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

ADVANCES IN GEODESY AND GEOINFORMATION
(formerly GEODESY AND CARTOGRAPHY)
Vol. 71, no. 1, article no. e22, 2022

https://doi.org/10.24425/gac.2022.141175

Original paper

Part of Topical collection:
"Advancements in Applied Geoinformatics"

Two variants of M_{split} estimation – similarities and differences

Patrycja Wyszkowska*, Robert Duchnowski

University of Warmia and Mazury, Olsztyn, Poland
e-mail: patrycja.wyszkowska@uwm.edu.pl; ORCID: http://orcid.org/0000-0002-0080-7897
e-mail: robert.duchnowski@uwm.edu.pl; ORCID: http://orcid.org/0000-0002-6331-8345

*Corresponding author: Patrycja Wyszkowska, e-mail: patrycja.wyszkowska@uwm.edu.pl

Received: 2021-02-24 / Accepted: 2022-04-15

Abstract: M_{split} estimation is a novel method developed to process observation sets that include two (or more) observation aggregations. The main objective of the method is to estimate the location parameters of each aggregation without any preliminary assumption concerning the division of the observation set into respective subsets. Up to now, two different variants of M_{split} estimation have been derived. The first and basic variant is the squared M_{split} estimation, which can be derived from the assumption about the normal distribution of observations. The second variant is the absolute M_{split} estimation, which generally refers to the least absolute deviation method. The main objective of the paper is to compare both variants of M_{split} estimation by showing similarities and differences between the methods. The main dissimilarity stems from the different influence functions, making the absolute M_{split} estimation less sensitive to gross errors of moderate magnitude. The empirical analyses presented confirm that conclusion and show that the accuracy of the methods is similar, in general. The absolute M_{split} estimation is more accurate than the squared M_{split} estimation for less accurate observations. In contrast, the squared M_{split} estimation is more accurate when the number of observations in aggregations differs much. Concerning all advantages and disadvantages of M_{split} estimation variants, we recommend using the absolute M_{split} estimation in most geodetic applications.

Keywords: accuracy, influence function, absolute M_{split} estimation, squared M_{split} estimation

1. Introduction

M_{split} estimation was introduced by (Wisniewski, 2009) as a development of the maximum likelihood method. It is based on the general assumption that an observation set might be a mixture of realizations of two random variables that differ in location parameters.

The Author(s). 2022 Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Such an assumption leads to splitting the functional model into two competitive ones. It also results in two competitive versions of the parameter and two versions of observation errors. The main aim of M_{split} estimation is to estimate such parameters without dividing the observation set into the subsets preliminarily. Assignment of each observation to either of the competitive functional models is automatic and done during an iterative process. The first variant of M_{split} estimation refers to the general assumptions of the least squares method; hence, it is called the squared M_{split} estimation (SMS). The general formula of the objective function can be derived from the assumption that the normal distribution is the stochastic model of observation errors. This variant was successfully applied to several geodetic problems (Wisniewski, 2009; Janicka and Rapinski, 2013; Li et al., 2013; Czaplewski et al., 2019; Nowel, 2019; Guo et al., 2020; Janicka et al., 2020; Zienkiewicz and Baryla, 2020; Baselga et al., 2021; Zienkiewicz, 2022). The main problem arises because SMS estimation is sensitive to the outlying observations (observations that do not match two main observation groups); this fact was addressed in (Duchnowski and Wisniewski, 2020). Thus, there was a need for a new variant of M_{split} estimation, which was less sensitive to outliers. For that reason, the second variant of M_{split} estimation, called the absolute M_{split} estimation (AMS), was introduced (Wyszkowska and Duchnowski, 2019). In general, that variant refers to the fundamental assumptions of the least absolute deviation method. Respective influence functions make the new variant less sensitive to the observations, which errors have a more significant magnitude.

The main purpose of the paper is to present the main similarities and differences between the methods mentioned. It is essential to understand the fundamental differences in features and computation ways of both variants of M_{split} estimation. It helps to understand what observation sets can be processed by applying M_{split} estimation and what outcomes one can expect. Such knowledge would define the possible practical applications of both variants. Up to now, such a comparison was not presented in one paper.

2. General assumptions

M_{split} estimation is a modern development of the maximum likelihood estimation based on a fundamental assumption that an observation set can consist of several groups of observations. One can list different reasons for such a case. The differences between the observation groups might stem from the measurement technique applied, different accuracy of measurements, the occurrence of outliers, etc. Generally, we can assume that the observation set is an unrecognized mixture of realizations of two or more random variables, which differ at least in location parameters. It means that the classical functional model should be split into two (or more) competitive ones:

$$\mathbf{y} = \mathbf{AX} + \mathbf{v} \;\Rightarrow\; \mathbf{y} = \mathbf{AX}_{(l)} + \mathbf{v}_{(l)} \tag{1}$$

where: $\mathbf{y}$ – observation vector, $\mathbf{A}$ – coefficient matrix, $\mathbf{X}$ – parameter vector, $\mathbf{v}$ – measurement error vector, $\mathbf{X}_{(l)}$, $\mathbf{v}_{(l)}$ – variants of parameter vector or measurement error vector, respectively. The most applications of M_{split} estimation assume that there are two competitive models, hence $l = 1$ or 2 (Wyszkowska and Duchnowski, 2019; Zienkiewicz

