CIVA FairPlay System Working Group Proposal – 2017

The development work discussed here was carried out by the FPS Working Group as a response to French proposal NP #2017 introduced at the CIVA 2016 plenary conference and Matthieu Roulet's paper "MRoulet Rebalancing CIVA Scoring Method", which together questioned whether a scoring system could be developed so that the competitors scores obtained in a programme would not depend upon the scoring of other competitors.

Summary
Though primarily concerned with the FairPlay statistical system the working group has reviewed many of the simpler methods used by other major sporting regimes such as swimming, gymnastics, high-board diving etc. which deliver fixed or stable individual scores from judges opinions, in case any could simplify or render obsolete the more sophisticated techniques employed by FPS to provide the fairest possible result taking into account all relevant data. In every case however by setting a "fixed" result as the primary criterion the balance and coordination of the underlying judgements are handled with less regard for bias and fairness, and for CIVA the outcome would very likely be unacceptable. Doug Lovell's IAC publication "Throwing Away High and Low" illustrates many examples from this minefield, though IAC continue with raw averaging at all US events except their annual Nationals, where FPS and ACRO have been employed throughout for several years.

By contrast a fundamental requirement of FPS methodology is that all stored scoring data within each programme should be assessed every time results are calculated in order to provide the most robust, fair and unbiased result possible, and therefore individual pilot score and rank refinements are inevitable as more data becomes available. It follows that if the existing FPS algorithms can be developed to stabilise these minor variations this would present a more appropriate solution than moving to a radically different system offering an inherently inferior level of fairness and accuracy.

FPS developments
Using existing championship contest files we have found that a significantly improved level of stability between successive results calculations as more pilots are added into the database can be achieved by replacing the existing Go/No-Go method of data rejection with immediate substitution of Fitted Values by one that provides smooth and proportional Fitted Value substitution over a range of uncertainty values. It can also be argued that this revised approach more closely mimics the normal human response to matters of judgement, in which personal deliberations normally move through a gradual transition from sure to unsure.

In the FairPlay process other items that can cause instability between successive calculations are:

a) The FPS "60% Rule" stipulates that if more than 60% of judges' marks for any pilot/figure combination fail their acceptance test then all must be rejected and replaced by their Fitted Value. This assessment is influenced by all other FPS considerations within each group, and on the margin can occasionally slip from one state to the other. The logic for this item is not

at fault, and in this area too a more flexible and proportionate solution has been found that provides improved stability.

b) When more than 30 pilots have flown a re-calculation is triggered if any scores are below 60% (in Knowns) or 50% (in Unknowns) that temporarily excludes the lower performing pilots and thereby avoids any influence that they might have on pilots with scores above these levels. Minor variations in scores can on occasion change the selection of pilots in the above/below sections, and inevitably there is some effect on the results.

     o Revising the 30 value to 11 for example would be more logical and also encompass smaller (e.g. glider) events as well as the larger power championships.

     o Discarding this system entirely is also a supportable move, as in reality the judges' opinions are influenced by everything they see during each programme and thus arbitrarily excluding some data for some of the time is not logically secure.

Further work to resolve these features are still being assessed to provide a recommended solution. Whatever is decided most appropriate will not however affect the underlying statistical evaluations that FPS uses to detect and address mark and score anomalies.

The W/G therefore proposes that the 'proportional' FPS method (FPS2) should be adopted by CIVA for use from 2018 forwards.
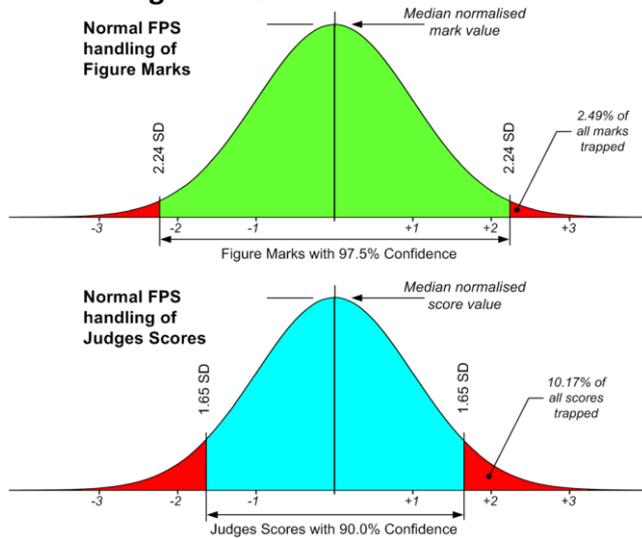
An explanation of the FPS judges mark and score handling process
The judges' marks are initially collated into groups within which the figures have notable similarities so that as far as possible 'like' is compared to 'like' and logical data comparisons can be established.

