Deep QSP Modeling: Leveraging Machine Learning for QSP Model Development and Evaluation

Ahmed Elmokadem[1], Timothy Knab[1], Eric Jordie[1], and Yuezhe Li[1]
[1]*Metrum Research Group, Tariffville, CT, USA*

Abstract

Objectives: QSP models characterize the interaction between biological systems and pharmacologic therapies. As such, they require a wealth of mechanistic knowledge of the system of interest, which is oftentimes limited, uncertain, and incomplete. Universal differential equations (UDEs) provide a framework for integrating scientific modeling with data-driven machine learning approaches, such as deep learning. Within such a framework, the scientific model provides enough structural knowledge that alleviates the need for large amounts of data, typically required by deep learning methods. This work introduces the concept of Deep QSP (DQSP) modeling that applies UDEs to QSP models using open-source Julia tools. This framework employs deep learning to fill in the gap in our knowledge of the biological system.

Methods: A simple QSP model of the effect of the antipsychotic drug, remoxipride, on the lactotroph-prolactin system, was used to demonstrate the proposed framework [1, 2]. The model assumed a positive feedback effect of plasma prolactin on prolactin production in lactotrophs. This positive feedback was considered unknown and replaced by an artificial neural network (ANN) in the DQSP framework. Symbolic regression, using sparse identification of nonlinear dynamics (SINDy), was used to retrieve the actual 'unknown' positive feedback effect on the lactotroph-prolactin system dynamics from the trained ANN. An integrated Bayesian DQSP (BaDQSP) framework was also introduced to quantify the uncertainty around the missing dynamics and the retrieved symbolic term.

Results: Observed versus predicted profiles, loss convergence, and error estimates showed that the ANN was efficiently trained to replace the unknown dynamics using a limited training dataset. Symbolic regression was able to retrieve the actual underlying missing term, which was a saturable positive feedback effect of plasma prolactin on lactotroph prolactin production. BaDQSP was able to quantify the uncertainty around the recovered term and assigned it a probability of 0.712 of being the true missing term.

Conclusions: DQSP/BaDQSP framework was introduced in this work using open-source tools in Julia. A QSP model for the effect of remoxipride on the lactotroph-prolactin system was used as a case study to demonstrate the training of an ANN to replace unknown dynamics, retrieving the actual dynamics using symbolic regression, and integrating with Bayesian inference to quantify the uncertainty around the retrieved symbolic term. The convenience of the proposed DQSP/BaDQSP framework makes it accessible to the audience interested in applying deep learning and Bayesian analysis to QSP models.

Methods

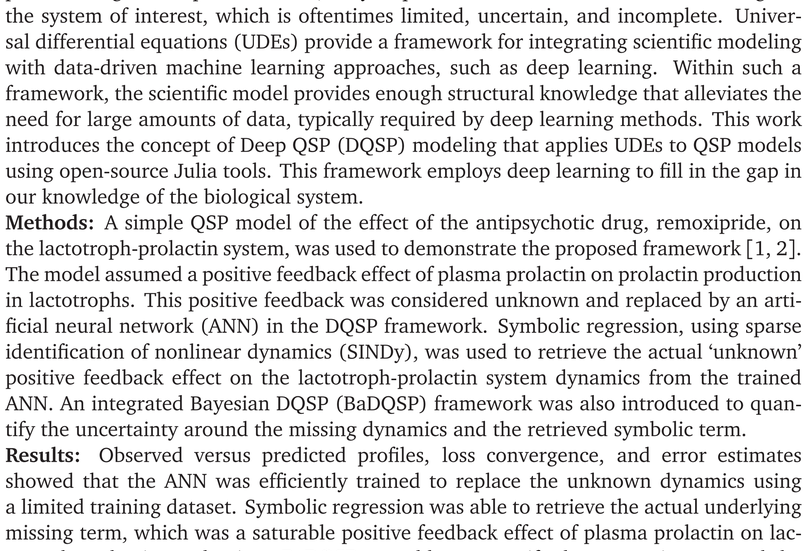

Figure 1. Generic DQSP workflow. The generic DQSP workflow depends on building the QSP UDE model based on the domain-specific knowledge of the reaction network and ANNs to learn the unknown/missing fluxes. Symbolic regression can then follow to recover the unknown flux underlying dynamics. Bayesian analysis is then used to obtain posterior probabilities for the ANN and the recovered term (BaDQSP).

