Nano World Journal

https://doi.org/10.17756/nwj.2022-s1-022

Review Article

Open Access

Comparative Analysis of Prediction Models for Melt Flow Rate of C2 and C3 Polymers Synthesized using Nanocatalysts

Devanshi Ledwani[1], Ishika Thakur[1], and Vaibhav Bhatnagar[2*]

[1]*Department of Computer and Communication Engineering, Manipal University Jaipur, India*

[2]*Department of Computer Applications, Manipal University Jaipur, India*

***Correspondence to:**
Vaibhav Bhatnagar
Department of Computer Applications, Manipal University Jaipur, India
E-mail: Vaibhav.bhatnagar15@gmail.com

Received: August 23, 2022
Accepted: October 27, 2022
Published: October 29, 2022

Citation: Ledwani D, Thakur I, Bhatnagar V. 2022. Comparative Analysis of Prediction Models for Melt Flow Rate of C2 and C3 Polymers Synthesized using Nanocatalysts. *NanoWorld J* 8(S1): S123-S127.

Published by United Scientific Group

Abstract

Analysis of product quality is extremely important in any industry. The polymer industry is gaining significant importance in recent years. C2 and C3 are polymers that are made up of monomeric units with two and three carbon atoms, respectively. Nanotechnology plays an important role in polymerization reactions which subsequently improves the quality and quantity of polymer products. The importance of measuring product quality in online polymerization industries poses significant difficulties. There are no online tools for measuring resin features that represent polymer quality, such as flow melting indicator (melt flow index; MFI) and congestion. MFI should invariably be tested in time-consuming and labor-intensive lab analysis. In many processing units, MFI is measured occasionally rather than regularly during a day using manual analysis. This paper presents models to predict melt flow rate using models of Linear regression, Support Vector Regression (SVR) and Adaboost Regressor. Adaboost regressor fittest with the accuracy of 69.2%. The primary objective of this research was to identify the best machine learning (ML) algorithm that can accurately predict the melt flow rate.

Keywords

Nanocatalyst, Regression, Model, Support vector regression

Introduction

A polymer is nothing but a class of natural or synthetic material made up of very large size molecules, known as macromolecules, which are duplicates of simple chemical units called monomers [1]. Polymers (polyethylene, polypropylene, polystyrene, etc.) are made from chemical building blocks called monomers (C2, C3, C4, C5, and others) in gas phase and liquid slurry reactors. The use of nanoparticles in polymerization reactions can significantly improve the product output. Nanocatalysis is a rapidly growing field that involves the use of nanomaterials as catalysts for various homogeneous and heterogeneous catalytic applications. Heterogeneous catalysis represents one of the oldest commercial practices of nanoscience; metal nanoparticles, semiconductors [2], oxides, and other compounds have been widely used for important chemical reactions. The monomers are introduced into a catalyst under precisely regulated pressure and temperature conditions to start a reaction that develops the polymer chains. To inhibit the expansion of the polymer chain, hydrogen acts as a chain transfer agent [3]. The polymer becomes excessively viscous for manufacture in films, injection molding, or other applications if the polymer chains lengthen too much [4]. The polymer is soft and lacks the strength needed for the specific application, such as a plastic bag, a vehicle bumper, or a washing machine drum, if the polymer chains are too short. To maintain the polymer at the proper viscosity for each

specific grade, routine lab samples taken from the reactor are employed [5].

MFI is a measure of the melting point of a thermoplastic polymer [6]. This can be understood as the polymer weight measured in grams flowing for 10 minutes with a particular length and a specific diameter with applied pressure at a certain temperature to the given weight. This technology is generally applied in most plastic industries to control the quality of thermoplastic. MFI indicates the flowability of the thermoplastic materials. This is a well-known melting property of thermoplastic material checked by equipment called a melt flow tester Figure 1.

