

Model Recurent Neural Network untuk Peramalan Produksi Tebu Nasional

Aziz Kustiyo ^{1,*}, Mukhlis Mukhlis ², Aries Suharso ³

¹ Ilmu Komputer; Institut Pertanian Bogor; Kampus IPB Dramaga, Bogor;
e-mail: azizku@apps.ipb.ac.id

² Informatika; Universitas Bhayangkara Jaya; Jl. Raya Perjuangan Bekasi Utara, Kota Bekasi;
e-mail: clisfadhlmukhlis@apps.ipb.ac.id

³ Teknik Informatika; Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo,
Telukjambe Timur, Karawang; e-mail: aries.suharso@unsika.ac.id

* Korespondensi: e-mail: azizku@apps.ipb.ac.id

Diterima: 21 Mei 2022; Review: 02 Juni 2022; Disetujui: 03 Juni 2022

Cara sitasi: Kustiyo A, Mukhlis M, Suharso A. 2022. Model Recurent Neural Network Untuk Peramalan Produksi Tebu Nasional. Bina Insani ICT Journal. Vol.9 (1): 1-10.

Abstrak: Produksi tebu di Indonesia tersebar di beberapa wilayah yang mengakibatkan variabilitas yang tinggi dari variabel-variabel yang mempengaruhi produksi tebu nasional. Di samping itu, tidak mudah untuk mendapatkan data-data tersebut dalam waktu yang cukup panjang. Oleh karena itu peramalan produksi tebu nasional berdasarkan variabel-variabel tersebut sangat sulit dilakukan. Sebagai solusi dari masalah tersebut, maka peramalan produksi tebu nasional dilakukan berdasarkan data historisnya. Penelitian ini bertujuan untuk mengembangkan model *recurrent neural networks* (RNN) untuk peramalan produksi tebu nasional berdasarkan data historisnya. Data yang digunakan adalah data produksi tebu nasional dari tahun 1967 sampai dengan tahun 2019 dalam satuan ton. Sebagai data latih digunakan data tahun 1967 sampai dengan tahun 2006 dan sisanya dipakai sebagai data uji. Pada penelitian ini dilakukan percobaan untuk mengetahui pengaruh panjang deret waktu dan ukuran *batch* terhadap kinerja model RNN dengan tiga ulangan. Hasil penelitian menunjukkan bahwa model RNN dengan panjang deret waktu 4 dan ukuran *batch* 16 menghasilkan nilai *mean absolut percentage error* (MAPE) sebesar 9.0% dengan nilai korelasi 0.77. Secara umum, model RNN yang dibangun mampu menangkap pola produksi tebu nasional dengan tingkat kesalahan yang masih dapat ditoleransi.

Kata kunci: deret waktu, peramalan, produksi tebu, *recurrent neural networks*

Abstract: Sugarcane production in Indonesia is spread over several regions. This condition results in high variability of the variables that affect national sugarcane production. In addition, it is not easy to obtain these data over a long period. As a result, it is very difficult to forecast the production of national sugarcane based on the influencing variables. Therefore, the forecasting was based on historical data of the national sugarcane production. This study aims to develop a recurrent neural networks (RNN) model for forecasting national sugarcane production based on historical data. The data used is national sugarcane production data from 1967 to 2019 in tons. As training data, data from 1967 to 2006 were used and the rest was used as test data. In this study, an experiment was conducted to determine the effect of time series length and batch size on the performance of the RNN model with three replications. The results showed that the RNN model with a time series length of 4 and a batch size of 16 produced a mean absolute percentage error (MAPE) of 9.0% with a correlation value of 0.77. In general, the RNN model is able to capture the national sugarcane production pattern with a tolerable error rate.

Keywords: forecasting, recurrent neural networks, sugarcane production, time series

1. Pendahuluan

Salah satu permasalahan yang dihadapi Indonesia sampai dengan saat ini adalah belum terpenuhinya kebutuhan gula dalam negeri. Sejak tahun 1986, Indonesia menjadi negara pengimpor gula dari negara lain [1]. Besarnya impor gula menunjukkan kurangnya jumlah produksi dalam negeri dibandingkan dengan jumlah konsumsi yang terus meningkat antara lain karena bertambahnya jumlah penduduk. Pada tahun 2007, volume impor gula Indonesia mencapai 2,973 juta ton dengan nilai 1,040 USD [2]. Nilai tersebut merupakan nilai impor tertinggi sejak tahun 1986. Kurangnya produksi gula tersebut salah satunya merupakan akibat dari kurangnya produksi tebu sebagai bahan utama pembuatan gula. Oleh karena itu peramalan produksi tebu nasional sangat penting dilakukan agar pemerintah dapat merencanakan kebijakan terkait dengan defisit gula yang akan terjadi pada masa mendatang.

Peramalan produksi tebu dapat dilakukan berdasarkan data historis produksi tebu maupun berdasarkan variabel-variabel yang mempengaruhi produksi tebu. Secara umum, peramalan produksi tebu untuk wilayah yang kecil seperti kabupaten atau wilayah yang mencakup beberapa kabupaten dapat dilakukan berdasarkan variabel-variabel yang mempengaruhi produksi tebu. Hal ini dapat dilakukan karena ketersediaan data dari variabel-variabel tersebut. Di samping itu untuk wilayah yang kecil, variabel-variabel lingkungan yang mempengaruhi produksi tebu seperti cuaca dan kondisi tanah pada umumnya relatif homogen. Untuk wilayah yang lebih luas seperti peramalan produksi tebu nasional, peramalan berdasarkan data historis produksi tebu lebih memungkinkan untuk dilakukan. Produksi tebu di Indonesia tersebar di beberapa wilayah seperti Sumatera Utara, Sumatera Selatan, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur, Gorontalo dan Sulawesi Selatan [3]. Produksi tebu di wilayah tersebut disajikan pada Tabel 1. Hal ini menyebabkan keragaman yang tinggi terhadap variabel-variabel yang mempengaruhi produksi tebu nasional. Di samping itu, tidak mudah untuk mendapatkan data-data tersebut dalam waktu yang cukup panjang. Oleh karena itu peramalan produksi tebu nasional berdasarkan variabel-variabel tersebut sangat sulit dilakukan.

