Reviewer #1  
  
The paper should be considered for publication after revision, as it attempts to fill an important knowledge gap in the urban climate space. A few recent publications (e.g. Rosenzweig et al, 2018) attempt to address the question of geographic scope and thematic coverage, but this paper provides a more comprehensive and up to date review of the literature.

**General Comments**

One of the primary shortcomings of the paper, is its depth coverage and lack of nuanced analysis of cities from the global south (identified as a gap), but there has been an expansion of the literature in this space in the AR5 and post-AR5 period.

The choice of Web of Science and Scopus as the core source of literature may be a limiting factor as both have limited coverage of the highly interdisciplinary space of urban climate change (line 41). A test of this method may have been its expansion using Google Scholar that has a wider pool of literature in this area. It is also unclear whether the authors have examined the fairly extensive language literature on East Asian and Latin American cities.

The reviewer raises some important issues. However, in the first instance, we must emphasise the scale of the analysis being conducted here: in the initial submission we surveyed the coverage and content of 3,440 case study articles. While this survey is undoubtedly incomplete, it is far more comprehensive in scope, if not in depth, than all efforts we are currently aware of. We would argue that limitations in literature scope should be considered in light of this new level of ambition.

Of course, systematic omissions are possible. Indeed, a central thread in our article pertains to literature biases, therefore we must ensure a good balance of coverage. To do so we have taken three further actions: (1) we expanded our search to google scholar; (2) we conducted our search in French and Spanish; and (3) we conducted new analysis looking at the post-AR5 literature expansion.

**1. Google scholar**

Google Scholar (GS) is a very problematic database to work with. First of all, it has a much poorer functionality for literature searches compared to Web of Science (WoS) and Scopus. Second, it typically returns a large volume of imprecise results (Haddaway *et al* 2015). And third, unlike other databases, GS actively prohibits users from bulk downloading document meta-data, particularly abstracts. Together this renders GS almost useless for the type of bibliometric analysis we perform in this article – which is based on the automated screening of thousands of abstracts for city names.

Nonetheless, we are interested in whether GS results offer more cases on Southern cities, as suggested. Therefore we manually screened the results from a GS search, importing 100 relevant articles (33 additional, non-relevant articles had to be discarded during this process). Many of these documents are books, book chapters and reports. In these cases we had to search for introductions or summaries that describe the case study locations, importing these to our database instead of abstracts. After this process we identify from the GS set a further 27 case studies, with the regional breakdown shown in table 1.

The main systematic difference we note in the GS sample is the higher proportion of cases on Latin American & Caribbean cities. Hence the assertion that we overlook cases on this region may be correct. However, we find that the total number of additional case study documents to be obtained is low – due to the limited number of actual cases we obtain from the GS search (out of the 133 “most relevant” search results, we obtain only 7 Latin American cases).

|  |  |  |  |
| --- | --- | --- | --- |
| **Region** | **n (GS sample)** | **Region % in GS sample** | **Region % in total sample (WoS, Scopus)** |
| AFRICA | 1 | 1.7 | 3.7 |
| ASIA | 21 | 36.2 | 37.9 |
| EUROPE | 16 | 27.6 | 25.5 |
| LATIN AMERICA AND THE CARIBBEAN | 7 | 12.1 | 4.7 |
| NORTH AMERICA | 12 | 20.7 | 23.8 |
| OCEANIA | 1 | 1.7 | 3.7 |

Table 1: results for a Google Scholar search

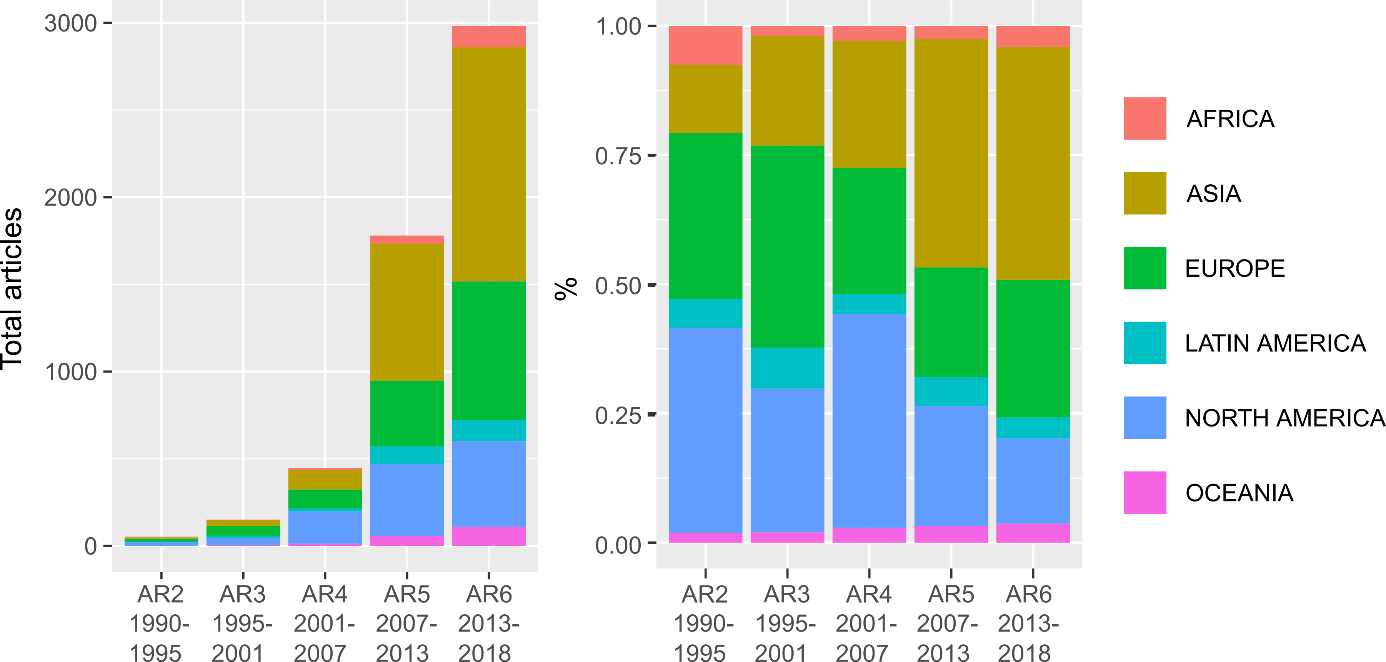
**2. Other languages**

We translated our search query into Spanish and French and applied it in WoS, Scopus and Google Scholar. This derives no extra results for WoS and Scopus – since these platforms already translate and catalogue all abstracts and titles in English, all relevant foreign language articles are captured by our original query.

In GS we find an additional 18 publications in French and 78 publications in Spanish. None of these are in fact peer reviewed articles. Considering this surprisingly small number of documents (of which we would expect approximately two dozen case studies), the aforementioned difficulties with GS, and the fact that we would have to manually obtain and translate the abstracts in these documents, we have decided not to import them into our database at this time.

**3. Post-AR5 literature expansion**

Figure 1 below shows the growth and fractions of case studies by region and AR period. This demonstrates that we indeed capture a large (and increasing) share of literature on Asian cities, but shows no evidence for a relative growth of articles African and Latin American cities in WoS and Scopus.



**Figure 2: Total articles and regional proportions of case study literature by IPCC Assessment Period**

Overall, our judgement based on these extensions in search strategy, languages and analysis is that few further cases can be easily identified for cities in the global South. We believe we have captured a majority of the relevant peer reviewed literature – and substantially more than comparable efforts. Further gains will be increasingly hard to obtain (i.e. there is a basic trade-off between comprehensiveness and accuracy).

In the manuscript, we have taken the following actions:

* We note the limitations of this analysis in the methods section and underline our focus on the peer reviewed literature:

Our results are not fully comprehensive, since further databases are available (e.g. Google Scholar) and relevant non-English language articles may exist. However, on translating our query into Spanish and French, and applying all versions in Google Scholar, we find very few further results, even from the “most relevant” documents – and no further peer-reviewed documents were identified. More in-depth studies, e.g. systematic maps on specific topics or cities, could consider additional databases and non-peer reviewed sources. (Line 431-436)

* We reference the additional analysis on the post-AR5 literature expansion in the main text, and place the figure in the SI Text:

Although we do observe an uptake in case studies on Asian cities since the IPCC 5th Assessment Report, there has been only a very limited expansion of cases on African and Latin American cities in recent years (SI Text Fig 4). (Line 95-97)

* We provide the full list of cases as a spreadsheet in this re-submission. This will be published with the article. It can be searched by city and topic, downloaded for analysis, and contributed to. We see this as an important starting point for on-going community efforts at in-depth case study mapping and synthesis.

