Decision Science Letters 4 (2015) 373–378

Contents lists available at GrowingScience

Decision Science Letters

homepage: www.GrowingScience.com/dsl

Optimization of continuous ranked probability score using PSO

Seyedeh Atefeh Mohammadi[a], Morteza Rahmani[a,b*] and Majid Azadi[c]

[a]*Industrial engineering department, Technology development institute (ACECR), Tehran, Iran*
[b]*University of Science and Culture, Faculty of Basic and Advanced Technologies in Biology, Tehran, Iran*
[c]*Dynamic and synoptic meteorology research group, Atmospheric science and meteorological research center, Tehran, Iran*

C H R O N I C L E

Article history:
Received December 10, 2014
Received in revised format: March 2, 2015
Accepted March 23, 2015
Available online April 1 2015

Keywords:
CRPS
PSO
BFGS
Weather prediction

A B S T R A C T

Weather forecast has been a major concern in various industries such as agriculture, aviation, maritime, tourism, transportation, etc. A good weather prediction may reduce natural disasters and unexpected events. This paper presents an empirical investigation to predict weather temperature using minimization of continuous ranked probability score (CRPS). The mean and standard deviation of normal density function are linear combination of the components of ensemble system. The resulted optimization model has been solved using particle swarm optimization (PSO) and the results are compared with Broyden–Fletcher–Goldfarb–Shanno (BFGS) method. The preliminary results indicate that the proposed PSO provides better results in terms of CRPS deviation criteria than the alternative BFGS method.

1. Introduction

Weather forecast has been a major concern in various industries such as agriculture, aviation, maritime, tourism, transportation, etc. A good weather prediction may reduce natural disasters and unexpected events. Since 1950, with the advent of computers with high computing power, numerical weather prediction models have been presented. Numerical weather prediction models involve differential equations, which describe the physical laws and dynamics of atmosphere and they mainly depend on boundary conditions. Due to approximate solutions of numerical methods, results obtained from the implementation of numerical models approximate the real-world conditions and always contain some errors (Palmer, 1998; Zeng et al., 1993). Of course, the use of numerical methods is not the only cause of the error in the output but there are also some other sources of error. Among them the presence of errors in the initial/boundary values, the lack of data in some areas and chaotic nature of the dynamical system of the atmosphere are the most important.

* Corresponding author.
E-mail address: rahmanimr@jdsharif.ac.ir (M. Rahmani)

doi: 10.5267/j.dsl.2015.4.001

According to Eckel et al. (2012) “Ambiguity is uncertainty in the prediction of forecast uncertainty, or in the forecast probability of a specific event, associated with random error in an ensemble forecast probability density function. In ensemble forecasting ambiguity arises from finite sampling and deficient simulation of the various sources of forecast uncertainty”.

In a deterministic forecast, only one initial condition is considered as the input to the model regardless of the inherent uncertainty of the atmosphere. But in an ensemble prediction system the forecast is produced by drawing a finite sample from the probability distribution describing the uncertainty of the initial state of the atmosphere and the weather is forecasted by a probability distribution function. Since each ensemble member is a deterministic forecast, it involves systematic errors and needs to be post processed. There are some methods to remove bias and produce a probability distribution function such as rank histogram (Hamill, 2001), dressing method (Roulston & & Smith, 2003), Bayesian model averaging (BMA) (Raftery, et al., 2005), ensemble regression (Unger et al., 2009), variance inflation technique (Johnson & Bowler, 2009), ensemble model output statistics model (EMOS) (Gneiting et al., 2005) and ensemble kernel density model output statistics (EKDMOS) (Veenhuis, 2013).

In this research, an attempt is made to estimate a normal density function to forecast the surface temperature over Iran through minimization of continuous ranked probability score (CRPS) using particle swarm optimization (PSO) algorithm in the training period.

2. The proposed study

This paper presents an empirical investigation to predict weather temperature using continuous ranked probability score (CRPS) (Gneiting et al., 2005; Toth & Kalnay, 1997; Gneiting & Raftery, 2007). Let $X_1, \dots, X_\mathrm{m}$ be the components of ensemble system. In addition, let S^2 be the variance of the components of ensemble system. Therefore, the linear combinations of ensemble system is defined as follows,

$$N(a + b_1 X_1 + \cdots + b_m X_m, c + dS^2) \tag{1}$$

where $a + b_1 X_1 + \cdots + b_m X_m$ and $c + dS^2$ are mean and variance of the ensemble system. Let $b_1, \dots, b_\mathrm{m}$ be the coefficients that present the performance of each ensemble member and c and d represent the coefficients of variance and are estimated in training period. The CRPS function is defined as follows,

$$crps(F, y) = \int_{-\infty}^{+\infty} (F(t) - H(t - y))^2 dt \tag{2}$$

where $H(t - y)$ is a Heaviside function, which is defined as follows,

$$H(t - y) = \begin{cases} 0 & t < y \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

Let F be a normal density distribution function with mean and standard deviation of μ and σ. Therefore we have,

$$crps(\mathcal{N}(\mu, \sigma^2), y) = \sigma \left\{ \frac{y - \mu}{\sigma} \left(2\Phi\left(\frac{y - \mu}{\sigma}\right) \right) + 2\varphi\left(\frac{y - \mu}{\sigma}\right) - \frac{1}{\sqrt{\pi}} \right\} \tag{4}$$

where $\varphi\left(\frac{y-\mu}{\sigma}\right)$ and $\Phi\left(\frac{y-\mu}{\sigma}\right)$ are the probability density and distribution functions, respectively. When F_i represents the deterministic prediction, $CRPS = \frac{1}{n}\sum_{i=1}^{n} crps(F_i, y_i)$ will be changed to mean absolute value. In fact, the average score is the mean absolute error, which is the generalized model and as it tends to zero the prediction will be more desirable.

3. Particle Swarm Optimization

Particle swarm optimization (PSO) is an evolutionary technique for optimizing functions designed based on social behavior of birds by Kennedy (2010). In this technique, a group of particles, as the variables of an optimization problem, are scattered in the search environment. More precisely, some particles have better positions than others do. Therefore, based on aggregative particles' behavior, other particles may try to raise their position to the prior particles' positions. The proposed PSO uses the following notations,

$t, t=1, \ldots, T$	Index for repetition
$i, i=1, \ldots, I$	Index for particles
$d, d=1, \ldots, D$	Index for dimensions
u	Uniform distribution number [0, 1]
p_{id}	Position of pbest particle *i* in dimension *d*
p_{gd}	Position of *gbest* in dimension *d*
c_p	Acceleration factor of *pbest*
c_g	Acceleration factor of *gbest*
X^{min}	The minimum value of position
X^{max}	The maximum value of position
w	Static weight
$v_{id}(t)$	Velocity of i^{th} particle in dimension *d* and repetition *t*
$x_{id}(t)$	Position of i^{th} particle in dimension *d* and repetition *t*
X_i	Vector of position $[x_{i1}, x_{i2}, \ldots, x_{iD}]$
V_i	Vector of velocity $[v_{i1}, v_{i2}, \ldots, v_{iD}]$
p_i	Vector of position of pbest $[p_{i1}, p_{i2}, \ldots, p_{iD}]$
p_g	Vector of position of gbest $[p_{g1}, p_{g2}, \ldots, p_{gD}]$
$z(x_i)$	The competence of X_i

The proposed study uses the following to adjust the speed and the position for each particle.

$$v(t+1) = wv(t) + c_p u\big(x(t) - x_{pbest}\big) + c_g u\big(x(t) - x_{gbest}\big), \tag{5}$$

$$x(t+1) = x(t) + v(t+1). \tag{6}$$

PSO Algorithm

Step 1. Initial values

Generate X_i in the interval [X^{min}, X^{max}], $i=1,\ldots, I$

Set $P_g = X_1$, $P_i = X_i$, $V_i = 0$

Step 2. Obtain $Z(X_i)$ for i=1,…, I

Step 3. Update *pbest*, *if* $z(x_i) < z(p_i)$ *then* $P_i = X_i$, $i=1,\ldots, I$

Step 4. Update *gbest*, *if* $z(p_i) < z(p_i)$ *then* $P_i = P_i$, $i=1,\ldots, I$

Step 5. Update velocity and position of each particle according to Eq. (5) and Eq. (6)

If $x_{id}(t+1) > X^{max}$ *then* $x_{id}(t+1) = X^{max}$ and $v_{id}(t+1) = 0$

If $x_{id}(t+1) < X^{min}$ *then* $x_{id}(t+1) = X^{min}$ and $v_{id}(t+1) = 0$

Step 6. If t=T Stop, otherwise *t = t + 1 go to Step 2.*

The proposed PSO method of this paper uses a vector for *x(t)* with the length of 12 where the first cell contains *a*, the second to tenth cells are devoted to b_1 to b_m and the last two items are devoted to *c* and *d*, respectively. Similarly, *v(t)* is also a vector of 12 cells, which contains the speed of each particle. X^{min} and X^{max} are as follows,

X^{min}	-10	0	0	0	0	0	0	0	0	0	0	0
X^{max}	+10	1	1	1	1	1	1	1	1	1	5	0.1

In our survey, I=50, T=100, $w = 0.5$, $c_p = c_g = 2$. Training program in this study is a rolling horizon of 60 days. In our survey, the ensemble system consists of 9 members, where each member contains weather temperature under various physical environments. For prediction, Weather Research and Forecasting (WRF) (Skamarock et al., 2005) model with initial Global forecast System (GFS) is used. The data is gathered from September, 4th, 2011 to February, 4th, 2012 leading us to have 160 data.

3. The results

In order to measure the performance of the proposed study, we compare the results with Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Fletcher, 1970). MAE and RMSE have been calculated for both mean of raw ensemble and mean of bias corrected ensemble and compared with MAE and RMSE of the mean of the normal probability density function.

Table 1
The summary of comparison of errors

Description	MAE	RMSE
Member 2	6.38	7.51
Member 2 with the bias-corrected	4.07	5.04
Member 8	3.61	4.42
Member 8 with the bias-corrected	2.45	3.04
Mean of ensemble system with raw data	3.97	4.38
Mean of ensemble system with the bias-corrected	2.54	2.97
Average normal probability density function	1.64	2.14

Ensemble members have been post processed by moving average method with a 15-day training period. Results show the minimum amounts obtained for the implementation of PSO on CRPS versus BFGS method are 1.19 and 1.21, respectively.

Table 1 demonstrates the results of our survey when simple mean error is used for raw data and bias-corrected data in the test period. As we can observe from the results of Table 1, the mean of the normal probability density function has MAE 59% less and RMSE 51% less.

4. Conclusion

This paper has presented a survey to forecast weather temperature using continuous ranked probability score (CRPS). The mean and standard deviation of normal density function were linear combination of the components of ensemble system. The resulted optimization model has been solved using particle swarm optimization (PSO) and the results are compared with BFGS method. The preliminary results indicate that the proposed PSO has provided better results in terms of CRPS deviation criteria than the alternative BFGS method.

Acknowledgement

The authors would like to thank the anonymous referees for constructive comments on earlier version of this paper.

References

Eckel, F. A., Allen, M. S., & Sittel, M. C. (2012). Estimation of ambiguity in ensemble forecasts. *Weather and Forecasting*, *27*(1), 50-69.

Estephenson, D. (2007). An introduction to probability forecasting. In *Seasonal Climate: Forecasting and Managing Risk*, Netherlands: Springer, pp. 241-264.

Fletcher, R. (1970). A new approach to variable metric algorithms. *The computer journal*, *13*(3), 317-322.

Gneiting, T., Raftery, A. E., Westveld III, A. H., & Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, *133*(5), 1098-1118.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359-378.

Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, *129*(3), 550-560.

Johnson, C., & Bowler, N. (2009). On the reliability and calibration of ensemble forecasts. *Monthly Weather Review*, *137*(5), 1717-1720.

Kennedy, J. (2010). Particle swarm optimization. In *Encyclopedia of Machine Learning* (pp. 760-766). Springer US.

Palmer, T. N., Gelaro, R., Barkmeijer, J., & Buizza, R. (1998). Singular vectors, metrics, and adaptive observations. *Journal of the Atmospheric Sciences*, *55*(4), 633-653.

Roulston, M. S., & Smith, L. A. (2003). Combining dynamical and statistical ensembles. *Tellus A*, *55*(1), 16-30.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., & Powers, J. G. (2005). *A description of the advanced research WRF version 2* (No. NCAR/TN-468+ STR). National Center For Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div.

Toth, Z., & Kalnay, E. (1997). Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, *125*(12), 3297-3319.

Veenhuis, B. (2013). Spread Calibration of Ensemble MOS Forecasts. *Monthly Weather Review*, 141, 2468-2482.

Zeng, X., Pielke, R. A., & Eykholt, R. (1993). Chaos theory and its applications to the atmosphere. *Bulletin of the American Meteorological Society*, *74*(4), 631-644.