

Geometric versus Geographic Models for the Estimation of an FTTH Deployment

Attila Mitcsenkov¹, Mirosław Kantor², Koen Casier³, Bart Lannoo³, Krzysztof Wajda², Jiajia Chen⁴, Lena Wosinska⁴

¹ *Budapest University of Technology and Economics, Dept. of Telecomm. and Media Informatics, High Speed Networks Laboratory (HSNLab), Magyar tudósok krt. 2, 1117, Budapest, Hungary*

E-mail: mitcsenkov@tmit.bme.hu

² *AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland*

E-mail: {kantor|wajda}@kt.agh.edu.pl

³ *Dept. of Information Technology, Ghent University - IBBT, Gaston Crommenlaan 8, box 201, B-9050 Gent, Belgium*

E-mail: {koen.casier| bart.lannoo}@intec.ugent.be

⁴ *Royal Institute of Technology KTH/ICT, Isafjordsgatan 22, Electrum 229, 16440 Kista, Sweden*

E-mail: {jiajiac|wosinska}@kth.se

Abstract Optical access networks provide a future proof platform for a wide range of services, and today, several operators are deploying fibre to the home (FTTH) networks. Installing an FTTH infrastructure, however, involves very high investment cost. Therefore, a good estimation of the investment cost is important for building a successful business strategy and, consequently, to speed up the FTTH penetration. In this paper, for calculating the amount of cable and fibre in the outside plant together with the associated civil works, and the number of required network elements, two different approaches are investigated: (1) geometric modelling of the fibre plant based on approximate mathematical models and (2) geographic modelling of the fibre plant based on map-based geospatial data. The results obtained from these two approaches can then be used as input for preliminary investment cost calculations and/or techno-economic evaluations. Compared to more complex and accurate geographic modelling, we verify that especially with uneven population density and irregular street system, simple geometric models do not provide accurate results. However, if no geospatial data is available or a fast calculation is desired for a first estimation, geometric models definitely have their relevance. Based on the case studies presented in this paper, we propose some important guidelines to improve the accuracy of the geometric models by eliminating their main distortion factors.

Keywords *Fibre to the Home (FTTH), Cost estimation, Techno-economic, Network models, Topology*

1 Introduction

Fibre to the home (FTTH) has been widely recognized as a future-proof solution for access networks due to its capability to meet the increasing bandwidth demand of the end users. Therefore, FTTH deployment is currently experiencing a fast growth all over the world. For the 2020 time horizon, the European requirement on the residential peak data rate is not less than 1Gbit/s [1]. In contrast to many existing broadband technologies, such as digital subscriber line (DSL) and wireless access, fibre access can easily fulfil such a huge bandwidth requirement on a per customer basis. Several FTTH network architectures have been developed over the years, e.g., point-to-point (P2P), active optical network (AON) and passive optical network (PON) [4]. The methodology presented in this paper is technology agnostic and can be applied to any type of fibre access networks. In the results section, however, we focus on PON as an example.

The deployment of FTTH requires a large investment in fibre infrastructure while the cost of access networks is shared by a relatively low number of users. It is therefore essential to identify the most cost-efficient way to provide FTTH. For such techno-economic assessments e.g. in [6, 7], a good estimation of the investment cost is crucial. Moreover, it is shown that the deployment cost of the fibre infrastructure is the dominating component of the capital expenditures (CapEx) [8]. Therefore, inaccuracy occurred in the deployment cost estimation of the fiber infrastructure may lead to the wrong conclusion of the techno-economic assessment.

To rapidly estimate the amount of the outside plant infrastructure, several geometric network models have been developed, e.g. [9,10], which are based on a set of statistic parameters describing the considered area, such as average population density and distance from the end users to the central office (CO). These models were widely used in the past decades because they are easy to use, offer fast calculations and don't require detailed information about the deployment area.

On the other hand, area-wide average parameters hide the impact of an uneven user population or irregular building sizes, which may have a negative effect on the overall cost estimation. To address this issue, a geographic approach based on the detailed geospatial data was proposed, that aims to design the FTTH outside plant infrastructure prior to deployment cost estimations [12]. This approach

offers much higher accuracy than geometric models does, in particular in uneven populated areas, but on the expense of higher complexity and longer computation time..

It is obvious that the higher level of detail in the input data set leads to more accurate results. However, the strategic network design process for the cost estimation is not trivial. Recent improvements in geospatial data technology and computer performance together with efficient heuristic algorithms support the automatic topology design, which was not available earlier.

While work in [10-13] addresses geometric and geographic modelling, this paper focuses on a comparison of them trying to clarify where the higher complexity of geographic, network design based modelling pays off, and also investigates in which way one can increase accuracy of the geometric models. Therefore, in this paper we extend our preliminary work presented in [15] and carry out several case studies covering both urban and suburban cases. It is shown that there are significant differences between the results obtained by the geographic approach and the geometric models. In particular, it is confirmed that the geometric models applied to a realistic – typically irregular – service area, may lead to large over- or underestimations. In this work we quantify the benefits of the geographic approach that can efficiently improve estimation of the fibre infrastructure cost in an unevenly populated area. Moreover, some general guidelines are given for using the geometric models in a more reliable way, since they can still be beneficial for a preliminary estimation.

The remainder of this paper is organized as follows. Section 2 describes the existing calculation models. More precisely, it presents a general framework for techno-economic evaluations indicating the role and importance of the network design phase and it summarizes the geometric and geographical models for dimensioning the outside plant. Then, Section 3 selects a set of representative case studies with respect to various types of populated areas. They will be used to point out the differences between the geometric and geographical models. Section 4 applies both models to the different cases assuming a PON technology. The results and an extensive evaluation of both models is presented, and their applicability under different circumstances is discussed. Finally, Section 5 gives the conclusions and summarizes some potential approaches to improve the accuracy of geometric models.

2 Calculations models

In this section, a techno-economic framework is presented to highlight the importance of the fibre plant dimensioning. Furthermore, two geometric models and the geographic model for dimensioning the fibre layout are described.

Techno-economic framework

A general techno-economic framework defines the context for the addressed network dimensioning. Fig. 1 gives an overview of such a general framework that can be used for performing a complete techno-economic evaluation, not solely for FTTx networks [13-14]. Note that a detailed comparison of the economic viability of different access network scenarios is crucial for operators due to the high cost of this network segment.