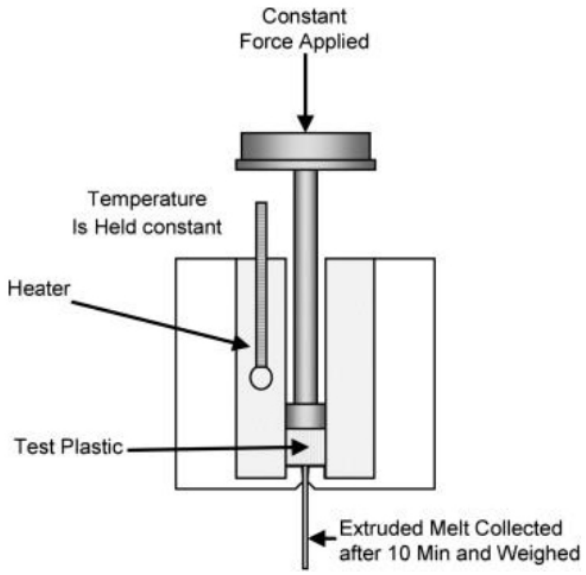

Figure 1: Melt flow index test apparatus.

The conditions of the test depend upon the type of the polymer used. One does not want a temperature so high that the polymers in the plastic decompose. The melt flow rate is an indirect measure of molecular weight, with high melt flow rate corresponding to low molecular weight. Synonyms of MFI are melted flow rate and melt index, which are commonly abbreviated as: MFR and MI, respectively. When the volume of the extrudate is measured, the melt volume rate is reported.

The objective of this paper is to identify the best supervised ML algorithm that can predict the melt flow rate of C2 and C3 polymers. In this paper regression, SVR and Adaboost algorithms are compared on the dataset.

The organization of this article is as follows: after the Introduction, second part is related work, third part is data collection, fourth part is data description, fifth part is implementation of ML algorithms and last sixth part is conclusion and future work.

Related Work

In this section, state of the art literature is presented where different ML algorithms are applied in different datasets. These algorithms are long short-term memory (LSTM), extreme gradient boosting (XGBoost), Support Vector Machine (SVM), etc.

A succinct account of the rheological and thermal impacts of the flow through the hot end in material extrusion AM is provided by Marcin P. Serdeczny et al. in the year 2020 [7]. An experiment was conducted to examine the effects of the liquefier length, temperature, liquefier length, and filament feeding rate on the extrudate swell and filament feeding force.

Ziyang Zhang et al. in the year 2021, devised a ML based data driven predictive model that helps calculate the accurate melt pool temperature during DED [8]. The predictive model was developed using XGBoost and LSTM. The experiment concludes that both the ML algorithms are one of the most optimal choices for predicting the melt pool temperature with higher accuracy. The investigation witnessed that LSTM is less efficient then XGBoost but gains higher prediction accuracy and is more robust then XGBoost.

Gholam Hossain Varshouee et al. in the year 2019, the main motive of this study was to identify a mathematical model based on the method of estimating population and determine the response of the reaction temperature and the amount of hydrogen on the final storage components of the product which includes molecular weight measurement and polydispersity index as well as the flow profile of polypropylene and profile degree of polymer and to find the best working condition using the model verified test data [9]. The paper concludes that increasing the reaction temperature up till a certain limit was useful and improving other indicators for the final product and subsequently an increase in reaction temperature has a detrimental effect on the indices.

This paper's major goal was to investigate polyethylene's rheological behavior [10]. Analysis of the cellular weight data for three varieties of polyethylene, including HDPE, LLDPE, and LDPE. In this investigation, a melting flow indicator and capillary rheometer were employed to determine the rheological behavior of the polymer. These three distinct polyethylene polymers' viscosities were compared to the sheer size. It was found that shear stress increased along with rising shear levels. Additionally, it was discovered that when the shear rate increased, the viscosity of HDPE, LLDPE, and LDPE decreased. The weight of the cells and their distribution have a significant impact on the rheological behavior of polymers, as well as the number of extensions with reduced viscosity.

Amulya K. Pervaje in the year 2022 [11], the multiscale approach was presented in this paper to develop a basic model for predicting the mechanical properties of polypropylene fibers based on chemical and physical propertiesc. The modeling of this method was based on validating the imperial stress-strain curves from nine isotactic polypropylene fibers with multi-molecular weight characteristics, Hermans shape features, and crystallinity.

Combined cell models are constructed using the simulation of molecular dynamics (MD) by atomic models. VehbiEmrahAtasoy in the year 2022, further investigated the relationship between greenhouse gas emissions (EGT) and the different performance parameters of the aeroengine of the navigation phase [12]. EGT predictions were made with a variety of models, including deep reading (DL) and SVM, using a set of flight data history, over 1300. To achieve this

goal, EGT was considered an output parameter. Input parameters were taken as the most important variables for predicting EGT.

