BUILD-IT2023 Workshop

Book of Abstracts

Andrea Serani and Cecilia Leotardi

BUILding a Digital Twin: requirements, methods, and applications

BUILD-IT 2023 WORKSHOP

19-20 October, Rome, Italy

ISBN: 978-88-7617-056-0 (e-book)
October 2023
Editors: Andrea Serani and Cecilia Leotardi
Publisher: National Research Council-Institute of Marine Engineering (CNR-INM)

ISTITUTO DI INGEGNERIA DEL MARE
INSTITUTE OF MARINE ENGINEERING

Il Centenario del CNR è realizzato con il contributo della Presidenza del Consiglio dei Ministri e con il Patrocinio di Rai

PREFACE

Digital twins are virtual representations of physical systems or processes that are used for modeling, simulation, and control. They enable a wide range of applications, from product design and optimization to predictive maintenance and smart cities. With the increasing availability of data and computational resources, digital twins are becoming more powerful and versatile, and are transforming many industries.

A digital twin must accurately represent the physical object, system, or process it is modeling. This requires gathering data from sensors, machines, and other sources in real-time and integrating it into a cohesive digital model. They have to monitor and analyze real-time data to provide insights and predict future behavior. This requires advanced analytics and machine learning algorithms that can process large volumes of data and identify patterns and anomalies. Finally, a digital twin should be designed to be interoperable with other systems and platforms. This allows data to be shared and integrated with other systems and tools, enabling more efficient workflows and better decision-making.

This Book of Abstracts collects the experiences and ideas gathered during the BUILD-IT workshop, which brought together researchers and practitioners working on digital twins for product design and manufacturing, product and process life cycles, health monitoring, predictive maintenance and fault diagnosis, product and process energy efficiency, urban planning and management, citizen-centered service design and optimization, connected and automated mobility, smart industries and infrastructure, communication network management, cultural heritage, healthcare and biomedicine, environment, biology, ecology, and ecosystems, fostering discussions on digital twin challenges and limitations.

SCIENTIFIC COMMITTEE

Andrea Ballarino	CNR-STIIMA
Gabriella Bretti	CNR-IAC
Raffaele Bruno	CNR-IIT
Mario Ciampi	CNR-ICAR
Giampaolo Coro	CNR-ISTI
Ugo Cortesi	CNR-IFAC
Michela Costa	CNR-STEMS
Ludovico Danza	CNR-IRC
Riccardo De Benedictis	CNR-ISTC
Matteo Diez	CNR-INM
Barbara Masini	CNR-IEIIT
Michela Mortara	CNR-IMATI
Alessandro Oggioni	CNR-IREA
Daniele Pellegrini	CNR-ISTI
Marco Sacco	CNR-STIIMA
Andrea Serani	CNR-INM
Giuseppe Stecca	CNR-IASI
Lorenzo Tamellini	CNR-IMATI

ORGANIZING COMMITTEE

Cecilia Leotardi	CNR-INM
Andrea Serani	CNR-INM

TABLE OF CONTENTS

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

ESSENCE OF DIGITAL TWINS

Laura Mainini
Politecnico di Torino, Torino, Italy
e-mail: laura.mainini@polito.it

Matteo Diez
CNR-INM, Consiglio Nazionale delle Ricerche–Istituto di Ingegneria del Mare, Rome, Italy
e-mail: matteo.diez@cnr.it

This extended abstract discusses the mathematical aspects that underpin the concept (and the essence) of digital twins as purpose-driven adaptive models (or virtual representations) and that are essential to their formalization for applications to science and engineering including product design, development, maintenance, and operations.

Keywords: *digital twins, systems design and development, systems maintenance, operations*

1. Introduction

Digital twins are widely considered as enablers of groundbreaking changes in the development, operation, and maintenance of novel generation of products. In this context, "a Digital Twin is a set of virtual information constructs that mimics the structure, context, and behaviour of an individual/unique physical asset, is dynamically updated with data from its physical twin throughout its lifecycle, and informs decisions that realize value" [1].

While there is a popular—frequently misleading—understanding of digital twins as the highest-fidelity virtual representation of all aspects of the real system of interest, digital twins are rather purpose-driven virtual representations. Therefore, those digital counterparts of an individual real artifact are not unique but would assume different forms depending on the purpose [2]. At the same time, all digital twins share a key property that distinguish them from the parent family of digital models: digital twins are physics-based and adaptive in nature, and are conceived to continuously learn from data. As such, digital twins are inevitably characterized by a mathematical soul to combine data streams and physics-based representations in a principled and efficient way.

Methods are rooted at the intersection of scientific computing and machine learning, and span the world of multi-fidelity and multi-source information fusion or calibration, data assimilation, surrogate and reduced order modelling, uncertainty quantification, optimal data selection and acquisition. Major research open challenges relate to the rapidity and the reliability of responses and predictions from the digital representations.

This extended abstract discusses the mathematical aspects that underpin the concept (and essence) of digital twins as purpose-driven adaptive models (or virtual representations) and that are essential to their formalization for applications to science and engineering including product design, development, maintenance, and operations. Based on the definitions of digital twin proposed by international communities over the years, this work aims at providing an overview of the mathematical formulations and computational methods associated to digital twinning. Considerations will also be drawn upon the perspectives of a multidisciplinary and cross-domain forum to identify the main challenges met to enable responsive, predictive and reliable digital twins, and the associated research avenues.

2. Purpose-driven adaptive models

A digital twin is a virtual model of a system or process that progressively adapts and specializes by learning from data from the real counterpart. The value introduced by the availability and adoption of digital twins relates to their usefulness—all models are wrong, some are useful [3]—in providing reliable and timely predictions to inform decisions along the entire product life cycle. The time- and resource-efficiency requirements impose digital twins of a given physical system or process not to be unique, but rather multifaceted and purpose-driven adaptive models, because there is not such a *twin that rules them all*. The principle of model usefulness discards the definition of digital twin as a "high-fidelity model of the system which can be used to emulate the actual system" [4] in favour of a synthesis of "the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its corresponding physical twin" [5].

These features require digital twins to be endowed with fine mathematical souls tasked to formulate, compute and update the virtual models as dynamic syntheses of physics formalizations and data streams to inform decisions with reliable predictions. In particular, data assimilation methods are pivotal to the realization of models that continuously morph by learning from data. Whether the assimilation is achieved via calibration or fusion approaches, computational methods for multisource information synthesis are essential to realize digital twins whose data sources can be diverse – sensor measurements, signal acquisitions, experimental databases, (*50-shades-of-grey*) models evaluations, physics based simulations. Among those, multifidelity methods also acknowledge that multiple representations of given physical systems and processes are possible at different levels of accuracy and costs, which offer tremendous opportunities to play across multiple levels of abstractions to maximize the usefulness with respect to the specific decision tasks to support [6].

The same usefulness rationale motivates the role of (non-intrusive) surrogate modelling for digital twins, and the specific need of formulations that allow to learn models from a limited amount of observations (small data) and somehow characterize the reliability of the estimated predictions. This requirement poses major limitations to the straight-forward use of fashionable deep learning and purely data-driven methods—which are data intensive and of questionable reliability—and demands for advanced approaches that could ideally embed physical constraints in the learning process. Examples include formulations for data-driven operator inference [7], projection based model reduction for physics-based machine learning [8], physics informed neural networks [9], and domain aware active learning [10].

In addition, the development and adoption of digital twins for platforms and systems is often affected by the challenges associated with the high-regret scenarios faced over the operational life, where badly informed decisions can lead to catastrophic consequences. This demands mathematical methods to characterize the reliability of the predictions provided by the digital twins in support of the decisional processes. On the one hand, computational methods for uncertainty quantification, characterization and propagation [11] are essential to approach these open questions and equip the predictions with forms of reliability measures or robustness bounds [12]. On the other hand, research efforts demonstrated the importance of the quality of source data over their quantity [13] to improve reliability and robustness of the predictions, which motivates the recommendation to pay larger attention to the field of optimal sensor placement, and optimal data selection and acquisition for purpose-driven digital twins [14].

As the digital twins ideally evolve (or degrade) together with the real/physical system along its life cycle, any mandate to represent the as-built system is intrinsically relaxed. Moreover, whether or not a digital twin could exist prior to (without) the corresponding physical twin is still open debate. Indeed, the twins are ideally continuously informed and enabled through the digital thread that links all the stages of a system life [15] with forward and backward feeds: limiting their existence to specific phases would introduce ontological and taxonomic inconsistencies.

3. System design, development, maintenance, and operations

This section briefly outlines how the digital twins have been/can be used to support decisions at design, manufacturing, operational, maintenance stages, even when the physical system does not yet exist.

Digital twins for systems design and development. Digital twins are used in system design and development to simulate, test, and refine/optimize new products or processes [16]. Digital twins are used to explore design spaces and advance the development of products. In this phase the digital twin lacks its physical counterpart, nevertheless it shall include all the relevant feature of its physical twin once the latter will be brought to life. The purpose of the digital twin could be to predict the product performance once in operation (design for performance) or to assess its manufacturability and prediction costs (design for manufacturability). Depending on the application, these can be achieved by a product digital shadow (focusing on the mathematical modeling of the relevant physical attributes) or digital replica (an automatic projection of the system construction), respectively [17].

Digital twins for systems maintenance. Maintenance practices benefit from digital twinning by using virtual representations of physical assets and systems, which integrate real-time data from sensors into the relevant mathematical models. By providing real-time, data-driven models of asset state and performance, digital twins enable predictive [18] and prescriptive maintenance [19], preventing downtime and reducing maintenance costs. Digital twins may be used to optimize maintenance schedules, troubleshoot issues, and develop more efficient maintenance procedures. Digital twins can be used to explore different maintenance scenarios with the aim of finding the most effective solutions before application in the physical world.

Digital twins for operations. In the context of optimizing operations, digital twins can play an essential role by providing insights into how assets and systems are operating in real-time, allowing for proactive decision making [20]. The latter can be extended to model-predictive control by incorporating system state forecasts provided by digital twins [21]. Moreover, digital twins enable virtual testing in different operating scenarios, reducing the need for physical testing, the risk of damaging the physical asset, and the testing overall cost. Environment digital twins are a research frontier for a fully integrated digital twin including the asset and the environment of operations [22].

REFERENCES

1. Committee, A. D. E. I., et al. Digital twin: Definition & value, *AIAA and AIA Position Paper*, (2020).
2. Rasheed, A., San, O. and Kvamsdal, T. Digital twin: Values, challenges and enablers from a modeling perspective, *Ieee Access*, **8**, 21980–22012, (2020).
3. Box, G. All models are wrong, but some are useful, *Robustness in Statistics*, **202** (1979), 549, (1979).
4. Board, S. E., (2022), *The Guide to the Systems Engineering Body of Knowledge (SEBoK), v. 2.7*.
5. West, T. D. and Blackburn, M. Is digital thread/digital twin affordable? a systemic assessment of the cost of dod's latest manhattan project, *Procedia computer science*, **114**, 47–56, (2017).
6. Beran, P. S., Bryson, D., Thelen, A. S., Diez, M. and Serani, A. Comparison of multi-fidelity approaches for military vehicle design, *AIAA AVIATION 2020 FORUM*, p. 3158, (2020).
7. Peherstorfer, B. and Willcox, K. Data-driven operator inference for nonintrusive projection-based model reduction, *Computer Methods in Applied Mechanics and Engineering*, **306**, 196–215, (2016).
8. Swischuk, R., Mainini, L., Peherstorfer, B. and Willcox, K. Projection-based model reduction: Formulations for physics-based machine learning, *Computers & Fluids*, **179**, 704–717, (2019).
9. Raissi, M., Perdikaris, P. and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational physics*, **378**, 686–707, (2019).
10. Di Fiore, F., Maggiore, P. and Mainini, L. Multifidelity domain-aware learning for the design of re-entry vehicles, *Structural and Multidisciplinary Optimization*, **64**, 3017–3035, (2021).
11. Le Maître, O. and Knio, O. M., *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*, Springer Science & Business Media (2010).

12. Yao, W., Chen, X., Luo, W., Van Tooren, M. and Guo, J. Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles, *Progress in Aerospace Sciences*, **47** (6), 450–479, (2011).

13. Mainini, L. and Willcox, K. E. Data to decisions: Real-time structural assessment from sparse measurements affected by uncertainty, *Computers & Structures*, **182**, 296–312, (2017).

14. Mainini, L. and Willcox, K. E. Sensor placement strategy to inform decisions, *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, p. 3820, (2017).

15. West, T. D., (2019), *Hopes, Dreams, and Challenges of Digital Nirvana: The State of the Art and the Art of the Possible in Digital Twin and Digital Thread*. American Institute of Aeronautics and Astronautics, Inc.

16. Jones, D., Snider, C., Nassehi, A., Yon, J. and Hicks, B. Characterising the digital twin: A systematic literature review, *CIRP Journal of Manufacturing Science and Technology*, **29**, 36–52, (2020).

17. Fuller, A., Fan, Z., Day, C. and Barlow, C. Digital twin: Enabling technologies, challenges and open research, *IEEE Access*, **8**, 108952–108971, (2020).

18. van Dinter, R., Tekinerdogan, B. and Catal, C. Predictive maintenance using digital twins: A systematic literature review, *Information and Software Technology*, p. 107008, (2022).

19. Errandonea, I., Beltrán, S. and Arrizabalaga, S. Digital twin for maintenance: A literature review, *Computers in Industry*, **123**, 103316, (2020).

20. Gonzalez, M., Salgado, O., Croes, J., Pluymers, B. and Desmet, W. A digital twin for operational evaluation of vertical transportation systems, *Ieee Access*, **8**, 114389–114400, (2020).

21. McClellan, A., Lorenzetti, J., Pavone, M. and Farhat, C. A physics-based digital twin for model predictive control of autonomous unmanned aerial vehicle landing, *Philosophical Transactions of the Royal Society A*, **380** (2229), 20210204, (2022).

22. Lee, J.-H., Nam, Y.-S., Kim, Y., Liu, Y., Lee, J. and Yang, H. Real-time digital twin for ship operation in waves, *Ocean Engineering*, **266**, 112867, (2022).

REDUCED ORDER MODELLING FOR DIGITAL TWIN

Gianluigi Rozza
SISSA, Trieste, Italy
e-mail: grozza@sissa.it

An increasing number of disruptive innovations with high economic and social impact are characterizing the digitalization processes occurring in the last years. Despite the benefits of these developments, their speed and extent are limited by the available simulation technologies to handle complex models. The key idea of Digital Twin (DT) lies exactly in the creation of a digital counterpart of the physical asset able to replicate its behaviour. The necessity of a continuous and fast interaction between the physical and digital environments through background simulations has led to the introduction of Reduced Order Modeling (ROM) as crucial technology to simplify this simulation phase. ROM can indeed increase the speed of model execution while maintaining good level of accuracy and reducing the system degrees of freedom. In this way, DT can be applied to new complex fields, providing also more accurate information on the real counterpart thanks to the presence of ROM techniques.

Keywords: *model order reduction, digital twin, fluid dynamics*

---

1. Introduction

Digital twin (DT) [1, 2, 3] is one of the emerging technologies of Industry 4.0, aiming at the creation of smart products, processes, factories, and facilities. The primary concept behind Digital Twin involves developing a collection of interconnected models that can replicate the behavior of a physical asset. These models aim at offering dependable and swift estimations of all the relevant variables throughout the asset's entire operational lifespan, and make them accessible to users through a suitable interface [4]. The possibility of integrating a real asset with a virtual one leads thus to the creation of a Digital Twin able to interact with its physical counterpart by only receiving data from it or by a bidirectional exchange of data. The application of DT to some problems related to structural mechanics and fluid dynamics proposes new challenges. The request of accurate predictions of physical phenomena is essential for DT to get consistent results. However, these results can be obtained through high computational and time expensive simulations, leading to the infeasibility of simulating complex phenomena in real-time. A methodology to overcome this difficulty is represented by Reduced Order Modelling (ROM), which trades speed with a modest loss of accuracy [5]. The main benefit of ROMs comes from the presence of the offline-online paradigm. Firstly, in the offline phase the high fidelity simulations are performed for some given combinations of input parameters and the dominant dynamics are extracted to create the ROM. Based on the model constructed, in the online phase the evaluation of the magnitudes of interest simplifies in a linear superimposition of the modes multiplied by proper weights.

The benefits of the integration of ROMs and Digital Twin is thus twofold. On one side, ROMs enable to obtain more accurate information about the real counterpart, requiring low computational resources and real time simulations. On the other hand, ROMs can help in improving the quality of the products during the development phase by performing processes which are based on multi-query, e.g. optimization, uncertainty quantification, and controls. In this way, the coupling of DT with ROMs leads to a methodology deployable in many engineering contexts for several applications. Figures 1 and 1 provide some useful example of test cases. Figure 1 represents an example of a model used for a process of structural optimization, while in Fig. 2 the corresponding displacement field for a given loading condition is described.

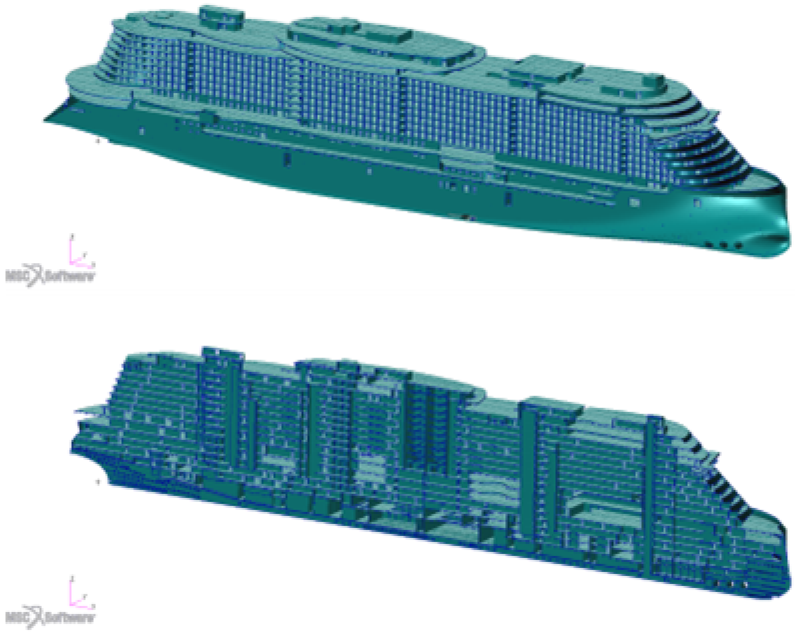


Figure 1: Example of a structural model in the naval engineering field. On the top the entire hull, while on the bottom a sectional view.

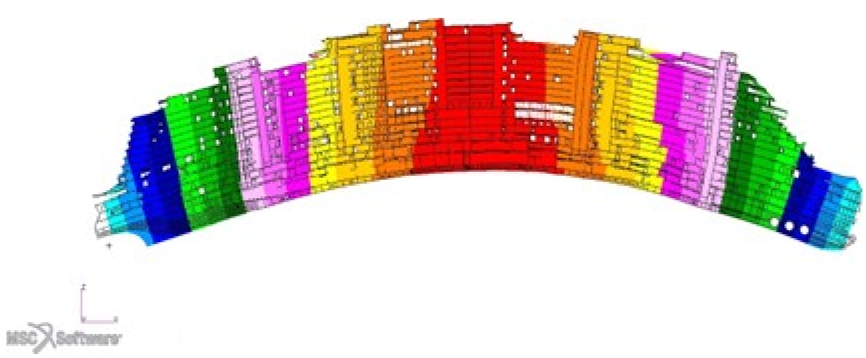


Figure 2: A modern cruise ship example of displacement field for the hogging loading condition.

REFERENCES

1. Rasheed, A., San, O. and Kvamsdal, T. Digital twin: Values, challenges and enablers from a modeling perspective, *Ieee Access*, **8**, 21980–22012, (2020).

2. Fuller, A., Fan, Z., Day, C. and Barlow, C. Digital twin: Enabling technologies, challenges and open research, *IEEE access*, **8**, 108952–108971, (2020).

3. Mashaly, M. Connecting the twins: A review on digital twin technology & its networking requirements, *Procedia Computer Science*, **184**, 299–305, (2021).

4. Hartmann, D. and Van der Auweraer, H. Digital twins, Cruz, M., Parés, C. and Quintela, P. (Eds.), *Progress in Industrial Mathematics: Success Stories*, Cham, pp. 3–17, Springer International Publishing, (2021).

5. Hartmann, D., Herz, M. and Wever, U. Model order reduction a key technology for digital twins, *Reduced-Order Modeling (ROM) for Simulation and Optimization: Powerful Algorithms as Key Enablers for Scientific Computing*, pp. 167–179, (2018).

DIGITAL TWIN ENABLED FRAMEWORK FOR INDUSTRIAL HUMAN-ROBOT COLLABORATION

Carmelo Mineo *and* Ignazio Infantino
National Research Council (CNR), Institute for High-Performance Computing and Networking (ICAR) Palermo, Italy
e-mail: carmelo.mineo@icar.cnr.it

Riccardo Siragusa *and* Filippo D'Ippolito
University of Palermo, Palermo, Italy

Industrial robots are systems used for manufacturing. Industrial robots allow for accomplishing automated tasks with high endurance, speed, and precision. However, industrial robots are typically enclosed within fences to guarantee human safety. A new generation of collaborative robots (cobots) is designed to collaborate with humans in an open workspace. Significant advances have regarded critical functional requirements and mechanisms for safety, collaboration variants and standardisations. As a result, cobots have become robust and suitable to be progressively introduced in manufacturing applications. Cobots are attractive for large, small and medium enterprises. Their capability to work alongside humans removes the requirement to confine them within a fenced space and makes them flexible manufacturing tools. In order to unlock the potential of cobots and speed up the exploitation of collaborative robotics in flexible manufacturing environments, research is needed to continue empowering cobots with ergonomic features, real-time sensing capabilities and controls, human-robot interfaces, and intuitive and task-driven programming. This work describes a framework being developed to facilitate the employment of human-robot collaboration in industrial assembly applications. Such a framework is enabled by a digital twin of the system.

Keywords: *cobots, robots, smart industry, digital twin*

1. Introduction

The technological development can be partially reflected by the scope and level of mechanisation and automation in replacing humans for operations, and decision-making supports [1]. Robots have been widely applied to relieve humans from tedious, repetitive, and risky tasks. However, for most applications, robots are still not advanced enough, and human assistance is needed to tackle the changes in tasks and environmental factors. Humans have advantages over machines regarding intuitive decisions, responsiveness, agility, and adaptability. Therefore, a new generation of collaborative robots (cobots) has been developed to enable humans to work with robots. Typically cobots are industrial robots outfitted with several sensors. Cobots are meant to put humans at the centre of a manufacturing task and free workers from various tasks without difficulty, danger, or dullness. Humans do not disappear from production; Cobots work alongside operators rather than replacing them. Whereas traditional industrial robots have found significant application in large factories due to the required space and investments, cobots are potentially attractive for large and small and medium enterprises. Therefore the cobot sector is expected to proliferate in the upcoming years [2]. Humanity is witnessing technological and economic changes related to integrating innovations such as artificial intelligence (AI), blockchain, the Internet of Things, cryptocurrencies, automation tools, etc. Whereas the main problem of the fourth industrial revolution (Industry 4.0) is about adding autonomous behaviours to machines, the fifth industrial revolution (Industry 5.0) assumes synergy between people and autonomous machines. The fifth industrial revolution has been defined as the revolution that will return workers to production plants, combining human intelligence and creativity with the capabilities of machines to increase the efficiency of processes [3]. The current vision is that the autonomous workforce will be receptive and informed of human intentions and desires, resulting in an exceptionally efficient value-added production process, flourishing reliable autonomy, and reducing waste and production costs. Assembly tasks that require human intervention are widespread in industrial manufacturing. A typical future scenario could consist of a person working alongside one or more cobots on assembling a specific industrial product made of several parts. The human worker may start the task with the robotic system monitoring the process using one or multiple overhead cameras operating as system eyes. The human-robot system is also managed by a data acquisition and control unit that captures the camera frames, performs image processing, and examines patterns through machine learning algorithms. Such a system may simultaneously monitor the person and the environment and infers what assembly phase the operator is performing at any time, using an analysis of human intentions based on artificial intelligence. Besides traditional RGB vision and stereo cameras, functional near-infrared spectroscopy (FNIRS) [4] may also suitably monitor human intentions. While forecasting human intention with sufficient accuracy, the cobots could help the human actor complete the assembly task, increasing efficiency. The worker could perceive that as similar to how another person standing next to him/her may help and give clues. For example, once the system predicts that the human operator will use a specific component in the next step of the assembly task, it commands the cobot to take that part in advance and deliver it to a convenient position for the operator. The current research challenge consists in making the afore-described vision a reality and creating non-invasive work companions so that human workers can focus on their tasks, feeling helped out safely. This paper introduces the authors' ongoing work to establish a flexible and future-proof framework for human-robot collaboration in industrial assembly applications. Crucially, such a framework builds on top of a digital representation of a real-world system, and this foundation will enable the implementation of multiple Industry 5.0 paradigms.

2. Framework components

The cobot-assisted system for industrial assembly tasks is being developed according to the framework shown in Figure 1. A computer runs a bespoke modular software package comprising the module for real-time communication and control of the cobot, the module for controlling the robotic gripper, the sensor data processing, the artificial intelligence algorithms and the graphical user interface (GUI).

The GUI is also composed of multiple modules. A page is designed for the user to input all system settings (e.g. the parameters to connect to the cobot and the gripper, the preferred speed and acceleration of the robotic movements, the gripping power and the safe robot distances for approach and retraction movements). Another GUI module is dedicated to defining the initial composition of the stock of components (part identification numbers and quantities). The most extensive GUI module is the digital 3D environment developed to support the offline simulations of the real-world system and function as a digital twin when connected to the system. Such a digital environment may enable an Extended Reality (XR) modality in human-robot interaction. The digital environment is initially populated with the components the user defines on the component stock page, but the stock is updated

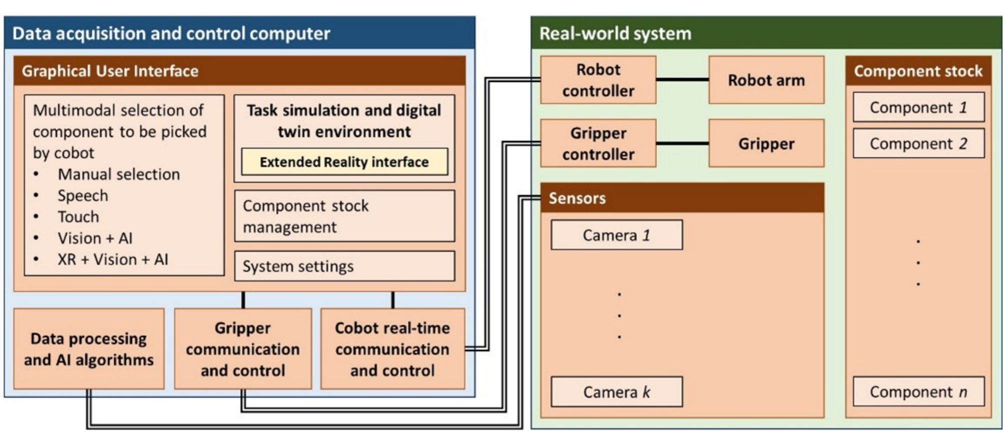


Figure 1: Proposed framework for industrial human-robot collaboration.

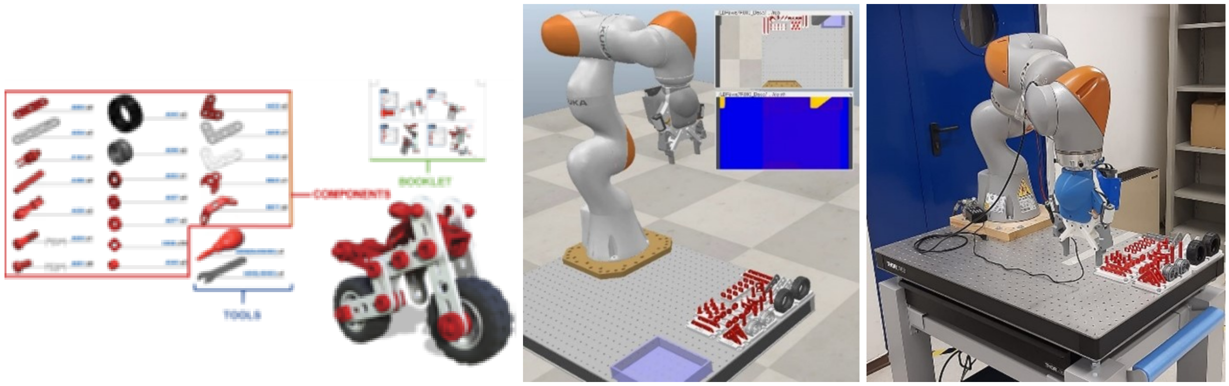


Figure 2: MECCANO toy kit (a), digital setup (b) and real-world setup (c).

during the execution of an assembly task to reflect the actual status of the real-world stock. Finally, the GUI comprises a multimodal approach to selecting the component the cobot needs to pick from the stock. The user can manually request a single component type by selecting its name or serial number and a specific number of units (via drop-down lists) or clicking on the component picture reproduced within the GUI. A user wishing to perform a manual selection can also trigger the robotic system to pick all the components required to complete a specific phase of the assembly task. Indeed, the assembly phases can be specified on the GUI setting page. Moving from manual selection towards more advanced triggering of the robotic system, the GUI will incorporate bespoke features to set up speech control, touch control, vision sensors, AI algorithms and XR interaction features. In turn, these features can be used to enable triggering the system through the worker's voice (him/her asking the system to pick a specific component), through the worker tapping the cobot wrist and through the autonomous inference of the worker's actions operated via the vision data, the AI algorithms and the XR interface.

3. The current state of the ongoing development

The system framework described above is a vision for which this paper's authors are working. Although most of the multimodal control features are still not developed, the underpinning digital simulation module, digital twin environment, and the communication links between the computer and the actual hardware have already been established and tested. The chosen application test case is based on a MEC-CANO toy kit. The kit contains all components required to assemble a toy model of a motorbike (see Fig. 2a). It is composed of 49 components with different shapes and sizes. Similarly to what happens in an industrial scenario, the subjects interact with tools such as a screwdriver and a wrench, as well as with tiny objects such as screws and bolts while executing a task involving sequential actions (e.g., take the wrench, tighten the bolt, put down wrench). Even though this scenario simplifies what can be found in real-world industrial settings, it is reasonably complex. Moreover, this is the same toy kit used by the authors of [5], where a benchmark to study human behaviour in an industrial-like scenario is proposed. Furthermore, the datasets generated in [5] are publicly available and may be used in the upcoming phases of this ongoing work. The robotic system used here is based on a KUKA iiwa 7 R800 collaborative robot and an in-house developed robotic gripper (consisting of a LEGO NXT Mindstorms controller, motors and gears [6] and 3D-printed body parts). Figure 2b-c shows the digital and real-world representation of the setup.

REFERENCES

1. Bi, Z. M., Luo, C., Miao, Z., Zhang, B., Zhang, W. and Wang, L. Safety assurance mechanisms of collaborative robotic systems in manufacturing, *Robotics and Computer-Integrated Manufacturing*, **67**, 102022, (2021).

2. Javaid, M., Haleem, A., Singh, R. P., Rab, S. and Suman, R. Significant applications of cobots in the field of manufacturing, *Cognitive Robotics*, **2**, 222–233, (2022).

3. Nahavandi, S. Industry 5.0—a human-centric solution, *Sustainability*, **11** (16), 4371, (2019).

4. Chen, W.-L., et al. Functional near-infrared spectroscopy and its clinical application in the field of neuroscience: advances and future directions, *Frontiers in neuroscience*, **14**, 724, (2020).

5. Ragusa, F., Furnari, A. and Farinella, G. M. Meccano: A multimodal egocentric dataset for humans behavior understanding in the industrial-like domain, *Computer Vision and Image Understanding*, p. 103764, (2023).

6. Gonçalves, J., Lima, J. and Costa, P. Rapid prototyping of mobile robots extending lego mindstorms platform, *IFAC Proceedings Volumes*, **39** (6), 47–52, (2006).

A GLOBAL SENSITIVITY ANALYSIS OF A HYBRID PDE–ODE MODEL FOR CANCER-ON-CHIP EXPERIMENTS

Elio Campanile
Fondazione The Microsoft Research University of Trento,
Centre for Computational and Systems Biology (COSBI), Rovereto, Italy
Department of Mathematics, University of Trento, Trento, Italy
e-mail: elio.campanile@unitn.it

Annachiara Colombi
Department of Mathematical Sciences G. L. Lagrange, Politecnico di Torino, Torino, Italy
e-mail: annachiara.colombi@polito.it

Gabriella Bretti
Istituto per le Applicazioni del Calcolo Mauro Picone, Roma, Italy
e-mail: gabriella.bretti@cnr.it

Keywords: *cancer-on-chip, cell migration, global sensitivity analysis, Morris method, Sobol method*

In recent years, the employment of digital twins to support the study of biological phenomena has widely diffused. In this context, digital twins able to reproduce complex phenomena such as organogenesis or tumorigenesis, possibly involve a large number of parameters, which may be also not directly measurable *in vivo/vitro*, and thus require a careful calibration. In this perspective, the global sensitivity analysis offers useful instruments to either identify regions in the space of parameters that result in reasonable scenarios, and to understand how the parameters affect the variability of the outputs of the model.

In the present work, we focus on the hybrid model proposed in [1] to reproduce Cancer-on-chip experiments where tumor cells, treated with chemotherapy drug, secrete chemical signals stimulating the response of immune cells. Specifically, the model consists of a coupled PDE-ODE system where cells are described as point particles while the chemical is characterized by the spatial distribution of their concentration. To investigate this digital twin, a global sensitivity analysis is performed by considering a series of target outputs, properly defined to characterize both the spatial distribution and the dynamics of immune cells. Among all the model parameters, we analyze the role of 13 parameters by performing a first screening using the method of Morris [2, 3], since each numerical simulations is a bit computationally expensive. This analysis confirms that (i) the selected target outputs are actually able to capture both typical and anomalous behaviors of the system; (ii) the considered ranges of parameters result in feasible scenarios (as the one shown in Fig. 1, left panel); (iii) the variability of both cell spatial distribution and their dynamics are mainly affected by 6 parameters (as shown in Fig. 1, right panel). These parameters include all the parameters characterizing the evolution of the chemical signal secreted by the tumor cells, i.e., the diffusion coefficient, the growth rate and the consumption rate; and some of the parameters regulating cell dynamics, i.e., the coefficient of the chemotactic effect, the damping coefficient and the drift velocity.

This is a first step in the investigation of this digital twin that suggests to continue our analysis by focusing on these 6 parameters only and applying more expensive methods able to quantify how much they affect the variance of the target outputs of our interest. In this perspective, we opt for the variance-decomposition based method of Sobol [4, 5] as it is able to quantify the effect of variations in the value of a parameter either one-by-one (main effect) or in combination with other parameters (total effect).

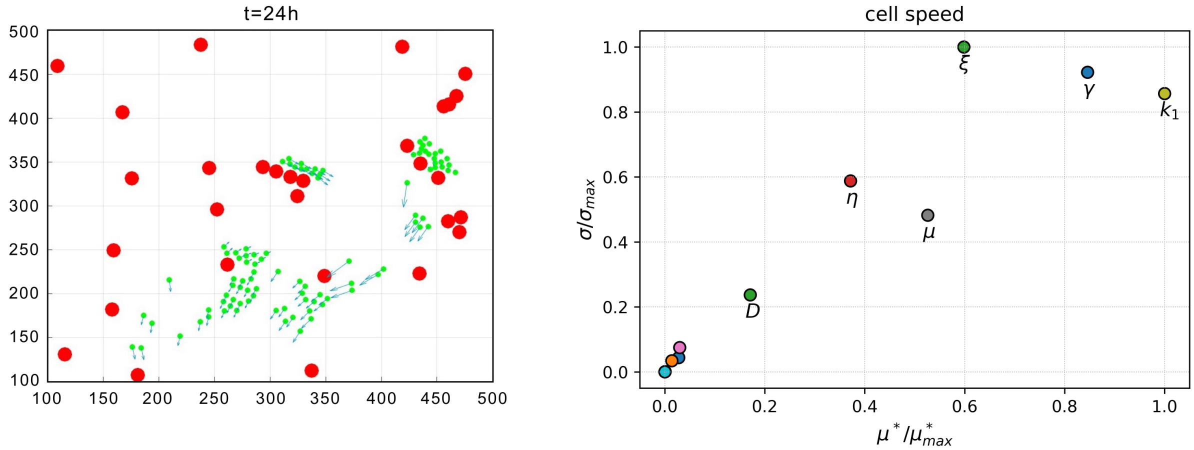


Figure 1: Left panel: Simulated dynamics of immune cells (green circles) in the chip environment with tumor cells (red circles). Right panel: Representative results showing how the parameters affect the velocity of immune cells.

REFERENCES

1. Bretti, G., De Ninno, A., Natalini, R., Peri, D. and Roselli, N. Estimation algorithm for a hybrid pde–ode model inspired by immunocompetent cancer-on-chip experiment, *Axioms*, **10** (4), (2021).

2. Morris, M. D. Factorial sampling plans for preliminary computational experiments, *Technometrics*, **33** (2), 161–174, (1991).

3. Campolongo, F., Cariboni, J. and Saltelli, A. An effective screening design for sensitivity analysis of large models, *Environmental Modelling & Software*, **22** (10), 1509–1518, (2007).

4. Saltelli, A. and Sobol, I. Sensitivity analysis for nonlinear mathematical models: Numerical experience, *Matematicheskoe Modelirovanie*, **7**, (1995).

5. Sobol, I. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, *Mathematics and Computers in Simulation*, **55** (1), 271–280, the Second IMACS Seminar on Monte Carlo Methods, (2001).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

HUMAN DIGITAL TWINS: MULTISCALE MODELING AND SIMULATIONS OF CELL DYNAMICS

Marta Menci
Università Campus Bio-Medico di Roma, Rome, Italy
e-mail: m.menci@unicampus.it

In the last years, the concept of digital twins is attracting the interests of different research fields. In particular human digital twins represent a promising tool to build customizable models of organs, and a way to simulate biological phenomena in order to predict outcomes, avoiding unnecessary surgery and reducing the costs of laboratory experiments. In this talk recent results on mathematical models inspired by Organ-on-chip experiments will be presented. The focus will be on a hybrid differential model for cell migrations due to mechanical and chemical interaction. Starting from the microscopic scale, numerical simulations of the derived kinetic and macroscopic models will be presented.

Keywords: *mathematical biology, cell migration, differential models, human digital twins*

1. Outline

Human digital twins represent an increasingly promising strategy to provide methods of analyzing healthcare practices. The complexity of human organisms and the difficult biological mechanism of systems within the human body are still challenging aspects of the building process.

In this talk we focus on a general class of hybrid mathematical models designed for collective motions of cells due to chemical stimuli, describing the behavior of immune cells and the concentration of chemicals in Cancer-on-chip environment [1]. In this context, cells are modeled as discrete entities and their dynamics is given by ODEs, while the chemical signal influencing the motion is considered as a continuous signal which solves a diffusive equation. The generical structure of the hybrid approach envisaged can be summarized as follows: consider on $\mathbb{R}^{2dN} \ni ((x_i(t))_{i=1,\dots,N}, (v_i(t))_{i=1,\dots,N}) := (X(t), V(t))$ the following vector field

$$\begin{cases} \dot{x}_i(t) = v_i \\ \dot{v}_i(t) = F_i(t, X(t), V(t)) \end{cases} \qquad i = 1, \dots, N, \tag{1}$$

where

$$F_i(t, X, V) = \sum_{j=1}^{N} \gamma(v_i - v_j, x_i - x_j) + \eta \nabla_x \varphi^t(x_i) \tag{2}$$

Here x_i, v_i are the position and velocity of the i-th cell, function γ models the mechanical interactions among cells and φ stands for a generic chemical signal produced by the cells themselves and such that the cells are attracted towards the direction where $\nabla_x \varphi$ is growing. In particular, φ satisfies the equation

$$\partial_s \varphi^s(x) = D\Delta_x \varphi - \kappa\varphi + f(x, X(s)),\ s \in [0, t], \tag{3}$$

for some $\kappa, D, \eta \geq 0$ and function f of the form

$$f(x, X) = \frac{1}{N} \sum_{j=1}^{N} \chi(x - x_i),\ \ \chi \in \mathcal{C}_c^1. \tag{4}$$

Numerical and analytical results on hybrid models in a two-dimensional domain ($d = 2$) can be found in [2], whereas application to Cancer-on-chip experiments in [3].

The final goal of building digital twins is to predict the behaviour of real organs, hence on the scale of billions of cells, using the data from the Organ-on-chip experiments. To this end, it is crucial to design a macroscopic model, which allows to overcome the computation complexity of a microscopic approach, keeping the microscopic experimental information.

From the analytical point of view, a pressureless nonlocal Euler-type system has been derived for the class of model (1)-(2), and a rigorous equivalence between kinetic and macroscopic scale has been proved assuming monokinetic initial data [4]. Denoting with μ the macroscopic density, and u the velocity field, the system reads:

$$\begin{cases} \partial_t \mu^t + \nabla(u^t \mu^t) = 0 \\ \partial_t(\mu^t u^t) + \nabla(\mu^t (u^t)^{\otimes 2}) = \mu^t \int \gamma(\cdot - y, u^t(\cdot) - u^t(y))\mu^t(y)dy + \eta\mu^t \nabla\psi^t - \alpha\mu^t u^t \\ \partial_s \psi^s = D\Delta\psi - \kappa\psi + \chi * \mu^s,\ s \in [0,t], \end{cases}$$

Concerning generic and more realistic initial configurations, a numerical study has been performed [5], as a first step in the study of the model at different scale. Two main aspects will be debated: the difference with respect to Euler system of equation accounting for pressure, and the role of the non-local integral term. Numerical simulations, investigating the role of key parameters of the model in different scenarios, will be shown. In particular, a comparison between microscopic, kinetic and macroscopic scale will be highlighted.

This talk is based on joint works and ongoing collaborations in particular with Gabriella Bretti, Roberto Natalini (IAC-CNR) and Thierry Paul (CNRS-Sorbonne Universitè) .

REFERENCES

1. De Ninno, A., Bertani, F. R., Gerardino, A., Schiavoni, G., Musella, M., Galassi, C., Mattei, F., Sistigu, A. and Businaro, L. Microfluidic co-culture models for dissecting the immune response in in vitro tumor microenvironments, *JoVE (Journal of Visualized Experiments)*, (170), e61895, (2021).

2. Costanzo, E. D., Menci, M., Messina, E., Natalini, R. and Vecchio, A. A hybrid model of collective motion of discrete particles under alignment and continuum chemotaxis., *Discrete & Continuous Dynamical Systems-Series B*, **25** (1), (2020).

3. Bretti, G., De Ninno, A., Natalini, R., Peri, D. and Roselli, N. Estimation algorithm for a hybrid pde–ode model inspired by immunocompetent cancer-on-chip experiment, *Axioms*, **10** (4), 243, (2021).

4. Natalini, R. and Paul, T. The mean-field limit for hybrid models of collective motions with chemotaxis, *SIAM Journal on Mathematical Analysis*, **55** (2), 900–928, (2023).

5. Menci, M., Natalini, R. and Paul, T. Microscopic, kinetic and hydrodynamic hybrid models of collective motions with chemotaxis: a numerical study, *arXiv preprint arXiv:2306.12835*, (2023).

PHYSIOLOGY-INFORMED MACHINE LEARNING FOR PATIENT-SPECIFIC INTRAOCULAR PRESSURE AND BLOOD PRESSURE MANAGEMENT IN GLAUCOMA

Lorena Bociu
Mathematics, North Carolina State University, Raleigh, NC, United States

Giovanna Guidoboni
Electrical and Computer Engineering, University of Maine, Orono, ME, United States

Alon Harris *and* Brent Siesky
Ophthalmology, Icahn School of Medicine at Mount Sinai, New York, NY, United States

Daniele Prada
CNR–IMATI, Pavia, Italy
e-mail: daniele.prada@imati.cnr.it

Riccardo Sacco
Mathematics, Politecnico di Milano, Milano, Italy

Vision loss in many glaucoma patients continues despite successful treatment aimed at lowering intraocular pressure. The reasons why this happens remain elusive. Here, we show that mathematical modeling and machine learning can help identify subgroups of glaucoma patients whose disease process entails different relative contributions of risk factors, particularly intraocular pressure and blood pressure. This work may enable the development of novel precision medical approaches for glaucoma care.

Keywords: *physiology-informed machine learning, glaucoma, mathematical modeling, patient-specific, precision medicine*

1. Purpose

Vision loss in many open-angle glaucoma (OAG) patients continues despite successful treatment lowering intraocular pressure (IOP). The reasons remain elusive. Here, we show that mathematical modeling and machine learning (ML) can help identify subgroups of OAG patients whose disease process entails different relative contributions of IOP and blood pressure (BP).

2. Methods

115 OAG patients were assessed every 6 months over a 7-year period for IOP, systolic and diastolic blood pressures (SBP, DBP), heart rate (HR), structural and hemodynamic evaluations via ocular coherence tomography (OCT), Heidelberg Retinal Tomography (HRT), Heidelberg Retinal Flowmetry (HRF), and Color Doppler Imaging (CDI). Fuzzy c-means (FCM) clustering was applied to the dataset comprising: (i) IOP, SBP, DBP, HR measured at the first visit for each patient in the IGPS study; and (ii) patient-specific estimates of vascular pressures and resistances, mechanical stresses and strains, obtained via validated mathematical models [1, 2]. FCM is part of ML and the mathematical models are based on physiology, leading to physiology-informed ML. Follow-up visits and data from OCT, HRT, HRF and CDI were not used for clustering.

