

10th MALAYSIA STATISTICS CONFERENCE 2023
Looking Beyond GDP: Toward Social Well-being and
Environmental Sustainability

26th September 2023
Sasana Kijang, Bank Negara Malaysia

**ADVANCING ENVIRONMENTALLY SUSTAINABLE GROWTH: A STATISTICAL
NARRATIVE**

**Analisis Risiko Pencemaran Udara Ekstrem di Klang, Malaysia:
Pendekatan Kopula-Vain**

Mohd Sabri Ismail¹; Nurulkamal Masseran¹

¹ Jabatan Sains Matematik, Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia (sabrismohd@92gmail.com; kamalmsn@ukm.edu.my)

Abstrak:

Kajian ini mengusulkan pendekatan kopula-vain dalam menilai risiko keberlakuan peristiwa pencemaran udara yang ekstrem yang boleh mengganggu pertumbuhan dari segi ekonomi, sosial dan alam sekitar. Kajian kes dijalankan terhadap kawasan Klang, Malaysia. Tiga ciri pencemaran udara, iaitu i) keparahan, ii) tempoh masa, dan iii) intensiti dimodelkan menggunakan pendekatan kopula-vain. Kopula-vain dipilih kerana model ini lebih realistik dan fleksibel, berbanding model taburan multivariat mahupun model kopula multivariat biasa. Kajian ini mendapati kopula-vain yang terbaik adalah berkompunen dengan kopula bivariat Joe, Tawn-2 dengan pusingan 180 darjah, BB8 dengan pusingan 90 darjah. Model ini memberikan korelasi Kendall yang positif, bermaksud tiga ciri tersebut cenderung bergerak seiringan. Ukuran ekor atasnya juga menunjukkan kebarangkalian sederhana untuk pencemaran udara ekstrem berlaku. Seterusnya, kajian simulasi daripada model kopula-vain menunjukkan bahawa kebarangkalian pencemaran udara ekstrem masih berada di tahap yang rendah, mengimplikasikan keadaan yang masih terkawal. Namun begitu, langkah waspada dan penilaian risiko berterusan perlu dilakukan untuk memastikan kesinambungan pertumbuhan mampan adalah terjamin.

Kata Kunci:

Permodelan statistik multivariat; Kopula-vain; Pengurusan risiko; Pencemaran udara

1. Pengenalan:

Pencemaran udara boleh menyebabkan banyak masalah seperti krisis kesihatan awam, kemerosotan kewangan, komplikasi psikologi, dan ketidakstabilan sosial [1]. Oleh itu, ia perlu dikawal, terutamanya pencemaran udara ekstrem yang dapat membesarkan lagi impak masalah tersebut. Maka, kaedah penilaian risiko pencemaran udara ekstrem sangat diperlukan untuk memastikan pertumbuhan mampan, dari segi ekonomi, sosial, dan alam sekitar.

Oleh kerana pencemaran udara ekstrem diwakili oleh taburan ekor atas, maka model statistik yang bersifat fleksibel dan sarat dengan sifat-sifat taburan ekor diperlukan. Sifat yang dinyatakan tiada pada model piawai yang simetri (seperti taburan normal dan

taburan t), tetapi ada pada model kopula multivariat [2]. Hal ini kerana kopula multivariat mempunyai fungsi kopula yang fleksibel dan tidak dipengaruhi oleh kepelbagaian taburan marginal yang berbeza.

Namun begitu, masih terdapat kelemahan pada model kopula multivariat kerana ianya bergantung pada hanya satu parameter untuk menerangkan struktur kebergantungan yang kompleks. Oleh itu, untuk mengatasi masalah ini, pendekatan kopula-vain yang lebih realistik dan fleksibel telah diperkenalkan. Hal ini kerana, permodelan kopula-vain menggunakan kopula bivariat (bersyarat) sebagai blok-blok pembina yang mempunyai parameter tertentu bagi setiap pasangan pemboleh ubah yang terlibat [3]. Maka, hubungan antara setiap pasangan pemboleh ubah dapat dimodelkan dengan lebih baik dan seterusnya ukuran pengukuran risiko yang lebih tepat dapat diperolehi.