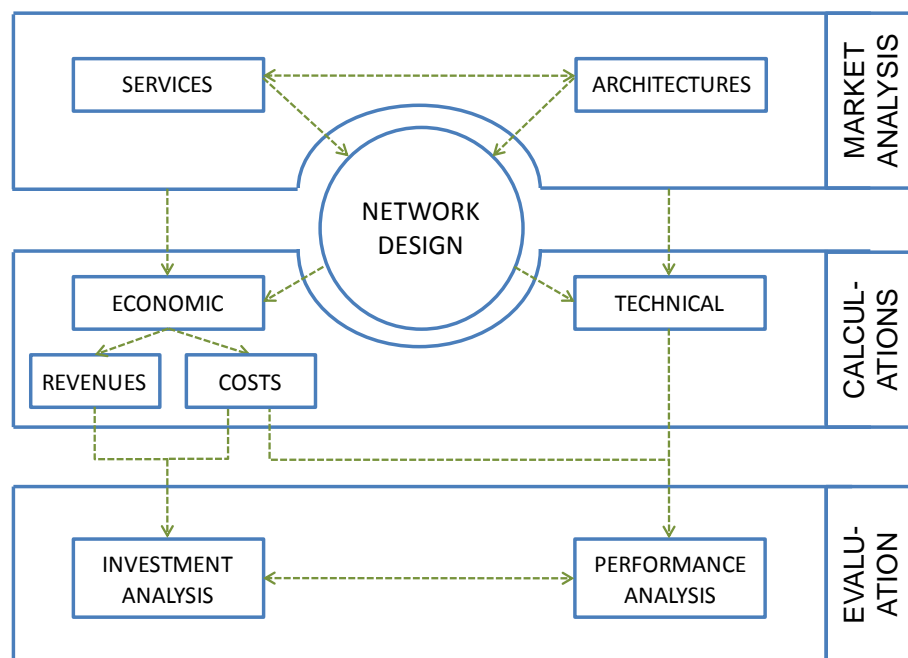


Fig. 1 Framework for a techno-economic evaluation [13]

A techno-economic evaluation starts from determining the scope of the problem and detailing the inputs for the study based on a market analysis. The most important outcomes here are indicated by the building blocks “services” and “architectures”. They contain all input information necessary for building the techno-economic model in the second step. Often in a telecom project the network

is the central piece and contains most optimization opportunities. Moreover, a proper network design also reflects suitability of a certain network infrastructure to the considered scenario, and therefore it supports the optimal choice among the competing technologies. As such, the network design is given a central position in Fig. 1 as the link between the market analysis and the calculations block.

In the calculations block, we make a distinction between economic calculations, in which we estimate costs and revenues; and technical calculations, in which we estimate the performance metrics of the proposed network solution (e.g. reliability, peak and sustainable bandwidth, reach, energy, etc.). In the final step, an evaluation will be based on the outcomes – economic and technical – of the calculations step. This step is split between investment analysis and performance analysis. In the first part, we make an estimation of the (expected) profitability of the project. In the second part, we compare different alternatives and make trade-offs of costs vs. performance. Both results are the final outcome of a well-balanced techno-economic study.

The work of this paper is focussing on the network design block, which plays a central role in the techno-economic framework. The physical network infrastructure has a fundamental impact on deployment costs and future economic value; therefore it needs a reliable and good estimation. Since no “best” network architecture exists in general, finding the most suitable solution for a given service area and service set under given economic conditions plays a crucial role regarding profitability of the network operator.

As already mentioned in Section 1, the physical access network could be represented by analytical or geometric models [9] using area-wide average parameters of the selected area (e.g. population density, diameter); or the results of a more realistic, geographic network design process could be used directly, supposed that the required geospatial information is available [12]. Both approaches are discussed in the following sections.

Geometric Models

To estimate the amount of fibre, ducts and civil works (trenching) required for an access network design, geometric models can be used. They assume uniformly spread customers in a given area, represented as a polygon, and build a repetitive structure for which the analytical formulations exist. Two generally accepted

geometric models, namely Triangle Model (TM) and Simplified Street Length Model (SSL) are described in this section, and compared later to the geographic model presented in this paper.

In the following subsections both models are described. Both models are described in a comparable way by using similar parameters for the number of buildings and customers. The main parameters in these two models are:

- N_B : the number of buildings or houses in the polygon
- N_H : the number of households or customers in the polygon
- d : building density in the polygon

Triangle Model

The Triangle Model (TM) is a polygon-based model. The model initially proposed in [9], was also used in the RACE 2087/TITAN [16] project, in the ACTS 226 OPTIMUM project as a part of the techno-economic methodology [17], and later (improved) in TONIC [18] and TERA [19] projects. The triangle model has been also applied in some techno-economic analysis [20].

Fig. 2 illustrates the model showing the polygonal structure surrounding the hub and representing the distribution area. The dispatching boxes FP1 (boxes B, C, D and E) and distribution cabinets FP2 (points F) are symmetrically located at the gravity centres of the elementary triangles. In the model a uniform subscriber density over the whole distribution zone is assumed.

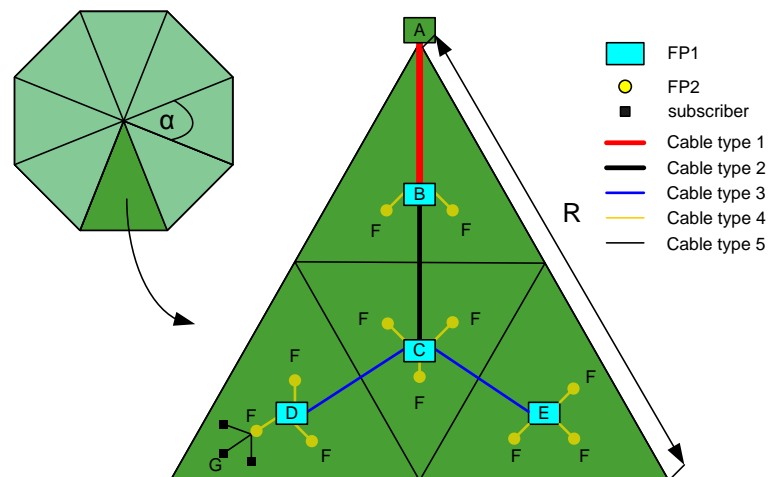


Fig. 2 Triangle Model (TM) [16]

The radius of the polygon is given by the following formula:

$$R^2 = \frac{2 \cdot N_B}{d \cdot f \cdot \sin(2 \cdot \pi/f)} \quad (1)$$

where f denotes the number of fibres connected to the hub (this is also the rank of the polygon).

The length of the feeder trenching considered in the TM model is calculated using the following formulas:

$$|\overline{AB}| = |\overline{BC}| = (R/3)\cos(\pi/f) \quad (2)$$

$$|\overline{CD}| = |\overline{CE}| = (R/6)\sqrt{1 + 8\sin^2(\pi/f)} \quad (3)$$

$$|\overline{DF}| = R(0.132 + 0.336/f) \quad (4)$$

Equations 1, 2 and 3 are directly derived from Fig. 2 whereas equation 4 is obtained by simulation.

The total feeder trenching length can be calculated by summing up these distances for all triangles:

$$L_{FT} = f(|\overline{AB}| + |\overline{BC}| + |\overline{CD}| + |\overline{CE}|) + N_{sp}|\overline{DF}| \quad (5)$$

where N_{sp} is the number of splitters. The total feeder fibre length can be calculated by using the following formula:

$$L_{FF} = \frac{N_{sp}}{4f} (4|\overline{AB}| + 3|\overline{BC}| + |\overline{CD}| + |\overline{CE}|) + N_{sp}|\overline{DF}| \quad (6)$$

The average distance b between the branching box F and building entrance G gives the formula:

$$b = \frac{2}{3} \sqrt{\frac{N_B}{12 \cdot f \cdot \pi \cdot d}} \quad (7)$$

where we assume 12 branching boxes per triangle (see Fig. 2).

