Computational Prediction of Intracellular Signaling Behavior via Machine Learning

Pamela Romero, Juliette Gourdon, Romain Yvinec, Frederic Jean-Alphonse, Misbah Razzaq

Computational Prediction of Intracellular Signaling Behavior via Machine Learning

Pamela Romero[1], Juliette Gourdon[1], Romain Yvinec[1,2], Frederic Jean-Alphonse[1,2], and Misbah Razzaq[1,*]

[1] INRAE, CNRS, Université de Tours, PRC, 37380, Nouzilly, France.
{pamela.romero-jofre, juliette.gourdon, romain.yvinec, frederic.jean-alphonse, misbah.razzaq}@inrae.fr
[2] Inria, Inria Saclay-lle-de-France, 91120 Palaiseau, France.
*corresponding author

Abstract. Assessing the dynamics of intracellular signaling processes under various conditions such as protein-protein interactions, protein conformational changes, dose-dependent drug effects, and ligand binding is crucial for biologists and pharmacologists. However, generating such data can be time-consuming and costly. In this study, we used data obtained by Bioluminescence Resonance Energy Transfer (BRET) in live cells to develop models to predict the temporal dynamics of intracellular second messenger cAMP (cyclic adenosine monophosphate). We formulated the task of predicting time series in a supervised manner and employed decision tree, random forest, and XGBoost models within a multiple-input, multiple-output framework, enabling simultaneous forecasting of multiple future time steps. Our findings demonstrate that our model achieves, in the main experiment, a mean absolute error of less than 0.018 on the testing set. This model is computationally efficient, requiring only ≈ 16 seconds for training, validation, and testing. Finally, we identified that BRET signal intensities are the most important feature for predictions among the set of features in the models. These results highlight the potential of machine learning models for advancing research in dynamical biological processes, particularly in cellular signaling and pharmacological systems.

Keywords: Machine learning · Random forest · Decision tree · Bioluminescence Resonance Energy Transfer (BRET) · Intracellular signalling · Time series.

1 Introduction

Cell surface G-protein-coupled receptors (GPCRs) are a critical class of pharmacological targets, representing about 30% of FDA-approved drugs. The functional selectivity (biased signaling) of GPCRs is an important matter regarding pharmacological discovery, and is currently an active area of research with many statistical and mathematical developments [17,12,7,13,24]. In this study,

we aimed to predict the cAMP (cyclic adenosine monophosphate) signaling dynamics downstream of the luteinizing hormone receptor (LHR) and the follicle-stimulating hormone receptor (FSHR), two GPCRs essential for regulating reproductive physiology in mammals. cAMP is an intracellular second messenger, whose concentration is controlled by Gs-coupled receptors and the transmembrane adenylate cyclases (ADCY). cAMP dynamics are shaped by two main pathways: (i) the receptor activation by its hormone and the subsequently Gs/ADCY mediated cAMP concentration elevation, and (ii) the cAMP degradation by intracellular phosphodiesterases (PDEs). Bioluminescence Resonance Energy Transfer (BRET) is widely used to measure dynamic cellular events such as signaling, protein-protein interactions, ligand binding, and protein conformational changes in live cells. This technique thus allows gathering important information and knowledge regarding normal or pathological biological processes. To assess cAMP dynamics in live cells, we then used a cAMP BRET sensor transfected into cells expressing the Gs-coupled GPCRs of interest [17].

Time series prediction is a well-established area of research across various applications [23]. Time series prediction can be categorized by forecast horizon into single-step and multi-step forecasting. In single-step forecasting, a one-to-one prediction is made where the present input generates a single future output. In contrast, multi-step forecasting is feedback-based, incorporating prediction information from previous time steps. While this method requires less information, it may accumulate errors since it uses approximated inputs instead of actual values [1]. We adopt a multiple-input, multiple-output strategy that forecasts several future time steps simultaneously [21,22]. In this approach, the model is trained on an initial segment of time series data to predict subsequent segments. For simplicity, we refer to this as the "multi-step model". We focus on decision tree, random forest, and XGBoost models due to their flexibility in capturing non-linear relationships and complex interactions between variables like ligand, dose, receptor, and perturbation. These models are well-suited for predictions in dynamic biological systems, where data can be noisy and complex. The feature importance is calculated to understand how different biological factors influence the prediction of the model, such as receptor activation, Gs-coupled adenylate cyclase activity, or phosphodiesterase-mediated degradation of cAMP, which can ultimately help to understand the signaling components that affect the intracellular second messenger dynamics measured by the cAMP BRET sensor.

