

BAB II

KAJIAN PUSTAKA DAN LANDASAN TEORI

2.1 Tinjauan Pustaka

Tinjauan pustaka dilakukan dengan mengkaji penelitian-penelitian sebelumnya melalui identifikasi keterbatasan yang ada serta eksplorasi peluang integrasi antar metode. Analisis ini bertujuan untuk mengungkap kesenjangan penelitian yang masih terbuka dan membangun landasan bagi pengembangan suatu metode hibrid yang lebih adaptif, stabil, dan robust.

Sejumlah penelitian telah mengintegrasikan pendekatan Multi-Criteria Decision-Making (MCDM) ke dalam sistem pemeringkatan universitas dengan berbagai variasi metode dan sudut pandang. Gul dan Yucesan (2022), misalnya, mengusulkan model Bayesian BWM–TOPSIS dengan memanfaatkan indikator institusional yang bersumber dari TÜBİTAK untuk mengevaluasi kinerja universitas (Gul & Yucesan, 2022). Selanjutnya, Zhang et al. (2022) mengembangkan kerangka MCDM–NRSDEA tiga tahap yang dikombinasikan dengan teknik *bootstrapping* untuk menilai efisiensi penelitian universitas di Tiongkok (Zhang et al., 2022). Selanjutnya penelitian Ayyildiz et al. (2023) menggabungkan hierarchical clustering, IVN-AHP, dan VIKOR dengan pendekatan berbasis perspektif mahasiswa (Ayyildiz et al., 2023), sementara Trung Do (2024) meneliti pengaruh berbagai metode pembobotan terhadap hasil pemeringkatan sepuluh universitas teratas di Vietnam (Trung Do, 2024).

Selain itu, kerangka PROMETHEE juga telah diterapkan dalam konteks evaluasi kinerja pendidikan tinggi oleh Živković et al. (2017), antara lain untuk perumusan strategi universitas (Živković et al., 2017), serta analisis aktivitas transfer pengetahuan di lingkungan akademik oleh Ishizaka et al. (2020). Penerapan-penerapan tersebut menunjukkan fleksibilitas metode MCDM dalam berbagai konteks evaluasi institusi pendidikan tinggi (Ishizaka et al., 2020).

Meskipun demikian, terlepas dari keberagaman pendekatan yang digunakan, penelitian-penelitian tersebut masih memiliki kelemahan metodologis yang mendasar. Fokus analisis umumnya masih bertumpu pada indikator institusional.

yang bersifat umum, sementara tantangan spesifik yang melekat pada data scientometric, seperti keberadaan outlier dan efek kompensatori antar kriteria, belum ditangani secara eksplisit. Akibatnya, kontribusi utama penelitian-penelitian tersebut lebih banyak terbatas pada penerapan baru dari metode yang telah mapan, bukan pada pengembangan kerangka metodologis yang secara khusus dirancang untuk mengatasi karakteristik dan permasalahan data scientometric dalam pemeringkatan universitas.

Penggunaan indikator scientometric untuk mengukur dan memeringkat universitas dalam kerangka Multi-Criteria Decision-Making (MCDM) telah diusulkan oleh Maral (2024), yang mengombinasikan metode pembobotan objektif, yaitu Entropy, MEREC, dan CRITIC, dengan model pemeringkatan linier seperti TOPSIS, MABAC, dan ARAS, yang selanjutnya diagregasikan menggunakan metode Borda count (Maral, 2024). Pendekatan ini menunjukkan inovasi dalam mengintegrasikan indikator scientometric ke dalam kerangka MCDM dan memperluas pemanfaatan metode pengambilan keputusan multikriteria dalam konteks pemeringkatan universitas, namun, meskipun bersifat inovatif, pendekatan tersebut masih memiliki keterbatasan metodologis yang signifikan. Penggunaan pembobotan yang sepenuhnya bersifat objektif berisiko mengabaikan relevansi substantif dari masing-masing kriteria, sehingga kepentingan relatif antar indikator tidak selalu tercermin secara memadai. Kondisi ini dapat menyebabkan distorsi dalam proses evaluasi, di mana bobot kriteria lebih merepresentasikan karakteristik statistik data daripada signifikansi konseptual indikator yang dinilai (Chen et al., 2025; Ezell et al., 2021; Zoraghi et al., 2013).

Secara khusus, metode Entropy memiliki keterbatasan yang cukup menonjol dalam penentuan bobot kriteria, terutama karena tingkat sensitivitasnya yang tinggi terhadap keberadaan outlier dan kurangnya robustnes statistik. Formulasi Entropy yang berbasis probabilitas menyebabkan perubahan kecil pada data dapat menghasilkan pergeseran bobot yang besar, sehingga metode ini cenderung tidak stabil ketika diterapkan pada dataset yang tidak teratur, bersifat noisy, atau memiliki distribusi yang sangat skewed, seperti data scientometric (Chen et al., 2025; Erbey et al., 2025; Li & Zhang, 2023).