**4. Additional changes to analysis**

In line with the reviewer’s comments below, that adaptation itself is a whole field of study – one that likely merits its own review – we adjusted our search query to filter out adaptation-only case studies. We found these were being included due to a combination of keywords for “climate policy”. We adjusted the search to remove this combination, while preserving mitigation relevant combinations (e.g. “climate mitigation policy”). As a result “adaptation” no longer appears as a topic in the analysis in the second section, and the manuscript has been adjusted throughout. Note that many studies combining adaptation and mitigation are still included, particularly those with a strong governance and policy focus.

After this adjustment, and updating for recent publications during the review phase of this article, we now capture 4,051 articles, covering 5,570 cases on individual cities. The relevant text and figures have been updated throughout.

A further consequence of this adjustment is that we capture even more studies on East Asian cities, which are rapidly growing in number (see Figure 2 above). This changes the weight of topics in our sample, giving even more detail on the large sub-literature on urban emissions accounting in China. Hence our topic modelling analysis generates a further topic on these studies, capturing a particular set of phrases that are used to motivate their analysis i.e. regarding low-carbon urban development. Since this is less content focused, and specific to this set of studies, we manually remove it from the main manuscript for clarity, but document it in the SI text Table 1.

Finally, on review we found that our analysis of average citations at a regional level was not robust, as results one region (Oceania) were driven by a small set of extremely highly cited papers. Since this result was only briefly mentioned, we have simply removed it from the main manuscript and SI text.

The suggested quantitative typology of cities does not appear to be robust and may be influenced by the bias in the sample used. For example, heating degree days is a significant variable, but because cooling needs are not examined, cities in tropical or composite climates do not apparently appear in the clustering. Fuel price data masks a complex regime of subsidies and carbon taxation in multiple geographies. Similarly, energy poverty is a defining characteristic of Asian and African cities. Have the authors not found this in the literature or is there a methodological gap in the search terms that they are using?

Note that the development of the typology is NOT a contribution of this manuscript. Instead we highlight the promise of linking case-study evidence with typologies for an enhanced learning on urban climate solutions. For demonstrative purposes we use an EXISTING typology from the literature (Creutzig *et al* 2015). Therefore, we refer the reviewer to the original publication for any issues of robustness of the typology, where the underlying algorithm is substantiated with checks for tree size and clustering (Creutzig *et al* 2015).

The reviewer is right that the city sample is biased and underrepresents the cooling dimension. However, it should also be noted that until recently cooling energy demand was relatively low even in hot climates due to low average income levels (e.g. in India). Yet, we argue that concerns over details of this typology are not central for our use here. We address the energy poverty point in specific comment #5 below, and the wider issue of linking case studies with typologies in the next comment.

A useful analysis of co-benefits i.e. synergies and trade-offs between mitigation and adaptation that appears embedded within the clustering of the three most important themes by city in Figure 5 is not drawn upon. The conflation of the quantitative typology in the upper part of Figure 5 with the case study literature below is potentially confusing at worst or dense at best.

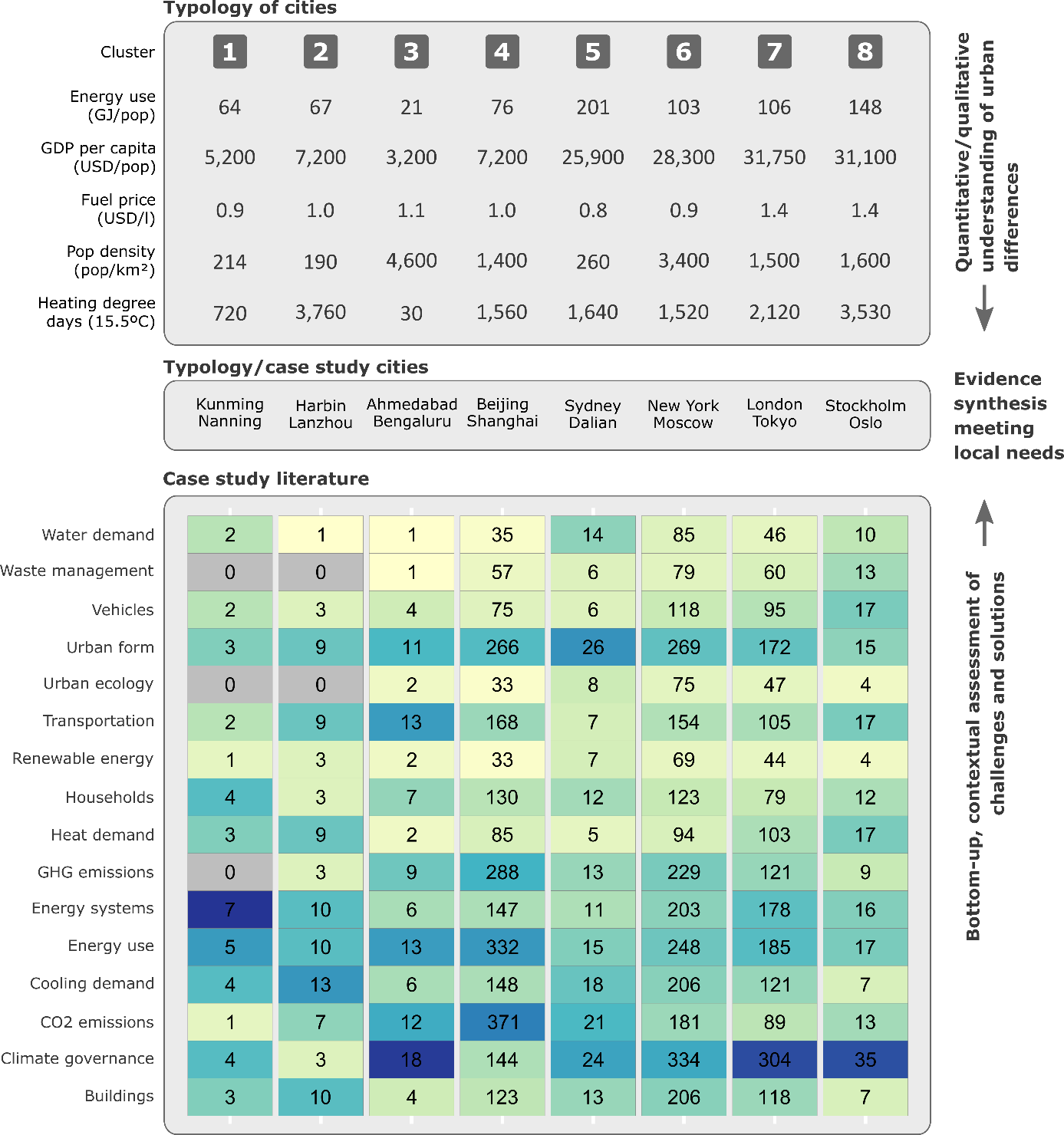
As discussed previously, we have removed adaptation-only studies, so this specific result is lost.

Our motivation for linking a typology with the case study literature follows from the general questions posed in our piece: What types of cities and topics are being studied? And what are the different ways to learn from this evidence? We see that typologies are commonly used in the urban literature, and can structure assessments of policies and comparative case studies. For instance, the Atkins Future Proofing Cities report organises policies by a qualitative typology of cities, while an ICLEI report assesses infrastructure transitions in the specific urban context of post-industrial cities.

Hence we make a specific suggestion: build maps of case study evidence that are searchable by typology, as well as by city size, region and topic (as we do elsewhere in the paper). In doing so, we see more clearly that specific urban types (as defined by their drivers of energy use) are very poorly researched; while others have masses of evidence yet to be consolidated. Further, we observe congruence in the types of topics that are researched relative to those factors identified as important in the typology. The linking therefore articulates a much more actionable agenda when it comes to identifying research needs and the next opportunities for evidence synthesis.

A series of research gaps were raised in the reviews to this article: cooling needs, energy poverty, pricing mechanisms, further detail on adaptation, and so forth. We share the concern that certain research may be overlooked. Our conceptual advance is designed to address this, by showing the extent of case study literature relative to the most pressing issues faced by cities. This has obvious policy relevance. We state it as an idea that should be further explored, regarding: (1) different qualitative and quantitative typologies; and (2) opportunities for structuring synthesis projects that bridge the current gap between global quantitative urban studies and the body of qualitative contextually rich case studies.