Tabel 1. Produksi tebu di beberapa provinsi di Indonesia (ton)

No	Provinsi	2015	2016	2017
1	Sumatera Utara	29680	17936	9582
2	Sumatera Selatan	204506	112837	89010
3	Lampung	743883	676443	632321
4	Jawa Barat	84899	84770	72580
6	Jawa Tengah	231662	172537	173857
7	Jawa Timur	1207333	1047414	1023514
8	Gorontalo	49059	30678	52791
9	Sulawesi Selatan	34805	39727	42999

Sumber: Kementan (2018)

Peramalan produksi tebu berdasarkan variabel-variabel yang mempengaruhinya untuk wilayah yang kecil memungkinkan untuk dilakukan seperti peramalan produksi tebu di PT. Perkebunan Nusantara IX (PTPN IX) [4]. Variabel-variabel yang digunakan dalam penelitian [4] adalah curah hujan dan produksi tebu dalam dua tahun terakhir. Data yang digunakan merupakan data curah hujan dan data produksi tebu tahun 1979-2010 yang diambil dari Pabrik Gula Mojo (PG Mojo), salah satu pabrik gula dari PTPN IX. Wilayah kerja PTPN IX meliputi provinsi Jawa Tengah dan memiliki 4 pabrik gula [5]. Dengan demikian, wilayah yang diteliti oleh [4] dapat disebut sebagai wilayah yang kecil. Pada penelitian [4] model yang digunakan untuk peramalan produksi tebu adalah model jaringan syaraf tiruan propagasi balik. Hasil penelitian [4] menunjukkan kinerja model yang baik dengan nilai *Mean Absolut Percentage Error* (MAPE) sebesar 2,6547%.

Penelitian terkait peramalan untuk wilayah yang luas, yaitu produksi tebu nasional, antara lain dilakukan oleh [6] yang menggunakan model *autoregressive integrated moving average* (ARIMA) untuk wilayah Tamilnadu. Data yang digunakan dalam [6] adalah data tahun 1950-2007. Penelitian [7] menggunakan data tahun 1947-2018 untuk peramalan produksi tebu di Pakistan menggunakan ARIMA. Untuk India, data tahun 1950-2018 digunakan untuk peramalan produksi tebu nasional dengan mengaplikasikan metode ARIMA [8]. Ketiga penelitian tersebut menghasilkan model terbaik yang berbeda-beda di mana penelitian [6] menghasilkan model terbaik ARIMA (1,1,1), ARIMA (2,1,1) penelitian [7], dan penelitian [8] menghasilkan beberapa model terbaik di antaranya adalah ARIMA (2,1,3).

Selain model ARIMA, model yang banyak digunakan dalam peramalan adalah model jaringan syaraf tiruan. Secara umum, arsitektur model jaringan syaraf tiruan (JST) dapat

dibedakan menjadi jaringan syaraf umpan maju (*feedforward neural networks*, FNN) dan jaringan syaraf berulang (*recurrent neural networks*, RNN) [9]. Pada FNN, tidak mengasumsikan adanya ketergantungan antara data ke- n dengan data ke- $(n-1)$ sedangkan pada RNN terdapat asumsi ketergantungan antar data. Dengan demikian, model RNN telah memasukkan konsep waktu ke dalam model, sehingga model RNN lebih sesuai untuk data deret waktu (*time series*) [10].

Arsitektur awal model RNN diperkenalkan oleh Jordan pada tahun 1986 dan Elman tahun 1990. Model RNN Jordan merupakan model FNN dengan satu lapisan tersembunyi yang ditambah dengan dengan unit khusus. Unit khusus ini akan menyimpan nilai keluaran jaringan pada waktu ke- t , yang kemudian akan mengumpukan nilai-nilai tersebut ke simpul tersembunyi pada waktu ke- $(t+1)$. Arsitektur RNN yang diperkenalkan oleh Elman lebih sederhana dibandingkan arsitektur RNN Jordan. RNN Elman dilatih menggunakan propagasi balik (*backpropagation*) dan menunjukkan bahwa jaringan dapat mempelajari ketergantungan waktu [10].

Proses pembelajaran (*learning*) pada RNN menggunakan propagasi balik dan variasinya menimbulkan masalah yang disebut sebagai *vanishing gradients* atau *exploding gradients*. Hal ini terutama untuk pembelajaran dengan ketergantungan yang cukup panjang (*long-range dependencies*). Masalah tersebut kemudian dapat diatasi dengan diperkenalkannya *Long Short-Term Memory* (LSTM) oleh Hochreiter dan Schmidhuber tahun 1997. Pada LSTM, terdapat unit baru yang disebut sebagai sel memori. Unit ini merupakan unit komputasi yang menggantikan simpul tradisional di lapisan tersembunyi dari RNN. Dengan sel memori ini, jaringan dapat mengatasi kesulitan pelatihan yang dihadapi oleh RNN terdahulu. Model RNN ini menyerupai jaringan saraf berulang standar dengan lapisan tersembunyi, tetapi setiap node biasa di lapisan tersembunyi digantikan oleh sel memori. Setiap sel memori berisi simpul dengan tepi berulang yang terhubung sendiri dengan bobot tetap, memastikan bahwa gradien dapat melewati banyak langkah waktu tanpa adanya *vanishing gradients* atau *exploding gradients* [10].

