Multi-dimensional extrapolation on use of multi multi-layer neural networks

Seisho Oshige*, Tomoo Aoyama*, and Umpei Nagashima**
*The Faculty of Engineering, Miyazaki University, Miyazaki, Japan
(Fax : +81-0985-58-7411; E-mail: tgb310u@student.miyazaki-u.ac.jp)
** Grid Technology Research Center, National Institute of Advanced Industrial Science and Technology
(Tel : +81-3-5246-6204; Fax : +81-3-5246-6208; E-mail: u.nagashima@aist.go.jp)

Abstract: It is an interest problem to predict substance distributions in three-dimensional space. Recently, a research field as Geostatistics is advanced. It is a kind of inter- or extrapolation mathematically. Some useful means for the inter- and extrapolation are known, in which slide window method with neural networks is hopeful one. We propose multi-dimensional extrapolation using multi-layer neural networks and the slide-window method. The multi-dimensional extrapolation is not similar to one-dimension. It has plural algorithms. We researched line predictors and local-plain predictors I two-dimensional space. The both predictors are equivalent; however, in multi-dimensional extrapolation, it is very important to find the direction of predictions. Especially, since the slide window method requires information to predict the future in sampling data, if they are not ordered appropriately in the direction, the predictor cannot operate. We tested the extrapolation for typical two-dimensional functions, and found an excellent character of slide-window method based on local-plain. By using the method, we can extrapolate the function until twice-outer regions of the definitions.

Keywords: multi-dimensional extrapolation, slide window method, multi multi-layer neural networks, time series prediction

1. INTRODUCTION

Slide window method is a method to predict time series phenomena [1, 2-3]. It is based on a multi-layer neural network. The observations of phenomenon are processed as a set; and it is divided into partial sets that have same element number. The partial sets are learned by a neural network. The fragments of observations are stored as patterns in the network. Neural networks have a function of pattern recognition originally. Therefore, the neural network responses for an unknown time series pattern by using the recognition function. This is the principle of the slide window method. The slide window method is one of extrapolation mathematically, and can be expanded for many applications. Recently, a statistics for spatial observations is advanced. Geostatistics is such a new spatial statistics, which predict 3-dimensional distributions of mineral deposits [4]. Using observations at boring places, it predicts amounts of the mineral at other un-boring places. The geostatistics is applied to many fields now, which are environmental science, meteorology, fisheries, and maintenance of forests and river. Prediction methods of the geostatistics are called "kringing." The kringing is a kind of weighted mean, and it requires probabilities, variograms. We seem that it is an interpolation. Thus, if we could propose a spatial extrapolation, it would have wider applications. Our objective is development of the spatial extrapolation on use of neural networks.

2. ONE-DIMENSIONAL EXTRAPOLATION

Since slide window method is one-dimensional extrapolation, it is easy to make it to multi-dimensional expression. It is based on an idea that a plain can be constructed of many curves.
The one-dimensional time series observation are written,

$$X=\{X_0,X_1,X_2,\ldots,X_n\}. \qquad (1)$$

We make $Y_0,Y_1,\ldots,Y_m$ that are subsets of X. Assuming Y_i and Y_j have same number of elements, we define Y_i as following.

$$Y_0=\{X_0,X_1,\ldots,X_m\},\ m<n,\ n=even \qquad (2)$$

$$Y_1=\{X_1,X_2,\ldots,X_{m+1}\},$$

$$Y_2=\{X_2,X_3,\ldots,X_{m+2}\},\ldots \qquad (3)$$

$$Y_L=\{X_L,X_{L+1},\ldots,X_n\},\ L=n-m$$

We define scalar variables, T_i, as follows.

$$T_0=X_{m+1},\ T_1=X_{m+2},\ldots,T_L=X_{m+L+1}=X_{m+n-m+1}=X_{n+1} \qquad (4)$$