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

and Baryla, 2020; Wyszkowska et al., 2021). The main objective of M_{split} estimation is to estimate the competitive versions of the parameters of the model of Eq. (1) by solving the general optimization problem by minimization of the objective function in the following general form (Wisniewski, 2009; Wyszkowska and Duchnowski, 2020):

$$\varphi\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)=\sum_{i=1}^{n}\rho\left(v_{i(1)},v_{i(2)}\right)=\sum_{i=1}^{n}\prod_{l=1}^{2}\rho_{(l)}\left(v_{i(l)}\right) \tag{2}$$

where: ρ and $\rho_{(l)}$ might result from the probabilistic model of observation errors or be arbitrarily chosen functions. We might assume that the accuracy of all observations is the same; however, if the observations differ from each other in *a priori* accuracy, then the standardized errors should be the arguments of functions in Eq. (2) (Wisniewski, 2010). We can use the Newton method to solve the optimization problem and find the competitive estimates $\hat{\mathbf{X}}_{(1)},\hat{\mathbf{X}}_{(2)}$ (Wisniewski, 2010). We have two possible iterative processes in such a context. The traditional iterative process is defined as follows (Wisniewski, 2009):

$$\begin{aligned}\mathbf{X}_{(1)}^{j}&=\mathbf{X}_{(1)}^{j-1}+d\mathbf{X}_{(1)}^{j}=\mathbf{X}_{(1)}^{j-1}-\left[\mathbf{H}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\\ \mathbf{X}_{(2)}^{j}&=\mathbf{X}_{(2)}^{j-1}+d\mathbf{X}_{(2)}^{j}=\mathbf{X}_{(2)}^{j-1}-\left[\mathbf{H}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right)\end{aligned} \tag{3}$$

where: $\mathbf{H}_{(l)}(\circ)$ – Hessian, $\mathbf{g}_{(l)}(\circ)$ – gradient. The respective Hessians and gradients are computed by applying the respective weight functions $w_{(1)}(v_{(1)},v_{(2)})$ and $w_{(2)}(v_{(1)},v_{(2)})$:

$$\begin{aligned}&\begin{cases}\mathbf{H}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)=2\mathbf{A}^{T}\mathbf{w}_{(1)}\left(\mathbf{v}_{(1)}^{j-1},\mathbf{v}_{(2)}^{j-1}\right)\mathbf{A}\\ \mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)=-2\mathbf{A}^{T}\mathbf{w}_{(1)}\left(\mathbf{v}_{(1)}^{j-1},\mathbf{v}_{(2)}^{j-1}\right)\mathbf{v}_{(1)}^{j-1}\\ \mathbf{w}_{(1)}\left(\mathbf{v}_{(1)}^{j-1},\mathbf{v}_{(2)}^{j-1}\right)=\text{Diag}\left[w_{(1)}\left(v_{1(1)}^{j-1},v_{1(2)}^{j-1}\right),\ldots,w_{(1)}\left(v_{n(1)}^{j-1},v_{n(2)}^{j-1}\right)\right]\end{cases}\\ &\begin{cases}\mathbf{H}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right)=2\mathbf{A}^{T}\mathbf{w}_{(2)}\left(\mathbf{v}_{(1)}^{j},\mathbf{v}_{(2)}^{j-1}\right)\mathbf{A}\\ \mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right)=-2\mathbf{A}^{T}\mathbf{w}_{(2)}\left(\mathbf{v}_{(1)}^{j},\mathbf{v}_{(2)}^{j-1}\right)\mathbf{v}_{(2)}^{j-1}\\ \mathbf{w}_{(2)}\left(\mathbf{v}_{(1)}^{j},\mathbf{v}_{(2)}^{j-1}\right)=\text{Diag}\left[w_{(2)}\left(v_{1(1)}^{j},v_{1(2)}^{j-1}\right),\ldots,w_{(2)}\left(v_{n(1)}^{j},v_{n(2)}^{j-1}\right)\right]\end{cases}\end{aligned} \tag{4}$$

The parallel iterative process can be written as (Wyszkowska and Duchnowski, 2019):

$$\begin{aligned}\mathbf{X}_{(1)}^{j}&=\mathbf{X}_{(1)}^{j-1}+d\mathbf{X}_{(1)}^{j}=\mathbf{X}_{(1)}^{j-1}-\left[\mathbf{H}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\\ \mathbf{X}_{(2)}^{j}&=\mathbf{X}_{(2)}^{j-1}+d\mathbf{X}_{(2)}^{j}=\mathbf{X}_{(2)}^{j-1}-\left[\mathbf{H}_{(2)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\end{aligned} \tag{5}$$

and

$$\begin{cases} \mathbf{H}_{(a)}\left(\mathbf{X}_{(a)}^{j-1},\mathbf{X}_{(b)}^{j-1}\right) = 2\mathbf{A}^T\mathbf{w}_{(a)}\left(\mathbf{v}_{(a)}^{j-1},\mathbf{v}_{(b)}^{j-1}\right)\mathbf{A} \\ \mathbf{g}_{(a)}\left(\mathbf{X}_{(a)}^{j-1},\mathbf{X}_{(b)}^{j-1}\right) = -2\mathbf{A}^T\mathbf{w}_{(a)}\left(\mathbf{v}_{(a)}^{j-1},\mathbf{v}_{(b)}^{j-1}\right)\mathbf{v}_{(a)}^{j-1} \\ \mathbf{w}_{(a)}\left(\mathbf{v}_{(a)}^{j-1},\mathbf{v}_{(b)}^{j-1}\right) = \mathrm{Diag}\left[w_{(a)}\left(v_{1(a)}^{j-1},v_{1(b)}^{j-1}\right),\ldots,w_{(a)}\left(v_{n(a)}^{j-1},v_{n(b)}^{j-1}\right)\right] \end{cases} \tag{6}$$

where $a = 1$ and $b = 2$ or $a = 2$ and $b = 1$, respectively. Both iterative processes end when both gradients $\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{k-1},\mathbf{X}_{(2)}^{k-1}\right)$, $\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{k-1},\mathbf{X}_{(2)}^{k-1}\right)$ are equal to 0 (or at least when each parameter changes no more than ε between subsequent iterative steps, where ε – an assumed small positive number, e.g., 0.01% of parameter values). Hence finally, $\hat{\mathbf{X}}_{(1)} = \mathbf{X}_{(1)}^k$ and $\hat{\mathbf{X}}_{(2)} = \mathbf{X}_{(2)}^k$, where k – the number of the last iterative step. The general schemes of both iterative processes are presented in Figure 1.

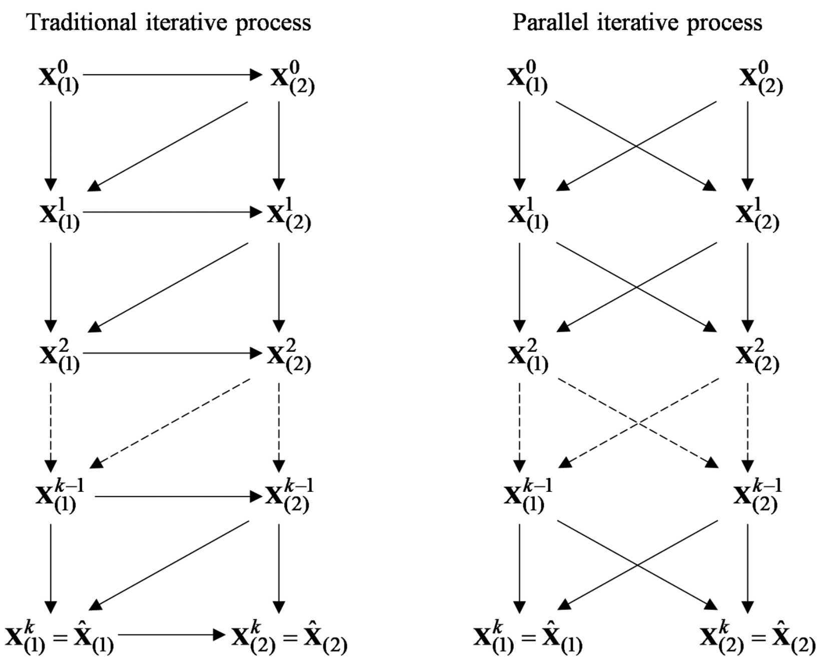

Fig. 1. Schemes of both iterative processes

In both iterative processes, the weight functions play a significant role. What is more, the choice of either of the algorithms depends on the properties of the functions mentioned (Wyszkowska and Duchnowski, 2019). Considering the general formulas of the weight functions (they are presented in the next section), one can recommend the traditional iterative process for SMS estimation and the parallel iterative process for AMS estimation (Wyszkowska and Duchnowski, 2020).

3. Comparison of squared and absolute $\mathbf{M_{split}}$ estimation

The first difference between SMS and AMS estimations is given in the previous section, and it concerns the type of iterative process convenient for each method. However, it is not the most essential difference. The crucial theoretical dissimilarities stem from the assumed objective functions, hence also from the influence and weight functions. It also influences the practical use of both methods. To compare the basic theoretical features and practical usages of both methods, let us assume the following univariate functional model:

$$\mathbf{y} = \mathbf{1}_n X + \mathbf{v} \quad \Rightarrow \quad \begin{cases} \mathbf{y} = \mathbf{1}_n X_{(1)} + \mathbf{v}_{(1)} \\ \mathbf{y} = \mathbf{1}_n X_{(2)} + \mathbf{v}_{(2)} \end{cases} \tag{7}$$

where: $\mathbf{1}_n$ – all-ones vector (n – observation number), X – parameter, $X_{(l)}$ – its competitive versions. Such an assumption lets us compare the methods and present results more clearly, and the general conclusion can be applied to the multivariate case. Referring to Eq. (2) and taking the model of Eq. (7), one can write the following objective functions of SMS or AMS estimations (Wisniewski, 2009; Wyszkowska and Duchnowski, 2019):

$$\begin{aligned} \varphi_{\text{SMS}}\left(X_{(1)}, X_{(2)}\right) &= \sum_{i=1}^{n} \rho_{\text{SMS}}\left(v_{i(1)}, v_{i(2)}\right) = \sum_{i=1}^{n} v_{i(1)}^2 v_{i(2)}^2 \\ \varphi_{\text{AMS}}\left(X_{(1)}, X_{(2)}\right) &= \sum_{i=1}^{n} \rho_{\text{AMS}}\left(v_{i(1)}, v_{i(2)}\right) = \sum_{i=1}^{n} \left|v_{i(1)}\right| \left|v_{i(2)}\right| \end{aligned} \tag{8}$$

The difference between such functions is noticeable; however, their graphs seem similar in shape generally (see, Fig. 2).

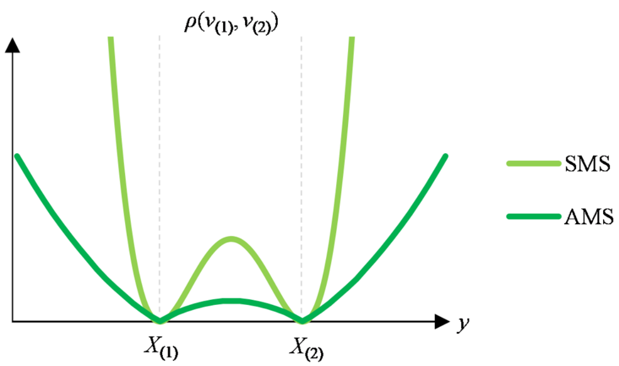

Fig. 2. The component of the objective function $\rho(v_{(1)}, v_{(2)})$ of SMS and AMS estimations

The dissimilarity between the respective objective functions yields much more critical differences in the influence functions, defined as $\psi_{(l)}(v_{i(l)}; y_i, X_{(k)}) =$

$\partial\rho(v_{i(1)}, v_{i(2)})/\partial v_{i(l)}$, where $k = 1$ or 2 and $k \neq l$, and presented as follows:

$$\psi_{\mathrm{SMS}(1)}\left(v_{i(1)}; y_i, X_{(2)}\right) = 2v_{i(1)}v_{i(2)}^2 \quad \text{and}$$
$$\psi_{\mathrm{SMS}(2)}\left(v_{i(2)}; y_i, X_{(1)}\right) = 2v_{i(2)}v_{i(1)}^2,$$
$$\psi_{\mathrm{AMS}(1)}\left(v_{i(1)}; y_i, X_{(2)}\right) = \begin{cases} -\left|v_{i(2)}\right| & \text{for } v_{i(1)} < 0 \\ \left|v_{i(2)}\right| & \text{for } v_{i(1)} > 0 \end{cases} \quad \text{and} \tag{9}$$
$$\psi_{\mathrm{AMS}(2)}\left(v_{i(2)}; y_i, X_{(1)}\right) = \begin{cases} -\left|v_{i(1)}\right| & \text{for } v_{i(2)} < 0 \\ \left|v_{i(1)}\right| & \text{for } v_{i(2)} > 0 \end{cases}$$

The values of the influence function of AMS estimation for the moderate and high absolute values of $v_{(1)}$ (the left panel of Fig. 3) are much smaller than the values of the influence function of SMS estimation (the same is for the $v_{(2)}$ – presented in the right panel of Fig. 3). It means that AMS estimation is less sensitive to such errors, including gross errors of moderate magnitude. However, it should be mentioned that the influence functions of both variants of $\mathrm{M_{split}}$ estimation are not bounded; hence, the methods cannot be classified as robust against outliers.

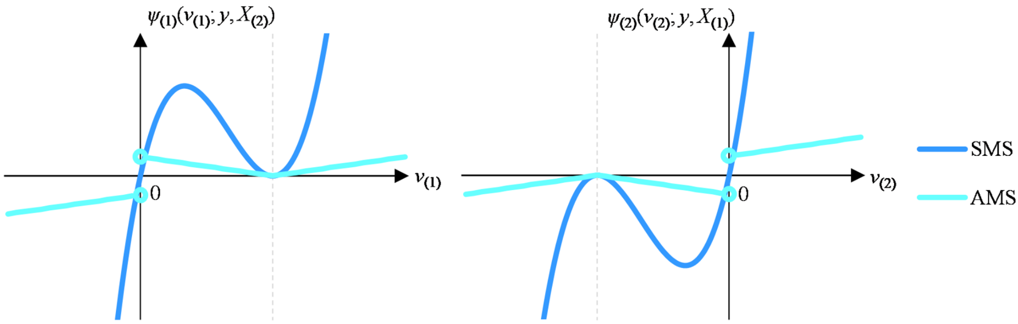

Fig. 3. The influence functions $\psi_{(1)}(v_{i(1)}; y_i, X_{(2)})$ and $\psi_{(2)}(v_{i(2)}; y_i, X_{(1)})$ of SMS and AMS estimations

