



**HAL**  
open science

# Application of Topological Network Measures to Identify Critical Gas Transmission Network Components

Yiping Fang, Enrico Zio

► **To cite this version:**

Yiping Fang, Enrico Zio. Application of Topological Network Measures to Identify Critical Gas Transmission Network Components. [Research Report] CentraleSupélec. 2018. hal-01924426

**HAL Id: hal-01924426**

**<https://hal.science/hal-01924426v1>**

Submitted on 15 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



CentraleSupélec

---

## Application of Topological Network Measures to Identify Critical Gas Transmission Network Components

---

Yiping FANG and Enrico ZIO

27 MAI 2018

CHAIR ON SYSTEMS SCIENCE AND THE ENERGY CHALLENGE (SSEC)  
3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France

**Disclaimer:**

*The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.*

## Table of content

Abstract .....	3
Executive Summary .....	4
1. Introduction.....	6
2. Criticality measures for gas infrastructure networks .....	8
2.1 Criticality of transmission components .....	8
2.1.1 Shortest path betweenness centrality .....	9
2.1.2 Max-flow betweenness centrality .....	10
2.1.3 Generalized information centrality .....	11
2.2 Criticality of supply.....	12
3. Application to a realistic European gas network.....	13
3.1 Description of the analyzed European gas network .....	13
3.2 Criticality of gas transmission components .....	15
3.3 Criticality of gas supply .....	21
4. Discussions .....	24
5. Conclusions.....	25
References.....	25

## Abstract

This study introduces three generalized component importance measures for identifying critical components of heterogeneous substrate infrastructure networks. The proposed metrics are applied to a realistic European gas transmission network and different sets of critical components are obtained. The criticality of gas supply and the resilience of demand nodes are also analyzed based on the multiple sources and sinks maximum flow problem. The weakness of the network structure under some particular supply disruption scenarios is highlighted. These results of this type of analysis can be of great use in supporting network design and operation, e.g., for deciding which components should be protected and/or upgraded under a limited budget.

## Executive Summary

The study aims to identify critical gas transmission network components by application of three generalized component importance metrics, i.e., the shortest path-based betweenness centrality  $C^{B-SP}$ , the maximum flow-based betweenness criticality  $C^{B-MF}$  and the information centrality  $C^{I-G}$ . These metrics are complementary in the sense that they capture different aspects of the role that a component plays in the network. Identification of critical network components is highly linked to system resilience as loss of a critical component might significantly undermine system's ability to perform its functions, i.e. ensure adequate gas supply to its customers. The key results and conclusions drawn from the application of the proposed criticality measures to the realistic European gas transmission network of the case study are as follows:

- There is a large positive correlation of the critical components identified by  $C^{B-SP}$  and  $C^{I-G}$ , since the two metrics are both based on the shortest paths problem. One advantage of  $C^{I-G}$  comparing to  $C^{B-SP}$  is that it is able to characterize the criticalities of gas supply nodes which are related to their geographical distances to all the demand nodes.
- The measure  $C^{B-MF}$  can identify components that are critical in terms of both their structural location and transmission capacities. E.g., nodes 30, 31, 27 and links 30-31, 27-31 which have relatively low values of  $C^{B-SP}$  and  $C^{I-G}$  have actually high values of  $C^{B-MF}$ , as they are essential for transmitting gas from sources 29 and 38 to the rest of the network.
- Particular attention should be paid to components jointly identified as critical by all the three measures, e.g. node 12 that serves as a hub to connect different subsets of components loosely linked in the gas transmission network.
- The criticality of gas supply depends not only on the capacity of the gas source but also on its topological location in the network. For instance, disruptions at node 10 of relatively large supply capacity do not reduce the global network throughput significantly, since this node is originally designed as a redundant gas supply (backup LNG source).
- Large values of coefficient of variation of throughput at some demand nodes indicate that these nodes are particularly vulnerable to certain types of supply disruptions, which should be paid particular attention to by the system operators.



## 1. Introduction

Reliable energy (e.g., electrical power and natural gas) supply is essential for many of the services of our society. Disturbances in the energy supply have the potential of severely disrupting these services. Examples of unforeseen energy crises, with particular reference to gas supply, include the disputes between Russia and Ukraine over the price of natural gas in 2009 ([Yafimava 2009](#)), the terrorist attack on the Amenas gas plant that affected more than 10% of Algerian production of natural gas in 2013 ([Chrisafis, Borger et al. 2013](#)), and the gas supply shortage in 2013 when the UK had only 6 hours worth of gas left in storage as a buffer ([Gill and Guy 2013](#)). These events have risen significant concern about the reliability and resilience to disturbances and failures of our energy infrastructure systems, with a corresponding demand for methods capable of analyzing the vulnerabilities of these systems ([Zio 2009](#), [Zio and Piccinelli 2010](#)).

In particular, the security of supply of natural gas, a fossil fuel that accounts for 24% of energy consumption in OECD-Europe, has become one of the top priorities for the European Commission (EC). The recently adopted European Union (EU) Regulation 2017/1938 concerning measures to safeguard the security of gas supply and repealing Regulation (EU) No. 994/2010 has highlighted the importance of *a Framework Strategy for a Resilient Energy Union* ([EU Regulation 2010](#), [Praks, Kopustinskias et al. 2015](#)). Resilience is an emerging term to describe system's ability to withstand and recover quickly from an accident or disruptive event, which may be known or unknown ([Fang, Pedroni et al. 2016](#), [Zio 2018](#)). This is strongly reflected in the Energy Union package ([EU regulation 2017](#)).

Gas transport is expensive and is done mainly over a geographically large-scale pipeline network, i.e., a gas transmission network, which is usually understood as a critical infrastructure (CI) ([Zio 2016](#)). The protection of CI has been recently addressed by various initiatives from research institutions and governments worldwide. The EC has taken the initiative to organize a network consisting of research and technology organizations within the EU with interests and capabilities in CI protection ([Lewis, Ward et al. 2013](#)).



Identification of critical network components is highly linked to system vulnerability and resilience, as loss of a critical component might significantly undermine a system's ability to perform its functions, i.e., ensure adequate gas supply to its customers. In this context, the network analysis paradigm set up to study the dynamics of the relations in social networks has been previously utilized to analyze the vulnerability of various critical infrastructures, including electric power systems ([Holmgren 2006](#), [Zio, Petrescu et al. 2008](#), [Fang and Zio 2013](#)) and gas transmission networks ([Carvalho, Buzna et al. 2009](#)). The focus of these types of studies is typically on analyzing the structural properties of the system from a topological point of view, i.e., considering merely the connectivity properties of the network ([Freeman 1978](#), [Latora and Marchiori 2007](#)). The related measures in this thread are usually based on the assumption of shortest path routing and neglecting the physical characteristics of the connections (e.g., transmission link capacity), which may lead to misleading results ([Hines and Blumsack 2008](#)).

To overcome the major drawback of component importance measures relying only on topological information, a number of studies have proposed component criticality measures that take into account the capacities of the transmission elements and examine the different transmission routes available to the network flow. For example, [Freeman, Borgatti et al. \(1991\)](#) introduced a flow-based betweenness centrality measure based on the idea of maximum network flow. [Jenelius, Petersen et al. \(2006\)](#) suggested several vulnerability-based importance measures for transportation networks. [Zio and Piccinelli \(2010\)](#) proposed a randomized flow model-based centrality measure specifically for electrical power networks.

In this study, we adopt the two above mentioned complementary points of view aiming at identifying the critical components of a gas transmission network. We present this study with reference to a real gas transmission infrastructure of three European countries. Three component criticality measures, i.e., the shortest path-based betweenness centrality ([Freeman 1978](#)), maximum flow-based betweenness centrality ([Freeman, Borgatti et al. 1991](#)) and information centrality ([Latora and Marchiori 2007](#)), are carefully chosen and generalized to be applicable to the context of heterogeneous substrate gas transmission networks, where gas supply, transport and consumption occur. For further understand the roles that the different suppliers play in

determining the global network throughput and the resilience of sink nodes under possible disruptions, we adopt the multiple sources and sinks maximum flow (MSS-Max-Flow) problem formulation ([Deo 2017](#)) for modeling the gas pipeline operation, based on which a set of supply disruption scenarios that are relevant for understanding security of supply effects are investigated.

