Validation of a geographic weighted regression analysis as a tool for area-wide integrated pest management programs for *Ceratitis capitata* Wiedemann (Diptera: Tephritidae) on Terceira Island, Azores

Pimentel, R.¹; Lopes, D.J.H¹; Mexia, A.M.M.² & Mumford, J.D.³

¹ Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group and Universidade dos Açores, Departamento de Ciências Agrárias, Rua Capitão João d’Ávila, 9700-042 Angra do Heroísmo, Tel. (+351) 295402420 (ext 3218)
reinaldo.ms.pimentel@uac.pt
² Universidade de Lisboa, Instituto Superior de Agronomia, Departamento de Ciências e Engenharia de Biossistemas, Tapada da Ajuda; e-mail: amexia@isa.utl.pt;
³ Centre for Environmental Policy, Imperial College London, UK; e-mail: j.mumford@imperial.ac.uk

Adress to Correspondence
Reinaldo Pimentel, Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group and Universidade dos Açores, Departamento de Ciências Agrárias, Rua Capitão João d’Ávila, 9700-042 Angra do Heroísmo, Azores, Portugal; (+351) 295 402 200 (ext 3218); Fax (+351) 295 402 209; e-mail: reinaldo.ms.pimentel@uac.pt;
Abstract

Extrapolation from quantitative sampling to the surrounding areas is an essential feature of many aspects of real world applications in pest management decision making. However, these decisions are only as good as the accuracy of the methods that provided the information. The problem of estimating Mediterranean fruit fly population densities from trap grids is a specific case.

The efficiency of three methods to estimate fruit flies trapped per day (FTD) values for non-sampled areas in Terceira Island is evaluated, the Inverse Distance Weighted (IDW), Ordinary Kriging (OK) and the Geographic Weighted Regression (GWR). Each method has its own specificities and merits.

The results demonstrate that the GWR method is capable of estimating hotspots for the next season and can be used to identify ecological corridors over a non-sampled area. The high spatial heterogeneity and topographical conditions present on Terceira Island may explain why a more mathematically complex method is more reliable than simpler methods for use in possible future wide-area control program for Medfly.

Keywords: Ceratitis capitata; medfly; Geographic Weighted Regression; Inverse Distance Weighted; Ordinary Kriging;
Introduction

Extrapolation from quantitative sampling to the surrounding areas is an essential feature of many aspects of real world applications in pest management decision making. Knowing where pest populations are in time and space is indispensable information needed to effectively plan, implement and evaluate area-wide integrated pest management (AW-IPM) programmes (Hendrichs et al. 2007), and because of this awareness there is a never-ending search for the best tool to fill the spatial gaps between the sampling points in order to get the most efficient overview of the pest status and even to predict when and where future outbreaks will occur. Li et al. (2009), Adeva et al. (2012) and Lux (2013) have achieved this goal by using stochastic models where, according to the authors, they obtained results suitable to support decision making.

Stochastic algorithms usually incorporate random elements or stochasticity into the strategy for searching the space of possible solutions. While stochastic algorithms can find solutions to problems in large search spaces, it can be at the expense of repeatability because of the random search method. In some cases, the convergence on solutions can take a long time, especially because in the initial stages the starting population is not known (Stockwell 1999). For example, in most stochastic models used to predict fruit fly populations in the field, the spatial distribution is considered to be very close to random (Adeva & Reynolds 2012). On the other hand, deterministic models assume that the future response of the system is completely determined by knowledge of the present state and future measured inputs (Jorgensen & Fath 2011).

Independent of the method to be used (stochastic or deterministic), spatial forecast will always be difficult to achieve with some precision due to the existence of unknown ecological heterogeneity. This is even more pertinent since some studies have shown spatial aggregation of fruit flies and particular host-plant pest relationships that show distributions are not random. For example, Dickens et al. (1990) concluded that green leaf volatiles can enhance the attraction responses of male pheromone in several insect orders and such effects have been detected in both Tephritidae and Drosophilidae.

