Modelling the vulnerability of urban settings to WUI fires in Chile.

Paula Aguirre[1], Jorge León[2], Constanza González-Mathiesen[3], Randy Román[4], Manuela Penas[5], Alonso Ogueda[6]

[1] Instituto de Ingeniería Matemática y Computacional, Escuela de Ingeniería, Pontificia Universidad Católica de Chile. Santiago, Chile.
[2] Departamento de Arquitectura, Universidad Técnica Federico Santa María, Valparaíso, Chile.
[3] Facultad de Arquitectura, Universidad del Desarrollo, Concepción, Chile.
[4] Instituto de Estudios Territoriales y Urbanos, Pontificia Universidad Católica de Chile, Santiago, Chile.
[5] Independent researcher, Valparaíso, Chile.
[6] George Mason University, College of Science, Virginia, USA.

Correspondence to: Paula Aguirre (paaguirr@uc.cl)

Abstract.

Wildland-urban interface (WUI) regions are particularly vulnerable to wildfires due to their proximity to both nature and urban developments, posing significant risks to lives and property. To enhance our understanding of the risk profiles in WUI areas, we analysed seven fire case studies in central Chile. We developed a mixed-methods approach for conducting local-scale analyses which involved field surveys, remote-sensing through satellite and drone imagery, and GIS-based analysis of the collected data. The methodology led to the generation of a georeferenced dataset of damaged and undamaged dwellings, including 16 variables representing their physical characteristics, spatial arrangement, and the availability of fire suppression resources. A binary classification model was then used to assess the relative importance of these attributes as indicators of vulnerability. The analysis revealed that spatial arrangement factors have a greater impact on damage prediction than the structural conditions and fire preparedness of individual units. Specifically, factors such as dwelling proximity to neighbours, distance to vegetation, proximity to the border of dwelling groups, and distance from the origin of the fire substantially contribute to the prediction of fire damage. Other structural attributes associated with less affluent homes may also increase the likelihood of damage, although further data is required for confirmation. This study provides insights for the design, planning, and governance of WUI areas in Chile, aiding the development of risk mitigation strategies for both built structures and the broader territorial area.

1 Introduction

1.1 General Problem: fires are an increasing issue in the wild-urban interface

Wildfires are a natural hazard that refers to an out-of-control wildland fire burning over a large area (Coppola 2015). Every year, more than two million small wildfire events are registered worldwide, with most of them having no significant impacts

and not evolving into disasters (Bowman et al. 2017a). However, a small proportion of these events escalate into very large incidents that have significant ecological and socioeconomic impacts (Bowman et al. 2017a; Tedim et al. 2018), including fatalities, physical and psychological injuries, property loss, and diverse environmental impacts (March et al. 2020a).

35 Furthermore, 96% of extreme wildfire disasters are associated with anomalous meteorological or climatic conditions (Bowman et al. 2017b). Climate change has already led to a worldwide increase in the severity and frequency of weather conditions conducive to fires, a trend that is expected to persist (Jones 2020). For instance, Ellis et al. (2022) pointed out that forests in boreal and mid-latitude regions have experienced significant increases in fire activity, driven by more frequent occurrences of extreme fire-prone weather and longer fire seasons ~~as a result~~because of anthropogenic climate change.

40

Wildland Urban Interface (WUI) areas imply a higher wildfire risk because human lives and material goods are more exposed to fire (Gill & Stephens, 2009). Wildfires in WUI areas constitute an "unresolved and growing problem given population dynamics, climate change and, in some cases, increasing forest areas" (Sarricolea et al. 2020, p.2). In these areas, the likelihood of housing loss due to wildfires is significantly influenced by housing arrangement and location (Syphard et al. ~~2012)~~2012a).

45 Moreover, settlement patterns, such as the expansion of low-density urban sprawl and rural-residential developments that encroach into fire-prone areas, can also affect the frequency and severity of catastrophic wildfires (Butt et al. 2009). Therefore, it is important that the current methods of WUI development be revised (Bowman et al. 2017b; González-Mathiesen and March 2018a; Moritz et al. 2014).

1.2 Wildfire risk assessment

50 To revise the current methods of WUI development and reduce or mitigate risk, it is of foremost importance that a wildfire threat assessment precede development. The need to integrate wildfire considerations into urban planning systems is well-established in the literature (e.g., González-Mathiesen and March 2018; March et al. 2020b). The first step toward integrating wildfire considerations into urban planning systems is to ~~analyze the~~carry out wildfire risk analyses (AIDR 2020). ~~This analysis,~~ which should provide valuable insights for settlement expansion decisions, ensuring that new development does not

55 perpetuate or increase current risks. Furthermore, conducting assessments of already developed areas is also the first step for the targeted implementation of mitigation strategies.

Wildfire assessments can emphasize different perspectives, including risk, hazard, and vulnerability (Galiana-Martin 2017). In its general definition, the term *hazard* refers to the process or phenomenon that may cause loss of life or injuries, property damage, social and economic disruption, or environmental degradation (Goldammer et al. 2017). The wildfire hazard is a

60 combination of likelihood, defined as the annual probability of ignition in a specific location, and intensity, which refers to the expansion or energy expected from a wildfire (USDA 2024). Usually, wildfire hazard is associated with vegetation or available fuel, topography, weather/climate conditions, ignition likelihood, and suppression capabilities (eg., March et al. 2020b). The assessment of *risk* aims to quantify the potential losses caused by fire and its spread (eg. Jappiot et al. 2009) over a given period and spatial region, and is therefore a product of three components: (1) the hazard, (2) exposure, or inventory of

2

65 population, built assets, ecosystem services, economic activities or other valuable elements located in hazard-prone areas (UNISDR 2009), and (3) the vulnerability, which represents the propensity of the exposed elements to be damaged if a wildfire occurs. The vulnerability is determined for example by the socioeconomic factors, building conditions, and population demographics of exposed areas (Oom et al. 2022; USDA 2024)

Hazard assessments ~~are~~have been often preferred over risk assessments in urban planning as they tend to be simpler and faster
70 to conduct and can be easier to implement. For example, land use planning in Victoria, Australia, is determined by hazard-based zoning that identifies the wildfire hazards rather than the risk (DTPLI 2013). However, this approach assumes that the fire hazard in urban areas is low, which could imply an underestimation of the risk in WUI areas. This suggests that risk assessment techniques that are simple to conduct and implement would promote the use of risk analysis over hazards analysis in urban planning contexts.

75 For the purposes of risk reduction and mitigation, wildfire risk assessments should be approached with a medium-term risk analysis perspective. ~~This corresponds to the likely ignition and behavior of fire, which are influenced by structural factors that evolve slowly over an extended period (Sivrikaya et al. 2014). Therefore, assessments should be developed for longer time frames, such as~~, considering for example time frames from two to ten years (Jappiot et al. 2009). Several countries have developed ~~. Medium-term risk is usually assessed considering two components: (1) likelihood, which refers to both the~~
80 ~~likelihood of ignition and of fire expansion and; (2) vulnerability or consequence, which refers to the impact of a fire and its spread (Jappiot et al. 2009; Sivrikaya et al. 2014). Furthermore, the~~ WUI risk assessment methodologies at varying spatial scales (i.e. local, landscape, regional, national), and with different data inputs and information purposes, but overall the process usually follows a procedural approach similar to the one established by ISO 31000 (2018~~).~~, which integrates hazard, exposure, and vulnerability components. ~~According to March~~ (see, e.g. Oom et al. 2022; San-Miguel-Ayanz et al. 2017; Syphard et al.
85 2012b; Mitsopoulos, Mallinis, and Arianoutsou 2015; Calkin, Owen Price, and Salis 2019; Caggiano et al. ~~(2020b), wildfire hazard is associated with vegetation or available fuel, topography, weather/climate conditions, ignition likelihood~~2020; Zong, Tian, and Fang 2022; Sakellariou et al. 2022). Other novel approaches aim to extend probabilistic assessments to include WUI resilience; for example, Tampekis et al. (2023) propose a holistic theoretical framework based on performance-based engineering that expands the components of risk to include potential social-ecological impacts~~suppression capabilities.~~
90 ~~Exposure relates to the population, built assets, ecosystem services~~ and ~~economic activities exposed to the hazard. Vulnerability refers to socioeconomic factors, building conditions, population demographics~~interactions, and ~~assets exposed to the hazard.~~to evaluate social-ecological resilience.

Spatial mapping techniques are an important part of wildfire risk analysis and management, evolving alongside technology,
95 especially in terms of Geographic Information Systems (GIS). These techniques aim to quantify the risk as well as its spatial magnitude (Atkinson et al. 2010). During the 1990s, these analysis methods were generally qualitative and focused on components such as hazard (involving estimates of fuel load based on vegetation) and risk (associated with ignition risk to valuable natural and man-made assets). These components were put into a matrix and assigned low, medium, or high-risk

values, which were then overlaid with maps to determine high-risk areas (Atkinson et al. 2010). As technology and data quality have improved, wildfire risk analysis processes have evolved and become more complex and multi-faceted. Spatial mapping techniques, such as GIS, are becoming important tools that can greatly contribute to reducing the hazards' impact on society (Chen, Blong, and Jacobson 2003). For instance, Romero-Calcerrada et al. (2008) used GIS-based spatial models to assess the main causes of wildfire ignition (and to predict ignition risk) in the Southwest of Madrid, Spain. Their results indicated that proximity to urban areas and roads are the most significant causal factors. Along these lines, studies that explore ways to assess wildfire risk in WUI areas usually distinguish between land cover types. These studies record vegetated land cover classes, in addition to one or a few categories of urban land covers (such as high-density residential, low-density residential, and commercial/industrial) that cover large areas (e.g., Bar Massada et al. 2009; Zhu et al. 2012). Based on this information, researchers then construct a burn probability map to identify the probability of structures being affected by a wildfire, highlighting those at higher risk. Additionally, some studies also include housing density as one of the assessment criteria (e.g., Lu et al. 2009; Sirca et al. 2017).

