

CHALLENGES AND PROBLEMS IN TODAY'S ECONOMIC FORECASTS

Georgi Georgiev, Vladislava Georgieva

University of agribusiness and rural development, Plovdiv, Bulgaria

Abstract: The purpose of the publication is to present to the academic community and practitioners in our country the achieved forecast accuracy in the economic field by applying various modern methods. The main challenges and problems related to time series forecasting and the adequacy of the forecasting models are considered.

Keywords: time series forecasting, symmetric mean absolute percentage error, machine learning, forecasting model adequacy, overfitting, underfitting.

ПРЕДИЗВИКАТЕЛСТВА И ПРОБЛЕМИ ПРИ СЪВРЕМЕННИТЕ ИКОНОМИЧЕСКИТЕ ПРОГНОЗИ

Георги Георгиев, Владислава Георгиева

Висше училище по агробизнес и развитие на регионите, Пловдив, България

Резюме: Целта на настоящата публикация е да запознае академичната общественост и практикуващите прогностици у нас с постиганата прогнозна точност при динамични редове в икономическата област чрез прилагане на различните съвременни методи. Разгледани са основните предизвикателствата и проблеми, свързани с прогнозиране на ключови икономически променливи и адекватността на прилаганите прогностични модели.

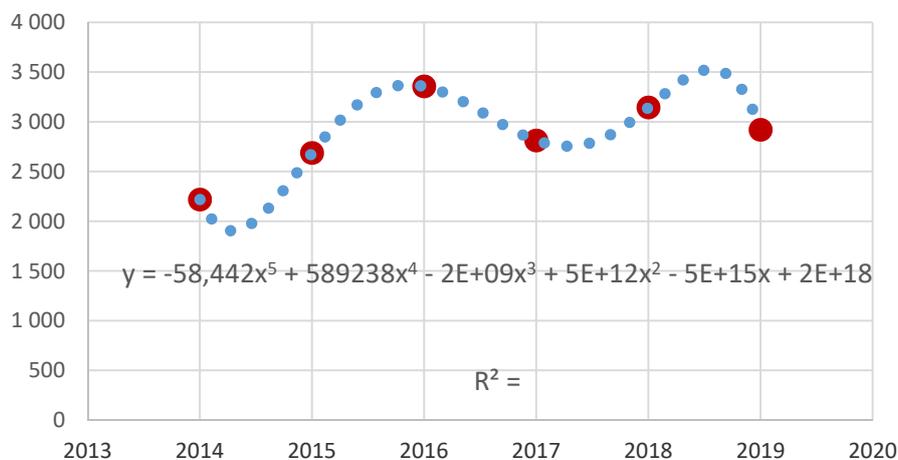
Ключови думи: прогнозиране на динамични редове, симетрична средна абсолютна процентна грешка, машинно обучение, адекватност на прогностичен модел.

Обикновено, прогнозите на определени бизнес променливи се основават на исторически данни, които трябва да отговарят на определени количествени и качествени критерии. Тези критерии са специфични и зависят от избрания метод за прогнозиране. В нашето съвремие използваните методи за прогнозиране от гледна точка на използваната технология могат да бъдат разграничени на три отделни групи: статистически, методи за машинно обучение и комбинирани.

Недостигът на достатъчно количество емпирични данни би могло да доведе до избор на подвеждащи стохастични модели или такива за машинно обучение, при които се наблюдава ниска прогнозна точност. В статистическата теория и практика проблемите, свързани с избора на подходящ модел за прогнозиране се асоциират с понятията „пренапасване“ (**overfitting**) и „неадекватност на модела“ (**underfitting**).

