Milano Chemometrics and QSAR Research Group

Roberto Todeschini
Viviana Consonni
Davide Ballabio
Francesca Grisoni
Miriam Cortinovis

Department of Environmental Sciences
University of Milano - Bicocca
P.za della Scienza, 1 - 20126 Milano (Italy)
Website: michem.disat.unimib.it/chm/
Tournament tables, Power-Weakness Ratio and Hasse diagrams: an informative combination for multi-criteria decision-making.

Roberto Todeschini
Viviana Consonni
Davide Ballabio
Francesca Grisoni
Miriam Cortinovis
The starting points of this work are our two previous papers published in 2015:

Weighted power-weakness ratio for multi-criteria decision making. 
*Chemometrics and Intelligent Laboratory Systems*, 146, 329-336.

F. Grisoni, V. Consonni, S. Nembri, R. Todeschini (2015)
How to weight Hasse diagrams and reduce incomparabilities. 
*Chemometrics and Intelligent Laboratory Systems*, 147, 95-104.
The results of a Round Robin tournament of N players can be conveniently expressed by mean of a tournament table (dominance matrix) as:

\[ t_{ij} + t_{ji} = 1 \]

H.A. David (1971)
Ranking the Players in a Round Robin Tournament.
The results of a Round Robin tournament of \( N \) players can be conveniently expressed by mean of a tournament table (dominance matrix) as:

\[
t_{ij} + t_{ji} = 1
\]

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>...</th>
<th>...</th>
<th>PN</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0</td>
<td>1</td>
<td>( t_{13} )</td>
<td>...</td>
<td>...</td>
<td>( t_{1N} )</td>
</tr>
<tr>
<td>P2</td>
<td>0</td>
<td>0</td>
<td>( t_{23} )</td>
<td>...</td>
<td>...</td>
<td>( t_{2N} )</td>
</tr>
<tr>
<td>P3</td>
<td>( t_{31} )</td>
<td>( t_{32} )</td>
<td>0</td>
<td>...</td>
<td>...</td>
<td>( t_{3N} )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>PN</td>
<td>( t_{N1} )</td>
<td>( t_{N2} )</td>
<td>( t_{N3} )</td>
<td>...</td>
<td>...</td>
<td>0</td>
</tr>
</tbody>
</table>

H.A. David (1971)
Ranking the Players in a Round Robin Tournament.
## Tournament table

Tournament table $T_1$. For each $t_{ij}$: 1 if the $P_i$ player won over $P_j$, 0 if $P_j$ won, 0.5 if they drew the match.

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>P2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>P3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>P5</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

$t_{ij} + t_{ji} = 1$

Sometimes, some conflicting rings arise:
1) Ranking cannot be decided
2) Transitivity property is lost

$P1 > P3 > P5 > P1$

The row sum (Copeland score) can used for ranking:

- P1: 2.5
- P2 = P3 = P5 = 2
- P4 = 1.5

... but the ranking power can be low!
A tournament table can be derived from any data matrix $X (N, p)$, where $N$ is the number of objects and $p$ the number of variables, i.e. the considered criteria.

$$X \rightarrow T_w$$

The general expression for this transform is defined by comparing objects pairwise:

$$t_{ij}^W = \sum_{k=1}^{p} \mathbf{w}_k \cdot \delta_{ij,k} \quad \text{where} \quad \delta_{ij,k} = \begin{cases} 1 & \text{if } x_{ik} \succ x_{jk} \\ 0.5 & \text{if } x_{ik} = x_{jk} \\ 0 & \text{if } x_{ik} \prec x_{jk} \end{cases}$$

and $\sum_{k=1}^{p} w_k = 1$

... where the main differences with respect to the Hasse approach are ...
A tournament matrix can be derived from any data matrix $X$ ($N, p$), where $N$ is the number of objects and $p$ the number of variables, i.e. the considered criteria.

$$X \rightarrow T_W$$

The general expression for this transform is defined by comparing objects pairwise:

$$t_{ij}^W = \sum_{k=1}^{p} w_k \cdot \delta_{ij,k} \quad \text{where} \quad \delta_{ij,k} = \begin{cases} 1 & \text{if } x_{ik} \succ x_{jk} \\ 0.5 & \text{if } x_{ik} \cong x_{jk} \\ 0 & \text{if } x_{ik} \prec x_{jk} \end{cases} \quad \text{and} \quad \sum_{k=1}^{p} w_k = 1$$

