PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 7-9, 2022
SGP-TR-223

Physics-Guided Deep Learning for Prediction of Geothermal Reservoir Performance

Zhen Qin (USC), Anyue Jiang (USC), Dave Faulder (Cyrq Energy Inc.), Trenton T. Cladouhos (Cyrq Energy Inc.), and Behnam Jafarpour (USC)

Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California, USA

zhenq@usc.edu

Keywords: machine learning, physics-informed, neural networks, geothermal reservoir

ABSTRACT

Data-driven predictive models have recently been developed and applied to predict energy production from geothermal reservoirs. Traditional physics-based simulation models that are often used for predicting the performance of geothermal reservoirs require detailed descriptions of reservoir conditions and properties, in-depth technical knowledge, and extensive computational efforts. Data-driven models, on the other hand, extract complex statistical patterns from representative training data and use them for prediction. Deep learning-based architectures offer powerful and fast prediction models that are amenable to flexible implementation and convenient deployment and application. Despite their efficiency, data-driven models suffer from three main limitations: 1) extensive data requirements, 2) lack of physical plausibility and interpretability, and 3) poor generalizability beyond training data. These limitations undermine the use of data-driven models for applications in complex science and engineering problems.

We develop a novel physics-informed machine learning approach that integrates physics-based representations into recurrent neural network (RNN) models. A typical approach for enabling the incorporation of physics is to regularize the learning of trainable parameters by adding a physics-constrained loss function as a regularization term. In this framework, the architecture of the deep learning model is designed based on the governing equations as well as expert knowledge of the physical system. We present a variant of the RNN as a physics-guided deep-learning model for predicting the dynamical responses of geothermal reservoirs. Inspired by the lumped-parameter modelling, the architecture of the predictive model is designed based on simplified governing equations (i.e., conservation of energy). Different components of heat-transfer terms are incorporated into the architecture of the physics-constrained RNN to approximately represent the behavior of the underlying dynamics. We present the physics-constrained RNN approach in detail and demonstrate its prediction performance by applying it to data from a field-scale geothermal reservoir model.

1. INTRODUCTION

Prediction of production behavior from geothermal reservoirs using numerical simulation is traditionally used to facilitate field development and management strategies. Traditional physics-based reservoir simulation approaches leverage the physics of energy and mass transport and provide reliable predictions of the responses of geothermal reservoirs to alternative development/operation plans. However, a simulation model requires detailed description of reservoir conditions and properties and therefore is not trivial to develop. On the other hand, data-driven predictive models learn complex statistical patterns from training data and use them for fast prediction. An important limitation of purely data-driven models that work as "black box" is that they are prone to generate physically inconsistent predictions and often fail to extrapolate beyond the training samples.

One of the major challenges for geothermal reservoir engineering is the lack of an efficient and reliable method that can accurately make predictions of production performance based on historical data. The conventional predictive methods based on historical information from geothermal reservoirs include (1) numerical reservoir simulation, (2) lumped-parameter models, and (3) decline curve analysis (Ciriaco et al., 2020). Similar to numerical simulation, lumped-parameter models are built based on governing equations but are much less discretized and more simplified. Compared with numerical models that consist of many (>100-10^6) grid blocks, the lumped parameter model of a geothermal system contains only a small number of homogeneous blocks and mainly focuses on well blocks (Sarak et al., 2005). Similarly, in lumped parameter models, a fully implicit Newton-Raphson procedure is used to solve the nonlinear differential equations for each time step (Tureyen and Akyapı, 2011). The iterative nature of the Newton-Raphson method leads to slow runtime, especially with an increasing number of wells. While decline curve analysis is used for the prediction of geothermal reservoirs with faster runtime, this approach does not consider the dynamic production behavior.

Deep learning-based predictive models have been increasingly applied to the prediction of production in subsurface flow systems in recent years. The recurrent neural networks (RNNs), as variants of neural networks (NN), are commonly used for predicting sequences of dynamical time-series data. RNNs have been applied in the geothermal domain for permanent downhole gauge data prediction (Tian and Horne, 2017), tracer concentration prediction (Gudmundsdottir and Horne, 2020), and production performance prediction (Jiang et al., 2021; Shi et al., 2021). As an efficient and powerful alternative, deep learning-based models are amenable to flexible implementation and convenient deployment and application.

One major drawback of data-driven models is their inability to extrapolate beyond the training data (Karniadakis et al., 2021; Willard et al., 2021). Novel methodologies are introduced to solve these issues by integrating physical principles and expert knowledge with data-

driven models. A common approach for incorporation of physics is to introduce the governing equations and/or expert knowledge into the loss function during training (Gao et al., 2021; Raissi et al., 2019; Ranade et al., 2021; Sirignano and Spiliopoulos, 2018). The physics-based loss functions work as a regularization term and penalize the data-driven model for physically inconsistent predictions. Another typical approach is to design specialized NN architectures that implicitly embed physical knowledge by specifying node connections. The resulting physics-based architectures capture physical dependencies among variables and, therefore, make the data-driven model interpretable. Recent studies have designed architectures of RNNs based on physics or expert knowledge in other applications, including seismic inverse problems (Sun et al., 2020) and modeling lake temperature (Daw et al., 2020). Another approach is to combine physics-based proxy models with purely data-driven models. An example of a hybrid approach is presented in (Long et al., 2018) by including the fluid dynamical system and heat convection-diffusion system.

In this paper, we present a hybrid predictive model that combines Gated Recurrent Unit (GRU) with a physics-based proxy model to predict the dynamical production data from a geothermal reservoir. In our example, we use data from a field-scale geothermal reservoir simulation model to enable validation. Similar to the lumped parameter method, the physics-based proxy model is designed based on simplified physics that describes the geothermal reservoir system. The proxy model, however, relies on several more approximations that allow it to be represented as an RNN unit. Within the hybrid model, the GRU layer and the proxy are designed to predict the bottom-hole pressure (BHP) and brine temperature, respectively. The GRU layer is fed with rate control and predicts the BHP of production and injection wells. The rate control and predicted BHP are then sent to the proxy model as input, where the mapping between inputs and outputs is designed to mimic different components of energy conservation. Compared with lumped parameter model, the physics-guided hybrid approach decouples two primary variables (pressure and temperature) and circumvents the Newton-Raphson iterations. In the remainder of this paper, we present the physics-informed RNN approach in detail and demonstrate its prediction performance by applying it to simulated dataset from a field-scale geothermal reservoir model.