Model RNN dengan konsep LSTM telah memasukkan faktor ketergantungan jangka panjang (*long-range dependencies*) antar data dalam modelnya [10]. Di samping itu, model RNN memiliki kinerja lebih baik dibandingkan rantai Markov untuk peramalan data sekuensial [11]. Terkait dengan peramalan produksi tebu nasional, model RNN tersebut telah digunakan untuk peramalan berbagai komoditas pertanian seperti gandum [12], kedelai [13], jagung [14] dan pisang [15] dengan kinerja yang cukup baik. Model RNN mampu memodelkan data sekuensial seperti halnya data historis produksi tebu nasional yang merupakan data deret waktu (*time series*).

Penelitian perbandingan antara model LSTM dan ARIMA telah dilakukan antara lain oleh [16] dan [17]. Data yang digunakan oleh [16] adalah data deret waktu dari beberapa indeks harga saham dan data ekonomi dari bulan Januari tahun 1985 sampai dengan bulan Agustus 2018. Hasil penelitian [16] menunjukkan bahwa dibandingkan model ARIMA, model LSTM mampu mereduksi nilai *root mean square error* (RMSE) sebesar 87.445% untuk prediksi harga saham dan 84.394% untuk data ekonomi. Penelitian [17] berkaitan dengan prediksi kecepatan angin per jam. Data yang digunakan adalah data kecepatan angin per jam dari tanggal 1 Mei 2021 sampai dengan 20 Juni 2021. Penelitian yang dilakukan oleh [17] menunjukkan bahwa LSTM menghasilkan nilai RMSE dan *mean absolute error* (MAE) yang lebih kecil dibandingkan dengan ARIMA. Dengan demikian, model LSTM memiliki kinerja yang lebih baik dibandingkan model ARIMA.

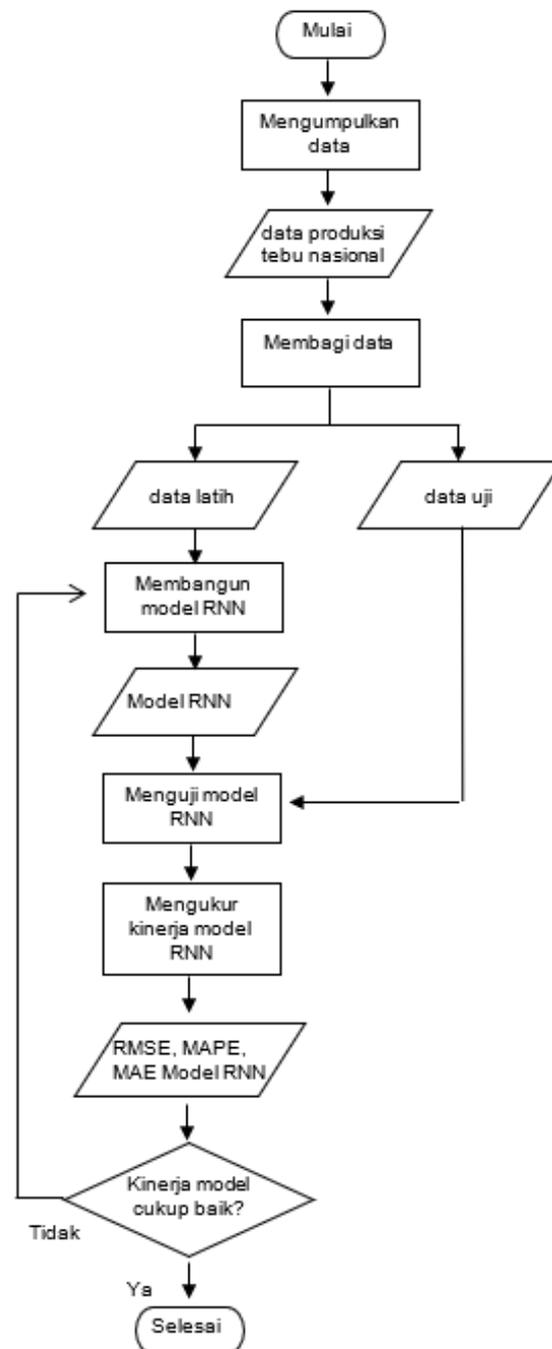
Berdasarkan uraian di atas, tujuan dari penelitian ini adalah mengembangkan model RNN dengan konsep LSTM [10] dalam peramalan hasil produksi tebu nasional. Hasil yang diperoleh diharapkan dapat memberikan gambaran kinerja model RNN dalam peramalan produksi tebu nasional.

2. Metode Penelitian

Tahapan penelitian disajikan pada gambar 1. Secara garis besar, penelitian diawali dengan mengumpulkan data, membagi data menjadi data latih dan data uji, membuat model RNN dari data latih, menguji model RNN menggunakan data uji dan mengukur kinerja model yang telah dibuat.

