SUPPLEMENTARY INFORMATION FOR THE PAPER "Uncertainty Quantification in Dynamic Models of Biological Systems Using Conformal Prediction"

Alberto Portela[1], Julio R. Banga[1,*] and Marcos Matabuena[2,*]

[1]Computational Biology Lab, MBG-CSIC, Spanish National Research Council, Pontevedra, Spain
[2]Department of Biostatistics, Harvard University, MA, USA
[*]Corresponding authors: mmatabuena@hsph.harvard.edu, j.r.banga@csic.es

October 1, 2024

1 Theoretical justification of the algorithms

For this theorem, and all results that follow, all probabilities are stated with respect to the distribution of the training data points $(t_1, Y_1), \ldots, (t_n, Y_n)$ and the test data point (t_{n+1}, Y_{n+1}) drawn i.i.d. from an arbitrary distribution P, and we assume implicitly that the regression method m is invariant to the ordering of the data–invariant to permutations. We will treat the sample size $n \geq 2$ and the target coverage level $\alpha \in [0, 1]$ as fixed throughout.

Proposition: The conformal jacknife prediction interval algorithms satisfies:

$$\mathbb{P}(Y_{n+1} \in \widehat{C}^{\alpha}(t_{n+1}))\} \geq 1 - 2\alpha \tag{1}$$

Consequence from Barber *et al.* (2021). The target in the inequality is $1 - \alpha$ that is reached often except in some non-trivial mathematical examples.

2 Matlab implementation of the algorithms

We implemented our `CUQDyn1` and `CUQDyn2` algorithms in Matlab. Parameter estimations were formulated as the minimization of a least squares cost function subject the dynamics (described by the model ODEs) and parameter bounds. These problems are non-convex and were solved using a global hybrid method, enhanced scatter search (`eSS`) due to its good performance an robustness (Villaverde *et al.*, 2019). `eSS` is available in Matlab as part of the `MEIGO` optimization toolbox (Egea *et al.*, 2014). Our code also has dependencies with the the Optimization Toolbox and the Parallel Computing Toolbox. The software for the methodology and the reproduction of the results is available at Zenodo (10.5281/*zenodo*.13644870). All computations were carried out on a PC DELL Precision 7920 workstation with dual Intel Xeon Silver 4210R processors.

3 Comparison with a Bayesian method

Bayesian methods are a classical approach for performing automated uncertainty quantifications by estimating the posterior distribution $P(\theta \mid \mathcal{D}_n)$, where θ represents the parameter of interest and $\mathcal{D}_n = \{X_i\}_{i=1}^n$ denotes the observed data. The key components in Bayesian analysis are the prior distribution $P(\theta)$, which encapsulates our initial beliefs about θ, and the likelihood function $P(\mathcal{D}_n \mid \theta)$, which represents the probability of observing the data $\mathcal{D}_n$ given the parameter θ.

In many practical scenarios, computing the posterior distribution analytically is challenging. Markov Chain provide Monte Carlo (MCMC) methods provides a general and powerful techniques used to estimate the posterior distribution by generating samples from it. Notable MCMC algorithms include Metropolis-Hastings and Gibbs sampling.

Nowadays, there are general software tools available for implementing Bayesian inference and MCMC methods. One such tool is STAN. To use STAN, one writes a model in the STAN modeling language, which involves defining the data, parameters, and model (i.e., prior and likelihood). STAN can be seamlessly integrated with `R` through the `rstan` package (Guo *et al.*, 2020), allowing users to perform Bayesian analyses within the R environment. The `rstan` package provides functions to compile STAN models, fit them to data, and extract samples for posterior analysis. Our implementations of the different case studies are also available in the Zenodo link above.

4 Case studies

Below, we present and discuss these case studies where we have considered synthetic datasets that have been generated considering a noise model described by Equation (2).

$$\tilde{y}_{k,i} = y_k(t_i) + \epsilon_k(t_i) = m_k(x(t_i), \theta) + \epsilon_{k,i}, i = 1, \ldots, n; k = 1, \ldots, n_y. \tag{2}$$

For simplicity, we assumed the errors to be normally distributed, centered around the noise-free data sample, and we have adopted a homoscedastic model, i.e. the variance remains constant across each dimension of the dataset. More specifically,

$$\tilde{y}_{k,i} \sim \mathcal{N}(y_k(t_i), \sigma^2_{k,i}), \quad i = 1, \ldots, n; \ k = 1, \ldots, n_y, \tag{3}$$

where, for any $i \neq j$, $\sigma_{k,i} = \sigma_{k,j} = \sigma_k = \epsilon\mu_k$, with $\mu_k = \sum_{i=1}^{n} y_k(t_i)/n$ capturing the mean value of state k, and ϵ representing the percentage of added noise.

4.1 Case I: Logistic growth model

As our initial case study we considered the well-known logistic model (Tsoularis and Wallace, 2002), governed by a single differential equation with two unknown parameters. This model is frequently used in population growth and epidemic spread modeling.

$$\dot{x} = rx\left(1 - \frac{x}{K}\right). \tag{4}$$

Here, r represents the growth rate, and K denotes the carrying capacity. The initial condition considered in the generation of the datasets was $x(0) = 10$. Additionally, the values of the parameters used were $r = 0.1$ and $K = 100$. The initial condition is assumed to be known across all case studies considered. Since this logistic model has an analytical solution, it facilitates the comparison of our methods' performance with other established conformal methods for algebraic models, such as the jackknife+ (Barber *et al.*, 2021).

To evaluate the performance of our methods on this case study, we considered various scenarios with different noise levels (0%, 1%, 5% and 10%) and dataset sizes (10, 20, 50 and 100 data points). For each combination of noise level and dataset size, we generated 50 different synthetic datasets, totaling 800 unique datasets. By generating multiple datasets for each scenario, we were able to obtain a robust estimate of the methods' behavior and assess their consistency across different realizations of the data.

The comparative analysis of the logistic growth model, as shown in Figure 1, highlights the robustness of the proposed methods `CUQDyn1` and `CUQDyn2` compared to conventional methodologies such as the Bayesian approach implemented with STAN. For a 10-point synthetic dataset with a 10 percent noise level, the predictive regions obtained by both conformal methods showed good coverage without requiring prior calibration of the models, unlike the Bayesian approach. Moreover, both `CUQDyn1` and `CUQDyn2` yield predictive regions comparable to those generated by the jackknife+ method; however, in this particular case, the `CUQDyn1` method shows superior performance.

In terms of computational efficiency, the conformal methods proved to be marginally faster than STAN, even for a problem of this small size, with differences on the order of a few seconds. This makes them more suitable for real-time applications.

To examine the marginal coverage $\mathbb{P}(Y_{n+1} \in \widehat{C}^{\alpha}(X_{n+1}))$ for $\alpha = 0.05, 0.1, 0.5$ of our first algorithm `CUQDyn1`, see Figure 2 for different signal noises and sample sizes. The figure indicates the good empirical performance of our algorithm, achieving the desired nominal level in expectation.

