Part 2

UNI:sv2414

NOTE:

1.This program writes the final output to file. Check “Part2_output.csv” for all values described here

2.The actual code can be viewed in the Rmd file “Part 2.Rmd”

3.Calculations here are based on Uncurtailed Power Generation values obtained in Part 1 (cleaning by correction). The same calculations can also be done using Actual Power Generation values.

Check if packages are installed, install if required, and load

```
## Loading required package: plyr
## Loading required package: ggplot2
## Loading required package: reshape2
```

Import CSV file containing demand data

***Summary of “dd” data frame is shown below*:**

```
##                 datetime      demand.kWh
##  2013-11-03 01:00:00:   2   Min.   :    0
##  2013-01-01 00:00:00:   1   1st Qu.: 6719
##  2013-01-01 01:00:00:   1   Median : 7877
##  2013-01-01 02:00:00:   1   Mean   : 7926
##  2013-01-01 03:00:00:   1   3rd Qu.: 9174
##  2013-01-01 04:00:00:   1   Max.   :11769
##  (Other)            :8753
```

Rename columns

Updated summary:

```
##                 DateTime        Demand
##  2013-11-03 01:00:00:   2   Min.   :    0
##  2013-01-01 00:00:00:   1   1st Qu.: 6719
##  2013-01-01 01:00:00:   1   Median : 7877
##  2013-01-01 02:00:00:   1   Mean   : 7926
##  2013-01-01 03:00:00:   1   3rd Qu.: 9174
##  2013-01-01 04:00:00:   1   Max.   :11769
##  (Other)            :8753
```

Import CSV file containing supply data (from part 1)

Summary of "sd" data frame is shown below:

```
##        X                       DateTime      AvgWindSpeed
##  Min.   :    1   2013-01-01 00:00:00:    1   Min.   : 0.00
##  1st Qu.:13132   2013-01-01 00:10:00:    1   1st Qu.: 5.60
##  Median :26254   2013-01-01 00:20:00:    1   Median : 8.50
##  Mean   :26271   2013-01-01 00:30:00:    1   Mean   : 8.08
##  3rd Qu.:39437   2013-01-01 00:40:00:    1   3rd Qu.:10.90
##  Max.   :52560   2013-01-01 00:50:00:    1   Max.   :18.80
##                  (Other)            :52486
##   MeterReading       tenminkwh   tenminmpcurve   tenminbetz
##  Min.   :     78   Min.   :  0   Min.   :  0   Min.   :   0
##  1st Qu.:2740574   1st Qu.:217   1st Qu.:132   1st Qu.: 158
##  Median :5151092   Median :417   Median :461   Median : 553
##  Mean   :5051828   Mean   :404   Mean   :467   Mean   : 776
##  3rd Qu.:7153463   3rd Qu.:584   3rd Qu.:799   3rd Qu.:1166
##  Max.   :9999699   Max.   :964   Max.   :992   Max.   :5980
##
##   tenminKEwind      eqPower       mpcWind       finalWSvalue
##  Min.   :    0   Min.   :  0   Min.   : 3.13   Min.   : 3.13
##  1st Qu.:  267   1st Qu.:186   1st Qu.: 6.77   1st Qu.: 7.00
##  Median :  932   Median :357   Median : 8.21   Median : 9.00
##  Mean   : 1309   Mean   :347   Mean   : 7.79   Mean   : 8.87
##  3rd Qu.: 1966   3rd Qu.:501   3rd Qu.: 9.25   3rd Qu.:11.00
##  Max.   :10085   Max.   :826   Max.   :13.04   Max.   :18.80
##
##  ActualGenerationkWh UncurtailedGenerationkWh    KEinWind
##  Min.   :  0         Min.   :  0              Min.   :   47
##  1st Qu.:217         1st Qu.:234              1st Qu.:  521
##  Median :417         Median :540              Median : 1106
##  Mean   :404         Mean   :520              Mean   : 1428
##  3rd Qu.:584         3rd Qu.:810              3rd Qu.: 2020
##  Max.   :964         Max.   :992              Max.   :10085
##
##  TurbineEfficiency
##  Min.   : 0.0
##  1st Qu.:22.7
##  Median :37.9
##  Mean   :33.6
##  3rd Qu.:47.2
##  Max.   :49.6
##
```

