

Systems biology

Bayesian modelling of time series data (BayModTS) – a FAIR workflow to process sparse and highly variable data

Sebastian Höpfl ¹, Mohamed Albady^{2,3}, Uta Dahmen², Karl-Heinz Herrmann⁴,
Eva Marie Kindler⁵, Matthias König⁶, Jürgen Rainer Reichenbach⁴, Hans-Michael Tautenhahn⁷,
Weiwei Wei², Wan-Ting Zhao⁴, Nicole Erika Radde ^{1,*}

¹Institute for Stochastics and Applications, University of Stuttgart, 70569 Stuttgart, Germany

²Experimental Transplantation Surgery, Department of General, Vascular and Visceral Surgery, University Hospital Jena, 07745 Jena, Germany

³Department of Pathology, Faculty of Veterinary Medicine, Menoufia University, Shebin Elkom, Menoufia, Egypt

⁴Medical Physics Group, Institute for Diagnostic and Interventional Radiology, University Hospital Jena, 07743 Jena, Germany

⁵Clinic for General, Visceral and Vascular Surgery, Jena University Hospital, 07747 Jena, Germany

⁶Institute for Biology, Faculty of Life Sciences, Humboldt-University Berlin, 10115 Berlin, Germany

⁷Clinic for Visceral, Transplantation, Thoracic and Vascular Surgery, Leipzig University Hospital, 04103 Leipzig, Germany

*Corresponding author. Institute for Stochastics and Applications, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany.
E-mail: nicole.radde@simtech.uni-stuttgart.de (N.E.R.)

Associate Editor: Anthony Mathelier

Abstract

Motivation: Systems biology aims to better understand living systems through mathematical modelling of experimental and clinical data. A pervasive challenge in quantitative dynamical modelling is the integration of time series measurements, which often have high variability and low sampling resolution. Approaches are required to utilize such information while consistently handling uncertainties.

Results: We present BayModTS (Bayesian modelling of time series data), a new FAIR (findable, accessible, interoperable, and reusable) workflow for processing and analysing sparse and highly variable time series data. BayModTS consistently transfers uncertainties from data to model predictions, including process knowledge via parameterized models. Further, credible differences in the dynamics of different conditions can be identified by filtering noise. To demonstrate the power and versatility of BayModTS, we applied it to three hepatic datasets gathered from three different species and with different measurement techniques: (i) blood perfusion measurements by magnetic resonance imaging in rat livers after portal vein ligation, (ii) pharmacokinetic time series of different drugs in normal and steatotic mice, and (iii) CT-based volumetric assessment of human liver remnants after clinical liver resection.

Availability and implementation: The BayModTS codebase is available on GitHub at <https://github.com/Systems-Theory-in-Systems-Biology/BayModTS>. The repository contains a Python script for the executable BayModTS workflow and a widely applicable SBML (systems biology markup language) model for retarded transient functions. In addition, all examples from the paper are included in the repository. Data and code of the application examples are stored on DaRUS: <https://doi.org/10.18419/darus-3876>. The raw MRI ROI voxel data were uploaded to DaRUS: <https://doi.org/10.18419/darus-3878>. The steatosis metabolite data are published on FairdomHub: 10.15490/fairdomhub.1.study.1070.1.

1 Introduction

Biology is a field of extremes with extensive high-throughput data (e.g. omics data) on the one hand and sparse data (e.g. western blots) on the other. These data are generally characterized by high variability caused by inherent biological variation and measurement uncertainties. Due to cost and ethical aspects, small sample sizes further complicate the calculation of reliable statistics. In such a setting, it is essential to use the acquired data wisely to make progress in the field and steer biological and medical research in promising directions.

A recurring challenge in biomedical data analysis is comparing the temporal response under different conditions, i.e. comparing different time series. The current standards in biology for investigating differences in time series data are pairwise hypothesis tests, testing each time point separately

(Huang *et al.* 2022, Ko *et al.* 2023) or a comparison of derived parameters such as the area under the curve, e.g. in pharmacokinetics (Scheff *et al.* 2011, Chen *et al.* 2022). Both methods are frequency-oriented and have the problem of extensive multiple testing if each time point is used as a single sample. Moreover, information about the dynamics is lost when using a summary statistic, such as the area under the curve. Methods that take the dynamics of the data into account are, e.g. semi metric ensemble time series (SMETS) (Tapinos and Mendes 2013) and the use of confidence bands (Korpela *et al.* 2014). SMETS can compare multivariate time series by measuring their pairwise distance, and nonparametric frequentistic confidence bands compute a minimum envelope of a time series but were not yet used to compare different time series.

Received: 29 January 2024; Revised: 11 April 2024; Editorial Decision: 30 April 2024; Accepted: 13 May 2024

© The Author(s) 2024. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Another challenge with sparse and noisy time series data is dealing with outliers, which can drastically distort the results. As defined by Hawkins, an outlier is an observation that deviates so strongly from other observations that it is suspected of being caused by a different mechanism (Hawkins 1980). However, the question of suspicion is not easy to answer if only a few replicates are available. The z -score, the Grubbs test, the Tietjen–Moore test, and Dixon’s Q test are existing methods for detecting single-point outliers. Further, time series filtering (e.g. bandpass filters) and dynamic linear models (Campagnoli et al. 2009) can smooth data series and identify outliers by considering the dynamics. We argue against outlier classification in sparse datasettings without incorporating process knowledge. The potential pitfall is a substantial information loss by excluding outliers in sparse datasettings. Using Bayesian inference with all available data, outliers are inherently corrected by incorporating process knowledge.

