



Real-Time Tornado Forecasting Using SLHGN

Benny Benyamin Nasution^(✉), Rahmat Widia Sembiring,
Muhammad Syahrudin, Nursiah Mustari,
Abdul Rahman Dalimunthe, Nisfan Bahri, Berta br Ginting,
and Zulkifli Lubis

Politeknik Negeri Medan, Medan 20155, Indonesia
benny.nasution@polmed.ac.id

Abstract. The architecture of mHGN has been improved and become Single Layer Hierarchical Graph Neuron (SLHGN). The speed of this new architecture for recognizing multidimensional patterns is faster than the one of mHGN. It is therefore more suitable for forecasting multidimensional and complex process of tornado's genesis in real-time. Additionally, two important issues related to data handlings of non-accurate recorded data and data handlings of complex weather data have been solved. These improvements have given significant and positive quality of SLHGN in forecasting tornado. Although the accuracy and the forecasting performance cannot be calculated properly, due to the fact that weather data is not always available, the specific characteristics of the SLHGN experiment results show very promising values. This results suggest that tornado can be forecasted at least 5 h before it occurs. People in the to-be-hit area will then have adequate time to be evacuated or to escape. The deployment of SLHGN in risky areas of tornados can then be expected as a tool for reducing damages, losses, and costs. Several improvements in weather station distribution still need to be carried out in order to improve the quality of tornado forecasting using SLHGN.

Keywords: Hierarchical Graph Neuron (HGN) · Multidimensional Hierarchical Graph Neuron (mHGN) · Single Layer Hierarchical Graph Neuron (SLHGN) · Natural disaster forecast · Tornado forecast

1 Introduction

It has been suggested that tornadoes are one of the most powerful weather events associated with destructive forces of nature. The frequency of occurrence of tornadoes is highest in North America especially in the US [1, 2]. Canada is second only to the US and approximately 80 occurrences are reported annually [3]. The ones that occurred in several countries across the gulf part of American continent are clear evidences that those disasters are real, and they will occur regularly [4]. Some evidences even show that tornados may occur in unusual locations. October 6, 2010, is the day when tornadoes occurred in Arizona. This day shouts out loud and clear that tornadoes indeed happen in the Grand Canyon State. On this day, Arizona experienced its largest, single-day tornado outbreak in its recorded history [5].

Major challenges toward improving the warning operations include obtaining observations of severe weather for real-time forecasting and post-event verification,

educating forecasters, and having access to state-of-the-art forecaster workstations. An additional challenge is in communicating anticipated or ongoing severe tornados, both internationally between national weather bodies such as National Hydro-Meteorological Services (NHMSs), and nationally with media and emergency authorities [6]. This effort is also important to increase the accuracy of the forecasting. The reason why the occurrence of natural disasters cannot be accurately forecasted is due to its randomness. Furthermore, many people are not aware of the precondition of its occurrence, and they are not prepared to that.

Currently, many European NHMSs are actively developing their severe thunderstorm forecast process and warning services with 26 (79%) of respondent countries issuing severe thunderstorm warnings and 8 (24%) issuing tornado warnings [6]. Both warning criteria and methodologies used in the warning process vary from country to country. Lead-times range from 30 min to 96 h, indicating a range of different warning philosophies for each country [6]. Given that tornados are hard to predict and the warnings give a very brief window of opportunity to prepare for evacuation to a secure underground or other location, each activity in the detection and warning phases is critically important to enable effective response actions to be taken before impacts on lives and property occur [3].

Yet, in the US tornados are not entirely random. The area of occurrences is generally located in the southeast area of the US, and most of tornados occurred between April and June. Nevertheless, every single tornado appears very suddenly and prior to its appearance there is no sign that can be recognized beforehand, for instance at least 5 h earlier, so that people have adequate time to get away of it. As the result, damages, losses, and costs totally cannot be predicted.

The most difficult part dealing with a natural disaster, such as a tornado, is to forecast it. For long time researchers have worked on ways to forecast the occurrence of a natural disaster. Some of them are at the stage of now-casting [7–11], not yet forecasting. To handle tornados properly, real-time seems to be an important aspect in forecasting its occurrences. So far researchers that have worked on the handling of natural disasters have developed some kind of disaster management that deals with prior, while, and post disaster. It is true that some approaches of early warning system has been built in various countries. However, the most time distance between the warning and the occurrence of tornado is very short. The reliable result is usually less than 30 min [6]. Such time distance is inadequate for people to protect themselves from the tornado's force. Some researchers believe that the most difficult part of forecasting natural disasters is producing the mathematical formulas of it [12].

Due to its complexity and its randomness, it is a strong sign that methods other than mathematical formulas for forecasting a tornado are required. Not only the complexity and the randomness have been a concern, the real-time capability of the forecaster may improve the time distance to hours before the occurrence and still maintain the accuracy. A number of researchers [12] have developed an artificial intelligent technology to forecast an upcoming tornado. This approach has been taken into account in order to avoid complex mathematical formulas. Although there is no definite mathematical functions that can be used to determine the condition of a tornado, wind-speed, wind-direction, air-temperature, and air-pressure that constitute a tornado, are all caused by physical states [9]. It means that the time-series of several physical values of

wind-speed, wind-direction, air-temperature, and air-pressure determine particular tornado condition. It can then be figured out that the occurrence of a tornado can generally produce particular physical patterns.

The Single Layer Hierarchical Graph Neuron (SLHGN) has been developed as an improvement of Multidimensional Hierarchical Graph Neuron (mHGN) [12]. At this stage, SLHGN begins to be ready for forecasting tornados, whereas mHGN is able to recognize incomplete patterns [13]. There are two major issues that have been resolved so that SLHGN is ready to forecast tornados. First, the physical values of weather data are represented using sophisticated scheme so, that false positive and true negative rates can be reduced. Second, the occurrence of a tornado can be determined based on the location where the measuring sensors of weather for wind-speed, wind-direction, air-temperature, and air-pressure are located. Therefore, with this new approach patterns from previous occurrences of tornados can be used as the training pattern for future purposes. In this paper, those two issues are discussed in order to show the improvement of mHGN in forecasting tornados.

2 The Need for a Real-Time Tornado Forecasting System

It is very clear that the genesis of a tornado happens suddenly. To deal with such a rapid occurrence a real-time forecasting system would be required. Many countries have taken this issue very seriously. In their study [6], they summarize the current severe thunderstorm warning and forecast operations in different European National Hydro-Meteorological Services (NHMSs). They also suggest various ways for countries developing their own warning service to learn from experiences from other countries, including the warning operations from the United States of America, the longest-lived severe thunderstorm warning operations in the world. Their study is based on a questionnaire sent to 39 European NHMSs of which thirty-three (85%) responded.

Currently, many European NHMSs are actively developing their severe thunderstorm forecast process and warning services with twenty six (79%) of respondent countries issuing severe thunderstorm warnings and eight (24%) of respondent countries issuing tornado warnings [6]. Both warning criteria and methodologies used in the warning process vary from country to country. Lead-times range from 30 min to 96 h, indicating a range of different warning philosophies for each country. Major challenges toward improving the warning systems include obtaining observations of severe weather for real-time forecasting and post-event verification, educating forecasters, and having access to state-of-the-art forecaster workstations [6]. An additional challenge is in communicating anticipated or ongoing severe thunderstorms, both internationally between NHMSs and nationally with media and emergency authorities. This is clear that those European countries have the same need in such real-time forecasting systems.

