A MODEL PROPOSAL FOR DIGITAL TWIN DEVELOPMENT

Matheus Antônio Nogueira de Andrade[a], Herman Augusto Lepikson[b], Carlos Alberto Tosta Machado[c], Yasmim Thasla Santos Ferreira[d], Sara Freitas Santos Vasconcelos[e]

a Universidade Federal da Bahia, Av. Adhemar de Barros, s/nº - Ondina, Salvador - BA, 40170-110, Brazil, matheus.andrade.lp@gmail.com
b Senai Cimatec University Center, Orlando Gomes Avenue, 1845 - Piatã, Salvador, BA, 41650-010, Brazil, herman.lepikson@fieb.org.br
c Senai Cimatec University Center, Orlando Gomes Avenue, 1845 - Piatã, Salvador, BA, 41650-010, Brazil, carlos.tosta@uol.com.br
d Senai Cimatec University Center, Orlando Gomes Avenue, 1845 - Piatã, Salvador, BA, 41650-010, Brazil, yasmim.thasla@gmail.com
e Senai Cimatec University Center, Orlando Gomes Avenue, 1845 - Piatã, Salvador, BA, 41650-010, Brazil, sarasantoos500@gmail.com

Abstract: Digital twins are becoming a powerful tool to enhance industrial processes worldwide. This paper proposes a model for the creation of industrial processes' digital twins, using a steam distillation process for essential oil extraction as a case study. A grey box modeling is suggested combining a machine learning based model with physical modeling to improve the process. Real time simulation and a hybrid control strategy are used, linked to reinforcement learning and PID control, focusing on the yield increase and optimization. A creation method is elaborated to support other applications of digital twins in industrial processes in the future.

Keywords: digital twin; creation method; essential oil; grey box modeling; real time simulation.

UMA PROPOSTA DE MODELO PARA DESENVOLVIMENTO DE GÊMEO DIGITAL

Resumo: Gêmeos digitais estão se tornando uma ferramenta poderosa para melhorar os processos industriais ao redor do mundo. Este artigo propõe um modelo para a criação de gêmeos digitais de processos industriais, usando a extração de óleos essenciais por destilação a vapor como um estudo de caso. Uma modelagem caixa cinza é sugerida combinando um modelo baseado em aprendizado de máquina com uma modelagem física para melhorar o processo, usando simulação em tempo real e uma estratégia de controle híbrida, usando aprendizado por reforço e controle PID, focando no aumento e otimização do rendimento. Um método de criação será elaborado, baseado neste estudo, para auxiliar outras aplicações de gêmeos digitais de processos industriais no futuro.

Palavras-chave: gêmeo digital; método de criação; óleos essenciais; modelagem caixa cinza, simulação em tempo real.

1. INTRODUCTION

The fourth industrial revolution, which started during the first years of the 21st century, introduced a new phase for industries, impacted by technological innovations that came from automation and Information Technology (IT). New technologies, such as Cyber-physical Systems, Internet of Things (IoT), Artificial Intelligence (AI) and robotics, are continuously bringing new possibilities to support systems that can be autonomous and customizable. This new era is commonly known as Industry 4.0 [1].

In this new environment arises the concept of digital twin, which is the digital representation of a physical asset or system used for simulation in order to find possible improvements to the physical entity [2]. A digital twin can also be interpreted as a group of virtual information that describes a physical object [3]. The physical object can be a product, a system or a process. Digital twins of process are able to anticipate the behavior and operation of a system that involves various machines, stages and procedures before implementing it in the physical world, saving time and other resources [4]. Companies all over the world are seeking ways to use less and less resources to produce more, therefore, simulation is a great ally because of its capability to predict different scenarios [5].

The twin concept was introduced when NASA (National Aeronautics and Space Administration) decided to create the physical twin of a space aircraft within the Apollo Program to reproduce its behavior in space [6]. After that, the digital twin concept shows up for the first time in 2002 referring to the existence of a twin that was digital instead of physical [3]. More recently, NASA has begun to use digital twins to develop new solutions motivated by the results they achieved with the previous twins [6]. A way to look at digital twins is to think about them as an instrument to support PLM (Product Lifecycle Management), from its creation until discard, involving all the stages between them, like design, manufacturing, usage and so on [3]. Digital twin may also refer to a digital history profile and the current state of a physical asset or process that evolves through time and becomes better, thus contributing to its development [7].

