A Fundamental Study of Neural Network System for Conjecturing Weather in Local Area Using GPS

Shigenori Tanaka*, Hitoshi Furuta*, Michiyuki Hirokane*
Tomonori Takada**, Tatsunori Sada**, and Kousei Uchino***

Abstract: The GPS is used in surveying recently. The results of GPS surveying contain errors due to the delay of GPS radio waves caused by vapor in the atmosphere. Developed in this study is a system which forecasts local weather by making use of GPS data under the influence of the vapor. The neural network technology was used to relate GPS data to weather conditions. In a preliminary experiment, the simulation of weather forecasting was carried out to select GPS data items to be used in the main experiment. A pilot system was constructed on the GPS data items of the preliminary experiment. The pilot system forecast sunny and rainy weathers correctly. Then the main experiment was carried out, the results of the preliminary one taken into account. In the main experiment, meteorological factors in addition to GPS data were used to construct a forecasting system, which forecast all weathers correctly except a mixture of sunny and cloudy weathers.

Keywords: Short-Range Weather Forecast, Local Area, GPS, Neural Network, and Identification Problem

1. Introduction

The GPS (Global Positioning System) is a satellite navigation system for ships and airplanes to find their own locations. It has recently come to be used in surveying in particular [1]. Errors due to the delay of GPS radio waves are inherent in the results of GPS surveying. Factors of observation errors are the locations of satellites, the errors of receivers and those of clocks on satellites, and the delay of radio waves due to the ionosphere, multipaths, vapor in the atmosphere, etc. All the factors but the vapor in the atmosphere can be removed to a certain degree. Because the vapor in the atmosphere is in a phantasmagoric dynamic state which cannot completely be grasped with the present technology, it is difficult to remove the errors due to the vapor-caused delay of radio waves [2]. On the other hand, research is being made to observe the vapor in the atmosphere by using the GPS[3].

The object of the present study is the development of a local short-range weather forecast system which makes use of the vapor-caused errors inherent in observation data by the GPS [4] and the neural network technology.

A preliminary experiment was carried out to search for the possibility of constructing a GPS weather forecast system. Then the main experiment was conducted to develop a

* Faculty of Informatics, Kansai University (2-1-1, Ryouzenji, Takatsuki, Osaka 569-1095, Japan)
** Mitsui Construction Ltd., (518-1, komagi, nagareyama, Chiba 270-0132, Japan)
*** Graduate School of Informatics, Kansai University (2-1-1, Ryouzenji, Takatsuki, Osaka 569-1095, Japan)

practical weather forecast system. The neural network technology was used to relate GPS data to weather conditions.

2. Preliminary Experiment

A pilot system was constructed to make local short-range weather forecasts based on GPS data, as shown in Fig.1.

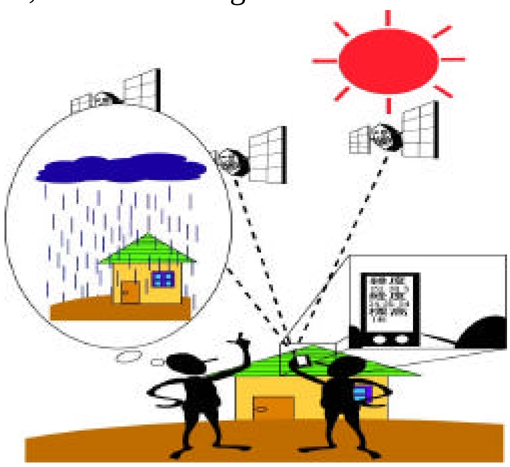


Fig.1 Schematic Illustration of Study

2.1 Plan of Preliminary Experiment

It was necessary for local short-range weather forecasting to find GPS data items which effectively relate the errors in GPS data to weather. Accordingly a preliminary experiment was carried out, of which the process is shown in Fig.2.

GPS surveying was carried out at the same two points at the same time for six months. The errors of, or the dispersion in, the measured values of the slope distance between the two points caused by the delay of radio waves due to the vapor in the atmosphere was related to the local weather by the neural network technology to construct the pilot system for local short-range weather forecasting. Going back to the times of surveys, the local weather was forecasted with the pilot system to check its performance.

Data were collected by surveying for one hour from 2:00 p.m. to 3:00 p.m. every day for six months. The slope distance between the two points, the base-line vectors dx, dy, dz, and dh, the ratio indicating the quality and accuracy of the slope distance, the satellite deployment index PDOP, and the local weather were recorded. The pilot system was developed by using these data.

