

ANALISIS PENGARUH FAKTOR-FAKTOR PENGUKURAN TERHADAP EN SCORE DAN ZETA SCORE PADA UJI BANDING LABORATORIUM DENGAN PENDEKATAN REGRESI LINEAR BERGANDA

Arum Setianingsih¹, Elin Herlinawati²

^{1,2} Program Studi Matematika, Universitas Terbuka, Tangerang Selatan, Indonesia

e-mail: 049468432@ecampus.ut.ac.id

ABSTRAK

Uji banding antar laboratorium merupakan mekanisme penting untuk memastikan kesetaraan atau kesesuaian kinerja antar laboratorium sesuai ISO 17025:2017. Evaluasi kesesuaian dan konsistensi hasil menggunakan metode *En score* dan *Zeta score* sesuai rekomendasi ISO 13528:2015. Artikel ini bertujuan untuk mengevaluasi kinerja antar laboratorium dan menganalisis faktor-faktor yang mempengaruhi variasi *En score* dan *Zeta score* pada uji banding bejana ukur dengan metode gravimetri, khususnya dengan memasukkan variabel kontrol yang berasal dari parameter pengukuran dan komponen ketidakpastian yang dilaporkan dengan pendekatan regresi linear berganda. Hasil pembahasan pada artikel menunjukkan bahwa 31 peserta dinyatakan setara dan konsisten dalam pelaporan unjuk kinerjanya, namun terdapat satu laboratorium yang *outlier* yaitu Lab 20 dengan *En score* (−3.02) dan *Zeta score* (−6.028), yang mengindikasikan adanya bias signifikan. Pemodelan regresi menunjukkan R^2 sebesar 0.59 dapat diartikan 59% variabel kontrol dapat dijelaskan oleh variabel kontrol dalam artikel ini. Hasil yang diperoleh untuk variabel resolusi timbangan memiliki pengaruh paling signifikan terhadap *En score* dan *Zeta score*, dengan koefisien masing-masing −7.261 ($p < 0.01$) dan −14.492 ($p < 0.01$). Temuan ini menegaskan bahwa penggunaan timbangan dengan resolusi rendah (*low precision*) secara langsung meningkatkan deviasi pengukuran dan menurunkan konsistensi hasil. Selain itu, komponen ketidakpastian dari anak timbangan (u_{timb}) juga signifikan dan berkontribusi negatif pada kedua model dengan koefisien regresi −1.178 untuk *En score* dan −2.355 untuk *Zeta score* (keduanya signifikan pada $p < 0.01$) yang berarti bahwa pelaporan ketidakpastian yang akurat dan objektif membantu mengurangi risiko *outlier* akibat pelaporan ketidakpastian yang *underestimation* atau *overestimation*.

Kata Kunci : Uji banding, Bejana Ukur, *En score* *Zeta score*, Regresi Linear Berganda

1. PENDAHULUAN

Kegiatan uji banding antar laboratorium atau dikenal juga dengan interkomparasi merupakan suatu kegiatan yang dimaksudkan untuk evaluasi kinerja laboratorium dalam pengujian atau kalibrasi sekaligus pemantauan kinerja laboratorium secara berkesinambungan. Kegiatan ini dimaksudkan untuk menilai kesesuaian serta kemampuan kinerja laboratorium dalam melaksanakan pengujian kalibrasi, sehingga kredibilitas pengukuran yang dilakukan dapat dipastikan dan terjamin. Interkomparasi atau uji banding antar laboratorium merupakan salah satu mekanisme penting dalam menjamin keabsahan hasil pengujian dan kalibrasi. Pada dokumen ISO/IEC 17025:2017 merekomendasikan laboratorium untuk mengevaluasi dan memantau unjuk kinerjanya melalui perbandingan hasil dengan laboratorium lain. Partisipasi laboratorium dalam kegiatan uji banding dimaksudkan sebagai sarana evaluasi terhadap kinerja pengujian, membantu

mendeteksi adanya bias dan ketidaksesuaian pengujian, serta mendorong peningkatan kompetensi teknis laboratorium Durgut, (2021). Studi menunjukkan bahwa interkomparasi memiliki peran strategis dalam menjaga standar laboratorium. Hal ini selaras dengan penelitian dari Oktari *et al.* (2023) bahwa pelaksanaan uji banding yang dilakukan oleh laboratorium Instalasi Bahan Bakar Nuklir, meskipun partisipasi hanya 43% dari cakupan yang ditargetkan, pelaksanaan uji banding dapat meningkatkan kualitas. Temuan ini menunjukkan bahwa kegiatan uji banding merupakan mekanisme pemenuhan persyaratan akreditasi dan secara sistematis meningkatkan kualitas layanan. Hal ini mendukung gagasan bahwa uji banding berperan dalam pengontrolan kualitas layanan yang diberikan oleh laboratorium.

Balai Standardisasi Metrologi Legal Regional II (BSML Regional II) merupakan unit pelaksana teknis pemerintah yang menaungi 154 Unit Metrologi Legal (UML) Kabupaten/Kota di wilayah Jawa, Bali, dan Nusa Tenggara. Berdasarkan Peraturan Menteri Perdagangan Nomor 28 Tahun 2025 tentang Organisasi dan Tata Kerja Unit Pelaksana Teknis di Lingkungan Kementerian Perdagangan, BSML Regional II memiliki fungsi untuk melaksanakan penelaahan standar ukuran metrologi legal dan pengelolaan standar ukuran miliknya dan milik UML Kabupaten/Kota. Pengelolaan standar yang dimaksud salah satunya yakni melakukan penjaminan kesesuaian hasil pengukuran berupa penyelenggaraan interkomparasi yang diikuti oleh UML Kabupaten/Kota. Tujuan dari interkomparasi ini adalah untuk mengukur kemampuan Unit Metrologi Legal (UML) Kabupaten/Kota dalam melakukan pengelolaan standar ukuran metrologi legal untuk membangun saling pengakuan yang didasarkan pada keseragaman hasil pengukuran. Hal ini sejalan dengan penerapan ISO 17025:2017 di BSML Regional II yang menuntut laboratorium untuk secara berkelanjutan membuktikan kompetensi teknis dan keandalan hasilnya melalui pengendalian mutu eksternal.

Sebagaimana diatur dalam ISO/IEC 17043, kegiatan interkomparasi dilakukan untuk memastikan tidak hanya kesesuaian hasil pengukuran antar-laboratorium tetapi juga kompetensi teknis dan ketelusuran hasil pengukuran. Standar ini menetapkan prosedur umum yang harus diikuti oleh penyelenggara uji profisiensi (PT). Prosedur ini mencakup perencanaan, pelaksanaan, dan evaluasi hasil antar-laboratorium. Studi oleh Cordeiro *et al.*, (2022) meneliti apakah *z score* cukup untuk menilai kinerja peserta untuk menemukan adanya koreksi tersembunyi dalam kegiatan uji banding laboratorium menggunakan grafis intuitif yang disebut Naji plot yaitu dengan mengkombinasikan *zcore* dan *zeta score* untuk mengidentifikasi adanya ketidaksesuaian pelaporan ketidakpastian pengukuran mengacu pada ISO/IEC 17043:2010 dan ISO 13528:2015. Aydemir & Fank (2012).

