Deep Learning in Simulative Sciences.

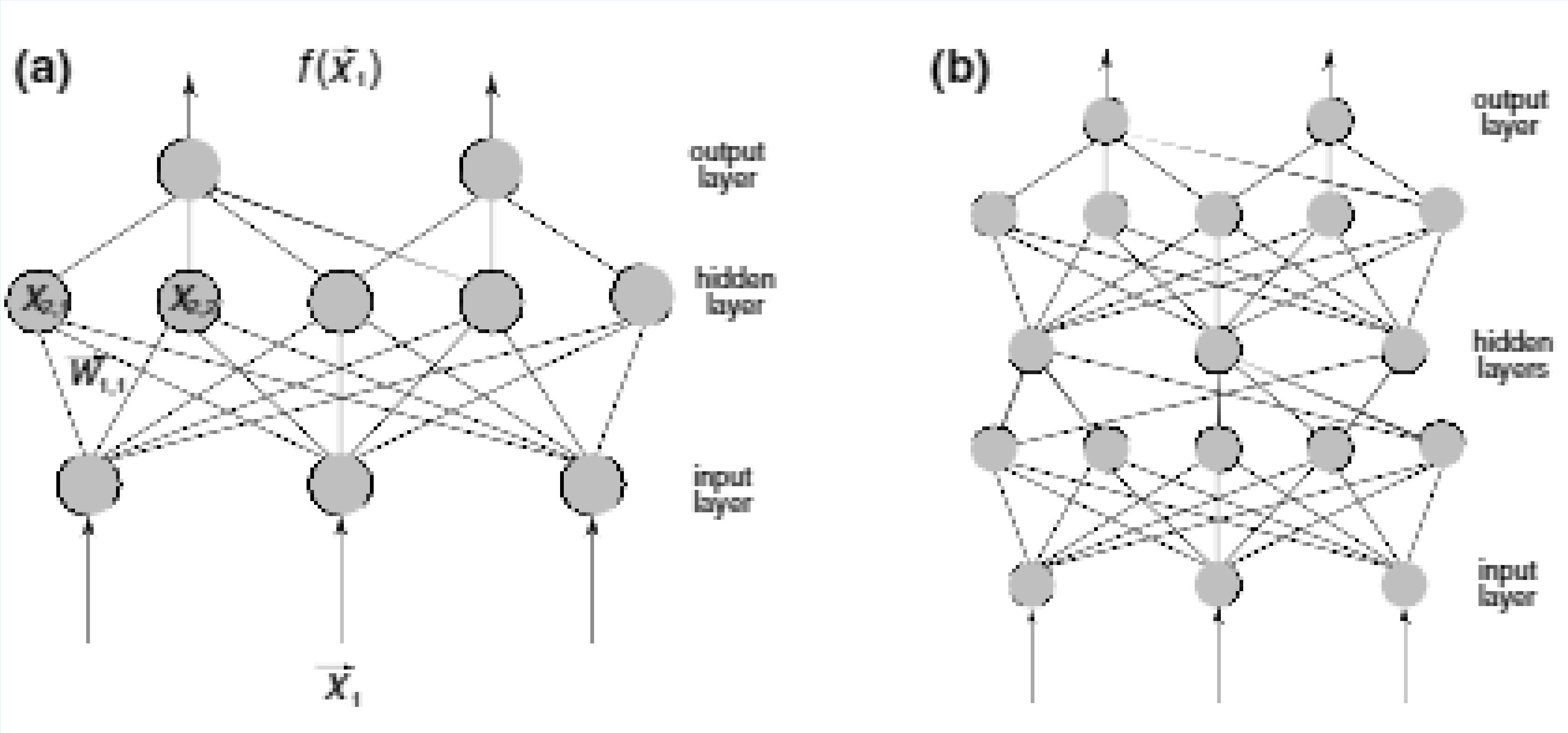
The case of the Met-Net2 model

Lucia Guerrisi

Department of Cognitive Science
University of Messina,
lucia.guerrisi@unime.it

SMARTEST 1st workshop, Noto - September 10, 2024

- The DL allows to overcome the traditionally equation-based methods. Given the growing usage of Deep Learning (DL)modelling in scientific contexts traditionally considered simulative:
- Examine whether and to what extent DL modelling can be acknowledged, from an epistemological and methodological point of view, as a **simulative method**.
- Analyse the *reliability* of DL as simulative method



The major difference between Artificial Neural Networks(ANN) **(a)** and DL (**b**) is the use of more than three layers of neurons.

