

Received: 2024/07/09  
Revised: 2024/07/15  
Accepted: 2024/08/27  
Published: 2024/09/30

**\*Corresponding Author:**

**Chang Woo Hong**

Dept. of Mechanical System Engineering,  
Republic of Korea Naval Academy  
1 Jungwon-ro, Jinhae-gu, Changwon-si,  
Gyungsangnam-do, 51704, Republic of Korea  
Tel: +82-55-907-5307  
E-mail: spearw@navy.ac.kr

# 다변량 데이터 분석을 통한 효과적인 전력부하 예측: Gas Carrier 사례 연구

## Effective Power Load Prediction Through Multivariate Data Analysis: Gas Carrier Case Study

홍창우<sup>1\*</sup>, 김소연<sup>2</sup>

<sup>1</sup>해군 소령/해군사관학교 기계시스템공학과 조교수

<sup>2</sup>해군 중령/해군사관학교 전기전자공학과 교수

**Chang Woo Hong<sup>1\*</sup>, So Yeon Kim<sup>2</sup>**

<sup>1</sup>LCDR, ROK Navy/Assistant professor, Dept. of Mechanical System  
Engineering, Republic of Korea Naval Academy

<sup>2</sup>CDR, ROK Navy/Professor, Dept. of Electrical Engineering, Republic of Korea  
Naval Academy

### Abstract

본 논문에서는 다변량 시계열 데이터의 특징을 보이는 gas carrier의 전력부하를 인공신경망 네트워크 아키텍처를 구성하여 예측한다. 네트워크는 CNN-RNN 기반으로 구성하며 효과적인 학습과 예측 정확도 향상을 위해 차원 축소인 주성분분석을 사용하였으며, CNN 알고리즘에서 dilation rate를 달리하며 학습의 결과를 비교하였다. 학습 결과 주성분분석과 Dilated CNN을 적절히 사용하여 다변량 데이터를 높은 정확도로 예측할 수 있음을 보였다.

In this paper, the power load of a gas carrier, which shows the characteristics of multivariate time series data, is predicted by configuring an artificial neural network architecture. The network was constructed based on CNN-RNN, and principal component analysis, a dimensionality reduction, was used to improve effective learning and prediction accuracy, and the learning results were compared by varying the dilation rate in the CNN algorithm. The learning results showed that multivariate data can be predicted with high accuracy by appropriately using PCA and Dilated CNN.

### Keywords

인공신경망(Artificial Neural Network),  
확장된 합성곱 신경망(Dilated CNN),  
가스 운반선(Gas Carrier),  
주성분분석(Principal Component Analysis)

## 1. 서론

해양환경 오염 방지에 대한 관심이 계속 높아지고 있다. 국제해사기구(IMO, International Maritime Organization)에서는 전 세계 모든 선박(400 t급 이상)의 배기가스의 황산화물(SOx) 함유율을 현행 '3.5 % 미만'에서 '0.5 % 미만'으로 낮춘 저유황유를 사용하도록 강제하는 IMO 2020을 2020년부터 시행했다[1]. 해운 업계는 연료유를 바꾸거나 배출 가스의 황산화물 저감장치(스크러버)를 장착하는 등의 대책을 강구하고 있다. 혹은 전기추진선박을 활용하거나 gas carrier의 경우 BOG(boil off gas)를 사용하여 환경 오염물질 배출을 줄이면서 연료 소모효율을 높이는 방식으로 친환경 정책을 따라가고 있다[2].

특히 전기추진선박 혹은 전력 부하가 큰 선박에서는 전력 부하에 대한 예측과 배터리 등을 사용한 peak shaving 방법을 통해 연료 소모량을 절감할 수 있다. 하지만 다양한 장비에서 전력을 소모하며 운전 조건에 따른 부하의 변동이 큰 선박에서는 전력부하 예측을 정확하게 하기 어렵다는 문제점이 있다. 또한 선박의 운영 특성상 다양한 환경요소 하에서 다양한 부하의 불규칙적인 변동이 있으므로 다변량 시계열 예측 문제가 가지는 어려움을 포함하게 된다. 따라서 여러 가지 운전 조건 등의 센서 데이터를 얼마나 잘 정제하여 정확한 예측을 하는지가 중요한 문제라고 할 수 있다.

