## References

- Bartoszuk, M. and Gagolewski, M. (2014). A fuzzy R code similarity detection algorithm. In Laurent, A. et al., editors, Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part III, volume 444 of Communications in Computer and Information Science, pages 21–30. Springer. doi:10.1007/978-3-319-08852-5 3.
- Bartoszuk, M. and Gagolewski, M. (2015). Detecting similarity of R functions via a fusion of multiple heuristic methods. In Alonso, J., Bustince, H., and Reformat, M., editors, *Proc. IFSA/EUSFLAT'15*, pages 419–426. Atlantis Press. doi:10.2991/ifsa-eusflat-15.2015.61.
- Bartoszuk, M. and Gagolewski, M. (2017). Binary aggregation functions in software plagiarism detection. In *Proc. FUZZ-IEEE'17*. IEEE. doi:10.1109/FUZZ-IEEE.2017.8015582. no. 8015582.
- Bartoszuk, M. and Gagolewski, M. (2020). SimilaR: R code clone and plagiarism detection. R Journal, 12(1), 367-385. doi:10.32614/RJ-2020-017. URL https://CRAN.R-project.org/package=SimilaR.
- Bartoszuk, M. and Gagolewski, M. (2021). T-norms or t-conorms? How to aggregate similarity degrees for plagiarism detection. *Knowledge-Based Systems*, **231**, 107427. doi:10.1016/j.knosys.2021.107427.
- Bartoszuk, M., Beliakov, G., Gagolewski, M., and James, S. (2016a). Fitting aggregation functions to data: Part I Linearization and regularization. In Carvalho, J. et al., editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Part II, volume 611 of Communications in Computer and Information Science, pages 767–779. Springer. doi:10.1007/978-3-319-40581-0 62.
- Bartoszuk, M., Beliakov, G., Gagolewski, M., and James, S. (2016b). Fitting aggregation functions to data: Part II Idempotization. In Carvalho, J. et al., editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II*, volume 611 of *Communications in Computer and Information Science*, pages 780–789. Springer. doi:10.1007/978-3-319-40581-0 63.
- Beliakov, G., Gagolewski, M., and James, S. (2016). Penalty-based and other representations of economic inequality. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **24(Suppl.1)**, 1–23. doi:10.1142/S0218488516400018.
- Beliakov, G., Gagolewski, M., and James, S. (2018). Least median of squares (LMS) and least trimmed squares (LTS) fitting for the weighted arithmetic mean. In Medina, J. et al., editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Theory and Foundations, pages 367–378. Springer. doi:10.1007/978-3-319-91476-3 31.
- Beliakov, G., Gagolewski, M., and James, S. (2019a). Aggregation on ordinal scales with the Sugeno integral for biomedical applications. *Information Sciences*, **501**, 377–387. doi:10.1016/j.ins.2019.06.023.
- Beliakov, G., Gagolewski, M., James, S., Pace, S., Pastorello, N., Thilliez, E., and Vasa, R. (2019b). Measuring traffic congestion: An approach based on learning weighted inequality, spread and aggregation indices from comparison data. *Applied Soft Computing*, **67**, 910–919. doi:10.1016/j.asoc.2017.07.014.
- Beliakov, G., Gagolewski, M., and James, S. (2020a). Robust fitting for the Sugeno integral with respect to general fuzzy measures. *Information Sciences*, **514**, 449–461. doi:10.1016/j.ins.2019.11.024.
- Beliakov, G., Gagolewski, M., and James, S. (2020b). DC optimization for constructing discrete Sugeno integrals and learning nonadditive measures. *Optimization*, **69**(12), 2515–2534. doi:10.1080/02331934.2019.1705300.
- Beliakov, G., Gagolewski, M., and James, S. (2022). Hierarchical data fusion processes involving the Möbius representation of capacities. Fuzzy Sets and Systems, 433, 1–21. doi:10.1016/j.fss.2021.02.006.