Based on their reports [14], a tornado hit the north-eastern suburbs of Hamburg, Germany, on 7 June 2016. It had an estimated strength of upper end F1 on the Fujita scale and was short-lived with an approximate duration of only 13 min and a path length of just about 1.3 km. They demonstrate that such a small-scale, extreme event

can be observed and forecasted accurately by a low-cost radar and by an atmospheric model with low computational costs, respectively. Observations from a low-cost single polarized X-band radar covering the urban area of Hamburg with 60 m spatial and 30 s temporal resolution are analysed with respect to their ability to capture the development as well as the track of the tornado. In contrast to the national C-band radar network, the X-band radar is capable of capturing the hook echo of the tornado as well as the circular pattern in rain rates, because of its higher resolution in space and time [14].

In their [14] research, High-resolution forecasts of the tornado event are conducted with the computational efficient Conformal Cubic Atmosphere Model (CCAM) in order to test the capability of predicting the tornado with a lead time of a few hours. A three step downscaling method is used to obtain a spatial resolution of 1 km with initial conditions taken from the NCEP analysis. Calculated severe weather indices clearly indicate a potential for a tornado. CCAM cannot explicitly resolve small scale tornadic features but the model simulates a strong convective cell only a few kilometres apart from the tornadic thunderstorm observed by the radar [14]. This is another indication that the complexity of tornado requires highly computational calculations. Although they claim that the CCAM is computationally efficient, the system is not yet capable of resolving small spatial and temporal scale of tornado time series measurements. Due to the rapid development of a tornado genesis, small spatial and temporal scale of data is the fundamental requirement for a real-time tornado forecasting.

Furthermore, the randomness of a tornado has been discussed by some researchers [5]. They reported about the occurrence of tornados on October 6, 2010 in Arizona, in the Grand Canyon State. On this day, Arizona experienced its largest, single-day tornado outbreak in its recorded history. Eight tornadoes were officially recorded in northern Arizona. This day further proved that, tornadoes are not only possible in Arizona, but, they can even be dangerous to both life and property.

Since a real-time forecasting system is not yet ready, losses from all natural hazards have increased steadily over the past three decades. A continuous cycle of Presidential disaster declarations was generated as communities rebuilt and recovered from these often devastating events [4]. Using a 50 year record, the paper examines the temporal variability and spatial distribution of tornado hazards in the United States. Tornado hazards are defined very specifically as any reported tornado that resulted in human injury, human fatality, or some amount of economic loss. The results suggest that, while the actual number of tornadoes (tornado segments) doubled over the entire time period, there was a smaller overall increase in the number of tornado hazards from 1950 to 2000 [4].

Another report [15] presents the climatology of Illinois tornadoes based on data from the 1916–1969 period, and offers a variety of general interest tornado facts. Illinois ranks eighth nationally in the number of tornadoes, but first in deaths and second in tornado damages. On the average, there are 10 tornadoes per year, occurring on five days. The annual average death rate from these storms is slightly over 19 with an injured average of 110 people. A majority (65%) of Illinois tornadoes occur during March through June, with 15–21 April being the prime 7-day period. Over 40% occur between 1500 and 1800 CST, and 65% take place from 1400–2000 CST. Five of the outstanding Illinois tornado days of the 1916–1969 period are discussed in detail,

including the famed Tri-State tornado of 18 March 1925, the most devastating tornado in the United States since systematic collection of tornado data began in 1916 [15].

However, in their paper [4] it is reported that the ratio of tornado hazards to all tornadoes has remained relatively constant since the 1960s. There has been a steady decline in fatalities and reductions in injuries caused by tornado hazards. Losses are more variable over the past 50 years, but the 1990s showed near record lows, in terms of both total dollar losses and mean losses per tornado hazard event. The statistical centre of tornado hazard activity is in south-central Missouri, southeast of the statistical centre of tornadoes identified by previous research. The density of tornado hazards has expanded outward from the historic ‘Tornado Alley’ region. This is again an indication to the randomness of tornado. The distribution of other high-density regions suggests additional tornado hazard regions in Florida, the lower Mississippi Valley, the Gulf Coast, and in the Carolinas [4].

Researchers [3] believe that communities are impacted only when and if a tornado touches down on the ground. Therefore, early recognition of tornadoes and proper communication of warnings at the pre-touch down phase would help the public to be ready and respond appropriately and effectively. Given that tornadoes are hard to predict and the warnings give a very brief window of opportunity to prepare for evacuation to a secure underground or other location, each activity in the detection and warning phases is critically important to enable effective response actions to be taken before impacts on lives and property occur [3].

Furthermore, the paper [3] presents a detailed analysis of the tornado detection and warning system in Canada. The sequence of activities, their interrelationships in the tornado detection, warning and communication system are identified and developed as a network taking the City of Calgary, Alberta as a case study. In their system, the time durations of activities in the network are estimated and represented via triangular probability distributions. Developing the activity network is a continuous process of refinement based on information gathered from different sources such as Environment Canada and emergency management officials at provincial and local levels based on how they are associated with tornado detection, warning and communication. Their network is modelled using the simulation-based schedule networking tool DSSS in the Symphony software. Based on the simulation output results, improvements to the existing tornado detection, warning and communication system in Canada are proposed. Again, highly computational calculation is required in this approach, which may not be suitable for a real-time forecasting system [3].

It is also reported in [16] that the amount of forecast skill involved when issuing tornado and severe thunderstorm warnings is closely related to the type of storm that causes the severe weather. Storms from eight tornado outbreaks are classified and correlated with tornado warnings and severe thunderstorm warnings. These warnings were verified, missed, or shown to be false alarms by relating them with storm reports that match temporally and spatially with those in the Storm Prediction Center’s database. Certain forecast parameters, including the critical success index (CSI), probability of detection (POD), false alarm ratio (FAR), and warning lead time are calculated for each storm type and for each type of warning. Because it was not practical to manually classify these storms (~50,000 entities), a decision tree was trained on a subset of manually classified storms using Quinlan’s C4.5 algorithm. The decision tree was then

used to automatically classify storms as being of one of four types: super cellular, linear, pulse or unorganized.

It was found in [16] that both tornado warnings and severe thunderstorm warnings issued for isolated supercells and convective line storms have higher CSI, higher POD, and lower FAR scores than those issued for pulse and non-organized storms. Lead times were consistently longer for supercell and line storms, while usually very short for pulse and non-organized storms. They conclude that measures of forecast skill are particularly sensitive to the type of storm. Thus, any measurement of forecast skill, such as the year-over-year skill measure of an individual forecast office, has to take into account the types of storms in that office's warning area in the time period considered. However, as mentioned in the previous paragraphs most of tornados and storms are pulse and non-organized ones. Since many parameters, measurements, and conditions need to be considered, many issues related to forecasting tornados need to be investigated further.

The possibility of the occurrence of a natural disaster is not always constant, but it is still and always there in various parts of the world. Some countries have experienced natural disasters more than other countries [12]. Such a situation has been the main reason why a country like the US has spent many efforts to deal with natural disasters. However, this does not mean that only the US must be concerned with the occurrence of tornados. The randomness of the occurrence of a natural disaster is not only in terms of the location, but also of the time and the severity. However, the location and the time (season) of tornados to occur is generally the same. Previous tornados in the US occurred between April and June, and the most places that have been hit are those in the southeast area of the US.

Although many researchers in opinion that the severity and the average magnitude of natural disasters have increased since the last decade. However, it is still not clear how severe future natural disasters might be. The impossibility of measuring, or predicting the severity of natural disasters, has been the major cause of the difficulties in anticipating their occurrences. Many other researchers have suggested that, one way to deal with the randomness of the occurrence of natural disasters is through a real-time disaster forecaster, as many early warning system [2] and now-casting [1, 7, 9, 10, 17, 18] that have been investigated and developed are not yet able to help people avoiding and mitigating natural disaster.

