PROJECT CEER-TCB18

Pan-European cost-efficiency benchmark for gas transmission system operators

MAIN REPORT

2019-07-17 V1.2

Disclaimer

This is the final report of a CEER project on cost efficiency benchmarking that involves data collection, validation and calculation of various efficiency indicators. Respecting the confidentiality of the submitted data and the prerogatives of each national regulatory authority to use or not the information produced in review of network tariffs or other monitoring, the report does not contain details for individual operators, nor comments or recommendations concerning the application of the results in regulation. In addition to this open report, each regulator and participating operator has also received a more detailed confidential analysis.

Pan-European Cost Efficiency Benchmark for Gas Transmission System Operators
Final report. Open. Project no: 370 / CEER-TCB18
Release date: 2019-07-17

Executive Summary

The Transmission Cost Benchmarking project 2018 (TCB18) is an initiative by the Council of European Energy Regulators (CEER) to initiate a stable and regular process for performance assessment of energy transmission system operators. The project covers both electricity and gas transmission and involves in total 46 operators from 16 countries in Europe. The project is the most ambitious regulatory benchmarking project documented so far, mobilizing national regulatory authorities (NRA), transmission system operators (TSO) and consultants in a joint effort to develop robust and comprehensive data and models. The project lasted from December 2017 to June 2019, involving five workshops and three successive stages of project setup, data collection and validation, followed by calculation and reporting.

Comparability
The primary challenge of any benchmarking is assuring comparability among observations emanating from operators with differences in organization, task scope and asset base. This challenge is addressed by (i) limiting the scope to comparable activities in transport and capacity provision, (ii) controlling to systematic differences in labor costs, (iii) standardizing the asset life-times and capital costs to equal conditions, (iv) excluding country-specific cost factors (land, taxes), (v) controlling for joint assets and cost-sharing, (vi) adjusting capital costs for inflation effects.

Reliability
The benchmarking is performed on NRA collected data, subject to a multi-stage data quality assurance process and using state-of-the art benchmarking methods such as Data Envelopment Analysis (DEA). The reliability and replicability of DEA results are immediate, since the method does not depend on any *ad hoc* parameters, but relies on the input data and linear programming. The environmental, economic and technical parameters and indices used have been collected from public sources based on clear techno-economic arguments. The sensitivity analysis shows that the results are robust to these latter assumptions. Globally the reliability of the method and the results is very good.

Verifiability
The quality of the data material in the project is a key determinant of the precision of the project results. The project addresses this criterion (i) by issuing and validating data collection guides and templates to avoid the use of incomparable data sources at an early stage, (ii) by defining a clear NRA validating procedure, (iii) by organizing a cross validation process for both technical and economic data through the consultant, (iv) by fully disclosing all processed data to each respective operator for control and confirmation to avoid misinterpretations and error, (v) by organizing interactive workshops to enable questions, and (vi) by providing online support on the project platform for submitting operators and NRAs.

Confidentiality
The data involved in the study go deeply into the operational efficiency of the participating operators. As this data are of crucial economic importance to the enterprises, the integrity and confidentiality of the data are taken seriously in the project both from structural, procedural and organizational viewpoints. Although transparency has advantages in data validation and interpretation of the results, the current project setup respects the concerns of operators not wishing to reveal the individual information or scores.

Approach
The methodological approach in the study has been to proceed independently with the estimation of a proxy for the diversified asset base of the operators, called the normalized grid or NormGrid. This system, constructed by international transmission system engineers based on transmission cost functions, provides a totex-relevant proxy for comparing operators in terms of size. The resulting metric was then tested by another team on the actual data, confirming the strong explicative value of the NormGrid. Quality provision was subject to a specific survey to assess potential indicators, but the results from this survey could not be directly applied to the model.

Environmental factors
The engineering team continued to develop testable hypothesis for the cost impact of various relevant environmental factors. After collection of such data, partially using a very detailed GIS-supported data set for each TSO, an analysis was made to enhance the NormGrid parameter with an environmental correction multiplier to adjust for heterogenous operating conditions. Other parameters were tested and included if not covered by correlation to the already incorporated factors or the grid in itself (NormGrid).

Activity model
Based on a multi-dimension performance model, additional parameters were selected based on their statistical and techno-economic significance to form a final model with one input, totex and four output parameters; NormGrid corrected for topography (slope class), total compressor power, total number of connection points and the pipeline length corrected for humidity class. The final model caters for all three performance categories; transportation work, capacity provision and customer service.

Benchmarking results
The model shows that the gas transmission system operators had a mean cost efficiency of 79% for 2017, with six frontier outlier operators and four best-practice peers. The results confirm earlier findings both in terms of level and distribution of scores, meaning that there likely is an efficiency potential corresponding to about 20% of total comparable expenditure. The result corrects for salary differences, heterogenous opening balances, unequal length of investment streams and overhead cost allocation rules.

Robustness
The results show a stable rank order with respect to the parameter interest rate and very low sensitivity in general to changes in the NormGrid system weights. The outlier identification procedure limits also the impact of operators with very specific cost structures that might be non-replicable for non-peers.

Table of Contents

1. Project objectives and organization

In this Chapter we state the project objectives, the organization and the report outline.

1.1 Main objectives

1.01 The main objective with the CEER TSO Cost efficiency Benchmark 2018 (project TCB18) is to produce a robust and methodologically sound platform for deriving cost efficiency estimates for transmission system operators, under process and data quality requirements allowing use of the results to inform regulatory oversight of the operators. In the project, best practice TSOs (forming the so-called frontier) are identified and related to other TSOs in a pan-European and regulatory context. Ultimately this is the purpose of TCB18.

1.02 TCB18 succeeds the E3GRID project in 2012/2013 and the E2GAS study of 2015/2016, combining in a single project a benchmark of gas TSOs and electricity TSOs. This report deals with the gas study. The electricity part is described in a separate report.

1.2 Project management

1.03 TCB18 is owned and initiated for regulatory purposes by CEER, the Counsil of European Energy Regulators. CEER has hired Sumicsid for advise and to perform parts of the benchmark study, notably analysis, modelling, and reporting.

1.04 Daily management of TCB18 is done by a project steering group (PSG) that consisted of representatives from ACM (Dutch NRA), BNetzA (German NRA), CNMC (Spanish NRA), NVE (Norwegian NRA), PUC (Latvian NRA), and Sumicsid (consultant). The PSG held regular meetings about every two weeks plus ad hoc meetings to discuss and decide about issues.

1.3 Project deliverables

1.05 The project produced two deliverables to document the results and the process:

1.06 **Final reports**:
This document for gas constitutes the final report documenting the process, model, methods, data requests, parameters, calculations and average results, including sensitivity analysis and robustness analysis. The report is intended for open publication and does not contain any data or results that could be linked to individual participants.

1.07 **TSO-specific reports:**
Clear and informative report on all used data, parameters and calculations leading to individual results, decomposed as useful for the understanding. The report only contains data, results and analyses pertaining to a single TSO. The confidential report was uploaded in an electronic version to each authorized NRA on the platform.

1.4 Reading guide

1.08 Chapter 2 provides a short summary of the project organization, followed by Chapter 3 outlining the data collection and validation process. Chapter 4 covers the full methodology for the activity analysis, the standardization of operating and capital expenditure, the benchmarking method, the model specification and the outlier detection. Chapter 5 reports on the results for the final model, including a robustness analysis. The results of the complementary survey on service quality are summarized in Chapter 6. Chapter 7 closes the study with a discussion of main findings, some perspectives and future work.

1.5 Appendix

1.09 The Appendix is released as a separate file. It contains the following documentation, not covered in the report but essential for the comprehension of the project:

A. Gas asset reporting guide, 2018-03-08

B. Financial reporting guide, 2018-03-08

C. Special conditions reporting guide, 2018-09-13

D. Method to treat upgrading, refurbishing and rehabilitation of assets, 2017-12-19

E. Modelling opening balances and missing initial investments, 2018-01-11

F. Norm Grid Development Technical Report, 2019-02-27 V1.3

2. Benchmarking process

In this Chapter the benchmarking process is summarized, including list of participants and the different points of interaction in the project.

2.1 Project phases

2.01 The project is organized into three phases as in Figure 2-1, described below. The time axis in this picture refers to the original plan. Dates mentioned below Figure 2-1 are realized dates.

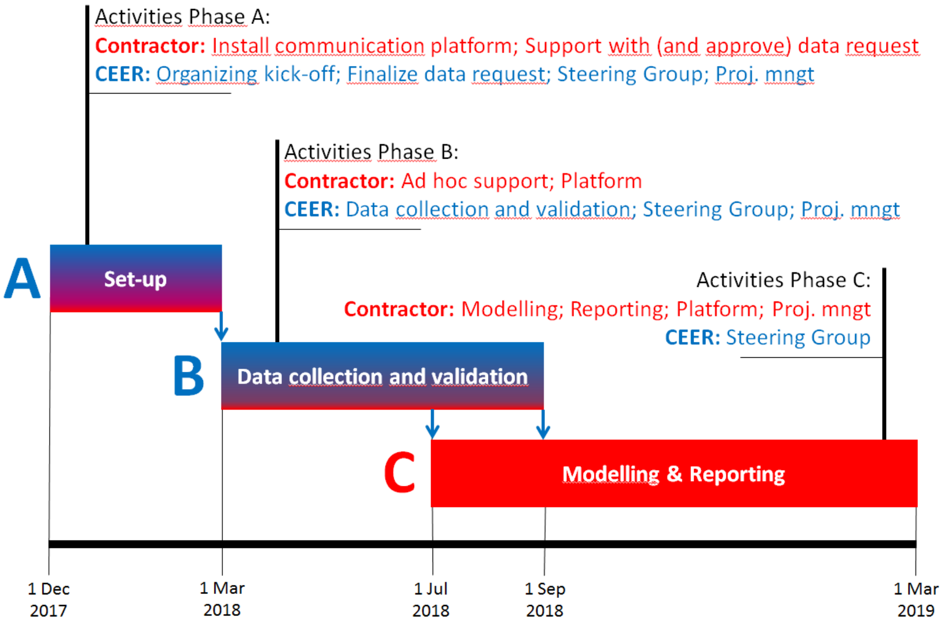

Figure 2-1 Project phases (original dates)

Phase A

2.02 The initial phase is devoted to the launch, detailed planning and preparation for the operational part of the project in the next two phases.

2.03 Duration: 01/12/2017 – 28/02/2018

2.04 Key events:

1) Project management setup
2) Kick off workshop W1
3) Project platform setup
4) Revision and final release of data definition guides and Excel templates

Phase B

2.05 The data collection and validation phase is mainly in the hands of CEER and the NRAs, the consultant act as support and coordinator of the project platform.

2.06 Duration: 01/03/2018 – 30/08/2018

2.07 Key events:

1) Data collection
2) Data validation (NRA)
3) Cross validation of data (consultant)
4) Workshop W2 on data collection
5) Collection of environmental public parameters (consultant)

Phase C

2.08 The last project phase contains the model specification, verification, calculations, outlier identification, sensitivity analyses, documentation, presentation and report editing for CEER and the individual NRAs.

2.09 Duration: 01/09/2018 – 30/06/2019

2.10 Key events:

1) NormGrid development
2) Workshop W3 on NormGrid models and environmental factors
3) Model specification
4) Workshop W4 on model specification
5) Release of individual TSO-specific data sheets pre-run
6) Efficiency analyses
7) Robustness analyses
8) Workshop W5 on final results
9) Editing of final report
10) Editing of individual TSO-specific score sheet

2.2 Project Team assignments

2.11 The consultant is organized in four teams (CENTRAL, ECON, TECH-GAS, TECH-ELEC). The Sumicsid project members include Prof.dr. AGRELL and Prof. dr. BOGETOFT, with a long experience in methodological and applied benchmarking of energy networks, as well as Ir. BEAUSSANT and Ir. TALARMIN, international expert engineers in electricity and gas, respectively, all with extensive experience in transmission system analysis and benchmarking.

2.3 Project documentation

2.12 The documentation for the project, data calls, instructions and workshop material as well as methodological notes, were published at a project platform only. Likewise, all data and validation material were up- and downloaded from the project platform, avoiding versioning and security problems associated with email. The platform contained private and public areas for all, electricity and gas transmission operators, respectively.

2.13 The project initially aimed at transparency for, at least, aggregate data and results. However, no consensus could be reached among the TSO participants to share data generally in the project. In consequence, all detailed data and results were disclosed uniquely to the participating TSO and their respective NRA.

2.4 Workshops

2.14 Since for an important part the project is focused at TSO-NRA interaction, a number of workshops were organized (cf. Table 2-1). All project participants, TSOs and NRAs, were invited to the workshops, from which all documentation and minutes were published on the project platform.

Table 2-1 Project workshops GAS

Workshop	Phase	Date
W1 Kickoff	A	2018-01-16
W2 Method, data validation	B	2018-04-26
W3 Normgrid and environment	C	2018-10-11
W4 Model specification	C	2018-11-27
W5 Final results	C	2019-04-05

2.5 Project participants

2.15 The following TSOs and NRAs took part in the project (cf. Table 2-2):

Table 2-2 TCB18 participants GAS.

TSO	Country	NRA
Amber Grid	LT	NCC
Bayernets	DE	BnetzA
Conexus	LV	PUC
DESFA	GR	RAE
Elering	EE	ECA
Enagas	ES	CNMC
Energinet.dk	DK	DUR
Fluxys SA	BE	CREG
Fluxys Deutschland	DE	BnetzA
Fluxys TENP	DE	BnetzA
GASCADE Gastransport	DE	BnetzA
Gasum	FI	EV
Gasunie Deutschland Transport Service	DE	BnetzA
GRTgaz Deutschland	DE	BnetzA
Gastransport Nord	DE	BnetzA
GTS	NL	ACM
jordgas Transport	DE	BnetzA
Lubmin-Brandov Gastransport	DE	BnetzA
NEL Gastransport	DE	BnetzA
National Grid Gas Transmission	UK	OFGEM
Nowega	DE	BnetzA
Open Grid Europe	DE	BnetzA
ONTRAS Gastransport	DE	BnetzA
OPAL Gastransport	DE	BnetzA
Plinovodi	SI	EA
Reganosa	ES	CNMC
REN	PT	ERSE
terranets bw	DE	BnetzA
Thyssengas	DE	BnetzA

3. Data collection

In this chapter, the data collection and the data validation process are discussed.

3.1 Procedure (guide and collection)

3.01 For TCB18 data definition guides, one for asset data (Appendix A) and one for financial data (Appendix B), were developed in a separate project that preceded TCB18. That preceding project started in February 2017 and ended about six weeks after the kick off of TCB18 (so there was actually a slight overlap). Part of that were two workshops, one in May 2017 (W0a) and one in October 2017 (W0b).

3.02 TSOs received the final data definition guides (Appendix A and B) early March 2018 and were asked to deliver data in the middle of May 2018. In that period CEER organized the second TCB18 workshop (W2), dedicated to data collection. That workshop was meant to discuss the progress of data collection by TSOs and to identify and solve issues with it. NRAs had the time to validate TSO data until the end of June. After the second TCB18 workshop CEER decided to extend "softly" the deadline for delivering data by TSOs to the end of June. By "softly" was meant that TSOs were asked to agree with their NRAs a time path for delivering data in such a way that by the end of June the data was delivered by the TSOs and validated by the NRAs. Eventually, most data was delivered and validated nationally on time. However, not for all TSOs, imposing some stress on subsequent stages of TCB18.

