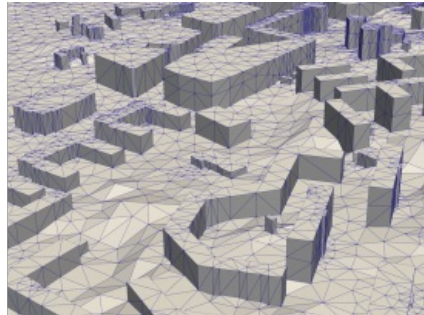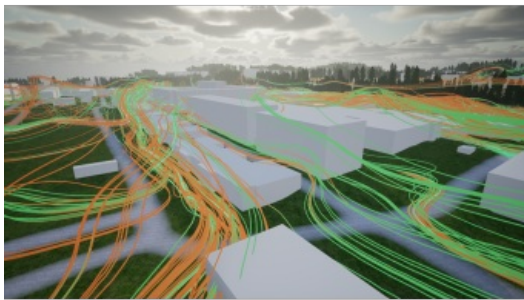## FROM RAW DATA TO END PRODUCT

Modelling and simulating cities with digital twins


The digital twin concept has virtually exploded in recent years, but within the built environment the traditional term has long been '3D city model'. However, the digital twin is increasingly gaining acceptance as a useful concept that extends beyond 3D city models for not only modelling but also simulating cities. So what are digital twins, how are they being used, and what challenges are involved?
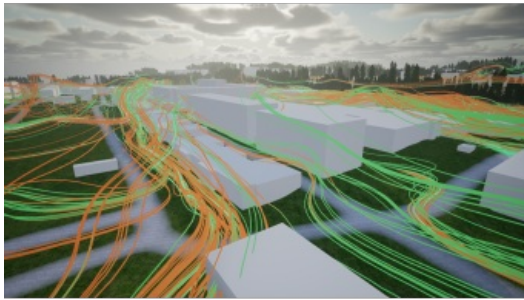
The digital twin concept has virtually exploded in recent years, as evidenced by the exponentially growing number of scientific articles making use of the concept. The concept originates from the manufacturing industry where the use of CAD models enables the creation of exact digital replicas of components and products. The earliest use of the term dates to 2003 and is often credited to Grieves and Vickers, but earlier references to the concept can be found; certainly, the understanding that mathematical and, more recently, digital models of physical systems are of enormous importance to both science and engineering dates back centuries.
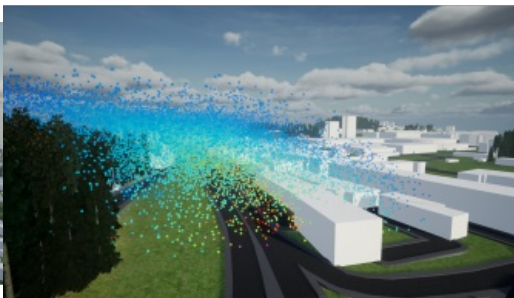
Defining the digital twin
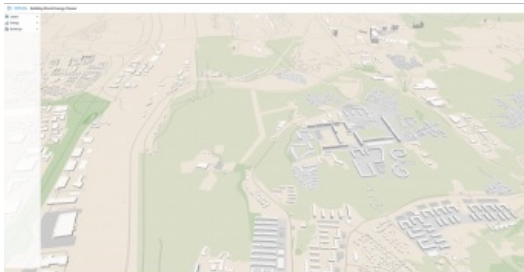
What is then a digital twin? In both the scientific literature and, even more so, in commercial narratives, 'digital twin' is a quite elastic concept used to label technologies or systems that may or may not live up to all the criteria of a digital twin. Does a digital twin need to include a 3D model? Does a digital twin need to include real-time sensor data? Does a digital twin need to include mathematical modelling and simulation?

It is instructive and interesting to look at some of the many definitions that have been proposed for the digital twin concept, since there seems to be some convergence towards a universally accepted definition. For example, most definitions now agree that a digital twin is a *model* of a physical system that mirrors the physical system in real time and enables analysis and prediction of the physical system. The digital twin can thus be used to both analyse the physical system ('what is') and to predict its future behaviour under given assumptions ('what may be').

This definition is partly overlapping with that of Rasheed et al. (2020): "A digital twin is defined as a virtual representation of a physical asset enabled through data and simulators for real-time prediction, optimization, monitoring, controlling and improved decision-making." A similar definition is used by IBM: "A digital twin is a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision-making." These latter two definitions emphasize two technologies that may be used to enable the predictive function of the digital twin: simulation and machine learning.