Remove columns I don't need

Updated summary:

```
##                 DateTime     ActualGenerationkWh UncurtailedGenerationkWh
##  2013-01-01 00:00:00:    1   Min.   :  0         Min.   :  0
##  2013-01-01 00:10:00:    1   1st Qu.:217         1st Qu.:234
##  2013-01-01 00:20:00:    1   Median :417         Median :540
##  2013-01-01 00:30:00:    1   Mean   :404         Mean   :520
##  2013-01-01 00:40:00:    1   3rd Qu.:584         3rd Qu.:810
##  2013-01-01 00:50:00:    1   Max.   :964         Max.   :992
##  (Other)             :52486
```

Convert DateTime to Date-Time values

There are missing rows in the supply data, I need to correct this

I create a new data frame “fulldf” with all 8760times6 timestamps

Summary of “fulldf” data frame is shown below:

```
##       full
##  Min.   :2013-01-01 00:00:00
##  1st Qu.:2013-04-02 06:57:30
##  Median :2013-07-02 12:55:00
##  Mean   :2013-07-02 12:55:00
##  3rd Qu.:2013-10-01 18:52:30
##  Max.   :2013-12-31 23:50:00
```

Rename column

Updated summary

```
##     DateTime
##  Min.   :2013-01-01 00:00:00
##  1st Qu.:2013-04-02 06:57:30
##  Median :2013-07-02 12:55:00
##  Mean   :2013-07-02 12:55:00
##  3rd Qu.:2013-10-01 18:52:30
##  Max.   :2013-12-31 23:50:00
```

I now merge “fulldf” with “sd” to get a complete supply data frame with all timestamps

Summary of “compsd” data frame is shown below:

```
##     DateTime                    ActualGenerationkWh
##  Min.   :2013-01-01 00:00:00   Min.   :  0
##  1st Qu.:2013-04-02 06:57:30   1st Qu.:217
##  Median :2013-07-02 12:55:00   Median :417
##  Mean   :2013-07-02 12:55:00   Mean   :404
##  3rd Qu.:2013-10-01 18:52:30   3rd Qu.:584
##  Max.   :2013-12-31 23:50:00   Max.   :964
##                                NA's   :68
##  UncurtailedGenerationkWh
##  Min.   :  0
##  1st Qu.:234
##  Median :540
##  Mean   :520
##  3rd Qu.:810
##  Max.   :992
##  NA's   :68
```

Convert NAs to 0s in compsd

Updated summary:

```
##     DateTime                    ActualGenerationkWh
##  Min.   :2013-01-01 00:00:00   Min.   :  0
##  1st Qu.:2013-04-02 06:57:30   1st Qu.:216
##  Median :2013-07-02 12:55:00   Median :417
##  Mean   :2013-07-02 12:55:00   Mean   :404
##  3rd Qu.:2013-10-01 18:52:30   3rd Qu.:584
##  Max.   :2013-12-31 23:50:00   Max.   :964
##  UncurtailedGenerationkWh
##  Min.   :  0
##  1st Qu.:233
##  Median :539
##  Mean   :520
##  3rd Qu.:808
##  Max.   :992
```

Now, I'm ready to start summing 10min intervals to get hourly intervals

Create new data frame "ad" for all data

I add demand (from "dd") and supply data (from"compsd") to "ad"

Summary of "ad" data frame is shown below:

```
##                  DateTime         Demand      UncurtailedSupply
##  2013-11-03 01:00:00:   2   Min.   :    0   Min.   :   0
##  2013-01-01 00:00:00:   1   1st Qu.: 6719   1st Qu.:1410
##  2013-01-01 01:00:00:   1   Median : 7877   Median :3251
##  2013-01-01 02:00:00:   1   Mean   : 7926   Mean   :3119
##  2013-01-01 03:00:00:   1   3rd Qu.: 9174   3rd Qu.:4819
##  2013-01-01 04:00:00:   1   Max.   :11769   Max.   :5939
##  (Other)             :8753
```