## 2. Criticality measures for gas infrastructure networks

For characterizing the role that a component (node or link) plays in a network, various measures of the importance of a network element, namely, of the relevance of its location in the network with respect to a given network performance, have been proposed. Classical topological centrality measures include the degree centrality, the closeness centrality, the information centrality, the betweenness centrality, and the Eigenvector centrality ([Freeman 1978](#), [Latora and Marchiori 2007](#)).

The above-mentioned measures are originally defined based on the assumption that the considered network is homogeneous, i.e., network nodes/links are all of the same kind. This is not the case for a gas transmission network which is in fact substrate, where gas flows from sources to sinks through components laid out heterogeneously in geographical space. In this respect, we carefully choose several relevant measures, i.e., the shortest path-based betweenness centrality, maximum flow-based betweenness criticality and the information centrality, and generalize them to be applicable to the context of heterogeneous substrate energy infrastructure networks. In addition, it is also important to study the vulnerability of gas infrastructure networks from the point of view of security of supply, as emerged by a number of recent energy supply disruptions due to economic, political or technical reasons. For this purpose, we adopt the multiple sources and sinks maximum flow-based model to quantify the network reaction to various gas supply disruption scenarios.

### 2.1 Criticality of transmission components

This section introduces three generalized betweenness centrality/criticality measures for characterizing the importance of transmission nodes and links in a gas transmission network.

### 2.1.1 Shortest path betweenness centrality

Betweenness includes node and link betweenness; it reflects the global connection importance of the component in the entire network. The classical betweenness centrality is defined based on the shortest path because, in practice, the network design and route planning usually consider the shortest path as an important reference. Consider a substrate network  $G = (V, E)$  with node set  $V$  and edge set  $E$ . The betweenness centrality of node  $i \in V$  is defined as the relative number of shortest paths between all pairs of nodes which pass through node  $i$ ,

$$C^B(i) = \sum_{s,t \in V, s \neq t \neq i} \frac{n_{st}(i)}{n_{st}}, \quad (1)$$

where  $n_{st}$  is the number of shortest paths from node  $s$  to node  $t$  and  $n_{st}(i)$  is the number of these paths passing through node  $i$ . Similarly, the betweenness centrality for link  $e_{ij} \in E$  is defined by

$$C^B(e_{ij}) = \sum_{s,t \in V, s \neq t} \frac{n_{st}(e_{ij})}{n_{st}}, \quad (2)$$

where  $n_{st}(e_{ij})$  is the number of these paths containing link  $e_{ij}$ .

Betweenness centrality is relevant in man-made networks which deliver products, substances, or materials as cost constraints on these networks condition transportation to occur along shortest paths. The classical node and link betweenness centralities defined by Eqs. (1) and (2) count all the network nodes as potential sources and sinks, which is not the case for gas transmission networks where there is only a small set of supply nodes. Here, we propose a generalization of betweenness centrality in the context of given specific sets of sources and sinks. Consider a gas network  $G = (V, E)$  whose nodes are composed by supply nodes  $V_S$ , transport nodes  $V_T$  and demand nodes  $V_D$ , i.e.,  $V = V_S \cup V_D \cup V_T$ : the generalized shortest path-based node and link betweenness centrality measures are given by respectively

$$C^{B-SP}(i) = \sum_{s \in V_S, t \in V_D} \frac{n_{st}(i)}{n_{st}}, \quad (3)$$

$$C^{B-SP}(e_{ij}) = \sum_{s \in V_S, t \in V_D} \frac{n_{st}(e_{ij})}{n_{st}}. \quad (4)$$

### 2.1.2 Max-flow betweenness centrality

The betweenness centrality measures introduced before assume that transport occurs along the shortest paths in geographical space. However, they do not capture the capacities of supply nodes and transmission links. To overcome this methodology, we consider the maximum flow problem as follows. In a network with given supply and link capacities, we aim to transmit the maximum flow possible between two particular nodes, a source and a sink, without exceeding the capacity of the source and the capacity of any link along the connecting path. Formally, we denote the capacity of source nodes and transmission links by a capacity function  $c : E \cup V_S \rightarrow \mathbb{R}^+$ . The maximum source-terminal (s-t) flow is defined by the following linear programming (LP) problem:

$$F_{st}(G, c) = \max_f \left[ \sum_{i: e_{it} \in E} f(e_{it}) - \sum_{j: e_{tj} \in E} f(e_{tj}) \right] \quad (5)$$

subject to

$$-c(e_{ij}) \leq f(e_{ij}) \leq c(e_{ij}), \forall e_{ij} \in E \quad (6)$$

$$\sum_{j: e_{sj} \in E} f(e_{sj}) - \sum_{i: e_{is} \in E} f(e_{is}) \leq c(s) \quad (7)$$

$$\sum_{j: e_{ji} \in E} f(e_{ji}) = \sum_{j: e_{ij} \in E} f(e_{ij}), \forall i \in V \setminus \{s, t\} \quad (8)$$

where inequalities (6) and (7) are link and supply capacity constraints, respectively, and equations (8) are the flow conservation constraints.

The question to address is how does the maximum flow between all sources and sinks change when a transport node  $i \in V \setminus \{s, t\}$  or a link  $e_{ij} \in E$  are removed from the network. Based on the model (5)-(8), we can calculate the flow that is lost when a transport node  $i$  or a link  $e_{ij}$  is removed from operation, assuming that the network is working at its maximum capacity. Then, the generalized max-flow node and link betweenness centralities are defined as

$$C^{B-MF}(i) = \sum_{s \in V_S, t \in V_D} \frac{F_{st}(G, c) - F_{st}(G \setminus i, c)}{F_{st}(G, c)}, \quad (9)$$

$$C^{B-MF}(e_{ij}) = \sum_{s \in V_S, t \in V_D} \frac{F_{st}(G, c) - F_{st}(G \setminus e_{ij}, c)}{F_{st}(G, c)}. \quad (10)$$

where  $[F_{st}(G, c) - F_{st}(G \setminus i, c)]$  and  $[F_{st}(G, c) - F_{st}(G \setminus e_{ij}, c)]$  represent the amount of flow which must go through node  $i$  and link  $e_{ij}$ , respectively, when the network is operating at maximum capacity, and  $F_{st}(G, c)$  is the maximum s-t flow in  $G$  calculated through (5)-(8).

### 2.1.3 Generalized information centrality

The original topological information centrality,  $C^I$ , relates a node/link importance to the ability of the network to respond to the disruption of the node/link. In this view, the network performance is measured by the network topological efficiency  $E[G]$  defined as:

$$E[G] = \sum_{s, t \in V, s \neq t} \varepsilon_{st}, \quad (11)$$

where  $\varepsilon_{st} = 1/d_{st}$  is the efficiency of the connection between nodes  $s$  and  $t$  measured as the inverse of the shortest path distance connecting them. Similar to the definition of the generalized betweenness centrality, we use

$$E'[G] = \sum_{s \in V_S, t \in V_D} \varepsilon_{st} \quad (12)$$

to characterize the network efficiency of a gas transmission network with given sets of sources and sinks.

Then, the generalized information centrality of node  $i$  and edge  $e_{ij}$  are defined as the relative drop in the network efficiency caused by the removal of the node and the edge, respectively, given by

$$C^{I-G}(i) = \frac{E'(G) - E'(G \setminus i)}{E'(G)}, \quad (13)$$

$$C^{I-G}(e_{ij}) = \frac{E'(G) - E'(G \setminus e_{ij})}{E'(G)}. \quad (14)$$

## 2.2 Criticality of supply

According to their definitions, the generalized shortest path and max-flow betweenness centrality measures are appropriate for characterizing the criticalities of network links and transport nodes, and the generalized information centrality is suitable for quantifying the topological roles of different supply nodes from the perspective of shortest path routing.