Thanks to

Geoffrey Hinton

and

Ruslan Salakhutdinov

ANN

↓

studying the human brain – modeling cognitive processes

DL

↓

focus on practical applications – separation from biological neurons

Deep Learning in Science

- The convolutional DL model used for simulate **medical images**(Shen et al., 2017)
- The **predictive maintenance** of the industrial plants (Liu, 2021 – Serradilla, 2022)
- The **MetNet-2** (Meteorological Neural Network 2) DL model, the first featuring a forecasting range of up to 12 hours of lead time (Espeholt et al., 2021)

The simulative method (broad sense)

In a *broad sense*, a *simulation* involves the process of:

▶ *designing* a mathematical model;

▶ *translating* the mathematical model into an algorithm;

▶ *implementing* the algorithm in a computational model;

▶ *executing* the computational model on a machine, the latter often referred to as the artefact;

▶ *studying* the resulting outputs (usually through visualization).

MetNet-2 and NWP (numerical weather prediction) represent two different approaches to weather forecasting

Numerical Weather Prediction models (NWP) consists in physics-based simulations:

▶ the mathematical model consists in a set of differential equations involving the desired atmospheric variables;

▶ the computational model is a computer program aimed at providing numerical solutions to the mathematical model;

▶ the computational model is implemented in supercomputers, here the artefacts, dislocated around the globe which run continuously and frequently update the forecasts given the latest observation.

The MetNet-2 model

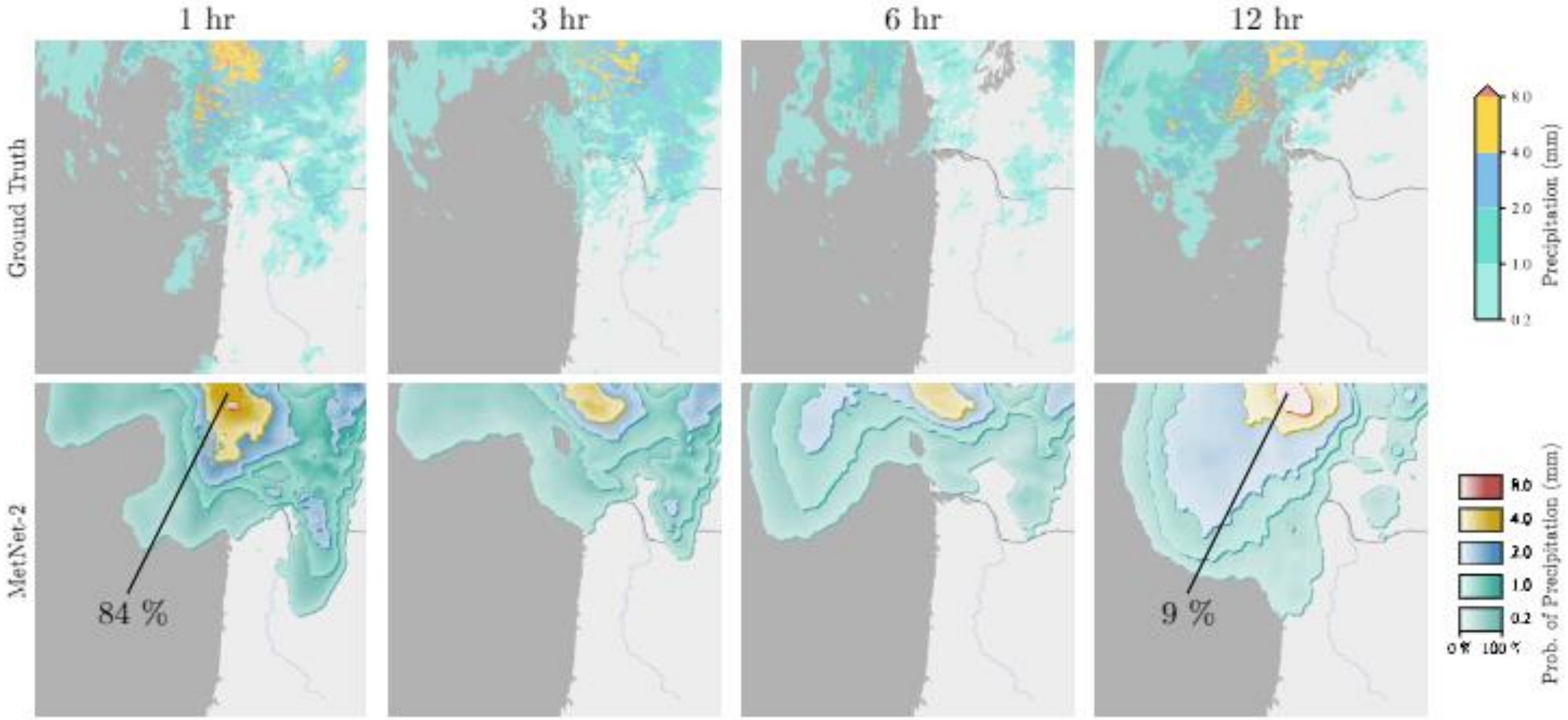
- The Meteorological Neural Network 2 (MetNet-2) is a probabilistic weather forecast model built using DL networks:

▶ MetNet-2 allows precipitation forecasting of 12 hours of lead time, with a 2 minute update frequency.

▶ The training dataset is provided by precipitation measure sobtained through radars dislocated in the Continental United States (CONUS).

The MetNet-2 Model Architecture

- Input encoder
- Second Convolutional Network
- Crop and Predictions



The varying intensities correspond to MetNet-2's predicted probability for the respective rate of precipitation showing the probabilistic structure of the forecast.

MetNet-2 Model as a Simulative Method (**Broad Sense**)

- Designing a DL model with its architecture (**the mathematical model**).
- Implementing the DL model into a high-level programming language (**the computational model**).
- The MetNet-2 model is executed on a computer using a specialized hardware setup: a 16-parallel TPU processor, which is **the artefact** used for running the Deep Learning program.
- Studying the resultant data (**through visualization**).

Is There a Real Simulation?

- In the traditional **NWP model**, the initial values of the involved variables represent the system at time t; the computed solutions to the dynamical system allow for examining how the system evolves over time.

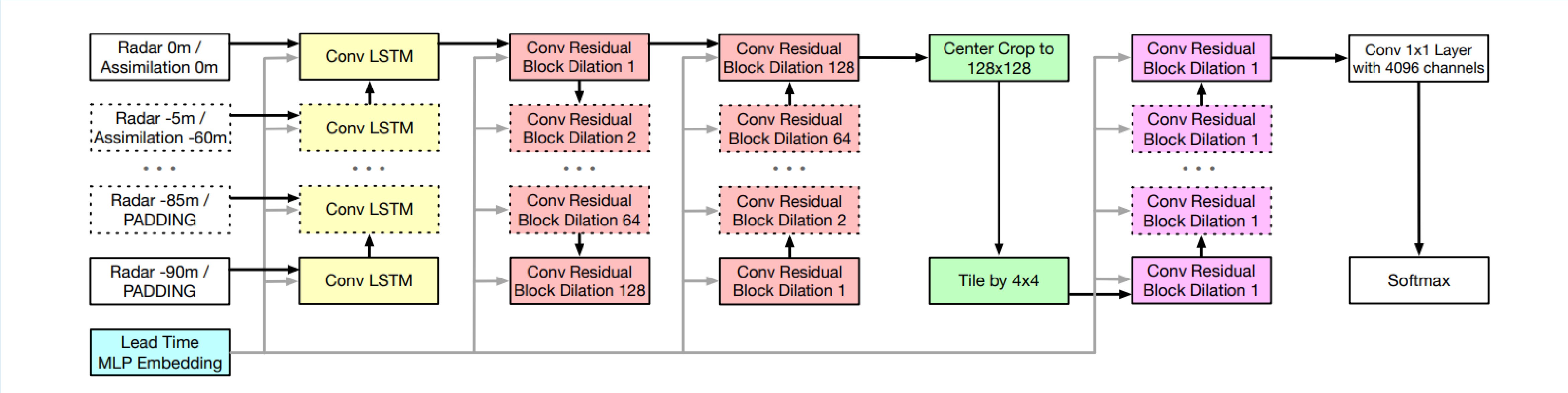
- In the **MetNet-2 model**, no system evolution is represented; rather, the network learns the relationships between initial and final values of the involved variables from the training dataset.