Dalam kegiatan interkomparasi, evaluasi kecakapan laboratorium dilakukan sesuai dengan pedoman ISO 13528:2015, yang menetapkan berbagai metode penilaian, termasuk *En score*, *Z-score*, *Z'-score*, dan *Zeta score*. Standar ISO 13528:2015 ini secara khusus merekomendasikan penggunaan *En score* dan *Zeta score* untuk skema interkomparasi yang mempertimbangkan ketidakpastian pengukuran, baik dari laboratorium referensi maupun laboratorium peserta. Pendekatan tersebut memungkinkan penilaian kinerja yang tidak hanya membandingkan hasil laboratorium dengan nilai acuan, tetapi juga mengintegrasikan ketidakpastian pengukuran yang dilaporkan, sehingga evaluasi menjadi lebih akurat, realistis, dan objektif bagi seluruh peserta. Pelaporan ketidakpastian peserta menjadi hal yang perlu dianalisis, hal ini disebabkan ketidakpastian mempengaruhi unjuk kerja dari laboratorium peserta.

Pada skema evaluasi kegiatan uji banding antar laboratorium menggunakan *En score* untuk menilai kesetaraan hasil dengan nilai acuan, hal ini sangat dipengaruhi oleh besarnya ketidakpastian yang dilaporkan: semakin besar nilai ketidakpastian, maka *En score* cenderung

kecil sehingga hasil dari pelaporan tampak “setara”. Namun, jika semakin kecil laporan nilai ketidakpastian, maka *En score* bisa meningkat sehingga hasil pelaporan terlihat “tidak setara” walaupun selisihnya terlihat wajar. Untuk memberikan gambaran yang lebih objektif, terutama pada skema interkomparasi dengan peserta terbatas, maka digunakan evaluasi Zeta score untuk menilai konsistensi statistik laboratorium. Penelitian Skrzypczak et al., (2021) menggunakan kombinasi zscore, zeta score dan En Score untuk mengevaluasi kesesuaian maupun konsistensi pelaporan ketidakpastian dari peserta uji banding dengan material uji banding berupa beton siap pakai. Berdasarkan hasil penelitian zscore diterapkan untuk mengevaluasi kemampuan laboratorium, sedangkan zeta score dan En score untuk meningkatkan kinerja laboratorium dalam pengambilan tindakan korektif untuk menentukan penyebab tidak konsistensi campuran beton yang digunakan, Metode ini sejalan dengan tujuan utama yang ingin dibahas pada artikel ini untuk mengevaluasi dan mengambil tindakan korektif pada kegiatan uji banding laboratorium.

Arrhenius et al. (2024) menggunakan *Zeta score* sebagai alat evaluasi performa laboratorium, namun belum menilai hubungan antara komponen ketidakpastian yang dilaporkan oleh peserta dengan hasil evaluasi tersebut. Dalam skema evaluasi interkomparasi *Zeta score* terbukti efektif untuk mengevaluasi kinerja laboratorium karena secara eksplisit mempertimbangkan ketidakpastian pengukuran yang dilaporkan peserta, sehingga potensi bias yang tersembunyi dapat diidentifikasi—hal yang mungkin tidak terdeteksi dengan *Zscore* saja Cordeiro et al., (2022). Pendekatan ini sejalan dengan artikel yang penulis bahas, dimana perlunya penilaian terhadap konsistensi statistik peserta dan deteksi laboratorium yang melaporkan ketidakpastian berlebihan atau kurang.

Sementara itu, Li et al. (2019) meneliti penyebab kegagalan *proficiency testing* di laboratorium kimia secara deskriptif berdasarkan kategori kesalahan teknis, alat, dan metodologi. Walaupun penelitian tersebut memberikan gambaran umum mengenai sumber ketidaksesuaian, namun pendekatannya belum mengkuantifikasi hubungan antara ketidakpastian pengukuran dan performa laboratorium secara statistik sebagaimana yang disarankan dalam ISO 13528 dan ISO/IEC 17043. Oleh karena itu, artikel ini mengusulkan pendekatan kuantitatif untuk menganalisis faktor penentu *outlier* dalam kegiatan interkomparasi bejana ukur berdasarkan ketidakpastian pengukuran menggunakan pendekatan regresi linear berganda.

Pendekatan regresi linear berganda digunakan untuk menganalisis faktor-faktor yang dapat mempengaruhi evaluasi *En score* dan *Zeta score*, sekaligus menganalisa faktor dominan dan besar kontribusinya terhadap kemungkinan munculnya hasil *outlier* dengan berbantuan aplikasi STATA versi 17 yang mendukung penghitungan regresi linear berganda. Selain itu, STATA menyediakan fungsi diagnostik seperti *Variance Inflation Factor* (VIF) untuk mendeteksi multikolinearitas, *standardized beta coefficients* untuk membandingkan kontribusi antar variabel, serta *marginal effects* untuk menginterpretasikan pengaruh probabilitas komponen ketidakpastian terhadap hasil pengujian.

Artikel ini memberikan kontribusi keilmuan yang lebih mendalam mengenai faktor-faktor kritis dalam kegiatan uji banding laboratorium, serta menyoroti aspek-aspek yang perlu diperhatikan untuk meningkatkan kualitas dan konsistensi pengukuran. Dengan demikian, artikel ini tidak hanya menilai kesesuaian hasil antar-laboratorium, tetapi juga memberikan model analitik berbasis data untuk menjelaskan bagaimana masing-masing sumber ketidakpastian berkontribusi terhadap hasil evaluasi performa laboratorium. Pembahasan pada artikel ini diharapkan dapat mampu memberikan rekomendasi pada peningkatan kesesuaian dan konsistensi hasil uji banding antar-laboratorium di masa mendatang.

2. METODE

2.1 Data

Data kegiatan uji banding pada artikel ini merupakan studi kegiatan uji banding artefak bejana ukur standar nominal 10L yang diselenggarakan oleh Balai Standardisasi Metrologi Legal Regional II (BSML Regional II) yang diikuti oleh 32 peserta. Artefak terdiri dari dua unit yaitu Artefak A dan Artefak B yang masing-masing artefak diuji oleh 16 laboratorium peserta. Adapun pelaksanaan uji banding dilakukan dalam lima siklus uji banding, dimana tiap siklus diikuti oleh 4 sampai dengan 5 laboratorium peserta dengan mekanisme tiap siklus yaitu dimulai dengan pengujian dari laboratorium penyelenggara untuk mengambil nilai acuan awal, kemudian didistribusikan ke peserta1, peserta2, peserta3, peserta4, dan kembali lagi ke penyelenggara sebagai akhir dari satu siklus untuk pengambilan nilai acuan akhir. Pengujian artefak dilaksanakan dengan metode gravimetri pengujian basah menggunakan timbangan elektronik pada suhu referensi 28°C, dengan nilai ketidakpastian hasil pengukuran dilaporkan pada tingkat kepercayaan 95%(k=2). Tabel 1 mencantumkan definisi operasional, pengukuran variabel, dan statistika deskriptif yang digunakan dalam artikel ini, sedangkan Tabel 2 mencantumkan statistika deskriptif yang digunakan dalam artikel ini.

