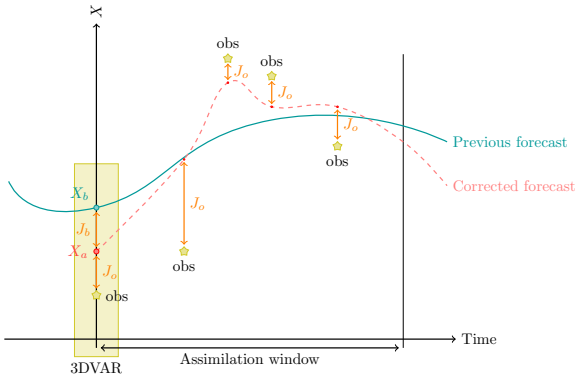
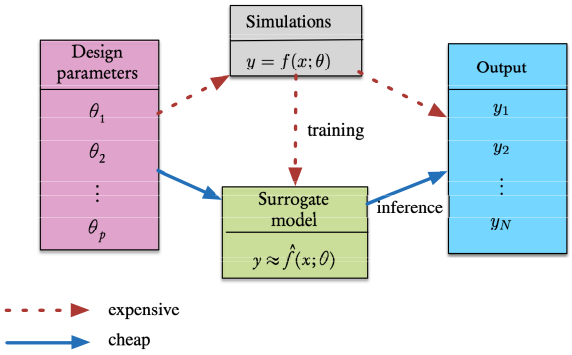
DA and SciML Advanced Course Guide

Mark Asch - CSU/IMU/VLP, Philippines - 2023

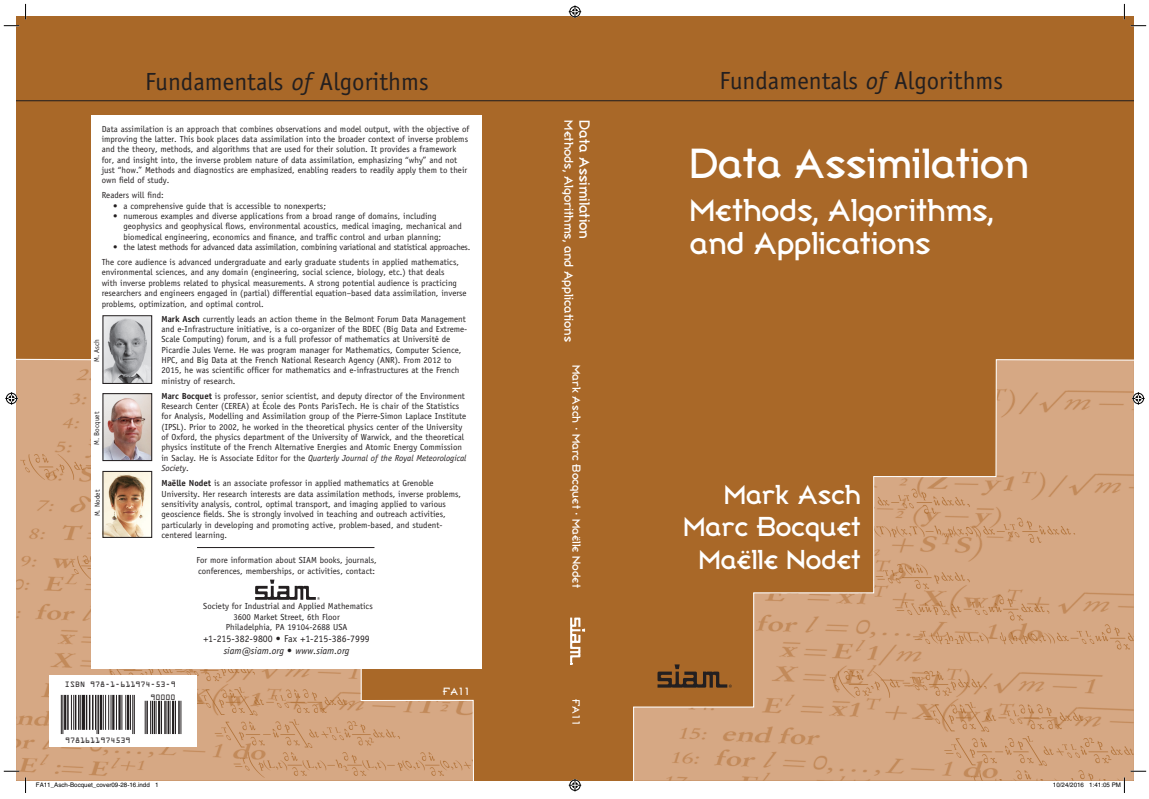




Advanced Course Lecture program

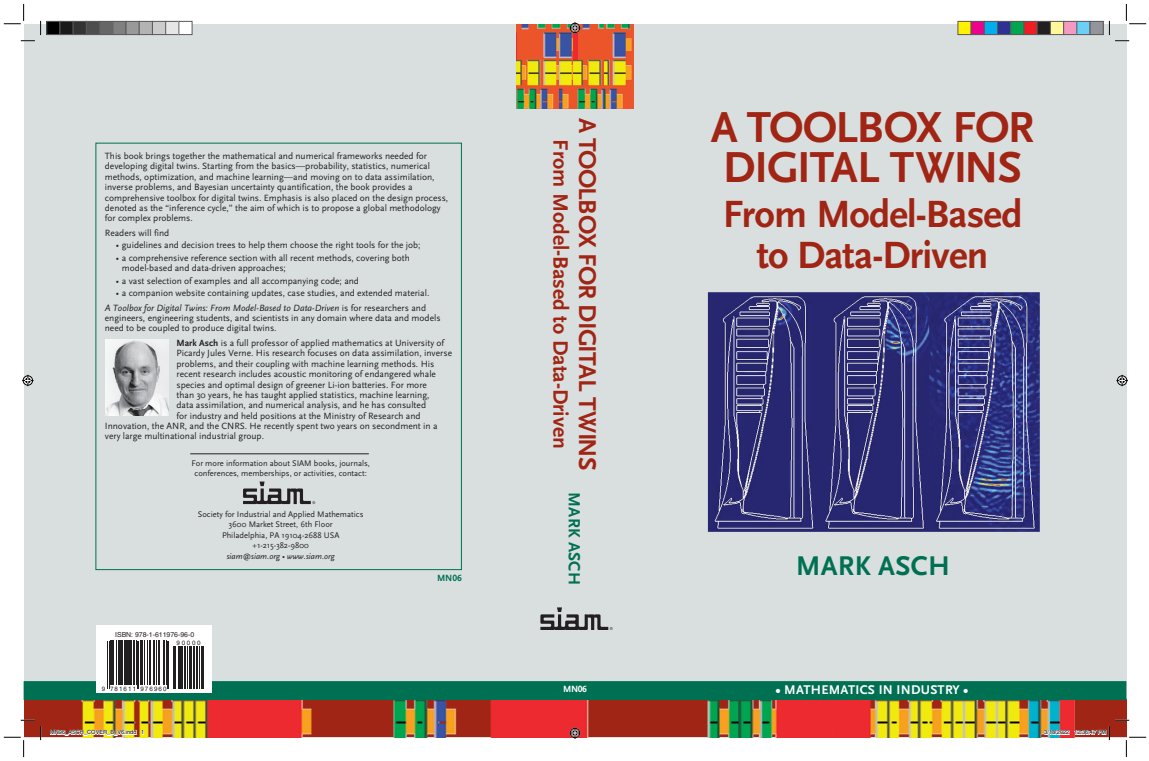
1. Introduction to Scientific Machine Learning (3h)
2. Optimization theory and practice for SciML (3h)
3. Machine Learning methods for SciML (6h)
4. Automatic differentiation for SciML (3h)
5. Scientific Machine Learning approaches (6h)
6. Principles and Ethics of Scientific Machine Learning (3h)
7. Advanced Data Assimilation methods. (6h)

Reference Book I



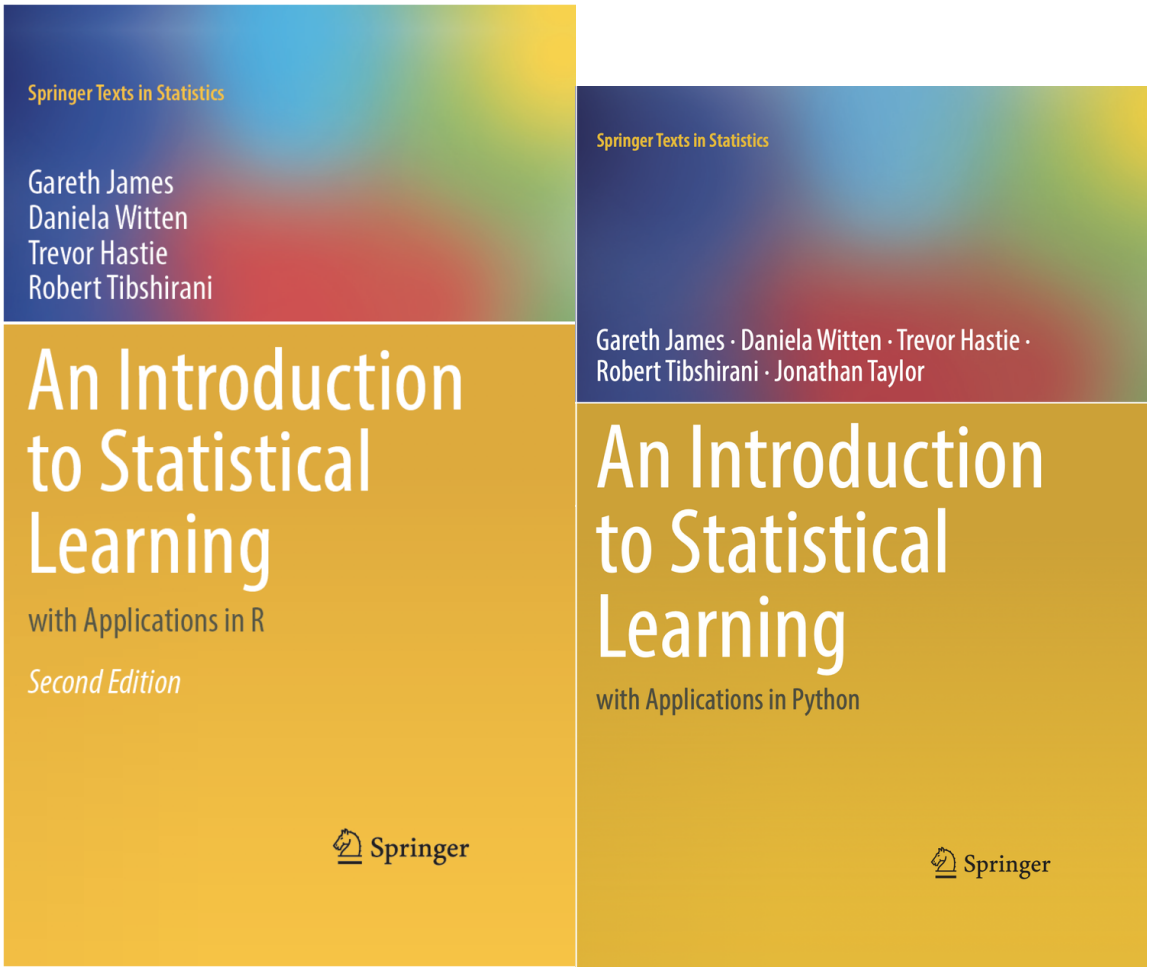
- Extracts available from Google Books...
- Complete Chapters 1 to 3 are provided.

Reference Book II



- Covers ML and DA, but contains a LOT more material
- All software codes available.

Reference Book III



- Covers ML in a concrete, but mathematically solid way.
- Exists in two, freely available versions: R and

Python (see references below)

- All software codes and data available.
- Website: https://www.statlearning.com/

Website

All the lectures and supplementary material can be found at the accompanying website and GitHub pages:

- https://sites.google.com/view/csu2023/
- https://github.com/markasch/CSU-IMU-2023

Note

The website is the basis of this course—please consult it regularly and use <Shift-Reload> to ensure that you have the latest versions of each page.

Examples and Exercises

- every lecture has accommpanying examples that illustrate the contents

- these examples should be used

  $\Rightarrow$ to understand the theory
  $\Rightarrow$ as exercises to learn the effects of modifying the parameters
  $\Rightarrow$ as a basis for yor own research

- there is a CodeLab associated with each lecture that provides a list of examples and pointers to the website and/or github pages

Projects and Prizes

- The final objective of the course: each participant prepares a code project that implements one or more of the scientific machine learning concepts that we have learned, including the issue of ethics.

- This project can take one of the following forms:

  $\Rightarrow$ extending one or more of the examples, or broadening their application
  $\Rightarrow$ attempting to reproduce some known results from a research project or paper
  $\Rightarrow$ trying to apply SciML to a new research topic that has not yet been considered

- During the afternoon workshops, we will discuss the potential topics and guide you in your choice.

- Each project will be presented and dicussed in the final two workshop sessions.

- A jury will award prizes (in the form of textbooks, kindly donated by SIAM) to the best projects.

References

1. https://en.wikipedia.org/wiki/Data_assimilation

2. G. Evensen. *Data assimilation, The Ensemble Kalman Filter*, 2nd ed., Springer, 2009.

3. E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 2003.

4. K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation. A Mathematical Introduction.* Springer, 2015.

5. A. Tarantola. *Inverse problem theory and methods for model parameter estimation.* SIAM. 2005.

6. G. James, D. Witten, T. Hastie, R. Tibshirani. *An Introduction to Statistical Learning with Applications in R.* Springer. 2013.

http://www.statlearning.com
https://hastie.su.domains/ISLR2/ISLRv2_corrected_June_2023.pdf

7. G. James, D. Witten, T. Hastie, R. Tibshirani. *An Introduction to Statistical Learning with Applications in Python.* Springer. 2023. https://hastie.su.domains/ISLP/ISLP_website.pdf

8. Rachel Schutt and Cathy O'Neil. *Doing Data Science.* O'Reilly. 2014.

9. I. Goodfellow, Y. Bengio, A. Courville. *Deep Learning.* MIT Press. 2016.

   http://www.deeplearningbook.org

10. M. Kuhn, K. Johnson. Applied Predictive Modeling. Springer 2018.
    http://appliedpredictivemodeling.com/

Software for DA

Various open-source repositories and codes are available for both academic and operational data assimilation.

1. DARC: `https://research.reading.ac.uk/met-darc/` from Reading, UK.

2. DAPPER: `https://github.com/nansencenter/DAPPER` from Nansen, Norway.

3. DART: `https://dart.ucar.edu/` from NCAR, US, specialized in ensemble DA.

4. OpenDA: `https://www.openda.org/`.

5. Verdandi: `http://verdandi.sourceforge.net/` from INRIA, France.

6. PyDA: `https://github.com/Shady-Ahmed/PyDA`, a Python implementation for academic use.

7. Filterpy: `https://github.com/rlabbe/filterpy`, dedicated to KF variants.

8. EnKF; `https://enkf.nersc.no/`, the original Ensemble KF from Geir Evensen.

Software for ML

1. R:
   https://cran.r-project.org/

2. scikit-learn:
   https://scikit-learn.org/stable/

3. PyTorch:
   https://pytorch.org/get-started/locally/