VOL.39 Nº4 2022
https://doi.org/1022119\IMECTA.9278.en

Methodology proposal for temporal evaluation of leveling network stations using Graph Theory

Jessica Caroline dos Santos Siqueira[a], Leonardo Castro de Oliveira[b], Ivanildo Barbosa[c]

Cartographic Engineering - IME
[a]jcssiqueira.jj@gmail.com
[b]leonardo@ime.eb.br
[c]ivanildo@ime.eb.br

RESUMO: A remedição dos desníveis das seções por meio da técnica de nivelamento geométrico é uma atividade comum em redes de nivelamento e que gera várias observações de uma mesma seção. Esse conjunto de observações precisa ter sua consistência checada a fim de minimizar a inserção de possíveis outliers aos dados. Neste trabalho é proposta, para avaliação temporal dos dados de nivelamento, uma metodologia que: otimiza o método usual de análise; explora alternativas quando esse não é passível de aplicação, com o aproveitamento das informações disponíveis; adicionalmente, visa tornar a avaliação mais homogênea e não subjetiva. Para atingir tal objetivo, a rede foi tratada como um grafo, assim, foram empregados conceitos e algoritmos oriundos da Teoria dos Grafos, como todos os ciclos e caminhos. Esse paradigma possibilitou a geração de um conjunto de insumos para análise dos dados. A aplicação da metodologia na Rede Altimétrica Brasileira – RAAP – demonstrou resultados consistentes, desde que a rede possua redundância suficiente que permita a detecção e a confirmação de possíveis abalos.

PALAVRAS-CHAVE: Consistência das observações. Nivelamento. Outliers. Movimentação temporal. Ciclos e caminhos em grafos.

ABSTRACT: The remeasurement of height differences of the sections through the geometric leveling technique is an everyday activity in leveling networks. It generates many observations of the same section. This set of observations needs to be checked for consistency to minimize possible outliers' insertion into the data. This work proposes, for the temporal evaluation of the leveling data, a methodology that: optimizes the usual method of analysis: explores alternatives when it cannot be applied taking advantage of available information; additionally it aims to make the assessment more homogeneous and non-subjective. To achieve this goal, the network was treated as a graph. Thus, concepts and algorithms from Graph Theory were used, like all cycles and paths. This paradigma made it possible to generate a set of inputs for data analysis. The application of the methodology in the Brazilian Altimetric Network – RAAP – showed consistent results, as long as the network has enough redundancy to allow detection and confirmation of possible temporal movements.

KEYWORDS: Consistency of observations. Leveling. Outliers. Temporal movements. Cycles and paths in Graphs.

1. Introduction

A leveling network, or an Altimetric Reference Network (RRA), consists of a set of Level References (RRNN) materialized on the ground, where measurements of differences in level and distance between them are performed. The networks can be local or national and both are intended to provide altitude values with adequate accuracy for their purposes in a given region. National networks have the advantage of providing a single and homogeneous reference for the most diverse activities that require an adequate vertical positioning, such as: dam monitoring processes, irrigation and drainage projects, infrastructure works such as road construction and sanitation, among others.

It is common for RRA, especially national ones, to be built in stages and, occasionally, densifications are carried out, due to the inherent slowness of the measurement method by geometric leveling. With this process, the need arises to connect new levels to the structures of the existing networks, in addition to the existence of cases in which new lines pass close to the old ones, resulting in the RRNN of the old lines found visually in good condition being leveled. In this work, these sections remeasured at different times will be called check section or check edges.

https://doi.org/1022119\IMECTA.9278.en

These sections require a specific evaluation of such data – this process being called temporal evaluation.

Temporal evaluation is a step that contributes to the robust analysis of altimetric networks, as any RN that has been moved can generate outlier(s) in that region and even compromise the quality of the network. In addition, the process allows validating whether there was a consistent connection between leveling performed at different times. It is important to emphasize that this phase of analysis of an RRA precedes the least squares adjustment and the application of statistical and robust methods to identify outliers. The evaluations are carried out directly on the primary data and there is no influence of any adjustment.