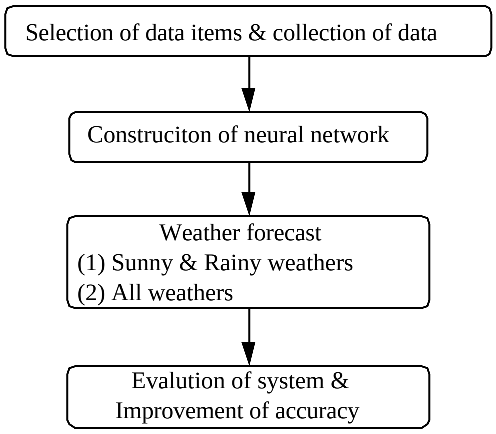


Fig.2 Process of Preliminary Experiment

2.2 Construction of Pilot System

The input items (slope distance between the two points, base-line vectors, ratio, and PDOP) and the output item (local weather information at the time of surveying) were related to each other by the neural network technology. All the input data were normalized, their value being 0 to 1. As shown in Fig.3, weather conditions were defined by five-grade fuzzy quantity. A three-layer network was adopted. The input layer was of seven units and the output layer was of one unit.

Used in the preliminary experiment were 121 observation data which were collected during half a year from November 11, 1998 to May 26, 1999. The numbers of data were standardized through all the weather conditions [5] to construct the pilot system. Table 1 shows the number of data of each weather condition.

Because only seven data were collected for rainy weather, the numbers of data of the other weather conditions were standardized to seven. For each weather condition, the deviations of the data from their mean were calculated and seven data with the smallest to seventh smallest deviations were chosen.

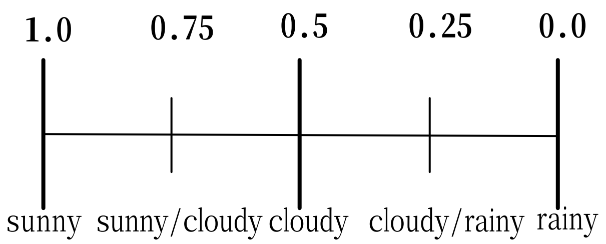


Fig.3 Fuzzy Quantities of Weather

Table 1 Numbers of Data of Weather Conditions

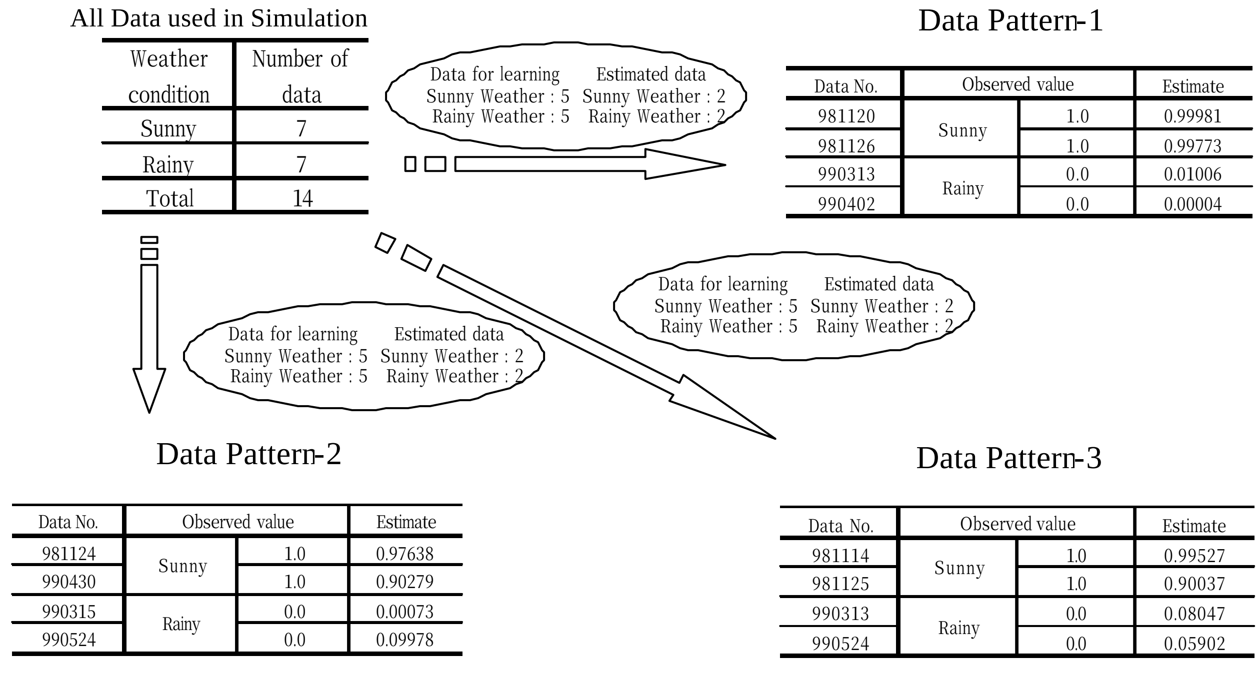
Weather	Collected data	Adopted data
Sunny	39	7
Sunny/cloudy	36	7
Cloudy	30	7
Cloudy/rainy	9	7
Rainy	7	7
Total	121	35