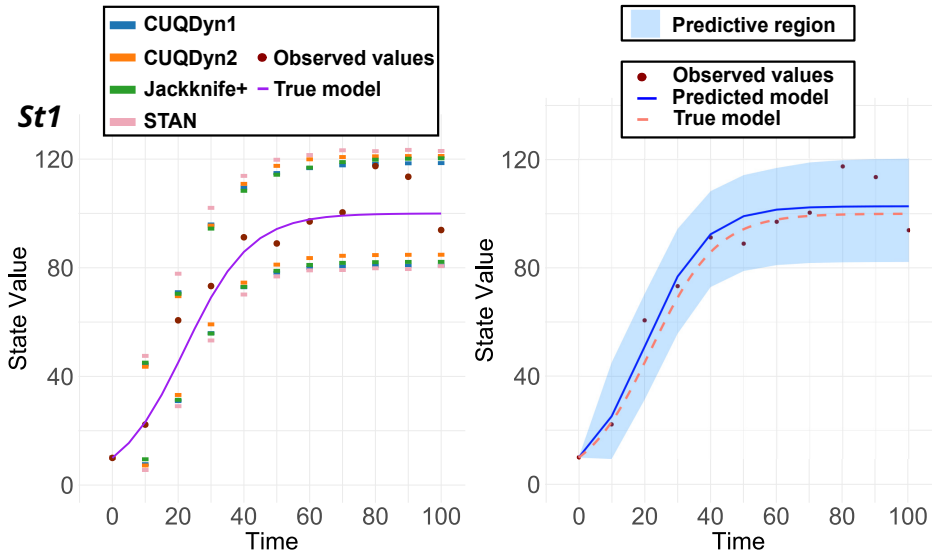


Figure 1: Comparative analysis of the Logistic model predictive regions. This figure presents the 95% predictive regions obtained from a 10-point dataset subjected to 10% noise. The left subplot showcases results using four different methodologies: our two proposed methods (`CUQDyn1` and `CUQDyn2`), the original jackknife+ method and a Bayesian approach implemented with STAN. The right subplot shows the predictive region and the predicted model for the `CUQDyn1` algorithm applied to the same dataset.

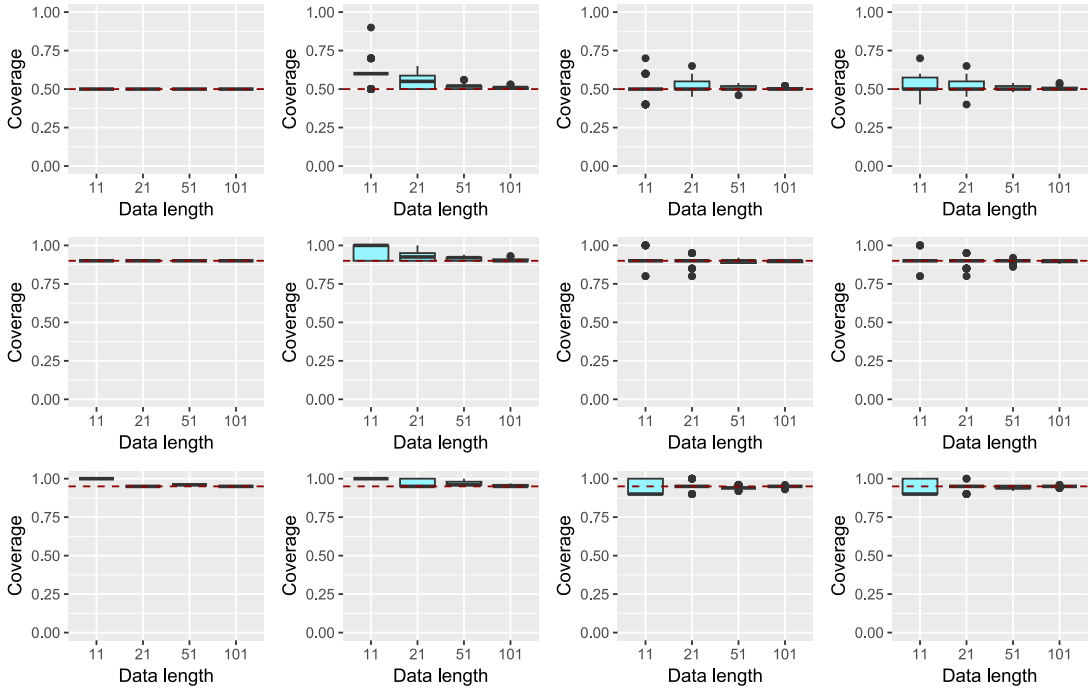


Figure 2: Boxplot of marginal coverage $\mathbb{P}(Y_{n+1} \in \widehat{C}^{\alpha}(X_{n+1}))$ for different sample sizes and $\alpha = 0.05$, 0.1, and 0.5 of our first algorithm, `CUQDyn1`, is presented for different noise levels (0%, 1%, 5%, and 10%) across different columns. The results remain very stable across all examined cases.

Table 1: Numerical results corresponding to the comparative analysis of the Logistic model predictive regions. This table presents the 95% lower and upper predictive bounds (LPB and UPB, respectively) obtained for each time point (t) in a 10-point dataset subjected to 10% noise. The observed data (y) and the true state value (x_{nom}) are also shown. The results are reported for four methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods, the original jackknife+ approach, and a Bayesian method implemented with STAN.

			CUQDyn1		CUQDyn2		Jackknife+		STAN	
t	y	x_{nom}	LPB	UPB	LPB	UPB	LPB	UPB	LPB	UPB
0	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01
10	2.226e+01	2.320e+01	7.698e+00	4.440e+01	7.213e+00	4.357e+01	9.493e+00	4.507e+01	5.545e+00	4.756e+01
20	6.066e+01	4.509e+01	3.092e+01	7.100e+01	3.318e+01	6.954e+01	3.140e+01	7.043e+01	2.904e+01	7.781e+01
30	7.327e+01	6.906e+01	5.593e+01	9.594e+01	5.916e+01	9.552e+01	5.578e+01	9.442e+01	5.324e+01	1.021e+02
40	9.123e+01	8.585e+01	7.284e+01	1.095e+02	7.453e+01	1.109e+02	7.296e+01	1.083e+02	7.015e+01	1.138e+02
50	8.895e+01	9.428e+01	7.816e+01	1.149e+02	8.116e+01	1.175e+02	7.888e+01	1.143e+02	7.681e+01	1.198e+02
60	9.703e+01	9.782e+01	7.997e+01	1.167e+02	8.357e+01	1.199e+02	8.105e+01	1.169e+02	7.905e+01	1.215e+02
70	1.004e+02	9.919e+01	8.055e+01	1.177e+02	8.441e+01	1.208e+02	8.181e+01	1.189e+02	7.917e+01	1.232e+02
80	1.174e+02	9.970e+01	8.073e+01	1.183e+02	8.468e+01	1.210e+02	8.207e+01	1.198e+02	7.982e+01	1.229e+02
90	1.135e+02	9.989e+01	8.079e+01	1.185e+02	8.477e+01	1.211e+02	8.215e+01	1.201e+02	7.953e+01	1.234e+02
100	9.390e+01	9.996e+01	8.081e+01	1.185e+02	8.480e+01	1.212e+02	8.218e+01	1.203e+02	8.060e+01	1.230e+02

4.2 Case II: Lotka-Volterra model

As a second case study, we considered a two species Lotka-Volterra model (Wangersky, 1978), often referred to as the predator-prey model. This model provides a fundamental framework for studying the dynamics between two interacting species. In its simplest form, it describes the interactions between a predator species and a prey species through a set of coupled differential equations with four unknown parameters:

$$\begin{aligned} \dot{x}_1 &= x_1(\alpha - \beta x_2), \\ \dot{x}_2 &= -x_2(\gamma - \delta x_1). \end{aligned} \tag{5}$$

Here, x_1 and x_2 represent the populations of the prey and predator, respectively. The parameters α, β, γ and δ are positive constants representing the interactions between the two species. Specifically, this parameters dictate the growth rates and interaction strengths, capturing the essence of biological interactions such as predation and competition. The initial conditions considered in the generation of the datasets were $x(0) = (10, 5)$. Additionally, the values of the parameters used were $\alpha = \gamma = 0.5$ and $\beta = \delta = 0.02$.