The use of contemporary modelling techniques that map the fire risk for structures in WUI areas could play a significant role in improving risk reduction and mitigation through urban planning and governance. However, research that explores mapping wildfire risk in WUI areas is still in its seminal stages and is typically conducted at a territorial scale. Remarkably, the consideration of detailed-scale physical characteristics of settlements and structures, along with their nuances, remains an emerging field of research ~~(e.g., Papathoma Köhle et al. 2022). This area of study is still very limited, especially when compared to the sophisticated models currently available for assessing wildfire risk in rural contexts.~~, which is commonly addressed through the examination of post-disaster scenarios with data collected through field surveys. Examples of this type of approach include, for instance, Papathoma-Köhle et al. (2022), who developed a physical vulnerability index for buildings subject to wildfire. To do this, they used data (13 building structural and surrounding features) from 423 built units affected by the Mati fire occurred in Greece in 2018 and applied the random-forest-based Boruta algorithm to estimate the feature importance of each of those elements on the degree of damage. In turn, (Dossi et al. 2022) used a similar approach (but only considering built features and not characteristics from the immediate surroundings) a large sample of 17,500 buildings exposed to wildfires in California, USA (between 2013 and 2017) and of 1,190 buildings exposed to the 2017 Pedrógão Grande Fire Complex in Portugal. They applied statistical dependence tests to this database, to examine the possible correlations between the dependent variable (i.e., the surveyed damage level) and the independent ones (i.e., each building feature).

In comparison to the sophisticated models currently available for assessing wildfire risk in rural contexts, the understanding of the feature importance of a built unit's physical characteristics as predictors of wildfire damage is still an emerging field of research. This is evidence of a fundamental shortcoming of risk assessment and mapping techniques that requires attention, particularly for the mitigation and prevention of wildfire risk in WUI areas.

4

1.3 The case of Chile

Chile's fire regime is dominated by frequent, low-intensity fires and some rare, but exceptionally intense events, all concentrated in the South-Central territory. Fire occurrence is highly seasonal and closely associated with precipitation and temperature variations (González et al. 2011). However, the occurrence of fire weather is increasing due to climate change (Urrutia-Jalabert et al. 2018). In this respect, Sarricolea et al. (2020) underline that climate change (with increasing temperatures and decreasing precipitation) has increased the probability, intensity, and speed of wildfire propagation, especially in subtropical and Mediterranean regions like the south-central area of Chile. As they point out, this trend was demonstrated by the tragic summer wildfires of 2016-2017, where the total burned area was more than 500,000 hectares (largely exceeding the records of the previous two decades, with roughly 50,000 to 100,000 burned hectares per season). Moreover, the 2022-2023 summer season proved to be equally catastrophic, with roughly 440,000 damaged hectares and approximately 5,700 recorded fires (CONAF, 2023). Historical statistics show that only a small number of fires, representing approximately 0.6% to 0.9% of all fires, are responsible for 60% of the burned areas in the country (CONAF 2018). Furthermore, the central and southern areas are the most impacted by wildfires in terms of the number of events and affected surface area (Castillo, Julio-Alvear, and Garfias 2014). These regions also correspond to the most populated areas of Chile, with Sarricolea et al. (2020) reporting that almost 3 million people live in WUI areas. Furthermore, fire ignition is mostly associated with human activities (CONAF 2018).

WUI areas impacted by wildfires are a relatively new – and increasingly common – challenge for Chile. The first recorded wildfire that impacted the WUI can be traced to an event that took place in Viña del Mar on January 22, 1968. More recently, in 2012, the Ñuble fire affected several small towns and resulted in the destruction of 162 dwellings and 24,000 burned hectares (Muñoz et al. 2012). In 2014, the 'Great Valparaiso Fire' impacted 11 of the 42 hills in Valparaíso, destroyed more than 2,900 houses, burned 926 hectares, caused 15 fatalities, injured 500 people and displaced over 12,500 (Reszka and Fuentes 2015). Furthermore, in 2017, the fire event called *Tormenta de Fuego*, the most devastating wildfire event recorded in Chile, burned 546,677 hectares, destroyed 2,831 buildings, displaced over 8,129 people, and caused 11 fatalities. These recent events have increased local awareness of the fact that most populated areas in the south-central territory might be at risk for wildfires (e.g., Castillo 2013; Castillo et al. 2014; Úbeda and Sarricolea 2016) and that this needs to be addressed through risk-based modelling approaches in the planning, governance, and development of WUI areas (e.g. Arana et al. 2018; Gómez-González et al. 2018; Reszka & Fuentes 2015).

In Chile, several wildfire risk analyses have been conducted in the WUI. For example, Garfias et al. (2012) used a socioeconomic approach to analyse wildfire-prone areas within the Valparaíso-Viña del Mar WUI through a semi-structured questionnaire survey that was answered by 403 households in an area of approximately 14,000 hectares. The survey was focused on assessing each dwelling's basic attributes (e.g., total built area), but also on examining its occupants' perception of the surrounding natural environment, including the fire hazard it poses to them. Castillo et al. (2014) examined (using high-resolution satellite images) the fire vulnerability in the Valparaíso WUI area using a socio-spatial approach that considered 15

variables that comprised the physical characteristics of buildings and settlements (e.g., number of houses per hectare, average dwelling size, construction material), and also socioeconomic variables (e.g., risk ratio of individuals in each household, average number of adults and children per household, economic loss factors) of the dwellings located there. When combined with environmental characteristics such as land slope and vegetation, this analysis allowed them to model the vulnerability of the territory in relation to forest fires. Sarricolea et al. (2020) analysed the wildfire occurrences using satellite images to identify WUI areas and determine the population residing in areas that may be affected by wildfires. In their study, they considered data regarding the monthly burned areas, land use/land cover (LULC), and census block and population data. The LULC data distinguished land covers such as grasslands, croplands, wetlands, and diverse types of forests (at 500 m spatial resolution for the period between 2000 and 2017, and 30 m spatial resolution for the year 2014). It distinguished urban and built-up land as one type of LULC.

1.4 Research gap and questions

The aforementioned studies contribute new and valuable insights into the wildfire threat in WUI areas in Chile and around the world. However, it should be noted that the analyses were undertaken mostly at a territorial scale that does not commonly allow the detailed physical characteristics and distribution of buildings and settlements or their impact on wildfire risk to be distinguished.; this more spatially focused approach is still an emerging field of research globally, not yet carried out in Chile. This highlights the necessity of smaller small-scale analyses that can provide nuanced information about how the different physical characteristics of structures and settlement patterns affect the risk profiles of WUI areas vulnerable to wildfires in the Chilean context. These nuanced understandings could contribute valuable knowledge to guide the design, planning, and governance of settlements in WUI areas in order to reduce and mitigate wildfire risk. Moreover, as several WUI areas in Chile include informally developed zones, these findings could also help to inform retrofitting policies with this objective. Also, we enhance the accuracy and speed of the collection of post-fire field data by using remote sensing through satellite and drone imagery. Along these lines, this study aims to answer the following research questions: What are the main vulnerability parameters that contribute to wildfire risk in WUI areas in Chile? How can these parameters be rapidly surveyed and assessed in the aftermath of destructive wildfires?

To answer these questions, we proposed a mixed-methods approach that combines remote-sensed satellite and drone imagery (captured in the field in the aftermath of destructive WUI wildfires), field reconnaissance and GIS-based analysis of the collected data, followed by a multivariate classification analysis of this information. This approach allowed us to assess the relative importance of each parameter as an indicator of wildfire vulnerability in the examined WUI areas. In turn, these conclusions could be used to propose risk mitigation strategies for both the built structures and the larger territory.

The rest of this paper is as follows. Section 2 describes the methodology, which encompasses a GIS-based examination of seven wildfire case studies in Chile, followed by a multivariate regressive analysis of vulnerability indicators to WUI fire hazards. Section 3 presents the research results, which we discuss in Section 4. Lastly, Section 5 provides the study's main conclusions and proposes paths for future investigation.

2 Methods

2.1 Case studies: collection and pre-analysis of damage data for WUI fire events

Starting in the summer fire season of 2019 and continuing until December 2022, our fieldwork team conducted data collection
campaigns in the aftermath of seven WUI events that occurred in the central region of Chile. Table 1 summarizes the case
200 studies, including information about the start date and location of each wildfire.

Table 1: WUI fire case studies considered in this study.

Case study	Date of wildfire	Location	S2 image date	S2 image tile
Limache	~~January 9,~~ 2019-01-09	32°59'21" S; 71°13'50" W	2018-12-15	T19HBD
Rodelillo	~~November 18,~~ 2019-11-18	33°3'~~12~~03'12" S; 71°~~35'0~~35'00" W	2019-11-15	T19HBD
Rocuant	~~December 24,~~ 2019-12-24	33°4'~~27~~04'27" S; 71°35'22" W	2019-12-20	T19HBD
Quilpué	~~January 15,~~ 2021-01-15	33°4'~~51~~04'51" S; 71°~~25'0~~25'00" W	2021-01-13	T19HBD
Laguna Verde	~~March 30,~~ 2022-03-30	33°8'~~24~~08'24" S; 71°40'58" W	2022-03-29	T18HYJ
Melipilla	~~December 11,~~ 2022-12-11	33°47'58" S; 71°~~18'2"~~18'02" W	2022-12-09	T19HBC
Viña del Mar	~~December 22,~~ 2022-12-22	33°2'~~44~~02'44" S; 71°33'10" W	2022-12-04	T19HBD

205 Each field survey (typically conducted one or two days after the fire) entailed a drone flight over the burnt WUI to obtain high-
resolution optical orthomosaics and digital surface models. For this task, we used two small aircrafts that could be easily
operated in urban scenarios, a DJI Mavic 2 Pro and a DJI Phantom 4. We also used the PIX4DCapture application
(https://www.pix4d.com/product/pix4dcapture/) for flight planning and drone operation, and all flights were conducted by a
licensed pilot following the Chilean Civil Aeronautic General Direction guidelines. We processed the aerial images and
210 generated the RGB orthomosaics, digital surface models (DSM), and digital terrain models (DTM) with the aid of PIX4D
Mapper software (https://www.pix4d.com/product/pix4dmapper-photogrammetry-software/). Additionally, we collected
photographic evidence of affected dwellings, and made in-situ observations of dwelling characteristics (e.g., materials, age),
urban maintenance conditions (e.g., presence of garbage dumps, road blockages), data on fire suppression resource availability

(e.g., water tanks, fire hydrants), and witnesses' accounts of the origin and progression of the fire. The combination of drone imagery and direct visual inspections allowed us to examine each built unit within the study areas and to classify it into three possible categories: (1) *no damage*; (2) *partial damage*; (3) *total damage*. All data was georeferenced using mobile GIS systems.