To tackle the problem of criticality of suppliers from the network flow perspective, we formulate the following multiple sources and sinks maximum flow (MSS-Max-Flow) problem to model the gas pipeline operation:

$$F(G, c) = \max_f \left[ \sum_{t \in V_D} \left( \sum_{i: e_{it} \in E} f(e_{it}) - \sum_{j: e_{tj} \in E} f(e_{tj}) \right) \right] \quad (15)$$

subject to

$$-c(e_{ij}) \leq f(e_{ij}) \leq c(e_{ij}), \forall e_{ij} \in E \quad (16)$$

$$\sum_{j: e_{sj} \in E} f(e_{sj}) - \sum_{i: e_{is} \in E} f(e_{is}) \leq c(s), \forall s \in V_S \quad (17)$$

$$\sum_{j: e_{ji} \in E} f(e_{ji}) = \sum_{j: e_{ij} \in E} f(e_{ij}), \forall i \in V_T. \quad (18)$$

We define the  $F(G, c)$  as the global network throughput, which is the sum of the maximum flow received at all sinks. It can be seen that the MSS-Max-Flow problem is similar to the maximum s-t flow problem (5)-(8). Actually, it can be shown that it can be straightforwardly converted into a one-source and one-sink problem (Prentice Hall 2008). Let us introduce a supersource  $s$  (virtual source node) with edges (of unlimited capacity) directed from this supersource  $s$  to all source nodes  $s_1, s_2, \dots, s_k$ . Furthermore, let us introduce a supersink  $t$  (virtual sink node) with edges (also of unlimited capacity) directed from all sink nodes  $t_1, t_2, \dots, t_r$  to the supersink  $t$ . Then, the problem of maximizing the total value of the flow from all sources is equivalent to that of maximizing the value of the flow from  $s$  to  $t$ .

A set of supply disruption scenarios significant for the security of supply effects are defined,  $\Phi = \{\phi_1, \phi_2, \dots, \phi_p\}$ , for analyzing the consequences (i.e., reductions of the global network throughput)

of loss of one or more key gas sources. Then, the criticality of disruption scenario  $\phi_i$  can be defined as the relative drop in the global network throughput that it causes

$$C^S(\phi_i) = \frac{F(G, c) - F(G, c|\phi_i)}{F(G, c)}. \quad (19)$$

where  $F(G, c|\phi_i)$  is the global network throughput under the disruption scenario  $\phi_i$ .

Based on the MSS-Max-Flow model and the set of disruption scenarios, we can also characterize the supply resilience of the sink node by analyzing the variation of throughput at a sink node across different scenarios: the supply at the sink node is resilient if it combines high throughput across different scenarios with a low coefficient of variation of throughput.

### 3. Application to a realistic European gas network

#### 3.1 Description of the analyzed European gas network

Fig. 1 shows the topology of the European gas transmission network analyzed in our study. This test case is based on the real gas transmission network of three European countries. The presented supply and demand datasets are realistic. However, the geographical topology is not disclosed for sensitivity reasons. The network contains the following elements: pipelines, compressor stations and the LNG terminal (node 10).

The analyzed network has 56 nodes and 74 transmission links, and there are 5 supply nodes: 2, 10, 19, 29 and 38 (see Table 1). All numbers are expressed in millions of cubic metre per day (mcm/d). The gas source at Node 10 represents an LNG terminal, with a maximum designed capacity of 10.2 mcm/d. The Pipeline diameters and their lengths were obtained from the gas operators. Consequently, the respective capacities have been estimated from pipelines diameters. The detailed properties of the transmission gas pipelines are summarized in Table 2. It is noted that nodes 28 and 51 in the gas network are export nodes and they will be supplied only when all other nodes in the network are supplied, i.e., in case of disruption, internal network sink nodes are supplied first, and nodes 28 and 51 only in case there is gas available. Therefore, the distances of

the links to the two nodes, i.e., 7-51, 8-51 and 21-28, are artificially set very large for the calculations of the shortest path betweenness and information betweenness centralities.

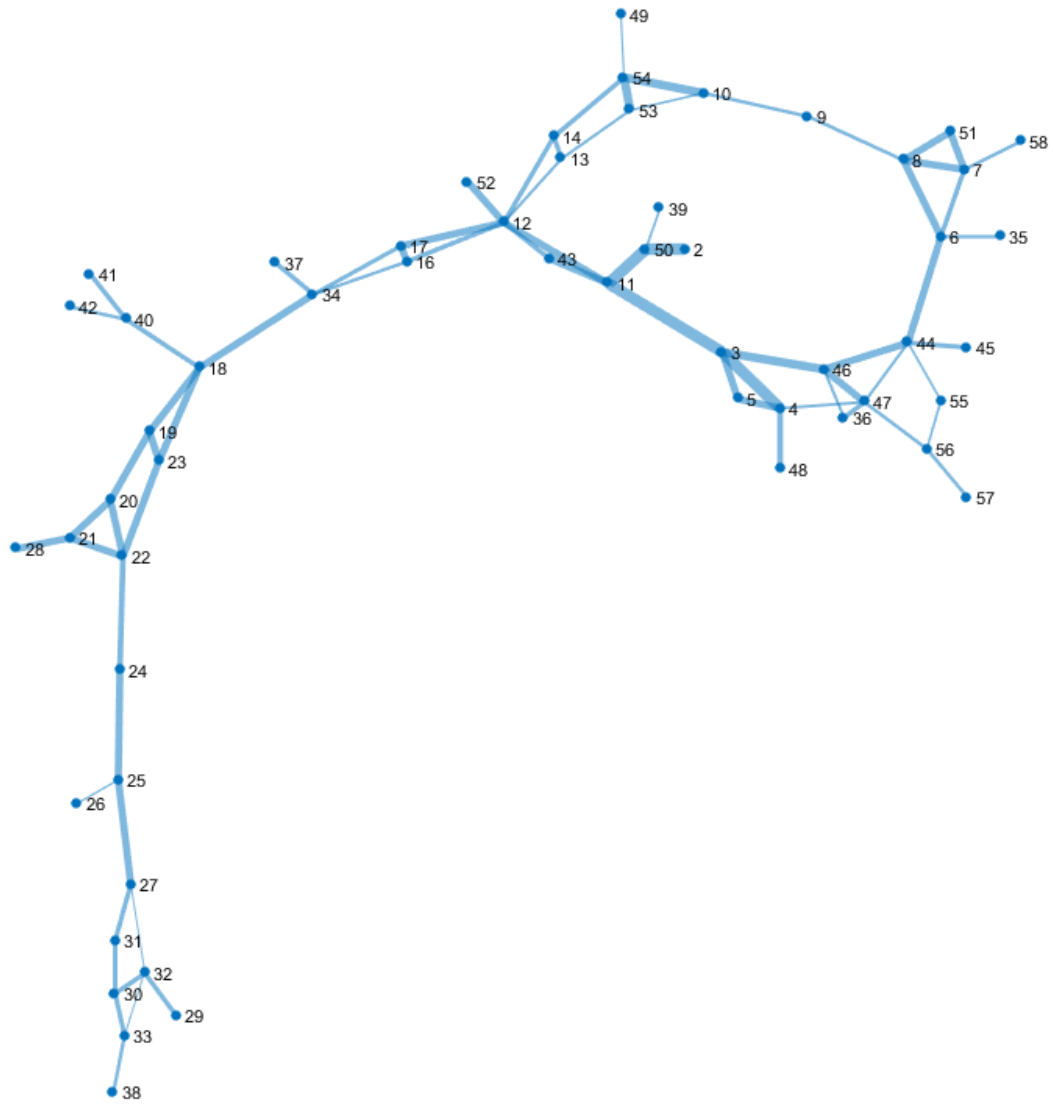


Figure 1. The topology of the analyzed European gas network. The linewidth is proportional to the pipeline capacity.

Table 1. Properties of the gas sources of the gas transmission network.

Source nodes	Capacity (mcm/d)
2	31.2
10	10.2
19	30.0
29	4.3
38	2.7



Table 2. Properties of transmission links of the gas network.