Souta Miyamoto investigated [13] the effectiveness of a ML approach to learn the constitutive relationship of well-connected polymer melts. Specifically, we use Gaussian process regression on training data (e.g., stress, strain rate, and number of entanglements) obtained from a microscopic Slip-link simulation under various flow conditions. The learned constitutive relations (stated in differential form) were then used in macroscopic simulations of the smooth particle hydrodynamics of standard flows (i.e., parallel plates and contraction expansion).

Alesadi et al. [14] presented an integrated framework to predict the glass transition temperature of conjugated polymers having diverse chemistry through the integration of ML, MD simulations, and experiments. The predictive model takes simplified "geometry" of six key chemical building blocks as molecular features to predict glass transition temperature.

Data Collection

The process of information collection and information measuring for the variable of interest is known as data collection [15]. The data must be collected in a systematic way which helps to answer the research question, evaluate the idea and results analysis The meta data was shown in Table 1 from where the data was used [16].

Data Description

The process of data mining and data analysis is a step of pre-data processing. The data preprocessing takes raw data as input and transforms into a well understood form that can be analyzed by the computer system [17].

Firstly, the most important step is data cleaning which involves removing NA values wherever they exist. Many ML algorithms fail if the dataset contains missing values. After null values are replaced with mean of the dataset next step is a Descriptive Analysis of data which was done with the help of the following mathematical methods [18]. The probability density function of different input variables was shown below (Figure 2).

Correlation is often used in ML to identify multicollinearity, which occurs when two or more predictor variables are highly correlated with each other. Multicollinearity can affect the accuracy of predictive models, so it's important to identify and eliminate predictors that are associated with high

Table 1: Parameter analysis to predict Melt Flow Rate of Polymer.

S. N.	Labels	Description	Data Type	Role
1	Time	Timestamp of the measurements	Time Series	Ignored in the Model
2	Temp	Reactor Temperature	Integer	Input Variable
3	Cat	Catalyst Feed Rate (kg/hr)	Integer	Input Variable
4	Level	Reactor Bed Level (m)	Integer	Input Variable
5	H2R	Hydrogen to C3= Ratio	Integer	Input Variable
6	C2=	Ethylene (C2=) Flow (kg/hr)	Integer	Input Variable
8	Pressure	Reactor Pressure (bar)	Integer	Input Variable
9	C3=	Propylene (C3=) Feed Rate (kg/hr)	Integer	Input Variable
10	MFR	Melt Flow Rate (gm/10min)	Integer	Output Variable

