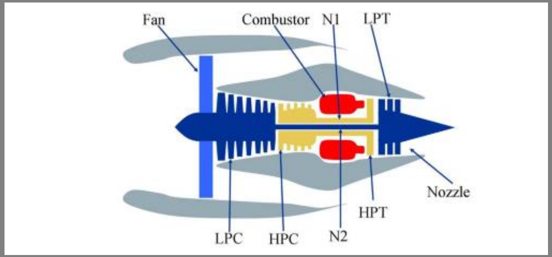
REMAINING USEFUL LIFE PREDICTION OF AN AIRCRAFT ENGINE

PHM08 Challenge - NASA Challenge Data
Prognostics Center of Excellence Data Repository
https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/



PREDICT RUL ON AN AIRCRAFT ENGINE – README FILE

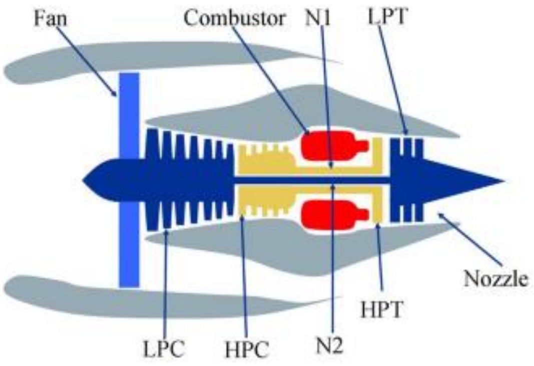
Experimental Scenario
Data sets consists of multiple multivariate time series. Each data set is further divided into training and test subsets. Each time series is from a different engine – i.e., the data can be considered to be from a fleet of engines of the same type. Each engine starts with different degrees of initial wear and manufacturing variation which is unknown to the user. This wear and variation is considered normal, i.e., it is not considered a fault condition. There are three operational settings that have a substantial effect on engine performance. These settings are also included in the data. The data is contaminated with sensor noise.

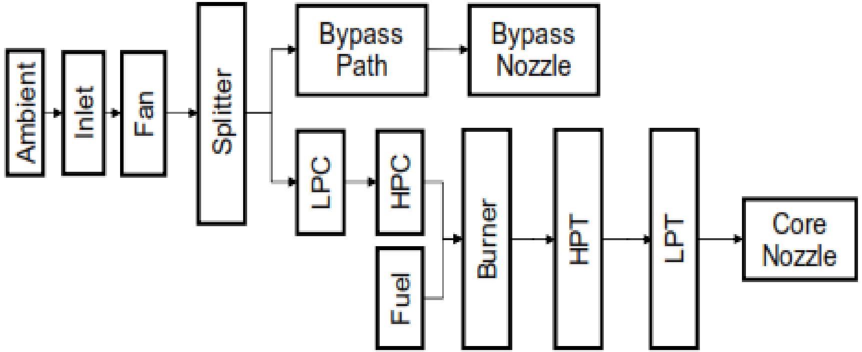
The engine is operating normally at the start of each time series, and develops a fault at some point during the series. In the training set, the fault grows in magnitude until system failure. In the test set, the time series ends some time prior to system failure. The objective of the competition is to predict the number of remaining operational cycles before failure in the test set, i.e., the number of operational cycles after the last cycle that the engine will continue to operate.

The data is provided in a csv file with 26 columns of numbers. Each row is a snapshot of data taken during a single operational cycle, each column is a different variable. The columns correspond to: 1)unit number 2)time, in cycles 3)operational setting 1 4)operational setting 2 5)operational setting 3 6) sensor measurement 1 7) sensor measurement 2 ... 26) sensor measurement 26

Reference: A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation”, in the Proceedings of the Ist International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.

PREDICT REMAINING USEFUL LIFE ON AN AIRCRAFT ENGINE



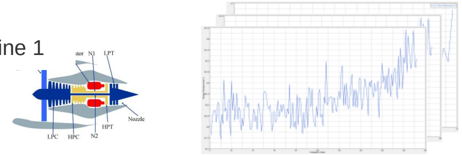


Symbol	*Description*	*Units*
Parameters available to participants as sensor data		
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--

PROBLEM STATEMENT

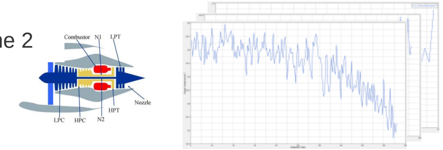
Given

Engine 1



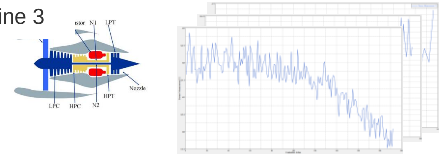
TUL = 192

Engine 2



TUL = 245

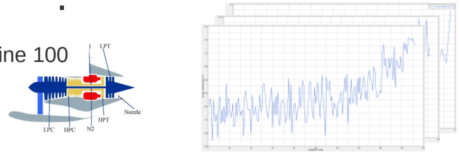
Engine 3



TUL = 135

.
.
.

