Dear Dr. Toth,

Thank you for the opportunity to submit a revised manuscript. We appreciate all the comments we have received and have made a great effort to incorporate each of them into our revision.

As suggested, the manuscript has been thoroughly reconsidered and amended to provide better clarity on both the method and data used in this analysis. In particular, more details have been added with respect to the method and data, as follows:

- The Methods section has been expanded to include more specific information on the SOMs method (with equations), as have the error measures and Davies-Bouldin index.
- Data sources have been appropriately cited, and a list of previous papers produced using the same data has been provided. The timeline of the study, interpretation of the placement of the cities on the maps, and the necessary assumption of flood protection level have been further discussed.

All points recommended by the reviewers and editor have been systematically addressed and responded to. Please see the 'response to reviewer' documents submitted concurrently for a detailed response to each query about the method and/or data. We hope the information added to the Data and Methods sections, as well as further detail in the Introduction and Conclusion, will remove any further confusion regarding the pertinence of the method and validity of the data in this study.

The vocabulary used in the manuscript has been checked and updated. The term 'spatiotemporal' has been removed, and other terms mentioned by reviewer 1 have also been revised. The references have been updated as suggested.

A few paragraphs have been added to the end of the introduction to explicitly describe what types of 'patterns' and 'clusters' are expected to be found in this study and why they are of interest. This is also reiterated at the beginning of the Discussion. The conclusion has been expanded to include more discussion of shortcomings of this analysis, and areas that could be improved in future studies.

We hope you find this revised manuscript much improved.

Kind regards,

Stephanie Clark

Reviewer 1:

We thank you for your time and effort in providing comments on this paper.

- Specific changes that have been made to the revised manuscript based on these comments are described in the dot points following each comment.

The paper explores the global dataset on fluvial flood risk prepared and discussed by Philip Ward and co-authors on city level. This is potentially a good idea; although the data has been explored in a series of papers by these authors there are still room for further studies. In its scope the current paper is rather close to Winsemius et al (2016); the current paper focuses on cities while the original paper focuses on river catchments.

As mentioned, Winsemius et al (2016) analyses the data set with respect to river basin and income level. Our analysis is based on city-level flood impacts. Whilst similar, these studies are not the same, and the insights revealed in this paper do not overlap with those provided in Winsemius et al (2016).

- This is already stated in the Introduction, paragraph 5.

There are however a number of important shortcommings that should be addressed before I potentially can recommend publication.

First of all I would like to see a discussion of what new knowledge the authors think they can gain by considering cities rather than river catchments. At least the findings of the current paper should be compared to Winsemius et al (2016).

The comparison of cities allows an understanding of conditions that can be expected on a city-level, and identification of cities facing similar circumstances. As climate change, development, and urban administrations transcend river basin boundaries, it is sensible to investigate impacts and determine potential mitigation strategies at the city level (as well of course as basin level, but that has been covered in Winsemius et al (2016)).

This is discussed in the conclusion of this paper, p19 lines 12-26, which is reproduced here:

> 'This study adds to the understanding of natural hazards in a global context, which is an important aspect of regional disaster risk management due to the dependency of local situations on global processes (UNISDR, 2015). The complex nonlinear socio-environmental relationships make it difficult to foresee local responses to global changes (UNISDR, 2015), and therefore this study focuses on risk communication (the process between risk perception and adaptation planning (Cardona et al., 2012)) to provide a visual analysis of the global patterns of evolving flood impacts, socioeconomic development and climate change, and the local city-level consequences of these changes.
>
> Cities have major implications for climate change mitigation and adaptation (Revi et al., 2014). Unplanned development and urban migration are increasing vulnerabilities to natural hazards (UNEP, 2016) and land cover change and greenhouse gas emissions are intensifying urban hydrology. Understanding the relationship between flood impacts and social vulnerability is a necessary step for prioritizing flood mitigation and prevention strategies (Doocy et al., 2013). Whether the main driver of increased urban flood impacts is development or climate change,

cities will benefit from development restrictions and planning standards for urban expansion, sustainable land development, management of population distribution and migration, and early warning systems and preparedness (Revi et al., 2014; UN-DESA, 2014; Doocy et al., 2013).'

- A discussion of the interest of flood impacts in cities already exists in the conclusion, and this point has also been added to the final paragraph of the introduction.

Next, the introduction is rather long and gives a thorough introduction to flooding. Unfortunately the authors does not distinguish between the different types of flooding that can occur, i.e. sea surges, groundwater induced flooding, pluvial flooding and, fluvial flooding.

It is stated on p4, lines 13-14, that this paper only considers fluvial flooding:

'This data is solely related to the influence of fluvial flooding on metropolitan areas, and does not include coastal or pluvial flooding.'

If this statement would be better placed in the introduction, it can be moved, however we are conscious that the introduction is already 'rather long'.

- A sentence has been added to the Introduction, in paragraph 4, to reiterate that only fluvial flooding is considered. The phrase 'river flooding' is used seven times in the abstract and introduction.

The paper is based on a specific data set that only considers fluvial flooding. Hence references to authors that specifically refer to pluvial flooding should be removed, e.g. Willems et al (2012).

We refer to Willems a few times, for example in the following statement (p2 lines 36-37):

'Increases in rainfall intensity at urban hydrology scales of up to 60% are anticipated by 2100 (Willems, 2012).'

Though Willems may be generally interested in pluvial flooding, any references made here concern Willems' statements on increases in rainfall, not flooding.

Perhaps differences in fluvial flood risk between cities can be explained by different exposures to other types of flood risk? If the authors do not wish to enter such discussions they should stick to considering only one type of flooding.

The comment 'if the authors do not wish to enter such discussions they should stick to considering only one type of flooding' is confusing for three reasons.

First, we have made it clear we are only considering one type of flooding - fluvial flooding (see previous comment).

Second, we have 'entered such discussions'. The exposures of various cities to other types of flood risk is discussed in the 'discussion' section of the paper (eg. p18 lines 22-28, as reproduced here). Unfortunately, the reviewer has stated (below) that they have not read this part of the paper.

'Though this study does not consider coastal flooding, it may be noted that due to their locations near river mouths, many of the cities in the lower left of the map that are projected to experience high increases in impacts from river flooding are also at risk of increased coastal flooding from

intensified storms and sea level rise due to climate change. Mumbai, Guangzhou, Shanghai, Ho Chi Minh City, Kolkata, Bangkok, and Dhaka are 7 of the top 14 cities (out of 136) ranked by current population exposure to coastal flooding. These same cities also comprise the top 7 cities (in this order: Kolkata, Dhaka, Mumbai, Guangzhou, Ho Chi Minh City, Shanghai, Bangkok) ranked by future (2070) estimated population exposed to coastal flooding (UNEP, 2016; Nicholls et al., 2008).'

Third, the reviewer has implied in a previous comment that it is unfortunate different types of flooding are not distinguished: 'Unfortunately the authors does not distinguish between the different types of flooding that can occur, i.e. sea surges, groundwater induced flooding, pluvial flooding and, fluvial flooding.'

Therefore, it is unclear what the reviewer is expecting to result from this comment.

Next, the literature on whether flood risk is stationary or has an increasing trend is quite abundant. The findings differ, primarily as a function of the framing, i.e. if the models include corrections for changes in socio-economic development, vulnerability, etc. The author uses the terms 'impacts', 'risks' and 'material damage' more or less inter-changeably throughout the paper. When referring to IDSR and other recognized frameworks I would expect more stringent use of the terminology and more transparent explanations and assumptions.

- In paragraph 3 of the Data section, we have indicated that one variable of the data set will be referred to in this paper as 'damages'. It has been defined here as 'urban property damages costs (in US dollars)', and this term has only been used for this.
- Use of the term 'risk' (dictionary definition: *To expose (someone or something valued) to danger, harm, or loss. The possibility that something unpleasant or unwelcome will happen*.) has been checked within the manuscript and we believe it has been used adequately in each instance. If there is a specific occurrence we should reconsider, please list it explicitly.
- Use of the term 'impacts' (dictionary definition: *A marked effect or influence. Have a strong effect on someone or something*.) has been checked within the manuscript. This term has been used in the manuscript to describe the influence that flooding will have on the population and urban property. In some instances, it could perhaps be replaced with 'risk' but in these cases we believe the word 'impact' also fits. If there are specific occurances of use the reviewer would prefer changed, please list them as we are happy to attempt for the greatest clarity possible.

This includes a description of what the scenario for development is between 2010 and 2030 and the rationale for choosing the approximately 100 cities in Table 1, that are then later reduced to 80 cities (without mentioning which of the cities are excluded).

The development scenarios are discussed in the Data section (p4 lines 18-28) with references to the models that were used to produce the data and accompanying documentation. The specific scenario is not described explicitly, however enough detail is provided for an interested reader to find the specific details:

'Three separate scenarios of climate change and socioeconomic development (optimistic, business-as-usual, and pessimistic) are given in Aqueduct, and in this study we use data from the business-as-usual case for our future flood impact scenario.'

The rationale for choosing the cities is explicitly given on p4, lines 6-8:

> 'The selection of cities used here is based on a list provided by the Lincoln Institute of Land Policy's Atlas of Urban Expansion (Angel et al., 2010, website 1), spanning all continents except Antarctica, encompassing four economic levels and four population levels.'

The reason that some of these cities have been excluded is explicitly state on p5, lines 5-6:

> 'Cities with no flood impacts in both 2010 and 2030 were removed (22 cities), though cities with no flood impacts in 2010 but with flood impacts in 2030 have been kept in the study.'

The reviewer is wondering why there is no mention of which cities were excluded - it was decided that a list of cities that are NOT included in this study would not be of interest to the readers hence they have not been listed.

- The reference to Table 1 in the text has been moved until after the discussion as to why some cities were not included. Hopefully this will make it obvious that only the cities in Table 1 are included in this study.

However, my most important concern is that I cannot see what the authors are doing with the data. The closest to an explanation is that the authors state that the calculations are carried out in Matlab using a modified version of the package SOM. Modified in which way?

The entire Methods section is describing what we have done with the data. Clearly the use of the SOM is causing some confusion, so this will be elaborated on.

- The description of the SOM method has been expanded to include equations and more detailed explanations.

Why are the data transformed the way they are if the SOM approach is particularly good in dealing with non-linear relationships,

The data is log transformed, following Agarwal and Skupin's (2008) recommendation that highly skewed variable distributions may benefit from log transformation before use in the SOM. *[Agarwal, P. and A. Skupin (2008). Self-organising maps: Applications in geographic information science, John Wiley & Sons.]* This information and reference will be added to the methods section.

- This information and reference have been added to the end of the data section.

how do you treat taking logarithm of the value of zero,

The treatment of the logarithm of values between -1 and 1 is carefully described on p15, lines 12-13.

what are the definitions of QE, TE, DRR, and the Davies-Bouldin Index,

QE, TE, and DRR are defined on p7, lines 16-19. A reference is provided for the Davies-Bouldin Index with the information that it is used to determine the number of clusters (p7, line 31).

- These definitions have been expanded with further explanations and equations.

and how is a U-matrix visually verified? Without this information the reader will have to do a complete reanalysis to (perhaps) get the same results as the authors.

Further SOM references have been provided for readers who are unfamiliar with the method, though it has been summarised in the Methods section. In an effort to avoid repetition of previously published material, the SOM method has not been reproduced in detail. As the papers and books outlining the SOM method have already been cited over 40,000 times, it was not felt that this method needed further repetition. With any paper, it is hard to know how much background knowledge to expect the readers to possess, and therefore we provide direction to further reading for those unfamiliar with certain aspects of the background.

- A paragraph has been added further describing the U-matrix.

Since I expect a thorough revision of the method section before perhaps resubmitting the paper I have not read the results and discussions sections in detail.

We are disappointed to hear this. I expect the reviewer would have found some answers to their questions had they read the paper.

More detailed comments:

I disagree with the use of the term 'spatio-temporal'. It usually denotes something where there is explicit reference to a spatial dimension, in the current example perhaps the physical distance between the cities. I think the authors should find another term to describe the characteristics of their data.

The term 'spatiotemporal' is commonly used to refer to data in which there is a cross-sectional structure to the data as well as a temporal one. In this case these are: 1) the state of flooding at each timestep, 2) and the timesteps themselves. The spatial dimension does not always refer to geographic distance.

- The term spatiotemporal no longer appears in the manuscript.

The data on the figures cannot be read in the pdf-version of the paper.

- Figures 1b), 2b), 3a) and 3b) have been reproduced with larger text, and the background colours have been lessened in intensity to make the text appear more clearly. We have tried a variety of text sizes on Figures 1a) and 2a), and believe the current text size makes it more clear precisely where each city is located on the map than a larger text size would. A note has been added to the caption referring the reader to the online version which can be zoomed in on. Figure b) maps with larger text could also be referred to if there is an issue reading the a) maps. The following is a screen shot of the pdf version of 1a), first at actual size, and then zoomed in. We believe both versions are readable.

a)

Amount of population affected
Shanghai
Kolkata
Cairo
Seoul
Paris
Los Angeles
Philadelphia
Bangkok
Buenos Aires
Tokyo
Cincinnati
Mexico City
Chicago
Houston
Mumbai
Ahvaz
Dhaka
Sao Paulo
Wien
Manila
London
Minneapolis
Palembang
Tijuana
Pittsburgh
Modesto
Guangzhou
Hyderabad
Pune
Kanpur
Moscow
Springfield
Casablanca
Santiago
Yulin
Astrakhan
Ho Chi Minh City
Bamako
Bandung
Kuala Lumpur
Warsaw
Rajshahi
Tacoma
Anqing
Vijayawada
Tel Aviv
Kigali
Beijing
Yiyang
Zhengzhou
Jequie
Le Mans
Yerevan
Kuwait City
Shimkent
Leshan
Accra
Chinju
Algiers
Castellon
Ulan Bator
Port Sudan
Budapest
Madrid
Oktyabrsky
Marrakech
Gorgan
Changzhi
Leipzig
Aswan
Fukuoka
Jalna
Malatya
Montevideo
Alexandria
Ouagadougou
Harare
Sheffield
Caracas
Sanaa
Ibadan
Pusan
Guadalajara
Istanbul
Baku
Zugdidi
Sydney
Tebessa
Thessaloniki
Valledupar
Ribeirao Preto
Johannesburg
Ansan
Teheran
St. Catharines
Songkhla
Kampala

I cannot follow the discussion on the cluster. Perhaps it is just me not being able to see the same patterns as the authors.

