



EVALUASI KINERJA *BICLUSTERING* ALGORITMA *CHENG AND CHURCH* DAN QUBIC PADA DATA PRODUKSI PERIKANAN BUDIDAYA DI INDONESIA

DEWI ZULYANI POMALINGO



PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Evaluasi Kinerja *Biclustering* Algoritma *Cheng and Church* dan QUBIC pada Data Produksi Perikanan Budidaya di Indonesia” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Dewi Zulyani Pomalingo
G1501221020



RINGKASAN

DEWI ZULYANI POMALINGO. Evaluasi Kinerja *Biclustering* Algoritma *Cheng and Church* dan QUBIC pada Data Produksi Perikanan Budidaya di Indonesia. Dibimbing oleh MUHAMMAD NUR AIDI dan KUSMAN SADIK.

Biclustering merupakan sebuah metode penggerombolan objek dan peubah secara simultan sehingga mampu menemukan pola lokal yang tidak selalu terlihat melalui metode *clustering* klasik. Beragam algoritma *biclustering* memiliki cara kerja dan tujuan optimasi yang berbeda, sementara belum terdapat pedoman khusus yang dapat dijadikan acuan untuk memilih algoritma paling sesuai. Penelitian ini bertujuan untuk mengevaluasi kinerja dua algoritma *biclustering*, yaitu CC (*Cheng and Church*) dan QUBIC (*qualitative biclustering*) pada data simulasi dan data produksi perikanan budidaya di Indonesia tahun 2023. Algoritma CC mengidentifikasi bicluster berdasarkan homogenitas numerik melalui minimisasi *Mean Squared Residue* (MSR), sedangkan QUBIC mengidentifikasi kesamaan pola kualitatif melalui proses diskretisasi.

Data simulasi dibangun dalam matriks berukuran 34×14 dengan *background* berdistribusi normal baku. Dua bicluster dengan model konstan memiliki rata-rata 5 dan 10 serta ragam 0,25 disisipkan pada matriks tersebut. Skenario simulasi terdiri atas ukuran bicluster kecil 10×4 , sedang 12×5 , dan besar 15×6 , serta tingkat tumpang tindih 0%, 10%, dan 20%. Nilai pada daerah tumpang tindih ditentukan menggunakan rata-rata terbobot. Setiap kombinasi skenario diulang sebanyak 100 kali, kemudian baris dan kolom matriks diacak sebelum algoritma CC dan QUBIC diterapkan. Kemampuan algoritma dalam mengidentifikasi *bicluster* aktual dievaluasi menggunakan indeks Liu dan Wang. Hasil simulasi menunjukkan bahwa QUBIC menghasilkan indeks Liu dan Wang yang lebih tinggi daripada CC pada seluruh skenario, dengan rata-rata keseluruhan masing-masing sebesar 0,870 dan 0,382. Uji Wilcoxon berpasangan menunjukkan perbedaan kinerja yang signifikan, sehingga QUBIC secara umum lebih baik dalam mengidentifikasi *bicluster* aktual. CC memberikan kinerja terbaik pada *bicluster* berukuran besar, sedangkan QUBIC memberikan kinerja terbaik pada ukuran kecil dan sedang.

Data empiris berupa volume produksi 14 jenis perikanan budidaya pada 34 provinsi telah distandardisasi. Pada algoritma CC, threshold optimal ditentukan melalui pengujian beberapa nilai ambang batas δ yang menghasilkan ASR rendah dengan kestabilan struktur *bicluster*. CC menghasilkan 7 bicluster pada threshold optimal $\delta = 0,009$, mencakup 33 provinsi dengan nilai *Average Squared Residue* (ASR) sebesar 0,0052. Pada algoritma QUBIC, pemilihan kombinasi parameter dilakukan dengan mempertimbangkan jumlah bicluster *nontrivial*, rata-rata volume bicluster, dan nilai ASR. QUBIC menghasilkan 8 *bicluster* pada kombinasi parameter $r = 1$, $q = 0,06$, dan $c = 0,75$ yang mencakup 31 provinsi dengan ASR sebesar 0,4360. Nilai ASR CC yang lebih rendah dibanding QUBIC menunjukkan bahwa CC lebih unggul dalam membentuk *bicluster* yang homogen pada data produksi perikanan budidaya. Hal ini juga terlihat pada plot profil bicluster CC yang cenderung lebih berimpit dan sejajar, sedangkan profil bicluster QUBIC lebih menyebar, dan umumnya menunjukkan pola naik turun yang searah.