2. METHODOLOGY

2.1 Governing Equation

The most common approach for predicting the performance of a geothermal reservoir is numerical simulation. The development of geothermal reservoir numerical simulation is based on governing equations, material properties, boundary conditions, initial conditions, and discretization. In this section, only a brief introduction to governing equations of the reservoir will be provided. A more detailed description of geothermal reservoir modeling can be found in (Jung et al., 2018) and (Vinsome and Shook, 1993). The mathematical representation of the lumped parameter method resembles that of the numerical simulation method, with the addition of several simplifications. A detailed discussion of lumped parameter method can be found in (Tureyen and Akyapı, 2011).

In this work, the studied geothermal reservoir model is a single-component and single-phase system. A general form of mass- and energy-balance equations for non-isothermal flow in the geothermal reservoir can be written as

$$\frac{\partial}{\partial t}M^{\kappa} = -\nabla\ \mathbf{F}^{\kappa} + q^{k} \tag{1}$$

where M^{κ} represents the mass or energy component κ per volume within an arbitrary subdomain. The energy term is treated as a pseudo component. The left-hand side represents the mass or energy accumulation term. The notation $\mathbf{F}^{\kappa}$ denotes mass or heat flux term, and q^{k} is sink/source term, which is the mass flow rate for the mass-balance equation, and the energy flow rate for the energy-balance equation. The notation ∇ is the divergence of a vector field.

The accumulation terms for mass- and energy-balance for component κ are written as

$$M^{\kappa} = \phi\rho_f \tag{2}$$

$$M^{\kappa} = (1-\phi)\rho_r c_r T + \phi\rho_f c_f T \tag{3}$$

where ϕ is the porosity, ρ_f is the fluid density, ρ_r is the rock grain density, c_r is the specific heat of rock grains, T is temperature, and c_f is the specific heat of fluid.

For mass-balance equation, advective mass flux term can be defined by Darcy's law:

$$\mathbf{F} = \rho_f \mathbf{u} = -k\frac{\rho_f}{\mu}\left(\nabla P - \rho_f \mathbf{g}\right) \tag{4}$$

where $\mathbf{u}$ is the Darcy velocity, k is the absolute permeability, μ is the dynamic viscosity, P is pressure, and $\mathbf{g}$ is the vector of gravitational acceleration.

In the energy balance equation, the heat flux term includes conductive, convective, and radiative components. By ignoring the radiative component, the heat flux is written as

$$\mathbf{F}^{\kappa} = -\lambda\nabla T + h_f \mathbf{F} \tag{5}$$

where λ is thermal conductivity, and h_f is the specific enthalpy for fluid.

2.2 Architecture of the Predictive Model

The schematic of the hybrid predictive model is shown in Figure 1 and consists of two stacked layers. The first layer of the model contains two main components: a convolution neural network (CNN) and an RNN. The CNN-RNN layer is implemented with the open-source deep learning framework TensorFlow (Ababi et al., 2016). The second layer is a physics-based proxy model designed to approximate the energy-balance equation. The detailed structures of the two layers are depicted in Figures 2~4.

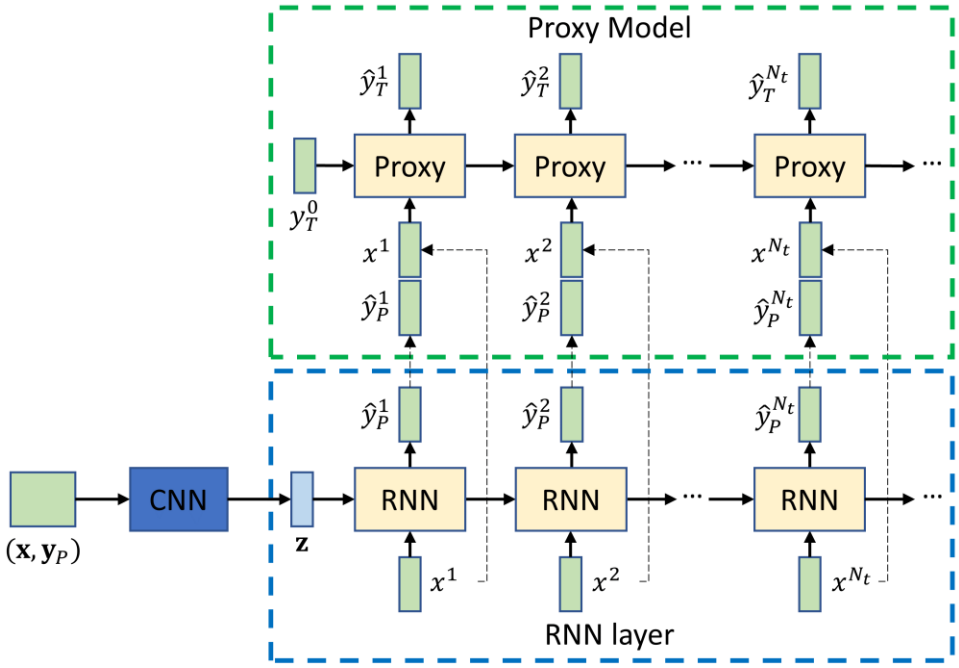


Figure 1: Schematic of the physics-guided predictive model.

The type of CNN used in the first layer is the one-dimension (1D) convolution operation (denoted as *conv1D* and shown in Figure 2), which can be applied to the time-series data (Goodfellow et al., 2016). In the proposed model, the 1D-convolution network is fed with a short sequence of historical data ($\mathbf{x}, \mathbf{y}_P$) and maps it to a low-dimensional latent space, $\mathbf{z} \in R^{1\times N_z}$. The latent variable $\mathbf{z}$ from the *conv1d* layer is then sent to the RNN layer as the initial state. The historical data includes the control variable, $\mathbf{x} \in R^{N_{t0}\times N_x}$, and BHP response data, $\mathbf{y}_P \in R^{N_{t0}\times N_P}$, where subscript P denotes pressure. The superscripts N_{t0}, N_x, and N_P denote the historical data timesteps, input feature dimension, and BHP data feature dimension, respectively. More specifically, $\mathbf{x}$ and $\mathbf{y}_P$ are the rate control and BHP of all active wells (production and injection), respectively. Therefore, the values of N_x and N_P equal to the number of wells from the geothermal reservoir under study.

Two common variants of RNN models are the long short-term memory (LSTM) and the gated recurrent unit (GRU). Both LSTM (Hochreiter and Schmidhuber, 1997) and GRU (Cho et al., 2014) leverage gate mechanisms to regulate the flow of information and enhance the dependency on previous states over a long period. The GRU unit (Figure 3) is used in the proposed model since it has a similar performance to the LSTM but has fewer weights to train. The RNN layer is fed with future control, $\mathbf{x} \in R^{N_{t1}\times N_x}$, as input and latent variable $\mathbf{z}$ as the initial state. The output from the RNN layer is the predicted BHP response, $\hat{\mathbf{y}}_P \in R^{N_{t1}\times N_P}$. The superscript N_{t1} is the future control timesteps.

