www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK

GEODESY AND CARTOGRAPHY
Vol. 62, No 1, 2013, pp. 23-31

DOI: 10.2478/geocart-2013-0002

Analysis of the horizontal structure of a measurement and control geodetic network based on entropy

Maria Mrówczyńska

University of Zielona Gora
Faculty of Land and Environment Engineering
Institute of Building Engineering
1 Szafrana St., 65-516 Zielona Gora, Poland
e-mail: m.mrowczynska@ib.uz.zgora.pl

Received: 14 September 2012 / Accepted: 3 January 2013

Abstract: The paper attempts to determine an optimum structure of a directional measurement and control network intended for investigating horizontal displacements. For this purpose it uses the notion of entropy as a logarithmical measure of probability of the state of a particular observation system. An optimum number of observations results from the difference of the entropy of the vector of parameters $\Delta H_{\hat{\mathbf{X}}}(\mathbf{X})$ corresponding to one extra observation. An increment of entropy interpreted as an increment of the amount of information about the state of the system determines the adoption or rejection of another extra observation to be carried out.

Keywords: entropy, amount of information, geodetic network.

1. Introduction

While designing the structure of geodetic control networks an important factor is the optimisation of the structure with possibly the highest amount of information. This course of action provides possibilities of satisfying assumptions about the accuracy characteristics of the parameters being determined and the sensitivity of a model of displacements (Sharif et al., 2012). The identification of movement of points by means of the geodetic method through classic measurements (distance, directions) requires the satisfaction of geometrical conditions for geodetic networks. The importance of geometrical and technical conditions (Nowak, 1993) is unquestionable from the point of view of the accuracy of determining the position of controlled points.

The article presents a suggestion for a construction of a horizontal linear geodetic network and a directional one according to a criterion of changes of information entropy, which provides a particular amount of information about the state of the observation system.

2. Information entropy

The mathematical measure of lack of information about the internal state of a system is information entropy defined as a logarithm with a negative average value P of the probabilities of all the possible states of the system according to the formula (Bilgi et al., 2008, Philips el al., 2006).

$$H(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_r p_i \tag{1}$$

where: p_i denotes the probability of occurrence of an *i-th* result from all the n possible states, r denotes a choice of an entropy measurement unit. Using a logarithm with a base of 2 we express entropy in bits (a bit – an entropy measurement unit used in technology) (Brillouin, 1969). Entropy as a measure of "lost information" can be determined only for systems all the states of which create an absolute system i.e. when

$$\sum_{i=1}^{n} p_i = 1\,. \tag{2}$$

Entropy is a positive value or equal zero $H \geq 0$, which reaches a minimum when the system reaches only one state with the probability $p_i = 1$. In the case of a step change in the states of the system with an equal probability of their occurrence entropy reaches a maximum because (Wu et al. 2004)

$$H = H_{\max} = -\sum_{i=1}^{n} p_i \log p_i = -n\frac{1}{n}\log\frac{1}{n} = \log n \tag{3}$$

For systems the change of states of which occurs continuously entropy reaches a maximum (Peng et al. 2006) when the probability of their occurrence is in normal distribution, i.e.

$$\begin{aligned} H_{\max} &= -\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{-x^2}{2\sigma^2}} \ln\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}\right) dx = \ln\left(\sigma\sqrt{2\pi e}\right) = \\ &= \ln 4{,}13266[nat] = 1[bit] \end{aligned} \tag{4}$$

where: x – random variable, σ – standard deviation of a probability density function, *nat* – information amount unit (1 $bit = 1{,}4427$ *nat*).

Entropy can be discussed in terms of elementary events and in terms of complex events. If two events α and β are independent, then the value of entropy, as an additive value, can be written as a sum of the entropy of the event α and the event β in the form (Gil, 1997; Shannon and Weaver, 1949)

$$H(\alpha\beta) = H(\alpha) + H(\beta). \tag{5}$$

The dependence of the events α and β leads to the conclusion that the event β will occur on condition that the event α occurs or vice versa, and then

$$H(\alpha\beta) = H(\alpha) + H_\alpha(\beta) \tag{6}$$

or

$$H(\alpha\beta) = H(\beta) + H_\beta(\alpha). \tag{7}$$

The dependence of events decreases the value of unconditional entropy.

In practical problems the most important is an increment of entropy as a difference between the entropy of the state under discussion and the state of the reference system.

