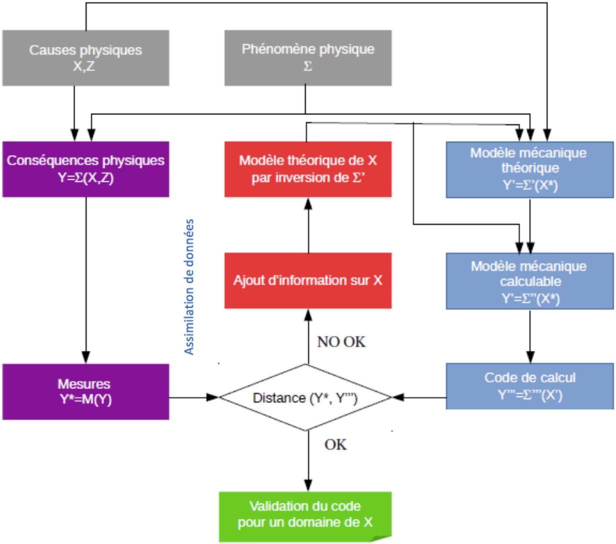
Making new knowledge emerging at unresolved scales for physics

- E.g., merging real data at lowest scale and simulated outputs for downscaling ?

Beyond the climatic framework, AI tools are more and more studied for working in interaction with causal (physically-based) complex models

Better understanding and modeling the uncertainties that affect these models and their outputs should be a prerequisite for

Effects on trust in AI, safety / reliability questionings, etc.


Causes physiques
X,Z
Phénomène physique
Σ
Conséquences physiques
Y=Σ(X,Z)
Modèle théorique de X
par inversion de Σ'
Modèle mécanique
théorique
Y'=Σ'(X*)
Assimilation de données
Ajout d'information sur X
Modèle mécanique
calculable
Y'=Σ''(X*)
NO OK
Mesures
Y*=M(Y)
Distance (Y*, Y''')
Code de calcul
Y'''=Σ'''(X')
OK
Validation du code
pour un domaine de X

The treatment of uncertainties in computer models is characterized by a lack of clear definitions

1. "Aleatoric uncertainty (natural,stochastic...) is due to the randomness or natural variability of a physical phenomenon (the values are accurate but different due to natural variations). Generally related to measurable quantities / objective knowledge and considered *irreducible* since it is inherent in the natural variability of physical phenomena" [Winkler 1996] .

2. "Epistemic uncertainty is due to the imprecise nature of knowledge or linked to a lack of knowledge. It is generally related to non-measurable quantities, and is considered reducible [Winkler 1996] in the sense that new knowledge could reduce or even eliminate this type of uncertainty. It is mainly present in the case of subjective data based on beliefs (expert opinion) and may be qualitative or quantitative.[20].

The treatment of uncertainties in computer models is characterized by a lack of clear definitions

1. "Aleatoric uncertainty (natural,stochastic...) is due to the randomness or natural variability of a physical phenomenon (the values are accurate but different due to natural variations). Generally related to measurable quantities / objective knowledge and considered *irreducible* since it is inherent in the natural variability of physical phenomena" [Winkler 1996] .

2. "Epistemic uncertainty is due to the imprecise nature of knowledge or linked to a lack of knowledge. It is generally related to non-measurable quantities, and is considered reducible [Winkler 1996] in the sense that new knowledge could reduce or even eliminate this type of uncertainty. It is mainly present in the case of subjective data based on beliefs (expert opinion) and may be qualitative or quantitative.[20].

The treatment of uncertainties in computer models is characterized by a lack of clear definitions

1. "Aleatoric uncertainty (natural,stochastic...) is due to the randomness or natural variability of a physical phenomenon (the values are accurate but different due to natural variations). Generally related to measurable quantities / objective knowledge and considered *irreducible* since it is inherent in the natural variability of physical phenomena" [Winkler 1996] .

2. "Epistemic uncertainty is due to the imprecise nature of knowledge or linked to a lack of knowledge. It is generally related to non-measurable quantities, and is considered reducible [Winkler 1996] in the sense that new knowledge could reduce or even eliminate this type of uncertainty. It is mainly present in the case of subjective data based on beliefs (expert opinion) and may be qualitative or quantitative.[20].

