

혼돈 이론을 이용한 일적산 일사량의 예측†

Prediction of Daily Solar Irradiance Based on Chaos Theory

조성인*	배영민*	윤진일**	박은우***	황헌****
정회원	정회원			정회원
S. I. Cho	Y. M. Bae	J. I. Yun	E. W. Park	H. Hwang

SUMMARY

A forecasting scheme for daily solar irradiance on agricultural fields is proposed by application of chaos theory to a long term observation data. It was conducted by reconstruction of phase space, attractor analysis, and Lyapunov analysis. Using the methodology, it was determined whether evolution of the five climatic data such as daily air temperature, water temperature, relative humidity, solar radiation, and wind speed are chaotic or not. The climatic data were collected for three years by an automated weather station at Hwasung-gun, Kyonggi-province. The result showed that the evolution of solar radiation was chaotic, and could be predicted. The prediction of the evolution of the solar radiation data was executed by using 'local optimal linear reconstruction' algorithm. The RMS value of the prediction for the solar radiation evolution was 4.32 MJ/m² · day. Therefore, it was feasible to predict the daily solar radiation based on the chaos theory.

주요용어(Key Words) : 혼돈 이론(Chaos theory), 기상 데이터(Climatic data), 예측(Prediction)

1. 서 론

농업 환경을 포함한 자연계에서 나타나는 동력학계의 상태 변수 대부분은 비선형성을 가지고 있으며, 동력학계의 초기 조건에 민감하게 반응하여 장기적으로 예측 불가능한 경우가 많다. 따라서, 기존의 동력학계 해석 방법은 비교적 선형적인 부분에 대해 '선형 근사화'하고 나머지 부분에 대해서는 미지의 분야로 남겨 두어야 했다. 그러나, 최근에는 동력학계에 내재되어 있는 결정론적 메커니즘(deterministic mechanism)을 인정하게 됨에 따라, 동력학계의 비선형적 특성에 조금씩 접근할

수 있게 되었다. 이렇게 비선형 동력학에 대해서 내재적 결정론적 메커니즘을 인정하여 해석하는 분야를 혼돈 이론(chaos theory)이라 하며, 이러한 혼돈 이론의 도입으로 인해 기존에는 임의적 현상으로 인정되었던 동력학계가 결정론적 메커니즘을 가진 혼돈 현상으로 인정받는 경우도 나타나고 있다(Cavalieri와 Kosak, 1994).

그러나, 혼돈 이론은 물리학과 수학에서 도출된 이론 및 알고리즘으로, 혼돈 이론을 응용하고자 하는 학자들의 입장에서는 난해한 부분들이 많이 있다. 따라서, 혼돈 이론을 자연 현상의 여러 방면에 적용하기 위한 체계를 세울 필요가 있으며, 혼

† 본 연구는 교육부 1998년도 교육부 학술연구 조성비(농업 과학 : A-95-6) 지원으로 수행되었음.

* 서울대학교 생물자원공학과

** 경희대학교 생명과학부

*** 서울대학교 응용생물화학부

**** 성균관대학교 생물기전공학과

돈 이론의 적용을 위한 도구가 필요하게 되었다. 이에 혼돈 이론의 입문에 대한 문헌들이 나왔으며 (Doscher, 1994), 이를 자연 현상 또는 사회 현상에 적용한 연구가 진행되었다(Roush 등(1994), Chavas 등(1991)). 그러나, 이들 문헌 역시 해가 알려진 동력학의 해석에 초점을 맞추고 있고, 실험 데이터에 대한 실제 분석 체계에 대해서는 미약하게 설명되어 있다. 따라서, 이를 보완하여 시계열 실험 데이터에 대한 혼돈 이론의 응용 체계가 세워져야 할 것이다.

한편 농업 방면에서 벼의 도열병, 잎집무늬마름병 등의 병 예찰을 위해서는 재배 지역의 각종 기상 데이터들이 필요하다. 그런데, 기상청에서 기온, 상대 습도, 풍속 등의 기상 데이터에 대해서는 예보를 생산하고 있으나, 일사량 같은 기상 데이터에 대해서 고려해야 할 변수가 많고, 비선형성이 강하여, 예보를 생산하지 못하고 있다. 다만 흐림, 비, 맑음 등의 정성적인 하늘 상태만을 예보하고 있기 때문에, 이 기상 데이터의 정량적 예보는 농업 방면의 작물의 병 예찰에 유용한 정보가 될 것이다.