3. Results

While the data for IOP and mean arterial pressure (MAP) for the first visit of each IGPS patient do not exhibit any particular pattern, the physiology-informed ML method revealed distinct clusters, which are associated with different clinical outcomes after 4 years (p values obtained with the 2-sample paired Wilcoxon signed rank test for medians). Clusters show minimal glaucoma progression, marked structural progression accompanied by significant hemodynamic changes and/or significant changes only in HRT and HRF markers, but not in OCT and CDI.

4. Conclusions

This study suggests that the proposed physiology-informed ML approach can identify and quantify the relative contributions of IOP and BP on the OAG risk for patient subgroups. Thus, this approach may enable precision medical approaches of IOP and BP management in OAG.

REFERENCES

1. Guidoboni, G., Harris, A., Cassani, S., Arciero, J., Siesky, B., Amireskandari, A., Tobe, L., Egan, P., Januleviciene, I. and Park, J. Intraocular pressure, blood pressure, and retinal blood flow autoregulation: A mathematical model to clarify their relationship and clinical relevance, *IOVS*, **55**, 4105–4118, (2014).

2. Causin, P., Guidoboni, G., Harris, A., Prada, D., Sacco, R. and Terragni, S. A poroelastic model for the perfusion of the lamina cribrosa in the optic nerve head, *Mathematical Biosciences*, **257**, 33–41, (2014).

TOWARDS A DIGITAL TWIN FOR PERSONALIZED DIABETES PREVENTION: THE PRAESIIDIUM PROJECT

Alessia Paglialonga, Marta Lenatti, Davide Simeone, Pierluigi Francesco De Paola, Alberto Carlevaro, Maurizio Mongelli, *and* Fabrizio Dabbene
Cnr-Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (CNR-IEIIT), Italy
e-mail: alessia.paglialonga@cnr.it

Filippo Castiglione
Technology Innovation Institute, Biotech Research Center, Abu Dhabi, UAE and Cnr-Istituto per le Applicazioni del Calcolo (CNR-IAC), Rome, Italy

Maria Concetta Palumbo, Paola Stolfi, *and* Paolo Tieri
Cnr-Istituto per le Applicazioni del Calcolo (CNR-IAC), Rome, Italy

This contribution outlines current research aimed at developing models for personalized type 2 diabetes mellitus (T2D) prevention in the framework of the European project PRAESIIDIUM (Physics Informed Machine Learning-Based Prediction and Reversion of Impaired Fasting Glucose Management) aimed at building a digital twin for preventing T2D in patients at risk. Specifically, the modelling approaches include both a multiscale, hybrid computational model of the human metaflammatory (metabolic and inflammatory) status, and data-driven models of the risk of developing T2D able to generate personalized recommendations for mitigating the individual risk. The prediction algorithm will draw on a rich set of information for training, derived from prior clinical data, the individual's family history, and prospective clinical trials including clinical variables, wearable sensors, and a tracking mobile app (for diet, physical activity, and lifestyle). The models developed within the project will be the basis for building a platform for healthcare professionals and patients to estimate and monitor the individual risk of T2D in real time, thus potentially supporting personalized prevention and patient engagement.

Keywords: *counterfactuals, multiscale models, physics-informed machine learning, prediabetes, type 2 diabetes*

1. Introduction

Diabetes is a chronic disease affecting about 10.5% of the global population and characterized by abnormally increased blood glucose level and associated with complications such as cardiovascular and neurological diseases. Type-2 diabetes (T2D) is characterised by insulin resistance (i.e., the insulin hormone cannot effectively regulate the blood glucose levels) and/or a failure in compensatory mechanisms for insulin secretion (i.e., the body does not correctly use the insulin produced) [1]. Prediabetes is an early-stage condition in the "healthy-to-T2D transition" but, unlike T2D, it can be reversed, for example by lifestyle modification. However, despite high-quality prevention guidelines, prevention of T2D remains a global challenge. Developing prevention strategies to reduce the risk of T2D is of paramount importance to limit the burden of disease. Individualized recommendations can be more effective than 'one size fits all' approaches in supporting patient motivation and behaviour change.

To support personalized prevention in patients at risk of T2D, data-driven approaches using digital twin technology can be helpful to support patient awareness and engagement. For example, digital twins can be used to monitor the individual risk in real time, thus helping the patient understand how the risk decreases as a function of preventive strategies (e.g., diet, physical activity, improved lifestyle, medications). Recently, the European project PRAESIIDIUM[1] (Physics Informed Machine Learning-Based Prediction and Reversion of Impaired Fasting Glu-

[1] https://praesiidium.spindoxlabs.com/

cose Management), including 11 participants from seven countries, was launched with the objective to develop an AI-based tool coupled with multi-scale, multi-organ integrated mathematical equations for the real-time prediction of the prediabetes risk of an individual and the identification of personalized recommendations to reduce and monitor the individual risk in real time. Within the framework of PRAESIIDIUM, we have developed a combination of methods, including multiscale modelling and data-driven modelling, as a basis for a digital twin able to characterize the individual risk of developing T2D and identify personalized countermeasures to reduce the risk.

2. Multiscale modelling: Mission-T2D (MT2D)

MT2D is a computational model of the human metabolic and inflammatory status that is determined by the individual dietary and activity habits[2]. MT2D embraces, to a certain degree of sophistication, different levels of description from the molecular/cellular to organs and the whole-body and is based on the sub-jects-specific features to achieve greater generalization and user-customization.

It includes a model for the food intake, stomach emptying and gut absorption of a mixed meal [2] (with the three macronutrients proteins, carbohydrates, fats), a component to account for the effects of physical activity on the hormones' regulation [3], a description of the secretion of Interleukine-6 both by the skeletal muscle during physical activity [4] and by the adipose tissue in resting conditions [5], giving evidence of its dual nature as an adipokine (i.e., adipose tissue-derived cytokine) and as a myokine (i.e., muscle-derived cytokine), and a characterization of energy intake-expenditure balance leading to gaining/losing weight and finally the detailing of the immunological scenario of the subject [6]. All these components are merged into a single, integrated simulation tool with the aim to provide and explore the systemic picture of the metaflammatory status of an individual, that can be potentially exploited proactively to prevent the onset of T2D.

3. Data-driven modelling: Counterfactual explanations and multi-input multi-output dynamic models

Counterfactual explanations (CE) are a local eXplainable AI technique and are defined as the set of minimal changes that, applied to the input features related to a specific instance, can change its predicted class. We developed a new method for generating CE using a Support Vector Data Description classifier to define personalized recommendations to reduce the risk of T2D [7, 8, 9]. Using routinely collected biomarkers extracted from a balanced dataset of Electronic Medical Records (EMRs, derived from the Canadian Primary Care Sentinel Surveillance Network, CPCSSN[3]) of 5582 patients at low/high risk of developing T2D, we determined low-T2D-risk regions with varying specificity, and we assessed the related CE using quantitative performance metrics (e.g., availability, validity, actionability) and a qualitative survey administered to expert clinicians. The minimum viable changes implied a significant modification of fasting blood sugar, systolic blood pressure, triglycerides, and high-density lipoprotein to lower the risk of T2D, particularly in hypertensive patients. The surveyed experts were overall satisfied with the proposed method and provided suggestions for building more meaningful recommendations. Future research will focus on a larger set of input features (e.g., medications) and on the extension to multi-class problems [9] (e.g., to model the risk of multiple conditions, or health-prediabetes-T2D transitions).

In addition to CE, dynamic models were developed to investigate the long-term trajectories of biomarkers in patients at high/low risk of developing T2D and assess the individual risk several years before the onset. We have developed a multi-input multi-output approach based on a Multivariate Gaussian Process Model with an autoregressive structure [10]. Using routinely collected data (biomarkers, comorbidities, and medications) extracted from CPCCSN3 EMRs from three or more consecutive years from 667 T2D patients (before diagnosis) and from 25094 no-T2D patients, we observed that the slow dynamics of biomarkers (e.g., blood pressure, body mass index, lipids, fasting blood sugar) were able to model accurately the long-term evolution in real patients and were consistent with the literature. Future research includes developments of more specific models (e.g., for male/females, for patients

[2] https://kraken.iac.rm.cnr.it/T2DM/
[3] http://cpcssn.ca

with/without comorbidities or risk factors), validation on large real-world datasets, and individualized modelling of the short- and long-term effects of physical activity on T2D risk [11].

4. Conclusions

The combination of multiscale and data-driven models, backed by data from real-world retrospective databases and data from ad hoc, prospective clinical trials in the project PRAESIIDIUM can help in building personalized algorithms for estimating, monitoring, and reducing the individual risk of developing T2D. The proposed modelling approach can be the basis for building a digital twin for T2D prevention that can support individuals in gaining awareness of the factors influencing the risk, thus potentially engaging patients in safeguarding their own health.

ACKNOWLEDGEMENT

This work was supported in part by the European Union, Grant Agreement n. 101095672 (PRAESIIDIUM "Physics informed machine learning-based prediction and reversion of impaired fasting glucose management", call HORIZON-HLTH-2022-STAYHLTH-02). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them. M. Lenatti is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVIII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma. D. Simeone, P.F. De Paola, and A. Carlevaro are PhD students at Politecnico di Milano (XXXVIII cycle), Politecnico di Bari (National PhD Program in Autonomous Systems, XXXVIII cycle), and Università di Genova (XXXVI cycle), respectively.

REFERENCES

1. Magliano, D. J., Boyko, E. J., Atlas, I. D., et al., (2021), What is diabetes? *IDF DIABETES ATLAS [Internet]. 10th edition*, International Diabetes Federation.

2. Palumbo, M. C., de Graaf, A. A., Morettini, M., Tieri, P., Krishnan, S. and Castiglione, F. A computational model of the effects of macronutrients absorption and physical exercise on hormonal regulation and metabolic homeostasis, *Computers in Biology and Medicine*, p. 107158, (2023).

3. Palumbo, M. C., Morettini, M., Tieri, P., Diele, F., Sacchetti, M. and Castiglione, F. Personalizing physical exercise in a computational model of fuel homeostasis, *PLoS computational biology*, **14** (4), e1006073, (2018).

4. Morettini, M., Palumbo, M. C., Sacchetti, M., Castiglione, F. and Mazza, C. A system model of the effects of exercise on plasma interleukin-6 dynamics in healthy individuals: role of skeletal muscle and adipose tissue, *PloS one*, **12** (7), e0181224, (2017).

5. Prana, V., Tieri, P., Palumbo, M. C., Mancini, E., Castiglione, F., et al. Modeling the effect of high calorie diet on the interplay between adipose tissue, inflammation, and diabetes, *Computational and mathematical methods in medicine*, **2019**, (2019).

6. Bernaschi, M. and Castiglione, F. Design and implementation of an immune system simulator, *Computers in biology and medicine*, **31** (5), 303–331, (2001).

7. Lenatti, M., Carlevaro, A., Guergachi, A., Keshavjee, K., Mongelli, M. and Paglialonga, A. A novel method to derive personalized minimum viable recommendations for type 2 diabetes prevention based on counterfactual explanations, *Plos one*, **17** (11), e0272825, (2022).

8. Carlevaro, A., Lenatti, M., Paglialonga, A. and Mongelli, M. Counterfactual building and evaluation via explainable support vector data description, *IEEE Access*, **10**, 60849–60861, (2022).

9. Carlevaro, A., Lenatti, M., Paglialonga, A. and Mongelli, M. Multi-class counterfactual explanations using support vector data description, (2023).

10. Simeone, D., Lenatti, M., Lagoa, C., Keshavjee, K., Guergachi, A., Dabbene, F. and Paglialonga, A. Multi-input multi-output dynamic modelling of type 2 diabetes progression, *Submitted to the European Federation of Medical Informatics Special Topic Conference (EFMI STC 2023), Oct 25-26, Torino, Italy*, (2023).

11. De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F. and Borri, A. The long-term effects of physical activity on blood glucose regulation: a model to unravel diabetes progression, *IEEE Control Systems Letters*, (2023).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

ALTERNATIVE APPROACHES TO CANCER TREATMENT: TOWARDS STOCHASTIC TUMOUR CONTROL

Alessandro Borri
Institute for Systems Analysis and Computer Science 'A. Ruberti' (IASI), National Research Council of Italy, L'Aquila, Italy
e-mail: alessandro.borri@iasi.cnr.it

Elena Mitkova Nikolova
DISIM Department, University of L'Aquila, L'Aquila, Italy

Pasquale Palumbo
Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy

Federico Papa
Institute for Systems Analysis and Computer Science 'A. Ruberti' (IASI), National Research Council of Italy, Rome, Italy

In this study, we address the challenge of controlling tumor growth by leveraging a low-dimensional model based on the Chemical Reaction Network (CRN) formalism. Designed to function in both deterministic and stochastic frameworks, our findings demonstrate that the deterministic approach adequately characterizes the system behavior in presence of high number of tumor cells, particularly for the purpose of control planning. In this context, we propose different control strategies involving constant or variable treatment plans, exploiting complete or partial knowledge of the system state, and providing asymptotical guarantees of tumour eradication. However, after the tumour mass has been consistently reduced (namely when the number of tumor cells becomes relatively low), random fluctuations are not negligible any more, implying that a stochastic formalization can be more accurate. We preliminarily show by numerical simulations that, due to the properties of the underlying Continuous-Time Markov Process, finite-time tumour eradication can be obtained in the stochastic framework with statistical guarantees.

Keywords: *tumour growth and treatment, control theory, qualitative behavior analysis, numerical simulation, stochastic control*

1. Introduction

Model-based control has gained increasing interest in recent decades due to its ability to design sophisticated feedback regulations that consider the inherent dynamics of the system under investigation. In biomedical applications, minimal models are often utilized as they capture the basic relationships among variables without explicitly detailing all the physical or molecular mechanisms. These models can be easily identified through standard perturbation experiments and enable the synthesis of affordable and readily implementable control laws. Starting from the seminal work by Hahnfeldt et al. [1], proposing a low-dimensional minimally parametrized Ordinary Differential Equation (ODE) model for vascular tumor growth, several theoretical and experimental advancements have been made, including model extensions [2, 3] and investigations into closed-loop and open-loop anti-angiogenic drugging combined with chemotherapy treatments [4, 5, 6, 7, 8, 9, 10], both in the deterministic and in the stochastic setting [11, 12].

More recently, tumor growth models based on the Chemical Reaction Network (CRN) formalism have been proposed [13, 14, 15]. As described in [16], the chemical players considered by this formulation are the growing cancer cells X_1, the necrotic cancer cells X_2, and the drug molecules X_3, subject to the following set of reactions:

$$\begin{array}{ll} R_1 - \text{proliferation}: & X_1 \to 2X_1, \\ R_2 - \text{necrosis}: & X_1 \to X_2, \\ R_3 - \text{dead cell washout}: & X_2 \to \emptyset, \\ R_4 - \text{drug clearance}: & X_3 \to \emptyset, \\ R_5 - \text{drug action}: & X_1 + X_3 \to X_2, \\ R_6 - \text{drug administration}: & \emptyset \to X_3. \end{array} \tag{1}$$

The advantage of this approach lies in its ability to model CRNs in a stochastic framework exploiting Continuous-Time Markov Chains (CTMCs) or, alternatively, in a mean-field ODE model approximating its average dynamics [17], which is more manageable from a computational viewpoint and can be effectively utilized when the copy number of chemical players is high. The contribution [18] builds upon the qualitative analysis presented in [16] for the ODE model associated with the CRN and further investigates deterministic feedback control laws, with possibly partial information, highlighting their advantages compared to constant administration therapies.

With respect to [18], which is devoted to deterministic control, in this study we perform a preliminary simulative investigation of the potential of feedback in a stochastic sense, i.e. by using stochastic models. To this end, we build a stochastic differential equation (SDE) version of model [13], exploiting the formalism of the Chemical Langevin Equation [19]. This approach has the computational advantage of accounting for the stochastic nature of the phenomenon at hand (in an approximate sense) while preserving the low dimensionality of the ODE model. Since all the states of the CRN reaction graph (Figure 1, left panel) with $X_1 = 0$ are absorbing states, we are able to obtain a statistics of the eradication time over $1,000$ random paths (Figure 1, right panel), showing the potential of this approach for future investigations.

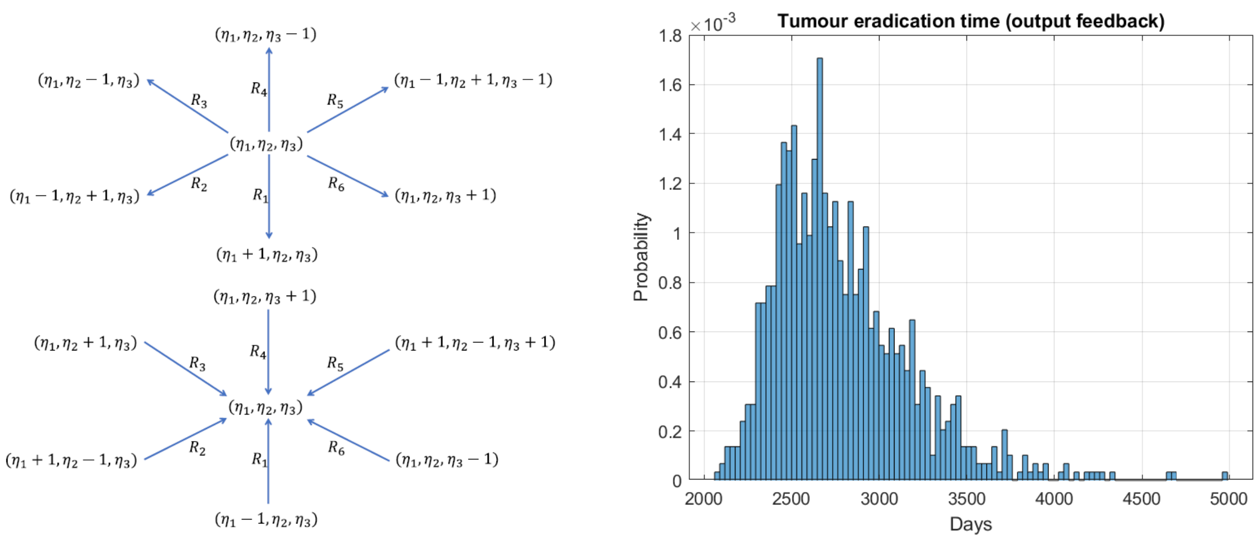


Figure 1: Left panel: State transition diagram for the CRN tumour model. Right panel: example of finite-time eradication in stochastic tumour control.

REFERENCES

1. Hahnfeldt, P., Panigrahy, D., Folkman, J. and Hlatky, L. Tumor development under angiogenic signaling, *Cancer research*, **59** (19), 4770–4775, (1999).

2. d'Onofrio, A. and Gandolfi, A. Tumour eradication by antiangiogenic therapy: Analysis and extensions of the model by hahnfeldt et al. (1999), *Mathematical Biosciences*, **191**, 159–184, (2004).

3. d'Onofrio, A. and Gandolfi, A. Chemotherapy of vascularised tumours: Role of vessel density and the effect of vascular "pruning", *Journal of Theoretical Biology*, **264**, 253–265, (2010).

4. Cacace, F., Cusimano, V., Germani, A., Palumbo, P. and Papa, P. Closed-loop control of tumor growth by means of anti-angiogenic administration, *Mathematical Biosciences and Engineering*, **15** (4), (2018).

5. Cacace, F., Cusimano, V., Germani, A. and Palumbo, P. Optimal impulsive control with application to antiangiogenic tumor therapy, *IEEE Trans. Control Systems Technology*, pp. 1–12, (2018).

6. Drexler, D., Sápi, J. and Kovács, L. Optimal discrete time control of antiangiogenic tumor therapy, *Proc. 20th IFAC World Congress*, pp. 14046–14051, (2017).

7. Ledzewicz, U. and Schättler, H. Optimal and suboptimal protocols for a class of mathematical models of tumor anti-angiogenesis, *Journal of Theoretical Biology*, **252** (2), 295–312, (2008).

8. Sápi, J., Drexler, D., Harmati, I., Sápi, Z. and Kovács, L. Qualitative analysis of tumor growth model under antiangiogenic therapy–choosing the effective operating point and design parameters for controller design, *Optim Control Appl Methods*, **37** (5), 848–866, (2016).

9. d'Onofrio, A., Ledzewicz, U., Maurer, H. and Schättler, H. On optimal delivery of combination therapy for tumors, *Mathematical biosciences*, **222** (1), 13–26, (2009).

10. Ledzewicz, U., Maurer, H. and Schaettler, H. Optimal and suboptimal protocols for a mathematical model for tumor anti-angiogenesis in combination with chemotherapy, *Mathematical Biosciences and Engineering*, **8** (2), 307–323, (2011).

11. Preziosi, L., Toscani, G. and Zanella, M. Control of tumor growth distributions through kinetic methods, *Journal of Theoretical Biology*, **514**, 110579, (2021).

12. Medaglia, A., Colelli, G., Farina, L., Bacila, A., Bini, P., Marchioni, E., Figini, S., Pichiecchio, A. and Zanella, M. Uncertainty quantification and control of kinetic models of tumour growth under clinical uncertainties, *International Journal of Non-Linear Mechanics*, **141**, 103933, (2022).

13. Drexler, D. A., Ferenci, T., Lovrics, A. and Kovács, L. Modeling of tumor growth incorporating the effect of pegylated liposomal doxorubicin, *IEEE 23rd Int. Conf. on Intelligent Engineering Systems*, pp. 369–374, (2019).

14. Drexler, D. A., Sápi, J. and Kovács, L. A minimal model of tumor growth with angiogenic inhibition using bevacizumab, *Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics*, pp. 185–190, (2017).

15. Drexler, D. A., Sápi, J. and Kovács, L. Modeling of tumor growth incorporating the effects of necrosis and the effect of bevacizumab, *Complexity*, pp. 1–11, (2017).

16. Borri, A., Palumbo, P. and Papa, F. Deterministic vs stochastic formulations and qualitative analysis of a recent tumour growth model, *IFAC-PapersOnLine*, **53** (2), 16418–16423, (2020).

17. van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, 3rd Ed., The Nederlands: Elsevier (2007).

18. Papa, F., Borri, A. and Palumbo, P. Tumour growth control: analysis of alternative approaches, *Journal of Theoretical Biology*, **562**, 111420, (2023).

19. Gillespie, D. T. The chemical Langevin equation, *The Journal of Chemical Physics*, **113** (1), 297–306, (2000).

DIGITAL TWINS AND VIRTUAL POPULATIONS: APPLICATIONS IN QUANTITATIVE SYSTEMS PHARMACOLOGY

Federico Reali, Alessio Paris, *and* Stefano Giampiccolo
Fondazione The Microsoft Research - University of Trento Centre for Computational and Systems Biology, Rovereto, Italy

Elena Righetti *and* Luca Marchetti
Fondazione The Microsoft Research University of Trento Centre for Computational and Systems Biology, Rovereto, Italy
Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Trento , Italy
e-mail: {reali, paris, giampiccolo, righetti, marchetti}@cosbi.eu, luca.marchetti@unitn.it

Quantitative Systems Pharmacology (QSP) models face challenges in estimating parameters from patient-specific data to obtain their digital counterparts. To this end, global and local fitting strategies involving optimization methods such as least squares and Bayesian inference are employed. In addition, identifiability and uncertainty quantification techniques can be instrumental in developing robust QSP models. Virtual populations generated by statistical approaches and Monte Carlo simulations can capture patient variability and incorporate genetic, demographic, and treatment factors. Digital patients (or twins) and virtual populations can help optimize dosing, inform clinical trials, and aid in understanding diseases. Here, we report two examples of their applications in studying neurofilament trafficking in spinal muscular atrophy patients and the rare Gaucher disease type 1, showing promising results. Overall, QSP combined with digital patients and virtual populations has the potential to push drug development toward personalized medicine.

Keywords: *QSP, mathematical modeling, virtual populations, neurodegenerative diseases, rare diseases*

1. Introduction

Quantitative systems pharmacology (QSP) is an approach that utilizes mathematical modeling to understand and predict the behavior of complex biological systems and their interactions with pharmacological treatments. It integrates data from multiple sources, from literature, *in-vitro*, preclinical and clinical experiments, and encompasses different biological levels [1]. Model-informed drug development (MIDD) is a framework that leverages pharmacokinetics, pharmacodynamics, and QSP models to facilitate decision-making throughout the drug development process [2]. The combination of QSP and MIDD has significantly accelerated the pace of introducing new drugs into clinical practice.

One of the primary challenges in QSP and MIDD lies in accounting for the variability and uncertainty in model parameters and outputs. Addressing this challenge involves employing individualized calibrations, which are techniques used to estimate model parameters for specific individuals based on observed data [3]. Additionally, digital twins and virtual populations are utilized to tackle this challenge. Digital twins are virtual replicas of individuals that can simulate their responses to various interventions or scenarios. On the other hand, virtual populations consist of collections of digital twins that represent the diversity and heterogeneity of a target population. By incorporating these methods, the accuracy, reliability, and applicability of QSP models can be enhanced, ultimately leading to more efficient and effective drug development processes.

In certain cases, such as rare or neurodegenerative diseases, virtual representations of patients can provide more comprehensive datasets that aid in understanding disease mechanisms and predicting treatment efficacy. This information is especially valuable when the number of patients is limited or when there are significant barriers to accessing the disease site. Furthermore, employing digital patients and virtual populations allows for the generalization of results to different populations, considering factors such as age or disease severity. This work exemplifies how digital patients and virtual populations were utilized within this context.

2. Digital patients and virtual populations in QSP modeling

One of the primary challenges in QSP modeling, which encompass multiple biological scales, is to estimate model parameters using patient-specific time series data, including biomarkers, pharmacokinetics data, and clinical manifestations or outcomes [1]. When sufficient data is available, this process enables the creation of digital patients, where specific model parameters are linked to the measured endpoints of individual patients. This approach helps to account for inter-individual variability and optimize dosing regimens. However, there are many adopted fitting strategies in QSP, and they may depend on various factors such as data availability and quality, model complexity, and computational resources. Different methods can be used to fit patient-specific time series data, such as least squares, heuristic evolutionary strategies, and Bayesian inference such as Markov Chain Monte Carlo. These methods have different advantages and disadvantages for both local and global optimization tasks, and they are essential for creating a digital representation of the patient [4].

Another key aspect of determining digital patients is the identifiability and uncertainty quantification of model parameters, which reflect the biological variability and measurement errors in the data. Parameter identifiability refers to the ability to estimate unique and precise values of model parameters from the available data, while uncertainty quantification measures the range and confidence of parameter estimates. Both are essential for ensuring the robustness, validity, and predictive power of QSP models and to assure that the model parameters can correctly identify a reliable digital twin. Several methods have been proposed and applied for parameter identifiability and uncertainty quantification in QSP, such as sensitivity analysis, Bayesian inference, profile likelihood, and bootstrap methods [3]. These methods have different advantages and limitations depending on the complexity and structure of the model, the quality and quantity of the data, and the available computational resources [3].

Once enough digital patients are collected, it is possible to define virtual populations (VPs) that can capture the inherent variability present in real patient populations. In addition, virtual populations are often created by incorporating various sources of variability, such as genetic factors, demographic characteristics, disease subtypes or severity, physiological parameters, or different treatments. Among the various methods used to generate virtual populations, statistical approaches are commonly used to infer distributions of model parameter estimates, while Monte Carlo simulation is employed to randomly sample parameter values from these distributions. This general framework inspired several effective methods, such as those in [5, 6], which incorporate biologically inspired constraints to generate more realistic virtual populations. Also in this case, the choice of the method depends on available data, model complexity, computational resources, and multiple methods that may be combined for comprehensive representation.

3. Applications in QSP modeling

Virtual populations in QSP modeling can be used to optimize dosing regimens in subpopulations, considering factors like age, gender, disease severity, and comorbidities. They provide insights into expected variability in clinical trials, allowing researchers to anticipate diverse patient responses and design robust trials. Here, we report two examples, studying rare or hard-to-reach populations, such as patients with rare and neurodegenerative diseases, and children. In these cases, virtual populations can help to understand the physiological processes and the effects of treatments.

In the works [7, 8], a quantitative systems pharmacology (QSP) model was developed to examine neurofilament trafficking in both healthy individuals and patients with spinal muscular atrophy (SMA). SMA is a rare genetic disorder affecting motor neurons and resulting in muscle weakness and atrophy, which is most impactful in pediatric ages, with high severity and lethality. Neurofilaments, which play a role in neuronal structure and transport, serve as biomarkers for disease progression in SMA and other neurodegenerative conditions.

The QSP model captures the dynamics of neurofilament production, transport, degradation, and release throughout the nervous system and periphery. It combines a previously established mathematical model of neurofilament trafficking in healthy adults [7] with detailed characterizations of pediatric physiological processes, such as organ volumes and postnatal nervous system development. The authors used data from clinical trials that enrolled SMA patients treated with nusinersen to build a virtual population of untreated subjects. The parameters related to the treatment were estimated by fitting individual time series of SMA patients followed during the treatment. The combination of the estimated treatment parameters and the untreated virtual populations allowed the generation of virtual populations of treated patients with different ages and treatment protocols. The predicted neurofilament concentrations of these populations were compared to the time series of SMA patients reserved for model validation. This validation strategy confirmed the predictive performance of the model, making it a valuable tool for investigating neurofilament as a biomarker for disease progression and treatment response in SMA.

Another promising application is in Gaucher disease type 1 (GD1), a rare genetic disorder that can lead to significant morbidity and mortality through clinical manifestations such as splenomegaly, hematological complications, and bone disease. In [9], the authors developed a QSP model to represent the effects of eliglustat, a novel substrate reduction therapy for GD1, on treatment-naïve or enzyme replacement therapy (ERT) stabilized adult GD1. The model was informed by a data-driven genotype-phenotype relationship to represent a wide spectrum of patients with mild to severe GD1, even beyond the initial patients. The model was applied to predict ERT and eliglustat responses in virtual populations of adult patients with GD1, representing patients across a spectrum of disease severity as defined by genotype-phenotype relationships. The QSP model and the virtual populations provide a mechanistic platform for predicting different treatment responses within the heterogeneous GD1 patient population.

As shown, digital twins and virtual populations can play a crucial role in advancing QSP by addressing patient heterogeneity, optimizing treatment strategies, informing clinical trial design, and enhancing our understanding of complex biological systems. As more accurate methods emerge for characterizing virtual patients and digital twins, research is steadily moving toward the realm of *in-silico* modeling. This progress can change drug development, paving the way for increasingly personalized medicine approaches.

REFERENCES

1. Azer, K., et al. History and future perspectives on the discipline of quantitative systems pharmacology modeling and its applications, *Frontiers in Physiology*, **12**, 127, (2021).

2. Marshall, S. F., et al. Good practices in model-informed drug discovery and development: Practice, application, and documentation, *CPT: Pharmacometrics and Systems Pharmacology*, **5**, 93–122, (2016).

3. Sher, A., Niederer, S. A., Mirams, G. R., Kirpichnikova, A., Allen, R., Pathmanathan, P., Gavaghan, D. J., van der Graaf, P. H. and Noble, D. A quantitative systems pharmacology perspective on the importance of parameter identifiability, *Bulletin of Mathematical Biology*, **84**, 1–15, (2022).

4. Reali, F., Priami, C. and Marchetti, L. Optimization algorithms for computational systems biology, *Frontiers in Applied Mathematics and Statistics*, **3**, (2017).

5. Allen, R. J., Rieger, T. R. and Musante, C. J. Efficient generation and selection of virtual populations in quantitative systems pharmacology models, *CPT: Pharmacometrics and Systems Pharmacology*, **5**, 140–146, (2016).

6. Cheng, Y., Straube, R., Alnaif, A. E., Huang, L., Leil, T. A. and Schmidt, B. J. Virtual populations for quantitative systems pharmacology models., *Methods in molecular biology*, **2486**, 129–179, (2022).

7. Paris, A., et al. An age-dependent mathematical model of neurofilament trafficking in healthy conditions, *CPT: pharmacometrics & systems pharmacology*, **11** (4), 447—457, (2022).

8. Paris, A., et al. A pediatric quantitative systems pharmacology model of neurofilament trafficking in spinal muscular atrophy treated with the antisense oligonucleotide nusinersen, *CPT: Pharmacometrics & Systems Pharmacology*, **12** (2), 196–206, (2023).

9. Abrams, R., et al. A quantitative systems pharmacology model of gaucher disease type 1 provides mechanistic insight into the response to substrate reduction therapy with eliglustat, *CPT: Pharmacometrics & Systems Pharmacology*, **9**, 374–383, (2020).

TOWARDS INTELLIGENT URBAN DECISION SUPPORT: COGNITIVE DUALITY AND DIGITAL TWINS

Riccardo De Benedictis, Gloria Beraldo, Amedeo Cesta, *and* Gabriella Cortellessa
Institute of Cognitive Sciences and Technologies (ISTC) - National Research Council of Italy (CNR), Rome, Italy
e-mail: riccardo.debenedictis@istc.cnr.it

Urban intelligence is an emerging research area focused on leveraging advanced technologies and data analysis techniques to enhance urban efficiency, sustainability, and livability. Decision support plays a crucial role within urban intelligence, utilizing automated reasoning to both plan activities on the urban landscape and react to its dynamic changes. Drawing inspiration from dual-process cognitive theories, this paper explores the integration of automated planning and rule-based systems to facilitate decision-making in urban management.

Keywords: *timeline-based planning, plan execution, urban intelligence, decision support system*

---

1. Introduction

Urban areas are intricate and ever-changing, demanding continuous management to uphold their efficiency, sustainability, and livability. With the expansion and diversity of urban populations, the complexities of urban management intensify [1, 2]. In recent times, there has been a growing interest in utilizing advanced technologies and data analysis techniques to bolster decision-making in urban management. Urban intelligence, an emerging and evolving research domain, aims to harness these advancements to enhance the effectiveness, sustainability, and livability of urban areas [3]. Urban intelligence encompasses decision support as a crucial element, enabling informed choices for urban development and management. A specific type of decision support system (DSS) called a cognitive decision support system (CDSS) incorporates cognitive theories and models of human decision-making to aid in complex decision-making tasks [4]. The dual processing theory [5], among the various cognitive theories available, has found widespread application in the development of CDSS [6, 7]. This theory proposes two distinct cognitive systems known as System 1 and System 2. System 1 entails fast, automatic, and intuitive thinking relying on heuristics, mental shortcuts, and past experiences to make judgments and decisions. It operates at an unconscious or intuitive level being involved in perception, pattern recognition, and emotional responses. On the other hand, System 2 involves slower, deliberate, and analytical thinking based on logic, reasoning, and conscious effort to make judgments and decisions. It operates at a conscious or rational level and is engaged in problem-solving, planning, and decision-making activities. Inspired by the dual processing theory, this paper presents COCO (COmbined deduction and abduCtiOn logic reasoner), a cognitive architecture that serves as a digital twin of urban decision makers. COCO integrates a rule-based system to replicate the characteristics of System 1 and a timeline-based planner, enhanced with semantic reasoning capabilities, to emulate the cognitive processes of System 2. By embodying the decision-making traits of urban decision makers, COCO aims to enhance decision-making in urban management, allowing the simulation and analysis of various scenarios, conducting what-if analyses, and even simulating suboptimal decision-making behaviors.

2. Thinking, Fast and Slow, Logically: COCO

Rule-based systems are a type of Artificial Intelligence (AI) systems that use predefined “if-then” rules to make fast and reliable decisions in various situations. For instance, a rule like “IF temperature is above 30 degrees Celsius AND humidity is above 70%, THEN activate the sprinkler system in the park” exemplifies a rule-based system. By introducing facts such as "current temperature is 35 degrees Celsius" and "current humidity is 75%", the system activates the corresponding rule and executes the action. These systems store rules and facts in a knowledge base, allowing quick assessment and response without extensive analysis. The transparency of rule-based systems enhances trust and accountability in decision-making. COCO adopts the CLIPS[1] rule-based system to recognize familiar patterns and make intuitive decisions, similar to System 1 processes.

Automated planning is another field of AI focused on creating computer programs that generate plans and strategies to accomplish specific goals or tasks [8]. One approach within automated planning is timeline-based planning [9, 10], in which activities are organized over timelines, considering temporal constraints and dependencies. Such activities are represented through tokens, consisting of a predicate symbol (denoted as n) and parameters (denoted as $x_0, \ldots, x_i$), which can include temporal details. The tokens can be classified as either facts or goals. Rules play a crucial role in achieving goals. They consist of a head (e.g., $n\left(x_0, \ldots, x_k\right)$) and a body ($\mathbf{r}$), defining the requirements (i.e., other tokens, token constraints, conjunctions or disjunctions of requirements) for reaching a specific goal. Tokens can also be associated with timelines, identified by a special object variable τ. Different tokens with the same τ value belong to the same timeline and may interact with each other based on the nature of the timeline. There are different types of timelines, such as state-variable timelines, where tokens on the same state-variable cannot overlap temporally, and reusable-resource timelines, where tokens representing resource usages can overlap as long as the concurrent usage remains within the resource's capacity. In timeline-based planning, the objective is to find a set of tokens that satisfy all the constraints, rules, and a requirement defined within the planning problem. This is typically achieved through a process of reasoning and decision-making, similar to the deliberate and intentional decision-making of System 2 thinking.

To enable effective decision-making in dynamic environments, the COCO system integrates a rule-based system as a sequencing (reactive) tier, and a timeline-based planner[2] as a deliberative tier, applying production rules to deduce new information or actions while reasoning backward to find action sequences that achieve goals. Notably, the sequencing tier has the capability to interact with both the environment and introspectively with higher-level reasoning. The intrinsic motivations and higher-level tasks generated during plan execution by the deliberative tier serve as suggestions rather than mandatory components for the system's autonomy, influencing the choice of actions made by the sequencing tier. By combining the reactivity and explainability of the rule-based system with the planning capabilities of the timeline-based planner, this integration harnesses the strengths of both components. The architecture, depicted in Figure 1, consists of a deliberative tier for plan generation and adaptation, a sequencing tier for executing actions, and sensing and controlling tiers for interpreting sensor data and actuator commands. By functioning as a digital twin of the urban decision makers, this integration provides a comprehensive support by addressing System 1 tasks such as abstraction and low-level command generation, as well as System 2 tasks including reasoning and high-level plan adaptation based on dynamic environmental information. This enables the system to facilitate what-if analysis and offers decision makers a powerful tool to simulate and evaluate different scenarios, thereby enhancing their decision-making capabilities in urban management.

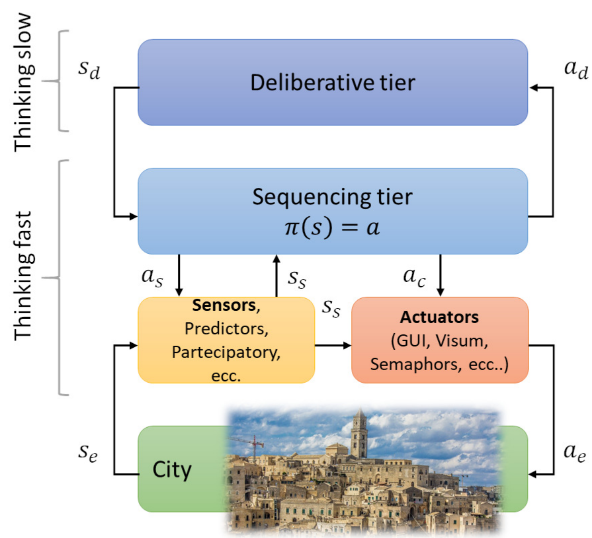


Figure 1: The COCO three-layer architecture.

[1]https://www.clipsrules.net
[2]https://github.com/ratioSolver/oRatio

3. Conclusions

The COCO system, introduced in this paper, combines a rule-based system and automated planning to support decision-making in urban management. Rule-based systems excel at responding to environmental changes, while automated planning generates task-oriented plans for achieving desired goals. By merging these two approaches, urban managers and decision-makers gain a comprehensive and efficient solution for managing diverse aspects of urban life. The rule-based system utilizes indexing techniques to generate reactive behaviors efficiently, while the deliberative component handles scenarios, plan adaptation during execution, and what-if analyses, despite its computational demands. The effectiveness of the COCO system has been initially evaluated by assessing the efficiency of the reasoners' resolution processes. This evaluation has involved benchmark problems of increasing complexity to estimate the resolution times as the problems grew in size. The goal was to measure how quickly COCO could respond when urban decision-makers used the system to analyze and modify generated solutions by adding additional constraints for what-if analyses. Impressively, even with the exponential complexity of the problem, the resolution times consistently remained within a few seconds, even with larger numbers of activities. It's important to note that both System 1 and System 2 heavily rely on defining rules, which can be a challenging and time-consuming task. To simplify and automate this process, future efforts will explore machine learning techniques such as induction and decision tree learning.

ACKNOWLEDGMENTS

This work is partially supported by the project "House of emerging technologies of Matera" (CTE-MT) funded by the Ministry of Economic Development of Italy, CUP I14E20000020001.

REFERENCES

1. Moreno-Monroy, A. I., Schiavina, M. and Veneri, P. Metropolitan areas in the world. delineation and population trends, *Journal of Urban Economics*, **125**, 103242, delineation of Urban Areas, (2021).

2. OECD and Commission, E., *Cities in the World* (2020).

3. Castelli, G., et al. Urban intelligence: a modular, fully integrated, and evolving model for cities digital twinning, *2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT)*, pp. 033–037, (2019).

4. Niu, L., Lu, J. and Zhang, G., *Cognition-driven decision support for business intelligence: models, techniques, systems and applications / Li Niu, Jie Lu, and Guangquan Zhang*, Springer Verlag Berlin (2009).

5. Kahneman, D., *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York (2011).

6. Arnott, D. and Gao, S. Behavioral economics for decision support systems researchers, *Decision Support Systems*, **122**, 113063, (2019).

7. Tsalatsanis, A., Hozo, I., Kumar, A. and Djulbegovic, B. Dual processing model for medical decision-making: An extension to diagnostic testing, *PLoS One*, **10** (8), (2015).

8. Ghallab, M., Nau, D. and Traverso, P., *Automated Planning: Theory and Practice*, Morgan Kaufmann Publishers Inc. (2004).

9. Muscettola, N., (1994), HSTS: Integrating Planning and Scheduling. Zweben, M. and Fox, M.S. (Ed.), *Intelligent Scheduling*, pp. 169–212, Morgan Kauffmann.

10. De Benedictis, R. and Cesta, A. Lifted Heuristics for Timeline-based Planning, *ECAI-2020, 24th European Conference on Artificial Intelligence*, Santiago de Compostela, Spain, August, pp. 498–2337, (2020).

HUMAN AND SOCIAL ASPECTS IN THE DEVELOPMENT OF URBAN DIGITAL TWINS

Roberto Malvezzi *and* Giordana Castelli
Engineering, ICT and Technology for Energy and Transportation Department (DIITET), CNR, Rome, Italy
email: roberto.malvezzi@amministrazione.cnr.it

The CNR strategic project "Urban Intelligence" (UI) pursues the development of Urban Digital Twins (UDT) able to support the urban governance by integrating the various dimensions of a city, and thus by overcoming the limitations of the smart city paradigm still dominated by sectoral approaches. A relevant asset for UI relies in the capacity of integrating "community knowledge" within the UDT development, with the aim of completing the objective and top-down information typical of the urban sciences, and of extending range, scope and depth of the operations carried out in the UDT (e.g. analysis, simulation, prevision, optimization, decisions support). To this end, UI adopts a participatory method rooted in the "community mapping" theory, and evolves it in the implementation of social innovation processes aimed on the one hand, at exploring community knowledge, and on the other, at empowering the capacity of communities of tapping it for envisioning future scenarios. In particular, a hybrid "phygital" participatory approach is adopted, where analogical engagement for-mats are intertwined with innovative tools for GIS-based community surveys. Finally, a description is provided of the potential contributions of the participatory domain of UI to the UDT development, with special regards to ICT innovations.

Keywords: *urban digital twin, community mapping, stakeholders engagement, experiential knowledge*

1. Motivation and problem addressed

The making of an Urban Digital Twin (UDT) should rely on a direct involvement of the local community, to the extents that the inhabitants of a city are also elements of its physical system, actors and users which determine the success or failure of its functions, sources and receptors of urban dynamics, knowledge, information and services. The strategic project "Urban Intelligence" (UI) of the CNR, by conceiving UDTs as cyber-physical-social systems [1], requires a full connection and integration among these three dimensions.

Paba tells the experiences of Patrick Geddes, one of the founders of the participatory approach in urban planning, with special reference to the "Outlook Tower" [2]: "Geddes invited its visitors to quickly climbing the five flights of stairs, so to reach the basis of the octagonal tower, and to directly climb another wooden stair therein which led to the small terrace on the top, from which it was finally possible to gain a 360° overview on the city from a 80 feet height. This was the first grade of interpretation: the city vision from the height, synoptic, holistic, the city seen in its entirety, from each part towards the horizon". This was followed by a visit to a darkroom, where the city was seen through its image reflected upon a concave mirror. Thanks to a lever it was possible to change the angle of the mirror, and observe the city under every perspective.