Not formal definitions, but qualitative characterization (exhibition of a property)

Semantics used too extensive

Require formalization

- **parametric uncertainty**
  - comes from the requirement to estimate parameters from a finite amount of information (ex : assimilated data)
  - traduced by an error between the estimated model $\Sigma"'$ and the formal model Σ'
- **model uncertainty** traduced as an error between the real phenomenon Σ and the theoretical model Σ' ;
- **completness uncertainty**
  - lack of knowledge of the real exhaustiveness of a model, limited by the choice of its input parameters, with regard to the real first name
  - traduced by an error between Σ and Σ' over the circonscription of the model perimeter through the choice of its parameters

[**Error** *(Oberkampf 2002)*.] An error in the sense of Oberkampf (2002) is an identifiable imprecision that is not due to a lack of knowledge. It can be voluntary (e. g. simplification of a mesh used to accelerate calculations) or unintentional (e. g. programming error)

In the concrete problem of numerical modelling of a phenomenon of a mechanistic nature, the addition of knowledge is expressed in practice by :

1. a refinement in the execution of the program implementing the theoretical model (e.g., downscaling) ;
2. a refinement in the algorithmic description of the phenomenon (for example, using additional structural parameters and equations).

In the concrete problem of numerical modelling of a phenomenon of a mechanistic nature, the addition of knowledge is expressed in practice by :

1. a refinement in the execution of the program implementing the theoretical model (e.g., downscaling) ;
2. a refinement in the algorithmic description of the phenomenon (for example, using additional structural parameters and equations).

In the concrete problem of numerical modelling of a phenomenon of a mechanistic nature, the addition of knowledge is expressed in practice by :

1. a refinement in the execution of the program implementing the theoretical model (e.g., downscaling) ;
2. a refinement in the algorithmic description of the phenomenon (for example, using additional structural parameters and equations).

Expert judgment

Another key sub-theme in epistemic uncertainty is the use of expert judgment for assessing uncertain information when there is a lack of experimental data (or other objective source of information)

This lesson is driven by the questions :

Why and how stochastic modeling can be a relevant tool for using expert judgment, and more generally dealing with epistemic uncertainty ?

Another key sub-theme in epistemic uncertainty is the use of expert judgment for assessing uncertain information when there is a lack of experimental data (or other objective source of information)

This lesson is driven by the questions :

Why and how stochastic modeling can be a relevant tool for using expert judgment, and more generally dealing with epistemic uncertainty ?

Another key sub-theme in epistemic uncertainty is the use of expert judgment for assessing uncertain information when there is a lack of experimental data (or other objective source of information)

This lesson is driven by the questions :

Why and how stochastic modeling can be a relevant tool for using expert judgment, and more generally dealing with epistemic uncertainty ?

Beyond scientific computing and uncertainty propagation, expert judgment has a foremost role in decision-making

Beyond scientific computing and uncertainty propagation, expert judgment has a foremost role in decision-making

- guiding designs of experiments, ordering scientific results [3, 36, 19]
-
-

Beyond scientific computing and uncertainty propagation, expert judgment has a foremost role in decision-making

- guiding designs of experiments, ordering scientific results [3, 36, 19]
- enriching economic [18] and actuarial studies [33] on the impact of financial risks
-

Beyond scientific computing and uncertainty propagation, expert judgment has a foremost role in decision-making

- guiding designs of experiments, ordering scientific results [3, 36, 19]
- enriching economic [18] and actuarial studies [33] on the impact of financial risks
- being determining in legal arbitration, public policies [22] or environmental governance [21, 7]

Beyond scientific computing and uncertainty propagation, expert judgment has a foremost role in decision-making

- guiding designs of experiments, ordering scientific results [3, 36, 19]
- enriching economic [18] and actuarial studies [33] on the impact of financial risks
- being determining in legal arbitration, public policies [22] or environmental governance [21, 7]

Its influence on technological, economic, societal or personal choices when elaborating strategies of gain-winning is explored by many epistemological and psychological authors [11, 19, 10]

An infinite number of conceptions

Among them, two main kinds of experts for [36] :

1. (performative expertise)
2.

An infinite number of conceptions

Among them, two main kinds of experts for [36] :

1. *those who expertise is a function of what they do (performative expertise)*
2.