3.2 Data quality strategy

3.03 For TCB18 CEER developed and laid down (workshop W2) a clear strategy for safeguarding the quality of the benchmark data that enters the benchmark, see Figure 3-1 below. For the German TSOs, participating through their NRA, particular attention was paid to screen and analyze any potential differences between the reporting instructions in the previous benchmark, the national validation performed and potential sources of errors. This latter process resulted in a positive analysis for the gas TSO, deemed to offer comparable data of high validated quality subject to a clear regulatory task description.

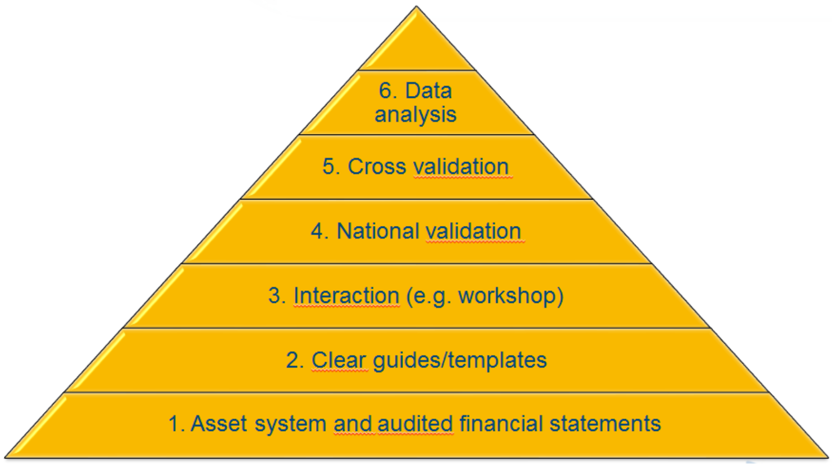

Figure 3-1 Data quality strategy.

3.04 The data quality strategy consists of six layers:

1) The first- or base-layer is the asset system and audited financial statements of TSOs. The data quality strategy is founded on the principal that TSOs have a proper asset system and audited financial statements.
2) The second layer consists of reporting guides and templates, see Paragraph 3.1. For a year CEER, TSOs, and the consultant have interactively worked on clear data definitions to translate the base-layer (asset system and audited financial statements) into benchmark data.
3) In all steps of the process there was interaction between TSOs, NRAs and the consultant, notably through many workshops. The interaction helped in the correct interpretation of definitions among participating TSOs and NRAs.
4) After data collection, national validation at NRA level has been performed. The goal of national validation is to assure that data is complete, consistent, correct and plausible.
5) After National validation, cross validation was done by the consultant. The goal of cross validation is that remaining misinterpretation of definitions amongst countries are detected and corrected for. In an ideal world it should not be necessary, but practice is unruly and a cross validation is necessary.
6) Finally, data analysis has been done by Sumicsid to develop a benchmark model. This is seen as part of the data quality strategy as data analysis may reveal errors in the data that was not picked up by national or cross validation. So actually, the validation (i.e. the previous layer) did not have a well-defined ending, it continued as long as the analysis and modelling were in progress.

3.05 TSOs were not asked to audit their data formally by an independent auditor. A first reason for that is that the data definitions take the audited annual accounts as starting point. Furthermore, NRAs will also check data against sources like regulatory data, which are often audited and validated before. Also, an audit often focuses on just a part of the data, mostly the financial accounts. So an explicit audit on the benchmark data for each TSO was not seen as a necessary part of the data quality strategy.

3.06 Final data checks were done in March/April 2019. All TSOs and NRAs received a dump of asset and financial files that they could check on missing or incorrect data. For many TSOs a few final corrections have been made, leading to data sets of good quality.

3.07 Although no strategy will be fully safe, CEER believes that its structured approach was indeed vital in securing a successful benchmark project.

3.3 Environmental data

3.08 The TCB18 benchmark model addresses several environmental factors, like landuse, slope, humidity etc. To do this data is required about such factors. In E2GAS this data was collected by asking TSOs to specify the operating conditions at asset level. The main drawback of that approach was that it stimulated strategic reporting. Also, item-wise reporting assumed all environmental effects and their combinations to be known beforehand, making statistical analysis difficult and the results too dependent on the engineering assumptions. Finally, the capacity and resources necessary from the TSOs to estimate the different factors vary and depend on the importance assigned to the benchmarking results in the respective countries. All these reasons made the E2GAS approach less attractive.

3.09 In E3GRID (the CEER 2012-13 electricity TSO benchmarking), the consultants collected some aggregate indicators at country level, e.g. population density, that were used as proxies for environmental complexities. This approach is exogenous and "equitable", but the resulting adjustment for environmental conditions is rather crude, prompting various technical measures in the benchmarking techniques to avoid absurd results. The E3GRID approach was therefore judged to be unsatisfactory for the new benchmarking.

3.10 TCB18 is not only a one-shot project to arrive at a unique model. It is one step towards a structured development of periodic regulatory benchmarking. As such, the priority is also to provide structurally and incentive-analytically sound solutions for future repetitions. An ideal solution would be to organize external collection of all environmental conditions from public established databases based on the actual asset locations for all participants. In subsequent runs these reporting restrictions and the format for delivery and processing of environmental data could be developed as an add-on project to TCB18, leading to several interesting applications also for the TSOs own use. Combining open databases for landuse, soil type, humidity, topography et.c. into a platform where the environmental complexity could be objectively assessed without any manual intervention by operators or regulators would be a desired outcome of this process.

3.11 The process proceeded initially by an independent identification of the relevant environmental factors by type of energy (gas, electricity), the assets concerned by factor, the economic rationale of impact and the hypothesized magnitude (See Appendix F NormGrid Model). The consultants thereafter identified and collected the corresponding data items from the available data bases, subjecting the data to statistical tests for impact using the reported data.

3.12 The sources in Table 3-1 were used for analysis, in particular the Copernicus and CORINE GIS-based metrics derived for each TSO.

Table 3-1 Data sources for environmental factors.

Condition	Source	Granularity
Landuse (agricultural, urban, …)	EUROSTAT	Country
Landuse (type of use)	CORINE (GIS)	TSO
Vegetation (shrubs, grass, …)	EUROSTAT	Country
Area (forests, lakes, mountains, ...)	EUROSTAT, OECD	Country
Climate (wind, icing, salt, extreme temperature)	WeatherOnline, Geographic	City
Road infrastructure	OECD	Country
Topography (ruggedness, coastal area)	Puga et al. (2012)	Country
Topography (slope)	Copernicus (GIS)	TSO
Soil conditions (subsurface features)	Copernicus (GIS)	TSO
Humidity conditions (wetness, water)	Copernicus (GIS)	TSO
Soil conditions (humidity, subsurface features)	Copernicus (GIS)	TSO

3.13 The granularity of the GIS-based data is very good. As an example, the slope factor (a key factor in the construction costs for major infrastructure projects over land) is estimated in Copernicus from cells with a side of 25m, providing height data with a vertical accuracy of 7m, based on satellite imagery and geographical modelling. The data allows detailed calculations of the share of any area within given ranges of slopes, defining the concepts as 'hilly', 'undulating' , 'mountainous' etc. objectively and with high scientific validity.

3.4 Special conditions

3.14 During the project TSOs were given an opportunity to signal conditions that are not taken into account by the benchmark model, but they think should have been. Such conditions are referred to as special conditions and may call for correction of benchmarked scope or data, or the benchmark model. The concept of special conditions evolves from the concept of so-called Z-factors in previous CEER benchmarks.

3.15 Defining and implementing special conditions is meant to get closer to the purpose of the benchmark, i.e. to define best practices. As all TSOs in the sample will be related to frontier companies, it is therefore important that special conditions should only be labelled as such if they stand a number of criteria:

Complementarity

3.16 This criterion is meant to distinct conditions that are already sufficiently dealt with by the benchmark model from conditions that are not and may need complementary treatment. For example, if the condition can be dealt with by building additional standard assets, and if the model would "credit" TSOs for their asset base, then the condition is likely to be already taken into account sufficiently by the model. There can actually be two reasons for complementary treatment. First of all, this could be the case if the benchmark model is insufficiently specified. A typical example of complementary treatment in such case would be the change or addition of a modelling parameter. Secondly, complementary treatment may be called for if the claimed condition is something very specific that only one or few TSOs in the sample have to live with, i.e. the condition is relatively unique to the claimant. At all times and most importantly, complementary treatment will only be done if doing so fits the purpose of the benchmark.

Objectification

3.17 A special condition is something that, so to say, overcomes a TSO, i.e. it can reasonably not be held against the TSO and this should not be arguable. Special conditions must not be defined in terms of the (subjective) strategy to deal with the condition. So a claim cannot be formulated like "we do A because of condition C", because A would only refer to a choice made by the TSO that may be up for efficiency analysis. Instead a claim should be formatted like "we are faced with condition C and dealing with it inevitably comes with a disadvantage (compared to not having C)." So, both the condition C and the unavoidability of a disadvantage must fully and inarguably be beyond control of the TSO. Objectivity also implies that the condition is conceptually simple, obvious, and transparent, even to less informed public.

Durability

3.18 Incidents do not qualify as special conditions, think e.g. of a flooding in a certain year. Instead, special conditions are supposed either to exist over a substantial part of the reporting period, i.e. many years, or to exist for many years in the future impacting operations in the past. No explicit norm for this has been set as it may depend on the precise nature of the condition (geographical, technical, economical, etc.). At any rate, this criterion is meant to separate structural circumstances from incidents.

Materiality

3.19 Special conditions can only be recognized as such if they come with a well-defined and significant cost impact. The cost impact of a special condition is defined as the minimum unavoidable cost to deal with the condition. This is what is seen as the value of the claim. Put differently, the value of the claim is the cost difference between the lowest cost alternative to deal with the condition (this is not per se the alternative that is actually implemented) and the cost that would have been made if the condition would not exist. At any rate, the cost impact of a special condition must be clearly quantifiable. If quantification is ambiguous or poorly documented, it will be difficult to correct in the benchmark for the condition. Moreover, it would signal that the condition does not have (had) the explicit attention of management as such, which makes the condition being a special one less credible. Also, the (monetary) value of the claim must be significant, i.e. it must be big enough to significantly impact the outcome of the benchmark. A soft norm for this is about 5 percent of the benchmarked gross investment stream of the claimant or, if the claim is about expenses only, about 5 percent of its benchmarked expenses. This is important to avoid erosion of the best practice frontier by relatively small peculiarities of which all TSOs will have some, some fortunately, some unfortunately.

3.20 These criteria are cumulative, forming a firewall to improper claims in order to protect the hygiene of the best practice frontier, which is in the interest of all TSOs. Individual interests can only impact the benchmark if this is reasonable to all. Nevertheless, as the benchmark can be used in regulation, individual interests are of course quite relevant, think of a severe unfortunate incident in the reference year, strong political pressure on the TSO, legacy, or regulatory decisions. However, such cases boil down to interpretation of an individual benchmark score, which is a national affair between individual NRAs and TSOs, just like with implementation of benchmark results afterwards in regulatory decisions. So it is important to bear in mind that there is a cut-off point where international benchmarking stops and national interpretation and implementation starts. The benchmark model defines that point and the criteria for special conditions are instrumental to that.

3.21 The text in the above was part of a special conditions reporting guide of which a first draft was consulted in July 2018 (Appendix C). The final version of September 2018 was almost the same as the draft. TSOs were given time until early January 2019 to submit claims.

3.22 5 TSOs submitted in total 15 claims of which 8 were rejected by the PSG and 7 were put under investigation. The rejected claims, including the reason for rejection read:

Table 3-2 Operator specific claims rejected with motivation.

TSO	Claim	Grounds for rejection
GTS	In contrast to e.g. German TSOs GTS provides unconditional capacity.	This is a re-claim from previous benchmarks. GTS did not provide evidence at that time and does not present new evidence or circumstances this time.
GTS	Joint ventures are not correctly addressed in the benchmark model.	The benchmark model corrects for allocation of shares.
GTS	Demands by the authorities lead to higher TOTEX.	The submitted claims show that all TSOs face certain obligations, even though this differs country wise. In fact, also TSOs that did not claim anything in this area face many obligations. Therefore, correcting this only for TSOs that claimed in this area would bias the benchmark result.
Plinovodi	Obligations to provide DSO tasks.	Insofar these tasks are in scope of the benchmark, both the cost and outputs of it are included in the model.
Fluxys	Soil cleaning up.	Not unique, nor material.
Fluxys	Severe Belgian gas law.	The submitted claims show that all TSOs face certain obligations, even though this differs country wise. In fact, also TSOs that did not claim anything in this area face many obligations. Therefore, correcting this only for TSOs that claimed in this area would bias the benchmark result.
Fluxys	Archeological studies and excavations.	Not unique, nor material.
Amber Grid	For security of supply an LNG terminal including connections to existing grid had to be built.	This claim is quite common, showing that all TSOs need to secure supply as part of their business. Claims of this kind do not convince that such obligations are much more severe in some country than in others. Moreover, LNG is not benchmarked.

3.23 Claims that were put under investigation are listed in Table 3-3:

Table 3-3 Investigated operator-specific claims.

TSO	Claim	Consideration in model
GTS	Routes for the network are suboptimal due to high population density.	Population density best covered by landuse factors (incl. GIS-level density areas). NG_Area is 99% correlated with NG_Slope for gas, leading to a choice where Slope makes a stronger techno-economic sense.
GTS	The more humid (wet) soil, the higher the construction and maintenance costs.	Soil factors under GIS tested for inclusion, leading to the inclusion of soil humidity as model output.
GTS	Complexity of GTS is higher than complexity of other TSOs, which may lead to incomparable input/output ratios.	Complexity not well defined for testing beyond the dimensions (area, slope, connections, capacity) already in the model.
Fluxys	Horizontal drilling (ordered by environmental obligations).	Engineering choice occurring also for other TSOs, no significant difference to justify exemption.
Enagas	Copernicus has higher resolution than D. Puga (25 versus 1000 meters) for topography.	Implemented: Copernicus used for all TSOs as a complement to Ruggedness (Puga).
Enagas	No European source available for subsoil type, IGN should be used for Spain.	Using ESDAC at GIS-level, detailed data for soil and subsoil.
Enagas	Low density of network leads to higher cost for grid maintenance centers.	See comment for high density above. Population density best covered by landuse factors (incl. GIS-level density areas). NG_Area is 99% correlated with NG_Slope for gas, leading to a choice where Slope makes a stronger techno-economic sense.

3.24 Putting the claims in Table 3-3 under investigation means that the impact of the claim was tested for in the cost driver analysis. Note that most of these claims have to do with modelling environmental conditions. Specifically for gas, the GIS-data for landuse (area) and slope (topography) classes were strongly correlated, meaning that the main effects for either one are captured by the inclusion of any of the two.

4. Methodology

This Chapter is devoted to the discussion of the methodological approach that has been used in the TSO benchmarking, including the important preparation in terms of activity analysis, cost standardization, asset aggregation and correction for structural comparability. The Chapter then addresses model specification and method choice.

