Bridging Physics and Learning: application to ocean dynamics

R. Fablet et al.
ronan.fablet@imt-atlantique.fr
web: rfablet.github.io

Webinar IMT Data & AI, July 2020
Why the ocean(s) as sandbox?

ML/DL & Physics

Focus on Learning variational models for inverse problems

Beyond Ocean Dynamics
Context: No observation / simulation system to resolve all scales and processes simultaneously

General question: how could data-driven/learning-based tools contribute to solving sampling gaps and higher-level information?
Illustration of satellite-derived sea surface observations

- Altimeters
- Radiometers
- Ocean Colour
- SWOT
- SAR
- SMOS
Deepwater horizon [2010]
What about AI to solve sampling gaps and infer higher-level information?
Context: Data-driven and learning-based approaches for ocean monitoring & surveillance
Learning & Geoscience: nothing new?

Empirical Orthogonal Functions and Statistical Weather Prediction

by EDWARD N. LORENZ

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF METEOROLOGY
Cambridge, Massachusetts

DECEMBER 1956

Scientific Report No. 1
STATISTICAL FORECASTING PROJECT

EDWARD N. LORENZ

Determined Flow

Analogos / Nearest-neighbors

EOF/PCA
Learning & Geoscience: Data-driven approaches for data assimilation

The analog data assimilation \[\text{[Lguensat et al., 2017]}\]
- Combination of analog forecasting strategies and EnKF assimilation schemes
- Extension to 2D+t geophysical dynamics

Open questions
- Bridging model-driven and data-driven paradigms
- Learning data-driven representations from real observation data
Bridging ML/DL paradigms and Physics?
Bridging Physics & AI: Expected breakthroughs

Model-data synergies

AI toolbox
- Machine learning
- Reinforcement learning
- Automatic differentiation
- GPU

Smart observing systems
Bridging Physics & AI: Expected breakthroughs

Physical model
\( \frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u \)

Representation learning

AI toolbox
- Machine learning (Deep learning)
- Computational acceleration
- Big data
- Automatic differentiation
- GPU

Model-data synergies

Data-driven representation

Smart observing systems
Direct applications of DL schemes to physics-related issues

Discovering governing equations from data by sparse identification of nonlinear dynamical systems

Deep learning at scale for the construction of galaxy catalogs in the Dark Energy Survey

Classification of the global Sentinel-1 SAR vignettes for ocean surface process studies

Deep learning to represent subgrid processes in climate models
Bridging physics & AI: Expected breakthroughs

Physical model
\[ \frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u \]

Representation learning

Data-driven representation

Making the most of AI and Physics Theory

- Model-Driven/Theory-Guided & Data-Constrained schemes
- Data-Driven & Physically-Sound schemes (eg, Ouala et al., 2019)
Bridging physics & AI: Expected breakthroughs

Physical model
\[
\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u
\]

Representation learning

Data-driven representation

Making the most of AI and Physics Theory

- Model-Driven/Theory-Guided & Data-Constrained schemes
- Data-Driven & Physically-Sound schemes (eg, Oualal et al., 2019)
DL representations for ODEs/PDEs (Neural ODE)

An example: Residual RK4 Bilinear Network \([\text{Fablet et al., 2018}]\)

\[ \frac{dx(t)}{dt} = \sigma (y(t) - x(t)) \]
\[ \frac{dy(t)}{dt} = x(t) (\rho - z(t)) - y(t) \]
\[ \frac{dz(t)}{dt} = x(t) y(t) - \beta z(t) \]

Lorenz-63 equations

<table>
<thead>
<tr>
<th>Noise-free training data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasting time step</td>
</tr>
<tr>
<td>Analog forecasting</td>
</tr>
<tr>
<td>Sparse regression</td>
</tr>
<tr>
<td>MLP</td>
</tr>
<tr>
<td>Bi-NN(4)</td>
</tr>
</tbody>
</table>
NN Generator from Symbolic PDEs (Pannekoucke et al., 2020)

\[ \partial_t u + u \partial_x u = \kappa \partial_x^2 u \]

**Symbolic calculus** (Simpy)

**PDE-GenNet** (keras)

\[
\begin{pmatrix}
\mu_u(t) \\
\Sigma_u(t)
\end{pmatrix} \xrightarrow{ResNet} \begin{pmatrix}
\mu_u(t+1) \\
\Sigma_u(t+1)
\end{pmatrix}
\]