3. Entropy of a geodetic network

A certain indeterminacy results from the random character of the vector of geodetic observations $\mathbf{L}=(l_1,l_2,\ldots,l_n)$, which concerns the real values of its components. The vector $\mathbf{L}$ is weighted by the random vector $\boldsymbol{\mu}\tilde{}N(0,\sigma^2)$, and its real value is $\mathbf{Z}=\mathbf{L}+\boldsymbol{\mu}$. The equalization of the vector of observations $\mathbf{L}$ will result in weighting the components $\hat{\mathbf{X}}=(x_1,x_2,\ldots,x_n)$, of the vector of parameters $\mathbf{X}=(x_1,x_2,\ldots,x_n)$ with a certain rate of indeterminacy (Lehmann and Scheffler, 2011).

As additional observations are carried out there occurs a change in the state of the observation system. The difference of the entropies of two states, described with a normal distribution with standard deviations σ_1 and σ_2 is (cf. formula (4))

$$H_2 - H_1 = \log\left(\sigma_2\sqrt{2\pi e}\right) - \log\left(\sigma_1\sqrt{2\pi e}\right) = -\log\frac{\sigma_1}{\sigma_2} \tag{8}$$

The difference of the entropies is an increment of information understood as the difference of the initial and the final rate of information about the state of the system (Wang 2010).

The information contributed by a single additional observation to the general information of the state of the observation system, written in line b of the matrix $\mathbf{A}$, including the elements

$$b\left(a_{m+1,1}, a_{m+1,2}, \ldots, a_{m+1,n}\right) \tag{9}$$

has been described by the author of the article (Neuman, 1965) with the dependence

$$\Delta h = \log_2 \frac{g(m,n)\left\{1 + trace\left[\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(b^T b\right)\right]\right\}}{g(m+1,n)\prod_{j=1}^{n}\sqrt{1+\frac{a_{m+1,j}^2}{\sum_{i=1}^{m} a_{ij}^2}}} \tag{10}$$

where:

$$g(m,n) = \frac{(m-n)^{\frac{n}{2}} B\left(\frac{m-n}{2},\frac{n}{2}\right)\exp\left\{\frac{m}{2},\left[\Psi\left(\frac{m}{2}\right)-\Psi\left(\frac{m-n}{2}\right)\right]\right\}\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \tag{11}$$

and

- a_{ij} – elements of the matrix **A**,
- $B\left(\frac{m-n}{2},\frac{n}{2}\right)$ – beta function,
- $\Psi\left(\frac{m}{2}\right),\Psi\left(\frac{m-n}{2}\right)$ – psi Euler function (algorithmic derivative of the gamma function),
- $\Gamma\left(\frac{n}{2}\right)$ – gamma function.

The same author (Neuman, 1965) also invented a formula for the entropy of the estimator $\hat{\mathbf{X}}$ of the vector of parameters **X** in the form

$$H_{\hat{\mathbf{X}}}(\mathbf{X}) = \log_2\left[\frac{g(m,n)m_0^n \prod_{j=1}^{n}\sqrt{\sum_{i=1}^{m} a_{ij}^2}}{\prod_{j=1}^{n}\varepsilon_j \det\left(\mathbf{A}^T\mathbf{A}\right)}\right] \tag{12}$$

where:

- m_0 – mean *a priori* error of a single observation [mm],
- $\varepsilon_j = m_0\sqrt{\frac{\delta_0}{\lambda_j}}$ – sensitivity of the system along a particular axis of coordinates,

and

- δ_0 – non-centrality parameter,
- λ_j – own value of the matrix $\mathbf{A}^T\mathbf{A}$.

4. Numerical example

The entropy of an observation system can be discussed in the aspect of the vector of observations **L** in relations to the vector **Z**, and in the aspect of the entropy of the vector of parameters $\hat{\mathbf{X}}$ in relation to its real values **X**. Because of the limited volume of this paper we will only discuss the second case.

In order to present possibilities of using entropy for designing an optimum structure of a directional geodetic network, a network consisting of 16 points and 124 observed directions was analysed (Figure 1). In order to satisfy the technological correctness requirement, 8 observations from a total of 124 were eliminated.