4.1 Background

4.01 The benchmarking model is pivotal in incentive based regulation of natural monopolies. By essence, benchmarking is a relative performance evaluation. The performance of a TSO is compared against the actual performance of other TSOs rather than against what is theoretically possible. In this way, benchmarking substitutes for real market competition.

4.02 Of course, the extent to which a regulator can rely on such pseudo competition depends on the quality of the benchmarking model. This means that there is no simple and mechanical formula translating the benchmarking results into for example revenue caps. Rather, regulatory discretion – or explicit or implicit negotiations between the regulator, the industry and other interest groups – is called for.

4.2 Steps in a benchmarking study

4.03 The development of a regulatory benchmarking model is a considerable task due to the diversity of the TSOs involved and the potential economic consequences of the models. Some of the important steps in model development are:

4.04 **Choice of variable standardizations:** Choices of accounting standards, cost allocation rules, in/out of scope rules, asset definitions and operating standards are necessary to ensure a good data set from TSOs with different internal practices.

4.05 **Choice of variable aggregations:** Choices of aggregation parameters, such as interest and inflation rates, for the calculation of standardized capital costs and the search for relevant combined cost drivers, using, for example, engineering models, are necessary to reduce the dimensionality of potentially relevant data.

4.06 **Initial data cleaning:** Data collection is an iterative process where definitions are likely to be adjusted and refined and where collected data is constantly monitored by comparing simple Key Performance Indicators (KPIs) across TSOs and using more advanced econometric outlier - detection methods.

4.07 **Average model specification:** To complement expert and engineering model results, econometric model specification methods are used to investigate which cost drivers best explain cost and how many cost drivers are necessary.

4.08 **Frontier model estimations:** To determine the relevant DEA (and depending on data availability SFA) models, they must be estimated, evaluated and tested on full-scale data sets. The starting point is the cost drivers derived from the model specification stage, but the role and significance of these cost drivers must be examined in the frontier models, and alternative specifications derived from using alternative substitutes for the cost drivers must be investigated, considering the outlier-detecting mechanisms.

4.09 **Model validation:** Extensive second-stage analyses shall be undertaken to see if any of the non-included variables should be included. The second-stage analyses are typically done using graphical inspection, non-parametric Kruskal-Wallis tests for ordinal differences and truncated Tobit regressions for cardinal variables. In addition to second stage control for possibly missing variables, it is desirable to perform extensive robustness runs to ensure that the outcome is not too sensitive to the parameters used in the aggregations.

4.10 It is worth emphasizing that model development is not a linear process but rather an iterative one. During the frontier model estimation, for example, we identified extreme observations resulting from a data error not captured by the initial data cleaning. In turn this may lead to renewed data collection and data corrections. Such discoveries make it necessary to redo most steps in an iterative manner.

4.3 Activity analysis and scope

4.11 Benchmarking relies crucially on the structural comparability of the operators constituting the reference set. Differences in structure primarily result from differences in (i) assigned transport tasks, (ii) interfaces with other regulated or non-regulated providers and (iii) asset configuration. The identification of the main functions is the first action in a benchmarking context since various operators cover diverse functions and therefore cannot directly be compared at an aggregate level. The identification is also crucial since different regulations and usages of the performance evaluations may require different perspectives.

4.12 Just as electricity TSOs perform a range of functions from market facilitation to grid ownership, the gas TSOs demonstrate a portfolio of transport and terminal tasks, also including specific functions related to storage, LNG terminals, trading and balancing. The task here is twofold; first to make a systematic and relevant aggregation of the different activities and to map them to existing or obtainable data that could be reliably used in an international benchmarking. Second, the scope must be judged against the types of benchmarking methods and data material realistically available. E.g. if the activity (say planning) yields output for a horizon way beyond the existing data, the activity is not in the relevant scope for a short-term benchmarking.

4.13 The common core task for the gas TSOs here is defined as providing and operating the assets for transport and transit of energy. More specifically, we focus on (i) services: transport to downstream exit and transit to a cross-border point, (ii) assets: a pipeline network with its control system and (iii) activities: grid planning, grid financing/ownership, grid construction, grid maintenance, and grid metering. Other elements, notably storage and LNG services/assets and system operations and market facilitation, are out of scope in TCB18. For more discussion of the definition of relevant scope, see the PE2GAS study (2014, Chapter 3).

4.4 Grid transmission activities

4.14 The fundamental objective of a transmission system operator is to transport energy to distribution networks and other connected clients.

4.15 By distinguishing activities, the autonomy and independency of an operator may be put in a correct context to enable, among other things, performance assessments. The activities are listed below.

4.16 Note that in previous benchmarking, activities such as Grid construction (C) or Grid financing (F) were listed and defined. In this project, these activities are no longer informative for validation or comparability. In practice, almost all activities of construction are capitalized and the activity has no assets, staff or costs in the accounts of the typical TSO. Likewise, the financial activities related to grid operations are not susceptible to standardization.

4.5 T Transport

4.17 The transport activity includes the operation of the injection, transport and delivery of energy through the transmission system, from defined injection points to connection points interfacing a client, a downstream network, a storage facility (gas) or an interconnection to another transmission network. The transport activity is enabled by the operations of grid assets for transport (lines and transformers for electricity, pipelines and compressors for gas). The transport activity thus comprises the day-to-day activities of real-time flow control, metering and operational control and communication.

4.18 The assets utilized for transport constitute the pipeline system characterizing the TSO. The operational expenses for transport include staffing control centers, inspections, safety and related activities, including direct costs for products and services as well as staff.

4.19 The cost for energy used in transport (covering internal consumption and losses) is reported separately under T to control for structural comparability

4.6 M Grid maintenance

4.20 The maintenance of a given grid involves the preventive and reactive service of assets, the staffing of facilities and the incremental replacement of degraded or faulty equipment. Both planned and prompted maintenance are included, as well as the direct costs of time, material and other resources to maintain the grid installations. It includes routine planned and scheduled work to maintain the equipment operating qualities to avoid failures, field assessment and reporting of actual condition of equipment, planning and reporting of work and eventual observations, supervision on equipment condition, planning of operations and data-collection/evaluation, and emergency action.

4.21 The activity may have assets (spare parts) and operating costs (direct, staff and outsourced services).

4.7 P Grid planning

4.22 The analysis, planning and drafting of gas network expansion and network installations involve the internal and/or external human and technical resources, including access to technical consultants, legal advice, communication advisors and possible interaction with European, governmental and regional agencies for preapproval granting.

4.23 Grid planning also covers the general competence acquisition by the TSO to perform system-wide coordination, in line with the IEM directive, the TEN corridors and the associated ENTSO tasks. Consequently, costs for research, development and testing, both performed in-house and subcontracted, related to functioning of the transmission system, coordination with other grids and stakeholders are reported specified under grid planning P.

4.24 The activity has no assets and operating costs (direct, staff and services). In the case internal planning costs are capitalized, this is noted in the investment stream.

4.8 I Indirect support

4.25 With indirect services, we refer to services related to the general management of the undertaking, the support functions (legal, human resources, regulatory affairs, IT, facilities services etc.) that are not directly assigned to an activity above. Central management, including CEO, Board of directors and equivalent is also explicitly included.

4.26 In principle, the residual assets for a gas transmission system operator (e.g. office buildings, general infrastructure) could be considered as assets for Indirect support.

4.27 However, to the extent that this entails the incorporation of land, land installations and non-grid buildings in the analysis, all of which are susceptible to be country specific investments, such elements are excluded from the benchmarking.

4.9 S System operations

4.28 Within system operations for gas transmission, ancillary services are retained as defined in 2009/73/EC and congestion management (compliant with the ENTSO-G classification). Ancillary services include all services related to access to and operation of gas networks, gas storage and LNG installations, including local balancing, blending and injection of inert gases, but exclude "facilities reserved exclusively for transmission system operators carrying out their functions", 2009/73/EC Art 2(14).

4.29 ENTSO-G further considers the transparency in data exchange with the purpose of interoperability as a specific point in system operations. In consequence, costs related to this activity per se are to be considered as system operations.

4.30 If part of the services above are delegated to subordinate (regional) transmission coordinators with limited decision rights, the associated costs are included in system operations.

4.31 System operations has no assigned assets, the costs are direct costs for services and staff.

4.10 X Market Facilitation

4.32 Market facilitation includes all direct involvement in energy exchanges through information provision or contractual relationships. This comprises regulated tasks through procurement of renewable power, residual buyer obligations or capacity allocation mechanisms, capacity auctioning mechanisms, and work on coordination of feed-in tariffs.

4.33 The market facilitation activity is composed uniquely of direct expenses related to the contractual relations excluding transport and storage, primarily information costs and energy purchases for other purposes than the consumption in their own grid.

4.34 The activity has no eligible assets and no staff costs.

4.11 TO Offshore transport

4.35 The transport and transit of gas through offshore assets (i.e. subsea pipelines and subsea interconnectors, see Asset reporting guide GAS, art 17) are considered as offshore gas transmission activities.

4.12 G Gas storage

4.36 The operation of gas storage facilities, including their maintenance and internal energy consumption, can be considered as separate service of gas storage, analogous to that of non-TSOs.

4.37 Costs concerning gas storage are separable according to the Directive 2009/73/EC Art 23 §1 (principle), Art 30§3 (obligation) and Art 41 §1(f), 6(a) (NRA authority to request data), both in terms of ownership of assets and their operation.

4.13 L LNG terminals

4.38 The operation and maintenance of LNG terminals and peak-shaving plants, the interfaces with ports and other infrastructure, the administration and specific actions necessary to enable such operations are considered part of a specific service.

4.39 Costs concerning LNG terminals are in principle separable according to the Directive 2009/73/EC Art 23 §1 (principle), Art 30§3 (obligation) and Art 41 §1(f), 6(a) (NRA authority to request data), both in terms of ownership of assets and their operation.

4.14 O Other activities

4.40 A gas TSO may have marginal activities that are not covered by the classification above, such as external operator training, field testing for manufacturers, leasing of land and assets for non-transport use. These activities should be listed, the costs and assets should be specified and excluded from the benchmarking.

4.15 Scope

4.41 Based on the analysis of common factors in cost reporting, the variability and homogeneity of the data and the separability of the activity, it was decided to define the benchmarked scope as the structurally comparable core activities of the transmission operator, i.e. T, M, P, and I (partially), see Figure 4-1 below. Planning (P) was included as it was present in all TSOs and considered as a techno-economic necessity, inseparable from the investment and operational activities.

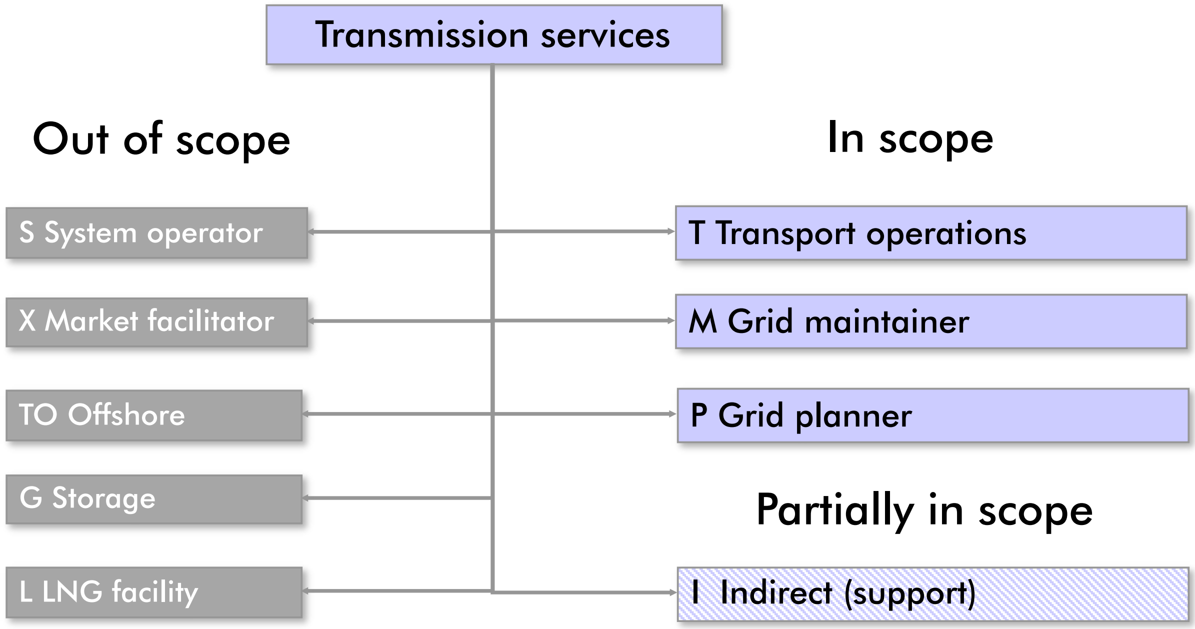

Figure 4-1 Benchmarked actvities and scope.

4.42 To permit a mapping of the P&L onto the activities, the operators also report the activities S, X, TO, and, if applicable, O, G, and L. These activities are to be validated to avoid cost leakage, but are not in the planned benchmarking scope.

4.16 Cost definitions and standardization

4.43 Benchmarking models can be grouped into two alternative designs with an effect on the scope of the benchmarked costs:

A. A short-run maintenance model, in which the efficiency of the operator is judged-based on the operating expenditures (Opex) incurred relative to the outputs produced, which in this case would be represented by the characteristics of the network as well as the typical customer services.

B. A long-run service model, in which the efficiency of the operator is judged-based on the total cost (Totex) incurred relative to the outputs produced, which in this case would be represented by the services provided by the operator.

4.44 From the point of view of incentive provision, a Totex based approach is usually preferred. It provides incentives for the TSOs to balance Opex and Capex solutions optimally. In this study, the focus is therefore on Totex benchmarking.

4.45 The standardization of costs plays a crucial role in any benchmarking study, especially, when the study is international. Below we discuss the derivations of the benchmarked operating and capital cost, leading to the final benchmarked dependent variable; the benchmarked Totex.

4.17 Benchmarked OPEX

4.46 There are various steps involved in order to derive the respective benchmarked Opex for the benchmarked functions in scope below, see Figure 4-2 below.

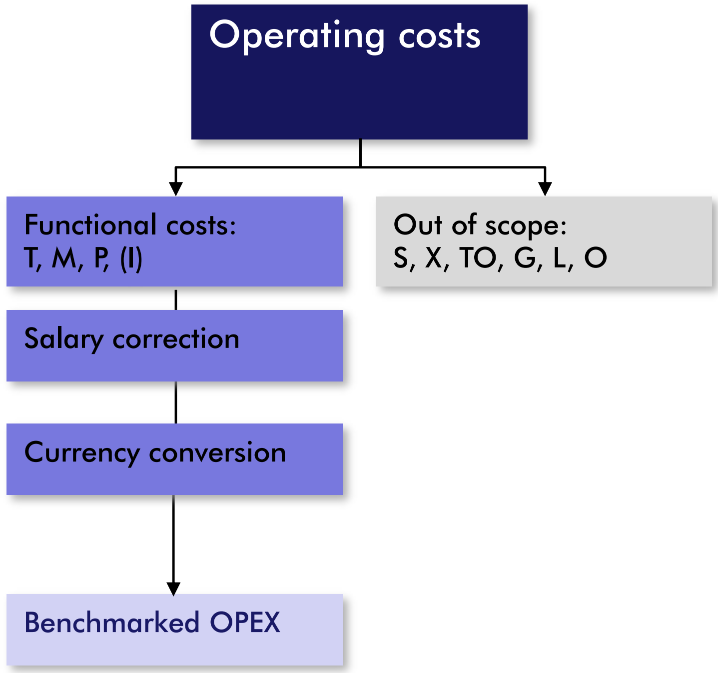

Figure 4-2 Steps in deriving benchmarked OPEX.

