SciGRID_gas: The raw INET data set

Release 1.1

J.C. Diettrich & A. Pluta & W. Medjroubi

Mar 16, 2021

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Summary

The goal of SciGRID_gas is to develop methods to create an automated process that can generate a gas transmission network data set for Europe. Gas transmission networks are fundamental for simulations by the gas transmission modelling community, to derive major dynamic characteristics. Such simulations have a large scope of application, for example, they can be used to perform case scenarios, to model the gas consumption, to minimize leakages and to optimize overall gas distribution strategies. The focus of SciGRID_gas will be on the European transmission gas network, but the principal methods will also be applicable to other geographic regions.

Data required for such models are the gas facilities, such as compressor stations, LNG terminals, pipelines, etc. One needs to know their locations, in addition to a large range of attributes, such as pipeline diameter and capacity, compressor capacity, configuration, etc. Most of this data is not freely available. However throughout the SciGRID_gas project it was determined, that data can be found and grouped into two fundamental different groups: a) OSM data, and b) non-OSM data. The OSM data consists of geo-referenced facility data that is stored in the OpenStreetMap (OSM) data base, and is freely available. However, the OSM data set currently contains hardly any other information than the location of the facilities. The Non-OSM data set can fill some of those gaps, by supplying information such as pipeline diameter, compressor capacity and more. Part of the SciGRID_gas project is to mine and collate such data, and combine it with the OSM data set. In addition heuristic tools are required to fill data gaps, so that a complete gas network data set can be generated.

This document here describes the so called "InternetData" data set (INET), which is one of the fundamental building blocks of the SciGRID_gas project. The INET data set is the first data set to be published as part of the SciGRID_gas project. This document explains the origin and structure of the data set.

In this document, the chapter "Introduction" will supply some background information on the SciGRID_gas project, followed by the chapter "Data structure", which gives a detailed description of the data structure that is being used in the SciGRID_gas project. Chapter "Data sources" describes the INET data sets.

The appendix contains a glossary, references, location name alterations convention and finishes with the table of country abbreviation.

INTRODUCTION

SciGRID_gas is a three-year project funded by the German Federal Ministry for Economic Affairs and Energy [BMWi20] within the funding of the 6. Energieforschungsprogramm der Bundesregierung [BMWi11].

The goal of SciGRID_gas is to develop methods to generate and provide an open-source gas network data set and corresponding code. Gas transmission network data sets are fundamental for the simulations of the gas transmission within a network. Such simulations have a large scope of application, for example, they can be used to preform case scenarios, to model the gas consumption, to detect leaks and to optimize overall gas distribution strategies. The focus of SciGRID_gas will be the generation of a data set for the European Gas Transmission Network, but the principal methods will also be applicable to other geographic regions.

Both the resulting method code and the derived data will be published free of charge under appropriate open-source licenses in the course of the project. This transparent data policy shall also help new potential actors in gas transmission modelling, which currently do not possess reliable data of the European Gas Transmission Network. It is further planned to create an interface to SciGRID_power [MMK16] or heat transmission networks. Simulations on coupled networks are of major importance to the realization of the German *Energiewende*. They will help to understand mutual influences between energy networks, increase their general performance and minimize possible outages to name just a few applications.

This project was initiated, and is managed and conducted by DLR Institute for Networked Energy Systems.

1.1 Project information

- **Project title**: Open Source Reference Model of European Gas Transport Networks for Scientific Studies on Sector Coupling (*Offenes Referenzmodell europäischer Gastransportnetze für wissenschaftliche Untersuchungen zur Sektorkopplung*)

- **Acronym**: SciGRID_gas (Scientific GRID gas)

- **Funding period**: January 2018 - July 2021

- **Funding agency**: Federal Ministry for Economic Affairs and Energy (*Bundesministerium für Wirtschaft und Energie*), Germany

- **Funding code**: Funding Code: 03ET4063

- **Project partner**: DLR Institute for Networked Energy Systems

1.2 Background

As of today, only limited data of the facilities of the European Gas Transmission Networks is publicly available, even for non-commercial research and related purposes. The lack of such data renders attempts to verify, compare and validate high resolution energy system models, if not impossible. The main reason for such sparse gas facility data is often the unwillingness of transmission system operators (TSOs) to release such commercially sensitive data. Regulations by EU and other lawmakers are forcing the TSOs to release some data. However, such data is sparse and too often not clearly understandable for non-commercial users, such as scientists.

Hence, details of the gas transmission network facilities and their properties are currently only integrated in in-house gas transmission models which are not publicly available. Thus, assumptions, simplifications and the degree of abstraction involved in such models are unknown and often undocumented. However, for scientific research those data sets and assumptions are needed, and consequently the learning curve in the construction of public available network models is rather low. In addition, the commercially sensitivity also hampers any (scientific) discussion on the underlying modelling approaches, procedures and simulation optimization results. At the same time, the outputs of energy system models take an important role in the decision-making process concerning future sustainable technologies and energy strategies. Recent examples of such strategies are the ones under debate and discussion for the Energiewende [BundesregierungDeutschland20] in Germany.

In this framework, the SciGRID_gas project initiated by the research centre DLR Institute of Networked Energy Systems in Oldenburg (Germany) aims to build an open source model of the European Gas Transmission network. Releasing SciGRID_gas as open-source is an attempt to make reliable data on the gas transmission network available. Appropriate (open) licenses attached to gas transmission network data ensures that established models and their assumptions can be published, discussed and validated in a well-defined and self-consistent manner. In addition to the gas transmission network data, the Python software developed for building the model SciGRID_gas will be published under the GPLv3 license.

The main purpose of the SciGRID_gas project is to provide freely available and well-documented data on the European gas transmission network. Further, with the documentation and the Python code, users should be able to generate the data on their own computers.

The input data itself is based on data available from openstreetmap.org (OSM) under the Open Database License (ODbL) as well as Non-OSM data gathered from different sources, such as Wikipedia pages, fact sheets from TSOs or even newspaper articles.

The main workload of this project is to:

- retrieve the OSM and Non-OSM data sets for the gas infrastructure

- merge all available data sets

- build a gas transmission component data set

- generate missing data using heuristic methods

- document the process and the output.

The first step of the project was to collate a Non-OSM data set by searching the web for metadata that will be useful for the project. This included information, such as pipelines, compressors, LNG terminals, and their attributes, such as diameters, capacities etc.

Here the first single data source will be introduced, the **InternetDaten** data set (INET). At later stage, descriptions of other data sources will follow, and will be made available on the project webpage.

This multi-stage release will allow us to easily and effectively incorporate feedback from potential users during the lifetime of the project. Those releases can be downloaded from the SciGRID_gas webpage with documentation, and can be seen as a snapshot of the current research project state.

Further information on the project can be found on the SciGRID_gas web page: https://www.gas.scigrid.de/pages/imprint.html.

The web page is maintained throughout the project lifetime, and will contain information on:

- General project information

- Contact details

- Presentations

- Bug/data fixes

- Data, code and documentation releases

- Publications.

As part of the SciGRID_gas webpage, one can also sign up to the SciGRID_gas newsletter by sending an email to

news.gas-subscribe@scigrid.de

1.3 Project goal

The overall goals of the SciGRID_gas project are:

- **Data output**: Creation of customisable gas transmission network data sets.

- **Open source**: Any one can download the data, make changes to it, pass it on to others, or even use it in commercial projects, as long as the SciGRID_gas project is mentioned as the original source of the data (CC by).

- **Application**: The outcome of the project can be used for a variety of scientific applications (e.g. sector coupling, entry-exit models etc.).

- **Transparency**: The Python code, the documentation and the data (that can be passed on under copyright licences) is supplied.

- **Extendability**: Every user can extend the software code to their needs. However, we would encourage users to update and maintain the original git-repository and documentation for others.

- **Feedback**: Through constant data releases, it is hoped that the output data set will improve in quality and quantity by constantly incorporating feedback from the research community.

1.4 Document overview

This is an overview of this SciGRID_gas documentation, as this will help the user to better understand the overall project, its aims and the steps that were taken to obtain/model the resulting data set.

SciGRID_gas has been coded in Python, and hence, with that came the overall data structure that was selected for the project. As this is the most fundamental aspect for anyone wanting to use the data and the code, it is described first. In the chapter **Data Structure** we define the terms *Components*, *Elements*, and *Attributes*. We also give an overview on the internal workings of the SciGRID_gas source code.

A fundamental building block for the SciGRID_gas project is the data itself. Overall, we have classified the data into two groups: OSM and non-OSM data. The chapter **Data Sources** contains background information on InterDaten (INET) data set only. Information is supplied, such as how the data was collated and how it was implemented into the SciGRID_gas data sets structure. In addition, an overview of the extent of the data will also be given for the data, e.g. the number of elements or the list of attributes that the data contains.

The document also contains the chapter **Appendix** that contains sub-sections, such as *Glossary*, *References* etc.

DATA STRUCTURE

A well designed and documented data structure is fundamental in any large-scale project. Good data structure in combination with tools, based on algorithms, improve the performance of any project output.

This structure needs to represent the gas flow facilities as good as possible. Hence, it needs to include components, such as pipelines, compressors etc. A finite number of components have been identified, that are required as building blocks of a gas network. In addition, each component will contain attributes, such as pipeline diameter, maximal operating pressure, maximal capacity, number of turbines etc.

It is anticipated, that the adopted data structure can be implemented in different types of gas flow models and will be used by the research community for topics, such as sector coupling or identifying gas transmission bottlenecks.

Within the SciGRID_gas project, the structure of the data model is part of classes defined within the Python code. Alterations may occur over the duration of the project, but it is envisaged, that those will be small, and that compatibility will be assured.

The goal of this section is to describe in details the data structure that has been adopted and implemented into the Python code. This will be important in understanding other aspects of this document, such as exporting the data into CSV files or generating missing values.

2.1 Data structure description

This section contains information on the SciGRID_gas data structure, the format, and the code that can be used to import publicly available data into the project, so that it can be used in subsequent steps. Paramount for an understanding of the data structure is a good understanding of the terminology used throughout this section and the document in general. Hence, terminology will be introduced in the following sub-section.

2.1.1 Terminology

Throughout this document certain terms will be used, which will be described below and have been summarized in Figure 2.1.

