

KIOST-TR-2022-0186

머신러닝 기반 남해연구소 부두 변위 예측  
Displacement Prediction for the pier of South Sea  
Research Institute based on Machine learning

KIOST

KOREA INSTITUTE OF  
OCEAN SCIENCE & TECHNOLOGY

2022. 09. 23.

한 국 해 양 과 학 기 술 원

## 요 약

제 목	국문	머신러닝 기반 남해연구소 부두 변위 예측
	영문	Displacement Prediction for the pier of South Sea Research Institute based on Machine learning
관련과제명	관련 연구과제 없음	
저 자 명	강한구	
작성일자	2022년 09월 23일	
유형	기술동향자료(O), 기술업무메뉴얼( ), 특허허지적재산권( ), 결함보고서( )	
주제분야	해양물리( ), 해양화학( ), 해양생물( ), 해양지질( ), 해양공학(O) 해양정책( ), 융복합( ), 기타( )	
Keyword	인공지능, 머신러닝, 기계학습, 예측, 회귀분석	
초록	<p>본 기술보고서는 우리원 남해연구소의 잔교식 부두 변위 예측에 관한 내용을 다룬다. 2018년 이사부호 정박을 위하여 남해연구소 잔교식 부두를 확장하였다. 잔교식 부두는 4개의 블록으로 설계 및 건축되었으며, 이러한 이유로 각 블록에 변위가 발생한다. 따라서 본 기술보고서에서는 남해연구소 잔교식 부두의 변위 관측 시스템, 관측 자료, 변위 유발 요소 및 환경에 따른 변위 예측에 관하여 기술한다. 우선, 확장 건설된 남해연구소 잔교식 부두를 설명하고 다음으로 변위 관측 시스템, 변위 관측 자료, 변위 유발 요소, 변위 예측 및 변위 예측 모델의 평가를 기술한다. 본 기술보고서의 연구 결과는 확장된 남해연구소 부두의 관리를 위해 활용이 가능하며, 변위 예측 모델로 적용한 여러 모델을 회귀분석이 필요한 연구에 적용할 수 있음을 보여준다.</p>	

## 목 차

1. 개요 .....	1
2. 부두 변위 관측 시스템 .....	3
3. 부두 변위 관측 자료 .....	7
4. 부두 변위 유발 요소 .....	8
5. 부두 변위 예측 .....	10
6. 부두 변위 예측 모델 평가 .....	20
7. 결론 .....	21
8. 참고문헌 .....	22

KIOST

KOREA INSTITUTE OF  
OCEAN SCIENCE & TECHNOLOGY



## 1. 개요

본 기술보고서의 목표는 2018년 이사부호 정박을 위해서 확장 건설된 남해연구소 잔교식 부두의 변위 관측 시스템, 변위 관측 자료, 변위 유발 요소, 변위 예측 및 변위 예측 모델의 평가에 관하여 기술한다.

잔교식 부두란 그림 1과 같이 기둥 형상의 구조물로 부두 블록을 지지하는 고조의 부두를 말한다. 그림 2와 같이 남해연구소 확장 부두는 신규 대형 종합연구선 이사부호의 정박을 위하여 그림 3와 같이 4개의 부두 블록으로 확장 건설되었다. 확장된 부두는 기존의 부두 길이 대비 두 배 이상 길이가 길어졌으며, 또한 부두의 폭 또한 기존 대비 두 배로 증가하였다. 이에 따라 원활한 연구선의 정박지 확보뿐 아니라 부두를 활용한 연구개발 사업이 활발하게 진행되고 있다.

하지만 확장 건설된 남해연구소 잔교식 부두의 준공 후 계속해서 미세하게 변위가 발생한다는 보고가 있었다. 이에 따라 변위 관측 시스템을 부두 블록에 설치하여 실제로 관측하였다. 이렇게 관측된 변위 관측 자료를 바탕으로 변위 유발 요소 추정 및 변위 예측 연구를 수행하였다.



그림 1. 잔교식 부두



그림 2. 확장 건설된 남해연구소 잔교식 부두

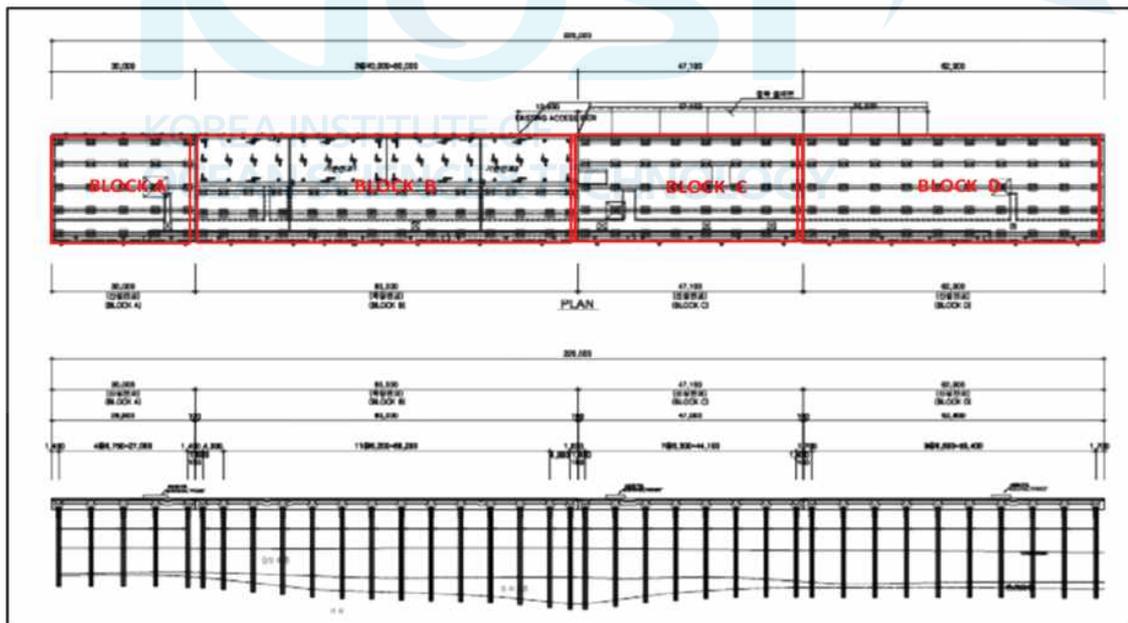


그림 3. 남해연구소 잔교식 부두 도면: Top view & Side view

## 2. 부두 변위 관측 시스템

남해연구소 확장 부두의 블록 간 변위 관측을 위하여 그림 4와 같이 3개소의 위치에 변위 관측 시스템을 설치하였다. 그림 5는 설치된 변위 관측 시스템을 보여주고 있다. 그림 6과 그림 7은 설치된 변위 센서의 사양과 도면을 보여준다. 설치된 센서의 모델은 CDP-100이며, 최대 100mm의 변위를 측정할 수 있다. 그림 8은 부두 블록 간 발생하는 변위의 수직 수평 방향에 관하여 보여주고 있다. 실제로 부두의 변위는 3차원으로 발생할 수 있으나 본 연구에서는 2차원의 변위만을 관측하고 다루었다.