We regard pairs, $(Y_0,T_0), (Y_1,T_1),\ldots,(Y_{L-1},T_{L-1})$ as learning data of a neural network. Where the condition, $D_{ij}||Y_i-Y_j||\sim 0$, is not allowed. Hereafter, we write the neural network as $NN()$. For the completed the neural network, we get,

$$T_i=NN(Y_i), \qquad (5)$$

then,

$$T_{L-1}=NN(Y_{L-1}). \qquad (6)$$

Therefore, we obtain,

$$T_L\sim NN(Y_L). \qquad (7)$$

Here, considering the next relation,

$$Y_{L+1}=\{X_{L+1},X_{L+2},\ldots,X_n,X_{n+1}\}\sim\{X_{L+1},X_{L+2},\ldots X_n,T_L\}, \qquad (8)$$

the extrapolation can be calculated iteratively, as,

$$T_{L+1}\sim NN(Y_{L+1}),\ Y_{L+2},\ldots . \qquad (9)$$

This calculation scheme is called "slide window method", because Y_i is thought like slide window. The future can be accurately predicted by using the slide window method meeting the requirement that past observations include a factor by which the future is predicted.

The following idea has been proposed. Since Xn+1 is already observed when $T_{L+1}\sim NN(Y_{L+1})$ is calculated, $T_{L+1}\sim NN(Y_{L+1})$ is calculated by $Y_{L+1}=\{X_{L+1},X_{L+2},\ldots X_n,X_{n+1}\}$.
There is no necessary that the prediction is done by $Y_{L+1}\sim\{X_{L+1}, X_{L+2},\ldots,X_n,T_L\}$. Since a previous observation is taken into the iterations, the accuracy is higher. We call it "1-step prediction." The former is often called "long-term prediction" and the latter is called "short-term prediction."

3. TWO-DIMENSIONAL EXTRAPOLATION

3.1 One-dimension extrapolation

We consider the method to extrapolation the two-dimension function in L^2 times area $0<x<L,\ 0<y<L,$ when this function is assumed $z=f(x,y),\ 0<x<1,\ 0<y<1$. The one-dimension extrapolation slide window method can be defined along an arbitrary curve (A of Figure 1) in the two-dimension plane of Cartesian coordinate.

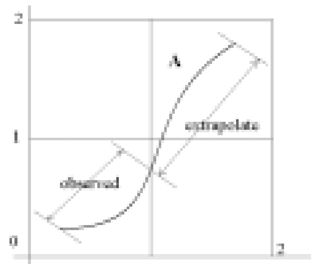

Fig.1 Learning and extrapolation section

The learning section is $0<x<1$ and $0<y<1$. We made [0,1]-interval scaling for learning data. Other spaces are the extrapolation sections (In this figure, to make it easy to see, it was assumed $L=2$. This value is meaningless and other values can be replaced with it).

We can define the one-dimensional slide window method along arbitrary ***curve-A*** in the above-mentioned. We can consider more curves, and calculate the extrapolation along them. That is a character of multi-dimensional slide-windows method. It depends what curve group only has to be applied on the character of the function (In this thesis, a straight line group being parallel to x - and y-axis is applied). Since the selection of a curve group relates to the extrapolation accuracy, it is necessary to investigate whether there is an extrapolation possibility by the selection. One investigation method is described.

The data observed along ***curve-A*** in the plane is assumed to be $\{A_0,A_1,...,A_n\}$. We defined subset, $Y_0=\{A_0,A_1,...A_m\}$, $Y_1=\{A_1,A_2,...A_{m+1}\},Y_2,...,Y_L$, which are used for slide-windows method. The window-width is "$m+1$"; so, it is a vector of $m+1$ elements, and it is a point in $m+1$ dimensional space. Here, we introduce a distance, D_{ij}. This distance progression, $D_{01},D_{12},D_{23},...D_{L-1,L},D_{L,0,}$ is plotted. Then, it becomes $D_{01}\sim D_{12}\sim D_{23}\sim....D_{L-1,L}$. That is, when you take suitable value C, it is $||D_{ij}-C||<\delta$. δ is very small value (<0.05). If $||D_{L0}-C||<\delta$, and necessary information for extrapolation is found in y-set, when the slide window method has a possibility of extrapolation. On the other hand, if $||D_{L0}-C||>\delta$, the possibility is lower as increasing of δ. However, since the neural network has second-order's extrapolation, functions are extrapolated in a short term. The precision of multi-dimensional extrapolation depends on appropriate selection of the ***curve-A*** in figure 1. The appropriateness relates to information in ***Y***-set.