After the judge styles have been normalised (balanced) within each group, each data point is assessed by its Standard Deviation (SD) to measure its "fit" within the group with respect to each pilot and judge. In the existing FPS solution marks with SD greater than 2.24 (97.5% confidence) or scores with SD greater than 1.65 (90.0% confidence) are deemed unacceptable; these are replaced by a Fitted Value that is calculated after the rejected grade(s) have been removed, to match the judge's style.

In the proposed FPS2 solution however the transition from normalised mark to Fitted Value is made gradually, and in each case the final mark is the result of a smooth change from one to the other. Marks and scores are still subject to careful analysis throughout, but successive results iterations are significantly more stable and we have found that as more scoring data is added into the system the stability of pilots' scores is considerably improved and hence their rank positions are less likely to be affected. Whereas in the current FPS approach some 2.5% of all marks and 10.2% of all scores are likely to be rejected and the Fitted Value abruptly substituted, in the proportional FPS2 version with the SD's pairs that we have found to be most suitable just 1.7% of marks and 5.5% of scores transition smoothly from their normalised or 'balanced' value to their Fitted Value, and with these settings we have found the influence on results stability at its most effective.

The existing FPS solution

Normal FPS handling of Figure Marks
Median normalised mark value
2.24 SD | 2.24 SD
2.49% of all marks trapped
Figure Marks with 97.5% Confidence

Normal FPS handling of Judges Scores
Median normalised score value
1.65 SD | 1.65 SD
10.17% of all scores trapped
Judges Scores with 90.0% Confidence

Proportional FPS solution (FPS2)

Proportional FPS handling of Judges Figure Marks
■ Marks
■ Fitted Values
SD's 1.96 to 2.43
1.96 SD | 2.43 SD
100% Normalised
100% Fitted Value
2.96% of all marks trapped

Proportional FPS handling of Judges scores
■ Scores
■ Fitted Values
SD's 1.65 to 2.43
1.65 SD | 2.43 SD
100% Normalised
100% Fitted Value
3.98% of all scores trapped
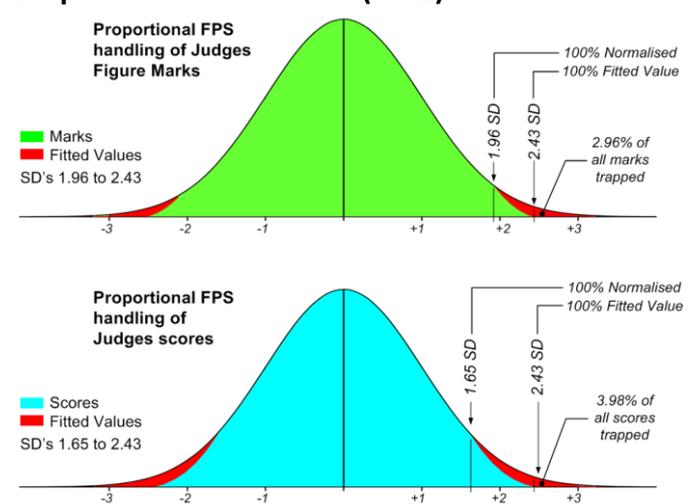
In our development of the proportional FPS2 method, from many trials we have found the following SD ranges to offer the most stable and repeatable advantages –

