Introduction

The fast tracking of the invasion spatial patterns of alien species is crucial for the implementation of preventive and management strategies of those species. Recently, a honeybee pest, the small hive beetle *Aethina tumida* (hereafter SHB), has been reported in Italy, where it colonized more than 50 apiaries in an area of about 300 km². SHB is a nest parasite and scavenger of honeybee colonies native of Sub-Saharan Africa. Likely being helped by the globalization of apiculture, SHB underwent several invasions in the last twenty years, causing locally relevant economic impact. While many features of its biology have been addressed, an important knowledge gap concerns the spatial invasion dynamics in invaded areas. In this paper we coupled two spatial analysis techniques (geographic profiling and a density-based spatial clustering algorithm) to uncover the possible invasion pattern of SHB in Italy. We identified the port town of Gioia Tauro as the most likely point from which SHB may have spread and suggested the possible successive axes of diffusion. These putative diffusion paths suggest that the SHB spread in south Italy might have been due to a mix of natural dispersal between close apiaries and longer distance movement through faster, likely human-mediated, communication routes.

© 2018 Published by Elsevier Editora Ltda. on behalf of Sociedade Brasileira de Entomologia. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
colonies in South Italy, near the coastline of region of Calabria, where it has probably accidentally introduced from Africa (Granato et al., 2017). SHB was detected in 58 apiaries and one natural colony, spanning over an area of about 316 km² (Mutinelli et al., 2014). New infested apiaries have been detected in 2015, 2016 and 2017 (http://www.izsvenezie.it/aethina-tumida-in-italy/), and the invasion is thus not resolved.

Despite research efforts strongly increased the knowledge on SHB during the last few years, many knowledge gaps still remain (Neuman et al., 2016). Among them, one of the top priorities is to understand the invasion dynamics, in particularly the spatial pattern of diffusion after the arrival in new areas (Neumann and Elzen, 2004).

Here, we propose a two-step spatial analysis in order to identify the possible post-invasion spatial dynamics of IAS starting from presence-only reports, using the recent invasion of small hive beetle Aethina tumida Murray, 1867 (Coleoptera: Nitidulidae, hereafter SHB) as a model. The detection of spreading centres, in addition to facilitate the prevention of recurrent IAS introduction, may also allow the identification of the possible idiosyncratic features that boosted the invasion (e.g. ecological requirements). Starting from the spreading centre(s), diffusion may then take place in several spatial ways, being influenced by the ecological needs and life history traits of the IAS as well as the environmental heterogeneity and possible man-assisted displacement (Wilson et al., 2009). Our analysis aims to (a) identify the most likely spreading centre, (b) infer the likely early post-invasion diffusion pathways of SHB invasion in Italy and (c) provide a model for other invasions of A. tumida in other geographical areas. IAS spreading centre(s) are those invaded areas which more likely allowed the first successful establishment of the IAS and thanks to which the IAS then spread to other areas.

In order to infer the possible invasion dynamics of SHB in Italy from presence reports, we coupled Geographic profiling – hereafter GP – (Rossmo, 2000), a modelling technique widely used to infer spreading centre of invasive species, together with a density-based spatial clustering algorithm (DBSCAN, Ester et al., 1996). Geographic profiling (GP) is an analytic tool for the identification of the geometrical origin of linked events, which has been recently applied to a range of ecological issues, among which the identification of the source population(s) of invasive species using the known positions of their current populations (Stevenson et al., 2012; Cini et al., 2014; Faulkner et al., 2017; Papini et al., 2013, 2017a,b), while several enhancements of the method are under study in order to refine the reliability of the analysis (Papini and Santosuosso, 2017; Santosuosso and Papini, 2018).

The DBSCAN is a clustering algorithm defined as density based as it produces clusters of points (on a plane) that are grouped together according to the concept that each point of the cluster must be close to at least a given number p of other points of the clusters within a given radius r (Xu and Wunsch, 2005). DBSCAN has several features that seem advantageous in biological invasions, since it is quite robust to outliers (Smith et al., 2002), minimally affected by points ordering and it requires little a priori information about the dataset parameters values, such as number and shape of the clusters (Smitti and Elouidi, 2013), a case which is typical of biological invasions in their first years after IAS introduction.