2.3 Simulation of Weather Forecasting

2.3.1 Sunny and Rainy Weather Conditions

Simulation of weather forecasting was carried out with only the data for sunny and rainy weather conditions. Because the number of data of every weather condition were standardized to that of the rainy weather condition, the numbers of data of all weather conditions for the simulation were limited to only seven. Accordingly three sets of data patterns, each consisting of four pairs of data for learning and data estimated, were randomly chosen to check the reliability of the pilot system.

As shown in Fig.4, the pilot system forecast the local weather nearly perfectly based on the GPS data in the simulation. Thus, the relation between the GPS data and sunny and rainy weather conditions was established.



All Data used in Simulation

Weather condition	Number of data
Sunny	7
Rainy	7
Total	14

Data Pattern-1

Data No.	Observed value		Estimate
981120	Sunny	1.0	0.99981
981126		1.0	0.99773
990313	Rainy	0.0	0.01006
990402		0.0	0.00004

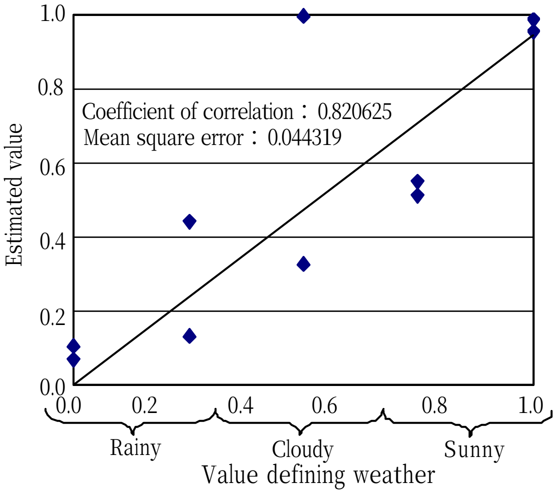
Data Pattern-2

Data No.	Observed value		Estimate
981124	Sunny	1.0	0.97638
990430		1.0	0.90279
990315	Rainy	0.0	0.00073
990524		0.0	0.09978

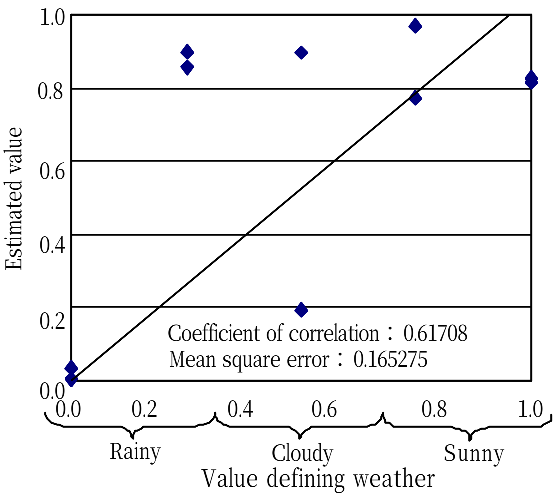
Data Pattern-3

Data No.	Observed value		Estimate
981114	Sunny	1.0	0.99527
981125		1.0	0.90037
990313	Rainy	0.0	0.08047
990524		0.0	0.05902

