

A Systematic Review of Zero-Inflated Count Time Series Models: Techniques, Developments, Challenges and Future Directions

Saurabh Kumar

Central University of Rajasthan

Suraj Prakash

surajpat2002@gmail.com

Central University of Rajasthan <https://orcid.org/0009-0002-3509-736X>

Systematic Review

Keywords: Zero-Inflated, Count Time series, INAR, Bibliometric Analysis, PRISMA

Posted Date: October 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11247897/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

A Systematic Review of Zero-Inflated Count Time Series Models: Techniques, Developments, Challenges and Future Directions

Saurabh Kumar¹ and Suraj Prakash^{1*}

Email: sonysaurabh123@gmail.com; surajpat2002@gmail.com.

¹Department of Statistics, Central University of Rajasthan, Bandarseendri, Kishangarh, Ajmer, Rajasthan, India, Pin- 305817.

Abbreviations: ZINAR, Zero Inflated Integer-valued autoregressive model; ZIP, Zero Inflated Poisson model; ZINB, Zero Inflated Negative Binomial model; ARMA, Auto-regressive moving average; ZINGARCH, Zero Inflated Integer-valued generalized autoregressive conditional heteroscedasticity; SLR, Systematic Literature Review; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses ; MSE, Mean squared error; AIC, Akaike information criterion; MADE, Mean absolute deviation error; BIC, Bayesian information criterion; RMSE, Root mean squared error; PRMSE, Predicted root mean squared error; PMAD, Predicted mean absolute deviation; PTP, Percentage of true prediction; CAIC, Consistent Akaike information criterion; HQIC, Hannan-Quinn information criterion; FMAE, Forecast mean absolute error; FMSE, Forecast mean squared Error; SMAPE, Symmetric mean absolute percentage error; MAAPE, Mean arctangent absolute percentage error; MAB, Mean absolute bias; MAE, Mean absolute error; MAEM, Mean absolute error of the median; WAIC, Watanabe-Akaike Information Criterion; SMC, Sequential Monte Carlo; MCMC, Markov Chain Monte Carlo; MLE, Maximum Likelihood Estimator; QMLE, Quasi-Maximum Likelihood Estimator.

Abstract

Count time series is crucial for analyzing a discrete time series. It shows temporal dependence, variation, and structural irregularities in the data. These irregularities may arise from having too many zeros. This model has become prominent because it can handle two data-generating mechanisms. These involve both a zero-state process and a traditional count process. This is the first comprehensive SLR on Zero-Inflated INAR models, emphasizing methodological developments, estimation strategies, and applied research involving Zero-Inflated Integer-valued autoregressive (INAR) count time series, as well as a Bibliometric analysis from 1992 to 2025. Using the PRISMA framework, this study draws on research articles from the Web of Science and arXiv databases. This review article addresses the following keywords by reviewing 40 articles: model evolution, estimation techniques, diagnostic practices, domain applications, and future challenges. The expected findings of this article indicate that although methodological diversity has grown considerably, encompassing Zero-Inflated models and state-space models, issues such as overdispersion, computational burden, and nonstationarity remain open challenges for researchers. We also found that some of the authors' recommendations for future research may emphasize Bayesian hierarchical extensions and Structural Breaks in models.

Key Words: (Zero-Inflated; Count Time series; INAR; Bibliometric Analysis; PRISMA).

***Corresponding Author: Suraj Prakash**, Department of Statistics, Central University of Rajasthan, Bandarseendri, Kishangarh, Ajmer, Rajasthan, India, Pin- 305817.

1. Introduction

Count time series data are mainly observed in the fields of epidemiology, finance, environmental science, transportation, and public healths, where observations represent discrete counts recorded over time (Fokianos, 2011; Davis et al., 2021). Traditional count time series, including Poisson Autoregressive and integer-valued Autoregressive models (INAR) assume equi-dispersion and lack the ability to handle excessive zeros, such type of a common phenomenon found in real-world dataset (Zeger, 1988; Weiß, 2018). When number of zero occurrences substantially exceed in observations then model expectations, conventional approaches lead to produced biased inference, estimated variance is underestimate, and shows poor forecasting performance (Lambert, 1992; Ridout et al., 1998).

To address above said issues, zero-inflated models (ZIMs) were introduced, originally for cross-sectional count data, and later extended to another model with time-dependent structures (Lambert, 1992; Li et al., 1999). Zero-inflated count time series models incorporate a dual generative process: one for structural zeros and another for the count-generating mechanism which is Poisson or Negative Binomial distributions (Lambert, 1992; Zhu, 2012). Over time, these models evolved into multiple branches including zero-inflated Poisson auto regression (ZIP-AR), ZINB time series, ZI-INAR, Markov-switching Zero Inflation, and state-space based models with excess zero components (Li et al., 1999; Yang et al., 2013; Liu et al., 2020). In the present scenario most researchers had proposed new techniques in count time series such as introducing zero-inflated, state space, thinning operator and some distributions with different statistical inference techniques motivate me to explore more this area and developed as my research interest. Our main objective of this study is to introducing the most useful techniques and models to address the real world problems and also for forecasting. In that way, we have conducted a structured and documented study of published papers following the rules as defined by the method of Systematic Literature Review (SLR) (Kitchenham et al., 2009). The SLR method generally used in various areas, Medicine, Management and also in theoretical framework which is statistical model also in time series. Its goal is to examine a rigorous and well-structured search for published papers in order to collect and evaluate evidences on this research area. The main advantages of this is to perform a wide search, recovering not only well-known papers but also other related documents like published thesis, book, and edited book chapters ones. Papers collected in SLR can be either directly read and analyzed or organized as a Systematic Mapping, as presented in Petersen K & Mattson, (2008), Durelli & Delamaro, (2010). In this Review, the paper envisages strong evidences of the importance of Count Time series and also identifies the most common popular methods considered for this purpose, the frequency of published papers, the relation between the number of papers and the type of publication, and the most important gaps of each method. This paper is presented as follows: Section 2 outlines concepts; Section 3 describes the search parameters; Section 4 analyzes the selected documents; Section 5 reviews common Count Time series techniques; Section 6 provides conclusions and future scope of study. A complete summary of all 40 papers are given in table- A.5, with additional results presented in the Appendix. Flow Chart for this study is given as Fig-1.

2. Methodology

In conducting this Review on Zero Inflated count time-series methods for Developments and Challenges; we followed a rigorous methodology which is delineated into specific components outlined in the subsequent sections. Kitchenham et al., (2009) describe evidence as the collection of strong, scientifically sound studies available on a specific topic. They highlight the SLR as the vital method for gathering such evidence because it provides clear steps for reviewing research in a structured and unbiased way. Still, an SLR cannot fully replace a traditional literature review. Before conducting a systematic literature review (SLR), researchers should have a basic understanding of the topic. This knowledge helps those select appropriate keywords and determines which studies to include or exclude. A traditional review can reveal important ideas or papers that the keyword-based search in an SLR might overlook. Since SLRs rely heavily on predefined terms, some important studies may not show up in the results. To address this issue, researchers often take a recursive approach by checking the references of selected papers to find more relevant work. In this study, all analysis was done using Microsoft Excel and VoS-Viewer software.

To build a clear and thorough understanding of the latest developments, methods, and challenges in count time series, especially in areas like zero-inflation and INAR models, we followed the SLR guidelines proposed by Kitchenham et al., (2009) and Endo & Silva (2010). The process was organized into three phases. The first phase, Search, focused on defining how to find relevant studies. In the second phase, Analysis, we carefully examined and rated the quality of the collected papers. The final phase brings together these findings and outlines the key conclusions from the review. Each phase is explained in detail in the next sections, along with a flowchart that visually summarizes the process.

