Battery Capacity Estimation: Algorithm and Implementation

By – Pratinav Kashyap

Objective

- The project aims to find a robust capacity estimation algorithm that can be incorporated into the battery management system of Li-ion battery based electric vehicle.

Introduction

- The battery capacity is essential for the applications such as cell balancing and SoC estimation.
- As the cell ages, the total capacity of the cell degrades. The lithium ion cells have side reactions that occur during charge and discharge cycles. These reactions consume lithium that could have been otherwise used during charge and discharge cycles, and structural deterioration of the electrode active materials that eliminates lithium storage sites.
- The degradation over aging is called capacity fading and the objective of this project is to explore the algorithms that track capacity fade.

The problems with capacity estimation

Sensitivity of capacity to the voltage measurement can be found as:

$$S^{Q}_{v_k} = \frac{Q}{V_k}\frac{dv_k}{dQ} = \frac{Q}{V_k}\frac{d}{dQ}\left(OCV(\frac{Q}{v_k}\frac{d}{dQ}OCv(z_k) + Mh_k + M_0 s_k - \sum_i R_i i_{R_{i_k}} - i_k R_0)\right)$$

The only capacity dependent element in the above equation is the OCV

We evaluate using the chain rule:

$$= \frac{dOCV(z_k)}{dQ} = \frac{\partial OCV(z_k)}{\partial z_k}\frac{dz_k}{dQ}$$

We expand the total derivative:

$$= \frac{dz_k}{dQ} = \frac{dz_{z-1}}{dQ} - \vartheta_{k-1} i_{k-1}\Delta t\frac{d(\frac{1}{Q})}{dQ} = \frac{dz_{z-1}}{dQ} - \frac{\vartheta_{k-1} i_{k-1}\Delta t}{Q^2}$$

Voltage sensitivity to Capacity

The $\frac{dOCV(z_k)}{dz_k}$ is very small during the regular SoC range of operation as seen during the figure 1. Moving on to the $\frac{dz_k}{dQ}$:

We expand the total derivative:

$$= \frac{dz_k}{dQ} = \frac{dz_{z-1}}{dQ} - \boldsymbol{\eta}_{k-1} i_{k-1} \Delta t \frac{d(\frac{1}{Q})}{dQ} = \frac{dz_{z-1}}{dQ} - \frac{\boldsymbol{\eta}_{k-1} i_{k-1} \Delta t}{Q^2}$$

The $\frac{dz_{z-1}}{dQ}$ can be initialized and calculated recursively. The second term has a Q^2 in the denominator and as Q is in terms of coulomb making the term large. This makes $\frac{\boldsymbol{\eta}_{k-1} i_{k-1} \Delta t}{Q^2}$ small. Hence the capacity is

The individual voltage measurements has very little information regarding capacity. We have to use multiple voltage measurements to estimate total capacity accurately.

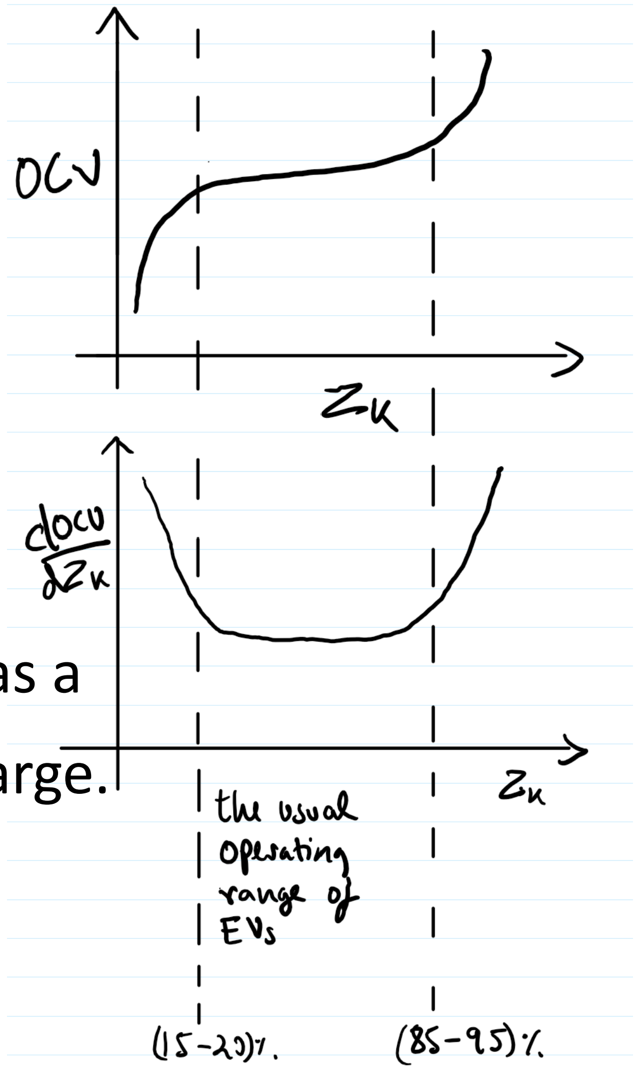


Figure - 1

Using least squares estimation
Why?

- Relation between SoC and Capacity by the coulomb counting formula:

$$z(t_1) = z(t_2) + \frac{1}{Q}\int_{t_1}^{t_2} \frac{\eta_k i(\tau)}{3600}\, d\tau$$

$$\underbrace{\int_{t_1}^{t_2} \frac{\eta_k i(\tau)}{3600}\, d\tau}_{y} = Q\underbrace{(z(t_2) - z(t_1))}_{x}$$

$$y = Q \times x$$

Using least squares estimation
Why?

- The measurement of current (used to calculate Ah) and SoC have noise and we can model them by Δy and Δx respectively.

$$y\ + \Delta y = Q \times (x + \Delta x)$$

- As seen above, the noise in the current measurement and voltage measurement have noise that must be accounted for and using standard liner squares linear regression results provides inaccurate results. (example will be shown during simulation)
- One way could be to use a good SoC estimator for accurate SoC estimate and hence reducing Δx.

Weighted Ordinary total least squares

- This method assumes noise in the current measurement.

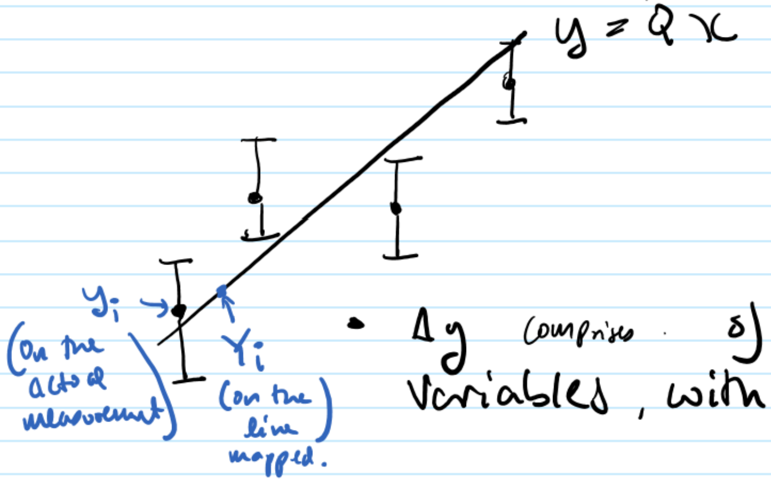
$$y + \Delta y = Q \times x$$

- Δy is considered to have zero mean gaussian random variables. (the variances are assumed to be known).
- The merit/cost function of the WLS approach is:

$$\chi^2_{\mathrm{WLS}} = \sum_{i=1}^{N} \frac{(y_i - Y_i)^2}{\sigma^2_{y_i}} = \sum_{i=1}^{N} \frac{\left(y_i - \hat{Q}x_i\right)^2}{\sigma^2_{y_i}}.$$

