



# 11<sup>th</sup> MALAYSIA STATISTICS CONFERENCE 2024

Data and Artificial Intelligence: Empowering the Future

Sasana Kijang, Bank Negara

19<sup>th</sup> September 2024

## Pertumbuhan Ekonomi dan Analisis Data Bertopologi

Mohd Sabri Ismail<sup>1</sup> dan Mohd Salmi Md Noorani<sup>1</sup>

<sup>1</sup> Jabatan Sains Matematik, Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia

### Abstrak:

Kajian ini mengusulkan homologi gigih, iaitu satu kaedah utama dalam bidang matematik baharu yang dikenali sebagai analisis data bertopologi untuk menganalisis dan meramal keadaan pertumbuhan ekonomi Malaysia. Untuk itu, siri masa KDNK bagi tempoh masa 1 Januari 2016 sehingga 1 Januari 2024 yang dalamnya terkandung satu peristiwa kemelesetan ekonomi akibat pandemik Covid-19 dan krisis politik tanah air digunakan. Menggunakan siri masa KDNK tersebut, jentera homologi gigih digunakan untuk menghasilkan siri masa norma- $L_1$ . Siri masa norma- $L_1$  digunakan sebagai satu perwakilan baharu untuk menganalisis dan meramal pertumbuhan ekonomi. Hasil kajian menunjukkan siri masa norma- $L_1$  memaparkan satu trend peningkatan yang besar dan ketara, dan meningkat sehingga ke satu puncak tertinggi ketika berlakunya peristiwa kemelesetan ekonomi. Hal ini menggambarkan bahawa analisis data bertopologi dapat digunakan sebagai satu ukuran ekonometrik baharu bagi menggambarkan keadaan pertumbuhan ekonomi negara dan menjangkakan kemunculan kemelesetan ekonomi yang mungkin wujud akibat terdapatnya gangguan dalam pasaran. Oleh itu, menggunakan kaedah ini, pihak berkepentingan seperti kerajaan, peniaga, pelabur, dan pelanggan dapat membuat keputusan yang lebih baik berdasarkan penganalisisan dan peramalan pertumbuhan ekonomi yang lebih dipercayai.

### Kata Kunci:

Pertumbuhan ekonomi; KDNK; Sistem Dinamik; Analisis Data Bertopologi; Homologi Gigih

### 1. Pengenalan:

Analisis dan peramalan keadaan pertumbuhan ekonomi sesebuah negara sangat penting untuk pihak berkepentingan seperti kerajaan, peniaga, pelabur, dan pelanggan supaya mereka dapat membuat keputusan yang lebih baik untuk menguruskan aktiviti ekonomi mereka. Tugas ini biasanya melibatkan kefahaman kepada perilaku kedinamikan penunjuk utama bagi pertumbuhan ekonomi yang biasanya digambarkan oleh siri masa keluaran dalam negeri kasar (KDNK). Siri masa KDNK adalah jujukan bagi ukuran jumlah nilai barangan dan perkhidmatan yang dihasilkan oleh sesebuah negara pada tempoh masa tentu. Penunjuk ini sering menjadi bahan berita ekonomi dan diambil

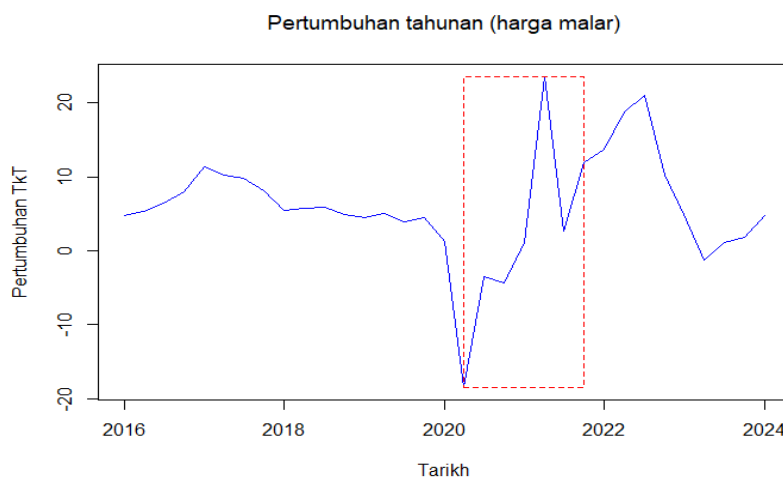
kira sebagai salah satu penanda aras dalam pembentukan polisi kerajaan bagi menjamin kestabilan ekonomi sehingga matlamat makroekonomi tercapai.

