China, the outlier of world research system?

The stray ICR (international citation ratio) of China

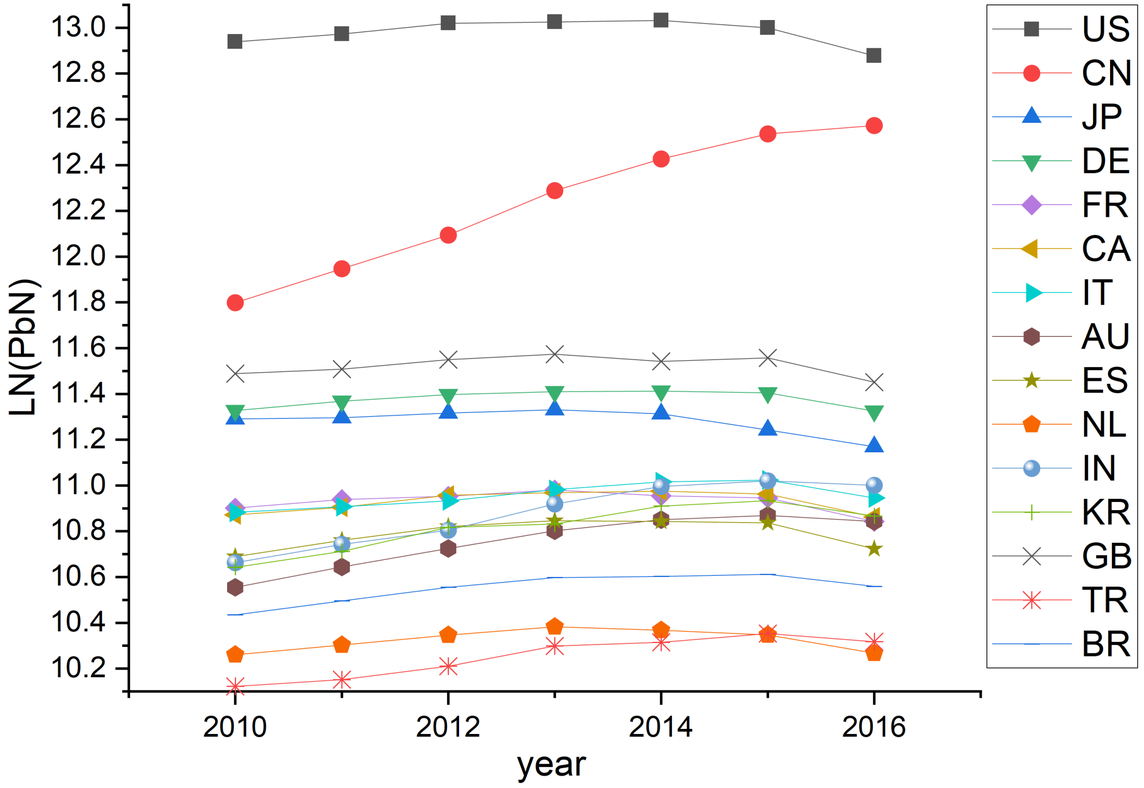
Weihan Wang
3/3/2021

Contents

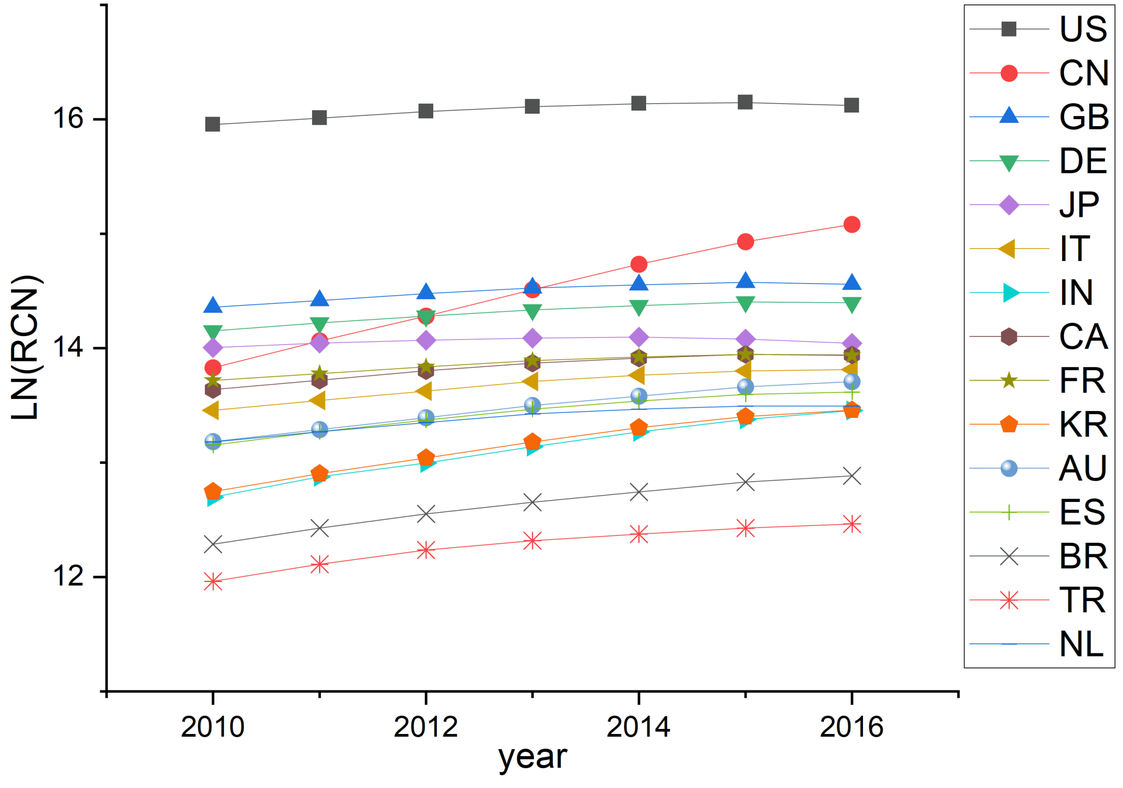
*All data is from WOS (Clarivate Analytics@ Web of Science) database

China, a rising star

- Significant increase in both publication and citation received.
- Second only to the US in publication and citation quantity.



