

ACCEPTED MANUSCRIPT • OPEN ACCESS

A two-stage clustering approach to investigate lifestyle carbon footprints in two Australian cities

To cite this article before publication: Andreas Froemelt *et al* 2020 *Environ. Res. Lett.* in press <https://doi.org/10.1088/1748-9326/abb502>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2020 The Author(s). Published by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

A two-stage clustering approach to investigate lifestyle carbon footprints in two Australian cities

Andreas Froemelt^{1*} and Thomas Wiedmann¹

¹ Sustainability Assessment Program, School of Civil and Environmental Engineering, UNSW Sydney, NSW, Australia
* Corresponding Author; current address: Ecological Systems Design, Institute of Environmental Engineering, ETH Zurich, Switzerland; E-mail: froemelt@ifu.baug.ethz.ch

ORCID Andreas Froemelt: 0000-0001-9388-7816
ORCID Thomas Wiedmann: 0000-0002-6395-8887

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

Given the key role of households in driving global emissions and resource uses, a change in their consumption behaviours towards more sustainable levels is essential to reduce world-wide adverse environmental impacts. Thereby, focusing on cities is especially important because of today's large share of the global population living in cities and because local authorities are close to the needs of their residents. However, devising targeted and effective policy measures implies a thorough understanding of prevailing consumption patterns and associated environmental consequences. The goal of this article is to investigate and compare household behaviours and lifestyle-induced carbon footprints in Sydney and Melbourne in order to enhance today's understanding of household consumption in cities of a high-income, high-emission country. For this purpose, we employed a two-stage clustering with a Self-Organising Map and a subsequent Ward-clustering. This allowed for including expenditure data along with socio-economic attributes and thus for recognising lifestyle-archetypes. These emerging archetypes represent households with similar characteristics and comparable consumption patterns. Analysing the archetypes in detail and performing a city-comparison based on multi-dimensional scaling revealed similarities and dissimilarities between the two metropolises. "Older" archetypes seem to behave more alike across cities but show different carbon footprints emphasising the importance of regionalised environmental assessments and of city-specific supply chains. Distinct patterns especially emerged in the high- and low-income segments highlighting the different importance of different lifestyles in each city. Socio-economically similar family-archetypes were found in both cities, but some of them showed diverging consumption behaviours. This article showed that studying household-induced environmental impacts in cities should not rely on macro-trends but should rather be based on city-specific analyses that capture local peculiarities and consider socio-economic characteristics and consumption data simultaneously.

Keywords: household consumption, carbon footprint, data mining, input-output life cycle assessment, urban households

1. Introduction

Given the fact that more than half of the World's population lives in urban areas and facing ongoing rapid global urbanisation [1], cities have been identified as key actors to curb adverse environmental impacts (e.g., [2,3,12–14,4–11]). Local governments are well-positioned to establish and support more sustainable production systems within their municipal boundaries, but they are also especially qualified to abate negative consequences of household consumption behaviours [2,9–13,15–18]. Households are main drivers of the economy and are thus responsible for a high share of worldwide emissions and resource use along global supply chains [19]. A large body of literature (e.g. [2,9,25–29,10,12,17,20–24]) emphasises the inevitability of changing lifestyles towards more sustainable consumption patterns in order to keep anthropogenic impacts within the carrying capacity of the Earth [30]. This will consequently mitigate risks for humans and environment and hence help to achieve political targets such as the United Nations Sustainable Development Goals [31], the Paris Agreement [32], or limiting global warming to 1.5°C [11,33]. City authorities can provide financial incentives, introduce regulatory codes, offer best-practice guidance, or "lock in" households by stipulating renewable energy supply [2,9,11,13,16,17]. Being close to the needs of their residents, they are especially fit to devise "soft" policy instruments such as awareness raising campaigns or personalized messages [2,10–13,16,17,34,35]. However, the development of such targeted measures that are tailored to the actual city inhabitants require a deep understanding of the variability and peculiarities of local consumption patterns [2,5,27,36–38,6–10,13,18,20].

A novel approach for investigating household behaviours in a certain region and thus for delivering data in support of environmental policymaking was developed by Froemelt and colleagues [39]. Using data mining techniques these authors exploited the Swiss household budget survey and were able to group households according to their socio-economic characteristics and consumption behaviours. Previous household consumption studies (e.g., [25,27,47,48,38,40–46]) have also been able to provide insightful and highly interesting findings, but they are often based on pre-defined socio-economic household segments. However, several articles point out that there is still a large variability of behaviour present within socio-economic cohorts [9,25,40–42]. Froemelt et al.'s [39] approach attempts to advance current methodologies in this regard and thus also to respond to recent calls for a better exploration of lifestyles and the inclusion of behavioural anomalies [9,10,20,34]. By letting consumption archetypes emerge from consumer expenditure data, the proposed two-stage clustering has the potential to recognise important behavioural patterns and thus to efficiently focus on a set of household groups that represent distinct behaviours and socio-economic circumstances occurring in a certain city. Furthermore, Froemelt et al.'s approach [39] has also been successfully integrated in the context of large-scale models [49].

However, despite the importance of cities, Froemelt et al.'s methodology [39] has never been applied in a city-context. Moreover, to understand household consumption behaviour more broadly, it is essential to apply this methodology also in an international context and not only to Swiss data.

The two Australian metropolises Sydney and Melbourne provide interesting case studies for these purposes [5,14,50–52]. Both cities show commitment to reduce their carbon footprints to some extent [5,50]. However, together, they are currently contributing about 30% to the total carbon footprint of Australia [5], which, from a per-capita perspective, is among the highest in the world [19,53].

The goal of this article is thus to investigate consumption patterns in Sydney and Melbourne in order to learn, understand, and compare different lifestyles in two cities of a high-income, high-emission country and thus to improve today's knowledge of household consumption behaviours.

By way of analysing two cities simultaneously, we also aim at providing answers to important questions that might further our understanding of household consumption; e.g., Do similar household segments behave similarly across the cities or do we find city-specific behaviour? These questions are particularly interesting in the context of Melbourne and Sydney since both cities show – except for being located in somewhat different climate zones – many similarities at first glance: same nation, similar population size, high economic importance within the country [5], and both rank among the top most liveable cities in the world [54–56].

2. Methods

In order to analyse different lifestyles and to assess their induced environmental consequences in Sydney and Melbourne, we proceeded in a similar way as suggested by Froemelt et al. [39]. Their two-stage clustering approach revealed interesting household consumption patterns for Switzerland and can be considered a well-suited procedure for the current study. Furthermore, these kinds of pattern recognition techniques also prove their usefulness in other research fields and perform especially well and robust for high-dimensional and noisy data [57–60].

A synopsis of our applied methodology is provided as a flow scheme in Figure 1 and is detailed in the following sub-sections. The Supplementary Information (SI) holds more information to enhance the reproducibility of our study.

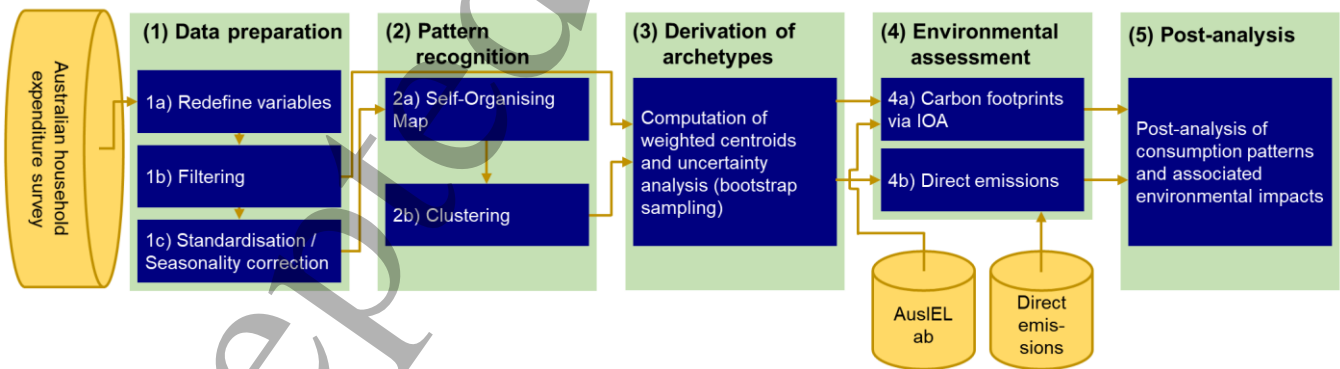


Figure 1: Simplified flow scheme of the applied methodology. [IOA = input-output analysis; AusIELab is the implementation of IELab for Australia [61]]