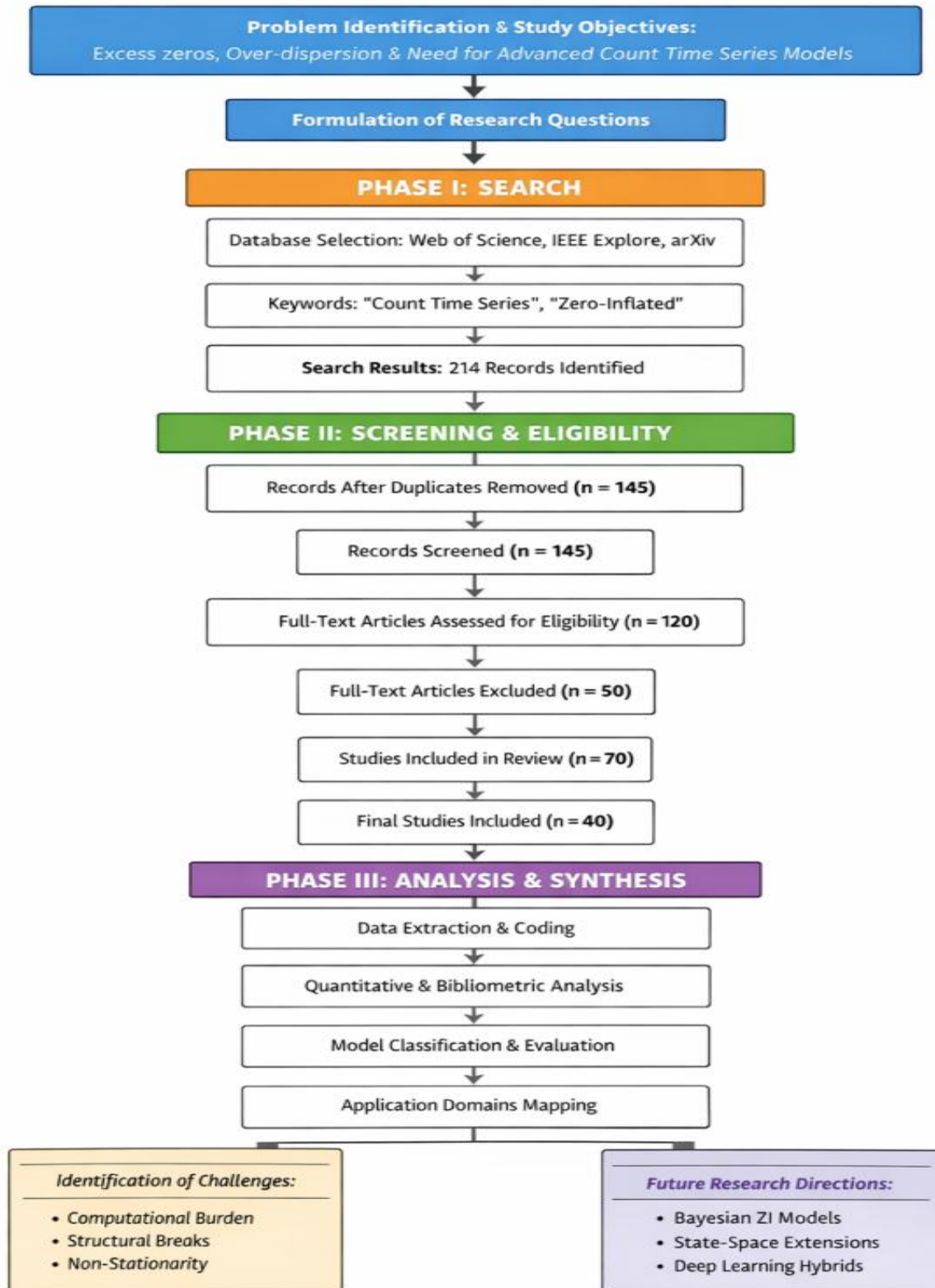


Fig.1: Flow chart of this study

3. Phase I- Search:

In first phase, we have defined the overall scope of this study by deciding on the criteria that determine if a paper is relevant to our research. This involved clarifying the research objective, creating the main and secondary research questions, selecting search repositories, choosing the standard language for the review, identifying keywords, building the search query, and setting the inclusion and exclusion criteria. We also outlined the common procedure for the review. Our main goal is to explore existing techniques, recent developments, key challenges, and future research opportunities in count time series, especially focusing on zero-inflation. With this purpose in mind, we formulated the Main Research Question (MRQ) for the study is as:

What are the techniques, Advancements and Challenges in Count time series particularly Zero-Inflated?

Establishing a series of Secondary Questions (SQs) that validate the study's direction and depth is crucial to supporting the primary research topic. These inquiries aid in determining useful applications, evaluating current methods, and spotting patterns in published work. The following are the secondary questions:

SQ.1. How can count time series be used in practical applications?

SQ.2. How was methods assessed?

SQ.3. Why is count time series used?

SQ.4. What is the primary methods used to assess the performance of count time series models?

SQ.5. What major methodological advancements have been introduced in count time series modeling over time?

SQ.6. In which real-world fields or domains are count time series models most commonly used?

SQ.7. How many papers are published each year?

SQ.8. Who is the leading researchers in this field?

SQ.9. What is limitations for model implementation, evaluation, or computational feasibility?

SQ.10. What is the emerging challenges and future research directions?

After establishing the research questions, our next task in this initial phase was to identify suitable search repositories from which to collect relevant studies. We have selected repositories that offer reliable web-based search engines, support keyword-driven queries, and are widely recognized and used within the scientific community. Considering these criteria, the following repositories were selected:

Web of Science (www.webofscience.com), IEEE Explore Digital Library (<http://ieeexplore.ieee.org/>) and Arxiv (<https://arxiv.org/>).

English was selected as the standard language for this systematic review, meaning that studies published in any other language were excluded. Following this, we identified a set of keywords based on the research hypothesis and the main research question. These keywords are:

- (i) Data organization: Count Time series;
- (ii) Goals: Zero-Inflated, filter;
- (iii) Results: Zero-Inflated Count time series.

We also chose to include the term “filter” as a synonym for time series, since earlier studies often use this term to describe techniques involving zero-inflation and thinning operators in count time series. Although many additional synonyms could be added as keywords, doing so would dramatically increase the number of search results and make the review unmanageable.

Additionally, this strategy complies with SLR regulations. According to Kitchenham et al., (2009), unique research that presents a novel approach or technique within a certain discipline is referred to as a main study. In contrast, reviews of existing methods such as SLRs or methodological surveys are classified as secondary studies. Carefully selecting and limiting the number of keywords helps keep the systematic search manageable and ensures that the review can be carried out effectively. The visualization of the keywords in the papers given by the authors is given below in fig-2.

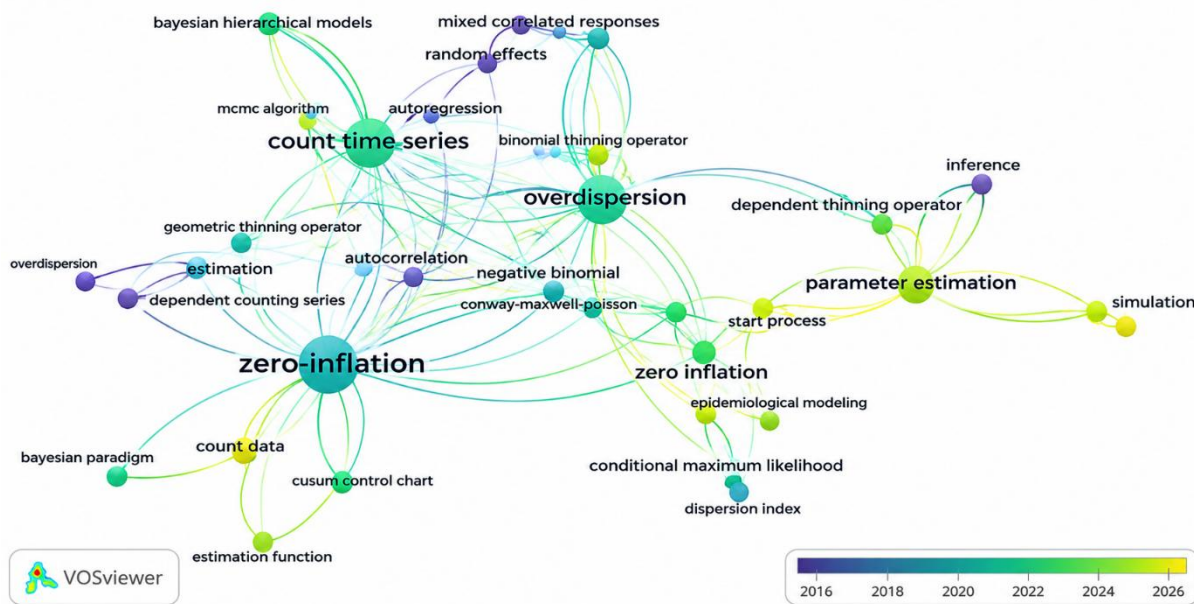


Fig-2: visualization of keywords of the Papers given by authors.

Using the selected keywords, we have formulated the following search query: ("Count time series") AND (Zero-Inflated) AND (Zero-Inflated Count Time Series). We applied this search query across the selected repositories to gather relevant studies. Even with a limited set of keywords, the search returned a plenty of papers, many of which were unrelated to the focus of the SLR. To narrow down the results, we established a set of inclusion and exclusion criteria. A study was included in the SLR only if it met the following conditions:

- (a) Does the study clearly distinguish zero-inflated count time series from standard time series?
- (b) Is it a primary study?
- (c) Does this study specifically address the modeling of zero-inflated count time series?

On the other side, a paper was excluded from the review if it met any of the following criteria:

- (a) The proposed technique was too narrowly tailored to a single, highly specific problem;
- (b) The study did not provide a clearly defined analytical model;
- (c) The evaluation of the technique was inadequate or insufficient;
- (d) The paper lacked a thorough and meaningful review of related literature.

After establishing all criteria, the first phase of the SLR was conducted as following these steps:

- (i) The formulated search query was applied across the selected repositories;
- (ii) Each paper's title and abstract were screened using the inclusion and exclusion criteria;
- (iii) Duplicate or ineffective papers were removed;
- (iv) All remaining papers were fully reviewed in detail.

It is important to note that a paper included during the initial screening may still be excluded in step-4 if, after a full reading, it meets any of the exclusion criteria. Similarly, if the references cited in selected papers are found to be relevant, they are added to the SLR, allowing the review to capture studies that were not retrieved through the repository search. We present the next section which is analysis of the papers selected at the initial phase.

4. Phase II : Analysis

The search query has defined in the first phase was applied to the selected repositories, as a result total of 214 papers. Which were distributed in repositories as presented in Table 1. However, after removing the invalid papers the remaining papers is 145 then after, the PRISMA is applying for Selection of papers across the repositories is shown in Table 2. However, several retrieved references were found to be invalid, as they included conference proceedings, editorial

information, or duplicate papers. After filtering out the invalid references, the final set of valid papers was reduced to 120. Each of these papers was then examined in detail by reviewing those titles and abstracts by applying the established inclusion criteria. Following the screening process, 70 papers were selected for a full reading and detailed evaluation. Subsequently, after applying the exclusion criteria, some of these papers were discarded.