An infinite number of conceptions

Among them, two main kinds of experts for [36] :

1. *those who expertise is a function of what they do (performative expertise)*
2. *those who expertise is a function of what they know (epistemic expertise)*

An infinite number of conceptions

Among them, two main kinds of experts for [36] :

1. *those who expertise is a function of what they do (performative expertise)*
2. *those who expertise is a function of what they know (epistemic expertise)*

An infinite number of conceptions

Among them, two main kinds of experts for [36] :

1. *those who expertise is a function of what they do (performative expertise)*
2. *those who expertise is a function of what they know (epistemic expertise)*

An usual view, with the ability of explaining and transmitting. Furthermore, acccording to Luntley [19] :

I argue that what differentiates the epistemic standpoint of experts is not what or how they know [...], but their capacity for learning

Today's question is in fact "what is *formally* an expert ?"

We should rather talk about **"expert systems delivering new knowledge"**

Typically :

- implicit cognitive systems
  - humans
  - some artificial intelligences
- explicit causal systems
  - phenomenological models and their numerical implementation (simulation models)

Capacity for proving expertness ⇔ capacity of predicting adequately

Capacity for learning ⇔ capacity of inferring (processing) coherently when new data arrive

Eliciting = assessing her/his/its relevant epistemic information on the behavior of a magnitude of interest $X \in \chi$

elicio, eliciere : to extract from, to drawout (*ex aliquo verbum elicere*)

Immediate difficulties

- bias
- impact of subjectivity in the delivery process
- lack of correct or sharp information
- ...

resulting in epistemic uncertainty

Our work : formalizing the most adapted measure of uncertainty, highlighting clearly the subjective and objective parts of the modeling

Eliciting = assessing her/his/its relevant epistemic information on the behavior of a magnitude of interest $X \in \chi$

elicio, eliciere : to extract from, to drawout (*ex aliquo verbum elicere*)

Immediate difficulties

- bias
- impact of subjectivity in the delivery process
- lack of correct or sharp information
- ...

resulting in epistemic uncertainty

Our work : formalizing the most adapted measure of uncertainty, highlighting clearly the subjective and objective parts of the modeling

Prior information = *information whose the value of truth is justified by considerations independent on experiment on focus* [24]

- other trial results (e.g., on mock-ups)
- technical running specification
- physical bounds
- literature corpus
- and of course, human experts

Often incomplete, always uncertain, because

- of the non-existence of a system allowing a priori if the expertness is complete or not
- of the non-existence of a system precise enough to specify that $X = x_0$ exactly (except in rare cases)

What means “uncertainty" and especially “epistemic uncertainty" ?

Why probabilities for dealing with uncertainty ?

If we are ok with probabilities, how choosing the probability distributions ?

What means “uncertainty" and especially “epistemic uncertainty" ?
Hard philosophical question ! Providing answering attempts here

Why probabilities for dealing with uncertainty ?

If we are ok with probabilities, how choosing the probability distributions ?

What means “uncertainty" and especially “epistemic uncertainty" ?
Hard philosophical question ! Providing answering attempts here

Why probabilities for dealing with uncertainty ?
Many practical advantages, but how proving they are theoretically relevant ?

If we are ok with probabilities, how choosing the probability distributions ?

What means “uncertainty" and especially “epistemic uncertainty" ?
Hard philosophical question ! Providing answering attempts here

Why probabilities for dealing with uncertainty ?
Many practical advantages, but how proving they are theoretically relevant ?
Raises the question of **auditability** of mathematical procedures = growingly increasing concern

If we are ok with probabilities, how choosing the probability distributions ?

What means “uncertainty" and especially “epistemic uncertainty" ?
Hard philosophical question ! Providing answering attempts here

Why probabilities for dealing with uncertainty ?
Many practical advantages, but how proving they are theoretically relevant ?
Raises the question of **auditability** of mathematical procedures = growingly increasing concern

If we are ok with probabilities, how choosing the probability distributions ?
Use the help of important Bayesian prior modeling techniques

Treating prior information from implicit cognitive systems

If we were omniscient, a causal model could be

$$X \quad = \quad g(Z)$$

where :

- Z is a hidden property of the experiment
- g is a model of information production

The value of Z could be explained by another transformation $\tilde{g}$ of another hidden property $\tilde{\theta}$, etc.