Pada tahap pengumpulan data, digunakan data sekunder yaitu data produksi tebu nasional diperoleh dari Kementerian Pertanian [3]. Data tersebut merupakan data produksi tebu

nasional dari tahun 1967 sampai dengan tahun 2019 dalam satuan ton. Data produksi tebu nasional dari tahun 2010 sampai dengan 2019 disajikan pada Tabel 2. Pada penelitian ini yang digunakan sebagai data latih adalah data dari tahun 1967-2006 (40 data) dan sisanya menjadi data uji.



Sumber: Hasil Penelitian (2022)

Gambar 1. Tahapan penelitian

Pada tahap pembuatan model RNN, dilakukan percobaan dengan faktor berupa panjang periode dan ukuran *batch* dengan tiga ulangan. Arsitektur model RNN disajikan pada gambar 2. Percobaan dilakukan dengan mengimplementasikan model RNN menggunakan model sequensial dari *library keras* melalui *google colab*. Panjang periode yang dicobakan adalah 4,6 dan 8 sedangkan ukuran *batch* adalah 10 dan 16.

Tabel 2. Luas areal dan produksi tebu berdasar pengusahaan tahun 2010-2019

Tahun	Luas areal				Produksi			
	PR	PBN	PBS	Total	PR	PBN	PBS	Total
2010	275908	68141	110062	454111	1295319	315174	679623	2290116
2011	278733	67020	106035	451788	1366294	295635	605958	2267887
2012	266379	77719	107157	451255	1543411	336288	711988	2591687
2013	291132	67434	110661	469227	1561047	294069	695911	2551026
2014	290967	77504	109638	478108	1516551	354733	707889	2579173
2015	267064	66715	120391	454171	1440260	339129	718609	2497997
2016	271564	52150	121360	445075	1286910	229374	688335	2204619
2017	247883	73015	109466	430363	1220780	268892	631999	2121671
2018*)	235925	69719	111932	417576	1254455	271291	648654	2174400
2019**)	253582	78686	120970	453238	1404099	339593	706308	2450000

Keterangan:

PR : perkebunan rakyat,

PBN: perkebunan besar negara,

PBS: perkebunan besar swasta

Sumber: Kementan (2018)

Kinerja model RNN diukur menggunakan RMSE, *mean absolute percentage error* (MAPE), *mean absolute error* (MAE) dan korelasi (*r*). MAPE digunakan untuk mengetahui persentase kesalahan prediksi terhadap data sebenarnya, sedangkan MAE digunakan untuk mengukur nilai kesalahan prediksi yang sebenarnya (dalam ton). MAPE dapat digunakan untuk membandingkan hasil penelitian ini dengan hasil penelitian lain terutama ketika nilai produksi tebu memiliki rentang yang berbeda. MAE dapat digunakan untuk menentukan kesalahan peramalan secara riil sesuai data yang ada. Korelasi digunakan untuk mengetahui keeratan hubungan linier antara dua variabel [18]. Nilai korelasi mendekati 1 berarti bahwa nilai hasil peramalan model memiliki kecenderungan dan nilai yang mendekati nilai aktual. Nilai korelasi sebesar 0 berarti bahwa pola hasil peramalan tidak dapat mengikuti pola data aktual, sedangkan nilai korelasi mendekati -1 mengandung makna bahwa nilai hasil peramalan memiliki kecenderungan yang berlawanan dengan data aktual. Persamaan 1 sampai dengan persamaan 4 merupakan persamaan untuk menghitung RMSE, MAPE, MAE dan *r*.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \dots\dots\dots (1)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right| \dots\dots\dots (2)$$

$$MAE = \frac{\sum_{i=1}^N |x_i - \hat{x}_i|}{N} \dots\dots\dots (3)$$

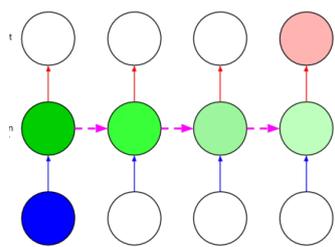
$$r = \frac{N \sum_{i=1}^N x_i \hat{x}_i - (\sum_{i=1}^N x_i)(\sum_{i=1}^N \hat{x}_i)}{\sqrt{(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2)(N \sum_{i=1}^N \hat{x}_i^2 - (\sum_{i=1}^N \hat{x}_i)^2)}} \dots\dots\dots (4)$$

dengan

x_i : nilai data ke - *i*

\hat{x}_i : nilai prediksi data ke - *i*

N : banyaknya data

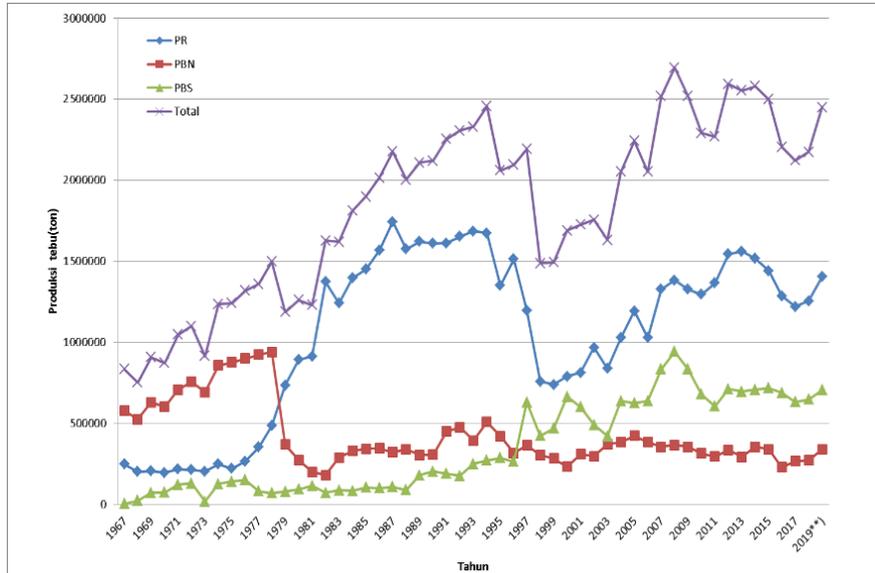


Sumber: Lipton et. al (2015)