Fortunately, time series data are an example where knowledge of the underlying dynamics is often available. Here, we present BayModTS (Bayesian modelling of time series data), a novel findable, accessible, interoperable, and reusable (FAIR) workflow for processing time series data that incorporates process knowledge. BayModTS is designed for sparse data with low temporal resolution, a small number of replicates and high variability between replicates. BayModTS is based on a simulation model, representing the underlying data generation process. This simulation model can be an ordinary differential equation (ODE), a time-parameterized function, or any other dynamic modelling approach. For model calibration, posterior distributions of the parameters are sampled using a Bayesian approach. Ensembles of these posterior distributions are then used for forward simulations that can be compared to the data. This framework allows us to investigate whether conditions are credibly different, using measures from statistics and information theory such as credibility intervals (CIs). Importantly, BayModTS is based on systems biology modelling standards and is easily understandable and accessible for experts and novices.

We demonstrate the broad applicability of BayModTS using three datasets that investigate different characteristics of the liver in animals and humans. These data were collected using fundamentally different measurement methods: (i) quantification of blood perfusion changes in the liver lobes of rats after portal vein ligation (PVL) using magnetic resonance imaging (MRI), (ii) quantification of metabolic drug concentrations in mice via ultra-performance liquid chromatography–tandem mass spectrometry, and (iii) computed tomography (CT)-based volumetric assessment of the liver remnant in human patients after clinical liver resections (Supplement ‘Postoperative liver regeneration after resection’). We use BayModTS to statistically compare the time-series data of different conditions in these application scenarios. In particular, BayModTS deals with intra- and inter-individual variation in the studies by focusing on condition differences rather than population heterogeneity. The results are discussed in their respective biological and medical contexts.

2 The BayModTS workflow

The main applications of BayModTS are (i) to statistically test whether different datasets stem from the same data

generating process and (ii) to process time series data and create continuous input functions with uncertainties for other models. In brief, BayModTS infers the dynamics of time series data via Retarded Transient Functions (RTFs) (Fig. 1, step 1). Bayesian parameter estimation via Markov Chain Monte Carlo (MCMC) sampling is used to infer the posterior distribution, given the model and the data (Fig. 1, step 2). Parameter ensembles are simulated from the posterior distribution to transfer the uncertainty from the parameter to the data space (Fig. 1, step 3). These steps are repeated for n conditions to compare their dynamics (Fig. 1, step 4). A more detailed description of the workflow follows.

Inputs to BayModTS are (i) serial data $D = \{m^{(j)}(t_k)\}$, with time points $k = 1, \dots, T$ which do not have to be equidistant, and $j = 1, \dots, N$ replicates per time point (Fig. 1A), (ii) an associated simulation model $s(t, \theta)$ with parameters θ (Fig. 1B), and (iii) a prior distribution $p(\theta)$ on the model parameters.

The simulation model s is defined in SBML (systems biology markup language) (Keating et al. 2020) and leverages PEstab (Schmiester et al. 2021) to integrate the experimental data and an error model. BayModTS uses the explicit RTF (Kreutz 2020)

$$f_{TF}(t) = \underbrace{A_{\text{sus}}(1 - e^{-\frac{t}{\tau_1}})}_{\text{sustained response}} + \underbrace{A_{\text{trans}}(1 - e^{-\frac{t}{\tau_{11}}})e^{-\frac{t}{\tau_2}}}_{\text{transient response}} + p_0 \quad (1)$$

with transformed time

$$t = \log_{10} \left(10^{t_{\text{real}} \frac{10}{T_{\text{range}}}} + 10^{T_{\text{shift}}} \right) - \log_{10}(1 + 10^{T_{\text{shift}}}) \quad (2)$$

as a universal simulation model $s(t, \theta)$ with parameters

$$\theta = (A_{\text{sus}}, t_1, A_{\text{trans}}, t_{11}, t_2, T_{\text{shift}}, p_0). \quad (3)$$

Equation (1) consists of a transient and a persistent activation term. RTFs can describe the responses of many biological processes to various inputs without knowing the system’s detailed interactions a priori. RTF responses can be delayed by a shift parameter T_{shift} via a nonlinear time transformation that leaves $t = 0$ invariant (Equation 2). An immediate response corresponds to $T_{\text{shift}} = -2$ (Kreutz 2020), which is a valid assumption for all examples in this paper and was therefore fixed a priori for all application examples. The parameter T_{range} is fixed to the range of the observation interval. Even if the system under investigation can be described by an ODE model, RTFs can serve as computationally advantageous surrogate models because they can be evaluated directly without numerical integration.

PEstab is a reproducible format for parameter estimation problems. It contains information about the estimated parameters θ , the experimental conditions, the observed data D and the error model needed to define the likelihood function $L_D(\theta)$. Here, we assume additive, independent, and identically distributed (i.i.d.) Gaussian measurement errors with equal variances σ^2 for all measurements,

$$m(t_k) = s(t_k, \theta) + \varepsilon \quad \varepsilon \sim \mathcal{N}(0, \sigma^2), \quad (4)$$

