The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

A data analysis of the Chilean housing stock and the estimation of uncertainty in indoor air quality in Chilean houses

Constanza Molina Carvallo

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

September 2019

Abstract

Air pollution is currently one of the leading causes of mortality and morbidity worldwide and forecasts predict that it will remain so until 2040. Exposures to ambient fine particles ($PM_{2.5}$) were responsible for 103.1 million disability–adjusted life–years (DALYs) in 2015. However, indoor concentrations of pollutants, including $PM_{2.5}$, can be higher than those found outside, and so the indoor environment may have a significant effect on personal exposures. Dwellings are particularly important because people spend over 70% of their time in them.

The Chilean housing stock comprises 6.4 million dwellings of great diversity, because of the country's variable geography. Dwellings change with the local climate, occupants' lifestyles and behaviours, and with the different materials they are constructed from. They evolve as technologies are used to save energy and improve occupants' quality of life, and improve their performance in response to new policies and standards.

This thesis explores the indoor air quality (IAQ) across the Chilean housing stock, by developing a set of statistically representative *archetypal* dwellings and quantifying uncertainty in their characteristics using available data sources, such as national censuses. Eight archetypes, representing 35% of the stock, are modelled using `CONTAM` and simulated using a probabilistic sampling approach to generate distributions of indoor $PM_{2.5}$ concentrations, ventilation rates, and associated heat losses. A sensitivity analysis is used to identify the most influential inputs.

The results show that the variability in the physical and environmental parameters and windows use influence $PM_{2.5}$ exposures, ventilation rates, and energy losses. Therefore, it is important to understand how much these results vary depending on these parameters. This will help to establish effective national standards and guidelines for IAQ and energy demand reduction in order to avoid negative health impacts at a population scale.

This work contributes to knowledge by (i) characterising the Chilean housing stock and presenting a set of archetypal buildings to represent it; (ii) presenting a model and modelling framework for evaluating the Chilean housing stock probabilistically; (iii) predicting uncertainties in occupant exposures to $PM_{2.5}$, and dwelling ventilation rates and energy losses across the Chilean housing stock; (iv) identifying the most important parameters that affect the predictions; (v) contextualizing and interpreting the results; (vi) showing how the model and its predictions can be used to inform and evaluate the impacts of policies, and improve the IAQ and environmental performance of dwellings; and by (vii) identifying the need for future measuring, surveying, and data gathering exercises.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr Benjamin Jones, for his continued support and encouragement throughout these four years, for challenging me to perform at my best, for all the time and patience during the supervision meetings, and for sharing his curiosity about the world and science. Special thanks should be given to Professor Ian P. Hall, for his enormous kindness, words of wisdom, and for always sharing his knowledge, and to Professor Max H. Sherman, for his guidance and constructive feedback on this project.

I would also like to extend my thanks to Michael and Sergio, for their cooperation and critical review in the early stages of this project, to Gustavo, Felipe, and Cristian who provided me with technical support.

I am deeply thankful to Mom, Dad, María Jesús, Guillermo, and Andrés for supporting me and keeping me close to their hearts. Also, to Mandy and Ken for all their love, the special meetings, the million moments of joy we shared, for making me feel at home, and being my family during these four years and without whom this work would not have been possible; to Verónica and Cristina for their infinite words of encouragement from the very beginning and for their companionship.

I am very grateful to CONICYT/Becas Chile for funding this research.

Publications

Jones, B & Molina, C. 2017. Indoor air quality. *Reference Module in Earth Systems and Environmental Sciences*, 2, 197-207.

Molina, C, Jones, B, Kent, M & Hall, IP. 2017. The Development of Archetypes to Represent the Chilean Housing Stock. *In: 38th AIVC Conference: 6th TightVent & 4th Venticool Conference*, Nottingham, UK.

Molina, C, Jones, B, Kent, M & Hall, IP. 2018. A stochastic approach to estimate uncertainty in pollutant concentrations in an archetypal Chilean house. *In: 39th AIVC Conference: 7th TightVent & 5th Venticool Conference*, Antibes Juan-Les-Pins, France.

Jones, B, Phillips, G, O'Leary, C, Molina, C, Hall, I & Sherman, M. 2018. Diagnostic barriers to using $PM_{2.5}$ concentrations as metrics of indoor air quality. *In: 39th AIVC Conference: 7th TightVent & 5th Venticool Conference*, Antibes Juan–Les–Pins, France.

Molina, C, Jackson, A, & Jones, B. 2019. The evaluation of real-time indoor environment parameters measured in 297 Chilean dwellings. *In: 40th AIVC Conference: 8th TightVent & 6th Venticool Conference*, Ghent, Belgium.

Molina, C, Jones, B, Kent, M & Hall, IP. 2020. A data analysis of the Chilean housing stock and the development of modelling archetypes. *Energy and Buildings, 206*.

Contents

List of Figures

List of Tables

Nomenclature

$\dot{Q}$	Airflow rate	m/s
μ	Population standard deviation	
ϕ	Decay rate	s^{-1}
σ	Population or sample standard deviation	
$\tilde{\mu}$	Sample mean	
E_t	Total exposure	
F_{INF}	Infiltration factor	n.d.
H_I	Heat loss due to exfiltration	
n	Flow exponent	
C_{in}	Indoor concentration	$\mu g/m^3$
C_{out}	Outdoor concentration	$\mu g/m^3$
a	Air exchange rate	s^{-1}
C	Flow coefficient	$m^3 s^{-1} Pa^{-n}$
P	Penetration factor	n.d.
S	Source strength	mg
ACH	Air changes per hour	
ASHRAE	American Society of Heating, Refrigerating, and Air–Conditioning Engineers	
BC	Black carbon	
BP	Building Permits	
BTEX	Benzene, Toluene, Ethyl-benzene and Xylene	

C_P Wind pressure coefficient

C_V Coefficient of variation

CASEN Encuesta de Caracterización Socioeconómica Nacional

CDF Computational Fluid Dynamics

COPD Chronic Obstructive Pulmonary Disease

DALYs Disability-Adjusted Life Years

DMC Department of Meteorology of Chile

EC Elemental carbons

ENUT Encuesta de Uso del Tiempo (Use of Time National Survey)

ER Emission rate

ETS Environmental Tobacco Smoke

GBD Global Burden of Disease

GDP Gross Domestic Product

GHG Green House Gas

GIS Geographic Information System

GUI Graphical user interface

HAP Household air pollution

HSEM Housing stock energy model

I/O Indoor/Outdoor ratio

IAQ Indoor Air Quality

IEA International Energy Agency

INE National Institute of Statistics

LCA Life Cycle Assessment

LHS Latin Hypercube Sampling

MINVU Ministerio de Vivienda y Urbanismo, the Chilean Ministry of Housing

N Sample size

NAAQS National Ambient Air Quality Standards

NHST Null hypothesis significance testing

NIST National Institute of Standards and Technology

NL Normalised leakage

NRC National Research Council Canada

OC Organic carbons

RENAM Red Nacional de Monitoreo de Viviendas

SBS Sick Building Syndrome

SDGs Sustainable Development Goals

SE Standard error or standard deviation of the mean

USEPA United States Environmental Protection Agency

VOC Volatile Organic Compound

WHO World Health Organization

CHAPTER 1

Introduction

1.1 Energy and Green House Gas emissions in Chile

As countries try to reduce their Green House Gas (GHG) emissions to meet climate change targets, governments tend to seek different ways to optimise efforts through coordinate actions and resources required by every end–use sector. Chile, a South American country with a population of 17.6 million people (INE, 2018a) and a population growth rate of <1% (The World Bank, 2018), is reaching that state of development. In January 2017, Chilean authorities ratified the Paris Agreement and set two energy–related targets as part of its Nationally Determined Contribution: these are (i) to reduce the carbon intensity of its economy to 30% below its 2007 level by 2030; and (ii) to reduce all GHG emissions by 35–45%.

Between 2003 and 2013, the carbon intensity of its economy (tonnes of CO_2 equivalent per million US dollars of GDP) decreased from 198 $tCO_2e/mUSD$ to 81 $tCO_2e/mUSD$ (Center), 2018), although latterly the decline has ceased in contrast to the decreasing trends of many International Energy Agency (IEA) member countries (IEA, 2018). This is attributed to the increasing carbon intensity of electricity generation, which surpassed the IEA average in 2011. The largest consumer of electricity is the Industry sector (62.6% of total consumption), followed by the Transport (17.6%), and Residential (17.4%) sectors. The same three sectors are responsible for the majority of the nation's Total Final Consumption (the aggregate of all of the energy used to provide energy services): Industry (43.0%); Transport (33.6%); and Residential (15.4%). These values are consistent with the world average. In IEA member countries, in contrast, the residential sector contributes up to 28% of their total energy use and GHG emissions, attributed to an increasing and more demanding population and an increase in the time people spend indoors (Pérez-Lombard *et al.* , 2008). These tendencies have also been seen in Chile and so the energy use in the residential sector is expected to rise in the future, intensified by the economic growth.

Whilst pragmatic policies designed to reduce national energy demand and GHG emissions –and so to de–couple the energy demand from economic growth– might first be directed at the Industry and Transport sectors, the IEA (2018) believes that building energy codes and standards are "essential for improving their energy performance and comfort", and for providing long term energy savings and health benefits for householders.

1.2 The ambient air

The composition and concentration of contaminants in the ambient air vary temporally and spatially according to several factors, such as: topography, wind condition, sources of emission in the area or nearby, time of day, season, among others. The air can be considered contaminated "when a substance is present where it should not be or at concentrations above background" (Chapman, 2007). And pollution is "contamination that results in or can result in adverse biological effects to resident communities" (Chapman, 2007). Many countries have their own air quality standards and indices. In Chile, five criteria pollutants are monitored in different cities: Particle Matter below 10 micrometers in diameter (PM_{10}), Sulfur Dioxide (SO_2), Nitrogen Oxides (NO_X, NO, NO_2), Ozone (O_3) and Carbon Monoxide (CO). These contaminants are used for declaring states of ambient emergency when limits are exceeded. In those cases, some actions take place for reducing their concentration, such as restricting vehicle use and industry activities, or banning the use of household wood–stoves.

The effect of ambient air pollution in population health is well documented, and is now considered one of the major causes of mortality and morbidity globally. One example is the association of lung cancer, respiratory and cardiovascular diseases with exposures to fine particles (particles below 2.5 microns in diameter, $PM_{2.5}$). The health effects of this exposure on the global population has been estimated at 103.1 million disability–adjusted life–years DALYs in 2015 (Cohen *et al.* , 2017). $PM_{2.5}$ is generally emitted by the combustion of solid and liquid fuels. The main sources of $PM_{2.5}$ in urban areas are industry, vehicles and buildings. When higher comfort requirements due to higher household incomes are not accompanied by a more energy efficient building stock or a shift to cleaner energy, there will inevitably be an increase in the energy demand, which impacts on both air pollution and population's health (Schueftan & González, 2015).

Although much has been improved in Chile regarding the outdoor air quality during the last decade, some dense cities in the centre and south still show high, and for some of them increasing, levels of air pollution (Molina *et al.* , 2017). $PM_{2.5}$ concentrations are especially high in winter and autumn where combustion of wood in

household stoves for cooking and heating are still highly prevalent, whereas Ozone, which is produced by chemical and photochemical reactions is higher in spring and summer due to higher temperatures.

1.3 Housing policy in Chile

Chile is one of the most seismically active countries in the world. During the last century, building codes have evolved in order to give new building structures greater resistance to earthquakes and to reduce failures. The Chilean construction sector has learned from experience, starting with the Chillán earthquake of 1928, since when, new structural typologies and stricter structural calculations have been required by law in order to reduce the earthquake vulnerability of structures.

The continental territory has a length of 4,300 km (from north to south) and an average width of 177 km. It is bounded on the West by the Pacific Ocean and on the East by the Andes mountain range. All of which results in 15 different climates, ranging from *warm dessert* in the north to *Tundra* in the south. Because of the different latitudes and topographies, the main temperature gradients towards the four cardinal points can range from 6°C(E–W) up to more than 15°C (N–S). The same applies to the relative humidity, and so different design strategies are necessary to provide hygrothermal comfort and to promote energy savings, and strategies that could affect the indoor air quality. The country is geographically and administratively divided into 15 regions, within which the Metropolitan region is the main and most populated one.

Figure 1.1 shows some important events in history of the Chilean housing stock. Historically, Chilean housing policy has focused its efforts and public resources on eradicating the quantitative deficit of low–income households (particularly on families living in a house with irreparable damage and houses with more than one household). More recently, it has moved to eradicate the qualitative deficit, by providing basic services, decreasing overcrowding, and improving the quality of building structures and elements. Included in the latter is the improvement of the energy performance of the new stock. To tackle it, a national code governing thermal regulation was implemented in 2000, which includes a minimum requirement for the thermal insulation for the envelope of new dwellings, and a maximum of glazing area percentage (MINVU, 2007). Dwellings built before 2000 had no thermal requirements and are generally considered to be as uninsulated, although some may have been insulated. This requirement varies along the length and width of the country depending on local climatic factors. And for houses built before the thermal regulation was issued, a subsidised weatherisation program has been carried out since 2009 by the Chilean government to insulate low–income residential buildings and to improve their air-

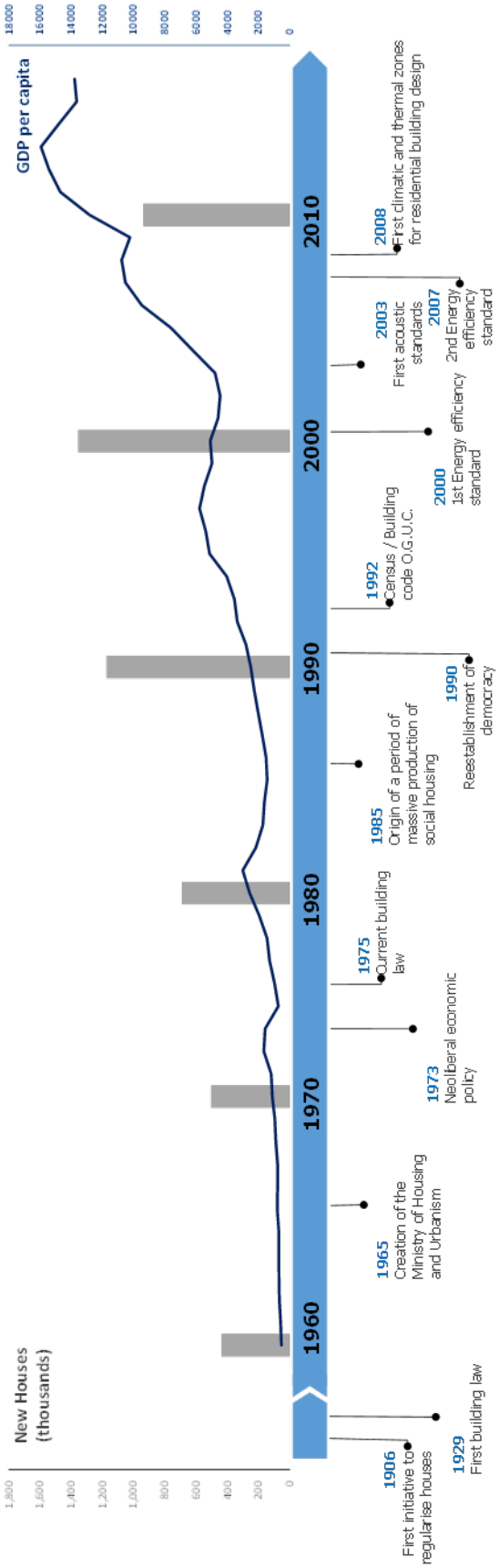


Figure 1.1: Timeline of the Chilean housing development. Source: Image by Author, GDP data from The World Bank (2018).

tightness. These regulations solely concern the housing stock, and energy efficiency standards now exist for social and private dwellings, applying to around 80% and 20% of all newly constructed dwellings respectively (IEA, 2018).

Although those improvements might be beneficial to their occupants and the environment, a greater proportion of houses were built before these regulations were adopted or are not eligible for the subsidised weatherisation program and so they need to be part of any study that aims to predict or improve the air quality of the Chilean housing stock.

1.4 Energy demand and heat loss in buildings

The main energy demand in buildings as end–use form is during their operational phase, which accounts for 90–95% of the total energy used (Sartori & Hestnes, 2007) and 70–90% of the total environmental impact during their life cycle.

In dwellings, the primary energy demand is for heating. A low–energy building can be considered to have an operational energy demand below 70 kWh/m^2 year for heating and an overall demand of around 200 kWh/m^2 year (Sartori & Hestnes, 2007). The energy demand in Chilean houses is mainly for hot water (33%), heating (19%) and cooking (15%) (CDT, 2010). The estimation for heating is at 37 kWh/m^2/year, which is approximately half of the heating demand in a low–energy building, and 197 kWh/m^2 year for the overall energy demand. Moreover, this low energy demand can be associated to low comfort levels. Recent data from 293 houses across the country (RENAM, 2018) shows that indoor air temperature in the heating season is commonly lower than the WHO (World Health Organization) recommendation of 18°C. A strong relationship of the indoor temperatures with the household income may suggest that these houses suffer from fuel poverty. Income inequality is high in Chile compared to other countries. Although it has decreased during the past three decades, from a 56.2 Gini index in 1987 to a 47.7 in 2015, Chile is still ranked 23rd in inequality of 158 measured countries (The World Bank, 2018).

Common interventions in the Chilean housing stock for reducing energy demand are by limiting the heat loss through the envelope and increasing the efficiency of their services. The heat lost through an element of the envelope due to conduction depends on the conductivity of the material, its surface area and the temperature difference between the indoor and outdoor. The Chilean building code, as presented above, regulates this heat loss by limiting the thermal transmittance of the buildings elements (MINVU, 2007).

Apart from losing heat through the building materials, buildings also suffer loses

through ventilation and infiltration, defined as the controlled and uncontrolled exchange of air, respectively, between a building and its local environment. Ventilation is required to maintain acceptable levels of indoor air quality by diluting air contaminants to improve occupant thermal comfort and satisfaction with respect to the indoor air, and to provide oxygen for combustion. In Chile, ventilation is only included in the design of rooms with gas appliances, by the regulation for the installation of gas services and gas meters, DS N° 66 (SEC, 2007).

Preventing infiltration and reducing ventilation rates in buildings are also common strategies to reduce energy demand and to reduce pollutants from outdoor sources coming into a house, but it is also subject to high levels of uncertainty and unclear effect on other parameters, such as on health (Hamilton *et al.* , 2015) and on air quality (Derbez *et al.* , 2018; Milner *et al.* , 2014). Under–ventilated rooms may have a high concentration of pollutants, which may lead to health issues. Ventilation and infiltration in naturally ventilated buildings is driven by two natural forces: wind pressure and temperature differences. Both parameters are difficult to control and there is little knowledge of airflow rates in houses at stock scale. Infiltration, however, can be a primary source of outdoor air and sometimes the only form of ventilation, especially in naturally ventilated houses during the heating season, when windows tend to be closed to save energy. A knowledge of ventilation and infiltration in houses can be used to estimate heat loses during the heating season, associated CO_2e emissions, and costs.

1.5 Indoor air quality in buildings

According to the USA standard ASHRAE 62.2 for ventilation and air quality, an acceptable indoor air quality is the ”air toward which a substantial majority of occupants express no dissatisfaction with respect to odour and sensory irritation and in which there are not likely to be contaminants at concentrations that are known to pose a health risk”. Contaminants are defined as ”a constituent of air that may reduce the acceptability of that air” (ASHRAE 62.2, 2016). Therefore, to know more about the indoor air quality in an enclosed volume, the concentration of individual pollutants must be known. However, the composition and toxicity of the indoor air change from house to house and over time according to occupants' lifestyles and habits, building technologies and appliances (Prasauskas *et al.* , 2014). Therefore, housing and public health interventions that aim to improve indoor air should account for the heterogeneity of occupants, houses and the air within them. So far, there is no general agreement in Chile on either the levels considered to be acceptable or on criteria pollutants, and international health–related guidelines and reference values are continuously varying

as new scientific information becomes available.

The Commission of Social Determinants of Health, organization established by the World Health Organization (WHO) to study health equity worldwide, states that home and work conditions are two of many social determinants of health (Marmot *et al.* , 2008). People spend around 70% of their time in their homes, where $PM_{2.5}$ is considered the most important pollutant (Borsboom *et al.* , 2016) because of the doses and its potential damage to human health. However, evidence on the toxicity and on health effects of exposures to indoor particles matter is still lacking and so far the impacts have been estimated using methods and parameters for the outdoor air. The most effective mitigation strategy against occupant exposure to indoor pollutants is *source control*, but this is not always possible for a number of reasons, such as cultural traditions or economic circumstances. Cooking is a predominant source for $PM_{2.5}$, but in Chile, it is common to heat houses using stoves (42%), gas heaters (24%) and paraffin (10%). Many of them vent indoors, likely to reduce the indoor air acceptability , and so their contribution to the total exposure need to be quantified. Modern stoves currently in use in Chile have a relatively tight smoke chamber to minimise emissions into the house (Schueftan & González, 2015). Although the smoke is intended to be completely emitted outside, it could enter the house through infiltration and ventilation openings, contributing to most of the exposures to outdoor contaminants. Moreover, even if a heater vents to outdoors it still contributes to exposure when people are outdoors. Here, low ventilation rates might benefit both energy demand and indoor air quality.

Therefore, it is essential to quantify the contribution of known emitters to indoor pollutant concentrations and to test remediation measures, such as trickle vents, cooker hoods or windows opening. To do so, this work will ignore outdoor contaminants and concentrate on the contribution of indoor sources only.

Changes made to a buildings' structure to reduce its energy demand can produce negative impacts on indoor air quality (see Figure 1.2). When considering the building sector, dwellings are a focus of attention because people spend most of their lives in them (McCurdy & Graham, 2003). Also, concentrations of some pollutants are found to be higher indoor than outdoors (Chen & Zhao, 2011; Cometto-Muñiz & Abraham, 2015), and so there is likely to be a greater degree of exposure to them indoors (e.g. Bruinen de Bruin *et al.* (2008); Guo *et al.* (2004)), such as $PM_{2.5}$, formaldehyde, and benzene. Those that could pose a public–health risk are known as *contaminants of concern.* Air quality and energy demand in buildings depend on a range of factors such as the building design, use, and environment. Indoor exposures and energy demand are then affected by three systems, which makes the analysis a complex task because they involve: (i) a heterogeneous, and sometimes peculiar, building stock;

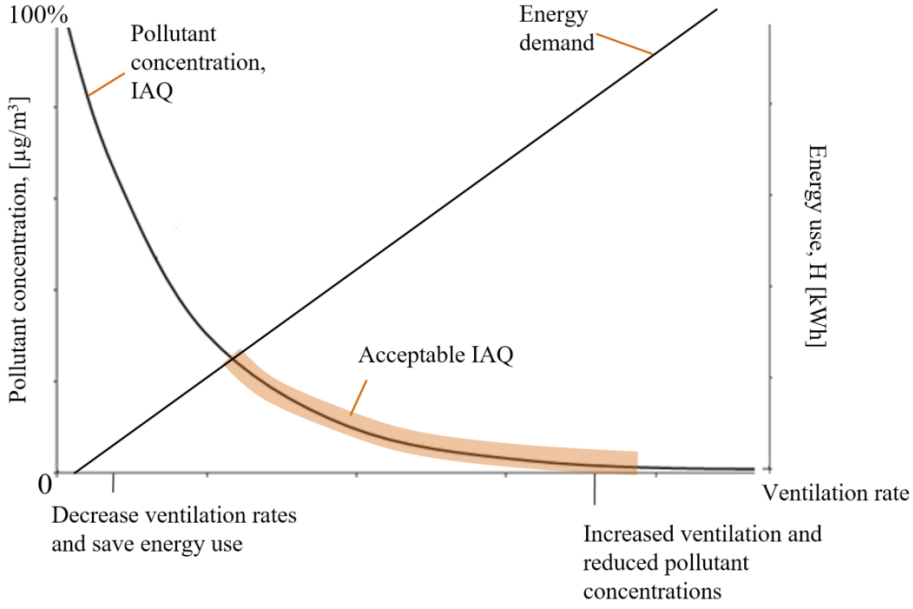


Figure 1.2: Indoor air quality and energy demand in relation with ventilation rates in a building (Image based on Liddament & Air Infiltration, Agence internationale de l'énergie (1996)).

(ii) a population with varied characteristics, composition and behaviours; and (iii) the environmental parameters that influence the building and occupant behaviours which are both difficult to identify and quantify. These three systems constantly interact with each other, and their parameters and characteristics need to be considered in order to analyse their relationships.

1.6 The Chilean housing stock

The Chilean housing stock comprises 6.4 million residential units of great diversity. Its characteristics vary according to the local weather conditions and the availability and the affordability of building materials. This variation is considerable because of Chile's geography. Furthermore, dwelling stocks change and evolve over time according to multiple internal and external factors. The building design adapts to the climate and to its occupants lifestyle and characteristics, it incorporates new materials and technologies, and change its overall performance according to new standards and demands.

Chilean new energy targets for the housing stock represent a great challenge (see Section 1.1). The unintended impacts that those interventions could have on other indoor variables and the population health require a deeper understanding of the current status of the indoor air quality in the stock and the establishment of benchmarks

for a set of indicators.

Therefore the research question raised by this study is what indoor concentrations of $PM_{2.5}$ are to be expected in the Chilean housing stock? A deeper understanding of the behaviour of the indoor air quality in the current housing stock and a comparison against international standards will enable us to know how likely the Chilean houses are to reach unacceptable levels of pollutants and so to have an impact on occupant health.

1.7 Aim and objectives

The aim of this thesis is to develop a tool capable of estimating uncertainty in pollutant concentrations in the Chilean housing stock. In order to achieve this, the study has the following objectives and steps:

The **first objective** is to characterise the housing stock of Chile by defining a set of representative houses (i.e. *archetypes*). In order to achieve this objective, an understanding of the variability of housing types and properties is required. All available information on Chilean houses will be gathered and combined and used to characterise the stock, and to define archetypes that can be used to model it.

The **second objective** is to estimate uncertainty in pollutant concentrations in the archetypes and, therefore, across the Chilean housing stock. This requires an appropriate model of pollutant transport, and a suitable statistical approach that captures stock variability and parametric uncertainty. The estimated concentration can then be compared against a suitable health metric for each pollutant.

The **third objective** explores the relationship between ventilation and infiltration rates required to dilute pollutants and the energy demand required to condition indoor air. If the Chilean housing stock is to develop sustainably, it must simultaneously minimise its energy demand and its impact on occupant health.

The **fourth objective** is to analyse the relative importance of the model inputs, such as building geometry, occupant behaviour, and local environmental conditions. A sensitivity analysis is used to rank the inputs by their effect on predicted pollutant concentrations, ventilation rates, and energy demand. The quality of data for the important inputs can be assessed and, where necessary, targeted by future data gathering exercises.

1.8 Thesis outline

The thesis outline will present the different stages that were planned for this research in order to fulfil the aim and objectives presented above. This thesis is structured as follows:

Chapter 1 Introduction: presents an overview of the field of study, the research problem and objectives of the study.

Chapter 2 Literature review: presents the existing knowledge on indoor air quality in dwellings and methods for assessing the indoor air quality in a stock of houses.

Chapter 3 Data and Methods: identifies existing sources of data that can be used to describe elements of the Chilean housing stock and to form the archetypes. It also highlights areas of data paucity, thus informing future surveys and research. It presents the methods for data acquisition, for characterising the Chilean stock, for developing the archetypes and for simulating the mass transfer and heat loss in houses.

Chapter 4 Developing the Chilean national archetypes: presents the process to development of the archetypal dwellings for the country, by categorising the information on the houses according to relevant factors and weighted by type. The archetypes are finally reported as a documented database containing the building archetypes with statistically representative values of parameters related to indoor air quality (IAQ) and energy demand that were used in the simulation process. It finishes by identifying and describing the dwelling archetypes that were modelled in the current study.

Chapter 5 Modelling process: describes the stochastic modelling, the data processing and the sensitivity analysis used to show those parameters that most influence model outputs.

Chapter 6 Modelling and analysis outcomes: summarises the simulations and the sensitivity analysis results. The main estimates and their statistics are shown for both indoor air quality and heat loss.

Chapter 7 Discussion and Conclusions: Discusses the findings and how the archetypes might then be used to predict and evaluate the impacts of this exposures on the energy demand, GHG emissions, air quality, and occupant health for a stock of houses or the whole country. It analyses the key drivers that were targeted for indoor air quality remediation and the ones that can be the subject of future field surveys to better define the uncertainties in their actual values.

CHAPTER 2

Literature Review

2.1 Introduction

This chapter presents the existing knowledge relevant to indoor air quality in dwellings and methods for assessing the indoor air in a stock of houses. It is organised into four sections. Section 2.2 identifies the ambient air pollution as an environmental and health hazard, and as an object of research. It also presents some criteria pollutants used as indicators of its quality and for regulating related activities, their emissions and concentrations. Section 2.3 describes the personal exposure approach to the air pollution and its variability among different microenvironments. Section 2.4 The Indoor Air, describes relevant characteristics of the indoor air and factors that might affect its quality. Section 2.5 Modelling the indoor air, presents different methods for modelling the indoor air in a house. Particular focus is placed on multi–zone modelling. Finally, Section 2.6 presents methods for modelling the indoor air for a housing stock.

In this literature review, the focus was over the modelling approach for indoor air quality studies rather than on the measuring techniques. Thus, current issues related to the equipment and the way the data is interpreted are not considered neither as part of this project nor in this review.

2.2 Ambient air pollution

The ambient air is composed of a complex mixture of compounds with different concentrations and toxicities. And it has been associated as one of the leading risk factor for mortality and morbidity in the world. Nearly 99% of the volume of the air we breathe consists of nitrogen and oxygen. The remaining 1% corresponds to other gases and particles. Some of them are common constituents (argon, carbon dioxide, or neon) and harmless in normal concentrations, while others (called *contaminants*) may

damage the environment and human health even in relatively low concentrations.

The ambient air quality is usually reported using a list of pollutants (also called *criteria pollutants* due to their known effect on health and welfare): particulate matter, also known as *particles*, ground-level Ozone (O_3), carbon monoxide (CO), Sulphur Dioxide (SO_2) and nitrogen dioxide (NO_2). Particles, a type of pollutant rather than a single compound, are normally reported according to their mass concentration. For example, using the mass fraction of particles with aerodynamic diameters smaller than 2.5 or 10 μm, as $PM_{2.5}$ or PM_{10}, respectively.

To simplify the information on the impacts of exposures to ambient contaminants on the population, some integrated indices have been developed. They usually convert time–series data on air pollutant concentrations into a common and single index and their associated health outcomes. One example is the Air quality index (*AQI*) developed by The U.S. Environmental Protection Agency (USEPA). The AQI is computed for specific locations using the the EPA standard (EPA, 2018). Concentration of five ambient air pollutants ($PM_{2.5}$, PM_{10}, O_3, NO_2 and CO) are converted into a common and dimensionless value that goes between 0 and 500. In Santiago, Chile, the annual average AQI is 70 (interpreted by the standard as *moderate*) and varies annually from 10 (*good*) to 142 (*Unhealthy for sensitive groups*). Within the air components, $PM_{2.5}$ and O_3 are the two main contributors to the air pollution. Although the AQI index might be useful for alerting the general population when the ambient air reaches short–term unhealthy levels, it relies on the information and location of the stationary monitoring stations, which might not relate to personal exposures (Bravo-Linares *et al.* , 2016). Additionally, reporting a single value by converting the known outcomes of only five pollutants into a common index for short–term health impact does not provide a comprehensive impact assessment, or information for chronic impacts.

Moreover, hundreds of different components can compose PMs. Among them, typical components of the ambient PM are: Sulphates, Nitrates, ammoniums, sodium and chloride ions, elemental carbons EC, organic carbons OC, and a range of mineral materials and water-soluble components. Fractions of ambient particles vary according to the sources, location, and meteorological conditions. They can be emitted from a wide range of sources (primary sources) or formed in the atmosphere (secondary particles), so the strategies to control them may be diverse (Tucker, 2000). For instance, most ambient particles in the fine fraction (PM with an aerodynamic diameter below 2.5μm) are emitted during high–temperature metallurgical and combustion processes and formed within the atmosphere, whereas coarse particles (PM with a diameter between 2.5 – 10μm) generally come from windblown and road dust, pollen and spores, and some industrial particles (Tucker, 2000). They can also change in size and shape through physical and chemical processes, such as nucleation, coagulation or conden-

sation. Ultrafine particles (PM with a diameter bellow 0.3μm) are more in number, but fine and coarse particles are higher in mass (Heal *et al.* , 2012) and they are commonly treated separately.

2.2.1 Health impacts of exposures to ambient air pollution

There is strong evidence supporting the impact of air pollution on people's health (Cohen *et al.* , 2017; Landrigan, 2017). The diversity in size, composition and behaviour make their health impact assessment through toxicology a complex task (Cassee *et al.* , 2013). Epidemiology, normally associates time–series ambient concentrations of pollutants (assumed as the population exposure) to daily health outcomes, sometimes including a certain exposure time. This had led to greater attention to fine particles. They are considered more harmful because they penetrate deeper into the lung system and so legislation and policies are mainly focused on controlling and reduce their emission. Results of epidemiological studies that give exposure–response coefficients are analysed and summarised by the World Health Organization (WHO) (Gakidou *et al.* , 2017).

Concentrations of *criteria pollutants* exceed the WHO annual threshold in some cities in Chile, and show negative health impacts on both mortality and morbidity. Cifuentes *et al.* (2000) show ambient concentrations of criteria pollutants were significantly associated with daily mortality rates in the general population of Santiago. The effect ranged from 4% to 11%, which depends on the pollutant, season and meteorological conditions. Ilabaca *et al.* (1999) show a positive relationship between ambient concentrations of criteria pollutants and children respiratory–related emergency visits in Santiago. The strongest relationship was with $PM_{2.5}$ levels, which ranged from 10 to 156 $\mu g/m^3$. An increase of 45 $\mu g/m^3$ of $PM_{2.5}$ daily mean was related to an increase of 2.7–6.7% in emergency visits. Later, Pino *et al.* (2004) showed that an increase of 10 $\mu g/m^3$ of ambient $PM_{2.5}$ was related to a 5–9% increased risk of wheezing bronchitis in young children.

Cakmak *et al.* (2007, 2009) estimated the effects of acute increases of PM_{10}, Ozone, Sulphur dioxide, CO, NO_2, and $PM_{2.5}$ components on mortality in Santiago's population. They found a positive and statistically significant relationship between pollutant concentrations and related daily mortality, particularly in the elderly (>85 years). Here, the strongest association was seen for elemental carbon EC, a $PM_{2.5}$ component. Leiva G *et al.* (2013) estimated that the risk of hospital admissions in Santiago for cerebrovascular causes increased by 1.28% for 10 $\mu g/m^3$ increase of $PM_{2.5}$ using data between 2002 and 2006; risk in winter was found to be higher, with 48.0 $\mu g/m^3$ increase in $PM_{2.5}$ associated with 6.34%. Conversely, international studies, found an increase

of 1.78%, 1.04% and 1.56% in The USA, Finland and Denmark respectively. Finally, Prieto-Parra *et al.* (2017) studied the prevalence of respiratory symptoms in children and exposure to air pollution in the winters of 2011 and 2012. Ambient concentrations of criteria pollutants and $PM_{2.5}$ components were obtained from government monitoring stations. Here, $PM_{2.5}$, NO_2, and O_3 were found to have a strong association with exacerbation of respiratory symptoms in asthmatic and non–asthmatic children. The mean and standard deviation of ambient concentration of $PM_{2.5}$ and NO_2 were $30 \pm 12.5\,\mu g/m^3$ and 40.2 ± 14.7 ppb (7 days averages) respectively, below the current Chilean standard but above the WHO guidelines (see Table 2.1).

A well–known example of the impact of air pollution on population health is The Global Burden of Disease (GBD) study. The 2018 up–date with data from 2016 showed that ischemic heart disease, stroke, lower respiratory infection, COPD, and lung cancer were the 1st, 2nd, 3rd, 9th, and 13th leading causes of years of life lost YLL worldwide. All of which have been linked to air pollution as a risk factor. Furthermore, the first three are forecasted to remain in the same position by 2040, whereas COPD is expected to rise to 4th position, and lung cancer to 9th. Cohen *et al.* (2017) used the data on $PM_{2.5}$ ambient concentrations gathered by this study and its association with those four health outcomes given by the epidemiological cohort studies. Nevertheless, these exposure–response relationships are estimated for low ambient concentrations and short–term exposures that might be exceeded in many locations or during certain activities or events, such as cooking, cleaning, wood burning from chimneys, and wildfires. Moreover, the composition of fine particles are not accounted for in the assessment.

Burnett *et al.* (2014) developed an integrated relative risk (RR) model for long–term higher concentrations (means ranging from 38 to 166 $\mu g/m^3$). It included information on RR of four combustion types (four known sources of $PM_{2.5}$): Ambient air pollution, second hand smoke, household air pollution and active smoking, with four associated causes of mortality in adults: Ischaemic heart disease (IHD), stroke, chronic obstructive pulmonary disease (COPD), and lung cancer. The study also gave estimates of population attributable fractions for every country included in the burden of disease study. The Organisation for Economic Co–operation and Development (OECD) forecasted the economic consequences of outdoor air pollution (including $PM_{2.5}$ and ozone projections) on population health and agriculture by 2060 (OECD, 2016). The report projected an increase of pollutant emissions if the current economic growth and energy demand continue to rise at the current pace. For the estimation, health impacts in the form of health care and premature deaths, lost working days (reduction of productivity), increase in GDP costs, and lower crop yields were included. Both $PM_{2.5}$ as black carbon (BC) and ozone emissions are estimated to increase, especially in South

and East Asia. Indoor air pollution in dwellings was not included in the report but it projects a reduction in organic carbon (OC) from a lower energy demand and the use of cleaner and more efficient technologies for cooking and heating. This assumption considered the use of LPG or ethanol for cooking and heating instead of open fires. In Chile, the majority of the housing stock use LPG or gas for cooking (CASEN, 2015), and so major improvements might be seen only from switching from wood–stoves for heating.

2.2.2 Chilean regulation

Concentrations of criteria pollutants in Chile are regulated by the Ministry of Environment. Primary regulations are established for protecting public health. The current primary standards and limits given in Table 2.1:

Table 2.1: Chilean standards and the WHO recommendations for ambient concentration of criteria pollutants. Time–series data on ambient concentrations from are publicly available for each criteria pollutant and others pollutants

Contaminant	National Limit (mean)	National Standard / year of publication	International recommendation (WHO *et al.* , 2000, 2006) (mean)
CO	30 $\mu g/m^3$ P_{99}th 1 h **	DS 115/2002	100 mg/m^3 15 min
	10 $\mu g/m^3$ 8 h **		60 mg/m^3 30 min
			30mg/m^3 1 h
			10mg/m^3 8 h
Pb (Lead)	0.5 $\mu g/m^3$ yearly *	DS136/2001	0.5 $\mu g/m^3$ yearly
NO_2	100 $\mu g/m^3$ yearly **	DS114/2003	40 $\mu g/m^3$ yearly
	400 $\mu g/m^3$ 1 h **		200 $\mu g/m^3$ 1 h
O_3	120 $\mu g/m^3$ 8 h **	DS112/2002	100 $\mu g/m^3$8 h
PM_{10}	50 $\mu g/m^3$ yearly **	DS59/1998	20 $\mu g/m^3$ yearly
	150 $\mu g/m^3$ 1 h **		50 $\mu g/m^3$ 24 h
$PM_{2.5}$	20 $\mu g/m^3$ yearly **	DS12/2012	10 $\mu g/m^3$ yearly
	50 $\mu g/m^3$ 1 h **		25 $\mu g/m^3$ 24 h
SO_2	50 $\mu g/m^3$ yearly **	DS 113/2002	20 $\mu g/m^3$ 24 h
	250 $\mu g/m^3$ 24 h **		500 $\mu g/m^3$ 10 min

* during the last 2 years.

** during the last 3 years.

The last standard to be incorporated was $DS12/2012$, that regulates concentrations of $PM_{2.5}$ in the ambient air. Supporting scientific evidence on chronic and acute health impacts at national and international levels were used during its development. Additionally, a study on the economic impacts (benefit–cost analysis) of decreasing ambient fine particles concentrations was requested by the authorities from a third party in order to compare different scenarios based on 2009 values. Net benefits related to the new limits were estimated to be \$30.6 billion dollars a year (50th centile; CI(90): $-$M\$5 $-$ 61,700) per year over a period of 30 years. Positive changes in premature deaths (es-

timated in 11 months), number of hospitalisations, health care visits and working days lost were contrasted against intervention and mitigation costs by type of source identified in a group of cities (Dictuc, 2009). The limits established in the current standard are considered the first stage in providing a better quality of life, and will be reviewed in the near future.

Source identification is a useful tool for policy–making and source control. Sax *et al.* (2007) took daily samples of $PM_{2.5}$ in Santiago during two monitoring campaigns in 1998 and 2003. Annual means decreased over the two times period (from 41.8 to 35.4 $\mu g/m^3$), yet concentrations during winter season remained high –three times higher than in warmer seasons– and above recommended values. The composition of fine particles showed a significant change, associated to the implementation of remediation policies, mostly related to transport. The authors advised stricter control of transport and industrial emissions in order to keep improving the ambient air quality.

Toro *et al.* (2014) identified three factors that lead to high concentrations of particles in Santiago in winter: the meteorological conditions, combustion from heating, and wood burning. Another study was carried out by Kavouras *et al.* (2001) in five regional cities in Chile. They measured $PM_{10}$and $PM_{2.5}$ concentrations throughout 1998, in order to identify and quantify their sources and contaminant species. Copper smelters and oil refineries were identified as the most important contributors to $PM_{2.5}$, followed by vehicles and wood burning. Annual mean $PM_{2.5}$ and PM_{10} concentrations exceeded the WHO recommendation in all cities.

Since then, more actions have been taken in order to reduce $PM_{2.5}$ concentrations. Tailored measures are taken after a city is declared "saturated zone", when concentrations of criteria pollutants surpass the established limits. Particularly, $PM_{2.5}$ concentrations in Santiago have decreased over the last two decades (Toro *et al.* , 2014). Koutrakis *et al.* (2005) analysed and measured fine and coarse PMs ($PM_{2.5}$, $PM_{2.5--10}$ and PM_{10}) daily sampling means in four locations in Santiago. They found that $PM_{2.5}$ levels have decreased by half from 1989 and 2001. They also found seasonal and weekly patterns –concentrations were higher during the colder seasons and from Tuesday to Friday– mainly affected by traffic and industries. The change in the source also meant that the composition of the particles also changed, and now particles are lower in sulfur, lead and other anthropogenic elements.

Similar results were estimated by Jorquera *et al.* (2000). They showed that the most considerable concentration drops were seen after the implementation of the new transportation system in 1989, when the annual averaged $PM_{2.5}$ concentrations started to decrease between -5% and -7%.

Jorquera & Barraza (2012) characterised and compared $PM_{2.5}$ in Santiago between 1999 and 2004; 95 samples were taken in 1999 and 117 in 2004. The contribution of each source to $PM_{2.5}$ was assumed by identifying their key species. Important reductions in the absolute $PM_{2.5}$ concentrations were seen between the two samples. Stricter source control might have helped this reduction. Results showed no statistically significant change in the relative composition of the fine particles, and that the major contributor continued to be vehicles (28% in 1999 to 31% in 2004) and wood burning (25% in 1999 to 29% in 2004). It also showed that sources surrounding Santiago are important, such as wood burning and metallurgical smelters. Similar analysis was carried out by Jorquera & Barraza (2013) in a northern industrial city, Antofagasta. Their analysis showed that a cement plant and soil dust were the two greatest contributor to ambient $PM_{2.5}$. Because the city has desert climate, the collected soil dust carries multiple species, such as heavy metals and marine aerosol elements.

2.3 Personal exposures

According to the Haber's Law of total exposure, total cumulative exposures depend on the concentration of the pollutant, or quantity of element present, at a specific location i (also called a *microenvironment*) and the time spent in it t_i, i.e. $C_i \cdot t_i$ (Gaylor, 2000; McCurdy, 2015). Some type of contact between the pollutant and the person is needed, so transmission of the element may occur. Because the concentration varies with time, a more accurate expression is:

$$E_t = \sum_i E_i + \sum_j E_{pact} \tag{2.3.1}$$

$$E_t = \frac{1}{T}(\sum_i \int C_i dt_i + \sum_j \int C_{pact,j} dt_j) \tag{2.3.2}$$

where E_t is the total personal exposure to any pollutant in all microenvironments, E_i is the time-weighted personal exposure in each microenvironment i, and E_{pact} is the time-weighted personal exposure to the pollutant emitted by personal activities j other than in the house. T is the total time spent in each microenvironment and in each personal activity, and C are the concentrations of the pollutant recorded in each location (Wilson *et al.* , 2000).

Normally, exposure models are defined for a specific age and gender due to their relationship with the activity. McCurdy & Graham (2003) analysed US activity data and identified cohort variables that showed a statistically significant effect of the time spent indoors, outdoors and in vehicles for each cohort. They recommend that exposure studies take into account: age and gender firstly, followed by one level of physical

activity of individuals. The daily maximum temperature, months of the year, and the day of the week are also important drivers.

Klepeis *et al.* (2001) analysed time activity data for US citizens ($N = 5{,}678$). They classified the data into 10 locations. Figure 2.1 shows the variability of the location given by the respondents. It also shows that most of the people spend most of their time indoors (average 87%), commonly at home (average 69%) but that it can vary from 34 to 98%. Koehler *et al.* (2018) also studied personal exposures to air pollutants, and its variability by microenvironment or location (home, work, transit, eateries, and others). 44 adults of different ages and gender were involved with a total of 373 sampling days. Results showed great variability within-person, which depends on the location. While the highest exposures to $PM_{2.5}$ were in transit and in eateries due to proximity to traffic and cooking sources, the higher total integrated exposure were at home.

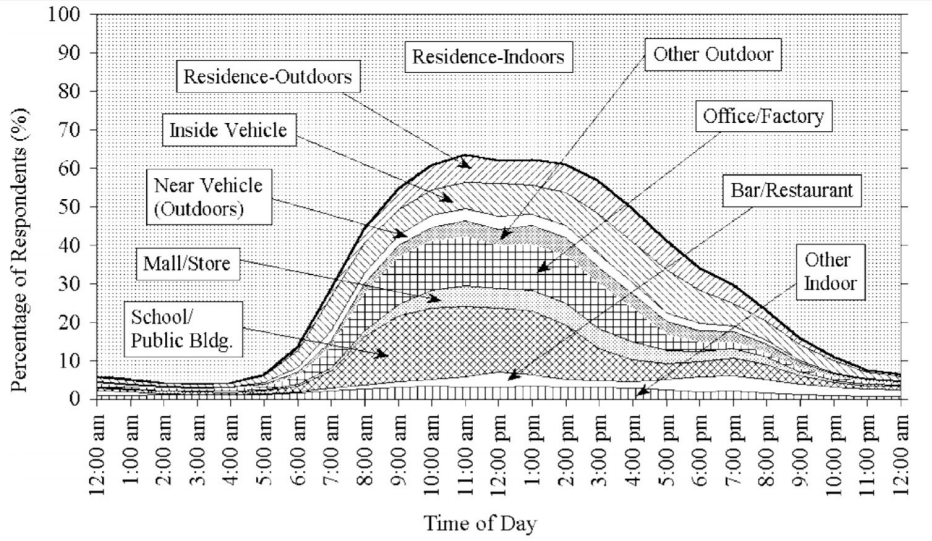


Figure 2.1: Stacked plot with the percentage of people at each location over the day. The bigger area corresponds to "Residence–Indoors". It shows that the percentage of the respondents in their residence varied from approximately 38% near midday to 95% around midnight (Klepeis *et al.* , 2001).

Some examples of how variable personal exposures can be are shown in Delgado-Saborit (2012) and Steinle *et al.* (2015). They studied personal exposures to black carbon BC (markers of combustion–related PMs) and $PM_{2.5}$, respectively, in different locations using non–expensive real–time sensors. Both associated the highest exposures with cooking and commuting activities, and indoor/outdoor (I/O) ratios above one. Figure 2.2 shows individual concentrations of $PM_{2.5}$ during a) May and b)November, of 2013 Steinle *et al.* (2015).

Both field and modelling studies might account for variability of pollutant concentration according to the location, yet time–activity data is limited. Moreover, when

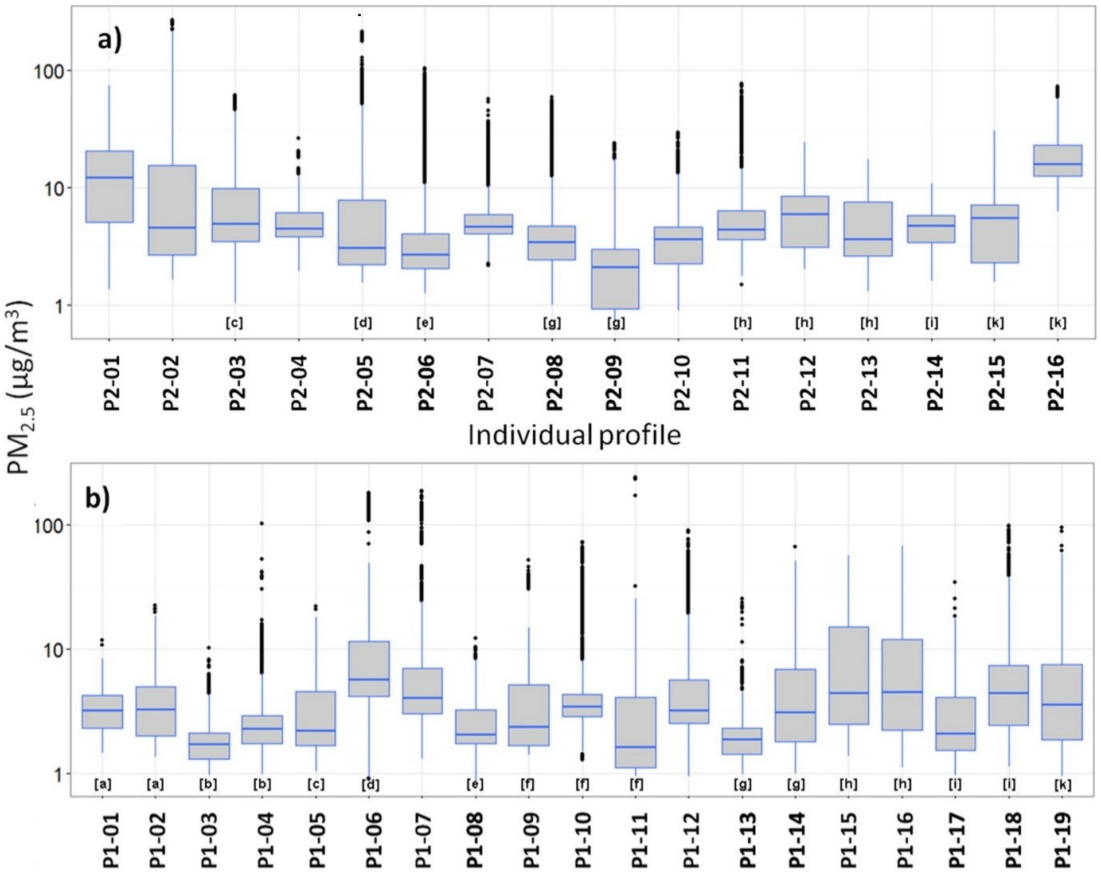


Figure 2.2: Distribution of personal exposures to $PM_{2.5}$ for different subjects during a) Spring and b) Autumn. The higher concentration levels (averages and standard deviations) were found in the residential buildings and the lowest at work. (Steinle *et al.* , 2015).

analysing health impacts toxicologically, time–series data on personal exposures are not the end point. These data are then used to calculate personal doses according to several factors, such as age, health status, metabolic rate and breathing rate, and follow the path from the exposure to the pollutant to the effect in the organism, until the adverse effect on a specific or several organs are obtained (e.g. a non–linear dose-effect curve). Finally, dose–response curves are obtained from this analysis at a population scale. To date, a detailed time–activity analysis of the Chilean population has not yet been undertaken. This lack of knowledge restrict the analysis of individual and community exposures, and the impacts of such exposures on occupants' health.

2.4 Indoor air quality in dwellings

Several studies state that we spend the majority of our time indoors, mostly in dwellings (e.g. Burke *et al.* (2001); Klepeis *et al.* (2001); McCurdy & Graham (2003); Yang *et al.* (2011)). This makes indoor air the main source of exposure to air pollution. The Com-

mission of Social Determinants of Health, organisation established by the WHO to study health equity worldwide, states that home and work conditions are two of many social determinants of health (Marmot *et al.* , 2008). Household air pollution has been an important subject of interest because it can reach unhealthy levels, often higher than outdoors. Besides, houses are where children, the elderly and those in poor health, the most susceptible fraction of the population, spend most of their time. Nevertheless, indoor air pollution is generally not regulated beyond the chemical composition of some household products. The United Nations' Sustainable Development Goals (SDGs) are a set of 17 international priorities to provide "peace and prosperity for people and the planet, now and into the future" (UN, 2019). Goal 8 aims to to ensure healthy life and wellbeing for people of all ages, which includes a health–related indicator of air pollution mortality. The 2018 up–date of the indicators gave an estimate of progress and the likelihood that each (195) country will achieve the goals by 2030. Chile shows a 71% and 94% of progress on air pollution mortality ($PM_{2.5}$) and on household air pollution, respectively, but only 14% on tobacco smoking. They estimated the probability of Chile attaining the target related to household air pollution by 11% (Lozano *et al.* , 2018). It is worth noting that household air pollution in this metric only accounts for exposures to solid fuels used for cooking, but in Chile 42% of the stock uses wood–burner stoves for heating.

A mass balance equation has frequently been used to relate the indoor concentration of particles with certain parameters (Ott *et al.* , 2006). For a well-mixed single zone:

$$\frac{\mathrm{d}C_{in}}{\mathrm{dt}} = P \cdot a \cdot C_{out} - (a + \phi) \cdot C_{in} + S/V \tag{2.4.1}$$

where P is the particle penetration factor. It is assumed to be 1 when particles are not lost by infiltration, or when windows are opened. a is the total air exchange rate through the envelope, C_{out} the outdoor particle concentration in mass per unit of volume, ϕ the total decay rate, C_{in} the indoor concentration, and S the source(s) strength. Here, C_{out}, C_{in}, a and S variables are time dependant, whereas P and ϕ depend on the particle fraction and normally considered constant over time. In this equation, the ambient and non–ambient particles are considered to be equivalent, and chemical reactions or re–suspension are not included.

Among all indoor pollutants, particulate matter PMs are considered the most important (Logue *et al.* , 2011a). Indoor particles differ from ambient particles in a number of important ways, such as in sources, composition, toxicity, and concentrations. This has been shown through temporal patterns and spatial variability. Given these varied properties and conditions, and until associations between ambient particles and indoor–generated particles are better understood, a separate treatment in the analysis is advised (Wilson *et al.* , 2000). To do so, the first term of Equation 2.3.2 must

account for both exposures to ambient pollution and exposures to non–ambient pollution, depending on whether the pollutants have outdoor or indoor origin. Exposures to the ambient air occurs both outside and inside when pollutants of outdoor origin are transported indoors via infiltration and ventilation (Wilson *et al.* , 2000). In order to separate the analysis of indoor and outdoor sources, either the first or third term of equation 2.4.1 must be omitted (no outdoor or indoor sources are included, and so either C_{out} or S/V are cancelled, respectively). Thus, assuming steady state conditions and no indoor sources, the fraction of the ambient PMs that passes through the envelope is known as the *infiltration factor*, F_{INF}, which is, in some cases, equal to the I/O ratio. I/O ratios give a simplify idea of the indoor and outdoor concentration relationship:

$$F_{INF} = \frac{C_{in}}{C_{out}} = \frac{P \cdot a}{a + \phi} \tag{2.4.2}$$

Finally, using the third term of the equation 2.4.1, the contribution of an indoor source to the indoor concentration is:

$$C_{in} = \frac{S/V}{a + \phi} \tag{2.4.3}$$

All these terms can be highly variable and uncertain, and so there is much uncertainty in the relationship between the building's characteristics, the environment and the composition and interaction of a mixture of pollutants were various sources are present.

2.4.1 International studies on indoor pollutant concentrations

Indoor contaminants can be classified according to their physical state (gaseous or particulate), their origin (primary or secondary), their known impact on health (criteria or non–criteria pollutants), or their type of source (*outdoor sources*, such as infiltration or underground soil, *indoor sources*, such as stoves or heaters, or *re–suspended*, such as by occupant's movements or cleaning activities). Sources are defined by ASHRAE 62.2 as "*an indoor object, person, or activity from which indoor air contaminants are released, or a route of entry of contaminants from outdoors or sub-building soil*" (ASHRAE 62.2, 2016). Similarly, indoor air can contain multiple contaminants emitted by different sources (primary contaminants) or formed within the indoor atmosphere (secondary contaminants).

During the last 70 years, new materials and technologies have been incorporated in the building industry in order to improve buildings' performance. Some of these changes, along with new lifestyles, have impacted the indoor air quality. Weschler

(2009) presents an overview of contemporary indoor sources and building characteristics that have affected the indoor air since the 1950s. He describes how the presence of new chemicals emitted by new materials and consumer products, changes in personal habits and changes in building air–tightness have altered the composition of the air and also occupants' exposures. For instance, low infiltration in airtight buildings can retain pollutants for longer and increase their concentrations if they are under–ventilated. Air conditioning systems that recirculate air without filtration, increase moisture problems and pollutants concentration. Conversely, restrictions on the use of certain chemicals in building materials or consumer products and changes in occupant activities (such as formaldehyde and indoor smoking) have showed a decrease of some pollutants.

Among the pollutants identified as hazards in residential buildings, some gases, particulate matter, and Volatile Organic Compounds (VOCs) are considered the most common. VOCs are organic chemical compounds whose composition makes it possible for them to evaporate under normal indoor atmospheric conditions of temperature and pressure. VOCs are normally emitted from sources such as construction materials, tap water, household products and human breath. Most VOCs are present indoors in low concentrations, often below the limit of detection of most sensors.

Taking measurement is difficult and so field studies often only report the concentration of a restricted number of contaminants and only for limited time. Meta–analysis research, which examine the results of independent studies, has been useful to show trends and the variability of the quality of indoor air in buildings. Dawson & McAlary (2009) studied background indoor air quality. They compiled and gave summarised statistics of 18 studies in North American dwellings reporting VOCs concentrations. The age of the studies ranged from 1970 to 2005. They excluded studies of smokers and personal exposures, and they did not differentiate between urban, suburban, or rural locations, or sampling method. The data showed trends in the composition and concentrations of pollutants over time. In general, newer studies reported lower concentrations than older studies, implying some improvements in air quality, and large variability between houses. Concentrations may be a function of the variation in ventilation rates, changes to, and the ageing of, building materials, and occupant preferences and habits. They identified commonly reported VOCs, such as BTEX compounds (Benzene, Toluene, Ethyl-benzene and Xylene) with concentrations similar to the USEPA health–based standard (NAAQS).

Wang *et al.* (2017) identified the concentration of some VOCs over five days of sampling from 25 houses; 19 diverse houses in London, and 9 houses of similar age and design in York, UK. They concluded that the two most abundant and common VOCs are *D*-limonene and α-pinene. Both pollutants are related to occupant activities and

consumer products, and are formaldehyde (HCHO) precursors. However, the study does not report ventilation rates, so its ability to control air quality could not be assessed. Huang *et al.* (2018) studied concentrations of HCHO, other VOCs, $PM_{2.5}$ and CO_2 in 21 houses in northeast China. They also investigated the effects of window opening, outdoor air temperature, furniture surface area, indoor temperature and indoor humidity on pollutant concentrations. Results showed a stronger relationship of HCHO with variations of temperature and relative humidity (plots (c) and (d) in Figure 2.3) than with window opening and furniture area (plots (a) and (b) in Figure 2.3).

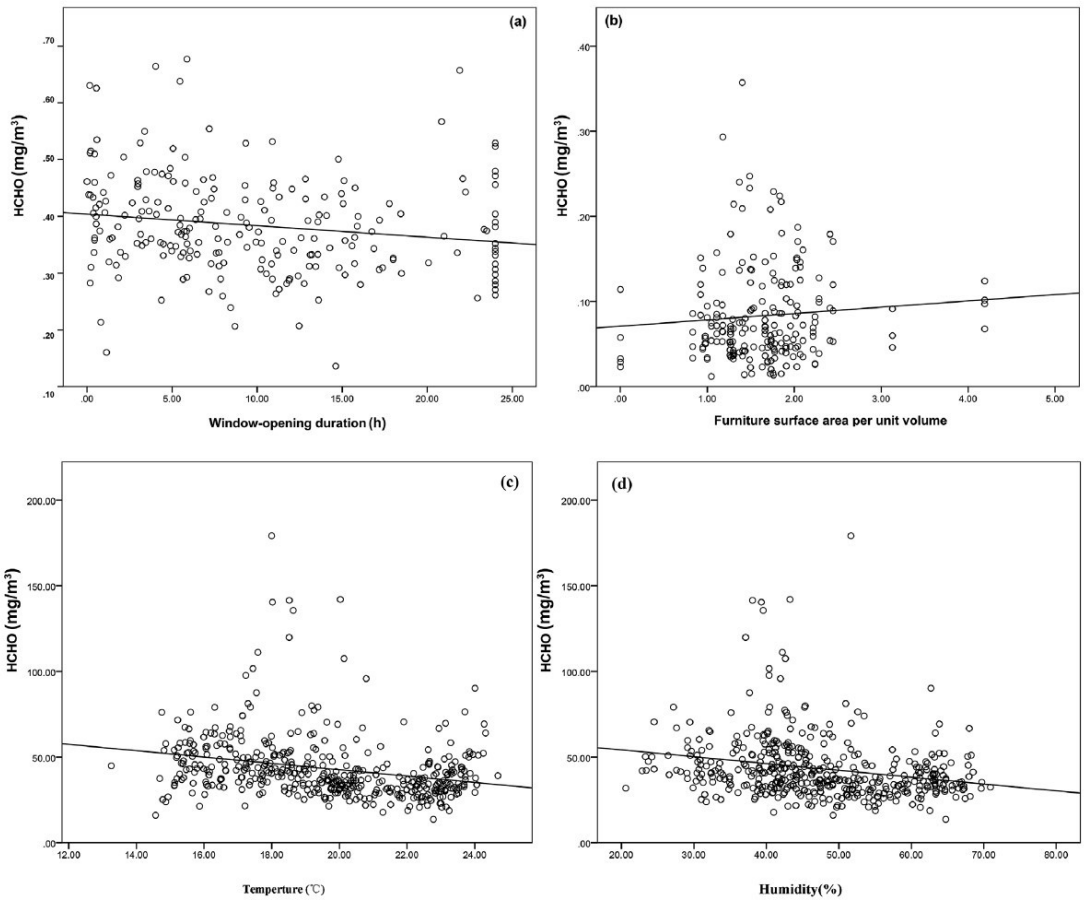


Figure 2.3: Effect of indoor parameters and HCHO concentrations (Huang *et al.* , 2018) in 6 houses in China.

Logue *et al.* (2011b) quantified concentrations of multiple indoor contaminants reported in field studies between 1995–2010 that were carried out in US dwellings and in other countries with similar lifestyles. Representative indoor concentrations for criteria pollutants were compared to the WHO and the USEPA standards (NAAQS). Both the $PM_{2.5}$ median and NO_2 $95th$ percentile were above one or both of these limits. The top three non-criteria hazardous volatile organic pollutants with the highest acute and chronic impact were acrolein, formaldehyde, and benzene. Indoor $PM_{2.5}$ is generally emitted by combustion of solid and liquid fuels, infiltrates from outside, or is

re–suspended during cleaning activities; acrolein is frequently released during cooking and smoking; and formaldehyde is emitted by smoking, released by composite–wood products or formed by chemical reaction in the indoor air. These data came from houses that in general were designed to have low emission events or sources.

Long *et al.* (2000) produced time–series data on particles concentration (see Figure 2.4) and size distributions (see Figure 2.5) in nine houses in Boston, USA, over more than 7 days of measurements during different seasons. They found that indoor $PM_{2.5}$ emission events tend to be brief (5 min on average), intermittent, and with highly variable peaks. Several sources of pollution were identified by linking time–activity data with measurements. Cooking activities were linked to the highest $PM_{2.5}$ concentrations. Figure 2.4 shows the $PM_{2.5}$ concentrations in one house over two days; the dotted line represents the two–day average concentration. The size distribution analysis, see Figure 2.5, showed that most of the indoor emissions (in volume) are within the $PM_{2.5}$ fraction, represented by the black, hatched and white portions of the stack bars. Factors that could explain the variability of the results include the type of source and their condition, house condition, and the proximity of the monitoring device to the source. Among house conditions, hourly air exchange rates were highly variable according to the season and ventilation conditions (ranging from 0.11 to 20.40 h^{-1}). Window and door opening were the main determinant factors of it, but the authors also highlighted the effect of the age of the buildings and the indoor–outdoor temperature difference, especially in winter. Low correlations were found between indoor and outdoor concentrations, noting the need to separate their analyses; indoor and outdoor exposures cannot be assumed to be equivalent. I/O ratios calculated by $PM_{2.5}$ component were dissimilar; typically above one for Organic Carbon *OC* and bellow one for elemental carbon *EC*. I/O ratios vary mainly according to window opening and season.

Fazli & Stephens (2018) summarised I/O ratios, indoor concentrations of $PM_{2.5}$, and concentrations of other pollutants reported in 21 studies carried out in The USA and Canada. Although their mean and medians are 1.07 and 0.99, respectively, they ranged from 0.73 to 1.7 (not accounting for the sample size). For indoor concentrations, the median of the reported means in these studies was 13.9 $\mu g/m^3$, ranging from 5.5 $\mu g/m^3$ in Saskatchewan (Canada) in winter, to 20.9 $\mu g/m^3$ in winter in New York City. Infiltration factors, F_{INF}, are reported in 15 buildings and range between 0.43 and 0.74 (Media = 0.58).

Jones *et al.* (2007) studied the spatial variability of particles inside of 78 Australian houses occupied by young children. PMs were monitored at two heights in the living rooms, bedrooms, and in cots. There was not a statistically significant difference between the two heights, suggesting that occupant activities helps to mix the air well

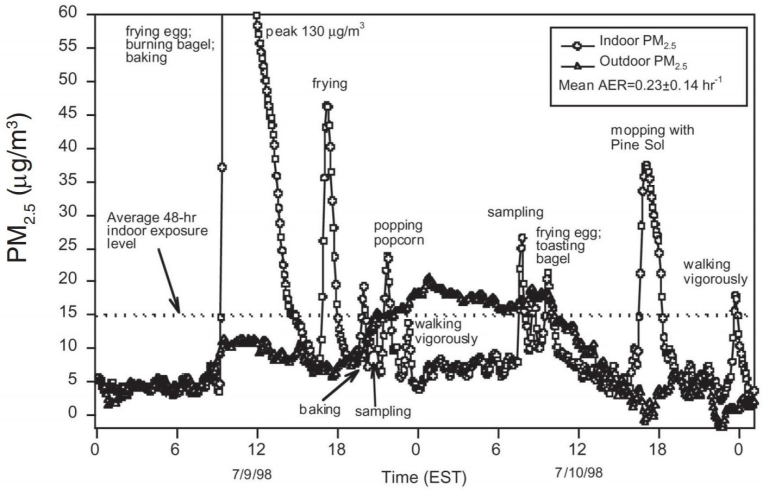


Figure 2.4: Inside and outside 20–min averaged $PM_{2.5}$ time–series data for one house monitored by Long *et al.* (2000).

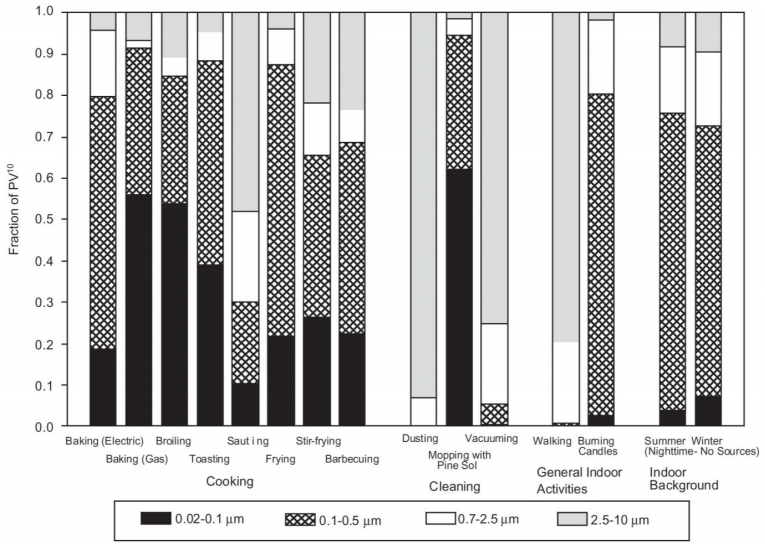


Figure 2.5: Volumetric particle size distributions. $PM_{2.5}$ accounted by black, hatched, and white portions, are higher in volume (Long *et al.* , 2000).

in each of the rooms. To the best of my knowledge, there are no comparisons of similarities and differences in Chilean, or other country's, houses or lifestyles, so a direct extrapolation cannot be made.

2.4.2 Chilean studies

Section 2.3 shows that exposures can be highly dependant on the cohort variables, such as their living spaces and activities carried out. For example, cooking and cleaning activities, two of the most influential sources of fine particles, are more related to adults than to children. Some studies on indoor air quality in the most populated cities of Chile have been carried out, mostly in low-socioeconomic and young population.

An example of this can be seen in Rojas-Bracho *et al.* (2002). They studied children personal exposures to PM_{10}, $PM_{2.5}$ and NO_2 in 18 houses in Santiago during five winter days in 1999 and compared them to indoor and outdoor concentrations. For this cohort (children between 10 and 12 years old), living in different locations in Santiago and different type of houses, outdoor and indoor exposures were strongly associated. NO_2 concentrations were higher than similar studies from other countries. For the $PM_{2.5}$ fraction, personal exposures were more strongly related to indoor concentrations than outdoors (See figure 2.6). The lower ratios between concentrations compared to other studies are likely to respond to the activities that the children carried out during the study; although no information on time spent indoors, activities and house characteristics were given.

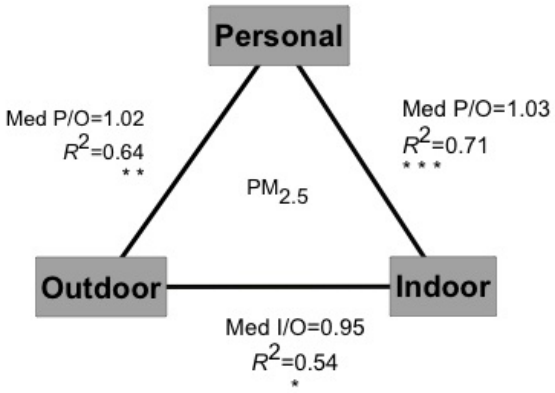


Figure 2.6: Relationship between personal exposures, indoor and outdoor concentrations for $PM_{2.5}$. Based on Rojas-Bracho *et al.* (2002). * indicates the level of statistical significance.

