Bulgarian Journal of Meteorology and Hydrology

Национален институт по метеорология и хидрология

National Institute of Meteorology and Hydrology

Bul. J. Meteo & Hydro 23/2 (2019) 20-30

Use of some types of errors to improve the forecast of temperatures

Hristo Hristov*, Andrey Bogatchev

National Institute of Meteorology and Hydrology,
Tsarigradsko shose 66, 1784 Sofia, Bulgaria

(Received: 16 Oct 2019, accepted: 23 Dec 2019)

Abstract. In the present work, an assessment is made of the issued forecasts for the minimum and maximum temperatures, by stations. Different types of errors are calculated, special attention is paid to the mean error (systematic). The idea is to deliver this information (location, value and sign of the mean error) as a feedback to the forecaster and to help him/her to reduce the errors, and thus, to increase the accuracy of the forecasts. It is shown how the decrease of the systematic error increases the accuracy, which is the purpose of the work. It is also shown how the accuracy of the forecast of temperatures increases.

Keywords: verification, forecast, assessment, evaluation, extreme temperature, Mean error, Mean Absolute error

1. INTRODUCTION

The aim of the present work is to give a differentiated information (in current points of the issued forecast) for Mean Absolute error (*MAE*), for the percent correct (*PC*) and, mostly for the Mean (systematic) error (*ME*). Using this information, and mainly the systematic error (location, value and sing) the forecaster try to reduce it – at least in the points where it is the biggest. This will lead to an increase of the accuracy of the forecast. Even better results could be expected when personal errors will be calculated.

An integral assessment for the whole country was given by Hristov and Bogatchev (2018). In this first stage of the work the authors present the time series by months and years for the period 2009-2014, but the spacial distribution of the information is lost. Namely the latter, will give us the valuable feedback, that will enact the purpose set

* hristo.hristov@meteo.bg

of this work, which is continuation of mentioned above. The final target of the whole investigation is to improve the accuracy of the temperature's forecasts. The spacial distribution of the forecast will be drawn out by making evaluation by stations. This assessment is presented in the paper. The information, necessary to correct the systematic error and to make the forecasts better, is shown. The references from different authors are used: Fajman (2011), Murphy (1993), Brier and Allen (1951), Riply (2002), Wilk (2005), Wilk (2007). The results from two Bulgarian authors, Spiridonov (1987) and Bogatchev (1994), are used. They calculated MAE for Sofia (Bulgaria) for previous periods.

The Guidance of the World Meteorological Organization (WMO) (http://www.wmo.int/pages/prog/amp/pwsp/pdf/TD-1023.pdf) is used to calculate different kinds of errors and the percent correct of the forecasts. An interval of ±2 °C or ±3 °C is taken as a limit when calculating the percent correct. This assessment is very suitable to be used for the general public or other users. This study uses a margin of admissible error up to 2 °C, i.e. ±2 °C for calculating the percent correct.

2. METHODOLOGY

The short-range weather forecast of National Institute of Meteorology and Hydrology (NIMH) is up to 36-48 hours ahead and it is prepared till 11 AM every day. It consists of a forecast for today and a forecast for the next 24-h (that will be assessed). Besides of the text forecast, a numerical forecast is issued as well. It includes minimum and maximum temperatures and a symbol for the phenomena expected (sun, clouds, rain, snow, thunders, fog etc.). The numerical forecast is elaborated for 68 points in Bulgaria, 37 of which coincide with synoptic stations. For these 37 points namely (shown on Figure 1) the assessment will be carried out. Four of these 37 points are mountainous. The period for assessing is 2009-2014.

Mean error (*ME*) or bias. It can be interpreted as systematic error. The mean error indicates the difference between the forecast and the observations for the day. For a definite period of time, the sign and the magnitude of the systematic error can be seen.

$$ME = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i) \tag{1}$$

where F_i and O_i are respectively the values of the forecast and observations for day i.

