Identifiability of pharmacological models for online individualization

Wahlquist, Ylva; Gojak, Amina; Soltesz, Kristian

Published in:
IFAC-PapersOnLine

DOI:
10.1016/j.ifacol.2021.10.226

2021

Document Version:
Peer reviewed version (aka post-print)

Identifiability of pharmacological models for online individualization

Ylva Wahlquist * **Amina Gojak** * **Kristian Soltesz** *

* *Lund University, Dept. Automatic Control, Sweden*
(e-mail: ylva.wahlquist@control.lth.se)

Abstract: There is a large variability between individuals in the response to anesthetic drugs, that seriously limits the achievable performance of closed-loop controlled drug dosing. Full individualization of patient models based on early clinical response data has been suggested as a means to improve performance with maintained robustness (safety). We use estimation theoretic analysis and realization theory to characterize practical identifiability of the standard pharmacological model structure from anesthetic induction phase data and conclude that such approaches are not practically feasible.

Keywords: Individualized anesthesia, Drug delivery, Identifiability, Pharmacokinetic/ pharmacodynamic modeling, Sensitivity analysis

1. INTRODUCTION

In general anesthesia, hypnotic depth—reciprocal to the level of awareness—is controlled through the addition of anesthetic drugs distributed across tissues and eventually metabolized. This paper considers the prospect of automatically controlling the hypnotic depth through real-time electroencephalogram (EEG) measurements, from which the depth can be estimated, and intravenous infusion of the anesthetic drug propofol. This setting is well-studied in the literature and has been the subject of a handful of clinical studies. Without attempting a comprehensive survey, we suggest the survey Ghita et al. (2000) as a starting point for the interested reader.

As with many drugs, there is a large inter-patient variability in the response to propofol. It can partly be explained using mixed-effect population models as described in Eleveld et al. (2018). Even so, the remaining variability remains large, and importantly limits the performance of a safe (robust) feedback controller, as we have illustrated in for example Gonzales-Cava et al. (2020).

Consequently, several simulation studies have suggested to individualize the controller, based on data from the induction phase of anesthesia, comprising of the transition from the fully aware state to a desired hypnotic depth. In Soltesz et al. (2013) we obtained promising results when considering an idealized simulation case. Here, we return to the same idea and investigate how including a measurement noise model changes the prospect of individualized therapy based induction phase data. Particularly, we investigate the practical identifiability of the standard pharmacological model used to describe the effect of propofol on the hypnotic depth from induction phase time series. This is what would be required in a clinical setting since the point of the individualization is to have the model available for updating the drug dosing controller as early as possible during the ongoing treatment.

2. DYNAMICAL MODEL

2.1. Pharmacological model

The relation between propofol infusion u [mass/time] and measured clinical effect y is commonly modeled by a pharmacokinetic (PK) model relating u to the blood plasma concentration C_p [mass/volume], a pharmacodynamic (PD) model relating C_p to the (normalized) clinical effect $\bar{y}$, and finally an observation model relating $\bar{y}$ to the measured clinical effect y.

The combined PK and PD (PKPD) model comprises a linear and time invariant (LTI) mammillary three-compartment model with C_p as output, connected in series with a lowpass filter, that can be thought of as an effect site compartment of infinitesimal volume. The output of the series connection is the effect site concentration C_e [mass/volume]. Here we will normalize C_e by EC_{50} [mass/volume], being the C_e value corresponding to the midpoint between no and full clinical effect. We thus introduce $z = C_e/EC_{50}$. The combined LTI parts of the PKPD models can be expressed on state space form as

$$\dot{x} = \underbrace{\begin{bmatrix} -(k_{10}+k_{12}+k_{13}) & k_{12} & k_{13} & 0 \\ k_{21} & -k_{21} & 0 & 0 \\ k_{31} & 0 & -k_{31} & 0 \\ k_{e0} & 0 & 0 & -k_{e0} \end{bmatrix}}_{A} x + \underbrace{\begin{bmatrix} \frac{1}{\eta} \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{B} u, \quad (1)$$

$$z = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}}_{C} x, \quad (2)$$

where $\eta = v_1 EC_{50}$ is the product of the central compartment volume v_1 [volume] and EC_{50}.