- Hopefully the lack of clarity on 'clusters' has been resolved by the addition of information regarding the clustering with the SOM that has been added to the Method section, the Discussion and in other locations throughout the manuscript. Please see the 'response to reviewer 2' for more info on this.

The list of references should be improved. Just from browsing it I can see dubious referencing to e.g. IPCC (2014) (Use author names) and Willems (2012) (several authors), and Kohonen (2001) (incomplete reference). There are more errors than the ones I have pointed out.

- The references have been checked and updated where necessary. The following, and others, have been updated as suggested:
    - Pachauri, RK, Allen, MR, Barros, VR, Broome, J, Cramer, W, Christ, R, … & Dubash, NK. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (p. 151). IPCC.

- Willems, P, Olsson, J, Arnbjerg-Nielsen, K, Beecham, S, Pathirana, A, Gregersen, IB, & Madsen, H. (Eds.). (2012). Impacts of climate change on rainfall extremes and urban drainage systems. IWA publishing.
- Kohonen, T. (2001). Self-organizing maps, Volume 30 of Series in Information Sciences.

Reviewer 2:

Thank you very much for your time in reviewing this paper, Dr Larsson. We appreciate all comments and will address each one as thoroughly as possible to produce an improved paper. It appears your main concern rests with the data source used in the paper. We understand that more information is needed in the referencing of this data set, and hope that the information provided below will allay these concerns.

- Specific changes that have been made to the revised manuscript based on these comments are described in the dot points following each comment.

General comments:

The topic of this paper is clearly both interesting and of great importance. The paper is well written and has a clear structure.

However, it seems to me that the paper does not really fulfil the promises implicit in the title. Which are the patterns in flood impacts on cities that are revealed in the paper?

The 'patterns' are the key characteristics extracted for each cluster of cities sharing similar baseline and projected flood conditions. Each city in each cluster is matched to the 'pattern' represented by the cluster centroid. As described on p6 lines 1-3, the values of the prototype vectors (map nodes) come to represent the prevalent patterns in the data after they have self-organized amongst the data items. For example, cities in cluster 10 in the upper right of Fig.3 share the 'pattern' of low baseline flooding impacts, with small increases in population impacts projected primarily due to climate change, and increases in material damages projected due to both climate change and development. For Fig3, the text on p13 and 15, and Fig 4 on p14 interpret these patterns in detail. The Method section (p6) may benefit from expansion, as it appears more detail is needed on the process of pattern extraction by the SOM.

- A further description of the 'patterns' expected to be revealed with this analysis has been added to the final paragraph of the introduction. In the Methods section, the terms 'pattern' and 'cluster' have been defined in terms of the SOM method (which has been expanded upon), as well as the use of the terms with respect to this study. A paragraph has been added to the beginning of the Discussion section to again describe how the terms 'pattern' and 'cluster' relate to the results of this analysis. Hopefully this will clarify any confusion regarding references to 'patterns'.

There are some problems with the method used, and there is a problem with the general approach. The latter problem is related to the presumption that it is possible to draw conclusions about individual cities based on a global model. It is also doubtful whether any meaningful patterns can emerge from such a rather superficial study. The method issues are covered below.

It is not within the scope of our study to assess whether the model used to produce the data is comprehensive, but rather to provide a visualisation of the available data. The data set as published is

immense and difficult to comprehend without a reduction and visualisation, which we are providing. Any conclusions about individual cities are based solely on the data items presented, and comparisons are based on similarities between data items, be them cities or merely numeric vectors.

- General issues with the method will hopefully be cleared up as we have added a significant amount of explanation and equations to the Methods section to better describe the SOM for readers unfamiliar with this method.

Specific comments:

(p.3, lines41-44) The objective is given only in rather general terms. It would have helped the reader to get some more precise information about which types of "global patterns and relationships" that is meant to emerge from the study.

This sentence had previously been more long-winded, but was reduced to this phrase for brevity with the knowledge that the 'patterns and relationships' would be explained in detail throughout the paper. This sentence can certainly be expanded to detail the expected patterns and relationships.

- The final paragraph of the introduction (now the second-last paragraph) has been expanded to better describe the 'patterns and relationships' that are expected to emerge from the study, with specific examples of patterns and relationships described.

(p.4, lines 11-20) The main problem with the method is the data source used in the paper. This source consists of output from a global tool produced by the World Resources Institute. In the paper this model is described as made up of "global hydrologic and hydraulic models" and more. It raises some concern about the accuracy of such models when they are "global". Unfortunately, the only reference in the paper leads to a web site, which in turn refers to the name of a model but with no proper literature reference. So, it is quite difficult for the reader to judge for herself how useful the data used in the paper is. Considering the spatial resolution given in the said website the usefulness is doubtful. As an example of interest for the reviewer: the country of Sweden is presented as one "basin" !

Thank you for raising this issue. Perhaps the simple website address is not sufficient as a reference for this data, and that will be amended. Specific information on the creation of the database can be accessed through hitting on the 'i' in the upper right corner of the cited website, which leads to this page describing the models:

http://floods.wri.org/#/

This page cites the sources as:

> Winsemius, H. C., et al. "A framework for global river flood risk assessments." Hydrology and Earth System Sciences 17.5 (2013): 1871-1892.

and

> Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., van Beek, R., Bierkens, M. F., ... & Winsemius, H. C. (2013). Assessing flood risk at the global scale: model setup, results, and sensitivity. Environmental research letters, 8(4), 044019.

As the first of these papers is in fact published by HESS, we have deemed this data to be trustworthy.

Further papers published with this data include:

> Winsemius, Hessel C., et al. "Global drivers of future river flood risk." Nature Climate Change 6.4 (2016): 381-385.
>
> Jongman, B., Winsemius, H. C., Aerts, J. C., de Perez, E. C., van Aalst, M. K., Kron, W., & Ward, P. J. (2015). Declining vulnerability to river floods and the global benefits of adaptation. Proceedings of the National Academy of Sciences, 112(18), E2271-E2280.
>
> Muis, S., Güneralp, B., Jongman, B., Aerts, J. C., & Ward, P. J. (2015). Flood risk and adaptation strategies under climate change and urban expansion: A probabilistic analysis using global data. Science of the Total Environment, 538, 445-457.

Yes, it is a concern that Sweden is represented as a single basin in this database! We have chosen to work with the city information instead of basins, so fortunately this is basin does not appear in our analysis.

- This information and these references have been added to the data section.

(p.4, lines 29-34) Another weakness in the method is that the actual, real flood protection level is not included. Instead flood protection level is based on the assumption of proportionality with national income level.

This is a shortcoming of the method, as the real flood protection levels for each city are not known (by us) for each city. However, the assumptions made to overcome this are in line with current practice, as outlined in lines 33-37 of p4 with references to contemporary literature:

> '…we assign an assumed flood protection level to each city based on the country's World Bank income level (as in the World Resource Institute's Aqueduct Global Flood Risk Country Rankings, website 6) due to a lack of information on each city's actual protection level. This method follows recommendations based on the rational that higher standards of protection against flooding may be expected in higher income countries (Jongman et al., 2012; Nicholls et al., 2008), and findings by Doocy et al. (2013) that flood impacts are significantly associated with classification of income level by the World Bank. '

This method is necessary as standardised, current information on the nearly 100 cities in the data set is not readily available. This shortcoming is acknowledged in the literature, as Winsemius et al. (2016) state that 'currently installed flood protection is an important missing link in the assessment of global flood risk'. If there is a better practise that we have missed, we would be very happy to learn of it.

- A paragraph discussing this shortcoming has been added to the conclusion.

(p.4, line 18-22, 38) On the one hand: changes due to socioeconomic development are driven by population and economy. On the other hand: protection level, which is assumed proportional to national income level is kept constant over time. This is inconsistent.

True, this seems inconsistent. However, 'changes due to socioeconomic development' by nature consist of changes in population and economy, which are thoroughly measurable for current conditions and reliably estimated for future scenarios based on comprehensive models. On the other hand, the current flood protection levels are estimates and the changes in flood protection in the 20 years of the study

would also be estimates, based on estimated projections of income levels (which are not readily available), increasing the level of uncertainty in a variable value that is already an assumption. Therefore, we have deemed it reasonable to retain a constant flood protection level for each city over the 20 years.

(p. 6, lines 16-19) The present situation is in the paper characterized based on conditions 2010 while effects of climate change is based on projections for 2030. This time interval, 20 years, is a bit short for meaningful comparisons.

Whilst this is a short timeline, it is restricted by the data that is available. Studies at a global scale have been traditionally limited due to lack of cohesive data sets, and therefore this data set is valuable for the fact that it spans a global set of cities and provides an opportunity for comparison. As the data is only provided for 2010 and 2030, there was no choice for us to use different dates. Whilst it may not provide a long-term outlook, at the very least an insight into the current and near-future conditions can be gained.

- This is a good point you have made, and we have added a discussion of it in the conclusion.

(p.8, 3 lines 5-15; Fig. 1a) While the colouring of the maps allows for different absolute impact on people and material damage for an individual city, the location of all cities are exactly the same on both people and damage maps. This latter fact seems to indicate that the relative difference in impact between cities remains almost identically the same for both people and damage.

The relative placement of the cities on the map is the main map characteristic providing insight into the features of the data. The map is created based on both population and material damage data, organizing the cities with respect to each other based on both of these factors. The same map is repeated in Fig. 1a, with the background colouring indicating levels of population damage (on top left) and material damage (on top right). Therefore, the relative distance between cities on the map indicates the differences in a *combination* of population and material damages (which can be discerned from the colouring). For example, Cincinnati (top right) incurs high material damages costs, and medium population affected, whereas Ulan Bator (mid left) has similar population affected to Cincinnati, but much lower material damages costs. This interpretation is explained on p8 in lines 5-23.

- This explanation has been added here.

(p.17, lines (23-38) The actual, on the ground flood management measures and other socioeconomic development are presented for Marrakech. However, in the model producing the source data for this paper only some of these factors were incorporated. See previous comment re page 4.

The analysis in this paper is based solely on the data provided. We are not aware of the extent to which the on-the-ground flood management measures are incorporated into the socioeconomic models which produced this data. The discussion on p17 is included as a point of interest to help characterise a couple of individual cities, providing possible reasons why they may fit into the map where they do, for the readers' interest.

- A paragraph has been added after paragraph 4 of the discussion to explain this, and references have been added to the data to reinforce this point.

Technical corrections:

(p.1) The title is“.....changes on river...”should be“.....changes in river...”

- Good point! This has been changed, thank you.

(p.7: 25) “.....Whist...” should be “.....Whilst ...”

- Thank you, this has been updated.

(Fig. 1) Fig. 1a) is difficult to read because of the many cities and the small font; however still ok. Fig. 1b) is too blurry due to combination of background colours and text. More or less impossible printed; still difficult on the screen.

- Figures 1b), 2b), 3a) and 3b) have been reproduced with larger text, and the background colours have been lessened in intensity to make the text appear more clearly. We have tried a variety of text sizes on Figures 1a) and 2a), and believe the current text size makes it more clear precisely where each city is located on the map than a larger text size would. A note has been added to the caption referring the reader to the online version which can be zoomed in on. Figure b) maps with larger text could also be referred to if there is an issue reading the a) maps. The following is a screen shot of the pdf version of 1a), we believe it is readable.

(Fig.2) Similar problems as with Fig. 1

- Figures 1b), 2b), 3a) and 3b) have been reproduced with larger text, and the background colours have been lessened in intensity to make the text appear more clearly.

(Fig.3) Similar problems as with Fig. 1 &2. Difficulties reading this figure exacerbated by all the information crammed into one page. The gradients indicated along the axes are impossible to interpret.

- Figures 1b), 2b), 3a) and 3b) have been reproduced with larger text, and the background colours have been lessened in intensity to make the text appear more clearly. A screen shot of the pdf of Figure 3a) is shown below. Though the gradients aren't too large, a general idea of their magnitude in the different regions of the map can be discerned, which was the intention. The figure has been reformatted with the legend at the bottom, which has allowed the figure room to expand sideways on the page, making it larger. Again, a note has been added that the reader can refer to the online version to zoom in on the figure.

(p.13:1) “.....reduction on...” should be “.....reduction of ...”

- Thank you, this has been updated.

P.16:16) “.....loses...” should be “.....losses ...”

- Thank you, this has been updated.

Patterns and comparisons of human-induced changes in river flood impacts in cities

Stephanie Clark[1*], Ashish Sharma[2], Scott A. Sisson[1]

1: School of Mathematics and Statistics, University of New South Wales, Sydney, Australia
2: School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia
***** *corresponding author*

ABSTRACT

This study investigates patterns of current conditions and anticipated future changes in city-level flood impacts driven by urbanisation and climate change. Global patterns relating urban river flood impacts to socioeconomic development and changing hydrologic conditions are established, and world cities are matched to these patterns. Comparisons are provided between 98 individual cities. We use a novel adaption of the self-organizing map method to establish and present patterns in the nonlinearly-related environmental and social variables. Output maps of prevalent patterns compare baseline and changing trends of city-specific exposures of population and property to river flooding, revealing relationships between the cities based on their relative map placements. Cities experiencing high (or low) baseline flood impacts on population and/or property that are expected to improve (or worsen), as a result of anticipated climate change and development, are identified and compared. This paper condenses and conveys large amounts of information through visual communication to accelerate the understanding of relationships between local urban conditions and global processes, and to potentially motivate knowledge transfer between decision makers facing similar circumstances.