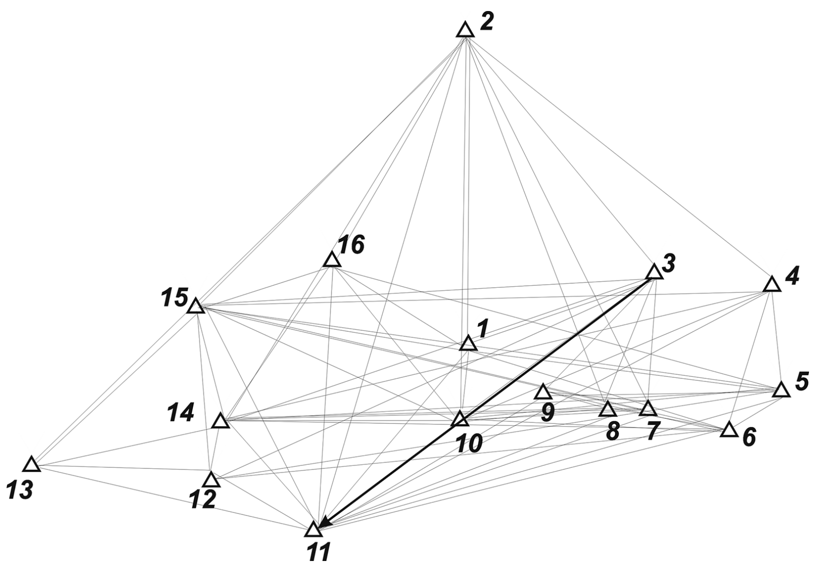

Fig. 1. General diagram of a directional network

The network was equalized with minimum restrictions on rates of freedom while making points 3 and 11 stable and the mean *a priori* error of a single observation $m_{obs}=\pm 4^{cc}$, which after equalization reached the value $m_0=1,17$.

The equalization process provides a number of valuable directions about the perception of entropy of an observation system. The first approach can be the identification of mean errors of observations with a particular amount of information in order to omit unpromising data in further discussion. Observations with the largest amount of information do not usually satisfy the technological condition $m_v \geq 0,7$ about the internal reliability of a network (Nowak and Prószyński, 1990). It was found that the amount of information in particular observations (Figure 2) depended on the assumptions adopted for the elimination of the external defect of the network, but the amount of information in the whole system remained approximately stable independently of the assumptions adopted.

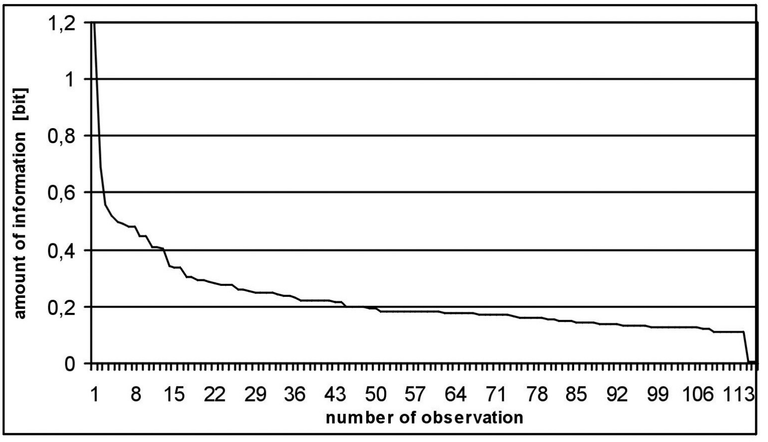

Fig. 2. Amount of information in particular observations

It results from the equalization of the system (116 observations) with minimum restrictions on rates of freedom that the sum of the amount of information is $[\Delta h] = 25{,}9982$, the entropy of the vector of parameters is $H_{\hat{\mathbf{X}}}(\mathbf{X}) = 22{,}7367$ [bit], and there is a reversely proportional dependence between the amount of information in particular observations and the mean residuum error (Figure 3), which may be used for diagnostic purposes in the network (Nowak and Prószyński, 1990). If all the possible observations in the network were considered (240 observations), then the entropy of the vector of parameters **X** would increase to the value $H_{\hat{\mathbf{X}}}(\mathbf{X}) = 32{,}2102$ [bit].

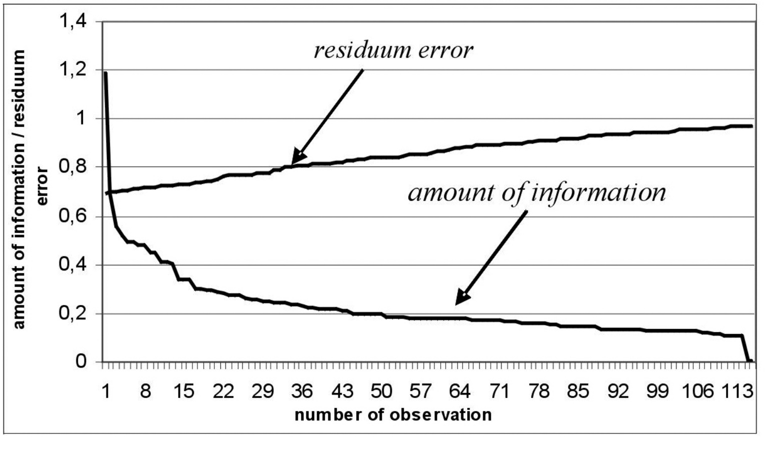

Fig. 3. Amount of information and the residuum error

It was found from the analysis that the minimum of entropy occurs when the number of observations is 101 (Figure 4), then:

– entropy of the vector of parameters $H_{\hat{\mathbf{X}}}(\mathbf{X}) = 22{,}3491$,
– mean error from equalization with minimum restrictions on rates of freedom $m_0 = 1{,}0206$.