The transfer function of (1)–(2) from u to z is

$$G_{z,u}(s) = C(sI - A)^{-1}B = K\frac{(s+z_1)(s+z_2)}{(s+p_1)(s+p_2)(s+p_3)(s+k_{e0})}, \quad (3)$$

with static gain $K = k_{e0}/\eta$, zeros $-z_1 = -k_{21}, -z_2 = -k_{31}$, and where the poles $-p_1, -p_2, -p_3$ solve the characteristic equations

$$p_1 + p_2 + p_3 = k_{10} + k_{12} + k_{13} + k_{21} + k_{31} \quad (4a)$$

$$p_1p_2 + p_1p_3 + p_2p_3 = k_{10}k_{21} + k_{10}k_{31} + k_{12}k_{31} + k_{13}k_{21} + k_{21}k_{31} \quad (4b)$$

$$p_1p_2p_3 = k_{10}k_{21}k_{31}. \quad (4c)$$

The nonlinear output part of the PD model is characterized by a sigmoidal function h, relating $z = C_e/EC_{50}$ to the normalized clinical effect:

$$\bar{y} = h(z, \gamma) = 1 - \frac{1}{1 + z^\gamma}, \quad (5)$$

where $\bar{y} = 0$ and $\bar{y} = 1$ signify full awareness, and the maximal hypnotic depth, respectively.

The eight parameters k_{10}, k_{12}, k_{13}, k_{21}, k_{31}, k_{e0}, γ, η are assumed strictly positive, with $\gamma \geq 1$. The six rate parameters $k_{\cdot}$ [1/time] define the diffusion between the compartments; η has unit [mass]; γ is unitless and defines the shape of (5). A more in-depth explanation of the combined model (1)–(5) is provided in Copot (2019).

2.2. Actuation model

The input u is actuated using a computer-controlled infusion pump. For any reasonable syringe drug concentration quantization error in the magnitude of $u(t)$ is negligible, as is quantization in t introduced by zero-order-hold (ZOH) sampling [1] Furthermore, since the input is ZOH, ZOH sampling of (1) at T_s is approximation-free and used for all simulations herein.

2.3. Observation model

Clinicians are used to the awareness being reported using the bispectral index (BIS) or other awareness estimates reported on the same scale [2] , where

$$y^0 = E_0(1 - \bar{y}), \quad E_0 = 100. \quad (6)$$

We will here consider $E_0 = 100$ in the equilibrium state $x = 0$, but it could alternatively be considered a free parameter that accounts for some patients generating a monitor reading of e.g. 90 BIS (by setting $E_0 = 90$) in absence of drug, as observed and modeled in van Heusden et al. (2013).

The clinical effect measure y is an *estimate* of the clinical effect y^0, derived from EEG measurements using a monitoring device, with a ZOH output, updated at $T_s = 1$ s. The observation model employed here is that the monitor introduces additive independent and identically distributed (IID) Gaussian noise $w \sim \mathcal{N}(0, \sigma^2)$,

$$y = y^0 + w, \quad (7)$$

where $\sigma^2 = 9^2$ BIS was identified for the NeuroSense EEG monitor in Soltesz et al. (2012). Herein we use the more optimistic $\sigma^2 = 1$ BIS [3] .

3. SIMULATED EXAMPLE

To generate a representative virtual patient, we have used the population model of Eleveld et al. (2018) that, based on openly available data sets, defines (nominal) individual models on the form (1)–(5) as functions of patient demographics. Particularly for our investigation, we have used the reference patient model (35 year old male, 1.70 m, 70 kg) with parameter values θ^o in table 1.

The experiment is commenced in the stationary state $x = 0$ of (1), corresponding to the drug-free equilibrium. At time $t = 0$ a bolus dose (an impulse) $\alpha\delta(0)$ is delivered and a steady infusion β is commenced. The input is thus

$$u(t) = \begin{cases} 0, & t < 0, \\ \alpha\delta(0) + \beta, & t \geq 0 \end{cases} \quad (8)$$

The steady infusion was chosen to yield $y^0 = 50$ BIS $\Leftrightarrow \bar{y} = 1/2 \Leftrightarrow z = 1$ in stationarity. With system matrices defined through (1)–(2) this corresponds to solving $-CA^{-1}B\beta = 1$, resulting in $\beta = k_{10}\eta = 5.5$ mg/min.

In our simulations, the bolus was approximated by a high-rate infusion [4] and the time during which the bolus dose was delivered was chosen (through bisection search) to limit the response overshoot to $\bar{y}(t) \geq 0.6$ as motivated in Agrawal et al. (2010).

The input u and resulting effect y^0 are shown in Fig. 4 (black) together with the noise-corrupted observation y (gray).

4. LOCAL IDENTIFIABILITY

4.1. Practical identifiability at the true parameter

Unless pole-zero cancellations occurs (as further explained in Sec. 5.2) the system (1)–(2) is structurally identifiable from u, z. Since $h(z)$ of (5) is strictly monotonous in z it also holds that the dynamics are structurally identifiable from u, y^0.

However, the input u of Fig. 4 comprises merely two step changes. Furthermore, the experiment duration is less than one sixth of the slowest pole time constant of the true model (cf. table 1).

Increasing input excitation or prolonging the experiment would both be associated with ethical concerns related to patient safety. It is therefore natural to ask whether the parameters of (1)–(5) are *practically* identifiable from u, y.