Namun begitu, menganalisis dan meramal perilaku siri masa KDNK dengan tepat bukanlah sesuatu tugas yang mudah. Hal ini kerana siri masa ini adalah suatu sistem dinamik yang bersifat kompleks, tak linear, mempunyai ralat dan kalut [1]. Maka, satu kaedah yang teguh perlu digunakan untuk memberikan satu analisis dan peramalan yang bermakna. Untuk itu, kaedah teguh seperti homologi gigih (HG) boleh digunakan untuk menganalisis dan meramal perilaku siri masa KDNK [2]. HG adalah salah satu kaedah teguh yang diperkenalkan dalam satu bidang matematik baharu yang dikenali sebagai analisis data bertopologi (ADbT) [3]. Pendekatan ADbT sangat menarik kerana ianya memfokuskan pemodelan rupa bentuk disebalik siri masa tak linear, mempunyai ralat dan kalut, sehinggakan sifat invarian topologi, iaitu komponen-komponen berkait yang wujud dalam siri masa tersebut dapat dicam dan dihitung untuk tujuan analisis dan peramalan.

Dengan mengfokuskan kepada sifat invarian topologi yang teguh terhadap ralat data, HG dapat menghasilkan satu perwakilan baharu, iaitu siri masa norma- $L_1$  yang dapat menggambarkan perilaku pertumbuhan ekonomi negara dan mengesan kemungkinan berlakunya kemelesetan ekonomi yang boleh menyebabkan ketidakstabilan ekonomi berlaku. Seterusnya, penulisan makalah ini akan berkisarkan kepada kaedah (Bahagian 2), hasil dan perbincangan (Bahagian 3), dan kesimpulan (Bahagian 4).

## 2. Kaedah:

Kajian ini menggunakan data siri masa KDNK Malaysia yang diperoleh daripada platform digital OpenDOSM yang dinyatakan secara berkala secara suku tahunan bagi tempoh 1 Januari 2016 sehingga 1 Januari 2024. Selain itu, tempoh masa bagi kemelesetan ekonomi yang direkodkan oleh portal tersebut yang berlaku pada suku kedua tahun 2020 (01-04-2020) sehingga suku keempat tahun 2021 (01-10-2021) juga dirujuk dalam kajian ini. Kemelesetan ekonomi ini berlaku akibat pandemik Covid-19 dan krisis politik tanah air [4, 5]. Siri masa KDNK yang digunakan dan peristiwa kemelesetan ekonomi yang diperhatikan ditunjukkan dalam Rajah 1.



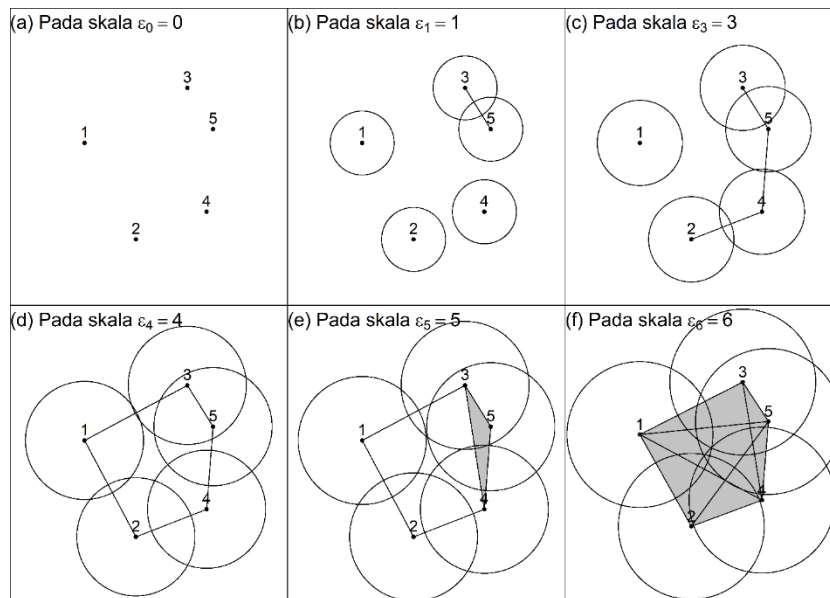
**Rajah 1.** Siri masa KDNK Malaysia (garisan biru) dan peristiwa kemelesetan ekonomi (dalam kotak merah putus-putus).