2.1 Data description and preparation

The Australian household expenditure survey (HES) conducted by the Australian Bureau of Statistics (ABS) in the 2015-16 cycle represents the main dataset for our study [62,63]. In the HES, detailed data on expenditures, household and personal characteristics, sources of income, as well as net wealth and possession of properties of 10,046 individual households is collected. The full dataset comprises about 400 attributes (mainly characteristics, income categories and possession) for each household and each household member, respectively, and subdivides expenditure into 709 consumption categories at the most detailed level of the Australian household expenditure classification system (HEC) [62,63]. Note that data accessible for research is partially confidentialised [63]. In our case, this especially affected the geographical resolution of the data and set the city boundaries of Sydney and Melbourne according to the "Greater Capital City Statistical Area" (GCCSA) defined by ABS [64].

In a first step, the HES microdata needed to undergo basic pre-processing to enable the subsequent application of data mining techniques. This included the aggregation of person-specific data at the household level and the preparation of categorical variables of household characteristics.

As outlined in [39], the clustering of households is only meaningful if it is based on attributes that enable the detection of similar characteristics and behaviour. Therefore, the HES dataset was filtered for such variables and various highly detailed expenditure categories were aggregated to broader consumption areas. Subsequently, all 108 variables destined for the clustering process were put on an even footing by standardising and by correcting for seasonality in cases where the survey time period was found to be statistically significant by one-way ANOVA-tests [65].

Besides a more detailed description of this preparatory step, the SI provides also a tabular overview of all HES categories and the variables used in the next steps.

2.2 Two-stage pattern recognition

In the first stage of this two-tiered approach, a Self-Organising Map (SOM) is trained [66,67]. SOMs belong to the class of unsupervised artificial neural networks and are well-suited to enable the efficient subsequent application of clustering techniques by preconditioning noisy and high-dimensional datasets [58–60]. The SOM usually consists of a 2D-lattice whose nodes represent the neurons. Attached to each neuron is a prototype vector with the same dimension as the original dataset. During the training phase, the data points of the original dataset are exposed to the SOM. The nearest neuron as well as adjacent neurons are activated and move closer to the data point under consideration. The underlying learning procedure allows for preserving the data topology (neighbouring neurons have similar prototype vectors) and for projecting the original dataset to a 2D-map. After several repetitions of the learning process, the prototype vectors approximate an optimised set of substitutes that are representative of the original data points. The resulting vectors are not only reduced in numbers compared to the original data points, but are also smoothed with regard to noise and enable thus an enhanced pattern recognition for clustering algorithms and even for the human eye [66–68].

After extracting the households living in Sydney (1,778 households) and Melbourne (2,004 households) from the prepared dataset, we trained a SOM separately for both cities. The training followed literature recommendations

Journal **XX** (XXXX) XXXXXX

Froemelt and Wiedmann

[66–70] and the tuning parameters were selected based on the resulting topographic error and quantization error (see SI for details) [69]. The final map for Sydney comprises 420 neurons in a 14:30-arrangement, while 434 neurons on a 14:31-grid were considered optimal for Melbourne.

For the clustering of neurons in the second step, a hierarchical agglomerative approach [71,72] was deployed on top of the two SOMs. The decision for this type of clustering was taken mainly because of the possibility to retain the connectivity among the neurons [58,59,73] and due to its good performance in other studies [39,57]. However, to find an optimal clustering, we systematically evaluated different numbers of clusters as well as different affinity metrics and linkage criteria. To judge the performance of clustering parameter combinations, we relied on the Silhouette coefficient [74] as well as on the representation of the so-called U-Matrix which visually summarises a SOM by showing the distances between the neurons and which is well-suited to assess different clusterings [59,75,76]. Based on these considerations and including ANOVA-tests [65] to obtain a sense of the statistical importance of the attributes for the clustering, we fixed the clustering parameters (Ward-linkage with Euclidean distances [77] for both cities) and decided for an upper and a lower number of clusters. In a last step, we consulted the so-called Hits-Map and included all clusters that would be formed by the hierarchical clustering between the upper and the lower number of clusters and that do not violate "zero-hits"-boundaries [58].

Among the formed household groups, we excluded clusters with less than 50 households due to concerns of representativeness [78]. The whole pattern recognition procedure described above and in the SI finally found 13 clusters for Sydney and 14 for Melbourne.

2.3 Computation of centroids and uncertainty analysis

For computing a representative vector for the different household clusters, all variables from the HES dataset (also those excluded during the clustering) were averaged according to the recommendations of ABS [62]. These guidelines suggest considering the representativeness-weight given by ABS for each household [63].

As a valuable addition to the approach of [39], we developed an uncertainty analysis framework by introducing a bootstrap-sampling [79]. The latter randomly re-samples the households in each cluster and, by applying above weighted-average procedure, is able to estimate the uncertainty of the cluster's centroid [79]. In the present case, 1000 vectors were computed for each cluster.

2.4 Environmental assessment

For assessing the carbon footprint of the different household behaviours, we set the temporal boundaries to one year of household consumption (the reference year corresponds to the financial year 2015-16 in accordance with the HES dataset [63]).

For computing indirect, upstream life-cycle greenhouse gas (GHG) emissions caused by households, the IELab [61,80,81] was used to build a highly detailed multi-region-supply-use-table for the year 2015 with corresponding GHG-satellites (taken from [82]). Our table distinguishes 10 regions (among them also the GCCSAs of Sydney and Melbourne) and details 797 economic sectors in each region. Together with a m:n-correspondence table established

for [83], the CO₂-eq-intensity-factors for 593 consumption categories and thus for the highest possible HEC-resolution could be computed. Thanks to the regionalisation, different factors apply for Melbourne and Sydney (see SI for details).

The inclusion of a constraint for the HES in the reconciliation process of the table ensured good correspondence of the HES data with the final demand vectors (12.9% and 24.2% deviation of the resulting table for Sydney and Melbourne respectively). Furthermore, reasonable agreement of the resulting carbon footprints with existing studies were found (deviations are between 2.2% and 13.3%; see SI).

Direct household emissions from housing (heating and cooking) and private transport were retrieved from [82,84] for Victoria and New South Wales. The respective state-wide household expenditures [63] for residential energy and transport fuels allowed for estimating intensity factors to compute direct emissions caused by the clusters based on their spending. The total carbon footprint of households is the sum of life-cycle and direct emissions.

2.5 Post-analysis

The SI discloses fully detailed consumption profiles of all archetypes and contains various plots and heatmaps to characterise and study the household consumption behaviours found in Sydney and Melbourne. In addition, a simplified, but intuitive characterisation of these archetypes is provided in Figure 2, and further described in Table 1 and Table 2. The indicative titles for each cluster in these tables are based on a semi-automated analysis that detects expenditure behaviour of individual archetypes that is most distinct from other archetypes in the same city (based on the average distance of the expenditure under consideration from the archetype under consideration to all other archetypes in the respective city; see SI).

Finally, in order to investigate and compare consumption behaviours and carbon footprint compositions in the two cities, metric multi-dimensional scaling (MDS) was applied [73,85]. The employed technique projects the high-dimensional dataset to a 2D-plane and attempts to simultaneously preserve the original distances between the data points in the high-dimensional space. This results in a 2D-plot that reveals which consumption patterns in Sydney and Melbourne are similar within, but also across these two cities. Note that we deliberately choose to apply MDS only to the 20 main consumption categories of the HEC-classification and thus to exclude socio-economic aspects and more detailed spending. This will allow for comparing more generic consumption behaviours across the cities rather than socio-economic differences and city-specific purchases. Likewise, when comparing carbon footprint composition with MDS, we focused on eight main categories.