Metode MEREC juga menunjukkan keterbatasan ketika diterapkan pada data dengan variabilitas yang tinggi atau mengandung outlier ekstrem. Karena MEREC menghitung bobot berdasarkan deviasi absolut terhadap kondisi ideal, keberadaan nilai ekstrem pada satu kriteria dapat memberikan pengaruh yang tidak proporsional terhadap struktur pembobotan secara keseluruhan. Akibatnya, bobot yang dihasilkan dapat terdistorsi dan tidak lagi mencerminkan kontribusi relatif kriteria secara seimbang (Chen et al., 2025; Nguyen & Nhieu, 2025; Pala, 2024).

Demikian pula, metode CRITIC memiliki beberapa keterbatasan dalam konteks dataset dengan variabilitas tinggi atau yang mengandung outlier. CRITIC menetapkan bobot kriteria berdasarkan kombinasi varian dan korelasi antar kriteria, sehingga kriteria dengan tingkat dispersi yang tinggi cenderung memperoleh bobot yang lebih besar. Dalam kondisi data scientometric yang sangat heterogen, mekanisme ini menjadikan CRITIC sensitif terhadap nilai ekstrem, yang memungkinkan kriteria tertentu mendominasi proses pembobotan secara tidak proporsional (Alrababah & Gan, 2023; Chen et al., 2025; X. Li et al., 2022).

Keterbatasan metode pembobotan objektif tersebut menunjukkan perlunya pendekatan alternatif yang tidak hanya mempertimbangkan karakteristik statistik data, tetapi juga mampu menghasilkan bobot yang lebih stabil, konsisten, dan relevan secara substantif, khususnya dalam konteks pemeringkatan universitas berbasis indikator scientometric.

Lebih lanjut pendekatan agregasi yang bersifat kompensatori juga menghadapi permasalahan serius akibat distorsi yang dipicu oleh outlier, karena struktur agregasi linier aditif yang digunakan tidak mampu merepresentasikan perilaku preferensi yang bersifat non-kompensatori. Dalam konteks ini, penerapan metode ARAS pada dataset yang tidak berdistribusi normal atau terkontaminasi oleh outlier perlu dilakukan dengan sangat hati-hati. Struktur ARAS yang bersifat aditif dan *fully compensatory*, dikombinasikan dengan proses penilaian berbasis normalisasi, menjadikan metode ini sangat sensitif terhadap dispersi data dan heterogenitas skala antar kriteria. Akibatnya, nilai ekstrem (*outlier*) dapat meningkatkan utilitas agregat suatu alternatif secara tidak proporsional, sehingga menutupi kelemahan pada

kriteria lainnya dan pada akhirnya menghasilkan peringkat yang tidak stabil atau terdistorsi (Liu & Xu, 2021; Turskis & Keršulienė, 2024).

Metode TOPSIS juga memiliki struktur agregasi yang bersifat kompensatori, di mana kinerja yang sangat tinggi pada satu indikator dapat mengompensasi kelemahan yang signifikan pada indikator lain. Selain itu, TOPSIS mengasumsikan independensi penuh antar kriteria, sehingga tidak mampu menangkap adanya interaksi, ketergantungan, atau hubungan nonlinier dan asimetris antar indikator. Dalam kondisi dataset yang kompleks secara struktural, penetapan bobot kriteria saja tidak cukup untuk mengatasi keterbatasan tersebut, sehingga metode ini menjadi kurang robust ketika diterapkan pada data yang heterogen dan mengandung outlier (Baydaş et al., 2024; Jamwal et al., 2021; Rahman et al., 2024). Dalam konteks data scientometric yang bersifat heavy-tailed, mekanisme kompensatori pada TOPSIS berpotensi menyamarkan ketidakseimbangan kinerja antar indikator penelitian.

Sementara itu, metode MABAC juga tergolong sebagai metode *fully compensatory*. Struktur perhitungannya yang didasarkan pada deviasi aditif melalui matriks Q menjadikan MABAC sangat sensitif terhadap prosedur normalisasi dan perbedaan skala antar kriteria. Akibatnya, kriteria yang memiliki varians lebih besar atau rentang nilai yang lebih luas cenderung mendominasi hasil akhir pemeringkatan. Lebih lanjut, MABAC tidak menyediakan mekanisme eksplisit untuk mengendalikan pengaruh outlier maupun untuk menangani distribusi data yang sangat *skewed* (Ayan et al., 2023; Shi et al., 2022; Torkayesh et al., 2023). Kondisi ini menjadikan metode MABAC rentan terhadap distorsi ketika diterapkan pada data scientometric yang memiliki karakteristik *heavy-tailed* dan heterogen.

