Synthetic gas networks for the statistical assessment of low-carbon distribution systems

Synthetic gas networks for the statistical assessment of low-carbon distribution systems

Enrico Vaccariello*[1,2], Riccardo Trinchero[2], Pierluigi Leone[1], Igor S. Stievano[2]

[1] Department of Energy "*Galileo Ferraris*" (DENERG)
[2] Department of Electronics and Telecommunications (DET)
Politecnico di Torino – 24 C.so Duca degli Abruzzi, 10129 Torino, Italy
*Corresponding author at enrico.vaccariello@polito.it

keywords: synthetic network models**,** distribution gas grids**,** gas network design**,** integrated energy systems, low-carbon and renewable gases

Abstract. Most of the simulation studies on energy networks, including gas grids, derive their results from a limited number of network models. The findings of these works are therefore affected by a substantial case-specificity, which partially limits their validity and prevents their generalisation. To overcome this limitation, the present work proposes a novel statistical-based approach for studying distribution gas networks, enabled by a generator of random gas grids with accurate technical designs and structural features. Ten-thousand random and unique networks are produced in three different tests, where increasingly tight constraints are applied to the synthetisation process for a higher control over the generated grids. The experiments verify the accuracy of the tool and highlight that substantial variations can be found in the hydraulic behaviour (pressures and gas velocities) and structural properties (pipe diameters and network volumes) of real-world gas networks. The observed 10,000 gas grids evidence the information gain offered by statistical-based approaches with respect to traditional case-specific studies. The tool opens a broad range of applications which include, but are not limited to, statistical analyses on the distributed injection of alternative gases, like hydrogen, in integrated, low-carbon, energy systems.

I. Introduction

The ongoing integration of distributed renewable sources of energy (RES) has drastically changed the functioning paradigms of distribution system operators (DSO). At present time, the electrical grid has undoubtedly experienced the strongest transition. DSOs in the power sector switched from passive to active operational models and witnessed increased risks of voltage and current violations as non-dispatchable distributed generators (DG) proliferate at fast pace. Nevertheless, not only the power grid is expected to play an important role in the accommodation of electrical RES: as a matter of fact, increasing attention is being devoted to the cross-sectorial integration among energy subsystems (e.g., power, gas, heating) and infrastructures, to achieve an overall higher system efficiency and flexibility [1]. One of the most promising coupling options among the power and gas sectors, as indicated in the European Strategy for Energy System Integration [2], is represented by electrolysis. Electrolysers produce a gaseous fuel, namely hydrogen, exploiting electrical power. Among the possible options, hydrogen produced in this way can be deployed as an industrial feedstock or as a fuel for mobility [3]. In other cases, it can substitute conventional gaseous fuels (e.g., methane) in pure, blended or methanised (i.e., synthetic natural gas, SNG) forms [4,5]. The practice of injecting and distributing hydrogen within the existing gas networks in blended form is suggested in the European Hydrogen Strategy, especially in a transitional and scaling up stage [6,7]. Hydrogen is already being deployed in the gas pipelines in more than 20 ongoing power-to-gas (PtG) projects [4,5,8] and the number of pilot and full-scale demonstrators is expected to raise in the next future, as the European Commission envisages up to 470 billion euros of investments and several funding programmes for the hydrogen sector [6,9].

The forthcoming escalation of the use of hydrogen within the gas networks and, in general, of the distributed injection of renewable gases like SNG, biogas and biomethane [2], requires assessing the readiness of the existing grids to non-conventional operational schemes. In this regard, the scientific literature is rich in gas network modelling tools and assessments of the compatibility of gas supply systems with a diversity of gas sources. Abeysekera et al. [10] evaluated the response of a distribution gas network considering the deployment of hydrogen and upgraded biogas. Similarly, the introduction of up to 10% of hydrogen in a distribution grid was studied in [11]. In Ref. [12], gas quality limits were considered to assess the maximum allowable injection of biogas within a medium pressure distribution grid. Ref. [13] considered the steady-state injection of hydrogen and upgraded biogas in a transmission network, and investigated the effects on pressures, temperatures, and nodal gas qualities. The dynamic modelling of a transmission pipeline in presence of H_2 injection was addressed in [14], while a method for the transient tracking of gas quality in networked systems is found in [15]. The computation of gas flows in multi-pressure systems was provided by [16]. Finally, papers [17,18] proposed the optimal sizing of a PtG unit for grid injection applications at distribution level, with attention to the maximum hydrogen concentration in the network.

As a common practice, the aforementioned contributions implement either real-world or ad-hoc fictitious network models as testbeds of fluid-dynamic simulations. The quantitative results provided by these studies are therefore affected by a case-specificity which may limit the extent of their validity. This limitation could be overcome by observing the findings of a given experiment over several testbeds (i.e., network models), thus shifting from a case-specific approach to a statistical analysis conducted over several networks. Nevertheless, the main factor hindering this practice is data availability: publicly available energy network models are very limited due to privacy, security, and industrial secrecy issues. The scarcity of data is even more severe for gas grids, given that IEEE benchmarks and other real-world power grid models are accessible from several sources [19–21].

A number of solutions have been proposed to produce synthetic energy network models (mostly power grids) to overcome the lack of available data and to tackle data secrecy issues. Some tools make use of geographic information system (GIS) data to estimate and plan the location and the topology of the power grid within a target geographical area [22–24]. Other algorithms do not account for geographical information, but rather implement spatial and distance-based criteria to establish the topology of the synthetic grids [25,26]. Alternative approaches neglect the spatial embeddedness of the networks and establish grid topologies based on statistical observations [27–29] or inspired by reference structures like small-world networks [30]. Most of the aforementioned contributions produce network models readily functional for simulation (electrical power flow) purposes.

Concerning gas grids, extensive research has been carried out on hydraulic simulation and optimization methods. However, tools for the automated generation of synthetic network models are still lacking. Work accomplished in this field has mostly regarded optimization problems applied to the grid topology [31,32] and the selection of network components [33–35].

Based on the above considerations, this work proposes a novel tool for the random generation of finished distribution gas grid models, that can be readily deployed for simulation purposes. The algorithm implements our previous work for the automated creation of correct network topologies [36] and accomplishes the technical sizing of the networks based on target design parameters. Using a reference input grid, the tool can synthetise a virtually infinite number of distribution gas networks, thus enabling the accomplishment of statistical studies that may include, but are not limited to, the distributed injection of renewable gases in multi-vector energy systems. As discussed within the paper, the algorithm can be also deployed to produce anonym twins of real-world sensible networks, that can be disclosed without incurring into privacy, security, and industrial secrecy issues.

In the next sections, we extensively describe the tool (Sections IIand III) and demonstrate its capability to generate ten-thousand finished models of synthetic gas networks with consistent technical and structural properties (“Case A” and “Case B”). We also show that a higher control over the properties of the generated networks can be obtained by imposing a-priori constraints (“Case B”) or a-posteriori enforcements (“Case C”). The hydraulic response of the grids (nodal pressures and gas velocity in pipes) and their structural characteristics (pipe diameters and network volumes) are analysed. The results highlight that the network properties can feature substantial variations that cannot be taken into account in case-specific studies. Accordingly, the discussion and conclusions emphasise the added value of statistical studies on gas networks, in perspective applications involving the adoption of alternative fuels and/or unconventional operational schemes.

List of acronyms and abbreviations

CDF - Cumulative distribution function
DMP - Design minimum pressure
DSO – Distribution system operator
EOS - Equation of state
GIS – Geographic information systems
KS - Kolmogorov-Smirnov
KSstat2 - Two-sample Kolmogorov-Smirnov statistic
MP - Medium pressure
NG - Natural gas

PDF - Probability distribution function
PtG – Power-to-gas
RES – Renewable energy source
SNG – Synthetic natural gas

II. Generation of statistically correct gas network topologies

The first step towards the generation of synthetic models of gas networks is the establishment of the network topologies. This stage is inspired from a novel methodology described in [36], which acknowledges the spatial embeddedness of the grids. The synthetic networks mimic the topological properties and the spatial density of an input reference gas grid. In the original paper, the methodology is proposed for infrastructures operated at multiple pressure levels interfaced by pressure reduction stations. In the current application only single-pressure networks are contemplated. Figure 1 provides a high-level diagram of the procedure, which is articulated in three main stages: the first stage establishes the spatial location of the N synthetic nodes; the second stage ensures the connectivity of the synthetic network by creating N-1 links among the nodes, thus forming a spanning tree; in the third and last stage, the network is reinforced including additional edges which form L loops.

