THE ROYAL SOCIETY
PUBLISHING

ROYAL SOCIETY
OPEN SCIENCE

Using Fisher information to assess stability in the performance of public transportation systems

Nasir Ahmad, Sybil Derrible and Heriberto Cabezas

Review History

RSOS-160920.R0 (Original submission)

Review form: Reviewer 1

Is the manuscript scientifically sound in its present form?
Yes

Are the interpretations and conclusions justified by the results?
Yes

Is the language acceptable?
Yes

Is it clear how to access all supporting data?
Yes

Do you have any ethical concerns with this paper?
No

Have you any concerns about statistical analyses in this paper?
No

Recommendation?
Major revision is needed (please make suggestions in comments)

Comments to the Author(s)
This article applies Fisher Information to the study of transportation systems across the United States. I was initially very intrigued by the approach and its potential, but ultimately had trouble figuring out exactly what benefits this method brings over other methods of analyzing transit systems.
My main critique of the article, therefore, is that the utility from a transportation planning perspective is poorly motivated. How does this method compare or improve upon other more commonly used metrics? What is the benefit of this method as compared to other trend clustering methods (i.e. latent curve modeling) for analyzing time series data? I understand that it can be used for multiple variables, but the limitation I see is that we can't discern which variable is causing the change? Is it a supply or demand problem?

Overall, I didn't find that the method revealed anything unknown or extraordinary about US transit system. Rather, it confirmed in a very descriptive manner rather obvious trends about these systems (such as, ridership changes follow the opening of a new light rail line).

The authors really need to convey more convincingly that something novel was uncovered about transit systems that could not have been detected with other, previously used metrics.

Review form: Reviewer 2 (Valerio De Martinis)

Is the manuscript scientifically sound in its present form?
Yes

Are the interpretations and conclusions justified by the results?
Yes

Is the language acceptable?
Yes

Is it clear how to access all supporting data?
Yes

Do you have any ethical concerns with this paper?
No

Have you any concerns about statistical analyses in this paper?
No

Recommendation?
Accept with minor revision (please list in comments)

Comments to the Author(s)
In this paper, the authors propose a method to assess the overall stability and order of public transit systems over time using elements of Information Theory and adopting Fisher Information (FI) concepts. Results shown on a very large dataset identifies eight different patterns in the calculation of FI and assesses an overall stable state of public transport systems or, in some cases, a search towards this stable state.

The work is overall well written and well organized. Only a few comments from my side to be considered as minor reviews.

•   The introduction is very clear, however, it must be improved clarifying the main innovative contributions to the field (a couple of sentences or a small paragraph at the end of section 1 would be enough).

•   At the end of the literature review, the research gap to fill/the alternative method you proposed together with possible benefits (e.g. accuracy, computation speed, easy implementation) should be clearly stated, in order to position your paper in the current literature. Since you already published something related to FI, it would be also desirable to add a couple of sentences on the advances of your research in the proposed paper starting from reference [13].

•   The link https://github.com/csunlab/fisher-information , could it be added as a reference? Since it is absolutely that. But you can check it during the editorial process for publication.

•   My opinion is that section 3 (very short and still part of the "background" needed for readers) can be merged with section 2, with a small introduction to the Sustainable Regimes Hypothesis.

•   In the discussion section, mainly trends descriptions are reported. I would add also some specific explanations. Why rail over the other transport modes? why the jump-drop of NY in 2004-09? What's behind that?

•   Another important issue to consider is the role of the latest technology. Probably the discussion on autonomous vehicles is not yet influencing PTS dynamics, since very few applications around the world are available (e.g. Autonomous Post bus in Sion, Switzerland). It would be interesting to have also some consideration on this aspect, your expectation, etc.

Decision letter (RSOS-160920)

09-Mar-2017

Dear Mr Ahmad

On behalf of the Editors, I am pleased to inform you that your Manuscript RSOS-160920 entitled "Assessing Performance Stability in Public Transportation Systems using Fisher Information" has been accepted for publication in Royal Society Open Science subject to minor revision in accordance with the referee suggestions. Please find the referees' comments at the end of this email.

The reviewers and handling editors have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the comments and revise your manuscript.

• Ethics statement
If your study uses humans or animals please include details of the ethical approval received, including the name of the committee that granted approval. For human studies please also detail whether informed consent was obtained. For field studies on animals please include details of all permissions, licences and/or approvals granted to carry out the fieldwork.