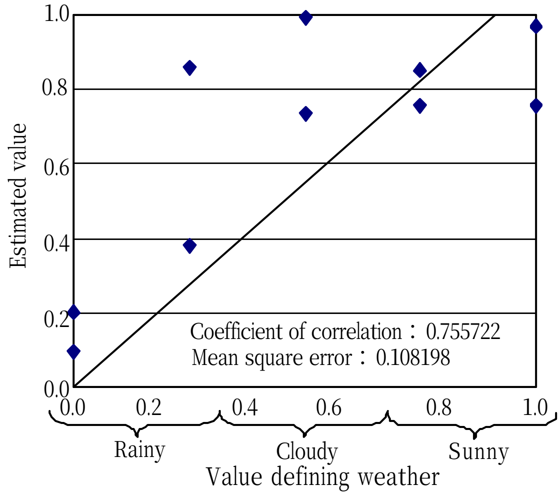
Fig.4 Results of Weather-Forecasting Simulation; Sunny and Rainy



(a) Data Pattern-1



(b) Data Pattern-2



(c) Data Pattern-3

Fig.5 Results of Weather-Forecsting Simulation; All Weather Conditions

2.3.2 All Weather Conditions

The pilot system was reconstructed to handle all the five weather conditions. Again three sets of data patterns were chosen to check the reliability of the pilot system.

Fig.5 shows the results of weather-forecasting simulation by the system. The ordinate represents the fuzzy quantities estimated; the abscissa, those defining weather conditions. The coefficient of correlation and the mean square error are also shown.

Regarding sunny and rainy weather conditions, the simulation turned out to be rather satisfactory. With respect to the other weather (sunny/cloudy, cloudy, and cloudy/rainy) conditions, correct weather forecasting was hardly made.

2.4 Discussion on Results of Preliminary Experiment

The pilot system constructed in the preliminary experiment proved itself to be capable of forecasting local weather with respect to sunny and rainy conditions. Regarding sunny/cloudy, cloudy, and cloudy/rainy conditions, however, the system failed to make acceptable forecast. To raise the forecasting accuracy of the system, the following matters had to be addressed.

- More data had to be collected.

- The base-line in the preliminary experiment was approximately 100 m in length, and the dispersion of its measured vales was very small. The difference in elevation between the two points, to which the error due to vapor is attributed, was approximately 20 m. Accordingly, desirable was to fix a base-line with as large difference in elevation as possible so that the effect of vapor can be observed easily.

- The base-line length was determined from collected data by correcting the errors in the troposphere to a certain degree in the analysis stage. The pilot system was constructed by using the corrected length of the base-line. It was considered that the accuracy of the system could be increased further by not correcting the base-line length.

- It seemed difficult to achieve enough forecasting accuracy based on GPS data alone. Accordingly other factors such as temperature, atmospheric pressure, humidity, and direction and velocity of wind had to be considered.

3. Main Experiment

The main experiment was carried out based on the results of the preliminary experiment to construct a practical system for forecasting local weather in a short range.

3.1 Plan of Main Experiment

Fig.6 shows the process of the main experiment. When a system is built by using a neural network, the quality and the quantity of data determine the accuracy of the system. Accordingly GPS surveying was again conducted at the same two points at the same time for half a year. Data were collected, the problems awaiting solution found in the preliminary experiment considered, attention paid to the seven points below.

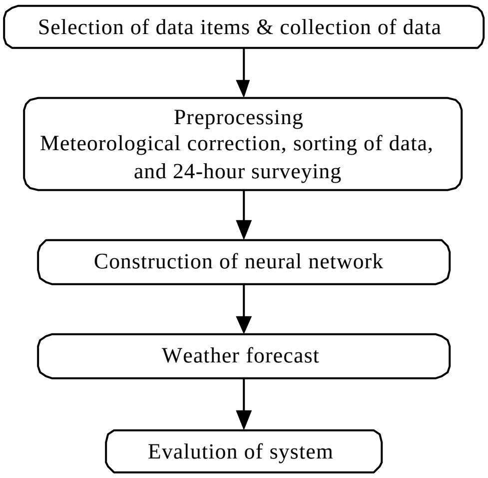


Fig.6 Process of Main Experiment

- Surveying and observation were made for one hour from 9:00 a.m., from Monday to Friday, for construction sites demand weather reports early in the morning.

- Surveyed was a base-line with length of approximately 600 m and difference in elevation of 70 m to determine the errors due to vapor.