	SD low	SD high	Confidence %
Figures:	1.96	2.43	95.0 to 98.5
Scores:	1.65	2.43	90.0 to 98.5

100% Normalised
100% Fitted Value
1.65 SD
2.43 SD
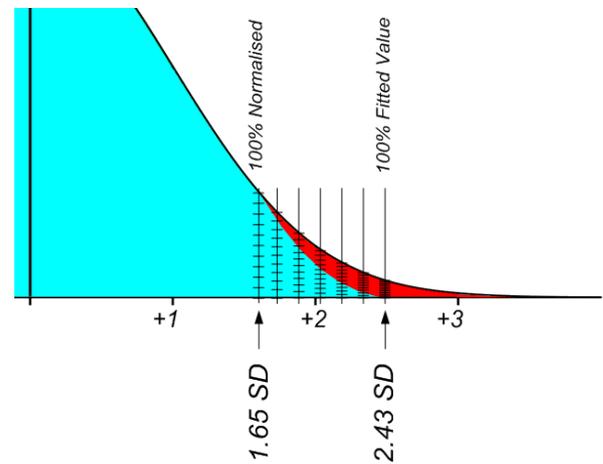
The broadened scope of this method is such that at each successive calculation of the results the additional judging data is unlikely to cause the marks and score SD's to vary significantly from their prior value and thus their influence on score (and hence rank) changes is of reduced concern.

Appendix

On the next pages are some graphic illustrations of the ranking output as produced by a modified version of the ACRO scoring software, using typical data from programmes at recent glider and power championships. They show the variations in scores and hence the rank positions of each flown pilot at each possible stage of the programme, starting with just 3 flights completed (the minimum that FPS can handle) right through to the last. Pilots have been added according to the Flying Order established at the event, so that each one appears in the results in the correct order.

When using any system that reviews the value of all available data every time it is run to create, as far as possible, a balanced and bias-free overall result, it is inevitable that successive re-evaluations of the data as the pool content increases will reach similar but slightly different conclusions. The Working Group is satisfied however that this modified process provides a robust improvement over the existing system, which has remained virtually unchanged since its first use in 2005.

In each case here it is very clear that the use of the proportional FPS2 method makes a significant contribution to the stability of calculated successive rank orders by comparison with the established

FPS system, and markedly reduces instances of rank exchanges between pilots with similar scores. It can also be seen that the final scores provided by the proportional system are very similar to those from current FPS methodology.

We have also concluded that the proportional FairPlay system described here is likely to provide a marginally more realistic assessment of all the available data, and can therefore be accepted by CIVA as a beneficial change that implies no impropriety toward prior established results.

Note that when the revised logic of FPS2 is incorporated into CIVA's scoring systems it will be necessary that the software retains the original FPS process so that this can be used to repeat championship results from years prior to 2018.