Add a new row and calculate curtailment

Summary of "Curtailment" column is shown below:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0       0      20       0    5920
```

Create data frame for storing 12*24 curtailment values

For consistency, ensure DateTime is of class POSIXct

Now, I calculate mean hourly curtailment for each hour slot for each month

(I do this by converting the DateTime stamps to POSIXlt, which is stored as a list and individual elements (month, hour) can be accessed)

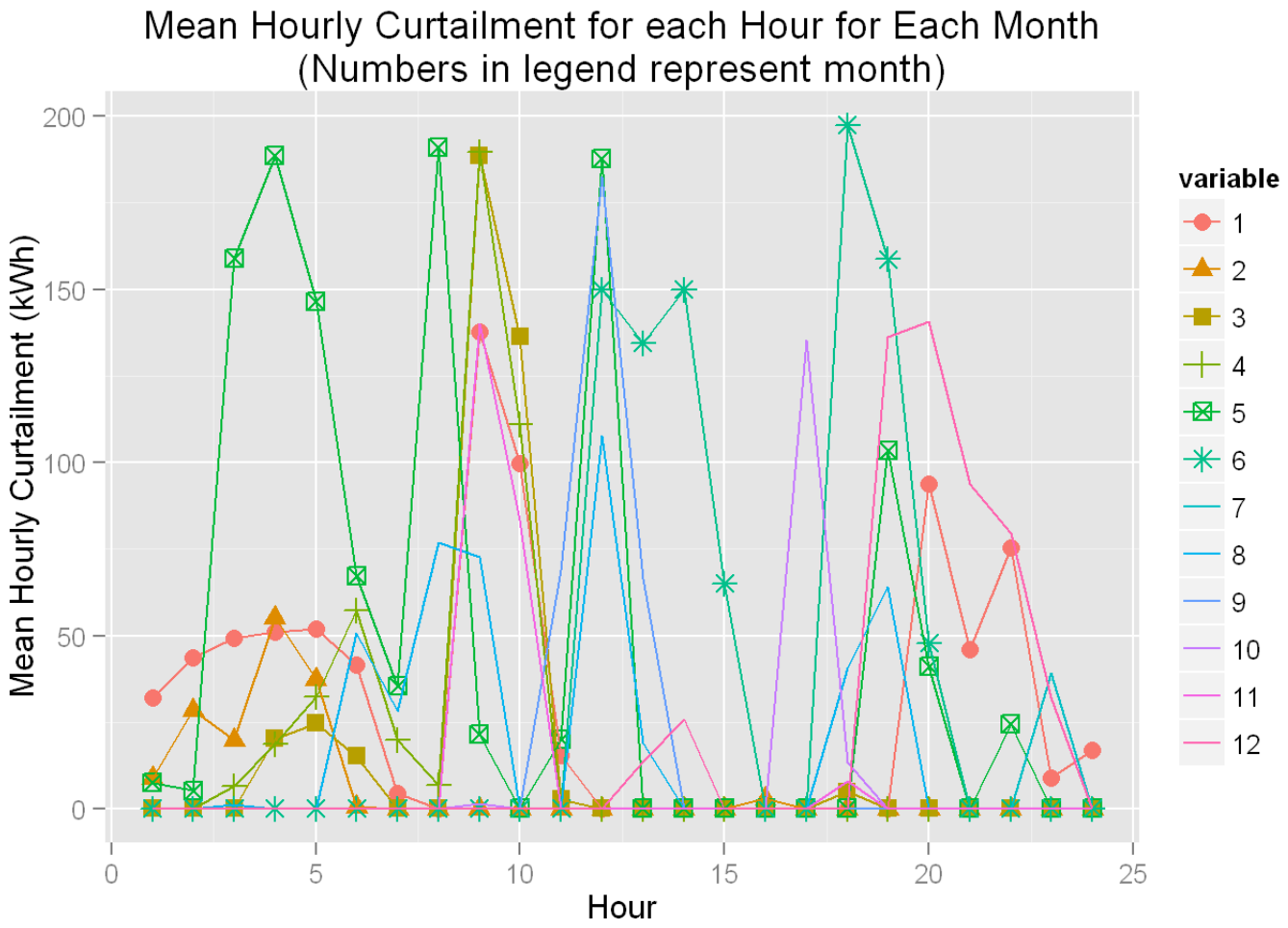
Mean hourly curtailment for each hour slot for each month is shown:

```
##          1       2       3       4       5      6     7       8      9
## 1   32.023  8.8882   0.000   0.000   7.513   0.00  0.00   0.000   0.00
## 2   43.510 28.4434   0.000   0.000   5.207   0.00  0.00   0.000   0.00
## 3   49.130 19.8397   0.000   6.470 158.870   0.00  0.00   1.033   0.00
## 4   51.029 55.2230  20.315  18.626 188.572   0.00  0.00   0.000   0.00
## 5   51.861 37.3778  24.694  32.351 146.497   0.00  0.00   0.000   0.00
## 6   41.680  0.5339  15.254  57.319  67.255   0.00  0.00  50.791   0.00
## 7    4.451  0.0000   0.000  19.881  35.350   0.00  0.00  28.103   0.00
## 8    0.000  0.0000   0.000   6.724 191.148   0.00  0.00  77.010   0.00
## 9  137.833  0.0000 188.700 189.581  21.468   0.00  0.00  72.613   0.00
## 10  99.791  0.0000 136.296 110.948   0.000   0.00  0.00   0.000   0.00
## 11  15.309  0.0000   2.686   0.000  19.950   0.00  0.00   0.000  68.88
## 12   0.000  0.0000   0.000   0.000 187.865 150.03  0.00 107.656 182.51
## 13   0.000  0.0000   0.000   0.000   0.000 134.53  0.00  18.968  66.52
## 14   0.000  0.0000   0.000   0.000   0.000 150.00  0.00   0.000   0.00
## 15   0.000  0.0000   0.000   0.000   0.000  65.07  0.00   0.000   0.00
## 16   0.000  2.9026   0.000   0.000   0.000   0.00  0.00   0.000   0.00
## 17   0.000  0.0000   0.000   0.000   0.000   0.00  0.00   0.000   0.00
## 18   0.000  0.0000   5.015   0.000   0.000 197.45  0.00  40.727   0.00
## 19   0.000  0.0000   0.000   0.000 103.444 158.77  0.00  64.227   0.00
## 20  93.661  0.0000   0.000   0.000  41.097  47.85  0.00   0.000   0.00
## 21  45.883  0.0000   0.000   0.000   0.000   0.00  0.00   0.000   0.00
## 22  75.274  0.0000   0.000   0.000  24.466   0.00  0.00   0.000   0.00
## 23   8.753  0.0000   0.000   0.000   0.000   0.00 39.06   0.000   0.00
## 24  16.780  0.0000   0.000   0.000   0.000   0.00  0.00   0.000   0.00
##          10      11     12
## 1    0.0000   0.000   0.00
## 2    0.0000   0.000   0.00
## 3    0.0000   0.000   0.00
## 4    0.0000   0.000   0.00
## 5    0.0000   0.000   0.00
## 6    0.0000   0.000   0.00
## 7    0.0000   0.000   0.00
## 8    0.0000   0.000   0.00
## 9    1.3923 139.966   0.00
## 10   0.0000  83.310   0.00
## 11   0.4229   0.000   0.00
## 12   0.0000   0.000   0.00
## 13   0.0000   0.000  13.71
## 14   0.0000   0.000  25.84
## 15   0.0000   0.000   0.00
## 16   0.0000   0.000   0.00
```

```
## 17 135.3746   0.000   0.00
## 18  13.3517   8.006   0.00
## 19   0.0000   0.000 136.17
## 20   0.0000   0.000 140.64
## 21   0.0000   0.000  93.72
## 22   0.0000   0.000  79.55
## 23   0.0000   0.000  31.98
## 24   0.0000   0.000   0.00
```

Plot mean hourly curtailment values

