

기계학습 기법을 이용한 실선 추진성능의 확률 모델 개발

유병현* · 김진환*
*한국과학기술원 기계공학과

Probabilistic modeling of powering performance of a full-scale ship using machine learning techniques

Byunghyun Yoo*, Jinwhan Kim*
*Department of mechanical engineering, KAIST, Korea

KEY WORDS: Ship powering performance 선박 추진성능, Machine learning 기계 학습, Regression model 회귀 모델, Physical knowledge 물리적 지식

ABSTRACT: With the increasing interest in reducing environmental pollution and global warming, regulations for fuel efficiency have been strengthened and much research has been conducted to improve fuel efficiency. For this, the powering performance of a full-scale ship under real-sea weather and operating conditions needs to be accurately estimated. However, existing empirical and statistical methods do not provide satisfactory estimation performance of full-scale ships under operational uncertainties due to environmental disturbances. In this study, machine-learning techniques are applied to model the ship's powering performance using on-board measurement and weather forecast data. To demonstrate the validity of the proposed method, the proposed model is compared with ISO15016 and the results are discussed.

1. 서 론

최근 환경오염에 대한 심각성이 증대되면서 이에 대한 환경 규제 방안들이 강화되고 있다. 국제해사기구(IMO)는 에너지 효율설계지표(EEDI)와 선박에너지효율관리계획서(SEEMP)와 같은 환경 규제 관련 기준을 제시하고 이를 준수하기를 권고하고 있고, 이에 따라 이산화탄소 배출량과 밀접한 관련이 있는 연료 소비량을 줄이기 위한 연구의 필요성이 증대되고 있다.

연료 소비량을 줄이기 위한 다양한 방안 중 운항 효율을 향상 시킴으로서 전반적인 연비 성능을 높이는 방법에 대한 연구의 중요도가 증가되고 있으며, 이러한 방법을 적용하기 위해서는 일차적으로 선박 운항 효율에 대한 정확한 추정 및 예측이 필요하다. 그러나 선박의 운항 효율에는 다양한 인자들이 영향을 미치고, 현재까지의 해석적인 방법이나 경험식 기반의 방법들로는 정확한 결과를 얻어내는 것이 쉽지 않다. 수조 시험을 통한 추정 기법도 시험 상에서의 조건과 실제 해상 환경에서의 운용 조건이 상당히 다르다는 한계점을 가진다.

이러한 한계점을 극복하기 위해서 기존 방법과는 달리, 실제 선박에서 계측된 데이터를 기반으로 통계적 모델을 구성하는 연구들이 점차 수행되고 있다. 엔진의 유효 출력과 연료 소비율의 데이터에 기반을 둔 선박의 운항 효율 평가에 관련된 연구(Borkowski, et al., 2011)가 진행된 바 있으며, 실선 운항 기록을 활용하여 선박의 속도 변화를 추정한 연구(Sogihara et al., 2010, Sasa et al., 2015) 또한 수행된 바 있다.

또한, 최근 성공적인 연구 결과를 통해 각광받고 있는 기계학습 기법 중 인공신경망(Artificial Neural Network) 기법과 가우시안 프로세스(Gaussian process)기법을 이용해 연료 효율을 추정한 연구가 진행된 바 있다(Petersen., et al., 2012, 유병현 등, 2013).

본 연구에서는 실선 운항 기록을 이용하여 가우시안 프로세스 기법을 통한 회귀 분석을 통해 운항 성능의 대표 인자인 선속과 마력을 추정하되, 계측 물리량들의 역학적 관계를 적용한 운항 성능 모델을 구성함으로써, 전반적인 회귀 모델의 구성 방식 및 배경 지식의 적절한 활용 방안에 대한 일련의 절차를 제시하고자 한다.

2. 실선운항기록 분석

2.1 실선운항기록 개요 및 전처리

본 연구에서는 컨테이너선의 실해역 운항 기록을 사용하였다. 실제 운항 상태에서의 환경 조건과 관련된 기록은 계측 데이터와 예보 데이터를 활용하였다. 데이터 내의 잡음 제거와 시간 동기화를 위해 전체 데이터의 5분 평균값을 계산해 사용하였다. 또한, 본 연구에서는 정상상태 구간에서의 추진성능 파악을 목적으로 하고 있기 때문에, 5분 평균 데이터 중 정상상태 구간의 데이터를 추출하여 사용하였다. 정상상태 구간은 RPM과 타각이 일정 시간 동안 유지되는 구간을 의미한다.

