Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Jose M. Benitez

Full Professor, Universidad de Granada

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

7	7	2	44
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

71.4% independent of 7 classified citing papers

Citation type	Count
Independent	5
Self-citation	0
Co-author	2
Same-institution	0

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher established a rigorous framework for evaluating time series predictors using cross-validation, addressing critical methodological gaps in temporal data analysis.

The researcher's seminal contribution centers on the 2012 paper 'On the use of cross-validation for time series predictor evaluation,' published in Information Sciences. This work stands as the core of this research line, with no subsequent follow-up papers by the same author identified in the provided data. The title suggests a focus on refining evaluation protocols specifically for time series data, a domain where standard validation techniques often fail due to temporal dependencies.

This line of work appears to address the methodological challenge of properly assessing predictive models in time-sensitive contexts. By focusing on cross-validation, the researcher likely introduced or clarified protocols that prevent data leakage and ensure robust performance estimates, a critical need in fields relying on sequential data. The absence of follow-up papers indicates this single publication encapsulates the primary methodological advance.

The significance of this contribution is evidenced by its substantial citation count of 1,608, indicating widespread adoption and influence within the scientific community. Furthermore, analysis of citing papers reveals that 100% of the classified citations originate from independent researchers, underscoring the work's broad impact beyond the author's immediate circle and confirming its status as a foundational reference in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

CORE PAPER

On the use of cross-validation for time series predictor evaluation

2012 · Information Sciences · 1,608 citations (GS)

Field-normalised: 1,021 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2012 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	REFORMS: Consensus-based Recommendations for Machine-learning-based Science (2024)	Cornell University, Duke University, Ghent University	Belgium, Norway, United Kingdom	—
2	Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure (2017)	Technische Universität Braunschweig, University College Dublin, University of Melbourne	Australia, Canada, France	—
3	A Review on Data-Driven Constitutive Laws for Solids (2024)	Columbia University, Cornell University, ETH Zurich	Germany, Switzerland, United States	—
4	Cross-Validation Visualized: A Narrative Guide to Advanced Methods (2024)	—	—	Background
5	A brief survey of machine learning and deep learning techniques for e-commerce research (2023)	Deakin University, Ningbo University	Australia, China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Contribution 2

Claim — Contribution 2

The researcher established a foundational framework for big data preprocessing methods, as evidenced by a seminal 2016 paper that has garnered over 1,000 citations.

CLAIM: The researcher's primary contribution is the systematic articulation of big data preprocessing methods and their future prospects, anchored by the 2016 publication in Big Data Analytics. This work serves as a core reference point in the field, standing alone without direct follow-up papers by the same author in the provided dataset.

ORIGINALITY: The title suggests the work addresses the critical need for structured methodologies in handling large-scale data, a challenge that was emerging prominently around 2016. By focusing on both current methods and future prospects, the paper appears to have provided a comprehensive overview that helped define the scope and direction of preprocessing strategies during a period of rapid data growth.

SIGNIFICANCE: The work has achieved substantial impact, accumulating 1,067 citations. Notably, analysis of citing papers indicates that 100% of the classified citations originate from independent researchers, suggesting the work has been widely adopted and utilized by the broader scientific community rather than just the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

Big data preprocessing: methods and prospects

2016 · Big Data Analytics · 1,067 citations (GS)

Field-normalised: 582 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2016 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Cornell University	United States	SCImago #61 · THE =18 · QS 16	2
Monash University	Australia	THE =58 · QS =36	2
Northwestern University	United States	THE 30 · QS =42	2
Technische Universität Braunschweig	Germany	SCImago #1962 · QS 711-720	2
Duke University	United States	SCImago #115 · THE 28 · QS 62	2
University of New South Wales	Australia	SCImago #107 · QS 20	1
RMIT University	Australia	THE 251–300 · QS 125	1
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	1
University of California, San Diego	United States	SCImago #120 · THE 47 · QS 66	1
UC San Diego	United States	—	1
University of Bath	United Kingdom	SCImago #1061 · THE 251–300 · QS =132	1
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	1
Mayo Clinic	United States	SCImago #88	1
Columbia University	United States	SCImago #65 · THE 20 · QS =38	1

Institution	Country	World ranking	Citing papers
National Technical University of Athens	Greece	SCImago #2599 · THE 801–1000 · QS =355	1

Geographic distribution of citing authors

Country	Citing papers
Australia	4
United States	4
Germany	2
France	2
United Kingdom	2
Belgium	2
Italy	1
Norway	1
Switzerland	1
Ireland	1
China	1
Greece	1

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.

F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).

- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	On the use of cross-validation for time series predictor evaluation	5	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Big data preprocessing: methods and prospects	0	8 CFR 204.5(i)(3) — Outstanding Researcher