256 shades of grey – comparing OD matrices using image quality assessment techniques

Tom van Vuren
Tim Day-Pollard
tom.vanvuren@mottmac.com

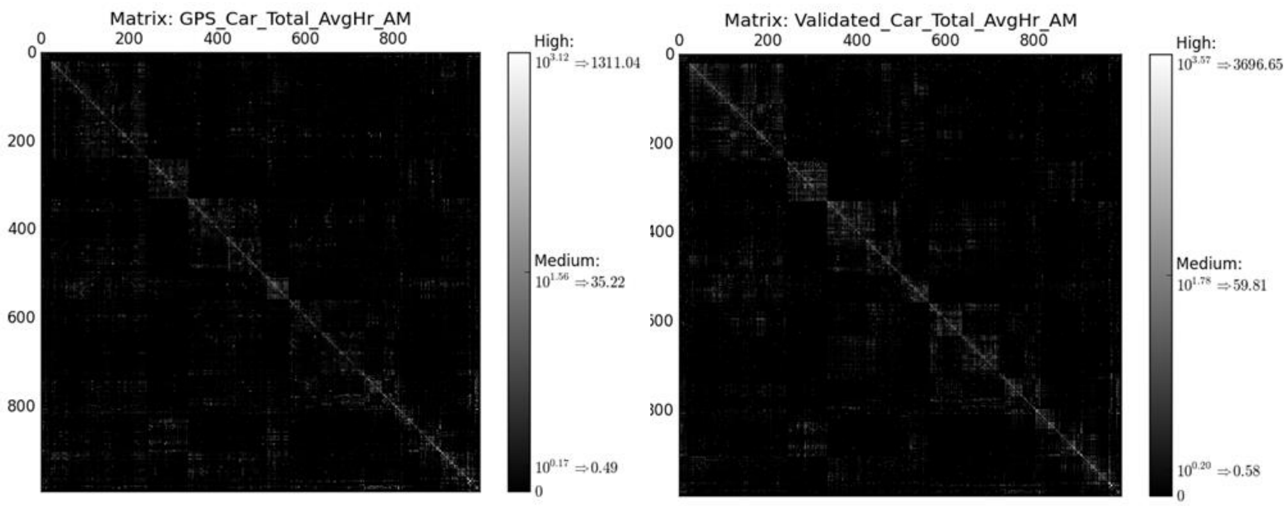
1. OUR PROPOSITION

Imagine a trip matrix, an OD matrix. Usually square, consisting of anything between a few tens of cells for a simple junction to millions of cells for a strategic, regional transport model.

If the matrix is sparse, for example the result of a roadside interview (RSI), the matrix will contain mainly zeroes and a few quite large numbers (interviews expanded to observed flows). If the matrix is based on Big Data, such as mobile phone data or TrafficMaster GPS data, most of the cells in the matrix will be filled, probably with quite small numbers.

Now colour the cells with 256 shades of grey – the zeroes are black, the maximum value is white and use the 254 shades in-between in equal size blocks. Look at the matrix now, and it has become a black and white picture. Matrices that are similar will look very similar. Image processing techniques can now be used to compare them.

Figure 1: Two similar OD-matrices (a GPS component matrix and the final validated matrix) compared as grey-scale images



2. WHY IS SUCH A TECHNIQUE HELPFUL?

OD matrices reflect dynamic spatial movements of travellers but forced into a discrete cell-based structure determined by subjective zoning systems and time period definitions. Both spatially and temporally, these movements are continuous. They differ between days and surveys can only observe a small overall sample whilst estimation techniques tend to be inevitably crude.

Standard statistical techniques do not really reflects such complexities. For example, the r-squared statistic (as recommended in DfT, 2014) can neither recognise the spatial or temporal proximity of movements in adjacent cells. Think again of the image processing analogy. Squinting through your eye-lashes you will recognise that two photographs of which one has moved a fraction left, right, up or down are essentially the same (or at least structurally similar); the r-squared statistic cannot do that.

What we are looking for is a comparator that does not just look at the cell values in isolation, but also considers adjacent cells and cells adjacent to these (in space and potentially in time). The comparator must also be able to deal with some of the idiosyncrasies of OD matrices, for example comparing sparse matrices with dense matrices.