2.2 독립변수 선정

과도한 수의 독립변수는 과적합(Over-fitting)의 주요한 원인이 될 수 있으므로, 필수적인 독립변수를 선정하는 과정을 거쳤다. RPM과 초기 트림 및 홀수는 선속과 마력 추정 시 주요한 변수로 알려져 있기 때문에 우선적으로 독립 변수로 선정하였다. 환경 외란 항 중에는 파도, 바람, 조류, 수온 등의 여러 항 중에 비교적 더 중요하다고 판단되는 파도와 바람을 독립 변수로 선정하였다. 파도와 바람의 경우 크기와 방향으로 주어진 데이터를, 효과적인 모델링을 위해 정면 방향과 측면 방향의 요소로 변환하여 사용하였다. 정면 방향의 파고(Head-wave, head-swell)와 정면 방향의 바람(Headwind)은 선박의 저항에 직접적인 영향을 주기 때문에 독립변수로 선정하였다. 그러나 측면 방향의 환경 외란은 선박의 저항에 직접적인 영향을 주기 보다는 선박의 경로에 영향을 주게 되고 그로 인해 타각의 변화를 가져오게 된다. 불필요한 독립 변수의 수를 줄이기 위해 측면 방향의 환경 외란(Cross-wave, cross-swell, crosswind)대신에 타각을 독립변수에 포함시켰다.

3. 회귀 분석 모델

3.1 가우시안 프로세스 회귀 분석

가우시안 프로세스(Gaussian process)는 유한개의 확률 변수를 추출하였을 때 그 변수들이 다중 정규 분포를 따르는 확률 과정을 의미한다. 가우시안 프로세스를 회귀 분석에 적용하는 경우 계산하고자 하는 모델의 함수 형태가 가우시안 프로세스를 따른다고 가정하며, 일반적인 회귀 분석 시와 마찬가지로 입력과 출력 데이터를 이용하여 함수를 계산한다. 가우시안 프로세스에서는 데이터를 기반으로 구하고자 하는 함수의 평균과 분산을 계산하고, 이 때 함수의 평균과 분산은 커널 함수(Kernel function)에 의해서 정의된다. 학습 데이터에서의 입력 데이터를 X , 출력 데이터를 f , 검증 데이터에서의 입력 데이터를 X_* , 출력 데이터를 f_* , 커널 함수를 K 라고 정의하면, f 와 f_* 는 식 (1)과 같이 가정된다.

$$\begin{bmatrix} f \\ f_* \end{bmatrix} = \begin{bmatrix} 0, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix} \end{bmatrix} \quad (1)$$

확률론적인 관점에서 가우시안 프로세스를 통해 결과적으로 계산하게 되는 분포는 $f_*|X_*, X, f$ 이며, 자세한 식은 식 (2)과 같이 표현된다.

$$P(f_*|X_*, X, f) \sim N(K(X_*, X)K(X, X)^{-1}f, K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*)) \quad (2)$$

커널 함수로는 다양한 형태의 함수가 사용될 수 있으나, 본 연구에서는 가장 일반적으로 사용 되는 squared exponential 형태의 함수를 사용하였다. 식 (3)은 squared exponential 커널 함수를 나타낸다.

$$k(x, x') = \sigma_{f_0}^2 \exp\left(-\sum_{i=1}^d \frac{(x_i - x'_i)^2}{2\sigma_{f_i}^2}\right) \quad (3)$$

$\sigma_{f_i}^2, i = 0, \dots, d$ 는 모델의 파라미터이며, 데이터에 의해 이 파라미터들은 최적화되고, 최적화 된 파라미터가 결정되면 모델의 형태가 결정된다.

3.2 선속 모델

선속(Speed through water; STW) 모델의 경우 가우시안 프로세스 회귀 분석 모델을 통해 구성되었다. 입력으로 사용되는 독립변수는 RPM, 초기 트림, 홀수, 정면 방향의 파고와 풍속, 타각이며 출력 변수는 선박의 속도이다. 일반적으로 가우시안 프로세스 회귀 분석에서는 주변우도(Marginal likelihood)를 최대화하는 파라미터 θ 를 쥘레기울기법(Conjugate gradient method)과 같이 반복적으로 찾는 형태로 최적화를 수행한다. 그러나 선박의 추진성능 추정에서는, 다른 독립 변수에 비해 RPM의 영향이 지배적이기 때문에 RPM에 대한 과적합이 발생할 가능성이 크다. 본 연구에서는 이러한 문제점을 막기 위해 단순히 우도를 최대화하는 것뿐만 아니라 물리적 지식을 반영하는 정규화(Regularization) 항을 목적 함수에 추가하여 파라미터를 계산한다. 본 연구에서 적용된 물리적 지식은 RPM과 선속이 비례한다는 배경 지식이며, 수학적으로는 RPM에 대한 선속의 2차 도함수의 적분 값으로 표현된다. 식 (4)는 선속 모델의 목적 함수를 나타낸다.