If the lead time of detecting a tornado is short, people will not be able to save themselves away from the tornado. For instance, the forecasting approach that they [8, 11, 19] have attempted is able to forecast the disaster, but the lead time is only one hour. Despite those efforts that have been taken by researchers, Sorensen [2] argues that in terms of prediction and forecasting, still no radical breakthroughs have occurred in the past twenty years. Due to its complexities, most natural disaster researchers are working on technologies that are not focusing on the forecasting techniques. They have tried to find an appropriate approach for working on three areas: natural disaster forecaster, now-casters, or early warning systems. However, they [2, 17, 20] also still integrate their system with disaster management systems. Even Doong et al. [20] suggest that the success of a disaster mitigation concept lies in the quality of the disaster management. This shows that their approach alone is not yet adequate to handle natural disasters nor tornados. The potential reason to this case is the fact that a

system for handling tornados requires very complicated mathematical analysis. So many parameters and values need to be considered and included in their calculation [1, 8, 9], and it is time consuming [18], but the system must run fast [1], that can be used to warn people as early as possible. Therefore, a real-time forecasting system needs to be used to deal with it.

Daily Weather History Graph

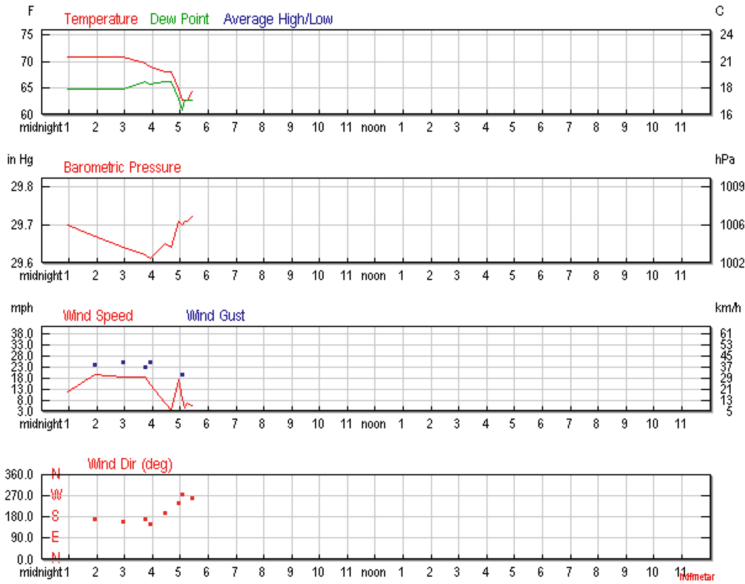


Fig. 1. Incomplete weather data

As already mentioned, although the complexity and the randomness of the occurrences of a natural disaster have caused difficulties in handling it, the development of every natural disaster still follows natural science characteristics and rules. Each tornado owns specific location, time, patterns and characters. Yet, the difficulty to gain the measured values of those characteristics has become a new challenge in recognizing tornados before it turns up. The steps that a tornado builds before its strong and winding wind can be treated as a pattern. It means that the recorded data from previous tornado disasters plays a big role in recognizing it. Therefore, the data must be kept properly. The data is the important source of clue for researchers to analyse the pattern of a tornado. When patterns of tornados can be recorded, it is a strong possibility that when one of the patterns is about to turn up, a system that can recognize patterns can be used to recognize a tornado early before it becomes a strong and destructive one. Such patterns are the most important part of SLHGN for forecasting tornados hours before they strike. Unfortunately, the data provided by weather stations such as the weather

data of NOAA is not so accurate (see Fig. 1). Furthermore, the location of weather sensors is not exactly where the previous tornados have occurred (see Figs. 2 and 3).

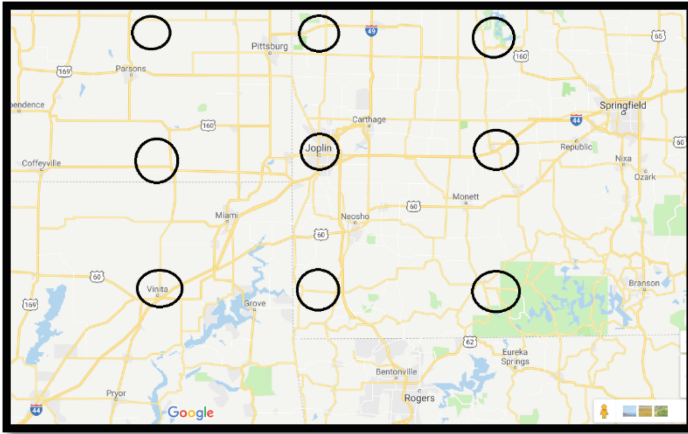


Fig. 2. Ideal positions of weather sensors for Joplin

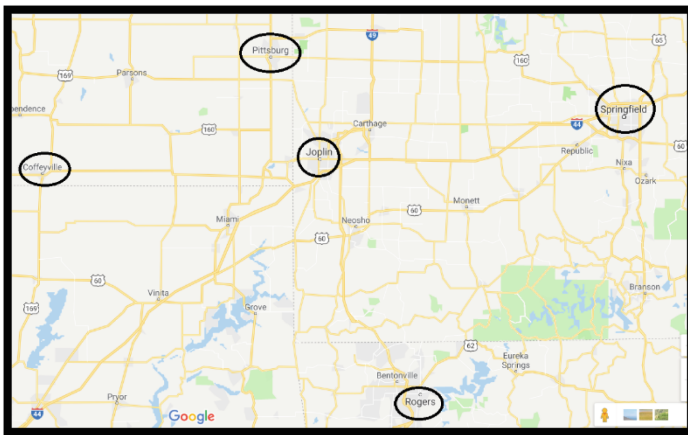


Fig. 3. Real positions of weather sensors for Joplin

3 Single Layer Hierarchical Graph Neuron (SLHGN)

As already mentioned, the SLHGN is an improvement of mHGN [13]. There are some important features that have been implemented in SLHGN compared to mHGN through which the accuracy and speed have been increased. All the features in mHGN

are still available in SLHGN, for instance its capability of recognizing multidimensional patterns.

Again, the SLHGN has been developed for solving multidimensional problems that has been discussed since a long time ago. Not only multidimensional, the process must be real-time without affecting the accuracy. Fortunately, researchers are aware of that solving complex problems numerous dimensions must be considered and calculated. Otherwise, analysing using just a few parameters cannot be considered correct. Such a condition will normally produce very high false positive or true negative error rates. Another issue when solving multidimensional problems is the method that will be implemented. Not only the number of dimensions is large, but how all the dimensions are interrelated to each other, or independent on one another, is often not clear.

The reason why single layer and real-time issues are required due to the fact that natural disasters is real-time problems and multidimensional problems. Therefore, forecasting natural disasters like tornados can also be considered as solving multidimensional problems. In this case, not only the longitude and the latitude determine the condition of a tornado, air-temperature, air-pressure, air-humidity, wind-direction, and wind-speed also play a big role in causing tornados. A problem that still exists is the interdependency amongst those tangible and with intangible values such as: industrial development, people movement, etc. It is so difficult to figure out mathematical formulas that constitute such interdependency. This is, therefore, a strong indication that such multidimensional problems may be solved using SLHGN.

3.1 Experiment Results

The procedure for testing SLHGN architecture is pretty much the same as the one for mHGN architecture. For the experiment, in SLHGN composition a number of neurons (GN) are operated by a single thread only, whereas in mHGN composition one neuron is operated by a thread. Various 2D-, 3D-, 4D- and 5D-pattern recognition have been scrutinized. The compositions used in the experiment are: 15×15 SLHGN, $5 \times 15 \times 15$ SLHGN, $5 \times 5 \times 15 \times 15$ SLHGN, and $5 \times 5 \times 5 \times 15 \times 15$ SLHGN respectively. For instance, in the 15×15 pattern recognition the SLHGN composition requires only 225 threads (see Fig. 5), whereas mHGN composition requires 1009 threads (see Fig. 4). Both compositions of SLHGN and mHGN contains 2018 neurons. This shows how the composition of SLHGN requires very much smaller number of threads, which reduces the operational time and makes it more suitable for a real-time application. As for creating patterns, binary data is used, then two values (i.e. 0 and 1) of data are required. Therefore, 450 threads are required in the 15×15 SLHGN composition. So, 450 threads have been run in parallel during this 2D pattern recognition. By using threads, the activity of neurons is simulated so that the functionalities are close to the real neuron functionalities.