*PbN=Publication Number

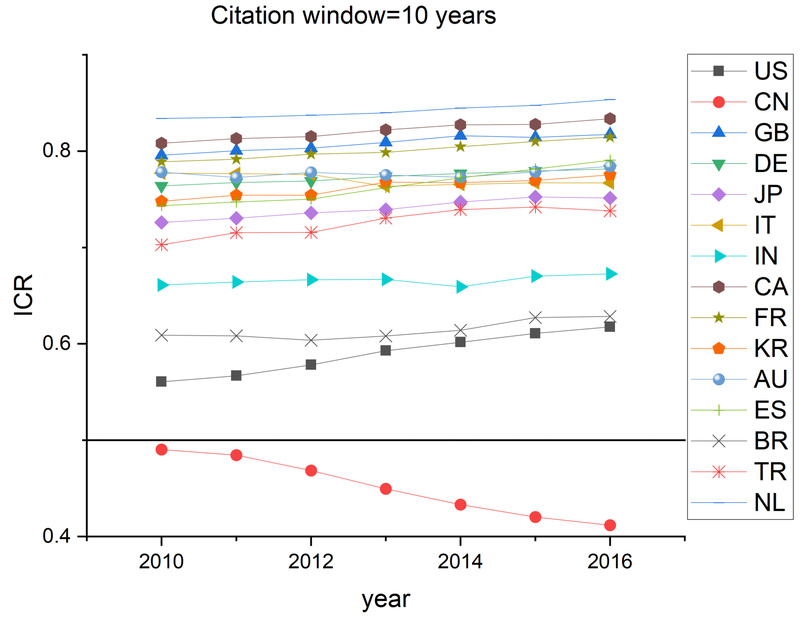


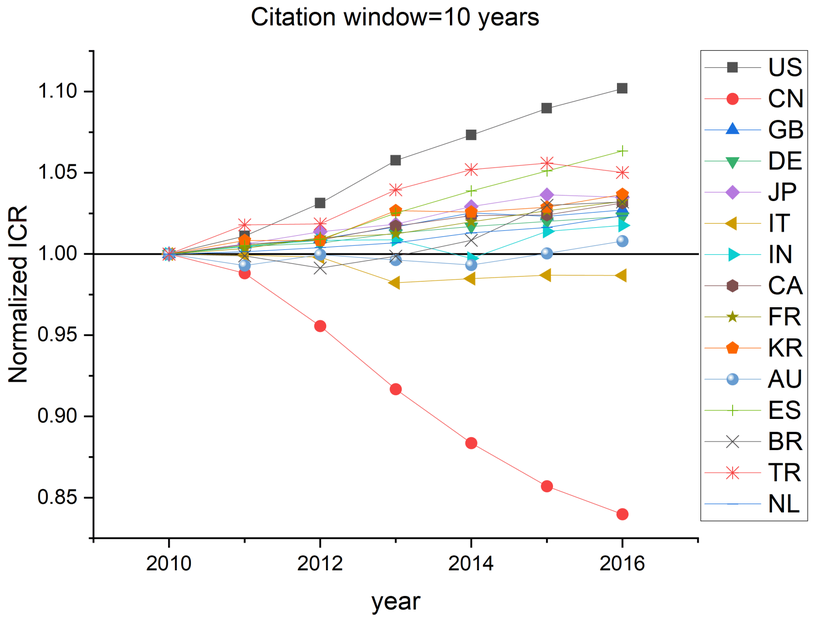
*RCN=Received Citation Number

Citation Window=10 years

International citation ratio

- $$F = \frac{International\ citations\ received}{Total\ citations\ received}$$
- Reflect the international impact to a certain extend.
- The stray performance (significant continuous decline) of China's F compared with other main countries*.

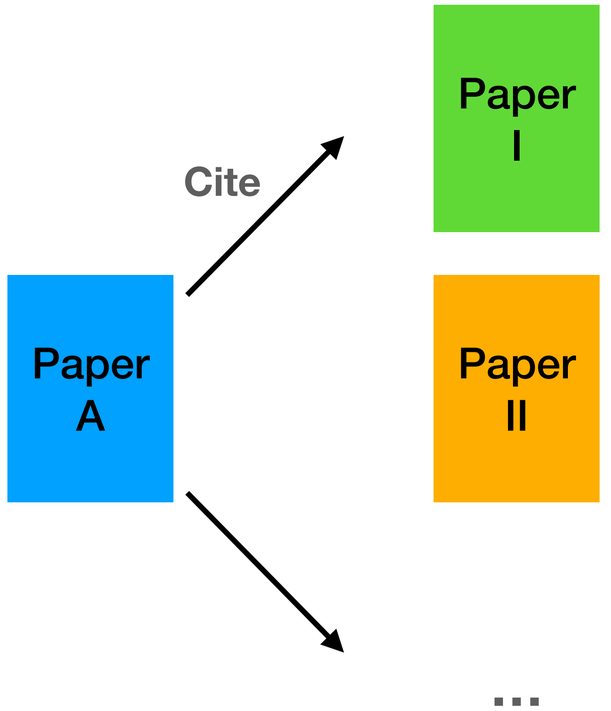




*Top 15 countries in publication number from 2010 to 2016.

Citation process

- Paper A is published with a reference list.
- Paper I, II, etc. appeared in paper A's reference list.
- Paper A has cited paper I and II while paper I and II have received citation from paper A.
- The first address* of a paper determines its nationality.
- Citation window is a time window from which you cite papers.
- Expected citation probability is the corresponding probability distribution in citation window.