따라서, 본 연구에서는 시계열 실험 데이터의 분석을 위한 혼돈 이론 적용 체계를 명확하게 제안하고, 각 적용 단계에 요구되는 지식 및 알고리즘을 컴퓨터 프로그램으로 구현하고자 한다. 그리고, 벼의 주요 병 예찰에 이용되는 수온, 기온, 풍속, 상대습도, 일사량 등의 기상 데이터의 일별 변화 양상을 혼돈 이론을 통해 분석하고, 진화의 예측을 시도하고자 한다.

2. 실험 재료 및 방법

가. 시계열 데이터의 혼돈 양상 분석

혼돈 양상은 동력학계에 내재되어 있는 결정론적 메커니즘에 의해서 결정되기 때문에, 동력학계의 여러 가지 인자들의 영향에 의해서 그 진화 방향이 결정되지만, 그 진화 형태의 걸보기는 임의적 양상으로 보여질 수 있다. 따라서, 진화 양상의 혼돈 양상 판별을 위해서는 동력학계의 양상을 기술할 수 있는 양을 계산하여 판별하여야 하며, 이러한 시계열 데이터의 혼돈 양상 판별 과정을 혼

돈 양상 분석이라 할 수 있다. 본 논문에서는 혼돈 양상 분석을 위상 공간(phase space)의 재구성, 끌개 분석(attractor analysis), 리아프노프 분석(Lyapunov analysis)의 3 단계로 제안하였다.

(1) 위상 공간(embedding dimension)의 재구성

혼돈 양상 분석을 위해서 우선 대상 시계열 자료를 위상 공간 안에 재구성해야 하며, 이는 Takens' Embedding 정리에 의해서 식 1로 설명될 수 있다(Takens, 1981).

Vector reconstructed in phase space =

$$(y_{s \times 1}, y_{s \times (t + \tau \times 1)}, y_{s \times (t + \tau \times 2)}, \dots, y_{s \times (t + \tau \times (D - 1))})$$

y_N : time-series data at time N

s : sampling time (1)

τ : time delay

D : embedding dimension

이와 같이, 시계열 자료를 위상 공간에 재구성하기 위해서는 동력학계의 끌개를 완전히 포함할 수 있는 위상 공간의 차원(embedding dimension)을 결정하여야 하며, 이는 식 2와 같이 상관 차원(correlation dimension)을 계산함으로써 이루어진다.

$$D > 2d + 1$$

D : embedding dimension (2)

d : correlation dimension

상관 차원은 Grassberger와 Procaccia(1983)가 처음으로 그 개념을 제안하였으며, 수학적으로는 끌개를 포함할 수 있는 최소 차원을 의미한다. 이후, 동력학계의 시계열 자료에 대한 상관 차원 계산 알고리즘이 출현했으며(Albano 등, 1988), 이들은 모두 Grassberger와 Procaccia(1983)가 제안한 알고리즘(G-P 알고리즘)을 근간으로 한다. G-P 알고리즘은 식 3에서 표현된 $\log(r)$ 에 대한 D_2 의 스펙트럼을 계산하고, 위상 공간의 차원이 증가함에 따라 수렴하는 D_2 스펙트럼의 기울기를 상관 차원으로 결정한다. 따라서, 본 연구에서는 상관 차원의 분석을 위해서 G-P 알고리즘을 선택하였으며, D_2 스펙트럼을 계산하는 "Dim_all.exe"이라는 컴퓨터 프로그램을 작성하였다.

$$D_2(M;r) = \frac{d \log_e C_M(r)}{d \log_e(r)} \quad (3)$$

$C_M(r)$: correlation integral

M : dimension of phase space

(2) 끌개 분석(attractor analysis)