Legend:
Component
Element
Attribute
Attribute value

gas data set

Border Point

Pipelines

Compressors

LNG terminal Rotterdam

LNGs

LNG terminal Stade

LNG terminal Amsterdam

LNG terminal Florence

Storage capacity: 450001

Gasification rate: 231

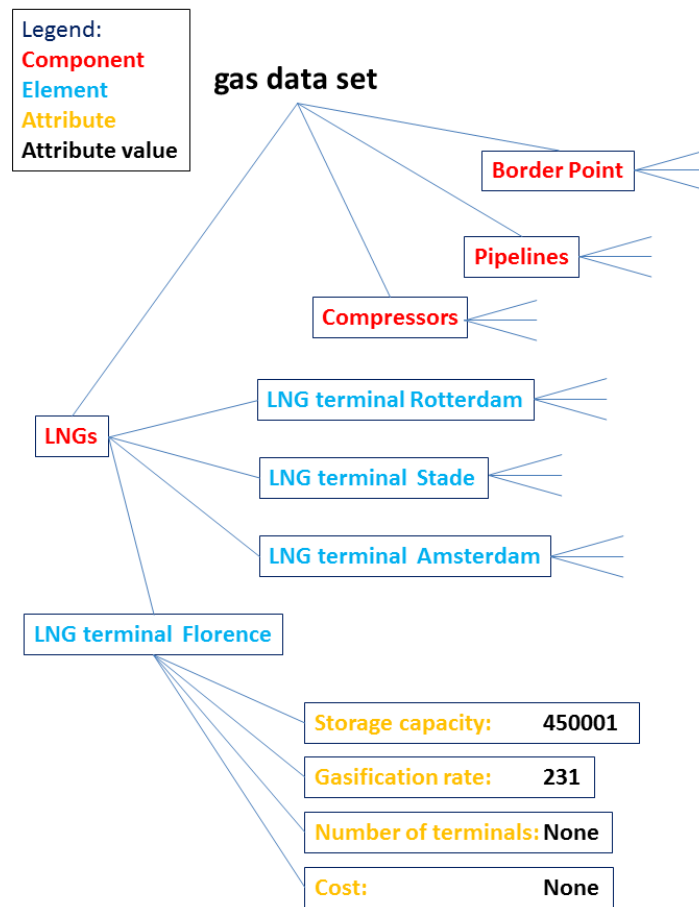**Number of terminals:** None

Cost: None

Figure 2.1: Data structure for the SciGRID_gas data set.

Gas transmission network

The term "gas transmission network" describes the physical gas transmission grid. This does not include the distribution of gas through gas distribution companies, but includes the long distance transmission of gas from producer countries to consumer countries, as carried out by the Transmission System Operators (TSO) [Wik20g]. In addition, throughout this document, the terms "transportation" and "transmission" are seen as interchangeable, and hence, will both be used describing the same.

Gas component data set

The term "gas component data set" is used for all raw data of objects/facilities that have been loaded using SciGRID_gas tools into a Python environment. Gas component data sets are used as input into our SciGRID_gas project. Several data sources can be loaded as gas component data sets, and then combined into a single gas component data set. However, not all elements (e.g. compressors) must be connected to pipelines. Hence, such a data set is referred to as a "gas component data set".

Gas network data set

A "gas component data set" can be converted into a "gas network data set", by connecting all non-pipeline elements to nodes and all nodes are connected to pipelines, and as part of the process all network islands have been connected or removed, resulting in a single network. Therefore, the network contains nodes and edges which are coherently connected, and all objects with the exception of pipelines are associated with nodes in this network, whereas pipelines are associated with edges.

Component

There are several component types in a gas transmission network, such as compressors, LNG terminals, or pipelines. In Figure 2.1 they are coloured red. Hence, whenever the word "component" is mentioned, it refers to one of these components. There are roughly a dozen different components that will form a gas network data set. They will be briefly explained below.

Element

The term "element" refers to individual facilities, e.g. the LNG terminal in Rotterdam, or the compressor in Radeland. In Figure 2.1 they are coloured blue. The first one is an element of the component LNG terminals, whereas the second one is an element of the component compressors. Hence, many elements make up a component. However, all elements are referring to different facilities by default. This means in a single network, one cannot have two elements of a component describing the same facility. The structure of elements is described below.

Attribute

"Attribute" is a term that is being used for the individual parameters that are associated with the elements. Examples of this term are gas "pipeline diameter", "maximum capacity", "max gas pipeline pressure", to name just a few and in Figure 2.1 they are coloured yellow. Overall there will be several hundred attributes in the SciGRID_gas project. However, the same attributes can occur in more than one component, e.g. "max flow capacity" exist for pipelines and also for compressors. Throughout the project, we have tried to keep the units of such attributes the same, so that there is no unit conversion required.

Attribute value

Each attribute has a value, most likely a number or a string. In Figure 2.1 they are coloured black. While booleans (*True*/*False*) are also allowed, more likely a "1" will stand for *True* and "0" for *False*. However, not all attribute values are given. Therefore, a no value attribute value needs to be specified. In the SciGRID_gas Python code it is *None*.

The Figure 2.1 depicts the relationships between the terms "gas data set", "component", "element", "attribute", and "attribute value". As can be seen, a single gas data set consists of several components. On the next level, each component contains several elements. Further, each element has several attributes, where each attribute has a single or several values. The heuristic processes described in this document at a later stage will fill all missing values with heuristically generated values.

Gas component types

A gas transmission network consists of different components, such as pipelines, compressors etc. For the SciGRID_gas project a hand-full of components have been implemented, and will be described here briefly:

- *Nodes*: In a gas network, gas flows from one point to another point, which are given through their coordinates. All elements of all other components (such as compressor stations and power plants) have an associated node, which allows for the geo-referencing of each element. Overall the term *Nodes* will be used throughout this document, as it aligns with graph theory aspects.

- *PipeLines*: *PipeLines* allow for the transmission of the gas from one node to another. *PipeLines* are georeferenced by an ordered list of nodes.

- *PipeSegments*: *PipeSegments* are almost identical to *PipeLines*. However, are only allowed to connect two nodes. Hence, any *PipeLines* element (with 3 or more nodes) can easily be converted into multiple *PipeSegments* elements.

- *Compressors*: *Compressors* represent compressor stations, which increases the pressure of the gas, and hence, allows the gas to flow from one node to another node. A gas compressor station contains several gas compressors units (turbines).

- *LNGs*: *LNGs* is the acronym for the LNG terminals and LNG storages, which there are several in Europe, as some gas gets transported to Europe via ships.

- *Storages*: *Storages* are a further network component. Surplus gas can be stored underground (e.g. in old gas fields or salt caverns), and used during low supply or high demand periods.

- *Consumers*: *Consumers* is the term used for gas users, which can be households, industry, power plants or others.

- *Production*: These can be wells inside a country where gas is pumped out of the ground. Most of the gas used in Europe comes from outside of the EU. However, there are several smaller gas production sites scattered throughout Europe.

- *BorderPoints*: *BorderPoints* are facilities at borders between countries, which are mainly used to meter the gas flow from one country to another.

Element structure

As described above, elements are describing individual facilities, such as compressors or LNG terminals. However, the overall structure of those elements is the same for all elements of all components, and is described as follows:

- *id*: A string that is the ID of the element, and must be unique.

- *name*: A string that is the name of the facility, such as "Compressor Radeland". In most cases this is not supplied.

- *source_id*: A list of strings that are the data sources of the element. As several elements from different sources could have been combined into a single element, one might need to know the original data sources.

- *node_id*: The ID of a geo-referenced node to which an element of the network is associated to. For a compressor, this will be just a single *node_id*. However, for a gas pipeline this entry would be a list of at least two *node_id* values: the starts node id and the end node id.

- *lat*: The latitude value of an element. For elements of type *PipeLines* and *PipeSegments*, *lat* is a list of latitude values. Throughout the SciGRID_gas project the projection World Geodetic system 1984 (epsg:4326) will be used.

- *long*: The longitude, analogue to lat.

- *country_code*: A string indicating the 2-digit ISO country code (Alpha-2 code, see Chapter 4.6 for list of countries and their codes) of the associated node of elements or list of nodes in case of *PipeLines* or *PipeSegments*.

- *comment*: An arbitrary comment that is associated with the element. In most cases this is not supplied.

- *tags*: This dictionary is reserved for OpenStreetMap data. It contains all associated key:value-pairs of an OpenStreetMap item.

In addition, there are three further groups of attributes to each element, which have been coded as Python "dictionaries", named:

- *param*

- *method*

- *uncertainty*.

The structure within each dictionary is the same. The dictionary *param* (short for "parameter") contains a list of attributes and their values. This list of attributes will be different for each component. For the component *PipeLines* they might be pipeline diameter, max pipeline pressure, and max pipeline capacity. For the component *Compressors* they might be, a number of turbines, overall turbine power, energy source of turbine or other.

The other two attribute dictionaries are *method* and *uncertainty*. Each of those two dictionaries contains exactly the same list of attributes as the *param* dictionary. However, their attribute values reflect the name of the dictionary. E.g. the attributes in the dictionary *method* contain the information on the method used to derive the attribute value that is stored in the *param* dictionary. Here methods of value generation can include heuristic methods names (in form of strings) that have been implemented in the SciGRID_gas project. However, if attribute values are not being generated by the SciGRID_gas project, but originate from one of the input data sources, then the attribute values in the *method* dictionary is set to "raw".

See example below, for an *LNGs* element with the following entries:

- "make_Attrib(const)": the attributes *end_year*, and *is_H_gas* have been set to a constant value

- "raw", indicating that the two attributes *max_cap_store2pipe_M_m3_per_d* and *start_year* contain original values

- "Lasso(max_cap_store2pipe_M_m3_per_d)", here for the attribute *median_cap_store2pipe_M_m3_per_d* a method was used that is based on the lasso method and uses the attribute *max_cap_store2pipe_M_m3_per_d* as input.

Similar is the content of the *uncertainty* dictionary. It contains information on the uncertainty of the attributes from the *param* dictionary of that component. Again, all attributes listed in the *param* dictionary are also present in the *uncertainty* dictionary. The attribute values here reflect the uncertainty of the attribute. Here, it is assumed, that attributes with a method of "raw" have an uncertainty of zero. Only for those attributes, which were generated during heuristic SciGRID_gas methods an uncertainty larger than zero will be specified.

2.2 Summary

The SciGRID_gas software is designed to construct a gas transmission network data set form different open and not-open source gas component data sets. The gas transmission data set needs to be available and stored in a precise and predefined way, which was described in this section. We have identified several *component*-types of a gas transmission network grid, like pipelines, compressor stations, LNG-terminals etc. Each specific facility that falls under such a component is considered an *element* of that component. Each element is described by a list of *attributes* and correspondent *attribute values*, including information on the uncertainty of the attribute value and the way the attribute value was generated.