Cáceres *et al.* (2001), Ruiz *et al.* (2010) and Burgos *et al.* (2013) gave more information about social gaps related to indoor air quality in Santiago. Cáceres *et al.* (2001) measured pollutant concentrations, in 24 houses in a low socioeconomic area of Santiago, Chile. Most of cases had concentrations above the ambient limits, being the type of heater the main factor. Ruiz *et al.* (2010) measured $PM_{2.5}$, its components, and other pollutants in 16 houses and flats located in a middle–class area and in a mid–to–high area of Santiago in winter. Buildings had relatively the same construction period. Cases were selected by the type of gas heater used (compressed natural gas (CNG), liquefied petroleum gas (LPG), and kerosene), and were compared to houses using electric heaters or central heating, and to outdoor concentrations (see figure 2.8). Indoor concentrations were higher than the outdoor air in 60% of the cases, with kerosene heaters giving the highest concentrations for $PM_{2.5}$ (86.3 $\mu g/m^3$ on average), $PM_{2.5}$ components and other pollutants. Figure 2.7 shows 24 h indoor and outdoor $PM_{2.5}$

concentrations in a house using a kerosene heater, and Figure 2.8 shows means and standard deviations for the four types of heaters. Houses using electric heaters or central heating ("control" group) had lower $PM_{2.5}$ levels (42.1 $\mu g/m^3$ average), yet above the current WHO limit for ambient air. Two regression models were used to compare the effect of the outdoor air and the type of heater on the overall concentrations, and the type of house (either house or flat) and home activities (i.e. heating, cooking and cleaning) on the hourly $PM_{2.5}$ components concentration. Both cooking and cleaning were statistically significant at $p < 0.05$. Outdoor air was statistically significant for indoor concentrations in all of them at $p < 0.001$; $PM_{2.5}$ levels were high (55.9 $\mu g/m^3$ on average).

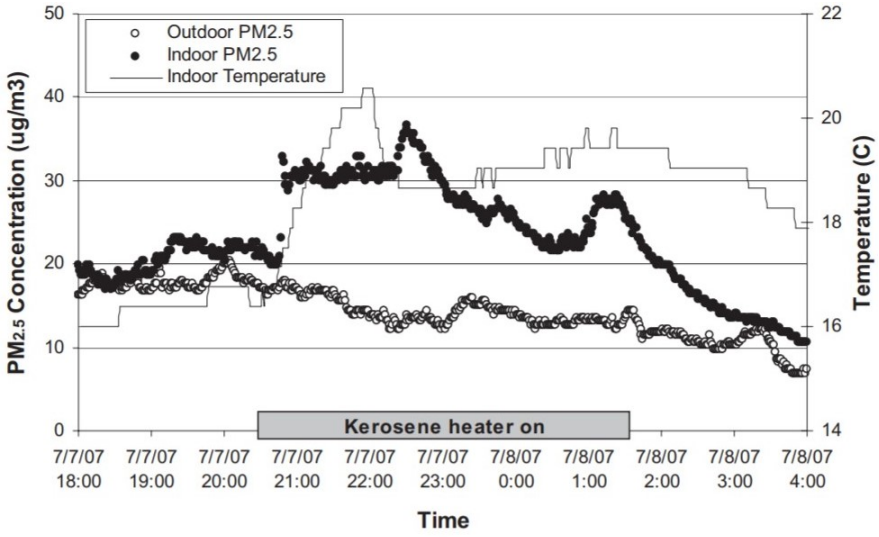


Figure 2.7: $PM_{2.5}$ concentrations for the outdoor and indoor air in a house using a kerosene heater (Ruiz *et al.* , 2010).

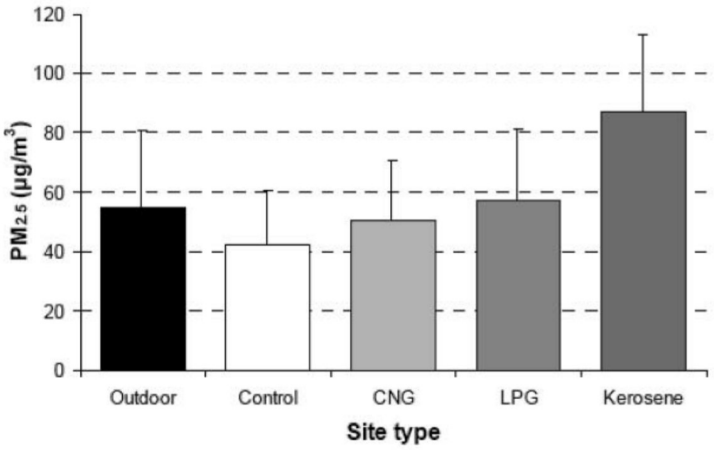


Figure 2.8: $PM_{2.5}$ concentrations for outdoor air, and for the four groups studied (houses using electric or central heater as the control group, and three different fuel types) (Ruiz *et al.* , 2010).

Burgos *et al.* (2013) compared the indoor air quality between two groups of houses with children <8 years; one group living in a slum and the other group in a social housing building. Indoor $PM_{2.5}$ concentrations in children bedrooms and sources were identified, along with information on house conditions and determinants. The median indoor concentrations for the whole group were higher than outdoors in both groups, and both above the WHO ambient recommendation. Slum families showed to be at greater risk. Slum houses had significantly higher indoor and outdoor concentrations than social houses, due to source differences. Outdoor air (R^2=0.26; p <0.00; regression coefficient β= 0.5), type of heating fuel (R^2=0.053; p <0.00), and indoor smoking (R^2=0.048; p= 0.01) were identified as the main factors. House type and condition were not evaluated. A limitation of this type of study is the time constraint of measurements (24 h). Other indoor activities, different lengths, or the influence of outdoor environmental conditions might affect the variability of the results. Finally, there is no information on how representative the houses were compared to other slums or social buildings of the country.

A similar study was carried out by Barría *et al.* (2016). They studied $PM_{2.5}$ concentrations in 207 houses of new–borns, in two cities of Chile, during 24 h in winter 2007 and 2008. Indoor levels were higher than in Santiago (median= 107.5 $\mu g/m^3$; range= 13.8 $\mu g/m^3$ (bedroom) – 373.9 $\mu g/m^3$ (living room)), but they did not measure or mention any outdoor concentrations for comparison (see Figure 2.9). Variations between and within groups were found. Statistically significant factors were: time of use and type of heater, family history of asthma, indoor smoking, cleaning activities (higher when using a broom than a wet cloth), ventilation time (higher when ventilating) and the monitored room (higher levels in the living room or in houses with only one room). Ventilation was related to an increase of $PM_{2.5}$ concentrations (p = 0.0039), which may suggest greater contribution of outdoor pollution, total deposition rates or physicochemical reactions.

A few studies with focus on $PM_{2.5}$ particles have been carried out in Chile. Most of the reviewed studies were carried out before the implementation of the decontamination plans and before the $PM_{2.5}$ regulation came into force. Since then, outdoor pollutant concentrations have decreased in most cases (see Section 1b Ambient pollution), which led to a reduction of its contribution to the indoor pollution. The relative importance of the indoor sources is then becoming even more relevant. The following research were conducted specifically on $PM_{2.5}$ particles within Chilean residences from year 2000 onwards.

Jorquera *et al.* (2018) measured and characterised the indoor and outdoor $PM_{2.5}$ of 64 households in a highly polluted city in Chile, Temuco, during 24 h in winter season. They estimated that outdoor pollution contributed to indoor pollution by 68% – a

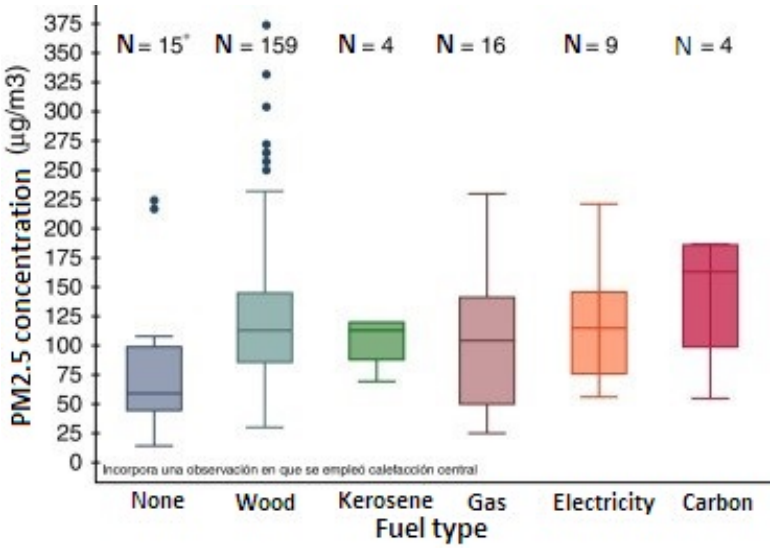


Figure 2.9: $PM_{2.5}$ concentrations for houses using four different fuel types (Barría *et al.* , 2016).

strong effect size Ferguson (2009)– giving indoor and outdoor median concentrations of 44.4 and 41.8 $\mu g/m^3$, respectively. They exceeded the WHO annual recommendation of 10 $\mu g/m^3$. The analysis also shed light on the household characteristics and design factors that most affect the indoor air quality in winter season: the type of cooker and heater, whether the house was insulated or not, the presence of windows in the tested room, and floor area, showed to have a statistically significant effect on $PM_{2.5}$ concentrations. $PM_{2.5}$ concentrations were found to be higher indoor than outdoor during 7 hours of the day, which are likely to be the hours with maximum occupancy and probably when cooking, cleaning and heating activities took place.

Bravo-Linares *et al.* (2016) measured $PM_{2.5}$ concentrations in another polluted city, Valdivia. They identified the main sources and indoor/outdoor ratios for 12 houses in two consecutive years during winter and autumn season. Daily mean of ambient concentrations (range = 5 – 367 $\mu g/m^3$; mean= 85 $\mu g/m^3$) were higher than indoors (range = 6 – 194 $\mu g/m^3$; mean = 72 $\mu g/m^3$) for the majority of the houses, except for the ones located in rural or semirural areas. Figure 2.10 shows the variability and relationship of the indoor and outdoor concentrations. The outdoor concentrations were measured at the sites using the same portable instrument used indoors.

A regression analysis (R^2) showed that the outdoor concentrations accounted for 48% of the variation in the indoor concentrations, which is a *moderate* effect size (Ferguson, 2009). Data on the left hand side of the 1:1 agreement line indicate indoor concentrations that are higher than those outdoors. In all cases, the mean of the indoor concentrations exceeded the WHO daily threshold for ambient air, and some of them also exceeded the Chilean national daily mean threshold of 50 $\mu g/m^3$ daily mean (WHO *et al.* , 2006). Interestingly, the indoor concentrations were above the national

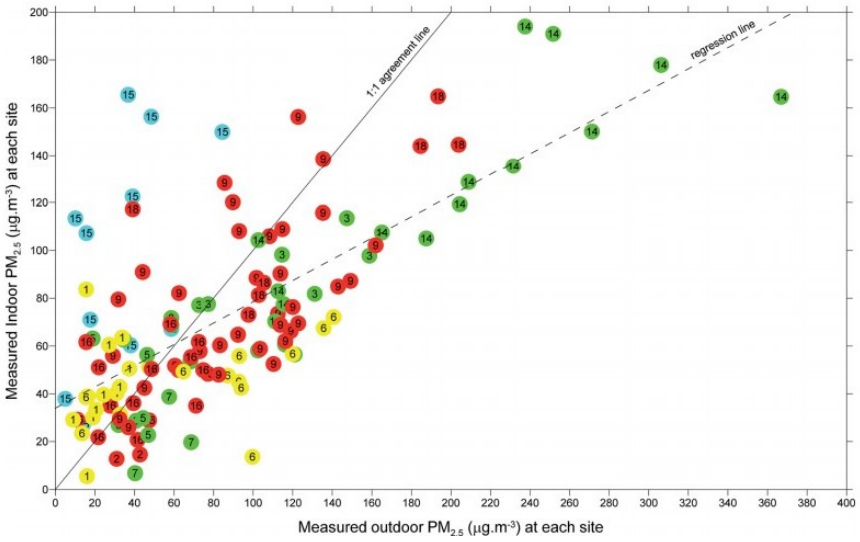


Figure 2.10: $PM_{2.5}$ concentrations measured inside and outside the houses (N = 135). (Bravo-Linares *et al.* , 2016). The 1:1 agreement line corresponds to a I/O ratio of 1

threshold for ambient air more often (∼70% of the time) than the outdoor (60%). Additionally, outdoor concentrations measured outside each dwelling were compared with values reported by the government monitoring station. A weak agreement was found between the two values; see figure 2.11. Accordingly, outdoor concentrations of houses located closer to the monitoring station showed a better agreement than houses at more distant locations. And, some association with housing density was found and so the use of the outdoor data to make assumptions about community exposures must be used with care. The monitor was placed in the room the occupants reported that they used the most, normally the living room or dining room. However, an analysis of occupant exposures to pollutants in their house depends on their concentrations in the other inhabited rooms as well.

In both Temuco and Valdivia, the outdoor air, cooking, smoking, and space heating, were considered to have the greatest impact on indoor air pollution. Jorquera *et al.* (2018) also identified cooking and heating as the two main indoor sources of $PM_{2.5}$, with some emission rates estimated to be up to about 100 $\mu g/m^3$ during the evening.

This section emphasises the need for more information about indoor air quality, which can differ enormously according to the location of the building, the building design, occupant activities, and the type of indoor sources. For example, the relative importance of indoor sources compared to outdoor sources are expected to vary in less polluted or better ventilated locations, or in houses with lower emission rates. Further work is required to investigate these areas of uncertainty at a larger scale, while also accounting for the local characteristics of houses.

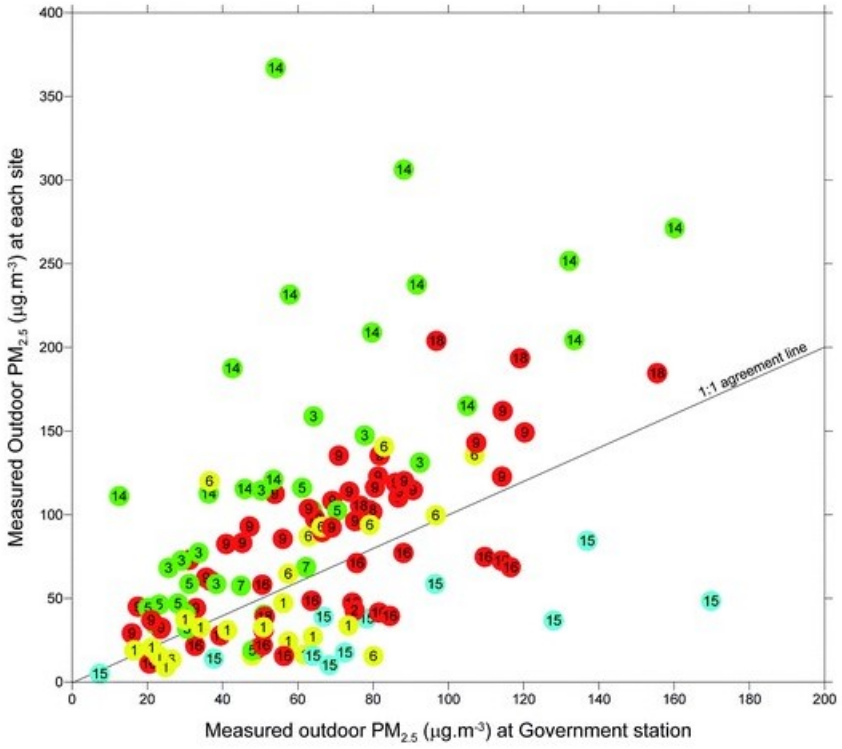


Figure 2.11: $PM_{2.5}$ concentrations measured outside of the house and in the nearest monitoring station (Bravo-Linares *et al.* , 2016).

2.4.3 Health effects of exposure to indoor pollutants

Section 2.4 shows that fundamental dissimilarities may exist between pollutants found outside and inside. Consequently, the health impacts of exposures to ambient pollutants may be different from those to pollutants generated inside. Figure 2.12 shows some health impacts of indoor and outdoor exposures to ambient $PM_{2.5}$ and PM_{10} using time–activity data, field measurements, and infiltration factors from The USA, Europe, and China. It shows that indoor exposures (in black) contribute the most to the increased mortality risk. Here, the time spent outdoors was a more influential parameter than the infiltration factor (Ji & Zhao, 2015).

Logue *et al.* (2011a) ranked the pollutants included in Logue *et al.* (2011b) according to their potential chronic health risks. To calculate the contribution of indoor exposures to the total exposure, a scaling factor of 0.7 was applied to the indoor concentration because people in the US spend around 70% of their day in their houses (Klepeis *et al.* , 2001). Here, concentration–response functions were used instead of dose–effect curves, and so they only report the risk to which the occupants are likely to be exposed. Confidence intervals of population impacts of exposures were reported; see Figure 2.13.

The impacts of both household features, such as those included in 2.4.1 and others,

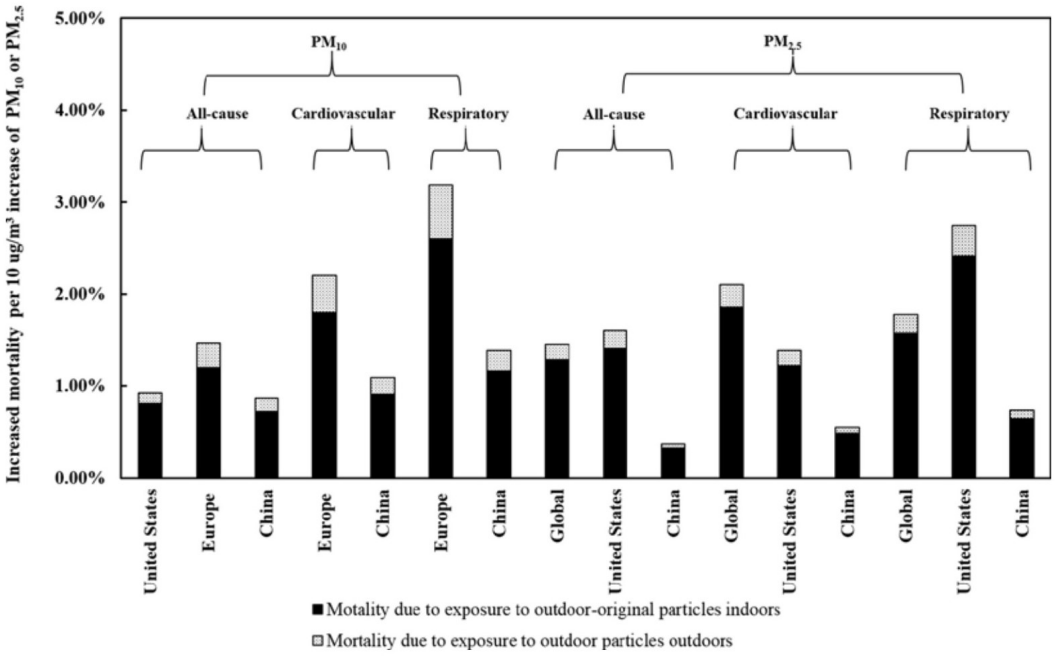


Figure 2.12: Increased mortality risk due to exposures to ambient particles in both indoor (black portion) and outdoor (grey portion) (Ji & Zhao, 2015).

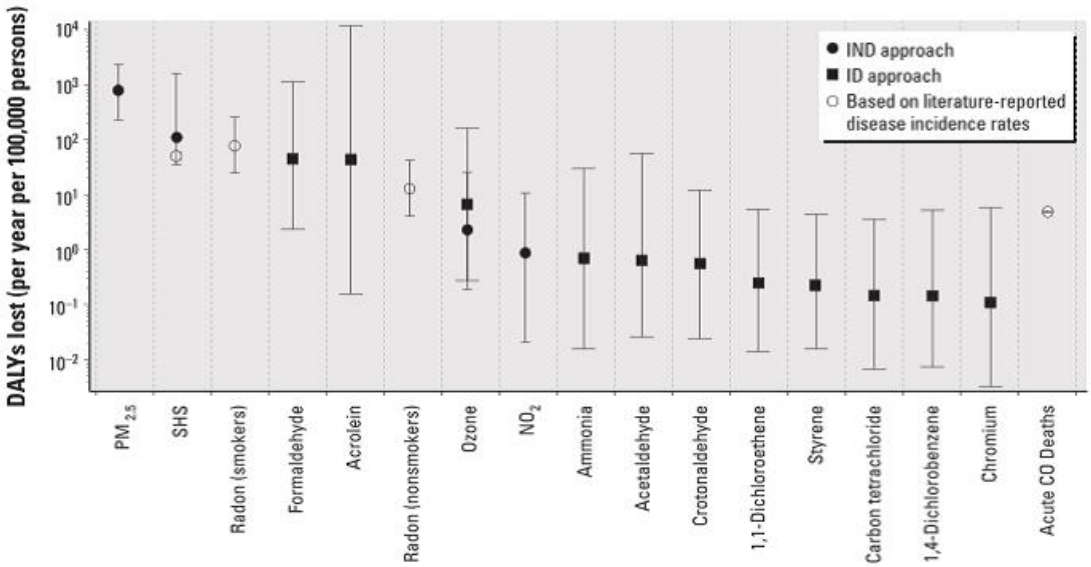


Figure 2.13: The impact of indoor pollutants on occupants' health measured in DALYs (Logue *et al.* , 2011a).

and interventions on occupant exposures and on occupant health remain unclear. The European project HEALTHVENT estimated the potential health benefits of control policies for decreasing indoor exposures to six pollutants and home dampness in 26 European countries. They compared health effects, measured in disability–adjusted life years (DALYs), between the current scenario in 2010, and three potential scenarios: (i) a change in ventilation rates from 0.1 to $5\,l\,s^{-1}$ per person, (ii) the use of mechanical ventilation with filtration, and (iii) source control with a minimum ventilation rate.

The results showed that the first scenario (dilution) would reduce the burden of disease by 20%, the second scenario (adding filtration) by 38%, and source control by 44%. The reductions largely depended on the location of the houses because of variable ambient pollution concentrations, and the type of source reduction. The minimum health impact was achieved with $4.4\,l\,s^{-1}$ per person, which is low when compared with the base scenario that estimated the ventilation rate for the housing stock to be $17\,l\,s^{-1}$ per person (Asikainen *et al.* , 2016), and higher than ASHRAE 62.2 guideline recommendation of $3.5\,l\,s^{-1}$ per person (ASHRAE 62.2, 2016).

Singleton *et al.* (2017) studied the relationship between reported incidences of asthma in children, indoor pollutant concentrations (over four consecutive days) and housing characteristics in 63 houses in Alaska, USA. Relevant indoor parameters were identified as increasing the risk of asthma. Among them, the occupancy level (used to indicate overcrowding), reduced ventilation, indoor sources of pollution, and occupant behaviour were the most important. Chin *et al.* (2014) also studied the houses of children with asthma symptoms in Michigan, USA. The concentrations of some VOCs were measured for 7 day periods in 126 houses during at least 2 or 3 seasons. Results were consistent with other studies carried out in the country and with Logue *et al.* (2011b) meta–analysis. Figure 2.14 shows a variance component analysis. Nearly half of the variance in the measured VOCs concentrations is explained by house parameters (49$\pm$11%), followed by seasonal variation (34$\pm$12%), room differences (10$\pm$6), and replicated measurements (7$\pm$6%). Indoor temperature, the degree of mixing of the indoor air within the house, the location of the VOC sources, and the air change rates were also considered in a correlation analysis. The three most important sources identified were solvents used in building materials, vehicle emissions, and household products.

Sundell *et al.* (2011) reviewed the literature looking for associations between ventilation rates and health effects in different building types. In this review, odour perception was excluded, because it was argued that this variable was already included in standard requirements and "it does not directly represent a heath outcome". Papers considering serious health impacts were excluded. The panel noted that some reduced health issues, such as SBS–related symptoms, risk of allergies and asthma

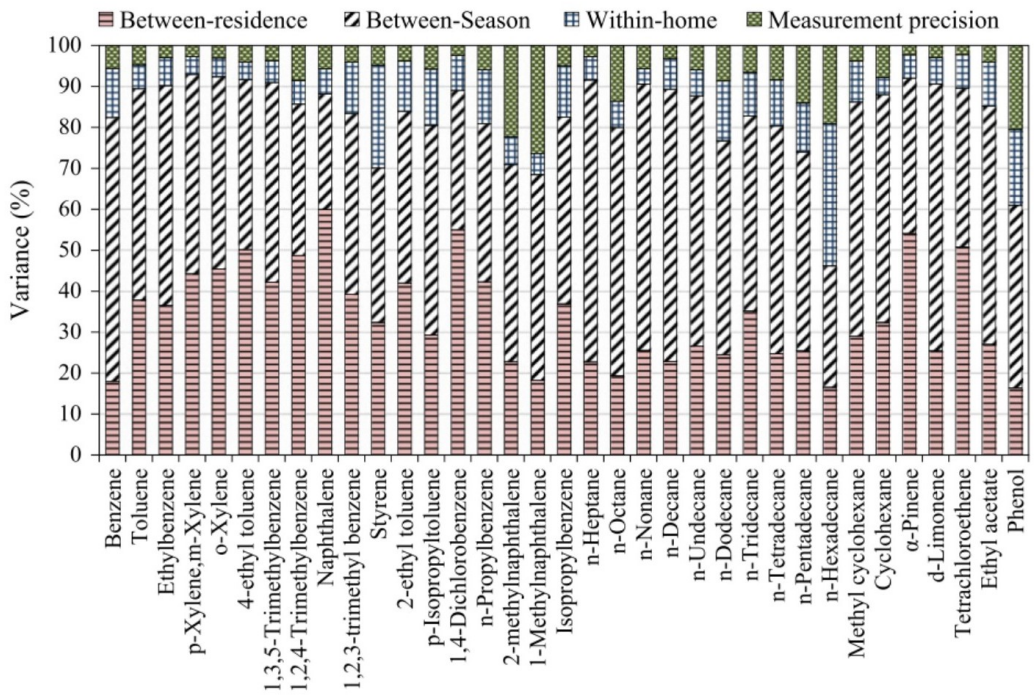


Figure 2.14: Variance in VOCs concentrations measured in Michigan houses (Chin *et al.* , 2014).

symptoms, especially in children, were found in offices with higher ventilation rates (up to $25\,l\,s^{-1}$ per person) and in houses with ventilation rates above $0.5\,h^{-1}$. The reviewed studies did not give pollutant concentrations, some only reported CO_2 concentrations, which was used as an indicator of ventilation rates or associated with other compounds. The review only included four studies on indoor air quality in residencies, and only from Scandinavian countries. One key conclusion was the lack of information on outdoor air quality in the specific locations of the buildings and on the indoor pollutants sources. They do not give a minimum ventilation rate or a threshold, only giving recommendations to increase ventilation rates to improving the indoor environmental quality and, consequently, human health.

In Chile, health outcomes from exposure to ambient and household air pollution were estimated to be on the top ten leading causes of death and of premature deaths (IHME, 2017): ischemic heart disease (1st), stroke (2nd), lung cancer (7th), three non–communicable diseases, and lower respiratory infections (12th) a communicable disease. Moreover, air pollution is the 8th risk factor for most of those deaths and disabilities combined. Siddharthan *et al.* (2018) analysed data on 13 low and mid–income countries (according to the World Bank classification). Cities were chosen for their use of biomass materials as a primary fuel source. In Chile, the data came from The Pulmonary Risk in South America (PRISA) study. Results showed a strong association between exposure to biomass fuel smoke by adults (a household air pollutant) and

COPD; see Figure 2.15. In Chile, the prevalence was higher for women than men, in contrast to the general tendency of the other 12 countries. No dose–response relationship was given.

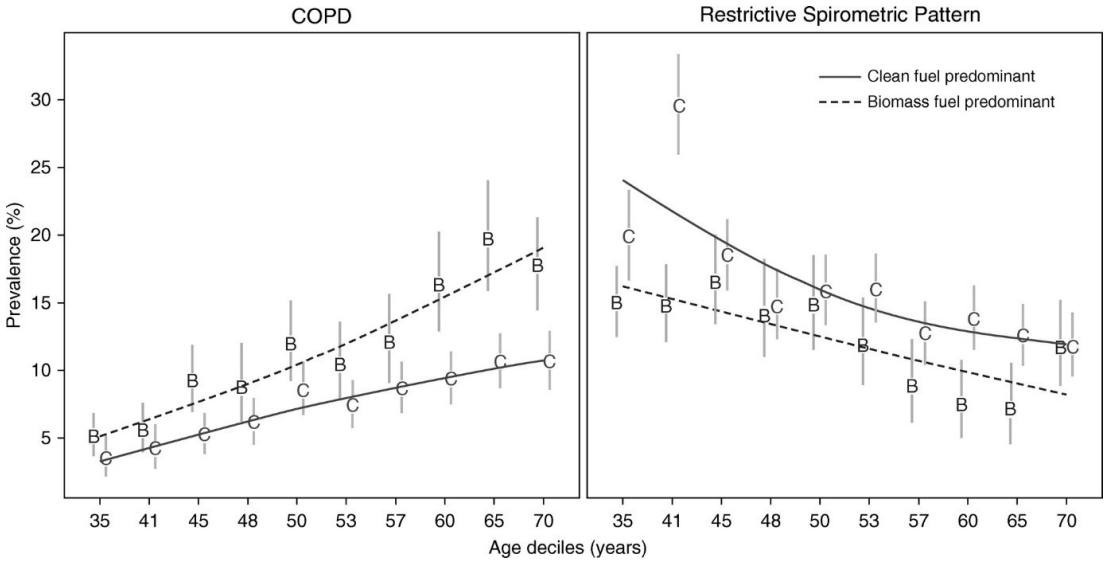


Figure 2.15: Prevalence of COPD and restrictive spirometric patterns among adults of a rural location in Chile (Siddharthan *et al.* , 2018).

.

Rivas *et al.* (2008) surveyed 355 mothers from low–income families, with children under 5 years old to identify associations between occupant habits (factors) and the prevalence of respiratory diseases (episodes). Factors that showed a significant difference in the absolute frequency of respiratory episodes were drying laundry in the kitchen (p=0.04), fuelwood use (p 0.05), and cleaning with wet cloth (p 0.79).

Chilean data on indoor exposures to household air pollutants is lacking and may be skewed towards high–risk communities so the burden of disease estimates related to this risk could also be segregated. The Prospective Urban Rural Epidemiology (PURE) study is an ongoing multi–country cohort study that aims to increase information on $PM_{2.5}$ exposure–response relationship based on varying cooking fuel types and conditions for more accurate estimates of global health impacts (Arku *et al.* , 2018). Within the 18 participant countries, they selected 10 of them reporting >10% of households using solid fuels for cooking and heating in a rural location. These findings may help us to understand Chilean sources of emission and household characteristics.

Because the composition and toxicity of particles found in indoor and outdoor air could differ, separate toxicological studies are needed for comparison and to establish dose–response relationships. Furthermore, more information about the variability of contaminants found in indoor air across the housing stock is required. An economic and quicker way to address this lack of information is through using a modelling ap-

proach. This method is capable of predicting indoor concentrations accounting for information on pollutants, state–of-the-art building technologies, and occupant factors. Nevertheless, models are more reliable and accurate when that information is regularly updated and is applicable to the community being studied. The following two sections present modelling approaches for predicting indoor concentrations for single or multi–zone buildings (Section 2.5), and for a larger proportion of the stock (Section 2.6).

2.5 Modelling indoor air

Two approaches can be used to assess personal exposures, directly or indirectly. Direct, or *in–situ*, methods use personal monitoring or biomonitoring to obtain data on actual exposures to air pollution, either by using a personal monitor or biomarker samples; see Section 2.3. In contrast, indirect methods use available monitoring data and time–activity patterns to model the concentrations within a volume and predict personal exposures. Direct methods are more accurate in reflecting individual exposures and the studies can be simpler to design. They give a better description of the magnitude and duration of events, and can give information on source emission rates. Indirect methods allow individual and population exposures to be modelled, and multiple exposure scenarios to be evaluated (Branco *et al.* , 2014).

However, direct methods are constrained by time and costs. Monitoring a large sample requires a good description of the input variables and can be impractical, and so they are limited for a specific cohort, or representative sample, for short periods. Moreover, there are still high uncertainties in the measurements and devices that could mislead the readings (Jones *et al.* , 2018; O'Leary *et al.* , 2019a; Singer & Delp, 2018). The detection of individual sources of pollution and parameters under real conditions requires complex analyses and identification of relevant pollutants from each source. Conversely, indirect approaches rely on assumptions to simplify any model, and must consider uncertainty in input parameters and model accuracy. Depending on the method, assumptions can oversimplify temporal and spatial patterns, and exposure distributions (such as statistical regression models) when predictions cannot be extrapolated to the general population scale (such as micro–environmental modelling or computational fluid dynamics (CFD) tools), so they need extra validation (Graham & McCurdy, 2004).

Several approaches have been used to address these issues. For example, both methods can complement one another so that the model inputs are better informed by the measured data, and results can be validated and recalibrated. The accuracy of the model will then depend on the availability and quality of the input data, and

the ability of the model to predict indoor air pollutant concentrations. Another option is to use representative scenarios, such as *representative buildings*, so that the results can then be extrapolated to wider groups. When considering the degree to which the indoor airspace is partitioned to give a required level of spatial detail, the modelling approach can be classified into three methods. Going from the use of simple models with (i) a single and well-mixed zone, continuing with (ii) multi–zone models, and (iii) for more complex models, using a non–uniform distribution of the pollutants with Computational Fluid Dynamics (CFD) simulation tools.

The definition of the method will define the level of complexity, resolution and detail to represent the indoor air in a building. The three methods all have their own advantages and disadvantages. Well–mixed single–zone models consider just one volume of air, normally with an homogeneous temperature, air pressure, and contaminant concentration. Multi–zone models consider more than one zone (or *nodes*), which are connected by airflow paths. Using these two approaches, the average pollutant concentrations of a whole building, or each zone, are predicted over time. Air flows between each zone and the outdoor air are calculated using mass balance equations, which are solved iteratively until the mass of air through the building is conserved at each time-step. Finally, a CFD model can be used to analyse the distribution of a pollutant within a room and identify flow directions, giving a greater level of spatial detail; see Figure 2.16. Nevertheless, the time–cost of this type of simulation is higher, the modelling is more complex and time-consuming, and it requires more computational resources (Yang *et al.* , 2014).

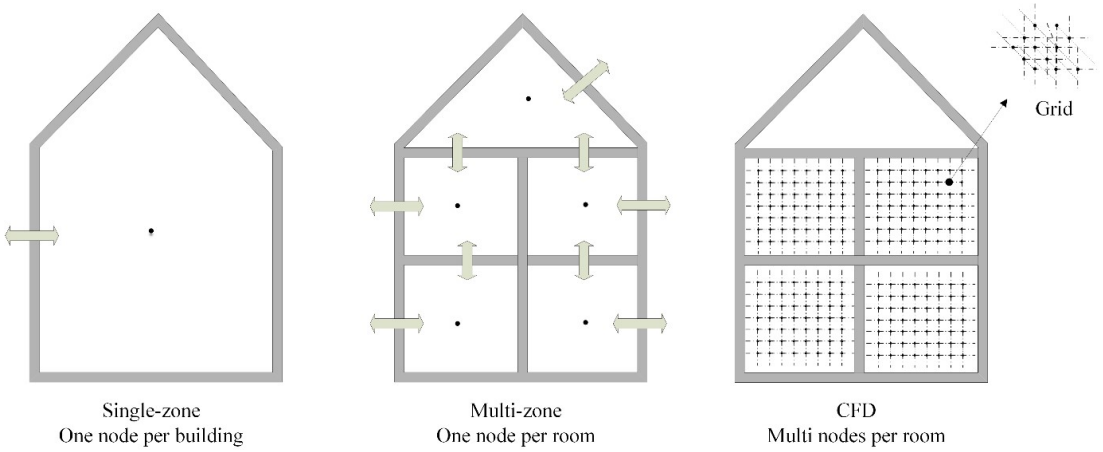


Figure 2.16: Three different approaches to model the IAQ in a building. Left: Single well-mixed models; Middle: multi–zone well-mixed models; Right: CFD model. Each dot or node represents a well-mixed volume. Based on Axley (2007).

A single source can emit multiple pollutants (see S_1 of Figure 2.17) and a pollutant can be associated to multiple sources (see P_2 of Figure 2.17).

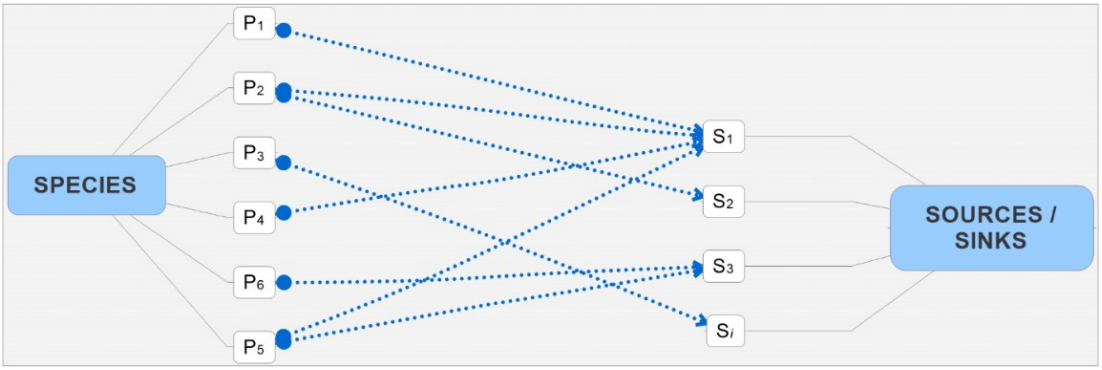


Figure 2.17: Source and species composition. Indoor sources, *S*, can account for one or more pollutants emissions, *P*, also known as *species*.

An example of a simulation program capable of all this is CONTAM, a freely available multi–zone ventilation and pollutant transport model. The program is developed by the National Institute of Standards and Technology (NIST) (Dols & Polidoro, 2015) and has been validated by several studies of different building types and locations (Ng *et al.* , 2012), and pollutants (Bastani *et al.* , 2012; Underhill *et al.* , 2018; Yu *et al.* , 2015). A recent example of the use of CONTAM is provided by Underhill *et al.* (2018). They modelled a low–income multi–family building using CONTAM to analyse the impacts of retrofitting interventions on energy demand and indoor $PM_{2.5}$ and NO_2 concentrations. They designed multiple scenarios, using factorial design, by combining different stages that the retrofit could undertake, different occupancy and indoor activities, and seasons. Building elements were modelled using ASHRAE air leakage data (ASHRAE 62.2, 2001). Pollutant sources included a gas stove, cooking activities, infiltration, and indoor smoking for some scenarios. No heaters were included in the models. Using this approach they found that the weatherisation programs alone do not necessarily improve the quality of the indoor air, and when other elements are added, such as ventilation improvements or filtration, occupant activities play an important role in the final outcome. Thus, to better understand the possible indoor air quality in any occupied building, human behaviour must to be included in the models.

2.6 Modelling the housing stock

An understanding and evaluation of building performance at a large scale needs good information and the definition of target indicators. The constant evolution in the number, composition, and characteristics of a building stock over time requires models that allow the stock to be outlined and assessed according to certain indicators, and the benefits, drawbacks and trade–offs of potential interventions to be tested before

their implementation. This section presents the housing stock as an object of research, the methods that are currently used to evaluate them, and some examples and applications.

Modelling a housing stock has many advantages. A robust model should be able to estimate the baseline performance accurately so that it can be used to explore other technical and economic outcomes, and to inform policymakers, the market, and academic researchers. It can also be used to estimate the impact of future climate change on a stock of buildings and on occupants' health and comfort.

Housing stock models (HSM) are generally categorised as top–down or bottom–up, working at aggregated and disaggregated levels, respectively. Both methods can be sub–categorised by their application of statistical (e.g. regression techniques) or physical modelling methods (Kavgic *et al.* , 2010). For example, top–down models might predict the changes in the energy demand or pollutant concentrations of houses from changes in a number of factors, such as variations in energy price or weather conditions, but they cannot explain those changes in detail. They are normally carried out by fitting historical or time–series data at an aggregated level but are limited to the given variables. New drivers, such as new technologies, require a new model. Conversely, a bottom–up model might use empirical data from surveys, field measurements and, in the absence of data, assumptions to describe each component of a dwelling using a physical model. Thus, the model could be used to assess the impacts of a new intervention policy on a metric from a change in the conditions of a component, such as a U-value, or in a set of components by comparing building type A to B.

Different techniques have been used to simplify the representation of stocks. The UK has a 25 year history of building models to support policy development and so it is used for comparison herein. Its models of dwelling energy and airflow systems are summarised by Sousa *et al.* (2017). The majority of UK Housing stock energy models (HSEMs) are bottom–up simplified steady–state models of physical phenomena. Many suffer from over–simplification, a lack of model and data transparency, and poor modularity Sousa *et al.* (2017). However, other recent models have applied advanced sampling methods that generate distributions of predictions and quantify the uncertainty in both inputs and outputs, and use global sensitivity analyses to identify key drivers that can be targeted for remediation or by future field surveys Das *et al.* (2014); Jones *et al.* (2015).

Stock models have shown to be useful and informative, but uncertainties still exist in their development process and the information used to inform the models, which may limit their usefulness. Booth *et al.* (2012) identified three sources of uncertainties in HSEMs and some options for handling them. The first relates to heterogeneity

within the building population. The second is due to random variability, or first–order uncertainty, meaning that different outcomes are possible to obtain within the same building. The third is due to parameter, or *epistemic*, uncertainty over their values, or second–order uncertainty. A wide range of inputs are required to accurately model heat and mass transfer in houses. However, this information may not exist or may only exist at an aggregated level. Furthermore, the collation and processing of data can be time consuming and can introduce systematic errors when they are not tracked.

In order to obtain a proper representation of a stock, and to limit the first–order uncertainty, normally a large sample of buildings is required. This can involve an intense use of time and computational resources to collect and process the data. In order to address these issues, some countries have developed and used a set of archetypal buildings to represent their stock, which in turn requires and depends on an extended database (Mata *et al.* , 2014) of information on that stock.

Relevant parameters that describe the performance of a building or building stock can be gathered either by *in–situ* measurements or by using building physics models to predict it. Modelling an individual building, or a housing stock, is a quicker and more economic method of assessment when compared to *in–situ* measurements. Furthermore, modelling a set of individual buildings representing a proportion of the stock and applying a probabilistic distributions is preferred due to its capabilities in assessing multiple scenarios, re–diversifying the results, and so handling second–order uncertainty. Finally, an up–date or calibration process can be carried out to minimise the error between the expected and real values by using the model results along with newly available data or *in–situ* measurements (Balaras *et al.* , 2016; Booth *et al.* , 2012; Cerezo *et al.* , 2017; Sokol *et al.* , 2017).

Many countries have applied bottom–up techniques, based on building physics modelling, to develop their national representative buildings, also called *archetypes* or *building typologies*, and have used them to predict energy performance and to calibrate the model; see Persily *et al.* (2006), Sokol *et al.* (2017) and Ghiassi & Mahdavi (2017) for example. Available data is used to classify the stock into groups sharing the same or similar attributes. Data entries sharing similar or equal category are grouped or *clustered*, and are used to describe the archetypes. After the classification, the attributes of each group are used to define the archetypes. Each archetype then represents a proportion of the stock, and so the larger the number of archetypes used, the more representative of the stock they are and the more widespread are the conclusions derived from the results. The outputs for each archetype are then extrapolated to a whole stock of dwellings.

2.6.1 The use of representative buildings in the literature

Archetypes have been widely used, for instance, to model the energy demand of existing buildings, as tools to assist during the transition to a more energy efficient stock, for Life Cycle Assessment (LCA), and for indoor air quality assessment. One example is the TABULA project whose aim was to create a set of reference buildings for 21 European countries in order to study the energy demand of their stocks and to assess potential energy savings. Three methods were available to researchers: the development of a real example by a group of experts; the selection of an averaged building using statistical analysis; or the creation of a set of archetypes (Ballarini *et al.* , 2014). The number of buildings required to represent each stock ranged from 11 for Hungary to 62 for Germany.

Novikova *et al.* (2018) present the archetypes for Albania ($N = 20$), Montenegro ($N = 15$) and Serbia ($N = 24$) developed by The TABULA project. The performance of these buildings was assessed by modelling a current scenario, and then moderate and advanced low–carbon high–thermal–comfort scenarios. Several assumptions about the evolution of the existing and future housing stock were made to estimate the energy demand by end–use and rank priorities for stakeholders at a national level.

Filogamo *et al.* (2014) present the development of 12 archetypes to represent the Sicilian housing stock. The archetypes were then used as an evaluation tool for energy planning. The characterisation process involved (i) a statistical analysis of the stock by relevant variables, (ii) classification of the stock into clusters by geometrical properties, (iii) a definition of the archetypes' properties, (iv) a further subdivision of the clusters by other variables of interest, and (v) an allocation of the archetypes by defined locations. After the stock was characterised, the energy demand of the archetypes was calculated and extrapolated to the whole stock by the frequency of occurrence of each archetype. Finally, the results were compared against the actual energy demand during a specific year. The difference was around 8%, showing that the method is a reliable tool for predicting energy use or for evaluating different energy saving scenarios in a large stock.

Mata *et al.* (2014) represented the French, German, Spanish and the UK buildings stocks (residential and non–residential) for similar purposes. The results are similar to those reported by Filogamo *et al.* (2014). Differences between the predicted energy demand of the whole stocks, extrapolated from the archetypes, and nationally available statistics, were between -6% to +2%. The study also included a sensitivity analysis of the house-related input variables that most affected the outcome, which were ranked by their relative importance. The minimum indoor air temperature, *U*-value, mechanical ventilation rate, and envelope surface area were ranked among the

top four.

Oikonomou *et al.* (2012) studied overheating of the London housing stock using 15 representative buildings representing 76% of the stock at two locations within the city. They modelled indoor temperatures under different scenarios and conditions. The results showed the effect of the urban heat island on the stock (ranging between 1°C and 4.4°C) and that the most important determining factors are those related to the fabric design rather than to the location of the building.

Persily *et al.* (2006) developed a set of 209 houses that represent 80.2% of the US housing stock. National databases on residential buildings were analysed statistically and used to describe the stock using a number of variables, such as location, range of floor area, household size, year built, and their categories. Using categorical data clustering, 848 archetypes were formed by combining one category from each variable. This is also called *factorial design*. Then, each building was classified into one archetype, thus forming a cluster. The weighting factor of each archetype corresponds to the number of buildings fraction of the stock that it represents and those with the highest number are the most common. Finally, 209 cells or archetypes were retained that accounted for 80% of the stock.

The 209 US archetypes were then modelled by Persily *et al.* (2010) in `CONTAM` to predict a distribution of infiltration rates for the entire US housing stock; see Figure 2.18. Nineteen cities were chosen to represent the US climates and reflect its variability across the country. The archetypes' weighting factors were used to account for the total representation of the stock. Modelled distributions generally agreed well with previously measured data, especially studies with larger samples and larger representation. Some deviations were found in studies that included window opening or different sampling criteria. This comparison shows the importance of having a good representative sample and a wide coverage of scenarios for both building types and environmental conditions, and the usefulness of the study for understanding building behaviour.

Fazli & Stephens (2018) also used the US archetypes to predict the annual energy demand attributable to HVAC operation, and indoor concentrations of several pollutants; see Figures 2.19 and 2.20. The archetypes were modified to fit a single–zone fully–mixed model in Building Energy Optimization Tool (*BEopt*) software and then co–simulated using `EnergyPlus`. Predicted concentrations agreed relatively well with historical data and were within the measured ranges reported in other studies of US houses. The total energy consumption estimated by the model was only 2% lower than the annual national estimates for 2009, although they ranged between -65% and 6% depending on the type of fuel. This discrepancy reflects one of the limitations of the study, which was that the fuel was chosen using the most common type

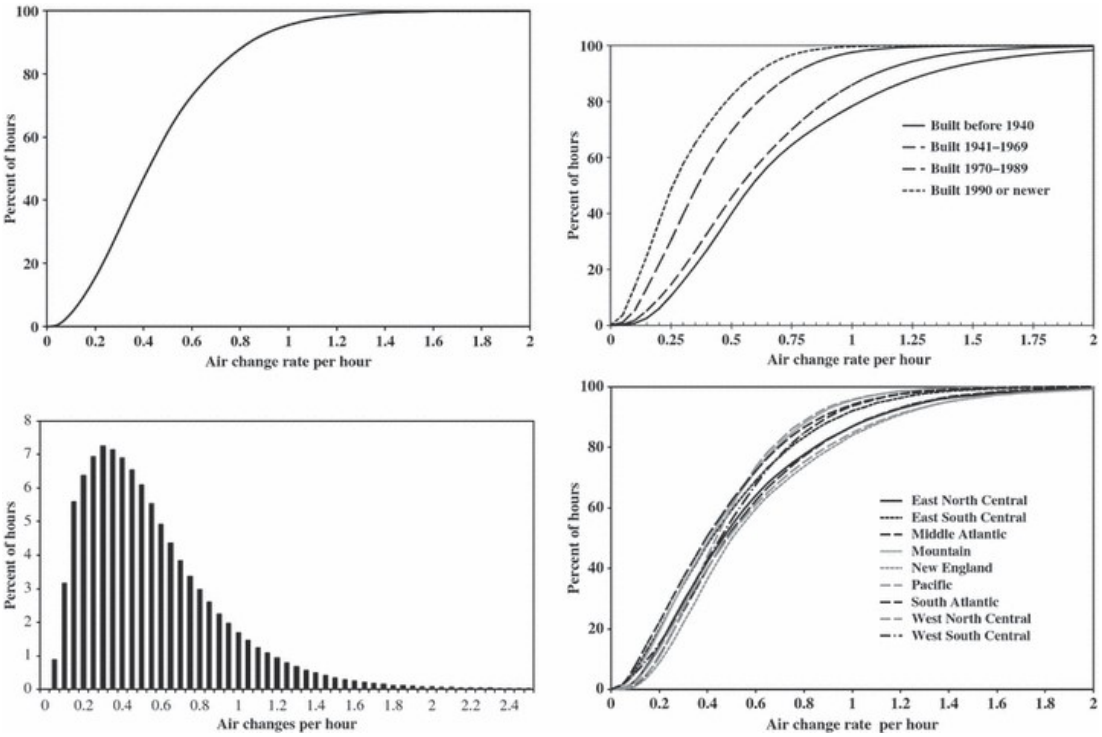


Figure 2.18: Predicted distribution of averaged annual air changes per hour in detached US houses attributable to air infiltration and air handling systems (Persily *et al.* , 2010)

used in each region, and did not account for its variability. The same criteria was used for some emission rates and activity schedules. Nevertheless, the authors acknowledge these limitations arguing that the decisions were made to reduce computational time. Discrete and constant whole–house emission rates, normalised by house volume rather than by source, were used for some aldehydes and VOCs. For $PM_{2.5}$, NO_2, and ultra–fine particles (UFP), only fixed transient emissions from cook stoves and cooking events were used. Secondary emissions were shown to contribute less than 1% to the modelled indoor $PM_{2.5}$ concentrations; international modelling studies estimated them to be between 0.4 and 3% in China and up to 6% in the USA (Fazli & Stephens, 2018).

2.7 Summary

This chapter has highlighted similarities and differences between the composition of the outdoor air and the indoor air found in a dwelling stock. Exposures to high levels of contaminants have been linked to increased risks of negative health outcomes, both at national and international levels. Much of the Chilean studies have focused on the outdoor air in polluted cities. Only in the past decades has there an increased literature on indoor air quality, but generally limited to houses with children and elderly people,

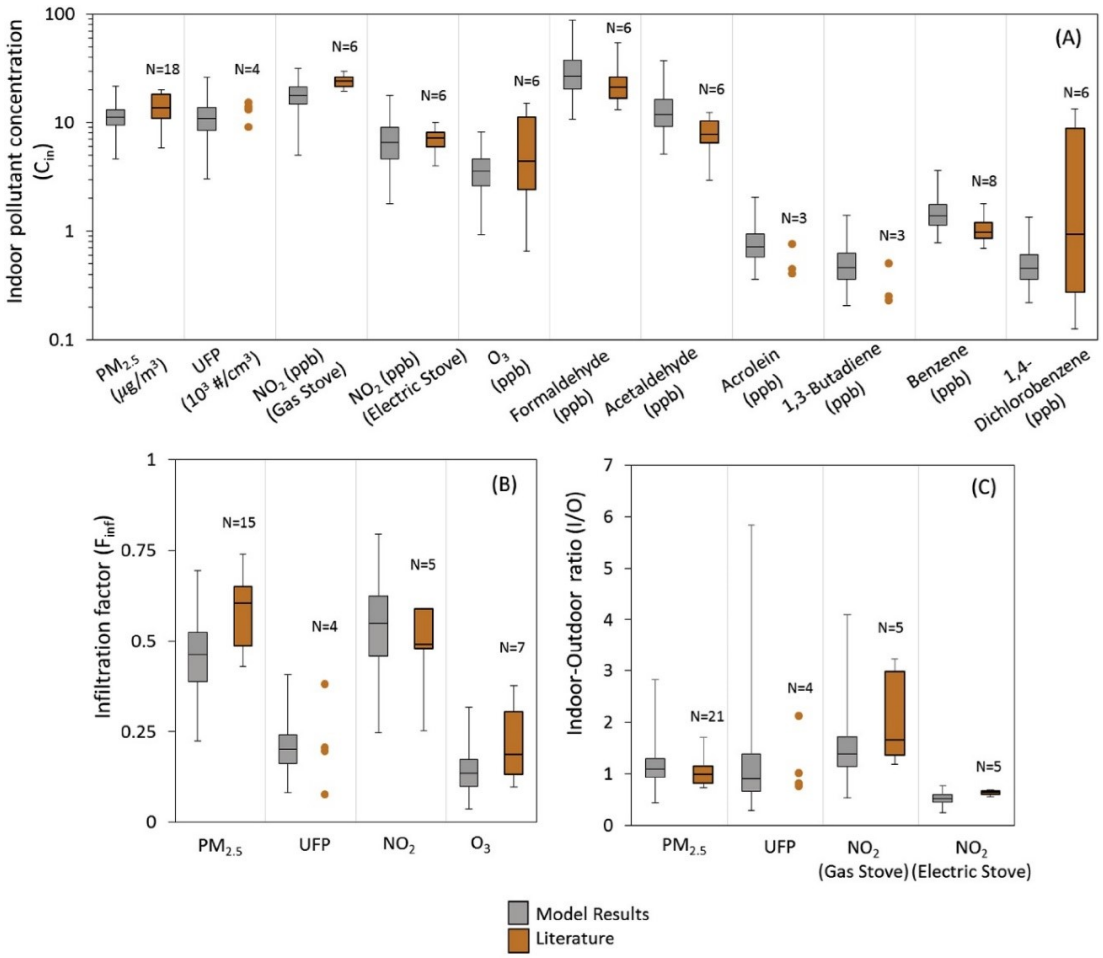


Figure 2.19: Top: Simulated distributions of the annual hourly mean of pollutants in houses before population weighting (grey), and average values using the number of published field studies N (orange).
Bottom: Infiltration factor F_{INF} and Indoor/Outdoor I/O ratios.
The box bounds the inter–quartile range and the whiskers represent minimum and maximum values (Fazli & Stephens, 2018).

and in low–income households. Recently, attention has turned to the use of polluting household appliances. And, while these field studies have been useful for reporting and describing the conditions in which vulnerable people may be living, their particularities limit the strengths and generalisation of the findings in other populations, settings, and locations.

The review of national and international literature has shown high spatial and temporal variability of pollutant concentrations and personal exposures. This is the result of a combination of factors, such as the local and surrounding environment, the building design and use, and occupant behaviour. Cross sectional and longitudinal studies carried out around the world emphasise the importance of the indoor air quality in the home, as it is the place where the majority of people spend most of the time, and where most exposures occur.

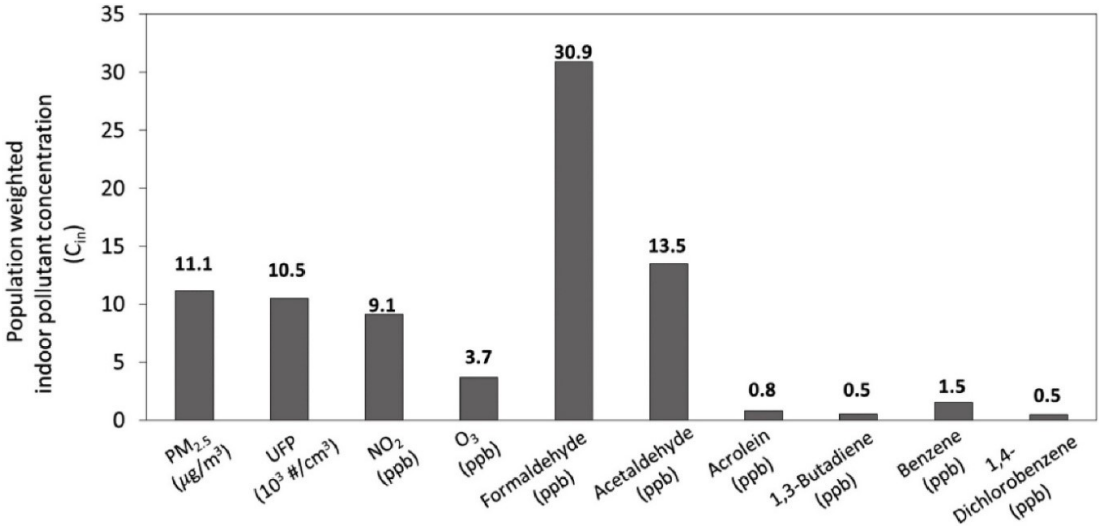


Figure 2.20: Population–weighted average concentrations of modelled pollutants over a year (Fazli & Stephens, 2018).

Different methods can be used to study air in a single dwelling and those of a stock. They can be classified into two groups, direct and indirect. Direct methods include monitoring or field measurements using personal or stationary equipment, and biomonitoring using biomarkers. Indirect methods comprise modelling and simulation techniques. Overall, simulations are preferred to field studies. Nevertheless, they rely on developed models that replicate building physics and dynamics. Therefore, their predictions then differ from measured data. When sufficient high quality input data is available, probabilistic modelling and sensitivity analysis methods can use that data to better inform about the inputs, validate and calibrate predictions. Their outputs have been used as tools to inform interested parties about the *status quo* and hypothetical scenarios. By modelling factors that contribute to indoor air pollution following a bottom–up approach enables the modeller to evaluate each factor individually, and to carry out a sensitivity analysis of the outputs to each of the inputs. This method has proven to be beneficial, particularly in countries where data is limited. In the Chilean context, indoor air modelling studies for its building stocks are scarce.

Although the indoor air in a stock of houses can be modelled with a high level of accuracy, there are inevitably uncertainties in the predictions. Some are related to the model used to represent the phenomenon, some to the heterogeneity of all possible scenarios and cases under investigation, and some to the random variation of the inputs value, either due to lack of knowledge (epistemic uncertainty), or due to the natural randomness of the input (aleatory variability). Ways of handling these uncertainties include the use of clustering techniques, which reflect the variability between groups, and Monte–Carlo sampling methods that account for variability in the descriptive parameters within groups.

Finally, the way predictions are analysed, presented, and interpreted is of critical importance to stakeholders and policy–makers. Therefore, the delivery of either simulations predictions or field measurements must account for uncertainties involved in each process. They should not hide the spectrum of possible results or measured values, but show which values are more likely to occur and those that are extreme and, therefore, unlikely.

CHAPTER 3

Data analysis of the Chilean housing stock

In Chapter 2 the importance of providing a better description of the indoor air in houses and its variability in different stocks was identified. High levels of pollutants are considered as a risk to public health and to the environment. Exposures in buildings, especially in houses, are of greater concern due to the time people spend there, in which they are exposed to a mixture of known and unknown pollutants. Those which are of concern should be given priority, well supported by scientific evidence, and so necessary actions can be taken to better inform and protect the public. They are known as *criteria pollutants* and herein considered as such.

In order to advance knowledge in this area, firstly, the uncertainty in indoor air quality exposures in the stock has to be identified and acknowledged. A diagnostic tool capable of describing a scenarios and contexts of interest and delivering a probabilistic assessment can be an advantageous and cost–effective way to better guiding and informing decision–making and research.

A wide range of inputs are required to accurately simulate mass transfer in houses. However, this information may not exist, or may only exist exist at an aggregated level. Furthermore, the collation and processing of data can be time consuming and can introduce systematic errors when they are not tracked. Any tool used to model a stock must be informed by the most reliable sources of information. Accordingly, this chapter addresses this need in Section 3.1 by examining all available data on the Chilean housing stock that could be used to inform a model. Furthermore, areas of paucity are highlighted, thus informing future surveys and research. Then, in Section 3.2, the information is combined to characterise the stock considering parameters such as building type and geometry, age, size, occupancy, and potential sources of pollution.

3.1 Sources of data

This section identifies sources of data held by the various Ministries and Institutions of the Chilean government, or relevant non–governmental organisations. They include population scale censuses carried out at large intervals, a tri–annual socioeconomic survey of a sample of the population, information on all newly constructed houses, a nationwide network of measurements of indoor environment parameters made in houses, and weather data that represents national climate variations.

Table 3.1: A comparison of variables and categories given by databases.

VARIABLE	**DATABASES** (Number of Categories)			
Source	Census 2002	Census 2012	Building Permits	CASEN Survey 2015
Frequency (size)	10 years (15million)		Monthly (712k)	2–3 years (70k)
Geometry				
Type of house or No. of shared walls	2	4	4	4
Number of storeys	✗	✗	✓	✗
Floor Area	✗	✗	✓	6
Location (rural/urban)	✓	✓	✓	✓
Construction material				
External walls	6	5	12	6
Roof	9	6	14	8
Flooring	9	7	13	6
Partitions	✗	✗	9	✗
Zones				
Number of rooms	✓	✓	✓	✗
Number of bedrooms	✓	✓	✗	✓
Main use of rooms	✓	✗	✗	✗
Number of showers or bathrooms	✓	✗	✗	✓
Occupants				
Total	✓	✓	✗	✓
Men	✓	✓	✗	✓
Women	✓	✓	✗	✓
Energy sources				
Cooking fuel	8	8	✗	8
Heating fuel	✗	8	✗	8
Hot water	✗	8	✗	8

✓, included; x, not included. Integers describe the number of categories for variables considered useful for modelling purposes.

3.1.1 National Census

A census aims to record information about an entire population rather than a sample of it, and so it is an accurate and valuable source of data. The Chilean Population and Housing Census has been conducted approximately every 10 years since 1952; Table 3.2 gives the total entries for each. The data is summarised for public dissemination by the National Institute of Statistics (INE) and the raw data is available upon

request. Both the summary documents and the raw data from the 2002 and 2012 censuses (INE, 2003, 2013) were analysed in this study.

Table 3.2: A comparison of variables and categories given by databases.

Census	Occupied Dwellings	Households	People	Omissions (%)
1952	1,091,446	n/a	5,932,995	6.1
1960	1,365,577	n/a	7,374,115	4.5
1970	1,724,951	1,715,937	8,884,768	7.5
1982	2,320,246	2,466,653	11,329,736	1.5
1992	3,111,392	3,293,779	13,348,401	1.9
2002	3,915,963	4,141,427	15,116,435	3.8
2012	4,924,926	5,035,637	16, 634,603	9.6
2017	5,804,375	5,651,637	17,574,003	n/a

Table 3.2 shows that the 2002 census has higher coverage than the 2012 census and the percentage of omissions increased from 3.8% to 9.6%. There are also significant differences in the methods and questionnaires used by the two censuses. Therefore, a technical audit (INE, 2014) was carried out immediately after the 2012 census to analyse the data and evaluate whether it achieved its aim. Several inconsistencies were identified, such as significant changes in the gender ratio that was inconsistent with birth and death rates, which suggested that the fieldwork was deficient and the data is of limited use (INE, 2014). A further audit found the 2012 census to be unhelpful in developing new public policies or for estimating the demand on public services. To understand the evolution of the housing stock, it is important to able to compare censuses to identify changes. However, the two censuses cannot be merged because of differences in the categorical data recorded and types of questions included for some variables. For example, in the 2012 census the categories for the *type of building* variable were *flat/apartment with elevator, flat/apartment without elevator, detached, attached (semi–detached or terraced)*, or *other*, whereas the 2002 census used categories of *flat/apartment* or *house*. Herein, the term apartment is used. Furthermore, the 2012 census asks fewer questions about the number of rooms in a dwelling and their use. Consequently, the 2002 census is considered to be the more robust of the two. [1]

The censuses ask a number of questions that help to characterise types of dwellings and household and are useful for understanding the stock (see Table 3.1 and Sec-

[1] A new census was conducted in 2017 to correct the problems with the 2012 census, although the raw data is not yet publically available. This new census used categories and questions that were similar to those used in 2002 but in a abbreviated version and so the analyses done with these data may not be possible to be replicated until a full version is available.

tions 3.2) or for modelling it (see Chapter 4).

For each dwelling the number of rooms and their use, the presence of bathrooms (including the number of showers) and a kitchen, the dwelling type (*house, apartment/flat* or *other*), and services present (boiler, water supply, sewage, heating system (only in the 2012 census), photo–voltaic solar panels, and appliances, such as a cooker (with fuel type), washing machine, or fridge) are all recorded. All Chilean censuses use the *de facto* definition of residency so that all occupants of a dwelling are interviewed at the time of the census. Occupants identified as guests are counted in their own household rather than the surveyed household. If a dwelling is empty, it is considered as occupied but absent. In addition to the number of people, their gender and age, employment status, general health (including disabilities), and social–economic status (using a scale ranging from one for the first decile or most vulnerable houses to 10 corresponding to the most wealthy) are also recorded. Regional information is included so that the distribution of the population around the country can be determined thus helping to government to provide adequate utilities and public services.

However, there are a number of parameters and metrics that are not surveyed by the census that would be useful to know in order to construct basic steady–state energy demand models, similar to the UK's Cambridge Housing Model (see Sousa *et al.* (2017) and Jones *et al.* (2015) for more details). These include dwelling properties, such as geometries (floor area and volume), window area and glazing type, year of construction, number of floors, insulation level, internal temperature (perhaps indicated by a thermostat setting), orientation, and heating system fuel. It would also be useful to understand some the type and frequency of occupant activities, such as appliance use, cooking, or tobacco consumption, to improve estimations of energy demand, GHG emissions, IAQ, and occupant health risks. However, these parameters may not fall within the scope of the census, and so further data sources are required.

3.1.2 Building Permits

More detailed building information is contained within the Building Permits' datasets (INE, 2016a). This data has been collected since 1929 in Santiago and since the 1980s for the rest of the country. The raw dataset obtained for this study covers the period between 1990 and 2015 and provides information on 3.1 m houses, approximately 54% of the current national housing stock. Information about a proposed building is submitted to the local government (municipality) when requesting a construction permit. It is then forwarded to the INE, which compiles and publishes summary statistics annually.

Although this dataset provides additional information, some values of physical

dimensions suggest that either the data gathering or the transcription processes introduced errors. Therefore, the modeller must decide whether to exclude these data from the study or to clean the data set. In this study, all the questionable data was removed and the variables were re–coded so that they can be used in conjunction with other databases; see Section 4.1. Approximately 2% of records were removed so that the number of dwellings in the Santiago Metropolitan region after cleaning was 1,278,328, which corresponds to 57% of the current housing stock.

3.1.3 National Socioeconomic Characterisation Survey

The National Socioeconomic Characterisation Survey (CASEN) is a cross–sectional survey of education, employment, income, and health status that has been conducted by interview every two or three years since 1987 by the Ministry of Social Development. In 2015, the sample size was 83,887 households and 266,968 people living in 15 regions of the country, and is considered nationally representative (CASEN, 2015; GHDx, 2018). Of particular interest is its reporting of dwellings (housing materials and conditions), the household (family size, number of children, education, employment, family composition, hours worked, income, living conditions, cooking fuels, and household air pollution), utilities (electricity, household heat, and water supply), and occupant health (general health status and mobility, and prevalence of asthma, chronic obstructive pulmonary disorder, cancers, and ischemic heart disease).

3.1.4 Use of Time national survey

The way people behave in a dwelling is usually more significant in determining energy demand than either the size of the dwelling or the household size (Palmer & Cooper, 2013). Accordingly, it is important to understand how people use their time. The Use of Time National Survey (ENUT) database is a nationwide cross–sectional study. The data comprises a randomly sampled group of 15,312 dwellings using a self–reported questionnaire. A limited number of activities and their duration during one normal week and weekend day are recorded (INE, 2016b). However, personal location is not recorded and must be assumed by an activity. For example, the average time working people in the capital region spend at home is 14.5 hours (58.77% of the day), the time in–transit or outside is 5.9 hours (2.9%) and the time at work is 8.7 hours (35.38%). The survey only focusses on an activity and so the duration does not add up to 24 hours and the percentages to 100%.

Longitudinal data is required to understand how personal activities vary throughout a year and the determinant or explanatory variables that affect decisions. If the season, location, and weather conditions are given, it is possible to describe occupant

profiles more accurately and thus estimate more realistic exposures to pollutants and their impacts on health. Therefore, in the short term, the cross–sectional ENUT data must be adapted to infer longitudinal activity patterns that can be used by modellers (McCurdy & Graham, 2003).

3.1.5 National Housing Monitoring Network

The Ministry of Housing, known colloquially as the Ministerio de Vivienda y Urbanismo (MINVU), is running an ongoing program of *in–situ* indoor and outdoor environment monitoring in houses, which started in May 2017, known as the National Housing Monitoring Network or colloquially as the Red Nacional de Monitoreo de Viviendas (RENAM) (RENAM, 2018). RENAM has installed real–time sensors in 299 houses (all concurrently) located in the north (N = 28), south (N = 60) and centre (N = 211) of the country in five cities. Most (N = 150) are located in the metropolitan and capital region of Santiago de Chile, which is inhabited by 41% of the population and whose stock is estimated to comprise 2.4 million houses (INE, 2003). The monitoring is preceded by a survey that provides information about the location (region, city, and commune), the dwelling (type, construction year, storeys, floor area, orientation, envelope materials, number of windows, glazing properties, and heating system), the householders (income, energy bills, health characteristics issues), and behaviours (heating months, weekday and weekend occupancy, smoking). The platform is accessible online (RENAM, 2018) to registered users and contains the location of each sensor and all surveyed and measured data.

Three different devices are used. The first is the Netatmo weather station (Netatmo, 2018), which comprises indoor and outdoor modules. Both record air temperature and relative humidity, but the indoor module also measures sound level, and CO_2; see Table 3.3. The majority of sensors are located in the living room, although some have been placed in a bedroom, and measurements are time–averaged over 30 minute periods. The second is a Plantower (Yong, 2016) which is capable of measuring the concentration of particles in air with a diameter of 0.3–2.5 μm (known as $PM_{2.5}$), and the third is a Wenu Work smart meter that records the electricity demand in houses that use electric heaters exclusively (N= 50).

There are known limitations of consumer–grade IAQ monitors; see Singer & Delp (2018). All of the sensors require annual calibration. Most optical CO_2 sensors are prone to zero–drift and so the Netatmo periodically self–calibrates when the measured concentration is both low and steady, and this assumes minimum to be predefined ambient concentration. Therefore, its readings are imprecise and heavily dependent on the differences between the actual ambient, defined calibration, and measured con-

Table 3.3: Monitoring network: data and metrics.