- The PDOPs at the two survey points were recorded. Although the PDOP value does not vary largely because the deployment of the satellites have been completed by now, it changes in accordance with the always-changing geometrical positions of satellites. Therefore, the PDOP was recorded at 8:59, 9:00, and 9:01 a.m. The mean of the recorded PDOP values was used in the construction of the system.

- Meteorological factors such as temperature, atmospheric pressure, humidity, and direction and velocity of wind were recorded seven times, every minute from

9:15 a.m. to 9:21 a.m. The mean of the seven recorded values for each factor was used in the construction of the system.

- Weather was observed for 30 minutes before and after the observation time, and also at 3:00 p.m. and 9:00 p.m.

- The weather conditions were classified into 13 categories, as shown in Table 2, instead of 5 for the preliminary experiment.

- Environmental conditions were recorded in detail.

Table 2 Weather Classification

Weather at the time of data collection	
(1) Fine	(2) Sunny
(3) Sunny, cloudy later	(4) Cloudy, sunny later
(5) Sunny, occasionally cloudy	(6)Cloudy, occasionally sunny
(7) Cloudy	
(8) Cloudy, rainy later	(9) Cloudy, occasionally rainy
(10) Rainy, cloudy later	(11) Rainy, occasionally cloudy
(12) Rainy	(13) Heavy rain

Table 3 Outline of Preliminary and Main Experiments

Item	Preliminary	Main
Base-line length	Approx. 121 m	Approx. 642 m
Data-collection Period	Nov. 11, '98 to May. 26, '99	Sep. 8, '99 to Feb. 28, 2000
Total number of data	About 121	About 80
Data items	Slope distance dx of base-line vector dy of base-line vector dz of base-line vector dh of base-line vector Ratio PDOP Weather at surveying	Slope distance dx of base-line vector dy of base-line vector dz of base-line vector dh of base-line vector Ratio PDOP at point 1 PDOP at point 2 Temperature Atmospheric Pressure Humidity Wind velocity Weather at surveying

Table 4 A Part of the Observation Data

Number of data	Slope distance	Ratio	dx	dy	dz	dh	PDOP		Temperature	Atmospheric	Humidity	Wind velocity
							point1	point2	%	hpa	°C	m/s
990908	642.667	5.0	465.905	380.849	225.626	74.713	5.9	2.4	70.9	984	27.0	1.8
990909	642.664	5.7	465.908	380.841	225.624	74.706	2.5	2.6	63.5	985	25.5	1.6
990910	642.666	13.0	465.912	380.839	225.627	74.705	2.5	2.5	59.5	984	26.5	0.7
990913	642.668	14.2	465.904	380.850	225.631	74.719	2.4	2.3	70.0	986	27.0	0.7
990914	642.673	21.4	465.914	380.846	225.629	74.709	2.4	2.3	79.4	986	25.2	0.9
990916	642.674	10.2	465.908	380.852	225.635	74.719	2.5	2.5	69.7	994	23.0	0.1
990917	642.674	14.8	465.911	380.846	225.639	74.715	2.5	2.5	64.1	990	25.5	0.9
990921	642.674	12.0	465.913	380.843	225.641	74.714	2.8	2.8	85.0	986	24.5	0.1

Table 5 A Part of Observation Data for Delete

Number of data	Slope distance	Ratio	dx	dy	dz	dh	PDOP		Temperature	Atmospheric	Humidity	Wind velocity
							point1	point2	%	hpa	°C	m/s
991022	642.752	3.9	465.546	381.213	225.996	75.347	2.7	2.7	65.7	1001	26.5	0.4
991125	658.752	6.9	414.271	431.445	276.023	162.839	2.7	2.5	84.0	987	13.5	0.6
991220	642.677	22.0	465.881	380.874	225.663	74.763	—	2.0	55.9	1000	2.6	0.9

Paying attention to the above seven points, data were collected. Data items for the preliminary experiment and those for the main one are compared in Table 3, Table 4 shows the part of the collected data.

3.2 Preprocessing of Data Adopted

For smooth learning processing in a neural network system, it is necessary to construct the system after data are processed [6]. Made in the main experiment was the preprocessing of data sorting, cancel of meteorological correction, and correction for the positional characteristics of GPS satellites.