In terms of actions, we have shortened and re-written this section to better communicate these messages. Please refer to the section “Synthesizing urban typologies and case study evidence” in the manuscript (too large to quote here). Additionally, we have re-drawn figure 5 explicitly communicate the quantitative typology data and underlying case literature, while bringing attention to the intended outcomes of the typology/case matching (see text and arrows on the right hand side):



Similarly, economic and pricing incentives seems to be an unexplored theme in the paper. Is this because of the underlying literature or a limited search for these terms that are otherwise significant in the mitigation literature?

We can only answer this question with reference to content we find in these case studies. Our topic modelling analysis suggests economic and pricing policies are not widely discussed as headline approaches to urban mitigation. For instance, the set of keywords gathered under the main policy topic (“climate governance”) – and their associated documents (the spreadsheet attached to this submission can be sorted by this topic to show the most relevant ones) – mainly refer to generic governance issues, mitigation and adaptation action plans, as well as multi-level coordination. Instead, keywords such as “income” and “price” are found under the “households” topic (SI Table 1), showing the importance of economic incentives for the analysis of household behaviours in the case study literature.

This is now noted in the manuscript:

Policy focused cases associated with the “climate governance” topic tend to refer to mitigation action plans, multi-level coordination and generic governance issues (SI Text Table 1). Pricing and economic analysis appears to be only a marginal topic in the urban case study literature (as pointed out by a reviewer of this manuscript), despite the importance of fuel prices as a long-run determinant of urban form and GHG emissions (Creutzig *et al* 2015, Creutzig 2014). (Line 171-175)

It may be useful to highlight in the abstract the finding that world regions and city scales with most future relevance are systematically unrepresented in the literature. The limited scope of comparative regional/urban case research is an important finding that does not find its way to the conclusions/abstract. That case study analysis is an underdeveloped field that has limited generalisability, is buried dep in the text.

We edited and included the following sentence in the abstract:

We find that cities in world regions with the highest future mitigation relevance are underrepresented, while comparative analysis between cities and efforts to synthesize the existing case study evidence remain limited in scope. (Line 7-9)

The paper provides a useful analysis of both knowledge gaps and multiple ways forward, crisper articulation in the conclusions may enhance its utility to both scientists and policymakers.

**Detailed/Technical Comments**

1. Line 48: GDP per capita is a measure of income not wealth

We have replaced “wealth” with “income”.

2. Line 48: infrastructure development is different from access, which is identified as an important variable in AR5

We have changed the text to:

Cities vary in many dimensions, including size, income and levels of access to different infrastructures (such as public transportation and energy services). (Line 56-57)

3. Line 53: there is still an open debate on the definition of urban agglomerations, cities and urban areas. Why confound this by adding metrpoles in the mix? There is no consistency in national definitions of urban areas. It may be useful to reference the urban frame work (UN, 2014) that is used by the paper

We have replaced “metropoles” with “cities”. Please note the suggested citation is already provided in the caption to the figure, which is referred to in this paragraph.

4. Line 80: both infrastructure and building transformations are important as identified by the paper. Why privilege only one? Does the literature actually support this statement? The case study shows extensive literature on East Asian cities, which are currently significant contributors to urban and global emissions. There are large middle-income countries in this region.

We have changed the text to:

With the window on the 1.5°C and 2°C goals rapidly closing it is essential to immediately initiate urban infrastructure transformations and comprehensive low-carbon retrofitting ~~in wealthier Northern cities~~, worldwide. (Line 87 - 88)

The point of this paragraph is to emphasise the trade-offs between different geographical foci in urban case studies: cases on wealthier and high-emitting cities are needed, especially considering their current and historical responsibility for carbon emissions; yet, much more work should be done on up-coming cities, where a large potential to ‘lock-in’ emissions exists. This has been clarified with the following changes:

The current focus of case study research on wealthier and high-emitting cities is congruent with current debates in climate ethics: that responsibility for drastic mitigation action rests on the shoulders of high-emitters. Nonetheless, the majority of future urban emissions growth will originate from up-coming Asian and African cities, where ongoing processes of urbanization and infrastructure development provide a window of opportunity for establishing urban designs consistent with low-carbon mode choices and building use. Guiding these growing cities towards compact, low-carbon urban forms requires a major shift in research focus. (Line 88-95)

5. Lines 85 and 86: Energy poverty is a defining characteristic of Asian and African cities. Have the authors not found this in the literature or is there a methodological gap in the search terms that they are using?

Actually we find only 14 documents that refer to “energy poverty” directly in the title or abstract. These are mainly situated in India, Bangladesh and South Africa. We consider this topic out of scope for our review, as it is large enough to require a dedicated study, particularly to capture non-urban (rural) case studies.

6. Line 93: a more careful use of the term agglomeration may be useful

We have changed the text to:

We can safely presume that coverage is even worse for small cities, even though data on the number of ~~small agglomerations~~ these remains incomplete. (Line 103-104)

7. Line 98: economic and pricing incentives seems to be an unexplored theme in the paper. Is this because of the underlying literature or a limited search for these terms that are otherwise significant in the mitigation literature?

We address this issue in the general comments above.

8. Line 99: a definition of ‘well-being’ or relevant citation that helps clarify what the authors mean may be useful. This is a very ‘catholic’ term, open to a wide range of interpretation

This is an introductory sentence, with no firm definitions in mind. Our point is to frame this section by suggesting that ‘urban climate mitigation’ research encompasses a wide variety of topics – and methods are needed to uncover such content across large literatures. We have replaced ‘well-being’ with ‘social and environmental sustainability’, and reference this.

9. Line 103: missing cooling needs in tropical regions is a major omission, given that a core theme of the paper is climate change

After the changes we make to the literature set, described in the general comments above, cooling demand indeed emerges as a core topic in the analysis; it is particularly prevalent in studies on North American and Latin American cities.

10. Line 113: the claim of ‘unsupervised learning’ being less subjective, is not substantiated in the paper. This is a complex and contested territory that may be avoided. For a sense of the complexity involved see the IPCC guidance notes on uncertainty and the considerable literature in this space (e.g. Mastrandrea et al, 2010)

We have deleted this sentence.

11. Lines 118 to 124: Many key themes identified in IPCC AR5 are missing in the analysis: energy efficiency, energy storage, smart grids. Unclear whether this is a gap is in the method used or the literature. It would have been useful if piecewise continuity from the systematic reviews in this space (many of them cited) and the thematic analysis were maintained through the paper, else key threads/mitigation options may be missed.

We are confident that these themes are found in the documents we identify: energy efficiency is captured under the ‘energy use’ topic; energy storage and smart grids in the ‘energy systems’ topic. These can be seen in the stemmed keywords in SI Table 1, as well as the case studies marked up by these topics in the data provided with this article.

Beyond the categorisation described briefly in this section, we have only limited space to discuss specific topics in detail. Furthermore, our focus and methods in the article are aimed at a broader scope – i.e. which cities do we know about? What topics do we know about? What comparative and secondary analysis is there of cases? And how can generalizable knowledge be derived from urban cases? We clarify this intention in the introduction, with following phrase:

As a starting point to our analysis, we obtain a sample of urban mitigation articles using a search query that combines synonyms for “urban” and “mitigation” in the Web of Science and Scopus literature databases (see methods). Our interpretation of case study research is straightforward: if an article mentions a city name in the abstract or title, we assume it is a case study located in the city (or cities) mentioned. We validate this assumption subsequently for a representative sample of cases. Our dataset for the proceeding analysis consists of 4,051 case study publications. We use language processing methods to extract relevant meta-data on case study locations, topics, review studies, and other information. In our article, this ‘mapping’ of the case literature is limited in analytical depth, but it is highly comprehensive in scope. The database of case studies we provide with this article thus enables researchers to ‘zoom in’ on particular topics or places of interest, with the intention of stimulating future work, refinement, and large-scale assessments. (Line 44-54)

12. Line 125: Climate adaptation is a field of knowledge in itself. A clear analysis of its relationship to mitigation may be useful. Conflation as just another term isn’t very useful.

We address this issue in the general comments above.

13. Lines 128 to 130: unsubstantiated statement. There are many other possible explanations of the importance given to emissions and urban form in East Asia, including government policy. Is there more substantial evidence on this from the regional literature?