A set of thresholds are also derived from the tournament table:

$$X \rightarrow T_W \rightarrow \{t_1, t_2, \ldots, t_k\}$$
Analyzing thresholds of the tournament table

$$0.5 \leq t^* \leq 1$$

The following transforms are performed:

- **Tournament table $Tw(t^*)$**
  - $1 - t^*$: unchanged
  - $0.5$: unchanged

- **Regularized Hasse matrix($t^*$)**
  - $1 - t^*$: $-1$
  - $0$: $0$
  - $t^*$: $+1$
Power-Weakness Ratio

For any squared asymmetrical matrix, the \textit{Perron-Frobenius theorem} guarantees the existence of a positive eigenvalue associated with an eigenvector \( e \) having positive values.

\textbf{Tournament table} \( Tw \)

Kendall (1955) proposed to use the \textit{eigenvector values} to rank the objects, thus also removing possible lost of transitivity:

\[
Tw \rightarrow e
\]

Ramanujacharyulu (1964) proposed to use also the eigenvector values calculated on the transpose of \( Tw \):

\[
Tw^T \rightarrow e^*
\]
... then the $PWR$ of the $i$-th object was defined as:

$$PWR_i = \frac{e_i}{e^*_i}$$

Indeed, the first eigenvector awards good players able to win with other good players, while the second eigenvector characterizes bad players which loss with other bad players.
Power-Weakness Ratio

Tournament table $T_1$. For each $t_{ij}$: 1 if the $P_i$ player won over $P_j$, 0 if $P_j$ won, 0.5 if they drew the match.

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>P2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>P3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>P4</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>P5</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

Results of PWR scoring on table $T_1$. Entries of the Perron–Frobenius eigenvector calculated on tournament table ($e_{PF}$) and on its transpose ($e_{PF}^*$) for each player are also reported.

<table>
<thead>
<tr>
<th>Players</th>
<th>$e_{PF}$</th>
<th>$e_{PF}^*$</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.529</td>
<td>0.370</td>
<td>1.368</td>
</tr>
<tr>
<td>P2</td>
<td>0.430</td>
<td>0.442</td>
<td>0.976</td>
</tr>
<tr>
<td>P3</td>
<td>0.426</td>
<td>0.414</td>
<td>1.025</td>
</tr>
<tr>
<td>P4</td>
<td>0.364</td>
<td>0.535</td>
<td>0.714</td>
</tr>
<tr>
<td>P5</td>
<td>0.471</td>
<td>0.459</td>
<td>1.023</td>
</tr>
</tbody>
</table>

Eigenvector ($T_W$)  Eigenvector ($T_W^T$)
Tw transform

Data $X$ goes through weights to $T_W$, resulting in Tournament table.

Threshhlds $t_1, t_2, \ldots, t_k$ lead to PWR ranks $PWR_1, PWR_2, \ldots, PWR_k$.

PWR Diagrams $PWRD_1, PWRD_2, \ldots, PWRD_k$.
Hasse transform

\[ H^R(t^*) \]

\[
\begin{cases}
  +1 & \text{if } t_{ij}^W \geq t^* \\
  -1 & \text{if } t_{ij}^W \leq 1 - t^* \\
  0 & \text{otherwise}
\end{cases}
\]

\[ 0.50 < t^* \leq 1 \]
Summary

$X(\mathbb{N}, p)$

$\{w_1, w_2, \ldots, w_p\}$

Tw

$\{t_1, t_2, \ldots, t_k\}$

$T'w$

Tw transform

Hasse transform

HD

RHD

PWR

PWRD

Hasse diagrams

Regularized Hasse diagrams

PWR diagrams
Comparisons of classification methods

- 32 data sets
- Validation procedure: leave-one-out
- Parameter: Non-Error-Rate (NER%)

10 CLASSIFIERS

<table>
<thead>
<tr>
<th>N3</th>
<th>PLS-DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNN</td>
<td>CAIMAN</td>
</tr>
<tr>
<td>KNN</td>
<td>CART</td>
</tr>
<tr>
<td>LDA</td>
<td>SVM/LIN</td>
</tr>
<tr>
<td>QDA</td>
<td>SVM/RBF</td>
</tr>
</tbody>
</table>
## Comparisons of classification methods