2 Data and Methods

2.1 Biological protocol for dataset generation

HEK293A or mLTC-1 cells per well were seeded in 96-well plates previously coated with 0.01% polylysine. Cells were transiently transfected in suspension using Metafectene Pro transfection reagent (Biontex Laboratories, München, Germany) according to the manufacturer's instructions, using DNA quantities specified. Nb37 was transfected under the relevant experimental conditions. Forty-eight hours post-transfection, BRET measurements were conducted upon

addition of 5μM coelenterazine-H (Interchim) diluted in Ca^{2+}/Mg^{2+}-free PBS, containing various concentrations of ligand, and with the appropriate concentration of chemical drug for concerned conditions. For experiments performed in presence of Dyngo4a, PitStop2, ES9-17, or YM254890, cells were pre-incubated 20-35 minutes in presence of drug-containing buffers before measurements. For experiments performed in the presence of PTX, this was added to cells 16 hours before the experiment. Signals were recorded for at least 60 minutes with a Mithras LB 943 plate reader (Berthold Technologies GmbH & Co., Wildbad, Germany), and BRET ratio (480nm/540nm) was calculated.

2.2 Dataset processing

Our dataset includes temporal changes in ligand-induced cAMP BRET signals under different conditions. A time series signal consists of continuous values, and a time series with k data points can be represented as:

$$T_c = (t_c^1, t_c^2, \ldots, t_c^k) \tag{1}$$

where T_c represents the time series for a given condition c, and c denotes a specific condition determined by the dose, ligand, receptor, and perturbation. The dose refers to the concentration of the ligand in nanomolar units [nM]. The dataset includes ten dose values. Ligands are molecules that bind to receptors. Our dataset includes three types of ligands: hCG, LH, and FSH [7]. We also have five types of receptors: hLHR, hLHR-T (a truncated version of hLHR missing 17 amino acids in the distal C-terminus, which affects its post-endocytic trafficking and receptor recycling [11]), mLHR, hFSHR, and mFSHR. These receptors are from two species (human and mouse) and are associated with three main hormones: hCG and LH for LHR, and FSH for FSHR. The final category, perturbations, includes pharmacological compounds (Dyngo4a, PitStop2, Es9-17, YM254890) that alter signaling pathways, or genomic modifications that allow for measuring cAMP at specific sub-cellular locations. mCherry was used as a control for compartmentalized conditions.

Each BRET signal time series corresponds to a case with varying features: time, BRET signal, and conditions. The dataset includes three replicates of each case, and we use the average of these replicates to reduce potential bias and data leakage. We performed a main experiment and a five-fold cross-validation. In the main experiment, we divided the dataset into three subsets: (i) a training set for model training, (ii) a validation set for hyperparameter tuning, and (iii) a testing set to evaluate the model on unseen data. The dataset contains 330 training cases, 47 validation cases, and 95 test cases. In the five-fold cross-validation experiment, the complete dataset was partitioned into five subsets, each subset used once as the test set (≈ 95) and the remaining four for training (≈ 378).

Normalization: The measurements vary across different ranges. Min-max normalization scheme scales the measurements to a common range, specifically

between 0 and 1, to eliminate bias toward larger or smaller values [5]. Equation 2 is used to perform normalization in this study:

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2}$$

Where x denotes an original value and x_{norm} is the normalized value, $\max(x)$ and $\min(x)$ are the maximum and minimum values of the input data. In this study, both the BRET signal and the dose are normalized to avoid introducing bias. The normalized signals retain the aspect of the original signals.

Transformation of different conditions: We applied different transformation schemes, using one-hot encoding for ligands, receptors, and perturbations to avoid introducing ordinal relationships between categories [25]. The ligand variable has four possible values. With one-hot encoding, each input is mapped to a 4-dimensional vector, with a 1 in the position corresponding to the input value and 0s in the others. The same scheme is applied to both receptors and perturbations.

2.3 Predicting signalling processes using machine learning models

In this work, we address the task of time series prediction using decision tree, random forest, and XGBoost algorithms. These models are all tree-based but differ in their complexity and underlying techniques. The decision tree is the simplest, while XGBoost is a more advanced method. Random forest employs bagging [2], an ensemble technique that aggregates multiple decision trees to improve robustness and reduce overfitting. In contrast, XGBoost [6] (Extreme Gradient Boosting) uses boosting [8], a sequential ensemble method that builds trees iteratively to minimize a loss function using gradient descent optimization [9]. Random forest is widely used for regression tasks due to its robustness. A decision tree makes predictions by learning a set of rules from the training data [10]. Gradient boosting trees, like XGBoost, combine multiple weak learners to form a strong predictive model.

In this study, we developed a multi-step prediction model to forecast multiple future time steps of the signal simultaneously. Given a fixed number of features, the model predicts not only the next time step but also several subsequent steps. To formulate the time series prediction task, we transformed the sequential signal data into a supervised learning problem by using past observations as input features to predict future time points. Specifically, the input to the model consists of the first 50% of the signal from each time series, and the model predicts the remaining 50% (see Figure 1). This approach requires all cases to have the same length of sequences.