The experiment works on all the patterns of 26 alphabetical figures. Following the composition of the threads, the alphabet patterns consist of 15×15 pixels. For the training purpose, the SLHGN is first fed one-cycle with all the 26 non-distorted patterns. The pattern order during the training phase has been determined randomly. Then, to acquire the recognition results the SLHGN is fed with a number of randomly

distorted patterns of alphabetical figures. The recognizing accuracy is taken by calculating the average value of the results.

During the experiment, 20 distorted patterns for each alphabetical figure have been generated. After gaining the results, the experiment is re-run until in total 10 times with the same steps, but in every run the SLHGN is still trained with 26 patterns of alphabetical figures but with randomly different order. So, for each alphabetical figure for particular percentage of distortion, in total 200 distorted patterns have been generated as testing patterns.

There are 7 percentage levels of distortion that have been tested, they are: 1.3%, 2.7%, 4.4%, 6.7%, 8.0%, 8.9%, and 10.7%. These percentage levels have been so chosen based on the number of pixels that have been distorted. The sizes of distorted pixels represent the factor and the non-factor of the dimension of the patterns. By doing so, we can observe all the possibilities of distortion. So, in total there are 5200 ($26 \times 20 \times 10$) randomly distorted testing patterns.

The following shows some results taken from testing 4.4% randomly distorted patterns, and the SLHGN was previously stored with alphabetical figure patterns, and the order was IEFXMQYJHPDKTORZCUALBGVWNS. The value on the right side of each alphabet show the portion (percentage) of the pattern that is recognizable as the corresponding alphabet.

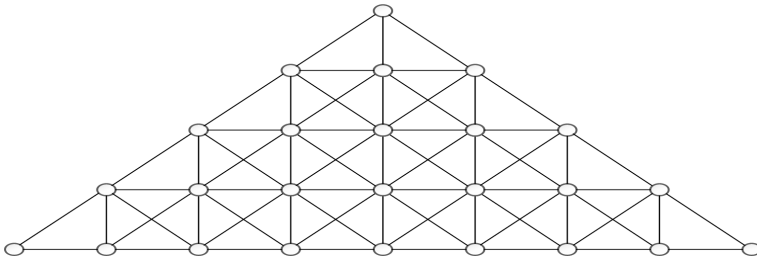


Fig. 4. One-Dimensional 25 neurons mHGN run by 25 threads

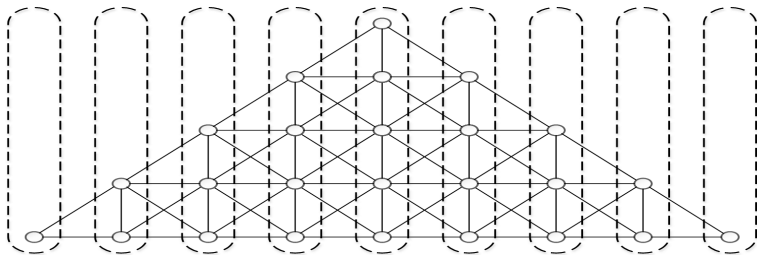


Fig. 5. One-Dimensional 25 neurons SLHGN run by 9 threads

After collecting the results taken from testing 5200 patterns we can summarize how accurate the SLHGN is, in recognizing different percentage levels of distortion of 26 alphabets. The summary is taken based on the average accuracy values from all the steps. The following shows the summarized result taken from testing distorted patterns using five-dimensional $5 \times 5 \times 5 \times 15 \times 15$ SLHGN. Important to note that these SLHGN results must be the same as those of mHGN. By doing so, it confirms the SLHGN architecture functions correctly (Figs. 6 and 7).

PATTERNS RANDOMLY DISTORTED 4.4 %																								
Patterns Stored	Distorted Pattern	Recognised patterns and their recognized portion (%) from 20 different randomly distorted patterns																				Recognised Correctly		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			
Order Type 0	I	A	9A	8A	8A	54A	13A	3A	8A	13A	16A	6A	15A	14A	15A	14A	14A	2A	1A	54A	17A	15	20	
	E	B	3B	4B	36B	37B	11B	36B	11B	11B	37B	10B	5B	36B	13B	10B	12B	11B	6B	9B	11B	36	20	
	F	C	C	40C	7C	36C	15C	8C	39C	8C	7C	7C	8C	15C	7C	0C	14C	15C	40C	14C	1C	14C	39	20
	X	D	D	26D	10D	9D	10D	10D	10D	2D	26D	12D	25D	11D	5D	12D	6D	9D	9D	10D	11D	10	20	
	M	E	E	10E	8E	30E	29E	12E	11E	11E	10E	10E	29E	11E	10E	10E	11E	12E	29E	31E	6E	10E	9	20
	Q	F	F	20F	4F	19F	3F	20F	9F	7F	8F	4F	1F	7F	21F	19F	22F	8F	9F	21F	8F	4F	8	19
	Y	G	H	8H	14H	54H	13H	15H	51H	14H	50H	7G	51G	14G	15G	6G	15G	15G	7G	3G	14G	6G	7	20
	H	H	H	8H	14H	54H	13H	15H	56H	15H	14H	7H	54H	13H	3H	16H	55H	14H	54H	9H	15H	55H	54	20
	H	I	I	54I	6I	1I	53I	6I	14I	14I	55I	15I	15I	54I	14I	14I	15I	54I	6I	15I	14I	16I	5	20
	P	J	J	54J	5J	5J	14J	15J	13J	54J	14J	55J	6J	55J	15J	15J	6J	54J	55J	14J	13J	14J	3	20
	D	K	K	4K	7K	8K	8K	8K	8K	8K	8K	22K	9K	8K	7K	6K	4K	2H	22K	9K	7K	5	18	
	K	L	L	9L	11L	5L	39L	11L	11L	11L	39L	6L	40L	12L	39L	10L	6L	5L	10L	7L	10L	12L	39	20
	T	M	M	22M	19M	4M	21M	22M	8M	4M	8M	6M	7M	6M	5M	7M	5M	4M	7M	7M	6M	7	20	
	O	N	N	19N	4N	3N	7H	19N	8N	8N	4N	7N	6N	7N	8N	18N	8N	6N	7N	8N	5N	7N	5	18
	R	O	O	54O	16O	15O	14O	14O	7O	56O	14O	54O	8O	3O	16O	14O	16O	55O	54O	6O	15O	7O	54	20
	Z	P	P	10P	32P	10P	9P	9P	10P	5P	6P	5P	33P	10P	0P	8P	10P	10P	32P	5P	9P	11P	11	20
	C	Q	Q	44Q	15Q	4Q	15Q	45Q	16Q	13Q	14Q	13Q	6Q	14Q	15Q	15Q	15Q	45Q	6Q	44Q	14Q	45	20	
	U	R	R	11R	12R	12R	5R	12R	11R	38R	5R	6R	10R	10R	7R	37R	10R	37R	38R	11R	2R	12	20	
	A	S	S	14S	15S	8S	14S	13S	13S	3S	7S	13S	14S	46S	46S	13S	15S	46S	7S	13S	14S	15	20	
	L	T	T	14T	15T	14T	15T	15T	16T	13T	7T	15T	58T	15T	14T	15T	55T	16T	54T	55T	54T	14T	14	20
	B	U	U	4U	7U	10U	10U	11U	11U	9U	11U	5U	30U	11U	5U	30U	4U	10U	10U	10U	11U	2	20	
	G	V	V	56V	14V	54V	15V	54V	8V	55V	55V	6V	14V	14V	9V	2V	15V	1V	56V	16V	14V	16V	15	20
	V	W	W	6W	14W	30W	29W	13W	13W	3W	14W	6W	12W	14W	11W	13W	13W	7W	30W	30W	11W	12	20	
	W	X	X	15X	53X	16X	15X	15X	14X	15X	14X	14X	14X	7X	6X	16X	54X	15X	14X	54X	14X	13X	54	20
	N	Y	Y	15Y	16Y	15Y	13Y	14Y	51Y	8Y	2Y	51Y	13Y	51Y	15Y	14Y	13Y	52Y	16Y	15Y	15Y	5	15	20
	S	Z	Z	7Z	15Z	15Z	14Z	13Z	14Z	16Z	17Z	0Z	51Z	14Z	51Z	50Z	7Z	6Z	7Z	51Z	14Z	13Z	51	20

Fig. 6. The result of all the 26 alphabetical patterns that are twenty times 4.4% randomly distorted.