A key point in the success of any Sterile Insect Technique (SIT) operation is the ability of the released males to locate and compete for the wild females (Hendrichs et al. 2002; Gavriel et al. 2012). This necessity could be well improved if the sterile insects were to be released at the locations where the wild population is most abundant. But, according to Hendrichs et al. (2007) the ecological heterogeneity within field, within farm, and at broader spatial scales profoundly affects the population dynamics of pests. This is where Geostatistics applied to the study of the spatial distribution of pest
insects measuring the spatial fluctuations of the variables under study based on rigorous sampling, can improve any program of integrated pest management (Duarte et al. 2015).

The Medfly (*Ceratitis capitata* Wiedemann (Diptera: Tephritidae)) is one of the most important threats to the trade of fresh fruits in the world (Carvalho & Aguiar 1997; Sciarretta & Trematerra 2011; Radonjic et al. 2013) and according to Dickens et al. (1990), green leaf volatiles enhance the attraction of females of this species to male pheromone.

In addition to pest aggregation and host-plant pest relationships, weather conditions such temperature, air humidity and rain fall (soil humidity) will also have an impact on the Medfly lifecycle and its ability to disperse (Bodenheimer 1951; Liu et al. 1995; Papadopoulos et al. 2001; Powell 2003; Estay et al. 2009; Li et al. 2009; Pimentel 2010).

According to Azevedo (1996), the topography of the Azorean Islands has a great influence on weather conditions causing the existence of microclimates, which, according to Pimentel (2010) and Pimentel et al. (2014), can contribute to pest aggregation or dispersion.

Therefore, as a way to try to understand spatial distribution of Medfly adults on Terceira Island (Azores Archipelago), Pimentel et al. (2014) used two different deterministic approaches. The first approach tested was the Ordinary Least Square (OLS) method which provides a global model of the variable or process under survey to understand or predict. It creates a single regression equation to represent that process. In this approach, among other variables, the host-plants variable (presence/absence of host-plants) was taken into consideration. The resulting model could not explain the spatial distribution. As argued by Pimentel et al. (2014), this fact might be related to the pest aggregation strategy of this species and to the local heterogeneity in hosts-plants. Probably that is why only the Geographic Weighted Regression (GWR) analysis could in fact explain the spatial distribution of this species (Pimentel et al. 2014). Considering the countless fruit trees across Terceira Island (abandoned or not) as well as invasive host-plants (like *Solanum mauritianum*), the results obtained by Pimentel et al. (2014) could be reflecting the suitability for fruit trees (or plants with fruits) as well as the existence of microclimates most suitable to Medfly adult aggregation spots. Similar results were also obtained by Lin & Wen (2011), but for the Dengue vectors. These authors managed to successfully detect the geographical heterogeneity by using the GWR analysis, while the OLS method was not so successful.
Despite the differences between the insect pests of both studies, both studies evidenced the same conclusion. For any intervention, the GWR analysis could in fact provide more insights into spatial targeting.

Therefore, the main goal of this study is to compare the geographic weighted regression analysis for Medfly, proposed by Lin & Wen (2011) and Pimentel et al. (2014), to other simpler and more commonly used methods and validate it as an alternative tool for area-wide integrated pest management programs.

**Material and Methods**

In order to validate the methodology proposed by Pimentel et al. (2014) to estimate non-sampled values between sampling points and to compare the efficiency of each method, the estimated values from each method must be statistically compared with real sampled information.

Two different locations were selected for the validation process. One is located at the North side of Terceira Island (Biscoitos parish) with 10 monitoring sites (orchards/backyards). At each site there was a McPhail trap baited with Biolure MedFly* lure (ammonium acetate, putrescine, and trimethylamine) and these traps were located in the control sites of a lufenuron survey conducted by Lopes et al. (2011). The other location is at the South part of Terceira Island (S.Pedro parish). At this site, there were 30 Jackson traps, baited with trimedlure, evenly distributed and placed in two concentric circles. This spatial configuration of traps followed the protocol of Lopes et al. (2008) and it was used to evaluate the spread ability of the released sterile males in a survey conducted by Pimentel et al. (2016).

For the validation process, comparisons are made of sampled abundance, expressed as average Flies per trap per day (FTD) values of females from Lopes et al. (2011) (Figure 1), and the FTD values of wild males sampled during a sterile male Medfly survey (Pimentel et al. 2016) within a research project (MAC/3/A163) CABMEDMAC (2010-2014) (Figure 2).