Secara keseluruhan, keterbatasan metode ARAS, TOPSIS, dan MABAC dalam menghadapi *outlier* dan efek kompensatori menunjukkan bahwa pendekatan agregasi linier yang sepenuhnya kompensatori kurang memadai untuk pemeringkatan universitas berbasis indikator scientometric. Hal ini menegaskan perlunya pengembangan kerangka pemeringkatan yang mampu mengendalikan efek kompensatori dan meredam pengaruh nilai ekstrem, sehingga hasil

pemeringkatan yang dihasilkan lebih stabil, adil, dan representatif terhadap kinerja multidimensional institusi.

Tantangan-tantangan tersebut semakin diperparah oleh keterbatasan teknik normalisasi data yang umum digunakan dalam evaluasi scientometric. Pendekatan normalisasi dalam domain scientometric hingga saat ini masih banyak mengandalkan metode konvensional, seperti min-max scaling, standardisasi z-score, dan transformasi logaritmik, yang juga lazim digunakan dalam berbagai implementasi MCDM. Namun demikian, metode-metode tersebut belum sepenuhnya mampu meredam pengaruh outlier secara efektif (Gagolewski et al., 2022; Lovakov & Teixeira da Silva, 2025).

Pada metode min-max scaling, keberadaan satu nilai ekstrem dapat memperlebar rentang data secara signifikan, sehingga nilai-nilai lainnya menjadi terkompresi secara tidak proporsional. Kondisi ini menyebabkan perbedaan relatif antar alternatif menjadi kurang terlihat dan berpotensi mendistorsi hasil evaluasi. Sementara itu, standardisasi z-score sangat sensitif terhadap nilai rata-rata dan simpangan baku, yang keduanya dapat dipengaruhi secara signifikan oleh outlier. Akibatnya, nilai ekstrem dapat menggeser pusat distribusi dan memperbesar varians, sehingga hasil normalisasi menjadi tidak stabil (Mahmoudi et al., 2020).

Adapun transformasi logaritmik memang mampu mengurangi skala data, namun tidak sepenuhnya menghilangkan dominasi nilai ekstrem, terutama pada distribusi data yang bersifat *heavy-tailed*. Dalam kondisi tersebut, nilai ekstrem tetap memiliki pengaruh yang besar terhadap hasil akhir meskipun telah dilakukan transformasi. Oleh karena itu, keterbatasan metode normalisasi konvensional ini menunjukkan bahwa penanganan *outlier* dalam data scientometric masih menjadi tantangan metodologis yang signifikan dan memerlukan pendekatan yang lebih adaptif dan robust (Gagolewski et al., 2022; Lovakov & Teixeira da Silva, 2025).

Untuk mengatasi berbagai keterbatasan metodologis tersebut, diperlukan suatu kerangka evaluasi yang lebih resilien yang mampu menangani tantangan mendasar berupa ketidakstabilan pembobotan, sensitivitas normalisasi terhadap outlier, serta sifat kompensatori dari model pemeringkatan tradisional. Penanganan tantangan pertama menuntut adanya mekanisme pembobotan yang tidak terpengaruh oleh

distorsi data atau distribusi *heavy-tailed*. Dalam konteks ini, Best–Worst Method (BWM) menawarkan tingkat efisiensi dan konsistensi yang lebih tinggi dalam penentuan bobot kriteria, karena hanya memerlukan perbandingan berpasangan antara kriteria terbaik dan terburuk dengan tetap mempertimbangkan tingkat kepentingan relatifnya (Rezaei, 2015). Berbeda dengan metode pembobotan objektif seperti Entropy, CRITIC, atau MEREC yang menurunkan bobot secara langsung dari dispersi data dan karenanya sangat sensitif terhadap skewness, varians, serta keberadaan outlier, BWM sepenuhnya independen terhadap struktur data, karena bobot dalam BWM diperoleh dari penilaian pakar dan bukan dari karakteristik statistik dataset, metode ini mampu menghindari distorsi yang disebabkan oleh nilai ekstrem, menjaga stabilitas bobot pada berbagai kondisi data, serta menghasilkan bobot yang lebih akurat dalam merepresentasikan tingkat kepentingan indikator (Goodarzi et al., 2022; Keenan, 2024; Raed et al., 2025). Oleh karena itu, BWM sangat sesuai untuk evaluasi scientometric, di mana distribusi *heavy-tailed* dan variabilitas yang dipicu oleh outlier sering kali melemahkan reliabilitas pendekatan pembobotan yang bergantung pada data.

