Method to detect malfunctioning traffic count stations

By: Juan de Oña, Penélope Gómez and Enrique Mérida-Casermeiro

Juan de Oña, Penélope Gómez and Enrique Mérida-Casermeiro (2012) *Method to detect malfunctioning traffic count stations*. **IET Intelligent Transport Systems, 6(4), 364-371.**

Direct access to the published version: http://dx.doi.org/10.1049/iet-its.2011.0102

Method to Detect Malfunctioning Traffic Count Stations

Juan de Oña[a], Penélope Gómez[a] and E. Mérida-Casermeiro[b]

[a]Transportation Engineering, Dept. of Civil Engineering. University of Granada, Granada, Spain

[b]School of Computer Science, Dept. of Applied Mathematics. University of Málaga, Málaga, Spain.

E-mail:jdona@ugr.es

Abstract

This paper presents a method for the automatic detection of malfunctioning Traffic Count Stations (TCS) in a transport system. First, double linear optimization is used to detect inadmissible errors in the recordings of a series of TCS and next, the TCS that are most likely to be failing are identified. The method has been applied to an urban traffic network showing success rates up to 93% in identifying the TCS that are failing.

Keywords: Traffic count errors, Linear optimization, Transport planning, Data consistency.

1 Introduction

In traffic operation management and control field, accurate estimates of the density of vehicle flow density in road networks are very important. Information on traffic density may be ascertained from

gross counts taken by loop detectors and other detection devices. However, the counts available may be incorrect due to an improper collection process and errors.

When counting the number of vehicles that travel on a road, two types of errors can be committed:

- **Admissible:** In general, admissible errors are the errors that are within the measuring device's tolerance and, therefore, they depend on the precision defined for each device by the manufacturer. For instance, if the manufacturer of the detectors in the traffic counts stations (TCS) indicates 3% reliability, it means that if one of the measurements is $x^{obs} = 784$, the real value $x^* \in [784(1 - 0.03), 784(1 + 0.03)]$. In practice, the admissible boundary of error tends to be somewhat higher, since margins tend to increase with use and over time.

- **Inadmissible:** These are errors that not only give erroneous information, but also invalidate the work done. They can be due to detector malfunctioning (failure to record passing vehicles, constant recording of non-existent vehicles, always counting an arbitrary number, etc.) or to failure on the part of the person who handles the detector (failure to set the counter to zero, erroneous readings, etc.)

Let consider an intersection with two in (x_1 and x_2) and three out movements (x_3, x_4 and x_5) the principle of flow conservation should verify that:

$$x_1 + x_2 = x_3 + x_4 + x_5$$

Let the measurements be taken and the following is obtained:

- **Case 1:** $x_1^{obs} = 800$, $x_2^{obs} = 1200$, $x_3^{obs} = 600$, $x_4^{obs} = 700$ and $x_5^{obs} = 740$.

  It is found that the above-mentioned condition is not verified, since: $x_1^{obs} + x_2^{obs} = 2000$, whereas $x_3^{obs} + x_4^{obs} + x_5^{obs} = 2040$. Are the measurements reliable and therefore they can provide relevant

information? Or are they indicating that a detector is failing and giving inadmissible measurements? In this case, and assuming that 3% of errors is admissible, we can indicate the existence of a set of values for the measurements that verifies the condition of conservation flow and is within the tolerance range: $x_1^{adj} = 808$, $x_2^{adj} = 1212$, $x_3^{adj} = 594$, $x_4^{adj} = 693$ and $x_5^{adj} = 733$. Therefore, they should be close to the real values.

- **Case 2:** $x_1^{obs} = 800$, $x_2^{obs} = 1200$, $x_3^{obs} = 1600$, $x_4^{obs} = 700$ and $x_5^{obs} = 740$.

  It is found that the above condition is not verified either, since: $x_1 + x_2 = 2000$, whereas $x_3 + x_4 + x_5 = 3040$. However, at present no combination of x_i^{adj} values verifies flow conservation and falls within the 3% tolerance range. The inference would be that one of the measurements was erroneous and a detector must be repaired or replaced (unless there was a human error in the installation, reading or recording of the data).

A number of studies ([1], [2], [3], [4] and [5]) attempt to find a solution to Case 1 (admissible errors) to obtain adjusted data that are consistent with flow conservation laws.

For Case 2 (inadmissible errors), several approaches ([6], [7], [8] and [9]) have been attempted to resolve or diminish count errors after they have been detected, but they do not address how they can be detected.

The methods for trying to detect errors may be classified according to the consistency criterion [10]:

- *Fundamental consistency:* Data should be consistent with basic notions of traffic theory and should be physically plausible; establishes upper and lower boundaries for traffic values (e.g. negative values and vehicle volumes that exceed the road's capacity cannot be measured).

- *Network consistency:* Data should be related to measurements that are close in space and time. It

is based on flow conservation when several connected nodes in a transport network are studied. This is the type of consistency shown in the preceding example.

- *Historical consistency:* Historical observations can provide insight as to the plausibility of current data. Practice tells us that the values measured on a road are almost always given for an interval. Values outside of the interval may be plausible, but they indicate outliers, an anomaly that should alert the control service. The historical values constitute a basis for determining the boundaries of the interval in which normally consistent values must be found.

In current traffic control centres, detecting a malfunctioning count station is pseudo-automated because historical consistency marks the value interval each observation should have. If a measurement is not within that interval, an alarm is triggered, indicating a potential error in one of the TCS.

The problem arises when no historical values are available or when they exist but may indicate measurements as erroneous when they are actually correct. An incident on the network - repair work, accidents and weather issues, for instance - may alter track conditions significantly and cause outliers in the above-mentioned measurements without presupposing that the detector has failed, in fact there is a research field on this issue (among others [11], [12], [13] and [14]).

The bibliography [10] indicates several error detection techniques based solely on historical consistency. They do not take nearby detectors, that is, network consistency, into consideration. Other approach is to incorporate observations from adjacent detectors ([4]). This paper presents a method that is complementary to the existing ones, where basic consistency and network consistency are taken into consideration.

The method automatically detects a TCS that is failing, by only considering the data observed by the network detectors as input data. Once the detector that is failing has been detected, the procedure can be

repeated to see if the remaining measurements are consistent and free of errors.

This paper is organized as follows: Section 2 describes the method and the computational issues; in Section 3 the method is applied to an urban network; Section 4 discuss the effect of the model's variables on the results; and, finally, Section 5 presents the main conclusions of the paper.

2 Methodology

The method presented in this paper to detect and identify a malfunctioning detector is based on the resolution of a linear programming problem (LP). In general terms, the $\mathbb{R}^n$ region that meets certain restrictions is known as the LP's feasible region. That is what will be built for the problem posed in this paper.

2.1 Feasible region.

Let a series of measurements be taken $\{x_i^{obs}\}$ and that the tolerance indicated for each measurement is α_i. This tolerance is usually expressed as a percentage of the measured value, since it is reasonable to assume that any absolute errors incurred will be lower for small magnitudes than for larger ones, assuming the detectors function under the conditions specified by the manufacturer: $\forall i; x_i^* \in [a_i, b_i]$, where $a_i = x_i^{obs} - \alpha_i x_i^{obs}$ and $b_i = x_i^{obs} + \alpha_i x_i^{obs}$. In example 1, a 3% error was considered admissible for all the measurements, and therefore we would take $\forall i, \alpha_i = 3\%$, although in other cases a different error for each detector could be considered.