위상 공간의 차원을 결정하면, 시계열 자료의 진화를 위상 공간 안에 나타낼 수 있으며, 그 진화 궤적을 끌개(attractor)라 부른다. 일반적으로 끌개 궤적의 형태에 따라 시계열 자료의 특성을 분석할 수 있는데, 수렴 양상의 시계열 자료는 한 점으로 수렴하는 형태의 끌개로 나타나며, 주기적, 준주기적 양상의 시계열 자료는 원형으로 수렴하는 끌개로 나타난다. 그리고, 혼돈 양상 데이터는 끌개의 궤적이 서로 교차하지 않으며 무한대로 진행되는 형태(strange attractor)를 띤다(Rough 등, 1994). 따라서, 끌개의 형태 분석은 가시적으로 혼돈 양상을 판정할 수 있는 장점이 있지만, 위상공간의 차원이 4차원 이상일 경우에는 가시적 분석이 불가능하며, 실험에서 발생하는 시계열 자료의 끌개는 그 양상이 쉽게 드러나지 않는다는 단점이 있다.

(3) 리아프노프 분석(Lyapunov analysis)

리아프노프 지수는 끌개 분석의 단점을 보완하기 위해, 위상 공간 안에서 나타나는 끌개 궤적의 팽창 또는 수축 정도를 정량화한 수치이며, 동력학계의 리아프노프 지수를 계산하여 혼돈 양상을 판별하는 것을 리아프노프 분석(Lyapunov analysis)이라 한다(Nayfeh와 Balachandran, 1995).

리아프노프 지수의 수학적 정의는 식 4와 같으며, 이는 동력학계의 초기값에 대한 t 시간 진화된 동력학계의 비(ratio)의 대수값을 의미한다. 따라서, 동력학계의 리아프노프 지수가 양의 값이라는 것은 끌개가 계속적으로 팽창한다는 것을 의미하며, 이러한 동력학계는 혼돈 양상으로 판정할 수 있다. 그리고, 시계열 데이터의 진화 양상을 위상 공간 안에 재구성하였을 때 각 차원의 리아프노프 지수 중에서 하나만이라도 양의 값을 가지면, 혼돈 양상으로 판정할 수 있다.

In system given as $y(t) = \Phi(t) y(0)$

$$\bar{\lambda}_i = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\|y(t)\|}{\|y(0)\|} \quad (4)$$

where $\bar{\lambda}_i$ is Lyapunov exponent.

본 연구에서는 Eckmann 등(1986)이 소개한 알고리즘에 기초하여 “Lyap_exp.exe” 라는 리아프노프 지수 계산 프로그램을 작성하였다.

나. 혼돈 양상 데이터의 예측

리아프노프 지수가 양의 값인 혼돈 양상 데이터는 초기값에 대한 데이터의 진화가 각종 내재적 원인에 의해서 기하급수적으로 확장하여 중장기적으로는 예측 불능인 상태로 진화하기 때문에 (Nayfeh와 Balachandran, 1995), 혼돈 양상 데이터의 중장기적 예측은 불가능한 것으로 알려져 있다. 따라서, 최근에는 혼돈 데이터 예측에 관한 연구는 결정론적 메커니즘에 따른 단기 예측 알고리즘의 개발에 집중되어 있다(Farmer와 Sidorowich (1987), Casdagli(1988), Jimenez 등(1992)). 현재까지 개발된 예측 알고리즘은 대부분 서로 비슷한 과정에 따라 전개되고 있으며, 신경회로망의 비선형적 적합 특성을 이용하여 혼돈 양상 데이터의 예측 알고리즘을 개발한 예도 있다(Mees 등, 1992).

본 연구에서는 Jimenez 등(1992)이 제안한 local optimal linear-reconstruction method를 예측 알고리

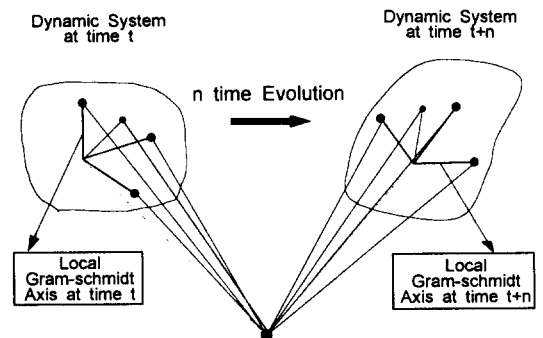


Fig. 1 The schematic diagram showing the local optimal linear reconstruction method.