Andaikan siri masa KDNK adalah set  $X = \{x_t \in \mathbb{R} | t \in I\}$ , di mana  $I$  adalah indeks suku tahunan. Perihal statistik (ukuran min, sisihan piawai, kepencongan, dan kurtosis) bagi set  $X$  dihitung untuk menyatakan maklumat awal mengenainya. Seterusnya, prapemprosesan set  $X$  dijalankan untuk menyediakan data input kepada kaedah HG

untuk menganalisis dan meramal kedinamikannya. Untuk itu, set  $X$  dipetakan kepada suatu set  $Y$  yang berdimensi 3, dengan menggunakan kaedah pemetaan Takens dengan lat masa satu suku tahun. Hasilnya adalah set  $Y = \{x_t = (x_t, x_{t-1}, x_{t-2}) \in \mathbb{R}^3 | t \in J\}$ , dimana indeks  $J \subset I$ .

Kemudian, set  $Y$  dipecahkan kepada subset-subsetnya menggunakan kaedah tettingkap gelonsor dengan saiz tettingkap satu tahun (bersamaan empat data suku tahunan). Hasilnya, subset  $Y_t = \{x_t, x_{t-1}, \dots, x_{t-3} | t \in K\} \subset Y$ , di mana indeks  $K \subset J$ . Setiap subset  $Y_t$  adalah input data kepada HG untuk menganalisis dan meramal kedinamikan siri masa KDNK Malaysia.

Untuk setiap subset  $Y_t$ , HG digunakan untuk memodelkan rupa bentuk ruang geometri yang terdapat di sebalik subset tersebut menggunakan pemodelan kompleks simpleks Rips pada skala pembinaan model yang berbeza dan menggunakan homologi untuk mengecam dan menghitung sifat invarian topologi, iaitu komponen-komponen berkait yang wujud sepanjang selang masa skala pembinaan model tersebut. Rajah 2 menunjukkan pemodelan kompleks simpleks Rips pada skala pembinaan model yang berbeza untuk satu contoh set data yang simpel.



**Rajah 2.** Pemodelan kompleks simpleks Rips pada skala pembinaan model yang berbeza untuk satu contoh set data yang simpel.

Kemudian, tempoh jangka hayat bagi komponen-komponen berkait yang dihitung HG bagi set subset  $Y_t$  boleh dinyatakan sebagai set  $JH = \{(l_i, m_i) | i \in M\}$ , di mana  $M$  adalah indeks integer positif terhingga, dan  $l_i$  dan  $m_i$  adalah tempoh lahir dan mati bagi suatu komponen berkait ke- $i$ . Menggunakan set  $JH$  yang diperolehi, nilai norma- $L_1$  boleh dihitung daripada landskap kegigihan yang diterbitkan daripada set  $JH$ .

Diberi set  $JH$ , terdapat satu fungsi cebis demi cebis, iaitu

$$g_{(l_i, m_i)}(x) = \begin{cases} x - l_i, & l_i \leq x \leq (l_i + m_i)/2 \\ l_i - x, & (l_i + m_i)/2 < x < m_i \\ 0, & \text{selainnya} \end{cases}$$