Tabel 1. Definisi Variabel Operasional

Variabel	Jenis Variabel	Definisi Operasional	Skala Pengukuran (Satuan)
En-score	Dependen	Indeks evaluasi kesesuaian hasil kegiatan uji banding	Rasio (tanpa satuan)
Zeta-score	Dependen	Indeks evaluasi konsistensi pelaporan hasil uji banding	Rasio (tanpa satuan)
Xref	Independen	Nilai acuan pengujian bejana ukur standar dari laboratorium penyelenggara	Rasio (mL)
u_ref	Independen	Ketidakpastian gabungan yang diperluas dari laboratorium penyelenggara	Rasio (mL)
Xlab	Independen	Nilai hasil pengujian bejana ukur standar dari laboratorium peserta	Rasio (mL)
u_lab	Independen	Ketidakpastian gabungan yang diperluas dari laboratorium peserta	Rasio (mL)
u_repeat	Independen	Komponen ketidakpastian akibat pengulangan pengukuran	Rasio (mL)
u_timb	Independen	Komponen ketidakpastian akibat anak timbangan standar	Rasio (mL)
u_meniskus	Independen	Komponen ketidakpastian akibat pembacaan meniskus	Rasio (mL)
resolusi	Independen	Resolusi alat ukur yang digunakan oleh laboratorium peserta	Rasio (mL)
suhu_air	Independen	Suhu air saat pengujian bejana ukur standar	Rasio (°C)
rho_air	Independen	Massa jenis air saat pengujian	Rasio (g/mL)

Tabel 2. Statistika Deskriptif

Variabel	Definisi	Obs	Mean	Std. dev.	Min	Max
Kode_Artefak	kode artefak bejana ukur standar	32	16.5	9.380	1.000	32.000
Kode_lab	kode lab peserta	32	16.5	9.381	1.000	32.000
Xref	nilai acuan laboratorium penyelenggara	32	10000.908	1.596	9999.020	10002.710
u_ref	ketidakpastian gabungan yang diperluas laboratorium penyelenggara	32	0.77	0.186	0.610	1.300
uc_ref	ketidakpastian gabungan laboratorium penyelenggara	32	0.385	0.093	0.305	0.650
Xlab	nilai hasil pengujian laboratorium peserta	32	10000.541	1,552	9996.750	10003.740
u_lab	ketidakpastian gabungan yang diperluas laboratorium peserta	32	0.866	0.290	0.580	1.600
uc_lab	ketidakpastian gabungan laboratorium peserta	32	0.437	0.148	0.295	0.799
u_repeat	ketidakpastian perulangan laboratorium peserta	32	0.181	0.124	0.000	0.517
u_timb	ketidakpastian anak timbangan standar laboratorium peserta	32	0.179	0.181	0.003	0.686
u_meniskus	ketidakpastian meniskus laboratorium peserta	32	0.303	0.075	0.167	0.577
resolusi	resolusi timbangan yang dipakai oleh peserta	32	0.11	0.073	0.010	0.500
suhu_air	suhu air pengujian bejana standar	32	22.032	1.853	18.110	25.730
rho_air	massa jenis pengujian bejana standar peserta	32	0.998	0.000	0.997	0.999
en	evaluasi hasil uji banding	32	-0.288	0.727	-3.020	0.930
zeta	evaluasi hasil uji banding	32	-0.573	1.451	-6.028	1,853

Tujuan utama analisis artikel ini adalah mengevaluasi kesesuaian hasil pengukuran menggunakan *En score* dan konsistensi peserta menggunakan *Zeta score*, serta memeriksa faktor-faktor teknis yang dapat mempengaruhi deviasi pengukuran dan outlier dari kegiatan uji banding. Meskipun *En score* dan *Zeta score* dapat diklasifikasikan ke dalam kategori biner, penelitian ini tidak bertujuan untuk melakukan klasifikasi hasil, melainkan untuk mengkaji variasi numerik kedua indikator tersebut sebagai representasi kinerja laboratorium. Oleh karena itu, pendekatan analisis yang digunakan adalah regresi linear berganda dengan variabel dependen berupa skor kontinu. Pendekatan ini dipilih agar informasi kuantitatif dalam data tetap terjaga dan hasil estimasi yang diperoleh lebih stabil serta bermakna secara statistik.

Untuk langkah evaluasi tersebut diperlukan beberapa variabel kontrol berupa seperangkat kovariat yang berpengaruh pada nilai hasil pengujian dan ketidakpastian yang dilaporkan oleh peserta. Variabel kontrol yang digunakan dalam analisis nilai hasil pengujian bejana ukur metode gravimetri yaitu hasil penimbangan dalam hal ini diwakili oleh resolusi timbangan, suhu air, massa jenis air, dan massa jenis udara. Dikarenakan massa jenis udara diasumsikan nilainya 0.0012 g/m³, maka massa jenis udara tidak digunakan sebagai variabel kontrol. Variabel kontrol yang digunakan

dalam analisis ketidakpastian hasil pengujian bejana ukur metode gravimetri yaitu komponen ketidakpastian perulangan pengukuran, ketidakpastian anak timbangan yang digunakan sebagai alat standar pengujian timbangan, dan ketidakpastian pembacaan meniskus bejana ukur standar.

Pada artikel ini pemodelan regresi dilakukan dua pemodelan regresi linear bergada yakni terhadap En score dan Zeta score. Meskipun kedua model menggunakan variabel independen dan variabel kontrol yang sama, En score dan Zeta score merupakan indikator yang dihitung dengan formula berbeda serta merepresentasikan dimensi evaluasi yang tidak identik, yaitu kesesuaian hasil pengukuran dan konsistensi pelaporan ketidakpastian. Perbedaan ini menyebabkan karakteristik distribusi dan struktur galat kedua skor tidak dapat diasumsikan sama, sehingga penggabungan keduanya dalam satu model regresi berpotensi menghasilkan interpretasi yang bias. Oleh karena itu, analisis dilakukan menggunakan dua model regresi linear terpisah agar hubungan antara variabel prediktor dan masing-masing indikator dapat diinterpretasikan secara lebih tepat dan bermakna secara statistik.

2.2 Spesifikasi Model

2.2.1 Penentuan Nilai Acuan Uji Banding

Dalam mengevaluasi uji banding laboratorium, mekanisme penentuan nilai acuan uji banding (x_{reff}) dan ketidakpastiannya (u_{reff}) menjadi tanggungjawab penyelenggara, (International Organization for Standardization, 2010). Penentuan nilai acuan (x_{reff}) dilakukan dengan mengambil nilai rata-rata dari nilai awal (x_{awal}) dan nilai akhir (x_{akhir}) pada satu siklus pengujian dan ketidakpastian acuan seperti dengan menggunakan persamaan matematis sebagai berikut:

$$x_{reff} = \frac{x_{awal} + x_{akhir}}{2} \quad (1)$$

$$u_{reff} = 2\sqrt{u_{awal}^2 + u_{stab}^2} \quad (2)$$

Terkait dengan pelaksanaan kegiatan uji banding, penyelenggara bertanggungjawab melakukan pemantauan stabilitas terhadap artefak yang digunakan dalam kegiatan uji banding, yang secara matematis dituliskan:

$$u_{stab} = \frac{drift}{\sqrt{3}}, drift = |x_{awal} - x_{akhir}| \quad (3)$$

Uji stabilitas (*drift*) dimaksudkan untuk memantau artefak memiliki karakteristik metrologis yang tidak berubah selama periode pengujian. Selisih antara nilai sebelum dan sesudah artefak didistribusikan dihitung sebagai *drift*, $\Delta x = |x_{akhir} - x_{awal}|$, dimana artefak dinyatakan stabil apabila perubahan tersebut tidak melebihi dua kali ketidakpastian pengukuran awal ($\Delta x \leq 2u_{awal}$) dan ketidakpastian stabilitas $u_{stab} < \frac{1}{3}U_{reff}$.