In conclusion, 40 papers were finally identified and classified as strongly relevant to the topic of this study. The Inclusion and Exclusion Criteria is given in Table3 and Table4.

PRISMA Methodology – Flow Chart for Study Screening

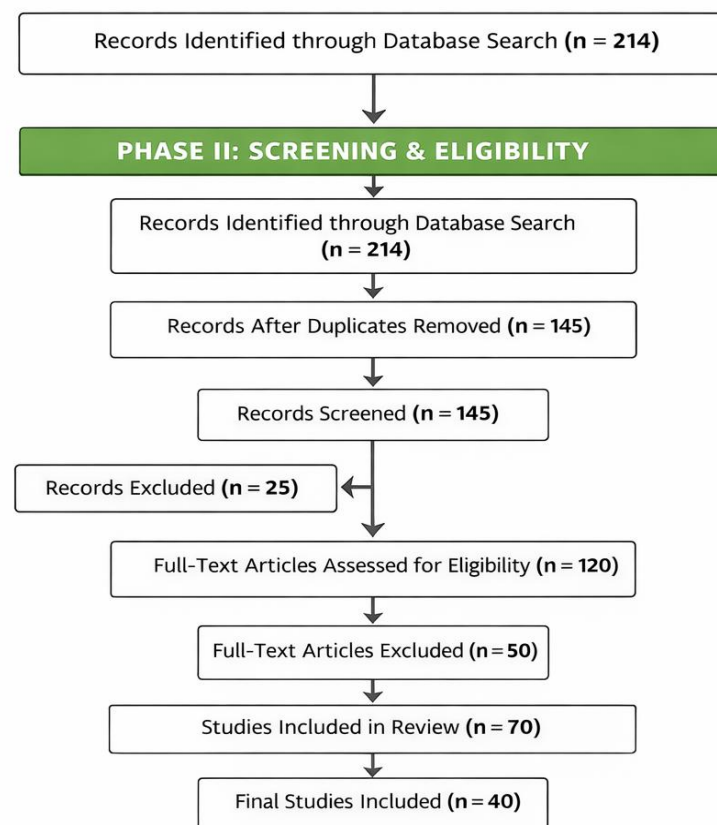


Fig.3: PRISMA Diagram for selection of papers.

Table1: Number of Papers across various repositories and also recovered papers.

Repository	Number of papers
IEEE	10
Web of science	203
Arxiv	1
Total	214
Recovered papers	145
After inclusion criteria	70
After exclusion criteria	40

Table2: PRISMA table for Selection of relevant paper.

PRISMA Stage	Criteria Applied	Output
Identification	Search performed using keywords ("zero-inflated", ZINB, ZIP, INAR, etc.) in WoS	145 papers retrieved
Screening	Review of titles and abstracts; exclude cross-sectional or non-modeling papers	120 papers remain
Eligibility	Full-text screening; keep only papers proposing/applied zero-inflated time-series models	70 papers
Included	Exclude those with no ZI modeling or not time-dependent	40 final papers

Table3: Inclusion Criteria for choosing the related papers.

Code	Inclusion Criteria
IC1	Must model count time series (not cross-sectional data).
IC2	Must apply a zero-inflated component (ZIP / ZINB / ZI-INAR / Bayesian ZI).
IC3	Must contain original methodology or empirical implementation.
IC4	Written in English and full text available.

Table4: Exclusion Criteria for excluding the papers.

Code	Exclusion Criteria
EC1	Pure cross-sectional zero-inflated studies (no time component).
EC2	Reviews, editorials, book chapters, conceptual articles.
EC3	Paper only mentions zero-inflation but does not estimate or model it.
EC4	Duplicates or inaccessible full text.

Note: IC- Inclusion Criteria, EC- Exclusion Criteria

First, we examined the 40 selected papers to determine the number of authors by country and the number of publications attributed to each author. These results have summarized in Table 5 and Table 6. All subsequent analyses were conducted using these 40 papers. The visualization of author's country is given in given in fig-4.

Table-5: Country wise No. of Authors.

Country	No. of Authors	Country	No. of Authors
USA	12	Brazil	5
China	10	Malaysia	3
India	8	Saudi Arab	3
Germany	6	Canada	3
Iran	6	France	5
South Korea	5		

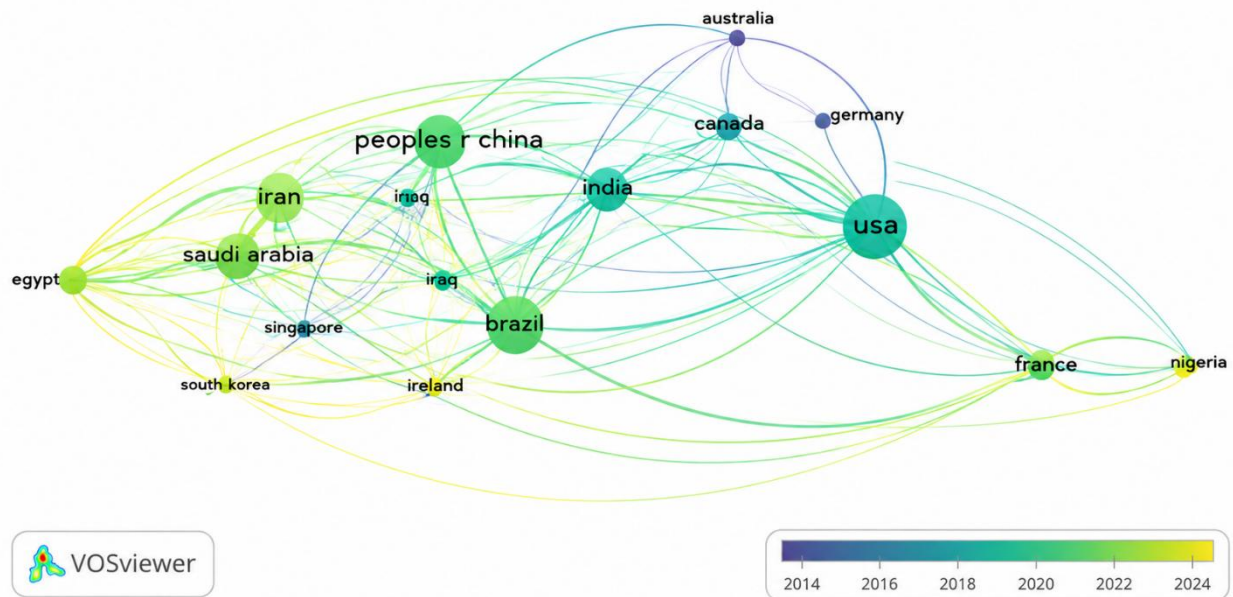


Fig-4: Visualization of Author's Country

Now, proceed with table-6 provides another meaningful insight by highlighting the researchers who have contributed the most papers on zero-inflated count time series, particularly in terms of challenges, techniques, and developments. This analysis helps address Secondary Question SQ.8, which focuses on identifying the leading researchers in this field.

Table-6: Number of Publications of Authors.

Researcher	Publications	Researcher	Publications
Christian Weiss	4	Li, Qi	2
Zhu, Fukang	4	Mohammadpour, M.	2
Lee, Sangyeol	3	Möller, Tobias	2
Chen, Cathy W. S.	3	Samani, EhsanBahrami	2
Wang, Dehui	3	Efstathiosaparoditis	2
KonstantinosFokianos	3	Shirozhan, Masoumeh	1
Ristic, Miroslav M.	3	Wagner Barreto-Souza	1
Kim, Dongwon	2	Lee, Sangyeol	1
Bakouch, Hassan S.	2		

Table 7 shows the number of papers published each year, providing an answer to Secondary Question SQ.7. According to this table, 2021 recorded the highest number of publications.

Table-7: Numbers of paper published per year.

Year	Publications	Year	Publications
1992	1	2014	1
1998	2	2017	3
2003	3	2019	3
2005	1	2020	4
2007	1	2021	6
2008	3	2023	3
2010	1	2024	4
2012	2	2025	2

Table-8: Number of types of selected papers Journal.

Journal Name	Frequency
Journal of Statistical Computation and Simulation	8
Journal of Time Series Analysis	6
Journal of Applied Statistics	5
Journal of the Korean Statistical Society	4
Entropy	3
Communications in Statistics - Simulation and Computation	4
Statistics	2
Computational Statistics & Data Analysis	3
Journal of Statistical Planning and Inference	2
Statistical Modeling	3

Table 8 presents the frequency distribution of the selected papers based on their publication type. From this, it is evident that the majority of studies were published in reputable journals. Following this, Table 9 provides an analysis of how many of these papers qualify as primary studies.

Table-9: Types of papers and its frequency.

