## Abstract

Here is the abstract.

## Introduction

Kinetic models consist of ordinary differential equations (ODEs) that specify the time-based evolution of biological species. For large biological models, parameter estimation can prove to be challenging for multiple reasons. The non-convex nature of the parameter space makes the search of a global minima computationally intensive. Most biological ODEs are stiff equations where many computational integrators can struggle to perform integrations, resulting in failed parameter searches. Recent work has established a multi-start local gradient search as one of the most robust strategies for parameter estimation, but even such methods require large computational times for an exhaustive parameter search due to the problems listed above (17). test[1] . test[1,2]

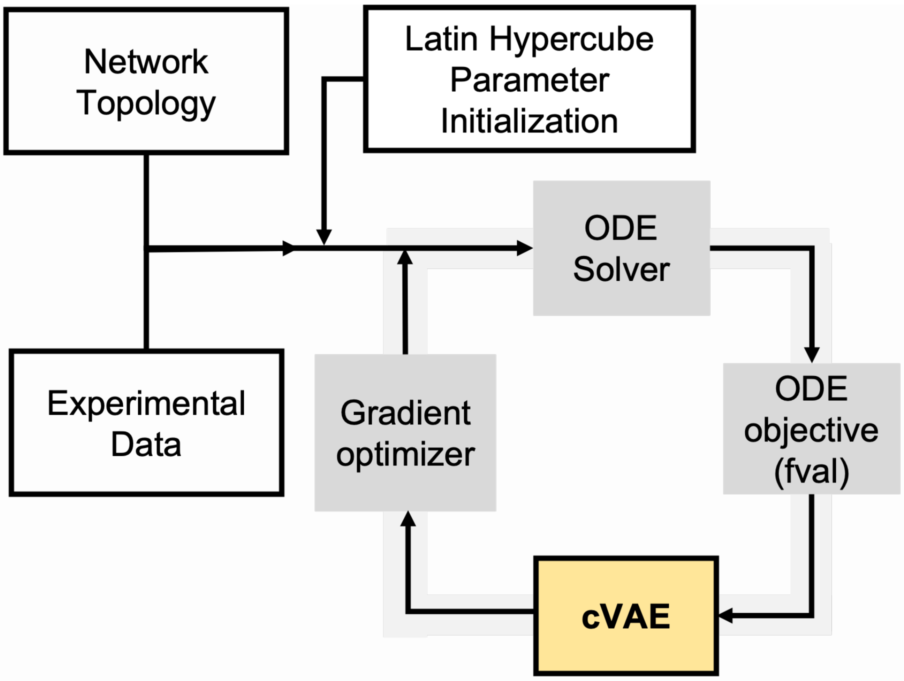
We explore a specific application of machine learning in metabolic modeling, namely the use of a conditional variational autoencoder in assisting the parameter estimation process of a biological kinetic model. A variational autoencoder (VAE) is a form of autoencoder that generates a latent space from trained embeddings of input vectors (10). The latent space allows the sampling of new embeddings which can be decoded through the decoder to provide new data that retains the properties of the original input data. Thus, VAEs can serve as a form of generator models, which has been shown to be beneficial in various biological applications (11–14). Conditional VAEs (cVAEs) are a modification to the VAE architecture wherein a conditional variable is also provided with the input data to modulate the latent space generation wherein each input is conditioned along with the accompanying label (15, 16). This allows the generation of new data where the accompanying properties can be controlled with a desired label. This feature of cVAE was exploited in the specific problem of parameter estimation of a kinetic model.

In this study, we present a novel application of the cVAE architecture to improve the kinetic model parameter estimation process when used in conjunction with existing search methods. We trained a cVAE on parameter vector search data generated from the parameter estimation process for a metabolic kientic model. The objective of such an application of cVAEs was to train a model that can learn the parameter landscape sufficiently to then reliably suggest better parameter vectors for kinetic models. Each parameter vector encountered during the multi-start local gradient search was provided as training input to the cVAE along with a conditional label. The label in this case was the value of the objective function used in kinetic model training, which captures the mismatch between a model's predictions and the experimental measurements it is being trained on as a weighted sum of squares. Upon training the cVAE, we observed a latent space with well-formed clusters, where high and low values of the objective function were particularly well separated. To test the capability of the trained cVAE in assisting new parameter estimation efforts, we generated novel parameter vectors from sampling the latent space for a range of conditional labels (which represented the objective function value) and computed the objective function again for each of them. The objective function values for the low conditional label were remarkably low with very few generated vectors proving to be stiff. On the other hand, parameter vectors generated for the high conditional labels resulted in higher objective values with a larger fraction of these simulations failing due to stiffness in the vectors. Our results suggest that a well-trained cVAE can serve as a better recommender for kinetic model parameters than current multi-start local searches which often rely on uniform sampling methods such as Latin Hypercube sampling (17).