For this case study we generated datasets with the same noise levels (0%, 1%, 5% and 10%) as in the previous example and three different sizes (30, 60 and 120 points). Additionally, for each combination of noise level and dataset size, we generated 50 different synthetic datasets, resulting in a total of 600 unique datasets.

Figure 3 shows the results in a 30-point Lotka-Volterra dataset, indicating that the predictive regions generated by the conformal methods and STAN are similar in terms of coverage. However, as in the previous case, `CUQDyn1` and `CUQDyn2` offer the advantage of not requiring extensive hyperparameter tuning, while also being more computationally efficient. In this particular example, while the bayesian method obtains results within a timeframe on the order of minutes, both conformal methods achieve this in a significantly shorter span, on the order of seconds.

4.3 Case III: Isomerization of α-Pinene

As a third case study, we examined the α-pinene isomerization model. The isomerization process of α-pinene is significant in industry, especially in the production of synthetic fragrances and flavors. These complex biochemical reactions can be effectively modeled using a system of five differential equations with five unknown parameters. The resulting kinetic model has been a classical example in the analysis of multiresponse data (Box *et al.*, 1973). The kinetic equations encapsulate the transformation dynamics of α-pinene into its various isomers through a series of reaction steps:

$$\begin{aligned} \dot{x}_1 &= -(p_1 + p_2)x_1, \\ \dot{x}_2 &= p_1 x_1, \\ \dot{x}_3 &= p_2 x_1 - (p_3 + p_4)x_3 + p_5 x_5, \\ \dot{x}_4 &= p_3 x_3, \\ \dot{x}_5 &= p_4 x_3 - p_5 x_5. \end{aligned} \tag{6}$$

In the equations above, each $p_i \in [0, 1]$, $i = 1, \ldots, 5$ represents a different rate of reaction, defining the conversion speed from one isomer to another. The initial conditions considered in the generation of the datasets were $x(0) = (100, 0, 0, 0, 0)$. Additionally, the values of the parameters used were $p = (5.93e-05, 2.96e-05, 2.05e-05, 2.75e-04, 4.00e-05)$. The dataset generation procedure for this case study mirrored that used for the Logistic model, employing the same noise

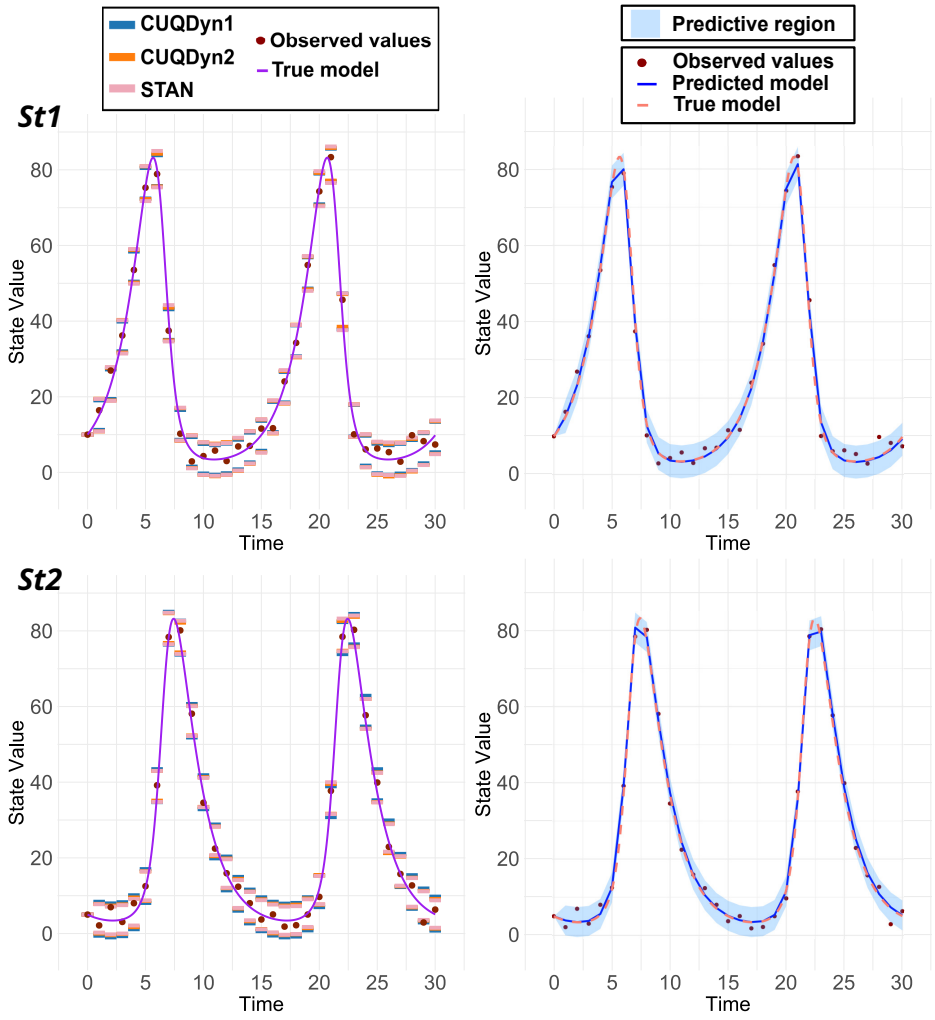


Figure 3: Comparative analysis of the Lotka-Volterra model predictive regions. This figure presents the 95% predictive regions obtained from a 30-point dataset subjected to 10% noise. The two left subplots showcase results using three different methodologies: our two proposed methods (`CUQDyn1` and `CUQDyn2`) and a Bayesian approach implemented with STAN. The two right subplots show the predictive region and the predicted model for the `CUQDyn2` algorithm applied to the same dataset.