Variable	Measurement range	Data unit	Sensor accuracy
Indoor temperature	0 to 50	°C	±0.3
Outdoor temperature	-40 to 50	°C	± 0.3
Indoor relative humidity	0 to 100	%	±3
Outdoor relative humidity	0 to 100	%	±3
Sound pressure level	35 to 120	dB	-
Indoor CO_2	0 to 5,000	ppm	±50
Indoor $PM_{2.5}$	0 to 500	$\mu g/m^3$	±10

centrations. There are no studies quantifying the deterioration in accuracy of low–cost $PM_{2.5}$ sensors over time, but their precision depends on the difference between the size fractions they are sensitive to and those emitted by a source. This is compounded by differences in the refractive index of particles from varying sources, and so for accurate measurements a source must be identified and a calibration factor applied, which can range from 0.016 to 12 (Fischer & Koshland, 2007; O'Leary *et al.* , 2019a). Therefore, the outputs of these devices can be considered indicative rather than exact (Jones *et al.* , 2018).

The sample is randomly selected but cannot be said to be statistically representative of the wider housing stock, but statistical methods can be applied to generalise findings. The data is expected to identify broad trends in indoor environment quality and highlight areas worthy of more detailed investigations. Furthermore, a top–down modelling approach could use the categorical survey data to identify trends in specific groups of dwellings defined by their type, geography, or the socioeconomic status of their occupants.

3.1.6 Airtightness

The permeability of a dwelling's thermal envelope is conventionally assessed using a blower–door, which artificially and systematically increases the difference between the internal and external air pressures ΔP (Pa) and measures the airflow rate through adventitious openings located in it $\dot{V}$ (m^3/h); see Jones *et al.* (2015). These parameters are related by a power law:

$$\dot{V} = C\{\Delta P\}^n \tag{3.1.1}$$

where C ($m^3\,h^{-1}\,Pa^{-b}$) is a flow coefficient and n is a flow exponent. It is common to report $\dot{V}$ at 50 Pa, interpolated from measurements, when it is known as an air leakage

rate, $\dot{V}_{50}$ (m^3/h). In order to compare the air leakage rates in different dwellings, it is normalised by a common parameter, such as the thermal envelope area when it becomes an Air Permeability, Q_{50} ($m^3\,h^{-1}\,m^{-2}$), or the building's volume to give N_{50} (h^{-1}). These values of airtightness are often used to infer an infiltration rate, the rate unconditioned ambient air passes through adventitious openings located in the thermal envelope of a dwelling.

Measurements of airtightness are not a legal requirement in Chile and there is no a database equivalent to that compiled in the U.S. (Chan *et al.* , 2013), which contains more than 160,000 measurements. The most significant number of measurements have been made in 187 dwellings built in between 2007 and 2010 (Citec UBB & Decon UC, 2013), immediately after the implementation of national regulation of fabric thermal performance (MINVU, 2007), and so this sample was selected to be statistically representative of them. Dwellings were selected by common parameters, such as the year of construction, the main structural material, and the dwelling type. For reference, the frequency of Q_{50} values is presented in Figure 3.1, which shows that 20% of the tested dwellings are over the UK limit of $10\,m^3\,h^{-1}\,m^{-2}$.

Several studies (Murray & Burmaster, 1995; Persily *et al.* , 2010; Shi *et al.* , 2015), find good correlations between N_{50} values measured in a housing stock and a logarithmic function. Figure 3.1 shows a positively skewed frequency distribution (skewness=2.62) with a log–normal probability density function PDF fitted to it with a mean of μ =1.8 $m^3\,h^{-1}\,m^{-2}$ and a standard deviation σ =0.653 $m^3\,h^{-1}\,m^{-2}$ as estimated parameters; see Figure 3.1. Similar process is carried out with N_{50} values, giving a μ and σ of 6.24 h^{-1} and 0.85 h^{-1} respectively.

This project has access to 65 measurements of air permeability made in Santiago, where Q_{50} and n are reported with dwelling geometry and structural materials. Data is aggregated by the construction materials and the building type in Figures 3.2b and 3.2a, respectively.

Because Q_{50} values do not meet the normality assumption, a Kruskal–Wallis test is carried out to test similarity in the means when data is grouped by categorical variables. This is a non–parametric test similar to an ANOVA test for normal data, which tests whether the difference between groups is statistically significant; see Section 7.6 about the interpretation of p–values. According to the test, Q_{50}s show significant differences by the housing type ($p = 2.54e-06$) and by the main structural material of the envelope ($p = 6.39e-11$). Multiple comparisons of the mean ranks between groups (a *post–hoc* procedure after Kruskal–Wallis test, where *difference* corresponds to the difference between the mean ranks of two groups irrespective of the measured values) showed that the mean rank of Q_{50} for terraced houses was significantly lower than for both detached houses (mean ranks *difference* of 50.1) and semi–

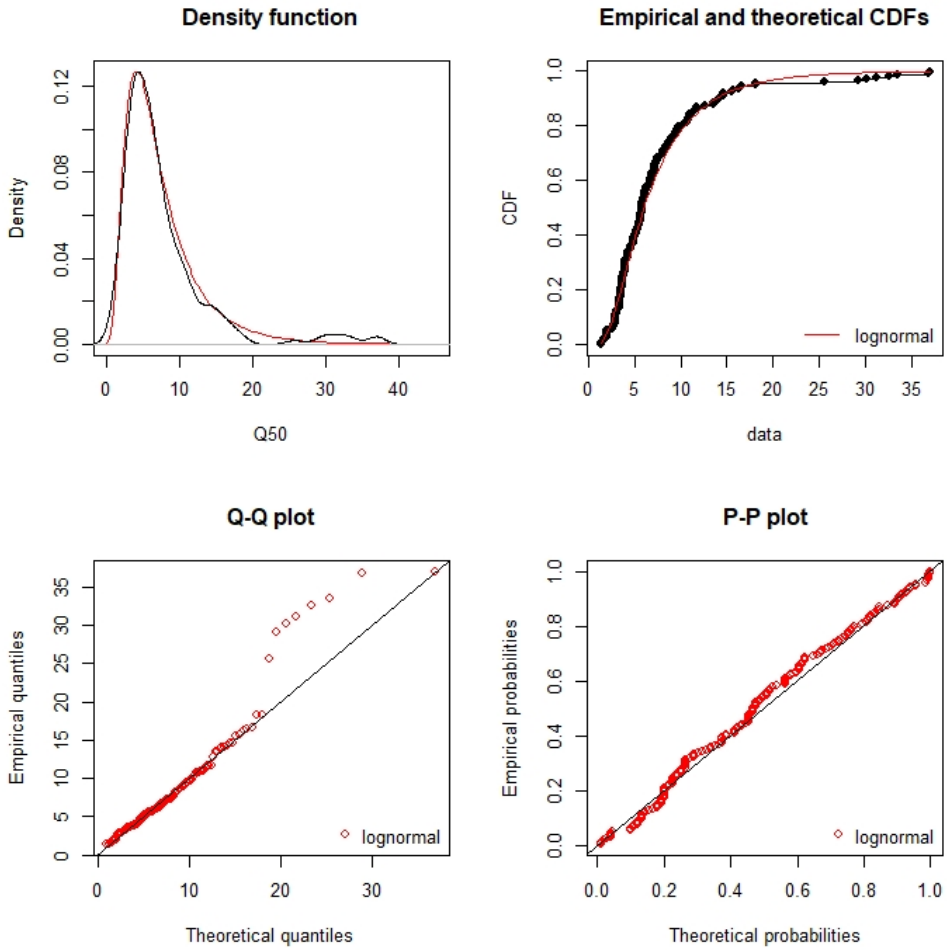
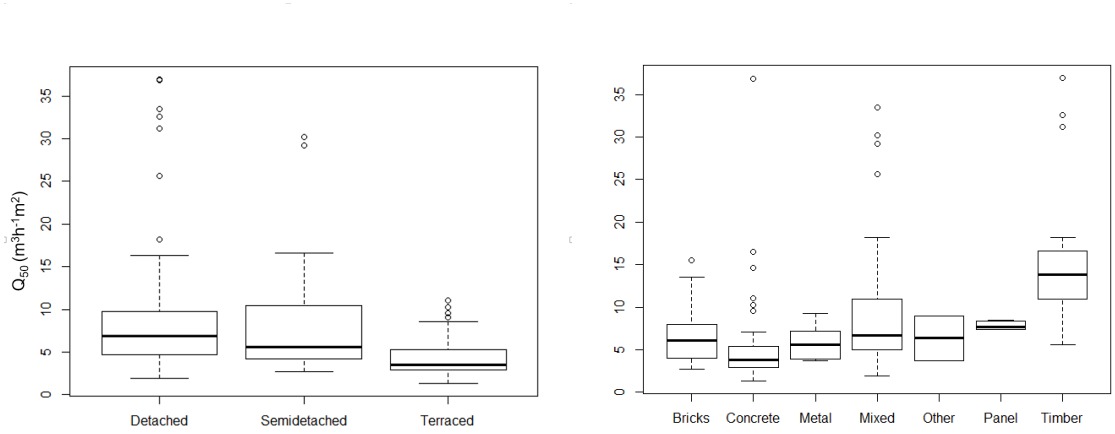


Figure 3.1: 187 Blower Door tests. The red line in the top–left and top–right plots correspond to a lognormal distribution that best fit the data. Red dots in the plots at the bottom correspond to the empirical data. P–P correlation coefficient value for the lognormal distribution R^2 is 0.973. Analysis and plots made by the author using data retrieved from Citec UBB & Decon UC (2013).

detached houses (*difference* of 40.8). When considering the main structural material of the envelope, masonry has a significantly higher Q_{50} than concrete (*difference* of 33.7) but is lower than timber (*difference* of 67.3). Q_{50} was significantly higher in timber houses than mixed structures (*difference* of 49.6) and concrete (*difference* of 101.0), and mixed structures with concrete (*difference* of 51.4). Due to the inequality in the sample sizes of the groups, the critical difference ($\alpha = 0.05$) was calculated for each comparison individually. This information provides some empirical data on the difference in air leakage for new houses and some insights into its relationship with structural materials and house types.

Infiltration cannot be assumed to be a constant regardless of time of year and age of the building since it varies with pressure difference and ageing of the envelope (Chan



(a) **(b)**

Figure 3.2: Q_{50} measured in 187 Chilean houses aggregated by **(a)** building type, and **(b)** main structural material. The whiskers represent the 5th and 95th percentiles, and the dots show outliers. The boxes bound the interquartile–range, and the bold lines are the 50th percentiles or median (Citec UBB & Decon UC, 2013).

et al. , 2015; Liddament, 1986; Persily *et al.* , 2010), and so these models need to account for this variation. To date, there has good agreement on the best distribution of the airflow exponent n in Eq. 3.1.1. Flow exponents from measurements in 1,290 houses located in North America, New Zealand, the UK and the Netherlands, are found to be normally distributed with a mean of 0.65, although no standard deviation is given (Liddament, 1986). In Sherman & Dickerhoff (1998) (N = 1,942), n values are normally distributed, $N(0.651, 0.077)$. Similarly, Chan *et al.* (2013) find $N(0.646,0.057)$, but are only for detached houses in the USA. Values close to 0.5 are associated with fully developed turbulent flow and 1 with laminar flow.

Figure 3.3 shows the distribution of n for the 65 Chilean houses compared to those of Chan *et al.* (2013) and Sherman & Dickerhoff (1998). Lower values of n might indicate that infiltration in the Chilean houses is mainly through large sharp edged openings that cause turbulent airflow.

A second dataset comprises 58 social houses located in the centre–south and extreme south of the country where space heating is often required (Fissore A, 2013). These houses are a statistically representative sample of 15,000 low–income and uninsulated houses that received insulation and improved airtightness interventions from an ongoing weatherisation program that started in 2009 and is subsidised by the Chilean Government. Only N_{50} values were reported and Figure 3.4 shows the changes in each house and their variability. Most houses have a significantly reduced N_{50} after the intervention, and the more substantial reductions are seen in the highest leakage homes (see data points below the line of Figure 3.4–right), M_{dn}=36.0h^{-1}; mean

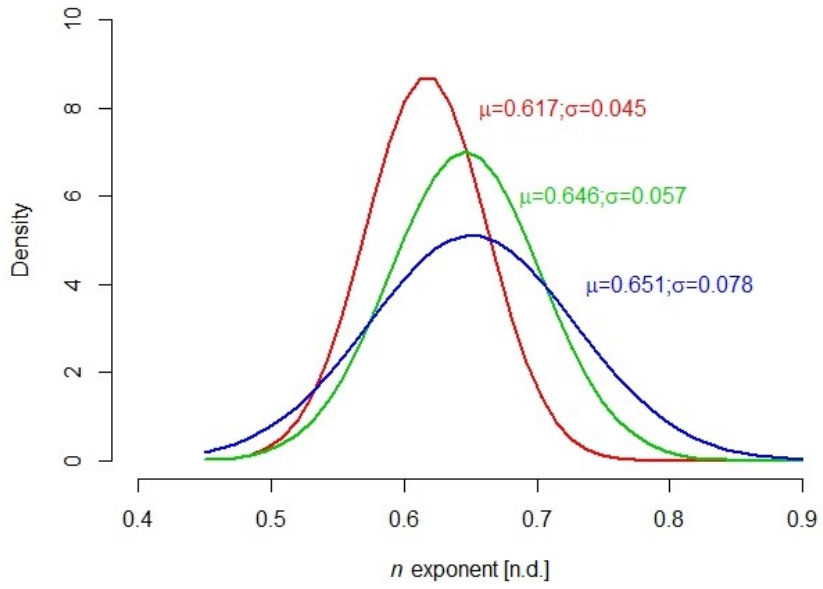


Figure 3.3: Density distributions of airflow exponents, n, for Blower Door tests of Chilean houses (N = 65), red line, and U.S. dwellings reported by Chan *et al.* (2013) ($N = 134,000$), green line, and Sherman & Dickerhoff (1998) ($N = 1,942$), blue line.

$\tilde{\mu}$=41.9 h^{-1}; σ=23.9 h^{-1}) than before it (M_{dn}=39.0h^{-1}; $\tilde{\mu}$=51.2h^{-1}, σ=34.6h^{-1}). When performing a sign test, the results show that the difference between the medians of the two groups is statistically significant: $Z(58) = 3.2 - 2.54$; $p \leq 0.05$; Cohen's effect size index, $\delta = 0.39$ (small).

These are very small samples, which do not capture the variability of N_{50} across the stock. Infiltration is responsible for a significant proportion of the energy demands and GHG emissions of national housing stocks (Jones *et al.* , 2015). Therefore, future field work is required to develop a database that can be used to identify construction quality, and infiltration's contribution to Chilean national energy demand and GHG emissions.

3.1.7 Weather

Chile has a continental territory with a north to south length of 4,300Km and an average width of 177 km. It is bounded on the West by the Pacific Ocean and on the East by the Andes mountain range, and so elevations range from sea level to 6.8 km. The mean annual air temperature varies by up to 6°C laterally and by more than 15°C longitudinally, and differences in relative humidity are similarly well defined (Castillo, 2001).

Chile is divided into 15 regions located in 9 different climate zones (INN, 2016). Weather data for Chile is available from three sources. The first is the American So-

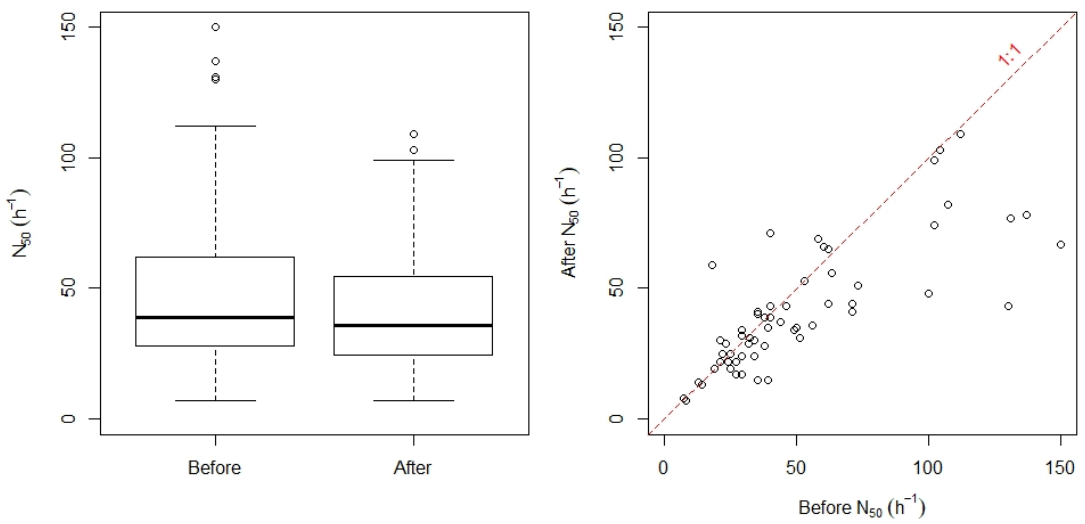


Figure 3.4: Changes in permeability after the weatherisation program (Fissore A, 2013). **Left**: Box plots showing the low values, first quartile, median, third quartile, high values and outliers of N_{50} before and after the intervention. **Right**: Paired information. Dashed red line indicates no difference. Figures by Author using data retrieved from Fissore A (2013).

ciety of Heating, Refrigerating and Air–Conditioning Engineers (ASHRAE) International Weather for Energy Calculation version 2.0 (IWEC2) database (ASHRAE, 2018), which provides weather data for 14 locations in Chile. The data is averaged over 12 to 25 years giving wind speed and direction, sky cover, visibility, dry–bulb temperature, dew–point temperature, atmospheric pressure, and liquid precipitation at hourly intervals. A second database is provided by the Department of Meteorology of Chile (DMC), which contains climate averaged data from 24 weather stations for a 30–year period between 1970 and 2000, organised by month and year. A third source is the Meteonorm meteorological database that uses Meteonorm software (Meteonorm, 2017) to stochastically generate time–series data for typical years. It does this by interpolating from long term monthly means for a location by combining a database of ground station measurements with those from five geostationary satellites. The root mean square error for ambient air temperatures is 1.2°C (Meteotest, 2017).

There are currently no weather files that can be used to predict future climate change scenarios, although guidance could be sought elsewhere; for example, the future weather files of the Chartered Institution of Building Services Engineers (CIBSE, 2014).

3.2 Stock characterisation

This section draws on the data source described in Section 3.1 and uses it to characterise the housing stock and its occupants.

3.2.1 Dwelling quantity, type, and geometry

The Chilean housing stock comprises 6.5 million residential units, where 79.5% are houses and 17.5% are apartments (INE, 2018b). The remaining 3% includes mobile homes, uninsulated timber *Mediaguas* used for temporary or emergency accommodation, and self–constructed (using local materials) *Ranchos*. The characteristics of the stock vary according to the local weather conditions and the availability and the affordability of building materials. This variation is considerable because of Chile's varying geography (see Section 3.1.7).

The best level of detail on dwelling type is given by the Building Permit dataset; see Table 3.1 and Figure 3.5. It registers the new houses as (i) detached, (ii) semi–detached or (iii) terrace, and the apartments according to the number of building blocks (from 1 to 9 or more). As both semi–detached and end–terrace houses share one wall with the adjacent, they can be considered as belonging to the same class. Therefore, using this data source, three categories can be considered according to the number of attachments or shared walls: (i) detached, (ii) semi–detached and end-terrace, and (iii) mid–terrace. None of the databases include information about the position of a dwelling within a terraced block of houses (as end or mid-terrace) or the number of houses in a block. Therefore, the proportions of end– and mid–terrace buildings are estimated according to the probability of obtaining a given housing type. For example, the maximum number of houses in a block is arbitrarily limited to 20, and when a terrace contains 3 houses, 1 is classed as terraced whereas the other 2 houses are classed as semi–detached.

3.2.2 Year of construction

In Chile, the age of buildings is currently undocumented. An estimation can be made using a Building Permit date (see Table 3.4), which is the date when the building was approved for construction, rather than the date of actual construction. This is an area of epistemic uncertainty and future surveys should aim to estimate the age of dwellings.

The envelope structure and design of a house affects its thermal performance and airtightness.

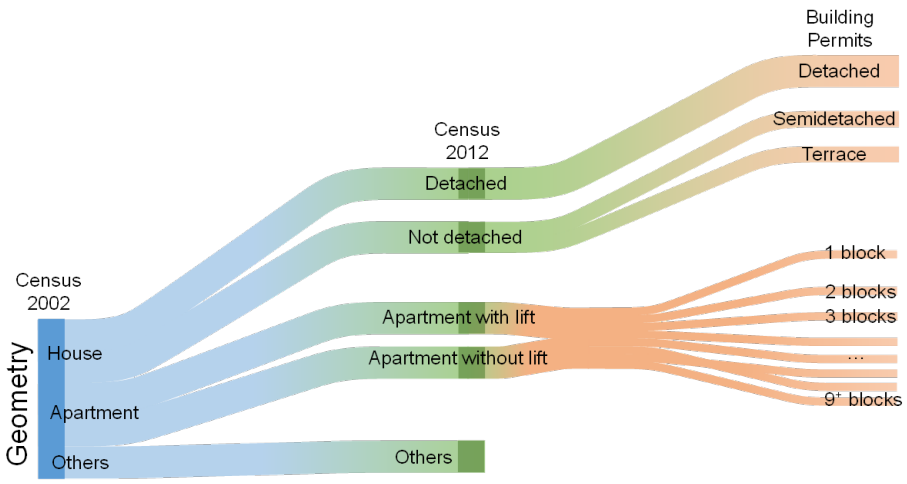


Figure 3.5: Categories of type of house given by three sources of data INE (2003, 2013, 2016a).

Major changes in the building code for new houses started in 2000. Dwellings built before 2000 had no thermal requirements and are generally considered as non–insulated, although some may have been. Dwellings designed and approved for construction between 2000 and 2007 are required to meet a maximum thermal transmittance for the roof, but not for the walls and windows. After 2007, building codes were upgraded and all new dwellings must meet a maximum thermal transmittance for all envelope components in contact with the ambient air. The requirement varies along the length of the country depending on the number of heating degree days. New changes to the building codes are about to be adopted, and so the database will need to be updated in the near future to include this stock and new archetypes may be required.

Accordingly, the housing stock here is divided into two age–related groups: those constructed before 2008 when little or no insulation was required by law, and those built thereafter with high levels of insulation; see Table 3.4. Houses constructed before 2008 that have been weatherised are not included. The Building Permits are used to obtain the proportion of the stock belonging to each group because they contain the total number of dwellings registered per year (INE, 2016a).

Table 3.4: Proportion of dwellings by construction period (INE, 2016a).

Construction period	Number of Dwellings	Percentage
Before 2008	2,057,142	66.2%
From 2008	1,052,429	33.7%
Total	3,122,950	100%

3.2.3 Number of rooms and floor areas

Floor area is a key parameter and an independent variable with relevance to both energy and mass transfer models. Building Permits register the total floor area for every new dwelling, and the CASEN survey categorises the useful floor area into six bands. Buildings Permits (INE, 2016a) issued between 1990–2015 show that the average floor area of new dwellings has increased by around 40% over the last 27 years from 57 m^2 in 1990 to 82 m^2 in 2015; see Figure 3.6. This is important because floor area is often correlated with energy demand (Palmer & Cooper, 2013), and suggests an increase in energy demand and carbon emissions attributable to the housing stock over the same period. However, the national mean floor area of $\mu = 65.5\,m^2$ ($\sigma = 46.1\,m^2$) is modest when compared to those of many other nations; see Figure 3.7. The relationship between house size and the number of rooms has remained relatively stable over time, but has started to increase over the past decade; see Figure 3.8. Therefore, the energy demand and carbon emissions of the housing stock are very likely to continue to increase in the future if houses sizes approach international norms, unless steps are taken to mitigate them.

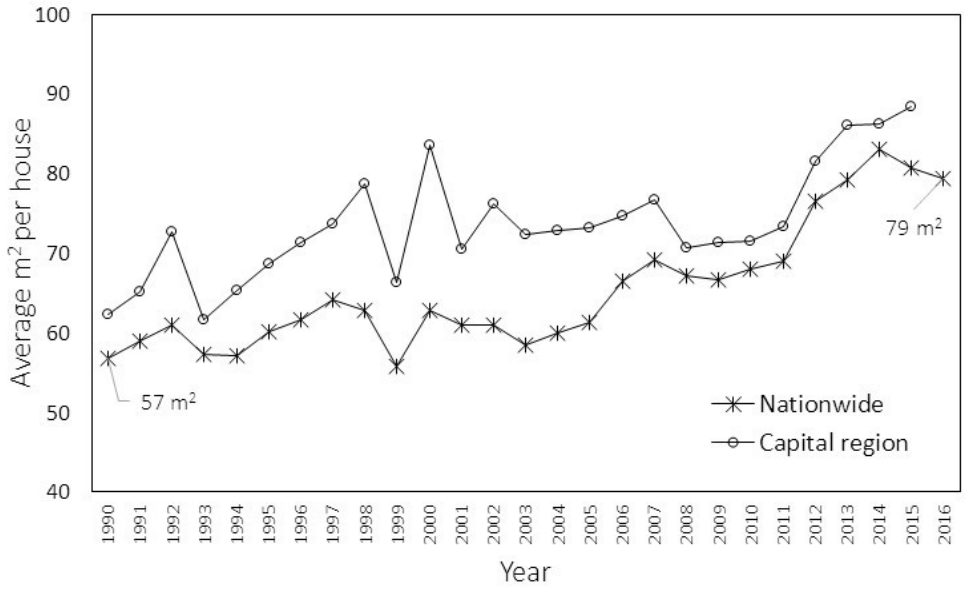


Figure 3.6: Mean floor area of Chilean houses over time (INE, 2016a).

3.2.4 Occupancy

The number of occupants, their location and activities are required when modelling IAQ and energy demands to correctly allocate pollutant sources and sinks, determine occupant exposures, and to determine the demand of services. The mean number of occupants per dwelling is 3.64 persons (INE, 2003), which is high when compared

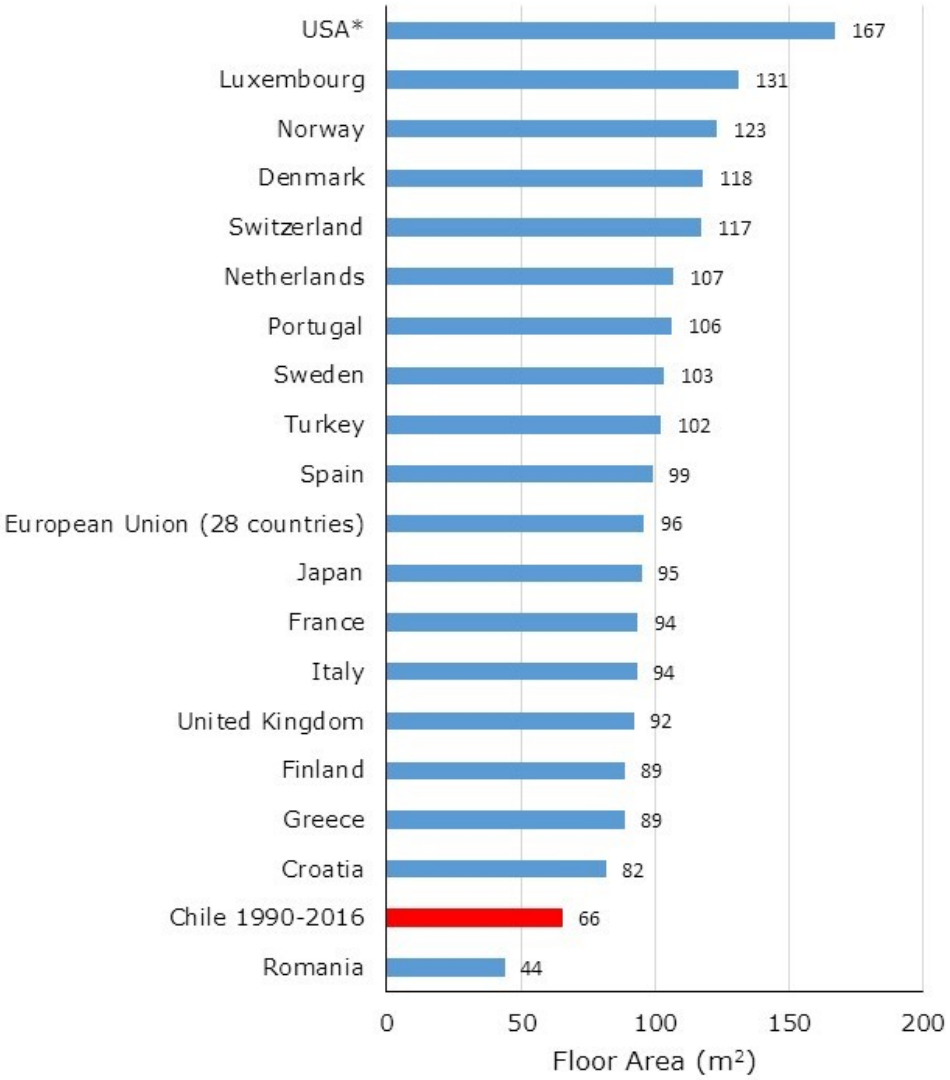


Figure 3.7: Mean floor area of houses by country. Chilean value for houses built between 1990 and 2016 (INE (2016a); Eurostat, 2012; Moura et al., 2015).
*, median value.

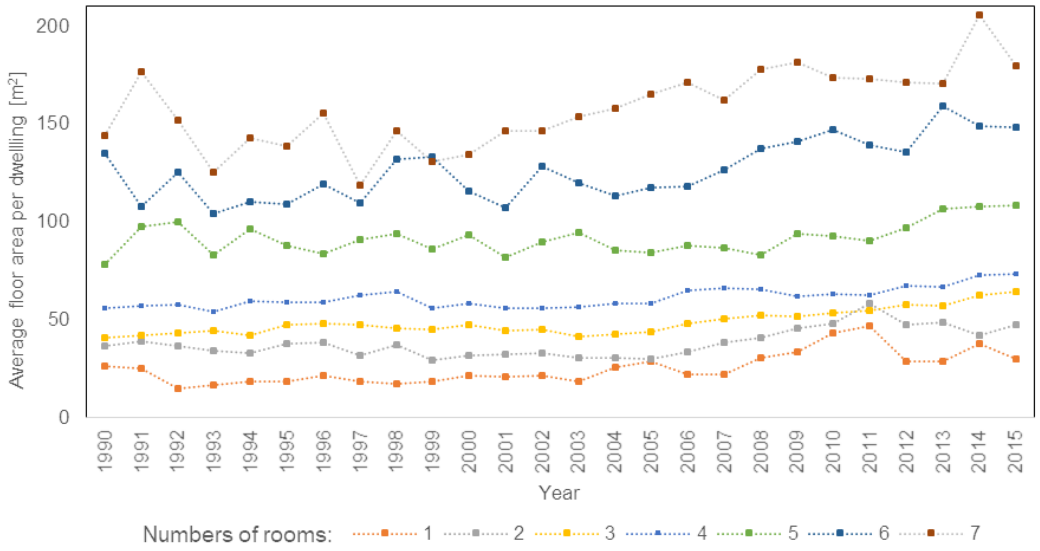


Figure 3.8: Average floor area of Chilean dwellings by number of rooms over time in houses with up to 7 rooms, accounting for 98% of the dataset (INE, 2003, 2016a).

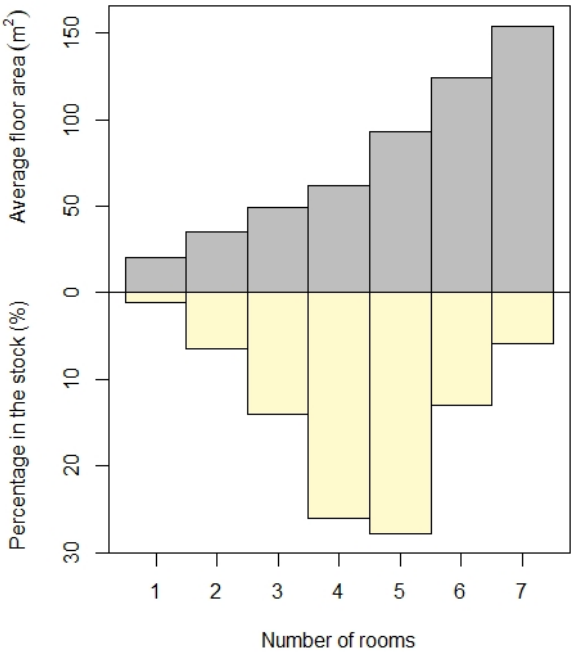


Figure 3.9: Average floor area of Chilean dwellings by number of rooms (top) and its proportion of the dwelling stock (bottom), considering dwellings with up to 7 rooms (accounts for 98% of the dataset). Data: (INE, 2003, 2016a)

to the UK's mean of 2.35 occupants per dwelling. The mean occupancy density in Chilean dwellings is 18.2 m^2 per person.

The total number of rooms in a dwelling is commonly used to determine overcrowding. In Chile, overcrowding is defined as more than 2.4 persons per bedroom, and occurs in 11% of dwellings surveyed by the 2015 CASEN, in 15% in 2002 census, and 7% in 2017 census; see Table 3.5. The European Union, the UK Government, and the American Crowding Index deem a house to be overcrowded when there is more than 1 person per room, where a room is defined as any enclosed space with habitable conditions (Eurostat, 2012), and thus excludes sanitary rooms and circulation spaces. Using this metric, overcrowding occurs in 19.6% of Chilean dwellings that contain only one household, which is high when compared to 3% in England, where English households comprise approximately 80% of all UK households (Wilson and Barton, 2018).

Some dwellings are unoccupied, either temporarily (unfurnished and ready for sale or rent or furnished but occupants are absent) or permanently (abandoned). It is important to understand the type of occupancy of a dwelling because it affects its energy demand and the health of its occupants. However, only the total number of houses corresponding to each occupancy types is reported by the censuses (see Section 3.1.1) and so this information is not available for individual dwellings.

Table 3.5: Change in household size, bedrooms, and room density over time in Chile reported by census data. n.i.: no information. *: see calculation criteria in Section 3.2.4 (page 61).

Census	Household size [People per dwelling]	Bedrooms per dwelling	Room density or crowding* [People per bedroom]
1952	5.4	n.i.	n.i.
1960	5.4	n.i.	n.i.
1970	5.1	n.i.	n.i.
1982	4.6	3.2	1.4
1992	4.3	2.5	1.7
2002	3.6	2.2	1.6

3.2.5 Cooking and heating

The combustion of fuels used for heating and cooking are significant contributors to the total energy demand of houses and their carbon emissions. The most common heating systems are stoves, where 45% are fuelled by wood, 25% by bottled LPG, propane, and butane gases, and 10% by kerosene and paraffin; see Figure 3.10

(CASEN, 2015). This diversity is not found in the UK where 90% of dwellings have a gas fired central heating system with a mean efficiency of 82.5% (Sousa et al., 2017; Jones et al., 2015).

Heating and cooking systems, and the cooking of foods, are known sources of pollutants associated with elevated incidences of asthma, wheeze, airway obstruction and lung cancer (Borsboom *et al.* , 2016). The most common cooking fuels used in Chilean houses are bottled gas (86%) and wood (13%).

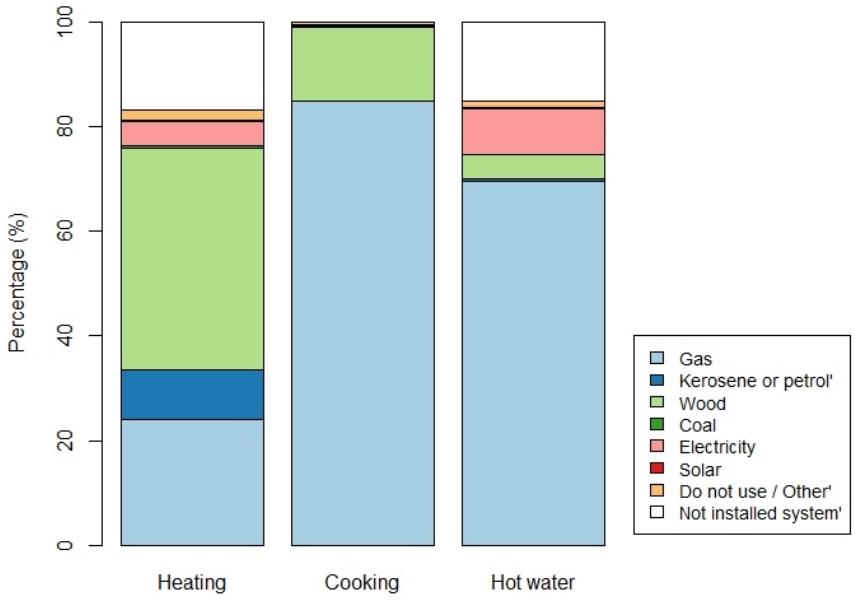


Figure 3.10: Types of fuel used by activity (CASEN, 2015).

3.2.6 Construction and finishing materials

Information on predominant structural materials used in walls, roofs and flooring are reported in the censuses. The prevalence, availability, and affordability of materials varies by region; for example, concrete block construction is more common in the north and centre of the country, brick work is most frequently used in the centre and centre–south, and wood is most commonly used in the south and extreme south (INE, 2003). More recently, new construction typologies and technologies have been increasingly incorporated into the stock, and traditional or vernacular architecture has given way to light–weight and mixed systems. More detailed information is given in the Building Permits, which describes up to three types of finishing materials per building element and their percentage of use. This information can be used to inform energy and mass transfer models.

3.3 Summary

In this chapter, a range of data sources are identified that give an understanding of the Chilean housing stock, especially about the parameters relevant to indoor air quality and energy demand.

The analysis highlights a paucity of information about categorical descriptors, such as dwelling ages and types, ownership, year of construction, and heating and cooking fuels. There is also insufficient granularity in physical, such as dwelling floor area and volume, window area and glazing type, air permeability, and number of floors, insulation level, and orientation. Very little is known about occupant behaviours, such as occupancy patterns, appliance use, cooking, indoor air temperatures and thermostat settings, or tobacco consumption.

There is always uncertainty in data. Therefore, when using this for modelling, a stochastic approach is required to capture both stock variability and parametric uncertainty. To do so, this study followed the method of the archetypes, derived from the data processing of the sources of information presented in this Chapter. The next chapter illustrates the process to develop the national archetypes as a generic tool that can be used for different studies, and Chapter 5 applies them to a physical model in order to predict the air quality of the Chilean housing stock.

The implications of the use of this information on the models is discussed in detail in Section 7.3.

CHAPTER 4

Representative Archetypes

It is quicker and cheaper to model the performance of a stock of dwellings than it is to measure it *in situ*. Furthermore, a model can be used to consider the consequences of future changes to the stock. However, the size and diversity of large stocks of dwellings often makes modelling individual buildings impractical, and so they can be grouped together by common factors into *archetypes* to make the dataset and the number of models more manageable. The unique properties of each dwelling are replaced by representative values when they are allocated to an archetype, which increases uncertainty in the stock model and makes tracking data sources important. The performance of the whole stock is then considered by extrapolation (Sousa *et al.* , 2017). The common factors used to create archetypes are determined by the model's outputs or performance indicators derived from them. The change in the size, composition, and characteristics of the stock over time requires archetypes that can be updated quickly and easily.

The information given in Chapter 3 is used here to apply bottom–up techniques to develop a set of nationally representative archetypal dwellings for Chile and to model them using building physics models (see Chapter 5). These archetypes are commonly developed using *clustering* techniques that group dwellings so that the houses in each *cell* or *cluster* are more similar to each other than those in other groups. For example, the English housing stock (a subset of the UK stock) of 22.3 million dwellings is represented by over 14,000 archetypes using a statistically representative cross sectional survey of the stock and a clustering method. Each archetype is weighted where the sum of all weights is the number of dwellings in the stock. It has been applied by seven independent models (Sousa *et al.* , 2017) to investigate energy related research questions, and recently reduced to just over 1000 archetypes by considering 8 parameters of interest and used to investigate housing stock decarbonisation strategies (Sousa *et al.* , 2018). Similarly, over 200 archetypes are used to represent 80% of the US housing stock (Persily *et al.* , 2006) by considering 10 parameters of interest, and 593 archetypes

have been developed for four European countries (Mata *et al.* , 2014). Further examples are given in Chapter 2. However, the location data for surveyed houses is currently too aggregated to apply *geographic information system* (GIS) techniques, such as those employed by Ghiassi & Mahdavi (2017).

In this Chapter, current information about the Chilean housing stock is analysed in order to develop and define a set of archetypal dwellings (see Sections 4.1 to 4.3) and their attributes to inform simulations and estimate uncertainty in contaminant exposures across the stock (see Section 4.4 onwards). Figure 4.1 shows how gathered *background* information is synthesised to inform model *inputs*. Inputs are defined using available *data*, and when data is either limited or non–existent, *assumptions* are defined probabilistically so that uncertainty in them is explored by a model.

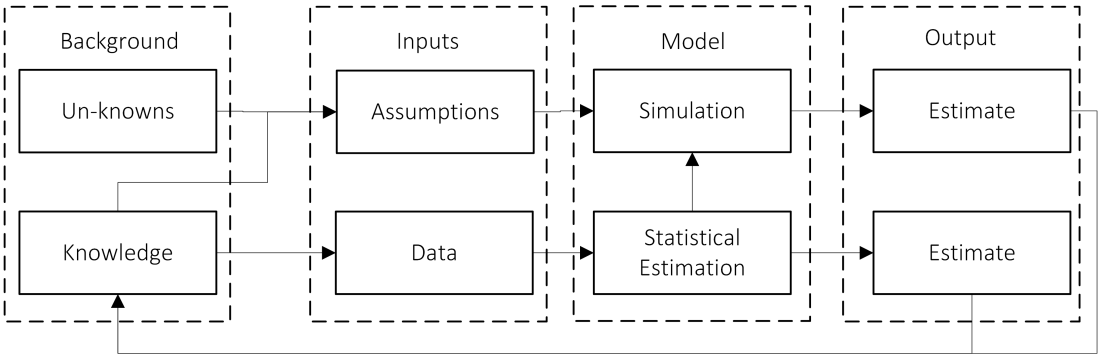


Figure 4.1: How the model used in this study combines the knowledge and estimates of the stock to represent and simulate the status quo.

4.1 Method

This section uses the terminology of Persily *et al.* (2006) and broadly follows the methods of Persily *et al.* (2006), Mata *et al.* (2014) and Shi *et al.* (2015). The majority of the data given in Chapter 3 are either discrete (number of rooms) or categorical (building type), and so a set of archetypes are developed by the aggregation of relevant characteristics, known as *factors*. Each factor may have a number of discrete levels (number of rooms, storeys) or may be continuous (total floor area). By allowing some factors to be vary within known limits, the uncertainty in the dataset can be assessed when simulating the models using an appropriate sampling method, such as Monte Carlo; see (Das *et al.* , 2014). Factors are chosen according to their influence on both indoor air pollution concentrations and the energy demand of a dwelling, and by their availability in the dataset. Predominant factor levels or values in the housing stock are selected to aggregate the data entries into *clusters*, or *cells*, where houses share common charac-

teristics. Some factor values occur more frequently than others in the dataset causing the formation of larger cells.

When selecting data sources, those considered reliable in Chapter 3 are given priority. For the purpose of this study, the census entries are selected by occupancy (occupied houses only), number of households per house (equal to 1), type of house (houses and flats/apartments) and maximum number of occupants and enclosed areas, arbitrarily set to 7 and 8, respectively. By recoding the variables they are easier to control and missing data can be removed, including houses with absent occupants, empty houses and cases with more than 1 home per building.

Data entries from the 2002 census and Building Permits were aggregated and used to develop the archetypes. Key descriptive data from the 2002 census data, dwelling geometry (2 levels: *apartment* or *house*), number of zones (8 levels), and number of bedrooms (levels 0–7 for houses and 0–6 for apartments, whereby 0 is a studio), were used to create 120 cells in a factorial manner, comprising $8 \times 8 = 64$ for houses and $8 \times 7 = 56$ for apartments. The same method is applied to Building Permits to determine their type (3 house levels: *detached*, *semi–detached*, *terraced*; 1 *apartment* level), construction period (2 levels: $\leq$2007 and $>$2007), and building size (8 levels: *number of zones*) were used to create a total of $4 \times 2 \times 8 = 64$ cells.

Following the aggregation of both databases, the census data is given priority because it is a more comprehensive data source (see Section 3.1.1). This is done by weighting the aggregated data by the dwelling type and construction period so that $(64 \times 3 \times 2) + (56 \times 1 \times 2) = 496$ cells are defined; see Figure 4.2. The analysis is performed using bespoke `R` Version 3.5.0 code (R Core Team, 2018) and the `aggregate` and `subset` functions.

The *building size* parameter is contained in both datasets and, although they each use different descriptive factors, it can be used to unify them. Here, the 2002 census gives the number of rooms whereas the Buildings Permit gives the *total floor area*. The *number of rooms* is used as an estimator of the building size; see Table 4.6 and Table A.2 in Appendix A for the values assigned to each number of rooms. Means and median values found within each cell are used to define levels. The *number of bedrooms*, *living rooms*, and *kitchen* were assigned according to the total number of rooms. The *number of floors*, *occupants* (from 1 to 5), and *bathrooms* were selected from the mode. Table 4.2 presents the variables assigned to the first 29 archetypes.

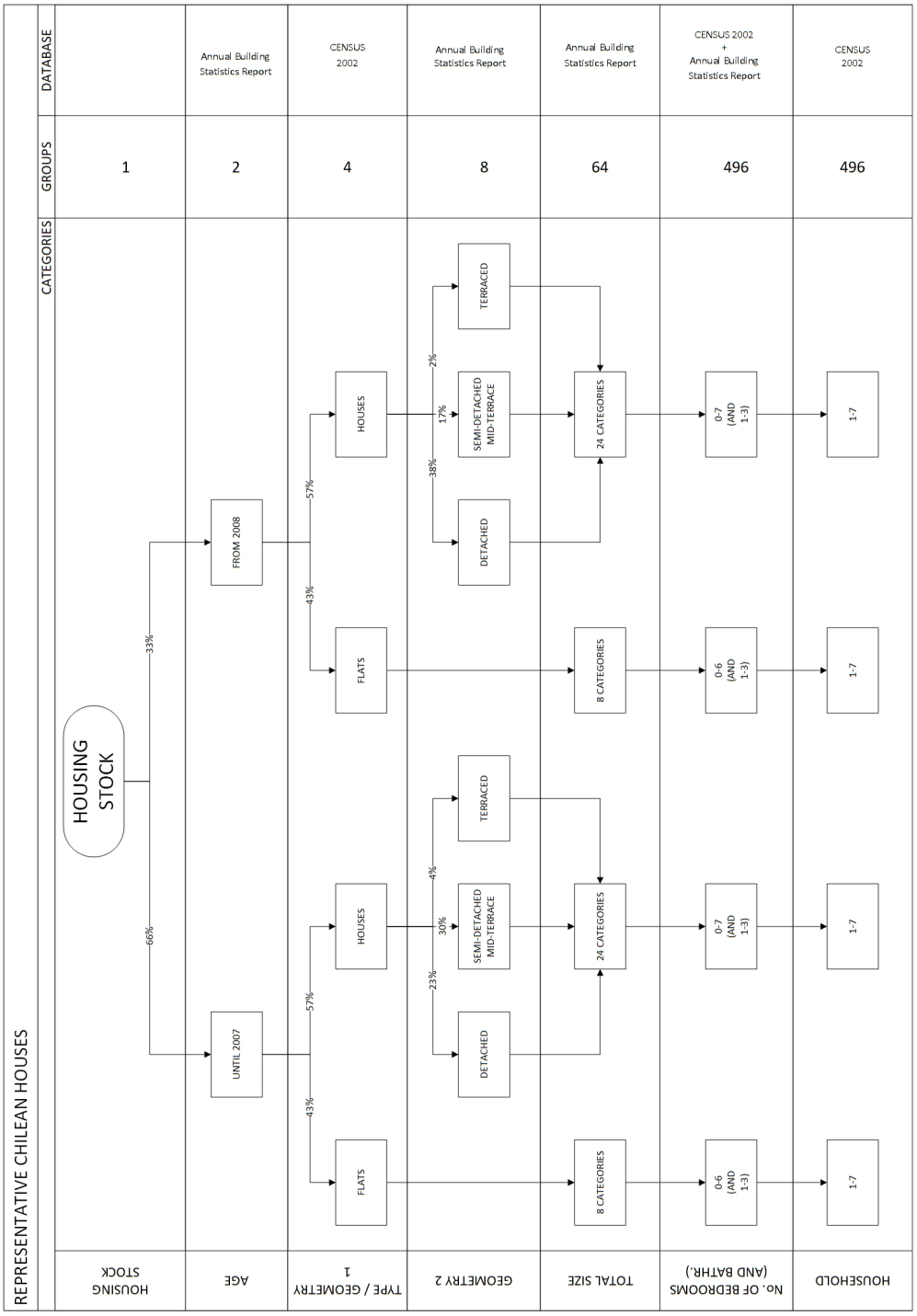


Figure 4.2: Hierarchy tree. The disaggregation of the Chilean stock into cells showing factors considered (left column) and source of information (right column).

The 496 cells can be used to classify the stock as dwelling archetypes and weighted using the number of dwellings in each cell (see figure 4.2). However, this is a significant number of archetypes, which has limitations on computational runtime. Sousa *et al.* (2018) use a high performance computer to simulate their 1,000 archetypes, whereas Persily *et al.* (2010) and Mata *et al.* (2014) only use a few hundred archetypes making simulation computationally less expensive. Therefore, it may be convenient to use fewer archetypes and acknowledge that the loss of resolution.

To estimate proportion of the stock represented by each archetype, Null Hypothesis Significance Testing is used. The differences in observed frequencies between two archetypes were compared using the chi–square (χ^2) test of statistical significance (Field, 2013). Emphasis is placed on the magnitude of the differences to represent a set of archetypes. This method allows each archetype to be removed from the whole set, by comparing the largest –the archetype with the highest observed frequency– to other archetypes in a descending order of frequency. When comparisons made against the largest archetype frequency were of a practically relevant effect size (as denoted by classified thresholds), all comparisons made prior were used to represent the overall stock. Hence, less importance was placed on the statistical significance. The effect size, ϕ, is given by (Kim, 2017)

$$\phi = \sqrt{\frac{\chi^2}{N}} \tag{4.1.1}$$

where N is the total number of observations considered for each statistical comparison. Thresholds can be used to interpret the magnitude of the effect size and are given by Ferguson (2009) as *small* ($\phi \leq 0.2$), *moderate* ($\phi \leq 0.5$), or *strong* ($\phi \leq 0.8$). Effect sizes of $\phi < 0.2$ are considered to be negligible (meaning no practically relevant difference), and greater than 0.5 are considered to be relevant to this study. Each comparison with the (χ^2) test does not require the testing of the same hypothesis or data repeatedly, and so Bonferroni corrections were not applied to control the Type I error rate across multiple analyses (Cabin & Mitchell, 2000).

4.2 Definition and selection of the archetypes

Figure 4.3 shows the number of archetypes required to represent each centile of the building stock and the number of required to meet each effect size classification. The archetypes are ranked by their observed frequency in descending order, beginning with the most common archetype, which is used as a reference for comparison. Figure 4.3 shows that a set of 496 archetypes can be used to represent the entire national housing stock and 90 archetypes represent 95% of the stock (marked by the blue data

point), indicating a large number of small outlier archetypes. A χ^2 test is used to compare the change in effect size between the most common archetype and each lower ranked archetype. Figure 4.3 shows that sets of 2, 8, and 29 archetypes, represent 13%, 35% and 70% of the entire stock, respectively (data points in orange), and correspond to the effect size classifications defined in Section 4.1 of small, moderate, and strong; see Table 4.1. The gradient of the line shows a law of diminishing returns as archetypes with ever decreasing weights are added because they do not significantly increase the proportion of the building stock represented. Therefore, Figure 4.3 can be used to choose a number of archetypes that balances the proportion of the stock represented with the time taken to create and simulate models.

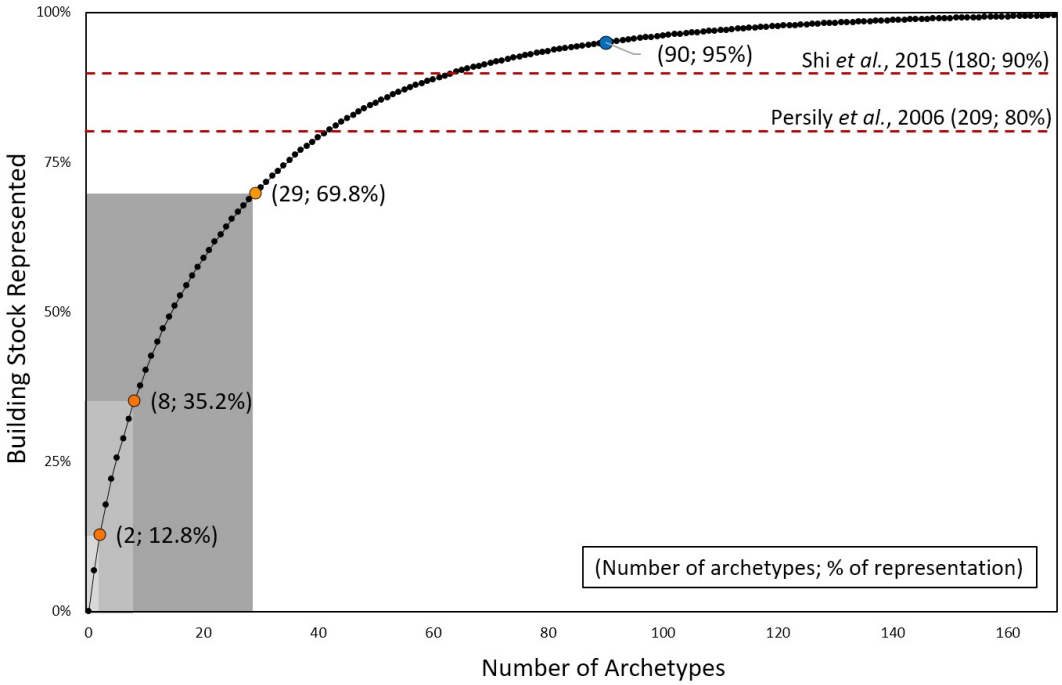


Figure 4.3: Cumulative frequency distribution of percentage of the housing stock represented by archetypes. Orange points, effect size thresholds; Blue point, 90th centile.

Table 4.1: Classification of archetype effect sizes. An effect size of moderate or greater than 0.5 was considered relevant to this study.

Ferguson's Effect Size thresholds	**Number of archetypes retained**	**Percentage of the stock represented**
Small ($\phi \leq 0.20$)	2	12.80%
Moderate ($\phi \leq 0.50$)	8	35.20%
Strong ($\phi \leq 0.80$)	29	69.80%

4.3 Description of archetypes

This section presents the model's input parameters needed to inform the archetypes for simulating the housing stock. Data inputs are classified into four sections: Section 4.4 presents parameters related to the environment and the location of the houses. Section 4.5 presents the physical parameters, such as the envelope air permeability and airflow elements. Section 4.6 gives the parameters related to the occupancy of the houses, and Section 4.7 those to the indoor pollutants.

Table 4.2 shows key factors and levels for the first 29 archetypes, which represent a strong effect size. This group comprises 15 single–story detached houses and 14 semi–detached houses (6 are two–story) with floor areas of 30–139m^2. Further information is given in Table A.1 in Appendix A. It also shows that the archetype with the highest weighting represents a 6.8% of the stock. It is a detached single–storey uninsulated house, constructed from clay bricks when located in central cities of the country, with prefabricated panels in northern cities, and with wooden panels when in southern cities. It has two bedrooms, two bathrooms and a separate kitchen. Its heating and cooking fuel is generally gas. Its mean household is 3.4 persons whose average socioeconomic status ranges from the 2nd decile (by population) in the centre–south of Chile to the 7th decile in the capital region.

Each archetype is further described by the predominant structural materials used in walls, roofs and flooring, cooking fuel and socioeconomic status, variables that were assigned according to the observed frequency in each region, and so the re–diversification of each of them over the stock can be used in the simulations. Data is presented in Appendix A, Table A.1.

Tables 4.2 and A.1 show that there is considerable variability in the parameters that describe each archetype. Accordingly, they should not be simulated deterministically when each archetype is modelled, simulated once, and the output scaled by its weight. Instead, the uncertainty in each parameter can be considered by varying them between known limits and running multiple simulations to give a range of outputs, following Das *et al.* (2014), Jones *et al.* (2015) and Sousa *et al.* (2018). Hence, the following sections present the input data used for describing and modelling the archetypes, and a method of obtaining the distributions is given. Inputs are presented depending on whether they are related to the environment (in Section 4.4), to the physical properties of the houses (Section 4.5), to the occupancy of the houses (Section 4.6), or to the pollutants (see Section 4.7).

Table 4.2: Characteristics of the 29 top–ranked archetypes.

ID Number	Percentage of Representation	Weighting Factor	Construction Period	Type of dwelling	Number of Zones	Number of Bedrooms	Floor area [m2] (**)	Number of storeys (*)	Number of bathrooms (**)	Number of occupants (**)
27	6.8	395,957	1	D	4	2	53	1	1.6	3.4
36	5.9	345,019	1	D	5	3	79	1	1.6	4.1
91	5.0	290,569	1	S	4	2	51	1	1.6	3.4
100	4.4	253,189	1	S	5	3	65	2	1.6	4.1
275	3.5	201,567	2	D	4	2	64	1	1.6	3.4
35	3.3	189,834	1	D	5	2	79	1	1.6	3.4
19	3.3	189,033	1	D	3	2	42	1	1.6	3.4
284	3.0	175,637	2	D	5	3	103	1	1.6	4.1
339	2.5	147,918	2	S	4	2	60	1	1.6	3.4
44	2.5	146,807	1	D	6	3	112	1	1.6	4.1
99	2.4	139,308	1	S	5	2	65	2	1.6	3.4
83	2.4	138,720	1	S	3	2	40	1	1.6	3.4
348	2.2	128,889	2	S	5	3	81	2	1.6	4.1
108	1.9	107,733	1	S	6	3	75	2	1.6	4.1
18	1.9	107,386	1	D	3	1	42	1	2.1	2.4
10	1.8	102,290	1	D	2	1	30	1	2.1	2.4
283	1.7	96,637	2	D	5	2	103	1	1.6	3.4
267	1.7	96,230	2	D	3	2	54	1	1.6	3.4
28	1.5	88,401	1	D	4	3	53	1	1.6	4.1
26	1.4	79,864	1	D	4	1	53	1	2.1	2.4
82	1.4	78,804	1	S	3	1	40	1	2.1	2.4
45	1.4	78,500	1	D	6	4	112	1	1.8	4.7
74	1.3	75,064	1	S	2	1	30	1	2.1	2.4
292	1.3	74,734	2	D	6	3	139	1	1.6	4.1
347	1.2	70,917	2	S	5	2	81	2	1.6	3.4
331	1.2	70,617	2	S	3	2	54	1	1.6	3.4
92	1.1	64,873	1	S	4	3	51	1	1.6	4.1
90	1.0	58,607	1	S	4	1	51	1	2.1	2.4
109	1.0	57,606	1	S	6	4	75	2	1.8	4.7

(*) Mode (**) Mean (***) median D: Detached; A: Apartment; S/E: Semidetached or End–terrace; M: Mid–terrace. Construction period: 1, $\leq$2007; 2, >2007.

4.4 Environmental inputs

Chile is geographically and administratively divided into 15 regions. Given the diversity in the conditions and characteristics of their stocks, the use of the regional classification of the country is a practical and simple way of dividing the country for data analysis, as well as reducing simulation time and achieving some granularity in the results. It should be noted that many of the data presented in Chapter 3 are presented in finer granularity, and so the archetypes could be developed using that data in the future. However, other information, such as weather data, or the surrounding conditions (shielding and terrain) of the houses are not available, and so major assumptions are required now, or additional data gathering is be required in the future, to achieve this additional granularity. Ways of handling these types of uncertainties are presented when describing the inputs in the following sections or when modelling the building in Chapter 5.

4.4.1 Geographic location and terrain type

Census data (INE, 2002) gives a location of each house. The observed frequency of the archetypes by region, given in Chapter A and Table A.1, is used to scale up the simulation results to a national level.

Coordinates (latitude and longitude), altitude, and the main climatic zone for each regional capital city is obtained. For simplicity, the location of every house modelled within each region is assumed to correspond to its capital city. The altitude of the house is used to calculate the local atmospheric pressure, as described by Liddament & Air Infiltration, Agence internationale de l'énergie (1996), and the air density used for air mass flows.

By knowing the number of houses located in rural or urban areas (INE, 2003), the discrete probability of each shielding condition, by region, is assumed to be:

- Country with scattered wind breaks: 20% of the rural stock of the region.
- Rural: 80% of the rural stock of the region.
- Suburban: 33% of the urban stock of the region.
- Urban: 33% of the urban stock of the region.
- City: 33% of the urban stock of the region.

The main assumption is made for rural areas (20:80), whereas for urban areas, terrain types are considered as equally distributed (33:33:33). Table 4.3 shows the regional distribution disaggregated by terrain type. This input data is deemed to be a source of epistemic uncertainty and may need to be revised in case this parameter is shown

to significantly affect the results by a sensitivity analysis; see Chapter 6.

Once a geographical location is assigned to each house, terrain data is used to scale the local wind speed using a *wind speed modifier*, which is a function of roof height and the local terrain type, following BS5925 (1991) and the LBL Infiltration Model, described by Orme *et al.* (1994). The four BSI terrain types and the local wind pressure shielding coefficients of Deru & Burns (2003) are mapped to the six BSI terrain types with the format BSI{Deru & Burns} as shown in Table 4.4. The values of the terrain and shielding coefficients (K and a, respectively) depend on the local environment of the house following Equation 4.4.1, which is allocated to each model according to the probability of it being located in each shielding type.

Table 4.3: Terrain types and their regional distributions. Based on INE (2002).

	Terrain Type {Shielding type}					
Region	Open flat country {None}	Country with scattered wind breaks {Light}	Rural {Moderate}	Suburban {Heavy}	Urban {Heavy}	City {Very heavy}
1	0	0.006	0.026	0.234	0.234	0.234
2	0	0.003	0.012	0.282	0.282	0.282
3	0	0.013	0.052	0.278	0.278	0.278
4	0	0.039	0.158	0.238	0.238	0.238
5	0	0.015	0.059	0.243	0.243	0.243
6	0	0.052	0.208	0.196	0.196	0.196
7	0	0.061	0.244	0.199	0.199	0.199
8	0	0.034	0.137	0.236	0.236	0.236
9	0	0.062	0.248	0.2	0.2	0.2
10	0	0.063	0.253	0.199	0.199	0.199
11	0	0.042	0.167	0.238	0.238	0.238
12	0	0.011	0.043	0.298	0.298	0.298
13	0	0.005	0.022	0.238	0.238	0.238
14	0	0.061	0.244	0.199	0.199	0.199
15	0	0.008	0.033	0.271	0.271	0.271

Table 4.4: Terrain types and the wind pressure coefficient scaling factors. Constants K and a are assigned to each house to scale the wind speed following equation 4.4.1. Sources: BS5925 (1991); Deru & Burns (2003)

	Terrain Type {Shielding type}					
Region	Open flat country {None}	Country with scattered wind breaks {Light}	Rural {Moderate}	Suburban {Heavy}	Urban {Heavy}	City {Very heavy}
Coefficient (K)	0.68	0.52	0.35	0.35	0.35	0.21
Exponent (a)	0.17	0.2	0.25	0.25	0.25	0.33

4.4.2 Block aspect ratio and orientation

The block aspect ratio S is the quotient of the length L and width W of the house or building block, assuming the houses have a rectangular shape. This value is required for calculating the wind pressure profiles for each side of the building block; see Section 4.4.3.

The orientation of Chilean houses is a source of epistemic uncertainty. While some assumptions can be made from urban planning and block layouts, they are limited. Therefore, the orientation of each archetype is a uniform distribution with values between 0° and 360°, following (Das *et al.* , 2014; Jones *et al.* , 2015).

4.4.3 Wind speed and wind pressure profile

The wind speed is obtained from the nearest meteorological station, u_m, is normally reported at 10 m height, and is scaled to the house roof height according as a function of the terrain type using a the power law model (Liddament & Air Infiltration, Agence internationale de l'énergie, 1996):

$$v_r = u_m \cdot K \cdot r^a \tag{4.4.1}$$

Here, v_r (m/s) is the wind speed at roof height, u_m (m/s) is the wind speed at the meteorological station, r (m) is the roof height of the house, and K and a are terrain coefficients; see Table 4.4.

The movement of the wind over and around a building is an important parameter that can affect air infiltration and exfiltration rates. The wind pressure profile around a building is a function of the geometry of the block, local shielding, and the wind's angle of incidence.

Two types of wall are considered; one for the long side and one for the short side of the house. For calculating the wind pressures profiles of each wall, the Swami & Chandra (1987) model is applied following Jones *et al.* (2015). The Swami & Chandra (1987) wind pressure coefficients are a function of the block aspect ratio (S) and the terrain constants given in Table 4.4. The model is used to calculate the wind pressure coefficients C_P for a set of wind angles over all walls. The application of this model is described in more detail in Chapter 5, Section 5.2.2.

4.4.4 Weather inputs data

Hourly weather data were obtained for Chilean locations to run transient simulations. 35 weather files were extracted using Meteonorm software (see Section 3.1.7) (Meteonorm, 2017). For this study, data files from periods between 1961–1990 and 2000–

2009 were processed. Finally, the nearest weather station to the location of each city centre and a random period is used to allocate the corresponding weather file.

4.5 Physical properties of the dwellings

Elements presented in this section are selected to be used in simulating the Chilean housing stock. They are based on the literature, on empirical data data, or from information of elements available in the Chilean market.

4.5.1 Airflow elements

The indoor spaces in Chilean houses are predominantly naturally ventilated and so the dominant drivers of ventilation are the wind and temperature differences between zones and the ambient environment. Airflow elements include windows and doors, air leakage paths, and mechanical extractor fans in some cases. These fans are located in the kitchen and bathrooms, and assumed from information on the local market to have a constant flow rate of 48 l/s (200 cfm) and 14 l/s (30 cfm), respectively, and operate when cooking or showering. These are manufacturers claimed flows –not what might actually happen in a real installation. Window and door geometries were obtained from personal correspondence with Chilean suppliers. External doors are 0.9 m x 2 m (width by height), with a gap under the door of 9 mm x 900 mm. Internal doors are 0.75 m x 2 m with a gap under the door of 2 mm x 750 mm. Windows are all assumed to be of sash type with cross–sectional areas of 0.36 m^2, 0.84 m^2, or 2.4 m^2, and are placed in bathrooms, kitchens or bedrooms, and living rooms, respectively. This dimensions meet the current construction code that governs the glazed proportion of the envelope (MINVU, 2007). The only vent considered by the models is located in the kitchen, an opening for combustion gases. The cross–sectional area of the opening is 100 cm^2, which complies with the current standard, DS N^{o}66 (SEC, 2007), for rooms with a cooker range hood.

4.5.2 Envelope air permeability

This section describes the air permeability of the envelope in the Chilean housing stock using several explanatory variables that have a significant impact on air leakage rates (Chan *et al.* , 2013). There are important differences in construction practices in houses built before and after 2008; see Section 3.1.6. Thus, both groups are expected to perform differently. Therefore, two different models were used to develop distributions of air permeability; one for each construction period.

The US multivariate regression model is used to predict a distribution of Normalised Leakage (NL) for Chilean houses built either before or after 2008, using the empirical data presented in Section 3.1.6 and following the procedure used by Chan *et al.* (2013). Here, the natural log of NL is given by

$$ln(NL) = \beta_{area} \cdot Area + \beta_h \cdot H + \beta_{year} \cdot I_{year} + \beta_{LI} \cdot I_{LI} + \beta_e \cdot I_e + \beta_{cz} \cdot I_{cz} + \beta_{floor} \cdot I_{floor} \tag{4.5.1}$$

where $Area$ is the house floor area (m^2), H is the house height (m), I_{year} is the house construction year category, I_{LI}, and I_e is the energy performance corresponding to low–income (LI) and energy efficient houses respectively, I_{cz} is the climate zone, and I_{floor} is the air leakage of the house floor.

Chilean data is used when the distribution of the input in the stock is known, such as floor area and climate zone. Coefficients are assumed to have the value that best represent the stock, with a value of 1 if true or 0 if not. In order to find the β_{year} coefficient that best fit the empirical data, all estimates of β shown in Eq. 4.5.1 are kept unchanged except for β_{year}; see Table 4.5.

Data on floor area is only available for houses built after 1990 (INE, 2003). Thus, distribution of floor areas for the Chilean stock is estimated using the number of rooms and their floor areas shown in Table 4.6, and extrapolated to the whole stock using the 2002 census. The mean and median floor areas are 81 m^2 and 79 m^2. Here, only the median is retained for the sampling process, as the best measure of central tendency. Building height is assumed to be 3 m (2.5 m + 0.5 m for roof space), following Chan *et al.* (2013).

Due to the differences in construction practices and standards between Chile and the USA, and the lack of information on the number of energy efficient houses in the stock, I_{LI} and I_e are assumed to be 1 and 0, respectively, namely all considered to be low–income USA houses. β_{floor} for all houses are assumed to have a concrete slab due to the lack of reliable data.

Climate was one of the most influential parameter in the Chan model. To include the climate factor here, the International Energy Conservation Code (IECC) classification is used to match the Köppen classification for Chilean and the USA climate zones and to associate a β_{cz} with each region; see Table 4.7. The results of β (SE) for houses built before and after 2008 are 3.480 (0.719) and 1.469 (0.845), respectively.

To obtain a national distribution, floor areas for each climate zone are sampled according from their distributions across the stock using Monte Carlo sampling. Since the accuracy of the prediction improves with the sample size, the sample is increased until the absolute differences of the mean and standard deviation between one set of

Table 4.5: Model coefficients used for predicting the air leakage distribution of the Chilean dwelling stock. Climates follow the IECC classification.

Parameter	**Coefficient Estimate β**	β Std. Error
Floor Area	-0.00208	0.0000179
Height	0.06380	0.00125
Year		
<2008	3.4898366	0.7190268
≥2008	1.469268284	0.844718303
Energy performance		
Low–income (LI)	0.42	0.00428
Energy efficient	-0.3840	0.00453
Climatic zone		
A_Humid_1_2	0.47300	0.01020
A_Humid_3	0.25300	0.00653
A_Humid_4	0.3260	0.00586
A_Humid_5	0.11200	0.00551
A_Humid_6_7	0.00	0.00
B_Dry_2_3	-0.03760	0.00759
B_Dry_4_5	-0.00877	0.00684
B_Dry_6	0.01940	0.00988
C_Marine_3	0.04830	0.01410
C_Marine_4	0.25800	0.01130
AK_7	0.02560	0.00589
AK_8	-0.51200	0.00938
Floor		
Foundation Slab	-0.036992	0.007092
Unvented crawlspace	0.108713	0.004923
Vented crawlspace	0.180352	0.005768
Duct – Conditioned Space	-0.123810	0.025460
Duct – Attic or Basement (Unconditioned)	0.07126	0.03387
Duct–Vented crawlspace	0.18072	0.03826

Table 4.6: Distribution of floor areas by number of rooms. Only houses with up to 7 rooms were included, which represent 95% of the stock (INE, 2003).