<table>
<thead>
<tr>
<th>Id</th>
<th>Data set</th>
<th>N3</th>
<th>BNN</th>
<th>KNN</th>
<th>LDA</th>
<th>QDA</th>
<th>PLSDA</th>
<th>CART</th>
<th>CAIMAN</th>
<th>SVM /LIN</th>
<th>SVM /RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IRIS</td>
<td>96.0</td>
<td>96.7</td>
<td>96.7</td>
<td>98.0</td>
<td>97.3</td>
<td>90.2</td>
<td>94.0</td>
<td>98.0</td>
<td>97.3</td>
<td>97.3</td>
</tr>
<tr>
<td>2</td>
<td>WINES</td>
<td>96.2</td>
<td>98.6</td>
<td>97.7</td>
<td>99.1</td>
<td>99.5</td>
<td>99.5</td>
<td>86.2</td>
<td>98.7</td>
<td>99.1</td>
<td>99.5</td>
</tr>
<tr>
<td>3</td>
<td>PERPOT</td>
<td>99.0</td>
<td>99.0</td>
<td>99.0</td>
<td>85.0</td>
<td>92.0</td>
<td>86.0</td>
<td>97.0</td>
<td>97.0</td>
<td>87.0</td>
<td>100.0</td>
</tr>
<tr>
<td>4</td>
<td>ITAOILS</td>
<td>96.2</td>
<td>95.2</td>
<td>94.7</td>
<td>94.7</td>
<td>95.9</td>
<td>95.9</td>
<td>87.2</td>
<td>82.8</td>
<td>94.7</td>
<td>95.9</td>
</tr>
<tr>
<td>5</td>
<td>SULFA</td>
<td>77.4</td>
<td>73.8</td>
<td>73.8</td>
<td>45.2</td>
<td>69.4</td>
<td>74.0</td>
<td>81.5</td>
<td>58.7</td>
<td>50.0</td>
<td>88.7</td>
</tr>
<tr>
<td>6</td>
<td>DIABETES</td>
<td>73.6</td>
<td>71.1</td>
<td>70.5</td>
<td>72.7</td>
<td>69.6</td>
<td>75.1</td>
<td>68.8</td>
<td>73.5</td>
<td>72.3</td>
<td>72.3</td>
</tr>
<tr>
<td>7</td>
<td>BLOOD</td>
<td>67.9</td>
<td>62.2</td>
<td>62.3</td>
<td>53.7</td>
<td>54.5</td>
<td>68.7</td>
<td>62.1</td>
<td>59.3</td>
<td>50.0</td>
<td>64.1</td>
</tr>
<tr>
<td>8</td>
<td>VERTEBRAL</td>
<td>80.8</td>
<td>81.6</td>
<td>80.2</td>
<td>80.7</td>
<td>84.0</td>
<td>82.1</td>
<td>76.9</td>
<td>56.0</td>
<td>83.3</td>
<td>84.3</td>
</tr>
<tr>
<td>9</td>
<td>SEDIMENTS</td>
<td>88.9</td>
<td>88.9</td>
<td>89.9</td>
<td>66.9</td>
<td>69.4</td>
<td>79.4</td>
<td>84.3</td>
<td>61.1</td>
<td>50.5</td>
<td>69.9</td>
</tr>
<tr>
<td>10</td>
<td>BIODEG</td>
<td>84.5</td>
<td>85.3</td>
<td>85.4</td>
<td>77.0</td>
<td>78.6</td>
<td>79.9</td>
<td>79.6</td>
<td>65.6</td>
<td>81.5</td>
<td>83.8</td>
</tr>
<tr>
<td>11</td>
<td>DIGITS</td>
<td>74.2</td>
<td>72.3</td>
<td>73.6</td>
<td>74.0</td>
<td>68.6</td>
<td>41.0</td>
<td>65.2</td>
<td>77.3</td>
<td>74.9</td>
<td>74.5</td>
</tr>
<tr>
<td>12</td>
<td>APPLE</td>
<td>94.0</td>
<td>92.3</td>
<td>91.9</td>
<td>91.9</td>
<td>87.6</td>
<td>95.4</td>
<td>92.1</td>
<td>83.9</td>
<td>94.4</td>
<td>92.3</td>
</tr>
<tr>
<td>13</td>
<td>TOBACCO</td>
<td>92.3</td>
<td>92.3</td>
<td>92.3</td>
<td>84.6</td>
<td>80.8</td>
<td>88.5</td>
<td>96.2</td>
<td>92.3</td>
<td>92.3</td>
<td>92.3</td>
</tr>
<tr>
<td>14</td>
<td>SCHOOL</td>
<td>95.3</td>
<td>96.6</td>
<td>96.2</td>
<td>90.8</td>
<td>95.2</td>
<td>89.4</td>
<td>86.8</td>
<td>95.0</td>
<td>94.0</td>
<td>96.4</td>
</tr>
<tr>
<td>15</td>
<td>BANK</td>
<td>86.9</td>
<td>91.2</td>
<td>86.9</td>
<td>86.5</td>
<td>88.5</td>
<td>84.9</td>
<td>86.5</td>
<td>88.5</td>
<td>88.5</td>
<td>88.9</td>
</tr>
<tr>
<td>16</td>
<td>HIRSUTISM</td>
<td>88.3</td>
<td>90.1</td>
<td>90.0</td>
<td>55.4</td>