Keywords: urban hydrology; self-organizing maps; SPATIOTEMPORAL patterns; clustering; city-scale; neural networks; Anthropocene

1 INTRODUCTION

Through urban development and climate change, humans are progressively generating (and being on the receiving end of) increased hydrologic impacts, with these anthropogenically induced changes becoming particularly evident in cities (Revi et al., 2014, Mills 2007, Kreimer et al., 2003; Willems et al., 2012). With high densities of urban populations, infrastructure, property and industry, cities are both substantial drivers and receivers of environmental impacts. River flooding, the environmental event affecting more people than any other natural hazard (Doocy et al., 2013; Desai et al., 2015; Sofia et al., 2016), currently poses a threat to almost 380 million urban residents (UN-Habitat, 2014). Globally, hydrologic regimes leading to urban flooding are varying with climate change (Desai et al., 2015; UNEP, 2016; Willems et al., 2012), and locally, socioeconomic factors associated with urban development (variations in population growth, development, land use and urban density) are uniquely altering each city's individual response to these changing flood levels and frequencies (Desai et al., 2015). In the next few decades, cities will need to anticipate and adapt to this combination of shifting quantities of water and city features (Revi et al. 2014; Doocy et al., 2013). In this study, we aim to develop an understanding of the prevalent global patterns of human-environmental relationships influencing city-level river flooding, and discover how a global set of individual cities fits into these patterns.

Climate change and urbanization are combining to force more frequent flooding and higher flood peaks in cities, though the influence of each factor varies spatially and temporally (Desai et al., 2015). Historically, cities have formed near rivers and population density is still highest, globally, where the closest water feature is a large river. As cities grow, the proximity of population and property to these water courses increases (Kummu et al., 2011). It is estimated that 70% of the world's population will live in cities by 2050 (UN-Habitat, 2010), up from 54% in 2015 (UN-DESA, 2015). With this rapid urbanization, highly populated areas are experiencing an increase in flood vulnerability (Kreimer et al., 2003), as unplanned expansion often leads to migration into urban flood plains (Jongman et al., 2012; Revi et al., 2014). Global urban land cover is increasing at a rate over double that of urban population growth (Angel et al., 2010a) and is projected to increase three-fold by 2030 (Pachauri et al., 2014). More impervious areas and encroachment into the surrounding countryside are forcing faster concentrations of rainfall in urban rivers during storm events, as well as higher flood peaks (Desai et al., 2015; Doocy et al., 2013; Kreimer et al., 2003).

Hydrology in cities is also affected by increased surface temperatures associated with climate change. Already, increases in the frequency and intensity of precipitation (Frich, et al. 2002; Desai et al., 2015; UNEP, 2016), changes in spatial and temporal storm patterns (Wasko & Sharma, 2015) and changing snow melt conditions (Schiermeier, 2011; Barnett et al., 2005; Immerzeel et al., 2010) are leading to variations in the magnitude, frequency and timing of urban river floods, with higher peak flows and shorter response times (Shiermeier, 2011; Cunderlik, 2009). These changing patterns of precipitation and runoff are complex and not uniformly spatially distributed (Meehl et al., 2005; Desai et al., 2015; Wentz et al., 2007; Frich et al., 2002). In the future, cities in particular are predicted to become even more vulnerable to extreme hydrologic events as a result of climate change (Pachauri et al., 2014; Willems et al., 2012; Revi et al., 2014, Sofia et al., 2016). Increases in rainfall intensity at urban hydrology scales of up to 60% are anticipated by 2100 (Willems et al., 2012), and the micro-climates of cities are expected to interact with climate change in a variety of ways, potentially exacerbating flood effects (Revi et al., 2014).

In this paper, a comparison is made amongst a selection of cities based on their current and projected future urban river flood impacts on population and property, resulting from an anticipated combination of climate change and development. It should be noted that fluvial flooding is the only type of flooding that is considered here, and this study does not include an analysis of cities subject to coastal or pluvial flooding. Analysing data with city-specific projections of changes in hydrology, population and development levels (based on future

climate scenarios, projected development pathways, and a best assumption of flood protection standards) we produce an analysis and visualisation of the patterns of baseline conditions and anticipated changes in city-level river flooding impacts to the year 2030. We establish the prevalent global spatial and temporal patterns of urban flood impacts, explore these impacts as resulting from both developmental and hydrological drivers, and match the cities to their most similar pattern. The patterns are established through dimension reduction, clustering and visualisation of multivariate data with an adaptation of the self-organizing map (SOM) technique. The SOM is an artificial neural network useful for exploring nonlinearly related variables, and is popular for investigating potentially difficult-to-define environmental responses to human influences (e.g. Shanmuganathan et al., 2006; Vaclavik et al., 2013; Clark et al., 2016b) as well as providing comparisons between geographic areas (Kaski & Kohonen, 1996; Clark et al., 2015; Clark et al., 2016). We begin by presenting analyses of patterns of urban flood conditions (as measured by the amount of population affected and urban damages costs) for a baseline global snapshot (2010), then investigate projected temporal changes (up to 2030), and finally combine this information into a global temporal analysis of the cities. As individual cities are matched to their closest patterns at each stage, we discover clusters of cities with similar urban flooding characteristics and projected trends.

A growing body of research is investigating the impact of anthropogenic changes on urban flooding at regional and global scales, however we have found no literature comparing specific cities in terms of changing city-level flood impacts on populations and property. The Intergovernmental Panel on Climate Change's 5th Assessment Report Chapter 8 'Urban Areas' (Revi et al., 2014) discusses the vulnerabilities and resilience of cities to climate change in general, noting that the analysis is based on economic losses and would differ if a human component is included. Jongman et al. (2012) investigated global trends of coastal and river flooding based on changing regional population densities and land use. Increased vulnerability to flooding is attributed to population growth or increases in wealth, though the modelling does not include changing hydrology due to climate change. Jongman et al. (2015) estimated regional trends in human and economic river flooding vulnerabilities by income level, through hazard and exposure calculations. Kunkel et al., (1999) investigated the increasing trend of economic losses and fatalities in the USA due to increasing vulnerability to floods, however the climate change contribution to this increase was not possible to quantify due to a lack of data. Winsemius et al. (2016) produced the first projections of global future flood risk that consider separate impacts of climate change and socioeconomic development, with results discussed by geographic region (river basin) and economic level. The investigation of the connection between coastal flooding and climate change (increasing storms combined with sea level rise) is more common in the literature than the connection between river flooding and climate change (Nicholls et al., 2008; Nature, 2016) due to better data availability. Most existing river flood assessments are at a local or regional scale (as in Muis et al., 2015), limiting the possibility to compare between multiple cities, as studies at a global scale have traditionally been limited by a lack of datasets and methods. Sofia et al (2016) emphasize that analyses of climate change and socio-economic development as both drivers and receptors of flood risk is needed. Muis et al. (2015) call for an investigation between the combination of land use change and hydrologic change on future flood risk. (Jongman et al., 2012) highlight that due to population growth and climate change, global methods incorporating both spatial and temporal dynamics to investigate inland flooding at the city scale are necessary for global development studies and estimating costs associated with climate change. To date, a global examination of changing flood conditions at the city level resulting from urban development and climate change, including a direct comparison between specific cities, has not been made. The analysis we present here corresponds directly to this gap in the literature.

General patterns as well as specific relationships can be extracted from the output maps in this paper. In the interest of channelling the 'potential of visual communication to accelerate social learning and motivate implementation of changes' (Sheppard, 2005) the aim of the method used here is to discover and demonstrate

potentially interesting global patterns and relationships that would not otherwise be evident in the data, for example: clusters of cities which are currently experiencing high flood impacts that are projected to greatly increase in the future, and to what extent this may be due to climate change (or socioeconomic development) within each city; which cities not currently experiencing notable effects of flooding may expect to in the future; which cities are projected to mitigate potentially adverse flood effects from climate change with reductions in flooding due to socioeconomic factors; which cities are projected to experience an increased flood vulnerability driven by socioeconomic factors alone; and the relationship between the changes in vulnerability of the population and urban damages costs for each city.

The comparison of individual cities in this study (rather than river catchments) allows a blending of environmental and social information which reinforces the co-dependence of humans and their natural environment, a relationship which is often easily overlooked by urban dwellers. Explicitly visualising the role that urbanisation may have on the environmental conditions experienced by urban citizens is an essential reminder of this connection. Cities potentially facing similar circumstances and challenges are identified in this study, suggesting possibilities for a sharing of strategies. As climate change, development, and urban administrations transcend river basin boundaries, an investigation of impacts and determination of potential mitigation strategies at the city level as well as the basin level expands the potential for decision makers to be presented with all the available, relevant data for consideration.

2 Data and method

Data

The data set used in this study combines city-level estimates of annual expected urban river flood impacts on population and urban damages costs (2010), projections of future changes in flood impacts attributed to climate change and/or development (up to 2030), and socioeconomic data for a globally distributed set of cities.

The selection of cities used here is based on a list provided by the Lincoln Institute of Land Policy's Atlas of Urban Expansion (Angel et al., 2010, website 1), spanning all continents except Antarctica, encompassing four economic levels and four population levels. City population data (2010) and future population estimates (2030) are from the UN Department of Economic and Social Affairs (UN-DESA, 2015), and GDP per country are from the World Bank's World Development Indicators database (website 2).

Annual river flood impact estimates are obtained from the global dataset of fluvial flood risk published in the World Resources Institute's Aqueduct Global Flood Analyzer Tool (herein referred to as Aqueduct) (Winsemius et al., 2013; Ward et al., 2013; website 3). Released in 2015, this data set comprises the first unified global set of fluvial flood risk data at the city level. As this data is solely related to the influence of fluvial flooding on metropolitan areas, it does not include coastal or pluvial flood risks. In this data set, Aqueduct provides separate estimates of annual impacts on the number of affected population (people exposed to flood waters) and urban property damages costs (in US dollars), which will be referred to in this paper as 'population' and 'damages' impacts.

Global hydrologic and hydraulic models, inundation modelling, and spatial data sets of population, land use and infrastructure are used within Aqueduct to quantify flood risk in each city. Aqueduct identifies future anticipated changes in urban flood vulnerabilities as driven by climate change (altered hydrology), socioeconomic development (population, land use and economic changes), or in most cases a combination of both. Either of these drivers may increase or decrease the frequency and intensity of flooding, and the resulting flood impacts, for a given city. Three separate scenarios of climate change and socioeconomic development (optimistic,

business-as-usual, and pessimistic) are given in Aqueduct, and in this study we use data from the business-as-usual case for our future flood impact scenario. Future hydrologic and hydraulic estimates in Aqueduct are based on global circulation model data from the ISIMIP project (website 4) and changes in population and economic development are based on Shared Socioeconomic Pathways data with a downscaling procedure that differentiates between urban and rural growth (website 5; Samir & Lutz, 2014). Recent papers published with this data include Winsemius et al. (2016), Jongman et al. (2015) and Muis et al. (2015).

Expected flood impacts are provided by Aqueduct for nine possible levels of city-wide flood protection, from protection against the 2-year average return interval (ARI) flood to the 1000-year ARI flood. This protection level indicates how well protected the area is against flood damage, based on the standard or capacity of flood protection measures such as dikes, levees or dams. In this study, we assign an assumed flood protection level to each city based on the country's World Bank income level (as in the World Resource Institute's Aqueduct Global Flood Risk Country Rankings, website 6) due to a lack of information on each city's actual protection level. This method follows recommendations based on the rational that higher standards of protection against flooding may be expected in higher income countries (Jongman et al., 2012; Nicholls et al., 2008), and findings by Doocy et al. (2013) that flood impacts are significantly associated with classification of income level by the World Bank. We assume each city's flood protection level remains the same during the timeline of this study.

To allow for a comparison between cities of greatly differing sizes and hydrologic conditions, the wide-ranging data values were log-transformed. The data set was then standardized by transforming these values linearly into the range 0-1 (with the lowest value becoming 0 and the highest value becoming 1) for each variable (population affected, urban damages, etc). The data is log transformed, following recommendation by Agarwal & Skupin (2008) that highly skewed variable distributions may benefit from log transformation before use in the SOM. Cities with no flood impacts in both 2010 and 2030 were removed (22 cities), though cities with no flood impacts in 2010 but with flood impacts in 2030 have been kept in the study. The final list of cities is presented in **TABLE 1.**

TABLE 1: CITY LIST - alphabetically by region.

Eastern Asia & the Pacific	
Anqing	China
Ansan	Rep. of Korea
Beijing	China
Changzhi	China
Chinju	Rep. of Korea
Fukuoka	Japan
Guangzhou	China
Leshan	China
Pusan	Rep. of Korea
Seoul	Rep. of Korea
Shanghai	China
Sydney	Australia
Tokyo	Japan
Ulan Bator	Mongolia
Yiyang	China
Yulin	China
Zhengzhou	China

Southeast Asia	
Bandung	Indonesia
Bangkok	Thailand
Ho Chi Minh City	Vietnam
Kuala Lumpur	Malaysia
Manila	Philippines
Palembang	Indonesia
Songkhla	Thailand

South Asia	
Dhaka	Bangladesh
Hyderabad	India
Jalna	India
Kanpur	India
Kolkata	India
Mumbai	India
Puna	India
Rajshahi	Bangladesh
Vijayawada	India

Western & Central Asia	
Ahvaz	Iran
Astrakhan	Russian Fed.
Baku	Azerbaijan
Gorgan	Iran
Istanbul	Turkey
Kuwait City	Kuwait
Malatya	Turkey
Moscow	Russian Fed.
Oktyabrsky	Russian Fed.
Sanaa	Yemen
Shimkent	Kazakhstan
Teheran	Iran
Tel Aviv	Israel
Yerevan	Armenia
Zugdidi	Georgia

North Africa	
Alexandria	Egypt
Algiers	Algeria
Aswan	Egypt
Cairo	Egypt
Casablanca	Morocco
Marrakech	Morocco
Port Sudan	Sudan
Tebessa	Algeria

Sub-Saharan Africa	
Accra	Ghana
Bamako	Mali
Harare	Zimbabwe
Ibadan	Nigeria
Johannesburg	South Africa
Kampala	Uganda
Kigali	Rwanda
Ouagadougou	Burkina Faso

Latin America & the Caribbean	
Buenos Aires	Argentina
Caracas	Venezuela

Guadalajara	Mexico	Cincinnati	United States	Castellon	Spain
Ilheus	Brazil	Houston	United States	Le Mans	France
Jequie	Brazil	Los Angeles	United States	Leipzig	Germany
Mexico City	Mexico	Minneapolis	United States	London	UK
Montevideo	Uruguay	Modesto	United States	Madrid	Spain
Ribeirao Preto	Brazil	Philadelphia	United States	Paris	France
Santiago	Chile	Pittsburgh	United States	Sheffield	UK
Sao Paulo	Brazil	Springfield	United States	Thessaloniki	Greece
Tijuana	Mexico	St. Catharine's	Canada	Warsaw	Poland
Valledupar	Colombia	Tacoma	United States	Wien	Austria

North America		Europe	
Chicago	United States	Budapest	Hungary

METHOD

We use an extension to the self-organizing map method to determine patterns and similarities in the impacts, changes and drivers of urban flooding amongst the cities. The self-organizing map (SOM, Kohonen, 2001) is an unsupervised learning algorithm from the family of artificial neural networks that discovers patterns in multivariate data sets with nonlinear inter-variable relationships.