Mean Absolute error (*MAE*). This error determines the accuracy of the forecast. It does not take into account the direction of the error, only its magnitude.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F_i - O_i| \tag{2}$$

Percent correct is determined as follows:

$$PC = \frac{1}{n}\sum_{i=1}^{n}\begin{Bmatrix} correct, if\ (F_i - O_i) \le 2°C \\ not\ correct, if\ (F_i - O_i) > 2°C \end{Bmatrix} \times\ 100\ \% \quad (3)$$

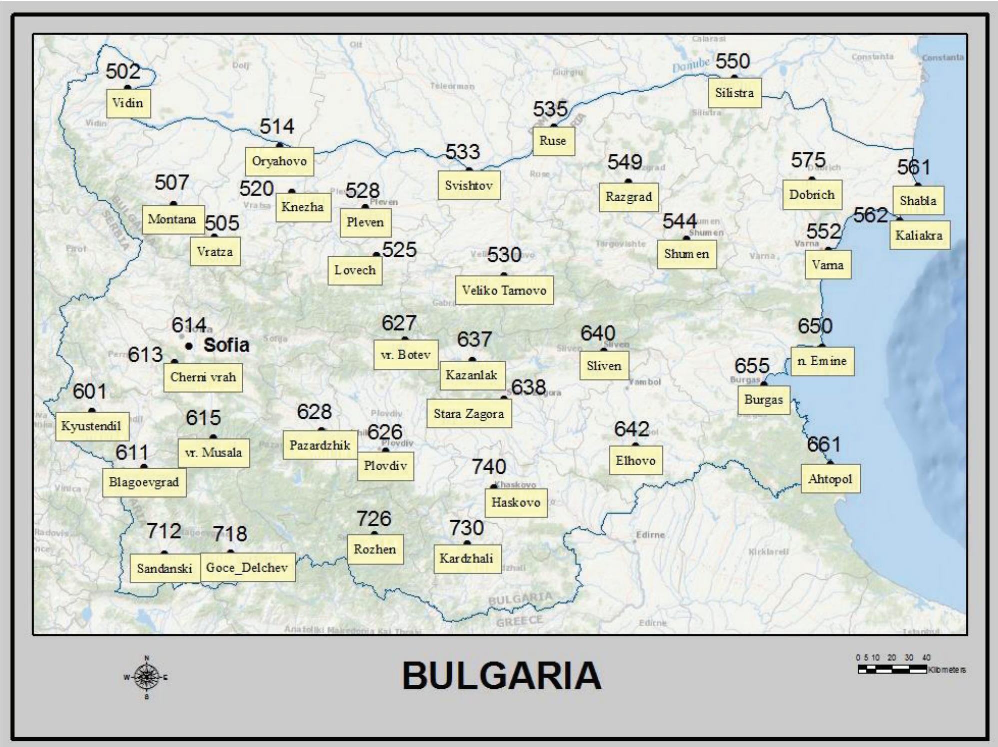

Fig. 1. Map of the points, where the assessment will be carried out, names of the stations and their WMO code. First 2 digits (15) are omitted.

3. RESULTS

Unlike the integral method, where information for different points is lost, in this investigation we use a differentiation of the stations. Evaluation by stations is made by calculating for the period 2009 - 2014 for the current stations: mean error, mean absolute error and percent correct. The purpose here is to look in details in the spatial distribution of the errors (of the temperatures), where (in which point) we detect bigger errors (*MAE*) and, respectilely, the sing of these errors (*ME*). Namely the latter will be the most useful for our studies.

3.1. Mean absolute error (*MAE*)

MAE for the minimum and the maximum temperatures for the period 2009 – 2014 are presented on Figure 2. The mean absolute error in the minimum temperatures is the least for all high mountainous stations (615, 613, and 627) and for Burgas station (655). The forecasts in these points are the most accurate. Also, in the minimum temperatures the least accuracy is for Plovdiv (626), Stara Zagora (638), and Kazanluk (637). From this graph, it is seen that the maximum temperature's error is bigger (for most of the points) than the minimum temperature's error (where accuracy is greater).