Tantangan kedua berkaitan dengan proses normalisasi, yang harus mampu menstabilkan distribusi data scientometric yang bersifat skewed. Kondisi ini mendorong penggunaan Grey Relational Analysis (GRA). Metode GRA menyediakan mekanisme praproses data melalui transformasi Grey Relational Coefficient (GRC), yang berfungsi untuk mengurangi dominasi nilai-nilai ekstrem serta menstabilkan distribusi data yang mendasari proses evaluasi. Selain itu, GRA dapat diintegrasikan dengan berbagai metode MCDM, sehingga memberikan fleksibilitas dalam membangun kerangka evaluasi yang lebih efektif (Esangbedo et al., 2024; Zheng et al., 2025).

Sejumlah penelitian telah menunjukkan efektivitas GRA pada berbagai domain. Sebagai contoh Wu et al. (2025) mengembangkan contribution-driven weighted GRA (Wu et al., 2025), kemudian Zhu et al. (2024) menerapkan model coupling-GRA untuk penilaian keberlanjutan universitas (Zhu et al., 2024), dan Li (2024) menggunakan pendekatan GRA–TOPSIS dalam evaluasi pendidikan vokasi (Li, 2024). Studi komparatif antara Entropy-weighted GRA dan PROMETHEE juga

telah dilaporkan oleh Elevli dan Elevli (2024) menghasilkan komparasi ranking dan dianalisa kekuatan masing-masing metode (Elevli & Elevli, 2024). Pendekatan GRA–DEMATEL berbasis PROMETHEE untuk pemilihan energi terbarukan dilakukan oleh Li et al. (2022), menunjukkan potensi integrasi GRA, meskipun fokusnya masih terbatas pada dataset berbasis linguistic (P. Li et al., 2022).

Secara konseptual, Grey Relational Analysis (GRA) tidak hanya mampu menangani ketidakpastian informasi, tetapi juga efektif dalam meredam pengaruh nilai *outlier* (Başaran & Ighagbon, 2024). GRA menawarkan mekanisme normalisasi relasional yang sesuai untuk menangani data yang bersifat skewed atau terkontaminasi outlier dengan menghitung koefisien relasional grey sebagai ukuran kedekatan absolut terhadap referensi ideal (Başaran & Ighagbon, 2024; Zheng et al., 2025). Mekanisme ini memungkinkan perbandingan alternatif yang konsisten dan, karena berbasis pada perbedaan relatif, lebih tahan terhadap distorsi yang disebabkan oleh nilai ekstrem (Esangbedo & Wei, 2023; Paradowski et al., 2025).

Namun, meskipun pembobotan telah dibuat stabil dan proses normalisasi telah bersifat robust, tantangan ketiga masih tetap muncul, yaitu perilaku kompensatori dari model agregasi tradisional dan metode-metode sebelumnya, di mana kinerja yang sangat tinggi pada satu atau dua indikator dapat menutupi kelemahan pada indikator lainnya. Untuk mengatasi permasalahan ini, PROMETHEE II menawarkan mekanisme *outranking* yang bersifat non-kompensatori, di mana alternatif dibandingkan berdasarkan hubungan preferensi bukan agregasi numerik semata. Dalam kerangka non-kompensatori, kinerja yang sangat tinggi pada satu atau dua kriteria tidak dapat sepenuhnya mengompensasi kekurangan yang signifikan pada kriteria lain. Konsep *outranking* sendiri merujuk pada perbandingan berbasis preferensi yang menunjukkan bahwa suatu alternatif memiliki dukungan yang cukup untuk lebih dipilih dibandingkan alternatif lainnya (Deng et al., 2022; Gul et al., 2018). Dengan demikian, PROMETHEE II secara inheren mampu membatasi kompensasi yang tidak diinginkan antar kriteria.

Pembatasan efek kompensatori tersebut dicapai melalui penggunaan fungsi preferensi berpasangan (*pairwise preference functions*) serta penerapan threshold *indifference* dan *preference* secara eksplisit, yang memungkinkan perbedaan

antara perbedaan kinerja yang tidak signifikan dan perbedaan yang benar-benar bermakna secara preferensial (Papapostolou et al., 2024; Wulf et al., 2021). Selain itu, PROMETHEE II memiliki fleksibilitas metodologis yang tinggi dan dapat diintegrasikan dengan berbagai prosedur analitis lainnya, sehingga memungkinkan penerapannya dalam beragam kerangka MCDM hybrid (Ishizaka & Resce, 2021; Oubahman & Duleba, 2024; Pinochet et al., 2023; Pohl & Geldermann, 2024).

