

## 주급수 유량의 유효 모델(커널 회귀)에 대한 연구

양학진<sup>1</sup>, 김성근<sup>2\*</sup>

<sup>1</sup>제주관광대학교 메카트로닉스과, <sup>2</sup>호서대학교 기계자동차공학부

### A Study of the Valid Model(Kernel Regression) of Main Feed-Water for Turbine Cycle

Hac-Jin Yang<sup>1</sup>, Seong-Kun Kim<sup>2\*</sup>

<sup>1</sup>Dept. of Mechatronics, Jeju Tourism University

<sup>2</sup>Division of Mechanical and Automotive Engineering, Hoseo University

**요약** 터빈 사이클 보정 열 성능 분석은 발전소의 현재 성능을 결정하고 향상된 경제성 운전을 위해 요구된다. 본 연구에서는 신뢰성있는 성능 분석을 위해서 산업 표준인 ASME(American Society of Mechanical Engineers) PTC(Performance Test Code)를 기본으로 성능 분석에서 우선적으로 중요하게 적용되는 주급수 유량을 대상으로 영역별 판정 알고리즘을 개발하고 각 영역별로 현재의 터빈 사이클 성능을 추정하는 알고리즘을 개발하였다. 추정 알고리즘은 측정 상태량의 상관 관계를 기반으로 영역별로 형상 분류를 제시하고, 이를 기반으로 커널 회귀 모델을 이용하여 학습된 추정 모델을 구성하였으며, 커널 회귀 모델링의 우수성을 검증하기 위하여 신경 회로망 모델의 학습 결과와 비교하였다. 주급수 유량의 형상 특성에 따른 분류 및 추정 모델은 터빈 사이클에서 정확한 보정 열 성능 분석을 제공함으로써 성능 분석의 신뢰도를 증가시킬 수 있었으며 다른 성능 결정 변수에 대한 학습 및 검증 모델로 사용될 수 있다.

**Abstract** Corrective thermal performance analysis is required for power plants' turbine cycles to determine the performance status of the cycle and improve the economic operation of the power plant. We developed a sectional classification method for the main feed-water flow to make precise corrections for the performance analysis based on the Performance Test Code (PTC) of the American Society of Mechanical Engineers (ASME). The method was developed for the estimation of the turbine cycle performance in a classified section. The classification is based on feature identification of the correlation status of the main feed-water flow measurements. We also developed predictive algorithms for the corrected main feed-water through a Kernel Regression (KR) model for each classified feature area. The method was compared with estimation using an Artificial Neural Network (ANN). The feature classification and predictive model provided more practical and reliable methods for the corrective thermal performance analysis of a turbine cycle.

**Keywords** : Main Feed-Water, Valid Model, Artificial Neural Network, Kernel Regression, Turbine Cycle.

---

이 논문은 중소벤처기업부 '중소기업연구인력지원사업'의 재원으로 한국산업기술진흥원(KIAT)의 지원을 받아 수행된 연구임.  
(2019년 기업연계형연구개발인력양성사업, 과제번호 : S2755592)

이 논문은 2019년도 호서대학교 연구년 결과물로 제출되었음.

\*Corresponding Author : Seong-Kun Kim(Hoseo Univ.)

email: kimskun@hoseo.edu

Received September 29, 2019

Revised November 7, 2019

Accepted December 6, 2019

Published December 31, 2019

## 1. 서론

발전설비의 실시간 상황에서의 효율적인 에너지관리를 위해서는 수시로 발전시스템의 열평형 계산을 할 수 있어야 하며, 단위기기의 성능계산이 신속하고 정확하게 이루어져야 한다. 이를 위해서는 해당 발전 시스템의 설계 열평형 특성에 기반한 성능 분석 프로그램을 활용할 필요가 있다. 성능 시뮬레이션 프로그램은 성능분석에 필요한 계산 모듈과 단위기기들의 특성 정보를 갖고 있기 때문에 검증 및 진단 시뮬레이션 프로그램을 사용하여 개별기기들의 모델을 보정하고 연계·통합함으로써 개별기기의 성능변화가 전체시스템에 미치는 영향을 현재 상태에 적합하고 정확하게 분석할 수 있다. 성능 분석의 목적은 발전소 운전성능 및 기기성능의 현재 상태를 파악하고 지속적으로 분석하는 것이다. 발전소 운전상태 성능 분석에 필수적인 열평형(HB: Heat Balance) 성능을 계산하기 위하여 열유체 물성치를 계측하고 측정 데이터로부터 보일러(화력) 또는 증기발전기(원자력), 터빈, 습분분리기(원자력), 복수기, 급수가열기 등 발전소 주요 계통의 성능을 계산·분석함으로써 실시간으로 성능분석이 가능하게 하며, 정확성, 편리성, 신속 대응성 등을 제공하는 보다 정확한 성능 분석을 할 수 있다.