Our aim was to draw the likely spatial pattern of SHB invasion in Italy, thus setting the stage for further implementation of specific monitoring protocols in this area.

Material and methods

Data collection

SHB presence data were retrieved from Mutinelli et al. (2014) which provides the data of the National Reference Centre for Apiculture (available at http://www.izsvenezie.it/aethina-tumida-in-italia/), collected during the apiary surveillance programme after the discovery of SHB in Italy. Briefly, a subsample of the apiary present in a radius of 20 and 100 km from the first report were inspected, and apiary were considered infested if at least one hive showed the presence of SHB, in any developmental stage. As georeferentiation of infested apiaries is not freely available due to privacy reasons, we georeferenced apiary site directly from the map. Using this procedure estimated spatial accuracy was 25 metres, which was thus not expected to affect the resolution at which our spatial analysis was performed (tens to hundreds of kilometres). As we focus on the very early post-invasion dynamics in the arrival zone (Calabria region) and we aimed at understanding possible axes of diffusion using only early post-invasions data, we did not take into account the presence report from Sicily (one infected colonies in a 56-hives migratory apiary, likely due to human hive translocation, Mutinelli et al., 2014), nor the presence data referring to 2015 or 2016 surveys, an approach already used (Cini et al., 2014). Moreover, 2015 and 2016 data have been collected after sanitary measures, such as hive destruction, were performed, so that they might not properly represent the initial spatial dynamics of SHB. Data were mapped on the South Italy map (obtained by the open access website www.openstreetmap.org).

Identification of the most likely spreading centre: geographic profiling

GP uses coordinates on a map of linked events (e.g. locations where an IAS has been reported) to generate a probability surface, which will be superimposed on the original map to produce the so-called geoprofile. Such geoprofiles, based only on presence data, do not provide the exact origin of the events, but produces decreasing probability density, thus prioritizing geographical areas (Rossmo, 2000) where monitoring protocols should be implemented and targeted researches carried out. We used the model for Geoprofiling analysis described by Rossmo (2000) and modified according to Papini et al. (2013). For each pixel with coordinates (i,j) of the target area, the score function (p) is calculated as follows (redrawn from Stevenson et al., 2012):

\[
P_x = k \sum_{n=1}^{r} \left( \frac{\phi}{\sqrt{(x_i - x_n)^2 + (y_i - y_n)^2}} \right)^2 + \left( 1 - \phi \right) B^{f-x} \]

where:

\[
\phi = \begin{cases} 
1, & \text{if distance } > B \\
0, & \text{otherwise}
\end{cases}
\]

For a point P of coordinates (i, j) the formula sums the probability across all the locations where the invading organism was found. This model is based on two components: a distance decay function (in which the probability of a presence report decreases as distance from the invasion site increases) and a buffer zone of radius B, within which the likelihood increases when distance increases.
The distance-decay function can be explained by dispersal costs while the buffer zone may be due to habitat unsuitability of invasion sites. Stevenson et al. (2012) suggested that it is conceivable to use general, taxon or habitat specific values for the model parameters in cases where no data on the target species exist. In our case, we chose model parameters according to previously published analyses on invertebrate invasive species, such as Stevenson et al. (2012) and Cini et al. (2014). Since we do not have any information about the possible real values of the crucial buffer zone radius (B) for this or related species, and since it might depend on several factors (such as dispersal ability, habitat heterogeneity, etc.), we selected the values obtained for the only available insect species from Stevenson et al. (2012) (B values: 0.38–0.42), which, scaled to our map was about B = 10 (it is not straightforward to obtain this data, since Stevenson et al. did not indicate how large is B in km or the resolution of their map). In order to evaluate the robustness of the model when B varies, we ran the analysis using B values ranging from B = 5 to B = 20 (corresponding to a range of about 0.125–5 km), which spanned across the values found for other terrestrial flying insects (Stevenson et al., 2012, adapted to the different map scale). For this purpose we used the python script written by the authors: Geoprot2.0_5csv.py on Bitbucket (https://bitbucket.org/ugosnt/al_and_ugo/).