2.2   Generation of burnt area and vegetation layers.

For each case study, the burnt area was manually delimited based on visual analysis of the drone RGB orthomosaics. In addition, satellite images obtained by the Multi Spectral Instrument (MSI) on board Sentinel-2 (S2) were used to characterize the distribution and density of vegetation in each study region through the Normalized Difference Vegetation Index (NDVI; Rouse et al. 1973). This index quantifies the density of plan growth as the normalized difference between reflection at near-infrared (NIR) and optical (red) wavelengths, and in case of MSI/Sentinel-2 data is calculated as $NDVI = (B8 - B4)/(B8 + B4)$, where B8 and B4 are the NIR and optical red band, respectively; both with a spatial resolution of 10 m (Drusch et al. 2012). Pixels with NDVI > 0.6 were selected to generate a vegetation raster layer (De Fioravante et al. 2021). The details of image dates and tiles used in each case study are given in the last columns of Table 1.

2.2 2.3   Quantification of context and landscape variables in WUI fire scenarios

To study the vulnerability of housing units to WUI fires, we defined a set of 16 physical and spatial features to examine as potential predictors (i.e., independent variables) of the probability of total or partial damage. These variables, summarized in Table 2, were selected to encompass both the structural and preparedness conditions of individual units, the spatial arrangement of units relative to the surrounding landscape and to each other, and the availability and extent of fire suppression resources. To measure the variables, each built unit within the surveyed area was manually digitized from the drone-based orthomosaic maps and elevation models to obtain its geographic position, roof footprint, and ground height. The digital terrain models constructed from drone imaging were also used to generate the terrain slope and orientation layers. Building materials were inferred from a visual inspection conducted during the field visit, and from pre-event street view imaging. Additionally, pre-event Copernicus Sentinel 2 multispectral satellite data, public geodatabases (see Table 2), and orthomosaics were used to generate a set of vectorial layers to represent the distribution of vegetation, location of garbage dumps, evacuation roads, and availability of water supplies (e.g., water network, water hydrants). The likely point of ignition was established based on information provided by inhabitants of the burnt areas, and/or by emergency management entities like the National Service for Disaster Prevention and Response (SENAPRED) and firefighters. The levels of preparedness and maintenance of dwellings within the study areas were evaluated using our survey data, and high-resolution satellite and street view imaging accessed through Google Earth.

Table 2: ~~List~~Description of spatial variables evaluated for each housing unit in the study areas.

| ID	Variable name	Description	Source	Units/classification
Target	*damage*	Target variable. Categorical variable that identifies the degree of damage experienced by each built unit as result of the wildfire.	Orthomosaic Field survey	No damage Partial damage Total damage
1	*elevation*	The geographical elevation at each dwelling's location, relative to the lowest dwelling in the study area	DTM	Metres
2	*orientation*	Categorical variable that represents the geographical orientation of each dwelling	DTM Orthomosaic	NW / NE / SW / SE
3	*slope*	Average slope of the dwelling's terrain	DTM Orthomosaic	Degrees
4	*n_storeys*	Number of storeys of each dwelling.	Field survey	n
5	*material*	Categorical variable that classifies the dwelling's structural qualities and building materials. We define three categories, *Light*: wood, light wood, other informal materials. *Solid*: concrete, masonry. *Mixed*: combination of light and solid materials.	Field survey	Light ~~(light wood, other)~~ Solid ~~(concrete, masonry)~~ Mixed
6	*floor_area*	Ground floor area of each dwelling	Orthomosaic	Square metres
7	*area_parcel_ratio*	Measurement of the dwelling's floor area in relation to the size of the lot/parcel that it is located on	Orthomosaic	Ratio
8	*preparedness*	Binary variable that classifies the management of the dwelling's surrounding area to reduce vegetation and other potentially inflammable materials. We define two categories, *Good:* the dwelling's parcel is clear from vegetation or other visible sources of flammable material (e.g. gas tanks).	Orthomosaic Google Earth Field survey	Good ~~/poor~~/Poor

		Poor: the dwelling's parcel is occupied by vegetation, or other visible sources of flammable material (e.g. gas tanks)		
9	*maintenance*	Categorical variable that identifies whether the dwelling'sdwelling is in an acceptable overall material condition. We define two categories, *Good:* the dwelling is clear from surrounding garbage dumps or debris and does not appear to be evidently abandoned. *Poor:* the dwelling is seen to be surrounded by garbage dumps or debris, or looks evidently abandoned.	Orthomosaic Google Earth Field survey, Google Street View imaging	Yes / noGood/Poor
10	*dist_fire*	Euclidean distance between each dwelling and the likely point of fire ignition, as informed by local inhabitants and/or by emergency management entities.	Orthomosaic Field survey	Metres
11	*dist_veg*	Euclidean distance between each dwelling and the boundary of the vegetation layer delimited from pre-event multispectral satellite imagery	Sentinel 2 pre-event multispectral imagery	Metres
12	*dist_group*	Euclidean distance to the nearest border of the dwelling group	Orthomosaic	Metres
13	*dist_neigh*	Euclidean distance from each dwelling to its closest neighbour	Orthomosaic	Metres
14	*water_supply*	Binary variable that identifies if the dwelling is connected to the water supply network	Database, Chilean Superintendency of Sanitary Services	Yes/No
15	*emergency_access*	Binary variable that identifies whether the dwelling is directly located on an access road	Orthomosaic	Yes/No

16	*supression_access*	Euclidean distance between the dwelling and its closest access road	Orthomosaic	Metres

245

The areas surveyed by the drone encompass a larger area than the region of interest for the analysis of WUI fire vulnerability, and the generation of a control dataset of undamaged dwellings requires the spatial identification of the urban region that is effectively *exposed* to the hazard. Given the absence of a formal or standardized definition of the area exposed to the fire, and to avoid statistical biases, we tested six different GIS-based strategies for the selection of undamaged dwellings, considering
250   i) the full perimeter of the burnt area (BA) as mapped from our drone imagery, and ii) the convex hull (CH) of damaged units, plus buffers of 50, 100 and 150 meters around each boundary. These area and buffer definitions are illustrated in Figure 1 for the Rocuant wildfire case study. By combining all the case studies, we obtained six training datasets that differ only in the number and location of undamaged units, which were used for sensitivity analysis as explained in the next section.
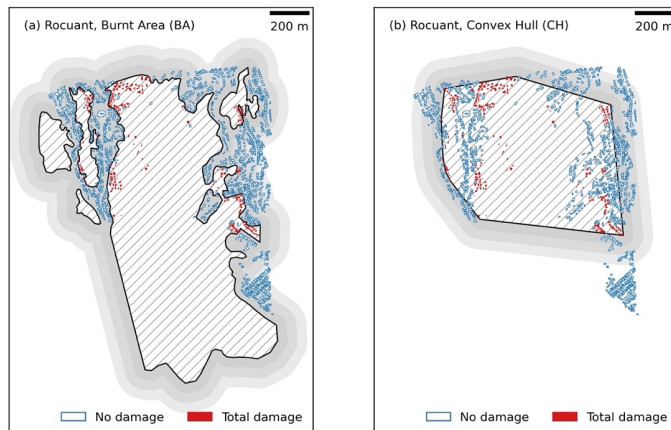
255
Figure 1: The figure illustrates the different spatial approaches considered to identify the areas affected by each fire and the undamaged dwellings for the Rocuant case study. (a) Identification of the burnt area (BA, hatched region), and buffers of 50, 100 and 150 meters (grey scale, ranging from dark to light gray). (b) Identification of the convex hull (CH) polygon of the damaged dwellings, with its respective buffers. In both panels, filled red, and non-filled blue polygons identify units with total and no damage,
260   **respectively.**

In the context of vulnerability analysis, our goal was to identify attributes of built units that could affect their probability of being damaged when exposed to WUI fires. The problem can be posed as one of supervised learning and specifically of multivariate classification, where a model is trained to predict whether a given housing unit is expected to suffer total, partial or no damage given a set of observed input variables. A trained classifier with good generalization properties can make

11

265 predictions with new data, but such models also provide valuable insights into the contribution of the individual features to the model's predictive prowess. In this study, the focus was on achieving model explainability to understand the relative importance of each dwelling attribute (individual and environmental) in determining the likely damage classification.