열평형 분석은 발전소별로 차이가 있으나 Fig. 1에서와 같이 터빈 사이클 전체에서 압력, 온도, 유량 등의 데이터를 측정하고, 각 트레인 별로 유동 라인에서의 유체 상태량 데이터를 측정한다. 그리고 이 데이터를 이용하여 기기별 성능 계산, 발전 효율 그리고 열 소비율 등을 계산한다.

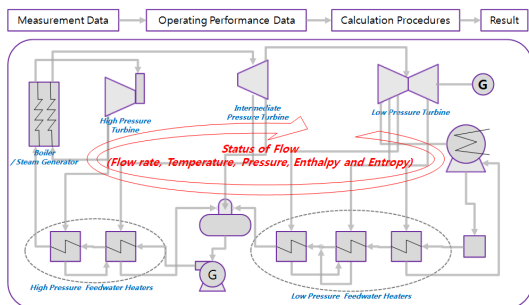


Fig. 1. Cycle of the Power Plant System.

지난 연구[1,2]에서 터빈 사이클 성능 분석은 ASME(American Society of Mechanical Engineers) PTC(Performance Test Code) 6[3], PTC 6A[4],

PTC 12[5,6]를 토대로 터빈 사이클 열평형 계산 및 보정 계산을 하였다. 상업용 열평형 소프트웨어 프로그램인 PEPSE (Performance Evaluation of Power System Efficiency) 모델을 사용하여 영광 3, 4호기 운전 열성능 데이터의 부족분을 열평형 계산을 통하여 구하였으며, 성능 계산 모듈의 입력용 기본 자료로 활용하고 미측정 항목값에 대한 예측 모듈 및 열성능 예측 모듈 개발에 활용하여 열평형 분석을 위한 성능 계산 절차 및 분석 시스템을 개발하였다. 본 연구에서는 이러한 절차를 기반으로 보정 성능을 정확히 결정할 수 있도록, 측정 데이터의 입력 검증 모듈을 개발하고, 열 성능 분석 시스템의 성능을 향상시키고자 하였다. 대표적인 중요 변수인 주급수 유량을 대상으로 영역별 판정 알고리즘을 개발하고 각 영역별로 현재의 터빈 사이클 성능을 추정하는 알고리즘을 개발하였다. 추정 알고리즘은 측정 상태량의 상관관계를 기반으로 형상 분류를 제시하고, 이를 기반으로 커널 회귀모델을 이용하여 추정 모델을 구성하였으며, 커널 회귀모델링의 우수성을 검증하기 위하여 신경 회로망 모델과 비교하였다.

## 2. 모델의 구성

### 2.1 측정 데이터

성능 계산에 필요한 데이터의 약 80~90%가 계측기를 이용하고 있으며, 발전소의 유량계는 주로 차압식 유량계를 사용한다. Table 1은 국내 가동 원자력 발전의 주급수관 유량계 설치 현황이며, 영광 3, 4호기인 경우는 이물질 침적 현상으로 전기출력 감소 사례가 발생한 적이 있다. 일반적인 차압식 유량계는 제작이 비교적 간편하고 대용량 유량 측정일수록 설치비용이 저렴하고 유지보수가 간편하다. 또한 고온·고압에 잘 견디며 내구성이 좋다. 하지만 다른 유량계보다 압력손실이 크고 정밀도가 떨어지며 배관의 형태 또는 유동의 형태에 따라 유출계수 및 정확도에 민감한 영향을 받는다.

Table 1. Installation of Flowmeter in Operating Power Plant

Plant	Flowmeter
YeungGwang 3, 4 Units	Venturi
UJin 1, 2 Units	Orifice / Venturi
WolSeong 2, 3, 4 Units	Flow Nozzle
Gori 1, 2, 3, 4 Units	Venturi

유량과 관련된 출력신호가 선형이 아니고 유량의 제곱에 비례하기 때문에 선형출력 신호를 갖는 유량계와 비교시 유량계의 측정 범위가 약 3:1 ~ 5:1 정도로 제한된다. 요구되는 직관부의 길이가 매우 길며, D/P 셀 매니폴드 등이 필요하고 온도, 압력, 보정 장치가 필요하다. 오리피스 가공단면 등의 마모 및 노후화로 정확도가 떨어지고, 이에 대한 영향이 정량화되지 않았다.