Given a set of observed values $\{x_i^{obs}\}$, $i \in \mathcal{I}$, (where $\mathcal{I}$ is a set of indexes) each with a tolerance of α_i, we define the *admissible region* as the set $\mathcal{A} \subset \mathbb{R}^n$, such that $\forall \vec{x} = \{x_i\} \in \mathcal{A}$ where the following conditions are satisfied:

1. $x_i^{obs} - \alpha_i x_i^{obs} \leq x_i \leq x_i^{obs} + \alpha_i x_i^{obs}$.

2. Vector $\vec{x}$ verifies flow conservation laws.

Attention should be paid to the fact that the cardinal of the set of observed values and the number of variables, n, do not necessarily coincide. Thus, to continue with the previous example, let the set of observed values be $x_1^{obs} = 800$, $x_2^{obs} = 1200$, $x_3^{obs} = 600$ and $x_4^{obs} = 700$, which would give the admissible region:

$$\mathcal{A} = \{\vec{x} \in \mathbb{R}^5 / 776 \leq x_1 \leq 824, 1164 \leq x_2 \leq 1236, 582 \leq x_3 \leq 618, 679 \leq x_4 \leq 721, x_5 = x_1 + x_2 - x_3 - x_4\}$$

where the first 4 intervals are obtained by $x_i = x_i^{obs} \pm \alpha_i x_i^{obs} = x_i^{obs}(1 \pm \alpha_i)$, adding the flow conservation law: $x_1 + x_2 = x_3 + x_4 + x_5$.

Theorem 1 *If all the detectors function properly, the feasible region is not empty ($\mathcal{A} \neq \emptyset$).*

Obviously, if all the detectors give admissible errors, then the true values vector, $\vec{x}^*$, belongs to the feasible region ($\vec{x}^* \in \mathcal{A}$).

Therefore, the inference is:

Corolary 1 *Si $\mathcal{A} = \emptyset$, one of the detectors is giving an inadmissible error.*

Corolary 1 provides a method for detecting incorrect measurements by taking into consideration fundamental inconsistencies and network inconsistencies. Although the converse theorem is not true, that is:

A detector may produce an inadmissible error, but the remaining detectors' margins permit admissible values and, therefore, $\mathcal{A} \neq \emptyset$. In practice, this means that although there exists out of range measures,

it is possible even to obtain a consistent vector. So, if a detector is severely malfunctioning it will be impossible to generate consistent traffic counts.

We should also consider that if there are several vectors in $\mathcal{A}$, ($\mathcal{A} \neq \emptyset$), some are more plausible than others, insofar as they are closer to the observed values. So, for a vector $\vec{x} \in \mathcal{A}$ we can associate another vector $\vec{h} = \{h_i\}$ such that the verisimilitude of the i-th component is:

$$h_i^* = 1 - \frac{|x_i - x_i^{obs}|}{\alpha_i |x_i^{obs}|}, \qquad h_i = \max\{0, h_i^*\} \tag{1}$$

Figure 1 shows the verisimilitude of assigning a value x_i when x_i^{obs} with reliability α_i has been observed.

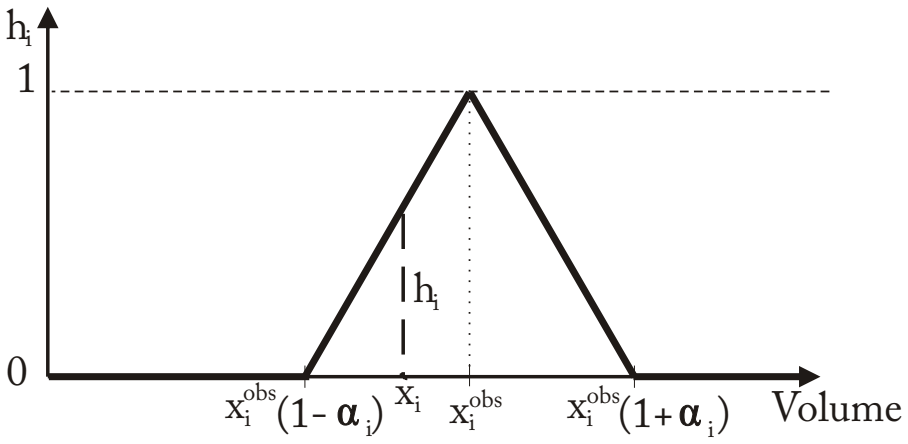


Figure 1. Verisimilitude function for a single observation.

For the sake of simplicity, a triangular shape function has been chosen since the function shape is not an important issue, since the aim is to check if the adjusted value is in or out of the feasible region and simplicity of linear decay allows it to be solved by linear programming. However, other polygonal function could be used, as it is stated in [1].

Assuming $x_i^{obs} > 0, \forall i \in \mathcal{I}$ and making the relevant transformations in equation 1, finding out whether

an admissible set of values exists becomes a problem of finding out whether a solution to the linear optimization problem exists:

Problem 1

$$\textit{Maximize: } \sum_{i \in \mathcal{I}} h_i$$

$$\textit{Subject to: } \begin{cases} 0 \leq h_i \leq 1, \ x_i \geq 0 \\ x_i + \alpha_i x_i h_i \leq x_i^{obs}(\alpha_i + 1) \\ -x_i + \alpha_i x_i h_i \leq x_i^{obs}(\alpha_i - 1) \\ M\vec{x} = 0 \end{cases}$$

where x_i, h_i and h are the variables that can be considered adjusted (consistent) values, variable verisimilitude and minimal verisimilitude, respectively, and where the flow conservation laws are represented by the homogeneous linear system $M\vec{x} = 0$. Thus, for case 1 with the single conservation law: $x_1 + x_2 - x_3 - x_4 - x_5 = 0$, the matrix $M = (1, 1, -1, -1, -1)$. In general, the matrix M will have as many rows as existing flow conservation equations. Very different target functions could have been selected for this task, but this will also serve the second aim of this paper: To determine which detector is producing erroneous values. The benefit of transforming the problem into a linear programming problem is being able to count on multiple and optimized routines for the solution. See [15]. It is easy to amend the above method by considering different margins to the right and to the left of the observed values, i.e. $x_i^{obs} \in (x_i^{obs} - \alpha_i^L x_i^{obs}, x_i^{obs} - \alpha_i^R x_i^{obs})$.