As mentioned in the previous section, their weight functions are essential in practical computations. Taking $w_{(l)}(v_{i(l)}, v_{i(k)}) = w_{(l)}(v_{i(l)}; y_i, X_{(k)}) = \psi_{(l)}(v_{i(l)}; y_i, X_{(k)})/2v_{i(l)}$, one can write them as follows:

$$w_{\mathrm{SMS}(a)}\left(v_{i(a)}, v_{i(b)}\right) = v_{i(b)}^2$$
$$w_{\mathrm{AMS}(a)}\left(v_{i(a)}, v_{i(b)}\right) = \begin{cases} \dfrac{\left|v_{i(b)}\right|}{2\left|v_{i(a)}\right|} & \text{for } \left|v_{i(a)}\right| \geq d \\ \dfrac{\left|v_{i(b)}\right|}{2d} & \text{for } \left|v_{i(a)}\right| < d \end{cases} \tag{10}$$

where: d – a small positive constant (usually very close to 0, e.g., 0.001). Here, such a constant is introduced to avoid possible singularity (the original weight functions of

AMS estimation can be found in (Wyszkowska and Duchnowski, 2019)). The respective graphs are presented in Figure 4.

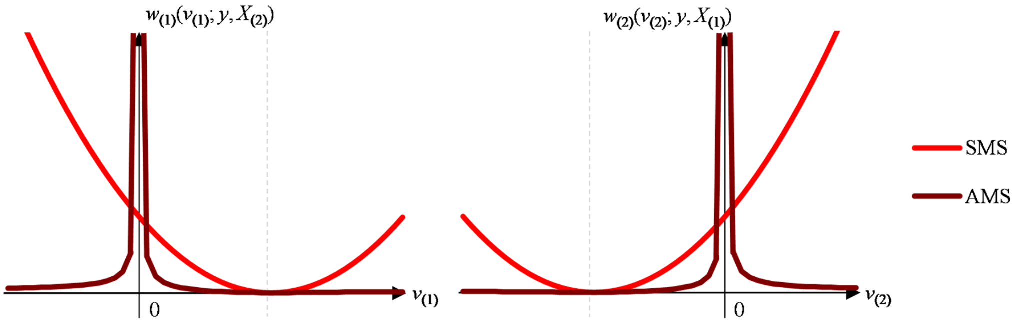

Fig. 4. The weight functions $w_{(1)}(v_{(1)}; y, X_{(2)})$ and $w_{(2)}(v_{(2)}; y, X_{(1)})$ of SMS and AMS estimations

The differences between the weight functions of both methods are significant. In AMS estimation, the weight functions have the highest values for minimal absolute values of $v_{(1)}$ or $v_{(2)}$, respectively. The values obtained for bigger absolute errors are much smaller. In the case of SMS estimation, the range of the weight functions for small and high absolute $v_{(1)}$ (or $v_{(2)}$) is much smaller. One can say that AMS estimation "prefers" the observations with small errors.

The elemental analysis of the influence or weight functions gives theoretical clues about method sensitivity to outlying observations. This method feature can also be described empirically by applying the empirical influence function (*EIF*). In the paper context, one can use *EIF* in the following form (Rousseeuw and Verboven, 2002)

$$EIF(x) = T_n\left(y_1, y_2, y_3, \ldots, y_{n-1}, x\right) \tag{11}$$

where: T_n – the estimator under examination. The other forms, which might be advisable in other problems, can be found in (e.g., Duchnowski and Wyszkowska, 2018, 2022). Let SMS and AMS estimates be examined by assuming that $X_{(1)} = 0$ and $X_{(2)} = 5$, and the observation errors are normally distributed with the expected value equal to 0 and the standard deviation $\sigma = 1$. Since the sensitivity to outlying observations depends on the size of the observation set, here, we consider four variants that differ from each other in the number and location of observations (n_1 – number of observations in the first group of observations, n_2 – number of observations in the second group of observations):

– Variant I: $n = 10,\ n_1 = 5,\ n_2 = 5$;
– Variant II: $n = 20,\ n_1 = 10,\ n_2 = 10$;
– Variant III: $n = 20,\ n_1 = 15,\ n_2 = 5$;
– Variant IV: $n = 13,\ n_1 = 3,\ n_2 = 10$.