The SOM reduces the dimensionality of the data set by creating a (in this case) two-dimensional map grid which, through an iterative process, is essentially bent and stretched over the data set until it best characterizes the shape of the data cloud. The numerous data items become represented by a (usually) much smaller number of map nodes, known as prototypes. The map nodes, or prototypes, move iteratively into position amongst the data whilst maintaining their grid formation, establishing a higher density of prototypes in areas of higher data density. Once in position, the prototypes represent the most prevalent patterns in the data. Each data item is then matched to its closest prototype, creating clusters of similar data items.

The SOM algorithm consists of a two-step iterative process of comparing the map and the data, and then updating the map to better represent the data. The method begins with a calculation of distances in data space (in this case we use Euclidean distance) between each data item, x_i (where $i = 1{:}\,N$), and each map node, m_j (where $j = 1{:}\,M$). Data and map nodes vectors are all of the same dimension, d. The goal of the comparison stage is to find the nearest map node to each data item (commonly referred to as the best matching unit, BMU), which is then given the index c, using the following calculation:

$$\|\mathrm{x_i} - \mathrm{m_c}\| = \min_\mathrm{j}\{\|\mathrm{x_i} - \mathrm{m_j}\|\}.$$

This partitions the data into subsets of items sharing the same nearest node, $\mathrm{m_c}$. Next, the locations of the map nodes are adjusted to become closer to their nearby data items. Application of a smoothing 'neighbourhood' kernel during this stage produces a smoother map by updating neighbouring nodes to a similar extent based on the nearby data. That is, the location of each map unit, m_j, becomes updated based on a weighted average of the data items matching itself as well as its neighbouring nodes, where the weighting is given by the neighbourhood kernel. The size of the kernel decreases with each iteration to include fewer nodes. We use a Gaussian shaped neighbourhood kernel, where h_{ij} (the neighbourhood kernel element indicating the influence of each data item, x_i, on the updating of node m_j) is defined at iteration t as:

$$\mathrm{h_{ij}(t)} = \exp\left(\frac{-(m_c - m_j)^2}{2\sigma^2(\mathrm{t})}\right)$$

where σ is the kernel radius. At each iteration (t), the updated node locations are found as in (Kohonen, 2013):

$$m_j(t+1) = \frac{\sum_{i=1}^{N} h_{ij}(t)\, x_i}{\sum_{i=1}^{N} h_{ij}(t)}.$$

After map training is complete, the map node vectors each represent a unique combination of variables in the data, according the final location of the map nodes in data space. Each of these unique combinations of variables represent a characteristic pattern in the data. The data items are once again matched to their closest map node, forming clusters of data that best match each pattern.

In this study, the 'patterns' are the key characteristics represented by each map node vector (such as specific baseline and/or projected flood conditions, and the drivers of change). The 'cluster' members are the cities that match the pattern represented by their nearest map node better than they match the patterns of any other nodes.

As the SOM is an unsupervised learning algorithm, there is no subjectivity in the resulting cluster memberships. The iterative training process discovers the principal curves of the data set (the nonlinear directions of maximum variance) and aligns the map coordinate system with these, so that the two axes of the map generally follow the first two principal curves of the data. When the map is presented in its two-dimensional form, with data items located at their nearest map node, similar data ends up in close proximity on the map and dissimilar data is far apart. Through the SOM creation process the prevalent data patterns are identified by the nodes, data items become grouped into clusters around these patterns, and the clusters are ordered by similarity on the map. For a more detailed summary of the SOM method, refer to e.g. Clark et al. (2015).

In this study, the data set is split into two subsets ('baseline' data and 'projected future changes') for each city, allowing a progressive investigation of spatial and temporal patterns of urban flooding. A series of three separate SOMs (also referred to as maps) are created with prevalent global patterns and city similarities established separately on each map through colouring and labels, as follows:

- SOM1 explores the spatial properties of the baseline data set, enabling a comparison of the state of urban river flood impacts in each city at a snapshot in time (2010).
- SOM2 explores patterns of projected temporal changes in impacts of urban flooding on population and property (to 2030), incorporating the drivers of climate change and urban development, and
- SOM3 portrays the temporal relationships between the cities in a type of longitudinal exploratory data analysis, clustering cities that are similar in the baseline situation and are also projected to trend similarly in response to each driver in the future.

SOM1, the baseline map, depicts prevalent global spatial patterns and identifies urban flooding conditions in each city based on two variables: 1) the total population affected annually by river flooding, and 2) annual urban property damages costs incurred by river flooding. The map is created based on these two variables, though by projecting new variables onto the trained map it is also used to show: 3) the percentage of each city's population affected, and 4) the percentage of the country's GDP affected. Usually used with higher-dimensional input data, the SOM method is useful here for creating a map with two variables as the nonlinear projection establishes the relationships between cities in alignment with the directions of maximum variance (ie. the directions of most importance) in the data. It also allows for the results to be used as input into SOM3 later.

SOM2, the future projected changes map, describes the anticipated alterations in urban river flooding in each city by 2030. This map is based on four variables of projected changes and their associated drivers: 1) the projected change in population affected annually, 2) the projected change in annual urban damages costs, 3) the proportion of change in population affected that is anticipated to be attributable to climate change, and 4) the proportion of change in urban damages costs that is anticipated to be attributable to climate change. The

remainder of the increase or decrease in impacts is attributed to socioeconomic causes (such as population change, urban density change, increased city footprint, and changes in urban land cover).

SOM3, the temporal map, uses the location of each city along the axes of the two-dimensional baseline and future projected changes maps (which essentially delineate the first two principle curves in each higher dimensional data subset) as input data. In creating SOM1 and SOM2, the baseline and future data subsets have already been reduced to their two most prominent dimensions respectively (which have become the axes of these maps), and each of these four dimensions is considered equally when placing the cities on the temporal map. This method is based on the method used in Clark et al. (2015) to investigate individual data items transitioning through a self-organizing-time-map, and has been modified for the comparison of patterns on two-dimensional maps of differing sizes and shapes that have been created separately based on different variables.

Distinct patterns that have emerged through the process of training the three maps are represented by the nodes of SOM3. These patterns are the most relevant combinations of dynamic city flood impacts, socioeconomic, and climate change characteristics in the overall data set. SOM3 is clustered, coloured and labelled to indicate the relationships between the cities in terms of similar or differing baseline situations *and* projected changes. Cities with relatively close locations on both the baseline and future projected changes maps are considered to have parallel temporal paths, and will be found close together on the temporal map. Those with converging trends (dissimilar baseline conditions, but similar future projected changes) and diverging trends (close baseline conditions, but dissimilar future projected changes) are also identifiable on this map.

In the creation of each map, grid size and shape have been determined using quantization, topographic and dimension range representation error measures (QE, TE, and DRR) with comparisons between the data set and the map.

The QE (Kohonen, 2001) measures how well the map nodes represent the data items using the sum of squared Euclidean distances between each data item, x_i, and the node closest to it, m_c, averaged over all data points:

$$QE = \frac{1}{N}\sum_i \|\mathrm{m_c} - \mathrm{x_i}\| = \frac{1}{N}\sum_i \sqrt{({m_c}^2 + {x_i}^2 - 2m_c x_i)}\,.$$

The TE (Kiviluoto, 1996) indicates how well the topography of the data set is preserved on the map, giving higher error values for maps that are unnecessarily bent or twisted. The BMU and second BMU for each data point are checked to determine if they are adjacent ($u_{x_i} = 1$ if the first and second BMUs of x_i are neighbours, 0 otherwise), and TE is calculated as:

$$TE = \frac{1}{N}\sum_{i=1}^{N} u_{x_i}$$

The DRR (Clark et al., 2015) measures how well the map represents each variable of the data set to ensure even coverage of the dimensions. The maximum intra-cluster spread of data items in each dimension, d, that become represented by a single map node, x_i (as a proportion of the overall data range in that dimension) is determined. The DRR is calculated as follows, where $x_i(d)$ are data values in dimension d, and $x_{ij}(d)$ are the data values in dimension d that are assigned to map unit j:

$$DRR(d) = \max_j \frac{\max_{ij}(x_{ij}(d)) - \min_{ij}(x_{ij}(d))}{\max_i(x_i(d)) - \min_i(x_i(d))}$$

For the baseline map, a 10*7 grid is found to be the optimum shape to represent the data based on the error measures. An 8*8 map is fitted to the future projected changes data set. After finding these optimum side ratios, the maps are increased in size preserving their side ratios (to 20*14 and 18*18) to allow the data items to spread out until most cities are placed individually, allowing the relationships between all cities to become evident (as in Skupin & Hagelman, 2005). The temporal map is sized at 25*17 nodes. Whilst the input data for the baseline and future projected changes maps were standardized into the range 0-1 before training, the input data for the temporal map is not standardised in order to preserve the ratios between the lengths of the first two principal curves in each of the first two data subsets.

Prevalent cluster characteristics are determined using a 'second level' clustering of the nodes of the SOM (as in Vesanto & Alhoniemi, 2000; Skupin & Hagelman, 2005), performed using Ward's clustering method (Ward, 1963) with the number of clusters determined using the Davies-Bouldin index (Davies & Bouldin, 1979). The Davies-Bouldin index reports the ratio of within cluster scatter (S_j for cluster j) to inter-cluster distances, looking at each cluster and its most similar one, (M_{jk}), with a lower ratio (S/M) indicating a better estimate of the number of clusters of interest present in the data. Ward's minimum variance method is a hierarchical clustering algorithm based on minimizing the total within-cluster variance. With this second-level clustering, each data item of the original data set becomes a member of the same final cluster as its closest node (Vesanto & Alhoniemi, 2000).

The final clustering is visually verified with a SOM 'U-matrix' (Ultsch, 2003). The U-matrix visualises distances in data space between immediately neighbouring nodes, indicating these distances by colour on a grid of the same size as the SOM. By computing how close adjacent map nodes are in data space, the U-matrix is able to provide an indication of cluster boundaries based on large dissimilarities between neighbouring nodes. A greater change in relative distance between the locations of the nodes in data space than in map space is displayed in a lighter colour on the grid, and lesser distances in darker shades. The darker regions of the grid then indicate the cluster centres, separated by lighter coloured boundary areas.

By reducing the information from this multivariate data set into the two most prominent dimensions and finding relationships between the data items at each of these three stages, spatial and temporal information about global patterns of urban flooding is abstracted, and similarities and differences between the cities are clearly portrayed. This method extracts two levels of information:

(1) the most characteristic socio-environmental patterns in the data are found, and
(2) cities are compared to each other with respect to their relative flooding conditions.

The simulations are run in Matlab with use of the SOM Toolbox (website 7) with variables and map sizes as described above.

3 Results

Three SOMs are presented sequentially to reveal three unique sets of patterns in the data, where the term 'patterns' refers to combinations of variables that characterise a specific set of conditions. The cities are clustered into groups with conditions matching these patterns, based only on the given data. The maps each have different sizes, shapes and colours as they represent different subsets of input data.

SOM1: Baseline urban flood impacts

Patterns of urban flood conditions in 2010 are shown on the baseline map, SOM1, in Figure 1. The placement of city labels indicates the relationship of each city to each other in terms of river flood impacts on population and urban damages costs. The map is created by organizing the cities with respect to each other based on both of

these factors. Cities close together are more similar in the amount of population affected and urban damages costs, and cities located far apart are less similar.

The relative placement of the cities on the map is the main map characteristic providing insight into the features of the data, indicating differences in a *combination* of the variables which can be discerned from the colouring of Figure 1(a). Each map node has a four-component vector (representing the value of each of the four variables at the location of the node in data space). The four images in Figure 1(a) show SOM1's city labels over grids coloured separately by the values of each of the four variables (white is low, purple is high). For each city, the relative value of each of the variables can be seen. For example, Cincinnati (top right) incurs high material damages costs, and medium population affected, whereas Ulan Bator (mid left) has similar population affected to Cincinnati, but much lower material damages costs.

The nonlinearity of the relationships between the variables is evident, as is the smooth transition of the values of each variable along the map. General information about the prevalent baseline global patterns and the relative flood conditions in the specific cities can be gained from inspection of these map labels and coloured grids.

Each area of the grid represents a general pattern, or combination of variables in the data, some of which are indicated by annotations on Figure 1(b). In general, higher amounts of population affected and urban damages costs resulting from river flooding are represented by areas towards the top of the map, and these variables decrease in value down the map. Values of affected population are lowest just in from the lower left corner and undulate along the bottom of the map, sweeping upwards to a maximum at the upper left corner. Urban damage values are lowest in the lower left corner and increase in concentric arcs up to the upper right corner. Generally, the left of the map contains patterns involving higher impacts on populations than on property, and the right of the map higher impacts on property than on populations.

From Figure 1(b), relationships can be discerned between regions, as well as between cities in the same region. For instance, cities in North Africa, Sub-Saharan Africa and West & Central Asia are predominantly located in the lower portion of the map, corresponding to a prevalent pattern of low flood impacts on both population and property. Cities in Southeast and South Asia generally correspond to the patterns of high impacts on population and property found in the upper left of the map. Cities in Europe stretch from the top to the bottom of the map, ranging from high overall flood effects (Paris) to no flood effects at all (Thessaloniki). North American cities are matched to patterns that represent more significant impacts on property than on population (down the right side of the map), and are split between those with high property damages (Philadelphia, LA, etc. – in the top right) and those with low damages (St. Catherine's – in the bottom right).

Impacts on GDP and the proportion of the cities' populations affected are shown in the two lower maps of Figure 1(a), though these variables were not used to position the cities on the map. Cities in which river-related urban flooding is estimated to highly affect the country's GDP are coloured on the lower left map. Kigali, in particular, which incurs medium-high flood impacts, sees a large impact on Rwanda's GDP, perhaps because Kigali is the main city in this relatively small country (Kreimer et al., 2003). GDP is most affected by flooding in: Kigali, Bangkok, Yerevan, Dhaka, Bamako and Cairo. Cities in which the flood-affected population forms a significant proportion of the city's population are coloured on the lower right map, predominantly in a horizontal strip across the centre. The highest proportions are in: Jequie (15%), Kigali (7%), Chinju (6%), Le Mans (5%) and Tacoma (3%).