2.2 Detection of inadmissible measurements

Let the problem of resolving linear programming 1 in section 2.1 be posed and that there is no solution, since $\mathcal{A} = \emptyset$. We would be in the case of Corolary 1, which indicates that one of the measurements is

inadmissible. Unfeasible should not be confused with outliers, since the latter may be correct and due to traffic anomalies (an accident, repairs, etc.) but consistent with flow conservation laws. To detect an incorrect measurement, we relax the manufacturer's α_i margins, multiplying them by a constant $K \gg 0$ so the new linear optimization problem will have a non-empty admissible region. That is:

Problem 2

$$\textit{Maximize: } \sum_{i \in \mathcal{I}} h_i$$

$$\textit{Subjec to: } \begin{cases} 0 \leq h_i \leq 1, \ x_i \geq 0 \\ x_i + K\alpha_i x_i h_i \leq x_i^{obs}(K\alpha_i + 1) \\ -x_i + K\alpha_i x_i h_i \leq x_i^{obs}(K\alpha_i - 1) \\ M\vec{x} = 0 \end{cases}$$

It is known that one property of the 'maxsum' objective function is that it gives high values to most variables at the expense of giving low values to a few variables [16]. In this case, its effect is to assign values very close to the observed ones (high verisimilitude) to the detriment of assigning very distant values to a few (low verisimilitude). The measurement that produces $h = \min\{h_i\}$ in problem 2 will be proposed as inadmissible. We can always obtain a K that is large enough to make $\mathcal{A} \neq \emptyset$; since its effect is to increase the variables' admissible margin. In an extreme case, any measurement x_i would fit into the $(x_i^{obs} \pm K\alpha_i x_i^{obs})$ interval. It could be assumed that selecting K would modify the solution obtained, but the following theorem shows that such is not the case:

Theorem 2 *If the problem 2 is solved by using two different values for K ($K_1 \neq K_2$), performing both feasible solutions, then optimum solutions for K_1 and K_2 verify:*

1. *The optimum vector $\vec{x}^{(1)^*}$ for K_1 is also optimum vector for K_2: $\vec{x}^{(2)^*} = \vec{x}^{(1)^*}$*

2. *The index of observation with minimum value for h_i is the same for both constants:* $\arg\min_i\{h_i^{(1)}\} = \arg\min_i\{h_i^{(2)}\}$

Proof is given in Appendix.

2.3 Proposed algorithm

From previous considerations, next algorithm is proposed.

Algorithm 1 (Erroneous sensor detector)

1) Read values for α_i^L, α_i^R y x_i^{obs}

2) Represent the flow conservation laws by matrix M.

3) Repeat until all $h_i^* > 0$:

a) Represent all inequalities by matrix A and vector $\vec{b}$:

$$x_i + \alpha_i^R x_i h_i^* \leq x_i^{obs}(\alpha_i^R + 1)$$

$$-x_i + \alpha_i^L x_i h_i^* \leq x_i^{obs}(\alpha_i^L - 1)$$

b) Express restrictions $x_i \geq 0$.

c) Solve LP with the target function Maximize $\sum_i h_i^*$.

d) If all $h_i^* \geq 0$, go to step 4, else:

d1) Evaluate $h^* = \min_i h_i^*$, $K = \frac{1-h^*}{0.9}$

d2) Replace $\alpha_i^R \leftarrow K\alpha_i^R$ and $\alpha_i^L \leftarrow K\alpha_i^L$, into A and $\vec{b}$ (step 2a).

d3) Solve LP with the objective function Maximize $\sum_i h_i^*$.

d4) The index k that produces $h_k^* = \min_i h_i^*$ is obtained.

d5) Observation x_k^{obs} is ellipsed and considered as erroneous.

d6) Return to step 3). With initial values of α_i^R and α_i^L, but the ellipsed one.

4) Finish. (Ellipsed observations are considered as inadmissible ones.)

The algorithm is focused on detecting inadmissible observations from the network consistency viewpoint. However, it is easy to incorporate any available additional information. For instance, by changing the upper bound of any variable (adding the restriction $x_i \leq U_i$ in step 3b), or by changing the lower bound of any variable, which by default is 0 ($x_i \geq L_i$), etc. This fact allows making it suitable to perform fundamental consistency, generally expressed by bounds.

This method could be complementary to standard pre-process that analyzes historical consistency [10]. That is, observed variables must be into a real interval, in other case the observation is considered an outliers. An outliers must be analyzed separately since it can be produced by anomalous traffic, but be correct.

Perhaps the algorithm 1 was only executed to verify that the detectors were working properly, but it is usually part of the study on a region's traffic. In such case, the next step would be to obtain the adjusted data, that is, the consistent data that most closely resembles the observed data. Any data deemed inadmissible during the pre-process will have been eliminated from the observed data using one of the procedures suggested by other authors ([1], [2], [3] and [4]).

3 Application to an urban network

3.1 Road network data

The method is applied to the urban network shows in Fig. 2.

The network has seven intersections, of which four have twelve movements (intersections D, E, F and G), two have six movements (A and C), while the last one has five potential movements (intersection B). So, in total, there are 86 unknown variables. Since it is impossible to guarantee that a set of true data