However, there is still a model error between the true values of X and $g(Z)$, since nor g neither Z are known (completely or not)

Hypothesis 1 (epistemological) by Lakatos [17]

- Information on the world is hidden and partially revealed by a consensual theory (*in the sense of Popper [25] : by mutual decision of protagonists*) defining objectivity [12]
- Knowledge is “filtered" from information
- Filtering is performed through the intervention of symbols, or signs, in order to transmit it or even implement it

Hypothesis 2 (arising from neurosciences) [28, 27, 26, 14, 6, 2]

- Face to situations where uncertain information is mobilized, human reasoning produces probabilistic inferences
- Difficulties appear when trying to explicit this inferred knowledge by an interpretative language $\Rightarrow$ providing usable expertness

Hypothesis 1 (epistemological) by Lakatos [17]

- Information on the world is hidden and partially revealed by a consensual theory (*in the sense of Popper [25] : by mutual decision of protagonists*) defining objectivity [12]
- Knowledge is “filtered" from information
- Filtering is performed through the intervention of symbols, or signs, in order to transmit it or even implement it

Hypothesis 2 (arising from neurosciences) [28, 27, 26, 14, 6, 2]

- Face to situations where uncertain information is mobilized, human reasoning produces probabilistic inferences
- Difficulties appear when trying to explicit this inferred knowledge by an interpretative language $\Rightarrow$ providing usable expertness

We don't know what is the "deconvolution" transforming uncertain knowledge backwards into uncertain information, following Lakatos' hypothesis

But we can have ideas about the impact of the addition of uncertain but useful knowledge in the problem of determining X

It should traduce by the increasing of information on X = inference (updating)

$\Rightarrow$ this inference should stands on a reasoning principle

$\Rightarrow$ this principle should stand on a logic = set of **formal rules**

We don't know what is the "deconvolution" transforming uncertain knowledge backwards into uncertain information, following Lakatos' hypothesis

But we can have ideas about the impact of the addition of uncertain but useful knowledge in the problem of determining X

Desirable properties [34]

- Sorting *atomic* assertions of type $X = x_0$ at each addition of information (*exclusive logic*)
  - an initial situation (premise) is less informative than a conclusion (updating)
- Allowing *uncertain* information
  - not only true or false situations can be sorted (*non-boolean logic*)

Definition [34]

Denote S_X a set of atomic propositions of type $X = x_i$. The set B_X of all possible *compound propositions* generated by

$$\begin{aligned} \neg X = x_i, &\quad X = x_i \wedge X = x_j, \\ X = x_i \vee X = x_j, &\quad X = x_i \Rightarrow X = x_j \\ \text{and} &\quad X = x_i \Leftrightarrow X = x_j \end{aligned}$$

is called a state of information, with $\mathrm{Dom}(B_X) =$ logical closure of S_X

The state of information B_X summarizes the existing information on a set of propositions about X

The same logic should guide how B_X evolves : it is growing following a given metric when information on X is increasing

Definition [34]

Denote S_X a set of atomic propositions of type $X = x_i$. The set B_X of all possible *compound propositions* generated by

$$\begin{aligned} \neg X = x_i, &\quad X = x_i \wedge X = x_j, \\ X = x_i \vee X = x_j, &\quad X = x_i \Rightarrow X = x_j \\ \text{and} &\quad X = x_i \Leftrightarrow X = x_j \end{aligned}$$

is called a state of information, with $\mathrm{Dom}(B_X) =$ logical closure of S_X

The state of information B_X summarizes the existing information on a set of propositions about X

The same logic should guide how B_X evolves : it is growing following a given metric when information on X is increasing

Definition [34]

Denote S_X a set of atomic propositions of type $X = x_i$. The set B_X of all possible *compound propositions* generated by

$$\begin{aligned} \neg X = x_i, &\quad X = x_i \wedge X = x_j, \\ X = x_i \vee X = x_j, &\quad X = x_i \Rightarrow X = x_j \\ \text{and} &\quad X = x_i \Leftrightarrow X = x_j \end{aligned}$$

is called a state of information, with $\mathrm{Dom}(B_X) =$ logical closure of S_X

The state of information B_X summarizes the existing information on a set of propositions about X

The same logic should guide how B_X evolves : it is growing following a given metric when information on X is increasing

Definition

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

- **Consistency** : B_X is consistent if there is no proposition A for which both $[A|B_X] = T$ and $\neg[A|B_X] = T$
- **Propositional calculus** :

  (*i*) If $A = A'$ then $[A|B_X] \Leftrightarrow [A'|B_X]$
  (*ii*) $[A|B_X, C_X, D_X] = [A|(B_X \wedge C_X), D_X]$
  (*iii*) If B_X consistent and $\neg[A|B_X] < T$, then $A \cup B_X$ is consistent