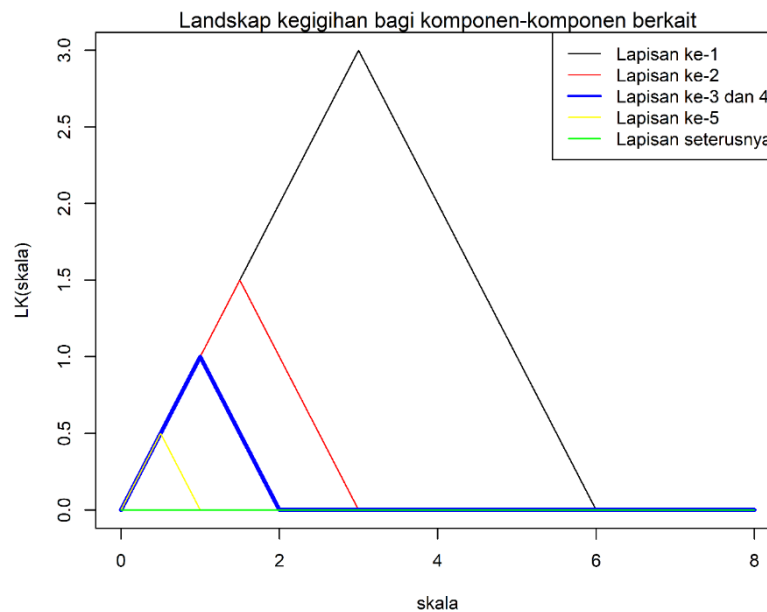
Kemudian, lapisan ke- $k$  bagi landskap kegigihan boleh ditakrifkan sebagai

$$\lambda_{JH}^k(x) = k - \max_{(l_i, m_i) \in JH} g_{(l_i, m_i)}(x),$$

di mana  $k - \max$  adalah nilai maksimum ke- $k$  bagi set terhingga semua fungsi  $g_{(l_i, m_i)}(x)$ .

Jika nilai  $k - \max$  tidak wujud bagi  $k = r$ , maka  $\lambda_{JH}^k(x) = 0$  untuk semua  $k \geq r$ . Maka,

landskap kegigihan boleh dinyatakan sebagai suatu set  $\lambda = \{\lambda_{j_H}^1(x), \lambda_{j_H}^2(x), \dots\}$  [6]. Rajah 3 menunjukkan landskap kegigihan bagi komponen-komponen berkait yang terbit daripada Rajah 2 sebelumnya.

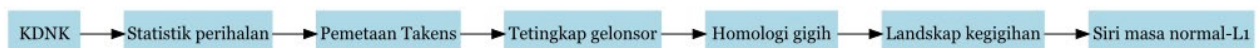


**Rajah 3.** Landskap kegigihan bagi komponen-komponen berkait yang terbit daripada Rajah 2 sebelumnya.

Untuk satu set  $\lambda$ , norma- $L_1$  bagi  $\lambda$  boleh ditakrifkan seperti berikut:

$$\|\lambda\| = \left[ \sum_{k=1}^{\infty} \int |\lambda_{j_H}^k(x)| dx \right].$$

Maka, dapatan akhir HG adalah satu siri masa baru, iaitu siri masa norma- $L_1$  yang boleh diwakilkan sebagai  $Z = \{\|\lambda\|_t \in \mathbb{R} | t \in K\}$ , di mana  $K$  adalah indeks suku tahunan yang sama di takrif sebelumnya. Rajah 4 menyatakan carta alir kaedah analisis dan peramalan yang digunakan dalam kajian ini.



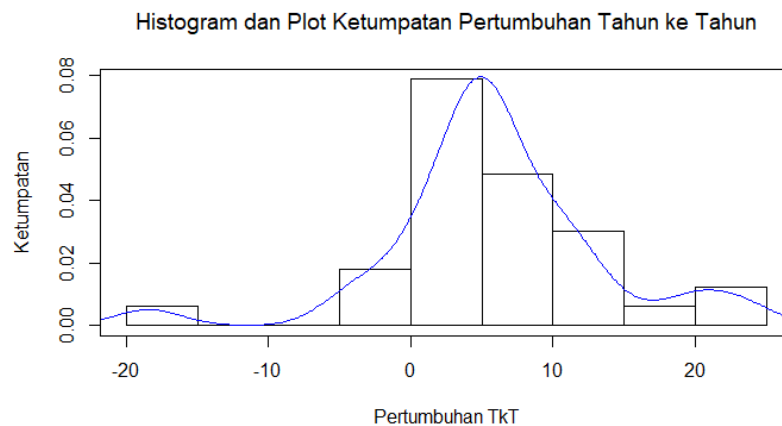
**Rajah 4.** Carta alir kaedah analisis dan peramalan yang digunakan dalam kajian ini.