[1] The resolution in t is typically $\approx$ 1 s and the resolution in u for a standard 10 mg/ml propofol solution is $\approx$ 0.02 mg/min, to be compared with the input magnitude and time scales of a representative therapeutic profile in Fig. 4.

[2] In the literature, the input matrix of (1) is often multiplied by EC_{50}, while (5) and (6) are combined into the standard Hill function form $y^0 = 100(E_0 - C_e^\gamma/(EC_{50}^\gamma + C_e^\gamma))$. The input-output behavior becomes equivalent, but the normalization of the internal signals z and $h(z)$ is lost.

[3] The aforementioned discrepancy between system time constants and sampling period allows for lowpass filtering before subjecting y to feedback control, making $\sigma^2 = 9^2$ BIS a somewhat pessimistic assumption in that context.

[4] The infusion rate corresponds to 1200 ml/h (being the max rate of e.g. the Alaris TIVA infusion pump, (BD, Franklin Lakes, NJ)) with 10 mg/ml propofol (being a common emulsion concentration).

Table 1. Parameters (left) together with time constants $1/p_:$ and $1/z_:$ (right) of the system poles and zeros of the model (1)–(5) corresponding to Fig. 2–3.

	k_{10}	k_{12}	k_{13}	k_{21}	k_{31}	k_{e0}	γ	η	$1/p_1$	$1/p_2$	$1/p_3$	$1/k_{e0}$	$1/z_1$	$1/z_2$
Unit			10^3/s				1	μg			min			
θ^o	4.8	8.6	2.9	2.1	0.068	2.4	1.47	19342	0.95	18	406	6.8	7.9	246
Δk_{10}	475	7626	243	213	0.72	1.3	1.66	267	0.0020	0.92	35	13	0.078	23
Δk_{12}	0.94	0.086	0.17	2.9	0.54	18	1.48	1.4e5	5.5	14	41	0.93	5.7	31
Δk_{13}	12	5.6	0.029	0.79	0.79	1.3	1.48	10857	0.92	31	96	12	21	96
Δk_{21}	0.95	9.2e-10	0.19	212	0.79	18	1.48	1.4e5	0.079	12	31	0.92	0.08	21
Δk_{31}	1.0	0.18	0.63	0.47	6.8	16	1.45	1.3e5	2.2	14	46	1.0	2.5	35
Δk_{e0}	1.5e-11	1.2e-8	0.53	2647	0.030	243	5.1	2.0e5	0.0063	30	2.1e13	0.069	0.0063	558
$\Delta\gamma$	17	1222	0.25	26	0.49	25	147	4733	0.014	32	47	0.66	0.65	34
$\Delta\eta$	687	60600	316	1124	0.77	1.4	1.51	193	3.0e-4	0.91	32	12	0.015	22

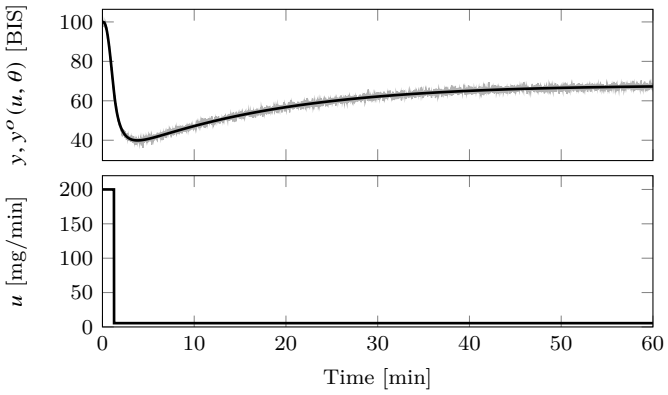

Fig. 1. *Top:* Simulated true effect $y^0 = g^o(u, \theta^o)$ (black) and noise-corrupted observation y (gray). *Bottom:* Propofol infusion profile u.

Let therefore $y = [y_1 \ \dots \ y_n]^\top$ be a vector of the simulated observations, with $u = [u_1 \ \dots \ u_n]^\top$ being the corresponding infusion rates. Since u remains constant between consecutive samples, we can express the observation model as

$$y = \underbrace{g^o(u, \theta^o)}_{y^o} + w, \tag{9}$$

where g^o denotes the *true* model structure dynamics of (1)–(5), θ is a vector comprising of the parameters of table 1 (in that order), and θ^o signifies the true parameter values. Given the known initial condition $x(0) = 0$, g^o can be evaluated approximation-free at u, θ^o through zero-order-hold (ZOH) sampling.