2.4 Configuration of the hyperparameters

This study uses the grid search algorithm from the scikit-learn library to determine the optimal configuration of hyperparameters, including the maximum

Time	BRET	Prediction P = 50%
T_0	x_0	
T_1	x_1	
T_2	x_2	
T_3	x_3	
T_4		$\hat{x}_4$
T_5		$\hat{x}_5$
T_6		$\hat{x}_6$
T_7		$\hat{x}_7$

Fig. 1. An example of the multi-step model predicting future time steps of the signals is shown. The blue bounding box contains the input to the model (T_0 to T_3 represents time steps and x_0 to x_3 are the signal values of these time steps), while the red bounding box represents the predicted signal ($\hat{x}_4$ to $\hat{x}_7$). T_4 to T_7 symbolizes the future time steps.

depth of the tree, the minimum number of samples required to be at a leaf node, the number of trees in the forest, and the learning rate. In the decision tree model, the hyperparameters found were: `max_depth = 10`, `min_samples_leaf = 1`. In the random forest model, the optimal hyperparameters were: `max_depth = 10`, `min_samples_leaf = 1`, `n_estimators = 10`. We selected the following hyperparameters for the XGBoost model: `learning_rate = 0.5`, `max_depth = 3`, and `n_estimators = 10`.

2.5 Performance evaluation

We use three metrics to evaluate the performance of our machine learning models: Mean Absolute Error (MAE), Area Under the Curve (AUC), and time-weighted Root Mean Square Error (RMSE). MAE quantifies the average absolute difference between the actual and predicted values of the time series and is calculated as shown in Equation 3. The time-weighted RMSE applies a weighting factor that increases with time to emphasize errors occurring later in the series, as defined in Equation 4. This weighting reflects the biological importance of accurately predicting later stages of the signal dynamics. In this study, the weights increase approximately linearly from 0 at the start of the time series to a maximum value near 1 at the final time point. For both MAE and time-weighted RMSE, an ideal value is zero, indicating perfect predictions.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

where n is the total number of data points, y_i is the observed value, and $\hat{y}_i$ represents the predicted value.

$$\text{Time-weighted RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{w_i(y_i-\hat{y}_i)^2}{\sum w_i}} \tag{4}$$

where n, y_i, and $\hat{y}_i$ have the same meanings as in Equation 3, and w_i denotes the weighting factor.

The Area Under the Curve (AUC) is calculated using the Trapezoidal rule [19]. The predicted and ground truth AUC values represent the total area under their respective curves plotted over time. Ideally, these values should be as close as possible, indicating that the model accurately captures the overall temporal dynamics of the signal. The AUC is computed according to Equation 5:

$$\text{AUC} = \frac{1}{2}\sum_{i=1}^{n-1}(t_{i+1}-t_i)\left(\mathring{y}_i+\mathring{y}_{i+1}\right), \tag{5}$$

where n is the total number of data points, t_i denotes the time at the ith point, and $\mathring{y}_i$ represents the time series value at t_i, which can be either the observed value y_i or the predicted value $\hat{y}_i$.

2.6 Feature importance

We employ three methods to study feature importance: mean decrease in impurity (MDI) [3], the permutation feature importance technique [3], and SHapley Additive exPlanations (SHAP) values [16]. MDI uses statistics from the training dataset to compute feature importance; however, these features may not necessarily be informative on the testing dataset. The permutation method addresses this issue, as it can be computed on unseen data [18]. SHAP values are a model-agnostic method based on game theory that quantify the contribution of each feature to an individual prediction.

3 Results

3.1 Comparative models

We trained a decision tree, random forest, and XGBoost models using multiple feature vectors and their targets. The implementation of the model and the processing of the data is performed using the scikit-learn library [18]. The codes are available at this link: GitHub.

In figures 2, 4, and 6, we illustrate the predictions of our models on the testing set of the main experiment, different colors indicate the doses, the solid lines show the ground truth and dashed lines show model predictions. The results of this method are highly satisfactory, with the predicted BRET signal values closely matching the ground truth. Figures 3, 5, and 7 show the absolute differences between the ground truth and the model predictions shown in the figures 2, 4, and 6; the continuous line illustrates the absolute differences between the ground

truth and predictions, and the ideal value of this difference is 0. In the random forest model, a significant discrepancy emerges around minute 55 for the 0.3 nM dose. For the 10 nM dose, a notable difference occurs between minutes 30 and 40. These discrepancies may be attributed to abrupt changes in the signal behavior over time. Importantly, one of the challenges in learning local patterns in this time series prediction task is that two points with similar signal values may correspond to different underlying states, depending on the trend (ascending or descending) or variations in the signal slope [15]. In contrast, for doses of 0.1, 1, 3, and 30 nM, the signal remains relatively stable over time, resulting in minimal differences between the ground truth and predictions. A similar pattern is observed with the decision tree model, where greater differences appear for the 0.3 and 10 nM doses, while the discrepancies are smaller for the other doses. The XGBoost model also shows a behavior close to the random forest model for the 0.3 and 10 nM doses. Moreover, we observe a difference between the ground truth and the predicted signal for the 30 nM dose, since the prediction shows a decreasing behavior compared to the original signal.