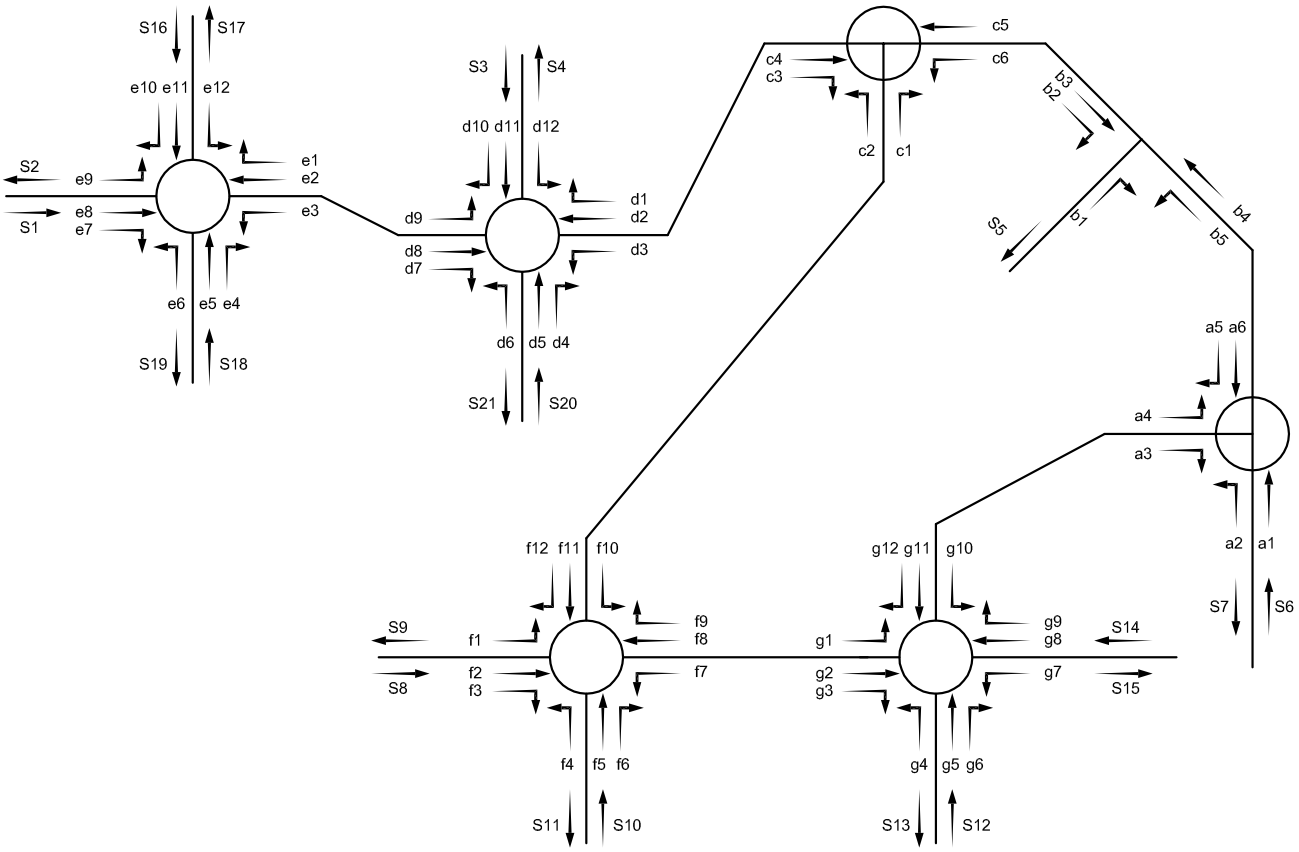


Figure 2. Example of an urban network

will always be available, the initial set of data will be a set of consistent data that is very close to the observed data.

Consider the situation shown in Fig. 2, in which consistent true data are available (Theoretical Values - TV), where the data that comply with flow conservation in the traffic network concerned is deemed to be consistent. In other words, the sum of incoming vehicles is equal to the sum of outgoing vehicles at any network intersection.

This consistent data-base is used to randomly deform values, by an uniform distribution, with a tolerance of $\pm 3\%$, which is the tolerance shown by the count stations most commonly used in urban networks [17]. This is not deterministic, however, because if the detector was of another type or had a different

tolerance, a value other than $\pm 3\%$ could be considered. The model allows a different α for each observed datum to be defined (several types of detectors with different tolerances). As shown in section 2, it even permits the definition of asymmetric feasible regions.

Having obtained a randomly distorted data base within the above-mentioned tolerance, it could then be considered as the data that would be obtained in an ideal counting campaign in which all 86 potential movements would be measured. Therefore, it could be taken as the series of observed data in an urban network (Observed Values - OV). In this case, the values would not be consistent, according to the above definition (the sum of incoming vehicles would not be equal to the sum of outgoing vehicles).

The fact that a base of consistent true data is used and subsequently randomly distorted permits a comparison between the results obtained and real life, and verification of the goodness of the method proposed.

3.2 Results

OV obtained randomly from TV with a tolerance of $\pm 3\%$ is used to verify the goodness of the model. Next, a datum is randomly selected and distorted to simulate a detector error that exceeds the error specified by the manufacturer or, in other words, a deviation from the detector's allowed tolerance. Specifically, deviations of 75%, 50%, 25% and 10% from OV are simulated.

This deformation gives an initial data base for each example generated (each of which contains an erroneous datum). For each one of the 4 deviations, 500 examples are randomly generated from OV. In all, 2,000 examples are executed. Table 1 shows the results for the random examples.

Column 1 in Table 1 shows the simulated error in a randomly selected measurement apparatus. Columns 2-3 show the number of times an error is detected in all the random samples. That is, the

Percentage simulated error	Error is detected		Error is pointed out and gets by h_{min}		Error is pointed out and gets by 2^{nd} h_{min}		Error is pointed out in total
	Number of times (A)	Proportion A/500	Number of times (B)	Proportion B/500	Number of times (C)	Proportion C/500	Success proportion (B+C)/500
75	491	0,982	443	0,886	23	0,046	0,932
50	486	0,972	386	0,772	37	0,074	0,846
25	438	0,876	269	0,538	27	0,054	0,592
10	266	0,532	137	0,274	24	0,048	0,322

Table 1. Obtained results with a simulated error of 75, 50, 25 and 10%.

number of times $\mathcal{A} = \emptyset$ is obtained applying theorem 1. Column 2 points out the number of times $\mathcal{A} = \emptyset$ is obtained for the random examples, which is when the adjusted value lies outside of the detector's allowed tolerance, and outside of the set boundaries of the feasible region. This indicates that a detector is giving a value that is higher than the allowed deviation, which in turn means that a detector is failing. Column 3 shows the same thing in relative terms.

By increasing all α_i from 0.03 (see d1) in algorithm 1) in a two steps process, the feasible region is extended in order to allow $\mathcal{A} = \emptyset$ to be found for every i. This value was selected because it produces all $h_i^* > 0$ at next step, as can be deduced from equation 2 in proof of Theorem 2.

Table 1, column 4 shows the number of times the index i that produces $h = \min h_i$, coincides with the failing TCS. Columns 6-7 show the number of times (and proportion, respectively) in which the failing TCS is the one that shows the second lowest value. So, when a TCS perform a 75% of error, it coincides with the error obtained by the second minor value of h_i in 5% of cases.

Column 8 shows the proportion of times that the model is able to detect the failing detector (i.e. adding the number of times it detects the detector that fails, whether it is the h_i minimum or the value immediately above it). This result points out the proportion of times at which it indicates a detector that is failing, out of all the random examples. This is the model's proportion of success, and it is calculated by adding column 4 and column 6, and dividing by the total number of examples simulated. For an error of 75%, the success rate is 93%. For the remaining cases, the model finds that there is a malfunctioning detector, but it does not point it out in the first or second places.

Table 1 shows that the model's success increases in the same measure as the device's error increases and worsens as the error diminishes, and the closer it is to the measurement device's tolerance range.

If the ratio (r) is expressed as the proportion of times that an error is detected compared to the number of examples executed (Table 1, column 2), the ratio of cases in which a failing detector is detected for each simulated error can be compared.

In other words, if N random examples have been executed (in this case, $N = 500$) and A times errors have been detected (Table 1, column 2), the estimated ratio obtained experimentally is $r = \frac{A}{N}$ (Table 1, column 3). In this manner, for an error of 75%, the error is detected in 98% of cases; for an error of 50%, in 97% of cases; for an error of 25%, in 88% of cases; and finally, for a simulated error of 10%, an error is detected in 53% of cases.

4 Sensibility analysis to different variables

First, the effect of the situation of the failing traffic counts will be analyzed. Second, what happens when certain points of the network have not been counted?. Finally, the sensitivity to the number of not counted data in the network will be analyzed (with approximately 50% more and 50% less points not

Percentage simulated error	Error is detected		Error is pointed out and gets by h_{min}		Error is pointed out and gets by 2^{nd} h_{min}	
	center	edge	center	edge	center	edge
75	492	500	426	500	24	0
50	474	500	340	486	49	6
25	419	500	227	428	32	25
10	205	466	90	249	11	45

Table 2. Results for center and edge detectors with a simulated error of 75, 50, 25 and 10%.

counted).

4.1 Effect of the situation of the failing detector

How does the sensitivity depend on which detector is malfunctioning? In order to analyze the method's sensitivity to the detector position, a selective choosing of the malfunctioning detector has been made. At first stage, for each scenario, the model was forced to choose an edge detector, (S_1, S_2,..., S_{21}, or b_1 in Figure 2), and at second stage the central ones (the remaining detectors) have been chosen to be failing. Table 2 shows the results.

