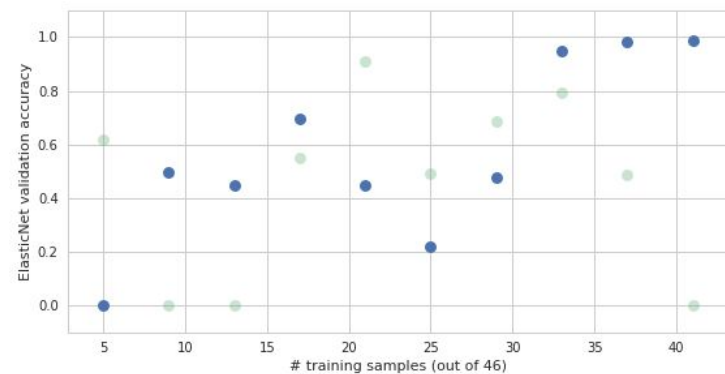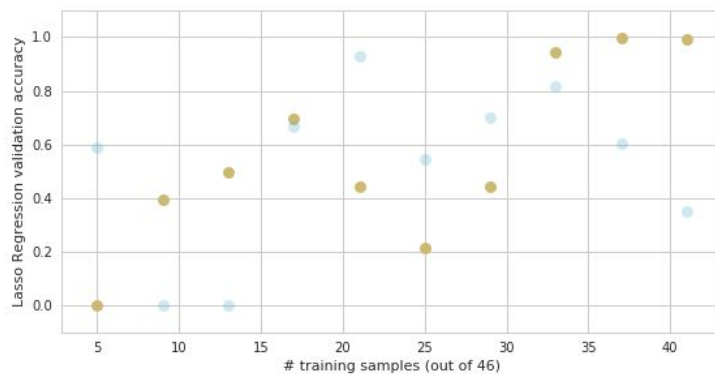**Below:** Linear regression performance (validation accuracy, up to 1.0) over multiple cross-validation splits, with models trained on polynomial (all polynomial combinations of) features, up to degree=|features|

Features = [Time, Load, Cycle #, Heating, Cooling, Tail, Belly]

Linear regression performance on Area prediction, with polynomial features
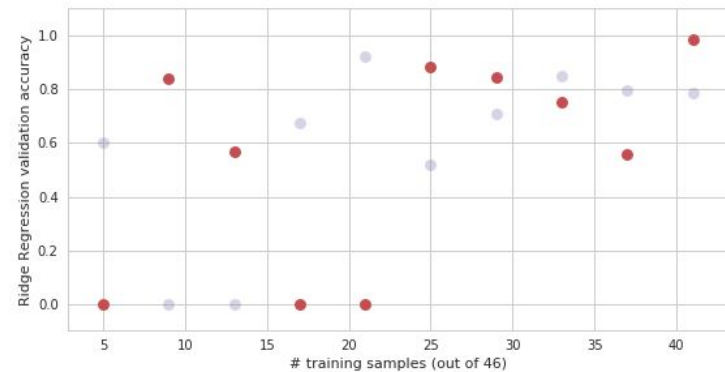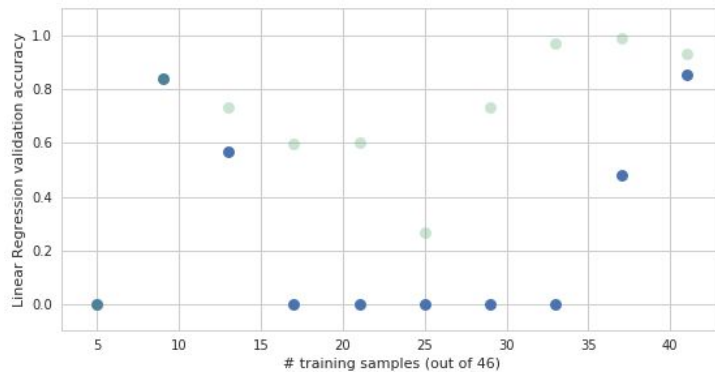(Performance on original features in faded colors)

Below: Linear regression performance (validation accuracy, up to 1.0) over multiple cross-validation splits, with models trained on polynomial features **including skew and kurtosis.**

Validation error of models with a regularization term (e.g. all methods except standard linear regression) noticeably decreases (validation accuracy increases).