Rooms	**Floor Area (m^2)** INE (2016a)	**Proportion (%)** INE (2003)
1	20	1.1
2	35	6.5
3	49	14.1
4	62	26.1
5	93	27.8
6	124	13.0
7	154	5.9

samples and the previous set is less than 1e−6. To compare two bands of construction year, two different data sets are used; one for *old* houses and one for *new*. The model predicts NL 95% CI $[9.91 - 106.59]$ for old houses and 95% CI $[1.39 - 15.90]$ for new houses. Figure 4.4 shows both distributions.

To generate separate distributions of NL for each climate zone, the same method is conducted by sampling random values from the normal distribution of each β coefficient shown in Table 4.5. Finally, the predicted regional cumulative probability distributions are generated to be used in the models representing either the old (Figure 4.5a) or the new stock of houses (Figure 4.5b).

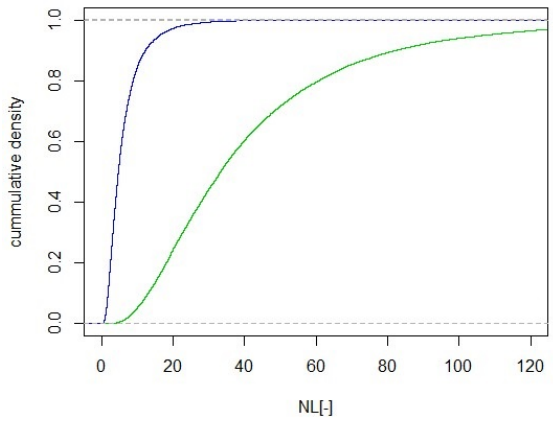
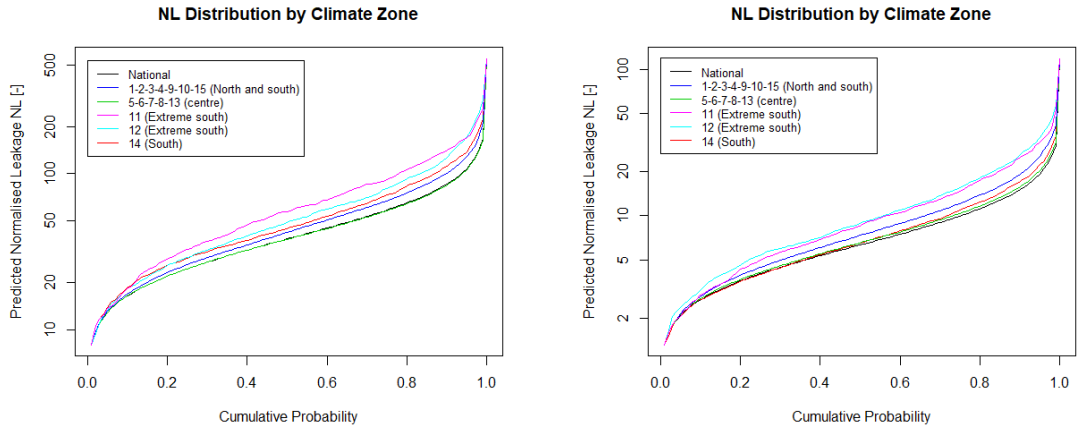


Figure 4.4: National cumulative density of NL for old (green line) and new houses (blue line).

Finally, a comparison between the NL distributions predicted by the model and the empirical data presented in Section 3.1.6 is carried out to evaluate the performance of the model; see Figures 4.6a and 4.6b. For old houses, linear regression between the measured and predicted indicates a *strong* or *high* correlation (coefficient of determi-

Table 4.7: Chilean climatic zones according to Köppen classification and matching US IECC climatic zones. Sources: Chan *et al.* (2013); INE (2003); Rioseco & Tesser (2014).

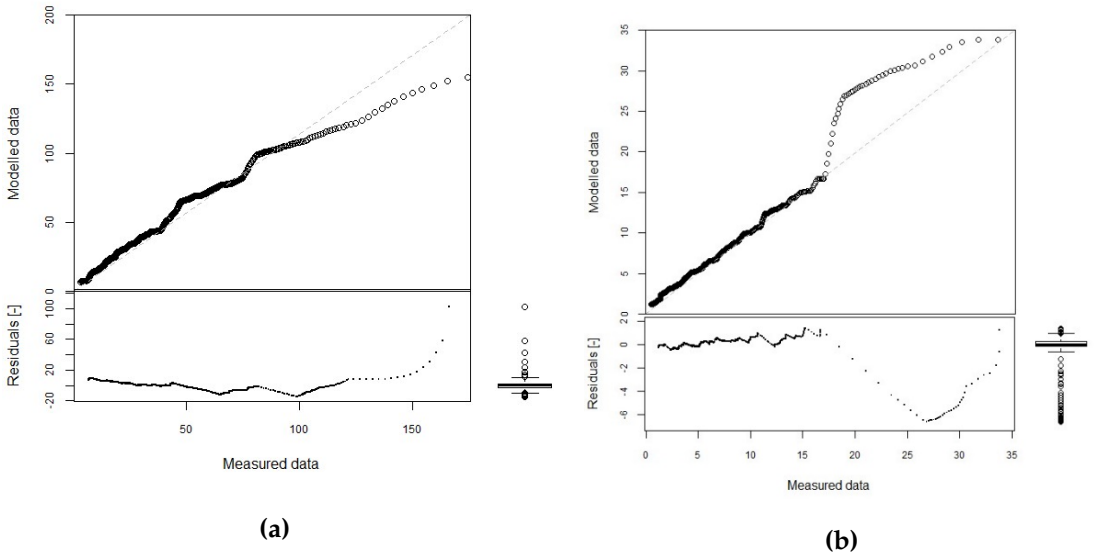
Region	Region name	Capital city	Köppen nomenclature	Köppen Classification	USA classification IECC	Proportion in the national stock
1	Tarapacá	Iquique	BWn	Arid/Desert/foggy	B5	1.338589
2	Antofagasta	Antofagasta	BWn	Arid/Desert/foggy	B5	3.949158
3	Atacama	Copiapó	BWk	Arid/Desert/cold	B5	2.414491
4	Coquimbo	La Serena	BWn	Arid/Desert/foggy	B5	4.176639
5	Valparaíso	Valparaíso	Csbn	Temperate/DrySummer/Warm summer/ Rain in winter/cloudy	B3	4.969534
6	O'Higgins	Rancagua	Csb	Temperate/DrySummer/Warm summer/ Rain in winter	B3	5.158549
7	Maule	Talca	Csb	Temperate/Dry Summer/Warm summer/ Rain in winter	B3	6.403527
8	Bio–Bío	Concepcion	Csbn's	Temperate/Dry Summer/Warm summer/ Rain in winter	B3	11.01014
9	Araucanía	Temuco	Cfb	Temperate/Without dry season/Warm summer/ Rain all year around	B4	6.310351
10	Los Lagos	PuertoMontt	Cfbs	Temperate/Dry Summer/Warm summer	B4	4.930199
11	Aysén	Coyhaique	Cfc	Temperate/Dry Summer/Cold summer/ Cold	C4	0.832992
12	Magallanes	Punta Arenas	BSk's	Semi–Arid/Steppe/Very Cold/ Rain in winter	B6	1.444968
13	Santiago	Santiago	Csb	Temperate/Dry Summer/Warm summer/ Rain in winter	B3	33.82966
14	Los Ríos	Valdivia	Cfb	Temperate/Without dry season/Warm summer/ Rain all year around	A4	2.529508
15	Arica y Parinacota	Arica	BWn	Arid and cloudy	B5	1.312891



(a) **(b)**

Figure 4.5: Predicted normalised leakage (NL) (n.d.) distribution for (a) old and (b) new Chilean houses by climate zone, and nationwide in black.

nation R^2 of .62 and a correlation coefficient R of .79 (Taylor, 1990), 95% CI $[.70 - .85]$) meaning that 62.0% of the variance in NL in the housing stock is described by the model and its explanatory variables. Similarly NL predictions for new houses have R^2= of .57 and a R of .75, 95% CI $[.66 - .83]$, indicating that 57% of the variance in NL values in the new stock can be described by the model.



(a) **(b)**

Figure 4.6: Predicted normalised leakage (NL) (n.d.) distribution versus empirical data for **(a)** old and **(b)** new Chilean houses. Boxplots show the residuals, with a $\tilde{\mu} = -0.42$ and $\sigma = 6.54$ for old houses, and $\tilde{\mu} = -0.086$ and $\sigma = 1.14$ for new houses.

4.6 Occupancy and activity data

Data from the 2015 ENUT survey (see Section 3.1.4) are used to inform the models about the duration of the emissions and activities in Chilean houses. Special attention is given to those related to heating and cooking activities that will influence the indoor air quality and exposure analysis. For example, questions *o*_11_1_2 and *o*_11_2_2 ask about the time spent preparing and heating meals. These questions are used to inform the schedule of $PM_{2.5}$ and moisture emissions in the kitchen. Nationwide, cooking activities have a mean duration of 1 h 06 min in weekdays and 1 h 12 min in weekends. Sleeping duration are reported in questions *q*11_1_2 and *q*11_2_2 for weekdays and weekends, respectively. Table 4.8 shows the average sleeping time aggregated by region and for the entire country.

Table 4.8: Sleeping time by region number and nationwide. (INE, 2016b).

Region Number	**Sleeping hours Weekday**		**Sleeping hours Weekend**	
	Hours	Minutes	Hours	Minutes
1	7	11	7	48
2	7	5	7	47
3	7	18	7	32
4	7	12	7	43
5	7	22	7	49
6	7	19	8	00
7	7	49	8	26
8	7	19	7	54
9	7	29	8	02
10	7	49	8	26
11	7	43	8	37
12	7	3	7	40
13	7	10	7	50
14	7	26	7	55
15	7	18	8	01
Nationwide	7	21	7	57
Associated schedule				
Awake time weekdays	6	21		
Sleeping start time weekdays	23	0		
Awake time weekends			7	57
Sleeping start time weekends			24	0

Tables B.1 to B.3 in Appendix B gives the duration statistics for some other activities aggregated by region, and Table 4.9 summarises them for the country. Table 4.9

Table 4.9: Average time corresponds to the mean time between the activity duration in weekdays (WD) and in weekends (WE). The proportions are rounded so that their sum is 100% (INE, 2016b).

Room	Activity	WD		WE		Average Time (h)	Time in the Room (h)	Percentage (%)
		H	M	H	M			
Bedroom	Sleeping hours	7	21	7	57	7.65	9.26	38
	Showering / getting dressed	0	36	0	37	0.61		
Kitchen	Preparing meals	1	6	1	12	1.15	2.33	10
	Washing up	0	19	0	22	0.34		
	Cleaning the kitchen	0	28	0	31	0.49		
	Eating Breakfast	0	19	0	23	0.35		
Family	Eating Lunch	0	34	0	42	0.63	12.4	52
room	Eating Tea	0	30	0	33	0.53		
	Eating Dinner	0	33	0	36	0.58		

shows that some activities can be related to a room, such as *cooking* to the kitchen or *sleeping* to the bedroom, so that the total time spent in each room can be calculated. Using the activities presented in Table 4.9 the ratio of the time a family spends time in the kitchen, bedroom, and family room is $10 : 38 : 52$, respectively. This ratio can be used to inform the method used to analyse modelling or measurement studies; see Hamilton *et al.* (2015). A different approach is used in this study, by making further assumptions about the occupancy pattern in rooms and, consequently, about the exposure of occupants to indoor pollutants. Further details are presented in Section 5.5.2.

4.7 Pollutant inputs

This section describes the pollutants selected to be modelled, including the physical parameters of each species, and their emissions sources and sinks.

4.7.1 Species

Nine species are included in the archetypes:

- Fine Particles, $PM_{2.5}$
- Environmental Tobacco Smoke – Solanesol, ETS
- Formaldehyde, HCHO
- Acrolein, C_3H_4O
- Nitrogen dioxide, NO_2
- Carbon monoxide, CO
- Nitrogen monoxide, NO
- Sulfur dioxide, SO_2
- Moisture as water vapour, H_2O

The first five contaminants are identified by Logue *et al.* (2011a) within the top seven most harmful air pollutants found in the US dwellings. Although Logue *et al.* also include radon and ozone, they are excluded from this analysis.

Radon (^{222}Ra) sources comprises the building substratum, tap water, and building materials. According to the literature, see Chapter 2, the concentrations of radon in Chilean dwellings are found to be below the WHO's recommended threshold. A survey carried out in winter 1992 and 1993 in 119 houses in Santiago (Stuardo, 1996) found radon concentrations in the bedrooms that ranged between 4 to 86 Bq m^{-3} ($\tilde{\mu}$=25 Bq m^{-3}); below the WHO's current recommendation of a 100 Bq m^{-3} annual average. Furthermore, Zielinskia & Chambers (2008) published the concentrations of residential radon concentrations in several countries around the world, showing that Chile has some of the lowest indoor concentrations. Therefore, in this study, radon is not considered to be a contaminant of concern and is not modelled.

In contrast, ozone is a *criteria pollutant* found in high concentrations in Chile. It is emitted by many electric appliances in homes, enters from the outside, and reacts with other indoor pollutants. However, there is no information on expected ambient concentrations in Chile and because this has been shown to be such as important source in other countries, ozone is not modelled here. However, ozone emission could be added to the model in the future if required and when data becomes available.

The physical properties of each species are needed so that the modelling tool can simulate their transport between a building's rooms and with the local environment. Table 4.10 shows the main properties of each species considered in this study. Because different names can be used for the same chemical component, the CAS number is given to identify each species. Also given are their chemical formula, molecular weight, and the mean diameter for particles and ETS. ETS is a mixture of chemicals. Solanesol is an alcohol present in the tobacco leaf and has been used as a tracer for the particle phase of tobacco smoking emissions (Tang *et al.* , 1990), and acrolein is used as a tracer of its volatile organic compounds emissions (Howard-Reed *et al.* , 2003).

4.7.2 Deposition rates

It is important to model the potential loss of indoor particles due to their deposition onto, or their reaction with, indoor surfaces. Therefore, a probability distribution of deposition rates reported in the literature is included. The product loading ratio is the quotient between the surface area A and the air volume V, and may change according to the dimensions of the rooms and the surface area of its furniture. Here, only one loading ratio of $A/V = 2$ is included in accordance with the references. Table 4.11 shows the deposition rate and deposition velocity associated with each species used

Table 4.10: Species properties. n/a: not–applicable (Klepeis *et al.* , 2001; Tang *et al.* , 1990)

Name	CAS Number	Linear or Empirical formula	Molecular_Weight (g/mol)	Mean_Diameter (μm)
Acrolein	107-02-8	C3H40	56.1	0.00E+00
Formaldehyde	50-00-0	HCHO	30.0	0.00E+00
Carbon monoxide	630-08-0	CO	28.0	0.00E+00
Dry Air	n/a	n/a	29.0	0.00E+00
ETS-Solanesol from tobacco	13190-97-1	C45H74O	631.0	2.00E-07
Water vapour	7732-18-5	H2O	18.0	0.00E+00
Nitric oxide	10102-43-9	NO	30.0	0.00E+00
Nitrous oxide	10024-97-2	N2O	44.0	0.00E+00
Fine particles $PM_{2.5}$	n/a	n/a	1.4	3.00E-07

in this study.

Table 4.11: Deposition rates for species used in this study.

Species	**Deposition velocity** v_d	**Deposition Rate** v_g $v_d \cdot 4.10$	**Reference**
NO	0.108 (m/h)	0.000–0.0123 (s^{-1})	Miyazaki, T., 1984
NO_2	0.0003–0.12 $(cm \cdot s^{-1})$	0.00123–0.492 (s^{-1})	Miyazaki, T., 1984
HCHO	0.005 ± 0.003 $(cm \cdot s^{-1})$	0.0205$\pm$0.0123 (s^{-1})	(Traynor *et al.* , 1982)
NO_2	$(0.6–3.6)e-4$ $(cm \cdot s^{-1})$	Not used	(Nazaroff *et al.* , 1993)
$PM_{2.5}$	Not used	0.39 ± 0.16 (h^{-1})	(Oezkaynak *et al.* , 1996)
FA	Not used	0.0205$\pm$0.123 (s^{-1})	(Traynor *et al.* , 1982)

4.7.3 Emission rates from cooking

In order to model the uncertainty in the emission rates from cooking, a synthetic cumulative distribution function of emission rates is developed. Data is collected from four studies reporting $PM_{2.5}$ emission rate means and standard deviations, or giving the emission rate for an individual test (Dacunto *et al.* , 2013; He *et al.* , 2004; O'Leary, 2018; Olson & Burke, 2006).

To incorporate them into the occupants' daily activities, meals are classified into two groups: emissions from cooking activities other than *toasting bread* are classified as *main meals*, and *toasting bread* emission rates as *breakfast*. Cumulative distributions are inferred by bootstrap sampling until the mean and standard deviation converge. Emission rates are sampled randomly with replacement from their own normal distributions, truncated to positive numbers, with μ and σ shown in Table 4.12 and Figure 4.7. Each type of meal is assumed to occur with equal probability, regardless the sample size. The sample sizes were increased until the absolute change in its $\tilde{\mu}$ and σ was $< 10^{-6}$ mg/min.

Results for main meals ($N = 15,650$) show that their emission rates have a $\tilde{\mu}$ of 2.56 mg/min, a σ of 4.4 mg/min, and 90% CI $[0.047 - 15.2]$ mg/min; see Figure 4.8a. Results for breakfast ($N = 4,165$) show that their emission rates have a $\mu = 4.32$ mg/min, a $\sigma = 7.42$ mg/min, and 90 % CI $[0.072 - 21.77]$ mg/min; see Figure 4.8b.

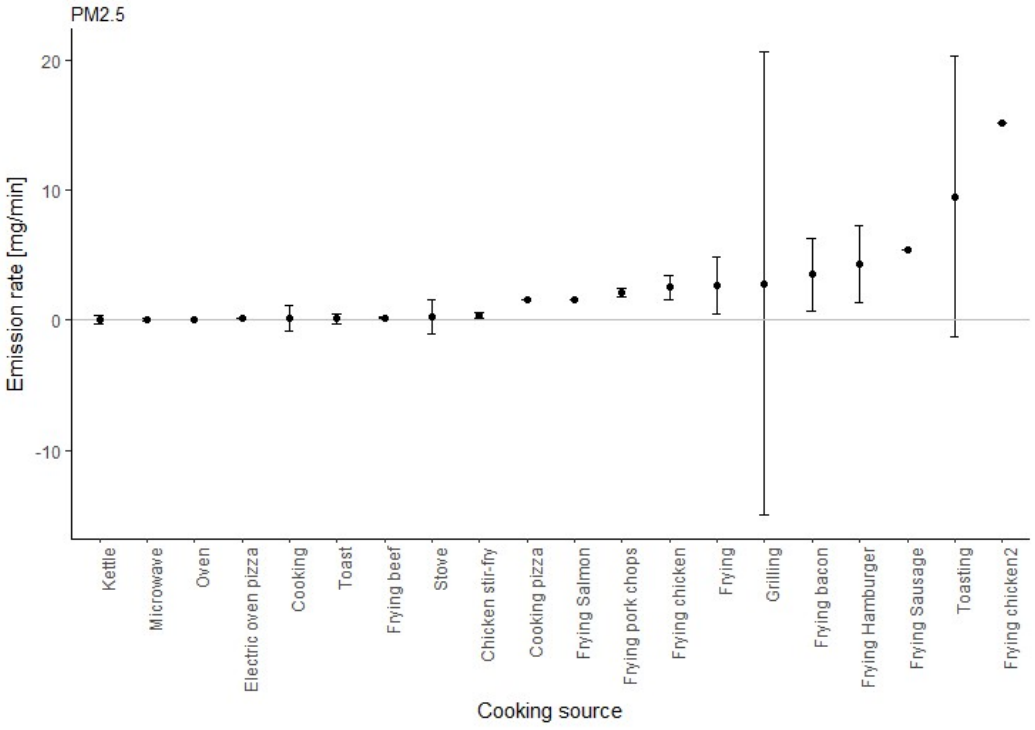


Figure 4.7: $PM_{2.5}$ emission rates from cooking activities. Based on data from Dacunto *et al.* (2013); He *et al.* (2004)

4.7.4 Emission rates from heaters

In Chile, there are six common types of heaters that burn gas, paraffin, and wood, whose pollutant emission rates to inside have been measured by CENMA (2011); see Table 4.13. Emissions to outside were not measured and not included in this analysis. Moreover, the prevalence of heating fuel varies across the country (see Section 3.2.5). Therefore, the probability of presence of each heater type can be allocated by region, whereby the emission rate is a constant determined from the fuel type and measurements. Accordingly, the proportion of the stock using each heater type is assumed from the heating fuel given in the CASEN survey, where the data is recorded as variable *v36b* (CASEN, 2015); see Table 4.14.

The total number of heating hours per day corresponds to those required to maintain an indoor temperature of at least 16°C. This is calculated using the Meteonorm weather files. Following the assumption that the indoor temperature in unheated houses is found to be approximately 3°C above the external temperature (Jones *et al.*

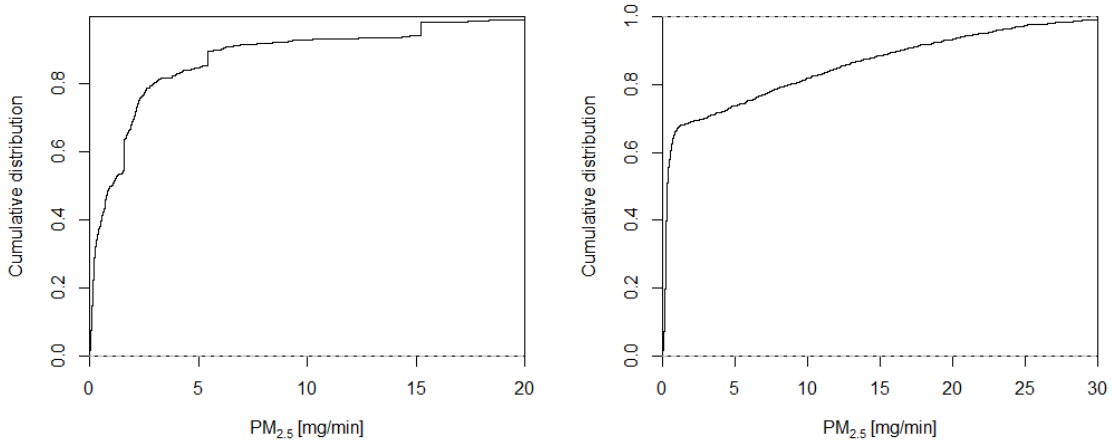
Table 4.12: $PM_{2.5}$ Emission rates from cooking and other activities. Sources: (1) He *et al.* (2004); (2) Dacunto *et al.* (2013); (3) O'Leary (2018).

Meal type	N	$\tilde{\mu}$ (mg/min)	σ (mg/min)	Reference	Statistics
Main meals					
Cooking	24	0.11	0.99	1	Median
Cooking pizza	1	1.59	0	1	Median
Frying	4	2.68	2.18	1	Median
Grilling	6	2.78	17.8	1	Median
Kettle	25	0.03	0.31	1	Median
Microwave	18	0.03	0.11	1	Median
Oven	6	0.03	0.03	1	Median
Stove	4	0.24	1.29	1	Median
Frying chicken	6	2.5	0.9	2	Mean
Frying bacon	6	3.5	2.8	2	Mean
Frying burger	4	4.3	3	2	Mean
Frying pork chops	2	2.1	0.3	2	Mean
Chicken stir fry	2	0.4	0.2	2	Mean
Frying beef	2	0.17	0.05	2	Mean
Frying chicken	1	15.2	0	2	Mean
Frying Sausage links	1	5.4	0	2	Mean
Frying Salmon	1	1.6	0	2	Mean
Electric oven pizza	1	0.1	0	2	Mean
Meal 1	6	0.1554	0.04	3	Mean
Meal 2	6	0.1734	0.08	3	Mean
Meal 3	6	0.612	0.30	3	Mean
Meal 4	6	2.046	0.24	3	Mean
Meal 1 – Margarine	5	0.714	0.05	3	Mean
Meal 1 – Stainless Steel	5	1.656	0.62	3	Mean
Meal 1 – Salt	5	0.2292	0.10	3	Mean
Breakfast					
Toasting	18	0.11	0.37	1	Median
Toasting	3	9.5	10.8	2	Mean
Toasting	40	0.22	0.06	3	Mean
Other sources					
Smoking	6	0.99	0.81	1	Mean
Smoking cigarette	17	3.8	0.9	2	Mean

, 2015), the average time during the winter season when the indoor temperature is below 16°C is calculated and shown in Table 4.15.

4.7.5 Moisture emission

Moisture emitted from the preparation of meals is 4 g/min and shower moisture is 43.33 g/min following TenWolde & Pilon (2007).



(a) **(b)**

Figure 4.8: Inferred cumulative distributions of $PM_{2.5}$ emission rates from cooking activities for modelling (a) main meals and (b) breakfast.

Table 4.13: Contaminants emission rate (mg/min) from Chilean heaters.ND: non–detectable

Tecnology	**Commercial Brand**	**PM_{10}**	**$PM_{2.5}$**	**NO**	**NO_2**	**NO_x**	**CO**	**SO_2**
Gas radiant	Fenza FEL 1430	ND	ND	ND	1.20	1.20	24.44	ND
Gas forced convection	Rinnai Dynamo 15	ND	ND	0.16	0.65	0.81	8.02	ND
Kerosene standard wick	Mademsa Foguita	0.37	0.29	0.20	1.92	2.12	26.20	1.24
Kerosene special wick	Toyotomi Omni 230	0.19	0.14	5.47	1.61	7.08	5.32	2.52
Kerosene laser	Toyotomi LC–43	0.15	0.15	3.74	1.80	5.54	8.00	2.29
Wood double chamber	Amesti Doble Cámara	0.05	0.05	0.24	0.11	0.35	2.55	ND

Table 4.14: Heater types distribution factor across the country.
Regions are numbered from north to south, except for regions 13th (capital region), 14th, and 15th.

Region	**Heater type and name**						
	1	2	3	4	5	6	7
	Gas 1	Gas 2	Kerosene 1	Kerosene 2	Kerosene 3	Wood	None
1	0.054	0.054	0.012	0.012	0.012	0.010	0.842
2	0.076	0.076	0.003	0.003	0.003	0.011	0.827
3	0.094	0.094	0.005	0.005	0.005	0.016	0.780
4	0.109	0.109	0.006	0.006	0.006	0.053	0.711
5	0.212	0.212	0.017	0.017	0.017	0.239	0.279
6	0.091	0.091	0.049	0.049	0.049	0.600	0.057
7	0.074	0.074	0.029	0.029	0.029	0.717	0.038
8	0.053	0.053	0.020	0.020	0.020	0.784	0.043
9	0.025	0.025	0.007	0.007	0.007	0.908	0.017
10	0.014	0.014	0.006	0.006	0.006	0.945	0.009
11	0.011	0.011	0.004	0.004	0.004	0.961	0.003
12	0.469	0.469	0.001	0.001	0.001	0.059	0.001
13	0.279	0.279	0.083	0.083	0.083	0.049	0.116
14	0.015	0.015	0.006	0.006	0.006	0.940	0.011
15	0.023	0.023	0.001	0.001	0.001	0.007	0.944

Table 4.15: Estimated use of heaters during winter season in hours and minutes by region number.

Region	**Hours**	**Minutes**
1	3	24
2	9	48
3	11	54
4	14	54
5	15	6
6	15	48
7	16	18
8	16	24
9	17	0
10	17	0
11	17	0
12	17	0
13	15	36
14	17	0
15	2	18

4.8 Summary

The robustness of key data related to the Chilean housing stock is analysed and the most reliable sources chosen to develop the Chilean archetypes. There is a sufficient level of detail to allow the archetypes to be developed by applying a bottom–up technique and to model them using a physical model. In order to limit the computational time and to allow the assessment of the variability of the stock, sets of 2, 8, and 29 archetypes are selected to represent 13%, 35%, and 70% of the entire stock. Data entries are aggregated to create 496 archetypes. Archetypes are sorted by their observed frequency and compared using the chi–square (χ^2) test of statistical significance. Groups are chosen by their change in the effect size of χ^2 and using Cohen's ϕ thresholds of 0.2, 0.5, and 0.8, respectively.

In addition, some parameters needed for modelling the archetypes are described using the data presented in Chapter 3, while others must be assumed by using expert judgement, the literature, or observation. Inputs are classified according to their relationships with the house environment, to the house design, to the pollutants, or to the occupants.

Houses are allocated across the country and inputs about their physical surroundings are given. Properties of the building elements, such as window sizes and vents comply with the current national building code (MINVU, 2007). National data was applied to predict the air permeability of the envelope for the Chilean housing stock following the procedure by Chan *et al.* (2013). Two distributions are presented, one for each construction period. The model predicts a NL 90% CI $[9.91 - 106.59]$ for old houses and a NL 90% CI $[1.39 - 15.90]$ for new houses. 62.0% and 57% of the variance in NL of old and new houses, respectively, is described by the model and its explanatory variables.

Information on nine contaminants, some of their indoor sources, and deposition rates are presented so they can be applied in the models. The total mean time ratio the cook family member spends in the kitchen, bedroom, and family room is found to be $10 : 38 : 52$, respectively. Activity durations are given to model the occupancy of the rooms.

The archetypes developed here can also inform field studies by describing buildings that are predicted to represent a large proportion of the stock.

A detailed discussion about the development of this archetypes is presented in Section 7.2.

CHAPTER 5

Model Development

In Chapter 4 the most common archetypes in the Chilean housing stock were identified. In this chapter a validated IAQ and ventilation analysis tool, CONTAM, is applied to model the first 8 archetypes, which represent 35% of the national stock; see Figure 4.3. The models include the input data as presented in Chapter 4. The diversity of the parameters most directly related to house design are accounted for in the archetypes, and so they can be used to reduce the primary uncertainty in them. Others, which have primary (aleatory) or secondary (epistemic) uncertainty, must be determined differently.

Specific aspects of CONTAM, such as restrictions or boundaries, and modelling assumptions are presented here. Simulations are made under these conditions to enable the analysis of some aspects of indoor air quality across the Chilean housing stock, to compare housing characteristics, and to evaluate their the relative importance of model inputs. In the future, the models can be used, for example, to assess remediation actions or to evaluate impacts of human behaviour, clothing, and habits that are typical of the Chilean people in their houses.

5.1 CONTAM modelling and analysis tool

CONTAM (Dols & Polidoro, 2015) is a freely available multi–zone indoor ventilation and pollutant transport tool that models airflows between multiple indoor zones, and between the external environment. It has been validated by comparing its performance against other modelling tools (Dols *et al.* , 2016), against measurements in controlled environments (Chang & Guo, 1992), and against field studies (Das *et al.* , 2014). CONTAM has been used to model different types of building (Ng *et al.* , 2012) and for evaluating input parameters (Chen *et al.* , 2014) and pollutant concentrations (Bastani *et al.* , 2012; Underhill *et al.* , 2018; Yu *et al.* , 2015).

CONTAM comprises ContamW, a graphical user interface (GUI), and ContamX, a

numerical solver (Lorenzetti *et al.* , 2013). The GUI allows a user to sketch building elements and store them in a project file, which ContamX simulates. Any necessary changes to a modelled archetype and to its environment is carried out by modifying the project file. To do so, R code is used to modify a generic archetype project file, create a new file, and overwrite its information with a new sets of inputs before a model is simulated.

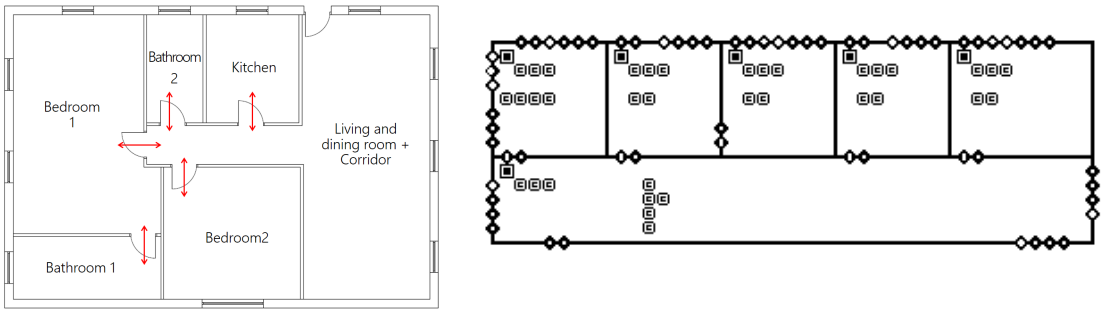
CONTAM is selected over other simulation tools because it can include multiple pollutants, multiple sources and sinks, and multiple emission and removal models. This enables the modeller to have a more accurate and detailed representation of the indoor air. In contrast, EnergyPlus only models a single and generic pollutant, which it seldom occurs in reality. CONTAM has 12 different emission and deposition models, and the possibility of adding an almost infinite number of species, which can all be used simultaneously. Furthermore, emission parameters provided by public databases, or by the literature, can be incorporated. Finally, CONTAM has the ability to couple its numeric solver to a CDF modelling method in one zone of the building. However, CONTAM lacks the capability of modelling building thermal dynamics, and so internal air temperatures can either be included in the models as input data or CONTAM must be coupled with a dynamic thermal model, such as EnergyPlus or TRANSYS (Dols *et al.* , 2016).

5.2 Modelling the archetype

A generic model is developed for each archetype accounting for their characteristics using ContamX. Figure 5.1a shows the first archetype, as an example. It is a detached single storey house with 6 rooms, two bedrooms and two bathrooms, a self–contained kitchen, and a living–dining room. All rooms are connected to a common family room so that the living room, dining room, and corridor can be considered a single volume, to which all other rooms are connected via doors. Rooms and doors are then represented in ContamW models by well–mixed zones (seen as rectangles) and airflow paths (seen as diamond–shaped dots); see Figure 5.1b.

ContamX calculates the rate of air pollutant transport through airflow paths using information entered for each one. Thus, it uses information about the actual layout of the house (Figure 5.1a) rather than the highly approximate CONTAM sketch. CONTAM models require that each external element is designated an orientation, or azimuth angle, so the wind pressures at its location can be determined. These are assigned for each generic archetype. For instance, a 0° azimuth angle is assigned to all airflow paths located in north facing walls. For an identical archetype that does not face north, these azimuth values can then overwritten by creating a new project file, and adding

the new orientation to the azimuth angle of each external airflow element.



(a) **(b)**

Figure 5.1: First archetype (a) layout and (b) CONTAM model.

Total floor areas are given by Building Permits; see Section 4.3. However, they are not given by room and so they are allocated to each rooms by the modeller, making them a source of uncertainty. Room volumes and wall surface areas are then calculated by assuming that the rooms have a square floor plan and a room height of 2.4 m, which is a standard height given in the Chilean Building Code (MINVU, 2007). External walls from each room are then used to determine the total envelope area following the layout.

5.2.1 Airflow paths

For simplicity, each archetype has fixed room volumes and floor areas, but the properties of each airflow path are variables. Figure 5.2 shows the CONTAM model used to simulate one of the archetypes; Archetype ID27. Airflow paths (in red) and sources/sinks (in green) are arranged following the same order in every room. From the top–left of the diagram, airleakage paths are modelled using a single element to represent those in a room's ceiling and another to represent the floor (Jones *et al.* , 2015), followed by a window element, and three airflow paths are used to model airleakage paths in a wall, assumed to be uniformly porous, by locating at its top, mid–point, and bottom following Jones *et al.* (2013). In bathrooms, an exhaust fan is incorporated next to the window element. When a room is located in a corner, like the first bathroom in Figure 5.2, six airleakage paths are modelled; three for each wall.

The flow elements are adjusted so they account for the corresponding surface area of the element they represent. CONTAM allows this by adding a *multiplier* to the element. Airflow through floors is not accounted for in any of the archetypes modelled for this study, and so their airleakage paths are ignored by giving them a multiplier with a value of zero. However, they remain in the model so that they can be applied

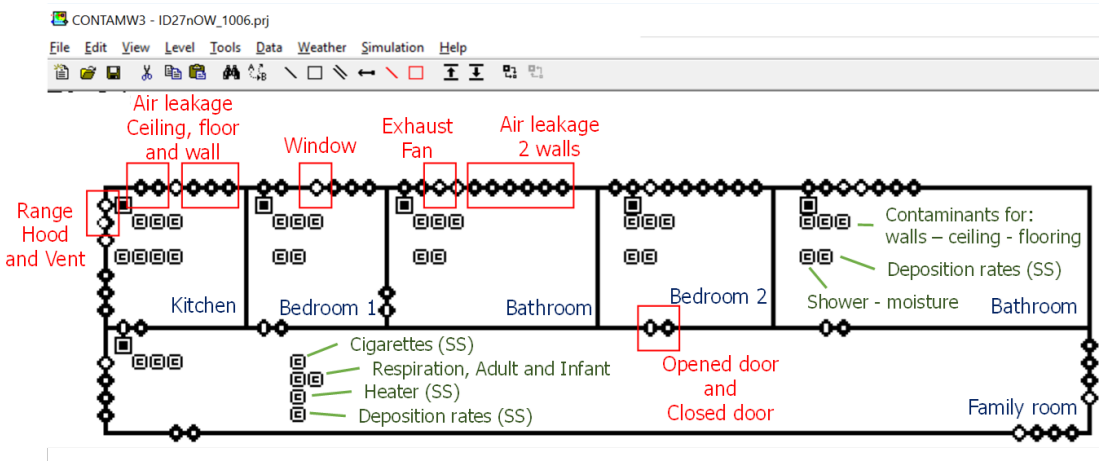


Figure 5.2: CONTAM elements for archetype ID27.

in the future, if required. Wall area (*multipliers*) are equally distributed into the three paths. Paths are modelled using the power law model given in Equation 3.1.1. To calculate the airleakage rate normalised by the thermal envelope $\dot{Q}_{50}$, NL values are sampled from the corresponding NL distribution, which depends on the region number and climatic zone, and converted to $\dot{Q}_{50}$ values using

$$NL = 1000 \cdot \frac{ELA_{4Pa}}{Area} \cdot (\frac{H}{2.5m})^{0.3} \tag{5.2.1}$$

$$ELA_{4Pa} = \sqrt{\frac{\rho}{2(4Pa)}} \cdot Q_{50} \cdot (\frac{4Pa}{50Pa})^{n} \tag{5.2.2}$$

where, ELA_{4Pa} is the effective leakage area at 4 Pa (m^2), $Area$ is the house floor area (m^2), H is the house height assumed to be 3 m, n is the non–dimensional flow exponent, and ρ is the air density (kg/m^3). The flow exponents n are sampled from a normal distribution truncated between 0.5 and 1, with $N(0.651, 0.077)$ following Sherman & Dickerhoff (1998). The corresponding $\dot{Q}_{50}$ values for each geographic region are shown in Figures 5.3a and 5.3b. This value is then entered into CONTAM with a coefficient equivalent to the corresponding section of the envelope area (for each third of a wall, and for ceilings and floors). Party walls and floors are assumed to be impermeable.

Windows are all modelled as sash types, with a rectangular section and a fixed cross–sectional area. They are modelled using the one–way orifice equation

$$Q = C_d \cdot A_f \sqrt{\frac{2\Delta P_a}{\rho}} \tag{5.2.3}$$

where Q ($m^3\,s^{-1}$) is the airflow rate through the opening, C_d is the discharge coefficient, A_f (m^2) is the opening free area, and ΔP_a (Pa) is the pressure drop across the

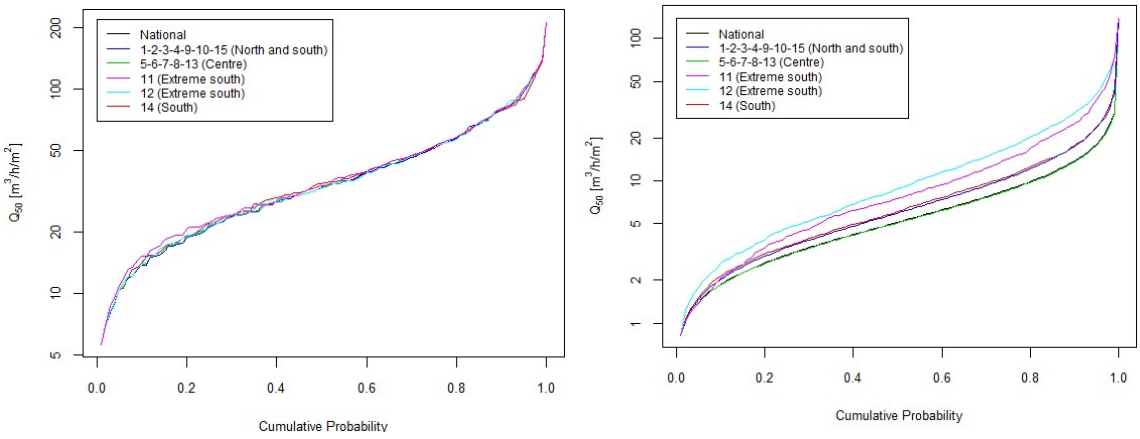


Figure 5.3: Predicted air permeability normalised by envelope area $\dot{Q}_{50}$ to model (a) old and (b) new Chilean houses by climate zone, and nationwide in black.

opening. Their discharge coefficient C_d is 0.6 and their relative elevation is at the mid–point of the window opening (Das *et al.* , 2014).

Open internal doors are modelled using the two–way flow two–opening model suggested in Dols & Polidoro (2015), with a discharge coefficient of 0.78 and its relative elevation is at the bottom of the door. When closed, doors are modelled by one–way flow power law as rectangular sections with a discharge coefficient of 0.68.

Simulations of a blower door test at 50 Pa were run in `CONTAM` for each archetype in order to ensure that the model's external airleakage rate was correct.

5.2.2 Weather data

To include the local weather conditions in the models, winter data is extracted from the Meteonorm files and converted into `CONTAM`'s format (`*.wth` files) using `R` software following Dols & Polidoro (2015). Data is reported hourly giving the date and time, ground temperature, atmospheric pressure, wind velocity, and absolute humidity. Atmospheric pressure (Pa) is calculated as a function of the altitude of the nearest capital city, following Liddament & Air Infiltration, Agence internationale de l'énergie (1996). The same weather file is used in the simulations for an entire region. Data is selected so that only the winter season, between June 21st and September 21st, are retained.

The wind speed modifier specific to the location of each house is calculated following Liddament & Air Infiltration, Agence internationale de l'énergie (1996). The terrain type is randomly sampled using the procedure described in Section 5.4, which gives a coefficient K and an exponent a that are applied to Equation 4.4.1. The wind

pressure coefficients C_P (see Section 4.4.1) are then calculated using the Swami & Chandra (1987) model. Wind pressure profiles for each house are a function of a sampled block aspect ratio S (see Section 4.4.2) and a sampled sheltering coefficient. C_P at different angles are assumed to be between 0° to 360° and specified for each side of a house. The corresponding wall type is assigned to each airflow path when required. Figure 5.4 shows the wind pressure profile for the longest wall of one house.

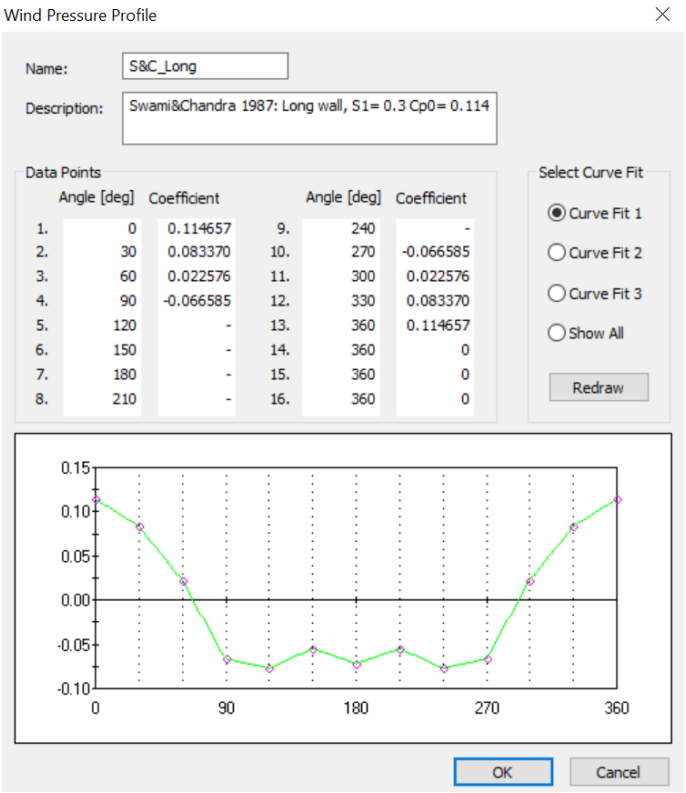


Figure 5.4: Example of a wind pressure profile for the long wall of a house to be used in the CONTAM simulations.

5.2.3 Indoor temperatures

CONTAM is not a thermal model and so the internal air temperatures must be specified. Indoor temperatures are chosen from a normal distribution of N(21.1°C, 2.5°C), following the empirical data analysis of Shipworth *et al.* (2010), which is used elsewhere (Jones *et al.* , 2015). Temperatures are considered constant and equal to each room, and so they are not included in daily and weekly schedules. This temperature is different from the temperatures used to calculate the use of heaters in Section 4.7.4..

5.2.4 Species, sources and sinks

CONTAM requires an emission rate and a deposition rate for pollutants, and because the sources are not used constantly, an emission rate schedule is also required. CONTAM only accounts for the dynamic processes of aerosols and gases associated to emissions

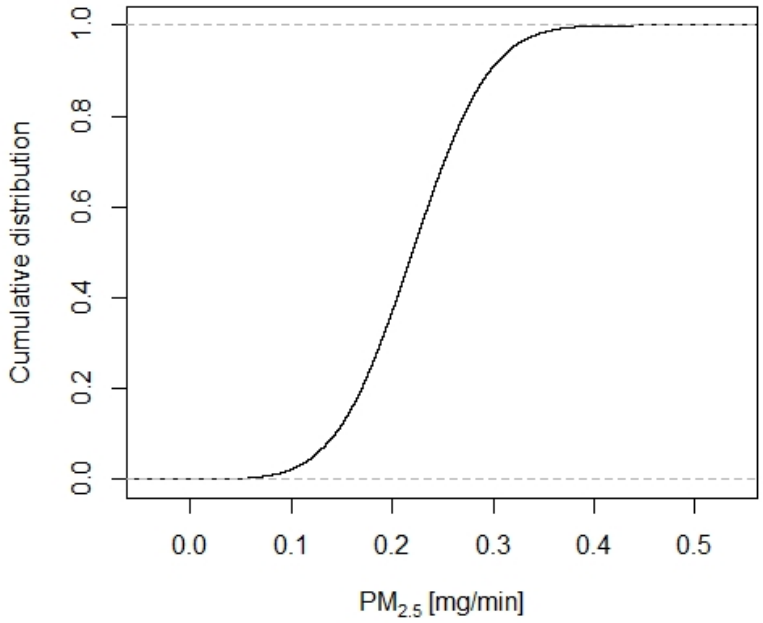


Figure 5.5: Cumulative density function of breakfast $PM_{2.5}$ emission rates.

from primary sources and their deposition onto indoor surfaces. Losses are also possible through purpose provided openings, exfiltration, and mechanical extract fans without recirculation (see Section 4.5.1). Although two highly reactive gases are modelled, NO and NO_2, their chemical reaction are not considered. Only *Water Vapour* and *Dry Air* are considered as non–trace contaminants, or constituents of the air, following Dols & Polidoro (2015).

$PM_{2.5}$ emission rates from cooking meals follows the synthetic probability distribution given in Section 4.7.3. Table 4.12 can be populated as more data become available, and so the probability distribution may be subject to change in the future. Table 4.12 shows that the first and second studies of the toasting of bread are subject of higher uncertainty in their results (having both coefficient of variance above 1), whereas the third study has a higher sample size, a lower variability, and coefficient of variance below 1. Thus, a normal distribution of $N(0.218, 0.0648)$ is selected to model breakfast preparation in the archetypes. Emission rates are all assumed to be constant, $S(t)$, during all cooking events (O'Leary *et al.* , 2019a).

`CONTAM` allows the definition of *Super* sources and sinks, *SS*, where a set of species are assigned to a single source element, as shown in Figure 2.17, and so only one icon and schedule is needed. Eight sources and sinks are designated, three of them in the form of super sources (*SS*): one to model the emissions from heaters, one for the emissions from cigarette smoking, and one for all deposition rates. Table 5.1 shows

the sources and sinks included.

Table 5.1: Sources and sinks included in the archetypes. *SS*: super source or Super sink; **ccf**: constant coefficient model; **brs**: burst source model; **drs**: deposition rate model.

Source	**Species**	**Model**	**Name**
Construction materials	HCHO	ccf	Ceiling_FA
Construction materials	HCHO	ccf	Flooring_FA
Construction materials	HCHO	ccf	Paint_Wall_FA
Cigarette		*SS*	
	C_3H_40	brs	Cigarette_Acr
	ETS	drs	Cigarette_ETS_Deposition
	ETS	ccf	Cigarette_ETS_Emission
	$PM_{2.5}$	brs	Cigarette_PM2.5
Cooking Main Meals	$PM_{2.5}$	ccf	Cooking_PM2.5
	H_2O	ccf	Moisture_1meal
Heater		*SS*	
	CO	ccf	Heater_CO
	NO	ccf	Heater_NO
	NO_2	ccf	Heater_NO2
	$PM_{2.5}$	ccf	Heater_PM2.5
Deposition rates		*SS*	
	$PM_{2.5}$	drs	DepPM2.5
	HCHO	drs	Dep_FA
	NO	drs	Dep_NO
	NO_2	drs	Dep_NO2
Respiration – Adult	H_2O	ccf	RespAd_moisture
Respiration – Infant	H_2O	ccf	RespIn_moisture
Showering	H_2O	ccf	Shower_moisture
Cooking breakfast	$PM_{2.5}$	ccf	WeeklyBF_PM2.5

This study aims to estimate the contribution of indoor sources to the total exposure, and so it assumes that both the background and internal initial concentrations are zero. There are additional reasons for this decision, which are further discussed in Section 7.4. Firstly, indoor and outdoor air composition may differ and so the exposure analysis is carried out separately (see Chapter2). Secondly, infiltration factors are also subject of high uncertainty, which must be incorporated stochastically, followed by the definition of its probability distribution. Thirdly, large differences can occur between ambient concentrations in different locations due to, for instance, their proximity to the sources, or local sheltering. Finally, high resolution data about outdoor concentrations in Chile is very limited, significantly affecting their value, but they can be applied in the future.

Deposition rates play an important role in determining the removal rates and indoor concentrations of pollutants, especially when ventilation is limited. In these models, deposition rates are sampled from a normal distribution truncated at the origin and with whose values of $\tilde{\mu}$ and σ given in Table 4.11. They are modelled using constant deposition rate models ϕ(t). Nevertheless, these values are uncertain (Nazaroff *et al.* , 1993), for instance, due to room dimensions, furniture area, and air velocity. Although all these parameters change by house, this model simplifies the deposition process and input values by following the literature and using a surface to volume ratio $S/V = 2$ (Thatcher *et al.* , 2002).

5.2.5 Activity schedules

The occupants of the archetypes generally consist of 2 adults and 0–3 children; see Section 4.6. In the example shown in this chapter, the house considers a family of 2 adults and 2 infants. Both adults and both children occupy bedrooms 1 and 2, respectively. To account for the use of the rooms and to calculate occupant exposures, a fixed daily schedule is developed using the data presented in Section 4.6. Activities of interest (cooking, heating etc.) were selected from the 2015 ENUT survey and their statistics were computed for each region and nationwide and are given in Appendix B. The use of the rooms is assumed to be the same for all houses, since the effect of human behaviour are not accounted. Further analysis of activity patterns are required in order to test this factor in a manner that is appropriate to answer social research questions.

Daily schedules account for the total time that the 2015 ENUT interviewees (considered here to be the Chilean population) reported as their average sleeping hours; see Section 4.6. The sleeping time is considered to be the core activity and the national average is selected to represent the entire Chilean population; see Table 4.8. Consequently, sleeping is scheduled from 11:00 pm to 6:21 am for weekdays and from 12 am to 7:57 am for weekends; see Figure 5.6.

Meals are used to inform the time when the kitchen door is opened and the cooking periods. During a normal working day, breakfast is eaten at 7:20 am, and has a duration of 1 h 15 min. Lunch time is scheduled from 12 pm to 1 pm, *tea time* from 5 pm to 6 pm, and dinner from 8 pm to 9 pm; see Figure 5.7. At weekends breakfast starts at 9am and has a duration of 1h 20min. Lunch is scheduled from 12:30 pm to 3 pm, and a type of meal that joins *tea time* and dinner (known colloquially as *once–comida*) is scheduled from from 8 pm to 10 pm. Breakfasts are assumed to be eaten in the kitchen, whereas the other meals are eaten in the family room.

Meal preparation is assumed to have a duration of 1 h on normal week days and 1 h during the weekends. Times and durations are included within the time the kitchen

door is opened (see Table B.4).

The emission rate of moisture from occupants' respiration and transpiration (see Section 4.7.5) follows the bedroom door schedules. Occupants emit water vapour in the bedroom during sleeping hours when the door is closed, and are considered to be out of the room when the door is open. During the rest of the day, one adult and one infant are assumed to remain at home, and so they are modelled as sources of moisture in the living rooms.

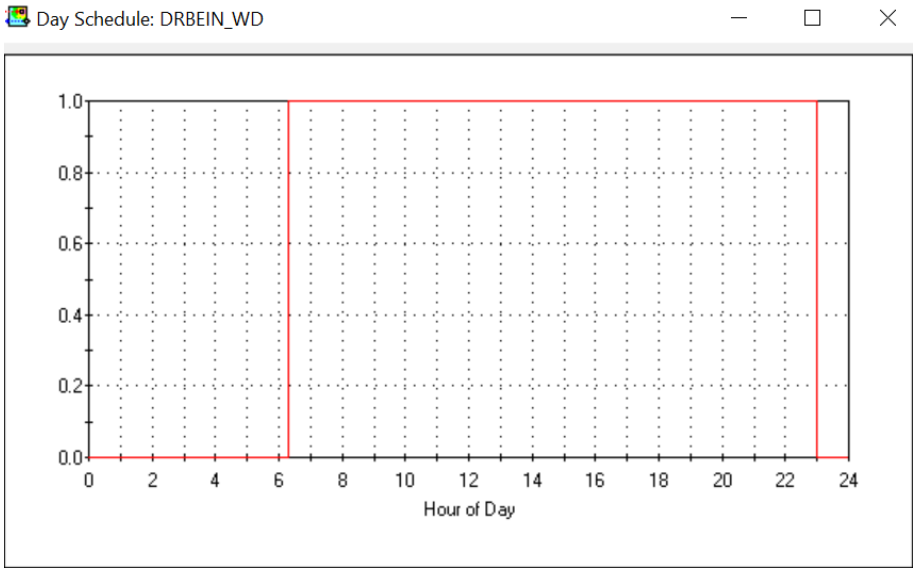


Figure 5.6: Example of the schedule followed by the bedroom doors during a normal working day. A value of 1 is used when the door is opened and a value of 0 when closed.

Bathroom mechanical extractor fans are modelled *ON* from the wake–up time and when *showering* and *getting dressed*; both activities reported in the 2015 ENUT (INE, 2016b). Kitchen range hoods are considered *ON* when cooking a meal (from the time a meal starts) and remain in operation for one hour. The capture efficiency of the range hood is not considered by `CONTAM`. Moreover, there is no information about the use of range hoods in Chile or about their capture efficiency. To be included in the study, a known or assumed probability distribution in the stock would be required in order to test the effect of this factor in the results. This implies that the contaminants are perfectly and instantaneously mixed at each time–step, concentrations that are used to calculate the mass removed by the exhaust fan and by the rest of the airflow elements. A capture efficiency could be added in the future as a work–around by amending the emission rate.

The heater is considered to be *ON* from 7 am, and to remain *ON* for the same num-

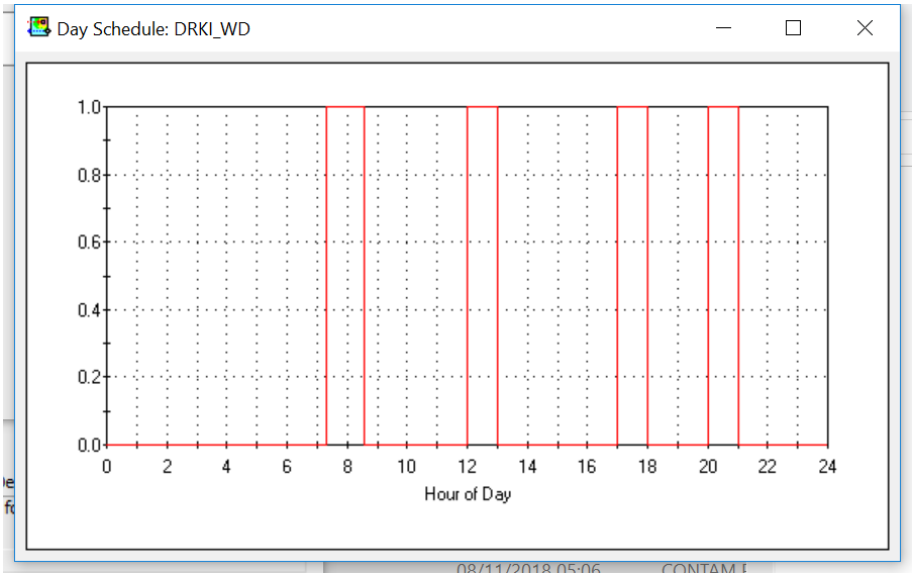


Figure 5.7: Example of the schedule followed by the kitchen door during a normal working day. A value of 1 is used when the door is opened and a value of 0 when closed.

ber of hours every day during the winter season, which depends on the location of the house; see Section 4.7.4 and Table 4.15.

Bedroom doors are always open except at night, and the kitchen door is closed except when cooking (Milner *et al.* , 2014); see Section 5.2.5. Doors are never partially opened, so that their schedules are modelled with rectangular shapes rather than trapezoidal; see Figure 5.7.

5.3 Case scenarios

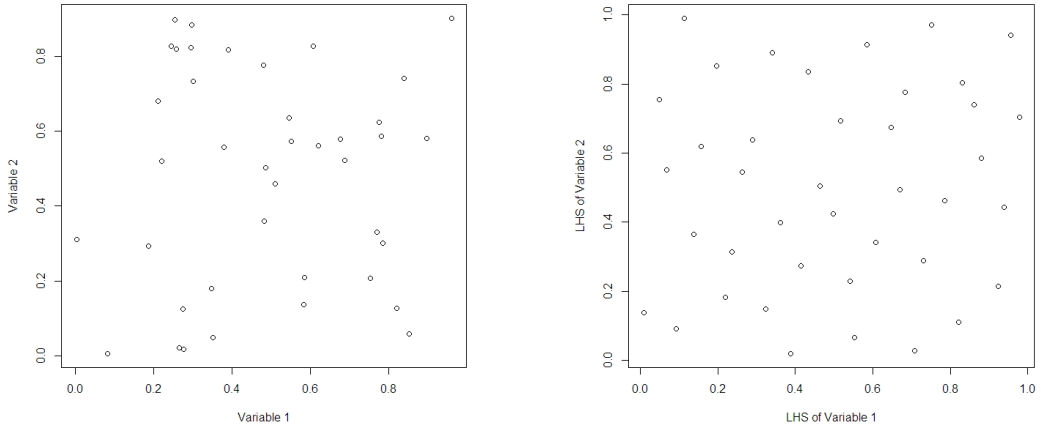
There are no reported measurements in the literature of windows opening behaviour in Chilean houses at the stock scale. In the absence of knowledge, the models consider a *multiplier* for every window where 1 is fully open and 0 is fully closed. These two extremes are used to test extremes: a fully open and a fully closed scenario.

5.4 Sampling method

The present model has input parameters that are specified deterministically, or described by discrete or continuous probability distributions. A set of input values are applied to `CONTAM` to predict pollutant concentrations over time during the winter season. By systematically varying each set of `CONTAM` inputs and running multiple

simulations, distributions of output variables are generated that quantify uncertainty in them; see Chapter 6.

The values of probabilistic input parameters are obtained using Latin Hypercube Sampling (LHS), using bespoke *R* Version 3.5.0 code (R Core Team, 2018) and the `optimumLHS` function. LHS is chosen over the random sampling because it has the advantage of improving the stratification of the samples within a hyper–grid (Helton & Davis, 2003), and so it reduces the number of simulation iterations required for the predictions to converge; see Figures 5.8a and 5.8b.



(a) **(b)**

Figure 5.8: A sample of 40 values for 2 different parameters, requiring 2 *dimensions*, using (a) a random sampling with uniform distribution and (b) a LHS.

There are 11 direct probabilistic inputs:

1. Block aspect ratio (Section 4.4.2),
2. Δ temperature (Section 4.5.2),
3. Relative north (Section 4.4.2),
4. Air permeability (Section 4.5.2),
5. n exponent (Section 5.2.1),
6. $PM_{2.5}$ deposition rate (Section 4.7.2),
7. Formaldehyde deposition rate (Section 4.7.2),
8. NO deposition rate (Section 4.7.2),
9. NO_2 deposition rate (Section 4.7.2), and
10. Breakfast emission rate (Section 4.7.3),
11. Cooking meal emission rate (Section 4.7.3).

The LHS generates a value for each sample between 0 and 1 that is assigned to an input and considered to be the probability of it occurring. When applied to an

inverse cumulative distribution function (CDF) of the input, a value is generated; see Figure 5.9 for an example of a single variable. Here, the LHS chooses 40 samples, one of which is a value of 0.7 assigned to variable 2, the $PM_{2.5}$ emission rate. When it is applied to the y–axis of a CDF for $PM_{2.5}$ emissions, reading from the x–axis gives an emission rate of 2.1 mg/min. Ten sets of these inputs are chosen at a time, following Das *et al.* (2014).

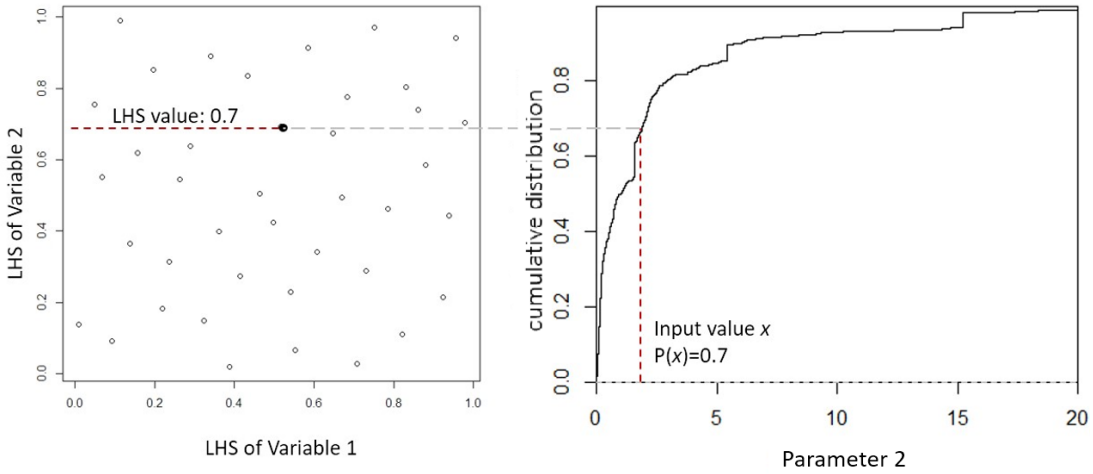


Figure 5.9: Example of how an input parameter (right plot) is obtained using the LHS sampling procedure (left plot).

After each set of samples is simulated, the total sample size increases by the set size and its $\overline{PM}_{2.5}$ mean ($\tilde{\mu}$) and standard deviation (σ) is calculated. The total number of samples is deemed to have converged, and the *stopping criteria* met, if the change in $\tilde{\mu}$ and σ from one set of 10 samples to the next is less than 0.5%. This stopping criterion is chosen to reflect the lower limit of accuracy of a good Indoor Air Quality (IAQ) sensor following Jones *et al.* (2015).

Simulations are run individually and converged for each region number (15 geographic regions). The process was repeated using the 8 archetypes and the 2 scenarios. Therefore, $8x15x2 = 240$ different processes are carried out. The 240 processes described herein are carried out in batch mode and run by `R` until the stopping criteria is met.

5.5 Processing model predictions

Criteria pollutants for the Chilean houses follow Logue *et al.* (2011a) and are listed in Section 4.7.1. They used DALYs as the metric of health impact, which allows the health outcomes caused by different pollutants to be compared using a single unit. Fig-

ure 2.13 shows that $PM_{2.5}$ are predicted to be the most harmful pollutant in dwellings by one order of magnitude. Thus in this study the stopping criterion for the simulation process, the post–analysis of the results, and the discussion will be focused only on the first–ranked pollutant, $PM_{2.5}$. This choice also reduces computational and processing time. However, other outputs, such as other pollutant sources, could be used in the future.

From each set of outputs, the room–weighted pollutant concentration ($\overline{PM}_{2.5}$) is calculated. The average $PM_{2.5}$ concentration over the simulation period is calculated for each room. Then, the average concentrations in the kitchen, bedrooms, and in the family room are weighted by the daily time the cook family member spend in them. Here, a ratio of 10 : 38 : 52 is used for the *kitchen : bedroom : family room*, respectively; see Section 4.6.

Then, the time–series data of the ventilation rates (ACH) for each building, the $PM_{2.5}$ concentrations for each of the rooms, and the indoor–outdoor temperature difference (ΔT) are obtained. ACH and $PM_{2.5}$ data are extracted from the simulation output files, and ΔT is computed using the indoor temperature input value and the weather data.

From this data, the median ACH, the total $PM_{2.5}$ exposure levels, and total heat loss due to infiltration (H_I) are computed for each simulation. These metrics are recorded and compiled into a single file labelled with the simulation file ID and region number.

5.5.1 Predictions at national level

To identify the uncertainty of the three outcomes at a national level, the 15 cumulative distributions of the simulations produced at regional level are used. Values are chosen by randomly sampling from the CDFs according to the corresponding regional weighting factor, and so giving a stock–weighted national curve. To compare both extreme scenarios, two curves are produced, which are seen in Figures 6.13 and 6.15 of Chapter 6.2, and discussed in Section 7.5.2.

5.5.2 Exposure analysis

There are two main approaches for quantifying exposures, direct and indirect; see Section 2.5. This study uses an indirect approach by predicting the indoor pollutant concentration over time and making assumptions about occupant behaviour.

Hourly average particle concentration profiles are predicted over the winter time and used for the exposure assessment, following the Oezkaynak *et al.* (1996) method of calculating the population–weighted exposure.

In order to calculate the magnitude of the exposures, the predicted $PM_{2.5}$ concentration in each room over time, and the contact times of one family member are used. Table 4.9 is used to set the activity pattern followed by *the cook* family member; the time of the day the occupant is in each room. Pollutant concentrations are selected from `CONTAM`'s contaminant files (`.ncr`) using `simread3.exe`, data analysis software developed by NIST.

Time–weighted $PM_{2.5}$ average exposure concentrations are then calculated using the first term of Equation 2.3.1, where E_i is the exposure concentration in each of the rooms (units of concentration) at the times presented in Tables 5.2 and 5.3 for houses with and without a separate kitchen, respectively.

Table 5.2: Time schedule followed by the cook family member, used to calculate the $PM_{2.5}$ exposure levels.

Room	**Time**
Bedroom	23:00 - 06:00
Kitchen	07:00 - 09:00
	12:00 - 13:00
	17:00 - 18:00
	20:00 - 21:00
Living room	06:00 - 07:00
	09:00 - 12:00
	13:00 - 17:00
	18:00 - 20:00
	21:00 - 23:00

Table 5.3: Time schedule followed by the cook family member, used to calculate the $PM_{2.5}$ exposure levels in houses with no separate kitchen.

Room	**Time**
Bedroom	23:00 - 06:00
Kitchen & Living room	06:00 - 23:00

5.5.3 Ventilation rates

`CONTAM`'s *Whole Building Air Change Rate* files (`.ach`) are processed by reading them as text files using `R` and selecting the columns containing the time date and ventilation rates. Hourly averaged, and median values, over the winter time were obtained for each project file.

5.5.4 Heat loss

Heat loss due to exfiltration (H_I) is calculated using the transient predictions where

$$H_I = \int \dot{V}_I(t) \cdot \overline{\rho}(t) \cdot c \cdot \Delta T(t) \cdot dt \tag{5.5.1}$$

Here, V_I (m^3/s) is the infiltration or airleakage rate, $\overline{\rho}$ (kg/m^3) is the mean air density between the indoor and outdoor air densities, c ($J/kg/K$) is the specific heat capacity of air, and ΔT (K) is the absolute difference between the indoor and outdoor temperatures. The air density is a function of the air pressure ph and temperature, which are obtained the relevant weather file, and calculated using Equation 5.5.2.

$$\rho = ph/(RT) \tag{5.5.2}$$

where R is the specific gas constant for dry air (J/(kg K))(Dols & Polidoro, 2015).

5.5.5 Statistical analyses

To analyse the data, three statistical tests are used. As the three model outputs are expected to follow a non–normal distribution, the tests are chosen to account for this.

To test the occurrence of an effect, and to find its location, in categorical variables, a Kruskal–Wallis H test and *post–hoc* pairwise multiple comparison tests are carried out for all archetypes and regions, following Field *et al.* (2012). The H test determines whether differences between the medians of two samples are significant, based on the null hypothesis that all medians are equal. The *post–hoc* tests identify which pairs of samples are significantly different of each other, testing the null hypothesis that each pair of medians are equal. These tests are carried out using `R`'s `kruskal.test` and `kruskalmc` functions. Additionally, Levene's tests is executed with `R`'s `leveneTest` function to test the homogeneity of variance. Chapter 6 shows that the results are non–normal and so the median was used for comparison. Finally, effect sizes are used to identify the size of the difference between two samples, following Ferguson (2009) and using the group difference index, also known as Cohen's d

$$d = \frac{M_{ed1} - M_{ed2}}{\sigma_{pop}} \tag{5.5.3}$$

Here, M_{ed1} and M_{ed2} are the outcome median for the two groups compared, σ_{pop} is the standard deviation for the entire population, in this case, calculated by sampling from the national distribution given in in Table 6.3. The following thresholds have been used for labelling the effects:

Table 5.4: Effect sizes thresholds (Ferguson, 2009; Sullivan & Feinn, 2012)

Magnitude	***d*, threshold**
Small	0.2
Medium (Cohen's benchmark)	0.5
Large	0.8
Very large	1.3

5.6 Sensitivity analyses of the input and output data

Twelve inputs and three outputs are retained for the sensitivity analysis. A fundamental assumption of the sensitivity analysis is that all the tested inputs are independent of each other, and so any that are themselves correlated are combined. Therefore, nine inputs are used directly and three are scaled using house characteristics to avoid multicollinearity.

All inputs and outputs are unique for each house, except for the *heater emission rate* and *envelope area : volume* ratio because they relate to a specific household appliance and archetype, respectively; see sections 4.7.4 and 4.1. To compute representative values for the wind speed, the median wind speed at the meteo stations is taken from the winter weather files and scaled at the house site as described in Section 4.4.1. Δ T uses the sampled indoor temperature and median outdoor temperature from the weather files. Finally, the total permeable area is calculated from the sum of the individual air-leakage areas and window areas, when they are modelled *open*. Table 5.5 summarises the 12 inputs used for the sensitivity analyses.