Jumlah keanggotaan provinsi yang tercakup menunjukkan bahwa kedua algoritma mampu menangkap pola pada sebagian besar provinsi di Indonesia, meskipun CC memiliki cakupan yang sedikit lebih luas. Nilai indeks Liu dan Wang sebesar 0,1951 menunjukkan bahwa struktur keanggotaan hasil *bicluster* CC dan QUBIC memiliki tingkat kemiripan yang relatif rendah. Dengan demikian, meskipun kedua algoritma diterapkan pada data yang sama, yakni produksi perikanan budidaya, CC dan QUBIC cenderung menghasilkan penggerombolan yang berbeda.

Kata kunci: *average squared residue, biclustering, Cheng and Church, perikanan budidaya, qualitative biclustering*

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



SUMMARY

DEWI ZULYANI POMALINGO. Performance Evaluation of the Cheng and Church and QUBIC Biclustering Algorithms on Indonesian Aquaculture Production Data. Supervised by MUHAMMAD NUR AIDI and KUSMAN SADIK.

Biclustering is a method that simultaneously clusters objects and variables, thereby enabling the identification of local patterns that may not be detected by conventional clustering methods. Various biclustering algorithms have different mechanisms and optimization objectives, while no specific guideline is currently available for selecting the most appropriate algorithm. This study aims to evaluate the performance of two biclustering algorithms, namely CC (Cheng and Church) and QUBIC (Qualitative Biclustering), using simulated data and Indonesian aquaculture production data from 2023. The CC algorithm identifies biclusters based on numerical homogeneity through the minimization of Mean Squared Residue (MSR), whereas QUBIC identifies similarities in qualitative patterns through a discretization process.

The simulated data were generated as a 34×14 matrix with a standard normally distributed background. Two constant-model biclusters with means of 5 and 10 and a variance of 0,25 were embedded in the matrix. The simulation scenarios consisted of small (10×4), medium (12×5), and large (15×6) biclusters, with overlap levels of 0%, 10%, and 20%. Values in the overlapping regions were determined using a weighted average. Each combination of scenarios was replicated 100 times, after which the rows and columns of the matrix were randomly permuted before the CC and QUBIC algorithms were applied. The ability of the algorithms to identify the true biclusters was evaluated using the Liu and Wang index. The simulation results showed that QUBIC produced higher Liu and Wang index values than CC across all scenarios, with overall means of 0,870 and 0,382, respectively. The paired Wilcoxon test indicated a significant difference in performance, demonstrating that QUBIC was generally better at identifying the true biclusters. CC performed best for large biclusters, whereas QUBIC performed best for small and medium-sized biclusters.

The empirical data consisted of the production volumes of 14 aquaculture commodities across 34 provinces and were standardized prior to analysis. For the CC algorithm, the optimal threshold was determined by evaluating several values of the threshold parameter δ , with selection based on a low ASR value and the stability of the resulting bicluster structure. CC produced 7 biclusters at the optimal threshold of $\delta=0,009$, covering 33 provinces, with an Average Squared Residue (ASR) value of 0,0052. For QUBIC, the parameter combination was selected by considering the number of nontrivial biclusters, the average bicluster volume, and the ASR value. QUBIC produced 8 biclusters using the parameter combination $r=1$, $q=0,06$, and $c=0,75$, covering 31 provinces, with an ASR value of 0,4360. The lower ASR value obtained by CC indicates that CC was superior to QUBIC in forming numerically homogeneous biclusters in the aquaculture production data. This is also reflected in the bicluster profile plots, where CC shows more aligned profiles, while QUBIC shows more dispersed but similar

trends. The number of provinces included in the biclusters indicates that both algorithms were able to capture patterns across most provinces in Indonesia, although CC provided slightly broader coverage. The Liu and Wang index value of 0,1951 indicates that the membership structures of the biclusters produced by CC and QUBIC had a relatively low level of similarity. Therefore, although both algorithms were applied to the same aquaculture production data, CC and QUBIC tended to produce different clustering structures.