In conclusion, in the quantitative results, the random forest model presents better results in the different metrics, followed by the decision tree model, and finally XGBoost. Meanwhile, in the qualitative results, the prediction of the XGBoost model tends to try to capture the trend (ascending or descending) of the signal. In contrast, the random forest and decision tree models present predictions with more linear behavior.

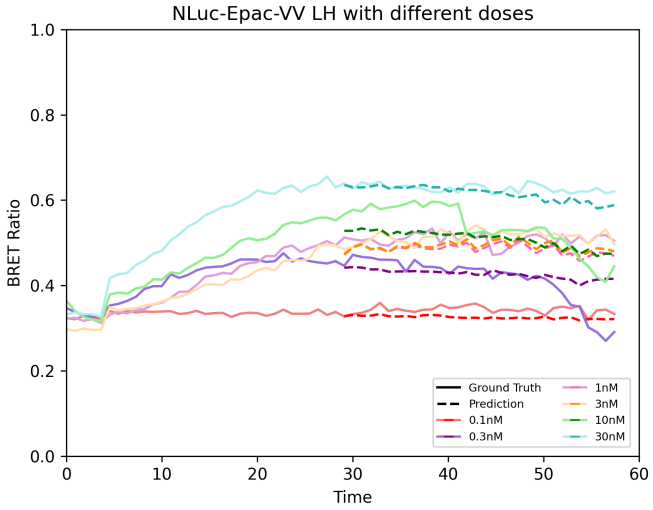

Fig. 2. Random forest predictions. Solid lines show the ground truth and dashed lines show model predictions.

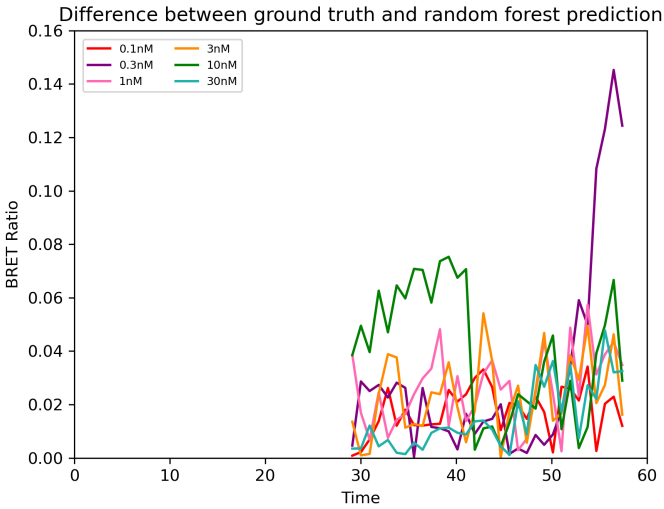

Fig. 3. Random forest difference. The continuous line illustrates the absolute differences between the ground truth and predictions.

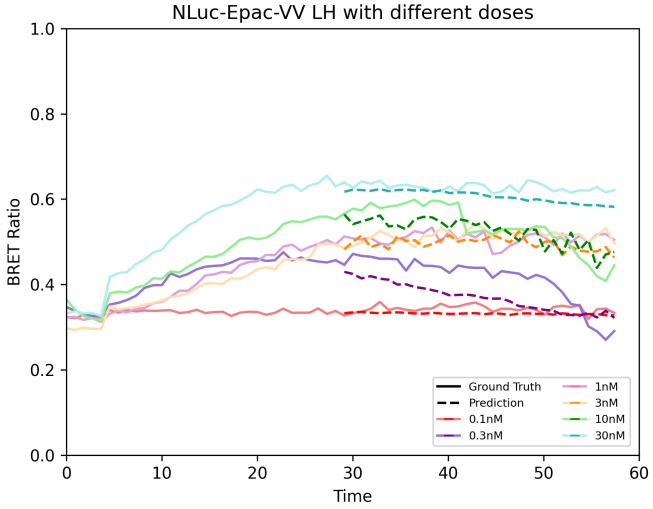

Fig. 4. Decision tree prediction. Solid lines show the ground truth and dashed lines show model predictions.

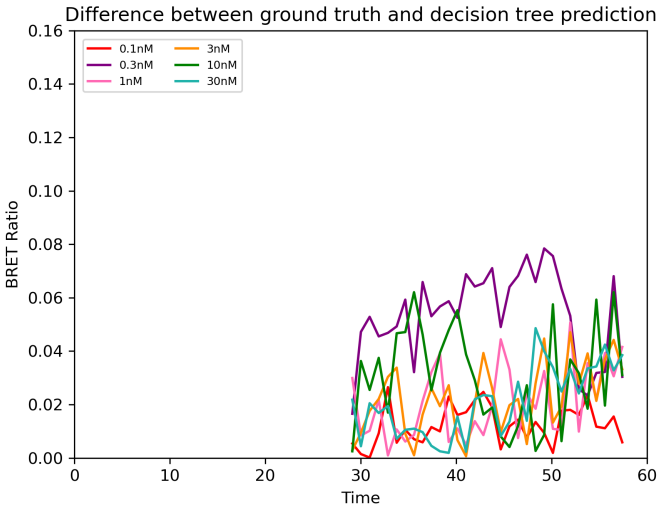

Fig. 5. Decision tree difference. The continuous line illustrates the absolute differences between the ground truth and predictions.