With different technologies and levels of integration involved, digital twins may be created for different purposes and different focus areas, such as layout planning, product lifecycle, manufacturing, maintenance, process design and PPC (Production Planning and Control) [8]. Digital twins focused on PPC are generally created to help the production planning by providing statistical assumptions and detailed diagnosis so all this data can be used to optimize the plant's operation [9].

The main components needed to build a digital twin are sensors/actuators, historical process data, connection between the physical and the digital worlds and data analysis. The data from the sensors are integrated to the historical data, the simulation runs in real time, mirroring the physical state, and the data analysis decides the actions that need to be taken [7].

A digital twin also needs a model that describes the real world. Industrial processes may be modeled in two different ways, using either white box or black box modeling. The white box modeling is based on first principles and physical laws that can describe the system completely and indicate its behavior. The black box is obtained from processing data to create a model that comes from correlations detected by computer analysis [10]. There is a third type of model, also referred as hybrid model

or grey box, which attempts to combine the advantages of white and black box modeling [11].

This paper presents a digital twin model instantiated in a steam distillation for essential oil extraction. Essential oils are extracted from plants and are made up mostly of low molecular weight substances. Thus, these oils have high volatility and can be extracted by steam distillation [12]. During the steam distillation process, the steam acts as a means of transportation for the essential oil particles when it passes through the botanical material. Steam and oil are later condensed producing a two-phase liquid, one phase is the oil and the other is a mixture of components including water, called hydrosol [13]. This study is proposing a grey box modeling approach for the system using machine learning and a physical modeling to achieve this goal.

2. METHODOLOGY

A small pilot plant design was created in order to be the physical object mirrored by the digital twin. The steam distillation is a relatively simple process when compared to others. It begins when the steam gets inside the distiller, passing through the mass, and is later condensed. The figure 1 is the P&ID (Piping and Instrumentation Diagram) of the process.
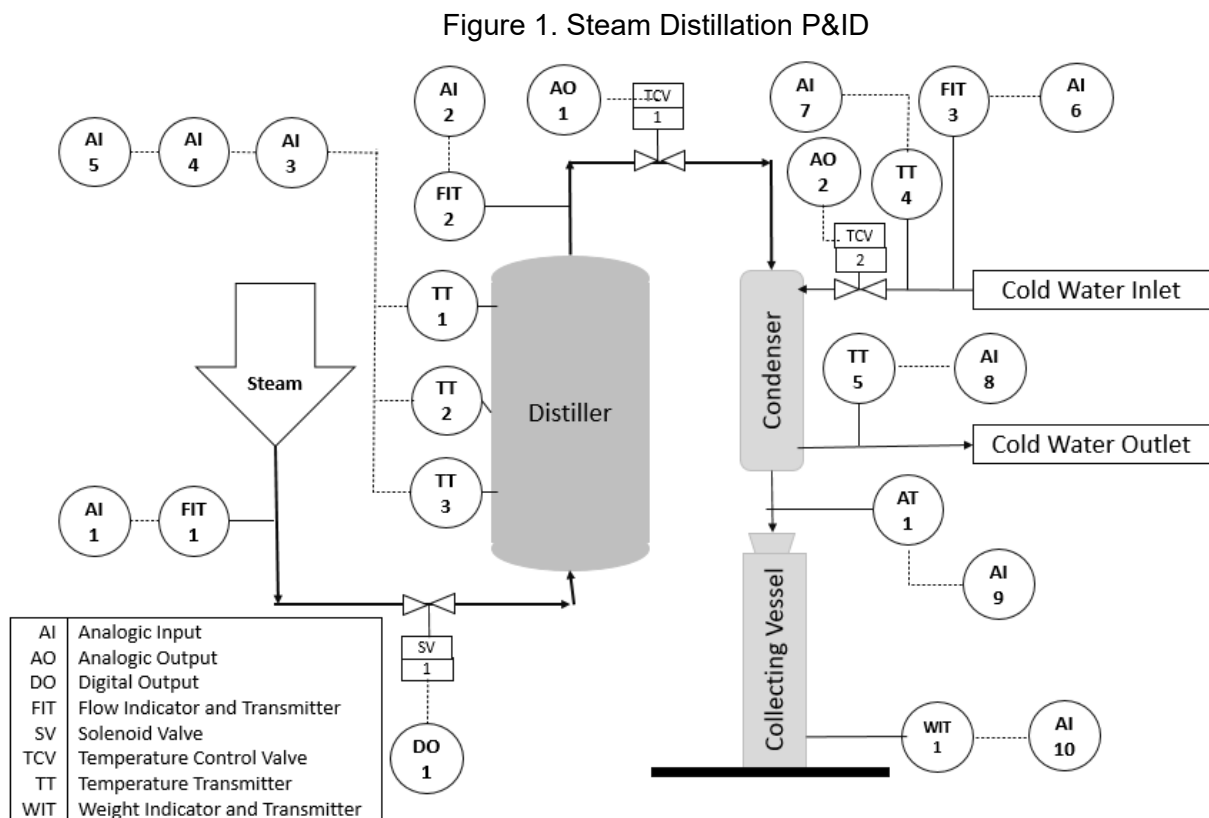
Figure 1. Steam Distillation P&ID