3.2 Two-dimension extrapolation

Observation of plane function $z=f(x,y)$ are assumed to be procession

$Z=\{Z_{00},Z_{01},Z_{02},...,Z_{0n},Z_{10},Z_{11},Z_{12},...,Z_{1n},..., Z_{n0},Z_{n1},Z_{n2},...,Z_{nn}\}$

At this time, following subsets procession form Y_{ij} are made.

$Y_{00}=\{Z_{00},Z_{01},Z_{02},..,Z_{0m},Z_{10},Z_{11},Z_{12},..,Z_{1m},..,Z_{m0},Z_{m1},Z_{m2},..,Z_{mm}\}$,

$Y_{01}=\{Z_{01},Z_{02},Z_{03},..,Z_{0m+1},Z_{11},Z_{12},Z_{13}..,,Z_{1m+1}..,,Z_{m1},Z_{m2},Z_{m3}..,,Z_{m,m+1}\}$,

$Y_{10}=\{Z_{10},Z_{11},Z_{12},..,Z_{1m},$
$Z_{20},Z_{21},Z_{22},..,Z_{2m},..,Z_{m+1},0,Z_{m+1},1,Z_{m+1},2,..,Z_{m+1},m\}$,

...

$Y_{LL}=\{Z_L,L,Z_L,L+1,Z_L,L+2,..,Z_L,L+n,$
$Z_{L+1},L,Z_{L+1},L+1,Z_{L+1},L+2,..,Z_{L+1},n,..,Z_{nL},Zn,L+1,Z_n,L+2,..Z_{nn}\}$

Next, scalar T_{ij} corresponding to Y_{ij} is defined as follows.

$T_{00}=Z_{m+1},m+1,$

$T_{01}=Z_{m+1},m+2,$

$T_{10}=Z_{m+2},m+1,$

...

$T_{LL}=Z_{n+1},n+1$

$\{Y_{ij}\}$ and $\{T_{ij}\}$ are the learning data of the neural network. If it is not $||Y_{ij}-Y_{kl}||=0$, the learning is completed. It does as well as the one-dimensional slide window method as follows and slide window method of the two dimensions is possible (Figure 2).

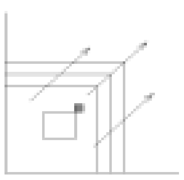

Fig.2 Concept figure of the two-dimensional slide window method

In the figure smallest quadrangle of which is painted is teacher data (scalar). Another quadrangle that has come in contact with the quadrangle is input data (procession). This both are one learning data of the neural network. The most internal square among of the triple quadrangles where the learning data has been included is a learning section. The outside the two quadrangles show the extrapolation section. The extrapolation section can be enhanced to infinity. The arrows show that the extrapolation outside is down outside one by one. It is unnecessary to select of two-dimension appropriate ***curve-A***, but one-dimension. In Figure 2, it was defined that the learning data is in position that is diagonal 1step ahead of the input and teacher data. This necessity doesn't exist, and any position is allowed. Moreover, input data need not be defined in a quadrangle area, and you may change it properly according to the character of the function. Therefore, two- dimension extrapolation that specializes in the direction of various function and extrapolation can be defined. We can define not only three-dimensional but also N-dimensional extrapolation as well as two-dimension.

4. Numeric calculation

To examine the effectiveness of one-dimensional and two-dimensional extrapolation with neural network, the extrapolation accuracy was calculated by valuable two-dimensional functions with different characters.

4.1 Learning area and neural network

We assumed that learning area is $0<x<2\pi$, $0<y<2\pi$, and scaled and defined that learning area is ***0<scaled x<1, 0<scaled y <1.*** We assumed that extrapolation area is $0<x<4\pi$, $0<y<4\pi$ (learning area is excluded) , and scaled as the above. In the case that the extrapolation could be done up to twice at learning (observation) period, the practicality of the method was judged enough. If the function type is uncertain, we hardly do extrapolation up to twice at the observation period in the extrapolation by usual function. We divided $[0,4\pi]$ section at 41 equal interval, and calculated the value by point and let references for the learning data of the neural network and the accuracy judgment of its calculation result. High accuracy simulation can be done as number of division N is large. On the other hand, because the calculation time increased by $N**3$, we let 41 be N as a compromise. The number of division in the learning section is $(41/2)+1=21$. Then, eight is a limit in the width of slide window. What the width of slide window is eight is too small in the case that the function changes rapidly, and it is necessary to make it be larger. But, to do so, the number of data (the number of divisions of learning sections- the width of window) used to learning becomes small. It causes the decrease in the forecast function.

4.2 Selection of two-dimension function

(1) Keep abreast of sine wave

Prototype slide window method extrapolates accurately the sine function. We examine how this character becomes in two-dimensional extrapolation. We made the sine function the two-dimension as follows.

$z=sin(a*x)+sin(a*y), a=5/4$ **(1)**

$z=sin(a*x)*sin(a*y)$ **(2)**

In this function the function of expression Eq.**(1)** becomes a sine function on a parallel straight line to x-axis and y-axis, the extrapolation of high accuracy is expected. Since the wave seems to do the keep abreast of movement parallel to x-axis and y-axis, so we said it the keep abreast of sine wave in this thesis. Though coefficient ***a=1*** if the neural network learns this sine wave for one period, we let ***a=5/4=1.25*** be it because of revision what the effective learning section becomes short (It is ***8/21=38%*** decrease in setting of this thesis) according as the amount of the width window. Moreover, Eq.***(2)*** is a kind of the Keep abreast of sine wave and its character is also same (Differ as the wave crest becomes like the checker).

(2) 16 keep abreast of Gaussian

Functions whose extrapolation is more difficult than the sine wave might be linear uniting of the Gaussian. The center ***C*** is in the Gaussian, so we defined as follow.

$r_{ij}^2=(x-C_{ijx})2+(y-C_{ijy})2$ **(3)**

C_{ijx} shows x coordinates at center C_{ij}. R_{ij} shows the distance of point ***(x,y)*** from center C_{ij}. Linear uniting ***z*** in the Gaussian is shown as follow.

$z=\Sigma_{ij}exp(-a*r_{ij}^2)$, $a=-0.1$ **(4)**

Coefficient a relates to the size of the distribution of the Gaussian. If the value of ***a*** is enlarged when extrapolation is examined, ***z*** value between centers of the Gaussian extremely becomes small. Then, the Gaussian becomes independent existence mutually, so the distribution of one Gaussian is learned and the distribution of the following Gaussian is not predictable. It is meaningless. Then, ***z*** value between Gaussian was graphed out, and ***a*** value was searched for suitable ***z*** value and was assumed 0.1. We chose 16 points like the lattice so that the center of the Gaussian covered the extrapolation section

$C_{ij}=\{ (-\pi,-\pi), (-\pi,\pi), (-\pi,3\pi), (-\pi,5\pi), (\pi,-\pi), (\pi,\pi), (\pi,3\pi), (\pi,5\pi),(3\pi,-\pi), (3\pi,\pi), (3\pi,3\pi), (3\pi,5\pi), (5\pi,-\pi), (5\pi,\pi), (5\pi,3\pi), (5\pi,5\pi)\}$ **(5)**

(3) Series Gaussian

$r_{ij}^2=(x-C_{ijx})^2+(y-C_{ijy})^2$

$z=\Sigma_{ij}exp(-a*r_{ij}^2)$, $a=-0.05$

$C_{ij}=\{ (-\pi,-\pi), (\pi,\pi), (3\pi,3\pi), (5\pi,5\pi)\}$

(4) Sine wave

$r_2=(x-C_x)2+(y-C_y)2$, $C=(\pi,\pi)$

$z=sin(r)$

4.3 The feature of each function

(1) The keep abreast of sine wave:

The one-dimension sine function has been only enhanced to the two-dimension. The learning section was made 5/4 cycles so that highly accurate extrapolation was possible.

(2) 16 keep abreast of Gaussian:

Though the shape of the function is similar to function (1), the shape of the slope is not more suitable for extrapolation than that of the sine function. We shortened the learning section has been shortened more than (1). This is more difficult extrapolation than (1).

(3) Series Gaussian:

Function that pulls out opposite element of function (2). From Figure 3, It is the case of difficult extrapolation in the one dimension slide window method. We selected it to see the effect of two dimension slide window method, it selected it.

(4) The sine wave:

It is the function that a ripple diffuses from center ***C*** to the concentric circle. And the function cannot be extrapolated by one-dimension slide window method along a parallel ***curve-A*** to x-axis and y-axis. In the two-dimension slide window method, we dare to set that the learning section is not enough for 1/2 period. So the extrapolation of this is a very difficult.

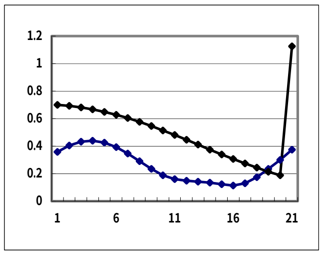

Fig.3 Examination of extrapolation possibility

An upper curve is a distance between learning (input) data of the series Gaussian. A lower curve is a distance of 16 keeps abreast of Gaussian. In the former, the right end part is separated. It is shown that it does not satisfy condition of 3.1 passage and the extrapolation possibility is now.

5. Calculation result

5.1 Keep abreast of sine wave

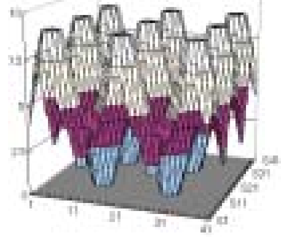

Fig.4 Contour line chart of function value of keep abreast of sine function

In this figure, the learning area is 1~21, and S1~S21 of the axis scales of in two directions. Scale 1 on the axis shows 0, and 41 shows 4π. In a true function value the scale is done ***[0,1]*** section, and the spindle is increased by a factor of ten.

This function is accurately extrapolated by one- and two- dimension slide window method. We calculated the error range between the extrapolation value and the true value with %, and plotted it in the plane. The spindle shows the size of the error range (%). The 1~21 and S1~S21 section of the bottom are the study learning areas, and the rest is the extrapolation areas. Scale 1 on the axis shows 0, and 41 shows 4π.

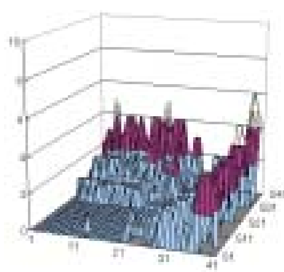

Fig.5 Error range of the one-dimension and short-term forecasts

The maximum error range is within 2%. It is accurately extrapolated.

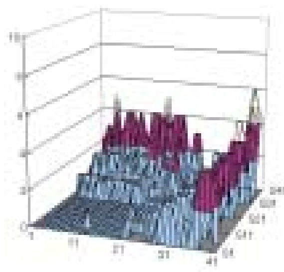

Fig.6 Error range of the one-dimension and long-term forecasts

The maximum error range is within 6%. The error range is growing at the edge point of extrapolation. The forecast increases to 2.8 times that of observation (learning) section in the edge point. It is clear that the error range is accumulated from the side in figure.

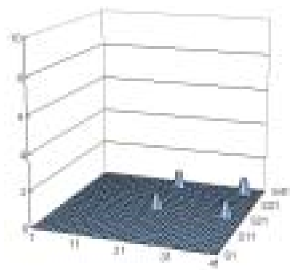

Fig.7 Error range of the two-dimension and short-term forecasts

The maximum error range is within 1%. It is more accurately extrapolated than one-dimensional extrapolation.

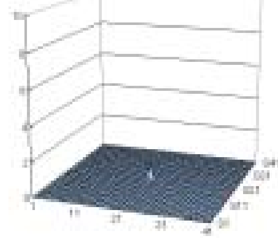

Fig.8 Error range of the two-dimension and long-term forecasts

The maximum error range is within 1%. It is more accurately extrapolated more than one-dimensional extrapolation. The error range doesn't accumulate in the long-term forecast because the learning section of this function is very large for 5/4 period. It is one sample that showed the effectiveness of two-dimension slide window method.

5.2 16 keep abreast of Gaussian

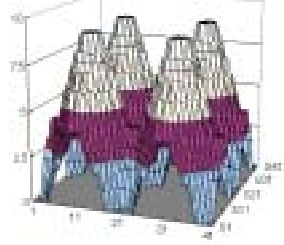

Fig.9 Contour line chart of function value of 16 keep abreast of Gaussian

In this figure, the learning area is 1~21, and S1~S21 of the axis scales of in two directions. Scale 1 on the axis shows 0, and 41 shows 4π. In a true function value the scale is done ***[0,1]*** section, and the spindle is increased by a factor of ten.

Though this function looks like the keep abreast of sine function this is linear uniting of the Gaussian with distribution originally in the limited part. Therefore, there is a problem that the extrapolation of another limited part Gaussian can be done again after the distribution of the Gaussian approaches 0. So, this is more difficult extrapolation than the keep abreast of sine function. We defined that the learning section was a place where just one Gaussian value was minimized. The extrapolation result of one- and two- dimensions slide window method of these 16 keep abreast of Gaussian (The error range with a true value is expressed with %) is shown. The spindle shows the size of the error range (%). The 1~21 and S1~S21 section of the bottom are the study learning areas, and the rest is the extrapolation areas. Scale 1 on the axis shows 0, and 41 shows 4**π**.

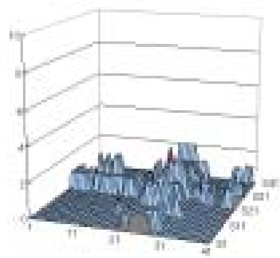

Fig.10 Error range of the one-dimension and short-term forecasts.

The maximum error range is within 3%. It is accurately extrapolated.

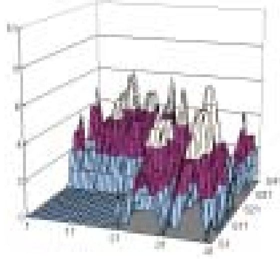

Fig.11 Error range of the one-dimension and long-term forecasts

The maximum error range is within 7%. There is no remarkable error in the learning section. However, in the extrapolation section the errors grow rapidly, and they are saturated about 6%. The errors near scale 31 on the horizontal axis are small, where the extrapolated value has changed negative from positive (or oppositely). Therefore, the errors are small superficially. The approximation level would be accepted on practical use though the accuracy is lower than that of sine-wave. Where, learning the slope shapes of a Gaussian, the neural network predicts another adjacent Gaussian on use of the shape-information. This is an example of ideal selections of ***curve-A***.

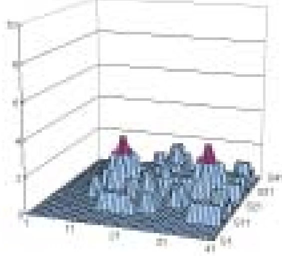

Fig.12 Error range of the two-dimension and short-term forecasts

The maximum error range is within 3%. From axis scale S1 to S8 is an area where extrapolation cannot be done because of the width of window of the two-dimension. It is quite unquestionable on practical use though accuracy has decreased more slightly than the one-dimension short-term projection. Generally, it is hardly thought that the two-dimension slide window method accuracy decreases more than the first former extrapolation in this examination, because the selection of standard curve for one-dimensional extrapolation is ideal , use guessed that this happened.

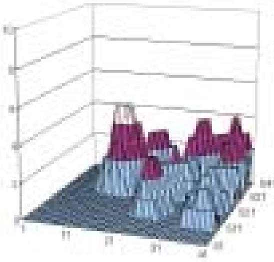

Fig.13 Error range of the two-dimension and long-term forecasts

The maximum error range is within 5%. There is no error range in the learning section (The neural network reproduces the Gaussian well). However, in the extrapolation section the error range rapidly grows, and it doesn't grow any more when it reaches about 4%. The error range about at scale 31 on the axis seems to be small, because the extrapolation value has changed negative from positive (or oppositely) and it seems that the error range has become small superficially. This is more accurately extrapolated than the result of the one-dimension slide window method (Figure 11). An advantage of extrapolation with neural network is the point that the error range doesn't increase rapidly.

5.3 Gaussians located on series

We examined the one-dimensional slide window method in difficult case of selection ***curve-A***. The most simplified example is to put Gaussians diagonally, and the ***curve-A*** is a line that is parallel to x- or y-axis. Figure 14 shows the Gaussians.

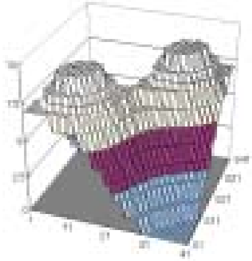

Fig.14 Contour line chart of t he series Gaussian

In this figure, the learning area is 1~21, and S1~S21 of the axis scales of in two directions. Scale 1 on the axis shows 0, and 41 shows 4π. In a true function value the scale is done ***[0,1]*** section, and the spindle is increased by a factor of ten. This is a difficult extrapolation problem because there are no information to predict a deep valley in the vicinity of X-axis scale 40 and Y-axis scale S1. The extrapolation in the direction of the diagonal is easy, but it is not extrapolated when a diagonal carve A standard cannot be discovered. Since the error range of this extrapolation was large, we showed the extrapolation value of the neural network in figure.

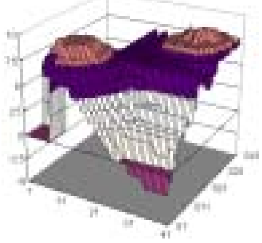

Fig.15 Values of the one-dimension and short-term extrapolation

The extrapolate shape is deformed, and predictions through diagonal directions are so wrong. It is caused by insufficient learning of Gaussian along the diagonal line. Moreover, the neural network predicts very deep valley near scale 40 on y-axis, which is deeper than the real. The error range increases rapidly in the extrapolation section. This situation is a limit of one-dimension slide window method without finding an appropriate ***curve-A*** standard. In the one-dimension and long-term extrapolation the error range got larger the shape of the unlearning Gaussian cannot be reappeared, and the depth of the valley was predicted to be –10 or more (the true value is 0).

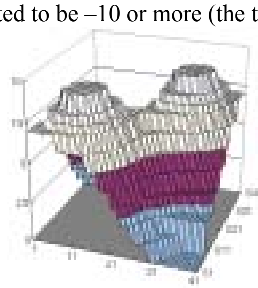

Fig.16 Values of the two-dimension and short-term extrapolation

The learning Gaussian for diagonal one from learning section is predicated in accuracy good. Its shape is also correctly forecast. The valley in the vicinity of x-axis scale 40 and y-axis scale S1 is correctly predicated also as for the depth. It is an example that two-dimensional slide window method effectively works.

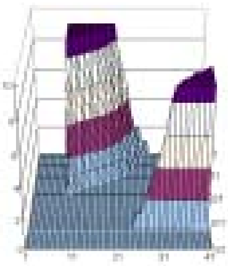

Fig.17 Error range of the two-dimension and short-term forecasts

This figure corresponds to Figure 16. In order to make the situation of the error range easy to judge, we rotated the aspect and drew it. The maximum error range exceeds 10%. In some areas from S1 to S8 1--8, the extrapolation cannot be done because of the width of window of the two-dimension. It is understood that the error range occurred in the part in the valley on both sides.

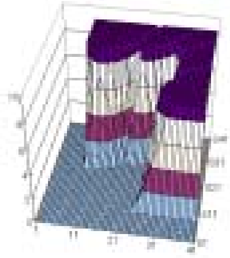

Fig.18 Error range of the two-dimension and long-term forecast

The maximum error range exceeds 9%. In some areas from S1 to S8 1--8, the extrapolation cannot be done because of the width of window of the two- dimension. The error range is 0 in the learning section. The range where prediction is possible with two-dimension slide window method is only 3--6 step ahead from the learning section. The forecast section has somewhat extended to a diagonal interior.