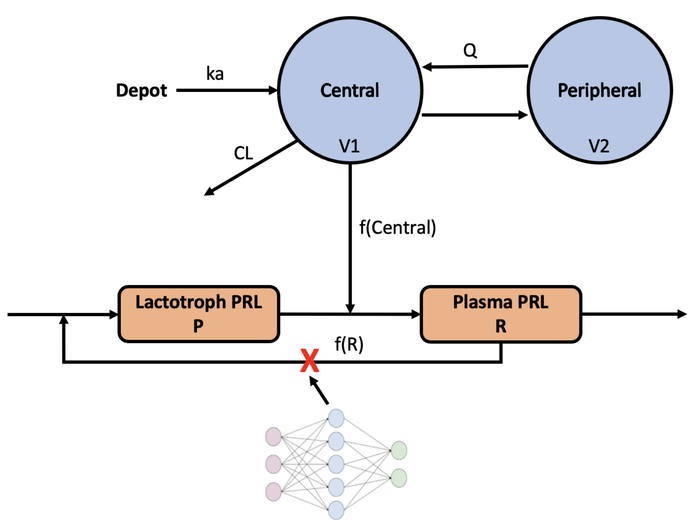

Figure 2. Prolactin UDE Model Structure. The UDE model structure for prolactin (PRL) demonstrates the 2-compartment PK of remoxipride and its effect on stimulating PRL release from lactotrophs to plasma. A nonlinear positive feedback loop, where plasma PRL stimulates the production of lactotroph PRL was included in the original model, and synthetic data were generated from this model. The feedback loop was then assumed missing, and an ANN was used to learn that missing flux in the UDE model structure. The model was solved using the Julia open-source package DifferentialEquations.jl.

ANN structure

- an input layer made of 4 nodes (Central, P, R, and R-R0, where R0 = baseline plasma PRL)
- a hidden layer made of 5 nodes
- an output layer of 1 node representing the feedback flux

The *swish* activation function was used. The ANN was built using the Julia open-source packages Lux.jl and DiffEqFlux.jl.

Optimization

Optimization of the ANN was carried out in two consecutive steps where an initial optimization was carried out using the adaptive moment ADAM optimizer for 5000 iterations to get closer to the global minimum then the final parameters were used to initialize a second round of optimization using the limited-memory Broyden-Fletcher-Goldfarb-Shanno optimizer (LBFGS) for 1000 iterations. Optimization was done using the Julia open-source package Optimization.jl and the loss function computed least squares between the data and the predictions.

Symbolic regression

Recovering the unknown flux was carried out using SINDy. Automatically generated basis functions were generated and included second order polynomials of the input variables as well as a control signal representing the nonlinear dynamics involved in the actual feedback loop. Symbolic regression was done using the Julia open-source package DataDrivenDiffEq.jl.

Bayesian analysis

Bayesian analysis was carried out using the stochastic gradient Langevin dynamics (SGLD) algorithm. 1000 warmup and 1000 sampling iterations were utilized.

Results

DQSP framework successfully implemented a UDE that characterized the PK of remoxipride and its effect on PRL release from lactotrophs to plasma. The model also characterized the positive feedback effect of plasma PRL on lactotroph PRL stimulation using an ANN. The QSP model missing the feedback effect was not able to characterize the plasma PRL data (Figure 3), while the trained UDE model was able to characterize the PRL data (Figure 4). The two-step optimization of the ANN parameters led to a consistent decrease in the loss function value throughout a total of 6000 iterations (Figure 5). Since the data was synthetic, we had the opportunity to overlay the trained ANN time-course output to the actual 'unknown' flux and the ANN was able to successfully learn to *be* that unknown flux (Figure 6). Bayesian analysis as part of BaDQSP was able to characterize the uncertainty around the ANN as depicted by the sample losses and posterior predictive checks (Figures 7 and 8, respectively). The posterior probability around the recovered term is demonstrated (Table 1) where the correct nonlinear flux that was dependent on $R-R0$ had a probability of 0.712 while a second potential flux had a probability of 0.288. Nevertheless, the lower probability term still included the correct nonlinear flux.

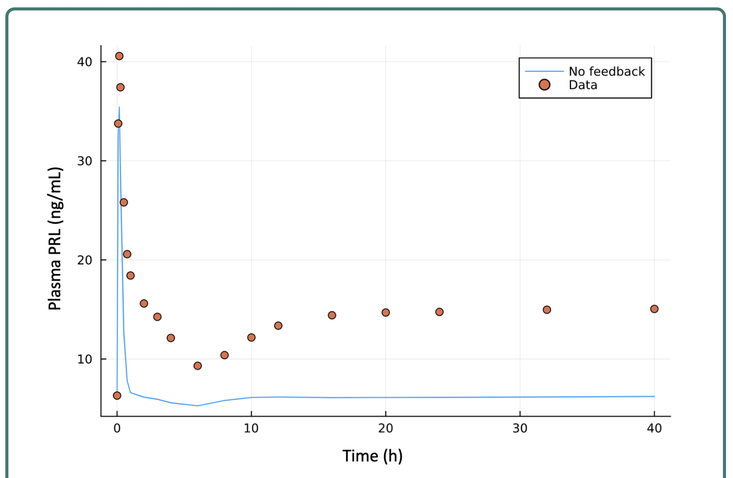

Figure 3. Data versus initial QSP model predictions without the plasma PRL feedback. Data and predictions are represented by orange dots and blue line, respectively.