- Cena, A. and Gagolewski, M. (2013a). OM3: Ordered maxitive, minitive, and modular aggregation operators Part I: Axiomatic analysis under arity-dependence. In Bustince, H. et al., editors, Aggregation Functions in Theory and in Practise, volume 228 of Advances in Intelligent Systems and Computing, pages 93–103. Springer. doi:10.1007/978-3-642-39165-1\_13.
- Cena, A. and Gagolewski, M. (2013b). OM3: Ordered maxitive, minitive, and modular aggregation operators Part II: A simulation study. In Bustince, H. et al., editors, Aggregation Functions in Theory and in Practise, volume 228 of Advances in Intelligent Systems and Computing, pages 105–115. Springer. doi:10.1007/978-3-642-39165-1 14.

- Cena, A. and Gagolewski, M. (2015a). Clustering and aggregation of informetric data sets. In *Computational methods in data analysis (Proc. ITRIA'15 vol. 1)*, pages 5–26. Institute of Computer Science, Polish Academy of Sciences.
- Cena, A. and Gagolewski, M. (2015b). Aggregation and soft clustering of informetric data. In Baczyński, M., De Baets, B., and Mesiar, R., editors, *Proc. 8th International Summer School on Aggregation Operators* (AGOP 2015), pages 79–84. University of Silesia, Katowice, Poland. ISBN 978-83-8012-519-3.
- Cena, A. and Gagolewski, M. (2015c). A K-means-like algorithm for informetric data clustering. In Alonso, J., Bustince, H., and Reformat, M., editors, *Proc. IFSA/EUSFLAT'15*, pages 536–543. Atlantis Press. doi:10.2991/ifsa-eusflat-15.2015.77.
- Cena, A. and Gagolewski, M. (2015d). OM3: Ordered maxitive, minitive, and modular aggregation operators Axiomatic and probabilistic properties in an arity-monotonic setting. Fuzzy Sets and Systems, 264, 138–159. doi:10.1016/j.fss.2014.04.001.
- Cena, A. and Gagolewski, M. (2016). Fuzzy k-minpen clustering and k-nearest-minpen classification procedures incorporating generic distance-based penalty minimizers. In Carvalho, J. et al., editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, *Part II*, volume 611 of *Communications in Computer and Information Science*, pages 445–456. Springer. doi:10.1007/978-3-319-40581-0 36.
- Cena, A. and Gagolewski, M. (2017). OWA-based linkage and the Genie correction for hierarchical clustering. In *Proc. FUZZ-IEEE'17*. IEEE. doi:10.1109/FUZZ-IEEE.2017.8015652. no. 8015652.
- Cena, A. and Gagolewski, M. (2020). Genie+OWA: Robustifying hierarchical clustering with OWA-based linkages. *Information Sciences*, **520**, 324–336. doi:10.1016/j.ins.2020.02.025.
- Cena, A., Gagolewski, M., and Mesiar, R. (2015). Problems and challenges of information resources producers' clustering. *Journal of Informetrics*, **9**(2). doi:10.1016/j.joi.2015.02.005.
- Cena, A., Gagolewski, M., Siudem, G., and Zogała Siudem, B. (2022). Validating citation models by proxy indices. *Journal of Informetrics*. in press.
- Coroianu, L. and Gagolewski, M. (2019). Penalty-based data aggregation in real normed vector spaces. In Halaš, R. et al., editors, New Trends in Aggregation Theory, volume 981 of Advances in Intelligent Systems and Computing, pages 160–171. Springer. doi:10.1007/978-3-030-19494-9\_15.
- Coroianu, L., Gagolewski, M., and Grzegorzewski, P. (2013). Nearest piecewise linear approximation of fuzzy numbers. Fuzzy Sets and Systems, 233, 26-51. doi:10.1016/j.fss.2013.02.005. URL https://CRAN.R-project.org/package=FuzzyNumbers.
- Coroianu, L., Gagolewski, M., Grzegorzewski, P., Adabitabar Firozja, M., and Houlari, T. (2014). Piecewise linear approximation of fuzzy numbers preserving the support and core. In Laurent, A. et al., editors, Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II, volume 443 of Communications in Computer and Information Science, pages 244–254. Springer. doi:10.1007/978-3-319-08855-6 25.