The art of city surveying entails the merging of the "top view", which attains the objective knowledge as it may come from sources such as data fluxes of sensor networks, with the "bottom view", which wanders across streets listening to people and collecting stories and emotions. Between these two levels a substantial epistemological difference arises, and thus, of approach. Limiting the exploration to a perspective from the top, or in "third person", induces a behaviourist approach aimed at modeling the community life by means of a mere description of what is externally perceived, and favours the use of Cartesian-like maps which do not correspond to the experience of the city as it is seen in "first person". Assuming such "bottom-up" approach leads to take in a higher consideration the substratum of senses and meanings attributed by people to the spaces of their experiences, which are caught through a ground perspective constrained by the horizon of the visible. Moving within this mode, people develop “subjective geographies” which, according to [3], are «informed by background knowledge organized into patterns. We rely on these schemas or pattern to acquire new knowledge». Such schemas, known also with the term of ”cognitive maps”, ”are not just a set of spatial mental structures denoting relative position, they contain attributive values and meanings” [4] which pre-orient our relation with the external world, causing potential mismatches with previsions overlooking them [5]. Some relevant indications for the development of UDTs stem from this analysis: first of all, their role as tools for empowering local communities in their pursue of a sense of identity, intended as an essential precondition for consciously shaping their city’s future. As underlined by [6], «the re/construction of an image is just apparently a frivolous activity: it involves various symbolic, cultural and even political codes, influencing the wellbeing of who is self-portrayed, but also of those who are in the presence of the image, without being able to intervening on it.» The places of a city are not just filled by the stories occurring in them: their physical settings affect the life of the inhabitants, by opening or closing possibilities, by favouring or hampering connections, by offering or denying opportunities. There is a close link between the physicality of a place, its social value, and its usage. Crucial to this objective is thus the involvement of communities from the very first phases of development of the UDTs, aimed at building an experiential knowledge of the city with three goals: (i) highlighting the modes through which people live their city and assign values to it; (ii) enhancing interactions among local actors in order to generate shared added-values, to co-create shared scenarios of urban evolution, and to activate new social practices for enhancing their implementation. On these premises, Urban Intelligence pursues the transition from a “city-of-use” toward a “city-of-sense” where the human factor is put at the heart of the urban planning and programming processes.

2. Methodology and results

The participatory methodology adopted for UI sees in the development of UDTs a ”social innovation process” aimed at empowering local communities to govern their future. This objective can be pursued by running “open-ended processes” where participants undergo some kind of changes about their current practices, aims or modalities [7] and where sustainability of the results achieved is pursued [8]. To this aim, the two fundamental concepts of seeding and infrastructuring are followed, already introduced by the Design-Driven Innovation Studies (DDIS) [9]. ”Seeding” refers to the metaphor introduced by Manzini [10] for underlining how the social innovation solutions act just like fragments of a “social sourcing code” which need to be cultivated accurately and adaptively in order to verify the kind of fruits they can produce in a given context. ”Infrastructuring” refers to the necessity of coordinating the “innovation seeds” within solid process structures; this concept can be traced back to the Meta-Design Studies (MDS), which are focused not on the final output, but on a reflection about the very process architecture which could guide toward the best outputs [11], as well as to Giddens’ ”institutionalism” [12], according to which enduring social interactions end up generating recurrent patterns and clusters. These two concepts offer the idea of ”process” as an immaterial infrastructure, the constitution of which is an autonomous result of participation, which can be spent for multiple issues and goals. Structuring an innovation process is thus a primary design challenge for the development of an UDT, at the core of which it should be placed the activation and management of connections between actors [13] aimed at establishing an enduring space of mutual dialogue [14]. In particular, UI adopt the ”community mapping” principles for understanding cities, territories and communities as socio-economic integrated ecosystems, each characterized by a specific DNA [15]; such mapping engages the various social groups and networks of a community by implementing a set of dedicated participatory approaches aimed at achieving the following specific objectives:

- exploring the city from surveying its material and immaterial heritage [16, 17];
- awakening the latent knowledge of citizens starting from the listening of persons;
- describing the vision of a city and the urban values dispersed in the stories linked to places of life;
- reading the spatial network of the sense connections for representing new urban geographies;
- building a circularity between the social actions and their cultural representations.

From an operative point of view, a hybrid cyber-physical participatory approach is performed [18], based on an interaction between analogical and digital tools allowing for both on-line/off-line, synchronous/asynchronous engagement activities. Among the digital tools, of particular interest are the PPGIS (Public Participation GIS), which enable the development of GIS databases (e.g.. shapefiles) annotated through metadata, and thus, a direct integration of the participatory knowledge with the other knowledge levels of the UDT.

3. Contribution to the field

Many are the potential contributions that the research and experimentation on the experiential knowledge can offer to the technological development of the UDTs and more in general, to the advancement of the strategic process ”Urban Intelligence”, with special regards to the following ICT fields:

A. Data for the construction of a knowledge base of the UDT extended to the urban community: definition of typologies and formats for participatory data and metadata in view of their integration with the UDT tools and services; development of innovative representations the ”urban structures of sense” as the immaterial heritage.

B. Tools for the management of the variegated thematic fields addressed y the UDT: development of tools conceived for supporting hybrid participatory approach aimed at enlarging the community engagement and at enhancing the efficacy of the social innovation process; co-design paths of the use-cases aimed at ideating and calibrating the tools for the UDT for a multiplicity of profiles, and at increasing their diffusion, inclusion and usability towards a wide range of potential users (from technicians, to enterprises, to citizens); development of user interfaces shaped on the needs of the different potential users; immersive 3 exploration of the city, including community contents, and related to both the actual status, as well as evolutive and transformative proposals.

C. Services in support to planning programming and governing the city: multi-disciplinary approaches or the analysis, simulation, prevision and optimization of urban dynamics, integrating community behaviors, experiences and expectations, social and cultural assets; decision support systems aimed to individuating optima scenarios of intervention reflecting the preference structures of the different actors involved in the urban governance.

REFERENCES

1. Tomko, M. and Winter, S. Beyond digital twins–a commentary, *Environment and Planning B: Urban Analytics and City Science*, **46** (2), 395–399, (2019).

2. Paba, G., et al. Dall’outlook tower alla casa della città, *La nuova città*, **1**, 4–7, (2013).

3. Gallagher, S., *The inordinance of time*, Northwestern University Press (1998).

4. Kitchin, R. M. Cognitive maps: What are they and why study them?, *Journal of environmental psychology*, **14** (1), 1–19, (1994).

5. Malvezzi, R. Per un’urbanistica cognitiva: il percorso d’ascolto per il documento preliminare d’indirizzo di borbona, *Per un’urbanistica cognitiva: il percorso d’ascolto per il documento preliminare d’indirizzo di Borbona*, pp. 113–122, (2021).

6. Marson, A., *Archetipi di territorio*, Alinea Editrice (2008).

7. Meroni, A., et al. Strategic design: where are we now? reflection around the foundations of a recent discipline, *Strategic design research journal*, **1**, 31–38, (2008).

8. Murray, R., Caulier-Grice, J., Mulgan, G., et al., *The open book of social innovation*, vol. 24, Nesta London (2010).

9. de Mello Freire, K., Del Gaudio, C. and Franzato, C. Design-driven strategies for creative social innovation ecosystems, *International Journal of Knowledge Engineering and Management*, **6** (16), 46–69, (2017).

10. Manzini, E., (2008), *On Service Design, Presentation held at the Service Design Symposium, Copenhagen, CIID*.

11. Giaccardi, E. Metadesign as an emergent design culture, *Leonardo*, **38** (4), 342–349, (2005).

12. Giddens, A., (2004), The constitution of society: Outline of the theory of structuration: Elements of the theory of structuration. *Practicing History*, pp. 121–142, Routledge.

13. Hillgren, P.-A., Seravalli, A. and Emilson, A. Prototyping and infrastructuring in design for social innovation, *CoDesign*, **7** (3-4), 169–183, (2011).

14. Cohen, M. D., Riolo, R. L. and Axelrod, R. The role of social structure in the maintenance of cooperative regimes, *Rationality and Society*, **13** (1), 5–32, (2001).

15. Magnaghi, A., *Montespertoli. Le mappe di comunità per lo statuto del territorio*, Alinea Editrice (2010).

16. UNESCO, (2003), *Convenzione per la salvaguardia del patrimonio culturale immateriale*.

17. EUROPE, C. O., (2005), *Convention on the Value of Cultural Heritage for Society (Faro Convention)*.

18. Malvezzi, R. and Castelli, G. Gemelli digitali urbani per lo sviluppo di comunità partecipanti: il caso di matera, *Atti della XXIV Conferenza Nazionale SIU - Società Italiana degli Urbanisti, Planum, 5, 60-66*, (2023).

AN IT PLATFORM FOR MANAGING AND SHARING DATA AMONG DIGITAL TWINS OF AN URBAN INTELLIGENCE SYSTEM

Stefano Silvestri *and* Mario Ciampi
National Research Council of Italy, Institute for High Performance Computing and Networking, Via Pietro Castellino, 111 - 80131 Naples, Italy
e-mail: {stefano.silvestri, mario.ciampi}@icar.cnr.it

This paper describes a novel architecture devoted to the implementation of an IT platform for the deployment and integration of the different Digital Twin subsystems that compose a complex Urban Intelligence system. The proposed IT platform has the following main purposes: i) facilitating the deployment of the subsystems in a cloud environment; ii) effectively storing, integrating, managing, and sharing the huge amount of heterogeneous data acquired and produced by each subsystem using a data lake; iii) supporting data exchange and sharing; iv) managing and executing workflows, to automatically coordinate and run processes; and v) provide and visualize the required information. A prototype of the IT platform has been implemented leveraging open source frameworks, allowing to test its functionalities and performance. The results of the tests confirmed that the proposed architecture can efficiently and easily support the deployment and integration of heterogeneous subsystems, making them able to share and integrate their data and to select, extract, and visualize the information required by a user, also promoting the integration with other external systems.

Keywords: *urban intelligence, IT platform, data lake, data management*

1. Introduction

A comprehensive Urban Intelligence (UI) system for a smart city includes multiple virtual representations of heterogeneous subsystems (Digital Twins, DTs) that model complex interacting assets, infrastructures, and social aspects of the city, such as mobility, social and computer networks, weather, pollution, tourist flow, and so on [1, 2]. To allow a correct functioning of the DTs according to an integrated approach, several IT paradigms and technologies should be integrated in an UI system, such as dedicated IoT sensor networks, network infrastructures, information systems, high-level web services, cloud or edge-computing environments [3], able to collect and manage the heterogeneous data of the city and provide the hardware and software infrastructures to implement and run the required models and analytics, visualization, and user interfaces functionalities [4, 5].

In light of such considerations, the complexity and heterogeneity of the DTs require an advanced enabling IT platform that facilitates their deployment, enables their intercommunication and integration, provides data management and sharing capabilities, supports information searching and retrieving, and offers visualization and user interface capabilities, allowing to fully integrate the DTs.

In this contribution, the architecture of an IT platform able to support the management and integration of DTs for UI systems is presented. The proposed IT platform has the following main purposes: i) facilitating the deployment of the DTs in a cloud environment; ii) effectively storing, integrating, managing, and sharing the huge amount of heterogeneous data acquired and produced by each DT; iii) integrating and supporting data exchange and sharing; and iv) managing workflows, to automatically execute the requested processes and provide and visualize the required information. A prototype of the proposed architecture has been implemented, relying on a cloud infrastructure, a data lake which acts as both a data collector and repository, a set of specifically designed APIs, a workflow engine, and a user interface. The preliminary experiments showed that the proposed IT platform can be effectively adopted for the deployment in real UI projects.

2. IT Platform Architecture

The layers of the proposed IT platform architecture are depicted in Figure 1.

In the lower layer, the *Cloud computing* layer provides and dynamically manages and coordinates the hardware and software resources in the cloud environment. This layer is also devoted to container deployment and orchestration, facilitating the installation and execution of containerized software modules that implement each user DT subsystem in any cloud-based environment, as well as facilitating their integration with other systems.

The next layer is the *Data lake* layer, whose purpose is to acquire, store, integrate, query, and retrieve data collected or produced by any UI subsystem, as well as obtained from external sources. The data lake is based on a Distributed File System and a NoSQL database, able to manage and process structured, semi-structured, and unstructured raw data. To the end of preventing the data swamp degeneration, a relational database technology is integrated into the data lake, with the purpose of enriching the heterogeneous data with specific metadata and a catalog. This way, it supports their retrieval and integration, enriches the structured and unstructured data with additional information, and offers basic Extract, Transform, and Load (ETL) functionalities for data aggregation. Moreover, the data lake layer also acts as a common data collector for the DTs, allowing them to easily share their own data and communicate with the other subsystems. Asynchronous notification services are included in this layer, to the end of notifying specific modules of possible data updates in the data lake according to the publish/subscribe pattern, thus optimizing bandwidth and performances.

The *Workflow engine* layer includes a workflow management and execution component aimed to provide the UI system with an automatic business process engine, whose purpose is to easily implement and execute the workflows that are used to model each analysis requested, where the interaction of more subsystems needs. The workflow engine coordinates the process and data flows through the various subsystems, in order to obtain the required information or perform the requested operation.

The upper *Data visualization and user interface* layer provides the visualization and user interface features (including dashboards), adopting a web-based approach.

Finally, the communication with the data lake and the other UI subsystems exploits the transversal *Communication interface* layer, where a set of dedicated APIs are available to provide common and standard way to get or send data and/or access the exposed functionalities. The Communication interface layer also offers the possibility to connect the IT platform with external systems of the city (like local databases or IT services) and the authentication and authorization functionalities, able to implement single-sign on through standardized claims.

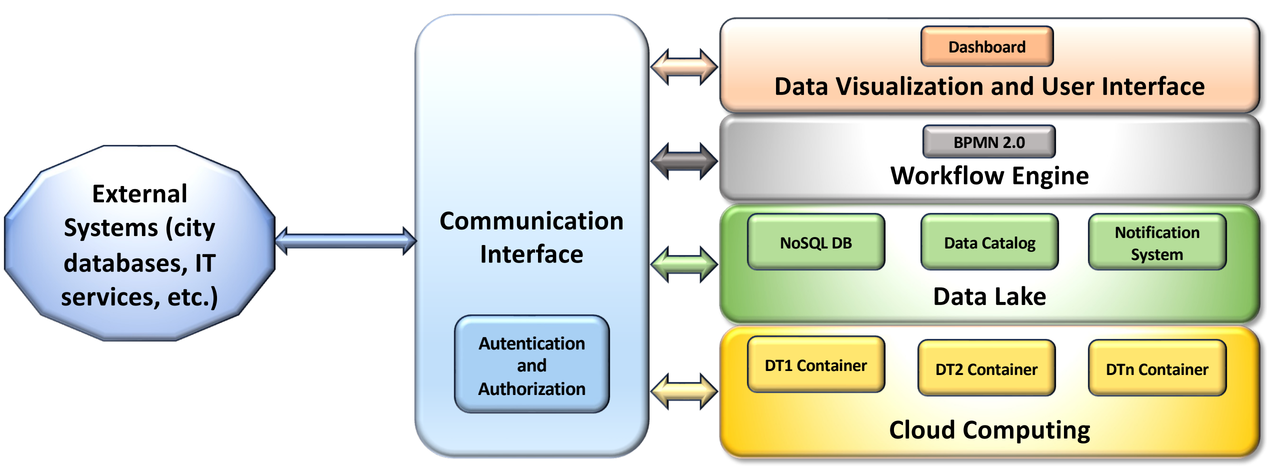


Figure 1: Layers of the proposed IT platform architecture.

3. Implementation and Testing Details

A prototype of the proposed architecture has been implemented exploiting different open source frameworks.

The cloud environment for DTs deployment adopts Docker for module containerization and Kubernetes for their orchestration. This approach makes the IT platform easily deployable, customizable, and embedded within different environments and other external systems.

The data lake and its data catalog are implemented using a MongoDB NoSQL database and a GridFS distributed file system (the latter is devoted to the storage of large raw data). These frameworks ensure high-performance levels, scalability, and the capability of managing very large and heterogeneous data collections. Moreover, the NoSQL database does not require a strict data schema, allowing a simple extension of data models and the catalog. The asynchronous notification of data updates has been implemented adopting a publish/subscribe approach based on Message Queue Telemetry Transport (MQTT) standard protocol, allowing the selection of data that requires notification and the corresponding subscriber modules that receive such notifications. The interfaces to the data lake layer have been developed as RESTful APIs on the HTTPS application protocol, leveraging Python Flask framework and PyMongo library. The authentication and authorization mechanisms for the data lake and other modules use claims based on Java Web Token, leveraging the OAuth 2.0 and OpenID Connect protocols.

The workflow engine layer leverages a Business Process Management engine, which allows the workflows to be easily designed according to the BPMN 2.0 standard. Finally, the visualization and user interface layer is obtained by using a web-based data visualization software developed in JavaScript.

The implemented prototype made it possible to carry out some preliminary tests, exploiting the data and subsystems of two Italian UI projects in the cities of Matera and Catania. The functional and performance tests carried out demonstrated that the proposed IT platform can acquire from the sensor network of a city high rates of data, with a rate of 5 minutes, sharing them at the same time among the UI DTs. It is also able to store and integrate the huge volume of data from both internal and external sources. RESTful interfaces effectively facilitate intercommunication and data sharing between UI subsystems, while workflows can be easily configured and executed.

4. Conclusions

This paper presented a general architecture of a novel IT platform that allows for the deployment and integration of DT subsystems of an UI system. The IT platform leverages the cloud environment and container technology for DTs deployment, adopts a data lake approach and dedicated APIs to collect, store, retrieve, integrate, and share the data produced inside and outside the UI system. It also includes a workflow engine to design and execute the requested processes, a visualization tool, and a user interface. A prototype of the architecture has been implemented and tested in real UI projects, demonstrating that it effectively supports the deployment and the integration of DT subsystems for UI.

REFERENCES

1. Deng, T., Zhang, K. and Shen, Z.-J. M. A systematic review of a digital twin city: A new pattern of urban governance toward smart cities, *Journal of Management Science and Engineering*, **6** (2), 125–134, (2021).

2. White, G., Zink, A., Codecá, L. and Clarke, S. A digital twin smart city for citizen feedback, *Cities*, **110**, 103064, (2021).

3. Habibzadeh, H., Kaptan, C., Soyata, T., Kantarci, B. and Boukerche, A. Smart city system design: A comprehensive study of the application and data planes, *ACM Comput. Surv.*, **52** (2), (2019).

4. Allam, Z. and Dhunny, Z. A. On big data, artificial intelligence and smart cities, *Cities*, **89**, 80–91, (2019).

5. Han, Q., Nesi, P., Pantaleo, G. and Paoli, I. Smart city dashboards: Design, development, and evaluation, *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, pp. 1–4, (2020).

TOWARDS IMMERSIVE URBAN DIGITAL TWINS

Brigida Bonino, Daniela Cabiddu, Michela Mortara, Katia Lupinetti *and* Simone Pittaluga
CNR-IMATI, Genova, Italy
e-mail: brigida.bonino@ge.imati.cnr.it

Urban digital twins are virtual representations of physical systems, subsystems, and processes of a city. The construction of a three-dimensional geometric model of the city provides the basis for the development of an urban digital twin. On top of the 3D model, heterogeneous data describing or simulating one or more systems and subsystems can be embedded and visualized. In this context, we focus on the possibility to exploit Virtual Reality (VR) technologies to enhance the visualization of these 3D representations and to provide an innovative immersive approach to interact with urban digital twins. We describe a pipeline for the design of a 3D model of an urban area suitable for an immersive visualization and a natural interaction so that the users can have a sense of reality as if they are in the real world. Preliminary results are shown related to our case study in Matera, in Southern Italy.

Keywords: *virtual reality, 3D modeling, 3D visualization*

1. Introduction

In the context of creating an Urban Digital Twin, the geometric representation of the morphology and physical features (either built or natural) of the city, possibly annotated with additional contextual information, has a high potential in describing, monitoring and/or predicting urban processes [1, 2]. Focusing on smart tourism management, for instance, the geometric model can be useful to design a virtual visit to non-accessible sites (due to overcrowding, closing hours, or visitor disabilities) or to show a preview for the tourist to select the must-see spots in their day visit, simulate emergency egress, plan temporary events according to the capacity and morphology of the site, etc. Another example is represented by the public engagement, indeed an accessible 3D model of the city can foster citizen engagement and participation. People can virtually explore the city, understand proposed developed projects and provide constructive feedback. This creates a more inclusive and transparent decision-making process, which allows citizens to have an active voice in building their community.

Currently, urban digital representations are still visualized on 2D displays, providing a limited perception of the real tridimensionality of the environment. The availability of realistic 3D representations of real objects embedded with heterogeneous data (i.e. colors, normals, semantically relevant attributes) opens the doors to new visualization methodologies, conveying the 3D content and its associated information in an intuitive way. In this context, Virtual Reality (VR) technology has gained much interest in the last years, since it allows innovative and intuitive ways to explore and interact with 3D models [3].

In this contribution, we focus on the design and development of a virtual immersive visit of an urban area, starting from the data acquisition survey to the 3D model generation and its embedding into a Virtual Reality (VR) environment. Our pipeline seems promising in strengthening the visualization capabilities of digital twins and providing realistic experiences and natural interactions.

2. Methodology

Our main long-term objective is to immerse individuals in the city digital twin with a 1:1 scale, and thus allow the users to explore and learn about an urban environment by moving in it as if they are in the real city.

To this aim, the first step is generating a 3D surface representation of the city to be embedded in VR. Our 3D generation method assumes the input point cloud to represent the area of interest with a high level of details and associated colors. This is necessary in order to guarantee an authentic experience with realism. When such a representation is not available, data acquisition surveys are performed by exploiting proper technologies able to capture a wide area at fine geometric and appearance detail [4]. Since point clouds coming from LiDAR/laser scanning and photogrammetric technologies are often subjected to noise and other types of artifacts and defects, a pre-processing pipeline is often necessary to remove undesired elements. Acquisitions performed by these techniques return dense point clouds (i.e. resolution is often of one point per millimeter). To use these representations in VR, it is important to choose a good trade-off between the high level of detail and realism and the complexity of the model. The pre-processed point cloud is then given as an input to triangulation methods (such as the well-known marching cubes algorithm [5]). The result is a 3D model representing the morphology of the real world enriched with color attributes on vertices, where the color is an interpolation of neighbor points in the decimated point cloud.

Once the 3D digital model has been reconstructed, its exploration in VR arises two main challenges. On the one hand, the overcome of the issues associated with importing a georeferenced 3D model in a game engine (e.g. Unity) to correctly position and orient it, as well as give the model appropriate colors or textures to ensure realism. On the other hand, the development of natural and intuitive navigation techniques by which users can freely navigate in the reconstructed environment in a natural way. We addressed the former issues, specifically related to repositioning the 3D model into Unity world. Our solution takes into account the different coordinate system used in Unity, that is a left-handed, Y-Up coordinate system. Since Unity does not properly support geographic coordinates, the 3D model is rigidly translated so that the center of its bounding box is centered on the origin of the Unity coordinate system. Then, our approach tries to identify the terrain level in the model in order to properly set the camera position. Notice that, as for the appearance of the 3D model, when imported in Unity it is transparent by default. It is thus first necessary to assign the built-in material and set the occlusion properties.

3. Preliminary Results

As a preliminary result, we describe the development of a virtual replica of the main square (well-known as "Piazza Vittorio Veneto") in Matera together with the Hypogeum located underneath. The Hypogeum, including the main water reservoir (well-known as "Palombaro Lungo"), is only partially open to the public, due to some environmental issues, such as the cave environment, the steep steps, and the low illumination. This is to demonstrate the benefits of 3D urban Digital Twin for improving accessibility and better overall experience for visitors.

Figure 1 shows an example of the results in the pipeline on our use case, from the reality to the generated triangulation. In particular, in the proposed case study, the original point cloud has 256M points and it was decimated to 65M points after some cleaning. The final triangulation is made of 75K colored vertices.

Figure 1: Left: view of the Vittorio Veneto square (the hypogeum is underneath). Middle: Acquired points cloud. Right: Reconstructed triangle mesh.

The preliminary VR environment representing Piazza Vittorio Veneto in Matera and the Palombaro cistern is shown in Figure 2. It enables an immersive virtual exploration of these sites moving between the different underneath levels in a quite realistic way, also allowing the visit to the narrow, steep, or closed to the public areas. Preliminary visualization has been tested accessing the VR environment by wearing a Meta Quest 2 head-mounted display (HMD).

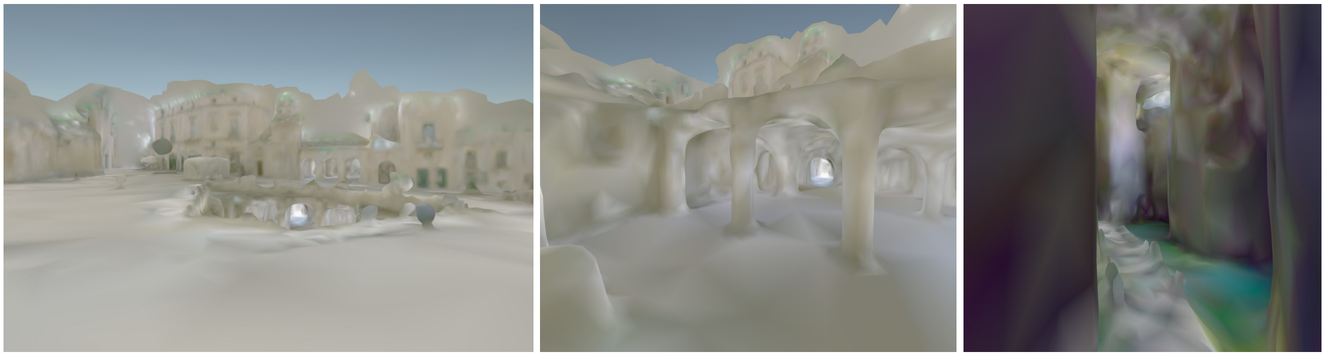


Figure 2: Different areas of the VR environment representing the Vittorio Veneto square, the Hypogeum access, and the Palombaro cistern (from left to right).

4. Discussion and Future Works

One of the primary objectives is to achieve a reconstruction by the 3D acquisition described previously that ensures an authentic and realistic experience balancing the level of details and the model size so that it does not affect the performance. To this aim, our approach generates a 3D model with a certain color associated for each vertex. However, some work should be done to improve realism, possibly by defining and using more specific textures.

As future work, we aim to study and develop interactive navigation techniques of the virtual 3D environment. For instance, users can freely navigate the city simply by teleporting them into different parts of the 3D reconstructed areas, or naturally walking around the urban area. In the last case, an analysis of the height variations of the floor and a camera adaptation has to be taken into account to improve the sense of realism.

ACKNOWLEDGMENTS

This work was supported by the national project "House of emerging technologies in Matera", CTEMT, I14E20000020001. The instrumentation BLK2Go was made available by the Italian Competence Centre START4.0.

REFERENCES

1. Castelli, G., et al. Urban intelligence: a modular, fully integrated, and evolving model for cities digital twinning, *2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT)*, pp. 033–037, IEEE, (2019).

2. Scalas, A., Cabiddu, D., Mortara, M. and Spagnuolo, M. Potential of the geometric layer in urban digital twins, *ISPRS International Journal of Geo-Information*, **11** (6), 343, (2022).

3. Lupinetti, K., Bonino, B., Giannini, F. and Monti, M. Exploring the benefits of the virtual reality technologies for assembly retrieval applications, *Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019, Santa Maria al Bagno, Italy, June 24–27, 2019, Proceedings, Part I 6*, pp. 43–59, Springer, (2019).

4. Scalas, A., Cabiddu, D., Mortara, M., Pittaluga, S. and Spagnuolo, M. Mobile Laser Scanning of Challenging Urban Sites: a Case Study in Matera, Ponchio, F. and Pintus, R. (Eds.), *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, (2022).

5. Lorensen, W. E. and Cline, H. E. Marching cubes: A high resolution 3D surface construction algorithm, *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, (1987).

TOWARDS DIGITAL TWINS OF TERRITORIES THROUGH SEMANTIC STORY MAPS

Valentina Bartalesi, Gianpaolo Coro, Emanuele Lenzi, Nicolò Pratelli *and* Pasquale Pagano
Institute of Information Science and Technologies "Alessandro Faedo" (ISTI) of the National Research Council of Italy (CNR), Via Moruzzi 1, 56124, Pisa, Italy
e-mail: valentina.bartalesi@isti.cnr.it

Digital maps greatly support storytelling about territories, especially when enriched with data describing cultural, societal, and ecological aspects, conveying emotional messages that describe the territory as a whole. Story maps are interactive online digital narratives that can describe a territory beyond its map by enriching the map with text, pictures, videos, and other multimedia information. This paper outlines how online story maps can fill the gap between a map and a territory in narratives to create a digital twin of different territories as inter-connected semantic stories.

Keywords: *story maps, semantic web, linked open data, digital twin of territories*

1. Introduction

A narrative is a method of articulating life experiences in a meaningful way and a conceptual basis of collective human understanding [1]. Humans use narratives, or stories, to explain and communicate abstract experiences, describe a particular point of view on some domain of interest, and share meanings among different communities. Maps have always geographically supported stories and storytelling, especially to represent the geospatial components of narratives. However, as Korzybski pointed out in his General Semantics [2], there is a perceptual gap between the territory and its map in narratives. The *perceptive* cartographic challenge for a map is when it tries to represent also the life, emotions, reality, fiction, legends, and expectations associated with the described territory. Story maps are computer science realisations of narratives based on interactive online maps enriched with text, pictures, videos, data, and other multimedia information. Story maps building platforms support diverse users to participate and possibly collaborate in the map-based story telling. Story maps can be used both to enhance the perception of the geographic space cited in narratives and also to overcome Korzybski's perceptual gap, enriching maps with media that communicate emotional messages associated with the described territory, e.g. digital audio/video material to describe the territorial complexity. Story maps satisfy these properties creating a digital twin of the territory, going beyond the standard map representation. Digitally interconnecting different story maps within a knowledge base can reveal implicit information hidden within and between the territories and allows answering previously new, complex questions.

This paper briefly outlines how online story maps can fill the gap between a map and a territory in narratives to create a digital twin of different territories as inter-connected stories. We start from a specific domain (mountain territories and value chains) and propose a general software solution for a semi-automatic transformation from text to narrative. In particular, we build a workflow to transform the textual descriptions of the value chains and territories of 23 rural European areas from 16 countries involved in the MOVING (MOuntain Valorisation through INterconnectedness and Green growth) European project [3]. MOVING aims at building co-development policy frameworks across Europe that can contribute to the resilience and sustainability of mountain areas. Community data on rural areas come in unstructured textual formats. They summarise the emotional, cultural, and societal contents to attract tourism and customers. Digitalising this information and assigning it to map locations is complex, but is crucial for extending the information reachability, effectively communicate contents, and harmonise the data to produce one overall narrative of territories.

Our semi-automatic workflow transforms the unstructured information about a territory into a story map. The workflow uses natural language processing algorithms [4] to extract terms (persons, locations, organisations, and keywords) with high importance to understand the text (named entities). Successively, it associates the extracted terms with Wikidata entries and geographic coordinates. Then, it produces and visualises as story maps a sequence of story events enriched with titles, descriptions, entities, coordinates, and multimedia data. Finally, the workflow feeds a semantic knowledge base, based on an ontology for representing narratives [5], to enable data discovery and integration. The workflow design is independent of the specific application context. Furthermore, differently from other story-map creation software [6, 7], the workflow components are open-source, and oriented to the Open Science directives of transparency and reproducibility of data, results, and processes. Moreover, different from other story-visualisation software [8, 9, 10, 11, 12], the workflow provides a semantic knowledge base that automatically inter-connects the created stories, extracts and discover knowledge from them, and comply with the Linked Open Data paradigm. This way, our workflow points towards satisfying the creation of a digital twin of the territory.

2. Methodology

2.1 Data pre-processing

We pre-processed the unstructured textual data of 23 European regions to prepare one new textual document - in the MS Excel format - for each region. For each story, we organised the new MS Excel document by describing one story event in each row. Specifically, each row reported (i) a title, (ii) a description, (iii) one representative image (optionally), (iv) hyperlinks to online multimedia material (optionally), and (v) the event *type*. The row sequence represented the events' sequence of the region's story. An event sequence reports the same concepts expressed by the paragraphs of the corresponding region description by the MOVING regional experts. The data associated with each region from the experts' descriptions and original data were mapped onto the new MS Excel documents by attaching them to the most appropriate events. The 23 newly prepared documents were sequentially passed to our workflow as input data.

2.2 Story-structure building

A story-structure building module processes the text of the input MS Excel rows (events) to enrich them with information on associated entities and locations. In particular, it uses the NLPHub [4] and the Wikidata services to extract valid entities with possible associated spatial coordinates. To keep the entity locations focused on a geographically consistent region, our module traces a bi-variate log-normal distribution on the longitude-latitude pairs. Then, it sets the boundaries to the upper and lower log-normal confidence limits over the axes, and marks all coordinates outside of these boundaries as outliers. For the entities whose coordinates are not present in Wikidata, the algorithm adds by default the coordinates of the Local Administrative Unit in which the story occurs.

2.3 Story map creation

Each MS Excel files enriched with event and coordinate information are further processed to finalise the story map representation through multimedia and hyperlink information and a structured format. First, each story event is associated with the event type and multimedia hyperlinks specified in the original input text. Second, all acronyms are expanded through a reference domain-specific dictionary to make descriptions less technical. Third, images are linked to the events if referred in the original input text; otherwise, the first image associated with the event entities on Wikidata, ordered by their position in the text, is retrieved and linked. Finally, the story's event sequence, with all associated entities, images, and links is described in the JSON format, according to the schema used by the Story Map Building and Visualising Tool (SMBVT) [13]. This JSON document is an offline realisation of the story map. The document it stored on a PostgreSQL-JSON database used by the SMBVT for fast online visualisation (Section 2.4). Finally, the script invokes a semantic triplification software that translates the JSON document into a Web Ontology Language (OWL) graph and stores it in an Apache Jena Fuseki triple store. This graph complies with the Narrative Ontology model [5], and its scope is to populate the SMBVT knowledge base and enable semantic queries for knowledge extraction.

2.4 Story Map Building and Visualising Tool

SMBVT [14, 15] is a open-source software that represents narratives as a network of spatiotemporal events related by semantic relations (part-of, temporal, spatial, and causal relations) enriched by event components, i.e., the entities that take part in the event (e.g., persons, objects, places, concepts). For the present experiment, we used a free-to-use SMBVT instance hosted on the D4Science e-Infrastructure accessible after free registration to the platform [16]. Following the Semantic Web approach [17], SMBVT assigns to each event and event component an IRI [18]. These IRIs are mainly extracted from Wikidata [19], which SMBVT uses as an external reference knowledge base. SMBVT retrieves events, entities and all associated data from a PostgreSQL-JSON database for visualisation. It synchronises this database with an OWL-graph representation of the stories stored on an Apache Jena Fuseki server, compliant with the Narrative Ontology model. The Fuseki server provides a SPARQL endpoint to query the complete graph of collected stories. The graph automatically connects the stories through the entities shared between the events. The server allows executing SPARQL data-extraction queries on the entire story graph within or across the stories. This feature allows connecting the stories to other knowledge bases published as Linked Open Data [20] (e.g., Europeana, [21]). In particular, based on the SPARQL server, SMBVT offers an entity-search functionality through a Web interface that allows querying the entire knowledge base in a user-friendly way. This feature is crucial to explore story inter-connections, for example, to retrieve (i) the events involving a specific entity (e.g., sheep, beer, etc.), (ii) the nations sharing the same entities (e.g., products, export locations, etc.), and (iv) the most frequent entities across the stories (e.g., the most common products). SMVBT provides an online graphical interface to create and manage the stories. This interface facilitates story-event building and event sequencing and contextualisation. SMBVT visualises the produced stories as story maps placing the narrative events on an interactive map that respects an event browsing order based on the user-defined plot.

The SMBVT story map publication process returns a public link for each story map. Each publication operation overwrites the previously published application so that the public link always points to the latest story-map version.

3. Conclusions

Through semantic queries, our approach contributes to discover new knowledge from the data; for example, the territories sharing the same environmental characteristics (e.g., rivers, lakes, vineyards, chestnut trees) and issues (e.g., depopulation, emigration, climate change problems), or providing similar products (e.g., cow or sheep milk cheese). Discovering new knowledge from the data is particularly useful for mountain ecosystems to design sustainable environmental management pathways and contribute to long-term cities' ecological sustainability. The extracted knowledge therefore represent a digital twin of the territories through the interconnection of their representations as story maps. The constructed semantic knowledge base can indeed help understand and dam the vanishing of essential services in rural areas due to constant depopulation trends in rural communities. Therefore, it helps answering crucial questions about the territory socio-economic and ecosystem status. The importance of describing a territory beyond its maps is more understandable if we consider that the world's rural population will likely pass from 47% (today) to 30% by 2050 [22]. This trend poses problems due to vanishing traditional and cultural heritage aspects and human health, welfare, and cities' ecological sustainability, monitored by the United Nations' Sustainable Development Goal 11 ("sustainable cities and communities"). In this context, story maps and knowledge discovery can help citizens, investors, and governmental authorities to better understand the ecosystem and support sustainability strategies. The MOVING project has indeed adopted story maps (available at `https://www.moving-h2020.eu/story-maps/`) as effective tools for this task because the experts judged them to appropriately convey "information going beyond the map" for scientists, stakeholders, and the general public.

REFERENCES

1. Wertsch, J. V. and Roediger, H. L. Collective memory: Conceptual foundations and theoretical approaches, *Memory*, **16** (3), 318–326, (2008).

2. Korzybski, A., (1933), A non-aristotelian system and its necessity for rigour in mathematics and physics. *Science and sanity: an introduction to non-Aristotelian systems and general semantics*, Lancaster.

3. MOVING, (2020), *The MOVING European project - Mountain Valorisation through Interconnectedness and Green Growth*. Accessed 4 January 2023 https://www.moving-h2020.eu/.

4. Coro, G., Panichi, G., Pagano, P. and Perrone, E. Nlphub: An e-infrastructure-based text mining hub, *Concurrency and Computation: Practice and Experience*, **33** (5), e5986, (2021).

5. Meghini, C., Bartalesi, V. and Metilli, D. Representing narratives in digital libraries: The narrative ontology, *Semantic Web*, **12** (2), 241–264, (2021).

6. ArcGIS, (2022), *ArcGIS Story Map Building Tool*. https://storymaps.arcgis.com.

7. Timescape, (2020), *Timescape Story Map Building Tool*. https://www.recombine.net/projects/timescape.html.

8. Knightlab, (2020), *StoryMap JS tool for narrative building*. https://storymap.knightlab.com.

9. Frenvik Sveen, A., (2020), *Leaflet Storymaps template*. https://github.com/atlefren/storymap.

10. Map Box, (2020), *Map Box Storymaps template*. https://www.mapbox.com/solutions/interactive-storytelling.

11. Becker, T., Köbben, B. and Blok, C. Timemapper : visualizing moving object data using wms time and svg smil interactive animations, *Proceedings SVGOpen 2009*, SVGOpen, 7th international conference on scalable vector graphics (SVGOpen 2009), 02-04-10-2009, Mountain View, United States, (2009).

12. Odyssey, (2020), *Odyssey Story Maps template*. http://cartodb.github.io/odyssey.js/.

13. Bartalesi, V., Coro, G., Lenzi, E., Pratelli, N. and Pagano, P. ISTI Research Report, MOVING, D3.3, 2022, Moving d3.3 - tools for science-society-policy interfaces. using semantic story maps to describe a territory beyond its map, (2022).

14. Bartalesi, V., Coro, G., Lenzi, E., Pagano, P. and Pratelli, N. From unstructured texts to semantic story maps, *International Journal of Digital Earth*, **16** (1), 234–250, (2023).

15. Bartalesi, V., Lenzi, E. and Pratelli, N. A web tool to create and visualise semantic story maps, *Text2Story 2023 Sixth Workshop on Narrative Extraction From Texts*, CEUR, (2023).

16. Assante, M., et al. Enacting open science by d4science, *Future Generation Computer Systems*, **101**, 555–563, (2019).

17. Berners-Lee, T., Hendler, J. and Lassila, O. The semantic web, *Scientific american*, **284** (5), 34–43, (2001).

18. Dürst, M. and Suignard, M. W3C, Internationalized resource identifiers (iris), (2005).

19. Vrandecic, D. The rise of wikidata, *IEEE Intelligent Systems*, **28** (4), 90–95, (2013).

20. Bizer, C., Heath, T. and Berners-Lee, T., (2011), Linked data: The story so far. *Semantic services, interoperability and web applications: emerging concepts*, pp. 205–227, IGI global.

21. Isaac, A. and Haslhofer, B. Europeana linked open data–data. europeana. eu, *Semantic Web*, **4** (3), 291–297, (2013).

22. United Nations, (2018), *68% of the world population projected to live in urban areas by 2050, says UN*. on-line, https://www.un.org/en/desa/68-world-population-projected-live-urban-areas-2050-says-un.

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

3D FEATURE RECOGNITION FOR THE ASSESSMENT OF BUILDINGS' ENERGY EFFICIENCY

Daniela Cabiddu, Michela Mortara, Chiara Romanengo, *and* Andreas Scalas
CNR-IMATI, Genova, Italy

Alice Bellazzi, Lorenzo Belussi, Ludovico Danza, *and* Matteo Ghellere
CNR-ITC, Milano, Italy,
e-mail: chiara.romanengo@ge.imati.cnr.it

The real assets, procedures, systems, and subsystems of a city can be virtually represented through an urban digital twin (DT), which integrates heterogeneous data to learn and evolve with the physical city, offering support to monitor the current status and predict possible future scenarios. A DT of a city can be organized into layers, which represent specific facets of the city and cooperate to address specific issues. In this work, we present an application scenario in which a geometric layer, representing the 3D morphology of the urban environment, cooperates with an energy consumption layer, providing knowledge of the peculiarities of the building urban area and in particular of the built fabric, to assess their impact in terms of energy efficiency. The analysis of the urban geometries provides quantitative measures as useful input, for instance, to define heat leakage.

Keywords: *urban intelligence*, *geometric layer*, *semantic enrichment*, *energy efficiency*, *urban mapping*

1. Introduction

The geometric representation of the city's physical space, through the enrichment of relevant elements with heterogeneous data, becomes the common place where the integration of different information, "views", or topics actually takes place. Accompanying salient urban elements (single building, block, street, etc.) with all kinds of information related to them allows reasoning on multiple levels and inferring qualitative or quantitative attributes about the state of the element. In this work, we focus on the *geometric layer*, which represents the morphology and physical features (either built or natural) of the city, and its integration with the *energy consumption layer* to compute and map the building energy efficiency, e.g., to support urban planning.

2. Methodology

The energy efficiency assessment of the built heritage within the DT passes through the interaction between the geometric layer -the 3D model of the city- and the body of knowledge related to the energetic analysis of the built heritage, that is, the energy consumption layer. The integration concretely happens through the mechanism of *annotation*. The energy layer defines the properties of elements that concur to the computation of efficiency (e.g., building exposed surface, roof orientation) and interrogates the geometric layer to get these properties. The geometric layer runs geometry analysis algorithms to automatically compute these dimensions.

The energy consumption layer aggregates the heterogeneous knowledge of the built heritage that is necessary for the energy efficiency assessment[1]. Building knowledge comes from different sources as periodic censuses, historical and thematic urban maps and other previous studies. The information layer considers both general data and data derived from building and urban morphology. In the first case, general data refers to year/period of construction and the building's intended use, while the latter is the S/V ratio, which indicates how large the surface area S (such as wall, ceiling, roof and window surface areas) is in relation to the building volume V, the urban context surrounding the building and the incidence of window area on the external envelope of the building.

The geometric layer consists of a 3D model - a triangle mesh representation of buildings, generated either by reconstruction from hi-res point clouds produced by specific acquisition surveys, or from low-res data available at national level [2]. In either case, the salient entities, e.g., buildings, are selected and annotated within the undifferentiated triangles of the mesh, using information about the building footprints (e.g., from OpenStreetMap or from cadastral documents). Hence, we compute specific geometric attributes for a class of elements. For the sake of energy assessment, we focus on the following attributes: (i) exposed surface of buildings; (ii) volume of buildings; (iii) surface, orientation and slope of roof pitches; (iv) window areas.

Concerning the exposed surface, the external area of a single building can be computed straightforwardly by summing up the triangle area annotated as that building. Note that adjacent buildings share walls to some extent and this is an important feature for assessing buildings' heat loss. Therefore, the adjacency relation among buildings can be easily obtained by detecting common vertices in building footprints. The approximated shared surface can be computed as the distance between common vertices pairs multiplied by the difference in the buildings' height; then, the total surface minus the shared surface provides the estimated of exposed surface. Regarding the volume computation, it is sufficient to close each building's annotated portion and apply volume computation as defined in [3] (i.e. creating a tetrahedron for each triangle by connecting each of its vertices to the origin and summing up the corresponding *signed volume*). We refer to figure 1 for a depiction of city geometric models with annotations and characterization.

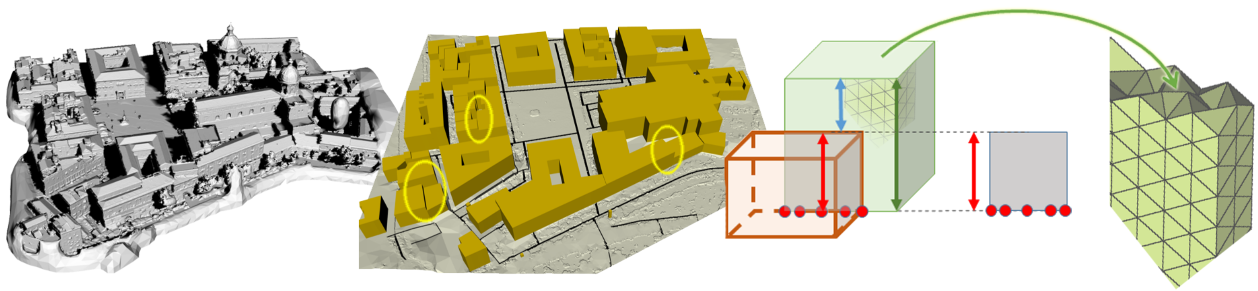


Figure 1: 3D models of Piazza Duomo in Catania, Italy. From left to right: detailed reconstruction from hi-res data; block reconstruction from low-res data and semantic annotation of buildings (in yellow) and streets (black); estimated shared surface between adjacent buildings; volume computation on the corresponding tetrahedral mesh.