일반적인 유량계의 교체 주기는 약 5년 정도이며, 안정성을 확보해야 하는 발전소인 경우는 주기가 더 짧아진다. 또한 화력 발전소의 경우 1차 증기 유량의 측정 오차가 1%에 이를 때 열속(Heat Flux)은 1% 증가하며 재열 터빈 효율은 1.4% 정도 감소하는데, 교체 주기 동안 거의 모든 센서는 약 2 ~ 3%의 편차를 보이고 있다. 그 결과 발전소 가동 폐널티는 상당하다고 할 수 있다. 측정 데이터에 있어서, 위와 같은 계측기의 일반적인 오차 및 계측기의 오작동, 파이프의 마모 및 이물질 침적현상 등에 의하여 측정량이 정확하지 않다면, 운전성능 계산 결과에 대한 신뢰성은 보장받지 못할 것이다.

터빈 사이클을 구성하고 있는 터빈(고압, 중압, 저압)과 복수기, 급수 가열기, 펌프류의 설계를 위한 수학적 모델과 사이클 설계를 위한 절차는 공급 업체(General Electric의 GER 2007C)에서 제공한 설계 절차로서 이미 알려져 있는 설계 절차에 따른 성능은 인수 시험 때 확인하게 된다. 그러나 이러한 설계 데이터와 조건은 설계 및 인수 성능 시험시에만 의미가 있을 뿐, 인수 후 운전과정이 진행된 현재 운전 상태에서는 부적절한 데이터가 되며, 같은 설계에 의한 사이클이라도 운전 과정에 따라 데이터가 달라진다. 현재의 사이클 상태에 대한 진단과 이를 바탕으로 한 최적화 과정을 수행하기 위해서는 역설계(IDP: Inverse Design Procedure) 모델이 필요하다.

측정 데이터에는 두 가지 그룹이 있게 되는데, 한 그룹은 발전소 정밀 성능 분석을 위해서 특별히 설치한 계측기에서 측정되는 데이터이며, 또 한 그룹은 이미 발전소 운전을 위해 상시 설치된 계측기로부터 측정된 성능 관련 데이터이다. 이러한 데이터에 대해서는 수일에 걸쳐서 많은 분량의 시료 채취가 이루어지며, 이러한 샘플 데이터에 대한 평균화 작업, 필터링 및 순간 측정 오류가 분명한 데이터에 대한 처리가 이루어져서 한 개의 측정값으로 구성된다. 본 연구에서는 이러한 샘플링 및 필터링에 의한 센싱 검증이 아니라, 정상 측정되고, 평균화 및 필터링된 데이터 측정값에 대하여 발전소 전체 사이클 정보를 사용하여, 유효화하고 검증하는 것이다. 그러나

발전소 터빈 사이클 자체가 상호 연결된 사이클 유동 네트워크이기 때문에 이러한 80여개의 측정 데이터 중 어느 하나가 잘못된 수치를 가지고 있는 것인지를 파악하는 것은 쉽지 않은 문제가 된다.

## 2.2 성능 측정 데이터 유효화 검증

하나의 측정 데이터가 특정 값으로 변화하였다면, 사이클 유동 네트워크의 특성상 다른 측정값에서도 변화 흔적이 나타날 것이라고 가정하였으며, 성능 측정 데이터의 유효화를 위해서는 측정 데이터로부터 시스템 반응을 정확히 정의하는 것과 시스템 반응으로부터 시스템 상태를 정확히 파악하는 방법이 있다. 첫 번째 단계에서는 통계 처리 방식에 의한 평균, 표준 편차와 센싱 데이터 분포 처리 및 필터링에 기반을 두고 있으며, 두 번째 단계를 위해서는 발전소 기기에 대한 엔지니어링 모델과 인수 데이터, 설계 데이터 및 측정 데이터를 이용한 구성 방법이다.

현재 발전소에서 입력 검증의 대상이 되는 기기는 동적인 기기가 아니라 의사 정적인(pseudo-static) 기기에 가깝고, 정확한 모델링 샘플 데이터가 제공 될 수 있는 것이 한정되어 있기 때문에 이러한 방법을 전체에 적용하여 시스템을 구성하는 것은 적절하지 못하다고 판단된다. 따라서 보다 실용적인 시스템 구축을 위하여 발전소에서 제공될 수 있는 최소한의 유효 데이터, 즉 정확도를 보장할 수 있는 열평형 데이터를 근거로 유효화 모델을 구성하는 것이 타당하다고 판단된다. 본 연구에서는 영광 원자력 3,4호기의 데이터를 기반으로 학습 모델을 구성하였다.

## 2.3 상관 인자 선정

상관 인자의 선정은 최적의 추정 모델을 설정하기 위한 목적에 있으며, 상관 인자의 적절하지 못한 선택은 다음과 같은 문제점을 야기한다.

