Simulating organic thin film transistors using multilayer perceptron regression models to enable circuit design

Laurie E Calvet, Sami El-Nakouzi, Zonglong Li, Yerin Kim, Amer Zaibi, Patryk Golec, Mei Bhattacharyya, Yvan Bonnassieux, Lina Kadura, Benjamin Iniguez

Simulating organic thin film transistors using multilayer perceptron regression models to enable circuit design

Laurie E. Calvet, Sami El-Nakouzi, Zonglong Li, Yerin Kim, Amer Zaibi, Patryk Golec, Ie Mei Bhattacharyya, Yvan Bonnassieux, Lina Kadura, Benjamin Iniguez*

11 November 2024

(accepted for publication in the special issue of *Advanced Electronic Materials* for NOD2023)

Laurie E. Calvet, Sami El-Nakouzi, Zonglong Li, Yerin Kim, Ie Mei Bhattacharyya, Yvan Bonnassieux
LPICM (UMR 7647),
CNRS-Ecole Polytechnique,
France
E-mail : laurie.calvet@cnrs.fr

Zonglong Li,
Université Paris-Saclay, Palaiseau, France

Patryk Golec
Advanced Technology Institute, School of Computer Science and Electronic Engineering, University of Surrey, Guildford, U.K

Amer Zaibi, Benjamin Iniguez
University Rovira i Virgili, Tarragona, Spain

Lina Kadura
CEA-LITEN, Université Grenoble-Alpes,38000 Grenoble, France.

Keywords: organic electronics, device modelling, circuit simulations, artificial neuron circuits

Abstract

There is increasing interest in using specialized circuits based on emerging technologies to develop a new generation of smart devices. The process and device variability exhibited by such materials, however, can present substantial challenges for designing circuits. We consider the use of three models: a physical compact model, an empirical look up table and an empirical surrogate model based on a multilayer perceptron (MLP) regression. Each is fit to measured discrete organic thin film transistors in the low voltage regime. We show that the models provide consistent results when designing artificial neuron circuits, but that the MLP regression provides the highest accuracy and is much simpler to fit compared to the compact model. The targeted technology exhibits non-ideal behavior such as variable threshold voltage and hysteresis. Using the multiplayer perceptron model, we compare the effect of such variability on the performance of the neuron circuit. We find that these effects alter the neuron firing rate and change the time spent in the on/off states but do not change the basic operation.

1 Introduction

Recent research on emerging flexible electronics has shown that robust transistors and digital circuits can be realized with a low-cost, and a greener manufacturing profile, but at the expense of device performance [1,2]. For many applications, such as continuous health monitoring devices [3] or environmental sensors, the end products target wide and relatively short term use. The question of sustainability and cost can therefore be more important than state-of-the art performance. In this paper we focus on low power organic devices, which are most promising for wearable devices that integrate with the human body.

In conventional silicon research and development, mature physical compact models are able to capture the device behavior essential for designing circuits[4]. While there is a large literature on modeling Organic Thin Film Transistors (OTFTs) [5], it is a challenge is to develop a single model that can capture the wide variety of physical effects possible in the different technologies used for device fabrication. Most notably, the interfaces resulting from different techniques cause trap states, which can vary widely in their characteristics. In addition, modeling is complicated by changes due to light, moisture and history that strongly influence OTFT behavior. These difficulties are especially important for circuit design at lower operating voltages ($\leq$ 5V) and typically models and circuits focus on higher voltage regimes.

Fitting devices to a physical compact model for a given technology has been widely investigated, but the process of determining the fitting parameters is often time consuming. In this paper we first compare three different OTFT modeling techniques: a physical compact model, an empirical Look Up Table (LUT) and an empirical model based on the Artificial Intelligence (AI) algorithm known as a multi-layer perceptron. The compact model is based on a quasi-static charge model [6], which solves Poisson's equation self-consistently using an appropriate density of states and an electron/hole transport model based on variable range hopping. The fitting procedure takes into account both transfer (drain current, I_d, versus gate-

source voltage, V_{gs}) and output (I_d versus drain-source voltage, V_{ds}) and curves. It involves about 20 fitting parameters. Nevertheless, the fitting procedure can be long, making it very difficult to explore how changes in transistor variability within a technology can impact circuits.

One possibility to overcome this difficulty is the use of empirical compact models. In this case, a large data set of OTFT transfer and output curves that includes the range of the targeted circuit operation is obtained. The simplest empirical model is to use LUT, where empirical data is placed in a table and a linear interpolation between the data points enables a model of its behavior for circuit design. LUTs are easily implemented in standard design tools such as Cadence Spectre. Nevertheless, this simplistic model is not accurate in the presence of hysteretic effects, where the I_d at fixed V_{ds} and fixed V_{gs} can be different depending on the history of the device. To overcome the difficulties related to the hysteresis observed in this dataset, the LUT used here is based on the transfer (I_d versus V_{gs}) curves.

Here we propose a model based on well-known AI algorithm, the multi-layer perceptron (MLP) regression, illustrated in **Figure 1**. In this method, the input consists of the particular function of V_{ds} and V_{gs} and the output is $\ln(I_d)$. To develop the parameters for the input vector, we first considered a linear regression to a typical diffusion transport in the sub-threshold regime[5]. It is given by:

$$I_{sub} = -I_0 \exp\left(\frac{Ln(10)}{S}\,(V_{GS} - V_0)\right)\left[1 - exp\left(\frac{qV_{DS}}{k_B T}\right)\right]$$

where I_0 is the prefactor normally dependent on the width, length and mobility of the transistor, S is the subthreshold factor, V_{th} is the threshold voltage, k_B is Boltzmann's constant and T is temperature. Noting that the logarithm of this curve is linear, we first considered a linear regression of I_d versus $\ln(V_{gs})$ characteristics. The data, as shown in Figure 2, is clearly non-linear. We therefore extended the linear fit to a polynomial of V_{ds} and V_{gs}. The results improved, but to obtain an even better result an MLP regression was used. Based on this initial analysis, we tested an input vector for the best accuracy using a polynomial of variables V_{ds} and V_{gs}. The best results were found for V_g, V_g^2, V_d, V_d^2, and V_gV_d, as depicted in Figure 1. The input layer is connected to two hidden layers, also each in the form of a vector, and finally a single output $\ln(I_d)$. Connections between the layers are matrices, whose values, known as weights, are optimized during the training process. During training, input values corresponding to known current output for all of the empirical data are used. The values include the ~ 14,000 current points from both transfer and output characteristics. A stochastic gradient descent method optimizes the weights. The final architecture with the trained values is encoded into VerilogA format that can then be read by a circuit simulator such as Cadence Spectre. Further details are provided in greater detail in the Methods section.

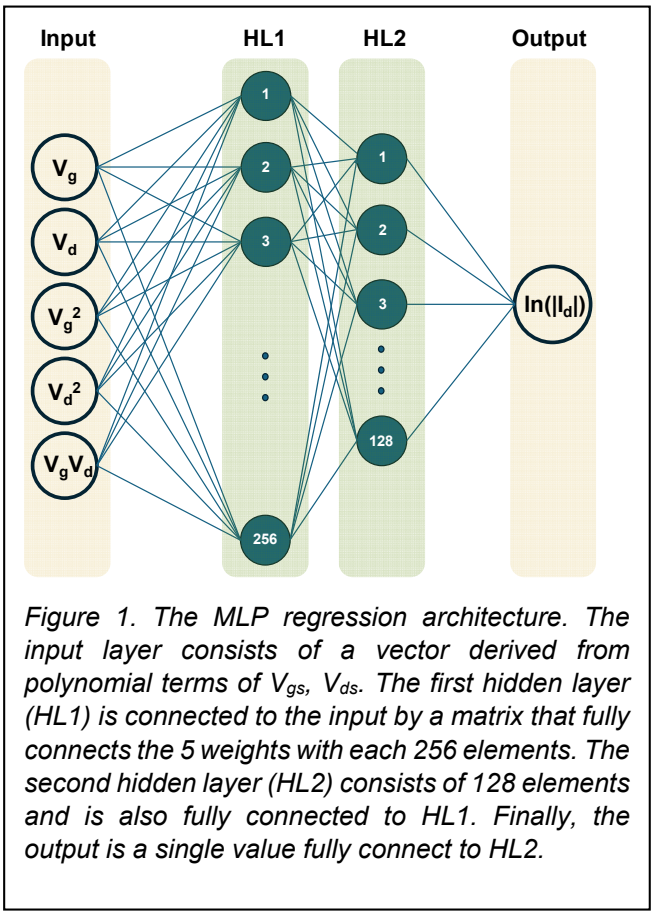

Figure 1. The MLP regression architecture. The input layer consists of a vector derived from polynomial terms of V_{gs}, V_{ds}. The first hidden layer (HL1) is connected to the input by a matrix that fully connects the 5 weights with each 256 elements. The second hidden layer (HL2) consists of 128 elements and is also fully connected to HL1. Finally, the output is a single value fully connect to HL2.