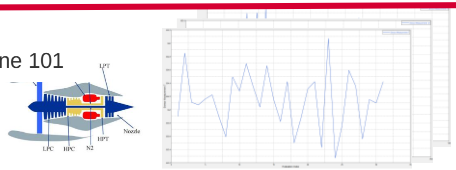
Engine 100



TUL = 212

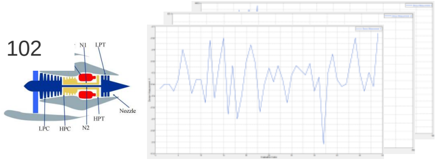
Find

Engine 101



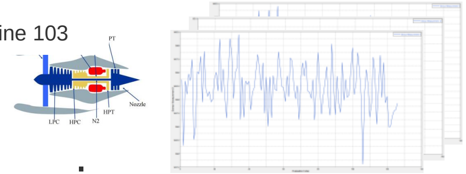
RUL = ? given 31 cycles

Engine 102



RUL = ? given 49 cycles

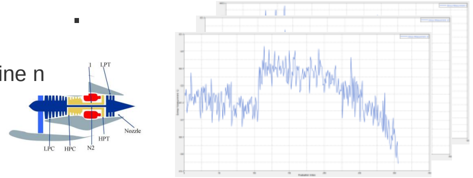
Engine 103



RUL = ? given 112 cycles

.
.
.

Engine n



RUL = ? given 317 cycles

PREDICTION SUMMARY

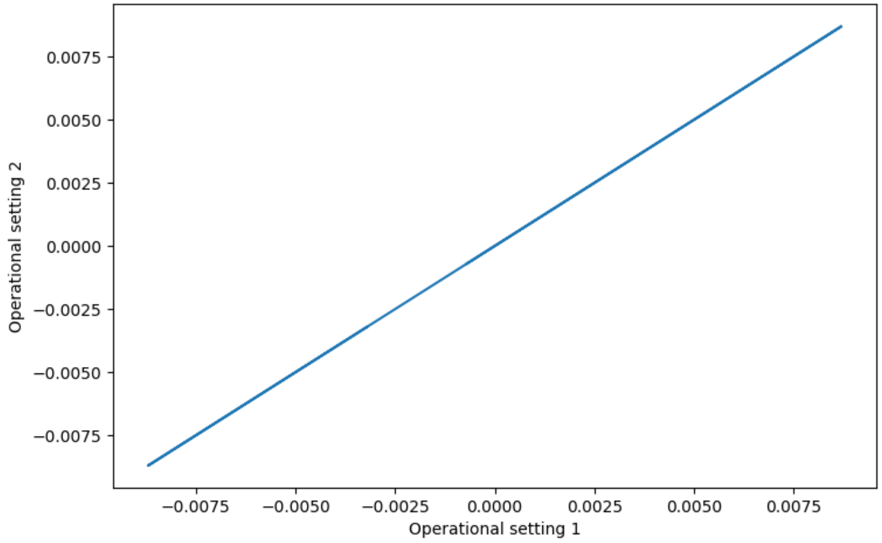
Engine No.	Current Cycle	Actual RUL (cycles)	Predicted RUL (cycles)
101	31	112	139
102	49	98	125
103	126	69	64
104	106	82	80
124	186	20	40
125	48	145	126
134	199	7	34