Gambar 2. Arsitektur model RNN

3. Hasil dan Pembahasan

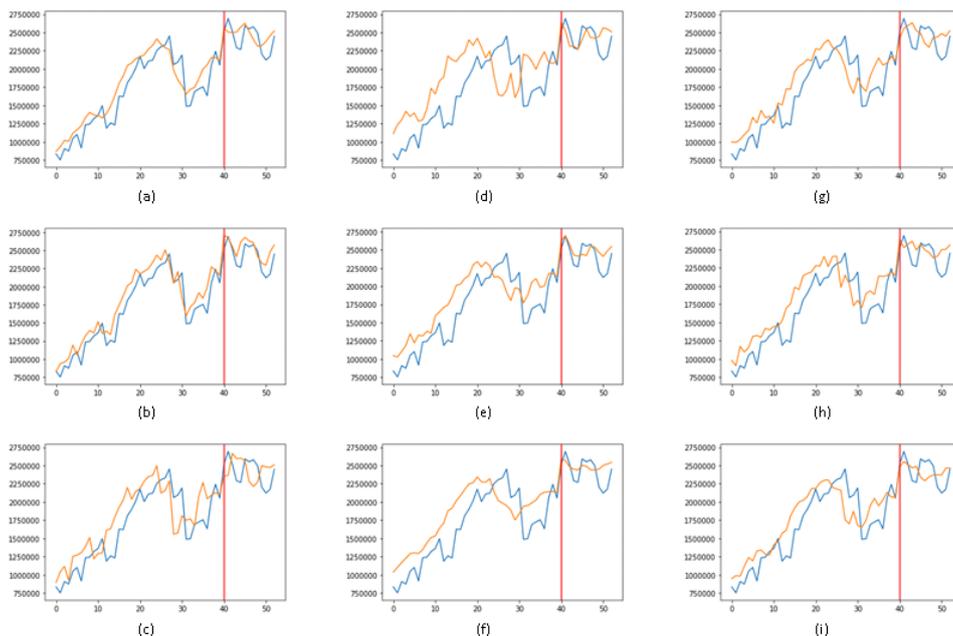
Data produksi tebu nasional diperoleh dari Kementerian Pertanian [4] yang mencatat produksi tebu nasional dari tahun 1967 sampai dengan tahun 2019 dalam satuan ton. Data tahun 2018 dan 2019 merupakan data hasil perkiraan. Data tebu nasional tersebut merupakan penjumlahan dari produksi tebu perkebunan rakyat (PR), perkebunan besar negara (PBN) dan perkebunan besar swasta (PBS). Grafik data deret waktu produksi tebu disajikan pada Gambar 3 yang menunjukkan bahwa secara umum proporsi produksi tebu nasional (total) masih didominasi oleh PR.



Sumber: Kementan (2018)

Gambar 3. Produksi tebu nasional tahun 1967-2019

Setelah dilakukan pelatihan dengan 100 epoch, model RNN yang sudah dibangun, digunakan untuk memprediksi produksi tebu nasional dari data uji. Kinerja model RNN untuk ukuran batch sebesar 10 disajikan pada Gambar 4. Garis merah vertikal pada Gambar 4 menunjukkan batas antara data latih dan data uji. Data latih berada di sebelah kiri garis merah, dan data uji di sebelah kanan.



Sumber: Hasil Penelitian (2022)

Gambar 4. Grafik produksi tebu nasional dan prediksinya untuk ukuran batch =10.

(a),(b),(c) untuk $T=4$, (d),(e),(f) untuk $T=6$ dan (g),(h),(i) untuk $T=8$.

Pada Gambar 4 dapat dilihat bahwa pola hasil prediksi RNN yang paling mendekati pola data sebenarnya adalah pada panjang periode (T) 4 dan ulangan kedua (Gambar 4b). Sementara itu, berdasarkan pengamatan untuk ukuran batch sebesar 16, model dengan pola yang mirip dengan data sebenarnya adalah model RNN untuk T=4 pada ulangan pertama. Secara visual, model RNN untuk T=4 mampu menangkap pola produksi tebu nasional dengan baik khususnya untuk ulangan kedua (Gambar 4b).