Table 2: Numerical results corresponding to the comparative analysis of the Lotka-Volterra model predictive regions for the first state. This table presents the 95% lower and upper predictive bounds (LPB and UPB, respectively) obtained for each time point (t) in a 30-point dataset subjected to 10% noise. The observed data (y) and the true state value (x_{nom}) are also shown. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

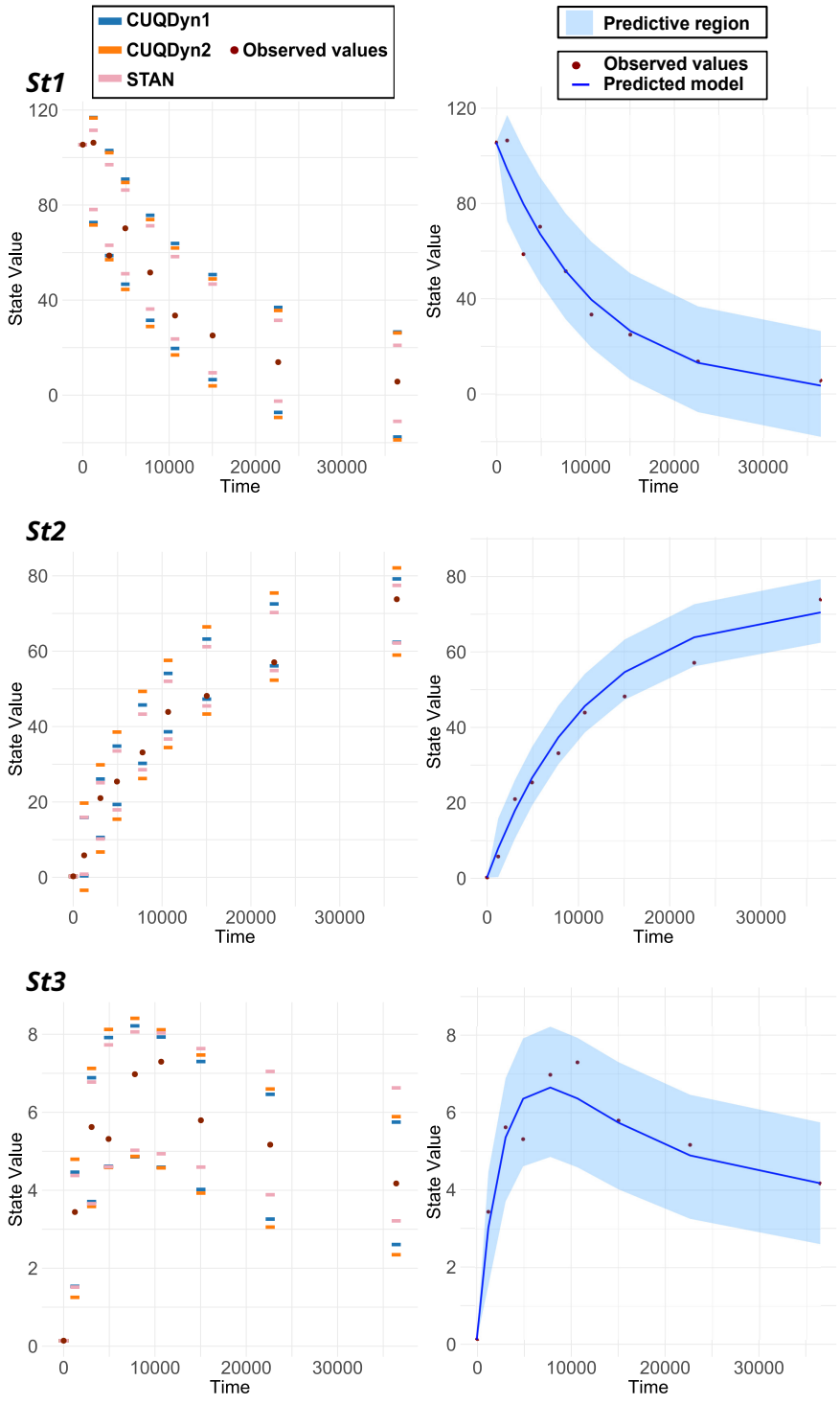
			CUQDyn1		CUQDyn2		STAN	
t	y	x_{nom}	LPB	UPB	LPB	UPB	LPB	UPB
0	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01	1.000e+01
1	1.627e+01	1.510e+01	9.974e+00	2.025e+01	9.535e+00	2.068e+01	1.022e+01	1.984e+01
2	2.597e+01	2.317e+01	1.804e+01	2.831e+01	1.760e+01	2.874e+01	1.832e+01	2.832e+01
3	3.373e+01	3.562e+01	3.046e+01	4.073e+01	3.000e+01	4.114e+01	3.069e+01	4.046e+01
4	5.374e+01	5.379e+01	4.854e+01	5.882e+01	4.807e+01	5.921e+01	4.837e+01	5.925e+01
5	7.607e+01	7.551e+01	7.009e+01	8.036e+01	6.961e+01	8.075e+01	6.965e+01	8.039e+01
6	7.493e+01	8.002e+01	7.344e+01	8.429e+01	7.304e+01	8.418e+01	7.317e+01	8.386e+01
7	4.005e+01	4.053e+01	3.381e+01	4.477e+01	3.347e+01	4.461e+01	3.364e+01	4.447e+01
8	1.435e+01	1.316e+01	7.461e+00	1.773e+01	7.070e+00	1.821e+01	7.415e+00	1.762e+01
9	6.658e+00	5.733e+00	2.295e-01	1.050e+01	-1.836e-01	1.096e+01	6.027e-01	1.044e+01
10	1.580e+00	3.786e+00	-1.595e+00	8.677e+00	-2.018e+00	9.122e+00	-1.310e+00	8.689e+00
11	6.848e+00	3.424e+00	-1.915e+00	8.357e+00	-2.340e+00	8.800e+00	-1.494e+00	8.226e+00
12	6.422e+00	3.821e+00	-1.528e+00	8.744e+00	-1.953e+00	9.187e+00	-1.257e+00	8.470e+00
13	4.250e+00	4.878e+00	-5.077e-01	9.764e+00	-9.314e-01	1.021e+01	-1.349e-01	9.607e+00
14	4.422e+00	6.777e+00	1.309e+00	1.158e+01	8.889e-01	1.203e+01	1.878e+00	1.148e+01
15	8.831e+00	9.921e+00	4.314e+00	1.459e+01	3.899e+00	1.504e+01	4.873e+00	1.474e+01
16	1.784e+01	1.498e+01	9.152e+00	1.942e+01	8.744e+00	1.988e+01	9.879e+00	1.965e+01
17	2.148e+01	2.297e+01	1.681e+01	2.708e+01	1.641e+01	2.755e+01	1.756e+01	2.741e+01
18	3.687e+01	3.532e+01	2.866e+01	3.893e+01	2.828e+01	3.942e+01	2.959e+01	3.937e+01
19	5.202e+01	5.337e+01	4.619e+01	5.667e+01	4.582e+01	5.696e+01	4.771e+01	5.748e+01
20	7.378e+01	7.512e+01	6.812e+01	7.858e+01	6.774e+01	7.888e+01	6.851e+01	7.887e+01
21	8.134e+01	8.039e+01	7.673e+01	8.701e+01	7.624e+01	8.738e+01	7.399e+01	8.524e+01
22	4.315e+01	4.139e+01	3.909e+01	4.999e+01	3.918e+01	5.032e+01	3.616e+01	4.739e+01
23	1.409e+01	1.343e+01	8.884e+00	1.916e+01	8.355e+00	1.950e+01	8.841e+00	1.841e+01
24	1.085e+01	5.804e+00	5.953e-01	1.087e+01	1.348e-01	1.128e+01	1.022e+00	1.065e+01
25	4.580e+00	3.804e+00	-1.537e+00	8.734e+00	-1.974e+00	9.166e+00	-1.097e+00	8.501e+00
26	-2.494e+00	3.423e+00	-1.973e+00	8.298e+00	-2.405e+00	8.736e+00	-1.860e+00	8.504e+00
27	1.040e+00	3.807e+00	-1.669e+00	8.602e+00	-2.098e+00	9.043e+00	-1.317e+00	8.542e+00
28	1.479e+00	4.849e+00	-7.498e-01	9.522e+00	-1.177e+00	9.964e+00	-2.282e-01	9.313e+00
29	1.030e+01	6.729e+00	9.260e-01	1.120e+01	5.132e-01	1.165e+01	1.641e+00	1.140e+01
30	9.117e+00	9.842e+00	3.710e+00	1.398e+01	3.310e+00	1.445e+01	4.792e+00	1.448e+01