<td>81.4</td>
<td>84.1</td>
<td>70.5</td>
<td>52.9</td>
<td>72.7</td>
<td>93.8</td>
</tr>
<tr>
<td>17</td>
<td>THIOPHENE</td>
<td>83.3</td>
<td>83.3</td>
<td>83.3</td>
<td>79.2</td>
<td>79.2</td>
<td>90.5</td>
<td>58.3</td>
<td>83.3</td>
<td>83.3</td>
<td>83.3</td>
</tr>
<tr>
<td>18</td>
<td>SUNFLOWERS</td>
<td>92.3</td>
<td>90.4</td>
<td>91.2</td>
<td>87.8</td>
<td>90.8</td>
<td>92.7</td>
<td>82.1</td>
<td>88.9</td>
<td>90.8</td>
<td>96.9</td>
</tr>
<tr>
<td>19</td>
<td>VINEGRES</td>
<td>100.0</td>
<td>91.7</td>
<td>95.8</td>
<td>100.0</td>
<td>87.5</td>
<td>100.0</td>
<td>67.3</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>20</td>
<td>CHEESE</td>
<td>76.1</td>
<td>78.3</td>
<td>78.1</td>
<td>78.8</td>
<td>82.9</td>
<td>84.7</td>
<td>63.9</td>
<td>77.5</td>
<td>76.2</td>
<td>85.6</td>
</tr>
<tr>
<td>21</td>
<td>ORUJOS</td>
<td>98.2</td>
<td>98.4</td>
<td>98.2</td>
<td>92.6</td>
<td>94.1</td>
<td>93.9</td>
<td>88.4</td>
<td>62.5</td>
<td>95.7</td>
<td>98.2</td>
</tr>
<tr>
<td>22</td>
<td>MEMBRANE</td>
<td>94.4</td>
<td>94.4</td>
<td>94.4</td>
<td>88.9</td>
<td>94.4</td>
<td>96.7</td>
<td>91.7</td>
<td>94.4</td>
<td>91.7</td>
<td>94.4</td>
</tr>
<tr>
<td>23</td>
<td>METHACYCLINE</td>
<td>82.5</td>
<td>86.7</td>
<td>82.5</td>
<td>45.8</td>
<td>81.7</td>
<td>55.8</td>
<td>65.8</td>
<td>80.0</td>
<td>54.2</td>
<td>82.5</td>
</tr>
<tr>
<td>24</td>
<td>SIMUL4</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>28.1</td>
<td>100.0</td>
<td>46.9</td>
<td>90.6</td>
<td>93.8</td>
<td>34.4</td>
<td>100.0</td>
</tr>
<tr>
<td>25</td>
<td>VEGOIL</td>
<td>99.0</td>
<td>100.0</td>
<td>99.0</td>
<td>98.0</td>
<td>82.2</td>
<td>99.0</td>
<td>99.3</td>
<td>89.9</td>
<td>99.3</td>
<td>100.0</td>
</tr>
<tr>
<td>26</td>
<td>CRUDEOIL</td>
<td>89.2</td>
<td>84.8</td>
<td>87.9</td>
<td>85.2</td>
<td>73.6</td>
<td>89.7</td>
<td>64.9</td>
<td>78.4</td>
<td>85.3</td>
<td>84.8</td>
</tr>
<tr>
<td>27</td>
<td>SAND</td>
<td>93.9</td>
<td>94.9</td>
<td>93.9</td>
<td>93.9</td>
<td>93.9</td>
<td>93.9</td>
<td>81.9</td>
<td>93.9</td>
<td>94.9</td>
<td>94.9</td>
</tr>
<tr>
<td>28</td>
<td>HEMOPHILIA</td>
<td>85.6</td>
<td>85.6</td>
<td>82.8</td>
<td>85.6</td>
<td>83.9</td>
<td>85.6</td>
<td>78.9</td>
<td>86.7</td>
<td>86.7</td>
<td>85.6</td>
</tr>
<tr>
<td>29</td>
<td>COFFEE</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>92.9</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>30</td>
<td>OLITOS</td>
<td>89.1</td>
<td>73.6</td>
<td>70.4</td>
<td>83.1</td>
<td>80.0</td>
<td>94.0</td>
<td>58.0</td>
<td>77.2</td>
<td>87.6</td>
<td>87.6</td>
</tr>
<tr>
<td>31</td>
<td>FISH</td>
<td>92.6</td>
<td>92.9</td>
<td>92.9</td>
<td>96.4</td>
<td>85.2</td>
<td>100.0</td>
<td>88.7</td>
<td>89.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>32</td>
<td>HEARTDISEASE</td>
<td>69.9</td>
<td>65.2</td>
<td>63.2</td>
<td>68.8</td>
<td>66.2</td>
<td>69.7</td>
<td>66.1</td>
<td>67.3</td>
<td>68.0</td>
<td>68.0</td>
</tr>
</tbody>
</table>
Comparisons of classification methods