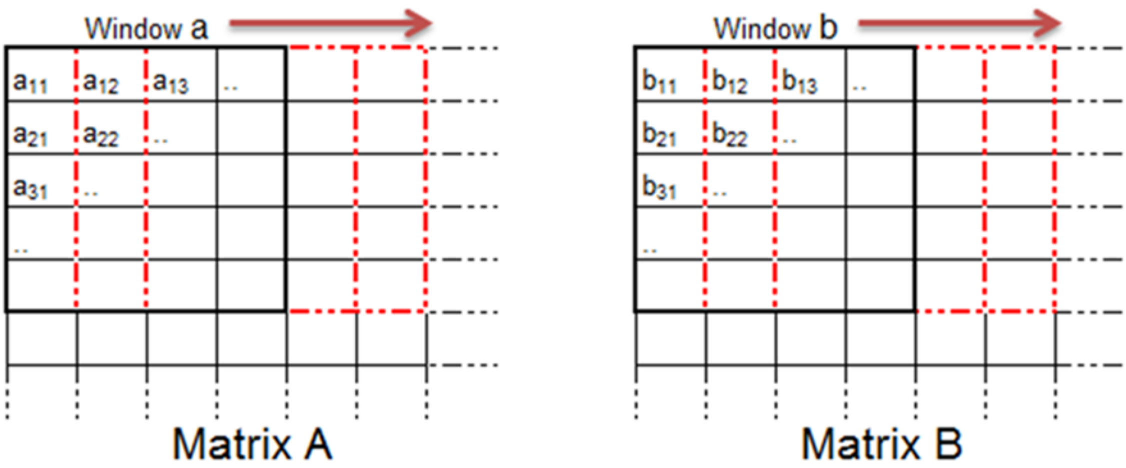
3. WHAT HAS BEEN DONE BEFORE?

In previous publications (Pollard et al, 2013; Day-Pollard and Van Vuren, 2014) we explained how we discovered in existing literature an image processing measure of comparison, the MSSIM (mean structural similarity index) which looks promising for use in OD-matrix comparisons too. The index was originally discovered by Tamara Djukic (Djukic et al, 2013) who refers back to Zhou et al (2004) as the source. Others have also used the MSSIM in OD-matrix comparisons, such as Bringardner et al (2014) and Ruiz de Villa et al (2014).

4. HOW DO I INTERPRET THE MSSIM?

The MSSIM is calculated by summing and averaging SSIM values, as illustrated below, across a whole matrix. The SSIM is calculated over a part or square block of the matrix, generally a few cells wide by a few cells high. Pollard et al (2013) describe the mechanics, illustrated in Figure 2 below.

Figure 2: Calculating the MSSIM by summing and averaging the SSIM value for consecutive blocks in matrices a and b, moving across the whole matrix



What is important to recognise is that the index calculates three characteristics of OD-matrices a and b as in Equation (1) and sums and averages as in Equation (2):

Equation 1

$$SSIM(a,b) = \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)}$$

- The mean value within this part of the matrix, represented by μ_a and μ_b
- The variance (in pixels or OD values), represented by σ_a and σ_b
- The covariance between pixels or OD values, represented by σ_{ab}

Equation 2

$$MSSIM = \frac{1}{N}\sum_{i=1}^{N} SSIM(a_i, b_i)$$

In the language of pictures, μ, σ_{ab} and σ are interpreted as luminosity, contrast and structure, and the SSIM can be simplified to the form in Equation (3):

Equation 3

$$SSIM(\boldsymbol{x},\boldsymbol{y}) = [l(\boldsymbol{x},\boldsymbol{y})]^{\propto} \cdot [c(\boldsymbol{x},\boldsymbol{y})]^{\beta} \cdot [s(\boldsymbol{x},\boldsymbol{y})]^{\gamma}$$

The calculation involves two constants C1 and C2, for which limited guidance is provided in Zhou et al (2004). We understand that their main job is to avoid divisions by zero when values for μ and σ are zero (or sufficiently close to zero to cause computational errors), and we have derived an alternative formulation without these constants.

5. PRACTICAL ISSUES WHEN USING MSSIM

There are a number of immediate issues that arise when comparing OD-matrices using the MSSIM indicator:

1. How big should the block of matrix cells be that are included in the SSIM comparison? This in turn raises the question of how to determine whether zones are close in matrices with different and user-specific zone numbering systems?
2. How to compare sparse and dense matrices?
3. What is an acceptable MSSIM value, or in other words, when are OD-matrices similar enough?