- 1) 인자의 수가 많아지면 계산이 복잡해져서 학습 및 인식 속도가 느려진다.
- 2) 불필요한 인자는 학습을 어렵게 만든다.
- 3) 부적합한 인자의 추가로 인하여 모델의 정밀도가 약화될 수 있다.

이렇듯 분별력 있는 데이터 분석이 되기 위해서는 일반적으로 복잡한 데이터를 단순한 데이터로 정제시켜야

한다. 즉, 추정 모델의 학습 단계 이전의 전처리 과정으로서, 원형 데이터에 대한 적절한 가지치기 작업을 해야 한다. 또한 정제시킨 데이터라 할지라도 앞에서 언급했던 문제점이 야기되므로, 적절한 입력 인자 및 인자 수에 대한 선정이 진행되어야 한다.

데이터 축소란 대량의 변수 군을 소량의 변수 군으로 병합하는 과정을 의미하는 데, 일반적으로 많이 사용하는 인자 축소 방식인 주성분 분석은 인자들의 상관 분석을 통하여 고유 벡터를 도출하고 큰 고유 벡터를 사상시킴으로서 축소된 인자를 추출하는 방식이다. 하지만 원용 데이터에서 모든 데이터를 사용할 경우 부적합한 인자도 적용되기 때문에 모델의 정밀도는 약화될 수 있다.

상관 인자의 선정은 크게 (1) 모델을 기반으로 하는 방식(Model-based Method)과 (2) 모델을 기반으로 하지 않는 방식(Model-free Method)으로 나눌 수 있다. 모델 기반 방식은 선택된 인자를 이용하여 학습을 통한 최적의 설정 변수를 찾는 방식으로서, 인자를 변경시키면서 반복 과정에 의하여 최적의 인자를 선택하는 과정이다. 그리고 모델을 기반으로 하지 않는 방식은 학습 데이터의 인자와 이에 대한 출력의 통계적 관계를 바탕으로 수행하는 방식이다.

본 연구에서는 상관 인자와 출력의 강도를 알아보기 위하여 피어슨 상관 계수를 사용하여 변수간의 영향력을 파악하였다.

$$r_{ij} = \frac{s_{ij}^2}{\sqrt{s_{ii}s_{jj}}} \quad (1)$$

$$\text{where } s_{ij}^2 = \frac{\sum(y_i - \bar{y}_i)(y_j - \bar{y}_j)}{N-1}$$

데이터 기반 모델에서의 학습 단계란 상관 인자로 구성된 학습 데이터로부터 객체의 특징(feature)을 정의하는 과정으로서 학습을 통하여 속이 빈 벡터를 완전한 특징 벡터로 만드는 단계이며, 이러한 벡터는 학습 데이터의 반복 계산에 의하여 이루어지기 때문에 학습 데이터의 영향력이 크다. 이렇게 완성된 특징 벡터를 이용하여 예측값을 추정하게 되며, 추정값이 실제 값과 흡사할 때 일반화 능력이 우수하다고 한다. 데이터 기반 모델에 대한 일반화 능력은 특징 벡터에 의하여 핸들링 되며, 학습 데이터에 의존하게 된다. 즉 추정 값에 대한 상관 인자의 데이터들이 학습 데이터의 특징과 비슷한 경향이 있어야 실제 값과의 오차를 줄일 수 있다고 판단된다[7].

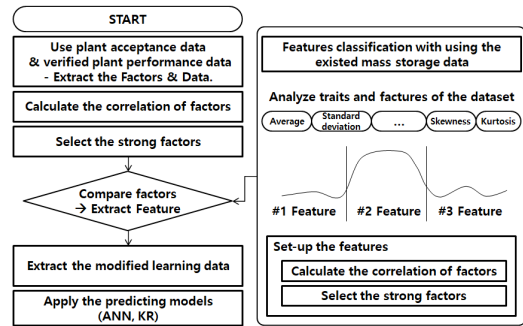


Fig. 2. the Procedures of the whole Algorithm.

본 연구에서는 추정값에 대한 추세를 알아보고, 이에 대한 강인한 상관 인자로 조합된 학습 데이터를 구성하고자 한다. 추정 값에 대한 경향은 기존에 수집된 데이터를 이용하여 추정 항목에 대한 특징 영역을 분류하고, 추정 값의 우수한 상관 인자와 특징 영역에서의 강인한 상관 인자를 비교하여, 추정 값에 대한 특징 영역, 즉 형상을 파악하는 것이다(Fig. 2).