본 연구에서는 다변량 시계열 데이터의 특징을 가진 gas carrier의 실 운전 데이터를 바탕으로 정확한 예측을 할 수 있는 네트워크를 제안한다. 시계열 데이터의 분석에 적합한 CNN-RNN 네트워크 구조를 바탕으로 하며, 보다 정확한 특징 추출을 위하여 Dilated CNN을 적용하여 그 결과를 비교한다. 또한, 다변량 입력 변수의 차원을 줄이기 위해 주성분 분석을 사용하였다.

## 2. 연구 배경

다변량 시계열 데이터의 예측은 딥러닝이라고도 불리는 심층 신경망과 여러 알고리즘을 통해 가능하다. 우리가 흔히 하는 날씨나 주가 예측, 건강 상태의 예측 등을 심층 신경망 모델을 통해 수행할 수 있으며 심층 신경망 모델을 구성하는 알고리즘으로는 CNN(convolutional neural network)과 RNN(recurrent neural network)을 대표적으로 들 수 있다. CNN 알고리즘은 합성곱(convolution)을 이용하여 특징을 추출하는 역할로 주로 이미지나 영상에 대한 분류(classification)의 용도로 사용된다. 하지만 다변량 시계열의 데이터의 경우 많은 입력 데이터로부터 불필요한 데이터와 잡음 등을 제거하고, 특징을 추출하는 데 사용될 수 있다. RNN 알고리즘은 시계열 관점에서 데이터의 다양한 추세 변화를 기억하여 시간적 변화를 추종할 수 있다.

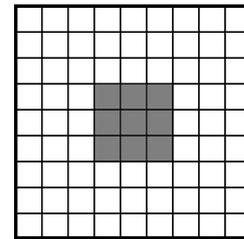
CNN 알고리즘과 RNN 알고리즘은 각각 장점과 단점을 가지고 있으며, 데이터의 특성과 학습의 목적에 따라 적절한 네트워크를 구성하여 사용할 수 있다. CNN-RNN 기반 네트워크는 CNN의 특징 추출과 RNN의 시간적 변화 추종의 장점을 모두 가져온 네트워크로 특히 다변량 시계열 예측에서 효과적으로 사용될 수 있다. 이미지 추출 분야에서 CNN-RNN 네트워크가 처음 제안된 이후[3], 다변량 데이터를 활용한 날씨 예측[4], 의학적 예지 분야[5]까지 여러 방면으로 활용되고 있다.

본 연구에서는 CNN-RNN 네트워크를 활용하여 gas carrier의 전력 부하 예측을 수행한다. CNN-RNN 네트워크를 사용함에 있어 입력 변수가 많은 점에 착안해 CNN 레이어 부분에서 Dilated CNN을 적용하고, 차원 축소 기법으로 입력 변수를 줄이는 주성분분석을 적용하였다.

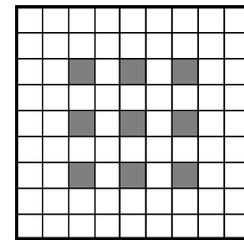
### 2.1 Dilated CNN

CNN 알고리즘은 필터와 합성곱을 사용하여 입력 데이터에서 공간적인 특징을 추출한다. 특징을 잘 추출하기 위해서는 필터가 특징을 추출하는 면적인 receptive field를 넓게 봐야 하는데, Dilated CNN은 필터의 크기를 늘리지 않으면서 receptive field를 늘려줄 수 있는 장점이 있다[6].

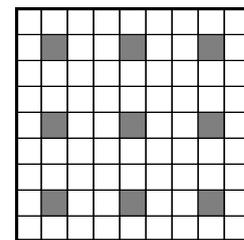
Fig. 1은 dilation rate에 따른 feceptive field의 면적을 나타낸다. Dilation을 잘 활용하면 연산량을 늘리지 않고, 과적합을 방지하면서 보다 특징을 잘 추출할 수 있다.