• Data accessibility
It is a condition of publication that all supporting data are made available either as supplementary information or preferably in a suitable permanent repository. The data accessibility section should state where the article's supporting data can be accessed. This section should also include details, where possible of where to access other relevant research materials such as statistical tools, protocols, software etc can be accessed. If the data has been deposited in

an external repository this section should list the database, accession number and link to the DOI for all data from the article that has been made publicly available. Data sets that have been deposited in an external repository and have a DOI should also be appropriately cited in the manuscript and included in the reference list.

If you wish to submit your supporting data or code to Dryad (http://datadryad.org/), or modify your current submission to dryad, please use the following link:
http://datadryad.org/submit?journalID=RSOS&manu=RSOS-160920

• Competing interests
Please declare any financial or non-financial competing interests, or state that you have no competing interests.
• Authors' contributions
All submissions, other than those with a single author, must include an Authors' Contributions section which individually lists the specific contribution of each author. The list of Authors should meet all of the following criteria; 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published.

All contributors who do not meet all of these criteria should be included in the acknowledgements.

We suggest the following format:
AB carried out the molecular lab work, participated in data analysis, carried out sequence alignments, participated in the design of the study and drafted the manuscript; CD carried out the statistical analyses; EF collected field data; GH conceived of the study, designed the study, coordinated the study and helped draft the manuscript. All authors gave final approval for publication.

• Acknowledgements
Please acknowledge anyone who contributed to the study but did not meet the authorship criteria.

• Funding statement
Please list the source of funding for each author.

Please note that we cannot publish your manuscript without these end statements included. We have included a screenshot example of the end statements for reference. If you feel that a given heading is not relevant to your paper, please nevertheless include the heading and explicitly state that it is not relevant to your work.

Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days (i.e. by the 18-Mar-2017). If you do not think you will be able to meet this date please let me know immediately.

To revise your manuscript, log into https://mc.manuscriptcentral.com/rsos and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions". Under "Actions," click on "Create a Revision." You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referees and upload a file "Response to Referees" in "Section 6 - File Upload". You can use this to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the referees.

When uploading your revised files please make sure that you have:

1) A text file of the manuscript (tex, txt, rtf, docx or doc), references, tables (including captions) and figure captions. Do not upload a PDF as your "Main Document".
2) A separate electronic file of each figure (EPS or print-quality PDF preferred (either format should be produced directly from original creation package), or original software format)
3) Included a 100 word media summary of your paper when requested at submission. Please ensure you have entered correct contact details (email, institution and telephone) in your user account
4) Included the raw data to support the claims made in your paper. You can either include your data as electronic supplementary material or upload to a repository and include the relevant doi within your manuscript
5) All supplementary materials accompanying an accepted article will be treated as in their final form. Note that the Royal Society will neither edit nor typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details where possible (authors, article title, journal name).

Supplementary files will be published alongside the paper on the journal website and posted on the online figshare repository (https://figshare.com). The heading and legend provided for each supplementary file during the submission process will be used to create the figshare page, so please ensure these are accurate and informative so that your files can be found in searches. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Once again, thank you for submitting your manuscript to Royal Society Open Science and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Kind regards,
Alice Power
Editorial Coordinator
Royal Society Open Science
openscience@royalsociety.org

on behalf of R. Kerry Rowe
Subject Editor, Royal Society Open Science
openscience@royalsociety.org

Associate Editor Comments to Author:

Comments to the Author:
The paper has received positive reviews with suggestions for revisions. Please undertake those revisions. However, the first reviewer has expressed concern that the findings related to transportation engineering are not novel. Since that kind of novelty and practical value are not requirements for RSOS papers, you do not need to address those concerns. However, I still encourage you to add novel engineering content if you can (since it will greatly increase the interest in your work and its value).

Reviewer comments to Author:
Reviewer: 1

Comments to the Author(s)
This article applies Fisher Information to the study of transportation systems across the United States. I was initially very intrigued by the approach and its potential, but ultimately had trouble figuring out exactly what benefits this method brings over other methods of analyzing transit systems.
My main critique of the article, therefore, is that the utility from a transportation planning perspective is poorly motivated. How does this method compare or improve upon other more commonly used metrics? What is the benefit of this method as compared to other trend clustering methods (i.e. latent curve modeling) for analyzing time series data? I understand that it can be used for multiple variables, but the limitation I see is that we can't discern which variable is causing the change? Is it a supply or demand problem?