Table 3: Numerical results corresponding to the comparative analysis of the Lotka-Volterra model predictive regions for the second state. This table presents the 95% lower and upper predictive bounds (LPB and UPB, respectively) obtained for each time point (t) in a 30-point dataset subjected to 10% noise. The observed data (y) and the true state value (x_{nom}) are also shown. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

			CUQDyn1		CUQDyn2		STAN	
t	y	x_{nom}	LPB	UPB	LPB	UPB	LPB	UPB
0	5.000e+00	5.000e+00	5.000e+00	5.000e+00	5.000e+00	5.000e+00	5.000e+00	5.000e+00
1	-3.196e-02	3.883e+00	-1.135e+00	8.952e+00	-2.394e+00	1.019e+01	-1.270e+00	8.999e+00
2	5.589e+00	3.432e+00	-1.566e+00	8.521e+00	-2.830e+00	9.755e+00	-1.996e+00	8.629e+00
3	8.552e+00	3.716e+00	-1.242e+00	8.845e+00	-2.512e+00	1.007e+01	-1.394e+00	8.545e+00
4	4.930e+00	5.456e+00	5.887e-01	1.068e+01	-6.879e-01	1.190e+01	5.693e-02	1.060e+01
5	1.029e+01	1.208e+01	7.423e+00	1.751e+01	6.165e+00	1.875e+01	7.167e+00	1.778e+01
6	4.178e+01	3.702e+01	3.351e+01	4.366e+01	3.252e+01	4.510e+01	3.298e+01	4.470e+01
7	7.783e+01	7.820e+01	7.447e+01	8.551e+01	7.439e+01	8.697e+01	7.480e+01	8.636e+01
8	7.844e+01	7.709e+01	7.236e+01	8.332e+01	7.230e+01	8.488e+01	7.263e+01	8.453e+01
9	5.618e+01	5.556e+01	5.085e+01	6.130e+01	5.027e+01	6.286e+01	5.121e+01	6.225e+01
10	3.780e+01	3.691e+01	3.207e+01	4.217e+01	3.113e+01	4.371e+01	3.188e+01	4.271e+01
11	2.325e+01	2.402e+01	1.912e+01	2.921e+01	1.798e+01	3.056e+01	1.900e+01	2.969e+01
12	1.937e+01	1.565e+01	1.070e+01	2.079e+01	9.467e+00	2.205e+01	1.030e+01	2.094e+01
13	1.317e+01	1.034e+01	5.354e+00	1.544e+01	4.086e+00	1.667e+01	4.917e+00	1.561e+01
14	8.096e-01	7.036e+00	2.019e+00	1.211e+01	7.399e-01	1.333e+01	1.946e+00	1.230e+01
15	2.381e+00	5.030e+00	-1.277e-02	1.007e+01	-1.292e+00	1.129e+01	-2.094e-01	1.021e+01
16	1.747e+00	3.898e+00	-1.170e+00	8.917e+00	-2.445e+00	1.014e+01	-1.033e+00	9.236e+00
17	2.904e+00	3.435e+00	-1.674e+00	8.413e+00	-2.940e+00	9.645e+00	-1.793e+00	8.531e+00
18	1.409e+00	3.700e+00	-1.480e+00	8.607e+00	-2.732e+00	9.853e+00	-1.505e+00	9.062e+00
19	9.303e+00	5.394e+00	-9.628e-03	1.008e+01	-1.227e+00	1.136e+01	2.618e-01	1.056e+01
20	1.114e+01	1.184e+01	5.597e+00	1.568e+01	4.578e+00	1.716e+01	6.865e+00	1.730e+01
21	3.388e+01	3.622e+01	2.765e+01	3.862e+01	2.740e+01	3.998e+01	3.055e+01	4.150e+01
22	8.328e+01	7.771e+01	7.197e+01	8.306e+01	7.193e+01	8.452e+01	7.272e+01	8.509e+01
23	8.053e+01	7.745e+01	7.579e+01	8.654e+01	7.556e+01	8.814e+01	7.390e+01	8.533e+01
24	5.640e+01	5.599e+01	5.433e+01	6.441e+01	5.317e+01	6.576e+01	5.213e+01	6.339e+01
25	3.562e+01	3.722e+01	3.459e+01	4.468e+01	3.316e+01	4.575e+01	3.307e+01	4.380e+01
26	2.464e+01	2.423e+01	2.078e+01	3.087e+01	1.930e+01	3.188e+01	1.956e+01	3.038e+01
27	1.785e+01	1.578e+01	1.176e+01	2.185e+01	1.029e+01	2.288e+01	1.096e+01	2.153e+01
28	1.065e+01	1.042e+01	6.021e+00	1.611e+01	4.582e+00	1.717e+01	5.093e+00	1.595e+01
29	6.680e+00	7.087e+00	2.427e+00	1.251e+01	1.029e+00	1.361e+01	2.178e+00	1.252e+01
30	4.822e+00	5.060e+00	2.266e-01	1.031e+01	-1.136e+00	1.145e+01	7.373e-02	1.025e+01

levels and dataset sizes. Although we generated synthetic datasets to assess the method's behavior, we illustrated this behavior with a real dataset from Box *et al.* (1973).

Figure 4 shows the resulting regions of the isomerization of α-Pinene by applying the different algorithms to the 9-point real dataset. The results are once again consistent between both conformal algorithms and closely align with the regions obtained using STAN. In terms of computational cost, the conformal algorithms are notably more efficient, requiring less than a minute to compute the regions, whereas the Bayesian approach takes several minutes.