Our simulations focus on the low voltage regime of device operation because our goal is to simulate spiking neuron circuits. Such low power circuits

take advantage of the exponential dependence of the sub-threshold region.[7,8] A strong impetus for developing neuromorphic hardware in flexible materials lies in the observation that biological brains perform very well at very low speeds, low energies and with hardware prone to variability as in these technologies. While the first realization of neuromorphic hardware in organic technologies was considered more than a decade ago [9,10], this field has recently blossomed following several very important breakthroughs[11]. The majority of recent demonstrations[12–15] have used organic electrochemical transistors (OECTs)[16]. The gate action of these devices is driven by ionic doping/dedoping of the semiconductor, which has a direct analogy with the coupled ion/electron transport in biological neurons. OECTs offer facile integration with biological systems, and provide an unparalleled opportunity in health sensors.[13,15,17,18] While circuits have been realized in OECTs [19,20], the main drawback is the slow speed of the devices, at best 10s-100s of milliseconds, typically due to the thickness of the electrolyte in the gate.

Although organic thin film transistor (OTFT) technologies are comparatively well-established compared to OECTs, neuromorphic functionalities lag in comparison[21–25]. Previous demonstrations have shown spiking organic neurons in the μWs and in the ~Hz range. OTFTs, however, can operate up to MHz frequencies[26]. One target of these studies was the Axon-Hillock (AH) circuit, first proposed by Mead in 1989 [7]. Biologically the AH neuron is where the membrane potentials from the synaptic inputs are summed before being transmitted to the axon. In circuit form it is a relatively simple to realize and therefore amenable to OTFT implementations. It is a self-resetting circuit where spikes are generated above a threshold that is dependent on the geometry of the transistors and their properties. Two demonstrations of organic neuron circuits have been reported[25,27]. In one case a complementary OTFT technology used relatively low cost (vacuum evaporation) and simple processing, however very large differences in the n- and p- transistor characteristics limited its performance [27]. In

the second case, pOTFT were fabricated using solution processing and discrete resistors and capacitors were used. Spiking was observed when voltage spikes were applied to the input. Improving performance by modeling is one potential way to advance this line of research. While the first paper developed a novel modeling methodology, parameter extraction was shown to be quite time consuming. The second paper used a simple MOSFET model with parameters corresponding to the organic devices. Little discussion was provided as to the reproducibility of these circuits.

To surmount these difficulties, we explore empirical modeling of OTFTs for circuit design. We focus on devices fabricated in a well-developed organic technology using flexible Thin Organic Large Area Electronics TOLAE processes. Current research has already demonstrated many capabilities of such technologies including analog front-end [28] and multiplexing circuits [29], and Serial-Parallel conversion [30]. To date the majority of applications

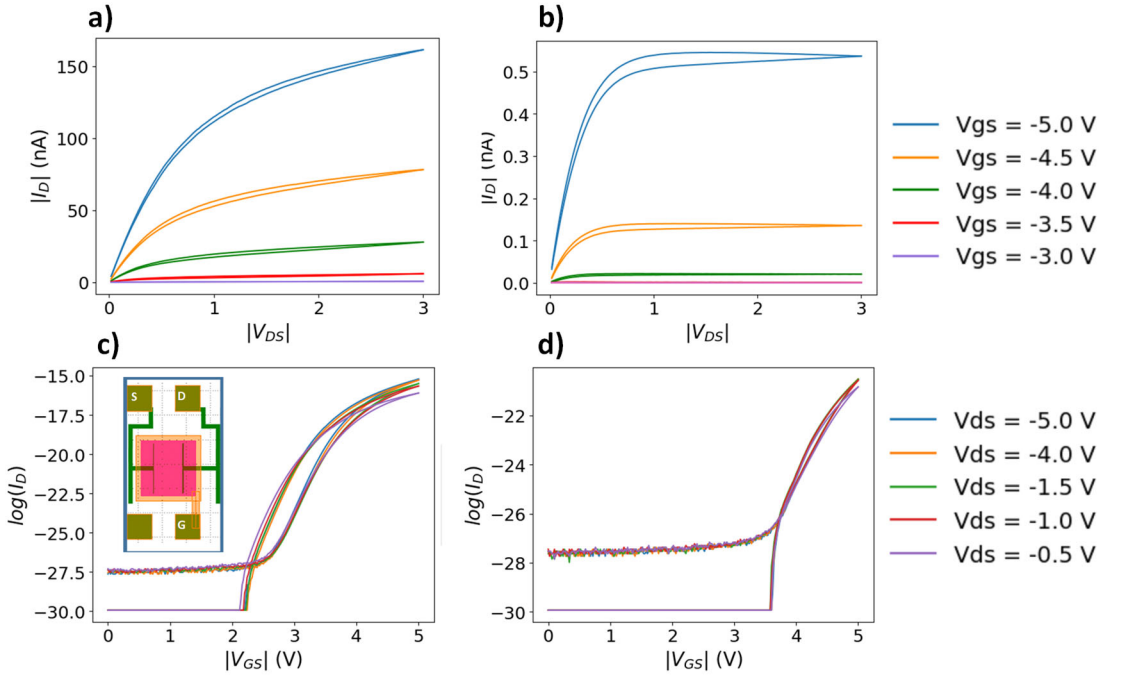

Figure 2. I_d versus V_{ds} for the a) 10 µm channel length from die 1, and b) 400 µm channel length transistors from die 1. All transistors have a 1 mm width. I_d versus V_g characteristics for the c) 10 µm channel length and d) 400 µm channel length transistors. The inset depicts a device schematic. The fabrication methods of these devices are discussed in the methods section. Data for all 10 die for these two devices are provided in Figs SI1-4.

in organic hardware have focused on using conventional digital approaches, but there are much

fewer explorations of analog and low power operation.

2. Results

2.1 Characterization and modeling of the OTFTs

We performed a detailed characterization at low voltages of a large set of p-type OTFTs consisting of 10 different size transistors repeated on 10 different die. Figure 2 shows a sample of the characteristics of two devices with 10 µm and 400 µm channel lengths with 1 mm widths. We observe an important hysteresis that complicates the modeling effort, where the voltage from zero towards greater negative voltages (forward sweep) have a lower absolute value of current compared to the sweep towards zero (reverse sweep). For the smaller length transistors, the hysteresis is larger but recovers more quickly as can be seen when comparing the hysteresis response of the I_d vs V_{ds} and I_d vs V_{gs} in the two transistor sizes. Hysteresis is likely due to charge trapping. Holes are trapped at defects in the forward sweep and then remain trapped causing a shift in the threshold voltage in the return sweep. The dependence of the hysteresis and the recovery time on the channel length suggests that the traps are dependent on the field due to the V_{ds}, possibly due to traps at the source/drain to channel interface. A complete understanding would require detailed measurements of the time dependence, which is outside the scope of this paper.