본 연구에서는 기존의 연구[8,9]에서 제시되었던 상관 관계를 기반으로 구성하였다. 강인한 상관인자란 Table

Table 2. Classification Table of Strong Factors

Factor	Feature				
	#1	#2	#3	#4	#5
ZHPTBNP	1M	1L	0	0	1L
ZMSP	0	0	0	0	0
ZMSF	1L	1H	1L	0	0
ZFWHT	1M	1L	0	0	1L
ZLTUBT	0	0	0	0	0
ZHTUBT	0	0	0	0	0
ZBSCALDT	0	1M	0	0	0
ZBDELTDT	0	1L	0	0	0
ZBTFSPDT	0	0	0	0	0
ZSFDP	0	0	0	1L	0
ZRCPDP	1L	0	0	0	0
ZFWHSGP	1L	1M	0	0	0
ZBSBD	0	0	0	0	0
ZBSBT	0	0	0	0	0
ZGENPO	1L	0	0	0	0
ZSTMf1	0	1H	1L	0	0
ZSTMf2	1L	1M	0	0	0
ZPNFW1	1L	1H	1H	1H	0
ZPNFW2	0	1H	1H	1H	0
ZDELt	1L	1M	0	0	0
ZMSROPA	1M	0	0	0	1L
ZMSROPB	1M	1L	0	0	1L
ZMSRTA	0	0	0	0	0
ZMSRTB	0	0	0	0	0
ZLPTP	0	0	1L	0	0

Note)  $-0.5 \leq r < 0.5$  : none factors(0),  $-0.7 \leq r < -0.5$  or  $0.5 \leq r < 0.7$  : Low strong factor(1L),  $-0.8 \leq r < -0.7$  or  $0.7 \leq r < 0.8$  : Middle Strong factor(1M),  $-1.0 \leq r < -0.8$  or  $0.8 \leq r < 1$  : High strong factor(1H).

2처럼 기존의 빅데이터를 이용하여 미리 추출한 특징 영역[10]의 상관인자들이다. 이러한 상관인자들은 형상별 인자 및 강인성을 갖고 있는 인자 수가 다르다.

상관 계수가 0.7 이상인 경우, 강인성을 갖고 있는 인자수가 40개가 넘는 경우도 발생하며, 적을 때는 10개 미만인 경우도 나타난다. 상관 인자 수가 많은 경우는 가동 시간이 길어지는 단점이 있어서, 신경회로망 모델에서는 입력 인자 수를 정해서 진행하는 연구가 많았다.

본 연구에서는 학습 데이터를 이용하여 모든 인자의 상관 계수를 계산하고, 인자의 강인성(0, 1)을 비교하여 형상을 구별하며, 인자의 순위별 그룹(low, middle, high)을 정하여, 해당 그룹까지 포함되는 인자를 입력 인자로 설정한다. 예를 들어 학습 데이터의 모든 인자별 상관 계수를 계산하여, 강인성을 부여하고 Table 2와 비교하였을 때, #5 형상을 판별되었고, 임의의 인자 수를 3으로 하였을 때 순위에 포함되는 그룹은 Low 그룹이다. 그런 경우에는 Low 그룹까지 포함하는 인자는 모두 입력 인자로 설정하는 것이다.

### 3. 커널 회귀 모델링 결과

#### 3.1 커널 회귀 알고리즘

어떤 인자에 대한 난수의 데이터 집합에 대하여 최대한 가장 정확히 추정할 수 있는 회귀 함수  $f(x_1, x_2, \dots)$  를 생성하고자 할 때, 신뢰할 수 있는 추정법으로는 모수 추정(parameters estimation) 방법과 비모수 추정 방법으로 대별될 수 있다. 모수 추정법에서는 분포함수가 미지의 모수를 갖는 고정된 함수의 형태를 따른다고 가정하는 반면, 비모수적 추정 법에서는 특정한 분포를 제한하지 않는다. 따라서 모수 추정법에서는 모델의 선정이 주요 사항으로 다루어진다.

만약 잘못된 모델을 선정할 경우 모수의 정확한 추정으로는 보상하기 힘든 모델의 편기가 발생할 수 있기 때문이다. 하지만 비모수 추정 방법은 일정한 분포를 제한하지 않고 주어진 자료에서 직접 함수를 추정하게 된다. 이러한 모수추정 방법으로는 선형회귀분석, 다항회귀 분석등이 있으며, 비모수 추정 방법으로는 커널 밀도 추정, k-NNR 추정, 비모수 베이지안 모델, 커널 회귀[11] 등이 있다.