*Provided by WOS according to the first institute.

Citation process

- When you decide to cite a paper, you are influenced by many factors.
- **Subjective factors**
- personal view, research relationship, religious belief, ideology…
- **Objective factors**
- type(article, review, letter), year, field, quality, language…
- **Too many factors, some are hard to define and quantify! Just like a thermodynamic system whose micro laws are difficult to observe and summarize.**

Citation process

- Many studies have pointed out that complex micro behaviors tend to show certain consistency at the macro level.

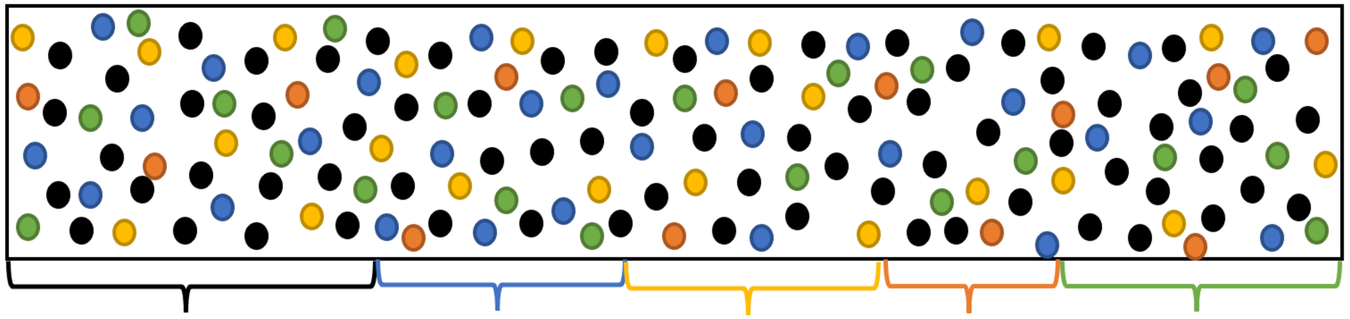
Micro-level (individual) citation behaviour

Aggregation

Macro-level (country) citation behaviour

‘Picking Ball’ model

- From the macro point of view, we try to explain the phenomenon by assuming a random citation minimalist model.
- Assumption: random citation
- The citation process is just like picking ‘balls’ from a ‘box’.
- Published papers—Existing balls in box.
- Expected probability of citation—Probability distribution of balls (same for all)
- Citation number—Number of balls you pick



$$F_i^{yr,cw} = \frac{number\ of\ non-i\ balls\ picked\ by\ i}{number\ of\ balls\ picked\ by\ i}$$

‘Picking Ball’ model

Total publication number in yr :

$$PBN^{yr} = n_0^{yr} + n_1^{yr} + n_2^{yr} + \ldots + n_N^{yr}$$

Expected citation probability of i :

$$P_i^{yr,cw} = \frac{\sum_{y=yr-cw}^{y=yr-1} n_i^y}{\sum_{y=yr-cw}^{y=yr-1} PBN^y}$$

Citation from j to i :

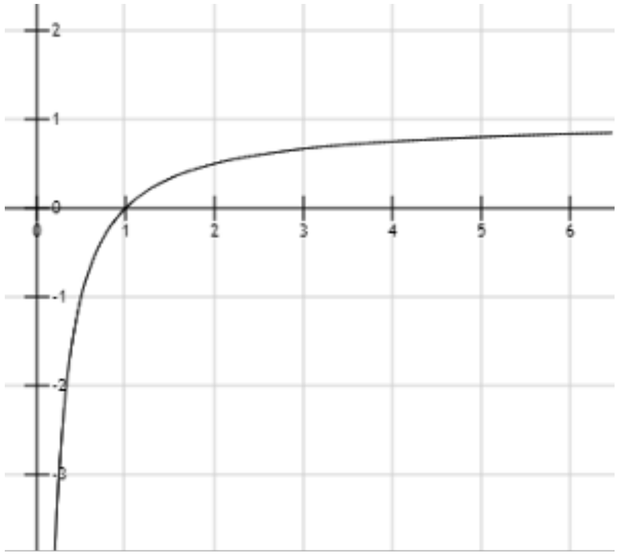
$$C_{ji}^{yr,cw} = C_j^{yr,cw} * P_{ji}^{yr,cw} = (n_j^{yr} * RL_j^{yr,cw}) * (\frac{\sum_{y=yr-cw}^{y=yr-1} n_i^y}{\sum_{y=yr-cw}^{y=yr-1} PBN^y})$$

International citation ratio of i :

$$F_i^{yr,cw} = \frac{\sum_j C_{ji}^{yr,cw} - C_{ii}^{yr,cw}}{\sum_j C_{ji}^{yr,cw}} = 1 - \frac{C_{ii}^{yr,cw}}{\sum_j C_{ji}^{yr,cw}} \longrightarrow f_i^{yr,cw} = \frac{C_{ii}^{yr,cw}}{\sum_j C_{ji}^{yr,cw}} > 0$$

$$F_i^{yr,cw} = 1 - f_i^{yr,cw} = 1 - \frac{1}{(\frac{1}{f_i^{yr,cw}})}$$

The graphic of function $1 - \frac{1}{X}(X > 0)$:



yr : year

cw: citation window

RL:average reference list length

‘Picking Ball’ model

$$\frac{1}{f_i^{yr,cw}} = \frac{\sum_j [(n_j^{yr} * RL_j^{yr,cw}) * (\frac{\sum_{y=yr-cw}^{y=yr-1} n_i^y}{\sum_{y=yr-cw}^{y=yr-1} N^y})]}{(n_i^{yr} * RL_i^{yr,cw}) * (\frac{\sum_{y=yr-cw}^{y=yr-1} n_i^y}{\sum_{y=yr-cw}^{y=yr-1} N^y})} = \frac{\sum_j [(n_j^{yr} * RL_j^{yr,cw})]}{(n_i^{yr} * RL_i^{yr,cw})} \qquad F_i^{yr,cw} \propto \frac{\sum_j [(n_j^{yr} * RL_j^{yr,cw})]}{(n_i^{yr} * RL_i^{yr,cw})} = \frac{\sum_j C_j^{yr,cw}}{C_i^{yr,cw}}$$