The sensitivity analyses are used to test the dependence of each output on the inputs. Here, the method of Jones *et al.* (2015) and Das *et al.* (2014) is followed, which tests for linear, monotonic, and non-monotonic relationships between the inputs and outputs. Their MATLAB code is accessed via Research Gate [1]. The tests for linear relationship are: (i) Kendall's τ rank, (ii) Pearson's r product moment correlation coefficient, and (iii) linear regression. Monotonic relationships are tested using: (iv) Spearman's ρ rank correlation coefficient, (v) rank-transformed standardised variables. Non–monotonic relationships are tested using: (vi) Kolmogorov–Smirnov and (vii) Kruskal–Wallis quantile tests.

The input and output data are not transformed, and all outliers are included. Outputs and inputs for both scenarios are merged and tested together. Coefficients and p–values are obtained for each test, and the inputs are ranked according to the magnitude of the coefficient.

[1] DOI: 10.13140/RG.2.2.30311.39844

Table 5.5: Inputs retained and outputs computed for the sensitivity analysis.

Inputs			Output
House Design	**Pollutant Species and Source**	**Environment**	
1. Block aspect ratio	7. $PM_{2.5}$ Deposition rate	11. W_s	1. $PM_{2.5}$ exposures
2. Orientation	8. ER from breakfasts	12. $T_{int} - T_{ext}$	2. ACH
3. Permeability	9. ER from main meals		3. H_I
4. n exponent	10. ER from the heater		
5. Envelope : Vol ratio			
6. Total permeable area			

5.7 Summary

This chapter describes the modelling process used to predict pollutant concentrations and ventilation rates in houses of the Chilean housing stock and the sensitivity analysis is used to test the relative importance of model inputs using a sensitivity analysis.

Indoor sources are modelled in eight archetypal houses that represent a 35% of the stock using the multi–zone indoor air quality and ventilation analysis tool `CONTAM`. The models account for the effects of the Chilean environment, local house design, different airflow elements, and occupancy patterns.

A probabilistic approach is applied to investigate uncertainty in model predictions. Inputs are created using a Latin Hypercube sampling (LHS) method. Their values are chosen so that they follow a cumulative density function when their distribution is known. When data is unavailable assumptions are made about their distribution by assigning a fixed value, a value within a range, or a known distribution reported in the literature.

Key pollutant sources are defined as *criteria pollutants* coming from cooking, combustion heaters, building materials, and some occupant activities. Ambient pollutants are excluded.

The models do not account for all possible pollutant behaviours, and so the contaminants *species* are modelled as inert gases, so that the chemical interactions with each other, or with other gases, are not considered. Other processes such as adsorption, resuspension, and filtration are also excluded from the models. Pollutant removal and dilution methods include ventilation through mechanical extractor fans, infiltration, and deposition onto indoor surfaces. To assess the effect of natural ventilation through windows, two extreme scenarios are considered; one in which all the windows remain closed over the entire winter season, and one in which the windows are fully opened.

Model outputs are limited to fine particle concentrations, ventilation rate, and heat loss due to infiltration over time. However, they can be updated in the future as new information and data becomes available. The models are now be simulated to give distributions of outputs that can be used to understand the behaviour of the model and to determine indoor air quality in Chilean houses. These results are presented in Chapter 6.

CHAPTER 6

Model Predictions

Predicted pollutant concentrations and sensitivity analysis

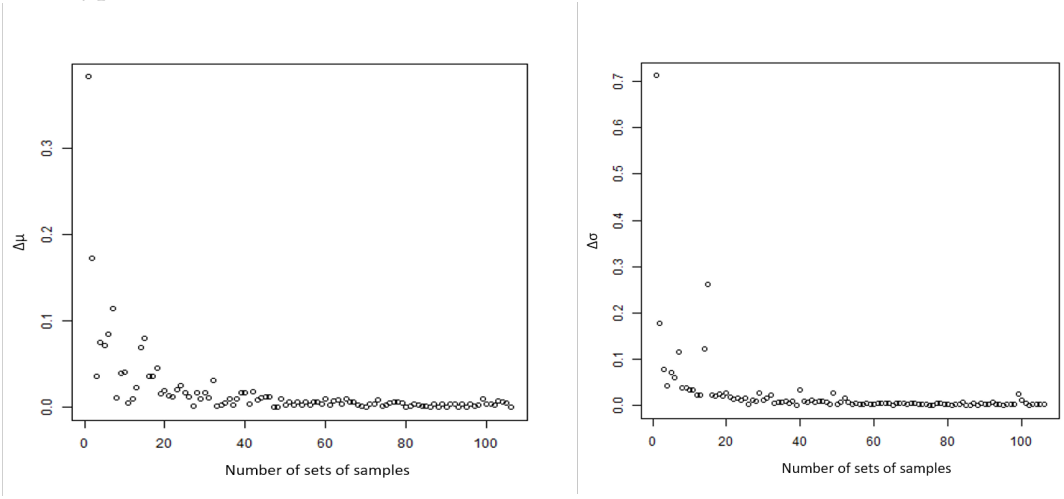
Chapter 4 described how 8 archetypes can be defined which cover 35% of the Chilean housing stock. Here, therefore these archetypes are used to investigate the effects of altering key parameters which might influence exposure to indoor $PM_{2.5}$ using the modelling approach described in Chapter 5. Two extreme scenarios were applied to model each archetype separately; with all windows closed (scenario 1) and with all windows open (scenario 2). The exposure analysis was limited to fine particle concentrations, and so the number of simulations increased until the time–weighted $PM_{2.5}$ concentrations converged into a steady mean and standard deviation.

In this chapter, the results of the predictions are presented. They are divided into four sections. Section 6.1 describes the predicted time–series data of the simulations; Section 6.2 gives the hourly results, and Section 6.3 aggregated over the winter season. Section 6.4 shows the relationship between $PM_{2.5}$ concentrations, heat loss, and ventilation rates, and presents the ventilation rates required to achieve *acceptable* long–term exposures to $PM_{2.5}$ in at least 95% of the national stock; and Section 6.5 presents the sensitivity analysis to show those inputs that most influenced the model predictions.

Following the second objective of this study (Section 1.7), the predictions are used to describe the uncertainty in the indoor air quality of the Chilean housing stock. Thus, the results are presented using probability distributions in addition to descriptive statistics. Moreover, to better understand the relationship between the indoor air quality and the energy demand in the Chilean stock, and some of the implications that intervention measures could lead to, ventilation rates and heat loss by exfiltration are also obtained from each house. Hence, the three modelling outcomes are presented in each section: (i) the exposure to fine particles, (ii) the ventilation rates, and (iii) the heat loss due to exfiltration.

Figures 6.1a and 6.1b show an example of the values of $\overline{PM}_{2.5}$ used for testing

convergence, and to determine the end of the simulation process; see Section 5.4. As the sample size increases by adding new sets of simulations, $\tilde{\mu}$ and σ of $\overline{PM}_{2.5}$ become more stable; see Section 5.4 for a detail description of the method and Section 7.2 for a discussion of this stopping criteria. Table 6.1 presents the number of simulations per archetype required to achieve convergence; approximately 2,100 project files per archetype and scenario.



(a) **(b)**

Figure 6.1: Changes in (a) $\Delta\tilde{\mu}$ and (b) $\Delta\sigma$ of $\overline{PM}_{2.5}$ as sample size increases. The x–axis indicates the number of sets of samples in dozens.

Table 6.1: Final number of simulations per archetype using a stopping criteria of 0.5% change in both $\Delta\tilde{\mu}$ and $\Delta\sigma$ of $\overline{PM}_{2.5}$.

Archetype ID	**Number of simulations**	
	Windows closed	Windows open
27	2,390	2,390
36	2,440	2,320
91	2,050	1,880
100	1,930	2,120
275	1,510	2,150
35	2,870	1,970
284	2,050	2,480
19	1,990	2,080
TOTAL	**17,230**	**17,390**

6.1 Time – series data

The hourly data is obtained from each `CONTAM`'s simulation file (`.sim`) and read using `simread3.exe`; see Section 5.5.2. Each simulation file records hourly averaged data (24 h) over the wintertime (93 days), giving a total of 2,232 data points. Figure 6.2 shows a snapshot of the airflow transient data through some airflow paths operat-

ing in a kitchen for one day. Three types of air path are shown: the kitchen vent, the kitchen fan, and the three airleakage paths located in one of the walls. This example corresponds to a house simulated with all windows closed (scenario 1). CONTAM draws lines from the transient airflow data with one–hour resolution; thus, the airflows for the kitchen fan are visually misrepresented. It shows triangular shapes rather than their corresponding rectangular shape. Nevertheless, this is only a visual misrepresentation, not an error in the models, and so this does not affect the results. Furthermore, the area under the curve, which gives the total airflow, is unaffected. Figure 6.3 shows a house of the same archetype simulated in the second scenario, namely all windows open, and so the plot includes the airflow through one of the kitchen windows.

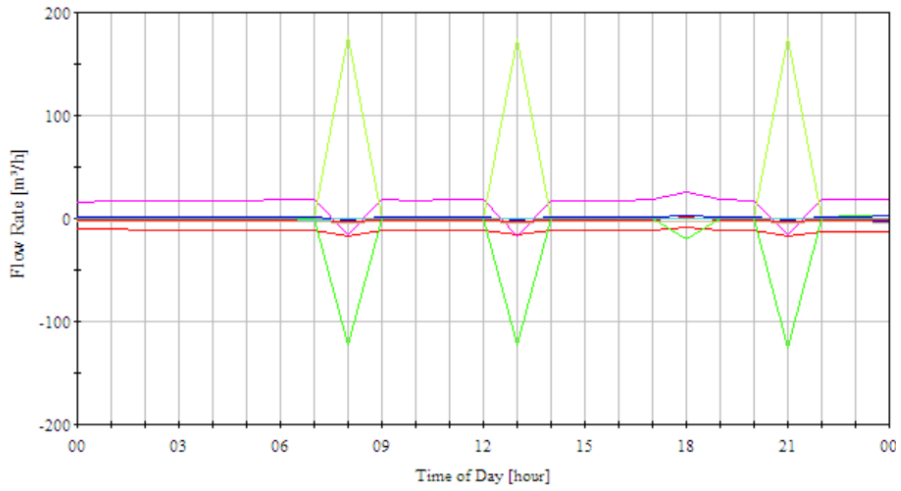


Figure 6.2: Windows closed scenario. 24 h airflows through the flow paths located in the kitchen of a simulated house in the capital region: (Red and green) windows; (Pink) vent; (light green) fan; (blue) kitchen door, (blue, brown, and red) envelope airleakages at 2.4 m, 1.2 m, and 0 m height of the walls. The Q_{50} of this house is 26 $m^3\,h^{-1}\,m^{-2}$, −31% of the average Q_{50} of the archetype.

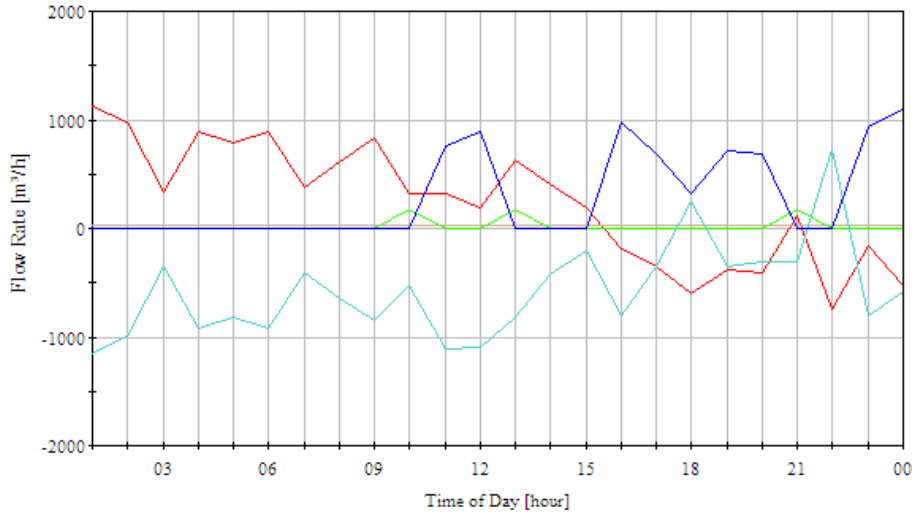


Figure 6.3: Windows open scenario. 24 h airflows through the flow paths located in the kitchen of a simulated house in the capital region: (Red and light blue) windows; (light green) fan; (blue) kitchen door. Here, Q_{50} =48 $m^3\,h^{-1}\,m^{-2}$, +30% higher than the average Q_{50} of the archetype.

Figures 6.4 and 6.5 show the transient $PM_{2.5}$ concentrations in the three rooms used for the exposure analysis: (i) the kitchen, (ii) the bedroom, and (iii) the living–room, for windows closed and windows open scenarios, respectively. Section 5.5.2 described the method and how the data is used, and Section 7.4 discusses the use of the `CONTAM` model for the predictions.

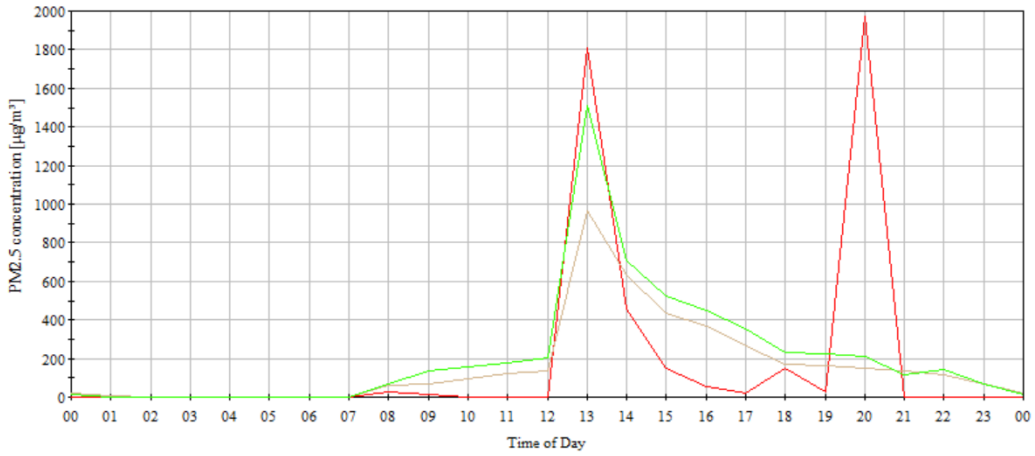


Figure 6.4: Windows closed scenario. 24 h $PM_{2.5}$ exposure profiles in a simulated house located in the capital region. The red line corresponds to the kitchen, the brown line to the bedroom, and green to the family–room.

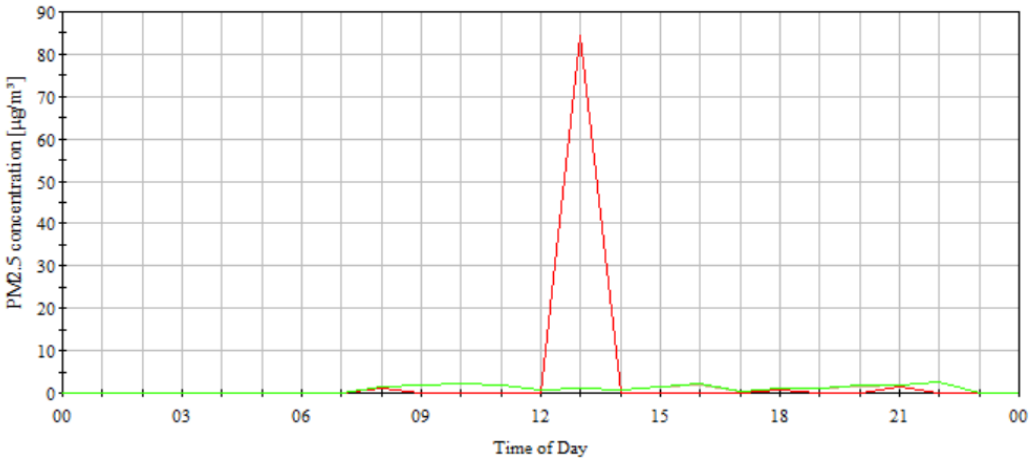


Figure 6.5: Windows open scenario. 24 h $PM_{2.5}$ exposure profiles in a simulated house located in the capital region. The red line corresponds to the kitchen, the brown line to the bedroom, and green to the family–room.

6.2 Hourly data over the winter season

Data over the wintertime (24 h, 93 days) is read using `R`, and summary statistics for $PM_{2.5}$ exposures, ventilation rates, and heat loss due to infiltration are computed.

A cumulative distribution function (CDF) shows the probability of a variable x being less or equal to a certain value of a predictor, Y, or $P(x \leq Y)$, and so they are used to represent the uncertainty in the predictions. Figures 6.10a to 6.10o show the cumulative distributions of hourly $PM_{2.5}$ exposures for one archetype aggregated by region. The x–axes are kept fixed to make the visual comparison between regions easier. Figures 6.6 and 6.7 show the aggregated probability distribution of hourly exposures to $PM_{2.5}$ and hourly ventilation rates for windows closed scenario. This is discussed in Section 7.5.4. The heat loss is not shown because this metric is computed for the entire simulation time, and so the hourly data are not included.

6.2.1 Results by archetype

Tables 6.2a and 6.2b present the simulation results aggregated by archetype. Median hourly exposures for windows closed scenario are generally higher in archetypes 275 and 284, which represent the newer and more airtight houses. In contrast, windows open scenario shows that there are negligible differences between archetypes, and exposures are close to background levels. Pollutant removal by natural ventilation, here the consequence of window opening, is discussed in Section 7.5.2.

(a) Statistics computed for the median hourly $PM_{2.5}$ exposures ($\mu g/m^3$) of windows closed scenario, aggregated by archetype ID.

ID	Median	Mean	SD
27	2.75	9.94	9.14
36	2.22	6.12	5.90
91	2.81	12.68	11.88
100	1.09	4.24	4.32
275	17.52	25.46	18.45
35	0.99	5.44	5.31
19	4.10	11.76	11.02
284	7.61	16.20	12.68

(b) Statistics computed for the median hourly ventilation rates (h^{-1}) of windows closed scenario, aggregated by archetype ID.

ID	Median	Mean	SD
27	0.8	0.9	0.3
36	1.1	1.0	0.3
91	0.7	0.7	0.2
100	1.0	1.0	0.3
275	0.1	0.1	0.0
35	1.0	1.1	0.3
19	0.8	0.8	0.2
284	0.2	0.2	0.1

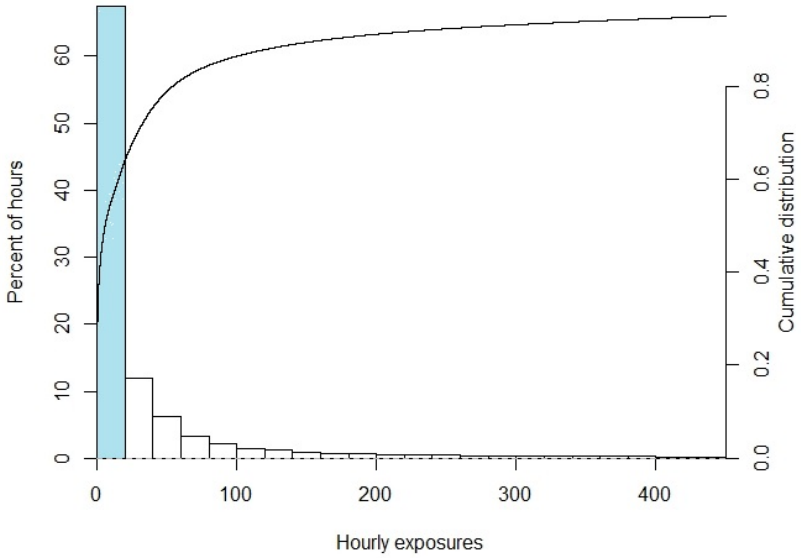


Figure 6.6: Windows closed scenario. Distribution of the hourly exposures of an occupant to $PM_{2.5}$ over the wintertime. The upper limit of the graph corresponds to the 95th quantile (449 $\mu g/m^3$) and the median, or 50th quantile, is 5.8 $\mu g/m^3$. The blue area shows the proportion of winter hours that the stock is below the WHO 24 h recommendation of 25 $\mu g/m^3$.

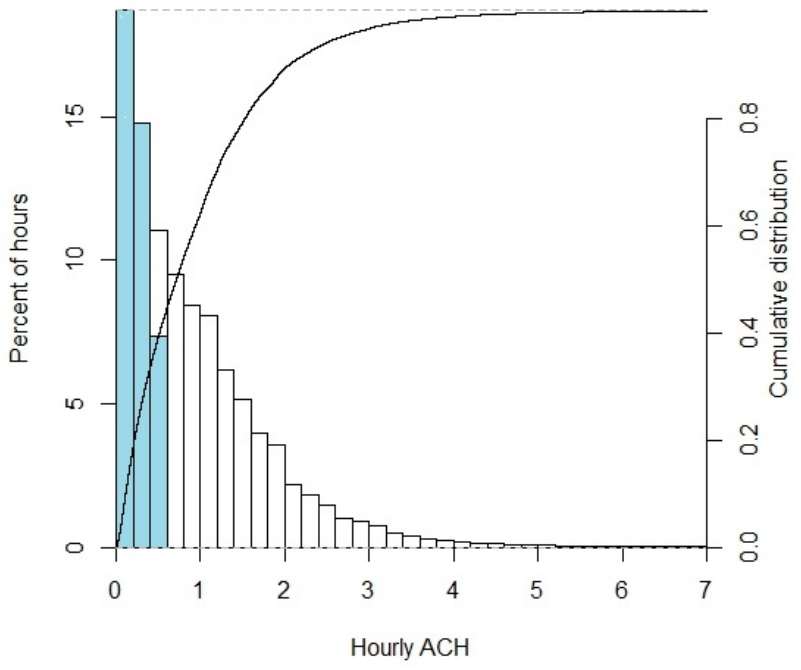


Figure 6.7: Windows closed scenario. Distribution of hourly airchange rates in simulated houses during the winter, as the portion of air volume of a room that is removed or added in each hour (h^{-1}). The blue area shows the proportion of the winter hours that the stock is below the international recommendation of 0.5 h^{-1}.

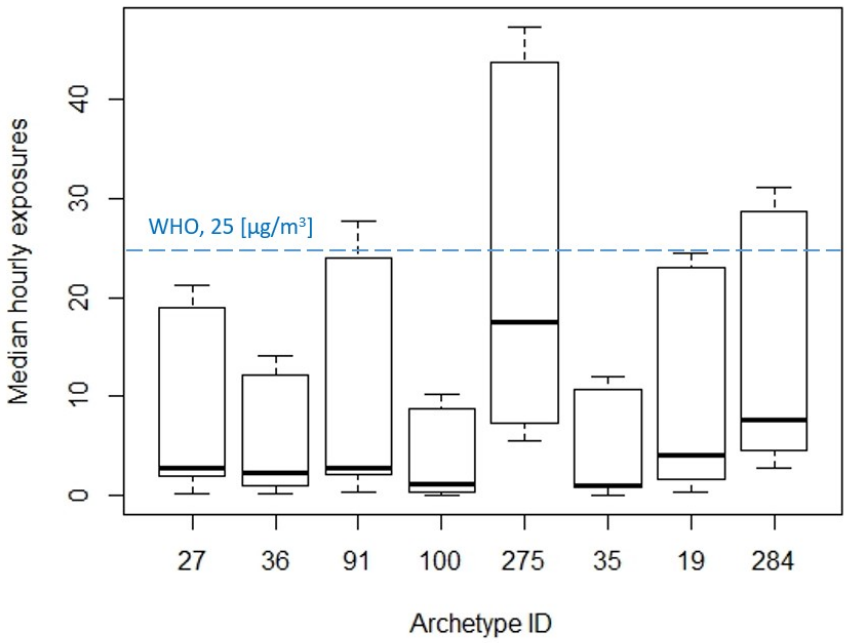


Figure 6.8: Windows closed scenario. Distribution of the medians of hourly $PM_{2.5}$ exposures of all simulated houses by archetype. The green dashed line shows the WHO's 24 h recommendation of 25 $\mu g/m^3$.

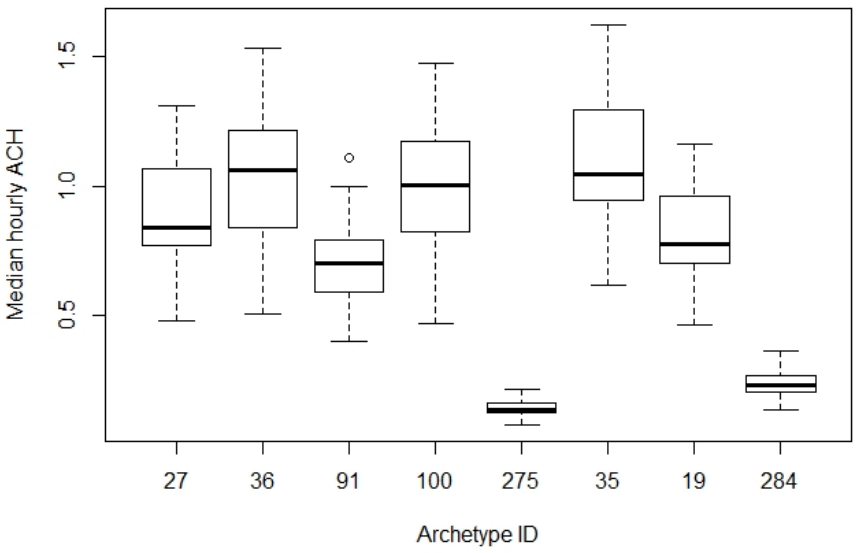
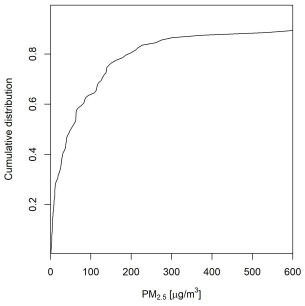
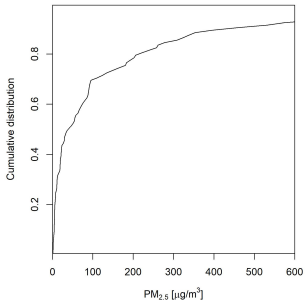


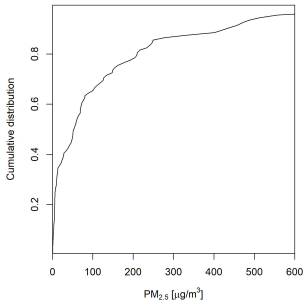
Figure 6.9: Windows closed scenario. Distribution of the medians of hourly ventilation rates (h^{-1}) of all simulated houses.



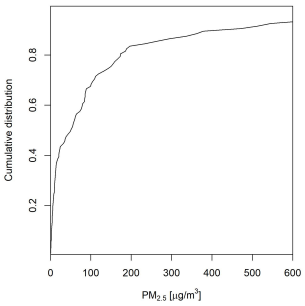
(a) Region 1



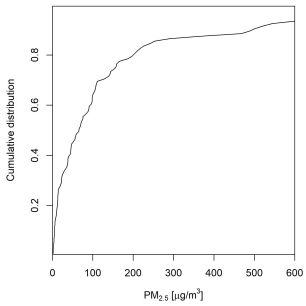
(b) Region 2



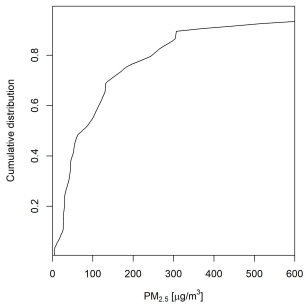
(c) Region 3



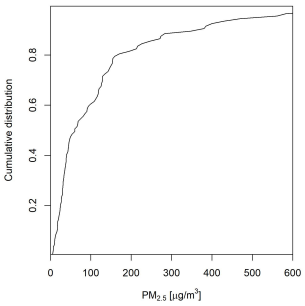
(d) Region 4



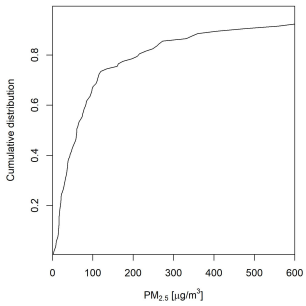
(e) Region 5



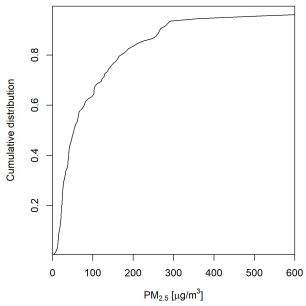
(f) Region 6



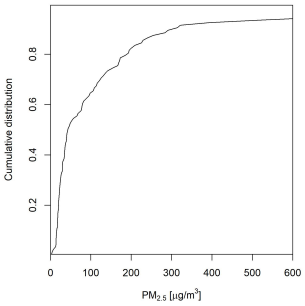
(g) Region 7



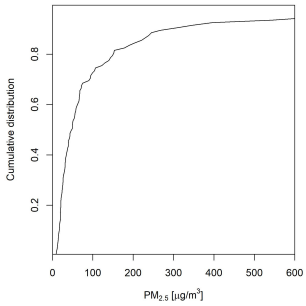
(h) Region 8



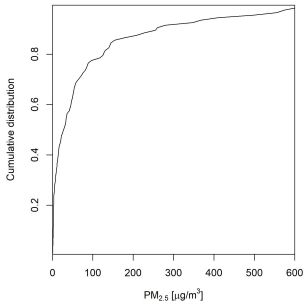
(i) Region 9



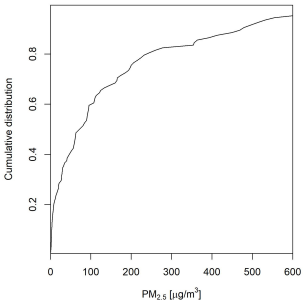
(j) Region 10



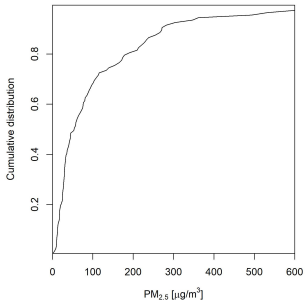
(k) Region 11



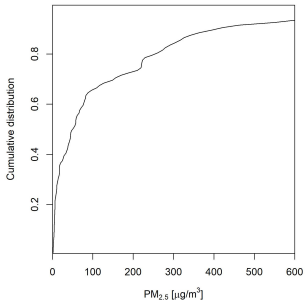
(l) Region 12



(m) Region 13



(n) Region 14



(o) Region 15

Figure 6.10: Windows closed scenario. Predicted distribution of hourly $PM_{2.5}$ exposures for archetype ID27, for each region.

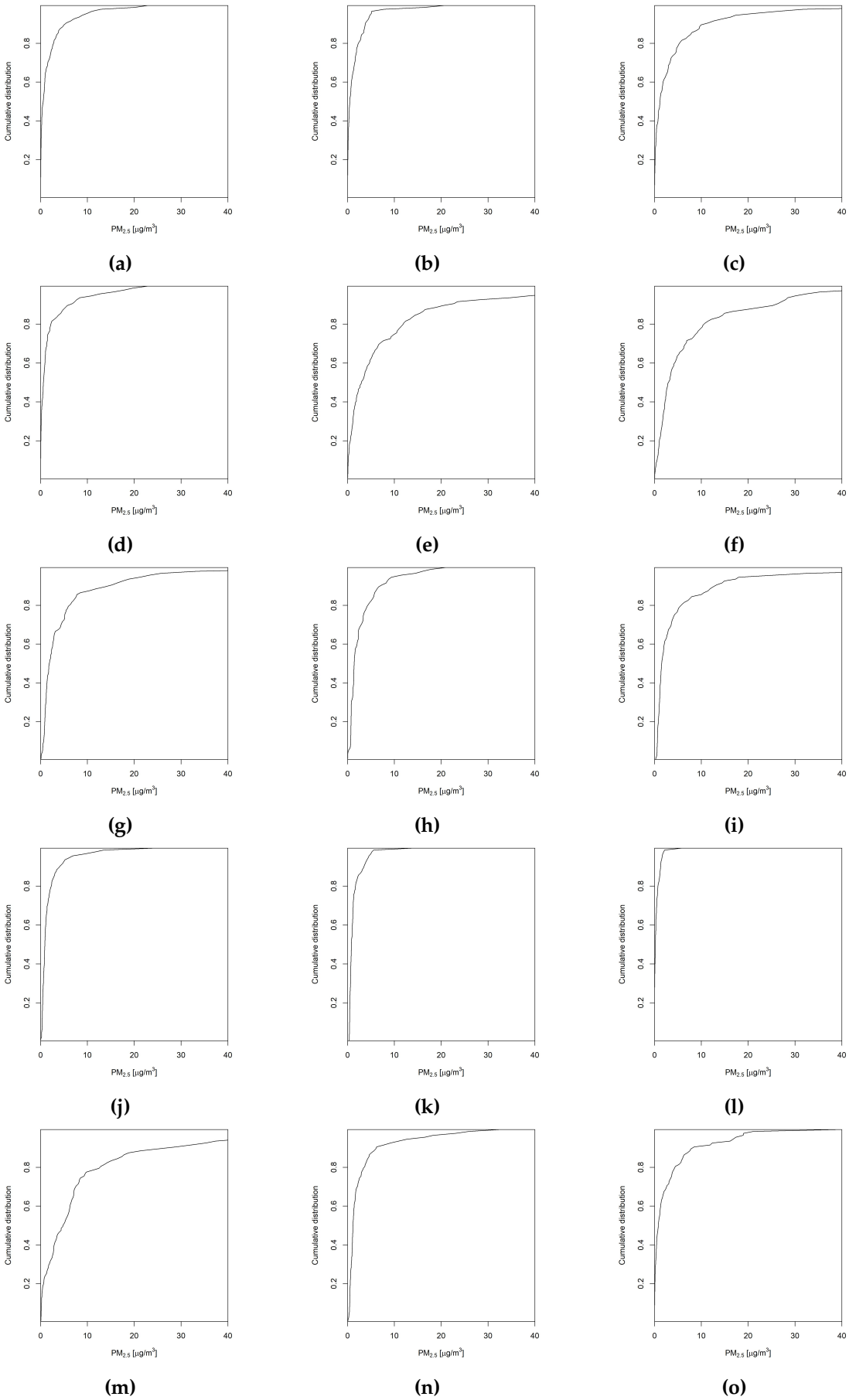
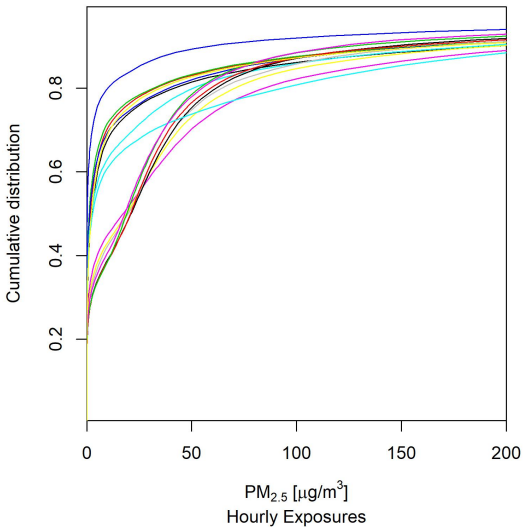
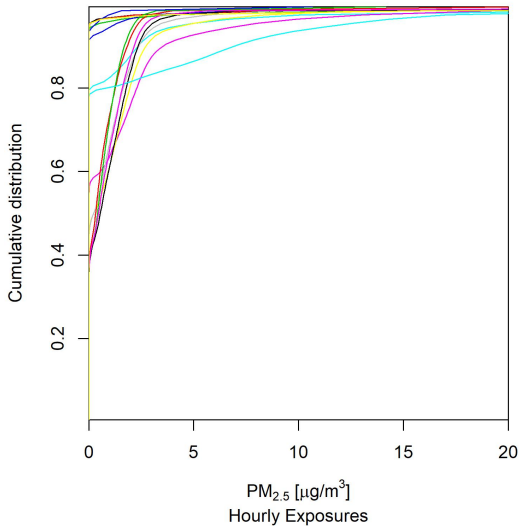


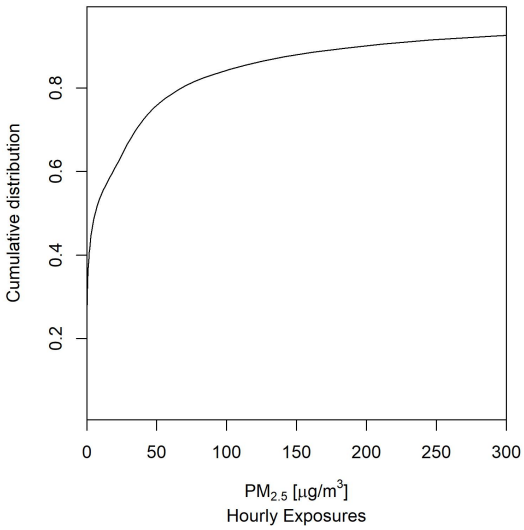
Figure 6.11: Windows open scenario. Predicted distribution of hourly $PM_{2.5}$ exposures for archetype ID27, for each region.



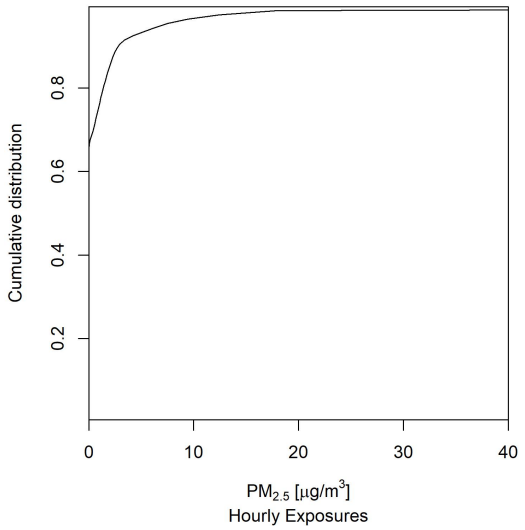
(a) Windows closed scenario: Archetype ID 27, by region.



(b) Windows open scenario: Archetype ID 27, by region.



(c) Windows closed scenario: Archetype ID 27, nationwide, regional stock–weighted.



(d) Windows open scenario: Archetype ID 27, nationwide, regional stock–weighted.

Figure 6.12: Example of one archetype, ID 27. Stock–weighted distributions of the predicted hourly exposures to $PM_{2.5}$.

6.2.2 Regional and nationwide outcome distributions for one of the archetypes

For each archetype, the national distribution was obtained by sampling values from the regional distributions and using their weighting factors; method presented in Section 5.5.1. Here, one example is shown. Figures 6.12a and 6.12b show the 15 regional cumulative distributions of $PM_{2.5}$ exposure for the first archetype. Figures 6.12c and 6.12d show the simulation results scaled up to the national stock.

6.2.3 Nationwide distributions

Similarly, the national distributions for windows closed and open scenarios were obtained using the eight archetypes. Figure 6.13 shows the national CDFs of the predicted hourly $PM_{2.5}$ concentrations for the two extreme scenarios. Given that neither of these two assumptions is likely to represent the actual scenario, the real line is expected to be somewhere between these two. This graph can be used to visualise the boundaries of the problem in the Chilean context, and to see the impact of the window opening behaviour on hourly exposures.

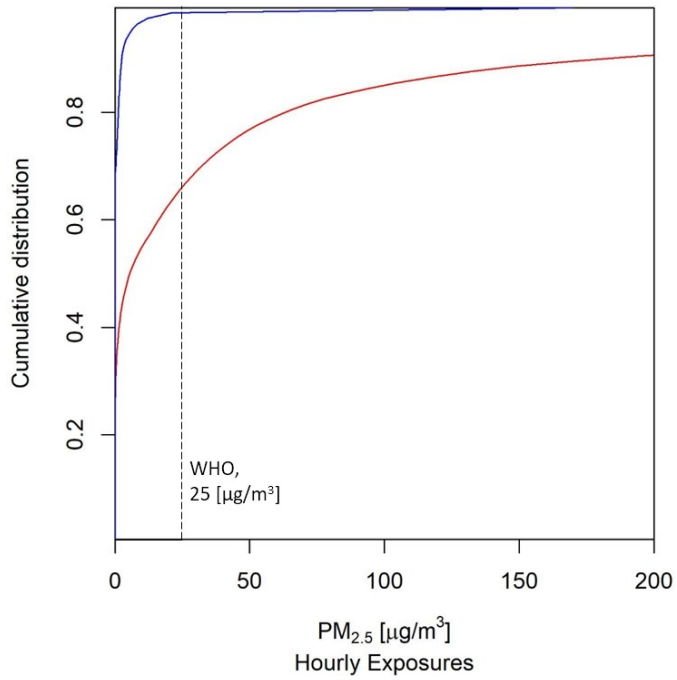
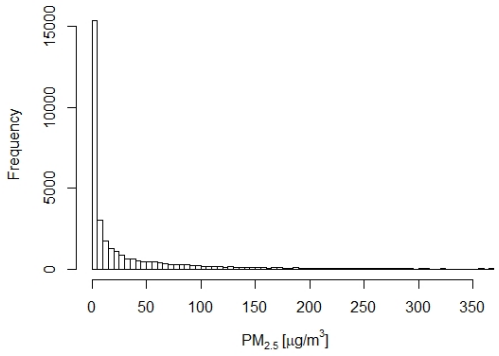


Figure 6.13: Nationwide hourly exposures to $PM_{2.5}$ during the winter season. Windows closed scenario in red line and Windows open scenario in blue. The dashed line shows the WHO's 24 h recommendation of 25 $\mu g/m^3$.

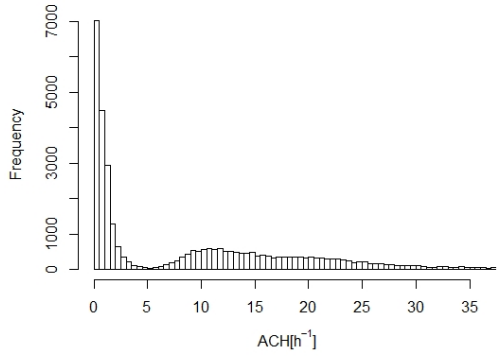
6.3 Winter data

To have a single value for each of the projects' outcomes, the total exposure to $PM_{2.5}$, the median ventilation rate, and the total heat loss over the wintertime were calculated for each house. Figures 6.14a, 6.14b, and 6.14c show the distribution of the results, including windows closed and windows open scenarios, before being stock–weighted.

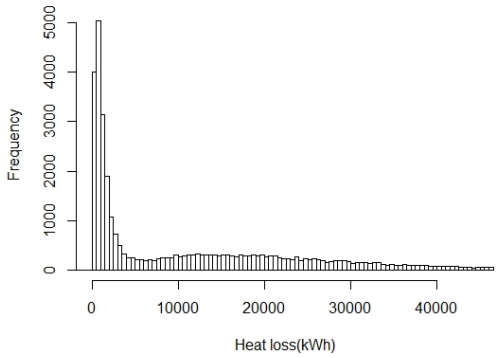
Figures 6.15a, to 6.15f show the distribution of the outcomes after they were weighted according to their frequency in the stock, and Table 6.3 presents some descriptive statistics. 90% intervals are presented to show the lower and upper limits of the pre-



(a) Distribution of total winter exposures to $PM_{2.5}$.



(b) Distribution of median ventilation rates over the winter



(c) Heat loss.

Figure 6.14: Distribution of the simulations for the three outcomes analysed in this study before weighting them. The x–axis were trimmed to display up to the 95% quantile.

dicted exposure levels, ventilation rates, and heat loss, and exclude those that are less likely to find. The coefficient of variation, C_V, is a descriptive statistic that measures the variability of any value and, due to its independence of the unit of measurement, can be used to compare different distributions. It is computed by dividing the standard deviation by the mean, or $\frac{\sigma}{\mu}$. For example, results for winter exposures are shown to be more variable than those for ventilation rates or heat loss, and the difference in variation between windows closed and open scenarios is similar across all outcomes. Interestingly, the lowest variability is seen in ventilation rates in the second scenario, with a C_V of 0.50, whereas the largest is in exposures, also of the second scenario, with a C_V of 1.85. The skewness and kurtosis describe the shape of the distributions, where 0 correspond to a normal distribution. The positive skewness and kurtosis for the three outcomes indicate that their distributions are all positively skewed and heavily–

tailed. This suggests that the use of the mean, instead of the median, as an indicator in policy–making or benchmarking could be problematic. This is further discussed in Section 7.5.2.

Table 6.3: Descriptive statistics for winter $PM_{2.5}$ exposures, ventilation rates, and heat loss, nationwide. Outcomes of the eight archetypes were scaled up to the national level using their regional and archetypal weighting factors. Scenario 1: Windows closed. Scenario 2: Windows open. C_V: coefficient of variation. P: percentile.

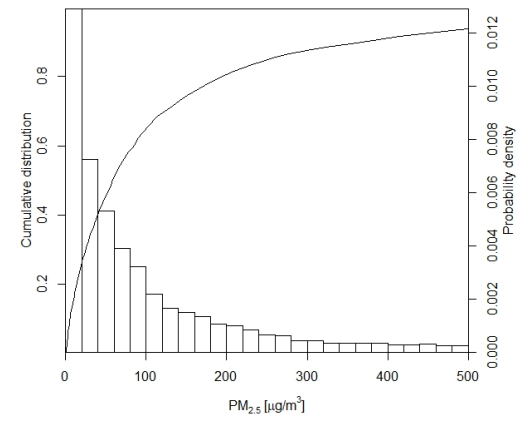
	Exposures (μg/m^3)		**Ventilation rates (h^{-1})**		**Heat loss (kWh)**	
Statistic	Scenario 1	Scenario 2	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Mean, μ	134.47	6.64	0.89	15.80	1283.85	19526.08
Median, M_{ed}	58.65	2.30	0.75	13.57	947.44	16681.37
Standard deviation, σ	210.28	12.26	0.79	7.91	1190.02	12556.3
90% CI	[2.58; 548.72]	[0.08; 29.81]	[0.08; 2.40]	[8.20, 29.75]	[252.6; 3471.2]	[5736; 42342]
P_{10}	5.20	0.22	0.13	8.85	319.3	7614
P_{25}	19.06	0.77	0.31	10.42	508.8	11053
P_{75}	154.20	6.77	1.16	19.14	1624.5	24686
P_{90}	367.06	16.94	1.84	24.75	2585.7	34295
C_V	1.56	1.85	0.89	0.50	0.93	0.64
Skewness	4.12	4.26	2.17	2.47	3.00	2.18
Kurtosis	34.62	26.36	9.19	11.31	16.07	9.48

6.3.1 Winter results by region and by archetype

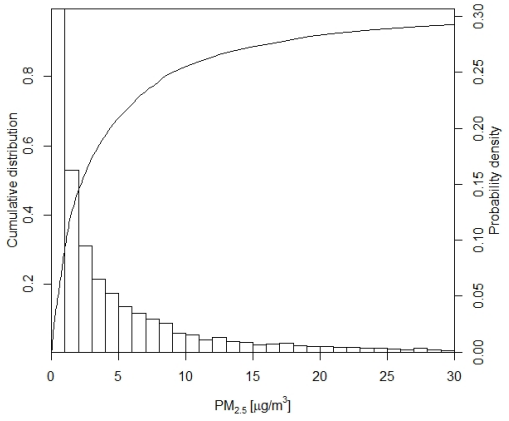
Simulation results were aggregated by region number, archetype ID, and heater type to show and test the variability between groups. Descriptive statistics for each outcome are shown in Tables 6.4 to 6.6, aggregated by region number, and in Tables 6.7 to 6.9 by archetype ID. Group comparison test statistics are reported in Section 6.3.3. A large difference between the averages and the medians of winter exposures is seen in the results of the archetypes and regions, indicating that their distributions are asymmetric and right–skewed. This difference is less pronounced in the results of heat loss, and not seen in the results of ventilation rates or heat loss. Figure 6.18 shows the distribution of the results by type of heater.

6.3.2 Mean values of the sampling distribution

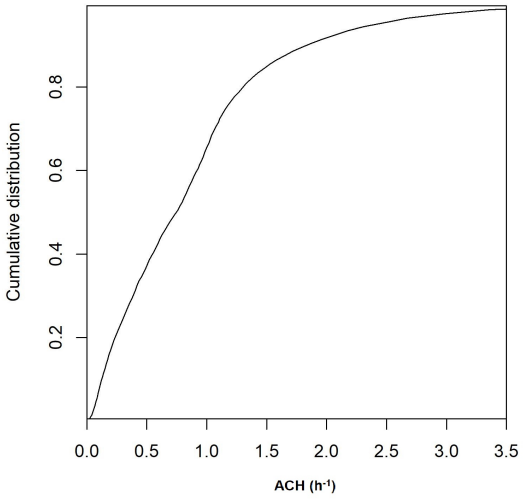
The uncertainty in the means of the three outcomes across the stock was identified by sampling multiple values from the national distribution. The samples size corresponds to the total housing stock. Due to the central limit theorem, the sampling distribution of the sample mean is expected to follow a normal distribution (Hughes & Hase, 2010). The number of samples is then increased until the test statistics of the Shapiro Test for normality become $W > 0.995$ and $p > 0.05$.



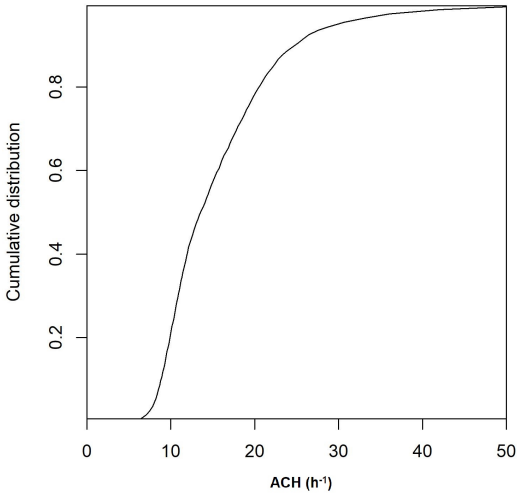
(a) $PM_{2.5}$, Windows closed scenario.



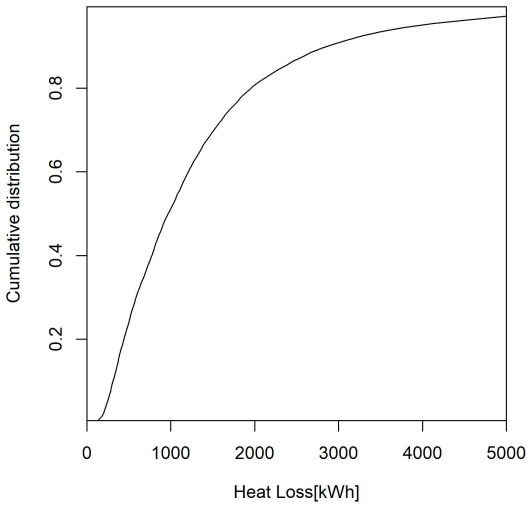
(b) $PM_{2.5}$, Windows open scenario



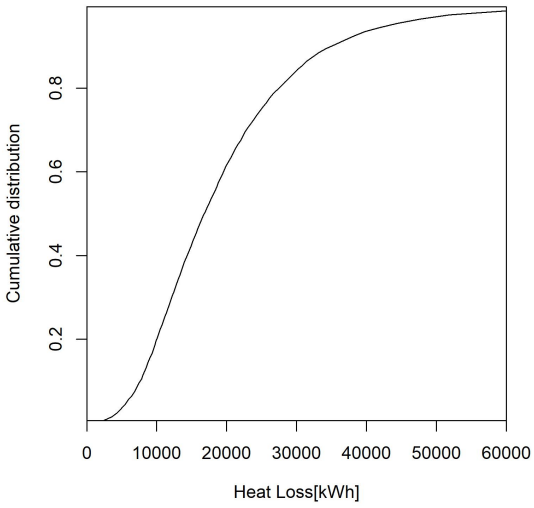
(c) Median ventilation rates, windows closed scenario.



(d) Median ventilation rates, windows open scenario.

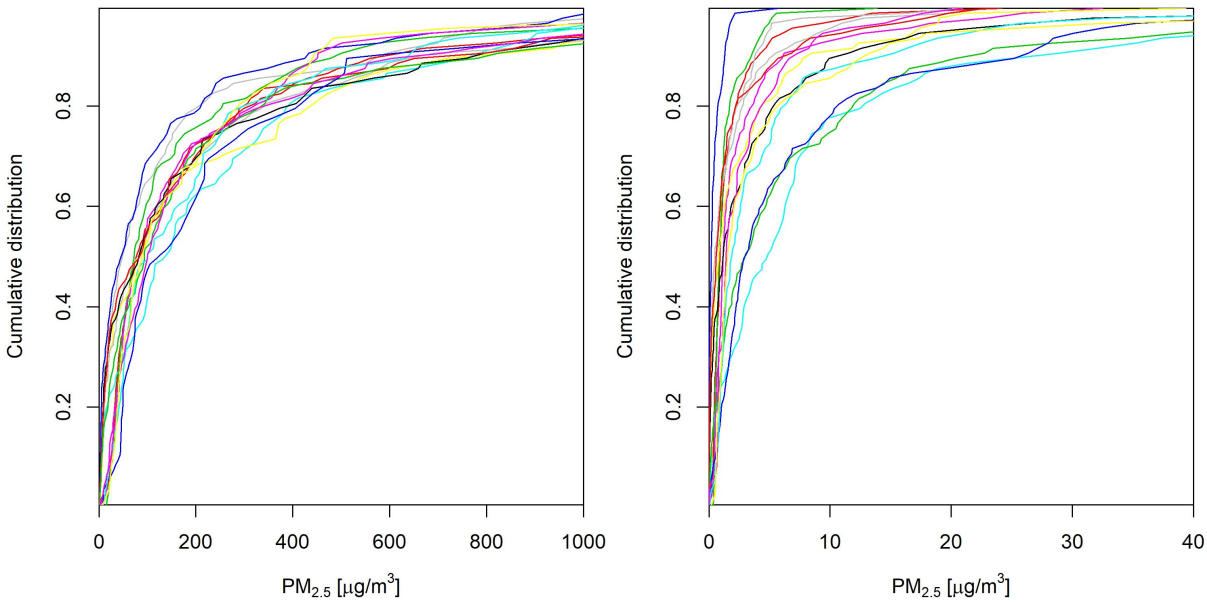


(e) Total heat loss over the winter, windows closed scenario.



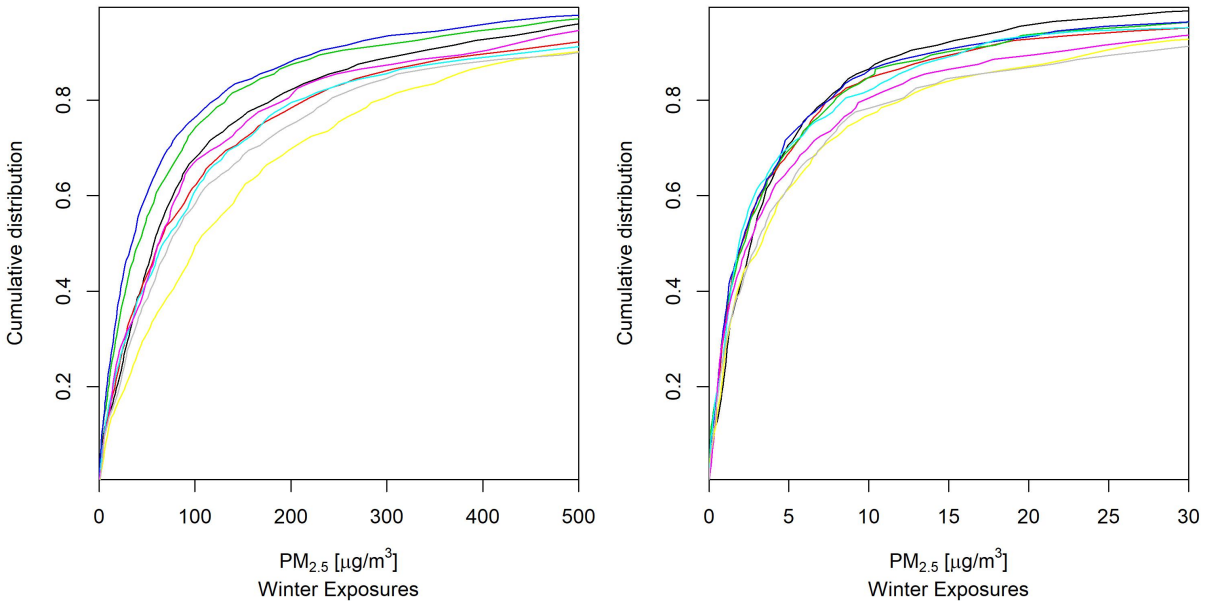
(f) Total heat loss over the winter, windows open scenario.

Figure 6.15: Nationwide cumulative distribution of the simulations for the three outcomes analysed in this study after weighting them.



(a) Windows closed scenario. **(b)** Windows open scenario.

Figure 6.16: Hourly exposures to $PM_{2.5}$ by region.



(a) Windows closed scenario. **(b)** Windows open scenario.

Figure 6.17: Exposures to $PM_{2.5}$ by archetype ID. ID 27 shown in red, the most common house; ID36 in blue; ID 91 in light blue; ID 100 in pink; ID 275 in yellow; ID 35 in green, ID19 in black; ID 284 in grey. Extreme values: In yellow and grey are ID 275 and ID284, the only two simulated archetypes from the second construction period and so air tighter; ID 275 is the smallest archetype of those two. In blue and green are the two of the largest houses by volume, from the first construction period.

Table 6.4: $PM_{2.5}$ exposures. Regional results.

Region	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Scenario 1															
μ	138.78	134.29	129.70	133.93	135.90	143.83	135.62	139.16	125.94	119.49	116.67	88.26	137.83	119.31	141.57
Med	48.29	46.76	51.04	51.73	55.89	70.02	62.13	60.41	55.90	48.69	49.61	30.53	63.03	50.04	53.65
σ	237.26	221.83	198.83	220.52	225.25	204.02	186.90	197.82	186.71	179.82	174.17	159.96	224.08	182.33	241.79
Scenario 2															
μ	2.49	1.91	4.45	2.33	8.87	7.95	5.3	3.21	5.26	2.74	2.07	0.83	9.95	3.84	3.41
Med	0.67	0.55	1.40	0.65	3.46	3.29	2.20	1.53	1.72	1.09	1.06	0.24	4.35	1.44	0.96
σ	5.08	3.55	9.52	4.50	15.15	13.32	9.32	4.13	10.13	4.96	2.96	1.73	15.52	6.78	7.73

Table 6.5: Ventilation rates. Regional results.

Region	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Scenario 1															
μ	0.59	0.60	0.74	0.70	0.86	0.86	0.88	0.81	0.93	1.00	1.12	1.24	0.94	1.14	0.47
Median	0.45	0.52	0.63	0.59	0.76	0.72	0.76	0.71	0.78	0.86	0.97	1.06	0.81	0.99	0.38
σ	0.68	0.55	0.66	0.60	0.72	0.78	0.75	0.65	0.79	0.82	0.88	1.04	0.86	0.97	0.41
Scenario 2															
μ	18.23	18.12	13.55	17.07	12.34	13.28	16.04	19.08	19.08	27.53	26.54	31.49	12.04	21.77	15.83
Median	17.52	17.49	12.90	16.84	11.36	11.91	14.87	17.96	18.61	24.17	25.70	29.25	10.91	20.70	14.65
SD	7.38	6.94	4.67	5.95	3.91	4.83	5.15	6.23	8.16	15.22	10.11	12.74	3.71	10.33	6.67

Table 6.6: Heat loss by ventilation. Regional results.

Region	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Scenario 1															
μ	980.80	935.26	1075.29	1035.02	1247.37	1299.59	1232.75	1178.93	1337.54	1425.55	1544.67	1672.90	1362.00	1582.67	803.70
Median	684.20	744.84	862.47	846.90	948.55	925.22	932.18	934.76	1003.69	1094.81	1198.05	1293.45	974.90	1195.36	617.40
σ	1150.04	814.04	910.37	835.41	1094.09	1216.32	1097.95	978.17	1132.46	1233.96	1254.82	1533.16	1330.46	1447.22	657.34
Scenario 2															
μ	22506.90	21624.46	16639.19	20282.04	15322.73	16580.19	19843.33	24595.73	23654.81	33727.24	32759.57	39123.06	14624.20	27145.27	19301.21
Median	20961.60	19266.54	15327.15	18708.38	13929.14	15010.62	18831.14	23019.82	20327.13	27610.76	29705.12	33674.12	13415.51	23586.90	17487.41
σ	11973.82	11730.28	8877.13	10831.40	7803.75	8940.46	10037.87	12071.80	14037.15	22032.91	17342.07	22758.61	7049.18	16618.80	10617.56

Table 6.7: $PM_{2.5}$ exposures. Archetype results.

Archetype ID	**27**	**36**	**91**	**100**	**275**	**35**	**19**	**284**
Scenario 1								
μ	154.24	83.75	159.60	124.73	186.13	93.60	116.37	167.70
Median	62.63	34.20	66.64	62.14	101.92	41.45	57.07	73.54
σ	267.92	133.80	248.38	166.32	222.88	142.21	153.91	240.88
Scenario 2								
μ	6.30	5.56	6.46	7.68	8.45	5.74	4.92	9.11
Median	2.03	2.02	1.89	2.47	3.28	2.25	2.63	3.02
σ	12.00	10.33	12.56	13.52	13.76	9.92	6.66	17.23

Table 6.8: Ventilation rates. Archetype results.

Archetype ID	**27**	**36**	**91**	**100**	**275**	**35**	**19**	**284**
Scenario 1								
μ	0.95	1.30	0.83	1.00	0.15	1.29	0.89	0.26
Median	0.80	1.02	0.69	1.00	0.12	1.05	0.74	0.20
σ	0.70	1.09	0.59	0.18	0.10	0.99	0.65	0.20
Scenario 2								
μ	17.68	11.91	13.81	19.64	13.70	13.37	23.69	13.38
Median	14.40	10.07	12.13	19.06	10.88	11.39	21.51	10.87
σ	9.57	5.38	5.38	4.53	7.61	6.07	8.69	7.00

Table 6.9: Heat loss by ventilation. Regional results.

Archetype ID	**27**	**36**	**91**	**100**	**275**	**35**	**19**	**284**
Scenario 1								
μ	1120.92	2078.32	623.58	1537.34	368.60	2127.30	1440.06	779.71
Median	911.25	1561.00	499.01	1483.47	343.48	1603.01	1202.91	683.38
σ	740.21	1779.12	434.90	507.93	135.59	1684.55	924.26	415.38
Scenario 2								
μ	18100.14	18201.86	9194.45	23918.00	15838.53	20364.67	31512.01	26418.84
Median	15304.56	15923.77	8258.63	22631.24	12904.58	18081.72	28806.49	22000.75
σ	10769.93	9849.44	5503.07	10069.06	9643.45	11290.55	14649.54	16442.10

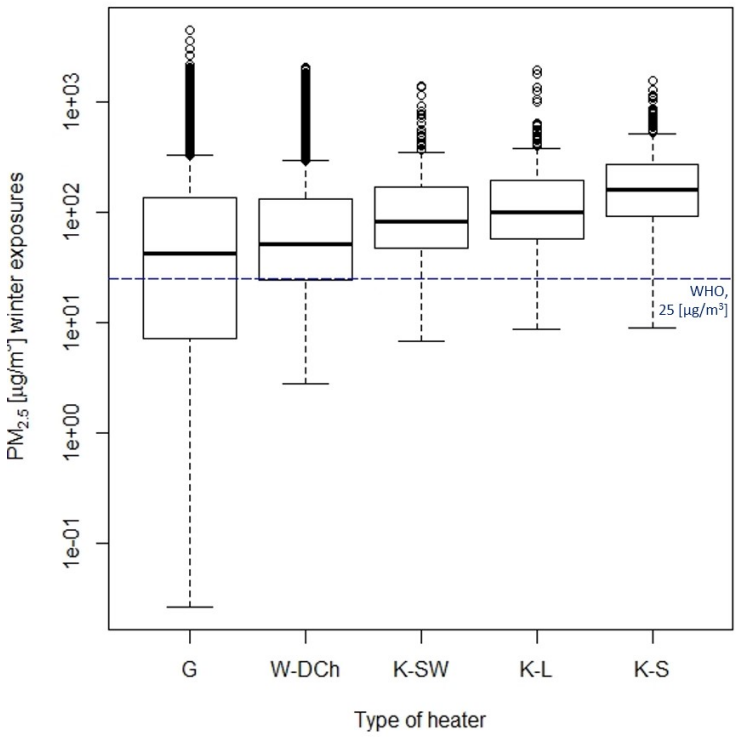
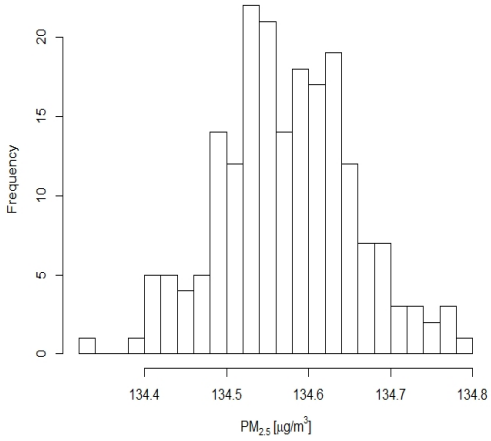
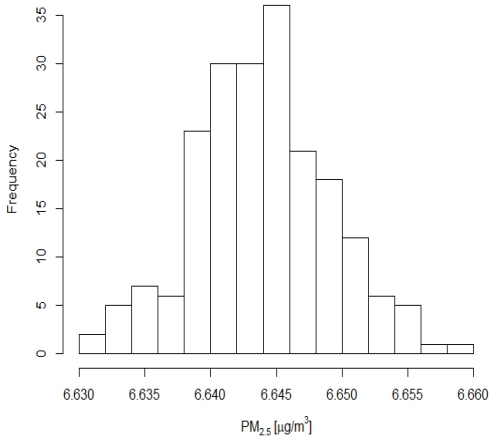


Figure 6.18: Windows closed scenario. Distribution of the medians of winter $PM_{2.5}$ exposures for all simulated houses by heater type. y–axis in log–scale. G: gas; W-DCh: wood double chamber; K-SW: kerosene special wick; K-L: kerosene laser; K-S: kerosene standard. The dashed line shows the WHO's 24 h recommendation of $25\,\mu g/m^3$.

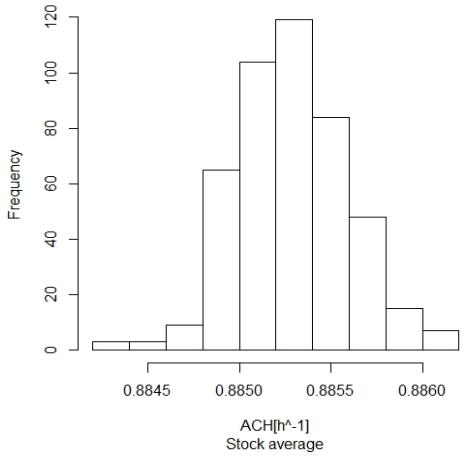
Figures 6.19a to 6.19f show the distribution of the sampling means from which the population mean for the outcomes are estimated. The mean $PM_{2.5}$ concentration in a Chilean house is estimated to be $\mu = 134.6\,\mu g/m^3$ ($SE = 0.08\,\mu g/m^3$) with all windows closed and $\mu = 6.65\,\mu g/m^3$ ($SE = 0.0049\,\mu g/m^3$) with all windows open. The mean ventilation rate of a Chilean house is estimated to be $\mu = 0.885\,h^{-1}$ ($SE = 3.12e-04\,h^{-1}$) with all windows closed and $\mu = 15.8\,h^{-1}$ ($SE = 3.14e-03\,h^{-1}$) with all windows open. The mean heat loss (by exfiltration) per house is estimated to be $\mu = 1,284$ kWh ($SE =$ kWh) and 19,510 kWh ($SE = 5.36$ kWh) for windows closed and open scenarios, respectively. Finally, the total heat loss due to exfiltration in the Chilean stock is estimated to be 8.16 TWh for windows closed scenario, and 124 TWh for windows open scenario. Some negative values were obtained in some houses showing heat gains. Negative values were set to zero to account only for the heat loss.These figures are compared against those from other studies in Section 7.7, and against the international guidelines and benchmarks in Section 7.5.3.



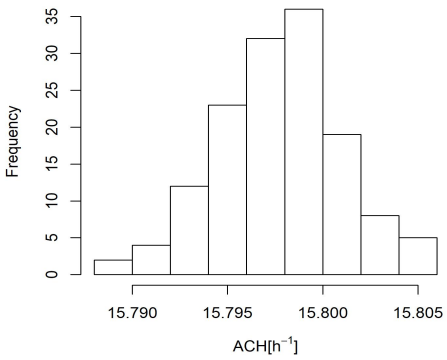
(a) Average $PM_{2.5}$ concentration per house across the stock. Windows closed scenario.



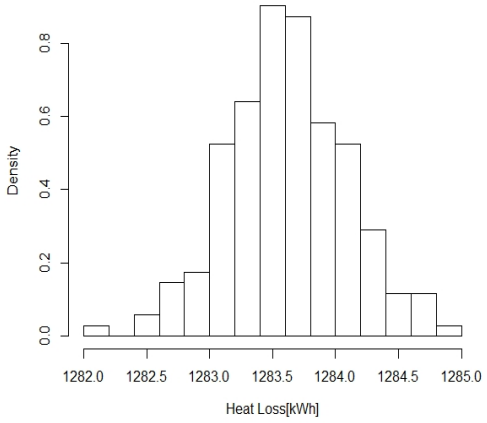
(b) Average $PM_{2.5}$ concentration per house across the stock. Windows open scenario.



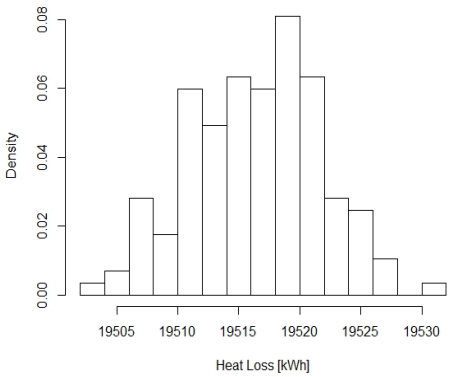
(c) Average ventilation rate per house across the stock. Windows closed scenario.



(d) Average ventilation rate per house across the stock. Windows open scenario.



(e) Average heat loss per house across the stock. Windows closed scenario.



(f) Average heat loss per house across the stock. Windows open scenario.

Figure 6.19: Distribution of the sampling mean for the three outcomes and the two extreme scenarios.

6.3.3 Statistical tests for group comparison

A series of statistical tests are performed to compare $PM_{2.5}$ exposures, ventilation rates, and heat loss when analysed by archetype ID and region, and when compared against each other. Kruskal–Wallis tests by ranks H showed that there is a significant difference in the distributions between the groups, and the Levene's tests showed that there is a significant difference in the variances in the groups. The LHS sampling method has the advantage of testing the inputs in all the archetypes equally, and so similar variability in the outcomes could be expected. Thus, these results of the Levene's tests for the three outcomes mean that the model inputs affect the archetypes differently, giving different inter–group variability.