To detect roof parts for each building, we extended the feature-recognition method in [4] to the urban context. This approach is designed to recognise geometric primitives (planes, cylinders, cones, spheres, tori) and their associated parameters in point clouds representing CAD objects; it can be naturally extended to the urban context, since the main urban elements can be identified by such types of surfaces. It can be also used to find relationships among the recognised parts (see Figure 2), and it is robust to noise and outliers in the geometric data. In the energy analysis scenario, it recognises mesh vertices composing the same roof pitch and provides the fitting plane normal direction. Then, the computation of roof surface, orientation and slope is straightforward. The same approach cannot be applied to windows, as the window plane mainly coincides with the external walls. Thus, windows are annotated manually or by images.

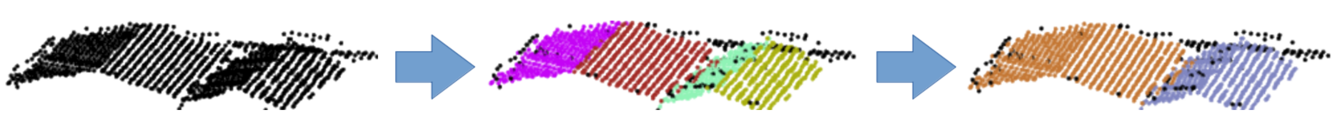


Figure 2: Feature-recognition method applied to urban context. (left) part of the point cloud representing a city. (center) the points that belong to different recognised planes are highlighted in several colours. (right) the geometric parameters associated with the recognised planes can be used to determine if two planes belong to the same roof.

3. Results

While the geometric layer can provide information such as the extent of dispersing surfaces, air-conditioned volume, type of roof, number of pitches, and orientation/exposure, the association of historical data allows deriving for each building HVAC (Heating, Ventilation and Air Conditioning) and lighting systems data as layout, yearly hours of operation, system efficiency, energy carrier used and building envelope thermal properties.

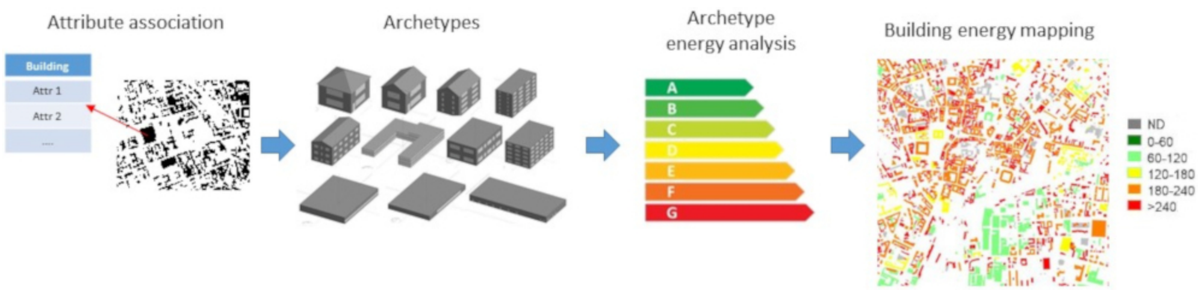


Figure 3: Building energy efficiency analysis after the semantic enrichment.

Considering the unique thermal characteristics of each real building and the number of buildings belonging to the city, some simplifications and a minimum level of energy calculation uncertainty are necessary for mapping purposes. On the basis of morphological and non-morphological annotated data, a set of building archetypes representing the building stock is defined. With the previous data, the estimation of the primary energy need (PE) and of the CO2 emissions (MCO2) produced per square meter of building archetype floor area is assessed through building energy simulation (BES).

A final additional semantic association links each building of the DT to the correspondent archetype, inheriting PE and MCO2 values and allowing their punctual mapping over the city territory.

From the integration of this information at a higher scale, an estimation of the energy consumption of all buildings, or in a portion of it, in the municipal area can be generated to assess potential energy savings as a result of urban regeneration management policies.

4. Conclusions

The proposed methodology provides the DT of the built heritage a mapping tool for the energy efficiency of the city buildings. The goals of the DT are twofold. Firstly, starting from the 3D model, to carry out an analysis of the peculiarities of the built fabric from the energy consumption point of view. Secondly, the DT aims to support public authorities in evaluating the impact of land transformation decisions and driving future urban policies. We point out that the accuracy of the computed quantitative properties depends on the resolution of the 3D model, and in turn, of the input point cloud. The current research is seeking to recognise and annotate more features, such as windows, as automatically as possible.

ACKNOWLEDGEMENTS

This work is funded by the research project DIT.AD016.097 UISH “Urban Intelligence Science Hub for City Network”– CUP B51B21000430001. It is part of the specific objective II.1 - "Pilot Projects" - II.1 - "Innovative Actions 3" of the complementary Operational Program “Metropolitan Cities” 2014-2020.

REFERENCES

1. Ghellere, M., Belussi, L., Barozzi, B., Bellazzi, A., Danza, L., Devitofrancesco, A. and Salamone, F. Energy and environmental assessment of urban areas: an integrated approach for urban planning, *Building Simulation 2021*, vol. 17, pp. 77–85, IBPSA, (2021).

2. Scalas, A., Cabiddu, D., Mortara, M. and Spagnuolo, M. Potential of the geometric layer in urban digital twins, *ISPRS International Journal of Geo-Information*, **11** (6), 343, (2022).

3. Zhang, C. and Chen, T. Efficient feature extraction for 2D/3D objects in mesh representation, *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, vol. 3, pp. 935–938 vol.3, (2001).

4. Romanengo, C., Raffo, A., Biasotti, S. and Falcidieno, B. Recognizing geometric primitives in 3D point clouds of mechanical CAD objects, *Computer-Aided Design*, **157**, 103479, (2023).

DIGITAL TWINS OF CIVIL STRUCTURES USING NEURAL NETWORKS AND PROBABILISTIC GRAPHICAL MODELS

Matteo Torzoni *and* Stefano Mariani
Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Milan, Italy
e-mail: matteo.torzoni@polimi.it

Andrea Manzoni
MOX, Dipartimento di Matematica, Politecnico di Milano, Milan, Italy

Marco Tezzele *and* Karen E. Willcox
Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, United States

This work proposes a predictive digital twin approach to the health monitoring and management planning of civil structures. The asset-twin dynamical system and its evolution over time are encoded by means of a probabilistic graphical model, adopted to rule the observations-to-decisions flow and quantify the related uncertainty. Deep learning models are adopted to assimilate observational data, and provide structural health diagnostics in real-time. The digital state is updated in a sequential Bayesian inference fashion, to inform an optimal planning of maintenance and management actions. A preliminary offline phase involves the population of training datasets through a reduced-order numerical model, and the computation of a health-dependent control policy.

Keywords: *digital twin, structural health monitoring, Bayesian network, deep learning, model order reduction*

1. Introduction

The digital twin (DT) concept represents the most exciting opportunity to move forward predictive maintenance practices, and thus increase the safety and availability of civil structures. This is nowadays possible as the installation of data collecting systems has become affordable, and thanks to the advances in learning methodologies.

This work proposes a DT framework for civil structures. The asset-twin dynamical system is encoded by means of a dynamic Bayesian network (DBN) inspired by [1]. The observations-to-decisions flow is encoded as:

- *From physical to digital.* Observational data, or measurements, are gathered from the physical system and assimilated with deep learning (DL) models, see e.g. [2], to estimate the structural health parameters underlying the digital state and describing the variability of the physical asset. This first estimate of the digital state is then exploited to estimate an updated digital state, according to control-dependent transition dynamics models describing how the structural health is expected to evolve.
- *From digital to physical.* The updated digital state is employed to predict the future evolution of the physical system, thereby enabling predictive decision making about maintenance and management actions.

The DT framework is made computationally efficient through a preliminary offline phase that involves: (i) the population of training datasets through a reduced-order numerical model, see e.g. [3], exploiting the physics-based knowledge about the system response. This is useful to overcome the lack of experimental data typical of civil engineering applications. (ii) training the DL models underlying the structural health identification. This allows for automating the selection and extraction of optimized damage-sensitive features, to ultimately relate them with the corresponding structural states in real-time. (iii) learning the health-dependent control policy to be applied at each time step of the online phase, to map the belief over the digital state onto actions feeding back to the physical asset.

The strategy is assessed on the simulated monitoring of a railway bridge, demonstrating the capabilities of health-aware DTs of accurately tracking the evolution of structural health parameters under varying operational conditions, and promptly suggesting the most appropriate control input with relatively low uncertainty.

2. Methodology

The DT assimilates vibration recordings shaped as multivariate time histories $\mathbf{U}(\boldsymbol{\mu}) = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_u}] \in \mathbb{R}^{L \times N_u}$, consisting of N_u time series made of L sensor measurements. To this aim, a simulation-based strategy is adopted to train DL models on vibration response data, exploiting physics-based models. The asset to be monitored is modeled as a linear-elastic continuum, discretized in space through finite elements. Its dynamic response to the applied loadings is described by the semi-discretized form of elasto-dynamics. The model is parametrized through a vector $\boldsymbol{\mu} \in \mathbb{R}^{N_{\text{par}}}$ of N_{par} parameters ruling the operational and damage conditions. Damage is modeled as a localized reduction of the material stiffness, obtained by means of two variables $y \in \mathbb{N}$ and $\delta \in \mathbb{R}$, respectively describing its position, among a set of predefined locations $y = 0, \ldots, N_y$, and its magnitude. To speed-up the generation of synthetic datasets, a reduced-order model, relying on a reduced basis method, is adopted in place of the finite element model. The training dataset $\mathcal{D}$ is then populated with I instances as $\mathcal{D} = \{(\mathbf{U}_i, y_i, \delta_i)\}_{i=1}^{I}$.

In order to detect, locate, and quantify the presence of structural damage, a classification DL model $\text{NN}_{\text{CL}} : \mathbf{U} \to y$ is adopted to address damage detection/localization, and regression DL models $\text{NN}_{\text{RG}}^{j} : \mathbf{U} \to \delta$, with $j = 1, \ldots, N_y$, are subsequently adopted to address damage quantification.

The DBN defining the asset-twin dynamical system is sketched in Fig. 1. The physical state $S_t \sim P(s_t)$, encapsulates the variability in the state of the asset. The digital state $D_t \sim P(d_t)$ is instead characterized by the structural health parameters adopted to capture such a variability. The observed data $O_t = o_t$ are assimilated with the DL models, to provide a first estimate of the digital state $D_t^{\text{NN}} \sim P(d_t^{\text{NN}})$. This is then adopted in a Bayesian inference fashion, to update the prior belief D_{t-1} and estimate an updated digital state D_t. This can thus be exploited to compute quantities of interest $Q_t \sim p(q_t)$ and to suggest the next control input. $U_t \sim P(u_t)$ and $U_t^A = u_t^A$ denote the belief about what action to take and the control input effectively enacted, respectively. U_t is estimated according to a control policy, that is computed offline by solving the planning problem induced by the expected evolution of the structural health. This involves maximizing the reward $R_t \sim p(r_t)$ quantifying the asset performance over the planning horizon. By exploiting the conditional independence resulting from the graph topology and the Bayes rule, the joint distribution over variables is factorized up to the current time step t_c, as:

$$\begin{aligned} &p(D_0^{\text{NN}}, \ldots, D_{t_c}^{\text{NN}}, D_0, \ldots, D_{t_c}, Q_0, \ldots, Q_{t_c}, R_0, \ldots, R_{t_c}, U_0, \ldots, U_{t_c} | o_0, \ldots, o_{t_c}, u_0^A, \ldots, u_{t_c}^A) \\ &\propto \prod_{t=0}^{t_c} \left[\phi_t^{\text{data}} \phi_t^{\text{history}} \phi_t^{\text{NN}} \phi_t^{\text{QoI}} \phi_t^{\text{control}} \phi_t^{\text{reward}} \right], \end{aligned} \tag{1}$$

$$\begin{aligned} &\phi_t^{\text{data}} = P(O_t = o_t | D_t^{\text{NN}}), && \phi_t^{\text{history}} = P(D_t | D_{t-1}, U_{t-1}^A = u_{t-1}^A), && \phi_t^{\text{NN}} = P(D_t | D_t^{\text{NN}}), \\ &\phi_t^{\text{QoI}} = p(Q_t | D_t), && \phi_t^{\text{reward}} = P(R_t | D_t, U_t^A = u_t^A), && \phi_t^{\text{control}} = P(U_t | D_t). \end{aligned} \tag{2}$$

Since the spaces of the unobserved variables is discrete, we can propagate and update the relative belief exactly with a single pass of the sum-product algorithm. Starting from the updated digital state D_{t_c}, prediction in future is then achieved by unrolling until a desired prediction time the portion of the graph relative to D_t, Q_t, R_t and U_t.

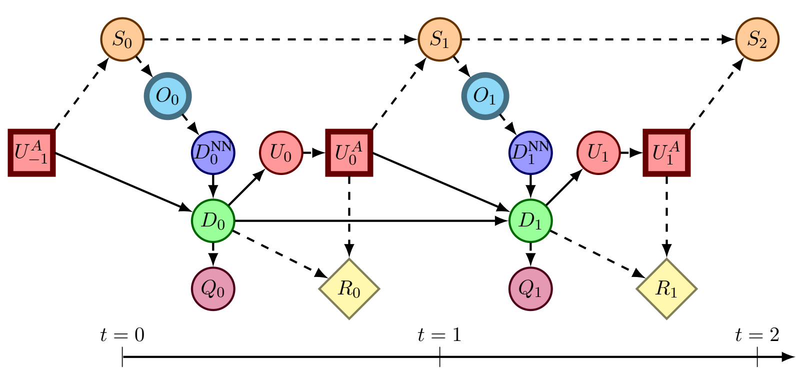
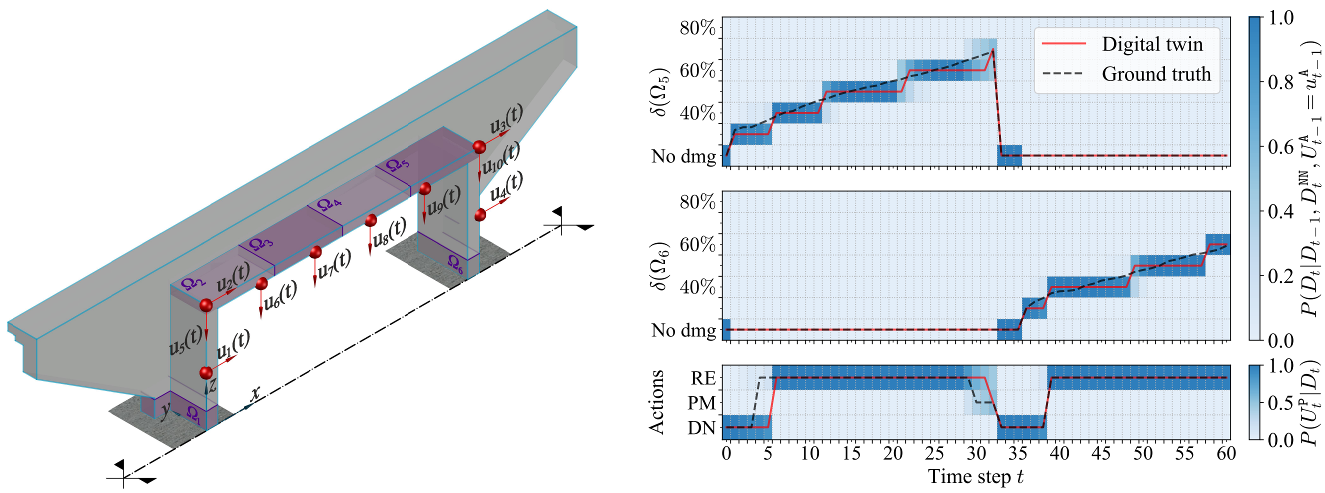


Figure 1: Adopted DBN: circle nodes denote random variables, square nodes actions, and diamond nodes the objective function. Bold outlines denote observed quantities, thin outlines estimated quantities. Solid edges represent dependencies encoded via conditional probability tables, dashed edges encode out-of-the-graph computations.

3. Simulated experiment

The DT framework is applied to the simulated monitoring of a railway bridge. Synthetic vibration recordings $\mathbf{U}$ are obtained from $N_u = 10$ sensors deployed as depicted in Fig. 2a. In addition to the undamaged condition, damage is accounted for by means of a stiffness reduction that can take place within $N_y = 6$ predefined subdomains Ω_j, $j = 1, \dots, N_y$, with magnitude $\delta \in [30\%, 80\%]$. The structural health parameters underlying the digital state are therefore $d = (y, \delta)$. We consider the following control inputs: do nothing (DN) – the physical state evolves according to a stochastic deterioration process; perfect maintenance (PM) – the asset returns to the damage-free state; restrict operational conditions (RE) – only light weight trains are allowed to travel across the bridge, yielding a lower deterioration rate and a lower revenue. Fig. 2b reports a sample DT simulation, in which the variation in the digital state is estimated every time a train travels across the bridge. Damage initially develops within Ω_5. The RE action is suggested as soon as the DT estimates a $\delta \geq 35\%$, after which point the DT keeps on tracking the structural health evolving with a lower deterioration rate. A PM action is finally suggested due to an excessively compromised structural state. A similar behavior can be observed for the following damage scenario affecting Ω_6.



(a) (b)

Figure 2: (a) details of synthetic recordings related to displacements $u_1(t), \dots, u_{10}(t)$, and predefined damageable regions $\Omega_1, \dots, \Omega_6$; (b) probabilistic and best point estimates of (top) digital state evolution against the ground truth digital state, and (bottom) informed control inputs against the optimal control input under ground truth.

4. Conclusions

In this work, we have proposed a predictive digital twin approach to the health monitoring of civil structures, to move forward predictive maintenance practices. The obtained results have demonstrated the digital twin capabilities of tracking the digital state with relatively low uncertainty, and promptly suggesting the appropriate control input.

REFERENCES

1. Kapteyn, M. G., Pretorius, J. V. and Willcox, K. E. A probabilistic graphical model foundation for enabling predictive digital twins at scale, *Nat Comput Sci*, **1** (5), 337–347, (2021).
2. Torzoni, M., Manzoni, A. and Mariani, S. Structural health monitoring of civil structures: a diagnostic framework powered by deep metric learning, *Comput Struct*, **271**, 106858, (2022).
3. Torzoni, M., Manzoni, A. and Mariani, S. A multi-fidelity surrogate model for structural health monitoring exploiting model order reduction and artificial neural networks, *Mech Syst Signal Process*, **197**, 110376, (2023).

A DIGITAL TWIN-BASED PLATFORM FOR STRUCTURAL HEALTH MONITORING: PRELIMINARY RESULTS

Maria Girardi, Cristina Padovani, *and* Daniele Pellegrini
Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", Consiglio Nazionale delle Ricerche, Pisa, Italy
e-mail: daniele.pellegrini@isti.cnr.it

This paper describes the preliminary results of the REVOLUTION project, focused on designing and implementing an open-source, digital twin-based platform for the structural health monitoring of architectural buildings and infrastructures. The platform leverages several technologies, such as 2D/3D digitalization techniques, sensor networks, numerical modelling, and damage detection algorithms. Integrating such technologies will allow us to build accurate digital twins of historical buildings, monitor their dynamic properties, model their mechanical behaviour, detect damages and anomalies and plan intervention strategies.

Keywords: *historical buildings, structural health monitoring, numerical simulations, digital twin, open-source*

1. Introduction

The idea of digital twin dates back to the first years of 2000 [1], and its application to architectural heritage is relatively recent and has yet to be thoroughly exploited. A review of the literature on using the digital twin approach to monitor, maintain and protect architectural constructions is given in [2]. This paper presents the preliminary results of the REVOLUTION project aimed at designing and developing an open-source digital twin platform for buildings of historical importance and infrastructures. After briefly outlining the platform under development, the paper describes an application of the digital twin paradigm to the Guinigi tower in Lucca. The results of the long-term dynamic monitoring campaign conducted on the tower via a sensor network are briefly recalled and the importance of experimental data for calibrating a finite element model of the tower and assessing its structural safety is emphasised.

2. The REVOLUTION project and its open-source platform

Digital twins are virtual representations of physical systems that use sensor networks, hardware and data to provide probabilistic and deterministic previsions on such systems' actual and future states. Implementing a digital twin platform for structural health monitoring is based on the (i) development of flexible and efficient protocols for 3D digitalization and CAD modelling of complex scenario and (ii) integration of data provided by advanced numerical simulations and data measured by the monitoring system and development of tools to make predictive assessments and support the decision process. Algorithms for detecting damages induced by the environmental impact, human activities and seismic actions and highlighting anomalies in the dynamic behaviour of the structure under examination will complement the platform's capabilities. The platform's core is NOSA-ITACA, a code developed by ISTI-CNR to disseminate the use of mathematical models and numerical tools in the field of Cultural Heritage. NOSA-ITACA [3] combines the finite element solver NOSA and SALOME [SAL], an open-source platform for pre- and post-processing operations and enables static and dynamic analysis of structures made of linear elastic and masonry mate-rials, thermomechanical analysis in the presence of thermal loads and can be applied to modelling restoration and reinforcement operations on constructions of architectural interest. The code adopts the constitutive equation of masonry-like materials and models masonry as a homogeneous isotropic nonlinear elastic material with zero or weak tensile strength and infinite or bounded compressive strength. Recently, numerical

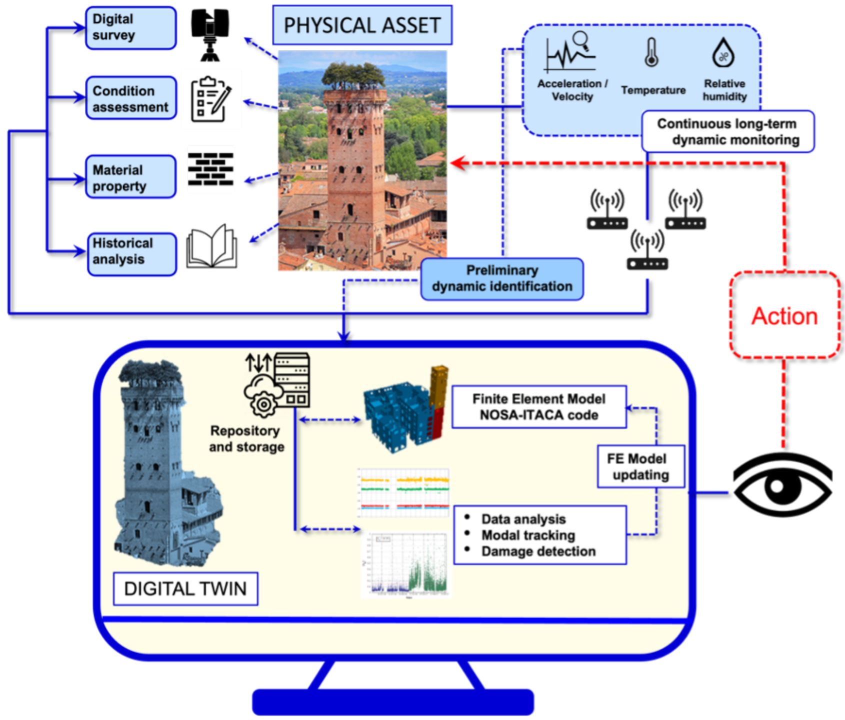


Figure 1: Sketch of the platform implemented within the REVOLUTION project.

methods for constrained generalized eigenvalue problems have been implemented in NOSA-ITACA to address the modal analysis of linear elastic structures. The latest developments of the code focused on the integration of numerical simulations with experimental tests. Thus, algorithms for the finite element model updating, aimed at calibrating the FE model of a structure using its experimental frequencies and mode shapes, have been implemented in NOSA-ITACA, along with a new numerical procedure, which relies on linear perturbation and allows to model the influence of cracks on the dynamical properties of a masonry structure. A description of the main features of the NOSA-ITACA code is provided in [3], where several case studies are presented. Within the framework of the REVOLUTION project, NOSA-ITACA will be modified in such a way as to work as a platform able to interface with other tools necessary to build and manage the digital twin of historical structures and infrastructures. The main functionalities of the platform developed within the REVOLUTION project are sketched in Figure 1.

In particular, the platform will feature the following activities.

1. To carry out the 3D digital acquisition of the under-study structure's geometry via laser scanner techniques. A semiautomatic algorithm will be developed to process and transform the point cloud into a finite element mesh for structural analyses. The acquisition phase of the construction will be conducted by the European Research Infrastructure for Heritage Science (https://www.e-rihs.it) coordinated by the CNR.
2. To apply OMA techniques to the data recorded by the sensor network installed on the structure and calculate the modal parameters of the structure (frequencies, modal shapes, damping ratios). The monitoring system will measure the vibrations induced on the structures by natural and anthropic actions (wind, earthquakes, traffic, crowd movements, etc.) and some environmental parameters (temperature, humidity, wind velocity and directions, etc.). The recorded data will be sent via the Internet to the platform hosted on a remote workstation at ISTI-CNR, where the processing and analysis operations will be carried out. Modal tracking and damage detection tools will be implemented.
3. To calibrate the finite element model of the structure described in 1. FE model updating algorithms will make it possible to determine some unknown parameters of the model (mechanical properties, boundary conditions, etc.) using the experimental data obtained in 2. The updated model will be used to simulate its behavior in the presence of dynamic loads and temperature variations and assess its structural safety.

The Guinigi tower in Lucca (Figure 2), dating back to the XIV century, was chosen as test bed of the platform.

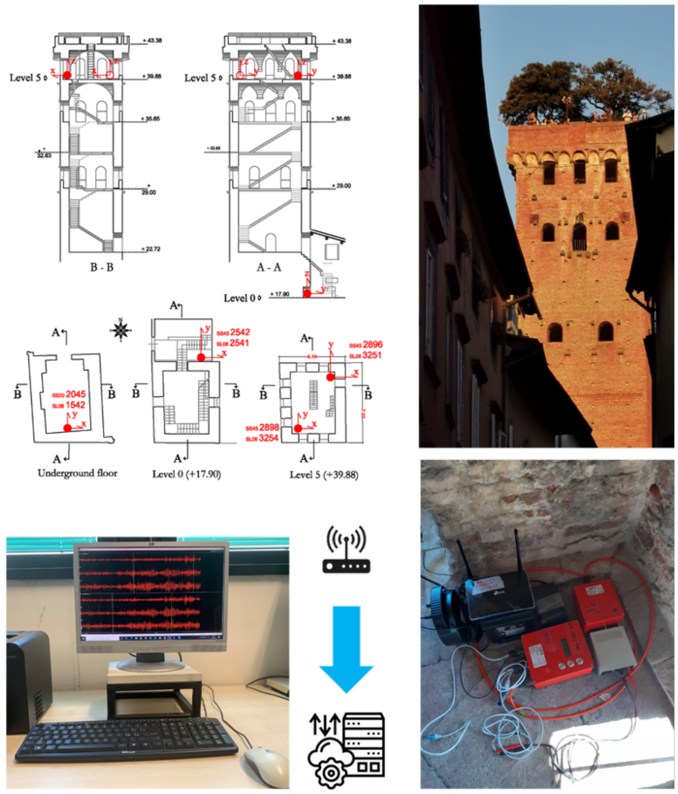


Figure 2: The Guinigi tower

The tower, one of the most iconic monuments in Lucca, was continuously monitored from June 2021 to October 2022 by high-sensitive seismic stations that recorded the structure's response to the dynamic actions of the surrounding environment. A Virtual Private Network allowed sending the data record-ed from the instruments to a server hosted at ISTI-CNR for storage and processing. A detailed analysis of the velocities recorded on the tower and the experimental frequencies calculated via the SSI/Cov algorithm is reported in [4], where the dependence of frequencies on temperature and humidity, as well as the effect of the visitors on the tower's dynamic behaviour, were investigated. A finite element model of the structure was calibrated via the optimization algorithm implemented in NOSA-ITACA aimed at minimizing the distance between the structure's natural frequencies evaluated experimentally and their numerical counterparts evaluated by modal analysis, in a feasible set.

REFERENCES

1. Grieves, M. and Vickers, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems, *Transdisciplinary perspectives on complex systems: New findings and approaches*, pp. 85–113, (2017).

2. Vuoto, A., Funari, M. F. and Lourenço, P. B. On the use of the digital twin concept for the structural integrity protection of architectural heritage, *Infrastructures*, **8** (5), 86, (2023).

3. Girardi, M., Padovani, C., Pellegrini, D., Porcelli, M. and Robol, L. Numerical modelling of historical masonry structures with the finite element code NOSA-ITACA, *INdAM Workshop: Mathematical modeling and Analysis of degradation and restoration in Cultural Heritage*, pp. 133–152, Springer, (2021).

4. Azzara, R., Girardi., M., Padovani, C. and Pellegrini, D., (in press), From structural health monitoring to finite el-ement modelling of heritage structures: The medieval towers of lucca. *Data driven method for civil Structural Health Monitoring and resilience: latest developments and applications, M. Noori, C. Rainieri, M. Domaneschi, V. Sarhosis (Eds)*, CRC Press –Taylor & Francis Group.

DIGITAL TWIN DEVELOPMENT OF A HISTORICAL BUILDING

Carlo Rainieri *and* Luigi Cieri
Construction Technologies Institute, National Research Council of Italy, Naples, Italy
e-mail: rainieri@itc.cnr.it

Ilenia Rosati
Department of Biosciences and Territory, University of Molise, Campobasso, Italy
Visiting fellow at Construction Technologies Institute, National Research Council of Italy, Naples, Italy

Effective tools to characterize the structural response under operational conditions are critical to ensure the reliability of structural health monitoring (SHM) systems. The current structural condition as well as relevant historical data have to be taken into account to have an up-to-date representation of the actual physical system in operation. Digital twins can play a primary role in the evaluation of the current condition of the structure and in anomaly detection to support structural maintenance. Building Information Modelling (BIM) is currently showing a potential in this perspective thanks to the possibility of collecting miscellaneous data and information in a unified platform. So, an effective integration between the two is highly desirable. Relevant information coming from SHM systems can be integrated into a BIM model of a structure to make them available for interrogation and further analyses. The development of a digital twin of a historical structure combining SHM data, BIM, and finite element model updating is illustrated pointing out its promising applicative perspectives for structural maintenance.

Keywords: *digital twin, operational modal analysis, building information modelling, model updating*

1. Introduction

SHM is currently recognized as a key technology for structural assessment and maintenance of civil structures [1]. Remote and automated damage detection can take advantage of the integration of SHM with advanced numerical modelling and simulation tools to enhance its capabilities in the interpretation of data [2]. BIM is also increasingly applied in civil engineering because of its potential, recognized by recent codes [3, 4], in enhancing the exchange of information between the technicians involved in the design and con-struction process, and in optimizing construction time and costs. While the potential of BIM for newly built structures is currently well-established, its applicative perspectives in the field of management, maintenance and rehabilitation of existing structures are still under investigated. Taking into account that advanced SHM technologies can effectively support structural maintenance of existing structures and infrastructures, the integration of data and information from advanced SHM systems can further enhance the effectiveness of 6D BIM in the management of structures [5, 6, 7]. The present study focuses the attention on such SHM-BIM integration for the development of the digital twin (DT) of a historical structure, also exploiting the opportunities of model updating to obtain representative structural models in operational conditions.

2. Digital Twin development

The investigated case study is the Tower of the Nations, a historical building located in the area of the Mostra D'Oltremare urban park in Naples, Italy. The reinforced concrete structure is characterized by two blind and two completely see-through parallel façades, with elevator shafts and stairs located in the central part. Most of the levels are characterized by alternate floors, which cover just a half of the imprint area of the building. The structural system is very original because, in spite of consideration of gravity loads only in structural design, it shows several

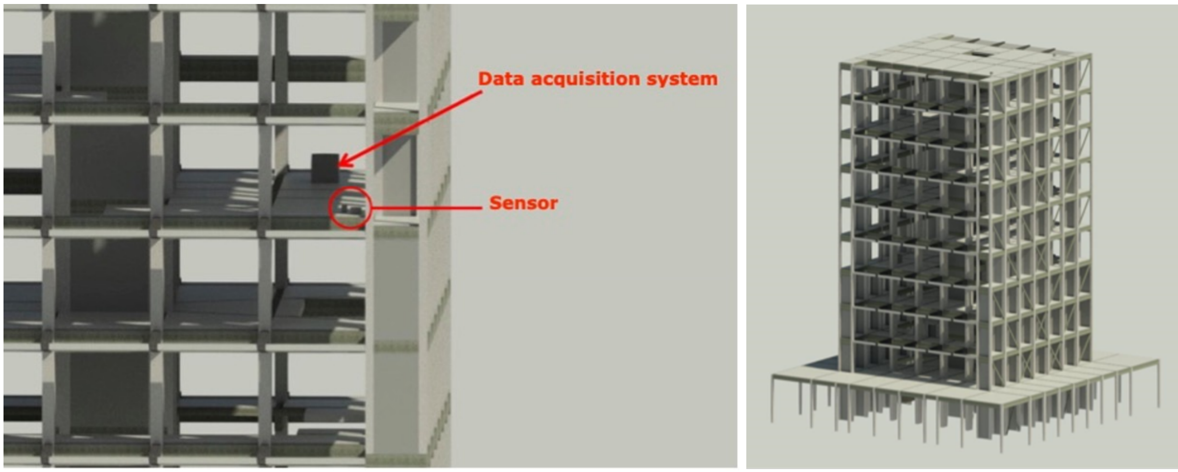


Figure 1: Families created in Revit for sensors and data acquisition system (a), BIM model (b).

elements (walls, braces) able to increase the lateral stiffness of the structure and to provide some capacity against horizontal actions. The mechanical characteristics of materials were obtained from drill cores and non-destructive tests, and the modal properties of the structure by OMA tests [8] have been considered in this application, together with original design drawings and results of geometric and structural surveys, for the development of the BIM model and the DT. For an effective integration between BIM and SHM and to provide the BIM model with some data processing capabilities, the open-source visual programming language Dynamo was used. The first step concerned the creation of new Revit Families allowing the insertion of new information of different nature in the same work environment; specifically, the newly created families contained technical information about the SHM equipment. This operation allows defining the positions of SHM equipment and assessing possible interferences with the ordinary activities. Families for sensors and data acquisition system have been created, and the corresponding elements (Fig. 1a) have been symbolically introduced into the BIM model (Fig. 1b) and coupled with the actual physical devices as described hereafter.

The data are transferred from the SHM system to the BIM model bidirectionally, by creating a direct correspondence between the database of the BIM and that of the SHM system collecting the experimental data. The database of the SHM system for the present application was a relational MySQL database collecting simulated data about the evolution of the natural frequencies of the fundamental modes of the structure over time. To assess

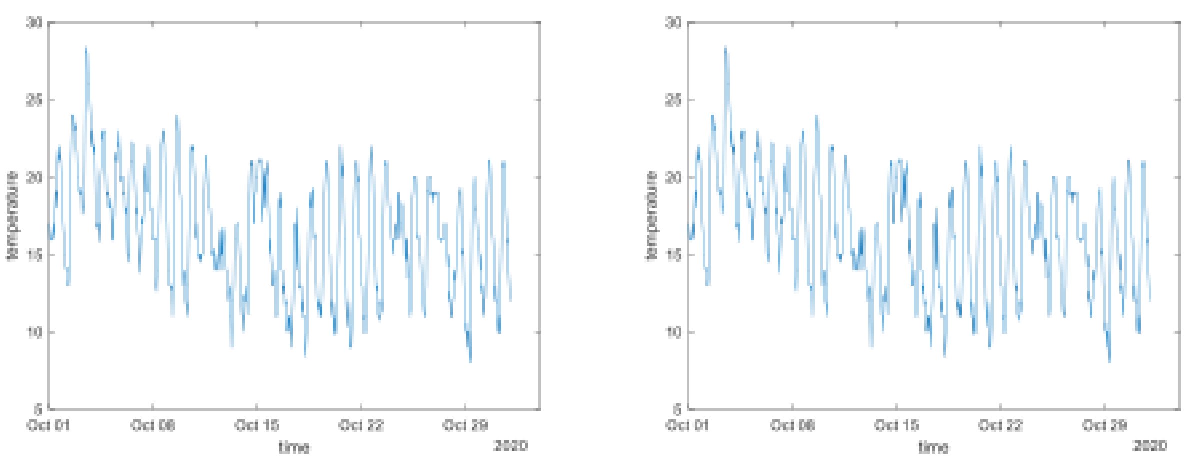


Figure 2: Temperature time series recorded in October 2020 in Naples (a), estimated Young's modulus of concrete by continuous model updating (b).

the effectiveness of the integrated solution for DT developed in the context of the present study, continuous modal parameter monitoring results have been simulated starting from the results of some OMA tests carried out on the structure in the past [9]. The simulated data were representative of the variability of the modal properties with temperature [9] (Fig. 2a). Random noise has been also added to the simulated data to make the time series as much realistic as possible. The opportunity of using SHM data for continuous model updating has been verified afterwards. A DT representative of the structural behavior in operation and able to replicate the structural response under varying environmental conditions was obtained (Fig. 2b).

3. Conclusions

The present study discussed how a DT of a historical structure can be effectively set by taking advantage of the integration of different technologies into a unified environment through the implementation of an appropriate code for continuous model updating based on SHM data. Encouraging results have been obtained in terms of capability of the DT to indirectly follow the evolution of material properties due to environmental effects.

REFERENCES

1. , (2006), *UNI/TR 11634:2016, Linee Guida per il monitoraggio strutturale*.
2. Cabboi, A., Gentile, C. and Saisi, A. From continuous vibration monitoring to fem-based damage assessment: Application on a stone-masonry tower, *Construction and Building Materials*, **156**, 252–265, (2017).
3. , (2017), *UNI 11337-1:2017, Edilizia e opere di ingegneria civile - Gestione digitale dei processi informativi delle costruzioni - Parte 1: Modelli, elaborati e oggetti informativi per prodotti e processi*.
4. , (2017), *D.M. 560/2017, Rome, Italy*.
5. Chen, J., Bulbul, T., Taylor, J. E. and Olgun, G. A case study of embedding real-time infrastructure sensor data to bim, *Construction research congress 2014: construction in a global network*, pp. 269–278, (2014).
6. Rio, J., Ferreira, B. and Poças-Martins, J. Expansion of ifc model with structural sensors, *Informes de la Construcción*, **65** (530), 219–228, (2013).
7. Sternal, M. and Dragos, M. Bim-based modeling of structural health monitoring systems using the ifc standard, *Proceedings of the 28th Forum Bauinformatik, Hanover, Germany*, vol. 20, (2016).
8. Rainieri, C., Fabbrocino, G. and Verderame, G. Non-destructive characterization and dynamic identification of a modern heritage building for serviceability seismic analyses, *NDT & E International*, **60**, 17–31, (2013).
9. Jiao, Y., Liu, H., Wang, X., Zhang, Y., Luo, G., Gong, Y., et al. Temperature effect on mechanical properties and damage identification of concrete structure, *Advances in Materials Science and Engineering*, **2014**, (2014).

BUILD-IT2023 Workshop

19-20 October, Rome, Italy

MODELING AND SIMULATION OF A FARMER ROBOT FOR INFIELD VINEYARD MONITORING

Arianna Rana, Antonio Petitti, *and* Annalisa Milella

Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing - National Research Council, Bari, Italy

e-mail: {arianna.rana, antonoi.petitti, annalisa.milella}@stiima.cnr.it

Digital twin technology has opened new avenues for enhancing agricultural practices through advanced simulation and control systems. In this study, we present the development and implementation of a digital twin for an outdoor mobile robot specifically designed for agricultural tasks. The digital twin comprehensively represents the robot's mechanical system, including its sensors and actuators. The digital twin simulates the robot's movement, perception, and interaction with the surrounding environment. The results of our simulations aim at demonstrating the effectiveness of the digital twin-based simulation approach in improving the performance and productivity of the outdoor mobile robot in agricultural settings.

Keywords: *digital twin, agricultural robotics, simulation frameworks*

1. Introduction

The Digital Twin (DT), which represents a virtual counterpart of a physical product, system, or process, is a rapidly emerging technology that has gained popularity due to its versatility across various application fields such as manufacturing, smart cities, healthcare, and agriculture [1] [2] [3]. In agriculture, one of the open issues to face is to develop autonomous navigation algorithms as a fundamental prerequisite for achieving full autonomy in performing agricultural tasks [4]. The availability of digital version of real platforms allows algorithms to be tested in advance and to observe the platform behavior under different operating conditions. In [5], a DT has been used during the initial phases to test the navigation algorithm for an AMR (Autonomous Mobile Robot) vehicle in a production hall. In [6] the DT of an underwater platform has been presented and is used to validate the feasibility of control algorithms for unmanned systems in real marine environments.

This work deals with the development of the digital twin of an all-terrain robotic platform operating in a vineyard. The implementation of a realistic model of the agricultural scenario and of the mobile robot is presented. The primary objective of the digital twin is to enhance the development and validation of autonomous navigation algorithms in a simulated world, allowing for thorough testing prior to deployment in actual operational conditions. The virtual representation of the robotic platform and the agricultural environment has been generated using Gazebo 11 [7], an open-source 3D robotics simulator. This enables the testing and validation of algorithms developed within the ROS (Robot Operating System) framework.

2. The mobile robotic platform CNRbot

This work introduces the digital twin of a research robotic platform here referred to as CNRbot, depicted in Fig.1. The CNRbot is an all-terrain mobile robot equipped with a locomotion system consisting of four steering and driving wheels, as well as two swing arms with suspensions. Its versatile design allows it to be used in a wide range of applications, both indoors and outdoors, including agriculture and industrial settings. Thanks to its robust suspension system, the vehicle delivers excellent performance even on rough terrains and during heavy-duty

Figure 1: The unmanned ground vehicle available at CNR-STIIMA.

Figure 2: The simulated environment developed in Gazebo composed by a vineyard row and the digital twin of the robotic platform equipped with several sensors.

tasks. Additionally, the robot exhibits omni-directional movement capabilities, as its four wheels can rotate by 180 degrees around their vertical axis.

3. Design and development of the digital twin

This section will discuss the digital twin of the robotic platform, introduced in Sec.2. The digital version of the robot has been developed using ROS and Gazebo. ROS is an open-source framework designed to assist developers in creating and managing robot applications. Gazebo refers to a simulator engine that enables the configuration of customized environments and the simulation of real robotic devices. As depicted in Fig. 2, the model includes a simulated environment, which consists of a vineyard row, and the robot model. The 3D model of the vineyard was created using a 3D modeling software, namely Blender, from which the DAE (Digital Asset Exchange) file was obtained. Then, the simulated environment in Gazebo has been created by adding the DAE file and the texture representing the vineyard row into a custom WORLD file.

Since the CNRbot is a customized robot, there was no existing model available in Gazebo. Therefore, a URDF (Unified Robot Description Format) file was generated using a SolidWorks plugin, starting from the CAD model. The URDF file incorporates information about the robot's links, such as the base and the various components including the wheels, and how those components are related to each other. Furthermore, physics engine like Gazebo require physical properties, such as inertial characteristics, to be specified. Gazebo allows also simulation of sensors, which can be added as a geometric component of the robot. To enable the publication of sensor data, it is required to incorporate the corresponding plugin for the used sensors into the robot model. In our case study, the robot is equipped with sensors that allow it to perceive the environment and perform self-localization, in detail, sensory device capable of creating a 3D model of the surrounding environment, such as an RGB-D camera or a lidar. In order to simulate properly, it is necessary to specify the transmission attribute for each non-fixed joint, specifically for each of the robot's wheels, as well as the speed controller. Without loss of generality, we assume that the robot moves with a speed controlled by a skid steering drive controller, despite having steering wheels.

Finally, in this simulated environment, it is possible to estimate the robot's relative pose (i.e., distance and orientation with respect to the row) [8], test and validate the row-following algorithm [9], and observe the performance of the system prior to deployment in actual operational conditions.

4. Conclusion

This paper presented the development of a digital twin built using ROS and Gazebo 11 for an all-terrain robotic platform operating in a vineyard. The implementation included the creation of a realistic model of the agricultural scenario and of a customized agricultural robot.

By leveraging the capabilities of the digital twin, it is possible to optimize algorithms, assess their performance, and identify potential challenges in a safe and controlled virtual environment. This approach significantly reduces the risks associated with field testing, while improving the overall reliability and robustness of the autonomous navigation system.

REFERENCES

1. Cimino, C., Negri, E. and Fumagalli, L. Review of digital twin applications in manufacturing, *Computers in Industry*, **113**, 103130, (2019).

2. Erol, T., Mendi, A. F. and Doğan, D. The digital twin revolution in healthcare, *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pp. 1–7, (2020).

3. Pylianidis, C., Osinga, S. and Athanasiadis, I. N. Introducing digital twins to agriculture, *Computers and Electronics in Agriculture*, **184**, 105942, (2021).

4. Bai, Y., Zhang, B., Xu, N., Zhou, J., Shi, J. and Diao, Z. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review, *Computers and Electronics in Agriculture*, **205**, 107584, (2023).

5. Stączek, P., Pizoń, J., Danilczuk, W. and Gola, A. A digital twin approach for the improvement of an autonomous mobile robots (amrrsquo;s) operating environmentmdash;a case study, *Sensors*, **21** (23), (2021).

6. Hu, S., Liang, Q., Huang, H. and Yang, C. Construction of a digital twin system for the blended-wing-body underwater glider, *Ocean Engineering*, **270**, 113610, (2023).

7. *Gazebo*. `https://staging.gazebosim.org/home`.

8. Rana, A., Vulpi, F., Galati, R., Milella, A. and Petitti, A. A pose estimation algorithm for agricultural mobile robots using an rgb-d camera, *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pp. 1–5, (2022).