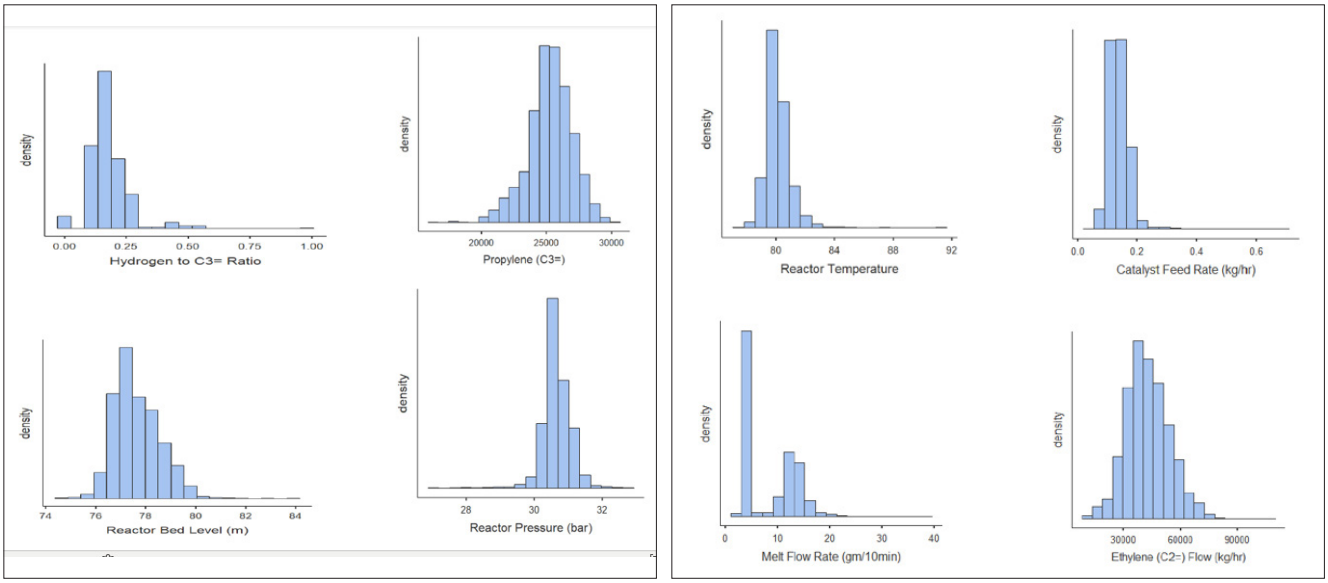

Figure 2: Probability density function of inputs of MFR value.

multicollinearity. Correlation heat maps are a type of plot that visualize the strength of relationships between numerical variables. A correlation plot contains several numerical variables, with each variable represented by a column. The rows represent the relationship between each pair of variables. The values in the cells specify the strength of the relationship, with positive values indicating a positive relationship and negative values indicating a negative relationship. Correlation heatmaps can be used to find potential relationships between variables and to understand the strength of these relationships. In addition, they can be used to identify outliers and to detect linear and nonlinear relationships. The color-coding of the cells makes it easy to identify relationships between variables at a glance. Correlation heatmaps can be used to find both linear and nonlinear relationships between variables [19]. The above heat map gives the matrix of correlation of different variables. The values of correlation in the correlation matrix can take up any values between -1 and 1. Correlation among different variables was shown in Figure 3 and the descriptive analysis of different variables in Table 2.

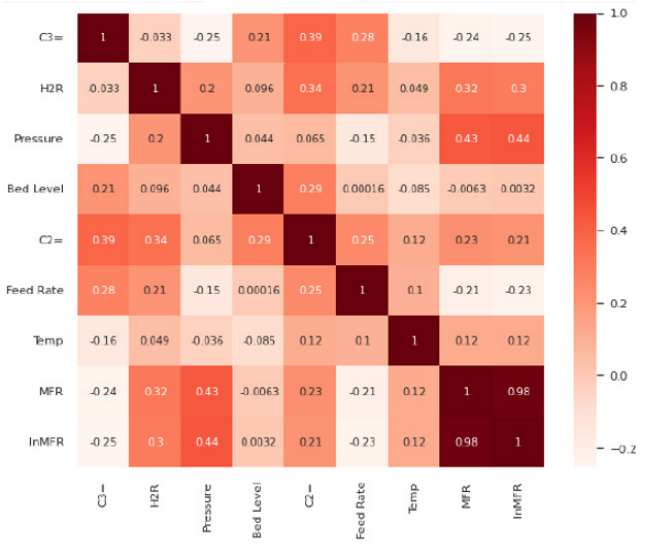

Figure 3: Heat map.

Table 2: Descriptive analysis of parameters.

S. N.	Name of Variable	Mean	Standard Error Mean	Standard Deviation
1	C3=	25306	34.2	1706
2	H2R	0.178	0.00153	0.0775
3	Pressure	30.7	0.008	0.423
4	Level	77.7	0.0185	0.920
5	C2=	42525	227	11332
6	Cat	0.139	8.4^{e-4}	0.0419
7	Temp	80.1	0.0165	0.82
8	MFR	8.19	0.100	5.09

Implementation of Machine Learning Algorithms

Python 3 version was used to identify the threshold value and for applying different ML algorithms.

Linear regression

It predicts the amount of variability based on the other variables value. The two variables in linear regression are the dependent variable which will be predicted, and another is the independent variable which helps to get the predicted value for some variables. In this study the MFR was the dependent variable. In this dataset, linear regression was applied. The dependent variable was MFR and other variables were taken as independent variables [20].