The total length of the distribution trenching can be calculated according to the following formula:

$$L_{DT} = N_B b \quad (8)$$

The length of the distribution fibre can be obtained from the following formula:

$$L_{DF} = N_H b \quad (9)$$

Simplified Street Length Model

The second geometric model considered in this paper, is called the simplified street length (SSL) model [22]. In this model, the potential customer base is uniformly distributed over a squared area (Fig. 3, left). One side of the square contains n houses or buildings ($\sqrt{N_B}$) and the distance between two houses is

indicated by l . The building density d is then $1/l^2$. The CO is always situated in the middle of the square. The model assumes that all houses can be connected in one line through the middle of the house (Fig. 3, right). All streets are connected using one divider street.

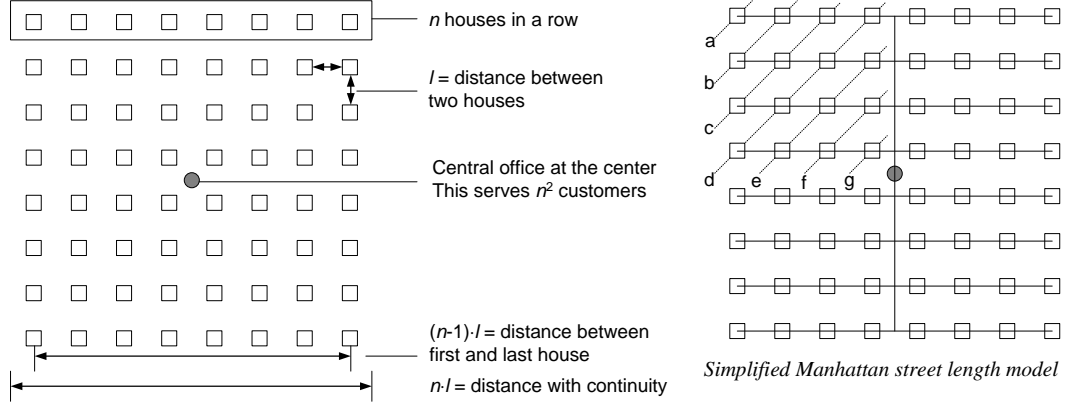


Fig. 3 Illustration of the SSL model for calculating trenching and fibre lengths in the distribution part

The SSL model is first described for a one-stage structure, where the feeder and distribution part are not split. We present this as a feeder fibre to each house. Regarding the (feeder) trenching length L_{FT} , each row of houses requires a trench with a length of $(n-1) \cdot l$, and there are n rows (resulting in a length of $n \cdot (n-1) \cdot l$). The divider street requires a trenching length of $(n-1) \cdot l$. Combined this gives a total (feeder) trenching length as given in (10).

$$L_{FT} = n \cdot (n - 1) \cdot l + (n - 1) \cdot l = (n^2 - 1) \cdot l \quad (10)$$

For the fibre calculation, we start from the number of households N_H , and we define k as the number of households per building (N_H/N_B) as a separate fibre per household is required. Regarding the (feeder) fibre length L_{FF} , the structure, as seen from the CO is fully symmetric and there are four equal quadrants. When we focus on one quadrant, we find for the houses in the categories indicated on Fig. 3 (right) the following lengths: $a=(n-1) \cdot l$, $b=(n-2) \cdot l$, ... $g=l$. At the beginning, the number of houses per category is increasing with one each step. Once the diagonal line crosses at the half of the quadrant the number of houses per category is decreasing with one each step. The (feeder) fibre length as such is given in (11).

$$L_{FF} = 4 \cdot k \cdot l \cdot \sum_{i=1}^{n-1} [\min(i, n - i) \cdot (n - i)] \quad (11)$$

The general SSL model can be extended to a two-stage structure consisting of a feeder and distribution part with one aggregation point in between, indicated as remote node (RN). For each network part (feeder, distribution), we can calculate the total fibre length between the endpoints (i.e. the number of fibres multiplexed in the RN) and the trenching length. With formula (11), we can calculate the total fibre length in each network part, if we adapt the formula for L_{FF} so that RNs are connected instead of households. The total trenching length, however, has to be calculated in another way, since there is not always a separate trench between two endpoints. If we consider a green-field situation, some trenches are used for both of the feeder and distribution network segments. To take this into account we first calculate the trenching length of the feeder part L_{FT} , to connect all RNs, based on formula (10). Finally, we calculate the trenching length to pass all the houses (without taking into account the RNs) and subtract the feeder part to define the (additional) trenching length of the distribution part L_{DT} .

Geographic Model

Using the already designed network topology to estimate the deployment costs is a completely different approach from geometric modelling. Difficulties of acquiring the necessary GIS data and the complexity and time of calculations are the price for the higher accuracy. The resulting network layout and data support a complete bill-of-material calculation for the cable plant and network equipments (e.g. splitters). Such a methodology is adaptive not only to the area-wide average descriptors, but also to the local characteristics of the service area, e.g. the uneven population density or irregularities of the street system, or e.g. a river that cuts the area. These are not considered in the geometric models, since these are using only the area-wide average parameters, and not their local characteristics.

The higher level of detail in the input data obviously supports a more accurate cost estimation. However, appropriate efficient heuristics, and the necessary technical background, e.g. the GIS databases were not available until the recent years. The existence and availability of geospatial information raise several practical difficulties in themselves. Moreover, access network design is a highly

complex algorithmic problem. In addition, realistic scenarios are large-scale, typically the graphs contain tens of thousands of nodes. These scalability issues, coupled with NP-hardness make it an extremely difficult algorithmic challenge that requires sophisticated decomposition techniques and highly specialized and efficient heuristics, as described in [24].

An overview of the methodology is given here, the reader is referred to [25] for further details of the AccessPlan Framework, which is an experimental project, for evaluating the concept. The main phases of the calculations are depicted on Fig. 4, where the layers show the different process steps.