5X5X5X15X15 Patterns	Distortion (%)						
	1.3	2.7	4.4	6.7	8.0	8.9	10.7
A	100	100	100	100	100	100	100
B	100	100	100	100	98	97	94
C	100	100	100	100	100	96	100
D	100	100	100	100	100	100	98
E	100	100	100	100	100	100	100
F	100	99	94	89	83	85	74
G	100	100	100	100	100	100	100
H	100	100	89	67	48	50	55
I	100	100	100	100	100	100	100
J	100	100	100	100	100	100	100
K	100	100	98	81	70	72	67
L	100	100	100	100	100	100	100
M	100	100	93	76	55	66	49
N	100	100	97	77	63	60	55
O	100	100	100	100	100	100	100
P	100	99	87	79	80	81	81
Q	100	100	100	100	100	94	99
R	100	100	100	95	100	99	95
S	100	100	100	100	100	100	100
T	100	100	100	100	100	100	100
U	100	100	100	100	100	100	100
V	100	100	100	100	99	98	92
X	100	100	100	100	100	100	100
Y	100	100	100	100	100	100	100
Z	100	100	100	100	100	100	100
Average	100	100	98	95	92	92	91

Fig. 7. The summary of the result using $5 \times 5 \times 5 \times 15 \times 15$ SLHGN [21]

The following figure shows the differences of recognition accuracy amongst 15×15 , $5 \times 15 \times 15$, $5 \times 5 \times 15 \times 15$, and $5 \times 5 \times 5 \times 15 \times 15$ SLHGN architectures when recognizing 10.7% distorted patterns of alphabets (Fig. 8).

Comparison Result		Distortion = 10.7 %			
		15X15	5X15X15	5X5X15X15	5X5X5X15X15
Recognition Accuracy for Each Pattern (%)	A	99	100	100	100
	B	58	69	92	94
	C	67	93	94	100
	D	78	92	94	98
	E	85	80	100	100
	F	61	71	81	74
	G	87	98	100	100
	H	23	63	69	55
	I	95	100	100	100
	J	77	95	100	100
	K	68	59	84	67
	L	50	80	100	100
	M	38	36	35	49
	N	53	42	63	55
	O	100	100	100	100
	P	61	59	75	81
	Q	63	73	73	99
	R	79	90	95	95
	S	78	97	100	100
	T	93	95	100	100
	U	89	84	85	100
	V	100	100	100	100
	W	75	82	98	92
	X	85	100	100	100
	Y	100	100	100	100
	Z	99	100	100	100
Average		75	83	90	91

Fig. 8. Differences of recognition accuracy amongst four different architectures

3.2 Time-Series in Pattern Recognition

Recognizing patterns of time series problem utilizes data that have previously been recorded regularly in timely manner [21]. For instance, if the parameter that needs to be recorded is a single value, and the recording step is every six hours, then there will be 4 values recorded every day. In order to constructs the recorded values as a pattern, the data representation of the recorded values need to be developed so, that they can fit into a pattern recognition architecture. The following figure shows six ways of representing recorded data for 8 levels of measurement.

Definition: The distance between two values is the number of different bits between them. For instance, the distance between 001 and 110 is 3, whereas the distance between 00000001 and 10000000 is 2. It is known as Hamming Distance.

It can be seen from Fig. 9 that the data (value) is represented using binary numbers, and there are six possible data representations (Ver1 till Ver6). For the Ver1 (Binary Code Decimal), the distance between adjacent values varies. It is therefore not suitable for SLHGN. For the Ver2 (Grey Code), the distance between two adjacent values is constant, which is one. However, the distance between Value 2 and 7 is also one. This is not suitable, as for SLHGN the distance one also means that the two values are very

Value	Ver1	Ver2	Ver3	Ver4	Ver5	Ver6
1	000	000	00000001	10000000	00000001	10000000
2	001	001	00000010	01000000	00000011	11000000
3	010	011	00000100	00100000	00000111	11100000
4	011	010	00001000	00010000	00001111	11110000
5	100	110	00010000	00001000	00011111	11111000
6	101	111	00100000	00000100	00111111	11111100
7	110	101	01000000	00000010	01111111	11111110
8	111	100	10000000	00000001	11111111	11111111

Fig. 9. Six examples of data representation for 8 levels of value

close to each other. In fact, value 2 and 7 are very different and very far from each other. Again, this is not suitable for SLHGN. For the Ver3 and Ver4 (Ring Counter), the distance between any two values is constant, which is two. These are again not suitable for SLHGN. For the Ver5 and Ver6 (Johnson Counter) the distance between adjacent values is constant, which is one. Additionally, the distance between any two values is linear with the value differences.

It seems to be that the Ver5 and Ver6 are the most suitable data representation that can be used with SLHGN. However, such data representation will not maximally utilize the binary combination. With 3-bit data, only $3/8$ or 0.375 is the occupation ratio. For 4-bit data is the occupation ratio $4/16$ or 0.25 . The occupation ratio is $5/32$ or 0.15625 for 5-bit value. This shows that the Ver5 and Ver6 data representation will produce less occupation ratio, the more number of bits is used. This is an indication that due to such an occupation ratio the pattern recognizer will have less recognition accuracy, the more number of bits is used. So, these are again not suitable for SLHGN.

The following is a better data representation that solves those above issues.

Value	3-bit	4-bit	5-bit
1	101	0101	00101
2	100	0100	00100
3	110	0110	00110
4	010	1110	01110
5	011	1111	01111
6	001	1101	01101
7		1001	01001
8		1000	01000
9		1010	01010
10		0010	11010
11		0011	11011
12		0001	11001
13			11101
14			11100
15			11110
16			10110
17			10111
18			10101
19			10001
20			10000
21			10010
22			00010
23			00011
24			00001

Fig. 10. Three examples of a better data representation for 3-, 4-, and 5-bit binary values

In Fig. 10 there are three examples of 3-bit, 4-bit, and 5-bit data representation. It is shown that the distance between any adjacent levels in all samples is constant, which is one. Furthermore, between any two values which have value difference of two, the distance is also constant, which is two. Last, the distance between any two values which have value difference of three, the distance is again constant, which is three. For SLHGN, such constant distances of 1, 2, and 3 are adequate to be used for tornado recognition. Another characteristic of these data representations is that the representation is cyclic. It means that, if it is required the order of binary representation can be modified circularly without affecting the distances. Using such better data representations, in all examples is the occupation ratio constantly 0.75. With such a constant occupation ratio the pattern recognizer will have constant recognition accuracy, any number of bits in it is used. Important to mention that the SLHGN will use the bold values in the cluster.

The following figure shows the proof that the bold values have a distance of 1 only to adjacent values, otherwise more than 1 to other values (Fig. 11).