From	To	Capacity (mcm/d)	Length (km)	From	To	Capacity (mcm/d)	Length (km)
2	50	31	23	18	34	12.11	43
3	4	59.16	0.01	18	40	5.05	148
3	5	12.11	32	19	20	12.11	60
3	11	30.6	29	19	23	59.16	0.01
3	46	17.13	22	20	21	12.11	90
4	5	12.11	32	20	22	59.16	0.01
4	47	2	22	21	22	12.11	90
4	48	8.11	2	21	28	12	10086
6	7	5.05	80	22	23	12.11	60
6	8	12.11	80	22	24	7.5	86
6	35	2.83	30	24	25	12.11	86
6	44	12.11	11.6	25	26	0.83	46
7	8	59.16	0.01	25	27	12.11	100
7	51	12.11	10200	27	31	59.16	0.01
7	58	2.83	50	27	32	0.47	70
8	9	2.83	25	29	32	5.05	195
8	51	12.11	10200	30	31	5.05	70
9	10	2.83	162	30	32	59.16	0.01
10	53	1.34	144	30	33	5.05	60
10	54	17.13	144	32	33	0.47	60
11	12	12.11	103	33	38	2.83	60
11	43	12.11	34	34	37	5.05	200
11	50	31	31	36	46	2	24
12	13	2	85	36	47	5.05	24
12	14	5.05	85	39	50	1.34	106
12	16	5.05	62	40	41	5.05	32
12	17	12.11	62	40	42	2.83	63
12	43	5.05	132	44	45	5.05	1
12	52	12.11	10	44	46	12.11	23
13	14	59.16	0.01	44	47	2	23
13	53	2	30	44	55	0	30
14	54	5.05	30	46	47	59.16	0.01
16	17	59.16	0.01	47	56	2.83	37
16	34	2.5	24	49	54	0.83	40
17	34	4	24	53	54	59.16	0.01
18	19	12.11	43	55	56	1.34	15
18	23	12.11	43	56	57	2.83	18

### 3.2 Criticality of gas transmission components

The three component criticality indicators introduced in Section 2.1, namely, the shortest path betweenness centrality  $C^{B-SP}$ , the max-flow betweenness centrality  $C^{B-MF}$ , and the generalized information centrality  $C^{I-G}$  are used to identify the most critical components with respect to the different network features they measure. Fig. 2 shows the results of the node criticality indicators

rankings, where only the 30 most critical nodes are reported. Note that all the results in this section are normalized for better illustration and comparison. According to Fig. 2, nodes 18, 11, 12 and 34 turn out to be among the top five most critical nodes with respect to both  $\mathcal{C}^{B-SP}$  and  $\mathcal{C}^{I-G}$ , whereas according to  $\mathcal{C}^{B-MF}$  nodes 18, 11, and 34 are less important than other nodes, for example, nodes 30 and 31; this is due to the fact that the latter two nodes constitute the main channel to transmit gas from source nodes 29 and 38 to the upper parts of the network. It is noted that the two channels 30-31-27 and 32-27 are the only two ways to transmit gas from sources 29 and 38 to the demand nodes at the upper parts, e.g., 27, 25, 24, etc., and the channel 30-31-27 has a much larger transmission capacity (5.05 mcm/d) than the channel 32-27 (0.47 mcm/d). Therefore, nodes 30 and 31 are identified as the most critical components with respect to  $\mathcal{C}^{B-MF}$ , which takes into account the characteristics of the link transmission capacities.

Fig. 3 reports the results of the criticality indicators rankings for the network links. It is shown that link 18-34 is the most critical one with respect to  $\mathcal{C}^{B-SP}$  and  $\mathcal{C}^{I-G}$ , whereas links 27-31 and 30-31 are the most critical ones according to  $\mathcal{C}^{B-MF}$ . The ranking agreement among the  $\mathcal{C}^{B-SP}$  and  $\mathcal{C}^{I-G}$  indicators both for nodes and links is somewhat not unexpected since they are both based on the network shortest path problem. As a matter of fact, the correlation coefficients between  $\mathcal{C}^{B-SP}$  and  $\mathcal{C}^{I-G}$  are  $r(\mathcal{C}_n^{B-SP}, \mathcal{C}_n^{I-G}) = 0.65$  for nodes and  $r(\mathcal{C}_e^{B-SP}, \mathcal{C}_e^{I-G}) = 0.53$  for links.

To better understand the roles of the identified critical components in the whole network, Fig. 4 (a)-(c) and Fig. 5(a)-(c) illustrate the top 10 most critical components (nodes and links) identified by  $\mathcal{C}^{B-SP}$ ,  $\mathcal{C}^{B-MF}$  and  $\mathcal{C}^{I-G}$ , respectively. It can be observed that nodes and links with high shortest-path betweenness centrality, e.g., nodes 18, 12, 34 in Fig. 4(a) and link 18-34 in Fig. 5(a), are often near the network barycenter, whose location is given by  $\mathbf{x}_G = \sum_i \mathbf{x}_i / N$ . Although (part of) these components are identified as critical according to  $\mathcal{C}^{I-G}$  when comparing Fig. 4(a) with Fig. 4(c) and Fig. 5(a) with Fig. 5(c), the generalized information centrality indicator gives emphasis also to the gas source nodes and the components close to the sources. For example, gas supply nodes 2, 10, 19 in Fig. 4(c) and link 2-50 in Fig. 5(c), which serves as the only connection to supply node 2, have been identified by  $\mathcal{C}^{I-G}$  as important components for the transmission network. As a matter of fact, being able to identify critical suppliers in terms of their distances to all the demand

nodes is a significant feature of  $C^{I-G}$  compared to  $C^{B-SP}$  which is only defined for transmission components.

