Risk-based regionalization approach for area-wide management of HLB vectors in the Mediterranean Basin

Anaïs Galvañ1, Renato Beozo Bassanezi2, Weiqi Luo3,4, Pilar Vanaclocha1, Antonio Vicent1 and Elena Lázaro1*

1Centre de Protecció Vegetal i Biotecnologia, Institut Valencià d’Investigacions Agràries (IVIA), Moncada, Spain, 2Departamento de Pesquisa e Desenvolvimento, Fundo de Defesa da Citricultura, Araquara, SP, Brazil, 3Agricultural Research Service, U.S. Department of Agriculture, Fort Pierce, FL, United States, 4Center for Integrated Pest Management, North Carolina State University, Raleigh, NC, United States

Huanglongbing (HLB) is one of the most devastating citrus diseases worldwide. It is associated with the non-culture bacteria Candidatus Liberibacter spp., which can be transmitted by grafting and/or the psyllid vectors Diaphorina citri (ACP) and Trioza erytreae (AfCP). Although HLB has not been reported in the Mediterranean Basin to date, both vectors are present, and thus represent a serious threat to the citrus industry in this region. Resistant citrus cultivars or effective therapeutic treatments are not currently available for HLB. Nevertheless, area-wide pest management via coordinated management efforts over large areas has been implemented in Brazil, China and the USA for HLB control. This study proposes an open access flexible methodology to address area-wide management of both HLB vectors in the Mediterranean Basin. Based on a risk-based approach which considers climatic information and other variables that may influence vector introduction and spread, such as conventional, organic, abandoned and residential citrus areas as well as transportation corridors, an area-wide management division in pest management areas (PMAs) is proposed. The size and location of these PMAs were estimated by means of a hierarchical clustering algorithm with spatial constraints whose performance was assessed under different configuration scenarios. This proposal may assist policymakers and the citrus industry of the citrus-growing areas of the Mediterranean Basin in risk management planning in the case of the spread of HLB vectors or a possible introduction of the disease. Additionally, it may be a valuable resource to inform opinion dynamic models, enabling the identification of pivotal factors for the success of control measures.

KEYWORDS
Huanglongbing, Spain, area-wide pest management, Diaphorina citri, Trioza erytreae, regionalization, ClustGeo
1 Introduction

Citrus huanglongbing (HLB), or citrus greening, is considered the most devastating bacterial disease of citrus worldwide (Bove, 2006; McCollum and Baldwin, 2016; Jeger et al., 2023). The disease affects citrus trees, causing an overall decline which eventually leads to their death (Hendrichs et al., 2021). HLB is mainly associated with the phloem-limited bacteria Candidatus Liberibacter asiaticus (CLas), Ca. L. africanus (CLaf) and Ca. L. americanus (CLam). The CLas species is the most aggressive form of HLB and is widespread in Asia, North America, South America and Africa (Bove, 2006; Gasparoto et al., 2022). CLaf is only present in Africa, while CLam, initially identified in Brazil, is diminishing and being taken over by CLas (Teixeira et al., 2005; Gasparoto et al., 2012; Wang, 2020). HLB-associated Ca. Liberibacter spp. (CLs) can be transmitted by graft propagation (Bove, 2006), but their natural spread is primarily mediated by two psyllid vectors: the Asian citrus psyllid (ACP), Diaphorina citri Kuwayama (Hemiptera: Psyllidae), and the African citrus psyllid (AfCP), Triozoa erytreae (Del Guercio, 1918) (Hemiptera: Trioziidae), native to the Asian and African continents, respectively (Teixeira et al., 2005; Bove, 2006; da Graça et al., 2016; Alquezar et al., 2022).

The pathosystem of HLB involves the pathogen, the plant host and the psyllid vector, interacting in an environmental background that affects the biology of each component and its interactions (Stelinski, 2019). The three species of the genus Ca. Liberibacter associated with HLB have a persistent propagative relationship with these psyllid vectors (Carmo-Sousa et al., 2020). As these bacterial pathogens are confined to the phloem, acquisition happens when these psyllid vectors feed on the phloem sap of infected citrus plants, with inoculation occurring via salivation after invading the salivary glands. After acquisition, depending on the psyllid species, transmission can occur within 7-12 days (Canale et al., 2017). Infected adult and/or nymph vectors transmit the bacterium when feeding on healthy flushes, while emerging nymphs acquire it when feeding of them (da Graça et al., 2016; Shimwela et al., 2016). Once the psyllid vector becomes inoculative (able to inoculate plants), it retains this ability for weeks or its whole life, characterizing the transmission as persistent. The HLB-associated bacteria then develop in the tree phloem vascular system, and symptoms take at least four months to appear depending on the tree’s age (Coletta-Filho et al., 2010). The unpredictable lag period between the acquisition of the pathogen and transmission by the psyllid vectors, and for the development of visual symptoms, make HLB eradication by the visual detection of HLB-symptomatic trees problematic.

The CLas-AfCP combination is widespread in Asia and America whereas CLaf-AfCP is the most prevalent in Africa (Alquezar et al., 2022). CLas is more heat tolerant than CLaf and can develop under relatively low humidity and high temperature conditions. Similarly, AfCP is able to adapt to these same climatic conditions. Transmission of CLas is closely linked to insect development, where acquisition must occur with a higher efficiency during the nymphal stage (Inoue et al., 2009; Coletta-Filho et al., 2014). Thus, temperature ranges influence transmission rate and disease spread. For instance, in the USA, lower temperatures in California have been suggested to slow down the life cycle of ACP resulting in lower infection rates, whereas the higher temperature range in Florida is more favorable for ACP mating and continued disease progression (McRoberts et al., 2021; Igwe et al., 2022; Hosseinzadeh and Heck, 2023). Conversely, both CLaf and AfCP are sensitive to hot and dry conditions, and are thus restricted to areas with relatively cool and humid environments (Catling, 1969; Liu and Tsai, 2000; Bove, 2006; Dala-Paula et al., 2019). Both AfCP and AfCP can transmit CLas and CLaf under experimental conditions (Reynaud et al., 2022), but more research is needed to better understand the epidemiological implications of the CLas/AfCP and CLaf/AfCP combinations.

Nowadays, HLB is widespread in some of the major citrus producing areas worldwide, such as Argentina, Brazil, China, India, Mexico, South Africa and the USA, among others. No therapeutic treatments or citrus cultivars resistant to HLB are currently available. A three-pronged strategy (TPS), consisting of (i) planting of certified healthy citrus material; (ii) removal of inoculum sources-infected trees; and (iii) application of insecticide treatments to control psyllid populations, has been proposed to be key to managing citrus HLB (Bassanezi and Gottwald, 2009; Bove, 2012). Experiences in both Brazil and the USA have highlighted the futility to control HLB without removing infected trees (Alquezar et al., 2022). Consequently, the task of controlling psyllid vectors goes beyond managing current HLB outbreaks; it encompasses the critical mission of preventing future infections. By diminishing the population of infected trees and thereby reducing the reservoir of the bacterium, vector control becomes an essential part of the long-term management of this disease. In essence, vector control is a cornerstone of HLB management because it directly addresses the mode of disease transmission.

Nevertheless, effective control of vector populations can only be achieved through area-wide pest management (Gottwald et al., 2007; Bassanezi et al., 2013). This strategy consists in implementing pest management over extensive areas as a group rather than on individual orchards (Klassen, 2008). It therefore needs the coordination and teamwork of the whole industry including growers, advisers and governmental authorities, among others. This type of management is particularly effective for migratory pest species such as some insect vectors (Dalal et al., 2017; Drmić et al., 2017; Singerman et al., 2017; García-Figuera et al., 2022; International Atomic Energy Agency (IAEA), 2022). Successful cases of area-wide pest management consist of time-coordinated pesticide sprays or other means of control applied simultaneously in an extensive area (Bassanezi et al., 2013; Navarro-Llopis et al., 2014; Garcia-Figuera et al., 2021b; International Atomic Energy Agency (IAEA), 2022).