Empirical influence functions for these variants are presented in Figure 5, which shows how estimates change for the growing value of x. SMS estimates are affected by changing x in all variants. The correct and accepted results are obtained only when $x \in \langle -2, 7\rangle$. It means that the method is locally robust (is not sensitive to the observations between the main observation groups (Duchnowski and Wisniewski, 2020)). AMS estimation provides correct results in the first three variants. The changing value of x influences the results only in a small way. One can say that AMS estimation is robust against outlying observations for $x \in \langle -15, 20 \rangle$. Considering the previous publications (Wyszkowska and Duchnowski, 2019; Duchnowski, 2021), one can say that for more significant values of x, results of AMS estimation would also be badly affected and unacceptable. In the last variant, the disproportion between the numbers of observations in both groups is higher, and the first observation group includes only a few observations. In such a case, AMS estimation does not bring good results for all values of x. It mainly concerns the estimates of $X_{(1)}$. The results of SMS estimation are like in the previous variants, and they seem worse than the results of AMS estimation.

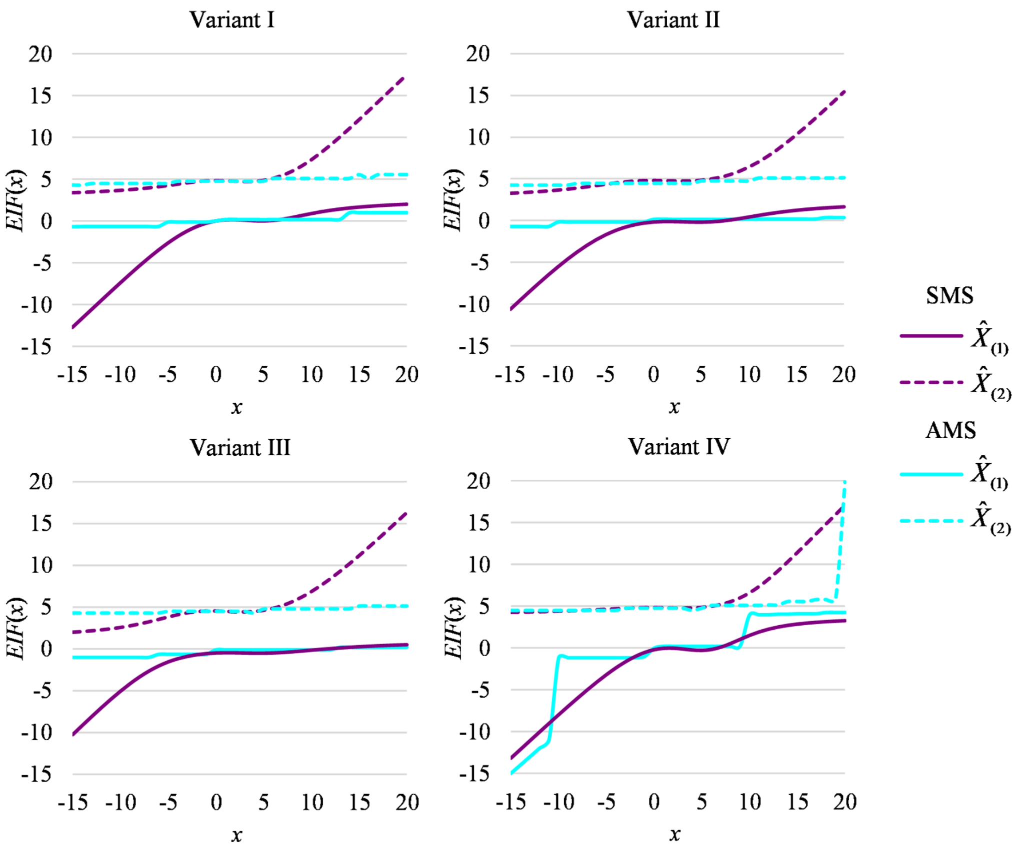

Fig. 5. Empirical influence functions $EIF(x)$

In the previous section, two different iterative processes are introduced. Let us check the length of the iterative processes by taking the assumptions of Variant I and simulating 1000 observation sets. The mean number of iterative steps for the traditional iterative process and SMS estimation is about 4, whereas, for the parallel iterative process and AMS estimation, it is about 20. It is also interesting that the maximum number of iterative steps is 5 for SMS estimation and more than 600 for AMS estimation. Such a discrepancy stems from the general features of the methods in question, especially from the differences in values of the weight functions. Generally, the iterative process of AMS estimation is much longer than the traditional iterative process. The processing of the observation sets presented here is four times longer for AMS estimation than for SMS estimation. In the case of the single observation sets, the time difference has no practical significance. A more detailed examination of both iterative processes can be found in (Wyszkowska and Duchnowski, 2020).

From a practical point of view, the accuracy of the estimates seems quite essential. By applying the Monte Carlo simulations, we can compare the root-mean-square deviations (*RMSD*) of both variants of M_{split} estimation. Figure 6 presents their *RMSD*s obtained for the set containing ten observations ($n = 10$) and assuming $X_{(1)} = 0$ and $X_{(2)} = 5$ in two variants: for different values of the observation standard deviation σ and the equal number of observations in aggregations, namely $n_1 = n_2 = 5$ (panel a); for a different number of observations in the aggregations, namely for $n = 10 = n_1 + n_2$ (panel b).