The simulative method (narrow sense)

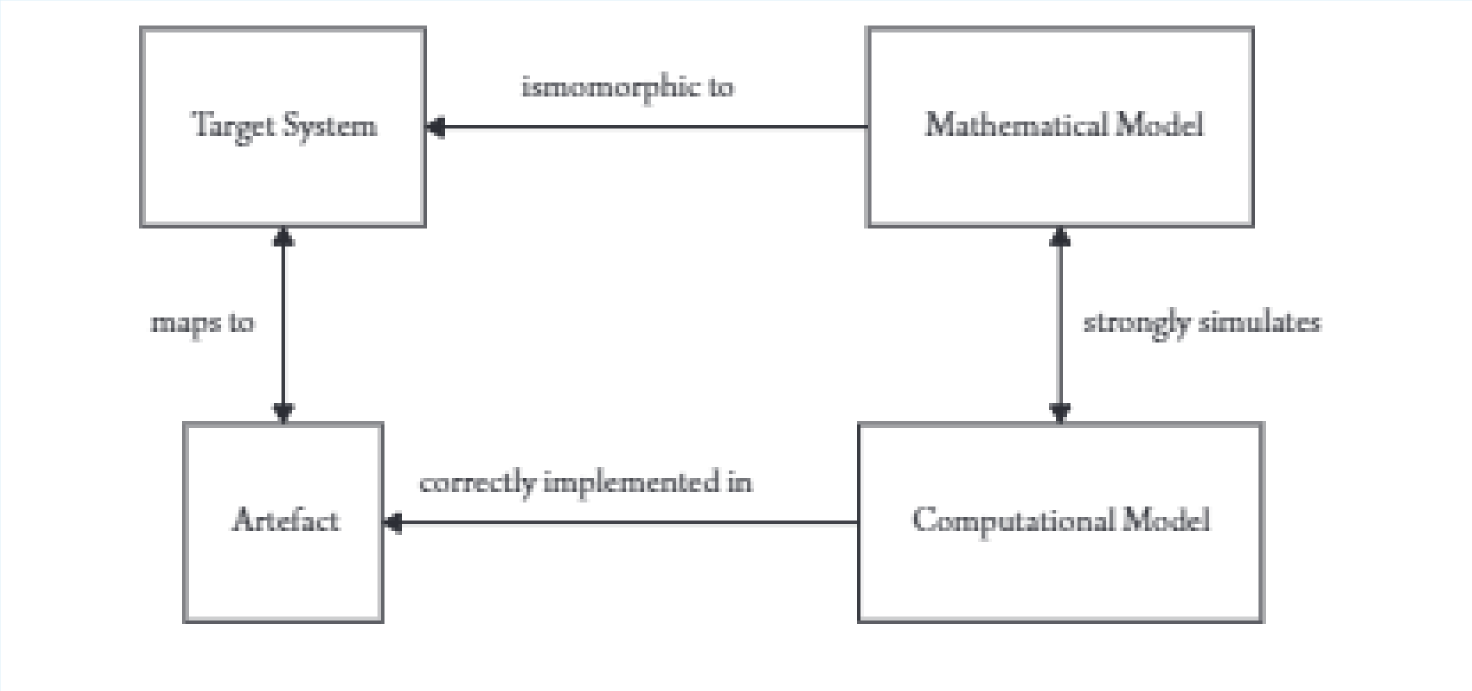
- *A system S provides a core simulation of an object or process B just in case S is a concrete computational devise that produces, via a temporal process, solutions to a computational model [...] that correctly represents B, either dynamically or statically. If in addition the computational model used by S correctly represents the structure of the real system R, then S provides a core simulation of system R with respect to B* (Humphreys 2004).

MetNet-2 as a *Strong Simulation*?