즘으로 하여 혼돈 데이터 예측 프로그램을 개발하였다. 알고리즘의 기본 개념은 그림 1에서 표현된 바와 같이, 혼돈 양상의 동역학계에서는 현재 시간(t)에서의 상태 변수 벡터는 인접 상태 변수 벡터(neighborhood vectors)들에 의해서 형성된 Gram-Schmidt 기저들에 표현될 수 있으며, 짧은 시간(n) 동안의 진화 후에는 진화된 인접 벡터들에 의해서 형성된 Gram-Schmidt 기저에 의해서 미래(t+n)의 값이 표현될 수 있다는 것이다.

본 연구에서는 이 알고리즘에 기초하여 "prdchaos.exe"라는 혼돈 양상 데이터의 단기 예측 프로그램을 작성하였다.

다. 농업 기상 데이터의 수집

경기도 화성군 농촌진흥원 소속의 벼 재배 실험 포장에 설치되어 있는 무인기상관측기를 통해서 각종 기상 데이터를 수집하였다. 수집된 데이터는 1995년 겨울부터(18일째부터) 1997년 가을까지의 일별 데이터이며 총 996일 동안의 기록이다. 이 중, 벼의 병해충 예찰 모형의 입력 변수로 이용되는 기온, 수온, 상대습도, 일사량, 풍속 등의 5가지 변수를 선정하였으며, 이에 대한 혼돈 양상 분석을 수행하였다.

그림 2는 각 기상 데이터 중, 일사량의 변화 양상이다. 데이터의 진화는 전반적으로 1년의 주기

로 변화하고 있으며, 세부적으로는 잡음에 오염된 것처럼 보이는 임의적 변화 양상을 나타내고 있는 것으로 판단되었다.

3. 결과 및 고찰

가. 위상공간의 재구성

기상데이터에 대한 리아프노프 분석을 수행하기 위해서, 먼저 위상 공간 안에 기상데이터를 재구성해야 하며, 이를 위해서 상관 차원을 계산하고 위상 공간의 차원을 결정하였다.

상관 차원의 계산은 위상 공간의 차원별로 식 1에 의해서 계산되는 D_2 의 스펙트럼 분석을 통해서 이루어졌다. 이를 위해서 본 연구에서 작성된 "Dim_all.exe" 프로그램을 이용하여 각 기상 데이터별로 D_2 의 스펙트럼을 계산하였다. 기온, 수온, 상대습도, 일사량에 대해서는 위상 공간의 차원을 15까지 확장시켜 가면서 스펙트럼을 구하였으며, 풍속의 경우에는 위상 공간의 차원을 22까지 확장시켜 가면서 스펙트럼을 계산하였다. 그림 3은 일사량에 대한 D_2 스펙트럼의 기울기의 수렴 양상이다. 상관 차원은 D_2 의 스펙트럼에서 위상 공간의 차원이 증가함에 따라 수렴하는 D_2 의 기울기값으로 결정된다. 그런데, 위의 각 기상 데이터별 D_2 스펙트럼에서는 $\log(r)$ 의 작은 값(<-2)에서는 수렴

