**Proceedings Paper:**
Mclean, J.H., Dervilis, N. and Rogers, T.J. (2024) Convolution models for output only linear structural system identification and the problem of identifiability. In: Journal of Physics: Conference Series. XII International Conference on Structural Dynamics, 03-05 Jul 2023, Delft, Netherlands. IOP Publishing. Article no: 192023. ISSN: 1742-6588. EISSN: 1742-6596.

https://doi.org/10.1088/1742-6596/2647/19/192023

Journal of Physics: Conference Series



# Convolution models for output only linear structural system identification and the problem of identifiability.

To cite this article: J.H. Mclean *et al* 2024 *J. Phys.: Conf. Ser.* **2647** 192023

# Convolution models for output only linear structural system identification and the problem of identifiability.

**J.H. Mclean, N Dervilis, T.J. Rogers**

Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

E-mail: `jmclean1@sheffield.ac.uk`

**Abstract.** This paper investigates the use of the Gaussian Process Convolution Model (GPCM) as an output only system identification tool for structural systems. The form of the model assumes a priori that the observed data arise as the result of a convolution between an unknown linear filter and an unobserved white noise process, where each of these are modelled as a GP. The GPCM infers both the linear time filter (which is the impulse response function, i.e. Green's function, of the system) and driving white noise process in a Bayesian probabilistic fashion with an approximate variational posterior over both signals. It will be shown that although the model structure is intuitive and sensible priors are applied, the GPCM falls short in recovering the linear impulse response of interest response due to the problem of identifiability. This is an interesting result indicating that physically informed kernel structures alone are not enough to recover the true impulse response in similar non-parametric probabilistic models. Despite this, the avenue of research remains highly promising, and several ideas are proposed to improve the model as a system identification tool.

## 1. Introduction

System identification on many operational structures can be made difficult by the lack of information regarding the structure and the input force. Output-only system identification methods exist as a way of inferring the system and (in some cases) input force, given the output of the system which can be readily measured. Within the context of identifying structural systems, the process of determining the properties of a system from output-only data falls under the banner of Operational Modal Analysis (OMA).

The lack of information available leads to the need for several some complicated choices in OMA, ranging from assumed excitation signals to various model choices and structures. There are however, several ways in which the impact of these choices can be mitigated against. First amongst these strategies is leveraging a probabilistic approach; the ability to recover uncertainty for values of interest is highly valuable in a field where these values were calculated based on uncertain assumptions. The increased richness of information can also be propagated forward into further dynamic analysis which can also capitalise on this uncertainty.

The second mitigating method is encoding physics into the models being used. This is beneficial because it can help make black box models more interpretable by grounding them in physical principles, as well as allowing for the extraction of physical properties of interest (which is a key goal in OMA).

These mitigating methods tie in nicely with the Gaussian Process Convolution Model (GPCM) [1]. The GPCM leverages the structure of the convolution of linear systems, and places a probabilistic prior on all stages of the convolution (impulse response, excitation and output). This allows for the excitation and impulse response to be inferred from output only data.

The GPCM was created as a method of learning the kernel function representing the output function, it does this by leveraging the structure of linear systems, but is not explicitly a system identification tool. In section 2, a brief introduction to the necessary theory of the GPCM will be provided. In section 3, it will be demonstrated that while each component of the GPCM can more effectively model its respective signal in isolation compared to non-dynamically informed counterparts, the GPCM falls short in recovering the impulse response of interest. Rather, it correctly identifies a filter function and excitation that correspond with the kernel of the output being learnt, but not necessarily the impulse response of the system at hand. Finally, in section 4 suggestions are put forth to enhance the performance of the GPCM as a tool for system identification.

### *1.1. Related Work*

Gaussian processes have been used in structural system identification due to the probabilistic and flexible nature of the models. A number of works utilise GPs in OMA in particular, such as [2] and [3]. In [4], a latent force GP is used to infer both the excitation and dynamic system, while in [5] a hierarchical GP is used in modal analysis.

Similar model structures to the GPCM exist in system identification literature, such as [6], which similarly places a GP over the impulse. However this method implements the GP naively and requires full excitation data. Other similar works include [7], which finds a physics based kernel to place over the impulse response and [8], which despite being a latent force model uses a convolutional structure similar to the GPCM often used in multi-output GPs.

For simplicity, this work introduces the GPCM as shown in the original publication. However, significant further work on the GPCM has also been conducted, including the causal and rough GPCM models as well as a variational posterior based on a Gibbs sampler in [9].