The method detects an error on the edge of the network better than when the detector is situated in the center. This is logical due to the following reason: when an edge detector is getting an inadmissible error, while the rest adjacent measurements are corrects, must significantly modify its value in order to reach network consistency. That is because a small amount of adjacent detectors exists which can be modified within the margin established by the feasible region. On the other hand, a major modification of these adjacent detectors makes the constraints able to be affected; therefore the $\sum_{\mathbf{i}} \mathbf{h_i}$ is reduced. The

target function forces to modify the one that is giving an erroneous measurement.

While, for center detectors, the measurements are linked to more variables that can be modified within the feasible region. So, for an inadmissible small error (around 10%) is easier to count on the adjacent values margin and move all of them, in order to get all measures within its feasible region, than a big change in the malfunctioning detector.

4.2 Effect of points that are not counted

In this subsection the effect of movements that have not been counted is analyzed.

Presumably, the network in Fig. 2 shows seven movements that have not been counted (movements c_2, c_3, c_4, d_9, d_{10}, d_{11} and d_{12}). This implies around 8% of all the movements in the network. This percentage is considered to be normal in counting campaigns in a traffic network [18]. A case consisting of 500 random examples is simulated below, in which the number of not measured movements is increased 50% (10 not measured movements), followed by a case in which the number of not measured points is diminished in 50% (4 not measured movements).

Table 3 and Figure 3 show a comparison between the results obtained in the study with 4 hypotheses (all measured data, 4, 7, and 10 not measured data). In Figure 3 the x-axis represents the simulated distortions for the measurement device and the y-axis represents the proportion of times the error is detected.

Taking column 3 ($A/500$) in Tables 1 and 3 into consideration, a comparison can be made about the number of times an error is detected in each case. Table 3 shows that the ratio of errors detected for the simulated scenarios gradually diminish when there is less measured data available (i.e. less information).

From Figure 3, Tables 1 and 3, it is possible to analyze the model's sensitivity to the number of not

Percentage simulated error	Error is detected		Error is pointed out and gets by h_{min}		Error is pointed out and gets by 2^{nd} h_{min}		Error is pointed out in total
	Number of times (A)	Proportion A/500	Number of times (B)	Proportion B/500	Number of times (C)	Proportion C/500	Success proportion (B+C)/500
4 not measured movements							
75	466	0,932	397	0,794	27	0,054	0,848
50	447	0,894	353	0,706	36	0,072	0,778
25	391	0,782	234	0,468	38	0,076	0,544
10	242	0,484	95	0,190	29	0,058	0,248
7 not measured movements							
75	449	0,898	374	0,748	30	0,060	0,808
50	425	0,850	315	0,630	42	0,084	0,714
25	351	0,702	201	0,402	28	0,056	0,458
10	226	0,452	100	0,200	30	0,060	0,260
10 not measured movements							
75	418	0,836	360	0,720	22	0,044	0,764
50	372	0,744	282	0,564	35	0,070	0,634
25	338	0,676	211	0,422	29	0,058	0,480
10	203	0,406	87	0,174	28	0,056	0,230

Table 3. Results with a simulated error of 75, 50, 25 and 10% with four, seven and ten not measured movements.

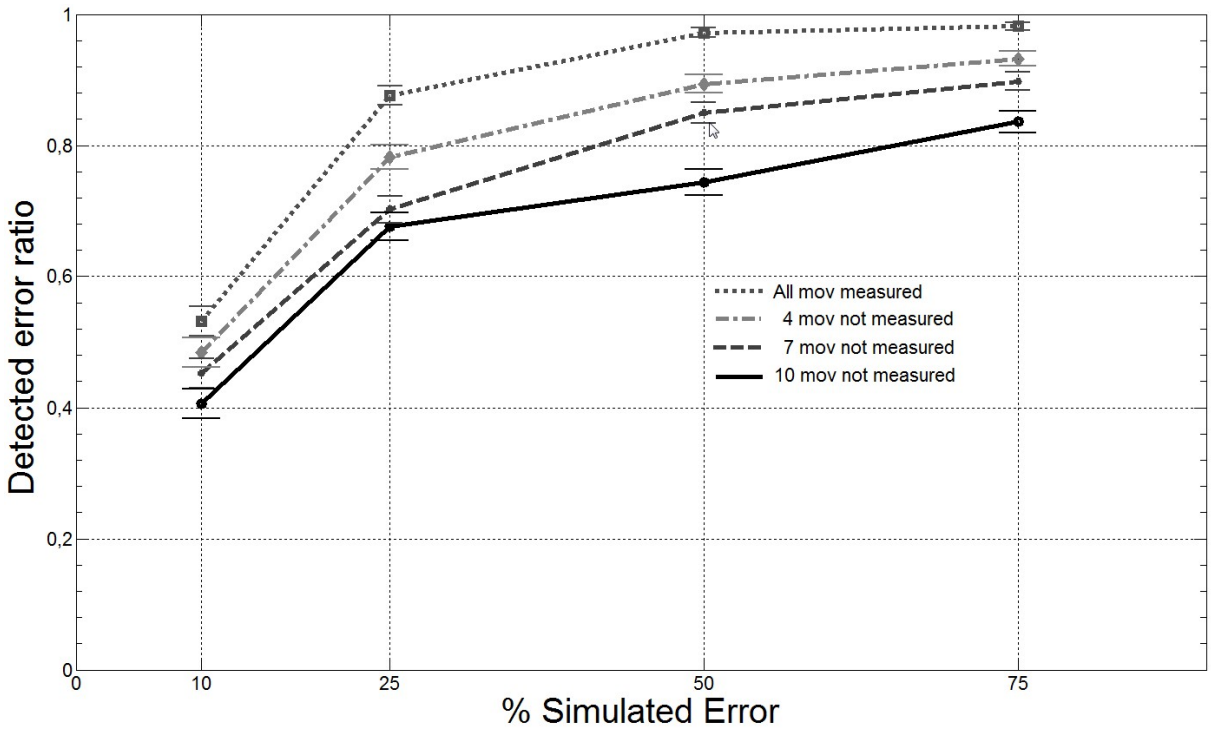


Figure 3. Sensitivity analysis of success versus increase in the number of data not measured.

measured movements in a case where all the data from all the TCS (i.e. all measured data) is available. The x-axis represents the simulated percentage of the device error (10, 25, 50 and 75%) and the y-axis data shows the percentage of success for every case, in comparison with the one in which all the data are measured. In the event that a 75% error occurs in a detector, for instance, the chart will show that the model presented in this paper is 93% successful if 4 network data are not measured, 90% if 7 network data are not measured, and 84% if 10 data are not measured.

Thus, the conclusion would be that the model gives good success results even when the number of not measured data increases, although, obviously, when more data is available, the more it improves.