Where N is the sample length. Y_i is the point on the line $Q \times x$

Weighted Ordinary total least squares



- Δy comprises of zero mean Gaussian random variables, with known variances $\sigma_{y_i}^2$

The cost function will be differentiated equated to 0 and solved for Q:

$$\frac{\partial \chi^2_{\text{WLS}}}{\partial \hat{Q}} = -2\sum_{i=1}^{N} \frac{x_i(y_i - \hat{Q}x_i)}{\sigma_{y_i^2}} = 0$$

$$\text{Q}\sum_{i=1}^{N} \frac{x_i^2}{\sigma_{y_i}^2} = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_{y_i}^2}$$

Weighted Ordinary total least squares

- $$Q\sum_{i=1}^{N}\frac{x_i^2}{\sigma_{y_i}^2} = \sum_{i=1}^{N}\frac{x_i y_i}{\sigma_{y_i}^2}$$
- $$Q = \frac{\sum_{i=1}^{N}\frac{x_i y_i}{\sigma_{y_i}^2}}{\sum_{i=1}^{N}\frac{x_i^2}{\sigma_{y_i}^2}}$$
- $$c_{1,n} = \sum_{i=1}^{n}\frac{x_i^2}{\sigma_{y_i}^2} \quad c_{2,n} = \sum_{i=1}^{n}\frac{x_i y_i}{\sigma_{y_i}^2}$$
- $$c_{1,n} = c_{1,n-1} + \frac{x_n^2}{\sigma_{y_n}^2} \text{ and } c_{2,n} = c_{2,n-1} + \frac{x_n y_n}{\sigma_{y_n}^2}$$
- $$Q_n = \frac{c_{2,n}}{c_{1,n}}$$
- On implementing the algorithm $c_{1,n}$ and $c_{2,n}$ are initialized with 0 if the initial capacity is unknown. If the nominal capacity is known then in a range of SoC, i.e, 1, we can define $x_0 = 1$ and $y_0 = Q_{nom}$. The $c_{1,0} = \frac{1}{\sigma_{y_0}^2}$ and $c_{2,0} = \frac{Q_{nom}}{\sigma_{y_0}^2}$

Weighted Ordinary total least squares
"Account for the fading memory"

- Fading Memory: As discussed, the capacity fades over time and if measurement does not account for this phenomenon then a bias will occur in forming the regression line. The cost function considers the capacity as a constant quantity, to add more importance to the most recent measurement of capacity, we include a weighting factor (γ) in the cost function:

$$\sum_{i=1}^{N} \gamma^{n-i} \frac{(y_i - Qx_i)^2}{\sigma_{y_i}^2}$$

Weighted Ordinary total least squares

"Account for the fading memory"

- $$Q = \frac{\sum_{i=1}^{N} \gamma^{n-i} \frac{x_i y_i}{\sigma_{y_i}^2}}{\sum_{i=1}^{N} \gamma^{n-i} \frac{x_i^2}{\sigma_{y_i}^2}}$$
- $$c_{1,n} = \sum_{i=1}^{n} \gamma^{n-i} \frac{x_i^2}{\sigma_{y_i}^2} \quad c_{2,n} = \sum_{i=1}^{n} \gamma^{n-i} \frac{x_i y_i}{\sigma_{y_i}^2}$$
- $c_{1,n} = \gamma c_{1,n-1} + \frac{x_n^2}{\sigma_{y_n}^2}$ and $c_{2,n} = \gamma c_{2,n-1} + \frac{x_n y_n}{\sigma_{y_n}^2}$
- $$Q_n = \frac{c_{2,n}}{c_{1,n}}$$

Hessian for WTLS

- The double derivative (Hessian) of the cost function is found to be:
- H = $\frac{\partial^2 \chi^2_{WLS}}{\partial Q_n^2} = 2 \sum_{i=1}^{n} \gamma^{n-i} \frac{x_i^2}{\sigma_{y_i}^2} = 2c_{1,n}$
- The hessian shall be used later for error bounds calculation

Properties of Weighted Ordinary total least squares

- Simple to compute as it is recursive
- Fading memory can be added easily

Problem with Weighted Ordinary least squares

- The noise in SoC estimate is ignored that leads to a bias.
- We shall consider the noise of both parameters in “Total Least Squares” based approach

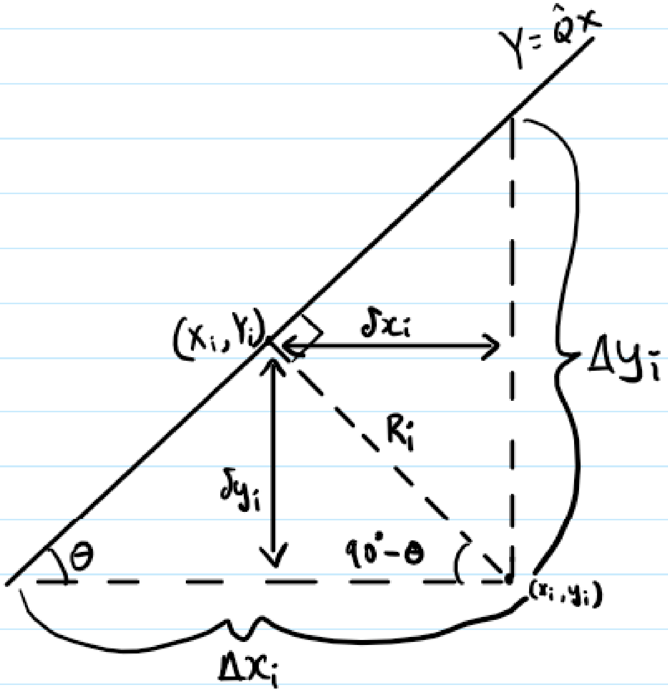
Approximate Weighted Total Least Squares (AWTLS)

- Here we consider noises on both the parameters:

The AWTLS cost function is defined as:

$$\chi^2_{AWTLS} = \sum_{i=1}^{N} \frac{\partial {x_i}^2}{\sigma^2_{x_i}} + \frac{\partial {y_i}^2}{\sigma^2_{y_i}}$$

AWLTS



$\theta = \tan^{-1} \hat{Q}_n \qquad \tan\theta = Q$

$$R_i = \Delta y_i \cos\theta = \frac{\Delta y_i}{\sqrt{1+\tan^2\theta}} = \frac{\Delta y_i}{\sqrt{1+\hat{Q}_n^2}}$$

$$\delta x_i = R_i \cos\theta = \frac{\Delta y_i}{\sqrt{1+Q_n^2}} \times \frac{1}{\sqrt{1+Q_n^2}}$$

$$\delta y_i = R_i \sin\theta = \frac{\Delta y_i}{\sqrt{1+Q_n^2}} \times \frac{Q_n}{\sqrt{1+Q_n^2}}$$

(where $\rightarrow \Delta y = y_i - Q_n x_i$)

$$\frac{\sin\theta}{\cos\theta} = Q$$

$$\frac{\sin^2\theta}{\cos^2\theta} = Q^2$$

$$1 - \cos^2\theta = Q^2\cos^2\theta$$

$$1 = (1 + Q^2)\cos^2\theta$$

$$\cos\theta = \frac{1}{\sqrt{1+Q_n^2}}$$

$$\sin^2\theta = Q_n^2 - Q_n^2 \sin\theta$$

$$\sin\theta = \frac{Q_n}{\sqrt{1+Q_n^2}}$$

Approximate Weighted Total Least Squares (AWTLS)

$$\chi^2_{AWTLS} = \sum_{i=1}^{n} \frac{\partial {x_i}^2}{\sigma^2_{x_i}} + \frac{\partial {y_i}^2}{\sigma^2_{y_i}}$$

$$\delta x_i^2 = \left(\frac{\Delta y_i^2}{1+\widehat{Q}_n^2}\right)\left(\frac{\widehat{Q}_n^2}{1+\widehat{Q}_n^2}\right)$$

$$\delta y_i^2 = \left(\frac{\Delta y_i^2}{1+\widehat{Q}_n^2}\right)\left(\frac{1}{1+\widehat{Q}_n^2}\right)$$