The change rate of citation sent out :

$$rC_i^{yr,cw} = \frac{C_i^{yr+1,cw} - C_i^{yr,cw}}{C_i^{yr,cw}} \qquad \Delta F_i^{yr \to yr+1,cw} = (F_i^{yr+1,cw} - F_i^{yr,cw}) \propto \sum_j [\frac{C_j^{yr,cw}}{C_i^{yr,cw}} (\frac{rC_j^{yr,cw} + 1}{rC_i^{yr,cw} + 1} - 1)]$$

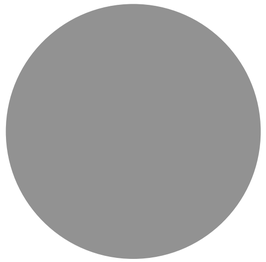
The impact from j to i in yr :

$$Impt_{ji}^{yr,cw} = \frac{C_j^{yr,cw}}{C_i^{yr,cw}} (\frac{rC_j^{yr,cw} + 1}{rC_i^{yr,cw} + 1} - 1) \begin{cases} > 0, & rC_j^{yr,cw} > rC_i^{yr,cw} \\ = 0, & rC_j^{yr,cw} = rC_i^{yr,cw} \\ < 0, & rC_j^{yr,cw} < rC_i^{yr,cw} \end{cases}$$

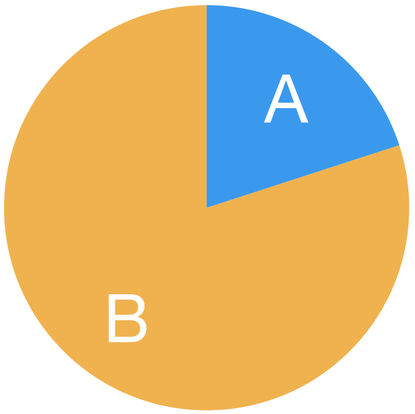
‘Picking Ball’ model

$$F_i^{yr,cw} = \frac{number\ of\ non-i\ balls\ picked\ by\ i}{number\ of\ balls\ picked\ by\ i}$$

$$Impt_{ji}^{yr,cw} = \frac{C_j^{yr,cw}}{C_i^{yr,cw}}\left(\frac{rC_j^{yr,cw}+1}{rC_i^{yr,cw}+1}-1\right)\begin{cases} >0, & rC_j^{yr,cw} > rC_i^{yr,cw} \\ =0, & rC_j^{yr,cw} = rC_i^{yr,cw} \\ <0, & rC_j^{yr,cw} < rC_i^{yr,cw} \end{cases}$$

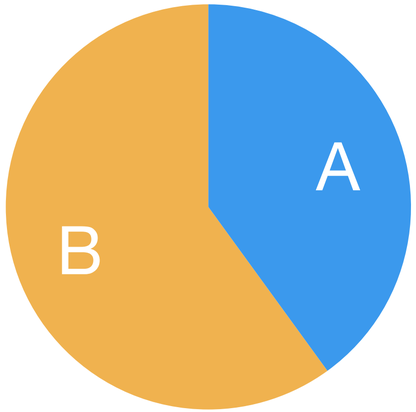


Total citation sent out



Citation distribution
yr

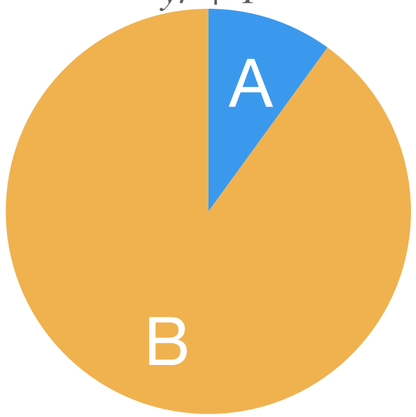
$rC_A^{yr,cw} > rC_B^{yr,cw}$



Citation distribution
$yr+1$

$$\begin{cases} F_A^{yr+1,cw} = \dfrac{C_B^{yr+1,cw} * P_A^{yr+1,cw+1}}{(C_A^{yr+1,cw}+C_B^{yr+1,cw}) * P_A^{yr+1,cw+1}} & decrease \\ F_B^{yr+1,cw} = \dfrac{C_A^{yr+1,cw} * P_B^{yr+1,cw+1}}{(C_A^{yr+1,cw}+C_B^{yr+1,cw}) * P_B^{yr+1,cw+1}} & increase \end{cases}$$