Linear regression performance on Area prediction, with polynomial features
(Performance on original features in faded colors)
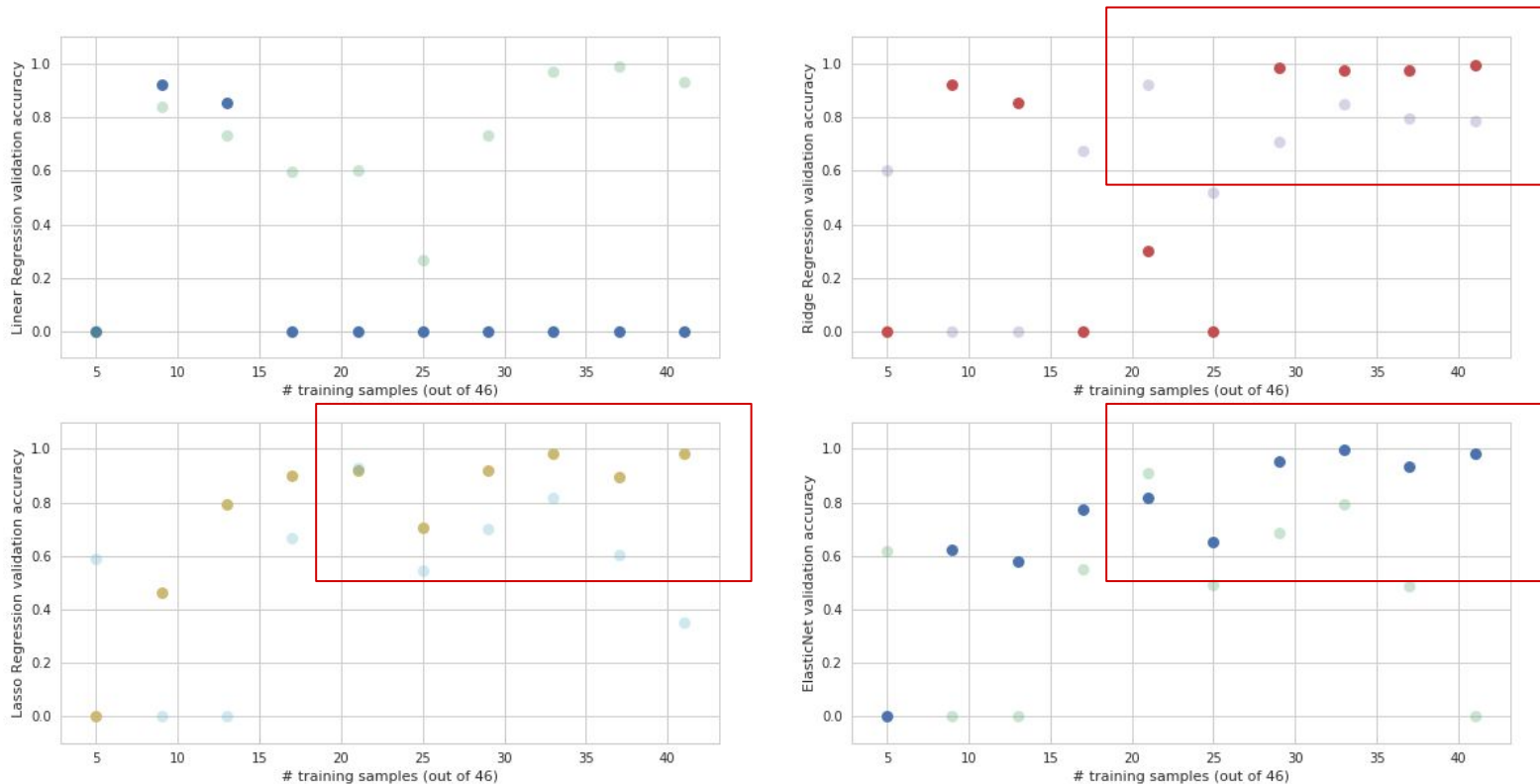
Question: What features are most heavily impacting model performance? (i.e. which features are most strongly weighted by each regression method?)

Answer: `Load, Heating time, Kurtosis`
Then (less strongly) `Cycle #, Cooling time, Skew, Tail (time <10% max load)`

```
In [165]:  for clf in clfs:
               rfe = RFE(clf, n_features_to_select=5)
               rfe.fit(X_train, y_train) # use the most recent training datasets
               print(rfe.ranking_, rfe.support_)
               print(X_train.columns[rfe.support_])
```

```
[3 1 2 1 1 4 1 1] [False  True False  True  True False  True  True]
Index(['Load', 'Heating', 'Cooling', 'Skew', 'Kurt'], dtype='object')
[4 1 1 1 2 3 1 1] [False  True  True  True False False  True  True]
Index(['Load', 'Cycle', 'Heating', 'Skew', 'Kurt'], dtype='object')
[3 1 1 1 1 2 4 1] [False  True  True  True  True False False  True]
Index(['Load', 'Cycle', 'Heating', 'Cooling', 'Kurt'], dtype='object')
[4 1 1 1 1 2 3 1] [False  True  True  True  True False False  True]
Index(['Load', 'Cycle', 'Heating', 'Cooling', 'Kurt'], dtype='object')
```

Time-series analysis using `fbprophet` **library:** Observed (ground truth) Load in black; predicted/trend values in blue. `fbprophet` is a time-series analysis/forecasting library designed to work with seasonal/periodic data.

Unfortunately, a common issue encountered so far in both LSTM and autoregressive approaches is non-periodic cycles, so `fbprophet` fails on the parameters attempted so far.