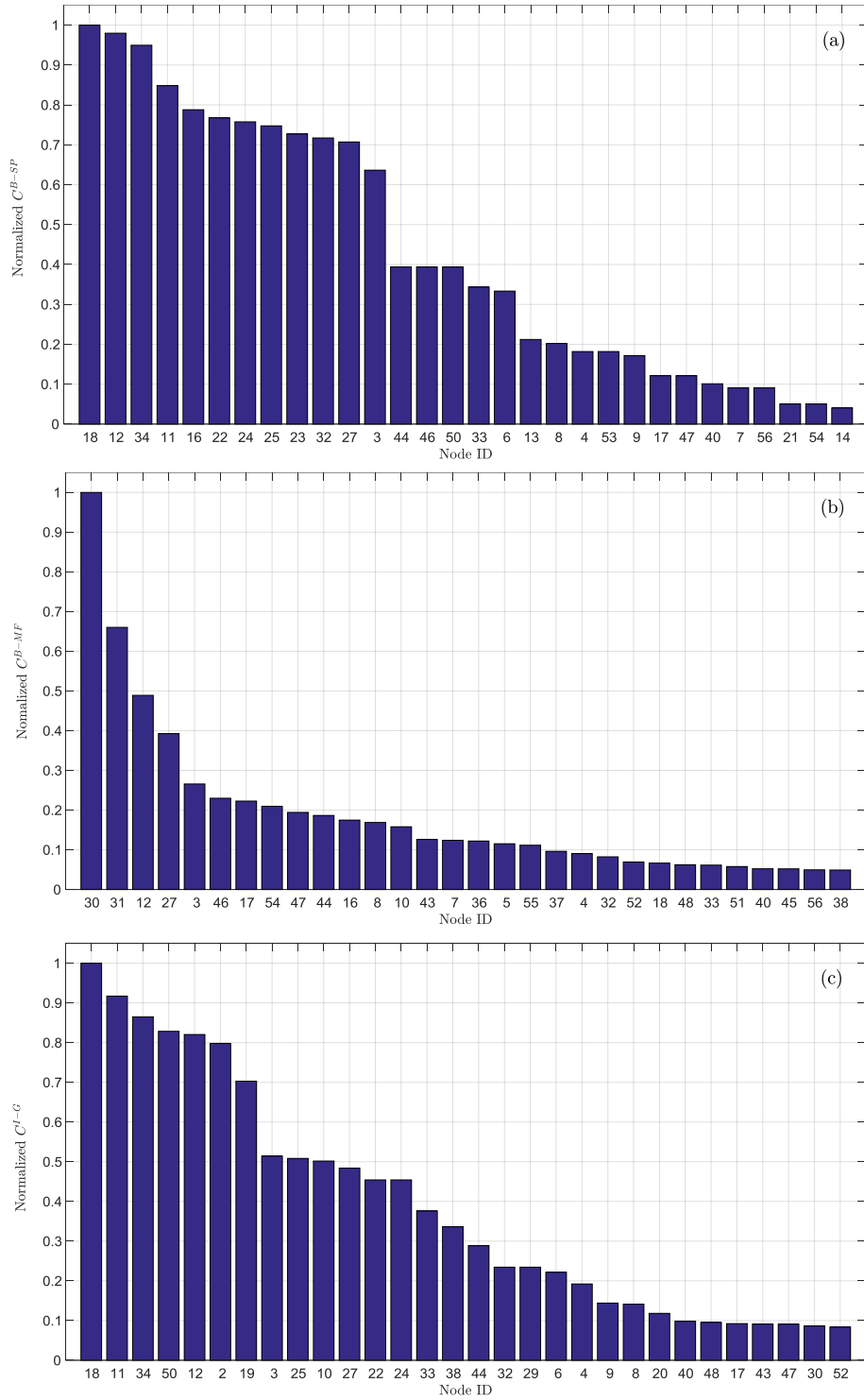


Figure 2. Results of the node criticality indicators rankings for normalized (a) shortest path betweenness centrality  $C^{B-SP}$ , (b) max-flow betweenness centrality  $C^{B-MF}$ , and (c) generalized information centrality  $C^{I-G}$ ; Only the 30 most critical nodes are reported.

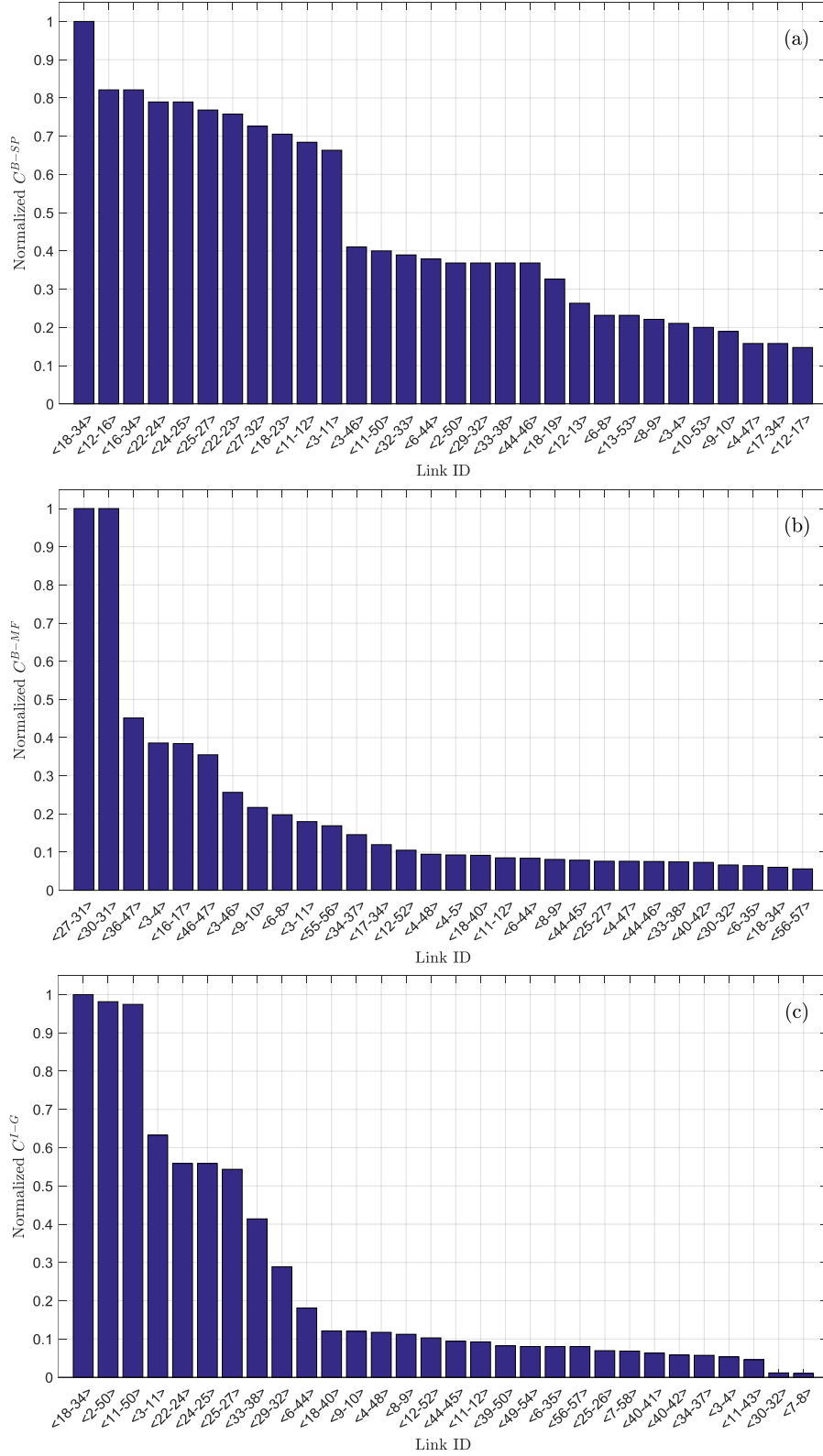


Figure 3. Results of the link criticality indicators rankings for normalized (a) shortest path betweenness centrality  $C^{B-SP}$ , (b) max-flow betweenness centrality  $C^{B-MF}$ , and (c) generalized information centrality  $C^{I-G}$ ; Only the 30 most critical links are reported.