$$J = -L + \lambda S, L = \log P(y|X), S = \int_0^{n_f} \frac{d^2 V}{dn^2} dn \quad (4)$$

J 는 목적 함수, L 은 로그주변우도, S 는 정규화 항, λ 는 정규화 항의 영향의 정도를 정하는 파라미터, X 는 입력 데이터, y 는 출력 데이터를 의미한다.

3.3 마력 모델

마력 모델의 경우도 원칙적으로 가우시안 프로세스 회귀 분석을 사용할 수 있으나, 식 (5)와 같이 물리적 법칙에 기반한 마력 계산식이 알려져 있으므로 이 식을 이용한다.

$$P_B = Q_B n_B = \frac{1}{\eta_s} Q_p n_B = \frac{1}{\eta_s} (K_Q \rho D^5 n_B^2) n_B \quad (5)$$

P_B 는 엔진 마력, Q_B 는 엔진 토크, n_B 는 엔진 RPM, η_s 는 축 효율, Q_p 는 프로펠러 토크, K_Q 는 토크 계수, ρ 는 물의 밀도, D 는 프로펠러의 직경을 나타낸다. 토크 계수의 경우 전진비(J)에 대한 함수로 알려져 있고, 본 연구에서는 토크 계수를 전진비에 대한 1차 함수로 가정하였다. 전진비의 식은 아래 식 (6)과 같다.

$$J = \frac{V(1 - w_f)}{nD} \quad (6)$$

V 는 선박의 속도, w_f 는 반류 계수를 의미한다. 반류 계수, 프

로펠러 직경, 물의 밀도를 상수라 가정하면 식 (5)와 (6)을 통해 마력 계산식인 식 (7)을 도출할 수 있다.

$$\begin{aligned}
 P_B &= \frac{1}{\eta_s} K_Q \rho D^5 n_B^3 = \frac{1}{\eta_s} (K_{Q1} J + K_{Q2}) \rho D^5 n_B^3 \\
 &= \frac{1}{\eta_s} (K_{Q1} \frac{V(1-w_f)}{nD} + K_{Q2}) \rho D^5 n_B^3 \\
 &= (C_{Q1} \frac{V}{n} + C_{Q2}) \rho D^5 n_B^3 \\
 C_{Q1} &= \frac{K_{Q1}(1-w_f)}{\eta_s D}, C_{Q2} = \frac{K_{Q2}}{\eta_s}
 \end{aligned} \tag{7}$$

식 (7)의 C_{Q1} 과 C_{Q2} 외의 항들은 데이터 상에 모두 주어진 값이기 때문에 미지수인 C_{Q1} 과 C_{Q2} 를 통계적으로 추정할 수 있다.

4. 분석 결과

모델의 성능은 평균 제곱근 오차(Root mean square error; RMSE) 및 표준화 된 평균 제곱근 오차(Normalized root mean square error; NRMSE)를 통해 나타낸다. 식 (8)은 평균 제곱근 오차와 표준화 된 평균 제곱근 오차의 식을 나타낸다.

$$\begin{aligned}
 RMSE &= \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \\
 NRMSE(\%) &= \frac{RMSE}{\bar{y}} \times 100
 \end{aligned} \tag{8}$$

4.1 선속 모델

선속 모델의 경우 정규화 항을 적용하기 전, 후에 따른 성능과 단순 다항식 기반 모델의 성능을 비교하였다.

Table 1의 결과와 같이, 학습 단계에서는 정규화 항을 적용하기 전의 결과가 더 정확하였으나, 검증 단계에서는 정규화 항을 적용한 모델이 더 정확한 결과를 얻었다. 이러한 결과는 단순한 파라미터의 최적화가 과적합 현상을 야기할 수 있다는 것을 나타낸다. 과적합 현상을 더 자세히 살펴보기 위해 단일 변수에 따른 선속 모델의 결과 값의 변화를 살펴보았다.