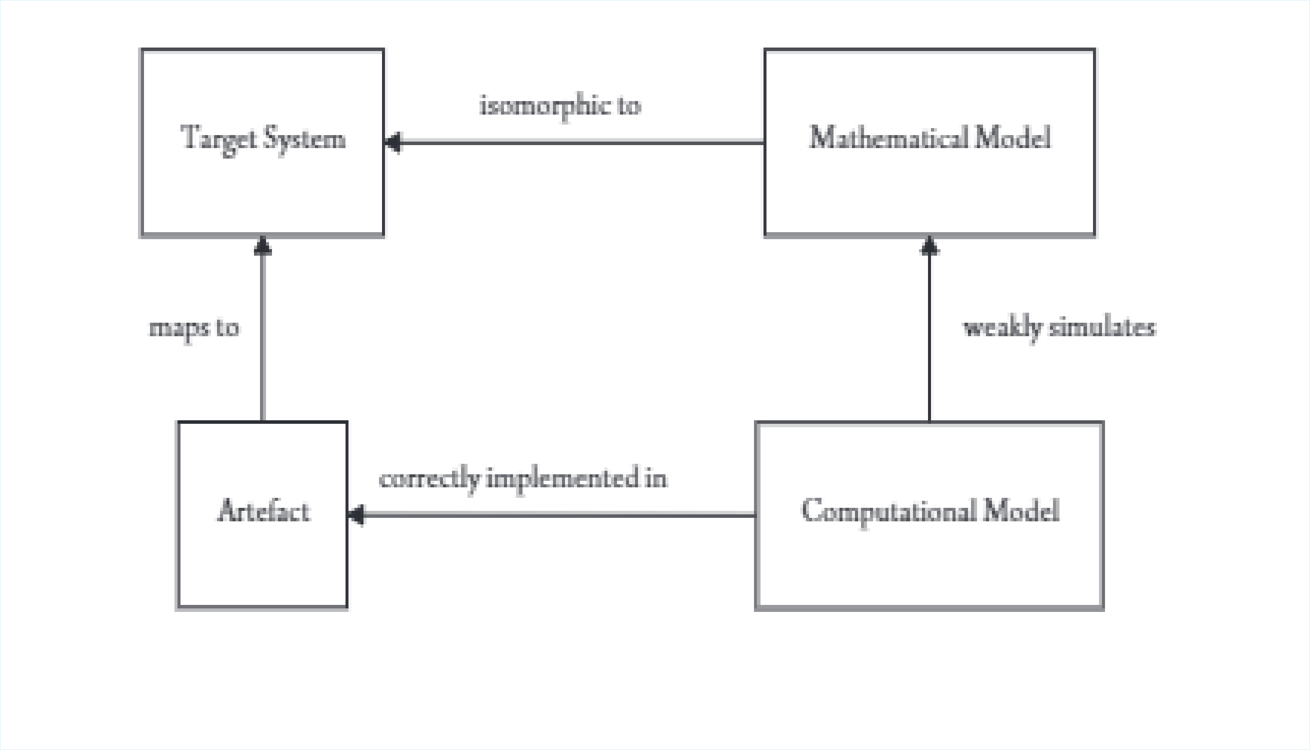
Can we talk about *isomorphism*?