(a) Rate = 0



(b) Rate = 1



(c) Rate = 2

**Fig. 1.** The receptive field of dilated convolution

$$(F * k)(p) = \sum_{s+t=p} F(s)k(t) \quad (1)$$

$$(F *_l k)(p) = \sum_{s+lt=p} F(s)k(t) \quad (2)$$

식 (1)은 일반적인 합성곱이며 \*는 합성곱 연산자이다. 식 (2)에서 \*<sub>l</sub>이 dilated convolution이며 이때 l이 dilation factor가 된다. 본 연구에서는 다변량 시계열 데이터의 특징을 보다 잘 추출하기 위해 일반적인 CNN 알고리즘과 dilation rate를 달리하였으며, Dilated CNN을 적용하여 정확도를 비교 분석하였다.

2.2. 차원 축소와 주성분 분석

입력 데이터의 차원이 너무 크거나 연산 부담을 줄여야 할 때 차원 축소(dimension reduction) 기법을 사용한다. 여러 차원 축소 기법 중 본 연구에서는 주성분분석과 독립성분분석 기법을 적용하였다. 주성분분석(PCA, principal component analysis)은 상관관계가 있는 다수의 변수로부터 상관관계가 없는 소수의 변수로 불규칙한 분포를 최대한의 상관관계로 표현하는 주성분(principal component)으로 합성하는 분석 기법이다. 불필요한 변수들이 학습될 경우 예측 정확도가 낮아지거나, 비슷한 변수들이 학습될 경우 다중공선성(multicollinearity) 문제가 발생하므로, 여러 가지 변수를 합쳐서 압축시켜 데이터를 가장 잘 설명할 수 있는 축을 찾는 역할을 하게 된다.

3. 실험 연구

3.1 인공신경망 네트워크 아키텍처

학습에 사용한 인공신경망 네트워크의 구성은 Fig. 2와 같다. 네트워크는 CNN 알고리즘 레이어와 RNN 알고리즘 레이어를 동시에 사용하여 총 6개 레이어로 구성되었다. 입력 데이터의 변수가 80개 이상으로 차원 축소 기법인 주성분분석 또는 독립성분분석을 사용하여 축소된 차원의 데이터를 입력한다.

입력된 데이터는 3개의 CNN 레이어를 거친다. CNN 레이어는 여러 변수의 공간적인 특징 추출이 가능하다. 또한, 합성곱을 통해 노이즈를 제거하고 데이터를 간략화하여 원하는 특징을 추출할 수 있다. 이때 dilation을 적용한 CNN 레이어의 정확도를 일반 CNN 레이어와 비교한다.

CNN 레이어 이후는 RNN 알고리즘을 적용하여 gas carrier의 전력부하에 대한 시간적인 추세 요인을 학습한다. 전체적인 네트워크에는 각 레이어 이후 출력 일부를 임의로 누락하여 과적합 방지가 가능한 dropout 기법을 적용하였다. 이때 dropout rate는 20 %이다. 네트워크의 CNN 레이어 이후에는 최초의 입력 데이터를 합하여 추가적인 연산량 증가 없이 입력 데이터의 특징을 학습하는 residual network를 사용하여 정확도 향상에 기여한다.