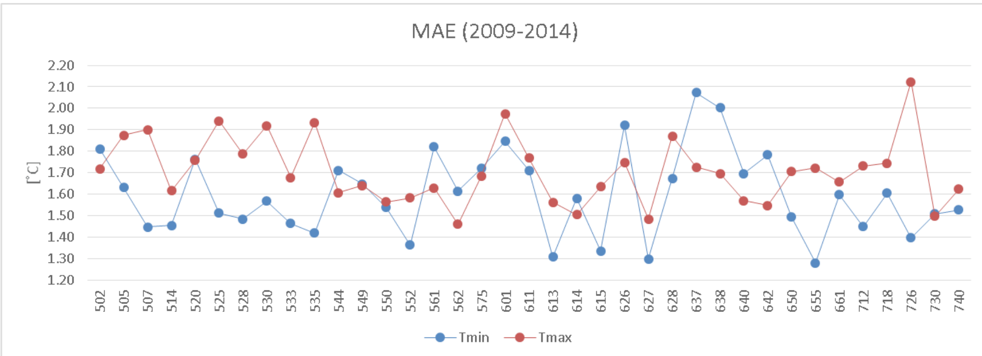

Fig. 2. *MAE* for the minimum and the maximum temperatures in the stations for the whole period 2009-2014 (station ID according to WMO code, first 2 digits (15) omitted).

Independently, Spiridonov (1987) and Bogatchev (1994) made calculations of *MAE* for Sofia. The results of their investigation in comparison with ours are shown on Figure 3. To be precise, we have to mention that in the past (before 2006), forecasters did not predict the minimum temperatures in the summer season (15 May - 15 September), unlike recent years. That is why, to be possible to make comparison, the data for minimum temperatures from May, 15 till September, 15 for the period 2009 – 2014 are excluded. During the warm months *MAE* is the least, that is why the real error in *Tmin* should be lower than shown on Figure 3. A comparison of *MAE* is made, and it can be seen how accuracy changes through period of 30 years.

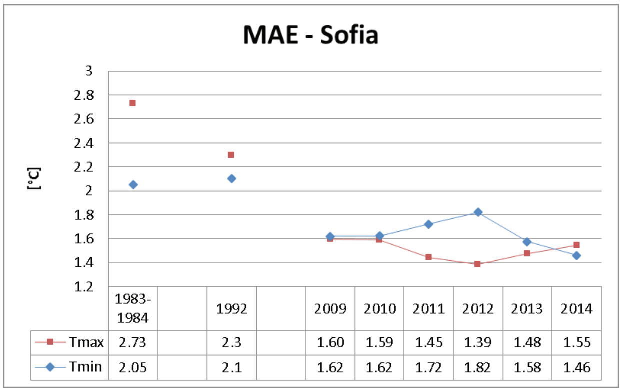

Fig. 3. *MAE* for the minimum and the maximum temperatures for human forecasts in the station 15614 (Sofia) for 2009-2014, 1983-1984 and 11.1991-10.1992.

Few conclusions could be drawn from Figure 3:

- *MAE* for the maximum temperature decreases significantly through the years: from 2.73 °C in 1983/1984 to about 1.5 °C. This is a big improvement of the accuracy of the forecast and it is mostly due to the integration in operational mode and progress of the numerical models. The forecasters experience probably did not change significantly.
- In the minimum temperatures, *MAE* decreases from 2.1 °C to a little below to 1.5 °C, i.e. here the initial values are considerably lower (comparing to the maximum temperatures) and they are due to the bigger influence of the persistence error (data used are with a day difference between observation and forecast, unlike to the maximum temperatures, where this difference is 2 days). Also at the end of the period, there is little influence from the numerical models, and this leads to a little higher accuracy during the recent years.

3.2. Percent Correct (*PC*)

The realization is shown on Figure 4:

- In the minimum temperatures, with some little exceptions, the percent correct is higher.
- In the minimum temperatures there are maximums on peak Botev (627), Cherni vruh (613), peak Musala (615), Varna (552), and Burgas (655), the greastest values are on peak Botev and Burgas. The lowest percent correct is in Stara Zagora (638) and Kazanluk (637), but for the latter for the years 2009 and 2010 data are missing, i.e. the line is shorter.
- In the maximum temperatures, the highest percent correct is on the cape Kaliakra (562) and the peak Botev (627), there are some quite big minimum in Lovech

(525), Ruse (535), and Kustendil (601), but the lowest percent correct is in Rozhen (726).