커널 회귀 방법은 입력 층에 대한 추정 값을 알고자 할 때, 표본 데이터와 커널 함수의 적절한 곱으로 표현할

수 있는데, 커널 함수는 표본 데이터와 추정값에 대한 입력 층의 공간적 거리에 대한 가중치를 부여하는 함수이며, 거리 함수로는 유클리디안 거리 함수를 이용한다. 거리 함수가 작을수록 오차가 작다는 것을 의미하는데, 이러한 데이터에 대한 값을 상대적으로 더 크게 부여하기 위하여 커널 함수를 적용한다. 이러한 커널 함수에는 역거리(Inverse Distance Weighting) 함수, 가우시안 커널(Gaussian Kernel) 함수 등이 있으며, 커널 함수의 가중치를 부여하는 것이 커널 폭  $h$ 이다.

보편적으로 최적의 커널 폭을 찾기 위하여, 초기 설정 이후 일정량씩 증가시키면서 실제 값 대비 추정 값에 대한 오차를 계산하며, 기존 오차와 비교한다. 기존 오차보다 클 경우, 기존 오차가 최소 오차라고 판단하여, 이에 대한 최적의 변수를 설정하는데, 이는 학습 시간을 줄이기 위한 장점도 있지만, 전역 최적점이 아닌 극소 최소점으로 수렴될 요지가 있다. 따라서 본 연구에서는 적용 대상인 주급수 유량의 정상 상태에서의 일부 윈도우(학습 데이터 60주기, 검증 데이터 10주기)를 추출하고, 커널 함수 중에서 보편적으로 사용하는 가우시안 커널 함수를 이용하여 커널 폭의 변화에 따른 오차율을 알아보았다.

모든 경우에서, 일정 반복 회수 이후부터는 오차율의 변화가 거의 없는 것을 알 수 있었다.

#### 3.2 커널 회귀 알고리즘 절차

커널 회귀 알고리즘은 표본 데이터를 갖고 있는 상태에서 근사치 함수를 구하는 방법으로서, 주어진 학습 데이터를 표본 데이터와 하위 학습 데이터로 세분화하고, 하위 학습 데이터의 최소 오차를 갖는 커널 폭을 결정하여, 추정 값을 예측한다.

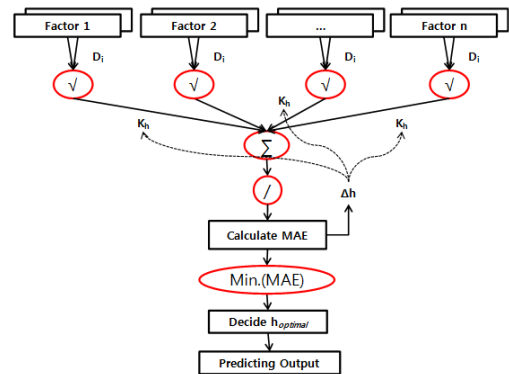


Fig. 3 Kernel Regression Algorithms.

- 1) 표본데이터의 입력 층과 하위 학습 데이터의 입력 층을 이용하여 유클리드 거리를 계산한다.
- 2) 커널 폭을 증가시키면서 커널 함수를 이용하여 가중치를 구한다.
- 3) 커널 함수의 가중치 및 가중치와 하위 학습 데이터의 입력층 곱의 합을 이용하여 하위 학습 데이터의 추정 값을 구하고, 하위 학습 데이터의 출력층과 비교한다.
- 4) 하위 학습 데이터의 각각에 대한 출력의 오차 (MSE)를 계산한다. (2)에서 (4)까지를 적절한 반복 횟수까지 반복하여 오차가 가장 작은 커널 폭 h를 찾는다.
- 5) 커널 폭을 갖고, 표본 데이터의 입력층 및 출력층과 하위 학습 데이터의 입력층 대신 검증 데이터의 입력층을 가지고, 유클리드 거리, 커널 함수를 통과하여 추정 값을 예측한다.

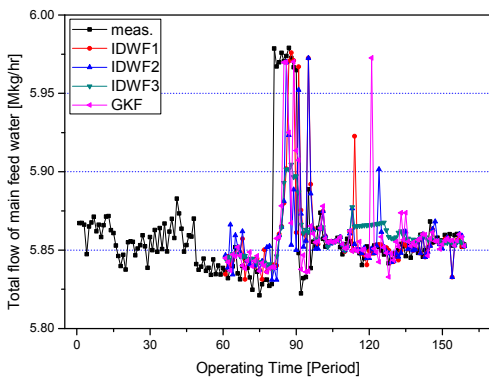


Fig. 4. Calculating Result with Kernel Regressions.