	00101	00100	00110	01110	01111	01101	01001	01000	01010	11010	11011	11001	11101	11110	10110	10111	10101	10001	10000	10010	00010	00011	00001	
00101	0	1	2	3	2	1	2	3	4	5	4	3	2	3	4	3	2	1	2	3	4	3	2	1
00100	1	0	1	2	3	2	3	2	3	4	5	4	3	2	3	2	3	2	3	2	3	2	3	2
00110	2	1	0	1	2	3	4	3	2	3	4	5	4	3	2	1	2	3	4	3	2	1	2	3
01110	3	2	1	0	1	2	3	2	1	2	3	4	3	2	1	2	3	4	5	4	3	2	3	4
01111	2	3	2	1	0	1	2	3	2	3	2	3	2	3	2	3	2	3	4	5	4	3	2	3
01101	1	2	3	2	1	0	1	2	3	4	3	2	1	2	3	4	3	2	3	4	5	4	3	2
01001	2	3	4	3	2	1	0	1	2	3	2	1	2	3	4	5	4	3	2	3	4	3	2	1
01000	3	2	3	2	3	2	1	0	1	2	3	2	3	2	3	4	5	4	3	2	3	2	3	2
01010	4	3	2	1	2	3	2	1	0	1	2	3	4	3	2	3	4	5	4	3	2	1	2	3
11010	5	4	3	2	3	4	3	2	1	0	1	2	3	2	1	2	3	4	3	2	1	2	3	4
11011	4	5	4	3	2	3	2	3	2	1	0	1	2	3	2	3	2	3	2	3	2	3	2	3
11001	3	4	5	4	3	2	1	2	3	2	1	0	1	2	3	4	3	2	1	2	3	4	3	2
11101	2	3	4	3	2	1	2	3	4	3	2	1	0	1	2	3	2	1	2	3	4	5	4	3
11100	3	2	3	2	3	2	3	2	3	2	3	2	1	0	1	2	3	2	3	2	3	4	5	4
11110	4	3	2	1	2	3	4	3	2	1	2	3	2	1	0	1	2	3	4	3	2	3	4	5
10110	3	2	1	2	3	4	5	4	3	2	3	4	3	2	1	0	1	2	3	2	1	2	3	4
10111	2	3	2	3	2	3	4	5	4	3	2	3	2	3	2	1	0	1	2	3	2	3	2	3
10101	1	2	3	4	3	2	3	4	5	4	3	2	1	2	3	2	1	0	1	2	3	4	3	2
10001	2	3	4	5	4	3	2	3	4	3	2	1	2	3	4	3	2	1	0	1	2	3	2	1
10000	3	2	3	4	5	4	3	2	3	2	3	2	3	2	3	2	3	2	1	0	1	2	3	2
10010	4	3	2	3	4	5	4	3	2	1	2	3	4	3	2	1	2	3	2	1	0	1	2	3
00010	3	2	1	2	3	4	3	2	1	2	3	4	5	4	3	2	3	4	3	2	1	0	1	2
00011	2	3	2	3	2	3	2	3	2	3	2	3	4	5	4	3	2	3	2	3	2	1	0	1
00001	1	2	3	4	3	2	1	2	3	4	3	2	3	4	5	4	3	2	1	2	3	2	1	0

Fig. 11. Matrix of distances of bold binary values of better data representation

The following figure shows an example of recorded data taken from a single location measurement and each value has 3×5 -bit (15 bits) values.

It can be seen from Fig. 12 that the recorded values from parameter of 15-bit data construct a two-dimensional pattern of 10×15 architecture. Utilizing these recorded data, the SLHGN can forecast a tornado 6 h earlier, when the same tornado will occur again. It means that if values have been recorded and the same pattern is recognized by the pattern recognizer, then the tornado is forecasted to occur again within 6-h time with around 90% of accuracy.

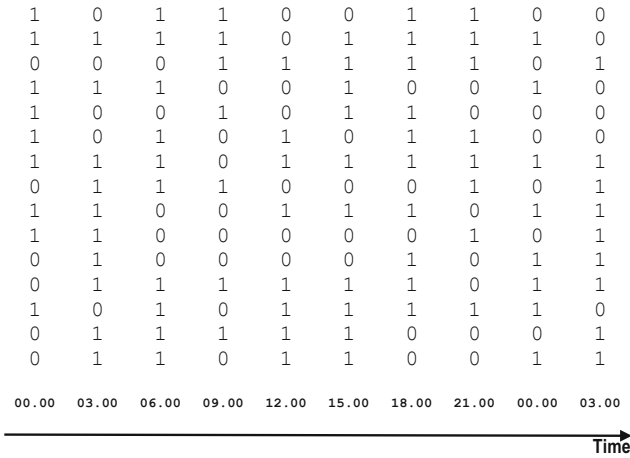


Fig. 12. Data of 15-bit values build a 2D-Pattern

4 Single Layer Hierarchical Graph Neuron for Tornado Forecasting

In the previous section, time series value is described and represented so, that it can be forecasted through utilizing a pattern recognition, such as SLHGN. In case of tornado forecasting, single parameter in a location, such as air-pressure, is not the only value that determine the occurrence of a tornado in the location within 6-h time. Several other parameters, such as wind-speed, wind-direction, air-temperature, and air-humidity, play a big role in the occurrences as well. It means that the number of levels or a measured value will increase according to the number of parameters. In case 5 parameters need to be measured and each parameter contains 5-bit value, the required pattern structure would be 10×25 .

Also described in the previous section that measuring a parameter at particular point of location for several periods of time will generate a two dimensional pattern. If a series of points of the location need to be measured for several period of time, then the measured values will become a three dimensional pattern. The following figure depicts how some part of it will look like.

Also described in the previous section that measuring parameters at particular point of location for several periods of time will generate a two dimensional pattern. If a series and linear of locations need to be measured for several periods of time, then the measured values will become a three dimensional pattern. If the location that need to be measured is an 2D area, then the measured values will generate a 4D pattern. Furthermore, if the location that need to be measured is a 3D area, then the measured values will generate a 5D pattern (Fig. 13).

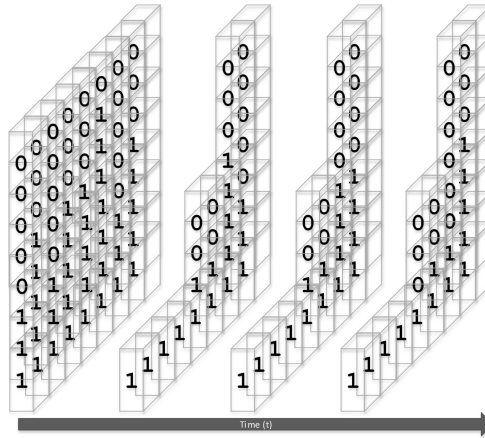


Fig. 13. A row of data of 8-bit value build a 3D-Pattern

4.1 The Architecture of SLHGN for Time-Series Tornado Data

The utilization of SLHGN has introduced a new approach that a local tornado forecast can be operated using small and cheap components. The values of air-temperature, air-humidity, air-pressure, wind-speed, and wind-direction can be gained through ordinary sensors. The area that is covered by those sensors can be a 3D area, because such small sensors can be easily mounted in valleys or hills, or even vehicles. The sensors can be embedded in a tiny computer, such as Raspberry Pi. The tiny computer will be responsible to run several GNs. The values taken from the sensors will then be worked out within the GNs. The connectivity of neurons is developed within a tiny computer and through the interconnectivity of the tiny computers.

In short, to build a tornado forecast for particular location, five parameters need to be measured. They are: wind-speed, wind-direction, air-temperature, air-humidity, and air-pressure. So, if one parameter is represented through 5-bit binary data, then for the measurement of 5 parameters 25-bit data is needed. For the time series, 15 series of measurement will be carried out. For an area that needs to be protected by SLHGN, $3 \times 3 \times 3$ measurement points will be deployed. So, the SLHGN dimension will be $3 \times 3 \times 3 \times 25 \times 15$.