Figure 1: Detail of the boundary of a 3D tetrahedral volume mesh generated from cadastral and point cloud data.

A definition often seen in earlier literature on digital twins is that by Glaessgen and Stargel (2012): "A digital twin is an integrated multiphysics, multiscale, probabilistic simulation of an as-built [...] system that uses the best available physical models, sensor updates, [...], to mirror the life of its corresponding [physical] twin." A somewhat simpler definition is given on Wikipedia: "A digital twin is a virtual representation that serves as the real-time digital counterpart of a physical object or process."

In the domain of digital cities, Stoter et al. (2021) emphasize the use of 3D city models as an essential part of a digital twin: "[A digital twin] should be based on 3D city models, containing objects with geometric and semantic information; it should contain real-time sensor data; and it should integrate a variety of analyses and simulations to be able to make the best design, planning and intervention decisions." This definition is a reminder of the long tradition within the built environment of creating 3D models of cities and buildings, which may be enriched with semantic data and used as a basis for analysis, including, e.g., daylight and energy analysis, as well as simulation of things like traffic, wind comfort or air quality. Within the built environment, the traditional term is '3D city model', and it is only very recently that the digital twin concept has started to gain acceptance as a useful concept and something that extends beyond 3D city models.

Raw data

The starting point for the creation of a digital twin of a city is access to raw data. This data may be created from aerial scans in the form of point clouds. The point clouds are then processed to create 2D or 3D city models. Access to data varies between countries and may not always be open or freely available. In Sweden, Lantmäteriet, the Swedish mapping, cadastral and land registration authority, provides (for a fee) a range of datasets including both point clouds and 2D maps for the whole of Sweden. Meanwhile, more detailed and higher-quality datasets, including 3D models, are owned by the local municipalities. In the Netherlands, the situation is different. The 3D Baseregister Addresses and Buildings (BAG) provides free and open access to 3D models for all ten million buildings in the country. Moreover, the dataset is regularly and automatically rebuilt to provide an up-to-date 3D model of the entire country.

Figure 2: Preliminary results from the simulation of urban wind comfort by the DTCC (rendered in Unreal Engine).

Data models

To build a digital twin of some complexity and use, it is essential to consider which data model is used to define the digital twin. Note that this is different from the mathematical model(s) used for simulation and prediction. The choice of data model dictates which data may be represented, and which use cases that may be supported by the digital twin. The data model is the implementation of a certain ontology, defined either explicitly or implicitly by the implementation. The ontology defines how the data of the digital twin may be described and understood, in terms of classes, attributes and relations. Several data models and corresponding exchange formats have been proposed for city modelling. One of the most prominent is CityGML, which is a standard of the Open Geospatial Consortium (OGC). The related CityJSON format (also an OGC standard) is a simplified and much more programmer-friendly encoding of the CityGML model.

Common to many data models for city modelling is the concept of level of detail (LOD). This concept enables the data model to store different representations of the city with varying levels of detail (geometric resolution) for different purposes. The coexistence of several levels of detail in a digital twin emphasizes that the digital twin is indeed a model of the physical system it mirrors, and that the digital representation as well as its accuracy are dictated by both the use cases the digital twin is designed for, the quality of data and the available computational resources.

Figure 3: Visualization of a digital twin of the campus of Chalmers University of Technology in Gothenburg (rendered in Unreal Engine).

Data generation

Different use cases of a digital twin will often require very different data representations. For the modelling of a city, the understanding of what constitutes a high-quality 3D model may differ widely if one were to ask an architect or a computational scientist. For the architect, a high-quality 3D model may mean a detailed set of surface meshes describing the topography of the city and the geometry of its buildings. The surface meshes may be both non-conforming and non-matching since the meshes will mostly be used for visualization and simple calculations like daylight analysis. For a computational scientist, on the other hand, a high-quality 3D model may mean a low-resolution, boundary-fitted and conforming volume mesh that may be used to run things like computational flow dynamics (CFD) simulations.

The team at the Digital Twin Cities Centre (DTCC) in Sweden is currently developing an open-source platform for representation and generation of high-quality data models for digital twins of cities. One of the key steps is the high-performance generation of high-quality surface meshes and tetrahedral volume meshes from cadastral and point cloud data (Figure 1). This allows for simple and efficient generation of 3D models for any part of Sweden (or any other part of the world that has compatible data). The mesh generation is currently limited to LOD1 models, meaning that the buildings are represented as polygonal prisms (flat rooftops). However, work is underway to extend the mesh generation to LOD2 models, including non-flat rooftop shapes based on segmentation of rooftops from orthophotos using machine learning techniques.