$rC_A^{yr,cw} < rC_B^{yr,cw}$

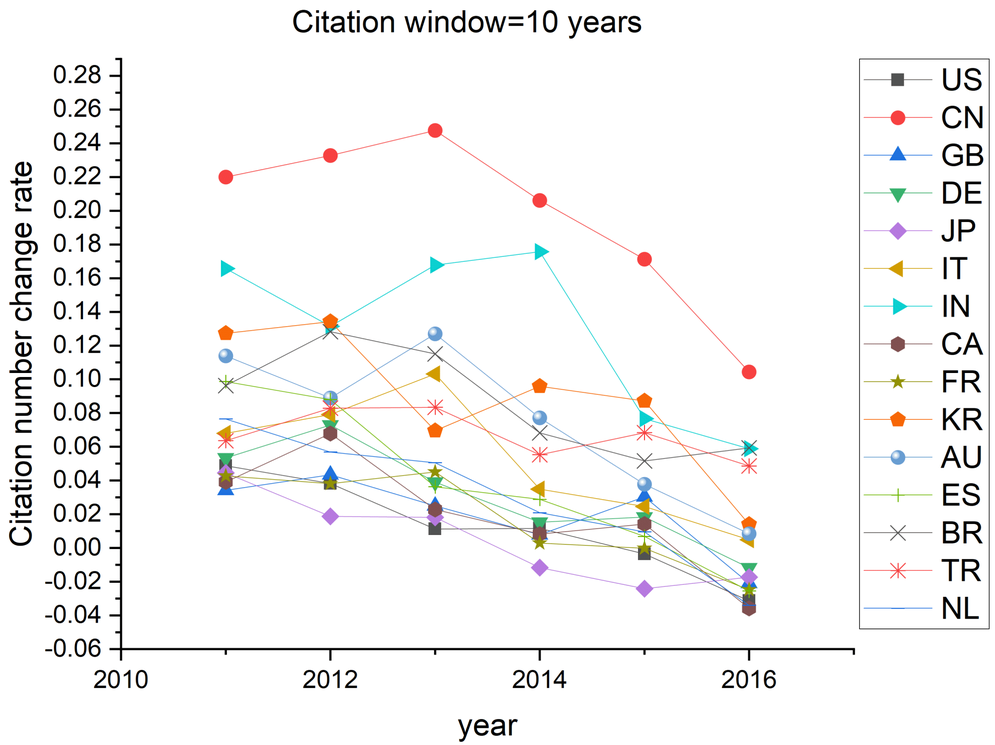


Citation distribution
$yr+1$

$$\begin{cases} F_A^{yr+1,cw} = \dfrac{C_B^{yr+1,cw} * P_A^{yr+1,cw+1}}{(C_A^{yr+1,cw}+C_B^{yr+1,cw}) * P_A^{yr+1,cw+1}} & increase \\ F_B^{yr+1,cw} = \dfrac{C_A^{yr+1,cw} * P_B^{yr+1,cw+1}}{(C_A^{yr+1,cw}+C_B^{yr+1,cw}) * P_B^{yr+1,cw+1}} & decrease \end{cases}$$

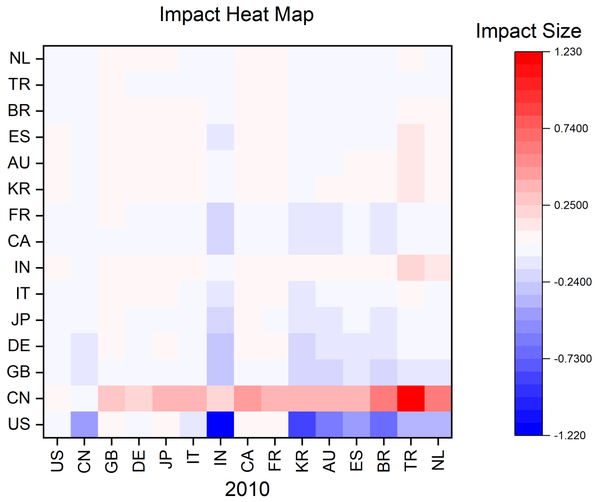
Model validation

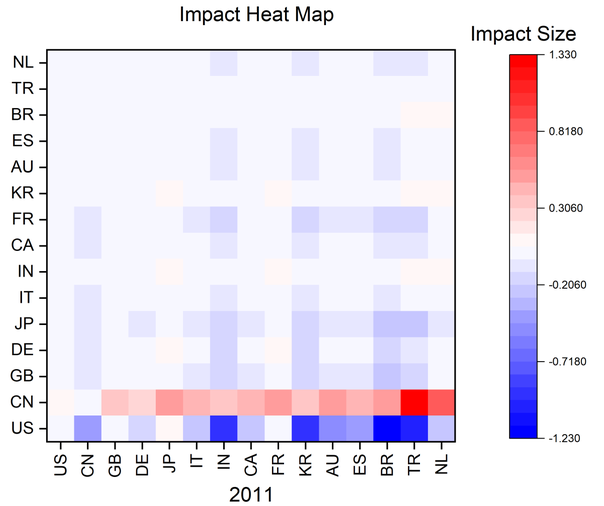
- China has the highest citation growth rate among the main countries in the empirical data.
- China has a positive impact on other countries' F while the others have negative impact on that of China.
- The model predicts the decline of China's F with such a high growth rate of citation sent out.

