A Model Proposal For Digital Twin Development: Essential Oil Extraction Perspective

Matheus Antonio Nogueira de Andrade

UFBA: Universidade Federal da Bahia

Herman Augusto Lepikson

SENAI CIMATEC Technology College: Faculdade de Tecnologia SENAI CIMATEC

Carlos Alberto Tosta Machado (✉ carlos.tosta@uol.com.br)

SENAI CIMATEC Technology College: Faculdade de Tecnologia SENAI CIMATEC

<https://orcid.org/0000-0002-6773-5314>

Case Report

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

Posted Date: March 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-300217/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Dear editors,

The present article is part of the special volume **“Post Covid 19 challenges: AI technology and innovation (SIINTEC VI)”**.

Best Regards,

The authors:

Matheus Antônio Nogueira de Andrade

Herman Augusto Lepikson

Carlos Alberto Tosta Machado

A MODEL PROPOSAL FOR DIGITAL TWIN DEVELOPMENT: ESSENTIAL OIL EXTRACTION PERSPECTIVE

Matheus Antônio Nogueira de Andrade ^a, Herman Augusto Lepikson ^b, Carlos Alberto Tosta Machado ^c

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

Abstract:

- **Introduction:** 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.
- **Case Description:** 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 proportional integral derivative control, focusing on the yield increase and optimization. Computer Vision and Artificial Intelligence enhancements were suggested.
- **Discussion and Evaluation:** Digital twins, in combination with Artificial Intelligence can be of great help to support companies with the decision-making challenges. Furthermore, some benefits that Artificial Intelligence can bring to the process were enlightened. Computer Vision approaches were also discussed.
- **Conclusions:** A creation method is elaborated to support other applications of digital twins in industrial processes in the future. In order to apply it to different processes, generalization capabilities must be proved.

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

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 [14].

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 [29]. A digital twin can also be interpreted as a group of virtual information that describes a physical object [12]. 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 [27]. 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 [1].

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 [11]. 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, being composed of a physical part, a digital part and the connection between them through which the data flows [12]. More recently, NASA has begun to use digital twins to develop new solutions motivated by the results they achieved with the previous twins [11]. 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 [12]. 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 [22].

An important aspect which should be considered while discussing about digital twins is the level of integration between the physical and the digital worlds [18]. The digital twins can be classified in three subcategories according to the levels of integration as follows:

- **Model Twin:** a digital representation of a physical object which does not use any type of automated data exchange with the physical counterpart.
- **Shadow Twin:** a step further in comparison with the model twin, it has a way of exchanging data automatically with the physical object, but the data only goes in one direction, meaning that a change of state in the physical object represents a change of state in the digital object, but not vice versa.
- **Digital Twin:** the higher level of integration, the data exchange happens in both directions, changes of states noted by the sensors in the physical object are reflected in the digital object. The digital object is also capable of changing the state of the physical object through the actuators changing the state of the physical object.

In the literature it is common to find different digital models and shadows being called by twins contributing to the lack of a common definition. Shadow and model twins do not have the third element proposed in 2002 when the term was coined. It is evident that digital twins

are in its infancy and it is still necessary to reach a higher level of maturity in order to create a proper definition. [10, 18, 31].

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) [18]. 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 [25].

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 [22].

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 [26]. There is a third type of model, also referred as hybrid model or grey box [24]. 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 doesn't require the same amount of data used to create and train a black box model [3]. 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 [17]. 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.

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 [9]. 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 [16]. Despite of the existence of other production processes for essential oils, like solvent extraction and critical fluid extraction, the steam distillation process is the most used for essential oil extraction [23], representing 93% of the oil extracted worldwide [20]. This study is proposing a grey box modeling approach for the system using machine learning and a physical modeling to achieve this goal.

The Physical Asset

A small pilot plant 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. Figure 1 shows the P&ID (Piping and Instrumentation Diagram) of the process, with all its sensors and actuators used to monitor and control the plant.

Figure 1. Steam Distillation P&ID

The distiller has a volume of $0,05 \text{ m}^3$ and can be filled with a maximum of 10 kg of plant material. Before the process starts, the bottom of the distiller is filled with water so the electrical heating element will be immersed in water avoiding overheating problems which could make it burn out. A low-level switch is placed on the bottom of the distiller to make sure it has the appropriate amount of water inside of it, if the switch detects the water is dropping, the level control valve opens and the water flows into the bottom of the distiller. A flow indicator, represented by FIT 1 in the Figure 1, is placed before the valve indicating the water flow rate.