2.2.2 Evaluasi En Score dan Evaluasi Zeta Score

Metode evaluasi pelaporan hasil uji banding laboratorium peserta terhadap pengujian artefak menggunakan *En score* dan *Zeta score* sesuai ISO 17043:2010 (International Organization for Standardization, 2010) dan ISO 13528:2015 (International Organization for Standardization, 2015). Kedua parameter tersebut digunakan untuk menilai kesesuaian hasil pengukuran laboratorium terhadap nilai acuan (*assigned value*) dengan mempertimbangkan ketidakpastian pengukuran yang dilaporkan. Namun, *En score* dan *Zeta score* memiliki fokus analisis yang berbeda dan saling melengkapi. Evaluasi dari *En score* dan *Zeta score* merepresentasikan pelaporan ketidakpastian oleh laboratorium peserta yaitu realistis, terlalu optimis atau terlalu konservatif. Penelitian Skrzypczak et al., (2021) menunjukkan evaluasi *En score* dimaksudkan untuk menguji apakah hasil laboratorium peserta masih berada dalam batas toleransi gabungan

dari nilai acuan pada tingkat kepercayaan 95%, sehingga *En score* menilai apakah laboratorium peserta masih setara atau tidak. Kriteria penilaian berdasarkan pada ketentuan berikut: $|E_n| \leq 1$ = pelaporan setara dan $|E_n| > 1$ = pelaporan tidak setara. Secara matematis, *En score* didefinisikan sebagai:

$$E_n = \frac{x_{lab} - x_{ref}}{\sqrt{U_{lab}^2 + U_{ref}^2}} \quad (4)$$

dengan:

- x_{lab} = hasil pengukuran laboratorium peserta,
- x_{ref} = nilai acuan
- U_{lab} = *expanded uncertainty* laboratorium peserta
- U_{ref} = *expanded uncertainty* acuan

Berbeda dengan *En score*, *Zeta score* digunakan untuk menilai konsistensi hasil secara statistik, dengan mempertimbangkan *standard uncertainty* (u) tanpa memperluasnya menggunakan faktor cakupan (k). Evaluasi *Zeta score* untuk melihat perbedaan antara hasil laboratorium dan nilai acuan sesuai oleh besarnya ketidakpastian yang dilaporkan Arrhenius et al., (2024). Analisa Kriteria penilaian berdasarkan pada ketentuan berikut: $|z| \leq 2,0$ = hasil pelaporan konsisten, $2,0 < |z| < 3,0$ = pelaporan perlu evaluasi lebih lanjut, dan $|z| \geq 3,0$ = pelaporan tidak konsisten. Berdasarkan ISO 13528:2015, *Zeta score* didefinisikan sebagai:

$$\zeta = \frac{x_{lab} - x_{ref}}{\sqrt{u_{lab}^2 + u_{ref}^2}} \quad (5)$$

dengan:

- x_{lab} = hasil pengukuran laboratorium peserta,
- x_{ref} = nilai acuan
- u_{lab} = *standard uncertainty* laboratorium peserta
- u_{ref} = *standard uncertainty* acuan

2.2.3 Analisis Pengaruh Faktor-Faktor Pengukuran

Analisa faktor-faktor pengukuran yang berpengaruh terhadap evaluasi *En score* dan *Zeta score* menggunakan pendekatan regresi linear berganda seperti yang dilakukan oleh Susanti & Saumi (2022) dalam penelitiannya untuk mengatasi masalah multikolinearitas pada Indeks Pembangunan Manusia di Aceh Tamiang. Regresi linear berganda mampu mengukur bahwa kontribusi kenaikan indeks pembangunan manusia (IPM) di Aceh Tamiang 3,257% dipengaruhi oleh penambahan pengeluaran, 1,656% dipengaruhi oleh angka harapan hidup dan 1,658% dipengaruhi oleh rata-rata lama sekolah. Pendekatan regresi linear berganda secara umum dirumuskan dalam bentuk:

$$Y_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ji} + \varepsilon_i \quad (6)$$

pengaplikasian dalam artikel ini dituliskan dalam format:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon \quad (7)$$

dengan:

Y = *En score* / *Zeta score*

β_0 = konstanta regresi

$\beta_{1,2,3,4,5,6}$ = koefisien regresi untuk masing-masing komponen ketidakpastian

x_1 = suhu air, x_2 = rho air, x_3 = resolusi, x_4 = ketidakpastian perulangan, x_5 = ketidakpastian akibat anak timbangan, x_6 = ketidakpastian pembacaan meniskus

ε_i = galat residual diasumsikan berdistribusi normal dengan rata-rata nol

Dalam analisis regresi linear berganda, model harus memenuhi beberapa asumsi statistik agar hasil estimasi koefisien dapat dipercaya (valid). Oleh karena itu, penelitian ini menerapkan tiga uji asumsi klasik, yaitu uji normalitas, multikolinearitas, dan heteroskedastisitas. Uji normalitas dilakukan untuk memastikan bahwa residual (selisih antara nilai pengamatan dan nilai prediksi model) berdistribusi normal. Pada penelitian ini digunakan Shapiro–Wilk test, yang secara umum digunakan untuk sampel kecil hingga menengah. Kriteria interpretasi: $p > 0.05$ menunjukkan residual berdistribusi normal (asumsi terpenuhi) dan untuk $p \leq 0.05$ menunjukkan terdapat residual tidak normal. Uji multikolinearitas untuk memastikan tidak adanya hubungan linear yang sangat kuat antarsesama variabel independen. Penelitian ini menggunakan pendekatan *Variance Inflation Factor* (VIF). Dengan kriteria teknis, $VIF < 10$ menunjukkan tidak ada multikolinearitas, $VIF 5\text{--}10$ menunjukkan perlu kehati-hatian, $VIF > 10$ menandakan terjadi multikolinearitas serius (model bermasalah), dan $1/VIF$ menunjukkan *tolerance* (semakin kecil, semakin buruk). Uji heteroskedisitas menggunakan Breusch-Pagan untuk melihat apakah variansi residual konstan, dengan interpretasi $p > 0.05$ menggambarkan tidak ada heteroskedastisitas (asumsi terpenuhi) dan $p \leq 0.05$ menggambarkan terdapat heteroskedastisitas pada variabel kontrol yang digunakan.

Selanjutnya adalah uji hipotesis model regresi yaitu meliputi uji simultan (uji F), uji parsial (uji t), dan uji koefisien determinasi R^2 . Uji F digunakan untuk mengetahui apakah semua variabel independen secara bersama-sama memiliki pengaruh yang signifikan terhadap variabel dependen. Dengan asumsi $p < 0.05$ menyatakan model signifikan sehingga variabel independen secara simultan memengaruhi Y dan $p \geq 0.05$ menyatakan model tidak signifikan. Uji t digunakan untuk menilai pengaruh masing-masing variabel independen secara individual terhadap variabel dependen. Dengan kriteria koefisien positif menunjukkan variabel kontrol meningkatkan *En score* atau *Zeta score*, dan koefisien negatif menunjukkan variabel kontrol menurunkan *En score* atau *Zeta score*, dengan $p < 0.05$ menyatakan variabel berpengaruh signifikan dan $p \geq 0.05$ menyatakan variabel tidak signifikan. Selanjutnya uji koefisien determinasi R^2 menunjukkan seberapa besar proporsi variasi variabel dependen yang dapat dijelaskan oleh variabel independen dalam model. Dengan asumsi nilai $0.0 < R^2 < 0.3$ menyatakan model lemah, $0.3 < R^2 < 0.6$ menyatakan model cukup baik, $0.6 < R^2 < 0.8$ menyatakan model kuat, dan $R^2 > 0.8$ menyatakan model sangat kuat.