Overall, I didn't find that the method revealed anything unknown or extraordinary about US transit system. Rather, it confirmed in a very descriptive manner rather obvious trends about these systems (such as, ridership changes follow the opening of a new light rail line).

The authors really need to convey more convincingly that something novel was uncovered about transit systems that could not have been detected with other, previously used metrics.

Reviewer: 2

Comments to the Author(s)
In this paper, the authors propose a method to assess the overall stability and order of public transit systems over time using elements of Information Theory and adopting Fisher Information (FI) concepts. Results shown on a very large dataset identifies eight different patterns in the calculation of FI and assesses an overall stable state of public transport systems or, in some cases, a search towards this stable state.
The work is overall well written and well organized. Only a few comments from my side to be considered as minor reviews.
•   The introduction is very clear, however, it must be improved clarifying the main innovative contributions to the field (a couple of sentences or a small paragraph at the end of section 1 would be enough).
•   At the end of the literature review, the research gap to fill/the alternative method you proposed together with possible benefits (e.g. accuracy, computation speed, easy implementation) should be clearly stated, in order to position your paper in the current literature. Since you already published something related to FI, it would be also desirable to add a couple of sentences on the advances of your research in the proposed paper starting from reference [13].
•   The link https://github.com/csunlab/fisher-information , could it be added as a reference? Since it is absolutely that. But you can check it during the editorial process for publication.
•   My opinion is that section 3 (very short and still part of the "background" needed for readers) can be merged with section 2, with a small introduction to the Sustainable Regimes Hypothesis.
•   In the discussion section, mainly trends descriptions are reported. I would add also some specific explanations. Why rail over the other transport modes? why the jump-drop of NY in 2004-09? What's behind that?
•   Another important issue to consider is the role of the latest technology. Probably the discussion on autonomous vehicles is not yet influencing PTS dynamics, since very few applications around the world are available (e.g. Autonomous Post bus in Sion, Switzerland). It would be interesting to have also some consideration on this aspect, your expectation, etc.

Author's Response to Decision Letter for (RSOS-160920)

See Appendix A.

Decision letter (RSOS-160920.R1)

28-Mar-2017

Dear Mr Ahmad,

I am pleased to inform you that your manuscript entitled "Assessing Performance Stability in Public Transportation Systems using Fisher Information" is now accepted for publication in Royal Society Open Science.

You can expect to receive a proof of your article in the near future. Please contact the editorial office (openscience_proofs@royalsociety.org and openscience@royalsociety.org) to let us know if you are likely to be away from e-mail contact. Due to rapid publication and an extremely tight schedule, if comments are not received, your paper may experience a delay in publication.

Royal Society Open Science operates under a continuous publication model (http://bit.ly/cpFAQ). Your article will be published straight into the next open issue and this will be the final version of the paper. As such, it can be cited immediately by other researchers. As the issue version of your paper will be the only version to be published I would advise you to check your proofs thoroughly as changes cannot be made once the paper is published.

In order to raise the profile of your paper once it is published, we can send through a PDF of your paper to selected colleagues. If you wish to take advantage of this, please reply to this email with the name and email addresses of up to 10 people who you feel would wish to read your article.

On behalf of the Editors of Royal Society Open Science, we look forward to your continued contributions to the Journal.

Best wishes,
Alice Power
Editorial Coordinator
Royal Society Open Science
openscience@royalsociety.org

Appendix A

Dear Editor,

We have attached the revised manuscript, "Assessing Performance Stability in Public Transportation Systems using Fisher Information" (RSOS-160920). The reviewers' comments were very helpful in revising the manuscript and our responses are summarized below.

Moreover, we updated the manuscript with the latest datasets available, and the time period studied now extends to 2016.

Reviewer 1

1. My main critique of the article, therefore, is that the utility from a transportation planning perspective is poorly motivated.

Thank you for highlighting this issue. We see three notable utilities of using Fisher Information (FI) for transit planners and we added a paragraph in the introduction. We also reiterated the benefits towards the end of the discussion.

2. How does this method compare or improve upon other more commonly used metrics? What is the benefit of this method as compared to other trend clustering methods (i.e. latent curve modeling) for analyzing time series data? I understand that it can be used for multiple variables, but the limitation I see is that we can't discern which variable is causing the change? Is it a supply or demand problem?