Several of the input variables are categorical, making the data appropriate for decision tree classification models. These models
270 divide the predictor space into several simple regions based on splitting that can be summarised in a tree scheme (James et al. 2013). In this study, we used the LightGBM classification model (Ke et al. 2017). This model was selected due to its ability to deliver shorter running times without compromising accuracy, and because it is well-suited for managing independent variables without extensive pre-processing. For model training and validation, we applied a k-fold cross validation strategy (Mosteller and Tukey 1968), stratified relative to the target variable (i.e., the damage classification of each dwelling), with 120
275 iterations to avoid overfitting and to assess performance uncertainties. These repetitions represent the outcome of splitting the labeled data into three training partitions, and one testing partition (75/25), with 30 crossed validations. As general performance metrics, we used accuracy and recall. The former assesses the overall fraction of correct predictions for a binary or multilabel dataset, and the latter measures the completeness of positive predictions (Pedregosa et al. 2012), i.e., the ability of the classifier to correctly identify dwellings that are likely to experience damage. Lastly, we used SHAP values (SHapley Additive
280 exPlanations) to interpret the model's results. According to Lundberg and Lee (2017), SHAP values, which align with human intuition, contribute to explaining a model's predictions and enhancing the interpretation of its complexity. SHAP regression studies examine the importance of each independent variable in the presence of multicollinearity. To do this, a model is trained for each independent variable, both including and excluding it. Subsequently, the predictions from these two approaches are compared. Finally, an importance value can be assigned to each feature, representing its effect on the model predictions. We
285 developed an ad hoc Python package to carry out our classification analyses, using the data analysis libraries NumPy (Berg et al. 2020), pandas (McKinney et al. 2020), scikit-learn (Pedregosa et al. 2012) and SHAP (Lundberg 2020).

3 Results

### 3.1 Summary of case studies

The field surveys conducted in this study allowed us to collect georeferenced damage cadastres and high-resolution drone
290 imagery for all wildfire events, except the event that occurred in Viña del Mar. In the case of Viña del Mar, the persistence of smoke and continued fire suppression operations precluded the operation of small drones. However, we were granted access to raw imaging captured by a larger drone operated by professionals from the Ministry of Public Works. It should also be noted that in the Melipilla case study, the fire had two different focus points, so two separate flights were required to cover the affected WUI areas. The drone imagery data collected during the field surveys were used to generate 30-cm resolution
295 orthomosaics and digital surface models for each case study. Through the analysis of these image products, we delineated the burnt areas and identified a total of 729 houses affected by fire. Of these, 648 suffered total damage, while 81 were partially

12

damaged. Additionally, we digitized 5332 built units located within the drone survey areas, but that were not impacted by the fire. In Figure 2 we show the obtained drone imaging, the footprints of all digitized dwellings, and the delimitation of the burnt areas. The extent of the burnt areas and the distribution of damaged and undamaged dwellings for each case study are detailed in Table 3. To the best of our knowledge, there are no other georeferenced, detailed, and consistent damage cadastres for these events, so we can only compare total damage statistics to general figures reported in the national press or in communications from public institutions such as the municipalities or SENAPRED. When available, these data align with the order of magnitude recorded in our results (e.g., SENAPRED 2023; Ríos 2022; Cooperativa.cl 2022). Some discrepancies among public sources, as well as between these sources and the data collected in this study are expected. These differences may be attributed to factors such as the rapid demolition and removal of debris from totally damaged dwellings, which makes them undetectable in post-event imaging. Additionally, the incompleteness of damage cadastres conducted by different governmental entities, as well as the reluctance of informal settlers to report their losses to the authorities may also contribute to the discrepancies.

Table 3: Burnt areas and distribution of damaged and undamaged dwellings for each case study.

Case study	Burnt area (Ha)	Partial damage	Total damage	No damage
Limache	12.23	1	29	449
Rodelillo	1.5	1	5	634
Rocuant	74.33	20	221	1250
Quilpué	19.78	1	5	216
Laguna Verde	36.32	14	32	500
Melipilla	87.57	23	42	477
Viña del Mar	117.53	21	314	1806
TOTAL	**349.26**	**81**	**648**	**5332**

The sixteen spatial variables detailed in Table 2 were evaluated for all dwellings reported in Table 3 and plotted in Figure 2. Analysis of the attribute distribution for affected dwellings shows that most of them were built with light materials (~70%), typically had one floor (~80%), and were mostly in poor preparedness and maintenance conditions (~60%). The multivariate statistical characterization of dwellings with partial and total damage were similar and given that in the combined dataset the number of records with partial damage was significantly lower, the *partial damage* and *total damage* categories were merged into a single *damaged* category (N=729). A similar approach was adopted by Penman et al. (2019), who examined damage data from several fire seasons in Australia and found that destroyed houses outnumbered damaged ones by a ratio of 4:1, and therefore merged them into a single category.

13
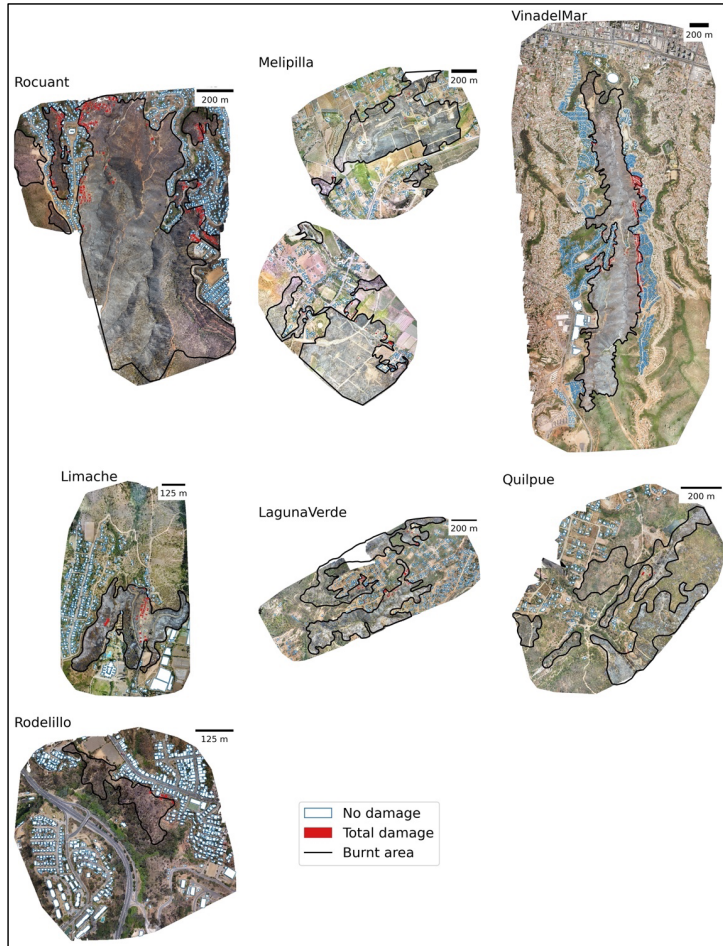
Figure 2: Drone-based orthomosaics and identification of exposed and affected building units for each case study. The orthomosaics generated from drone mapping are plotted at a spatial resolution of 1m. Damaged and undamaged units are plotted as filled red and unfilled blue polygons, respectively. The black contours delimit the total burnt area. In all images, north is up, and east is right.

3.2 Model training and sensitivity analysis

As described in section 2.3, the number and distribution of the spatial attributes for units undamaged by the fire depends on the chosen delimitation of the study area. We defined six alternative datasets for model training considering the burnt area (BA) and convex hull (CH) boundaries, with buffers of 50, 100 and 150 meters (BA50, BA100, BA150, CH50, CH100, CH150). We then carried out a sensitivity analysis to study the impact of the different delineations of the affected area on the imbalance of training datasets, the performance of the trained models, and the assessment of feature importance. As shown in Figure 3, we found that all models yielded an accuracy range of ~60-70% and had recalls ranging from ~80-90%. We note ~~than~~that in all cases, the accuracy and recall values obtained in the training dataset were only slightly higher than in the test dataset, so there is no evidence of overfitting. In both the CH and BA approaches, increasing the buffer distance naturally resulted in more imbalanced datasets with a higher proportion of *no damage* records, a trained model with increased median accuracy (+33% for BA and +13% for CH, when the buffer was increased from 50 to 150 metres), and a slight decrease in recall (-5% for BA and -3% for CH, when the buffer was increased from 50 to 150 metres). As the study region expanded, the accuracy increased, driven by a higher number of true negatives in the confusion matrix. This is misleading in terms of model performance given that its main purpose is to correctly predict and explain true positives, i.e., damage. This effect was stronger when the BA approach was used, particularly in the case of large fires (e.g., Viña del Mar and Rocuant), as it considered buildings further from the urban areas where damage effectively occurred. For example, Figure 2 shows the groups of undamaged dwellings located on the southern margins of the burnt areas in Rocuant and Viña del Mar. On the other hand, we found that recall provides a more stable and significant criteria for the selection of the final dataset to be used as a reference for assessing feature importance. Considering the trade-off between the data imbalance, accuracy, and recall of the trained model, we chose the CH100 scenario as our working definition of the study region and labelled datasets accordingly for each wildfire.
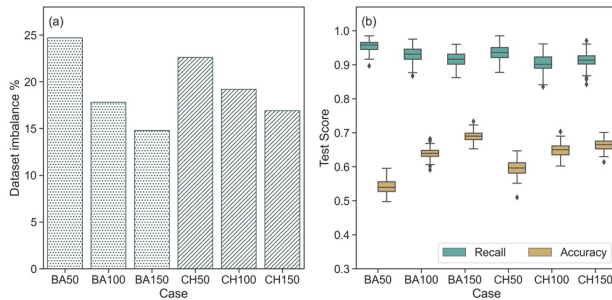

Figure 3: Comparison of dataset imbalance and model performance for six different definitions of the study regions. (a) Imbalance of the labeled dataset, expressed as the percentage of undamaged records over the total dataset. (b) Boxplots of accuracy and recall scores obtained for the test datasets with a k-fold cross validation strategy.
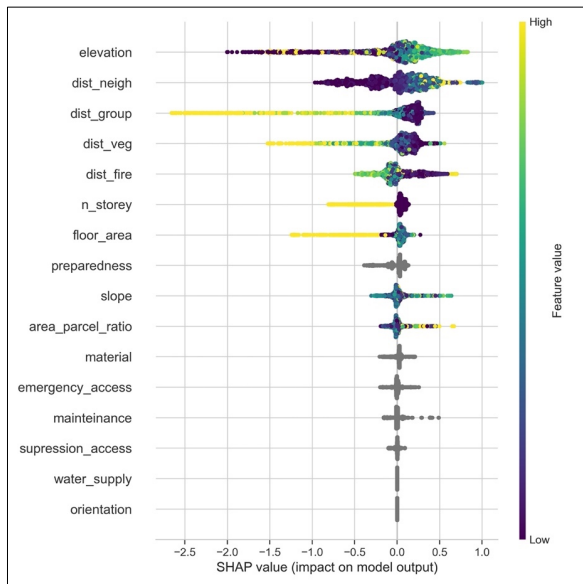
15

Figure 4: Beeswarm summary plot for the CH100 model. The y-axis indicates the feature names, ordered by importance from top to bottom, and the x-axis represents the SHAP value, or change in log odds of the target variable. Each point represents a row of data from the original dataset, and the color of each point on the graph represents the relative value of the corresponding feature, from high values (red) to low values (blue). Categorical variables were not classified as high/low, so they are plotted in gray.