leading to

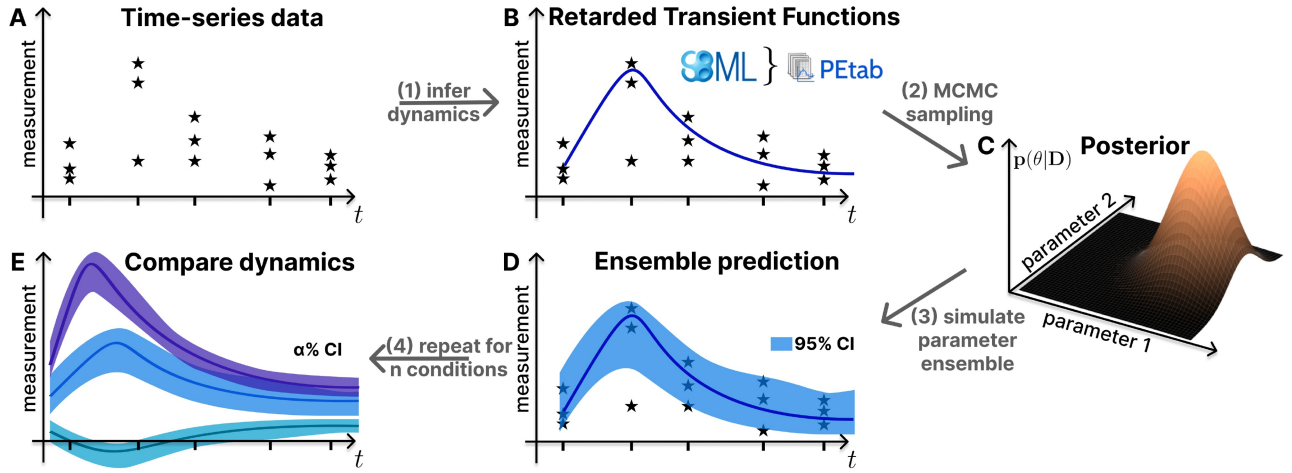


Fig. 1. The FAIR BayModTS workflow reveals credible regions in the data space and allows statistical testing whether different datasets stem from the same data generating process. BayModTS applies a statistical Bayesian framework and a simulation model $s(t, \theta)$ to transform sparse and highly variable time series data into less noisy time courses with uncertainty estimates for model states. (A) Time series data and (B) retarded transient functions as a universal simulation model $s(t, \theta)$ are used as input together with parameter priors. (C) The Posterior distribution $p(\theta|D)$ is inferred using Markov Chain Monte Carlo sampling. (D) Simulated parameter posterior sample trajectories (Ensembles) are used to infer summary statistics that quantify model state uncertainties. We use 95% CI tubes. (E) Dynamics under different conditions can be visually compared via CI tubes.

$$L_D(\theta) = \prod_{k=1}^T \prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(s(t_k, \theta) - m^{(j)}(t_k))^2}{2\sigma^2}\right). \quad (5)$$

The variance σ^2 can be treated as an additional unknown parameter that has to be estimated in the inverse problem or set to a fixed value for each data point or measurement series. Different approaches exist for formulating an appropriate error model and choosing hyperparameters for this model (Kreutz *et al.* 2007, Thomaseth and Radde 2023). Our pragmatic approach pools all data for estimating σ , assuming that the noise is mainly determined by the measurement method and does not vary much between time points. PETab can be used for maximum likelihood estimation or sampling-based approaches. Here, we infer parameters from the posterior distribution (Fig. 1C)

$$p(\theta|D) \propto p(D|\theta) \cdot p(\theta) \quad (6)$$

in a statistical Bayesian setting to quantify the uncertainty in the parameter space. Inference is based on MCMC sampling techniques to generate posterior samples $\theta^{(i)}$, $i = 1, \dots, P$ (Fig. 1, step 2). For RTFs, uniform priors with wide bounds around the measurement data can be used (Kreutz 2020) and adapted for other scenarios.

Posterior predictive distributions

$$p(\tilde{D}|D) = \int p(\tilde{D}, \theta|D) d\theta \quad (7)$$

propagate the uncertainty from the parameter into the data space. Using factorization of the joint density $p(\tilde{D}, \theta|D)$ and exploiting that the posterior predictive distribution (PPD) of any dataset \tilde{D} is independent of D given the model parameters, we can reformulate Equation (7),

$$p(\tilde{D}|D) = \int p(\tilde{D}|\theta, D) p(\theta|D) d\theta \quad (8a)$$

$$= \int p(\tilde{D}|\theta) p(\theta|D) d\theta. \quad (8b)$$

Using the posterior samples $\theta^{(i)}$, these PPDs can be estimated via Monte Carlo integration,

$$p(\tilde{D}|D) \approx \frac{1}{P} \sum_{i=1}^P p(\tilde{D}|\theta^{(i)}). \quad (9)$$

A lower bound for the variability of the PPD at time t can be obtained by a posterior parameter ensemble prediction of simulation model trajectories $s(t, \theta)$ evaluated with posterior samples $\theta^{(i)}$,

$$s^{(i)}(t) = s(t, \theta^{(i)}), \quad (10)$$

without adding measurement noise. Summary statistics derived from these sample trajectories represent the remaining uncertainties of the model states $s(t, \theta)$. In BayModTS, $(1-\alpha) \cdot 100\%$ CI tubes of the model states are used to quantify the uncertainty. CIs are obtained for each time point by calculation of the percentile ranges $[z_{\frac{\alpha}{2}}, z_{1-\frac{\alpha}{2}}]$ from all sample trajectories $s^{(i)}(t)$ (Fig. 1, step 3). The point-wise CIs are linearly interpolated for visualization and can be interpreted as model-informed noise filters. Data points far outside those tubes might be seen as suspicious points that need further investigation or can be classified as outliers based on the model assumption. The procedure can be repeated if the dataset includes time courses from multiple conditions. CI tubes can be used to compare dynamics under different conditions (Fig. 1, step 4 and 1E).

Regarding FAIR principles, BayModTS is interoperable and reusable using the PETab format, which contains all information about the inference problem. Findability and accessibility are ensured by uploading the analysis results to an open public repository and assigning a digital object identifier (DOI). We provide a PETab file containing the RTF SBML model and executable code

for the Bayesian analysis on GitHub (<https://github.com/Systems-Theory-in-Systems-Biology/BayModTS>). Both can be easily adapted for either more complex models, different conditions, and individual data. The posterior is sampled via algorithms of the PyPesto toolbox (Schälte et al. 2023). PyPesto is well received and maintained by the systems biology model community and includes, besides others, an adaptive Metropolis-Hastings algorithm, a parallel tempering algorithm, and a wrapper to the ensemble sampler emcee. We recommend examining the sampling convergence via measures such as the effective sample size (ESS), visually via traces, or via the Gelman–Rubin diagnostic for multi-chain chains.