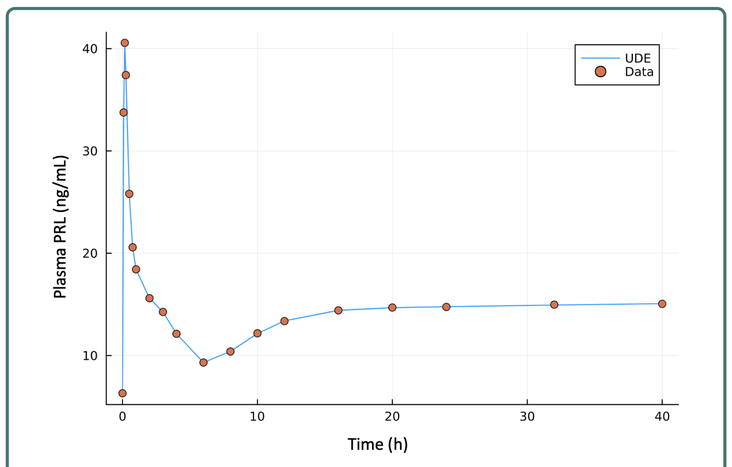

Figure 4. Data versus trained UDE model predictions. Data and predictions are represented by orange dots and blue line, respectively.

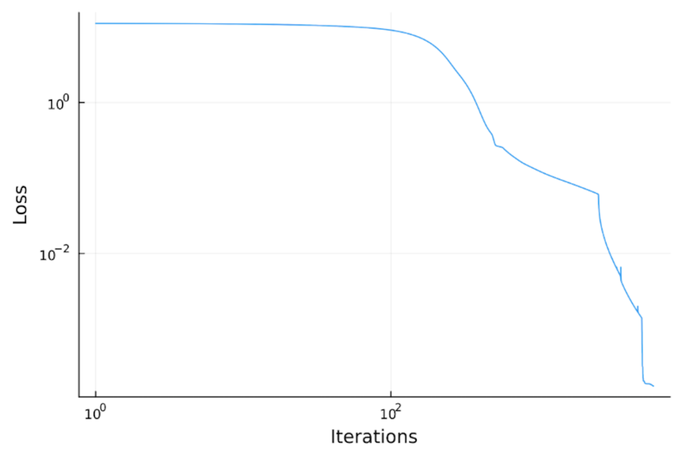

Figure 5. The loss function value over optimization iterations. Loss function values decreased for the 6000 iterations of optimization (5000 initial iterations with ADAM and 1000 subsequent iterations with LBFGS).

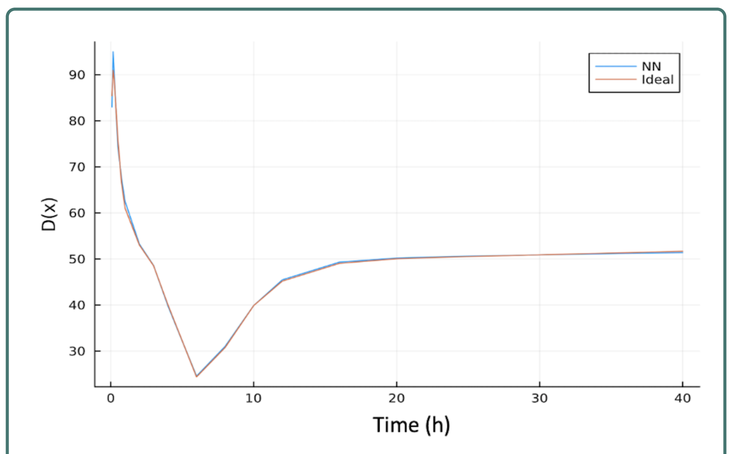

Figure 6. ANN prediction versus the ideal flux. The trained ANN time-course prediction overlaid with the ideal missing positive feedback flux.