The FTD is a population index that estimates the average number of flies captured in one trap in one day that the trap is exposed in the field ((IAEA) International Atomic Energy Agency 2003). The function of this population index is to have a relative measure of the size of the adult pest population in a given space and time. The formula of the FTD for site $n$ is as follows:

$$FTD_n = \frac{f}{d}$$

where,

- $f$ is the total number of flies;
- $d$ is the number of days of field exposure.
The selected studies for the validation purpose are from different spatial scales and from different periods. There are several geostatistical methods usable for area-wide integrated pest management programs, but entomologists frequently use spatial interpolation methods to characterize the spatial distribution of pests. Two of the most commonly used methods in insect studies are Inverse Distance Weighted (Lopes et al. 2005; Lopes, Pimentel, Dantas, et al. 2006; Pimentel et al. 2006; Bonsignore et al. 2008; Lopes et al. 2008; Lopes, Macedo, et al. 2009; Lopes et al. 2010; Pimentel 2010; Sciarretta & Trematerra 2011; Rhodes et al. 2011) and Kriging (Alemany et al. 2006; Sciarretta et al. 2008; Castrignanò et al. 2012; Tabilio et al. 2014).

The applicability of any these interpolation methods to study the dispersion of insect pests was recommended by the International Atomic Energy Agency (2006) to managers of area-wide pest management programmes as a working method to interpolate point data over a region.

The Inverse Distance Weighted (IDW) method was introduced by Isaaks & Srivastava (1989), and according to Kemp (2008) is one of the oldest and simplest approaches. It is perhaps the most readily available method. IDW assumes that each sampled point is under a local influence that diminishes with distance. Thus, points in the near neighbourhood are given high weights, whereas points at a far distance are given small weights (Isaaks & Srivastava 1989; Shekhar & Hui 2007). For values of non-sampled
points, there is an assumption that these can be approximated as a weighted average of values at sampled points according to the distances from that point, or from a given number of the closest sampled points (Kemp 2008). Weights are usually inversely proportional to a power of distance, and the most common choice by researchers is the power of 2 (Pimentel et al. 2006; Bonsignore et al. 2008; Lopes et al. 2008; Lopes et al. 2010; Sciarretta & Trematerra 2011; Lopes et al. 2012).

The Ordinary Kriging (OK) method has applications in spatial prediction and automatic contouring. According to Shekhar and Hui (2007), this method was fully developed by Georges Matheson, who named the technique after D.J. Krige, a South African mining engineer who did some of the early work on the topic. The OK methodology is regarded as the best linear unbiased estimator (Du et al. 2010). The distinguishing feature of OK is to minimize the error variance (Isaaks & Srivastava 1989).

However, according to Nestel (2004), all these interpolation methods do not take into consideration the spatial and geographic heterogeneity of the environment.

The Geographic Weighted Regression (GWR) method was introduced by Brunsdon et al. (1996) and this methodology was only possible due to the substantially increased processing power of computers at that time. It carries out multiple regression calculations across each sampling point, which makes this method a very intensive process for a central processing unit. The GWR is a powerful exploratory method in spatial data analysis, yet a relatively simple and effective technique for exploring spatial non-stationarity data as it serves for detecting local variations in spatial behavior (Shekhar and Hui 2007; Charlton et al. 2009). The detection of these local variations allows a better understanding of local details, which may be masked by global regression models (Shekhar and Hui 2007; Charlton et al. 2009). The identification of where and how such spatial heterogeneity in the processes appears on maps is a key in understanding complex geographical phenomena (Dent 2000; Kemp 2008).