In this work, he proposes a methodology for identifying and evaluating these check sections based on Graph Theory concepts. It is expected to automatically generate consistent inputs that enable the identification of possible temporal shocks in the leveling networks, thus contributing to the process of refinement and processing of these networks.

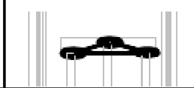
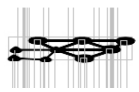
2. Graph Theory And Altimetric Networks

For the temporal evaluation of check sections using Graph Theory, it is necessary to understand the leveling network as a graph, that is, to associate the elements of a leveling network with the elements inherent in Graph Theory. This association is essential, as it allows for proper correspondence between the two areas – in this case, Geodesy and the Graph – and to employ concepts and algorithms from Graph Theory in modeling the proposal for evaluating the check sections.

Graph Theory is a branch of discrete mathematics that emerged at the beginning of the 18th century and has been used in the search for the solution of several practical problems. A graph is characterized by its non-empty set of vertices, a set of edges and a function (Ψ) that associates each edge to an unordered pair of vertices belonging to the graph, these vertices being distinct or not [1].

Table 1 presents the main elements of the altimetric network used in the modeling of this proposal and their corresponding correspondence in the universe of Graph Theory. Additional information about the elements of an altimetric network can be obtained in [2], as well as on graphs [1] and [3] are recommended.

Tab. 1 - Correspondence Between Elements Of A Leveling Network And Graph Theory

Altimetric Network	Theory of graphs	Illustration
RN or nodal RN	Vertex	
Internodal Section or Line	Edge	
Check section	Parallel Edge	
Circuit	Cycle	
Uni-Nodal Circuit (points like triangles)	Self-loops	

Source: Authors.

3. Temporal Evaluation And Check Edges

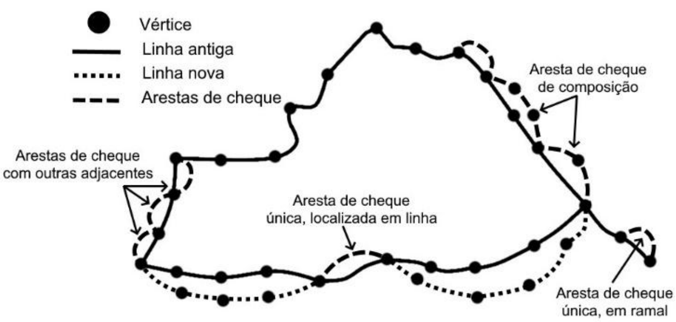
The temporal evaluation is the procedure of verifying the consistency of the unevenness measured at different times, in order to guarantee that there was no variation in the altimetric coordinate. This variation may be due to a vertical or even horizontal movement of the level reference above a tolerance established for the network under analysis. There are 4 basic situations in which a section or check edge appears in a network. The first situation would be the case in which the check edge has other adjacent check edges – in this situation the temporal evaluation is performed directly in order to identify whether a shock occurred in the RRNN involved. This is the evaluation usually carried out, recommended and described in the

Specifications and Standards for Geodetic Surveys associated with the Brazilian Geodetic System (SGB) [2]. The second case is when the adjacent check edges are a composition of gaps and distances measured between RRNN in the region, caused in most cases by the destruction of stations between one leveling or another. In this work, the maximum composition adopted was of the order of 10 km of leveled distance. It was decided to limit this distance in the composition of the sections due to the propagation of errors, as more extensive compositions can influence the temporal evaluation.

The third case considers that there was a leveling of just a single session in the network and this section is on a line. The fourth case is when the only leveled section is in a branch. The second and the third cases are situations that do not fit the standard recommendation for temporal evaluation, but which are being considered in this work, as they also have the potential to provide consistent information for identifying temporal disturbances in stations. The use of Graph Theory in these two cases makes this process more objective and optimizes the evaluation time, since it automatically identifies and generates the analyses.