DATA SOURCES

Original data sets describing gas transmission networks are the property of the transmission system operators (TSOs) and are generally not freely available in the form and depth that is required for modelling purposes. The major reason for the difficulty of obtaining of such data is that most of the gas network infrastructure, namely pipelines, is buried underground. Thus, a pipeline diameter is hard to estimate locally. In addition, almost all of the data is commercially sensitive.

Nevertheless, some data is made available by gas transmission network operators, through different channels. E.g. information on the size and number of compressors could be made public through a press release, as part of a refurbishment. An example is given below (https://www.maz-online.de/Lokales/Teltow-Flaeming/Neue-Verdichterstation-entsteht-in-Radeland):

> "Die Eugal-Pipeline dient dazu, Gas aus der neuen Ostseepipeline Nord Stream 2 bis zur tschechischen Grenze zu leiten. 275 Kilometer von ihr verlaufen in Brandenburg. Grundsätzlich soll die neue Leitung parallel zur bestehenden Opal-Pipeline gebaut werden."

In addition, some information can be found on company web pages, (https://www.open-grid-europe.com/cps/rde/SID-752BB6B5-E0A975F2/oge-internet-preview/hs.xsl/NewsDetail.htm?rdeLocaleAttr=en&newsId=50190C3B-E14F-4685-9E64-E40EEAB57A28):

> "Open Grid Europe (OGE) is investing roughly EUR 150 million at its compressor station in Werne to improve the security and flexibility of energy supply for North Rhine-Westphalia and Germany. The upgrade of the station, which is one of the hubs of the pipeline network, will allow gas flows to be switched (reversed) from north to south and south to north. In addition, OGE is preparing the station for the upcoming transition from L- to H-gas. Through this fitness programme, the station's transmission capacity will increase by about 500,000 to 6.5 million m^3/h, which is equivalent to the annual consumption of more than 2,100 single-family homes. The project, which is due for completion at the end of 2018, is fully on track."

However, there is a public drive to gather such data and subsequently make it available. The major platform through which this is occurring is the Open Street Map database [Hel18]. OSM is a geo-referenced database through which people can supply geo-referenced information on all man-made and natural structures, ranging from mountains to buildings. To achieve this, people throughout the world wander the globe and geo-reference everything that they can find. This also includes gas-pipeline markers, compressor stations or LNG terminals. However, the major problem remains that one cannot measure or estimate the diameter of the underground pipelines, or the number and size of the compressor turbines, as compressors are within buildings, which are fenced off. Hence, such information is hardly supplied to the OSM platform.

For the reasons mentioned above, the available data can be separated into two different groups:

- OSM data: Data can be found in the OSM data base. OSM data is well geo-referenced, but contains little meta-information (information on the facility attributes, such as pipeline diameter or pipeline capacity). OSM data is very helpful to obtain accurate routes of pipelines.

- Non-OSM data: Non-OSM data have in general lower geographical accuracy but contain a lot of meta-information. Unfortunately, such information is only known for a few facilities. One exception to this rule

are shapefiles from TSOs. They are rare, but well geo-referenced. However, the resolution of the meta information can vary from TSO to TSO.

The following section will introduce non-OSM data sets, and at a later stage, this will be followed by a section on the OSM data.

3.1 Non-OSM data

Non-OSM data includes data from internet research, TSO press releases, TSO transparency platform, TSO public data, national open-source gas network data sets[1] etc.

Some of the TSO information had to be made available due to EU-regulations. Other information has been made public as part of a company's self-presentation and advertisement. The information used by the SciGRID_gas project focuses on:

- the quality of the data

- the format of the data

- the level of representation of the data

- and the copyright restrictions on the data.

In addition, each data source is unique. Source specific tools need to be developed, so that all data sources can be made accessible for the SciGRID_gas project.

A significant portion of the project was spent on finding non-OSM data sets. Further data sources might be available, but unknown to the authors. If the authors are made aware of additional sources, the project will try to incorporate those, as this would only increase the depth of the data available and increase the applicability of the gas network data set and model.

Non-OSM data sources are very specific, addressing only certain aspects of the entire gas infrastructure. E.g. the GIE [GasIEurop20] data set supplies information on the daily gas flow in and out of gas storages in LNG terminals. However, they fall short on specifying the fundamental information of the actual physical location. Other data sets, such as the LKD [FMWP+17] data set is quite detailed in respect of pipelines, compressors and consumptions, however, only available for Germany.

Hence, the main task is to look closely at each data source, distil which data attribute values can be used, how it can be downloaded and incorporated into the SciGRID_gas model, and identify the copyright restrictions on the data source.

Due to copyright regulations, there are roughly two groups of data:

- Non-copyright restrictive data (N-CRRD): Here the copyright does not restrict the download, use and distribution of the data.

- Copyright restrictive data (CRRD): Here the data can be downloaded and used internally, but not re-distributed to others.

The following is a list of the data sources that will be used throughout the project and an identification into which group of copyright restriction they fall:

- **OSM** (https://www.openstreetmap.org) (N-CRRD)

- **GB** (https://www.nationalgridgas.com/land-and-assets/network-route-maps) (CRRD)

- **NO** (https://www.npd.no/en/about-us/information-services/available-data/map-services/) (N-CRRD)

- **LKD** (https://tu-dresden.de/bu/wirtschaft/ee2/forschung/projekte/lkd-eu) (N-CRRD)

- **ENTSOG** (https://transparency.entsog.eu/) (CRRD)

---

[1] An entire gas network data set is only available from the UK, see https://www.nationalgridgas.com/land-and-assets/network-route-maps'.

- **EMAP** (https://www.entsog.eu/sites/default/files/2020-01/ENTSOG_CAP_2019_A0_1189x841_FULL_401.pdf) (CRRD)

- **GIE** (https://www.gie.eu/) (N-CRRD)

- **GSE** (https://www.gie.eu/index.php/gie-publications/databases/storage-database) (N-CRRD)

- **IGU** (https://www.igu.org/) (CRRD)

- **INET** (see Chapter 4.4) (N-CRRD)

- **CONS** (N-CRRD).

Each data set and source comes with different copyright regulations. The copyright can be rather non-restrictive (e.g. INET) or can be restrictive (IGU). It is attempted to use only freely available data, so that such data can be re-distributed. In more restrictive data cases (IGU, GB), it is not allowed to download the data and distribute it to others. However, it is allowed to let other potential users know of the location of such data and supply them with tools that allow them to carry out the same data download and subsequent incorporation of the data into a gas network data set.

Note:

In case that other users are aware of other data sources, that might be useful to this project, please get in touch and supply us with a brief description of the data and the location of such data, so that additional tools can be developed to incorporate the data in this project. Please use the following email address: developers.gas(at)scigrid.de

3.2 The InternetDaten (INET) data set

This section contains information on the generated content and nature of the so called **InternetDaten** data set (**INET**).

The INET data set is a special data set, as it was collated from many www sources and the information has been collated into CSV files. Please note, that throughout the project, the separator within CSV files will need to be ";". This section here will give an overview on the INET data set, its components and how the data is stored in INET specific CSV files. Further the processing of the data in Python will be described.

Prior to the description of those processes, a general overview of the INET data set is given first, so that the reader has a better understanding of the size and depth of the data.

3.2.1 Overview of the INET data set

The INET data set contains geographical and meta information on gas facilities that were found through Internet searches. The data originated from www pages, such as Wikipedia, gas transmission system operators, fact sheets, press releases and more. Hence, most of the data had to be extracted manually out of text pages. To make this data available throughout the project, the data is being stored in CSV files. This also allows others to add additional properties and values to the INET data set at any stage. Tools have been written to load the INET from those CSV files and make them accessible throughout the project[1].

The Table 3.1 summarises the number of elements for each component that has been found so far. However, this does not imply that there is no missing data. In contrary, this data set comes with a lot of missing data.

Table 3.1: INET component summary.

Component Name	Count
BorderPoints	111
Compressors	249
Consumers	556
LNGs	33
Nodes	1391
PipeSegments	1050
Storages	199

In addition, a map (see Figure 3.1) visualizes these components for Europe.

3.2.2 Origin of the data

As has been stated before, the resulting INET data set originated from text sources found on the www. Here, for the pipeline JAGAL [Wik20h] an example from a Wikipedia page is given (https://en.wikipedia.org/wiki/JAGAL) is given in Figure 3.2:

As one can see, some information is given, such as location name of the compressor (Mallnow), total pipeline length (338 km), pipeline diameter (1200 mm) and maximum pipeline capacity (24 billion $\mathrm{m}^3\mathrm{a}^{-1}$). This is the information that is manually extracted from such pages and put into the CSV files.

To collate the data in an orderly manner, a system of CSV files has been created and will be described below.

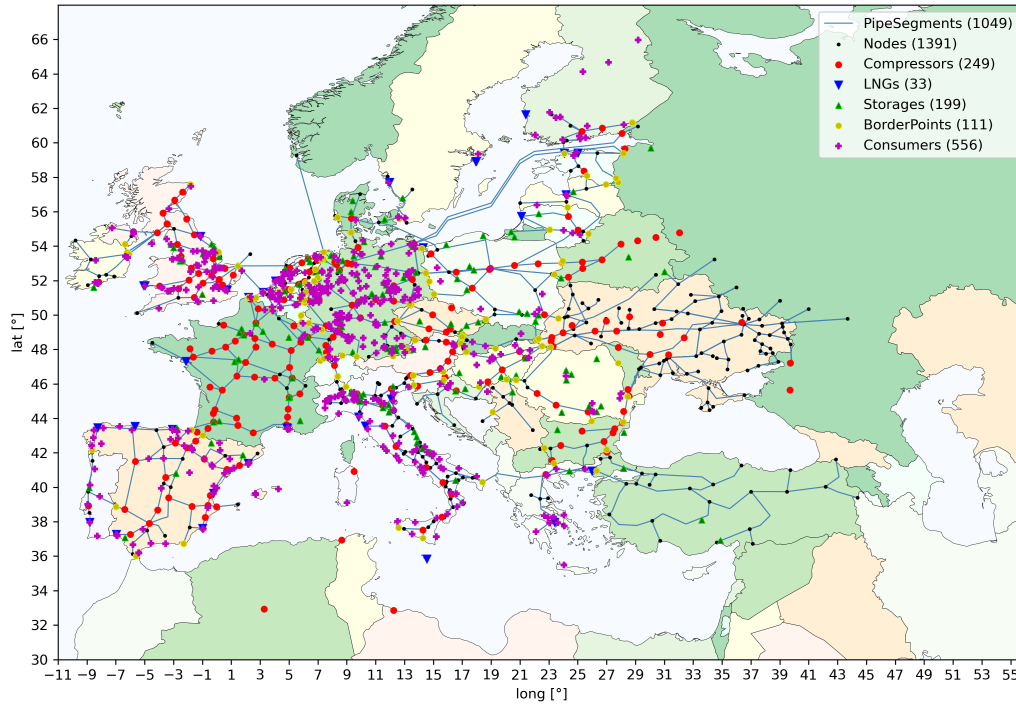[1] These tools will be made available during an upcoming release.