그림 9는 변위 유발 요소로 추정되는 기상 자료를 관측하기 위한 시스템이 설치되어 있는 것을 보여주고 있다. 변위 유발 요소 추정 시 부두 위의 기상관측 시스템을 관측한 온도, 습도, 대기압, 풍향 및 풍속 자료를 활용하였다.

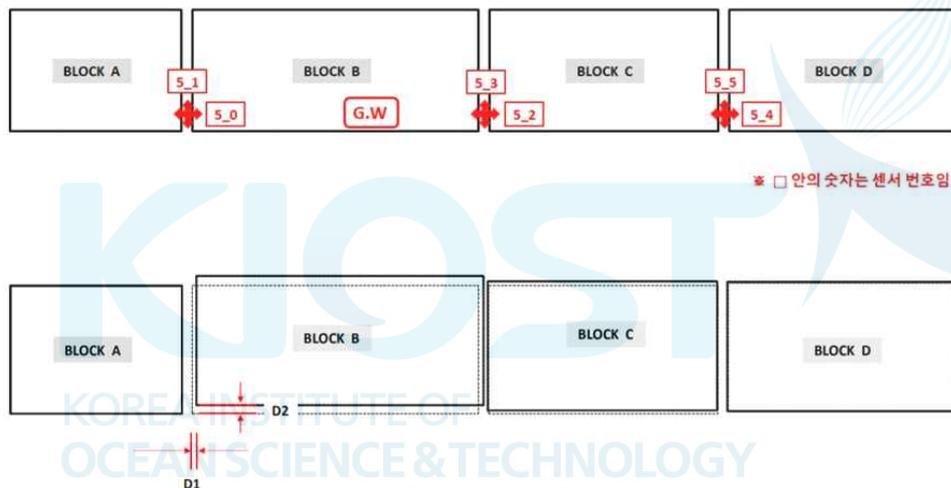


그림 4. 부두 변위 관측 시스템 설치 위치



그림 5. 부두 변위 관측 시스템

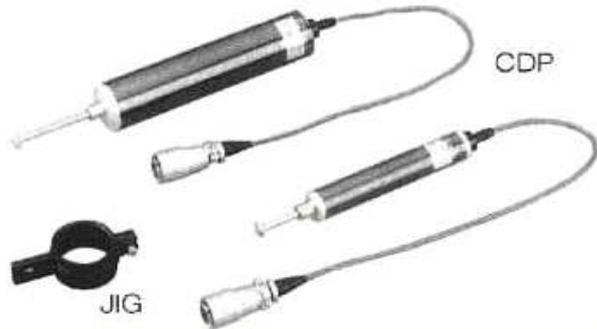
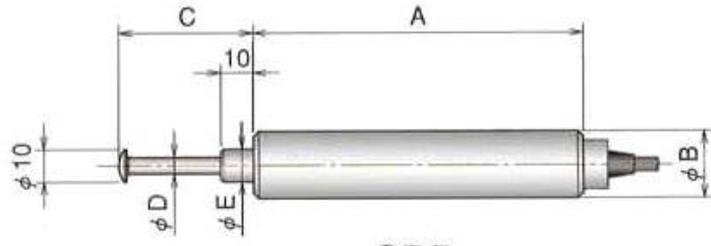
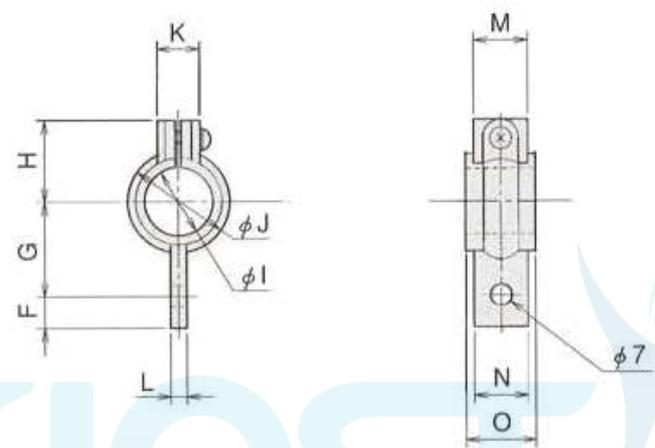
																																																																																					
<b>설 명</b>	Compact 한 형상으로 취급이 용이한 변위계이다. 출력이 크고, 안정성에도 뛰어나기 때문에 고정도의 측정을 할 수 있다. 정적측정에는 물론, 동적 측정에도 사용할 수 있다.																																																																																				
<b>특 징</b>	<ul style="list-style-type: none"> <li>- 고정도, 고효율</li> <li>- 뛰어난 응답성</li> <li>- 취급이 간단</li> </ul>																																																																																				
<b>사 양</b>	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>MODEL</th> <th>CDP-5</th> <th>CDP-10</th> <th>CDP-25</th> <th>CDP-50</th> <th>CDP-100</th> </tr> </thead> <tbody> <tr> <td>용량(mm)</td> <td>5</td> <td>10</td> <td>25</td> <td>50</td> <td>100</td> </tr> <tr> <td>정격출력</td> <td colspan="2">5mV/V (10,000×10<sup>-6</sup>)</td> <td>6.25mV/V (12,500×10<sup>-6</sup>)</td> <td colspan="2">5mV/V (10,000×10<sup>-6</sup>)</td> </tr> <tr> <td>감도(×10<sup>-6</sup>/mm)</td> <td>2,000</td> <td>1,000</td> <td>500</td> <td>200</td> <td>100</td> </tr> <tr> <td>비직선성</td> <td colspan="2">0.15% RO</td> <td colspan="3">0.1% RO</td> </tr> <tr> <td>측정력</td> <td colspan="2">6.4N (650gf)</td> <td colspan="2">3.4N (350gf)</td> <td>4.9N (500gf)</td> </tr> <tr> <td>응답주파수(Hz)</td> <td>40</td> <td>12</td> <td>8</td> <td>6</td> <td>3</td> </tr> <tr> <td>영점의 온도영향</td> <td colspan="5">1×10<sup>-6</sup>/c(0 ~ 40c)</td> </tr> <tr> <td>허용온도범위</td> <td colspan="5">0 ~ 60c</td> </tr> <tr> <td>입출력 저항</td> <td colspan="5">350Ω</td> </tr> <tr> <td>권장인가전압</td> <td colspan="5">Less than 2V</td> </tr> <tr> <td>허용인가전압</td> <td colspan="5">10V</td> </tr> <tr> <td>케이블</td> <td colspan="5">JR-9510(∅6mm, 0.3mm<sup>2</sup>, 4CORE SHIELD, 10m)</td> </tr> <tr> <td>HOLDER</td> <td colspan="3">1 pc. Supplied</td> <td colspan="2">2 pcs. Supplied</td> </tr> </tbody> </table>	MODEL	CDP-5	CDP-10	CDP-25	CDP-50	CDP-100	용량(mm)	5	10	25	50	100	정격출력	5mV/V (10,000×10 <sup>-6</sup> )		6.25mV/V (12,500×10 <sup>-6</sup> )	5mV/V (10,000×10 <sup>-6</sup> )		감도(×10 <sup>-6</sup> /mm)	2,000	1,000	500	200	100	비직선성	0.15% RO		0.1% RO			측정력	6.4N (650gf)		3.4N (350gf)		4.9N (500gf)	응답주파수(Hz)	40	12	8	6	3	영점의 온도영향	1×10 <sup>-6</sup> /c(0 ~ 40c)					허용온도범위	0 ~ 60c					입출력 저항	350Ω					권장인가전압	Less than 2V					허용인가전압	10V					케이블	JR-9510(∅6mm, 0.3mm <sup>2</sup> , 4CORE SHIELD, 10m)					HOLDER	1 pc. Supplied			2 pcs. Supplied	
MODEL	CDP-5	CDP-10	CDP-25	CDP-50	CDP-100																																																																																
용량(mm)	5	10	25	50	100																																																																																
정격출력	5mV/V (10,000×10 <sup>-6</sup> )		6.25mV/V (12,500×10 <sup>-6</sup> )	5mV/V (10,000×10 <sup>-6</sup> )																																																																																	
감도(×10 <sup>-6</sup> /mm)	2,000	1,000	500	200	100																																																																																
비직선성	0.15% RO		0.1% RO																																																																																		
측정력	6.4N (650gf)		3.4N (350gf)		4.9N (500gf)																																																																																
응답주파수(Hz)	40	12	8	6	3																																																																																
영점의 온도영향	1×10 <sup>-6</sup> /c(0 ~ 40c)																																																																																				
허용온도범위	0 ~ 60c																																																																																				
입출력 저항	350Ω																																																																																				
권장인가전압	Less than 2V																																																																																				
허용인가전압	10V																																																																																				
케이블	JR-9510(∅6mm, 0.3mm <sup>2</sup> , 4CORE SHIELD, 10m)																																																																																				
HOLDER	1 pc. Supplied			2 pcs. Supplied																																																																																	