Dalam kajian lampau, kopula-vain telah digunakan dalam pelbagai aplikasi dunia nyata dalam pelbagai bidang. Sebagai contoh, mengoptimalkan pulangan portfolio pelaburan (kewangan) [4], analisis kualiti air (pencemaran) [5], meramal kebarangkalian pair bocor (kejuruteraan) [6], memahami kemungkinan risiko banjir (bencana) [7], mengkaji kelangsungan makanan ruji (keterjaminan makanan) [8]. Dengan semangat yang sama, kajian ini dijalankan untuk meneroka aplikasi kopula-vain dalam konteks risiko pencemaran udara ekstrem di Klang, Malaysia. Lanjutan penulisan ini akan berkisarkan kepada kaedah (sub-topik 2), hasil dan perbincangan (sub-topik 3), dan kesimpulan (sub-topik 4).

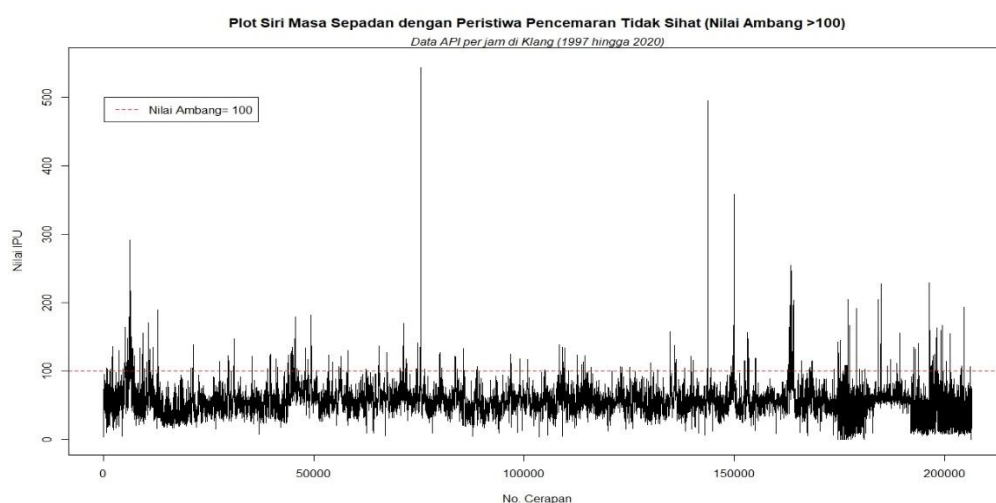
2. Kaedah:

Kajian ini menggunakan data indeks pencemaran udara (IPU) per jam, yang diperoleh daripada Jabatan Alam Sekitar, bagi tempoh 1 Januari 1993 sehingga 31 Ogos 2020. Rajah 1 menunjukkan plot siri masa bagi data IPU. Seterusnya, daripada data cerapan, 301 peristiwa tidak sihat yang melebihi nilai ambang IPU 100 telah dikenal pasti. Untuk setiap tempoh peristiwa tidak sihat ke- j , $T_j = \{t | x_t > 100\}$, yang mana x_t adalah nilai IPU pada masa t , ukuran keparahan, tempoh, dan intensiti diperoleh menggunakan formula-formula berikut:

$$kep_j = \sum_{t \in T_j} x_t, \quad (1)$$

$$tem_j = |T_j|, \quad (2)$$

$$int_j = \max_{t \in T_j} \{x_t\}. \quad (3)$$



Rajah 1. Indeks pencemaran udara untuk tahun 1997 sehingga 2020.

Oleh itu, diperoleh tiga set data untuk keparahan, tempoh, dan intensiti, iaitu $X_1 = \{x_{1,j} | j \in I\}$, $X_2 = \{x_{2,j} | j \in I\}$, dan $X_3 = \{x_{3,j} | j \in I\}$, dengan set indeks $I = \{1, 2, \dots, 301\}$. Untuk $k = 1, 2, 3$ semua unsur ke- j bagi setiap set ke- k dijelmakan kepada unsur kopula dengan menggunakan fungsi taburan empirik, iaitu

$$u_{k,j} = \hat{F}_k(x_{k,j}) = \frac{1}{N+1} \sum_{j=1}^N 1_{\{x \leq x_{k,j}\}}, \text{ for all } x \in X_k. \quad (4)$$