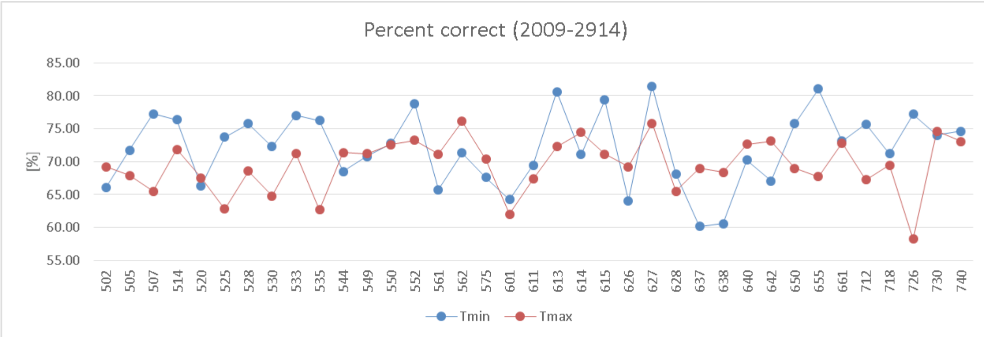

Fig. 4. Percent correct for the minimum and the maximum temperatures in the stations for the whole period 2009-2014 (station ID according to WMO code, first 2 digits (15) omitted).

Figure 5 is similar to Figure 3 in *MAE*, but for the percent correct for Sofia, only the study of Spiridonov (1987) being included, as it is the only available. To be possible to make comparison, from *PC* for the minimum temperatures the data from May, 15 till September, 15 are removed.

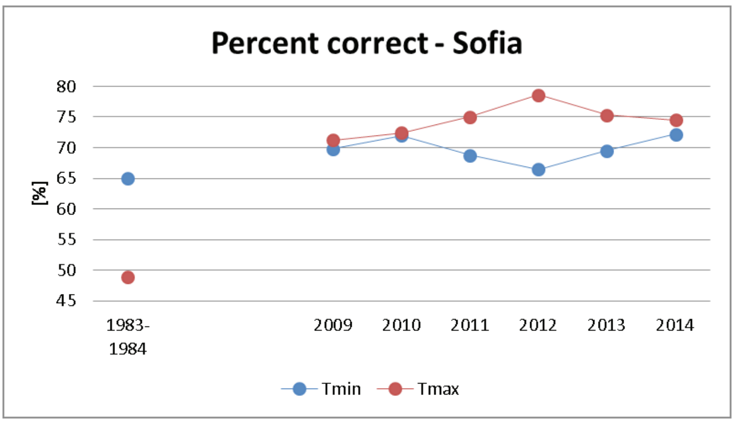

Fig. 5. Percent correct for the minimum and the maximum temperatures for human forecasts in the station 614 (Sofia) for 2009-2014 and 1983-1984.

- There is a big increase of *PC* for the maximum temperatures, from 49% to about 75%. This is due mainly to the integration in operational mode and progress of the numerical models.
- In minimum temperatures, there is an increase from 65% to about 70% (exception is the year 2012, when *PC* is lower), and the improvement here is due again, mostly, to the numerical models. The influence of the persistence error is bigger

in the minimum temperatures, that is why in 1983/1984 there is a bigger *PC* regarding the maximum temperatures.

- For 30-year period, the *PC* grows up: in the minimum temperatures with about 5%, in the maximum temperatures – with more than 25%.

3.3. Mean error (*ME*)

Finally, we will show the results of the mean error (the most important for us). This is the error that could give the valuable feedback to the forecaster, i.e. it could be used for correction of the forecast and for improving of its accuracy.

Up to now, we knew that there are different systematic errors in the different points of issuing of the forecast, but it was not clear what their magnitude was. The exact values of the Mean error for the period 2009-2014 (as well as of the error by years) are given in Table 1. The table is colored (lifted) – red for overestimate, and blue – for underestimate. We shall accept that *ME* below 0.2 °C is small enough and will be in green. The mean error for the maximum and the minimum temperatures for the whole period 2009-2014 is presented on Figure 6.