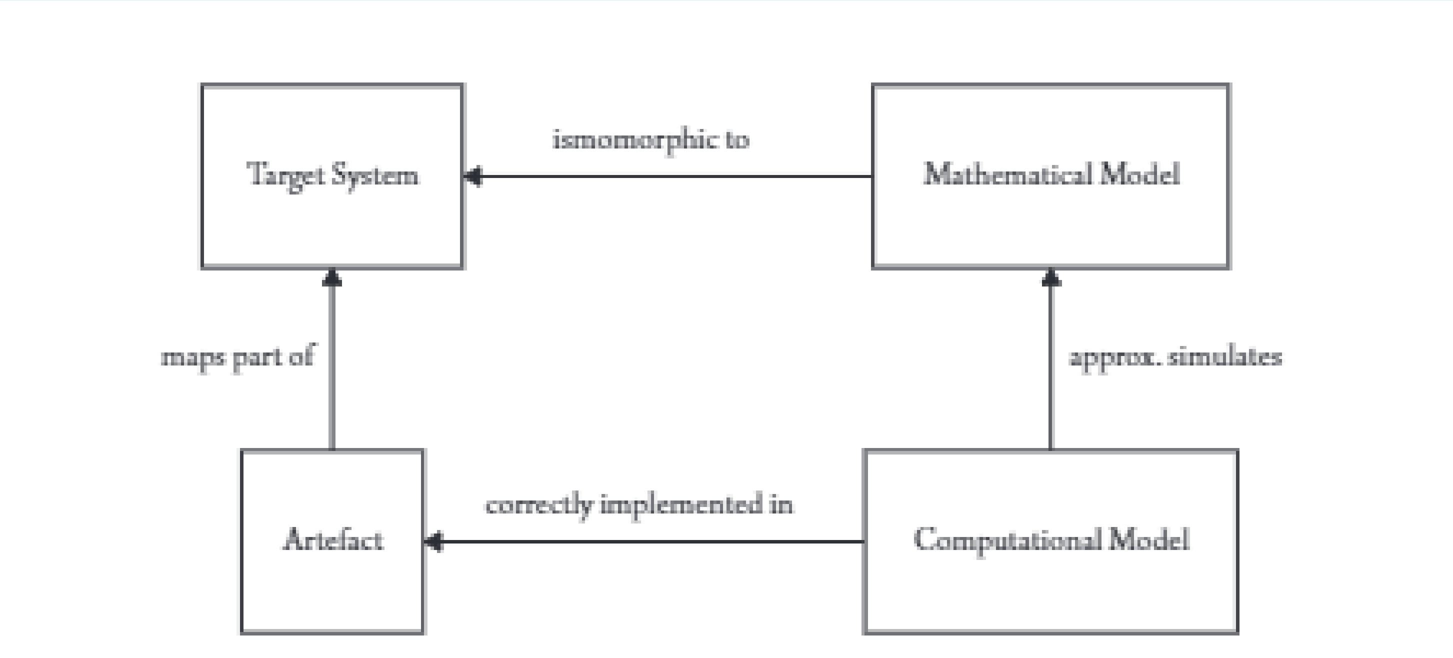
Strong simulation (from Primiero, 2020)



Weak simulation



Approximate simulation



Mathematical and Computational MetNet-2 Model

- The mathematical model in DL simulations already includes its algorithms. The computational model is simply an **implementation** of the model in a high-level programming language (Python).
- Tractability is no longer a primary concern in DL. Computational models are **not approximated** to avoid tractability limitations (as extensively examined by Humphreys 2004).
- The focus in DL simulations shifts to the relationship between the **mathematical model** and the **target system.**