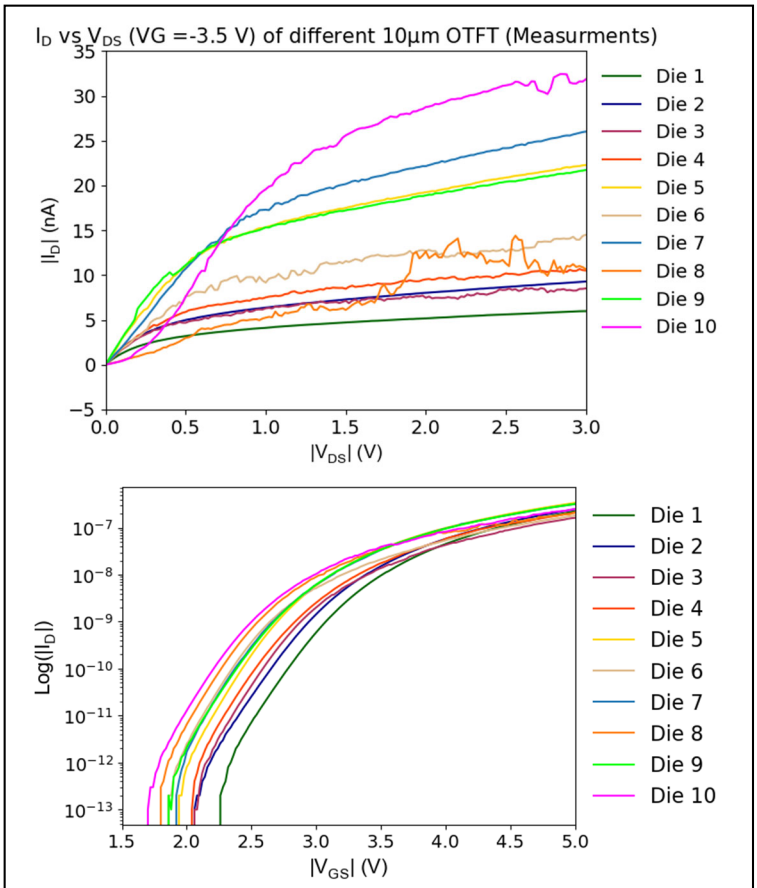

Figure 3: a) I_d vs V_{ds} at V_g = -3.5 V and b) I_d versus V_g at V_{ds} =-3.5 V measurements for all die for the 10 µm channel length devices.

We next consider the variability in the technology. Figs S1-S4 show the large variability for the two different device sizes for 10 transistors that are nominally the same size but are located on different die. Note that the size of the hysteresis varies from device to device. **Figure 3** compares the transfer and output characteristics for ten 10 µm channel length devices on different dies for fixed biases. We observe important changes in the threshold voltage and turn-on V_{on}, defined as the value of V_g at which the transistor current exceeds the background (in this case, when $|I_d| > 1 \times 10^{-13}$ A).

This large variability is an important concern when designing circuits. First, it prevents the use of a simple scaling rule to describe the technology, which is typically used in physical compact models. Instead, a model is needed for every device size and in fact for very device.

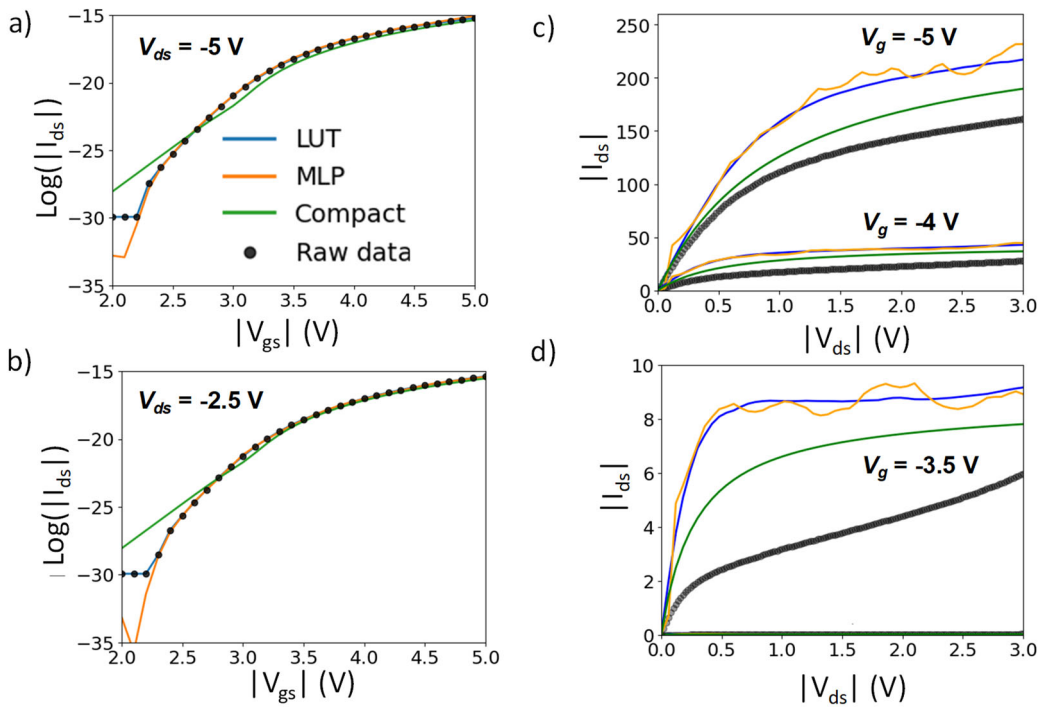

Figure 4. Transfer (a), (b) and output (c), (d) characteristics and fits to the three models for 10 µm and 40 µm channel lengths respectively. The fits used the forward sweep. Die 1 was used for both device sizes.

A second concern is the impact of the large variation in V_{on} on the circuits.

The dataset describing the devices are done using DC measurements and the lack of time dependent measurements precludes realization of an AC model. This situation is not unique. It may be encountered when trying to use data provided by a foundry to develop a model for a device that is operating in a transport regime that is not mainstream, for instance at low voltage values. In this case, time dependent measurements may not be possible to obtain and yet circuit designs need to be realized based on the DC transistor measurements. For this reason, we develop a methodology that can be used to understand the impact of the observed hysteresis during circuit design based on the partial information provided by a forward and reverse sweep.

Our modeling analysis begins with **Figure 4,** which compares the fitting of the three

	I_dV_{gs}				I_dV_{ds}		
	Compact	MLP Regression	LUT		Compact	MLP Regression	LUT
$\|V_{ds}\|$	Root mean square error (nA)	Root mean square error (nA)	Root mean square error (nA)	$\|Vg_s\|$	Root mean square error (nA)	Root mean square error (nA)	Root mean square error (nA)
5	11.40	7.03	0	5	135	162	160
4.5	9.91	2.78	0	4.5	71.5	84.6	85.8
4	8.73	1.67	0	4	28.7	32.9	34.3
3.5	7.68	2.3	0	3.5	6.76	7.75	8.52
3	7.40	1.88	0	3	0.654	0.727	1
2.5	7.50	1.34	0	2.5	0.0858	0.0859	0.0859
2	7.96	9.67	0	2	0.0987	0.0972	0.105
2.5	8.72	1.22	0	2.5	0.108	0.108	0.108
1	9.04	1.72	0	1	0.132	0.132	0.132

Table 1. Root mean square error of the 3 different models for a sampling of transfer and output curves for the 10 µm channel length device from die 1.

models using the data from forward sweep (from 0 towards negative bias). There is significantly

more data for the transfer characteristics and as a result the MLP regression and LUT (based only on the I_d-V_{gs} curves) fit better than the compact model. However, for the I_d - V_{ds}, where less data is available, the compact model is able to better capture the transport because it is based on physical reasoning. The time to simulate the transistor characteristics in cadence is ~ 0.5 s and does not substantially differ with the three models, indicating that it is negligible compare to starting the Cadence software.

