Yacht Hydrodynamics

Russell Shomberg

Abstract

Machine learning using polynomial regression implimented in python
to predict residual resistance based yacht geometry.

Introduction

Yacht Hydrodynamics

Sailing yachts are considered slender floating bodies. Therefore, their movement
through water can be assumed to be soley resisted by frictional forces and
wave-making forces.

Resistance on such a body is known to be a function of both Reynolds Number
and Froude Number. Both are dimensionless numbers that relate ratios of forces.
Reynolds Number represents a ratio of inertial and frictional forces. Froude
Number represents a ratio of inertial and wave making forces.

Actual forces on ships can be predicted by determining the forces that occur
on models with the similar dimensionless parameters. Unfortunately, creating
models that match both Froude and Reynolds similarity requires conditions that
do not exist in the real world.

Froude's Hypothesis

Froude's hypotheis states resistance on a ship is actually a sum of two forces.
One force, Frictional Resistance, is completely a function of Reynolds number.
The other, Residual Resistance, is only a function of Froud number.

Frictional resistance can be easily determined by assuming the ship or model
is a flat plate of equal Reynolds number. Many emperical experiments have
been performed allowing the approximation of Frictional Resistance. Thus by
subtracting the Frictional Resistance from total resistance on the model, the
Residual Resistance is obtained. This value can then be scaled up to a represent
the Residule Resistance on a full scale ship.

Residual Resistance is known to be a function of Froude Number. However, it is also influenced by other factors in the model's geometry and cannot be calculated analytically.

Yacht Geometry

The models are built to the same length and water line lengths (1.0 m). Additionally, this report is primarily conserned with the following adimensional geomtric ratios.

1. Longitudinal position of the center of buoyancy
2. Prismatic coefficient
3. Length-displacement ratio
4. Beam-draught ratio
5. Length-beam ratio
6. Froude number[1]

Experimental Methods

A series of experiments was carried out by J. Gerritsma, R. Onnink, and A. Versluis known as the Delft Systematic Yacht Hull Form Experiments[2]. The experiment series consisted of 308 experiments using geometry variations of model sailing yachts. Each experiment varied the following adimensional factors until all posibilities were covered.

Variations	Geometry
5	Center of Bouyancy
10	Prismatic Coefficient
8	Length / Displacement
17	Beam / Draught
10	Length / Beam
14	Froude Number

The experiment measured force on the model. Then in accordance with Froude's Hypothesis subtracted out the calculated Frictional Resistance. This resulted in the measured variable of Residual Resistance.

---

[1] dimensionless representation of speed

[2] Gerritsma, Onnink, and Versluis (1981)

# Results

Raw Data Analysis

Data is arranged into the following columns.

1. Longitudinal position of the center of buoyancy, adimensional.
2. Prismatic coefficient, adimensional.
3. Length-displacement ratio, adimensional.
4. Beam-draught ratio, adimensional.
5. Length-beam ratio, adimensional.
6. Froude number, adimensional.

The measured variable is the residuary resistance per unit weight of displacement:

7. Residuary resistance per unit weight of displacement, adimensional.

Largest Contributing Factor
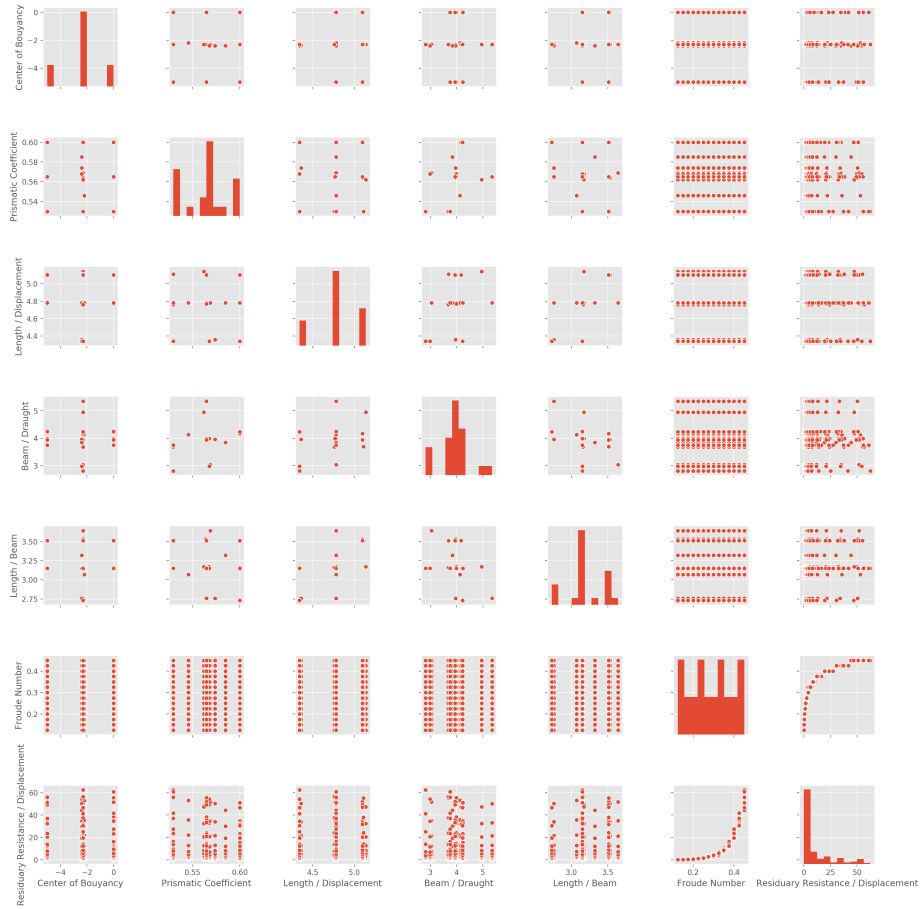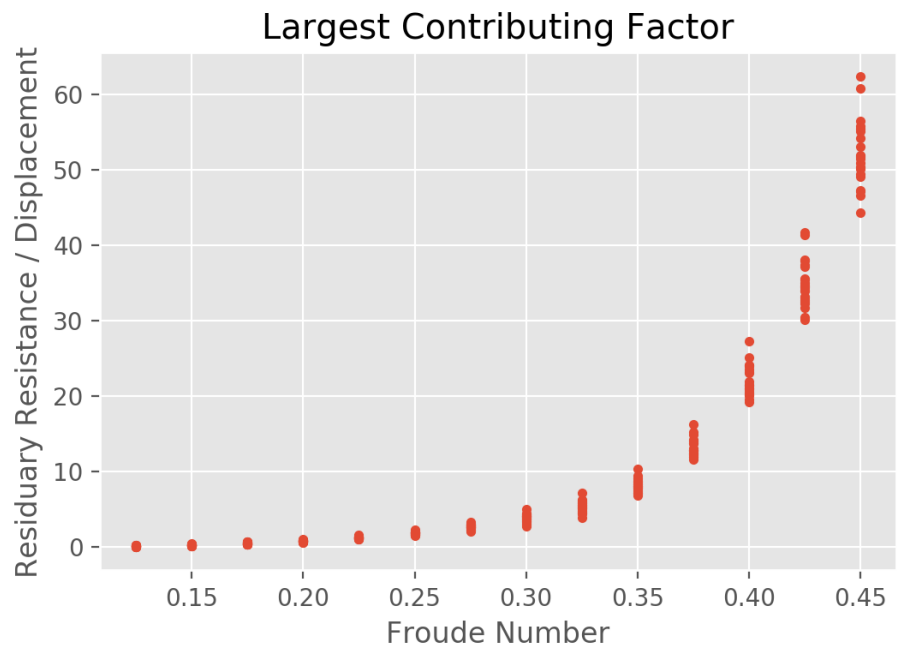
The residuary resistance to displacement ratio is most correlated to the Froude Number. However, it is not completely dependent on the Froude Number. The relationship is basically exponential, can be seen to be influenced by other factors.
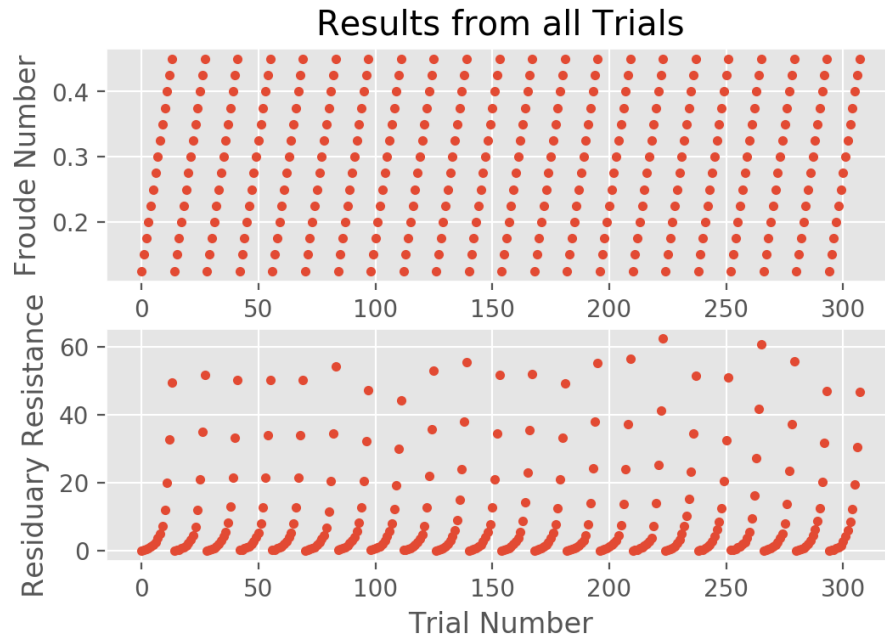
Figure shows the individual results of every trial. The systematic procedure of the experiments can be seen here.