In the next sections we will show that the first question can be answered by using a proximity measure between OD-cells that we developed in Day-Pollard and Van Vuren (2015). The third question remains an open one, and can only be answered by more testing with more matrices and more networks. We are always interested in hearing from you if you are willing to test MSSIM in your own transport modelling work.

6. WHAT IS A NEARBY OD-PAIR?

Whereas in standard images physically nearby pixels are also of greatest MSSIM interest, this is not necessarily the case when comparing OD-matrices. For example, in a hierarchical zoning system based on administrative areas, zones i and i+1 may well be quite distant. There is no guarantee that adjacency in the zone numbering also implies adjacency spatially. Also, and importantly, the comparison we want is between OD-pairs, not origin or destination zones.

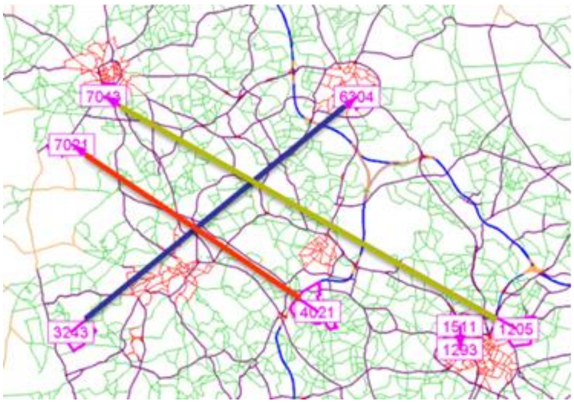
Fortunately, there is a way in which proximity between OD-pairs can be calculated. Every OD-pair has an origin and destination centroid, each of which has x- and y-coordinates. Hence, every OD-pair can be defined by (x_O, y_O, x_D, y_D). The proximity between two OD-pairs can now be established by the Euclidean distance between these coordinates (a_1, a_2, a_3, a_4) for OD pair a and (b_1, b_2, b_3, b_4) for OD pair b, as follows:

Equation 4

$$distance = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 + (a_4 - b_4)^2}$$

As an example, compare in the map below OD-pair 1511-7043 (bottom right to top left) with four other OD pairs, two running parallel and of similar length, one almost perpendicular and one very short but with a shared origin. The Euclidean distances are sensible – providing confidence that this measure rather than the zone numbering is appropriate for determining which OD pairs to consider in the SSIM calculations.

Figure 3: Example comparison of several OD-pairs on the basis of their Euclidean distance

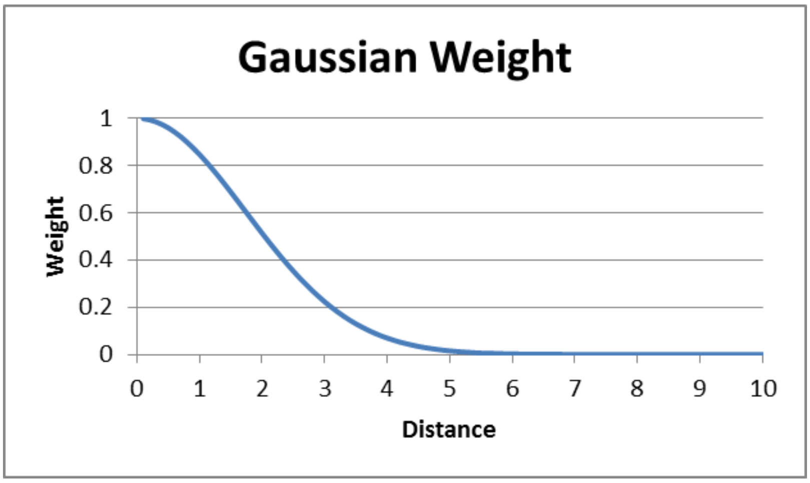


OD Pair	Distance (km)
1205-7043	2.5
3243-6304	19.3
4021-7021	6.9
1511-1293	18.4

As the Euclidean distance as calculated above does not just determine which OD-pairs are near but also *how* near, the actual Euclidean distance can be used as a weight for the contribution of *any* OD-pair to the SSIM value. In our implementation we use a Gaussian weight as advised by Zhou et al (2004), with further OD-pairs contributing less. In Zhou et al (2004) the Gaussian function is chosen because of the comparatively high weight it gives to the closest pixels, and then after a pre-defined boundary the weight drops significantly. Applying such a function directly to a matrix picture such as Figure 1 would not do justice to the spatial peculiarities of OD-movements for the reasons already discussed. In our implementation we have adjusted the formula to take the Euclidean distance as an input rather than position a proxy grid-picture.