The inputs to the physics-based proxy model are the future control $\mathbf{x}$ and predicted BHP $\hat{\mathbf{y}}_P$. The output is the predicted production temperature $\hat{\mathbf{y}}_T \in R^{N_{t1}\times N_T}$, where T and N_T denote production temperature and the dimension of the temperature input feature. To be consistent with physical properties inside the proxy, variables $\mathbf{x}$, $\hat{\mathbf{y}}_P$, and $\hat{\mathbf{y}}_T$ are re-written as production mass rate m_p^k, predicted BHP $\hat{P}^k$, and predicted production temperature $\hat{T}_p^k$ for the k^{th} time step, respectively. Only the production rate is used here since the proxy model only predicts the production temperature. Therefore, the injection rate in $\mathbf{x}$ is disregarded by the proxy model. The proxy model has three components to approximate the heat conduction term, the sink term, and the heat convection term (Figure 4). In the heat conduction term, the temperature of rock grains T_R is assumed to be constant. In the sink term, the extracted enthalpy is the product of mass flow rate and specific enthalpy. The specific enthalpy is simplified to be proportional to brine temperature. In the convection term, the energy flux between a production well i and an injection well j is simplified based on the relation from Tureyen and Akyapı (2011), which is written as

$$q_{ij} = a_{ij} h_j(t) \left(P_i(t) - P_j(t) \right) \tag{6}$$

where q_{ij} is the energy flow between a production well i and an injection well j, a_{ij} is recharge index constant or conductance (Li et al., 2017), derived from Darcy's law; h_j is the specific enthalpy for injection well j, which is assumed constant. Another assumption is that

the processing of the temperature variable is in an explicit form since temperature change over each time step is typically small. Based on these assumptions, the calculation of temperature decline consists of a series of algebraic operations and is amenable to implementation as a neural network.

The CNN-RNN layer and the physics-based proxy model are connected and jointly trained using the following loss function:

$$\mathcal{L}(\theta) = \sum^{N} \|\mathbf{y}_P - \hat{\mathbf{y}}_P\|_2^2 + \sum^{N} \|\mathbf{y}_T - \hat{\mathbf{y}}_T\|_2^2 \tag{7}$$

where θ is the weight of the predictive model. To train the resulting architectures, the Adam optimizer is used (Kingma and Ba, 2017).

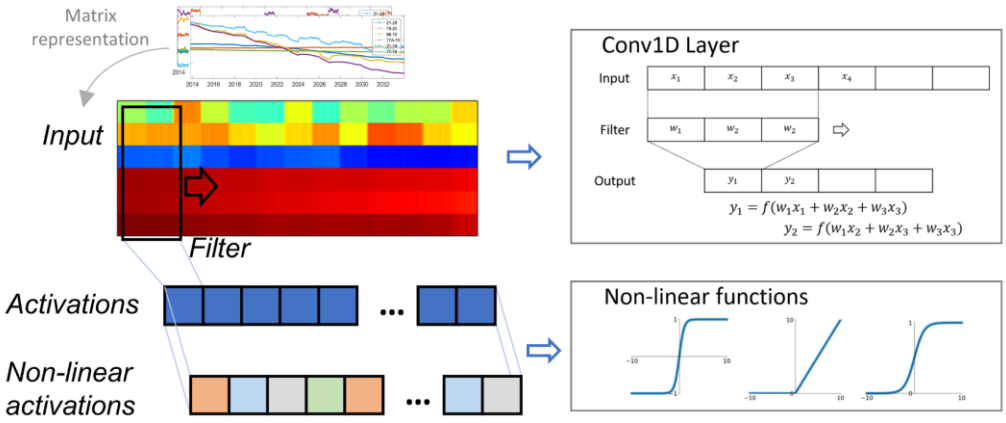


Figure 2: Schematic of 1D convolution operation.

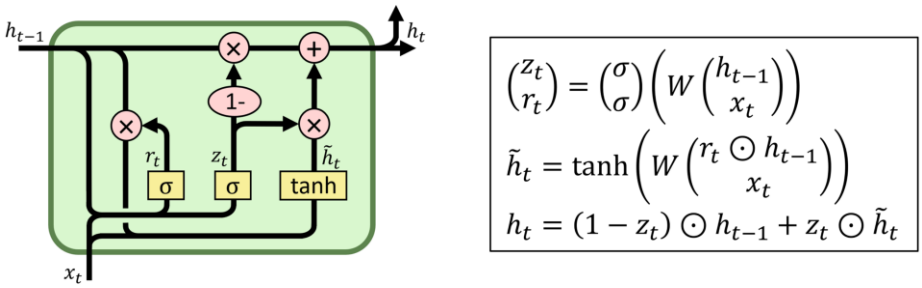


Figure 3. The structure of GRU cell and its mathematical representation.

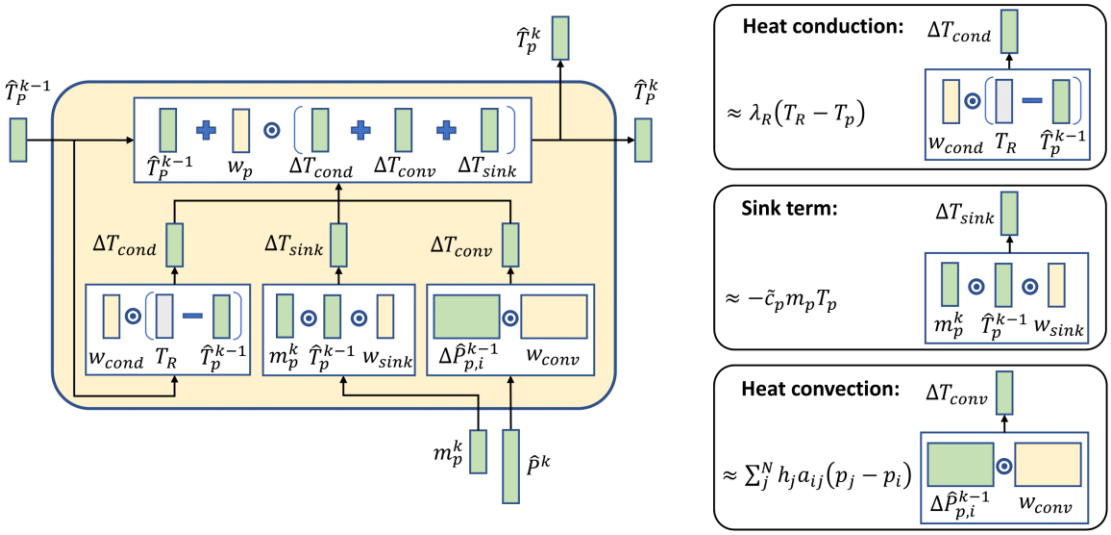


Figure 4. The structure of the physics-based proxy model.