Besides Froude Number, all remaining factors influence by at least an order of magnitude less. The influence of the other factors shows a less clear realationship. We will use polynomial regression with machine learning to determine the relationship between all the inputs and Residuary Resistance to Displacement ratio.

```
X = yachtData[[
    'Center of Bouyancy'
    ,'Prismatic Coefficient'
    ,'Length / Displacement'
    ,'Beam / Draught'
    ,'Length / Beam'
    ,'Froude Number'
    ]].values

Y = yachtData[['Residuary Resistance / Displacement']].values

testSize = 0.4
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=testSize)
```

No simple analytical solution exists to finding residuary resistance from geometry and Froude number. However, solutions can be found numerically with finite

element analysis. Clear relationships are expected to occur based on fluid dynamics. For example, Froude number is a representation of speed and therefore resistance can be expected to increase with the square. Additionally, wider, deeper, and generally less hydrodynamic forms should increase the observed resistance. While the trends are known, the exact influence and influence weighting is not.

Intuition and the highly structured nature of the data, support the idea that supervised machine learning in the form of regression will be highly suited to this task. Because of this judgement, a test size of 0.4 has been chosen.
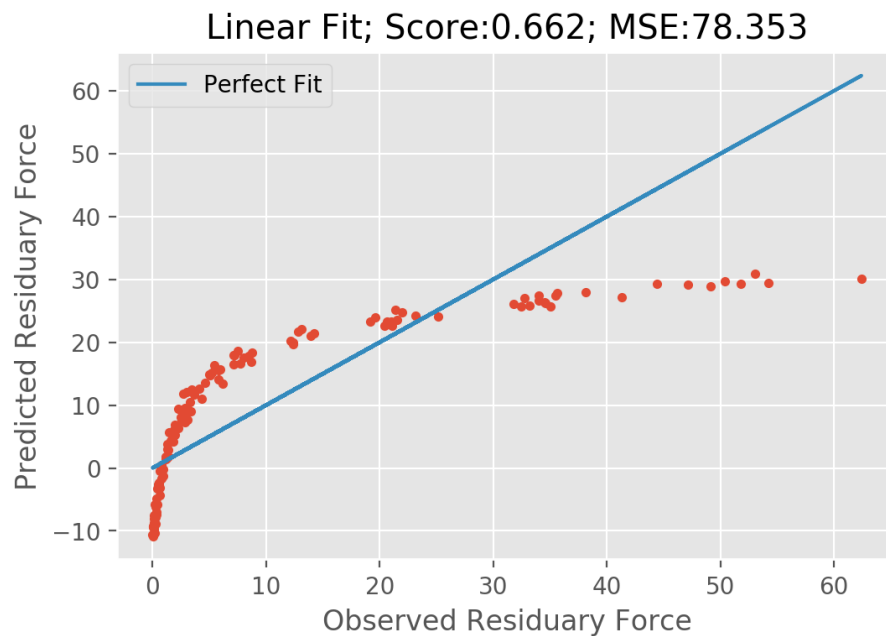
Linear Regression

Linear regression is a simple solution to the problem which should be used if possible to reduce complexity. In this form of machine learning, an offset and a weighting coefficient is assigned for each input factor.

```
lm = LinearRegression()
lm.fit(X_train,y_train)

y_pred = lm.predict(X_test)

score = r2_score(y_test,y_pred)
mse = mean_squared_error(y_test,y_pred)
```

The linear regression is clearly not a good predictor of the results. This was expected as previously mentioned due to the clear non-linear influence of Froude number on the results.