4.47 The relevant cost items for OPEX, derived directly from the TSOs' data per activity are added together (cf Cost reporting guide, Appendix B).

4.48 Depreciation of grid related assets is excluded from this list, as this is covered by the benchmarked CAPEX.

4.49 The cost of energy is deducted from benchmarked OPEX at this step.

OPEX: Labor cost adjustments

4.50 In order to make the operating costs comparable between countries a correction for differences in national salary cost levels has been applied. Otherwise TSOs would be held responsible for cost effects, e.g. high wage level, which is not controllable by them.[1]

[1] We note that there is some simplification involved in the logic of salary cost adjustment. Had the respective operator truly had lower (or higher) salary cost then it may in practice also have chosen a different mix of production factors - e.g. operate less (or more) capital intensively. However, we do not consider this in the context of salary cost adjustments.

The basis for the labor cost adjustment is the labor cost, not the data collected on FTE (full time equivalent employees) by function, since these data were less reliable.

4.51 The salary adjustment consists of two steps:

1) *Step 1 – adjustment of direct manpower costs* by increasing/decreasing the direct manpower costs of the companies using the respective salary index.
2) *Step 2 – reversal of part of salary* adjustment. Step 1 applies to a gross value, while the Opex entering the benchmarking is a net value after deducting direct revenues (for services outside the scope of the benchmark). Hence, some part of the salary adjustment has to be reversed considering that the share of direct manpower costs is proportionally smaller in the Opex used for benchmarking.

4.52 The correction for systematic salary cost differences can be made by several indexes, see Table 4-1 for those collected and tested in the study. The general indexes, such as the EUROSTAT index for all services (LCIS) correlates poorly to the actual salary differences observed among the TSOs, primarily since the basis for the index involves services not involved in transmission. Figure 4-3 illustrates three indexes, whereof the PLICI index was chosen since its scope (civil engineering services) corresponded the best to the differences between the salaries paid and European average. Compared to previous studies using general indexes, the current approach provides a lower variance in the estimation, better fitting the real differences.

Table 4-1 Labor cost indexes tested (PLICI selected).

Index	Source	Type	Scope
Plits	EUROSTAT	Price level index	Services
Plitg	EUROSTAT	Price level index	Goods
Plico	EUROSTAT	Price level index	Construction
Plici	EUROSTAT	Price level index	Civil eng
Lcis	EUROSTAT	Labor cost index	Services
Lcig	EUROSTAT	Labor cost index	Goods
Lcic2	EUROSTAT	Labor cost index	Construction F
Lciusm	Fed Bank	Purchasing parity	Manufacturing
Coc	EUROSTAT	Price level index	Construction

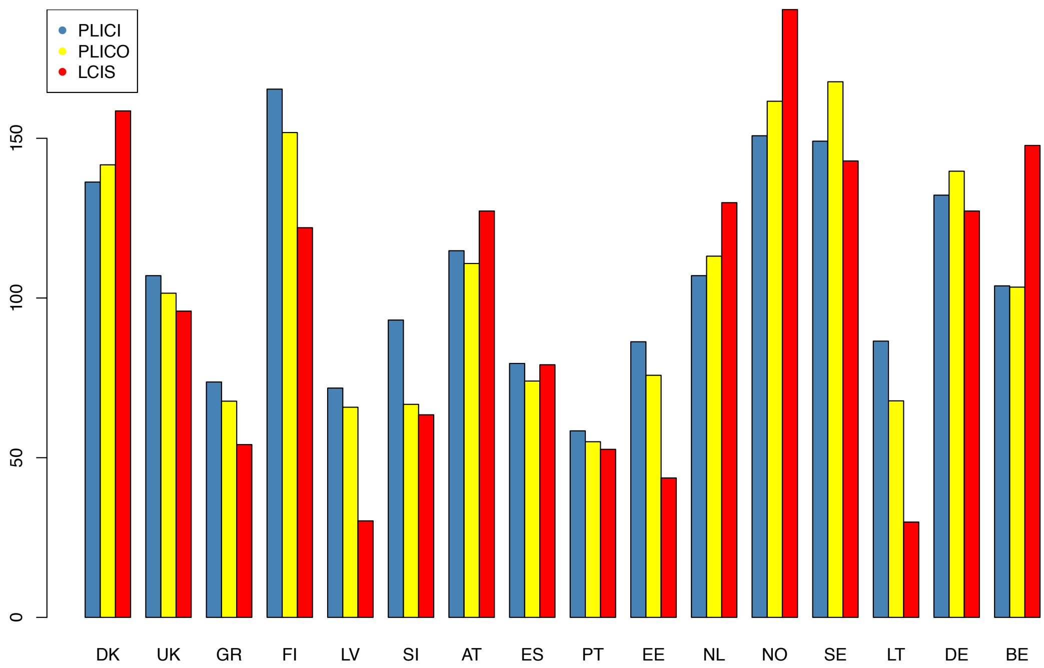

Figure 4-3 Labor cost indexes (LCIS = EUROSTAT, PLICI=Civil engineering, PLICO = Construction)

OPEX: Inflation adjustment

4.53 Opex data has been collected for 2013-2017 (70 observations). Hence, an indexation to a base year is necessary to make the costs comparable over the years. As for CAPEX, the harmonized price index for overall goods (HICPOG) is used, defining 2017 as the base year.

OPEX: Currency conversion

4.54 All national currencies are converted to EUR in 2017 by the average annual exchange rate.

4.18 Benchmarked CAPEX

4.55 As accounting procedures, depreciation patterns, asset ages and capital cost calculations differ between countries and sometimes even between operators depending on their ownership structure, the CAPEX needs to be completely rebuilt from the initial investment stream and up. In addition, a real annuity must be used since the application of nominal depreciations (even standardized) would immediately introduce a bias towards late investments. The steps involved in the calculation of benchmarked CAPEX are given in Figure 4-4 below.

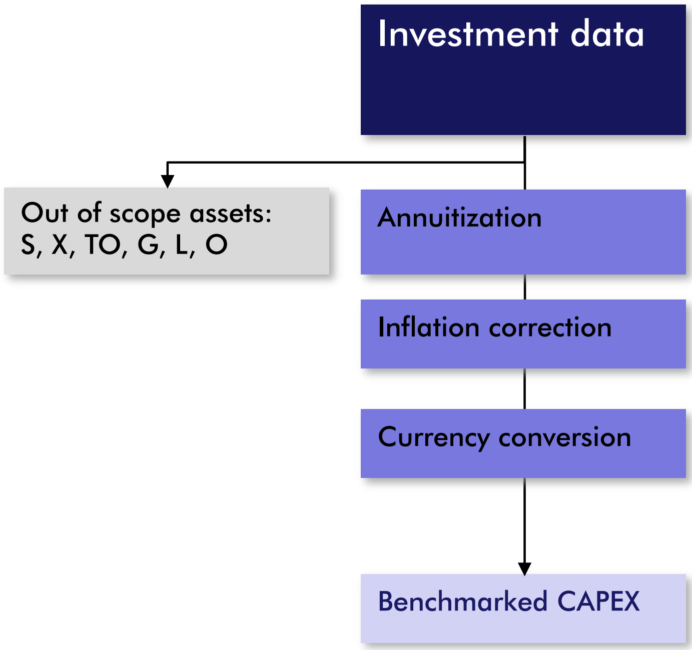

Figure 4-4 Steps in deriving benchmarked CAPEX.

CAPEX: Investment stream data

4.56 The starting point is the full investment stream reported by the operators from 1973 to 2017. Separating assets related to activities out of scope (S, X, TO, G, L, O), the residual investment stream is divided by type of asset as:

1) Pipelines,
2) Regulators,
3) Compressors,
4) Connections points,
5) Meter stations,
6) Control centers,
7) Other equipment.

CAPEX: Standard life times

4.57 The differentiation in investment is subject to different techno-economic life times, i.e. the standard real annuities constituting CAPEX.

4.58 The standard life times per asset class are given in Table 4-2 below.

Table 4-2 Standard techno-economic life times.

Asset class	Life time (yrs)
Pipelines	60
Pressure regulators	30
Metering stations	30
Connection points	30
Compressors	30
Control centers	20
Other assets	20
Other equipment	10

4.59 Assets acquired as used of any asset class are collected with original commissioning year or the expected remaining life time. The reported residual life is used for the annuity calculation for used assets, bounded above at the standard life time in Table 4-2 for new assets.

CAPEX: upgraded or (significantly) rehabilitated assets

4.60 In case the asset has been significantly rehabilitated the rehabilitation year also needs to be provided. Significant rehabilitation means a large incremental investment into an existing asset without change of any characteristics (i.e. its dimensions and properties). Large is defined as at least 25% of the (real) initial investment. Regular preventive and reactive maintenance, e.g. replacement of system components at or before their lifetime is not counted as a "rehabilitation". See also the note in Appendix D.

4.61 Investments changing the characteristics are considered as "upgrades" and not as rehabilitation.

4.62 Investments linked to upgrading assets that change asset class are counted as new investments. Thus, the original asset is replaced in the asset data with the new asset.

CAPEX: corrections

4.63 The following items are used for the correction of the investment stream prior to the calculation of the annuities:

1) Capitalized costs for out-of-scope assets (see Cost reporting guide)
2) Capitalized costs for financial costs (construction interest)
3) Capitalized taxes, fees and levies
4) Direct subsidies, exceptional direct depreciation and internal labor as direct expense.

4.64 Capitalized cost for out-of-scope assets, financial costs and taxes etc. are deducted from the gross investment stream.

4.65 Direct subsidies and exceptional depreciation are added to the gross investment stream.

CAPEX: Real annuities

4.66 Capex consists of depreciation and a return on capital. The actual investment streams are annuitized using a standard annuity factor α (r,T), where r stands for a real interest rate; and T stands for the average life-time of the investments in the respective year, calculated from the shares in art 4.56. The annual investments from the investment stream data are multiplied with the annual standard annuity factor α (r,T).

4.67 The numerical values for the annuity factors are provided to each TSO in a specific file.

CAPEX: Real interest rate

4.68 The real interest rate in the TCB18 project is set to 3% for the base run. The sensitivity with respect to this parameter is subject to an analysis reported in art 5.25 below.

CAPEX: Inflation adjustment

4.69 The current value of the past investments relative to the reference year is calculated using inflation indexes. Ideally, a sector-relevant index would capture both differences in the cost development of capital goods and services, but also the possible quality differences in standard investments. However, such index does not exist to our best knowledge. Several indexes have been collected from EUROSTAT and OECD, see Table 4-3. In this study, contrary to earlier projects, a Harmonized Inflation Index for overall goods and services has been used, HICPOG. The index is specifically developed for international comparisons, which is not the case with conventional indexes such as CPI and PPI. This provision is ensured by selecting comparable services and goods for the index, rather than those potentially only being used domestically.

Table 4-3 Inflation correction indexes tested (HICPOG used).

Index	Source	Type	Scope
Cpio	OECD	CPI	General
Cpiw	WorldBank	CPI	General
PPI	OECD	PPI	Producer goods
Hicpg	EUROSTAT	HICP	General
Hicpog	EUROSTAT	HICP	Overall goods
Hicpig	EUROSTAT	HICP	Industrial goods
Hicpmh	EUROSTAT	HICP	Maintenance

4.70 In addition, we have evaluated further indexes (CPI and other harmonized indexes) in the sensitivity analysis. Sector-specific indexes only exist for a handful of countries and require additional assumptions to be used for countries outside of their definition.

CAPEX: Currency conversion

4.71 As for OPEX, all amounts are converted to EUR values in 2017 using the average exchange rates. The exchange rates (annual averages of daily rates) used are provided among the public parameter files.

CAPEX: Old Capex

4.72 Investment stream data prior to 1973 are not required and by default are excluded, since they do not always exist or being of lower quality. However, without any correction this would create a bias towards operators with later opening investments, since these also include earlier assets. Thus, the calculation of the comparable Capex includes a residual element in 2017 corresponding to the pre-1973 assets still in the asset base. The calculation is equivalent to a Capex Break for 1973, that is the Capex unit cost from

1973 to 2017 is assumed prevail also up until 1973. In this manner, the inclusion of pre-1973 assets do not change the Capex-efficiency, but assures comparability. The calculated value, CapexOld, is capped by the sum of incumbent investments if known and validated. The methodology for the CapexBreak is described in Appendix E.

4.19 Benchmarked TOTEX

4.73 Summing up in Figure 4-5, we obtain the benchmarked Totex as the sum of Opex and Capex where C_{ft} is the total OPEX for firm f and time t after currency correction, I_{fs} is the investment stream for firm f and time s after inflation and currency correction, and $a(r,T)$ is the annuity factor for asset with life time T and real interest rate r.

Figure 4-5 Benchmarked Totex = Opex + Capex

4.20 Normalized Grid

4.74 Technically, the relevant scope is provided by an asset base consisting of:

1) Pipeline system
2) Compressor system
3) Pressure regulators and metering stations
4) Connection points
5) Control centers

4.75 A very detailed dataset was collected for the four asset categories above. Naturally, it does not make sense just to sum the different asset together since they correspond to different dimensions, pressure levels, material choices and capacities. Likewise, the geographical nature of the pipeline system makes it ideal to capture the environmental challenges through the following factors (see Appendix F):

1) Land use
2) Subsurface features
3) Topography
4) Soil humidity

4.76 Based on the data specification, a cost-norm for the construction costs for the standard assets above was developed, including the cost increases due to the environmental factors above. The result is an asset aggregate that we call the Normalized Grid (NormGrid; NG). Note that this detailed cost norm is independent of the actual costs and investments of the individual operator; it provides average costs rather than best-practice (or worst-practice) estimates. However, it is more general than a simple cost catalogue since it provides a complete system of complexity factors that explain the ratio of cost between any two type of assets, irrespective of which year, currency or context it is applied to (within reasonable bounds of course).

4.77 The exact formulae for the NormGrid system are documented in Appendix F, accompanied by an Excel calculator made available for all project participants on the project platform. In addition, workshop W3 was specifically devoted to the development of the norm grid metrics.

4.78 The NormGrid measure for all assets is adjusted for joint ventures by scaling with the share of ownership reported. The same approach is also used for output indicators related to assets in joint ownership, e.g. towers, connection points and power measures.

4.79 The size of the grid as measured by the Normalized Grid (NormGrid; NG) is naturally a key driver for Opex and Capex. The NormGrid is the sum of Capex and Opex components, proportional to the same effects in the total expenditure.

4.80 The NormGrid Opex component is simply the weighted sum of assets in use at a given time, irrespective of their age:

$$NormGrid_{OPEX} = \sum_t \sum_a N_{at} w_a$$

where

N_{at} Number of assets of type a in use, acquired at time t

w_a OPEX weight for assets of type of type a.