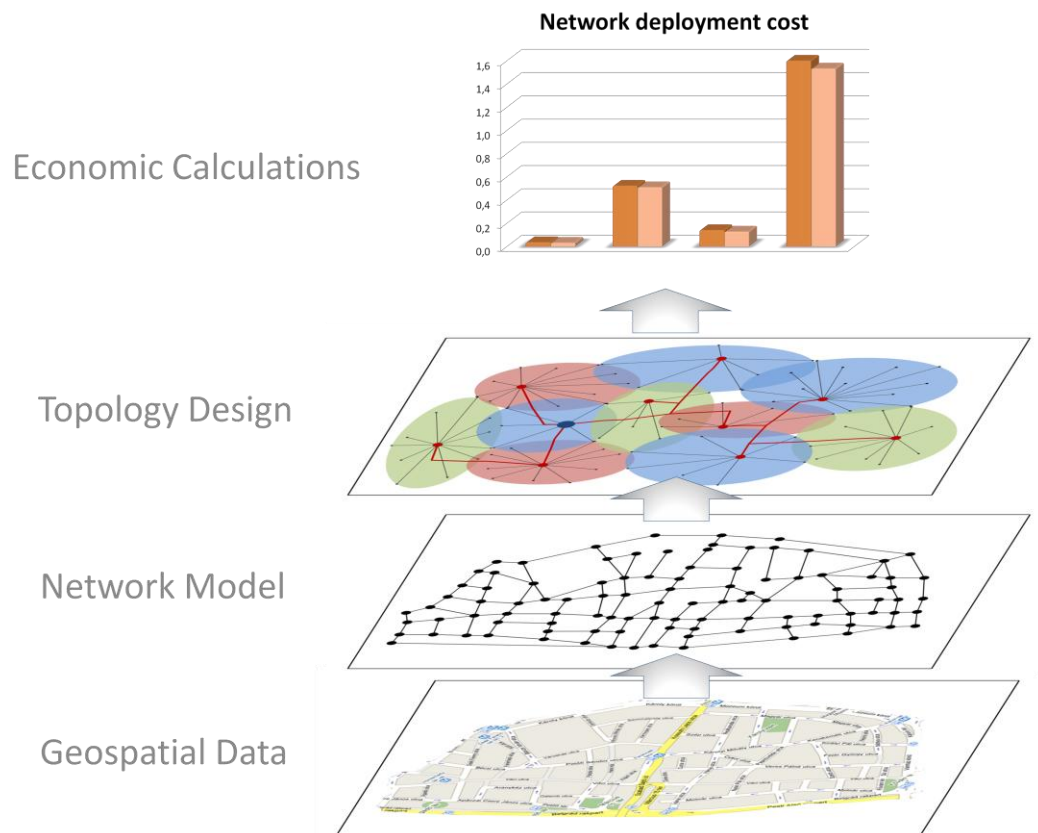


Fig. 4 AccessPlan Framework Workflow

As the first step, the required geospatial data is gathered and processed. The digital map on the bottom level of the figure includes the subscriber (household) data and the infrastructural information as well.

The next, intermediate task is the modelling phase, where these inputs are transformed to a network graph model. Combined with the cost values and physical constraints, the mathematical problem is formulated.

In the optimization phase, a set of heuristic algorithms is working on the graph model. These are high performance algorithms, each of them specialized for a

particular FTTH technology. The algorithmic background exceeds the scope of this paper, and the reader is referred to [24] for further details on this aspect. Finally we get the optimized network topology, which is then transformed from the graph model to the map again. The topology is analyzed carefully, in order to calculate the deployment cost. At this point comprehensive network information is available, which serves as input for the techno-economic analysis and comparison.

3 Case Studies

Four representative case studies have been chosen, in order to evaluate the difference between the results from the geographic and geometric modelling, and their accuracy under various circumstances.

The first two scenarios are sparsely populated: two smaller towns, one of them in the countryside, the other one in the agglomeration of a capital city. Due to geographic conditions and the distance of neighbouring settlements, both of them are individual service areas of a single Central Office (CO). Further two densely populated areas were chosen within a capital city, one suburban and one typical urban area. An overview of the selected scenarios can be found in Table 1.

Table 1 Scenario characteristics

DATA	Countryside	Agglomeration	Suburban	Urban
Area (km ²)	7,21	4,42	5,89	6,56
# Buildings	3067	2134	2796	3662
Building density (1/km ²)	430	480	470	560
# Households	3067	2714	14836	26015
Household density (1/km ²)	430	610	2520	3970
Avg. household / building	1,0	1,3	5,3	7,1
Less populated 10% area vs. average population	14%	9%	11%	6%
Most populated 10% area vs. average population	196%	175%	219%	226%
Variance in population density	133%	125%	89%	68%

The measures used for characterizing the scenarios are the standard measures of the area, and the population (household) density – these are also required for the geometric models. Additionally, the distinction between households and buildings is necessary due to the addressed FTTH scenario, as the building/household proportion gives the amount of fibres entering the building.

“Countryside” scenario

The first selected scenario is a smaller settlement in the agglomeration of a capital city. We typically find a few main roads, connecting the town to its neighbours. Along these main roads, we have a slightly higher population density; further away typically family houses are built, resulting in approximately 1.0 household per building.

The settlement structure is relatively simple; the map overview and a figure showing the population density variance are found below.

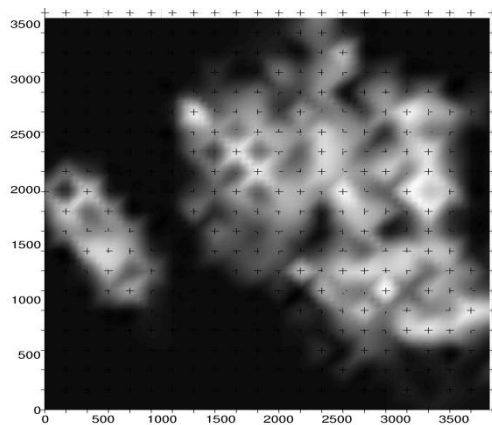


Fig. 5 Population density (Countryside scenario)

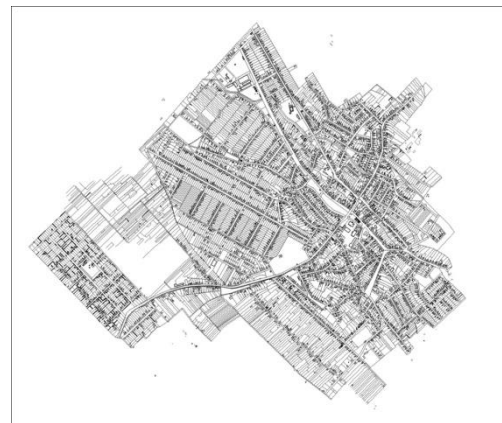


Fig. 6 Map overview (Countryside scenario)

“Agglomeration” scenario

The second selected scenario is still a smaller settlement in the agglomeration of a capital city. In this case we are facing some mid-size buildings, not only family houses. The structure is slightly different, mostly due to the typical rapid growth in the agglomeration during the last decade. Even though the average population density is similar to the countryside scenario, its distribution across the service area is different. Moreover, some parts are still in the construction phase (south-

west and northern areas in this case), which is also typical for an agglomeration settlement.

The map overview and the population density figures are given below.

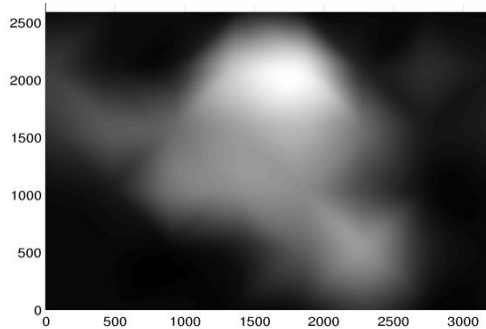


Fig. 7 Population density
(Agglomeration scenario)

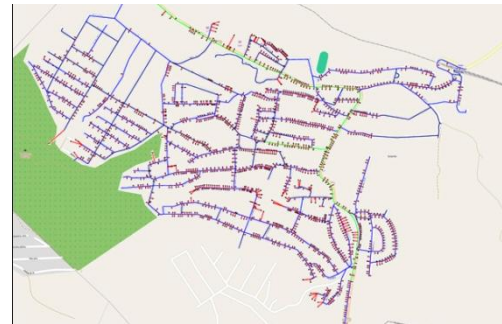


Fig. 8 Map overview
(Agglomeration scenario) Source: OSM [26]

“Suburban” scenario

We have also chosen an area inside a capital city, but not in its mostly populated part. It features a really widely distributed mix of building types and residential zones, along with a small nature reserve area (park), and a cemetery – such suburban regions typically have a few special parts like these.