The data distribution is defined by the probability density function (PDF) $p(y|\theta)$ and denotes the likelihood of observing y given θ. Therefore, the data distribution is also referred to as the likelihood of θ, $L(\theta) = p(y|\theta)$. Note that both g and θ of (9) are assumed to be deterministic, and so the likelihood of θ given the data, or observation, y is defined by the PDF of w.

Since w is IID with $w \sim \mathcal{N}(0, \sigma^2)$, we have that

$$L(\theta) = p(y|\theta) = \prod_{k=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{(y_k - g_k^o(u,\theta))}{\sigma}\right)^2\right), \tag{10}$$

where $g_k^o(u, \theta)$ denotes the k^{th} element of the vector $g^o(u, \theta)$. Maximizing the likelihood $L(\theta)$ of (10) is equivalent to maximizing the log-likelihood

$$l(\theta) = \log L(\theta) = n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - V(\theta), \tag{11}$$

where the loss (or cost) function V is given by

$$V(\theta) = \frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - g_k^o(u, \theta))^2. \tag{12}$$

If $\hat{\theta}$ is an unbiased estimator of θ, meaning that $\mathbb{E}\,\hat{\theta} = \theta^o$, then the Cramér Rao Lower Bound (CRLB) expresses a lower bound on its variance:

$$\operatorname{Cov} \hat{\theta} \geq I(\theta)^{-1}, \tag{13}$$

where the Fisher Information Matrix (FIM) is given by

$$I(\theta) = H(-l, \theta) = H(V, \theta), \tag{14}$$

and where the Hessian H of a function f with respect to an argument x is a matrix with elements

$$H_{r,c}(f, x) = \frac{\partial^2 f}{\partial x_r \partial x_c}(x). \tag{15}$$

For the maximizer $\theta = \theta^o$ of $l(\theta)$ the CRLB is tight. See e.g. Kay (1993) for derivation of the CRLB and discussion of necessary, and here fulfilled, regularity conditions for it to hold.

We can note that V of (12) conveys information about how well the model output $\hat{y}(u, \theta) = g^o(u, \theta)$ resembles the observation y. Particularly, the root mean square (RMS) error is $\sigma\sqrt{2V(\theta)/n}$.

The Hessian (14) also appears in the Taylor series expansion

$$V(\theta^o + \delta) - V(\theta^o) = \underbrace{\nabla V(\theta^o)}_{0}\,\delta + \frac{1}{2}\delta^\top H(V, \theta^o)\delta + r(\delta), \tag{16}$$

where the true parameter θ^o has been perturbed by δ, and where the residual $r(\delta)$ is a linear combination of

monomials in the components of δ, each with degree of at least 3. If $\|\delta\|_2$ is small, the contribution of $r(\delta)$ to V is small.

The function $g^o(\cdot, \theta)$ has continuous second derivatives with respect to components of θ, resulting in H being symmetric with singular value decomposition (SVD) $H(V)(\theta^o) = \frac{1}{\sigma^2} U\Sigma U^\top$. The CRLB states that the principal parameter space direction variance at the true parameter is lower bounded by the diagonal elements of $\sigma^2\Sigma^{-1}$, since $I(\theta^o) = \sigma^2(U\Sigma U^\top)^{-1} = \sigma^2 U\Sigma^{-1}U^\top$. (If any diagonal element of Σ is zero, the estimate variance in the principal parameter space direction defined by the corresponding column of U will thus be infinite.)

For convenience of interpreting the sensitivities, we have normalized the parameter vector by the true parameter values, so that the true normalized parameter vector corresponding to (17)–(18) is $\theta^{o'} = 1_{8\times1}$, where $'$ signifies the normalization and $1_{8\times1}$ is a row vector of eight ones. Upon this normalization, the SVD matrices at the true parameter values become

$$U = \begin{bmatrix} -0.30 & -0.73 & -0.28 & 0.24 & -0.03 & -0.35 & -0.29 & -0.16 \\ -0.03 & -0.15 & -0.13 & -0.45 & 0.12 & -0.01 & -0.28 & 0.82 \\ -0.16 & -0.41 & -0.16 & 0.10 & 0.02 & 0.58 & 0.63 & 0.17 \\ 0.04 & 0.11 & 0.22 & 0.81 & 0.31 & -0.02 & -0.10 & 0.42 \\ 0.02 & 0.01 & 0.03 & -0.06 & 0.04 & -0.73 & 0.66 & 0.18 \\ 0.00 & 0.04 & 0.06 & 0.21 & -0.94 & -0.02 & -0.02 & 0.26 \\ 0.49 & 0.18 & -0.84 & 0.17 & -0.0 & -0.02 & 0.01 & 0.02 \\ -0.80 & 0.48 & -0.35 & 0.05 & 0.02 & -0.01 & 0.01 & 0.03 \end{bmatrix} \tag{17}$$

$$\mathrm{diag}(\Sigma)^\top = \begin{bmatrix} 273352 & 82915 & 79814 & 55607 & 19398 & 6297 & 4508 & 845 \end{bmatrix}, \tag{18}$$

where the parameter order is that of table 1. The condition number of the FIM, being the ratio between the largest and smallest element in $\sigma^2\Sigma^{-1}$ is large, $\mathrm{cond}\, I(\theta^{o'}) > 300\sigma^2$, indicating high noise sensitivity and the last column of U in (17) reveals that the least certain parameter space direction at $\theta^{o'}$ is roughly $0.8k'_{12} + 0.4k'_{21} + 0.3k'_{e0}$.