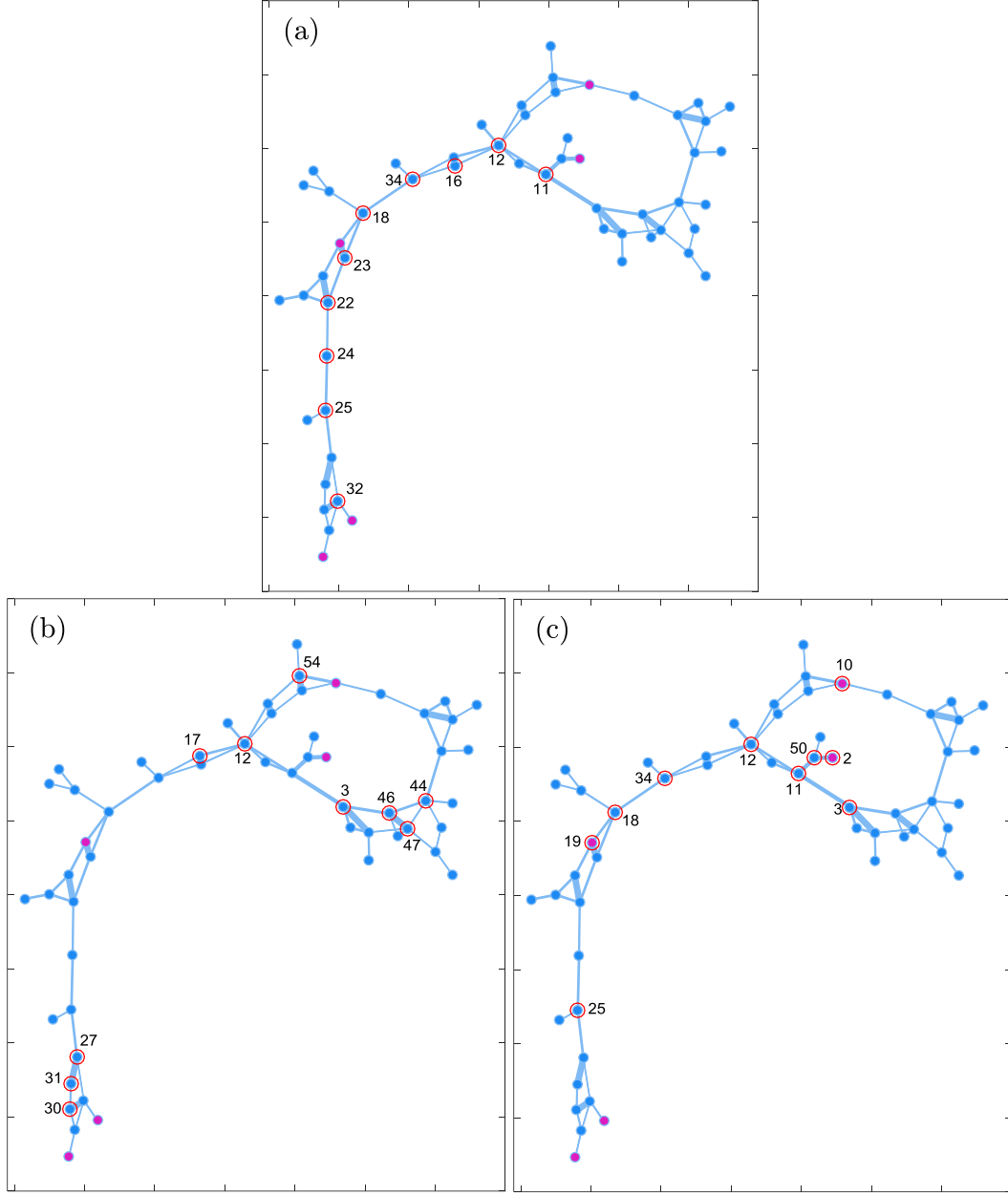


Figure 4. Illustration of the identified top 10 most critical nodes by (a) shortest path betweenness centrality  $C^{B-SP}$ , (b) max-flow betweenness centrality  $C^{B-MF}$ , and (c) generalized information centrality  $C^{I-G}$ ; Identified critical nodes are highlighted by red circles; magenta nodes are suppliers.

On the other hand, the critical components identified by the generalized max-flow betweenness centrality  $C^{B-MF}$  are very different from those identified by  $C^{B-SP}$  and  $C^{I-G}$ : the correlation coefficients of  $C^{B-MF}$  with  $C^{B-SP}$  and  $C^{I-G}$  are  $r(C_n^{B-MF}, C_n^{B-SP}) = 0.08$ ,  $r(C_n^{B-MF}, C_n^{I-G}) = -0.008$  for nodes, and  $r(C_e^{B-SP}, C_e^{B-MF}) = -0.13$ ,  $r(C_e^{B-MF}, C_e^{I-G}) = -0.08$  for links. It can be seen from Fig. 4(b) and Fig. 5(b) that the critical components obtained based on  $C^{B-MF}$  are those

along the periphery of the network, close to multiple sources and having relatively large transmission capacities (for links), e.g., nodes 30, 31, 27 and links 30-31, 27-31, which are close to source nodes 29 and 38 and, thus, they are usually essential elements for the transmission of the gas available from multiple source nodes to the whole network.

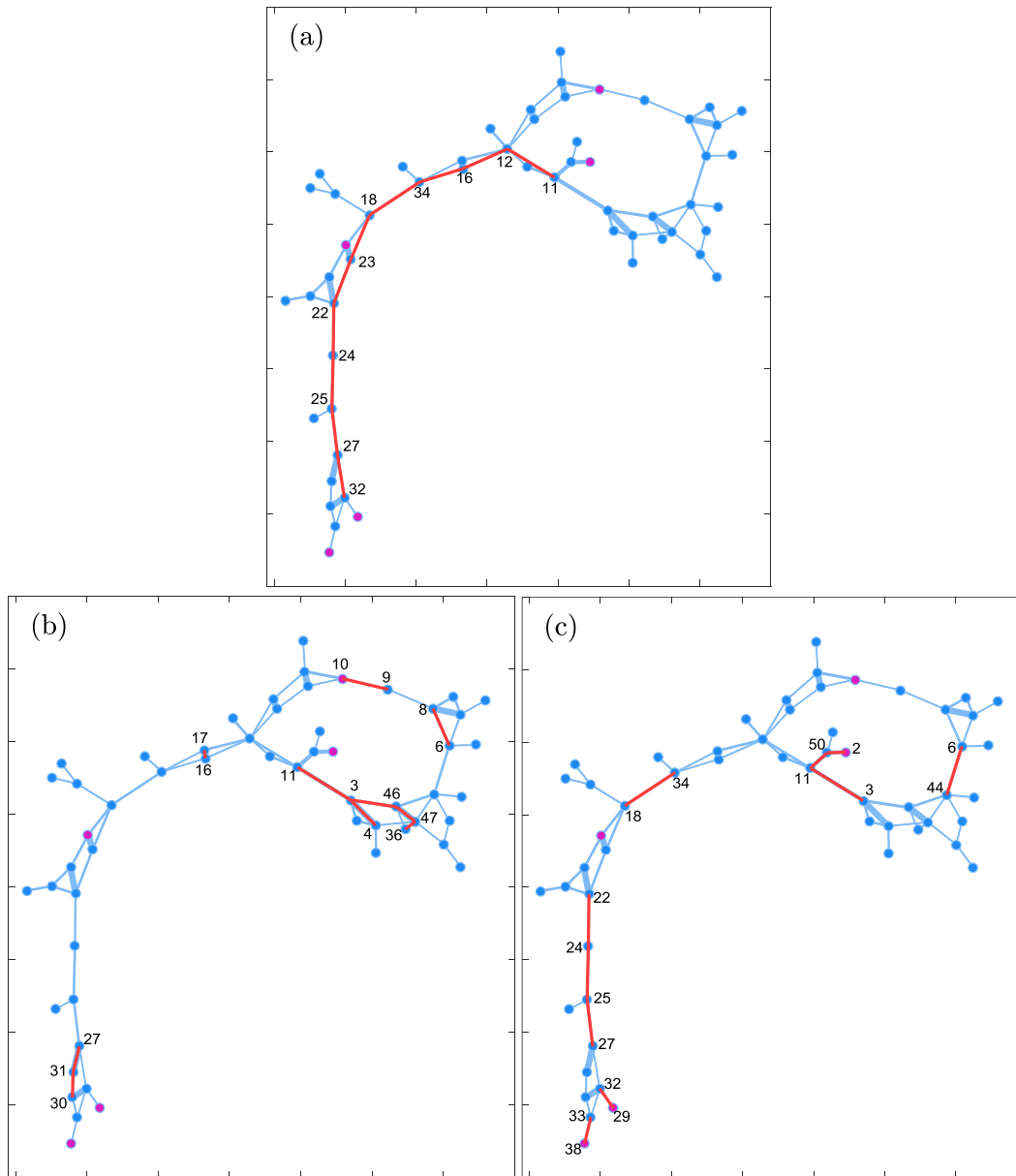


Figure 5. Illustration of the identified top 10 most critical links by (a) shortest path betweenness centrality  $C^{B-SP}$ , (b) max-flow betweenness centrality  $C^{B-MF}$ , and (c) generalized information centrality  $C^{I-G}$ ; Identified critical links are highlighted with red lines; magenta nodes are suppliers.

It is important to note that there are some components that are identified as critical by all three indicators, e.g., node 12, as highlighted in Fig. 4(a), (b) and (c); this is due to the fact that node 12 serves as a hub to connect different loosely linked subsets of components, namely, between the upper-right ring-like sub-network and the lower-left transmission backbone, and thus its failure would completely disconnect the two parts of the network and largely worsen the performance of the network in terms of either topological connection or capability of flow transmission.

### 3.3 Criticality of gas supply

In order to tackle the problem of criticality of suppliers and simulate the test network reaction to various gas supply disruptions, the following six sets of scenarios are defined: the Case  $\phi_1$  represents the gas system working in normal conditions (business as usual); the Case  $\phi_2(\varpi = 100\%)$  represents earlier version of the system, in which the LNG terminal is missing; the other selected scenarios have been chosen in order to analyze the consequences of gradual losses of each gas source node, i.e., cases  $\phi_3 \sim \phi_6$ .