Polynomial Regression

In contrast to linear regression, polynomial regression is significantly more complex. For this type of supervised learning an entire weighted polynomial curve is assigned to each factor with coefficients for every term up to the max degree used. However, if a smallest acceptable maximum degree is used, complexity can be minimized. The following code, tests polynomial regression for several maximum degrees including the linear and constant cases and compares their errors against the test data.

```python
maxDegree = 10
degrees = np.arange(maxDegree)
r2Scores = np.zeros(maxDegree)
meanSquareErrors = np.zeros(maxDegree)

for degree in degrees:
    polynomial_features= PolynomialFeatures(degree=degree)
    x_train_poly = polynomial_features.fit_transform(X_train)

    model = LinearRegression()
    model.fit(x_train_poly, y_train)

    y_test_predict = model.predict(polynomial_features.fit_transform(X_test))

    r2 = r2_score(y_test,y_test_predict)
    r2Scores[degree] = r2

    mse = mean_squared_error(y_test,y_test_predict)
    meanSquareErrors[degree] = mse
```
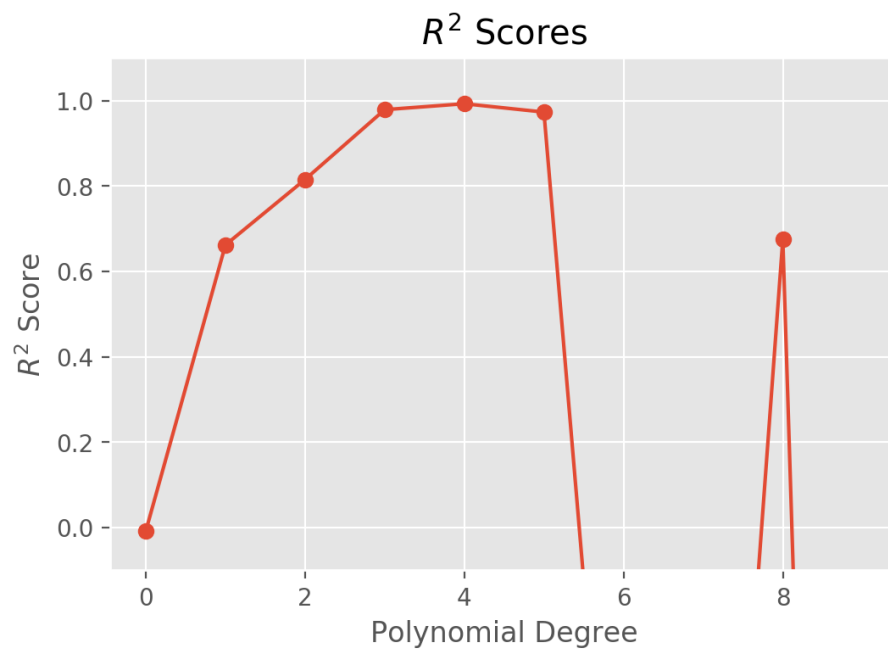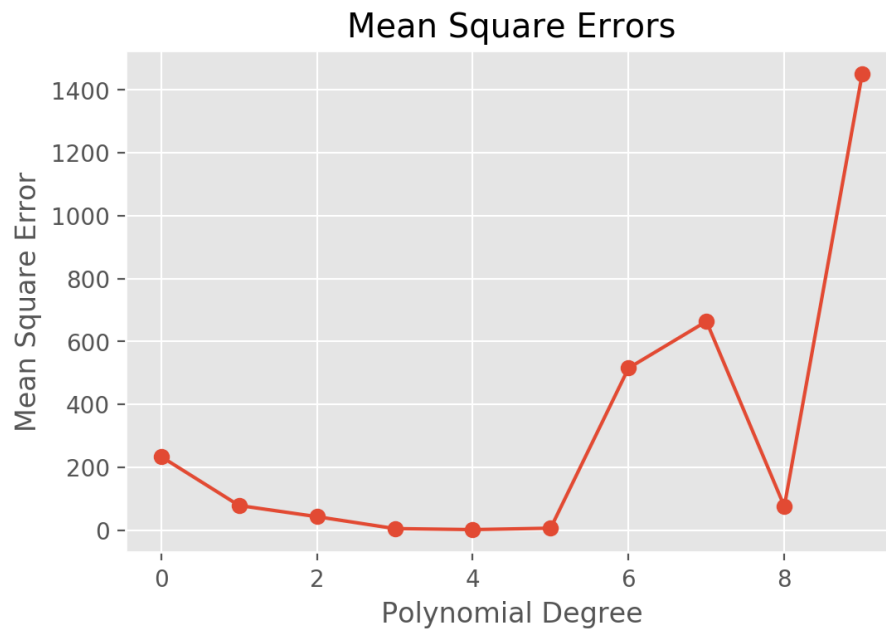
R^2 Scores