1 a)

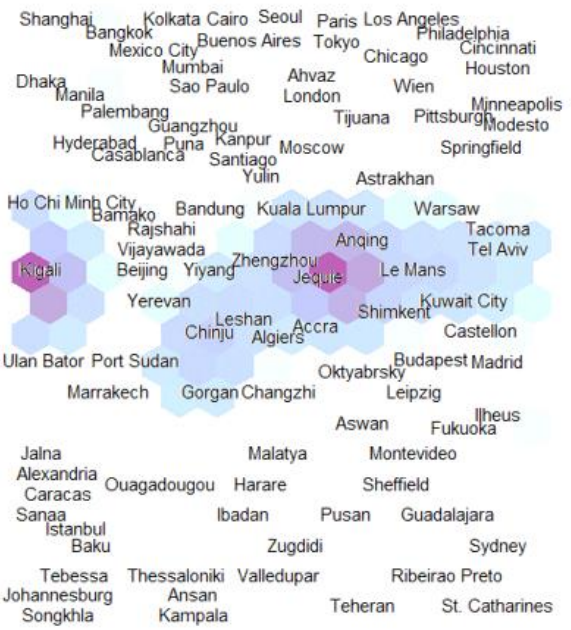
Amount of population affected

Urban damages costs

Low impact High impact

GDP affected

Percent of city's population affected



2

b)

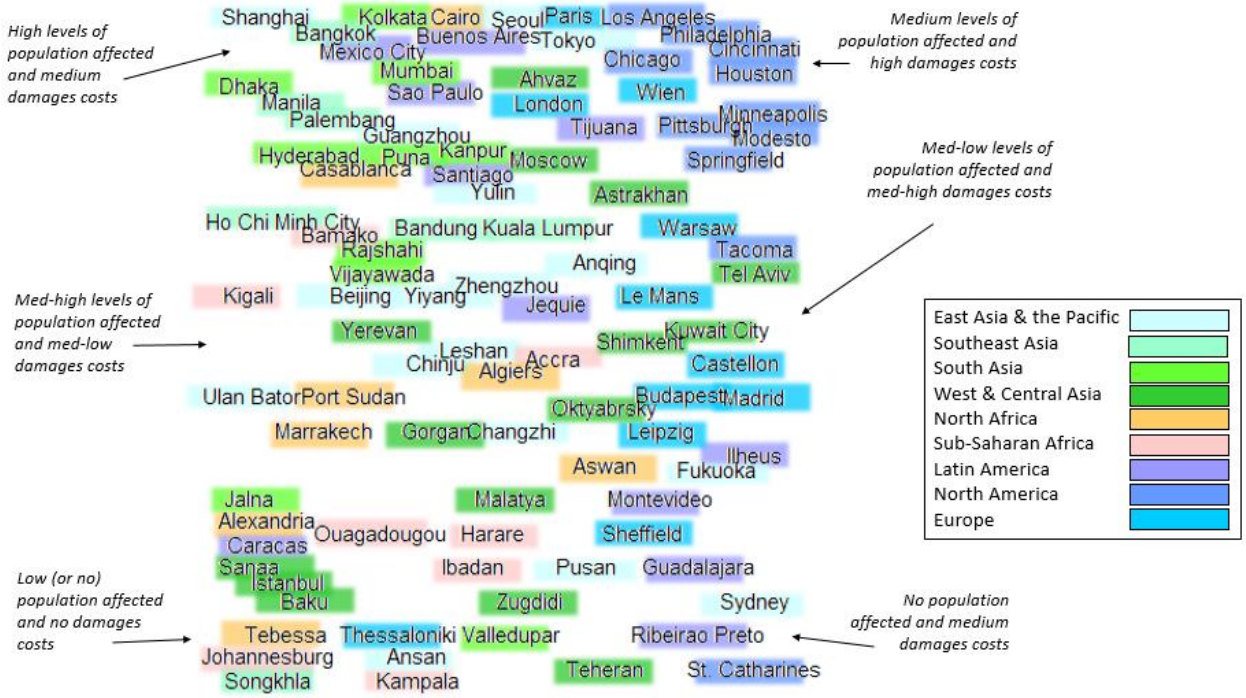


FIGURE 1: SOM1 - BASELINE (2010) URBAN FLOOD CONDITIONS. Cities are placed relative to each other based on annual river flooding impacts on population and urban damages costs. a) The same map is repeated for each of four variables, with colouring indicating low (white) and high (purple) values. b) The city labels are coloured by region (see Table 1), and characteristic patterns of general areas of the map are annotated. The reader may refer to the online version to zoom in on text if required.

SOM2: PROJECTED CHANGES IN URBAN FLOOD IMPACTS (TO 2030)

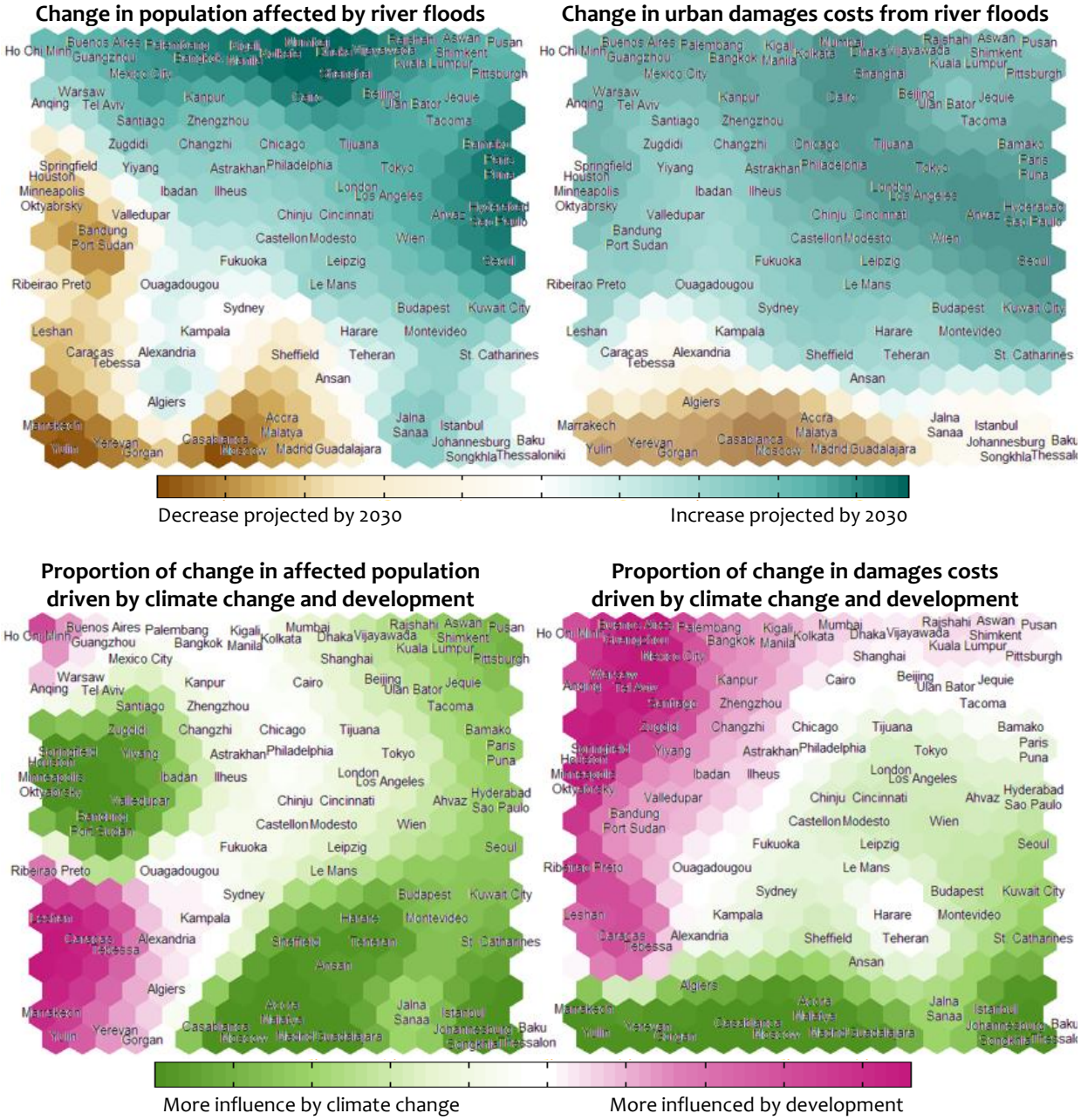
SOM2 identifies the projected patterns of evolving river flood conditions in the cities (between 2010 and 2030), based on city-specific projections of increasing or decreasing flood impacts on population and damages costs, and whether these changes are anticipated to be driven more by climate change or development (Figure 2).

In Figure 2(a), regions of the map representing projected increases in flood impacts on either populations or damages costs are coloured blue and reductions in flood impacts are coloured brown (in the top row), with white indicating no projected change. Projected changes primarily driven by socioeconomic development are coloured purple (in the lower row), and green indicates that the primary driver is climate change. White represents a mid-point in which both climate change and development are predicted impact future flood conditions relatively equally. Areas of the map representing patterns of increased flood impacts predominantly due to climate change or development can be located on Figure 2(b).

Investigating SOM2, we see that climate change is projected to be predominantly responsible for increases in population vulnerability in all cities besides those in the top left corner (around Ho Chi Minh City). Climate change is anticipated to decrease flood damages costs in cities located at the bottom of the map (around Madrid), and decrease impacts on populations in cities in the mid-left (around Minneapolis) and mid-lower (again around Madrid) portions of the map. Socioeconomic development is projected to be the main driver increasing flood

damages costs in cities on the upper-left triangle of the map (roughly from Mumbai down to Tebessa). Only in Ho Chi Minh City is development anticipated to be almost completely responsible for all increases in river flood impacts, all other cities in this study are at least partially affected by climate change. Development is not projected to play any part in a decrease in flood damages costs in any cities in this study (Caracas and Tebessa have no change in damages costs on the upper map, though it is attributed to development on the lower map).

a)



b)

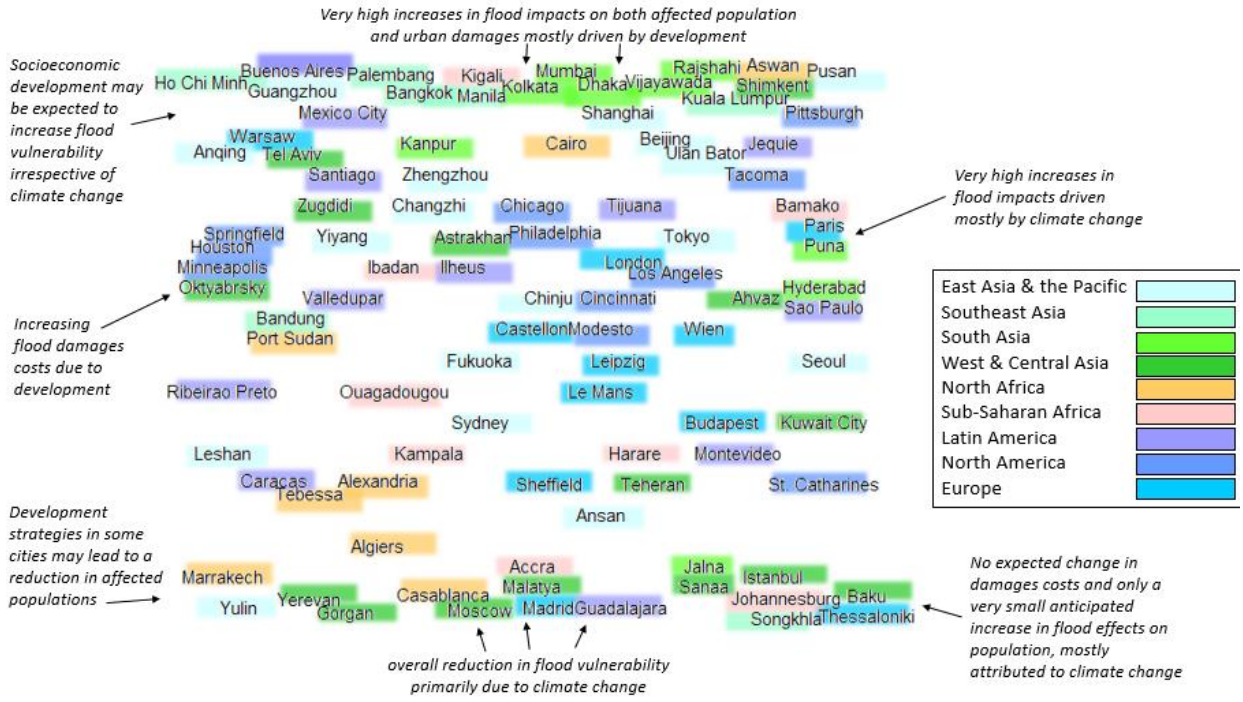


FIGURE 2: SOM2 - PROJECTED CHANGES IN RIVER FLOOD IMPACTS WITH ASSOCIATED DRIVERS. River flooding in individual cities will be affected separately by climate change and development between 2010 and 2030. Cities that are anticipated to experience similar pressures and responses in terms of river flooding impacts are located nearby on the map. a) City labels are placed over coloured copies of the map showing the relative values of each variable. b) City labels are coloured by region, and characteristic patterns of general areas of the map are annotated. The reader may refer to the online version to zoom in on text if required.

Geographic regions are shown on Figure 2(b) with coloured text backgrounds. Cities in Southeast Asia are almost all found at the top of the map indicating high projected increases in overall flood impacts. South Asian cities are mostly located in the two areas of the map with patterns of very high increases in flood impacts, split between those most affected by development (around Mumbai, top middle) and those most affected by climate change (around Puna, mid right). Many North African cities are located in the lower left, indicating anticipated reductions in flooding due to socioeconomic development. North American cities are spread across the middle of the map indicating a wide range of projected changes.

