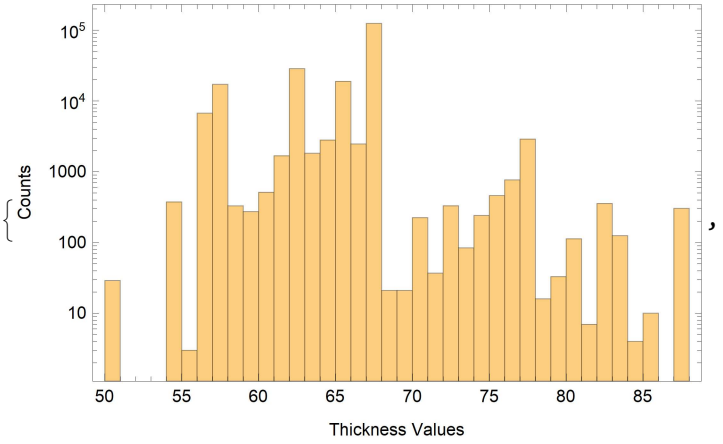
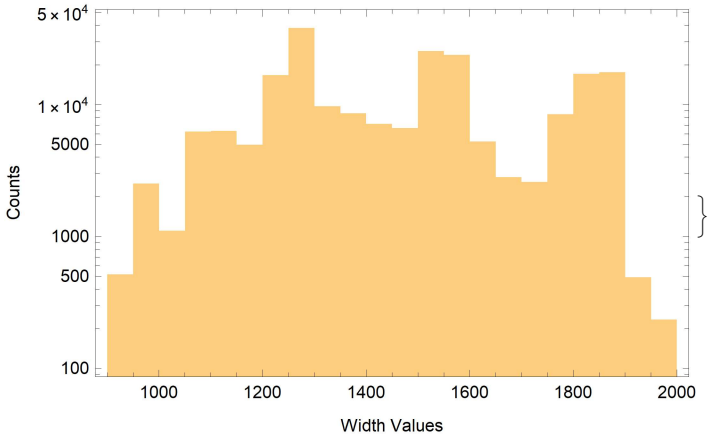
```
In[ ]:= SetDirectory[
          "C:/Users/serha/OneDrive/Masaüstü/MyRepo/master_thesis_MMT003/210421_OR_model_and
            _other_lines_sliding"];

In[ ]:= datafull = Import["../data/CSP_num.csv", HeaderLines → 2];
        datafull[[1]]
        Dimensions@datafull

Out[ ]= {1660, CCM1, 17078201-01, 239846, 49, 17078201,
         5, 1240, 16.766, 10.466, 66, 6.45848, 30.06.17 20:11}

Out[ ]= {212902, 13}

In[ ]:= nullpos = Position[datafull[[All, 8]], _?(Head@# == String &)];
        datafull = Delete[datafull, nullpos];

In[ ]:= mostlyzeropos = Position[datafull[[All, 9]], _?(# == 0 &)][[{2, 3}]];
        datafull = Delete[datafull, mostlyzeropos];

In[ ]:= deletepos = Position[datafull[[All, 11]], _?(1500 < # &)];
        (* Position[datafull[[All,8]],_?(35000<#&)] *)
        datafull = Delete[datafull, deletepos];

In[ ]:= deletepos2 = Position[datafull[[All, 11]], _?(10 > # &)];
        (* Position[datafull[[All,8]],_?(200>#&)] *)
        datafull = Delete[datafull, deletepos2];
```

In[ ]:=
```
{Histogram[datafull[[All, 11]], ScalingFunctions → "Log", PlotRange → Full,
  Frame → True, FrameLabel → {"Thickness Values", "Counts"}, ImageSize → Medium],
 Histogram[datafull[[All, 8]], ScalingFunctions → "Log", PlotRange → Full,
  Frame → True, FrameLabel → {"Width Values", "Counts"}, ImageSize → Medium]}
```

Out[ ]=



,



In[ ]:=
```
(* marginal density rows deletion *)
datafull = Delete[datafull, {{28 855}, {196 519}, {11 398}, {44 971}, {67 986}}];
```

In[ ]:=
```
(* remained negative density rows deletion *)
datafull =
  Delete[datafull, {{80 527}, {82 593}, {86 698}, {88 884}, {98 206}, {98 207}, {159 853},
    {159 854}, {172 604}, {172 605}, {172 606}, {184 960}, {189 744}, {189 752}, {195 041},
    {195 976}, {199 764}, {200 365}, {200 740}, {200 741}, {200 993}, {200 994}, {200 995},
    {201 012}, {201 566}, {201 581}, {201 582}, {201 584}, {201 589}, {201 591},
    {201 593}, {201 594}, {201 604}, {201 606}, {202 974}, {204 090}, {204 869}}];
```

irrelevant density rows deletion

In[ ]:=
```
density[thick_, width_, length_, weight_] := N@weight / (thick * width * length);
```

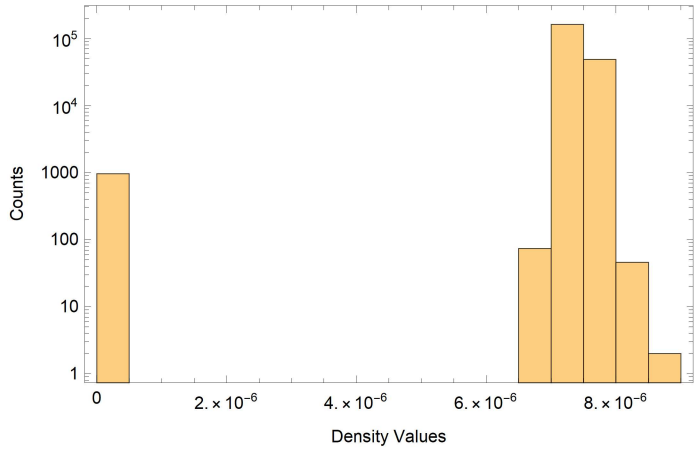
In[ ]:=
```
thickvaluesthkpos = datafull[[All, 11]];
widthvaluesthkpos = datafull[[All, 8]];
lengthvaluesthkpos = datafull[[All, 9]];
weightvaluesthkpos = datafull[[All, 10]];
densities = Quiet@Table[density[thickvaluesthkpos[[i]], widthvaluesthkpos[[i]],
      lengthvaluesthkpos[[i]], weightvaluesthkpos[[i]]], {i, Length@thickvaluesthkpos}];
densities = densities /. {Indeterminate → 0., ComplexInfinity → 0.};
KeySort@Counts@densities
```

Out[ ]=
$$\langle | 0. \to 968,\ 6.6087 \times 10^{-6} \to 1,\ 6.64335 \times 10^{-6} \to 1,\ 6.71005 \times 10^{-6} \to 1,$$
$$6.74438 \times 10^{-6} \to 1,\ 6.75775 \times 10^{-6} \to 1,\ \cdots 134\,194 \cdots,\ 8.46405 \times 10^{-6} \to 1,$$
$$8.47798 \times 10^{-6} \to 1,\ 8.48358 \times 10^{-6} \to 1,\ 8.57632 \times 10^{-6} \to 1,\ 8.58354 \times 10^{-6} \to 1 | \rangle$$

large output | show less | show more | show all | set size limit...

In[ ]:=
```
Histogram[densities, ScalingFunctions → "Log", PlotRange → Full,
 Frame → True, FrameLabel → {"Density Values", "Counts"}, ImageSize → Medium]
```