3.3   Analysis of feature importance

To represent the relative importance of features for the classifier trained with the CH100 dataset, we used beeswarm plots, which summarize the SHAP values of each feature for the input dataset. The plot is presented in Figure 4. From the collection of trained models, we identified five variables that were consistently more important for predicting damage: the elevation relative ~~geographical elevation~~to the lowest unit in the surveyed area (*elevation*), the distance to the closest neighbour (*dist_neigh*), the distance to the nearest border of the dwelling group (*dist_group*), the distance to the closest vegetated area (*dist_veg*), and the distance to the likely fire ignition point (*dist_fire*). On the opposite end of the beeswarm plot are six variables that have a negligible impact on the classification outcome: *material, emergency_access, maintenance, suppression_access, water_supply,* and *orientation*. The low importance of the material feature is explained by socioeconomic context of the studied

16

regions, which correspond to underprivileged developments where a large majority of dwellings (~70-80%) are built from light materials. In an intermediate tier, five variables were grouped together (*n_storey*, *floor_area*, *preparedness*, *slope*, *area_parcel_ratio*), whose relative order of importance for the model depends on the buffer distance used in the construction of the training dataset. To provide a more detailed analysis and qualitative interpretation of the trained classifier, Figure 5 includes individual dependence scatter plots illustrating the effect of each feature on the model's prediction. This applies to all features except the six at the bottom of Figure 4. In the next section, we discuss our findings for each variable in relation to previous studies in the literature.
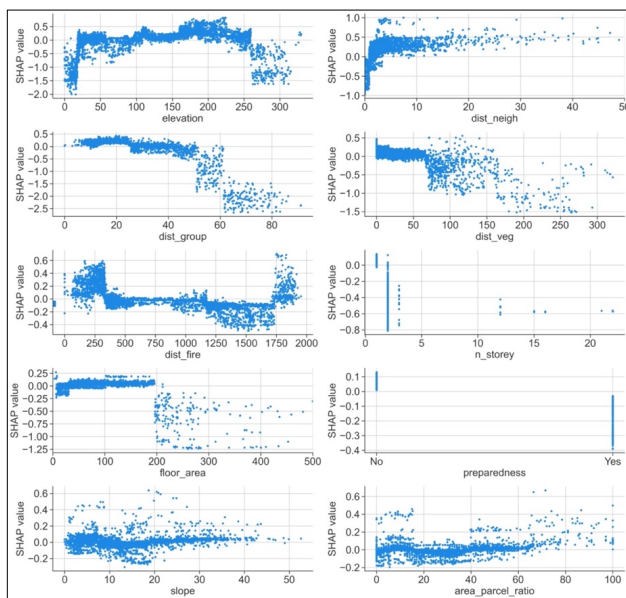

Figure 5: Data scatterplots showing the distribution of SHAP values compared with the ten most important independent variables. The x-axis represents the value of the feature, using the corresponding units provided in Table 2. The y-axis represents the corresponding SHAP value, in log-odd units

4. Discussion

Few studies have systematically examined how the physical characteristics of individual buildings relate to their vulnerability to wildfires, and furthermore, have done so systematically for several wildfire events or very large datasets. As relevant examples, As relevant examples, Dossi et al. (2022) examined a sample of 18,690 buildings exposed to wildfires in California,

17

USA (2013-2017) and Pedrógão Grande Fire Complex, Portugal (2017). Papathoma-Köhle et al. (2022) developed a physical vulnerability index based on building structural and surrounding features recorded for 423 buildings affected by the Mati fire (Greece, 2018). Penman et al. (2019) analysed building-level data for 309 units damaged by wildfires in Australia between 2001 and 2009, while Knapp et al. (2021) studied the associations between the distance to the nearest destroyed structure and vegetation for a sample of 400 dwellings in northern California affected by the 2018 Camp Fire. To the best of our knowledge, no similar studies have been conducted in Latin America and Chile, so this study contributes the first unique dataset that i) pertains to a geographical region where physical vulnerability to WUI fires has not been systematically studied; ii) includes a set of building-level variables relative to the individual structures, their location relative to vegetation, the source of the hazard, and to the urban environment, that are quantified for *both damaged and undamaged* dwellings; iii) comprises several case studies that correspond to a total of 729 damaged units, which is significantly larger than comparable studies.most of comparable studies; and iv) was constructed through an intensive use of remote sensing and image analysis methods that could be further advanced and scaled using machine learning approaches. The construction of this dataset is based on a mixed-methods approach that combines satellite and drone imagery, field survey, and a GIS-based analysis of the collected images and infrastructure data provided by public institutions. Such methods can be largely automated and scaled by leveraging open platforms for spatial analysis such as Google Earth Engine, building databases like the Microsoft Building Footprints, and state-of-the art segmentation algorithms like the Segment Anything Model (SAM). However, high-resolution post-event imaging remains a key element that requires improvement. Our experience shows that commercially available drones are an efficient alternative for generating optical orthomosaics and digital elevation models for fires under ~100 Ha, but in the case of more extensive burned areas, larger crafts are required to ensure geographical coverage and operational safety. These resources are available within different government entities that typically deploy emergency teams in the aftermath of disasters, such as municipalities and sectorial ministries. Therefore, there is a pivotal opportunity for institutions to collaborate in gathering comprehensive data for the purpose of risk assessment. As a long-term goal, a systematic mapping of the WUI based on spatially resolved aerial or satellite imagery would enable a more detailed characterization of damaged and undamaged dwellings in case of a future fire. Additionally, this mapping could include construction dates, which can be used to assess the impact of evolving building regulations and construction practices on vulnerability (e.g., Knapp et al. 2021). This is especially relevant for informal urban developments, which cannot be traced in official databases such as property tax cadastres.

Regarding model training, the accuracies obtained for the training datasets considered were in the 65-70% range, but the recall metrics were significantly higher, reaching close to 90% in all cases. Since the purpose of the model is to identify dwellings that are more susceptible to damage when exposed to WUI fires, the consequences of type II errors (false negatives) are more undesirable that type I errors (false positives). In the former, the application of the trained model to a new urban setting would imply that a dwelling likely to be damaged is incorrectly predicted to survive a wildfire and may thus be excluded from maps of vulnerable assets used as input for risk assessment and decision making. On the contrary, incorrectly predicting damage for a given unit that is not likely to suffer it would only lead to redundant safety considerations. Therefore, from an operational

18

perspective, recall provides a more appropriate performance metric for the classifier and supports using the trained model as a tool for prospective assessment of the vulnerability of urban developments close to the WUI, assuming data availability.

Moreover, the analysis of SHAP values provides a ranking of features in terms of their importance for predicting damage, which we compared to previous findings in the literature. In general, we found that the group of variables that describe the spatial arrangement of dwellings have a greater effect on damage prediction than other features, relative to the structural and preparedness conditions of the individual units, or to fire suppression resources. Although the *elevation* of each house (relative to the lowest unit in the surveyed area) appears to be an important feature in Figure 4, the average recorded elevation of dwellings (damaged and undamaged) from the Viña del Mar and Rocuant fires were higher than in other case studies. Since these accounted for a large fraction of damaged dwellings in the combined dataset, they may be driving the observed trend. In fact, although the summary of SHAP values suggests a correlation between higher elevations and damage, Figure 5 shows that this relationship is inconsistent across the case studies, with several of them with a SHAP value of 0 likely influenced by their local characteristics (similar to the case of the *slope* variable). While elevation as a fire-vulnerability feature has been examined in the literature (see Andersen and Sugg 2019 and Penman et al. 2019), these studies typically focus on large, geographical scales where elevation can influence fire-related factors like moisture and the type of vegetation, so their results are not directly comparable to our findings.

Regarding *d_neigh* (which corresponds to the distance to the closest neighbouring structure), we found that low values tend to diminish the odds of fire damage. Cohen (2000) highlighted that homes can be victims of wildland fire but also contribute to its propagation, especially when the distance between buildings is less than 40 meters. This is also supported by Gibbons et al. (2012) and Penman et al. (2019), who suggested ensuring proper separation between houses (based on their findings in Eastern Australia case studies). However, Papathoma-Köhle et al. (2022) stated that the presence of neighbouring buildings was not relevant in the construction of their Physical Vulnerability Index (PVI), which was statistically developed from the outcomes of a July 2018 wildfire in Mati, Greece, and included 423 buildings. In this respect, other research shows that higher building densities might lead to a 'protection' effect among houses. In line with this, our results for *d_neigh* can be related to Opie et al.'s (2014), which pointed out that areas with higher housing densities were less likely to be damaged during the 2009 wildfires in Bendigo, Australia. Similarly, drawing upon a study in south-eastern France, Lampin-Maillet et al. (2010) noted that fire ignition risk increased significantly when housing density decreased; therefore, they encouraged compact urban development and housing densification in the WUI.

To the best of our knowledge, the third most important feature, *d_group* (i.e., the distance of a housing unit to the nearest border of the dwelling group), has not been addressed previously in the literature. In this respect, our results show that shorter distances result in positive SHAP values, and the odds of material loss tend to increase if this distance is less than 40 m. This suggests that buildings located within a housing arrangement are less vulnerable than those along the perimeter. This further supports the argument of compact development as a protective measure.

