Referee #1

The study looks at the spatial accessibility of emergency medical service facilities in Beijing resolved according to weather situations, days and time-of-the-day. The authors demonstrate an empirical approach for linking spatially resolved accessibility decreases to weather situations, and manage to point out spatial inequalities on a sub-urban level. The approach, despite the reviewer having some reservations regarding certain over-simplifications along the way, shows a way forward to combine empirical data to yield insights into an under-studied topic (EMS accessibility) while shedding light on a dimension of social inequity.

General

1. L 28 not clear at that stage what is meant with « coverage rate ». Is it the total area covered within a 15 mins response time? i.e. 13 % reduction refers to a km2 number? why then « rate » (I do know that it is explained later on in the text, however the abstract should be understandable before this)?

Response: We thank for reviewer's comment. Using "coverage rate" in the abstract may cause confusion to readers, in the revised manuscript, we have rephrased the sentence in Lines 27-29, Page 2:

Under inclement weather scenario, the area in the citywide that could get EMSs within 15 minutes would decreased by 13% compared to normal scenario. And in some suburban townships, the population that could get 15-min EMS would decrease by 40%.

2. L 45-55: It is good that the authors give detailed insight into what parts make up EMS, especially tailored to the case study context. However, there seems to be - strictly speaking - some inconsistency in the use of the term services: One case includes the transport to the EMS facilities within the service definition (aka, the transportation service), the other case not (only treatment service).

Further, please elaborate how this goes together with the definition in L 72-74: It seems that this definition covers only the case of responders starting out from an EMS facility, getting to the scene, and transferring back to a respective facility (e.g. via ambulance). From the initial description above, the reader might think the authors will cover, however, also the case where patients transfer directly from the scene to an EMS facility (e.g. via private transportation). Hence, it might be good to explicitly mention again that this latter case will not be covered, even though described above for reasons of completeness.

Response: We thank for reviewer's comment. The EMSs includes both pre-hospital and in-hospital services, and in our study, we focus on the accessibility of EMSs (only the transportation part). And we divide the process into 2 parts: one is starting out from an EMS facility and getting to the scene (as in 3.2), the other is transferring the patient from the scene to hospitals (via ambulance, not private transportation, as in 3.3). So, we assume that the patients get the EMSs when the ambulance arrives the scene, and the EMSs is complete when the ambulance transfers the patient to hospital, and both parts were covered in our study. The case where patients transfer directly from the scene to an EMS facility via private transportation was not covered in this study. We apologize for the unclarity, and we mentioned it in the revised manuscript in Lines 193-195, Page 9:

The case where patients transfer directly from the scene to an EMS facility via private transportation will not be considered in this study.

3. L 58: To the non-local reader, it is unclear whether 1.5 to 2 hours response time is significantly longer than normally. One would assume this, but it would be helpful to provide average response times during normal conditions for comparison.

Response: We thank for reviewer's comment. As suggested by the reviewer, we have added the comparison of normal condition in Lines 61-62, Page 3. And this information comes from the same news.

while the transfer time wouldn't be more than 1 hour on usual

4. L 78: Reference to the 2SFCA method?

Response: We thank for reviewer's comment. We have added the reference to 2SFCA methods in Lines 79-80, Page 4:

Chen, X., and Jia, P.: A comparative analysis of accessibility measures by the two-step floating catchment area (2sfca) method, INT J GEOGR INF SCI, 33, 1739-1758,2019.
Kanuganti, S., Sarkar, A. K., and Singh, A. P.: Quantifying accessibility to health care using two-step floating catchment area method (2sfca): a case study in rajasthan, Transportation Research Procedia, 17, 391-399,2016.
Li, M., Kwan, M., Chen, J., Wang, J., Yin, J., and Yu, D.: Measuring emergency medical service (ems) accessibility with the effect of city dynamics in a 100-year pluvial flood scenario, CITIES, 117, http://doi.org/10.1016/j.cities.2021.103314, 2021b.
Luo, W., and Qi, Y.: An enhanced two-step floating catchment area (e2sfca) method for measuring spatial accessibility to primary care physicians, HEALTH PLACE, 15, 1100-1107,2009.

5. L 74-L104: it is good that you provide some literature revolving around the topic of study. However, this very brief listing-style does not make it very clear how those papers are related to the concrete problem statement, or not summarized into coherent areas of challenges, seeming a bit randomly aggregated.