3 Application examples

Biological context: Hepatocellular carcinoma (HCC) is among the leading causes of cancer-related deaths worldwide and has the fastest-growing number of cancer-related deaths in the United States (El-Serag and Rudolph 2007). Still, Africa and Asia are showing the highest incidences of HCC as Hepatitis B and C are most prevalent in these continents and the leading cause of chronic liver diseases and HCC (Ferlay et al. 2015). Based on the evidence, liver resection is the most promising treatment option for HCC, but randomized phase III trials are ongoing (Vogel et al. 2018, Elderkin et al. 2023). However, liver resections are complex and pose the risk of major complications (Abreu et al. 2020). Portal vein embolization (PVE), the clinical pendant to PVL in animal models, is often performed to induce growth of the future liver remnant. In PVE, the portal influx of diseased liver lobes is ligated, leading to hypoperfusion and undersupply of nutrients in these lobes. In contrast, the nonligated lobes are hyperperfused, inducing regeneration and hypertrophy. However, the relationship between the liver mass gain of the nonligated liver lobes and liver function is still poorly understood (Christ et al. 2021). Moreover, comorbidities such as a high degree of hepatic steatosis might impair the proper functioning and recovery of the remaining liver volume after resection (Veteläinen et al. 2007).

3.1 Quantification of blood perfusion changes in the liver lobes of rats after PVL

This study explored *in vivo* liver perfusion in rats' ligated and nonligated liver lobes. In total, 25 rats were subjected to 60% PVL. The experimental procedure involved ligation of the left stem of the portal vein supplying the left median lobe (LML) and left lateral lobe (LLL) and the right stem supplying the right lobe (RL) (Fig. 2A). As a result, the remaining nonligated caudate lobe (CL) and right median lobe (RML) of the liver were hyperperfused. BayModTS was used to compare the perfusion dynamics of the different lobes and to calculate continuous perfusion changes over time with uncertainties. The RTF with an immediate response was used to describe perfusion changes over time.

Changes in local tissue perfusion over time were monitored and quantified using MRI in a cross-sectional study design. Perfusion was quantitatively assessed at the voxel level. A voxel is an MRI image pixel representing a defined volume's perfusion. The voxel has the same *xy*-axis size as the pixel, while the *z*-axis corresponds to the slice thickness. Voxels were assessed using arterial spin labelling (ASL) (Jahng et al. 2014), which is based on labelling spins of arterial blood by pulsed inversion and subtraction of the resulting image from

a control image acquired without labelling. Negative values resulting from this subtraction were excluded as they have no physiological relevance.

Regions of interest (ROIs) that are representative of the perfusion were defined for each liver lobe. ROI selection is shown exemplary for the RML (Fig. 2B). The increased perfusion observed on postoperative day one (POD1) in the RML (Fig. 2B second picture) is visually recognizable by the increasing yellow to red colouration in the annotated region. Voxel-level perfusion values within the ROI of each lobe are shown in Fig. 2C for the different observation times. The observed variability in individual voxel values is large due to blood flow and pulsation in larger vessels. Further reasons for the high variability are that different ROIs differ in the number and size of vessels they contain and the biological variation between animals. This variability makes it difficult to identify a consistent trend over time by visually inspecting these raw data.

The current standard processing of MRI ROI data is averaging the individual voxel data per liver lobe in the animals and visualizing it as box plots (Fig. 2D). Inspection of the averaged data shows that the nonligated lobes (i.e. RML and CL) were transiently hyperperfused on POD1 and POD2, whereas perfusion in the ligated lobes decreased after POD1.

BayModTS processing of the individual voxel data results in time-continuous and smooth CI tubes (Fig. 2E), considerably reducing the variability, and revealing that the highest uncertainty is between POD0 and POD1. The CI tubes provide a credible and clear visual distinction between the dynamics of the hyper- and the hypoperfused lobes. For all three ligated lobes (LML, LLL, RL), the average perfusion values decrease by about 30%–50% within one day and remain at low values. At the same time, the nonligated lobes (RML, CL) are hyperperfused, which is most pronounced at POD1. The median perfusion of CL is already back to pre-PVL levels by POD2, which occurs more slowly in the RML. Overall, the BayModTS analysis shows that the liver adapted to ligation-induced perfusion changes within two days.

The LML showed the greatest variability in perfusion compared to the other liver lobes. Reasons for the increased uncertainty of the LML are its proximity to the heart, causing motion that affects measurements, and the putative development of collaterals leading to alternative portal inflow. In addition, the heart, liver and lung tissues are not equally susceptible to the magnetic field, distorting the MRI images. As a result, slice positions close to the lungs and heart, such as the LML, usually have poor image quality. Consequently, cardiac motion and B0 inhomogeneities reduce the usable ROI size, increasing the uncertainty reflected in the CI tubes' size. Consistent with this, the 95% CI of the noise parameter σ (Fig. 2F) is also the largest for LML, reflecting higher measurement noise.