3.2.1 Sorting of Data

While the data were being sorted, found were some data items of which no data had been collected or for which incorrect data had been collected. For example, because the values of some ratios were too small in the base-line analysis of data, the integer bias could not be determine and, hence, the optimum solution could not be found, or the result of the base-line analysis of some data differed from the mean by ten odd meters. Accordingly incomplete and defective data, as shown in Table 5, were not used in the main experiment.

3.2.2 Cancel of Meteorological Correction

The correction of GPS radio waves in the troposphere was incorporated in the base-line analysis software used in this study [7]. It is called Hopfield's model (a meteorological model). The correction in the troposphere by Hopfield's model was cancelled in this study. The results of base-line analysis were influenced by the vapor in the troposphere.

Fig.7 shows the measured lengths of the base-line with and without the application of Hopfield's model. Obviously, the effect of vapor can be grasped better without the application of the model. Accordingly the data without correction were adopted in the main experiment.

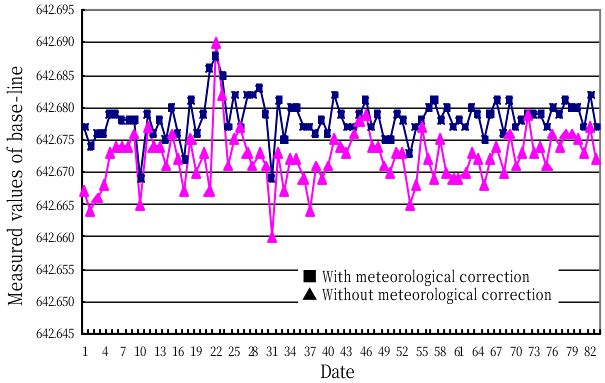


Fig.7 Effect of Meteorological Correction

3.2.3 Correction for Positional Characteristics of GPS Satellites

Although the orbiting period of the GPS satellites is known as 0.5 sidereal days [8], there occurs a slight delay. They delay four minutes a day, or two hours a month. The error in the length of a base-line in GPS surveying is attributable to the positions of the satellites among other factors. By removing the errors due to the positional characteristics of the satellites, values of base-line length containing larger errors due to vapor can be obtained.

However, the errors due to the positional characteristics of the satellites cannot be determined accurately. Accordingly a 24-hour survey was carried out on a reference day to make a model of the variation of length of the base-line and reduce the errors in the length due to the positional characteristics.

The 24-hour survey was carried out on

Table 6 A Part of the Result Survey

12/13	Slop distance	Ratio	dh	Base month	12/13～12/14	Slop distance	Ratio	dh	Base month
AM09:00	642.679	20.7	74.795	Dec. '99	PM21:00	642.682	11.4	74.794	Jun. '00
AM10:00	642.678	6.3	74.797		PM22:00	642.682	4.1	74.775	
AM11:00	642.682	20.6	74.800	Jan, '00	PM23:00	642.681	11.7	74.798	Jul. '00
AM12:00	642.681	21.9	74.799		PM24:00	642.681	27.2	74.800	
PM13:00	642.680	11.8	74.794	Feb, '00	AM01:00	642.683	13.1	74.792	Aug. '00
PM14:00	642.685	5.2	74.806		AM02:00	642.682	7.4	74.793	
PM15:00	642.677	8.1	74.796	Mar, '00	AM03:00	642.681	23.0	74.798	Sep. '00
PM16:00	642.683	15.6	74.796		AM04:00	642.680	30.4	74.800	
PM17:00	642.685	5.4	74.794	Apr, '00	AM05:00	642.680	9.7	74.799	Oct. '00
PM18:00	642.681	27.7	74.796		AM06:00	642.686	4.9	74.810	
PM19:00	642.680	7.8	74.801	May, '00	AM07:00	642.681	12.4	74.799	Nov. '00
PM20:00	642.683	11.7	74.802		AM08:00	642.680	24.6	74.794	

December 13, 1999. The data collected at 9:00 a.m. were regarded as the data of December 1999; those collected at 11:00 a.m., as those of January 2000; those collected at 1:00 p.m., as those of February 2000; and so on, as shown in Table 6. Thus the reference value of the base-line length for each month was determined.

The difference between the measured length of the base-line of each day in a month and the reference length for the same month was used in constructing the weather forecast system, as shown in Table 7.