- **Coherence** : there exists a non-increasing function S_0 such that, for all x and consistent B_X

$$\neg[A|B_X] = S_0([A|B_X])$$

- **Density** : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Definition

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

- **Consistency** : B_X is consistent if there is no proposition A for which both $[A|B_X] = T$ and $\neg[A|B_X] = T$
- **Propositional calculus** :

$$\begin{array}{ll} (i) & \text{If } A = A' \text{ then } [A|B_X] \Leftrightarrow [A'|B_X] \\ (ii) & [A|B_X, C_X, D_X] = [A|(B_X \wedge C_X), D_X] \\ (iii) & \text{If } B_X \text{ consistent and } \neg[A|B_X] < T, \text{ then } A \cup B_X \text{ is consistent} \end{array}$$

- **Coherence** : there exists a non-increasing function S_0 such that, for all x and consistent B_X

$$\neg[A|B_X] = S_0([A|B_X])$$

- **Density** : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Definition

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

- **Consistency** : B_X is consistent if there is no proposition A for which both $[A|B_X] = T$ and $\neg[A|B_X] = T$
- **Propositional calculus** : applicable to any problem domain for which we can formulate useful propositions

$$\begin{array}{ll} (i) & \text{If } A = A' \text{ then } [A|B_X] \Leftrightarrow [A'|B_X] \\ (ii) & [A|B_X, C_X, D_X] = [A|(B_X \wedge C_X), D_X] \\ (iii) & \text{If } B_X \text{ consistent and } \neg[A|B_X] < T \text{, then } A \cup B_X \text{ is consistent} \end{array}$$

- **Coherence** : there exists a non-increasing function S_0 such that, for all x and consistent B_X

$$\neg[A|B_X] = S_0([A|B_X])$$

- **Density** : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Definition

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

- **Consistency** : B_X is consistent if there is no proposition A for which both $[A|B_X] = T$ and $\neg[A|B_X] = T$
- **Propositional calculus** :

$$\begin{array}{ll} (i) & \text{If } A = A' \text{ then } [A|B_X] \Leftrightarrow [A'|B_X] \\ (ii) & [A|B_X, C_X, D_X] = [A|(B_X \wedge C_X), D_X] \\ (iii) & \text{If } B_X \text{ consistent and } \neg[A|B_X] < T\text{, then } A \cup B_X \text{ is consistent} \end{array}$$

- **Coherence** : there exists a non-increasing function S_0 such that, for all x and consistent B_X

$$\neg[A|B_X] = S_0([A|B_X])$$

- **Density** : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Definition

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

- **Consistency** : B_X is consistent if there is no proposition A for which both $[A|B_X] = T$ and $\neg[A|B_X] = T$
- **Propositional calculus** :

$$\begin{array}{ll} (i) & \text{If } A = A' \text{ then } [A|B_X] \Leftrightarrow [A'|B_X] \\ (ii) & [A|B_X, C_X, D_X] = [A|(B_X \wedge C_X), D_X] \\ (iii) & \text{If } B_X \text{ consistent and } \neg[A|B_X] < T\text{, then } A \cup B_X \text{ is consistent} \end{array}$$

- **Coherence** : there exists a non-increasing function S_0 such that, for all x and consistent B_X

$$\neg[A|B_X] = S_0([A|B_X])$$

- **Density** : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Axiom

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

This axiom of non-ambiguity is particularly important

This is an assumption of *universal comparability*

Consequence : an additional information (not a knowledge) can only increase or decrease the plausibility of a proposition

Axiom

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

This axiom of non-ambiguity is particularly important

This is an assumption of *universal comparability*

As seen later, the differences between probabilistic logic and extra-probabilistic logics arises from the agreement or disagreement with this assumption

Jaynes [16] argues for its validity on pragmatic grounds

Axiom

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

It is supported when we talk about quantities X with physical meanings and taking a unique value at each instant (possibly given a finite measurement precision)

It may be not supported if we talk about :

- magnitudes considered at the quantum scale (e.g., in neutronics)
- imaginary magnitudes (e.g., latent variables)

Remember that we are dealing with objective information on X, not interpreted knowledge !