Types of Papers	Frequency
Methodological	19
Theoretical	6
Applied	11
Computational / Software	4

Table 10 looks at the relationship between the chosen papers and their particular goals for using different strategies in order to answer Secondary Question SQ.3. Although these evaluations provide a general summary of the papers found by the SLR, more investigation of each study is necessary. As a result, we carefully examined all of the chosen articles' content. A series of inquiries intended to address Secondary inquiries SQ.1, SQ.2, SQ.4, SQ.5, SQ.6, and SQ.10 served as the basis for this thorough investigation. For this reason, the following questions have been defined:

- (i) What types of real-world applications may take advantage for this count time series?
- (ii) How techniques are evaluated?
- (iii) What are the key techniques commonly used to assess models in count time series?
- (iv) What major methodological advancements have been introduced in count time series modeling over time?
- (v) In which real-world domains are count time series models most frequently applied?
- (vi) What are the emerging challenges and future research directions?

Table-10: Frequency of Published Papers and their Goals.

Goal	Frequency
Developing new models	5
Improving estimation methods	4
Modeling excess zeros & over-dispersion	8
Model comparison studies	3
Group/spatial modeling	2
Forecasting & predictive modeling	5
Temporal dependence modeling	3
Applied case-study papers	2
Copula-based dependence	8

Table-11: Table for what techniques and software were used as per papers.

Model Type	No. of studies using model	Popular estimation approach	Popular software
Zero-Inflated Poisson (ZIP)	15	MLE	R
Zero-Inflated Negative Binomial (ZINB)	10	EM algorithm	Python
Zero-Inflated INAR	8	Bayesian	R (tscount, INAR)
Hybrid / Deep Learning	7	Simulation-based	Python / TensorFlow

In table-10 & 11 mentioned the answer of both questions which is what is the main Goal? And How is the new techniques evaluated?

The overall findings from this analysis which are summarized in the final table presented at the end of this paper. By these results we answer the SQ.1. Which is as Count time series models are especially useful in fields where events are rare and data contain many zeros such as, epidemiology, they support forecasting and monitoring of disease incidence, as shown in time-varying ZIP-AR models for outbreak dynamics (Abiodun et al., 2019) and zero-inflated INAR models for policy-driven regime shifts (Cui et al., 2025). In ecology, they help estimate abundance of rare or spatially dispersed species, including ZIP/ZINB state-space approaches for rare species (Sidumo et al., 2024), space-time models for harbor seals Hoef & Jansen, (2007), and ZI regression for reconstructing historical tree abundance (Nolan et al., 2021). ZI forecasting frameworks for sparse event prediction and zero-modified Markov processes for systems with sporadic activity are two examples of practical applications in operational forecasting and risk analysis (Balakrishna, 2024). All things considered, these models are useful in any field that needs to analyze in frequent, over-dispersed, and zero-heavy event counts robustly.

Techniques are assessed using fit diagnostics and model comparison metrics including AIC/BIC, deviation, Chi-square tests, and residual checks in order to provide the answers to SQ.2 and SQ.4. Accuracy metrics such as RMSE, MAPE, MAD, cross-validation, and, for Bayesian models, posterior predictive checks and DIC/WAIC are used to evaluate predictive performance (Dagne, 2010; Sidumo et al., 2024; Nolan et al., 2021). The introduction of zero-inflated and hurdle models to deal with excess zeros (Lambert, 1992; Ridout et al., 1998) and the creation of autoregressive count models like Poisson AR and INAR processes for temporal dependence Zeger, 1988; McCabe & Martin, (2005) are two significant methodological developments in count time series modeling. More recent progress includes state-space and GARCH-type count models for time-varying volatility (Zhu, 2012; Ratnayake et al., 2023), copula-based dependence structures (Alqawba, 2020), Bayesian hierarchical and spatio-temporal ZINB models (He et al., 2024), and deep learning based zero-inflated forecasting approaches (Kim et al., 2023), along with models allowing structural breaks and time-varying parameters (Cui et al., 2025; Mao et al., 2024). In Continuing to SQ.6, the answer of this is that Count time series models are most frequently applied in epidemiology for modeling disease incidence (Abiodun et al., 2019; Cui et al., 2025), ecology for species abundance and wildlife monitoring Sidumo et al., 2024; Hoef & Jansen, (2007), and environmental science for pollution and

climate-related event counts (Nolan et al., 2021). They are also widely used in transportation safety, public policy, and operational forecasting for rare events (Quddus, 2008).

Finally, we answer the SQ.10 to determine the challenges and future directions in which includes modeling zero-inflation with complex dependence, handling time-varying parameters, and integrating structural breaks, nonlinear dynamics, and spatial effects (Mao et al., 2024; Cui et al., 2025; He et al., 2024). Future research is moving toward Bayesian framework computation, copula and GARCH-based dependence, and deep learning hybrid models for improved forecasting in sparse, irregular, and high-dimensional settings (Alqawba 2020; Ratnayake et al., 2023; Kim et al., 2023).

5. Phase-III: Conclusion

In this section we look into the final stage of this study, in which we summarize the most widely applicable techniques for it. Although numerous approaches were identified during the review (see **Table A.5**), we focus here on a selected group of methods that demonstrated the strongest performance. The discussion provided below is essential for addressing the primary research questions of this study, which aims to examine the developments, techniques and challenges in Count Time series especially Zero-Inflated.

5.1. Zero-Inflated Models

It has become important method for analyzing count time series characterized by a combination of over-dispersion and structural zeros. Early techniques, such as the ZIP and ZINB frameworks, introduced a dual-process formulation to separate the zero-generation mechanism from the underlying count process (Lambert, 1992; Hall, 2000). Building on these foundations, more advanced time-series extensions including ZI-AR, ZI-INGARCH and state-space ZI models deals with temporal dependence and latent dynamics to provide better representation of real-world processes (Ferland et al., 2006; Abiodun et al., 2019). By allowing for complete uncertainty quantification and taking into account nonlinear or time-varying effects, Bayesian variations utilizing MCMC, particle filtering, and data cloning further improve model adaptability Douc et al., 2013; Al-Wahsh & Hussein, (2019). Zero-inflated models still have a number of methodological issues in spite of these developments. It is still challenging to accurately discriminate between sampling zeros and real structural zeros, especially in very sparse series or small samples. Due to complicated likelihoods, identification issues, and multimodality in zero-inflated posterior distributions, estimation can also be computationally taxing Ridout et al., 2001; Min & Agresti, (2005). Furthermore, recent advancements in break-detection ZI models, time-varying parameter ZI models, and Markov-switching ZI structures are motivated by the fact that classical ZI frameworks frequently fail to capture regime shifts, long-memory behavior, or covariate-driven changes in zero proportions (Cui et al., 2025; Balakrishna, 2024).

A distinct trend toward more adaptable, hybrid, and high-capacity modeling approaches may be seen in recent advancements. While ecological and epidemiological studies are

increasingly using spatial temporal ZI frameworks to represent complicated dependencies, state-space ZI models now incorporate sequential Monte Carlo or Kalman filtering to capture latent processes (Hoef & Jansen, (2007); Nolan et al., 2021). Although they frequently come at the expense of interpretability and theoretical assurances, machine-learning extensions such as random forests, deep recurrent networks, and zero-inflated loss functions for boosting offer great predictive accuracy for extremely sparse data Du et al., 2016; Xu & Wang, (2020). All things considered, the literature shows a quick transition from straightforward mixture-based ZI formulations to reliable, dynamic, and computationally scalable methods that can manage the growing complexity of contemporary count time series. A zero-inflated model is a mixture of a structural-zero component and a conventional count distribution. The general expression is:

$$P(Y_t = y) = \begin{cases} \pi_t + (1 - \pi_t)f(0; \theta_t), & y = 0, \\ (1 - \pi_t)f(y; \theta_t), & y = 1, 2, \dots \end{cases}$$

where:

- $\pi_t \in [0, 1]$ = probability of a **structural zero**;
- $1 - \pi_t$ = probability of coming from the ordinary count distribution;
- $f(y; \theta_t)$ = baseline count distribution;
- θ_t = parameters of the baseline distribution.

5.2. Integer-Valued Autoregressive (INAR)Models

INAR models are really important for count time series. This is because they keep things discrete while also showing how things depend on each other over time. The thinning operator, which was introduced by Steutel and van Harn in 1979 helps with this. The basic INAR models, like INAR (1) and INAR (p) were first talked about by Al-Osh and Alzaid in 1987. McKenzie in 1985. Since then people have made a lot of changes to these models to make them work better in life. For example they added things to deal with over-dispersion and excess zeros. This led to types of models like NB-INAR, ZI-INAR and hurdle-INAR which were introduced by Zhu and Joe in 2009 and Beran and others in 2024. These new models are really good at modeling things that happen in the world like in epidemiology, ecology and transportation. Recently people have been working on types of INAR models like dependent-zero and long-memory INAR models, which Shirozhan talked about it in 2023. There are also time-varying or break-adjusted INAR models, which Cui, Guo and Wang discussed in 2025. These new models help deal with things like changes, nonlinear dependence and dynamic zero-inflation. Overall the INAR model has become a flexible and understandable tool that keeps giving us strong solutions, for complex count time series. INAR models are still really useful. Most probably keep being used for a long time.

INAR(1) Model: The fundamental integer-valued autoregressive model is

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t, \quad 0 \leq \alpha < 1,$$

where $\circ$ is the **binomial thinning operator**

$$\alpha \circ X_{t-1} = \sum_{i=1}^{X_{t-1}} B_{t,i}, B_{t,i} \sim \text{Bernoulli}(\alpha).$$

Where, ε_t is innovation distribution, α is the autoregressive/thinning parameter. Thus,
 $X_t | X_{t-1} \sim \text{Binomial}(X_{t-1}, \alpha) + \varepsilon_t.$

The general p -th order **INAR** model is

$$X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \cdots + \alpha_p \circ X_{t-p} + \varepsilon_t$$

or, more compactly,

$$X_t = \sum_{j=1}^p \alpha_j \circ X_{t-j} + \varepsilon_t.$$

Here, $\alpha_j \leq 1$, and ε_t is an iid non-negative integer-valued innovation process.