- $\phi_1$ : Base case, LNG at Node 10 has an upper limit capacity of 10.2 mcm/d. No external disruption; i.e. all gas supply nodes are supplied as contracted.
- $\phi_2$ : External (partial) disruption. The gas supply at node 2 loses its supply capacity by  $\varpi$  percentage, where  $\varpi = 100\%$  represents a full disruption of supply node 2.
- $\phi_3$ : The LNG at node 10 has only  $(1 - \varpi)$  percent of its current capacity.
- $\phi_4$ : External disruption. The gas supply at node 19 loses its supply capacity by  $\varpi$  percentage.
- $\phi_5$ : External disruption. The gas supply at node 29 loses its supply capacity by  $\varpi$  percentage.
- $\phi_6$ : External disruption. The gas supply at node 38 loses its supply capacity by  $\varpi$  percentage.

Fig. 6 shows the results of the criticalities,  $C^S(\phi_i)$ , of all the considered supply disruption scenarios  $\phi_2 \sim \phi_6$  under different values of capacity reduction rate  $\varpi$ . We can see that the criticality increases (the global network throughput decreases) linearly with the increase of the capacity reduction  $\varpi$  for each case, which is not unexpected since the MSS-Max-Flow problem is a linear

programming problem. Furthermore, the cases of  $\phi_2$  and  $\phi_4$  are the two most critical scenarios, and even small disruptions at nodes 2 and 19 would cause large reductions in the global network throughput. This is probably because the two nodes are the main gas supplies of the whole network, the supply capacities of the two nodes 2 and 19 being 31.2mcm/d and 30mcm/d, respectively. On the other hand, the cases of  $\phi_5$  and  $\phi_6$  are relatively not critical since the supply capacities of nodes 29 and 38 are quite small (i.e., 4.3mcm/d and 2.7mcm/d, respectively): disturbances on these two nodes would not have significant impacts on the system throughput.

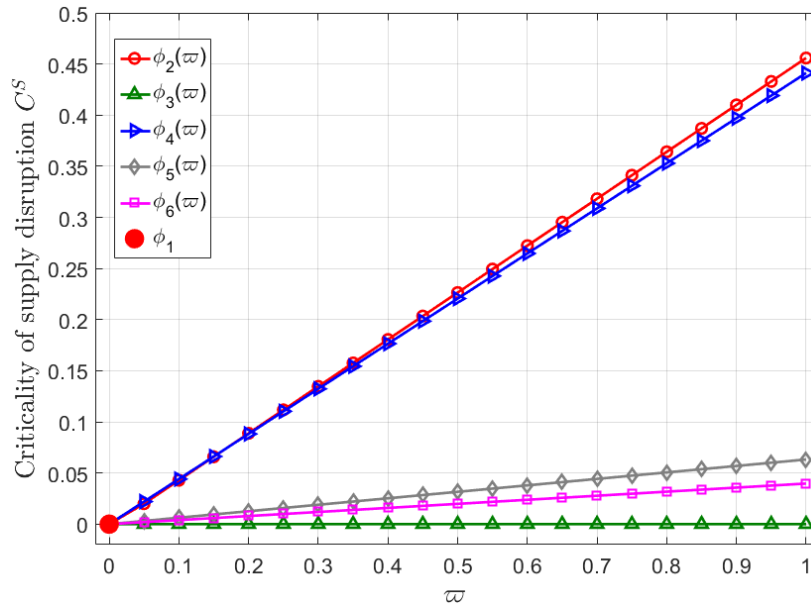


Figure 6. Criticality of supply disruption  $C^S(\phi_i)$  for cases  $\phi_1 \sim \phi_6$  under different values of capacity reduction rate  $\varpi$ .

However, it is important to note that the criticalities of disruptions at node 10 (i.e., cases  $\phi_3$ ) are all very small even though the supply capacity at this node is large, i.e., 10.2mcm/d, compared to nodes 29 and 38. In other words, disruptions only at node 10 would not reduce the global network throughput significantly. This is probably because the earlier version network without LNG at node 10 was designed (in terms of link capacities) to be able to deliver its gas supply to all its demand nodes: therefore, the addition of LNG at node 10 has not increased the throughput of the network largely. In turn, disruptions at node 10 would not worsen the system throughput considerably. These results show that the criticality of gas supply depends on both the capacity of the supply source and its topological location in the network.



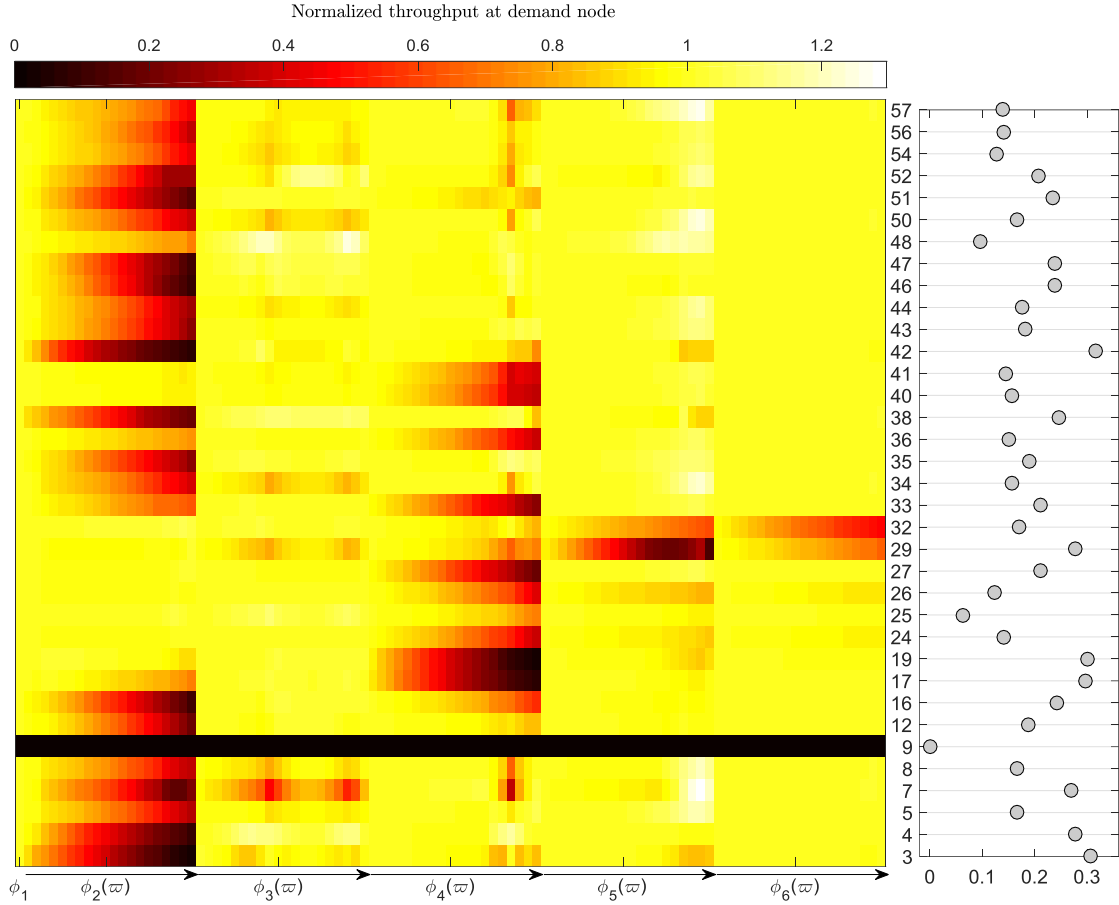


Figure 7. (Left panel) Heat-map of normalized throughput at each demand node across all the considered supply disruption scenarios, and (Right panel) coefficient of variation of normalized throughput at each demand node.

Finally, we address the resilience of demand nodes (which usually represent a geographical area) to supply disruptions. For this, we study the signatures in the scenario space given by the throughput at each demand node in each of the 5 sets of disruption scenarios  $\phi_2 \sim \phi_6$  and compare to those of the base case scenario. The heat-map in the left panel of Fig. 7 shows the normalized throughput for each pair of demand nodes and scenarios. The coefficient of variation, shown in the right panel of Fig. 7 measures the normalized dispersion of demand node throughput using the mean as a measure of scale. Larger values indicate that the throughput accessible to the sink node varies across the scenarios. We can see from Fig. 7 that some demand nodes have relatively unwavering throughput (nodes with relatively constant yellow row in the heat map and small coefficient variation) across all the scenarios, e.g., nodes 25 and 48, which may be due to a combination of effects: diversity of supply; good access to network capacity (strategic geographical location); and a relatively small throughput in the base case ([Carvalho, Buzna et al.](#)

2014). On the other hand, demand nodes like 3, 17, 19 and 42 have larger values of coefficient of variation of throughput, indicating that these nodes are particularly vulnerable to certain type of supply disruptions: nodes 3 and 42 are vulnerable to disruption scenarios  $\phi_2(\varpi)$  with  $\varpi \rightarrow 1$ , whereas nodes 17 and 19 are vulnerable to disruption scenarios  $\phi_4(\varpi)$  with  $\varpi \rightarrow 1$ .