4.81 The NormGrid component for Capex below, differs in two respects from the Opex component: first, it only concerns assets that are within their techno-economic life (=their annuity depreciation period), second, the weights are multiplied with the same annuity factors as for the corresponding investments:

$$NormGrid_{CAPEX} = \sum_t \sum_a n_{at} v_a \alpha(r, T_a)$$

where

n_{at} Number of assets of type a, acquired at time t and in prime age.

v_a CAPEX weight for assets of type of type a

r Real interest rate

T_a Techno-economic standard life for assets of type of type a

$\alpha()$ Real annuity function

4.21 Model specification

4.82 Any efficiency comparison should account for differences in the outputs and the structural environment of the companies. A key challenge is to identify a set of variables:

1) that describe the tasks (the cost drivers) that most accurately and comprehensively explain the costs of the TSOs;
2) that affect costs but cannot be controlled by the firm (environmental factors); and
3) for which data can be collected consistently across all firms and with a reasonable effort.

4.83 Conceptually, it is useful to think of the benchmarking model as in Figure 4-6 below. A gas TSO transforms resources X into services Y. This transformation is affected by the environment Z. The aim of the benchmarking is to evaluate the efficiency of this transformation. The more efficient TSOs are able to provide more services using less resources and in environments that are more difficult.

4.84 The inputs X are typically thought of as Opex, Capex, or Totex. In any benchmarking study and in an international benchmarking study in particular, it requires a considerable effort to make costs comparable. We have found in previous studies that a careful cost reporting guide is important to make sure that out-of-scope is interpreted uniformly, and that differences in depreciation practices, that taxes, land prices, labor prices etc. are neutralized.

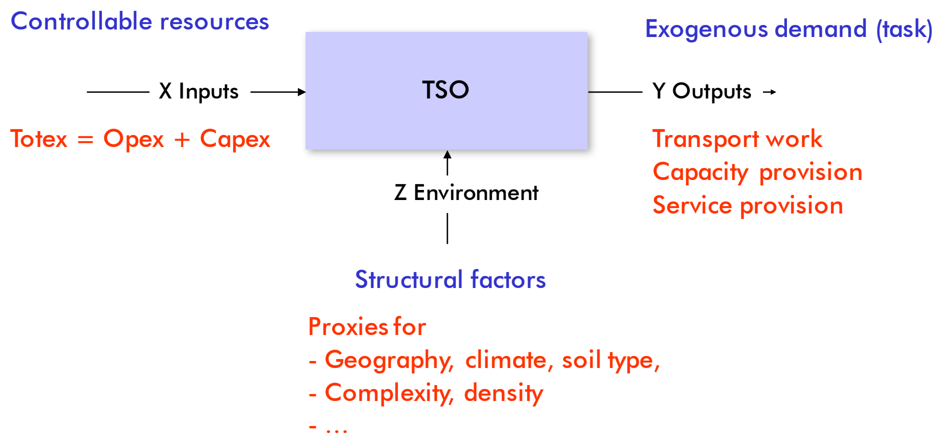

Figure 4-6 Conceptual benchmarking model

4.85 The outputs Y are made of exogenous indicators for the results of the regulated task, such as typically variables related to the transportation work (energy delivered etc.), capacity provision (storage volume, peak load, coverage in area etc.) and service provision (number of connections, customers etc.). Ideally, the output measures the services directly. In practice, however, outputs are often substituted by proxies constructed as functions of the assets base, like km of pipelines, number of connections, installed power of compressors etc. One hereby runs the risk that a TSO could play the benchmarking-based regulation by installing unnecessary assets. In practice, however, we have found that this is not a major risk in the early stages of the regulation and that the advantages of using such output indicators outweigh the risk. We shall therefore think more generally of the outputs as the cost drivers.

4.86 The class of structural variables Z contains parameters that may have a non-controllable influence on operating or capital costs without being differentiated as a client output. In this class we may often find indicators of geography (topology, obstacles), climate (temperature, humidity, salinity), soil (type, slope, zoning) and density (sprawl, imposed feed-in locations). One challenge with this class of parameters is that they may be difficult to validate statistically in a small data sample. Their role of potential complicating factors will therefore have to be validated by other studies or in a process of individual claims from the TSOs. Another challenge is that in a small dataset, the explicit inclusion of many complicating factors will put pressure on the degrees of freedom in a statistical sense. This is also the approach we have taken in this study. We have used elaborate engineering weight systems of the grid assets to reflect the investment and operating conditions. In this way, Z factors can to a large extent be captured by the traditional Y factors.

4.87 To ensure that the model specification is trustworthy, it is important to decide on some general principles as well as some specific steps. Based on our experience from other projects, we have in this project focused on the following generic criteria:

1) **Exogeneity** – Output and structural parameters should ideally be exogenous, i. e. outside the influence of the TSOs.
2) **Completeness** – The output and structural parameters should ideally cover the tasks of the TSOs under consideration as completely as reasonable.
3) **Operability** – The parameters used must be clearly defined and they should be measurable or quantifiable.
4) **Non-Redundancy** – The parameters should be reduced to the essential aspects, thus avoiding duplication and effects of statistical multi-collinearity and interdependencies that would affect the clear interpretation of results.

4.88 In reality, it is not possible to stick to these principles entirely. In particular, exogeneity must be partly dispensed with since the network assets are endogenous but also in many applications providing good approximations of the exogenous conditions. To rely entirely on exogenous conditions would require a project framework that far exceeds the present both economically and time wise.

4.89 The process of parameter selection combines engineering and statistical analysis. We have in this project used the following steps:

1) **Definition of parameter candidates**. In a first step we established a list of parameter candidates which may have an impact on the costs of TSOs. The relationships between indicators and costs must be plausible from an engineering or business process perspective.
2) **Statistical analysis of parameter candidates**. Statistical analysis was then used to test the hypotheses for cost impacts for different parameter candidates and their combinations. The main advantage of statistical analysis is that it allows us to explore a large number of candidate parameters and to evaluate how they individually and in combination allow us to explain as much as possible of the cost variation.
3) **Plausibility checks of final parameters**. The final parameters from the statistical analysis are finally checked for plausibility. This plausibility check is based *inter alia* on engineering expertise.

4.90 The model specification steps above have supported the model specification process. However, model development in transmission operation benchmarking is not a datamining exercise that follows blindly from statistical analyses aiming at predictive models. It may be that some parameters that help explain average costs have little techno-economic sense or explanatory power in the frontier-based benchmarking model and vice versa. The model specification steps have therefore been combined with careful second stage analysis to ensure that no frontier relevant parameters have been left out.

4.22 Benchmarking methods

4.91 Econometrics has provided a portfolio of techniques to estimate the cost models for networks, illustrated in Table 4-4 below. Depending on the assumption regarding the data generating process, we divide the techniques in *deterministic* and *stochastic*, and further depending on the functional form into *parametric* and *non-parametric* techniques. These techniques are usually considered state of the art and are advocated in regulatory applications provided sufficient data is available.

Table 4-4 Model taxonomy.

	Deterministic	Stochastic
Parametric	Corrected Ordinary Least Square (COLS) Greene (1997), Lovell (1993), Aigner and Chu (1968)	Stochastic Frontier Analysis (SFA) Aigner, Lovell and Schmidt (1977), Battese and Coelli (1992), Coelli, Rao and Battese (1998)
Non-Parametric	Data Envelopment Analysis (DEA) Charnes, Cooper and Rhodes (1978), Deprins, Simar and Tulkens (1984)	Stochastic Data Envelopment Analysis (SDEA) Land, Lovell and Thore (1993), Olesen and Petersen (1995)

4.92 In a study of European gas TSOs, the number of observations is too small for a full-scale application of SFA as main instrument. We have therefore used DEA as our base estimation approach, in line with regulatory best practice and earlier studies such as E2GAS and E3GRID. The DEA method is by now well established in the scientific literature as well as in regulatory applications, and we shall therefore not provide a theoretical description of it here. Further details are provided in e.g. Bogetoft and Otto (2011)

4.23 Frontier outlier analysis

4.93 *Outlier analysis* consists of screening extreme observations in the frontier model against average performance. Depending on the approach chosen (OLS, DEA, SFA), frontier outliers may have different impact. In DEA, particular emphasis is put on the quality of observations that define best practice. The outlier analysis in DEA can use statistical methods as well as the dual formulation, where marginal substitution ratios can reveal whether an observation is likely to contain errors. In SFA, outliers may distort the estimation of the curvature and increase the magnitude of the idiosyncratic error term, thus increasing average efficiency estimates in the sample. In particular, observations that have a disproportionate impact (influence or leverage) on the sign, size and significance of estimated coefficients are reviewed using a number of methods (cf. Agrell and Niknazar, 2014).

4.94 In non-parametric methods, extreme observations are such that dominate a large part of the sample directly or through convex combinations. Usually, if erroneous, they are fairly few and may be detected using direct review of multiplier weights and peeling techniques. The outliers are then systematically reviewed in all input and output dimensions to verify whether the observations are attached with errors in data. The occurrence and impact of outliers in non-parametric settings is mitigated with the enlargement of the sample size.

Outlier detection in DEA

4.95 In frontier analysis, the observation included in a reference or evaluation set is called a Decision Making Unit (DMU). A DMU can be an observation of (inputs,outputs) for a firm at a given time (cross section) or at other time periods (panel data). Outlier DMU may belong to a different technology either by errors in data, or unobserved quantities or qualities for inputs or outputs. The identification of DMUs to check more carefully has used in particular two approaches.

4.96 The outlier detection used in the final runs follows the German Ordinance for Incentive Regulation and the notion of DEA outliers herein (ARegV, annex 3). The invoked criteria are consistent with the method proposed and used in Agrell and Bogetoft (2007), representing a systematic and useful device to improve the reliability of regulatory benchmarking without resorting to *ad hoc* approaches. The idea is to use a dual screening device to pick out units that are doing extreme as individual observations and that are having an extreme impact on the evaluation of the remaining units. To do so, we use a super efficiency criterion similar to the Banker and Chang (2005) approach, although we let the cut-off level be determined from the empirical distribution of the super efficiency scores. In addition, we use a sums-of-squares deviation indicator similar to what is commonly seen in parametric statistics.

4.97 Let Ω be the set of n TSO in the data set and k be a potential outlier. Then define $E(h, \Omega)$ be the efficiency of a TSO h when all TSO are used to estimate the technology and let $E(h, \Omega/k)$ be the efficiency when TSO k does not enter the estimation. We can therefore evaluate the impact on the average efficiency by

$$\frac{\sum_{h\in \Omega/k}(E(h,\Omega/k)-1)^2}{\sum_{h\in \Omega/k}(E(h,\Omega)-1)^2}$$

4.98 Large values of this as evaluated in a $F(n\text{-}1,n\text{-}1)$ distribution, cf. Banker (1996), will be an indication that k is an outlier.

4.99 Using also the super-efficiency criteria of the Ordinance (ARegV), we shall classify an entity k as an outlier to be eliminated if

$$E(k,\Omega/k) > q(0.75) + 1.5(q(0.75) - q(0.25))$$

4.100 where $q(\alpha)$ is the α-fractile of the distribution of super-efficiencies, such that e.g. q(0.75) is the super-efficiency value that 75% has a value below. Hence, this criterion indicates if there are units that are having much higher super-efficiencies than the other units. If the distribution is uniform between 0 and 1 in a large sample, for example, all other units are evenly distributed between 0 and 1, a candidate unit must have a super efficiency above 0.75+1.5*(0.75-0.25) =1.5 to be an outlier.

4.24 Measures for incorporation of passive TSOs

4.101 In the gas benchmarking, the 16 German TSOs participated only through their NRA, making previously validated data for the year 2015 available for the study. In the data collection of environmental data, the same provisions as for other TSOs prevailed. However, certain data specifications differing from E2GAS were not exploited for these TSOs. This includes

1) Control centers
2) Upgraded / rehabilitated assets
3) Indirect cost allocations
4) Compressor power per individual compressor

4.102 The absence of data for the points above could entail an underestimation of the cost efficiency for the passive TSOs. Nevertheless, the presence of the data improves the benchmarking by permitting a validation of the shape and significance of the terms, without necessarily fully explaining the residual term for themselves.

4.25 Allocation key for indirect costs

4.103 Several allocation methods were tested for indirect cost onto benchmarked functions. The staff data intensity was considered biased since it excludes external services. Thus, the retained key is based on direct costs, excluding energy and depreciation, for the respective activities, including out-of-scope and non-benchmarked activities.

5. Benchmarking results

This Chapter provides some general and average results from the benchmarking, without providing any information that may lead to the identification of individual operators and their results. The results from the robustness analysis are also included and commented.

5.1 Model specification

5.01 Based on conceptual thinking and a statistical analysis reported during Workshops W4 and W5, the final model specification in the TCB18 project includes four cost drivers as shown in Table 5-1 below.

Table 5-1 Model specification: Final model GAS.

Variable	Definition
INPUT	
dTotex.cb.hicpog_plici	Totex excl energy, Capex break, inflation index HICPOG, labor cost adjusted in OPEX with PLICI
OUTPUT	
yNG_zSlope	NormGrid assets weighted by slope zSlope (% of service area) x complexity factors per class
yConnections_tot	Total number of connection points
yCompressors.power_tot	Total installed compressor capacity (MW)
yPipes_Landhumidity	Total pipeline length, weighted by wetness factor zLandhumidity (% of service area) x complexity factor

Input in the model is total expenditure (Totex). It is calculated as standardized capital costs using real annuities and after correcting for inflation and currency differences plus standardized operating costs excluding cost of energy, out-of-scope activities. See the explicit formula in Chapter 4 on methods. Labor cost expenditures in Opex are adjusted to average European costs by the PLICI labor cost index. The final model is using four outputs: normalized grid (weighted sum of all grid components as explained in section 4.20) weighted with the slope class complexity factor, total number of connection points, the total capacity (measured as compressor capacity) and the pipeline length weighted with humidity severity factors. These parameters capture both the investment (capital expenditure) dimension through the normalized grid and the operating cost dimension through the connections and capacity, leading to good explanatory results for the average cost in the sample. In general, the strongest candidate in the frontier models is the normalized grid. The next strongest cost driver candidate is the connections and the weakest candidate statistically is the humidity factor. Together the factors form a very strong explanatory base for the transmission system operators.

5.02 An initial proposal presented at Workshop W5 had a different parameter to capture capacity provision, total injection volume of H and L gas in nm^3. After discussions with project participants and additional analysis, it was decided to replace this parameter with compressor power, which is a parameter independent on exogenous events (temperature and business cycles) and related to actual deployed capacity rather than recent investments (i.e. the normalized grid component).

5.03 The final model has one parameter more than the E2GAS model[2] (three parameters), which reflects both a larger reference set (29 vs 22), but also a more advanced consideration of environmental conditions through GIS-collected data for both slope classes and soil humidity. It is therefore logical that the new model should explain costs to an even higher extent than in the previous study. Nevertheless, it is interesting to note that the same classes and types of parameters are found in the two models, differing primarily on the capacity provision dimension (compressor power vs nominal maximal capacity). The logic of the model specification with respect to the earlier categories is illustrated in Figure 5-1 below.