The GWR method by itself does not provide any calculation of interpolations, as opposed to the IDW and OK methods. This method only evaluates how coherent the introduced variables are to justify the spatial heterogeneity of the sampled data, and calculate parameters for each location (sampled points) in order to adjust each model to each point of sampled data. Therefore, these parameters later must be interpolated for the surrounding non-sampled points, and these calculations can easily be done by using a simpler method like the IDW. As for the variables to be introduced in this analysis, these are the same as those proposed by Pimentel et al. (2014): altitude, potential (theoretical by location but not accounting for cloud cover) monthly average solar radiation, exposure and slope of each sampling point.
According to Azevedo (1996), these variables determine the local weather conditions in all the Azorean Islands. As the population dynamics of *C. capitata* are linked to biotic and abiotic conditions (Bodenheimer 1951; Fletcher 1989; Meats 1989; Liu *et al.* 1995; Aluja & Rull 2009; Vayssières *et al.* 2009) as well as to the migration flux of adults from other populations nearby (Nestel *et al.* 2004; Israely *et al.* 2005; Shelly *et al.* 2014), these variables might provide a deep relationship between flies abundance and the topographic characteristics. Also according to Batista *et al.* (2006) the geographic location and topographic conditions of each Azorean island provides a different climate condition that affects fruit production.

All the IDW and OK interpolations were done using the Geographic Information System (GIS) for Terceira Island developed by Pimentel *et al.* (2005; 2006). This GIS has the ESRI software ArcMap 9.3 (Environmental Systems Resource Institute 2008) connected to a MySQL database, which contains all the spatial information about Medfly on Terceira Island (Figure 3, Figure 4, Figure 5 and Figure 6).
For the GWR calculations, the GWR parameters were first calculated using the software Spatial Analysis in Macroecology (Rangel et al. 2010) and then exported to ESRI software ArcMap 9.3 to calculate the interpolated constant values using the IDW method. Afterwards, raster calculations with the calculated GWR parameters of each variable were done to obtain the GWR estimated values for non-sampled points (Figure 7 and Figure 8).
The regression statistics were performed by Microsoft Excel 2010 using the previous interpolation values calculated in ESRI software ArcMap 9.3.

**Results**

Table 1 and Table 2 present the FTD values from orchards belonging to a control set (Lopes *et al.* 2011) from Biscoitos parish and the FTD values of wild males from a network of traps belonging to a sterile male release survey from 2014, conducted in São Pedro parish, respectively. In both tables the estimated FTD values of each method for each sampling point are also shown.
Table 1 - FTD values of Medfly females in Biscoitos parish.
* 2009 Average FTD values of Medfly females from Lopes et al. (2011).
** Calculated values from 2006-2008 average FTD values of females from Lopes et al. (2009).

<table>
<thead>
<tr>
<th>Plots</th>
<th>FTD values*</th>
<th>IDW**</th>
<th>OK**</th>
<th>GWR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA1</td>
<td>0.24</td>
<td>16.5</td>
<td>10.9</td>
<td>19.9</td>
</tr>
<tr>
<td>SA2</td>
<td>0.30</td>
<td>25.4</td>
<td>12.0</td>
<td>22.4</td>
</tr>
<tr>
<td>SA3</td>
<td>0.71</td>
<td>21.1</td>
<td>11.6</td>
<td>38.1</td>
</tr>
<tr>
<td>SA4</td>
<td>0.08</td>
<td>18.5</td>
<td>11.2</td>
<td>4.6</td>
</tr>
<tr>
<td>SA5</td>
<td>0.16</td>
<td>6.4</td>
<td>9.2</td>
<td>5.9</td>
</tr>
<tr>
<td>SA6</td>
<td>0.13</td>
<td>11.5</td>
<td>9.9</td>
<td>2.7</td>
</tr>
<tr>
<td>SA7</td>
<td>0.11</td>
<td>13.1</td>
<td>9.7</td>
<td>1.9</td>
</tr>
<tr>
<td>SA8</td>
<td>0.01</td>
<td>9.6</td>
<td>9.2</td>
<td>0.0</td>
</tr>
<tr>
<td>SA9</td>
<td>0.12</td>
<td>5.3</td>
<td>10.1</td>
<td>0.0</td>
</tr>
<tr>
<td>SA10</td>
<td>0.49</td>
<td>19.2</td>
<td>11.0</td>
<td>24.4</td>
</tr>
</tbody>
</table>

Table 2 - FTD values of Medfly males in Sao Pedro parish.
* Captures of wild Medfly males (24 hours after sterile male insect release).
** Calculated values from 2011-2014 average FTD values from field activities of research project (MAC/3/A163) CABMEDMAC (2010-2014).