From Table 3, the RMSE value obtained was 3.85 which was satisfactory in nature. The P value of the overall model obtained was below 0.001 which was less than 0.05, hence there is a significant relationship between input all input variables and output variable that is MFR [21]

Support vector regression

SVR is one of the supervised learning algorithms that predict different values [22]. SVR uses similar principles as SVMs. SVR's focus is to find the relevant line. Hyperplane is the most appropriate line with a greater number of points. An approach for supervised learning called SVR is used to forecast discrete values. The SVMs and SVR both operate on the same theory. Finding the optimum fit line is the fundamental tenet of SVR. The hyperplane with the most points is the best-fitting line in SVR. The SVR seeks to match the best line within a threshold value, in contrast to other regression models that aim to reduce the error between the real and projected value. The distance between the boundary line and the hyperplane is the threshold value. SVR is difficult to scale to datasets with more than a few ten thousand samples since the fit time complexity is more than quadratic with the number of samples [23].

AdaBoost Regressor: The AdaBoost algorithm, short for Adaptive Boosting, is a boosting technique used as an Ensemble Method in ML. This is called adaptive boosting because weights are reassigned to each instance, with higher weights assigned to misclassified instances. Boosting is used to reduce bias as well as variance for supervised learning. It works on the principle of gradual growth of students. Except for the first, each subsequent student grew out of previously mature students. Simply put, weak students are transformed into strong ones. It is a technique in ML used as an Ensemble Method [22]. The most common algorithm used with AdaBoost is decision trees with one level that means with decision trees with only 1 split. These trees are also called Decision Stumps [24]. A meta-estimator called AdaBoost Regressor, starts by fitting

Table 3: Model fit measures.

						Overall Model Test			
Model	R	R^2	Adjusted R^2	RMSE	F	df1	df2	p
1	0.651	0.423	0.422	3.85	260	7	2476	< .001

a regressor on the original dataset, and then it fits subsequent copies of the regressor on the same dataset with the weights of the instances being changed in accordance with the error of the most recent prediction.

The efficiency of above-mentioned algorithms was shown in Table 4. Adaboost performs better than the Linear Regression and SVR.

Table 4: Machine learning algorithm efficiency analysis.

S. N.	Name of Algorithm	Efficiency Received (%)
1	Linear Regression	50.2
2	Support Vector Regression	28.0
3	Adaboost	69.2

Conclusion and Future Work

The paper provides a brief introduction and Prediction Models for MFR of C2 and C3 polymers. After a deep analysis Gradient Boost algorithm (AdaBoost) will be the best fit to classify the polymers according to MFR. (Accuracy= 69.2%). The observation concludes that the polymer was more viscous for manufacturing injections, molding films, and many other applications. This is due to the increasing growth of the polymer chain. Another observation states that the polymer does not have the strength for some specific application such as plastic bags, a car bumper, or washing machine drums. This is due to the short polymer chains and will make the polymer softer. In the future work, the efficiency of the ML model can be enhanced by fine tuning the different parameters.