Apart from these spots, the population is really varying: less than 1% of the population is located in 10% of the area, while on the other hand, in the most populated 10% of the area we have more than 20-25% of the population.

At the same time, the majority of the area has a more or less even distribution, as its relatively small variance value shows (Table 1), however two highly populated spots may be observed in the south-west and south-east parts (see Fig. 9).

Population distribution and maps are given below.

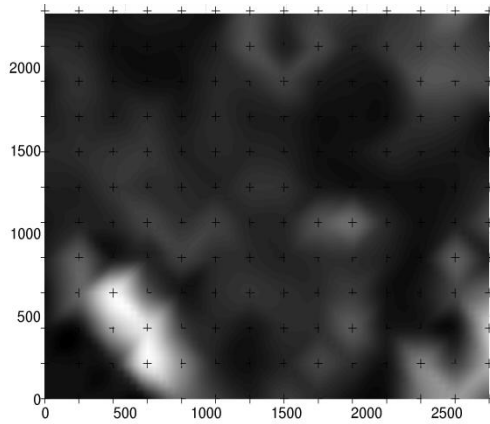


Fig. 9 Population density
(Suburban scenario)

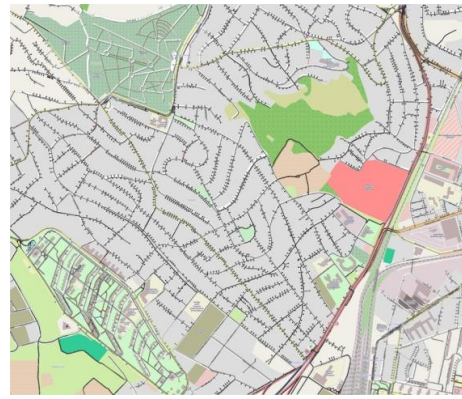


Fig. 10 Map overview
(Suburban scenario) Source: OSM [26]

“Urban” scenario

Finally a typical, densely populated urban scenario was chosen, with the highest and most evenly distributed population density among the presented scenarios. The highest household per building ratio belongs to this region, indicating an increased representation of large apartment houses.

As mentioned before, it has a very low variance and also the population density distribution (Fig. 11) shows relatively low local differences, even though a few small, high density spots may be observed, mainly around some high rise buildings. The latter leads to a new phenomenon, the “gravity effect” of those buildings in the splitter allocation phase: the closer the splitters are to those buildings, the more fibre saving is observed. The population density figure and a map overview are given below.

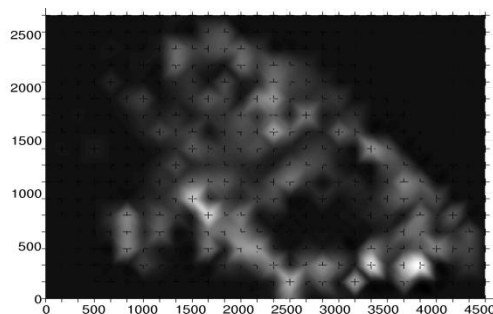


Fig. 11 Population density
(Urban scenario)



Fig. 12 Map overview
(Urban scenario) Source: OSM [26]

4 Results and evaluation

The presented geometric and geographic modelling approaches were applied on realistic scenarios, as introduced in Section 3. Based on the resulting splitter count, trenching and fibre lengths, a comparison of the geometric models to the geographic (benchmark) is made.

A typical network setup was chosen for comparison: a GPON access network, with 1:64 splitters, in a single level architecture (without primary and secondary splitting points) for sake of simplicity. A second setup, with 1:16 splitting ratio was also evaluated. However as its results were completely in line with the setup of 1:64 splitters they have been omitted from the paper.

Since these network modelling approaches are used as input values for the techno-economic calculations, we based the comparison on the raw outcome values in terms of trenching length, fibre length and splitter count. The estimated cost values were not used, in order to eliminate any distortion effect of the cost models (e.g. economies of scale and scope).

Numerical results

Results for the 1:64 setup are presented on Fig. 13 to Fig. 16. The calculated values are normalized in a way that in every case the particular result from the geographic model is used as the unit value for visualization.

Results of the geographic model are used as reference values because this methodology directly operates on the geospatial representation of the service area, provides valid access network topologies, and provides the most reliable and accurate base for trenching and fibre length and for the splitter count.

The absolute values are indicated right under the X-axis. Differences between the Triangle Model (TM) or Simplified Street Length (SSL) model and the geographic calculations are indicated as follows:

- small differences (less than 25%) - *(green) italic letters*
- moderate difference (between 25-50%) - **(orange) bold letters**
- significant difference (more than 50%) - **(red) bold, underline letters**