<table>
<thead>
<tr>
<th>TRAP NUMBER</th>
<th>FTD values*</th>
<th>IDW**</th>
<th>OK**</th>
<th>GWR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43.0</td>
<td>2.8</td>
<td>4.4</td>
<td>55.9</td>
</tr>
<tr>
<td>2</td>
<td>9.0</td>
<td>3.0</td>
<td>4.3</td>
<td>84.8</td>
</tr>
<tr>
<td>5</td>
<td>34.0</td>
<td>2.0</td>
<td>4.5</td>
<td>71.8</td>
</tr>
<tr>
<td>6</td>
<td>40.0</td>
<td>1.4</td>
<td>4.5</td>
<td>91.4</td>
</tr>
<tr>
<td>7</td>
<td>29.0</td>
<td>1.4</td>
<td>4.8</td>
<td>51.7</td>
</tr>
<tr>
<td>8</td>
<td>31.0</td>
<td>1.6</td>
<td>4.9</td>
<td>98.8</td>
</tr>
<tr>
<td>9</td>
<td>143.0</td>
<td>2.2</td>
<td>4.4</td>
<td>92.9</td>
</tr>
<tr>
<td>10</td>
<td>101.0</td>
<td>2.6</td>
<td>4.4</td>
<td>86.2</td>
</tr>
<tr>
<td>11</td>
<td>42.0</td>
<td>3.3</td>
<td>4.3</td>
<td>91.5</td>
</tr>
<tr>
<td>12</td>
<td>15.0</td>
<td>3.5</td>
<td>4.4</td>
<td>96.5</td>
</tr>
<tr>
<td>13</td>
<td>2.0</td>
<td>1.2</td>
<td>4.9</td>
<td>30.9</td>
</tr>
<tr>
<td>14</td>
<td>78.0</td>
<td>2.1</td>
<td>4.4</td>
<td>98.5</td>
</tr>
<tr>
<td>15</td>
<td>0.0</td>
<td>3.5</td>
<td>4.3</td>
<td>93.3</td>
</tr>
<tr>
<td>16</td>
<td>23.0</td>
<td>3.5</td>
<td>4.4</td>
<td>84.3</td>
</tr>
<tr>
<td>17</td>
<td>9.0</td>
<td>3.6</td>
<td>4.4</td>
<td>70.6</td>
</tr>
<tr>
<td>18</td>
<td>93.0</td>
<td>3.5</td>
<td>4.4</td>
<td>69.5</td>
</tr>
<tr>
<td>19</td>
<td>31.0</td>
<td>3.7</td>
<td>5.0</td>
<td>98.2</td>
</tr>
<tr>
<td>20</td>
<td>1.0</td>
<td>3.0</td>
<td>5.0</td>
<td>86.6</td>
</tr>
<tr>
<td>21</td>
<td>0.0</td>
<td>2.2</td>
<td>4.9</td>
<td>39.8</td>
</tr>
<tr>
<td>22</td>
<td>1.0</td>
<td>1.9</td>
<td>4.9</td>
<td>35.8</td>
</tr>
<tr>
<td>23</td>
<td>20.0</td>
<td>0.8</td>
<td>4.8</td>
<td>0.9</td>
</tr>
<tr>
<td>24</td>
<td>2.0</td>
<td>0.7</td>
<td>4.8</td>
<td>54.7</td>
</tr>
<tr>
<td>25</td>
<td>4.0</td>
<td>0.8</td>
<td>4.4</td>
<td>43.7</td>
</tr>
<tr>
<td>26</td>
<td>44.0</td>
<td>1.4</td>
<td>4.4</td>
<td>81.9</td>
</tr>
<tr>
<td>27</td>
<td>9.0</td>
<td>2.5</td>
<td>4.4</td>
<td>90.8</td>
</tr>
<tr>
<td>28</td>
<td>48.0</td>
<td>2.8</td>
<td>4.4</td>
<td>94.8</td>
</tr>
<tr>
<td>29</td>
<td>39.0</td>
<td>3.4</td>
<td>4.5</td>
<td>99.5</td>
</tr>
<tr>
<td>30</td>
<td>113.0</td>
<td>3.4</td>
<td>4.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>
From the values presented in Table 1, regression analyses were performed for each method. On the X axis are the average FTD values from trap counts in 2009, and on the Y axis are the calculated values of each method (Figure 9, Figure 10 and Figure 11).