Menggunakan unsur kopula, permodelan kopula-vain bagi keparahan, tempoh, dan intensiti dibangunkan. Modelnya adalah

$$c(u_{kep}, u_{tem}, u_{int}; \theta) = c_{int,tem;kep}(C_{int|kep}(u_{int}|u_{kep}), C_{tem|kep}(u_{tem}|u_{kep}); \theta_{int,dur;kep}), \quad (5) \\ \times c_{int,kep}(u_{int}, u_{kep}; \theta_{int,kep}) \times c_{tem,kep}(u_{tem}, u_{kep}; \theta_{tem,kep})$$

di mana komponen kopula bivariat (bersyarat) $c_{int,tem;kep}$, $c_{int,kep}$, dan $c_{tem,kep}$ yang terbaik perlu dikenal pasti. Dalam kajian ini, sebilangan kopula bivariat berparameter yang dibincangkan dalam Schepsmeier [9] telah dicuba untuk setiap komponen. Kopula bivariat terbaik dikenal pasti menggunakan anggaran kebolehdajian maksimum (AKM) dan kriteria maklumat Akaike (KMA). Seterusnya, kopula-vain berserta komponennya dioptimumkan secara keseluruhan menggunakan anggaran kebolehdajian maksimum bercantum (AKMb) [10].

Kemudian, kopula-vain ini digunakan untuk menghitung dua ukuran kebergantungan, iaitu ukuran korelasi Kendall dan ukuran ekor atas [10]. Dua ukuran ini dihitung daripada purata bagi nilai-nilai korelasi Kendall dan ukuran ekor atas komponen-komponennya. Dua ukuran ini masing-masing mewakili sifat korelasi untuk keseluruhan data dan sifat taburan ekor yang sinonim dengan peristiwa pencemaran udara ekstrem.

Kajian ini dilanjutkan dengan simulasi daripada kopula-vain dengan menggunakan penjelmaan Rosenbalt and fungsi sonsangannya [11]. Data dalam purata asal juga diperoleh daripada kaedah pensampelan diskret songsang. Untuk simulasi, 20 sampel bagi setiap data keparahan, tempoh, dan intensiti dibangunkan dan setiap sampel mengandungi 1000 unsur. Seterusnya, setiap sampel yang berkaitan digunakan untuk menghitung kebarangkalian bersyarat seperti berikut:

$$P_{kep|int,tem} = P(Kep \leq kep | Int \geq int, Tem \geq tem), \quad (6)$$

dan nilainya divariasikan seperti berikut: $kep \in \{30000, 35000\}$, $int \in \{340, 420, 500\}$, $tem \in \{133, 177, 221\}$. Seterusnya, untuk setiap kombinasi (kep, int, tem) , purata ukuran kebarangkalian bersyarat tersebut dihitung untuk tujuan penilaian risiko. Lebih tinggi purata kebarangkalian bersyarat, bermaksud kadar intensiti and tempoh tersebut berkemungkinan memberikan lebih keparahan akibat pencemaran udara.

Seterusnya, purata yang dihitung juga digunakan untuk mengkaji anggaran tempoh ulangan bagi kebarangkalian bersyarat tersebut, iaitu

$$TU_{kep|int,tem} = \frac{Jul(TU)}{P_{kep|int,tem}}, \quad (7)$$

di mana $Jul(TU)$ adalah purata tempoh ulangan daripada data lampau, iaitu 106 hari untuk kajian ini. Semakin rendah nilai anggaran tempoh ulangan ini, semakin kerap pencemaran udara dijangka akan berlaku di Klang, Malaysia.