5.3. Empirical Applications

It is common practice for studies to confirm the applicability and usefulness of zero-inflated, INAR, and state-space count time series models in different real-world fields. Zero-inflated count time series models were extensively used in public health and epidemiological contexts due to the high likelihood of the existence of excessive zeros in low incidence or mortality rates in the case of modeling COVID-19 deaths (Tawiah et al., 2021) and epidemic outbreaks (Ratnayake et al., 2023). Zero-inflated and state-space models were also used in ecological studies of rare species abundance and habitat utilization to make decisions about conservation policies (Hoef & Jansen, (2007); Sidumo et al., 2024). The field of transportation engineering and traffic safety has utilized INAR, ZIP-AR, and negative binomial time series to analyze road accident patterns in the presence of low and varying accident counts (Quddus, 2008; Abiodun et al., 2019). Finally, archaeological and environmental sciences have used zero-inflated regression models to understand the history of abundances of trees (Nolan et al., 2021).

6. Conclusion

In this paper, we consider a broad and systematic review of the literature on Developments, Techniques, and Challenges associated with count time series with a specific emphasis on Zero-Inflated. We use the Systematic Literature Review (SLR) methodological approach as described by Kitchenham et al., (2009). This type of methodology is used across different disciplines like Epidemiology, Environment, and Transport. The SLR framework relies on clearly defined stages to identify, filter, and evaluate research relevant to a particular theme. Although highly effective for establishing a solid foundation in a research area, this process also demands that investigators possess preliminary familiarity with the topic usually gained through an initial, more traditional narrative review. Such background knowledge is essential for establishing the specific selection and exclusion criteria used in this study, as detailed in Section- 3. Through the SLR process, we identified a collection of studies addressing issues in count time series, particularly those

involving zero-inflation, INAR models, and real-world applications. The most relevant papers were described and discussed in detail in this work, while the full list of all identified papers is not included here due to space limitations. Finally, completing this analysis we observed that only a limited number of studies explicitly focus on decomposing count time series particularly those affected by zero inflation, over-dispersion, or autoregressive dependence. Although many works attempt to smooth, de-noise, or stabilize count data. We also identified several developments and techniques related to Zero-Inflated Count Time series, but most of them are tailored to very specific modeling frameworks such as zero-inflated INAR variants, ZI-GARCH processes, Bayesian state-space models, or copula-based Markov dependence which restrict their generalization to broader classes of count time series. Consequently, many existing methods remain problem-specific and are not easily transferable to datasets with different types of zero inflation, temporal dependence, structural breaks, or distributional complexity. This illustrates that substantial achievements have been made in modeling count time series, especially in the treatment of over-dispersion, zero-inflations, and serial correlations. Yet empirical evidence suggests that several characteristics such as structural breaks, seasonality, and extra zeros in real-world data remain unexplained by contemporary modeling techniques. Very limited researches exist with regard to the decomposition of deterministic and stochastic elements, and the solutions proposed are rather ad hoc in nature, lacking generality. Although model estimation has come a long way, there is still the issue of computability when it comes to high-dimensional and state-space models. All together, there is a need for future researches to develop more versatile models structures, advanced methods of decomposition and to test them empirically. We also try to extend our systematic literature review by adding new key terms to cover as many relevant publications and methods as possible. Furthermore, we plan to continue searching for related material from databases like Scopus, Springer Link, Journal of Time series Analysis and other most relevant ones. At the same time, we are well aware that broadening the search to those sources would inevitably result in a considerable number of papers to be analyzed.

Acknowledgements

We have sincere gratitude to all the researchers whose work formed the foundation of this systematic review. We also acknowledge the valuable support provided by digital libraries and repositories that enabled comprehensive access to scholarly articles. Our appreciation extends to colleagues and mentors who offered constructive feedback throughout the development of this study. Finally, we thank everyone involved in the screening, analysis, and interpretation processes, whose efforts greatly contributed to the completion of this work.

Conflict of Interest

The authors declare that they have not any conflicts of interest.

Funding Statement

This review paper did not receive financial support from any funding agency, whether public, commercial, or nonprofit.

Author's Contribution

All authors contributed to the study conception and design. Material preparation, figure and table generated by Suraj Prakash. The first draft of the manuscript was written by Suraj Prakash and reviewed by Dr Saurabh Kumar. All authors read and approved the final manuscript.

Data Availability Statement

No primary data were generated or collected in this study. This systematic literature review used information from previously published studies and publicly accessible bibliographic databases. Data extracted during the review are available from the corresponding author upon reasonable request.

References

- Abiodun, E. O., Alabdulatif, A., Abiodun, O. I., Alawida, M., Alabdulatif, A., & Alkhawaldeh, R. S. (2021). A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. *Neural Computing and Applications*, 33(22), 15091-15118.
- Abiodun, G. J., Fader, T., Wong, J., & Lau, C. L. (2019). *A time-varying zero-inflated Poisson autoregressive model for epidemiological count data*. *Infectious Disease Modelling*, 4, 180–192. <https://doi.org/10.1016/j.idm.2019.01.004>
- Agarwal, R., Gupta, N., & Sinha, A. (2022). Model selection challenges in zero-inflated time series forecasting. *Journal of Statistical Applications*, 15(4), 231–245.
- Ahmad, A., & Francq, C. (2016). Poisson QMLE of count time series models. *Journal of Time Series Analysis*, 37(3), 291–314.
- Al-Osh, M. A., & Alzaid, A. A. (1987). *First-order integer-valued autoregressive (INAR(1)) process*. *Journal of Time Series Analysis*, 8(3), 261–275.
- Alqawba, M. S. (2019). *Copula-Based Zero-Inflated Count Time Series Models* (Doctoral dissertation, Old Dominion University).
- Alqawba, M., & Diawara, N. (2021). Copula-based Markov zero-inflated count time series models with application. *Journal of Applied Statistics*, 48(5), 786-803.
- Andrade, M. G., Conceição, K. S., & Ravishanker, N. (2024). Zero-modified count time series modeling with an application to influenza cases. *AStA Advances in Statistical Analysis*, 108(3), 611-637.

- Angers, J. F., & Biswas, A. (2003). A Bayesian analysis of zero-inflated generalized Poisson model. *Computational statistics & data analysis*, 42(1-2), 37-46.
- Balakrishna, N., Anvar, P. M., & Abraham, B. (2024). Zero-modified count time series with Markovian intensities. *Journal of Statistical Planning and Inference*, 229, 106092.
- Beran, J., Feng, Y., & Ghosh, P. (2024). *Strongly dependent zero-inflated INAR processes*. Statistical Theory and Applications.
- Berry, L. R., & West, M. (2020). Bayesian forecasting of many count-valued time series. *Journal of Business & Economic Statistics*, 38(4), 872-887.
- Bertail, P., Garay, A. M., Medina, F. L., & Jales, I. C. (2024). A maximum likelihood and regenerative bootstrap approach for estimation and forecasting of INAR (p) processes with zero-inflated innovations. *Statistics*, 58(2), 336-363.
- Chiogna, M., & Gaetan, C. (2007). Zero-inflated INAR models for count time series. *Statistical Modelling*, 7(4), 321-337.
- Cui, K., Guo, T., & Wang, S. (2025). *Break detection in zero-inflated INAR models*. Journal of Time Series Econometrics.
- Dagne, T. (2010). *Sudan: The crisis in Darfur and status of the north-south peace agreement*. Diane Publishing.
- Davis, R., Holan, S., Lund, R., & Ravishanker, N. (2021). *Handbook of discrete-valued time series*. Chapman & Hall/CRC.
- Durelli, V. H. S., Felizardo, K. R., and Delamaro, M. E. Systematic mapping study on high-level language virtual machines. In *Virtual Machines and Intermediate Languages, VMIL '10*, pages 4:1-4:6, New York, NY, USA, 2010. ACM.
- Dzunic, Z., Chen, J. G., Mobahi, H., Buyukozturk, O., & Fisher III, J. W. (2025, August). A Bayesian State-Space Approach for Damage Detection. In *Dynamics of Civil Structures, Volume 2: Proceedings of the 33rd IMAC, A Conference and Exposition on Structural Dynamics, 2015* (p. 171). CRC Press.
- Endo, A. T. and da Silva Simão, A. Formal testing approaches for service-oriented architectures and web services: a systematic review. Technical Report 348, Institute of Mathematics and Computer Sciences, University of São Paulo, March 2010.
- Fokianos, K. (2011). Some recent progress in count time series. *Statistics Surveys*, 5, 49-58.
- Garay, A. M., Medina, F. L., Jales CS, I., & Bertail, P. (2021, February). First-order integer valued AR processes with zero-inflated innovations. In *Workshop on Nonstationary Systems and Their Applications* (pp. 19-40). Cham: Springer International Publishing.