The heating element has 4 kW of power. When the water reaches the lower level accepted, the element starts heating the water so that it turns into steam and passes through the plant material inside 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. This phenomenon is called channeling. The steam should pass through all the botanical material equally. To identify these heat flow channels, several temperature sensors are placed inside the distiller: on the bottom, the middle and the top of it, represented in the Figure 1 as TT 1, TT 2, TT 3, TT 4, TT 5 and TT 6.

At the top of the distiller, there is a stepper motor connected to a structure of aluminum rods moves the plant material inside the distiller if any channeling is detected. Once the steam extracts the essential oil from the material, then both pass through the condenser transferring the latent heat to the cold water that is being injected by the pump. A flow indicator is placed in the condenser outlet to measure how much condensed hydrosol and essential oil are being produced. The amount of hydrosol is much greater than that of essential oil.

The temperature transmitters TT 7 and TT 8 are responsible for indicating the temperature of the water when it is getting in and out of the condenser. The flow indicator FIT 2 measures the water flow rate at the pump outlet. The water that once got out of the cold water

tank and passed through the condenser returns to the tank. The chiller is used to lower the water temperature. The switch represented by LSL 2, indicates if the water level is low and allows water to enter the tank. The hydrosol collector is filled during the process and the essential oil tube is monitored by an image processing algorithm.

The Computer Vision Approach

The image processing system is composed by a Raspberry Pi 3 Model B+, a small size single board computer, with its camera called Camera Module V2. The algorithm was created using a library called OpenCV [21] from the Python programming language. The image process system is the responsible for monitoring the amount of essential oil produced, in real time, while it leaves the condenser separates from the hydrosol. Color detection and finding contours functions were applied in the algorithm.

The essential oil color is selected by trackbars and the algorithm search for the right color inside the image. The oil contours are found within the image because of its color and the number of pixels with the oil's color can be converted to real measurements by a scale. Essential oils are commonly used in perfumes, aromatherapy and by the pharmaceutical industry. Therefore, the image processing system is a good approach to measure the essential oil production because of its non-invasive application.

The Digital Twin Creation Process Proposed

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 [4, 6]. 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 as a Matlab function block is added to Simulink in order to create the grey box model. Some regression algorithms are widely applied to model batch processes Partial Least Squares (PLS), Principal Component Analysis (PCA) and their variants are some of the most used ones [19]. The introduction of data-driven, analytical, and knowledge-based approaches enable better levels of safety, quality and reliability responses with undeniable effects on cost optimization and capacity utilization. Given the relevance of essential oils to important markets like cosmetic and pharmaceutical, the manufacturing excellence must be encompassed with business overall competitiveness. The proposed digital twin can be the starting point for the creation of several others being one of the first digital twin in the entire pharmaceutical industry [8].

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, the simulation runs in real time. The connection between the PLC and Matlab is established using an OPC UA (Open Platform Communications Unified Architecture) connection. It was created with the aim of establishing a connection between the most diverse types of computers and through several communication protocols [5]. An application that uses OPC UA can be configured to be as complex as necessary and therefore enables and contributes to the development of IoT applications [13].

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 adaptive 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 [30].

Discussion and Evaluation

Matlab has toolboxes that can add functions to the digital twin, such as optimization, model predictive control and CFD (Computational Fluid Dynamics) and it can be implemented 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 [7] and the used approach, combining data to create models and simulations, can be helpful.

The camera is providing to the system the final process output, the oil column measurement. AI can be applied, in conjunction with the image processing algorithm, to be a valuable source of information to the digital twin. CV (Computer Vision) and AI already proved

their capabilities when applied together to other industry sectors [15]. The color of the oil can be correlated to what is the raw material inside the distiller and can be used to detect undesirable changes providing additional information about the process dynamics. The quality of the raw material or the conditions in which the plant was harvested is also another correlation to be explored. Another possibility for the AI application is the automatic color and oil detection in different scenarios facilitating the implementation of the same code to various different places.