## 4. Discussions

The critical components identified by the three indicators, i.e., the shortest path betweenness centrality  $C^{B-SP}$ , the max-flow betweenness centrality  $C^{B-MF}$ , and the generalized information centrality  $C^{I-G}$  are complementary in the sense that they capture different aspects of the role that a component plays in the network.  $C^{B-SP}$  shows which components act as bridges between sources and sinks in a network; it does so by identifying all the shortest paths between all pairs of sources and sinks and, then, counting how many times each component belongs to one.  $C^{I-G}$  describes the components' influence on the network efficiency, as defined based on the lengths of the shortest paths between sources and sinks.  $C^{B-MF}$  characterizes the amount of flow which must go through one component when the network is operating at its maximum capacity. Since  $C^{B-SP}$  and  $C^{I-G}$  are both based on the shortest path problem, there is a relatively large positive correlation of the critical components (nodes and links) identified by them. However, one advantage of  $C^{I-G}$  comparing to  $C^{B-SP}$  is that it is able to characterize the criticalities of source nodes, which are related to their distances to the demand nodes. On the other hand, the max-flow betweenness centrality  $C^{B-MF}$  contributes to the results by being able to identify components that are important in terms of both their structural location and transmission capabilities. For example, nodes 30, 31, 27 and links 30-31, 27-31 which have relatively low values of  $C^{B-SP}$  and  $C^{I-G}$  are actually of high values of  $C^{B-MF}$ : they are essential for transmitting gas from sources 29 and 38 to the rest of the network.

The results show that the criticality of gas supply depends not only on the capacity of the gas source but also on its topological location in the network. Disruptions at node 10 of relatively large supply capacity do not reduce the global network throughput significantly since this node is originally designed as a redundant gas supply (backup LNG source). Furthermore, the large values of coefficient of variation of throughput at some demand nodes indicate that these nodes are

particularly vulnerable to certain types of supply disruptions and, thus, should be paid specific attention by the system operators.

One key assumption made in this study for analyzing the criticality of gas supply is that the MSS-Max-Flow problem is used for simulating the gas pipeline operation, which is actually a simplified model of the complicated gas transmission process in reality. From Fig. 7 one can find that some demand nodes always receive a very small or even zero amount of gas flow (e.g., node 9) in the MSS-Max-Flow model that aims to maximize the network throughput as a whole. It is an unfair solution from the user point of view. This problem can be solved by congestion control algorithms by achieving cost-effective and scalable network protocols that best utilize the network capacity, sharing it among users in a fair way ([Buzna and Carvalho 2017](#)). Besides, more realistic nonlinear gas transmission models can be used ([De Wolf and Smeers 2000](#), [Üster and Dilaveroğlu 2014](#)) instead of the simplified linear ones. These aspects can be studied in future work.

## 5. Conclusions

This study has introduced three generalized component importance metrics, i.e., the shortest path-based betweenness centrality, the maximum flow-based betweenness criticality and the information centrality, for identifying critical components of heterogeneous substrate infrastructure networks. They have been applied to a realistic European gas transmission network. The results obtained by the three metrics are complementary in the sense that they capture different aspects of the role that a component plays in the network. Furthermore, the criticality of gas supply and the resilience of demand nodes have been analyzed based on the MSS-Max-Flow problem. The weaknesses of the network structure under particular disruption scenarios have been highlighted. These results are useful in supporting network design and operation, i.e., deciding which components should be protected and/or upgraded under a limited budget.

## References

Buzna, L. and R. Carvalho (2017). "Controlling congestion on complex networks: fairness, efficiency and network structure." [Scientific reports](#) 7(1): 9152.

Carvalho, R., L. Buzna, F. Bono, E. Gutiérrez, W. Just and D. Arrowsmith (2009). "Robustness of trans-European gas networks." Physical review E **80**(1): 016106.

Carvalho, R., L. Buzna, F. Bono, M. Maserà, D. K. Arrowsmith and D. Helbing (2014). "Resilience of natural gas networks during conflicts, crises and disruptions." PloS one **9**(3): e90265.

Chrisafis, A., J. Borger, J. McCurry and T. Macalister (2013). "Algeria hostage crisis: the full story of the kidnapping in the desert." The Guardian **25**.

De Wolf, D. and Y. Smeers (2000). "The gas transmission problem solved by an extension of the simplex algorithm." Management Science **46**(11): 1454-1465.

Deo, N. (2017). Graph theory with applications to engineering and computer science, Courier Dover Publications.

EU Regulation (2010). "No 994/2010 of the European parliament and of the Council of 20 October 2010 "concerning measures to safeguard security of gas supply and repealing Council Directive 2004/67." EC" Google Scholar.

EU regulation (2017). REGULATION (EU) 2017/1938 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 October 2017 concerning measures to safeguard the security of gas supply and repealing Regulation (EU) No 994/2010, European Parliament.

Fang, Y.-P., N. Pedroni and E. Zio (2016). "Resilience-based component importance measures for critical infrastructure network systems." IEEE Transactions on Reliability **65**(2): 502-512.

Fang, Y.-P. and E. Zio (2013). "Unsupervised spectral clustering for hierarchical modelling and criticality analysis of complex networks." Reliability Engineering & System Safety **116**: 64-74.

Freeman, L. C. (1978). "Centrality in social networks conceptual clarification." Social Networks **1**(3): 215-239.

Freeman, L. C., S. P. Borgatti and D. R. White (1991). "Centrality in valued graphs: A measure of betweenness based on network flow." Social networks **13**(2): 141-154.

Gill, P. and C. Guy (2013). UK gas supply six hours from running out in March. The Financial Times

Hines, P. and S. Blumsack (2008). A centrality measure for electrical networks. Hawaii International Conference on System Sciences, Proceedings of the 41st Annual, IEEE.

Holmgren, Å. J. (2006). "Using graph models to analyze the vulnerability of electric power networks." Risk Analysis **26**(4): 955-969.

Jenelius, E., T. Petersen and L.-G. Mattsson (2006). "Importance and exposure in road network vulnerability analysis." Transportation Research Part A: Policy and Practice **40**(7): 537-560.

Latora, V. and M. Marchiori (2007). "A measure of centrality based on network efficiency." New Journal of Physics **9**(6): 188.

Lewis, A. M., D. Ward, L. Cyra and N. Kourti (2013). "European reference network for critical infrastructure protection." International journal of critical infrastructure protection **6**(1): 51-60.

- Praks, P., V. Kopustinskas and M. Masera (2015). "Probabilistic modelling of security of supply in gas networks and evaluation of new infrastructure." Reliability Engineering & System Safety **144**: 254-264.
- Üster, H. and Ş. Dilaveroğlu (2014). "Optimization for design and operation of natural gas transmission networks." Applied Energy **133**: 56-69.
- Yafimava, K. (2009). The Russo-Ukrainian gas dispute of January 2009: a comprehensive assessment.
- Zio, E. (2009). "Reliability engineering: Old problems and new challenges." Reliability Engineering & System Safety **94**(2): 125-141.
- Zio, E. (2016). "Challenges in the vulnerability and risk analysis of critical infrastructures." Reliability Engineering & System Safety **152**: 137-150.
- Zio, E. (2018). "The Future of Risk Assessment." Reliability Engineering & System Safety.
- Zio, E., C.-A. Petrescu and G. Sansavini (2008). Vulnerability analysis of a power transmission system.
- Zio, E. and R. Piccinelli (2010). "Randomized flow model and centrality measure for electrical power transmission network analysis." Reliability Engineering & System Safety **95**(4): 379-385.