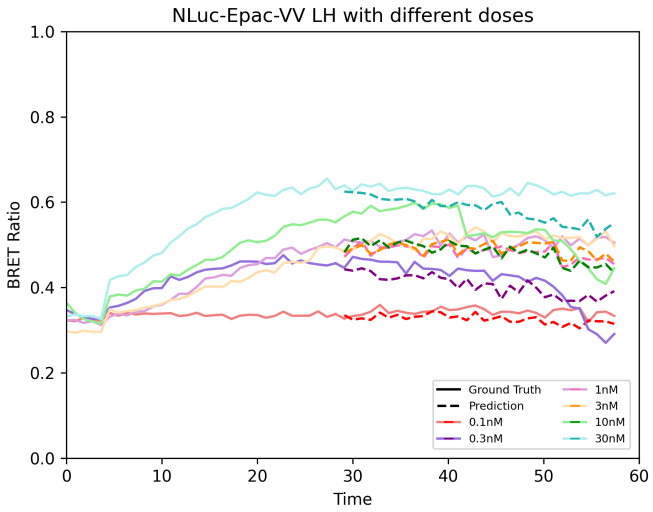

Fig. 6. XGBoost prediction. Solid lines show the ground truth and dashed lines show model predictions.

3.2 Evaluation of prediction accuracy

Figures 8 and 9 present the main experimental results for the evaluation metrics, with colors indicating the different dataset partitions: blue for training, green for

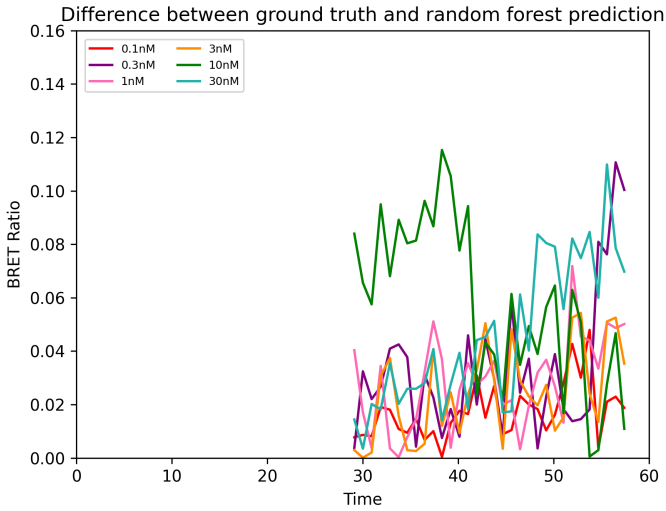

Fig. 7. XGBoost difference. The continuous line illustrates the absolute differences between the ground truth and predictions.

validation, and red for testing. Figure 8 illustrates the MAE results, where the random forest model achieved values of 0.0020, 0.0128, and 0.0140 for the training, validation, and testing sets, respectively. The decision tree model obtained MAE values of 0.0003, 0.0138, and 0.0158, while the XGBoost model yielded 0.0019, 0.0132, and 0.0173 in the respective sets. These errors are close to zero, with the lowest errors generally observed on the training set. Errors increase slightly in the validation and testing sets, but the differences across the sets are minimal. Figure 9 shows the results of the time-weighted RMSE, which was approximately 0 across all sets and models. Moreover, the three models produced very similar results in these evaluation metrics.

Figure 10 presents the AUC results for the prediction models. The decision tree and random forest models show comparable AUCs across the training, validation, and testing sets, with predicted and ground truth AUC values closely aligned. The XGBoost model exhibits a slightly larger gap between predicted and ground truth AUCs in the validation and test sets, which may reflect its higher sensitivity to data and hyperparameters.

In summary, these results indicate that the multi-step prediction approach proves to be both practical and efficient, allowing the prediction of multiple future steps based on the initial signal values, taking approximately 15 seconds to complete training, validation, and testing.

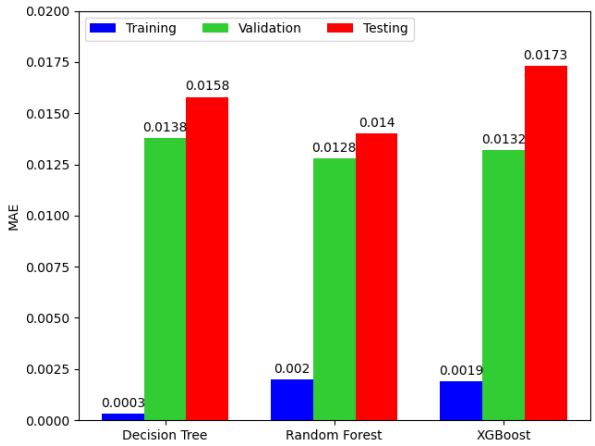

Fig. 8. MAE results.