Stage 1: Generation of synthetic nodes

The synthetic nodes replicate the spatial density of a reference real-world input network. The spatial distribution of the reference network is correlated to the underlying population and geographical properties [25,36]. The position of nodes can therefore be clustered using mixture models: Gaussian Mixture Models (GMM) are fitted to the node coordinates of the reference grid and in turn deployed for a probabilistic generation of the synthetic nodes.

Stage 2: Establishment of connectivity (spanning tree)

The connectivity of the network is established in probabilistic fashion, which favours the creation of connections among near pairs of nodes. The procedure iteratively selects one of the synthetic nodes and adds it to a growing structure of the tree. At each iteration, the probability for one node to be picked and added to the tree depends on its position and is established by a bivariate Gaussian-shaped probability distribution function (PDF), with mean $\overline{\mathbf{p}}$ (average position of the network nodes – x-y coordinates in meters) and standard deviation α. Parameter α regulates the width of the PDF and must be suitably calibrated. It is set equal to 93 according to [36]. Every time a new node is sampled, it is connected to the closest node in the growing tree structure. For a more comprehensive description of the procedures of Stages 1 and 2, we suggest referring to [36].

Stage 3: Creation of loops

Unlike electricity grids, distribution gas grids are typically designed and operated with a looped structure. Loops provide redundancy to the system in case of component failure and are beneficial to the network pressure profiles.

After the creation of a spanning tree, further edges are iteratively added to the synthetic graph for the establishment of loops. A novel approach is adopted in this stage: the extra edges must contribute to an efficient improvement of the connectivity of the growing network, providing shortcuts among regions that are topologically distant and spatially close. For this purpose, couples of nodes i and j are described in terms of mutual physical (Euclidean) distances $\left\| \mathbf{p}_i - \mathbf{p}_j \right\|$ and mutual topological distances $\mathbf{D}_{i,j}$ (node-to-node path length). Accordingly, at each iteration the network structure is integrated with a new edge among the couple of nodes featuring the highest $\frac{\mathbf{D}_{i,j}}{\left\| \mathbf{p}_i - \mathbf{p}_j \right\|}$ ratio.

The number of nodes *N* of the synthetic network is imposed to equal the reference grid. The number of edges must be at least equal to the number of nodes minus one to ensure a full connectivity through a spanning tree. The number of loops, named L in the diagram of Figure 1, can be randomly chosen based on ranges observed in real-world networks. An example is given in Section V.

The probabilistic approach that is adopted along the procedure ensures that each execution of the algorithm leads to a different and unique graph structure. Accordingly, in this stage the proposed tool can be deployed to generate a virtually infinite number of distribution gas network topologies, starting from a reference input grid.

The evaluation of the tool provided in [36] evidences that the generated topologies feature consistent properties compared to real-world case studies. Among these properties, it is shown that the grids feature realistic degree distributions, line lengths distributions, average path lengths and clustering coefficients. For sake of conciseness, these results are not replicated in the present work.

III. Technical sizing of distribution gas grids

Gas network topologies established as from the previous methodology require to be assigned with technical specifications, to constitute finished case studies readily exploitable for simulation applications. For this purpose, this section illustrates a systematic procedure to accomplish a technical sizing of the generated networks, satisfying arbitrary operational restrictions.

Gas grids are typically dimensioned to constrain their working conditions within target boundaries of operational states. In real-world applications, the hydraulic design of distribution gas systems needs to comply with a design minimum pressure (DMP) and a maximum gas flow velocity [37]. The DMP is the minimum pressure level to be guaranteed across the whole infrastructure to ensure the safe and correct operation of customer appliances, service regulators and intermediate pressure reduction stations. The maximum gas flow velocity constitutes a further constraint to prevent excessive mechanical stress, noise, dragging of impurities and corrosion of pipelines [37,38]. The above design parameters can differ among pressure tiers of the networks and national standards. As a matter of example, for medium-pressure (MP) networks Italian standards recommend maximum gas velocities ranging between 10 m/s to 25 m/s, depending on the pressure tier. For reference, pressure tiers comprised within the MP class vary between 0.04-0.5 bar_g ("6^{th} species"), 0.5-1.5 bar_g ("5^{th} species"), and 1.5-5 bar_g ("4^{th} species"), as stated in the national Network Code [39]. The technical sizing performed by the proposed tool replicates the typical design criteria of above. Other design considerations of real-world gas distribution systems (namely related to safety and regulation aspects) are not addressed by the model, as out of scope for the current applications.

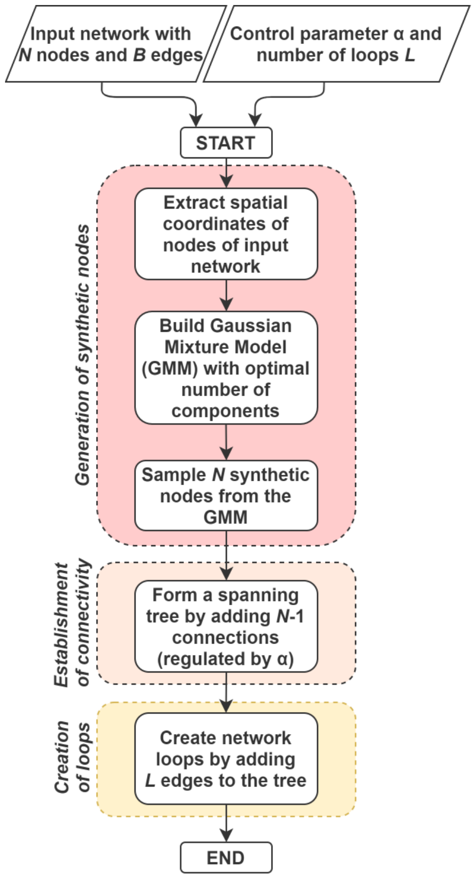

Figure 1: High-level diagram of the algorithm designed for the generation of synthetic network topologies.

Table 1: Inputs required by the tool for the technical design of distribution gas networks.

Inputs	Unit	Description
Nominal source pressure, p_{nom}	bar_g	Gas pressure at the supply node at design conditions
Design minimum pressure, p_{min}	bar_g	Minimum pressure to be guaranteed at all the nodes of the network
Maximum flow velocity, v_{max}	m/s	Maximum admitted velocity of gas flow in the pipelines
Location of gas supply point	-	Index of the network node which withdraws gas from a source and supplies it to the grid at the nominal source pressure p_{nom}
Nodal gas consumptions at design conditions	MW	Design gas loads (in thermal power) at the consumption nodes – convertible into Sm^3/h with known NG composition
NG composition[1]	-	21-component molar composition of NG, needed to compute gas properties (i.e., higher heating value, density, viscosity, compressibility factor, etc.)
Ground temperature	K	Temperature of the ground – assumed in thermal equilibrium with the temperature of the gas across the network
Roughness coefficient	µm	Internal roughness of the pipelines – based on the material

Although the model has been designed to offer a wider flexibility, its description and application in the present work is restricted to infrastructures operated at a single pressure tier and supplied by a single source node. The tool accomplishes the technical design of input network structures according to a set of custom parameters, that are listed in Table 1. Its major purpose is assigning the pipelines of the input network with suitable diameter sizes picked from a discrete set of off-the-shelf commercial sizes. The procedure ensures that, at design conditions, all the nodes are characterised by acceptable pressure values ($p \geq p_{min}$) and all the pipelines feature admitted velocities ($v \leq v_{max}$), when the gas supply at the source node occurs at the nominal pressure p_{nom}.

Nodal gas consumptions at design conditions (Table 1) represent the maximum values of gas demand expected at the network nodes. These values can also embed safety margins accounting for load variability issues, as well as for possible future system expansions.