Area-wide pest management has been attempted for CLas/AfCP in Brazil, China, Mexico, Argentina, California, Florida and Texas with different aims (Bassanezi et al., 2013; McCollum and Baldwin, 2016; Singerman et al., 2017; McRoberts et al., 2019; García-Figuera et al., 2021b; Hendrichs et al., 2021; Yuan et al., 2021; Alquezar et al., 2022; García-Figuera et al., 2022). In areas where both the vector and the disease are widespread, area-wide management aims to reduce disease intensity and sustain fruit production by combining vector management with the elimination of affected
The Mediterranean Basin accounts for around 12% of the world’s citrus-growing areas (McCollum and Baldwin, 2016; Siverio et al., 2017; Food and Agriculture Organization of the United Nations (FAO), 2020). Citrus in the Mediterranean are cultivated mainly in coastal areas, forming a virtually continuous belt along the basin. Spain has the largest citrus-growing area in the Mediterranean Basin, with ~23% of the total surface area (Food and Agriculture Organization of the United Nations (FAO), 2020). In Spain, commercial citrus crops cover nearly 300,000 ha with a production of ~6.6 million tonnes, making it the main producer in the EU and the sixth in the world (Siverio et al., 2017; Food and Agriculture Organization of the United Nations (FAO), 2020). With a 25% share, Spain is the world’s leading exporter of fresh citrus, with nearly 60% of the production destined to foreign markets. The main citrus region in Spain is the Valencian Autonomous Community, with 54% of the total Spanish citrus-growing area, followed by Andalusia (28%) and the Region of Murcia (13%) (Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2019; Food and Agriculture Organization of the United Nations (FAO), 2020; Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2022). Furthermore, citrus in Spain also has an important sociocultural component and ornamental citrus are widely planted in residential areas (Duarte et al., 2016; Arenas-Arenas et al., 2018; Torregrosa et al., 2019).

The Mediterranean Basin countries are currently free of HLB (da Graça et al., 2016; McCollum and Baldwin, 2016; European and Mediterranean Plant Protection Organization (EPPO), 2022a; European and Mediterranean Plant Protection Organization (EPPO), 2022b; European and Mediterranean Plant Protection Organization (EPPO), 2022c; Garcia-Figuera et al., 2022). However, in the European Union (EU), T. erytreae is currently spreading in the Iberian Peninsula and getting closer to the commercial citrus-growing areas of Spain and Portugal (Arenas-Arenas et al., 2019; Benhadi-Marin et al., 2020), and D. citri has recently been reported in Israel (European and Mediterranean Plant Protection Organization (EPPO), 2022) and Cyprus (European and Mediterranean Plant Protection Organization (EPPO), 2023). As indicated by Bové (2006), the introduction of the vector usually precedes subsequent outbreaks of HLB. In fact, during recent years the time lag between vector introduction and HLB outbreaks has been progressively reduced (Alquézar et al., 2022).

The HLB-associated CLs and their psyllid vectors ACP and AfCP have quarantine status in most countries of the Mediterranean Basin, including the EU, where CLs causing HLB are also included in the list of priority pests (European Union (EU), 2019a; European Union (EU), 2019b; Food and Agriculture Organization of the United Nations (FAO), 2022). These EU priority pests are the ones with the highest potential socioeconomic impact. They are subjected to additional mandatory phytosanitary measures including annual surveys, contingency and action plans, simulation exercises, and public information (European Union (EU), 2016; Aragón et al., 2022). Despite the phytosanitary regulations, over the last few years several interceptions of ACP at the EU borders have been reported. Furthermore, irregular imports of propagating plant material constitute another high-risk entry pathway (EFSA PLH Panel et al., 2021). This fact, together with the presence of AfCP in the Iberian Peninsula and ACP in Israel and Cyprus, represents a serious threat to the citrus-growing areas in the Mediterranean Basin (Gutierrez and Ponti, 2013).

The Mediterranean Basin is currently free of the main citrus-growing areas from certified nurseries (Bassanezi et al., 2013; Singerman et al., 2017; García-Figuera et al., 2021b; Alquézar et al., 2022). However, in areas where the vector is widespread but the disease is still geographically restricted, the aim of area-wide management is to limit the spread of the disease to neighboring areas (Grafton-Cardwell and García-Figuera, 2018; Garcia-Figuera et al., 2021b; Yuan et al., 2021; Alquézar et al., 2022).

Citrus-growing areas in Australia, New Zealand and the Mediterranean Basin are also included in the list of priority pests (European Union (EU), 2019b; Food and Agriculture Organization of the United Nations (FAO), 2022). These EU contingency plans were recently evaluated in a simulation exercise (Aragon et al., 2022) and have been translated to a specific regulation to address the control and eradication of the AfCP and to prevent the introduction of the ACP and CLs (Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2023b). Nevertheless, how to organize citrus growers in order to coordinate vector control through area-wide management is not specifically addressed in this contingency plan. The particularities of the citrus industry in the Mediterranean Basin, with small orchards (<1 ha on average in Spain) often managed independently by individual growers (Reig-Martínez and Picaizo-Tadeo, 2004; Fernández-Zamudio et al., 2005; Morell-Monzó et al., 2021), increases their exposure to primary infections, namely those coming from neighboring areas (Bassanezi et al., 2013; Alquézar et al., 2022). The different characteristics of the citrus industry in the Mediterranean Basin limit the extrapolation of the area-wide management programs designed for other citrus-growing areas, thereby making the implementation of area-wide management a challenge. Moreover, area-wide management has not been attempted for AfCP control elsewhere, despite being the vector currently spreading in the Iberian Peninsula.

Considering that ACP and AfCP are now present in the Mediterranean Basin and to prepare for its potential spread to the main citrus-growing areas in the region, the objectives of this study are:

i. To define an area-wide management framework for ACP and AfCP in the citrus-growing areas in the Mediterranean Basin. To this end, data from the Valencian Autonomous Community, Spain, and a series of risks potentially associated with the establishment and spread of ACP/ AfCP were considered (Flores-Sánchez et al., 2017; McRoberts et al., 2019).

ii. To optimize the size and location of area-wide management areas. For this purpose, the methodology developed by Gottwald et al (2014a; 2014b) was tailored by adapting a regionalization algorithm and assessing its performance under different configurations of risk homogeneity and spatial aggregation. The methods were implemented in R language (R Core Team, 2021) and, following open science
principles, all the underlying code and data necessary to run the tool are made available. This will allow the approach to be updated as new information becomes available, so that end users can adapt it to different epidemiological settings.

In addition, this proposal could provide valuable support to policy makers and the citrus industry in planning for risk management in the event of the spread of HLB vectors or the potential introduction of the disease into citrus-growing regions in the Mediterranean Basin. It can also be used by modelers to improve surveillance and management programs. By integrating epidemiological models with opinion dynamics models that incorporate the principles of area-wide management, it will be possible to determine how best to maximize the potential effectiveness of voluntary HLB control campaigns (Gottwald et al., 2020).

2 Materials and methods

The Valencian Autonomous Community (ES52 NUTS level 2 EU territorial unit for statistics) was selected as the study area. This region consists of three provinces: Alicante (ES521 NUTS 3), Castellón (ES522 NUTS 3) and Valencia (ES523 NUTS 3). The study area is the region with the largest extension dedicated to citrus in Spain, with characteristics similar to those of other citrus-growing areas in the Mediterranean Basin. Of the three provinces, Valencia has the biggest area devoted to citrus.

The resulting area-wide management zones in our study were defined as Pest Management Areas (PMAs). To facilitate regionalization of the PMAs, the study area was gridded into 24,048 cells of 1 km² using the European Environment Agency (EEA) reference grid downloaded from its official website. This grid is in the ETRS89 Lambert Azimuthal Equal Area (LAEA) projection coordinate reference system (ETRS89/ETRS-LAEA (EPSG:3035)) with a coding system identifying each grid cell with a unique ID (Peifer, 2011).

The original 24,048-cell grid was filtered to identify cells with commercial citrus coverage using a georeferenced dataset from the Spanish Agricultural Plots Geographic Information System (SIGPAC). The SIGPAC dataset provided comprehensive information on the size and shape of commercial citrus orchards and was intersected with the reference grid of the study area. Finally, 7,021 cells of 24,048 cells were identified as cells with commercial citrus coverage. Thus, the grid of the study area was defined by 7,021 cells of 1 km².

Based on previous studies in other citrus areas (Diaz-Padilla et al., 2014; Gottwald et al., 2014a; Gottwald et al., 2014b), the study area was characterized through different risk factors than can affect the introduction and spread of HLB vectors and, consequently, the disease epidemiology (see Supplementary Material A for further details). For this purpose, different georeferenced datasets were exploited to extract data on climatic variables, commercial citrus orchards, population census data and the network of the main transportation corridors. The extracted data were transformed to the ETRS89/ETRS-LAEA (EPSG:3035) reference system using R software version 3.6.0, https://www.R-project.org and aggregated at the cell level into the grid of the study area.

2.1 Climatic variables

Hourly data at 2 m for air temperature (K) and dew point temperature (K) at a spatial resolution of 0.1° x 0.1° (∼ 9 km x 9 km) were retrieved from ERA5-Land dataset from 2009 to 2018 in raster format through the Climate Data User (CDS) user interface.

