Normalizing Flow-Enhanced Variational Autoencoder with Neural ODEs for Population Modeling

Marcus Baaz*[1], Anders Sjöberg*[1,2], Mats Jirstrand[1,2]

1) Fraunhofer-Chalmers Centre, Gothenburg, Sweden
2) Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden
*Equal contribution, Marcus.Baaz@fcc.chalmers.se

Population modeling aims to mathematically represent and predict how individual variability shapes broader population dynamics. In this context, population modeling often deals with time-series data from multiple subjects who share the same model structure, but whose individual parameters are assumed to be samples from a probability distribution. In pharmacometrics, this takes the form of nonlinear mixed-effects (NLME) models, which capture both fixed population-level effects and random individual deviations amid complex nonlinearities and noisy data. The estimation of NLME parameters can be viewed as a form of bilevel optimization, where both population and individual-level objectives are intertwined. A widely used method for this is the Stochastic Approximation Expectation-Maximization (SAEM) algorithm, which iteratively combines stochastic simulation of latent variables with deterministic maximization steps, enabling efficient navigation of complex likelihood surfaces where direct optimization is intractable.

Machine learning, especially neural networks, is increasingly used to improve flexibility and predictive power in population modeling. A common approach integrates them into existing SAEM-based software. However, estimating large networks with SAEM is challenging due to repeated Markov chain Monte Carlo (MCMC) sampling and high computational and memory demands, making it inefficient or impractical at scale. Alternative strategies are therefore needed to incorporate neural architectures into population models.

To address this challenge, we propose a variational autoencoder (VAE) framework with normalizing flows for modeling population data. This approach decouples the estimation process from the SAEM algorithm by leveraging amortized inference, which replaces per-subject sampling with a learned encoder network. In our architecture, the encoder is a transformer model that captures individual variability by processing longitudinal data and mapping it into a structured latent space. The decoder is a neural ordinary differential equation (NODE) that models the temporal dynamics of the system, conditioned on the learned latent variables. The normalizing flow component enhances the expressiveness of the posterior approximation, allowing us to better capture complex latent structures and inter-individual variability.

On simulated data, our model demonstrates predictive accuracy comparable to that of a standard autoencoder. However, by learning a structured, probabilistic latent space, our approach additionally captures individual variability and population-level structure in a principled way. This allows the model to generalize effectively to unseen data and provides richer insight into inter-individual differences, even in the presence of noise and temporal complexity.