3. NUMERICAL EXPERIMENT

In this experiment, the proposed predictive model is trained and tested using simulated data. The purely data-driven model suffers from three main issues: 1) extensive data requirements, 2) lack of physical plausibility and interpretability, and 3) poor generalizability beyond training data (Willard et al., 2021). Since the interpretability is handled by the proposed model itself, we only focus on the other two issues in this work. Therefore, we classify the data based on data quantity and the range of training samples to better define the problem statement (Figure 5). The lengths of training samples N for training the predictive model are 400, 500, and 600, where the length is the number of data timesteps used for training. Another direction in the quad chart is the consistency of input range between the training and test data. The training samples are then divided into six cases: consistent range with N = 600, 500, 400 (denoted as A1, A2, and A3) and inconsistent range with $N = 600, 500, 400$ (B1, B2, and B3). Four training samples are randomly generated for the examples with consistent range. Six training samples with different control patterns are generated for the quadrants with inconsistent ranges, of which the total production rates for different patterns are shown in Figure 6. The simulated samples are generated over 20 years (from 2013/11/01 to 2034/01/06) with a total length of 1054 timesteps and a time interval of seven days.

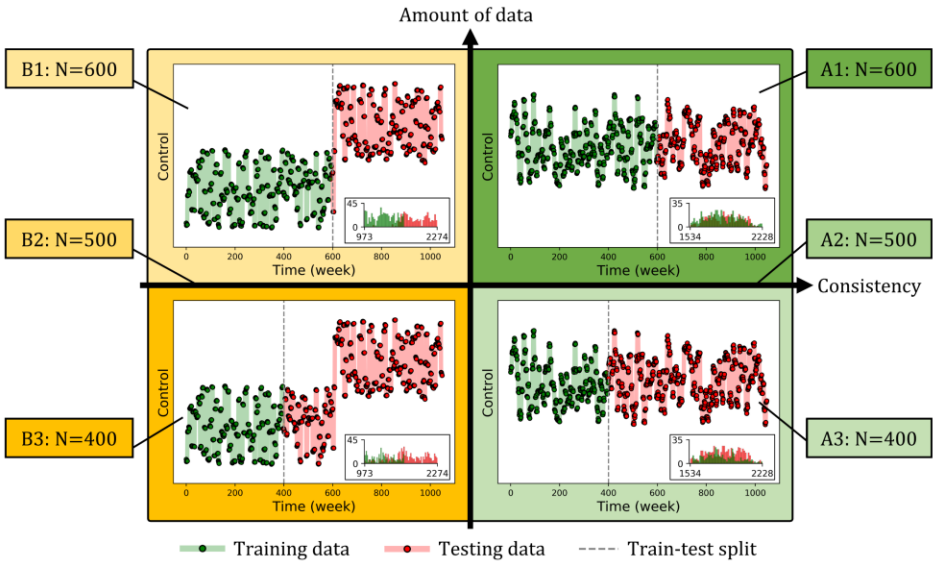


Figure 5. Classifications of input features based on consistency of sample range (A for consistent range and B for inconsistent range) and data quantity (N=400, 500, and 600).

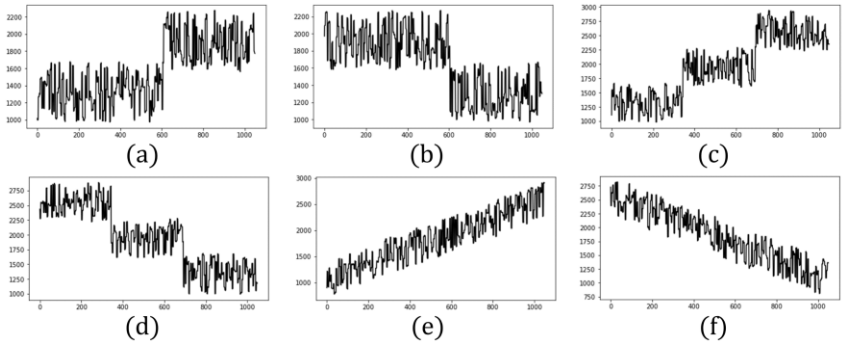


Figure 6. Total production rate for six control patterns with inconsistent range between training and testing set.

Two purely data-driven models are applied in the experiments for comparison with the proposed model: a standard GRU model and a variant of the GRU model. The GRU variant is designed to overcome the extrapolation issue by removing the multiplication weight $(1 - z_t)$ from the last step within the unit. This change is to convert the last step of calculation within the standard GRU from a weighted average between h_{t-1} and $\tilde{h}_t$ to a cumulative sum of h_{t-1} and $\tilde{h}_t$ (See Appendix A). The GRU variant is referred to as "Cumsum GRU", which represents the cumulative sum. Three predictive models are then trained with simulated training samples with different data amounts and ranges. Each predictive model is trained five times to get the average performance for each sample.

Figure 7 shows the prediction results for four samples with a consistent range from three models trained with the first 600 timesteps. The values on the y-axis are the normalized production temperature of six production wells. Wells 88-19, 77A-19 and 77-19 have obvious temperature decline, while Wells 21-28, 78-20 and 21-19 remain stable over 20 years. It is observed that predictions from the proposed model capture the future trends for four samples. The Cumsum GRU captures some of the trends, but shows instability in its performance, while the prediction for Wells 88-19, 77A-19 and 77-19 from the standard GRU flattens. Figure 8 shows the prediction results for six

samples with an inconsistent control range. The trends in the temperature features change since controls go out of range. The proposed predictive model captures the trends for all six control patterns, even when the control values go out of range. The standard GRU model fails to make accurate predictions for most examples. The Cumsum GRU works well for the control patterns "a", "e", and "f", but fails to provide correct prediction for "b" and "d".