Berdasarkan sintesis terhadap literatur yang ada, belum terdapat pendekatan sebelumnya yang secara simultan mampu menangani tiga tantangan metodologis utama yang melekat dalam evaluasi scientometric, yaitu kebutuhan akan pembobotan yang konsisten dan tidak terdistorsi oleh data ekstrem atau distribusi yang skewed, kebutuhan akan mekanisme normalisasi yang robust terhadap skewness dan outlier, serta kebutuhan akan metode pemeringkatan non-kompensatori yang mampu mencegah dominasi satu atau dua indikator tertentu. Ketiadaan kerangka yang mampu memenuhi ketiga persyaratan tersebut secara bersamaan menunjukkan adanya kesenjangan penelitian (*research gap*) yang jelas dalam pengembangan metodologi pemeringkatan scientometric yang lebih robust dan memiliki landasan teoritis yang kuat.

Untuk menjembatani kesenjangan penelitian tersebut, penelitian ini mengusulkan kerangka hibrid BWM–GRA–PROMETHEE (BGP). Perancangan terintegrasi ini secara sengaja memanfaatkan kekuatan komplementer dari masing-masing metode, di mana BWM menyediakan mekanisme pembobotan yang ketat dan konsisten serta independen terhadap ketidakberaturan data, GRA berperan sebagai lapisan praproses yang robust dalam menstabilkan distribusi data scientometric yang bersifat heavy-tailed dan rentan terhadap outlier, dan PROMETHEE II menyediakan logika *outranking* non-kompensatori yang mencegah dominasi satu atau dua indikator sehingga tidak menutupi kelemahan pada indikator lainnya.

Melalui integrasi ketiga komponen tersebut, kerangka hibrid BGP secara bersama-sama mampu mengatasi permasalahan yang selama ini belum ditangani secara terpadu dalam literatur scientometric. Dengan demikian, kerangka ini

diharapkan dapat menghasilkan pemeringkatan universitas yang lebih robust, stabil, dan reliabel dibandingkan pendekatan-pendekatan yang ada.

2.2 Landasan Teori

Pada bagian ini disajikan landasan teori yang mendukung pengembangan kerangka pemeringkatan universitas berbasis indikator scientometric menggunakan pendekatan BWM–GRA–PROMETHEE (BGP). Pembahasan teori dimaksudkan untuk memberikan dasar konseptual dan metodologis yang kuat terkait evaluasi kinerja penelitian universitas, karakteristik data scientometric, serta relevansi penerapan metode pengambilan keputusan multikriteria (*Multi-Criteria Decision Making/MCDM*).

2.2.1 Perguruan Tinggi

Perguruan tinggi merupakan lembaga pendidikan pada jenjang pendidikan tinggi yang memiliki peran penting dalam pengembangan ilmu pengetahuan, teknologi, dan inovasi. Menurut Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tentang Pendidikan Tinggi, perguruan tinggi adalah satuan pendidikan yang menyelenggarakan pendidikan tinggi dan dapat berbentuk universitas, institut, sekolah tinggi, politeknik, akademi, atau akademi komunitas. Perguruan tinggi berfungsi sebagai pusat pengembangan ilmu pengetahuan dan berperan dalam menghasilkan sumber daya manusia yang berkualitas serta berkontribusi terhadap pembangunan nasional.

Dalam sistem pendidikan tinggi, perguruan tinggi memiliki tiga fungsi utama yang dikenal sebagai Tri Dharma Perguruan Tinggi, yaitu pendidikan dan pengajaran, penelitian, serta pengabdian kepada masyarakat. Ketiga fungsi ini saling berkaitan dan menjadi landasan utama dalam penyelenggaraan kegiatan akademik di perguruan tinggi. Melalui kegiatan pendidikan, perguruan tinggi menghasilkan lulusan yang memiliki kompetensi akademik dan profesional. Sementara itu, kegiatan penelitian berperan dalam menghasilkan pengetahuan baru serta inovasi yang dapat dimanfaatkan oleh masyarakat. Pengabdian kepada masyarakat menjadi sarana penerapan hasil penelitian dan keilmuan untuk memberikan manfaat langsung bagi masyarakat.