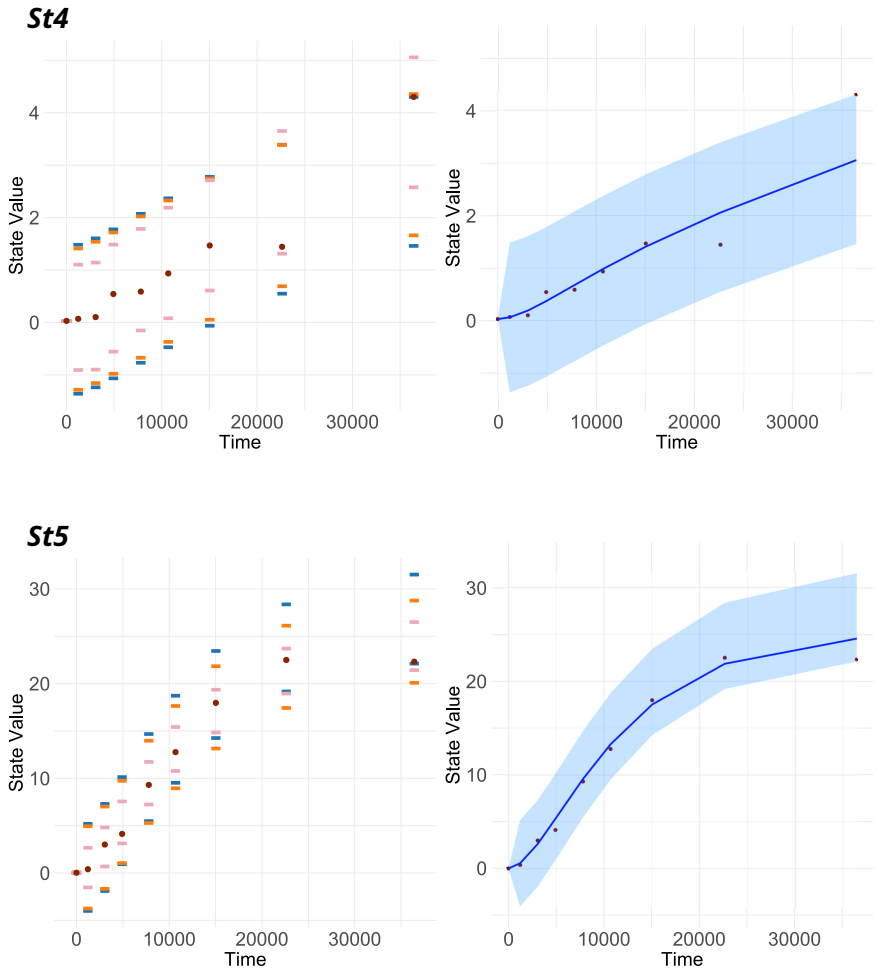


Figure 4: Comparative analysis of the α-pinene isomerization model predictive regions. This figure presents the 95% predictive regions obtained from a 9-point real dataset. It showcases the regions for the five states obtained by using three different methodologies: our two proposed methods (`CUQDyn1` and `CUQDyn2`) and a Bayesian approach implemented with STAN.

4.4 Case IV: NFKB signaling pathway

The Nuclear Factor Kappa-light-chain-enhancer of activated B cells (NFKB) signaling pathway plays a key role in the regulation of immune response, inflammation and cell survival. This pathway is activated in response to various stimuli, including cytokines, stress and microbial infections, leading yo the transcription of target genes involved in immune and inflammatory responses. Here we consider the dynamics of this pathway as described by a system of differential equations (Lipniacki *et al.*, 2004):

Table 4: Numerical results corresponding to the comparative analysis of the α-pinene isomerization model predictive regions for the first state. This table displays the observed data (y) alongside the 95% lower and upper predictive bounds (LPB and UPB, respectively) calculated for each time point (t) in a 9-point real dataset. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

		CUQDyn1		CUQDyn2		STAN	
t	y	LPB	UPB	LPB	UPB	LPB	UPB
0	1.054e+02	1.054e+02	1.054e+02	1.054e+02	1.054e+02	1.054e+02	1.054e+02
1230	1.063e+02	7.266e+01	1.169e+02	7.162e+01	1.167e+02	7.814e+01	1.115e+02
3060	5.875e+01	5.875e+01	1.030e+02	5.702e+01	1.020e+02	6.308e+01	9.701e+01
4920	7.023e+01	4.670e+01	9.091e+01	4.450e+01	8.953e+01	5.110e+01	8.633e+01
7800	5.161e+01	3.149e+01	7.570e+01	2.888e+01	7.391e+01	3.624e+01	7.128e+01
10680	3.350e+01	1.965e+01	6.385e+01	1.691e+01	6.194e+01	2.364e+01	5.830e+01
15030	2.510e+01	6.523e+00	5.073e+01	3.894e+00	4.892e+01	9.418e+00	4.676e+01
22620	1.388e+01	-7.284e+00	3.693e+01	-9.388e+00	3.564e+01	-2.531e+00	3.148e+01
36420	5.688e+00	-1.763e+01	2.658e+01	-1.883e+01	2.620e+01	-1.103e+01	2.096e+01

Table 5: Numerical results corresponding to the comparative analysis of the α-pinene isomerization model predictive regions for the second state. This table displays the observed data (y) alongside the 95% lower and upper predictive bounds (LPB and UPB, respectively) calculated for each time point (t) in a 9-point real dataset. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

		CUQDyn1		CUQDyn2		STAN	
t	y	LPB	UPB	LPB	UPB	LPB	UPB
0	2.769e-01	2.769e-01	2.769e-01	2.769e-01	2.769e-01	2.769e-01	2.769e-01
1230	5.842e+00	4.078e-01	1.589e+01	-3.424e+00	1.970e+01	8.377e-01	1.593e+01
3060	2.101e+01	1.059e+01	2.608e+01	6.730e+00	2.985e+01	1.020e+01	2.511e+01
4920	2.542e+01	1.934e+01	3.482e+01	1.543e+01	3.855e+01	1.792e+01	3.353e+01
7800	3.314e+01	3.024e+01	4.573e+01	2.621e+01	4.933e+01	2.855e+01	4.330e+01
10680	4.389e+01	3.862e+01	5.410e+01	3.444e+01	5.756e+01	3.669e+01	5.201e+01
15030	4.814e+01	4.729e+01	6.321e+01	4.333e+01	6.645e+01	4.547e+01	6.119e+01
22620	5.705e+01	5.611e+01	7.253e+01	5.231e+01	7.543e+01	5.487e+01	7.025e+01
36420	7.378e+01	6.237e+01	7.918e+01	5.897e+01	8.209e+01	6.217e+01	7.745e+01

Table 6: Numerical results corresponding to the comparative analysis of the α-pinene isomerization model predictive regions for the third state. This table displays the observed data (y) alongside the 95% lower and upper predictive bounds (LPB and UPB, respectively) calculated for each time point (t) in a 9-point real dataset. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

		CUQDyn1		CUQDyn2		STAN	
t	y	LPB	UPB	LPB	UPB	LPB	UPB
0	1.409e-01	1.409e-01	1.409e-01	1.409e-01	1.409e-01	1.409e-01	1.409e-01
1230	3.441e+00	1.535e+00	4.463e+00	1.253e+00	4.796e+00	1.515e+00	4.380e+00
3060	5.622e+00	3.709e+00	6.887e+00	3.584e+00	7.127e+00	3.656e+00	6.776e+00
4920	5.316e+00	4.615e+00	7.918e+00	4.583e+00	8.125e+00	4.601e+00	7.730e+00
7800	6.977e+00	4.858e+00	8.217e+00	4.869e+00	8.411e+00	5.025e+00	8.063e+00
10680	7.298e+00	4.590e+00	7.931e+00	4.574e+00	8.117e+00	4.936e+00	8.036e+00
15030	5.795e+00	4.021e+00	7.303e+00	3.928e+00	7.470e+00	4.595e+00	7.635e+00
22620	5.169e+00	3.261e+00	6.461e+00	3.055e+00	6.598e+00	3.885e+00	7.048e+00
36420	4.175e+00	2.608e+00	5.748e+00	2.346e+00	5.889e+00	3.215e+00	6.627e+00

Table 7: Numerical results corresponding to the comparative analysis of the α-pinene isomerization model predictive regions for the fourth state. This table displays the observed data (y) alongside the 95% lower and upper predictive bounds (LPB and UPB, respectively) calculated for each time point (t) in a 9-point real dataset. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