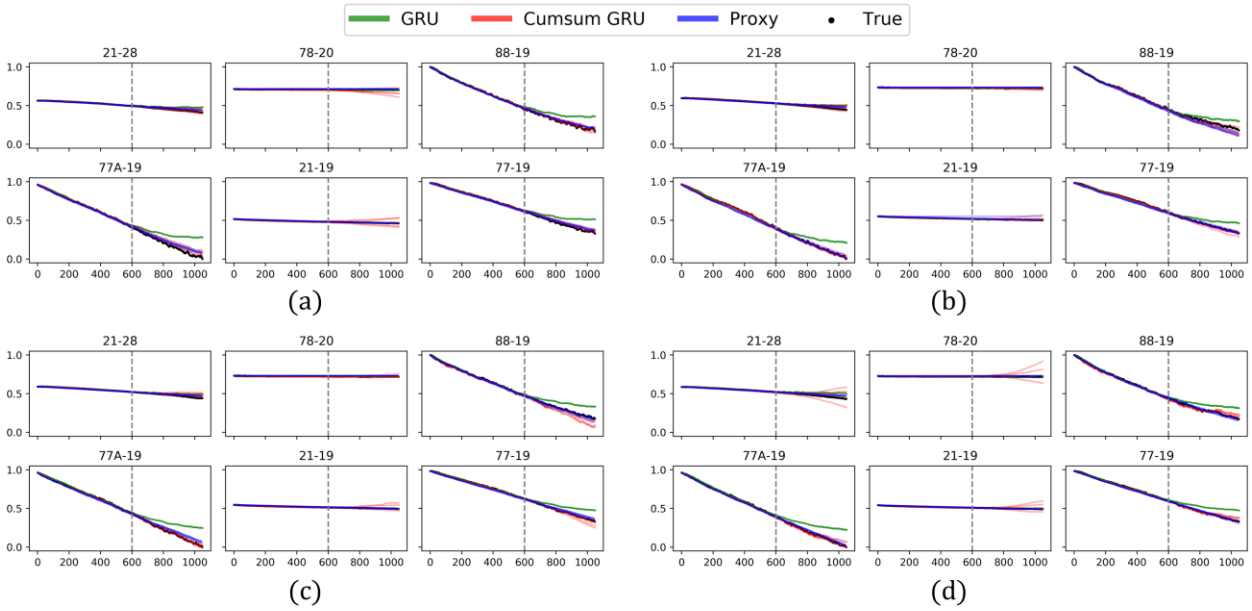


Figure 7. Results for example A1: samples with 600 steps and consistent range. "Proxy" refers to the proposed predictive model.

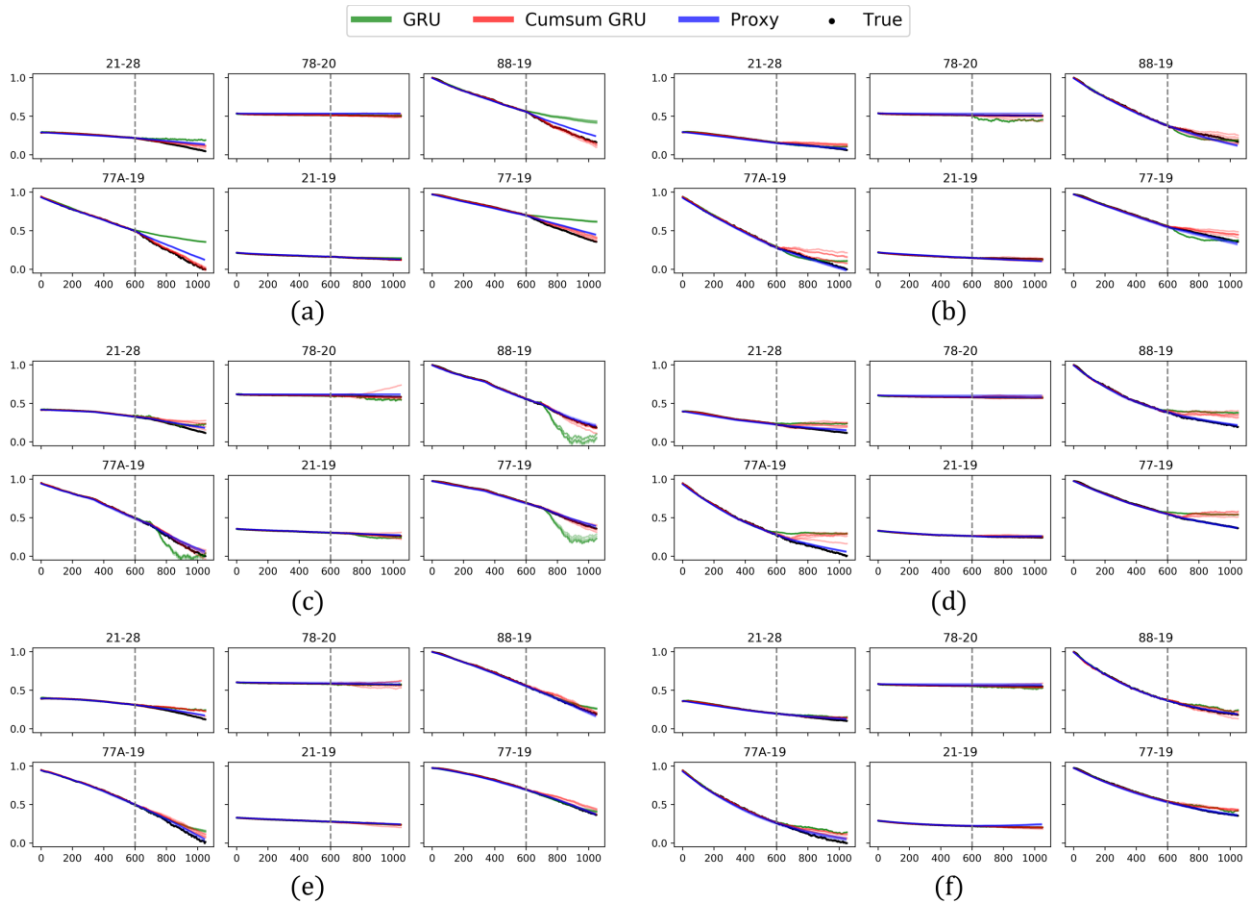


Figure 8. Results for example B1: samples with 600 steps and inconsistent range.

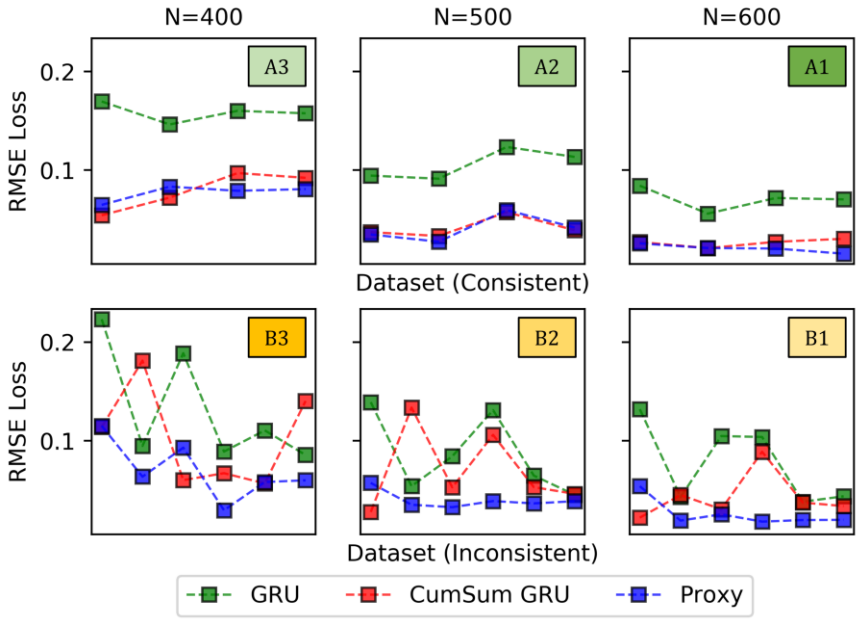


Figure 9. Summary of average RMSE losses for four consistent samples and six inconsistent samples.