실선은 선속 모델의 함수 값, 음영으로 표시된 영역은 가우시안 프로세스 모델에서 표현되는 함수의 불확실성을 의미한다. Fig. 1, 2를 통해 알 수 있듯이 배경 지식을 활용한 정규화 항이 없이 단순히 파라미터를 최적화하면 RPM에 대한 과적합 현상이 발생할 수 있다. 정규화 항은 파라미터 최적화 과정에서 이러한 과적합 현상을 막고, 위배되는 안 될 배경 지식의 영향을 유지할 수 있는 역할을 하게 된다.

4.2 마력 모델

마력 모델의 경우, 기존에 알려진 식을 기반으로 구성되었고 식의 계수만을 실제 선박에서 계측된 데이터로 추정하였다. 이러한 모델의 유용성을 검증하기 위해 모형 시험 데이터만으로 마력을 추정한 모델과 비교를 수행하였다.

Table 1 Error comparison in the model for STW

	RMSE(kn)		NRMSE(%)	
	Training	Test	Training	Test
w/o regularization	0.195	0.504	1.331	3.429
w/ regularization	0.290	0.319	1.973	2.171
polynomial	0.621	0.628	4.225	4.273

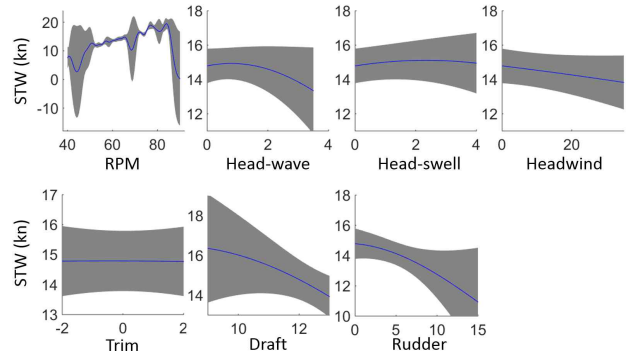


Fig. 1 Results of the GP model without regularization

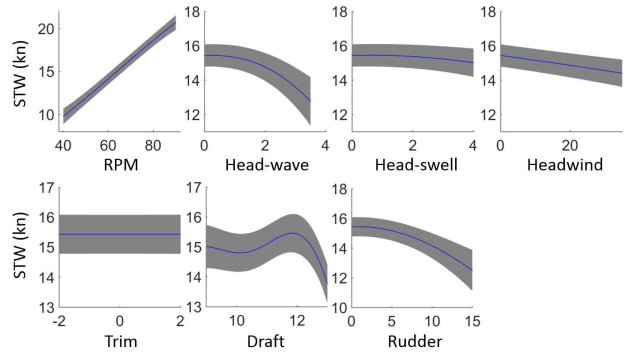


Fig. 2 Results of the GP model with regularization

Table 2 Error comparison in the model for engine power

	RMSE(kW)		NRMSE(%)	
	real-sea data	model test data	real-sea data	model test data
Training set	395.8	441.6	2.57	2.87
Test set	509.4	571.1	3.58	4.01

본 연구에서 사용된 데이터의 경우 실제 선박에서 계측되거나 예보된 데이터이지만, 실제 해상 환경에서 운항 중에 얻어진 데이터이므로 그만큼 잡음의 크기가 크다. 그러나 그러한 계측 잡음 및 환경 외란에 의한 불확실성에도 불구하고 모형 시험 데이터와 비교할 때 비교적 정확한 결과를 보임을 확인할 수 있었다. 모형 시험과 실선 운항 조건 사이에 축적 효과(scale effect)로 인한 추정 오차가 있을 수 있음을 감안할 때 운항기록을 바탕으로 추정이 실제에 더 가까운 것은 자연스런 결과라 하겠다.

나 본 연구에서 제안한 실선 데이터 기반의 마력 모델이 실제 선박의 성능 추정에 효과적 일수 있음을 보여줄 수 있는 근거가 된다.

4.3 ISO15016과의 비교

본 연구에서 제안하는 모델을 사용하여 환경 조건을 고려한 마력과 속력 추정 오차를 보정한 결과와 기존 속력 시운전 해석 지침인 ISO15016를 적용한 해석 결과와 비교하였다. 체계적인 비교를 위해 검증 데이터를 Beaufort number 1~7에 해당하는 데이터로 설정한 후 계측된 속도와 마력 대비 오차를 계산하여 비교하였다. Table 3은 제안하는 모델과 ISO15016에 의한 결과를 비교한 것이다.