```
## Warning: The shape palette can deal with a maximum of 6 discrete values
## because more than 6 becomes difficult to discriminate; you have
## 12. Consider specifying shapes manually. if you must have them.
## Warning: The shape palette can deal with a maximum of 6 discrete values
## because more than 6 becomes difficult to discriminate; you have
## 12. Consider specifying shapes manually. if you must have them.
## Warning: Removed 144 rows containing missing values (geom_point).
## Warning: The shape palette can deal with a maximum of 6 discrete values
## because more than 6 becomes difficult to discriminate; you have
## 12. Consider specifying shapes manually. if you must have them.
```



Now, to find how storage affects the system

I will use the following method:

PSEUDOCODE TO FIND HOW STORAGE AFFECTS SYSTEM:

For each hour, I make the following columns:

1. Demand
2. Supply
3. Difference (Supply minus Demand)
4. Positive Differences
5. Negative Differences
6. Storage (Minimum of 0kWh, maximum of 5000kWh)
7. Total kWh Supplied (From Generation+Storage)

Details of calculating each column:

1. Demand is known from demand data.
2. Supply is known from Part 1.
3. The difference between Supply and Demand can be calculated (Supply minus Demand). These values are positive or negative. If they are positive, Supply>Demand. If they are negative, Demand>Supply.
4. Positive differences from Difference column are copied over here. Other rows are assigned 0s.
5. Negative differences from Difference column are copied over here. Other rows are assigned 0s.
6. kWh in storage can be calculated by:
    a. If Difference (column 3) is Positive,
    Storage=Storage from previous row + 0.85*Positive Difference
    b. If Difference (column 3) is Negative,
    Storage=Storage from previous row + Negative Difference/0.9
    c. If, after the above operations, storage becomes negative,
    Storage<-0
    If after the above operations, storage becomes greater than 5000,
    Storage<-5000
7. Total kWh supplied in each hour (from generation+storage) can be calculated by:
    a. If Difference (column 3) is Positive,
    Total kWh Supplied=Demand (column 1)
    b. If Difference (column 3) is Negative,
    Total kWh Supplied=Supply (column 2) + [Storage (previous row) – Storage (this row)]*0.9

NOTE:

a. The Positive and Negative Difference columns are not really required. They just make it easier for me to verify the working of the code by manually scrolling and checking in the data frame.

b. The Positive Difference column is essentially the Curtailment column.

Add a column to find difference between supply and demand (supply minus demand)

Summary of "Diff" column is shown below:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -11400   -6550   -4670   -4810   -3030    5920
```

I separate the positive and negative differences in "Diff"

(I do this to visualise storage inputs and outputs easily)

***Summary of "PosDiff" and "NegDiff" columns are shown below*:**

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0       0      20       0    5920
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -11400   -6550   -4670   -4830   -3030       0
```

Add a column for how much is stored

(Everything that flows in (PosDiff) is multiplied by 0.85, and everything that flows out (NegDiff) is divided by 0.9)

(Limits of 0 and 5000 are maintained)

(Check pseudocode/code for details)

***Summary of "storage" column is shown below*:**

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0       0       0      39       0    5000
```

Add a column for total supplied kWh (from generation+storage)

(Here, I calculate what is effectively supplied to the grid each hour)

(Check pseudocode/code for details)

***Summary of "TotkWhSupplied" column is shown below*:**

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0    1390    3240    3110    4810    9590
```

Find Capacity Factors with and without storage

```
## [1] "Capacity factor WITHOUT storage is 52.09pc"
```

```
## [1] "Capacity factor WITH storage is 52.28pc"
```

Write data to file

```
## [1] "Printed to: Part2_output.csv"
```

FINAL COMMENTS:

1. Uncurtailed Capacity factors are calculated considering compatibility with grid demand, both with and without storage.

2. The CF considering demand WITHOUT storage is the least at 52.08%. This is expected because of "curtailment" losses.

3. The CF considering demand WITH storage is higher, at 52.28%. This is expected because "curtailment" losses are redirected back to the grid after storage (with small efficiency losses).

4. The CF from Part 1 was 52.4%. This, as expected, is the highest since “curtailment” losses are not considered.

5. However, as apparent, the difference between these CFs is small. This is because “curtailment” losses are small compared to total generation. (In most cases, demand is greater than supply.) Also, the effect of the storage, while present, is small, since it rarely gets filled. It is an improvement, but by a small margin.