Fig. 4는 측정 데이터 대비 역거리함수(IDWF #1 ~ IDFW#3) 및 가우시안 커널 (GKF) 추정값을 보여준 것이며, Table 3.은 커널 함수별 형상에 따른 오차율을 나타낸 것이다. IDWF #1 커널 함수를 살펴보면, #1 형상에서는 1% 이상의 오차를 갖는 추정 값이 있기는 허나, 모두 2% 미만이며, #2 형상에서 여러 개의 추정 값이 큰 오차를 나타나는 것을 알 수 있으나, #5 형상에서 1% 미만의 정확한 추정값을 제시하는 것이 많아서 평균 오차가 줄어드는 것으로 나타났다. IDWF #2 커널 함수를 살펴보면 IDWF #1과 마찬가지로 #2 형상에서는 비슷하나 전체적으로 IDWF #1 오차율보다 크게 나타났다.

Table 3. Comparison with Errors to Kernel Regression (IDWF #1-#3, GKF)

No.	Feature	Error [%]			
		IDWF #1	IDWF #2	IDWF #3	GKF
1	#1	1.01	1.21	1.01	0.94
2	#2	6.05	6.90	5.94	4.56
3	#3	0.82	0.59	0.94	0.51
4	#4	0.35	0.74	1.18	1.30
5	#5	0.55	0.69	0.47	0.49
6	Total	1.72	1.97	1.87	1.53

그리고 IDWF #3 은 IDWF #1, #2와 마찬가지로, 오차율이 큰 값이 많지 않고, 많은 값들이 1% ~ 2% 사이의 오차를 보이고 있다. 마지막으로 GKF 커널함수는 역거리 함수 (#1 ~ #3) 과 비슷한 경향을 보이고 있으나 추정값의 오차가 작은 것으로 나타났으며, 특히 정상 상태의 학습 데이터를 이용하여, 정상 상태의 검증 데이터를 검증할 때가 오차율이 작았다.

Table 4에서처럼 신경 회로망보다 커널 회귀 함수의 추정 값이 신뢰할 수 있었다. 특히 #3 형상으로 진입하는 구간 즉 강제 보정 데이터가 학습 데이터로 인식하여 추정값에 반영되는 구간에서 커널 회귀 함수가 빠르게 오차를 줄여서 추정값을 도출하는 것을 알 수 있었다.

Table 4. Comparison with Errors to ANN, KR (Sigmoid, Tanh, IDWF #1-#3, GKF)

Feature	Error [%]					
	ANN		KR			
	Sigmoid	Tanh	IDWF#1	IDWF#2	IDWF#3	GKF
#1	1.42	5.73	0.22	0.99	0.78	0.80
#1	0.75	5.93	0.53	0.46	0.52	0.46
#1	2.40	1.42	1.64	1.64	2.14	1.55
#2	13.10	13.81	13.80	14.01	13.88	13.99
#2	10.02	0.79	10.36	11.49	6.76	5.13
#2	11.80	13.45	2.33	2.31	3.03	0.34
#3	1.92	3.39	0.68	0.69	0.51	0.52
#3	0.39	0.02	0.05	0.03	0.02	0.06
#3	1.37	0.95	0.44	0.48	0.52	0.55
#4	1.65	1.58	0.36	1.18	1.75	0.87
#4	2.29	2.79	0.85	0.64	1.70	0.87
#4	1.18	2.67	0.23	0.93	0.69	0.42
#5	1.78	0.32	0.29	0.28	0.51	0.51
#5	0.13	0.78	0.85	0.85	0.67	0.64
#5	0.14	1.81	0.22	0.22	0.10	0.16
Avg.	3.36	3.70	2.19	2.41	2.24	1.79

#### 4. 결론

열 성능 분석 시스템이란 터빈 사이클의 열 성능을 분석하고 성능 변화 추이 뿐 아니라 성능저하 기기를 파악하고 이를 정량적으로 평가할 수 있는 시스템을 말한다. 현장 실무자가 터빈 사이클의 성능을 분석하는 데에는 많은 성능 변수가 필요하고 전체 사이클의 성능에서 기기별 차지하는 성능을 분석하기 위해서는 여러 개의 기기들이 터빈 출력에 영향을 미치므로 관련 기기의 열평형을 반복적으로 계산해야 하는데, 이러한 분석 시스템은 ASME PTC 6 코드의 계산 방법 및 터빈 제작사의 설계 성능 계산 지침서를 참조하여 계산되어 진다.

본 연구에서는 이러한 측정 데이터의 정확성을 얻기 위하여 유효화 검증을 하였으며, 신뢰성 있는 측정 데이터를 얻을 수 있었다.