- Coroianu, L., Gagolewski, M., and Grzegorzewski, P. (2019). Piecewise linear approximation of fuzzy numbers: Algorithms, arithmetic operations and stability of characteristics. *Soft Computing*, **23**(19), 9491-9505. doi:10.1007/s00500-019-03800-2. URL https://CRAN.R-project.org/package=FuzzyNumbers.
- Coroianu, L., Fullér, R., Gagolewski, M., and James, S. (2020). Constrained ordered weighted averaging aggregation with multiple comonotone constraints. Fuzzy Sets and Systems, 395, 21–39. doi:10.1016/j.fss.2019.09.006.
- Ferraro, M. B., Giordani, P., Vantaggi, B., Gagolewski, M., Ángeles Gil, M., Grzegorzewski, P., and Hryniewicz, O., editors (2017). Soft Methods for Data Science, volume 456 of Advances in Intelligent Systems and Computing. Springer. ISBN 978-3-319-42971-7. doi:10.1007/978-3-319-42972-4.
- Gagolewski, M. (2011a). Bibliometric impact assessment with R and the CITAN package. *Journal of Informet-* rics, 5(4), 678-692. doi:10.1016/j.joi.2011.06.006. URL https://CRAN.R-project.org/package=CITAN.
- Gagolewski, M. (2011b). Aggregation operators and their application in a formal model for quality evaluation system of scientific research (Wybrane operatory agregacji i ich zastosowanie w modelu formalnym systemu jakości w nauce). PhD thesis, Systems Research Institute, Polish Academy of Sciences. in Polish.
- Gagolewski, M. (2012). On the relation between effort-dominating and symmetric minitive aggregation operators. In Greco, S. et al., editors, Advances in Computational Intelligence, Part III, volume 299 of Communications in Computer and Information Science, pages 276–285. Springer. doi:10.1007/978-3-642-31718-7 29.

- Gagolewski, M. (2013a). Statistical hypothesis test for the difference between Hirsch indices of two Pareto-distributed random samples. In Kruse, R. et al., editors, Synergies of Soft Computing and Statistics for Intelligent Data Analysis, volume 190 of Advances in Intelligent Systems and Computing, pages 359–367. Springer. doi:10.1007/978-3-642-33042-1 39.
- Gagolewski, M. (2013b). Scientific impact assessment cannot be fair. *Journal of Informetrics*, **7**(4), 792–802. doi:10.1016/j.joi.2013.07.001.
- Gagolewski, M. (2013c). On the relationship between symmetric maxitive, minitive, and modular aggregation operators. *Information Sciences*, **221**, 170–180. doi:10.1016/j.ins.2012.09.005.
- Gagolewski, M. (2014). Programowanie w języku R. Analiza danych, obliczenia, symulacje (R Programming. Data Analysis, Computing, Simulations). Wydawnictwo Naukowe PWN, Warsaw, Poland, 1st edition. ISBN 978-83-01-17461-3. in Polish.
- Gagolewski, M. (2015a). Data Fusion: Theory, Methods, and Applications. Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland. ISBN 978-83-63159-20-7.
- Gagolewski, M. (2015b). Sugeno integral-based confidence intervals for the theoretical h-index. In Grzegorzewski, P. et al., editors, Strengthening Links Between Data Analysis and Soft Computing, volume 315 of Advances in Intelligent Systems and Computing, pages 233–240. Springer. doi:10.1007/978-3-319-10765-3 28.
- Gagolewski, M. (2015c). Some issues in aggregation of multidimensional data. In Baczyński, M., De Baets, B., and Mesiar, R., editors, *Proc. 8th International Summer School on Aggregation Operators (AGOP 2015)*, pages 127–132. University of Silesia, Katowice, Poland. ISBN 978-83-8012-519-3.
- Gagolewski, M. (2015d). Normalized  $\mathrm{WD}_p\mathrm{WAM}$  and  $\mathrm{WD}_p\mathrm{OWA}$  spread measures. In Alonso, J., Bustince, H., and Reformat, M., editors,  $Proc.\ IFSA/EUSFLAT'15$ , pages 210–216. Atlantis Press. doi:10.2991/ifsa-eusflat-15.2015.32.
