Data-driven modelling of pelvic floor muscles dynamics

Steffi Knorn∗, Damiano Varagnolo**, Ernesto Oliver Chiva**, Reinhilde Melles***, and Marieke Dewitte***

∗Department of Engineering Sciences, Uppsala University, Sweden
**Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden
***Department of Psychology and Neuroscience, Clinical Psychological Science, Behavioural Medicine, Maastricht Universitair Medisch Centrum, Maastricht, The Netherlands

Abstract: This paper proposes individualized, dynamical and data-driven models that describe pelvic floor muscle responses in women that use vaginal dilation. Specifically, the models describe how the volume of an inflatable balloon inserted at the vaginal introitus dynamically affects the aggregated pressure exerted by the pelvic floor muscles of the person. The paper inspects the approximation capabilities of different model structures, such as Hammerstein-Wiener and NARX, for this specific application, and finds the specific model structures and orders that best describe the recorded measurement data. Hence, although the current dataset is drawn from a sample of healthy volunteers, this paper is an initial step towards better understanding women’s responses to vaginal dilation and facilitating individualised medical vaginal dilation treatment.

Keywords: female sexual dysfunction, black box, nonlinear models, system identification

1. INTRODUCTION

It is estimated that 30-40% of women suffer from Genital pain / penetration disorders (GPPD), e.g., painful experiences during sexual intercourse, at least once in their life (Goldstein et al., 2009, Chap. 2). The problem can be caused by a variety or combination of physiological processes (e.g., complications after cervix cancer surgeries, vaginal radiotherapies, Mayer-Rokitansky-Küster-Hauser syndromes, male-to-female gender confirmation surgeries) and psychosocial processes (e.g., traumatic sexual experiences) (Goldstein et al., 2009, Chap. 3). It is assumed that psychological mechanisms (e.g., anxiety, catastrophising pain and avoidance of sexual intimacy) and interpersonal factors (e.g., hostile partner responses, relationship conflict) may maintain, prolong and exacerbate the suffering.

Treatments of GPPD combines psychological and physiological treatments. The latter often include stretching the vaginal duct, desensitizing the vestibulum, and relaxing the pelvic floor muscles (Binik et al., 2006; Bergeron et al., 2008; Goldstein et al., 2011) through vaginal dilators. However, since these therapies are perceived as invasive, lengthy and uncomfortable, several patients delay, avoid or stop treatment, which might be improved by individualising the vaginal dilation patterns. To the best of our knowledge it is nonetheless still unclear how to quantitatively perform this individualisation step. Thus, control-oriented models are needed for design purposes.

The medical literature comprises studies that analyse some implications with medical-oriented approaches. For instance, it is known that sexual arousal in women initiates genital blood flow, and that this leads to vasocongestion of the vestibular bulbs, Puppo (2013), and vaginal lubrication, Levin (2002); Boyer (2009). Further, it is known that inducing vibrations in a “suitable range” of the inner labia and the vestibular bulbs can facilitate and intensify orgasms, Puppo (2011). Also, both touch and pain perception thresholds increase with the physiological arousal levels once the patient is stimulated with vibrations, while for the thermal stimulation case the touch and pain perception thresholds do not seem to be greatly influenced by physiological arousal levels, Gruenwald et al. (2007). On the other hand, experiencing fear induces activity of the pelvic floor muscle, van der Velde et al. (2001); Both et al. (2012). Tense pelvic floor muscles before or at the beginning of the penetrative act may then lead to decreased blood flow and lubrication, Van Lunsen and Ramakers (2002); Binik et al. (2006). Hence, penetrative activities with no or little arousal or initial activity of the pelvic floor muscles due, e.g., to fear may cause vulvar pain, Brauer et al. (2006); Farmer and Meston (2007); ter Kuile et al. (2010). When it comes to understanding the behaviour of the pelvic floor muscles, several models exist as summarised in Li et al. (2010). However, they mainly focus on childbirth and do not describe women suffering from female GPPD.