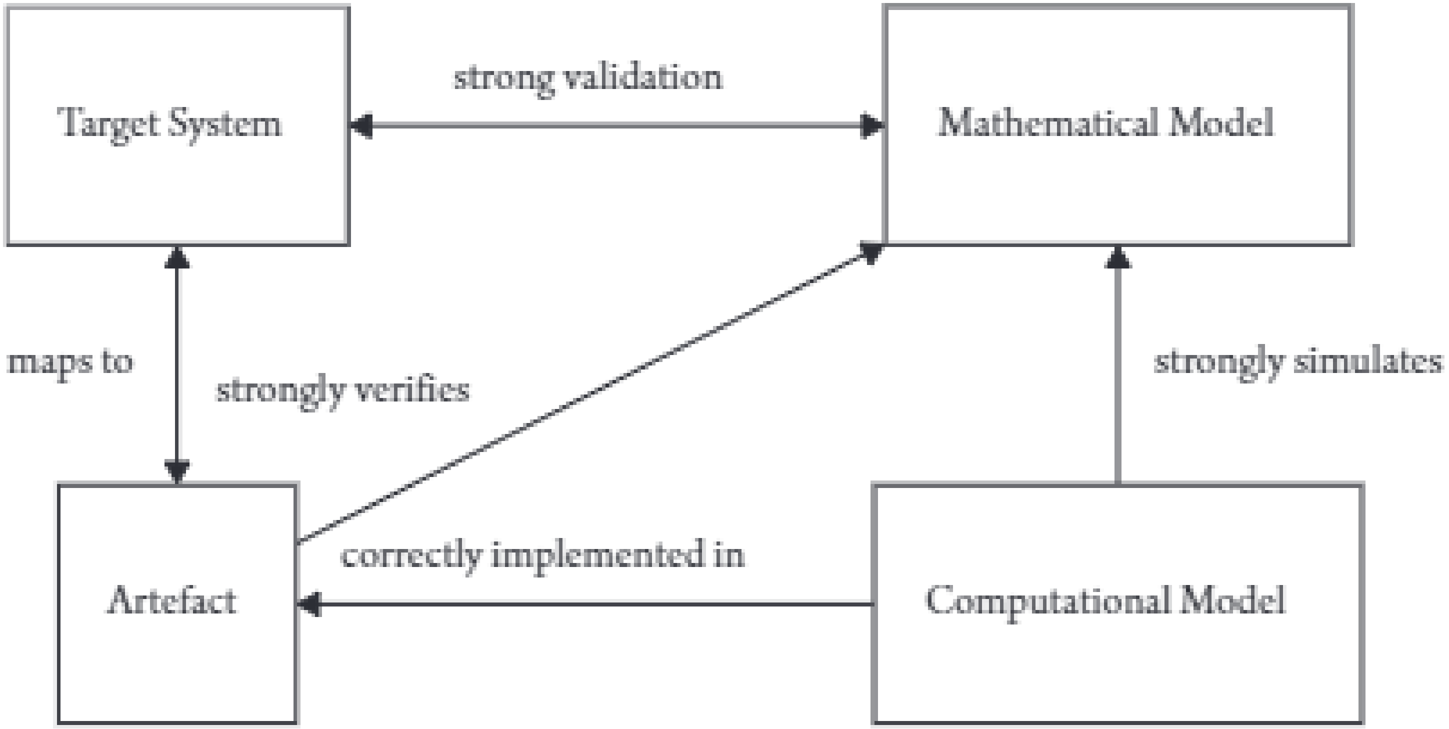
Relationship between Mathematical and Computational Models in Deep Learning (DL)

- No Approximation between Models
- Direct Implementation - pre-defined algorithms

Python/TensorFlow/Keras Implementation:

- **Layers**: defined using `Keras.layer`
- **Activation Functions and Optimizers**: direct implementations of mathematical functions.
- **Backpropagation**: automatically handled by TensorFlow

Validation and Verification



Approximate Validation/Verification of MetNet2 models

How *the mathematical model* relates to *the target system*?

- Improving on training data
- Performing well on new data
- **Overfitting** : in model fails on new data

Traditional verification: predefined functions

DL model verification: functions learned during training

Use Explainability Techniques (XAI) for approximate validation

- **Integrated Gradients** is an XAI technique that helps understand the influence of variables on the model's predictions.
- **MetNet-2 model** - XAI methods revealed that **absolute vorticity**
- This finding supports the **quasi-geostrophic theory**

Conclutions

The usage of DL modelling in scientific contexts can be considered an **approximate simulative** method in all those cases wherein structural similarities can be identified through the interpretation of the network.

References

- Choudhary, A., Lindner, J. F., Holliday, E. G., Miller, S. T., Sinha, S., and Ditto, W. L. (2020). Physics-enhanced neural networks learn order and chaos. Psychonomic Bulletin & Review, 27:217-236.
- Espeholt, L., Agrawal, S., S londerby, C., Kumar, M., Heek, J., Bromberg, C., Gazen, C., Hickey, J., Bell, A., and Kalchbrenner, N. (2021). Skillful twelve hour precipitation forecasts using large context neural networks. arXiv, abs/2111.07472.
- Humphreys, P. (2004). Extending ourselves: Computational science, empiricism, and scientific c method. Oxford University Press, Oxford (UK).
- Primiero, G. (2020). On the foundations of computing. Oxford University PressShen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, *19*(1), 221-248.
- Serradilla, O., Zugasti, E., Rodriguez, J., & Zurutuza, U. (2022). Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. *Applied Intelligence*, *52*(10), 10934-10964.
- Liu, Y., Ding, K., Zhang, J., Li, Y., Yang, Z., Zheng, W., & Chen, X. (2021). Fault diagnosis approach for photovoltaic array based on the stacked auto-encoder and clustering with IV curves. *Energy Conversion and Management*, *245*, 114603.