4.2 Data Handlings for Real Tornadoes

Two deadliest tornadoes occurred quite recently are the tornado that struck Joplin, Missouri on May 22, 2011 and the one in Hackleburg–Phil Campbell, Alabama on April 27, 2011. To store the circumstances, several parameters in these areas need to be stored in SLHGN. Fortunately, the National Oceanic and Atmospheric Administration (NOAA) provides lots of data of: air-temperature, air-humidity, air-pressure, wind-speed, wind-direction in most areas of the US. These data will be the major source for SLHGN to store previous occurrences of tornadoes.

The following is the list of tornados scale F5/EF5 (the strongest) occurred In the US. The indicator F stands for Fujita and EF stands for extended Fujita. The scale has been named the same as the meteorologist Ted Fujita, who developed the scale (Fig. 14).

- (1) May 4, 2007, Greensburg, Kansas
- (2) May 25, 2008, Parkersburg–New Hartford, Iowa
- (3) April 27, 2011, Philadelphia–Preston, Mississippi
- (4) April 27, 2011, Smithville, Mississippi
- (5) April 27 2011, Hackleburg–Phil Campbell, Alabama
- (6) April 27 2011, Tuscaloosa–Birmingham, Alabama
- (7) April 27, 2011, Rainsville–Sylvania, Alabama
- (8) May 22 2011, Joplin, Missouri
- (9) May 24, 2011, El Reno–Piedmont, Oklahoma
- (10) May 24, 2011, Chickasha–Blanchard–Newcastle, Oklahoma
- (11) May 24, 2011, Washington–Goldsby, Oklahoma
- (12) May 20, 2013, Moore, Oklahoma
- (13) May 31, 2013, El Reno, Oklahoma
- (14) April 27, 2014, Vilonia, Arkansas

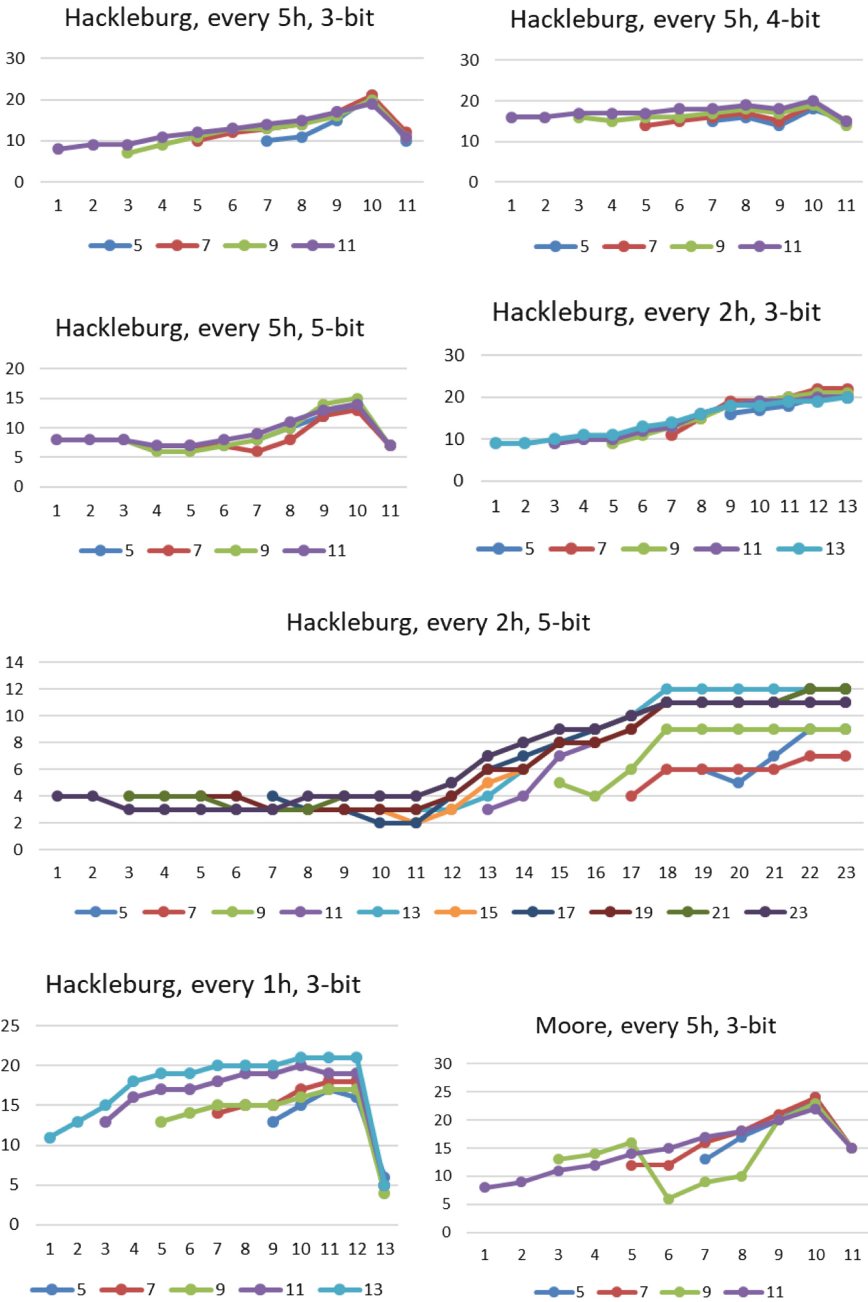


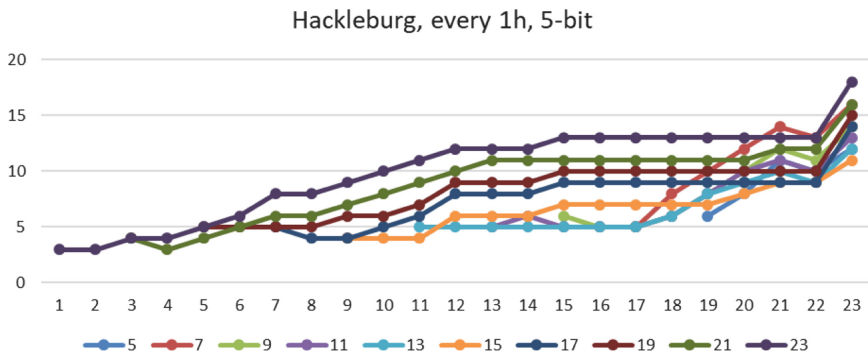
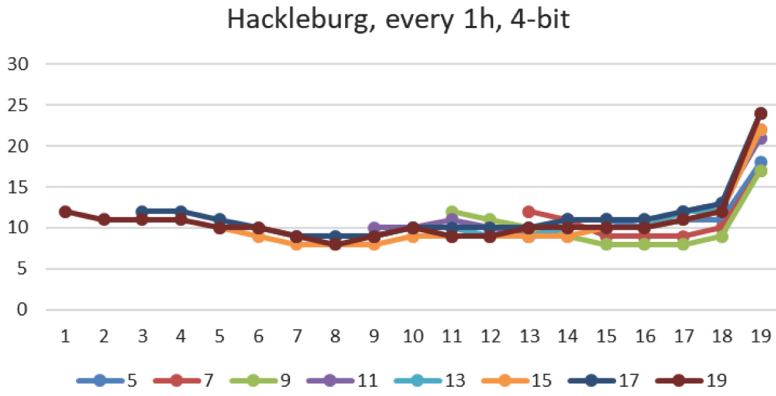
Fig. 14. The locations of F5/EF5 tornados

It can be seen from the map above, that based on the close location and the same time frame the data of Joplin’s tornado can be used to test the Oklahoma’s tornados (five tornados). Similarly, the data of Philadelphia-Preston’s tornado or Smithville’s tornado can be used to test Alabama’s tornados (three tornados).