The models presented above do not focus on connecting causes with effects for all variables involved in experiencing genital pain, or on describing the dynamics of these variables. To the best of our knowledge the unique model that describes the interplay of several key variables as a dynamic model (in contrast to static cause-effect relationships as described above) is derived in Varagnolo et al. (2017). Here the variables form two distinct loops, called
the Circle Of Fear (COF) and Circle Of Pleasure (COP). The COF captures the facts that: i) pelvic muscle activity before or at the beginning of penetration may lead to pain; ii) fear induces muscular tension; and iii) inducing positive erotic stimuli may reduce fear. The COP relies on the Basson model of the female sexual response, Basson (2000), and models that i) the physiological arousal increases if the patient is sexually stimulated and subjectively aroused; ii) the subjective arousal increases with sexually stimulation and pleasurable physical sensations; and iii) physiological arousal affects the subjective arousal indirectly via the intermediate state variable of physical pleasure. The model in Varagnolo et al. (2017), though, is solely based on known cause-effect relationships in the medical literature, informed guesses from experts in the field, and the objective to find a suitable deterministic mathematical model that strikes a balance between being able to accurately model some known relationships and being simple enough to be mathematically analysable. However, the model in Varagnolo et al. (2017) is neither directly based on specific medical tests nor measurement data, and is hence not validated.

Here, we do the first step towards closing this gap by deriving data-driven dynamical models of female response to vaginal dilation using time-series of pelvic floor pressure collected from healthy patients during ad-hoc medical trials. We investigate which type of model and model order are suitable to accurately describe the recorded data. Importantly, given our vision of providing tools for designing personalised vaginal dilation patterns, we focus specifically on models with control-oriented structures such as Hammerstein-Wiener and Nonlinear autoregressive exogenous (NARX) models, which have been shown to be suitable in other biomedical applications, see Bro and Medvedev (2017); Langdon et al. (2016).

The paper is organised as follows: Section 2 describes the experimental setup used to record the measurement data. Section 3 overviews the standard strategies of modelling generic muscular activity. Section 4 presents our identification results. Section 5 closes the paper by drawing some qualitative and quantitative conclusions.

2. MEDICAL TRIALS SETUP

To derive quantitative dynamical models of how the pelvic floor muscles respond to forced vaginal dilation we use the dataset recorded at Maastricht University Hospital and described in more detail in Melles et al. (2018). The data comprises participants’ responses to the gradual vaginal dilation forced by a Vaginal Pressure Inducer (VPI), an inflatable balloon to be inserted at the introitus as described in Figure 1, while watching sequences of 5-minutes long erotic or non-erotic movies in the (tentatively) neutral environment shown in Figure 2.

The study included 42 women without sexual problems, aged between 18 and 45 years, in a steady heterosexual relationship for at least 3 months, and being sexually active including coitus. Each individual participated in single sessions where, while using the VPI and watching movies sequences, they recorded their perceived level of comfort (on a scale from 0 to 100) with an opportune slider. As soon as the pressure felt unpleasant, participants could end the experiment and force the deflation of the balloon. The sessions started with the presentation of a neutral acclimatisation movie with pressure induction using the VPI. This was followed by showing one high-arousal sexual movie without inducing vaginal pressure, then followed by four randomised movies with inducing pressure (one high-arousal and sexual, one low-arousal and sexual, one high-arousal and nonsexual, and one neutral movie). A typical data set is shown in Figure 3.

3. MODELLING OF THE PHYSIOLOGICAL RESPONSE TO VAGINAL DILATION

We are interested in modelling the muscular pressure exerted by the body based on measurement data recorded in the test described in Section 2. We model the aggregated muscular pressure exerted by the pelvic floor muscles as the output of the system, while as inputs we consider the volume of the VPI and the perceived pleasure levels.

Notice that our modelling problem is reminiscent of the one of tying muscular stimulation levels with the corresponding pressure (or force) outputs, a general problem for which researchers developed many different generic models of different complexities. These include physiologically based models (that relate stimuli and corresponding forces as interactions of the fibers at a microscopic level, Huxley...
Hill-type models (that relate stimulation levels and corresponding forces through mechanically-inspired concepts, Hill (1977)), and black-box models (that relate input-output relations starting from numerical evidence). In this work we are interested in the last type of models because in contrast to the standard literature on muscle models, we are not provided with

1. measurements of the muscular stimulation signals or other signals that are known to correlate with them (e.g., Electromyography (EMG) levels);
2. measurements of the force or pressure exerted by a specific set of muscular fibers;
3. measurements of the mechanical parameters of the muscular fibers (such as thickness and length).

In other words, since the available data is not compatible with standard physiologically-based or Hill-type models; we thus follow a purely data-driven approach.