9. Rana, A., Milella, A. and Petitti, A. A row following algorithm for agricultural multi-robot systems, *2023 International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 1–6, to appear, (2023).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

DIGITAL TWINS OF THE OCEAN: 2D/3D/4D MICROPARTICLE FLOW VISUALISATIONS

Jacopo Fabrini, Harald Warmelink, *and* Marie Lhuissier
Breda University of Applied Sciences, Breda, the Netherlands
e-mail: fabrini.j@buas.nl

Ute Brönner
SINTEF Ocean, Trondheim, Norway

Within the context of the Iliad project, the authors present early design mock-ups and resulting technical challenges for a 2D/3D/4D geo-data visualisation application focused on microparticle flows. The Iliad – Digital Twins of the Ocean project (EU Horizon 2020) aims to develop a 'system of systems' for creating cutting-edge digital twins of specific sea and ocean areas for diverse purposes related to their sustainable use and protection. One of the Iliad pilots addresses the topic of water quality monitoring by creating an application offering dynamic 2D and 3D visualisations of specifically identified microparticles, initially observed by buoys/sensors deployed at specific locations and whose subsequent flows are modelled by separate software. The main upcoming technical challenges concern the data-driven approach, where the application's input data is completely obtained through external API-based services offering (near) real-time observed data from buoys/sensors and simulated data emanating from particle transport models.

Keywords: *ocean digital twin, real-time sensor data, 2D & 3D data visualisation, time series data visualisation*

1. Introduction

The Iliad – Digital Twins of the Ocean project (EU Horizon 2020) aims to develop a system of systems for creating cutting-edge digital twins of specific sea and ocean areas for diverse purposes, notably related to their sustainable use and protection. The project will fuse a large volume of data in a semantically rich and data-agnostic approach to enable simultaneous communication with real-world systems and models. Ontologies and a standard style-layered descriptor will facilitate semantic information and intuitive discovery of underlying information and knowledge to provide a seamless experience. The combination of geovisualisation, immersive visualisation and virtual or augmented reality allows users to explore, synthesize, present, and analyse the underlying geospatial data in an interactive manner. To develop and demonstrate its 'system of systems', the Iliad project relies strongly on 20+ pilots, i.e., actual ocean digital twin instances at specific areas for specific purposes [1].

One of the Iliad pilots addresses the topic of water quality monitoring by creating an application offering dynamic 2D and 3D visualisations of specifically identified microparticles, initially observed by buoys/sensors deployed at specific locations and whose subsequent flows are modelled by separate software. Different microparticles might be monitored through such an application, notably micro-organic or microplastics. Other so-called transport models might subsequently be used to extrapolate how the observed microparticles flow in and beyond the sea area around the buoy/sensor over the next hours, days or weeks. The end-user should be able to view this in both 2D (top-down) and 3D, and in both cases with time controls (the '4D' aspect). In the end, the application should also offer an alerting service, so end-users can also be notified of certain (concentration levels of) microparticles when they are observed by the buoy/sensor and base their use of the application on such an alert or notification.

Two of the Iliad partners involved in this pilot are SINTEF Ocean and NTNU. Based in Trondheim, Norway, they developed and deployed OceanLab, which includes a research buoy for essential ocean variables, a microparticle observing camera, a machine-learning-based particle recognition algorithm, and an infrastructure for (near) real-time data retrieval.

Breda University of Applied Sciences (BUas) supports this pilot by leading the development of new visualisation software. Based in Breda, the Netherlands, BUas' Academy of AI, Games Media offers internationally highly regarded educational and research programmes with innovative technologies (particularly game and media). Over the past seven years, several BUas' RD staff have developed more and more expertise in using geo-data from Geographical Information Systems (GIS) with game engines, notably Unity and Unreal, to create engaging, user-friendly, multi-user systems for 2D, 3D and 4D marine/maritime planning, simulations and visualisations. With this expertise and continued interest, BUas will develop a key system for interactive 2D/3D/4D microparticle visualisations within the Iliad 'system of systems' for any other interested pilot or future digital twin of the ocean.

In this extended abstract, the authors explain through early mock-ups the design they are considering for the interactive 2D/3D/4D microparticle visualisation application and the resulting technical challenges they expect to face. The latter is also the result of the data-driven approach taken in developing this application, which is also further explained.

2. First Interface Design Mock-ups

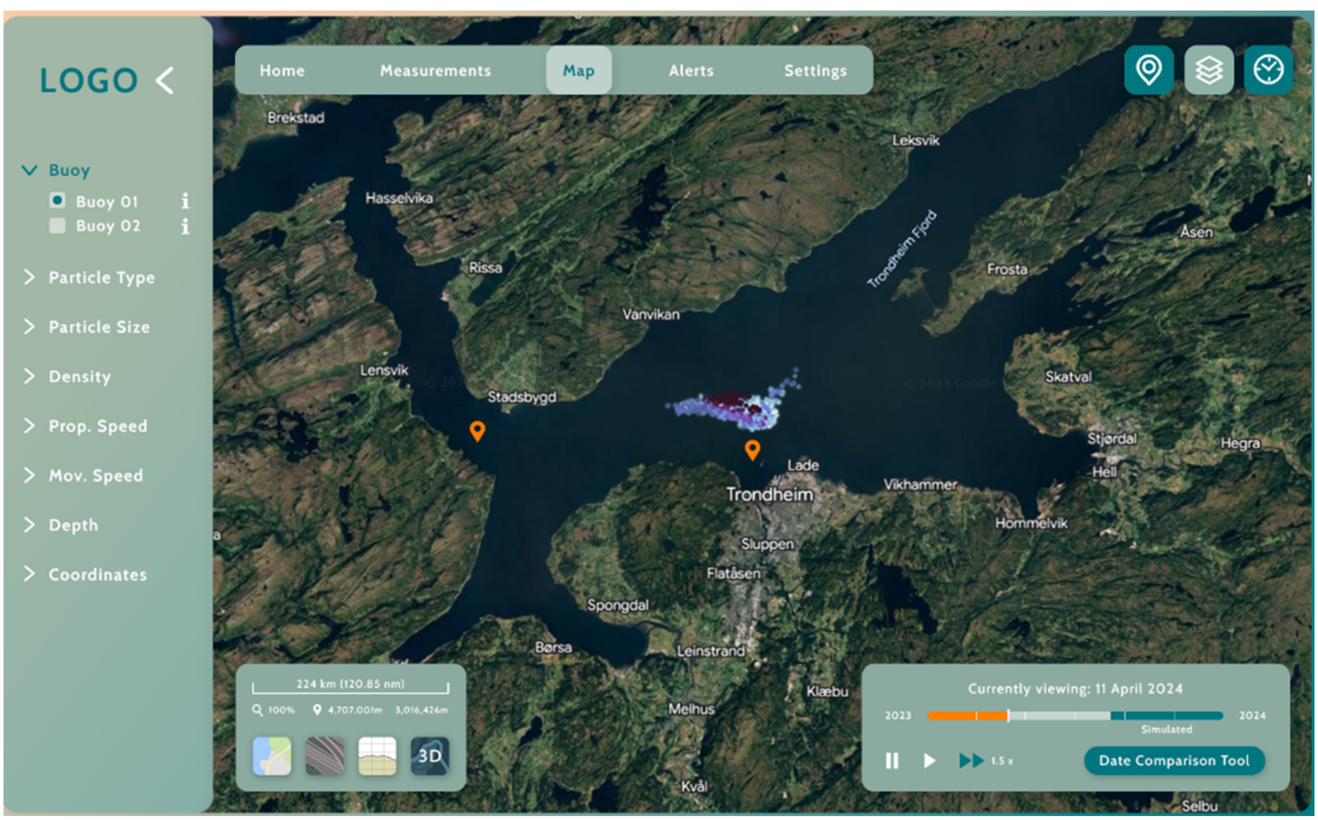


Figure 1: Mock-up of 2D/map-based microparticle visualisation interface, with time controls.

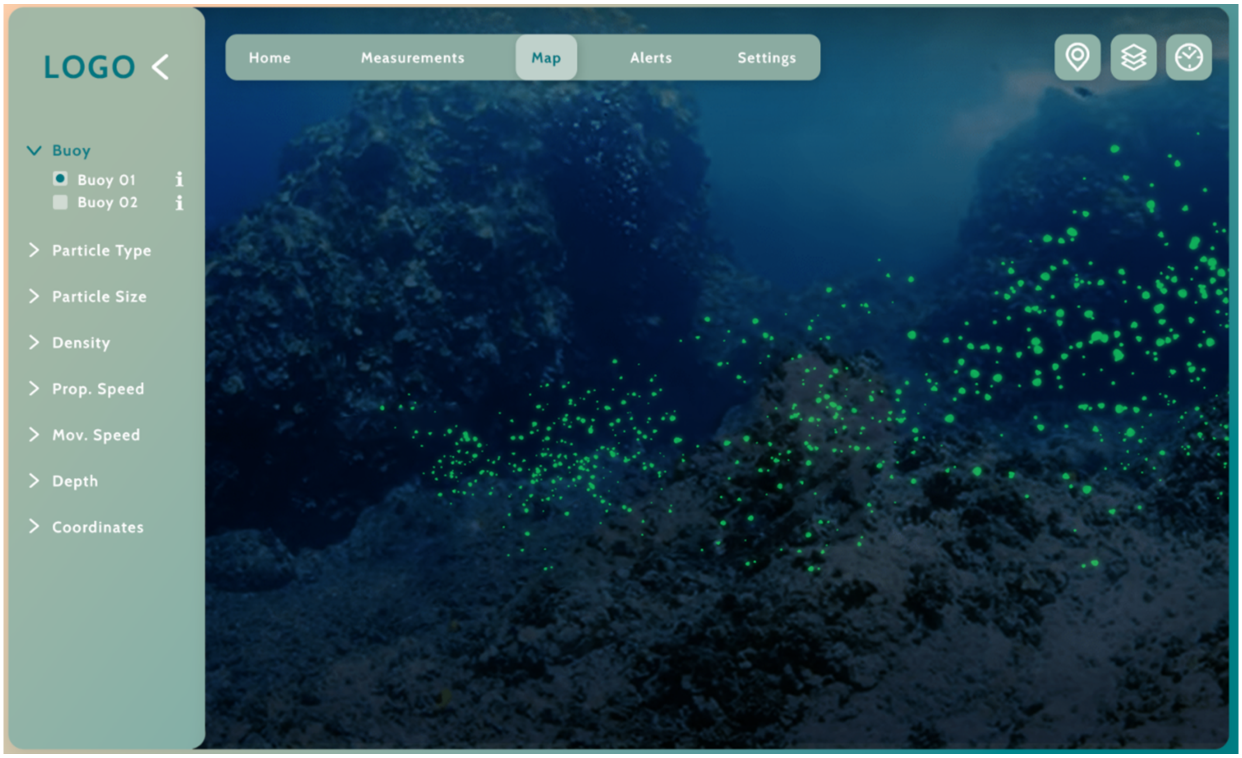


Figure 2: Mock-up of 3D microparticle visualisation interface.

3. Technical Challenges

The main objective here is to create a new bit of software that can take any (near) real-time observed and modelled flow of microparticles delivered through multiple services' APIs and offer 2D and 3D visualisations, also dynamically throughout an appropriate period (hence the 4D). Ideally, the trigger for using the application is not the users themselves but also the application alerting those users of certain types and concentrations of microparticles. Given our expertise with game engines and our wish to provide an interface with which end-users can quickly and seamlessly switch between 2D and 3D visualisations of high quality, we will develop this as a native Windows and macOS app using the Unity 3D game engine. Given this main objective and taking an agile/iterative approach to the development of this system, we are currently working on the following main technical challenges:

- Identifying appropriate 3D time-series-based geo-data standards and protocols serving as input data for this application, usable for both (near) real-time observed data emanating from buoys/sensors and simulated data obtained from particle transport models.
- Ensuring the input data provides complete and correct metadata of again appropriate standards so that the application has absolutely nothing hard-coded and can present the input data in a way that the user can fully understand it.
- Actually getting the application to on-the-fly obtain the input data in 'bursts' or streams from different sources, i.e., the (near) real-time observed sensor data and the output data of the particle transport model.
- Implementing key modular functionalities (particularly the alerting functionality) of the user experience as separate applications or services to add as another system within Iliad's 'system of systems'.
- Integrating the separate services or components seamlessly and with limited to no lag for an optimal end-user experience.

REFERENCES

1. , (2023), *Iliad - Digital Twins of the Ocean*. https://www.ocean-twin.eu/,lastaccessed31May.

HOLISTIC DIGITAL TWIN OF THE OCEAN

Emilio Fortunato Campana *and* Gianpaolo Coro
DIITET-CNR, Rome, Italy
e-mail: {emiliofortunato.campana, gianpaolo.coro}@cnr.it

We introduce the concept of holistic Digital Twin of the Ocean (*h*-DTO), a novel physics-based and data-driven integrated model, which uses AI and Decision Support Systems to model the complex relations between all systems composing a marine area. The *h*-DTO models a marine area as a system of systems. The AI components process data flows to simulate ecosystem functions and answer complex questions about the ecosystem's state, response to perturbations, and the future change of biodiversity and ecosystem services. The *h*-DTO overall aims to enhance the ecosystem understanding and improve our knowledge of the oceans. In this paper, we give indications of the main *h*-DTO features and propose design and implementation pathways.

Keywords: *holistic digital twin of the ocean*, *artificial intelligence*, *decision support systems*, *marine ecosystems*, *ecosystem functions*

1. Introduction

The unsustainable exploitation of ocean resources - with overfishing, chemical and physical pollution, and heavy maritime traffic - is threatening European oceans, seas, and coasts [1]. Climate change is exacerbating this issue. To mitigate this pressure, there is an urgent need for new intelligent digital technologies [2]. Integrated Environmental Assessment (IEA) systems and Ecosystem Models (EMs) can model causal links between anthropogenic driving forces, environmental pressures, and the impact on and response of ecosystems [3]. However, these models heavily rely on data interoperability and scalability capacity and use heuristic and non-automatic approaches with limited result transparency. Artificial Intelligence (AI) can help overcome these limitations and develop scalable and robust virtual representations as Digital Twins of the Ocean (DTOs), combining data from different, heterogeneous, and real-time data flows.

The present paper introduces a new physics-based and data-driven integrated model corresponding to a holistic concept of DTO (*h*-DTO). It empowers AI models with physical/biological laws to model the complex relations between the systems composing a marine area. The *h*-DTO considers the ocean as a system of systems (e.g., morphological, geological, chemical, physical, biological, and economical). Its AI components process data flows to simulate a functioning marine ecosystem and answer complex questions about its current state, its resilience and expected responses to perturbations, and the future change and impact on resident species and ecosystem services. The *h*-DTO adapts to the available data of a marine area - rather than being constrained to a minimum type and quantity of data - performs predictions and informs decision makers.

2. Research questions

An *h*-DTO addresses the following research questions:

1. Can we operate sustainably in a local marine area, given the interconnections between ecosystems and activities and existing constraints from complex natural and anthropogenic driving forces acting on the area?
2. Can we predict ecosystem responses to perturbations and disruptions?
3. Can we predict long-term changes to the ecosystem and blue economy due to climate change effects?

These macro-questions require answering more detailed research questions, such as:

1. Can we automate the discovery of natural relations between the ecosystem, environmental conditions, and anthropogenic stressors?
2. Which data are required within a Decision Support System (DSS) to generate reliable predictions and suggestions?
3. Can AI fill the possible gaps in these data?
4. How can the DSS scale?
5. Can the data and results produced for one area inform the DSS of another area?
6. How can consequences from AI-driven biases, assumptions and omissions be identified and mitigated?

3. Theoretical framework

The *h*-DTO adopts a Bayesian-like approach for decision support. It follows a marine environment over time through environmental monitoring and data accumulation. Meanwhile, it infers information about the exploitation and ecological sustainability of the marine area, the status of the ecosystem, and its capability to adapt to climate change (e.g., sea level rise, extreme events, and long-term change). The *h*-DTO constantly collects new data, integrates these with existing data - available from public and local providers, IoT, and possibly citizen-science networks -, assimilates past and new data with data-driven and physics-based computational models, and intelligently reuses historical heterogeneous data and models via cloud-based analyses. Overall, it can predict the evolution of a marine environment under the effects of heterogeneous driving forces acting on the areas and helps optimise and conserve resources while considering the policies in place.

An active *h*-DTO should live within a virtual environment of an Open Science-compliant e-Infrastructure to ensure repeatability, reproducibility, and reusability features for all models and publish, communicate, and disseminate maps and data through long-term sustainable catalogues. These features should guarantee access and sharing facilities for the marine research activities and campaigns, the re-processed data flows, and the forecasts. The catalogues and the virtual environments should promote the *h*-DTO predictions, assessments, knowledge gap filled, and the emerging possibilities to build productive interactions with the research and industrial sectors. Moreover, they should increase coastal communities' involvement, consensus, and sense of responsibility.

The recipe for building such a system requires working on four main dimensions: (i) Data census, which requires identifying available oceanographic, biodiversity, maritime, social, and fisheries data flow; (ii) strategies to standardise and gap-fill the identified data flows and improve their usability in AI models; (iii) design of AI models to automatically discover ecosystem functional connections; (iv) design of DSSs to simulate complex relations within the ecosystem and answer the research questions.

One critical aspect is model scalability. The *h*-DTO must have modularity, scalability, and adaptability features, such as:

- *spatial* scalability, i.e., generalisation from local to extensive areas;
- *complexity* adaptability, i.e., managing from uniform to heterogeneous data flows and drivers;
- *integration* scalability, i.e., managing from one area to different areas;
- *sensitivity* adaptability, i.e., adapting from data-poor to data-rich scenarios.

The number of satisfied scalability features is proportional to the predictions' robustness.

As for the models and processes, these should also be oriented to satisfying the research questions. Models should include (i) optimisation algorithms (e.g., operation research, simulated annealing) for resource conservation and optimisation; (ii) AI models (e.g., Machine- and Deep-learning, Maximum Entropy, and hybrid mechanistic correlative models) to simulate ecosystem functions and produce information for enhancing and speeding up traditional EMs (e.g., Ecopath with Ecosim); (iii) AI models combined with oceanographic and environmental forecast models (e.g., DIVA, ECHAM5) and signal processing models (e.g., Singular Spectrum Analysis, Fourier Analysis) for long-term projections; (iv) state-space models (e.g., Markov Chain Monte Carlo methods) and unsupervised models to unfold natural complex relations between the data; (v) feature selection and model sensitivity analysis to identify the data carrying the highest information amount; (vi) multidisciplinary AI approaches to fill data gaps;

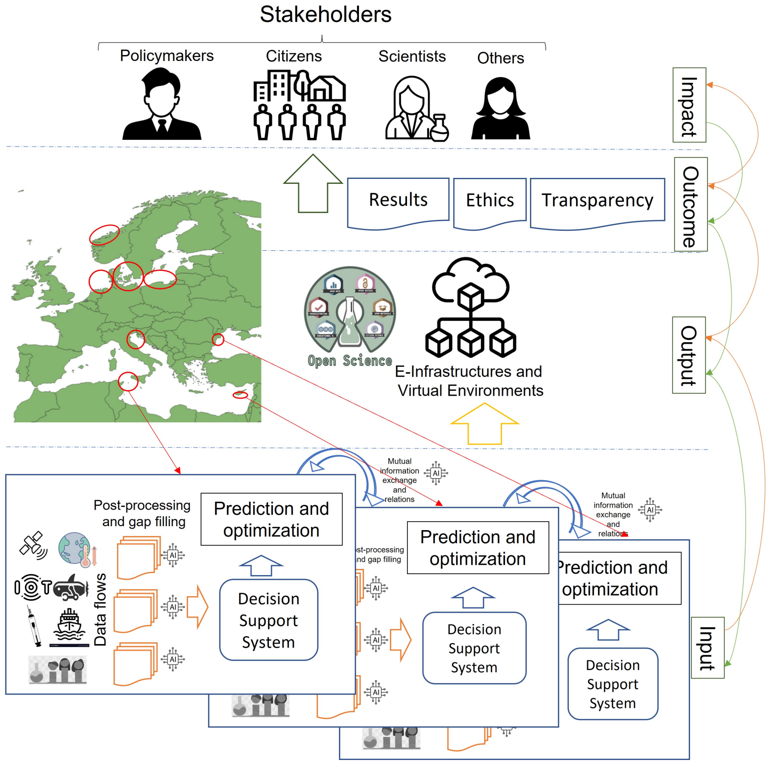


Figure 1: Conceptual schema of a multi-area *h*-DTO.

(vii) Open Science model and data reusability platforms to create interconnections between different DSSs.

4. Conclusions

The *h*-DTO is a novel, dynamic paradigm to unfold the understanding of essential ecosystem functions, the interactions between the driving forces acting on a marine area, and overall make oceans more understandable and predictable. The research questions addressed are indeed classic, unsolved questions in ecosystem modelling. A prototype *h*-DTO could be built on small PMAs guaranteeing easier control of the mathematical, numerical, and data interoperability challenges. The information available for each area should include data on marine biodiversity, cultural heritage, ship traffic, fishing and overfishing, plastic, artificial and organic waste, and climate change effects. The uncertainty of the DSSs should decrease with the increase in the completeness of the data flows. Multidisciplinarity is integral to developing an *h*-DTO because designing all components requires combining different skills from different disciplines. The realisation of an *h*-DTO project alone demonstrates the added value of multidisciplinary research actions for meeting environmental and societal challenges and improving the current ecosystem understanding at the levels of data, ecosystem functions, spatial planning, and predictions.

REFERENCES

1. Piante, C. and Ody, D. Blue growth in the mediterranean sea: the challenge of good environmental status, *MedTrends Project. WWF-France*, p. 192, (2015).

2. Campana, E. F., Ciappi, E. and Coro, G. The role of technology and digital innovation in sustainability and decarbonization of the blue economy, *Bulletin of Geophysics and Oceanography*, **123**, (2021).

3. Coro, G., Ellenbroek, A. and Pagano, P. An open science approach to infer fishing activity pressure on stocks and biodiversity from vessel tracking data, *Ecological Informatics*, **64**, 101384, (2021).

BUILD-IT2023 Workshop

19-20 October, Rome, Italy

DIGITAL TWINS OF THE OCEAN: AUTOGENERATED 3D ENVIRONMENTS FOR VALIDATING OFFSHORE WIND FARM OPERATIONS

Carlos Pereira Santos *and* Harald Warmelink

Breda University of Applied Sciences, Breda, the Netherlands

e-mail: santos.c@buas.nl

Martin Courchesne

CEA Saclay NanoInnov, Gif-sur-Yvette, France

Within the context of the Iliad project, the authors present technical challenges and the first results of having valid 3D scenes of (non-)existing offshore wind farms procedurally and automatically generated within either the Unreal or Unity game engine. The Iliad – Digital Twins of the Ocean project (EU Horizon 2020) aims to develop a 'system of systems' for creating cutting-edge digital twins of specific sea and ocean areas for diverse purposes related to their sustainable use and protection. One of the Iliad pilots addresses the topic of offshore floating wind farm construction or maintenance scenario testing and validation using the Unity 3D game engine. This work will speed up the development of these scenarios by procedurally and automatically creating the Uni-ty 3D scene rather than manually (which is done at present). The main technical challenges concern the data-driven approach, in which a JSON configuration file drives the scene creation. The first results show a base wind farm running in Unreal 5.1. The final product will be able to handle environmental conditions, biological conditions and specific human activities as input parameters.

Keywords: *ocean digital twin, procedural content generation, 3D scene generation, offshore wind farm, game engines*

1. Introduction

The Iliad – Digital Twin of the Ocean project (EU Horizon 2020) aims to develop a 'system of systems' for creating cutting-edge digital twins of specific sea and ocean areas for diverse purposes related to their sustainable use and protection. The project will fuse a large volume of data in a semantically rich and data-agnostic approach to enable simultaneous communication with real-world systems and models. Ontologies and a standard style-layered descriptor will facilitate semantic information and intuitive discovery of underlying information and knowledge to provide a seamless experience. To develop and demonstrate its 'system of systems', the Iliad project relies strongly on 20+ pilots, i.e., actual ocean digital twin in-stances at specific areas for specific purposes [1].

One Iliad pilot concerns the creation of a 3D interactive digital twin of a part of an offshore floating wind farm to test and ultimately validate specific complex construction or maintenance operations scenarios under particular conditions. This could be an existing or soon-to-be-developed farm. The main conditions of interest are weather conditions, which could entail current and extreme conditions observed over a certain time. In terms of scenarios, this pilot will typically deal with construction ships operating large and heavy cranes and other specialised construction equipment to install or upgrade wind turbine components or entire turbines. Since both the vessel and the turbine are floating, wave dynamics are highly important, and thus high-fidelity wave and related or responding physics simulations are paramount. The digital twin would be used by concept engineers and operators to validate and optimise the full operations ahead of time and to define operational windows related to environmental conditions. The primary and practically sole concern is the safety of involved personnel and assets.

One Iliad partner is CEA Saclay NanoInnov (The French Alternative Energies and Atomic Energy Commission). Based in France, CEA is a key player in several areas of research, development and innovation. In particular, the Interactive Simulation Lab develops interactive digital twins for industrial applications. As explained, CEA already uses and applies the Unity game engine to develop, test, and ultimately validate offshore floating wind farm operation scenarios. As such, they are the pilot's leader within the Iliad project. The Unity 3D scene is currently developed manually.

Another Iliad partner involved in this pilot is Breda University of Applied Sciences (BUas). Based in Breda, the Netherlands, BUas' Academy of AI, Games Media offers internationally highly regarded educational and research programmes with innovative technologies (particularly game and media). Over the past 10+ years, several BUas' Games staff developed more and more expertise with procedural and automatic generation of 3D environments using game engines, particularly Unreal, but also Unity. With this expertise and continued interest, BUas sees an opportunity to develop a key system for the Iliad pilot on offshore floating wind farm operation validation, as well as within the whole Iliad 'system of systems' for use by any other interested pilot or future digital twin of the ocean.

In this extended abstract, the authors explain and demonstrate through early prototypes the approach they are currently taking for the procedural and automatic generation of a 3D digital twin of a (non-)existing offshore wind farm with certain parameters working in the Unreal or Unity game engines. Specifically, they describe the main technical challenges in their prototyping with both Unreal and Unity. They subsequently show and briefly explain the first screenshots of a working prototype.

2. Technical Challenges

The main objective of our development work is to be able to have valid 3D scenes of (non-)existing offshore wind farms procedurally and automatically generated within either the Unreal or Unity game engine for a certain/limited set of subsequent use cases, first and foremost for our partner CEA to open up in the Unity editor itself so they can further develop and test the operation scenarios (with limited to no manual 3D scene editing required). Procedurally generated refers to the algorithmic generation of valid 3D content without precise definitions because the algorithms already possess or apply them. This approach leads to diverse results due to the multiple outcomes produced by the parametric algorithms.

The classic example in procedural 3D content generation is the creation of any physically valid bridge crossing a river. An example more relevant to this pilot could concern the generated weather. On the other hand, automatically generated pertains to generating more specific content with no requirement or intention for multiple outcomes or randomness. An example relevant to this pilot would be the use and placement of a 3D model of a wind turbine.

Given this main objective and taking an agile/iterative approach to the development of this system, we are currently working on the following main technical challenges:

- Defining and prioritising relevant input parameters of different kinds with different units of measurement in an appropriate format of ideally relevant and recent standards. Think of the shape of the wind farm area to be rendered, the number of turbines to position in that area with a certain distance between them, the depth of the sea in that area, etc. Also, think of the existing (GIS) standards for defining each input parameter and the total input parameter set.
- Defining the acceptance levels of the Unity or Unreal scene from audio-visual, user interaction, and coding perspectives. Think of answering the following questions:
  - Will it actually work for generating a scene for particular use cases? More specifically, will the end-user be able to at least explore the 3D scene, also using an appropriate VR headset? Or will the developer wanting to use the created scene be able to subsequently develop the operations scenario with limited to no manual scene adjustments?
  - Will the generated scene actually have the desired quality? More specifically, what quality should the water system, 3D models, animations, etc. attain?
- Developing the actual code to procedurally and automatically generate the Unity or Unreal scene from the input parameters mentioned above at the acceptance levels mentioned above. This challenge also covers dealing with assets, notably 3D model files, audio files, game engine plug-ins or extensions, and any licensing

issues involved.

- Validating the resulting Unity or Unreal scene:
  - o Internally, checking all input parameters to the scene output, especially concerning procedural generation techniques.
  - Externally, by having the automatic procedural generation process be evaluated and tested within their work and the scope of Iliad project, and reviewing and assessing what are the automatic generation and manual adjustments required. Second, by looking for other potential use cases and user feedback on how to apply our software to generate different scenarios and possible applications for offshore wind farms.

3. First Results

In terms of initial input parameters design, we are currently working with the following sets:

- Environmental conditions. Realistic depiction of the situation in a location: bathymetry, sediment layer, wave and weather conditions.
- Biological conditions: Realistic depiction of the fauna and flora based on biology charts and bio-mass information.
- Human Activities: Overview of human-related actions or operations in depicted marine regions. These activities encompass a wide range of human interactions.

To ensure that our work remains within scope, we only consider higher biological groups and specific sectors of human activities. Specifically, we are focusing on wind farms and shipping. Wind farms provide information on energy production using wind turbines, cabling, energy production, and ecological pressures such as noise. The shipping sector includes representation of shipping corridors, traffic, and shipping vessels. We have developed a first/preliminary version of a JSON schema with which to define these kinds of parameters. We subsequently wrote the code for Unreal 5.1 to read and interpret the JSON schema to generate the working scene (Fig. 1). Procedural aspects of the resulting 3D scene have so far concerned how to appropriately animate a 3D asset such as a fish or how to achieve a buoyancy effect with floating as-sets (notably ships, of course).

REFERENCES

1. , (2023), *Iliad - Digital Twins of the Ocean*. https://www.ocean-twin.eu/,lastaccessed31May.

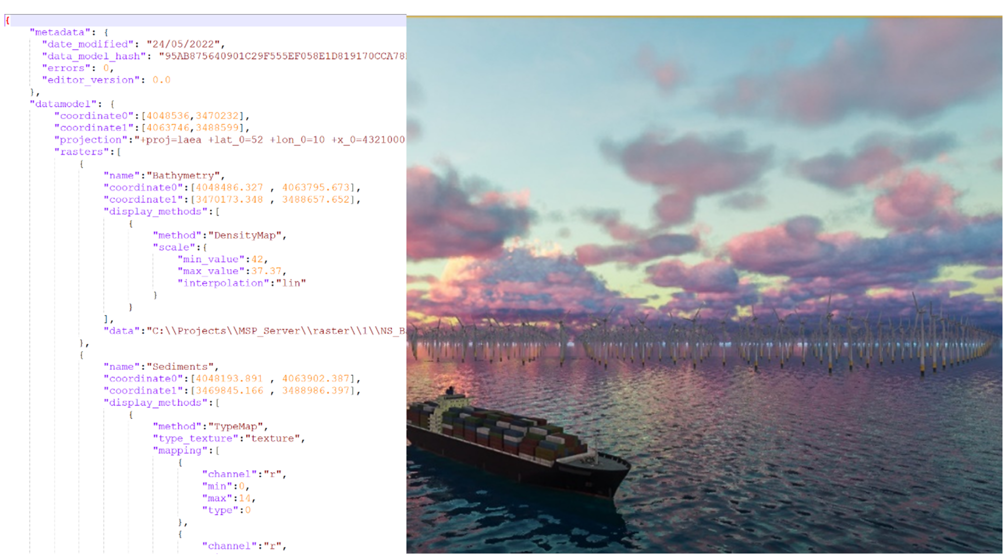


Figure 1: Excerpt of a JSON configuration file (left), leading to an Unreal scene (right).

A SHIP DIGITAL TWIN FOR SAFE AND SUSTAINABLE SHIP OPERATIONS

Mingyang Zhang *and* Spyros Hirdaris
Department of Mechanical Engineering, Marine Technology Group, Aalto University, Espoo, Finland
e-mail: mingyang.0.zhang@aalto.fi

Nikolaos Tsoulakos
Laskaridis Shipping Co. Ltd., Athens, Greece

This paper presents a novel digital twin that can predict ship motions and fuel consumption in real operational conditions. The analysis is based on two optimal Deep Learning Models (DLM) namely (a) a transformer neural network used for the analysis of ship motions and (b) a Long Short-Term Memory (LSTM) network for the prediction of ship fuel consumption. Comparisons of results against sea trial data suggest that subject to further testing and validation DLM could be used as part of a digital twin framework for safe and sustainable ship operations.

Keywords: *digital twins, ship motions, ship fuel consumption, big data science, deep learning*

1. Introduction

Shipping is responsible for more than 90% of the global trade. While it is a relatively safe and clean mode of transportation, it has a significant impact on the environment. Therefore, ensuring the safety and sustainability of ship operations through the proactive utilization of modern technology is a pressing concern.

Digital twins, originally conceptualized by NASA, enable the creation of a digital replica of an artifact to realize and diagnose, critical operational scenarios in real-time [1]. Digital twin technology has been adopted for integrated design and maintenance, as well as performance and safety improvements in modern and increasingly complex ship systems [2]. The traditional format of ship digital twins, which has been used to represent ships as physical entities during design, can be enhanced to monitor and improve ship safety and efficiency, prevent unnecessary outcomes, ensure environmental performance, and minimize downtime. However, it is important to note that the accuracy of physical models may be limited when dealing with complex operational conditions. This is because real operational conditions present unique challenges and factors that may not be adequately captured or accounted for in existing physical models.

Artificial Intelligence (AI) may offer improved predictive capabilities via the integration of deep learning algorithms for the prediction of ship motions [3, 4] and ship fuel consumption [5]. Challenges may arise in selecting a suitable model that matches the data streams, data compatibility and uncertainty modeling. This paper introduces a digital twin that incorporates two optimal deep learning models to achieve accurate predictions of ship motions and fuel consumption for safe and sustainable ship operations.

2. Deep Learning Methods (DLM)

The digital twin presented in this paper comprises two main deep learning Layers namely (a) Layer 1 used to idealise the ship energy system and predict fuel consumption and (b) Layer 2 that can be used to identify ship motions in real operational conditions (see Figure 1).

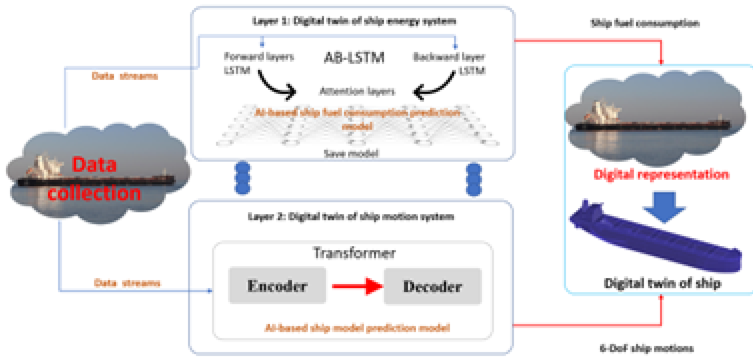


Figure 1: The overall procedure of AI-based digital twin.

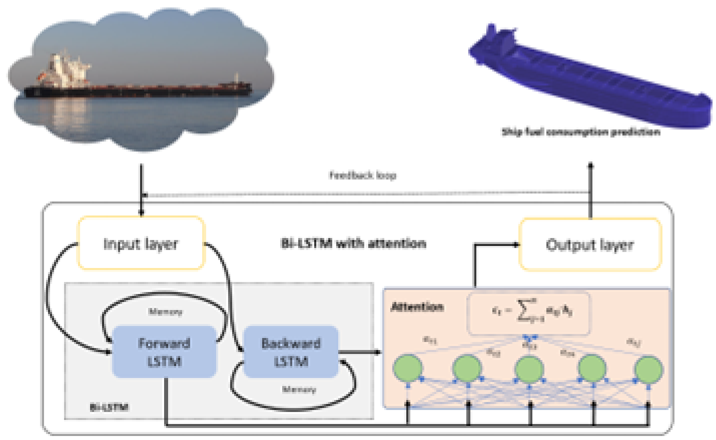


Figure 2: Ship fuel consumption prediction.

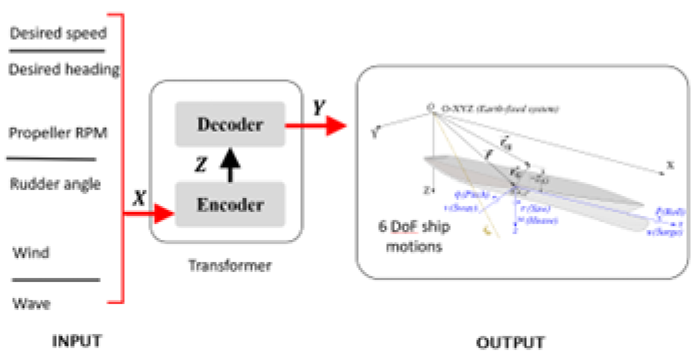


Figure 3: 6-DoF ship motion prediction.

In Layer 1 a wide range of sensors are used to collect navigation data (ship speed, course, etc.), operational/environmental condition data (e.g., draft, trim, wind speed and direction, wave height and direction, etc.), and engine data (fuel flow, propeller rpm, etc.). Then, an LSTM network is used to learn the influence of sailing speed, displacement/draft, trim, weather, sea conditions, etc. on ship fuel consumption (see Figure 2).

In Layer 2 the digital twin utilises Automatic Identification System (AIS), now-cast, and bathymetry da-ta records to extract motion trajectories in real hydro-meteorological conditions (see Figure 3). Then, a rapid Fluid-Structure Interaction (FSI) model and a transformer neural network [4] are employed to generate ship motion patterns while considering the influence of the surrounding water and ship-controlling devices.

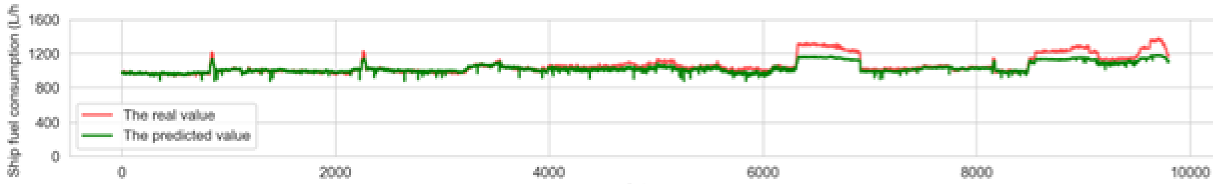


Figure 4: Comparison of fuel consumption from predicted (green line) and real (red line) data.

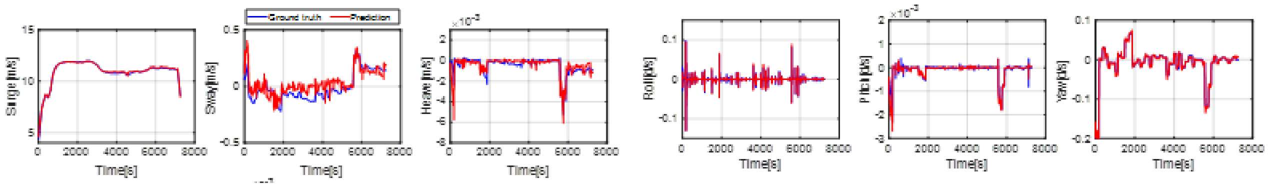


Figure 5: Prediction of ship motion dynamics (blue line is the real data; red line is the predicted results).

3. Results

Layer 1 was trained using big data streams from high frequency sea trial data of a Kamsarmax bulk carrier of Laskaridis Shipping. Comparisons of sea trial records sailing from Canada to Attu Island against DLM suggested that for the selected route the average error in the prediction of the ship fuel consumption during the whole voyage is 2.54% only. These results suggest that subject to further validation over a variety of ships and routes the DLM could be used as an efficient tool for fuel consumption forecasts. Layer 2 was trained by operational data corresponding to Ro-Pax ship voyages and hydro-meteorological conditions between two ports located in the Gulf of Finland. Ship motion predictions indicated that the trained DLM may capture well ship functional behaviour (see Figures 6 and 5).

4. Conclusions

AI-based digital twin models can be used to idealise ship motions and estimate ship fuel consumption. The preliminary results presented in this paper, however promising, are constraint to certain ship types and routes. Hence further testing and validation are necessary with the ultimate aim to develop a framework for the suitable use of digital twin technology for ship operations management (see Figure **??**).

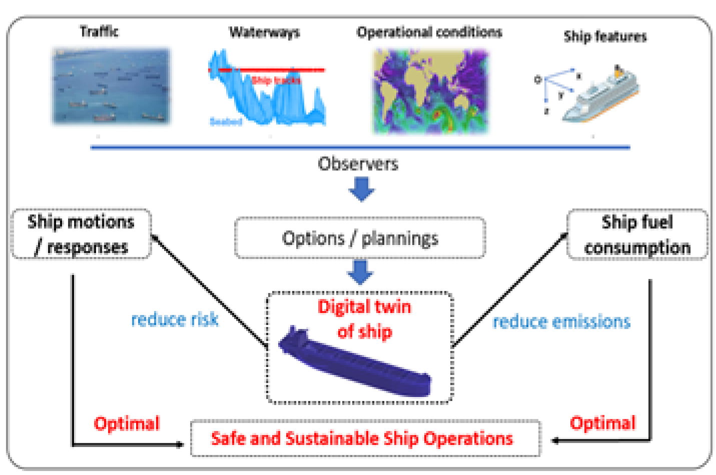


Figure 6: The iterations of digital twins for safe and sustainable ship operations.

ACKNOWLEDGMENTS

All authors acknowledge research funding received funding from Horizons EU 2020 project" RETROFIT SOLUTIONS TO ACHIEVE 55% GHG REDUCTION BY 2030 (RETROFIT55) – Project No.: 101096068".

REFERENCES

1. Allen, B. D. Digital twins and living models at nasa, *Digital Twin Summit*, (2021).

2. Lee, J.-H., Nam, Y.-S., Kim, Y., Liu, Y., Lee, J. and Yang, H. Real-time digital twin for ship operation in waves, *Ocean Engineering*, **266**, 112867, (2022).

3. Nielsen, R. E., Papageorgiou, D., Nalpantidis, L., Jensen, B. T. and Blanke, M. Machine learning enhancement of manoeuvring prediction for ship digital twin using full-scale recordings, *Ocean Engineering*, **257**, 111579, (2022).

4. Zhang, M., Taimuri, G., Zhang, J. and Hirdaris, S. A deep learning method for the prediction of 6-dof ship motions in real conditions, *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, p. 14750902231157852, (2023).

5. Chen, Z. S., Lam, J. S. L. and Xiao, Z. Prediction of harbour vessel fuel consumption based on machine learning approach, *Ocean Engineering*, **278**, 114483, (2023).

DIGITAL TWIN OF AUTONOMOUS SURFACE VEHICLES: FROM STANDARD METHODOLOGIES TOWARDS EXTENDED DATA-BASED MODELS

Marco Bibuli, Roberta Ferretti, Angelo Odetti, Simona Aracri *and* Massimo Caccia
CNR-INM, National Research Council-Institute of Marine Engineering, Genoa, Italy
e-mail: marco.bibuli@cnr.it

Matteo Diez *and* Andrea Serani
CNR-INM, National Research Council-Institute of Marine Engineering, Rome, Italy

The development of Autonomous Surface Vehicles (ASVs) has significantly advanced with the integration of digital twin technology. A digital twin serves as a virtual replica of a physical ASV, providing real-time monitoring, analysis, and control capabilities. This work explores the multifaceted nature of digital twin development for autonomous surface vehicles, highlighting the incorporation of modeling and identification techniques for robotic systems, deriving from the standard procedures employed in naval systems following the ITTC guidelines. Furthermore, extended models, including environmental condition reconstruction and prediction, as well as the integration of AI-based methodologies can be developed, on the basis of standard data gathering procedures and exploiting FAIR (Findable, Accessible, Interoperable, and Reusable) data management principles.

Keywords: *ASV, identification and modeling, standard data gathering, aggregated system, evolution prediction*

1. Introduction

Autonomous Surface Vehicles (ASVs) have emerged as valuable tools for various applications, ranging from marine research to surveillance and transportation to environmental observation and climate change tracking. The integration of digital twin technology enhances ASV capabilities in handling harsh, unforeseen, and fast-changing conditions, thus revolutionizing their operational applications and procedures. This work presents an overview of the different approaches involved in developing a digital twin of an ASV, exploiting standard methodologies such as the International Towing Tank Conference (ITTC) identification maneuvers for the characterization of the vehicle performance [1]. Such modeling allows for estimating the current status of the autonomous platform, as well as predicting the behavior in the near future as a function of the operational requests. High-precision digital twin models can be reached by the integration of environmental forecasting models and Artificial Intelligence (AI) algorithms, allowing an aggregated description of the framework and extending the model representation capabilities. All these approaches are evaluated and integrated to obtain a first example of a digital twin of the SWAMP ASV (Shallow Water Autonomous Multipurpose Platform). SWAMP is an autonomous surface robotic platform (see Fig. 1) specifically designed to work in very shallow water [2]. However, thanks to its reconfigurability and modularity characteristics, it has demonstrated its ability to be used in critical environments (like polar or remote areas) and for monitoring and observing the marine environment during numerous experimental campaigns [3].