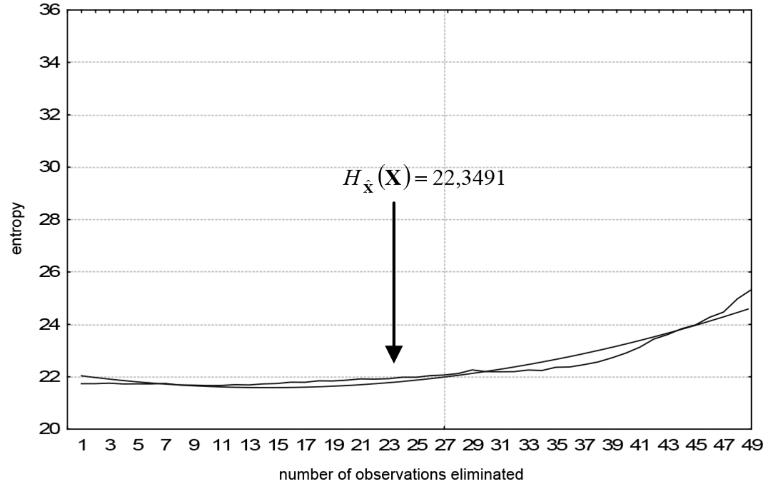

Fig. 4. Changes in entropy in the process of eliminating observations

While determining an optimum number of observations in a geodetic network, it is not the value of entropy itself that is important, but its increment, which is interpreted as an increase or a decrease in the amount of information about the state of the system. Each subsequent additional observation will result in an increase of entropy $\Delta H_{\hat{\mathbf{X}}}(\mathbf{X})$, which will be compared with the entropy determined for a system with the probabilities α_0 and $\beta_0 = 1 - \alpha_0$. The value of the probability α_0 is an adopted rate of importance of the estimator $\hat{\mathbf{X}}$ of the vector of parameters $\mathbf{X}$ before and after adding each subsequent additional observation. Using the notion of conditional entropy and the abovementioned course of action, a criterion for the adoption of another observation to be carried out will be described by the formula (Gil, 1999)

$$-\alpha_0 \log \alpha_0 - \beta_0 \log \beta_0 \geq \Delta H_{\hat{\mathbf{X}}}(\mathbf{X}). \tag{13}$$

For the test network consisting of 16 points, in which observations of directions were carried out, the number of measurements n is as follows:
– for $\alpha_0 = 0{,}10$ $\beta_0 = 0{,}90$ $\delta_0 = 6{,}2$ $n = 80$
– for $\alpha_0 = 0{,}05$ $\beta_0 = 0{,}95$ $\delta_0 = 7{,}6$ $n = 90$
– for $\alpha_0 = 0{,}01$ $\beta_0 = 0{,}99$ $\delta_0 = 11{,}6$ $n = 103$.

Detailed characteristics of the value of entropy obtained, dependently on the criterion used for determining an optimum number of observations, have been presented in Table 1.

Table 1. Results of calculations

Number of observations	Entropy [bit]
80 observations determined for $\beta_0=0{,}90$ $\alpha_0=0{,}10$ $\delta_0=6{,}2$	25,5777
90 observations determined for $\beta_0=0{,}95$ $\alpha_0=0{,}05$ $\delta_0=7{,}6$	23,1173
101 observations determined on the basis of the minimum of entropy	22,3491
103 observations determined for $\beta_0=0{,}99$ $\alpha_0=0{,}01$ $\delta_0=11{,}6$	22,6716
116 observations which were adopted for further calculations after the elimination of aberrant observations (initially 124)	22,7367
240 all the possible observations in the network	32,2102

5. Conclusions

The paper presents one of the possible methods of optimization, as far as the number of elements measured in measurement-control geodetic networks is concerned. On the basis of the calculations and analysis carried out, it is possible to say that the amount of information in particular observations has a considerable influence on the determination of an optimum structure of geodetic networks. When the amount of information, the value of entropy, and the trace of a covariance matrix are known, it is possible to choose the right number of observations so as to preserve favourable accuracy characteristics with less work necessary.