The procedure is detailed in Figure 2 and consists in an iterative process in which, at each iteration, pipelines serving directly or indirectly one or more critical nodes (i.e., nodes for which $p < p_{min}$) are enlarged. The diameters of the pipelines are initialised imposing a flow velocity equal to the maximum design value v_{max}, which ensures that any subsequent enlargement of the pipelines leads to acceptable gas velocities. At each iteration, a steady-state fluid-dynamic test identifies the node pressures that are still critical. The process stops when all the nodes have acceptable pressure levels. For meshed networks, each iteration firstly resizes a spanning tree extracted from the whole grid, while the

[1] The tool supports GERG-88 (as detailed in [41]) and GERG-2008 [42] virial equations of state (EOS) for the evaluation of the state variables and properties of NG, including the compressibility factor. The EOS are deployed in the module for the hydraulic verification of the network, and they are characterised by different complexities and computational performances. Both treat NG as a mixture of gases with given composition. GERG-88 EOS requires indicating up to 13 components for NG, while GERG-2008 requires 21 components. Because of its more efficient implementation, GERG-88 has been used in the execution of this work. For less complex cases and/or for applications demanding a higher accuracy, GERG-2008 can be readily deployed instead.

remaining unsized pipelines are later assigned with average diameter sizes computed within their respective loop.

Pipelines to be enlarged along the design procedure are identified with two main conditions, highlighted in the diagram of Figure 2:

1. a pipeline cannot be assigned with a diameter size that is larger than the diameter of its predecessor pipeline (i.e., the diameters of the pipes decrease from the supply node down to the terminals of the network). This rule is inspired from observations on real-world distribution gas systems, where most of the source-sink network paths have decreasing diameter values.
2. No more than the strictly necessary number of pipelines is resized at one iteration: if enlarging two or more pipelines causes multiple beneficial effects on some downstream nodes, then only one of the pipelines is resized. This is done to avoid overlapping of effects which may incur risks of system oversizing.

At each iteration, the fluid-dynamic verification of the network is carried out via a custom isothermal and stationary model that is fully described and validated in [40].

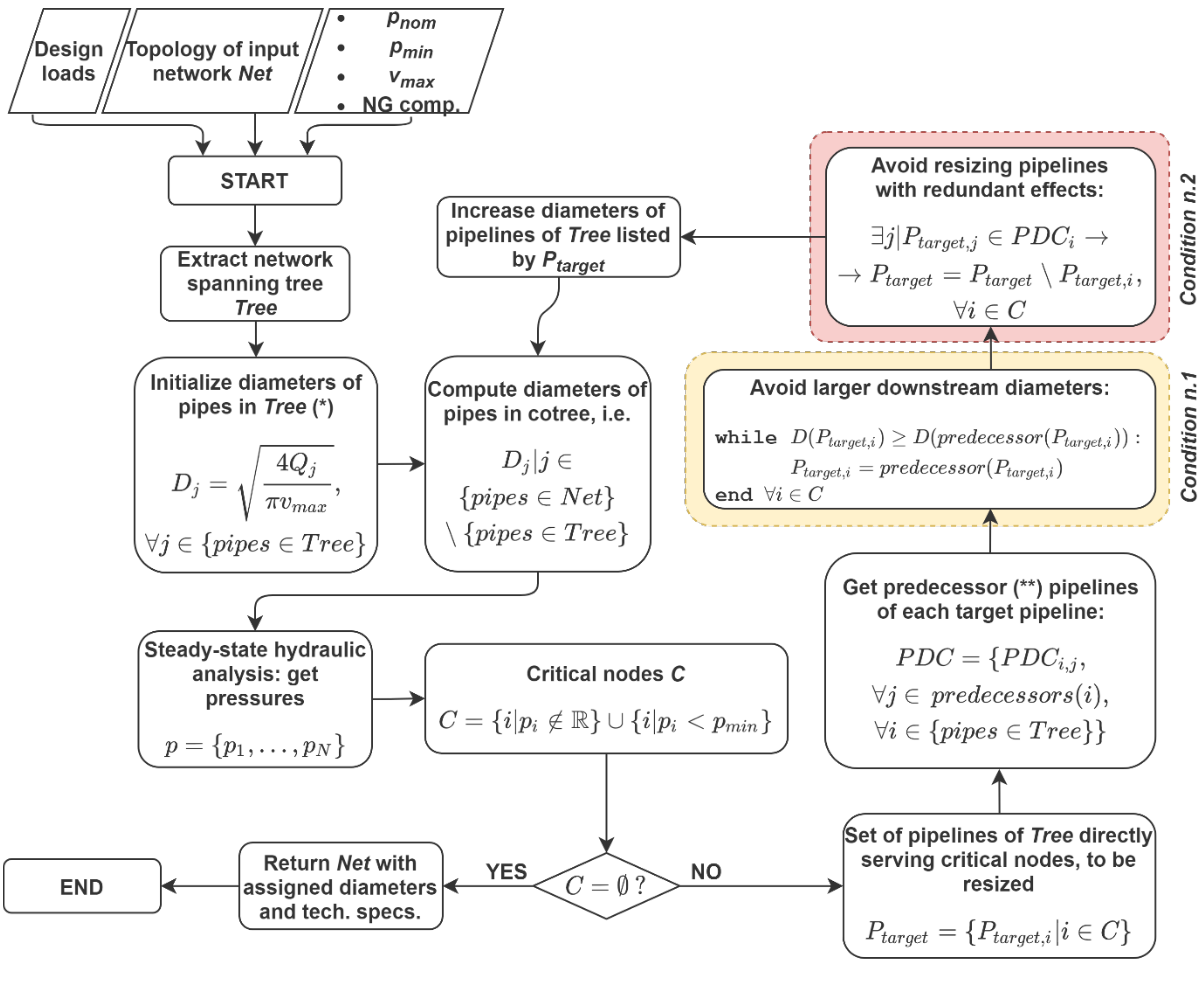

Figure 2: Algorithm accomplishing the technical design of distribution gas networks with given topology and target operational restrictions. The algorithm iteratively identifies and increases the diameters of the pipelines serving those critical

nodes that do not comply with operational restrictions. Conditions imposed for the identification of the pipelines are highlighted. Condition n.1 ensures that diameter values decay from the source node down to the sink nodes. Condition n.2 ensures that only the strictly necessary number of pipes are resized at each iteration, avoiding the risk of system oversizing.
(*) Q_j represents the volume flow rate of gas [m^3/s] streaming in the j-th pipe of the tree.
(**) The term "predecessors" of a pipeline is here used in a non-rigorous way to indicate the set of pipes that precede the given pipeline along the source-sink paths of the spanning tree "Tree". The "predecessor" (singular) of a pipeline is here intended as the pipeline directly preceding the given pipe.

IV. Generation of synthetic case studies

The technical designer presented in the previous section and the topology generator introduced in Section II have been sequentially combined in a MATLAB implementation to produce finished synthetic models of gas distribution grids. Taking advantage of the probabilistic nature of the presented tools and of the overall short execution times (around 1 minute per network), a virtually infinite number of different synthetic networks can be generated.

Accordingly, ten-thousand synthetic distribution grids have been created using as a reference a real-world MP distribution grid. The reference MP network is located in Italy and serves approximately 3400 delivery points (connected to both medium and low-pressure). Its total extension is 34 km, and its maximum total capacity is estimated at 47 MW, although the system may have been oversized to meet an actually lower gas demand. The model of the infrastructure, represented in Figure 3, comprises 373 nodes and 375 pipelines (edges) – and therefore features 3 loops. According to the Italian network code, the network is classified as of "4th species", with admitted pressure values ranging between 1.5 bar_g and 5 bar_g [39]. As illustrated below, these values are respectively adopted as design minimum pressure and design source pressure in the synthetisation process.