2. Digital Twin of the Robotic Platform

The basic step to obtain a digital twin of a marine surface platform is to analyze the maneuvering characteristics of the platform and to derive a suitable and reliable model. Such a model is then employed for the estimation/forecasting of the vehicle motion evolution, as well as a knowledge basis to design the guidance and control system

Figure 1: SWAMP ASV

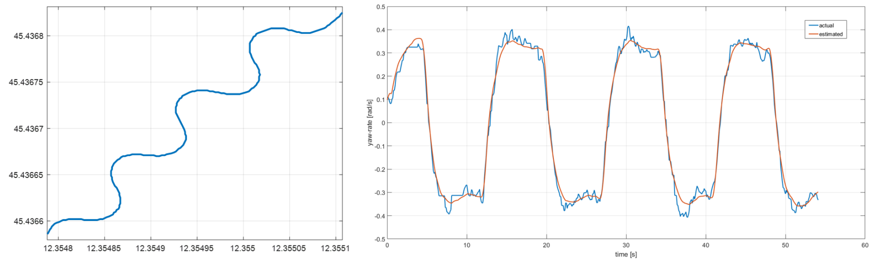


Figure 2: Yaw dynamics modeling of the SWAMP ASV. On the left, a zig-zag maneuver is executed to measure the maneuvering performance. On the right, the raw measurement is compared with the estimated signal generated by the designed mathematical model

needed to provide autonomous capabilities. With the aim of extending the digital twin development process to any general autonomous marine agent, the methodology must rely on standardized processes in such a way as to provide replicable operations. The ITTC procedures are the standard maneuvers widely employed for ship identification at sea and can be directly applied to such autonomous platforms, with appropriate adaptations related to the specificities of this class of robotic vehicles. Additionally, applying ITTC standard maneuverability procedures to groundbreaking marine technology supports the uptake of robotic sea vehicles in observational and commercial employments. These results are particularly important when a new vehicle is headed to operate in cooperation or in the proximity of commercial vessels. Moreover, the development of a digital twin for SWAMP will ensure the accurate planning and environmental mindfulness required when operating in pristine areas, whether these are located in the polar regions or simply in Marine Protected Areas (MPAs). The evaluation of the maneuvering performance is exploited to tailor-design a mathematical model able to represent such maneuvering characteristics with proper accuracy; an example of yaw dynamic modeling of the SWAMP ASV is reported in Fig. 2. The latter procedure is applied to all the measurable entities related to the vehicle motion capabilities so that the digital twin complete model embeds the global characteristics of the autonomous platform.

3. Data Management for Extended Digital Twin

Since the digital twin of an ASV relies on a vast amount of data collected from various sensors and systems, in order to provide data integrity, accessibility, and interoperability, a consistent data management approach must be employed. For such a reason, the data management guiding principles known as FAIR (Findable, Accessible, Interoperable, and Reusable) [4] are exploited. The principles emphasize machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with no or minimal human intervention), so that, if data are consistently gathered and tagged, computer-based systems can automatically aggregate the information, providing extended digital models. To this end, work is underway [5] to extend the SWAMP management architecture for data acquisition, both robotic and environmental, producing "FAIR by default" data. With this ambitious definition, we refer to the practice of developing procedures for data acquisition, management, and sharing integrated from the beginning into the design and operation of the robot. A first example of an automated procedure, allowing to transform raw data into standard data format, characterized by a high level of fairness, already containing the necessary metadata to understand and reuse them, with a few simple steps was implemented and tested for SWAMP. FAIR data management facilitates a seamless integration of data coming from different sources, enhances simulation capabilities, and enables more effective decision-making for optimizing the performance and operational adaptation of the SWAMP, including the development of an extended digital twin of the robotic platform, relying on data-based models. Thanks to data standardization, the digital twin can be validated against real-world scenarios, utilizing historical data, sensor observations, and environmental forecasts. This validation process ensures that the digital twin accurately represents the ASV's behavior, accounting for both internal and external factors. Furthermore, it supports collaboration and data sharing, driving innovation and advancements in autonomous surface robotic platforms thanks to the reliability of the corresponding digital twins.

The further outcome of FAIR-based data acquisition is the availability of a number of compatible data sets that can be integrated to enhance the modeling of both the autonomous platform and its sensors and the operational environment in an aggregated way. With this approach in mind, it would be possible to predict the evolution of the system as a whole, virtualizing the presence of the operators while at the same time increasing the promptness of the interventions. For instance, imagining the periodic observation of a coast of sandy beaches subject to strong storm surges, the development of an extended digital twin will allow the design of proper patrolling and data gathering schedules, focusing on the most sensitive spots in the area of interest (on the basis of the environmental evolution prediction) and the subsequent autonomous execution of monitoring of bathymetry evolution and sediment displacement by means of the ASV.

As a final step, beyond the current state of the art, the integration of AI-based methodologies for data analysis and performance improvement is under investigation. AI-based decision support systems are currently advancing at a hectic pace, revolutionizing the way humans interact with computer-based systems. A preliminary experiment focused on learning-by-imitation control of SWAMP ASV was carried out and provided encouraging results as reported in [6]. Finally, a high-level decision system can be developed in order to schedule and control the employment of one or more ASV, collecting environmental data and, in real-time, re-planning the operations on the basis of the online predicted evolution of the overall systems based on both the information gathered as well as on the forecast obtained from the extended digital twin. The ASV can leverage its digital twin to perform a task replanning. This consists of modifying the original mission in case of changes in the initial conditions. For example, if the battery is draining faster than expected due to currents, winds, or obstacles, the digital twin can estimate the remaining consumption and decide how to complete the mission. The AI-based decision support system together with a digital twin can replan in an efficient and effective way, that is, minimizing the energy costs and the navigation times and maximizing the satisfaction of the mission criteria. In this way, the vehicle can evaluate different alternatives and choose the best one based on predefined criteria.

ACKNOWLEDGMENTS

This work is developed within the framework of the National Recovery and Resilience Plan (PNRR), RAISE-Robotics and AI for Socioeconomic Empowerment, and the Horizon Europe Twinning Project MONUSEN - Montenegrin Centre for Underwater Sensor Networks.

REFERENCES

1. Ferretti, R., Bibuli, M., Bruzzone, G., Caccia, M., Aracri, S., Motta, C., Odetti, A., Diez, M. and Serani, A. Procedures for maneuverability characterization: From ships to marine robots, *Computer Applications and Information Technology in the Maritime Industries*, COMPIT 2023, (2023).

2. Odetti, A., Bruzzone, G., Altosole, M., Viviani, M. and Caccia, M. Swamp, an autonomous surface vehicle expressly designed for extremely shallow waters, *Ocean Engineering*, **216**, 108205, (2020).

3. Ferreira, F., et al. Heterogeneous marine robotic system for environmental monitoring missions, *2023 IEEE Underwater Technology (UT)*, pp. 1–5, IEEE, (2023).

4. Wilkinson, M. D., et al. The fair guiding principles for scientific data management and stewardship, *Scientific data*, **3** (1), 1–9, (2016).

5. Aracri, S., Ferretti, R., Corrado, M., Ferreira, F., Bibuli, M., de Pascalis, F., Odetti, A., Bruzzone, G. and Caccia, M. Open science in marine robotics, *BOOK OF ABSTRACTS – ICOD 2022*, pp. 96–100, (2022).

6. Odetti, A., Bibuli, M., Bruzzone, G., Cervellera, C., Ferretti, R., Gaggero, M., Zereik, E. and Caccia, M. A preliminary experiment combining marine robotics and citizenship engagement using imitation learning, *IFAC-PapersOnLine*, **53** (2), 14576–14581, (2020).

DRIVING SUSTAINABLE CHANGE: ANALYZING AND PROMOTING THE ADOPTION OF ELECTRIC VEHICLES

Chiara Ravazzi *and* Fabrizio Dabbene
CNR-IEIIT, c/o Politecnico di Torino, Italy
e-mail: chiara.ravazzi@cnr.it

Valentina Breschi, Mara Tanelli, *and* Silvia Strada
Politecnico di Milano (DEI), Milano, Italy

The mobility is undergoing significant changes, including the emergence of battery-operated vehicles, user-ship models, autonomous driving, and digital mobility services. To address these shifts, we propose a human-centric design framework for transitioning to electric vehicles (EVs). This framework empowers policymakers to shape the future of mobility by considering the interplay of social and personal factors influencing individual mobility behaviors. Using quantitative representations of socio-economic identities and data-driven social network analysis, we uncover the mechanisms driving the adoption of new mobility paradigms. This understanding facilitates the design of policies and services aligned with these insights. Applying our approach to real-world data from ICE vehicles in Italy, we demonstrate the framework's potential in designing effective policies that promote greener mobility habits. Policymakers can leverage this framework to implement policies fostering the adoption of greener mobility options, thereby contributing to the fight against climate change and promoting inclusiveness in the transition to sustainable mobility.

1. Introduction

Mobility is a vital aspect of our lives and cities, and it will continue to be crucial in the future of Smart Cities. Before the pandemic, transportation accounted for 20% of overall energy consumption [1], contributing significantly to greenhouse gas emissions. Addressing climate change requires reducing these emissions, and Electric Vehicles (EVs) play a key role in achieving greener mobility and pollution reduction [2]. However, obstacles hinder the widespread adoption of EVs, including ownership costs, driving range, and charging time [3]. Fuel prices, consumer characteristics, availability of charging stations, and social norms also influence adoption. Achieving full electrification is a complex process due to these factors.

To address this challenge, we propose an integrated and human-centered framework that considers the conflicting constraints arising from infrastructure limitations, management issues, and user needs. By examining both personal inclinations and social influences, our framework takes into account users' social connections to understand how interactions can promote sustainable mobility habits. This insight can inform the design of social-aware public policies that leverage users' social bonds to encourage acceptance of new mobility solutions.

Aligned with the aforementioned considerations, this study introduces a human-centered architecture designed to assist policymakers in three main areas: (i) analyzing the adoption of electric vehicles (EVs), (ii) developing intervention policies to encourage EV adoption, and (iii) quantitatively evaluating the costs and benefits of these policies using available data.

2. Main contribution

Figure 1 illustrates the conceptual framework, depicting the various stakeholders involved in a mobility system, including users, the environment or network, management, and infrastructure.

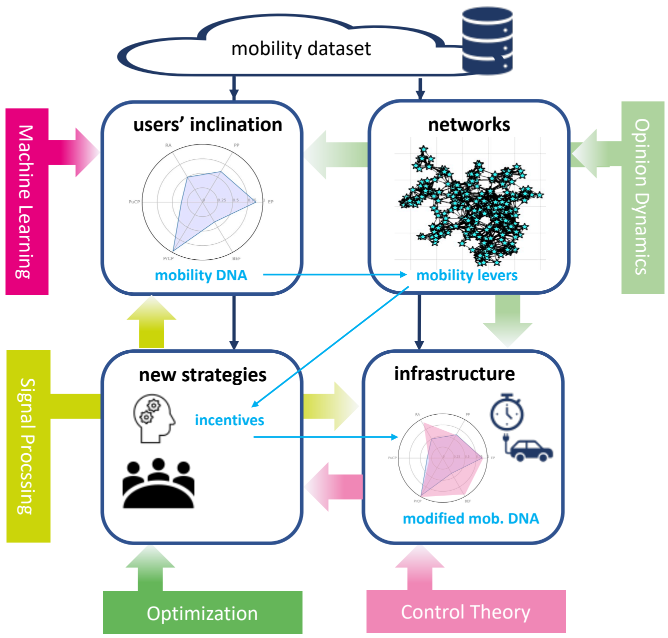
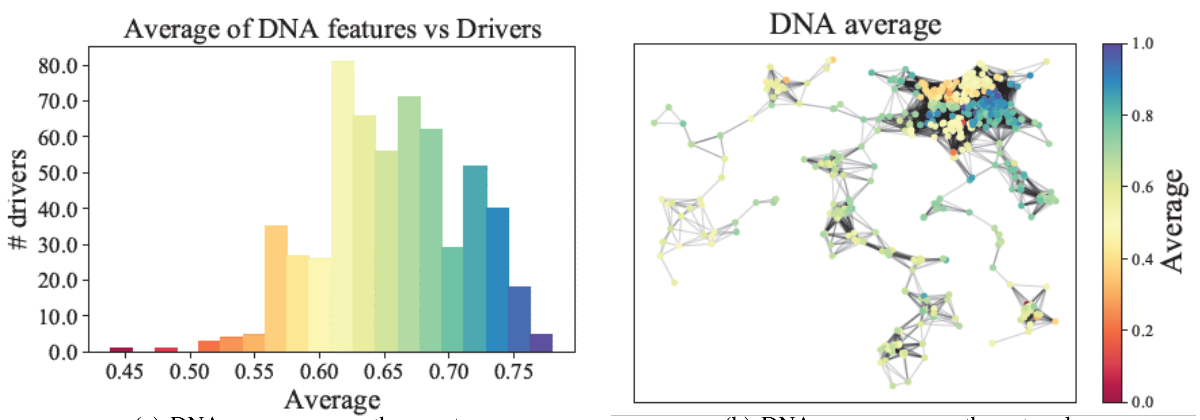


Figure 1: Interactions between different players and methodologies in the proposed human-centered design.



(a) DNAs averages over the agents.

(b) DNAs averages over the network.

Figure 2: DNAs average distribution over the network and the drivers.

By leveraging a dataset of anonymized trips, we progressively demonstrate how quantitative and unbiased information on individual mobility behaviors can provide insights into users' adaptability to EV adoption when combined with known limitations of EV technology and available infrastructure. Building upon the work of Fugiglando et al. [4], our contribution is the introduction of a novel concept known as the users' EV-adoptability DNA [5]. This compact indicator combines synthetic features that characterize the adaptability of each individual to an immediate switch to an EV based on their travel patterns. Unlike the user description used in our previous work [6], this granular representation enables policymakers to visualize key factors influencing the spread of EVs while allowing for flexibility in augmenting it with additional socio-economic features, if available.

Additionally, we propose an approach that utilizes the available mobility data to establish a proximity bond between agents, thereby enabling the assessment of the role played by homophily (similarity) in the adoption process. Through simulations, we mathematically describe the impact of individual adaptability to EVs and proximity-based relationships on the diffusion of EVs within a community. This formalizes the interplay between different layers of our proposed framework. To model the adoption phenomenon, we employ an irreversible cascade model [7] to characterize the effect of social contagion on EV adoption. This model results in a binary representation of individual attitudes, where a transition from one predisposition state to the other is driven by users' attributes surpassing a personal threshold. In contrast to McCoy and Lyons (2014, [8]), where individual thresholds are randomly assigned, our framework associates thresholds directly with each individual's resistance to adopting EV technology, thus connecting them to the EV-adoptability DNA. In our perspective, the thresholds represent the minimum number of EV-accepting neighbors required for an individual to consider EVs as a viable mobility solution through mutual influence alone.

To analyze and design policies, we propose the use of the cascade model within limited time periods in our framework. Through extensive simulations, we demonstrate how our framework can be applied to study the evolution of users' predispositions toward EVs over time, test various policies to promote EV adoption based on enforcing purchase power and the public charging potential, and quantitatively analyze their socio-economic and environmental impact.

To evaluate the effectiveness of the implemented policies, it is crucial to analyze their outcomes from a human-centered perspective, considering factors such as sustainability, environmental impact, and social inclusiveness. In order to achieve this, we introduce a comprehensive set of Key Performance Indicators (KPIs) that provide a self-contained framework for quantifying these seemingly subjective aspects. These indicators enable us to assess the policies based on various perspectives, including efficiency, effectiveness, and sustainability. By utilizing these KPIs, we can quantitatively compare different policies, extending the evaluation beyond the mere measurement of changes in individual attitudes towards EVs. This approach allows for a more holistic and comprehensive assessment of the policies' impact.

REFERENCES

1. Docherty, I., Marsden, G. and Anable, J. The governance of smart mobility, *Transportation Research, Part A: Policy and Practice*, **115**, (2017).

2. Popovich, R. D. T. E., N.D. and Phadke, A. Economic, environmental and grid-resilience benefits of converting diesel trains to battery-electric, *Nature Energy*, **6**, 1017–1025, (2021).

3. Helmus, J. R., Lees, M. H. and van den Hoed, R. A validated agent-based model for stress testing charging infrastructure utilization, *Transportation Research Part A: Policy and Practice*, **159**, 237–262, (2022).

4. Fugiglando, U., Santi, P., Milardo, S., Abida, K. and Ratti, C. Characterizing the "driver dna" through can bus data analysis, New York, NY, USA, pp. 37–41, CarSys '17, Association for Computing Machinery, (2017).

5. Breschi, V., Ravazzi, C., Strada, S., Dabbene, F. and Tanelli, M. Driving electric vehicles' mass adoption: An architecture for the design of human-centric policies to meet climate and societal goals, *Transportation Research Part A: Policy and Practice*, **171**, 103651, (2023).

6. Breschi, V., Ravazzi, C., Strada, S., Dabbene, F. and Tanelli, M. Fostering the mass adoption of electric vehicles: a network-based approach, *IEEE Transactions on Control of Network Systems*, pp. 1–1, (2022).

7. Acemoglu, D., Ozdaglar, A. and Yildiz, E. Diffusion of innovations in social networks, *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 2329–2334, (2011).

8. McCoy, D. and Lyons, S. The diffusion of electric vehicles: An agent-based microsimulation, *MPRA Paper 54560, University Library of Munich, Germany*, (2014).

A SIMULATION APPROACH TO ENHANCE WAREHOUSE LOGISTICS PERFORMANCE IN THE CERAMIC TILE SECTOR

Marco Taccini, Giulia Dotti, *and* Manuel Iori
University of Modena and Reggio Emilia, Reggio Emilia, Italy
e-mail: {marco.taccini, giulia.dotti, manuel.iori}@unimore.it

Anand Subramanian
Universidade Federal da Paraíba, Mangabeira João Pessoa–PB, Brazil
email: anand@ci.ufpb.br

This study aims to improve warehouse logistics performance for an international ceramic tile company based in Italy by identifying an efficient storage policy for the buffer area between the production plant and the logistics department. Due to the lack of homogeneity in the ceramic tile production process and the need for uniform tile orders, the policy must categorize products accordingly. The current policy classifies tiles based on technical properties, while the newly devised policy considers their downstream destination. A simulation using Salabim, a Python-based open-source software, demonstrated that the proposed policy outperformed the current one on all indicators. A sensitivity analysis showed consistent superiority of the devised policy under different scenarios, regardless of production increase. As a result, the company decided to implement the proposed policy, estimating a reduction in costs for the buffer area emptying process. This research contributes to simulation-based decision-making in material management and showcases Salabim's effectiveness in modeling complex systems.

Keywords: *discrete event simulation, ceramic tile, warehouse management.*

1. Introduction

In the past decade, the global ceramic tile market has experienced significant growth with 18.3 billion square meters produced worldwide in 2021, generating revenues of €6.2 billion in Italy [1, 2]. However, the tile sector faces challenges due to Lack of Homogeneity in the Product (LHP) [3], which conflicts with customer demands for uniform tile orders. To address this issue, tile companies must incorporate a classification stage in their production process to organize products into homogeneous subgroups and facilitate uniform orders. As a consequence, the classification process must be carefully evaluated because it can affect material handling performance. There is a vast literature about the tile industry, with many studies regarding the production process and the sustainability perspective. A growing body of literature studied the application of decision support systems in the ceramic sector, such as for inventory control and production planning and scheduling. Although considerable research has been done on different aspects of the tile industry, much less is known about the application of simulation in the sector. To the best of our knowledge, only a few papers investigated process simulation tools as performance measurement, and none of these studies applied Discrete Event Simulation (DES) using non-commercial software. As a consequence, Salabim [4], an open-source simulation software developed in Python, was chosen as the simulation tool due to its advantageous features such as the activate/passivate/hold paradigm, animation, queues, tracing, and statistical distributions. However, despite its potential, there is limited literature available on Salabim, with only a few scheduling and healthcare applications found. The extended version of this abstract can be found in [5].

2. Problem description

The study focuses on the buffer area placed between the production plant and the logistic department, schematized as a matrix in Figure 1. Each cell can contain four pallets, vertically stacked. After they are produced, the

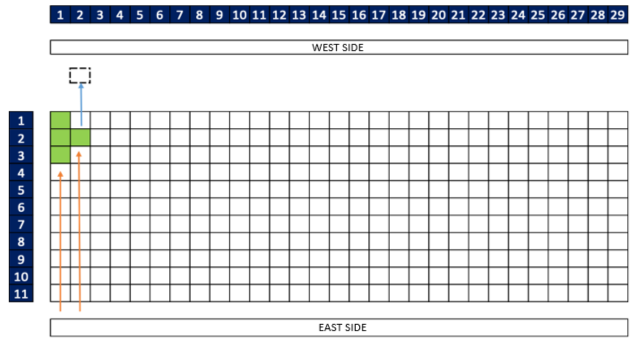


Figure 1: Storage area representation, taken from [5].

pallets are moved by automated vehicles to the buffer area. There, they are stored temporarily in the uppermost available cell until operators take them from the top and pack them into the logistics department. Each column in the area can contain a single class, and only when all columns are already allocated, the system allows for class mixing to avoid production downtime. However, mixing products of different classes leads to penalties because operators manually pick pallets from the buffer area, and their work efficiency is highly impacted by pallet positions. Consequently, the storage policy must carefully assign pallets to cells, dividing pallets into classes based on criteria to cluster them and facilitate picking activities. Each pallet of tiles is identified by a code that considers both commercial and technical characteristics. The current classification strategy clusters columns based on codes. The study focuses on the definition of a new policy that categorizes pallets according to their final downstream destination rather than their codes.

3. Methodology

The two classification policies were compared through a DES implemented with Salabim. The simulation model was verified and validated through statistical analysis to ensure real-world coherence. The simulation was run considering the data of 30 real-world production days and the values of performance indicators were collected for comparison through statistical tests. Since the logistics department works on a two-shift schedule during working days, while the production process never stops, two distinct sets of indicators were collected to evaluate both the described scenarios, without and with operators. In the first scenario the following performance indicators were selected:

- time (in minutes) elapsed before the first occurrence of the mixed columns;
- time (in minutes) elapsed before the area being filled;
- number of pallets lost due to lack of space.

The real-world system was designed to enable operators to handle the production flow without exceeding the available buffer storage area. Therefore, in the second scenario, the probability of the area being filled is not considered, and the focus moves to space utilization. As a consequence, the only indicator selected for this scenario is the maximum number of columns used at the same time. Due to positive market trends in the ceramic tile sector, it was crucial for the company to understand the new policy's impact on various increasing production scenarios. Therefore, seven different scenarios were simulated, incrementally increasing production quantity by 5%, from 0% to 30%. For each scenario, the simulation was run over 30 days.

4. Results

The results are presented in Figure 2. Focusing on shifts without operators, the mixed columns in the devised policy happen, on average, 37.6% minutes later compared to the current policy. This delay has a positive effect on worker activities as it significantly reduces the time needed for emptying operations. The second chart shows an average 5.2% delay in the time it takes to fill the area with the devised policy, resulting in increased capacity and decreased production downtime. Moreover, according to the third chart, with the proposed policy there was a significant decrease in the number of pallets lost on days when the current policy reaches its maximum storage capacity, consequently reducing the economic loss. Overall, the discussed results showed that the proposed classification policy enhances storage activity flexibility, enabling the company to reconsider required operators and shift distribution, leading to economic benefits. Taking shifts with operators into account, the fourth chart showcases a significant decrease in the maximum number of columns used at the same time, highlighting the potential to reduce the buffer area size by approximately 17.5% without affecting production capacity.

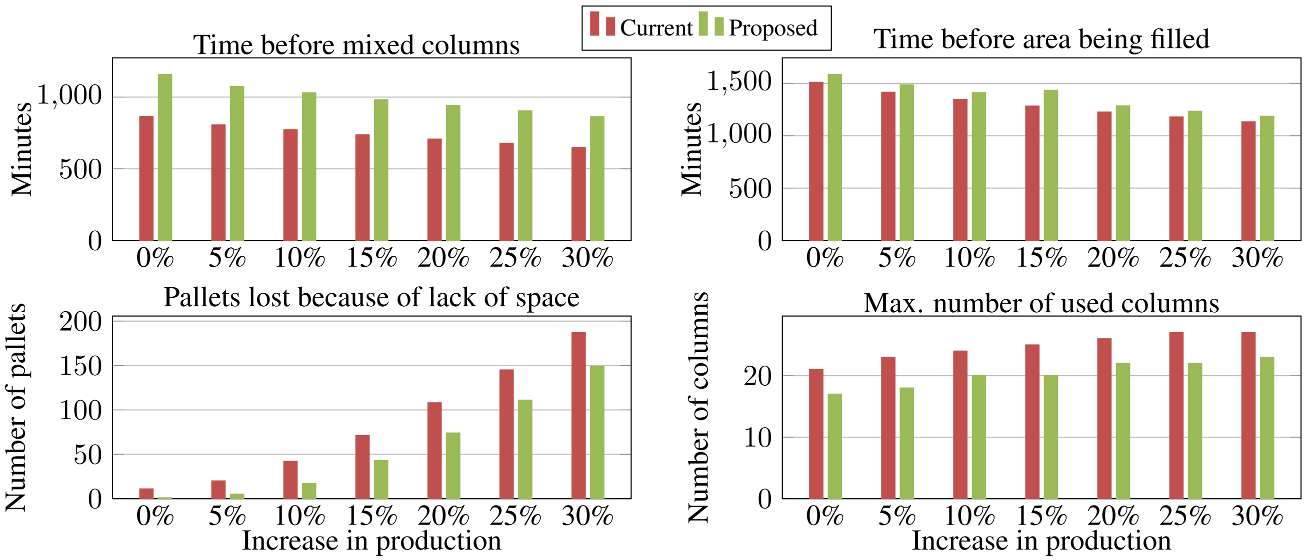


Figure 2: Simulation results, taken from [5].

5. Conclusions

The experimental results confirmed the benefits of the devised classification policy, both in terms of economic efficiency and flexibility. Further economic studies were conducted to estimate the potential cost savings associated with the buffer area emptying process and the analysis revealed that the proposed policy is expected to reduce the costs by 17%. As a consequence, the company decided to implement the new policy, providing strategic benefits to managerial decision-makers. Finally, the simulation model developed in this study is a valid representation of the process and can be utilized for further research regarding new classification policies for continuous improvement.

REFERENCES

1. Baraldi, L. World production and consumption of ceramic tiles, *Ceramic World Review*, **148**, 36–57, (2022).

2. ACIMAC Research Department. Manufacturing Economics Study, World production and consumption of ceramic tiles, (2022).

3. Boza, A., Alemany, M., Alarcón, F. and Cuenca, L. A model-driven dss architecture for delivery management in collaborative supply chains with lack of homogeneity in products, *Production Planning & Control*, **25** (8), 650–661, (2014).

4. van der Ham, R. Salabim: Discrete event simulation and animation in python, *Journal of Open Source Software*, **3** (27), 767–768, (2018).

5. Taccini, M., Dotti, G., Iori, M. and Subramanian, A. Improving buffer storage performance in ceramic tile industry via simulation, *Proceedings of the 2023 Winter Simulation Conference*, **forthcoming**, (2023).

BRINGING DIGITAL AND CIRCULAR ECONOMY TO RISK MANAGEMENT IN MAJOR-ACCIDENT HAZARDS RISK INSTALLATIONS

Carla Mazziotti Gomez de Teran
Department of Engineering, ICT and Technology for Energy and Transport (DIITET), Rome, Italy
email: carla.mazziotti@cnr.it

Massimo Mari
Institute of Atmospheric Pollution Research, Research Area Rome-1, Italy

Climate change affects many regions of the world. Southern Europe and the Mediterranean Area are nowadays severely hit by the adverse effects of climate change. Eminent studies (JRC) show the impressive growing projection of the huge economic impact caused by extreme weather events and expected on critical European infrastructures of the industrial and social sector for the following decades. Minimizing the negative impacts caused by so-called NaTech (Natural Hazards Triggering Technological Disasters) events is also an objective of Euro- pean policies. On the one hand, the necessity is for a flexible, accessible tool for digitalizing risk management. On the other hand, such automation, cognitive applications, advanced analytics, and pattern recognition at scale require new transversal skills. In this context, the requests for sound data and information foresee specific measurement and monitoring systems to pursue concrete actions and achieve measurable results. It may cooperate with such a vision, the circular economy approach driving towards innovation and the virtualization of activities. In risk management, a DT that entirely recreates state-of-the-art and possible alternatives can be effective in modelling emergencies, especially for heavy and hazardous industries. Also, the ISO 14001 standard remarks the opportunity for the companies to optimize the functionality and performance of the production sites by the commitment to protecting the environment.

Keywords: *risk analysis, Risk management, circular economy, digital twins, NaTech events*

1. Introduction

Global climate change shows each day its heavy environmental and economic impacts on the world. Despite the entire industrial development history and path, poor and developing countries are paying higher prices. At the same time, because of the infrastructure intensification and profound urbanization, the expected economic damages on critical infrastructures will be huge also for the European countries. Regarding the climate change impacts and NaTech events, looking at recent data collections and scenarios for the next future (Fig. 1, source: Euro- stat/EEA), Southern and Mediterranean Europe (Spain; Italy; France; Greece) must be considered as the region characterized by the highest level of risk for critical infrastructures and industries. In a global society still just oriented to prosperity production, if the climatic and environmental crisis would not be enough to move, the economic amount and the potential harm quantification should be able to sweep away any remaining obstacle on the way to a circular economic transition, that is nowadays so essential to permit a genuine cut of the anthropic impacts along the entire production and consumption cycles. With climate change affecting some natural hazards triggers and increasing human development (booming urbanization; rapid industrialization), NaTech risk is expected to increase soon.

2. Methods

European countries have adopted the Seveso Directive to manage and control the major accident hazard at industrial installations involving dangerous substances. National authorities shall implement effective safety policies

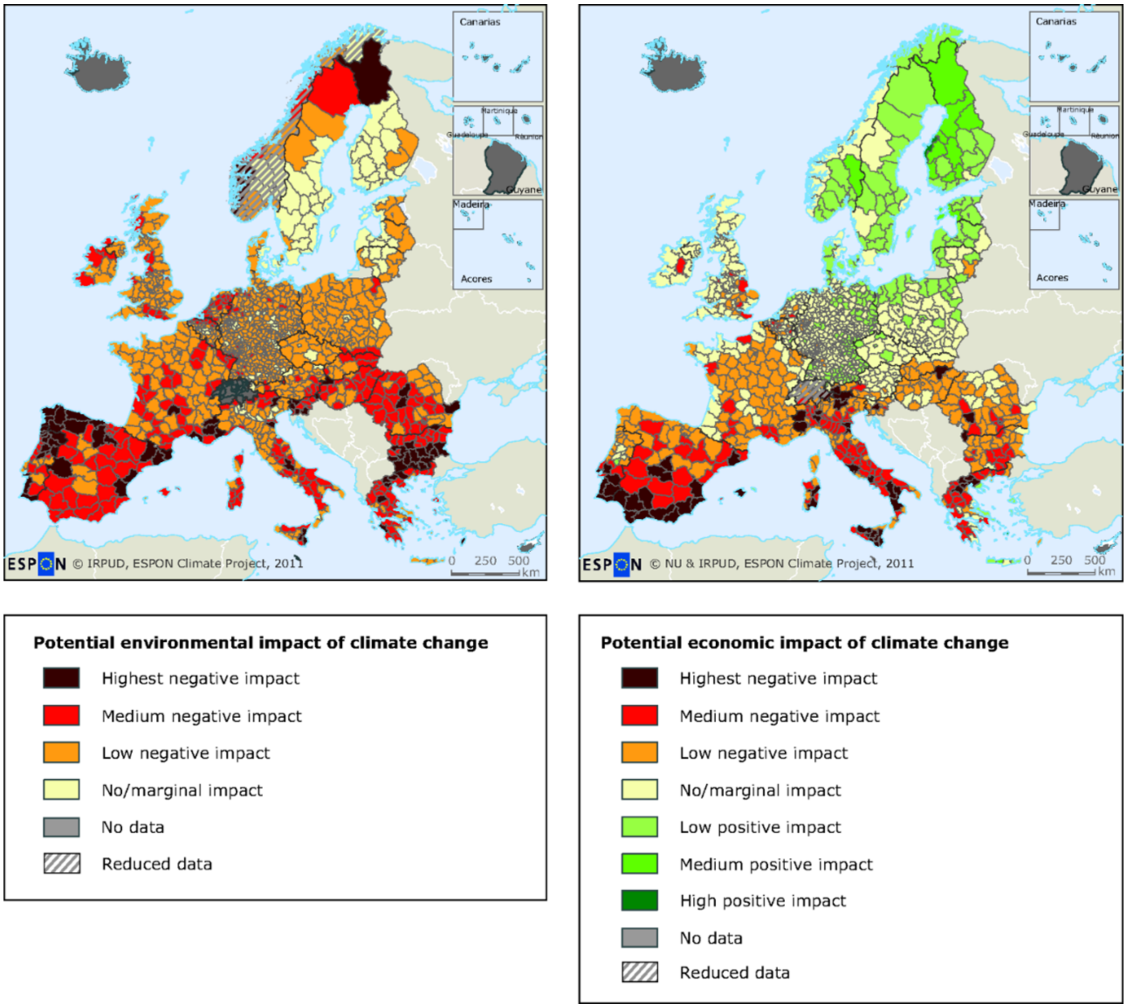


Figure 1: Potential impacts of climate change on environment and on economical infrastructures. Source: Eurostat/EEA; Impact calculated as combination of regional exposure to climatic changes and recent data on regional sensitivity.

by the last law application (Dir. n. 2012/18/EC). The circular economic transition should be able to reduce the impacts linked to production and consumption cycles and minimize risks connected to industrial/chemical plants. Regarding this, two main circular economy pitches could play an important role in NaTech risk management:

1. the introduction of biological elements (often from waste) in the production cycles to replace the use of traditional raw materials [1];
2. digital information, communication, and consultation.

This work aims to describe the fundamental role of digital information and communication in industrial risk management. Experiences show that many past NaTech accidents could have been prevented if only better awareness and understanding of the risk had existed [2].

As represented in the following figure describing the risk management process (Fig. 2, "the process of risk management", source: ISO 3100:2009), communication, consultation, monitoring, and review of data set and information is ever present in the considered risk management process. Moreover, access to the European market is allowed if the substances produced and marketed comply with the requirements of the REACH Regulation ("No data, no market"). For the chemical industry sector, the BAT conclusions, according to Directive 2010/75/EU, identify the best available techniques for the plant operation.

The circular economy allows exploiting "digital technologies" potential, NaTech accidents, with improved knowledge of potential risks.

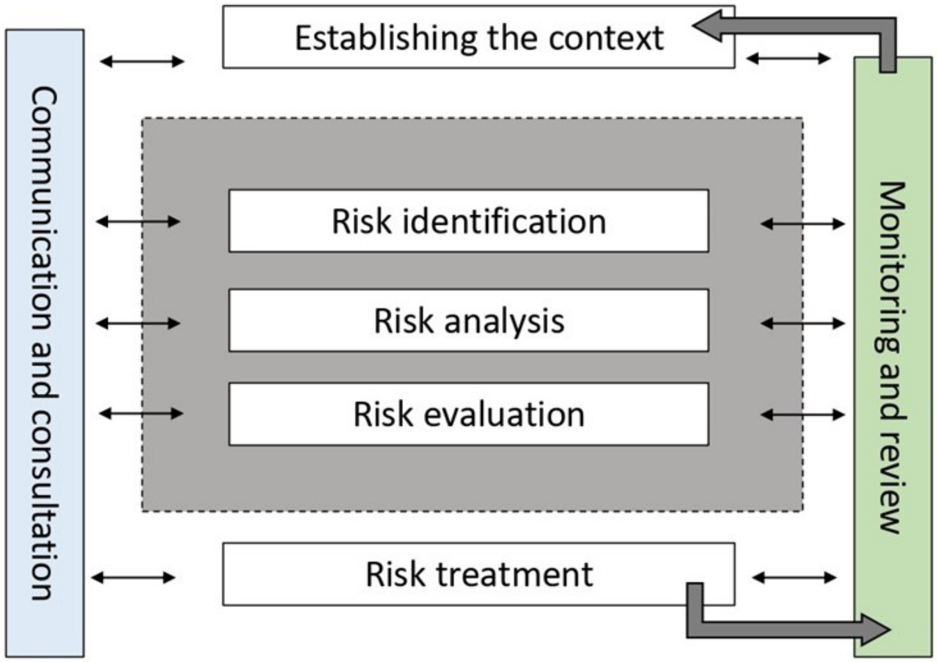


Figure 2: The process of risk management. Source: ISO 3100:2009; JRC Technical Report 2022 "NaTech risk management"

3. Discussion

The fundamental role of digital technologies and digital data is clear and undeniable in each phase represented of the risk management process:

1. Communication and consultation with internal and external stakeholders are essential to ensure taking into account all interests related to risk;
2. Establishing the context is headed by the definition of internal and external parameters to be considered in the risk analysis;
3. Risk assessment is a complex risk identification, analysis, and evaluation process. NaTech risk assessment requires an amount of input data, such as information on the natural hazard, the vulnerable equipment, damage models and data linking damage to releases, consequences analysis models, likelihood estimates, and information on the risk receptors [3];
4. Risk treatment aims at selecting and implementing actions to reduce risks; this phase requires an evaluation analysis of benefits/costs;
5. Monitoring and review are an integral part of the process, implemented by audits or inspections, ensure the effectiveness of controls and help to obtain additional information for risk assessment.

So digital risk management should help in each described phase of the process, providing in-depth analytics to help the organization to monitor the compliance status by the chosen indexes and current threat level for all risk factors. The correct digital risk management will provide, in fact, clear visibility of current threats as well as remediation actions that may be needed, shortening the time to respond to threats. Adequate controls and specific monitoring are activated to satisfy the requests for data and information foreseen by the law simultaneously as effective communication guided by an appropriate communication strategy. Also, the 2015 ISO 14001 standard cooperates with such a framework allowing the companies to optimize the functionality and performance of the production sites through the commitment to protecting the environment. The reporting must be addressed to the interested public and internal company operators to make them more aware and involved, for example, by disseminating information that non-experts can use. Therefore, an initial effort for monitoring is required, together with the necessary resources for acquiring and storing the essential data and information for assessing the impact of systems and infrastructure reliability. Virtual and Augmented Realities can effectively display emergencies and critical situations, especially for large industries. Training, system monitoring, preventive machine maintenance, logistics, and other uses may use digital applications may be used to improve the process [4]. While virtual reality uses an entirely computer-simulated environment and allows the evaluation of different options, augmented reality technology represents the overlay of digital information over the actual company background. Virtual Reality (VR) has brilliantly been used in some functional applications for risk and predictive management, enabling companies to plan maintenance tasks and minimize the consequences of accidents. Quality Assurance is also systematically provided by quick checks of the operating systems. The implementation of VR entails the improvement of the Operational Expenditures - OPEX budget based on virtual close supervision, sapient distribution of tasks, and cost reductions. On the other hand, augmented Reality (AR) guides operators through standardized processes by showing each step to follow with audio and visual instructions using video and animations, staying away from the site where the operations occur. It can be effectively used for manufacturing training [5] and in a more rapid change of productive activities. In complex installations, for instance, chemical installations, some safety devices for preventing accidents connected with the formation of explosive atmospheres, such as the systems for conveying and venting unwanted or excess gas streams to the oxidation system, also take the form of measures for preventing the release of polluting emissions into the atmosphere. Diffuse emissions will therefore be considered during the design phase of the new installations by assuming possible plant solutions and subsequent management activities (lifecycle thinking). Usually, in the chemical industry, the periodic monitoring of diffuse emissions is in the lines of accessible components, such as valves, safety valves, connectors, pumps, flanges, pipe ends, agitators, and compressors. The measurement applies the LDAR methodology according to the following technical standards: EPA 453/R-95-017 “Protocol for Equipment Leak Emission Estimates”; EPA Method 21 (Annex F of EPA protocol 453/R-95-017); UNI EN 15446 “Measurement of emissions from leaks of gaseous compounds from leaks from equipment and pipes”. With the increased possibilities of applying a system of sensors at reasonable costs, especially in critical operating systems. Such measurements can guide operators in Safety Inspections under challenging situations (network malfunctions) by preventing risky situations and providing useful indications for faster remediation actions [6]. However, Digital Technologies implies an increase in CAPEX (Capital Expenditure) budget. Finally, the combination of Digital Technologies (VR and AR) favours the choice of industrial processes that are ever less polluting, intending to prevent and eliminate pollution from production cycles as part of the general framework of the circular economy driving toward innovation and the virtualization of activities. It will provide transparency across the supply chain, ensuring traceability, ethical sourcing, and more effective material flows.

4. Conclusion

The present paper investigates the framework of the risk management process in light of the innovation and the virtualization of industrial activities. The circular economy aims to accelerate the transition to circular business models and practices. The exploitation of Digital Technologies brings higher levels of transparency and provides valuable information on risk for critical infrastructures and industries. Sustainable and digital options can provide relevance under risky challenging situations. From a circular economy perspective, Digital Technologies improvements can also minimize the negative impacts caused by so-called NaTech events.

REFERENCES

1. Mari, M., Millucci, L., Fardelli, A. and Mazziotti Gomez de Teran, C. Circular bioeconomy growth to face the increasing industrial risk, *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*, (2020).

2. Necci, A. and Krausmann, E., (2022), Natech risk management. *EUR 31122 EN*, Publications Office of the European Union, Luxembourg.

3. Krausmann, E., Salzano, E., et al., (2017), Lessons learned from natech events. *Natech Risk Assessment and Management-Reducing the Risk of Natural-Hazard Impact on Hazardous Installations*, pp. 33–52, Elsevier Inc.

4. Gualtieri, L., Revolti, A. and Dallasega, P. A human-centered conceptual model for integrating augmented reality and dynamic digital models to reduce occupational risks in industrial contexts, *Procedia Computer Science*, **217**, 765–773, (2023).

5. Arana-Landín, G., Laskurain-Iturbe, I., Iturrate, M. and Landeta-Manzano, B. Assessing the influence of industry 4.0 technologies on occupational health and safety, *Heliyon*, **9** (3), (2023).

6. Ramos-Hurtado, J., Muñoz-La Rivera, F., Mora-Serrano, J., Deraemaeker, A. and Valero, I. Proposal for the deployment of an augmented reality tool for construction safety inspection, *Buildings*, **12** (4), 500, (2022).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

BUILDING AN ECOSYSTEM OF INTEROPERABLE DIGITAL TWINS: CHALLENGES & EXPERIENCES

Marco Picone, Nicola Bicocchi, *and* Marco Mamei
University of Modena & Reggio Emilia, Italy
email: {marco.picone, nicola.bicocchi, marco.mamei}@unimore.it

Mattia Fogli *and* Carlo Giannelli
University of Ferrara, Italy
email: {mattia.fogli, carlo.giannelli}@unife.it

Paolo Bellavista
University of Bologna, Italy
email: paolo.bellavista@unibo.it

Digital Twins (DTs) are software replicas of physical entities that hold great potential for cyber-physical applications. However, existing solutions often focused on centralized and Cloud-centric approaches without interoperable solutions and operating in isolated silos. This works aims to explore the challenges and opportunities in establishing an interoperable ecosystem of connected DTs by sharing research, development, and deployment experiences. The objective is to enable DTs to actively engage and collaborate across distributed environments through a flexible and dynamic deployment of pervasive DTs spanning from edge to cloud to meet the diverse applications' requirements. By embracing a shared vision and promoting interoperable methodologies and investigating standardization opportunities, we foster the creation of an open ecosystem that empowers the development of the next generation of intelligent cyber-physical applications through multiple application domains.

Keywords: digital twins, pervasive twins, software, edge-cloud continuum, standardization

1. Introduction

Digital Twins (DTs) are undergoing a transformation and emerging as a cross-domain paradigm for designing and implementing cyber-physical applications. They create synchronized software replicas associated with physical devices, products, and organizations, and are denoted as Physical Twins (PTs). The adoption of DTs extends beyond industrial and manufacturing contexts, bridging the physical and digital realms with the help of the Internet of Things (IoT) and Web technologies, enabling their applicability in diverse scenarios. The evolution of DTs brings new challenges that call for new software architectures, development, and deployment approaches. Existing solutions often treat DTs as passive components, lacking interoperable modeling and relying on platform-specific silos. This independent model-building approach without standardized methodologies poses the risk of vendor lock-in and hampers the creation of an open ecosystem for seamless collaboration among PTs, DTs, and applications. To unlock the full potential of DTs, a shared and interoperable vision is necessary, empowering DTs to function as active entities with their own behaviors. This vision encourages seamless collaboration with applications and services across distributed environments through an edge-cloud computational continuum, as introduced in [1, 2].

The primary objective of this contribution is to explore existing challenges, share research experiences, and identify innovation opportunities in creating an interoperable ecosystem of connected DTs. This interdisciplinary

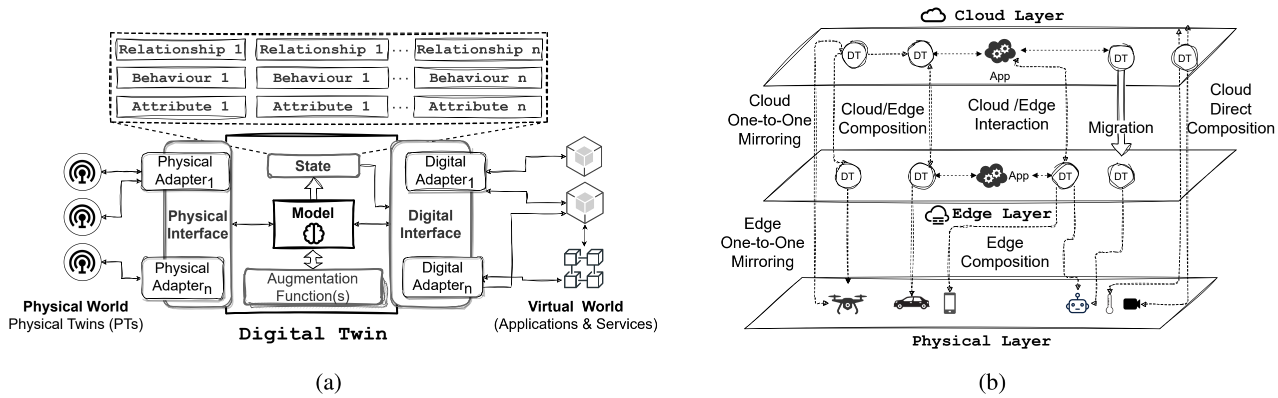


Figure 1: (a) A schematic representation of the Software blueprint architecture of a Digital Twin; (b) A hierarchical and multi-layer overview of multiple deployed twins across an edge-cloud continuum.

vision aims to facilitate the design, deployment, and maintenance of the next generation of intelligent cyber-physical applications. The focus is on pervasive and distributed DTs that can be flexibly and dynamically deployed across multiple network levels, ranging from edge to cloud, to meet specific needs and application requirements.