Figure 1 presents all the sensors and actuators to monitor and control the plant. The distiller used to represent in the digital twin has a volume of 0,05 m^3 and can be filled with a maximum of 10 kg of plant material. The steam used during the distillation

will be provided at a temperature a little bit higher than 373,15 K by a boiler. The distiller already has a condenser connected and the collecting vessel is responsible for the separation of the oil from the hydrosol. All the data generated by the sensors are stored and used to create a data-based model for the digital twin.

Temperature and steam flow are the most important measured properties in the process, which justifies the presence of some sensors alongside the pipes to monitor them. The solenoid valve is responsible for the amount of steam that gets inside the distiller. The temperature control valve helps controlling the temperature inside the distiller and the condenser cold water inlet. The weight sensor is responsible for measuring the amount of condensed matter that has already gotten out of the condenser. Inside of the distiller. The heat may find an easier way to flow from the bottom to the top, which is undesirable, since the heat might burn some of the material inside of the distiller while it might not pass through another part of it, reducing the amount of oil extracted. The steam should pass through all the botanical material equally. To identify these heat flow channels, several specific temperature sensors are placed inside the distiller: on the bottom, the middle and the top of it, represented in the Figure 2 as TT 1, TT 2 and TT 3.

In order to create the model to the digital twin, the chosen software is the Matlab from Mathworks, since Matlab has all the functions needed to build the digital twin avoiding connectivity problems between different software. The goal is to achieve yield increase and optimization by applying technology to the system. Figure 2 shows the structure of the digital twin to help understanding its creation.

Figure 2. Digital Twin Construction Flow
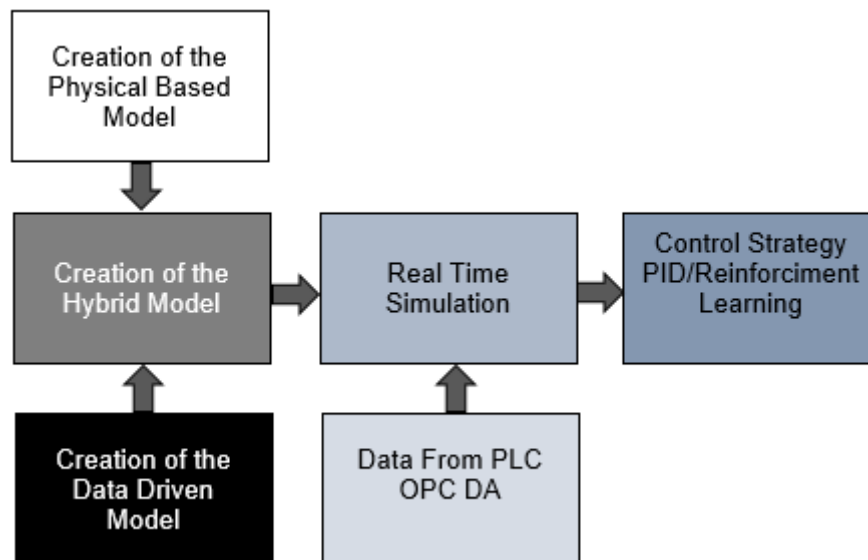
Initially, a white box model is created with Simulink in the Matlab. There are two steps considered for this. The first one is a mathematical model using the main Simulink library based on well-known physics laws, such as Fick's law, that describes the mass diffusion, and Darcy's law, that describes the flow of a fluid through a porous media. Both are already used to describe the phenomenon inside the distiller [14, 15]. The second is to use the Simscape library, from Simulink, to build a model based on