Percentage of simulated error	All movements are measured		4 not measured movements		7 not measured movements		10 not measured movements	
	Proportion	σ	Proportion	σ	Proportion	σ	Proportion	σ
75	0.982	0.006	0.932	0.011	0.898	0.014	0.836	0.017
50	0.972	0.007	0.894	0.014	0.850	0.016	0.744	0.020
25	0.876	0.015	0.782	0.018	0.702	0.020	0.676	0.021
10	0.532	0.022	0.484	0.022	0.452	0.022	0.406	0.022

Table 4. Ratios calculated for every scenario providing the standard deviation.

4.3 Combined effect of the size of the error and the number of points that are not counted

Figure 3 shows the ascending trend of the ratio when a detector's error increases in all the hypotheses. The trend is even more pronounced when it moves from an error close to the detector's tolerance range (such as 10%) to around 25%, after which the detector's behaviour is asymptotic, reaching an error ratio within the range 0.9-1 for the biggest device error simulated. In other words, when the error exceeds the threshold at around 25%, it can be asserted that the model succeeds in around 90% of the cases.

In figure 3 the $1-\sigma$ errors bars have been included in order to show conclusions do not owe to random.

Table 4 showed the ratios (or proportion of success, p_i) at which error is detected in every scenario. To demonstrate that the model's proportion of success increases when more data are measured ($p_{i+1} < p_i$) and that the observed results are not due to chance, a hypotheses of proportional difference was tested at a significance level of 5%, taking $N_{i+1} = N_i = 500$ ([19] and [20]).

Three statistical tests were conducted to compare the three hypotheses in groups of two. That is, firstly hypothesis of all movement measured was tested versus 4 not measured movements, the case of 4 not

measured movements versus 7 not measured data, and lastly, 7 not measured data were tested versus 10 not measured data. The $Z_{exp} = \frac{p_i - p_{i+1}}{\sigma}$ is calculated and compared with the $Z_{theoretical} = 1.645$, it determines the significant region ($Z_{exp} > 1.645$). The results are given in Table 5.

It is found that $p_{i+1} < p_i$ in all cases and statistically significant results are obtained for the cases of 75, 50 and 25% error in the first and second tests, and in the third one the results are significant after the 50% error.

Therefore, it can be asserted that the success proportion improves with the number of counted data and this fact is not due to chance.

5 Summary and Conclusions

This paper presents a method for detecting inadmissible errors in TCS and identifying which device is more likely to be failing. The method is based on a double linear optimization process that can easily be solved with existing software on the market, and which we consider highly useful for practitioners.

If the method detects the existence of an inadmissible error in the TCS' measurements when the first linear optimization is used, a second optimization can be used so the method can obtain the detector that is most likely to be failing (the one that obtains the $\min_i h_i$). This facilitates to replace or fix them for obtaining adjusted data.

Four different cases of potential errors were simulated in order to identify the effects on the method (deviations of 10%, 25%, 50% and 75%). The results show that the method works better with bigger errors (75%), which are more frequent when dealing with malfunctioning detectors, than with small errors (10%), close to the TCS's tolerance (3%). For deviations of around 25% of their theoretical value, the method is 88% efficient for detecting that there is an error in the measures. The efficiency

% error	p_i	p_{i+1}	Z_{exp}
All measured vs 4 not measured movements			
75	0,982	0,932	**3,927110655**
50	0,972	0,894	**4,993847666**
25	0,876	0,782	**3,978624361**
10	0,532	0,484	1,519839919
4 vs 7 not measured movements			
75	0,932	0,898	**1,931244679**
50	0,894	0,850	**2,086907096**
25	0,782	0,702	**2,903158213**
10	0,484	0,452	1,014529379
7 vs 10 not measured movements			
75	0,898	0,836	**2,898969254**
50	0,850	0,744	**4,20341134**
25	0,702	0,676	0,888433399
10	0,452	0,406	1,471128409

Table 5. Test of hypotheses.Significant cases in bold.

in identifying a failing detector can be considered good (over 90%) when the error is over 75% of the deviation, and diminishes as the errors become smaller.

The same tolerance was considered for all the TCS (3%), but the model is versatile and allows assigning a different tolerance to each detector according to its type and level of precision.

Finally, a statistical test has been conducted to demonstrate that the increase in the number of times an error is detected when more movement counts were obtained as opposed to a gradually decreasing number of times is not due to chance. This serves to assert that the results are significant and the size of the sample selected is sufficient to corroborate the conclusions arrived at in this paper.

Usually studies perform automated data checking by comparing measured data to historical data for consistency [10]. Sometimes, however, there are no historical data and only the observed database is available. This is when the method proposed in this paper becomes a good tool for detecting errors, since the only incoming data required are the observed data, with no need for preprocessing. Actually, both approaches could be considered as complementary: it is possible to use fundamental and network consistency for detecting inadmissible errors and, historical consistency as alarm signal.

Acknowledgement

The authors appreciate the reviewers' comments and effort in order to improve the paper.

References

[1] Kikuchi, S. and Miljkovic, D.: 'Method to Preprocess Observed Traffic Data for Consistency: Application of Fuzzy Optimization Concept.' Transportation Research Record, 1999, 1679, pp. 73-80.

[2] Wall, Z. R. and Dailey, D. J.: 'Algorithm for Detecting and Correcting Errors in Archived Traffic Data.' Transportation Research Record, 2003, 1855, pp. 183-190.

[3] Vanajahshi, L. and Rillet, L. R.: 'Loop Detector Data Diagnostics Based on Conservation-of Vehicles Principle.' Transportation Research Record, 2004, 1870, pp. 162-169.

[4] De Oña, J., Gómez, P. and Mérida-Casermeiro, E.: 'Bilevel Fuzzy Optimization to Pre-process Traffic Data to Satisfy the Law of Flow Conservation.' Transportation Research Record Part C, 19 (1), pp. 29-39.

[5] Schleifer, W. and Mannle, M.: 'Online error detection through observation of traffic self-similarity', IEE Proceedings-Communications, 2001, 148 (1) pp. 38-42.

[6] Nihan, N. L. and Davis, G. A.: 'Application of prediction-error minimization and maximum likelihood to estimate intersection O-D matrices from traffic counts', Transportation Science, 1987, 23, pp. 77-90.

[7] Nihan, N. L. and Davis, G. A.: 'Recursive estimation of origin-destination matrices from input/output counts', Transportation Research B, 1987, 21, 149-163.

[8] Nihan, N. L. And Davis, G. A.: 'Application of prediction-error minimization and maximum likelihood to estimate intersection O-D matrices from traffic counts', Transportation Science, 1989, 23, pp. 77-90.

[9] Tavana, H. and Mahmassani, H.: 'Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method'. Proceeding of the 80th Annual Meeting of the Transportation Research Board, CD ROM, 2000.

[10] Lin, D.-Y., Boyles, S., Valsaraj, V. and Waller, S.T.: 'Fuzzy Reliability Assessment for Traffic Data', Journal of the Chinese Institute of Engineers, (in press), 2011.

[11] Thomas, T. and van Berkum, E.C.: 'Detection of incidents and events in urban networks', IET Intelligent Transport Systems, 2009, pp. 198-205.

[12] Zhang, H-z, Wang, J. and Zi-hui Ren, Z-h.: 'Rough Sets and FCM-Based Neuro-fuzzy Inference System for Traffic Incident Detection'. ICNC'08. Fourth International Conference on Natural Computation. 2008, 7, pp. 260-264.