Figure 3.1: Map of the INET data set. The legend contains the number of elements for each component.

Figure 3.2: Screenshot of part of the Wikipedia page for the pipeline JAGAL.

3.2.3 INET CSV file description

Each component of the INET data set is represented by a single CSV file. Each of those files has a single header line, and it is very important to know, that entries in the first line should only be changed if one knows, what one is doing, as the first-row labels (the actual words) are imported and used as variable names in the SciGRID_gas project Python programs. Hence, if certain labels would be missing, the program would fail. In addition, each label needs to be unique within each file. It is advised to incorporate the units of the attributes into the label name, where possible.

Nodes.csv file

This is a unique file, and contains information on the nodes of the INET data set. Nodes are such entities, to which and from where pipelines can run, or to which other facilities can be associated to. Nodes supply information on a location including its name, its latitude and longitude, and the country in which it is located. Additionally, they supply information on the topological correctness of the lat/long values. The nodes component data is supplied to the SciGRID_gas data model only via a single CSV file, containing the following columns:

- *id*: A unique id of a node of type string. Most likely this will be the name of an element. White spaces are allowed in this string.

- *comment*: Here the user can place additional information on the location node.

- *country*: Here the user needs to write the 2 letter abbreviation of the country, in which this node is located (see Table 4.5 for a list of country codes used).

- *lat*: A number of the best estimate of the latitude of the location. Best latitude value (and long value) were attempted to be generated by using metadata of the facility node and satellite maps. Using the satellite data, address information etc., it was tried to visually find the facility of the node. Values need to be added as decimal degrees.

- *long*: This is analogue to *lat*.

- *node_id*: An identifier of a node.

- *source_id*: A unique identifier describing the source of the element. Here "INET" is the abbreviation for Inter- netDaten data set. Hence, all elements originating from the INET data set starts with the letters "INET".

- *name*: A string containing the name of the location. It is allowed to contain white spaces.

- *exact*: A number in the range of 1 to 5, indicating how accurate the lat/longs were supplied for the node. Options are as follow:

    - "1": The exact location of this node is known, as one was able to verify the facility through satellite data.

    - "2": Here the lat/long is not known exactly, however, one assumes that the location is within a small region (e.g. Krummhörn), hence, uncertainty not larger than 10 km.

    - "3": Here so little is known about the exact location, and one only knows, that the location is within a large region (e.g. Hamburg). Hence, the actual location could be out by 10 km or more, but less than 100 km.

    - "4": Here so little is known about the exact location, and one only knows, that the location is within a state (e.g. Niedersachsen). Hence, the actual location could be out by 100 km or more, but less than 1000 km.

    - "5": Here so little is known about the exact location, and one only knows, that the location is within a country (e.g. Ukraine). Hence, the actual location could be out by 1000 km or more.

All other components need two files, the location file and the metadata file, which will be described next.

Compressor CSV meta file

The compressor file ("Meta_Compressors.CSV") contains all the metadata for the known compressor stations.

In addition to the seven mandatory columns introduced above, the following columns are currently implemented:

- *end_year*: Integer number of the year, when the operation of this element stopped. If it contains a value of "None", then it is still operational.

- *start_year*: Integer number of the year, when the operation of this element started. If it contains a value of "None", then this year is not known.

- *operator_name*: A string, containing the name of the operator of the compressor station.

- *pipe_name*: A string containing the label of the pipeline that the compressor is connected to.

- *source*: Information on where the information of this element originated from.

- *is_H_gas*: A boolean, indicating if the gas is of high calorific gas type ("1") or of low calorific gas ("0").

- *max_cap_M_m3_per_h*: A number, which is the overall capacity of gas that can be compressed by the compressor station. Values need to be supplied in units of $[\mathrm{Mm^3h^{-1}}]$.

- *max_pressure_bar*: A number, which is the maximum pressure that the gas can be compressed to. Values need to be supplied in units of [bar].

- *max_power_MW*: A number, which is the sum of the power of all compressor units that are installed at the compressor station. Values need to be supplied in units of [MW].

- *num_turb*: The number of compressor turbines installed at the compressor facility. This number also includs the reserve turbine unit.

- *turbine_fuel_isGas_1*: A boolean, indicating if the turbine is powered by gas ("1"), or by electric ("0").

- *turbine_type_1*: A string containing additional information on the type of turbine unit, e.g. name of the turbine.

- *turbine_power_1_MW*: A number, indicating the power of the turbine unit. The value needs to be supplied in units of [MW].

- *turbine_fuel_isGas_2*: Information for the second turbine unit. Same as for *turbine_fuel_isGas_1* applies. Currently up to 6 individual units can be stored in the database, hence, the last digit in the identifier can be as large as 6.

- *turbine_type_2*: Information for the second turbine unit. Same as for *turbine_type_1* applies. Currently up to 6 individual units can be stored in the database, hence, the last digit in the identifier can be as large as 6.

- *turbine_power_2_MW*: Information for the second turbine unit. Same as for *turbine_power_1_MW* applies.

- ...

- *turbine_power_6_MW*: A number, indicating the power of the sixth turbine unit. The value needs to be supplied in units of [MW].

LNG CSV meta file

The LNG terminal metafile ("Meta_LNGs.CSV") contains all the metadata for the LNG terminals.

Next to the above described first seven columns the following columns are currently implemented:

- *end_year*: Integer number of the year, when the operation of this element stopped. If it contains a value of "None", then it is still operational.

- *start_year*: Integer number of the year, when the operation of this element started. If it contains a value of "None", then this year is not known. The values for this attribute dominantly originated from the King and Spalding report [KingSpalding18].

- *source*: Information on where the information of this element originated from.

- *max_workingGas_M_m3*: A number, indicating the maximum amount of liquid gas that can be stored, after having been brought in by ship. Values need to be supplied in units of $[\mathrm{Mm}^3]$ LNG. The values for this attribute dominantly originated from the King and Spalding report [KingSpalding18].

- *max_cap_store2pipe_M_m3_per_a*: A number, indicating the maximum amount of gas that can leave the LNG terminal. This gas is in gas phase. Values need to be supplied in units of $[\mathrm{Mm}^3\mathrm{a}^{-1}]$. The majority of those values originated from the EntsoB map [EntsoG20], and where missing were extracted from the King and Spalding report [KingSpalding18].

- *GCV_mean_kWh_per_m3*: This is the average Gross Calorific Value (GCV) of the gas flowing through this border point, retrieved from the Entso-G map [EntsoG20].

- *from_TSO*: This is a string of the name of the TSO that this LNG facility is associated to.

- *to_TSO*: This is a string of the name of the TSO that owns the pipe that the gas is fead into.

- *entsog_id*: A lot of the additional meta data originated from the data tables of the Entso-G map [EntsoG20]. Hence the EntsoG ID used on those tables has been place here, so that associating the meta data with the source is made easier.

- *max_vessel_size_LNG_m3*: Maximum size of ship that can land with size given in volume LNG.

BorderPoints CSV meta file

The metafile for *BorderPoints* elements ("Meta_BorderPoints.CSV") contains all the metadata for each border point.

Next to the above described first seven columns the following columns are currently implemented:

- *pipe_name*: A string, the name of the pipe that is passing the border point.

- *end_year*: Integer number of the year, when the operation of this element stopped. If it contains a value of "None", then it is still operational.

- *start_year*: Integer number of the year, when the operation of this element started. If it contains a value of "None", then this year is not known.

- *source*: Information on where some of the information of this element originated from.

- *entsog_id*: A lot of the additional meta data originated from the data tables of the Entso-G map [EntsoG20]. Hence the EntsoG ID used on those tables has been place here, so that associating the meta data with the source is made easier.

- *is_EU_to_EU*: Boolean (1 or 0) indicating if the border point is between two European Union nations or not.

- *from_country*: As the information content of the tables was flow of gas, a direction needed to be incorporated as well, to which the flow is related to. Here the from country is given.

- *to_country*: Here the to country is given, into which the gas is flowing and the gas flow numbers are related to.

- *from_TSO*: This is the name of the TSO from where the gas is being send from.

- *to_TSO*: This is the name of the receiving TSO.

- *max_cap_from_to_M_m3_per_d*: This is the maximum gas flow capacity in units of Mm^3/d, that is flowing from the from-country to the to-country.

- *max_cap_to_from_M_m3_per_d*: This is the maximum gas flow capacity in units of Mm^3/d, that is flowing from the to-country to the from-country. In case that it can be assumed that this border point allows only for uni-directional gas transmission, e.g. Nord Stream, then the value for *max_cap_to_from_M_m3_per_d* to zero. However, if the original data source was not clear enough on if there is gas flowing from the to-country to the from-country, then this field was left empty (None).

- *GCV_mean_kWh_per_m3*: This is the average Gross Calorific Value (GCV) of the gas flowing through this border point, retrieved from the Entso-G map [EntsoG20].

Storages CSV meta file

The metafile "Meta_Storages.CSV" contains all the metadata for gas storage elements within Europe.

Next to the above described first seven columns the following columns are currently implemented:

- *access_regime*: String indicating the access of the storage facility, TPA or not TPA (nTPA), and could be used for heuristic processes at a later stage.

- *end_year*: Integer number of the year, when the operation of this element stopped. If it contains a value of "None", then it is still operational.

- *start_year*: Integer number of the year, when the operation of this element started. If it contains a value of "None", then this year is not known. A value of 2050 was selected, if the site is in planing/construction, but not yet in operation.

- *store_type*: A string, indicating the type of storage, such as "Leeres Gas Feld" (empty gas field), "Salz Kaverne" (salt cavern) etc.

- *source*: Information on where the information of this element originated from.

- *is_H_gas*: A boolean, that indicates if the gas is of high calorific nature ("1") or of low calorific nature ("0").