3. Hasil dan Perbincangan:

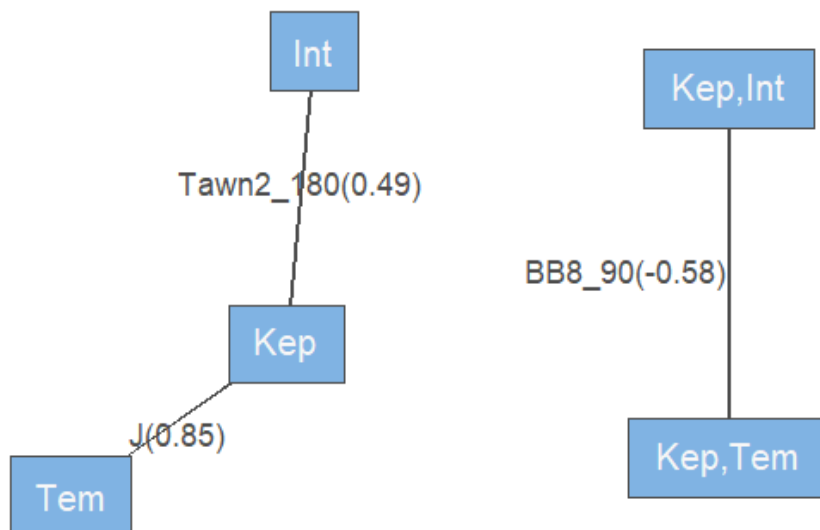
Untuk menganalisis risiko keparahan pencemaran udara, peristiwa tak sihat dikaji dengan fokus diberikan kepada tiga cirinya yang saling berhubung kait, iaitu keparahan,

intensiti dan tempoh. Statistik perihalan terhadap data keparahan, intensiti, dan tempoh dipaparkan dalam Jadual 1. Secara umumnya, taburan data tersebut didapati tidak menumpu pada nilai tengah, mencapah, pencong, dan tidak simetri. Sifat-sifat ini boleh dimodelkan menerusi pendekatan kopula-vain tak simetri.

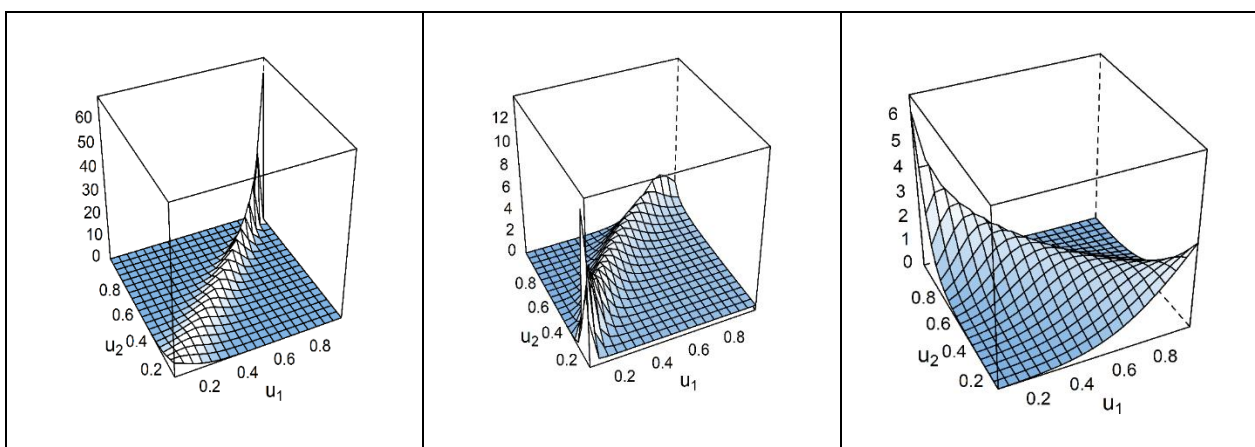
Jadual 1. Statistik perihalan terhadap data keparahan, intensiti, dan tempoh.

Data	Purata	Median	Nilai min.	Nilai maks.	Sisihan piawai	Kepencongan	Kurtosis
Keparahan	2241.76	231.27	100	36677	4948.3	3.92	20.92
Intensiti	125.11	112	100	543	44.77	5.61	44.97
Tempoh	16.74	2	1	224	31.91	3.24	15.73

Menggunakan kaedah permodelan kopula-vain yang dibincang dalam sub-topik 2 dan dinyatakan pada persamaan (5), komponen-komponen terbaik bagi $c_{tem,kep}$, $c_{int,kep}$, dan $c_{int,tem;kep}$ masing-masing adalah kopula bivariat Joe, Tawn 2 dengan pusingan 180 darjah, BB8 dengan pusingan 90 darjah. Struktur kopula-vain berserta komponennya ditunjukkan pada Rajah 2. Selain itu, plot ketumpatan bagi setiap komponen juga digambarkan pada Rajah 3.