Dalam konteks evaluasi kinerja perguruan tinggi, berbagai indikator sering digunakan untuk menilai kualitas dan produktivitas institusi. Salah satu pendekatan yang banyak digunakan adalah evaluasi berbasis scientometric, yaitu penilaian kinerja penelitian menggunakan indikator bibliometrik seperti jumlah publikasi ilmiah, jumlah sitasi, kualitas jurnal, serta tingkat kolaborasi internasional. Pendekatan ini digunakan oleh berbagai lembaga pemeringkatan global seperti URAP, SCImago Institutions Rankings, dan Leiden Ranking, yang memanfaatkan data publikasi ilmiah untuk menilai performa riset perguruan tinggi.

Oleh karena itu, analisis kinerja perguruan tinggi berbasis indikator scientometric menjadi penting dalam memahami posisi relatif suatu institusi dalam ekosistem penelitian global. Evaluasi semacam ini tidak hanya membantu memetakan kekuatan dan kelemahan perguruan tinggi, tetapi juga memberikan dasar bagi pengambilan keputusan strategis dalam pengembangan riset dan peningkatan kualitas akademik.

2.2.2 Konsep Pemeringkatan Universitas dan Evaluasi Kinerja Penelitian

Pemeringkatan universitas merupakan suatu pendekatan evaluatif yang digunakan untuk menilai dan membandingkan kinerja institusi pendidikan tinggi berdasarkan seperangkat indikator tertentu (Freyer, 2014; Lukić & Tumbas, 2019). Dalam konteks global pendidikan tinggi, pemeringkatan tidak hanya berfungsi sebagai alat perbandingan reputasi akademik, tetapi juga menjadi dasar dalam pengambilan keputusan kebijakan, alokasi pendanaan riset, serta peningkatan daya saing institusi. Salah satu dimensi utama dalam pemeringkatan universitas adalah kinerja penelitian, karena mencerminkan kemampuan institusi dalam menghasilkan pengetahuan ilmiah dan memberikan dampak akademik (Keenan, 2024; Maral, 2024).

Evaluasi kinerja penelitian umumnya dilakukan menggunakan indikator kuantitatif, seperti jumlah publikasi, jumlah sitasi, dan metrik kualitas jurnal, yang merepresentasikan produktivitas dan dampak ilmiah suatu universitas (Gagolewski et al., 2022; Lovakov & Teixeira da Silva, 2025). Namun, kinerja penelitian bersifat multidimensional, sehingga tidak dapat direpresentasikan secara memadai melalui

satu indikator tunggal atau agregasi linier sederhana (Daraio et al., 2023). Pendekatan agregatif semacam ini berpotensi menyederhanakan kompleksitas kinerja penelitian dan menimbulkan efek kompensatori, di mana keunggulan pada satu indikator dapat menutupi kelemahan pada indikator lainnya (Erbey et al., 2025).

Pemeringkatan universitas berbasis evaluasi kinerja penelitian memerlukan kerangka evaluasi yang lebih sistematis dan metodologis, yang mampu merepresentasikan kontribusi berbagai indikator secara proporsional serta menghasilkan pemeringkatan yang lebih, transparan, robust, stabil, dan reliabel.

2.2.3 Scientometric

Scientometric merupakan bidang kajian yang mempelajari pengukuran dan analisis aktivitas penelitian ilmiah melalui pendekatan kuantitatif. Scientometric digunakan untuk mengevaluasi produktivitas, dampak, dan pola perkembangan ilmu pengetahuan berdasarkan data publikasi, sitasi dan metrik lainnya (Gyorffy et al., 2023; Lopez-Rodriguez & Ceballos, 2022). Dalam konteks pemeringkatan universitas, scientometric berperan sebagai dasar objektif untuk menilai kinerja penelitian institusi pendidikan tinggi secara terukur dan dapat dibandingkan (Keenan, 2024).

Indikator scientometric yang umum digunakan meliputi jumlah publikasi, jumlah sitasi, kualitas jurnal, serta berbagai metrik turunan yang merepresentasikan produktivitas dan dampak ilmiah. Indikator-indikator tersebut memungkinkan evaluasi kinerja penelitian dilakukan secara sistematis dan berbasis data. Namun demikian, data scientometric memiliki karakteristik khusus, yaitu bersifat multidimensional, heterogen, dan sering kali berdistribusi tidak simetris (*skewed*) serta mengandung nilai ekstrem (*outlier*) (Lovakov & Teixeira da Silva, 2025; Loyola-Gonzalez et al., 2020; Schlögl et al., 2025).

Karakteristik tersebut menyebabkan analisis dan pemeringkatan berbasis scientometric rentan terhadap distorsi apabila tidak didukung oleh pendekatan metodologis yang tepat. Penggunaan indikator scientometric dalam evaluasi kinerja penelitian memerlukan kerangka analisis yang mampu menangani kompleksitas data, mengendalikan pengaruh outlier, serta merepresentasikan kontribusi relatif

setiap indikator secara proporsional dan konsisten (Bornmann, 2024; Schmoch, 2020).