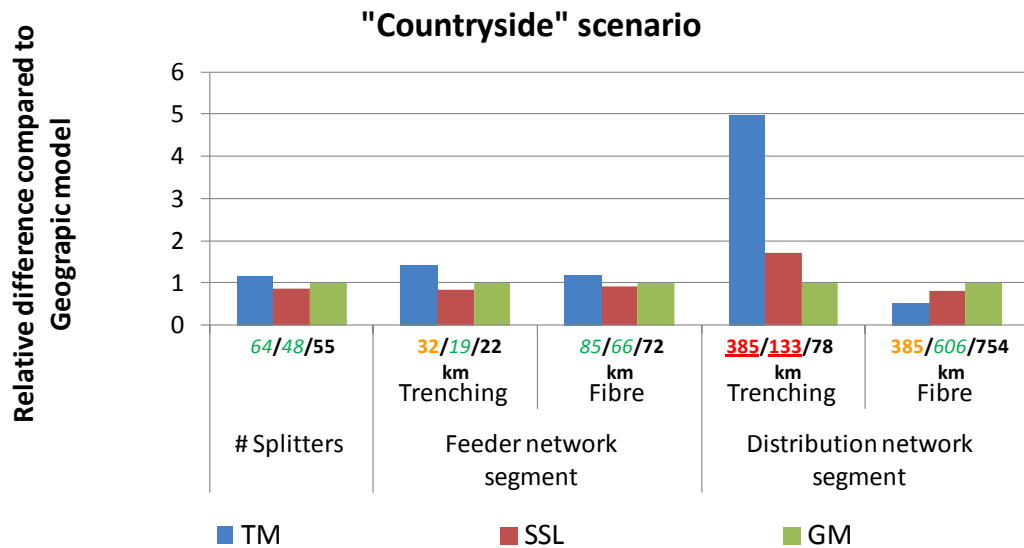


Fig. 13 Comparison of results for the **Countryside** scenario

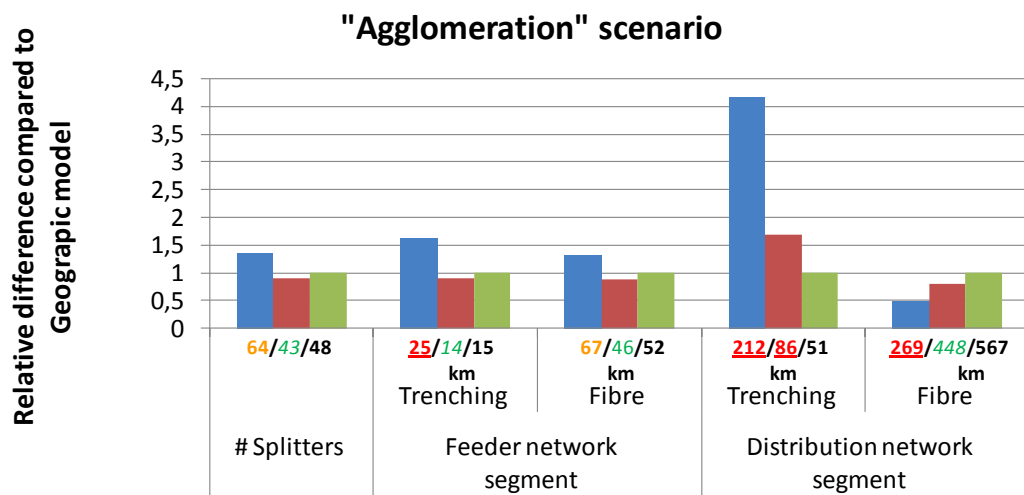


Fig. 14 Comparison of results for the **Agglomeration** scenario

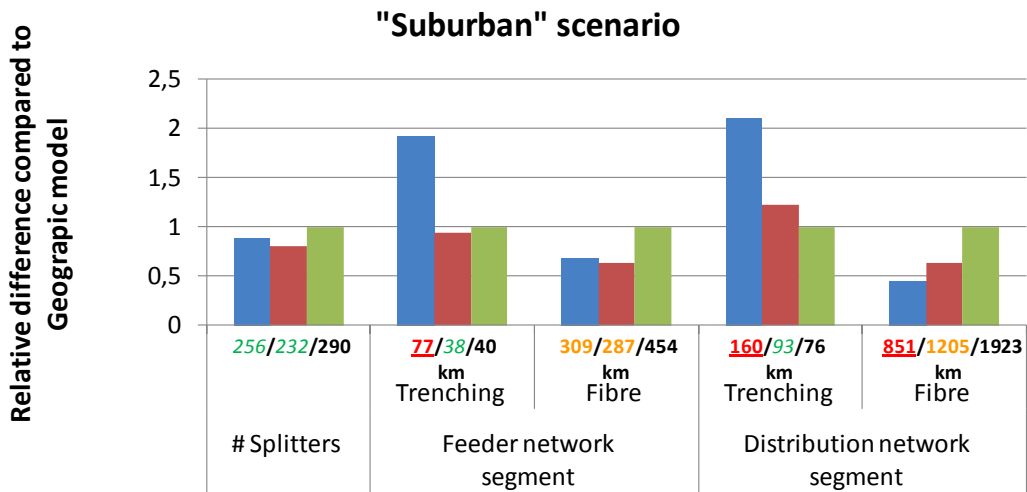


Fig. 15 Comparison of results for the **Suburban** scenario

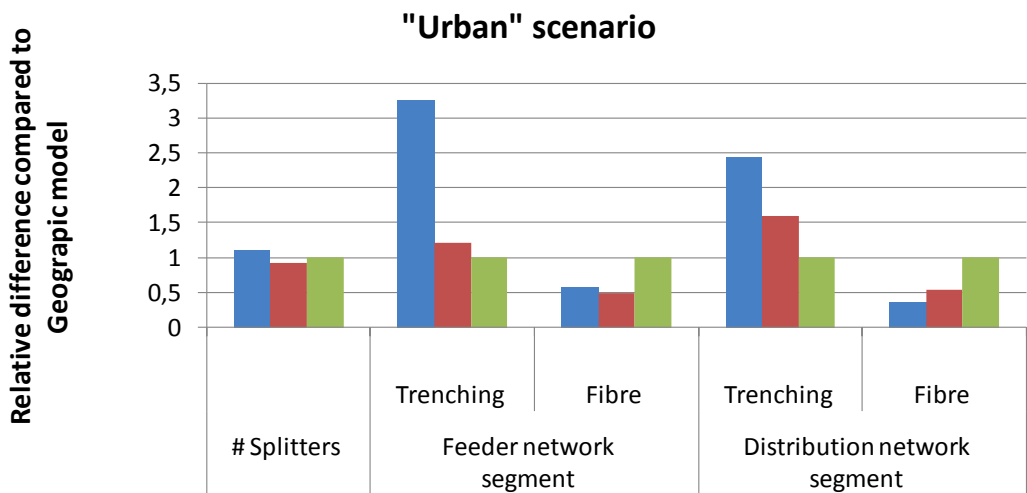


Fig. 16 Comparison of results for the **Urban** scenario

Model evaluation

Three different models were evaluated, i.e. two different geometric models, the *Triangle Model* (TM) and the *Simplified Street Length* (SSL) model were compared to the *Geographic Model* (GM). During the evaluation, five measures were considered: the amount of splitters, trenching and fibre usage in the feeder and distribution network segments, respectively.

The evaluation section is organized along the models (TM vs. GM, then SSL vs. GM). At the end of the section, a summary table is given, in order to help comparing the three models based on five measures.

Triangle Model vs. Geographic Model

We call back the schematic figure of the Triangle Model (Fig. 2). By its nature it has an obvious flaw: it was designed for estimating the cable plant in the absence of geospatial information, e.g. the street system. Therefore the connections between every pair of nodes are estimated with straight lines, instead of following the street system, even though a significant difference may be observed between the map-based distance, and the straight line distance: the latter was typically around 70-80% of the actual distance in the presented scenarios.

The street system has another, more important effect on the trenching: a large portion of the traces is shared among connections between neighbouring buildings and their respective splitter. The separate straight lines used to connect buildings individually to their splitters in the Triangle Model ignore this effect.

These two factors together are responsible for the overestimation of **trenching** by the TM: in the feeder network, 140-300% of the reference values (GM) was observed, mainly due to the straight line vs. street system distance anomalies, but also the supposed individual trenching adds to the overestimation, especially in the densely populated urban areas, where a large amount of splitters exist.

Overestimation due to individual trenches is more pronounced in the distribution network, particularly in sparsely populated areas, where longer parallel connections are present, i.e. more sharing of trenches is possible. The latter manifests in 4-5 times overestimated distribution trenching in Countryside and Agglomeration scenarios, and a 2-3 times overestimation was observed even in the urban regions.

Fibre usage in the feeder part is acceptable, the \pm 20-30% difference, especially in the feeder segment does not distort the results significantly. In the sparse areas, the amount of splitters is slightly overestimated, which leads to increased feeder fibre calculations (+30%), while the underestimation in the suburban and urban areas is due to the position of the CO.

In the Triangle Model the CO is right in the middle, while in the presented scenarios, the CO has a less central position, which leads to a higher feeder fibre need.

The distribution part has a higher contribution to the overall fibre usage, and in this segment a consequent underestimation is present, i.e. roughly 40-50% of the GM results. The main reason behind this phenomenon is again the difference between straight line and street measured distance, but it is not alone responsible for the anomalies. The presence of "irregular clusters" (households assigned to the same splitter) increases the difference even more: in a real topology we often find a few buildings falling further from the splitter than the average, increasing the fibre need. However, the circular clusters around splitters in a geometric model are not affected by these irregularities. These two effects altogether cause the 30-50% underestimation of fibre usage.