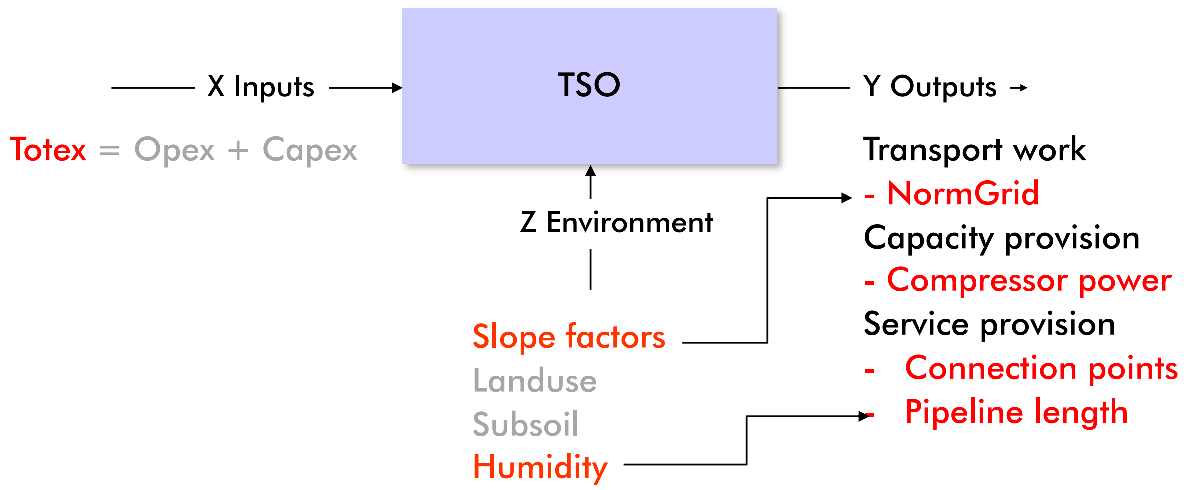

Figure 5-1 Final GAS model with service categories.

yNormGrid_zSlope

5.04 The NormGrid provides a Totex-relevant proxy for the total pipeline system, summing all relevant assets with weights corresponding to their Capex and Opex impact. As documented in the engineering study (Appendix F), certain environmental conditions influence the cost of constructing and operating the pipeline system. These factors include land use type, topography (slope), vegetation type, soil humidity, subsurface features (rockiness, stones), extreme temperatures and salinity. Extensive statistical tests revealed correlations and interaction between several of the factors, e.g. vegetation and landuse type, subsurface features and topography. The most important factor for pipelines was topography (slope class), relating to costs of construction (reinforcements, site access) and to operation (maintenance access). Most other factors, with the exception of humidity, correlates with the normalized grid slope-weighted parameter. Thus, this parameter was chosen as the primary variable, explaining by itself over 90% of the variance in Totex in robust regression (cf. Table 5-4 below).

yConnections_tot

5.05 The connection points in the transmission grid can be of different types, see Table 5-2. Each of these connections causes certain costs of operation, metering, monitoring etc.

[2] yNormGrid, yConnections_tot and yCapacity_max (max of injection and delivery peak flow).

Statistically, the sum of the connections, **yConnections_tot** is the preferred variable, adjusted for ownership asset by asset.

Table 5-2 Type of connection points (Appendix A, art 63)

Type	Description
I_N	Injection from upstream net/production/injection from biogas/LNG
I_P	Injection from production plant
I_S	Injection from storage
D_D	Delivery to downstream network
D_C	Delivery to customers, direct withdrawal
D_N	Delivery to neighbouring networks
D_S	Delivery to storage

yCompressors.power_tot

5.06 The transport capacity of the transmission system is measured through the sum of the installed power of the compressors units, adjusted for ownership, irrespective of type of compressor unit (Appendix A, art 58) and type of unit. The parameter is frequently used in international comparisons and correlates both to Capex and Opex in the sense that a higher capacity requires direct costs for operation and maintenance.

yPipes_Landhumidity

5.07 The pipeline installations specifically, not so much other assets, are subject to cost increases resulting from high humidity and wet soil. This results from more expensive construction site management, drainage, isolation not accounted for in the NormGrid and resources devoted to evacuation of water for repairs and preventive maintenance of segments. It was seen that dimensioning and age were not proportional to these costs. Thus, to capture this effect a technical parameter was created from the unweighted pipeline length combined with the landhumidity factors from the GIS and engineering calculations. Increasing the explanatory value, it forms a good complement to the primary parameter for transport work, yNormgrid_zSlope.

5.2 Summary statistics

5.08 Summary statistics of the costs and cost drivers in the base model is shown in Table 5-3 below. (Note that range values cannot be provided for confidentiality reasons). Q1 denotes first quartile, Q3 third quartile and Q2 the median.

Table 5-3 Summary statistics for model variables (2013-2017, full sample, n = 70)

Variable	Mean	Q1	Q2 (median)	Q3
dTotex.cb.hicpog_plici	151,486,489	31,977,516	55,917,669	261,211,605
yNG_zSlope	125,349,439	29,248,670	42,600,584	139,369,517
yConnections_tot	289	40	190	451
yCompressors.power_tot	242,866	42	22,824	425,697
yPipes_Landhumidity	3,576	1,073	1,385	4,480

5.09 We see that the gas TSOs in the sample vary in terms of size. The 25% largest gas TSOs are approximately 5 times larger than the 25% smallest TSOs. Also, we see that the mean values exceed the median values. This reflects that the size distributions have a relatively long right tail.

5.10 To get an initial understanding also of the ability of these cost drivers to explain the variation in average costs together and individually, Table 5-4 below shows the adjusted R2 (the conventional measure of regression fit) of three ordinary regression models with 1, 2, 3 and 4 cost drivers. We see that the adjusted R2 of a model with only **yNG_zSlope** is 96.7%. Adding **yConnections_tot** as a cost driver brings us to an adjusted R2 of 97.5%. When we add also **yCompressors.power_tot**, the adjusted R2 attains 97.9%. Finally, the addition of the humidity parameter **yPipes_Landhumidity** increases the explanatory power to 98.7%. Three TSOs were identified as statistical outliers in all regressions in this example.

Table 5-4 Explanatory power in 1, 2, 3 and 4-variable models, robust regressions.

Number of variables	Cost driver(s)	Adjusted R2
1	yNG_zSlope	0.967
2	yNG_zSlope + yConnections_tot	0.975
3	yNG_zSlope + yConnections_tot + yCompressors.power_tot	0.979
4	yNG_zSlope + yConnections_tot + yCompressors.power_tot + yPipes_Landhumidity	0.987

Outliers

5.11 The analyses of the raw data as well as the analysis of a series of model specifications, i.e. models with alternative costs drivers, suggest that one of the 29 TSOs almost always is an extreme outlier. This TSO has therefore been permanently removed from the reference set. In addition, five others have been identified using the model specific outlier detection tests explained in section 4.23, making in all six TSOs frontier outliers.

Returns to scale

5.12 For all possible model specifications, we have also tested which of the returns to scale assumptions in the DEA model fit data the best: variable returns to scale (VRS), increasing returns to scale (IRS), decreasing returns to scale (DRS), or constant returns to scale (CRS). We have done so using F-tests based on a goodness-of-fit measure as explained in the Method chapter. The general finding is that the IRS assumption (see Figure 5-2 below) is the best assumption to invoke. This is supported also by parametric analyses for a logarithmic model, where the coefficients sum to less than one for the selected parameters.

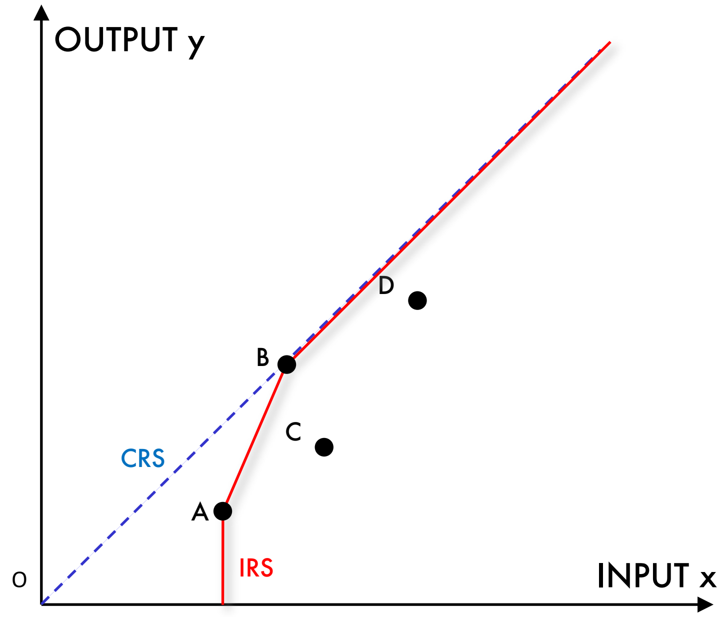

Figure 5-2 DEA frontier under increasing returns to scale (IRS).

5.13 The IRS assumption means that it can be a disadvantage to be a small TSO but not to be a large TSO. In Figure 5-2 the large TSO D is benchmarked against the most productive (CRS-efficient) TSO B, the somewhat smaller TSO C is gauged against the standard set by TSO B and A, whereas TSO A (smaller than B) forms a frontier unit for its scale class. This is also conceptually appealing. A TSO can be small due to the size of the country or by the service area it has to serve and there may be an element of fixed costs involved in the operation of any TSO. On the other hand, if a TSO is suffering from extra cost of being large, it is likely that a reorganization of the TSO to imitate a combination of smaller TSOs could improve cost efficiency.

5.3 Exclusion of significant rehabilitation

5.14 Although informed in the data specification and at workshops, only very few TSOs used the reporting options for significant rehabilitations (see Appendix D). Worse, of those reporting some TSOs reported proportions of their assets base under significant rehabilitation that do not correspond to any reasonable techno-economic policy. As stated in Appendix D, significant rehabilitations are not equivalent to preventive or corrective maintenance, but major overhauls with part of the initial asset retained in a different state. In order not to compromise the data quality, the PSG decided to exclude the significant rehabilitation from the benchmarking runs.

5.4 Efficiency scores

5.15 The efficiency scores are obtained using DEA on the final model described. The primary static result concerns the 2017 data for all except German TSOs, for which 2015 was used as benchmark year (costs were indexed to 2017).

Final model efficiencies

5.16 Summary statistics for the efficiency scores in the final TCB18 model are shown in Table 5-5 below. We see that the DEA model leads to average efficiencies of 79%, i.e. the model suggests that the gas TSOs on average can save 21% in benchmarked comparable Totex.

Table 5-5 Efficiency scores in final model GAS, static 2017/2015

	Mean	Q1	Q2 (median)	Q3
Final DEA (2017)	0.793	0.631	0.881	1.000
Peers (non-outliers)	4			
Outliers	6			

5.17 In Table 5-5 we see all the quartiles of the efficiency distribution and we note that there is a longer left tail in the sense that the median is now to the right of the mean value. This is also illustrated in Figure 5-3 below.

5.18 The full distribution of the efficiencies is shown in Figure 5-3. We note also here the relatively large number of fully efficient TSOs. This is not surprising since we are using a model with four cost drivers on a small sample and with cautious (aggressive) outlier elimination instruments. Indeed, in the base model there are six DEA outliers as also discussed in art 5.11. The decreasing number of outliers (three in the preliminary run) indicates that the data changes have led to convergence in the dataset.

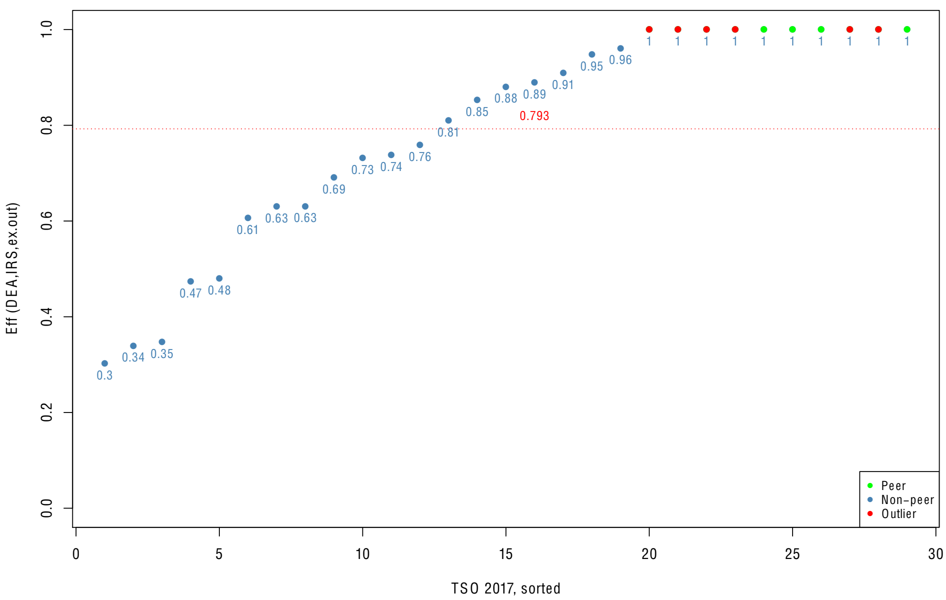

Figure 5-3 Distribution of scores 2017 in the final GAS model.

5.5 Robustness analysis

5.19 The final model provides a cautious estimate of the cost efficiency in gas transmission, in line with the E2GAS results in terms of level and distribution.

5.20 The revision of a preliminary model to incorporate compressor power rather than injection volume made the final model more stable and less dependent on exogenous factors.

5.21 Overall, the model constitutes an improvement in the consideration of economic, environmental and infrastructure factors. Although a selection has been made among the derived environmental factors, the correlations among them render the specification robust.

Sensitivity for model parameters

5.22 The results have been tested for changes with respect to the following model parameters:

1) Interest rate
2) Normgrid weight – calibration between Opex and Capex
3) Normgrid weight for pipelines vs other assets
4) Salary corrections for capitalized labor in investments

5.23 All analyses are relative to the impact of a parameter change, say q on the DEA score for the base case used in the final run, q_0. For each TSO k, the impact of q is measured as:

$$E(k|q) / E(k|q_0)$$

5.24 The illustrations below concern the mean effects on the 2017 dataset, i.e. the final scores. A negative slope for the function above would imply that increasing the parameter q would lead to a decrease in mean score, the vertical axis gives an indication of the percentage change in score expected.

Sensitivity to interest rate

5.25 The results for the sensitivity to interest rate changes show a relatively flat and predictable shape. Lowering the interest rate to 1.8% (-40% of the 3% base rate) would on average increase the DEA score by 5% (proportionally, the maximum change is 13% units), likewise an increase to 4.5% (+50% on base rate) would on average decrease the DEA scores by 5% (maximum unit change: -13%). The outcomes are illustrated in Figure 5-4 below. The vertical axis denotes the change in average DEA scores relative to the average DEA scores calculated with interest rate 3%.