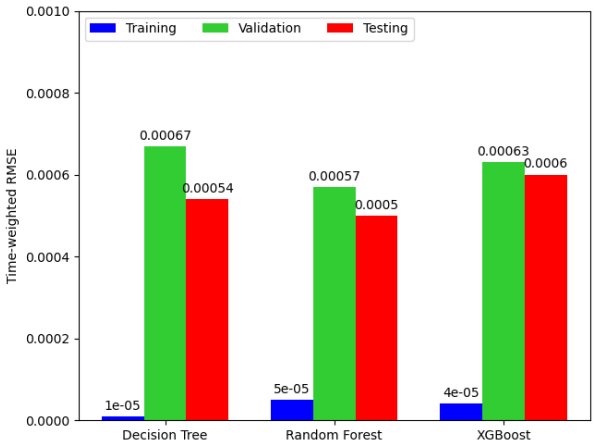

Fig. 9. Time-weighted RMSE results.

3.3 Model performance under cross-validation

Tables 1, 2, and Figure 11 show the cross-validation results on the test sets. The average MAE values (Table 1) were 0.026 for the random forest, 0.027 for both the decision tree and the XGBoost models. The time-weighted RMSE (Table 2) was approximately zero for all three models across all folds. The average AUC values (Figure 11) were similar between the predictions and ground truth across the models. However, fold three exhibited higher error values across all three metrics compared to the other folds, likely reflecting differences in the test set conditions and time series data compared to the training sets. Overall, the error values in all folds were close to zero, which is ideal, and the differences between predicted and ground truth AUCs were minimal. The multi-step prediction ap-

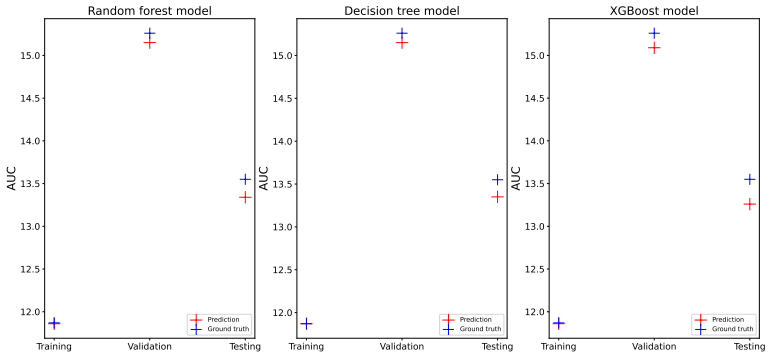

Fig. 10. Main experiment AUC results. Red represents the predicted signal, and blue represents the ground truth signal.

proach remains efficient under cross-validation, requiring less than 16 seconds for training, validation, and testing for each model. Furthermore, the performance of the model under cross-validation is consistent with that observed in the main experiment. Among the models, random forest achieved the best average performance, followed by decision tree and then XGBoost.

Table 1. MAE results in cross-validation.

Folds	Decision Tree	Random Forest	XGBoost
Fold 1	0.0175	0.0154	0.0175
Fold 2	0.0002	0.0002	0.0005
Fold 3	0.0936	0.0937	0.0929
Fold 4	0.0092	0.0092	0.0097
Fold 5	0.0166	0.0134	0.0158
Average	0.027	0.026	0.027

Table 2. Time-weighted RMSE results in cross-validation.

Folds	Decision Tree	Random Forest	XGBoost
Fold 1	0.00072	0.00066	0.00080
Fold 2	0.00000	0.00000	0.00002
Fold 3	0.00253	0.00252	0.00250
Fold 4	0.00030	0.00029	0.00033
Fold 5	0.00056	0.00049	0.00055
Average	0.00082	0.00079	0.00084

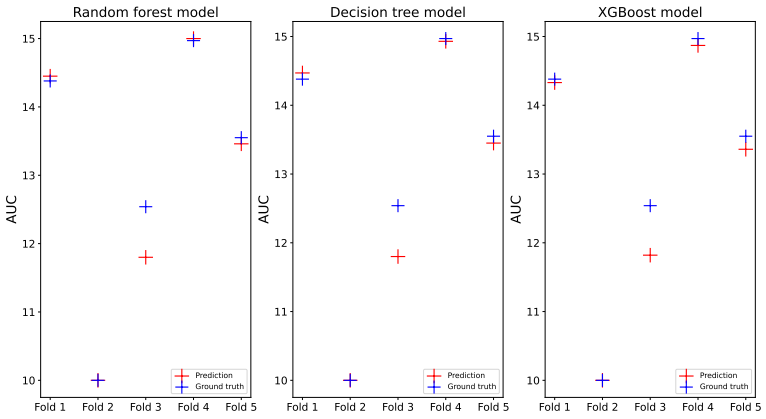

Fig. 11. Cross-validation average AUC results. Red represents the predicted signal, and blue represents the ground truth signal.