Table 7 Data for Construction of System

	Slope length	dx	dy	dz	dh
8.9.'99[*1]	642.667	465.905	380.849	225.626	74.713
9.'00[*2]	642.681	465.861	380.895	225.681	74.798
Differ.	0.014	0.044	0.046	0.055	0.085

[*1]Measured value on September 8, 1999

[*2]Reference value for September, 2000

3.3 Construction of System

The input items (slope distance, dx, dy, dz, dh, ratio, PDOP, temperature, atmospheric pressure, humidity, and wind velocity all measured) and the output item (weather information at the time of surveying) were related to each other by the neural network technology.

About 80 data were collected for half a year from September 8, 1999 to February 28, 2000. Table 8 shows the number of data for each weather category.

Table 8 Number of Data for each Category of Weather Conditions

Weather conditions	Number of measured data
Sunny	38
Sunny / cloudy	14
Cloudy	16
Cloudy / rainy	8
Rainy	3
Total	79

The weather forecast system constructed in this study consists of four subsystems as shown in Fig.8. Weather was represented by five-stage fuzzy quantity shown in Fig.9. The 13 categories of weather conditions observed at the time of collecting data were reclassified into these five fuzzy quantities. The first subsystem was for sunny weather; second, sunny/cloudy; third, cloudy; and fourth, cloudy/rainy. Each data set was first evaluated in the first subsystem. If it was not for sunny weather, it was forwarded to the second subsystem. If it was not for the sunny/cloudy weather, it was passed to the third subsystem. If it was not for cloudy weather, it was passed to the fourth subsystem (Fig.10). If it was not for the cloudy weather, it was considered to be for rainy weather.

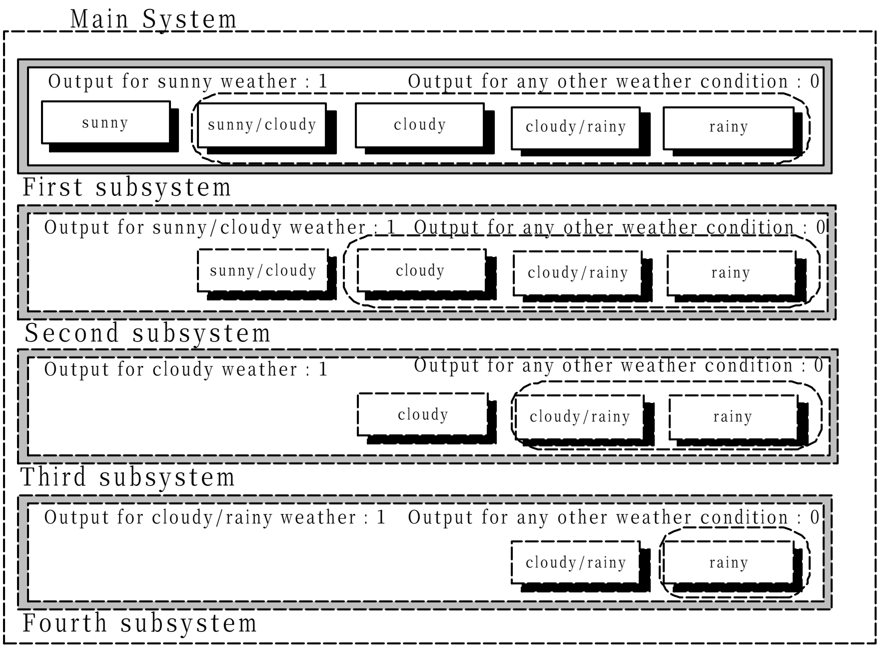


Fig.8 System Configuration

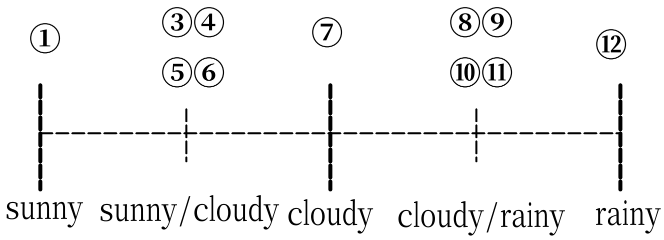


Fig.9 Fuzzy Quantities of Weather

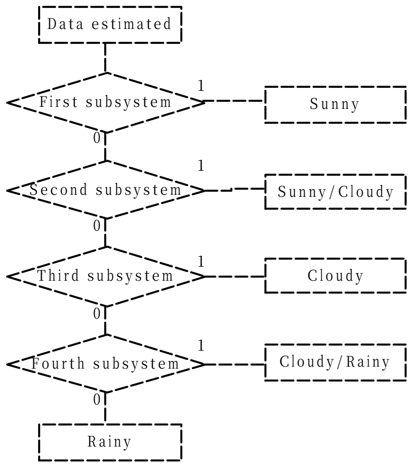


Fig.10 Operating Procedure of System

3.4 Result of Analysis

Weather was forecast in accordance with the operating procedure of the system. Table.9 shows the results.