Operators, in their daily duties face complex challenges to monitor and control process variables. Often, when these actions are performed manually, they are not effective and fast as necessary. Both the diagnosis and correction of undesired occurrences depend solely on empiricism and manual actions, with the risk of misjudgment and slowness. This complex multivariable panorama creates the proper field for AI application [28] and digital twins can be of a very good help to support the introduction of AI in the process.

The application of automatic learning and pattern recognition bring benefits that go far beyond the purely operational improvements: the business bottom line results. The search for promising correlations among process variables, the estimation of results and trends, the decision-making algorithms, among others, are the AI tools that can positively contribute to business key performance indicators enhancement [2]. This research brings an insight of the potential presented by this approach.

Conclusions

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 grey box modeling approach for digital twins focuses on the improvement of production process that may benefit from sensor fusion and image data analysis based on AI technics. In addition, CV algorithms fused with AI can contribute a lot to improve the process.

The next step is to implement the concept explored in this paper in order to prove other capabilities. In the future, the creation method contemplated by this work can be the solution to develop digital twins for various processes. For this purpose, further research will address the groundwork for modular, scalable and generalizable method for digital twins' creation in order to enhance its utilization could make the difference in this scenario.

List of Abbreviations

AI – Artificial Intelligence

CFD – Computational Fluid Dynamics

CV – Computer Vision

IoT – Internet of Things

IT – Information Technology

NASA – National Aeronautics and Space Administration

OPC UA – Open Platform Communications Unified Architecture

P&ID – Piping Instrumentation Diagram

PCA – Principal Component Analysis

PID – Proportional Integral Derivative

PLC – Programmable Logic Controller

PLM – Product Lifecycle Management

PLS – Partial Least Squares

PPC – Production Planning and Control

Declarations

- **Ethical Approval and Consent to Participate**

Not applicable.

- **Consent for Publication**

All authors authorize the publication of the article.

- **Availability of Supporting Data**

Not applicable.

- **Competing Interests**

The authors declare that they have no competing interests.

- **Funding**

The present work was carried out with the support of the Improvement Coordination of Higher Education Personnel - Brazil (CAPES), process number 88882.452862/2019-01.

- **Authors' contributions**

MANA wrote the major part of the article and designed the proposed digital twin model. HAL provided a solid base for the digital twin concept understanding and also wrote part of the article. CATM contributed with all the knowledge about essential oil extraction and contributed to the writing as well.

- **Acknowledgments**

Authors thank the laboratorial support from Senai Cimatec, in Salvador. Authors also thank the company from the essential oil extraction Linax that donated an essential oil distiller.

- **Authors' Information**

MANA is a mechanical engineer and master's degree student working with the creation of digital twins of industrial processes. HAL have a PhD in mechanical engineering and

is a researcher in the smart manufacturing context. CATM is a mechanical engineer and PhD student currently working with technology applied to essential oil extraction processes.

References

- ¹ Abreu CEM, Gonzaga DRB, dos Santos FJ, de Oliveira, JF, Oliveira KDM, Figueiredo LM, Gonçalves, GAS (2017) Indústria 4.0: Como as empresas estão utilizando a simulação para se preparar para o futuro. *Revista de Ciências Exatas e Tecnologia*, 12(12), 49-53.
- ² Ali JM, Hussain MA, Tade MO, Zhang J (2015) Artificial Intelligence techniques applied as estimator in chemical process systems—A literature survey. *Expert Systems with Applications*, 42(14), 5915-5931.
- ³ Arahal MR, Cirre CM, Berenguel M (2008) Serial grey-box model of a stratified thermal tank for hierarchical control of a solar plant. *Solar Energy*, 82(5), 441-451.
- ⁴ Berka-Zougali B, Besombes C, Allaf T, Allaf K (2014) Extraction of essential oils and volatile molecules. In *Instant Controlled Pressure Drop (DIC) in Food Processing* (pp. 97-126). Springer, New York, NY.
- ⁵ Braune A, Hennig S, Hegler S (2008) Evaluation of OPC UA secure communication in web browser applications. In *2008 6th IEEE International Conference on Industrial Informatics* (pp. 1660-1665). IEEE.
- ⁶ Cassel E, Vargas RMF, Martinez N, Lorenzo D, Dellacassa E (2009) Steam distillation modeling for essential oil extraction process. *Industrial crops and products*, 29(1), 171-176.
- ⁷ Chemat F, Abert-Vian M, Fabiano-Tixier AS, Strube J, Uhlenbrock L, Gunjevic V, Cravotto G (2019) Green extraction of natural products. Origins, current status, and future challenges. *TrAC Trends in Analytical Chemistry*, 118, 248-263.
- ⁸ Chen Y, Yang O, Sampat C, Bhalode P, Ramachandran R, Ierapetritou M (2020) Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review. *Processes*, 8(9), 1088.
- ⁹ Craveiro AA, Queiroz DC (1993) Óleos essenciais e química fina. *Química nova*, 16(3), 224-228.
- ¹⁰ Fuller A, Fan Z, Day C, Barlow C (2020) Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8, 108952-108971.
- ¹¹ Glaessgen E, Stargel D (2012) 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* (p. 1818).
- ¹² Grieves M., Vickers J (2017) Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems* (pp. 85-113). Springer, Cham.
- ¹³ Imtiaz J, Jasperneite J (2013) Scalability of OPC-UA down to the chip level enables “Internet of Things”. In *2013 11th IEEE International Conference on Industrial Informatics (INDIN)* (pp. 500-505). IEEE.
- ¹⁴ Kagermann H, Wahlster W, Helbig J (2013) Recommendations for implementing the strategic initiative Industrie 4.0: Final report of the Industrie 4.0 Working Group. *Forschungsunion: Berlin, Germany*.