Figure 4: Visualization of a road network as part of a digital twin of an area in central Gothenburg (rendered in Unreal Engine).

Modelling and simulation

With computational meshes readily available for any city, it is natural to consider the use of physics-based modelling and simulation to enable advanced analysis and prediction. Examples of physical phenomena that may be of relevance in the study of cities include urban wind comfort (wind conditions at street level), air quality, noise, and electromagnetic fields (for network coverage analysis).

One example of such a simulation currently being investigated at the DTCC is the simulation of urban wind comfort. The simulation uses

an immersed boundary method for the Reynolds Averaged Navier-Stokes (RANS) equations using IPS IBOFlow. The current focus is on verification and validation of the simulation results for a selection of benchmark cases of city wind simulation that have previously been studied in wind tunnels. Some preliminary results are shown in Figure 2. Other examples of physics-based modelling and simulation currently being investigated at the DTCC include simulation of air quality, street noise, crowd movement, and geotechnical simulation based on elastoplastic models of the soft clay that constitutes much of the underground of Gothenburg.

Visualization

Data visualization on the urban scale is itself a field of ongoing research. Physical information, such as wind flow and air quality (i.e., concentration of pollutants), needs to be represented in a way that is understandable to the end-user, but without overly simplifying scientific results. Effective communication of results requires several design iterations where both researchers, developers and end-users/stakeholders are involved. The DTCC actively collaborates with major stakeholders such as the Swedish Transport Agency in research projects that explore how to best communicate simulation results to different user groups. The ongoing research projects within the visualization domain focus on different solutions for data derivation, preparation, bundling, homogenization and propagation. Different graphical engines are tested and employed, e.g., Unreal Engine and OpenGL, as well as various web application implementations based on Mapbox, CesiumJS and Babylon.js.

Figure 5: Visualization of volumetric data as part of a digital twin of an area in central Gothenburg (rendered in Unreal Engine).

Technical challenges

There are many challenges involved in creating a digital twin of something so complex as a city. Since the city itself is a complex system involving not only the streets and buildings of the city, but also its inhabitants, the cars driving on the streets, the interaction with the surrounding environment (wind and water), as well as the infrastructure below ground – which is sometimes overlooked yet very substantial – it is only natural that the creation of a digital twin of the city is equally complex. The task of building the digital twin is therefore by necessity a project that must involve experts from many different disciplines. The technical challenges involved in building the digital twin will involve both interdisciplinary challenges in the collaboration between team members from vastly different disciplines, as well as already established intradisciplinary or domain-specific technical challenges such as how to implement a finite element solver most efficiently for one of the many mathematical models that together constitute the multiphysics model that is the digital twin.

Non-technical challenges

Setting the technical challenges aside, the main challenges experienced so far at the DTCC are all related to data:

- *Data ownership across organizations*: Data is often neither free nor open. Organizations, even municipalities, are reluctant to share their data freely since they at some point made a substantial investment in collecting and curating it. This differs in different parts of the world; in some cases (like in the Netherlands), the data is indeed free and open.
- *Data quality across disciplines*: As in the abovementioned example of a mesh for use by an architect versus a computational scientist, a certain dataset may be regarded as being of high quality for a particular use case but may be of very low quality for another use case.
- *Data sustainability across time*: The creation of a digital twin must be understood as a process rather than as a project. There are many examples of cities, municipalities and other organizations that invest in projects for the creation of a 3D model or even a digital twin, only to realize a few years (or even just months) later that the digital twin no longer mirrors the physical twin since reality is constantly changing. The only way to reconnect the digital twin with the physical twin is then to invest in a new and costly project. Therefore, the process of creating the digital twin must instead be automated so that it can be continuously rebuilt and regenerated.

Figure 6: Visualization of a digital twin of a Gothenburg suburb (rendered using WebGL).

Further reading

2020. Ketzler, V. Naserentin, F. Latino, C. Zangelidis, L. Thuvander, and A. Logg. Digital Twins for Cities: A State of the Art Review. *Built Environment*, 46(4):547–573, December 2020.
DOI: 10.2148/benv.46.4.547.
2021. Rasheed, O. San, and T. Kvamsdal. Digital Twin: Values, Challenges and Enablers From a Modeling Perspective. IEEE Access, 8:21980–22012, 2020. ISSN 2169-3536.
DOI: 10.1109/ACCESS.2020.2970143. Conference Name: IEEE Access.
2022. Stoter, K. A. Ohori, and F. Noardo. Digital Twins: A Comprehensive Solution or Hopeful Vision? *GIM International*, October 2021.