## 2. Theory

### *2.1. The Gaussian Process*

The Gaussian Process is a flexible non-parametric model designed to model regression problems of the form $y = f(\boldsymbol{x}) + \varepsilon$. Where $\boldsymbol{x}$ is some vector of inputs and $\boldsymbol{y}$ the target data with additive Gaussian noise $\varepsilon$ which has a zero mean and a variance of $\sigma_n^2$. The GP defines a prior distribution where any finite subset of the function has a multivariate Gaussian distribution. The GP is fully defined in Eq. 1, where $\mu(\boldsymbol{x})$ is the mean function and $k(\boldsymbol{x}, \boldsymbol{x}')$ is the covariance function (otherwise known as the kernel).

$$f \sim \mathcal{GP}(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \qquad \text{(Eq. 1)}$$

Kernel functions are off the shelf auto correlation functions that almost fully determine the properties of the generated functions. As such, the choice of kernel function allows for easy encoding of prior beliefs about the signal being modelled, and will significantly impact the results of the GP. For a more complete introduction to Gaussian Processes, the reader is referred to [10].

Journal of Physics: Conference Series **2647** (2024) 192023 doi:10.1088/1742-6596/2647/19/192023

### *2.2. Linear structural dynamics*

System identification aims to create a comprehensive model of a dynamic system, such that for any given input $\boldsymbol{x}$, the output $\boldsymbol{f}$ can be predicted. This comprehensive model is denoted $\boldsymbol{h}$, and is known as the impulse response.

For a linear system, the output $\boldsymbol{f}$ of a dynamic system is the product of the convolution between the structures impulse response $\boldsymbol{h}$ (also known as Green's function) and the excitation $\boldsymbol{x}$, as shown in Eq. 2. In OMA, it is a justifiable assumption (as long as the system is linear) that the output only data being analysed can be thought of as the product of this convolution.

$$f = \int_{\mathbb{R}} h(t-\tau)x(\tau)d\tau \qquad \text{(Eq. 2)}$$

It can be seen then that $\boldsymbol{h}$ completely characterises the system, such that if $\boldsymbol{h}$ is known, the output $\boldsymbol{f}$ for any given input $\boldsymbol{x}$ will also be known. Additionally, the dynamic properties of the structure being analysed can be extracted from the impulse response.

### *2.3. The Gaussian Process Convolution Model*

The GPCM proposes to model both the impulse response and the excitation as Gaussian Processes, shown in Eq. 3. The kernel $k_h$ is the squared exponential kernel pre and post multiplied by some decaying term in order to reflect the finite power and decaying nature of an impulse response, and the kernel for the corresponding white noise excitation is the delta-function covariance.

$$h \sim \mathcal{GP}(0, k_h(\boldsymbol{t}, \boldsymbol{t'})), \quad x \sim \mathcal{GP}(\mu(\boldsymbol{x}), \delta(\boldsymbol{t}, \boldsymbol{t'})) \qquad \text{(Eq. 3)}$$

$$k_h(t,t') = \sigma_h^2 e^{-\alpha t^2} e^{-\gamma(t-t')^2} e^{-\alpha t'^2} \qquad \text{(Eq. 4)}$$

If the output $\boldsymbol{f}$ is conditioned on the system $\boldsymbol{h}$, then $\boldsymbol{f}$ is a linear transformation of $\boldsymbol{x}$, and as such $f|h$ will be a Gaussian Process (given that $\boldsymbol{f}$ is now a combination of Gaussian variables), defined in Eq. 5. Where the covariance is defined as $k_{f|h}(t,t') = (h(t) * h(-t))(t)$, or the convolution of the filter and its mirror. The appeal of the GPCM outside of system identification is that it infers the kernel $k_{f|h}$ instead of requiring manual selection for a given signal, resulting in potential improvements in accuracy, time savings, better generalization, and reduced reliance on user expertise.

$$f|h \sim \mathcal{GP}(0, k_{f|h}(\boldsymbol{t}, \boldsymbol{t'})) \qquad \text{(Eq. 5)}$$

Since the GPCM only has access to the output of the system, the excitation and impulse response are not observable. This means there is no data with which to model a GP over the impulse or excitation. Instead, both the impulse and excitation are approximated through a variational posterior $q(u_x, u_h)$, where $\boldsymbol{u}_h$ and $\boldsymbol{u}_x$ are the inducing points of the filter and excitation to be learnt. This variational posterior needs to be optimised, and this is done so through the ELBO of the GP on the output, the form of which is shown in Eq. 6. The ELBO is then optimised through the users choice of gradient descent. The reader is refered to [1] and [9] for a comprehensive introduction to the GPCM.