physical connections of blocks, each block representing a part of the pilot plant, facilitating the construction of the physical model.

To build the black box model, the Statistics and Machine Learning Toolbox is used to build and train the model based on all the data from the pilot plant. Several algorithms are available to be used on Matlab for this purpose. All the data and the model created in the Matlab can be integrated to the white box model and a Matlab function block is added to Simulink in order to create the grey box model.

The Simulink Desktop Real-Time is the tool used to do the real time simulation. With the Simulink model and the real-time data coming from the PLC (Programmable Logic Controller) the simulation runs in real time. The connection between the PLC and Matlab is stablished using an OPC DA (Open Platform Communications Data Access) connection. The PLC available for this study is the Rockwell Micro850, a largely used solution in the industry.

The last step in order to create the digital twin is to design a controller to the plant to guarantee that the results from the simulations can be implemented through the actuators changing the current state of the physical asset. The pilot plant will be working initially with a PID (Proportional Integral Derivative) controller. The goal is to add reinforcement learning to the control strategy creating a hybrid controller. An adaptative PID controller with reinforcement learning, when applied to nonlinear systems, may be a robust option to control the plant that adapts very well to changing parameters [16].

3. RESULTS AND DISCUSSION

A grey box model combines the ease of creating a black box model from data and the physical principles already known about a system, from a white box. This modeling method is better for generalization purposes than the black box and doesn't require the same amount of data used to create and train a black box model [17]. Grey box models benefits', compared to the black box models, are related to the physical knowledge within the grey box. Compared to the white box modeling, the benefits are the fact that grey box models can avoid approximations, usually common while creating a white box model [18]. The idea of using a grey box approach is very interesting, since the data driven and the physically based modeling can be combined to develop a better model.

Matlab has lots of toolboxes that can add functions to the digital twin, such as optimization, model predictive control and CFD (Computational Fluid Dynamics), which weren't explored in this study. However, these tools can be implemented later to improve the model. The essential oil extraction processes lack technology applied to them and the energy consumption is an issue for this industry, compromising its profitability. The introduction of technology focusing on the optimization may bring good benefits [19] and Matlab, combining data to create models and simulations, can be helpful in this case.

This digital twin is being crested as a case study for a greater goal, the final objective is to develop a digital twin creation method. It should be modular, scalable and generalizable, so some questions should be answered to achieve it. The connection between the PLC and the software is one of them. The OPC DA is not the

best option in terms of connectivity, while OPC UA (Open Platform Communication Unified Architecture), a more recent and open version, could be a solution as it is a better bridge to the data, since it was created to serve as a way to connect many types of hardware and software [20].

4. CONCLUSION

Digital twins are one of the main technologies from Industry 4.0 and its implementation may bring lots of benefits to an industrial process. The optimization capabilities, the possibility of saving resources, the production control and the data analysis are some of the reasons the digital twins must be more discussed, creating better solutions for the different types of industries. The proposed digital twin focuses on the improvement of the steam distillation extraction of essential oil process proposing a grey box modeling, a type of modeling that is being used more and more nowadays. The next step is to implement the concept explored in this paper in order to prove its capabilities.

In the future, the creation method explored in this paper can be the solution to develop digital twins for various processes. For this purpose, there are some questions that should be addressed in order to create this groundwork for digital twins.

Acknowledgments

5. REFERENCES

[1] KAGERMANN, Henning; WAHLSTER, Wolfgang; HELBIG, Johannes. **Recommendations for implementing the strategic initiative Industrie 4.0:** Final report of the Industrie 4.0 Working Group. Forschungsunion: Berlin, 2013.