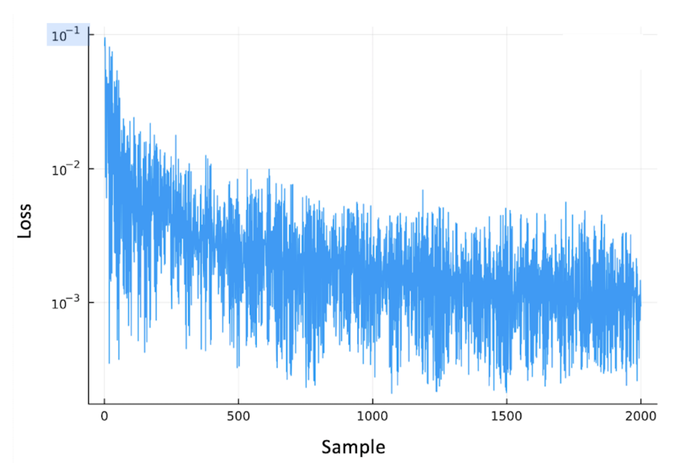

Figure 7. The Bayesian analysis loss function value over iterations. Loss function values for the SGLD Bayesian analysis over 2000 iterations (1000 warmup and 1000 sampling). The optimized parameters from the LBFGS run were used as initial values for the SGLD algorithm.

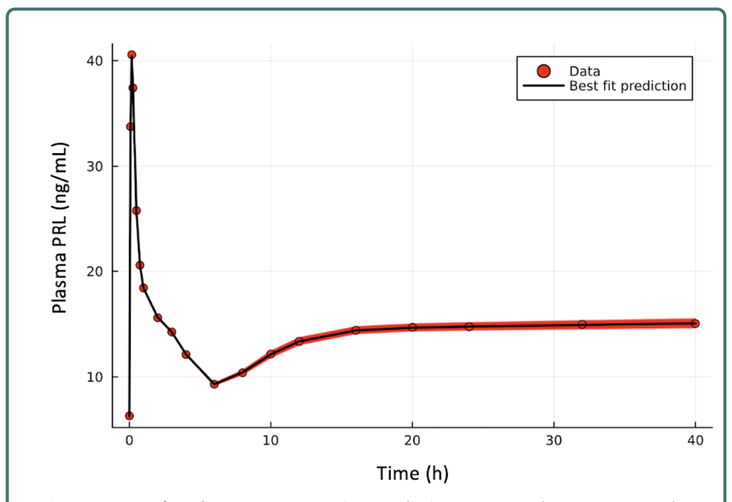

Figure 8. PPC for plasma PRL Bayesian analysis. Posterior plasma PRL samples (red) were overlaid on top of the data (dots). The black line represents the best fit prediction among the posterior samples.

Recovered term	Probability
$\varphi \sim \theta * \frac{Emax * (R - R_0)}{EC50 + (R - R_0)}$	0.712
$\varphi \sim \theta_1 * C^2 + \theta_2 * \frac{Emax * (R - R_0)}{EC50 + (R - R_0)}$	0.288

Table 1. Posterior probability for the possible missing terms. R_0 = baseline plasma PRL concentration; C = plasma concentration of remoxipride; $Emax$ = maximum positive feedback effect; $EC50$ = plasma PRL difference from baseline that achieves 50% of $Emax$.

Conclusion

- A DQSP framework was introduced in this work that integrates the domain-specific knowledge embedded in a QSP model and machine learning to fill in the gaps in that knowledge.
- The proposed framework successfully characterized an unknown positive feedback loop in the lactotroph-prolactin system using a UDE model structure as well as retrieving the actual dynamics from the trained ANN using symbolic regression.
- BaDQSP is an extension to that framework that applies Bayesian analysis to infer a posterior probability of the trained ANN and the underlying unknown term.
- The proposed DQSP/BaDQSP framework was carried out using open-source Julia tools, which makes it accessible to the audience interested in applying deep learning and Bayesian analysis to QSP models.

References

[1] Bakshi, S., de Lange, E.C., van der Graaf, P.H., Danhof, M. and Peletier, L.A. Understanding the Behavior of Systems Pharmacology Models Using Mathematical Analysis of Differential Equations: Prolactin Modeling as a Case Study. *CPT Pharmacometrics Syst Pharmacol* **5** (2016):339–351.

[2] Stevens, J., Ploeger, B.A., Hammarlund-Udenaes, M., Osswald, G., van der Graaf, P.H., Danhof, M. and de Lange, E.C.M. Mechanism-based PK-PD model for the prolactin biological system response following an acute dopamine inhibition challenge: quantitative extrapolation to humans. *J. Pharmacokinet. Pharmacodyn.* **39** (2012):463–477.