- *is_onShore*: A number, indicating if this gas store is on land or not. Options are "1" and the gas store is on land, whereas the second option "0" indicates, that the gas store is not on land, but off shore.

- *operator_name*": String, containing the name of the operator.

- *max_workingGas_M_m3*: A number indicating the maximum amount of gas that can be stored and worked with in that gas field. Values need to be supplied in units of $[\mathrm{Mm}^3]$.

- *max_cap_store2pipe_M_m3_per_d*: A number indicating the maximum amount of gas that can move from the gas store into a gas pipe. Values need to be supplied in units of $[\mathrm{Mm}^3\mathrm{d}^{-1}]$.

- *max_cap_pipe2store_M_m3_per_d*: A number indicating the maximum amount of gas that can move from the gas pipeline into a gas store. Values need to be supplied in units of $[\mathrm{Mm}^3\mathrm{d}^{-1}]$.

Konsumers CSV meta file

The metafile "Meta_Konsumers.CSV" contains all the metadata for gas power plants within Europe. The data used here is mainly a dump from the http://globalenergyobservatory.org.

Next to the above described first seven columns the following columns are currently implemented:

- *capacity_E_MW*: Value of installed electric power output in units of [MW].

- *capacity_TH_MW*: Value of installed thermal power output in units of [MW].

- *start_year*: Integer number of the year, when the operation of this element started. If it contains a value of "None", then this year is not known. A value of 2050 was selected, if the site is in planing/construction, but not yet in operation.

- *store_type*: A string, indicating the type of storage, such as "Leeres Gas Feld" (empty gas field), "Salz Kaverne" (salt cavern) etc.

- *source*: Information on where the information of this element originated from.

- *owner*: String, containing the name of the operator.

PipeSegments CSV meta file

The metafile "Meta_PipePoints.CSV" contains all the metadata for gas elements of type *PipeSegments* within Europe.

Next to the above described first seven columns the following columns are currently implemented:

- *is_bothDirection*: A boolean with value of '1' or '0'. If set to '1', then the gas pipeline can be operated in both directions, whereas if set to '0', then the gas can only flow from the start point to the end point. Hence, here the order of the *point_lables* in the pipes file is important.

- *length_km*: The overall length of the pipeline, and NOT of the segment. The value needs to be supplied in units of [km].

- *diameter_mm*: The diameter of the pipe in units of [mm].

- *max_pressure_bar*: The maximum pressure of the gas within the gas pipeline in units of [bar].

- *max_cap_M_m3_per_d*: The maximum annual gas volume that the pipe can transmit in units of $[\mathrm{Mm^3d^{-1}}]$.

- *num_compressor*: The number of compressors along the pipeline.

- *end_year*: Integer number of the year, when the operation of this element stopped. If it contains a value of "None", then it is still operational.

- *is_H_gas*: A boolean, that indicates if the gas is of high calorific nature ("1") or of low calorific nature ("0").

- *source*: Information on where the information of this element originated from.

- *lat_mean*: Average mean latitude value of the pipe-segment.

- *lat_mean*: Average mean longitude value of the pipe-segment.

3.2.4 INET data density

"Data density" is defined as the ratio of the number of usable attribute values (not missing, e.g. filled or raw values) over the number of all elements of the component. Supposedly the INET would have two LNG terminals. One of the facilities has a known storage volume, whereas the other one does not. Hence, the data density would be 50% for the attribute storage volume. Here, the data density for the most relevant attributes will be given next for all components. At a later stage, missing values will be estimated through heuristic processes, to complete the data set.

PipeSegments elements

Overall, there are 1050 *PipeSegments* elements in the INET data set.

Table 3.2 summarizes the data densities for the most important pipe-segment attributes:

Table 3.2: INET *PipeSegments* data density

Attribute name	Data density [%]
diameter_mm	49
is_H_gas	95
is_bothDirection	8
length_km	100
max_cap_M_m3_per_d	13
max_pressure_bar	26
num_compressor	3

Compressors elements

Overall, there are 249 *Compressors* elements in the INET data set. The data densities for the most important attributes is given in Table 3.3 below:

Table 3.3: INET *Compressors* data density

Attribute name	Data density [%]
is_H_gas	99
max_cap_M_m3_per_d	7
max_power_MW	16
max_pressure_bar	7
num_turb	15
operator_name	11
pipe_name	10
turbine_power_1_MW	15
turbine_power_2_MW	147
turbine_power_3_MW	9
turbine_power_4_MW	3
turbine_power_5_MW	1
turbine_power_6_MW	0
turbine_fuel_isGas_1	14
turbine_fuel_isGas_2	14
turbine_fuel_isGas_3	10
turbine_fuel_isGas_4	7
turbine_fuel_isGas_5	1
turbine_fuel_isGas_6	0
turbine_type_1	9
turbine_type_2	9
turbine_type_3	6
turbine_type_4	3
turbine_type_5	1
turbine_type_6	0

Nodes elements

Overall, there are 1394 nodes. As described above, the information supplied is an "id", latitude and longitude values, the country code and a value indicating the accuracy of the node location. Hence, Table 3.4 summarizes the relative number of nodes within the possible value range of 1 to 5:

Table 3.4: Summary for the attribute *exact* of component *Nodes* of the INET data set.

Exact value	ration of data with exact value [%]
1	60
2	27
3	5
4	3
5	4

Storages elements

Overall, there are 199 storage elements in the INET data set. The data densities for the most important attributes is given in Table 3.5 below:

Table 3.5: INET *Storages* data density

Attribute name	Data density [%]
access_regime	90
is_H_gas	17
is_onShore	96
max_cap_pipe2store_M_m3_per_d	77
max_cap_store2pipe_M_m3_per_d	77
max_workingGas_M_m3	81
operator_name	99
source	95
store_type	94

Consumers elements

Overall, there are 556 consumer elements in the INET data set. The data densities for the most important attributes is given in Table 3.6 below:

Table 3.6: INET *Consumers* data density

Attribute name	Data density [%]
est_generation_GWh	100
capacity_E_MW	100
is_gas_fuel1	100
PowerPlantType	34
capacity_TH_MW	8
is_gas_fuel2	9
is_gas_fuel3	0
is_gas_fuel4	0

A large portion of this data was derived from the Global Energy http:/globalenergyobservatory.org.

BorderPoints elements

Overall, there are 111 *BorderPoints* elements in the INET data set. The data densities for the most important attributes are given in Table 3.7 below:

Table 3.7: INET *BorderPoints* data density

Attribute name	Data density [%]
GCV_mean_kWh_per_m3	80
entsog_id	98
is_EU_to_EU	100
from_country	100
to_country	100
max_cap_from_to_M_m3_per_d	96
max_cap_to_from_M_m3_per_d	93
from_TSO	94
to_TSO	92

LNGs elements

Overall, there are 33 *LNGs* elements in the INET data set. The data densities for the most important attributes is given in Table 3.8 below:

Table 3.8: INET *LNGs* data density

Attribute name	Data density [%]
max_cap_store2pipe_M_m3_per_d	91
max_vessel_size_M_m3	67
max_workingGas_M_m3	94
GCV_mean_kWh_per_m3	64
entsog_id	67
from_TSO	67
to_TSO	67

Overall, a lot of data has been collated and is made available through the INET data set. However, as presented in the data density tables, a significant number of attributes have low data density. The following chapter in this document will demonstrate how missing values can be estimated, so that the generated SciGRID_gas data set has a data density of 100 % for each attribute.

3.2.5 Copyright and disclaimer for the INET data set

Copyright

the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons. org/licenses/by/4.0/.

A list of the sources used for the generation of the INET data set can be found in Chapter 4.4.

Disclaimer

The INET data set is supplied on a best-effort basis only, using available information as documented gathered from the Internet. While every effort is made to make sure the information is accurate and up-to-date, we do not accept any liability for any direct, indirect, or consequential loss or damage of any nature–however caused–which may be sustained as a result of reliance upon such information.

3.2.6 Improvements to previous release

Below major changes, that occured between version 1.0 and 1.1 are listed:

- The component *BorderPoints* was filled with additional information from the EntsoG map [EntsoG20].
- The component *LNGs* was filled with additional information from the EntsoG map [EntsoG20].
- Additional pipelines were added for the country of Ukrain.
- Additional pipelines were added for the country of Turkey.

3.3 Data summary

SciGRID_gas is based on open source data. To generate a gas pipeline network data set, one needs to access different data sets that were found throughout the project and presented here. Emphasis was given to depict the number of elements per component and the data density for each data set.

3.4 Summary

Gas component data sets come in different forms, licenses, formats and detail. The SciGRID_gas project can process such data and combine them to a consistent and reliable network data set.

The underlying gas component data sets were categorized into two different groups:

- OSM data: This is data originating from the OSM data base, containing well geo-referenced locations of gas facilities, such as pipe locations or gas storage facilities. However, it comes with very few meta information.

- Non-OSM data: These are all other data sources, which can "supply" detailed information on some of the gas facilities attributes. However, this information is sparse, as published only for a few facilities. Here, the INET data set was introduced as an example of the non-OSM data set, and the pathway of converting the raw data from the www into SciGRID_gas project component structure.

Here detailed information on one or several data sources have been given, and should be used as a reference for later data processes.

FOUR

APPENDIX

4.1 Glossary

Dataset abbreviations can be found in Table 4.1.

Table 4.1: Dataset abbreviations

Name	Abbreviation	Description
Raw InternetDaten data set	INET	Label/name for the raw InternetDaten data set
Raw Gas Infrastructure Europe data set	GIE	Label/name for the raw Gas Infrastructure Europe data set
Raw Gas Storage Europe data set	GSE	Label/name of the raw Gas Storage Europe data set
Raw Norwegian data set	NO	Label/name for the raw Norwegian data set
Raw Long-term planning and short-term optimization data set	LKD	Label/name for the raw Long-term planning and short-term optimization data set
Raw International Gas Union data set	IGU	Label/name for the raw International Gas Union data set
Raw EntsoG-Map data set	EMAP	Label/name for the raw EntsoG-Map data set
Merged and filled IGG data set	IGG	Filled data sets, for which the **INET**, **G**IE and **GSE** data sets were merged
Merged and filled IGGI data set	IGGI	Filled data sets, for which the **INET**, **G**IE, **GSE** and **IGU** data sets were merged
Merged and filled IGGIN data set	IGGIN	Filled data sets, for which the **INET**, **G**IE, **GSE**, **IGU** and the **NO** data sets were merged
Merged and filled IGGINL data set	IGGINL	Filled data sets, for which the **INET**, **G**IE, **GSE**, **IGU**, **NO** and the **LKD** data sets were merged
Merged and filled IGGIELGN data set	IGGIELGN	Filled data sets, for which the **INET**, **G**IE, **GSE**, **IGU**, **EMAP**, **LKD**, **G**B, and the **NO** data sets were merged

The glossary terms can be found in Table 4.2.