Keywords: aquaculture, average squared residue, biclustering, Cheng and Church, qualitative biclustering



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



EVALUASI KINERJA *BICLUSTERING* ALGORITMA *CHENG AND CHURCH* DAN QUBIC PADA DATA PRODUKSI PERIKANAN BUDIDAYA DI INDONESIA

DEWI ZULYANI POMALINGO

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister Sains pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



Penguji Luar Komisi Pembimbing pada Ujian Tesis:
Dr. Ir. Indahwati, M.Si.

@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

Judul Tesis : Evaluasi Kinerja *Biclustering* Algoritma *Cheng and Church* dan QUBIC pada Data Produksi Perikanan Budidaya di Indonesia

Nama : Dewi Zulyani Pomalingo

NIM : G1501221020

Disetujui oleh

Pembimbing 1:

Prof. Dr. Ir. Muhammad Nur Aidi, MS.



Pembimbing 2:

Dr. Kusman Sadik, M.Si.



Diketahui oleh

Ketua Program Studi:

Dr. Agus Mohamad Soleh, S.Si., M.T.

NIP 197503151999031004



Dekan Sekolah Sains Data, Matematika, dan Informatika:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.

NIP 196607021993021001



Tanggal Ujian: 5 Agustus 2026

Tanggal Lulus:



PRAKATA

Puji syukur penulis panjatkan atas kehadiran Allah subhanaahu wa ta'ala karena hanya atas izin dan kuasa-Nya penulis dapat menyelesaikan tesis ini yang berjudul “Evaluasi Kinerja *Biclustering* Algoritma *Cheng and Church* dan QUBIC pada Data Produksi Perikanan Budidaya di Indonesia”.

Tesis ini dapat terselesaikan berkat bantuan serta dukungan dari berbagai pihak. Oleh karena itu, penulis menyampaikan banyak terima kasih kepada:

1. Bapak Prof. Dr. Ir. Muhammad Nur Aidi, MS. dan Bapak Dr. Kusman Sadik, M.Si. selaku komisi pembimbing yang telah memberi bimbingan dan arahan kepada penulis sehingga dapat menyelesaikan tesis ini.
2. Keluarga Besar Departemen Statistika dan Sains Data IPB University atas ilmu dan bantuan dalam proses penyusunan tesis ini.
3. Mama, Papa dan adik penulis yang selalu memberikan doa, motivasi, dukungan dan kepercayaan kepada penulis dalam setiap proses yang dilalui.
4. Keluarga atas kasih sayang, dukungan dan semangat untuk penulis.
5. Teman-teman yang telah menjadi tempat berbagi, berdiskusi serta senantiasa membantu dan kebersamaan penulis selama proses penyelesaian tesis.

Penyusunan tesis ini masih memiliki berbagai keterbatasan, sehingga penulis sangat mengharapkan kritik dan saran yang membangun. Penulis berharap tesis ini dapat bermanfaat serta turut mendukung kemajuan ilmu pengetahuan.