„Пренапасването“ представлява прогнозна грешка при избора на математически или стохастичен модел, която е предизвикана от това, че дадена избрана функция твърде плътно пасва към ограничена извадка от данни, не кореспондира достатъчно близо спрямо други реални данни и по този начин прогнозите са ненадеждни. „Пренапасването“ на модела обикновено е под формата на създаване на твърде сложен модел, който включва освен тренда и случаен компонент („бял“ шум). Графика 1 илюстрира визуално „пренапасване“ на модела спрямо емпиричните данни. В конкретния случай е избран полином от 5 степен при 6

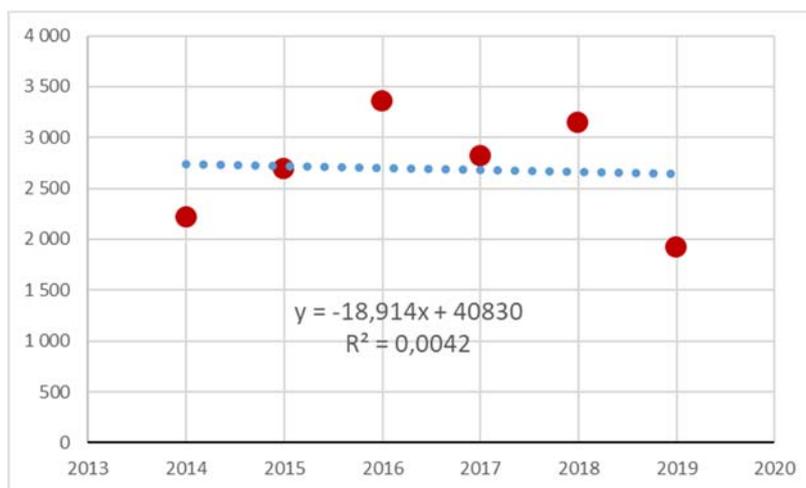
наблюдения и въпреки, че коефициента на детерминация е максимално висок (единица) прогнозната сила на модела е слаба.



Графика 1. „Пренапасване“ на модел

В действителност реалните данни съдържат освен определен модел на поведение на изследваното явление и „бял“ шум. „Белият шум“ представлява случайно вариране, което в моделите за прогнозиране е известно като остатъчна дисперсия или случайна грешка. Стремещт да се направи модела да пасва твърде точно спрямо реални данни може да зарази модела със значителни грешки и да намали неговата прогнозна сила¹.

Неадекватност на прогнозния модел (underfitting) се наблюдава, когато той не отразява коректно основната информационна структура на емпиричните данни. Моделът е неадекватен, ако в неговата рамка са пропуснати (липсват) някои важни параметри (независими променливи)². Типичен пример за неадекватен модел е използване на линеен модел за прогнозиране при нелинейни данни. Графика 2 илюстрира именно такъв случай. Видно от графиката е, че линейния модел не пасва на емпиричните данни и това се потвърждава от коефициента на детерминация, чиято стойност е 0,0042.



Графика 2. Неадекватен модел (underfitting)

¹ Прогнозната грешка произтича от наличието на бял шум и недостатъчната адекватност на стохастичния или машинно самообучаващ се модел.

² Everitt B.S., Skrondal A. Cambridge Dictionary of Statistics, Cambridge University Press 2010.

По аналогия на приетите дефиниции в областта на машинното самообучение (Machine learning) адекватността на моделите може да бъде дефинирана по определени критерии. Таблица 1 представя подобни критерии, които са често използвани при прогнозиране чрез машинно самообучение.

Таблица 1. Критерии за адекватност на моделите

адекватност на модела	критерий
пренапасване (overfitting)	$E_C < E_V$
недостатъчно адекватен модел (underfitting)	$E_C > E_V$
адекватен модел	$E_C \cong E_V$

където E_C е прогнозната грешка при конструкционната извадка, а E_V е прогнозна грешка на контролната извадка.

Във всички случаи е важно да се тества прогнозната сила модела върху емпирични данни, които са извън извадката, използвана за неговото конструиране.

Съвременният арсенал от подходи и методи за прогнозиране на икономически променливи е изключително разнообразен и се развива и дообогатяване, модифицира и усъвършенства непрекъснато. Какъв конкретен метод в рамките на определен подход ще използва даден мениджър зависи от голям брой фактори:

- количество и качество на емпиричните данни;
- данни (статични или времеви редове, сезонен компонент, непрекъснати или дискретни)
- компетентност и умения на мениджъра;
- предназначение и точност на прогнозите;
- прогнозен хоризонт;
- времето и наличните ресурси и др.