Table 4.2: Glossary

Name	Abbreviation	Description
component		A gas network consists of different components, such as: pipelines, compressors, LNG terminals, storages, entry points and production sites
element		Elements are instances of component. Hence, "10 compressor elements" refers to a data set that contains information for 10 compressor stations
attribute		Gas facilities, such as pipelines or compressors, can be described with a large set of parameters, such as pipeline diameter, or compressor capacity. Those parameters are referred to as attributes
facility		General term used for a gas appliance, such as a single compressor station, or a single LNG terminal
PipeLine		A gas pipeline entity, which has one start and one end point, however, can run via many nodes
PipeSegment		A gas pipeline, that has only one start and one end point, but no nodes in-between
LNG	LNG	Liquefied natural gas
CNG	CNG	Compressed natural gas
flow duration curve	FDC	It is the cumulative frequency curve that shows the percent of time specified flow were equal or exceeded during a given period. The temporal information, when certain events occur, is lost
Energiewende		German term for the change in using primary energies, the move away from coal to renewable energies, such as wind or solar
gas component data set		Raw input data, associated with components of the gas transmission grid
gas network data set		Output data, a coherent network of gas transmission components
OSM	OSM	Data that is available from the openstreetmap.org
non-OSM	Non-OSM	Data that is not part of the OSM data set
gas type		There are two types of gas High (H) and Low (L) calorific gas
mean absolute error	MAE	mean difference between input values and estimated values
data density		The ratio of the number of usable (not missing) attribute values over number elements of the component, in units of [%]
Transmission System Operators	TSO	An entity entrusted with the transportation of natural gas/electricity, as defined by the European Union
gas transmission network		This describes the physical gas transmission grid, however, it excludes any facilities/components that would be part of a distribution network and their facilities
gas component data set		The term "gas component data set" is used for raw data sets of gas network facilities. However, not all elements (e.g. compressors) need to be connected to pipelines, where the emphasis is on the term **component**
gas network data set		A "gas component data set" can be converted into a "gas network data set", by connecting all non-pipeline elements to nodes and all nodes are connected to pipelines. Hence, the emphasis here is on the term **network**

4.2 Unit conversions

Table 4.3: Unit conversions

From Unit	To Unit	MultiVal
LNG Mt	LNG Mm^3	2.47
gas tm^{3}h^{-1}	gas Mm^{3}d^{-1}	24/1000
LNG Mm^3	gas Mm^3	584
LNG t	gas Mm^3	1442.48
GWh (H)	gas Mm^3	0.0879757777
GWh (L)	gas Mm^3	0.1023541453

For some elements of some components, the calorific value was given through the references. Hence during the conversion process from GWh to M m^3, the elements calorific value was used, however, whereever the element specific calorific value was not known, the default values from Table 4.3 was used in dependence of the gas type of the element. If no gas type was known, then high calorific gas is assumed.

4.3 Attribute *exact*

Each element of type *Nodes* has an attribute *exact*. With this, the SciGRID_gas project is trying to let the user know, how well the actual location of the *Nodes* elements are known. The actual location (latitude longitude pair) can be spot on (verifiable through satellite imagery) or can be unknown by 10's or 100's of km, where city names or country names are known only. Here the attribute value for *exact* is being given, ranging from "1" to "5" as listed in Table 4.4 below.

Table 4.4: Unit conversions

Exact value	Description	Uncertainty [km]
1	The exact location of this node is known, as one was able to verify the facility through satellite data.	0
2	Here the lat/long is not known exactly. However, one assumes that the location is within a small region (e.g. Krummhörn). Hence, not being much larger than 10 km	10
3	Here so little is known about the exact location, and one only knows, that the location is within a large region (e.g. Hamburg). Hence, the actual location could be out by 10 km or more but less than 100 km	100
4	Here so little is known about the exact location, and one only knows, that the location is within a state (e.g. Niedersachsen). Hence, the actual location could be out by 100 km or more but less than 1000 km	1000
5	Here so little is known about the exact location, and one only knows, that the location is within a country (e.g. Ukraine). Hence, the actual location could be out by 1000 km or more.	> 1000

4.4 References for INET data set

Below a list of those sources used to generate the INET data set.

- http://belarus-tr.gazprom.ru/
- http://corporate.vattenfall.com/about-vattenfall/operations/market-transparency/gas-storage/
- http://en.gaz-system.pl//
- http://gaslager.energinet.dk/EN/Pages/default.aspx
- http://interfaxenergy.com/article/19138/lng-better-than-norway-pipeline-ex-polish-pm
- http://ir.gasplus.it/home/show.php?menu=00002
- http://italgasstorage.it/eng/progetto.html
- http://media.edfenergy.com/Misc/AboutUs.aspx
- http://mmbf.hu/en/company/gas_storage
- http://mndgsgermany.com/
- http://mysolar.cat.com/cda/files/870985/7/Solar_Turbines_5000_Gas_Compressor_News_Rel_Sent.pdf
- http://sse.com/whatwedo/ourprojectsandassets/thermal/Aldbrough/
- http://sse.com/whatwedo/wholesale/gasstorage/
- http://www.bayernugs.de/4-1-Home.html
- http://www.bendisenergy.com.tr/tr/Projelerimiz/24-toren-dogalgaz-depolama-ve-madencilik-as
- http://www.berliner-erdgasspeicher.de/en/Pages/default.aspx
- http://www.botas.gov.tr/
- http://www.bulgartransgaz.bg/en/pages/transstorge-110.html
- http://www.caythorpegasstorage.com/caythorpe/
- http://www.centrica-sl.co.uk/
- http://www.ceskaplynarenska.cz/en/ultimate-speed-underground-gas-storage
- http://www.dea-speicher.de/en
- http://www.depomures.ro/
- http://www.edisonstoccaggio.it/en
- http://www.ekb-storage.de/de/home/
- http://www.emplpipeline.com/en/the-gas-pipeline/
- http://www.emplpipeline.com/en/the-gas-pipeline/,Extra
- http://www.enagas.es/enagas/en/Transporte_de_gas/Almacenamientos_Subterraneos
- http://www.energystock.com/
- http://www.ewe-gasspeicher.de/english/index.php
- http://www.fluxys.com/belgium/en/Services/Storage/Storage
- http://www.gasnaturalfenosa.com/en/activities/lines+of+business/1285338591925/supply+and+transportation+of+gas.html
- http://www.gasspeicher-hannover.de/startseite.html

- http://www.gasstorage.cz/en/operation-information/available-firm-capacity/

- http://www.gasstoragebergermeer.com/gas-storage-bergermeer-2/

- http://www.gastrade.gr/en/the-company/the-project.aspx

- http://www.gas-union-storage.de/

- http://www.gatewaystorage.co.uk/

- http://www.gazprom.com/about/production/underground-storage/

- http://www.geogastock.it/ITA/Home.asp

- http://www.grtgaz.com/en/major-projects/beynes-compressor-station/presentation/news/
compressor-station-at-the-beynes-site.html

- http://www.grtgaz.com/fileadmin/plaquettes/en/2017/Essentiel-plaquette-institutionnelle-EN-2017.pdf

- http://www.gsa-services.ru/

- http://www.halite-energy.co.uk/our-project/project-overview/

- http://www.hradf.com/en/portfolio/south-kavala-natural-gas-storage

- http://www.humblyenergy.co.uk/about-us#useful-information

- http://www.islandmageestorage.com/

- http://www.kge-gasspeichergesellschaft.de/

- http://www.kgsp.co.uk/

- http://www.kingstreetenergy.com/

- http://www.kinsaleenergy.ie/gas-storage.html

- http://www.le.lt/index.php/projects-in-progress/syderiai-underground-gas-storage/535

- http://www.lg.lv/index.php?id=3376&lang=eng

- http://www.magyarfoldgaztarolo.hu/en/Lapok/default.aspx

- http://www.mnd.eu/en/2014-11-26-12-13-02/mnd-group-companies

- http://www.nafta.sk/en/about-gas-storage

- http://www.nam.nl/en/about-nam/facts-and-figures.html

- http://www.omv.com/portal/01/com/gas/storage

- http://www.petroceltic.com/operations/bulgaria.aspx

- http://www.pgnig.pl/reports/annualreport2012/en/ar-obrot-magazynowanie-2.html

- http://www.pozagas.sk/en/?PHPSESSID=8d83cc7890abb7980f6793cf56633a72

- http://www.psp.hr/home

- http://www.rag-energy-storage.at/en.html

- http://www.romgaz.ro/en/content/ugs-n-366-new-gas-storage-facility-romania

- http://www.romgaz.ro/en/content/ugs-n-371-sarmasel-storage-facility-upgrading

- http://www.romgaz.ro/en/inmagazinare

- http://www.rwe.com/web/cms/de/37110/rwe/presse-news/pressemitteilungen/?pmid=4005467,http:
//digitalnewsservice.net/clients/net4gas-nimmt-neue-hochdruck-gas-pipeline-gazelle-in-betrieb/

- http://www.rwe.com/web/cms/en/531750/rwe-gasspeicher/

- http://www.scottishpower.com/pages/hatfield_moor_gas_storage_facility.asp

- http://www.snam.it/en/transportation/Thermal_Year_Archive/Thermal_Year_2013_2014/Info-to-users/index.html

- http://www.sppstorage.cz/#

- http://www.srbijagas.com/naslovna.1.html

- http://www.stogit.it/en/index.html

- http://www.storengy.com/countries/deutschland/en/products-services.html

- http://www.storengy.com/countries/france/en/

- http://www.storengy.com/countries/unitedkingdom/en/oursites.html

- http://www.taqaglobal.com/our-regions/netherlands/gas-storage/peak-gas-installation/overview?sc_lang=en

- http://www.terranets-bw.de/en/gas-transmission/we-transport-your-natural-gas/

- http://www.thueringerenergie.de/Unternehmen/Ueber_uns/Geschaeftsfelder/Speicher.aspx