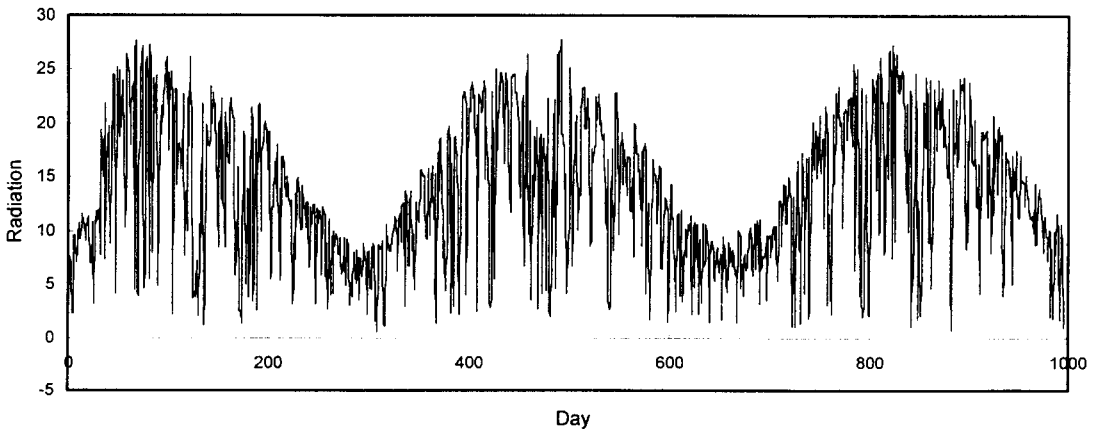


Fig. 2 The evolution of daily radiation obtained by AWS.

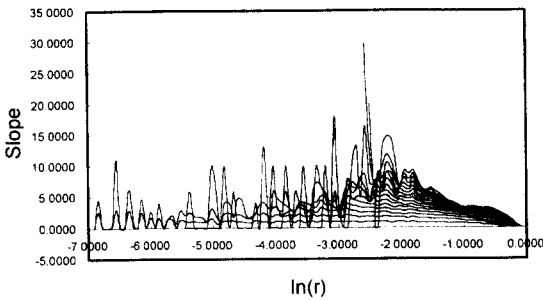


Fig. 3 The plot of slope versus $\log(r)$ according to delay dimension for daily radiation.

이 나타나지 않고 오히려 요동을 일으키고 있다. 이러한 양상은 미지의 동력학계의 실험 데이터에 포함될 수 있는 계측 오차 및 계산 오차 그리고 임의적 잡음 등에 의해서 발생하는 것으로 알려져 있다(Albano 등, 1988). 따라서, 본 연구에서는 $\log(r)$ 의 큰 값(>-2)에 대해서 국부적으로 수렴하는 평균 기울기를 찾아내려고 하였다.

각 기상데이터에 대해서 위상 공간 차원 증가에 따른 국부적 평균 기울기값의 수렴 양상을 그림 4에 나타내었다. 기온, 수온, 일사량 등의 수렴 양상은 양호한 형태를 나타내고 있으나, 상대습도와 풍속의 경우 그 수렴하는 정도가 완연하지 못한 경향이 있다. 이는 위상 공간의 차원이 증가함에 따라 분석 데이터 수의 부족으로 나타나는 현상으로 추후 더 많은 데이터를 이용한 분석이 필요할 것으로 판단된다(Albano 등, 1988). 따라서, 수렴 양상이 보이는 기온, 수온, 일사량에 대해서만 상관 차원과 위상 공간의 차원을 결정하였다(표 1).

나. 리아프노프 분석

위의 상관 차원 분석에서 도출한 위상 공간의 차원에 대해서 리아프노프 지수를 계산하였다. 본 연구에서 개발한 "lyap_exp.exe" 프로그램을 이용하여 기온, 수온, 일사량에 대한 리아프노프 지수를 구하였으며, 그림 5는 각 기상데이터에 대한 위상 공간의 차원별 리아프노프 지수값을 보여주는 리아프노프 스펙트럼(Lyapunov spectrum)이다.

일별 기온 및 수온의 경우 각 차원의 리아프노프 지수의 값이 모두 음의 값을 가진다. 이는 기

온 및 수온의 시계열 자료의 끝개가 수축한다는 것을 의미하며, 따라서, 혼돈 양상의 데이터가 아닌 것으로 판정하였다. 일별 일사량의 경우에는 1, 2, 3 차원에서 양의 값을 나타내고 있기 때문에, 혼돈 양상을 가지는 진화 양상으로 판정하였다.

다. 혼돈 데이터의 예측

위의 리아프노프 분석을 통해 혼돈 양상이 나타난다고 판단되는 일사량의 진화에 대해서, local optimal linear-reconstruction 알고리즘을 이용하여 혼돈 양상의 진화 예측을 시도하였다. 각 데이터 중 초기 900 일의 데이터를 근접 벡터 검색을 위한 기준 데이터로 활용하여 후기의 데이터에 대해서 예측을 시도하였다.