As in Galvañ et al. (2022), relative humidity (RH) as a percentage was calculated according to Wallace and Hobbs (2006) using air and dew point temperatures as follows:

\[
RH = 100 \cdot \left( \frac{e(Td)}{e(T)} \right)
\]

with \(e(Td)\) and \(e(T)\) denoting actual and saturation vapor pressures in hPa, respectively, with \(Td\) as the dew point temperature and \(T\) as the air temperature both in °C and for \(e(Td) \leq e(T)\). Actual and saturation vapor pressure were estimated following Bolton (1980) as:

\[
e(Td) = 6.12 \cdot \exp \left[ \frac{17.67 \cdot Td/(243.5 + Td)}{243.5 + Td} \right]
\]

\[
e(T) = 6.12 \cdot \exp \left[ \frac{17.67 \cdot T}{243.5 + T} \right]
\]

Daily maximum, minimum, mean temperature and minimum relative humidity were computed from the hourly variables using the function `apply.daily()` from the R package rts (Babak, 2023) and were extracted to the 1 km² grid of the study area by means of the function `extract()` from the R package raster (Hijmans, 2022). A total of 326 cells were excluded because ERA5-Land does not provide climatic data for land areas adjacent to the coast line (Pelosi et al., 2020) and so the final grid of the study area was defined by 6,695 cells.

2.2 Commercial, abandoned and organic citrus

Georeferenced data of organic and abandoned orchards were not available. Thus, the total surface area of organic and abandoned citrus orchards in the study area was obtained from official statistics (Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2019; Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2020). Based on these data, a total of 42.68 km² of citrus under organic production and 103.81 km² of abandoned citrus were randomly assigned to the commercial citrus orchards extracted from the SIGPAC and were first used to filter the grid of the study area. This was done using the function `sample()` from the R package base (Core Team, 2021). The number of organic/abandoned orchards was determined by dividing the total number of hectares of organic/abandoned by the average commercial orchard size. Consequently, each individual orchard was categorized as in-production, distinguishing between conventional vs. organic management, or abandoned, that is, unmanaged.
In order to characterize each of the cells comprising the study area’s grid in terms of the total surface area under commercial production, including conventional, organic, and unmanaged citrus areas, the ‘st_intersection()’ function from the R package sf was employed (Pebesma et al., 2023). Through this function, the study area grid was intersected with the georeferenced orchard database, and the total area for each type of orchard was computed by aggregating the surfaces of each orchard intersected by a cell.

2.3 Residential citrus

Two sources of data were exploited for residential citrus and subsequently combined: i) daily minimum temperatures, and ii) population census data (see Section 2.1 for further details).

To account for the detrimental effect of cold conditions in citrus growth, daily minimum temperatures in January were averaged for each year from 2009 to 2018. Based on Davies and Albrigo (1999) and Bijzet (2006), a temperature threshold of 2.5°C was chosen to draw an isotherm defining those parts of the study area in which the growth of citrus is not limited by cold temperatures 𝑇 ≥ 2.5°C.

Population census data at country level (ES NUTS 1) were extracted from the GEOSTAT population grid, which provides data on population density as the number of inhabitants per 1 km² cell at the EEA reference grid (EUROSTAT, 2016). The ratio between the residential citrus coverage and the number of inhabitants was estimated based on the census of citrus trees for the city of Seville (Cabanillas, 2020), in Andalusia, and its population (Instituto Nacional de Estadística (INE), 2020). This city was chosen for having the most complete and updated census on residential citrus trees in Spain. The number of residential citrus trees in Seville was transformed to residential citrus area, assuming a coverage area of 20 m² per tree. The estimated residential citrus area was divided by the total number of inhabitants, resulting in a ratio of 1.33·10⁻⁶ km² of residential citrus area per inhabitant.

This ratio was applied to the number of inhabitants in each 1 km² cell, but considering only those cells defined by the isotherm as suitable for citrus growth. Finally, only those cells classified as “suitable” within the grid of the study area were taken into account. As a result, the information of the residential citrus area was integrated into the study area cells, along with the information of the total citrus area in conventional and organic production and the “abandoned” citrus area.

2.4 Transportation corridors

The road transport network at country level (ES NUTS 1) was downloaded from the National Geographic Information Centre (CNIG) of Spain (Centro Nacional de Información Geográfica (CNIG), 2021). The main transportation corridors were extracted from this dataset choosing only motorways and highways in the study area, as they are considered the primary routes for the transport of citrus fruit and planting materials.

2.5 Pest management areas (PMAs)

2.5.1 Risk factors

CLIMATIC SUITABILITY. The risk factor associated with climatic suitability was calculated separately for ACP and AfCP. As with most insects, the psyllid vectors of HLB are strongly influenced by climatic conditions, particularly temperature and relative humidity. Two risk components related to the potential number of generations (R_{pg}) and the number of favorable days for development (R_{nfd}) were considered for both vectors (see Supplementary Material A for further details). To estimate these two risk components, daily data of temperature and relative humidity for the period 2009-2018 were used (see Section 2.1).

The potential number of generations per year R_{pg,i} for each cell i was calculated based on the average cumulative degree-days for the 10-year period (2009-2018) and dividing by the degree-days (DD) required for an egg of ACP/AfCP to become an adult above a developmental threshold temperature \( T_i \) as follows:

\[
R_{pg,i} = \frac{\sum_{j=1}^{nyears} T_{mean_{day,j}} \cdot (T_{min_{day,j}})}{DD} \quad (4)
\]

where \( T_{mean_{day,j}} \) denotes the daily mean temperature for day \( j \) of year \( k \) within cell \( i \) with \( j = 1, \ldots, 365, k = 1, \ldots, 10 \) and \( n_{days(years)} \) denoting the total number of days and years considered, respectively.

For ACP, \( T_i = 10.45°C \) and \( DD = 249.88°C \) were the values selected according to Liu and Tsai (2000). For AfCP, \( T_i = 10°C \) and \( DD = 270.00°C \) based on the work of Catling (1973) and Aidoo et al. (2022).

The number of favorable days \( R_{nfd,i} \) for each cell i was calculated considering only those days allowing the development of all the stages of the insect. For ACP, “a favorable day” (FD_{acp}) was defined considering that this insect species is able to develop with mean temperatures \( (T_{mean}) \) between 10°C, as a lower threshold (\( T_l \)), and 33°C, as an upper threshold (\( T_u \)) (Liu and Tsai, 2000):

\[
FD_{acp} = \begin{cases} 1 & \text{if } T_{mean} > 10°C \text{ and } T_{mean} < 33°C \\ 0 & \text{otherwise} \end{cases} \quad (5)
\]

According to Aidoo et al. (2022), a “favorable day” for AfCP (FD_{afcp}) was defined considering that it is able to develop at mean temperatures \( (T_{mean}) \) between 10°C, as \( T_l \) and 27°C, as \( T_u \), and also including a saturation deficit (SD) \( \leq 32.1 \text{ mmHg} \) (Catling, 1969; Catling, 1972). This insect species is particularly sensitive to the effects of high temperature and low relative humidity, which are combined in the SD index.

\[
FD_{afcp} = \begin{cases} 1 & \text{if } T_{mean} < 10°C \text{ and } T_{mean} > 27°C \text{ and } SD < 32.1 \text{ mmHg} \\ 0 & \text{otherwise} \end{cases} \quad (6)
\]

SD in hPa was estimated following Murray (1967) as,

\[
SD = \frac{(100 - RH_{min})}{100} \cdot e(T_{max}) \quad (7)
\]
where $RH_{\text{min}}$ and saturation vapor pressure ($e_s$) were estimated from hourly $RH$ following equations 1 and 2, respectively, but considering maximum daily temperature ($T_{\text{max}}$). To convert SD from hPa to mmHg, the estimates were multiplied by 0.7501.

Thus, the $R_{\text{adj},i,j}$ for each cell $i$ was computed as:

$$R_{\text{adj},i,j} = \sum \frac{FD_{\text{day},\text{year}}}{n_{\text{year}}},$$

where $FD_{\text{day},\text{year}}$ is an indicator variable of a “favorable day” for day $j$ of year $k$ within cell with $j = 1,2,\ldots,365$, $k = 1,2,\ldots,10$ and $n_{\text{year}}$ denoting the total number of years.

The psyllids ACP and AfCP lay eggs on leaf flushes and their immature stages feed on tender foliar tissues but not on mature leaves (Annecke and Cilliers, 1963; Moran and Buchan, 1975; van den Berg and Deacon, 1988; Hall, 2020; Monzó and Vanaclocha, 2023). Therefore, daily climate variables were filtered considering only the days from 15 February to 30 April, June and October, assuming that under Mediterranean conditions these are the three major leaf flushing periods (Garcia-Marí et al., 2002) (See Supplementary Material A for further details). Figures S1A, S1B (Supplementary Material A) describe the potential number of generations and the favorable days estimated for the study area.

COMMERCIAL CITRUS AREA. The risk factor associated with the commercial citrus area for each cell $i$, $R_{\text{com},i}$ was calculated as:

$$R_{\text{com},i} = CCA_i,$$

where $CCA_i$ denotes the commercial citrus area in km² for each cell $i$.