Acknowledgements

The article was written as statutory research carried out at the Department of Civil and Environmental Engineering at the University of Zielona Góra. The author of the article extends her gratitude to the reviewers, whose remarks considerably improved the quality of the article. She wants to especially thank Prof. Józef Gil, whose valuable knowledge and inexhaustible patience greatly contributed to the publication of this article.

References

Bilgi, S., Ipbuker, Ucer D. & Muhammed S. (2008). Map entropy analysis of topographic data used in disaster information systems. *Journal of Earthquake Engineering,* 12, Supplement 2, 23-36. DOI: 10.1080/13632460802013438.

Brillouin, L. (1969). *La science et la théorie de l'information. Polish edition: Nauka a teoria informacji*. Warszawa: PWN.

Gil, J. (1997). Analysis of horizontal geodetic network on the entropy basis (in Polish). In III Scientific – Technical Conference on Problems of Automation in Engineering Surveying, 20-21 March 1997 (pp. 93-107). Warsaw, Poland: Polish Academy of Sciences.

Gil, J. (1999). The problem of the enthropy o fan observation system in the research of displacement. The 9th FIG International Symposium on Deformation Measurements, 27-30 September 1999 (pp. 256-266) Olsztyn, Poland: Warmia and Mazury University.

Lehmann, R. & Scheffler T. (2011). Monte Carlo-based data snooping with application to a geodetic network. *Journal of Applied Geodesy*, 5(3-4), 123-134. DOI: 10.1515/JAG.2011.014.

Neuman, J. (1965). K analizu geodeziczeskich postrojenii. Izwestia wuzow. Moscow, Rusia: MIIZ.

Nowak, E. & Prószyński W. (1990). *Analysis of horizontal engineering survey networks theory and worked examples (in Polish)*. User Guide SYPO. Warsaw, Poland: SYPO.

Nowak, E. (1993). RMS Technique. Geometrical and technical factors (in Polish). In III Scientific – Technical Conference on Problems of Automation in Engineering Surveying, 20-21 March 1997 (pp. 177-190). Warsaw, Poland: Polish Academy of Sciences.

Peng, J., Guo Ch., Hang H. & Shong S. (2006). An aggregate constrain method for inequality – constrained lest squares problems. *Journal of Geodesy,* 79, 705-713. DOI: 10.1007/s00190-006-0026-z.

Philips, S.J., Anderson R.P. & Schapire R.E. (2006). Maximum entropy modeling of species geographic distribution. *Ecological Modelling,* 190 (2006), 231-259.

Schaffrin, B. & Cothren J. (2004). Hierarchical data fusion with photoframmetric applications. *Revista Brasileira de Cartografia,* 55/02, 25-34.

Shannon, C.E. & Weaver W. (1949). *The mathematical theory of communication*, Urbana: University of Illinois Press.

Sharif, S., Yusoff N.S. & Djauharin M.A. (2012). Network Topology of Foreign Exchange Rate. *Modern Applied Science*, 6(11), 35-43. DOI:10.5539/mas.v6n11p35.

Wang, H. (2010). An hierarchical information entropy model coverage estimation of coastal areas based on an adaptive Discrete Global Grid System. In Geoinformatics, 2010 18th International Conference, 18-20 June 2010 (pp.1-5). Beijing, China: IEEE Geoscience and Remote Sensing Society. DOI:10.109/GEOINFORMATICS.2010.5567822.

Wu, H., Zhu H. & Liu Y. (2004). A raster-based map information mearurment for QoS. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 35(2), 365-370.

Analiza struktury poziomej sieci geodezyjnej pomiarowo-kontrolnej na podstawie entropii

Maria Mrówczyńska

Uniwersytet Zielonogórski
Wydział Inżynierii Lądowej i Środowiska
Instytut Budownictwa
ul. Szafrana 1, 65-516 Zielona Góra
e-mail: m.mrowczynska@ib.uz.zgora.pl

Streszczenie

W pracy podjęto próbę określenia optymalnej struktury sieci kierunkowej pomiarowo-kontrolnej przeznaczonej do badań przemieszczeń poziomych. W tym celu wykorzystano pojęcie entropii jako logarytmicznej miary prawdopodobieństwa stanu określonego układu obserwacyjnego. Optymalna liczba realizowanych obserwacji wynika z różnicy entropii wektora parametrów $\Delta H_{\hat{\mathbf{X}}}(\mathbf{X})$ odpowiadającej jednej obserwacji nadliczbowej. Przyrost entropii interpretowany jako przyrost objętości informacji na temat stanu układu decyduje o przyjęciu względnie odrzuceniu do realizacji kolejnej obserwacji nadliczbowej.