- http://www.tigf.fr/en/what-we-can-offer/storage.html

- http://www.tpao.gov.tr/eng/?tp=m&id=31

- http://www.tpao.gov.tr/eng/?tp=m&id=84

- http://www.trianel-gasspeicher.com/

- http://www.ugs-katharina.de/en/unternehmen.html

- http://www.vng-gasspeicher.de/content/en/Speicher/index.html

- https://de.wikipedia.org/wiki/Baumgarten_an_der_March

- https://de.wikipedia.org/wiki/Erdgasleitung_Jamal%E2%80%93Europa,https://en.wikipedia.org/wiki/Yamal%E2%80%93Europe_pipeline

- https://de.wikipedia.org/wiki/Hungaria-Austria-Gasleitung

- https://de.wikipedia.org/wiki/MIDAL,https://www.gascade.de/netzinformationen/unser-leitungsnetz/midal/

- https://de.wikipedia.org/wiki/Mittel-Europ%C3%A4ische_Gasleitung,https://en.wikipedia.org/wiki/MEGAL_pipeline

- https://de.wikipedia.org/wiki/NEL_(Pipeline),https://en.wikipedia.org/wiki/NEL_pipeline,https://www.fluxys.com/nel/en/NELSystemInfo/AboutNEL

- https://de.wikipedia.org/wiki/Norddeutsche_Erdgas-Transversale,https://en.wikipedia.org/wiki/Netra

- https://de.wikipedia.org/wiki/OPAL_(Pipeline),https://en.wikipedia.org/wiki/OPAL_pipeline,https://www.opal-gastransport.de/netzinformationen/ostsee-pipeline-anbindungsleitung/

- https://de.wikipedia.org/wiki/Penta-West

- https://de.wikipedia.org/wiki/Rehden-Hamburg-Gasleitung,https://en.wikipedia.org/wiki/Rehden%E2%80%93Hamburg_gas_pipeline

- https://de.wikipedia.org/wiki/Trans_Austria_Gasleitung,https://en.wikipedia.org/wiki/Trans_Austria_Gas_Pipeline

- https://de.wikipedia.org/wiki/Trans-Adria-Pipeline,https://en.wikipedia.org/wiki/Trans_Adriatic_Pipeline

- https://de.wikipedia.org/wiki/Trans-Europa-Naturgas-Pipeline

- https://de.wikipedia.org/wiki/Transgas-Pipeline,Extra

- https://de.wikipedia.org/wiki/WEDAL,https://www.gascade.de/netzinformationen/unser-leitungsnetz/wedal/

- https://de.wikipedia.org/wiki/West-Austria-Gasleitung

- https://en.wikipedia.org/wiki/Arad%E2%80%93Szeged_pipeline

- https://en.wikipedia.org/wiki/BBL_Pipeline

- https://en.wikipedia.org/wiki/BRUA_Pipeline

- https://en.wikipedia.org/wiki/Europipe_II

- https://en.wikipedia.org/wiki/Giurgiu%E2%80%93Ruse_pipeline

- https://en.wikipedia.org/wiki/MEGAL_pipeline,https://de.wikipedia.org/wiki/Mittel-Europ%C3%A4ische_Gasleitung

- https://en.wikipedia.org/wiki/National_Transmission_System

- https://en.wikipedia.org/wiki/Scotland-Northern_Ireland_pipeline,E127

- https://en.wikipedia.org/wiki/South_Wales_Gas_Pipeline

- https://en.wikipedia.org/wiki/Transitgas_Pipeline,https://web.archive.org/web/20070808033932/http://www.swissgas.ch/en/3_2.php

- https://en.wikipedia.org/wiki/Transitgas_Pipeline,transitgassystem_en.gif,https://web.archive.org/web/20070808033932/http://www.swissgas.ch/en/3_2.php

- https://en.wikipedia.org/wiki/V%C3%A1rosf%C3%B6ld%E2%80%93Slobodnica_pipeline

- https://www.fnb-gas.de/media/FNB_GAS_Projekte_2014_02_17_anlage_6_nep-gas-2014_projekt-steckbriefe.pdf

- https://globallnghub.com/wp-content/uploads/2018/09/King.pdf

- https://globallnghub.com/wp-content/uploads/2018/09/King.pdf

- https://news.err.ee/610905/vopak-to-build-initially-4-000-cubic-meter-lng-terminal-at-muuga

- https://globallnghub.com/wp-content/uploads/2018/09/King.pdf

- https://www.lngworldnews.com/tallinna-sadam-alexela-to-work-on-paldiski-lng-terminal/

- https://globallnghub.com/wp-content/uploads/2018/09/King.pdf

- https://www.sourcewatch.org/index.php/Delimara_Malta_LNG_Terminal

- https://theodora.com/pipelines/france_and_belgium_pipelines.html

- https://theodora.com/pipelines/france_and_belgium_pipelines.html, E21

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E10

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E106

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E11

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E112

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E113

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E13

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E14

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E16_1

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E16_2

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E17_Skikda

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E22_Arzew_1

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E22_Arzew_2

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E48

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E64

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E65

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E78

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E80_1

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E80_2

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E81

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E85

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E86

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E87

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E88

- https://theodora.com/pipelines/france_and_belgium_pipelines.html,E91

- https://transparency.entsog.eu/

- https://www.astora.de/storage-locations/haidach-storage-facility.html?L=1

- https://www.astora.de/storage-locations/jemgum-storage-facility.html?L=1

- https://www.astora.de/storage-locations/jemgum-storage-facility.html?L=2

- https://www.astora.de/storage-locations/rehden-storage-facility.html?L=1

- https://www.baltic-pipe.eu/de/das-projekt/

- https://www.enagas.es/stfls/ENAGAS/Transporte%20de%20Gas/Documentos/CAT_English.pdf

- https://www.enbw.com/unternehmen/konzern/geschaeftsfelder/speicher/gasspeicher/zahlen-daten-fakten.html

- https://www.eneco.nl/over-ons/projecten/gasspeicher/voorraadniveau

- https://www.energiefachmagazin.de/Branchen-News/VNG-Gasspeicher-GmbH-legt-Erdgasspeicher-Buchholz-still

- https://www.enovos.de/industrie/ueber-uns/erdgasspeicher

- https://www.eugal.de/eugal-pipeline/,https://www.mdr.de/nachrichten/politik/regional/baubeginn-erdgastrasse-eugal-umstritten-100.html

- https://www.europipe.com/de/referenzen/referenzprojekte/

- https://www.eustream.sk/en_transmission-system/en_transmission-system

- https://www.fluxys.com/belgium/en/about%20fluxys/infrastructure/network/network

- https://www.gascade.de/en/our-network/compressor-stations/radeland/

- https://www.gascade.de/en/our-network/our-pipelines/jagal/

- https://www.gascade.de/netzinformationen/unser-leitungsnetz/stegal/

- https://www.gascade.de/netzinformationen/unser-leitungsnetz/stegal/,https://en.wikipedia.org/wiki/STEGAL, https://de.wikipedia.org/wiki/STEGAL

- https://www.gascade.de/presse/presseinformationen/pressemitt,http://www.friedrich-vorwerk.de/de/aktuell/projekte/neubau-nowal-dn-1000.htmleilung/news/bau-der-nord-west-anbindungsleitung-nowal-startet-anfang-maerz/

- https://www.gasinfocus.com/en/indicator/existing-and-planned-lng-terminals/

- https://nauticor.de/lng-terminal-nynaeshamn

- https://www.gazprom.com/f/posts/86/569604/portovaya_eng.pdf

- https://www.habau.at/de/projekte/erdgasleitung-wag-expansion-3

- https://www.habau.at/de/projekte/erdgasleitung-wag-plus-600-phase-1

- https://www.hydrocarbons-technology.com/projects/swedegas-lng-facility-port-gothenburg/

- https://www.ign.ren.pt/en/armazenamento-subterraneo3

- https://www.ign.ren.pt/en/armazenamento-subterraneo4

- https://www.innogy-gasstorage.cz/en/index/

- https://www.innogy-gasstorage.cz/en/stramberk/

- https://www.n-ergie.de/geschaeftskunden/produkte/erdgas/erdgasspeicher.html

- https://www.ontras.com/fileadmin/user_upload/Dokumente_Download/Publikationen/ONTRAS_Netzpufferanlage_Burggraf_Bernsdorf.pdf

- https://www.osm.pgnig.pl/en

- https://www.rnf.de/hintergrund-die-erm-gasleitung-von-ludwigshafen-nach-karlsruhe-58182/

- https://www.shz.de/lokales/landeszeitung/wissenschaftler-unter-zeitdruck-id5845016.html

- https://www.sourcewatch.org/index.php/Lithuania-Latvia_Interconnection_Gas_Pipeline

- https://www.storengy.com/countries/france/en/nos-sites/beynes.html

- https://www.storengy.com/countries/france/en/nos-sites/cere-la-ronde.html

- https://www.storengy.com/countries/france/en/nos-sites/manosque.html

- https://www.storengy.com/countries/france/en/our-sites/saint-clair-sur-epte.html

- https://www.swedegas.com/Our_services/services/Storage

- https://www.uniper-energy-storage.com/cps/rde/xchg/ust/hs.xsl/3252.htm?rdeLocaleAttr=en

- https://www.uniper-energy-storage.com/cps/rde/xchg/ust/hs.xsl/3437.htm?rdeLocaleAttr=en

- https://www.vng-gasspeicher.de/storage_locations

- https://www.wesernetz.de/netznutzung/bremen/gasnetz-speicheranlagen.php

- https://www.wingas.com/storage-uk-ltd/home.html

- https://www.PLE

- https://www.www.gascade.de/Verdichterstation_Bunde_2016.pdf

- https://www.www.gascade.de/Compressor_station_Eischleben_2016.pdf

- https://www.www.gascade.de/Verdichterstation_Lippe_72dpi060614.pdf

- https://www.www.gascade.de/Anlandestation_Greifswald_Lubmin_2016.pdf

- https://www.www.gascade.de/Verdichterstation_Mallnow_2016.pdf

- https://www.www.gascade.de/Compressor_station_Olbernhau_2016.pdf

- https://www.www.gascade.de/Verdichterstation_Reckrod_2016.pdf

- https://www.www.gascade.de/Verdichterstation_Rehden_2016.pdf

- https://www.www.gascade.de/Compressor_station_Rueckersdorf_2016.pdf

- https://www.www.gascade.de/Verdichterstation_Weisweiler_201609.pdf

- https://www.opal-gastransport.de/Compressor_Station_Radeland_72dpi.pdf

- https://www.opal-gastransport.de/Verdichterstation_Radeland_2016.pdf

- https://www.pipelinesystems.com/: "NETRA compressor station Wardenburg"

- https://www.grtgaz.com: Plan_decennal_2017-2026.pdf

- https://www.streicher.de

- https://www.open-grid-europe.com

- https://www.open-grid-europe.com

- https://www.open-grid-europe.com, Pressinformation, May 2017, "Herbstein compressor station new build project"

- https://www.fluxys.com/tenp/de

- Presseinfo Bayernets, WinGas, 19-Sep-2008

- NWZonline.de, 22-04-2010: "ExxonMobil gibt kräftig Gas" by Tanja Mikulski

- Porzerleben.de: 6-sep-2011, "Open Grid Europe investiert in europäischen Netzverbund"

- Christoph Edler, Bachelorarbeit PR 370005, Technische Universität Wien, "Das österreichische Gasnetz", Juli 2013.

