**Reviewer #1**

The authors propose a very interesting adjustment to the 2SFCA method. I’m still unsure that these adjustments are the panacea as suggested by the authors (i.e. inflation issues are ‘eliminated’) – I’d like to do my own testing (in the future); however, they are a worthy addition to the literature for further academic debate.

First of all, many thanks for your thoughtful response to our submission. If you are interested in conducting testing of your own, we would like to remark that the data and code are placed in a public GitHub repository as follows:

<https://github.com/paezha/Demand-and-Supply-Inflation-in-Floating-Catchment-Area-FCA-Methods->

In what follows we respond to your comments and indicate the actions that we have taken in response.

One fundamental problem throughout the paper is their referral to step 1 as if is measuring the actual or real demand on that supply point. Clearly that is not the case because, in most scenarios, populations will have multiple supply points that are accessible and thus their demand will be ‘averaged’ across these.

Thank you for this comment. This is something that we have clarified. “Demand” in a model of accessibility represents potential demand – since we do not really know how many members of the population actually have a family doctor. Even in a universal care system such as Canada, there are members of the public who do not have a regular health care provider – approximately 15.8% of the population in 2016, for example (Statistics Canada, 2017). Assuming, however, that all members of the population *do* demand health care (universal coverage as the design parameter, so to speak), one thing that becomes clear from our approach is that the “averaging” that intuitively takes place across the multiple supply points tends to be biased in earlier FCA approaches.

Typically the 2SFCA is measuring ‘potential demand’, but the authors effectively are presenting it as ‘actual demand’ (when in reality this is far from true). This very important distinction needs clarification upfront. The authors present an adjustment method that I distinguish as recalculating the ‘likely demand on supply points’. Similarly, accessibility (Step 2) is ‘averaged’ across multiple supply points. The effect of demand inflation and supply inflation to the 2sfca method really should be looking at the combined effect – i.e. at the Ai (accessibility) scores.

Many thanks for this comment. Indeed, in addition to clarifying the distinction between potential access (accessibility) and actual access, we note that our approach assumes that the system operates in an efficient way, by fully allocating both demand and level of service proportionally. To allow for system inefficiencies, we suggest the use of “slack factors”, similar to your suggestion of “likely demand on supply points”. Furthermore, we show with the example how the combined effect in the accessibility scores is more intuitive when demand and level of service are rectified to their global levels by means of our standardized weights.

Their simple examples in Figure 1 and 2 are highly extreme (wholly un-real) scenarios, which are very effective at demonstrating how the 2sfca method can fail. However, their extremeness also makes it hard for the reader to evaluate whether, in a more ‘normal’ scenario, the same deficiencies still exist or are a significant concern. I believe these examples need to introduced as such or similar: “These examples have been purposefully designed to demonstrate key weaknesses of the 2sfca method; we acknowledge these scenarios are not representative of typical demand and supply landscapes”

In this version of the paper, we recognize the artificial simplicity of the original example, and introduce a small but realistic simulation to complement our presentation.

Overall, I like the concept of the demand vector summing to equal the original population (demand) size – even after the application of impedance weights. In theory, this enables the generation of accessibility scores that more truly are interpretable ratios (e.g. health providers per population). Having said that, their adjusted weights – which divide the original weight by the sum of all weights for (potential) service points – still assumes that all service points are equally ‘attractive’. I strongly believe this is not the case at urban fringe locations where larger urban populations are unlikely to be attracted to nearby rural services, but the reverse scenario is very different.

In fact, points are equally attractive only when binary weights are used as the basis for obtaining a standardized set of weights. When the basis weights are from an enhanced (i.e., non-binary) impedance function, the attractiveness of more distant points declines, and more demand is allocated proportionally to closer points (same with the allocation of level of service). We hope that the new examples will help to clarify this point.

Similarly the concept of saturation of intervening opportunities and variable catchment sizes seem pertinent here (i.e. why would a population travel further if sufficient opportunities are available at nearer locations?). Some recent 2sfca methods papers have introduced variable catchment size modelling, which I believe are critical for larger geographic scale applications (e.g. for national modelling).

Our toy example and small simulation indeed assume a heterogeneous and isotropic landscape, and as a consequence all catchment areas are of fixed sizes. The empirical application, in contrast, is based on network distance, and therefore lead to adaptive catchment areas.

As seen in Figures 5 and 6, the demand inflation factor is mostly a factor of urbanized populations only. I didn’t understand the Accessibility scores in Figures 7 and 8 or 11 and 12 – how do you get an Accessibility ratio of 1 (i.e. equal supply/demand ratio)?

In the revised version of the paper, we clarify that accessibility is interpreted as the share of provider-to-population ratio allocated to a population center. Therefore, it would mean supply (i.e., doctors) divided by demand (number of patients).

Again, the values in Figure 9 and 13 make no sense to me. Are the authors truly claiming that ‘current’ accessibility scores (e.g. 2sfca, e2sfca) are inflated 4 to 9 times higher than they should be? In all of the published work on floating catchments, I’ve never seen any evidence that even vaguely suggests they are out by this factor level.

We have added information to explain this. At first glance, a factor of 4 to 9 is startling. However, when we compare the accessibility estimates obtained from current FCA methods, it turns out that they indeed tend to exceed the regional PPR by a very large factor (see Table 6 in new version of paper).