Next, we focused on *dist_veg,* which measures the distance between each house and the closest patch of dense vegetation. As seen in Figure 4, houses located shorter distances from vegetation have slightly higher odds of damage. A more disaggregated

19

460 examination in Figure 5 shows that this effect remains consistent up to distances of ~60 m. Beyond this range, its effect on damage prediction becomes increasingly negative. This finding is consistent with the literature, which reports that reducing physical vulnerability through a 'wildfire-defensible space' minimizes the probability of fire damage. Such spaces are characterized by the alteration (treatment, clearing, or reduction) of fuels and vegetation in the immediate vicinity of a structure. For example, Syphard, Brennan, and Keeley (2014), through regression analysis in their post-fire analysis of structures in San

465 Diego County, California, suggested that defensible space distances up to 30 m effectively prevent house destruction. Our capacity to identify this small-scale spatial effect in our data was hindered by the spatial resolution of the Sentinel 2 imagery (10 m) used in the mapping of vegetation. Additionally, Penman et al. (2019) suggested that the percentage of mapped forest within 100 m of an examined structure significantly influences the probability of fire damage. Similarly, $dist\_fire$ (i.e., the distance between an examined dwelling and the likely point of fire ignition) is an important factor. While its results appear

470 more scattered, as shown in Figure 5, they clearly demonstrate that proximity to the ignition point increases the probability of fire damage, which is an intuitive result. However, this variable demonstrated significant variability across our case studies, most likely due to local geophysical characteristics (topography, built environment) and the specific manner in which the fire began and spread.

As previously mentioned, in our model, the variables characterizing a dwelling's structural conditions had relatively less

475 influence on damage classification than the spatial metrics mentioned above. Specifically, the contribution of the variable $n\_storeys$ to the SHAP values was close to zero. However, a few exceptions corresponding to high buildings (>10 storeys) built from reinforced concrete (in compliance with the Chilean building code) consistently survived the fires. Regarding the $floor\_area$ variable, the overall trend showed a negative contribution to the damage prediction when surfaces were large ($\gtrsim$ 200 m^2). This could be associated with higher income and therefore better construction conditions. Analogously, the analysis

480 of the $area\_parcel\_ratio$ suggests an association between the damage prediction and high ratios (i.e., dwellings with little open surrounding space), which are more typical of lower-income dwellings. Again, this is related to the concept of defensible space. However, the $material$ variable had a negligible impact on damage classification. Although we found some hints that individual building attributes may be useful for characterizing vulnerability, the statistical analysis was hindered by the limited variability of our dataset, which predominately consisted of single-story dwellings (~88%) built with light materials (~79%).

485 For example, Papathoma-Köhle et al. (2022) used a more detailed characterization of structure type and roof material, which they identified as the two most relevant indicators of wildfire vulnerability. However, we were not able to confirm or test these results for the Chilean case with our current data. Regarding the number of storeys or building height, this metric has been scarcely examined in the literature, perhaps due to the same lack of data variability as in our case studies. Continued observation of WUI fires in future seasons may enable a statistically significant assessment of these features as vulnerability

490 indicators. The $preparedness$ variable, per our definition, (whether houses had managed surroundings to reduce vegetation and other potentially combustible materials) allowed only two answers, 'yes' or 'no' (see Fig. 5). Our data for this variable shows that most of the houses with no preparedness had an increased likelihood of fire damage. The characteristics that we assigned

20

to the *preparedness* variable relate to the definition of 'defensible space' discussed above, for which there is a consensus regarding its importance in reducing fire vulnerability.

495   Lastly, the importance of the *slope* variable also appears to be significantly dependent on the specific study context, as shown by the scatterplot in Fig. 5. As a result, a relationship cannot be inferred between the probability of damage and the average terrain slope of each dwelling's footprint. In this regard, several authors have examined the impact of terrain slope on wildfire vulnerability. For example, Papathoma- Köhle et al. (2022) pointed out that this variable, alongside roof material and structural type, becomes the most relevant indicator for predicting wildfire vulnerability when values exceed 10º. Andersen and Sugg

500   (2019) underlined a positive correlation between slope and wildfire risk. They also reported that the cases in which the opposite trend was observed may be related to increased exposure related to the populations' tendency to settle in terrains with gentler slopes. Finney et al. (2021) argued that wildfire can spread more rapidly on a steeper slope, implying a higher risk. These studies typically provided the source from which they extracted the slope values but did not delve into how they measured them.

505   In summary, our study aligns with previous research findings on the significance of different wildfire vulnerability indicators (e.g., *dist_veg*, *dist_neigh*, *preparedness*). However, in some cases, comparisons are limited by the volume and diversity of the data. These disparities in our analysis reveal the challenge of extrapolating the selection and weighting of physical wildfire vulnerability indicators to other regions and cases. Therefore, weighted indexes developed for specific geographic areas, whether based on one or a few large or small events, cannot be directly generalized to other regions, even if they share some

510   geographical and environmental conditions. A first key step towards constructing consistent global or regional vulnerability indexes is to standardize the variables that may constitute them, as well as the spatial scales and methodologies applied in their measurement. This standardization could lead to the development of large and comprehensive datasets for statistical modelling purposes.

Lastly, our findings can contribute to improving current regulations for building design, construction, and urban planning in

515   Chilean WUI areas. As González-Mathiesen and March (2018b) point out, countries including Australia, France, Spain, and the United States have implemented policies and guidelines to enhance the resilience of sites and subdivisions in wildfire-prone areas, aiming to increase both resistance and response capacities. In contrast, current Chilean spatial planning schemes do not adequately address wildfire risk reduction. While national-level planning regulations enable generic disaster risk reduction (with a focus on hazards rather than on vulnerability), it's important to note that wildfire is not explicitly included

520   as a threat. While some municipalities do incorporate wildfire risk management in their local plans, this does not occur systematically throughout the country (González-Mathiesen and March 2023).

5. Conclusions

This study presents a local-scale analysis of the impact of WUI fires on individual dwellings as a function of their physical characteristics, their spatial arrangement relative to the urban settlement and the natural environment, and the availability and

21

accessibility of fire suppression and emergency management resources. The aim is to enhance our understanding of risk profiles in exposed WUI areas by identifying and measuring key parameters of physical vulnerability that can be systematically surveyed and assessed. This contribution will enable continued risk monitoring and provide insights for the design, planning, and governance of WUI areas in Chile.

To conduct this analysis, we developed a mixed-methods approach that combines field surveys, remote-sensed satellite and drone imagery (captured in the field in the aftermath of destructive WUI wildfires), and GIS-based analysis of the collected data, followed by a multivariate classification analysis. The methodology was applied to seven WUI fire case studies that occurred in central Chile between 2019 and 2022, and led to the generation of a detailed georeferenced dataset of 729 dwellings that experienced partial or total damage. For each of these dwellings, we assessed 16 physical and spatial attributes that could potentially influence the damage outcome in the event of exposure to wildfires. Additionally, we constructed a similar dataset for dwellings that survived the fires with no damage. We proposed various spatial methods to select such units to test the performance of classification models for different delimitations of the exposed area. We concluded that the definition that represents the best compromise between class imbalance and recall is a 100 m buffer around the smallest convex polygon enclosing the impacted dwellings, referred to as CH100. To the best of our knowledge, these data represent the first detailed characterization of WUI settlements in Latin America, and with a volume and diversity that rivals or surpasses similar studies conducted in Greece, the USA, and Australia.

With the CH100 labelled dataset as input, we trained a LightGBM binary classification model (damage/no damage) and used SHAP values to explain its results and to assess the relative importance of 16 physical dwelling parameters as indicators of wildfire vulnerability. Our findings show that the group of variables describing the spatial arrangement of dwellings has more impact on damage prediction compared to other features related to the structural and preparedness conditions of the individual units, or to fire suppression resources. Specifically, our results show that the odds of damage are reduced when the dwelling's distance to its closest neighbour is under 40m, the distance to vegetation is over ~60 m, when it is a long distance from the border of the dwelling group, and when it is a long distance from the likely fire ignition point. The analysis of SHAP dependence values for variables characterizing individual dwellings (*n_storeys, material, floor_area. preparedness*) provides some hints that attributes typically associated with less affluent homes (e.g., light material, small floor area, lack of preparedness) may increase the likelihood of damage prediction. However, more data is required to robustly confirm this association.

While our data show clear correlations between the probability of damage and some of the variables (e.g., distance to the nearest border of the dwelling group, distance to dense vegetation), others have scattered results that may be influenced by each case study's specific context.

This approach allowed us to assess the relative importance of each parameter as an indicator of wildfire vulnerability in the WUI. These findings could be used to develop risk mitigation strategies for both the built structures and the broader territorial area, and to enhance current regulations for building design, construction, and urban planning in Chilean WUI areas, which are currently underdeveloped.

Code availability

560   All code generated in this study are available upon request.

Data availability

All data generated in this study are available upon request.

Author contribution.

PA: conceptualization, funding acquisition, methodology, writing original draft.

565   JL: conceptualization, funding acquisition, methodology, writing original draft.

CG: conceptualization, writing original draft.

RR, MP: data collection and curation, formal analysis.

AO: data curation, formal analysis, software

Competing interests.

570   The authors declare that they have no conflict of interest.

References

575   AIDR. 2020. "Land Use Planning for Disaster Resilient Communities."

Andersen, Lauren M., and Margaret M. Sugg. 2019. "Geographic Multi-Criteria Evaluation and Validation: A Case Study of Wildfire Vulnerability in Western North Carolina, USA Following the 2016 Wildfires." *International Journal of Disaster Risk Reduction* 39 (October): 101123. https://doi.org/10.1016/j.ijdrr.2019.101123.

Arana, M T, A Casagrande, A Cavieres, S Espejo, Gilabert, and J H. Leon. 2018. " Prevención de Incendios Forestales:

580   Propuestas Para Una Mirada Más Efectiva e Integral."

Atkinson, Dale, Mark Chladil, Volker Janssen, and Arko Lucieer. 2010. "Implementation of Quantitative Bushfire Risk Analysis in a GIS." *Int. J. Wildland Fire* 19 (5): 649.