Statement

Density : the set $[S_0(T), T]$ admits a non-void, dense and consistent subset

Can be false when the set of all propositions is finite (e.g., discrete and bounded) [15]

Could be partially removed by arguments provided by Snow [31], plaiding for infinite gradations of plausibility within even a single, finite domain

The source of an objective rational measure of belief is external to the cognitive apparatus of the believer. Its value is determined by the vagaries of the real world or by some idealized model of the world. There is no way to tell in advance just which values must arise, and each value may be graduated with arbitrary precision. Any such value can simply be adopted by the believer without recourse to unboundedly precise discrimination between affective states related to credibility... [31]

Working (as usually) with uncountable input spaces for X is not an issue :-)

1. Reproductibility rule : two equivalent assertions about X have the same plausibility
2. Non-contradiction rule : if it exists several ways of coming to the same conclusion about X, all have the same plausibility
3. Consistency rule : the logic cannot reach a conclusion contradicted by the common deductive rules (*e.g., transitivity*)
4. Integrity rule : the logic cannot disregard a part of information to reach to a conclusion about X to come to a conclusion
5. Monotony rule : the plausibility of the non-exclusive union of two assertions is at least equal to the upper plausibility of each
6. Product rule : the plausibility of the intersection of two assertions is at most equal to the lower plausibility of each

Originally proven (erroneously) by Cox [4], corrected by Jaynes [16], extended more rigorously by Paris [23], Van Horn [34] Dupré and Tipler [9] (among others) then finalized by Terenin and Draper [32]

Theorem

Under the previous assumptions, there exists a continuous, increasing function $\mathbb{P}$ such that, for every proposition A, C and consistent B_X,

(i) $\mathbb{P}([A|B_X]) = 0$ iif A is known to be false given the information in X

(ii) $\mathbb{P}([A|B_X]) = 1$ iif A is known to be true given the information in X

(iii) $0 \leq \mathbb{P}([A|B_X]) \leq 1$

(iv) $\mathbb{P}([A \wedge C|B_X]) = \mathbb{P}([A|B_X])\mathbb{P}([C|A, B_X])$

(v) $\mathbb{P}(\neg[A|B_X]) = 1 - \mathbb{P}([A|B_X])$

Any system of plausible reasoning, under the previous assumptions, is isomorphic to probability theory

Goertzel [13] proved that if the consistency rule is weakened, then plausibilities behave approximately like probabilities

The probability theory is relevant to account for uncertainties on a subject explored by a cognitive system (human or machine) which could be not completely consistent

Numerous authors in artificial intelligence [35], epistemology [1] or cognitive sciences [5] recognize the practical relevance of this axiomatic for extracting or updating information, using Bayes rule

Axiom

Consider any proposition A on X. Given B_X, the plausibility $[A|B_X]$ is a single real number, upperly bounded by a real (finite or infinite) T

Its mot common "relaxation" is the assumption that two dimensions are required to represent correctly the plausibility of a proposition

At the origin of belief theory [29, 30] and possibility theory [8]

Experiments show that such a relaxation is clearly supported when the plausibility is understood as the summary of a belief, or a *gamble* [34]

Nonetheless, this "relaxation" remains arbitrary, and usually stands on an interpretation of *the nature of knowledge* (expressed through a language), and not of the *nature of information* (expressed by physical reality or an idealized model of the reality) [31]

Treating uncertain prior information from causal models

Practical models used by engineers (e.g., implemented computer codes Σ''') can produce prior simulations of a phenomenon Σ

$$\begin{array}{ccccccc} \text{Real phenomenon} & \to & \text{Theoretical model} & \to & \text{Algorithmic model} & \to & \text{Implemented model} \\ \Sigma & & \Sigma' & & \Sigma'' & & \Sigma''' \end{array}$$

We want to define what is the conceptual nature of model uncertainty affecting Σ'''

We could ask the question otherwise : what is the conceptual nature of reduction of model uncertainty ?

We need also to define Σ'''

Program. Sequence of operations and instructions

Algorithm. Finite and non-ambiguous sequence of operations and instructions allowing for solving a problem that can be solved exhaustively

Self-delimiting program. A program that ends. Its ending is a command of the program itself

Happens at step Σ''

- Refining the algorithmic description Σ'' by adding new parameters and/or structural equations, necessarily based on improvement of Σ'
- Refining the execution of Σ''' (e.g., improving a tolerance)

Reducing model uncertainty implies to reduce model error

Maybe the nature of model error could say something about the nature of model uncertainty ?