# Results

## An iterative parameter estimation workflow using generative models

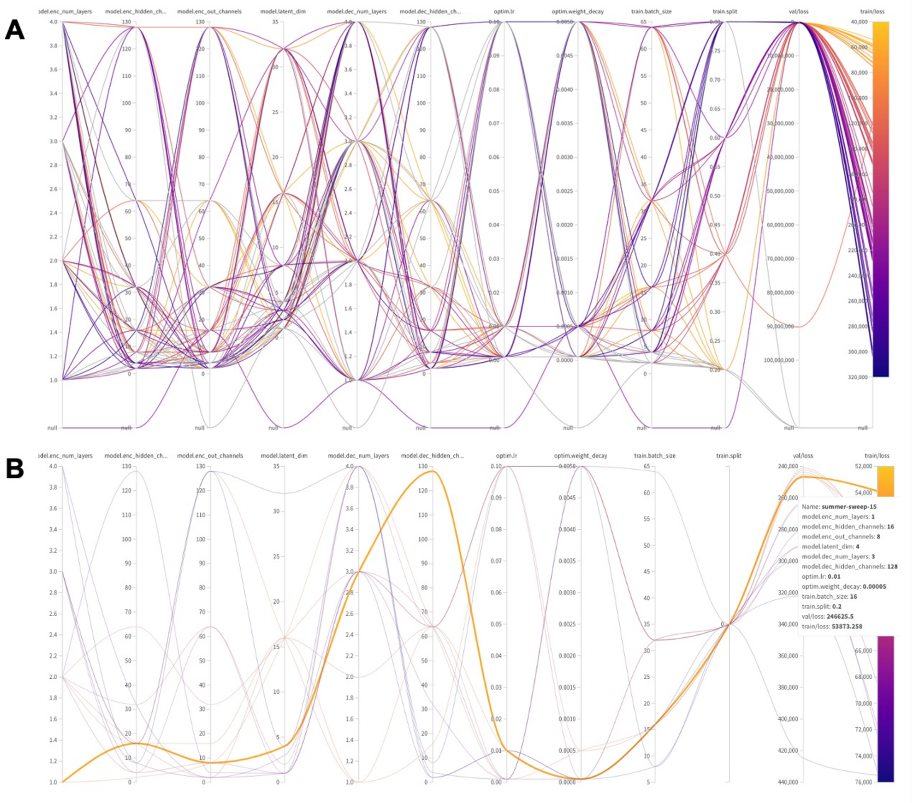
We trained a cVAE model on parameter vectors obtained from the multi-start local gradient search method. The multi-start method employed Latin Hypercube sampling to generate guess vectors, and parallel gradient searches were started from each such guess vector till a termination condition was encountered. The parameter vectors in the training data were used to integrate the differential equations of the kinetic model in time. The result of each such integration was a set of predicted values of the model variables as a function of time. The mismatch of predicted and measured values was captured in the value of an objective function, here the negative log likelihood function. This value was provided as a conditional label along with the parameter vectors themselves when training the cVAE. The parameter estimation process which includes the cVAE training is shown in Figure 1. The architecture of the cVAE model is shown in Figure S1.



cVAE iterative workflow

## Hyperparameter optimization sweeps (Supplementary Material?)

Hyperparameter optimization was performed on the cVAE training process, which included model architecture parameters such as number of layers in the encoder and decoder, latent space dimension, training split ratios, learning rates etc. to identify the best combination of hyperparameters for the model. The results of all the optimization tests are shown as a parallel lines plot in Figure S2-A, with the best combination of parameters displayed as the highlighted yellow line in Figure S2-B.



Name: **summer-sweep-15**
model.enc_num_layers: **1**
model.enc_hidden_channels: **16**
model.enc_out_channels: **8**
model.latent_dim: **4**
model.dec_num_layers: **3**
model.dec_hidden_channels: **128**
optim.lr: **0.01**
optim.weight_decay: **0.00005**
train.batch_size: **16**
train.split: **0.2**
val/loss: **246625.5**
train/loss: **53873.258**