$$\mathcal{F} = \mathbb{E}_q[\log p(y|h,x)] - \mathbb{KL}[q(\mathbb{u}_h)||p(\mathbb{u}_h)] - \mathbb{KL}[q(\mathbb{u}_x)||p(\mathbb{u}_x)] \qquad \text{(Eq. 6)}$$

$$\mathbb{E}_q[\log p(\mathrm{y}|h,x)] = -\frac{N}{2}\log(2\pi\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\mathbb{E}_q\left[\left(y(t_i) - \int_{\mathbb{R}} h(t_i-\tau)x(\tau)d\tau\right)^2\right] \qquad \text{(Eq. 7)}$$

$$\mathbb{E}_q\left[\int_{\mathbb{R}} h(t_i-\tau)x(\tau)d\tau\right] = \int_{\mathbb{R}} \mathbb{E}_{q(h,\mathbb{u}_h)}[h(t_i-\tau]\mathbb{E}_{q(x,\mathbb{u}_x)}[x(\tau)]d\tau \qquad \text{(Eq. 8)}$$

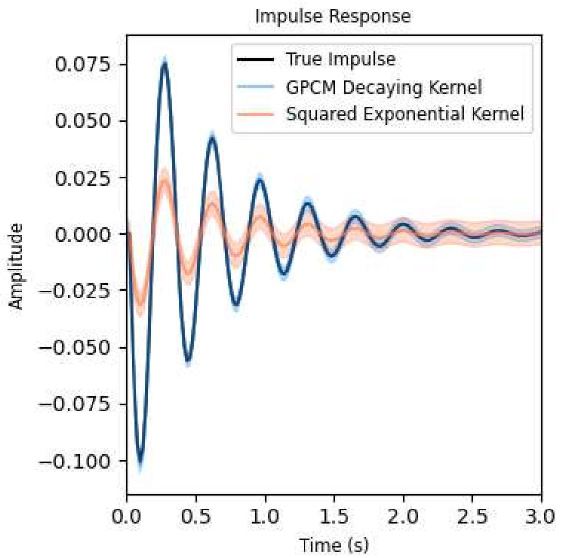
Journal of Physics: Conference Series **2647** (2024) 192023 doi:10.1088/1742-6596/2647/19/192023

## 3. Case study of GPCM on a synthetic linear system

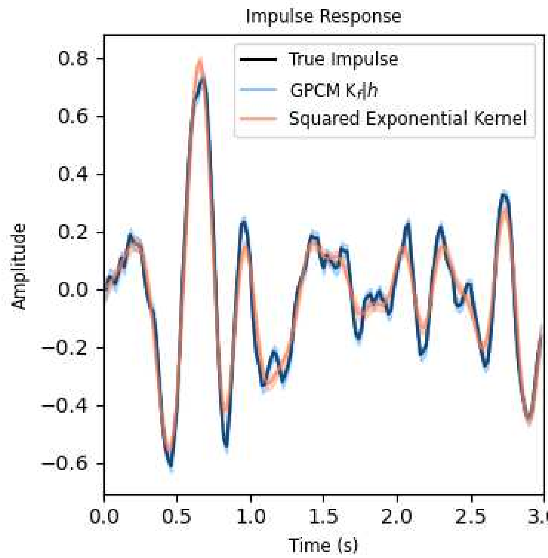
The draw for the GPCM in non-system identification settings is that it can recover the kernel $k_{f|h}$ of a given signal. However, there exist an unknown number of combinations of filter functions and excitations that can result in $k_{f|h}$. This means that the GPCM as outlined in the previous section will return some filter function, but it is unlikely this filter function will be the impulse response of the system being modelled.

What is interesting about this result is that the structure of the GPCM is both intuitive and effective, which indicates that no matter how realistic the probabilistic prior placed on the impulse response may be, it cannot be recovered. In order to demonstrate the structure of the GPCM, various stages of it are applied in isolation and compared to non-dynamically informed counterparts.

Figure 1a shows a known impulse response with arbitrary constants of a Single Degree of Freedom (SDOF) system. Two GPs model the impulse response, the first with an identical kernel to the GPCMs filter function, and the second a standard squared exponential kernel. It is clear the kernel for the GPCMs filter function can capture the impulse response almost perfectly. In contrast, the squared exponential kernel struggles to capture the decaying signal, providing a compelling illustration of the benefits that can be drawn from incorporating physical principles into models.