Nick Buckenham
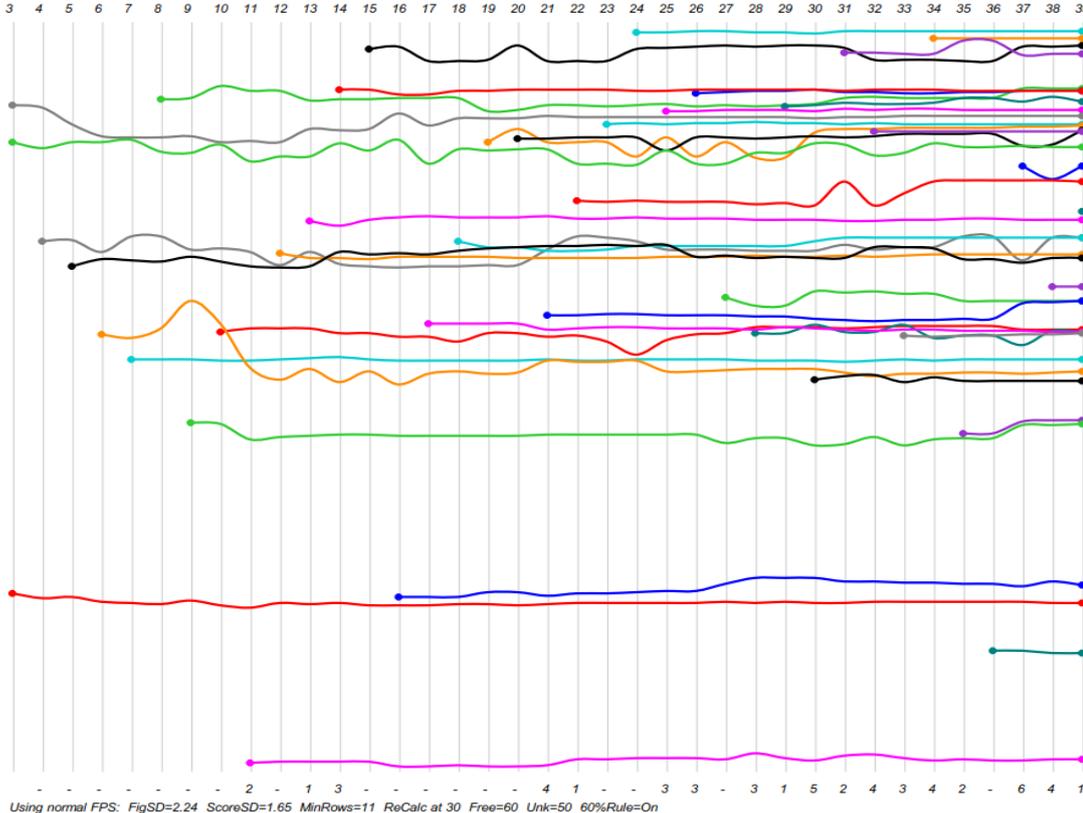CIVA FPS Working Group chairman

FPS W/G Members:
Gilles Guillemard   FRA
Doug Lovell   USA
Mikhail Mamistov   RUS
Vladimir Machula   CZE

Progression of Pilot Scores Report: Free Unknown, 7th WAGAC
Data is in scores order per cycle, F/O steps 3 to 39 of 39

FPS: Normal 2.24 & 1.65, Re-Calc @ 30 per CIVA

Rank	Scores	Pno	Name
1	1390.51	038	Sebastian JANSSON
2	1383.90	002	Ales FERRA
3	1376.91	015	Dávid JÓZSA
4	1369.00	011	Marvin WOLTERING
5	1336.87	030	Agata NYKAZA
6	1334.88	031	Miroslav WRZESNIEWSKI
7	1334.43	027	Michal KLIMASZEWSKI
8	1324.62	036	Jonas LANGENEGGER
9	1317.51	020	Daniele FERRARESE
10	1311.00	018	Claudio BRASAOLA
11	1303.16	005	Erwin GEORGE
12	1302.15	028	Piotr LEWANDOWSKI
13	1297.86	014	Miklós HOÓS
14	1297.05	035	Ciprian LUPAS
15	1282.70	012	Dániel CZITÁN
16	1264.90	003	Lis ARNEBERG
17	1250.27	004	Jukka PELTO-AHO
18	1221.94	006	Eric LANQUETIN
19	1214.66	022	Marcello TEDESCHI
20	1197.58	016	Péter SZABÓ
21	1197.09	034	Valentin HOTA
22	1181.64	019	Massimo CICERI
23	1178.64	039	Gustav SALMINEN
24	1151.79	021	Daniele MELONI
25	1138.65	033	Alin DROBOS
26	1138.02	009	Markus PÖNICKE
27	1111.34	037	Pekka HAVBRANDT
28	1109.54	026	Wojciech GALUSZKA
29	1109.32	024	Jacek BOGATKO
30	1108.32	013	Pál HÁRSFALVI
31	1083.72	017	Zsolt SZABÓ
32	1071.82	023	Takashi SAKAI
33	1063.46	007	Jürgen GRYGIER
34	1026.36	029	Paulina MATECKA
35	1023.05	001	Bernhard BEHR
36	872.13	010	Sandro RUTZ
37	856.12	032	Lorand DAROCZI
38	808.93	025	Slavomir CICHON
39	709.35	008	Dennis KONIECZNY