We call this enhancement 4D-MSSIM, the 4-dimensional mean structural similarity index. We aim for a shape similar to that in Figure 4, with a falling contribution from OD-pairs further away. The functional form in Equation (5) works well. The value in the denominator deserves some attention; it would make sense to link this to attributes of the network eg the average trip length or average zone diameter and would thus be mode and possibly purpose-specific. We have not yet analysed the effect of different assumptions in this respect. There is a calculation overhead associated with the 4D extension, as the nearest OD-pairs need to be calculated for every OD-pair in turn. Also, it may be reasonable to apply a cut-off in the contributions of further OD-pairs to the 4D-MSSIM value. So far we have limited our calculations to the nearest 625 OD pairs to aid run-time, and used a parameter σ of 7.5 in the Gauss-equation. This parameter was estimated based on the diameter of a sample of zones in the core model area.

Figure 4: Example shape of Gaussian weighting of contributions by OD-pairs to the 4D-MSSIM value as a function of their Euclidean distance



Equation 5

$$w = \exp\left({}^{-x^2}/_{2\sigma} \right)$$

7. COMPARING SPARSE AND DENSE OD-MATRICES

As explained earlier, the MSSIM compares structural similarity calculations of mean, variance and covariance in parts of the matrix, summed and averaged. Whereas the mean is a valid comparator beween sparse and dense matrices, the variances and covariances can be expected to be structurally rather different so that the standard MSSIM would not be able to capture the similarities that we seek to find.

As discussed in the context of Equation (3) the standard MSSIM comparator assesses similarities in mean or luminosity (l), structure (s) and contrast (c) in equal weight with α, β and γ = 1. When comparing sparse matrices, a greater weight might need to be placed on the mean, for example α = 1, β = 0.1 and γ = 0.1. We have carried out some comparative tests but have not yet concluded whether this aids in the comparisons and what appropriate values are for α, β and γ.

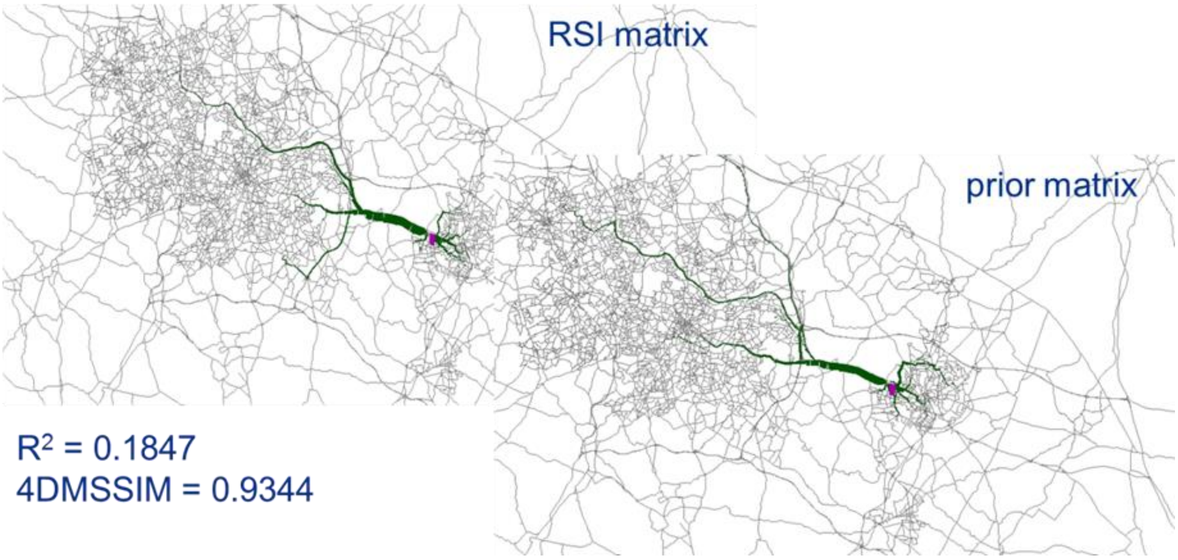
8. DOES IT WORK?