### 3. Hasil dan Perbincangan:

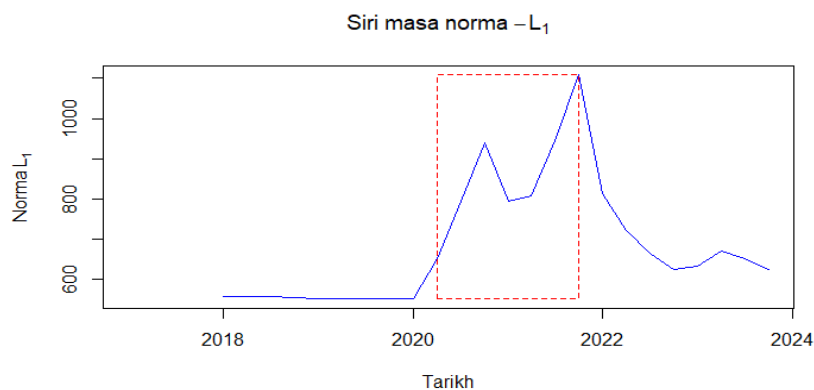
Untuk melihat keupayaan ADbT dalam menganalisis dan meramal keadaan kedinamikan KDNK Malaysia, siri masa KDNK Malaysia dan peristiwa kemelesetan ekonomi yang diperoleh daripada platform digital telah digunakan. Statistik perihalan terhadap siri masa KDNK tersebut dipaparkan dalam Jadual 1. Daripada Jadual 1, dapat dinyatakan bahawa min KDNK berpusat pada 5.87, sisihan piawai menunjukkan KDNK menyimpang dari min secara purata sebanyak 7.59, kepencongan negatif (-0.32) menunjukkan KDNK sedikit condong ke arah kiri, dan kurtosis (2.1) menunjukkan KDNK mempunyai taburan dengan ekor yang lebih ringan berbanding taburan normal. Ilustrasi histogram dan plot ketumpatan pertumbuhan tahun ke tahun untuk KDNK ditunjukkan dalam Rajah 5. Rajah 5 jelas menunjukkan bahawa taburan KDNK adalah tidak menumpu pada kadar pertumbuhan sifar, mencapah, pencong dan tidak simetri.

**Jadual 1.** Statistik perihalan terhadap data siri masa KDNK negara Malaysia.

| Minimum | Maksimum | Min  | Sisihan piawai | Kepencongan | Kurtosis |
|---------|----------|------|----------------|-------------|----------|
| -18.43  | 23.52    | 5.87 | 7.59           | -0.32       | 2.1      |

**Rajah 5.** Histogram dan plot ketumpatan pertumbuhan tahun ke tahun.

Walaupun statistik perihalan (min, sisihan piawai, kecondongan, dan kurtosis) dapat memberikan gambaran tentang ciri-ciri keseluruhan sesuatu siri masa, ia biasanya tidak mencukupi untuk meramal krisis seperti kemelesetan ekonomi yang mungkin akan berlaku pada masa hadapan. Meramalkan krisis dalam siri masa biasanya melibatkan teknik dan model yang lebih canggih seperti HG yang diperkenalkan dalam bidang baharu yang dinamakan ADbT ini. Menerusi HG, siri masa norma- $L_1$  dapat diperoleh dan digunakan untuk menggambarkan keadaan kedinamikan siri masa KDNK Malaysia. Rajah 6 menggambarkan siri masa norma- $L_1$  yang diperoleh daripada kajian ini.

**Rajah 6.** Siri masa norma- $L_1$ 

Daripada Rajah 6, didapati bahawa siri masa norma- $L_1$  jelas menunjukkan peningkatan nilai yang besar dan ketara sehingga ke puncak tertinggi, yang meningkat ketika berlakunya kemelesetan ekonomi di Malaysia yang disebabkan oleh Covid-19 dan krisis politik tanah air. Peningkatan ini adalah disebabkan jangka hayat kewujudan komponen-komponen berkait dalam KDNK lebih gigih ketika peristiwa kemelesetan ekonomi berbanding dari keadaan biasa yang stabil atau beransur pulih kepada keadaan stabil. Dengan kata lain, variasi dalam siri masa norma- $L_1$  dapat menerangkan keadaan semasa pertumbuhan ekonomi di Malaysia.