Кои от огромното разнообразие от разработени методи осигуряват най-висока точност при икономическите прогнози е въпрос, които към настоящия момент няма еднозначен отговор, но въпреки това благодарение на провежданите през последните четири десетилетия Makridakis състезания (известни още като М-състезания) могат да бъдат очертани група от определени прогнозни методи, които постигат по-висока точност в реални условия.

М-състезанията представляват поредица от открити състезания за научни работници и професионалисти с награден фонд, организирани от екипи на University of Nicosia, New York University, Department of Finance and Risk Engineering to Tandon School of Engineering, School of Electrical and Computer Engineering of the National Technical University of Athens ръководени от професор Spyros Makridakis с цел да оценят и сравнят точността на различните методи за прогнозиране. Четири открити състезания са проведени до сега (М-1 през 1982 г, М-2 през 1993, М-3 през 2000 г и последното М-4 през 2019 г.)³. В последното състезание М-4 за оценяване и сравняване прогнозната сила на моделите са използвани 100 000 динамични редове от емпирични данни от различни сектори на икономиката. Таблица 2 дава представа за количеството използвани динамични редове от различни сектори на реалната икономика.

³ www.mcompetitions.unic.ac.cy

Таблица 2. Използвани динамични редове на различна времева база по икономически сектори (в хил.)⁴

времеви интервал	Данни по икономически сектори						
	микроекономика	промишленост	макроекономика	финанси	демография	други	общо
години	6,538	3,716	3,903	6,519	1,088	1,236	23
3 месеция	6,02	4,637	5,315	5,305	1,858	865	24
месеци	10,975	10,017	10,016	10,987	5,728	277	48
седмици	112	6	41	164	24	12	359
дни	1,476	422	127	1,559	10	633	4,227
часове	0	0	0	0	0	414	414
общо	25,121	18,798	19,402	24,534	8,708	3,437	100

Както се вижда от таблица 2 за оценяване прогнозната точност на моделите на годишна база от сектор финанси са използвани 6 519, от промишленост 3 716 и т.н. т.е. огромен емпиричен материал за прецизно оценяване.

Оценяваните модели са класифицирани в две основни групи: статистически (стохастични) и машинно самообучение (Machine learning). Тяхната прогнозна точност е оценявана чрез следните показатели: симетрична средна абсолютна процентна грешка (Symmetric mean absolute percentage error – SMAPE), средна симетрична абсолютна процентна грешка (median symmetric absolute percentage error), Средна относителна абсолютна грешка (Median relative absolute error) и др.

Фигура 1 показва конкретните резултати от оценяване прогнозната точност на моделите от двете групи чрез симетрична средна абсолютна процентна грешка (SMAPE).

Симетрична средна абсолютна процентна грешка е изчислена по следната формула:

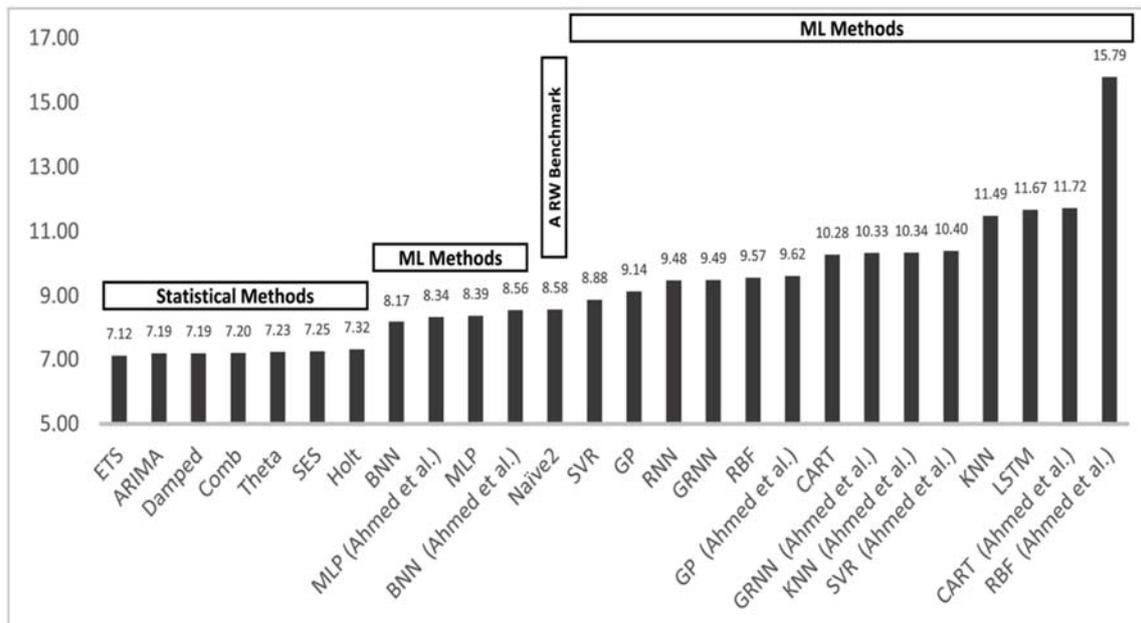
$$SMAPE = \sum \frac{|X-F|}{(X+F)/2} * 100,$$

където X е реалната стойност, а F е прогнозната.

Колкото по ниска е стойността на симетричната средна абсолютна процентна грешка, толкова по-точен е модела. Както се вижда от фигура 1 статистическите модели превъзхождат тези от групата на машинното самообучение⁵.

⁴ Makridakis Spyros, The Contributions of the M4 Competition to the Theory and Practice of Forecasting, 2019

⁵ Първите 7 най-точни модела са статистически на фигура 2.1.



Фигура 1. Оценяване прогнозната точност на моделите чрез SMAPE⁶

Представителите на бизнеса и академичните среди, които са участвали в последното състезание M-4 са използвали следните статистически модели:

- Naive 2 -случайно движение (random walk adjusted for seasonality)
- SES просто експоненциално изглаждане (simple exponential smoothing) без тренд
- Holt - експоненциално изглаждане на Holt
- Damped експоненциално изглаждане с намаляващ тренд
- Theta method
- ARIMA
- ETS (simple exponential smoothing with additive errors)
- Comb – комбинация (средна аритметична) между SES, Holt и Damped

Таблицы 3 и фигура 1, ни позволяват да оценим точността, постигната от всеки метод в множество хоризонти, и да вземем решение за тяхната целесъобразност за различни приложения.

Резултатите от всеки от двата метода за прогнозиране (ML и статистически) за хоризонт от 18 месеца са показани в таблица 2.3. За да се опрости представянето, резултатите са групирани в три прогнозни хоризонта: краткосрочен (от 1 до 6 месеца напред), средносрочен (7 до 12 месеца) и дългосрочен (13 до 18 месеца).

ML моделите са нелинейни функции, свързващи входовете и изходите на изкуствените неврони. Целта на мрежата е да се „научи“ да прогнозира чрез оптимизация т.е. избере набор от параметри или тегла, които минимизират прогнозната грешка. Същият тип оптимизация се прави и при моделите ARIMA (или при регресия). Следователно няма очевидна причина да се оправдае повече от 1,24% по-висок sMAPE на MLP, един от най-добрите методи на ML, в сравнение с този на ARIMA, или че sMAPE на този MLP е само 0,19% по-точен от Naive 2 (сезонно коригираният модел на случайно движение). Таблица 4 два представя за постигната средна точност при осемте представени в състезанието статистически модела.