5. GLOBAL IDENTIFIABILITY

5.1. *Flexibility of parameterization*

The analysis of Sec. 4.1 characterizes *local* identifiability of linear combinations of the parameters at θ^o.

To analyze global identifiability of the individual parameters, we fixed them—one at a time—to a value that was off by a factor of 100 to investigate how well the remaining parameters could compensate for this. The quadratic loss V of (12) was minimized using Newton's method with line search. Automatic differentiation was used to obtain approximation-free gradient and Hessian information, see Revels et al. (2016). After careful initialization, the model fits of Fig. 2 were obtained. The legend notation $\Delta_{\cdot}$ indicates which parameter has set to be off by a factor 100, and Δk_{13} corresponds to a reduced-order model, detailed in Sec. 5.2. Corresponding Bode diagrams are shown in Fig. 3.

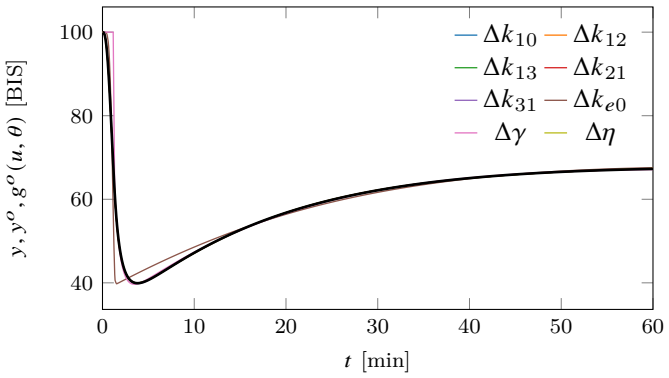

Fig. 2. Simulated true effect y^0 (black) where remaining curves show model outputs $g^o(u,\theta)$ where all parameters (elements of θ) have been optimized one at a time (see legend). k_{10}, k_{21}, k_{31}, k_{e0} and γ was fixed to a factor 100 times its true value and k_{12}, k_{13} and η to 1/100 of its true value.

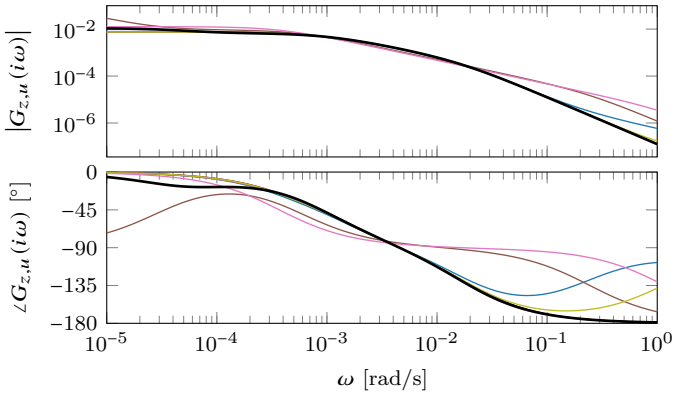

Fig. 3. Bode plots of the models in Fig. 2. See Fig. 2 for further specification of the individual models.

Considering a model $g(u,\theta)$ where $g \neq g^o$ or $\theta \neq \theta^o$ leads to a bias error $y_0(\theta)$:

$$y = g(u,\theta) + y_0(\theta) + w. \tag{19}$$

This results in a typically non-Gaussian residual $\epsilon = y_0(\theta) + w$, and consequently minimizing the loss V, corresponding to minimizing $\epsilon^\top\epsilon$, does in general no longer yield a θ that maximizes the log-likelihood l of the data. However, as seen in Fig. 2, the bias for all cases but Δ_{ke0} is negligible. Consequently, the relation (11) between V and l holds with good approximation for these cases. For Δk_{e0}, and to a lesser extent $\Delta\gamma$, the fitting errors in Fig. 2 can be visually detected. This corresponds to some degree of identifiability, but reducing the error factor to 8 for k_{e0} and 2 for γ enables fitting models that are visually indistinguishable from the true dynamics in Fig. 2.