그림 6은 일별 일사량 데이터의 1일 후의 진화 양상을 예측한 결과이다. 예측 결과의 성능을 평가하기 위해서 식 5에 의한 RMS 값을 계산하였으며, 일사량의 RMS값은 $4.32 \text{ MJ/m}^2 \cdot \text{day}$ 이었다.

$$\text{RMS} = \sqrt{\frac{\sum(x_i - y_i)^2}{N}} \quad (5)$$

x_i : real value

y_i : predicted value

그리고, 현재 시점에서의 예측 시점을 1일, 3일, 5일, 7일로 증가시켜가면서 진화 양상의 예측을 시도하였으며, 표 2는 그 결과를 보여 준다.

표 2에서 알 수 있듯이, 예측 기간(prediction interval)이 길어짐에 따라, RMS의 값이 점점 커지는 것을 알 수 있다. 이는 혼돈 양상 데이터의 일반적인 특징으로서, 리아프노프 지수의 값이 양임에 따라 시계열 데이터의 끝개가 계속적으로 팽창한다는 것을 의미한다(Jimenez 등, 1992).

하지만, 전반적으로 RMS의 값을 근거로 하여 예측 성능을 판정하였을 때, 대체적인 경향은 일치하나 구체적인 일별 차이는 RMS의 값이 $4.32 \text{ MJ/m}^2 \cdot \text{day}$ 의 오차를 보이고 있다. 이러한 이유는 두 가지로 요약할 수 있다. 첫 번째는 앞에서 언급하였다시피, 실험 데이터의 계측 오차, 계산 오차 및 임의적 양상의 오염이다. 일반적으로 실험 데이터는 필연적으로 계측 오차 또는 계산 오차를

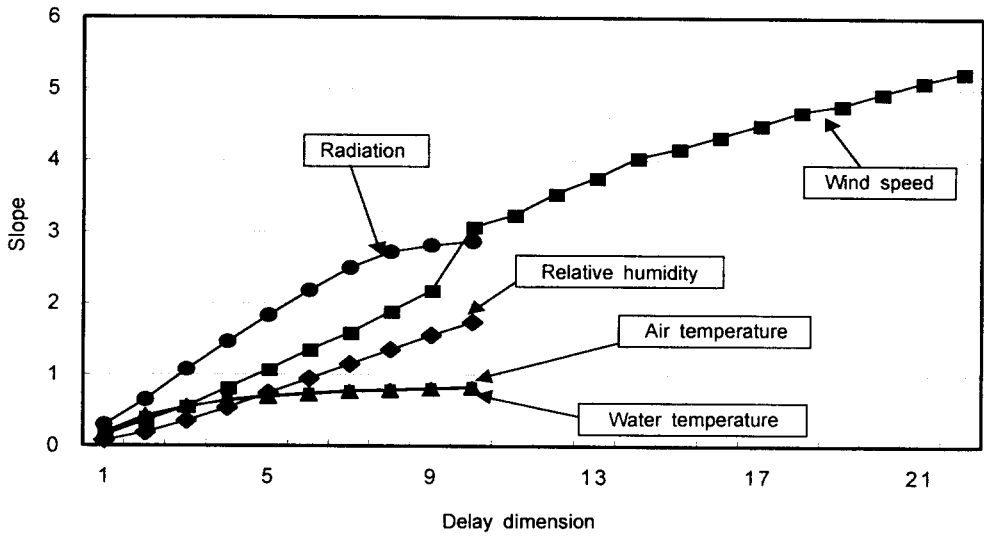


Fig. 4 The plot of delay dimension vs slope of D_2 spectra for climatic data (air temperature, water temperature, relative humidity, and radiation).