Response: We thank for reviewer's comment. We recomposed the whole part of review in the revised manuscript, please refer to Lines 73-105, Pages 4-5:

The spatial accessibility of EMSs is defined by the travel impedance (distance or time) between *service locations and the scene (Guagliardo, 2004). A large body of research on spatial accessibility is concerned with access to hospitals (Luo and Wang, 2003; Mao and Nekorchuk, 2013; Pan et al., 2018; Yang et al., 2020; Yin et al., 2021) and first-aid stations (Hashtarkhani et al., 2020; Jones and Bentham, 1995; Shin and Lee, 2018). To measure the EMSs accessibility, the two-step floating catchment area (2SFCA) method is one of the common methods (Chen and Jia, 2019; Kanuganti et al., 2016; Li et al., 2021; Luo and Qi, 2009). The 2SFCA method considers accessibility to be mediated by not only the distance decay but also the interactions between supply and demand (Chen and Jia, 2019), which is more suitable for normal scenarios. While for studies focusing on the influence of inclement weather on EMSs, people concern more about the transportation situation, instead of the interaction between supply and demand. The coverage analysis method (Coles et al., 2017; Green et al., 2017; Yu et al., 2020) or shortest path analysis method(Albano et al., 2014; Andersson and Stålhult, 2014) are more widely used. These methods could better characterize the reduction of accessibility caused by the road service degradation. For example, Yu et al. (2020) analyzed the accessibility of emergency service in England and identified vulnerability hotspots by quantifying the EMSs coverage of area and population within different time radii under different flood scenarios; Coles et al. (2017) measured the travel time and service area coverage of EMSs in York, UK under flood scenarios by using FloodMap-HydroInundation2D to model flood inundation; Yin et al. (2021) assessed the vulnerability of EMSs to surface water flooding in Shanghai, China by quantifying accessibility in terms of service area, response time, and population coverage, considering four temporal scenarios (nighttime, evening, daytime, and morning–evening peak) of average drive speeds based on a real-time traffic analysis from GPS big data; Andersson and Stålhult (2014) used network analysis methods to generate the shortest paths from hospitals to various administrative areas in Manila, Philippines, and evaluated the impact of different flood events on these paths. Most of these studies assumed that roads are impassable or traffic speed has a certain degree of reduction when the flooded water depth reaches a specific depth, and further evaluated the impact of rainstorm on EMSs accessibility. Due to insufficient recorded traffic data, relatively few studies have been performed to analyze the impact of road access capacity on EMSs accessibility according to actual traffic speed variation.*

6. Section 3.1:

Could you please elaborate the reason for averaging to daily speeds for the baseline constructions, since you later also look at rush-hours and non-rush-hours specifically?

In a similar line of argumentation, averaging hence-obtained speed reduction rates across all road sections within the city (L220f), seems to conceal congestion hotspots? Please elaborate why this was done and the potential limitations of this.

Also, it is not clear to me from this description, if days are simply divided in a binary manner into inclement and non-inclement weather days, irrespective of the precipitation magnitude? Please elaborate in more detail if this is the case, and what was the reasoning for and potential shortcoming of this.

Response: We thank for reviewer's comment. Actually, we averaged the traffic speed data of the selected period, not of the whole day (as in line 219 and line 221). We apologize for the unclarity, and we added "in the selected period" in Lines 207-215, Pages 10:

The average speed data of the four baseline days ***in the selected period*** *were then averaged as the baseline speed for the given precipitation day, and the traffic speed reduction rate was calculated by eq. (1):*

$$r_c = \frac{v_p - \frac{\sum_{j=0}^{m} v_{d_j}}{m}}{\frac{\sum_{j=0}^{m} v_{d_j}}{m}} \qquad (1)$$

where r_c *is the traffic speed reduction rate* ***in the selected period*** *of the precipitation day to its corresponding baseline days;* v_p *is the traffic speed* ***in the selected period*** *of the given precipitation day;* v_{d_j} *is the traffic speed* ***in the selected period*** *of a baseline day, and m is the number of baseline days. In this case, m equals 4.*

We admit that averaging hence-obtained speed reduction rates across all road sections within the city do conceal the congestion hotspots, which is a limitation of this study. This limitation could be improved in the further study on a small scale based on a higher resolution of precipitation. Nevertheless, averaging could give us the overall impact of precipitation on vehicle speed across the city. And we put Figure 4 to complement the spatially difference and find out the congestion hotspots. We also admit this limitation in Lines 487-492, Pages 25:

However, there are also some limitations in this study. First, we averaged the traffic speed reduction rate of all the roads in the city, as well as the precipitation data, which could conceal congestion hotspots. In further studies, with higher resolution precipitation, along with corresponding traffic data, we could narrow the scale to blocks, pay more attention to local congestions, and analyze the correlation of precipitation and traffic speed on a finer scale.