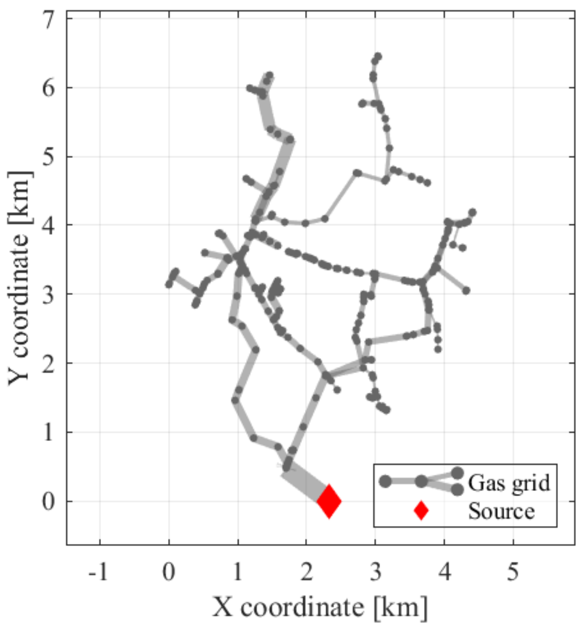

Figure 3: Real-world MP distribution grid used as reference. The width of the edges (pipelines) of the graph is proportional to the diameter values. For comparison: the largest pipeline (connected to the source node) has a diameter of 0.25 m.

The procedure leading to the creation of a single synthetic network is composed of the following stages, which are replicated for each of the 10,000 synthetised grids:

1. *Establishment of the spatial distribution and topology of the grid*: the synthetic nodes are randomly positioned on a two-dimensional plane, replicating the spatial distribution of the nodes in the reference grid (see grey dots in Figure 3); afterwards, the topology of the network is established by connecting the nodes. The methodology provides the output grid with correct topological and spatial properties and it is illustrated in Section II.
2. *Identification of the source node*: one source node is selected among all the nodes in the network having degree equal to 1 (i.e., nodes having only one connection).
3. *Identification of consumption nodes*: loads are identified in correspondence of all the nodes with degree equal to 1, excluding the source node.
4. *Assignment of load values to consumption nodes*: a total load equal to 47 MW (equal to the estimated reference network capacity) is distributed among the consumption nodes. Different allocation schemes are proposed in the next sections.
5. *Technical design of the grid*: network pipelines are sized and assigned with suitable technical properties, as extensively described in Section III. In this stage, the nominal source pressure is assumed to be equal to 5 bar_g (upper bound of the admitted range of pressures for the network class), while the design minimum pressure is taken to be 1.5 bar_g (lower bound). A conservative value of 20 m/s is adopted as maximum admitted velocity in the pipelines [38]. A complete list of the inputs adopted in this stage is given in Table 2.

Table 2: Design parameters used for the technical sizing of the synthetic networks.

Inputs	**Unit**
Nominal source pressure, p_{nom}	5 bar_g
Minimum design pressure, p_{min}	1.5 bar_g
Maximum flow velocity, v_{max}	20 m/s
Ground temperature	283.15 K
Roughness coefficient	140 µm
NG properties (based on composition)	
– Higher heating value	51.3 MJ/kg
– Specific gravity	0.64

To consider and replicate the variability of real-world distribution systems, networks are firstly generated with several degrees of freedom (“Case A”). In a second stage, a set of topological, spatial and technical constraints is applied to gain a higher control over the properties of the output networks (“Case B”). It is shown that the properties and behaviours of the obtained networks span within significant ranges of values, especially when no constraints are applied to the generated systems. In a subsequent section (“Case C”) an a-posteriori enforcement is proposed to demonstrate the possibility of selecting a reduced number of case studies with target properties from the large set of synthetised grids.

V. Case A: Constraint-free generation of synthetic networks

Ten thousand networks are generated in a constraint-free fashion. The procedure acknowledges the differences that may exist among distribution networks in real-world applications. The characteristics and the response of the distribution systems can be highly affected by the location of the supply node, the magnitude of the loads and their distance from the supply, and the presence of loops in the system. Accordingly, the following assumptions are formulated:

1. The number of loops in the grids varies randomly according to ranges of network cyclicity (here meant as ratio of number of loops over total network length) observed in real MP systems, which span between 0 and 0.29 km^{-1} [43].
2. The supply node can be located in any region of the network, and it is randomly identified in one node out of all the nodes with degree equal to 1 (i.e. nodes with only 1 connection).
3. Nodal gas loads are randomly established by distributing the total design load (47 MW) according to a suitable statistical model (Weibull distribution is here adopted).

A sample of 10 networks generated in this way is illustrated in Figure 4. As expected, the grids feature a spatial extension and density similar to the reference network of Figure 3. It can be observed how the location of the source nodes (in red) changes in each case study, as well as the number of loops. Due to the constraint-free allocation of the load values, the patterns of the gas loads against their distances from the supply point are noticeably variable (see Figure 5). Consequently, while in some networks nearly all the load is within 4 km from the source, in other cases it is concentrated in remote regions of the grid.

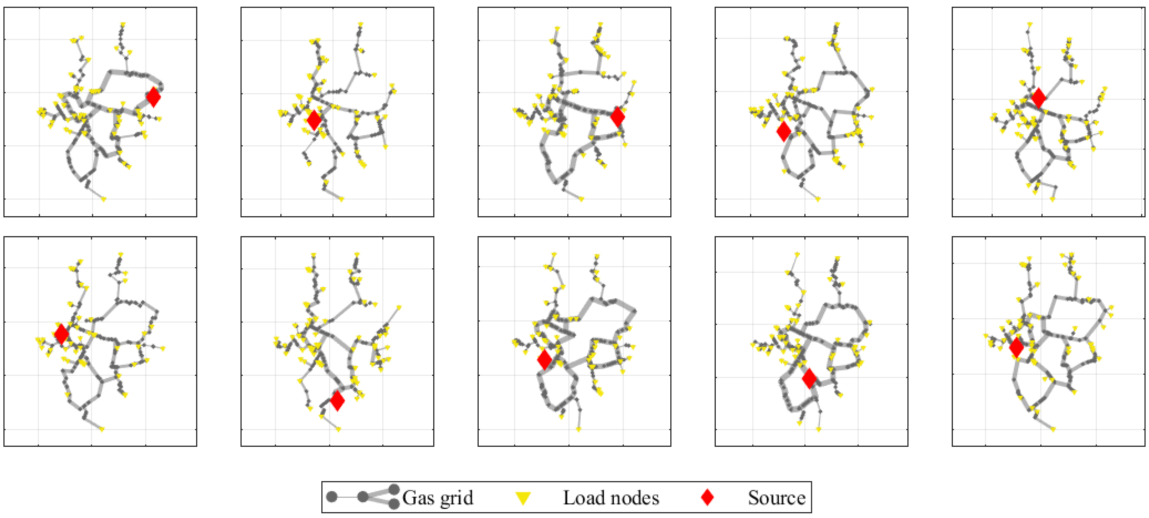

Figure 4: Ten random samples out of the 10,000 networks generated in the constraint-free synthetization process.

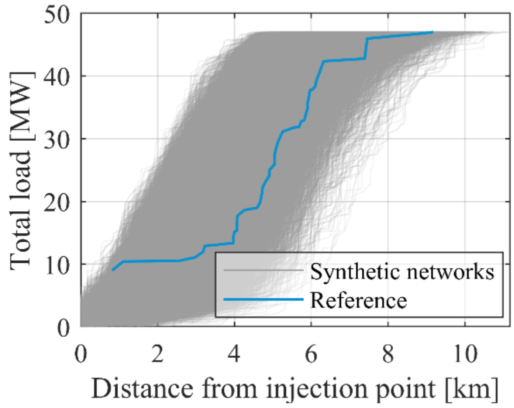

Figure 5: Cumulative gas loads against their distance from the supply node. The chart includes the information for the real-world network used as reference, for comparison.

Due to the many degrees of freedom discussed above, the characteristics of the networks feature a substantial variability. A statistical approach is therefore adopted in their representations in Figure 6, where attention is given to diameter sizes (A), network volumes (B), nodal pressures (C) and gas flow velocities in pipes (D) at steady-state design conditions. In charts A, C and D the network properties are described by cumulative distribution functions (CDF) displayed in grey. Each grey CDF refers to a single network, so that the resulting grey area indicates the band of values covered by the different networks. Blue boxplots are also included to highlight the variability of the CDF values at discrete intervals of the observed property. Thick blue rectangles of the boxplots cover the 2nd and 3rd percentiles of records, while white and black dots indicate the median values. Whiskers reach the minimum and maximum values of the observations. For graphical reasons, the bins used for the boxplots are different (wider) than the bins used for the CDF.

