Measuring the Performance of Urban Models

Mantelas A. Lefteris¹, Prastacos Poulicos²

Regional Analysis Division, Institute of Applied and Computational Mathematics, Foundation for Research and Technology, Vasilika Voutwn, GR 71110, Heraclion Crete, Greece
¹mantelas@iacm.forth.gr, ²poulicos@iacm.forth.gr

PROBLEM DESCRIPTION

"It is not the precise numerical path, but faith to the dynamic behavior generated by the model" (Meadows et al. 2005) that matters. Nevertheless, the only objective way to evaluate the performance of a model is through the accuracy of its results. In this respect, there is a great number of numerical error and fitting indicators that are used to evaluate the outcomes of urban growth and Land-Use change models. However, accurate results may not necessarily imply an accurate model and as a result comparing the accuracy of models’ results is not a valid mean to compare the different models.

Among existing fitting indicators, the Lee-Sallee (LS) index of agreement and the Kappa (K) coefficient are commonly used to measure the similarity between the model’s results (layer A) and the reality (layer B). The LS index is essentially the quotient of the two layers’ union divided by the two layers’ intersection and shows how correctly coincident are the results of modeling with the spatial shape of the actual urban area (Kim et al. 2006). Its values range from 0, indicating no fitting at all, to 1 indicating a perfect match. The K coefficient is also used to make pair wise comparisons between layers A and B while its values range from -1, indicating the absence of any similarity, to 1 indicating that layers A and B are identical. The value 0 represents the special case where the agreement is exactly equal to the agreement as can be expected by chance.

It should be noted that K coefficient is sensitive to the amount of change present in the maps (Hagoort et al. 2008) while the same applies to LS. This means that if the changes are small, any simulation may lead to very good results. “A high agreement therefore does not necessarily indicate an accurate model” (Jasper 2009). At the same time, previous studies showed that in a selection of published applications, the average LU change is less than 10% (Pontius & Malanson 2005) which means that (in average) the no-change model would score a similarity index greater than 90%. Actually, in some cases, the no-change model may appear to be more accurate than the model’s results (Pontius & Malanson 2005, Jasper 2009). It is evident that new indicators should be used; indicators that measure the performance of urban models taking into account the specific characteristics of each case study. To this end, these indicators should conduct pair wise comparisons between the initial reference map, the final reference map and the map resulting from the model’s simulation (Pontius et al. 2007).

THE GAIN & FIDELITY FACTORS

Towards this direction, certain indicators have been proposed such as $K_{simulation}$ (Jasper 2009). In this paper, we introduce the notion of the Gain and the Fidelity factors that are calculated based on the classical definition of LS. Let $U_1$ and $U_2$ be the urban cover of an area under study in $t_1$ and $t_2$ ($t_1 < t_2$ ) respectively and $U_m$ a model’s estimation of $U_2$ based on $U_1$. Additionally, let $LS_{null}$ be the $LS$ indicator calculated for the null model which assumes no change/growth (the coincidence between $U_1$ and $U_2$) and $LS_m$, the $LS$ indicator calculated for the model’s result (the coincidence between $U_m$ and $U_2$). Finally, note that for the case of a perfect model, $LS$ would equal 1. Using the above notions we may give the following definitions:

$$LS_{gain} := \frac{(LS_m - LS_{null})}{LS_m}$$
$$LS_{fidelity} := \frac{(LS_m - LS_{null})}{(1 - LS_{null})}$$

The Gain factor depicts how much of the results’ accuracy (using LS) is attributed to the model used or in other words how much accuracy we gain by applying the model instead of using the no-model approximation. The Fidelity factor on the other hand, compares the performance of the model to the performance of a perfect model, measuring hence how much of the change simulated by the
model is actually accurate and trustworthy. Assuming that $LS_m \geq LS_{null} > 0$, both factors’ values range from 0, indicating that the model used is no better than the null model, to 1 which indicates a perfect performance. On the other hand, if $LS_m < LS_{null}$, both factors are negative implying thus that the model used is worse than the null model. The Gain and Fidelity factors can be applied upon the $K$ coefficient as well, if both $K_m$ and $K_{null}$ are greater than zero, which is usually - but not always - the case.

In order to illustrate further the use of the herein proposed factors, we use them to measure the performance of the CaFe urban growth model (Mantelas et al. 2010). Specifically we use the simulations produced by CaFe for the east Attica region for three periods during which significant changes in urban cover occurred (Table 1 & 2 – Figure 1). Using only the classical $LS$ and $K$ indicators ($LS_m$ and $K_m$ in tables) it would appear that the model performs better for periods A and B rather than period C. Nevertheless, despite the fact that the results of the model for the periods A and B are closer to reality than the results for period C, this is not a valid conclusion. The Gain factors reveal that the end user has less information to gain by applying the model for periods A and B while for period C the gain in information is tripled. At the same time, the Fidelity factors indicate that the model, when applied for period C, performs 50% percent more accurately than when applied for period A and B.

Figure 1: The actual urban cover layers and the correspondent estimations of them used

<table>
<thead>
<tr>
<th>Period</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>Change%</th>
<th>$LS_{null}$%</th>
<th>$LS_m$%</th>
<th>LS gain%</th>
<th>LS fidelity%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1988</td>
<td>2000</td>
<td>66</td>
<td>61</td>
<td>72</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>B</td>
<td>2000</td>
<td>2007</td>
<td>66</td>
<td>60</td>
<td>72</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>C</td>
<td>1988</td>
<td>2007</td>
<td>177</td>
<td>36</td>
<td>63</td>
<td>43</td>
<td>42</td>
</tr>
</tbody>
</table>

Table 1: The Gain and Fidelity based on $LS$

<table>
<thead>
<tr>
<th>Period</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>Change%</th>
<th>$K_{null}$%</th>
<th>$K_m$%</th>
<th>K_gain%</th>
<th>K_fidelity%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1988</td>
<td>2000</td>
<td>66</td>
<td>74</td>
<td>82</td>
<td>10</td>
<td>31</td>
</tr>
<tr>
<td>B</td>
<td>2000</td>
<td>2007</td>
<td>66</td>
<td>71</td>
<td>80</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>C</td>
<td>1988</td>
<td>2007</td>
<td>177</td>
<td>48</td>
<td>72</td>
<td>33</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 1: The Gain and Fidelity based on $K$
CONCLUSIONS

While $LS$ and $K$ are advisable indicators to measure the fitting of a model’s result to reality, they fail to evaluate the model’s performance itself. For this reason, we propose two indicators that are very easy to calculate based on the classical definitions of the $LS$ and $K$ indicators under certain numerical assumptions. The Gain factor measures how much information the end user can gain by applying the model while the Fidelity factor evaluates the ability of the model to simulate the occurred change.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community’s Seventh Framework Programme FP7/2007-2013 under grant agreement n° 212034

BIBLIOGRAPHY