To give a better interpretation of the statistical significance of these results, the effect sizes are quantified using the medians. The results show varied magnitudes for the three outcomes. *Negligible* to *small* effect sizes are seen between pairs of archetypes and regions for the $PM_{2.5}$ exposures, and *negligible* to *large* effect sizes for the ventilation rates and heat loss. When *negligible* effect size is seen between two groups and in the three outcomes, the analysis and interventions could be carried out at both groups together with equal effect.

See Section 5.5.5 for a brief description of the method, and Sections 7.2 and 7.6.3 for the discussion of the tests and these results. The results of the statistical tests for the three outcomes are presented in greater detail in the following sections.

Exposures to $PM_{2.5}$ ($\mu g/m^3$):

The results of the Kruskal–Wallis test by ranks were highly significant for both archetypes and regions ($p < 2.2e-16$). When results were pairwise compared by region, 57 and 16 out of 105 combinations were non–significant for windows closed and open scenarios, respectively. When comparing windows closed scenario by pairs of archetype IDs, 20 out of 28 combinations were significant at 0.05 level. 16 were significant when comparing windows open scenario.

Levene's F test shows that the variances of the median winter exposures to $PM_{2.5}$ for windows closed and open scenarios were different in the archetypes, $F(7, 17222) = 38.79$, $p < 2.2e-16$ and $F(7, 17382) = 10.21$, $p = 7.868e-13$, respectively. Variances were also significantly different in the regions for windows closed and open scenarios, $F(14, 17215) = 5.60$, $p = 6.113e-11$ and $F(14, 17375) = 65.82$, $p < 2.2e-16$.

The magnitude of the effects of median $PM_{2.5}$ exposures between regions is *negligible* for all combinations for windows closed scenario, and for 85 out of 105 combinations for windows open scenario. The remaining 20 combinations showed a *small*

effect size, especially between the 13th (capital) region versus northern and southern regions (1–4, and 8–12). The largest effect size for each scenario is seen between the 6th and 12th region for windows closed scenario, and for the 12th and 13th for windows open scenario. For each scenario the effect sizes of $PM_{2.5}$ between pairs of archetypes are generally of *negligible* in magnitude, and *small* only when archetype ID 275 is compared against ID 35 and 36 for windows closed scenario.

Ventilation rates (h^{-1}) during winter:

The Kruskal–Wallis test by ranks H was highly significant ($p < 2.2e-16$) for both scenarios when compared by IDs and region ($p < 2.2e-16$), meaning that there is a difference between the groups. The *Post–hoc* test for multiple comparisons between pairs of regions showed that 75 out of 105 combinations were significantly different from each other, and 92 out of 105 combinations in windows open scenario. Between archetype IDs, ventilation rates in windows closed scenario are significantly different for all combinations except for three of them: ID 35 versus ID 36 and ID 100, and between ID 36 and ID 100. In windows open scenario, the only three combinations with non–significant difference are ID 35 versus ID 91 and ID 284, and between ID 91 and ID 284.

Levene's F test shows that the variances of ventilation rates for windows closed scenario were significantly different when grouping the results by either archetype ID and region, $F(7, 17222) = 428.61$ ($p < 2.2e-16$) and $F(14, 17215) = 53.942$ ($p < 2.2e-16$), respectively. Similarly for windows open scenario, variances of ventilation rates showed significant differences between archetypes and regions, $F(7, 17382) = 102.6$ ($p < 2.2e-16$) and $F(14, 17375) = 273.12$ ($p < 2.2e-16$), respectively.

The magnitude of the effects of ventilation rates between regions is between *negligible* to *medium* for combinations in windows closed scenario, and to *large* for windows open scenario. Effects of *large* magnitude are seen in 13 combinations, only in windows open scenario. *Medium* effects are shown in 8 and 18 out of 105 combinations in windows closed and open scenario, respectively.

Effect sizes of ventilation rates between pairs of archetypes are between *negligible* and *large* for windows closed scenario and all of *negligible* magnitude for windows open scenario. In windows closed scenario, 9 out of 28 combinations are *small*, 5 are *medium*, and 6 are *large*. *Large* effect is shown for archetypes 275 and 284 versus archetypes 35, 36, and 100.

Heat loss (kWh) by ventilation:

The Kruskal–Wallis test by ranks H was found to be highly significant ($p < 2.2e-16$) for both scenarios when comparing the distributions of the heat loss with archetypes and regions. *Post–hoc* test for multiple comparisons for windows closed scenario showed that 76 out of 105 pairs of regions were significantly different from each other, and 88 out of 105 for windows open scenario. For pairs of archetypes in windows closed scenario, only ID 35 versus ID 36 and 100, and ID 36 versus ID 100 were non–significant, and between ID 27 and ID 36, and between ID 100 and ID 285 for windows open scenario.

Levene's F test of homoscedasticity showed that the variances in heat loss for windows closed scenario were significantly different when grouping the results by either archetype or region; $F(7, 17222) = 468.34$ ($p < 2.2e-16$) and $F(14, 17215) = 38.11$ ($p < 2.2e-16$), respectively. Similarly for windows open scenario, variances of ventilation rates showed significant differences between archetypes and regions, $F(7, 17382) = 124.09$ ($p < 2.2e-16$) and $F(14, 17375) = 157.22$ ($p < 2.2e-16$), respectively. This means that the groups show significantly different medians and variances.

In windows closed scenario, effect sizes between pairs of regions ranged from *negligible* to *medium*. Medium effects were seen in region 11 versus regions 1, 2, 4, and 15. In windows open scenario, the magnitudes of the effect were between *negligible* to *large*; 45 cases had *small* magnitude, 18 *medium*, and 12 *large*. *Large* effects were seen in some combinations of regions 1–7 and 10–13 versus region 15. Effect sizes of heat loss between archetypes were more variable, ranging from *negligible* to *large* for windows closed and open scenarios.

6.4 Relationship between outcomes

The third objective of this study (Section 1.7) was to explore the relationship between ventilation rates, pollutant concentration, and energy loss in Chilean houses. The literature review (Section 1.5) presented the argument for and against an increase in ventilation rates. The two conflicting aims are illustrated in Figure 1.2 on page 8, which shows the monotonic relationship between ventilation rate and pollutant concentration and the linear relationship between ventilation rate and energy loss. The figure is replicated here for the Chilean stock to contextualise the results and to identify the nature of the relationships. Figure 6.20 shows the interaction between the three predicted outcomes in a scatter plot, where all predicted outputs are included.

Figure 6.21 uses the same data to show the P_{95} of both $PM_{2.5}$ and heat loss, versus the associated ventilation rate. To ensure that 95% of the stock is below 25 $\mu g/m^3$, the ventilation rates must be increased up to 13 h^{-1}, which can be cost–prohibitive. This

ventilation rate corresponds roughly to the median of windows open scenario and out of the range of windows closed scenario. These results show the range of scattering of the predictions for different conditions, and also the uncertainty in data so that it can be accounted for in any decision–making process. This is discussed in Section 7.5.2.

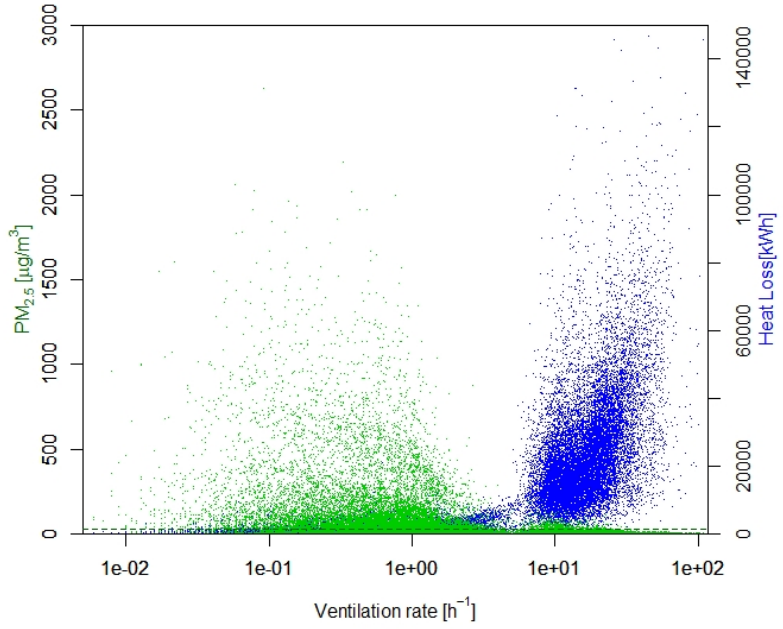


Figure 6.20: Scatter plot of the ventilation rates and the predicted winter exposures to $PM_{2.5}$ (in green) and total energy loss (in blue) for both scenarios. The green dashed line shows the WHO's 24 h recommendation of 25 $\mu g/m^3$.

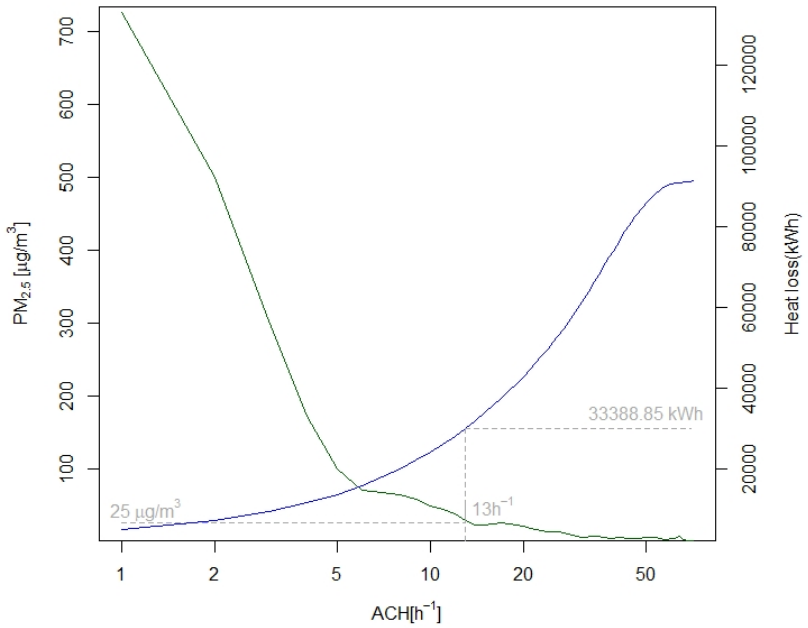


Figure 6.21: P_{95} of the predicted winter exposures to $PM_{2.5}$ (in green) and the total energy loss (in blue) versus the ventilation rates. Windows closed and open combined. The dashed lines show the WHO's annual recommendation of 25 $\mu g/m^3$ and the related heat loss.

6.5 Sensitivity Analyses

In this section the sensitivity of the model's three outputs to its inputs are evaluated; see Section 5.6 for a description of the method. The three output metrics computed for each house were: (i) the total exposure of one occupant to fine particles, (ii) the median ventilation rates, and (iii) the total heat loss over the winter. These metrics are used to measure the strength of the relationships. This analysis enabled the identification of the inputs that are most influential so that they can be targeted for attention when designing new dwellings or by future data gathering exercises when there is epistemic uncertainty in influential model inputs.

Figures 6.22, 6.23, and 6.24 show the relationships between the sampled inputs and the three outcomes, including scenarios 1 and 2. No clear relationship is shown between the three outputs and the aspect ratio or orientation of the building, whereas relationships between the three outputs and the total leakage area and Q_{50} look exponentially related. An approximately linear relationship is seen between the winter exposures and the cooking emission rates. The relationships between the winter exposures and the deposition rate, breakfast emission rates, and ΔT look parabolic, and more scattered for smaller values. No pattern is observed between winter exposures and wind speed, the S:V ratio, or the heater emission rates. A linear relationship is seen between ΔT and both the ventilation rate and heat loss. A random relationship is seen between the n exponent and the heat loss, whereas the relationship between the n exponent and the ventilation rate may be parabolic.

Tables 6.10—6.12 present the rankings of the inputs by sensitivity test, where 1 is the most sensitive. The test statistics used to compute the ranking along with their p–values are presented in Appendix D. Results of the tests show that $PM_{2.5}$ winter exposures are most strongly correlated with the cooking emission rates, followed by the envelope permeable area, and the heater emission rates. Ventilation rates are most strongly correlated with the permeable envelope area, Q_{50}, and ΔT. Finally, heat loss is also most sensitive to the permeable envelope area, but ΔT is the second–ranked input, and Q_{50} is the third.

The consequences of these results are discussed further in Section 7.6.

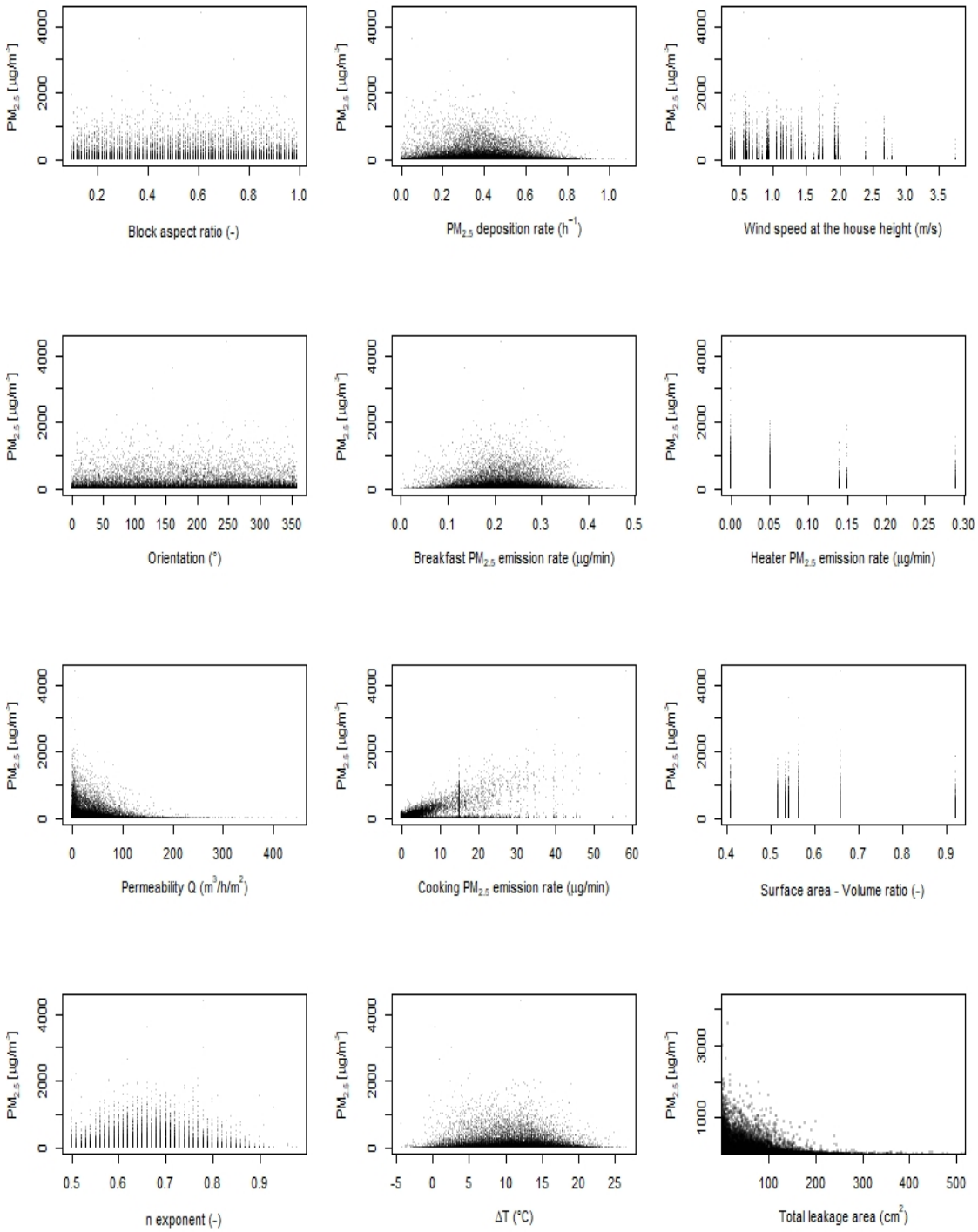


Figure 6.22: Scatter plot of the inputs and the predicted winter exposures to $PM_{2.5}$ ($N = 34,620$).

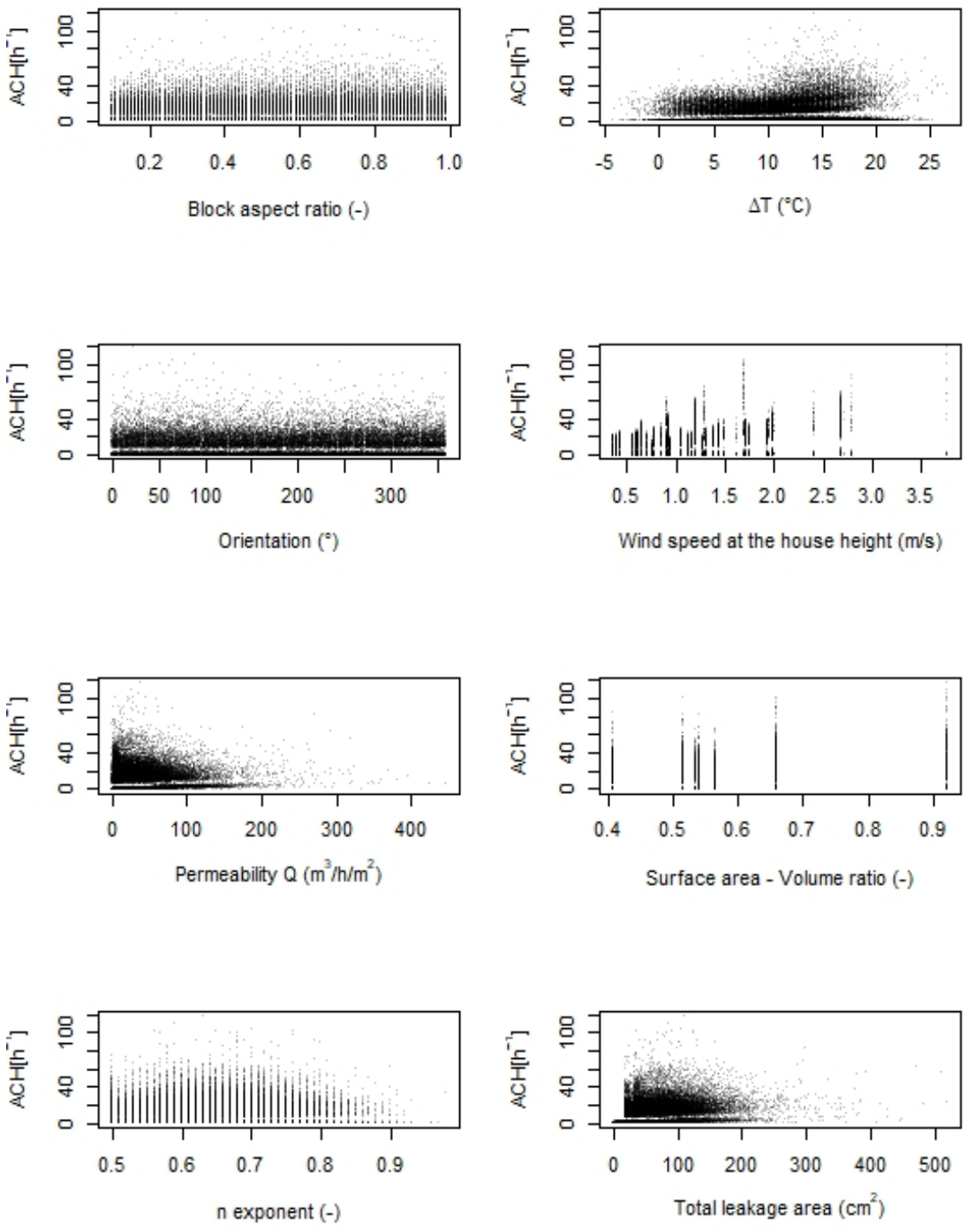


Figure 6.23: Scatter plot of the inputs and the predicted median ventilation rates over the winter.

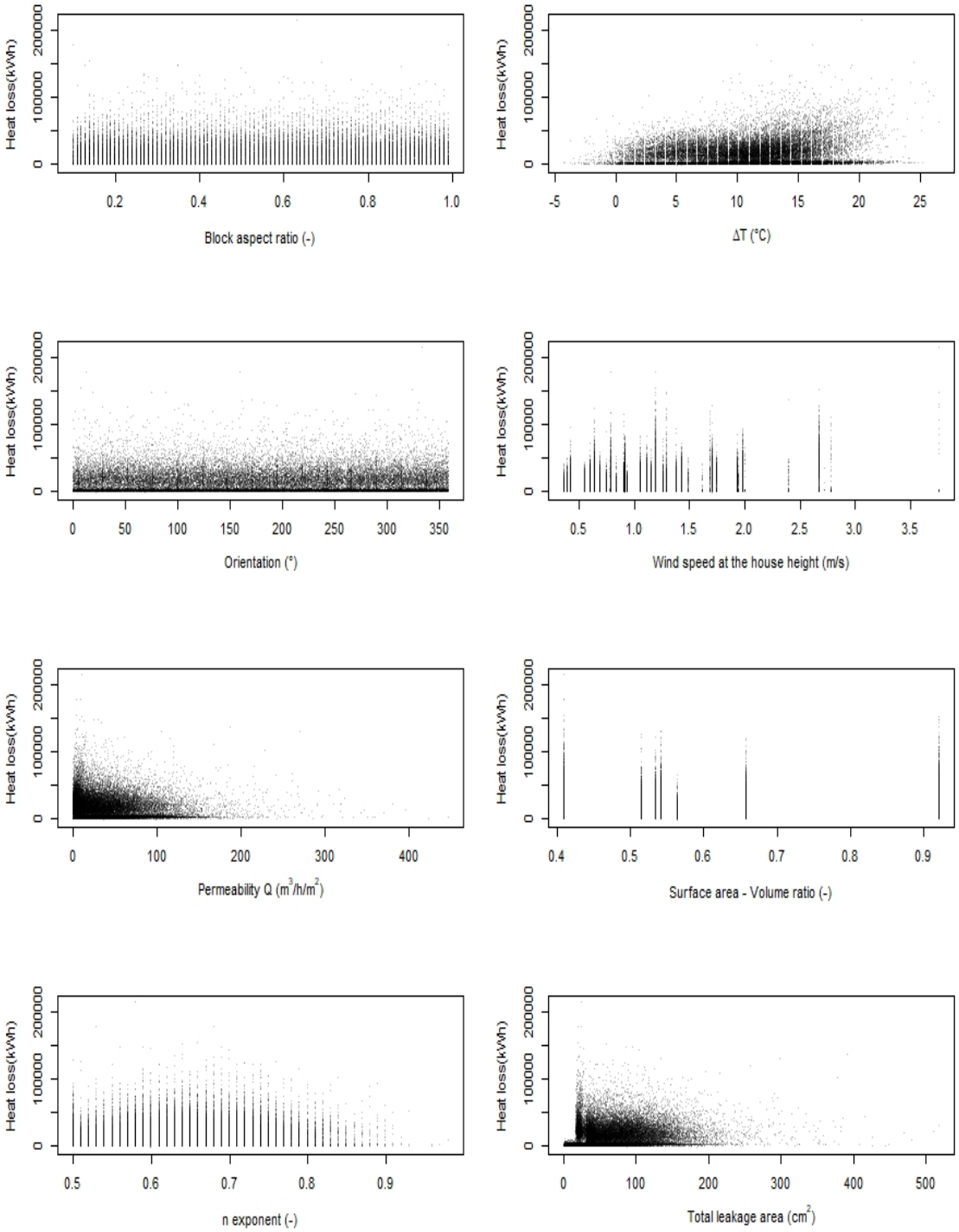


Figure 6.24: Scatter plot of the inputs and the predicted total heat loss over the winter season.

Table 6.10: Sensitivity analyses of the input parameters on the exposure to $PM_{2.5}$ over the winter.

Input — Test rank	Kendall τ	Pearson	Spearman	Regression	Rank Regression	K–S	K–W P_2	K–W P_5	K–W P_{10}	K–W P_{20}
$L:W$	10	9	10	9	10	11	11	11	10	12
Orientation	11	12	11	12	11	10	10	10	12	10
Permeability Q_{50}	5	3	5	3	5	6	5	5	6	6
n exponent	6	5	6	5	6	9	6	7	8	8
Deposition Rate of $PM_{2.5}$	7	4	7	4	7	8	7	9	7	9
Breakfast $PM_{2.5}$	12	11	12	11	12	12	12	12	11	11
Cooking $PM_{2.5}$	1	1	1	1	1	1	1	1	1	1
Δ T	8	7	8	7	8	7	9	6	5	5
Wind at House	4	10	4	10	4	5	3	4	4	4
Heater $PM_{2.5}$ER	3	6	3	6	3	4	4	3	3	3
$S:V$	9	8	9	8	9	3	8	8	9	7
Total permeable area	2	2	2	2	2	2	2	2	2	2

Table 6.11: Sensitivity analyses of the input parameters on ventilation rate.

Input — Test rank	Kendall τ	Pearson	Spearman	Regression	Rank Regression	K–S	K–W P_2	K–W P_5	K–W P_{10}	K–W P_{20}
$L:W$	7	5	7	5	7	7	8	7	7	7
Orientation	8	8	8	8	8	8	7	8	8	8
Permeability Q_{50}	2	7	2	7	2	2	2	2	2	2
n exponent	6	6	6	6	6	6	4	6	6	6
Δ T	3	4	3	4	3	5	3	5	5	5
Wind at House	4	2	5	2	5	4	6	3	3	3
$S:V$	5	3	4	3	4	3	5	4	4	4
Total permeable area	1	1	1	1	1	1	1	1	1	1

Table 6.12: Sensitivity analyses of the input parameters on heat loss.

Input — Test rank	Kendall τ	Pearson	Spearman	Regression	Rank Regression	K–S	K–W P_2	K–W P_5	K–W P_{10}	K–W P_{20}
$L:W$	7	6	7	6	7	7	8	8	7	8
Orientation	8	8	8	8	8	8	7	7	8	7
Permeability Q_{50}	3	5	3	5	3	3	2	2	2	2
n exponent	6	7	5	7	5	6	5	6	6	6
Δ T	2	2	2	2	2	5	3	3	3	3
Wind at House	4	3	4	3	4	4	4	4	4	4
$S:V$	5	4	6	4	6	2	6	5	5	5
Total permeable area	1	1	1	1	1	1	1	1	1	1

CHAPTER 7

Discussion

This chapter discusses the sources of information and archetypes presented in Chapters 3 and 4, and of the simulation model and the model inputs given in Chapter 5. It interprets the predictions of the Chilean housing stock model described in Chapter 6, and compares them with international recommendations and benchmarks.

7.1 Sources of information

The accuracy of the predictions of any model is a function of the quality of the input data. Chapter 3 shows that there are a number of data sources that describe Chilean houses and their occupants, and that the quality of the data is good enough to derive archetypes. However, there are some areas where the data can be improved. There is currently inconsistency between questionnaires, and so a collaborative approach between surveying organisations would make linking their outputs easier. Moreover, surveys can be improved to reduce uncertainty in some parameters.

There are also a number of parameters and metrics that are not surveyed by the census that would be useful to know to construct basic steady–state energy demand models, similar to the UK's Cambridge Housing Model (Jones *et al.* , 2015; Sousa *et al.* , 2018). These include dwelling properties, such as geometries (floor area and volume), window area and glazing type, year of construction, insulation level, internal air temperature (for instance, indicated by a thermostat setting), orientation, and heating system fuel. It would also be useful to understand some of the type and frequency of occupant activities, such as appliance use, cooking, or tobacco consumption, to improve estimations of energy demand, GHG emissions, IAQ, and occupant health risks.

Finally, future fieldwork is needed to develop a database that can be used to identify construction quality, and the contribution of infiltration to the Chilean national energy demand and GHG emissions.

7.2 Archetypes

The current Chilean housing stock is defined using archetypes in Chapter 4. The method used to develop them used all available sources of information that describe the houses.

The results show that 496 archetypes can be used to represent the entire stock. The number of archetypes reflects the level of resolution provided by the datasets and the characteristics of the chosen parameters. As the key parameters change or the available information about them improves, the number of cells, and so the number of archetypes, are expected to change accordingly. The identification of the key parameters is then relevant for the classification of the stock. In this study, indoor air quality and energy demand are taken into account in the selection process and so if they are to be used to assess other criteria this must be accounted for.

Because there was variability expected in some archetype parameters, and its impact varies according to the model used to make predictions, a stochastic approach is chosen to quantify the extent of this variability. The results are then valid for the entire stock the archetypes, but cannot be transferred to other stocks unless they share similar input values and distributions.

The uncertainty in each parameter is explored by varying them between known limits, and running multiple simulations to give a range of outputs. The Latin Hypercube Sampling method has the advantage of optimising the choice of samples, and so ensures that the full variance in each parameter is accounted for and also reduces the number of simulations required for solution convergence; see Section 5.4. The simulation process was stopped when the output of interest, $\overline{PM}_{2.5}$, converged so that its mean and standard deviation did not change as new data was added; see Section 5.4. Here, each region was tested separately. While this criteria is arguable, there is no established method in the literature. Consequently, other criteria could have been applied; for instance, to the whole country rather than to each region, or it could have used a different metric or threshold. The criteria regarding the metric and thresholds used in this study have been subject to peer–review, and were used by Jones *et al.* (2015), and because their approach has similarity to this study, it is considered adequate.

An inter–group and between group comparison analysis was carried out for each outcome using the medians, distributions, and variances. When comparing the model predictions by archetype ID and region, and being cautious about the interpretation of p–values for significance (see Section 7.6), the Levene's and Kruskal–Wallis statistical tests are all significant at $p \ll .001$ and $p \ll .05$ respectively (see Section 6.3.3), meaning that there is a significant difference in the variances within the groups and in

the distributions of groups. When calculating the effect sizes using the medians and Cohen's *d* thresholds, the indices are highly variable between pairs of IDs and regions, and for the scenarios. Thus, the statistical significance of the tests, in addition to the magnitude of the effects, generate confidence in the use of the archetypes as a tool for analysing different types of house. It would not be acceptable to aggregate the entire Chilean housing stock into a single archetype. Some negligible effects sizes are seen between some neighbouring regions for the three outcomes, see Section E (page 207) for a detailed list, suggesting that they could be joined by proximity and analysed together. This reduction may save time and computational resources. This also means that interventions can be targeted at all those houses and similar effects should be expected.

7.2.1 Comparison with MINVU's archetypes

The Ministry of Housing (MINVU) is currently developing its own housing archetypes. Information about the method and categorisation process was obtained by private communication and is used to compare them against the sets of archetypes presented (PUCV, 2019) in Chapter 4. MINVU found that 13 archetypes can represent the entire Chilean housing stock; 5 detached and 4 semi–detached/terrace houses, and 4 apartments. This represents a significant reduction in the number of archetypes proposed here. However, both studies share similar classification, categorisation, data manipulation, and data analysis methods. They also both found issues concerning the datasets, and similar typographical and coding errors, although the criterion for data cleaning and choosing the representative values differ. The geometry of the first 7 archetypes are similar for both studies, with some differences in the number of storeys.

Areas with important differences include the sources of information analysed and the categories used for aggregating the data. MINVU included two more datasets, the first from the council tax department (SII), and the second from houses that went through a voluntary qualification process to obtain an energy label according to their design (CEV). The first source did not provide any additional information, whereas the second gave information about windows (namely *window* : *wall* ratios and sunlight condition[1]) from 37,724 houses built after 2009, which corresponds to 0.7% of the current stock. The authors did not provide an analysis of the quality of this data set; however, some bias may be expected against older and less energy–efficient houses, and other locations and economic sectors where this qualification is in low demand. The use of this data may be problematic, and therefore an uncertainty analysis should be added.

[1] *Sunlight condition* is a metric based on the direct solar gain through windows. The method used to calculate it is unknown.

Section 4.1 outlines the classification process and shows that six categories or key parameters were used, whereas MINVU's process used nine; those that were only relevant to the energy use for heating. Finally, neither a literature review nor an explanation for selecting the key variables were given. The proposed hierarchy trees of the key parameters between the two studies are different; Figure 4.2 on page 70 shows the one used here. The impact of this is a change in the number of cells, as the categories and resolution of the variables vary. For instance, in the classification process MINVU used the *geometry* variable before the *year of construction*; so new houses gained greater prominence in comparison with those developed here.

This study gives a weighting factor to each archetype ID and geographical region, whereas MINVU's method used the thermal zone classification as the main key factor, and assigned the percentage of representation to each zone. The use of thermal zones[2] may be advantageous due to its influence in energy use but not for indoor air quality. Moreover, the lack of weighting factor by region may be problematic for carrying out a separate (regional) analysis. Chile is divided administratively and politically into geographical regions, and so public policies could be more easily applied and checked by employing them. At present, the simulations run with the archetypes include the environmental parameters using the local weather data, but future work might add the thermal zones in the regional database.

Another difference is that MINVU used three periods of construction instead of two. They added an age band for houses partly insulated, those built between 2001 and 2007 that comprises 8.5% of the current stock, although the distribution of the bands are not given explicitly for each archetype.

Conversely, the archetypes presented here include information about the number and type of rooms, number and socioeconomic status of occupants, and type of cooking and heating fuel, whereas MINVU's archetypes do not. Their lack of these variables may be because they used the problematic 2017 census (see Section 3.2), which did not record this data. Nevertheless, both sets of archetypes could be joined to complement one–another and represent information that is unique to each study.

7.3 Model inputs

This section discusses the limitations in the selection and use of model the input data and their effects on the results. CONTAM is a deterministic model that uses a single set of input values to calculate a single prediction. However, by systematically varying the inputs between known limits and making many calculations it is possible to

[2] *Thermal zones* are classified according to the number of heating degree days and local wind conditions.

explore the uncertainty in outputs, such as pollutant concentrations. Thus, to determine uncertainty in the outputs a knowledge of the uncertainty in the inputs is also required. However, a lack of knowledge has made necessary to make some assumptions about inputs.

7.3.1 Weather data and local environment

Given the limited locations with weather data, the simulations were run using the weather from the capital city of each regions, where most of the population is concentrated. Therefore, the ability to represent the diverse conditions in each region may be compromised by this simplification. Nevertheless, characterising a wider climatic context will require a higher number of simulations to meet the convergence criterion, which in turn means increased computational time and resources.

Major assumptions are made about the local environment of the dwellings, particularly their local environmental sheltering and terrain types, which need to be validated with empirical data. The parameters are applied to the Swami & Chandra's wind pressure coefficient models to define wind pressure profiles for each walls of a dwelling. The location is also used to scale the wind speed at dwelling height. This study uses the location of each dwelling, such as *urban* or *rural*, and then assumed that they are equally distributed into the sub–levels; see Section 4.4.1. Moreover, none of the references provide an exact description of each terrain type that can be used to classify site locations. In future studies, a description could be linked to measured data, such as to the spatial density and height of the buildings. The sensitivity analyses show wind speed as an important input parameter, ranked second for ventilation rates and heat loss, and fifth for exposures, and so these assumptions need to be verified with fieldwork to reduce uncertainty in the parameter and the predictions.

7.3.2 Envelope air permeability

The predicted air permeability of the envelope for Chilean houses follows Chan's model, a method for predicting the air permeability of USA houses (Chan *et al.* , 2013). These predictions are considered an approximate model that applies solely to the current dwelling stock rather than to individual houses, but it can be updated as new dwellings are added to the stock.

The results of the linear regression between the measured and predicted values gave the coefficients of determination, R^2, of 0.62 and 0.57 for old and new houses, which indicate a *strong* and *modest* correlation, respectively. Nevertheless, the error analysis shows that the model is less accurate for those houses at the extremes; see Figures 4.6a and 4.6b. It underestimates extremely large values for old houses and

overestimates values for leaky new houses. The error analysis also suggests that new coefficients are required, especially for the new stock, because additional explanatory variables could reduce errors in the models' predictions. For instance, Chan's model does not include the construction material as an explanatory variable, and it is not expected to have a significant effect on the permeability of USA dwellings. However, Section 3.1.6 shows that Blower Door pressurisation tests carried out in Chile show that structural materials of the envelope of the sampled houses have a significant effect on the total airtightness. Furthermore, finishing materials such as the use of air barriers may also influence the envelope air permeability, parameter that, to date, has not been explored in detail and data quality is poor. Finally, houses built after 2016 must be included as the stock evolves and the model is updated.

To fill this knowledge gap, more data gathering is required. Pressurisation tests using the Blower Door technique have been extensively used in other countries, such as France, USA, and UK. To do so, regulations, guidelines, and protocols are available; some of them are already in use in Chile. The establishment of an open–source national platform is also necessary to synthesise this data, the test results, and additional physical parameters of the housing stock.

7.3.3 Internal air temperatures

There are no large–scale measurements of indoor air temperature in Chilean houses reported in the literature. Indoor temperatures are assumed to follow the same distribution of temperatures found in English houses (Shipworth *et al.* , 2010). Although the study of Shipworth *et al.* (2010) considered a big sample size (N=195), the majority of the measured houses used gas or oil–fired central heating systems, which are only found in a small proportion of the Chilean stock. Houses using other heating systems, such as heater stoves, are likely to have different thermal comfort temperatures. Another limitation of the Shipworth study is that it included data for both summer and winter seasons, whereas this study only considers winter. Similarities between the English and Chilean weather and occupant behaviour has not been a subject of study, and so the difference between indoor air temperatures in the two countries cannot be considered with any accuracy.

Another important consideration is temporal temperature fluctuations during the simulation period. Here, due to the lack of reliable information, indoor temperatures are modelled as constant for the entire simulation time, which rarely occurs under reality. Since naturally ventilated houses tend to have greater temperature changes than those that are mechanically ventilated, indoor–outdoor temperature differences are also expected to be lower and so reduce the effect of buoyancy on the ventilation

rates. The magnitude of this uncertainty on the results in unknown, but it is likely to overestimate ventilation rates and heat losses, and to underestimate pollutant concentrations.

A data analysis of a sample of 297 Chilean houses located across the country was recently completed (Molina *et al.* , 2019). It showed that indoor temperatures are generally cold, following a normal distribution with $\mu = 19.4°C$ and $\sigma = 4.1°C$ when considered both seasons, and $\mu = 17.1°C$ and $\sigma = 3.1°C$ for winter days (Molina *et al.* , 2019). These temperatures are cold when compared against international standards and guidelines, including the European adaptive model of thermal comfort, EN 15251 (BSI, 2007), and the WHO's recommended boundaries of 18°C and 21°C for clothed sedentary occupants to avoid potential health risks (WHO *et al.* , 2007).

The sensitivity analysis presented in Chapter 6 (page 112) shows that temperature difference is the second most sensitive parameter to the winter exposures and highly significant in all statistical tests, which justifies the need for a better and more accurate representation, see Section 7.6 for a detailed discussion.

7.3.4 Floor areas

Although the variability of the houses floor area used to develop the archetypes are known (see Table A.2 on page 191), for simplicity and to reduce the number of simulations, this variable is fixed, and discrete values are assigned to the archetypes. Nevertheless, Equation 2.4.1 (page 20) shows that room volume is an important parameter used by indoor air quality models and so merits inclusion in future surveys.

7.3.5 Window model

Windows are modelled using the orifice equation; see Section 5.2.1 (page 95). Ongoing work at the University of Nottingham[3] suggests that the values used here for C_d and A_f may overestimate the actual effective area of the openings, which could lead to an inaccurate estimations of airflow rates. To reduce this epistemic uncertainty, the use of the *effective area* as the product of the first and second term of the Equation 5.2.3 (page 96) should be used instead, following Jones *et al.* (2016). Windows were modelled using one–way orifice equation. CONTAM has the capability to do two–way window flow, which could provide a better estimate of the airflow through these elements and tend to significantly reduce the flow.

[3]Jones B. Iddon C. 2016. Window Effective Area Calculator DOI: 10.13140/RG.2.2.10748.08323

7.3.6 Pollutant species

The nine species included in the CONTAM models are selected according to their health effects on occupants —ranks given by Logue *et al.* (2011a)— and due to their prevalence in Chilean houses; see Section 4.7 (page 85). Although there is no large–scale study about hazardous pollutants in the Chilean stock similar to that of Logue *et al.* (2011a), it is assumed that all the pollutants included here are of concern and could negatively affect public health.

Estimating exposures to all the pollutants included in the model was beyond the scope of this study. Due to time constraints, this thesis could not provide a comprehensive prediction of the overall *acceptability* of indoor air quality in the Chilean stock and so using the first criteria pollutant, fine particles. Nevertheless, the emission profiles of the other species may differ, and so they would need to be analysed individually.

7.3.7 Pollutant emission and deposition rates

Pollutant types and emission rates (ERs) associated with heating and cooking depend on the type of fuel, the condition and design of the heater or cooking stove, food types and cooking methods, and the frequency and durations of heating and cooking events (O'Leary *et al.* , 2019a). These factors are not all recorded by the national censuses or the ENUT, and so it would be advantageous if they were considered in future surveys. The heaters modelled here account only for a subset of those available in the Chilean market. The measurement study of CENMA (2011) applied here chose the heaters most commonly bought and ignored the rest, such as the old and less efficient. These are expected to have different ERs, and so they should be measured. Moreover, wood–burning stoves are known to emit particles when their doors are opened to add fuel. To develop better guidelines on heaters, more devices must be tested. Furthermore, the method used for the measurements only reports the average ERs, and there is no error analysis with which to estimate uncertainty in them.

There is also uncertainty in ERs from cooking. One of the studies used to produce the distribution of ERs from the cooking of meals (O'Leary, 2018) was included before data was published, and ERs have changed since then; see O'Leary *et al.* (2019a) for further information. The way the data were re–processed lead to different values that increased by up to 100%, and so these results will underestimate the concentrations of cooking $PM_{2.5}$ for some of the meals. Moreover, there is no convention on protocols or standards for designing an experiment to measure $PM_{2.5}$ emission rates, for defining appropriate equipment and for using data to estimate ERs from this type of source.

ERs associated with construction and finishing materials in Chilean houses are not

be available from public sources, and so a national database would be beneficial. It would be of particular interest to know how materials age and how their ERs change over time. These effects are not included in well–known international databases, such has as CONTAM–Link, the EPA, or NRC ER databases. Moreover, ERs from building materials are measured under controlled environment conditions that could differ from the actual conditions. Studies have shown that ERs under indoor conditions vary according to the temperature, humidity of the air, the water content of the substrate, and the air velocity in the room (Haghighat & De Bellis, 1998). Thus emission rates from building materials are not analysed in this thesis, and more experimental studies are required to provide a deeper understanding of these complex phenomenons so that they can be modelled in old and new houses.

Deposition rates used here, $N(0.39, 0.16)$, may underestimate those found in Chilean dwellings. The PTEAM study by Oezkaynak *et al.* (1996) recorded mean exchange rates of around $1\,\mathrm{h}^{-1}$, which is similar to the median ACH of windows closed scenario but well below those for windows open scenario. Field and chamber studies use smaller samples but have shown the effects of the ventilation rate on particles deposition. The results are currently inconsistent and so a better model that accounts for more factors, such as the origin of particles, the surface–to–volume ratio, surface roughness, and air turbulence is needed. This phenomenon is not currently accounted for in the CONTAM model, but it is possible to incorporate it manually by varying the deposition rate parameters during the simulations according to indoor environmental parameters. Consequently, this variable may be found to be more important than it currently is in the future.

7.4 The modelling tool

CONTAM model treats each room as a single node with well–mixed conditions, with a uniform air temperature, air pressure, and pollutant concentration. In kitchens, this assumption may not reflect the actual pollutant dispersion around the room. Emission rates are simulated using a constant emission rate over time while accounting for the total mass (known as the *source strength*) added to a zone over a period of time. However, exposure to pollutants in the kitchen depend on several other factors, such as the proximity of the cook to the source, the location of the range hood and its capture efficiency, and the time the cook spends at each position (O'Leary *et al.* , 2019a). Moreover, pollutant concentrations in the kitchen are not uniform, but rather have three–dimensional gradients, which may be affected by the use of the range hood over a cooker.

Kitchen range hoods with a capture efficiency of less than 1 were not modelled

here. Only standard exhaust fans were considered. This omission is likely to have overestimated exposures. Further data collection is required to determine with more accuracy how range hoods affect the (average or gradient) concentrations and about their use in Chile. This is a limitation that was not addressed in this study and so may need to be considered in future studies. A possible alternative could be the application of CFD methods to determine the spatial variation in concentrations . However, this type of analysis would be confined to a 2D analysis single room when using CONTAM, although repeated tests would indicate uncertainty in the predictions and their sensitivity to inputs. Finally, the capture efficiency of the range hood can be accounted for using multiplying the emission rates by 1 minus its value; see O'Leary *et al.* (2019b). Although this would enable the adjustment of the average concentration of particles in a room, it would not necessarily simulate the actual exposures of the cook, due to their position relative to the source and the range hood.

The models do not account for all the dynamic processes that could affect the indoor concentrations. Chemical reactions such as ozone interactions, adsorption and absorption, re–suspension, gas–to–particle conversion (such as condensation or agglomeration) can be incorporated in the models when data becomes available and the need arises.

Presently, background concentrations are assumed to be zero. However, ambient particles can be an important contributor to indoor concentrations, especially in polluted areas; see Section 2 (page 11). They can enter the house through adventitious and purpose provided openings located in the envelope of a dwelling. These particles from outdoor sources are excluded from this study due to the significant uncertainty in their concentrations, and a lack of knowledge in the composition and effects that a mixture of particles from different sources may pose to human health; see Section 2.4 (page 19) for further arguments. Moreover, data on pollutant concentrations from centrally located monitoring stations have poor agreement with those measured at *in–situ*; see Section 2.4.2 (page 25). Future work should attempt to quantify the harm of exposure to indoor sources of $PM_{2.5}$ at a population scale. Here, it is appropriate to exclude outdoor concentrations. However, there may be situations when it is necessary to include them, if new knowledge becomes available or a hypothetical concentration needs to be assessed. Then, it is possible to simply add outdoor concentrations using Equation 2.4.1 given on page 20. The ambient concentrations can be used in the exposure analysis by adding them directly to the indoor concentrations. When a penetration factor is used account for filtering as $PM_{2.5}$ pass through openings, the ambient concentration must multiplied by both the air change rate and the penetration factor, and then divided by the sum of the air change and deposition rate before being added to the indoor concentration. The levels of exposures presented

in Figure 6.21 (page 135) will then be modified by an increase to the exposure line in parallel with the y–axis. Finally, a separate analysis of each source is also possible, to differentiate between their contributions to the overall concentration, by adding an outdoor species of $PM_{2.5}$ to each model.

7.5 Predicted outcomes

The model outputs are analysed to identify those that are the most common and those that are extreme and therefore unlikely to occur. This data processing enables policy decisions to be made that affect an acceptable proportion of the stock so that an intervention can be made that works and is also financially and logistically achievable.

7.5.1 Daily average $PM_{2.5}$ concentrations

To generate a single value of concentration for each house, the concentrations in the kitchen, bedroom, and living room are weighted by the time spent in each zone using the Chilean data, a 10 : 38 : 52 ratio for the *kitchen : bedroom : family room*, instead of using the ratio suggested by Hamilton *et al.* (2015); see Section 4.6 (page 84). This estimate gives a room weighted indoor $PM_{2.5}$ concentration averaged over the winter season, $\overline{PM}_{2.5}$. In this study, the predictions of the total mass of pollutants at which the occupants are exposed do not account for the mass that is inhaled, which would require knowledge of the age and metabolic rate of each occupant, among other variables. Therefore, these values of $\overline{PM}_{2.5}$indicates a potential level of risk to which occupants may be exposed.

7.5.2 Predicted exposures to $PM_{2.5}$

Temporal data can be used to show when peak concentrations occur, and the magnitude and duration of events. Concentrations that occurred during occupied periods can be used to develop an exposure profile. The same data can also be used to derive information about occupants behaviour, habits, and patterns. Aggregated concentrations can be compared against exposure recommendations, and cross–compare against measurement made *in–situ* (e.g. for a meta–analysis), and used to predict health effects when an exposure–response relationship is known.

To obtain a better estimation of the exposures, concentrations were selected according to the occupancy pattern (by time and location) rather than using room averages. Figures 6.4 and 6.5 (page 116) show $PM_{2.5}$ concentration profiles for the three rooms used in the exposure analysis. Concentrations are shown to be generally high, espe-

cially during cooking periods, and not only in the kitchen but also in the living room and bedroom. Doors are assumed to be open when cooking activities are taking place, see Section 5.2.5 (page 101), which may contribute to the spread of pollutants. This is an important issue for future research. The impact of door opening and the variation of this behaviour on indoor air quality can be used for informing the occupants about these features.

Table 6.2a (page 117) and Figure 6.8 (page 119) show the statistics and distributions of the medians of hourly exposures to $PM_{2.5}$; namely, the upper limit concentrations at which the occupants are exposed for 50% of the hours. Archetypes IDs 275 and 284 have the highest values. These correspond to dwellings built after year 2008, which required a tighter envelope than those built before 2008. This indicates that interventions are necessary in newer houses to lower concentrations.

Figure 6.6 (page 118) shows the distribution of the hourly winter concentrations in the simulated houses when all the windows are closed. The CDF shows the percentage of the hours indoor concentrations are likely to be below or equal to particular values. Figure 6.13 (page 123) shows the effect of window opening on the hourly exposure concentrations, and Figures 6.15a and 6.15b (page 126) show their effect on the winter exposures. These curves can be used to communicate the magnitude of the impact of one type of intervention, in this case increasing the ventilation rates by opening windows. Although it is evident that the actual scenario will be somewhere between these two limits, it is more likely that occupants will keep the windows closed most of the time to conserve energy and maintain thermal comfort. However, a combination of the two scenarios can be complicated, but to model their effect, occupant behaviour data is needed.

Results show that both cases have different consequences. Figure 6.21 (page 135) shows that a low ventilation rate can lead to high $PM_{2.5}$ concentrations, which may impact occupants' health. Conversely, an increased ventilation rate leads to higher energy losses and the entry of ambient pollutants at a higher rate. These results reflect the understanding presented in Chapter 1 and Figure 1.2 (page 8), contextualising theory to the Chilean situation. The shape of the line of the ventilation rates versus the heat loss do not follow a linear relationship, as shown by Liddament & Air Infiltration, Agence internationale de l'énergie (1996), but appear exponential. This is explained by the use of the logarithmic scale due to significant range of predicted values.

Two important tools are used to assess exposure and when monitoring data: (i) reference values and (ii) exposure limits. They are useful for evaluating the degree of exposures of a population to $PM_{2.5}$, and to evaluate changes, such as interventions or remedial actions (Ewers *et al.* , 1999). A reference value for indoor air is the concentration at which the general population is exposed in a given location and time (EPA,

2018). They are normally calculated to characterise the upper margin of the current scenario using either the 90th or 95th percentiles of a representative sample. The predictions presented here can be used to set reference values for the general population; for example, those in Table 6.3 (page 125).

The winter season is used to represent a worst case scenario when occupants tend to keep their windows closed to maintain thermal comfort and save energy. However, information about some inputs is limited and so assumptions are required, see Chapter 3 and Sections 7.2–7.3 (pages 142–144). Given the uncertainty in occupants' behaviours, two extreme scenarios were proposed for evaluating the indoor exposures and used to characterise them. At one extreme is a scenario where dwellings keep their windows closed for the entire season, and at the other, occupants keep the windows fully open.

Values presented in Tables 6.6 and 6.7 (pages 128 to 129) can also be used as reference values if they are needed for subgroups of the stock; for example, to identify dwellings of a certain archetype or region with unexpectedly high levels of pollution. These values can also be revised periodically in order to evaluate changes and define new values.

Field studies presented in Section 2.4.2 (page 25) showed that $PM_{2.5}$ concentrations in Chilean dwellings are roughly 5–374 $\mu g/m^3$. Comparable results are shown for the 90th percentile of 367 $\mu g/m^3$ (only accounting for indoor sources of $PM_{2.5}$, outdoor $PM_{2.5}$ would increase this). Although these field studies are useful for comparative purposes, they represent minority subgroups of the Chilean stock or a vulnerable population. Ideally, housing data should be informed by large–scale fieldwork. Representative buildings, like those shown here, can guide future data gathering exercises.

Nationwide, the mean (μ) exposures for the two scenarios are predicted to be between 6.6–134.5 $\mu g/m^3$, and the meadian (M_{ed}) exposures are 2.3–58.65 $\mu g/m^3$, ranges that fall within representative values estimated by Logue *et al.* (2011b) for USA houses ($\mu = 15.9\,\mu g/m^3$; $M_{ed} = 15.7\,\mu g/m^3$). The 95th percentile of 86 $\mu g/m^3$ in USA houses corresponds to the 60th percentile of the Chilean stock, and so the health impacts of indoor exposures for the Chilean population are expected to be higher than those predicted by Logue *et al.* (2011b). The 95th percentile for the Chilean stock is estimated to be 549 $\mu g/m^3$.

Examining the central tendency of the results, the M_{ed} winter exposure for windows open scenario is 26 times lower than the M_{ed} for windows closed scenario and less variable, approaching background levels.

7.5.3 International guidelines and health impacts

The previous section presented the advantages of using a reference value for the stock and proposed a reference value for $PM_{2.5}$ exposures in Chilean houses. These values are especially useful when health effects are related to a mass concentration, taken as the integrated exposure over a certain time. Normally the average of these values is also used for an analysis of lifetime exposure. However, using a reference value has two main disadvantages: (i) it does not follow any criteria related to health, and so it is only useful for comparison, for meta–analyses, or when a healthy limit is known, and (ii) it is calculated from values that reflect concentrations at a specific time and location. Conversely, *exposure limits* are usually selected according to the health effects of exposures by the general population; for example, when a dose–response or exposure–response relationship is known. The CDFs presented in Chapter 6 (page 112) predict the uncertainty in reaching certain values indoors. Similarly, statistics can be used for benchmarking, monitoring, and for comparisons against international thresholds and exposure limits. In Chile and the rest of the world, indoor air quality is not regulated and guidelines are not enforced. Particles are only legislated for the ambient air; see Section 2.2.2 (page 15). Because there is limited information about the composition $PM_{2.5}$ from indoor sources, risk assessments can only be carried out using international recommendations at the present time. They must assume that particles of indoor origin are equally as harmful as those from outdoor sources, or they must identify the individual components of the indoor air and their toxicities.

Using the guidelines presented in Chapter 2 (page 11), the proportion of the stock with winter exposures above the international benchmarks are:

- EPA, NAAQS guidelines: 24h mean above $35\,\mu g/m^3$: **63%**
- Chilean standard DS12/2012: hourly mean above $50\,\mu g/m^3$: **55%** of the winter hours across the stock.
- EPACal, AAQS: 24h mean above $35\,\mu g/m^3$: **63%** of the winter hours across the stock.
- WHO: 24h mean over above $25\,\mu g/m^3$: **70%** of the stock.

The top graph in Figure 7.1 shows a CDF of predicted stock–weighted concentrations in the Chilean houses (windows closed scenario) in red, and population–weighted exposures in other countries in black. The lines of the bottom graph show

the association of integrated exposure to ambient $PM_{2.5}$ with outcomes including tracheal, broncheal, and lung cancers (LC in light blue), Chronic Obstructive Pulmonary Disease (COPD in green), ischemic heart disease (IHD in red), cardiovascular disease (CVD in black), and lower respiratory infections (LRI in blue). These include both mortality and morbidity risk estimates. As a limitation, the risks plotted in this figure were those assigned to a population between 25 and 29 years old, of both sex, and related to ambient pollution (Gakidou *et al.* , 2017).

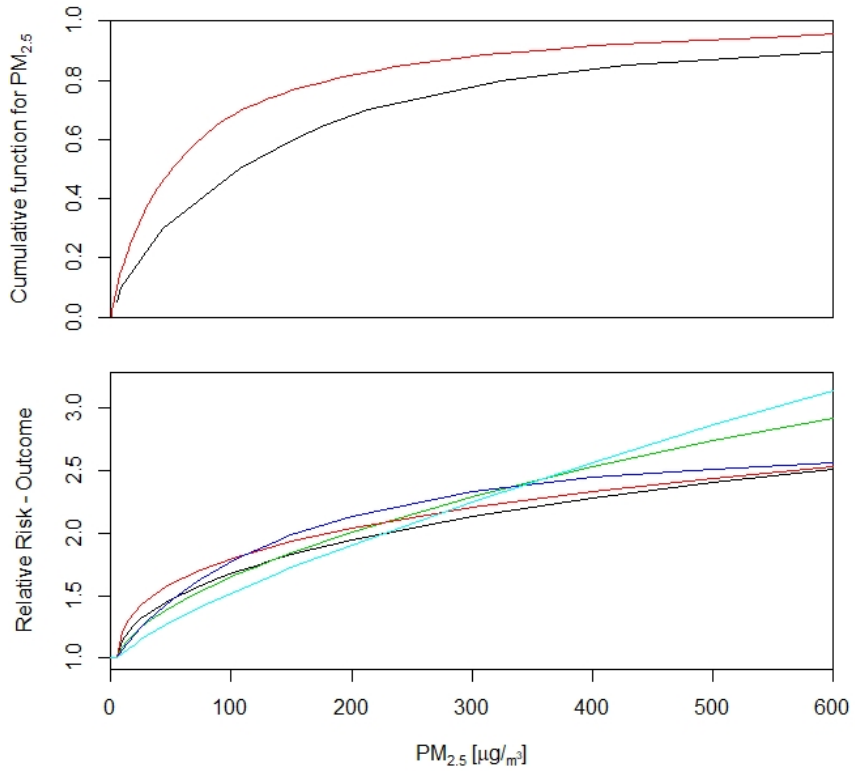


Figure 7.1: Top: Stock–weighted exposures to $PM_{2.5}$ in Chile (windows closed scenario), in red, and population–weighted exposures in other countries, in black. **Bottom:** relative risks associated to ambient $PM_{2.5}$. Tracheal, bronchus, and lung cancers (LC in light blue), Chronic Obstructive Pulmonary Disease (COPD in green), ischemic heart disease (IHD in red), cardiovascular disease (CVD in black), and lower respiratory infections (LRI in blue) (Gakidou *et al.* , 2017).

The relationships in Figure 7.1 show that the relative risk (RR) of developing IHD and LRI doubles when $PM_{2.5}$ concentrations are above 150 $\mu g/m^3$, and above 200 $\mu g/m^3$ for LC, CVD, and COPD. The proportion of the Chilean stock under these conditions, with time–weighted average concentrations above 150 and 200 $\mu g/m^3$, are 26% and 20%, respectively; see Figure 7.1 top. These percentages do not consider the time the occupants are outside of the house, and only the activities modelled here contribute

to these exposures.

7.5.4 Ventilation rates

Median (M_{ed}) ventilation rates for the Chilean stock during the wintertime are shown to be high. Table 6.8 presents the distribution of the median ventilation rates aggregated by archetype ID. Summary statistics in Table 6.2b indicate that new houses, represented by IDs 275 and 284, have ventilation rates between 7 to 11 times lower than those built before 2008.

Hourly data distributions and the median over the winter can be used to inform about the variability and uncertainty of this parameter. The predicted median is between 0.75 and 13.57 h^{-1} using windows closed and open scenarios, respectively. Windows closed scenario accounts only for air leakage through the envelope and extractor fans in bathrooms and kitchens. This scenario is predicted to have a μ of 0.89 h^{-1}, M_{ed} of 0.75 h^{-1} and σ of 0.79 h^{-1}.

Applying the international ventilation standard ASHRAE 62.2, equation 4.1b (ASHRAE 62.2, 2016), the minimum ventilation rate required for a Chilean house of average size and two bedrooms is 20 $l\,s^{-1}$, equivalent to 0.46 h^{-1}, which is comparable to the rule of thumb of 0.5 h^{-1}. Accordingly, the results show that ventilation rates are below 0.5 h^{-1} the 39% of the time, with a μ and σ of 0.95 h^{-1} and 0.90 h^{-1}, respectively.

Taking the median ventilation rate over the winter, 62.8% of the stock is predicted to meet the requirement. In contrast, the minimum (median) ventilation rate in the second scenario is 4.4 h^{-1}; see Figures 6.15c and 6.15d. If the actual scenario is expected to be between those two, Chilean houses are likely to be over–ventilated from the energy point of view and not considering the indoor air quality. The predicted median ventilation rate for an average house would provide 33 $l\,s^{-1}$ for windows closed scenario and 593 $l\,s^{-1}$ for windows open scenario.

Figure 6.21 shows that to ensure that 95% of the stock is below the WHO's recommendation of 25 $\mu g/m^3$, the ventilation rates must be up to 13 h^{-1}, which can be cost–prohibitive. This corresponds roughly to the median of windows open scenario and out of the range of windows closed scenario.

7.5.5 Ventilation heat losses

CDT (2010) estimated the total energy consumption of the housing stock to be 53.8 TWh per year, measuring a sample of 3,220 houses. This study estimates that the total heat loss for the entire country during winter, only by exfiltration and the use of extractor fans in kitchen and bathrooms is 8.2 TWh (windows closed scenario), and opening

the windows 124 TWh (windows open scenario), which accounts for 15%–230% of the estimated total. These values seem plausible, and given that both scenarios consider extreme conditions, the actual scenario is expected to be in that range. Comparing against the same study of CDT (2010), the average energy use for heating a single house in Chile is approximately 5,761 kWh/year (no uncertainties are given). This study, using the mean of the sampling means, estimates the stock averaged heat loss due to ventilation by 1,284 kWh/season. If it is considered that the house is heated only in wintertime, the energy demand required to condition indoor air predicted by this study accounts for 22.35% of the total energy use of a house. Comparing these figures against Sartori & Hestnes (2007) energy use reference for low–energy buildings located usually in cold climates, a Chilean average–size house should demand 4,585 kWh year for heating (20% less energy than the estimation made by CDT (2010) for the current stock) and should have an overall demand of 16,400 kWh year (-43%). This study predicts that an average house would consume at least 28% of that only for indoor air heating during winter, not considering the heat loss due to conduction. And if the windows are fully opened, a mean use could reach up to 19,510 kWh, 4.3 times the benchmark, only to condition indoor air.

7.6 Sensitivity analysis

The sensitivity analysis determine the relationships between each of the inputs and the outputs.

Scatter plot of inputs versus outputs illustrate the relationships between the individual inputs and the output for visual inspection. They are shown to be highly variable, especially when including both scenarios, and so the type of correlation becomes more difficult to interpret; see figures C.1 to C.6 in Appendix C for each separate scenario, and figures 6.22 to 6.24 in Chapter 6 for both scenarios plotted together. Nevertheless, the results are yet useful for identifying the inputs that are more important, more related, and contribute the most to the outputs.

p–values can be interpreted for significance, and so testing the relationship between the two sets. Given the large sample size, this interpretation can be meaningless, since the chances of finding significance increases with the sample size (Gigerenzer, 2004). Furthermore, the statistical significance of *p*–values is arbitrary, and so in this study the focus is on the nature and the magnitude of the effects (Fenton & Neil, 2012). In consequence, for reporting *p*–values the exact level of significance are given rather than their interpretation. For brevity, Chapter 6 only reported the resulted ranks; *p*–values and coefficients are found in Appendix D.

Linear, monotonic, and non–monotonic tests were applied to the inputs and the

three outcomes using correlation, regression, and sample comparison methods.

7.6.1 Correlations

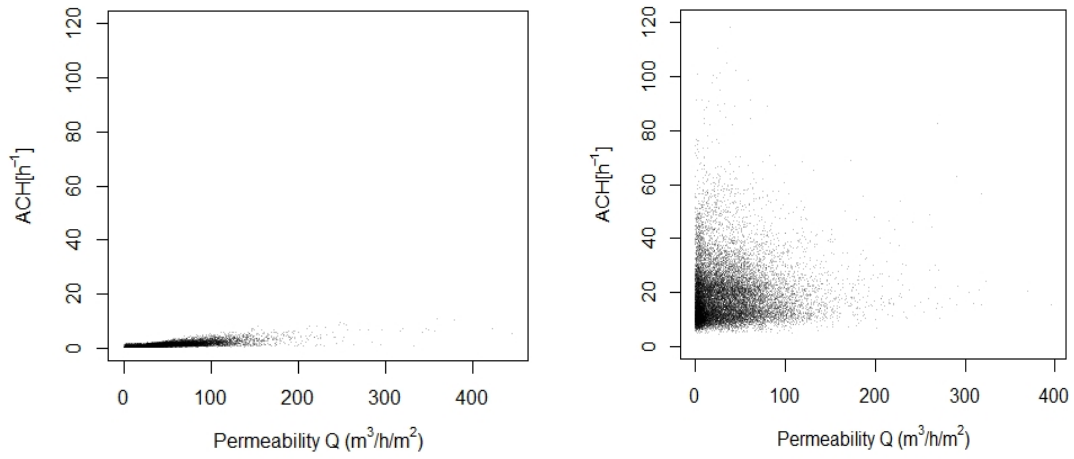
The three tests chosen for determining the strength of linear and monotonic association are: (i) Kendall's τ rank, (ii) Pearson's r, and (iii) Spearman rank. These three statistical tests measure the strength of the linear correlation between the two variables and its direction, giving a correlation coefficient between -1 and 1; weakening the strengths as approaching to zero. The correlation coefficient tells about nature and size effect of the correlation, or how strongly related are the two datasets.

Pearson's r is recommended for parametric data of both inputs and outputs, and continuous inputs. Nevertheless, not all the inputs and any of the outputs meet this requirement, and so their meaning is limited. Kendall τ rank and Spearman rank analyses do not require the data to be normally distributed. Moreover, Spearman can also be applied to ordinal data and because it is based on ranking the variables, is less sensitive to outliers. Thus, given the characteristics of the input and output data, the statistical test that is particularly meaningful for the correlation analyses regarding linear and monotonic relationship is the Spearman rank correlation test.

Results of Spearman's rank correlation coefficients show that exposures are most strongly correlated to cooking emission rates, followed by envelope permeable area and heater emission rates. The only input that indicates the wrong tendency to decrease the exposures is emissions from toasts, although it shows very weak correlation and p–values are non–significant at the 5% level.

For ventilation rates, the total permeable area of the envelope was the first–ranked input, followed by the Q_{50} and ΔT. Figures 7.2a and 7.2b show the positive relationship between the air permeability of the envelope and the ventilation rates for windows closed and open scenarios, respectively. Figure 7.2b shows more clearly the relationship between the two variables and that both have shifted distributions, meaning that for a single permeability value (taking any vertical slice), the possible ventilation rates have a non–normal, right–skewed and bimodal distributions. Similarly, for a specific ventilation rate, a broad range of Q_{50} values are possible. Nevertheless, the objective of this sensitivity analysis was to identify the relative importance of the inputs within the range of possible values by analysing both sets of results together. Thus the interpretation is made here as a whole.

Finally, heat loss and ventilation rates share similar sensitive inputs. The total permeable area is the input most strongly correlated and so first–ranked, followed by ΔT and Q_{50}.



(a) Windows closed scenario.

(b) Windows open scenario.

Figure 7.2: Scatter plots of the envelope air permeability for (a) windows closed scenario and (b) windows open scenario, and the predicted median ventilation rates over the winter.

7.6.2 Regressions

For testing the contribution of each independent variable to the model outcomes, two regression methods were applied: (i) the standardised regression coefficients (SRC) that tests the linear relationship between the two sets using the raw values, and (ii) the ranked regression coefficients (RRC) that tests the monotonic relationship using the variables ranked.

The magnitude of the regression coefficients, R, are used to rank the inputs and to show which have the greatest effect on the output. Furthermore, coefficients of determination, or R^2, are computed from them to show the input's contribution to the variability of the outputs (Taylor, 1990). Low R^2 indicates poor regression model, namely the input's variance is poorly reproduced by the (linear) regression model.

The highest regression coefficients in the total exposure is seen for emission rates from cooking, a modest 0.58, followed by a low -0.21 for the total permeable area. For the other two outputs the regression coefficients are generally low, with a maximum contribution of 0.31 and 0.20 for the total permeable area to the ventilation rates and heat loss, respectively.

Finally, all the coefficient of determinations are low (below 0.36) and many of them unimportant (below 0.1) for a linear relationship. The highest coefficient of determination, R^2, is seen for emission rates from cooking to the exposures, with a value of 0.34, meaning that a 34% of the variability of the exposures to $PM_{2.5}$ could be accounted for by the emission rates from cooking.

7.6.3 Sample comparison tests

Two non–parametric sample comparison tests were used, (i) the two–sample Kolmogorov–Smirnov (KS) and (ii) the many–sample Kruskal–Wallis (KW) with 2, 5, 10 and 20 sub–samples. Special attention is given to these studies due to their ability to detect non–monotonic relationships. The KS test calculates the maximum distance between the two empirical CDFs by partitioning the input's distributions, in this case, based on the median. It is a variance–based sensitivity index between 0 and 1, used to determine the influential inputs, or the parameter sensitivity rankings; the lower the value of the KS index the less influential or sensible is the input to the output. The sensitivity index can be used to cross–compare the sensitivity analyses results across different studies.

This test identifies that the most important variable for exposures is again the emission rates from cooking, followed by the total permeable area, and the S:V ratio; for ventilation: the total permeable area, the envelope air permeability, and S:V ratio; and for heat loss: the total permeable area, the S:V ratio, and the envelope air permeability.

In contrast, the KW test compares the variance of the data using the medians, similar to the ANOVA test, which is used as the parametric method. Here, not all the distribution parameters are known for the inputs, and the output data does not meet the distributional assumptions. Thus, the non–parametric method is applied. The outputs distribution is divided into 2, 5, 10 and 20 sub–samples using their quantiles, and their corresponding inputs are tested (Jones *et al.* , 2015). The test statistics or H–value is the horizontal distance between the sub–sample's inputs medians, and so the values are greater than 0; the higher the H value, the more dissimilar they are and so the more influential to the output. One disadvantage of this test is that it does not consider the entire probability distribution of the outputs, and so it cannot fully characterise their uncertainties. Although they inform that at least one section of the distribution of the outputs is influenced due to a difference in the inputs, it does not tell where the difference is, and so a *post–hoc* test is required for further analysis.

For exposure to $PM_{2.5}$, all the KW tests show equal ranks for the first two inputs, cooking emission rates and total permeable area, followed by the emission rates of heaters and wind speed at the house height in the 3rd or 4th position. For ventilation rates, all KW tests show the same two highest ranks: total permeable area and Q_{50}. After these two, wind speed, ΔT, and S:V ratio follow in importance at different number of sub–samples. Finally, heat loss is more variable due to changes in total permeable area, Q_{50}, and ΔT.

Table 7.1 presents the top five–ranked inputs accounting for monotonic and linear relationship tested, and classifies them according to their relationship with the house

design, the environment, or the sources.

Table 7.1: Inputs ranked according to their sensitivity to the outputs and classified according to the type of input.

Input			Output
House Design	**Environment**	**Sources**	
2. Total permeable area 5. Permeability Q_{50}	4. W_s	1. Cooking ER 3. Heater ER	$PM_{2.5}$ exposures
1. Total permeable area 2. Permeability Q_{50} 4.-5. S:V ratio	3. $T_{int} - T_{ext}$ 4.-5. W_s		ACH
1. Total permeable area 3. Permeability Q_{50} 5. n exponent 5. S:V ratio	2. $T_{int} - T_{ext}$ 4. W_s		H_I

On the lowest ranks, block aspect ratio and orientation of the house show little or no importance on any of the three outcomes.