ABANDONED/ORGANIC ORCHARDS. The risk factors associated with abandoned and organic orchards for each cell $i$, $R_{\text{abn/organic,i}}$ were computed considering the presence of abandoned or organic orchards within a cell $i$ and/or in its $j$ neighboring cells. The baseline risk associated for each cell $i$, $R_{\text{abn/organic},i}$ was computed as:

$$R_{\text{abn/organic},i} = \left\{ \begin{array}{ll} CCA_i & \text{if } A(O)CA_i = 0 \\ CCA_i \cdot F_{\text{abn/organic}} & \text{if } A(O)CA_i \neq 0 \end{array} \right.,$$

where $CCA_i$ and $A(O)CA_i$ denote the commercial and abandoned/organic citrus area in km² for each cell $i$, respectively; $F_{\text{abn/organic}}$ is an amplification coefficient that increases the risk in those cells $i$ in which abandoned/organic citrus orchards are present. The values for this coefficient were set based on the study by Martini et al. (2016). In this work, the overwintering abundance of adult ACP was evaluated in commercial citrus orchards under conventional or organic management as well as those that have been abandoned. The amplification coefficient was then calculated based on the average abundance of adults in organic (3) and abandoned orchards (19) relative to that in commercial orchards (0.50) as $F_{\text{abn}} = 38 = 19/0.50$ and $F_{\text{org}} = 6 = 3/0.50$. In the absence of this type of information for AfCP, the same values of ACP were used.

The $R_{\text{abn/organic},i}$ for each cell $i$ was calculated as:

$$R_{\text{abn/organic},i} = R_{\text{abn/organic},i} + \sum_{j=1}^{max} R_{\text{abn/organic},j},$$

where cell $i$ and each cell $j$ were considered neighbors if the distance between their centroids did not exceed 1,500 m based on van den Berg and Deacon (1988) and Lewis-Rosenblum et al. (2015) (see Supplementary Material A for further details). Thus, the maximum number of neighbor cells $j$ per cell $i$ was 8.

RESIDENTIAL CITRUS. The risk factor associated with the proximity of residential citrus ($R_{\text{res},i}$) for each cell $i$ was calculated following the same approach used for $R_{\text{abn/organic},i}$. The presence of residential citrus within a cell $i$ and/or the proximity of cells with residential citrus was considered using the same criterion of proximity between cells and, thus, the baseline risk associated $R_{\text{res},i}$ for each cell was computed as:

$$R_{\text{res},i} = \left\{ \begin{array}{ll} CCA_i & \text{if } RCA_i = 0 \\ CCA_i \cdot F_{\text{res}} & \text{if } RCA_i \neq 0 \end{array} \right.,$$

where $CCA_i$ and $RCA_i$ denote the area (%) covered by commercial and residential citrus for each cell $i$, respectively. Due to the lack of information, $F_{\text{res}}$ was set as 6, assuming that the presence/proximity of residential citrus would affect the introduction and spread of ACP and AfCP in a similar way to the presence/proximity to organic orchards.

The $R_{\text{res},i}$ for each cell $i$ was calculated as:

$$R_{\text{res},i} = R_{\text{res},i} + \sum_{j=1}^{max} R_{\text{res},j} \cdot F_{\text{res}},$$

TRANSPORTATION CORRIDORS. The spread of ACP and AfCP can be influenced by human activities, such as the transportation of infested plants and fruits. Trucks transporting citrus fruit from orchards to packing houses have been identified as a means of HLB-vector spread (Halbert et al., 2010). In our case, the motorways and highways within the study area were considered the main transportation corridors for commercial citrus fruit and planting material. The risk factor of transportation corridors, $R_{\text{tra},i}$, was estimated considering the presence of this type of roads within a cell $i$ and/or their close proximity. As in the other risk factors, the criterion of proximity between cells was also established considering a maximum distance between centroid cells $d_{i,j} \leq 1,500$ m. The associated baseline risk $R_{\text{tra},i}$ for each cell was computed as:

$$R_{\text{tra},i} = \left\{ \begin{array}{ll} CCA_i & \text{if } I_{\text{tra},i} = 0 \\ CCA_i \cdot F_{\text{tra}} & \text{if } I_{\text{tra},i} \neq 0 \end{array} \right.,$$

where $CCA_i$ denotes the area (%) covered by commercial citrus and $I_{\text{tra},i}$ is an indicator variable related to the presence/proximity to transportation corridors. The $I_{\text{tra},i}$ was calculated as:

$$R_{\text{tra},i} = R_{\text{tra},i} + \sum_{j=1}^{max} R_{\text{tra},j} \cdot F_{\text{tra}}.$$
**RISK FACTOR NORMALIZATION.** The values of the risk factors were scaled from 0 to 1 using the min-max normalization. For each risk factor, its minimum value was transformed into 0 and its maximum value into 1 and, thus, every other value was transformed into a decimal between 0 and 1 as follows:

$$R'_i = \frac{R_i - \min(R)}{\max(R) - \min(R)} \quad (16)$$

Here, $R_i$ is the risk value associated with a cell $i$, and $\min(R)$ and $\max(R)$ are the minimum and maximum values of the risk factor $R$, respectively, which define the range of this risk factor $R$ within the study area.

**OVERALL RISK.** To compute the overall risk, $OR$, for each cell $i$, two approaches were followed. In the absence of information on the relative importance of each risk factor, the first one consisted in averaging them:

$$OR_{1ij} = \frac{1}{M} \sum_{m=1}^{M} R^*_m$$

$$= \frac{1}{M} (R^*_{}\text{apc} + R^*_{}\text{afcp} + R^*_{}\text{org} + R^*_{}\text{abn} + R^*_{}\text{ned} + R^*_{}\text{nfd}).$$

(17)

When information is available from epidemiological studies or expert consultancy, a second approach may be used, assigning a weight ($w_m$) to each of the $m$-th normalized individual risks and calculating a weighted average such as follows:

$$OR_{2ij} = \sum_{m=1}^{M} w_m R^*_m \quad \text{with} \quad \sum_{m=1}^{M} w_m = 1$$

$$= (w_1 \cdot R^*_{}\text{apc} + w_2 \cdot R^*_{}\text{afcp} + w_3 \cdot R^*_{}\text{org} + w_4 \cdot R^*_{}\text{abn} + w_5 \cdot R^*_{}\text{nfd} + w_6 \cdot R^*_{}\text{nfd} + w_7 \cdot R^*_{}\text{nfd}),$$

(18)

with $w_1 = w_2 = 0.07, w_3 = 0.43, w_4 = 0.21, w_5 = 0.11$ and $w_7 = 0.04$. In our case, those values were set considering a higher weight for the risks associated to abandoned and organic orchards. The two above-mentioned approaches were estimated for both ACP and AfCP. Consequently, $OR_1$ and $OR_2$ were separately determined for ACP and AfCP, respectively. The variations in these estimates are due to the differences in the biology of these two insect vectors, which affect the calculation of $R^*_{\text{apc}}$ and $R^*_{\text{afcp}}$ as justified above in the section **CLIMATIC SUITABILITY.** The overall risk factors were also scaled from 0 to 1 following min-max normalization as described above.

### 2.5.2 PMAs size and location: regionalization algorithm

The Ward-like hierarchical clustering algorithm from the R package ClustGeo (Chavent et al., 2021) was used to set the size and location of the PMAs. This algorithm makes it possible to include spatial constraints optimizing the convex combination of two dissimilarity matrices $D_0$ and $D_1$ and a mixing parameter $\alpha \in [0,1]$, $D_{mix} = (1-\alpha)D_0 + \alpha D_1$. The $D_0$ matrix captures the dissimilarities in the “feature space” (i.e., estimated from the overall risk) whereas the $D_1$ gives the dissimilarities in the “constraint space” (i.e., estimated from the distance between observations). The parameter $\alpha$ quantifies the relative importance of $D_0$ as compared to $D_1$ in the clustering procedure. Specifically, when $\alpha = 0$, the spatial dissimilarities are not taken into account and when $\alpha = 1$, the overall risk distances are not considered and PMAs are estimated considering only the spatial distances (Chavent et al., 2018; Camelo Aguiar et al., 2020). In this respect, the package ClustGeo also includes a functionality to address the choice of the parameter $\alpha$.

Risk factors were calculated for the Valencian Autonomous Community (ES52 NUTS level 2), but the proposed methodology for PMA delineation was evaluated in the two most representative counties (Institut Valencià d’Estadística (IVE), 2022) within the study area in terms of commercial citrus areas. These two counties, “Camp de Túria” and “Ribera Alta”, are located in the province of Valencia (ES52 NUTS 3). A county (i.e., a comarca in Spanish) is a territorial division legally defined by each autonomous community in Spain (NUTS 2 regions) that groups together several municipalities [LAUs, (EUROSTAT, 2021)] within NUTS 3 regions with the aim of providing common public services. PMAs were set at county level owing to the smallholding structure of citrus in the study areas as well as the potential role of agricultural cooperatives to coordinate programs for pest control (Fernández-Zamudio et al., 2005; Melià-Martí, 2021), which operate mainly at county level.