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis pengaruh faktor-faktor pengukuran terhadap *En score* dan *Zeta score* pada uji banding laboratorium dengan pendekatan regresi linear berganda, pada penentuan nilai acuan (x_{reff}) dan uji stabilitas artefak disajikan pada Tabel 2a dan 2b. Pengujian stabilitas artefak diperlihatkan bahwa evaluasi *drift* yang terjadi pada artefak A berada pada rentang 0,08–0,65 sedangkan pada artefak B berada pada rentang 0,03–0,60. Nilai *drift* tersebut kemudian dikonversi menjadi ketidakpastian stabilitas menggunakan pendekatan distribusi segitiga. Hasil menunjukkan bahwa u_{stab} untuk artefak A berada pada rentang 0,02–0,18, dan untuk artefak B berada pada rentang 0,008–0,17. Evaluasi nilai *drift* secara keseluruhan pada dari siklus 1 sampai dengan siklus 5 memenuhi ambang batas yang ditetapkan penyelenggara yaitu ($\Delta x \leq 2u_{awal}$) seperti ditunjukkan pada Gambar 1, adapun proporsi u_{stab} terhadap ketidakpastian acuan u_{reff} diperlihatkan pada Gambar 2 yang menunjukkan nilainya tidak melebihi 30% pada seluruh *loop*, sehingga dapat disimpulkan bahwa *drift* tidak memberikan kontribusi signifikan terhadap ketidakpastian total. Berdasarkan hasil analisis nilai acuan dan uji stabilitas, kedua artefak

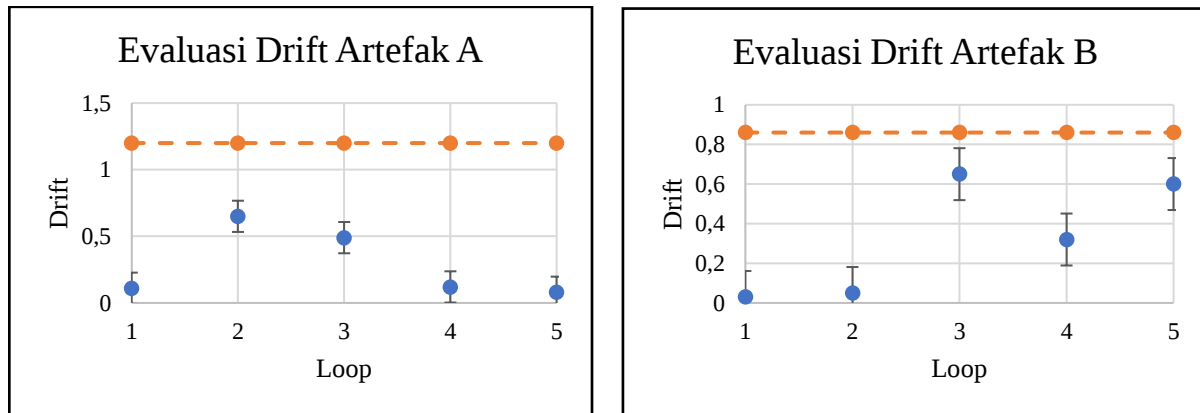
dinyatakan stabil selama proses pengujian, dan nilai acuan yang dihitung sebagai rata-rata nilai awal dan akhir dapat telah memenuhi ketentuan yang ditetapkan oleh penyelenggara.

Tabel 2a. Penentuan Nilai Acuan Artefak A

Loop	Nilai Awal	U_Awal	Nilai Akhir	U_Akhir	Drift	u_awal	u_akhir	u_stab	Nilai Acuan	U_acuan
I	10002.19	0.67	10002.3	1.2	0.11	0.335	0.6	0.031	10002.25	0.67
II	10002.3	1.2	10002.95	0.63	0.65	0.6	0.315	0.187	10002.6	1.3
III	10002.95	0.63	10002.46	0.67	0.49	0.315	0.335	0.141	10002.71	0.69
IV	10002.46	0.67	10002.34	0.67	0.12	0.335	0.335	0.035	10002.40	0.67
V	10002.34	0.67	10002.42	0.65	0.08	0.335	0.325	0.023	10002.38	0.67

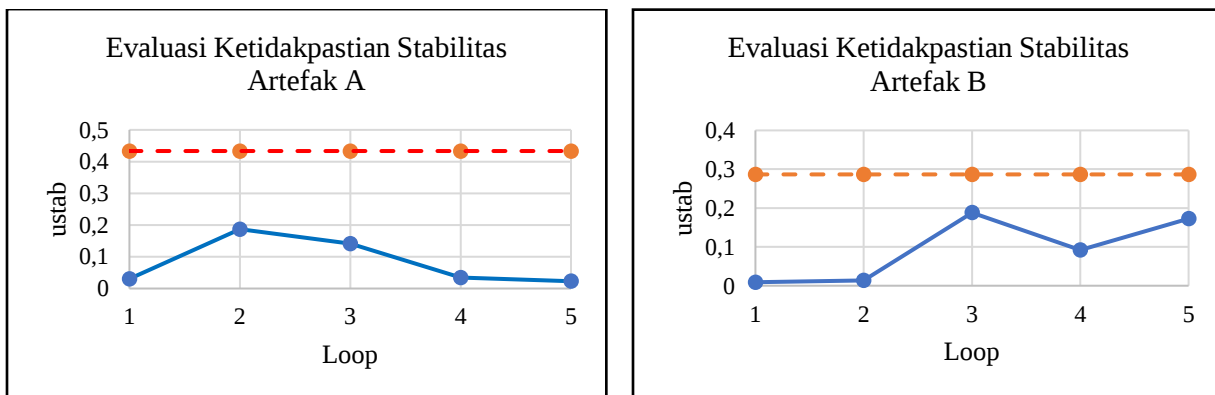
Tabel 2b. Penentuan Nilai Acuan Artefak B

Loop	Nilai Awal	U_Awal	Nilai Akhir	U_Akhir	Drift	u_awal	u_akhir	u_stab	Nilai Acuan	U_acuan
I	9999.67	0.61	9999.7	0.86	0.03	0.305	0.43	0.009	9999.69	0.61
II	9999.7	0.86	9999.65	0.68	0.05	0.43	0.34	0.014	9999.68	0.86
III	9999.65	0.68	9999	0.6	0.65	0.34	0.3	0.188	9999.33	0.78
IV	9999	0.6	9999.32	0.63	0.32	0.3	0.315	0.092	9999.16	0.63
V	9999.32	0.63	9998.72	0.73	0.6	0.315	0.365	0.173	9999.02	0.72



Gambar 1. Evaluasi Drift Artefak A dan Artefak B

*Garis merah menunjukkan batas nilai drift ($\Delta x \leq 2u_{awal}$)