In this study the days were simply divided in a binary manner into inclement and non-inclement weather days, the reason is that we only got the traffic data of 2019, and there are few days with precipitation. The data does not support an analysis considering precipitation magnitude. In the revised manuscript, we have acknowledged this shortcoming in the discussion section in lines 499-492, page 25:

Due to the data limitation, we could only analyze the EMSs accessibility in 2019, and the precipitation intensity in this year was not quite high. If with more precipitation and traffic data, we could analysis the impact of precipitation magnitude to the traffic and accessibility, instead of divided the days in a binary manner into inclement and non-inclement weather days.

7. Section 3.3: How does aggregation of the population grids to 1000m in a city distort potential travel patterns? Given that the topology of the road network within Beijing is at a much higher resolution, does this aggregation not lead to a very coarse estimation of what roads are taken, and which ones not?

Response: We thank for reviewer's comment. When we run the OD cost matrix analysis for each population grid, we use its centroid as the origin point, not the whole square, so it won't affect the resolution of the topology of the road network. This might be interpreted as a sampling survey. The choice of roads to the nearest hospital won't be greatly impact because the distance to the nearest to the nearest hospital is normally more than 1km in most areas. The reason of aggregation the population grids are mainly reducing the amount calculation, because the OD cost matrix analysis is computationally intensive. We added some explanation in Lines 256-258, Page 12:

This could be interpreted as a sampling method, because we use the centroid point of the grid to represent the other possible starting points in the grid, and we ignored the tolerance caused by the travel time inside the grids.

8. Section 4. Results:

L275-280: It is hard to understand to which scenarios / analysis areas the percentages belong. Do 38 and 40%, resp., refer to the city including suburban areas? And 77 and 83% refer to only the inner city (is that meant by Six Road Ring)? Please phrase it in a way that describes the area analyzed better to a non-local.

Response: We thank for reviewer's comment. The reviewer's comprehension is correct. And we revised the sentences in lines 273-279, page 13:

We divided the city into the inner city and suburban areas along the Sixth Ring Road. Taking the average 15-minute coverage of the area of all Mondays in November as an example: (1) in the whole city (both inner city and suburban), the coverage rate of EMSs is 38.72% in morning rush hours, compared with 40% (±0.3%) in the remaining periods; (2) in the inner city, the coverage rate is 77.37% in morning rush hours, compared with 83% (±0.6%) in the remaining periods.

9. L 283-288: The definition and selection of a precipitation event belongs to the methodology section.

Response: We agree with the reviewer's suggestion, and we moved this part to 2.2.2 in the revised manuscript. Please refer to Lines 155-161 in Pages 7-8.

10. 4.1.1. When I first read that you average out the total precipitation, across the grid cells, I was skeptical whether this conceals local effects, as one might assume that certain parts of the city could hence flood more, and cause over-proportional traffic delays.

Also, I do not see an analysis of total precipitation on traffic speed, which I could imagine has quite an impact (while it is certainly important how strongly it rains in a given hour, it surely also matters how long it rains for causing pluvial flooding). Please elaborate more on both those aspects.

Please also use figure 3 in justifying your assumptions / method, as indeed, it seems that without the (explained) outliers, there seems to be not much of an impact on how much it rains for causing travel reductions? This may be an argument in favor of your decision; however, it is somewhat unintuitive why there is so little effect.

In general, please comment more on the relationship between precipitation and urban pluvial flooding, and limitations of looking only at precipitation data without any hydrological modelling associated with it, that would link precipitation with the actual amount of water on the streets.

Response: We thank for reviewer's comment. First, we averaged the total precipitation because we intended to evaluate the correlation of precipitation and overall traffic speed reduction in the city. And the spatial difference of precipitation was illustrated in Figure 4. We admit that could lead to conceal the local effects. This limitation was mainly due to the lack of high-resolution precipitation data. There are only 175 grids of precipitation data in Beijing, and we couldn't analyze the local effect on that scale. And we added more discussion on that in lines 487-492, page 25:

There are also some limitations in this study. First, we averaged the traffic speed reduction rate of all the roads in the city, as well as the precipitation data, which could conceal congestion hotspots. In further studies, with higher resolution precipitation, along with corresponding traffic data, we could narrow the scale to blocks, pay more attention to local congestions, and analyze the correlation of precipitation and traffic speed on a finer scale.