[13] Srinivasan, D., Sanyal, S. and Sharma,V.: 'Freeway incident detection using hybrid fuzzy neural network', IET Intelligent Transport Systems, 2007, 1, (4), pp. 249-259.

[14] Tang, S. and Gao, H.: 'Traffic-incident detection-algorithm based on nonparametric regression', IEEE Trans. Intell. Transp.Syst., 2005, 6, (1), pp. 38-42.

[15] LINPROG. Available from: http://www.mathworks.com/help/toolbox/optim/ug/linprog.html Last accessed: March 22, 2011.

[16] Saameño Rodríguez, J.J., Guerrero García, C., Muñoz Pérez, J. and Mérida Casermeiro, E.: 'A General Model for the Undesirable Single Facility Location Problem.' Operations Research Letters, 2006, 34 (4), pp. 427-436.

[17] CEGASA. Available from: http://www.cegasatraffic.com/es/listado/b2/TomaDatos.html. Last accessed: September 30, 2010.

[18] Zhong M., Lingras P. and Sharma S.: 'Estimation of missing traffic counts using factor, genetic neural and regression techniques'. Transportation Research C, 2004, 12, pp. 139-166.

[19] Anderson, T.W. and Sclove, S. T.: 'The Statistical Analysis of Data'. The Scientific Press, USA, Second Edition 1986.

[20] Mendenhall, W. and Sincich, T.: 'Statistics for the engineering and computer sciences'. Dellen Publishing Company. San Francisco, California, USA. Second Edition 1988.

Appendix

Proof of Theorem 2

Let $\vec{x} \in \mathcal{A}$ and $\forall i$, $h_i^{(1)} = 1 - \frac{|x_i - x_i^{obs}|}{K_1 \alpha_i x_i^{obs}}$, $h_i^{(2)} = 1 - \frac{|x_i - x_i^{obs}|}{K_2 \alpha_i x_i^{obs}}$ then it is easily obtained:

$$K_1(h_i^{(1)} - 1) = K_2(h_i^{(2)} - 1) \tag{2}$$

and by naming m the number of observed variables:

$$\sum_{i \in \mathcal{I}} \{h_i^{(1)}\} - m = \sum_{i \in \mathcal{I}} \{h_i^{(1)} - 1\} = \frac{K_2}{K_1} \sum_{i \in \mathcal{I}} \{h_i^{(2)} - 1\} = \frac{K_2}{K_1} \left(\sum_{i \in \mathcal{I}} \{h_i^{(2)}\} - m \right) \tag{3}$$

From equation 3, it can be defined the monotonically increasing function: $S_1 = \frac{K_1}{K_2}(S_2 - m) + m$ between $S_1 = \sum_{i \in \mathcal{I}} h_i^{(1)}$ and $S_2 = \sum_{i \in \mathcal{I}} h_i^{(2)}$.

If we assume that $\vec{x}^{(1)^*} \in \mathcal{A}$ is the set of values that produces the optimal solution to problem 2 with K_1, producing values for target functions S_1^* and S_2^* for the constant K_2. Then, if a $\vec{x'} \neq \vec{x}^{(1)^*} \in \mathcal{A}$ exists and could provide a better solution to problem 2 with K_2 ($S_2' > S_2^*$), then the monotony of the equation produces that for this vector $\vec{x'}$, ($S_1' > S_1^*$), which is absurd, since no solution can be better than the optimal solution. Therefore, there cannot exist any vector $\vec{x'}$ giving a better value to the target function of problem 2 with K_2 than $\vec{x}^{(1)^*}$. This verifies the first part of the theorem. In addition, for equation 2, components $h_i^{(1)}$ and $h_i^{(2)}$ are related by an increasing monotonous function in such a way that the index of the function that produces the minimum in $\{h_i^{(1)}\}$ is the same one that produces the minimum in $\{h_i^{(2)}\}$. This proves the second part.

Appendix

Proof of Theorem 2

Let $\vec{x} \in \mathcal{A}$ and $\forall i$, $h_i^{(1)} = 1 - \frac{|x_i - x_i^{obs}|}{K_1 \alpha_i x_i^{obs}}$, $h_i^{(2)} = 1 - \frac{|x_i - x_i^{obs}|}{K_2 \alpha_i x_i^{obs}}$ then it is easily obtained:

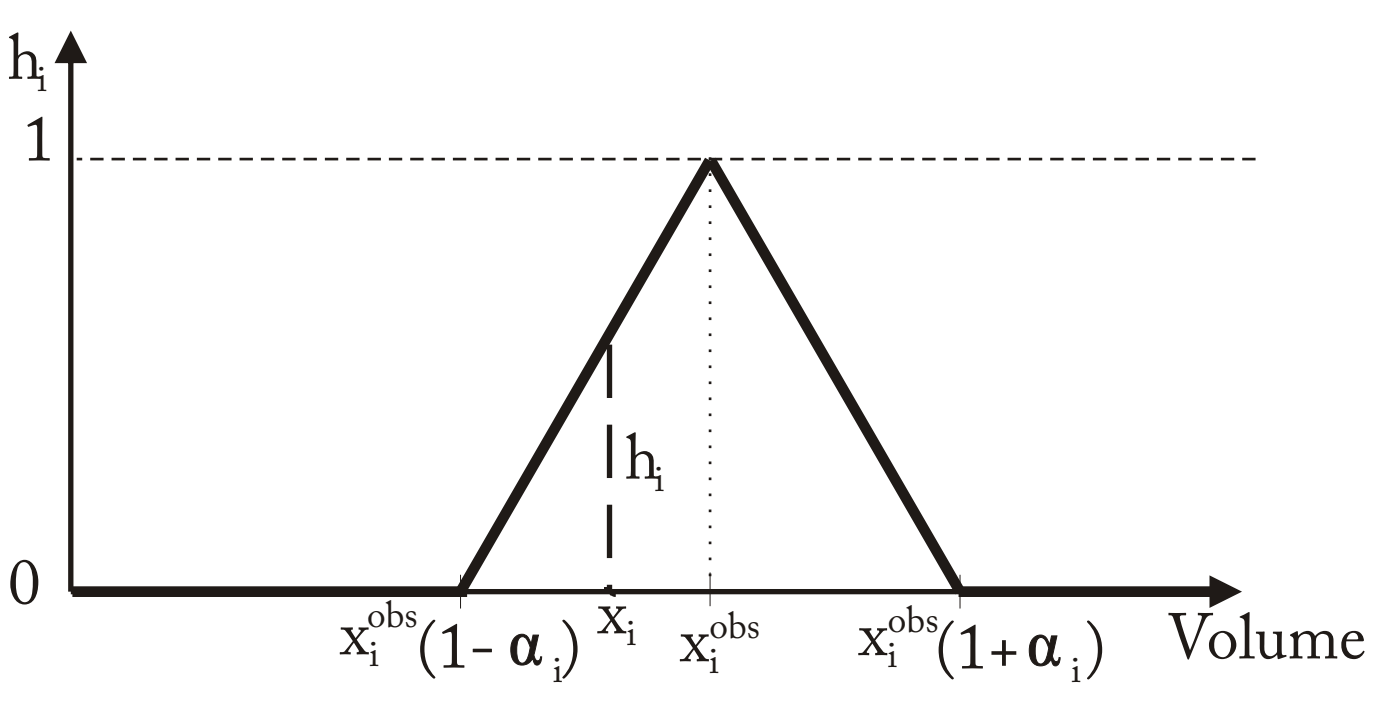
$$K_1(h_i^{(1)} - 1) = K_2(h_i^{(2)} - 1) \tag{2}$$