Our previous papers (Pollard et al, 2013; Day-Pollard and Van Vuren, 2015) contain some numerical comparisons between MSSIM and standard OD-matrix comparators such as the r-squared and Chi-squared statistics and the GEH of link flows after assignment. Our conclusions in these papers were, based on artificially constructed matrices:

- The MSSIM index can identify structural differences better than traditional measures such as r-squared or GEH comparisons;
- Conversely, the r-squared value can fail to pick up similarities that MSSIM can detect.

Since developing the 4D-MSSIM implementation we have carried out further tests, using real-life matrices from the PRISM strategic transport model for Greater Birmingham (Van Vuren, 2008); in particular the prior matrix before matrix estimation from counts and one of the (partial) roadside interview (RSI) matrices. We created an RSI equivalent matrix from the validated matrix through a select link analysis at the RSI site.

Figure 5: Assignment of observed roadside interview matrix and equivalent prior matrix in PRISM strategic model plus associated 4D-MSSIM and r-squared values between the two matrices



The results show:

- That the assignments of the two matrices are very similar, as expected;
- That the 4D-MSSIM detects great similarity between the two matrices (value close to 1)
- That the r-squared comparison detects very little similarity (value close to zero), possibly because of the sparseness of the RSI matrix

We also carried out a comparison between the prior matrix and the validated matrix, i.e. that post matrix estimation from counts. WebTAG advises an r-squared comparison, with a target value in excess of 0.95. In the PRISM case an r-squared value of 0.958 is achieved, i.e. a pass. However, it is interesting to note that the 4D-MSSIM comparator only achieves a mediocre 0.611 value.

When considering which is the better comparator for matrix-estimation we must consider the question, what level/type of similarity are we hoping to achieve after matrix estimation? This is particularly relevant given that we are hoping to achieve a different pattern of flows in the assignment. It is acknowledged that some significant level of change is expected and required on some of the OD pairs. The r-squared comparator tests the overall correlation i.e. the average change, whereas the 4D-MSSIM tests the similarity in groups of ODs. This raises the question: is the r-squared a good and sufficient comparator? Are the matrices really similar enough?

9. NEXT STEPS

Our MSSIM work has been carried out after work and in weekends, and as a result our analyses are limited and as yet inconclusive. However, we hope you agree that the results are promising; and in particular that existing matrix comparisons using r-squared or Chi-squared statistics may well be inappropriate; whilst GEH analyses of link flows after assignment are insufficient.

Further testing of the MSSIM or 4D-MSSIM implementations on a wider range of networks and under a wider set of circumstances should provide us insight a) whether the MSSIM index should have a place in our transport modellers' toolkit; b) which of the implementations is most suitable and c) which values we should aim for to be confident that matrices are similar enough. We are always keen to hear from those interested in testing MSSIM in their own work; please get in touch with us using the contact details in the paper.

REFERENCES

Bringardner, JW, Gemar, MD, Boyles SD and Machemehl RB (2014): Establishing the Variation of Dynamic Traffic Assignment Results Using Subnetwork Origin-Destination Matrices, presented at TRB Annual Meeting, Washington DC, January 2014

Day-Pollard T and Van Vuren T (2015): When are origin-destination matrices similar enough?, presented at 94th Annual TRB Meeting, Washington DC, 2015

Djukic, T, Hoogendoorn, SP and Lint, JWC van (2013) Reliability assessment of dynamic OD estimation methods based on structural similarity index, presented at TRB Annual Meeting, Washington DC, January 2013

DfT (2014) WebTAG Unit M3.1 – Highway Assignment Modelling, Department for Transport, January 2014, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427124/webtag-tag-unit-m3-1-highway-assignment-modelling.pdf, accessed June 2015

Pollard T, Taylor N and Van Vuren T (2013): Comparing the quality of OD matrices in time and between data sources, European Transport Conference (http://.etcproceedings.org), Frankfurt, 2013

Ruiz de Villa, A, Casas J and Breen M (2014): OD matrix structural similarity: Wasserstein metric, presented at TRB Annual Meeting, Washington DC, January 2014

Van Vuren, T (2008): A transport policy toolkit for the West Midlands, Proceedings of the Institution of Civil Engineers, Transport 161, pp 85-90, London, 2008.

Zhou, W, Bovik, AC, Sheikh, HR and Simoncelli, EP (2004): Image quality assessment: from error visibility to structural similarity, Image Processing, IEEE Transactions on, 13(4): p. 600-612, April 2004.