The synthetisation process produced networks with realistic structural properties and compliant hydraulic behaviours. Pipeline diameters are within a sensible range of sizes, varying between 2.5 and 16.0 cm (chart A in Figure 6). The band of their distributions is significantly wide, being the maximum recorded two-sample Kolmogorov-Smirnov (KS) statistic (KSstat2) equal to 0.97[2]. This aspect influences the distribution of the volumes as well – see chart B – whose standard deviation amounts to 49.6 m^3 against a mean value of 154.9 m^3.

The charts C and D of Figure 6 evidence that the algorithm has correctly sized all the 10,000 networks. Complying with the imposed design criteria, pressures are never lower than 1.5 bar$_g$ and velocities never exceed 20 m/s. The average pressure of the networks is 3.01 bar$_g$. In 60% of the cases, most of the nodal pressures are comprised within 1.5 and 3 bar$_g$. As evidenced by the grey band of CDF and the blue boxplots of chart C, pressure distributions are remarkably variable, being the maximum recorded KS statistic equal to 0.97. The mean gas flow velocity in pipes amounts to 6.8 m/s and, on average, 50.1%

[2] The two-sample KS statistic measures the fit between two empirical distributions. It is defined as the maximum absolute difference between the cumulative distributions of the two empirical distributions in question. Its value can span between 0 (high correlation) and 1 (low correlation).

of values are lower than 6 m/s. Only 6.1% of the velocities is higher than 15 m/s, indicating that most of the pipelines operate well below the imposed limits (20 m/s).

Results highlight that distribution networks designed with identical sizing criteria can be characterised by substantially dissimilar structures and hydraulic behaviours. These observations suggest that different impacts should be expected when alternative gases are to be injected into these systems. Accordingly, statistical-based approaches may offer a more suitable solution to capture this information, that would be otherwise lost in case-specific studies.

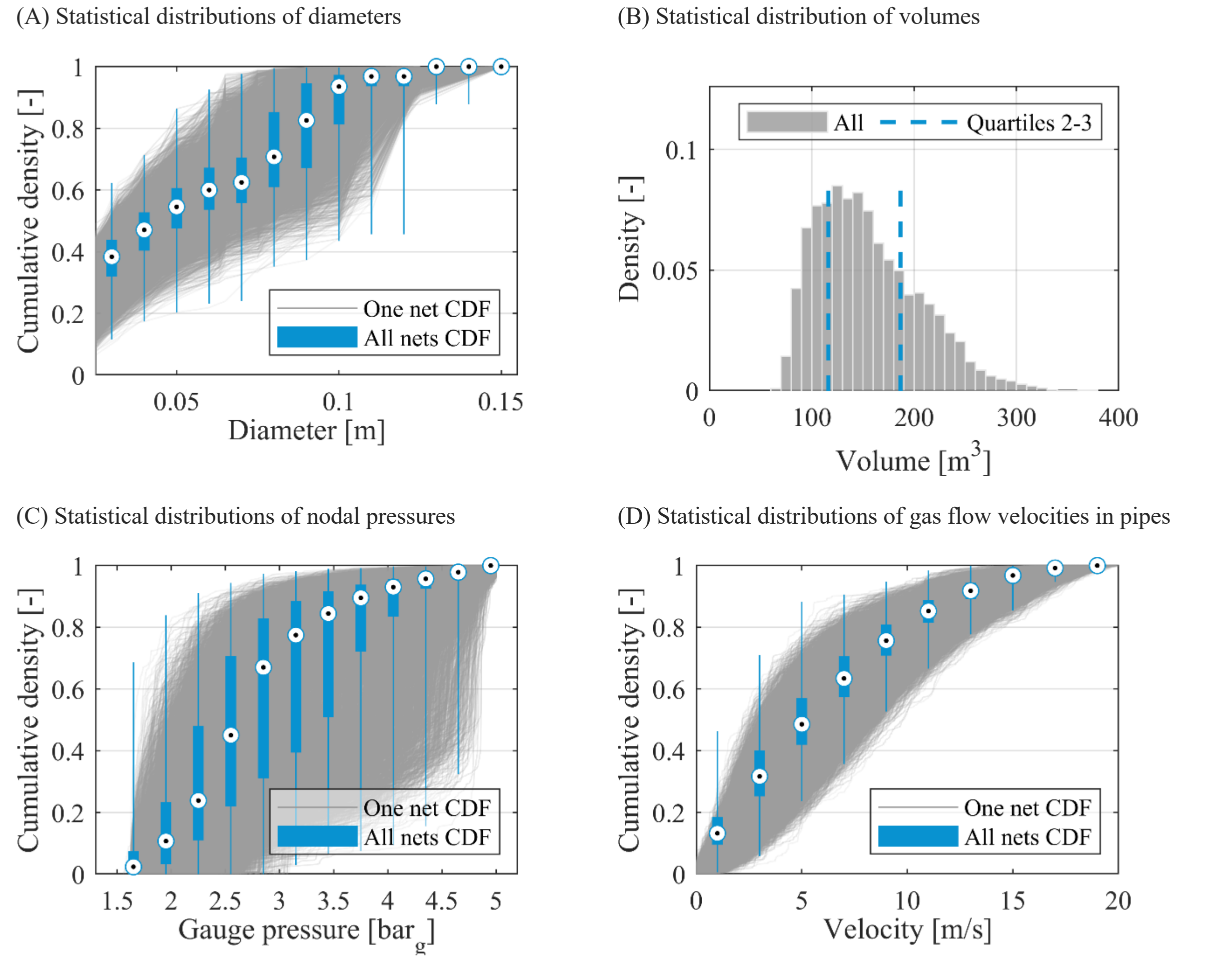

Figure 6: Structural properties and fluid-dynamic response of synthetic networks generated by the algorithm in the constraint-free process. Cumulative distribution functions (CDF) of diameters, pressures and velocities are illustrated as overlaid histograms. Boxplots are added for a complete description of the CDF values.

VI. Case B: Generation of synthetic networks with technical and topological constraints

In a second stage, a set of constraints has been exerted to the synthetisation process to gain a higher control over the configuration of the output networks. Overall, the steps leading to the synthetisation of the networks are unvaried from the previous case. Given that the overall gas demand and the design parameters are unchanged, the restrictions are applied over those remaining factors that mostly affect the properties of distribution grids. In particular, attention has been given to the location of the supply node, the allocation of gas consumptions and the number of network loops. Accordingly, the following input constraints are included:

1. The number of loops is the same in all the generated networks and it is equal to 3, as in the reference grid.
2. The location of the supply point of the synthetised networks is fixed and it is found in correspondence of the closest node to a given reference (the position of the source in the reference grid is used – see red marker in Figure 3).
3. The sum of the loads weighted by their distance from the supply node is similar in all the networks, and it is inspired from the reference real-world grid.

The configurations of the networks generated in this way exhibit limited variations. This behaviour can be qualitatively observed from Figure 7, which illustrates ten random samples extracted from the 10,000 synthetic networks. What emerges from the illustration is that the networks feature similar topology, spatial extension, and point of supply position. Also the diameter sizes (widths of the lines in the plot) follow comparable spatial patterns in all the grids. The effect of the constraint imposed over the allocation of gas loads is noticeable in Figure 8: load patterns are considerably more uniform with respect to the previous case (see Figure 5), and consistently replicate the adopted reference.