Figure 9 summarizes the average root mean square error (RMSE) losses for different examples with the three models. The proposed physics-embedded model outperforms the purely data-driven models in almost all examples. The Cumsum GRU has a similar performance to that of the physics-guided model on the consistent samples but fails to provide correct predictions when the controls during training and test cases are inconsistent. On the other hand, training the models with only 400 timesteps (samples) of data degrades the quality of the physics-guided predictive model, which shows overall higher RMSE losses for $N = 400$. The detailed results for $N = 400$ and 500 are shown in Appendix B.

4. DISCUSSION AND CONCLUSION

Effective development of a geothermal reservoir requires a reliable and efficient predictive model. A fast data-driven model facilitates the implementation of control optimization and other complex iterative workflows (e.g., inverse modelling and uncertainty quantification) that entail high runtime for physics-based simulation models. The main challenges of data-driven models are limited ability to extrapolate beyond training data and extensive data requirement. The prediction of production temperature versus time leads to an extrapolation problem since the temperature keeps declining with a continuous injection of cold water. Moreover, the control optimization implementation with data-driven models typically leads to control strategies that are likely to fall beyond the training samples (Yu and Jafarpour, 2021). On the other hand, data-driven models that are trained with insufficient samples are not likely to provide accurate predictions of future responses. These challenges motivate the need to develop novel architectures that integrate domain knowledge into data-driven models.

In this paper, the proposed deep learning model is augmented by simplified physics to predict future responses for different control patterns. The model consists of a purely data-driven component (i.e., the CNN-RNN layer) and a physics-based proxy. The data-driven layer captures the relationship between the control and the BHP response, while the physics-based proxy reliably approximates the heat flux and the sink term for predicting the temperature decline. The model is trained using simulated data from a field-scale model and is used to predict the BHP and temperature simultaneously. The proposed hybrid model for predicting simulated data outperforms the purely data-driven models. When the control inputs during prediction are out of range (i.e., inconsistent with those from training), the hybrid model shows promising extrapolation capability.

Compared with physics-based models, the proposed model decouples the interaction between temperature and pressure in the governing equations and circumvents the time-consuming Newton-Raphson iterations needed in numerical computation. Compared with data-driven models, the hybrid predictive model learns the mapping between inputs and outputs more efficiently and is constrained by the approximate physics that governs the system. Even though several simplifications are made, the hybrid predictive model is able to present the general correlations and trends for unseen ranges of inputs. Future studies on hybrid physics-guided predictive models include fine-tuning of the model and implementations of integrating physics in different ways.

ACKNOWLEDGEMENT

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Geothermal Technologies Office award number DE-EE0008765. The authors acknowledge the Energi Simulation support of the Center for Advanced Reservoir Characterization and Forecasting at USC. The simulated data in this work were generated using the TETRAD software.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: A System for Large-Scale Machine Learning. Presented at the 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), pp. 265–283.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv:1406.1078.

Ciriaco, A.E., Zarrouk, S.J., Zakeri, G., 2020. Geothermal resource and reserve assessment methodology: Overview, analysis and future directions. Renewable and Sustainable Energy Reviews 119, 109515.

Daw, A., Thomas, R.Q., Carey, C.C., Read, J.S., Appling, A.P., Karpatne, A., 2020. Physics-Guided Architecture (PGA) of Neural Networks for Quantifying Uncertainty in Lake Temperature Modeling, in: Proceedings of the 2020 SIAM International Conference on Data Mining (SDM), Proceedings. Society for Industrial and Applied Mathematics, pp. 532–540.

Gao, H., Sun, L., Wang, J.-X., 2021. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain. Journal of Computational Physics 428, 110079.

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

Gudmundsdottir, H., Horne, R.N., 2020. Prediction Modeling for Geothermal Reservoirs Using Deep Learning. PROCEEDINGS, 45th Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California, February 10-12, 2020, SGP-TR-216.

Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation 9, 1735–1780.

Jiang, A., Qin, Z., Cladouhos, T.T., Faulder, D., Jafarpour, B., 2021. Recurrent Neural Networks for Prediction of Geothermal Reservoir Performance. PROCEEDINGS, 46th Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California, February 15-17, 2021, SGP-TR-218.

Jung, Y., Pau, G.S.H., Finsterle, S., Doughty, C., 2018. TOUGH3 user's guide.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat Rev Phys 3, 422–440.

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Li, Y., Júlíusson, E., Pálsson, H., Stefánsson, H., Valfells, Á., 2017. Machine learning for creation of generalized lumped parameter tank models of low temperature geothermal reservoir systems. Geothermics 70, 62–84.

Long, Y., She, X., Mukhopadhyay, S., 2018. HybridNet: Integrating Model-based and Data-driven Learning to Predict Evolution of Dynamical Systems, in: Billard, A., Dragan, A., Peters, J., Morimoto, J. (Eds.), Proceedings of The 2nd Conference on Robot Learning, Proceedings of Machine Learning Research. PMLR, pp. 551–560.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707.

Ranade, R., Hill, C., Pathak, J., 2021. DiscretizationNet: A machine-learning based solver for Navier–Stokes equations using finite volume discretization. Computer Methods in Applied Mechanics and Engineering 378, 113722.

Sarak, H., Onur, M., Satman, A., 2005. Lumped-parameter models for low-temperature geothermal fields and their application. Geothermics 34, 728–755.

Shi, Y., Song, X., Song, G., 2021. Productivity prediction of a multilateral-well geothermal system based on a long short-term memory and multi-layer perceptron combinational neural network. Applied Energy 282, 116046.

Sirignano, J., Spiliopoulos, K., 2018. DGM: A deep learning algorithm for solving partial differential equations. Journal of Computational Physics 375, 1339–1364.

Sun, J., Niu, Z., Innanen, K.A., Li, J., Trad, D.O., 2020. A theory-guided deep-learning formulation and optimization of seismic waveform inversion. GEOPHYSICS 85, R87–R99.