Figure S2

## Evaluation/Benchmarking via Small-scaling study

### Hyperparameter optimization sweep

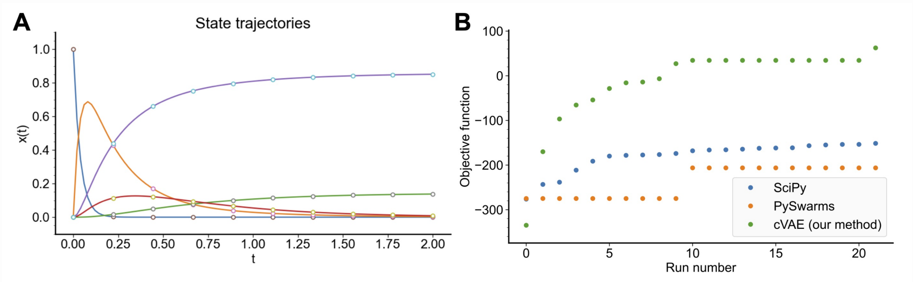
The cVAE method had so far been used to improve parameter estimation for a specific case of the lipid kinetic model. This involved hyperparameter optimization, training and testing of the cVAE model on the data generated during parameter estimation of the lipid model alone. To implement the cVAE pipeline on a broader class of models, the method was employed on a few chosen small-scale models from a suite of SBML models prepared for benchmarking and comparision of optimization algorithms.

Currently, methods of parameter estimation often pursue global optimization methods or hybrid global-local methods. In the former category, popular methods include genetic algorithms and particle swarm-based methods. In the latter, methods are developed that typically combine local gradient searches with global start points. To benchmark our method against contemporary methods, particle swarm was selected for the global optimization method and multi-start local gradient-search for the hybrid method (18, 24).

To perform extensive comparison tests of the three methods in question, a smallsized model of cortisol metabolism from the list of reproducible models was used, consisting of 4 reactions, 5 metabolites and 5 parameters (25). Selecting a model of such dimensions allowed for exhaustive coverage of the parameter space for comparison between the three optimization strategies. The cortisol metabolic model was simulated to steady state and the simulated data of metabolite trends used to represent the ground truth on which all other optimization strategies performed parameter estimation (Figure 4.10A).

While performing parameter estimation with the three chosen strategies, we had to decide on the optimal hyperparameter values and method settings that would lead to the best performance from each of the strategies. For the global method (particle swarm) and hybrid method (multi-start gradient search), we borrowed the method parameters suggested in the libraries accompanying these methods (namely PySwarms and SciPy optimizers). However, as no such heuristic existed for our developed cVAE method, we had to first repeat hyperparameter optimization for our pipeline on the chosen parameter estimation problem.

On performing this optimization, the lowest objective function values from each optimization run were identified, and sorted in an increasing manner and plotted for all three methods. The results of the top 21 runs from the different methods are shown in Figure 4.10B. The cVAE method achieved a better objective function value compared to the state-of-the-art methods. The particle swarm method showed its tendency to get stuck in local minima (observed as two different plateaus). The cVAE, while still operating similar to a global optimization method, managed to avoid getting stuck in similar local minima and instead discovered a lower objective function value. The improved performance of the cVAE pipeline on a standardized model chosen from the BioModels database adds more confidence to the potential of this method as a generally applicable parameter estimation method.



Cortisol model benchmarking results

# Conclusion

We showcased the application of a conditional variational autoencoder in improving the parameter estimation process of a metabolic kinetic model. Kinetic models are accompanied by non-convex parameter spaces that make the task of parameter estimation at a global optimum particularly challenging. The problem of stiff ODEs resulting in poor integration conditions also lends to the computational burden of the task. We showed that a cVAE can be trained on the parameter landscape where the objective function value of the kinetic model serves as a conditional label. On training the cVAE on just a fraction of the landscape, the model was able to generate new parameter vectors that closely matched their expected objective value. We showed that the cVAE did not simply resample the latent space associated with the lowest training data values. Instead, it sampled new regions of the parameter space while still displaying low objective values for the newly generated vectors (Figure 4.9). These vectors can serve as better starting points for local gradient methods of parameter estimation than uniform sampling methods. As the cVAE-generated parameter vectors also show a much lower fraction of stiff vectors that cause failed integrations (something not guaranteed with uniform sampling methods), cVAE-generated vectors can even serve to cut down on computational runtime and resources, thus significantly speeding up the parameter estimation process. The results of a benchmarking study performed on a kinetic model of cortisol demonstrated that the cVAE method was capable of discovering new minima with improved parameter vectors compared to previously established methods in global and hybrid optimization methods. Thus, the cVAE optimization method represents a new promising new method where machine learning techniques can assist in boosting the performance of mechanistic modeling methods.

## Methods

Here are the methods.

## References

1. Villaverde, A. F. *et al.* BioPreDyn-bench: A suite of benchmark problems for dynamic modelling in systems biology. *BMC Systems Biology* **9**, 8 (2015).
2. Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A. & Blom, J. G. Systems biology: Parameter estimation for biochemical models. *The FEBS Journal* **276**, 886–902 (2009).