Figures 14 and 15 are very interesting in that they suggest a very systematic bias between metropolitan and fringe/rural locations (strongest reds in the most metropolitan, strongest blues in the outer edges) – I suggest this is a key finding that needs further discussion.

We have revised the discussion to respond to this.

Again, we thank you for your constructive and thoughtful comments, and hope that you will find our responses and actions in revising the paper satisfactory.

**Reviewer #2**

This is a very interesting paper that presents an issue that I've grappled with in the past. Per PLOS ONE's reviewer policy, I'm going to unblind myself in an effort to promote a constructive dialog (and because the content of my review would likely do this anyways).

Many thanks for your comments and suggestions for revisions. This work was in good measure inspired by Delamater (2013), and so we are particularly pleased to receive your input.

Major concerns/issues/comments

L89: This equation is somewhat confusing as presented. Are these all from different metrics?

These are the same metric, just decomposed in different ways. For instance, D\_j = sum\_i D\_ij.

L104-105: Could this also be conceptualized as the cost required to overcome distance?... e.g., rather than making it about preferences, it seems like this passage should acknowledge that overcoming distance/time is a “cost” of using the service, rather than just framing it as preference-based.

We revised this as per your suggestion.

L347-349: I feel like I'm missing something here. The overall ratio of the system does not appear to be preserved, given my understanding of what the numbers represent. I think that they are (for 2SFCA): Prov/Pop ratio for A = 0.066 and for B = 0.006. In that scenario, the overall “system” ratio would be 0.46 (not 0.02) in this calculation. (If I'm interpreting the output incorrectly, I apologize. However, if that is the case, I think it would be extremely helpful to extend the example calculation to clearly demonstrate the output values and how the overall system ratio is preserved in the resulting values)

We have completely reworked the examples, and we hope that they will be clearer now.

Using the row standardized weight matrix to “allocate” the population to the various facilities in the first step appears to be very similar to the example I provided in Delamater, 2013 (citation 20 in the ms) on pages 33 and 34 and illustrated in System's #4 and #5. In this case, I was concerned about the overall effect of splitting populations into discrete “pieces” and it appears that this approach might have similar issues. If I'm understanding the authors' approach correctly, I think that the 2SFCA calculation for the Figure 1(I) system would result in an accessibility value of (0.03 + 0.03 + 0.03) = 0.09 facilities per person. This output doesn't really make sense to me, purely from a logical perspective – and it would concern me that the approach does not perform well in the most simple of systems.

We have expanded the discussion and added summaries of the results to illustrate the way different implementations of FCA operate. Hopefully you will be persuaded that our approach to proportionally allocate supply and level of service provides a more accurate representation of provider-to-population ratios, and therefore are easier to interpret.

My overall thought is that the problem of supply and demand overestimation might be “problem” when each is considered in a vacuum, but not in the output/results when the two steps are integrated. I would suggest that the Demand and Supply Inflation section be expanded to show more simple examples – and most importantly how inflation actually effects the final sp accessibility values (A) in these systems, as a “proof of concept”. As of right now, I can clearly see the issue in the two separate steps, but I'm having a hard time understanding what the demonstrated effects of this are on the final outcome [e.g., is A (sp accessibility) under/over estimated for Pop A?... Pop B?]. This was one of my concerns when I critiqued the Wan et al 3SFCA paper in Delamater, 2013... I was not sure that competition really “needed” to be accounted for to adequately capture potential sp accessibility. In this case, I'm not sure if the inflation happening in the two steps is having a detrimental effect on the results. I think that clearly demonstrating issues with the output calculations of the E2SFCA and 2SFCA (in more than a single system) would provide a much better justification for this approach (than is presented currently).

Thank you for your suggestion. We have reworked the paper to include examples of 3SFCA and M2SFCA. As you correctly point out, it is important to think whether the final outcome (the accessibility) is affected by the potential inflation, and if so how. Currently, our understanding is that 3SFCA deflates demand by stacking a second impedance function that is standardized to allocate proportionally the results of E2SFCA. This is done again when allocating the level of service in the accessibility calculations. As our examples show, this deflates both demand and level of service, but in an ad-hoc way that is not consistent with the global demand or level of service. M2SFCA, on the other hand, deflates the level of service by squaring the impedance function – in other words, by shrinking the catchment areas. The standardized weights that we provide do not affect the catchment areas, but rather ensure that both demand and level of service are allocated proportionally within them.

We recognize the possibility of suboptimal and/or inefficient systems, following in part the discussion in Delamater (2013). To this end, we discuss the selection of the base impedance weights and the use of what we call “slack factors” that can modulate the allocation of demand and/or level of service in a simple and intuitive way.

I did not do a deep dive into the Results yet, given my concerns about the approach. However, I do wonder if using a ratio to describe inflation is a good approach, especially given that the output has true units (people and supply/people), e.g., from 0.002 to 0.004 and 2 to 4 would have a similar inflation factor... but would have very different absolute differences.

In case that you would like to do a deep dive on the analysis, the manuscript, code, and data files to reproduce the paper are publicly available at:

<https://github.com/paezha/Demand-and-Supply-Inflation-in-Floating-Catchment-Area-FCA-Methods->

We thank you again for your thoughtful and constructive comments and look forward to hearing back from you with your impressions regarding this new version of the paper.