References

1. Verbeeten WMH, Peters GWM, Baaijens FPT. 2002. Viscoelastic analysis of complex polymer melt flows using the extended Pom–Pom model. *J Nonnewton Fluid Mech* 108(1-3): 301-326. https://doi.org/10.1016/S0377-0257(02)00136-2
2. Rao NP. 2010. Nanocatalysis: applications in the chemical industry. [https://www.nanowerk.com/spotlight/spotid=18846.php] Accessed on November 30, 2022
3. Park TC, Kim TY, Yeo YK. 2010. Prediction of the melt flow index using partial least squares and support vector regression in high-density polyethylene (HDPE) process. *Korean J Chem Eng* 27(6): 1662-1668. https://doi.org/10.1007/s11814-010-0280-x
4. Zhu CH, Zhang J. 2020. Developing soft sensors for polymer melt index in an industrial polymerization process using deep belief networks. *Int J Autom Comput* 17(1): 44-54. https://doi.org/10.1007/s11633-019-1203-x
5. Kim TY, Yeo YK. 2010. Development of polyethylene melt index inferential model. *Korean J Chem Eng* 27(6): 1669-1674. https://doi.org/10.1007/s11814-010-0290-8
6. Wang S, Capoen L, D'hooge DR, Cardon L. 2018. Can the melt flow index be used to predict the success of fused deposition modeling of commercial poly (lactic acid) filaments into 3D printed materials?. *Plastics, Rubber and Composites* 47(1): 9-16. https://doi.org/10.1080/14658011.2017.1397308
7. Serdeczny MP, Comminal R, Pedersen DB, Spangenberg J. 2020. Experimental and analytical study of the polymer melt flow through the hot-end in material extrusion additive manufacturing. *Addit Manuf* 32: 100997. https://doi.org/10.1016/j.addma.2019.100997
8. Zhang Z, Liu Z, Wu D. 2021. Prediction of melt pool temperature in directed energy deposition using machine learning. *Addit Manuf* 37: 101692. https://doi.org/10.1016/j.addma.2020.101692
9. Varshouee GH, Heydarinasab A, Vaziri A, Roozbahani B. 2019. A mathematical model for determining the best process conditions for average molecular weight and melt flow index of polypropylene. *Bull Chem Soc Ethiop* 33(1): 169-182. https://doi.org/10.4314/bcse.v33i1.17
10. Escócio VA, Vasques Pacheco EBA, da Silva ALN, de Paula Cavalcante A, Visconte LLY. 2015. Rheological behavior of renewable polyethylene (HDPE) composites and sponge gourd (*Luffa cylindrica*) residue. *Int J Polym Sci* 2015: 714352. https://doi.org/10.1155/2015/714352
11. Pervaje AK, Pasquinelli MA, Khan SA, Santiso EE. 2022. Multiscale constitutive modeling of the mechanical properties of polypropylene fibers from molecular simulation data. *Macromolecules* 55(3): 728-744. https://doi.org/10.1021/acs.macromol.1c00630
12. Atasoy VE, Suzer AE, Ekici S. 2022. A comparative analysis of exhaust gas temperature based on machine learning models for aviation applications. *J Energy Resour Technol* 144(8): 082101. https://doi.org/10.1115/1.4052771
13. Miyamoto S, Molina J, Taniguchi T. 2022. Application of machine-learned constitutive relations for well-entangled polymer melt flows. [https://ui.adsabs.harvard.edu/abs/2022APS..MARW16011M/abstract] Accessed on November 30, 2022
14. Alesadi A, Cao Z, Li Z, Zhang S, Zhao H, et al. 2022. Machine learning prediction of glass transition temperature of conjugated polymers from chemical structure. *Cell Reports Physical Science* 3(6): 100911. https://doi.org/10.1016/j.xcrp.2022.100911
15. Responsible Conduct in data management. [https://ori.hhs.gov/education/products/n_illinois_u/datamanagement/dctopic.html] Accessed on November 30, 2022
16. Polymer melt flow rate. [https://apmonitor.com/pds/index.php/Main/PolymerMeltFlowRate] Accessed on November 30, 2022
17. Mishra P, Biancolillo A, Roger JM, Marini F, Rutledge DN. 2020. New data preprocessing trends based on ensemble of multiple preprocessing techniques. *Trends Analyt Chem* 132: 116045. https://doi.org/10.1016/j.trac.2020.116045
18. Dean S, Illowsky B. 2018. Descriptive statistics: skewness and the mean, median, and mode. Rice University, Texas, USA.
19. McKeen L. 2012. Introduction to the physical, mechanical, and thermal properties of plastics and elastomers. In: McKeen L (ed) The effect of sterilization on plastics and elastomers. William Andrew Publishing, pp 57-84. https://doi.org/10.1016/B978-1-4557-2598-4.00003-4
20. Yao W, Li L. 2014. A new regression model: modal linear regression. *Scandinavian Journal of Statistics* 41(3): 656-671.
21. Unlocking the true power of support vector regression. [https://towardsdatascience.com/unlocking-the-true-power-of-support-vector-regression-847fd123a4a0] Accessed on November 30, 2022
22. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf AJ. 2020. A systematic review on supervised and unsupervised machine learning algorithms for data science. In: Berry MW, Mohamed A, Yap BW (eds) Supervised and unsupervised learning for data science. Springer, Cham, pp 3-21. https://doi.org/10.1007/978-3-030-22475-2_1
23. Zhang Z, Liu Z, Wu D. 2021. Prediction of melt pool temperature in directed energy deposition using machine learning. *Addit Manuf* 37: 101692. https://doi.org/10.1016/j.addma.2020.101692
24. AdaBoost algorithm - a complete guide for beginners - analytics vidhya. [https://www.analyticsvidhya.com/blog/2021/09/adaboost-algorithm-a-complete-guide-for-beginners/] Accessed on November 29, 2022