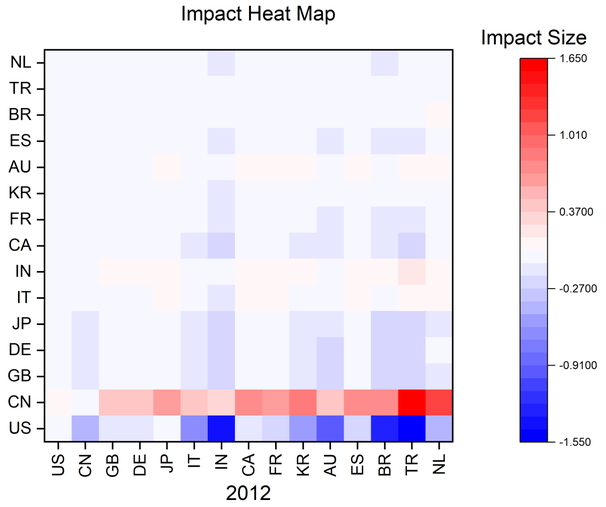


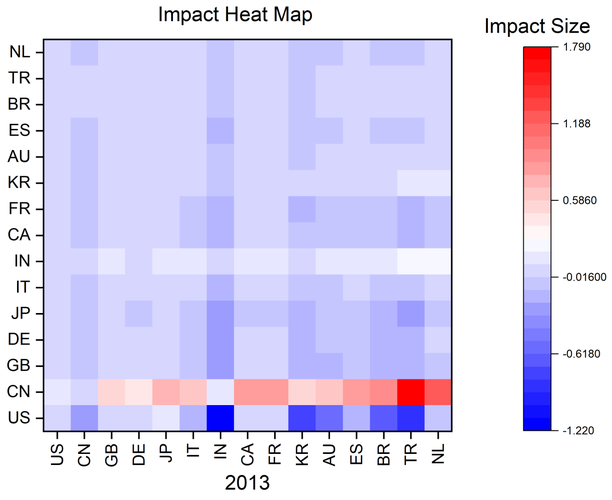
Model validation

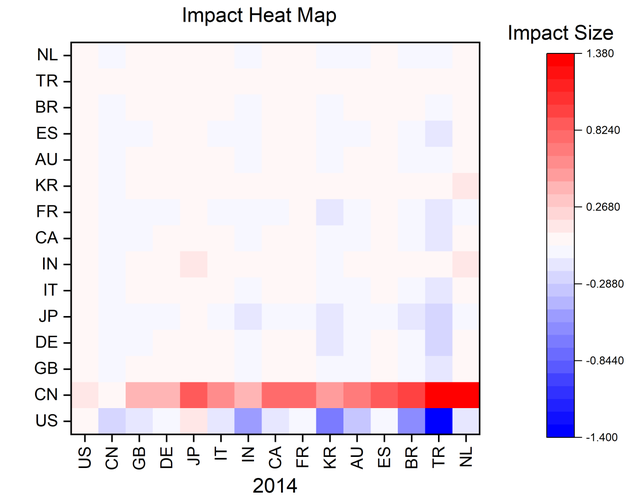
$$Impt_{ji}^{yr,cw} = \frac{C_j^{yr,cw}}{C_i^{yr,cw}}\left(\frac{rC_j^{yr,cw}+1}{rC_i^{yr,cw}+1}-1\right)\begin{cases} >0, & rC_j^{yr,cw} > rC_i^{yr,cw} \\ =0, & rC_j^{yr,cw} = rC_i^{yr,cw} \\ <0, & rC_j^{yr,cw} < rC_i^{yr,cw} \end{cases}$$

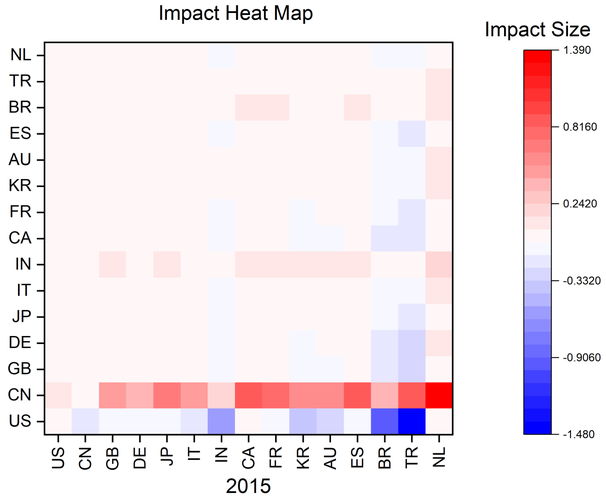












*Citation window=10 years

Model validation

- **Reflect the correctness of the model through five examples:**
- ① The change of China's F in the model.
- ② The change of China's F with half the citation sent out in first year and the same annual change rate.
- ③ The change of China's F with one-tenth the annual change rate as original and the same number of citation sent out in first year.
- ④ The F of India with and without China*.

①~③: base on random citation

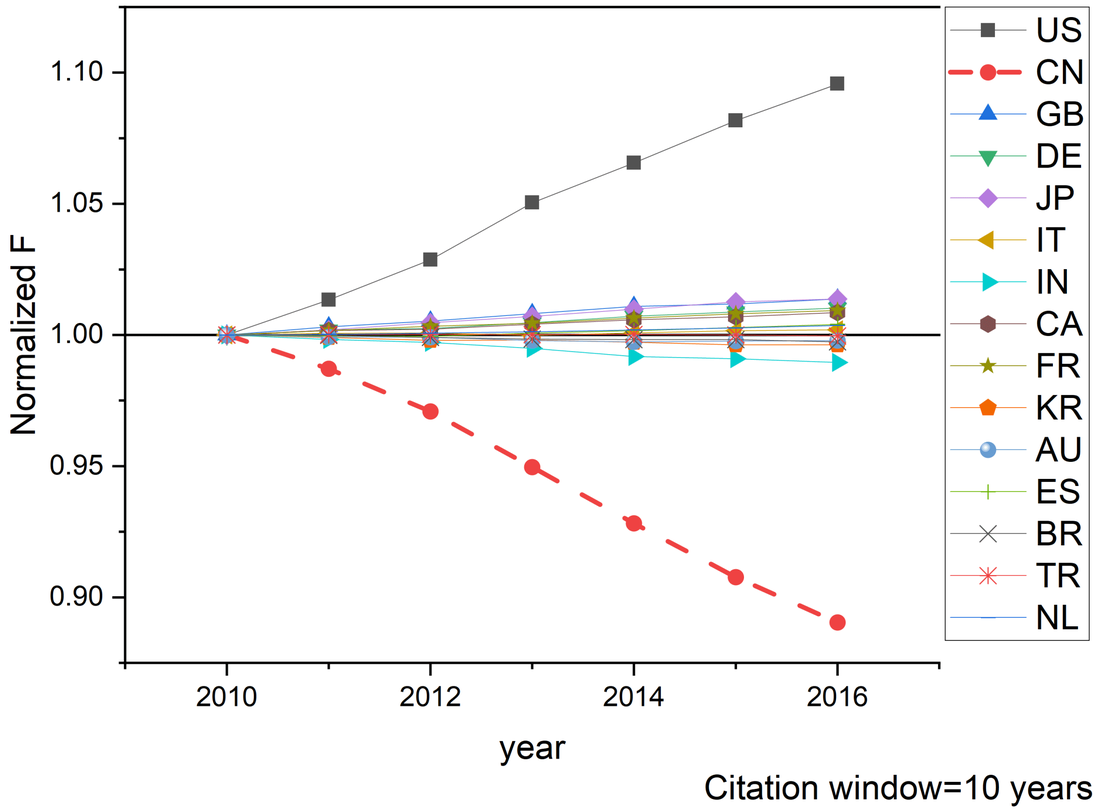
④: based on empirical data

*without China means all the citations sent out and received by China will be ignored