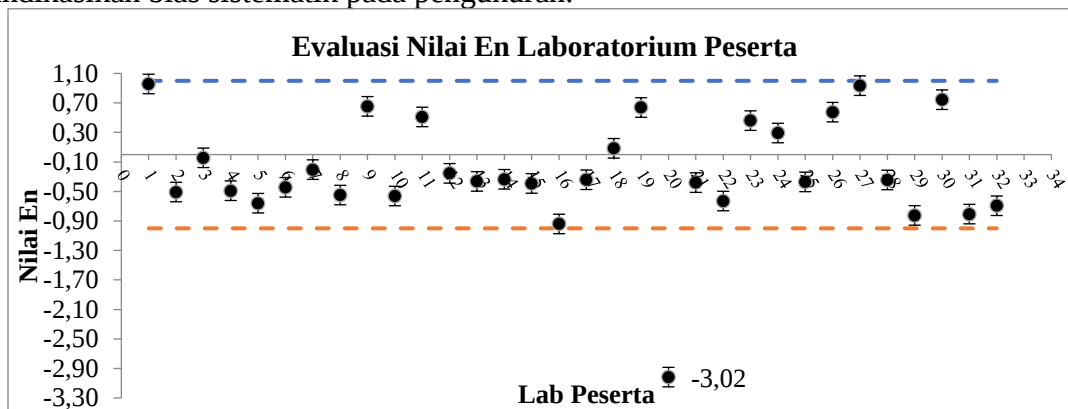


Gambar 2. Evaluasi Ketidakpastian Stabilitas Artefak A dan Artefak B

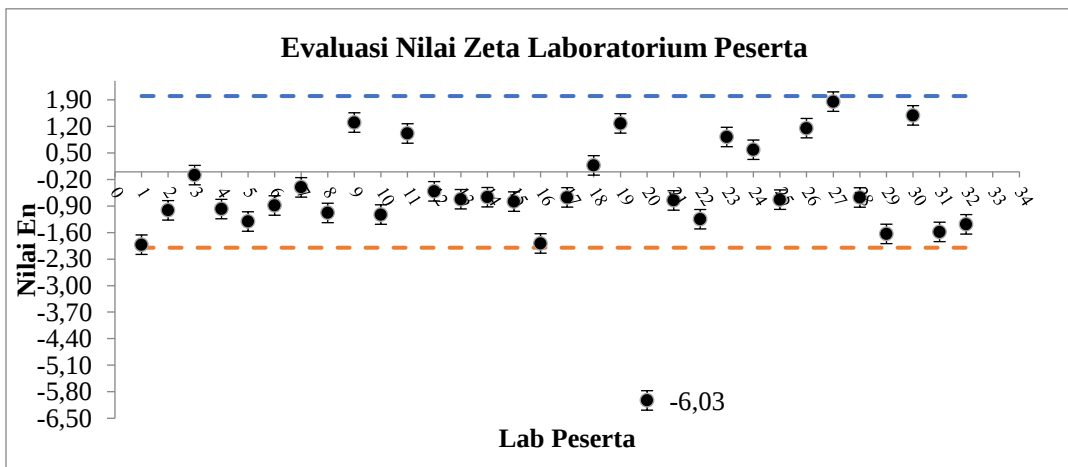
*Garis merah menunjukkan batas ketidakpastian stabilitas ($u_{stab} \leq \frac{1}{3} U_{reff}$)

Langkah selanjutnya setelah penetapan nilai acuan yakni evaluasi kesesuaian dan konsistensi pelaporan hasil uji banding menggunakan metode evaluasi *En score* dan *Zeta score* yang ditampilkan pada Gambar 3 dan Gambar 4. Berdasarkan hasil menunjukkan

bahwa hampir seluruh laboratorium peserta menunjukkan nilai dalam batas kesesuaian ($|En| \leq 1$ dan $|Zeta| \leq 2$). Namun, terdapat satu laboratorium yang menampilkan hasil berbeda secara signifikan, yaitu Lab 20. Laboratorium ini memperoleh $En\ score = -3.02$ (Tidak Setara) dan $Zeta\ score = -6.028$ (Tidak Konsisten). Nilai ini merupakan satu-satunya pasangan nilai $En\ score$ dan $Zeta\ score$ yang melampaui batas toleransi secara ekstrem dibanding seluruh peserta ($n = 32$). Penyimpangan ini menunjukkan bahwa hasil pengukuran laboratorium berbeda jauh dari nilai acuan (x_{reff}) meskipun ketidakpastian yang dilaporkan ($U_{lab} = 0.76$), hal ini tidak mencerminkan variasi aktual pengukurannya. Dengan kata lain, laboratorium ini tidak hanya menunjukkan bias pengukuran yang besar, tetapi juga melaporkan ketidakpastian yang kemungkinan terlalu optimis (*underestimated*), sehingga menghasilkan $En\ score$ dan $Zeta\ score$ yang sangat besar dan mengindikasikan bias sistematis pada pengukuran.



Gambar 3. Evaluasi $En\ score$ Laboratorium Peserta



Gambar 4. Evaluasi $Zeta\ score$ Laboratorium Peserta

Untuk mengkaji kemungkinan adanya bias sistematis yang muncul pada nilai $En\ score$ maupun $Zeta\ score$ dalam evaluasi kesesuaian hasil uji banding laboratorium, artikel ini menerapkan pendekatan regresi linear berganda berbantuan dengan aplikasi STATA versi 17. Pendekatan ini dimaksudkan untuk mengidentifikasi seberapa besar pengaruh dan kontribusi masing-masing variabel kontrol terhadap variasi nilai $En\ score$ maupun $Zeta\ score$. Dengan analisis ini, penulis dapat menilai apakah faktor seperti resolusi timbangan, kondisi lingkungan, dan besaran ketidakpastian laboratorium berperan signifikan dalam perbedaan hasil kesesuaian dan konsistensi antar laboratorium peserta.

Langkah-langkah yang diperlukan dalam analisis model regresi linear berganda terdapat beberapa tahap diantaranya uji normalitas, uji multikolinearitas, uji heteroskedastisitas, uji hipotesis (uji F, uji-t, koefisien regresi, R-squared) dan pemodelan regresi. Uji normalitas menggunakan Shapiro-Wilk, hasilnya sebagai berikut:

Tabel 3. Uji normalitas regresi terhadap residual_en dan residual_zeta

Shapiro-Wilk W test for normal data						Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	Z	Prob>z	Variable	Obs	W	V	Z	Prob>z
resid_en	32	0.975	0.815	-0.425	0.665	resid_zeta	32	0.974	0.838	-0.367	0.643

Berdasarkan uji normalitas Shapiro-Wilk terhadap residual model regresi *En score* dan *zeta-score*, diperoleh nilai $W = 0.975$ dengan p-value sebesar 0.665 untuk *En score* dan nilai $W = 0.974$ dengan p-value sebesar 0.643. Karena p-value dari kedua residual lebih besar dari $\alpha=0.05$, maka residual dianggap mengikuti distribusi normal. Dengan demikian, asumsi normalitas pada model regresi telah terpenuhi.

Tahap selanjutnya dari analisis regresi yakni uji multikolinearitas yang dimaksudkan untuk melihat variabel kontrol yang digunakan dalam pendekatan regresi tidak saling berkorelasi tinggi sehingga menyebabkan koefisien regresi menjadi tidak stabil dan nilai signifikansi menjadi bias. Berdasarkan hasil analisis multikolinearitas menggunakan STATA 17, nilai uji multikolinearitas untuk model regresi *En Score* dan *Zeta Score* nilainya adalah sama, hal ini disebabkan karena keduanya menggunakan variabel kontrol yang sama, yang membedakan dari evaluasi *Zeta score* dan *En score* adalah dari formula yang digunakan. Analisis uji multikolinearitas pada Tabel 4 menunjukkan nilai $VIF=1.45$ atau $VIF < 5$, hal ini dapat diartikan variabel kontrol pada model regresi tidak saling mempengaruhi terhadap evaluasi *En score* dan *Zeta score*, sehingga masing-masing variabel kontrol dapat dianalisis kontribusinya terhadap model regresi yang digunakan secara valid.