- ¹⁵ Kakani V, Nguyen VH, Kumar BP, Kim H, Pasupuleti VR (2020) A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033.
- ¹⁶ Kasuan N, Yunus M, Rahiman MHF, Aris SRS, Taib MN (2009) 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) (pp. 479-482). IEEE.
- ¹⁷ Kristensen NR, Madsen H, Jørgensen SB (2004) Parameter estimation in stochastic grey-box models. *Automatica*, 40(2), 225-237.
- ¹⁸ Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016-1022.
- ¹⁹ Marjanovic O, Lennox B, Sandoz D, Smith K, Crofts M (2006) Real-time monitoring of an industrial batch process. *Computers & chemical engineering*, 30(10-12), 1476-1481.
- ²⁰ Masango P (2005) Cleaner production of essential oils by steam distillation. *Journal of Cleaner Production*, 13(8), 833-839.
- ²¹ OpenCV. <https://opencv.org/about/>. Accessed 24 Feb 2021.
- ²² Parrott A, Warshaw L (2017) *Industry 4.0 and the digital twin*. Deloitte University Press, pp. 1-17.
- ²³ Reverchon E, Senatore F (1992) Isolation of rosemary oil: comparison between hydrodistillation and supercritical CO₂ extraction. *Flavour and fragrance journal*, 7(4), 227-230.
- ²⁴ Romijn R, Özkan L, Weiland S, Ludlage J, Marquardt W (2008) A grey-box modeling approach for the reduction of nonlinear systems. *Journal of Process Control*, 18(9), 906-914.
- ²⁵ Rosen R, Von Wichert G, Lo G, Bettenhausen KD (2015) About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3), 567-572.
- ²⁶ Sohlberg B, Jacobsen EW (2008) Grey Box modelling—branches and experiences. *IFAC Proceedings Volumes*, 41(2), 11415-11420.
- ²⁷ Tao F, Cheng J, Qi Q, Zhang M, Zhang H, Sui F (2018) Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, 94(9), 3563-3576.
- ²⁸ Uraikul V, Chan CW, Tontiwachwuthikul P (2007) Artificial intelligence for monitoring and supervisory control of process systems. *Engineering applications of artificial intelligence*, 20(2), 115-131.
- ²⁹ Vachálek J, Bartalský L, Rovný O, Šišmišová D, Morháč M, Lokšík M (2017) The digital twin of an industrial production line within the industry 4.0 concept. In 2017 21st international conference on process control (PC) (pp. 258-262). IEEE.
- ³⁰ Wang XS, Cheng YH, Wei S (2007) A proposal of adaptive PID controller based on reinforcement learning. *Journal of China University of Mining and Technology*, 17(1), 40-44.
- ³¹ Zhang H, Liu Q, Chen X, Zhang D, Leng J (2017) A digital twin-based approach for designing and decoupling of hollow glass production line. *IEEE Access*, 5, 26901-26911.

Figures

Figure 1

Steam Distillation P&ID

Figure 2

Digital Twin Construction Flow