Ini jelas menunjukkan bahawa ADbT boleh dijadikan satu ukuran ekonometrik baharu yang dapat menggambarkan keadaan pertumbuhan ekonomi negara Malaysia. Malah, kaedah ini juga berpotensi untuk dibangunkan sebagai satu tatacara untuk mengesan isyarat amaran awal bagi kemelesetan ekonomi Malaysia, seperti yang telah dibangunkan untuk kegunaan peramalan pergerakan harga dan kemungkinan krisis kewangan dalam pasaran saham dunia, termasuk di Amerika, Singapura dan Malaysia, serta pasaran matawang digital kripto [7-10].

## Kesimpulan:

Kajian ini mengusulkan ADbT, terutamanya kaedah HG untuk menganalisis dan meramal keadaan pertumbuhan ekonomi Malaysia menerusi siri masa KDNK. Kecanggihan pendekatan ini dalam penganalisisan dan peramalan siri masa KDNK boleh dinyatakan dengan dua sebab, iaitu (1) kaedah ini menggunakan kaedah pemetaan Takens dan tettingkap gelonsor untuk menghasilkan input data yang bermakna untuk digunakan dengan HG dan (2) kaedah ini menghasilkan siri masa baharu yang dikenali sebagai siri masa norma- $L_1$  untuk mewakili keadaan pertumbuhan ekonomi Malaysia. Secara spesifiknya, kajian ini mendedahkan bahawa HG menerusi siri masa norma- $L_1$  menunjukkan peningkatan yang besar dan ketara sehingga menerbitkan satu puncak yang tinggi ketika berlakunya peristiwa kemelesetan ekonomi disebabkan pandemik Covid-19 dan krisis politik di Malaysia. Kajian ini menunjukkan bahawa ADbT, terutamanya HG dapat menawarkan satu kaedah ekonometrik yang baharu dan teguh untuk menggambarkan keadaan pertumbuhan ekonomi di Malaysia dan mengesan kemungkinan berlakunya kemelesetan ekonomi sehingga menyebabkan ketidakstabilan dalam pertumbuhan ekonomi negara. Untuk kajian lanjutan, kaedah peramalan berasaskan pembelajaran mesin yang menggunakan corak kedinamikan siri masa norma- $L_1$  boleh diusulkan.

## Rujukan:

1. Kelsey, D. (1988). The Economics of Chaos or the Chaos of Economics. *Oxford Economic Papers*, 40(1), 1-31.
2. Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., & Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Science*, 6, 1-38.
3. Baas, N. A., Carlsson, G. E., Quick, G., Szymik, M., & Thauale, M. (2020). *Topological Data Analysis*. Springer.
4. Hunter, M. (2021). Malaysia in crisis: political instability and feelings of hopelessness. *The Round Table*, 110(5), 610-611.
5. Shah, A. U. M., Safri, S. N. A., Thevadas, R., Noordin, N. K., Abd Rahman, A., Sekawi, Z., Ideris, A., & Sultan, M. T. H. (2020). COVID-19 outbreak in Malaysia: Actions taken by the Malaysian government. *International journal of infectious diseases*, 97, 108-116.
6. Bubenik, P. (2020). The persistence landscape and some of its properties. *Topological Data Analysis: The Abel Symposium 2018*,
7. Ismail, M. S., Hussain, S. I., & Noorani, M. S. M. (2020). Detecting Early Warning Signals of Major Financial Crashes in Bitcoin Using Persistent Homology. *IEEE Access*, 8, 202042-202057.
8. Ismail, M. S., Md Noorani, M. S., Ismail, M., & Abdul Razak, F. (2022). Early Warning Signals of Financial Crises Using Persistent Homology and Critical Slowing Down: Evidence From Different Correlation Tests [Original Research]. *Frontiers in Applied Mathematics and Statistics*, 8.
9. Ismail, M. S., Noorani, M. S. M., Ismail, M., Razak, F. A., & Alias, M. A. (2020). Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange. *Applied Soft Computing*, 93, 106422.
10. Ismail, M. S., Noorani, M. S. M., Ismail, M., Razak, F. A., & Alias, M. A. (2022). Early warning signals of financial crises using persistent homology. *Physica A: Statistical Mechanics and its Applications*, 586, 126459.