We consider an illustrative example

Consider a real phenomenon Σ with output Y described by

$$\begin{aligned} \chi \times \chi_Z &\rightarrow \Upsilon \\ \Sigma : X, Z &\mapsto Y \end{aligned}$$

where X are known and treated variables, and Z are unknown or untreated variables

Consider a self-delimiting, calculable model of Σ

$$\begin{aligned} \chi_d &\rightarrow \Upsilon_d \\ \Sigma''' : X'' &\mapsto Y'' \end{aligned}$$

where

- $\chi_d \subsetneq \chi$ is the subset of χ that can be reached by a calculus
- $\Sigma(\chi_d, \chi_Z) = \Upsilon_d$ (Galerkin problem solving)

Assume the following hypotheses

$$\begin{aligned} (\mathrm{H1}) \quad &: \quad \mathrm{Card}(\chi_d) < \infty, \\ (\mathrm{H2}) \quad &: \quad \chi_Z \text{ is countable and } \mathrm{Card}(\chi_Z) < \infty. \end{aligned}$$

Assume that Υ_d is a metric space

It is possible to define a model error $\delta(x, z)$, through a measure $\mathcal{D}$ such that, for all couple $(x, z) \in \chi_d \times \chi_Z$,

$$\delta(x, z) \quad = \quad \mathcal{D}\left\{\Sigma''(x), \Sigma(x, z)\right\} \;\geq\; 0$$

with $\delta(x, z) = 0$ iif $\Sigma'''(x) = \Sigma(x, z)$

Proposition [B. and Denis 2017]

The model error $\delta(x, z)$ cannot be calculated $\forall (x, z) \in \chi_d \times \chi_Z$

Proof : based on tools of computational complexity theory

A more general result can be obtained using Turing's machines

The previous proposition (and its extensions) indicate that no algorithm is able to compute all the values of the model error $\delta(x, z)$

- We cannot prove that the error never exists
- Being cautious, we assume its existence

What would be the nature of the best reachable (computable) approximation $\tilde{\delta}(x,z)$ of $\delta(x,z)$?

- $\tilde{\delta}(x,z)$ should be computed by a self-delimiting program
- however there is no recursive function allowing to predict the next value of $\tilde{\delta}(x',z')$ at (x',z')

It comes that any finite sequence of $\tilde{\delta}(x_i,z_i)$ is exhaustively described only by itself

The adapted formalism to describe this property is the following

Kolmogorov's algorithmic complexity

Kolmogorov's complexity $H(s)$ of a program producing a sequence s is the length of the smallest program required to generate s.

A consequence of the impossibility of compressing the information in the sequence of $\tilde{\delta}(x_i, z_i)$ is the following : $\exists c \in \boldsymbol{R}$ such that

$$H(\tilde{\delta}(x_1, z_1), \ldots, \tilde{\delta}(x_n, z_n)) \quad \geq \quad n - c. \tag{1}$$

Result (1) implies that the sequence $\tilde{\delta}(x_i, z_i)$ is random in the sense of Chaitin-Levin

Proposition (B. and Denis 2017)

The best computable approximation of model error is random.

Randomness contamines the nature of all concepts incorporating model error

It is arguable to use probabilities for modeling epistemic model uncertainty

Stochastic prior modeling : examples and recipes

(if time ok)

Need to define, locate and separate the two sources of uncertainty

How modelling epistemic uncertainty ? Bayesian principles

- Good principles of modelling epistemic and aleatoric uncertainties in grey-box computer codes
- How about the other theories of uncertainties ?

A. Barberousse.
La valeur de la connaissance approchée. l'épistémologie de l'approximation d'émile borel.
Revue d'Histoire des Mathématiques, 14 :53–75, 2008.

S.C.Y. Chan, Y. Niv, and K.A. Norman.
A probability distribution over latent causes, in the orbifrontal cortex.
The Journal of Neuroscience, 36 :7817–7828, 2016.

R.M. Cooke.
Experts in Uncertainty : Opinion and Subjective Probability in Science.
Oxford University Press, 1991.

R.T. Cox.
Probability, frequency and Reasonable Expectation.
American Journal of Physics, 14 :1–10, 1946.

S. Dehaene.
The statistician brain : the Bayesian revolution in cognitive science.
Lecture at College de France : Chair in Experimental Cognitive Psychology, 2012.

S. Dehaene.
Consciousness and the Brain : Deciphering How the Brain Codes our Thoughts.
Viking Press, 2014.

M. Drescher, A.H. Perera, C.J. Johson, L.J. Buse, C.A. Drew, and M.A. Burgman.
Toward rigorous use of expert knowledge in ecological research.
Ecosphere, 4 :1–26, 2013.