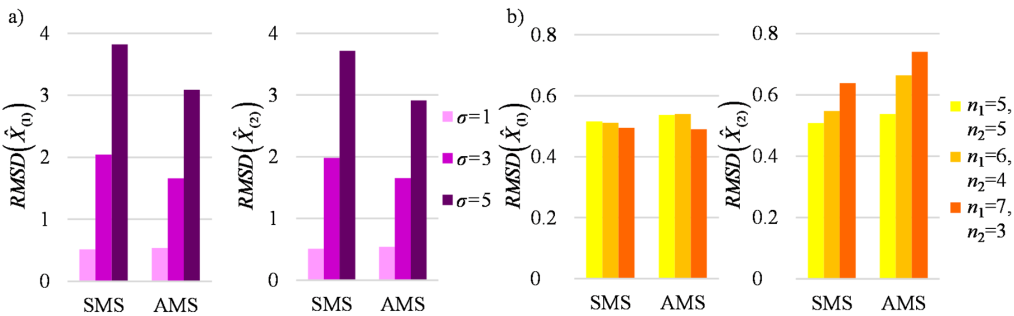

Fig. 6. *RMSD*s of SMS and AMS estimations

The differences between the accuracies of both M_{split} estimates are not high. Panel a) of Figure 6 shows that with the growing standard deviation of observations, the accuracy of all estimates decreases. One can conclude that AMS estimation is more accurate than SMS estimation for higher values of observation standard deviation. On the other hand, when the observation aggregations differ in the number of observations, SMS estimation provides a little bit more accurate results (Panel b) of Fig. 6).

4. Discussion

The paper presents a basic comparison of two variants of M_{split} estimation. The general assumption concerning an observation set for which the method in question should be applied is common for both variants. The main differences between the variants stem from the fundamental assumptions concerning the respective objective functions, hence also the influence and weight functions. The dissimilarity of the influence functions determines different sensitivity to observations with moderate errors. AMS estimation holds advantages over SMS estimation in such a context. Such a conclusion is also confirmed by empirical influence functions that show both methods' responses to observations of the mentioned type. The forms of the weight functions also determine the application of different iterative processes. The parallel iterative process is dedicated to AMS estimation (the traditional process is not applicable in such a context). SMS estimation might use both iterative processes; however, the traditional iterative process is recommended (Wyszkowska and Duchnowski, 2020). The accuracy of both estimation methods seems similar; however, AMS estimation is advisable when the observation accuracy is lower. A slight disadvantage of AMS estimation is its sensitivity to higher discrepancies in the numbers of observations in aggregations. The number of observations in many modern measurement techniques is high; thus, such a disadvantage seems meaningless.

The summary of similarities and differences between SMS and AMS estimations is presented in Table 1.

Table 1. Similarities and differences between SMS and AMS estimations

Similarities between SMS and AMS estimations	
Applicable when an observation set is a mixture of realizations of two random variables	
The same split functional model $\mathbf{y} = \mathbf{AX}_{(l)} + \mathbf{v}_{(l)}$	
They cannot be classified as robust against outliers in general	
Similar accuracy	
Differences	
SMS estimation	AMS estimation
Different objective, influence, and weight functions	
Traditional iterative process	Parallel iterative process
Shorter iterative process – less time-consuming	Longer iterative process – more time-consuming
More sensitive to gross errors	Less sensitive to gross errors of moderate magnitude
Less sensitive to different numbers of observations in aggregations	More sensitive to different numbers of observations in aggregations
Accuracy decreases faster with growing observation standard deviation	Accuracy decreases slower with growing observation standard deviation

5. Conclusions

The Introduction section shows that M_{split} estimation might be applied in processing geodetic observation sets in many practical problems. It especially concerns new measurement methods which provide big data sets, like, light detection and ranging systems (LiDAR) or global navigation satellite systems (GNSS). The respective observation sets are usually heterogeneous, stemming from measurements of different objects or occurrence of outlying observations; hence, they are predestinated to be processed by the method in question. The choice between SMS or AMS estimation should depend on the features of those variants presented in that paper. The advantage of AMS estimation is less sensitivity to outlying observations. On the other hand, SMS estimation is less time-consuming and less sensitive to discrepancies between the number of observations in main observation aggregations. In such a context, the choice between the variants should also consider the character of the observation set, e.g., a possible (expected) share of the outlying observations, a number of aggregations, etc.

Summing up the pros and cons of both M_{split} estimation variants, we suggest the application of AMS estimation, especially when we suspect that the observation set includes outlying observations. However, SMS estimation is recommended when there is a high discrepancy between the numbers of observations in aggregations.

M_{split} estimation is still a developing method, and it will allow us to create new modifications in the future. The characteristic of observations from modern measurement systems inclines us to consider new M_{split} estimation variants that would be robust against outlying observations. Such robustness should stem from *a priori* assumptions and guarantee the estimate robustness in a more general and traditional way than AMS estimation. Another future research may focus on developing AMS estimation for three (or more) functional models. The general idea of such a variant is known; however, it requires an investigation of a scheme of iterative process or example application computations.