Second, we apologize for the ambiguity of the precipitation unit. Actually, the index of precipitation we used is the total precipitation during the selected time period, instead of the intensity of precipitation. Because the unit of precipitation we took was mm/2h, and in case that we only focus on the morning rush hours period (7:00-9:00), the index represents the total precipitation in these 2 hours. To avoid confusion, we mentioned it again in lines 287-288, page 13:

The unit of precipitation data is mm/2h, which indicates the total precipitation in the 2 hours of morning rush hours.

Besides, how long it rains was hard to be accurate because the time resolution of precipitation is 0.5 hours, and the morning rush hours period is only 2 hours in total.

Third, Figure 3 shows that there are 4 days that has more speed reduction, and the rest days seems relatively normal. We think the reason of this result is that the days with heavy rain or snow were rare. But when the small rain encountered the specific date, such as teacher's day, the impact would be amplified. The limitation is that we only have one-year traffic data, we believe with a longer time series data, the further results could be more significant.

Fourth, we agree that run a hydrological pluvial flood simulation could better link the precipitation, waterlogging and traffic congestion. However, run a flood simulation need a high-resolution DSM data of the city, which is very hard to get in China. We therefore link the precipitation and traffic directly with the analysis of ground-truth data. And we add some discussion about this in lines 500-503, page 25-26:

Due to the lack of high-resolution DSM data, we didn't run a hydrological flood simulation in Beijing, which could reveal the relationship of precipitation and the actual amount of water on the streets. And this could be improved in the future studies with more high-resolution topographic data.

11. Figure 4 is basically not commented and further analyzed. Please elaborate more, what one can see.

Response: We thank for reviewer's comment. We have elaborated the description of Figure 4 in the revised manuscript in Lines 296-306, Pages 14.

Figure 4 illustrates the spatial difference of traffic speed reduction and distribution of precipitation on precipitation days. A large number of red roads (with traffic speed reduction over 10 km/h) can be observed in the 4 days mentioned above. By comparing the distribution of precipitation and traffic speed reduction on different dates in Figure 4, it can be found that the precipitation in the four days with the most severe speed reduction was moderate, and the precipitation distribution of the whole city was relatively uniform. Compared with other rain days, although the precipitation on July 5, August 9 and September 19 was larger and concentrated in the inner city, the traffic speed reduction of the whole city was not as serious as the four days mentioned above, which may be caused by the decrease of people's willingness to travel with the increase of rain.

Grammar / Style

1. In-line citations are ill-formatted (brackets around them), e.g. L 83, L 90, etc. Please format correctly. Also, some citations are CAPITALIZED.

Response: We thank for reviewer's comment. We have checked and corrected the format of citations.

2. L31: towns new sentence : Furthermore, …

Response: We thank for reviewer's comment. We have corrected the gramma mistake.

3. L 106: Could you briefly explain the term "waterlogging" (e.g. the saturation of ground with water), as it may not be clear to every single reader what is meant by this phenomenon.

Response: As suggested by the reviewer, we have added the explanation in Lines 68-69, Page 4 waterlogging (the phenomenon of a stagnant water disaster in an urban area due to heavy rainfall or continuous precipitation).

4. L145: Rather: "hit" by a rainstorm? They do not malevolently "attack".

Response: We thank for reviewer's comment. We have changed the verb from "attack" to "hit". (Line 56, Page 3)

5. L 145-150: You already gave quite a few examples of hazardous events in the introduction; also, this example does not fit the section "Study area". Please consider to delete it.

Response: We thank for reviewer's comment. We delete it.

6. Section labels are wrong. For instance, after 2.1 and 2.2. on p7, we see another section 2.1 and 2.2. on p8

Response: We thank for reviewer's comment. We have corrected the section labels.

7. L261: The term "population medical accessibility index" is a little bit cumbersome to read and understand. Could you perhaps think of a simpler, more descriptive term?

Response: We thank for reviewer's comment. We have changed the term to "total transfer time" in Lines 259-261, Page 12:

The total transfer time is introduced to quantify the cumulative transfer time for each population grid based on its population size, which is the number of potential users of EMSs.