2.2.4 MCDM sebagai Kerangka Evaluasi Multidimensi

Multi-Criteria Decision Making (MCDM) merupakan kerangka pengambilan keputusan yang digunakan untuk mengevaluasi dan memilih alternatif berdasarkan sejumlah kriteria yang saling berbeda dan sering kali memiliki tingkat kepentingan yang tidak sama (Ayyildiz et al., 2023; Zhang et al., 2022). MCDM dirancang untuk menangani permasalahan keputusan yang bersifat kompleks dan multidimensional, di mana satu kriteria saja tidak cukup untuk merepresentasikan kinerja atau kualitas suatu alternatif secara menyeluruh (Andjelković et al., 2024; Makki & Abdulaal, 2023).

Dalam konteks pemeringkatan universitas berbasis indikator scientometric, MCDM memungkinkan integrasi berbagai indikator kinerja penelitian ke dalam satu kerangka evaluasi yang sistematis dan eksplisit (Esangbedo et al., 2024; Zhang et al., 2022; Zoraghi et al., 2013). Melalui mekanisme pembobotan, MCDM dapat merepresentasikan perbedaan tingkat kepentingan antar indikator, sehingga evaluasi tidak lagi bergantung pada agregasi linier sederhana yang berpotensi menimbulkan bias.

Keunggulan utama MCDM terletak pada kemampuannya untuk mengakomodasi keragaman indikator serta mengendalikan interaksi antar kriteria dalam proses evaluasi. Dengan demikian, MCDM dipandang lebih sesuai untuk menilai kinerja penelitian universitas yang bersifat multidimensional, kompleks, dan heterogen, sekaligus menyediakan landasan metodologis yang lebih transparan dan dapat dipertanggungjawabkan dibandingkan pendekatan agregasi statistik konvensional (Erbey et al., 2025; Mukhametzyanov & Pamucar, 2018).

2.2.5 Delphi Method

Delphi Method merupakan suatu teknik pengambilan keputusan dan pengumpulan pendapat pakar yang dilakukan secara sistematis, terstruktur, dan iteratif dengan tujuan untuk memperoleh konsensus terhadap suatu permasalahan tertentu (Shang, 2023). Metode ini dikembangkan untuk mengintegrasikan

penilaian para ahli tanpa adanya interaksi langsung antar peserta, sehingga dapat mengurangi pengaruh dominasi individu dan bias sosial dalam proses pengambilan keputusan (Hasson et al., 2025; Olivero et al., 2022).

Dalam penerapannya, Delphi Method dilakukan di mana para pakar diminta untuk memberikan penilaian atau pendapat terhadap isu atau kriteria tertentu. Hasil dari setiap pakar dirangkum hingga mencapai tingkat konsensus yang diinginkan tercapai. Proses ini memungkinkan penyempurnaan penilaian secara bertahap dan terkontrol (Hesselink et al., 2024).

Dalam konteks pemeringkatan universitas berbasis indikator scientometric, Delphi Method digunakan untuk memvalidasi pemilihan dan relevansi kriteria yang digunakan dalam evaluasi kinerja penelitian. Dengan melibatkan pakar metode ini membantu memastikan bahwa indikator yang digunakan tidak hanya relevan secara statistik, tetapi juga memiliki dasar konseptual dan substantif yang kuat. Delphi Method berperan sebagai landasan awal yang penting dalam membangun kerangka evaluasi yang kredibel dan ilmiah.

2.2.6 Validasi Uji Sensitivitas

Uji sensitivitas merupakan prosedur validasi yang digunakan untuk menilai stabilitas dan robustnes hasil pemeringkatan terhadap perubahan parameter model. Dalam konteks *Multi-Criteria Decision Making* (MCDM), uji sensitivitas berfungsi untuk mengevaluasi sejauh mana hasil keputusan khususnya peringkat alternatif dipengaruhi oleh variasi bobot kriteria yang digunakan dalam proses evaluasi. Model pemeringkatan yang baik diharapkan menghasilkan peringkat yang relatif stabil meskipun terjadi perubahan bobot dalam rentang yang wajar (Doan & De Smet, 2018; Mukhametzyanov & Pamucar, 2018).

Pada penelitian ini, uji sensitivitas dilakukan melalui variasi bobot kriteria, dengan tujuan untuk menguji ketahanan kerangka BWM–GRA–PROMETHEE (BGP) terhadap perubahan preferensi kepentingan kriteria. Variasi bobot mencerminkan kemungkinan perbedaan penilaian pakar atau perubahan kebijakan evaluasi, yang secara realistis dapat terjadi dalam praktik pemeringkatan universitas. Dengan demikian, uji sensitivitas berbasis variasi bobot memungkinkan evaluasi apakah hasil pemeringkatan bersifat konsisten secara

struktural atau sangat bergantung pada konfigurasi bobot tertentu (Kabassi & Martinis, 2021; Mukhametzyanov & Pamucar, 2018).