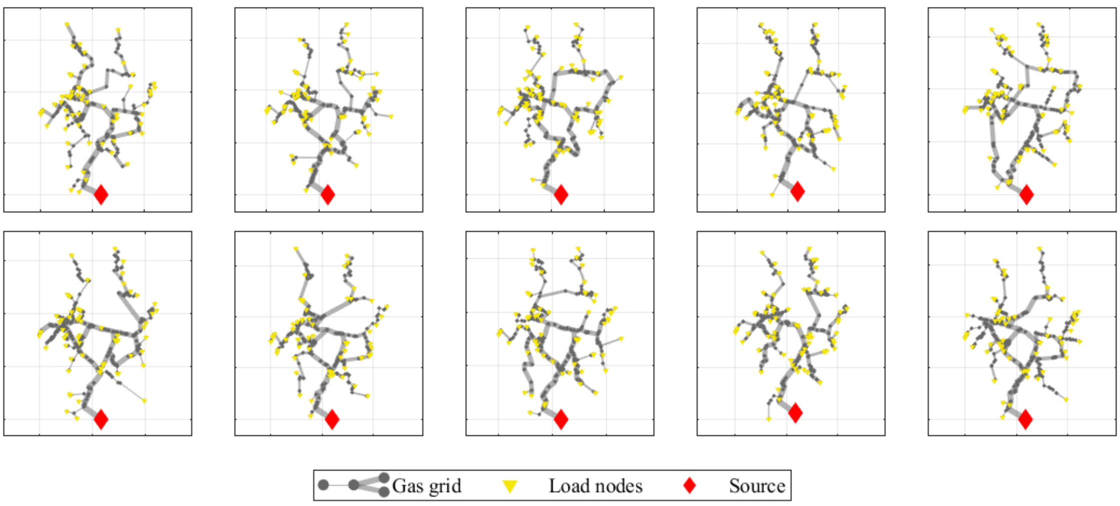

Figure 7: Ten random samples out of the 10,000 networks generated in the constrained synthetization process.

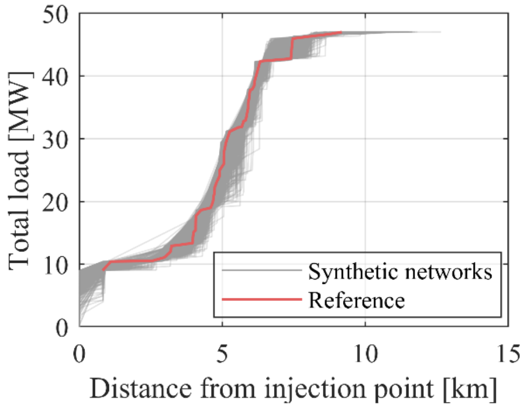

Figure 8: Cumulative gas loads against their distance from the supply node. The patterns of gas consumption within the synthetic networks are constrained to follow the behaviour of the reference real-world gas grid, as visible from the fine agreement of the grey curves and the reference red line.

As an effect of the application of the constraints, the properties of the networks – displayed in Figure 9 – feature a substantially reduced variability, providing evidence on the consistency of the tool. The maximum KS statistic recorded among the diameter distributions decreased from 0.97 to 0.52, while the standard deviation of network volumes was reduced from 49.6 m^3 to 19.0 m^3 (see charts A and B). The average volume amounts to 179.5 m^3, being higher than in Case A. In fact, because of the imposed peripherical location of the supply node, the algorithm deploys more frequently large pipelines with higher capacity and lower resistance to flow.

Average values of nodal pressures and gas flow velocities respectively decreased to 2.55 bar_g and 4.8 m/s. In virtually all the cases, most of the values are comprised between 1.5 and 3 bar_g. An appreciably higher uniformity characterises the pressure distributions of chart C, indicating that the applied constraints have proven effective. Nevertheless, non-negligible variations are still found for pressures lower than 3 bar_g, indicating that partially diverse hydraulic responses may be obtained even among very similar distribution systems. This aspect emphasises the value of statistical-based approaches in the study of gas grids.

Despite a significantly wide grey band, red boxplots in chart D indicate that most of gas flow velocities feature limited variations from their median values. Overall, the average gas velocity in pipelines is 4.8 m/s and in virtually all the networks most of the velocities are below 6 m/s. Velocities higher than 15 m/s represent only 5.5% of cases on average.

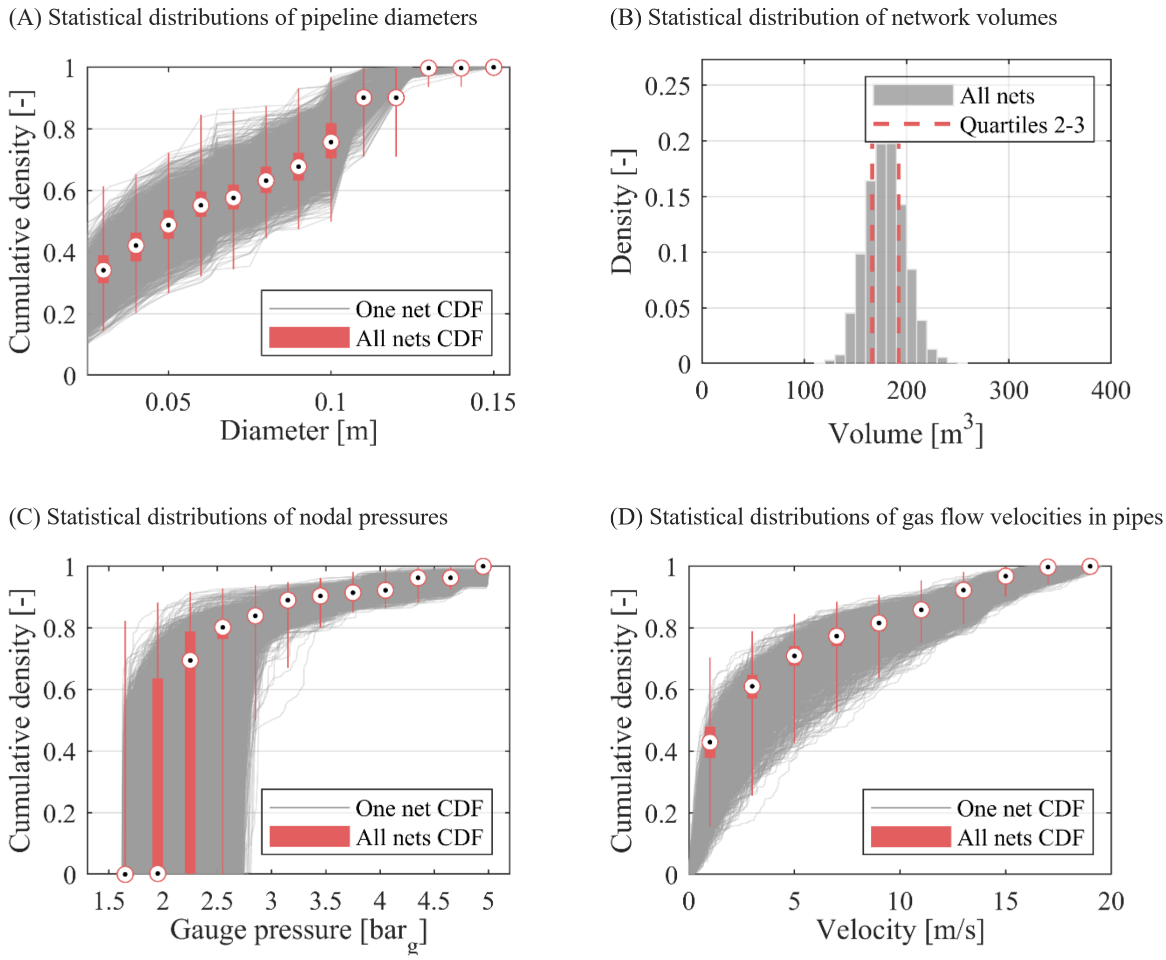

Figure 9: Structural properties and fluid-dynamic response of synthetic networks generated by the algorithm in the constrained process. Cumulative distributions functions (CDF) of diameters, pressures and velocities are illustrated as overlaid histograms. Boxplots are added for a complete description of the CDF values.

The tool demonstrates again to generate networks with consistent structures and correct technical designs. Under the applied constraints, the structural and hydraulic variations of the grids occur within significantly reduced, but in part non-negligible, ranges of values.