Using normal FPS: FigSD=2.24 ScoreSD=1.65 MinRows=11 ReCalc at 30 Free=60 Unk=50 60%Rule=On

Total rank changes: 52

This report created at 14:46 on Sunday 26 March 2017

Progression of Pilot Scores Report: Free Unknown, 7th WAGAC
Data is in scores order per cycle, F/O steps 3 to 39 of 39

FPS2: Figs 1.96 to 2.43, Scores 1.65 to 2.43, Re-Calc @ 11
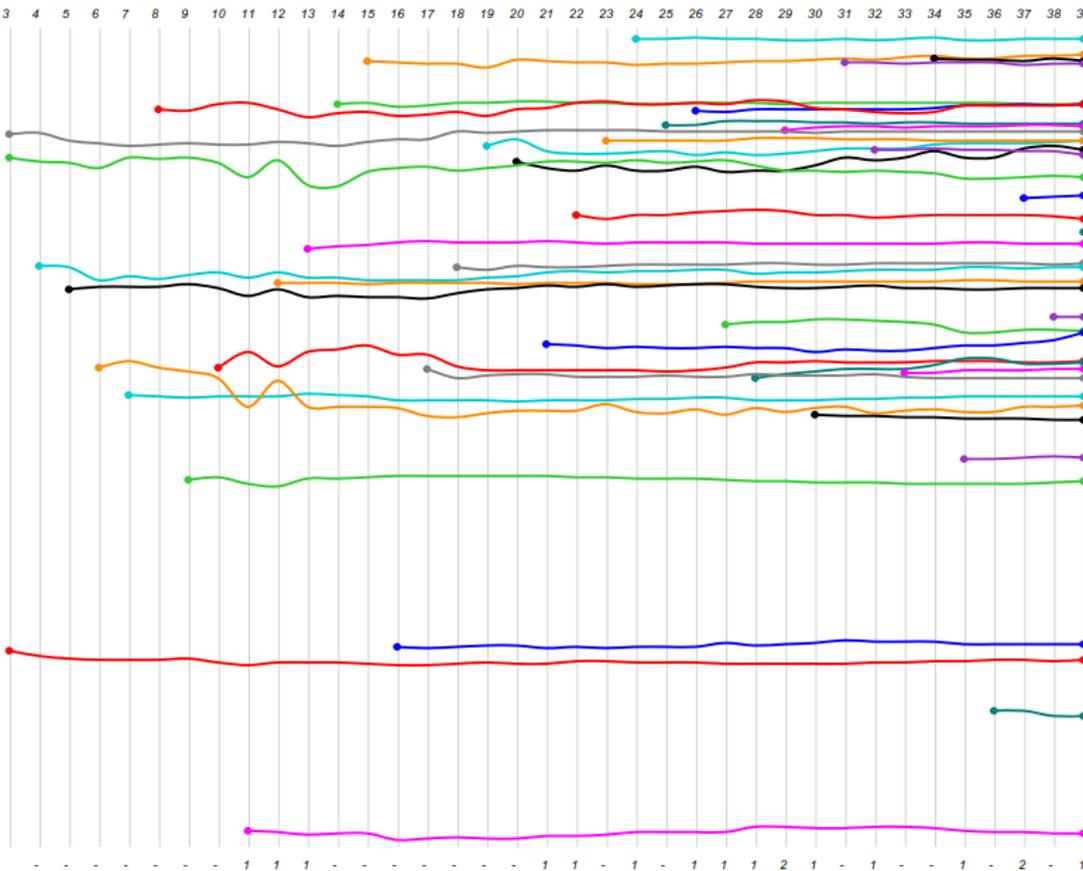
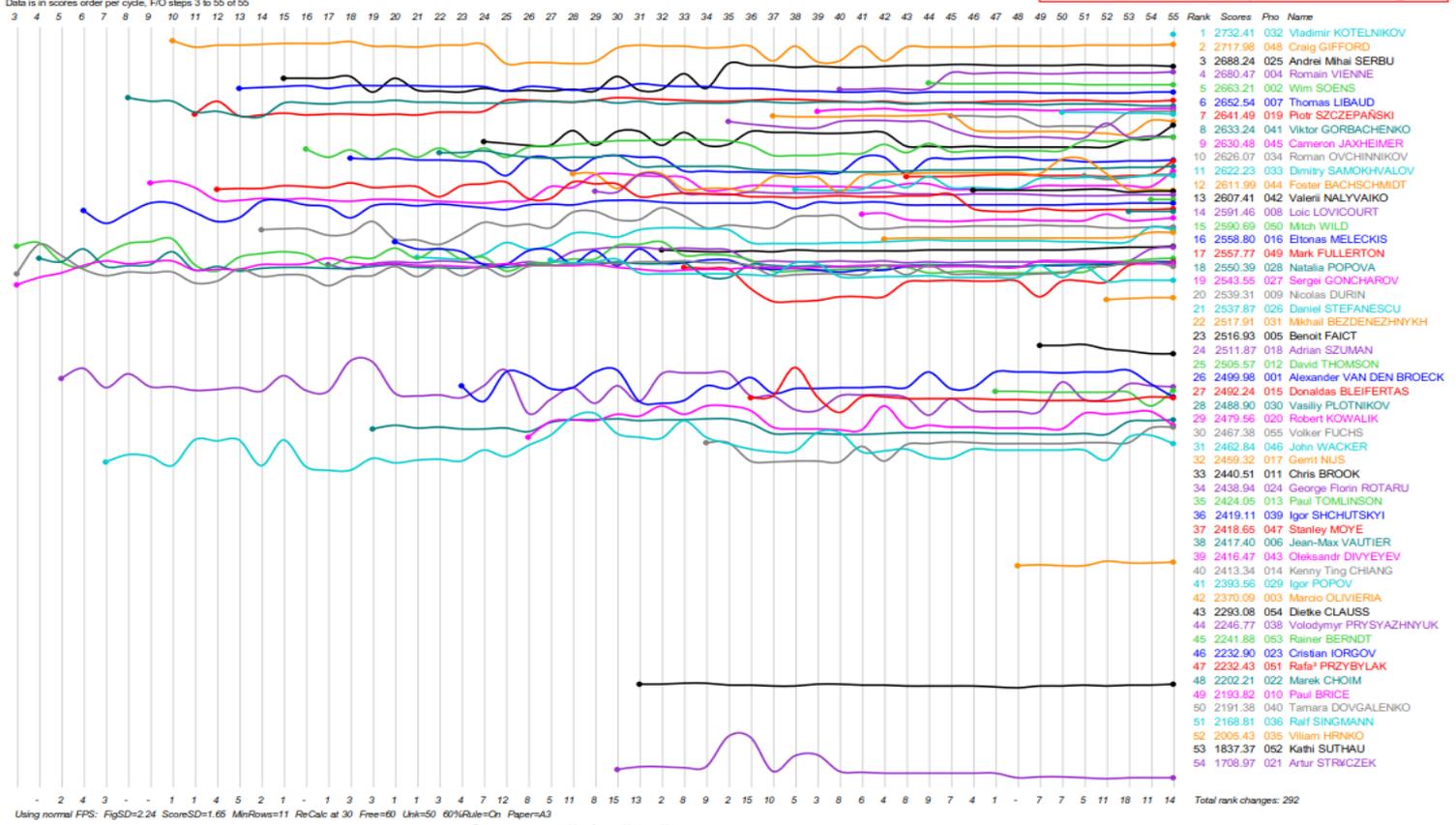| Rank	Scores	Pno	Name
1	1390.28	038	Sebastian JANSSON
2	1376.50	015	Dávid JÓZSA
3	1371.33	002	Ales FERRA
4	1368.80	011	Marvin WOLTERING
5	1334.78	027	Michal KLIMASZEWSKI
6	1334.55	031	Miroslav WRZESNIEWSKI
7	1334.29	030	Agata NYKAZA
8	1317.59	020	Daniele FERRARESE
9	1315.74	036	Jonas LANGENEGGER
10	1311.23	018	Claudio BRASAOLA
11	1303.31	028	Piotr LEWANDOWSKI
12	1303.10	005	Erwin GEORGE