3. Results and discussion

3.1 Consumption archetypes and environmental impacts in Sydney and Melbourne

Figure 2, Table 1 and Table 2 reveal a certain correlation between household size and income in Sydney, while in Melbourne, a similar trend is observed, but two archetypes (MEL-XIII and MEL-VI) emerge with low income but comparably large size. In general, the archetypes in Melbourne seem to be socio-economically more diverse (from very high to very low income and different combinations of income and household sizes). In Sydney, the very-low-income segment is subdivided into more archetypes suggesting that the behaviours in this income group might be more distinct than in Melbourne.

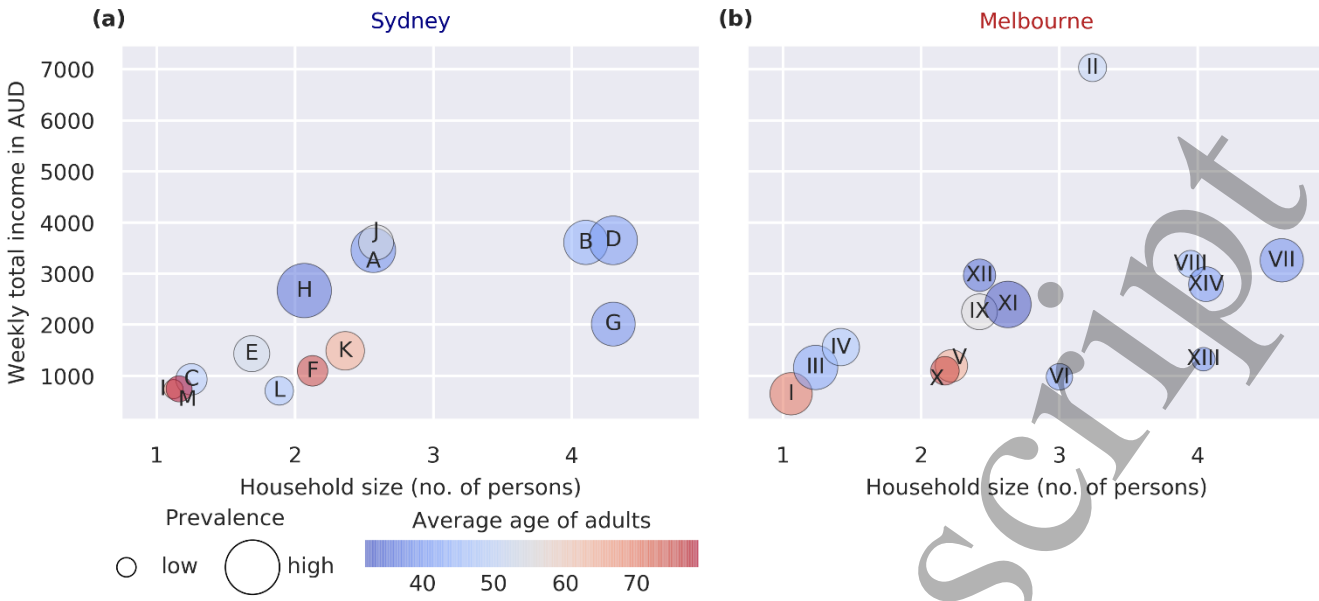


Figure 2: Brief overview of key socio-economic characteristics and prevalence of the clusters in (a) Sydney and (b) Melbourne. Further characterisation of these archetypes can be found in a simplified overview in Table 1 (Sydney) and Table 2 (Melbourne) and in full detail in the SI. The colours of the markers indicate the average age of adults in the archetype (all persons aged older than 18 years), while the size represents the prevalence of the clusters. In this article, prevalence is computed as the sum of the respective cluster's household-representativeness-weights provided by ABS [63]. Since the resulting absolute prevalence-values do not have a clear meaning, we reduced the legend to a relative low/high-comparison.

Table 1: Consumption archetypes found in Sydney. The different archetypes are randomly labelled with letters. The simplified subdivision into household sizes, income classes as well as the summarising titles help to better understand the nature of the archetypes and shall support the visualisation of socio-economic characteristics in Figure 2. Refer to the SI for a detailed characterisation of the archetypes. This table also provides the following information for each archetype: average age of adults (persons aged over 18 years), number of children (aged below 15 years) income in AUD per week; income per capita (p.c.); total carbon footprint (TCF) in tCO₂-eq per year; and total carbon footprint per capita. To aid the comparison, the latter three parameters show a colour scaling from blue (low) to red (high).

	Single/Small households (1 to 1.5 persons)	Two-person households (1.6 to 2.5)	Small families (2.6 to 3.5)	Families (>3.5)
Very high income (> 4000 AUD/wk)				
High income (3000 - 4000 AUD/wk)			SYD-J Wealthy couples before retirement Age of adults 55 No. kids (<15) 0.1 Income 3607 Income p.c. 1395 TCF 49 TCF p.c. 19	SYD-D Families with children > 15 years (in education) Age of adults 40 No. kids (<15) 0.3 Income 3648 Income p.c. 848 TCF 55 TCF p.c. 13
			SYD-A Young to middle-aged couples (some with children) Age of adults 38 No. kids (<15) 0.5 Income 3454 Income p.c. 1346 TCF 43 TCF p.c. 17	SYD-B (Rather wealthy) Families with children at primary/secondary school age Age of adults 43 No. kids (<15) 1.8 Income 3608 Income p.c. 879 TCF 56 TCF p.c. 14
Medium-high income (2500 - 3000 AUD/wk)		SYD-H Rather young couple (some in education) Age of adults 35 No. kids (<15) 0.1 Income 2665 Income p.c. 1289 TCF 31 TCF p.c. 15		
Medium income (2000 - 2500 AUD/wk)				SYD-G Young families with babies/toddlers Age of adults 38 No. kids (<15) 1.9 Income 2010 Income p.c. 467 TCF 33 TCF p.c. 8
Low income (1000 - 2000 AUD/wk)		SYD-K Retired two/three-person HH (males by trend) Age of adults 65 No. kids (<15) 0.0 Income 1486 Income p.c. 629 TCF 33 TCF p.c. 14		
		SYD-E Middle-aged persons (females by trend) Age of adults 53 No. kids (<15) 0.0 Income 1430 Income p.c. 847 TCF 25 TCF p.c. 15		
		SYD-F Retired, very old couple Age of adults 76 No. kids (<15) 0.0 Income 1094 Income p.c. 514 TCF 22 TCF p.c. 10		
Very low income (< 1000 AUD/wk)	SYD-C Middle-aged adults Age of adults 48 No. kids (<15) 0.1 Income 929 Income p.c. 742 TCF 16 TCF p.c. 12	SYD-L Probably poor, unemployed, single-parent households Age of adults 47 No. kids (<15) 0.5 Income 703 Income p.c. 373 TCF 15 TCF p.c. 8		
	SYD-M Very old persons (females by trend) Age of adults 79 No. kids (<15) 0.0 Income 742 Income p.c. 639 TCF 14 TCF p.c. 12			
	SYD-I Just retired, rather wealthy persons (females by trend) Age of adults 66 No. kids (<15) 0.0 Income 736 Income p.c. 656 TCF 17 TCF p.c. 15			

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The per-capita carbon footprint ranking of consumption archetypes in Figure 3b shows an intuitively logical order for Sydney by starting from high-income/small households to two-person/middle-aged households, then families, low-income single persons and low-income families. An interesting exception is cluster SYD-I (just retired, one-person households) with rather high per-capita footprint along with low per-capita income. However, the 95%-confidence interval for this archetype is large and these households have rather high net wealth. SYD-I is also the least prevalent archetype¹ and therefore only shows a minor contribution to the total carbon footprint of Sydney. This is depicted in Figure 3a that relates the archetypes' per-capita footprints to their shares in Sydney's prevalence-weighted total carbon footprint to express the archetypes' importance in the collective. This figure can also help to identify archetypes that might be promising targets to reduce the city's overall footprint (high per-capita footprint/high contribution to overall footprint). The highest shares in Sydney's carbon footprint originate from SYD-D and SYD-B, both high-income large families whose contributions are high due to their sheer abundance and household size (cf. Figure 3a). Yet, the per-capita footprint of these clusters rank medium compared to the smaller, high-income households (e.g. SYD-J and SYD-A) that also contribute significantly to the overall carbon footprint of Sydney.