그림 6. 변위 센서 사양



CDP



JIG CDPF-11

MODEL/DIM.	A	$\phi B$	C	$\phi D$	$\phi E$
CDP-5	99	20.5	20	5	10
CDP-10	99	20.5	24	5	10
CDP-25	114	20.5	40	5	10
CDP-50	154	33.5	65	5	10
CDP-100	274	41	118	6	12

MODEL/DIM.	적용 CDP	F	G	H	I	J	K	L	M	N	O
CDPF-11-25	CDP-5~25	10	28	25	20.5	30	13	5	15	15	20
CDPF-11-50	CDP-50	10	35	32	33.5	43	13	5	15	15	20
CDPF-11-100	CDP-100	11	36	40	41	50	17.5	8	14	16	25

그림 7. 변위 센서 도면



< 부두의 단차 변위량 측정 >

< 부두의 수평 변위량 측정 >

그림 8. 부두에 발생하는 변위의 방향



그림 9. 부두 기상관측 센서 설치 위치

### 3. 부두 변위 관측 자료

남해연구소 확장 부두 변위 관측 자료와 변위를 유발하는 기상 환경 관측 자료는 그림 10과 같다. 자료는 15분 단위로 로깅을 하였으며 본 연구에서는 2019-09-06 05:04:59 ~ 2019-09-25 01:06:02 구간의 자료를 활용하였다. 또한, 방대한 양의 관측 데이터로 인하여 본 연구에는 3개소에 설치된 변위 관측 시스템 중 그림 11과 같이 부두 블록 A와 B 사이에 설치된 변위 관측 시스템의 변위 관측 자료만을 활용하였다. 자료는 변위, 변위계 주변 온도, 변위계 주변 습도, 기상관측 탑의 기상 온도, 노점온도, 체감온도, 습도, 기압, 풍속, 풍향 및 풍속 정보를 포함하고 있다.

Index	변위 (mm)	변위계 온도 deg	변위계 습도 %	기상 온도 deg	노점온도 deg	체감온도 deg	습도 %	기압 hPa	풍속 m/s	풍향 deg
0	0.95399	30.556	89.018	27.4	26	29	89	1006	6	125
1	0.96289	30.566	89.018	27.4	26	29	89	1006	5	134
2	0.95695	30.556	89.018	27.4	26	29	89	1006	6	113
3	0.96585	30.556	89.018	27.3	25	29	89	1005	6	134
4	0.96289	30.536	89.018	27.4	25	29	89	1005	6	125

그림 10. 변위 및 기상 관측 자료

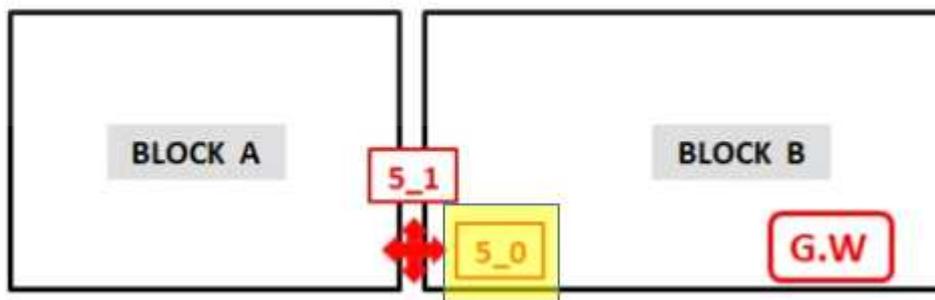


그림 11. 본 연구에 활용한 변위 관측 시스템 위치

## 4. 부두 변위 유발 요소

그림 12는 본 연구에 활용된 변위 관측 자료 및 기상관측 자료를 보여주고 있다. 각 관측 자료를 분리하여 보여주고 있으며, X축의 값을 시간으로 통일하여 비교해 보기 쉽게 그래프를 작성하였다. 그림 13은 Y축의 값을 0과 1 사이로 정규화한 값을 변위 관측 자료를 비교하여 보여주고 있다. 그림 13을 통해서 변위는 온도와 관련된 인자들을 추종하는 것을 육안으로 확인할 수 있다.

그래프만으로 변위 유발 요소를 육안으로 확인하는 것은 다소 어려울 수 있다. 또한 정보는 사람에 따라서 그 판단 정도가 달라진다. 따라서 Correlation Coefficient 값을 확인하여 어떤 인자가 변위를 유발하는지 수치로 확인하였다. 표 1을 통해서 변위 유발 요소를 수치로 확인 가능하며, 변위계 주변 온도, 기상관측 온도 및 노점온도(이슬점)이 변위와 유의미한 양의 상관 관계가 있는 것을 확인할 수 있다.