속도 추정의 경우에는 제안하는 모델이 모든 Beaufort number에 대해 더 정확한 결과를 나타냄을 확인할 수 있었다.

Table 3 Error comparison with ISO15016 according to Beaufort number

Beaufort number	Power prediction error		Speed prediction error	
	ISO15016	Proposed.	ISO15016	Proposed.
1	1.91	3.82	1.97	1.01
2	3.21	4.30	2.62	1.71
3	6.07	4.06	2.67	1.91
4	3.16	1.97	3.08	1.37
5	14.63	5.40	7.20	4.11
6	11.39	3.26	6.96	3.37
7	12.55	7.27	7.03	4.18
Mean	7.56	4.30	4.50	2.52

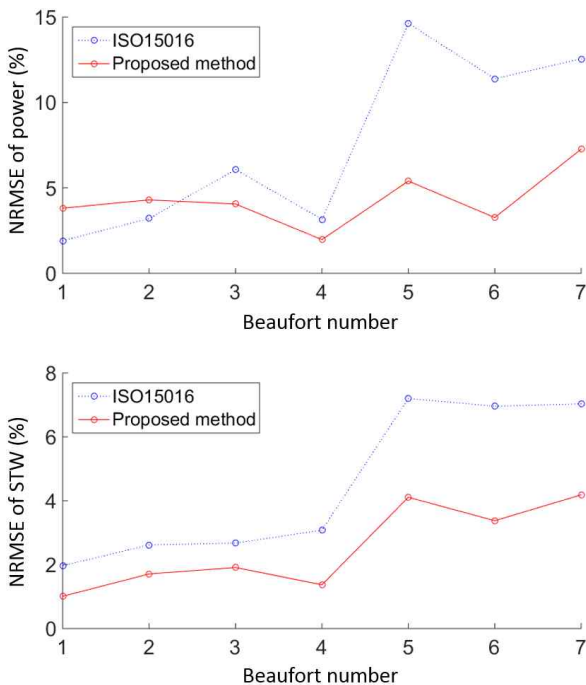


Fig. 3 Error plots of the proposed model with ISO15016 according to Beaufort number

마력 추정의 경우에는 Beaufort number 1, 2와 같이 비교적 환경 외란이 작은 조건에서는 풍동 시험, 수조 시험과 같은 추가적인 데이터를 사용하는 ISO15016의 결과가 더 정확하였으나, 제안하는 모델과의 오차가 크지 않았다. 오히려 그 외의 환경 외란이 큰 구간에서는 제안하는 모델의 결과가 ISO15016에 비해 상당히 정확한 결과를 보이는 것을 확인할 수 있었다.

5. 결 론

본 논문에서는 실선 운항 기록을 이용하여 실제 해상 환경에서의 운항 성능 모델을 가우시안 프로세스 회귀 분석을 적용하여 구성해보았다. 제안하는 모델의 정확도가 기존의 속력 시운전 해석 지침인 ISO15016에 비해 전반적으로 개선된 결과를 보이는 것을 확인할 수 있었고, 이를 통해 데이터 기반의 통계적인 방법의 유용성을 확인하였다.

후 기

이 논문은 2016년 해양수산부 재원으로 한국해양과학기술진흥원의 지원을 받아 수행된 연구임(고신뢰성 무인선 운용기술 및 인프라 구축).

참 고 문 헌

- 유병현, 김진환 (2013). “기계학습 기법을 이용한 실선 운항 성능 분석 및 모델링”, 2013 추계 대한 조선학회
- Borkowski, T., Kasyk, L. and Kowalak, P. (2011). “Assessment of ship’s engine effective power fuel consumption and emission using the vessel speed”, Journal of KONES Powertrain and Transport, Vol 18, No 2, pp 31-39.
- Petersen, J.P, Jacobsen D.J. and Winther, O. (2012). “Statistical modelling for ship propulsion efficiency”, Journal of Marine Science and Technology, Vol 2, No 1, pp 30-39.
- Sasa, K., Terada, D., Shigeaki S.S., Wakabayashi, N., Ikebuchi, T., Chen, C., Takayama, A. and Uchida, M. (2015). “Evaluation of ship performance in international maritime transportation using an onboard measurement system - in case of a bulk carrier in international voyages”, Ocean engineering, Vol 104, No 1, pp. 294-309.
- Sogihara, N., Ueno, M., Hoshino, K., Tsujimoto M., and Sasaki, N. (2010). “Verification of calculation method on ship performance by onboard measurement”, The International Society of Offshore and Polar Engineers 2010.