Conditions can also be compared by looking at the marginal posterior distributions of the model parameters. This is exemplified by the marginal posterior distributions of the parameter p_0 , which describes the preoperative perfusion in each of the lobes (Fig. 2G). Here again, the LML p_0 marginal distribution is the broadest, consistent with an overall increased uncertainty of the CI tubes. The preoperative perfusion of the RL is credibly lower than that of all other lobes (the upper limit of the RL CI is below the lower limit of all other CIs). Here, the relatively low RL perfusion is likely an

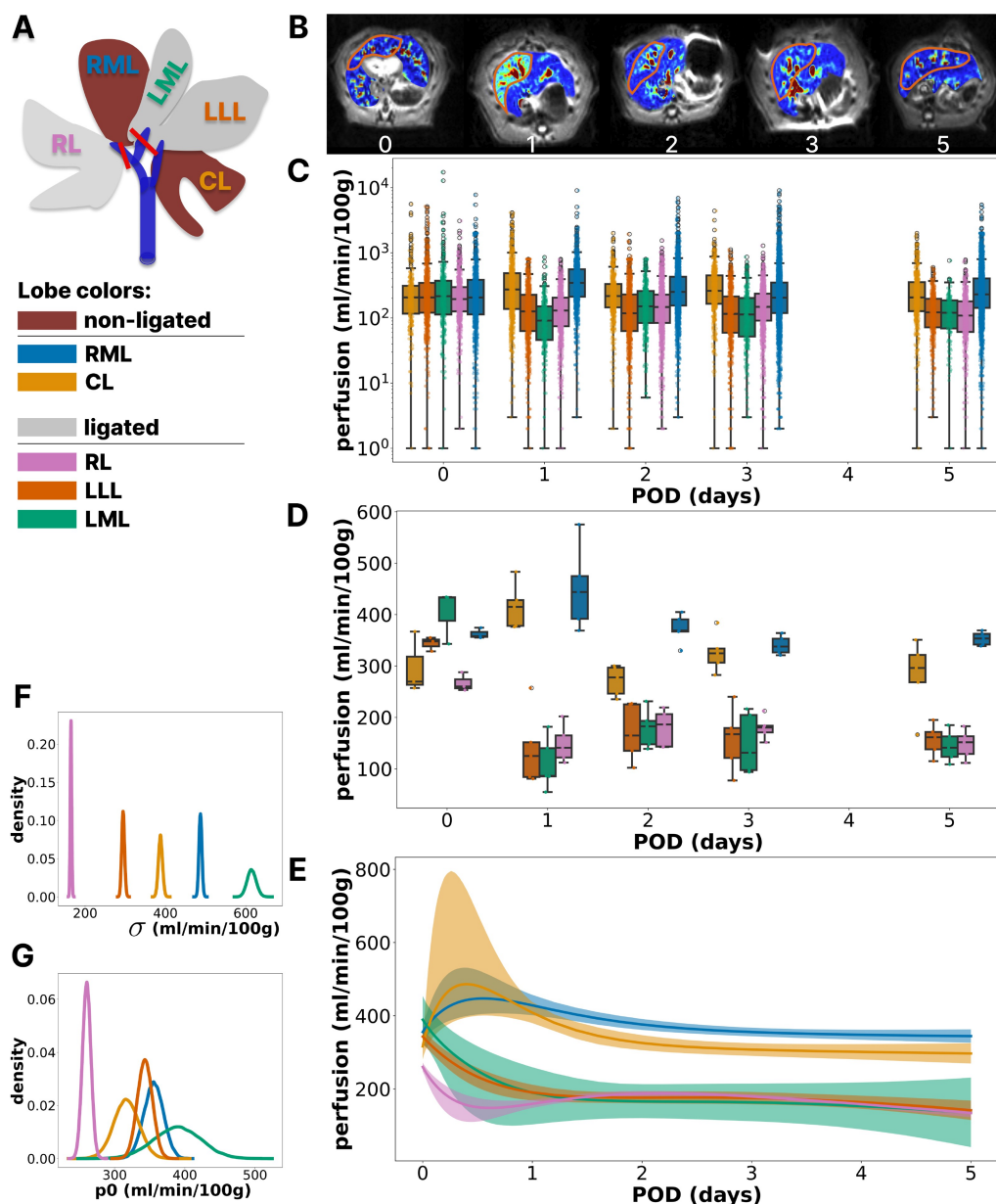


Fig. 2. BayModTS analysis calculates perfusion courses in ligated and nonligated liver lobes from MRI data after PVL. (A) Illustration of ligated (hypoperfused) and nonligated (hyperperfused) lobes of a rat liver subjected to PVL. (B) Exemplary MRI images for 0, 1, 2, 3, and 5 POD. The outlined area (orange) refers to the ROI annotation for RML. Perfusion is colour-coded, with yellow to red values indicating high values. (C) Boxplots of individual voxel values in the selected ROIs of $n = 5$ animals per time point (raw data). (D) State-of-the-art visualization of MRI ROI data. Boxplots are based on per-animal averaged ROI voxel data. (E) BayModTS ensemble perfusion predictions with median (solid lines) and 95% CI tubes for all lobes. (F, G) Marginal distributions of the parameters σ and p_0 per lobe, restricted to the 95% highest density interval.

artefact of having fewer vessels in the typical RL slice, which disappears in the postoperative measurements.

In conclusion, BayModTS provided a continuous and credible prediction of postoperative liver perfusion in the different lobes after PVL. The BayModTS inferred CI tubes are credibly different between ligated and nonligated lobes for all postoperative days. The credibility bounds allowed quantitative hypotheses about the timing of hyperperfusion and the time scales of adaptation. Furthermore, the continuity of the BayModTS results allows the PVL data to be used as input for physiological models. In the future, this information could be valuable for determining the correct timing of surgery in human patients without the need for repeated MRI scans.

3.2 Influence of steatosis on drug metabolism dynamics in mice

Hepatic steatosis is a common liver disease affecting up to 25% of the population in the Western world (Younossi *et al.* 2023). However, the effects of hepatic steatosis on drug metabolism remain poorly understood. Here, we apply BayModTS to investigate statistically if different steatosis degrees in mice change the metabolism dynamics of test drugs.

The study by Albadry *et al.* (2022) aimed to investigate the influence of different degrees of hepatic steatosis on the pharmacokinetics of several test drugs. Mice were fed a high-fat (HF) methionine-choline deficient (MCD) diet for 2 or 4 weeks, resulting in micro- and macrosteatosis (Fig. 3A).