13	1295.33	014	Miklós HOÓS
14	1290.90	035	Ciprian LUPAS
15	1272.19	012	Dániel CZITÁN
16	1256.23	003	Lis ARNEBERG
17	1235.31	004	Jukka PELTO-AHO
18	1224.73	006	Eric LANQUETIN
19	1214.69	022	Marcello TEDESCHI
20	1197.56	034	Valentin HOTA
21	1193.99	016	Péter SZABÓ
22	1182.12	019	Massimo CICERI
23	1176.46	039	Gustav SALMINEN
24	1151.64	021	Daniele MELONI
25	1139.70	033	Alin DROBOS
26	1137.76	009	Markus PÖNICKE
27	1113.26	037	Pekka HAVBRANDT
28	1112.23	026	Wojciech GALUSZKA
29	1107.20	013	Pál HÁRSFALVI
30	1099.33	024	Jacek BOGATKO
31	1083.81	017	Zsolt SZABÓ
32	1074.97	023	Takashi SAKAI
33	1063.24	007	Jürgen GRYGIER
34	1031.05	029	Paulina MATECKA
35	1010.12	001	Bernhard BEHR
36	870.48	010	Sandro RUTZ
37	856.41	032	Lorand DAROCZI
38	808.51	025	Slavomir CICHON
39	707.86	008	Dennis KONIECZNY