The sections of the fourth case have restricted evaluation, as they are located in branch lines. Extensions are elements that do not meet the concept of a network and do not have redundancy. In **figure** 1 illustrates the 4 situations that check edges can be found in a network.

Fig. 1 – The four basic situations that a check edge appears in a leveling network.



Source: Authors.

4. Methodology

In **figure** 2 presents a diagram containing the steps of the methodology implemented for the evaluation of the check sections of a network. The dashed rectangles highlight the phases of the methodology in which the Graph Theory concepts and algorithms are employed. The following concepts were used: vertex degree, adjacent vertices, weighted graphs, connected graphs, sub graphs, paths and cycles. All these are presented in detail by [1] and [3]. Some algorithms used were those for identifying shortest cycles, as discussed by [4], and algorithms for *depth-first search* (DFS) and shortest paths, explained by [5].

The characteristics of an RRA allow its modeling and storage as a connected and undirected graph, with the edges weighted by the leveled distance (in kilometers) of the section, with the RRNN being the vertices of the graph and the edges representing the existence of a measurement between two RRNN. The slope value of the section is the average between the flatness and counter-levelling. Leveling is the measurement of the difference in level between two RRNN in each direction and counter-leveling is the remeasurement of this difference in the opposite direction.

The script starts by reading the slope data and average distance of each section, as well as the date (month and year) of measurement. Then, it identifies the existing check edges and verifies if the differences in the slope value in the different measurements of the same section are within a tolerance value stipulated in the network planning. If they are, the average of the differences in level and distances in the different surveys is calculated and a text file is generated with all the calculated evaluations. If not, these different measurements of the same section are replaced by a single observation, where the difference in level is zeroed and the distance is assumed as the average between the surveys. The zeroed gap serves to mark in the gap file which sections will still be evaluated and the non-zeroed distance was a strategy to enable the internodal reduction step and, consequently, the identification of check edges obtained by the composition of other sections. Such conduct allows the network to be stored as a simple graph and from this to calculate the intermodal network, that is, the leveling part that meets the network requirements and allows evaluations to be carried out exploring its redundancy.

In this phase, files are generated with lists of extensions and identified self-loops. During the identification of the internodal network, the check edges generated from the composition of one or more sections and that have an accumulated distance less than or equal to 10 km are identified and separated.

The next step is to gather in a single file all the check edges found, those originating from the direct sections of the leveling and those obtained after the internodal composition, both those that are above and below the tolerance.

This new file of check edges is transformed into a new graph. This new file of check edges is transformed into a new graph. There are two groups, one with the components that have only 1 single check edge and another of the components with 2 or more check edges. Temporal analyzes and figures illustrating the graph of components with 2 or more check edges are generated, so that the analyst can have subsidy to make the most appropriate decision.

In the case of single check edges located on branches, these are printed in a separate list. In the case of single edges located in lines, the idea is to use the previously generated internodal network and calculate the minimum cycles in the region. With these cycles, extract those that have unique check edges and perform an evaluation with all paths and cycles from this set. This evaluation aims to generate additional inputs that allow the identification of possible shocks or even verify the consistency of the network with each observation independently.

In this phase, in which the cycles around the check edge being evaluated are used, it is possible to choose larger sets of cycles around it, as much as necessary. However, the more data used to generate all paths and all cycles, the greater the computational time required.

It is important to emphasize that in the literature an outlier is defined as an atypical observation in the sample and that, according to [6], it can be caused by any of the three common types of errors, such as gross, random and systematic, as well as by their composition. However, when an RN suffers a shock, it can generate an outlier in the sample, even the observations (the measured gaps) not having the presence of any of the three errors. This fact occurs because an RN, when shaken, no longer materializes the original position when it was implanted and leveled. Changing the original position generates a “new” level reference. In this context, evaluating this type of situation is essential to minimize sources of errors that often end up being ignored or unknown.