Table 1 The correlation dimension and embedding dimension for each climatic data

	Air temperature	Water temperature	Radiation
Correlation dimension	0.85	0.85	3.1
Embedding dimension	3	3	8

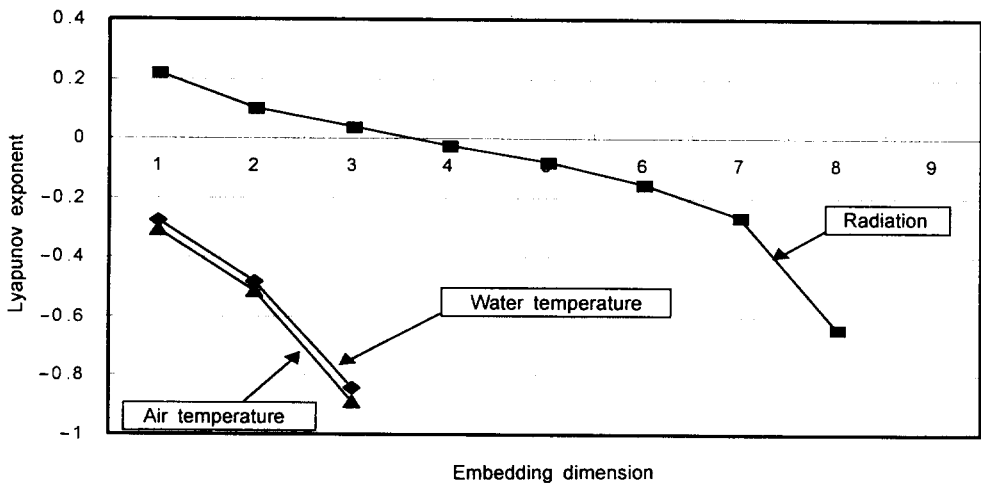


Fig. 5 The Lyapunov spectra of climatic data (air temperature, water temperature, radiation).

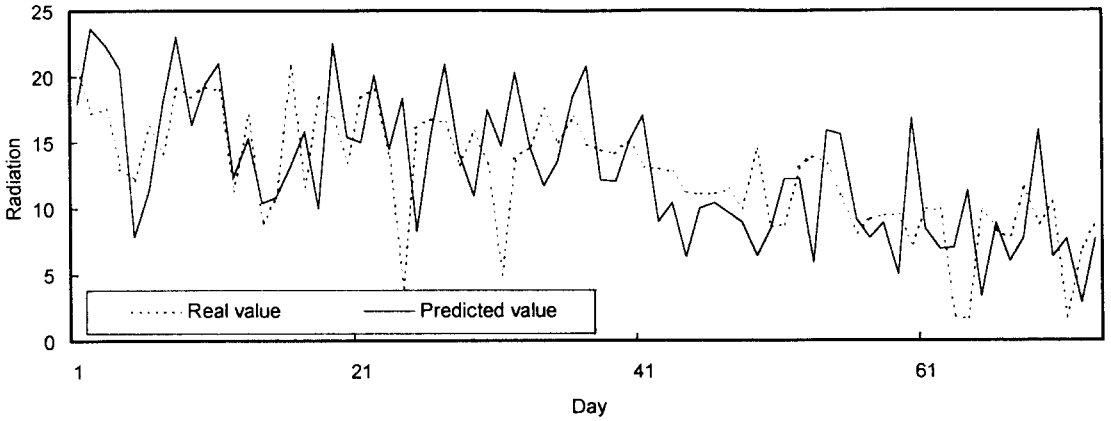


Fig. 6 The 1 day ahead step prediction for the evolution of radiation.

Table 2 The change in RMS of prediction according to increase of prediction step

Prediction step	Radiation (MJ/m ² · day)
1 day step	4.32
3 day step	8.97
5 day step	10.89
7 day step	8.48

포함하게 되고, 이들이 상관 차원의 수렴에 영향을 줄 수 있다. 따라서, 앞에서 상관 차원을 계산할 때, r 의 작은 값에서 D_2 가 수렴하지 않고 요동의 원인을 제공하며, 예측 알고리즘에서 인접 벡터를 선택하는 데 영향을 주게 되어 예측 성능을 저하시킨다. 두 번째 이유는 수집된 데이터 수의 부족이다. 실험 데이터 수의 제한은 상관 차원의 계산에서 r 의 큰 값 영역에서의 D_2 의 수렴을 어렵게 할뿐만 아니라, 시계열 데이터의 혼돈 진화 양상을 충분히 내재하지 못하여 혼돈 데이터의 예측에 영향을 줄 수 있다. 혼돈 양상 분석에 필요한 시계열 데이터의 수는 보통 $10^4(d : \text{embedding dimension})$ 이상인 것으로 알려져 있으나 (Doscher, 1994), 본 연구에서 분석한 실험 데이터의 수는 996개로서 혼돈 양상을 분석 및 예측 성

능을 향상시키기 위해서는 더 많은 데이터가 필요하다. 이러한 점은 일반적인 선형 회귀 분석 등에서도 동일하다 할 수 있다.