Figure 9 - Regression analysis using the FTD values from trap counts and the calculated values using the IDW interpolation presented in Table 1.

Figure 10 - Regression analysis using the FTD values from trap counts and the calculated values using the OK interpolation presented in Table 1.
From the three regression analyses, the only regression that presented a reliable coefficient of determination (89%) is the one related to the GWR method (Figure 11). The high coefficient of determination in that regression is supported by a p-value of 3.8E-05, which for a confidence level of 95%, means that the relationship between the estimated values (based on the 2006-2008 field data from Lopes et al. (2009)) and the measured values (2009 field data from Lopes et al. (2011)), is significant. The calculated values might not be accurate in terms of a real local populations, but according to the results, it might be possible to get a qualitative estimation of the presence of a hotspot over the tested area just using the average FTD values from Lopes et al. (2009).

As for the values presented in Table 2, regression analyses were also performed for each method. Again, on the X axis are the FTD values of wild males from a network of traps belonging to a sterile male release survey from 2014, and on the Y axis are the estimated values from each method. Figure 12, Figure 13 and Figure 14 show graphical representations of those regressions and their trendlines, along with the equation and coefficient of determination.
Figure 12 - Regression analysis using the FTD values from trap counts and the calculated values using the IDW interpolation presented in Table 2.

Figure 13 - Regression analysis using the FTD values from trap counts and the calculated values using the OK interpolation presented in Table 2.
In contrast to the previous tests, none of the regression analyses from Table 2 data showed any suitable coefficient of determination. Although all methods in the evaluation performed poorly, the GWR method got the highest coefficient of determination.

The main difference between the information presented in Table 1 and Table 2, apart from the location, is the average distance between sampling points. For Table 1 these traps are on average separated by 600 metres while the traps in the study related to Table 2 are on average separated by 70 metres from each other (Lopes et al. 2008). This closer spacing between sampling points demands a great amount of definition on any mathematical method. The sampling points in Table 1 are within 0.44 km² and in this situation the general spatial heterogeneity seems to be detected by the GWR method as it reasonably estimated values over a non-sampled area. But when it comes to evaluate the heterogeneity within only 0.13 km², this method, as well as the others, statistically fails. However, it is useful that the GWR method identified an ecological corridor over a non-sampled area (see the lighter shading across traps 19, 9, 8, 6 and 14 in Figure 8) where other methods did not. This corridor was already reported in previous field work done by Lopes et al. (2006) in the same area, where the main concentration of Medfly male abundance was identified across this corridor.

The IDW method in both tests did not perform as well as Bonsignore et al. (2008) reported in their work in Italy. These authors evaluated IDW field applicability on area-wide pest management for the estimate of non-sampled points and found this method to be very adaptable in forecasting estimates. However, the precision of those estimates, according to the same authors, varied between months.
One important feature of the interpolation methods is the ability to produce maps with graphical impact showing insect pest aggregation sites. Sciarreta et al. (2011) while studying the spatio-temporal distribution of medfly in a heterogeneous landscape concluded that this method will allow the sanitation and control methods to be directed toward areas where flies aggregate.

According to Alemany et al. (2006) the OK interpolation method provided the revelation of spatial heterogeneity and provided the information that would allow them to proceed to more efficient a pest management program by focusing control efforts on areas where target populations are at their greatest.

In order to get a map of Lobesia botrana spatial distribution inside the olive groves, Sciarreta et al. (2008) used the OK interpolation method and these authors, based in their results, suggests its use as a tool in precision farming control within integrated pest management programs.

While studying the spatio-temporal population dynamics of Bactrocera oleae, Castrignanò et al. (2012) concluded that despite the good theoretical results of the OK interpolation, such method should first be more tested and validated to the study area before its application is considered in any future area-wide integrated pest management program. This testing procedure is important because if the process under study is essentially spatially random, according to Tobin (2004) it should not be consequently mapped using a spatial interpolation method such as OK.