It is seen, for example for Rozhen (726), that **underestimating** in the maximum temperatures is more than 1 °C. For station Sliven (640) things are similar – **underestimate**, but this time, it concerns the minimum temperatures. For Lovech (525), Pleven (528), Veliko Turnovo (530), Shumen (544), Dobrich (575), Emine (650), Burgas (655), Kurdzali (730), Kustendil (601) – **underestimate** in the maximum temperatures. For Vidin (502) we have systematic **overestimating** in the minimum temperatures, as well as for Stara Zagora (638) and Akhtopol (661). For Silistra (550) and peak Botev (627) the **overestimate** in the maximum temperatures is almost 0.5 °C. For Haskovo (740) there is also an **overestimation**, but in the last year it turns to zero.

A systematic underestimation in the minimum and in the maximum temperatures in the most Bulgarian stations is shown on Figure 6.

Finally, we can conclude that for most of the stations there are systematic deviations (+ or -). But the underestimation (for the minimum and the maximum temperatures) for Sliven and Rozhen has to be acknowledged is more than 1 degree.

The benefit from all these calculations is that the systematic errors and their magnitude can be identified, i.e. the Table 1 with the *ME* can be used for correction of the forecasts – to decrease the systematic errors and, therefore, to increase the accuracy.

To confirm the above statement we show two figures – Figure 7 and Figure 8. They present *MAE* of human forecasted temperatures (minimum and maximum) corrected with *ME* over a 6-year period.

We can see on Figure 7 that big *ME*, corresponds to big corrected *MAE* or much better accuracy. The points with better accuracy are 514, 550, 562, and 640.

Table 1. *ME* for the minimum and the maximum temperatures for 2009-2014, and for each year of that period in the stations (station ID according to WMO code, first 2 digits (15) omitted). Green is from -0.19 to 0.19 (no systematic error), blue underestimate (< -0.2), red overestimate (> 0.2).