4. 요약

혼돈 이론을 시계열 데이터 분석에 활용하기 위해 위상 공간의 재구성, 끌개 분석, 리아프노프 분석의 단계로 이루어진 적용 체계를 제안하였으며, 각 단계별로 요구되는 도구를 컴퓨터 프로그램으로 구현하였다. 그리고, 수립된 적용 체계를 활용하여 기상 데이터의 진화 양상 분석을 시도하였다. 이를 위해서, 경기도 화성군 농촌진흥원내 실험포장에 설치된 무인기상관측기를 통해, 일별 기온, 수온, 상대습도, 일사량, 풍속 등의 기상데이터를 수집하였다. 위상 공간의 재구성 및 리아프노프 분석 등을 통해 일사량의 진화가 혼돈 양상을 가지는 것으로 판정하였다.

그리고, 혼돈 양상으로 판정된 일사량에 대해서 혼돈 양상 데이터의 단기 예측 알고리즘인 local optimal linear reconstruction법을 이용하여 진화 양상 예측을 시도하였고, 예보 선행 기간이 하루인 경우, RMS값은 4.32 MJ/m² · day이었다. 따라서, 혼돈이론을 이용하여 일별 일사량의 단기 예측의 가능성을 발견할 수 있었으며, 예측 성능을 향상시키기 위해서는 수집된 기상 자료의 정확성과 더 많은 기상 자료의 확보가 필요하다.

참 고 문 헌

1. Albano, A.M., J. Muench, C. Schwartz, A.I. Mees and P.E. Rapp. 1988. Singular-value decomposition and the Grassberger-Procaccia algorithm : in *Coping with chaos*. J. Willey, 1994.
2. Casdagli, M. 1988. Nonlinear prediction of chaotic time series : in *Coping with chaos*. J. Willey, 1994.
3. Cavalieri, L.F. and H. Koşcak. 1994. Chaos in biological control systems, *J. Theoretical Biology* 169:179-187.
4. Chavas, J.P. and M.T. Holt. 1991. On nonlinear dynamics: The case of pork cycle. *American Journal of Agricultural Economy* August:820-828.
5. Doscher, C. 1994. The application of nonlinear dynamics to complex systems. An ASAE Meeting Presentation Paper No. 943538.
6. Eckmann J.P., S. Oliggson Kamphorst, D. Ruelle and S. Ciliberto. 1986. Liapunov exponents from time series : in *Coping with chaos*. J. Willey, 1994.
7. Farmer, J.D. and J.J. Sidorowich. 1987. Predicting chaotic time series : in *Coping with chaos*. J. Willey, 1994.
8. Grassberger, P. and I. Procaccia. 1983. Characterization of strange attractors, *Physical review letters* 50(5):346-349.
9. Jimenez, J., J.A. Moreno and G.J. Ruggeri. 1992. Forecasting on chaotic time series: A local optimal linear-reconstruction method. *Physical review A* 45(6):3553-3558.
10. Mees, A., K. Aihara, M. Adachi, K. Judd, T. Ikeguchi and G. Matsumoto. 1992. Deterministic prediction and chaos in squid axon response. *Physics letters A* 169:41-45.
11. Nayfeh, A.H. and B. Balachandran. 1995. *Applied Nonlinear dynamics*. John Wiley & Sons, NewYork.
12. Roush, W.B., G.F. Barbato and T.L. Cravener. 1994. A nonlinear dynamical (chaos) approach to the analysis of broiler growth. *Poultry Science* 73:1183-1195.
13. Takens, F. 1981. Detecting strange attractors in turbulence : in *Lecture Notes in Mathematics*. D.A. Rand and I.S., Young eds., Springer-Verlag, NewYork, 366-381.