Climate change and development may lead to opposing changes in a city's flood impacts on population and property. A number of cities are predicted to have affected populations decreasing due to climate change, whilst damages costs increase due to socioeconomic factors (around Springfield and Port Sudan, in the mid-left). A decrease in flood effects on urban damages due to climate change, but an increase in affected population largely due to development is, out of the cities in this study, only projected for Algiers (in the lower left portion of the map).

In some cities, both drivers may generate changes in the same direction. For instance, in Marrakech, Yulin, Yerevan and Gorgan, climate change is projected to be responsible for a decrease in damages costs whilst socioeconomic development is anticipated to play a major role in the decrease in population affected, suggesting

that the reduction of population vulnerability due to development is complementing the direction of change instigated by climate change. In certain cities near the upper left of the map (Santiago, Zugdidi and Yiyang), an overall increase in flood impacts is expected, with increases in affected population almost completely attributed to climate change and increases in damages costs almost completely attributed to development.

SOM3: TEMPORAL PATTERNS

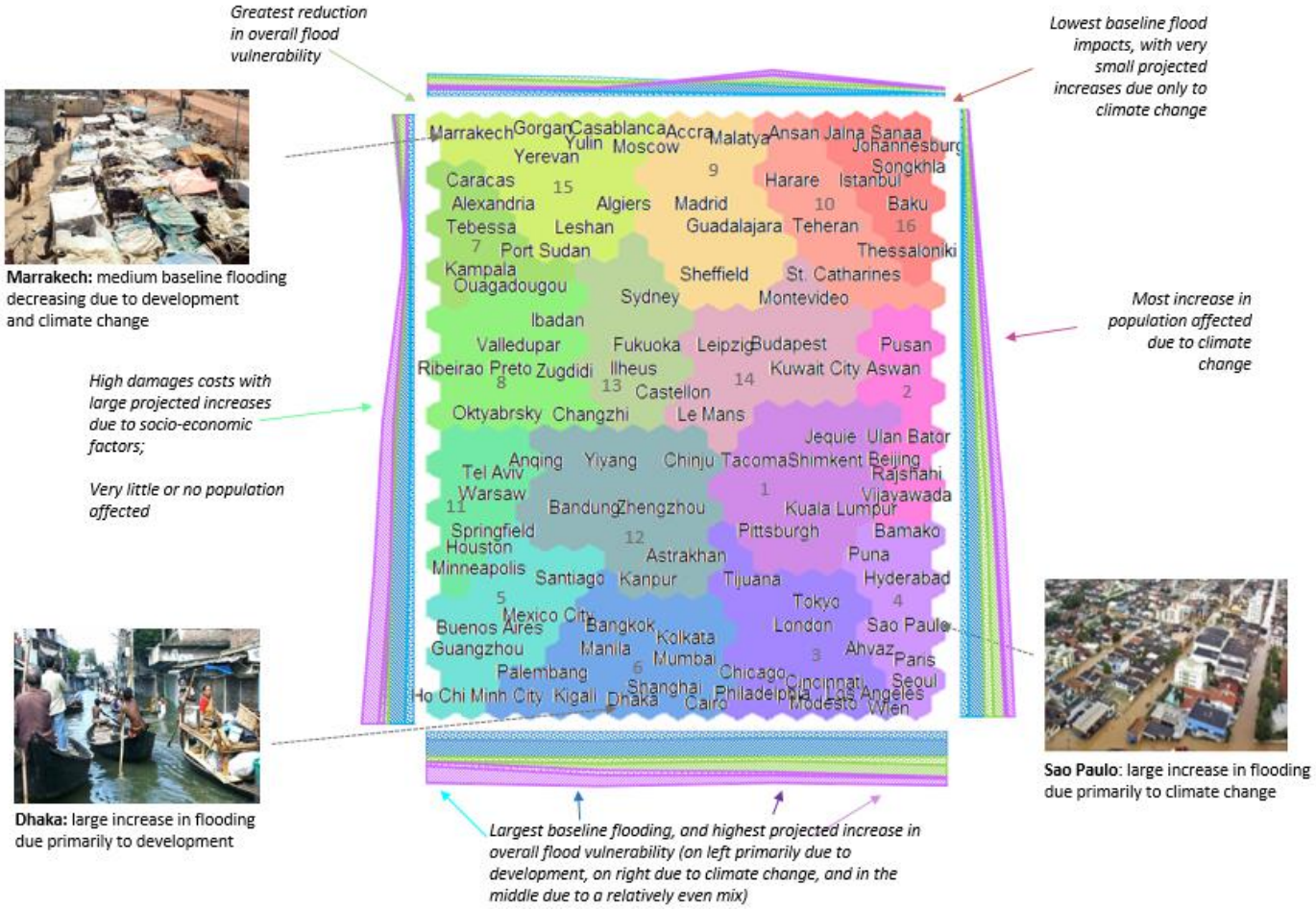
Relationships between the baseline characteristics and projected future changes of urban flooding in the individual cities are shown in Figures 1 and 2 respectively, however potentially similar temporal patterns between the cities are not evident from these maps. To link the information abstracted from the first two maps, we create a temporal map, SOM3, shown in Figure 3. SOM3 identifies which cities experience similar baseline flooding, are expected to incur comparable future hydrologic pressures from climate change and/or development, and are projected to respond in similar ways (or which cities may diverge in the future from similar baseline conditions).

Following the creation of SOM3 and the positioning of cities with respect to each other, we perform a second level clustering to colour the nodes, giving a visual separation to groups of more similar data. Clusters are numbered from 1 to 16 for reference. As the cities are placed on the temporal SOM based on their locations on the baseline and future projected changes SOMs (in which the values of the variables vary smoothly though not monotonically along the axes), again the characteristics of the cities will flow smoothly along the map though multiple peaks and troughs of each variable are possible. The gradients of the cluster characteristics are indicated along the axes in Figure 3(a), which are nonlinear in data space.

Broad overviews of the patterns represented by certain regions of the map are identified on Figure 3 with arrows. The largest increases in flood effects are generally represented by nodes in the lower half of the map, whilst the largest decreases in flood effects are represented by nodes in the top left. Climate change is predicted to be the main driver of changes in population vulnerability along the top and down the left and right sides of the map, and in urban damages on the top and right of the map; therefore, climate change is the leading driver of changes in flood impacts on both population and damages costs at the top of the map. Development is the main driver of changes in flood impacts on populations in the lower and upper left side of the map, and on urban damages in the lower left area of the map; therefore, development is the leading driver of changes in flood impacts on both population and damages costs in cities on the lower left side of the map.

On Figure 3(b) the city labels are coloured by geographic region. We see the cities of each geographical region are more spread out on SOM3 than on SOM1 where each region was generally contained in one or two broad areas of the map. For example, on SOM3 Cairo and Aswan are noticeably separated from other North African cities which are located close together. Although the cities of this region have differing baseline flood levels (as shown on SOM1), most are projected to incur some reduction in future flood impacts (as shown on SOM2), with the exception of Cairo and Aswan. These cities both have forecasts of increased flood impacts - for Aswan increased impacts on the population due to climate change and impacts on property due to development, and for Cairo future impacts are projected to increase due to a relatively even mixture of both drivers. For another example, cities in the USA (all of which have similar starting conditions) are in two well-separated clusters on SOM3 - those around Houston and those around Los Angeles. The cities clustered around Houston are characterised by low impacts on population but high damages costs projected to elevate due to development, implying the possibility for local redemption due to better planning or mitigation strategies. The cities clustered around Los Angeles, however, are characterised by high overall impacts projected to get higher predominantly due to climate change. Further, in Sub-Saharan Africa we see Kigali and Bamako (which have similar medium-high baseline flooding conditions) are both expected to see increased impacts, but the cities are separated by SOM3 as these flood increases are attributed to development in Kigali and climate change in Bamako.

1 a)


Greatest reduction in overall flood vulnerability
Lowest baseline flood impacts, with very small projected increases due only to climate change
Marrakech: medium baseline flooding decreasing due to development and climate change
High damages costs with large projected increases due to socio-economic factors;
Very little or no population affected
Most increase in population affected due to climate change
Dhaka: large increase in flooding due primarily to development
Sao Paulo: large increase in flooding due primarily to climate change
Largest baseline flooding, and highest projected increase in overall flood vulnerability (on left primarily due to development, on right due to climate change, and in the middle due to a relatively even mix)
Marrakech
Gorgan
Casablanca
Accra
Malatya
Ansan
Jalna
Sanaa
Yulin
Moscow
Johannesburg
Yerevan
Songkhla
Caracas
Harare
Istanbul
Alexandria
Algiers
Madrid
Baku
Tebessa
Leshan
Guadalajara
Teheran
Port Sudan
Thessaloniki
Kampala
Sheffield
St. Catharines
Ouagadougou
Sydney
Montevideo
Ibadan
Valledupar
Fukuoka
Leipzig
Budapest
Pusan
Ribeirao Preto
Zugdidi
Ilheus
Kuwait City
Aswan
Castellon
Oktyabrsky
Changzhi
Le Mans
Jequie
Ulan Bator
Anqing
Yiyang
Chinju
Tacoma
Shimkent
Beijing
Tel Aviv
Rajshahi
Warsaw
Vijayawada
Bandung
Zhengzhou
Kuala Lumpur
Springfield
Pittsburgh
Bamako
Houston
Astrakhan
Puna
Minneapolis
Santiago
Kanpur
Tijuana
Hyderabad
Tokyo
Mexico City
Buenos Aires
Bangkok
London
Sao Paulo
Guangzhou
Manila
Kolkata
Mumbai
Ahvaz
Paris
Palembang
Chicago
Cincinnati
Seoul
Ho Chi Minh City
Kigali
Dhaka
Cairo
Philadelphia
Los Angeles
Modesto
Wien
Baseline annual population affected
Baseline annual damages costs
Change in pop due to climate change
Change in damages due to climate change
Change in pop due to development
Change in damages due to development

2

3

4

b)

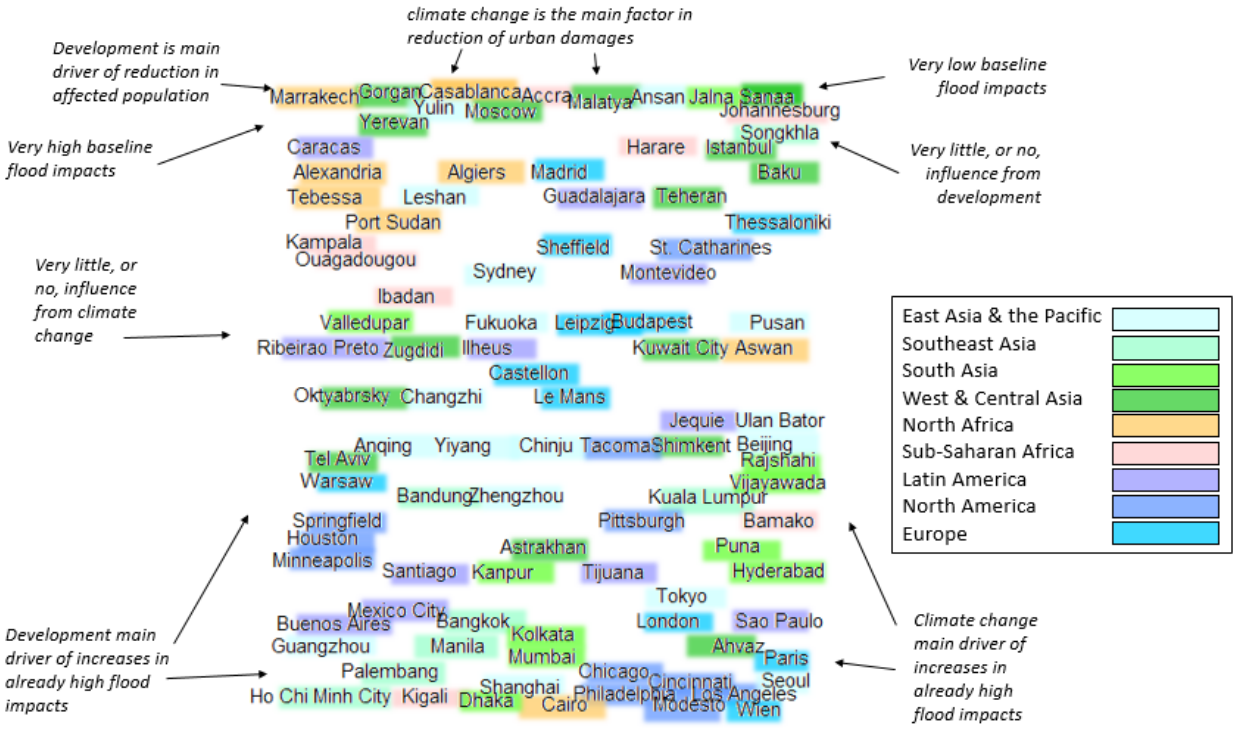


FIGURE 3: SOM3 - TEMPORAL PATTERNS – Cities are clustered close together that share similar baseline (2010) flood vulnerabilities as well as similar anticipated changes driven by climate change and development on population and urban damages costs by 2030. a) Locations of the cities are based on their individual relationships to the principal curves in the baseline and future projected changes data subsets - therefore, the axes represent the most important nonlinear gradients of flood vulnerabilities in the data set. Coloured bars along the axes indicate the average levels of each variable around the edges of the map. Cities are grouped into coloured clusters based on similarities. b) City labels are coloured by region, and characteristic patterns of general areas of the map are annotated. **The reader may refer to the online version to zoom in on text if required.**

To further analyse the characteristics of each cluster and the patterns found on SOM3, the properties of each city in the 16 clusters are shown in a radial plot in Figure 4. Baseline values of population affected (blue, units = number of people) and damages (orange, units = $US) are shown on a symmetrical logarithmic scale ranging from -8 (ie. signifying a value of -100,000,000) to 11 (100,000,000,000) with the region between -1 and 1 on the plot set as linear to avoid logarithmic discontinuities in the vicinity of zero. Zero is indicated by a dashed circumference, and each progressive ring is an exponentially higher (or lower) value. Changes in population affected and damages costs are shown on the same scale, in grey and yellow respectively. Values inside the dashed (zero) circle represent decreases in flood impacts, and values outside represent increases, with the size of the increase or decrease indicated by the distance from the dashed circle. The influence of climate change is shown (light green for population and dark green for damages) on a linear scale from the same zero circumference, in units of 'percentage of projected change attributable to climate change' (each progressive ring is 10%). Green lines closer to the outer ring than the centre therefore indicate that the flood impacts on the city are anticipated to be more influenced by climate change than by development. If the green lines are both in the middle of the segment, this indicates a relatively equal influence of both drivers on both population and

property. Diverging green lines indicate that either population or damages costs are more influenced by climate change, and the other by development.