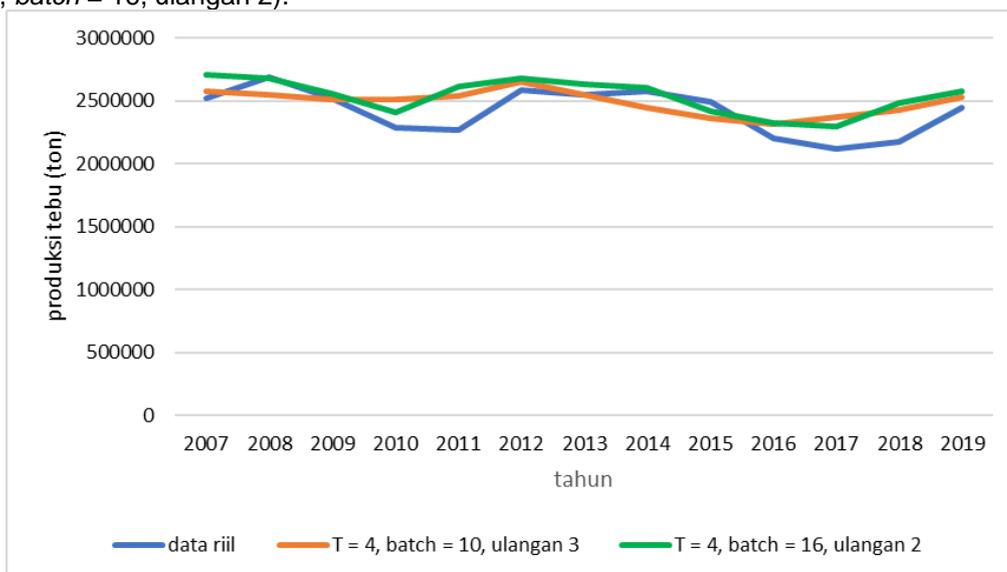
Ukuran kinerja model-model RNN yang dikembangkan, berupa RMSE, MAPE, MAE dan r disajikan pada Tabel 3. Berdasarkan Tabel 3, dapat dilihat bahwa model RNN untuk panjang periode sebesar 4 (T=4) memiliki RMSE, MAPE dan MAE yang lebih kecil dibandingkan dengan panjang periode yang lain. Hal ini terjadi baik pada ukuran *batch* 10 maupun 16. MAPE untuk T=4 mampu mencapai nilai kurang dari 10%, sementara untuk nilai T yang lain lebih dari 10%. Di samping itu terdapat kecenderungan, dengan meningkatnya nilai T, terjadi peningkatan nilai RMSE, MAPE dan MAE yang berarti terjadi penurunan kinerja. Untuk nilai korelasi, nilai tertinggi sebesar 0,77 dicapai pada saat T=4 (*batch* 16, ulangan ke-2), diikuti nilai korelasi 0.69 pada T=8 (*batch* 10, ulangan ke-1), nilai korelasi 0.68 pada T=6 (*batch* 10, ulangan ke-2) dan nilai 0.63 pada T=8 (*batch* 16, ulangan ke-2). Pada Tabel 3 juga dapat dilihat bahwa untuk nilai T=6, nilai RMSE, MAPE dan MAE nilainya relatif stabil untuk tiga ulangan baik untuk ukuran *batch* 10 maupun *batch* 16. Untuk T=8, ketiga ukuran kinerja RNN menghasilkan nilai-nilai yang sangat beragam (tidak stabil).

Tabel 3. Kinerja model-model RNN

Panjang periode (T)	Ukuran <i>batch</i> =10				Ukuran <i>batch</i> =16			
	RMSE	MAPE(%)	MAE	r	RMSE	MAPE(%)	MAE	r
4	250767.0	12.6	206622.0	-0.09	188841.3	9.7	158171.2	0.36
	189671.2	9.7	158335.2	0.33	174580.6	9.0	146709.2	0.77
	168402.1	8.8	142514.2	0.57	282973.7	15.0	246041.9	-0.41
6	255626.2	13.6	215207.0	0.21	252403.4	13.6	218206.4	-0.15
	229168.4	12.5	194500.5	0.68	255316.5	13.8	219496.3	0.20
	252326.4	13.4	211153.2	0.43	252864.9	12.8	207676.6	0.02
8	305059.3	16.7	249757.4	0.69	417609.7	22.9	346708.6	0.06
	367069.9	19.6	310749.6	-0.24	269656.3	15.2	231418.5	0.63
	419687.8	23.2	355422.7	-0.06	303687.0	17.2	263532.0	0.29

Sumber: Hasil Penelitian (2022)

Berdasarkan ukuran kinerja pada Tabel 3, terdapat dua kandidat model RNN terbaik yaitu model RNN dengan nilai MAPE 8.8% (T=4, *batch* = 10, ulangan 3) dan nilai MAPE 9.0% (T=4, *batch* = 16, ulangan 2).



Sumber: Kementan (2018)

Gambar 5. Hasil peramalan produksi tebu nasional 2007-2019

Hasil peramalan model RNN dengan MAPE sebesar 8.8% juga menghasilkan nilai-nilai RMSE, MAPE dan MAE yang terkecil namun demikian nilai korelasi yang dihasilkan hanya 0.57. Oleh karena itu, maka model RNN terbaik adalah model RNN untuk $T=4$, ukuran *batch* 16, ulangan ke-2 dengan nilai MAPE 9.0% dan korelasi (r) sebesar 0.77.

Pada Gambar 5 disajikan hasil peramalan dari model RNN terbaik ($T=4$, *batch*=16, ulangan 2) dan sebagai pembandingan disertakan juga hasil peramalan untuk model RNN dengan $T=4$, ukuran *batch* 10, dan ulangan 3. Dari Gambar 5 terlihat bahwa model RNN terbaik memiliki kesalahan yang relatif kecil dan lebih mendekati pola dari data asal. Untuk mengetahui lebih rinci hasil peramalan model RNN, pada Tabel 4 disajikan hasil peramalan model RNN beserta kesalahan (*error*) peramalannya. Secara umum, sebagian besar kesalahan bernilai kurang dari 10% dan hanya terdapat dua kesalahan peramalan yang cukup besar yang terjadi pada tahun 2011 (15.4%) dan tahun 2018 (14.2%).