The methodology was evaluated separately for the two overall risks described above and both insect vectors, resulting in 4 scenarios. In particular, PMAs regionalization for ACP and AfCP management was developed on the two overall risks computed ($OR_{1,\text{apc}}, OR_{1,\text{afcp}}$ and $OR_{2,\text{apc}}, OR_{2,\text{afcp}}$). The algorithm was run independently for both counties under the following conditions:

- $D_0$ matrix was computed from the Euclidean distance between the $n$ cells performed for the overall risk variable.
- $D_1$ matrix was computed from the shortest Euclidean distance in geographical space between the $n$ cell centroids.
- The choice of the maximum number $K$ of PMAs (i.e., clusters) was defined assuming an optimum PMA size of 25 cells, i.e., by dividing the total number of cells within a county by 25. Those counties with fewer than 25 cells of citrus were not considered in the regionalization process.

The performance of the regionalization method was evaluated by a sensitivity analysis defined on the basis on three values of the mixing parameter: $\alpha = 0.1, \alpha = 0.9$ and the optimal value suggested by Clustgeo that balances the loss of risk homogeneity and the gain in spatial aggregation. The choice of $\alpha = 0.1$ underscored the significance of risk distances, whereas $\alpha = 0.9$ highlighted the importance of spatial distances in the regionalization procedure. The exploration of various $\alpha$ values was aimed to provide a thorough evaluation of how this parameter influences clustering outcomes. Despite the package ClustGeo includes a functionality to address the choice of the parameter $\alpha$, it was proposed to explore these two extreme values, $\alpha = 0.1$ and $\alpha = 0.9$, against the proposed “optimal value” to indentify the value that makes the most practical sense.
Several internal validation measures were calculated to quantify compactness and/or separation (Leskovec et al., 2014; Kassambara, 2017). Compactness measures how close the objects are within the same PMA, estimated by the complete diameter as the distance between the two most remote cells within the same PMA and summarized by its maximum value. Separation measures how well separated a PMA is from the others, estimated by means of the complete linkage distance as the distance between the most remote cells belonging to two different PMAs and summarized by its minimum value. The Dunn index \((D)\) (Dunn, 1973; Dunn, 1974) was estimated from the max(complete.diameter) and the \(\min(\text{complete.linkage.distance})\) as follows:

\[
D = \frac{\min(\text{complete.linkage.distance})}{\max(\text{complete.diameter})}.
\]

Thus, those values of \(\alpha\) resulting in compact and well-separated PMAs will have a higher value of the Dunn index. The above-mentioned internal validation measures were computed using the R package clv (Nieweglowski, 2020). Compactness and separation were also evaluated visually by comparing overall risk distribution between PMAs using boxplots. Additionally, the normalized proportion of pseudo-inertia (Chavent et al., 2018) was computed for each \(\alpha\) value to quantify the loss of risk homogeneity and spatial aggregation in relation to the reference configurations \(\alpha = 0\) (i.e., no spatial information used to define PMAs and risk homogeneity is assumed to be 1 (100%)) and \(\alpha = 1\) (i.e., only spatial information used to define PMAs and spatial aggregation is assumed to be 1 (100%)), respectively.

3 Results

3.1 Description of study area

The study area was mapped onto a grid of 6,695 cells in accordance with the EEA reference grid, of which 3,550 (53.02%) are in the province of Valencia, 1,835 (27.41%) are in the province of Alicante and the remaining 1,310 (19.57%) are located in the province of Castellón. The study area has a total of 1,635.53 km\(^2\) of commercial citrus, with the largest area in Valencia, with 943.95 km\(^2\), followed by the province of Castellón, with 354.10 km\(^2\) and the remaining 1,310 (19.57%) are located in the province of Alicante, with 354.10 and 337.82 km\(^2\), respectively. Of the 1,635.53 km\(^2\) (100%) of commercial citrus-growing area in the province of Valencia are distributed into 862.94 km\(^2\) (91.42%) under conventional management, 23.47 km\(^2\) (2.49%) are organic and 57.55 km\(^2\) (6.10%) are abandoned. Of the 337.48 km\(^2\) of commercial citrus-growing area in Alicante, 307.74 km\(^2\) (91.19%) are under conventional management, 8.62 km\(^2\) (2.55%) are organic and 21.12 km\(^2\) (6.26%) are abandoned. A total of 2.95 km\(^2\) of residential citrus were estimated for the whole study area, with 0.49 km\(^2\) (16.43%) in the province of Castellón, 1.84 km\(^2\) (62.49%) in the province of Valencia and 0.62 km\(^2\) (21.08%) in the province of Alicante (Table 1).

3.2 Overall risk

The seven risk factors were summarized by two overall risk factors, \(OR_1\) and \(OR_2\), for ACP and AfCP in the study area (Figures 1, 2). In general, \(OR_2\) presented lower values than \(OR_1\) for both vectors, as observed in Figures 2B, D, F and 3B, D, F, where a greater predominance of cooler colors is observed in relation to Figures 2A, C, E and 3A, C, E. Figure 4 describes the range of the four distributions capturing \(OR_1\) and \(OR_2\) for both ACP and AfCP and shows the same trend as that noted above.

The distributions associated with \(OR_1\) (Figures 4A, B) have a smaller interquartile range, \(IQR = Q_3 - Q_1\), than those associated with \(OR_2\) (Figures 4C, D) for both insect vectors (Figures 4A, C for ACP and Figures 4B, D for AfCP). Hence, the differences between cells for \(OR_1\) risk are smaller than for \(OR_2\) risk. The median values \((\text{Md}_n)\) for \(OR_1\) are higher than those corresponding to \(OR_2\). However, for both overall risks, the differences between the distributions for ACP and AfCP are minor. Specifically, for ACP: \(\text{Md}_n\_\text{ACP} = 0.44\) and \(\text{Md}_n\_\text{AfCP} = 0.30\) (Figures 4A, C), whereas for AfCP: \(\text{Md}_n\_\text{ACP} = 0.43\) and \(\text{Md}_n\_\text{AfCP} = 0.30\) (Figures 4B, D). Likewise, the means of the distributions \((\mu)\) for ACP are \(\mu\_\text{ACP} = 0.47\) and \(\mu\_\text{AfCP} = 0.35\) (Figures 4A, C), whereas for AfCP \(\mu\_\text{ACP} = 0.42\) and \(\mu\_\text{AfCP} = 0.33\) (Figures 4B, D).

The overall risk factor trends described above for the study region are also reproduced at provincial level, when comparing between the two overall risk factors, \(OR_1\) vs. \(OR_2\), as well as between the two insect vectors, ACP and AfCP. The province of Alicante showed the smallest differences between cells in terms of overall risk values, with an IQR lower than the other two provinces. The province of Valencia, which has the largest percentage of the citrus area and therefore the highest number of cells (Table 1), presents a trend similar to that of the whole study area (Figure 4).

3.3 PMAs

3.3.1 Diaphorina citri

The geographical representation of the sensitivity analysis addressed in the PMA regionalization methodology for ACP on the two overall risks computed (\(OR_{1,\text{ACP}}\) and \(OR_{2,\text{ACP}}\) for "El Camp de Túria" and "La Ribera Alta" counties are shown in Figures 5, 6.
and Figures 7, 8, respectively. Additionally, the internal evaluation measures used to assess the performance of the different values of \( \alpha \) parameter are shown in Table 2.