5.4 Sine wave

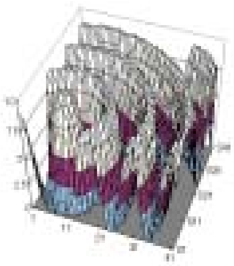

Fig.19 Contour line chart of sine wave

The learning area is a range of 1~21 and S1~S21 of the axis scales of figure in two directions. Scale 1 on the axis shows 0, and 41 shows $20\pi/3$. In a true function value, the scale is done in [0.1] section and the spindle scale is increased by a factor of ten. The learning sections area two inside mountains. Therefore, information for extrapolation exist the learning section.

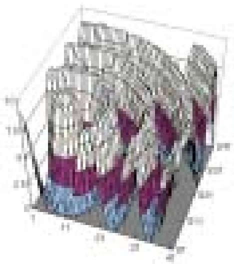

Fig.20 Values of the two-dimension and short-term extrapolation

We indicated the extrapolation function value. Ten function values are multiplied. A true function value is [0.1], and the extrapolation value is also the same.S1~S21 section of the bottom are the learning areas and the rest is the extrapolation areas. Scale 1 on the axis shows 0, and 41 shows $20\pi/3$. The crumble of the function type was not seen in the learning section, and the outside wall was forecast comparatively well also in the extrapolation section. From these, the advantage of the two-dimension slide window method is clear.

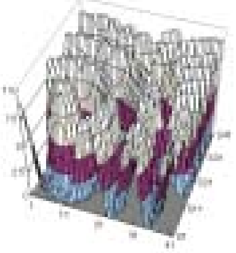

Fig.21 Values of the two-dimension and long-term extrapolation

We indicated the extrapolation function value. Ten function values are multiplied. A true function value was [0.1], and the extrapolation value became [-10, 10]. The prediction of most outside circular wall swerved greatly into negative direction. We cut 0 or less in this figure.
The 1~21 and S1~S21 section of the bottom are the learning areas, and the rest is the extrapolation areas. Scale 1 on the axis shows 0, and 41 shows $20\pi/3$. The crumble of the function type was not seen in the learning section, and the outside wall, especially the third wall, was forecast to some degree also in the extrapolation section.

6. CONCLUSION

On other statistical discussions, it is a known fact that multi-dimensional prediction is not equal to one-dimensional. The same fact is found on slide-window method with multi-layer neural networks. We researched two predictors in two-dimensional space, which are line- and local-plain-predictors. The both predictors are equivalent mathematically. However, in multi-dimensional inter/extrapolation, it is very important to find the direction of predictions. Especially, since the slide window method requires fragments to predict the future in sampling data, if they are not ordered appropriately along the direction, the predictor cannot operate accurately. We tested the as for the extrapolation for typical two-dimensional functions, and found an excellent character of slide-window method based on local-plain. That is, the local-plain method searches a reasonable direction automatically. If we could understand the direction beforehand, the line-method has a possibility to get accurate results. However, in practical usages, it is a rare case. The local-plain method is derived to three-dimensional expression easily. By using the method, we can extrapolate the function until twice-outer regions of the definitions. Thus, we believe that the multi-dimensional slide-window method is a useful predictor.

REFERENCES

[1] H.Zhu, T. Teshima, T. Aoyama, I. Yoshihara, "Long range forecasting for chaos by using multi-layer neural networks: Ikeda's chaos", *Proc. of 5th International Symposium on Artificial Life and Robotics*, Vol.5, pp.543-546, 2000.1.26.

[2] T. Aoyama, H. Zhu, I. Yoshihara, "Forecasting of the chaos by iterations including multi-layer neural-network", *Proc. of International Joint Conference on Neural Networks'2000,* CD-ROM (71_04.pdf), 2000.7.24-27.

[3] H. Wakuya and J. M. Zurada "Bi-directional computing architecture for time series prediction", *Neural Networks*, Volume 14, Issue 9, November 2001, Pages 1307-1321.

[4] S.Mase and J.Takeda, "Spatial Data Modeling: applications of the spatial statistics (in Japanese)" *Kyoritsu publishing Co. Ltd.*, (1999, Tokyo), ISBN4-320-12006-X