ME		2009-2014		2009		2010		2011		2012		2013		2014	
N	St_ID	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]	Tmin [°C]	Tmax [°C]
1	502	0.32	-0.17	0.59	0.01	0.15	0.2	0.14	-0.25	0.25	-0.49	0.48	-0.08	0.32	-0.48
2	505	-0.52	-0.13	-0.48	0.13	-0.33	0.12	-0.73	0.04	-0.76	-0.66	-0.42	-0.26	-0.41	-0.17
3	507	0.04	-0.26	0.09	0.12	-0.21	0.21	-0.26	-0.26	0.23	-0.84	0.62	-0.4	-0.23	-0.34
4	514	-0.78	0.13	-0.76	0.00	-0.69	0.39	-0.91	0.25	-0.92	-0.08	-0.72	0.28	-0.68	-0.07
5	520	0.04	-0.08	0.06	-0.14	-0.02	0.26	-0.22	-0.08	0.06	-0.35	0.19	0.02	0.17	-0.16
6	525	-0.37	-0.59	-0.36	-0.68	-0.37	-0.6	-0.64	-0.62	-0.45	-0.58	-0.14	-0.42	-0.28	-0.61
7	528	0.02	-0.29	0.04	-0.32	0.16	-0.21	-0.54	0.12	0.06	-0.54	0.19	-0.39	0.19	-0.42
8	530	-0.27	-0.76	-0.47	-1.18	-0.42	-0.79	-0.45	-0.87	-0.28	-0.87	0.03	-0.58	-0.05	-0.24
9	533	-0.59	0.05	-0.37	0.21	-0.34	0.31	-0.87	-0.04	-0.79	-0.18	-0.56	0.1	-0.61	-0.07
10	535	-0.41	-0.39	-0.27	-0.50	-0.22	-0.24	-0.74	-0.56	-0.62	-0.68	-0.28	-0.19	-0.35	-0.16
11	544	-0.13	-0.51	-0.04	-1.02	-0.26	-0.67	-0.3	-0.62	-0.11	-0.36	0.02	-0.19	-0.12	-0.23
12	549	-0.66	0.03	-0.88	-0.17	-0.76	-0.08	-0.84	0.1	-0.76	0.07	-0.42	0.23	-0.28	0.04
13	550	-0.82	0.43	-0.64	0.54	-0.7	0.63	-1.07	0.22	-0.83	0.34	-0.88	0.37	-0.78	0.45
14	552	-0.31	-0.19	-0.17	-0.53	-0.41	-0.4	-0.44	-0.25	-0.31	0.06	-0.46	-0.3	-0.1	0.26
15	561	0.13	-0.26	0.27	-0.07	-0.05	-0.08	0.11	-0.14	0.1	-0.5	0.07	-0.43	0.19	-0.11
16	562	-0.69	0.24	-0.83	0.25	-0.63	0.39	-0.61	0.48	-0.7	0.11	-0.89	0.02	-0.47	0.21
17	575	0.23	-0.62	0.34	-0.73	0.14	-0.54	0.02	-0.44	0.17	-0.66	0.19	-0.83	0.51	-0.5
18	601	-0.02	-0.81	-0.18	-0.98	-0.19	-1.23	0.04	-0.61	-0.09	-0.86	0.31	-0.52	0.08	-0.64
19	611	-0.12	-0.20	-0.26	-0.08	-0.35	-0.12	0.01	-0.11	-0.26	-0.52	0.1	-0.12	0.02	-0.27
20	613	-0.11	-0.05	-0.36	0.04	-0.04	0.21	-0.31	-0.22	-0.06	-0.09	0.1	0.06	0.02	-0.31
21	614	-0.34	-0.12	-0.29	-0.18	-0.62	-0.24	-0.39	-0.1	-0.52	-0.31	-0.18	0.13	-0.07	-0.01
22	615	0.20	-0.33	-0.03	-0.69	0.11	-0.35	0.23	-0.38	0.16	-0.15	0.44	-0.09	0.32	-0.29
23	626	0.14	-0.05	0.35	0.03	0.25	0.13	-0.12	-0.04	-0.06	-0.28	0.42	-0.07	-0.01	-0.06
24	627	-0.21	0.47	-0.36	0.43	-0.25	0.62	-0.15	0.53	-0.16	0.46	-0.01	0.63	-0.33	0.23
25	628	0.26	-0.33	0.46	-0.26	0.19	-0.07	0.02	-0.38	0.07	-0.57	0.56	-0.43	0.23	-0.33
26	637	0.21	-0.50					0.11	-0.58	0.14	-0.68	0.45	-0.52	0.2	-0.25
27	638	0.44	0.12	0.58	-0.01	0.87	0.85	0.1	-0.13	0.2	-0.19	0.61	0.09	0.29	0.11
28	640	-1.05	-0.42	-1.24	-0.55	-1.18	-0.34	-1.15	-0.47	-1.19	-0.57	-0.82	-0.5	-0.7	-0.09
29	642	0.32	-0.32	0.55	-0.26	0.05	-0.18	0.14	-0.26	0.38	-0.51	0.51	-0.36	0.31	-0.32
30	650	-0.42	-0.44	-0.45	-0.26	-0.32	-0.31	-0.39	-0.44	-0.25	-0.88	-0.68	-0.43	-0.4	-0.32
31	655	-0.09	0.13	-0.30	0.34	-0.3	0.31	-0.15	0.26	-0.1	0	0.21	-0.05	0.1	-0.09
32	661	0.23	0.17	0.32	0.18	0.06	0.23	-0.01	0.61	0.18	0.04	0.21	-0.08	0.57	0.03
33	712	-0.23	-0.44	-0.25	-0.45	-0.36	-0.45	-0.25	-0.38	-0.34	-0.52	-0.05	-0.38	-0.1	-0.54
34	718	-0.10	0.19	-0.35	0.09	-0.27	0.34	-0.05	0.35	-0.2	0.29	-0.04	0.34	0.28	-0.3
35	726	-0.20	-1.09	-0.46	-1.72	-0.16	-1.32	-0.17	-1.05	-0.13	-1.03	-0.08	-0.71	-0.23	-0.7
36	730	0.01	-0.37	0.21	-0.40	-0.08	-0.37	-0.17	-0.25	-0.25	-0.54	0.21	-0.29	0.16	-0.37
37	740	0.01	0.40	0.16	0.43	-0.12	0.3	-0.25	0.38	-0.06	0.35	0.14	0.26	0.08	0.1
средно		-0.16	-0.20	-0.16	-0.23	-0.21	-0.09	-0.30	-0.17	-0.22	-0.35	-0.02	-0.16	-0.06	-0.19