Tian, C., Horne, R.N., 2017. Recurrent Neural Networks for Permanent Downhole Gauge Data Analysis, in: Day 1 Mon, October 09, 2017. Presented at the SPE Annual Technical Conference and Exhibition, SPE, San Antonio, Texas, USA, p. D011S008R007.

Tureyen, O.I., Akyapı, E., 2011. A generalized non-isothermal tank model for liquid dominated geothermal reservoirs. Geothermics 40, 50–57.

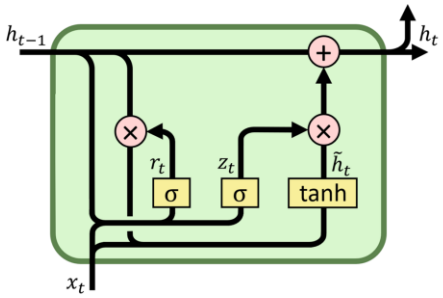
Vinsome, P.K.W., Shook, G.M., 1993. Multi-purpose simulation. Journal of Petroleum Science and Engineering 9, 29–38.

Willard, J., Jia, X., Xu, S., Steinbach, M., Kumar, V., 2021. Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. arXiv:2003.04919.

Yu, J., Jafarpour, B., 2021. Active Learning for Well Control Optimization with Surrogate Models. Presented at the AGU Fall Meeting 2021, AGU.

APPENDIX

Appendix A. GRU Variant



$$\begin{pmatrix} z_t \\ r_t \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \end{pmatrix} \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right)$$
$$\tilde{h}_t = \tanh\left(W \begin{pmatrix} r_t \odot h_{t-1} \\ x_t \end{pmatrix} \right)$$
$$h_t = h_{t-1} + z_t \odot \tilde{h}_t$$

Figure A-1. The structure of the variant of the GRU cell and its mathematical representation.

Appendix B. Results for N = 400 and N = 500

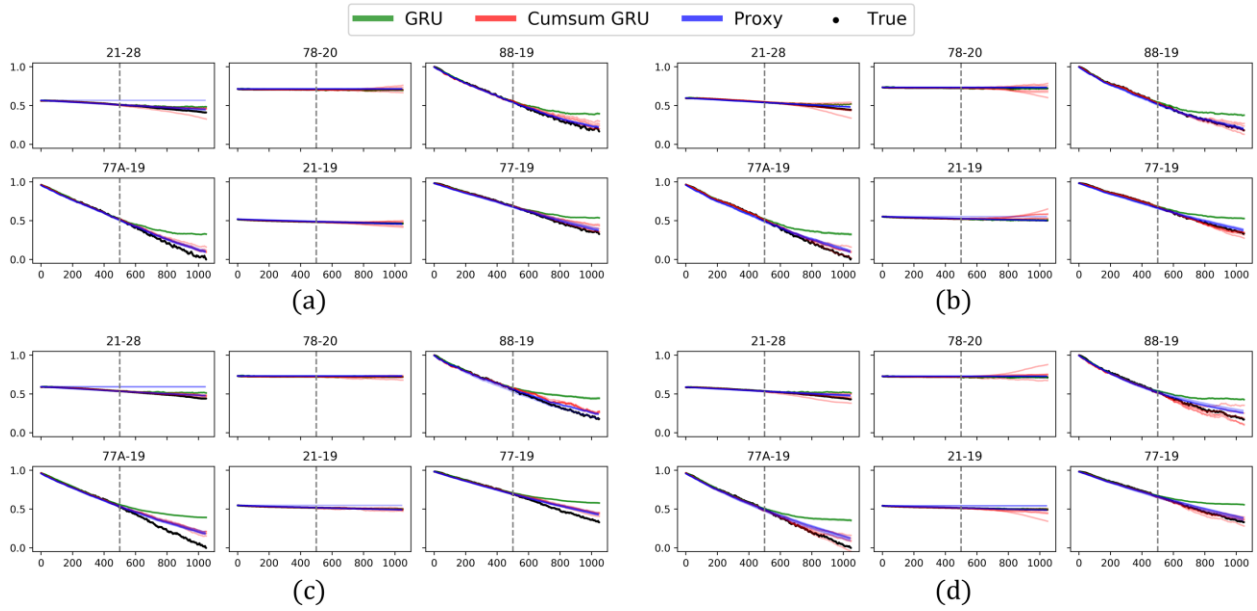


Figure B-1. Results for example A2: samples with 500 steps and consistent range.

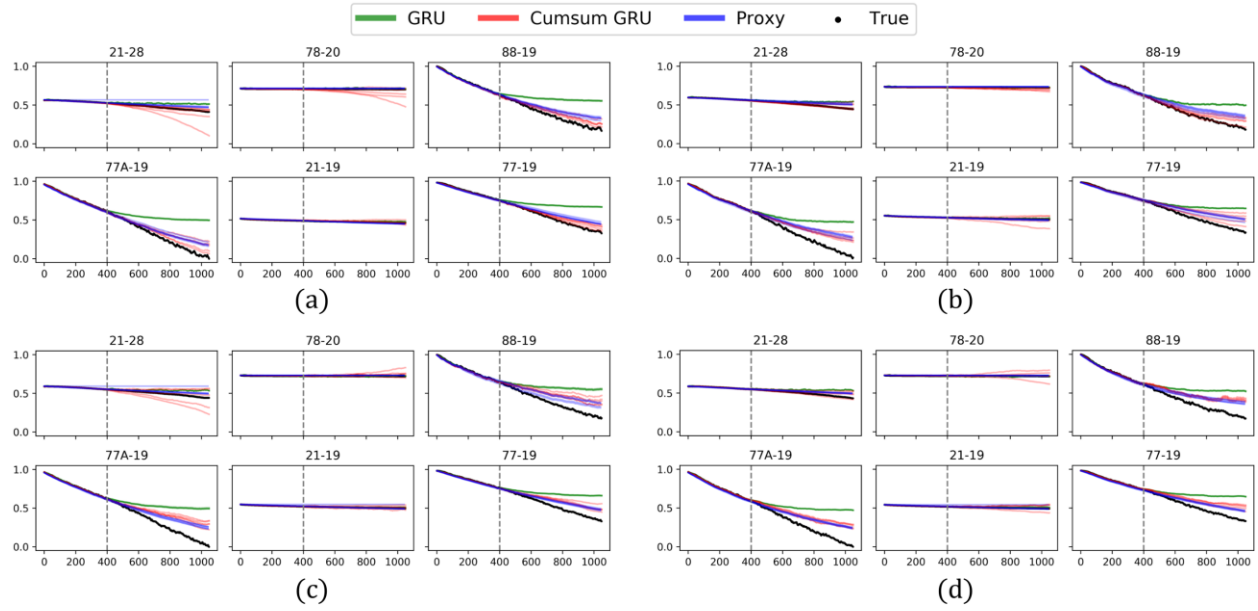


Figure B-2. Results for example A3: samples with 400 steps and consistent range.