Out[ ]=

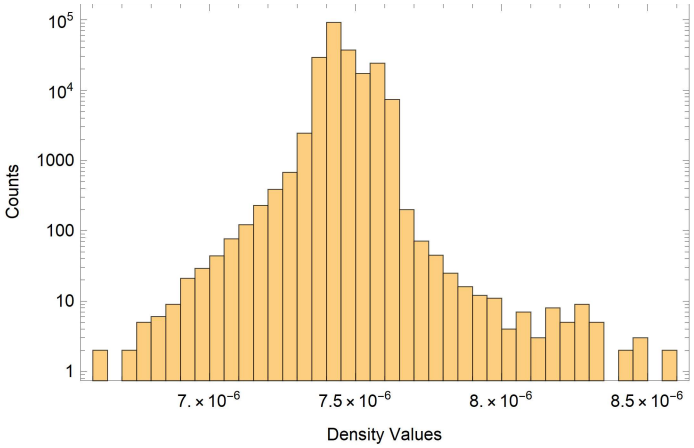


In[ ]:=
```
datafull = datafull[[Flatten@Position[densities, _?(# ≠ 0 &)]]];
```

In[ ]:=
```
thickvaluesthkpos = datafull[[All, 11]];
widthvaluesthkpos = datafull[[All, 8]];
lengthvaluesthkpos = datafull[[All, 9]];
weightvaluesthkpos = datafull[[All, 10]];
densities = Quiet@Table[density[thickvaluesthkpos[[i]], widthvaluesthkpos[[i]],
      lengthvaluesthkpos[[i]], weightvaluesthkpos[[i]]], {i, Length@thickvaluesthkpos}];
densities = densities /. {Indeterminate → 0., ComplexInfinity → 0.};
```

In[ ]:=
```
Histogram[densities, ScalingFunctions → "Log", PlotRange → Full,
 Frame → True, FrameLabel → {"Density Values", "Counts"}, ImageSize → Medium]
```

Out[ ]=



In[ ]:=
```
Length@datafull
```

Out[ ]=
```
211 512
```

In[ ]:=
```
Dimensions@datafull
```

Out[ ]=
```
{211 512, 13}
```

In[ ]:=
```
secondscolumn =
  Table[AbsoluteTime[{datafull[[i, 13]], {"Day", ".", "Month", ".", "YearShort",
      " ", "Hour", ":", "Minute"}}], {i, Length@datafull}];
datafull = Join[datafull, Partition[secondscolumn, 1], 2];
datafullsorted = Sort[datafull, #1[[1]] < #2[[1]] &];
```

In[ ]:=
```
(* checking if sequences are revealing consecutive *)
Length@Table[DeleteDuplicates@i, {i, Split[datafullsorted[[All, 1]], #2 == #1 &]}]
Length@DeleteDuplicates[datafullsorted[[All, 1]]]
```

Out[ ]=
```
1587
```

Out[ ]=
```
1587
```

Deletion of sequences less than 50

In[ ]:=
```
deletepos5 = Flatten@Table[Position[datafullsorted[[All, 1]], i],
    {i, Keys@Cases[Normal@Counts@datafullsorted[[All, 1]], _?(Values[#] < 50 &)]}];
datafullsorted = Delete[datafullsorted, Partition[deletepos5, 1]];
```

In[ ]:=
```
Dimensions@datafullsorted
```

Out[ ]=
```
{205 496, 14}
```

In[•]:=
```
programids = DeleteDuplicates@datafullsorted[[All, 1]];
Length@programids
datafullsortedfinal =
  Flatten[Table[Sort[Select[datafullsorted, #[[1]] == i &], #1[[14]] < #2[[14]] &],
    {i, programids}], 1];
Dimensions@datafullsortedfinal
```

Out[•]= 1357

Out[•]= {205 496, 14}

In[•]:=
```
datafullsortedfinal[[1]]
```

Out[•]= {1660, CCM1, 17078201-01, 239 846, 49, 17 078 201, 5,
1240, 16.766, 10.466, 66, 6.45848, 30.06.17 20:11, 3 707 842 260}

In[•]:=
```
data = Join[Partition[Range@Length@datafullsortedfinal, 1],
  Partition[datafullsortedfinal[[All, 1]], 1],
  ConstantArray[{0, 0, 0, 0, 0, 0}, Length@datafullsortedfinal],
  datafullsortedfinal[[All, {8, 11, 13, 14, 6, 5, 12, 4, 3}]], 2]
```

Out[•]=
```
{{1, 1660, 0, 0, 0, 0, 0, 0, 1240, 66, 30.06.17 20:11,
  3 707 842 260, 17 078 201, 49, 6.45848, 239 846, 17078201-01},
 ··· 205 494 ···, {205 496, 9X0007, 0, 0, 0, 0, 0, 0, 1278, 67, 21.09.19 00:51,
  3 778 015 860, 19 151 221, 26, 3.56109, 810 670, 19151221-06}}
```

large output | show less | show more | show all | set size limit...

In[•]:=
```
Dimensions@data
```

Out[•]= {205 496, 17}

In[•]:=
```
(*Export["csp_manipulated_205496.csv",data]*)

(*Export["../data/csp_manipulated_205496_rev1.csv",data];*)
```