⁶ <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194889>

Таблица 3. Оценяване прогнозната ефективността на ML и статистическите методи при различни времеви хоризонти чрез SMAPE⁷

Method	Short	Medium	Long	Average	CC
MLP Iterative	9.53	12.34	15.00	12.29	245.58
MLP Direct	10.72	13.55	16.20	13.49	438.53
MLP Multi	9.53	12.69	16.08	12.77	4006.82
BNN Iterative	9.39	12.08	14.80	12.09	141.91
BNN Multi	9.48	12.70	15.96	12.71	2046.49
Naive 2	10.78	12.46	15.08	12.77	1.48
SES	9.17	10.85	13.77	11.26	1.60
Holt	9.07	11.18	14.29	11.51	1.75
Damped	8.96	10.63	13.46	11.02	2.07
Comb	8.95	10.57	13.38	10.97	2.65
Theta	8.96	10.53	13.19	10.89	1.70
ARIMA	8.93	11.08	13.84	11.28	73.50
ETS	9.07	10.98	13.74	11.26	56.66

The bold numbers highlight the best performing method per forecasting horizon and computational complexity.

<https://doi.org/10.1371/journal.pone.0194889.t008>

Таблица 4. Оценяване прогнозната точност на статистическите модели

Method	sMAPE(%)	MASE	CC	MF
Naive 2	8.59	0.56	1.00	3.63
SES	7.36	0.49	1.53	2.37
Holt	7.41	0.48	2.31	2.35
Damped	7.30	0.48	3.96	2.34
Comb	7.27	0.48	6.88	2.32
Theta	7.31	0.48	5.84	2.34
ARIMA	7.34	0.47	43.96	2.53
ETS	7.19	0.47	34.07	2.28

Прогнозната точност е оценена чрез симетрична средна абсолютна процентна грешка (SMAPE), средна абсолютна грешка (MASE) и др. Най-ниска е прогнозната грешка (респективно най-висока прогнозна точност), както се вижда от таблица 4 е измерана чрез SMAPE е постигната при ETS моделите, следвани от комбинирания статистически метод, тета метод и т.н.

Основните обобщения, които могат да бъдат направени по отношение на прогнозирането към началото на 2019 г. въз основа на резултатите от M4 състезанието могат да бъдат обобщени в следното⁸:

- прогнозната точност може да бъде подобрена чрез комбиниране на модели;
- най-висока прогнозна точност осигуряват хибридни модели (между статистически и ML) и такива от комбиниран тип (средна аритметична стойност на три модела SES, Holt и Damped);
- традиционните статистически методи осигуряват поточни икономически прогнози в сравнение тези на машинно самообучение (Machine learning);
- методите или моделите, които най-добре пасват на емпиричните данни, не е задължително да доведат до по-точни прогнози извън извадката (общо досега вярване);
- прогнозите на простите статистически методи са най-малко толкова точни, колкото и по-сложните (като ARIMA и ARARMA модели).

⁷ <https://doi.org/10.1371/journal.pone.0194889.t008>

⁸ Makridakis Spyros, The Contributions of the M4 Competition to the Theory and Practice of Forecasting, 2019

Литература

1. Ahmed N., Amir F. Atiya, H. El-Shishiny, An Empirical Comparison of Machine Learning Models for Time Series Forecasting, *Econometric Reviews*, Volume 29, 2010.
2. Bontempi G., S. Ben Taieb, and Y. Le Borgne, *Machine Learning Strategies for Time Series Forecasting*, Springer-Verlag, Berlin Heidelberg, 2013.
3. Everitt B.S., Skrondal A. *Cambridge Dictionary of Statistics*, Cambridge University Press 2010.
4. Makridakis Spyros, *The Contributions of the M4 Competition to the Theory and Practice of Forecasting*, 2019.
5. Makridakis S., E. Spiliotis , V. Assimakopoulos, *Statistical and Machine Learning forecasting methods: Concerns and ways forward*, 2018.
6. Montgomery D. , Ch. Jennings and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons. Inc .. Hoboken. New Jersey, 2008.
7. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194889>
8. <https://doi.org/10.1371/journal.pone.0194889.t008>
9. www.mcompetitions.unic.ac.cy
10. <https://forecasters.org/resources/time-series-data/m3-competition>