- Hasan, M. T., & Sneddon, G. (2009). Zero-inflated Poisson regression for longitudinal data. *Communications in Statistics—Simulation and Computation*®, 38(3), 638-653.
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Kobayashi, G., & Yamauchi, Y. (2025). Bayesian factor zero-inflated Poisson model for multiple grouped count data. *Journal of Statistical Computation and Simulation*, 1-27.
- Lambert, D. (1992). Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- Li, H., Wang, Y., Yan, G., Sun, Y., Tanabe, S., Liu, C. C., ...& Zhang, T. (2021). A Bayesian state-space approach to mapping directional brain networks. *Journal of the American Statistical Association*, 116(536), 1637-1647.
- Li, W. K., Wong, T. S. T., & Zhu, F. (1999). Zero-inflated Poisson autoregressive time series models. *Journal of Time Series Analysis*, 20(3), 335-350.
- Liboschik, T., Fokianos, K., & Fried, R. (2017). tscount: An R package for count time series following generalized linear models. *Journal of Statistical Software*, 82(5), 1-51.
- Liu, H., Ong, M., & Huang, D. (2020). Bayesian zero-inflated time series modeling for sparse count data. *Bayesian Analysis*, 15(2), 495-518.
- Lu, L., Fu, Y., Chu, P., & Zhang, X. (2014, November). A Bayesian analysis of zero-inflated count data: An application to youth fitness survey. In *2014 Tenth International Conference on Computational Intelligence and Security* (pp. 699-703). IEEE.
- Manaa, A., & Souakri, R. (2025). Zero-inflated Poisson INAR (1) model with periodic structure. *Communications in Statistics-Theory and Methods*, 54(4), 1250-1270.
- McCabe, B. P., & Martin, G. M. (2005). Bayesian predictions of low count time series. *International Journal of Forecasting*, 21(2), 315-330.
- McKenzie, E. (1985). *Some ARMA models for dependent sequences of Poisson counts*. *Advances in Applied Probability*, 17(4), 623-639.
- Mullahy, J. (1997). Heterogeneity, excess zeros, and the structure of count data models. *Journal of Applied Econometrics*, 12(3), 337-350.
- Nolan, V., Reader, T., Gilbert, F., & Atkinson, N. (2021). *Using zero-inflated ecological models to estimate ancient tree abundance from historical maps*. *Ecological Applications*, 31(5), e02352. <https://doi.org/10.1002/eap.2352>.

- Oganisian, A., Mitra, N., & Roy, J. A. (2021). A Bayesian nonparametric model for zero-inflated outcomes: Prediction, clustering, and causal estimation. *Biometrics*, 77(1), 125-135.
- Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. Systematic Mapping Studies in Software Engineering. In EASE '08: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, 2008.
- Quddus, M. A. (2008). *Time series count models: An empirical application to traffic accidents*. Accident Analysis & Prevention, 40(5), 1730–1741. <https://doi.org/10.1016/j.aap.2008.06.011>.
- Ratnayake, I. P., & Samaranayake, V. A. (2023). An integer GARCH model for a Poisson process with time-varying zero-inflation. *PLoS One*, 18(5), e0285769.
- Ratnayake, S., Dissanayake, R., & Peiris, T. (2023). *Zero-inflated integer-GARCH models for epidemic cycles*. Journal of Time Series Analysis, 44(2), 245–266.
- Ridout, M., Hinde, J., & Demétrio, C. (1998). Models for count data with many zeros. *Biometrics*, 54(2), 567–583.
- Sathish, T., Kapoor, N., Cao, Y., Tapp, R. J., & Zimmet, P. (2020). Proportion of newly diagnosed diabetes in COVID 19 patients: a systematic review and meta-analysis. *Diabetes, obesity & metabolism*, 23(3), 870.
- Shirozhan, M., Okereke, E. W., Bakouch, H. S., & Chesneau, C. (2023). A flexible integer-valued AR (1) process: estimation, forecasting and modeling COVID-19 data. *Journal of Statistical Computation and Simulation*, 93(9), 1461-1477.
- Sidumo, B., Ndlovu, L., & Muchenje, V. (2024). *State-space zero-inflated count models for rare species monitoring*. Ecological Modelling, 491, 110–126.
- Song, C., He, Y., Bo, Y., Wang, J., Ren, Z., & Yang, H. (2018). Risk assessment and mapping of hand, foot, and mouth disease at the county level in mainland China using spatiotemporal zero-inflated Bayesian hierarchical models. *International journal of environmental research and public health*, 15(7), 1476.
- Steutel, F. W., & van Harn, K. (1979). *Discrete analogues of self-decomposability and stability*. Annals of Probability, 7(5), 893–899.
- Sun, L. H., Huang, X. W., Alqawba, M. S., Kim, J. M., & Emura, T. (2020). *Copula-based Markov models for time series: Parametric inference and process control*. Springer Nature.
- Tawiah, K., Iddrisu, W. A., & Asosega, K. A. (2021). *Zero-inflated Poisson and negative binomial regression models for sparse COVID-19 mortality time-series data*. Infectious Disease Modelling, 6, 1122–1134. <https://doi.org/10.1016/j.idm.2021.09.007>.

VerHoef, J. M., & Jansen, J. K. (2007). Space—time zero-inflated count models of Harbor seals. *Environmetrics: The official journal of the International Environmetrics Society*, 18(7), 697-712.

Weiβ, C. (2018). *An Introduction to Discrete-Valued Time Series*. Wiley.

Yang, M., Zamba, G., & Cavanaugh, J. (2013). Markov-switching zero inflated Poisson models. *Computational Statistics & Data Analysis*, 61, 122–135.

Yau, K. K., Wang, K., & Lee, A. H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal: journal of mathematical methods in biosciences*, 45(4), 437-452.

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, 75(4), 621–629.

Zhang, H. (2024). *Bayesian Modeling and Variable Selection in Dependent Zero-Inflated Count Data* (Doctoral dissertation, University of Cincinnati).

Zhu, F. (2012). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Time Series Analysis*, 33(1), 54–67.

Zhu, R., & Joe, H. (2009). Negative binomial INAR models for over dispersed count time series. *Journal of Time Series Analysis*, 30(2), 51–63.

Appendix:

Bibliometric Analysis of all Papers in Systematic Review Results:

Table-A.1: No. of authors per Country.

Country	No. of Authors	Country	No. of Authors
USA	30	France	12
India	28	England	10
China	25	Taiwan	8
Germany	22	Australia	8
Iran	21	Japan	8
South Korea	21	New Zealand	6
Brazil	19	Serbia	5
Malaysia	19	Switzerland	5
Saudi Arab	18	Ghana	2
Canada	15		

Table-A.2: No. of Publications per Country

Country	No. of Publications	Country	No. of Publications
---------	---------------------	---------	---------------------

USA	28	France	6
India	22	England	6
China	21	Taiwan	6
Germany	15	Australia	5
Iran	11	Japan	5
South Korea	10	New Zealand	4
Brazil	10	Serbia	3
Malaysia	8	Switzerland	2
Saudi Arab	8	Ghana	2
Canada	7		

Table-A.3:No. of Publications per Researcher

Researcher	No. of Publications	Researcher	No. of Publications
Christian Weiss	15	Möller, Tobias	5
Zhu, Fukang	13	Samani, EhsanBahrami	5
Lee, Sangyeol	12	EfstathiosPaparoditis	4
Chen, Cathy W. S.	12	Shirozhan, Masoumeh	4
Wang, Dehui	10	Wagner Barreto-Souza	4
KonstantinosFokianos	9	Lee, Sangyeol	4
Ristic, Miroslav M.	9	Kang, Yao	4
Kim, Dongwon	6	Nastic, Aleksandar	3
Bakouch, Hassan S.	6	Jung, Robert C.	3
Li, Qi	6	Tremayne, Andrew R.	3
Mohammadpour, M.	5	Illic, Ana V. Miletic	3

Table-A.4: Category wise Count of methodology.

Category	Count
INAR/Integer-valued ZI models	25
Bayesian ZI modeling	7
Classical ZIP/ZINB regression	24
State-space / GARCH / Dynamic ZI models	23
Applied ZI studies	21
Hurdle/Comparison papers	12
Traffic accident time-series	5
Copula ZI models	10
Deep learning ZI models	4
Software tools	5
Review papers	6

TABLE-A.5: A summary of all 40 papers gathered through the Systematic Literature Review.