On Replacing the values into the cost function

$$\chi^2_{AWTLS} = \sum_{i=1}^{n} \frac{(y_i - Q_n x_i)^2}{(1+Q_n)^2}\left(\frac{Q^2{}_n}{\sigma^2_{x_i}} + \frac{1}{\sigma^2_{y_i}}\right)$$

Approximate Weighted Total Least Squares (AWTLS)

We find the Jacobian of the cost function as:

$$\frac{\partial \chi^2_{\text{AWTLS}}}{\partial \widehat{Q}_n} = \frac{2}{(\widehat{Q}_n^2+1)^3} \sum_{i=1}^{n} \widehat{Q}_n^4 \left(\frac{x_i y_i}{\sigma_{x_i}^2}\right) + \widehat{Q}_n^3 \left(\frac{2x_i^2}{\sigma_{x_i}^2} - \frac{x_i^2}{\sigma_{y_i}^2} - \frac{y_i^2}{\sigma_{x_i}^2}\right) + \widehat{Q}_n^2 \left(\frac{3x_i y_i}{\sigma_{y_i}^2} - \frac{3x_i y_i}{\sigma_{x_i}^2}\right) + \widehat{Q}_n \left(\frac{x_i^2 - 2y_i^2}{\sigma_{y_i}^2} + \frac{y_i^2}{\sigma_{x_i}^2}\right) + \left(\frac{-x_i y_i}{\sigma_{y_i}^2}\right)$$

Where we have already seen $c_{1,n} = \sum_{i=1}^{n} \frac{x_i^2}{\sigma_{y_i}^2}$ $c_{2,n} = \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_{y_i}^2}$ the remaining can be defined as:

$$c_{4,n} = \sum_{i=1}^{n} \frac{x_i^2}{\sigma_{x_i}^2}, \qquad c_{5,n} = \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_{x_i}^2}, \qquad c_{6,n} = \sum_{i=1}^{n} \frac{y_i^2}{\sigma_{x_i}^2}.$$

we can define $x_0 = 1$ and $y_0 = Q_{nom}$

$$c_{1,0} = 1/\sigma_{y_0}^2 \qquad c_{4,0} = 1/\sigma_{x_0}^2$$
$$c_{2,0} = Q_{\text{nom}}/\sigma_{y_0}^2 \qquad c_{5,0} = Q_{\text{nom}}/\sigma_{x_0}^2$$
$$c_{3,0} = Q_{\text{nom}}^2/\sigma_{y_0}^2 \qquad c_{6,0} = Q_{\text{nom}}^2/\sigma_{x_0}^2,$$

After accounting for "fading memory" like before:

$$c_{1,n} = c_{1,n-1} + x_n^2/\sigma_{y_n}^2 \qquad c_{4,n} = c_{4,n-1} + x_n^2/\sigma_{x_n}^2$$
$$c_{2,n} = c_{2,n-1} + x_n y_n/\sigma_{y_n}^2 \qquad c_{5,n} = c_{5,n-1} + x_n y_n/\sigma_{x_n}^2$$
$$c_{3,n} = c_{3,n-1} + y_n^2/\sigma_{y_n}^2 \qquad c_{6,n} = c_{6,n-1} + y_n^2/\sigma_{x_n}^2.$$

We add the forgetting factor

$$c_{1,n} = \gamma c_{1,n-1} + \frac{x_n^2}{\sigma_{y_n}^2} \qquad c_{4,n} = \gamma c_{4,n-1} + \frac{x_n^2}{\sigma_{x_n}^2}$$

$$c_{2,n} = \gamma c_{2,n-1} + \frac{x_n y_n}{\sigma_{y_n}^2} \qquad c_{5,n} = \gamma c_{5,n-1} + \frac{x_n y_n}{\sigma_{x_n}^2}$$

$$c_{3,n} = \gamma c_{3,n-1} + \frac{y_n^2}{\sigma_{y_n}^2} \qquad c_{6,n} = \gamma c_{6,n-1} + \frac{y_n^2}{\sigma_{x_n}^2}$$

AWTLS
Solving for Q

We set the Jacobian to 0 to minimize the cost function:

$$c_5 Q^4 + (2c_4 - c_1 - c_6)Q^3 + (3c_2 - 3c_5)Q^2 + (c_1 - 2c_3 + c_6)\ Q^2 \text{-}\ c_2\ = 0$$

Dr. Plett introduces a approximate weighted TLS solution [1] in recursive form as:

$$\chi^2_{\text{AWTLS}} = \frac{1}{(\widehat{Q}_n^2 + 1)^2}\left(c_4\widehat{Q}_n^4 - 2c_5\widehat{Q}_n^3 + (c_1 + c_6)\widehat{Q}_n^2 - 2c_2\widehat{Q}_n + c_3\right)$$

Procedure:

- We solve the Jacobian set to 0 for Q
- We discard all the imaginary and negative values and retain the positive values.
- We feed the real and positive roots to the above approximate cost function and find the root that gives us the minimum value of the cost function and keep that capacity as the estimated capacity.

Hessian for AWTLS

- The double derivative (Hessian) of the cost function is found to be:

$$H = \frac{\partial^2 \chi^2_{AWTLS}}{\partial Q_n^2} = \frac{2}{(Q_n^2+1)^4}\Big(-2c_5Q_n^5 + (3c_1 - 6c_4 + 3c_6)Q_n^4 + (-12c_2 + 16c_5)Q_n^3 + (-8c_1 + 10c_3 + 6c_4 - 8c_6)Q_n^2 + (12c_2 - 6c_5)Q_n + (c_1 - 2c_3 + c_6)\Big)$$

- The hessian shall be used later for error bounds calculation

A brief about the WTLS method modified to PTLS

- Weighted least squares method is an improvement on the WLS/OLS method where the assumption is made that the variance in the SoC and the y (Ampere hour) measurement is proportional:

$$\sigma_{x_i}^2 = k\,\sigma_{y_i}^2 \quad \text{where k is a proportionality constant}$$

- This method is called proportional total-least-squares (PTLS) and the cost function for the same is:

$$\chi_{PTLS}^2 = \sum_{i=1}^{n} \frac{(y_i - Q_n x_i)^2}{\left(\underbrace{Q_n^2\,\sigma_{x_i}^2 + \sigma_{y_i}^2}_{\chi_{WTLS}^2}\right)^2} = \sum_{i=1}^{n} \frac{(y_i - Q_n x_i)^2}{\sigma_{y_i}^2\,(Q_n^2\,k^2+1)^2}$$

Problems of AWTLS

- The original Weighted total least squares method was used to derive the AWLTS and in doing do a condition where the deviation of SoC and ampere hour measurement are proportional:

$\sigma_{x_i}^2$ = $k\ \sigma_{y_i}^2$ where k is a proportionality constant

The AWLTS cost function does not form the similar cost function when the above assumption is placed. Hence a few changes have been suggested by Dr. Plett [1] to recover from this discrepancy.

Improvements…..