- 1) 열평형 분석을 하기 위한 운전 성능 데이터 중에서 측정 데이터와 상관성이 있다고 판단되는 성능 인자를 측정 데이터 및 성능 지침서의 엔지니어링 정보를 활용하여 인자 추출 및 데이터 쌍을 구성하였으며, 상관관계를 통하여 우수한 상관 인자 및 데이터 쌍을 추출하였다.
- 2) 기존의 빅데이터에서 데이터 집단의 특징을 활용하여 형상을 분리할 수 있었으며 각각의 형상별 강인한 상관 인자를 도출시켜 형상에 맞는 강인한 상관 인자를 활용하여 윈도우를 설정하였다. 또한 적절한 윈도우 수는 각각의 인자별 데이터 쌍에 대한 상관 계수를 통하여 파악할 수 있었으며, 이렇게 추출된 윈도우를 활용하여, 추정 모델에 맞는 학습 데이터로 재구성할 수 있었다. 인자의 순위별 그룹(Low, Middle, High)을 정하여, 해당 그룹까지 포함되는 인자를 입력 인자로 설정하였다.
- 3) 커널 회귀 모델에서는 주로 사용하는 커널 폭의 설정 방법에 대한 모순을 파악하여, 커널 폭의 증가 값과 구간을 설정하였으며 이를 통하여 최적의 커널 폭을 설정할 수 있었다. 또한 커널 폭을 제어할 수 있는 커널 함수인 역거리 함수와 가우시안 커널 함수의 비교·분석을 통하여 적절한 커널 함수를 활용할 수 있었다.

본 연구에서 학습 모델로 사용된 원자력 발전 성능 데이터를 화력 발전소 실시간 학습 데이터에 적용하는 것은 차후 연구 과제로 구성하고자 한다.

#### References

- [1] S. K. Kim, K. H. Choi, "Development of Performance Analysis System(NOPAS) for Turbine Cycle of Nuclear Power Plant", *Journal of Korean Nuclear Society*, Vol. 33, No. 1, pp.211-218, 2001.
- [2] S. K. Kim, K. H. Choi, "Development of Performance Analysis Methodology for Nuclear Power Plant Turbine Cycle using Validation Model of Performance Measurement", *Journal of KSME*, Vol. 24, No. 12, pp.1625-1634, 2001.  
DOI: <http://dx.doi.org/10.22634/KSME-B.2000.24.12.1625>
- [3] American Society of Mechanical Engineers, Performance Test Code 6, "Steam Turbines", 1976.
- [4] American Society of Mechanical Engineers, Performance Test Code 6A, "Appendix A to Test Code for Steam Turbine", 1982.
- [5] American Society of Mechanical Engineers, Performance Test Code 12.1, "Closed Feed Water Heaters", 1978.
- [6] American Society of Mechanical Engineers, Performance Test Code 12.2, "Steam Condensing Apparatus", 1983.
- [7] T. Fernando, H. Maier, G. Dnady, "Selection of input variables for data driven models: An average shifted histogram partial mutual information estimator approach", *Journal of Hydrology*, Vol. 367, pp.165-176, 2009.  
DOI: <https://doi.org/10.1016/j.jhydrol.2008.10.019>
- [8] H. J. Yang, S. K. Kim, K. H. Choi, "A Study of the Feature Classification and the Predictive Model of Main Feed-Water Flow for Turbine Cycle", *Journal of Energy Engineering*, Vol. 23, No. 4, pp.263-271, 2014.  
DOI: <https://doi.org/10.5855/ENERGY.2014.23.4.263>
- [9] H. J. Yang, S. K. Kim, "A Predictive Model of the Generator Output Based on the Learning of Performance Data in Power Plant", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol. 16, No. 12, pp.8753-8759, 2015.  
DOI: <https://doi.org/10.5762/KAIS.2015.16.12.8753>
- [10] H. J. Yang, *A study of the feature classification and the predictive algorithms of main feed-water for turbine cycle*, Ph.D Dissertation, Hoseo University, Asan, Korea, pp.85-88, 2011.
- [11] N. Cristianini, J. Shawe-Taylor, "Introduction to Support Vector Machines and other Kernel-based Learning Method", Cambridge University Press, 2000.

양 학 진(Hac-Jin Yang)

[정회원]



- 2002년 8월 : 호서대학교 기계공학과(석사)
- 2012년 8월 : 호서대학교 기계공학과(박사)
- 2014년 3월 ~ 2018년 2월 : 동양미래대학교 로봇자동화공학부 조교수
- 2019년 3월 ~ 현재 : 제주관광대학교 메카트로닉스과 조교수

<관심분야>

플랜트 성능 해석 및 예측, 장비 설계(CAD)

---

김 성 근(Seong-Kun Kim)

[정회원]



- 1987년 3월 : 서울대학교 기계설계학과(석사)
- 1992년 8월 : 서울대학교 기계설계학과(박사)
- 1995년 3월 ~ 현재 : 호서대학교 기계자동차공학부 교수

<관심분야>

플랜트 성능 해석 및 진단, 설계/가공