Table 1 compares the root mean square error (RMSE) (definition given in the methods) for the three different models. The LUT model has zero RMSE for the transfer characteristics because the values of the transfer curves were used in the look up table. The MLP has a significantly smaller RMSE compared to the compact model for the transfer curves. For the output characteristics, we see that the compact model always performs better. This is because there were much fewer points for the output characteristics and the compact model is able to work with fewer points in general. The MLP is able to capture the output characteristics better than the LUT for a large portion of the range. In the SI Table 1 we report other error measures such as the relative root mean square error and the relative mean square error, which are necessary to compare the technique with other methods. We find that the compact model performs as well as previously reported compact models [24]. The MLP perceptron performs much better than the compact model for the transverse curves and very similarly in the output curves.

The fitting score (R^2, explained in methods) for all the MLP regressions (10 die each with 10 different device sizes) are reported in Fig. SI5. The model for Die 1, 10 μm channel length has an excellent accuracy >0.999, which can be seen as a typical value. The worst model has overall accuracies of 0.98. We have chosen to use a fixed architecture for the MLP regression for devices, as depicted in Figure 1. In future work accuracies could be improved by optimizing the architecture for each individual device.

2.2 Inverter Circuit

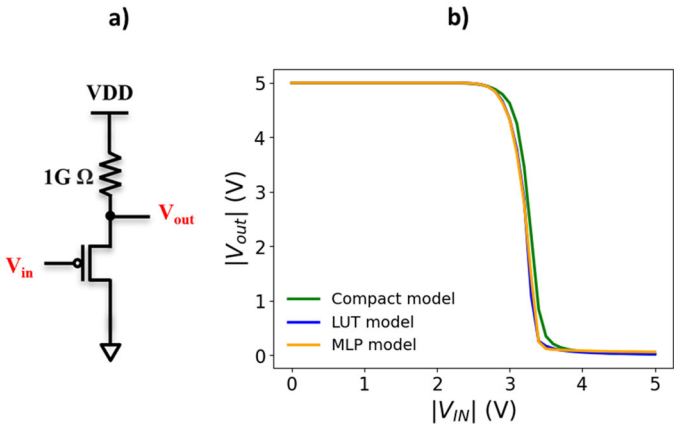

Figure 5. a) The inverter circuit used in the simulations. VDD is set to –5V. b) Simulations comparing the results of the 3 models using Cadence Spectre based on die 1 of the 10 µm channel length device and using the fit to the forward sweep model.

Figure 5 shows simulations for inverters consisting of a single transistor and a single resistor, which are important elements in the neuron circuit. A resistor loaded inverter has several disadvantages compared to a CMOS inverter, mainly a continuous power dissipation and a slower switching speed. However, here it allows us to reduce the variability of the circuit design by only using a single active transistor. An improved design could have a two resistor, one-transistor design where one resistor serves as the pull-up load and the other is used in series with the transistor to limit the power consumption. Figure 5b compares the simulation results for the three models. Despite the differences in fitting from the previous section and also the different origin of the models, they result in very similar characteristics. Because the inverter only uses a single transistor, the simulation times for this circuit were also ~ 0.5 s and it was hard to detect timing differences in the three models.

Next we consider the role of the hysteresis and variability using the MLP regression model. Figure 6 plots the variation of the inverter using the forward and reverse fits for all die from Figure 3. The inverter transition in the forward direction varies over 0.6 V and in the reverse direction by over 0.9 V. Note that there is a definite trend for V_{on} to shift to lower values as the die number increases. The hysteresis, shown in Fig 6d is relatively uniform around 0.5V. These changes may be due to changes in the homogeneity of the materials. Variations may also be due to small misalignments that are propagated in the sheet. Given these large variations, it

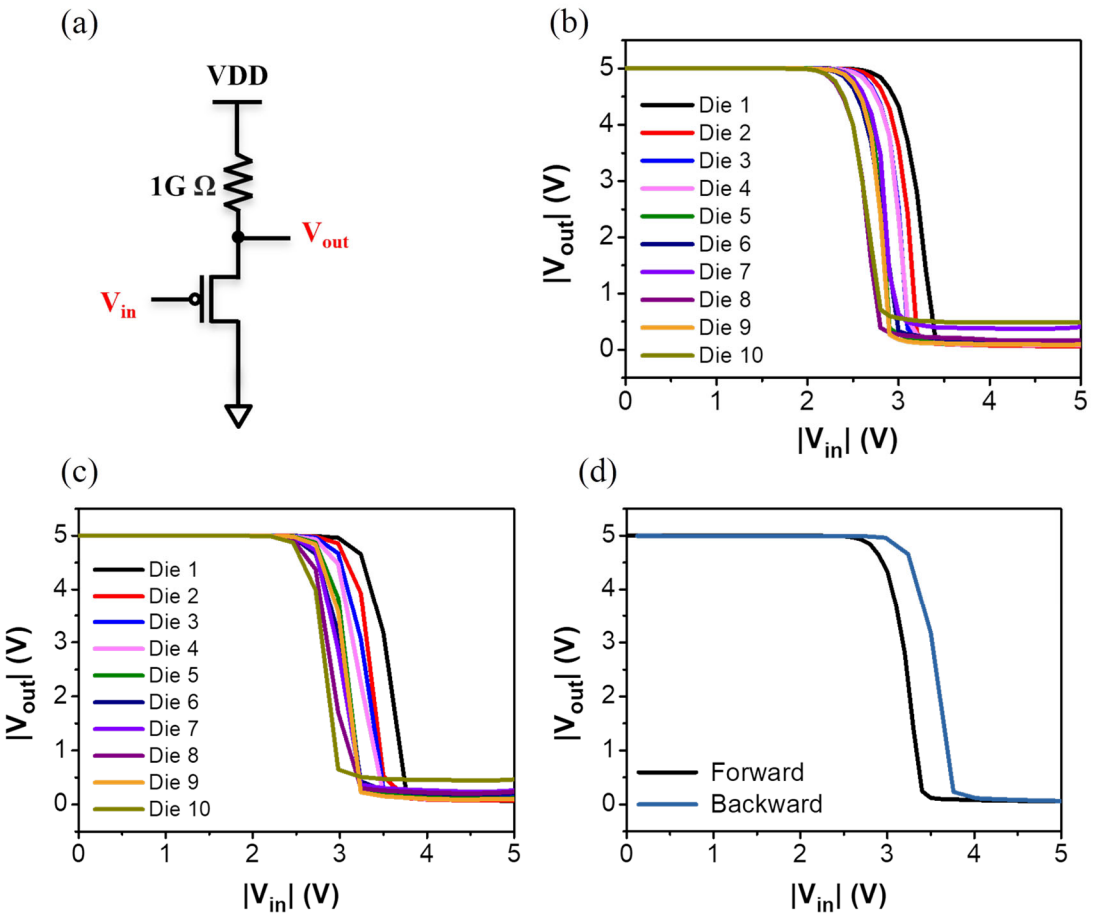

Figure 6 Variability and hysteresis of the inverter circuit (a). The forward sweep (b) and reverse sweep (c) is shown for all ten 10µm transistor length transistors. (d) The impact of the hysteresis on inverter characteristics are plotted for die 1. V_{DD} was set at -5 V.

is not clear how AH circuit will function in the presence of such effects.

2.2 Simulations of spiking neuron circuits

We consider a pOTFT implementation of the AH circuit based on the work Danneville et al[31], depicted in Figure 7a. The implementation is very appealing for organic electronics because all transistors are the same dimensions allowing us to avoid the shift to higher $|V_g|$ that is experienced by V_{on} in larger channel length devices. When no excitatory current is applied (I_{exc} = 0 A), both the output voltage V_{out} and the membrane voltage V_{mem} are 0 V. The capacitor C_f is not charged and the output of the first inverter is high. As I_{exc} is increased, C_f begins to charge, causing V_{mem} to increase. When V_{mem} reaches the threshold voltage to switch the inverters, both inverters change state and V_{out} rises towards the supply voltage VDD. A positive feedback loop

is then established through C_f, which raises V_{mem} to a value higher than V_{out}. When the reset current I_r through the transistor TR_r significantly exceeds I_{exc}, V_{mem} starts to decrease until it reaches the switching voltage of the first inverter. Consequently, the inverters switch again, bringing V_{out} back to 0V, and the cycle repeats.