23

Bar Massada, Avi, Volker C Radeloff, Susan I Stewart, and Todd J Hawbaker. 2009. "Wildfire Risk in the Wildland–Urban Interface: A Simulation in Northwestern Wisconsin." *For. Ecol. Manage.* 258 (9): 1990–99.

Berg, Sebastian, Jaime Fernández del Río, Ralf Gommers, Allan Haldane, Charles Harris, Stephan Hoyer, Matti Picus, et al. 2020. "NumPy." 2020. https://numpy.org/.

Bowman, David M J S, Grant J Williamson, John T Abatzoglou, Crystal A Kolden, Mark A Cochrane, and Alistair M S Smith. 2017a. "Human Exposure and Sensitivity to Globally Extreme Wildfire." *Nat. Ecol. Evol.* 1 (3): 58.

———. 2017b. "Human Exposure and Sensitivity to Globally Extreme Wildfire." *Nat. Ecol. Evol.* 1 (3): 58.

Butt, Andrew, Michael Buxton, Rachel Haynes, and Alexander Lechner. 2009. "Peri-Urban Growth, Planning and Bushfire in the Melbourne City-Region." In *State of Australian Cities (SOAC 2009)*, 14. Canning Bridge WA , Australia: Promaco Conventions.

Caggiano, Michael D., Todd J. Hawbaker, Benjamin M. Gannon, and Chad M. Hoffman. 2020. "Building Loss in WUI Disasters: Evaluating the Core Components of the Wildland–Urban Interface Definition." *Fire* 3 (4): 73. https://doi.org/10.3390/fire3040073.

Calkin, D., O. Owen Price, and M. Salis. 2019. "WUI Risk Assessment at the Landscape Level." In *Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires*, edited by Samuel L. Manzello, 1184–95. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-52090-2.

Castillo, M. 2013. "Integracion de Variables y Criterios Territoriales Como Apoyo a La Proteccion Contra Incendios Forestales. Area Piloto: Valparaiso - Chile Central." Universidad de Cordoba, España.

Castillo, M, G Julio-Alvear, and R Garfías. 2014. "Current Wildfire Risk Status and Forecast in Chile: Progress and Future Challenges." In *Wildfire Hazards, Risks, and Disasters*, edited by D Paton, J F Shroder, S McCaffrey, and F Tedim, 59–75. Elsevier Science.

Chen, Keping, Russell Blong, and Carol Jacobson. 2003. "Towards an Integrated Approach to Natural Hazards Risk Using GIS: With Reference to Bushfires." *Environ. Manage.* 31 (4): 546–60.

Cohen, Jack D. 2000. "Preventing Disaster: Home Ignitability in the Wildland-Urban Interface." *Journal of Forestry* 98 (3): 15–21.

CONAF. 2018. "CONAF."

———. 2023. "Número de Incendios Forestales y Superficie Afectada a La Fecha." 2023. https://www.conaf.cl/incendios-forestales/incendios-forestales-en-chile/estadistica-de-ocurrencia-diaria/.

Cooperativa.cl. 2022. "Tópicos: País | Desastres Naturales | Incendios Forestales Incendio En Melipilla Ha Consumido Sobre 600 Hectáreas y Dañado Cerca de 40 Casas." *Cooperativa.Cl*, 2022. https://cooperativa.cl/noticias/pais/desastres-naturales/incendios-forestales/incendio-en-melipilla-ha-consumido-sobre-600-hectareas-y-danado-cerca-de/2022-12-12/110511.html.

Coppola, Damon P. 2015. "The Management of Disasters." In *Introduction to International Disaster Management*, 1–39. Elsevier.

24

~~DPTLI~~Dossi, Simona, Birgitte Messerschmidt, Luís Mário Ribeiro, Miguel Almeida, and Guillermo Rein. 2022. "Relationships between Building Features and Wildfire Damage in California, USA and Pedrógão Grande, Portugal." *International Journal of Wildland Fire* 32 (2): 296–312. https://doi.org/10.1071/WF22095.

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, et al. 2012. "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120 (May): 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

DTPLI. 2013. ~~"Planning Advisory Note 46 ~~"Bushfire Management Overlay Mapping Methodology and Criteria~~."~~, Planning Advisory Note 46 ." Melbourne.

Ellis, Todd M., David M.J.S. Bowman, Piyush Jain, Mike D. Flannigan, and Grant J. Williamson. 2022. "Global Increase in Wildfire Risk Due to Climate-Driven Declines in Fuel Moisture." *Global Change Biology* 28 (4): 1544–59. https://doi.org/10.1111/GCB.16006.

Finney, Mark, Sara McAllister, Torben Grumstrup, and Jason Forthofer. 2021. *Wildland Fire Behaviour*. CSIRO Publishing. https://doi.org/10.1071/9781486309092.

Fioravante, Paolo De, Tania Luti, Alice Cavalli, Chiara Giuliani, Pasquale Dichicco, Marco Marchetti, Gherardo Chirici, Luca Congedo, and Michele Munafò. 2021. "Multispectral Sentinel-2 and SAR Sentinel-1 Integration for Automatic Land Cover Classification." *Land* 10 (6): 611. https://doi.org/10.3390/land10060611.

Galiana-~~Martín~~Mart\'\in, Luis. 2017. "Spatial Planning Experiences for Vulnerability Reduction in The-Urban Interface in Mediterranean European Countries." *Eur. Countrys.* 9 (3): 577–93.

Garfías, R, M Castillo, F Ruiz, G Julio, V Quintanilla, and J Antúnez. 2012. "Caracterización Socioeconómica de La Población En Áreas de Riesgos de Incendios Forestales. Estudio de Caso: Interfaz Urbano-Forestal, Provincia de Valparaíso. Chile." *Territorium* 19: 101–9.

Gill, A Malcolm, and Scott L Stephens. 2009. "Scientific and Social Challenges for the Management Of-Prone Wildland–Urban Interfaces." *Environ. Res. Lett.* 4 (3): 34014.

Goldammer, Johann, Ioannis Mitsopoulos, Giorgos Mallinis, and Martine Woolf. 2017. "Words into Action Guidelines: National Disaster Risk Assessment Hazard Specific Risk Assessment 6. Wildfire Hazard and Risk Assessment." https://www.undrr.org/publication/wildfire-hazard-and-risk-assessment.

Gómez-González, Susana, Fernando Ojeda, and Paulo M Fernandes. 2018. "Portugal and Chile: Longing for Sustainable Forestry While from the Ashes." *Environ. Sci. Policy* 81 (March): 104–7.

González, Mauro E, Antonio Lara, Rocío Urrutia, and Juvenal Bosnich. 2011. "Cambio Climático y Su Impacto Potencial En La Ocurrencia de incendios Forestales En La Zona Centro-Sur de Chile (33º - 42º)." *Bosque (Valdivia)* 32 (3): 215–19.

Gonzalez-Mathiesen, Constanza, and Alan March. ~~2018a~~2018. "Establishing Design Principles for Wildfire Resilient Urban." *Plan. Pr. Res.* 33 (2): 97–119.

González-Mathiesen, Constanza, and Alan March. ~~2018b~~2018. "Establishing Design Principles for Wildfire Resilient Urban." *Plan. Pr. Res.* 33 (2): 97–119.

———. 2023. "Long-Established Rules and Emergent Challenges: Spatial Planning and Wildfires in Chile." *International Planning Studies* 28 (1): 37–53. https://doi.org/10.1080/13563475.2022.2136629.

Google Earth. n.d. "Google Earth."

ISO 31000. 2018. *ISO 31000:2018, Risk Management - Guidelines.* International Organization for Standardization. https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Springer Texts in Statistics. New York, NY: Springer US. https://doi.org/10.1007/978-1-0716-1418-1.

Jappiot, M, J R Gonzalez-Olabarria, C Lampin-Maillet, and L Borgniet. 2009. "Evaluacion Temporal y Espacial Del Riesgo de Incendios Forestales." In *Convivir Con Los Incendios Forestales: Lo Que Nos Revela La Ciencia* , edited by Y Birot, 43–49. European Forest Institute.

Jappiot, Marielle, José Ramón González-Olabarria, Corinne Lampin-Maillet, and Laurent Borgniet. 2009. "Evaluación Temporal y Espacial Del Riesgo de Incendios Forestales." In *Convivir Con Los Incendios Forestales: Lo Que Nos Revela La Ciencia*, edited by Yves Birot. Joensuu, Finland: European Forest Institute.

Jones, Matthew. 2020. "Climate Change as an Enabler of Wildfire." In *The DRIVER+ Project Report*, 12–17. EU H2020 Project: Driving Innovation in Crisis Management for European Resilience.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree." In *Advances in Neural Information Processing Systems*, edited by I Guyon, U Von Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, and R Garnett. Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

Knapp, Eric E., Yana S. Valachovic, Stephen L. Quarles, and Nels G. Johnson. 2021. "Housing Arrangement and Vegetation Factors Associated with Single-Family Home Survival in the 2018 Camp Fire, California." *Fire Ecology* 17 (1): 25. https://doi.org/10.1186/s42408-021-00117-0.

Lampin-Maillet, Corinne, Marielle Jappiot, Marlène Long, Christophe Bouillon, Denis Morge, and Jean-Paul Ferrier. 2010. "Mapping Wildland-Urban Interfaces at Large Scales Integrating Housing Density and Vegetation Aggregation for Fire Prevention in the South of France." *Journal of Environmental Management* 91 (3): 732–41. https://doi.org/10.1016/j.jenvman.2009.10.001.

Lu, Yongmei, Lori Carter, and Pamela S. Showalter. 2009. "Wildfire Risk Analysis at the Wildland Urban Interface in Travis County, Texas." In *Geospatial Techniques in Urban Hazard and Disaster Analysis*, 203–27. Dordrecht: Springe. https://doi.org/10.1007/978-90-481-2238-7_10.