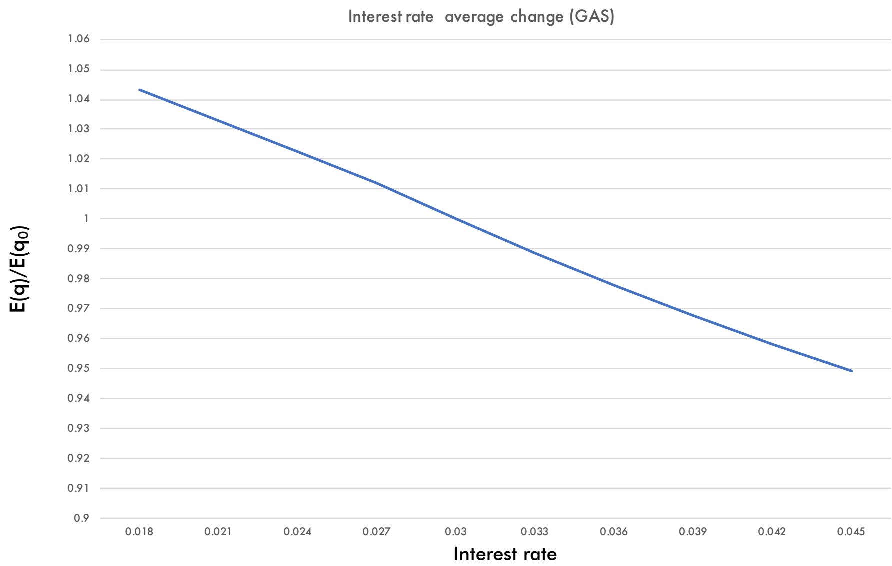

Figure 5-4 Sensivity to changes in real interest rate (proportional change in DEA score). .

Sensitivity to NormGrid weights

5.26 Each TSO has thousands of assets of different types and dimension, each assigned a specific value in the Normgrid system. Given the large number of assets and their dispersion, the impact of a change to an individual weight is of course minimal. But even systematic changes to the balance between Opex and Capex weights and to specific asset groups (here: pipelines) result in very small changes to the DEA scores, as seen in Figure 5-5 and Figure 5-6 below. The explanation for this stability is that the types of assets are relatively equally shared among the TSOs and the changes in absolute numbers hardly affect the relative ratios among the TSOs. The vertical axis denotes the change in average DEA scores relative to the average DEA scores calculated with the base values used in the NormGrid system (=1), multiplied with a factor ranging from 0.2 to 2.

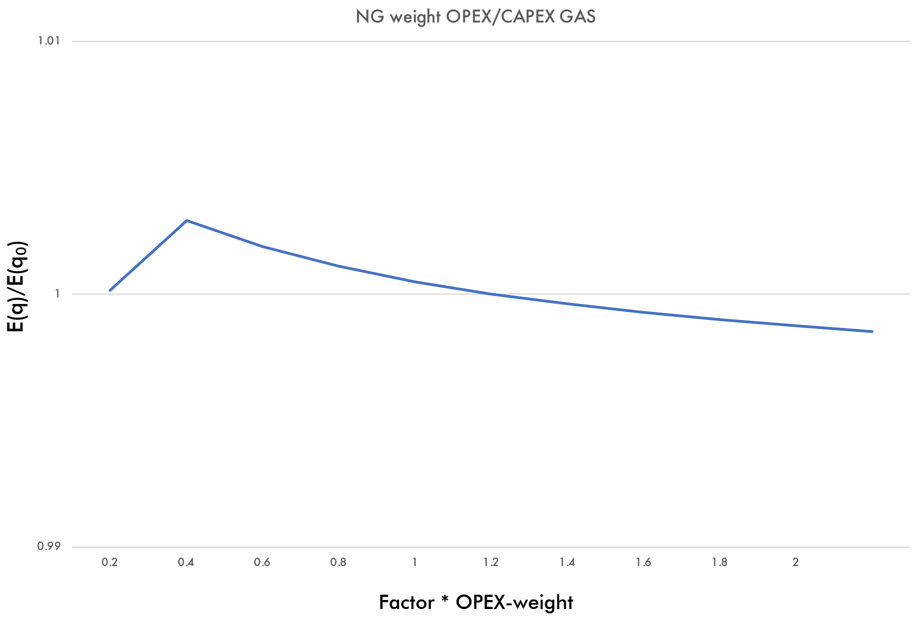

Figure 5-5 Sensitivity analysis wrt to NormGrid weights calibration Opex-Capex (change in DEA score).

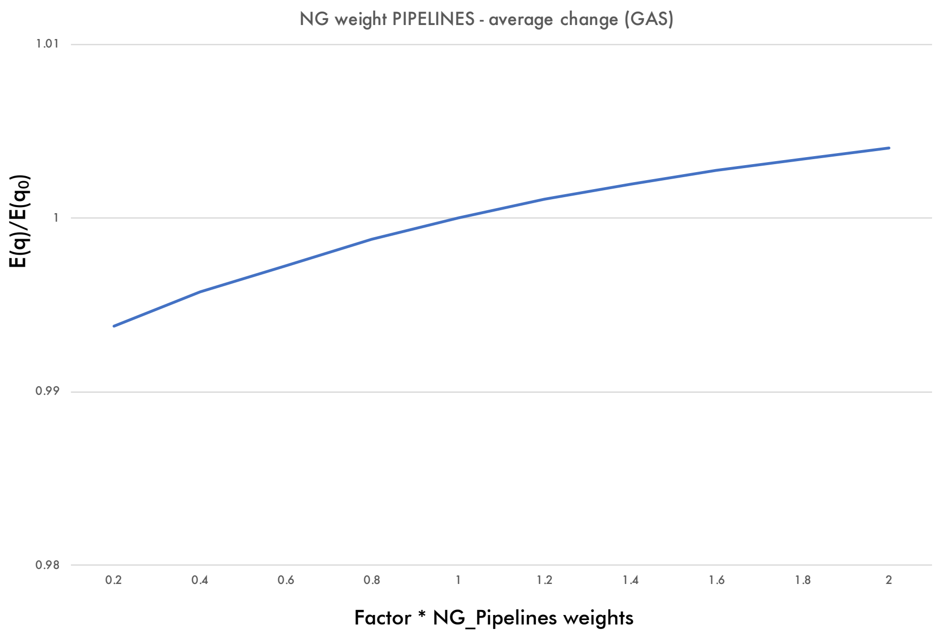

Figure 5-6 Sensitivity analysis of NormGrid weights for pipelines (factor of change) vs change in DEA score.

Sensitivity to salary corrections for investments

5.27 In E2GAS, a share of the investment stream was considered as local labor cost and subject to the same salary adjustment as in OPEX. In TCB18 this is not the case as the identification of the constructors in past investments is uncertain and the economic interpretation (closed markets) is in conflict with promoted best practice in other infrastructure areas. The sensitivity of the results with respect to this choice is illustrated in Figure 5-7 below. The average change is minimal, less than 1% for a 25% labor share, but the individual impact of course depends on the weight of investments in Totex and the salary correction factor compared. The maximum range of impact here in the interval (-8% to +4%) in percentage-units for the score confirms that even on an individual basis, the results are not primarily driven by country-specific labor cost differences.

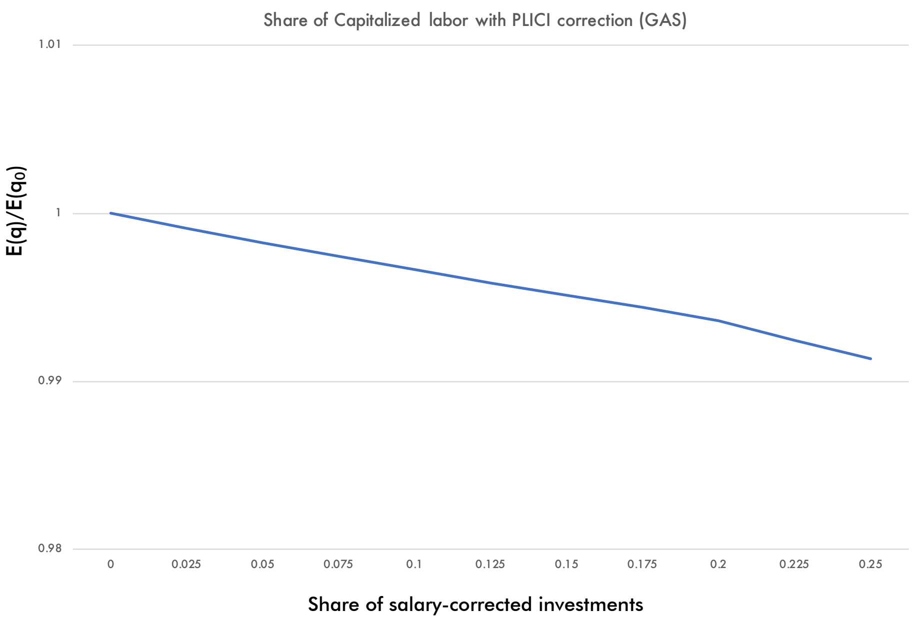

Figure 5-7 Sensitivity analysis for salary correction of capitalized labor in investments.

5.28 The sensitivity analysis confirms that the results are robust to changes in the model parameters (interest rate, norm grid parameters) or model assumptions (capitalized labor in investments).

5.6 Heterogeneity

5.29 The network operators in the study have different structure in terms of connection points, capacity and pipeline length. A concept often invoked is to distinguish between "point-to-point" (PTP) operators with a strong transport-transit focus (see Figure 5-8 below) and "meshed" operators with many intersection and connection points (see Figure 5-9), focusing diverse tasks: e.g., transit, storage and supply security.

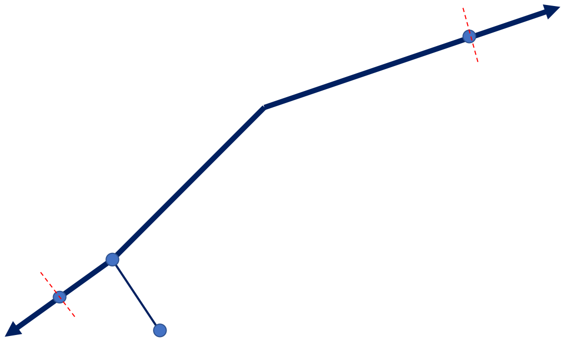

Figure 5-8 Example of "point-to-point" gas pipeline system.

5.30 In order to investigate whether the current model adequately addresses these structural heterogeneity, we created a proxy for heterogeneity as the number of connection points per pipeline km. This parameter varies considerably among the operators, but not over time, expressing an operator-dependent characteristics. If this parameter would explain the cost-efficiency score or even unit cost, this might indicate a problem of correcting for heterogeneity.

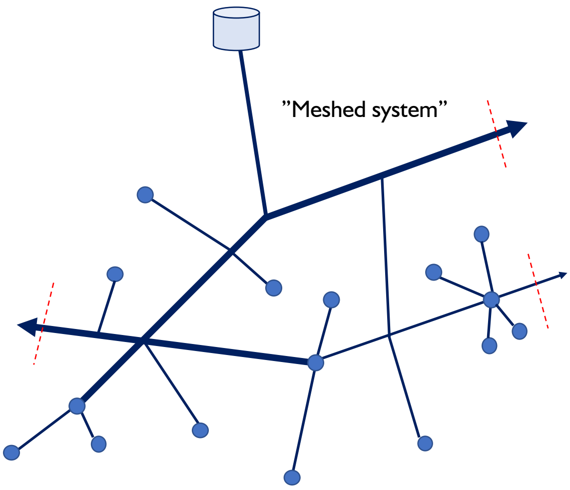

Figure 5-9 Example of dense "meshed" gas pipeline system.

5.31 The results from the analysis are clear: the heterogeneity proxy has no explanatory value on the DEA-score, nor on unit cost components in either ordinary or robust regression. This means that we can conclude that the current model is catering for the different types of operators through the model specification, in particular the explicit inclusion of the customer service (connection points), the physical network (pipeline length) and the dimensioning of the capacity (compressors). The PTP operators naturally have low output in terms of customer service and/or capacity provision, whereas the opposite is true for the meshed networks. The NormGrid warrants consideration of network assets even under low utilization, which also favors meshed operators.

6. Quality provision

In this chapter, the results from a survey on indicators and data for service quality in transmission.

6.1 Survey

6.01 So far measures of output quality have not been widely used in TSO cost efficiency studies. Still, CEER is open to add the aspect of quality to (future) benchmarks. For that, however, it is required to define the concept of quality, to find ways to measure it, and to be able to relate such measure(s) to benchmarked cost.

6.02 In order to get closer to answers, in October 2017 CEER initiated a survey among most TSOs participating in TCB18. The survey asked TSOs to suggest quality parameters that are of universal use, well defined, collectable, and verifiable with independent sources, and be as specific as possible regarding definition, interpretation, sourcing, availability, and verifiability.

6.03 CEER received responses to the survey from just one gas TSO.

6.04 The TSO remarks that in order to perform a benchmark on TSO efficiency with plausible and robust results, all non-efficiency related differences between TSOs should be controlled for, including differences in TSO quality parameters (quality of the network, products and services provided by TSO). This is especially important in a DEA benchmark, since in DEA all remaining differences between TSOs are labelled as inefficiency. The TSO concludes with suggesting a number of quality aspects: sustainability (e.g. compressing with green gas), environmental standards/norms, safety norms or other permits, standards for security of supply and peak delivery, delivery interruptions, type of capacity products offered (fixed vs. firm, etc.), monthly recalculation of technical capacity offered to customers, facilitating hub liquidity, balancing, connection task (universal service obligation, maintaining gas receiving stations), gas qualities transported, gas quality bound (Wobbe bandwidth), pressure levels at entry and exit points, odorization, implementation of network codes, and quality of IT systems (e.g. providing real time information).

6.05 The response was too low to work with in TCB18. Nevertheless, at the second TCB18 workshop of April 2018 the subject was found to be important enough to reinvestigate and it was agreed to revisit the survey, this time with a questionnaire that gives stronger guidance to what CEER is looking for. The survey was launched in October 2018 with a (extended) deadline of January 2019. The survey aimed at further exploring the business know-how at TSOs to investigate if quality of service provision could be defined meaningfully in terms of cost and cost efficiency. To that extent the survey focused on searching for concrete quality aspects and ways to measure these (parameters). CEER announced beforehand that the results of the survey were not meant to be used in the model of the current TCB18 benchmark. For the second survey CEER developed an Excel template to be filled in by TSOs and gave the following instructions in a separate guiding document.

6.06 First of all, CEER remarked in the guide that quality is not about *what* a TSO provides, but *how well* it is done. Therefore, CEER expects that a suggested quality aspect is of universal relevance. That is, if a quality aspect reflects the quality of a service that is not provided by all TSOs, the quality aspect may signal a benchmark scoping issue or something else rather than a quality issue.

6.07 Secondly, the quality aspects CEER is looking for:

1) must be interpretable, i.e. a quality aspect that has not at least an intuitive relation to cost will be difficult to use for the purpose of benchmarking. So, interpretability is more or less about the story behind the quality aspect in terms of cost and cost efficiency.
2) must be measurable as a parameter. For example, if the quality aspects is reliability, a parameter may be the number of service disruptions. It is important to define such parameters well, i.e. concrete, precise, and unambiguously.
3) must have a relation to cost. Apart from a more global interpretability of the quality aspect, it helps the analysis of the survey to understand the TSO's opinion on how specific cost parameters correlate to cost and asset components. The survey asked TSO's to link suggested quality parameters to cost items in the financial reporting sheet of the TCB18 data collection.
4) must be collectable. To use a quality parameter to interpret a benchmark result or to shape the benchmark model, the value of the parameter must be based on objective data that are collectable and verifiable.