The statement is already qualified with “perhaps”. It also provides quantitative evidence for a key part of the argument. We further qualify this statement by noting that the engineering focus not only translates into focus in research paper but also in government policy.

14. Lines 138 to 140: is this gap in the literature not an artefact of the sources used (Scopus and Web of Science) compared to the relatively rich regional literature in this space on these themes?

We are unable to answer this question, since there is no “comprehensive” body of case study literature to compare with our study. Indeed, we believe our attempt is the most comprehensive so far, especially in the academic literature. Nevertheless, we condition these statements with the following:

Emissions accounting dominates the case study work on top-tier cities in China (Beijing, Shanghai and Tianjin), while issues of water demand, urban ecology and climate governance receive less attention – at least in the mitigation focused literature we identify from Web of Science and Scopus. (Line 146-149)

15. Line 164: is the ‘uneven distribution’ of adaptation vs. mitigation literature in Africa and Latin America evidence of a differentiation in the key challenges faced there?

We address this adaptation point in the general comments above.

16. Lines 169 to 170: a more nuanced articulation of this finding in relation to the available IPCC AR6 chapter outlines may be useful. Other chapters than the ‘demand chapter’ may benefit from this attention.

We have simply removed this statement and now dedicate a new section to assessments (see ‘Reviews and assessments of cases’). We have limited space for a detailed outline of contributions, but give the following phrase:

In the upcoming IPCC AR6, urban case studies will be a key evidence base for both the urban and demand chapters in Working Group III (Mitigation), not to mention the regional chapters under Working Group II (Impacts & Adaptation). (Line 270-272)

17. Lines 188 to 194: more could be done with the chord diagram, in relation of existing analysis from other sources {IPCC AR5, Bartlett and Satterthwaite, 2016, Rosenzweig, 2018 etc.)

These references reflect the relevant urban assessment literature, each containing a few or many case studies. None of these studies systematically maps the case study literature, and the existing analysis operates on a different level. Hence a direct mapping of case studies from this books/assessment into this graph would not contribute much. Overall, we are therefore not sure what is being requested and cannot respond, beyond citing these obviously important literatures under the following text:

There is substantial comparative work analysing mitigation and adaptation plans across hundreds of cities (Reckien *et al* 2018, 2014, Castán Broto and Bulkeley 2013) and more detailed studies published as volumes (Bartlett and Satterthwaite 2016). Much attention is also paid to comparing urban carbon footprints (Sovacool and Brown 2010, Baiocchi *et al* 2015, Moran *et al* 2018). (Line 240-242)

18. Line 263: any useful inferences on behaviour change as a synthetic response that could find a place here?

We do not understand this comment and haven’t taken any actions.

19. Line 271: aren’t households, businesses etc. stakeholders rather than urban systems (energy, water, mobility etc.)

We have changed “systems” to “stakeholders”.

20. Lines 275 to 277: No relevant findings on city-region and system boundary definitions e.g. urban decarbonisation is linked to the decarbonisation of national electricity grids?

In our re-structuring of the article, this sentence was deleted.

21. Line 283: Impacts is a well-defined technical term in the climate literature, largely in the adaptation space. What kind of impacts are we speaking of here?

We have changed “impacts” to “emissions” for further clarity.

22. Lines 298 to 307: significant speculation in this paragraph. This is an extensive grey- and non- Web of Science/Scopus literature in this space, both from national (NDCs) and local governments and global urban networks (C-40, GCOM, ICLEI etc.). Without an analysis of this literature some of these extrapolations into the science-policy interface may be unfounded.

As discussed in the general comments above, we have re-written this section. It now includes two specific examples of reports (Atkins, ICLEI) using typologies of cases to structure policy recommendations:

Typologies are highly complementary with the case study method (Bennett and Elman 2006). They narrow the universe of cases that can be meaningfully compared, such that detailed case studies on specific types of cities can be generalised to a wider set of shared contexts. This holds the promise of stimulating learning within peer groups of cities, as was attempted in the Atkins report and in similar efforts led by ICLEI, where cases on infrastructure transitions have been assessed for post-industrial cities (ICLEI 2018b). But qualitative information from accessible and synthesized urban cases may also directly improve typology development itself. Linking together detailed case studies within a typology framework therefore allows a series of policy relevant questions to be asked, such as: what structural urban conditions shape energy demand? Within relevant peer groups, what intervention points exist for reducing energy demand in the short, medium and long term? Within relevant peer groups, which cities have enacted climate policies, were they successful, and why? (Line 326-336)

23. Line 308 and 309: IPCC AR5 (Ch 8, Wk Gp II) suggests an urban typology - a critical comment on that they be useful, if the authors consider it appropriate.

Unfortunately, we could not find a mentioning of typology/ies and/or archetypes in that IPCC chapter. There is a table on types of financing, but we could not find a specific recommendation to do typologies. Happy to correct that if pointed to the specific text.

24. Line 332: missing cooling days is a serious gap in this paper as noted earlier. Temperature elevation on top of urban heat island effects, is the core challenge in many regions.

As mentioned above, cooling days is now captured in our topic modelling, but not in the quantitative typology we use from Creutzig et al. 2015.

25. Line 337: a comment on non-C02 drivers from the literature may be useful esp. SLCPs/back carbon, N20 and O3

See next comment.

26. Line 359 to 370: It may be unfashionable, but is technical, economic and financial feasibility not emerging as concerns in the literature?

We feel these specific suggestions are out of scope for the research questions posed in our paper. Our aims are to (1) mapping the case literature in terms of cities, regions, topics (2) discuss the opportunities for systematic learning from these cases.

27. Line 373: Many of the measures identified in Fig 5 are supply-sided measures. Emphasis on demand-side measures an oversight or a systematic bias in the case study literature?

We are unsure whether the reviewer refers to the typology drivers, or the case study literature? To improve clarity here, we have expanded out the case literature as a heat map and have simply tabled the typology data instead of using bubbles (see figure above).

28. Lines 394 to 395: Would a sub-regionalisation of Asia help unbundle this finding? Clubbing oil-producing nations in the Gulf, with South Asia and East Asia may be a simplifying assumption but may mask the generalisability of the facts.

This is a fair point, but we have no space for further elaboration at a sub-regional level. Of course, our method for scoping case study evidence would allow this, even down to the city scale (as shown in Figure 4). The database of cases to be published with this article will allow such re-analysis.

29. Lines 402 to 403: Since both 1.5 to 2 C and adaptation have formed a subsidiary element in the narrative, would it be useful to comment on possible findings on these questions?

We have since removed adaptation from the analysis (see general comments above) and now cite the 1.5°C special report here.

30. Figure 1: not all national and regional capital cities in the world are 1-10 million in size. Unnecessary to conflate facts

We have changed the text to:

Figure 1 shows the spread of case study research across four different city sizes, from a small number of familiar ‘mega-cities’ (over 10m inhabitants), to large cities between 1-10m, and hundreds of medium (0.3-1m) and smaller (<0.3m) cities. The majority of research so far has focused on larger cities, with specific mega-cities receiving particular attention… (Line 60-63)

31. Figure 2: the salience of the data may be better represented if the bubble sizes were proportional to the share of global urban population (peach) and case studies (blue). Highlighting the world numbers would help with a visual benchmarking.

Thanks for this suggestion. We experimented with proportional bubble sizes and total numbers, but felt these additions distracted from the simple statement we wish to communicate: that small cities are under-researched relative to their peers. The share cases studies versus global urban population (in 2030) is provided in the SI text, Fig 3.

32. Fig 3: would be improved if a proper city population size scale were provided or a graphical indication of the proportion of

The size of the dots on each city are not scaled by population size, but number of case studies, as indicated in the caption text. We think that including population size would unnecessarily complicate this figure.

33. Fig 3: It is unclear what regionalisation the authors have used, the typical regionalisation schemed include: UN Population Division (?), IPCC, World Bank etc. Oceania is a rather mixed up category that mixes up Australia, New Zealand a number of SIDS from the Pacific. A footnote or citation may be in order. It may be useful, if the authors consider it appropriate to run an analysis between cities of fossil energy exporting cities vs. others

We use the UN Population Division regionalisation. The caption to Figure 2 now reads: “Population data and the United Nations Population Division regionalisation are from ref (UN DESA 2018), using agglomeration data where available.”

We will consider the fossil energy exporting categorisation for a further, more fine-grained publication, due to length constraints in this format.

34. Fig 4: the rationale for ordering the x and y axes is unclear. The visual patterns that emerge from a reclustering by increasing population size for example could be rather different. The mixing up of Adaptation with a range of mitigation options is confusing and possibly not appropriate, unless pulled out as a separate row?