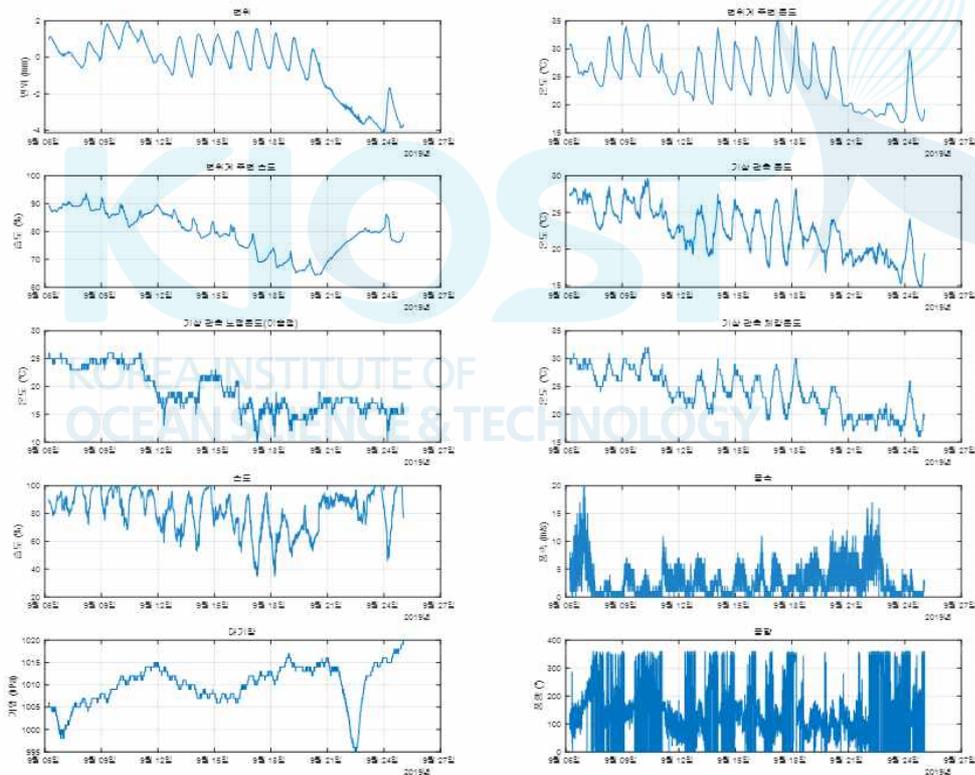
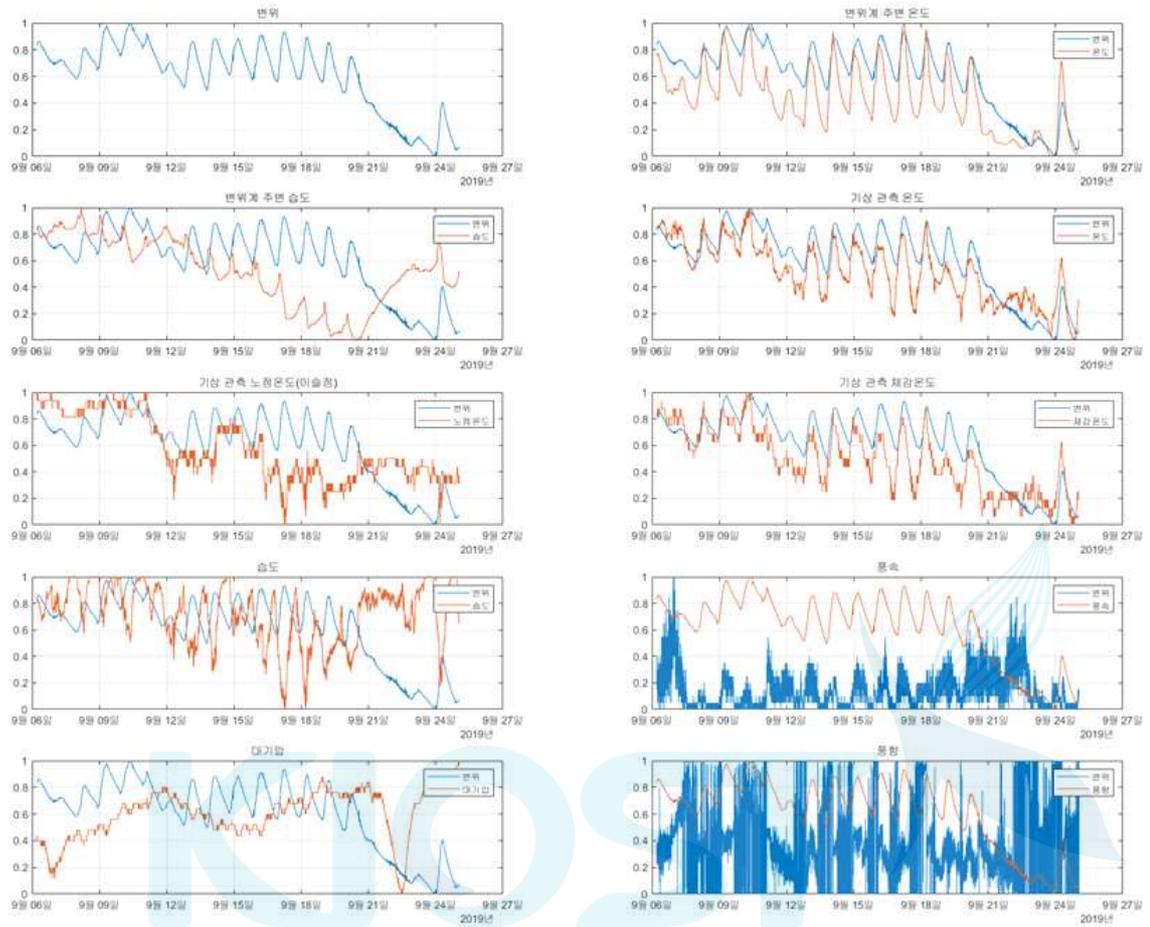


그림 12. 변위 및 기상 관측 자료



KOREA INSTITUTE OF OCEAN SCIENCE & TECHNOLOGY  
 그림 13. 정규화된 변위 및 기상 관측 자료

표 1. 부두 변위 유발 요소(Correlation Coefficient)