Rajah 2. Struktur kopula-vain berserta komponennya.



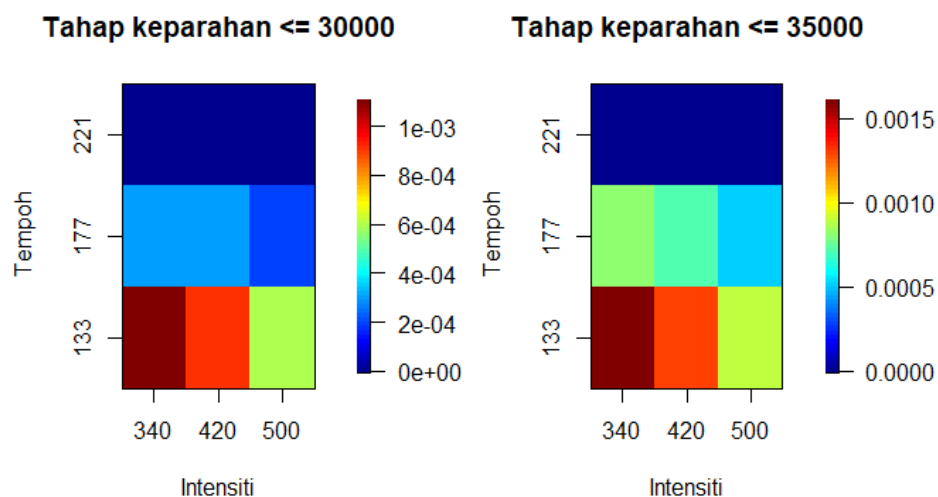
Rajah 3. Dari kiri ke kanan: Joe, Tawn 2 dengan pusingan 180 darjah, BB8 dengan pusingan 90 darjah.

Seterusnya, daripada kopula-vain yang diperoleh, dua ukuran kebergantungan, iaitu korelasi Kendall dan ukuran ekor atas boleh dihitung, di mana hasilnya dinyatakan pada Jadual 2. Korelasi Kendall (0.26) yang positif menunjukkan bahawa data keparahan, intensiti, dan tempoh cenderung untuk bergerak selari dalam arah yang sama. Malah, ukuran ekor atas (0.31) juga menantijahkan peristiwa-peristiwa ekstrem berlaku dalam kebarangkalian sederhana. Ini mengisyaratkan bahawa penilaian risiko perlu dilakukan untuk meneliti keadaan peristiwa ekstrem dengan lebih mendalam.

Jadual 2. Dapatan untuk dua ukuran kebergantungan.

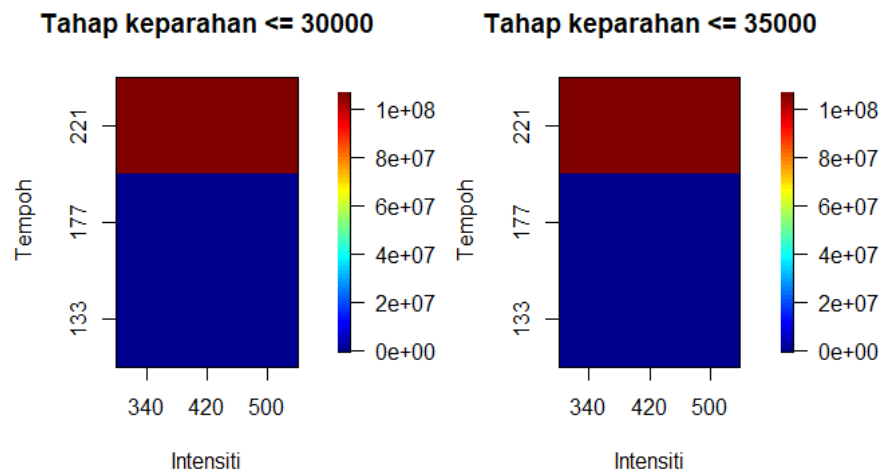
Model	Korelasi Kendall	Ukuran ekor atas
Kopula-vain	0.26	0.31

Untuk penilaian risiko yang lebih teliti, kajian simulasi dijalankan terhadap kopula-vain. Kebarangkalian bersyarat bagi keparahan, bersyarat kepada intensiti dan tempoh, dan anggaran tempoh ulangannya dihitung untuk menganalisis risiko keparahan pencemaran udara ekstrem di Klang, Malaysia. Dua dapatan terakhir ini masing-masing ditunjukkan pada Rajah 4 dan Rajah 5.