Please fill in the Predicted RUL column on the left based on your model

DESCRIPTION OF THE SOLUTION PROCESS

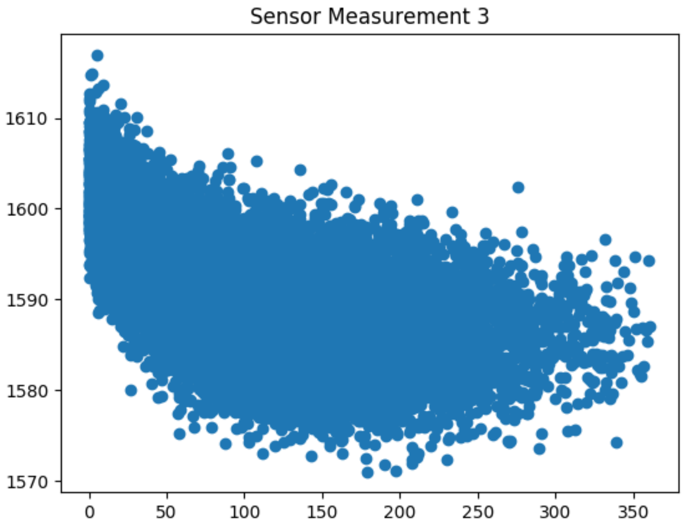
A similarity matching based RUL finding

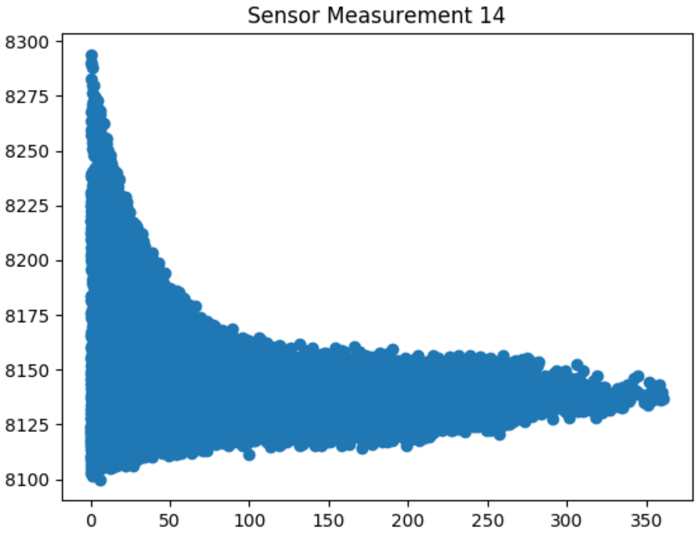
1. There are three operational settings and 26 sensor measurements. The three operational settings were plotted in 3D. The 3rd operational setting does not have any effect since it remains fixed at 100 throughout. Hence, we plotted only the first two operational settings with a view to finding any clustering. However, the following straight line doesn't convey any cluster. Thus, no clusters of operating regio



2. Individual sensor measurement plot was observed taking RUL in the X axis and measurement in the Y axis. Sensors that exhibit a regular pattern were taken into consideration and rest were ignored. For example, the sensor measurements that look like the left one was kept. But, since the right one was irregular, it was ignored. Finally, we taken into consideration the following three sets-

```
sen= [2,3,4,7,11,12,15,20,21]
sen = [2, 3, 4, 7, 11, 12, 15]
sen= [7, 12, 15]- gave the best result
```





3. Each regular sensor measurement was rescaled to the range [0,1]

4. Sensor measurements were transformed by a linear model to find out health index for each cycle-

$$[\beta_0 \quad \beta_1 \quad \beta_2 \ \dots\dots \ \beta_N]\,[1 \quad x_1 \quad x_2 \ \dots\dots x_N] = y$$

In the training phase, we intend to find the linear model. For this purpose, we assigned a linearly decaying 1 to 0 health index to the cycles of an engine, 1 being the healthiest and 0 being near failure

5. The obtained health indices give a time series for each engine. We fitted an exponentially decaying curve to each engine's time series-

$$y = a\,(1 - e^{bx})$$

where x= -RUL of each row of the engine. Thus, we obtain 100 [a,b] sets for 100 engines that are later used in testing phase.

6. Once we are done with training, i.e. finding linear model and exponential model, we want to fit potential exponential graphs to each unit's testing data. As it was done for training, each row of test data is linearly transformed using the model parameters obtained in training, to get health index. Now, we attempt to fit each of the 100 exponential graphs to the given test data and find distance between fitted data and linear model.

7. Thus, we have 100 distances and 100 potential RULs.

8. We remove the outliers by omitting RULs that fall outside the region below-

```
RUL_pred[i]>0 and RUL_pred[i]+r>150 and RUL_pred[i]<190
```

In other words, we are omitting unrealistic, very low, and very high RULs

9. After removing the outliers, we apply the following weighting rule to get final prediction-

```
RUL_final= 4/5*RUL_pos_pred[0]+1/5*RUL_pos_pred[-1]
```

which is a weighting between the least and most RULs.

10. We have taken mean absolute error of prediction to find out how good our model is. Following the above 9 steps, the obtained validation error is ~25 and test error is ~16. Thus, each RUL is predicted within ± 16 of the actual one. Therefore, we believe our model is reasonably correct.

11. Things to play with for furnishing result-

set of sensors

outlier removal criteria

weights in soft voting of RUL

exponential model

Reference: A Similarity-Based Prognostics Approach for Remaining Useful Life Estimation of Engineered Systems, Tianyi Wang, Jianbo Yu, David Siegel, and Jay Lee, 2008 INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT