

## Summary

- Multidisciplinary researcher ("investigator") with expertise in statistical modeling and data-driven simulations.
- Proven track record of delivering scalable solutions for high-impact challenges and actively contributing to cross-functional collaborations.

## Education

### University of North Carolina at Chapel Hill (UNC) | Ph.D. in Statistics and Operations Research

Chapel Hill, NC

Aug 2021 – May 2026

- Advisor: Dr. Vlasdas Pipiras (Qualification Exam: Probability theory, Statistical theory)
- [Cambanis-Hoeffding-Nicholson Award](#) for outstanding academic performance in first-year doctoral program.
- [Korean Government Scholarship for Overseas Study \(\\$80K\)](#): 5 students in Intelligent Infrastructure selected nationwide.

### Seoul National University (SNU) | M.S. in Statistics

Seoul, South Korea

Mar 2019 – Feb 2021

- Advisor: Dr. Hee-Seok Oh

### Seoul National University (SNU) | B.S. in Statistics, Minored in Computer Science and Engineering

Mar 2015 – Feb 2019

- [The Presidential Science Scholarship \(Tuition and incentives\)](#): 24 students in the Mathematics field selected nationwide.

## Work Experience

### Moloco | Software Engineer Research Intern | Seattle, WA

May 2023 – Aug 2023

- Extracted large-scale ML model data using **SQL**. Run simulation, draw inference, and visualize data using **Python**.
- Researched on *exploration strategies and distribution shifts* for real-time bidding price prediction for online ads ([link](#)).

### Lawrence Livermore National Laboratory | NSF Mathematical Sciences Graduate Intern | remote

May 2022 – Jul 2022

- Contributed to [libROM](#), a lightweight, scalable **C++** library for nonlinear dynamical problems. ([1], [3])
- Worked on hyper-reduction sampling in data-driven **physics simulation** applied to diffusion, hyperelasticity, hydrodynamics.

## Research Experience

### Statistical Machine Learning and Methodologies

- **Uncertainty quantification in multi-fidelity system** [3–6] (UNC, 21-24) **Probability density function estimation** of limited high-fidelity outputs by leveraging (cheap) surrogate models through (i) nonparametric kernel density estimation with *importance sampling* [4] and (ii) a parametric ensemble estimator approach [3], with emphasis on extreme outputs. Simulations via **R**.

### Spatio-temporal Modeling in Latent Space, Time-series Analysis

- **Physics-informed neural networks (PINNs)** [2] (UNC, '24) Developed reduced-order modeling for PINNs to approximate PDE solutions. Modeled latent-state dynamics with neural ODEs, employed *implicit neural representation* as a **decoder** and applied insights from 3D graphics to impose boundary conditions. AI research with **PyTorch** and **Python**, **HPC**, **GPU clusters**.
- **Spatio-temporal analysis of particular matter** (SNU, '21) Applied dynamic factor model and the EM algorithm for forecasting, missing data imputation, and [quantile analysis](#). Latent state modeling with linear projection and time-series model.
- **Multi-scale time-series clustering** [7] (SNU, 19–20) applied to step count data from wearable devices and COVID-19 data.

### Course Projects at UNC / Undergraduate Research at SNU

- [Generative diffusion models](#) for Advanced Machine Learning, [Talk](#) for High-Dimensional Time-Series class. (UNC 23-24)
- **MRI Stroke Lesion Segmentation** (SNU, '18) (SNU, '18) Used **ensemble** of FCN and **U-Net** architecture via **Keras** and **Python**.

## Teaching Experience

### Instructor UNC: STOR 155 Introduction to Data Models and Inference (24F)

### Teaching Assistant UNC: Introduction to Deep Learning (22Sp, Led labs on **SVM**, **CNN**, **LSTM**, **VAE**, **GAN**, etc), Methods of Data Analysis (22Sp, 21F), SNU: Sampling Design and Survey (20F), Design and Analysis of Experiments (20Sp), Lab (20Sp, 19F)

## Publications

- [1] S. Cheung et al., 'Survey on quadrature point selection in hyper-reduced order models for finite element methods' (In prep).
- [2] M. Kim, T. Wen, K. Lee, Y. Choi, (2024), '[Physics-informed reduced order model with conditional neural fields](#)', *NeurIPS 2024 Workshop on Machine Learning and the Physical Sciences*. **keywords**: implicit neural representation, physics-informed neural network, autoencoder
- [3] M. Kim, B. Brown, V. Pipiras, '[Parametric multi-fidelity Monte Carlo estimation with applications to extremes](#)', submitted to *Technometrics*.
- [4] M. Kim, K. O'Connor, V. Pipiras, T. Sapsis, (2024+), '[Sampling low-fidelity outputs for estimation of high-fidelity density and its tails](#)', *SIAM/ASA Journal on Uncertainty Quantification*. **keywords**: kernel density estimation, importance sampling, extreme value theory
- [5] M. Kim, V. Pipiras, T. Sapsis. (2024), 'Statistical Reduced-Order Modeling of Peaks of Vertical Bending Moment in Irregular Waves', *Proceedings of the 35th Symposium on Naval Hydrodynamics (SNH)*, Nates, France.
- [6] M. Kim, V. Pipiras, A. Reed, K. Weems, (2023), '[Calibration of low-fidelity ship motion programs through regressions of high-fidelity forces](#)', *Ocean Engineering* **290**, 116321.
- [7] M. Kim, H. Oh, and Y. Lim, (2023), '[Zero-Inflated Time-Series Clustering Via Ensemble Thick-Pen Transform](#)', *Journal of Classification* **40**.

## Technologies

**Programming and Tools**: Python, C++, Julia, R, SQL, Git, Bash, slurm, CUDA. **Machine Learning**: PyTorch, TensorFlow, dplyr, NumPy, ggplot2.