Inspired by the reduced-order models proposed in da Silva et al. (2012) we also minimized the quadratic loss V of (12) for a reduced-order model with three asymptotically stable real poles and one asymptotically stable zero. It is the result of this minimization plotted as Δk_{13} in Fig. 2, and it naturally raises the question of which reduced-order models correspond to (3) through pole-zero cancellation, and what flexibility of the parameterization such cancellation provides.

5.2. Pole-zero cancellations

Here we characterize conditions imposed on the poles and zeros of reduced-order transfer functions, for them to constitute a realization of (1)–(2) with $\theta > 0$ and what flexibility, i.e. loss of identifiability, the cancelled dynamics provide.

It follows directly from the compartmental structure of (1)–(2) that $\theta > 0$ implies that $G_{z,u}(s)$ has strictly positive static gain, and that all its poles and zeros are real negative numbers (asymptotically stable). Before commencing the analysis we also note that there is a symmetry in that indices 2 and 3 of (1) can be interchanged.

No cancelled dynamics Assuming that no two of k_{e0}, k_{21}, k_{31} are equal, the static gain, zeros, and effect site pole are uniquely determined by η, k_{21}, k_{31}, and k_{e0}, respectively. The characteristic equation (4c) then uniquely determines $k_{10} = p_1p_2p_3/(k_{31}k_{21})$, and we can rewrite (4a)–(4b) as

$$\underbrace{\begin{bmatrix} p_1 + p_2 + p_3 - k_{10} - k_{21} - k_{31} \\ p_1p_2 + p_1p_3 + p_2p_3 - k_{10}k_{21} - k_{10}k_{31} - k_{21}k_{31} \end{bmatrix}}_{b} = \underbrace{\begin{bmatrix} 1 & 1 \\ k_{31} & k_{21} \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} k_{12} \\ k_{13} \end{bmatrix}}_{k}. \tag{20}$$

It is necessary that $b = [b_1\ b_2]^\top > 0$ for all parameters to be positive. Since $k_{21} \neq k_{31}$, M has full rank and we are looking for a solution $k = [k_{12}\ k_{13}]^\top = M^{-1}b > 0$ of (20). We see from $b_1 = k_{12} + k_{13}$ that $k_{12} < b_1$ (we only treat one of the two symmetric cases of k_{12} and k_{13} explicitly) is a necessary condition for the existence of such solution. With $0 < k_{12} < b_1$ we fulfill $b_1 = k_{12} + k_{13}$ with $k_{13} = b_1 - k_{12} > 0$. We now rewrite the second characteristic equation as $b_2 = k_{31}k_{12} + k_{21}(b_1 - k_{12})$ and solve for $k_{12} = (b_2 - k_{21}b_1)/(k_{31} - k_{21})$. Due to symmetry we can exchange k_{21} with k_{31} to ensure $k_{12} > 0$. Combining the expression for k_{12} in known entities with the requirement $k_{12} < b_1$ finally yields the necessary and sufficient condition

$$b_1 > 0 \wedge b_2 > 0 \wedge |b_2 - k_{21}b_1| < b_1|k_{31} - k_{21}|. \tag{21}$$

It is straightforward to expand the inequalities conditions into $p_1, p_2, p_3, k_{10}, k_{12}, k_{13}, k_{21}, k_{23}$, but this has been omitted here in the interest of space and readability.

5.2.1. Single PK pole cancellation Setting $k_{21} = k_{31}$ results in the dynamics x_2 and x_3 of (1) becoming identical, and

$$G_{z,u}(s) = \frac{1}{\eta} \frac{(s + k_{21})}{(s + p_1)(s + p_2)} \frac{k_{e0}}{(s + k_{e0})} \tag{22}$$

with characteristic equations $p_1 + p_2 = k_{10} + k_{12} + k_{13} + k_{21}$ and $p_1p_2 = k_{10}k_{21}$. The static gain, remaining zero and effect site pole are specified as in the cancellation-free case. This leaves the remaining parameters to determine the two remaining poles. As before, (4c) determines $k_{10} = p_1p_2/k_{21}$. Inserting this expression, the first characteristic equation can be written $b = p_1 + p_2 - k_{21} - p_1p_2/k_{21} = k_{12} + k_{13}$, and as before $b > 0$ constitutes a necessary condition. With it fulfilled and with $k_{12} < b$, $0 < k_{13} = b - k_{12}$ solves the first characteristic equation. The necessary and sufficient condition is thus $0 < k_{12} < b > 0$. Relatedly, we see directly from (1) that cancelling two of the PK poles would require $k_{12} = k_{13} = 0$, which would void the strictly positive parameter requirement.