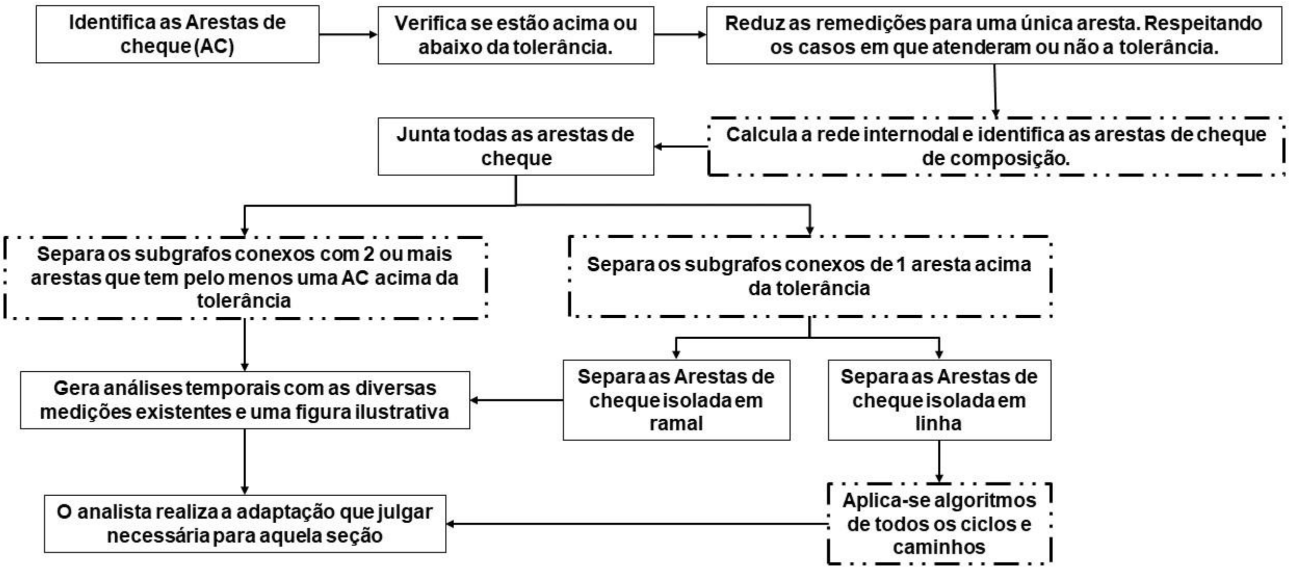
5. Script Development

The Python language was used to create the scripts. Two libraries written in this language were fundamental for the development of this proposal, Pandas [7] and NetworkX [8].

In the phase of generating all paths, the function already implemented in the NetworkX library, called *“all_simple_paths”*, was used. In the case of the script that generates all the cycles, it was created by the author and inspired both by the logic of the *“all_simple_paths”* function and by the algorithm presented by [9].

VOL.39 Nº4 2022
https://doi.org/1022119\IMECTA.9278.en

Fig. 2 – Diagram with the implemented phases of the methodology for temporal evaluation of level references (RRNN) using Graph Theory.



Source: Authors

6. Application Scenario

For the application of the proposed methodology, the High Precision Altimetric Network - RAAP, a basic altimetric reference for the country, which has been established and maintained since the 1940s by the Brazilian Institute of Geography and Statistics (IBGE), was chosen. Due to the continental dimensions of the country and its peculiarities, the RAAP has been densified over several years and this has generated the need for several remeasurements of sections. Therefore, it enabled the emergence of the 4 basic situations mentioned above, which makes the RAAP a favorable testing environment for the methodology.

7. Results And Discussion

The raw RAAP leveling file that was provided by IBGE for research purposes contains 71.384 sections (including remeasurements). After applying the methodology, 1.572 sections were identified that have two or more measurements. Of these, 1.522 are within the acceptable tolerance for the differences between the unevenness of each season. The tolerance adopted in the case of RAAPk was 4mm√k, being is the level distance of the section in kilometers. This value is in line with the survey accuracy adopted for most of the network, as described in [2], considering the age of the network and the times when most surveys took place.