Literature on black-box methods for modelling muscular dynamics can be divided in terms of which estimation tool is used for learning from the dataset. The most common strategies in this case use Hammerstein-Wiener or NARX models, including Neural network (NN) and fuzzy models. Physiological models of muscular dynamics are indeed typically non-linear, so that non-linear identification approaches tend to provide better results than linear ones.

Hammerstein models, that can be described as in Figure 4, comprise a static nonlinear map (a.k.a. the static recruitment in the specialised literature), an Autoregressive exogenous (ARX) model, and an additive disturbance that may account for temporary effects like fatigue and that can be modelled through another additional transfer function.

NARX models can be described in the general form

\[ y(t) = f \left( y(t-1), \ldots, y(t-T), u(t-1), \ldots, u(t-T), \theta \right) + d(t) \]

where \( y \) is the output, \( u \) the input, and \( d \) disturbances of the system. Notice that there exists a vast literature on how to determine both the structure of \( f(\cdot) \) and the best set of parameters \( \theta \). It is nonetheless known that the problem of selecting the structure of \( f \) is a difficult task specially when the size of the available dataset is small.

For an example of literature modelling muscular dynamics through NARX approaches see, e.g., Previdi (2002).

We notice that machine-learning inspired approaches such as NN and fuzzy models may achieve great generalization capabilities when modeling input-output muscle dynamics, but have also the drawback of being more difficult to be used for automatic control purposes (e.g., controlling the volume of the VPI so that the resulting vaginal pressure will likely follow a pre-specified pattern). On the other hand, Hammerstein-Wiener and more control-oriented NARX models have also been proven to be capable of high approximation capabilities for various medical applications (e.g., Bro and Medvedev (2017); Langdon et al. (2016)). In the paper we thus explore the approximation capabilities that Hammerstein-Wiener and some NARX strategies have on the available dataset.

4. RESULTS

For notational clarity, let \( i = 1, \ldots, 43 \) denote the patient ID, and \( D_i \) its associated dataset. Every \( D_i \), as described in Section 2, is composed by 4 time-series signals. Letting \( t \) denote the time index\(^1\), these signals are:

- \( m_i(t) \), boolean, indicating if patient \( i \) was watching a movie at time \( t \) or not;
- \( \ell_i(t) \), non-negative integer, indicating the perceived pleasure level of patient \( i \) at time \( t \) (this signal is nonzero only when a movie is being played);
- \( v_i(t) \), indicating the volume of the VPI at time \( t \) for patient \( i \);
- \( p_i(t) \), indicating the measured aggregate pelvic floor muscles pressure at time \( t \) for patient \( i \).

Recall then that during the collection of the generic dataset \( D_i \), patient \( i \) was exposed to 6 movies (see, e.g., Figure 3). For practical reasons in the following we divide each \( D_i \) in three parts: the first, composed by the first two movies, is neglected (the reason being that in this initial part of the trial the patient acclimatizes with the experiment during the first movie and the balloon was not inflated during the second movie); the second part, composed by the 3rd and 4th movie, is used for training purposes (and will be denoted with \( D_i^{(\text{train})} \)); the third part,

\(^1\) More precisely, all the various signals have sampling periods equal to 1 second.
composed by the 5th and 6th movie, is finally used for test purposes (and will be denoted with $\mathcal{D}_i^{\text{test}}$).

We are then interested in learning individual models of the generic form

$$p_i(t + 1) = \phi_i \left( p_i(t), \ldots, p_i(t - T), v_i(t), \ldots, v_i(t - T), \ell_i(t), \ldots, \ell_i(t - T); \theta_i \right)$$

using the individual training set $\mathcal{D}_i^{\text{train}}$. The quantities especially of interest for our purposes are:

- the functional structure of $\phi_i$, assumed to be selectable within a finite set of plausible functional structures denoted with
  $$\Phi := \left\{ \phi^{(1)}, \ldots, \phi^{(M)} \right\};$$
- the model order $T_i$, assumed to be selectable within a finite set of plausible orders denoted with
  $$T := \left\{ T^{(1)}, \ldots, T^{(N)} \right\};$$
- the vector of model parameters $\theta_i$ whose dimension depends on which structure $\phi_i$ and order $T_i$ is used.

As for the set of plausible functional structures $\Phi$ we consider the set of available alternative choices when using Matlab’s system identification toolbox—In practice, Hammerstein-Wiener models with different structures for the input and output nonlinearities, plus wavelet, treepartitioning, and sigmoid NARX models.