① Change in the model

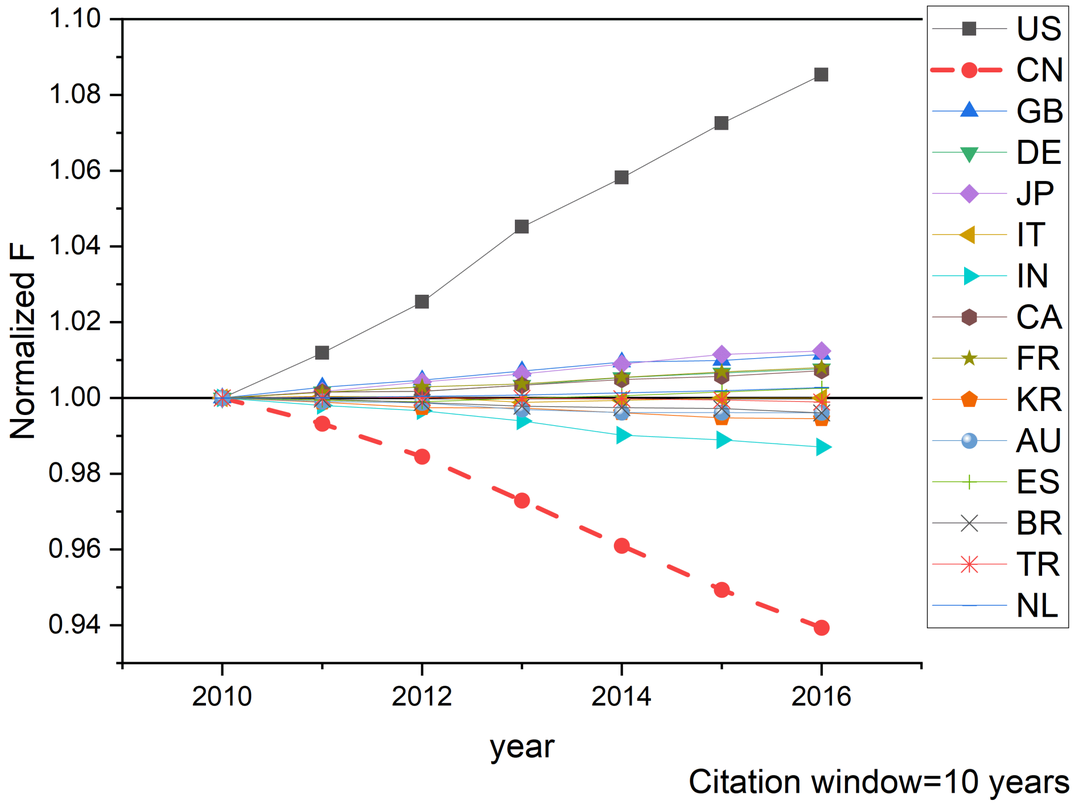
- In this example, the total citation sent out by each country will stay the same while the distribution of it will follow the expected citation probability according to the corresponding citation window (random citation).

- The result is obvious that China's F faces a significant decrease as well in the model with the same citation number and change rate.



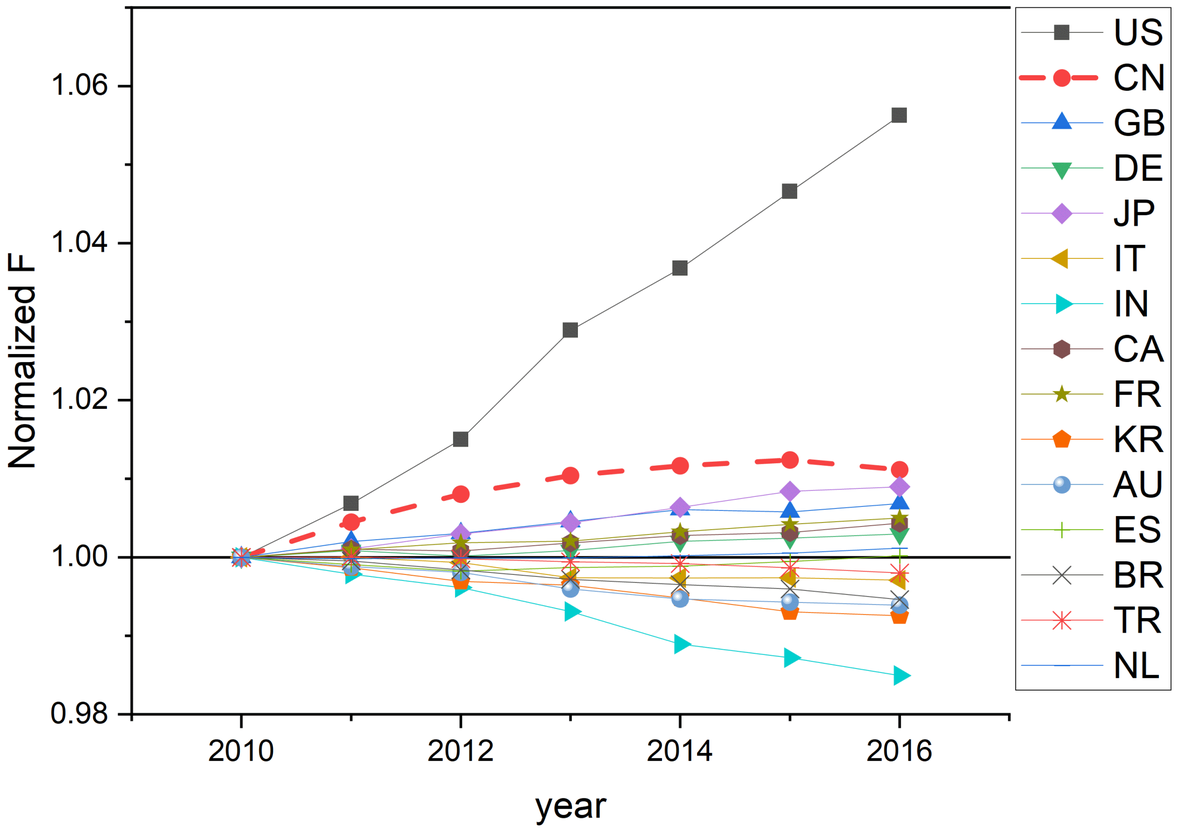
② Half the citation sent out in first year and the same annual change rate