- Gagolewski, M. (2015e). Spread measures and their relation to aggregation functions. European Journal of Operational Research, 241(2), 469–477. doi:10.1016/j.ejor.2014.08.034.
- Gagolewski, M. (2016a). Programowanie w języku R. Analiza danych, obliczenia, symulacje (R Programming. Data Analysis, Computing, Simulations). Wydawnictwo Naukowe PWN, Warsaw, Poland, 2nd edition. ISBN 978-83-01-18939-6. URL https://github.com/gagolews/Programowanie\_w\_jezyku\_R/. in Polish.
- Gagolewski, M. (2016b). Algorytmy i postawy programowania w języku C++ (Introduction to Algorithms and Programming in C++). e-book, Warsaw, Poland. URL https://github.com/gagolews/aipp. in Polish.
- Gagolewski, M. (2017). Penalty-based aggregation of multidimensional data. Fuzzy Sets and Systems, 325, 4–20. doi:10.1016/j.fss.2016.12.009.
- Gagolewski, M. (2021a). genieclust: Fast and robust hierarchical clustering. SoftwareX, 15, 100722. doi:10.1016/j.softx.2021.100722. URL https://genieclust.gagolewski.com.
- Gagolewski, M. (2021b). Lightweight Machine Learning Classics with R. e-book. doi:10.5281/zenodo.4539689. URL https://lmlcr.gagolewski.com/. draft v0.2.1.
- Gagolewski, M. (2022). stringi: Fast and portable character string processing in R. *Journal of Statistical Software*. URL https://stringi.gagolewski.com. in press.
- Gagolewski, M. and Grzegorzewski, P. (2009a). A geometric approach to the construction of scientific impact indices. *Scientometrics*, **81**(3), 617–634. doi:10.1007/s11192-008-2253-y.
- Gagolewski, M. and Grzegorzewski, P. (2009b). Possible and necessary h-indices. In Carvalho, J. P. et al., editors, *Proc. IFSA/EUSFLAT'09*, pages 1691–1695. IFSA.
- Gagolewski, M. and Grzegorzewski, P. (2009c). O pewnym uogólnieniu indeksu hirscha. In Kawalec, P. and Lipski, P., editors, *Kadry i infrastruktura nowoczesnej nauki: Teoria i praktyka, Proc. 1st Intl. Conf. Zarządzanie Nauką*, volume 2, pages 15–29. Wydawnictwo Lubelskiej Szkoły Biznesu, Lublin. ISBN 978-83-61671-12-1. in Polish.
- Gagolewski, M. and Grzegorzewski, P. (2010a). Arity-monotonic extended aggregation operators. In Hüllermeier, E. et al., editors, Information Processing and Management of Uncertainty in Knowledge-Based Systems, volume 80 of Communications in Computer and Information Science, pages 693–702. Springer. doi:10.1007/978-3-642-14055-6 73.

- Gagolewski, M. and Grzegorzewski, P. (2010b). Metody i problemy naukometrii (methods and problems of scientometrics). In Rowiński, T. and Tadeusiewicz, R., editors, *Psychologia i informatyka. Synergia i kontradykcje*, pages 103–125. Wyd. UKSW, Warsaw, Poland. ISBN 978-83-707-2679-9. in Polish.
- Gagolewski, M. and Grzegorzewski, P. (2010c). S-statistics and their basic properties. In Borgelt, C. et al., editors, Combining Soft Computing and Statistical Methods in Data Analysis, volume 77 of Advances in Intelligent and Soft Computing, pages 281–288. Springer. doi:10.1007/978-3-642-14746-3\_35.
- Gagolewski, M. and Grzegorzewski, P. (2011a). Axiomatic characterizations of (quasi-) L-statistics and S-statistics and the Producer Assessment Problem. In Galichet, S. et al., editors, *Proc. EUSFLAT/LFA'11*, pages 53–58. Atlantis Press. doi:10.2991/eusflat.2011.112.
- Gagolewski, M. and Grzegorzewski, P. (2011b). Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning*, **52**(9), 1312–1324. doi:10.1016/j.ijar.2011.01.010.