	변위	변위계 주변 온도	변위계 주변 습도	기상관측 온도	기상관측 체감온도	기상관측 노점온도(이슬점)	기상관측 습도	기상관측 대기압	기상관측 풍속	기상관측 풍향
변위	1	0.858555	0.1939139	0.8159881	0.4672747	0.8231826	-0.267469	-0.177884	-0.011671	-0.009827
변위계 주변 온도	0.858555	1	0.2145693	0.8770831	0.3432046	0.8772941	-0.473791	-0.178116	0.035782	0.0940572
변위계 주변 습도	0.1939139	0.2145693	1	0.4789888	0.7222367	0.5178419	0.4016121	-0.429824	-0.101084	0.1327484
기상관측 온도	0.8159881	0.8770831	0.4789888	1	0.6241801	0.9830196	-0.261298	-0.426775	0.1667648	0.1135995
기상관측 체감온도	0.4672747	0.3432046	0.7222367	0.6241801	1	0.6526062	0.5809189	-0.500694	-0.027672	0.0480925
기상관측 노점온도(이슬점)	0.8231826	0.8772941	0.5178419	0.9830196	0.6526062	1	-0.203041	-0.396375	0.0531663	0.11283
기상관측 습도	-0.267469	-0.473791	0.4016121	-0.261298	0.5809189	-0.203041	1	-0.171671	-0.225863	-0.060722
기상관측 대기압	-0.177884	-0.178116	-0.429824	-0.426775	-0.500694	-0.396375	-0.171671	1	-0.357588	-0.065588
기상관측 풍속	-0.011671	0.035782	-0.101084	0.1667648	-0.027672	0.0531663	-0.225863	-0.357588	1	-0.014706
기상관측 풍향	-0.009827	0.0940572	0.1327484	0.1135995	0.0480925	0.11283	-0.060722	-0.065588	-0.014706	1

## 5. 부두 변위 예측

부두의 변위를 예측하기 위해서 회귀 예측을 적용하였다. 회귀 예측에 Linear Regression, Lasso Regression, Ridge Regression, Random Forest, Gradient Boosting 및 ANN(Artificial Neural Network) Regression 모델을 적용하였다. 회귀 예측 모델 적용을 위해서 관측 자료를 파이썬 코드로 작성하였으며, 그림 14-15와 같이 학습데이터와 모델 테스트 데이터를 7:3 비율로 분할 하였다. 그림 16-30은 각 Regression 모델의 학습 코드 및 성능을 보여주고 있다.

```
data_path = 'data_9th_d0_weather.csv'  
data = pd.read_csv(data_path)  
data.head(5)
```

	0.95992	30.556	89.018	27.4	26	29	89	1006	5	132
0	0.95399	30.556	89.018	27.4	26	29	89	1006	6	125
1	0.96289	30.566	89.018	27.4	26	29	89	1006	5	134
2	0.95695	30.556	89.018	27.4	26	29	89	1006	6	113
3	0.96585	30.556	89.018	27.3	25	29	89	1005	6	134
4	0.96289	30.536	89.018	27.4	25	29	89	1005	6	125

KOREA INSTITUTE OF  
OCEANOGRAPHY

그림 14. 변위 및 기상 관측 자료 원본 확인 코드

```
np_data = np.array(data)
print(np_data[0:5,:])
```

```
[[9.5399e-01 3.0556e+01 8.9018e+01 2.7400e+01 2.6000e+01 2.9000e+01
 8.9000e+01 1.0060e+03 6.0000e+00 1.2500e+02]
 [9.6289e-01 3.0566e+01 8.9018e+01 2.7400e+01 2.6000e+01 2.9000e+01
 8.9000e+01 1.0060e+03 5.0000e+00 1.3400e+02]
 [9.5695e-01 3.0556e+01 8.9018e+01 2.7400e+01 2.6000e+01 2.9000e+01
 8.9000e+01 1.0060e+03 6.0000e+00 1.1300e+02]
 [9.6585e-01 3.0556e+01 8.9018e+01 2.7300e+01 2.5000e+01 2.9000e+01
 8.9000e+01 1.0050e+03 6.0000e+00 1.3400e+02]
 [9.6289e-01 3.0536e+01 8.9018e+01 2.7400e+01 2.5000e+01 2.9000e+01
 8.9000e+01 1.0050e+03 6.0000e+00 1.2500e+02]]
```

```
datax = np_data[:,1:10]
datay = np_data[:,0]
print(datax[0:5,:])
print(datay[0:10])
```

```
[[ 30.556  89.018  27.4   26.    29.    89.   1006.   6.
 125. ]
 [ 30.566  89.018  27.4   26.    29.    89.   1006.   5.
 134. ]
 [ 30.556  89.018  27.4   26.    29.    89.   1006.   6.
 113. ]
 [ 30.556  89.018  27.3   25.    29.    89.   1005.   6.
 134. ]
 [ 30.536  89.018  27.4   25.    29.    89.   1005.   6.
 125. ]]
[0.95399 0.96289 0.95695 0.96585 0.96289 0.96289 0.96289 0.96882 0.96882
 0.96585]
```

```
from sklearn.model_selection import train_test_split
trnx, tstx, trny, tsty = train_test_split(datax, datay, test_size=0.3)
print(trnx.shape, tstx.shape, trny.shape, tsty.shape)
```

```
(18755, 9) (8038, 9) (18755,) (8038,)
```

그림 15. 학습 데이터 및 모델 검증 데이터 분할 코드

```

: from sklearn.linear_model import LinearRegression
  model = LinearRegression()
  model.fit(X=trnx, y=trny)
  lr_pred = model.predict(X=tstx)

  from sklearn import metrics
  print('RMSE', np.sqrt(metrics.mean_squared_error(lr_pred, tsty)))
  print('R2', metrics.r2_score(tsty, lr_pred))

RMSE 0.6649995666353837
R2 0.8197161726500831

: model
: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)

```

그림 16. Linear Regression 학습 코드 및 결과

```

plt.plot(tsty[1:100])
plt.plot(lr_pred[1:100])
plt.legend(['test data', 'prediction result'], loc='upper right')
plt.title('Linear Regression')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()

```

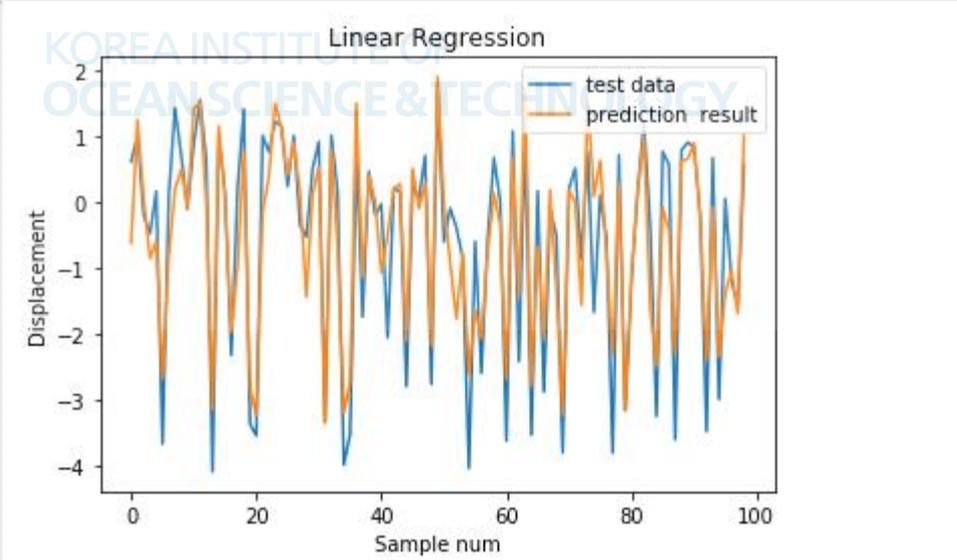


그림 17. Linear Regression 모델 테스트 결과

```

from sklearn.linear_model import Lasso
model_lasso = Lasso()
model_lasso.fit(X=trnx, y=trny)
lr_pred_lasso = model_lasso.predict(X=tstx)
print('RMSE:', np.sqrt(metrics.mean_squared_error(lr_pred_lasso, tsty)))
print('R2', metrics.r2_score(tsty, lr_pred_lasso))