Lundberg, Scott M. 2020. "SHAP (SHapley Additive ExPlanations)." 2020. https://github.com/slundberg/shap.

Lundberg, Scott M, and Su-In Lee. 2017. "A Unified Approach to Interpreting Model Predictions." In *Advances in Neural Information Processing Systems 30*, 4765–74. Long Beach, CA.

685 March, Alan, Graeme Riddell, Leonardo Nogueira de Moraes, Janet Stanley, Hedwig van Delden, Ruth Beilin, Stephen Dovers, and Holger Robert Maier. 2020a. "Urban Planning Capabilities for Bushfire: Treatment Categories and Scenario Testing." *Australian Journal of Emergency Management* 35: 32–40.

———. 2020b. "Urban Planning Capabilities for Bushfire: Treatment Categories and Scenario Testing." *Australian Journal of Emergency Management* 35: 32–40.

690 McKinney, Wes, Joris Van den Bossche, Tom Augspurger, Stephan Hoyer, Jeff Reback, and Phillip Cloud. 2020. "Pandas." 2020. https://pandas.pydata.org/.

Mitsopoulos, Ioannis, Giorgos Mallinis, and Margarita Arianoutsou. 2015. "Wildfire Risk Assessment in a Typical Mediterranean Wildland–Urban Interface of Greece." *Environmental Management* 55 (4): 900–915. https://doi.org/10.1007/s00267-014-0432-6.

695 Moritz, Max A, Enric Batllori, Ross A Bradstock, A Malcolm Gill, John Handmer, Paul F Hessburg, Justin Leonard, et al. 2014. "Learning to Coexist with Wildfire." *Nature* 515 (7525): 58–66.

Mosteller, F, and J.W Tukey. 1968. "Data Analysis, Including Statistics." In *Handbook of Social Psychology*, edited by G Lindzey and E Aronson. Reading, Mass.: Addison-Wesley Pub. Co.

Muñoz, P, A Guerrero, S Fuentealba, and M & Vega. 2012. "Gobierno Sospecha Que Incendio En Quillón Fue 700 Intencional."

Oom, D., D. de Rigo, H. Pfeiffer, A. Branco, D. Ferrari, R. Grecchi, T. Artés-Vivanco, et al. 2022. "Pan-European Wildfire Risk Assessment." Luxembourg. https://doi.org/10.2760/9429.

Opie, Kimberley, Alan March, Justin Leonard, and Glenn Newman. 2014. "Indicators of Fire Vulnerability: Risk Factors in Victorian Settlements." Melbourne.

705 Papathoma-Köhle, M., M. Schlögl, C. Garlichs, M. Diakakis, S. Mavroulis, and S. Fuchs. 2022. "A Wildfire Vulnerability Index for Buildings." *Scientific Reports* 12 (1): 6378. https://doi.org/10.1038/s41598-022-10479-3.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2012. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12 (January): 2825–30.

710 Penman, Sandra H., Owen F. Price, Trent D. Penman, and Ross A. Bradstock. 2019. "The Role of Defensible Space on the Likelihood of House Impact from Wildfires in Forested Landscapes of South Eastern Australia." *International Journal of Wildland Fire* 28 (1): 4. https://doi.org/10.1071/WF18046.

Reszka, Pedro, and Andrés Fuentes. 2015. "The Great Valparaiso Fire and Fire Safety Management in Chile." *Fire Technol.* 51 (4): 753–58.

715 Ríos, Diego. 2022. "Un Fallecido, 46 Viviendas Destruidas y 184 Personas Damnificadas: Balance Del Incendio Forestal En Melipilla Fuente: Emol.Com - Https://Www.Emol.Com/Noticias/Nacional/2022/12/13/1080891/Catastro-

Incendio-Melipilla.Html." *EMOL*, December 13, 2022. https://www.emol.com/noticias/Nacional/2022/12/13/1080891/catastro-incendio-melipilla.html.

Romero-Calcerrada, Raul, C. J. Novillo, J. D.A. Millington, and I. Gomez-Jimenez. 2008. "GIS Analysis of Spatial Patterns of Human-Caused Wildfire Ignition Risk in the SW of Madrid (Central Spain)." *Landscape Ecology* 23 (3): 341–54. https://doi.org/10.1007/S10980-008-9190-2/TABLES/5.

Rouse, J.W., R. H. Hass, J. A. Schell, and D. W. Deering. 1973. "Monitoring Vegetation Systems in the Great Plains with ERTS." In *3rd ERTS Symposium, NASA SP-351*. Washington DC.

Sakellariou, Stavros, Athanassios Sfougaris, Olga Christopoulou, and Stergios Tampekis. 2022. "Integrated Wildfire Risk Assessment of Natural and Anthropogenic Ecosystems Based on Simulation Modeling and Remotely Sensed Data Fusion." *International Journal of Disaster Risk Reduction* 78 (August): 103129. https://doi.org/10.1016/j.ijdrr.2022.103129.

San-Miguel-Ayanz, J., E. Chuvieco, J. Handmer, A. Moffat, C. Montiel-Molina, L. Sandahl, and D. Viegas. 2017. "Climatological Risk: Wildfires." In *Science for Disaster Risk Management 2017: Knowing Better and Losing Less*, edited by K. Poljanšek, M. Marín Ferrer, and T. Clark, I. De Groeve, 294–305. Publications Office of the European Union. https://purl.org/INRMM-MiD/c-14445352.

Sarricolea, Pablo, Roberto Serrano-Notivoli, Magdalena Fuentealba, Marina Hernández-Mora, Francisco de la Barrera, Pamela Smith, and Óliver Meseguer-Ruiz. 2020a. "Recent Wildfires in Central Chile: Detecting Links between burned Areas and Population Exposure in the Wildland Urban." *Sci. Total Environ.* 706 (135894): 135894.

———. 2020b. "Recent Wildfires in Central Chile: Detecting Links between Burned Areas and Population Exposure in the Wildland Urban Interface." *Science of the Total Environment* 706 (March): 135894. https://doi.org/10.1016/j.scitotenv.2019.135894.

SENAPRED. 2023. "Monitoreo Incendio Forestal – Comuna Viña Del Mar." https://senapred.cl/informate_conte/monitoreo-incendio-forestal-comuna-vina-del-mar/.

Sirca, Costantino, Franco Casula, Christophe Bouillon, Belén Fierro García, Monica Maria Fernández Ramiro, Beatriz Vila Molina, and Donatella Spano. 2017. "A Wildfire Risk Oriented GIS Tool for Mapping Rural-Urban Interfaces." *Environmental Modelling & Software* 94 (August): 36–47. https://doi.org/10.1016/J.ENVSOFT.2017.03.024.

Sivrikaya, Fatih, Bulent Saglam, Abdullah Akay, and Nuri Bozali. 2014. "Evaluation of Forest Fire Risk with GIS." *Polish Journal of Environmental Studies* 23 (1): 187–94.

Syphard, Alexandra D., Teresa J. Brennan, and Jon E. Keeley. 2014. "The Role of Defensible Space for Residential Structure Protection during Wildfires." *International Journal of Wildland Fire* 23 (8): 1165. https://doi.org/10.1071/WF13158.

28

Syphard, Alexandra D., Jon E. Keeley, Avi Bar Massada, Teresa J. Brennan, and Volker C. Radeloff. ~~2012~~2012a. "Housing Arrangement and Location Determine the Likelihood of Housing Loss Due to Wildfire." *PLoS ONE* 7 (3). https://doi.org/10.1371/journal.pone.0033954.

———. 2012b. "Housing Arrangement and Location Determine the Likelihood of Housing Loss Due to Wildfire." *PLoS ONE* 7 (3). https://doi.org/10.1371/journal.pone.0033954.

Tampekis, Stergios, Stavros Sakellariou, Palaiologos Palaiologou, Garyfallos Arabatzis, Apostolos Kantartzis, Chrisovalantis Malesios, Anastasia Stergiadou, Dimitrios Fafalis, and Evangelos Tsiaras. 2023. "Building Wildland–Urban Interface Zone Resilience through Performance-Based Wildfire Engineering. A Holistic Theoretical Framework." *Euro-Mediterranean Journal for Environmental Integration* 8 (3): 675–89. https://doi.org/10.1007/s41207-023-00385-z.

Tedim, Fantina, Vittorio Leone, Malik Amraoui, Christophe Bouillon, Michael Coughlan, Paulo Delogu Giuseppe and Fernandes, Carmen Ferreira, et al. 2018. "Defining Extreme Wildfire Events: Difficulties, Challenges, And." *Fire* 1 (1): 9.

Úbeda, Xavier, and Pablo Sarricolea. 2016. "Wildfires in Chile: A Review." *Glob. Planet. Change* 146 (November): 152–61.

UNISDR. 2009. "UNISDR Terminology on Disaster Risk Reduction." Geneva. UN. Office for Disaster Risk Reduction.

Urrutia-Jalabert, Roc\'\io, Mauro E González, Álvaro González-Reyes, Antonio Lara, and René Garreaud. 2018. "Climate Variability and Forest Fires in Central and south-central Chile." *Ecosphere* 9 (4): e02171.

USDA. 2024. "Wildfire Risk - Understand Risk." 2024. https://wildfirerisk.org/understand-risk/.

Zhu, Zhe, Curtis E Woodcock, John Rogan, and Josef Kellndorfer. 2012. "Assessment of Spectral, Polarimetric, Temporal, and Spatial for Urban and Peri-Urban Land Cover Classification Landsat and SAR Data." *Remote Sens. Environ.* 117 (February): 72–82.

Zong, Xuezheng, Xiaorui Tian, and Lei Fang. 2022. "Assessing Wildfire Risk and Mitigation Strategies in Qipanshan, China." *International Journal of Disaster Risk Reduction* 80 (October): 103237. https://doi.org/10.1016/j.ijdrr.2022.103237.