FI provides a means to monitor system dynamics by assessing changes in each individual variable. It can therefore be used to capture system-wide dynamics, but it cannot be used as a predictive model. As the reviewer mentions, the prime benefit of this method over others is its ability to handle multiple variables. As the computed FI contains information about all variables rather than a specific one, it can certainly be used as an input for any trend clustering methods. We also revised the introduction accordingly but we have not discussed other clustering methods at this point not to confuse readers who are not familiar with advanced clustering methods. We are happy to change this if the reviwer feels strongly about it.

In addition, we completely agree about the incapacity of FI to discern which variable is causing the change, and we added text on this point in the manuscript in section 2.2.

3. Overall, I didn't find that the method revealed anything unknown or extraordinary about US transit system. Rather, it confirmed in a very descriptive manner rather obvious trends about these systems (such as, ridership changes follow the opening of a new light rail line).

Regime shifts capture substantial changes in any system. For public transportation systems (PTS), the introduction or termination on any mode is an example of a major change, and but

these changes are obvious and an advanced method like FI is not required to capture them. Nevertheless, we capture a regime shift in the Sacramento rail system, where no transit mode was introduced or terminated in the 2002-2016 time period studied. To highlight this benefit, we replaced San-Diego with Sacramento in Figures 1, 2, and 4. Overall, we find regime shifts for only 254 PTS, out of 1146 different PTS. We also find decreasing FI patterns for 308 PTS and increasing FI patterns for 136 PTS. These results are interesting and non-trivial. After all, a decreasing FI trend can be alarming and capturing it ahead of time may help develop and make effective policies before any regime shifts. We added text throughout the text to better convey the general benefits of using FI.

Reviewer 2

1. The introduction is very clear, however, it must be improved clarifying the main innovative contributions to the field (a couple of sentences or a small paragraph at the end of section 1 would be enough)

Thank you. We revised the introduction according to your suggestion and added a paragraph showing the relevance of the work to transit planners. We also added information about FI in section 2.2

2. At the end of the literature review, the research gap to fill/the alternative method you proposed together with possible benefits (e.g. accuracy, computation speed, easy implementation) should be clearly stated, in order to position your paper in the current literature. Since you already published something related to FI, it would be also desirable to add a couple of sentences on the advances of your research in the proposed paper starting from reference [13].

Typically metrics related to ridership (e.g. unlinked passenger trips) and revenue (e.g. vehicle revenue miles) are used discretely to measure the performance of public transportation systems (PTS). Here, we introduce Fisher Information (FI) as a technique to combine all those performance metrics together to get a combined performance metrics, and we use UPT and VRM for our analysis. This is a substantial contribution in itself. As per your suggestion, we revised section 1 and we stated this aspect more clearly.

Moreover, our last paper on FI was offered a detailed procedure to compute FI. We simply implemented FI on a simple dataset to demonstrate how FI works. In contrast, here, we implement FI on a full scale to analyze all the PTS in the US, and we are able to directly extract information that can be used in practice.

3. The link https://github.com/csunlab/fisher-information , could it be added as a reference? Since it is absolutely that. But you can check it during the editorial process for publication.

We added it as a reference. We also added a link to the tutorial.

4. My opinion is that section 3 (very short and still part of the "background" needed for readers) can be merged with section 2, with a small introduction to the Sustainable Regimes Hypothesis.

Thank you. Sections 2 and 3 were merged together.

5. In the discussion section, mainly trends descriptions are reported. I would add also some specific explanations. Why rail over the other transport modes? why the jump-drop of NY in 2004-09? What's behind that?

We use UPT and VRM data for our analysis, we therefore do not explain why these changes occurred in the first place (the process would be much too long for the 1146 PTS). Although one of the limitations of FI is that it cannot discern which variable is causing the change, we tried to provide as much information as possible about how FI can be used in practice; i.e., overall trends in FI, eight different patterns, regime shift, etc.

6. Another important issue to consider is the role of the latest technology. Probably the discussion on autonomous vehicles is not yet influencing PTS dynamics, since very few applications around the world are available (e.g. Autonomous Post bus in Sion, Switzerland).

Agreed. Significant use of autonomous vehicle will definitely influence PTS dynamics, which can be captured through FI. While we may not want to highlight this aspect in the introduction (afterall, we are not directly studying autonomous vehicles), we added some text in the conclusion.