In Melbourne, the per-capita ranking (Figure 3d) starts with very-high-income households, leads over to high-income/small households, single persons, medium-income/small families to high-income/big families and finally families with low income. The most obvious difference to Sydney pertains to single-person households that rank higher in the per-capita footprints list in Melbourne. This is likely due to their generally higher per-capita income and thus probably higher consumption demand. Interestingly, archetypes with older household members, such as MEL-I, MEL-X, SYD-M and SYD-F, show relatively higher impacts in Melbourne than in Sydney (see also discussion below).

As a side note, we would like to pick the per-capita carbon footprints of MEL-X (old, retired couple), MEL-XI (young adults) and MEL-VII (large families) that have different compositions but add up to about the same total. This shows that households in different circumstances and with different consumption profiles might induce a similar amount of life-cycle GHG per person.

In addition to family-archetypes (such as e.g. MEL-VII), young-adult-households (MEL-XI) show also large contributions to Melbourne's overall footprint (cf. Figure 3c). In both cities small high-to-very-high-income households (MEL-II, SYD-A, SYD-J) have high shares in the overall footprints and simultaneously high per-capita footprints. This is in contrast to Froemelt et al.'s findings for Switzerland where very-high-income households had the highest per-capita footprints but minor shares in the nationwide total carbon footprint [39].

¹ Prevalence in this article is approximated by the sum of all representativeness-weights provided by ABS [63] and is visualised in Figure 2.

Table 2: Consumption archetypes found in Melbourne. The different archetypes are randomly labelled with roman numbers. The simplified subdivision into household sizes, income classes as well as the summarising titles help to better understand the nature of the archetypes and shall support the visualisation of socio-economic characteristics in Figure 2. Refer to the SI for a detailed characterisation of the archetypes. This table also provides the following information for each archetype: average age of adults (persons aged over 18 years), number of children (aged below 15 years), income in AUD per week; income per capita (p.c.); total carbon footprint (TCF) in tCO₂-eq per year; and total carbon footprint per capita. To enhance the comparison, the latter three parameters show a colour scaling from blue (low) to red (high).

	Single/Small households (1 to 1.5 persons)	Two-person households (1.6 to 2.5)	Small families (2.6 to 3.5)	Families (>3.5)
Very high income (> 4000 AUD/wk)			MEL-II Wealthy couples (some with children) Age of adults 50 No. kids (<15) 0.4 Income 7028 Income p.c. 2169 TCF 71 TCF p.c. 22	
				MEL-VII Families with children > 15 years (in education) Age of adults 38 No. kids (<15) 0.3 Income 3257 Income p.c. 707 TCF 60 TCF p.c. 13
				MEL-VIII Families with children at primary/secondary school age Age of adults 45 No. kids (<15) 1.4 Income 3189 Income p.c. 808 TCF 50 TCF p.c. 13
High income (3000 - 4000 AUD/wk)				
Medium-high income (2500 - 3000 AUD/wk)		MEL-XII Young adults (some with children) Age of adults 34 No. kids (<15) 0.4 Income 2964 Income p.c. 1224 TCF 46 TCF p.c. 19		MEL-XIV Young families with babies/toddlers Age of adults 39 No. kids (<15) 1.9 Income 2790 Income p.c. 687 TCF 47 TCF p.c. 12
Medium income (2000 - 2500 AUD/wk)		MEL-XI Young adults, old children (some in education) Age of adults 32 No. kids (<15) 0.3 Income 2393 Income p.c. 911 TCF 35 TCF p.c. 13		
		MEL-IX Couple before retirement Age of adults 55 No. kids (<15) 0.0 Income 2249 Income p.c. 929 TCF 39 TCF p.c. 16		
Low income (1000 - 2000 AUD/wk)	MEL-IV Mixed-aged adults (homeowners by trend) Age of adults 47 No. kids (<15) 0.1 Income 1559 Income p.c. 1098 TCF 26 TCF p.c. 18	MEL-V Retired couple Age of adults 65 No. kids (<15) 0.1 Income 1192 Income p.c. 537 TCF 30 TCF p.c. 14		MEL-XIII (Rather poor) young families with babies/toddlers Age of adults 39 No. kids (<15) 2.0 Income 1321 Income p.c. 327 TCF 32 TCF p.c. 8
	MEL-III Mixed-aged adults (tenants by trend) Age of adults 41 No. kids (<15) 0.0 Income 1160 Income p.c. 938 TCF 18 TCF p.c. 15	MEL-X Retired, very old couple Age of adults 76 No. kids (<15) 0.1 Income 1096 Income p.c. 504 TCF 29 TCF p.c. 13		
Very low income (< 1000 AUD/wk)	MEL-I Very old persons (females by trend) Age of adults 72 No. kids (<15) 0.0 Income 646 Income p.c. 609 TCF 15 TCF p.c. 14		MEL-VI Probably poor, unemployed, single-parent households Age of adults 41 No. kids (<15) 0.8 Income 978 Income p.c. 326 TCF 29 TCF p.c. 10	

Sydney's and Melbourne's prevalence-weighted overall footprints are comparable and in line with other studies [5,14,50]. For both cities, food (both: 25%-contribution), housing (Sydney: 23%; Melbourne: 27%) and transport (Sydney: 22%; Melbourne: 23%) are the environmentally most important consumption domains, although with a slightly different ranking in the two cities. While the most important contributions to the overall food impacts (expenditures in restaurants, for fast food, for dairy products, and more) and transport impacts (mainly direct emissions from car driving) are similar in both cases, the contributions to housing emissions differ. In Sydney, the use of electricity at home is by far the most important factor. In addition to electricity, Melbourne's direct emissions from burning natural gas is of similar importance. The differences are likely due to climatic conditions with more air conditioning required in Sydney in summer and more heating in winter in Melbourne.

Already in this overview perspective, it is apparent that different archetypes are of different importance in each city. Socio-economically, this manifests itself in small households with comparably lower incomes in Sydney and the archetypes of older population segments with relatively higher footprints in Melbourne. Furthermore, no very-high-income archetype was found in Sydney. Generally, the non-emergence of archetypes can be explained by a less distinct consumption behaviour of these households (becoming thus part of other archetypes) or a lower abundance (being either non-existent or less important and thus removed from consideration because the cluster was not regarded representative).

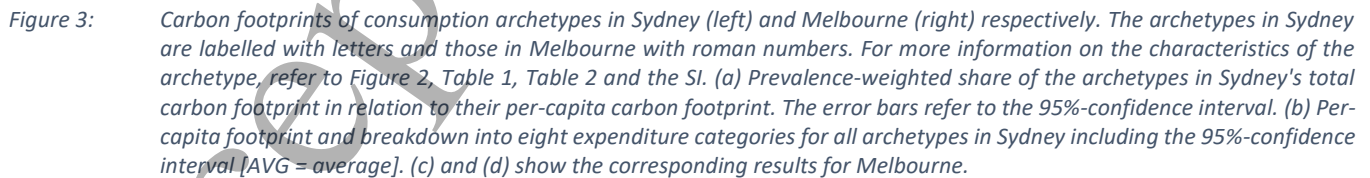


Figure 4 presents the results of the metric MDS and allows studying the similarities and dissimilarities between the two cities in more detail. While detailed data for all archetypes are provided in the SI, only a selection of the

most interesting comparisons is presented here. The choice of examples was based on often-investigated aspects like household age, income and size.