Also, while studying the spatio-temporal dispersion, but of the Anastrepha ludens over large flat topographic citrus areas, Vanoye-Eligio et al. (2015) obtained with success the occurrence probability maps for this specie.

Rhodes et al. (2011) used IDW and OK produced maps with similar accuracy, indicating that “hot spots” could be modeled using either interpolation method. One possible reason for this reported situation is because both methods rely on Tobler’s first law of geography, which states that things that are close are more related than things that are further away (Tobler 1970).

Midgarden et al. (2014) aside of pointing out the strengths and weakness of each method, show the importance of using the results of such interpolation methods on the decision making process of area-wide integrated pest management programs. This importance becomes even clearer by Enkerlin et al. (2015) while reviewing over 30 years of the successful containment program in Mexico, which culminated in the eradication of the C. capitata. These authors make a positive remark of such spatial analyses methods as it led to the increased program effectiveness as it these mathematical methods allowed them to identify and define infestation fronts and proceed accordingly.
The main reason the IDW and OK methods did not perform so well in our study is perhaps related to the Terceira Island topography. As indicated before, the Azorean topographic characteristics can have great importance in local weather conditions (Azevedo 1996) and the Medfly life cycle is closely linked to these (Bodenheimer 1951; Fletcher 1989; Meats 1989; Liu et al. 1995; Aluja & Rull 2009). Both methods basically rely on the location of the sampling points and the weight of each point (trap counts). There is no input of multiple variables for each point in the calculation. Therefore, as the results demonstrated, it will be hard for these methods to estimate hotspots over a large non-sampled area under Terceira Island topographic conditions with sufficient accuracy.

The underlying idea of the GWR method is that parameters may be estimated anywhere in the study area given a dependent variable and a set of one or more independent variables which have been measured at places whose location is known. Considering Tobler’s first law about proximity and resemblance we might expect that if we wish to estimate parameters for a model at some location, then observations which are nearer that location should have a greater weight in the estimation than observations which are further away (Charlton et al. 2009).

Conclusions

The scope of this work was to evaluate the efficiency of estimation from three methods over a large area for a future area-wide control program on Terceira Island. The results show that the GWR method is better than the others to estimate possible densities of Medfly abundance over a large non-sampled area.

The results also suggest that the methodology proposed by Pimentel et al. (2014), for Terceira Island, might be scale sensitive. However, the output resolution will obviously depend on the scale at which the GWR parameters were calculated and these will only depend on the density of the sampling points. It is also important to note that the GWR method can only provide this type of results because of its local multi-variate regression analysis characteristics. Within the parameters of the proposed methodology, with distances between sampling points of about 600 metres, the GWR method can estimate population densities with sufficient statistical significance over a non-sampled area. As demonstrated, even when it was not possible to statistically establish a significant relationship, the GWR map output might provide some insights about the Medfly spatial distribution, as illustrated by the apparent presence of an ecological corridor.
As an overall analysis, present results can provide two essential features for a future area-wide control program on Terceira Island. One is related to orienting the decision making process into non-sampled areas identified as of greater importance to this insect. The second feature is about having an initial state from which future stochastic models can start running, and therefore solving one of the main problems related to the initiating stage of stochastic models reported by Stockwell (1999).

While the results demonstrate that the GWR method is capable of estimating next season hotspots, or even identifying ecological corridors, these estimations might be time limited. This limitation might occur because pests and their interactions with host plants, ecological landscapes, parasitoids and predators are in a continuous flux, and therefore they are never static. Pest severity increases or decreases depending on environmental conditions and changes in production or pest control practices (Goodell et al. 2014).

Acknowledgments

This research is financed by Portuguese national funds through the Fundação para a Ciência e a Tecnologia (FCT - SFRH / BD / 81079 / 2011)
Bibliographic references


Bodenheimer FS. 1951. Citrus Entomology in the Middle East: With Special References to Egypt, Iran, Iraq, Palestine, Syria and Turkey (Hardback). 1st ed. Liverpool, Mers, United Kingdom: Springer.