(a) Modelling a known impulse response

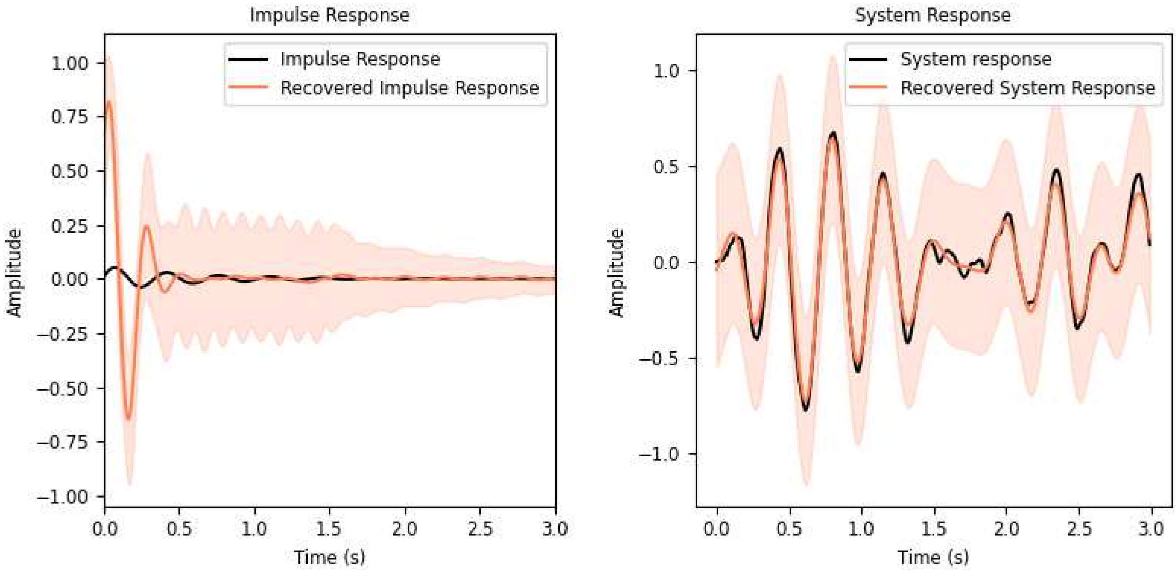



(b) Modelling the response of a known system

Figure 1: Comparison of GPCM vs normal GP on known system and response of system to random excitation

Figure 1b shows the response of the system in Figure 1a to some random noise. Again two GPs are used to model the response. The first is now a GP with the kernel $k_{f|h}$, where $h$ is the impulse response in Figure 1a and (in this case) known. The second is again a standard GP with a squared exponential kernel. It is again shown that the dynamically informed GP performs better.

As mentioned previously, the GPCM struggles to recover the impulse response of interest from output-only data. This is shown in Figure 2, where the full GPCM is trained on the response of same the system and excitation used for the results in Figure 1. It is clear from Figure 2 that while the GPCM does predict the output response (suggesting that it has correctly recovered the kernel $k_{f|h}$) it has not recovered a sensible impulse response corresponding to the system at hand (as seen in Figure 2a).




(a) Recovered vs true impulse response

(b) Recovered vs true system response to random excitation

Figure 2: GPCM modelling an SDOF system with random excitation

## 4. Concluding remarks and avenues for future research

This paper investigated the use of the GPCM in operational system identification. It was found that despite intuitive and effective kernel structures, the model was not able to meaningfully recover the true impulse response being investigated. This is a meaningful result because it shows that a strong physically informed model structure is not enough to adequately define the search space of possible filter functions.

As a result, it is likely that the best results for future work in integrating GPCMs into the system identification toolbox will come as a result of integrating some excitation data into the GPCM. It is shown in the supplemental material of [9] that the variational posterior of the impulse response is fully recoverable if the variational posterior of the excitation is known (and vice versa). The requirement of the full excitation data is highly undesirable in the context of system identification, however interesting work can be done in leveraging the non-parametric prior placed on the excitation in order to reduce the amount of excitation data required to recover the true impulse response. This is an interesting result for system identification on operational structures, where it can be possible to record the excitation over a very short time frame.

## Acknowledgments

The authors gratefully acknowledge the support of the Engineering Physical Sciences Research Council (EPSRC) through grant numbers EP/S001565/1 and EP/W002140/1.

## References

[1] Tobar F, Bui T D and Turner R E 2015 *Advances in Neural Information Processing Systems* **28**

[2] Zhu Y C and Au S K 2020 *Mechanical Systems and Signal Processing* **136** 106511
[3] Zeng J and Hu Z 2022 *Engineering Structures* **273** 115139
[4] Rogers T J, Worden K and Cross E J 2020 *Mechanical Systems and Signal Processing*
[5] Sedehi O, Katafygiotis L S and Papadimitriou C 2020 *Mechanical Systems and Signal Processing* **140** 106663
[6] Pillonetto G and De Nicolao G 2010 *Automatica* **46** 81–93
[7] Chen T 2018 *Automatica* **90** 109–122
[8] Alvarez M A, Luengo D and Lawrence N D 2013 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35** 2693–2705
[9] Bruinsma W P, Tegnér M and Turner R E 2022 *International Conference on Artificial Intelligence and Statistics* (PMLR) pp 5166–5195
[10] Rasmussen C E 2004 *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures* (Springer) pp 63–71