In “El Camp de Túria” county, the number of PMAs was estimated at 25 for both overall risks \((OR_{1,acp} \text{ and } OR_{2,acp})\), according to the restriction on the maximum PMA size (25 cells) imposed in the initial conditions of the algorithm. The value for the \( \alpha \) parameter that optimized the trade-off between the loss of risk homogeneity and the gain in spatial aggregation was estimated at \( \alpha = 0.2 \) and \( \alpha = 0.3 \) under \( OR_{1,acp} \) and \( OR_{2,acp} \). As observed in Figures 5A, C, E and 6A, C, E, higher \( \alpha \) values are related to a greater increase in the spatial aggregation of PMAs, which is interesting from a practical point of view of the coordinated management of ACP. Nevertheless, they are also the highest \( \alpha \) values that report the lowest intra-PMA homogeneity as well as the lowest inter-PMAs differences in terms of risk, as shown in Figures 5B, D, F and 6B, D, F. Box-and-whisker plots for each PMA belonging to the \( \alpha = 0.9 \) present wider distributions and overlap one another (Figures 5F, 6F). The internal validation measures that were estimated supported the graphic results (Table 2). Configurations associated with lower \( \alpha \) values show higher intra-PMAs compactness and PMA separation, and

<table>
<thead>
<tr>
<th>N cells (1 km(^2))</th>
<th>Commercial*</th>
<th>Conventional*</th>
<th>Organic*</th>
<th>Abandoned*</th>
<th>Residential*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castellón (ESS22 NUTS 3)</td>
<td>1,310 (19.57%)</td>
<td>354.10 (100%)</td>
<td>324.59 (91.67%)</td>
<td>8.62 (2.43%)</td>
<td>20.89 (5.90%)</td>
</tr>
<tr>
<td>Valencia (ESS23 NUTS 3)</td>
<td>3,550 (53.02%)</td>
<td>943.95 (100%)</td>
<td>862.94 (91.42%)</td>
<td>23.47 (2.49%)</td>
<td>57.55 (6.10%)</td>
</tr>
<tr>
<td>Alicante (ESS21 NUTS 3)</td>
<td>1,835 (27.41%)</td>
<td>337.48 (100%)</td>
<td>307.74 (91.19%)</td>
<td>8.62 (2.55%)</td>
<td>21.11 (6.26%)</td>
</tr>
<tr>
<td>Valencian A.C. (NUTS 2)</td>
<td>6,695 (100%)</td>
<td>1,635.53 (100%)</td>
<td>1,495.27 (91.42%)</td>
<td>40.71 (2.49%)</td>
<td>99.55 (6.09%)</td>
</tr>
</tbody>
</table>

* Area in km\(^2\)
therefore have lower values of the `max(complete.diameter)` and `min(complete.linkage.distance)`, respectively. However, the configurations \( \alpha = 0.2 \) (for \( OR_{1,\text{acp}} \)) and \( \alpha = 0.1 \) (for \( OR_{2,\text{acp}} \)) are the ones that maximize the Dunn index, with the best trade-off between compactness and separation. Regarding the loss of risk homogeneity, the \( \alpha = 0.9 \) configuration resulted in a loss of 25% and 24.1% for \( OR_{1,\text{acp}} \) and \( OR_{2,\text{acp}} \), respectively, in relation to the \( \alpha = 0 \) configuration (i.e., 100% risk homogeneity). Thus, they maintain a 75% and 75.9% of risk homogeneity in overall terms. The loss of spatial aggregation for the \( \alpha = 0.9 \) configuration is 0.2% and 0.6% for \( OR_{1,\text{acp}} \) and \( OR_{2,\text{acp}} \) in relation to the configuration of \( \alpha = 1 \) (i.e., 100% spatial aggregation).

In “La Ribera Alta” county, the number of PMAs was estimated at 31 for both overall risks (\( OR_{1,\text{acp}} \) and \( OR_{2,\text{acp}} \)), according to the restriction on the maximum PMA size (25 cells) imposed in the initial conditions of the algorithm. The value for the \( \alpha \) parameter that optimized the trade-off between the loss of risk homogeneity and the gain in spatial aggregation was estimated at \( \alpha = 0.4 \) for both risks. The influence of the \( \alpha \) parameter in the performance of the algorithm observed in “El Camp de Túria” was also observed in “La Ribera Alta”. As in “El Camp de Túria”, higher \( \alpha \) values increased the spatial aggregation of PMAs in “La Ribera Alta” (Figures 7A, C, E and 8A, C, E) but also reduced the intra/inter-PMA risk homogeneity (Figures 7B, D, F and 8B, D, F). Box-and-whisker plots for each PMA with \( \alpha = 0.1 \)
showed the narrowest distributions and lowest overlap (Figures 7B, 8B). The values of the internal validation measures were in line with the graphical results (Table 2). Higher intra-PMAs compactness and PMA separation values were obtained with $\alpha = 0.9$, although $\alpha = 0.4$ (for OR$_{1,acp}$) and $\alpha = 0.1$ (for OR$_{2,acp}$) resulted in the best trade-off between compactness and separation with the highest values of the Dunn index ($D = 0.432$ for OR$_{1,acp}$ and $D = 0.483$ for OR$_{2,acp}$). In relation to risk homogeneity, $\alpha = 0.9$ resulted in the greatest loss but without exceeding 11% (10.9% for OR$_{1,acp}$ and 11% for OR$_{2,acp}$), while resulting in a spatial aggregation of 99.8 (100-0.2)% for both overall risks.

Further description of the individual and overall distribution of the commercial citrus area (km$^2$) and number of cells in PMAs in “El Camp de Turia” and “La Ribera Alta” for OR$_{1,acp}$ and OR$_{2,acp}$ are described in Supplementary Figures SB1, SB2, respectively.

### 3.3.2 Trioza erytreae

Figures 9, 10 and Figures 11, 12 show a graphical overview of the results obtained for AICP on the two overall risks computed (OR$_{1,afcp}$ and OR$_{2,afcp}$) for “El Camp de Turia” and “La Ribera Alta” counties. The internal evaluation measures used to assess the performance of the different choices of the $\alpha$ parameter are displayed in Table 3.

In “El Camp de Turia” county the number of PMAs was estimated at 25 for both overall risks (OR$_{1,afcp}$ and OR$_{2,afcp}$), according to the
Higher intra-PMAs compactness and PMA separation values were resulted in higher intra-PMAs compactness and PMA separation values. Figures 9F, 10F. Box-and-whisker plots for each PMA with wider distributions and a higher overlap (Figures 9F, 10F). Lower values increased spatial aggregation but reduced the intra/inter-PMAs compactness and PMA separation in terms of risk homogeneity. Box-and-whisker plots for each PMA with $\alpha = 0.1$ showed the narrowest distributions with the lowest overlap (Figures 11B, 12B). Higher intra-PMAs compactness and PMA separation values were obtained with $\alpha = 0.9$, although $\alpha = 0.1$ showed the best trade-off between compactness and separation with the highest value of the Dunn index ($D = 0.617$ for $OR_{1,afcp}$ and $D = 0.504$ for $OR_{2,afcp}$) (Table 3). In relation to risk homogeneity, $\alpha = 0.9$ resulted in the highest reduction, 11.7% for $OR_{1,afcp}$ and 11.1% for $OR_{2,afcp}$ but resulted in a spatial aggregation of 99.8 (100-0.2)% for both risks.

Further description of the individual and overall distribution of the commercial citrus area (km$^2$) and number of cells in PMAs in “El Camp de Túria” and “La Ribera Alta” for $OR_{1,afcp}$ and $OR_{2,afcp}$ are described in Supplementary Figures SB3 and SB4, respectively.

4 Discussion

This study presents a risk-based approach to defining management areas for HLB vectors. The rationale of this approach is based on the regionalization of citrus-growing areas into Pest Management Areas (PMAs) that share a similar overall risk to the introduction and spread of insect vectors but considering...
spatial constraints to make it feasible from the viewpoint of implementation. The size and location of PMAs were estimated using the Ward-like hierarchical algorithm developed by Chavent et al. (2018). This spatial aggregation algorithm provides relevant elements for the grouping of areas under similar risk and allows the incorporation of geographical information to address the clustering process, which in the context of vector-borne plant diseases is a fundamental point to support the operational implementation (Bassanezi et al., 2013; Yuan et al., 2021). The Valencian Autonomous Community (ES52 NUTS 2), Spain, was selected as the study area because it is the largest citrus-growing region in Europe and is representative of others in the Mediterranean Basin. Specifically, PMAs were designed and optimized for “La Ribera Alta” and “El Camp de Túria” counties.

To date, several proposals for area-wide management of HLB have been developed, framed in different epidemiological settings (Bové, 2006; Manjunath et al., 2010; Kumagai et al., 2013; Flores-Sánchez et al., 2017; Badaracco et al., 2021; Garcia-Figuera et al., 2021a). In general, the rationale behind these approaches considers regional risk factors, estimated based on epidemiological variables such as cultivar susceptibility, host density, and vector presence and abundance (Gottwald et al.,

---

**FIGURE 5**

Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for OR_{acp} (A, C, E) in “El Camp de Túria” county with α of 0.1 (A, B), 0.3 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of OR_{acp} in “El Camp de Túria” county; n quantifies the size of each Pest Management area in terms of the number of 1 km² cells; central line represents the median value; box limits represent the first (Q₁) and third (Q₃) quantiles; upper whisker represents min(max(x), Q₃ +1.5*IQR); lower whisker represents max(min(x), Q₁ −1.5*IQR); IQR = Q₃ − Q₁; outliers are represented by dots.
Some of them also incorporate other factors related to human factors, such as the availability of infrastructure, human and economic resources in the citrus-producing regions, as in Mexico (Flores-Sánchez et al., 2017), or the social willingness to participate, as in California (McRoberts et al., 2019; García-Figuera et al., 2022). Nevertheless, although previous research has shown improved HLB control with the area-wide strategy (Bassanezi et al., 2013; García-Figuera et al., 2021a), the standard settings for determining the size and location of area-wide management have not been established (Flores-Sánchez et al., 2017).