- A proportionality factor is included in the solution of the AWLTS by considering a scaled measurement $y_i = k * y_i$ hence $\sigma_{y_0}^2 = k^2\sigma_{y_0}^2$
- On deriving the AWLTS again with the change:

Estimated Q = Q/k

Hessian H = k^2*H

“Goodness of model fit”

- The goodness of fit is found using the incomplete gamma function in Matlab.
- For the WLTS method degree of freedom is calculated based on the parameters mentioned in the paper “*Recursive approximate weighted total least squares estimation of battery cell total capacity*”

Error boundaries (Confidence Intervals)

- The variance is also estimated for the estimated capacity and a error bound is computed. This gives us an idea about the accuracy of the estimate.
- The Hessian of the cost function is calculated by replacing the capacity estimate found and the following identity is used to find the variance:
- $\sigma_Q^2 = 2\ (\frac{\partial^2\chi^2}{\partial Q^2})^{-1}$ for WTLS
- $\sigma_Q^2 = 2\ (\frac{\partial^2\chi^2}{\partial Q^2} * k\hat{}2)^{-1}$ for AWTLS

Simulation Consideration

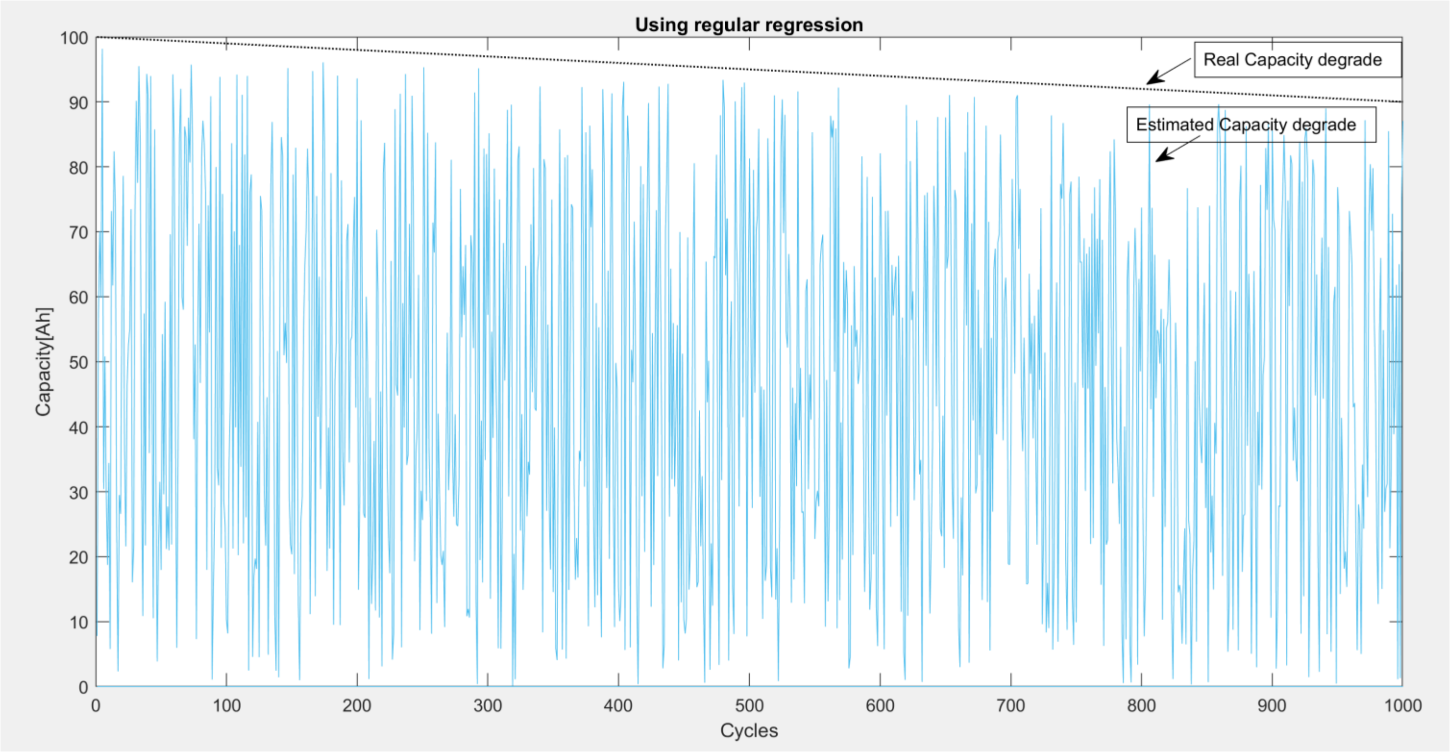
- To have control over application of the algorithm we shall generate samples of SoC and current.
- We shall simulate for a battery electric vehicle condition:
- We can get the end point SoC when fully charged $z(t_1)$ or $z(t_2)$ will be known precisely. The problem is that we cannot have a well defined updates and we can take updates only when we have the vehicle charging. For purposes of simulation, we consider the updates as a random variable with a gaussian distribution with mode of 0.5 and standard deviation of 0.6.
- A max SoC range is taken as 80%
- We consider a decade rate of 0.01 with a initial capacity of 100Ah
- We assume that the initial capacity is the nominal capacity
- A forgetting factor of 0.98 is taken
- Under the assumption that EKF is used to estimate the SoC, a SoC noise of 0.01 is taken
- A total of 1000 (n) sets of samples are considered.

Regular Regression

- The Simulated capacity degradation:

Capacity degrades from 100Ah to 90Ah in 1000 cycles

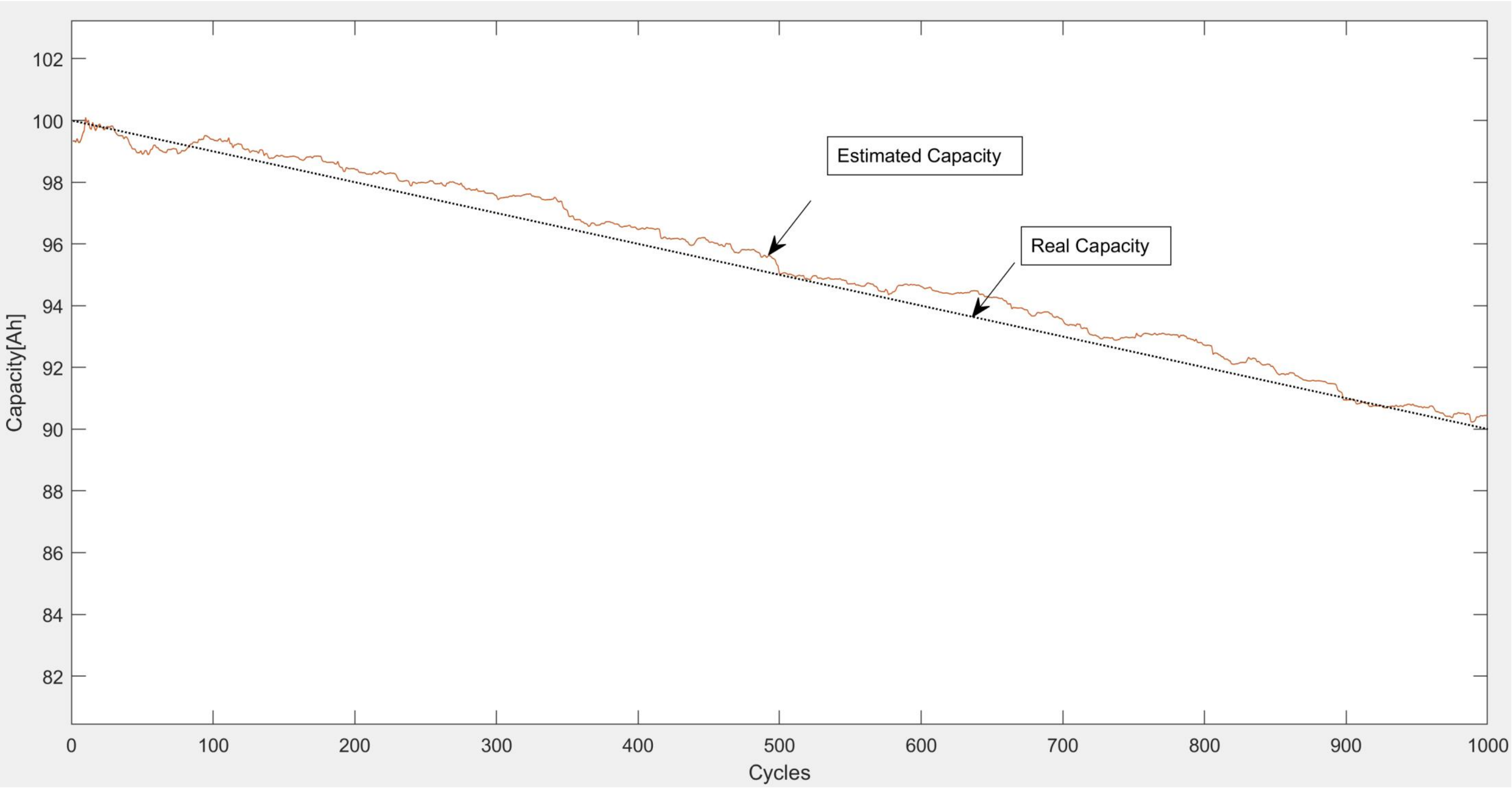
If regular regression is used (without considering noise). The following result is obtained:



$$Q = \frac{y}{x}$$ regular regression

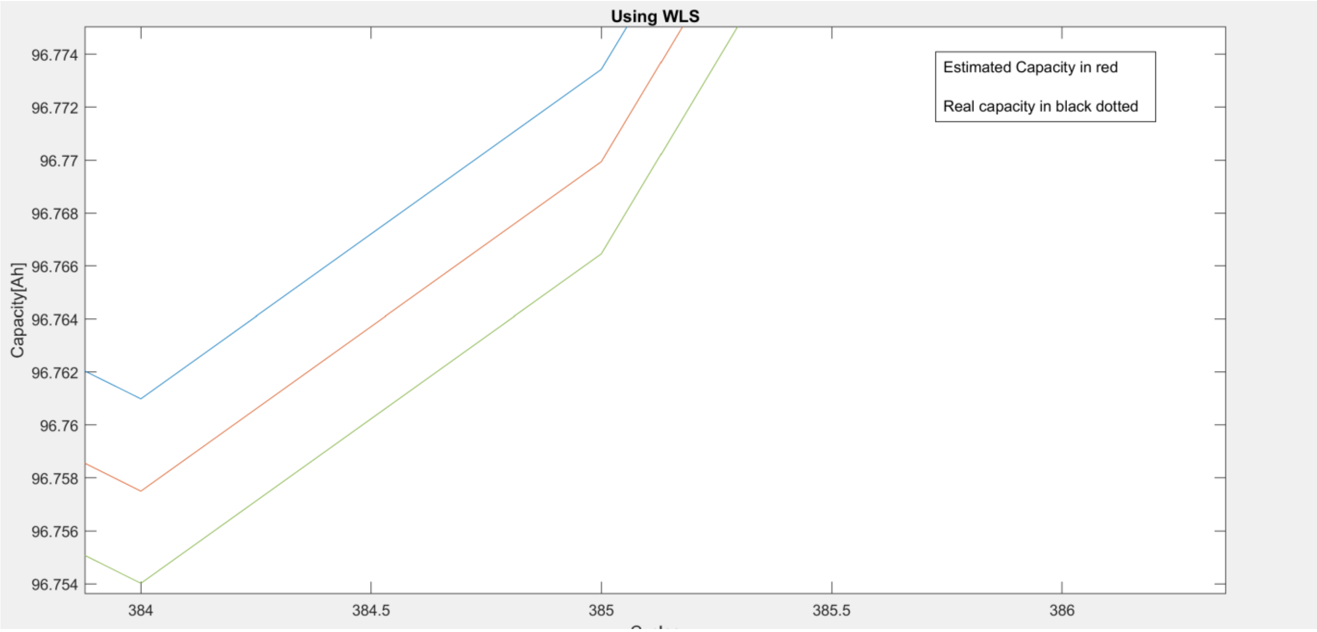
The noise in x and y measurements generates unusable data of capacity

With WLS Method

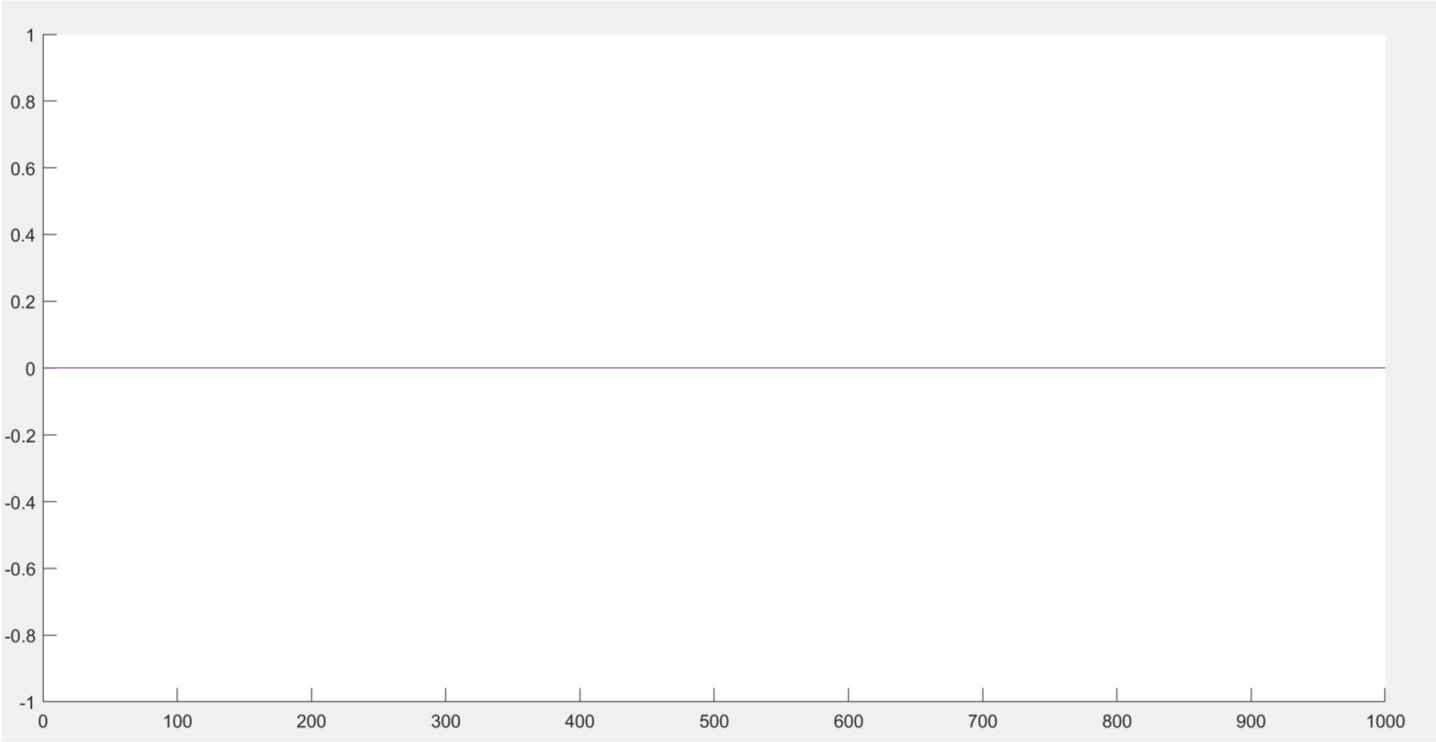


Comments and disadvantage

- WLS method was simple and accurate when applied to a data set of capacity degradation but on zooming in we see something odd.
- The error bounds show unreal confidence and hence we cannot be sure about the error of the method as it shows confidence intervals that are not relable.