Identification of the possible diffusion pathways: cluster analysis and density maps

The DBSCAN algorithm was used to produce clusters of points corresponding to observations on the used map. The points of a given cluster must be close to at least a given number p of other points within a given radius r (Xu and Wunsch, 2005) (Fig. 1). In order to separate different clusters of observations we used the following algorithms implemented in python: Kmeans (number of clusters assessed with Silhouette) (Python script Kmeans Silva 0.0.2.py) with the following space Voronoi tessellation as proposed in detail by Santonuosso and Papini (2016) and the Density-based spatial clustering (DBSCAN), as defined by Ester et al. (1996). DBSCAN was implemented in Python as dbscan 1.0.0.py and after preliminary trials we used r = 6 (pixels), corresponding to about 100 m and p = 5 (number of points to be found within the radius) as parameters. Finally, in order to infer possible spatial invasion dynamics, we coupled GP and DBSCAN results by drawing the putative invasion directions considering as a starting point the centre of the identified 95% spreading areas identified by the GP and the tip of each direction arrows as the centre of mass of each cluster identified by DBSCAN method.

Results

The GP analysis identified an area of about 27 km² on the coast, north/northeast of the town of Gioia Tauro as the 95% (in red) most likely spreading centre of SHB in South Italy (Fig. 2a) on the basis of the infected apiaries. Overlapping the area identified by GP as the spreading centre, using as input sources infected apiary only or both infected and non infected apiary (null model, Fig. 2b), showed that the second area of 95% probability (in red) of origin, was located in a more northern position with respect to the first (only infected apiaries), suggesting that the area identified by GP was not simply reflecting apiary density. The intersection between the two areas is shown in Fig. 2c and it corresponds to less than 4% of the total number of red pixels, which can be considered a significant difference. GP results are robust, as varying the B value, the result about the localization of the area of highest probability of spreading centre did not vary more than few pixels (data not shown).

DBSCAN clusterization assigned the 57 cases of infected apiaries to three clusters and identified 17 apiaries as outliers (i.e. not necessarily belonging to any cluster). Two of the three clusters fall indeed at the border of the spreading centre identified by the GP approach (compare Fig. 2 with Fig. 3). The third one (as purple ovals in Fig. 3) is about 10 km southward from Gioia Tauro and about 15 from the putative spreading centre, as identified by GP. Considering this coupled analysis, it seems very likely that a main axis of SHB spread, pointing southward, can be identified (Fig. 3a and b).

Discussion

The identified area of about 27 km² around the Port of Gioia Tauro as the most likely spreading centre of SHB in Italy supports the strict relationship of this invasion with the presence of one of the most important ports in South Italy.

The position of the three main clusters of SHB identified by the DBSCAN analysis showed that while two are close to the supposed spreading centre (tentatively identified with GP), the third one is positioned 15 km southward. This coupled approach thus suggests (a) that a port, Gioia Tauro, might be the entrance, and (b) that SHB might have had three main diffusion directions, with different easiness of penetration along different spatial gradients.

The importance of marine ports as likely points of IAS entry has been widely recognized (Keller et al., 2011; McCullough et al., 2006). Travelers’ baggage, cargo and trade items may easily introduce exotic invertebrates and plant weeds from all over the world. Several alien insects are indeed suspected of having spread into new areas diffusing from marine port town. It is the case of wood-boring beetles (Haack, 2001) and of the small fruit pest Drosophilidae suzuki (Cini et al., 2014). Indeed, the possibility that SHB might also enter through ports has been proposed for the USA East-coast and Australia invasions (Hood, 2004) and was already suggested for Italy (Mutinelli et al., 2014). Our GP approach strongly supports this suggestion.

The finding of two very close clusters and one more distant southern cluster suggests that SHB Italian invasion might have proceeded thanks to a mix of leading-edge dispersal (i.e. diffusive spread from the edge of the invasion range, likely driven by natural dispersal) and jump dispersal (long-distance dispersal over substantial distances likely due to accidental human mediated
transport), in agreement with previous findings (Hood, 2004; Spiewok et al., 2008). We must also consider that winds in Calabria blow most commonly northward from south and southeast regions (Fratianni and Acquavetta, 2017), and the third most distant cluster is positioned just south to the first two. Hence, dispersal appears to be in the opposite direction with respect to the sirocco wind. Moreover, in the area of the third cluster no main road is present, hence apparently the spreading was due to short distance local movements. The proposal of a spreading model with beekeepers infesting their other apiaries through ‘unintentional transfer’ was called “distance and ownership model” (EFSA, 2015) and this model of spreading would be supported by our data. On the contrary the first two northern clusters are in industrial areas where the human presence is very common: here the spreading occurred at relatively low distance (less than 1 km) and may be due partly to the natural movement of the pest and partly by accidental transport by human activity.