SYD-M and MEL-I are socio-demographically similar and constitute consumption patterns for elder single-person households with a little higher income in Sydney. Figure 4a and b clearly show that these two clusters are close, revealing similar consumption behaviours across cities. Both have a rather low transport demand, but a relatively high demand for health services, just as other "old-person" archetypes. Melburnians have somewhat higher expenditures on housing, while Sydneysiders spend a bit more on communications. SYD-F and MEL-X are even more alike from a socio-economic viewpoint. Both represent very old couples in similar financial situations. In Figure 4a and b, these clusters are again very close in both total and per-capita terms. However, MEL-X is still closer to MEL-V. The latter refers to younger people than MEL-X and SYD-F. The Melburnian consumption patterns being closer to each other than to the Sydney counter-part could be a hint that these consumption behaviours are city-specific to some extent. In any case, Figure 4c and d show clear differences among these four clusters with regard to footprint composition, especially in a per-capita view. The obvious segregation between Sydney clusters and Melbourne clusters in the MDS-footprint graphs (Figure 4c and d) leads to the conclusion that this is explainable by the different GHG-factors for Sydney and Melbourne. Thereby, the carbon intensities are higher for Melbourne by trend, which is at least partly attributable to the coal-based electricity production in Victoria.

In all four sub-plots of Figure 4, MEL-II (very-high-income, small families) is very distant to all other clusters. This is in line with the findings of Froemelt and colleagues [39]. In their study, the highest income segment had a special status and highest per-capita carbon footprints too. Transport demand causes the largest life-cycle GHG for this group, followed by food and housing. MEL-II's high expenditures on recreational activities induces also large environmental impacts. The most similar consumption behaviour is shown by SYD-J. Even though MEL-II represents younger adults and children together with higher incomes, SYD-J seems to be in a related financially comfortable situation. On the other end of the per-capita income distribution are SYD-L and MEL-VI. These archetypes are in a comparable occupational and social situation since many households belonging to these groups are unemployed and/or might be single-parents with low net wealth. The Sydney equivalent includes people of slightly higher age and households of a slightly lower size. However, in consumption terms, these archetypes are not as similar as could be expected. SYD-L behaves like small, low-income households (mostly single-person households such as MEL-I, MEL-III, SYD-C and SYD-I) and like poor families (MEL-XIII and SYD-G) in a per-capita view. Even though MEL-VI is not the closest cluster for SYD-L, they are obviously related in Figure 4a. However, from a total consumption view, many other consumption patterns are between the two archetypes with unemployed persons. Further investigations would be needed to clarify if these different behaviours are a consequence of living in two different cities.

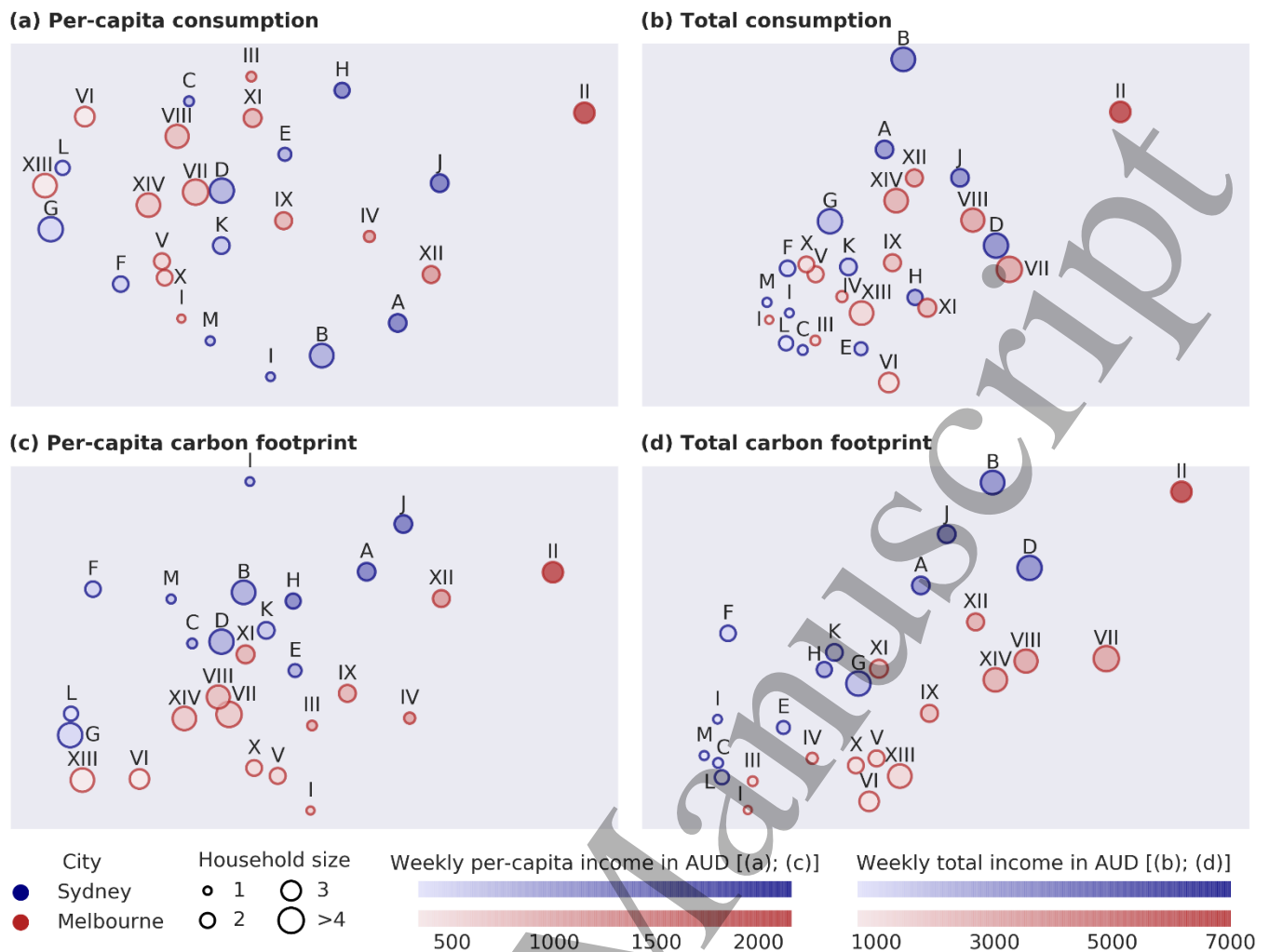


Figure 4: Results of the metric multi-dimensional scaling (MDS) of the archetypes' consumption behaviours and footprints in Sydney and Melbourne. The size of the markers visualises the average household size in the clusters, while the income is indicated by a colour shading scale in the respective city's colours (blue for Sydney, red for Melbourne). Note that MDS is a visualisation technique for multi-dimensional datasets and attempts to preserve the distances between the data points of the original, high-dimensional, data space in a 2D-plane. Therefore, the x- and y-axes do not have a specific meaning and do not correspond to each other in the different subplots. Remember that we excluded socio-economic aspects in these analyses to concentrate on investigating similar consumption profiles and footprints. Two clusters being close in these plots means that they show similar consumption behaviour or footprint compositions independent of their socio-economic circumstances. (a) Consumption behaviour per capita (shading follows per-capita income); (b) consumption behaviour per cluster (shading follows total income); (c) carbon footprint composition per capita (shading follows per-capita income); (d) carbon footprint composition per cluster (shading follows total income).

The family-archetypes in Sydney could be interpreted as a sequence from SYD-G to SYD-B and finally SYD-D with increasing age of children. These three archetypes spread over the whole space in Figure 4a and b, suggesting very different consumption behaviours of families when children are getting older. Among many different consumption categories, this is very obviously expressed by decreasing spendings for baby food and child care but increasing expenditures for education fees and recreation along this succession (see SI). However, it has to be pointed out that these clusters are also in different financial situations, which additionally affects their consumption patterns. With a very low per-capita income, SYD-G has the lowest per-capita footprint in Sydney, while SYD-B and SYD-D are placed in the middle field (cf. Figure 3b). Due to their prevalence and household size, all three clusters show substantial contributions to Sydney's overall footprint (Figure 3a).