The amount of **splitters**, due to careful use of the Triangle Model, relatively closely approximates our geospatial calculations; however this result was not trivial to achieve: the amount of polygons, triangles, "sub-triangles" and splitting points had to be chosen in a way that maintains geometrical regularity of the triangle partitioning.

Simplified Street Length (SSL) model vs. Geographic Model

The SSL model is an improved geometric model, in a sense that it considers the street system. We recall Fig. 3, where a regular grid structure is outlined, which represents the estimated network topology.

Thus estimation of **trenching** is accurate, particularly in the feeder segment, where the lack of street system was the main flaw of the TM: the estimated feeder trenching is within 89-127% of the geospatial calculations. The distribution trenching is slightly overestimated (150-160%), mostly due to depth-first-search (DFS) like operation of the topology designer algorithm in GM, which effectively

minimizes trenching via clusters of closely connected houses, which is not considered in the uniformly distributed grid model.

Fibre usage in the feeder segment is well estimated in the sparsely populated areas, but 50-60% underestimation is present in the suburban and urban scenarios. Location of the Central Office (CO) explains these trends: in the latter scenarios, the CO is not located directly in the middle of the area, and it increases CO-splitter distances, i.e. the fibre usage. We note that this underestimation may be eliminated if we place the CO not in the middle of the grid; however, in this case the calculations become more complex. The moderate underestimation of splitters (see below) also adds to the underestimation of feeder fibre: less splitters obviously require less fibre to connect.

The distribution segment, due to its larger fibre need, has higher importance. At the evaluation of the Triangle Model, two distortion factors were mentioned: the straight line connectivity and the irregularity of clusters. The former has almost disappeared with the street system of SSL, which leads to better results: the underestimation is less pronounced in this case. The remainder (60-80% instead of 40-50%) is mainly due to the already mentioned irregularity of clusters.

The **splitters** in the SSL model are placed for each 64 customers. Here we used the value of the average customers per building to select the size of the distribution area of buildings connected to one splitter (e.g. an average of 1.6 customers per building leads to a distribution area size of 40 buildings). As the customers are evenly spread, the SSL model can have a 100% fill rate for the splitter. This is however prone to an overestimated fill rate and could as such be further tuned to fit more closely the real fill rate.

Table 2 Summary of comparison results

Model		TM	SSL
# Splitters		89-130% (OK)	↓ 80-93% Reason: <ul style="list-style-type: none"> • Complete fulfilment of splitters
Feeder network segment	Trenching	↑ 1:64 - 140-300% (↑ 1:16: 200-1000%) Reason: <ul style="list-style-type: none"> • Straight line distance • Individual trenching 	89-127% (OK)
	Fibre	↑ Sparse 120-130% (OK) ↓ Dense 60-70% Reason: <ul style="list-style-type: none"> • Sparse: more splitters • Dense: non-central CO location 	↑ Sparse 110-120% (OK) ↓ Dense 48-63% Reason: <ul style="list-style-type: none"> • Sparse: more splitters • Dense: non-central CO location
Distribution network segment	Trenching	↑ Sparse: 400-500% ↑ Dense: 200-300% Reason: <ul style="list-style-type: none"> • Individual trenching • Straight line distance 	↑ 150-160% ↑ Suburban: 120% (OK) Reason: <ul style="list-style-type: none"> • Closely connected clusters (DFS)
	Fibre	↓ 1:64 - 37-51% Reason: <ul style="list-style-type: none"> • Straight line distance • Cluster irregularity 	↓ 1:64 - 54-80% Reason: <ul style="list-style-type: none"> • Cluster irregularity

Eliminating distortion effects

Following the careful study of the results, we found the main reasons behind the anomalies and inaccuracies of the geometric models and the obvious need for improvement of TM or SSL models. However, the question is if one can define guidelines for increasing accuracy to eliminate the over/underestimation inherent for the geometric models. The answer is not obvious. The distortions will never be completely eliminated, since the geometric models by nature do not consider local variance of geospatial data. However, in some cases these flaws may be handled by applying the TM or SSL models in the right way.

In the following, the distortion effects unveiled in the previous section are addressed and avoidance techniques are discussed.

Individual trenching

This is the major flaw of the Triangle Model, by its nature, it cannot incorporate shared trenching. The extent of distortion strongly depends on very local details of the service area, e.g. in the presented case studies, the overestimation varies between 200-500%, which unfortunately makes the TM inappropriate for estimation of trenching, especially in the distribution network segments.

Splitter fulfilment

The splitter usage, i.e. exactly 64 households for each splitter, leads to an underestimation of splitters in the SSL model. A straightforward compensation is provided by slightly scaling down the splitter usage, to approx. 85-90%, which works in most of the scenarios, and may be defined as a rule of thumb that improves quality in all reasonable scenarios.

However, it does not lead to the complete elimination of the phenomenon: during the topology design phase the optimal usage is determined for each splitter individually, based on its neighbourhood. A cost minimization decision is made, considering whether the increase in fibre usage when adding a few more (remote) households exceeds the savings in splitters, therefore the optimal fill ratio is part of the optimization and depends on the service area itself.

Straight line distance

Fibre usage was consequently underestimated in the TM, primarily due to the distance measure: instead of following the street system, the Euclidean distance of nodes was used.

Using the digital map, it is always possible to calculate the average difference between these distance measures for every pair of nodes, and then scale up the estimated fibre usage by this factor. Moreover, even in the absence of a digital map, this scaling factor may be estimated, as our calculations have shown, the Euclidean distance was typically around 70-75% of the actual street measured distance, therefore the estimated fibre usage may be multiplied by $1/75\% \sim 1,33$.

Irregular clusters

This is another factor that is really difficult, if not impossible to compensate: clustering of households around splitters strongly depends on really localized variance of the building sizes and street system.

In our point of view, the measured 54-80% underestimation is not representative in a sense that we cannot propose a confident scaling factor based on it for an arbitrary scenario, but the direction and magnitude of distortion is visible.

Non-central CO location

This is the difference between models and reality that is the easiest to compensate, at least theoretically. Although the central office is located in the middle of a symmetric service area in the geometric models, the effect of placing it out from its central position is clearly understandable.

The first option is to modify calculations, and put the CO somewhere else in the grid or the triangle structure. Breaking symmetry slightly complicates calculations, but summing up the small sections of triangles, or edges of squares this way is possible.

Another option follows the idea of Steiner's theorem about distance from the median: only the median-CO distance is what alters the results, therefore an analytical scaling could eliminate the inaccuracy.

5 Conclusions

The network layout and dimensioning (trenching, fibre need and splitter usage) has a large impact on the final investment costs of an FTTH network. In this paper, two modelling approaches for defining an FTTH network layout, i.e. geometric and geographic models, are presented and evaluated.

The geometric models provide a rapid approximation based on publicly available area-wide average parameters of the service area (diameter, area size, number of buildings and households). In contrast, in the geographic modelling a complete strategic topology design process is carried out, with respect to fibre layout and network equipments. It requires a geospatial representation of the service area, i.e. a digital map as well as the infrastructure and household data.