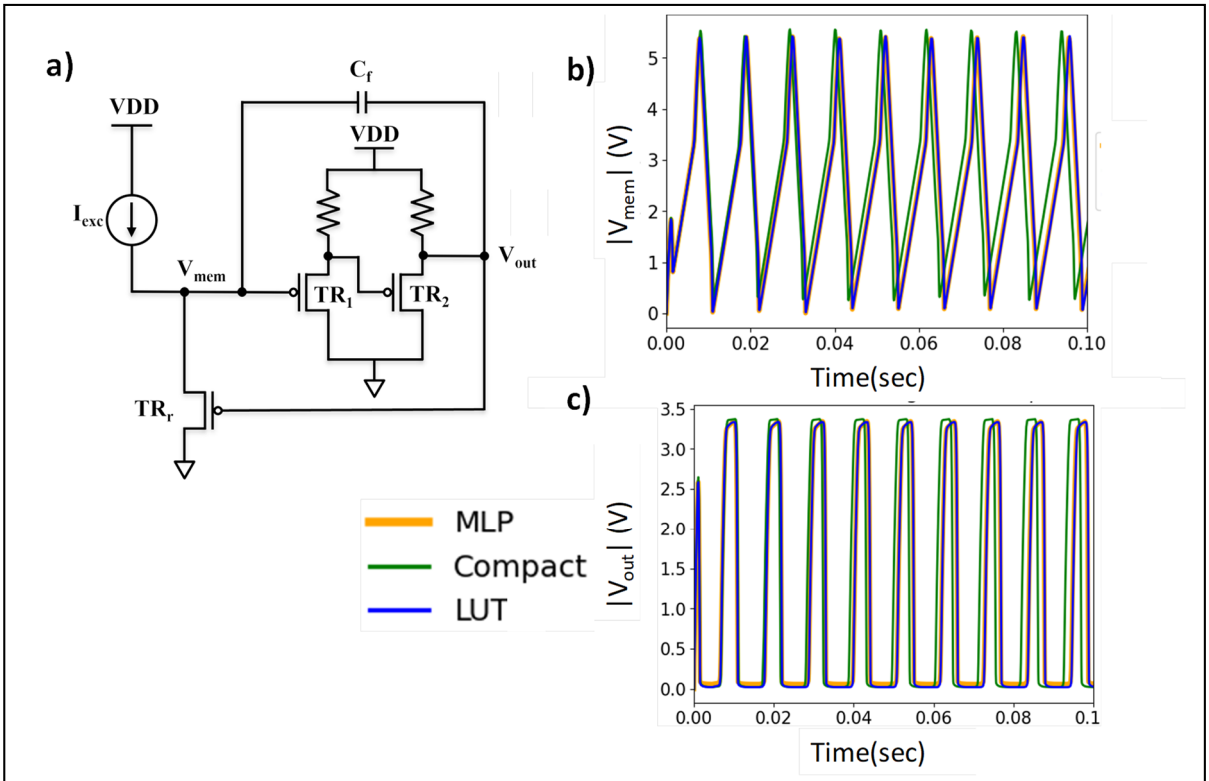

Figure 7. (a) The OTFT axon-hillock circuit. VDD is set to -5 V, the resistances are 1 GΩ, C_f = 1 pF, I_{exc} is -1 nA. The V_{mem} (b) and V_{out} (c) from simulations using the three models for the 10 µm transistor from die 1. All transistors use the same model based on forward sweep.

This circuit is different from the conventional AH circuit [7] which includes a capacitor C_{mem} connected from V_{mem} to ground. Here it is replaced by the parasitic input capacitance, which in this case corresponds to the capacitance of TR1, estimated to be ~ 0.4 pF. This value is included by setting the parasitic capacitance in the simulation parameters.

We first consider the ideal case where all the transistors use the same forward sweep model. The goal is to benchmark the MLP regression against the compact model and then it will be used more broadly to understand the expected circuit behavior. Figure 7b and 7c show

	Compact	LUT	MLP
Simulation time (s) (1 I_d-V_{gs})	0.5	2	30

Table 2. Simulation times in Cadence spectre for the AH circuit simulated for 100 ms for the three models. The simulation time for the inverter and the transistor characteristics was 0.5 s for all models.

a comparison of the AH circuits realized using the three models. As with the inverter, we find that the AH simulations do not exhibit significant variations for the different models. We do observe a slightly smaller spiking rate for the compact model, which is due to smaller resistances at larger voltages (see Figure 4). Unlike the inverter and the transistor characteristics, we find an important difference the simulation time of the three models, reported in Table 2, with the MLP model taking nearly 2 order magnitudes longer. For small circuits like the AH circuit, the time is still not a barrier. The power consumption in the simulated circuits is found to be < 5 nW and the firing rate is ~ 100 Hz. Having established that the MLP is representative of the transistors and provides similar results to the established compact models, we now consider simulations that explore the impact of the hysteresis and variability on the AH circuit.

L	Spikes/sec
5 µm	128
10 µm	87
20 µm	91
100 µm	11
200 µm	9

Table 3. Approximate firing rate when the AH circuit in Figure 7a is used for other transistor sizes. These simulations were done using die 1 for each size, the same transistor model in the forward sweep for TR1, TR2 and TRF and single transistor size. I_{exc} = -1 nA, the resistors were 1 GΩ and C_f = 1 pF.

We used the MLP model to fit other transistor sizes to test whether the model functions for other sizes and to test how the firing rate varies. Table 3 reports results on the firing rate for different size devices. Increasing transistor sizes while keeping the resistors the same alters the switching voltage of the inverters and also changes the charging time. Due to larger resistance with device length, the firing rate decreases.

Figure 8 considers how the variation of V_{on} impacts the firing rate by comparing the axon-hillock circuits using different 10 µm transistors and the forward sweep MLP model. Note that the transistors in the circuit (TR_1, TR_2, TR_r) use the same model. The firing rate varies from 80 Hz in Die 1 to 100 Hz in Die 10. This

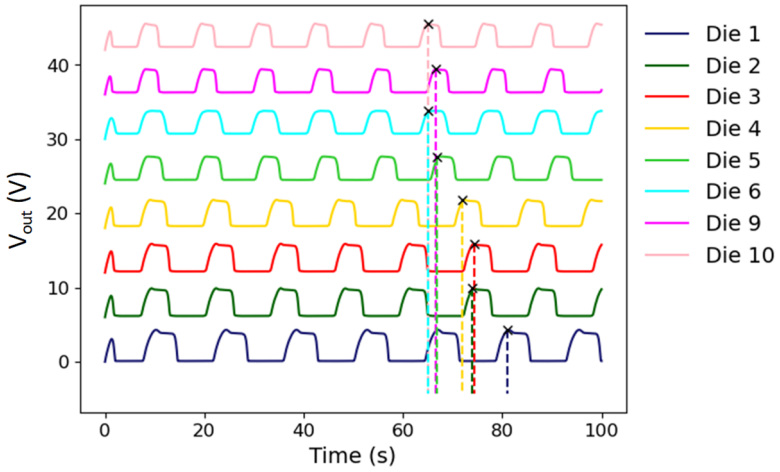

Figure 8. (a) Comparison of V_{out} of the axon-hillock circuit using the MLP regression empirical model for transistors on different die that are 10 µm length. Each curve is offset by 5V for clarity. The lines depict the change in the placement of the 7th or 8th spike in order to depict the variability due to changes in transistor characteristics.

shift corresponds to the shift in the transition in the inverter and also the shift of V_{on} of the discrete transistors. A second observation is that the faster spiking neurons have a smaller voltage difference between the ON/OFF states. The main impact of an increase in $|V_{on}|$ is to decreases the firing rate. Note that this effect is similar to what was observed with increasing channel length where devices also exhibit an increase in $|V_{on}|$ and slower firing rates.