2. Distributed Digital Twins Modeling

In the scientific literature [3], DTs have been recently studied to propose a unified conceptual framework applicable across various domains. In [4], we relied on these abstract capabilities extending their characterization to the industrial ecosystem with a specific focus on Software requirements and how DT-driven application design, implementation, and deployment. A DT is represented schematically as shown in Figure 1(a) and modeled as an independent and autonomous architectural component. Its *Model* component captures attributes, behaviors, and relationships of the physical counterpart accurately. The DT's *State* component structures and stores computed attributes and relationships, facilitating interaction with the physical layer. A *Physical Interface* with multiple *Physical Adapters* enables communication with connected devices bridging the cyber-physical gap while on the other hand, DTs interact with the digital layer through a *Digital Interface* composed of *Digital Adapters* to create an interoperable and standardized interaction point in terms of protocols, and data formats to enable effective collaboration in the digital space with applications, services, and other DTs.

In the next generation of DT-based architectures, scalability is crucial to meet diverse requirements and support a wide range of application scenarios. Deploying DTs across Edge, Fog, and Cloud facilities offers flexibility as illustrated in Figure 1(b). Edge DTs (EDTs) (introduced in [2]) enable real-time data processing, immediate response, and localized insights. Fog-based DTs balance local processing and Cloud connectivity, enabling complex analytics and collaborative decision-making. Cloud-based DTs provide scalability, computational power, and centralized control for handling massive data volumes and resource-intensive computing. The scalability of DTs allows organizations to choose the deployment strategy based on specific requirements, considering factors such as data sensitivity and network connectivity. A multi-tiered edge-to-cloud architecture empowers DTs to address localized to global-scale use cases, fostering innovation and adaptability in the dynamic landscape of cyber-physical systems.

3. A Modular Approach for IoT Digital Twin Development

The absence of standards and common agreements for DT design and development has led to the proliferation of platform-specific solutions, causing fragmentation and hindering the potential of DTs [5]. Open-source initiatives like the Eclipse IoT Foundation and Eclipse Ditto [1] project have emerged to address this issue by providing open

[1]Eclipse Ditto - https://www.eclipse.org/ditto/

platforms for DT management. However, these frameworks have limitations in terms of flexibility and modularity, as they follow a monolithic and centralized vision where all DTs are co-located in the same architectural component. To overcome these limitations, we introduced in [6] the *White Label Digital Twin* (WLDT) an open-source ([2]) Java library designed for maximum modularity and flexibility. WLDT simplifies DT design and development by offering core features and functionalities. It employs a multi-threading core engine, enabling independent components to shape the behavior of each DT. Built-in IoT features and modules facilitate the digitalization of smart objects using MQTT and CoAP. The library's internal software processing pipeline allows dynamic customization of data management, tailored to specific use cases. An internal caching system enhances performance and reduces communication response time. WLDT follows the design principle of modeling each DT as an independent and autonomous software component, facilitating microservices-oriented IoT architectures. It serves as a foundational building block for DT-driven applications targeting the IoT ecosystems through built-in MQTT and CoAP workers, or the definition of new modules and connectors to support specific protocols or data flows.

4. Interoperability & Standardization Opportunities

Standardizing the definition, responsibilities and description of DTs is vital for successful software implementation, enabling seamless communication and data exchange. A standardized definition should encompass properties, relationships, models, operations, protocols, and events, transcending application domains. Standardization organizations and consortia, focusing on IoT, IIoT, networking, and edge/cloud architectures, are actively shaping interoperable environments [7]. Efforts from the Digital Twin Consortium, Digital Twin Programme, and companies like Microsoft with their Digital Twin Definition Language (DTDL)[3] contribute to the standardization landscape[4]. However, the definition of DTs is in its early stages, with differences and a tendency to treat DTs as passive components. The recent exploration of DTs as active interoperable entities within the IoT ecosystem is being investigated by organizations like the European Telecommunications Standards Institute (ETSI) with a Specialist Task Force (STF) [5] with the aim to analyze DTs standardization opportunities within the IoT application domain and specifically within the SmartM2M ecosystem. The objective of this work is to address the missing elements in modeling and standardizing the communication concept of IoT DTs and their blueprint communication reference architecture. The focus is on identifying use cases and deployments where they can be effectively adopted, in order to determine the requirements and specifications associated with their functionalities. The aim is to establish horizontal cross-domain interoperability and standards, defining minimum requirements for the usability of professional and general public IoT services, regardless of their criticality. The work will map IoT Digital Twins within the oneM2M framework and contribute to ISO/JTC1/SC41 by defining new specifications that embrace new functionalities and effectively utilize existing features such as discoverability, security, and modularity.

REFERENCES

1. Ricci, A., Croatti, A., Mariani, S., Montagna, S. and Picone, M. Web of digital twins, *ACM Trans. Internet Technol.*, **22** (4), (2022).

2. Picone, M., Mamei, M. and Zambonelli, F. A flexible and modular architecture for edge digital twin: Implementation and evaluation, *ACM Trans. Internet Things*, **4** (1), (2023).

3. Minerva, R., Lee, G. M. and Crespi, N. Digital twin in the iot context: a survey on technical features, scenarios, and architectural models, *Proceedings of the IEEE*, **108** (10), 1785–1824, (2020).

[2]WLDT on GitHub: `https://github.com/wldt`
[3]Microsoft Azure DTDL: `https://github.com/Azure/opendigitaltwins-dtdl`
[4]Digital Twin Consortium: `https://www.digitaltwinconsortium.org/index.htm`
[5]ETSI STF 628 Digital Twins: `https://portal.etsi.org/XTFs/#/xTF/628`

4. Bellavista, P., Bicocchi, N., Fogli, M., Giannelli, C., Mamei, M. and Picone, M. Requirements and design patterns for adaptive, autonomous, and context-aware digital twins in industry 4.0 digital factories, *Computers in Industry*, **149**, 103918, (2023).

5. Tao, F. and Qi, Q. Make more digital twins, *Nature*, **573** (7775), 490–491, (2019).

6. Picone, M., Mamei, M. and Zambonelli, F. Wldt: A general purpose library to build iot digital twins, *SoftwareX*, **13**, 100661, (2021).

7. Zhou, C., Yang, H., Duan, X., Lopez, D., Pastor, A., Wu, Q., Boucadair, M. and Jacquenet, C. Internet Engineering Task Force, Digital Twin Network: Concepts and Reference Architecture. Work in Progress, (2022).

RANDOMIZED METHODS FOR UNDERDETERMINED LINEAR INVERSE PROBLEMS

Lorenzo Della Cioppa
Istituto di Scienze Marine - CNR, Roma, Italy

Michela Tartaglione
Università di Roma "La Sapienza", Roma, Italy

Annalisa Pascarella
Istituto per le Applicazioni del Calcolo "Mauro Picone" - CNR, Roma, Italy

Francesca Pitolli
Università di Roma "La Sapienza", Roma, Italy
e-mail: francesca.pitolli@uniroma1.it

We present a method based on random column sampling for solving underdetermined linear inverse problems. Random sampling allows reducing the dimensionality of the inverse problem and acts as a preconditioner for classical regularization methods. We provide error estimates when the random sampling method is coupled with Tichonov regularization. The proposed method achieves high accuracy while keeping a low computational cost.

Keywords: *underdetermined inverse problem, random sampling, Tichonov regularization*

1. Introduction

Many real-world problems require solving an underdetermined inverse problem. For example, reconstructing brain activity maps from MEG/EEG data requires solving a severely underdetermined inverse problem that is ill-posed and ill-conditioned. In such cases, regularization techniques must be used to make the solution feasible. In recent years, randomized methods have proven to be very efficient in reducing the dimensionality of large linear algebra problems. Initially used to construct low-rank approximations of large matrices, randomized methods were later successfully applied to regression problems as well (see [1, 2] and references therein). We present a method based on random sampling to solve underdetermined linear inverse problems. We observe that random sampling allows reducing the dimensionality of the inverse problem, thus acting as a preconditioner of classical regularization methods.

2. The random sampling method

Consider the discrete ill-posed inverse problem

$$A\,x = b\,, \tag{1}$$

where $A \in \mathbf{R}^{m\times n}$, with $m \ll n$, is a ill-conditioned (possibly large) matrix, $x \in \mathbf{R}^n$ is the quantity to be reconstructed and $b \in \mathbf{R}^m$ are the measurements affected by noise. Let us introduce the sampling matrix $S \in \mathbf{R}^{n\times c}$, with $c \ll n$, which randomly samples c columns of A, i.e.,

$$A_S = A\,S \in \mathbf{R}^{m\times c}. \tag{2}$$

Assuming that the columns of A are normalized so that they have unitary ℓ_2-norm, we can sample the columns drawing from the uniform probability distribution. Thus, the random sampling method consists in solving the reduced problem

$$A_S\, x_S = b\,, \qquad x_S \in \mathbf{R}^c. \tag{3}$$

3. Randomized Tichonov regularization

Tichonov regularization is a classical regularization method widely used in several fields [3]. It consists in solving the minimization problem

$$\min(\|A\,x - b\|^2 + \alpha\|x\|^2), \tag{4}$$

where $\alpha > 0$ is the regularization parameter. Let $A = U\Sigma V$ be the singular value decomposition (SVD) of A. It is well-known that this problem has a unique minimizer that can be written as

$$x^\alpha = V(\Sigma\Sigma^T + \alpha I)^{-1}\Sigma^T U b. \tag{5}$$

In the randomized version of the Tichonov regularization we used the sampled matrix A_S instead of A, i.e.,

$$x_S^\alpha = V_S(\Sigma_S\Sigma_S^T + \alpha I_S)^{-1}\Sigma_S^T U_S b, \tag{6}$$

where $A_S = U_S\Sigma_S V_S$ is the SVD of A_S. In case of uniform sampling with replacement it can be shown [**?**] that for the relative error $\mathcal{E} = \|x - S\,x_S\|/\|x\|$ the following estimate holds

$$\mathcal{E} \leq \frac{1}{\alpha}\max(\frac{n}{c}\nu - 1, 1)\sqrt{\nu}\,\kappa(A^T A + \alpha I)\sigma_{max}(A^T A + \alpha I), \tag{7}$$

where ν denotes the maximum number of times a column is drawn. $\kappa(A)$ and $\sigma_{max}(A)$ denote the conditioning and the maximum singular value of a matrix A, respectively. The estimate shows that when $c \ll n$ the error decreases fast as c increases reaching a minimum approximately when $\frac{n}{c}\nu = 2$. Then, the error increases as ν.

4. Conclusion

We applied the random sampling method to solve the MEG/EEG inverse problem [4, 5, 6]. The method can be easily applied to different devices and can be integrated with other methodologies. Numerical experiments show that the method achieves high accuracy while keeping a low computational cost, making it suitable for real-time applications with portable devices, such as brain-computer interface training or neurofeedback rehabilitation.

REFERENCES

1. Halko, N., Martinsson, P.-G. and Tropp, J. A. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, *SIAM review*, **53** (2), 217–288, (2011).

2. Drineas, P. and Mahoney, M. W. Lectures on randomized numerical linear algebra, *The Mathematics of Data*, **25** (1), (2018).

3. Engl, H. W., Hanke, M. and Neubauer, A., *Regularization of inverse problems*, vol. 375, Springer Science & Business Media (1996).

4. Della Cioppa, L., Tartaglione, M., Pascarella, A. and Pitolli, F. Solution of the eeg inverse problem by random dipole sampling, *submitted*, (2023).

5. Pascarella, A. and Pitolli, F. An inversion method based on random sampling for real-time meg neuroimaging, *Communications in Applied and Industrial Mathematics*, **10** (2), 25–34, (2018).

6. Campi, C., Pascarella, A. and Pitolli, F. Less is enough: Assessment of the random sampling method for the analysis of magnetoencephalography (meg) data, *Mathematical and Computational Applications*, **24** (4), 98, (2019).

CONVOLUTIONAL NEURAL NETWORKS FOR THE AUTOMATIC CONTROL OF CONSUMABLES FOR ANALYTICAL LABORATORIES

Meriam Zribi
Università degli Studi Roma La Sapienza, Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Rome, Italy
e-mail: meriam.zribi@uniroma1.it

Paolo Pagliuca
Institute of Cognitive Sciences and Technologies National Research Council (CNR), Rome, Italy

Francesca Pitolli
Università degli Studi Roma La Sapienza, Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Rome, Italy

In recent years, the need for advanced systems and technologies for industrial process optimization using computer vision and artificial intelligence (AI) techniques has become increasingly pervasive. The specific focus of this study is to introduce an AI-based monitoring system within a production chain involved in manufacturing plastic consumables for analytical laboratories, specifically targeting the control of vials containing an anticoagulant substance. Currently, the inspection process relies on manual visual inspection conducted on a sample basis, resulting in the potential discarding of entire production batches if the absence of the anticoagulant substance is detected in a single vial. To overcome the inefficiency of the manual system, a comprehensive method is proposed to verify the presence of the anticoagulant substance in all produced vials, leveraging advanced computer vision and AI techniques. This innovative monitoring system offers promising solutions for enhancing industrial processes by enabling accurate and real-time monitoring. Specifically, we present our model and some preliminary results showing the potentiality of the proposed approach.

Keywords: *automatic monitoring, green economy, deep learning, convolutional neural networks*

1. Introduction

In recent years, the application of computer vision and artificial intelligence (AI) techniques in the industrial domain has shown promising results. These methodologies enable the analysis of images captured during the production process and the extraction of valuable information for monitoring and control purposes. By utilizing deep learning algorithms such as convolutional neural networks (CNN), it becomes possible to identify patterns, detect defects or anomalies, and provide instant feedback on the process's performance. The scientific literature highlighted several successful cases of applying computer vision and AI-based monitoring systems [1, 2, 3, 4].

In this work, we focus on the development of a computer vision and AI-based monitoring system to replace the manual visual inspection of a specific stage in a production chain. The goal is to leverage the potential of computer vision techniques so as to identify process irregularities in real-time. Specifically, we design a deep network model able to detect the presence of an anticoagulant substance inside transparent tubes. We use real images acquired through a camera to train our model for the ability to distinguish between presence and absence of the reagent.

This approach aims to optimize resource utilization, increase operational efficiency, and reduce waste in industrial processes, in order to: (i) align with the principles of sustainable manufacturing and (ii) contribute to the achievement of environmental and economic goals. Moreover, it offers several advantages, including the ability to monitor processes without the need for expensive dedicated sensors and the capability of identifying hidden problems that may escape other monitoring methods. Additionally, the use of images provides an intuitive visualization of the process, facilitating the understanding and enabling prompt interventions when necessary.

2. Method and results

As we stated in Section 1, this work addresses a specific industrial application, i.e. the detection of the presence of an anticoagulant substance inside vials within a production chain involved in manufacturing plastic consumables for analytical laboratories. To this end, we used a Deep Network architecture constituted by two main blocks:

- a 3-layer CNN neural network extracting relevant features;
- a 4-layer fully-connected network that performs the classification.

The model parameters have been chosen from scratch through an empirical process. The values of parameters of our deep network model are provided in Table 1.

Table 1: Deep network parameters. The symbol "-" for the fully-connected layers (ids 4-7) indicates that the corresponding parameter is meaningless.

Layer id	# of Input channels	# of Output channels	Kernel size	Stride	Input size	Output size
1	3	10	3	1	400x400x3	398x398x10
2	10	20	3	1	398x398x10	396x396x20
3	20	30	3	1	396x396x20	394x394x30
4	-	-	-	-	394x394x30	50
5	-	-	-	-	50	15
6	-	-	-	-	15	10
7	-	-	-	-	10	2

Our model has been trained by using images of the vials acquired through a camera situated on the top of the pipeline. We collected images of resolution 400×400 pixels. An example of the images is shown in Figure 1:

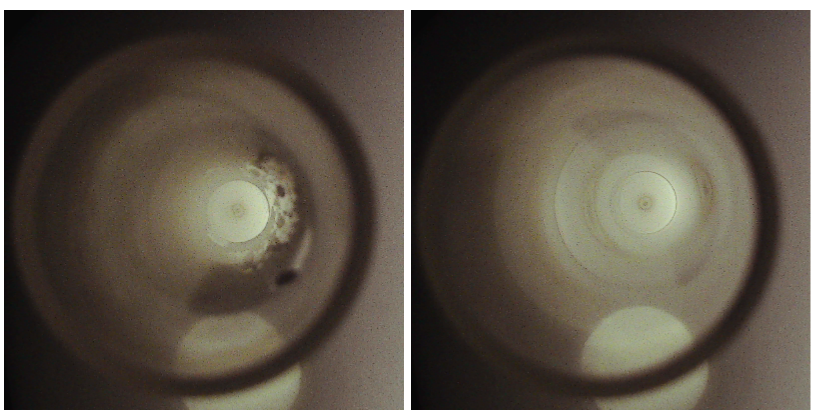


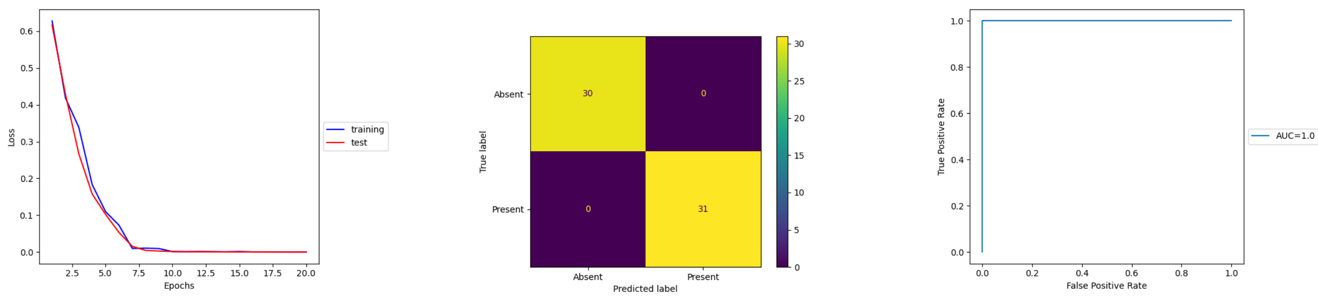
Figure 1: Two examples of images recorded by our system. **Left**: a tube containing anticoagulant. **Right**: an empty tube.

Specifically, we acquired 402 images split into a training set, which contains 341 images, and a test set including 61 images. In both sets, half of the images refer to tubes containing the anticoagulant substance, while the other half concerns empty tubes. The CNN block takes images as inputs and extracts features as outputs, which in turn will be used as inputs of the classification block. The first layer has a number of input channels corresponding to the basic colors (i.e., red, green and blue). For the other layers, the number of input channels is provided in Eq. 1:

$$num_{in\_channels}(l) = num_{out\_channels}(l-1), \quad l > 1 \tag{1}$$

where l indicates the layer id (see Table 1, first column).

The experiment has been replicated 10 times. Training lasted 20 epochs. We used the Adam optimizer [5] with weight decay. The learning rate has been set to 10^{-4} and the batch size to 16. With these settings, we achieved an average accuracy score of 100% over 10 replications of the experiment. This implies that all the trained models are able to correctly detect the presence/absence of the anticoagulant substance in the tube. The training and test errors are shown in Figure 2, left. By analyzing the detection ability of the best model, we can see that the confusion matrix (Figure 2, center) has no values outside the diagonal, i.e. no classification errors are performed. Furthermore, the ROC (Receiver Operating Characteristics) curve (Figure 2, right) corresponds to the ideal situation in which the classifier is able to distinguish between the positive class (presence of the anticoagulant) and the negative class (absence of the anticoagulant). Finally, the AUC (Area Under the Curve) score is equal to 1.0, thus indicating a perfect classifier.



(a) (b) (c)

Figure 2: (a) - Error curves during training. The blue curve represents the error on the training set, while the red curve indicates the error on the test set. We used the cross-entropy loss as an error metric. Data are obtained by averaging 10 replications of the experiment. (b) - Confusion matrix for classes "Present" and "Absent". The matrix represents the model capability of classifying images in the test set. (c) - ROC curve concerning the test set. Legend provides the AUC score. With respect to plots (b) and (c), data refer to the best model.

3. Conclusions

In this work, we describe an automated system able to correctly detect the presence/absence of an anticoagulant substance in vials. The model has been trained on a small dataset collected in a company dealing with plastic consumables. Preliminary results show that the approach is promising, as the system successfully classifies all images in the dataset. Nonetheless, real industrial applications deal with large amount of data. Future work should be devoted to validate this approach on a wider dataset. In addition, future research directions may focus on refining and optimizing the proposed computer vision and AI-based monitoring system, exploring its applicability in different industrial sectors, and investigating potential integration with other emerging technologies such as Internet of Things (IoT) and cloud computing for enhanced data analysis and decision-making processes.

REFERENCES

1. Agarwal, P., Aghaee, M., Tamer, M. and Budman, H. A novel unsupervised approach for batch process monitoring using deep learning, *Computers & Chemical Engineering*, **159**, 107694, (2022).

2. Lyu, Y., Chen, J. and Song, Z. Image-based process monitoring using deep learning framework, *Chemometrics and Intelligent Laboratory Systems*, **189**, 8–17, (2019).

3. Wu, H. and Zhao, J. Self-adaptive deep learning for multimode process monitoring, *Computers & chemical engineering*, **141**, 107024, (2020).

4. Yuan, J. and Tian, Y. A multiscale feature learning scheme based on deep learning for industrial process monitoring and fault diagnosis, *IEEE Access*, **7**, 151189–151202, (2019).

5. Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, (2014).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

CONVERGENCE BOUNDS FOR NODE SELECTION IN FEDERATED LEARNING

Francesco Malandrino
CNR-IEIIT, CNIT
email: francesco.malandrino@cnr.it

Carla Fabiana Chiasserini
Politecnico di Torino, Italy
CNR-IEIIT, CNIT

A priori, closed-form loss bounds have emerged as an important tool to estimate the performance of distributed learning and identifying the best learning nodes therein. Through real-world experiments, in this work we demonstrate how loss bounds are far away from the actual loss values; none the less, they can be useful for node selection purposes.

Keywords: *federated learning, loss bounds*

Federated learning (FL) [1], along with its derivative versions, is arguably the most promising distributed learning paradigm. Our work targets two of the main problems in FL, to wit:

- estimating the learning performance *before* running the learning process itself, and
- identifying the nodes that are likely to contribute the most to the learning.

We do so by performing a set of experiments leveraging the CIFAR dataset [2] and the LENET DNN [3]. We consider a total of 10 learning nodes, differing for the number of images in their local datasets (i.e., the *quantity* of data) and the classes represented therein (i.e., the *quality* of data). For the bounds, we leverage the work [4], where the expected loss at iteration t is bounded by

$$\frac{8L/\mu}{(t-1+8L/\mu)}\left(\frac{16G^2}{\mu}+4L\mathbb{E}\|\mathbf{w}_1-\mathbf{w}^*\|\right), \tag{1}$$

where

- μ is a non-negative quantity such that loss function F is μ-strongly convex;
- L is a non-negative quantity such that loss function F is L-smooth;
- G is a non-negative quantity such that the squared norm of the gradients of loss function F is bounded by G^2.

To begin with, we consider the actual and predicted loss bounds, represented in Fig. 1. Looking at the scales in the plot, it is immediately evident that the bound (1) is quite far away from the actual loss; therefore, it is of limited usefulness in directly predicting the evolution of the learning process.

We now turn our attention to the local values of the quantities appearing in (1), specifically, L, and seek to correlate them with the usefulness metric introduced in [5]. The results are summarized in Fig. 2.

It is extremely interesting to observe how, barring a single outlier, nodes with a high local value of L also have a good usefulness. These results suggest that, while bounds such as (1) may not be the best tool to quantitatively predict the evolution of learning, the quantities appearing therein – which themselves depend upon the local datasets – are an excellent tool to identify the most promising learning nodes.

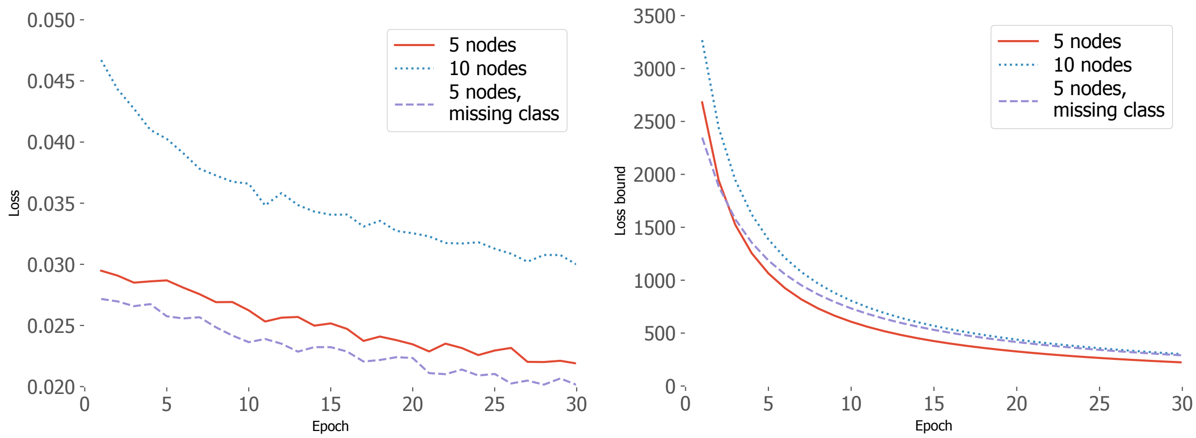


Figure 1: FL experiments: loss achieved during the training (left); bounds thereto (right).

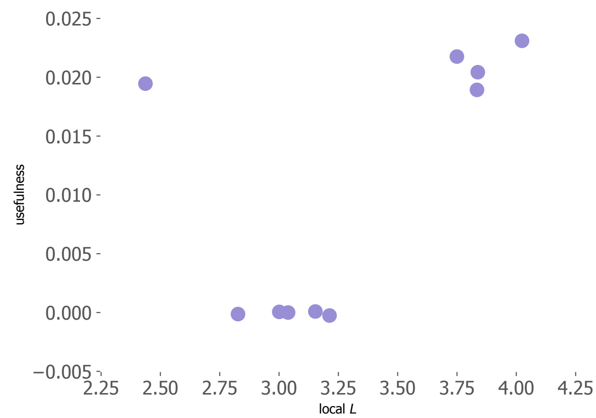


Figure 2: FL experiments: relationship between the node usefulness and the local values for the L.

REFERENCES

1. Konečný, J., McMahan, B. and Ramage, D. Federated optimization: Distributed optimization beyond the datacenter, *arXiv preprint arXiv:1511.03575*, (2015).

2. Krizhevsky, A., Hinton, G., et al. Learning multiple layers of features from tiny images, (2009).

3. Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, (1998).

4. Li, X., Huang, K., Yang, W., Wang, S. and Zhang, Z. On the convergence of fedavg on non-iid data, *International Conference on Learning Representations*, (2019).

5. Malandrino, F. and Chiasserini, C. F. Federated learning at the network edge: When not all nodes are created equal, *IEEE Communications Magazine*, (2021).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

ON SOME OPTIMIZATION PROBLEMS IN POWER ENERGY

Claudio Gentile, Tiziano Bacci, *and* Sara Mattia
IASI-CNR, Roma, Italy
e-mail: gentile@iasi.cnr.it

Antonio Frangioni
Dipartimento di Informatica, Università di Pisa, Pisa, Italy

Fabrizio Lacalandra
Autorità di Regolazione per Energia Reti e Ambiente (ARERA), Milano, Italy

Martina Gherardi
Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università di Bergamo, Bergamo, Italy

In the last decades the power energy systems underwent a complete organizational change. In many countries they were transformed from a monopolistic to a free market regime. This change introduced new actors each one with its particular view to be optimized. Moreover, the introduction of a relevant share of renewable sources (in particular, wind and photovoltaic) augmented the uncertainty in the energy production due to weather conditions. Therefore, the new setting increased the needs for optimization and required new levels of performances. We present some new advances in solving optimization problems in the power energy sector mainly considering exact approaches.

Keywords: *power energy production, power energy distribution, optimization, exact methods*

1. Introduction

In the last decade of the last century most countries of North America and Europe decided to change the power system management and the energy sector in general passing in each contry/region from monopolitic to free market regimes. Rules and governance may change from country to country but many elements are often common:

- a Day-Ahead Market (DAM) that handles the power energy offer and demand bids and establishes the *price of energy* according to the marginal price to maximize the general welfare of the system;
- Intra-day Markets that enable to modify the schedules defined by the DAM by new offer and demand bids taking into account a more exact behavior of production and consumption rates while approaching real time;
- bilateral agreements between power energy operators and large power energy consumers, such as manufacturing industries;
- Dispaching Services Market that deals with offers/demands of energy needed to keep balanced the electrical system, and correct voltage profiles thus allowing actual transmission from producers to consumers.

The energy markets are usually governed by two entities, one responsible for electricity market clearing/pricing, the Nominated Energy Market Operator (NEMO) and one responsible for the dispatching and balancing markets that take into account the transmission network, the Transmission System Operator (TSO).

2. The main optimization problems

The energy system nowadays sees the participation of many different actors: producers, traders, large consumers, small and residential consumers, small producers (also defined *prosumers*), energy communities [1], independent market operators, and others. Each actor manages its operations trying to maximize its own benefit;

therefore many optimization problems arise. However two issues, that were already present in the old monopolistic regime, are of fundamental importance and constraint many of the following operations: (i) where and when energy is produced, (ii) how energy is routed into the transmission grid. The first issue claims for a class of problems named "Unit Commitment problems" where one has to manage a set of production units and decide when and how much power each unit should produce. The second issue refers to the class of problems named "Optimal Power Flow" where one has to manage how the current flows into the grid to provide energy to the final consumers. Both classes of problems are of operational type, that is, they consider a short time horizon (from one day to a week).

3. The Unit Commitment Problem

The *Unit Commitment* (UC) problem considers a time horizon, a set of production units of different types, forecasted demands at each time period of the time horizon and its objective is to decide which units should produce and at which time periods in order to meet the demands at minimum production cost. Each production unit (thermal, nuclear, hydroelectric, solar, wind, and others) must consider its specific technical constraints. After a lot of simplifications, production units are usually partitioned into programmable units (mainly thermal, including nuclear ones, and hydroelectric units) and non-programmable units (solar, wind units). Programmable units can be managed, up to technical constraints, to enable the increase or decrease of production, thus changing the cost of the production. The technical working constraints need to be mathematically modeled with very interesting issues from an optimization and combinatorial point of view [2]. On the contrary, the energy produced by non-programmable units cannot be modified as one likes and largely depends on external and non-controllable factors (weather conditions in general). Consequently uncertainty issues arise that must be handled by correct optimization models (stochastic optimization, robust optimization) thus increasing the need for more efficient optimization algorithms [3].

4. The Optimal Power Flow Problem

The transmission of the power energy is another challenging optimization problem. The transmission grid is one of the main infrastructure of a country. From an optimization point of view, power must satisfy Ohm and Kirchoff's Laws, i.e., it cannot be routed in the network as water or other commodities. These laws define the basic transmission problem named *Optimal Power Flow* (OPF). Energy may be distributed according two different technologies: (i) direct current, (ii) alternate current. Alternate current is more suitable for transmission at long distances. According to optimization, direct current may be effectively modeled with linear programming models. On the contrary, alternate current must be modeled with high nonlinear models in complex variables. A recent survey on alternate current OPF models is given in [4]. With both transmission technologies, physical laws may induce phenomena such as the Braess paradox: increasing the number or capacity of transmission branches may not improve the transmission. Therefore one has to decide which transmission branches should be activated/deactivated (*Optimal Transmission Switching* problem [5]).

5. The Energy markets problems

The energy markets introduce two main types of problems: (i) market clearing problems to be solved by NEMOs, (ii) strategic bidding problems to be solved by production companies, traders and others involved actors. The former are used to find the offers and demands that maximize the social welfare according to the rules defined by market regulators; While in their simplest forms these market clearing problems are simple LPs (if the network constraints are suitably simplified), regulatory rules often make them more complex, e.g., due to the definition of zonal markets (i.e., aggregations of portions of the grid) and of possible "nonlinear" interactions among zones such as the unique purchase price in the Italian market, or "complex offers" in other European ones. Thus such market clearing problems often become more complex including integer variables. A recent need, due to the natural gas prices crisis, is related to the issue of (partly) decoupling the price of energy rewarded to producers with significant fuel costs (thereby subject to the large fluctuations in the hydrocarbon markets) from those relying on sources (mainly, but not exclusively, renewables) that do not suffer from the issue. A proposal in this sense [6] turns the market clearing problem into a bilevel / MPEC one, which is solvable with off-the-shelf tools for small-size instances but that may require specialised algorithmic developments for real-world deployments.

The strategic bidding problems are used by each single operator (produces, traders, large consumers) to define the bids to be presented at the different markets in order to optimize its own results (optimize profits or costs). This class of problems may be defined as bilevel programming problems [7].

6. The Generation and Transmission Expansion Planning Problem

Well known growing policy targets is prompting the development of effective strategies for a decarbonized energy system. Increasing the share of renewable generation requires also increasing the system flexibility (electrical storage of various kind) and integrating the electricity and gas systems (electricity-to-gas conversion by PtX and gas-to-electricity conversion by thermal power plants).

Generation and Transmission Expansion Planning (GTEP) problem determines the evolution of the eletric energy system over a long-term horizon, taking into account policy targets. While *decentralized models* consider multiple decision-makers with different objectives, in *centralized models* a single decision-maker (e.g., the authority or a suitable Ministry) determines the expansion plan that minimizes investment and operating costs: on the basis of this solution, policies and incentives can be identified to lead generation companies to invest in a socially efficient manner (*anticipative planning*). The variability of non-programmable generation can lead to higher operating costs for thermal units, which can be accurately estimated by considering a high level of temporal and technical detail. Long-term uncertainties, such as fuel costs and CO_2 costs, can also be considered. Recent GTEP approaches aim to determine the expansion plan in which generation units are guaranteed to recover all their costs, e.g. [8].

ACKNOWLEDGMENTS

This paper has been partially funded by the project PRIN2022BMBW2A "Large-scale optimization for sustainable and resilient energy systems".

REFERENCES

1. Fioriti, D., Frangioni, A. and Poli, D. Optimal sizing of energy communities with fair revenue sharing and exit clauses: Value, role and business model of aggregators and users, *Applied Energy*, **299**, 117328, (2021).

2. Bacci, T., Frangioni, A., Gentile, C. and Tavlaridis-Gyparakis, K. New minlp formulations for the unit commitment problem with ramping constraints, *Operations Research*, **to appear**, (2023).

3. van Ackooij, W., Lopez, I. D., Frangioni, A., Lacalandra, F. and Tahanan, M. Large-scale unit commitment under uncertainty: an updated literature survey, *Annals of Operations Research*, **271** (1), 11–85, (2018).

4. Bienstock, D., Escobar, M., Gentile, C. and Liberti, L. Mathematical programming formulations for the alternating current optimal power flow problem, *Annals of Operations Research*, **314**, 227–315, (2022).

5. Kocuk, B., Jeon, H., Dey, S. S., Linderoth, J., Luedtke, J. and Sun, X. A. A cycle-based formulation and valid inequalities for dc power transmission problems with switching, *Operations Research*, **64** (4), 922–938, (2016).

6. Frangioni, A. and Lacalandra, F. Dipartimento di Informatica, Università di Pisa, A bilevel programming approach to price decoupling in pay-as-clear markets, with application to day-ahead electricity market, (2022).

7. Kalashnikov, V., Dempe, S., Pérez-Valdés, G. A., Kalashnykova, N. and Camacho-Vallejo, J.-F. Bilevel programming and applications, *Mathematical Problems in Engineering*, **2015**, Article ID 310301, 16 pages, (2015).

8. Guo, C., Bodur, M. and Papageorgiou, D. J. Generation expansion planning with revenue adequacy constraints, *Computers & Operations Research*, **142**, 105736, (2022).

BUILD-IT2023 Workshop
19-20 October, Rome, Italy

AUTOMATIC DISCOVERY OF LOW-DIMENSIONAL DYNAMICS UNDERPINNING TIME-DEPENDENT PDES BY MEANS OF LATENT DYNAMICS NETWORKS

Francesco Regazzoni, Stefano Pagani *and* Luca Dede'
MOX, Department of Mathematics, Politecnico di Milano, Milan, Italy
e-mail: francesco.regazzoni@polimi.it

Matteo Salvador
Institute for Computational and Mathematical Engineering, Stanford University, California, USA

Alfio Quarteroni
MOX, Department of Mathematics, Politecnico di Milano, Milan, Italy
École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland (Professor Emeritus)

We present a novel Machine Learning technique able to learn differential equations that surrogate the solution of space-time-dependent problems. Our method exploits a finite number of latent variables, providing a compact representation of the system state, automatically discovered during training. It allows building, in a fully non-intrusive manner, surrogate models accounting for the dependence on parameters and time-dependent inputs. This work pushes forward a novel technology towards the construction of data-driven digital twins in various application fields.

Keywords: *data-driven modeling, scientific machine learning, surrogate modeling, latent dynamics networks*

1. Introduction

Predicting the evolution of systems that exhibit spatio-temporal dynamics in response to external stimuli is a key enabling technology fostering scientific innovation. Traditional equations-based approaches leverage first principles to yield predictions through the numerical approximation of high-dimensional systems of differential equations, thus calling for large-scale parallel computing platforms and requiring large computational costs. Data-driven approaches, instead, enable the description of systems evolution in low-dimensional latent spaces, by leveraging dimensionality reduction and deep learning algorithms [1, 2, 3, 4, 5, 6]. Through the data-driven modeling approach, our work provides a solid methodological foundation for creating digital twins, enabling predictive simulation in a fast and accurate manner [7].

2. Methods

We propose a novel architecture, named Latent Dynamics Network (LDNet), which is able to discover low-dimensional intrinsic dynamics of possibly non-Markovian dynamical systems, thus predicting the time evolution of space-dependent fields in response to external inputs [8]. Unlike popular approaches, in which the latent representation of the solution manifold is learned by means of auto-encoders that map a high-dimensional discretization of the system state into itself, LDNets automatically discover a low-dimensional manifold while learning the latent dynamics, without ever operating in the high-dimensional space. Furthermore, LDNets are meshless algorithms that do not reconstruct the output on a predetermined grid of points, but rather at any point of the domain, thus enabling weight-sharing across query-points. These features make LDNets lightweight and easy-to-train, with excellent accuracy and generalization properties, even in time-extrapolation regimes.

3. Results

We demonstrate the effectiveness of LDNets through several test cases. First, we consider a linear PDE model to analyze the ability of LDNets to extract a compact latent representation of models that are progressively less amenable to reduction. Then, we consider the time-dependent version of a benchmark problem in fluid dynamics. Finally, we compare LDNets with state-of-the-art methods in a challenging task, that is, learning the dynamics of the Aliev-Panfilov model [9], a highly non-linear excitation-propagation PDE model used in the field of cardiac electrophysiology modeling. In the latter test case (see Fig. 1) we show that LDNets outperform state-of-the-art methods in terms of accuracy (normalized error 5 times smaller), by employing a dramatically smaller number of trainable parameters (more than 10 times fewer).

We focus on synthetically generated data obtained by numerical approximation of differential models, thus allowing us to test LDNet predictions against ground-truth results. We evaluate the prediction accuracy of the trained models using two metrics: the normalized root-mean-square error (NRMSE) and the Pearson dissimilarity, $1 - \rho$, where ρ is the Pearson correlation coefficient. To tune hyperparameters, we employ a Bayesian approach, namely the Tree-structured Parzen Estimator algorithm [10], combined with Asynchronous Successive Halving scheduler to early terminate bad hyperparameters configurations [11].

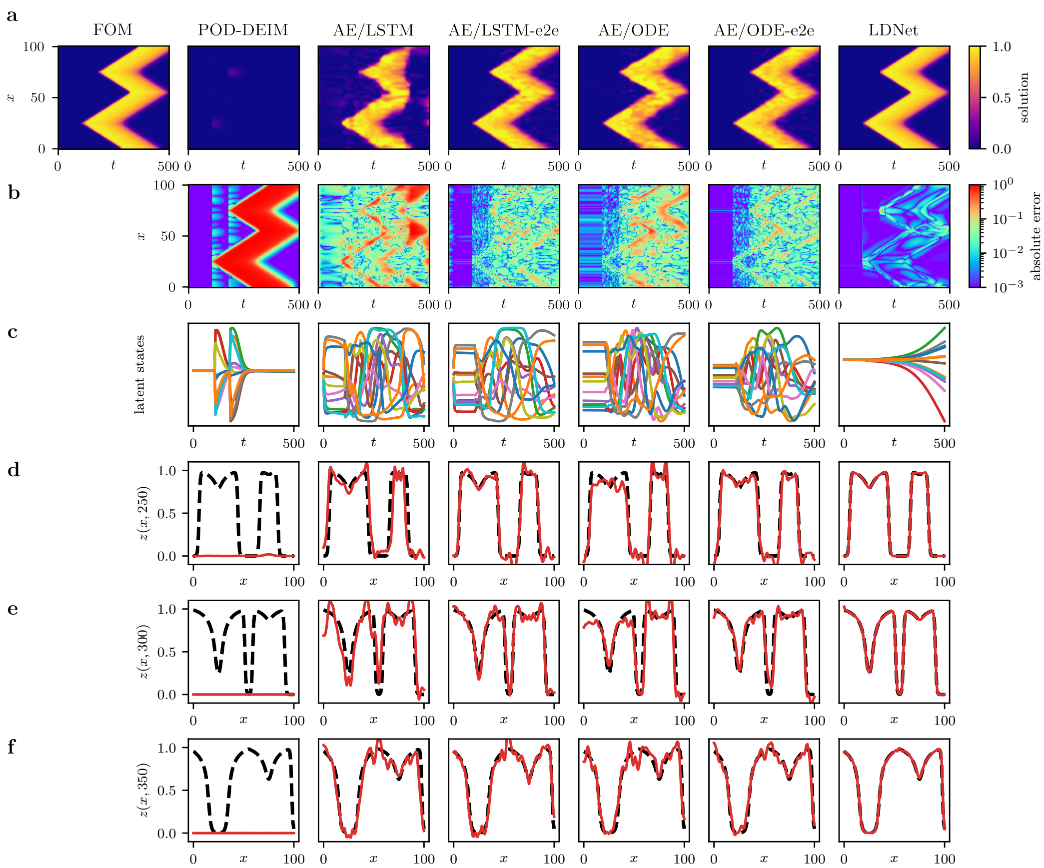


Figure 1: Results of the Aliev-Panfilov test case. We compare the results obtained different methods (reported in the captions) for a sample belonging to the test dataset. The left-most column reports the FOM solution of the AP model (the abscissa denotes time, the ordinate denotes space). For each method we report: **(a)** the space-time solution; **(b)** the space-time error with respect to the FOM solution; **(c)** the time-evolution of the 12 latent variables; **(d)**-**(e)**-**(f)** three snapshots of the space-dependent output field at $t = 250$, 300 and 350, in which we compare the predicted solution (red solid line) with the FOM solution (black dashed line).

4. Conclusions

LDNets represent, as proved by the results of this work, an innovative tool capable of learning spatio-temporal dynamics with great accuracy and by using a remarkably small number of trainable parameters. They are able to discover, simultaneously with the system dynamics, compact representations of the system state, as shown in Test Case 1 where the Fourier transform of a sinusoidal signal is automatically discovered. Once trained, LDNets provide predictions for unseen inputs with negligible computational effort (order of milliseconds for the considered Test Cases). LDNets provide a new flexible and powerful tool for data-driven digital twins that is open to a wide range of variations in the definition of the loss function (like, e.g., including physics-informed terms), in the training strategies, and, finally, in the NN architectures. The comparison with state-of-the-art methods on a challenging problem, such as predicting the excitation-propagation pattern of a biological tissue in response to external stimuli, highlights the full potential of LDNets, which outperform the accuracy of existing methods while still using a significantly lighter architecture.

REFERENCES

1. Bongard, J. and Lipson, H. Automated reverse engineering of nonlinear dynamical systems, *Proceedings of the National Academy of Sciences*, **104** (24), 9943–9948, (2007).
2. Schmidt, M. and Lipson, H. Distilling free-form natural laws from experimental data, *science*, **324** (5923), 81–85, (2009).
3. Peherstorfer, B., Gugercin, S. and Willcox, K. Data-driven reduced model construction with time-domain loewner models, *SIAM Journal on Scientific Computing*, **39** (5), A2152–A2178, (2017).
4. Rudy, S. H., Brunton, S. L., Proctor, J. L. and Kutz, J. N. Data-driven discovery of partial differential equations, *Science advances*, **3** (4), e1602614, (2017).
5. Bar-Sinai, Y., Hoyer, S., Hickey, J. and Brenner, M. P. Learning data-driven discretizations for partial differential equations, *Proceedings of the National Academy of Sciences*, **116** (31), 15344–15349, (2019).
6. Cenedese, M., Axås, J., Bäuerlein, B., Avila, K. and Haller, G. Data-driven modeling and prediction of non-linearizable dynamics via spectral submanifolds, *Nature communications*, **13** (1), 1–13, (2022).
7. Regazzoni, F., Salvador, M., Dedè, L. and Quarteroni, A. A machine learning method for real-time numerical simulations of cardiac electromechanics, *Computer Methods in Applied Mechanics and Engineering*, **393**, 114825, (2022).
8. Regazzoni, F., Pagani, S., Salvador, M., Dede, L. and Quarteroni, A. Latent dynamics networks (ldnets): learning the intrinsic dynamics of spatio-temporal processes, *arXiv preprint arXiv:2305.00094*, (2023).
9. Aliev, R. R. and Panfilov, A. V. A simple two-variable model of cardiac excitation, *Chaos, Solitons & Fractals*, **7** (3), 293–301, (1996).
10. Bergstra, J., Bardenet, R., Bengio, Y. and Kégl, B. Algorithms for hyper-parameter optimization, *Advances in neural information processing systems*, **24**, (2011).
11. Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Ben-Tzur, J., Hardt, M., Recht, B. and Talwalkar, A. A system for massively parallel hyperparameter tuning, *arXiv preprint arXiv:1810.05934*, (2020).