7.7 Comparisons with other studies

In this section the predictions are compared against three studies on housing stocks that used a similar method, a probabilistic sampling. Winter hourly ventilation rates for windows closed scenario are predicted to be below $0.5\,h^{-1}$ in 39% of the time, see Figure 6.6, lower than that of Persily *et al.* (2010) for US houses, which is estimated to be 57% of the time below this benchmark, see Figure 2.18. The reasons for this, apart from the disparities in infiltration properties between the US and Chilean houses, see Section 3.1.6, is that ventilation rates in this study are increased by the use of the extractor fans in kitchens and bathrooms during cooking and showering times.

The median ventilation rates over the winter are predicted to be between $0.08\,h^{-1}$ and $2.40\,h^{-1}$ (90%CI). Considering the use of fans, the upper value is roughly twice the USA predicted infiltration rates (90%CI: $0.10 - 1.21\,h^{-1}$) (Persily *et al.* , 2010), England's (96%CI: $0.02 - 1.24\,h^{-1}$) (Jones *et al.* , 2015), and also higher than Beijing's ($0.01 - 1.60\,h^{-1}$) (Shi *et al.* , 2015).

Jones *et al.* (2015) estimated that infiltration in the UK housing stock (similar to windows closed scenario) is responsible for 11–15% of the energy demand during heating hours. Compared against the UK total heat loss (Jones *et al.* , 2015), the Chilean heat loss is between 0.19 and 2.74 times the UK's, although some of the criteria differ between both studies.

Table 6.3 (page 125) shows that the predicted $PM_{2.5}$ exposure μ (σ) for the Chilean stock ranges between 6.64 and 134.47 $\mu g/m^3$ (12.26 – 210.28 $\mu g/m^3$) and the M_{ed} between 2.30 and 58.65 $\mu g/m^3$, considering windows open and closed, respectively. These ranges include the values predicted by Das *et al.* (2014) for the English stock. Das *et al.* considered kitchen windows opened between 0.01–10 times the duration of cooking, giving a μ of 12.7 $\mu g/m^3$, σ of 12.6 $\mu g/m^3$, and M_{ed} of 8.0 $\mu g/m^3$.

7.7.1 Other sensitivity analyses

The three most important variables for $PM_{2.5}$ exposures are the emissions from cooking, the total permeable area, and the emission rates of heaters. These are broadly consistent with that of Das *et al.* (2014): the opening area of the kitchen window, and the emission rate of internal $PM_{2.5}$. Jones *et al.* (2015) found that Q_{50}, n exponent, and S:V ratio were the three most sensitive inputs to the predicted infiltration in the UK houses. Here, infiltration is included in the total ventilation rate. For ventilation rates and heat loss, the three most important parameters are the total permeable area, indoor–outdoor temperature difference, and Q_{50}.

Finally, the negligible and non–significant sensitivity of the orientation of the house to the model outcomes is consistent with Shi *et al.* (2015), Das *et al.* (2014), and Jones *et al.* (2015).

Table 7.1 presents the first five inputs by considering both monotonic and linear relationships, classified by type of input. The two parameters related to the sources or the properties of the species that showed to be both influential and sensitive to exposures are the emission rates from cooking and from the heaters, which were first and second–ranked. This highlights the importance of source control strategies, which can have an immediate effect on the exposures to $PM_{2.5}$, and the need to provide information to the public on the risks factors associated with the house and on the behaviour to adopt in them. Concerning to house design, this rank can also inform about the design of future housing, public policies, and research. Total permeable area is the first–ranked input, and relates mainly to windows design, windows use, and the envelope's properties. Additionally, the envelope air permeability also plays an important role in the three outputs, along with the S:V ratio, and the n exponent. The S:V ratio shows a negative correlation, which shows that a building with a larger envelope surface relative to its volume will have better indoor air quality.

7.8 Applications of outcomes and future work

The characterisation of the Chilean housing stock is essential for an understanding of the diversity of the dwellings, which can be used to improve housing design, planning and, in turn, living conditions and public health.

Carrer *et al.* (2015) argue that epidemiological studies have shown that increased ventilation rates can avoid some indoor issues and health outcomes. Nevertheless, they have so far not succeeded in establishing a direct link between increased ventilation rates and health, mainly due to the lack of protocols on both data recording and exposure assessment. Field studies are subject to high uncertainty, given the diversity and variability of all of the factors involved. This makes it difficult to keep records of data, and so exposure–response and dose–response relationships are more difficult to specify or extrapolate to other populations. These limitations make modelling studies and intervention experiments more valuable. Modelling studies like this one, enable the testing of multiple scenarios quickly and at a lower cost than many other methods of research; see Chapter 2 Section 2.5 (page 36) for further information on those methods.

The data analysis of the national data sources highlights knowledge gaps in categorical descriptors and occupant behaviours, and poor granularity of physical data. These gaps should be filled by augmenting national surveys and complemented by fieldwork. The archetypes presented are used here as a tool of research, and can be applied to other modelling and field studies, such as the prediction and evaluation of the impacts of policies on indoor air quality and on energy demand, or to guide future data gathering exercises. This study has applied the `CONTAM` model to a set of archetypes to explore the indoor air quality in the stock. It has captured the variability of the entire stock using parameter distributions, deepening the understanding of the uncertainties involved in concentrations of fine particles, ventilation rates, and associated heat losses.

Uncertainty in the outcomes was explored using a probabilistic sampling method, and so a sensitivity analyses can be used to identify the key parameters that most affect the predictions of the model. Previous sections in this Chapter have used the results of these analyses to better judge the relative importance and the impact of epistemic uncertainty in some of the inputs.

Results of the simulations and sensitivity analysis can be used for direction and planning future research. In the following sections, areas for research are given in two–time horizons; short and medium term goals.

7.8.1 Short term

Field studies have identified subsets of the stock with poor indoor air quality; see Chapter 2 (page 11). This study took a step further by analysing the entire stock. It evaluated the uncertainty in this issue, examined the conditions that are likely to cause unacceptable outcomes, and identified the key parameters so the most important can be targeted objectively.

Suggested areas for immediate research based on these results are:

1. windows opening areas and schedules,
2. emission rates from cooking and heaters,
3. envelope air permeability, especially for houses built before 2007, and
4. indoor temperatures.

Information on window opening behaviour is most important, as it directly affects the three outputs studied here. Data could be gathered at low cost by installing sensors in windows to identify operation patters, and to compare them against measurements of environmental parameters, such as air temperatures. This is an area of research that should involve more than one discipline. For instance, sociology or anthropology may help to describe occupants more accurately, because human behaviour can be difficult to predict and categorise. These new insights would help to identify groups of the population that are more susceptible to living in dwellings with unacceptable indoor air quality.

A knowledge of the determinants of indoor activities can help to simulate the most probable scenarios. For example, natural ventilation through window openings, to amend the airflow rate or to improve air quality, have been associated with time of day and outdoor temperatures (Dubrul, 1988; Persily *et al.* , 2010). Moreover, it may also be an activity established in the daily routine of the occupants, that relates to their cultural behaviours, the local landscape around their house, and seasonal environmental conditions. Hence, the modelling of window opening using schedules and opening areas could help to account for these factors, which would give a more accurate spectrum of predictions.

Engineers can also help to provide a better description of a window's effective area; see Section 7.3.5. Schedules and window areas are clearly defined in the model code, meaning that any changes to them can be traced.

The prevalence of heating stoves varies widely across the country. Some studies suggest that this is a factor that can exacerbate socio–economic differences in indoor

air quality and thermal comfort (Ruiz *et al.* , 2010). The burden of disease study only includes the "use of solid fuels for cooking" as a risk factor in their calculations of health impacts due to exposures to household air pollution (HAP). This is mainly because nationally representative surveys or censuses do not include the presence of stoves for heating (Smith *et al.* , 2014). This excludes cooking stoves that use other fuels (88% of the Chilean stock), such as natural gas (10%) or LPG (77%), and exposures from other household uses, such as heating stoves. In Chile, two national surveys, the 2015 CASEN and the 2012 census, recorded heating fuel, but it was removed from the 2017 census (INE, 2018b). Thus, to facilitate a more comprehensive exposure assessment and evaluation of this risk factor, it is advisable to include this variable in the next census, which will be in 2027.

The implementation of a study to measure air leakage rates on a large scale is urgently required to obtain better descriptors of the relationship between leakage area and building characteristics, such as building materials and age of the building. This information would improve the model presented in Section 4.5.2 (page 78). It is particularly important to record the airflow exponent of the test, n, which was found to be of middle–ranked importance by the sensitivity analyses for all three outcomes.

Finally, air temperatures difference play an important role in the results, especially on ventilation rates (ranked 3rd), and on heat loss (ranked 2nd). This may be partially explained by the large variability in climates countrywide.

7.8.2 Medium term

In order to progress work in this field, it is necessary for academics, government ministries, and the private sector to work together. Academics must provide answers to research questions using multidisciplinary approaches and in an holistic manner, including perspectives from medicine, physics, chemistry, sociology, anthropology, architecture, and engineering. The public sector must provide the tools to enforce policies that benefit the population, by engaging with the Ministry of Health, the Ministry of Housing, the Ministry of Energy, and the Ministry of Environment. For instance, one important policy area to consider is the promotion of improved cooking and heating appliances, along with the use of cleaner fuels. It is necessary to involve industry because they are involved housing at every stage of a project, and they should be engaged to implement the precautions identified here when they design and construct dwellings to optimise performance.

Figure 6.21 (page 135) combines the three outcomes ($PM_{2.5}$ concentration, ventilation rates, and heat losses) to show their relationship with each other. It was produced using the 95th percentiles of each simulation. Furthermore, Section 6.5 (page 136)

presented relationships between the inputs and the three outputs. Several questions remain unanswered; for example, the post–analysis carried out here can be extended to each of the inputs to answer research questions such as: How large does a window or the purpose provided opening need to be so that the ventilation rate in the kitchen during/after the cooking time is enough to keep acceptable levels at a 95% level of confidence? Or, what is the nature and extent of the relationship between temperature difference, wind speed, and air permeability in Chilean houses?

Efforts should be made to implement source control strategies in housing in the first instance. This could be done by raising public awareness about IAQ and its impacts on health and well–being, by promoting and subsidising the use of cleaner fuels and more efficient stoves, and informing occupants about good cooking and housekeeping practices to reduce pollutant concentrations. The impact of source reduction strategies on population exposures to pollutants of concern can be estimated using this model. Secondly, house design and use should aim to optimise ventilation rates and heat loss in such a way that exposures to pollutants, in this case $PM_{2.5}$, are as low as possible. This study could then be expanded to include mechanical ventilation, particle filtration, and outdoor air quality.

This chapter discusses issues raised during this research and how they can be considered. Special attention was given to limitations of the work and how future studies are needed to understand the uncertainties in the parameters that are used to describe and model the stock in more detail.

This knowledge will have a positive impact by

- Characterising the Chilean housing stock and presenting a set of archetypal buildings to represent it (see Chapters 3 and 4);
- Presenting a model and modelling framework for evaluating the Chilean housing stock (see Chapter 5);
- Predicting uncertainties in occupant exposures to $PM_{2.5}$, and dwelling ventilation rates and energy losses across the Chilean housing stock (see Chapter 6);
- Identifying the most important parameters that affect the predictions (see Chapter 6);
- Contextualising and interpreting the results (see Chapter 7);
- Showing how the model its predictions can be used to inform and evaluate the impacts of new policies and improve the IAQ and environmental performance of dwellings (see Chapter 7), which will lead to improved performance;

- Identifying the need for future measuring, surveying, and data gathering exercises. (see Chapter 7).

CHAPTER 8

Conclusions

This thesis presents an exploratory and comparative study of the Chilean housing stock, and an estimation of uncertainty in predicted pollutant concentrations from indoor sources, ventilation rates, and heat loss in its dwellings. This new understanding creates a framework that can be used to evaluate and develop regulations for Chilean dwellings.

These conclusions are made by comparing outcomes, where applicable, against the objectives of this thesis defined in Section 1.7. The first objective was to define a set of buildings to represent the Chilean dwelling stock. This is achieved by developing a series of archetypal dwellings that are grouped together according to key parameters whose values are defined.

It is shown that 496 archetypes can be used to represent the entire stock and 90 can represent 95% of the stock. A selection of the first 2, 8, and 29 archetypes are statistically meaningful for academic and policy–making purposes. They represent 13%, 35%, and 70% of the stock, respectively, corresponding to *small*, *moderate*, and *strong* effect sizes.

The first 29 archetypes comprise 15 single–storey detached dwellings and 14 semi–detached dwellings (6 are two–storey) with floor areas of 30–139m^2. Archetypes are described by their construction period, dwelling type, number of rooms, bedrooms and bathrooms, mean, standard deviation, and median floor areas, number of storeys, and number of occupants. The predominant structural materials, cooking fuel, and socioeconomic status of the occupants are assigned to each archetype by geographical region. Distributions of the normalised leakage (NL) are developed for old and new dwellings. The NL for old dwellings is 95% CI $[9.91 - 106.59]$, and 95% CI $[1.39 - 15.90]$ for new dwellings. These distributions can be used to inform models of air infiltration and associated heating or cooling energy demand losses.

Important issues were raised during the data processing. Some parameters of interest are not included in public surveys, or they are not at a desirable resolution

and accuracy. The data analysis highlights a paucity of information about categorical descriptors, such as dwelling ages and types, year of construction, and heating fuels. There is also insufficient granularity in physical properties, such as dwelling floor area and volume (especially for dwellings built before 1990), window area and glazing type, air permeability (especially for dwellings built before 2007), insulation level, and local environmental conditions, such as sheltering and dwelling density. Very little is known about occupant behaviours, such as window opening, occupancy patterns, appliance use, cooking, indoor air temperatures and thermostat settings, or tobacco consumption. Future data gathering is required to fill these knowledge gaps and improve epistemic uncertainties.

From these factors, a good understanding of window areas and opening times, emission rates of internal sources, and indoor temperatures are the most important because they are required to improve the reliability of modelling predictions. Therefore, data gathering should focus on them first. It was also identified that a common platform for data gathering is required, especially for data on infiltration parameters, which could help improve analyses of its contribution to air exchange rates, dwelling energy demand, and indoor air quality. A better knowledge of infiltration factors would help to regulate dwelling energy demand and ventilation rates.

The absence of occupant activity data limits the ability of a model to quantify its effect on predicted outputs. In particular, it is not known when occupants open windows, and so two extreme scenarios were considered. *Windows closed scenario* accounted only for air leakage through the envelope and the periodic operation of extractor fans in bathrooms and kitchens, whereas *windows open scenario* also assumed that windows were fully opened all of the time. Time–use data used for assessing occupants' behaviours show that the time spent in the *kitchen : bedroom : family room* followed a $10 : 38 : 52$ ratio and a schedule was derived from it. The impact of these and other activities can be evaluated in more detail when more data becomes available. There is a definite need for activity start times and durations by region. This information can be used to develop customised information about Chilean occupants according to their location or demography. It could also be used to moderate occupant behaviour to improving the indoor air quality and preventing negative health effects.

The second objective was to estimate uncertainty in pollutant concentrations in the stock. The generic models of dwelling archetypes were developed using `CONTAM` and manipulated using an `R` code. They include nine contaminants emitted by construction materials, cigarettes, food preparation, heating stoves, human respiration, and showering, because they are directly or indirectly expected to lead to harm at a population scale. However, only $PM_{2.5}$ emissions from heater stoves located in living rooms and from the preparation of food in the kitchens are considered here. The remaining

sources can be included in future analyses. Dwelling and pollutant properties, such as windows opening area and schedules, or room volumes, are considered to be variable and can be manipulated by the code. They can also be updated in the future when new data becomes available.

The models were simulated probabilistically to estimate uncertainty in predictions. A total of 17,230 and 17,390 simulations were required to achieve consistent results for windows closed and open scenarios, respectively. The model estimates hourly and total exposures of one occupant to $PM_{2.5}$ during the winter season. Hourly and median ventilation rates are used to estimate winter heat loss due to exfiltration.

Air exchange rates and energy demand are higher for windows open scenario than for windows closed scenario, and consequently indoor pollutant concentrations are lower for windows open scenario.

The predicted M_{ed} ventilation rate of the stock is between 0.75 and 13.57 h^{-1}, the μ between 0.89 and 15.8 h^{-1}, the σ between 0.79 and 7.91 h^{-1}, and P_{95} between 2.4 and 29.75 h^{-1} using the two extreme scenarios. However, dwellings built after 2008 that are fully insulated are likely to be under–ventilated and more polluted when compared against the other archetypes, and so may require stricter source control measures and extra ventilation to lower health risks. The energy required to heat infiltrated air is predicted to be 22.35% of the total energy demand of an average Chilean dwelling.

Nationwide, the median (M_{ed}) exposures to $PM_{2.5}$ during winter is expected to be between 2.3 and 58.65 $\mu g/m^3$, the mean (μ) between 6.6 and 134.5 $\mu g/m^3$, and the P_{95} between 39.8 and 548.7 $\mu g/m^3$ using the two extreme scenarios.

Exposures, ventilation rates, and heat loss are significantly different when compared by archetype ID and region. The effects of ventilation rates and heat loss are substantive when the windows are open, but smaller when closed. The effect is also less noticeable on exposures for both scenarios.

When predicted concentrations are compared against international guidelines and the national ambient standard, Chilean dwellings are shown to be have unacceptable $PM_{2.5}$ concentrations >32% of the time, and so their occupants have an elevated risk of experiencing negative health outcomes, such as tracheal, bronchial, and lung cancers, Chronic Obstructive Pulmonary Disease, ischemic heart and cardiovascular diseases, and lower respiratory infections. Median ventilation rates for windows closed scenario are shown to be below 0.5 h^{-1} during 39% of the time, which is more than estimated times for dwellings in the USA, England, and Beijing.

The third objective was to determine the relationship between ventilation rates, the indoor exposures, and energy demand in the Chilean stock, taking into account the

diversity of dwellings and climates. To ensure that a 95% of the stock is below the WHO's guideline daily mean value of 25 $\mu g/m^3$, the ventilation rate should >13 h^{-1}. However, the associated energy cost to a maximum of 33.4 MWh for the winter season, which can be cost–prohibitive and so using a single ventilation rate in this way for all houses is sub–optimal. This ventilation rate can never be achieved by infiltration alone and so some window opening is always required. When opening windows (second 2) the WHO guideline value is achieved in around 50% of dwellings.

The fourth and final objective was to analyse the relative importance of the model inputs. The sensitivity analyses have shown that exposures to $PM_{2.5}$ are depend mostly on the emission rate of the internal sources and the total permeable envelope area (the sum of the normalised leakage area of the envelope and the total window area). Ventilation rates and heat losses are most affected by the total permeable area, the indoor–outdoor temperature difference, and the permeability of the envelope (Q_{50}).

These outcomes highlight the importance of, and urgency for, measures to achieve *acceptable* indoor air quality in Chilean dwellings, primarily by source control, but when this is not possible, by ventilation and dwelling design.

References

Arku, R E, Birch, A, Shupler, Mat, Yusuf, S, Hystad, P, & Brauer, M. 2018. Characterizing exposure to household air pollution within the Prospective Urban Rural Epidemiology (PURE) study. *Environment international*, **114**, 307–317.

ASHRAE 62.2. 2001. *Chapter 16: ventilation and infiltration*.

ASHRAE 62.2. 2016. *Standard 62.2 User's Manual: Based on ANSI/ASHRAE Standard 62.2-2016: Ventilation and Acceptable Indoor Air Quality in Residential Buildings*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Asikainen, A, Carrer, P, Kephalopoulos, S, de Oliveira Fernandes, E, Wargocki, P, & Hänninen, O. 2016. Reducing burden of disease from residential indoor air exposures in Europe (HEALTHVENT project). *Environmental Health*, **15**(1), S35.

Axley, J. 2007. Multizone airflow modeling in buildings: History and theory. *HVAC&R Research*, **13**(6), 907–928.

Balaras, C A, Dascalaki, E G, Droutsa, K G, & Kontoyiannidis, S. 2016. Empirical assessment of calculated and actual heating energy use in Hellenic residential buildings. *Applied Energy*, **164**, 115–132.

Ballarini, I, Corgnati, S P, & Corrado, V. 2014. Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project. *Energy policy*, **68**, 273–284.

Barría, R M, Calvo, M, & Pino, P. 2016. Contaminación intradomiciliaria por material particulado fino (MP2, 5) en hogares de recién nacidos. *Revista chilena de pediatría*, **87**(5), 343–350.

Bastani, A, Haghighat, F, & Kozinski, J A. 2012. Contaminant source identification within a building: toward design of immune buildings. *Building and Environment*, **51**, 320–329.

Booth, AT, Choudhary, R, & Spiegelhalter, DJ. 2012. Handling uncertainty in housing stock models. *Building and Environment*, **48**, 35–47.

Borsboom, W, De Gids, W, Logue, JM, Sherman, MH, & Wargocki, P. 2016. *TN 68: Residential Ventilation and Health*. Air Infiltration and Ventilation Centre, Brussels, Belgium.

Branco, PTBS, Alvim-Ferraz, MCM, Martins, FG, & Sousa, SIV. 2014. The microenvironmental modelling approach to assess children's exposure to air pollution–a review. *Environmental research*, **135**, 317–332.

Bravo-Linares, C, Ovando-Fuentealba, L, Orellana-Donoso, S, Gatica, S, Klerman, F, Mudge, S M, Gallardo, W, Pinaud, J P, & Loyola-Sepulveda, R. 2016. Source identification, apportionment and toxicity of indoor and outdoor PM 2.5 airborne particulates in a region characterised by wood burning. *Environmental Science: Processes & Impacts*, **18**(5), 575–589.

Bruinen de Bruin, Y, Koistinen, K, Kephalopoulos, S, Geiss, O, Tirendi, S, & Kotzias, D. 2008. Characterisation of urban inhalation exposures to benzene, formaldehyde and acetaldehyde in the European Union. *Environmental Science and Pollution Research*, **15**(5), 417–430.

BS5925, BSI. 1991. Code of practice for ventilation principles and designing for natural ventilation. *British Standard Institution*.

BSI. 2007. *BS EN 15251. Indoor Environment Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics.* Report. BSI.

Burgos, S, Ruiz, P, & Koifman, R. 2013. Changes to indoor air quality as a result of relocating families from slums to public housing. *Atmospheric environment*, **70**, 179–185.

Burke, J M, Zufall, M J, & ÃŰzkaynak, H. 2001. A population exposure model for particulate matter: case study results for PM2. 5 in Philadelphia, PA. *Journal of Exposure Science and Environmental Epidemiology*, **11**(6), 470.

Burnett, R T, Pope III, C A, Ezzati, M, Olives, C, Lim, S S, Mehta, S, Shin, H H, Singh, G, Hubbell, B, Brauer, M, *et al.* . 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental health perspectives*, **122**(4), 397–403.

Cabin, R J, & Mitchell, R J. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. *Bulletin of the Ecological Society of America*, **81**(3), 246–248.

Cáceres, D, Adonis, M, Retamal, C, Ancic, P, Valencia, M, Ramos, X, Olivares, N, & Gil, L. 2001. Indoor air pollution in a zone of extreme poverty of La Pintana, Santiago–Chile. *Revista Médica de Chile*, **129**(1), 33–42.

Cakmak, S, Dales, R E, & Blanco, C. 2007. Air pollution and mortality in Chile: susceptibility among the elderly. *Environmental Health Perspectives*, **115**(4), 524.

Cakmak, S, Dales, R E, & Blanco, C. 2009. Components of particulate air pollution and mortality in Chile. *International Journal of Occupational and Environmental Health*, **15**(2), 152–158.

Carrer, P, Wargocki, P, Fanetti, A, Bischof, W, De Oliveira, E, Hartmann, T, Kephalopoulos, S, Palkonen, S, & Seppänen, O. 2015. What does the scientific literature tell us about the ventilation–health relationship in public and residential buildings? *Building and Environment*, **94**, 273–286.

CASEN. 2015. *National Socioeconomic Characterization Survey (CASEN)*.

Cassee, F R, Héroux, M, Gerlofs-Nijland, M E, & Kelly, F J. 2013. Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission. *Inhalation Toxicology*, **25**(14), 802–812.

Castillo, C. 2001. *Estadística climatológica*.

CDT. 2010. *Estudio de usos finales y curva de oferta de la conservacion de la energía en el sector residencial.*

CENMA. 2011. *Evaluación de Impacto Atmosférico de Sistemas de Calefacción Domiciliaria*.

Center), CDIAC (Carbon Dioxide Information Analysis. 2018. *Fossil–Fuel CO2 Emissions by Nation*.

Cerezo, C, Sokol, J, AlKhaled, S, Reinhart, C, Al-Mumin, A, & Hajiah, A. 2017. Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City. *Energy and Buildings*, **154**, 321–334.

Chan, W R, Joh, J, & Sherman, M H. 2013. Analysis of air leakage measurements of US houses. *Energy and Buildings*, **66**, 616–625.

Chan, W R, Walker, I S, & Sherman, M H. 2015. Durable Airtightness in Single-Family Dwellings-Field Measurementsand Analysis. *International Journal of Ventilation*, **14**(1), 27–38.

Chang, J CS, & Guo, Z. 1992. Modeling of the fast organic emissions from a wood-finishing productâĂŤfloor wax. *Atmospheric Environment. Part A. General Topics*, **26**(13), 2365–2370.

Chapman, P M. 2007. Determining when contamination is pollutionâĂŤweight of evidence determinations for sediments and effluents. *Environment International*, **33**(4), 492–501.

Chen, C, & Zhao, B. 2011. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmospheric Environment*, **45**(2), 275–288.

Chen, W, Persily, A K, Hodgson, A T, Offermann, F J, Poppendieck, D, & Kumagai, K. 2014. Area-specific airflow rates for evaluating the impacts of VOC emissions in US single-family homes. *Building and Environment*, **71**, 204–211.

Chin, J-Y, Godwin, C, Parker, E, Robins, T, Lewis, T, Harbin, P, & Batterman, S. 2014. Levels and sources of volatile organic compounds in homes of children with asthma. *Indoor air*, **24**(4), 403–415.

Cifuentes, L A, Vega, J, Köpfer, K, & Lave, L B. 2000. Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality in Santiago, Chile. *Journal of the Air & Waste Management Association*, **50**(8), 1287–1298.

Citec UBB, & Decon UC. 2013. *Establecimiento de clases de infiltraciÃşn aceptables de edificios para Chile (Proyecto Fondef No D10 I 1025)*.

Cohen, A J, Brauer, M, Burnett, R, Anderson, H R, Frostad, J, Estep, K, Balakrishnan, K, Brunekreef, B, Dandona, L, Dandona, R, *et al.* . 2017. Estimates and 25–year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*, **389**(10082), 1907–1918.

Cometto-Muñiz, J E, & Abraham, M H. 2015. Compilation and analysis of types and concentrations of airborne chemicals measured in various indoor and outdoor human environments. *Chemosphere*, **127**, 70–86.

Dacunto, P J, Cheng, K-C, Acevedo-Bolton, V, Jiang, R-T, Klepeis, N E, Repace, J L, Ott, W R, & Hildemann, L M. 2013. Real-time particle monitor calibration factors and PM 2.5 emission factors for multiple indoor sources. *Environmental Science: Processes & Impacts*, **15**(8), 1511–1519.

Das, P, Shrubsole, C, Jones, B, Hamilton, I, Chalabi, Z, Davies, M, Mavrogianni, A, & Taylor, J. 2014. Using probabilistic sampling-based sensitivity analyses for indoor air quality modelling. *Building and environment*, **78**, 171–182.

Dawson, H E, & McAlary, T. 2009. A compilation of statistics for VOCs from post–1990 indoor air concentration studies in North American residences unaffected by subsurface vapor intrusion. *Groundwater Monitoring & Remediation*, **29**(1), 60–69.

Delgado-Saborit, j M. 2012. Use of real–time sensors to characterise human exposures to combustion related pollutants. *Journal of Environmental Monitoring*, **14**(7), 1824–1837.

Derbez, M, Wyart, G, Le Ponner, E, Ramalho, O, Ribéron, J, & Mandin, C. 2018. Indoor air quality in energy-efficient dwellings: Levels and sources of pollutants. *Indoor air*, **28**(2), 318–338.

Deru, M, & Burns, P. 2003. *Infiltration and natural ventilation model for whole-building energy simulation of residential buildings*. Tech. rept. National Renewable Energy Lab.(NREL), Golden, CO (United States).

Dictuc. 2009. Antecedentes para el Análisis General de Impacto Económico y Social del Anteproyecto de la Norma de Calidad Primaria para PM2.5 (AGIES).

Dols, W S, & Polidoro, B J. 2015. *CONTAM User Guide and Program Documentation Version 3.2*. Tech. rept. NIST.

Dols, W S, Emmerich, SJ, & Polidoro, B J. 2016. Coupling the multizone airflow and contaminant transport software CONTAM with EnergyPlus using co-simulation. *Pages 469–479 of: Building simulation*, vol. 9. Springer.

Dubrul, C. 1988. *Inhabitant behaviour with respect to ventilation-a summary report of IEA annex VIII*. Oscar Faber.

EPA. 2018. *About Air Data Reports*. Accessed: 2019-04-02.

Ewers, U, Krause, C, Schulz, Ch, & Wilhelm, M. 1999. Reference values and human biological monitoring values for environmental toxins. *International archives of occupational and environmental health*, **72**(4), 255–260.

Fazli, T, & Stephens, B. 2018. Development of a nationally representative set of combined building energy and indoor air quality models for US residences. *Building and Environment*, **136**, 198–212.

Fenton, N, & Neil, M. 2012. *Risk assessment and decision analysis with Bayesian networks*. Crc Press.

Ferguson, C J. 2009. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, **40**(5), 532.

Field, A, Miles, J, & Field, Z. 2012. *Discovering statistics using R*. Sage publications.

Filogamo, L, Peri, G, Rizzo, G, & Giaccone, A. 2014. On the classification of large residential buildings stocks by sample typologies for energy planning purposes. *Applied Energy*, **135**, 825–835.

Fischer, S L, & Koshland, C P. 2007. Field performance of a nephelometer in rural kitchens: effects of high humidity excursions and correlations to gravimetric analyses. *Journal of Exposure Science and Environmental Epidemiology*, **17**(2), 141.

Fissore A, Colonelli P. 2013. *Evaluación Independiente del Programa de Reacondicionamiento Térmico. Informe Final.*

Gakidou, E, Afshin, A, Abajobir, A A, Abate, K H, Abbafati, C, Abbas, K M, Abd-Allah, F, Abdulle, A M, Abera, S F, Aboyans, V, *et al.* . 2017. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, **390**(10100), 1345–1422.

Gaylor, D W. 2000. The use of HaberâĂŹs law in standard setting and risk assessment. *Toxicology*, **149**(1), 17–19.

GHDx. 2018. *Global Health Data Exchange*.

Ghiassi, N, & Mahdavi, A. 2017. Reductive bottom-up urban energy computing supported by multivariate cluster analysis. *Energy and Buildings*, **144**, 372–386.

Gigerenzer, G. 2004. Mindless statistics. *The Journal of Socio-Economics*, **33**(5), 587–606.

Graham, S E, & McCurdy, T. 2004. Developing meaningful cohorts for human exposure models. *Journal of Exposure Science and Environmental Epidemiology*, **14**(1), 23.

Guo, H, Lee, SC, Chan, LY, & Li, WM. 2004. Risk assessment of exposure to volatile organic compounds in different indoor environments. *Environmental Research*, **94**(1), 57–66.

Haghighat, F, & De Bellis, L. 1998. Material emission rates: literature review, and the impact of indoor air temperature and relative humidity. *Building and Environment*, **33**(5), 261–277.

Hamilton, I, Milner, J, Chalabi, Z, Das, P, Jones, B, Shrubsole, C, Davies, M, & Wilkinson, P. 2015. Health effects of home energy efficiency interventions in England: a modelling study. *BMJ open*, **5**(4), e007298.

He, C, Morawska, L, Hitchins, J, & Gilbert, D. 2004. Contribution from indoor sources to particle number and mass concentrations in residential houses. *Atmospheric environment*, **38**(21), 3405–3415.

Heal, M R, Kumar, P, & Harrison, R M. 2012. Particles, air quality, policy and health. *Chemical Society Reviews*, **41**(19), 6606–6630.

Helton, J C, & Davis, F J. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, **81**(1), 23–69.

Howard-Reed, Cynthia, Polidoro, Brian, & Dols, W Stuart. 2003. Development of IAQ Model Input Databases: Volatile Organic Compound Source Emission Rates. *Pages 1–14 of: Development of IAQ Model Input Databases: Volatile Organic Compound Source Emission Rates. Air and Waste Management Association Conference. Proceedings*. Citeseer.

Huang, Kailiang, Song, Jiasen, Feng, Guohui, Chang, Qunpeng, Jiang, Bian, Wang, Jun, Sun, Wen, Li, Huixing, Wang, Jinming, & Fang, Xianshi. 2018. Indoor air quality analysis of residential buildings in northeast China based on field measurements and longtime monitoring. *Building and Environment*, **144**, 171–183.

Hughes, Ifan, & Hase, Thomas. 2010. *Measurements and their uncertainties: a practical guide to modern error analysis*. Oxford University Press.

IEA. 2018. *Energy Policies Beyond IEA Countries*.

IHME, Institute for Health Metrics and Evaluation. 2017. *IHME*.

Ilabaca, Mauricio, Olaeta, Ignacio, Campos, Elizabeth, Villaire, Jeannette, Tellez-Rojo, Martha Maria, & Romieu, Isabelle. 1999. Association between levels of fine particulate and emergency visits for pneumonia and other respiratory illnesses among children in Santiago, Chile. *Journal of the Air & Waste Management Association*, **49**(9), 154–163.

INE. 2002. *Censo de Población y Vivienda 2002*.

INE. 2003. *Censo de Población y Vivienda 2002*.

INE. 2013. *Censo de Población y Vivienda 2012*.

INE. 2014. *Auditoría técnica a la base de datos del levantamiento censal año 2012*.

INE. 2016a. *Bases Edificación*.

INE. 2016b. *Encuesta Nacional sobre Uso del Tiempo*.

INE. 2018a. *Censo de Población y Vivienda 2017*.

INE. 2018b. *Censo de Población y Vivienda 2017*.

INN. 2016. *NCh1079 Of.2016 Arquitectura y Construcción- Zonificación climático habitacional para Chile*.

Ji, W, & Zhao, B. 2015. Estimating mortality derived from indoor exposure to particles of outdoor origin. *PLoS One*, **10**(4), e0124238.

Jones, B, Das, P, Chalabi, Z, Davies, M, Hamilton, I, Lowe, R, Milner, J, Ridley, I, Shrubsole, C, & Wilkinson, P. 2013. The effect of party wall permeability on estimations of infiltration from air leakage. *International Journal of Ventilation*, **12**(1), 17–30.

Jones, B, Das, P, Chalabi, Z, Davies, M, Hamilton, I, Lowe, R, Mavrogianni, A, Robinson, D, & Taylor, J. 2015. Assessing uncertainty in housing stock infiltration rates and associated heat loss: English and UK case studies. *Building and environment*, **92**, 644–656.

Jones, B, Phillips, G, O'Leary, C, Molina, C, Hall, I, & Sherman, M. 2018. Diagnostic barriers to using PM2. 5 concentrations as metrics of indoor air quality. *In: 39th AIVC Conference: 7th TightVent & 5th Venticool Conference. Ghent, Belgium.*

Jones, B M, Cook, M J, Fitzgerald, S D, & Iddon, C R. 2016. A review of ventilation opening area terminology. *Energy and Buildings*, **118**, 249–258.

Jones, J, Stick, S, Dingle, P, & Franklin, P. 2007. Spatial variability of particulates in homes: implications for infant exposure. *Science of the total environment*, **376**(1-3), 317–323.

Jorquera, H, & Barraza, F. 2012. Source apportionment of ambient PM2. 5 in Santiago, Chile: 1999 and 2004 results. *Science of the Total Environment*, **435**, 418–429.

Jorquera, H, & Barraza, F. 2013. Source apportionment of PM10 and PM2. 5 in a desert region in northern Chile. *Science of the Total Environment*, **444**, 327–335.

Jorquera, H, Palma, W, & Tapia, J. 2000. An intervention analysis of air quality data at Santiago, Chile. *Atmospheric Environment*, **34**(24), 4073–4084.

Jorquera, H, Barraza, F, Heyer, J, Valdivia, G, Schiappacasse, L, & Montoya, L. 2018. Indoor PM2.5 in an urban zone with heavy wood smoke pollution: The case of Temuco, Chile. *Environmental pollution (Barking, Essex : 1987)*, **236**(02), 477–487.

Kavgic, M, Mavrogianni, A, Mumovic, D, Summerfield, A, Stevanovic, Z, & Djurovic-Petrovic, M. 2010. A review of bottom-up building stock models for energy consumption in the residential sector. *Building and environment*, **45**(7), 1683–1697.

Kavouras, I G, Koutrakis, P, Cereceda-Balic, F, & Oyola, P. 2001. Source apportionment of PM10 and PM25 in five Chilean cities using factor analysis. *Journal of the Air & Waste Management Association*, **51**(3), 451–464.

Klepeis, N E, Nelson, W C, Ott, W R, Robinson, J P, Tsang, A M, Switzer, P, Behar, J V, Hern, S C, & Engelmann, W H. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science and Environmental Epidemiology*, **11**(3), 231.

Koehler, K, Good, N, Wilson, A, Mölter, A, Moore, B F, Carpenter, T, Peel, J L, & Volckens, J. 2018. The Fort Collins Commuter Study: Variability in Personal Exposure to Air Pollutants by Microenvironment. *Indoor air*.

Koutrakis, P, Sax, S N, Sarnat, J A, Coull, B, Demokritou, P, Demokritou, P, Oyola, P, Garcia, J, & Gramsch, E. 2005. Analysis of PM10, PM2. 5, and PM2. 5–10Concentrations in Santiago, Chile, from 1989 to 2001. *Journal of the Air & Waste Management Association*, **55**(3), 342–351.

Landrigan, P J. 2017. Air pollution and health. *The Lancet Public Health*, **2**(1), e4–e5.

Leiva G, M A., Santibañez, D A, Ibarra, S, Matus, P, & Seguel, R. 2013. A five-year study of particulate matter (PM2. 5) and cerebrovascular diseases. *Environmental pollution*, **181**, 1–6.

Liddament, M W. 1986. *Air infiltration calculation techniques: An applications guide*. Air infiltration and ventilation centre Berkshire, UK.

Liddament, M W, & Air Infiltration, Agence internationale de l'énergie. 1996. *A guide to energy efficient ventilation*. Air Infiltration and Ventilation Centre Coventry.

Logue, J M, Price, P N, Sherman, M H, & Singer, B C. 2011a. A method to estimate the chronic health impact of air pollutants in US residences. *Environmental Health Perspectives*, **120**(2), 216–222.

Logue, JM, McKone, TE, Sherman, MH, & Singer, BC. 2011b. Hazard assessment of chemical air contaminants measured in residences. *Indoor air*, **21**(2), 92–109.

Long, C M, Suh, H H, & Koutrakis, P. 2000. Characterization of indoor particle sources using continuous mass and size monitors. *Journal of the Air & Waste Management Association*, **50**(7), 1236–1250.

Lorenzetti, D M, Dols, W S, Persily, A K, & Sohn, M D. 2013. A stiff, variable time step transport solver for CONTAM. *Building and Environment*, **67**, 260–264.

Lozano, R, Fullman, N, Abate, D, Abay, S M, Abbafati, C, Abbasi, N, Abbastabar, H, Abd-Allah, F, Abdela, J, Abdelalim, A, *et al.* . 2018. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, **392**(10159), 2091–2138.

Marmot, M, Friel, S, Bell, R, Houweling, T AJ, Taylor, S, Commission on Social Determinants of Health, *et al.* . 2008. Closing the gap in a generation: health equity through action on the social determinants of health. *The lancet*, **372**(9650), 1661–1669.

Mata, É, Kalagasidis, A S, & Johnsson, F. 2014. Building-stock aggregation through archetype buildings: France, Germany, Spain and the UK. *Building and Environment*, **81**, 270–282.

McCurdy, T. 2015. *The dose profile is critical in understanding health risks associated with environmental exposures*. Tech. rept.

McCurdy, T, & Graham, S E. 2003. Using human activity data in exposure models: analysis of discriminating factors. *Journal of Exposure Science and Environmental Epidemiology*, **13**(4), 294.

Meteonorm. 2017. *METEOTEST (ed.) Version 7 ed. Switzerland*.

Milner, J, Shrubsole, C, Das, P, Jones, B, Ridley, I, Chalabi, Z, Hamilton, I, Armstrong, B, Davies, M, & Wilkinson, P. 2014. Home energy efficiency and radon related risk of lung cancer: modelling study. *Bmj*, **348**, f7493.

MINVU. 2007. *Ordenanza General de Urbanismo y Construcciones*.

Molina, C, Toro, R, Manzano, C, & Leiva-Guzmán, M A. 2017. Particulate matter in urban areas of south-central Chile exceeds air quality standards. *Air Quality, Atmosphere& Health*, **10**(5), 653–667.

Molina, C, Jackson, A, & Jones, B. 2019. The evaluation of real–time indoor environment parameters measured in 297 Chilean dwellings. *In: 40th AIVC Conference: 8th TightVent & 6th Venticool Conference. Ghent, Belgium.*

Murray, D M, & Burmaster, D E. 1995. Residential air exchange rates in the United States: empirical and estimated parametric distributions by season and climatic region. *Risk Analysis*, **15**(4), 459–465.

Nazaroff, W W, Gadgil, A J, & Weschler, C J. 1993. Critique of the use of deposition velocity in modeling indoor air quality. *In: Modeling of indoor air quality and exposure*. ASTM International.

Netatmo. 2018. *Personal weather station, howpublished = https://www.netatmo.com/en-GB/product/weather/weatherstation*. Accessed 02 July 2018.

Ng, L C, Musser, A, Persily, A K, & Emmerich, S J. 2012. Airflow and indoor air quality models of DOE reference commercial buildings. *Gaithersburg, MD, National Institute of Standards and Technology*, **163**.

Novikova, A, Csoknyai, T, & Szalay, Z. 2018. Low carbon scenarios for higher thermal comfort in the residential building sector of South Eastern Europe. *Energy Efficiency*, 1–31.

OECD. 2016. *Economic Consequences of Outdoor Air Pollution*. Organisation for Economic Co–operation and Development.

Oezkaynak, H, Xue, J, Weker, R, Butler, D, & Koutrakis, P. 1996. *Particle team (PTEAM) study: Analysis of the data. Final report, Volume 3*. Tech. rept. Harvard Univ., Boston, MA (United States). School of Public Health.

Oikonomou, E, Davies, M, Mavrogianni, A, Biddulph, P, Wilkinson, P, & Kolokotroni, M. 2012. Modelling the relative importance of the urban heat island and the thermal quality of dwellings for overheating in London. *Building and Environment*, **57**, 223–238.

O'Leary, C. 2018. Personal communication.

O'Leary, C, de Kluizenaar, Y, Jacobs, P, Borsboom, W, Hall, I, & Jones, B. 2019a. Investigating measurements of fine particle (PM 2.5) emissions from the cooking of meals and mitigating exposure using a cooker hood. *Indoor air*.

O'Leary, C, Jones, B, Dimitroulopoulou, S, & Hall, IP. 2019b. Setting the standard: The acceptability of kitchen ventilation for the English housing stock. *Building and Environment*, In.

Olson, D A, & Burke, J M. 2006. Distributions of PM2. 5 source strengths for cooking from the Research Triangle Park particulate matter panel study. *Environmental science & technology*, **40**(1), 163–169.

Orme, MS, Liddament, M, & Wilsom, A. 1994. An analysis and data summary of the AIVC's numerical database Technical Note 44. *Air Infiltration and Ventilation Centre, Coventry, England*.

Ott, W R, Steinemann, A C, & Wallace, L A. 2006. *Exposure Analysis*. CRC Press.

Palmer, J, & Cooper, I. 2013. United Kingdom housing energy fact file 2013. *Department of Energy & Climate Change, Prepared under contract to DECC by Cambridge Architectural Research, Eclipse Research Consultants and Cambridge Energy. The views expressed are not necessarily DECCâĂŹsp*, **172**.

Pérez-Lombard, L, Ortiz, J, & Pout, C. 2008. A review on buildings energy consumption information. *Energy and buildings*, **40**(3), 394–398.

Persily, A, Musser, A, & Leber, D. 2006. A collection of homes to represent the US housing stock. 2006. *National Institute of Standards and Technology: Washington DC*.

Persily, A, Musser, Amy, & Emmerich, Steven J. 2010. Modeled infiltration rate distributions for US housing. *Indoor Air*, **20**(6), 473–485.

Pino, P, Walter, T, Oyarzun, M, Villegas, R, & Romieu, I. 2004. Fine particulate matter and wheezing illnesses in the first year of life. *Epidemiology*, 702–708.

Prasauskas, T, Martuzevicius, D, Krugly, E, Ciuzas, D, Stasiulaitiene, I, Sidaraviciute, R, Kauneliene, V, Seduikyte, L, Jurelionis, A, & Haverinen-Shaughnessy, U. 2014. Spatial and temporal variations of particulate matter concentrations in multifamily apartment buildings. *Building and Environment*, **76**, 10–17.

Prieto-Parra, L, Yohannessen, K, Brea, C, Vidal, D, Ubilla, C A, & Ruiz-Rudolph, P. 2017. Air pollution, PM2.5 composition, source factors, and respiratory symptoms in asthmatic and nonasthmatic children in Santiago, Chile. *Environment international*, **101**, 190–200.

PUCV. 2019. *Estudio para caracterizar el parque construido de viviendas hasta el año 2018, en función de los principales parámetros que influyen en su demanda de energía*.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RENAM. 2018. *Red Nacional de Monitoreo de Viviendas*. Data retrieved from RENAM website, `http://renam.cl/site/acerca-de`.

Rioseco, R, & Tesser, C. 2014. Cartografía interactiva de los climas de Chile. *Instituto de Geografía, Pontificia Universidad Católica de Chile.[last accessed: 8 Sep 2014]. http://www7. uc. cl/sw_educ/geografia/cartografiainteractiva*.

Rivas, E, Barrios, S, Dorner, A, & Osorio, X. 2008. Fuentes de contaminación intradomiciliaria y enfermedad respiratoria en jardines infantiles y salas cunas de Temuco y Padre Las Casas, Chile. *Revista médica de Chile*, **136**(6), 767–774.

Rojas-Bracho, L, Suh, H H, Oyola, P, & Koutrakis, P. 2002. Measurements of children's exposures to particles and nitrogen dioxide in Santiago, Chile. *Science of the Total Environment*, **287**(3), 249–264.

Ruiz, P A, Toro, C, Cáceres, J, Lopez, G, Oyola, P, & Koutrakis, P. 2010. Effect of gas and kerosene space heaters on indoor air quality: a study in homes of Santiago, Chile. *Journal of the Air & Waste Management Association*, **60**(1), 98–108.

Sartori, I, & Hestnes, A G. 2007. Energy use in the life cycle of conventional and low-energy buildings: A review article. *Energy and buildings*, **39**(3), 249–257.

Sax, S N, Koutrakis, P, Ruiz-Rudolph, P A, Cereceda-Balic, F, Gramsch, E, & Oyola, P. 2007. Trends in the elemental composition of fine particulate matter in Santiago, Chile, from 1998 to 2003. *Journal of the Air & Waste Management Association*, **57**(7), 845–855.

Schueftan, A, & González, A D. 2015. Proposals to enhance thermal efficiency programs and air pollution control in south-central Chile. *Energy Policy*, **79**, 48–57.

SEC. 2007. *Decreto Supremo Nº 66 Reglamento de instalaciones interiores y medidores de gas.*

Sherman, M H, & Dickerhoff, D J. 1998. Air-tightness of US dwellings. *Transactions-American Society of Heating Refrigerating and Air Conditioning Engineers*, **104**, 1359–1367.

Shi, S, Chen, C, & Zhao, B. 2015. Air infiltration rate distributions of residences in Beijing. *Building and Environment*, **92**, 528–537.

Shipworth, M, Firth, S K, Gentry, M I, Wright, A J, Shipworth, D T, & Lomas, K J. 2010. Central heating thermostat settings and timing: building demographics. *Building Research & Information*, **38**(1), 50–69.

Siddharthan, T, Grigsby, M R, Goodman, D, Chowdhury, M, Rubinstein, A, Irazola, V, Gutierrez, L, Miranda, J J, Bernabe-Ortiz, A, Alam, D, *et al.* . 2018. Association between household air pollution exposure and chronic obstructive pulmonary disease outcomes in 13 low-and middle-income country settings. *American journal of respiratory and critical care medicine*, **197**(5), 611–620.

Singer, B C, & Delp, W W. 2018. Response of consumer and research grade indoor air quality monitors to residential sources of fine particles. *Indoor air*, **28**(4), 624–639.

Singleton, R, Salkoski, AJ, Bulkow, L, Fish, C, Dobson, J, Albertson, L, Skarada, J, Kovesi, T, McDonald, C, Hennessy, TW, *et al.* . 2017. Housing characteristics and indoor air quality in households of Alaska Native children with chronic lung conditions. *Indoor air*, **27**(2), 478–486.

Smith, K R., Bruce, N, Balakrishnan, K, Adair-Rohani, H, Balmes, J, Chafe, Z, Dherani, M, Hosgood, H D, Mehta, S, Pope, D, & Rehfuess, E. 2014. Millions Dead: How Do We Know and What Does It Mean? Methods Used in the Comparative Risk Assessment of Household Air Pollution. *Annual Review of Public Health*, **35**(1), 185–206. PMID: 24641558.

Sokol, J, Davila, C C, & Reinhart, C F. 2017. Validation of a Bayesian-based method for defining residential archetypes in urban building energy models. *Energy and Buildings*, **134**, 11–24.

Sousa, G, Jones, B M, Mirzaei, P A, & Robinson, D. 2017. A review and critique of UK housing stock energy models, modelling approaches and data sources. *Energy and Buildings*, **151**, 66–80.

Sousa, G, Jones, B M, Mirzaei, P A, & Robinson, D. 2018. An open-source simulation platform to support the formulation of housing stock decarbonisation strategies. *Energy and Buildings*, **172**, 459–477.

Steinle, S, Reis, S, Sabel, C E, Semple, S, Twigg, M M, Braban, C F, Leeson, S R, Heal, M R, Harrison, D, Lin, C, *et al.* . 2015. Personal exposure monitoring of PM2. 5 in indoor and outdoor microenvironments. *Science of the Total Environment*, **508**, 383–394.

Stuardo, E. 1996. Natural radiation measurements in Chile. *Radiation protection dosimetry*, **67**(2), 129–133.

Sullivan, G M, & Feinn, R. 2012. Using effect sizeâĂŤor why the P value is not enough. *Journal of graduate medical education*, **4**(3), 279–282.

Sundell, J, Levin, H, Nazaroff, W W, Cain, W S, Fisk, W J, Grimsrud, D T, Gyntelberg, F, Li, Y, Persily, AK, Pickering, AC, *et al.* . 2011. Ventilation rates and health: multidisciplinary review of the scientific literature. *Indoor air*, **21**(3), 191–204.

Swami, M V, & Chandra, S. 1987. Procedures for calculating natural ventilation airflow rates in buildings. *ASHRAE Final Report FSEC-CR-163-86, ASHRAE Research Project*.

Tang, H, Richards, G, Benner, C L, Tuominen, J P, Lee, M L, Lewis, E A, Hansen, L D, & Eatough, D J. 1990. Solanesol: a tracer for environmental tobacco smoke particles. *Environmental Science & Technology*, **24**(6), 848–852.

Taylor, R. 1990. Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, **6**(1), 35–39.

TenWolde, A, & Pilon, C L. 2007. The effect of indoor humidity on water vapor release in homes.

Thatcher, T L, Lai, A CK, Moreno-Jackson, R, Sextro, R G, & Nazaroff, W W. 2002. Effects of room furnishings and air speed on particle deposition rates indoors. *Atmospheric environment*, **36**(11), 1811–1819.

The World Bank. 2018. *World Bank Indicators*.

Toro, R, Canales, M, Gonzalez-Rojas, C, *et al.* . 2014. Inhaled and inspired particulates in Metropolitan Santiago Chile exceed air quality standards. *Building and Environment*, **79**, 115–123.

Traynor, GW, Apte, MG, Dillworth, JF, Hollowell, CD, & Sterling, EM. 1982. The effects of ventilation on residential air pollution due to emissions from a gas-fired range. *Environment International*, **8**(1-6), 447–452.

Tucker, WG. 2000. An overview of PM2.5 sources and control strategies. *Fuel Processing Technology*, **65**, 379–392.

UN. 2019. *Sustainable Development Goals*.

Underhill, LJ, Fabian, MP, Vermeer, K, Sandel, M, Adamkiewicz, G, Leibler, JH, & Levy, JI. 2018. Modeling the resiliency of energy-efficient retrofits in low-income multifamily housing. *Indoor air*, **28**(3), 459–468.

Wang, C Michelle, Barratt, B, Carslaw, N, Doutsi, A, Dunmore, R E, Ward, M W, & Lewis, A C. 2017. Unexpectedly high concentrations of monoterpenes in a study of UK homes. *Environmental Science: Processes & Impacts*, **19**(4), 528–537.

Weschler, CJ. 2009. Changes in indoor pollutants since the 1950s. *Atmospheric Environment*, **43**(1), 153–169.

WHO, *et al.* . 2000. Air quality guidelines for Europe.

WHO, *et al.* . 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide-Global update 2005-Summary of risk assessment, 2006. *Geneva: WHO*.

WHO, *et al.* . 2007. *Housing, energy and thermal comfort: a review of 10 countries within the WHO European Region*. Tech. rept. Copenhagen: WHO Regional Office for Europe.

Wilson, W E, Mage, D T, & Grant, L D. 2000. Estimating separately personal exposure to ambient and nonambient particulate matter for epidemiology and risk assessment: why and how. *Journal of the Air & Waste Management Association*, **50**(7), 1167–1183.

Yang, L, Ye, M, *et al.* . 2014. CFD simulation research on residential indoor air quality. *Science of the Total Environment*, **472**, 1137–1144.

Yang, W, Lee, K, Yoon, C, Yu, S, Park, K, & Choi, W. 2011. Determinants of residential indoor and transportation activity times in Korea. *Journal of exposure science and environmental epidemiology*, **21**(3), 310.

Yong, Z. 2016. *Digital universal particle concentration sensor pms5003 series data manual*.

Yu, S, He, L, & Feng, G. 2015. The Transient Simulation of Carbon Dioxide Emission from Human Body Based on CONTAM. *Procedia Engineering*, **121**, 1613–1619.

Zielinskia, J M, & Chambers, D B. 2008. Mapping of residential radon in the world.

APPENDIX A

Additional information about the housing stock

This appendix presents additional information obtained from the data sources described in Chapter 3, and expanded material used in the classification and description of the archetypes presented in Chapter 4.

A.1 Regional archetypes

Table A.1: Regional archetypes

Parameter No:	1	2	3	4	5	6	7	8	9	10	11	12	13
Parameter	ID	Region	Frequency	P(type)	Waighting factor	Proportion of total	Accum. Prop	Cum. Freq.	Walls	Roof	Floor	Cooking fuel	Predominant SDI
Raw database:	Geometric archetype	2002 census	2002 census	NA					2002 census	2002 census	2002 census	2002 census	2002 census
Variable name:		15 Regions 2017	f(Parameter2)	Type of Houses	(Parameter3) x (Parameter4)	f(Parameter3)	f(Parameter6)	f(Parameter7)	V4A	V4B	V4C	H12	CSE Decil
	27	1	5550	0.1954	4069	0.070%	0.1%	1084	3	5	6	2	6
	27	2	12209	0.1954	8952	0.154%	0.2%	3470	3	5	6	2	6
	27	3	9759	0.1954	7156	0.123%	0.3%	5377	3	5	6	2	5
	27	4	25697	0.1954	18842	0.325%	0.7%	10398	3	5	6	2	4
	27	5	61713	0.1954	45250	0.780%	1.5%	22457	3	5	3	2	5
	27	6	29827	0.1954	21870	0.377%	1.8%	28285	2	5	3	2	4
	27	7	42351	0.1954	31053	0.535%	2.4%	36561	2	5	3	2	3
	27	8	77261	0.1954	56651	0.976%	3.3%	51657	4	5	3	2	3
	27	9	34741	0.1954	25473	0.439%	3.8%	58446	4	4	3	4	2
	27	10	28062	0.1954	20576	0.354%	4.1%	63929	4	4	3	4	2
	27	11	3814	0.1954	2797	0.048%	4.2%	64674	4	4	3	4	3
	27	12	7289	0.1954	5345	0.092%	4.3%	66099	4	4	4	1	6
	27	13	180744	0.1954	132528	2.283%	6.6%	101416	2	5	3	2	7
	27	14	15481	0.1954	11351	0.196%	6.8%	104441	4	4	3	4	2
	27	15	5513	0.1954	4042	0.070%	6.8%	105518	3	5	6	2	5
	36	1	4712	0.1954	3455	0.060%	6.9%	921	3	5	5	2	7
	36	2	12549	0.1954	9201	0.159%	7.0%	3373	3	4	6	2	8
	36	3	8648	0.1954	6341	0.109%	7.1%	5063	3	5	6	2	6
	36	4	21188	0.1954	15536	0.268%	7.4%	9203	2	5	5	2	6
	36	5	52727	0.1954	38661	0.666%	8.1%	19506	2	5	3	2	6
	36	6	26412	0.1954	19366	0.334%	8.4%	24667	2	5	3	2	5
	36	7	26491	0.1954	19424	0.335%	8.8%	29843	2	5	3	2	4
	36	8	63963	0.1954	46900	0.808%	9.6%	42341	4	5	3	2	4
	36	9	30001	0.1954	21998	0.379%	9.9%	48203	4	4	3	4	3
	36	10	22710	0.1954	16652	0.287%	10.2%	52641	4	4	3	4	3
	36	11	2664	0.1954	1953	0.034%	10.3%	53162	4	4	3	4	4
	36	12	5694	0.1954	4175	0.072%	10.3%	54274	4	4	4	1	7
	36	13	174977	0.1954	128300	2.210%	12.5%	88465	2	5	3	2	8
	36	14	12967	0.1954	9508	0.164%	12.7%	90998	4	4	3	4	2
	36	15	4839	0.1954	3548	0.061%	12.8%	91944	3	5	6	2	6
	91	1	5550	0.1434	2986	0.051%	12.8%	796	3	5	6	2	6
	91	2	12209	0.1434	6569	0.113%	12.9%	2547	3	5	6	2	6
	91	3	9759	0.1434	5251	0.090%	13.0%	3946	3	5	6	2	5
	91	4	25697	0.1434	13827	0.238%	13.3%	7631	3	5	6	2	4
	91	5	61713	0.1434	33207	0.572%	13.8%	16480	3	5	3	2	5
	91	6	29827	0.1434	16049	0.277%	14.1%	20757	2	5	3	2	4
	91	7	42351	0.1434	22788	0.393%	14.5%	26830	2	5	3	2	3
	91	8	77261	0.1434	41573	0.716%	15.2%	37908	4	5	3	2	3
	91	9	34741	0.1434	18693	0.322%	15.5%	42890	4	4	3	4	2
	91	10	28062	0.1434	15100	0.260%	15.8%	46914	4	4	3	4	2
	91	11	3814	0.1434	2052	0.035%	15.8%	47461	4	4	3	4	3
	91	12	7289	0.1434	3922	0.068%	15.9%	48506	4	4	4	1	6
	91	13	180744	0.1434	97255	1.676%	17.6%	74423	2	5	3	2	7
	91	14	15481	0.1434	8330	0.144%	17.7%	76643	4	4	3	4	2
	91	15	5513	0.1434	2966	0.051%	17.8%	77434	3	5	6	2	5
	100	1	4712	0.1434	2535	0.044%	17.8%	676	3	5	5	2	7
	100	2	12549	0.1434	6752	0.116%	17.9%	2475	3	4	6	2	8
	100	3	8648	0.1434	4653	0.080%	18.0%	3715	3	5	6	2	6
	100	4	21188	0.1434	11401	0.196%	18.2%	6753	2	5	5	2	6
	100	5	52727	0.1434	28371	0.489%	18.7%	14314	2	5	3	2	6
	100	6	26412	0.1434	14212	0.245%	18.9%	18101	2	5	3	2	5
	100	7	26491	0.1434	14254	0.246%	19.2%	21900	2	5	3	2	4
	100	8	63963	0.1434	34417	0.593%	19.8%	31072	4	5	3	2	4
	100	9	30001	0.1434	16143	0.278%	20.1%	35374	4	4	3	4	3
	100	10	22710	0.1434	12220	0.211%	20.3%	38630	4	4	3	4	3
	100	11	2664	0.1434	1433	0.025%	20.3%	39012	4	4	3	4	4
	100	12	5694	0.1434	3064	0.053%	20.3%	39829	4	4	4	1	7
	100	13	174977	0.1434	94152	1.622%	22.0%	64919	2	5	3	2	8
	100	14	12967	0.1434	6977	0.120%	22.1%	66778	4	4	3	4	2

100	15	4839	0.1434	2604	0.045%	22.1%	67472	3	5	6	2	6
275	1	5550	0.0995	2072	0.036%	22.2%	552	3	5	6	2	6
275	2	12209	0.0995	4557	0.079%	22.2%	1767	3	5	6	2	6
275	3	9759	0.0995	3643	0.063%	22.3%	2737	3	5	6	2	5
275	4	25697	0.0995	9592	0.165%	22.5%	5293	3	5	6	2	4
275	5	61713	0.0995	23035	0.397%	22.9%	11432	3	5	3	2	5
275	6	29827	0.0995	11133	0.192%	23.1%	14399	2	5	3	2	4
275	7	42351	0.0995	15808	0.272%	23.3%	18612	2	5	3	2	3
275	8	77261	0.0995	28839	0.497%	23.8%	26297	4	5	3	2	3
275	9	34741	0.0995	12968	0.223%	24.1%	29753	4	4	3	4	2
275	10	28062	0.0995	10475	0.180%	24.2%	32544	4	4	3	4	2
275	11	3814	0.0995	1424	0.025%	24.3%	32923	4	4	3	4	3
275	12	7289	0.0995	2721	0.047%	24.3%	33648	4	4	4	1	6
275	13	180744	0.0995	67465	1.162%	25.5%	51627	2	5	3	2	7
275	14	15481	0.0995	5779	0.100%	25.6%	53167	4	4	3	4	2
275	15	5513	0.0995	2058	0.035%	25.6%	53715	3	5	6	2	5
35	1	2366	0.1954	1735	0.030%	25.6%	462	3	5	5	2	7
35	2	6503	0.1954	4768	0.082%	25.7%	1733	3	4	6	2	7
35	3	4707	0.1954	3451	0.059%	25.8%	2653	3	5	6	2	6
35	4	11077	0.1954	8122	0.140%	25.9%	4817	2	5	5	2	5
35	5	29860	0.1954	21894	0.377%	26.3%	10652	2	5	3	2	6
35	6	15053	0.1954	11037	0.190%	26.5%	13593	2	5	3	2	5
35	7	16338	0.1954	11980	0.206%	26.7%	16786	2	5	3	2	4
35	8	33794	0.1954	24779	0.427%	27.1%	23389	4	5	3	2	4
35	9	13331	0.1954	9775	0.168%	27.3%	25994	4	4	3	2	3
35	10	10972	0.1954	8045	0.139%	27.4%	28138	4	4	3	4	3
35	11	1155	0.1954	847	0.015%	27.4%	28364	4	4	3	4	4
35	12	3418	0.1954	2506	0.043%	27.5%	29031	4	4	4	1	6
35	13	101965	0.1954	74765	1.288%	28.8%	48955	2	5	3	2	8
35	14	5921	0.1954	4342	0.075%	28.8%	50112	4	4	3	4	3
35	15	2438	0.1954	1788	0.031%	28.9%	50589	3	5	6	2	6
19	1	4194	0.1954	3075	0.053%	28.9%	820	3	5	6	2	5
19	2	8969	0.1954	6576	0.113%	29.0%	2572	3	5	6	2	6
19	3	6096	0.1954	4470	0.077%	29.1%	3763	3	5	6	2	4
19	4	16575	0.1954	12153	0.209%	29.3%	7002	2	5	6	2	4
19	5	23395	0.1954	17154	0.296%	29.6%	11573	4	5	3	2	5
19	6	11814	0.1954	8662	0.149%	29.8%	13882	2	5	5	2	3
19	7	18172	0.1954	13324	0.230%	30.0%	17433	2	5	3	2	2
19	8	31947	0.1954	23425	0.404%	30.4%	23675	4	4	3	2	3
19	9	23638	0.1954	17332	0.299%	30.7%	28294	4	4	3	4	2
19	10	18501	0.1954	13566	0.234%	30.9%	31909	4	4	3	4	2
19	11	3843	0.1954	2818	0.049%	31.0%	32660	4	4	3	4	2
19	12	2491	0.1954	1826	0.031%	31.0%	33147	4	4	4	1	5
19	13	75443	0.1954	55318	0.953%	32.0%	47888	2	5	3	2	6
19	14	8819	0.1954	6466	0.111%	32.1%	49612	4	4	3	4	2
19	15	3909	0.1954	2866	0.049%	32.1%	50375	3	5	6	2	5
284	1	4712	0.0995	1759	0.030%	32.2%	469	3	5	5	2	7
284	2	12549	0.0995	4684	0.081%	32.2%	1717	3	4	6	2	8
284	3	8648	0.0995	3228	0.056%	32.3%	2577	3	5	6	2	6
284	4	21188	0.0995	7909	0.136%	32.4%	4685	2	5	5	2	6
284	5	52727	0.0995	19681	0.339%	32.8%	9930	2	5	3	2	6
284	6	26412	0.0995	9859	0.170%	32.9%	12557	2	5	3	2	5
284	7	26491	0.0995	9888	0.170%	33.1%	15192	2	5	3	2	4
284	8	63963	0.0995	23875	0.411%	33.5%	21554	4	5	3	2	4
284	9	30001	0.0995	11198	0.193%	33.7%	24539	4	4	3	4	3
284	10	22710	0.0995	8477	0.146%	33.9%	26798	4	4	3	4	3
284	11	2664	0.0995	994	0.017%	33.9%	27063	4	4	3	4	4
284	12	5694	0.0995	2125	0.037%	33.9%	27629	4	4	4	1	7
284	13	174977	0.0995	65313	1.125%	35.0%	45034	2	5	3	2	8
284	14	12967	0.0995	4840	0.083%	35.1%	46324	4	4	3	4	2
284	15	4839	0.0995	1806	0.031%	35.2%	46805	3	5	6	2	6

A.2 Size of houses

Table A.2: Floor area by type of house. Values assigned to each house according to the number of rooms. **Type A:** Detached house, age band 1. **Type B:** Semidetached or End–terrace house, age band 1. **Type C:** Mid-terrace house, age band 1. **Type D:** Flat, age band 1. **Type E:** Detached house, age band 2. **Type F:** Semidetached or End–terrace house, age band 2. **Type G:** Mid-terrace house, age band 2. **Type H:** Flat, age band 2.

Floor Areas				
House Type	N. of zones	Average	SD	Median
Type A				
	1	20.45	17.95	18
	2	27.36	15.57	25
	3	37.74	16.54	36
	4	44.39	24.20	41
	5	49.39	30.36	42
	6	53.68	35.99	43
	7	57.00	41.32	44
	8	60.10	47.67	45
Type B				
	1	16.97	6.82	15
	2	23.72	11.16	21
	3	35.86	12.78	39
	4	42.14	15.01	43
	5	43.83	16.83	44
	6	44.60	18.08	44
	7	44.71	18.45	44
	8	44.82	18.81	44
Type C				
	1	17.65	4.03	15
	2	27.17	10.90	27
	3	38.67	46.94	38
	4	41.46	40.99	40
	5	42.82	41.01	40
	6	43.44	41.36	40
	7	44.04	41.74	41
	8	44.16	41.92	41
Type D				
	1	96.98	178.90	50
	2	51.25	31.68	43
	3	60.87	28.61	49
	4	68.57	35.52	55
	5	75.54	43.89	60
	6	79.64	49.74	63
	7	81.02	51.98	64
	8	82.13	54.32	64
Type E				
	1	35.14	30.97	29
	2	39.64	25.16	38
	3	52.99	17.70	50
	4	59.77	20.09	56
	5	68.76	33.93	58
	6	72.04	38.55	59
	7	74.44	42.93	60
	8	76.66	47.61	60
Type F				
	1	25.52	24.45	15
	2	43.06	21.61	41
	3	53.20	11.98	55
	4	57.45	12.48	56
	5	60.02	17.27	56
	6	60.67	19.79	56
	7	61.04	21.10	56
	8	61.30	21.82	56
Type G				
	1	51.21	49.03	45
	2	45.54	26.18	45
	3	56.45	25.97	52
	4	60.91	23.11	56
	5	66.63	34.00	57
	6	68.72	37.21	58
	7	70.09	39.86	58
	8	71.02	41.73	58

A.3 Heating hours by geographic region

Table A.3: Heating hours for winter days by weather station.

Weather file name	Mean indoor temperature	SD of indoor temperature	Daily Heating hours
Antofagasta	16.91	1.83	12.91
Antofagasta city	17.80	1.86	9.83
Arica, Chacalluta	18.00	1.40	3.99
Arica city	19.81	1.57	2.3
Atacama	16.30	4.16	11.88
Calama El Loa	16.52	3.40	12.04
Chillán city	12.58	3.50	17.09
Chillán-Gral ohiggi	12.53	3.54	17.02
Concepción city	12.56	3.33	16.94
Concepción	12.71	3.38	17.34
Copiapo Chamonate	16.23	4.10	11.82
Coyhaique Teniente V.	9.03	4.42	18.01
Curico-General Fr	12.98	4.90	15.25
Iquique city	19.42	1.58	3.43
Iquique Diego Aracen	18.53	1.48	6.65
La Serena city	15.37	2.78	14.89
La Serena, La Florida	14.43	2.96	15.84
Osorno-Canal Bajo	10.25	3.62	17.93
Osorno city	10.22	3.73	17.98
Puerto Aysen Airport	9.56	4.38	17.96
Puerto Montt El Tepu	10.26	3.39	18.11
Puerto Montt city	10.27	3.49	18.02
Puerto Williams	5.90	2.91	18.3
Punta Arenas	5.82	3.00	18.3
Punta Arenas city	6.67	2.80	18.3
Rancagua city	12.42	4.98	15.77
Santiago	12.72	4.70	15.61
Talca city	12.66	4.21	16.31
Talcahuano city	13.52	3.39	16.39
Temuco Maquehue	11.08	4.13	17.3
Temuco city	11.06	4.10	17.3
Valdivia Pichoy	10.94	4.07	17.5
Valdivia city	10.77	3.87	17.34
Vallenar Airport	12.55	3.41	17.16
Valparaíso (USM)	13.50	4.81	15.09

APPENDIX B

Additional information on the occupant activities

B.1 Use of time data

Table B.1: Activity duration in hours by geographic region (INE, 2016b).