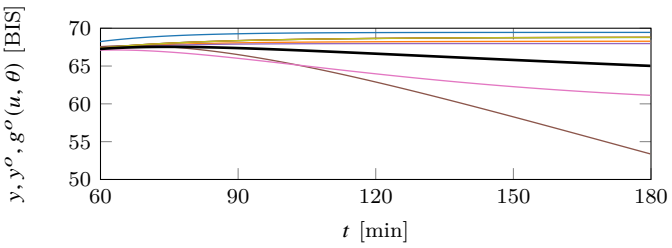

Fig. 4. Continuation of the simulations in Fig. 2 beyond the first hour used to fit the models. See Fig. 2 for further specification of the individual models.

5.2.2. Double cancellation Setting $k_{21} = k_{31} = k_{e0}$ results in two pole-zero cancellations in (3) and the transfer function

$$G_{z,u}(s) = \frac{k_{21}}{\eta} \frac{1}{(s + p_1)(s + p_2)} \tag{23}$$

with characteristic equations identical to the PK compartment cancellation case. Solving the second characteristic equation for k_{10} and inserting into the first results in the necessary and sufficient realization condition $b = p_1 + p_2 - p_1p_2/k_{21} - k_{21} > 0 \wedge b > k_{12}$ with $k_{13} = b - k_{12}$ (or symmetric case by exchanging k_{12} for k_{21}). The first inequality is quadratic in k_{21}, and the corresponding quadratic equation has solutions $k_{21} = p_1$ and $k_{21} = p_2$. Since all parameters are positive, the inequality can thus be written $\min(p_1, p_2) < k_{21} < \max(p_1, p_2)$. Solving $\partial b/\partial k_{21} = 0$ we can also see that $b(k_{21})$ is maximized at $k_{21} = (p_1 + p_2)/2$.

Again, it is straightforward to insert numerical values from the example to see within what bounds individual parameters could be varied without affecting $G_{z,u}$.

6. DISCUSSION

It should not come as a surprise that the eight individual parameters of θ cannot be reliably identified from an experiment with poor input excitation (essentially the sum of an impulse and a step) that is furthermore short compared to the slowest pole dynamic. Particularly, loss of identifiability of the static gain is expected from an experiment duration less than one eighth of the slowest pole time constant of the true system (see Fig. 4 and table 1). A clear illustration of this is provided in Fig. 4, that shows the continuations the experiments in Fig. 2 under constant u.

Yet, several factors worsen the prospect of identifiability from *real* data. Firstly, we have here assumed that the data was generated by dynamics that are in the set of the considered model class and that no disturbances act on the system in addition to a Gaussian IID measurement noise of known variance. Looking at EEG responses from actual patients, available through for example Eleveld et al. (2018) one quickly realizes that each of these assumptions is far from realistic. Co-administration of drugs also influences the anesthetic depth. Notably, opioids (administered for their analgesic effect) typically decrease y, see Vuyk (1997). The extent by which nociceptive stimuli affect y can to some extent be controlled using analgesic drugs, but not to an extent that their influence can be neglected.

Secondly, the observation model (7) is a simplification mainly in that EEG monitors introduce considerable time lag caused by internal signal processing. For the NeuroSense, the lag dynamics are LTI and constituted by ZOH sampling of a second order filter with real pole time constants of 8 s, at $T_s = 1$ s, as further explained in Bibian et al. (2011). For the much more commonly applied BIS monitor (Medtronic, Dublin, Ireland), the corresponding filtering dynamics are proprietary, and experiments reported in e.g. Lee et al. (2019); Bibian et al. (2011) have demonstrated that they exhibit nonlinear rather than LTI behavior.

The two principal alternatives available to improve the situation is to use simplified models, as proposed in for example da Silva et al. (2012), Wahlquist et al. (2020), van Heusden et al. (2013), and account for the associated bias by designing controllers with a higher degree of robustness to model uncertainty to account for this bias.

The described situation arises in many domains when complex dynamics are to be identified from non-informative data. In our case, identification of the full PKPD structure results in low bias (at least for simulated data) but high variance (uncertainty) of the parameter estimates. Reduced-orders models can reduce this variance, but at the cost of increased bias.

This brings up the central question of what the model will be used for. Historically the closed-loop anesthesia control research community has relied on the PKPD structure, perhaps mainly because models reported in the literature have that structure. However, if the purpose of the model is feedback controller design, there is no apparent benefit in sticking to the mechanically motivated model structure of (1)–(2). Therefore, it can be viewed as good news that comprehensive original experimental datasets for propofol PK and PD modeling have lately been compiled, perhaps most notably in the supplement of Eleveld et al. (2018). Setting out in such data it is for the first time possible to combine the activities of modelling and controller design, thus eliminating the reason to guess (or neglect) what uncertainties to associate with published PKPD models when considering them as a basis for controller design.