4.5 Location name alterations

Location names should be changed into the 26 letters used in the English language.

For names from the individual countries please follow the suggested approach:

- Germany/Austria: *Umlaute* to be replaced with the letter followed by an 'e', e.g.: ü = ue.

- France/Belgium: Omit accent de gues and accent de graphs, e.g.: ó = o.

- Sweden: Please change the last three letters of the Swedish alphabet and replace e.g.: ä = a.

- Poland: Please change any letter, that cannot be found in the English alphabet, knowing that for some letters, that one can only use a single letter instead of the three different letters used in the Polish alphabet, e.g.: z = z.

- Spain/Portugal: Please change any letter, that cannot be found in the English alphabet, e.g.: ñ = n.

- Greece: Please do not use Greek letters. Please try to write the Greek words with Latin letters.

- Denmark: Please change any letter that contains non-English letters, e.g.: "å" with "aa".

- Slovakia, Czech Republic, Hungary, Rumania, Latvia, Lithuania, Estonia, Bulgaria, Slovenia, Croatia: PLEASE use your common sense, based on the examples from the other countries above.

4.6 Country name abbreviations

For convenience we provide a short list of names and two-digit codes (see Table 4.5) for the probably most important countries associated with the European Transmission Grid.

Table 4.5: Country codes

Country name	Country code	Country name	Country code
Albania	AL	Kosovo	XK
Armenia	AM	Latvia	LV
Austria	AT	Liechtenstein	LI
Azerbaijan	AZ	Lithuania	LT
Belarus	BY	Luxembourg	LU
Belgium	BE	Malta	MT
Bosnia and Herzegovina	BA	Moldova	MD
Bulgaria	BG	Montenegro	ME
Croatia	HR	Netherlands	NL
Cyprus	CY	Norway	NO
Czech	CZ	Poland	PL
Denmark	DK	Portugal	PT
Estonia	EE	Romania	RO
Finland	FI	Serbia	RS
France	FR	Slovakia	SK
Georgia	GE	Slovenia	SI
Germany	DE	Spain	ES
Greece	GR	Sweden	SE
Hungary	HU	Switzerland	CH
Iceland	IS	Turkey	TR
Ireland and Northern Ireland	IE	Belarus	UA
Italy	IT	Great Britain	GB
Russia Federation	RU	Europe	EU
Ukrain	UA		

4.7 Acknowledgement

We acknowledge the contribution of Dr. Ontje Luensdorf from the DLR Institute of Networked Energy System to the SciGRID_gas project.

[AFW14]   M. Ahmed, B.T. Fasy, and C. Wenk. *New Techniques in Road Network Comparison*. Penguin Random House, New York, NY, 2014.

[AG99]   H. Alt and L. Guibas. *Discrete geometric shapes: matching, interpolation, and approximation-a survey*. Sack JR, Urrutia J, Handbook of Computational Geometry, Elsevier, New York, NY, 1999.

[DPDM20]   J.C. Diettrich, A. Pluta, J. Dasenabrock, and W. Medjroub. *SciGRID_gas: The combined IGGIN gas transmission network data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, 2020.

[DPM20a]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The INET gas transmission network data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, Aug 2020. URL: https://doi.org/10.5281/zenodo.4008975, doi:10.5281/zenodo.4008975.

[DPM20b]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The combined IGG gas transmission network data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, Aug 2020. URL: https://doi.org/10.5281/zenodo.4009129, doi:10.5281/zenodo.4009129.

[DPM20c]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The combined IGGI gas transmission network data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, Aug 2020.

[DPM20d]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The combined IGGINL gas transmission network data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, 2020.

[DPM20e]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The raw INET data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, Aug 2020. URL: https://doi.org/10.5281/zenodo.3985249, doi:10.5281/zenodo.3985249.

[DPM20f]   J.C. Diettrich, A. Pluta, and W. Medjroub. *SciGRID_gas: The raw LKD data set*. DLR-Institut für Vernetzte Energiesysteme e.V., Oldenburg, Germany, Aug 2020. URL: https://doi.org/10.5281/zenodo.3985271, doi:10.5281/zenodo.3985271.

[FMWP+17]   Kunz F., Kendziorski M., Schill W.-P., Weibezahn J., Zepter J., von Hirschhausen C., and Hauser P. *Electricity, Heat, and Gas Sector Data for Modeling the German System*. Deutsches Institut für Wirtschaftsforschung, Daten Documentation 92, Berlin, 2017.

[Hel18]   D. Helle. OpenStreetMap - Deutschland. https://www.openstreetmap.de/, 2018. Accessed: 2019-12-12.

[Kha13]   Y. Khalid. What is Pickle in python? https://pythontips.com/2013/08/02/what-is-pickle-in-python/, 2013. Accessed: 2019-10-10.

[KingSpalding18] King&Spalding. LNG in Europe 2018: An Overview of LNG Import Terminals in Europe. https://globallnghub.com/wp-content/uploads/2018/09/King.pdf, 2018. Accessed: 2018-09-01.

[LSS+19]   P. Lustenberger, F. Schumacher, M. Spada, P. Burgherr, and B. Stojadinovic. Assessing the performance of the european natural gas network for selected supply disruption scenarios using open-source information. *Energies*, 12(4685):1–28, 2019. doi:{10.3390/en12244685}.

[MMK16]   C. Matke, W. Medjroubi, and D. Kleinhans. SciGRID - An Open Source Reference Model for the European Transmission Network (v0.2). https://power.scigrid.de, 2016. Accessed: 2019-09-09.

[Mic20]   Microsoft. Bing maps api. https://www.microsoft.com/en-us/maps/licensing, 2020. Accessed: 2018 to 2021.

[Nis20]   A. Nisbet. Open topo data api. https://www.opentopodata.org, 2020. Accessed: 2020 to 2021.

[San19]   B. Sandvik. World Borders. http://thematicmapping.org/downloads/world_borders.php, 2019. Accessed: 2019-07-07.

[SAB+17]   M. Schmidt, D. Aßmann, R. Burlacu, J. Humpola, I. Joormann, N. Kanelakis, T. Koch, D. Oucherif, M.E. Pfetsch, L. Schewe, R. Schwarz, and M. Sirvent. *GasLib—A Library of Gas Network Instances*. 2017. doi:{10.3390/data2040040}.

[sl19]   scikit-learn. 1.1. Linear Models (scikit learn). https://scikit-learn.org/stable/modules/linear_model.html, 2019. Accessed: 2019-08-08.

[UoO14]   USA University of Oregon. Comparing distributions: Z Test. http://homework.uoregon.edu/pub/class/es202/ztest.html, 2014. Accessed: 2020-07-07.

[Wik20a]   Wikipedia. Bootstrapping (statistics). https://en.wikipedia.org/wiki/Bootstrapping_(statistics), 2020. Accessed: 2019-06-06.

[Wik20b]   Wikipedia. Cross-validation (statistics). https://en.wikipedia.org/wiki/Cross-validation_(statistics)#Exhaustive_cross-validation, 2020. Accessed: 2019-07-07.

[Wik20c]   Wikipedia. Jackknife resampling. https://en.wikipedia.org/wiki/Jackknife_resampling, 2020. Accessed: 2019-08-08.

[Wik20d]   Wikipedia. Lasso (statistics). https://en.wikipedia.org/wiki/Lasso_(statistics), 2020. Accessed: 2020-04-04.

[Wik20e]   Wikipedia. Limited-memory BFGS. https://en.wikipedia.org/wiki/Limited-memory_BFGS, 2020. Accessed: 2020-06-06.

[Wik20f]   Wikipedia. Out-of-bag error. https://en.wikipedia.org/wiki/Out-of-bag_error, 2020. Accessed: 2019-07-07.

[Wik20g]   Wikipedia. Transmission system operator. https://en.wikipedia.org/wiki/Transmission_system_operator/, 2020. Accessed: 2019-09-09.

[Wik20h]   Wikipedia. JAGAL. https://en.wikipedia.org/wiki/JAGAL, 2020. Accessed: 2020-01-01.

[BMWi11]   BMWi. Forschung für eine umweltschonende, zuverlässige und bezahlbare Energieversorgung. https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/6-energieforschungsprogramm-der-bundesregierung.pdf?__blob=publicationFile&v=12, 2011. Accessed: 2019-02-02.

[BMWi20]   BMWi. Home page of BMWi. https://www.bmwi.de/Navigation/DE/Home/home.html, 2020. Accessed: 2020-03-03.

[BundesregierungDeutschland20]   Bundesregierung Deutschland. Home page of Bundesregierung Deutschland. https://www.bundesregierung.de/Webs/Breg/DE/Themen/Energiewende/_node.html, 2020. Accessed: 2020-01-01.

[EntsoG20]   EntsoG. Home page of EntsoG. https://www.entsog.eu/, 2020. Accessed: 2020-03-03.

[GasIEurop20]   Gas Infrastructure Europ. Home page of Gas Infrastructure Europ. https://agsi.gie.eu, 2020. Accessed: 2020-01-01.

[GasSEurop20] Gas Storages Europ. Home page of Gas Storages Europ. https://www.gie.eu/index.php/transparency/gse-transparency-template, 2020. Accessed: 2020-01-01.

[Gassco20a] Gassco. Data page of facilities from Gassco. https://www.npd.no/en/about-us/information-services/available-data/map-services/, 2020. Accessed: 2020-01-01.

[Gassco20b] Gassco. Home page of Gassco Norway. https://www.gassco.no/en/, 2020. Accessed: 2020-01-01.

[IGU20] IGU. Home page of International Gas Union. https://www.igu.org/, 2020. Accessed: 2018-10-01.

[nationalGrid20] nationalGrid. Home page of National Grid UK. https://www.nationalgrid.com/uk/, 2020. Accessed: 2018-10-01.