D. Dubois and H. Prade.
Possibility Theory.
Springer, 2012.

M.J Dupré and F.J. Tipler.
New axioms for rigorous Bayesian probability.
Bayesian Analysis, 4 :599–606, 2009.

A. Eagle.
Philosophy of Probability : Contemporary Readings.
Routledge, 2011.

B. Fischhoff and D. MacGregor.
Subjective confidence in forecasts.
Journal of Forecasting, 1 :155–172, 1982.

A. Gelman and C. Hennig.
Beyond subjective and objective in statistics.
Journal of the Royal Statistical Society, Ser. A, 180 :1–31, 2017.

B. Goertzel.
Probability Theory Ensues from Assumptions of Approximate Consistency : A Simple Eerivation and its Implications for AGI.
Proceedings of AGI-13, Springer, 2013.

J.I. Gold and H.R. Heekeren.
Neural Mechanisms for Perceptual Decision Making.
In : Neuroeconomics (chapter 19), P.W. Glimcher and R. Fehr (eds), Second Edition, 2013.

J.Y. Halpern.
A counterexample to theorems of cox and fine.
Journal of Artificial Intelligence Research, 10 :67–85, 1999.

E.T. Jaynes.
Probability Theory : The Logic of Science.
2003.

I. Lakatos.
Falsification and the methodology of scientific research programmes. In : Criticism and the Growth of Knowledge, I. Lakatos and A. Musgrave (eds.).
Cambridge University Press, 1970.

J. Leal, S. Wordsworth, R. Legood, and E. Blair.
Eliciting expert opinion for economic models : an applied example.
Value Health, 10 :195–203, 2007.

M. Luntley.
Understanding Expertise.
Journal of Applied Philosophy, 26 :356–370, 2009.

C. Mercat Rommens, E. Chojnacki, and C. Baudrit.
Représentation and propagation de la connaissance imprécise : ce que les théories de l'incertain peuvent apporter aux sciences environnementales.
Colloque international Incertitude and Environnement, Arles (France). Editions EEISUE, collection Société d'Ecologie Humaine, 2005.

C.A. Miller and P.N. (eds) Edwards.
Changing the Atmosphere. Expert Knowledge and Environmental Governance.
MIT Press, 2001.

M.G. Morgan.
Use (and abuse) of expert elicitation in support of decision making for public policy.
Proceedings of the US National Academy of Sciences, 111 :7176–7184, 2014.

J.B. Paris.
The Uncertain Reasoner's Companion : a Mathematical Perspective.
Cambridge University Press, 1994.

M. Pegny.
Les deux formes de la thèse de church-turing et l'épistémologie du calcul.
Philosophia Scientiae, 16 :36–67, 2012.

K. Popper.
Objective Knowledge.
Oxford : Clarendon Pr, 1972.

A. Pouget, J.M. Beck, W.J. Ma, and P.E. Latham.
Probabilistic brains : knowns and unknowns.
Nature Neuroscience, 16 :1170–1178, 2016.

E. Salinas.
Prior and prejudice.
Nature Neuroscience, 14 :943–945, 2011.

L. Sanders.
The probabilistic mind.
Science News, 180 :18, 2011.

G. Shafer.
A mathematical theory of evidence.
Princeton University Press, 1976.

P. Smets.
The transferable belief model and other interpretations of Dempster-Schafer's model, pages 375–383.
Elsevier Science Publishers, 1991.

P. Snow.
On the correctness and reasonableness of cox's theorem for finite domains.
Computational Intelligence, 17 :178–192, 1998.

A. Terenin and D. Draper.
Rigorizing and extending the cox-jaynes derivation of probability : Implications for statistical practice.
Submitted. arXiv :1507.06597, 2015.

E.R.W. Tredger, J.T.H. Lo, S. Haria, H.H.K. Lay, N. Bonello, B. Hlavka, and C. Scullion.
Bias, guess and expert judgement in actuarial work.
British Actuarial Journal, 21 :545–578, 2016.

K.S. Van Horn.
Constructing a logic of plausible inference : a guide to cox's theorem.
International Journal of Approximate Reasoning, 34 :3–24, 2003.

P. Walley.
Measures of uncertainty in expert systems.
Artificial Intelligence, 83 :1–58, 1996.

B. Weinstein.
What is an expert ?
Theoretical Medicine, 14 :57–73, 1993.