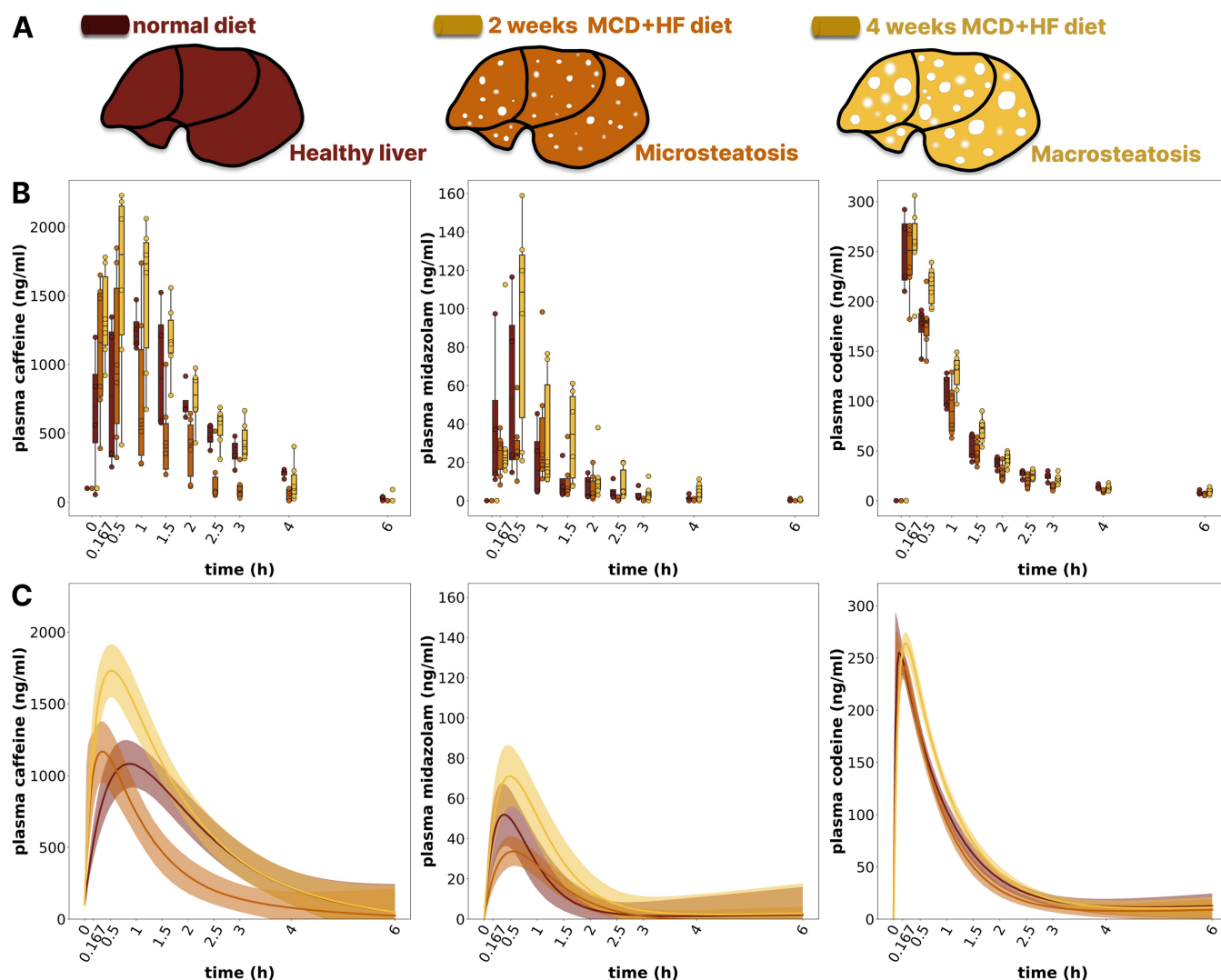


Fig. 3. Effect of severity and pattern of periportal steatosis on pharmacokinetics. (A) Mice were fed a normal diet for 2 weeks and a MCD and HF diet for 2 or 4 weeks. The MCD + HF diet induces hepatic microsteatosis after 2 weeks and macrosteatosis after 4 weeks. (B) Boxplots of plasma drug elimination time courses. The control group (red) consisted of $n=4$ animals, while the 2 weeks (orange) and 4 weeks (yellow) groups consisted of $n=6$ animals each. (C) Median ensemble prediction (dark lines) and 95% CI tubes of the BayModTS analysis with RTFs.

A drug cocktail consisting of codeine, caffeine and midazolam was administered, and whole blood samples were analysed to determine drug concentrations. Peak concentrations for all three test drugs occurred within 15–60 min, followed by elimination of the drug within 4–6 h. The time course measurements showed high inter-individual variability (Fig. 3B), with time-varying variances within a condition, which made it challenging to compare conditions over the entire period.

We used BayModTS to assess the elimination dynamics of all substrates as a function of steatosis degree. In the original study (Albadry *et al.* 2022), this was accomplished by assuming an exponential decay kinetics that mimics drug degradation, while we use the RTF with immediate response (Equation 1) here. The advantages of the RTFs compared to the exponential decay kinetics are thereby (i) a more flexible model approach that also allows dynamics to a different steady-state than the initial condition and (ii) a direct evaluation without numerical integration, remarkably improving the computation time. The uncertainty in the BayModTS inferred model states is considerably smaller than the

variability of the data (Fig. 3C). Our analysis suggests that macrovesicular periportal steatosis (4 weeks) delayed the clearance of all test drugs. Furthermore, peak concentrations of caffeine and midazolam were higher in macrovesicular steatosis. To validate the RTF dynamics in this example, we created a mechanistic two-compartment pharmacokinetic model for drug metabolism and repeated the BayModTS analysis with this model. The predicted medians and CI tubes for the RTF and PK approaches were similar, validating the RTF dynamics for this example (Supplementary Fig. S2).