We next consider how variability of the transistors (TR_1, TR_2, TR_r) in the axon hillock circuit impacts its output. Table 4 lists the dies and the distance between the $|V_{on}|$ of the different transistors. Figure 9 shows the resulting voltages of the circuits. We find that when TR_1 uses the model for Die 1, which has the largest $|V_{on}|$, the firing rate is

Curve No.	***Transistor die*** TR_1, TR_2, TR_r	ΔV_{ON} TR_1, TR_2	ΔV_{ON} TR_1, TR_r	ΔV_{ON} TR_2, TR_r
1	1, 2, 5	< 0.10 V	0.50 V	0.40 V
2	1, 10, 5	0.75 V	0.50 V	0.25 V
3	4, 6, 5	0.20 V	0.10 V	< 0.10 V

Table 4: Transistors and their ΔV_{on} (distance between the V_{on} of the two transistors indicated) used in in Figure 9.

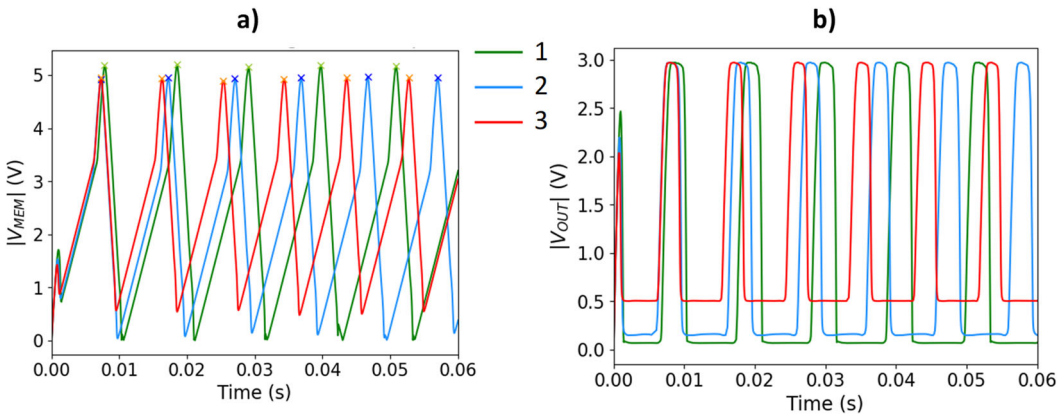

Figure 9. V_{mem} and V_{out} characteristics of the axon-hillock circuit when using different transistor models for TR_1, TR_2, TR_r. in the circuit

reduced (curve 1 and curve 2 compared to curve 3). However, the firing rate is increased by a smaller $|V_{on}|$ in TR_2. This is true for not only the firing rate in curves 1 and 2 in Figure 9, but also for curves 1 and 2 compared to the firing rates in Figure 8. We also observe that the faster firing rate in curve No. 3 (red) has a higher low state, as for the faster firing rates in Figure 8. Overall the impact of the variability in TR_1, TR_2, TR_r is very similar to the variability when the same model is used for all three transistors: it causes a change in the spike rate and can impact the difference between high and low spikes but the circuit is still seen to be functional.

Finally, we use the forward and reverse models to consider the impact of the hysteresis on the axon-hillock circuit. To do this we introduce two circuit models in Table 5 that account for the sweep states during circuit operation. The sweep states of the three transistors within the circuit differ depending on whether $|V_{out}|$ is high or low. When $|V_{out}|$ is low, $|V_{mem}|$ increases linearly until it reaches the switching threshold of TR_1. At this point TR_1 switches from low to high and TR_2 switches from high to low. This causes $|V_{out}|$ to reach VDD and switch TR_r for low to high. Therefore, in the spike up state TR_1 and TR_r undergo a forward sweep (circuit model 1), while

TR	Circuit Model 1	Circuit Model 2
TR_1	Forward model	Backward model
TR_2	Backward model	Forward model
TR_r	Forward model	Backward model

Table 5. Two circuit models that are used to understand how the hysteresis might impact the axon-hillock circuit.

TR_2 undergoes a reverse sweep. Conversely, during the spike down state, TR_1 and TR_r follow a reverse sweep, while TR_2 undergoes a forward sweep (circuit model 2).

Figure 10 shows the simulations for the two models. In a circuit with hysteresis, we expect the $|V_{out}|$ and $|V_{mem}|$ to follow the red lines during the spike up state, and the green line in the spike down state. As a result, although there is a difference in the firing dynamics between

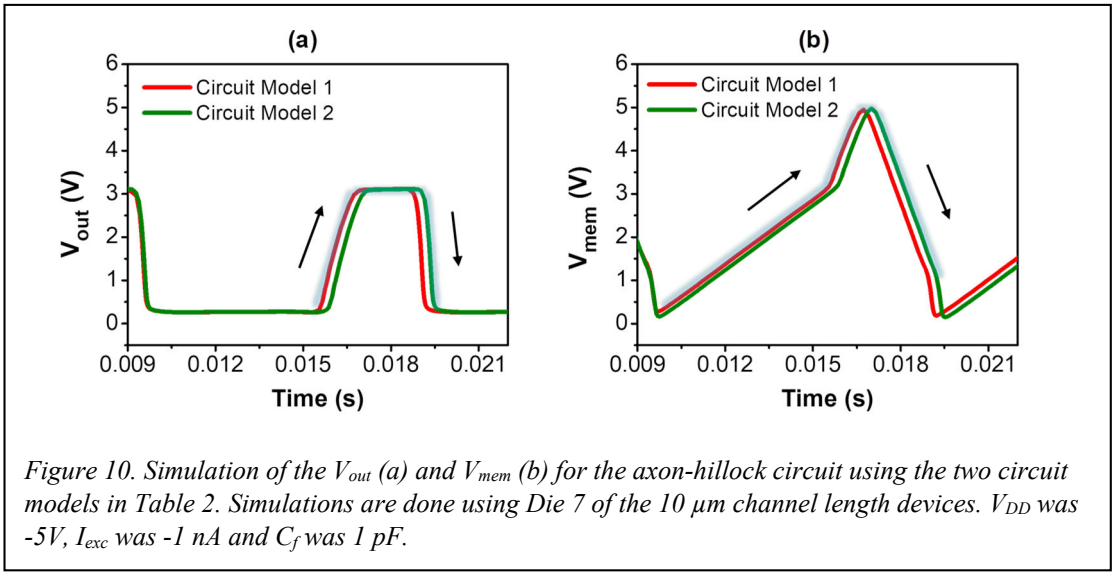

Figure 10. Simulation of the V_{out} (a) and V_{mem} (b) for the axon-hillock circuit using the two circuit models in Table 2. Simulations are done using Die 7 of the 10 µm channel length devices. V_{DD} was -5V, I_{exc} was -1 nA and C_f was 1 pF.

the circuit model 1 and 2, the overall shape of the $|V_{out}|$ spiking remains unchanged. This suggests that while hysteresis of OTFTs can introduce variability in the firing dynamics, and but it does not significantly alter the fundamental waveform characteristics of the output voltage. The resulting firing rate will remain unchanged but the relative time spent in high and low states will be different. We also see that if they hysteresis increases with sweep time, which is typically the case, the time spent in the high/low spike states will also change. Such behavior has important implications for the reliability of signal transmission, as it indicates that the axon hillock circuit can maintain consistent spiking patterns despite some hysteresis.

3. Discussion

3.1 Comparison of modeling

The three models considered here have advantages in different instances. The compact model, which is based on physics, allows us to understand the mechanisms at work in the devices. It

has much fewer parameters (~ 20) compared to the LUT (~ 12,000 parameters) and the MLP (34,176). The parameters of the compact model are physically relevant, such as the series resistance and the disorder parameter, which is used to model the variable range hopping in the semiconductor channel. The parameters can therefore be used to compare different technologies, different batches of devices in order to optimize the fabrication process. For a more mature technology the compact models can provide scalability and fitting the model to each size is not necessary.