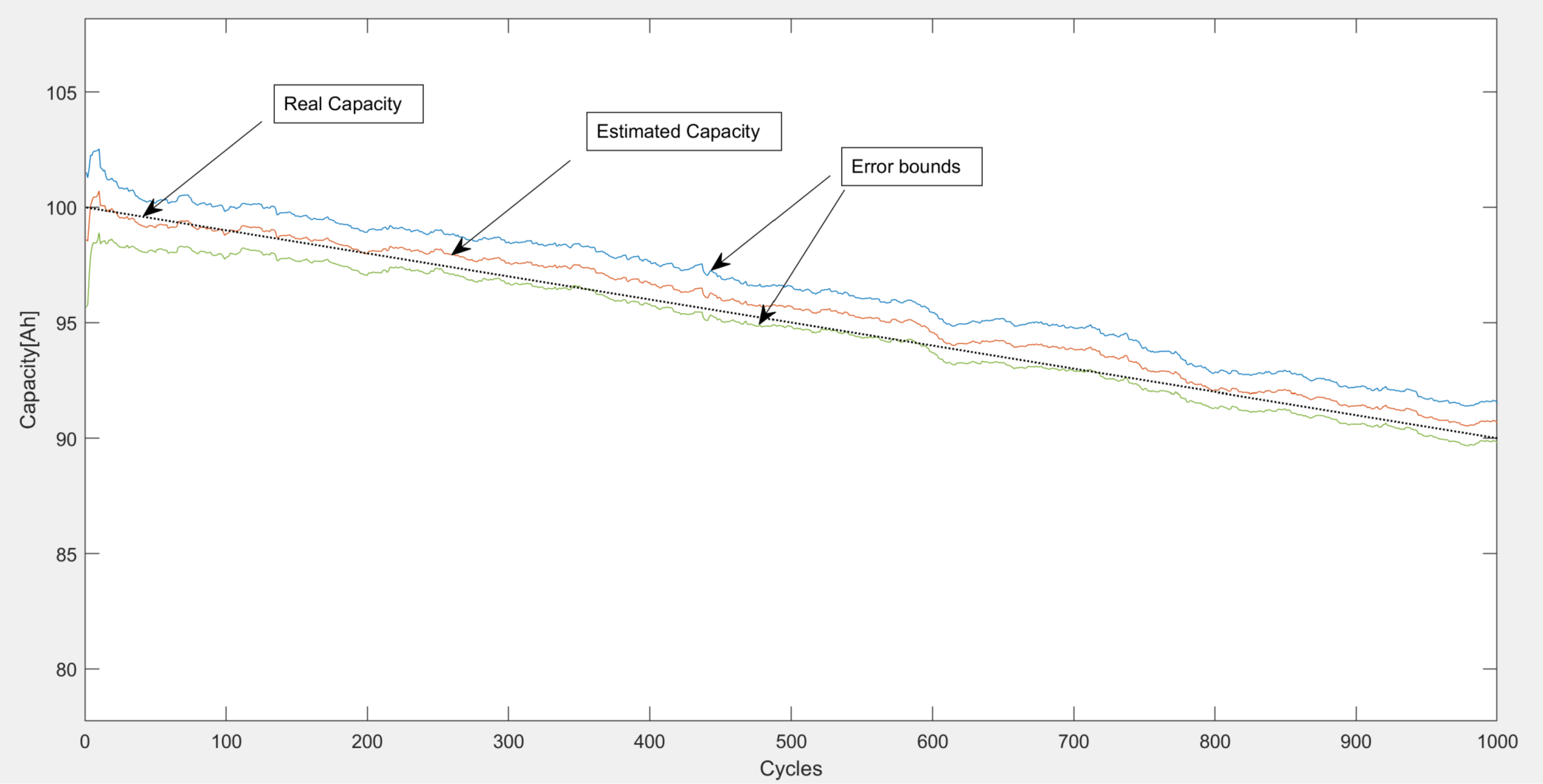


Goodness of fit for WLS



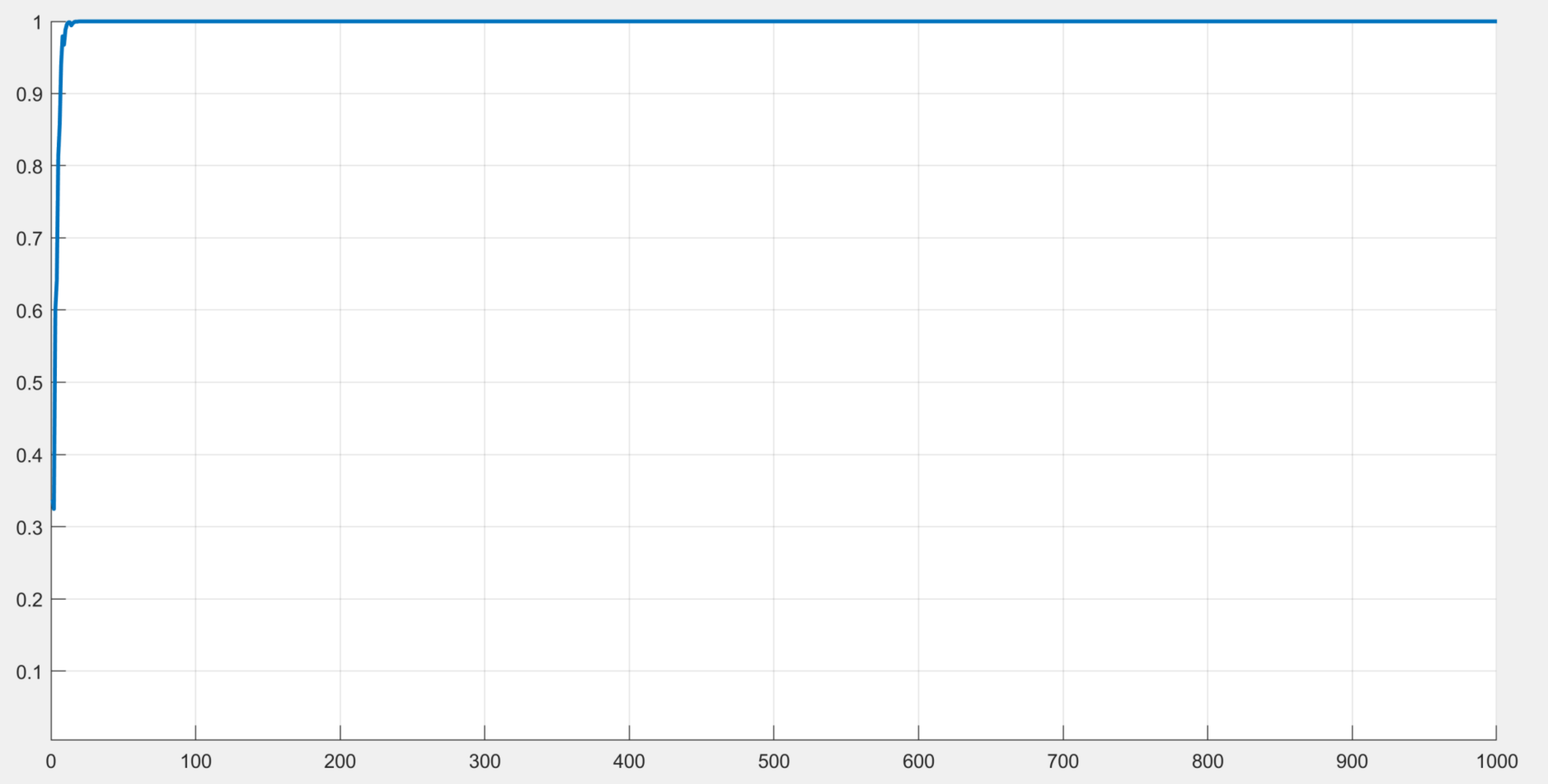
As expected, the goodness of fit shows 0. We can infer that WLS is not the best approach

With AWTLS Method



The Capacity seems to be estimated well and the error bounds are well distributed