Dalam kerangka MCDM, uji sensitivitas bobot umumnya dilakukan dengan menaikkan atau menurunkan bobot kriteria tertentu secara sistematis, kemudian mengamati perubahan posisi peringkat alternatif. Jika perubahan bobot menghasilkan pergeseran peringkat yang kecil atau terbatas pada alternatif tertentu, maka model dapat dikatakan robust. Sebaliknya, jika variasi bobot yang relatif kecil menyebabkan perubahan peringkat yang signifikan, maka hasil pemeringkatan dianggap kurang stabil dan rentan terhadap bias parameter (Demir et al., 2024; Kabassi & Martinis, 2021; Liu & Liu, 2024).

Oleh karena itu, validasi uji sensitivitas melalui variasi bobot menjadi langkah penting untuk memastikan bahwa hasil pemeringkatan yang dihasilkan oleh kerangka BGP tidak bersifat arbitrer, melainkan mencerminkan evaluasi kinerja yang konsisten dan dapat dipertanggungjawabkan secara metodologis. Uji ini juga memperkuat keandalan hasil penelitian dengan menunjukkan bahwa peringkat universitas yang dihasilkan tetap stabil di bawah skenario perubahan bobot kriteria yang masuk akal (Mahmoudi et al., 2020; Mukhametzyanov & Pamucar, 2018; Więckowski & Sałabun, 2023).

2.2.7 Uji Komparasi Antar Metode

Uji komparasi antar metode merupakan pendekatan evaluatif yang digunakan untuk membandingkan kinerja suatu metode yang diusulkan dengan metode-metode lain yang telah ada. Dalam konteks *Multi-Criteria Decision Making* (MCDM), uji komparasi bertujuan untuk menilai keunggulan relatif, konsistensi hasil, serta stabilitas pemeringkatan yang dihasilkan oleh suatu model dibandingkan dengan pendekatan alternatif. Analisis ini penting untuk menunjukkan bahwa metode yang diusulkan tidak hanya berbeda secara struktural, tetapi juga memberikan nilai tambah metodologis yang nyata (Shi et al., 2022; Wu et al., 2022).

Dalam penelitian pemeringkatan universitas berbasis indikator scientometric, uji komparasi antar metode umumnya dilakukan dengan membandingkan hasil pemeringkatan dari beberapa metode MCDM yang berbeda, seperti metode

agregatif atau kompensatori, terhadap metode yang diusulkan. Perbandingan tersebut dapat dilakukan dengan mengamati perbedaan posisi peringkat, tingkat perubahan peringkat antar alternatif, serta pola dominasi indikator tertentu. Melalui pendekatan ini, dapat dievaluasi sejauh mana suatu metode mampu merepresentasikan kinerja institusi secara seimbang dan tidak terdistorsi oleh karakteristik data tertentu (Baydaş et al., 2024; Coban et al., 2023).

Uji komparasi juga berperan penting dalam menilai robustnes metodologis, khususnya ketika data yang digunakan memiliki sifat kompleks, seperti distribusi yang *skewed* dan keberadaan *outlier*, sebagaimana umum dijumpai pada data scientometric. Metode yang menghasilkan pemeringkatan yang sangat berbeda atau tidak stabil dibandingkan metode lain dapat mengindikasikan adanya sensitivitas berlebih terhadap agregasi, pembobotan, atau efek kompensatori. Sebaliknya, metode yang menunjukkan hasil pemeringkatan yang lebih konsisten dan rasional dapat dianggap lebih andal (Ahmed et al., 2024; Cinelli et al., 2020; Wu et al., 2022, 2020).

Dalam penelitian ini, uji komparasi antar metode digunakan untuk mengevaluasi kinerja kerangka BWM–GRA–PROMETHEE (BGP) dibandingkan dengan metode pemeringkatan lain yang relevan. Analisis komparatif ini bertujuan untuk menunjukkan bahwa integrasi pembobotan yang konsisten, transformasi data yang robust, dan mekanisme pemeringkatan non-kompensatori dalam kerangka BGP mampu menghasilkan pemeringkatan yang lebih stabil dan representatif. Dengan demikian, uji komparasi antar metode menjadi bagian penting dalam pembuktian kontribusi ilmiah dan kebaruan metodologis dari penelitian ini.