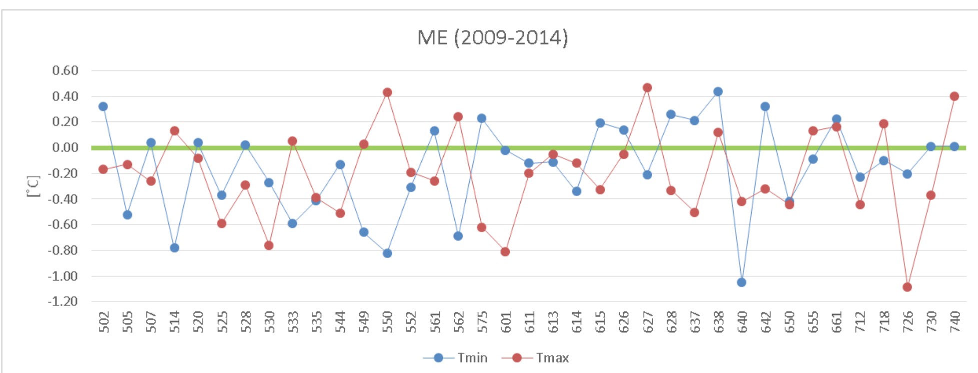

Fig. 6. *ME* of human forecast for minimum and maximum temperatures for the whole period 2009-2014.

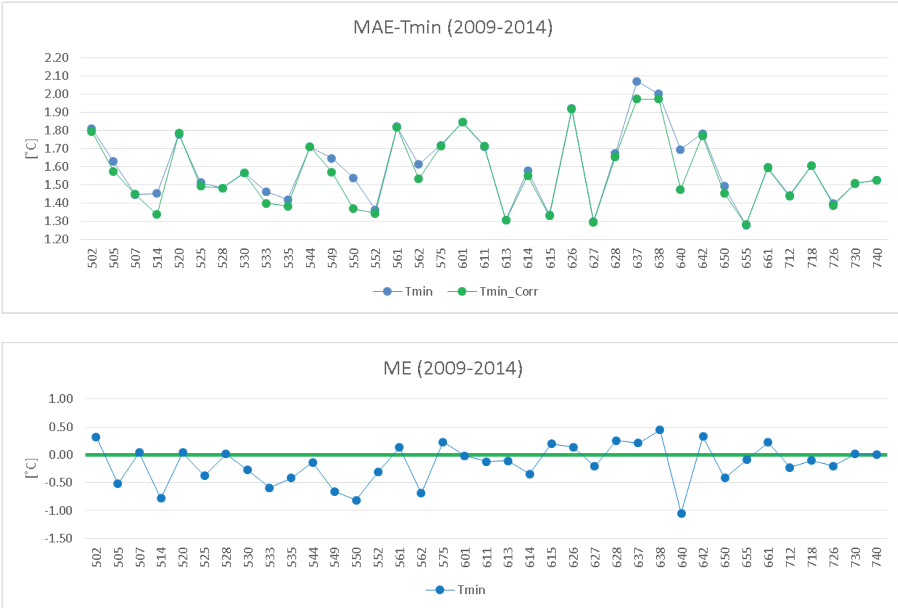

Fig. 7. *MAE* and corrected *MAE* (top) and *ME* (bottom) for the minimum temperatures for the whole period 2009-2014.