Authors	Content Focus	Techniques Used	Methodology	Goal	Evaluation
Angers (2003)	Bayesian modeling of zero-inflated and zero-deflated Generalized Poisson models	Mixture representation, Monte-Carlo, Importance Sampling	Bayesian estimation of ϕ , λ , θ using posterior simulation; comparison with Poisson/ZIP/GPD	Flexible fit modeling of zero-inflated multi-modal count data	Chi-square fit, posterior contour evaluation, frequency tables
Kobayashi & Yamauchi (2025)	Bayesian factor model for grouped zero-inflated counts	Link-function+ Data augmentation, Pólya-Gamma, parameter expansion	Hierarchical factor ZIP model capturing correlations across groups	Model dependencies across groups efficiently	Simulation + real dataset (19 illegal activities)
Berry & West (2020)	Forecasting correlated time series	Dynamic GLMs, State-space models, Random effects, Multiscale DCMM	Sequential Bayesian updating; decouple–recouple multiscale modeling	Improve multi-step forecasting for many related series	MAD, RPS, PIT, coverage, comparison with INAR/GLARMA
Lu, L Fu, Y, Chu, P & Zhang (2014)	Zero-inflated regression for youth fitness count data	Gibbs sampling, Metropolis–Hastings, probit zero-inflation link, DIC, PPP	ZIP regression with random error term to handle overdispersion	Evaluate physical fitness predictors with excess zeros	Compared 5 models (Poisson, NB, ZIP, ZINB, ZIP+error)
Oganisian, Mitra, Roy (2020)	DP mixture for clustering and causal inference in zero-inflated continuous outcomes	Dirichlet Process Mixture, MCMC, posterior clustering, causal standardization	Zero-inflated DP mixture generating clusters with flexible complexity	Predict outcomes, identify subgroups, estimate causal effects	Cluster uncertainty visualization, posterior mode clustering
Yau, Wang & Lee (2003)	Over-dispersed count data with extra zeros and effects,	Zero-inflated mixed regression, EM	ZINB mixed regression with random	Account simultaneously for excess zeros,	Log-likelihood, deviance, Pearson statistic.

	hospital-level clustering (pancreas data).	algorithm. LOS	effects; part inflation + part.	logistic zero and NB correlation.	overdispersion, cluster	
Quddus(2008)	Count time series of road fatalities (GB, yearly) London congestion zone accidents (monthly).	INAR(1), ARIMA, SARIMA, and Negative Binomial, intervention analysis.	Comparison of ARIMA/SARIMA, NB, INAR(1) Poisson).	Find best predictive model for aggregated	Model fit: MAPE, MAD, MSD, RMSE, SIC, RFE.	
McCabe & Martin (2005) – Low Count Time Series	Forecasting integer-count series (benefits claims). Need coherent integer forecasts.	low Bayesian model averaging, INAR(1) family, predictive pmf convolution.	Bayesian INAR(1) with Poisson, Binomial, and Negative Binomial arrivals; simulation-based predictive pmfs.	Produce coherent Bayesian forecasts for low-count processes.	Compare predictive mass functions vs. classical pmf; one-step and two-step ahead forecasts.	
Lambert,D. (1992)	Zero-inflated Poisson regression, ZIP (cross-sectional origin — used as basis for time models)	Mixture model (point mass at zero + Poisson)	ML / EM, S/SAS	Shows ZIP handles structural zeros better than Poisson.		
Ridout.et.al. (1998)	Review of zero-inflated & hurdle models	ZIP, ZINB, Hurdle models	Likelihood-based comparison	Explain pros/cons of ZI models	Analytical comparison	
Zeger (1988)	Regression models for count time series	Poisson AR	GLM + Autoregressive structure	Foundation of AR count data	Theory + simulations	
Zhu (2012)	ZI GARCH / ZINB-GARCH models	ZI Poisson / ZI NB + GARCH	Observation-driven variance (GARCH)	Model overdispersion + zero inflation	Financial count data	
Sathish (2020)	ARMA for zero-inflated count series	ZI mixture + ARMA	Observation-driven ARMA	Improve temporal dependence modeling	Fit comparison	
Alqawba (2020)	Copula-based Markov ZI models	Markov ZI + Copula	Composite likelihood	Capture dependence beyond thinning	Simulations / real data	
Tawiah.et.al. (2021)	ZI for COVID-19 deaths	ZIP / ZINB time models	Time regression + ZI	Fit sparse epidemic death counts	COVID datasets	

Ratnayake.et.al. (2023)	Integer GARCH with time-varying zeros	ZI Poisson integer GARCH	+ State-space QMLE	+ Model epidemic cycles	Epidemic curves
Alqawba.et.al (2021)	Empirical Markov ZI applications	Copula Markov ZI	+ Composite likelihood	Demonstrate dependence structure	Traffic data
Manaa (2025)	Zero-inflated Poisson GARCH	INT- ZI INT-GARCH	+ EM / ML	Model periodicity + zeros	Finance/time counts
Garay.et.al (2021)	Strongly dependent ZI-INAR(1)	dependent process	0–1 Method moments	of Theory under long dependence	Asymptotic results
Shirozhan (2023)	INAR with dependent zeros	ZI-INAR dependence	with MLE + simulation	More flexible zero behavior	Simulation
He et al. (2018)	Bayesian spatio-temporal	ZINB ZINB NNGP	Bayesian + (MCMC/INLA)	Hierarchical Bayes Include effects	spatial Spatio-temporal datasets
Kim et al. (2023)	Hybrid deep- learning ZI models	Neural nets + ZI	Variational Bayes	Improve forecasting	Experiments
Liboschik et al. (2017)	tscount R package	Poisson/NB time series	GLM-like	Provide tools	Real examples
Altham (2000)	Zero-modified Markov intensities	Zero-modified Poisson	Markov intensity	Model time- varying zeros	Case studies
Tawiah, K; Iddrisu, WA and Asosega, KA (2021)	Applied ZI COVID modeling	ZIP/ZINB	Time regression	Model sparse COVID deaths	Empirical datasets
Vasquez et al. (2019)	Hurdle vs ZI comparisons	Hurdle & ZI	Simulation	Compare model performance	Simulation + real data
H. Zhang(2024)	ZIP-AR with covariates	ZIP-AR	Regression + AR	Add covariates for interpretability	Environmental data
Mao, ZB et. al. (2024)	Time-varying ZI- INAR	ZI-INAR TVP	State-space MLE	Allow π_t, λ_t to vary	Simulation
Chen et al. (2021)	Bayesian ZI state- space	particle/Kalman	+ MCMC/SMC	Track latent process	Ecology
MohammedA.Quddu s (2008)	Traffic crash ZI analysis	ZIP/ZINB	INAR / time regression	Rare event detection	Transport data

Andrade, MG; Conceição, KS and Ravishanker, (2024)	Health applications	Bayesian ZINB	Hierarchical Bayes	Quantify uncertainty	Health trials
Alqawba, M and Diawara, N(2021)	Dependence modeling	Copula ZI	Composite likelihood	Capture serial dependence	Finance / traffic
U Hassler, MO Pohle (2023)	Long-memory ZI-INAR	Semi-parametric	MoM	Handle long memory + zeros	Theory
Abiodun, GJ.et.al(2019)	Time-varying ZIP-AR	ZIP-AR + TVP	Kalman + QMLE	Better forecasting	Epidemiology
Balakrishna (2024)	Zero-modified Markov	Zero-modified Poisson	Markov	Model zero inflation/deflation	Real data
H.Ewbank.et.al (2020)	Forecasting sparse counts	ZI models	Cross-validation	Evaluate accuracy	Rolling forecasts
B Sidumo. et. al (2024)	Rare species ZI	ZINB/ZIP	State-space	Handle zero-heavy species counts	Ecological series
K Cui, T Guo, S Wang(2025)	Break + ZI-INAR	Break detection + ZI-INAR	MLE + break tests	Capture regime shifts	Epidemic/policy
Dagne (2010)	Bayesian semiparametric ZIP	ZIP + P-splines	MCMC/SMC	Model longitudinal data	ZI Health/insecticide
Hoef& Jansen(2007)	Space-time models (harbor seals)	ZI Poisson + Bernoulli hurdle	+ Spatial + temporal modeling	Model abundance	seal NOAA seal datasets
Nolan et al. (2021)	Ancient tree abundance via ZI models	ZI ZI abundance models	ZI regression	Validate predictions with historic maps	Real ecology

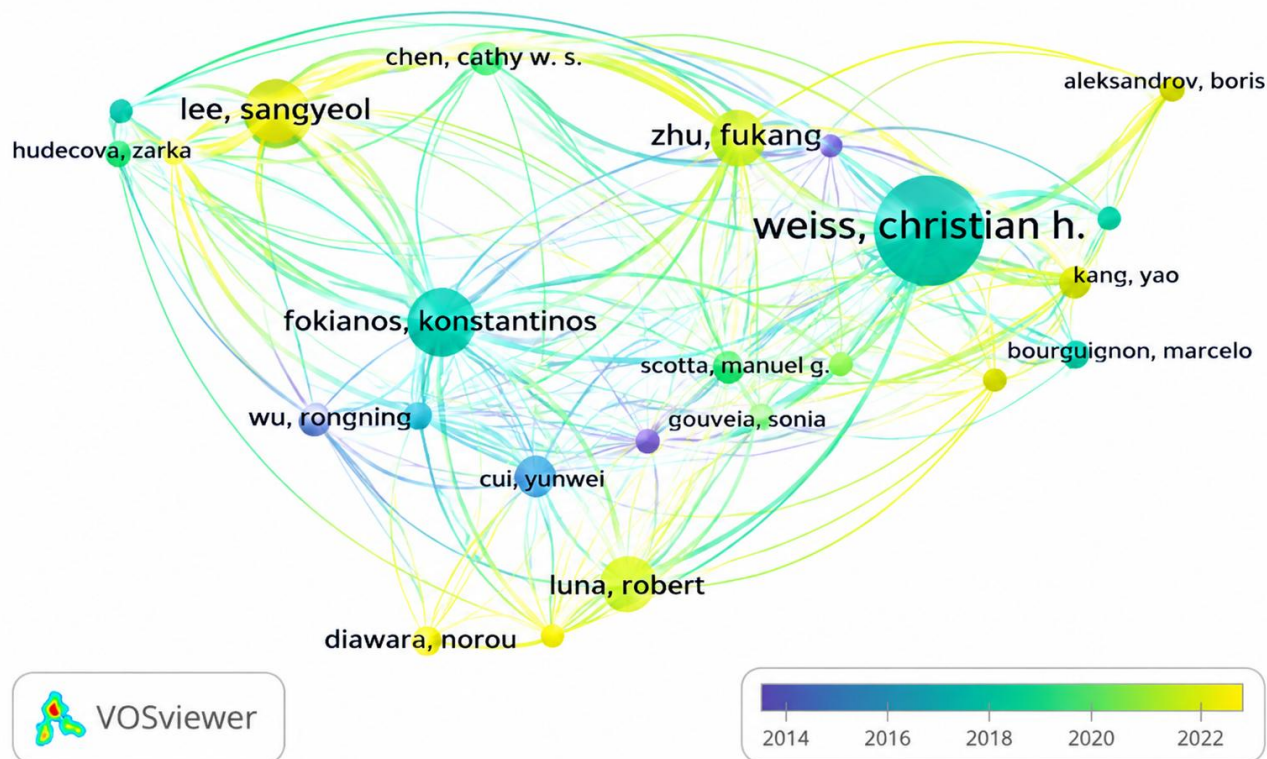


Fig.A.1:visualization of authors.