ARE DIGITAL TWINS SUITABLE TO DRIVE SAFE AI?

Alberto Carlevaro, Giacomo De Bernardi, Marta Lenatti, Sara Narteni, Marco Muselli, Alessia Paglialonga, Fabrizio Dabbene, *and* Maurizio Mongelli
CNR-IEIIT
email: maurizio.mongelli@cnr.it

The document deals with V&V of AI for autonomous vehicles that need to move and perform tasks in crowded environments. The adoption and extension of digital twin technology is discussed to pave the way towards AI certification.

Keywords: *smart mobility, autonomy, V&V, safe AI, SOTIF*

1. Introduction

The document is based upon and motivated by an ongoing H-EU project (REXASI-PRO, "REliable & eXplAinable Swarm Intelligence for People with Reduced mObility"[1]), in which AI is certified for autonomous wheelchairs for elderly and fragile people in indoor environments, such as stations (railway, airport) or museums (Fig. 1). Inherent applications involve complex interactions between humans and robots, such as: robots carrying material or autonomous stretchers in hospitals, material handling tasks in logistics, smart manufacturing, robot-aided physiotherapy or post trauma recovery. The rationale is to study how to develop a digital twin to expand the capacity of trials of the ecosystems like REXASI-PRO. According to traditional VV, the aim is to enumerate all safety issues, hazards and countermeasures, paving the ground for broader verification and certification applications (e.g., preventing collision of autonomous wheelchair through additional cameras and drones). The problem is however even more subtle; the certification needs the re-design of traditional VV when AI is in the loop.

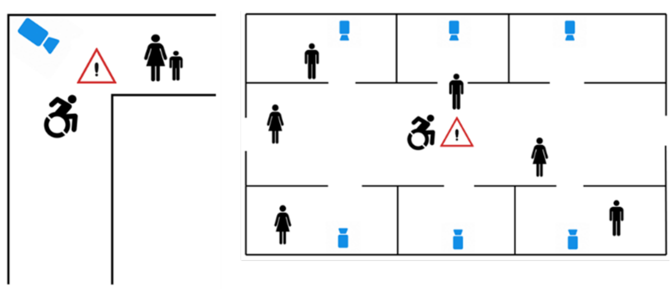


Figure 1: REXASI-PRO scope: autonomous wheelchair should move in the crowd fast and safely.

2. Why safe AI

The recent success of AI over traditional model-based methods is mainly due to its flexibility ("train, plug and play"). The wheelchair moves according to a neural control, trained through a collection of field data (Fig. 2). The database for training is based upon several registrations of passages of the wheelchair, guided by an operator, within a sufficient series of crowd states. The operator registers the most relevant passages with respect to safe and comfortable distance maintenance, as well as acceptable speed to reach the desired destinations.

[1] https://rexasi-pro.spindoxlabs.com/

The power of the approach is evident: just data recording and mapping the inherent movements into a neural network. The mapping between comfort, speed and collision avoidance is still something to be discovered from the trained neural network. More than this, the presumed “sufficiency” of crowd states highlights the problem of uncertainty. Is the trained neural network compatible with any other possible crowd scenario and corresponding field data acquisition? Under which operating conditions could it give rise to unexpected control spikes that make the actuation dangerous (close to a collision)? This leads to the need for a digital twin that incorporates the actuation engine and simulation of field data. AI itself then becomes an integrated tool in the digital twin as explained later on.

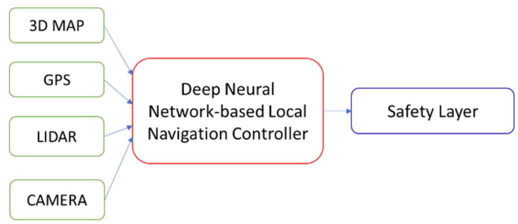


Figure 2: REXASI-PRO basic functional chain: sensing, neural controller, safety layer.

3. SOTIF

Although ISO 26262:2018 series of standards remains the foundation for providing safe hardware, safe software, and safe systems in the automotive industry, the high complexity and the real nature of AI creates a paradigm shift from safety assurance perspective, as safety failures may happen even in the absence of component failures. That is, hazards may result from the functional insufficiency and limitations of the technology in assuring the autonomous function, even in the absence of canonical hardware or software faults. This new safety challenge is presently referred to as safety of the intended functionality (SOTIF)[2], and is defined as the absence of unreasonable risk due to hazards resulting from functional insufficiency of the intended functionality or from misuse by users [1]. The final result is a static vision of possible dangerous situations (Fig. 1 is just one example).

4. Black swans

The problem is however far from being solved through a clear methodology and subsequent guidelines. As stated by the standard, the most critical challenge is to understand the so-called “black swans”, which means discovering hazards disregarded by the static risk analysis. Figure 3 helps understand. The risk associated with the introduction of autonomous functions is linked not only to the probability of incorrect execution of the safety-related functions (areas 1 and 4), but to the intrinsic uncertainty in system performance, in predicted (area 2) and unpredicted operating conditions (area 3). Methodologies for the systematic search of black swans constitute an unexplored area of research and this is where the digital twin comes into play.

Categorization of real-life driving scenarios	Known	Unknown
Safe	Area 1 Nominal behavior	Area 4 System robustness
Potentially hazardous	Area 2 Identified system limitations	Area 3 “Black swans”

Figure 3: SOTIF categorization of scenarios.

[2] https://www.iso.org/standard/70939.html

5. SOTIF and digital twin

The scheme in Fig. 4 helps summarize all the steps involved in the certification process. Static hazards are translated into specific requests to the simulation engine in order to investigate the realization of dangerous and critical events (e.g., collision/comfort of trajectory) under the simulated fragment of the operative design domain (ODD, e.g., no cameras, corridor corners, high speed and large crowd density). Finding the desired classes (safety versus criticalities) with respect to the dynamicity of the system (e.g., speed versus crowd density) is the final goal to pursue. The rationale of this investigation relies on the fact that static hazards may just simply suggest the categories of simulation runs, but cannot anticipate the exact mapping of the parameters leading to safety versus criticalities. It is rather necessary to perform several runs around the hypothesis made by SOTIF and extract knowledge in an automatic way. The approach follows explainable and reliable machine learning[3] [2] in order to improve the coverage of the scenarios considered, thus avoiding brute force simulation. The core of the methods is XAI as it may drive iteration with the experts in the field who, in turn, can understand AI reasoning and merge natural and artificial intelligence [3]. As to further details on functional architecture, risk analysis and preliminary results, the interested reader is invited to take a look at the extended version of the document available in the link below[4].

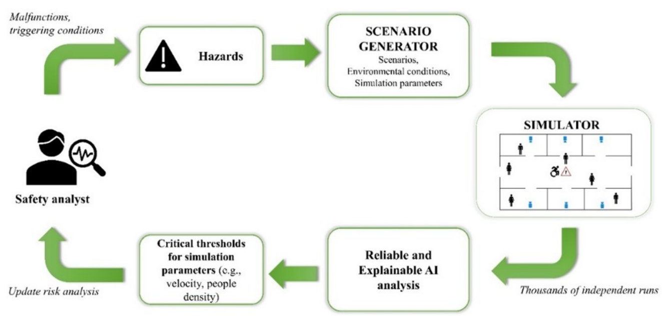


Figure 4: SOTIF and digital twin, empowered by AI.

ACKNOWLEDGEMENT

This work was supported in part by REXASI-PRO H-EU project, call HORIZON-CL4-2021-HUMAN-01-01, Grant agreement ID: 101070028.

REFERENCES

1. Kaiser, B. An integrative solution towards sotif and av safety, *IQPC SOTIF Conference*, (2019).
2. Narteni, S., Carlevaro, A., Dabbene, F., Muselli, M. and Mongelli, M. Confiderai: Conformal interpretable-by-design score function for explainable and reliable artificial intelligence, *Conformal and Probabilistic Prediction with Applications*, pp. 485–487, PMLR, (2023).
3. Lenatti, M., Carlevaro, A., Guergachi, A., Keshavjee, K., Mongelli, M. and Paglialonga, A. A novel method to derive personalized minimum viable recommendations for type 2 diabetes prevention based on counterfactual explanations, *Plos one*, **17** (11), e0272825, (2022).

[3] eXplainable AI (XAI): a machine learning (ML) model that is understandable by humans (e.g., expressed by rules of the if-then-else type). It lies into the transparency requirement of trustworthiness. Reliable AI (RAI): a ML model that it robust (often said resilient) to oscillations and attacks to its inputs in operation and keeps the model error under control.

[4] https://tinyurl.com/3nztxd7v

PHYSICS-AWARE SOFT-SENSORS FOR EMBEDDED DIGITAL TWINS

Fabio Marcuzzi, Giulio Giusteri, Erik Chinellato, *and* Laura Rinaldi
Dipartimento di Matematica, Universitá di Padova, Padova, Italy
e-mail: marcuzzi@math.unipd.it

Paolo Martin
SimNumerica s.r.l., Azzano Decimo (PN), Italy

Let us call a *Physics-aware soft sensor* the numerical algorithm that performs an indirect measurement by exploiting a physico-mathematical model plus a possible data-driven extension, within an estimation algorithm. In this sense, a physics-aware soft sensor solves a problem which is the combination of an inverse problem and a learning problem. It can conveniently be seen also as a generalization of an inverse problem, where part of the model is unknown and its construction is data driven. We will present two examples in mechanical vibrations analysis and in inverse heat transfer problems.

Keywords: *soft-sensors, inverse heat conduction problems, forcing term estimation, deep unfolding, sparse recovery*

1. Introduction

According to [1], Embedded Digital Twins, that is the virtual representation of physical systems that runs in an embedded system, are deployed on the edge within the embedded software stack to realize e.g. virtual sensors to enrich available information about physical variables and parameters that cannot be provided by direct physical measurements. These lacking measurements are estimated, at least roughly, by an algorithm that processes the available data, usually called a soft sensor. In the literature about soft-sensors, a strong emphasis is frequently put on black-box, data driven algorithmic techniques.

Let us call a *Physics-aware soft sensor* the numerical algorithm that performs an indirect measurement by exploiting a physico-mathematical model plus a possible data-driven extension, within an estimation algorithm. In this sense, a physics-aware soft sensor solves a problem which is the combination of an inverse problem and a learning problem. It can conveniently be seen also as a generalization of an inverse problem, where part of the model is unknown and its construction is data driven. We will present physics-aware soft-sensors of increasing complexity, pointing out that their complexity depends on:

- model complexity,
- interactions with the environment,
- centrality of physically measured variables in the virtual measurement process.

We start by briefly describing a couple of applications in electronic household appliances, involving a large number of items produced and tight computing resources. In the first one, i.e. indirect measurement of a mechanical load, the soft-sensor is physics-aware, while in the second one, indirect measurement of the load humidity during a drying process, the soft-sensor is data-driven (a neural network). We will briefly present also the technological aspects involved in computing and in model tuning of each single item at the production line. Then, we recall the recent, enormous improvement in computing performances given by low-cost electronics and consequent algorithmic possibilities available today for embedded digital twins, even exploiting models with distributed parameters, governed by PDEs. We will present two examples in mechanical vibrations analysis and in inverse heat transfer problems. We conclude by discussing some relevant issues in industry, when dealing with physics-aware soft-sensors, mainly:

- amount of experimental data needed to develop the application,
- effort needed at a posterior extension of the soft-sensor to product variants,

and show a clear practical advantage of physics-aware soft sensors compared with pure data-driven ones, when a physico-mathematical model can be set up, at least for some relevant parts of the real system.

2. Physical parameters estimation of a mechanical system from audio source-separation

The empirical estimation of physical parameters in a mechanical system is commonly performed in the Fourier domain, by computing the Frequency Response Function (FRF) or a Short-Time Fourier Transform (STFT). This is made usually from vibration measurements, given by accelerometer sensors, strongly fixed to the vibrating medium.

Using a microphone instead of the usual accelerometer, imposes to separate the acoustic source created by the process to be monitored, from the acoustics generated by the environment. In this setting, the source tracking property of Deep NMF [2] becomes crucial. Indeed, the *deep unfolding* paradigm, and precisely the Deep NMF, can be conveniently used for physical parameters estimation of a mechanical system, in terms of:

- performance with respect to the corresponding linear method (i.e. the NMF or Nonnegative Matrix Factorization);
- flexibility to create new model configurations, thanks to the physical explainability of the basis functions used in the learning process.

We will show results from an ad-hoc synthetic database and a real example involving experiments from a machine tool operation.

3. Internal voids/material-changes estimation from thermographic inspection

In the second example we will describe the results of an ongoing project about the embedded digital twin of a bread-making production line, focusing in particular on a soft-sensor that uses a Kalman Filter to estimate a distributed heat source on a thermal inverse problem [3].

Here, the formation of internal voids due to yeast activity is estimated by indirectly measuring an equivalent distributed heat source, whose shape can be analytically computed [3]. Hence, the reference model for the Kalman Filter is the coupling of a Finite Element model for the heat transfer within the known material (bread) and a data driven model for the unknown, distributed heat source, based on the analytical study [3].

We show how this device of interpreting a hidden material change as a fictitious heat source [4] gives better results than a general, data-driven optimization acting on the bare temperatures prediction error, and we show that, having a theoretical insight on some features of the shape of these fictitious heat sources [4], it is possible to tune the reference model of a Kalman Filter that allows to get reasonable estimates at a considerably reduced computational cost, that make it applicable to embedded digital twins.

REFERENCES

1. Hartmann, D. and der Auweraer, H. V. Digital twins, *CoRR*, **abs/2001.09747**, (2020).
2. Hershey, J. R., Roux, J. L. and Weninger, F. Deep unfolding: Model-based inspiration of novel deep architectures, *arXiv preprint arXiv:1409.2574*, (2014).
3. Rinaldi, L., Giusteri, G. G. and Marcuzzi, F. Replacing voids with fictitious sources in heat conduction problems, *submitted*, (2023).
4. Dessole, M. and Marcuzzi, F. Accurate detection of hidden material changes as fictitious heat sources, *Numerical Heat Transfer, Part B: Fundamentals*, pp. 1–24, (2023).

A NEW ILP MODEL FOR FINDING OPTIMAL EVACUATION PATHS IN EMERGENCY SCENARIOS

Paolo Ventura *and* Giuseppe Stecca
IASI-CNR, Rome, Italy
e-mail: paolo.ventura@iasi.cnr.it

We present a new time-indexed Integer Linear Programming model for the definition of optimal evacuation paths from an indoor area (museum, exhibition centre, etc.), in case of an emergency. The corresponding optimization algorithm has been integrated in a software tool that is powered with the data of the distribution of the visitors inside the considered site, provided by sensors or simulation tools. The output of the algorithm consists of a set of evacuation paths that will allow the visitors to reach the emergency exits in the shortest amount of time. In an *on-line* setting, such a tool provides the security department of the museum with safe evacuation plans that can be implemented in case of an emergency. Moreover, in a more *strategic* setting, it can be used as a simulation tool that can support the management to design the distribution of the exhibited works and to evaluate the maximum capacity of the rooms as well as the number and distribution of the emergency exits.

Keywords: *evacuation paths, integer linear programming, time-indexed formulations*

1. Introduction

The design, test, control and maintenance of safe and efficient evacuation plans for indoor spaces, both private or public, that are open to visitors, is a crucial issue for a municipal administration. Indeed, the problem regarding the definition of optimal evacuation paths has been widely studied in the literature, starting from the seminal work [1]. There, the author uses Integer Linear Programming (ILP) in order to model the problem as a maximum dynamic flow problem, where the capacity of each directed arc of the network depends on the amount of flow that uses the adjacent arcs (see [2] for a survey on the ILP-based approaches). More recently, in [3], an efficient algorithm has been proposed in order to calculate the evacuation paths from a building. The procedure minimizes the total evacuation time and assigns an optimal number of visitors to each of the defined paths. Unfortunately, the computational time rapidly grows with the dimension of the instance and the method cannot be used to get real-time solutions in real case scenarios. Other efficient approaches based on over-time flows or transshipment models can be found in [4, 5, 6]. Then, models and algorithms obtained by mixing optimization and simulation techniques have been defined in [7].

Here, we consider the problem of defining optimal paths for the evacuation of the visitors from a indoor area in case of an emergency (earthquake, fire, terroristic attack). In particular, we introduce a new time-index ILP formulation for the problem. The arcs of the directed graph that models the planimetry of the considered site are characterized by two given parameters assumed to be constant during the whole optimization process: the travel time and the max flow capacity in the time unit. This choice represented a good compromise among the accuracy of the results obtained and the computational time requested, that allows the whole procedure to be applicable in real case contexts.

2. The ILP model

2.1 Modeling the museum

Let R be the set of rooms/corridors (in the following, simply rooms) of the museum taken under consideration. Moreover, let $\underline{R} \subseteq R$ the set of rooms that are not available during the evacuation process. This can happen because of the emergency itself, or for different reasons (work in progress, maintenance, shortage of supervisory staff, etc.). For each room $r \in R$ we are given a capacity w_r, that is the maximum number of visitors that can simultaneously occupy the room at each instant of time. The planimetry of the museum is modelled by a directed graph $G = (N, A)$. The node set N is partitioned in the three subsets N_C, N_D and N_E. In particular: N_C are the *centroid* nodes, where the visitors are located at the beginning of the evacuation; N_D contains a couple of *passage* nodes for each door / gate among two rooms of the museum; N_E contains an *exit* node for each escape room of the museum. For each node $u \in N$, let $u(r) \in R$ be the room where the node belongs to and let $N(r)$ be set of nodes of each room $r \in R$. Clearly, $\cup_{r \in R} U(r) = N$. Then, for each arc $a = (u, v) \in A$, we are given: a travel time p_a and a capacity f_a, that is the maximum number of visitors that can traverse the arcs in a unit of time. Usually, the arcs of the graph G can be matched into symmetric couples (that is, for each $(u, v) \in A$, we also have $(v, u) \in A$. This rule admits two exceptions: i) one can enter in but not exit from an emergency exit; ii) one can exit from but not enter in an available rooms.

2.2 The ILP formulation

As we already mentioned, we introduce here a time-indexed ILP formulation for a new flow over time model of the problem. At the beginning of the evacuation process, we assume we are given the number v_u of visitors located in each centroid $u \in N_C$. Such a number can be provided by sensors located in the rooms, simulation models, estimations derived from historical data. Then, let T be the time horizon for the evacuation process. The variables of our formulation are the following

x_{ut} = number of visitors in node u at time t, $u \in N$, $t \in T$; y_{at} = number of visitors that start traversing arc a at time t, $a \in A$, $t \in T$;

$$q_{uv} = \begin{cases} 1 & \text{if arc } (u,v) \text{ is used by some visitor at during the evacuation process} \\ 0 & \text{otherwise} \end{cases} \qquad (u, v) \in A.$$

Moreover, in order to model the objective function, that is the minimization of the total evacuation time, we use the following binary variables

$$z_t = \begin{cases} 1 & \text{if all visitors reached the emergency exits at time } t \\ 0 & \text{otherwise} \end{cases} \qquad t \in T.$$

Then, we define a set of linear inequalities to model the following set of constraints:

- C1 *classical* flow conservation constraints that regulate the visitors located at each node of the graph at each instant of time;
- C2 at each instant of time, the room capacities cannot be exceeded;
- C3 at each instant of time t, the number of visitors that start traversing each arc $(u, v) \in A$ cannot exceed the flow capacity of the arc nor the number of visitors that are located in node u at time t;
- C4 at time t=1, the number of visitors in each node u is p_u;
- C5 if any arc (u, v) is used in the evacuation, then arc (v, u) is not.

3. The EVC-Suite

We used the ILP formulation defined in the previous section to solve the optimization problem of finding the evacuation paths that allows the visitors to reach the emergency exits of the museum in the smallest possible amount of time. In particular, we implemented a Python code that writes the ILP formulation which is then solved by the python callable libraries of the open source ILP optimization package COIN-OR. The whole optimization procedure is then embedded in a software package, called EVC-Suite and implemented by JustAnother s.r.l. EVC-Suite implements friendly routines that allows the user to easily: i) construct the model of the museum; ii) set the distribution of the visitors in the museum at the moment of the emergency; iii) solve the current optimal evacuation paths problem; iv) display the solution in a easily interpretable way. In Figure 3 a detail of the web app is depicted.

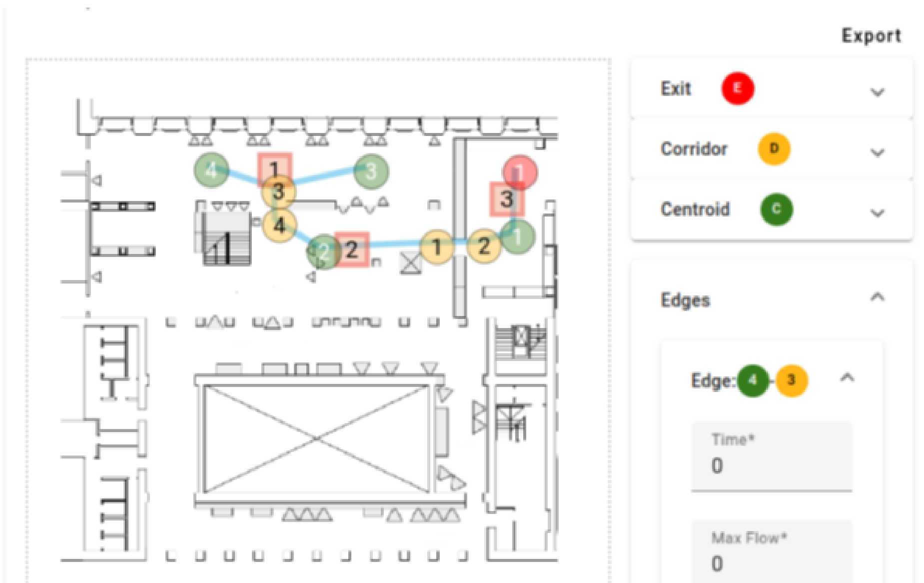


Figure 1: Screenshot of the ECV-Suite web app with details of the graph modeling the museum planimetry.

The overall approach has been tested on several instances that have been derived from the exhibition sites of the Museo Delle Civiltà in Rome (see here, the on-site experimentation conducted by CNR and NexSoft s.r.l. within the Smartour project [8]) and the S. Rocco Hub in Matera. All the computation experiments carried out gave evidence of the validity of the proposed approach, both in terms of quality of the solutions provided as well as of the computational times required.

REFERENCES

1. Choi, W., Hamacher, H. and Tufekci, S. Modeling of building evacuation problems by network flows with side constraints, *European Journal of Operational Research*, **1** (35), 98–110, (1988).
2. Hamacher, H. and Tjandra, S. Mathematical modelling of evacuation problems: A state of the art, *Pedestrian and Evacuation Dynamics*, pp. 227–266, (2002).
3. Chen, P. and Feng, F. A fast flow control algorithm for real-time emergency evacuation in large indoor areas, *Fire Safety Journal*, **4** (44), 732–740, (2009).
4. Dressler, D., Groß, M., Kappmeier, J.-P., Kelter, T., Kulbatzki, J., Plümpe, D., Schlechter, G., Schmidt, M., Skutella, M. and Temme, S. On the use of network flow techniques for assigning evacuees to exits, *Proceedings of the International Conference on Evacuation Modeling and Management*, (3), 205–215, (2010).
5. Dressler, D., Flötteröd, G., Lämmel, G., Nagel, K. and Skutella, M. Optimal evacuation solutions for large-scale scenarios, *Operations Research Proceedings*, pp. 239–244, (2011).
6. Schloter, M. and Skutella, M. Fast and memory-efficient algorithms for evacuation problems, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Mathematics*, pp. 821–840, (2017).
7. Abdelghany, A., Abdelghany, K., Mahmassani, H. and Alhalabi, W. Modeling framework for optimal evacuation of large-scale crowded pedestrian facilities, *European Journal of Operational Research*, **3** (273), 1105–1118, (2014).
8. Smart cities and communities (smartour). funding agency: Italian ministry of university and research, (2019-2023).

ENHANCED METHODOLOGY FOR DEVELOPING A SOFT SENSOR TO ESTIMATE SURFACE AREA AND ZETA POTENTIAL IN ADVANCED MANUFACTURING SYSTEMS

Roja P. Moghadam, Chinmay A. Shukla, *and* Vivek V. Ranade
Multiphase Reactors and Process Intensification Group
Bernal Institute, University of Limerick, Ireland
e-mail: Vivek.Ranade@ul.ie

Porous silica, with its unique structural and surface properties, has gained significant attention in the drug delivery field as a carrier material or a matrix for controlled and targeted drug release. Porous silica-based drug delivery systems offer high drug loading capacity, and protection of drugs. Silica-based excipients are commonly utilized to improve product performance due to the ease of modifying the silica surface through chemical reactions. The surface area and zeta potential of silica are critical factors in assessing the effectiveness of excipients. However, these critical quality attributes (CQAs) have to be measured using time-consuming offline methods. In this work, we have developed a quick method for estimating surface area and zeta potential of porous silica particles. We propose for the first time the use of dye adsorption and quick conductivity measurements combined with machine learning (ML) based soft sensor for estimating surface area and zeta potential. In order to adequately train the ML-based soft sensor, we have developed a phenomenological model and results obtained from it to complement available experimental data. The models, approach, and developed soft sensor presented will be useful in in-line CQAs estimation and will facilitate the development of decentralised 'Factory-in-a-Box' manufacturing of porous silica particles.

Keywords: *ML-based soft sensor, silica, excipient, surface area, zeta potential*

1. Introduction

Porous silica can be used to encapsulate drugs within its porous structure. The high surface area and large pore volume of silica allow for the efficient loading of drug molecules. The drug can be loaded into the pores through physical adsorption or by chemical conjugation [1]. Porous silica offers the advantage of controlled drug release. The porous structure allows for the diffusion of drugs out of the silica matrix in a sustained and controlled manner. By adjusting the pore size, pore volume, and surface properties of the silica, the release rate of the drug can be tailored to meet specific therapeutic needs [2, 3]. The surface area of silica particles in drug delivery systems significantly influences drug loading capacity, surface functionalization, drug release kinetics, interactions with biological systems, and stability. Understanding and optimizing the surface area is crucial for designing effective drug delivery platforms that can enhance therapeutic outcomes [1, 4]. A lower surface area may result in a slower release rate, which can be beneficial for sustained or controlled drug delivery. A larger surface area enhances the particle's ability to interact with target cells, improving bioavailability of drugs and internalization. The surface area of nanocarriers can be tuned based on synthesis methods for different drugs and targeting molecules [5]. The zeta potential of silica surface is another important surface property with a crucial role in drug delivery systems, especially those that utilize nanoparticles. Zeta potential refers to the electric potential difference between the surface of a particle and the surrounding fluid medium. It is a measure of the net electrical charge on the particle's surface. This potential arises from the adsorption of ions or charged molecules onto the particle surface, as well as the ionization of functional groups present on the particle's surface [6, 7]. surface area can indirectly influence zeta potential through factors like surface charge density or surface modification. For example, increasing the surface area of silica particles allows for more functional groups or surface modifications, which can alter the charge distribution on the surface and influence the resulting zeta potential [8, 9]. BET (Brunauer-Emmett-Teller) analysis, which is commonly used for determining specific surface area through gas adsorption, can be time-consuming due to the multiple steps involved in the measurement process [10]. Therefore, the use of soft sensor can be an excellent solution to solve this issue in the advanced manufacturing system. They are virtual sensors that estimate process variables or properties using mathematical models and available process data, eliminating the need for direct measurements. They can be employed to estimate CQAs such as surface area or other relevant parameters in real-time during manufacturing processes, providing timely information for process control and optimization [11, 12]. Overall, soft sensors based on machine learning (ML) tools play a crucial role in enabling real-time monitoring, control, adaptability, and estimations of CQAs to enhance operational efficiency within both digital twin and factory-in-a-box concepts. In this work, a methodology for developing a ML-based soft sensor for estimation of surface area and zeta potential is discussed with the aim of being used in advanced manufacturing.

2. Methodology

A stock solution of excipient was prepared using deionized water (DI) and a magnetic stirrer. The initial conductivity of the solution was measured. Various silica suspensions were prepared and sonicated. Through the addition of silica slurry to the excipient solution by syringe pump, the excipient within the solution underwent adsorption onto the porous silica particles, with the extent of adsorption being dependent on factors such as the surface area, zeta potential, and interactions between silica and excipient. The reduction in excipient concentration within the solution can be determined by measuring the conductivity of the solution. The achieved conductivity profiles were applied to develop the adsorption model and optimization of model parameters. By modeling excipient

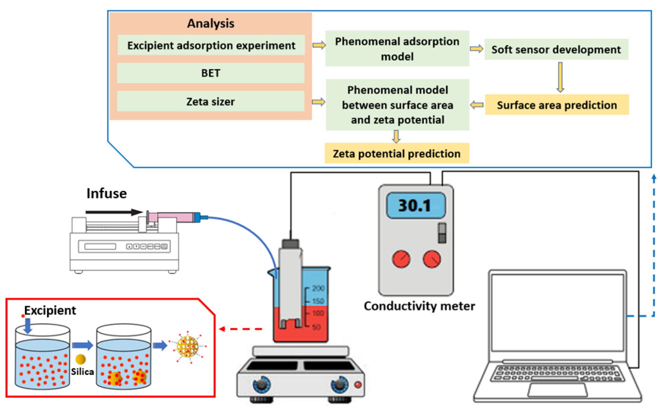


Figure 1: Concept of excipient adsorption experiment.

adsorption, a relationship between the concentration of dye and the surface area of silica can be established. This correlation enables us to predict various silica surface areas by considering the concentration of excipient. Since this relationship is grounded in physical principles, the data generated can be used to develop a soft sensor model, to provide highly accurate predictions. The initial differential equation governing the solute flux from the bulk to the liquid phase on the surface of the particle is directly proportional to the concentration difference between the bulk and the solute at the particle's surface in the liquid phase:

$$\frac{d(VC_D)}{dt}d = -Vk_{SL}\bar{a}(C_D - C_{Ds}). \tag{1}$$

Whereas V, k_{SL}, C_D, C_{Ds} and $\bar{a}$ are volume of solution, mass transfer coefficient, the concentration of dye in liquid, equilibrium concentration of excipient in liquid and surface of silica particles per volume of solution, respectively. The final adsorption model was achieved based on Eq. 1 and Langmuir isotherm model. The model-based profiles were applied for soft sensor training and development to estimate surface area. The predicted surface areas were adjacent to the real surface area and were used to estimate zeta potential with a proposed phenomenal model (a model based on surface charge density, surface area and zeta potential) between these two attributes. The soft sensor will be used for online estimation in bioinspired silica production process to control the CQAs.

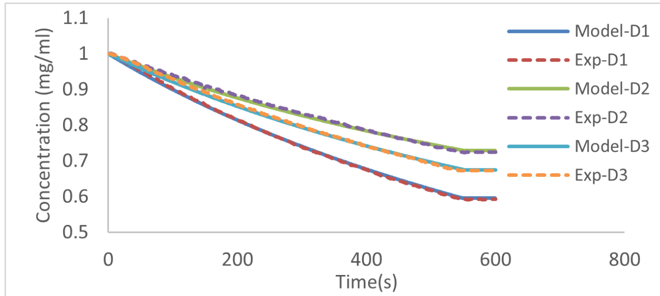


Figure 2: Comparison between adsorption experiment and model

3. Results

Outcome of this work will be practically useful for manufacturing process development and specially for characterizing any porous silica particles and developing a table top manufacturing of tailored silica for personalized medicine.

REFERENCES

1. Trzeciak, K., Chotera-Ouda, A., Bak-Sypien, I. I. and Potrzebowski, M. J. Mesoporous silica particles as drug delivery systems—the state of the art in loading methods and the recent progress in analytical techniques for monitoring these processes, *Pharmaceutics*, **13** (7), 950, (2021).

2. Bharti, C., Nagaich, U., Pal, A. K. and Gulati, N. Mesoporous silica nanoparticles in target drug delivery system: A review, *International journal of pharmaceutical investigation*, **5** (3), 124, (2015).

3. Djayanti, K., Maharjan, P., Cho, K. H., Jeong, S., Kim, M. S., Shin, M. C. and Min, K. A. Mesoporous silica nanoparticles as a potential nanoplatform: Therapeutic applications and considerations, *International Journal of Molecular Sciences*, **24** (7), 6349, (2023).

4. Peng, S., Huang, B., Lin, Y., Pei, G. and Zhang, L. Effect of surface functionalization and pore structure type on the release performance of mesoporous silica nanoparticles, *Microporous and Mesoporous Materials*, **336**, 111862, (2022).

5. Adepu, S. and Ramakrishna, S. Controlled drug delivery systems: current status and future directions, *Molecules*, **26** (19), 5905, (2021).

6. Honary, S. and Zahir, F. Effect of zeta potential on the properties of nano-drug delivery systems-a review (part 1), *Tropical journal of pharmaceutical research*, **12** (2), 255–264, (2013).

7. Thakur, P., Sonawane, S. S., Sonawane, S. H. and Bhanvase, B. A., (2020), Nanofluids-based delivery system, encapsulation of nanoparticles for stability to make stable nanofluids. *Encapsulation of active molecules and their delivery system*, pp. 141–152, Elsevier.

8. Ge, Z. and Wang, Y. Estimation of nanodiamond surface charge density from zeta potential and molecular dynamics simulations, *The Journal of Physical Chemistry B*, **121** (15), 3394–3402, (2017).

9. Pochapski, D. J., Carvalho dos Santos, C., Leite, G. W., Pulcinelli, S. H. and Santilli, C. V. Zeta potential and colloidal stability predictions for inorganic nanoparticle dispersions: Effects of experimental conditions and electrokinetic models on the interpretation of results, *Langmuir*, **37** (45), 13379–13389, (2021).

10. Ambroz, F., Macdonald, T. J., Martis, V. and Parkin, I. P. Evaluation of the bet theory for the characterization of meso and microporous mofs, *Small methods*, **2** (11), 1800173, (2018).

11. Parvizi Moghadam, R., Sadeghi, J. and Shahraki, F. Optimization of time-variable-parameter model for data-based soft sensor of industrial debutanizer, *Optimal Control Applications and Methods*, **41** (2), 381–394, (2020).

12. Iglesias Jr, C. F., Ristovski, M., Bolic, M. and Cuperlovic-Culf, M. raav manufacturing: The challenges of soft sensing during upstream processing, *Bioengineering*, **10** (2), 229, (2023).

SOME CHALLENGES AND METHODS FOR DIGITAL DEVELOPMENT OF NEW GENERATION FUELS

Francesco Saverio Marra
CNR-STEMS, Napoli, Italy
e-mail: francescosaverio.marra@sterms.cnr.it

New generation fuels, specifically those produced by renewable feedstocks and energy, are recognized as an unavoidable pillar to sustain the future economy, having important roles in long-term energy storage, clean energy production, heavy transport, energy-intensive industrial processes and even other applications. With the decline of fossil sources, traditional approaches for the study of the behaviour of new fuels are revealing inadequate for several reasons, among them the increased complexity of some new fuel molecules or because of the need to adopt new combustion regimes. Contemporary, both computational physics and mathematical methods, on one hand, and computational power and data management capabilities, are increasingly making it possible to develop new approaches for effective investigation of the properties of the new fuels and the design of the systems aiming at their adoption for energy conversion, both steps required for the development of digital twins of the new combustion systems. This presentation aims to give a brief overview of the challenges to face for an accurate representation of the new generation fuels and the correct reproduction by modelling their physical behaviour when employed in processes for energy conversion, together with an illustration of some new methodologies that are being developed at the CNR.

Keywords: *new generation fuels, detailed chemical mechanisms, bifurcation maps, community analysis, mechanisms reduction*

1. Introduction

The search for substitutes for fossil fuels is at the core of the energy transition dictated by the consequences of an increasing scarcity (for objective as well as political and economic factors) of these sources and the effect on the climate of their burning. Several programs are currently in development to find proper substitutes, and several options are on the table: use clean energy and abundant materials to produce small fuels molecules like hydrogen and ammonia, reprocess the sequestrated CO2 to produce fuels with medium size molecules like synthetic methane, or use/re-use renewable complex materials like bio-masses to produce complex fuel molecules, like drop-in substitutes of gasoline or jet fuels. In all these cases, technological challenges are open. For smaller molecules, new combustion processes must be envisaged for the clean combustion of the small molecules, whose characteristics are still far to be fully understood. In the case of complex molecules, the strict range of the properties required for the substitute fuels to be adopted in engines like jet turbofan is not easy to achieve and to stably maintain due to the continuous variation of the feedstock material. Not less important is the pursuit of a continuous improvement of the efficiency and emissions capabilities of the engines.

All these aspects claim for new and effective approaches to gain knowledge on the behaviour of the new combustion systems, and to design new combustors. With respect to the past, when refineries were optimized for the oil coming from a single well or a mix of well-established sources for decades, nowadays much more flexibility is required on both sources and engines that must accommodate varying properties of the final fuel product.

This challenge is further complicated by the need to adopt very detailed combustion mechanisms (up to thousands of species and tens of thousands of reactions) to catch the different behaviours of the new fuels and their mixing in the different operating conditions and mixture compositions. Figure 1 illustrates the evolution in the past years of the detailed chemical mechanisms developed for the description of combustion processes. Figure 1, left, indicates the number of species and reactions involved in the increasingly complex mechanisms, now easily extending to close to 10,000 species and 30,000 reactions. Figure 1, right, estimate the computational effort in computing the number of exponential function evaluations to compute the contributions in the balance equations for a multicomponent mixture coming from the detailed mechanism in a single state. It clearly results that even with the most powerful high-performance computing facilities, the tout-court adoption of detailed chemical mechanisms is prohibitive.

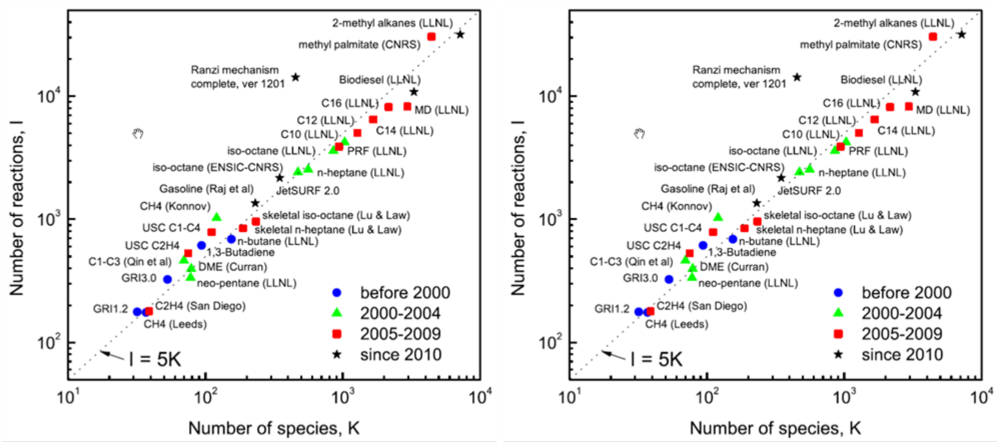


Figure 1: On the left, number of species and reactions of the detailed combustion reaction mechanisms. On the right, the corresponding computational complexity measured in terms of the computation of exponential functions for a single state determination. Source: ref. [1].

2. New directions of development

Several approaches are being developed to deal with these challenges. Most of them are devoted to reducing the computational effort by both reducing the dimension of the problem [2] or by making more efficient the computational efficiency [3]. Here the focus will be given to some original contributions developed at CNR-STEMS in cooperation with other scientific partners.

At the base for a correct reproduction of the physical phenomena by a virtual representation is a correct knowledge of its behaviour at least in the conditions aimed to be reproduced. In the case of combustion processes, the strong non-linearity of the chemical and physical interactions makes risky the attempt to derive unknown conditions as interpolation of known ones. This is essentially the procedure adopted even with the most advanced machine learning methods, also when based on experimental observations. This kind of procedure can be safely adopted only if prior knowledge of the regions in the parameters space where the system behaves sufficiently smooth has been gained, thus limiting the application of the virtual representation to the safe regions [4].

To gain this kind of knowledge, the bifurcation analysis offers the theoretical bases and the computational tools to systematically investigate the regions of the parameter spaces where “catastrophic” changes in the system behaviour occur, thus indicating the regions where “digital twins” can safely operate. In the case of systems represented by detailed chemical mechanisms, the use of tools like parametric continuation is not straightforward and is computationally extremely demanding. Special procedures have been developed to make affordable this type of analysis even for the most complex chemical mechanisms [5]. Even if they have been adopted up to now only for simple archetypal reactors, like the perfectly stirred reactor, thanks to the parallelism existent between this type of reactor and the computational cell of a finite volume CFD representation of a complex reactor, important information can be gained. In terms of knowledge gain, recognizing among such a large number of species and reactions, those mainly contributing to the observed behaviour of the system, is also important. Being manual inspection prohibitive, automatic methods must be envisaged. This aspect is being pursued by developing algorithms based on the representation of the detailed chemical mechanism as a bipartite network and thus employing concepts and methods of networks analysis, like community partitions and coarse-grained states. This approach appears, in conjunction with the bifurcation analysis, very promising to recognize how the evolution of a chemical mechanism changes while varying the actual conditions, allowing the identification of the regions where potentially very reduced mechanisms can be developed and safely applied to make affordable the computation of a digital twin [6].

The knowledge of the most critical regions in the parameter spaces also allows the development of robust and accurate reduced chemical mechanisms. This objective has been pursued by developing a systematic procedure for the reduction of detailed chemical mechanisms based on the weighting of the contribution of the species and reactions with respect to thermodynamic functions [7]. Specifically, the adoption of the entropy production rate as the weighting function has already permitted the development of efficient skeletal mechanisms for synthetic aviation fuels. The computational efficiency of this approach has been adopted also for developing the on-the-fly reduction of detailed mechanisms in CFD codes [8]. More recently, for the development of ultra-reduced mechanisms, named virtual chemistry [9], the possibility of reproducing with just a very few reactions the complex behaviour observed close to the ignition and extinction points of a reactive mixture, identified with the methods of bifurcation analysis, is being investigated.

3. Conclusions

Some of the challenges to face in some aspects of the development of Digital Twins for the design and control of energy systems involving combustion processes have been outlined. The presentation will deepen the here shortly illustrated arguments, furnishing practical examples of application of the proposed methods.

ACKNOWLEDGEMENTS

Several people contributed to these activities. I have to mention the most important contributions given by Dr Luigi Acampora in the development of the algorithms for the bifurcation analysis and the skeletal reduction, Prof. Emanuele Martelli, for his suggestions and support on on-the-fly skeletal reduction, Prof. Lin Ji for her invaluable contribution on the methods and algorithms for network analysis, Prof. Christos Frouzakis and Dr Mahdi Koosbaghi for the cooperation on the entropy reduction methods, Prof. Benoìt Fiorina and Prof. Nasser Darabiha for introducing me to the virtual chemistry approach.

REFERENCES

1. Law, C. K. Combustion at a crossroads: Status and prospects, *Proceedings of the Combustion Institute*, **31** (1), 1–29, (2007).

2. Turányi, T. and Tomlin, A. S., *Analysis of kinetic reaction mechanisms*, vol. 20, Springer (2014).

3. D'Alessio, G., Cuoci, A. and Parente, A. Feature extraction and artificial neural networks for the on-the-fly classification of high-dimensional thermochemical spaces in adaptive-chemistry simulations, *Data-Centric Engineering*, **2**, e2, (2021).

4. Acampora, L., Marra, F. and Martelli, E. Comparison of different ch4-air combustion mechanisms in a perfectly stirred reactor with oscillating residence times close to extinction, *Combustion Science and Technology*, **188** (4-5), 707–718, (2016).

5. Acampora, L. and Marra, F. S. Numerical strategies for the bifurcation analysis of perfectly stirred reactors with detailed combustion mechanisms, *Computers & Chemical Engineering*, **82**, 273–282, (2015).

6. Du, P., Li, M., Liang, S., Ji, L., Acampora, L. and Marra, F. S. Wide-parameter coarse-grained state mechanism analysis in the methane combustion system, *Reaction Chemistry & Engineering*, **8** (6), 1362–1375, (2023).

7. Acampora, L., Kooshkbaghi, M., Frouzakis, C. E. and Marra, F. S. Generalized entropy production analysis for mechanism reduction, *Combustion Theory and Modelling*, **23** (2), 197–209, (2019).

8. Marra, F. and Acampora, L. Implementation of the gepa method for on-the-fly reduction of detailed chemical mechanisms in openfoam, *The 18th OpenFOAM Workshop*, (2023).

9. Cailler, M., Darabiha, N. and Fiorina, B. Development of a virtual optimized chemistry method. application to hydrocarbon/air combustion, *Combustion and Flame*, **211**, 281–302, (2020).

AUTHORS INDEX