3.4 Importance of biological factors

According to SHAP values, the last four BRET signal values had the most significant impact on predictions in the random forest model, whereas in the decision tree model, the last two values of the sequence were most influential. In the XGBoost model, the last five values of the BRET signal were the most critical for prediction. Additionally, the XGBoost model showed greater variability in SHAP values, with the dose variable exhibiting slightly higher importance than the decision tree and random forest models. This diversity of important features in XGBoost reflects its more complex boosting architecture.

According to the MDI feature importance, the decision tree model prioritizes the last two values of the BRET signal, the random forest emphasizes the last four values of the sequence, and the XGBoost model highlights the last five BRET signal values as most important. In the permutation technique, the decision tree presents results similar to MDI, whereas the random forest model highlights the last two values of the BRET signal. In the XGBoost model, the permutation method suggests that the last BRET signal value is the most important feature, showing a modest shift from the MDI results.

These differences in importance likely stem from the distinct methodologies and biases inherent in different methods. All three explanation methods consistently indicate that later BRET signal values play the most critical role, while biological conditions (ligand, receptor, perturbation) did not play a significant role in any model. Meanwhile, dose was the only condition identified as slightly important using the SHAP method with the XGBoost model. The irrelevance

of conditions in the case of MDI may also result from the inflation of the importance of numerical features (i.e., features with high cardinality). Permutation importance can also be biased when features are correlated, as permuting one may not eliminate its information, which can still be inferred from others, leading to underestimation [18]. Although SHAP values minimally capture feature contributions of dose, they ignore causality, which is contradictory in our case since the biological conditions define the signal. Further investigation is necessary to understand the potential contribution of biological conditions under different experimental settings.

3.5 Comparison with traditional models

Traditionally, dynamic models are used to predict signal behavior, especially for GPCR data. To build these models, a mathematical framework must be chosen, such as continuous-time Markov chains (for stochastic dynamics) or ordinary differential equations (ODEs) (for deterministic dynamics) [24]. For example, Bridge et al. [4] used ODEs to model cAMP signaling in response to GPCR ligands. They showed that ligand and receptor concentrations strongly influence cAMP responses and derived ligand-specific kinetic parameters. Similarly, Leelawattanachai et al. [14] modeled GPCR dynamics using ODEs to predict drugs affect on G protein activation, with estimated parameters matching well with experimental observations.

Traditional models face several challenges compared to machine learning approaches. A major issue is parameter identification: parameter values taken from prior work may not reflect the current biological context, leading to errors. Computational costs also increases, particularly during parameter fitting and simulation, as the size and complexity of the network increase [20]. Moreover, ODE-based models are often tailored to specific experimental setups, making cross-study comparison difficult. Finally, while decision tree-based models can handle high-dimensional data through feature selection, they do not capture causal mechanisms. However, they provide a scalable, data-driven way to predict complex signaling dynamics such as cAMP temporal patterns. In this study, we did not perform empirical comparisons with traditional dynamic models, as constructing and calibrating such models requires extensive parameter identification for each experimental condition, which was beyond the scope of this work. Furthermore, deep learning approaches were not pursued due to the limited size of the available dataset.

4 Conclusion

In this work, we presented a machine learning approach to predict the dynamics of intracellular signaling processes. We employed decision tree, random forest, and XGBoost models in combination with a multi-step forecasting strategy. The proposed model effectively captures long-term dependency patterns and achieves a testing mean absolute error of less than 0.018. Although initially applied to

cAMP dynamics, it is quite versatile and and can be adopted to predict dynamics of other intracellular signaling pathways, such as PKA, ERK, AKT, other second messengers production and degradation dynamics (i.e. calcium), recruitment of effector proteins (G proteins and β-arrestins in the case of GPCRs), maybe also ligand binding to a target. It provides a practical framework for characterizing GPCR signaling dynamics and offers potential applications in drug discovery by aiding the characterization of new therapeutic compounds. To expand this methodology, we plan to incorporate additional biological conditions, including different types of ligands, genomic perturbations affecting receptor trafficking and signaling. It will help to understand receptor activity mechanisms and the dynamic signaling networks governing specific biological responses. Future work will focus on validating model predictions through laboratory experiments and developing a user-friendly interface for biologists to generate and validate predicted time courses under various experimental conditions. Additionally, we aim to overcome the current limitation of our multi-step model, i.e., the fixed length of time series by exploring deep learning approaches capable of handling variable-length sequences. This future direction will require collecting additional experimental datasets to sufficiently train and evaluate such models. Finally, we plan to investigate moving average normalization and improve the modeling approach by incorporating local trend information within the time series.