7. CONCLUSION

The classic PKPD model structure for propofol is not practically identifiable from induction phase data available in the clinic. Reliable individualization of drug delivery, beyond what is possible through compensating for demographic trends that can be determined *a priori*, will therefore need to be achieved by other means. The recent availability of comprehensive original propofol PKPD modeling datasets opens up for the identification of other model structures. If the objective is closed-loop control, it is more important for the dynamical patient model to be accurate around the intended closed-loop bandwidth than for it to maintain a physiologically motivated structure. Particularly, a closer investigation on how control performance is affected by employing less flexible models (with fewer parameters), is motivated in the light of the investigated identifiability issues.

8. ACKNOWLEDGEMENTS

The authors would like to acknowledge Johan Grönqvist, Lund University, for help with the implementation of the pharmacological model. This work was partially funded by the Swedish Research Council (grant 2017-04989). The authors are members of the ELLIIT Strategic Research Area at Lund University.

References

Agrawal, G., Bibian, S., and Zikov, T. (2010). Recommended clinical range for WAV_{CNS} index during general anesthesia. In *Proceedings of the American Society of Anesthesiologists (ASA) 2010 Annual Meeting*. San Diego, CA.

Bibian, S., Dumong, G., and Zikov, T. (2011). Dynamic behavior of BIS, M-entropy and neuroSENSE brain function monitors. *Journal of clinical monitoring and computing*, 25, 91–87. doi: 10.1007/s10877-010-9266-9.

Copot, D. (2019). *Automated drug delivery in anesthesia*. Elsevier, Amsterdam, Netherlands. doi:10.1016/B978-0-12-815975-0.00010-2.

da Silva, M., Wigren, T., and Mendonca, T. (2012). Nonlinear identification of a minimal neuromuscular blockade model in anesthesia. *IEEE Transactions on Control Systems Technology*, 20, 181–188. doi:10.1109/TCST.2011.2107742.

Eleveld, D., Colin, P., Absalom, A., and Struys, M. (2018). Pharmacokinetic–pharmacodynamic model for propofol for broad application in anaesthesia and sedation. *British Journal of Anaesthesia*, 120, 942–959. doi:10.1016/j.bja.2018.01.018.

Ghita, M., Neckebroek, M., Muresan, C., and Copot, D. (2000). Closed-loop control of anesthesia: Survey on actual trends, challenges and perspectives. *IEEE Access*, 8, 206264–206279. doi: 10.1109/ACCESS.2020.3037725.

Gonzales-Cava, J., Bagge Carlson, F., Troeng, O., Cervin, A., van Heusden, K., Dumont, G., and Soltesz, K. (2020). Robust PID control of propofol anaesthesia: uncertainty limits performance, not PID structure. *Computer Methods and Programs in Biomedicine*, 198. doi:10.1016/j.cmpb.2020.105783.

Kay, S. (1993). *Fundamentals of statistical signal processing. Volume I: Estimation Theory*. Prentice Hall, Upper Saddle River, NJ.

Lee, H., Rye, H., Park, Y., B.S., Y., Yang, S., H.-W., O., and Jung, C. (2019). Data driven investigation of bispectral index algorithm. *Scientific Reports*, 9. doi:10.1038/s41598-019-50391-x.

Revels, J., Lubin, M., and Papamarkou, T. (2016). Forward-mode automatic differentiation in Julia. *arXiv:1607.07892 [cs.MS]*. URL `https://arxiv.org/abs/1607.07892`.

Soltesz, K., Hahn, J.O., Hägglund, T., Dumont, G.A., and Ansermino, J.M. (2013). Individualized closed-loop control of propofol anesthesia: a preliminary study. *Biomedical Signal Processing and Control*, 8(6), 500–508. doi:10.1016/j.bspc.2013.04.005.

Soltesz, K., van Heusden, K., Dumont, G.A., Hägglund, T., Petersen, C., West, N., and Ansermino, J.M. (2012). Closed-loop anesthesia in children using a PID controller: a pilot study. *IFAC Proceedings Volumes*, 45, 317–322. doi:10.3182/20120328-3-IT-3014.00054.

van Heusden, K., Ansermino, J.M., Soltesz, K., Khosravi, S., West, N., and Dumont, G.A. (2013). Quantification of the variability in response to propofol administration in children. *IEEE Transactions on Biomedical Engineering*, 60(9), 2521–2529. doi: 10.1109/TBME.2013.2259592.

Vuyk, J. (1997). Pharmacokinetic and pharmacodynamic interactions between opioids and propofol. *Journal of clinical anesthesia*, 9, 23S–26S. doi:10.1016/s0952-8180(97)00117-7.

Wahlquist, Y., van Heusden, K., Dumont, G., and Soltesz, K. (2020). Individualized closed-loop anesthesia through patient model partitioning. In *Conference of the IEEE Engineering in Medicine and Biology Society*, 361–364. Quebec, Canada. doi: 10.1109/EMBC44109.2020.9176452.