Following the methodology, the 50 sections that had measurements that did not meet the tolerance were subdivided into three groups: 27 sections that have other adjacent edges (considering both the direct edges and those obtained by the composition of sections), making it possible to perform the standard temporal evaluation ; 13 sections without adjacent ones and which are located in branch lines; and 10 that are located in rows.

Figure 3 illustrates an excerpt from one of the text files generated with the methodology. More about the source textYou must provide the source text to see more information about the translationSend Feedback Side Panels It is important to point out that the temporal evaluation file generated considers the combination of analysis of all epochs for the same section, in this way the analyst will have as much information as possible for decision making. An example of this type of situation is also shown in **figure** 3.

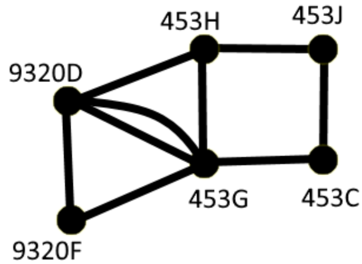
The evaluation of check sections when there are adjacent ones – cases 1 and 2 – must follow the usual procedure. The advantage of this methodology is that inputs are generated automatically with the maximum amount of information available in the analysis region, always seeking to make the process more objective and homogeneous. In the case of single sections located in branches, even if the usual procedure cannot be applied, their identification by scripts is important to register. This record can serve, for example, to indicate a reservation in the region in cases of studies of future densifications.

The case of single sections located in lines is what needs a more detailed evaluation, since, as these sections do not have other adjacent check sections that allow a usual evaluation, the adopted strategy is to exploit the redundancy of the network to generate the as many inputs as possible, in order to support decision-making. An interesting example that was detected after applying the methodology to identify the RAAP cross sections is illustrated in **figure** 4.

Section 9320D-453G has two measurements, one taken in 1960 and another in 1987. The difference between the flatnesses is -13.55mm, which leads to a closing error between the two measurements of -24.53 mm√k. As there are no otherremeasurements on adjacent edges that allow the scripts to generate data for a usual temporal evaluation, the proposed methodology for evaluating single check sections located in lines was followed.

The minimal circuits in the region were identified, separating those that were adjacent to the section under analysis (Fig. 4). All the paths that connected the 9320D-453G vertices were calculated and compared with the slope of the direct measurement of the 9320D-453G section in the two leveling epochs, one at a time. Subsequently, all the circuits of that set were generated. Depending on the magnitude of the temporal shock, the propagation of random errors in the sample may end up making the analysis unfeasible, so the evaluation was carried out considering the circuits closest to the verification section.

Fig. 4 – Illustration of the adjacent circuits of section 453G-9320D indicated by the methodology and used in the temporal evaluation. Each edge represents a measurement.



Source: Authors

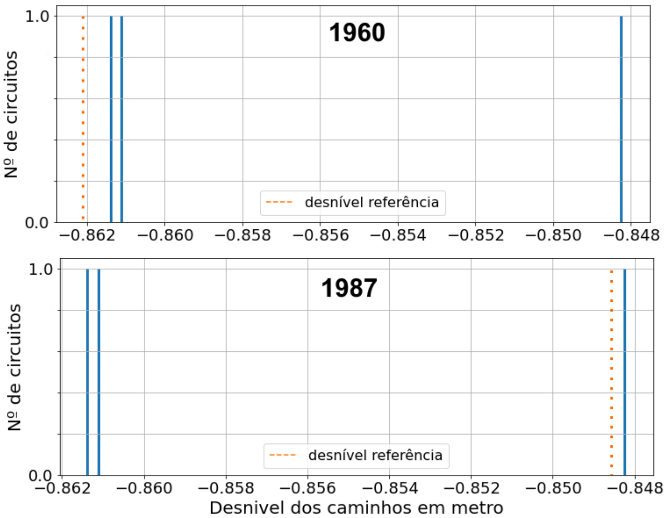
Fig. 3 - Illustrative excerpt from the temporal evaluation file when the check section has 3 or more remedies to enter the evaluation.