VII. Case C: A-posteriori selection of the synthetic networks

It has been observed that the proposed tool can produce large quantities of unique ready-made models of distribution gas grids. The synthetic networks can be characterised by many degrees of freedom (see Case A); alternatively, suitable a-priori constraints can be applied for a higher control over their properties (see Case B). As already argued, having a large set of networks with diverse properties is particularly suitable for statistical-based studies. However, in other specific applications it may be required that the generated grids feature peculiar structural and/or hydraulic properties. For these cases it is possible to perform an apposite a-posteriori selection of the networks based on the target properties of interest. The selection can be carried out excluding all those networks whose features are too different from a target reference. A graphical example is provided in Figure 10 and Figure 11, where it is assumed that the networks generated according to the procedure of Case A require featuring pressure distributions similar to the real-world reference grid (previously depicted in Figure 3). The correlation of the 10,000 synthetic networks and the reference real-world grid is assessed via the two-sample KS statistic (see Figure 10). Subsequently, only those 468 grids featuring KSstat2 values lower than 0.2 are kept to meet the similarity requirement (see Figure 11).

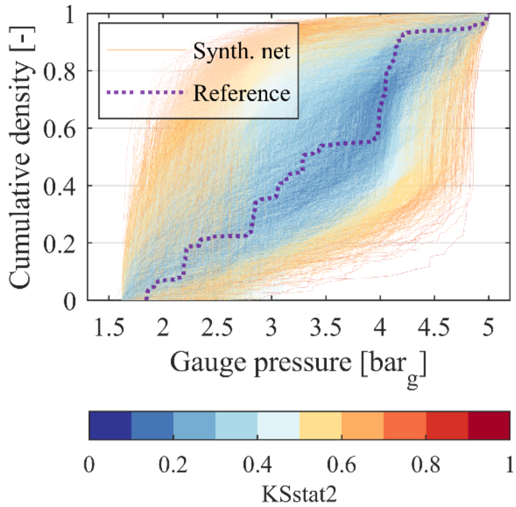

Figure 10: Statistical distribution (CDF) of pressures in the 10,000 synthetic networks generated in Case A (see panel C in Figure 6). Colours indicate the fit between the curves and the target reference, expressed by the two-sample KS statistic (KSstat2).

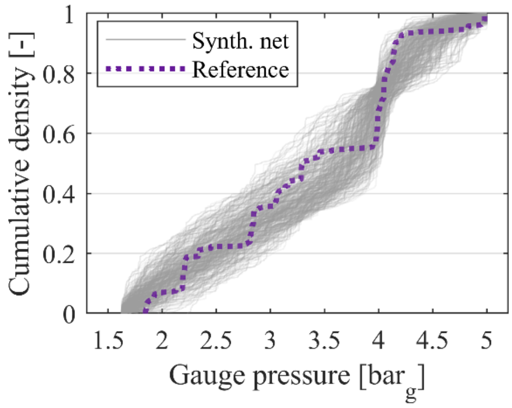

Figure 11: Resulting selection of the networks featuring the most similar pressure distributions to the target reference (KSstat2 ≤ 0.2). See Figure 10 for comparison.

Since real-world infrastructure models are often protected by non-disclosure agreements, we add that the capability of the tool illustrated above can be furtherly exploited to produce anonymised twins of the real-world network. Depending on the targeted property, the anonymised twin can ensure equivalent structural characteristics or hydraulic performances (or a combination of them) and can be disclosed without incurring in industrial secrecy, privacy and security issues.

VIII. Discussion and remarks

Previous results have provided a validation of the gas grid generator and a characterisation for networks featuring consistent topologies, spatial extensions, and equal design criteria. Substantial variations concerning the properties of the generated networks have been observed.

The structure of the grids and their fluid-dynamic responses span within significant ranges of values when many degrees of freedom are introduced in the synthetisation process, highlighting the variety of cases that may be found in real-world network systems. The subsequent application of constraints leads to reduced variations among the properties of the generated networks, evidencing the consistency of the proposed algorithm.

It has been shown that the generated grids never violate the required design criteria, in terms of minimum design pressure and maximum flow velocity. In the analysed cases, at most 6.1% of pipelines feature gas velocities higher than 15 m/s on average. Given that v_{max} is assumed 20 m/s, most networks can therefore adapt to changes in their operations (e.g., distributed injection of alternative fuels) with a low risk of violating maximum velocity constraints. At design conditions, the analysed grids feature a considerable number of nodes with low pressures, in the range of the DMP (1.5 bar_g) and 3 bar_g. This aspect indicates a potential vulnerability of the networks in cases of non-conventional operational conditions, where the deployment of alternative gas blends may give rise to violations of the DMP. Practices of system oversizing, which are rather common in the design of gas distribution systems, may increase the actual pressure of the networks at design conditions.

The distributions of pressures in the generated systems exhibit a significant variability, that is even appreciable in the constrained Case B for values in the range of 1.5-3.0 bar_g. This behaviour clearly evidences a limit of traditional studies around gas grids: the responses of similar infrastructures under identical simulation setups can potentially be very different, and traditional case-specific approaches, contrarily to statistical-based studies, cannot capture these deviations. As a last remark, it can be inferred from the observed network volumes and pressure profiles that the analysed grids offer variable capabilities of storing gas, being therefore characterised by different linepacks[3] at design conditions. Consequently, a statistical perspective may also offer a valuable support in analyses involving dynamic responses of the networks, as well as their utilisation as storage options for green hydrogen produced by intermittent electrical energy sources.

Ultimately, the synthetic gas network generator has proven to be functional for further applications. In Case C a fine selection of networks with arbitrary target properties has been proposed, demonstrating the suitability of the tool for the creation of anonymised twins of sensible networks, with equivalent structural and/or hydraulic properties.

IX. Conclusions

The work proposed the utilisation of an algorithm for the synthetic generation of distribution gas network models. The gas network generator has proved to generate a virtually infinite number of networks with similar topological, spatial, and technical properties. Accordingly, its suitability for the execution of statistical studies on gas networks, as well as for the anonymisation of sensible network data, has been discussed and demonstrated. The tool can be readily deployed with custom design parameters, and the generated networks have demonstrated to always satisfy custom target design requirements, including the design minimum pressure and maximum flow velocities in pipes. Two different experiments have proposed the synthetisation of ten thousand networks and their structural and fluid-dynamic analyses. The tests proved the accuracy and the consistency of the tool and highlighted significant variations in the properties of the generated networks. The results provide evidence on the added value offered by the proposed tool: in opposition to traditional case-specific

[3] The term "linepack" refers to the total amount of gas stored within a gas network in compressed form. It can be measured in mass or energy content of the gas.

investigations, the network generator enables statistical studies on gas networks carried out over a large basis of synthetic and realistic grids, accounting for the variety of responses found in real-world systems. In perspective applications, the tool is to be deployed to assess the injection of renewable gases in large numbers of synthetic grids, to provide a broader, generalised understanding of the performance of distribution systems under forthcoming operational schemes.

References

[1] Mancarella P. MES (multi-energy systems): An overview of concepts and evaluation models. Energy 2014;65:1–17. https://doi.org/10.1016/j.energy.2013.10.041.

[2] European Commission. Powering a climate-neutral economy: An EU Strategy for Energy System Integration - COM(2020) 299 final. Brussels: 2020.

[3] Davis SJ, Lewis NS, Shaner M, Aggarwal S, Arent D, Azevedo IL, et al. Net-zero emissions energy systems. Science (80-) 2018;360. https://doi.org/10.1126/science.aas9793.

[4] Quarton CJ, Samsatli S. Power-to-gas for injection into the gas grid: What can we learn from real-life projects, economic assessments and systems modelling? Renew Sustain Energy Rev 2018;98:302–16. https://doi.org/10.1016/j.rser.2018.09.007.

[5] Thema M, Bauer F, Sterner M. Power-to-Gas: Electrolysis and methanation status review. Renew Sustain Energy Rev 2019;112:775–87. https://doi.org/10.1016/j.rser.2019.06.030.

[6] European Commission. A hydrogen strategy for a climate-neutral Europe - COM(2020) 301 final. Brussels: 2020.

[7] Quarton CJ, Samsatli S. Should we inject hydrogen into gas grids? Practicalities and whole-system value chain optimisation. Appl Energy 2020;275:115172. https://doi.org/10.1016/j.apenergy.2020.115172.