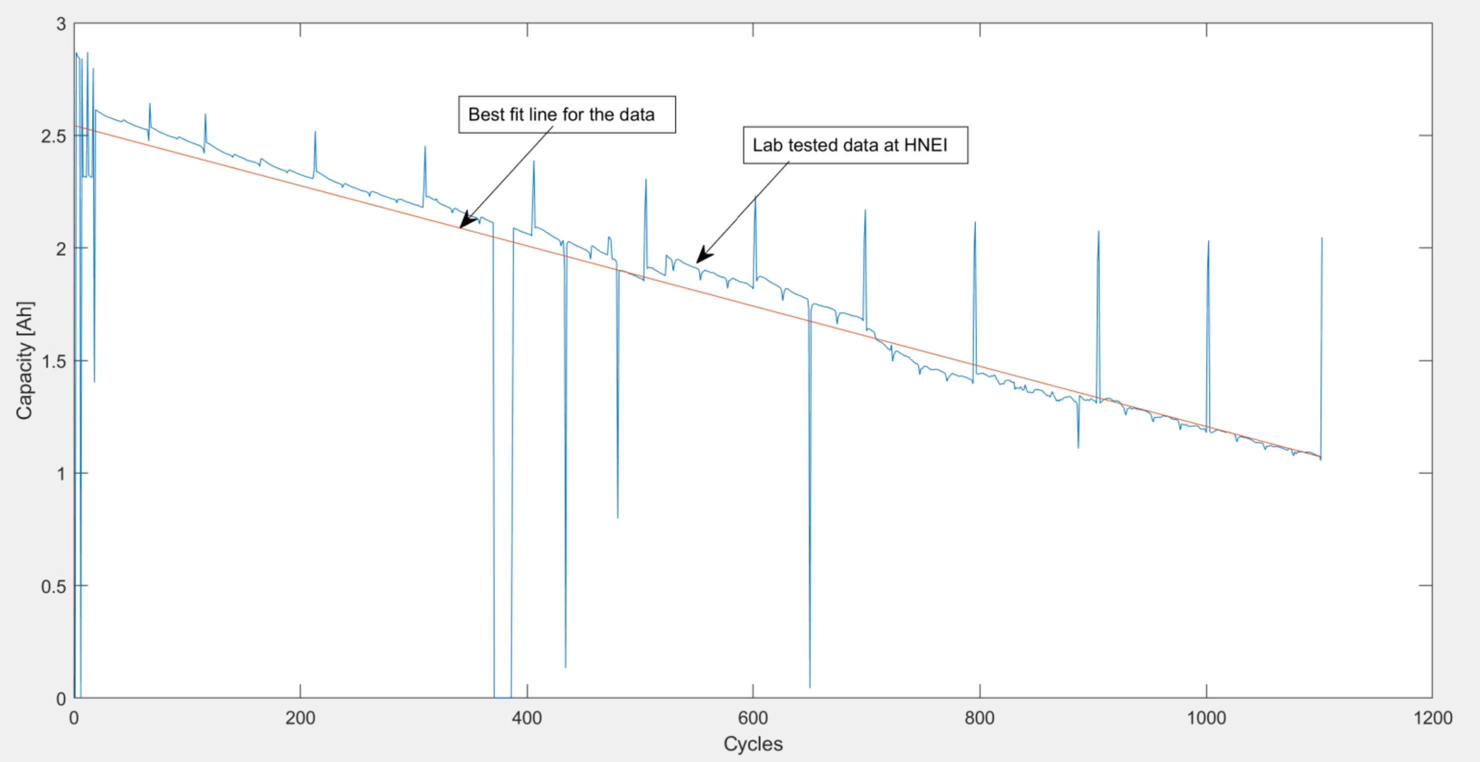
Goodness of Fit for AWTLS



As expected, the goodness of fit shows 1 as the error bounds are fair and the estimate is accurate. We can infer that WLS is not the best approach

Results with Real battery degrading data

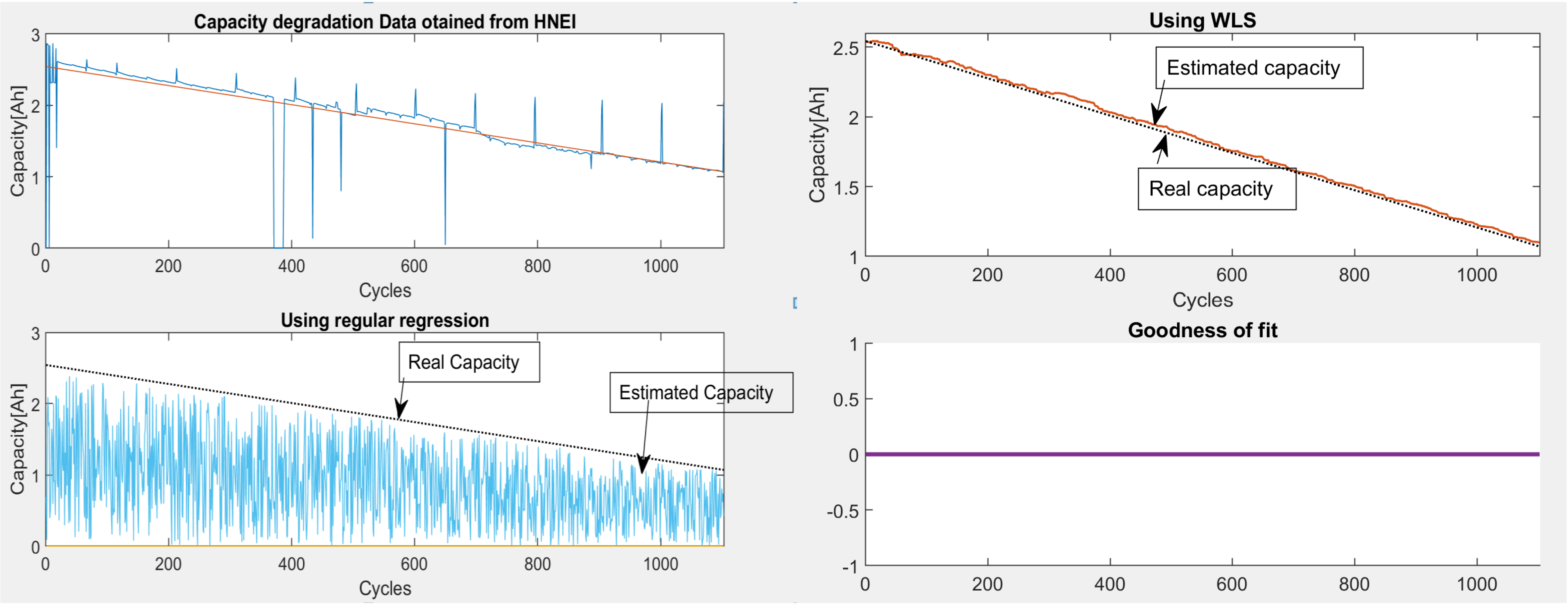
- A data set was taken from https://www.batteryarchive.org/index.html and a data set collected by the Hawaii Natural Energy Institute (HNEI) has been used as the real capacity degrading data [2].
- A 18650 Li-ion NMC cell was cycled for 1113 cycles and the capacity degrade was lab tested



The date obtained was noisy and hence a best fit line was generated to implement the algorithm.