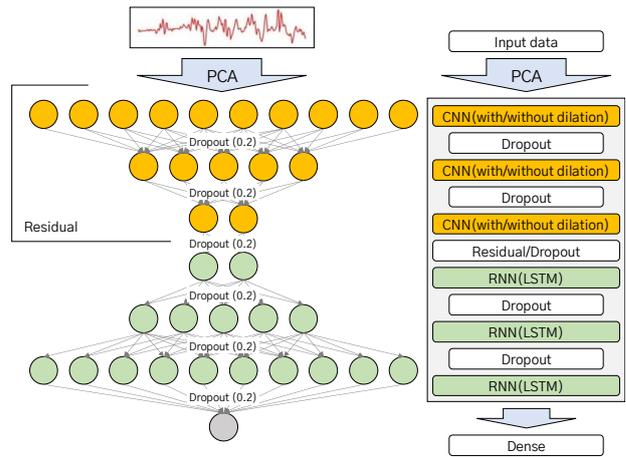


Fig. 2. The proposed CNN-RNN network architecture

3.2 하이퍼파라미터와 평가지표

Gas carrier의 전력부하 예측을 위한 입출력 레이어 유닛 구성을 포함한 하이퍼파라미터는 Table 1과 같다. 선박에 대한 상세한 제원은 생략하도록 한다. 입력 데이터는 BOG와 LNG 탱크, 발전기 등 전력부하 예측과 관련된 변수의 데이터를 모두 포함한 81개의 feature로 이루어져 있다. 임의로 변수를 삭제하지 않고 차원 축소를 적용하여 정확도를 비교하였다. 변수의 변화 추이를 학습하기 위해 데이터 입력 후에는 각 변수의 수치를 0과 1 사이의 값으로 정규화하였다. 학습과 시험 데이터셋은 7:3의 비율로 나누어 구성하고, 전체 데이터셋의 포인트는 약 30,000포인트이다. 총 반복횟수는 1,000회로 설정하며 학습 결과에 따른 조기 종료 가능한 early stopping 기법을 적용하였다. 예측 결과의 정확도는 3가지 지표인 mean squared error(MSE)와 mean absolute error(MAE), 그리고 root mean squared error(RMSE)로 나타냈다.

**Table 1.** Hyperparameters for experimental study

Hyperparameters	Value
Number of nodes	512-256-40-40-256-512
Normalization	Min-max normalization
Activation function	ReLU
Validation	10 % of data
Optimizer algorithm	Adam
Epochs	1,000(early stopping)

#### 4. 결과 분석

Gas carrier의 전력 부하 예측 결과는 Fig. 3와 같다. 본 연구에서 제안한 인공신경망 네트워크 아키텍처는 예측 데이터의 다이내믹한 부하 변동을 잘 추종하였음을 확인할 수 있다. 학습에서는 81개의 변수를 주성분분석을 통해 40개로 축소하여 예측을 진행하였다.

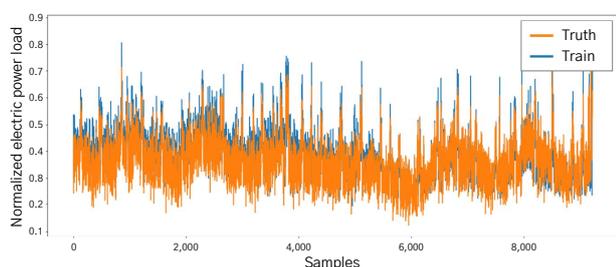
**Fig. 3.** Electric power load forecasting result(best case)

Table 2는 CNN-RNN 네트워크를 이용한 예측 결과와 Dilated CNN의 rate 변동에 따른 예측 결과를 나타낸 것으로, 각각 MSE, MAE, MAPE로 오차를 비교하였다. 일반적인 CNN-RNN 네트워크를 사용하였을 때보다 Dilated CNN을 사용하였을 때 정확도가 향상됨을 확인할 수 있다. 또한 dilation rate를

증가시키며 예측했을 때보다 정확한 것으로 나타났다. 이는 많은 변수에 대한 예측에 대해 dilation을 통해 연산량을 증가시키지 않으면서도 receptive field 확장으로 특징 추출을 보다 잘 수행하여 효과적인 예측을 가능하게 함을 보여준다. 또한, 학습에서 dropout 기법과 residual network를 사용하여 과적합을 방지하였다.

**Table 2.** Accuracy comparison by cases

Case	RMSE	MAE	MAPE
No dilation	0.1456	0.1209	36.018
Dilation rate = 1	0.1205	0.0977	27.865
Dilation rate = 2	0.0954	0.0767	22.391
Dilation rate = 3	0.0610	0.0480	13.897

#### 참고문헌

- [1] Li, Kevin, et al. "Determinants of ship operators' options for compliance with IMO 2020." *Transportation Research Part D: Transport and Environment* 86 (2020): 102459.
- [2] Uyanik, Tayfun, et al. "A data-driven approach for generator load prediction in shipboard microgrid: The Chemical Tanker Case Study." *Energies* 16.13 (2023): 5092.
- [3] Wang, Jiang, et al. "CNN-RNN: A unified framework for multi-label image classification." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- [4] Zhao, Bin, et al. "A CNN-RNN architecture for multi-label weather recognition." *Neurocomputing* 322 (2018): 47-57.
- [5] Zhou, Xiaokang, Yue Li, and Wei Liang. "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 18.3 (2020): 912-921.
- [6] Yu, Fisher, and Vladlen Koltun. "Multi-scale context aggregation by dilated convolutions." *arXiv preprint arXiv:1511.07122* (2015).