**Generated code**

```
# Example of computation of a derivative
kernel_Du_x_01 = np.array([[0, np.sin(pos[0]*2*np.pi)],
                           [0, 0, 0, 0]],
                          [0, 0, 0, 0, 0, 0, 0, 0])
Du_x_01 = DerivativeFactory((dim, 3), kernel=kernel_Du_x_01, name='Du_x_01')(u)
```

**Uncertainty propagation**

**Ensemble-based prediction**

**NN prediction**
Bridging physics & AI: Expected breakthroughs

\[
\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u
\]

Making the most of AI and Physics Theory

• Model-Driven/Theory-Guided & Data-Constrained schemes
• Data-Driven & Physically-Sound schemes (eg, Ouala et al., 2019)
Dealing with real systems, including Irregularly-sampled, noisy and/or partially-observed systems?
Can we learn directly from observation data?

End-to-end learning from irregularly-sampled data

[Nguyen et al., 2019; Fablet et al., 2019]

**Generic issue:**
Joint identification and inversion

**Dynamical model**
\[ X_t \xrightarrow{\partial_t X = F(X, \xi, t, \theta)} X_{t+1} \]

**Observation model**
\[ Y_t = H(X, \zeta, t, \phi) \]
An example for upcoming SWOT mission

Groundtruth

Proposed NN framework (Fablet et al., 2019)

State-of-the-art operational processing

(From Perez et al., 2018)
End-to-end learning for inverse problems (Fablet et al., 2020)

Model-driven schemes:
\[ \hat{x} = \arg \min_x \left\{ \lambda_1 \| x - y \|_2^2 + \lambda_2 \Phi(x) \right\}_{U_\Phi(x^{(k)}, y, \Omega)} \]

Gradient-based solver (adjoint/Euler-Lagrange method):
\[ x^{(k+1)} = x^{(k)} - \alpha \nabla_x U_\Phi(x^{(k)}, y, \Omega) \]

No control on the reconstruction error
\[ x^{true} \neq \arg \min_x U_\Phi(x^{(k)}, y, \Omega) \]

Variational cost for the true state
End-to-end learning for inverse problems (Fablet et al., 2020)

Model-driven schemes: \( \hat{x} = \arg \min_x \lambda_1 \|x - y\|_\Omega^2 + \lambda_2 \Phi(x) \)

Direct learning for inverse problems: \( \hat{x} = \Psi(y) \)

Examples of CNN architectures: Reaction-Diffusion architectures, ADMM-inspired architectures, ...

Good performance but possibly weak interpretability/generalization capacities of the solution beyond the training cases.
End-to-end learning for inverse problems (Fablet et al., 2020)

Model-driven schemes: 
\[ \hat{x} = \arg \min_x \lambda_1 \|x - y\|_\Omega^2 + \lambda_2 \Phi(x) \]

Direct learning for inverse problems: 
\[ \hat{x} = \Psi(y) \quad y \rightarrow \text{CNN} \rightarrow x \]

Proposed scheme: joint learning of the variational model and solver

- Theoretical bi-level optimization
\[ \arg \min_{\Phi} \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \arg \min_{x_n} U(\Phi, x_n, y_n, \Omega_n) \]
End-to-end learning for inverse problems (Fablet et al., 2020)

Model-driven schemes: \( \hat{x} = \arg \min_x \lambda_1 \|x - y\|^2_{\Omega} + \lambda_2 \Phi(x) \)

Direct learning for inverse problems: \( \hat{x} = \Psi(y) \quad y \rightarrow \text{CNN} \rightarrow x \)

Proposed scheme: joint learning of the variational model and solver

- Theoretical bi-level optimization
  \[ \arg \min_{\Phi} \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \arg \min_{x_n} U_{\Phi}(x_n, y_n, \Omega_n) \]

- Restated with a gradient-based NN solver for inner minimization
  \[ \arg \min_{\Phi, \Gamma} \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \Psi_{\Phi, \Gamma}(x_n^{(0)}, y_n, \Omega_n) \]