[2] VACHÁLEK, Ján et al. The digital twin of an industrial production line within the industry 4.0 concept. In: **2017 21st International Conference on Process Control (PC)**. IEEE, 2017. p. 258-262.

[3] GRIEVES, Michael; VICKERS, John. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: **Transdisciplinary perspectives on complex systems**. Springer, Cham, 2017. p. 85-113.

[4] TAO, Fei et al. Digital twin-driven product design, manufacturing and service with big data. **The International Journal of Advanced Manufacturing Technology**, v. 94, n. 9-12, p. 3563-3576, 2018.

[5] ABREU, Cleyde Evangelista Maia et al. Indústria 4.0: Como as Empresas Estão Utilizando a Simulação para se preparar para o Futuro. **Revista de Ciências Exatas e Tecnologia**, v. 12, n. 12, p. 49-53, 2018.

[6] GLAESSGEN, Edward; STARGEL, David. The digital twin paradigm for future NASA and US Air Force vehicles. In: **53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA**. 2012. p. 1818.

[7] PARROTT, Aaron; WARSHAW, Lane. Industry 4.0 and the digital twin. **Deloitte University Press**, pp. 1-17, 2017.

[8] KRITZINGER, Werner et al. Digital Twin in manufacturing: A categorical literature review and classification. **IFAC-PapersOnLine**, v. 51, n. 11, p. 1016-1022, 2018.

[9] ROSEN, Roland et al. About the importance of autonomy and digital twins for the future of manufacturing. **IFAC-PapersOnLine**, v. 48, n. 3, p. 567-572, 2015.

[10] SOHLBERG, Björn; JACOBSEN, Elling W. Grey Box modelling–branches and experiences. **IFAC Proceedings Volumes**, v. 41, n. 2, p. 11415-11420, 2008.

[11] ROMIJN, Reinout et al. A grey-box modeling approach for the reduction of nonlinear systems. **Journal of Process Control**, v. 18, n. 9, p. 906-914, 2008.

[12] CRAVEIRO, Afranio Aragao; QUEIROZ, D. C. Óleos essenciais e química fina. **Química nova**, v. 16, n. 3, p. 224-228, 1993.

[13] KASUAN, Nurhani et al. Essential oil composition of Kaffir lime: Comparative analysis between controlled steam distillation and hydrodistillation extraction process. In: **2009 IEEE Student Conference on Research and Development (SCOReD)**. IEEE, 2009. p. 479-482.

[14] CASSEL, E. et al. Steam distillation modeling for essential oil extraction process. **Industrial crops and products**, v. 29, n. 1, p. 171-176, 2009.

[15] BERKA-ZOUGALI, Baya et al. Extraction of essential oils and volatile molecules. In: **Instant Controlled Pressure Drop (DIC) in Food Processing**. Springer, New York, NY, 2014. p. 97-126.

[16] WANG, Xue-Song; CHENG, Yu-Hu; WEI, Sun. A proposal of adaptive PID controller based on reinforcement learning. **Journal of China University of Mining and Technology**, v. 17, n. 1, p. 40-44, 2007.

[17] ARAHAL, Manuel R.; CIRRE, Cristina M.; BERENGUEL, Manuel. Serial grey-box model of a stratified thermal tank for hierarchical control of a solar plant. **Solar Energy**, v. 82, n. 5, p. 441-451, 2008.

[18] KRISTENSEN, Niels Rode; MADSEN, Henrik; JØRGENSEN, Sten Bay. Parameter estimation in stochastic grey-box models. **Automatica**, v. 40, n. 2, p. 225-237, 2004.

[19] CHEMAT, Farid et al. Green extraction of natural products. Origins, current status, and future challenges. **TrAC Trends in Analytical Chemistry**, v. 118, p. 248-263, 2019.

[20] BRAUNE, Annerose; HENNIG, Stefan; HEGLER, Sebastian. Evaluation of OPC UA secure communication in web browser applications. In: **2008 6th IEEE International Conference on Industrial Informatics.** IEEE, 2008. p. 1660-1665.