Tabel 4. Hasil peramalan produksi tebu nasional tahun 2007-2019

Tahun	Produksi tebu nasional	Peramalan model RNN	Kesalahan	Kesalahan (%)
2007	2517374.0	2707317.5	189943.5	7.5
2008	2694227.0	2681954.0	-12273.0	-0.5
2009	2517374.0	2555035.0	37661.0	1.5
2010	2290116.0	2414181.8	124065.8	5.4
2011	2267887.0	2618124.8	350237.8	15.4
2012	2591687.0	2681850.8	90163.8	3.5
2013	2551026.0	2631331.8	80305.8	3.1
2014	2579173.0	2609537.8	30364.8	1.2
2015	2497997.0	2422530.8	-75466.2	-3.0
2016	2204619.0	2325155.8	120536.8	5.5
2017	2121671.0	2294764.2	173093.2	8.2
2018	2174400.0	2483119.0	308719.0	14.2
2019	2450000.0	2576927.2	126927.2	5.2

Sumber: Hasil Penelitian (2022)

Berdasarkan nilai MAPE dan nilai korelasi yang dihasilkan, model RNN yang telah dibangun memiliki akurasi tinggi dan korelasi yang kuat. Pada tahun 1982, Lewis membuat kriteria akurasi peramalan berdasarkan nilai MAPE [19]. Nilai MAPE kurang dari 10% menunjukkan peramalan dengan akurasi tinggi, nilai MAPE antara 10%-20% menunjukkan peramalan dengan kinerja yang baik, nilai MAPE antara 20%-50% menunjukkan kinerja peramalan masih dapat diterima, sedangkan nilai MAPE lebih dari 50% menunjukkan bahwa peramalan tersebut tidak akurat. Berdasarkan kriteria tersebut maka model RNN pada penelitian ini dapat dikatakan memiliki akurasi yang tinggi. Ditinjau dari nilai korelasi, model RNN pada penelitian ini menghasilkan korelasi sebesar 0.77. Berdasarkan [20], nilai RMSE 0.77 menunjukkan korelasi yang kuat (*strong correlation*). Hal ini berarti bahwa hasil peramalan model RNN yang telah dibangun, memiliki korelasi yang kuat dengan data tebu nasional.

Pada penelitian ini diperoleh kinerja model RNN terbaik menghasilkan MAPE sebesar 9.0%. Kinerja model RNN tersebut mendekati kinerja model ARIMA untuk wilayah Pakistan dengan nilai MAPE 8.92 % [7]. Data produksi tebu dalam penelitian [7] menunjukkan adanya pertambahan produksi tebu dengan pola yang linier. Kondisi ini berbeda dengan pola produksi tebu di Indonesia (Gambar 3). Pola produksi tebu di Indonesia menunjukkan peningkatan dari tahun 1967 sampai dengan tahun 1993 serta dari tahun 2000 sampai dengan tahun 2009. Pola penurunan produksi terjadi antara tahun 1993 sampai dengan tahun 2000 dan setelah tahun 2009. Dengan adanya pola tersebut, kemungkinan mengakibatkan nilai MAPE yang lebih besar dari nilai MAPE dengan pola produksi tebu yang linier.

Nilai MAPE model RNN pada penelitian ini lebih besar dibandingkan MAPE model RNN untuk gandum sebesar 5.28 % (Tabel 5). Terlebih lagi jika kinerja model ini dibandingkan dengan kinerja model LSTM untuk gandum dengan nilai MAPE sebesar 3.20%. Penelitian [12] menghasilkan MAPE yang kecil karena data gandum hanya memiliki pola tren dan menggunakan praproses berupa fungsi pemulusan (*smoothing function*) untuk mengatasi fluktuasi data. Fungsi pemulusan yang diterapkan adalah fungsi *robust-LOWESS* yang memberikan bobot rendah terhadap nilai pencilan (*outlier*). Fungsi tersebut menggunakan regresi linier lokal berupa *recursive least squares* (RLS) [21].

Tabel 5. Kinerja model prediksi penelitian-penelitian terkait

Wilayah	Komoditi	Periode data	Model	MAPE (%)	Sumber
Pakistan	Gandum	1902 – 2018	RNN	5.28	[12]
			LSTM	3.20	

Wilayah	Komoditi	Periode data	Model	MAPE (%)	Sumber
Pakistan	Tebu	1947 – 2018	ARIMA	8.92	[7]
Indonesia	Tebu	1967 - 2019	RNN	9.0	Penelitian ini

Sumber: Hasil Penelitian (2022)

Produksi tebu nasional merupakan gabungan dari produksi tebu dari PR, PBN dan PBS (gambar 3). Berdasarkan penelitian [22], produksi tebu dari PR semakin menurun karena berkompetisi dengan produksi gabah, sedangkan produksi tebu dari PBS menunjukkan tren semakin meningkat dan lebih efisien. Oleh karena itu, untuk penelitian berikutnya, peramalan dapat dilakukan terhadap ketiga komponen produksi tebu nasional tersebut secara terpisah. Berdasarkan pola produksi tebu pada Gambar 3, produksi tebu dari PR memiliki pola yang mirip dengan produksi tebu nasional. Produksi tebu dari PBS memiliki pola tren sehingga pola ini kemungkinan lebih mudah untuk dibuat modelnya seperti halnya penelitian [11]. Berbeda dengan produksi tebu dari kedua sumber tersebut, produksi tebu dari PBN memiliki pola yang cenderung stabil sejak tahun 1983. Dengan demikian, diharapkan model peramalan yang dikembangkan memiliki kinerja yang semakin baik.