An “individualized” learning process for every specific patient may then be performed implementing the following pseudo-code:

1. For every potential structure $\phi^{(j)} \in \Phi$ and model order $T \in 1, \ldots, 50$ learn the model using patient $i$’s training set $\mathcal{D}_i^{\text{train}}$.
2. Select the best model structure, order and parameters for patient $i$ as that triplet that leads to the best fit in the test set $\mathcal{D}_i^{\text{test}}$.

In general, this strategy leads to individual models that might differ in their types and/or orders. Different model structures, nonetheless, make the task of comparing and clustering different patients difficult. Using only one model type and model order for learning the individual model parameters, on the other hand, has the potential drawback of reducing the generalization capabilities of the estimated models.

To evaluate this trade-off quantitatively there is the need for solving the ancillary question of how to select the model type and order among the alternative competing choices. We thus consider the following strategy (the superscripts $m$ and $n$ are a mnemonic for remembering when quantities refer to specific model structures and orders):

1. For every patient $i$ learn $M \cdot N$ models (i.e., one for each couple $(\phi^{(m)}, T^{(n)})$ of potential alternative model types and order choices) using the individual training set $\mathcal{D}_i^{\text{train}}$. This means learning for each patient $i$ $M \cdot N$ different parameters vectors $\hat{\theta}_i^{(m,n)}$;
2. For every learned $\hat{\theta}_i^{(m,n)}$ (thus for every patient $i$, model type $m$ and model order $n$) compute the simulated pressure $\hat{p}_i^{(m,n)}$:
   $$\hat{p}_i^{(m,n)} := \phi_i^{(m,n)}(\hat{p}_i, v_i, \ell_i; \hat{\theta}_i^{(m,n)})$$
   where $\phi_i^{(m,n)}$ is a type of structure and $m$ and $n$ and we tacitly let the signals $v_i$ and $\ell_i$ above belong to the test set $\mathcal{D}_i^{\text{test}}$, omit writing the time delays, and assume that the initial conditions are known and set to be equal to the measurements;
3. For every simulated pressure $\hat{p}_i^{(m,n)}$ compute its fit in the test set as
   $$\mathcal{J}_i^{(m,n)} := 100 \cdot \left( 1 - \frac{\| p_i - \hat{p}_i^{(m,n)} \|}{\| p_i - \text{mean}(p_i) \|} \right);$$
4. For every couple $(m, n)$ of potential model type and model order compute its average fit over the set of all the patients, i.e., compute
   $$\overline{\mathcal{J}}^{(m,n)} := \frac{1}{43} \sum_{i=1}^{43} \mathcal{J}_i^{(m,n)}.$$
5. Select the “best” model type $m^*$ and model order $n^*$ as that couple that is associated to the highest average fit $\overline{\mathcal{J}}^{(m,n)}$.

Notice that since this procedure is reminiscent of a cross-validation approach we do not employ information criteria like Akaike or Bayesian for penalizing higher model orders.

Figure 5 shows then a summary of the results obtained following the procedure above. More precisely, it reports the average fits $\overline{\mathcal{J}}^{(m,n)}$ for a subset of potential model types, each associated to its best model order. Interestingly, the 3 model types returning the best average fits are all Hammerstein-Wiener models and all with similar functional structures in the input and output nonlinearities. This seems to indicate that for our specific framework of modelling pelvic floor muscular pressure as a function of vaginal dilation we recover the same functional structures that have been proposed in the literature for modelling.

\[ \text{Fig. 5. Average fits on the test sets for various potential model types.} \]


\[ \text{Legend:} A = \text{Hammerstein-Wiener (HW) piecewise (pw).linear - pw.linear, } B = \text{HW pw.linear-saturation, } C = \text{HW poly.-poly., } D = \text{sigmoidnet, } E = \text{ARMAX, } F = \text{HW pw.linear-deadzone, } G = \text{wavenet, } H = \text{treepartition.} \]
generic muscular force as a function of EMG levels, see for instance Hunt et al. (1998).