[8] Hydrogen Projects Database – Analysis - IEA n.d. https://www.iea.org/reports/hydrogen-projects-database (accessed January 11, 2021).

[9] Hydrogen Europe. Hydrogen in the EU's Economic Recovery Plans. 2020.

[10] Abeysekera M, Wu J, Jenkins N, Rees M. Steady state analysis of gas networks with distributed injection of alternative gas. Appl Energy 2016;164:991–1002. https://doi.org/10.1016/j.apenergy.2015.05.099.

[11] Gondal IA. Hydrogen integration in power-to-gas networks. Int J Hydrogen Energy 2019;44:1803–15. https://doi.org/10.1016/j.ijhydene.2018.11.164.

[12] Cavana M, Leone P. Biogas blending into the gas grid of a small municipality for the decarbonization of the heating sector. Biomass and Bioenergy 2019;127:105295. https://doi.org/10.1016/j.biombioe.2019.105295.

[13] Pellegrino S, Lanzini A, Leone P. Greening the gas network – The need for modelling the distributed injection of alternative fuels. Renew Sustain Energy Rev 2017;70:266–86. https://doi.org/10.1016/j.rser.2016.11.243.

[14] Guandalini G, Colbertaldo P, Campanari S. Dynamic modeling of natural gas quality within transport pipelines in presence of hydrogen injections. Appl Energy 2017;185:1712–23. https://doi.org/10.1016/j.apenergy.2016.03.006.

[15] Fan D, Gong J, Zhang S, Shi G, Kang Q, Xiao Y, et al. A transient composition tracking method for natural gas pipe networks. Energy 2021;215:119131. https://doi.org/10.1016/j.energy.2020.119131.

[16] Cavalieri F. Steady-state flow computation in gas distribution networks with multiple pressure levels. Energy 2017;121:781–91. https://doi.org/10.1016/j.energy.2017.01.062.

[17] Estermann T, Newborough M, Sterner M. Power-to-gas systems for absorbing excess solar power in electricity distribution networks. Int J Hydrogen Energy 2016;41:13950–9. https://doi.org/10.1016/j.ijhydene.2016.05.278.

[18] Mukherjee U, Elsholkami M, Walker S, Fowler M, Elkamel A, Hajimiragha A. Optimal sizing of an electrolytic hydrogen production system using an existing natural gas infrastructure. Int J Hydrogen Energy 2015;40:9760–72. https://doi.org/10.1016/j.ijhydene.2015.05.102.

[19] Resources | PES Test Feeder n.d. https://site.ieee.org/pes-testfeeders/resources/ (accessed January 11, 2021).

[20] Power Systems Test Case Archive - UWEE n.d. https://labs.ece.uw.edu/pstca/ (accessed January 11, 2021).

[21] Postigo Marcos F, Mateo Domingo C, Gómez San Román T, Palmintier B, Hodge B-M, Krishnan V, et al. A Review of Power Distribution Test Feeders in the United States and the Need for Synthetic Representative Networks. Energies 2017;10:1896. https://doi.org/10.3390/en10111896.

[22] Grzanic, Flammini, Prettico. Distribution Network Model Platform: A First Case Study. Energies 2019;12:4079. https://doi.org/10.3390/en12214079.

[23] Gershenson D, Rohrer B, Lerner A. Predictive model for accurate electrical grid mapping 2019. https://engineering.fb.com/2019/01/25/connectivity/electrical-grid-mapping/ (accessed January 11, 2021).

[24] Lichtinghagen J, Sieberichs M, Moser A, Kübler A. Medium voltage network planning considering the current network and geographical restrictions. 2017 6th Int. Conf. Clean Electr. Power Renew. Energy Resour. Impact, ICCEP 2017, Institute of Electrical and Electronics Engineers Inc.; 2017, p. 689–93. https://doi.org/10.1109/ICCEP.2017.8004765.

[25] Soltan S, Loh A, Zussman G. A learning-based method for generating synthetic power grids. IEEE Syst J 2019;13:625–34. https://doi.org/10.1109/JSYST.2018.2825785.

[26] Birchfield AB, Xu T, Gegner KM, Shetye KS, Overbye TJ. Grid Structural Characteristics as Validation Criteria for Synthetic Networks. IEEE Trans Power Syst 2017;32:3258–65. https://doi.org/10.1109/TPWRS.2016.2616385.

[27] Almeida D, Abeysinghe S, Ekanayake MP, Godaliyadda RI, Ekanayake J, Pasupuleti J. Generalized approach to assess and characterise the impact of solar PV on LV networks. Int J Electr Power Energy Syst 2020;121:106058. https://doi.org/10.1016/j.ijepes.2020.106058.

[28] Abeysinghe S. A statistical assessment tool for electricity distribution networks. Cardiff University, 2018.

[29] Schweitzer E, Scaglione A, Monti A, Pagani GA. Automated Generation Algorithm for Synthetic Medium Voltage Radial Distribution Systems. IEEE J Emerg Sel Top Circuits Syst 2017;7:271–84. https://doi.org/10.1109/JETCAS.2017.2682934.

[30] Sadeghian H, Wang Z. AutoSynGrid: A MATLAB-based toolkit for automatic generation of synthetic power grids. Int J Electr Power Energy Syst 2020;118:105757. https://doi.org/10.1016/j.ijepes.2019.105757.

[31] Mikolajková M, Haikarainen C, Saxén H, Pettersson F. Optimization of a natural gas distribution network

with potential future extensions. Energy 2017;125:848–59. https://doi.org/10.1016/j.energy.2016.11.090.

[32] Torkinejad M, Mahdavi I, Mahdavi Amiri N, Seyed Esfahani M. Topology Design and Component Selection in an Urban Gas Network: Simultaneous Optimization Approach. J Pipeline Syst Eng Pract 2019;10:04018035. https://doi.org/10.1061/(asce)ps.1949-1204.0000359.

[33] Fayez O, El-Mahdy M, Ezz M, Ahmed H, Metwalli S. Computer aided optimization of natural gas pipe networks using genetic algorithm. Appl Soft Comput 2010;10:1141–50. https://doi.org/10.1016/j.asoc.2010.05.010.

[34] Hansen CT, Madsen K, Nielsen HB. Optimization of pipe networks. Math Program 1991;52:45–58. https://doi.org/10.1007/BF01582879.

[35] Zhang J, Zhu D. A bilevel programming method for pipe network optimization. SIAM J Optim 1996;6:838–57. https://doi.org/10.1137/S1052623493260696.

[36] Vaccariello E, Leone P, Stievano IS. Generation of synthetic models of gas distribution networks with spatial and multi-level features. Int J Electr Power Energy Syst 2020;117:105656. https://doi.org/10.1016/j.ijepes.2019.105656.

[37] Nasr GG, Connor NE. Natural gas engineering and safety challenges: Downstream process, analysis, utilization and safety. vol. 9783319089485. Springer International Publishing; 2014. https://doi.org/10.1007/978-3-319-08948-5.

[38] UNI 9165:2020, Gas Infrastructures - Pipelines for maximum operating pressure up to and including 0,5 MPa (5 bar) - Design, construction, testing, operation, maintenance and rehabilitation 2020.

[39] Autorità di Regolazione per Energia Reti e Ambiente (ARERA). Codice di Rete per il servizio di distribuzione gas - CRDG. n.d.

[40] Vaccariello E, Leone P, Canavero FG, Stievano IS. Topological modelling of gas networks for co-simulation applications in multi-energy systems. Math Comput Simul 2020;183:244–53. https://doi.org/10.1016/j.matcom.2019.12.018.

[41] ISO - ISO 12213-3:1997 - Natural gas — Calculation of compression factor — Part 3: Calculation using physical properties n.d.

[42] Kunz O, Wagner W. The GERG-2008 wide-range equation of state for natural gases and other mixtures: An expansion of GERG-2004. J Chem Eng Data 2012;57:3032–91. https://doi.org/10.1021/je300655b.

[43] Vaccariello E. Synthetic models of distribution gas networks in low-carbon energy systems. Politecnico di Torino 2021. https://iris.polito.it/handle/11583/2912988