		CUQDyn1		CUQDyn2		STAN	
t	y	LPB	UPB	LPB	UPB	LPB	UPB
0	3.034e-02	3.034e-02	3.034e-02	3.034e-02	3.034e-02	3.034e-02	3.034e-02
1230	6.992e-02	-1.359e+00	1.482e+00	-1.283e+00	1.411e+00	-9.063e-01	1.103e+00
3060	1.027e-01	-1.237e+00	1.604e+00	-1.156e+00	1.539e+00	-8.984e-01	1.143e+00
4920	5.430e-01	-1.065e+00	1.776e+00	-9.783e-01	1.716e+00	-5.551e-01	1.485e+00
7800	5.887e-01	-7.679e-01	2.073e+00	-6.727e-01	2.022e+00	-1.521e-01	1.784e+00
10680	9.356e-01	-4.725e-01	2.368e+00	-3.696e-01	2.325e+00	8.084e-02	2.190e+00
15030	1.468e+00	-6.209e-02	2.779e+00	5.318e-02	2.747e+00	6.100e-01	2.711e+00
22620	1.443e+00	5.493e-01	3.390e+00	6.911e-01	3.385e+00	1.312e+00	3.652e+00
36420	4.301e+00	1.460e+00	4.301e+00	1.662e+00	4.356e+00	2.578e+00	5.056e+00

Table 8: Numerical results corresponding to the comparative analysis of the α-pinene isomerization model predictive regions for the fifth state. This table displays the observed data (y) alongside the 95% lower and upper predictive bounds (LPB and UPB, respectively) calculated for each time point (t) in a 9-point real dataset. The results are reported for three methodologies: the proposed `CUQDyn1` and `CUQDyn2` methods and a Bayesian method implemented with STAN.

		CUQDyn1		CUQDyn2		STAN	
t	y	LPB	UPB	LPB	UPB	LPB	UPB
0	3.252e-02	3.252e-02	3.252e-02	3.252e-02	3.252e-02	3.252e-02	3.252e-02
1230	4.046e-01	-4.002e+00	5.192e+00	-3.748e+00	4.940e+00	-1.514e+00	2.666e+00
3060	3.011e+00	-1.894e+00	7.300e+00	-1.668e+00	7.020e+00	6.917e-01	4.812e+00
4920	4.136e+00	9.374e-01	1.013e+01	1.051e+00	9.739e+00	3.127e+00	7.557e+00
7800	9.308e+00	5.485e+00	1.468e+01	5.285e+00	1.397e+01	7.238e+00	1.173e+01
10680	1.277e+01	9.528e+00	1.872e+01	8.952e+00	1.764e+01	1.078e+01	1.542e+01
15030	1.797e+01	1.426e+01	2.345e+01	1.315e+01	2.184e+01	1.484e+01	1.935e+01
22620	2.250e+01	1.917e+01	2.837e+01	1.743e+01	2.612e+01	1.895e+01	2.370e+01
36420	2.232e+01	2.212e+01	3.152e+01	2.008e+01	2.877e+01	2.142e+01	2.650e+01

$$
\begin{aligned}
\dot{IKKn} &= kprod - kdeg \cdot IKKn \\
&\quad - Tr \cdot k1 \cdot IKKn, \\
\dot{IKKa} &= Tr \cdot k1 \cdot IKKn - k3 \cdot IKKa \\
&\quad - Tr \cdot k2 \cdot IKKa \cdot A20 - kdeg \cdot IKKa \\
&\quad - a2 \cdot IKKa \cdot IkBa + t1 \cdot IKKaIkBa \\
&\quad - a3 \cdot IKKa \cdot IkBaNFkB \\
&\quad + t2 \cdot IKKaIkBaNFkB, \\
\dot{IKKi} &= k3 \cdot IKKa + Tr \cdot k2 \cdot IKKa \cdot A20 \\
&\quad - kdeg \cdot IKKi, \\
\dot{IKKaIkBa} &= a2 \cdot IKKa \cdot IkBa - t1 \cdot IKKaIkBa, \\
\dot{IKKaIkBaNFkB} &= a3 \cdot IKKa \cdot IkBaNFkB \\
&\quad - t2 \cdot IKKaIkBaNFkB, \\
\dot{NFkB} &= c6a \cdot IkBaNFkB - a1 \cdot NFkB \cdot IkBa \\
&\quad + t2 \cdot IKKaIkBaNFkB - i1 \cdot NFkB, \\
\dot{NFkBn} &= i1 \cdot kv \cdot NFkB - a1 \cdot IkBan \cdot NFkBn, \\
\dot{A20} &= c4 \cdot A20t - c5 \cdot A20, \\
\dot{A20t} &= c2 + c1 \cdot NFkBn - c3 \cdot A20t, \\
\dot{IkBa} &= -a2 \cdot IKKa \cdot IkBa \\
&\quad - a1 \cdot IkBa \cdot NFkB \\
&\quad + c4a \cdot IkBat - c5a \cdot IkBa - i1a \cdot IkBa \\
&\quad + e1a \cdot IkBan, \\
\dot{IkBan} &= -a1 \cdot IkBan \cdot NFkBn + i1a \cdot kv \cdot IkBa \\
&\quad - e1a \cdot kv \cdot IkBan, \\
\dot{IkBat} &= c2a + c1a \cdot NFkBn - c3a \cdot IkBat, \\
\dot{IkBaNFkB} &= a1 \cdot IkBa \cdot NFkB - c6a \cdot IkBaNFkB \\
&\quad - a3 \cdot IKKa \cdot IkBaNFkB \\
&\quad + e2a \cdot IkBanNFkBn
\end{aligned}
\tag{7}
$$

Table 9: Parameter estimation comparison for the first three case studies. Since the parameter estimation process is identical for both algorithms proposed in this paper, for simplicity, we will refer to the parameters obtained with them as CUQDyn. For the case study on the isomerization of α-pinene, a real dataset was considered, and thus the nominal parameters for this problem are unknown.

	Logistic		Lotka-Volterra				α-Pinene				
	r	K	α	β	γ	δ	p_1	p_2	p_3	p_4	p_5
True	1.000e-01	1.000e+02	5.000e-01	2.000e-02	5.000e-01	2.000e-02	NA	NA	NA	NA	NA
STAN	1.152e-01	1.021e+02	4.992e-01	2.004e-02	5.000e-01	2.001e-02	5.998e-05	2.797e-05	1.860e-05	2.826e-04	5.366e-05
CUQDyn	1.112e-01	1.027e+02	4.984e-01	2.001e-02	5.001e-01	2.000e-02	6.307e-05	2.841e-05	1.609e-05	2.729e-04	4.365e-05

In agreement with the scenario considered by , we assume that the available measurements are determined by the observation function $m : \mathbb{R}^{15} \to \mathbb{R}^6$, which is defined as follows:

$$\begin{aligned} m(\cdot) =& (NFkBn, IkBa + IkBaNFkB, A20t, \\ & IKKn + IKKa + IKKi, IKKa, IkBat). \end{aligned} \tag{8}$$