As shown in Table 9, weather was forecast accurately in all the subsystems but one for sunny/cloudy weather. With a very small mean square error, the subsystem for sunny weather in particular had a high forecasting accuracy.

3.5 Discussion on Results of Main Experiment

In the main experiment, a local collected forecast system was developed by using collected data. The estimated values for cloudy weather and rainy weather were unsatisfactory and those for the other weather conditions were satisfactory. The first subsystem in particular proved itself to be highly accurate for all the weather conditions.

Sunny/cloudy, cloudy, and cloudy/rainy conditions could not be forecast in the preliminary experiment, whereas these weather conditions could be accurately forecast in the main experiment because meteorological factors were taken into account. The forecast with respect to sunny/cloudy weather and cloudy/rainy weather was unsatisfactory, though better than in the preliminary experiment.

Table 9 Estimated Value

Number of data	Observed value	Subsystem for sunny		Subsystem for sunny/cloudy		Subsystem for cloudy		Subsystem for cloudy/rainy	
		Observed	Estimated value	Observed	Estimated value	Observed	Estimated value	Observed	Estimated value
990921	cloudy/rainy	0.0	0.00522	0.0	0.00355	0.0	0.01400	1.0	0.94470
000207	cloudy/rainy	0.0	0.00088	0.0	0.00395	0.0	0.22413	1.0	0.97204
000121	rainy	0.0	0.19197	0.0	0.98573	0.0	0.16938	0.0	0.29901
991209	cloudy	0.0	0.00438	0.0	0.55669	1.0	0.99731	-	-
000112	cloudy	0.0	0.00775	0.0	0.32186	1.0	0.99703	-	-
990909	sunny/cloudy	0.0	0.00771	1.0	0.99836	-	-	-	-
990908	sunny/cloudy	0.0	0.03273	1.0	0.95581	-	-	-	-
991001	sunny	1.0	0.99955	-	-	-	-	-	-
991216	sunny	1.0	0.99989	-	-	-	-	-	-
Mean square error		0.004232296		0.198163557		0.01582718		0.031082277	

4. Conclusions

In this study, a preliminary experiment was first conducted to correlate GSP data with weather conditions. The correlation was successful with respect to sunny and rainy weathers, but no meaningful correlation was made with respect to cloudy weather. In the main experiment, data in a wider range were collected, and a weather forecast system was developed. Measures were taken to grasp the maximum delay of GPS radio waves caused in the troposphere, and the system was constructed. The system forecast all weathers correctly except rainy and cloudy ones.

It was ascertained in this study that local weather can be forecast in a short range with GPS data. To make the system developed in this study practical, however, its forecasting accuracy has to be improved further. The authors will raise the accuracy by using the data of the Geographical Survey Institute of the Ministry of Construction. Besides the authors will study further to develop a method of forecasting weather over a longer range. This study was subsidized by the program, “Graduation Studies 1999,” of the Faculty of Informatics of Kansai University.

References

1) Tsuchiya, A. and Tsuji, H., “Easy, Simple GPS Surveying”, Surveying Association of Japan, 1997.

2) Thierry, G. and Blewitt, G. “The Effect of Weather Fronts on GPS Measurements”, GPSworld, 1998.

3) Naito, K., “GPS Meteorology; Measurement of Water Vapor”, Parity, 1998.

4) Otani, R., ”Application to GPS Meteorology”, Navigation Society of Japan, 1999.

5) Tanaka, S., “Practice of Knowledge and Information Processing for Construction Engineers”, Publishing Department, Kansai University, 1999.

6) Funabashi, M., “Introduction to Neurocomputer”, Ohm, 1992.

7) “WAVE Software User Guide”, Trimble Japan, 1996.

8) “Satellites-based Precision Positioning System”, Japanese Society for Geodesy, 1989.