A similar sequence also emerges in Melbourne with MEL-XIII, MEL-XIV, MEL-VIII and MEL-VII, whereby MEL-XIII and MEL-XIV are close in a socio-demographic view but differ in their income situation with MEL-XIV having a higher income. In many aspects, it seems as if SYD-G is a mix of MEL-XIV and MEL-XIII. It is interesting to see that the behaviours of these family types (young adults with babies/toddlers) led to two distinct behaviours in Melbourne but to one big archetype in Sydney. SYD-G being in-between the two corresponding Melburnian clusters could also explain the differences in behaviours observed in Figure 4a and b. While MEL-XIII and SYD-G are similar in a per-capita view, MEL-XIII obviously still behaves even more similar as SYD-L (the "unemployed" cluster above). MEL-XIV in this perspective is in transition from SYD-G to MEL-VII/SYD-D which seems intuitively correct. However, in a total perspective SYD-G has some proximity to MEL-XIV, but is closer to old two-person households, while MEL-XIII is surrounded by many different clusters. This leaves us with the conclusion that looking at socio-economic segments alone might fall short of supporting the development of targeted measures and highlights the importance of our approach that takes also the specific consumption behaviours of these segments into account.

Archetypes SYD-D and MEL-VII both represent families with children over 15 and act as a certain counter-example. They are not only socio-economically close but Figure 4a and b suggest also similar consumption behaviours in both cities. However, in turn, this is not the case for SYD-B and MEL-VIII both with children at school age. While MEL-VIII is understandably close to SYD-D in a total perspective, SYD-B is separated and takes a special status. The distance from SYD-B to other family-archetypes is even larger in a per-capita view. Despite its socio-economic counter-part in Melbourne, SYD-B obviously behaves very differently. This might be due to its high net wealth which also could explain a certain closeness to SYD-I (wealthy one-person households) in a per-capita consumption behaviour. Interestingly, the city-specific GHG-factors draw SYD-D and SYD-B closer together in the footprint-plots in Figure 4, while SYD-D and MEL-VII are pulled apart despite their similar expenditure patterns.

All these examples show that some archetypes only emerge in one of the cities, but there are also others that can be found in similar forms in both cities. However, the analyses reveal that seemingly similar archetypes, especially in a socio-economic view, might behave very differently in different cities or might even show unexpected similarities with clusters in other socio-economic circumstances. Furthermore, this study also emphasises the importance of regionalised GHG-factors. The examples of similar consumption patterns but with differing environmental consequences suggest that city-specific supply chains need to be taken into consideration, as was done in this study, when planning policy interventions.

3.3 Limitations of the study

While forming the clusters and the life cycle assessment of their consumption behaviour was done with due care and appraised with different performance metrics and evaluation techniques, our approach is still based on assumptions and simplifications that possibly influence the final results. An important issue in this regard pertains to the representativeness of the archetypes. The HES is not based on a minimum number of households and thus no official

recommendation from the ABS exists to compile a representative sample [62]. To adequately respond to this and alike issues, we developed an uncertainty framework.

Furthermore, previous studies point out that household expenditure surveys tend to underreport effective spending [42,86,87]. Besides that, we need to point out that the applied environmental assessment concentrates solely on expenditures, but neglects consequences of long-term investments. This is often encountered in input-output frameworks as is using solely monetary flows [27,88–91]. The latter assumption can lead to an overestimation of GHG when purchasers opt for better quality and thus high-price products instead of quantity [40]. Furthermore, by applying a multi-region-supply-use-table without a rest-of-world-region, we implicitly assume that domestic production technology and carbon intensity has been used for household expenditures outside Australia [92].

Finally, the MDS is purposely based on aggregated categories, but it is possible that different aggregations could lead to different pictures. Goodness-of-fit-evaluations of the MDS are presented in the SI. Thereby, it shall be mentioned that Figure 4a does not show an optimal representation of the distances, but can be considered suitable.

4. Conclusions and outlook

We derived distinctly separate lifestyle-archetypes for Sydney and Melbourne to study consumption behaviour patterns and associated environmental impacts within, but also across cities. The application of clustering techniques allowed for including expenditure data along with socio-economic parameters and hence for building household groups with comparable conditions of live and similar behaviour concurrently.

In our study, different archetypes emerged in the two cities and were of different importance. While household groups with older members were generally more consistent across cities, the lowest- and highest-income segments featured more distinct behaviours. Some socio-economically comparable groupings behaved differently in the two cities. For instance, family-archetypes with similar household-member-compositions could be found in both cities. This suggests that the age of children affects the expenditure patterns of households. However, some of those family-archetypes showed similar consumption patterns across cities, while others behaved completely differently.

In conclusion, we observed both: similar behaviour of related socio-economic groups as well as supposable city-specific behaviour. Further investigations are needed to find the causes of varying behaviours in different cities. While the search for such causal drivers is beyond the scope of the present article, the applied data-driven approach is able to quantify different lifestyles and associated environmental impacts in individual cities. Thereby, our study revealed that a detailed planning base for targeted measures in cities should not rely on general trends. A city-specific differentiation is important for the analysis of consumption behaviour as well as for GHG-factors; the latter also emphasising the importance of improving supply chains to cities. Nevertheless, some similar behaviours could be found suggesting that once an effective policy action has been designed for an individual archetype in a city, this measure could be used to target other household groups and applied even across cities. But to verify this, more research efforts are needed and psychological aspects and behavioural economics theories ought to be considered [16,35].

The application of the consumption-archetypes idea [39] to Sydney and Melbourne confirmed the importance of breaking traditional socio-economic household segments by including observed consumption data together with

socio-economic aspects when studying environmental impacts from household consumption patterns. Furthermore, we were able to advance Froemelt et al.'s approach [39] by developing an uncertainty framework. Our study might help to generate new insights into household consumption in general and particularly in cities. We also hope that this article may be a small contribution to support Australia towards the "Sustainable Transition" strongly suggested by Allen et al. [93] and we will further pursue the archetype-framework and aim for more international studies.

Acknowledgements

We thank the Swiss National Science Foundation (SNSF) for the Early Postdoc.Mobility grant (grant number: P2EZP2_184267), which supported this research. Thomas Wiedmann acknowledges funding from the Research Council of Norway under grant number 287690 (BYMARKA: Environmental Footprints of Cities). Furthermore, we would like to thank Arne Geschke and Manfred Lenzen for their valuable inputs and insightful discussions.