```

RMSE: 0.7940180894516775  
R2 0.7429753134200643

```
model_lasso
```

```

Lasso(alpha=1.0, copy_X=True, fit_intercept=True, max_iter=1000,
      normalize=False, positive=False, precompute=False, random_state=None,
      selection='cyclic', tol=0.0001, warm_start=False)

```

그림 18. Lasso Regression 학습 코드 및 결과

```

plt.plot(tsty[1:100])
plt.plot(lr_pred_lasso[1:100])
plt.legend(['test data', 'prediction result'], loc='upper right')
plt.title('Lasso Regression')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()

```

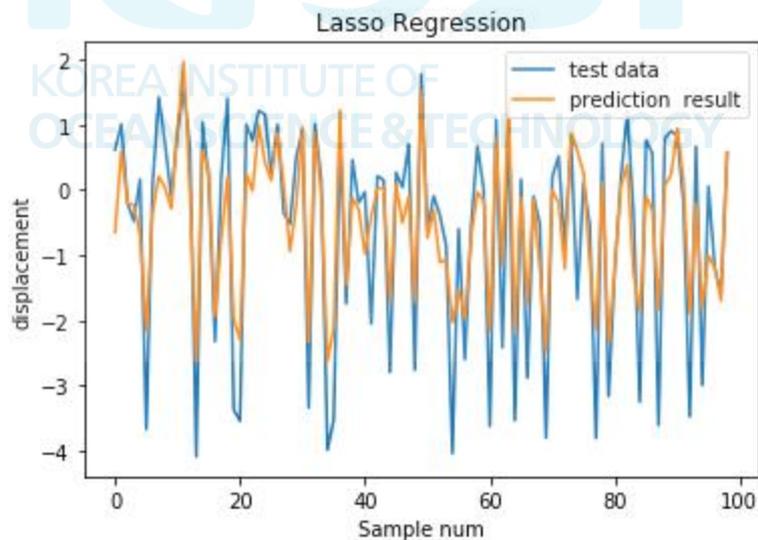


그림 19. Lasso Regression 모델 테스트 결과

```

from sklearn.linear_model import Ridge
model_ridge = Ridge()
model_ridge.fit(X=trnx, y=trny)
lr_pred_ridge = model_ridge.predict(X=tstx)
print('RMSE:', np.sqrt(metrics.mean_squared_error(lr_pred_ridge, tsty)))
print('R2', metrics.r2_score(tsty, lr_pred_ridge))

```

```

RMSE: 0.664999140868317
R2 0.8197164035040434

```

```
model_ridge
```

```

Ridge(alpha=1.0, copy_X=True, fit_intercept=True, max_iter=None,
      normalize=False, random_state=None, solver='auto', tol=0.001)

```

그림 20. Ridge Regression 학습 코드 및 결과

```

plt.plot(tsty[1:100])
plt.plot(lr_pred_ridge[1:100])
plt.legend(['test data', 'prediction result'], loc='upper right')
plt.title('Ridge Regression')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()

```

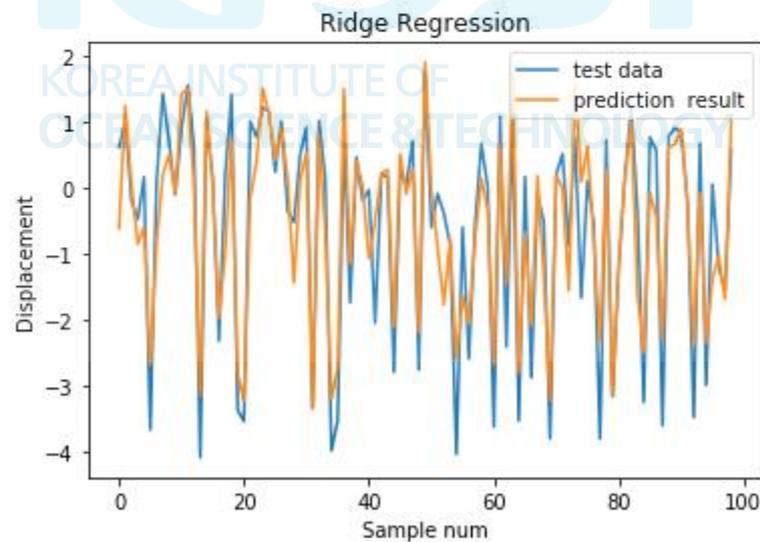


그림 21. Ridge Regression 모델 테스트 결과

```

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
scaler.fit(trnx)
trnx_scale = scaler.transform(trnx)
tstx_scale = scaler.transform(tstx)
# print(np.min(trnx_scale[:,0]), np.max(trnx_scale[:,0]))
# print(np.min(tstx_scale[:,0]), np.max(tstx_scale[:,0]))

from sklearn.ensemble import RandomForestRegressor
rf_model = RandomForestRegressor(max_depth=15, n_estimators=100, random_state=0)
rf_model.fit(X=trnx, y=trny)

rf_pred = rf_model.predict(X=tstx)

from sklearn import metrics
print('RMSE', np.sqrt(metrics.mean_squared_error(rf_pred, tsty)))
print('R2', metrics.r2_score(tsty, rf_pred))

RMSE 0.03383381256799827
R2 0.9995333235097635

```

```

rf_model
RandomForestRegressor(bootstrap=True, ccp_alpha=0.0, criterion='mse',
                       max_depth=15, max_features='auto', max_leaf_nodes=None,
                       max_samples=None, min_impurity_decrease=0.0,
                       min_impurity_split=None, min_samples_leaf=1,
                       min_samples_split=2, min_weight_fraction_leaf=0.0,
                       n_estimators=100, n_jobs=None, oob_score=False,
                       random_state=0, verbose=0, warm_start=False)

```

그림 22. Random Forest Regression 학습 코드 및 결과

```

plt.plot(tsty[1:100])
plt.plot(rf_pred[1:100])
plt.legend(['test data', 'prediction result'], loc='upper right')
plt.title('Ensemble_Random_Forest_Regression')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()