Also, different recommendations have been suggested concerning the area-wide size to manage HLB. Using empirical data, Bové (2012) proposed an area size of 500 ha in Brazil, whereas sizes in California range from 4,000 to 20,000 ha (Flores-Sánchez et al., 2017). Nonetheless, the farming systems in Brazil and California are usually composed of large orchards. This contrasts with the citrus-growing areas in the Mediterranean Basin like Spain, where farms are smaller and divided up into orchards of <1 ha often managed independently by individual growers (Fernández-Zamudio et al., 2006). Thus, PMAs aggregation will probably be more challenging from an extension point of view.

**FIGURE 6**
Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for OR$_{acp}$ (A, C, E) in “El Camp de Túria” county with $\alpha$ of 0.1 (A, B), 0.3 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of OR$_{acp}$ in “El Camp de Túria” county; n quantifies the size of each Pest Management Area in terms of the number of 1 km$^2$ cells; central line represents the median value; box limits represent the first ($Q_1$) and third ($Q_3$) quantiles; upper whisker represents $\text{min}(\text{max}(x), Q_3 + 1.5 \times IQR)$; lower whisker represents $\text{max}(\text{min}(x), Q_1 - 1.5 \times IQR)$; IQR = $Q_3 - Q_1$; outliers are represented by dots.
To implement our proposal, the study area was characterized by several risk factors at a grid resolution of 1 km² using data on climatic suitability, conventional, organic, abandoned and residential citrus areas, and transportation corridors. These risk factors have been associated with the introduction and spread of ACP, and in consequence with HLB epidemiology (Halbert et al., 2010; Richards et al., 2013; Gottwald et al., 2014a; Martini et al., 2015; Bayles et al., 2017; Flores-Sánchez et al., 2017). Our proposal considers the same approach for AfCP. While there have been extensive studies of ACP and AfCP in other regions, there is limited knowledge about both vectors’ biology, spread capacity and host range in the Mediterranean Basin. Likewise, the Valencian Autonomous Community remains free from HLB and from both ACP and AfCP and so it is uncertain how the vectors and the disease will behave.

The risk factors included in our study do not consider the wind effects (i.e., direction and speed) although previous studies on ACP spread dynamics and area-wide management often consider this factor to predict vector distribution in an invading scenario from an entry point. Antolínez et al. (2022) indicated that ACP adults are...
likely to spread across the prevailing wind direction and that wind speeds higher than 48 km/h are able to dislodge adults from citrus leaves and trigger passive long-distance spread events. Moreover, some studies have hypothesized that severe meteorological wind-related episodes, such as hurricanes in Florida, greatly influenced the major long-distance spread of ACP in Florida (Johnston et al., 2019; Monzo and Vanaclocha, 2023). Nevertheless, the availability of accurate wind data at large scales on a regional resolution is scarce given that wind observations vary on small space and time scales due to the fact that wind is affected by the local terrain, vegetation and buildings. Furthermore, recent studies have shown that other factors, such as orchard layout, have a stronger influence on HLB incidence and spread than wind (Benhadi-Marín et al., 2021; Primiano et al., 2023). Other risk factors integrated in previous studies, such as packing houses and nurseries (Gottwald et al., 2014a; Flores-Sánchez et al., 2017), were not included either. The relevance of these factors in the vector distribution also remains unclear. Similarly, host data (variable and cultivar) was not included because of the lack of available data on cultivar preference and susceptibility.

FIGURE 8 Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for ORacp in “La Ribera Alta” county with α of 0.1 (A, B), 0.4 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of ORacp in “El Camp de Túria” county; n quantifies the size of each Pest Management Area in terms of the number of 1 km² cells; central line represents the median value; box limits represent the first (Q₁) and third (Q₃) quantiles; upper whisker represents min(max(x), Q₃ + 1.5*IQR); lower whisker represents max(min(x), Q₁ − 1.5*IQR); IQR = Q₃ − Q₁; outliers are represented by dots.
Nevertheless, the proposed regionalization methodology for estimating the size and location of PMAs ensures sufficient flexibility to update the default risk factors considered or to add new ones as more information about the insect vectors/disease become available or as the epidemiological setting evolves. Similarly, although in our study two options have been proposed for estimating the overall risk, i) averaging its influence, and ii) giving greater importance to the risk posed by the presence of abandoned and organic orchards than to the rest of the risks, other criteria can be accommodated. In the same way, further refinements may be implemented for the risk factor normalization in the presence of extreme values.

The approach proposed in our study for defining the size and location of PMAs makes use of a hierarchical clustering algorithm that allows the introduction of spatial constraints (Chavent et al., 2018). This kind of clustering methodology is usually called regionalization. The term regionalization was defined by Guo (2008) as the process of aggregating a set of spatial entities into a reduced number of regions in a way that a predefined objective function is optimized. Spatial contiguity and homogeneity are the basic criteria underpinning regionalization. Spatial contiguity requires that spatial connectivity or spatial tightness should be met. Beyond the ClustGeo (Chavent et al., 2018), there are several important regionalization algorithms such as SKATER (Assunção et al., 2006) or REDCAP (Guo, 2008). However, ClustGeo was chosen because it enables the definition of the spatial relationship between the units intended for grouping through the assessment of similarity between geographical coordinates, without the necessity for these units to satisfy explicit spatial connectivity requirements, as, for instance, demanded by SKATER (Assunção et al., 2006). This functionality offered by ClustGeo was the decisive factor that led us to opt for this algorithm, given the specific spatial arrangement characteristics of our study area concerning the units to be grouped.

These regionalization techniques are used in several disciplines, including sociology, economics, urban planning, politics and health to identify areas with similar characteristics and to get useful information to support policy-making (Camelo Aguiar et al., 2020; Guo et al., 2022). However, to our knowledge, this is the first time that this methodology has been used in plant health to group areas with a similar level of risk of vector introduction and spread in order to address a coordinated management and, additionally, on an open-source implementation. The greatest potential of this algorithm in the frame of area-wide management is that it makes it possible to consider the spatial attributes of the units to be clustered and to modulate the importance of spatial aggregation versus the “risk” in the design of the solutions. This importance of spatial aggregation is defined in the implementation offered (i.e., the R package ClustGeo) through the setting of the $\alpha$ parameter. Specifically, when $\alpha = 0$ the spatial dissimilarities are not taken into account and when $\alpha = 1$ the “non-spatial attributes” distances are not considered and only the spatial

<table>
<thead>
<tr>
<th>County</th>
<th>OR</th>
<th>$\alpha$</th>
<th>Compactness $(A)^{a}$</th>
<th>Separation $(B)^{b}$</th>
<th>Dunn index $(D)/(B/A)^{c}$</th>
<th>$1-Q_0\text{norm} (d)^{d}$</th>
<th>$1-Q_1\text{norm} (e)^{e}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Camp de Túria</td>
<td>OR$_{ag}$</td>
<td>0.1</td>
<td>0.157</td>
<td>0.064</td>
<td>0.408</td>
<td>3.6</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
<td>0.153</td>
<td>0.085</td>
<td>0.556</td>
<td>7.3</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>0.351</td>
<td>0.126</td>
<td>0.359</td>
<td>25.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>OR$_{ag}$</td>
<td>0.1</td>
<td>0.127</td>
<td>0.053</td>
<td>0.417</td>
<td>2.4</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>0.252</td>
<td>0.092</td>
<td>0.365</td>
<td>5.8</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>0.458</td>
<td>0.150</td>
<td>0.328</td>
<td>24.1</td>
<td>0.4</td>
</tr>
<tr>
<td>La Ribera Alta</td>
<td>OR$_{ag}$</td>
<td>0.1</td>
<td>0.118</td>
<td>0.046</td>
<td>0.390</td>
<td>1.2</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.148</td>
<td>0.064</td>
<td>0.432</td>
<td>2.6</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>0.308</td>
<td>0.097</td>
<td>0.315</td>
<td>10.9</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>OR$_{ag}$</td>
<td>0.1</td>
<td>0.120</td>
<td>0.058</td>
<td>0.483</td>
<td>1.1</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.196</td>
<td>0.091</td>
<td>0.464</td>
<td>2.4</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>0.464</td>
<td>0.148</td>
<td>0.319</td>
<td>11.0</td>
<td>0.2</td>
</tr>
</tbody>
</table>

$^a$Estimated by the max(complete diameter).
$^b$Estimated by the min(complete linkage distance).
$^c$Estimated as the $D = \frac{\text{aggregated linkage distance}}{\text{complete linkage diameter}}$.
$^dQ_0\text{norm}(d)$ denotes the normalized proportion of pseudo-inertias explained by the overall risk.
$^eQ_1\text{norm}(e)$ denotes the normalized proportion of pseudo-inertias explained by the spatial aggregation.
distances are taken into account (Chavent et al., 2018; Camelo Aguiar et al., 2020).