From Figure 4, we can see the differences between neighbouring clusters, such as 10 and 16 located in the top right of the map. Both clusters are characterized by low baseline impacts of flooding on the population, with small increases in population impacts projected primarily due to climate change. However, cities in cluster 16 incur no flood damages costs at all in the baseline or future cases, yet in cluster 10 damages costs are projected to increase due to climate change and development. Therefore, development has little or no impact on cities in cluster 16 but does play a role in the increase in damages in cluster 10. In the top left of SOM3, we can now also discern the difference between clusters 9 and 15. In both clusters, development is projected to have no impact on the reduction of flood damages costs in most cities. Development does however play a strong role in the reduction of flood impacts on populations in cluster 15 (except for Moscow and Casablanca) but none on populations in cluster 9.

The relationship between the two drivers, climate change and development, can discerned from Figures 3 and 4. Climate change is projected to impact populations more than urban damages costs in clusters stretched across the centre of the map (clusters 8, 11, 5, 6, 12, 14, 1, 10, 1, 2, and 4 - in cities in these clusters, the proportion of change in the population affected attributed to climate change is higher than the proportion of change in damages costs attributed to climate change). In cluster 15, the population is projected to be more influenced by development than damages costs will be (a higher proportion of the change in damages costs is attributed to climate change than for population). In the remaining clusters, climate change (and development) are projected to affect the population and damages costs relatively similarly (clusters 7, 13, 9, 16 and 3). Some examples of diverging impacts on population and damages costs stand out on the radial plot in Figure 4. For instance, in Port Sudan, Sheffield and Bandung, significant reductions in affected population are projected to be 100% due to climate change, however large projected increases (~300 to 400%) in damages are due mostly to development. In Leshan, development is projected to slightly lower the amount of affected population and also to increase damages costs more than three-fold.

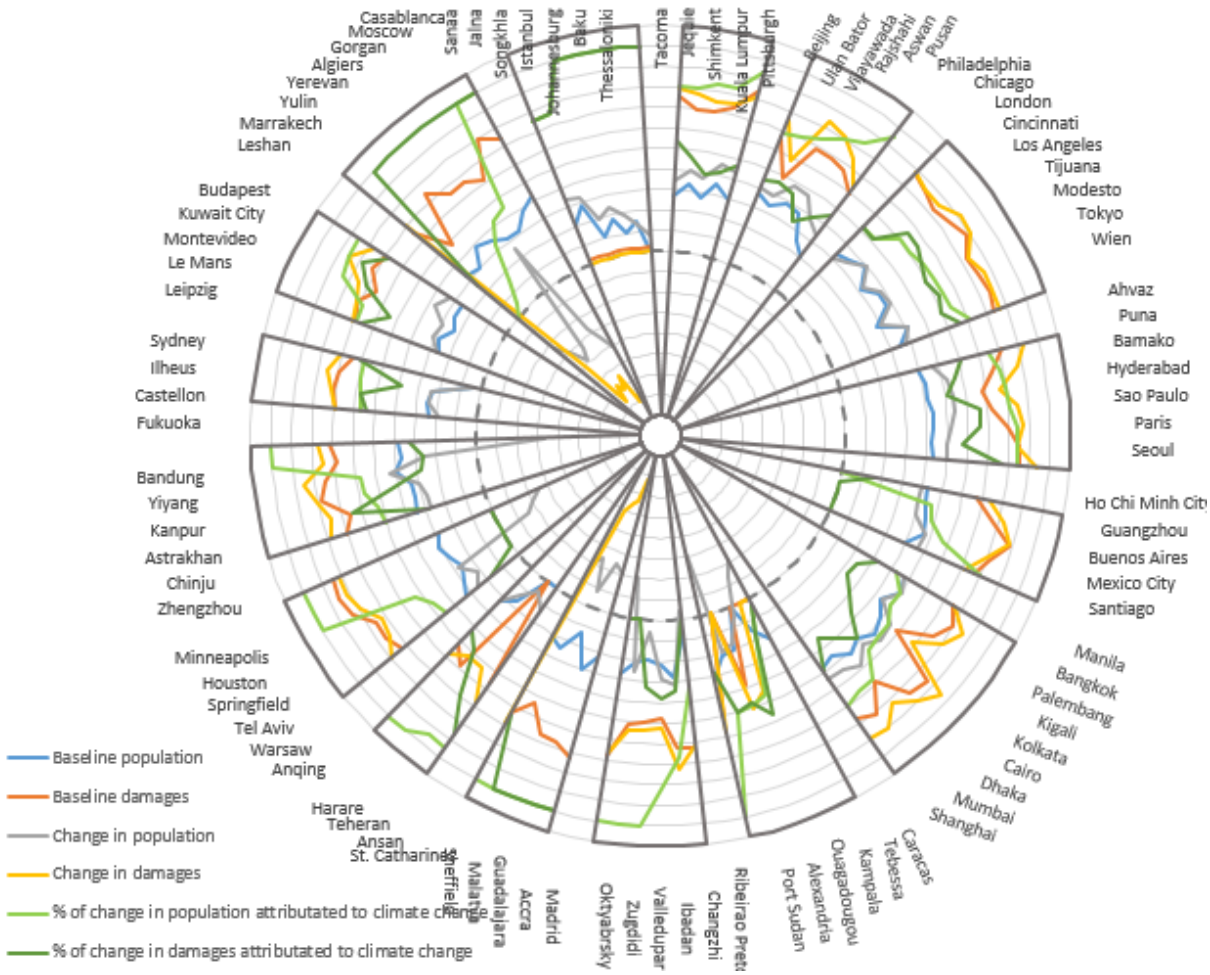


FIGURE 4: RADIAL PLOT OF CLUSTERS OF FIGURE 3 – The city members of the 16 clusters of Figure 3 are shown with their individual variable values. The scale is logarithmic for baseline and changes in population and damages, and linear for the percent of change attributed to climate change, with the dashed circle representing zero.

4 Discussion

In this study, the 'patterns' and 'clusters' in the data have been identified. The patterns, depicting key combinations of variables that are characteristic of the data set, have been extracted at three separate levels on SOM1, SOM2 and SOM3. For example, each pattern of SOM3 is a separate combination of levels of baseline and projected future flood conditions as well as projected influences of climate change and development. The clusters consist of groups of cities whose conditions are anticipated to be similar to these patterns, based on the given data. A discussion of a selection of these patterns and clusters is provided here.

Some cities already experiencing large flood effects are anticipated to incur great flood increases influenced predominantly by socioeconomic factors (migration, changing land use and unplanned development in flood zones). In the lower left region of SOM3, we see examples of cities in which climate change is playing a large role, and yet it is overshadowed by the magnitude of regional economic growth (UNEP, 2016; website 8). Many of these cities are in Asia, where the climate is experiencing warming trends, increasing temperature and precipitation extremes, and rapid glacial melting resulting from climate change (Pachauri et al., (2014) chapter

24: 'Asia'). However, socioeconomic growth in this area is projected to have even more of an impact on urban floods than climate change is. Flood risk and human and material losses are already heavily concentrated in India, Bangladesh and China (Pachauri et al., (2014), chapter 24: 'Asia'), and Jongman (2012) estimates the largest current and future economic exposure to river floods to be in Asia. As an example, we take a closer look at Dhaka which, with a GDP per capita of $1212 in 2015, already has one of the highest levels of population affected annually by flooding (over 130,000) and this number is projected to increase almost five-fold to over 630,000 by 2030. The greatest change predicted for Dhaka, though, is an almost 22-fold increase in annual damage costs (from $8 million to $175 million). Dhaka is subjected to regular flooding from surrounding rivers, with peak flows in the Brahmaputra and Ganges Rivers coinciding to exacerbate flood impacts. In the past, most low-lying areas of western Dhaka were infilled for residential and commercial use, causing a reduction in areas for flood water storage. Furthermore, uncontrolled and unplanned urban expansion is spreading rapidly across the floodplains in the east of the city placing more people in flood hazard zones (Kreimer et al., 2003). These hasty developmental changes are having more of an impact on the urban hydrology of Dhaka than the climate change is. Other examples of cities in similar situations include Kolkata (with the highest baseline affected population in this study), Mumbai (with a seven-fold increase in both population affected and damages due 40% and 60%, respectively, to development), Bangkok (with large increases 50-75% of which are attributed to development) and Ho Chi Minh City (with a 50% increase in affected population and an over five-fold increase in damages costs, almost entirely attributed to development).

Globally, migration trends are seeing more people moving into informal settlements in urban flood zones – the population exposed to river flooding increased by 2.6% more than total global population growth between 1970 and 2010 (Jongman et al., 2012). Most global population growth in the near future is projected to occur in cities of lower income countries, organically and through migration (Kreimer et al., 2003), with urban populations in these countries growing at a rate five times faster than in higher income countries (UN-DESA, 2015) and predicted to double in the next 30 years (Angel et al., 2010). The same regions experiencing such high urban population growth are also projected to triple their urban footprint in the same timeframe (Angel et al., 2010). These developmental changes are leading to, and will continue to produce, substantial effects on urban hydrology if not countered.

Developmental changes in some cities, however, appear to be effectively reducing impacts from river flooding. Marrakech, in cluster 15, is an example of this. The affected population level is projected to decrease mostly due to socioeconomic factors (website 3; Ward et al., 2013; Winsemius et al., 2013). Morocco is taking responsibility to make efforts countering global climate change, and through an 'Integrated Disaster Risk Management and Resilience Program for Morocco' (World Bank, April 2016-Dec 2021), is making its population more resilient to climate change, less vulnerable to natural hazards and ensuring a rapid transition to a low-carbon economy. Through Morocco's National Strategy for Sustainable Development, a commitment has been made to reduce national greenhouse gas emissions by 32% by 2030, through an increase in renewable energy sources to 50% by 2025, a reduction in energy consumption by 15% by 2030, as well as various agricultural, water, waste, forest, industry and housing initiatives (website 9). These housing initiatives in Marrakech include a slum clearance and relocation project, which has become part of urban policy (Ibrahim, 2016), reducing the amount of people inhabiting flood hazard zones. Alert systems in the valleys of the Atlas region above Marrakech have been improved, and the proportion of the population living in slums has decreased from over 8% in 2004 to less than 4% in 2010 (UN-Habitat website). The urbanization rate in Morocco is also projected to slow down towards 2030 (UN-Habitat website). This risk-prevention approach combining early warning systems, relocation of inhabitants out of the flood zone, and less urban expansion is expected to combine to reduce the impact of floods on the population of Marrakesh.

The analysis in this paper is based solely on the data provided in Aqueduct, regardless of the extent to which on-the-ground flood management measures are incorporated into the socioeconomic models which produced this data. A discussion characterizing individual cities is included here as a point of interest to relate the data to current national conditions, providing possible reasons why these cities may fit into the map where they do.

Current high flood impact conditions projected to get much greater primarily due to climate change are anticipated for cities in the lower right of SOM3, with high magnitude changes expected for impacts on both population and property. One of these cities, Sao Paulo, for instance, is expected to experience an almost seven-fold increase in both the number of population affected (to over 140,000 annually) and urban damages costs (to over $500,000,000 annually) by 2030. 15% of the change in population and 35% of the change in damages is attributed to development, but the majority of the change is projected to be caused by climate change. Sao Paulo is the largest city in Brazil, and the city footprint is projected to increase over 38% by 2030, by which time 22% of the urban area may be located in flood zones (Young, 2013). The IPCC (Pachauri et al., (2014) chapter 14 'Latin America') predicts the increase in temperature in central and south Brazil to be the largest projected increase in Latin America, which will be combined with a +10 to +15% increase in autumn precipitation, greatly affecting the hydrologic cycle in the region. The substantial change in development is therefore expected to be eclipsed by the even greater projected change in climate in Sao Paulo.

The anticipated reduction in flood damage costs caused by climate change (evident in Cluster 15) may be a result of changing snow melt conditions upstream of these cities. It has been shown that some global regions will experience a decreasing trend in the magnitude and frequency of snow melt floods as the climate warms, as well as a shift in the timing of these floods (Schiermeier, 2011; Barnett et al., 2005; Immerzeel et al., 2010). Although changing climate in some areas is projected to lessen regional flooding, development within urban flood zones may be severe enough to offset any reductions in flood impacts. This can be seen most prominently in a strip on the left of SOM3 stretching from Port Sudan down to Santiago.

Many high-income cities with already high current flood vulnerabilities have projections for large elevations in damage costs, but not increased levels of affected population. This can be seen in cities on SOM3 centred around London, Tokyo, LA and Vienna (cluster 3), and Sydney and Castellon (cluster 13). Through high levels of planning, preparedness and infrastructure, prosperous regions generally have systems in place to minimize flood impacts on the population, even though they may incur large economic losses (Desai et al., 2015; Kreimer et al., 2003). Almost half of the projected increases in these clusters are attributed to development, suggesting that these cities may have the capacity for lessening potentially elevated flood damage costs by concentrating on planning and mitigation policies.

Though this study does not consider coastal flooding, it may be noted that due to their locations near river mouths, many of the cities in the lower left of the map that are projected to experience high increases in impacts from river flooding are also at risk of increased coastal flooding from intensified storms and sea level rise due to climate change. Mumbai, Guangzhou, Shanghai, Ho Chi Minh City, Kolkata, Bangkok, and Dhaka are 7 of the top 14 cities (out of 136) ranked by current population exposure to coastal flooding. These same cities also comprise the top 7 cities (in this order: Kolkata, Dhaka, Mumbai, Guangzhou, Ho Chi Minh City, Shanghai, Bangkok) ranked by future (2070) estimated population exposed to coastal flooding (UNEP, 2016; Nicholls et al., 2008).