Iterative NN solver using automatic differentiation to compute gradient \( \nabla_{x} U_{\Phi}(x^{(k)}, y, \Omega) \)
End-to-end learning for inverse problems (Fablet et al., 2020)

Proposed scheme: associated NN architecture

- **Initial state** \( x^{(0)} \)
- **Observation data** \( y, \Omega \)
- **ResNet architecture**
  - **Residual Unit (RU)**
  - **LSTM or RNN cell**
  - **Automatic differentiation**
  - **NN model for** \( U_\Phi (x, y, \Omega) \)

\[ x^{(1)} \rightarrow RU \rightarrow x^{(2)} \rightarrow RU \rightarrow \ldots \rightarrow RU \rightarrow \tilde{x} \]

\[ \delta^{(k-1)} \rightarrow LSTM or RNN \rightarrow \delta^{(k)} \text{ with } x^{(k)} = x^{(k-1)} + \delta^{(k)} \]

\[ \nabla_x U_\Phi (x^{(k)}, y, \Omega) \]
End-to-end learning for inverse problems (Fablet et al., 2020)

Illustration on Lorenz-96 dynamics (Bilinear ODE)

Non-supervised setting (classic variational minimisation)

Supervised settings
End-to-end learning for inverse problems (Fablet et al., 2020)

Applications to the reconstruction of sea surface current from SWOT data

NB: preliminary results with a fixed-point Solver rather than a gradient-based solver
End-to-end learning for inverse problems (Fablet et al., 2020)

Key messages

• We can bridge DNN and variational models to solve inverse problems

• Learning both variational priors and solvers using groundtruthed (simulation) or observation-only data

• The best model may not be the TRUE one for inverse problems

• Generic formulation/architecture beyond space-time dynamics

Code: https://github.com/CIA-Oceanix
End-to-end learning from real observation data?

**Scarce time sampling**

Ouala, ICASSP’19

**Noisy and irregular sampling**

(From Perez et al., 2018)

Nguyen, ICASSP’20

**Partially-observed system**

Ouala, preprint 2019
Summary

• **NNs as numerical schemes for ODE/PDE/energy-based representations of geophysical flows**

• *Embedding geophysical priors in NN representations* (e.g., Lguensat et al., 2019; Ouala et al., 2019)

• *End-to-end architecture for jointly learning a representation (eg, ODE) and a solver* (e.g., Fablet et al., 2020)

• *Towards stochastic representations embedded in NN architectures* (e.g., Pannekoucke et al., 2020, Nguyen et al., 2020)
Beyond Ocean Dynamics

Learning stochastic hidden dynamics
Learning stochastic hidden dynamics [Nguyen et al., 2018]

The example of AIS Vessel trajectory data

- Millions of AIS positions daily
- Noisy data: irregular sampling, corrupted data

How can we learn from AIS data streams?
Learning stochastic hidden dynamics \cite{Nguyen2018}

Model training from noisy AIS streams using variational Bayesian approximation
Learning stochastic hidden dynamics [Nguyen et al., 2018]

- Vessel type recognition
- ~88% of correct recognition

Trajectory reconstruction

Abnormal behaviour detection
Beyond Ocean Dynamics

Dynamical System Theory for Deep Learning
Understanding DL models?

88% tabby cat → adversarial perturbation → 99% guacamole

Szegedy et al., 2015
Understanding ResNets [Rousseau et al., 2019]

ResNet [He et al., 2015] regarded as space registration machines

- Image registration examples

[Matlab tutorial]

[Dramms tutorial]
Understanding ResNets [Rousseau et al., 2019]

ResNet [He et al., 2015] regarded as space registration machines

Original feature space

Registered space to make feasible a linear separation between classes
AI Chair OceaniX 2020-2024

Physics-informed AI for Observation-Driven Ocean AnalytiX

PI: R. Fablet, Prof. IMT Atlantique, Brest

Internship, PhD and postdoc opportunities

(https://rfablet.github.io/)
Joint work with B. Chapron, F. Collard, L. Drumetz, J. Le Sommer, R. Lguensat, D. Nguyen, S. Ouala, A. Pascual, F. Rousseau, P. Tandeo, J. Verron, O. Pannekoucke, ...

More:
• Webpage: https://rfablet.github.io/
• Preprints: https://www.researchgate.net/profile/Ronan_Fablet

Thank you.