4.3 The Results of the Real Tornadoes

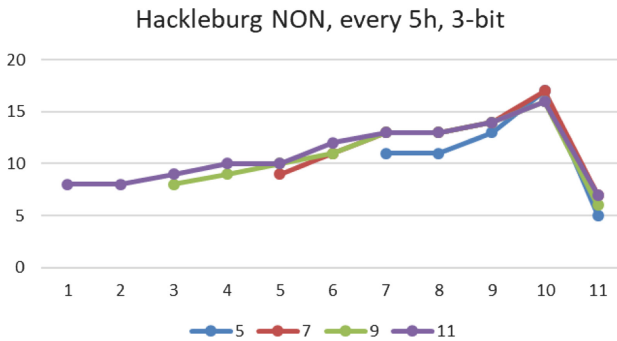
The following are the results of recognizing real tornados generated by SLHGN.

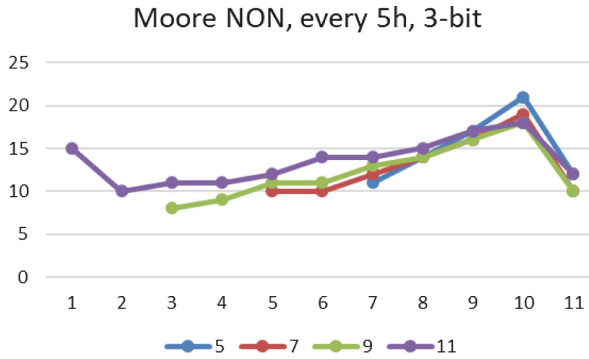




The results from the above diagrams show that SLHGN produces a steady incremental slope which means that a tornado is about to occur. Although the weather data is not complete, the slope clearly shows the tendency of a tornado. Of all the diagrams, the best one is produced by the three-bit data.

The following are the results for recognizing NON tornado generated by SLHGN.





The results from the above diagrams show that SLHGN produces a non-steady incremental slope which means that no tornado will occur. Again, the weather data is not complete, the slope can be analysed that the tendency suggests to no tornado. All the diagrams are produced by the three-bit data.

Still, the results of recognizing both tornados and non-tornados are not yet at the best. Refinement of the analysis as well as the data handling is still required.

5 Discussion

As is the case with pattern recognition of alphabets, patterns are more or less different to one another. However, in time series measurement data patterns, which are constructed from the measured values of the sensors, data can be very similar to one another. Therefore, data representation of measured values before data is fed to the architecture of SLHGN plays a big role in having very accurate results. So, the more tornado data can be gained, the more data can be used to find out which data can be used as the training data. The most challenging case with SLHGN is that the SLHGN can be trained one cycle only. The ideal data would be those taken from different cities and different countries. As SLHGN is trained one-cycle only, it is a challenge to choose which data is the right data for the training purpose. So for the decision to that challenge is the consolidated data from a number of tornados.

6 Conclusion

From the experiment results it is shown that SLHGN has the capability to recognize both tornado and non tornado patterns. For the tornado forecaster, we have presented results of up to 5D architecture. As already discussed in [22] and [13] there is no modification required if the architecture needs to be extended to bigger sizes of patterns. In the future this capability will be improved to the extent so, that multi oriented of multidimensional patterns will also be recognizable. At this stage it is also observed that SLHGN still use a single cycle memorization and recall operation. The scheme still

utilizes small response time (suitable for real-time) and it is insensitive to the increases in the number of stored patterns.

References

1. Cheng, R.T., Smith, R.E.: A Nowcast model for tides and tidal currents in San Francisco Bay, California. In: Ocean Community Conference, Baltimore, USA (1998)
2. Sorensen, J.H.: Hazard warning systems: review of 20 years of progress. *Nat. Haz. Rev.* **1**(2), 119–125 (2000)
3. Durage, S.W., Ruwanpura, J., Wirasinghe, S.: Analysis of the tornado detection, warning and communication system in Canada. In: Conference on building resilience, International Institute for Infrastructure Renewal and Reconstruction (IIIRR), Kumamoto, Japan (2012)
4. Boruff, B.J., Easoz, J.A., Jones, S.D., Landry, H.R., Mitchem, J.D., Cutter, S.L.: Tornado hazards in the United States. *Clim. Res.* **24**(1), 103–117 (2003)
5. Graves, M.: Cracking the AQ code. *Ariz. Dep. Environ. Qual.* **2**(8), 1–15 (2016)
6. Rauhala, J., Schultz, D.M.: Severe thunderstorm and tornado warnings in Europe. *Atmos. Res.* **93**(1), 369–380 (2009)
7. Kirschbaum, D.B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., Lerner-Lam, A.: Advances in landslide nowcasting: evaluation of a global and regional modeling approach. *Environ. Earth Sci.* **1**(1), 1–14 (2011)
8. Bellaire, S., Jamieson, B.: Nowcast with a forecast – snow cover simulations on slopes. *J. Something* **1**(1), 1–7 (2012)
9. Li, P., Lai, E.: Applications of radar-based nowcasting techniques for mesoscale weather forecasting in Hong Kong. *Meteorol. Appl.* **11**(1), 253–264 (2004)
10. Roy, S.S., Lakshmanan, V., Bhowmik, S.R., Thampi, S.: Doppler radar based nowcasting of cyclone Ogn. *J. Earth Syst. Sci.* **119**(2), 183–199 (2009)
11. Suzuki, Y., Michihiro, Y., Honma, M.: Development of strong wind nowcasting system. In: International Symposium on Extreme Weather and Cities, Tokyo, Japan (2013)
12. Nasution, B.B., et al.: Forecasting natural disasters of tornados using mHGN. In: Murayama, Y., Velev, D., Zlateva, P., Gonzalez, J. (eds.) ITDRR 2016. IFIPAICT, vol. 501, pp. 155–169. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68486-4_13
13. Nasution, B.B.: Towards real time multidimensional hierarchical graph neuron (mHGN). In: The 2nd International Conference on Computer and Information Sciences 2014 (ICCOINS 2014), Kuala Lumpur, Malaysia (2014)
14. Hoffmann, P., Merker, C., Lengfeld, K., Ament, F.: The Hamburg Tornado (7 June 2016) from the perspective of low-cost highresolution radar data and weather forecast model. *Atmos. Res.* **211**(1), 1–11 (2018)
15. Wilson, J.W., Stanley, J., Changnon, A.: Illinois Tornadoes. Illinois State Water Survey, Illinois (1971)
16. Guillot, E.M., et al.: Tornado and severe thunderstorm warning forecast skill and its relationship to storm type. In: 24th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology, New Orleans, USA (2008)
17. Winterrath, T., Rosenow, W.: The Radar-based Precipitation Nowcasting System RADVOR of Deutscher Wetterdienst for the support of meteorological and hydrological alert systems. In: 8th European Conference on Radar in Meteorology and Hydrology, Germany (2014)
18. Kryvasheyev, Y., et al.: Nowcasting Disaster Damage. arXiv preprint [arXiv:1504.06827](https://arxiv.org/abs/1504.06827) (2015)

19. Ogasawara, J., Tanimoto, K., Imaichi, O., Yoshimoto, M.: Disaster prevention and response support solutions. *Hitachi Rev.* **63**(1), 236–243 (2014)
20. Doong, D.J., et al.: Development of an operational coastal flooding early warning system. *Nat. Hazards Earth Syst. Sci.* **12**(1), 379–390 (2012)
21. Nasution, B.B., et al.: Realtime weather forecasting using multidimensional hierarchical graph neuron (mHGN). In: 16th International Conference on Neural Networks (NN 2015), Rome, Italy (2015)
22. Nasution, B.B., Khan, A.I.: A hierarchical graph neuron scheme for real-time pattern recognition. *IEEE Trans. Neural Netw.* **19**, 212–229 (2008)