Bogor, Agustus 2026

Dewi Zulyani Pomalingo



DAFTAR ISI

DAFTAR ISI	xi
DAFTAR TABEL	xii
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	3
II TINJAUAN PUSTAKA	4
2.1 <i>Biclustering</i>	4
2.2 Algoritma <i>Cheng and Church</i>	6
2.3 Algoritma <i>Qualitative Biclustering</i>	7
2.4 Evaluasi Kinerja Algoritma <i>Biclustering</i>	8
2.5 Analisis Ragam (ANOVA) Dua Arah	9
2.6 Uji Kruskal–Wallis	11
2.7 Duncan’s Multiple Range Test (DMRT)	12
2.8 Uji Wilcoxon Berpasangan	12
2.9 Perikanan Budidaya	14
III METODE	15
3.1 Data	15
3.2 Tahapan Analisis	16
IV HASIL DAN PEMBAHASAN	20
4.1 Eksplorasi Data Simulasi	20
4.2 Analisis Data Simulasi	21
4.3 Eksplorasi Data Empiris	24
4.4 Hasil <i>Biclustering</i> Algoritma <i>Cheng and Church</i>	28
4.5 Hasil <i>Biclustering</i> Algoritma QUBIC	31
4.6 Evaluasi Hasil <i>Biclustering</i>	36
V SIMPULAN DAN SARAN	38
5.1 Simpulan	38
5.2 Saran	38
DAFTAR PUSTAKA	39
LAMPIRAN	42
RIWAYAT HIDUP	48



DAFTAR TABEL

1	Data matriks <i>biclustering</i>	4
2	Tabel ANOVA dua arah	10
3	Skenario pembangkitan <i>bicluster</i>	15
4	Posisi pembangkitan <i>bicluster</i>	16
5	Matriks data penelitian	16
6	Hasil Kruskal-Wallis Indeks Liu dan Wang	23
7	Hasil uji Wilcoxon berpasangan pada seluruh skenario simulasi	24
8	Keanggotaan <i>bicluster</i> algoritma CC	30
9	Keanggotaan <i>bicluster</i> algoritma QUBIC	33
10	Evaluasi ASR algoritma CC dan QUBIC	36
11	Perbandingan evaluasi kinerja algoritma CC dan QUBIC	37

DAFTAR GAMBAR

1	Ilustrasi tipe <i>bicluster</i> (Pang 2022)	5
2	Struktur <i>bicluster</i> (Castanho <i>et al.</i> 2024)	5
3	Ilustrasi posisi tumpang tindih	6
4	Diagram alir data simulasi	17
5	Diagram alir data empiris	19
6	Skenario pembangkitan <i>bicluster</i>	20
7	Ilustrasi pengacakan <i>bicluster</i>	21
8	<i>Boxplot</i> sebaran nilai indeks Liu dan Wang	21
9	Plot profil rata-rata indeks Liu dan Wang CC dan QUBIC	22
10	<i>Heatmap</i> data produksi hasil standardisasi	24
11	<i>Boxplot</i> data produksi hasil standardisasi	25
12	Profil seluruh provinsi pada data hasil standardisasi	26
13	<i>Biplot</i> PCA data produksi hasil standardisasi	26
14	Sebaran nilai ASR pengujian awal ambang batas δ	28
15	Sebaran nilai ASR pengujian lanjutan ambang batas δ	28
16	Sebaran nilai indeks Liu dan Wang pada ambang batas yang diuji	29
17	Peta sebaran hasil <i>bicluster</i> CC (tanpa skala)	31
18	ASR kombinasi penyesuaian parameter r-q-c	32
19	Indeks Liu dan Wang kombinasi penyesuaian parameter	33
20	Peta sebaran hasil <i>bicluster</i> QUBIC (tanpa skala)	35
21	Plot profil <i>bicluster</i>	37



DAFTAR LAMPIRAN

1	Visualisasi skenario pembangkitan data simulasi	43
2	Jumlah ulangan keterhitungan nilai ASR pada identifikasi <i>bicluster</i> aktual	44
3	Hasil DMRT faktor ukuran dan tumpang tindih pada indeks Liu dan Wang	45
4	Hasil DMRT kombinasi ukuran dan tumpang tindih pada indeks Liu dan Wang	45
5	Profil <i>bicluster</i> algoritma CC	46
6	Profil <i>bicluster</i> algoritma QUBIC	47



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.