The Clustgeo algorithm presents a marked variability in the design of the results depending on the setting of parameter $k$ (i.e., maximum number of clusters or PMAs allowed) and depending on the value of parameter $\alpha$. Specifically, the choice of the maximum number of PMAs (i.e., clusters) was defined assuming a maximum PMA size of 25 cells based on expert criteria, but this parameter can be adjusted by the end user or optimized based on different criteria. Furthermore, a sensitivity analysis was performed to evaluate the effect of three different values of $\alpha$. For this purpose, in addition to the graphical results, several internal validation measures were computed to quantify compactness and/or separation as well as the trade-off between the loss of risk homogeneity and the loss of spatial aggregation (Leskovec et al., 2014; Kassambara, 2017). For $\alpha = 0.9$ under the four configurations evaluated ($OR_{1,\alpha<\alpha}$, $OR_{2,\alpha<\alpha}$, $OR_{1,\alpha<\alpha}$, $OR_{2,\alpha<\alpha}$), the results obtained in both counties, “La Ribera Alta” and “El Camp de Túria”, showed the greatest spatial aggregation in the design of the location of PMAs, which is fundamental for their implementation from a practical point of

![FIGURE 9](https://via.placeholder.com/150)

**FIGURE 9**
Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for $OR_{1,\alpha<\alpha}$ (A, C, E) in “El Camp de Túria” county with $\alpha$ of 0.1 (A, B), 0.3 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of $OR_{1,\alpha<\alpha}$ in “El Camp de Túria” county; n quantifies the size of each Pest Management Area in terms of the number of 1 km² cells; central line represents the median value; box limits represent the first ($Q_1$) and third ($Q_3$) quantiles; upper whisker represents $\min(\max(x), Q_3 + 1.5\times IQR)$; lower whisker represents $\max(\min(x), Q_1 - 1.5\times IQR)$; $IQR = Q_3 - Q_1$; outliers are represented by dots.
view. Furthermore, with $\alpha = 0.9$ the solutions also supported at least 72.4% homogeneity in terms of overall risk.

In sum, this PMAs design proposal, addressing two overall risk estimates for both ACP and ACP, offers a comprehensible, accessible and adaptable approach both to estimating the risk factors associated with the introduction and spread of the two HLB vectors and to establishing their corresponding management areas (PMAs). Although it was implemented at NUTS2 level due to administrative reasons, this framework can be upscaled at country and even at Mediterranean Basin levels to implement a standard management and minimize the possible migration of vectors among areas. Thus, the development of this PMAs design proposal means having a preventive tool at the service of the citrus industry. Its use can be extended beyond the coordination of vector/disease management, for instance, to the field of epidemiological surveillance as well as to the context of education and awareness-raising.

Nevertheless, future work would be needed to verify the citrus data at orchard level and its classification as

---

**FIGURE 10**
Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for $OR_{\alpha, a f c p}$ (A, C, E) in "El Camp de Túria" county with $\alpha$ of 0.1 (A, B), 0.3 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of $OR_{\alpha, a f c p}$ in "El Camp de Túria" county; $n$ quantifies the size of each Pest Management Area in terms of the number of 1 km² cells; central line represents the median value; box limits represent the first ($Q_1$) and third ($Q_3$) quantiles; upper whisker represents $\min(\max(x), Q_3 + 1.5\times IQR)$; lower whisker represents $\max(\min(x), Q_1 - 1.5\times IQR)$; $IQR = Q_3 - Q_1$; outliers are represented by dots.
conventional, organic, abandoned, residential, etc. Indeed, the locations of organic and abandoned citrus orchards were randomly allocated to reach the total area reported by Ministerio de Agricultura, Pesca y Alimentación (MAPA) (2019; 2020). The isotherm for citrus growth and population census data were used to estimate the residential citrus areas. Thus, it would also be necessary to update the default risk factors considered, and to add new ones when more information about the insect vectors/disease become available after an eventual introduction in the study area. Additionally, social willingness toward the adoption of PMAs should be properly addressed. Previous experiences from regions with the presence of disease and vector have demonstrated that the only successful action to control the spread of HLB is to control the vector on a large spatial scale (Singerman et al., 2017; Milne et al., 2018; Alquézar et al., 2022). Area-wide programs in countries with the presence of HLB typically rely on voluntary adoption (García-Figuera et al., 2021a; García-Figuera et al., 2022). To achieve voluntary adoption, in our case, the advantages of the coordinated treatments should be effectively communicated and make
Spatial (A, C, E) and overall distribution (B, D, F) of the delimited PMAs for OR$_{2,afcp}$ (A, C, E) in “La Ribera Alta” county with $\alpha$ of 0.1 (A, B), 0.4 (C, D) and 0.9 (E, F). In box-and-whisker plots red line represents the mean value of OR$_{2,afcp}$ in “El Camp de Túria” county; $n$ quantifies the size of each Pest Management Area in terms of the number of 1 km$^2$ cells; central line represents the median value; box limits represent the first ($Q_1$) and third ($Q_3$) quantiles; upper whisker represents $\text{min}(\text{max}(x), Q_3 + 1.5\text{IQR})$; lower whisker represents $\text{max}(\text{min}(x), Q_1 - 1.5\text{IQR})$; IQR = $Q_3 - Q_1$; outliers are represented by dots.

TABLE 3 Performance of the regionalization algorithm for the definition of the Pest Management Areas (PMAs) for Trioza erytreae (AfCP) under the three different specifications of the mixing parameter $\alpha$ for OR$_{1,afcp}$ and OR$_{2,afcp}$ in terms of compactness, separation, the Dunn index and the loss of risk homogeneity (1-$Q_0$ norm($\alpha$)) and increase in spatial aggregation (1-$Q_1$ norm($\alpha$)) in relation to the reference configurations $\alpha = 0$ and $\alpha = 1$, respectively.

<table>
<thead>
<tr>
<th>County</th>
<th>OR</th>
<th>$\alpha$</th>
<th>Compactness (A)$^a$</th>
<th>Separation (B)$^b$</th>
<th>Dunn index (D) (B/ A)$^c$</th>
<th>1-$Q_0$ norm ($\alpha$)$^d$</th>
<th>1-$Q_1$ norm ($\alpha$)$^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Camp de Túria</td>
<td>OR$_{1,afcp}$</td>
<td>0.1</td>
<td>0.128</td>
<td>0.060</td>
<td>0.469</td>
<td>3.1</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
<td>0.151</td>
<td>0.087</td>
<td>0.576</td>
<td>7.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>0.351</td>
<td>0.140</td>
<td>0.399</td>
<td>27.6</td>
<td>0.1</td>
</tr>
</tbody>
</table>

(Continued)
stakeholders trust the risk manager to keep them motivated and involved in the program (Klassen, 2008; Singerman et al., 2017; McRoberts et al., 2019; Monzo and Stansly, 2020; García-Figuera et al., 2021b; Pavone, 2022).

### Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author. Data curation and formal analysis have been implemented in the R statistical programming. All the generic code used in the research is open-source and available at https://zenodo.org/records/10079629.

### Author contributions

AG: Data curation, Writing – original draft, Writing – review & editing. RB: Conceptualization, Writing – review & editing. WL: Conceptualization, Writing – review & editing. PV: Methodology, Writing – original draft, Writing – review & editing. AV: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. EL: Data curation, Writing – original draft, Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing – review & editing.

### Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by PreHLB N°817526 grant of the European Union Horizon 2020 program. AG held an IVIA grant partially funded by the European Social Fund.

### Acknowledgments

We also thank César Monzó (IVIA), Aranzazu Moreno and Alberto Fereres (IAS-CSIC), Amilcar Duarte, Tomás Magalhães and Rita Poeria (Univ. Algarve), Jacinto Benhadi (Inst. P. Bragança), Silvio Lopes (Fundecitrus) and Alessandra Gentile (Univ. Catania) for their contribution in providing substantial insights and documented information concerning *D. citri* and *T. erytreae*.

### Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

### Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1256935/full#supplementary-material

#### SUPPLEMENTARY MATERIAL A

Climatic conditions, landscape architecture and citrus industry practices affecting introduction and spread of citrus psyllids.

#### SUPPLEMENTARY MATERIAL B

Pest Management Areas (PMAs) descriptive.


2.2 (Diptera: Tephritidae): edge effect in area-wide treatments. Remote Sens. 138, 403
Applicata. 18, 96
Available at: https://www.boe.es/eli/es/rd/2023/02/21/115.