Albadry *et al.* (2022) tested the differences of the steatosis conditions by calculating the area under the curve for each time series replicate. The four control, six 2 weeks MCD + HF, and six 4 weeks MCD + HF diet area under the curve values were analysed by a one-way ANOVA. The ANOVA calculated significant differences between the 2 and 4 weeks conditions for caffeine, midazolam and codeine. However, no significant difference between the control and the 4 weeks condition was detected (Albadry *et al.* 2022). ANOVA calculation on four to six data points only is critical. BayModTS improves the statistical analysis substantially as

(i) using the area under the curve as summary statistics can disguise differences in the metabolism dynamics if different effects compensate for each other over time, (ii) BayModTS takes all data points of the replicates into account, and (iii) BayModTS provides a visual comparison of the dynamics over time.

In summary, BayModTS helped to understand the data regarding differences in pharmacokinetics and indicated an impaired metabolism in mice with macrosteatosis.

4 Discussion

We have introduced BayModTS, a novel workflow for processing time series data using a FAIR Bayesian analysis. BayModTS can be applied in scenarios with sparse and noisy data. The processing is based on a simulation model that can mimic the measured time series. The statistical Bayesian framework enables a consistent transfer of data variability to uncertainties in model parameters, thereby acting as a noise filter and reducing the impact of potential outliers. BayModTS can process data in areas where purely data-driven approaches fail. The BayModTS-derived dynamics can be used to compare conditions or as time-continuous input equipped with uncertainty for computational models. Our examples show broad applicability in different contexts and for different kinds of data. Importantly, BayModTS contributes to reproducibility according to the FAIR principles by using established tools of the systems biology community and adhering to standard reporting guidelines (Tiwarei *et al.* 2021).

The RTFs are a flexible and computationally efficient simulation model which does not require knowledge of the underlying process. However, RTFs have two major limitations: (i) only one activation peak, and (ii) no oscillations can be modelled (Kreutz 2020). Further process knowledge is not incorporated in the model, which can lead to larger uncertainties than in mechanistic models. We showed that RTFs describe the underlying dynamics of drug metabolism in steatotic mice as well as a mechanistic pharmacokinetic model. We further demonstrate the flexibility of BayModTS with respect to user-defined SBML models by exchanging the RTF with a pharmacokinetic model incorporating domain-specific knowledge.

Properly describing observed data distributions in connection with deterministic simulation models is tricky in Bayesian settings. Forward simulations with posterior samples can filter noise because the model assumption restricts the uncertainty. Noise filtering works particularly well if the simulation model is a good description of the underlying process or is flexible enough to adapt to the course of the data. Part of the noise filtering property of BayModTS is based on the standard Likelihood (Equation 5) we use, which penalizes distances quadratically. Therefore, the uncertainty of the CI tubes of the application examples was smaller than the uncertainty in the data. In datasets having extreme values to higher and lower values simultaneously, penalizing the quadratic distance leads to regression towards the centre. The regression towards the centre is because the smaller distance to the one outlier cannot compensate for a disproportionately large distance to the other outlier. In cases where the measurement methods generate a lot of noise or a good estimator for the average population behaviour is to be found, this filter property can be beneficial. In BayModTS, the derived CI tubes are influenced by the choice of the error model, its hyperparameters and the prior. The CI tubes are currently calculated based on percentile intervals of the posterior

parameter ensemble predictions. The percentile intervals cut the lower and upper $\alpha/2\%$ of the ensemble predictions per evaluation time point, and upper and lower interval bounds are connected to form the credibility tube. Depending on the process and shape of the posterior distribution, percentile ensemble-based CIs can lead to broader or smaller credibility tubes compared to CI calculation via the highest posterior density. In the long run, calculating the credibility tubes on the posterior level would be desirable.

If only limited information about the parameters is known a-priori, noninformative priors can be used. In practice, sampling procedures and optimization algorithms usually require boundaries, which must be adjusted manually. These bounds should not restrict the sampling, which can be difficult with correlated or sloppy parameters. Model reduction techniques can be used here to improve the identifiability of the simulation model parameters (Eisenkolb *et al.* 2020).

Extensions of the workflow are easily possible for future applications. First, the method can also be applied to nontime series data. If there is only one dependent variable and serial data with known underlying dynamics, BayModTS can quantify the uncertainty and predict a credible range of the dynamics. To use serial nontime data, the user must define another dependent variable in the SBML file and include the corresponding data in the PÉtab measurement table. Second, BayModTS can be applied to user-defined SBML models as demonstrated with the pharmacokinetic model for the steatosis data. Third, the BayModTS analysis allows a time-continuous description of the experimental data with uncertainty. In the future, these time-continuous functions can be used as input modules for other computational models. For example, BayModTS-derived perfusion changes after PVL could be used as input for other liver models, thereby relating liver perfusion to liver function. Fourth, conditions can be compared on the parameter level. Some parameters might differ between conditions on the parameter level, others may not. This is different when looking at the posterior parameter ensemble prediction, which shows differences in the joint parameter level. Fifth, BayModTS can also compare different processes rather than different conditions. Here, different SBML models describing the underlying processes are used for comparison.

Acknowledgements

We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). Further, we acknowledge the support by the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the University of Tübingen, the state of Baden-Württemberg through bwHPC and the DFG through grant no INST 37/935-1 FUGG. We also acknowledge the use of Grammarly for language editing.