Author contributions

Conceptualization: P.W. and R.D.; data curation: P.W.; formal analysis: R.D.; investigation: P.W. and R.D.; methodology: P.W. and R.D.; software: P.W.; supervision: R.D.; visualization: P.W.; writing – original draft: P.W. and R.D.; writing – review and editing: P.W. and R.D.

Data availability statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

This research was funded by the Department of Geodesy, University of Warmia and Mazury in Olsztyn, Poland, statutory research no. 29.610.001-110.

References

Baselga, S., Klein, I., Suraci, S.S. et al. (2021). Global optimization of redescending robust estimators. *Math. Probl. Eng.*, 1–13. DOI: 10.1155/2021/9929892.

Czaplewski, K., Waz, M., and Zienkiewicz, M.H. (2019). A novel approach of using selected unconventional geodesic methods of estimation on VTS areas. *Mar. Geodesy*, 42, 447–468. DOI: 10.1080/01490419.2019.1645769.

Duchnowski, R. (2021). Vertical displacement analysis based on application of univariate model for several chosen estimation methods. In: Proceedings of FIG e-Working Week 2021, 20-25 June 2021 (pp. 1–13), Netherlands.

Duchnowski, R., and Wyszkowska, P. (2018). Empirical influence functions of Hodges-Lehmann weighted estimates applied in deformation analysis. In: Proceedings of 2018 Baltic Geodetic Congress (BGC Geomatics), 21-23 June 2018 (pp. 169–173). Olsztyn, Poland. DOI: 10.1109/BGC-Geomatics.2018.00038.

Duchnowski, R., and Wisniewski, Z. (2020). Robustness of squared $M_{split(q)}$ estimation: Empirical analyses. *Studia Geophys. et Geod.*, 64, 153–171. DOI: 10.1007/s11200-019-0356-y.

Duchnowski, R., and Wyszkowska, P. (2022). Empirical influence functions and their non-standard applications. *J.Appl. Geod.*, 16, 9–23. DOI: 10.1515/jag-2021-0012.

Guo, Y., Li, Z., He, H. et al. (2020). A squared M_{split} similarity transformation method for stable points selection of deformation monitoring network. *Acta Geod. et Cartogr. Sin.*, 49, 1419–1429. DOI: 10.11947/j.AGCS.2020.20200023.

Janicka, J., and Rapinski, J. (2013). M_{split} transformation of coordinates. *Surv. Rev.*, 45, 269–274. DOI: 10.1179/003962613X13726661625708.

Janicka, J., Rapinski, J., Blaszczak-Bąk, W. et al. (2020). Application of the M_{split} estimation method in the detection and dimensioning of the displacement of adjacent planes. *Remote Sens.*, 12, 3203. DOI: 10.3390/rs12193203.

Li, J., Wang, A., and Xinyuan, W. (2013). M_{split} estimate the relationship between LS and its application in gross error detection. *Mine Surv.*, 2, 57–59. DOI: 10.3969/j.issn.1001-358X.2013.02.20.

Nowel, K. (2019). Squared $M_{split(q)}$ S-transformation of control network deformations. *J. Geod.*, 93, 1025–1044. DOI: 10.1007/s00190-018-1221-4.

Rousseeuw, P.J., and Verboven, S. (2002). Robust estimation in very small samples. *Comput. Stat. Data Anal.*, 40, 741–758. DOI: 10.1016/S0167-9473(02)00078-6.

Wisniewski, Z. (2009). Estimation of parameters in a split functional model of geodetic observations (M_{split} estimation). *J. Geod.*, 83, 105–120. DOI: 10.1007/s00190-008-0241-x.

Wisniewski, Z. (2010). $M_{split(q)}$ estimation: estimation of parameters in a multi split functional model of geodetic observations. *J. Geod.*, 84, 355–372. DOI: 10.1007/s00190-010-0373-7.

Wyszkowska, P., and Duchnowski, R. (2019). M_{split} estimation based on L_1 norm condition. *J. Surv. Eng.*, 145, 04019006. DOI: 10.1061/(ASCE)SU.1943-5428.0000286.

Wyszkowska, P., and Duchnowski, R. (2020). Iterative process of $M_{split(q)}$ estimation. *J. Surv. Eng.*, 146, 06020002. DOI: 10.1061/(ASCE)SU.1943-5428.0000318.

Wyszkowska, P., Duchnowski, R., and Dumalski, A. (2021). Determination of terrain profile from TLS data by applying M_{split} estimation. *Remote Sens.*, 13, 31. DOI: 10.1061/(ASCE)SU.1943-5428.0000318.

Zienkiewicz, M.H. (2022). Identification of unstable reference points and estimation of displacements using squared M_{split} estimation. *Measurement*, 195, 111029. DOI: 10.1016/j.measurement.2022.111029.

Zienkiewicz, M.H., and Baryla, R. (2020). Determination of an adequate number of competitive functional models in the square $M_{split(q)}$ estimation with the use of a modified Baarda's approach. *Surv. Rev.*, 52, 13–23. DOI: 10.1080/00396265.2018.1507361.