RNP	RNC	Desn	Dist	mes	ano_aaaa	ano_decimal	datas	Diferenca_mm	Erro relativo (mm√km)
2806S	9351C	-4,69655	0,65	7	1985	1985,58			
2806S	9351C	-4,6936	0,645	6	1996	1996,50			
2806S	9351C	-4,69508	0,6475				7/1985-6/1996	-2,95	-3,67
2806S	9351C	-4,69655	0,65	7	1985	1985,58			
2806S	9351C	-4,6947	0,65	6	2003	2003,50			
2806S	9351C	-4,69563	0,65				7/1985-6/2003	-1,85	-2,29
2806S	9351C	-4,6936	0,645	6	1996	1996,50			
2806S	9351C	-4,6947	0,65	6	2003	2003,50			
2806S	9351C	-4,69415	0,6475				6/1996-6/2003	1,1	1,37

Source: Authors

In evaluating the paths, visualized by the histograms in **figure** 5, it can be seen that the 1960 measurement agrees with 2/3 of the detected paths, while the 1987 measurement agrees with 1/3. Seeking to evaluate this behavior, we set out to analyze the circuits.

Fig. 5 – Histograms with the calculated unevenness of the 3 paths detected plus the unevenness of edge 453G-9320D, in dashed lines.



Source: Authors

Table 2 presents the results of closing the circuits, considering the two measurements of section 453G-9320D and the survey times of the sections involved. And **figure** 6 illustrates the drawings of the circuits identified and calculated according to **table** 2.

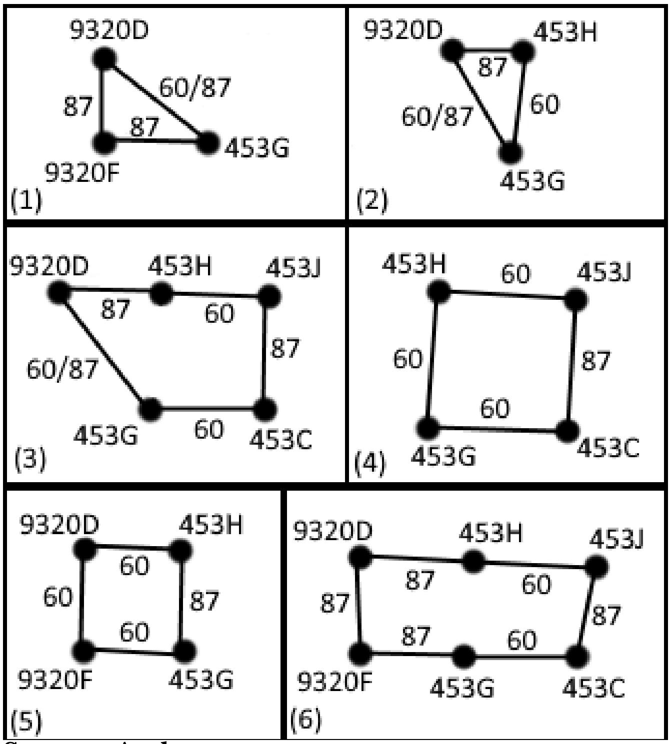
With these data whenever the circuits use measurements from 1960 and 1987 at the same time on the RN 453G, they present a closing error between 12 and 14 mm, while in the other cases the closing errors are smaller or equal to 1 mm. However, this fact does not occur with the 9320D, indicating that most likely the RN 453G was shaken between the two surveys. Therefore, even though it is not possible to apply the usual evaluation method, the inputs generated with the proposed methodology allowed, in this case, to indicate with a high possibility the existence of a shock to the RN 453G.