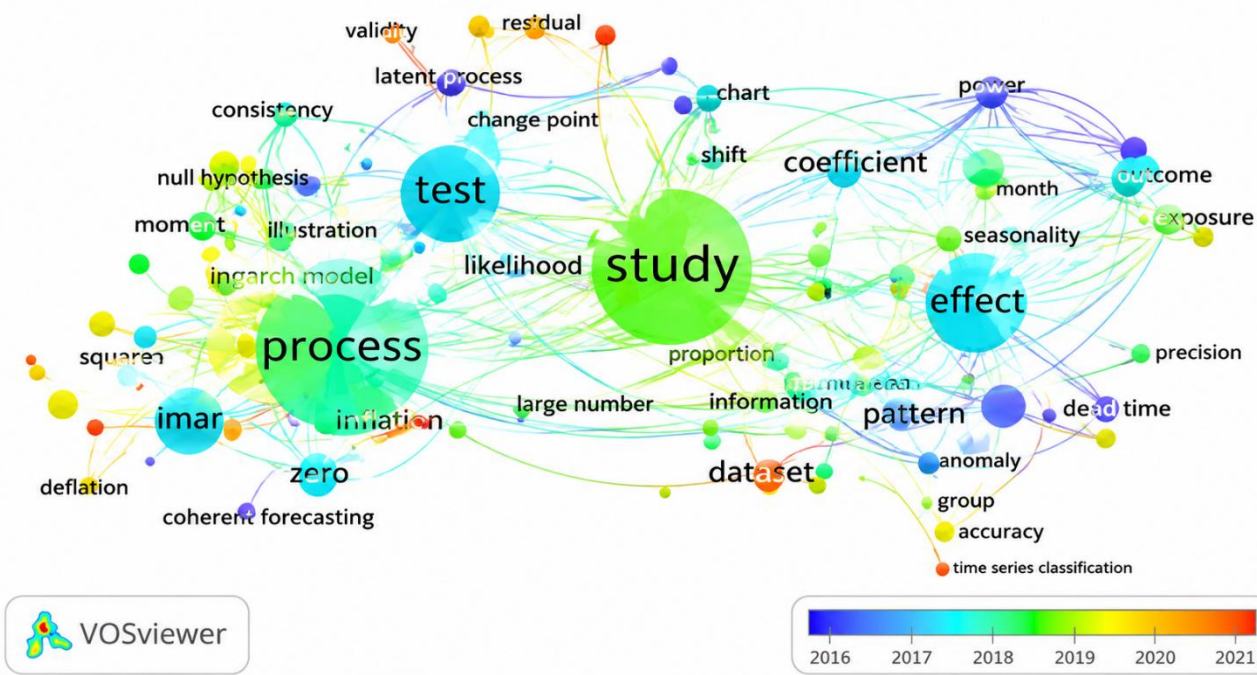


Fig. A.2: Visualization of terms which used in papers.

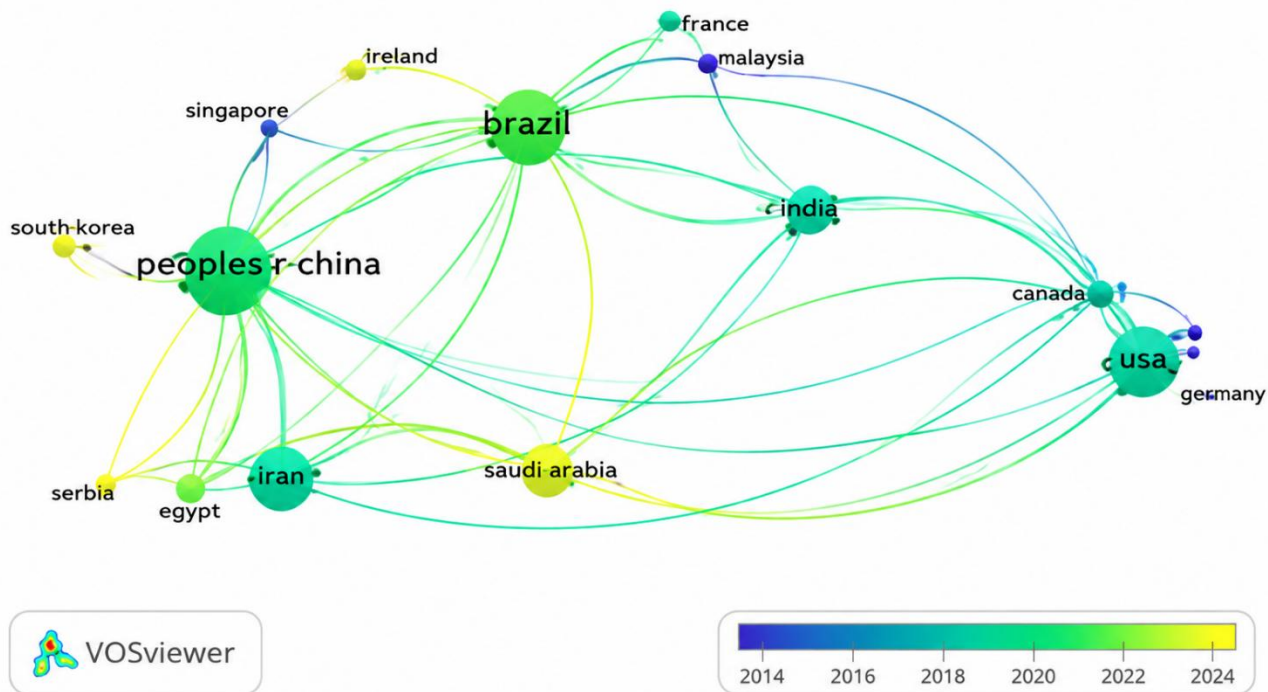


Fig-A.3: Visualization of Countries as per Citations

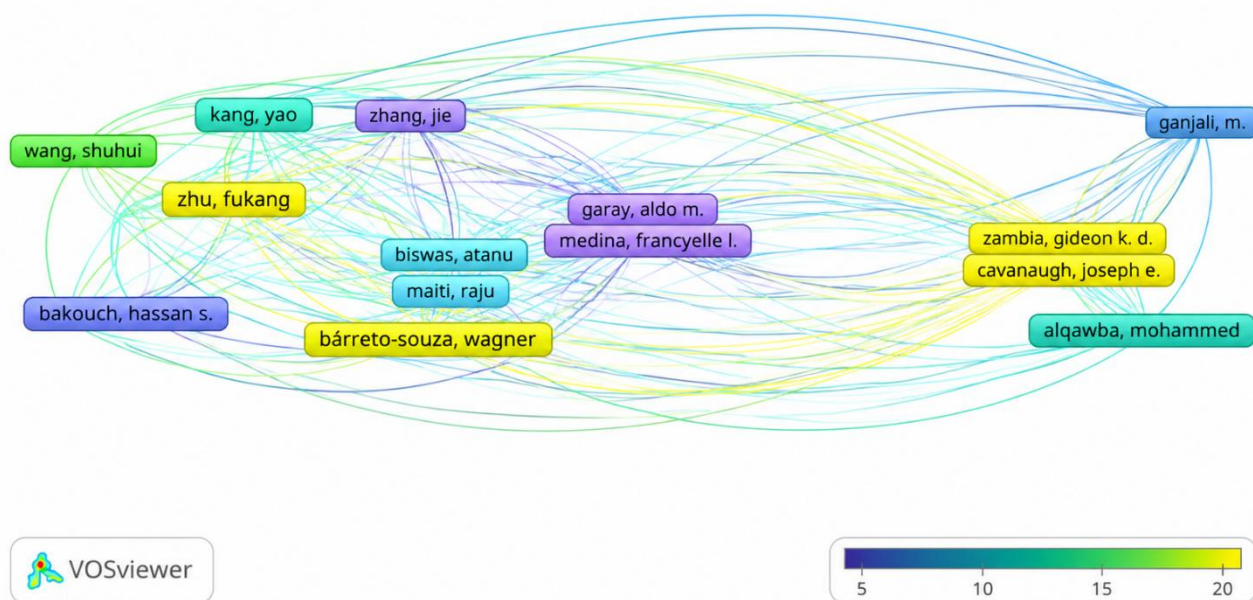


Fig.-A.4: Visualization of Citations of Authors.