- Gagolewski, M. and James, S. (2018). Fitting symmetric fuzzy measures for discrete Sugeno integration. In Kacprzyk, J. et al., editors, Advances in Fuzzy Logic and Technology 2017, volume 642 of Advances in Intelligent Systems and Computing, pages 104–116. Springer. doi:10.1007/978-3-319-66824-6 10.
- Gagolewski, M. and Lasek, J. (2015a). The use of fuzzy relations in the assessment of information resources producers' performance. In *Proc. 7th IEEE International Conference Intelligent Systems IS'2014, Vol. 2: Tools, Architectures, Systems, Applications*, volume 323 of *Advances in Intelligent Systems and Computing*, pages 289–300. Springer. doi:10.1007/978-3-319-11310-4\_25.
- Gagolewski, M. and Lasek, J. (2015b). Learning experts' preferences from informetric data. In Alonso, J., Bustince, H., and Reformat, M., editors, *Proc. IFSA/EUSFLAT'15*, pages 484–491. Atlantis Press. doi:10.2991/ifsa-eusflat-15.2015.70.
- Gagolewski, M. and Mesiar, R. (2012). Aggregating different paper quality measures with a generalized h-index. *Journal of Informetrics*, **6**(4), 566–579. doi:10.1016/j.joi.2012.05.001.
- Gagolewski, M. and Mesiar, R. (2014). Monotone measures and universal integrals in a uniform framework for the scientific impact assessment problem. *Information Sciences*, **263**, 166–174. doi:10.1016/j.ins.2013.12.004.
- Gagolewski, M., Dębski, M., and Nowakiewicz, M. (2013). Efficient algorithm for computing certain graph-based monotone integrals: The  $l_p$ -indices. In Mesiar, R. and Bacigal, T., editors, *Proc. Uncertainty Modeling*, pages 17–23. STU Bratislava. ISBN ISBN:978-80-227-4067-8.
- Gagolewski, M., Bartoszuk, M., and Cena, A. (2016a). Przetwarzanie i analiza danych w języku Python (Data Processing and Analysis in Python). Wydawnictwo Naukowe PWN, Warsaw, Poland. ISBN 978-83-01-18940-2. URL https://github.com/gagolews/Analiza\_danych\_w\_jezyku\_Python. in Polish.
- Gagolewski, M., Bartoszuk, M., and Cena, A. (2016b). Genie: A new, fast, and outlier-resistant hierarchical clustering algorithm. *Information Sciences*, **363**, 8-23. doi:10.1016/j.ins.2016.05.003. URL https://genieclust.gagolewski.com.
- Gagolewski, M., Cena, A., and Bartoszuk, M. (2016c). Hierarchical clustering via penalty-based aggregation and the Genie approach. In Torra, V. et al., editors, *Modeling Decisions for Artificial Intelligence*, volume 9880 of *Lecture Notes in Artificial Intelligence*, pages 191–202. Springer. doi:10.1007/978-3-319-45656-0 16.
- Gagolewski, M., James, S., and Beliakov, G. (2019). Supervised learning to aggregate data with the Sugeno integral. *IEEE Transactions on Fuzzy Systems*, **27**(4), 810–815. doi:10.1109/TFUZZ.2019.2895565.
- Gagolewski, M., Pérez-Fernández, R., and De Baets, B. (2020). An inherent difficulty in the aggregation of multidimensional data. *IEEE Transactions on Fuzzy Systems*, **28**, 602–606. doi:10.1109/TFUZZ.2019.2908135.
- Gagolewski, M., Bartoszuk, M., and Cena, A. (2021). Are cluster validity measures (in)valid? *Information Sciences*, **581**, 620-636. doi:10.1016/j.ins.2021.10.004. URL https://github.com/gagolews/optim\_cvi.
- Gagolewski, M., Żogała Siudem, B., Siudem, G., and Cena, A. (2022). Ockham's index of citation impact. Scientometrics. doi:10.1007/s11192-022-04345-2. in press.