Figures 2 and 3. Identification of the possible diffusion centre. (a) The 95% most likely spreading centre is in red and corresponds to the results of the geographic profiling (GP) applied only to the infected apiaries; (b) a null model, with GP results (the red areas) based on all the apiaries present in the area, both infected and uninfected. The red area is localized in a more northern area with respect to (a); (c) intersection of the red areas (a) and (b). Since the intersection is less than 5% (in our case the difference is 4% or less) of the total red area, we can consider the difference as significant. The blue points correspond to uninfected apiaries and red points to infected apiaries. (c) is positioned in the middle between (a) and (b) to better show the intersection. ×30,000.

Figures 2 and 3. Possible spreading pathways from the identified spreading centre are mapped in the invaded region on the basis of the algorithm DBSCAN. The pink and blue lines represent the Voronoi tessellation of the space after the Kmeans analysis for three clusters (checked with Silhouette method, see the text). The dashed arrows connect the centroid of the clusters and should give suggestions about the axes of diffusion. (a) Results of the GP analysis. (b) Representation of the recorded observations (infection cases) on the territory with the representatives of the different clusters highlighted with different colours. The blue points correspond to uninfected apiaries. ×250,000.

Policies to reduce the risk of further spreading of the pest may be hence centred mainly on the regulation of the activity of beekeepers proposing guidelines aiming to reduce the transfer of material from one hive to another or to transport hives from one area to another.

As with any modelling approach, GP has some limitations. First, its reliability reflects the accuracy of presence data. Unfortunately, standardized presence data are very difficult to obtain for invasive species right after their invasion. In the SHB case presented here, contrary to what usually happens with many IAS, the pest species was already recognized as a honeybee notifiable pest (Commission Decision 2004/216/EC), so that veterinary services rapidly put in place a standardized survey in a wide area around the first SHB reports (Mutinelli et al., 2014), which resulted in a reliable database,
thus providing strength to our modelling approach. A second possible source of biased sampling is represented by the ability of SHB to survive on non-honeybee related resources, such as rotten and fermenting fruits or other species (such as bumblebees, Hoffmann et al., 2008), which might represent a refuge for SHB allowing further re-infestations (Neumann and Elzen, 2004). However, so far no SHB has been detected on rotten fruits in a field investigation carried out in the infested area between 2014 and 2015 suggesting that fermenting fruits do not represent a possible alternative food resource for SHB in Calabria (Mutanelli et al., 2015). However, it has been shown that these alternative resources are less preferred and less successful food sources, that also reduce SHB fitness (Ellis et al., 2002), so that we are confident that the used dataset represent a reliable presence dataset.

Despite these possible limitations, we thus believe that our results provide a first reliable hypothesis of SHB invasion dynamics in Italy, supporting the hypothesis of a seaport mediated invasion (Mutanelli et al., 2014 for Italy, and Hood, 2004 for US and Australia), and further highlighting the role of seaports as IAS introduction sites (Mack et al., 2000). The Voronoi tessellation and the dashed lines linking groups of observations (see Fig. 3a) should represent an estimate of the spreading direction, while the distance from the hypothesized point of origin, related to the guessed time of arrival of SHB, should give an idea about the speed of the diffusion itself in the territory.

More generally, we believe our results show that the coupling of GP and DBSCAN might represent a promising and rapid approach to use these data to prioritize areas, inside IAS invaded ranges, in which to focus future monitoring and managing efforts.

This analysis may be used efficiently in areas where the invasion is still at the beginning as in Brazil (Al Toufailia et al., 2017), in order to better understand the causes of the invasion and its spreading directions in order to undertake suitable actions.

Compliance with ethical standards

This article does not contain any studies with human or other animal participants performed by any of the authors.

Author contribution statement

AC, US and AP conceived research. AC collected the data. AC, US and AP analyzed the data. AC wrote the manuscript. All authors read and approved the manuscript.

Conflict of interest

The authors declare no conflicts of interest.

References