The LUT should in theory be the most realistic model given sufficient data. However, because of the discrepancy between the I_d vs V_{gs} and the I_d vs V_{ds} characteristics; only the I_d vs V_{gs} curves are used. At low values of V_{gs} and V_{ds}, we expect the LUT to become the least accurate. It is important to note that the LUT is dependent on the linear interpolation between successive points. This requires a very fine grid of points in the more complex regions. The MLP, however is able to account for non-linearities between data points and is therefore more capable of advanced interpolations.

The main disadvantage of the MLP is that it is essentially a black box regression and the large number of parameters do not have a physical meaning. Nevertheless, it very easy to employ using standard python coding. For designing small circuits and trying to quickly understand how variability will impact circuits, it is a very time efficient method, especially if an individual model is needed for each device. It is able to capture some of the discrepancies due to the hysteresis and in some cases the fitting of the compact model may not be sufficient to obtain a working model.

We next consider the various metrics of the error. In Table 1 we considered the root mean square error (RMSE) to compare the three models. The unit of this error is the same as the data, making comparisons between it and the data very intuitive. The RMSE of the LUT in the transfer regime is of course a zero. In the output characteristics the LUT has a higher error

in a large part of the regime compared to the other two models, indicating it is important to include the output characteristics in the modeling.

We compared our relative MSE with the only known accuracy for modeling that we were able to find in published data[24], which refers to a single transfer curve. The errors in the transfer curves for compact model were the same (~ 0.02), exactly the same as this previous work, but the rel MSE of the MLP is much lower (0.001). We believe that the more appropriate measure of comparison is the relative root mean square error (RRMSE) metric because in the presence of larger errors the rel MSE can be very large. Here we find that the RRMSE is about 0.1 in the output characteristics, which means that they can reproduce the data within 10%, which is a very good model considering the changes in the transport. For this reason, we believe that the MLP is an excellent model that can be used to explore the variability of devices when needed.

3.2 Generalization of the MLP regression technique

We have explored how an MLP regression can be used to model pOTFT transistors in the low voltage regime. A natural question is whether this can this be generalized to the entire regime. Similar research has explored the use of a MLP regression in a silicon nanowire reconfigurable (r)FET [32]. They showed that it was convenient to use two MLPs, one for the sub-threshold and another for the saturation regime and to stich the two together using a mathematical function. This type of generalization could be an interesting solution for modeling complementary OTFT circuits where the n-type devices may be in the on state and the p-type devices in the sub-threshold. In this case, the model sizes should be chosen very carefully to ensure that simulation times are not excessive. From our simulations however, we do not expect that simulations times will be problematic because they are still quite small.

4. Conclusion

This paper considered modeling OTFTs in the low voltage regime using a physical compact

model, a look up table, and an empirical model based on the AI algorithm the multiplayer perceptron. The three models can be employed together to provide more insights into the devices physics or independently to facilitate modeling of a technology when all the effects cannot be described using the physical model. Our simulations indicate that organic biomimetic neurons using p-type OTFT technology are possible with firing rates >100 Hz and power consumption < 5 nW for 10 micron channel length devices. We found that the main impact of variability in $|V_{on}|$ results in variations in the firing rate. Hysteresis was found to change the relative time that V_{out} spends in the high and lows states. The axon-hillock neuron circuit is found to be very compatible with OTFT technologies but care should be taken to choose neuromorphic applications that are robust against the variability of spike rates and firing rate profile of high and low states. With improved design and a minimization of the parasitic capacitances, firing rates at least an order of magnitude larger should be possible than those previously reported. Finally, we believe that the variability in the device characteristics can be used advantageously to realize neuromorphic devices because in such systems, stochasticity is often a source of improvement via effects such as stochastic resonance.

5. Methods

The p-type OTFT transistors were realized using gravure printing on the Pilot Line at CEA-LITEN [33]. The source/drain contacts are 30 nm Au and are patterned by photolithography. It is subsequently coated with a self-assembled monolayer to improve the adhesion with the organic semiconductor. A 50 nm organic semiconductor (SP500 Lisicon® polymer from Merck) with a mobility of 2 cm^2/Vs is gravure printed. A 946 nm gate dielectric (Merck D320) of 2.2 is gravure printed. Interconnections between the source/drain and gate is achieved using vias. Finally, a conductive ink of PEDOT:PSS is screen printed on top of the stack to establish the gate electrode. Initial layouts of the axon-hillock circuit are 4 mm x 4mm in size and include

two discrete 1GΩ resistors and 1 discrete pF capacitor.

Measurements were done at ambient conditions in the dark. The semi-automatic prober is connected to a semiconductor parameter analyzer (HP B1500A) with triaxial cables. I_d-V_{ds} characteristics was measured in steps of -20 mV, sweeping both up to negative V_{ds} and back for V_{gs} with a uniform step size between 0 V to -5V of -0.5 V every 1s. Forward and reverse I_d-V_{gs} characteristics were measured in steps of -20mV with a V_{ds} step size of -0.1 V. This resulted in ~ 14,000 data points that enabled us to use a MLP model.

Ten devices for transistors lengths of 5 µm, 10 µm, 20 µm, 40 µm, 60 µm, 80 µm,100 µm, 200 µm, 400 µm, 600 µm) with 1000 µm width were measured in the sub-threshold region (V_{gs}, V_{ds} > -5 V). Typical mobilities were 1.8-2.0 cm²/V/sec with V_t ~ -5V.

The compact model is based on a DC and quasi-static charge model [6], which solve Poisson's equation self-consistently using an appropriate density of states and an electron/hole transport model based on variable range hopping. Parameters for the model are fit to the measured devices for each size and each die. The model cannot take into account directly the hysteresis. For this reason, the fits were done either using the sweep towards more negative voltages or its return. The hysteresis impeded us from using the physical compact model for larger device sizes because the currents in the I_d-V_{ds} characteristics where often not consistent with the currents in the I_d-V_g characteristics at the same biasing conditions. The devices reported in the figures are those where the model fit well. The total number of parameters used in the compact model is ~20. The parameter extraction and fitting was done in Matlab and the extracted parameters and final model were put into a verilogA code for simulations in Cadence Spectre.

In order to surmount the difficulties of the compact model, we developed a multilayer perceptron (MLP) regression to fit the data for each device. Simulations were performed using cadence and the machine learning library Sklearn[34]. Simulations were done using a single CPU

of Intel XEON microprocessor @ 3.2 GHz running on a Dell PowerEdge server. The training of an individual model requires ~ 2 minutes. We used a ReLu activation function and standard Adam for an optimized gradient descent.[35] All 56 I_d-V_g and all I_d-V_{ds} curves were included in the fit, resulting in a fit with ~ 14,000 experimental data points. We assumed that the device characteristics could be describe in terms of polynomials of V_{gs}, V_{ds}. To determine the best network architecture, we did a systematic exploration of the polynomial inputs (up to 4th order), number of hidden layers (up to 4) and number of neurons (up to 1056). The optimal network, with a typical accuracy score of 0.9996, was found using five parameters: V_g, V_g^2, V_d, V_d^2, and V_gV_d, two hidden layers (256 for the first and 128 for the second). To benchmark the regression, we used the score to determine the R^2 (Figure SI5). It is defined as:

$$R^2 = 1 - \frac{\sum_i^N (x_i - p_i)^2}{\sum_i^N (x_i - \bar{x_i})^2}$$

After training in python, the model was made compatible with VerilogA, where matrix multiplications are not available. In Cadence, we set C_{min} = 1 x 10^{-13} F to account for the parasitic capacitance.

To benchmark the different models, we used two measures: the root mean square error and 1- relative mean square error. These are defined as:

$$root\ mean\ square\ error = RMSE = \sqrt{\frac{1}{N}\sum_i^N \left(x_i - p_i\right)^2}$$

$$relative\ mean\ square\ error = Rel\ MSE = \sum_i^N \frac{\left(x_i - p_i\right)^2}{\sum_i^N x_i^2}$$

$$relative\ root\ mean\ square\ error = RRMSE = \sqrt{\frac{\frac{1}{N}\sum_i^N \left(x_i - p_i\right)^2}{\sum_i^N x_i^2}}$$

Acknowledgements

This work was supported by ANR contract ANR-21-FAI1-0006-01, and from the European Union's Horizon Europe Research and Innovation programme BAYFLEX under grant agreement 101099555.