6.08 The response to the second survey was again low in numbers, diverse, and in most cases not very concrete. A TSO mentions security of supply as quality aspect and suggest the N-1 criterion and the Import Route Diversification as parameters for it. The TSO also mentions the extent to which the risk of disruptions is mitigated as a quality aspect. For that aspect it suggests the number of interruptions and total system interruption duration as parameters for it. The same TSO mentions odorization (grams of odorant per nm3 of gas) and gas pressure regulation (outlet pressure) as quality aspects. A second TSO also mentions security of supply, parameterized with the level of interruptions. The third TSO that responded repeats its list of 17 aspects submitted in response to the first survey without additional information. Finally, the fourth TSO warns that relation to cost of quality aspects is often difficult to measure as many complexity factors play a role as well.

6.2 Analysis

6.09 The questionnaires were also sent in TCB18 to electricity TSOs. For electricity it seems clear that the reliability of transportation of energy (security of supply; measured by interruptions, energy not supplied, etc.), or actually the absence of it, appeals to what the users of the grid eminently experience as quality delivered by TSOs. The aspect has universal relevance. Given a metric for reliability that is consistently defined for all TSOs, sampled objectively, and for which the result of that is publicly available, its relation with cost could be tested for in a cost driver analysis. As to gas TSOs, reliability is a different story. There it is largely driven by low tolerance levels, often stemming from safety regulations. In general interruptions in gas transmission are much less frequent and last much shorter than with electricity. For gas, we also have the issue that interruption figures are much less distinctive, thus making statistical testing for a relation with cost difficult, even in the presence of perfect data.

6.10 Regarding other suggestions made, they seem less suitable to see these as quality aspects. Some suggestions, like facilitating hub liquidity, balancing, connection task, gas qualities transported, pressure levels at entry and exit points, or odorisation, are more about *what* a TSO does, not *how well* it is done. Other suggestions lack sufficient universal relevance, lack an obvious and practical metric, or seem of little relevance.

6.3 Conclusions

6.11 To conclude, CEER remains open to defining and implementing quality aspects, but sees on the basis of the responses to the surveys and available material currently no way to do this properly.

7. Summary and discussion

7.1 Main findings

7.01 The TCB18 project has established a comprehensive platform for cost efficiency assessments in gas transmission through a set of detailed data specifications for assets, activities, costs and environmental conditions. The specifications have been reviewed in several rounds by NRAs, TSOs and external experts to be as relevant and clear as possible. A new efficient organization of the data collection and validation has been implemented, managed by the PSG, more precise by the NRAs for cost and asset data and managed by consultants for the collection of environmental parameters mapped to the service areas of the operators. This process is forming a stable and powerful basis for periodic performance assessments and the systematic collection of data to gauge the development of the sector.

7.02 The collected data have been processed in order to derive a benchmarking model capturing the three main service dimensions (grid provision, capacity provision and customer service) considering heterogenous economic and environmental conditions and technical specifications. Using the normalized grid metric, the multiple assets of the power system have been included to form a Totex-relevant proxy for grid size, more predictive to cost than using conventional measures such as pipeline length or gas volumes transported. Using statistical methods to derive the most informative models, a final model with four outputs and one input, Totex, has been developed.

7.03 The cost efficiency results from the model present a mean cost efficiency for 2017 corresponding to 79% of relevant Totex. This result indicates an efficiency improvement potential that is on average about 21%. The potential appears to be a very conservative estimate of the true prospect for performance improvements in the sector, here excluding all effects of capacity utilization and energy consumption that could be added to the picture. However, the results do indicate examples of best practice to be analyzed and emulated, as well as providing information to regulators and operators about the sources of inefficient investments and operations.

7.2 Plausibility of the results

7.04 One way of looking at the results is to ask oneself if it is reasonable to believe that individual scores can come out as low as 80%, 70%, 60%, or even lower. The answer to that question is in our view YES for two important reasons. First of all, the TCB18 project itself has been performed with great care, i.e. extensively validating data, often making cautious assumptions when modelling, and verifying the results to the extent that the PSG cannot think of any reason why these could not be trusted. Often these steps were inspired by comments from TSOs, leading to the formulation and testing of additional hypothesis to rule out errors as much as possible.

7.05 Another interesting point of view is founded on the outcome of other benchmarking studies focusing on infrastructure sectors. Notably in gas and electricity many studies exist with similar outcomes as for TCB18. But also looking at a typical project in rail infrastructure efficiency made for the European Commission (Steer Davis Gleaves, 2015), one can see a considerable spread in raw cost efficiency, not explained by size, and in the DEA scores (that are particularly "soft" using a 2-input, 3-output model). Indeed, there are large differences in the way heavy infrastructure is planned, procured and operated - even if the operators use tendering and are incentivized (nationally).

7.06 We can even take this argument further, by looking at a non-infrastructure sector, like banking. To measure their efficiency banks commonly use the cost-to-income ratio. Seen as a unit cost efficiency measure, which is reasonable given that often banks focus on cost reductions to improve their ratio, we see banks worldwide having very low efficiencies. Even on a European Union scale, we see numbers as low as 50% in 2014, see https://m.theglobaleconomy.com/rankings/bank_cost_to_income .

7.07 Having observed this, it is important to realize that individually there can be many good reasons for very low or very high efficiency scores and that it is not the purpose of TCB18 to judge about that. With TCB18 a best practice frontier has been developed in a pan-European context, based on verifiable observations while maintaining a neutral position towards national circumstances.

7.3 Comparison with E2GAS

7.08 The current model has a number of similarities with the E2GAS (2016) model: the incorporation of environmental conditions in the analysis, the parameter for connection points to capture network density, and consideration of the capacity dimension. However, the model offers several enhancements for the assessment:

1) Exogenous reporting of environmental conditions by service area. Whereas the environmental factors were reported asset-by-asset by the TSOs, the data collection in TCB18 is based on a fully verifiable and repeatable data extraction from public data bases with high and fully equivalent precision. This approach guarantees the equity among operators and the due process in the validation across regulators.
2) Econometric validation of all environmental factors. In E2GAS, the complexity factors were implemented using a multiplicative form across all dimensions, whereas the TCB18 approach has inserted an additional model analysis step where the correlation between factors has been examined, leading to the inclusion of a reduced number of factors in the model specification.
3) Separable inclusion of soil humidity. Early analysis of potentially missing dimensions indicated that the soil humidity has a specific and significant effect on pipeline operations cost that is not directly proportional to the capex-construction cost. This led to the inclusion of a new output parameter that links the physical network length (not the NormGrid value) to the humidity factor.
4) Compressor power as capacity output. A preliminary model was presented at W4 using total injection flow as a capacity measure, motivated by statistical cost causality. In discussion with project participants and engineers, this output variable was replaced with the total compressor power to establish a more stable and network relevant indicator of capacity provision. Comparing with the peak flow measure from E2GAS, this output variable is independent of business cycle and weather conditions.

7.4 Limitations

7.09 Although state-of-the-art statistical techniques have been applied to determine the optimal combination of environmental factors for the final model, some conditions might apply to an individual or small group of operators passing undetected in the model specification. In the case the combined effects are significant, the systematic two-stage outlier detection in DEA would identify and remove the data. However, there might be cases of impact without being sufficient for outlier classification that merit the attention of the NRA in interpreting the results from the study and their potential use in informing regulatory decisions.

7.5 Future plans for benchmarking

7.10 Regulatory benchmarking has reached a certain maturity through this process and model development, signaling both procedural and numerical robustness. Drawing on the work, the definitions and data standards as well as the model, CEER can readily plan for a repeated regular benchmarking at a considerably lower cost in time and resources, to the benefit of all involved. Although the current model brings improvements in particular in environmental factors, the inflation and salary corrections and the NormGrid definitions, the relative symmetry with the earlier model from E2GAS can be seen as a confirmation of the type of parameters and approaches chosen, leading to stable and predictable results. In this manner, the future work can be directed towards further refinement of the activity scope and the interpretation of the results, rather than on the model development.

8. References

Agrell, P. J. and P. Bogetoft (2007) *Development of benchmarking models for German electricity and gas distribution*, Final report 2007-01-01, Bundesnetzagentur (BNetzA), SUMICSID AB.

https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/Netzentgelte/Gas/Effizienzvergleich%20Verteilernetzbetreiber/GA_Sumicsid_Effizienzvergleich.pdf?__blob=publicationFile&v=3

Agrell, P. J. and P. Bogetoft (2009b) *International Benchmarking of Electricity Transmission System Operators*. Project e^{3}GRID. SUMICSID AB. http://e3grid.sumicsid.com/pub/2009-03-09_e3grid_final_report_open_main.pdf

Agrell, P. J. and P. Bogetoft (2014), International Benchmarking of Electricity Transmission System Operators *Proceedings of European Energy Market Conference EEM14*, IEEE Proceedings, pp. 1-5 doi: 10.1109/EEM.2014.6861311

Agrell, P. J., P. Bogetoft, H. Beaussant, J. Talarmin (2014) *Benchmarking European Gas Transmission System Operators: A Feasibility Study*. Project PE2GAS. SUMICSID. Report for CEER. https://www.researchgate.net/publication/272094391_Benchmarking_European_Gas_Transmission_System_Operators_A_Feasibility_Study

Agrell, P. J., P. Bogetoft, C. Riechmann, A. Rodgarkia-Dara, C. Zimmer (2013) *e^{3}GRID 2012 – European TSO Benchmarking Study*. Project e^{3}GRID 2012. SUMICSID, Consentec and Frontier Economics. Report for CEER. https://www.researchgate.net/publication/265592571_E3GRID2012_European_TSO_Benchmarking_Study

Agrell, P.J., P. Bogetoft, and J. Tind (2005), DEA and Dynamic Yardstick Competition in Scandinavian Electricity Distribution, *Journal of Productivity Analysis*, 23, 173–201.

Agrell, P. J., P. Bogetoft, U. Trinker (2016) *Benchmarking European Gas Transmission System Operators*. Project E2GAS. SUMICSID and Swiss Economics. Report for CEER. https://www.ceer.eu/documents/104400/-/-/f919650e-0f9b-fb96-ebf8-47cac308667e

Agrell, P. J., and Niknazar, P. (2014). Structural and behavioral robustness in applied best-practice regulation. *Socio-Economic Planning Sciences*, 48(1), 89-103.

Aivazian V. A., J. L. Callen, M. W. L. Chan and D. C. Mountain (1987), Economies of Scale Versus Technological Change in the Natural Gas Transmission Industry, *The Review of Economics and Statistics*, 69(3), pp. 556-561.

Banker, R.D. (1996), Hypothesis Test Using Data Envelopment Analysis, *Journal of Productivity Analysis*, 7, pp. 139-159.

Bogetoft, P. (1997), DEA-Based Yardstick Competition: The Optimality of Best Practice Regulation, *Annals of Operations Research*, 73, pp. 277-298, 1997.

Banker, R.D. (1996), Hypothesis Test Using Data Envelopment Analysis, *Journal of Productivity Analysis*, 7, pp. 139-159.

Callen J. L. (1978) Production, Efficiency, and Welfare in the Natural Gas Transmission Industry, *The American Economic Review*, 68, pp. 311-323.

Carrington R., T. Coelli and E. Groom (2002) International Benchmarking for Monopoly Price Regulation: The Case of Australian Gas Distribution, *Journal of Regulation Economics*, 21:2, pp. 191-216.

Gordon, D. V., K. Gunsch and C. V. Pawluk (2003) A natural monopoly in natural gas transmission, *Energy Economics*, 25, pp. 473-485.

Granderson G. and C. Linvill (1996) The Impact of Regulation on Productivity Growth: An Application to the Transmission Sector of the Interstate Natural Gas Industry, *Journal of Regulation Economics*, 10, pp. 291-306.

Granderson, G. (2000) Regulation, Open-Access Transportation, and Productive Efficiency*. Review of Industrial Organization*, 16, pp. 251-266.

Lawrence, D., Kain, J., & Coelli, T. (2011). *Regulation of Suppliers of Gas Pipeline Services–Gas Sector Productivity*. Report for Commerce Commission, Economic Insights. http://www.economicinsights.com.au/reports/Economic_Insights_NZCC_Gas_TFP_Report_10Feb2011.pdf

Jamasb, T., Newbery, D., Pollitt, M., & Triebs, T. (2007) *International benchmarking and regulation of European gas transmission utilities*. Report prepared for the Council of European Energy Regulators (CEER). http://portal.e-control.at/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Gas/2006/ERGEG_Cost_benchmark_final_tcm7-107140.pdf

Jamasb, T., Pollitt, M., & Triebs, T. (2008) Productivity and efficiency of US gas transmission companies: A European regulatory perspective. *Energy Policy*, 36(9), 3398-3412.

MacAvoy, P. W. and R. Noll (1973) Relative Prices on Regulated Transactions of the Natural Gas Pipelines, *The Bell Journal of Economics and Management Science*, 4,(1), pp. 212-234.

Meyrick (2004) *Comparative Benchmarking of Gas Networks in Australia and New Zealand*, Report prepared for Commerce Commission, Wellington

Nieswand, M., Cullmann, A., & Neumann, A. (2010). Overcoming data limitations in Nonparametric Benchmarking: Applying PCA-DEA to natural gas transmission. *Review of Network Economics*, 9(2).

Sickles, R. C. and M. L. Streitwieser (1992) Technical Inefficiency and Productive Decline in the U.S. Interstate Natural Gas Pipeline Industry under the Natural Gas Policy Act, *Journal of Productivity Analysis*, 3, pp. 119-133.

Sickles R. C. and M. L. Streitwieser (1998) An Analysis of Technology, Productivity, and Regulatory Distortion in the Interstate Natural Gas Transmission Industry: 1977-1985, *Journal of Applied Econometrics*, 13, pp. 377-395.

Steer Davis Gleave (2015) Study on the Cost and Contribution of the Rail Sector. *Final Report prepared for the European Commission*. https://ec.europa.eu/transport/sites/transport/files/modes/rail/studies/doc/2015-09-study-on-the-cost-and-contribution-of-the-rail-sector.pdf

Appendix

A. Gas asset reporting guide 2018-03-08

B. Financial reporting guide 2018-03-08

C. Special conditions reporting guide, 2018-09-13

D. Method to treat upgrading, refurbishing and rehabilitation of assets in TCB18

E. Modelling opening balances and missing initial investments, 2018-01-11

F. Norm Grid Development Technical Report 2019-02-27 V1.3

SUMICSID

SUMICSID SPRL (GROUP)
Rue Maurice Lietart 56
B-1150 Brussels, BELGIUM
belgium @ sumicsid.com

Tel: +32 (0)2 675 15 23
Fax: +32 (0)2 706 53 58

http://www.sumicsid.com

CEER
Council of European Energy Regulators

Cours Saint-Michel 30a, box F (5th floor)
1040 Brussels, BELGIUM

Tel: +32 (0)2 788 73 30
Fax: +32 (0)2 788 73 50

www.ceer.eu