The trade-off between the complexity and data requirements of the accurate geographic modelling versus the simplicity of the geometric models is analyzed in this paper. Various case studies are presented in order to identify strengths and weaknesses of the considered approaches.

Using the geometric models gives an almost instant estimation of the costs for the outside plant and therefore can be useful for repetitive studies such as geomarketing, full economic optimization, technology set comparison, etc.

However, using geometric models only gives a rough estimate for the outside plant while the geographic model can offer an optimal solution if accuracy of the input data is high.

In the evaluation part of the paper we have not only analyzed inaccuracy brought by the geometric models, but also investigated the reasons behind the observed either over- or underestimations. As an important addition, we have provided guidelines to increase the accuracy of the two considered geometric models, avoiding the unveiled distortion effects.

In summary, as expected, the geographic calculations offer higher accuracy. It is shown that an uneven population density or an irregular street system further increases the gap between the accuracy of the geometric models and the more complex geographic calculations. Therefore geographic calculations are recommended when reliability of the cost estimation is the primary requirement, even at the expense of a more complex data collection and calculation process.

On the other hand, in the absence of geospatial data, or in cases when rapid calculations are necessary, the geometric models may provide valuable

information for a preliminary cost estimation. It is shown that the Simplified Street Length model with the proposed improvement guidelines offers an acceptable (20-30%) approximation for trenching, which is considered as the most important cost factor.

6 Acknowledgement

The authors would like to articulate special thanks to Péter Katzenberger (HSNLab) for the population density variance calculations and figures, created for his MSc Thesis, within the frame of the AccessPlan framework.

This work was partly supported by the Network of Excellence “Building the Future Optical Network in Europe” (BONE), funded by the European Commission through the 7th ICT-Framework Programme.

Part of the research presented in this paper was carried out in the frame of the NGAdesigner project, in cooperation with NETvisor Ltd, Hungary, issued by the Hungarian National Development Agency, supported by the EU.

7 References

- [1] 249025-ICT OASE Project, Requirements for European Next-Generation Optical Access Networks, D2.1.
- [2] IEEE 802.3ah Task Force; <http://www.ieee802.org/3/efm>
- [3] ITU-T G.984.x Series of Recommendations; <http://www.itu.int/rec/T-REC-G/e>
- [4] M. Ngo, A. Gravey, D. Bhadauria, “Controlling QoS in EPON-based FTTX access networks”, *Telecommunication Systems J.*, 2011, 48/1, p. 203-217
- [5] K. Grobe and J. P. Elbers, "PON in Adolescence: From TDMA to WDM-PON", *IEEE Communications Magazine*, vol. 46, pp. 26-34, Jan. 2008.
- [6] B.T. Olsen, D. Katsianis, D. Varoutas, K. Stordahl, J. Harno, N.K. Elnegaard, I. Welling, F. Loizillon, T. Monath, and P. Cadro, “Techno-economic Evaluation of the Major Telecommunication Investment Options for European Players”, *IEEE Network*, vol.20, pp. 6-15, Jul./Aug. 2006.
- [7] S. Verbrugge, D. Colle, M. Pickavet, P. Demeester, S. Pasqualini, A. Iselt, A. Kirstädter, R. Hülsermann, F.-J. Westphal, and M. Jäger, “Methodology and Input Availability Parameters for Calculating OpEx and CapEx Costs for Realistic Network Scenarios”, *OSA Journal of Optical Networking*, vol. 5, pp. 509-520, Jun. 2006.
- [8] J. Chen, C. M. Machuca, L. Wosinska, and M. Jäger, “Cost vs. Reliability Performance Study of Fiber Access Network Architectures”, *IEEE Communications Magazine*, vol. 48, pp.56-65, Feb. 2010.

- [9] D. Gardan, et al, "Techno-economics of advanced optical subscriber networks", in *Proc. IEEE GLOBECOM'89*, Dallas, USA, 27-30 November 1989.
- [10] Koen Casier, "Techno-Economic Evaluation of a Next Generation Access Network Deployment in a Competitive Setting", PhD degree at the Faculty of Engineering of the Ghent University obtained on October 8th 2009
- [11] Koen Casier et al. "Techno-economic evaluations of FTTH roll-out scenarios", in *Proc. NOC2008*, pp. 113-120, Krems, Austria, July 1-3, 2008
- [12] A. Mitcsenkov, G. Paksy, and T. Cinkler, "Topology Design and Capex Estimation for Passive Optical Networks", in *Proc. BroadNets2009*, Madrid, Spain, September 14-16, 2009.
- [13] M. Kantor, K. Wajda, B. Lannoo, K. Casier, S. Verbrugge, M. Pickavet, L. Wosinska, J. Chen, A. Mitcsenkov, "General framework for techno-economic analysis of next generation access networks", in *Proc. ICTON 2010*, Munich, Germany, 27 Jun - 1 Jul. 2010.
- [14] H. Jarmo et al, "Alternatives for mobile operators in the competitive 3G and beyond business", *Telecommunication Systems J.*, 2009, 41/2, p. 77-95.
- [15] A. Mitcsenkov, M. Kantor, K. Casier, B. Lannoo, K. Wajda, J. Chen and L. Wosinska, "Geographic Model for Cost Estimation of FTTH Deployment: Overcoming Inaccuracy in Uneven-populated Areas", in *Proc. IEEE ACP 2010*, IEEE, Shanghai, China, 2010
- [16] B. Olsen et al., "RACE 2087/TITAN: Tool for introduction scenarios and techno-economic studies for the Access Network", in *Proc. RACE Open Workshop on Broadband Access*, Nijmegen, the Netherlands, 7–8 June 1993.
- [17] Olsen, B.T. "OPTIMUM Methodology", In *proc. OPTIMUM workshop on Multimedia Techno-economics*, Aveiro, Portugal, 20-22 October, 1997.
- [18] T ONIC project, EU FP5 - <http://www-nrc.nokia.com/tonic/>
- [19] T ERA project, EU FP6-SSP-2005-006469 - <http://www2.dse.unibo.it/tera/>
- [20] Ims, L. A. (editor), "Broadband Access Networks - Introduction strategies and techno-economic analysis", Chapman-Hall, January 1998.
- [21] B. Olsen, "OPTIMUM – techno-economic tool", *Teletronikk*, vol. (2/3), 1999, pp. 239–250.
- [22] B. Lannoo, G. Das, M. De Groote, D. Colle, M. Pickavet, P. Demeester, "Techno-economic feasibility study of different WDM/TDM PON architectures", In *Proc. of ICTON 2010*, Munich, Germany, 27 Jun. – 1. Jul. 2010.
- [23] G. Fatma, E. Erhan, "Telecommunications network design with multiple technologies", *Telecommunication Systems J.*, 2011, 46/2, p. 149-161.
- [24] A. Mitcsenkov, G. Paksy, T. Cinkler, "Geography and Infrastructure Aware Topology Design Methodology for Broadband Access Networks (FTTx)", *J. Photonic Network Communications*, 2011, 21/3 (p. 253-266).
- [25] AccessPlan Framework – <http://safranek.tmit.bme.hu/accessplan/>
- [26] OpenStreetMap – <http://www.openstreetmap.org/>