Tabel 4. Uji Multikolinearitas Model Regresi *En Score* dan *Zeta Score*

Variable	VIF	1/VIF
suhu_air	2.15	0.465402
rho_air	2.09	0.478333
u_repeat	1.21	0.823316
u_timb	1.16	0.861406
Resolusi	1.05	0.952558
u_meniskus	1.02	0.981722
Mean VIF	1.45	

Tahap berikutnya adalah uji heteroskedastisitas model Breush-Pagan menggunakan aplikasi STATA 17 yang ditampilkan pada Tabel 5a. uji heteroskedastisitas terhadap model regresi *En Score* menunjukkan $\text{prob}>x^2$ sebesar $0.1361 > 0.05$ yang mengindikasikan bahwa model regresi terhadap evaluasi *En* tidak mengalami masalah heteroskedastisitas sehingga sebaran *error* dinyatakan homogen. Adapun pada Tabel 5b. uji heteroskedastisitas terhadap model regresi *Zeta Score* juga menunjukkan $\text{prob}>x^2$ yaitu $0.1348 > 0.05$ yang juga mengindikasikan bahwa model regresi *Zeta Score* tidak mengalami masalah heteroskedastisitas sehingga sebaran *error* dinyatakan homogen. Kedua model regresi memenuhi asumsi regresi linear berganda dan layak untuk dilakukan pemodelan regresinya

Tabel 5. Uji heteroskedastisitas

Tabel 5a. Uji Heterokedastisitas Model Regresi En Score

Tabel 5b. Uji Heterokedastisitas Model Regresi Zeta Score

Breusch–Pagan/Cook–Weisberg test for heteroskedasticity	Breusch–Pagan/Cook–Weisberg test for heteroskedasticity
Assumption: Normal error terms	Assumption: Normal error terms
Variable: Fitted values of ϵ	Variable: Fitted values of ζ
H0: Constant variance	H0: Constant variance
$\chi^2(1) = 2.22$	$\chi^2(1) = 2.24$
Prob > $\chi^2 = 0.1361$	Prob > $\chi^2 = 0.1348$

Selanjutnya adalah uji hipotesis untuk estimasi model regresi yaitu uji F, uji t dan R^2 pada tabel 6, tabel 6a merupakan uji hipotesis untuk model regresi *En Score*, hasil pengujian simultan menunjukkan nilai $F(6,25)$ sebesar 3,74 dengan p-value 0,0086, sehingga hipotesis nol dapat ditolak. Dengan demikian, variabel suhu_air, rho_air, urepeat, utimb, resolusi, dan umeniskus secara bersama-sama berpengaruh signifikan terhadap *En Score*, sehingga model regresi dinilai layak digunakan. Nilai koefisien determinasi (R^2) sebesar 0,4727 mengindikasikan bahwa 47,27% variasi *En Score* dapat dijelaskan oleh variabel-variabel dalam model, sementara sisanya dipengaruhi oleh faktor lain di luar model. Nilai Adjusted R^2 sebesar 0,3462 menunjukkan kemampuan penjelasan model setelah mempertimbangkan jumlah prediktor yang digunakan. Adapun nilai Root MSE sebesar 0,60503 mencerminkan rata-rata besarnya kesalahan prediksi model dalam satuan *En Score*. Pada tabel 6b merupakan uji hipotesis untuk model regresi terhadap model regresi *Zeta Score*. Berdasarkan hasil pengujian simultan diperoleh nilai $F(6,25)$ sebesar 6,00 dengan p-value 0,0005, yang lebih kecil dari tingkat signifikansi 0,05. Hal ini menunjukkan bahwa hipotesis nol dapat ditolak, sehingga seluruh variabel independen secara bersama-sama berpengaruh signifikan terhadap *Zeta Score*. Dengan demikian, model regresi yang digunakan dapat dinilai memiliki kelayakan yang baik. Koefisien determinasi (R^2) sebesar 0,59 mengindikasikan bahwa 59% variasi *Zeta Score* mampu dijelaskan oleh variabel-variabel dalam model, sedangkan 41% sisanya dipengaruhi oleh faktor lain di luar variabel yang diteliti. Nilai Adjusted R^2 sebesar 0,4916 mencerminkan kemampuan model dalam menjelaskan variasi data setelah memperhitungkan jumlah prediktor yang digunakan. Sementara itu, nilai Root MSE sebesar 0,51713 menunjukkan rata-rata deviasi antara nilai prediksi dan nilai aktual *Zeta Score* dalam model tersebut.

Tabel 6. Uji Hipotesis (Uji F dan R^2)

Tabel 6a. Uji Hipotesis Model Regresi En Score			Tabel 6b. Uji Hipotesis Model Regresi Zeta Score		
Number of obs	=	32	Number of obs	=	32
$F(6, 25)$	=	3.74	$F(6, 25)$	=	6.03
Prob > F	=	0.0086	Prob > F	=	0.0005
R-squared	=	0.4727	R-squared	=	0.5916
Adj R-squared	=	0.3462	Adj R-squared	=	0.4935
Root MSE	=	0.60503	Root MSE	=	0.51713

Pada uji t (parsial) dan koefisien pada model regresi ini dijelaskan pada tabel 7. Untuk regresi terhadap *En score* dan *Zeta score* memperlihatkan bahwa kontribusi variabel kontrol resolusi menunjukkan koefisien negatif terhadap *En score* dengan nilai -7.261 ($p < 0.01$) sebanding dengan koefisien terhadap *Zeta score* yaitu -14.492 ($p < 0.01$), hal ini menandakan bahwa semakin besar resolusi timbangan atau ketelitian lebih kasar yang digunakan oleh peserta maka deviasi kinerja yang tercermin pada *En score* dan *Zeta score* akan semakin besar. Dengan kata lain, resolusi timbangan merupakan variabel kontrol yang paling kuat mempengaruhi kualitas hasil

pelaporan uji banding laboratorium peserta, seperti pada peserta dengan kode Lab 20 yang menggunakan timbangan dengan resolusi 0.5g dengan hasil evaluasi *outlier*. Variabel kontrol yang berpengaruh berikutnya adalah ketidakpastian anak timbangan (u_{timb}) sebagai faktor utama pada koreksi timbangan. Ketidakpastian timbangan berkontribusi dominan terhadap *En score* dan *Zeta score* dimana variabel u_{timb} memiliki pengaruh negatif yang signifikan terhadap kedua indikator evaluasi, yaitu *En score* dan *Zeta score*. Koefisien regresi -1.178 untuk *En score* dan -2.355 untuk *Zeta score* (keduanya signifikan pada $p < 0.01$) menunjukkan bahwa semakin kecil nilai pelaporan komponen ketidakpastian u_{timb} yang dilaporkan peserta, semakin besar nilai evaluasi *En score* dan *Zeta score* seperti pada Lab 20 dengan hasil evaluasi *outlier* karena pelaporan ketidakpastian gabungan sebesar 0.67 yang relatif kecil. Dengan demikian, variabel kontrol yang signifikan terhadap hasil *outlier* pada kegiatan uji banding yang dibahas dalam artikel ini adalah resolusi dan ketidakpastian akibat anak timbangan. Adapun variabel kontrol lain seperti suhu air, ketidakpastian perulangan, dan ketidakpastian akibat pembacaan meniskus tidak signifikan untuk *En score* maupun *Zeta score*. Hal ini menunjukkan bahwa faktor-faktor tersebut tidak memiliki pengaruh dominan terhadap kesesuaian ataupun konsistensi dalam pelaksanaan uji banding pada artikel ini.