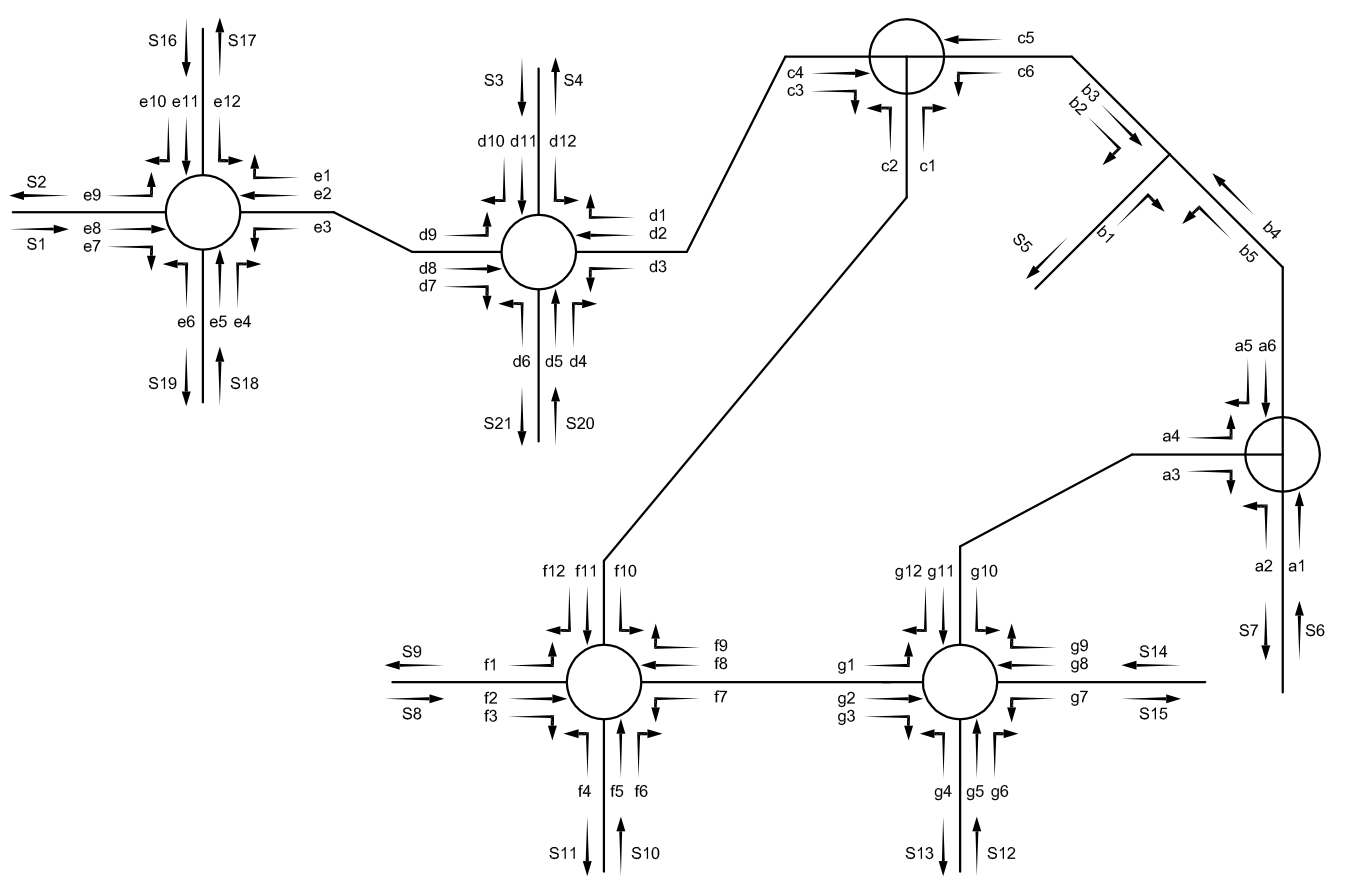
and by naming m the number of observed variables:

$$\sum_{i \in \mathcal{I}} \{h_i^{(1)}\} - m = \sum_{i \in \mathcal{I}} \{h_i^{(1)} - 1\} = \frac{K_2}{K_1} \sum_{i \in \mathcal{I}} \{h_i^{(2)} - 1\} = \frac{K_2}{K_1} \left(\sum_{i \in \mathcal{I}} \{h_i^{(2)}\} - m \right) \tag{3}$$

From equation 3, it can be defined the monotonically increasing function: $S_1 = \frac{K_1}{K_2}(S_2 - m) + m$ between $S_1 = \sum_{i \in \mathcal{I}} h_i^{(1)}$ and $S_2 = \sum_{i \in \mathcal{I}} h_i^{(2)}$.

If we assume that $\vec{x}^{(1)^*} \in \mathcal{A}$ is the set of values that produces the optimal solution to problem 2 with K_1, producing values for target functions S_1^* and S_2^* for the constant K_2. Then, if a $\vec{x'} \neq \vec{x}^{(1)^*} \in \mathcal{A}$ exists and could provide a better solution to problem 2 with K_2 ($S_2' > S_2^*$), then the monotony of the equation produces that for this vector $\vec{x'}$, ($S_1' > S_1^*$), which is absurd, since no solution can be better than the optimal solution. Therefore, there cannot exist any vector $\vec{x'}$ giving a better value to the target function of problem 2 with K_2 than $\vec{x}^{(1)^*}$. This verifies the first part of the theorem. In addition, for equation 2, components $h_i^{(1)}$ and $h_i^{(2)}$ are related by an increasing monotonous function in such a way that the index of the function that produces the minimum in $\{h_i^{(1)}\}$ is the same one that produces the minimum in $\{h_i^{(2)}\}$. This proves the second part.


h_i
1
0
h_i
x_i^obs(1- α_i)
x_i
x_i^obs
x_i^obs(1+ α_i)
Volume


S16
S17
e10 e11 e12
S2
e9
e1
e2
e8
e7
e3
S1
e6 e5 e4
S19
S18
S3
S4
d10 d11 d12
d1
d2
d9
d8
d7
d3
d6 d5 d4
S21
S20
c4
c3
c5
c6
c2 c1
b3
b2
b4
b5
b1
S5
a5 a6
a4
a3
a2 a1
S7
S6
f12 f11 f10
S9
f1
f9
f8
f2
f3
f7
S8
f4 f5 f6
S11
S10
g12 g11 g10
g1
g9
g8
S14
g2
g3
g7
S15
g4 g5 g6
S13
S12