Almost all projected changes in flooding in this data set are of a relatively similar order of magnitude to the original effects, as can be observed on Figure 4. That is, most cities that are only marginally affected by flooding in 2010 are projected to experience only small increases by 2030, whereas cities with larger flood effects can expect greater changes. A significant correlation exists between the magnitudes of the cities' baseline flooding

effects and the changes projected by 2030 (log-transformed absolute values for both variables) – an 88% correlation exists in the number of population affected and a 94% correlation for property damage costs. This supports the findings of Milly et al. (2002) who observed that the frequency of large flood events in large basins had increased substantially in the 20th century, but smaller floods had not.

5 Conclusion

Global patterns of urban flood responses to global and local changes in hydrology driven by climate change and development have been identified and visually communicated. Cities have been matched to these global patterns, and relationships between the individual cities have been discerned with respect to baseline flooding conditions and expected future changes. Information has been extracted from a large, recently released, global data set of city-level flood impacts relating hydrology and urban development, and combined with city-specific demographic information. The analysis and visual interpretation in this study has revealed interesting city-level patterns that are otherwise unobservable in the complex data set, and provides a comparison and distinction between individual cities that is not apparent in regional- or economic-level projections.

We have performed dimension reduction and clustering with a series of self-organizing maps to identify changing global patterns of city-level flood risks. The maps provide an indication of the predominant characteristics which determine the differences in urban river flood impacts between cities, and the cities occupy positions on the maps signifying their relative conditions. The method used here incorporates adaptions to the self-organising map technique for map shape selection and temporal pattern extraction, allowing two levels of information to emerge: the characteristic patterns of dynamic global urban flood vulnerabilities, and a comparison between the cities with respect to flood characteristics and trends.

A shortcoming of the method used here is the assignment of flood protection level based on an assumption of proportionality with national income level. As standardised, current information on the real flood protection levels of all the cities in the data set is not readily available, this assumption has been necessary and has been made in line with current practice. This limitation has been recently acknowledged in the literature, with Winsemius et al. (2016) noting that 'currently installed flood protection is an important missing link in the assessment of global flood risk'. Future studies may aim to include specific flood protection levels for each city.

Whilst the timeline of this study is short, it is restricted by the data that is available. Studies at a global scale have been traditionally limited due to lack of cohesive data sets, and therefore the data set provided by Aqueduct is valuable for the fact that it spans a global set of cities and provides a rare opportunity for comparison. As the data is only provided for 2010 and 2030, there was no prospect for a longer analysis. Whilst this analysis may not provide a long-term outlook, at the very least an important insight into the current and near-future conditions can be gained.

Cities have major implications for climate change mitigation and adaptation (Revi et al., 2014). Unplanned development and urban migration are increasing vulnerabilities to natural hazards (UNEP, 2016) and land cover change and greenhouse gas emissions are intensifying urban hydrology. Understanding the relationship between flood impacts and social vulnerability is a necessary step for prioritizing flood mitigation and prevention strategies (Doocy et al., 2013). Whether the main driver of increased urban flood impacts is development or climate change, cities will benefit from development restrictions and planning standards for urban expansion, sustainable land development, management of population distribution and migration, and early warning systems and preparedness (Revi et al., 2014; UN-DESA, 2014; Doocy et al., 2013).

This study adds to the understanding of natural hazards in a global context, which is an important aspect of regional disaster risk management due to the dependency of local situations on global processes (Desai et al., 2015). The complex nonlinear socio-environmental relationships make it difficult to foresee local responses to global changes (Desai et al., 2015), and therefore this study focuses on risk communication (the process between risk perception and adaptation planning (Cardona et al., 2012)) to provide a visual analysis of the global patterns of evolving flood impacts, socioeconomic development and climate change, and the local city-level consequences of these changes.

Future work may include the addition of greenhouse gas emissions data, geographic location, city sizes and densities to this study, to discern the relationships of these factors with urban flood changes. Greenhouse gas emissions are the largest contributor to global warming, leading to alterations in the intensity of the hydrologic cycle (Pachauri et al., 2014, Barnett et al., 2005; Wentz et al., 2007; Schiermeier, 2011), and cities are the major contributors of greenhouse gases, with a large proportion of global emissions produced by a small global land area (Mills, 2007; Angel et al., 2010; Revi et al., 2014). The addition of these elements could highlight the essential role cities could play in climate change mitigation and the reduction of urban flood impacts.

6 References

Agarwal, P, and Skupin, A. (2008). Self-organising maps: Applications in geographic information science, John Wiley & Sons.

Angel, S, Parent, J, Civco, D, & Blei, A. (2010a). Atlas of Urban Expansion, Cambridge MA: Lincoln Institute of Land Policy.

Angel, S, Parent, J, Civco, D, Blei, A, & Potere, D. (2010b). A Planet of Cities: Urban Land Cover Estimates and Projections for All Countries, 2000-2050. Lincoln Institute of Land Policy Working Paper.

Barnett, T, Adam, J & Lettenmaier, D. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, 438(7066), 303-309.

Clark, S, Sarlin, P, Sharma, A, & Sisson, SA. (2015). Increasing dependence on foreign water resources? An assessment of trends in global virtual water flows using a self-organizing time map. Ecological Informatics, 26, 192-202.

Clark, S, Sisson, SA, & Sharma, A. (2016). A dimension range representation (DRR) measure for self-organizing maps. Pattern Recognition, 53, 276-286.

Clark, S, Sisson, SA, & Sharma, A. (2016b, in press). Nonlinear manifold representation in natural systems. Environmental Modelling and Software.

Cunderlik, JM, & Ouarda, T. (2009). Trends in the timing and magnitude of floods in Canada. Journal of Hydrology, 375(3), 471-480.

Davies, DL, & Bouldin, DW. (1979). A cluster separation measure. Pattern Analysis and Machine Intelligence, IEEE Transactions (2), 224-227.

Desai, B, Maskrey, A, Peduzzi, P, De Bono, A, & Herold, C. (2015). Making Development Sustainable: The Future of Disaster Risk Management, Global Assessment Report on Disaster Risk Reduction. Geneva, Switzerland: United Nations Office for Disaster Risk Reduction.

Doocy, S, Daniels, A, Murray, S, & Kirsch, TD. The human impact of floods: A historical review of events 1980–2009 and systematic literature review. PLoS Curr, 16, 12.

Frich, P, Alexander, LV, Della-Marta, P, Gleason, B, Haylock, M, Tank, A, & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. Climate Research, 19(3), 193-212.

Ibrahim. (2016). Slum eradication policies in Marrakech, Morocco. World Bank Conference on Land and Poverty, Washington DC.

Immerzeel, WW, Van Beek, LPH, & Bierkens, MFP. (2010). Climate change will affect the Asian water towers. Science, 328(5984), 1382-1385.

Jiang, L, & O'Neill, BC. (2015). Global urbanization projections for the Shared Socioeconomic Pathways. Global Environmental Change.

Jongman, B, Ward, PJ, & Aerts, JC. (2012). Global exposure to river and coastal flooding: Long term trends and changes. Global Environmental Change, 22(4), 823-835.

Jongman B, Winsemius HC, Aerts JC, de Perez, EC, van Aalst, MK, Kron, W, Ward, PJ. (2015) Declining vulnerability to river floods and the global benefits of adaptation. Proceedings of the National Academy of Sciences. May 5 ;112 (18): E2271-80.

Kaski, S, & Kohonen, T. (1996). Exploratory data analysis by the self-organizing map: Structures of welfare and poverty in the world. Proceedings of the third international conference on Neural Networks in the Capital Markets.

Katz, RW, Parlange, MB, & Naveau, P. (2002). Statistics of extremes in hydrology. Advances in water resources, 25(8), 1287-1304.

Kiviluoto, K. (1996). Topology preservation in self-organizing maps. IEEE International Conference on Neural Networks.

Kohonen, T. (2001). Self-organizing maps, Volume 30 of Series in Information Sciences.

Kreimer, A, Arnold, M, & Carlin, A. (2003). Building safer cities: the future of disaster risk: World Bank Publications.

Kummu, M, De Moel, H, Ward, PJ, & Varis, O. (2011). How close do we live to water? A global analysis of population distance to freshwater bodies. PLoS One, 6(6), e20578.

Kunkel, KE, Pielke Jr, Roger A, & Changnon, SA. (1999). Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: A review. Bulletin of the American Meteorological Society, 80(6), 1077.

Meehl, GA, Arblaster, JM, & Tebaldi, C. (2005). Understanding future patterns of increased precipitation intensity in climate model simulations. Geophysical Research Letters, 32(18).

Mills, G. (2007). Cities as agents of global change. International Journal of Climatology, 27(14), 1849-1857.

Milly, PCD, Dunne, KA, & Vecchia, AV. (2005). Global pattern of trends in streamflow and water availability in a changing climate. Nature, 438(7066), 347-350.

Milly, PD, Wetherald, RT, Dunne, KA, & Delworth, TL. (2002). Increasing risk of great floods in a changing climate. Nature, 415(6871), 514-517.

Muis, S, Güneralp, B, Jongman, B, Aerts, JC, Ward, PJ. (2015) Flood risk and adaptation strategies under climate change and urban expansion: A probabilistic analysis using global data. Science of the Total Environment. 538: 445-57.

Nature. (2016). Waters encroaching. Nature Climate Change, 6(7), 635, editorial.

Nicholls, RJ, Hanson, S, Herweijer, C, Patmore, N, Hallegatte, S, Corfee-Morlot, J, Muir-Wood, R. (2008). Ranking port cities with high exposure and vulnerability to climate extremes.

Pachauri, RK, Allen, MR, Barros, VR, Broome, J, Cramer, W, Christ, R, ... & Dubash, NK. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (p. 151). IPCC.

Revi, A, Satterthwaite, DE, Aragón-Durand, F, Corfee-Morlot, J, Kiunsi, RBR, Pelling, M, Roberts, DC & Solecki, W. (2014). Urban areas. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. United Kingdom and New York, NY, USA: Cambridge University Press, Cambridge.

Samir, KC, & Lutz, W. (2014). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. Global Environmental Change.

Schiermeier, Q. (2011). Increased flood risk linked to global warming. Nature, 470(7334), 316.

Shanmuganathan, S, Sallis, P, & Buckeridge, J. (2006). Self-organising map methods in integrated modelling of environmental and economic systems. Environmental Modelling & Software, 21(9), 1247-1256. doi: 10.1016/j.envsoft.2005.04.011.

Sheppard, SRJ. (2005). Landscape visualisation and climate change: the potential for influencing perceptions and behaviour. Environmental Science & Policy, 8(6), 637-654.

Skupin, A, & Hagelman, R. (2005). Visualizing demographic trajectories with self-organizing maps. GeoInformatica, 9(2), 159-179.

Sofia, G, Roder, G, Dalla, FG, Tarolli, P. (2017) Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. Scientific reports.

Ultsch, A. (2003). U-matrix: a tool to visualize clusters in high dimensional data. Marburg: Fachbereich Mathematik und Informatik.

UN-DESA. (2015). World Urbanization Prospects: The 2014 Revision. New York: United Nations Department of Economic and Social Affairs.

UNEP. (2016). Summary of the sixth global environment outlook regional assessments: Key findings and policy messages. United Nations Environment Programme.

UN-Habitat. (2010). State of the world's cities. Earthscan.

UN-Habitat. (2014). Urban Equity in Development - Cities for Life. World Urban Forum 7. April 2014, Medellin, Colombia.

Václavík, T, Lautenbach, S, Kuemmerle, T, & Seppelt, R. (2013). Mapping global land system archetypes. Global Environmental Change, 23(6), 1637-1647.

Vesanto, J, & Alhoniemi, E. (2000). Clustering of the self-organizing map. IEEE Transactions on Neural Networks, 11(3), 586-600. doi: 10.1109/72.846731.

Ward, JHJr. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Ward, PJ, Jongman, B, Weiland, FS, Bouwman, A, van Beek, R, Bierkens, MF, ... & Winsemius, HC (2013). Assessing flood risk at the global scale: model setup, results, and sensitivity. Environmental research letters, 8(4), 044019.

Wasko, C, & Sharma, A. (2015). Steeper temporal distribution of rain intensity at higher temperatures within Australian storms. Nature Geoscience 8.7: 527-529.

Wentz, FJ, Ricciardulli, L, Hilburn, K, & Mears, C. (2007). How much more rain will global warming bring? Science, 317(5835), 233-235.

Willems, P, Olsson, J, Arnbjerg-Nielsen, K, Beecham, S, Pathirana, A, Gregersen, IB, & Madsen, H. (Eds.). (2012). Impacts of climate change on rainfall extremes and urban drainage systems. IWA publishing.

Winsemius, HC, Van Beek, LPH, Jongman, B, Ward, PJ, & Bouwman, A. (2013). A framework for global river flood risk assessments. Hydrology and Earth System Sciences, 17(5), 1871-1892.

Winsemius, HC, Aerts, JC, van Beek, LP, Bierkens, MF, Bouwman, A, Jongman, B, Kwadijk, JC, Ligtvoet, W, Lucas, PL, van Vuuren, DP, Ward, PJ. (2016) Global drivers of future river flood risk. Nature Climate Change, 6(4): 381-5.

Young, AF. (2013). Urban expansion and environmental risk in the São Paulo Metropolitan Area. Climate Research, 57(1), 73-80.

Websites:

1. **Atlas of Urban Expansion**: http://www.lincolninst.edu/subcenters/atlas-urban-expansion/Default.aspx
2. **World Bank's World Development Indicators database**: http://data.worldbank.org/data-catalog/world-development-indicators
3. **World Resources Institute's Aqueduct Global Flood Analyzer Tool**: http://floods.wri.org/#/
4. **ISIMIP**: https://www.pik-potsdam.de/research/climate-impacts-and-vulnerabilities/research/rd2-cross-cutting-activities/isi-mip/about

5. **Shared Socioeconomic Pathways**: https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about
6. **World Resources Institute's Aqueduct Global Flood Risk Country Rankings**: http://www.wri.org/resources/data-sets/aqueduct-global-flood-risk-country-rankings
7. **SOM Toolbox**: http://www.cis.hut.fi/somtoolbox
8. Scientific American: https://www. scientificamerican.com/article/extreme-rain-may-flood-54-million-people-by-2030/
9. **United Nations Framework Convention on Climate Change**: www4.unfccc.in

7 COMPETING INTERESTS

The authors declare that they have no conflicts of interest