A maximum R^2 of 1.0 shows perfect correlation. However, the R^2 score as calculated by python can become arbitrarily low. Very low negative numbers will occur when there is over fitting on the training set. Scores below 0 have been cut off of this figure to facilitate viewing the possible valid options.

Mean square error is an alternative metric for viewing error. Higher scores represent bad fitting. It generally agrees with the R^2 representation of error.

Discussion

Based on the R^2 score and mean square error, several polynomials degrees give great fits. However, when the degree gets too high, over-fitting becomes evident as the test error increases. In order to minimize complexity, we can choose a lowest degree polynomial which acts as a good prediction.

Polyfit: 4 Degrees; Score:0.993; MSE:1.621
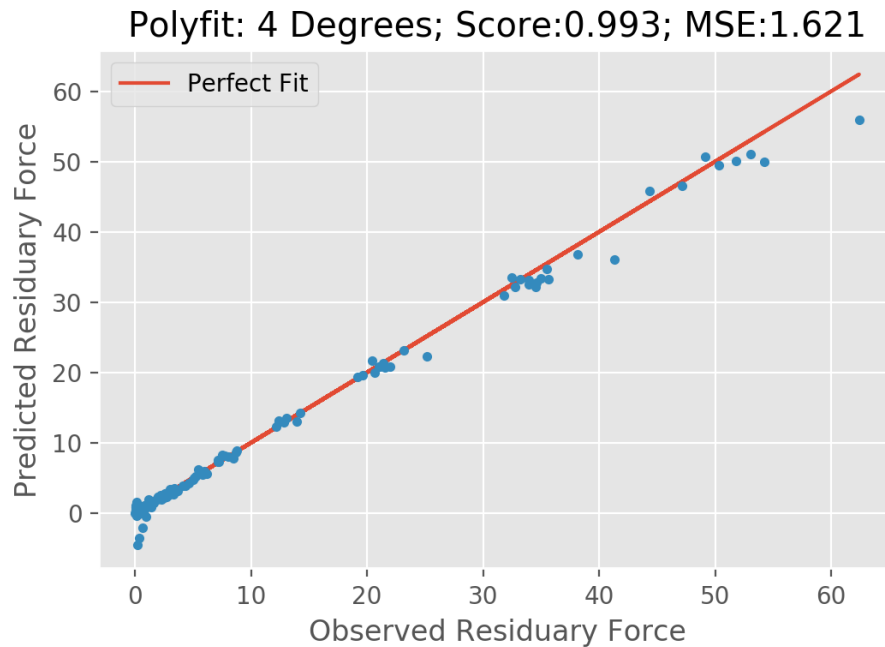
Based on the errors, a polynomial fit of degree 4 shows very good prediction results while also minimizing complexity. This results of such a fit's prediction results can be seen in the figure. The errors may have slightly different values since each new fit starts with a different random seed. The significant difference between 4 and 3 shows that one or more major factors has a significant degree-4 relationship with resistance.

Conclusion

Polynomial regression is a fast and relatively low complexity machine learning algorithm which can be used to make predictions about complicated data. It is paticularly well suited to the results of the Delft Yacht Series experiments because a continuous equation is expected to exist between the variables, but they cannot be easily visualized or analyzed using ordinary methods.

References

Gerritsma, J, R Onnink, and A Versluis. 1981. "Geometry, Resistance and Stability of the Delft Systematic Yacht Hull Series1." *International Shipbuilding Progress* 28 (December): 276–86. doi:10.3233/ISP-1981-2832801.