```

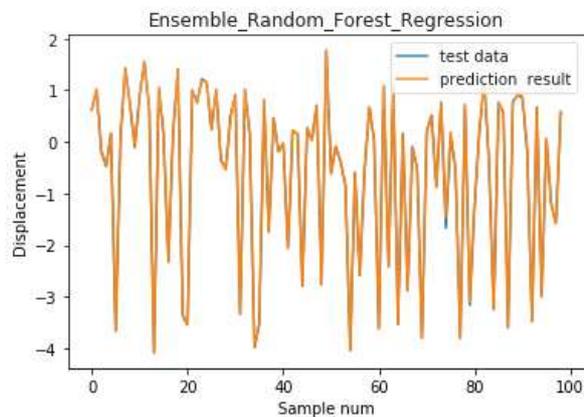


그림 23. Random Forest Regression 모델 테스트 결과

```

from sklearn.ensemble import GradientBoostingRegressor
gbm_model = GradientBoostingRegressor(max_depth=15, n_estimators=30, random_state=0)
gbm_model.fit(X=trnx, y=trny)

gbm_pred = gbm_model.predict(X=tstx)

from sklearn import metrics
print('RMSE', np.sqrt(metrics.mean_squared_error(gbm_pred, tsty)))
print('R2', metrics.r2_score(tsty, gbm_pred))

```

```

RMSE 0.07330425049205963
R2 0.9978093532073847

```

```
gbm_model
```

```

GradientBoostingRegressor(alpha=0.9, ccp_alpha=0.0, criterion='friedman_mse',
                           init=None, learning_rate=0.1, loss='ls', max_depth=15,
                           max_features=None, max_leaf_nodes=None,
                           min_impurity_decrease=0.0, min_impurity_split=None,
                           min_samples_leaf=1, min_samples_split=2,
                           min_weight_fraction_leaf=0.0, n_estimators=30,
                           n_iter_no_change=None, presort='deprecated',
                           random_state=0, subsample=1.0, tol=0.0001,
                           validation_fraction=0.1, verbose=0, warm_start=False)

```

그림 24. Gradient Boosting Regression 학습 코드 및 결과

```

plt.plot(tsty[1:100])
plt.plot(gbm_pred[1:100])
plt.legend(['test data', 'prediction result'], loc='upper right')
plt.title('Gradient Boosting')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()

```

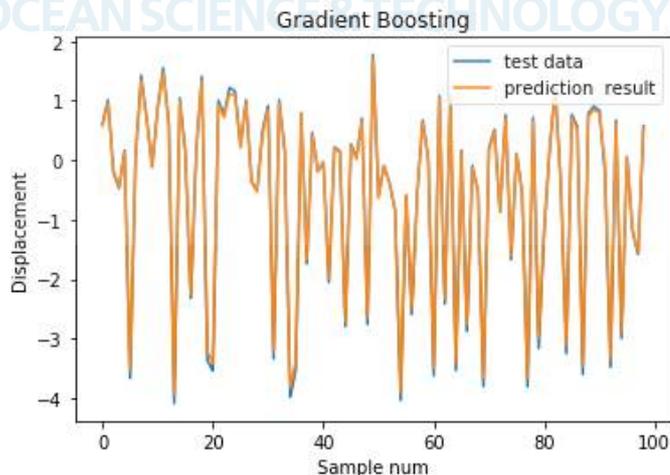


그림 25. Gradient Boosting Regression 모델 테스트 결과

```

mlp_model = models.Sequential()
mlp_model.add(layers.Dense(10, input_dim=9, kernel_initializer='normal', activation='relu'))
mlp_model.add(layers.Dense(30, activation='relu'))
mlp_model.add(layers.Dense(100, activation='relu'))
mlp_model.add(layers.Dense(30, activation='relu'))
mlp_model.add(layers.Dense(10, activation='relu'))
mlp_model.add(layers.Dense(1, activation='linear'))
mlp_model.summary()

```

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 10)	100
dense_1 (Dense)	(None, 30)	330
dense_2 (Dense)	(None, 100)	3100
dense_3 (Dense)	(None, 30)	3030
dense_4 (Dense)	(None, 10)	310
dense_5 (Dense)	(None, 1)	11

Total params: 6,881  
Trainable params: 6,881  
Non-trainable params: 0

그림 26. ANN(Artificial Neural Network) Regression 생성 모델

```

mlp_model.compile(loss='mse', optimizer='adam', metrics=['mse', 'mae'])
history = mlp_model.fit(trnx, trny, epochs=100, batch_size=50, verbose=1, validation_split=0.2)

```

Train on 15004 samples, validate on 3751 samples

Epoch 1/100  
15004/15004 [=====] - 1s 63us/sample - loss: 1.6453 - mse: 1.6453 - mae: 1.0007 - val\_loss: 0.7852 - val\_mse: 0.7852 - val\_mae: 0.6863

Epoch 2/100  
15004/15004 [=====] - 0s 27us/sample - loss: 0.6594 - mse: 0.6594 - mae: 0.6400 - val\_loss: 0.5347 - val\_mse: 0.5347 - val\_mae: 0.5821

Epoch 3/100  
15004/15004 [=====] - 0s 27us/sample - loss: 0.4987 - mse: 0.4987 - mae: 0.5396 - val\_loss: 0.4283 - val\_mse: 0.4283 - val\_mae: 0.4829

Epoch 4/100  
15004/15004 [=====] - 0s 23us/sample - loss: 0.3988 - mse: 0.3988 - mae: 0.4709 - val\_loss: 0.4565 - val\_mse: 0.4565 - val\_mae: 0.5255

Epoch 5/100  
15004/15004 [=====] - 0s 24us/sample - loss: 0.3369 - mse: 0.3369 - mae: 0.4320 - val\_loss: 0.2760 - val\_mse: 0.2760 - val\_mae: 0.3920

Epoch 6/100  
15004/15004 [=====] - 0s 26us/sample - loss: 0.3035 - mse: 0.3035 - mae: 0.4118 - val\_loss: 0.3615 - val\_mse: 0.3615 - val\_mae: 0.4374