Rajah 4. Peta warna untuk kebarangkalian bersyarat pada $kep = 30000$ (kiri) dan $kep = 35000$ (kanan).

Berdasarkan Rajah 4, kedua-dua peta warna yang diberi menunjukkan kebarangkalian bersyarat tidak melebihi 0.0016. Selain itu, perubahan warna keduanya juga berubah daripada merah ke biru, apabila tempoh dan intensiti meningkat. Ini menyatakan bahawa kebarangkalian bersyarat bagi keparahan dengan syarat tempoh dan intensiti semakin berkurang. Untuk anggaran tempoh ulangan, dua peta warna pada Rajah 5 pula menunjukkan bahawa tidak kurang daripada 66250 hari, dan angka ini jauh daripada 106 hari (purata tempoh ulangan daripada data lampau). Malah, warna di dalam dua peta warna tersebut juga berubah daripada warna biru kepada warna merah. Ini menunjukkan bahawa anggaran tempoh ulangan semakin membesar. Oleh itu, dapat disimpulkan bahawa pencemaran udara ekstrem di Klang, Malaysia masih berada di dalam tahap yang rendah dan keadaannya yang masih terkawal.



Rajah 5. Peta warna untuk anggaran tempoh ulangan pada $kep = 30000$ (kiri) dan $kep = 35000$ (kanan).

4. Kesimpulan:

Kajian ini mengusulkan pendekatan kopula-vain untuk memodelkan struktur kebergantungan untuk data keparahan, intensiti dan tempoh. Kemudian, ukuran kebergantungan (korelasi Kendall dan ukuran ekor atas) dan dua ukuran risiko daripada simulasi kopula-vain (kebarangkalian bersyarat dan anggaran tempoh ulangan) digunakan untuk menganalisis risiko keberlakuan pencemaran udara dengan keparahan ekstrem di Klang, Malaysia. Dapatan kajian ini menunjukkan bahawa pencemaran udara di Klang, Malaysia masih berada di dalam tahap yang rendah dan keadaannya yang masih terkawal. Namun begitu, langkah waspada dan penilaian berterusan perlu dilakukan untuk memastikan kesinambungan pertumbuhan mampan terjamin.

Rujukan:

1. Lu, J.G., Air pollution: A systematic review of its psychological, economic, and social effects. *Current Opinion in Psychology*, 2020. 32: p. 52-65.
2. Jaworski, P., et al., Copula theory and its applications. Vol. 198. 2010: Springer.
3. Czado, C. and T. Nagler, Vine copula based modeling. *Annual Review of Statistics and Its Application*, 2022. 9: p. 453-477.
4. Low, R.K.Y., et al., Canonical vine copulas in the context of modern portfolio management: Are they worth it? *Journal of Banking & Finance*, 2013. 37(8): p. 3085-3099.
5. Shan, B., et al., Vine copula and cloud model-based programming approach for agricultural water allocation under uncertainty. *Stochastic Environmental Research and Risk Assessment*, 2021. 35(9): p. 1895-1915.
6. Atique, F. and N. Attoh-Okine, Using copula method for pipe data analysis. *Construction and Building Materials*, 2016. 106: p. 140-148.
7. Tosunoglu, F., F. Gürbüz, and M.N. İspirli, Multivariate modeling of flood characteristics using Vine copulas. *Environmental Earth Sciences*, 2020. 79(19): p. 459.
8. Wu, H., et al., Agricultural Drought Prediction Based on Conditional Distributions of Vine Copulas. *Water Resources Research*, 2021. 57(8): p. e2021WR029562.
9. Schepsmeier, U., et al., Package 'vinecopula'. R package version, 2015. 2(5).
10. Czado, C., Analyzing dependent data with vine copulas. *Lecture Notes in Statistics*, Springer, 2019. 222.
11. Khashei-Siuki, A., et al., Simulation of potential evapotranspiration values based on vine copula. *Meteorological Applications*, 2021. 28(5): p. e2027.