Region	**Working hours**			**In-Transit**			**Sleeping hours WD**			**Sleeping hours WE**		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
1	8.29	8.00	2.55	1.52	0.58	8.33	7.19	7.00	2.14	7.81	8.00	2.27
2	8.44	8.00	2.79	4.35	0.83	24.88	7.10	7.00	1.73	7.79	8.00	1.99
3	8.83	9.00	2.51	0.99	0.67	1.59	7.31	7.00	1.99	7.54	8.00	2.04
4	8.04	8.00	2.86	5.07	0.67	27.74	7.21	7.00	1.80	7.72	8.00	1.95
5	8.19	8.00	2.69	3.12	0.67	19.73	7.37	7.00	1.81	7.83	8.00	2.10
6	8.31	8.00	2.64	3.60	0.67	20.03	7.33	7.00	1.95	8.01	8.00	2.18
7	8.20	8.00	2.25	1.65	0.67	10.45	7.83	8.00	1.60	8.44	9.00	1.86
8	7.87	8.00	2.65	1.19	0.67	5.46	7.33	7.00	1.78	7.91	8.00	1.98
9	7.91	8.00	2.47	2.09	0.67	14.99	7.49	8.00	1.73	8.04	8.00	1.94
10	7.96	8.00	2.61	2.06	0.58	14.85	7.82	8.00	1.79	8.45	8.00	1.95
11	8.13	8.00	2.15	0.64	0.50	0.85	7.72	8.00	1.62	8.63	9.00	2.11
12	8.30	8.00	2.34	1.84	0.50	14.32	7.06	7.00	1.56	7.67	8.00	1.87
13	8.31	9.00	2.45	2.34	1.17	12.89	7.18	7.00	1.81	7.84	8.00	2.02
14	7.64	8.00	2.55	0.78	0.58	0.94	7.44	7.75	1.85	7.93	8.00	1.96
15	8.24	9.00	2.95	12.76	0.67	44.19	7.31	7.00	1.54	8.03	8.00	1.73

Table B.2: Activity duration in hours by geographic region (INE, 2016b).

Region	**Shower/dressed - WD**		**Shower/dressed - WE**		**Breakfast - WD**		**Breakfast - WE**		**Lunch - WD**		**Lunch - WE**	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1	0.62	0.50	0.66	0.50	0.35	0.25	0.43	0.33	0.59	0.50	0.75	0.50
2	0.66	0.50	0.67	0.50	0.34	0.25	0.40	0.33	0.54	0.50	0.65	0.50
3	0.65	0.50	0.68	0.50	0.37	0.33	0.45	0.33	0.57	0.50	0.72	0.50
4	0.65	0.50	0.67	0.50	0.33	0.25	0.39	0.33	0.55	0.50	0.67	0.50
5	0.64	0.50	0.65	0.50	0.33	0.25	0.39	0.33	0.59	0.50	0.73	0.50
6	0.62	0.50	0.64	0.50	0.30	0.25	0.34	0.25	0.51	0.50	0.62	0.50
7	0.45	0.33	0.48	0.42	0.30	0.25	0.37	0.33	0.55	0.50	0.71	0.67
8	0.62	0.50	0.64	0.50	0.34	0.25	0.40	0.33	0.52	0.50	0.65	0.50
9	0.58	0.50	0.59	0.50	0.34	0.25	0.40	0.33	0.56	0.50	0.68	0.50
10	0.60	0.50	0.61	0.50	0.33	0.25	0.40	0.33	0.56	0.50	0.71	0.50
11	0.52	0.50	0.53	0.50	0.35	0.25	0.42	0.33	0.51	0.50	0.62	0.50
12	0.54	0.50	0.56	0.50	0.31	0.25	0.35	0.33	0.55	0.50	0.60	0.50
13	0.61	0.50	0.63	0.50	0.33	0.25	0.40	0.33	0.61	0.50	0.78	0.67
14	0.68	0.50	0.67	0.50	0.34	0.25	0.41	0.33	0.54	0.50	0.68	0.50
15	0.54	0.50	0.59	0.50	0.32	0.25	0.39	0.33	0.64	0.50	0.77	0.75

Table B.3: Activity duration in hours by geographic region (INE, 2016b).

Region	**Tea - WD**		**Tea - WE**		**Dinner - WD**		**Dinner - WE**	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1	0.52	0.50	0.55	0.50	0.55	0.50	0.58	0.50
2	0.51	0.50	0.54	0.50	0.54	0.50	0.57	0.50
3	0.56	0.50	0.54	0.50	0.56	0.50	0.61	0.50
4	0.46	0.33	0.50	0.50	0.48	0.50	0.50	0.50
5	0.51	0.50	0.57	0.50	0.55	0.50	0.61	0.50
6	0.45	0.33	0.49	0.50	0.48	0.33	0.51	0.42
7	0.57	0.50	0.65	0.50	0.52	0.35	0.54	0.42
8	0.49	0.50	0.55	0.50	0.51	0.50	0.62	0.50
9	0.49	0.50	0.54	0.50	0.56	0.50	0.59	0.50
10	0.52	0.50	0.57	0.50	0.57	0.50	0.64	0.50
11	0.51	0.50	0.56	0.50	0.52	0.50	0.60	0.50
12	0.43	0.50	0.45	0.50	0.57	0.50	0.63	0.50
13	0.50	0.50	0.55	0.50	0.60	0.50	0.65	0.50
14	0.57	0.50	0.59	0.50	0.54	0.50	0.62	0.50
15	0.47	0.50	0.51	0.50	0.58	0.50	0.62	0.50

B.2 Doors and fans daily schedules

Table B.4: Daily schedules for doors and use of fans. WD: Weekday; WE: Weekend; D: door; F: Fan; 1: on; 0: off.

Meals	**Time**	**Kitchen D**		**Kitchen F**		**Bedroom D**		**Bathroom D**		**Bathroom F**	
		WE	WD	WD	WE	WD	WE	WD	WE	WD	WE
	0:00:00	0	0	0	0	0	0	0	0	0	0
Meal 1 WD	7:00:00	0	1	0	0	1	0	1	0	1	0
	7:05:00	0	1	0	0	1	0	0	0	1	0
	7:20:00	0	1	1	0	1	0	0	0	1	0
	7:30:00	0	1	0	0	1	0	1	0	1	0
	7:35:00	0	1	0	0	1	0	0	0	1	0
	8:00:00	0	1	0	0	1	0	0	0	0	0
	8:30:00	0	1	0	0	1	0	0	1	0	1
	8:35:00	0	1	0	0	1	0	0	0	0	1
Meal 1 WE	9:00:00	1	0	0	1	1	0	0	1	0	1
	9:05:00	1	0	0	1	1	0	0	0	0	1
	9:20:00	1	0	0	0	1	0	0	0	0	1
	9:30:00	1	0	0	0	1	1	0	0	0	0
	10:00:00	0	0	0	0	1	1	0	0	0	0
	11:00:00	0	0	0	0	1	1	0	1	0	0
	11:05:00	0	0	0	0	1	1	0	0	0	0
Meal 2 WD & WE	12:00:00	1	1	1	0	1	1	0	1	0	0
	12:05:00	1	1	1	0	1	1	0	0	0	0
	12:30:00	1	1	1	1	1	1	0	0	0	0
	13:00:00	1	1	0	1	1	1	0	0	0	0
	13:30:00	1	1	0	0	1	1	0	0	0	0
	14:00:00	0	0	0	0	1	1	0	1	0	0
	14:05:00	0	0	0	0	1	1	0	0	0	0
Meal 3 WD	17:00:00	0	1	0	0	1	1	0	1	0	0
	17:05:00	0	1	0	0	1	1	0	0	0	0
Meal 3 WE	18:00:00	0	0	0	0	1	1	0	0	0	0
	19:30:00	0	0	1	0	1	1	0	0	0	0
	19:50:00	0	0	0	0	1	1	0	0	0	0
	20:00:00	1	0	0	1	1	1	0	0	0	0
Meal WD & WE	20:30:00	1	1	1	0	1	1	0	0	0	0
	21:00:00	1	0	0	0	1	1	0	1	1	0
	21:05:00	1	0	0	0	1	1	0	0	1	0
	21:30:00	1	0	0	0	1	1	1	0	0	0
	21:35:00	1	0	0	0	1	1	0	0	0	0
	22:00:00	0	0	0	0	0	1	1	1	0	1
	22:05:00	0	0	0	0	0	1	0	0	0	1
	22:30:00	0	0	0	0	0	1	0	0	0	0
	23:00:00	0	0	0	0	0	0	0	0	0	0
	24:00:00	0	0	0	0	0	0	0	1	0	0

APPENDIX C

Sensitivity Analyses

Scatter plots

This appendix presents additional plots of the sensitivity analyses presented in chapter 6. The scatter plots of each input versus the three outputs are shown for both scenario 1 and 2 separately.

C.1 Scatter plots, scenario 1

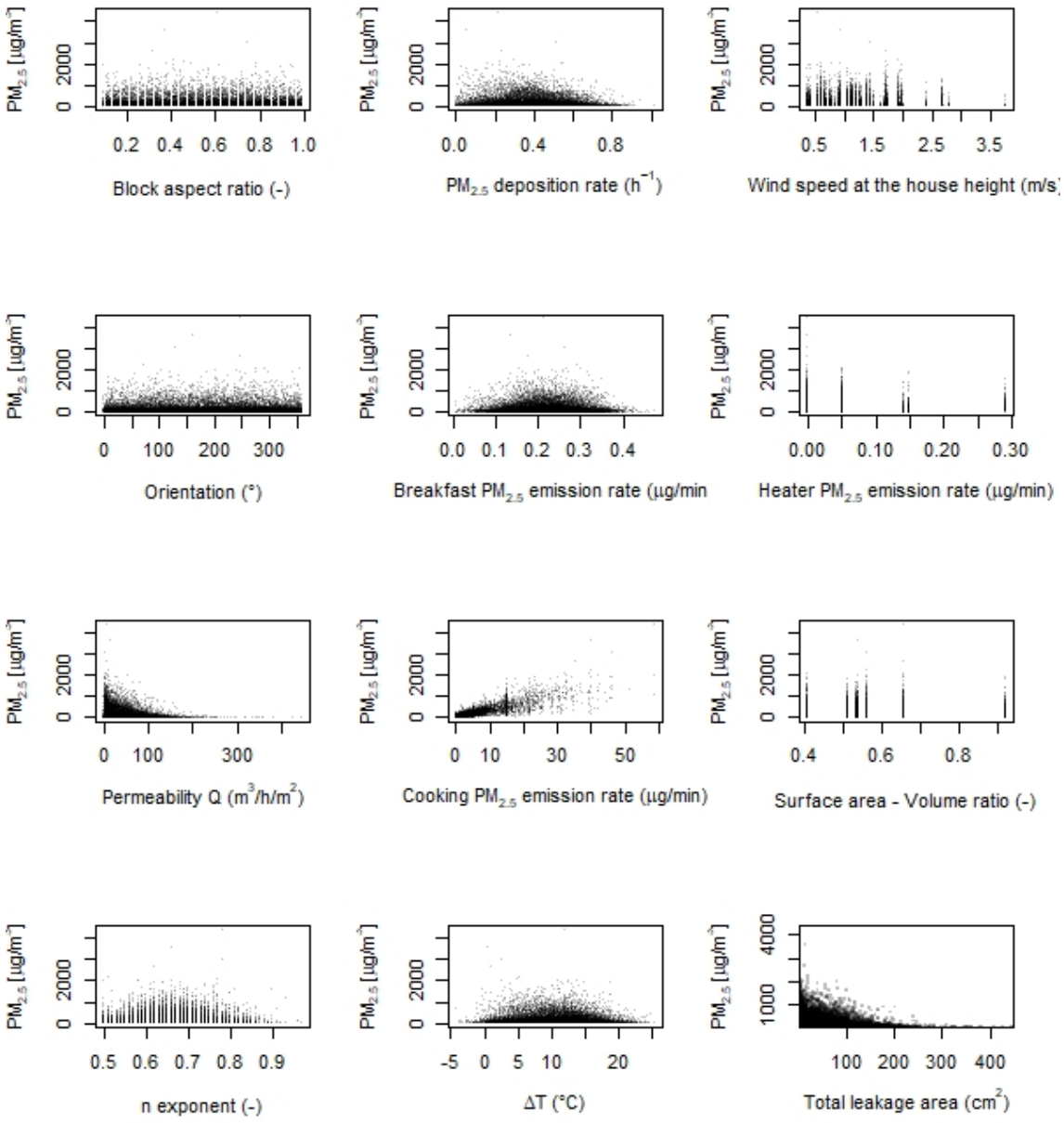


Figure C.1: Scenario 1. Inputs versus $PM_{2.5}$ exposures.

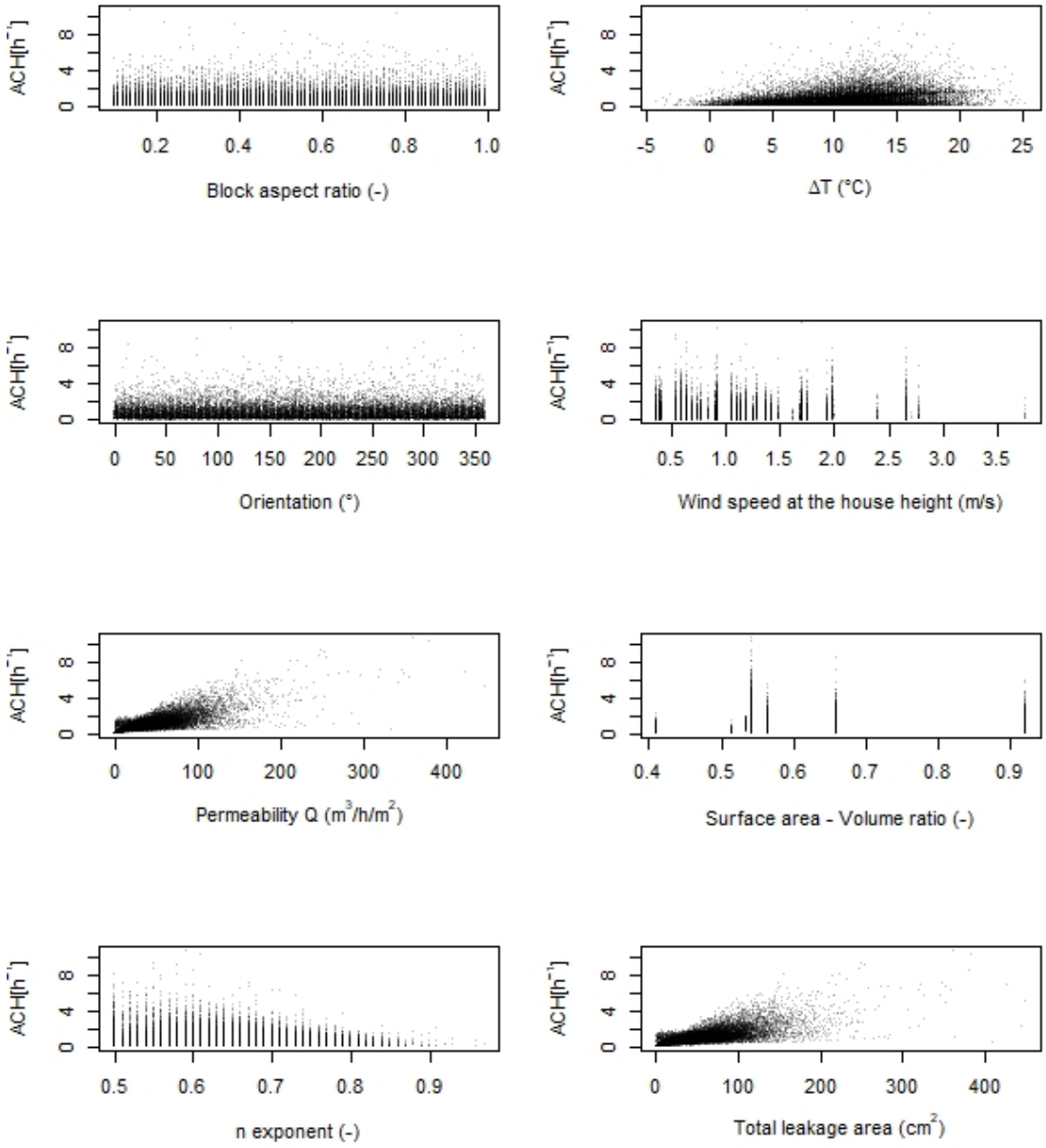


Figure C.2: Scenario 1. Inputs versus ventilation rates.

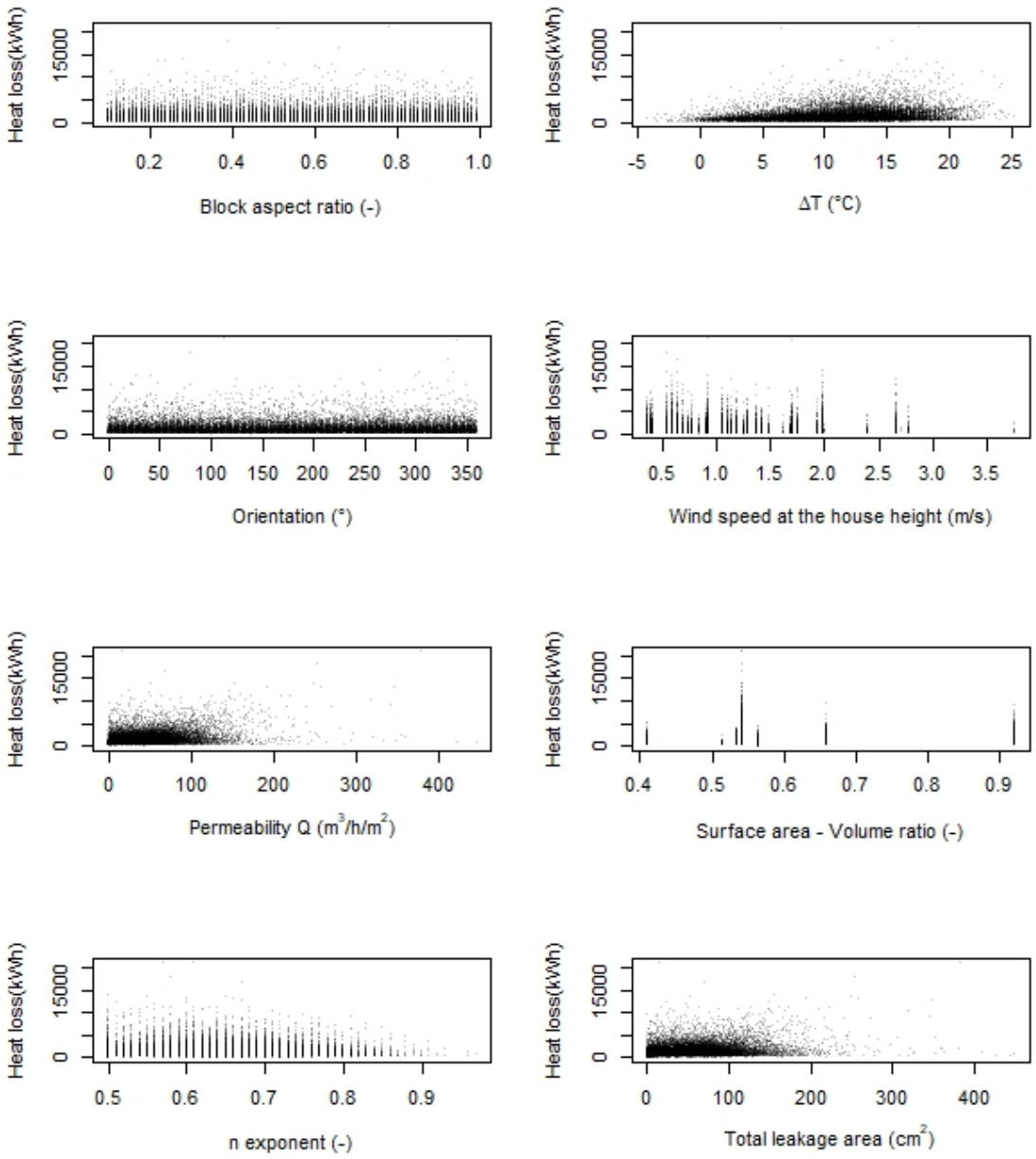


Figure C.3: Scenario 1. Inputs versus heat loss.

C.2 Scatter plots, scenario 2

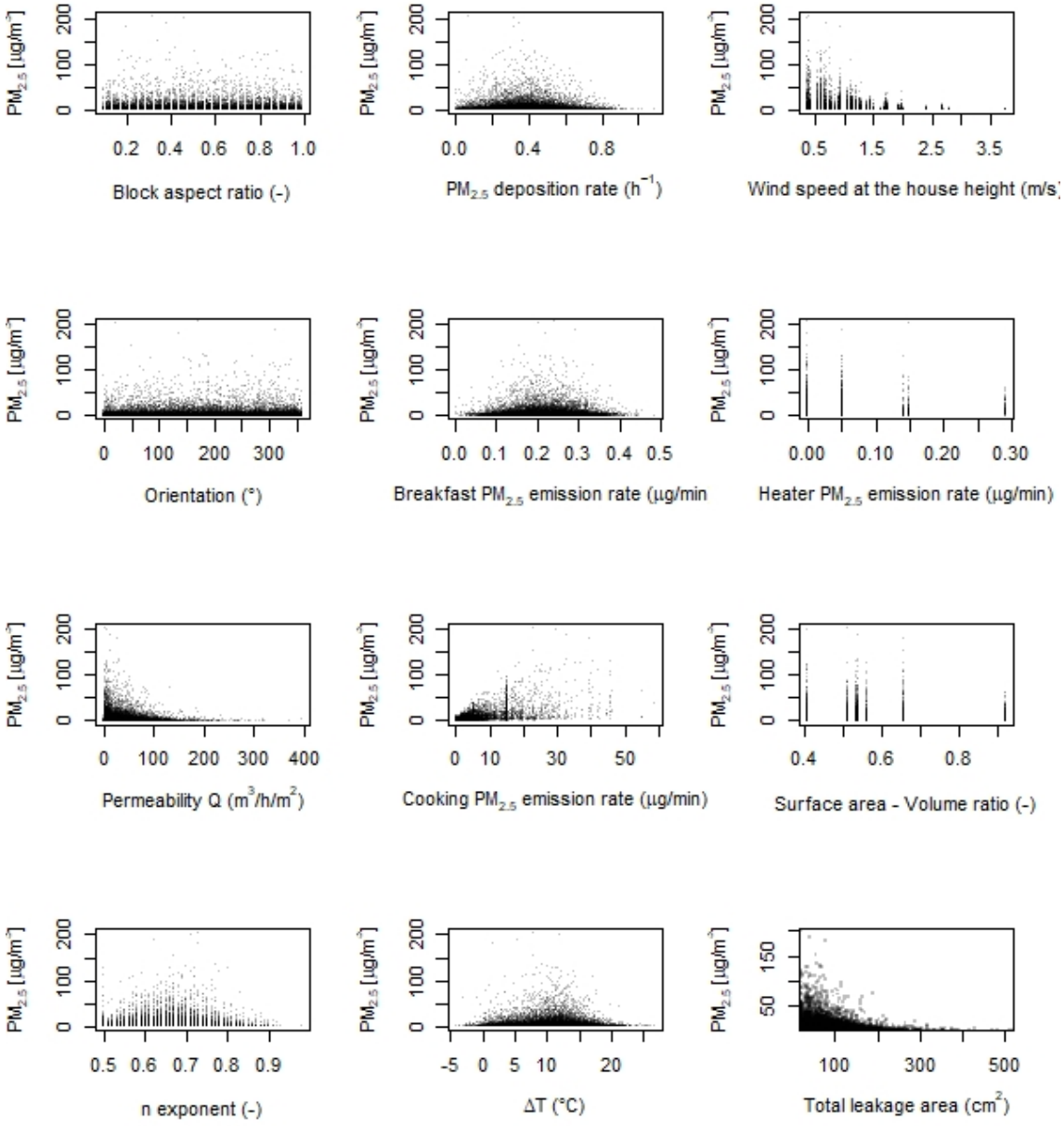


Figure C.4: Scenario 2. Inputs versus $PM_{2.5}$ exposures.

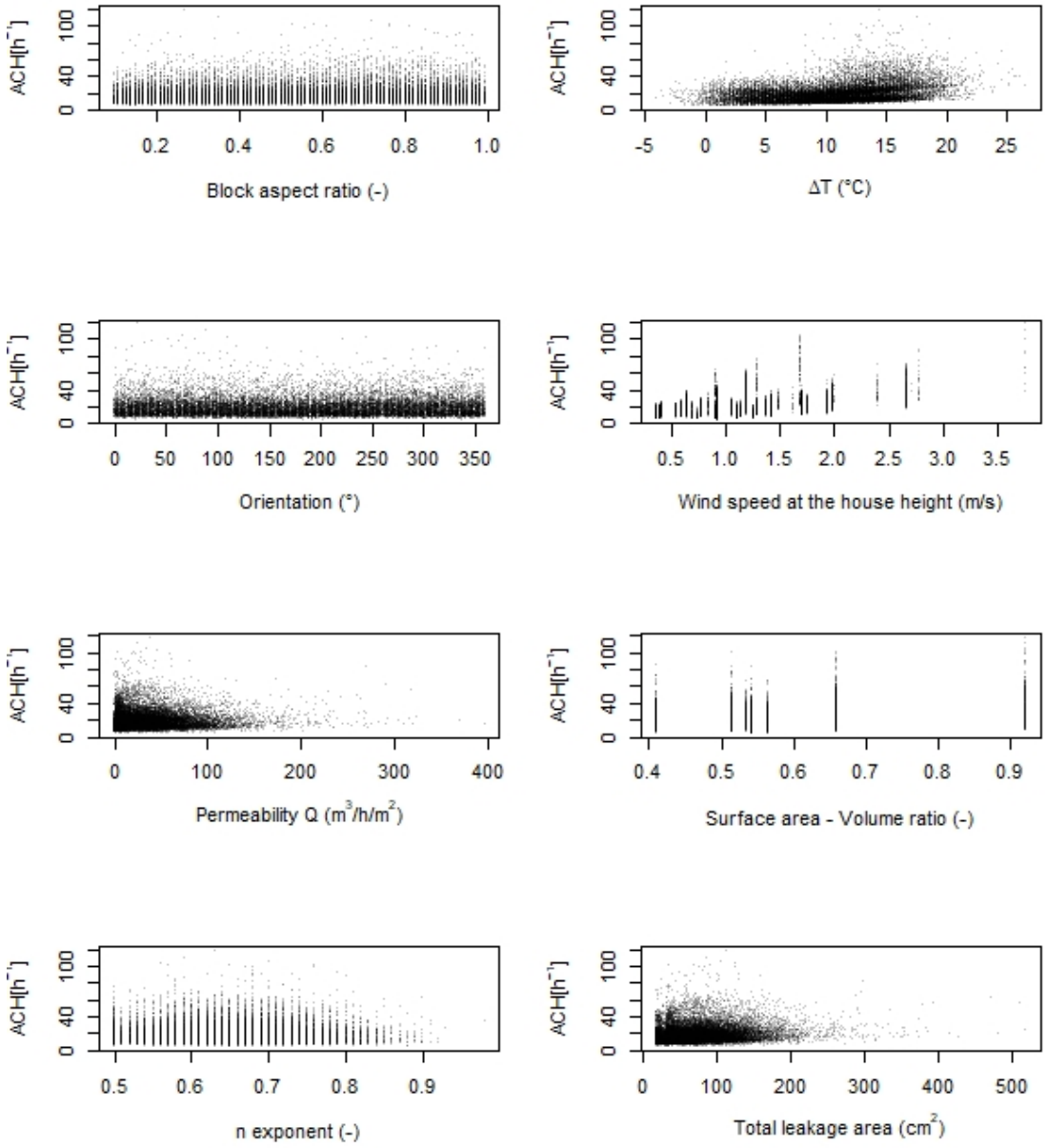


Figure C.5: Scenario 2. Inputs versus ventilation rates.

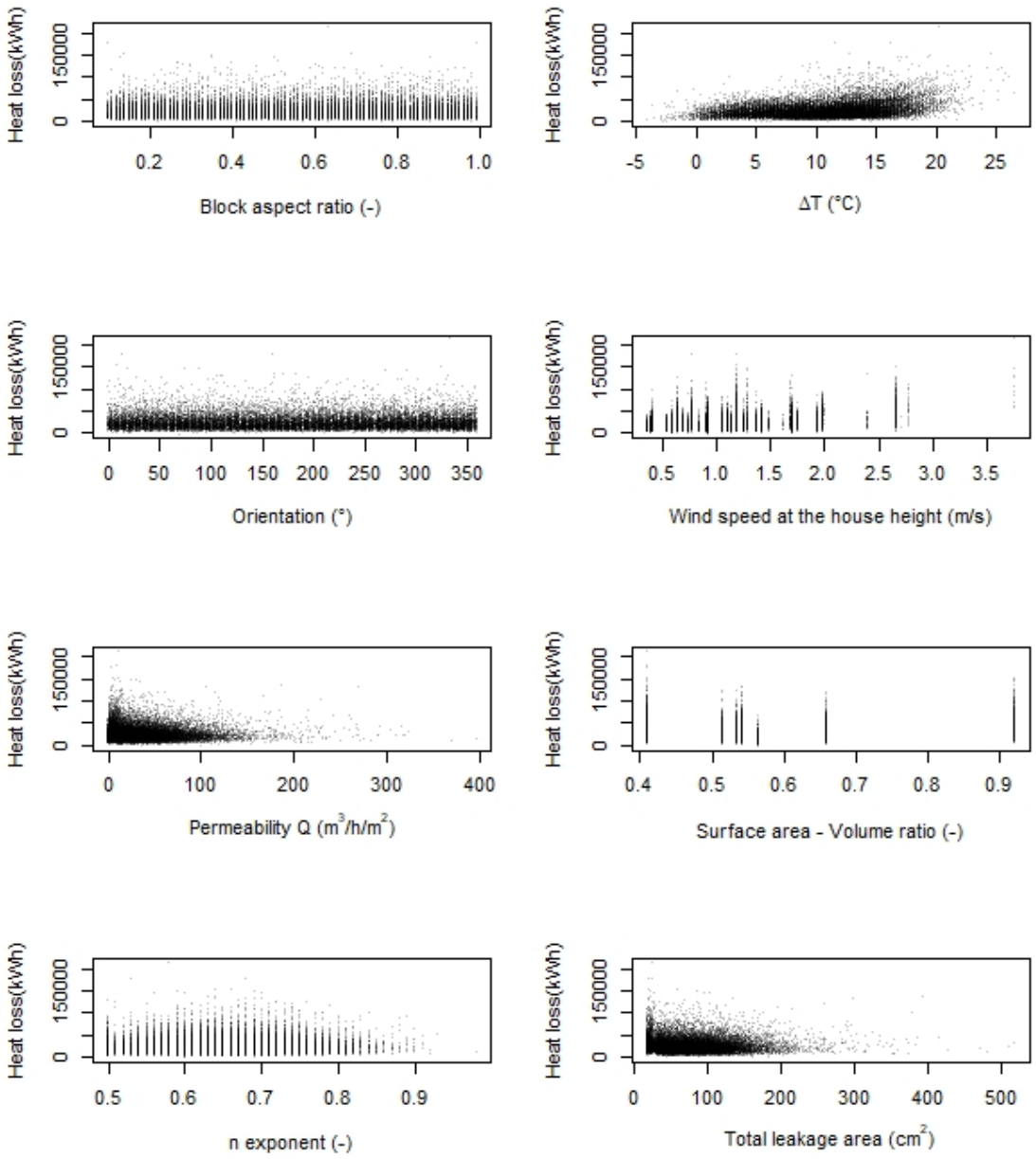


Figure C.6: Scenario 2. Inputs versus heat loss.

APPENDIX D

Sensitivity Analyses

Results

Appendix D provides the full tables of results of the sensitivity analyses for the three outcomes, including the tests statistics p–values, ranking coefficients and ranks. For further details on the implementation of this tests and its discussion see Chapters 5 and 7, respectively.

D.1 Exposure

Table D.1: Test statistics for correlation.

Input	Kendall	*p*-value	Rank	Pearson	*p*-value	Rank	Spearman	*p*-value	Rank
Length: width ratio	-0.00496	0.168418416	10	-0.00805	0.13426465	9	-0.00735	0.171317695	10
Orientation	0.00273	0.446234667	11	-0.00149	0.78137068	12	0.00414	0.441072388	11
Permeability	-0.05190	1.49951E-47	5	-0.10566	1.62093E-86	3	-0.07701	1.06774E-46	5
n exponent	0.02862	4.3866E-15	6	0.04577	1.58988E-17	5	0.04228	3.54264E-15	6
PM25 Dep. rate	-0.02366	4.07441E-11	7	-0.05745	1.03809E-26	4	-0.03556	3.64765E-11	7
Breakfast PM2.5	-0.00037	0.917887846	12	0.00315	0.557707582	11	-0.00058	0.913882897	12
Cooking PM2.5	0.38542	0	1	0.58379	0	1	0.53207	0	1
Delta T	0.01976	3.83606E-08	8	-0.03069	1.11527E-08	7	0.03099	8.07816E-09	8
Wind speed	-0.08765	2.3719E-126	4	-0.00518	0.334758995	10	-0.12551	1.503E-121	4
Heater PM2.5	0.13252	3.5903E-209	3	0.04008	8.60825E-14	6	0.16396	3.65E-207	3
SurfaceArea : Volume ratio	0.00914	0.017758492	9	-0.01894	0.000424256	8	0.01241	0.020928894	9
Total Permeable area	-0.23311	0	2	-0.21239	0	2	-0.34662	0	2

Table D.2: Test statistics for regression.

Input	Regression	Rank	R^2	Rank Regression	Rank
L:W ratio	-0.00805	9	0.00006	-0.00735	10
Orientation	-0.00149	12	0.00000	0.00414	11
Permeability	-0.10566	3	0.01116	-0.07701	5
n exponent	0.04577	5	0.00209	0.04228	6
PM25 Dep. rate	-0.05745	4	0.00330	-0.03556	7
Breakfast PM2.5	0.00315	11	0.00001	-0.00058	12
Cooking PM2.5	0.58379	1	0.34081	0.53207	1
Delta T	-0.03069	7	0.00094	0.03099	8
Wind speed	-0.00518	10	0.00003	-0.12551	4
Heater PM2.5	0.04008	6	0.00161	0.16396	3
SurfaceArea : Volume ratio	-0.01894	8	0.00036	0.01241	9
Total perm. area	-0.21239	2	0.04511	-0.34662	2

Table D.3: Test statistics for group comparison

Input	KW2	*p*-value	Rank	KW5	*p*-value	Rank	KW10	*p*-value	Rank	KW20	*p*-value	Rank	KS	*p*-value	Rank
L:W ratio	1.10	0.293453	11	1.85	0.76421274	11	15.78	0.071557663	10	17.16	0.578709363	12	0.01009	0.339531057	11
Orientation	1.96	0.161099	10	6.02	0.197665535	10	12.51	0.186129964	12	22.48	0.261247211	10	0.01024	0.322734561	10
Permeability	50.98	9.308E-13	5	329.21	5.39957E-70	5	534.19	2.7323E-109	6	565.50	9.4721E-108	6	0.07234	7.18587E-40	6
n exponent	17.86	2.3774E-05	6	75.10	1.89596E-15	7	93.08	3.90857E-16	8	111.11	4.96928E-15	8	0.01853	0.005166754	9
PM25 Dep. rate	16.94	3.8670E-05	7	46.59	1.85545E-09	9	94.79	1.77381E-16	7	105.62	5.09013E-14	9	0.03109	1.03422E-07	8
Breakfast PM2.5	0.10	0.74654	12	0.93	0.920904946	12	12.92	0.166322218	11	18.78	0.470912661	11	0.00737	0.733125267	12
Cooking PM2.5	5730.57	0	1	8533.80	0	1	11579.88	0	1	12672.53	0	1	0.35283	0	1
Delta T	5.43	0.01982	9	260.20	4.13783E-55	6	806.82	7.2234E-168	5	821.46	5.8791E-162	5	0.06869	5.43255E-36	7
Wind speed	773.67	2.8656E-170	3	1249.84	2.4981E-269	4	1384.27	1.9356E-292	4	1469.42	1.6209E-300	4	0.14735	4.3682E-164	5
Heater PM2.5	548.05	3.3479E-121	4	1929.56	0	3	3311.81	0	3	3380.31	0	3	0.16609	2.3519E-208	4
S:V ratio	9.84	0.00170	8	58.17	7.02223E-12	8	77.27	5.61818E-13	9	116.07	5.98815E-16	7	0.17692	2.4052E-236	3
Total perm. area	2232.24	0	2	4090.40	0	2	4436.93	0	2	4453.08	0	2	0.23017	0	2

D.2 Ventilation rates

Table D.4: Test statistics for correlation.

Input	Kendall	*p*-value	Rank	Pearson	*p*-value	Rank	Spearman	*p*-value	Rank
Length: width ratio	0.009608308	0.007664459	7	0.026513268	8.06173E-07	5	0.014300292	0.007795136	7
Orientation	-0.00162402	0.650852475	8	-0.004281721	0.425653615	8	-0.002454146	0.647948694	8
Permeability	0.156455981	0	2	0.019546349	0.000275727	7	0.192639284	1.1933E-286	2
n exponent	-0.046778575	1.29915E-37	6	-0.020236882	0.00016614	6	-0.068167787	6.0723E-37	6
Delta T	0.08309111	2.6249E-118	3	0.127131987	1.0904E-124	4	0.121481561	6.0681E-114	3
Wind speed	0.06919405	1.80102E-79	4	0.181326779	1.1422E-253	2	0.092352119	1.89164E-66	5
SurfaceArea : Volume ratio	0.06812354	8.33197E-70	5	0.143492513	1.2001E-158	3	0.098350875	3.7253E-75	4
Total Permeable area	0.386898496	0	1	0.313008908	0	1	0.536300648	0	1

Table D.5: Test statistics for regression.

Input	Regression	Rank	R^2	Rank Regression	Rank
Length: width ratio	0.026513268	5	0.00070295	0.014300292	7
Orientation	-0.004281721	8	1.8333E-05	-0.002454146	8
Permeability	0.019546349	7	0.00038206	0.192639284	2
n exponent	-0.020236882	6	0.00040953	-0.068167787	6
Delta T	0.127131987	4	0.01616254	0.121481561	3
Wind speed	0.181326779	2	0.0328794	0.092352119	5
SurfaceArea : Volume ratio	0.143492513	3	0.0205901	0.098350875	4
Total Permeable area	0.313008908	1	0.09797458	0.536300648	1

Table D.6: Test statistics for group comparison

Input	KW2	*p*-value	Rank	KW5	*p*-value	Rank	KW10	*p*-value	Rank	KW20	*p*-value	Rank	KS	*p*-value	Rank
Length: width ratio	0.14	0.711040825	8	31.27	2.69972E-06	7	60.14	1.2583E-09	7	62.82	1.37251E-06	7	0.025400926	2.74039E-05	7
Orientation	0.27	0.602294151	7	1.66	0.798498116	8	3.28	0.952146008	8	18.53	0.487207489	8	0.005808627	0.931471432	8
Permeability	2700.76	0	2	11149.07	0	2	10194.09	0	2	11568.63	0	2	0.326693309	0	2
n exponent	409.50	4.70789E-91	4	848.30	2.6407E-182	6	754.83	1.1193E-156	6	928.82	8.1208E-185	6	0.054460298	8.82102E-23	6
Delta T	535.13	2.161E-118	3	1712.92	0	5	1938.97	0	5	2834.69	0	5	0.114054123	1.78346E-98	5
Wind speed	46.15	1.09743E-11	6	3138.12	0	3	3475.05	0	3	4204.09	0	3	0.173230183	1.2939E-226	4
SurfaceArea : Volume ratio	115.97	4.8181E-27	5	2239.91	0	4	2450.66	0	4	3087.77	0	4	0.221281667	0	3
Total Permeable area	10879.00	0	1	14887.32	0	1	15068.06	0	1	15714.28	0	1	0.423171156	0	1

D.3 Heat loss

Table D.7: Test statistics for correlation.

Input	Kendall	*p*-value	Rank	Pearson	*p*-value	Rank	Spearman	*p*-value	Rank
Length: width ratio	0.001200887	0.738915526	7	0.0045271	0.399616399	6	0.001766343	0.742426135	7
Orientation	-0.000636156	0.859278784	8	-0.000713372	0.894408124	8	-0.00097369	0.856239937	8
Permeability	0.043843387	2.00063E-34	3	-0.023834386	9.19855E-06	5	0.060201258	3.59771E-29	3
n exponent	-0.012525911	0.00059806	6	-0.003832688	0.475781339	7	-0.018398525	0.000618251	5
Delta T	0.076395032	2.5422E-100	2	0.143286037	3.422E-158	2	0.112811537	1.98491E-98	2
Wind speed	0.03126401	1.47366E-17	4	0.127945754	2.8097E-126	3	0.042304384	3.41876E-15	4
SurfaceArea : Volume ratio	-0.015995261	3.36755E-05	5	0.064209961	5.81859E-33	4	-0.018108111	0.000753235	6
Total Permeable area	0.247284271	0	1	0.200521035	0	1	0.366575983	0	1

Table D.8: Test statistics for regression.

Input	Regression	Rank	R^2	Rank Regression	Rank
Length: width ratio	0.0045271	6	2.0495E-05	0.001766343	7
Orientation	-0.000713372	8	5.089E-07	-0.00097369	8
Permeability	-0.023834386	5	0.00056808	0.060201258	3
n exponent	-0.003832688	7	1.4689E-05	-0.018398525	5
Delta T	0.143286037	2	0.02053089	0.112811537	2
Wind speed	0.127945754	3	0.01637012	0.042304384	4
SurfaceArea : Volume ratio	0.064209961	4	0.00412292	-0.018108111	6
Total Permeable area	0.200521035	1	0.04020869	0.366575983	1

Table D.9: Test statistics for group comparison

Input	KW2	*p*-value	Rank	KW5	*p*-value	Rank	KW10	*p*-value	Rank	KW20	*p*-value	Rank	KS	*p*-value	Rank
Length: width ratio	0.00	0.962431365	8	1.88	0.758194243	8	8.45	0.489098981	7	13.59	0.807236812	8	0.006004917	0.913080391	7
Orientation	0.00	0.945083559	7	1.89	0.756784427	7	3.02	0.963376175	8	14.11	0.777374204	7	0.005414999	0.961038785	8
Permeability	659.35	2.0698E-145	2	2623.84	0	2	2693.03	0	2	2924.03	0	2	0.142469869	2.034E-153	3
n exponent	18.36	1.83269E-05	5	37.68	1.30272E-07	6	57.45	4.13859E-09	6	74.49	1.62631E-08	6	0.03798407	2.6673E-11	6
Delta T	361.47	1.3484E-80	3	1074.17	3.0082E-231	3	1460.36	0	3	1984.28	0	3	0.091869654	4.83817E-64	5
Wind speed	48.41	3.44922E-12	4	838.94	2.8158E-180	4	1238.51	5.8698E-261	4	1428.71	8.8411E-292	4	0.099579768	3.61969E-75	4
SurfaceArea : Volume ratio	8.68	0.003217368	6	495.84	5.3165E-106	5	595.54	1.8997E-122	5	631.06	1.392E-121	5	0.18633223	4.0082E-262	2
Total Permeable area	5121.11	0	1	5921.85	0	1	6129.70	0	1	6208.35	0	1	0.260963121	0	1

APPENDIX E

Effect sizes

Group difference index used:
Cohen's d = d = $(M_{ed1} - M_{ed2})$ / σ_{pop} Where $\sigma = \sigma_{pop}$

Thresholds: (Sullivan & Feinn, 2012)

- Small = 0.2
- Medium = 0.5
- Large = 0.8
- Very large = 1.3

Using the $PM_{2.5}$ σ of the populations:

- $PM_{2.5}$ Scenario 1: 210.28 $\mu g/m^3$
- $PM_{2.5}$ Scenario 2: 12.26 $\mu g/m^3$
- ACH Scenario 1: 0.8854 h^{-1}
- ACH Scenario 2: 15.8^{-1}
- Heat loss Scenario 1: 1,283 kWh
- Heat loss Scenario 2: 19,526 kWh

E.1 Winter exposures

Table E.1: Effect sizes using Cohen's d index for pairwise comparison of $PM_{2.5}$ exposures between regions. d < 0.2 = negligible; 0.2≤ d <0.5 = small; 0.5 ≤ d < 0.8 = medium; 0.8 ≤ d < 1.3 = large; d ≥ 1.3 = very large.

	SCENARIO 1		SCENARIO 2	
Regions	***d***	**Level**	***d***	**Level**
1 vs. 2	-0.01	Negligible	0.02	Negligible
1 vs. 3	-0.01	Negligible	-0.04	Negligible
1 vs. 4	0.00	Negligible	0.01	Negligible
1 vs. 5	-0.01	Negligible	-0.22	Small
1 vs. 6	-0.05	Negligible	-0.21	Small
1 vs. 7	-0.04	Negligible	-0.13	Negligible

Table E.1: Effect sizes. (continued)

1	vs.	8	-0.03	Negligible	-0.07	Negligible
1	vs.	9	-0.02	Negligible	-0.07	Negligible
1	vs.	10	0.00	Negligible	-0.03	Negligible
1	vs.	11	0.03	Negligible	-0.02	Negligible
1	vs.	12	0.08	Negligible	0.04	Negligible
1	vs.	13	-0.04	Negligible	-0.29	Small
1	vs.	14	0.02	Negligible	-0.06	Negligible
1	vs.	15	-0.01	Negligible	-0.01	Negligible
2	vs.	3	0.00	Negligible	-0.06	Negligible
2	vs.	4	0.01	Negligible	-0.01	Negligible
2	vs.	5	0.00	Negligible	-0.24	Small
2	vs.	6	-0.04	Negligible	-0.23	Small
2	vs.	7	-0.03	Negligible	-0.15	Negligible
2	vs.	8	-0.03	Negligible	-0.09	Negligible
2	vs.	9	-0.01	Negligible	-0.10	Negligible
2	vs.	10	0.00	Negligible	-0.05	Negligible
2	vs.	11	0.04	Negligible	-0.04	Negligible
2	vs.	12	0.09	Negligible	0.02	Negligible
2	vs.	13	-0.04	Negligible	-0.31	Small
2	vs.	14	0.02	Negligible	-0.08	Negligible
2	vs.	15	-0.01	Negligible	-0.04	Negligible
3	vs.	4	0.01	Negligible	0.05	Negligible
3	vs.	5	0.00	Negligible	-0.18	Negligible
3	vs.	6	-0.04	Negligible	-0.17	Negligible
3	vs.	7	-0.03	Negligible	-0.09	Negligible
3	vs.	8	-0.03	Negligible	-0.03	Negligible
3	vs.	9	-0.01	Negligible	-0.04	Negligible
3	vs.	10	0.00	Negligible	0.01	Negligible
3	vs.	11	0.04	Negligible	0.02	Negligible
3	vs.	12	0.09	Negligible	0.08	Negligible
3	vs.	13	-0.04	Negligible	-0.25	Small
3	vs.	14	0.02	Negligible	-0.02	Negligible
3	vs.	15	0.00	Negligible	0.02	Negligible
4	vs.	5	-0.01	Negligible	-0.23	Small
4	vs.	6	-0.05	Negligible	-0.22	Small
4	vs.	7	-0.05	Negligible	-0.14	Negligible
4	vs.	8	-0.04	Negligible	-0.08	Negligible
4	vs.	9	-0.03	Negligible	-0.09	Negligible
4	vs.	10	-0.01	Negligible	-0.04	Negligible
4	vs.	11	0.03	Negligible	-0.03	Negligible
4	vs.	12	0.08	Negligible	0.03	Negligible
4	vs.	13	-0.05	Negligible	-0.30	Small
4	vs.	14	0.01	Negligible	-0.07	Negligible
4	vs.	15	-0.02	Negligible	-0.03	Negligible
5	vs.	6	-0.04	Negligible	0.01	Negligible
5	vs.	7	-0.03	Negligible	0.09	Negligible
5	vs.	8	-0.03	Negligible	0.15	Negligible
5	vs.	9	-0.01	Negligible	0.14	Negligible
5	vs.	10	0.00	Negligible	0.19	Negligible
5	vs.	11	0.04	Negligible	0.20	Negligible
5	vs.	12	0.09	Negligible	0.26	Small
5	vs.	13	-0.04	Negligible	-0.07	Negligible
5	vs.	14	0.02	Negligible	0.16	Negligible
5	vs.	15	-0.01	Negligible	0.21	Small
6	vs.	7	0.01	Negligible	0.08	Negligible
6	vs.	8	0.01	Negligible	0.14	Negligible
6	vs.	9	0.03	Negligible	0.13	Negligible
6	vs.	10	0.04	Negligible	0.18	Negligible
6	vs.	11	0.08	Negligible	0.19	Negligible
6	vs.	12	0.13	Negligible	0.25	Small
6	vs.	13	0.00	Negligible	-0.08	Negligible
6	vs.	14	0.06	Negligible	0.15	Negligible
6	vs.	15	0.04	Negligible	0.19	Negligible
7	vs.	8	0.01	Negligible	0.06	Negligible
7	vs.	9	0.02	Negligible	0.05	Negligible
7	vs.	10	0.04	Negligible	0.10	Negligible
7	vs.	11	0.07	Negligible	0.11	Negligible
7	vs.	12	0.12	Negligible	0.17	Negligible
7	vs.	13	0.00	Negligible	-0.16	Negligible
7	vs.	14	0.06	Negligible	0.07	Negligible
7	vs.	15	0.03	Negligible	0.11	Negligible
8	vs.	9	0.01	Negligible	-0.01	Negligible
8	vs.	10	0.03	Negligible	0.04	Negligible

Table E.1: Effect sizes. (continued)

8	vs.	11	0.07	Negligible	0.05	Negligible
8	vs.	12	0.12	Negligible	0.11	Negligible
8	vs.	13	-0.01	Negligible	-0.22	Small
8	vs.	14	0.05	Negligible	0.01	Negligible
8	vs.	15	0.02	Negligible	0.05	Negligible
9	vs.	10	0.02	Negligible	0.05	Negligible
9	vs.	11	0.05	Negligible	0.06	Negligible
9	vs.	12	0.10	Negligible	0.11	Negligible
9	vs.	13	-0.02	Negligible	-0.21	Small
9	vs.	14	0.04	Negligible	0.02	Negligible
9	vs.	15	0.01	Negligible	0.06	Negligible
10	vs.	11	0.04	Negligible	0.01	Negligible
10	vs.	12	0.09	Negligible	0.07	Negligible
10	vs.	13	-0.04	Negligible	-0.26	Small
10	vs.	14	0.02	Negligible	-0.03	Negligible
10	vs.	15	-0.01	Negligible	0.01	Negligible
11	vs.	12	0.05	Negligible	0.06	Negligible
11	vs.	13	-0.07	Negligible	-0.27	Small
11	vs.	14	-0.01	Negligible	-0.04	Negligible
11	vs.	15	-0.04	Negligible	0.01	Negligible
12	vs.	13	-0.12	Negligible	-0.33	Small
12	vs.	14	-0.06	Negligible	-0.10	Negligible
12	vs.	15	-0.09	Negligible	-0.05	Negligible
13	vs.	14	0.06	Negligible	0.23	Small
13	vs.	15	0.03	Negligible	0.28	Small
14	vs.	15	-0.03	Negligible	0.04	Negligible

Table E.2: Effect sizes using Cohen's d index for pairwise comparison of $PM_{2.5}$ exposures between archetypes. $d < 0.2$ = negligible; $0.2 \leq d < 0.5$ = small; $0.5 \leq d < 0.8$ = medium; $0.8 \leq d < 1.3$ = large; $d \geq 1.3$ = very large.

			SCENARIO 1		SCENARIO 2	
Archetypes			***d***	**Level**	***d***	**Level**
19	vs.	27	0.01	Negligible	0.07	Negligible
19	vs.	35	0.09	Negligible	0.07	Negligible
19	vs.	36	0.10	Negligible	0.07	Negligible
19	vs.	91	-0.03	Negligible	0.06	Negligible
19	vs.	100	-0.01	Negligible	0.05	Negligible
19	vs.	275	-0.17	Negligible	0.04	Negligible
19	vs.	284	-0.06	Negligible	0.04	Negligible
27	vs.	35	0.08	Negligible	0.00	Negligible
27	vs.	36	0.09	Negligible	0.01	Negligible
27	vs.	91	-0.05	Negligible	-0.01	Negligible
27	vs.	100	-0.02	Negligible	-0.01	Negligible
27	vs.	275	-0.19	Negligible	-0.02	Negligible
27	vs.	284	-0.07	Negligible	-0.03	Negligible
35	vs.	36	0.01	Negligible	0.00	Negligible
35	vs.	91	-0.12	Negligible	-0.01	Negligible
35	vs.	100	-0.10	Negligible	-0.02	Negligible
35	vs.	275	-0.26	Small	-0.03	Negligible
35	vs.	284	-0.15	Negligible	-0.03	Negligible
36	vs.	91	-0.13	Negligible	-0.02	Negligible
36	vs.	100	-0.11	Negligible	-0.02	Negligible
36	vs.	275	-0.27	Small	-0.03	Negligible
36	vs.	284	-0.16	Negligible	-0.04	Negligible
91	vs.	100	0.02	Negligible	0.00	Negligible
91	vs.	275	-0.14	Negligible	-0.01	Negligible
91	vs.	284	-0.03	Negligible	-0.02	Negligible
100	vs.	275	-0.16	Negligible	-0.01	Negligible
100	vs.	284	-0.05	Negligible	-0.01	Negligible
275	vs.	284	0.11	Negligible	0.00	Negligible

E.2 Ventilation rates

Table E.3: Effect sizes using Cohen's d index for pairwise comparison of median ventilation rates between regions. $d < 0.2$ = negligible; $0.2 \leq d < 0.5$ = small; $0.5 \leq d < 0.8$ = medium; $0.8 \leq d < 1.3$ = large; $d \geq 1.3$ = very large.

	SCENARIO 1		SCENARIO 2	
Regions	***d***	**Level**	***d***	**Level**
1 vs. 2	-0.12	Negligible	-0.07	Negligible
1 vs. 3	-0.31	Small	0.38	Small
1 vs. 4	-0.14	Negligible	0.04	Negligible
1 vs. 5	-0.27	Small	0.33	Small
1 vs. 6	-0.48	Small	0.35	Small
1 vs. 7	-0.27	Small	0.15	Negligible
1 vs. 8	-0.27	Small	-0.09	Negligible
1 vs. 9	-0.32	Small	-0.12	Negligible
1 vs. 10	-0.33	Small	-0.51	Medium
1 vs. 11	-0.69	Medium	-0.42	Small
1 vs. 12	-0.49	Small	-0.83	Large
1 vs. 13	-0.34	Small	0.38	Small
1 vs. 14	-0.59	Medium	-0.23	Small
1 vs. 15	0.06	Negligible	0.15	Negligible
2 vs. 3	-0.19	Negligible	0.44	Small
2 vs. 4	-0.02	Negligible	0.11	Negligible
2 vs. 5	-0.15	Negligible	0.40	Small
2 vs. 6	-0.37	Small	0.42	Small
2 vs. 7	-0.15	Negligible	0.22	Small
2 vs. 8	-0.15	Negligible	-0.02	Negligible
2 vs. 9	-0.21	Small	-0.06	Negligible
2 vs. 10	-0.21	Small	-0.44	Small
2 vs. 11	-0.57	Medium	-0.35	Small
2 vs. 12	-0.37	Small	-0.76	Medium
2 vs. 13	-0.22	Small	0.45	Small
2 vs. 14	-0.48	Small	-0.16	Negligible
2 vs. 15	0.18	Negligible	0.22	Small
3 vs. 4	0.17	Negligible	-0.33	Small
3 vs. 5	0.04	Negligible	-0.04	Negligible
3 vs. 6	-0.17	Negligible	-0.02	Negligible
3 vs. 7	0.04	Negligible	-0.22	Small
3 vs. 8	0.04	Negligible	-0.46	Small
3 vs. 9	-0.02	Negligible	-0.50	Small
3 vs. 10	-0.02	Negligible	-0.88	Large
3 vs. 11	-0.38	Small	-0.79	Medium
3 vs. 12	-0.18	Negligible	-1.20	Large
3 vs. 13	-0.03	Negligible	0.01	Negligible
3 vs. 14	-0.29	Small	-0.60	Medium
3 vs. 15	0.37	Small	-0.22	Small
4 vs. 5	-0.13	Negligible	0.29	Small
4 vs. 6	-0.35	Small	0.31	Small
4 vs. 7	-0.13	Negligible	0.11	Negligible
4 vs. 8	-0.13	Negligible	-0.13	Negligible
4 vs. 9	-0.19	Negligible	-0.17	Negligible
4 vs. 10	-0.19	Negligible	-0.55	Medium
4 vs. 11	-0.55	Medium	-0.46	Small
4 vs. 12	-0.35	Small	-0.87	Large
4 vs. 13	-0.20	Small	0.34	Small
4 vs. 14	-0.46	Small	-0.27	Small
4 vs. 15	0.20	Small	0.11	Negligible
5 vs. 6	-0.22	Small	0.02	Negligible
5 vs. 7	0.00	Negligible	-0.18	Negligible
5 vs. 8	0.00	Negligible	-0.42	Small
5 vs. 9	-0.06	Negligible	-0.45	Small
5 vs. 10	-0.06	Negligible	-0.84	Large
5 vs. 11	-0.42	Small	-0.75	Medium
5 vs. 12	-0.22	Small	-1.16	Large
5 vs. 13	-0.07	Negligible	0.05	Negligible
5 vs. 14	-0.33	Small	-0.56	Medium
5 vs. 15	0.33	Small	-0.18	Negligible
6 vs. 7	0.22	Small	-0.20	Negligible
6 vs. 8	0.22	Small	-0.44	Small
6 vs. 9	0.16	Negligible	-0.47	Small
6 vs. 10	0.15	Negligible	-0.86	Large
6 vs. 11	-0.21	Small	-0.77	Medium
6 vs. 12	0.00	Negligible	-1.18	Large

Table E.3: Effect sizes. (continued)

6	vs.	13	0.15	Negligible	0.03	Negligible
6	vs.	14	-0.11	Negligible	-0.58	Medium
6	vs.	15	0.55	Medium	-0.20	Negligible
7	vs.	8	0.00	Negligible	-0.24	Small
7	vs.	9	-0.06	Negligible	-0.28	Small
7	vs.	10	-0.06	Negligible	-0.66	Medium
7	vs.	11	-0.42	Small	-0.57	Medium
7	vs.	12	-0.22	Small	-0.98	Large
7	vs.	13	-0.07	Negligible	0.23	Small
7	vs.	14	-0.33	Small	-0.38	Small
7	vs.	15	0.33	Small	0.00	Negligible
8	vs.	9	-0.06	Negligible	-0.03	Negligible
8	vs.	10	-0.06	Negligible	-0.42	Small
8	vs.	11	-0.42	Small	-0.33	Small
8	vs.	12	-0.22	Small	-0.74	Medium
8	vs.	13	-0.07	Negligible	0.47	Small
8	vs.	14	-0.33	Small	-0.14	Negligible
8	vs.	15	0.33	Small	0.24	Small
9	vs.	10	0.00	Negligible	-0.38	Small
9	vs.	11	-0.36	Small	-0.29	Small
9	vs.	12	-0.16	Negligible	-0.70	Medium
9	vs.	13	-0.01	Negligible	0.51	Medium
9	vs.	14	-0.27	Small	-0.10	Negligible
9	vs.	15	0.39	Small	0.27	Small
10	vs.	11	-0.36	Small	0.09	Negligible
10	vs.	12	-0.16	Negligible	-0.32	Small
10	vs.	13	-0.01	Negligible	0.89	Large
10	vs.	14	-0.27	Small	0.28	Small
10	vs.	15	0.39	Small	0.66	Medium
11	vs.	12	0.20	Small	-0.41	Small
11	vs.	13	0.35	Small	0.80	Large
11	vs.	14	0.09	Negligible	0.19	Negligible
11	vs.	15	0.75	Medium	0.57	Medium
12	vs.	13	0.15	Negligible	1.21	Large
12	vs.	14	-0.11	Negligible	0.60	Medium
12	vs.	15	0.55	Medium	0.98	Large
13	vs.	14	-0.26	Small	-0.61	Medium
13	vs.	15	0.40	Small	-0.23	Small
14	vs.	15	0.66	Medium	0.38	Small

Table E.4: Effect sizes using Cohen's d index for pairwise comparison of median ventilation rates between archetypes. $d < 0.2$ = negligible; $0.2 \leq d < 0.5$ = small; $0.5 \leq d < 0.8$ = medium; $0.8 \leq d < 1.3$ = large; $d \geq 1.3$ = very large.

			SCENARIO 1		**SCENARIO 2**	
Archetypes			***d***	**Level**	***d***	**Level**
19	vs.	27	-0.11	Negligible	-0.01	Negligible
19	vs.	35	-0.37	Small	-0.02	Negligible
19	vs.	36	-0.29	Small	-0.01	Negligible
19	vs.	91	0.10	Negligible	-0.01	Negligible
19	vs.	100	-0.27	Small	-0.03	Negligible
19	vs.	275	0.66	Medium	-0.01	Negligible
19	vs.	284	0.57	Medium	-0.01	Negligible
27	vs.	35	-0.26	Small	-0.01	Negligible
27	vs.	36	-0.18	Negligible	0.00	Negligible
27	vs.	91	0.21	Small	-0.01	Negligible
27	vs.	100	-0.15	Negligible	-0.02	Negligible
27	vs.	275	0.77	Medium	-0.01	Negligible
27	vs.	284	0.68	Medium	-0.01	Negligible
35	vs.	36	0.08	Negligible	0.01	Negligible
35	vs.	91	0.47	Small	0.00	Negligible
35	vs.	100	0.10	Negligible	-0.01	Negligible
35	vs.	275	1.03	Large	0.00	Negligible
35	vs.	284	0.94	Large	0.00	Negligible
36	vs.	91	0.40	Small	-0.01	Negligible
36	vs.	100	0.03	Negligible	-0.02	Negligible
36	vs.	275	0.95	Large	-0.01	Negligible
36	vs.	284	0.86	Large	-0.01	Negligible
91	vs.	100	-0.37	Small	-0.01	Negligible

Table E.3: Effect sizes. (continued)

91	vs.	275	0.56	Medium	0.00	Negligible
91	vs.	284	0.46	Small	0.00	Negligible
100	vs.	275	0.92	Large	0.01	Negligible
100	vs.	284	0.83	Large	0.01	Negligible
275	vs.	284	-0.09	Negligible	0.00	Negligible

E.3 Heat loss

Table E.5: Effect sizes using Cohen's d index for pairwise comparison of total heat loss between regions. $d < 0.2$ = negligible; $0.2 \leq d < 0.5$ = small; $0.5 \leq d < 0.8$ = medium; $0.8 \leq d < 1.3$ = large; $d \geq 1.3$ = very large.

			SCENARIO 1		SCENARIO 2	
Regions			***d***	**Level**	***d***	**Level**
1	vs.	2	-0.03	Negligible	0.09	Negligible
1	vs.	3	-0.09	Negligible	0.30	Small
1	vs.	4	-0.07	Negligible	0.11	Negligible
1	vs.	5	-0.17	Negligible	0.34	Small
1	vs.	6	-0.38	Small	0.32	Small
1	vs.	7	-0.20	Small	0.03	Negligible
1	vs.	8	-0.21	Small	-0.25	Small
1	vs.	9	-0.24	Small	-0.04	Negligible
1	vs.	10	-0.20	Negligible	-0.46	Small
1	vs.	11	-0.59	Medium	-0.45	Small
1	vs.	12	-0.31	Small	-0.84	Large
1	vs.	13	-0.24	Small	0.37	Small
1	vs.	14	-0.40	Small	-0.12	Negligible
1	vs.	15	0.07	Negligible	0.15	Negligible
2	vs.	3	-0.06	Negligible	0.21	Small
2	vs.	4	-0.04	Negligible	0.02	Negligible
2	vs.	5	-0.14	Negligible	0.25	Small
2	vs.	6	-0.35	Small	0.23	Small
2	vs.	7	-0.18	Negligible	-0.06	Negligible
2	vs.	8	-0.18	Negligible	-0.34	Small
2	vs.	9	-0.22	Small	-0.13	Negligible
2	vs.	10	-0.17	Negligible	-0.55	Medium
2	vs.	11	-0.56	Medium	-0.54	Medium
2	vs.	12	-0.28	Small	-0.93	Large
2	vs.	13	-0.21	Small	0.28	Small
2	vs.	14	-0.37	Small	-0.20	Small
2	vs.	15	0.10	Negligible	0.06	Negligible
3	vs.	4	0.02	Negligible	-0.19	Negligible
3	vs.	5	-0.08	Negligible	0.04	Negligible
3	vs.	6	-0.29	Small	0.02	Negligible
3	vs.	7	-0.11	Negligible	-0.27	Small
3	vs.	8	-0.12	Negligible	-0.55	Medium
3	vs.	9	-0.15	Negligible	-0.34	Small
3	vs.	10	-0.10	Negligible	-0.76	Medium
3	vs.	11	-0.49	Small	-0.75	Medium
3	vs.	12	-0.21	Small	-1.13	Large
3	vs.	13	-0.15	Negligible	0.07	Negligible
3	vs.	14	-0.31	Small	-0.41	Small
3	vs.	15	0.16	Negligible	-0.15	Negligible
4	vs.	5	-0.10	Negligible	0.23	Small
4	vs.	6	-0.31	Small	0.21	Small
4	vs.	7	-0.13	Negligible	-0.08	Negligible
4	vs.	8	-0.14	Negligible	-0.36	Small
4	vs.	9	-0.17	Negligible	-0.16	Negligible
4	vs.	10	-0.12	Negligible	-0.57	Medium
4	vs.	11	-0.51	Medium	-0.56	Medium
4	vs.	12	-0.23	Small	-0.95	Large
4	vs.	13	-0.17	Negligible	0.26	Small
4	vs.	14	-0.33	Small	-0.23	Small
4	vs.	15	0.14	Negligible	0.04	Negligible
5	vs.	6	-0.21	Small	-0.02	Negligible
5	vs.	7	-0.04	Negligible	-0.31	Small
5	vs.	8	-0.04	Negligible	-0.59	Medium

Table E.5: Effect sizes. (continued)

5	vs.	9	-0.08	Negligible	-0.38	Small
5	vs.	10	-0.03	Negligible	-0.80	Large
5	vs.	11	-0.42	Small	-0.79	Medium
5	vs.	12	-0.14	Negligible	-1.18	Large
5	vs.	13	-0.07	Negligible	0.03	Negligible
5	vs.	14	-0.23	Small	-0.45	Small
5	vs.	15	0.24	Small	-0.19	Negligible
6	vs.	7	0.17	Negligible	-0.29	Small
6	vs.	8	0.17	Negligible	-0.57	Medium
6	vs.	9	0.13	Negligible	-0.36	Small
6	vs.	10	0.18	Negligible	-0.78	Medium
6	vs.	11	-0.21	Small	-0.77	Medium
6	vs.	12	0.07	Negligible	-1.15	Large
6	vs.	13	0.14	Negligible	0.05	Negligible
6	vs.	14	-0.02	Negligible	-0.43	Small
6	vs.	15	0.45	Small	-0.17	Negligible
7	vs.	8	-0.01	Negligible	-0.28	Small
7	vs.	9	-0.04	Negligible	-0.08	Negligible
7	vs.	10	0.01	Negligible	-0.49	Small
7	vs.	11	-0.38	Small	-0.48	Small
7	vs.	12	-0.10	Negligible	-0.87	Large
7	vs.	13	-0.03	Negligible	0.34	Small
7	vs.	14	-0.19	Negligible	-0.15	Negligible
7	vs.	15	0.27	Small	0.12	Negligible
8	vs.	9	-0.03	Negligible	0.21	Small
8	vs.	10	0.01	Negligible	-0.21	Small
8	vs.	11	-0.38	Small	-0.20	Small
8	vs.	12	-0.10	Negligible	-0.59	Medium
8	vs.	13	-0.03	Negligible	0.62	Medium
8	vs.	14	-0.19	Negligible	0.13	Negligible
8	vs.	15	0.28	Small	0.40	Small
9	vs.	10	0.05	Negligible	-0.42	Small
9	vs.	11	-0.34	Small	-0.41	Small
9	vs.	12	-0.06	Negligible	-0.79	Medium
9	vs.	13	0.01	Negligible	0.41	Small
9	vs.	14	-0.15	Negligible	-0.07	Negligible
9	vs.	15	0.31	Small	0.19	Negligible
10	vs.	11	-0.39	Small	0.01	Negligible
10	vs.	12	-0.11	Negligible	-0.38	Small
10	vs.	13	-0.04	Negligible	0.83	Large
10	vs.	14	-0.20	Small	0.35	Small
10	vs.	15	0.27	Small	0.61	Medium
11	vs.	12	0.28	Small	-0.38	Small
11	vs.	13	0.35	Small	0.82	Large
11	vs.	14	0.19	Negligible	0.34	Small
11	vs.	15	0.66	Medium	0.60	Medium
12	vs.	13	0.07	Negligible	1.21	Large
12	vs.	14	-0.09	Negligible	0.72	Medium
12	vs.	15	0.38	Small	0.99	Large
13	vs.	14	-0.16	Negligible	-0.49	Small
13	vs.	15	0.31	Small	-0.22	Small
14	vs.	15	0.47	Small	0.27	Small

Table E.6: Effect sizes using Cohen's d index for pairwise comparison of total heat loss between archetypes. d < 0.2 = negligible; 0.2$\leq$ d <0.5 = small; 0.5 $\leq$ d < 0.8 = medium; 0.8 $\leq$ d < 1.3 = large; $d \geq$ 1.3 = very large.

			SCENARIO 1		SCENARIO 2	
Archetypes			***d***	**Level**	***d***	**Level**
19	vs.	27	0.17	Negligible	0.71	Medium
19	vs.	35	-0.37	Small	0.55	Medium
19	vs.	36	-0.27	Small	0.68	Medium
19	vs.	91	0.52	Medium	1.14	Large
19	vs.	100	-0.16	Negligible	0.34	Small
19	vs.	275	0.62	Medium	0.76	Medium
19	vs.	284	0.35	Small	0.28	Small
27	vs.	35	-0.54	Medium	-0.16	Negligible
27	vs.	36	-0.43	Small	-0.03	Negligible
27	vs.	91	0.36	Small	0.43	Small

Table E.6: Effect sizes. (continued)

27	vs.	100	-0.32	Small	-0.37	Small
27	vs.	275	0.46	Small	0.04	Negligible
27	vs.	284	0.19	Negligible	-0.44	Small
35	vs.	36	0.10	Negligible	0.13	Negligible
35	vs.	91	0.89	Large	0.59	Medium
35	vs.	100	0.21	Small	-0.21	Small
35	vs.	275	0.99	Large	0.21	Small
35	vs.	284	0.72	Medium	-0.27	Small
36	vs.	91	0.79	Medium	0.46	Small
36	vs.	100	0.11	Negligible	-0.34	Small
36	vs.	275	0.89	Large	0.07	Negligible
36	vs.	284	0.62	Medium	-0.41	Small
91	vs.	100	-0.68	Medium	-0.80	Large
91	vs.	275	0.10	Negligible	-0.39	Small
91	vs.	284	-0.17	Negligible	-0.87	Large
100	vs.	275	0.78	Medium	0.42	Small
100	vs.	284	0.51	Medium	-0.06	Negligible
275	vs.	284	-0.27	Small	-0.48	Small