그림 27. ANN(Artificial Neural Network) Regression 모델 학습 과정

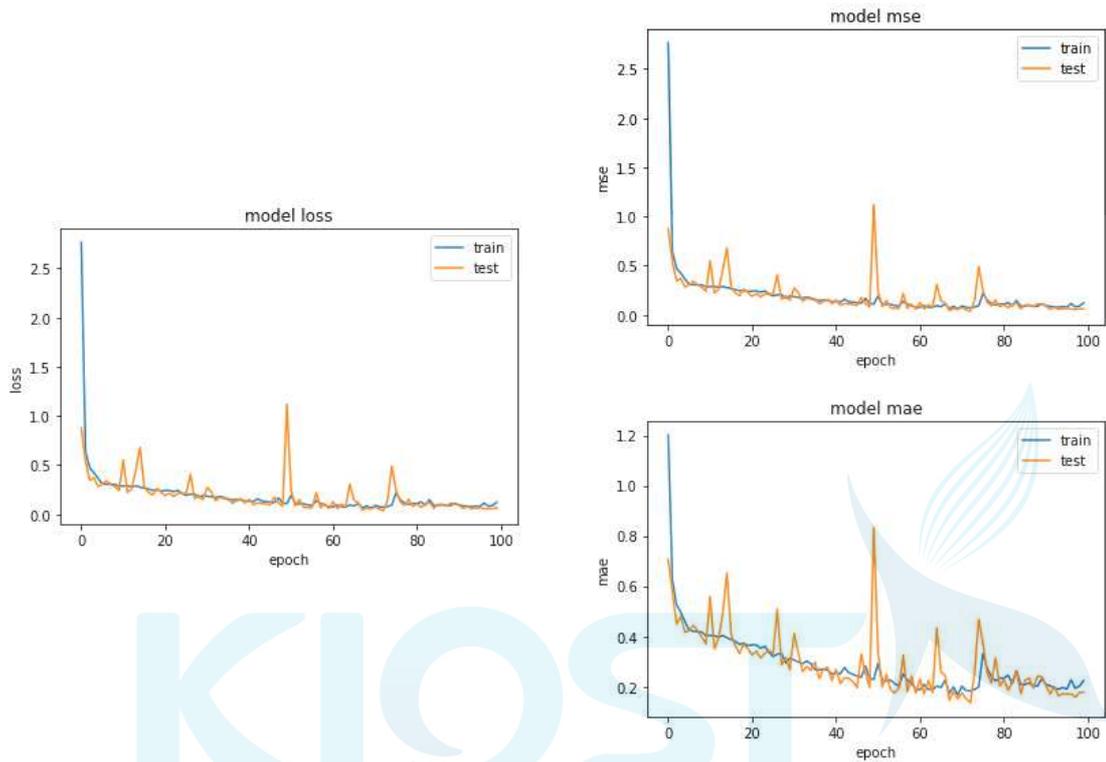


그림 28 ANN(Artificial Neural Network) Regression 모델 학습률 곡선

```

from sklearn import metrics
print('RMSE', np.sqrt(metrics.mean_squared_error(mlp_model, tsty)))
print('R2', metrics.r2_score(tsty, mlp_model))

```

RMSE 0.25829274891393317  
R2 0.972801905807386

그림 29 ANN(Artificial Neural Network) Regression 모델 성능

```
mlp_model = mlp_model.predict(tstx)
plt.plot(tsty[1:100])
plt.plot(mlp_model[1:100])
plt.legend(['test', 'prediction result'], loc='lower left')
plt.title('NN Regression')
plt.ylabel('Displacement')
plt.xlabel('Sample num')
plt.show()
```

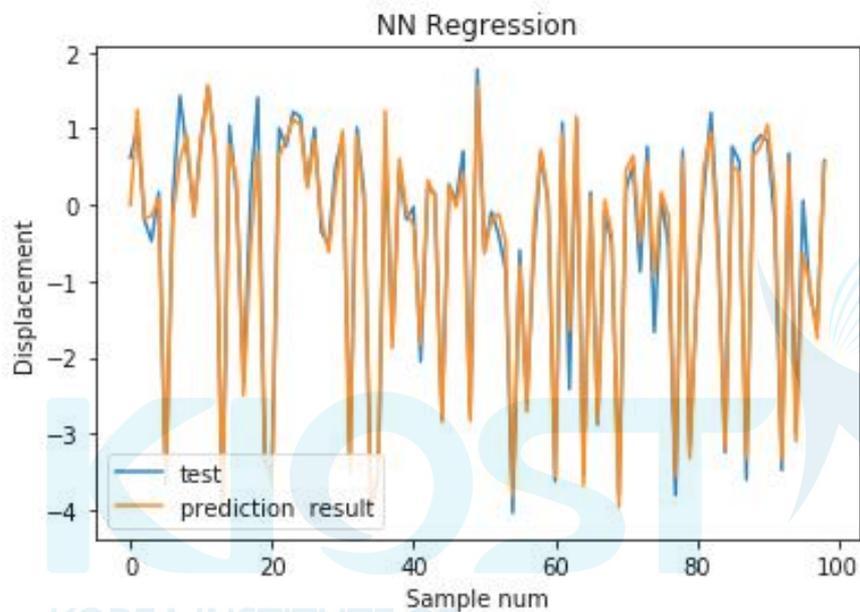


그림 30. ANN(Artificial Neural Network) Regression 모델 테스트 결과

## 6. 부두 변위 예측 모델 평가

표. 2는 부두 변위 예측 모델의 성능을 보여주고 있다. 성능 평가를 위한 값으로 RMSE와 R2 값을 계산하였으며 성능 비교 결과 Random Forest Regression 모델이 test dataset을 가장 잘 예측하는 성능을 보였으며, 다음으로 Gradient Boosting Regression 모델이 유사한 성능을 보였다. 그에 반하여 RMSE 값의 경우 Linear Regression 모델이 최저 성능을 보임을 확인하였으며, R2 값의 경우 Lasso Regression 모델이 최저 성능을 보임을 확인하였다.

표. 2 예측 모델 성능 비교표

Regression 모델	RMSE	R2
Linear	0.6649	0.8197
Lasso	0.7940	0.7429
Ridge	0.6649	0.8197
Gradient Boosting	0.0733	0.9978
Random Forest	0.03383	0.9995
ANN	0.2582	0.9728

KOREA INSTITUTE OF  
OCEAN SCIENCE & TECHNOLOGY

## 7. 결론

본 기술보고서의 연구 내용을 통해서 잔교식 부두 블록의 변위 및 기상관측 자료를 활용하여 변위 유발 요소의 상관관계 계수를 추정할 수 있었다. 부두 블록 변위 유발 요소로는 기온과 노점온도가 가장 큰 상관관계를 보임 정량적으로 확인하였다. 또한, 관측 Dataset을 활용하여 여러 Regression 모델을 생성하고 비교 평가하였으며, Random Forest Regression 모델이 Test dataset를 가장 잘 예측하는 것을 확인할 수 있었다.

추후 연구 내용으로 관측 기간 전체 Data를 이용한 분석 및 모델링, 선박 묘박 상태 (선박이 부두에 묶인 상태) Data를 포함하여 부두 변위 예측 모델을 생성 추진할 예정이다. 본 연구 결과는 부두 시설 안전 관리를 위한 근거 자료로 활용될 수 있으며, 활용된 여러 가지 Regression 모델을 활용하여 다양한 방식으로 회귀분석 연구에 활용 가능성을 보여주었다.



## 8. 참고문헌

### [Book]

Hart, Peter E., David G. Stork, and Richard O. Duda. Pattern classification. Hoboken: Wiley, 2000.

Mitchell, Tom M., and Tom M. Mitchell. Machine learning. Vol. 1. No. 9. New York: McGraw-hill, 1997.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