Tab. 2 - Closing Error Of The Circuits Of Figure 6.

	1960	**1987**
Circuit 01	-13,9 mm	-0,3 mm
Circuit 02	-0,7 mm	12,8 mm
Circuit 03	-1,0 mm	12,6 mm
Circuit 04	-0,3 mm	
Circuit 05	-13,2 mm	
Circuit 06	12,9 mm	

Source: Authors

Fig. 6 – Circuit design considering the two measurements of section 453G-9320D (1960 and 1987).



Source: Authors.

8. Conclusions

In this work, a methodology was proposed that optimizes the usual method, explores alternatives when it is not applicable, taking advantage of the available information, and which aims to make the temporal evaluation of leveling data more homogeneous and less subjective.

The methodology proved to be promising in generating inputs to support the analyst in the temporal evaluation process. When there is not enough information to allow the application of the usual temporal evaluation method to be applied, the methodology makes it possible to indicate the most consistent measurement with the surrounding network.

Additionally, as it is automated, the methodology makes it possible to explore possible compositions in the region to enhance the analysis, which can be seen as a gain over the usual method.

This research is part of the master's thesis in development, where Graph Theory is being applied to the process of evaluating primary data of leveling networks, with the purpose of detecting and identifying outliers and contributing to the process of refinement and processing of networks altimetric reference. Additional information about the dissertation in [10].

Thanks

The authors would like to thank the Brazilian Institute of Geography and Statistics (IBGE) for providing RAAP leveling data for the development of this work.

References

[1] J. A. Bondy, U. S. R. Murty. Graph Theory with Applications. Ontario: Elsevier Science Publishing Co., Inc, 1976. 270p.

[2] IBGE – Instituto Brasileiro de Geografia e Estatística. Especificações e Normas para Levantamentos Geodésicos associados ao Sistema Geodésico Brasileiro, Manual Técnico, Rio de Janeiro, 2017. 56p. Disponível em: <https://geoftp.ibge.gov.br/metodos_e_outros_documentos_de_referencia/normas/normas_levantamentos_geodesicos.pdf>

[3] G. O'Regan. Guide to Discrete Mathematics. [s.1.] Springer, 2016. 378p. DOI: 10.1007/978-3-319-44561-8_1.

[4] T. Kavitha. et al. An Õ(m^2n) Algorithm for Minimum Cycle Basis of Graphs Algorithmica 52, p. 333-349, 2008. DOI: 10.1007/s00453-007-9064-z.

[5] K. D. Lee, S. Hubbard. Data Structures and Algorithms with Python. Springer, 2015. 369p.

[6] S. S. Suraci. Aplicação das normas L1 e L∞ em redes altimétricas: identificação de outliers e construção do modelo estocástico. Dissertação (Mestrado em Engenharia Cartográfica). Instituto Militar de Engenharia. Brasil, 2018.

[7] Pandas. Pandas documentation. Versão 1.0.5. 2020. Disponível em: <https://pandas.pydata.org/pandas--docs/version/1.0.5/>.

[8] A. A. Hagberg, D. A. Schult, P. J. Swart. Exploring network structure dynamics, and function using NetworkX. Proceedings of the 7th Python in Science Conference (SciPy2008). G. Varoquaux, T. Vaught, J. Millman (Eds.) p. 11-15, 2008. Disponível em: <https://networkx.org/documentation/stable/index.html>.

[9] A. L. T. Penha. et al. Identificação de erros grosseiros em dados geodésicos unidimensionais. Projeto de fim de curso em Engenharia Cartográfica. Instituto Militar de Engenharia, Brasil, 2007.

[10] J. C. S. Siqueira. Teoria dos Grafos aplicada ao processo de avaliação dos dados primários de redes de nivelamento. Dissertação de Mestrado em desenvolvimento (Mestrado em Engenharia de Defesa). Instituto Militar de Engenharia, Brasil.