- In this example, the total citation sent out by China will become half as before while the annual change rate of citation sent out remains unchanged.
- The result shows that the decrease of total citation sent out won't change the downward trend of China's F.



③ One-tenth annual change rate and the same citation sent out in first year

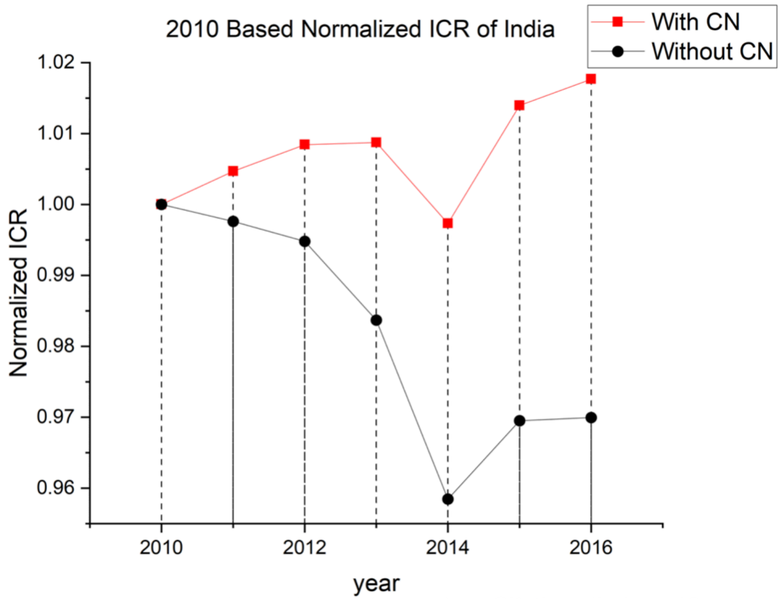
- In this example, the annual change rate of citation sent out by China will be one-tenth as before while the number of citation sent out in first year stays the same.
- The result demonstrates the correctness of the prediction that the downward trend of China's F can be reversed when its annual change rate of citation sent out is low enough compared with others.

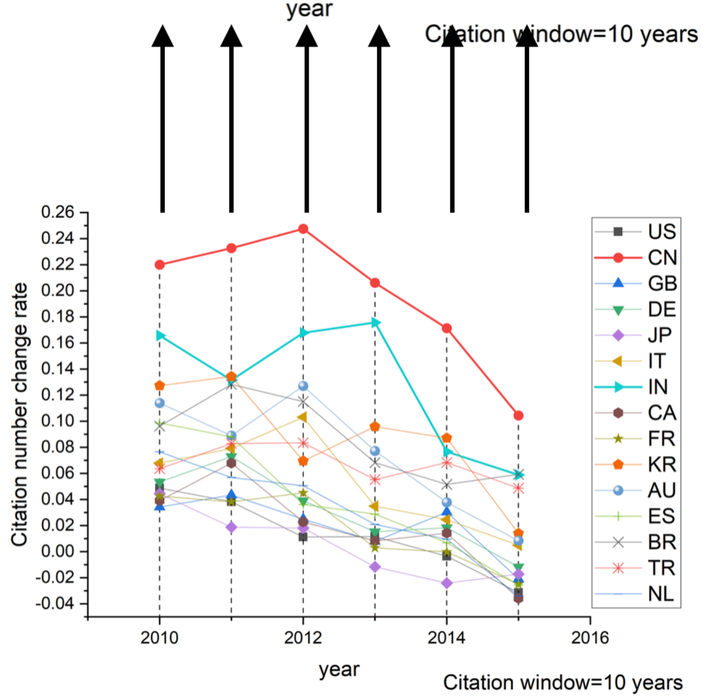


Citation window=10 years

④ The of India's F with and without China

- In this example, the F of India without China in the citation system has a significant decline compared with the original empirical result.
- The result reveals latent forecast that when China is removed from the system, India's F will have an inevitable downward trend as the country with highest annual change rate of citation sent out.





Conclusion

- There are many potential factors that may affect citation behavior while the aggregation of micro-level citation behavior seems to follow certain simple pattern on macro-level like international citation ratio.
- It is different from the previous view that we found international citation ratio fails to reflect the international impact through the minimalist model.
- Country with relatively high growth rate of citation sent out is more likely to have a decline trend of international citation ratio due to its positive impact on that of the others.

Research prospect

- The Prediction of the specific size and quantity change of international citation ratio.
- The potential future decline of China’s growth rate in quantity and the corresponding in international citation ratio provides great chance to discover the features of country-level citation behavior.
- Find an appropriate quantity to describe international academic influence.