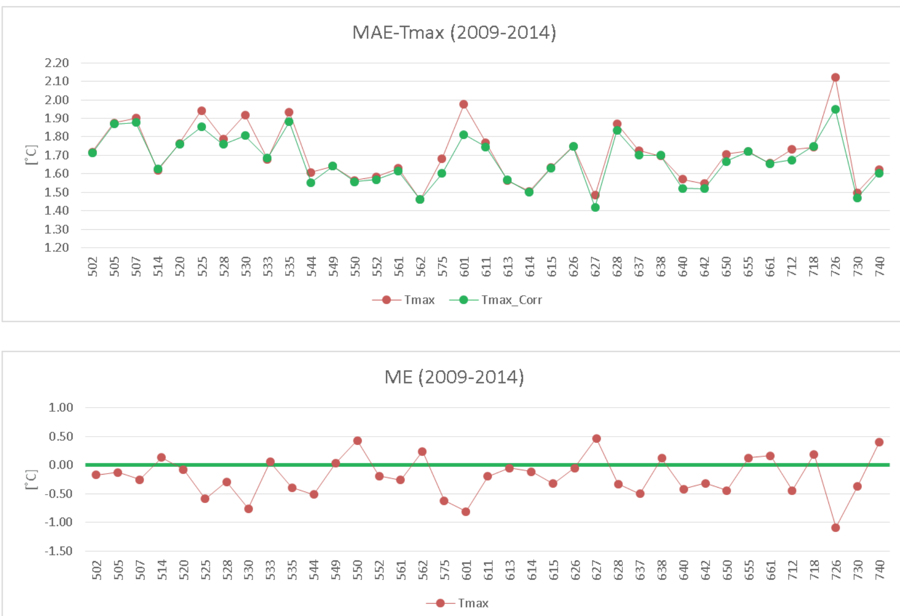

Fig. 8. *MAE* and corrected *MAE* (top) and *ME* (bottom) for the maximum temperatures for the whole period 2009-2014.

Similar results are seen on Figure 8 for the maximum temperatures. The biggest increase of accuracy is in points 530, 601 and 726, because in these points *ME* is very big, up to and over 1 °C – underestimation. Most of the other points also have increase of accuracy, but it is smaller (smaller systematic error - *ME*) – this applies both to minimum and maximum temperatures.

This confirms the statement that decrease of the systematic error (*ME*) leads to an increase of the accuracy. This technique was included in the numerical models (will be presented in a next investigation), where the increase of the accuracy is much bigger than in the human forecast. For example for *Tmax* on European Center for Medium-Range Weather Forecasts (ECMWF) the mean percent correct (for all stations) increases with 20%, while *MAE* decreases with 0.5°C.

Evaluation of the personal mean error will give more valuable feedback to the forecasters and would lead to even better results in diminishing the systematic error and, respectively, to increase the forecast accuracy. This is the reason to make personal

assessments every year and every 6 months (warm and cold half-year) since 2015. Thus, for the last few years, our forecasters are aware with their systematic errors (location, value and sing) and try to reduce them.

4. CONCLUSIONS

- Decreasing of the systematic error (*ME*) leads to an increasing of the accuracy.
- A valuable information is shown about the systematic errors – location, value and sign – necessary to increase the forecast's accuracy.
- For Sofia station, for the last 30 years, there is an increase of the percent correct: in the minimum temperatures – with about 5%, in the maximum temperatures – with more than 25%.

REFERENCES

Bogatchev, A., (1994): Statistical schemes for medium-range weather forecast, Abstract of a PHD thesis.

Brier, G.W. and R. A. Allen, (1951): Verification of weather forecasts, Compendium of meteorology, 841–848.

Fajman, P., (2011): Improved verification and analysis of national weather service point forecast matrices. For the Degree of Master of Science in the Graduate College at the University of Nebraska.

Murphy, A.H., (1993): What is a good forecast? an essay on the nature of goodness in weather forecasting, Weather and forecasting 8 (2), 281–293.

Ripley, E. and O. Archibold, (2002): Accuracy of Canadian short-and medium-range weather forecasts, Weather 57 (12), 448–457.

Spiridonov, V., (1987): Skill measures of daily extreme temperature forecast, Problems of Meteorology and Hydrology 2, 19–26.

Wilk, G.E., (2005): The Temperature and Precipitation Verification Program at WFO Corpus Christi Texas, United States Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service.

Wilk, G.E., (2007): Temperature and Precipitation Verification Results and Interpretation at WFO Corpus Christi and Other WFOS January 2004 Through June 2006, United States Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service.

Hristov, H., (2018): An assessment of daily extreme temperature forecasts – stations average view, IDŐJÁRÁS Vol. 122, No. 3, July – September, 237–257.

WMO: http://www.wmo.int/pages/prog/amp/pwsp/pdf/TD-1023.pdf Guidance of the World Meteorological Organization