In the following, assume thus the “best model type” $m^*$ to be “Hammerstein Wiener, piecewise linear input and output nonlinearities”, as indicated in Figure 5. We then quantify how much the performance of the individual estimators $\hat{f}^{(m^*,n)}_i$ may vary depending on the particular patient under consideration, given the “best” model type $m^*$ and model order $n^*$ as fixed (i.e., we analyze the spread of $\hat{f}^{(m^*,n)}_i$ over $i$). Instrumental to the final evaluation, we first check the sensitivity of the fit indexes on the model order. Consider thus the alternative strategy:

S1) select “best individual model order” that individual order $n^*_i$ that maximizes the individual fit $\hat{f}^{(m^*,n_i)}_i$ over $n$;

Figure 6 refers to strategy S1 above and plots how many patients had a certain individual model order as their best one. Unfortunately the plot does not give clear indications on what may be a suitable $n^*$ for strategy S1, in the sense that the curve is neither unimodal nor with a small overall spread. Considering not only how many patients have a specific model order as best fit (as shown in Figure 6) but also the actual fit values, the best order turns out to be $n^* = 4$.

Figure 7, instead, compares the histograms of two sets of fit indexes: the set of indexes $\hat{f}^{(m^*,n^*_i)}_i$ obtained by fixing the model type and order to be the same for the various patients (named “fixed order” in the figure), and the set of indexes $\hat{f}^{(m^*,n^*_i)}_i$ obtained using strategy S1, i.e., where the order of the models are individual variables (named “individual order” in the figure). As expected, the best fits can be obtained when allowing individual choices of model orders. The trade-off becomes thus the following: from practical reasons, allowing individual model orders might help getting models with better prediction capabilities. Restricting the models to have the same orders on the other hand allows to compare different estimated parameters $\hat{\theta}_i$ for different patients; this in its turn enables introducing generic algorithms for grouping and clustering the patients all together.

For the sake of completeness, we then consider the strategy the same model type and order for all patients (and more precisely the ones determined when computing Figure 5). We thus report in Figure 8 the simulation results for the patient associated to the worst fit against the simulations relative to the patient with the best fit. The simulation results shows that for the best case the learned model is qualitatively able to reproduce the features of the measured time series. For the worst case, instead, the simulated pressure has very poor approximation capabilities. This indicates that for some patients, considering additional data or modelling possible disturbances might be needed to obtain better models.

5. CONCLUSIONS

This paper presents results on data-driven modeling of the pelvic floor muscles dynamics for healthy patients undergoing vaginal dilation exercises through an inflatable balloon. For the available dataset, the pressure dynamics is qualitatively able to reproduce the features of the measured time series. For the worst case, instead, the simulated pressure has very poor approximation capabilities. This indicates that for some patients, considering additional data or modelling possible disturbances might be needed to obtain better models.

In the paper we specifically focused on understanding what are the effects from a system identification point of view...
of enforcing the model type and model order for capturing the dynamics of all the various patients: in a sense, we aimed at checking whether different patients share dynam- 
ics with approximately the same functional structure. The results have been though partially contradictory: even if, as said above, Hammerstein-Wiener structures seemed to capture the collected evidence for all the various patients, we haven’t been able to find a common order for the linear blocks of the various patients that led to satisfactory approximation capabilities for every patient. This is not ideal from a modelling perspective, since having the same model structure but different model orders for different patients prevents being able to compare (and thus group) the patients by means of comparing (and grouping) their estimated parameters.

In any case this works takes a step towards answering the problem of how to personalize vaginal dilation patterns by building dynamical models that are control-oriented. More precisely, personalizing the dilation patterns requires predictive models, i.e., models that can accurately forecast what will be the short- and long-term effects of applying specific dilation patterns to a specific patient in a specific condition. Here we obviously did not solve the entire problem, but rather focused on finding connections among few of the variables that are involved in the system. Future works thus include performing medical trials that involve more sources of information and extend the models derived up to now.

REFERENCES


Bergeron, S., Khalifé, S., Glazer, H. I., and Binik, Y. (2008). Surgical and Behavioral Treatments for Vestibu-
166.


Boyer, S. C. (2009). The relationship between pain sensitiv-
ity and vasocongestion due to sexual arousal in women with provoked vestibulodynia. PhD thesis, Queen’s University.


drived from a Postoperative Questionnaire. The Journal of Sexual Medicine, 3(5):923–931.


Langdon, R., Docherty, P., Chiew, Y.-S., Möller, K., and Chase, J. (2016). Use of basis functions within a non-
linear autoregressive model of pulmonary mechanics. Biomedical Signal Processing and Control, 27.


Van Lunsen, H. and Ramakers, M. (2002). The hyperac-