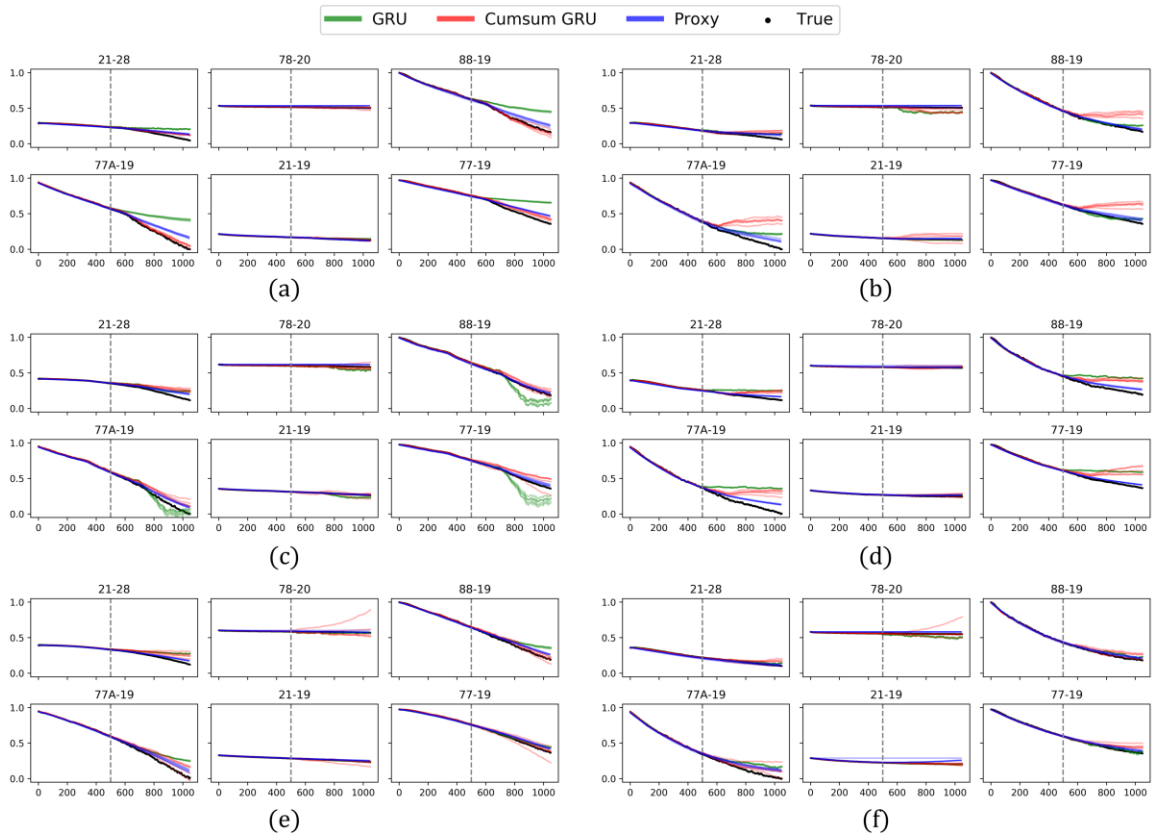


Figure B-3. Results for example B2: samples with 500 steps and inconsistent range.

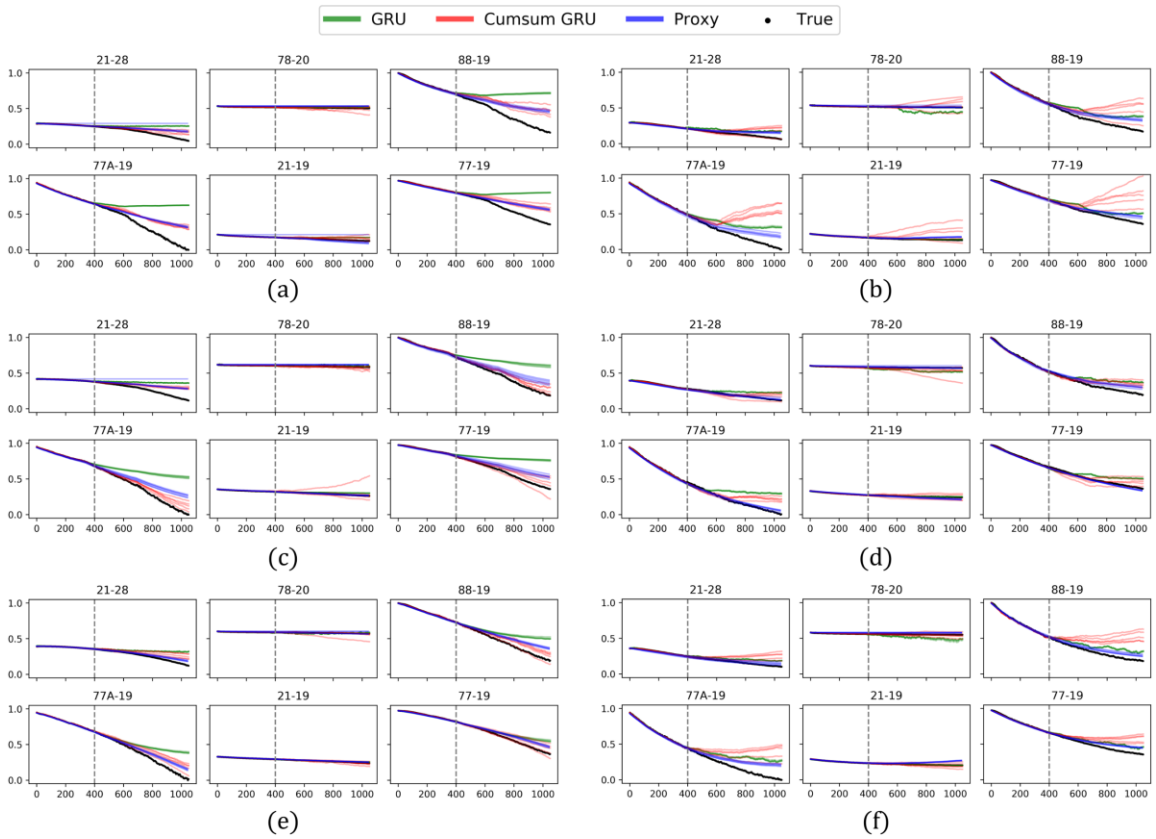


Figure B-4. Results for example B3: samples with 400 steps and inconsistent range.