References

- [1] United Nations. World Urbanization Prospects: The 2011 Revision. United Nations: New York; 2012.
- [2] Dawkins E, André K, Axelsson K, Benoist L, Swartling ÅG, Persson Å. Advancing sustainable consumption at the local government level: A literature review. *J Clean Prod.* 2019;**231** 1450–62.
- [3] Cheung G, Davies PJ, Trück S. Transforming urban energy systems: The role of local governments' regional energy master plan. *J Clean Prod.* 2019;**220** 655–67.
- [4] Fry J, Lenzen M, Jin Y, Wakiyama T, Baynes T, Wiedmann T, et al. Assessing carbon footprints of cities under limited information. *J Clean Prod.* 2018;**176** 1254–70.
- [5] Chen G, Hadjikakou M, Wiedmann T. Urban carbon transformations: unravelling spatial and inter-sectoral linkages for key city industries based on multi-region input–output analysis. *J Clean Prod.* 2017;**163** 224–40.
- [6] Chen G, Shan Y, Hu Y, Tong K, Wiedmann T, Ramaswami A, et al. Review on City-Level Carbon Accounting. *Environ Sci Technol.* 2019;**53** 10 5545–58.
- [7] Kennedy CA, Ibrahim N, Hoornweg D. Low-carbon infrastructure strategies for cities. *Nat Clim Chang.* 2014;**4** 5 343–6.
- [8] Ramaswami A, Chavez A, Ewing-Thiel J, Reeve KE. Two Approaches to Greenhouse Gas Emissions Foot-Printing at the City Scale. *Environ Sci Technol.* 2011;**45** 10 4205–6.
- [9] Dubois G, Sovacool B, Aall C, Nilsson M, Barbier C, Herrmann A, et al. It starts at home? Climate policies targeting household consumption and behavioral decisions are key to low-carbon futures. *Energy Res Soc Sci.* 2019;**52** 144–58.
- [10] Mundaca L, Ürge-Vorsatz D, Wilson C. Demand-side approaches for limiting global warming to 1.5 °C. *Energy Effic.* 2018;
- [11] Harris S, Weinzettel J, Bigano A, Källmén A. Low carbon cities in 2050? GHG emissions of European cities using production-based and consumption-based emission accounting methods. *J Clean Prod.* 2020;**248** 1–13.
- [12] Mirabella N, Allacker K, Sala S. Current trends and limitations of life cycle assessment applied to the urban scale: critical analysis and review of selected literature. *Int J Life Cycle Assess.* 2019;**24** 7 1174–93.
- [13] Moran D, Kanemoto K, Jiborn M, Wood R, Többen J, Seto KC. Carbon footprints of 13 000 cities. *Environ Res Lett.* 2018;**13** 6 064041.
- [14] Wiedmann TO, Chen G, Barrett J. The Concept of City Carbon Maps: A Case Study of Melbourne, Australia. *J Ind Ecol.* 2016;**20** 4 676–91.
- [15] Froemelt A, Mauchle M, Steubing B, Hellweg S. Greenhouse Gas Emissions Quantification and Reduction Efforts in a Rural Municipality. *J Ind Ecol.* 2018;**22** 1 92–105.
- [16] O'Rourke D, Lollo N. Transforming Consumption: From Decoupling, to Behavior Change, to System Changes for Sustainable Consumption. *Annu Rev Environ Resour.* 2015;**40** 233–59.
- [17] Schanes K, Giljum S, Hertwich E. Low carbon lifestyles: A framework to structure consumption strategies and options to reduce carbon footprints. *J Clean Prod.* 2016;**139** 1033–43.
- [18] Tukker A, Cohen MJ, Hubacek K, Mont O. The Impacts of Household Consumption and Options for Change. *J Ind Ecol.* 2010;**14** 1 13–30.
- [19] Ivanova D, Stadler K, Steen-Olsen K, Wood R, Vita G, Tukker A, et al. Environmental Impact Assessment of Household Consumption. *J Ind Ecol.* 2016;**20** 3 526–36.
- [20] Vita G, Lundström JR, Hertwich EG, Quist J, Ivanova D, Stadler K, et al. The Environmental Impact of Green Consumption and Sufficiency Lifestyles Scenarios in Europe: Connecting Local Sustainability Visions to Global Consequences. *Ecol Econ.* 2019;**164** 106322.
- [21] IPCC. IPCC Fifth Assessment Report. 2013.
- [22] Moran D, Wood R, Hertwich E, Mattson K, Rodriguez JFD, Schanes K, et al. Quantifying the potential for consumer-oriented policy to reduce European and foreign carbon emissions. *Clim Policy.* 2018;
- [23] Lekve Bjelle E, Steen-Olsen K, Wood R. Climate change mitigation potential of Norwegian households and the rebound effect. *J Clean Prod.* 2018;**172** 208–17.
- [24] Lorek S, Spangenberg JH. Sustainable consumption within a sustainable economy - Beyond green growth and green economies. *J Clean Prod.* 2014;**63** 33–44.

- [25] Jones CM, Kammen DM. Quantifying carbon footprint reduction opportunities for U.S. households and communities. *Environ Sci Technol.* 2011;**45** 9 4088–95.
- [26] Baiocchi G, Minx JC, Hubacek K. The Impact of Social Factors and Consumer Behavior on Carbon Dioxide Emissions in the United Kingdom. A Regression Based on Input-Output and Geodemographic Consumer Segmentation Data. *J Ind Ecol.* 2010;**14** 1 50–72.
- [27] Minx JC, Wiedmann T, Wood R, Peters GP, Lenzen M, Owen A, et al. Input-output analysis and carbon footprinting: An overview of applications. *Econ Syst Res.* 2009;**21** 3 187–216.
- [28] Mi Z, Zheng J, Meng J, Ou J, Hubacek K, Liu Z, et al. Economic development and converging household carbon footprints in China. *Nat Sustain.* 2020;**3** 7 529–37.
- [29] Wiedmann T, Lenzen M, Keyßer LT, Steinberger JK. Scientists' warning on affluence. *Nat Commun.* 2020;**11** 3107.
- [30] Steffen W, Richardson K, Rockström J, Cornell S, Fetzer I, Bennett E, et al. Planetary boundaries: Guiding human development on a changing planet. *Science (80-).* 2015;**347** 6223 1259855.
- [31] Secretary-General (United Nations). Progress towards the Sustainable Development Goals. Economic and Social Council. E/2017/66; United Nations: New York; 2017.
- [32] United Nations Framework Convention on Climate Change (UNFCCC). The Paris Agreement [Internet]. 2018 [cited 2018 Feb 15]. Available from: http://unfccc.int/paris_agreement/items/9485.php
- [33] Hoegh-Guldberg O, Jacob D, Taylor M, Guillén Bolaños T, Bindi M, Brown S, et al. The human imperative of stabilizing global climate change at 1.5°C. *Science (80-).* 2019;**365** 6459 eaaw6974.
- [34] D'Amato A, Giaccherini M, Zoli M. The Role of Information Sources and Providers in Shaping Green Behaviors. Evidence from Europe. *Ecol Econ.* 2019;**164** 106292.
- [35] Frederiks ER, Stenner K, Hobman E V. Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. *Renew Sustain Energy Rev.* 2015;**41** 1385–94.
- [36] Chen G, Zhu Y, Wiedmann T, Yao L, Xu L, Wang Y. Urban-rural disparities of household energy requirements and influence factors in China: Classification tree models. *Appl Energy.* 2019;**250** 1321–35.
- [37] Hubacek K, Baiocchi G, Feng K, Muñoz Castillo R, Sun L, Xue J. Global carbon inequality. *Energy, Ecol Environ.* 2017;**2** 6 361–9.
- [38] Duarte R, Feng K, Hubacek K, Sánchez-Chóliz J, Sarasa C, Sun L. Modeling the carbon consequences of pro-environmental consumer behavior. *Appl Energy.* 2016;**184** 1207–16.
- [39] Froemelt A, Dürrenmatt DJ, Hellweg S. Using Data Mining To Assess Environmental Impacts of Household Consumption Behaviors. *Environ Sci Technol.* 2018;**52** 15 8467–78.
- [40] Girod B, De Haan P. More or Better? A Model for Changes in Household Greenhouse Gas Emissions due to Higher Income. *J Ind Ecol.* 2010 Jan 27;**14** 1 31–49.
- [41] Girod B, De Haan P. GHG reduction potential of changes in consumption patterns and higher quality levels: Evidence from Swiss household consumption survey. *Energy Policy.* 2009 Dec;**37** 12 5650–61.
- [42] Weber CL, Matthews HS. Quantifying the global and distributional aspects of American household carbon footprint. *Ecol Econ.* 2008;**66** 2 379–91.
- [43] Druckman A, Jackson T. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. *Energy Policy.* 2008;**36** 8 3177–92.
- [44] Jones CM, Kammen DM. Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density. *Environ Sci Technol.* 2014;**48** 2 895–902.
- [45] Büchs M, Schnepf S V. Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO₂ emissions. *Ecol Econ.* 2013;**90** 114–23.
- [46] Bin S, Dowlatabadi H. Consumer lifestyle approach to US energy use and the related CO₂ emissions. *Energy Policy.* 2005;**33** 2 197–208.
- [47] Lenzen M, Wier M, Cohen C, Hayami H, Pachauri S, Schaeffer R. A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. *Energy.* 2006;**31** 2 181–207.
- [48] Wiedenhofer D, Guan D, Liu Z, Meng J, Zhang N, Wei Y-M. Unequal household carbon footprints in China. *Nat Clim Chang.* 2017;**7** 1 75–80.
- [49] Froemelt A, Buffat R, Hellweg S. Machine learning based modeling of households: A regionalized bottom-up approach to investigate consumption-induced environmental impacts. *J Ind Ecol.* 2020;**24** 3 639–52.
- [50] Chen G, Wiedmann T, Hadjikakou M, Rowley H. City carbon footprint networks. *Energies.* 2016;**9** 8 602.
- [51] Lenzen M, Peters GM. How City Dwellers Affect Their Resource Hinterland: A Spatial Impact Study of Australian Households. *J Ind Ecol.* 2010;**14** 1 73–90.
- [52] Baynes T, Lenzen M, Steinberger JK, Bai X. Comparison of household consumption and regional production approaches to assess urban energy use and implications for policy. *Energy Policy.* 2011;**39** 11 7298–309.
- [53] Hertwich EG, Peters GP. Carbon Footprint of Nations: A Global, Trade-Linked Analysis. *Environ Sci Technol.* 2009;**43** 16 6414–20.
- [54] Economist Intelligence Unit. The Global Liveability Index [Internet]. 2019 [cited 2020 May 7]. Available from: <https://www.eiu.com/topic/liveability>
- [55] Mercer. Quality of Living Ranking [Internet]. 2019 [cited 2020 May 7]. Available from: <https://www.mercer.com/newsroom/2019-quality-of-living-survey.html>
- [56] CNBC. Deutsche Bank Liveability Survey [Internet]. 2019 [cited 2020 May 7]. Available from: <https://www.cnbc.com/2019/05/20/these-cities-offer-the-best-quality-of-life-deutsche-bank-says.html>
- [57] Dürrenmatt DJ, Gujer W. Identification of industrial wastewater by clustering wastewater treatment plant influent ultraviolet visible spectra. *Water Sci Technol.* 2011;**63** 6 1153–9.
- [58] Vesanto J, Alhoniemi E. Clustering of the Self-Organizing Map. *IEEE Trans Neural Networks.* 2000;**11** 3 586–600.
- [59] Vesanto J, Sulkava M. Distance matrix based clustering of the Self-Organizing Map. In: *Artificial Neural Networks - ICANN 2002*. 2002.