Principal Component Analysis

PC 1 - EV = 51.68%

PC 2 - EV = 15.71%

Methods:
- BNN
- KNN
- SVM/RBF
- N3
- QDA
- CAIMAN
- PLS-DA
- SVM/LIN
- LDA
- CART
- W
Comparisons of classification methods

Minimum Spanning Tree

- LDA
- SVM/LIN
- SVM/RBF
- PLS-DA
- QDA
- CART
- BNN
- CAIMAN
- KNN
- N3
- B
Regularized Hasse diagrams

$t^* = 0.92$

$t^* = 0.81$

$t^* = 0.73$

$t^* = 0.64$

$t^* = 0.58$

$t^* = 0.55$
PWR diagrams ($t^* = 0.5$)
PWR diagrams (t* = 0.6)
PWR diagrams \((t^* = 0.8)\)
Anilines data set

45 anilines described by 4 criteria:
1. log Kow (octanol-water partition coeff.)
2. log VP (vapor pressure)
3. Biodegradability (1: yes; 2: no)
4. PNEC (Predicted No-Effect Concentration)

Study focused on:
1. Hasse diagram (HD)
2. From HD to MonteCarlo ranking
3. From HD to Average ranking

Figure 3. Hasse diagram of the 45 anilines based on the 4 descriptors given in table 2. The single compounds are identified through their ID (cf. table 2).
Anilines: PWR diagram ($t^* = 0.5$)

Anilines - PWR ($t = 0.5 - H = 0.803$)

- Equal weights

- Biodeg
- Not-biodeg
Anilines: PWR diagram (t* = 0.88)

Anilines - PWR (t = 0.88 - H = 0.606)

Equal weights
Anilines: ranks comparison

Ranks for adverse effects

- PWR (t* = 0.5)
- PWR (t* = 0.7)
- Reference ranking
- MonteCarlo Linear Extensions
- Average ranking
Conclusions

➢ Possibility to weight the criteria
➢ Threshold selection offers different opportunities to rank the objects
➢ The Hasse transform from tournament table produces a family of regularized Hasse diagrams, thus also allowing a reduction of incomparabilities
➢ PWR is able to rank objects by a well founded theory
➢ PWR can remove inconsistencies from the tournament table
➢ PWR diagrams introduce a quantitative axis
➢ PWR diagrams can recover several incomparabilities present in the Hasse diagrams
➢ Statistical analysis can be performed on both the family of regularized-Hasse diagrams and the set of PWR rankings