4. Kesimpulan

Berdasarkan kriteria dari nilai MAPE dan nilai korelasi, model RNN yang telah dibangun memiliki kinerja yang baik. Dengan nilai MAPE sebesar 9%, model RNN yang dihasilkan dari penelitian ini termasuk model yang memiliki akurasi tinggi. Nilai MAPE 9% mengandung makna bahwa kesalahan prediksinya secara rata-rata sebesar 9% dari nilai yang sebenarnya. Korelasi model RNN sebesar 0.77 menunjukkan adanya korelasi yang kuat antara hasil peramalan dengan data sebenarnya. Namun demikian, berdasarkan perbandingan dengan hasil penelitian lain maka kinerja model RNN masih dapat ditingkatkan antara lain menggunakan praproses berupa pemulusan (*smoothing*) data untuk mengatasi fluktuasi data. Penelitian lebih lanjut pengembangan model-model RNN untuk peramalan produksi tebu nasional masih sangat diperlukan. Model RNN dengan nilai MAPE yang semakin kecil akan semakin mempermudah pengambil kebijakan dalam menentukan kebijakan terkait pemenuhan kebutuhan tebu nasional.

Referensi

- [1] Pusdatin, "Outlook Komoditi Tebu", Pusat Data dan Sistem Informasi Pertanian, Sekretariat Jendral Kementerian Pertanian, 2014.
- [2] BPS, "Statistik Tebu Indonesia", Badan Pusat Statistik, 2015.
- [3] Kementan, "Statistik perkebunan indonesia 2017 – 2019", 2018.
- [4] I.W. Kusuma, dan A. M. Abadi, "Aplikasi Model Backpropagation Neural Network Untuk Perkiraan Produksi Tebu Pada PT. Perkebunan Nusantara IX". *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika*, pp. 97-108, 2011.
- [5] PTPN IX, "Laporan Tahunan 2019", PT Perkebunan Nusantara IX. www.ptpnix.co.id, 2019.
- [6] K. K. Suresh and S. R. Krishna Priya, "Forecasting Sugarcane Yield of Tamilnadu Using ARIMA Models", *Sugar Tech*, vol. 13, no. 1, pp. 23–26, 2011.
- [7] Q. Mehmood, M. Sial, R.M. Maqbool, N. Shaheen, "Forecasting The Production of Sugarcane Crop of Pakistan for The Year 2018-2030, Using Box-Jenkin's Methodology, *The Journal of Animal & Plant Sciences*, vol. 29, no. 5, pp. 1396-1401, 2019.
- [8] P. Mishra, A. M. G. Al Khatib, I. Sardar, J. Mohammed, K. Karakaya, A. Dash, M. Ray, L. Narsimhaiah, A. Dubey, "Modeling and Forecasting of Sugarcane Production in India", *Sugar Tech*, vol. 23, no. 6, pp. 1317-1324, 2021.
- [9] L. Fausset, *Fundamentals of neural networks*. Prentice Hall, Englewood Cliffs, 1994.
- [10] Z.C. Lipton, J. Berkowitz, C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequential Learning", <https://arXiv:1506.00019>, 2015.
- [11] V. Krakovna, and F. Doshi-Velez, "Increasing the interpretability of recurrent neural networks using hidden markov models", arXiv preprint arXiv:1606.05320, 2016.
- [12] S.A. Haider, S.R. Naqvi, T. Akram, G.A. Umar, A. Shahzad A, M.R. Sial, S. Khaliq, and M. Kamran, "RNN Neural Network Based Forecasting Model for Wheat Production in Pakistan". *Agronomy* vol. 9, no. 72, 2019.
- [13] A. S. Terliksiz, and D.T. Altylar, "Use Of Deep Neural Networks For Crop Yield Prediction: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA". 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), 2019.

- [14] A. Kaneko, T. Kennedy, L. Mei, C. Sintek, M. Burke, S. Ermon, D. Lobell, "Deep Learning For Crop Yield Prediction in Africa". *International Conference on Machine Learning: AI for Social Good Workshop*, 2019.
- [15] M.A. Rebortera and A.C. Fajardo, "An Enhanced Deep Learning Approach in Forecasting Banana Harvest Yields". *IJACSA* vol. 10, no. 9, 2019.
- [16] S. Siami-Namini, N. Tavakoli and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 1394-1401, 2018.
- [17] M. Elsarati, and A. Merabet, "A Comparative Analysis of the ARIMA and LSTM Predictive Models and Their Effectiveness for Predicting Wind Speed", *Energies*, vol.14, no. 6782, pp. 1-16, 2021.
- [18] R.E. Ronald, "Pengantar Statistika", Jakarta : PT. Gramedia Pustaka Utama, 1995.
- [19] H.K. Sharma, K Kumari, S Kar, "Forecasting Sugarcane Yield of India Based on Rough Set Combination Approach", *Decision Making: Applications in Management and Engineering* vol. 4, no. 2, pp. 163-177, 2021.
- [20] P. Schober, C. Boer, L.A. Schwarte, "Correlation Coefficients: Appropriate Use and Interpretation", *Anesthesia & Analgesia*, vol. 126, no. 5, pp. 1763-1768, 2018.
- [21] W.S. Cleveland, "Lowess: A program for smoothing scatterplots by robust locally weighted regression". *Am. Stat.* vol. 35, no 1, pp. 54, 1981.