- [60] Canetta L, Cheikhrouhou N, Glardon R. Applying two-stage SOM-based clustering approaches to industrial data analysis. *Prod Plan Control*. 2005;**16** 8 774–84.
- [61] Lenzen M, Geschke A, Malik A, Fry J, Lane J, Wiedmann T, et al. New multi-regional input–output databases for Australia—enabling timely and flexible regional analysis. *Econ Syst Res*. 2017;**29** 2 275–95.
- [62] Australian Bureau of Statistics (ABS). 6503.0 - Household Expenditure Survey and Survey of Income and Housing, User Guide, Australia, 2015–16 [Internet]. Canberra, ACT, Australia: Australian Bureau of Statistics (ABS); 2017 [cited 2019 Aug 1]. Available from: <https://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/6503.0?OpenDocument>
- [63] Australian Bureau of Statistics (ABS). 6540.0 - Microdata: Household Expenditure, Income and Housing, 2015–16, Basic Confidentialised Unit Record File (CURF), MicrodataDownload, Findings based on use of ABS Microdata. 2018.
- [64] Australian Bureau of Statistics (ABS). 1270.0.55.001 - Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas [Internet]. 2016 [cited 2019 Aug 1]. Available from: [https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/1270.0.55.001~July 2016~Main Features~Overview~1](https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Overview~1)
- [65] Fisher RA. Statistical methods for research workers. 14th ed. New York, NY, USA: Hafner; 1973. 362 p.
- [66] Kohonen TK. Self-Organizing Maps. Vol. 30, Springer Series in Information Sciences. Springer Berlin Heidelberg; 2001. 501 p.
- [67] Kohonen T. The Self-Organizing Map. *Proc IEEE*. 1990;**78** 9 1464–80.
- [68] Vesanto J, Himberg J, Alhoniemi E, Parhankangas J. SOM Toolbox for Matlab 5. Helsinki University of Technology, Technical Report A57. Technical Report A57; Helsinki University of Technology: Helsinki, Finland; 2000.
- [69] Tan H Sen, George SE. Investigating learning parameters in a standard 2-D SOM model to select good maps and avoid poor ones. In: *Australasian Joint Conference on Artificial Intelligence*. Springer Berlin Heidelberg; 2004. p. 425–37.
- [70] Vesanto J. Neural network tool for data mining: SOM toolbox. In: *Proceedings of symposium on tool environments and development methods for intelligent systems (TOOLMET 2000)*. 2000. p. 184–96.
- [71] Anderberg MR. Cluster analysis for applications. New York, NY, USA: Academic Press; 1973. 359 p.
- [72] Hartigan JA. Clustering algorithms. New York, NY, USA: Wiley; 1975.
- [73] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2012;**12** 2825–30.
- [74] Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;**20** 53–65.
- [75] Lötsch J, Ultsch A. Exploiting the Structures of the U-Matrix. In: *Advances in Self-Organizing Maps and Learning Vector Quantization*. Cham: Springer; 2014. p. 249–57.
- [76] Ultsch A. U*-Matrix: a Tool to visualize Clusters in high dimensional Data. Fachbereich Mathematik und Informatik. University of Marburg: Marburg, Germany; 2003.
- [77] Ward JH. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc*. 1963;**58** 301 236–44.
- [78] Bundesamt für Statistik (BFS). Haushaltsbudgeterhebung 2011: Kommentierte Ergebnisse und Tabellen. Bundesamt für Statistik (BFS): Neuchâtel, Switzerland; 2013.
- [79] Efron B, Tibshirani R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Stat Sci*. 1986;**1** 1 54–77.
- [80] Lenzen M, Geschke A, Wiedmann T, Lane J, Anderson N, Baynes T, et al. Compiling and using input-output frameworks through collaborative virtual laboratories. *Sci Total Environ*. 2014;**485–486** 1 241–51.
- [81] Geschke A, Hadjikakou M. Virtual laboratories and MRIO analysis—an introduction. *Econ Syst Res*. 2017;**29** 2 143–57.
- [82] Australian Government (Department of the Environment and Energy). Australian Greenhouse Emissions Information System (AGEIS) [Internet]. 2018 [cited 2019 Aug 1]. Available from: <https://www.environment.gov.au/climate-change/climate-science-data/greenhouse-gas-measurement/ageis>
- [83] Chen G, Hadjikakou M, Wiedmann T, Shi L. Global warming impact of suburbanization: The case of Sydney. *J Clean Prod*. 2018;**172** 287–301.
- [84] Australian Government (Department of the Environment and Energy). National Inventory by Economic Sector - Australian Greenhouse Emissions Information System. 2018.
- [85] Borg I, Groenen P. Modern Multidimensional Scaling - Theory and Applications. Springer Series in Statistics; 1997.
- [86] Steen-Olsen K, Wood R, Hertwich EG. The Carbon Footprint of Norwegian Household Consumption 1999–2012. *J Ind Ecol*. 2016;**20** 3 582–92.
- [87] Ivanova D, Vita G, Steen-Olsen K, Stadler K, Melo PC, Wood R, et al. Mapping the carbon footprint of EU regions. *Environ Res Lett*. 2017;**12** 5.
- [88] Barrett J, Peters G, Wiedmann T, Scott K, Lenzen M, Roelich K, et al. Consumption-based GHG emission accounting: a UK case study. *Clim Policy*. 2013;**13** 4 451–70.
- [89] Miller RE, Blair PD. Input-Output Analysis: Foundations and Extensions. 2nd ed. Cambridge University Press: Cambridge, United Kingdom; 2009.
- [90] Rodrigues JFD, Moran D, Wood R, Behrens P. Uncertainty of Consumption-Based Carbon Accounts. *Environ Sci Technol*. 2018;**52** 13 7577–86.
- [91] Wiedmann T, Wood R, Minx JC, Lenzen M, Guan D, Harris R. A carbon footprint time series of the UK - results from a multi-region input-output model. *Econ Syst Res*. 2010;**22** 1 19–42.
- [92] Wiedmann T. Editorial: Carbon footprint and input-output analysis - an introduction. *Econ Syst Res*. 2009;**21** 3 175–86.
- [93] Allen C, Metternicht G, Wiedmann T, Pedercini M. Greater gains for Australia by tackling all SDGs but the last steps will be the most challenging. *Nat Sustain*. 2019;**2** 11 1041–50.