Acknowledgments. We acknowledge the financial support from the iHEALTH, Inria, INRAE, MAbImprove Labex, CNRS, Le Studium Loire Valley, the Bill & Melinda Gates foundation CONTRABODY grant.

Disclosure of Interests. The authors declare that there is no conflict of interest.

References

1. Bao, Y., Xiong, T., Hu, Z.: Multi-step-ahead time series prediction using multiple-output support vector regression. Neurocomputing **129**, 482–493 (2014)
2. Breiman, L.: Bagging predictors. Machine learning **24**(2), 123–140 (1996)
3. Breiman, L.: Random forests. Machine learning **45**, 5–32 (2001)
4. Bridge, L., Chen, S., Jones, B.: Computational modelling of dynamic camp responses to gpcr agonists for exploration of glp-1r ligand effects in pancreatic β-cells and neurons. Cellular Signalling **119**, 111153 (2024)
5. Cabello-Solorzano, K., Ortigosa de Araujo, I., Peña, M., Correia, L., J. Tallón-Ballesteros, A.: The impact of data normalization on the accuracy of machine learning algorithms: A comparative analysis. In: International Conference on Soft Computing Models in Industrial and Environmental Applications. pp. 344–353. Springer (2023)
6. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp. 785–794 (2016)
7. De Pascali, F., Ayoub, M.A., Benevelli, R., Sposini, S., Lehoux, J., Gallay, N., Raynaud, P., Landomiel, F., Jean-Alphonse, F., Gauthier, C., et al.: Pharmacological characterization of low molecular weight biased agonists at the follicle stimulating hormone receptor. International journal of molecular sciences **22**(18), 9850 (2021)

8. Freund, Y.: Boosting a weak learning algorithm by majority. Information and computation **121**(2), 256–285 (1995)
9. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of statistics pp. 1189–1232 (2001)
10. James, G., Witten, D., Hastie, T., Tibshirani, R., et al.: An introduction to statistical learning, vol. 112. Springer (2013)
11. Jean-Alphonse, F., Bowersox, S., Chen, S., Beard, G., Puthenveedu, M.A., Hanyaloglu, A.C.: Spatially restricted g protein-coupled receptor activity via divergent endocytic compartments. Journal of Biological Chemistry **289**(7), 3960–3977 (2014)
12. Kenakin, T.: Biased receptor signaling in drug discovery. Pharmacological reviews **71**(2), 267–315 (2019)
13. Landomiel, F., De Pascali, F., Raynaud, P., Jean-Alphonse, F., Yvinec, R., Pellissier, L.P., Bozon, V., Bruneau, G., Crépieux, P., Reiter, E.: Biased signaling and allosteric modulation at the fshr. Frontiers in endocrinology **10**, 447467 (2019)
14. Leelawattanachai, J., Modchang, C., Triampo, W., Triampo, D., Lenbury, Y.: Modeling and genetic algorithm optimization of early events in signal transduction via dynamics of g-protein-coupled receptors: Internalization consideration. Applied mathematics and computation **207**(2), 528–544 (2009)
15. Lin, T., Guo, T., Aberer, K.: Hybrid neural networks for learning the trend in time series. In: Proceedings of the twenty-sixth international joint conference on artificial intelligence. pp. 2273–2279 (2017)
16. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)
17. Namkung, Y., Radresa, O., Armando, S., Devost, D., Beautrait, A., Le Gouill, C., Laporte, S.A.: Quantifying biased signaling in gpcrs using bret-based biosensors. Methods **92**, 5–10 (2016)
18. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in python. the Journal of machine Learning research **12**, 2825–2830 (2011)
19. Purves, R.D.: Optimum numerical integration methods for estimation of area-under-the-curve (auc) and area-under-the-moment-curve (aumc). Journal of pharmacokinetics and biopharmaceutics **20**(3), 211–226 (1992)
20. Reiter, E., Yvinec, R., Crépieux, P., Poupon, A.: Coupling of recognition and effect in gprc signaling computational modelling approaches. Encyclopedia of the molecular life sciences. In press, London (2016)
21. Taieb, S.B., Bontempi, G., Atiya, A.F., Sorjamaa, A.: A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. Expert systems with applications **39**(8), 7067–7083 (2012)
22. Taieb, S.B., Sorjamaa, A., Bontempi, G.: Multiple-output modeling for multi-step-ahead time series forecasting. Neurocomputing **73**(10-12), 1950–1957 (2010)
23. Tyralis, H., Papacharalampous, G.: Variable selection in time series forecasting using random forests. Algorithms **10**(4), 114 (2017)
24. Yvinec, R., Ayoub, M.A., De Pascali, F., Crépieux, P., Reiter, E., Poupon, A.: Workflow description to dynamically model β-arrestin signaling networks. Beta-Arrestins: Methods and Protocols pp. 195–215 (2019)
25. Zheng, A., Casari, A.: Feature engineering for machine learning: principles and techniques for data scientists. " O'Reilly Media, Inc." (2018)