Tabel 7. Uji t (parsial) variabel kontrol

	(1) En	(2) Zeta
suhu_air	0.114 (0.069)	0.226 (0.139)
rho_air	28.111 (261.313)	51.557 (523.740)
Resolusi	-7.261*** (0.404)	-14.492*** (0.809)
u_repeat	-0.128 (0.815)	-0.249 (1.633)
u_timb	-1.178*** (0.367)	-2.355*** (0.734)
u_meniskus	-0.855 (0.699)	-1.632 (1.365)
_cons	-29.559 (261.993)	-54.450 (525.106)
N	32	32

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Pemodelan regresi linear berganda untuk mengevaluasi kesesuaian dan konsistensi pada pelaporan hasil peserta pada kegiatan uji banding bejana ukur nominal 10 L serta analisis faktor yang berpengaruh pada hasil *outlier* peserta dapat dijelaskan dalam bentuk matematis untuk evaluasi *En score* adalah

$$Y = -29.559 + 0.114x_1 + 28.111x_2 - 7.261x_3 - 0.128x_4 - 1.178x_5 - 0.855x_6 + \varepsilon$$

Sedangkan untuk evaluasi *Zeta score* pemodelan matematis terhadap hasil uji banding laboratorium dituliskan dalam bentuk

$$Y = -54.450 + 0.226x_1 + 51.557x_2 - 14.492x_3 - 0.249x_4 - 2.335x_5 - 1.632x_6 + \varepsilon$$

4. KESIMPULAN

Artikel ini mengevaluasi dan menganalisis pelaksanaan kegiatan uji banding bejana ukur nominal 10 L menggunakan evaluasi *En score* dan *Zeta score* dengan hasil seluruh laboratorium peserta secara umum memenuhi kesesuaian dan konsisten terhadap nilai acuan, kecuali satu laboratorium peserta (Lab 20) yang teridentifikasi sebagai *outlier*. Pendekatan regresi linear berganda menunjukkan bahwa variabel resolusi dan ketidakpastian akibat anak timbangan (u_{timb}) merupakan faktor paling berpengaruh signifikan terhadap hasil evaluasi *En score* dan *Zeta score*. Hal ini sejalan dengan karakteristik pengukuran gravimetri atau penimbangan, di mana anak timbangan merupakan acuan utama dalam pembentukan koreksi dari timbangan yang digunakan dan ketidakpastian pengukuran pada uji banding bejana ukur, sementara resolusi timbangan menentukan sensitivitas pembacaan dan besarnya deviasi terhadap nilai acuan. Semakin besar atau rendah resolusi timbangan yang digunakan, semakin besar kemungkinan terjadinya deviasi hasil ukur, dan semakin besar pula kontribusi ketidakpastian dari anak timbangan yang digunakan. Secara keseluruhan, model regresi mampu menjelaskan sekitar 59% variasi *En score* dan *Zeta score* yang menunjukkan bahwa sebagian besar variasi dapat diterangkan oleh variabel kontrol dalam model, sementara sisanya dipengaruhi oleh faktor lain di luar ruang lingkup artikel ini. Terkait dengan hasil analisa bahwa ketidakpastian akibat anak timbangan berpengaruh signifikan terhadap *En score* dan *Zeta score*, maka penting untuk laboratorium peserta untuk tidak *underestimation* pada nilai ketidakpastian yang dilaporkan serta memastikan bahwa keseluruhan pelaporan ketidakpastian memenuhi metode yang digunakan. Kemudian terkait resolusi timbangan yang digunakan, maka perlu menggunakan timbangan dengan kelas resolusi yang sesuai dengan ketelitian bejana ukur yang diuji, agar mengurangi risiko bias dan memperbaiki konsistensi antar laboratorium. Berdasarkan temuan dari artikel ini bahwa sekitar 40% dari variasi *En score* dan *Zeta score* belum dapat dijelaskan oleh model regresi linear berganda, penelitian lanjutan disarankan untuk memasukkan variabel kontrol lain atau metode analisa lain yang dapat menjelaskan lebih lanjut dan terperinci terkait faktor-faktor lain yang berpengaruh signifikan pada uji banding laboratorium dengan artefak bejana ukur nominal 10 L.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada BSML Regional II yang telah mengizinkan dalam pengambilan data untuk penyusunan artikel ini, kepada rekan-rekan kerja sejawat di BSML Regional II terkhusus untuk saudari Tatik Sumiyati sudah menjadi rekan diskusi dalam penyusunan artikel ini, kepada keluarga dan seluruh pihak yang telah memberikan dukungan penuh dari persiapan sampai dengan penyelesaian artikel ini.

DAFTAR PUSTAKA

- Arrhenius, K., Morris, A., Hookham, M., Moore, N., Modugno, P., & Bacquart, T. (2024). An inter-laboratory comparison between 13 international laboratories for eight components relevant for hydrogen fuel quality assessment. *Measurement: Journal of the International Measurement Confederation*, 230. <https://doi.org/10.1016/j.measurement.2024.114553>
- Cordeiro, F., Emons, H., & Robouch, P. (2022). Is the z score sufficient to assess participants' performance in proficiency testing? The hidden corrective action. *Accreditation and Quality Assurance*, 27(3), 145–153. <https://doi.org/10.1007/s00769-022-01496-w>
- Direktorat Jenderal Perlindungan Konsumen dan Tertib Niaga. (2020). *Surat Keputusan Dirjen PKTN No. 124 Tahun 2020 tentang Standar Ukuran Metrologi Legal Besaran Volume*. Kementerian Perdagangan Republik Indonesia.

- Durgut, Y. (2021). Inter-laboratory Comparisons and Their Roles in Accreditation. *European Journal of Science and Technology*, October. <https://doi.org/10.31590/ejosat.1001957>
- International Organization for Standardization. (2010). *ISO/IEC 17043:2010: Conformity assessment – General requirements for proficiency testing*. ISO.
- International Organization for Standardization. (2015). *ISO 13528:2015: Statistical methods for use in proficiency testing by interlaboratory comparison*. ISO.
- Li, T., Zhao, H., Zhang, C., Wang, W., He, F., Zhong, K., Yuan, S., & Wang, Z. (2019). Reasons for proficiency testing failures in routine chemistry analysis in China. *Lab Medicine*, 50(1), 103–110. <https://doi.org/10.1093/labmed/lmy032>
- Oktari, Z., Masripah, Pramukti, S. R., & Fadila, M. (2023). Peningkatan Mutu Laboratorium Uji dan Kalibrasi IBBN Melalui Uji Profisiensi/Uji Banding Sesuai Persyaratan ISO/IEC 17025:2017. *Jurnal Penjaminan Mutu*, 9(1), 17–23.
- Skrzypczak, I., Leśniak, A., Ochab, P., Górka, M., Kokoszka, W., & Sikora, A. (2021). Interlaboratory comparative tests in ready-mixed concrete quality assessment. *Materials*, 14(13). <https://doi.org/10.3390/ma14133475>
- Susanti, Saumi F.(2022). Penerapan Metode Analisis Regresi Linear Berganda Untuk Mengatasi Masalah Multikolinearitas Pada Kasus Indeks Pembangunan Manusia (Ipm) Di Kabupaten Aceh Tamiang. <https://doi.org/10.33059/jgp.v4i2.5591>