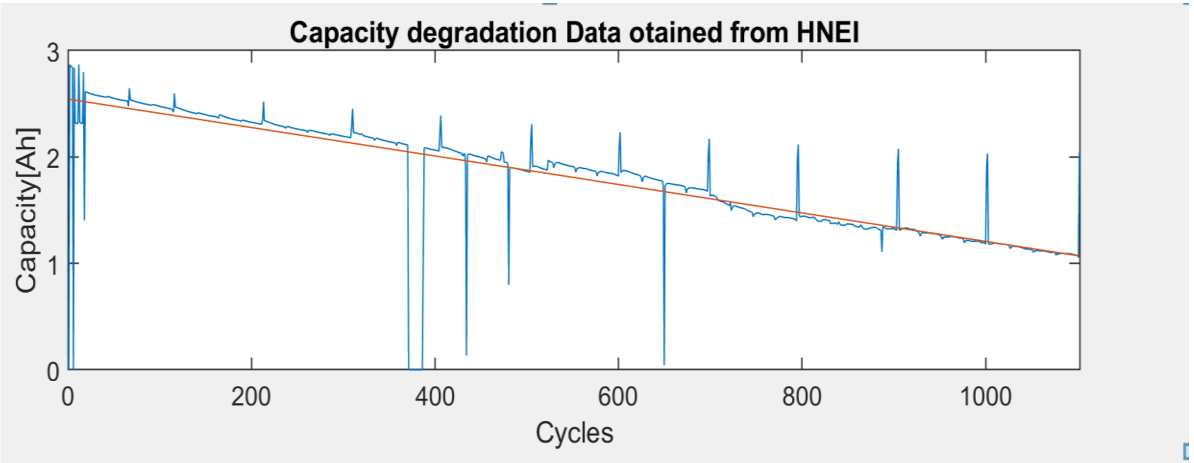
WLS on HNEI data

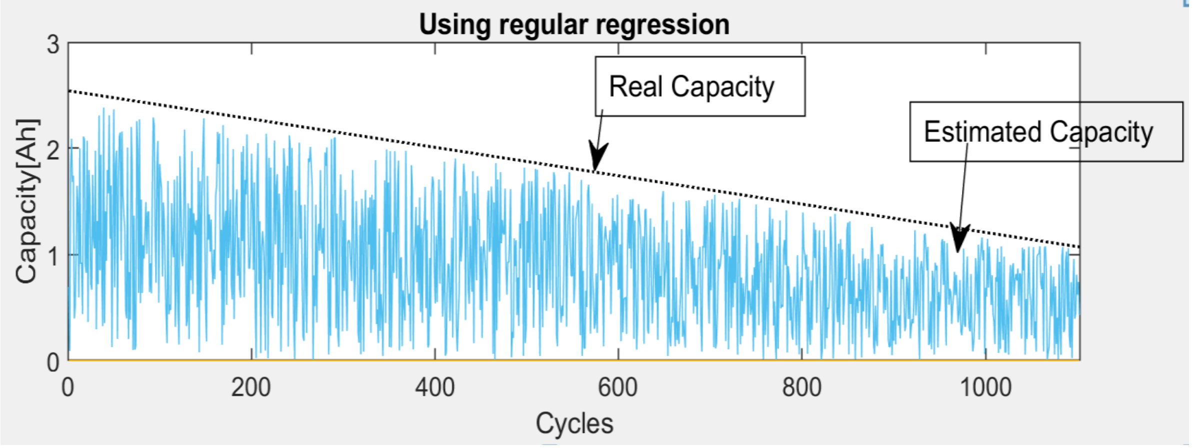
The lab test data was fed into the algorithm with a random SoC change across measurement intervals. Forgetting factor was taken as 0.93.

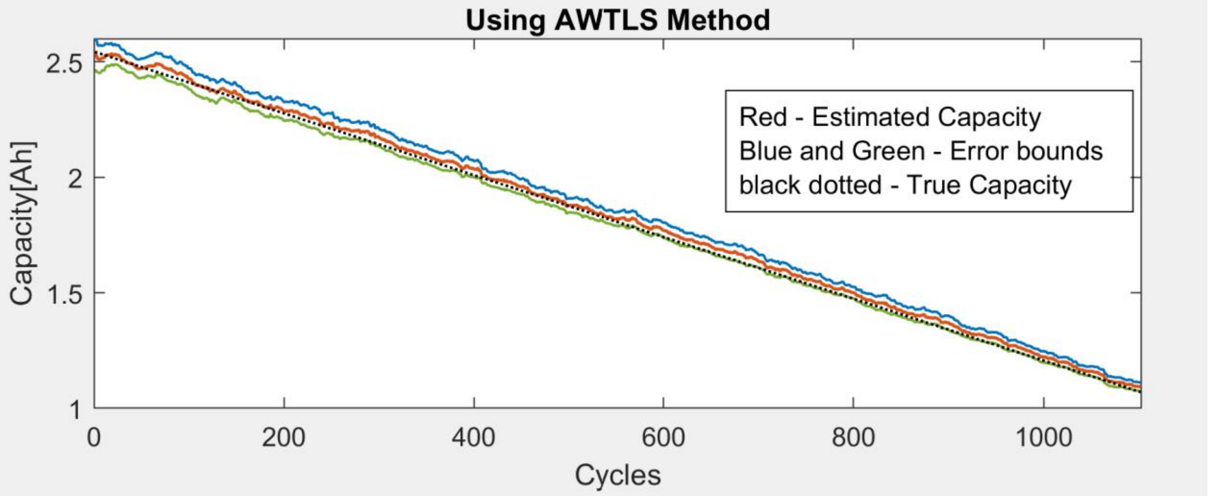


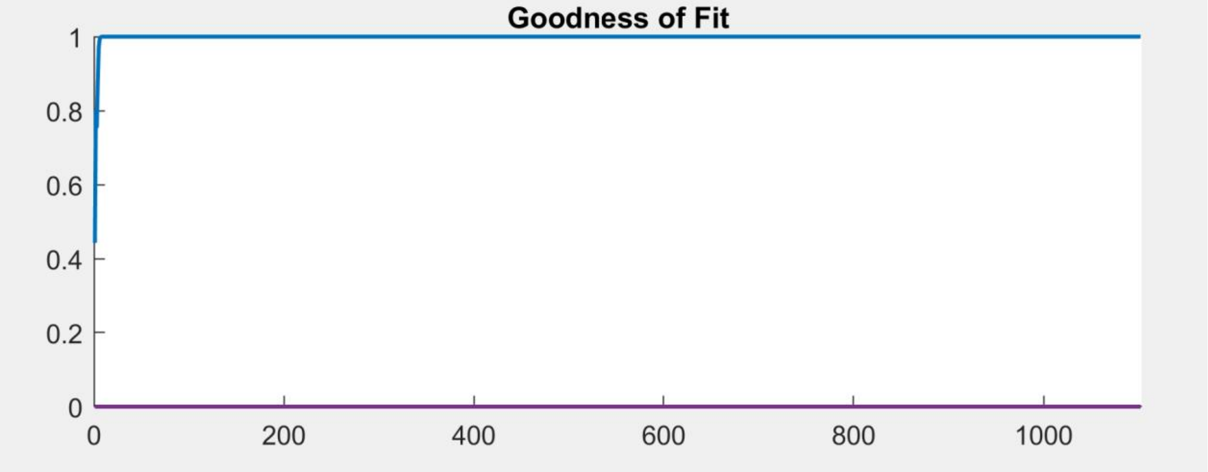
AWTLS on the HNEI data

The lab test data was fed into the algorithm with a random SoC change across measurement intervals. Forgetting factor was taken as 0.93.









Conclusion

- With the use of least squares based regression methods we were able to calculate the capacity degradation.
- The regular regression does not provide us with reliable data due to noise in both SoC and current measurements.
- Weighted Least squares (WLS) assumes a change in current measurement and assumes a noise free SoC measurement.
- The method works but is seen to have flaws in estimating error bounds and hence falls short in reliability.
- Approximated Weighted Least squares (AWTLS) takes into account both the noise and with it we get reliable estimates with reliable error bounds.
- On performing tests with a simulated data set AWTLS was found to be superior and with real lab tested data AWTLS remained the undefeated winner.
- We can conclude that AWTLS would be a robust algorithm with simple recursive computing, making it a reliable choice for capacity estimation.
- Simulation files have been attached

Reference

- [1] Plett, G.L.,"Recursive Approximate Weighted Total Least Squares Estimation of Battery Cell Total Capacity" Journal of Power Sources 196(4), 2011 https://doi.org/10.1016/j.jpowsour.2010.09.048
- [2] Devie, Arnaud, George Baure, and Matthieu Dubarry. 2018. "Intrinsic Variability in the Degradation of a Batch of Commercial 18650 Lithium-Ion Cells" *Energies* 11, no. 5: 1031. https://doi.org/10.3390/en11051031