We do not understand – the axes are categorical and no information should be inferred from their ordering. However, we clarify in the caption that the results in each column should be interpreted individually: “The colour scale is normalised by city (column), indicating the main topic focus of case study literature within each city.”

We address the adaptation point in the general response above.

35. In the supplementary data section, the sequence of graphs could be improved by moving: Fig 4 (total population) up front, followed by Fig 2 (growth rates), then Fig 5 (case study coverage), Figure 2 and 3.

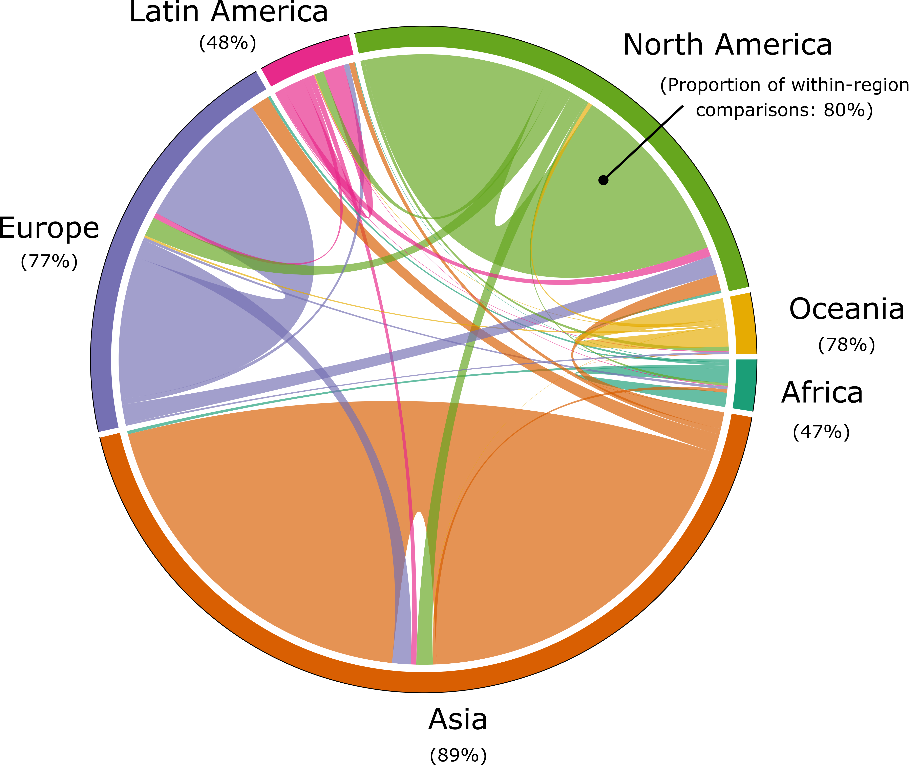
We are not sure why this new sequence improves understanding. We have chosen to keep them as before, so as not to mix up the sequential order of referencing in the main text.

36. Supplementary data Figure 3: is the global distribution of case studies vs. *population size*

As suggested, this now reads “Figure 2: The global distribution of urban case studies versus population size.”

37. Supplementary data Figure 6: is an important figure in terms of case comparisons but is unclear. For example, the referencing of cities within a region in the chord diagram e.g. North American cities referencing other North American cities requires extreme visual acuity to pull out. A better graphic scheme may be useful.

We hope the following representation does a better job, by placing a gap in the middle of within-region chords, and by clearly labelling these.



38. The relationship between Supplementary data Fig 6 and 7 is intriguing but unexplored, analytically and in the text.

Yes, this is interesting. We adapt the following text to highlight the limited number of internationally comparative studies with a high n:

Considering the total scope of the urban case study literature (4,051 studies), the subset that is comparative (702 studies), and internationally comparative across more than 2 cases (67 studies), is small. (Line 252-254)

We will make this data available with the publication, for interested users.

39. Aren’t Supplementary data Tables 3 and 4 using fraction and proportion interchangeably. Wouldn’t a simple % calculation be easier to read?

Yes, done.

Reviewer #2

This is an interesting paper and points to some clear gaps in the literature, and ways existing literature could be exploited to gain more information, and makes recommendations which are valid for furthering this field. The suggestions that case studies are done and used by researchers and urban policy makers, is a strong point, but could be made a little more compelling for both communities. Therefore, the main aspect which I would suggest strengthening is a question of how some of the recommendations which are made will contribute to or result in the stated changes at the science policy interface. I refer below to a few specific places where elaboration of this nature would strengthen the paper, as well as s few other small revisions, and would recommend publication with the suggested modifications.

**Primary comment:**   
Line 289- 297. How would review of cases which are confined to a geographically narrow scope, with the aim to initiate policy learning within that scope by drawing on a wider set of study designs and connected topics directly engage municipal stakeholders, integrate prioritizing agendas and enhance the quality of urban science policy. I think missing here is a link to the motivation and mechanisms by which the policy makers in the associated municipalities will receiving the synthesized information from the review, will it be ‘translated’ for best use for policy makers or will reviewers be done and published in ways which serve municipal authorities more directly? What will be motivations for municipalities to engage in the synthesized data. Also, a question of whether the synthesis of case studies being reviewed would be done in consultation with municipalities is important to explore and specify, as information can be lost of skewed in further processing.

We were inspired by the suggestion to reframe this piece around the science-policy interface and undertook a more fundamental re-structuring. The first half of the paper remains as before (i.e. section 1 & 2 on the geographic & topic distribution of case studies). The second half now begins with a section titled “Three ways to learn from case study evidence”, motivated by the following text:

Mapping out the literature provides a starting point for in-depth synthesis. Here we consider the various ways to best make use of the growing body of case studies: as individual studies, in comparative settings, and within reviews and assessments. Our review is guided by two questions: what are the different opportunities for generalisation in each approach? And how are cases actually used by urban stakeholders, researchers and policy makers? (Line 184-188)

We find this to be a useful entry point to various discussions. For instance, the conduct of individual cases has intrinsic value for learning among the participants involved. In addition, we see that science-policy groups (c40, ICLEI, UCCRN) are already undertaking comparative research and assessments of cases, closely based on their users’ needs. We then follow this with a final section on the complementary use of typologies and cases, discussed in the next response.

One consequence of this is that the quoted text to this comment was cut, to make space for these more concrete examples of how cases can be learned from and used in science-policy.

We have no specific suggestions as to how such information should be prepared and translated for individual municipalities. We think all the common routes for science-policy exchange should be left open – stakeholder engagement prior/during and after research; dissemination of results through reports and briefing notes; traditional media – but if conducting assessments of the literature, we emphasise that these should be systematic (i.e. transparent, reproducible) in order to secure their credibility:

Systematic evidence synthesis methods – those that deploy transparent and reproducible procedures for literature selection, quality assessment and synthesis – are the gold standard for generating a robust evidence base for policy (Berrang-Ford *et al* 2015, Sorrell 2007, Haddaway and Macura 2018). (Line 289-291)

We also now further emphasise the value of mapping out these cases for various communities in the conclusion:

But despite many recent commentaries highlighting the manifold challenges (and opportunities) of an urban focus in climate mitigation (Acuto *et al* 2018, Seto *et al* 2017, McPhearson *et al* 2016, Solecki *et al* 2018, Ürge-Vorsatz *et al* 2018, Nagendra *et al* 2018), there has been little reflection on how to generate insights from cases using additional layers of mapping, comparison and synthesis. Dedicated efforts in these areas, integrating cases in a wider project of learning, will generate substantial value for scientific assessment and urban practitioners of climate change mitigation.

Urban climate assessments have suffered from the anecdotal use of cases and a poor overview of the available evidence base. We identify 4,051 studies to date and make them available for systematic assessments in the future. Since we captured only ‘climate mitigation’ articles in a broad brush approach and limited our efforts to the Web of Science and Scopus as two of the largest scientific publication databases, this is still an underestimate that does not reveal the full extent of sector- and country-specific research (Lamb *et al* 2018). Manually scoping this number of studies to gain an overview of the field would be difficult and highly time consuming. But with computer assisted methods, including scientometrics and computational linguistics, a comprehensive evidence map can be developed more quickly and subsequently built upon as a community effort. (Line 387-401)

**Line 404-408**  
Why does this method result in more bottom up learning between cities? How does this method (above others) facilitate this and what would direct results/benefits be to both researchers and municipalities who are looking to learn from the reviews supplemented by city topology classification?

These are important issues. The basic premise is that maps of the literature useful and in demand from stakeholders and decision makers, who often ask ‘open framed’ questions like: How much evidence is there on X? What policies and measures have been studied? (James *et al* 2016). In the prior sections we have presented such a map (albeit at a highly aggregated level) and aimed to discuss its two potential purposes: to identify priorities for further research, and to identify opportunities for learning and synthesis.

On both counts, we claim in this final (now edited and more concise) section that such evidence maps can be made even more policy-relevant by linking them with urban typologies (What policies and measures have been studied, *in what contexts?).* We find that only specific types of cities have been studied. We also suggest that reviews of policies and measures could be usefully aggregated by typology, as is a common practice in the grey literature:

Urban typologies – taxonomies of city types – are powerful tools for comparing and aggregating knowledge on cities while preserving elements of context. Here we claim that typologies partnered with case studies could enhance learning on urban climate solutions, on the basis that similar structural conditions (i.e. captured through typologies of urban form, socio-economic development or types of governance) imply similar solution spaces. The Atkins *Future Proofing Cities* report pioneered such an approach by grouping 129 cities into 5 different types (based on qualitative criteria), linking these to more than 100 different policy options (Atkins 2013). In the academic literature typologies have proliferated, using newly compiled databases, as well as spatial, crowdsourced and ‘big’ data (Creutzig *et al* n.d., 2015, Baiocchi *et al* 2015, Louf and Barthelemy 2014). However, as yet there have been no systematic attempts to link typologies of cities to case study knowledge, and on this basis aggregate tailored policy advice.

Typologies are highly complementary with the case study method (Bennett and Elman 2006). They narrow the universe of cases that can be meaningfully compared, such that detailed case studies on specific types of cities can be generalised to a wider set of shared contexts. This holds the promise of stimulating learning within peer groups of cities, as was attempted in the Atkins report and in similar efforts led by ICLEI, where cases on infrastructure transitions have been assessed for post-industrial cities (ICLEI 2018b). But qualitative information from accessible and synthesized urban cases may also directly improve typology development itself. Linking together detailed case studies within a typology framework therefore allows a series of policy relevant questions to be asked, such as: what structural urban conditions shape energy demand? Within relevant peer groups, what intervention points exist for reducing energy demand in the short, medium and long term? Within relevant peer groups, which cities have enacted climate policies, were they successful, and why? (Line 316-336)

In summary, typologies enable learning between comparable cities, and hence overcome a crucial capacity challenge in evaluating city-specific policy options.

Additionally, this question of motivation and how this will be done could be incorporated more explicitly into the opening questions to be addressed in the paper, lines 32-36.

Thanks for this suggestion. We edited the introduction text to bring this out concisely:

A complementary strand of urban literature can be found in case studies: on individual city-scale reforms, as well as comparative studies across multiple urban settings (Lamb *et al* 2018). Using a variety of analytical tools, often qualitative, case studies can reveal the contextual and contingent nature of urban policy-making. Yet they remain under-represented in the scope of assessment literature on cities. The typical presentation of such evidence in assessments takes an anecdotal rather than analytical form – as dedicated boxed sections, as examples of particular phenomena, or within curated libraries of initiatives. A lack of rigorous literature selection procedures in assessments (and in reviews generally) means that potentially relevant cases remain overlooked (Minx *et al* 2017). Above all, the potential for coordination and mutual learning across epistemic communities, global and data-driven on the one hand, bottom-up and case-based on the other, has been neglected.

Urban case studies can add to our understanding of climate mitigation solutions, but an overview of the field is urgently needed. Which cities do we know about? What topics do we know about? What comparative and secondary analysis is there of cases? And how can generalizable knowledge be derived from urban cases? In this perspective we address these questions, using literature mapping approaches to discover and characterise the existing case studies. Overall we identify a rich and varied case literature, albeit one with regional and topic biases. To bring the policy relevance of cases to the forefront, we discuss the different opportunities they present for learning. We argue for more ambition in comparative studies, the need for systematic reviews of cases, and the potential to organise case study insights using quantitative typologies of cities. (Line 25-43)

**Additional comments:**

Table 1 is hard to read and I would restructure it so that the spacing between the different review titles is more clear. Currently the titles blend together.

Done, with light lines between the rows. It has also been moved to the SI Text.

Relating to figure 4, I think a stronger introduction should be given to what the figure is trying to convey in the text. I think it is a compelling figure, but it would be useful to state what can be gleaned and the main things within the figure which you would like the readers to be draw to. I think this could easily be done by extending the sentence on line 135, and adding a few additional sentence before going into the example on low-carbon transportation. Without an explanation this figure is a lot to take in.

We now introduce this paragraph with the following:

Where individual cities already have large literatures, particular topic trends can also be observed. Figure 4 presents the distribution of topics within the 10 most studied cities as a heat map. Low-carbon transportation case studies are well developed for Beijing, Shanghai and London, but scarcely researched in New York City, where cases have a greater focus on buildings. Emissions accounting dominates the case study work on top-tier cities in China (Beijing, Shanghai and Tianjin), while issues of water demand, urban ecology and climate governance receive less attention – at least in the mitigation focused literature we identify from Web of Science and Scopus. (Line 143-149)

We discussed it with a non-project colleague, who found this text sufficient, but we would be happy to take any suggestions.

Relevant to the explanation on lines 184-187, I think that it would be useful to have figure SI6 in the main text as this builds a bit of a foundation of the recommendations which are subsequently made, and the visual representation in the figure is clearer than the explanation currently given in the text.

The main problem here is the limited number of figures and items allowed in a perspective format. We also received feedback from Reviewer #1, who found this figure less clear. Therefore we have chosen to keep it in the SI Text.

We edited the text slightly, for clarity (and again checked it with a colleague not familiar with the project):

Inter-regional comparisons are relatively rare. Figure 7 in the SI text visualises the pairwise correlations of cities within abstracts, aggregating by region. Asian cities tend to be compared to other Asian cities, European cities to European cities, and likewise in North America. Studies on Latin American, African and Oceanian cities, on the other hand, are more frequently paired with cities in other regions, but are fewer in number. Considering the total scope of the urban case study literature (4,051 studies), the subset that is comparative (702 studies), and internationally comparative across more than 2 cases (67 studies), is small. (Line 248-254)

Is line 188 again referring to SI figure 6, not 5?

It was referring to 5, but we changed the order to make this sequential and less confusing in the main text.  
  
Relating to figure 5, there should be some inherent uncertainty in the size of the circles in the first box, on quantitative topology. To clarify the ‘scope’ of each circle, I think it could be useful to provide bounds for each circle size, (i.e. what is the highest amount of energy use which classifies Ahmedabad as low energy using cities, indicate by the point like circle in the energy use row. If there is not a strict range, then this gives me the indication that these circle represent and average of energy used by cities in a given typology, and therefore uncertainty in these values should be noted or illustrated in some way. Also relating to figure 5, in the text, line 329 ‘Figure 5’ is in bold, this should be corrected to normal text.

This is a good point. We just changed the circles to numbers and refer to uncertainties in the caption:

The typology and its data are from ref 5 (see also for uncertainty ranges); case study literature and topic analysis from this study. (Line 361-362)

Overall I think that this paper is well organised and provides an interesting perspective, which is not currently represented in the literature.  
  
  
Reviewer #3  
  
This paper opens up an important question but it does not provide a convincing answer. Some of the answers are very true (that the scholarship on cities and climate change focuses on large cities and particularly in the North) but these facts have been known for very long. It would be good to examine if measures have been proposed in, let’s say, the last decade and whether this pattern of research could be changed. Or whether we can identify empirical evidence of the reasons for these biases. Unfortunately, the authors do not seem to find this an important question, and they assume that calling for more research in smaller cities is sufficient to change current science drivers and orientations.

Regarding the novelty of this research, we provide 3 contributions. (1) We substantiate claims of bias with a comprehensive sample of urban case studies – the most comprehensive to date that we are currently aware of. (2) We extend the analysis of these biases into topics. As far as we know this is entirely novel, and there is a lot to learn (see for example the huge literature on energy/emissions accounting in China; or the mere handful of studies on buildings and heating demand in cold, wealthy cities - Figure 5). (3) We provide a detailed discussion on learning from case studies in the second half of the paper, which is currently missing from recent commentaries and discussions on progressing the field of urban sustainability research. (4) We provide the obtained case study literature as a searchable database, to facilitate future projects and in-depth studies.

Regarding the past decade of publications, we now plot the growth of literature by region and IPCC assessment period in Figure 4 (SI Text) and refer to it in the manuscript:

Although we do observe an uptake in case studies on Asian cities since the IPCC 5th Assessment Report, there has only a very limited expansion of cases on African and Latin American cities in recent years (SI Text Fig 4). (Line 95-97)

Regarding the reasons for these biases, in the original manuscript we made one suggestion on the topic bias in East Asian cities:

Scaling up the analysis from individual documents to groups of documents, we observe that emissions and energy accounting is a frequent subject of case study research situated in Asia (Figure 3), perhaps reflecting strong investments into engineering disciplines and education in China and South Korea (44% of all students in China graduate in science & engineering, compared with 16% in the US) (Gonzalez-Brambila *et al* 2016). (Line 136-142)

And we now add this text later in the manuscript, regarding general geographic biases in case study coverage:

But when it comes to selecting case study locations, practical concerns matter too. We cannot ascertain the motivations behind the case selections documented in this article, but their strong geographic and topic biases do suggest certain practical limitations. The dominance of North American and European universities in scientific publishing, as well as institutional funding biases, limited scientific capacities in the global South, and everyday travel limitations are all possible factors that confound the selection of cases on strong theoretical merit. (Line 204-209)

Strongly evidencing these claims is somewhat beyond the scope of this article, and potentially less interesting than the discussion on learning that we proceed to. Therefore we leave it this to further publications.

Some of the ideas of the paper are simply not right. I cannot recognise the characterisation of the urban studies field. The only references to major urban studies journals (such as Urban Studies and the International Journal of Urban Regional Research) are to the two very similar papers by Michael Storper and colleagues which represent a polemic view of the state of the field. These are polemic papers which seek to create debate and yet they are presented here as if they represent the state of the field.

We respond to the specific literature suggestions and concerns as they arise below. In general, our concern is not the urban studies field but the question of how to systematically map and learn from the aggregation of climate-mitigation focused case studies, as set out in the research questions (introduction):

Urban case studies can add to our understanding of climate mitigation solutions, but an overview of the field is urgently needed. Which cities do we know about? What topics do we know about? What comparative and secondary analysis is there of cases? And how can generalizable knowledge be derived from urban cases? (Line 35-38)

To our best understanding these questions are not (apart from the final one) guiding concerns of major urban studies journals. We have carefully gone through the article to stay focused on these questions and avoid creeping into characterisations of the broader urban field. In fact this led us to delete the sentence citing Storper. For the final question, we follow the reviewer’s specific suggestions, as discussed below.

Some of the more interesting conclusions such as the typology of cities are approaches which are used in the private sector such as Atkins Future Proofing Cities which already proposed a very similar typology six years ago (see references below). The two options proposed by the authors about future research (city-based and comparative) are already widely used. Moreover, these are not the only approaches! The diversity of the field is really important to create and develop innovation and sectoral approaches and comparison of the development of specific innovations and technologies across cities are really important to understand the transfer of ideas from location to location. The fact that many comparative studies look at similar policy approaches in different locations is ignored in this paper.

To address these concerns and bring the suggested diversity forward, we have re-organised the second half of the manuscript, structuring it around two sections: (1) Three ways to learn from case study research; and (2) Bridging urban typologies and case study evidence. The first of these is motivated by the following paragraph:

Mapping out the literature provides a starting point for in-depth synthesis. Here we consider the various ways to best make use of the growing body of case studies: as individual studies, in comparative settings, and within reviews and assessments. Our review is guided by two questions: what are the different opportunities for generalisation in each approach? And how are cases actually used by urban stakeholders, researchers and policy makers? (Line 184-188)

This allows us to address many of the reviewer’s concerns going forward, regarding the value of individual cases, the representation of comparative cases and recent advances in urban assessments.

The lack of originality of this paper relates with an inaccurate conceptualisation of the field of research on cities and climate change. There is a dearth of references to work in global environmental governance.

For example, I do not see any representation of studies that have looked systematically at climate change action in cities. A well-cited example that comes to mind is, for example, Broto, V.C. and Bulkeley, H., 2013. A survey of urban climate change experiments in 100 cities. Global environmental change, 23(1), pp.92-102. This kind of work is only superficially referred to, but it builds on local reporting of climate action and has had a strong impact on the UNFCCC approach to sub-national climate action (as demonstrated for example in the NAZCA platform for climate action).

Note that our effort covers more than 4000 cases in the literature. We cannot cite them all in this main manuscript, but provide the database as a resource to the community. While this is the most comprehensive map and analysis of urban cases we are aware of, we acknowledge in the manuscript that it will not be complete. Still, to respond to your comments and balance the manuscript’s text out, we begin with the following paragraphs in the ‘comparative cases’ section, citing these literatures:

A second common approach in the design of case studies is to implement a single methodology across multiple case locations. These comparative case studies have the advantage of enabling direct tests of generalisability, by identifying common patterns. For instance, evidence from a comparative study of 4 cities underscores the crucial need to restrict car traffic rather than rely on advancing public transit alone to decarbonize urban mobility (Creutzig *et al* 2012). Another study found that across 4 cities, urban leaders retain influence over the composition and magnitude of a large portion of upstream (consumption-based) greenhouse gas emissions (Pichler *et al* 2017).

Comparative research is also of deep interest to urban stakeholders, who may wish to understand how their peers commit to and overcome barriers in climate change mitigation. This is the basic premise behind urban cooperation groups such as C40 and ICLEI, which actively document and compare urban actions through case studies (C40 Cities Climate Leadership Group 2017, ICLEI 2018b, Nangini *et al* 2018); and action tracking initiatives such as the Global Climate Action NAZCA portal (UNFCCC 2018).

There is substantial comparative work analysing mitigation and adaptation plans across hundreds of cities (Reckien *et al* 2018, 2014, Castán Broto and Bulkeley 2013), and more detailed studies published as volumes (Bartlett and Satterthwaite 2016, Rosenzweig *et al* 2015). Much attention is also paid to comparing urban carbon footprints (Sovacool and Brown 2010, Baiocchi *et al* 2015, Moran *et al* 2018). Less known are the small-n comparative studies. (Line 228-242)

There is also a complete oversight over big data inspired research on cities, such as that impulse by the Yale Data-driven Research Initiative (<http://datadriven.yale.edu/> )

We do refer to big data inspired research, specifically in the typology section:

In the academic literature typologies have proliferated, using newly compiled databases, as well as spatial, crowdsourced and ‘big’ data (Creutzig *et al* n.d., 2015, Baiocchi *et al* 2015, Louf and Barthelemy 2014). (Line 322-325)

We are not sure, however, that it warrants specific attention given that our piece focuses on case studies.

Finally, there have been regionally-specific studies that the authors overlook. Many are focused on Europe or in the US, and yet, they provide an angle which is ignored here. Another well-known example is: Reckien, D., Flacke, J., Dawson, R.J., Heidrich, O., Olazabal, M., Foley, A., Hamann, J.P., Orru, H., Salvia, M., Hurtado, S.D.G. and Geneletti, D., 2014. Climate change response in Europe: what’s the reality? Analysis of adaptation and mitigation plans from 200 urban areas in 11 countries. Climatic change, 122(1-2), pp.331-340.

This study is cited in the above paragraphs on comparative research.

There are numerous sector-based studies some of which have been highlighted by the authors in their systematic review part, but they overlook those which do not do a systematic review, but rather use case studies to develop frameworks. There have been many example in the energy field in journals such as Environmental Innovation and Societal Transitions or Environmental Science and Policy. Overall, there seems to have been a problem with the use of keywords and the range of papers that the authors have unveiled, with less attention to well know studies which have had influence in policy circles. This underscores the limitations of big data analysis, in comparison with experience-based understandings of the field that reach better informed solutions with less data, as already explained by Bent Flyvberg in his analysis of the need for social research that takes seriously the principles of phronesis.

We cannot cite or give credit to the full range of studies found. Actually we received several comments on literature omissions (here on influential cases, from reviewer #1 on cooling, energy poverty, energy storage). This underlines exactly why the systematic mapping we conduct in this article is so necessary – in the end every reader will have their own judgement on the importance of different articles and streams of literature. We think the database to be provided with this article enables more in-depth reviews and assessments, where researchers and stakeholders judge these necessary (e.g. based on their experience of specific issues). More than this, our approach is transparent and can be easily reproduced from the methods. This means that further refinement can take place and on-going developments in the literature can be tracked.

Regarding the value of experience-based understandings, we thank the reviewer for this suggestion, as this led us to commit substantial (and we think useful) edits. We now dedicate one of the “three ways to learn from case study research” sections to the merit of conducting individual cases (Line 190-226). We reference Flyvberg and others, making an argument for the generalisability of cases when they are selected on theoretical merit, while also discussing the intrinsic value of cases for learning and phronesis. This is entirely complementary with our approach, which catalogues such cases and makes them available to the widest possible audience.

Also this paper ignores policy studies from UN-Habitat, UCCRN, OECD, the EU etc or trade reports such as ATKINS/DFID Future Cities Report, which contained one of the first attempts to compare systematically cities and develop city typologies using big data or classic work from ICLEI.

We now refer to major reviews and assessments in the third of our “three ways to learn from case studies” sections:

A third opportunity for learning from cases is through secondary analysis in literature reviews and assessments. Concise, policy-relevant advice based on the literature is in high demand from stakeholders. Such reviews and assessments can be particularly effective when stakeholders are involved at the design stage to communicate their knowledge needs (C40 Cities Climate Leadership Group 2017, ICLEI 2018a). In the area of urban climate mitigation, there have been several assessments so far and growing activity. The Global Energy Assessment provided one of the first dedicated reviews on urban energy systems and was followed by an urban chapter in the IPCC 5th Assessment Report (AR5) (Seto C. *et al* 2014, Grubler *et al* 2012) and two assessments by the Urban Climate Change Research Network (UCCRN) (Rosenzweig *et al* 2015, UCCRN 2011). In the upcoming IPCC AR6, urban case studies will be a key evidence base for both the urban and demand chapters in Working Group III (Mitigation), not to mention the regional chapters under Working Group II (Impacts & Adaptation). (Line 263-272)

The Atkins report is cited in the typology section, where it is of particular interest:

Urban typologies – taxonomies of city types – are powerful tools for comparing and aggregating knowledge on cities while preserving elements of context. Here we claim that typologies partnered with case studies could enhance learning on urban climate solutions, on the basis that similar structural conditions (i.e. captured through typologies of urban form, socio-economic development or types of governance) imply similar solution spaces. The Atkins *Future Proofing Cities* report pioneered such an approach by grouping 129 cities into 5 different types (based on qualitative criteria), linking these to more than 100 different policy options (Atkins 2013). In the academic literature typologies have proliferated, using newly compiled databases, as well as spatial, crowdsourced and ‘big’ data (Creutzig *et al* n.d., 2015, Baiocchi *et al* 2015, Louf and Barthelemy 2014). However, as yet there have been no systematic attempts to link typologies of cities to case study knowledge, and on this basis aggregate tailored policy advice. (Line 316-325)

The difference is that these used qualitative coding of case studies, but the insights maybe the same or in more depth than the ones presented here.

We are not sure why this article should be compared to major assessments and reports. Beyond the obvious constraint of space in a perspective article, detailed surveys of specific topics are beyond the scope of our research questions, as we set out in our responses above.

There are inaccuracies and inadequate statements in the text. Some examples:   
“Yet in the assessment of urban climate solutions, case studies are an unexploited resource.”- this is incorrect. In fact case studies are used very often, but in isolation, around ideas of best practice. It is more precise to say that a systematic analysis of 2500+ cases has not been done.

This was a typo – it should have read “…case studies are an underexploited resource”. We have since rephrased it to “Yet in the assessment of urban climate solutions, much more could be learned from case studies.” This statement has less commitment, but we think there is ample evidence only in the fact that no-one so far has attempted to map out the current cases. For more specific critiques that also corroborate the reviewer’s suggestion, we have edited the section on assessments:

Although the UCCRN generated many new cases, they are simply placed throughout the report in boxed sections. In other words, the case content is brought to the forefront, but isn’t synthesized across topics or locations. This pattern is repeated in the recent IPCC Special Report on 1.5°C, which dedicates multiple pages to urban case studies in boxed sections (e.g. boxes 4.1, 4.4, 4.5 and 4.9). A prominent example is Box 4.5 on congestion charges, which describes three successful cases of implementation (Singapore, Stockholm and London), but neither compares nor synthesizes these examples with cases on failed proposals, such as in New York City and Edinburgh (Gaunt *et al* 2007, Schaller 2010). As a result, little is learned about the circumstances for success or failure, despite a large underlying literature. (Line 280-287)

“There are strong size, regional and topic biases in case study coverage, and sparse efforts to conduct comparative analysis and systematic reviews.”- this is true, and it also relates to the structures of knowledge production; similar critiques are relevant in the climate change field more broadly.

We have taken no specific actions on this comment.

“For instance, urbanists often state that cities share common structural (political, economic, or geographic) characteristics that drive urban phenomena, leading to differing path dependencies in energy consumption.”- which urbanists? Actually most urbanists state exactly the opposite: that cities follow context-specific trajectories and that one city’s experience cannot be reduced to those of other cities. The statement on comparative urban studies made before overlooks the vibrant debate on comparison that is taking place within urban studies. Comparison in urban studies is not about finding commonalities (which is often considered an outdated approach) but about relating particularities with general insights. The classic article from Bent Flyvberg on the use of case studies is a good place to start. Unfortunately this is not appropriate, because it shows an outdated view of the urban studies field and it is this outdated view what is used to mount a criticism of it.

We agree this was poorly phrased – it has been deleted. Please refer to the re-structured section “Three ways to learn from case study evidence”, particularly the first of these (individual cases) where we discuss insights from Flyvberg and others.

“These consist of a wide spread of quantitative, qualitative and mixed review approaches that are well-documented in the health sciences literature”- but there is no recognition of the multiple studies in this area in environmental governance which is a lot of more relevant. Unfortunately, the authors do not demonstrate they know these.

We cite the governance studies suggested above in the comparative section. We are happy to cite further systematic reviews here if there are any specific suggestions (please note: systematic reviews of the peer reviewed literature, not reviews of policy plans, which count as primary research). We otherwise stand by our claim that very few such studies exist in this field of work.

The typology approach is something used, particularly among consultants. See the report of Atkins on Futrue Proofing Cities for very similar typologies of cities: <https://www.atkinsglobal.com/en-gb/group/sectors-and-services/services/future-proofing-cities>

Thanks for this suggestion. We now refer to it explicitly in the typology section:

Urban typologies – taxonomies of city types – are powerful tools for comparing and aggregating knowledge on cities while preserving elements of context. Here we claim that typologies partnered with case studies could enhance learning on urban climate solutions, on the basis that similar structural conditions (i.e. captured through typologies of urban form, socio-economic development or types of governance) imply similar solution spaces. The Atkins *Future Proofing Cities* report pioneered such an approach by grouping 129 cities into 5 different types (based on qualitative criteria), linking these to more than 100 different policy options (Atkins 2013). In the academic literature typologies have proliferated, using newly compiled databases, as well as spatial, crowdsourced and ‘big’ data (Creutzig *et al* n.d., 2015, Baiocchi *et al* 2015, Louf and Barthelemy 2014). However, as yet there have been no systematic attempts to link typologies of cities to case study knowledge, and on this basis aggregate tailored policy advice. (Line 316-325)

“Implicit in much sustainability research is an asymmetry of knowledge exchange: cases in the global North are relevant for the South, but not vice-versa” – this is simply incorrect. In fact, decades of sustainability research show that cases in the North are often not relevant for the South. Moreover, innovations from the South also spread to the North- and the cases are so many that I do not even need to explain them. You would need to make some sort of analysis to see what is more common but as things stands, this is not an acceptable assumption.

This phrase has been deleted. We are also aware of the recent and relevant article by Nagendra et al 2018 and cite it in our conclusion.

My view is that what drives research in the North are funding patterns and researchers’ interests and constraints. As I say above, there is not an analysis of that.

We now have a paragraph on this, but again consider any further analysis in this direction out of scope:

But when it comes to selecting case study locations, practical concerns matter too. We cannot ascertain the motivations behind the case selections documented in this article, but their strong geographic and topic biases do suggest certain practical limitations. The dominance of North American and European universities in scientific publishing, as well as institutional funding biases, limited scientific capacities in the global South, and everyday travel limitations are all possible factors that confound the selection of cases on strong theoretical merit. (Line 204-209)

If you use UNDESA data you need to use the 2018 revision, rather than the 2014 one.

Done.