Received: (30 June 2024)

Revised: (5 November 2024)

Published online: (())

References

[1] F. Liu, L. Lorenzelli, *Wearable Electron.* **2024**, *1*, 137.
[2] Y. Bonnassieux, C. J. Brabec, Y. Cao, T. B. Carmichael, M. L. Chabinyc, K.-T. Cheng, G. Cho, A. Chung, C. L. Cobb, A. Distler, H.-J. Egelhaaf, G. Grau, X. Guo, G. Haghiashtiani, T.-C. Huang, M. M. Hussain, B. Iniguez, T.-M. Lee, L. Li, Y. Ma, D. Ma, M. C. McAlpine, T. N. Ng, R. Österbacka, S. N. Patel, J. Peng, H. Peng, J. Rivnay, L. Shao, D. Steingart, R. A. Street, V. Subramanian, L. Torsi, Y. Wu, *Flex. Print. Electron.* **2021**, *6*, 023001.
[3] A. Pantelopoulos, N. G. Bourbakis, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2010**, *40*, 1.
[4] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*, 4th edition., Wiley-IEEE Press, Piscataway, NJ, **2019**.
[5] S. Jung, Y. Bonnassieux, G. Horowitz, S. Jung, B. Iñiguez, C.-H. Kim, *IEEE J. Electron Devices Soc.* **2020**, *8*, 1404.
[6] A. Nikolaou, G. Darbandy, J. Leise, J. Pruefer, J. W. Borchert, M. Geiger, H. Klauk, B. Iñiguez, A. Kloes, *IEEE Trans. Electron Devices* **2020**, *67*, 4667.
[7] C. Mead, *Analog VLSI and neural systems*, Addison-Wesley Longman Publishing Co., Inc., **1989**.
[8] G. Indiveri, B. Linares-Barranco, T. J. Hamilton, A. van Schaik, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Dudek, P. Häfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, K. Hynna, F. Folowosele, S. Saighi, T. Serrano-Gotarredona, J. Wijekoon, Y. Wang, K. Boahen, *Front. Neurosci.* **2011**, *5*.
[9] F. Alibart, S. Pleutin, D. Guérin, C. Novembre, S. Lenfant, K. Lmimouni, C. Gamrat, D. Vuillaume, *Adv. Funct. Mater.* **2010**, *20*, 330.
[10] V. Erokhin, T. Berzina, A. Smerieri, P. Camorani, S. Erokhina, M. P. Fontana, *Nano Commun. Netw.* **2010**, *1*, 108.
[11] Y. van de Burgt, A. Melianas, S. T. Keene, G. Malliaras, A. Salleo, *Nat. Electron.* **2018**, *1*, 386.
[12] P. Gkoupidenis, N. Schaefer, B. Garlan, G. G. Malliaras, *Adv. Mater.* **2015**, *27*, 7176.
[13] Y. van de Burgt, E. Lubberman, E. J. Fuller, S. T. Keene, G. C. Faria, S. Agarwal, M. J. Marinella, A. Alec Talin, A. Salleo, *Nat. Mater.* **2017**, *16*, 414.
[14] J. Y. Gerasimov, R. Gabrielsson, R. Forchheimer, E. Stavrinidou, D. T. Simon, M. Berggren, S. Fabiano, *Adv. Sci.* **2019**, *6*, 1801339.
[15] P. C. Harikesh, C.-Y. Yang, D. Tu, J. Y. Gerasimov, A. M. Dar, A. Armada-Moreira, M. Massetti, R. Kroon, D. Bliman, R. Olsson, E. Stavrinidou, M. Berggren, S. Fabiano, *Nat. Commun.* **2022**, *13*, 901.
[16] P. Gkoupidenis, Y. Zhang, H. Kleemann, H. Ling, F. Santoro, S. Fabiano, A. Salleo, Y. van de Burgt, *Nat. Rev. Mater.* **2024**, *9*, 134.
[17] C. Cea, G. D. Spyropoulos, P. Jastrzebska-Perfect, J. J. Ferrero, J. N. Gelinas, D. Khodagholy, *Nat. Mater.* **2020**, *19*, 679.
[18] S. T. Keene, C. Lubrano, S. Kazemzadeh, A. Melianas, Y. Tuchman, G. Polino, P. Scognamiglio, L. Cinà, A. Salleo, Y. van de Burgt, F. Santoro, *Nat. Mater.* **2020**, *19*, 969.
[19] P. Andersson Ersman, R. Lassnig, J. Strandberg, D. Tu, V. Keshmiri, R. Forchheimer, S. Fabiano, G. Gustafsson, M. Berggren, *Nat. Commun.* **2019**, *10*, 5053.
[20] R. B. Rashid, X. Ji, J. Rivnay, *Biosens. Bioelectron.* **2021**, *190*, 113461.
[21] Y. Yang, C. Bartolozzi, H. H. Zhang, R. A. Nawrocki, *Eng. Appl. Artif. Intell.* **2023**, *126*, 106838.
[22] R. A. Nawrocki, R. M. Voyles, S. E. Shaheen, *IEEE Trans. Electron Devices* **2014**, *61*, 3513.
[23] M. J. Mirshojaeian Hosseini, R. A. Nawrocki, *Micromachines* **2021**, *12*, 655.

[24] Y. Yang, M. J. M. Hosseini, W. Kruger, R. A. Nawrocki, *IEEE Trans. Circuits Syst. Regul. Pap.* **2023**, *70*, 1161.
[25] V. Tischler, P. Dudek, J. Wijekoon, L. A. Majewski, Y. Takeda, S. Tokito, M. L. Turner, *Org. Electron.* **2023**, *113*, 106685.
[26] J. W. Borchert, U. Zschieschang, F. Letzkus, M. Giorgio, R. T. Weitz, M. Caironi, J. N. Burghartz, S. Ludwigs, H. Klauk, *Sci. Adv.* **2020**, *6*, eaaz5156.
[27] M. J. M. Hosseini, E. Donati, T. Yokota, S. Lee, G. Indiveri, T. Someya, R. A. Nawrocki, *J. Phys. Appl. Phys.* **2020**, *54*, 104004.
[28] M. Sugiyama, T. Uemura, M. Kondo, M. Akiyama, N. Namba, S. Yoshimoto, Y. Noda, T. Araki, T. Sekitani, *Nat. Electron.* **2019**, *2*, 351.
[29] A. Morley, G. Lloyd, M. Charbonneau, D. Locatelli, S. Lombard, C. Laugier, L. Tournon, S. Bain, M. James, H. Wang, *SID Symp. Dig. Tech. Pap.* **2018**, *49*, 476.
[30] M. Fattori, J. Fijn, P. Harpe, M. Charbonneau, S. Lombard, K. Romanjek, D. Locatelli, L. Tournon, C. Laugier, E. Cantatore, *IEEE Electron Device Lett.* **2019**, *40*, 1682.
[31] F. Danneville, C. Loyez, K. Carpentier, I. Sourikopoulos, E. Mercier, A. Cappy, *Solid-State Electron.* **2019**, *153*, 88.
[32] M. Reuter, J. Wilm, A. Kramer, N. Bhattacharjee, C. Beyer, J. Trommer, T. Mikolajick, K. Hofmann, *IEEE J. Electron Devices Soc.* **2024**, *12*, 310.
[33] M. Charbonneau, D. Locatelli, S. Lombard, C. Serbutoviez, L. Tournon, F. Torricelli, S. Abdinia, E. Cantatore, M. Fattori, In *2018 48th European Solid-State Device Research Conference (ESSDERC)*, **2018**, pp. 70–73.
[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825.
[35] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv, **2017**.