Out of the system of 15 equations, only 6 observables, defined by the function m, are available. The parameter values used in the generation of the datasets are as follows:

$$\begin{aligned}
\mathrm{a1} &= 5e-01, & \mathrm{a2} &= 2e-01, & \mathrm{t1} &= 1e-01, \\
\mathrm{a3} &= 1e+00, & \mathrm{t2} &= 1e-01, & \mathrm{c1a} &= 5e-07, \\
\mathrm{c2a} &= 0e+00, & \mathrm{c3a} &= 4e-04, & \mathrm{c4a} &= 5e-01, \\
\mathrm{c5a} &= 1e-04, & \mathrm{c6a} &= 2e-05, & \mathrm{c1} &= 5e-07, \\
\mathrm{c2} &= 0e+00, & \mathrm{c3} &= 4e-04, & \mathrm{c4} &= 5e-01, \\
\mathrm{c5} &= 3e-04, & \mathrm{k1} &= 2.5e-03, & \mathrm{k2} &= 1e-01, \\
\mathrm{k3} &= 1.5e-03, & \mathrm{kprod} &= 2.5e-05, & \mathrm{kdeg} &= 1.25e-04, \\
\mathrm{kv} &= 5e+00, & \mathrm{i1} &= 2.5e-03, & \mathrm{e2a} &= 1e-02, \\
\mathrm{i1a} &= 1e-03, & \mathrm{e1a} &= 5e-04, & \mathrm{c1c} &= 5e-07, \\
\mathrm{c2c} &= 0e+00, & \mathrm{c3c} &= 4e-04.
\end{aligned}$$

It should be noted that in this problem, which involves 29 unknown parameters and 15 state variables, we only have access to 6 observable outputs. This discrepancy between the number of parameters and the available observables presents a challenge in the context of parameter identifiability, and is very commmon in systems biology applications. Identifiability refers to the ability to uniquely determine the model parameters based on the available data. When a system lacks identifiability, inferring unique parameter values from observable data becomes challenging, if not impossible. However, as mentioned in the introduction, by characterizing the impact of this lack of identifiability using appropriate uncertainty quantification (UQ) methods, it might still be possible to make useful predictions.

The dataset generation process was the same that the one used in the previous case studies.

Figure 5 shows the results of applying our two methods to a 13-point synthetic dataset. Both methods based on conformal inference yielded results that are in close agreement with each other. However, in this case, we were not able to obtain adequate predictive regions using STAN, even after many hours of computation, probably due to the partial lack of identifiability. Remarkably, our `CUQDyn1` and `CUQDyn2` algorithms can compute the regions in just a few minutes using a standard PC.

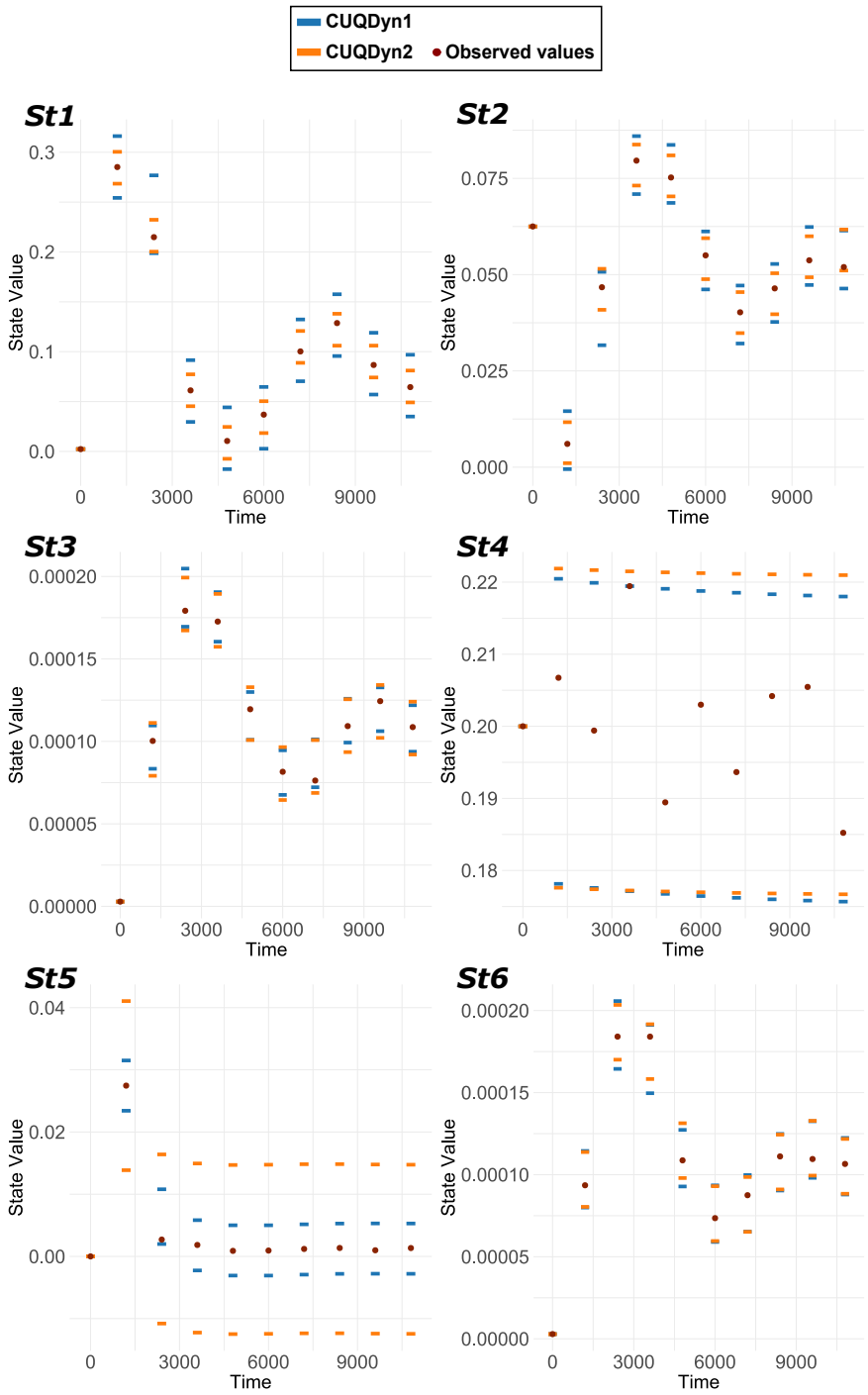


Figure 5: Comparative analysis of the NFkb signaling pathway model predictive regions. This figure presents the 95% predictive regions obtained from a 13-point synthetic dataset. It showcases the regions for the six observables obtained by using our two proposed methods: `CUQDyn1` and `CUQDyn2`.

References

Barber, R. F. *et al.* (2021). Predictive inference with the jackknife+. *The Annals of Statistics*, **49**(1), 486 – 507.

Box, G. *et al.* (1973). Some problems associated with the analysis of multiresponse data. *Technometrics*, **15**(1), 33–51.

Egea, J. A. *et al.* (2014). Meigo: an open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. *BMC bioinformatics*, **15**(1), 136.

Guo, J. *et al.* (2020). Package 'rstan'. *URL https://cran. r—project. org/web/packages/rstan/(2020 x y)*.

Lipniacki, T. *et al.* (2004). Mathematical model of NFκB regulatory module. *J. Theor. Biol.*, **228**, 195–215.

Tsoularis, A. and Wallace, J. (2002). Analysis of logistic growth models. *Mathematical Biosciences*, **179**(1), 21–55.

Villaverde, A. F. *et al.* (2019). Benchmarking optimization methods for parameter estimation in large kinetic models. *Bioinformatics*, **35**(5), 830–838.

Wangersky, P. J. (1978). Lotka-volterra population models. *Annual Review of Ecology and Systematics*, **9**, 189–218.