- Geras, A., Siudem, G., and Gagolewski, M. (2020). Should we introduce a dislike button for academic papers? Journal of the Association for Information Science and Technology, 71(2), 221–229. doi:10.1002/ASI.24231.

- Grzegorzewski, P., Gagolewski, M., and Bobecka-Wesołowska, K. (2014). Wnioskowanie statystyczne z wykorzystaniem środowiska R (Statistical Inference with R). Politechnika Warszawska, Warsaw, Poland. ISBN 978-83-93-72601-1. in Polish.
- Grzegorzewski, P., Gagolewski, M., Hryniewicz, O., and Ángeles Gil, M., editors (2015). Strengthening Links Between Data Analysis and Soft Computing, volume 315 of Advances in Intelligent Systems and Computing. Springer. ISBN 978-3-319-10764-6. doi:10.1007/978-3-319-10765-3.
- Halaš, R., Gagolewski, M., and Mesiar, R., editors (2019). New Trends in Aggregation Theory, volume 981 of Advances in Intelligent Systems and Computing. Springer. ISBN 978-3-030-19493-2. doi:10.1007/978-3-030-19494-9.
- Lasek, J. and Gagolewski, M. (2015a). Estimation of tournament metrics for association football league formats. In *Selected problems in information technologies (Proc. ITRIA '15 vol. 2)*, pages 67–78. Institute of Computer Science, Polish Academy of Sciences.
- Lasek, J. and Gagolewski, M. (2015b). The winning solution to the AAIA'15 Data Mining Competition: Tagging firefighter activities at a fire scene. In Ganzha, M., Maciaszek, L., and Paprzycki, M., editors, *Proc. FedCSIS'15*, pages 375–380. IEEE. doi:10.15439/2015F418.
- Lasek, J. and Gagolewski, M. (2018). The efficacy of league formats in ranking teams. *Statistical Modelling*, **18** (5–6), 411–435. doi:10.1177/1471082X18798426.
- Lasek, J. and Gagolewski, M. (2021). Interpretable sport team rating models based on the gradient descent algorithm. *International Journal of Forecasting*, **37**(3), 1061–1071. doi:10.1016/j.ijforecast.2020.11.008.
- Lasek, J., Szlavik, Z., Gagolewski, M., and Bhulai, S. (2016). How to improve a team's position in the FIFA ranking A simulation study. *Journal of Applied Statistics*, 43(7), 1349–1368. doi:10.1080/02664763.2015.1100593.
- Mesiar, R. and Gagolewski, M. (2016). H-index and other Sugeno integrals: Some defects and their compensation. *IEEE Transactions on Fuzzy Systems*, **24**(6), 1668–1672. doi:10.1109/TFUZZ.2016.2516579.
- Pérez-Fernández, R., De Baets, B., and Gagolewski, M. (2019). A taxonomy of monotonicity properties for the aggregation of multidimensional data. *Information Fusion*, **52**, 322–334. doi:10.1016/j.inffus.2019.05.006.
- Pérez-Fernández, R., Gagolewski, M., and De Baets, B. (2021). On the aggregation of compositional data. *Information Fusion*, **73**, 103–110. doi:10.1016/j.inffus.2021.02.021.
- Rowiński, T. and Gagolewski, M. (2007). Preferencje i postawy wobec pomocy online (attitudes towards online counselling and psychotherapy). Studia Psychologica UKSW, 7, 195–210. in Polish.
- Rowiński, T. and Gagolewski, M. (2011). Internet a kryzys. In Jankowska, M. and Starzomska, M., editors, Kryzys: Pułapka czy szansa?, pages 211–224. WN Akapit, Warsaw, Poland. ISBN 978-83-609-5885-8. in Polish.
- Siudem, G., Żogała Siudem, B., Cena, A., and Gagolewski, M. (2020). Three dimensions of scientific impact. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 117, 13896–13900. doi:10.1073/pnas.2001064117.
- Żogała Siudem, B., Siudem, G., Cena, A., and Gagolewski, M. (2016). Agent-based model for the bibliometric h-index Exact solution. European Physical Journal B, 89(21). doi:10.1140/epjb/e2015-60757-1.