Author contributions

S.H. and N.E.R. developed the methodology. S.H. conducted the Bayesian analyses in close consultation with N.E.R. M.K. supported FAIR principles and SBML modelling. K.-H.H., W.-T.Z., J.R.R., U.D., M.A., H.-M.T., and E.M.K. collected and curated the data used in this study and helped S.H., and N.E.R. to interpret the BayModTS results in a biomedical context. W.W. operated the animals of the PVL MRI measurements. All authors wrote and reviewed the manuscript.

Supplementary data

Supplementary data are available at *Bioinformatics* online.

Conflict of interest

None declared.

Funding

This work was supported by the German Research Foundation (DFG) within the Research Unit Program FOR 5151 ‘QuaLiPerF (Quantifying Liver Perfusion-Function Relationship in Complex Resection—A Systems Medicine Approach)’ [436883643 to S.H., M.A., U.D., K.-H.H., M.K., J.R.R., H.-M.T., W.-T.Z., N.E.R.]. S.H. and N.E.R. further acknowledge support by the DFG under Germany’s Excellence Strategy—EXC 2075 – 390740016. H.-M.T., E. M.K., M.K., and U.D. were supported by the DFG [465194077] (Priority Programme SPP 2311, Subproject SimLivA). E.M.K. was supported by the doctoral scholarship ‘Jena School for Ageing Medicine’ (JSAM) of the Else-Kröner-Fresenius Foundation. M.A. and U.D. were supported by DFG SteaPKMod [410848700]. M.K. and H.-M.T. were supported by the Federal Ministry of Education and Research (BMBF, Germany) within ATLAS [031L0304B]. W.W. was supported by the Interdisciplinary Center for Clinical Research (IZKF) of Jena University Hospital (Clinician Scientist Programm CSP 12).

References

- Abreu P, Ferreira R, Bussyguin DS *et al.* Liver resections for metastasis: surgical outcomes of a single center academic institution. *BMC Surg* 2020;20:254.
- Albadry M, Höpfel S, Ehteshamzad N *et al.* Periportal steatosis in mice affects distinct parameters of pericentral drug metabolism. *Sci Rep* 2022;12:21825.
- Campagnoli P, Petrone S, Petris G (eds.). *Dynamic Linear Models with R*. New York, NY: Springer, 2009.
- Chen C, Lavezzi SM, Iavarone L *et al.* The area under the effect curve as an efficacy determinant for anti-infectives. *CPT Pharmacometrics Syst Pharmacol* 2022;11:1029–44.
- Christ B, Collatz M, Dahmen U *et al.* Hepatectomy-induced alterations in hepatic perfusion and function – toward multi-scale computational modeling for a better prediction of post-hepatectomy liver function. *Front Physiol* 2021;12:733868.
- Eisenkolb I, Jensch A, Eisenkolb K *et al.* Modeling of biocatalytic reactions: a workflow for model calibration, selection, and validation using bayesian statistics. *AICbE J* 2020;66:1–13.
- El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007;132:2557–76.
- Elderkin J, Al Hallak N, Azmi AS *et al.* Hepatocellular carcinoma: surveillance, diagnosis, evaluation and management. *Cancers (Basel)* 2023;15:5118.
- Ferlay J, Soerjomataram I, Dikshit R *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.
- Hawkins DM. *Identification of Outliers*. Dordrecht: Springer Netherlands, 1980.
- Huang H, Zhang E-B, Yi O-Y *et al.* Sex-related differences in safety profiles, pharmacokinetics and tissue distribution of sinomenine hydrochloride in rats. *Arch Toxicol* 2022;96:3245–55.
- Jahng G-H, Li K-L, Ostergaard L *et al.* Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques. *Korean J Radiol* 2014;15:554–77.
- Keating SM, Waltemath D, König M *et al.*; SBML Level 3 Community Members. SBML level 3: an extensible format for the exchange and reuse of biological models. *Mol Syst Biol* 2020;16:e9110.
- Ko J, Yoo C, Xing D *et al.* Pharmacokinetic analyses of liposomal and non-liposomal multivitamin/mineral formulations. *Nutrients* 2023; 15:3073.
- Korpela J, Puolamäki K, Gionis A *et al.* Confidence bands for time series data. *Data Min Knowl Disc* 2014;28:1530–53.
- Kreutz C. A new approximation approach for transient differential equation models. *Front Phys* 2020;8:70.
- Kreutz C, Bartolome Rodriguez MM, Maiwald T *et al.* An error model for protein quantification. *Bioinformatics* 2007;23:2747–53.
- Schälte Y, Fröhlich F, Jost PJ *et al.* pyPESTO: a modular and scalable tool for parameter estimation for dynamic models. *Bioinformatics* 2023;39:foae011.
- Scheff JD, Almon RR, Dubois DC *et al.* Assessment of pharmacologic area under the curve when baselines are variable. *Pharm Res* 2011; 28:1081–9.
- Schmiester L, Schälte Y, Bergmann FT *et al.* PETab-Interoperable specification of parameter estimation problems in systems biology. *PLoS Comput Biol* 2021;17:e1008646.
- Tapinos A, Mendes P. A method for comparing multivariate time series with different dimensions. *PLoS One* 2013;8:e54201.
- Thomaseth C, Radde N. The effect of normalisation and error model choice on the distribution of the maximum likelihood estimator for a biochemical reaction. *IET Syst Biol* 2023;17:1–13.
- Tiwari K, Kananathan S, Roberts MG *et al.* Reproducibility in systems biology modelling. *Mol Syst Biol* 2021;17:e9982.
- Veteläinen R *et al.* Steatosis as a risk factor in liver surgery. *Ann Surg* 2007;245:20–30.
- Vogel A, Cervantes A, Chau I *et al.*; ESMO Guidelines Committee. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol Off J Eur Soc Med Oncol* 2018;29:iv238–55.
- Younossi ZM, Golabi P, Paik JM *et al.* The global epidemiology of non-alcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. *Hepatology* 2023;77:1335–47.