Using proportional FPS: FigSDLo=1.96 FigSDHi=2.43 ScoreSDLo=1.65 ScoreSDHi=2.43 MinRows=11 ReCalc at 11 Free=60 Unk=50

Total rank changes: 17

This report created at 17:03 on Monday 9 October 2017

FPS: Normal 2.24 & 1.65 per CIVA, Re-Calc @ 30

Rank	Scores	Pno	Name
1	2732.41	032	Vladimir KOTELNIKOV
2	2717.98	048	Craig GIFFORD
3	2688.24	025	Andrei Mihai SERBU
4	2680.47	004	Romain VIENNE
5	2663.21	002	Wim SOENS
6	2652.54	007	Thomas LIBAUD
7	2641.49	019	Piotr SZCZEPAŃSKI
8	2633.24	041	Viktor GORBACHENKO
9	2630.48	045	Cameron JAXHEIMER
10	2626.07	034	Roman OVCHINNIKOV
11	2622.23	033	Dimitry SAMOKHVALOV
12	2611.99	044	Foster BACHSCHMIDT
13	2607.41	042	Valerii NALYVAIKO
14	2591.46	008	Loic LOVICOURT
15	2590.69	050	Mitch WILD
16	2558.80	016	Eltonas MELECKIS
17	2557.77	049	Mark FULLERTON
18	2550.39	028	Natalia POPOVA
19	2543.55	027	Sergei GONCHAROV
20	2539.31	009	Nicolas DURIN
21	2537.87	026	Daniel STEFANESCU
22	2517.91	031	Mikhail BEZDENEZHNYKH
23	2516.93	005	Benoit FAICT
24	2511.87	018	Adrian SZUMAN
25	2505.57	012	David THOMSON
26	2499.98	001	Alexander VAN DEN BROECK
27	2492.24	015	Donaldas BLEIFERTAS
28	2488.90	030	Vasiliy PLOTNIKOV
29	2479.56	020	Robert KOWALIK
30	2467.38	055	Volker FUCHS
31	2462.84	046	John WACKER
32	2459.32	017	Gerrit NIJS
33	2440.51	011	Chris BROOK
34	2438.94	024	George Florin ROTARU
35	2424.05	013	Paul TOMLINSON
36	2419.11	039	Igor SHCHUTSKYI
37	2418.65	047	Stanley MOYE
38	2417.40	006	Jean-Max VAUTIER
39	2416.47	043	Oleksandr DIVYEYEV
40	2413.34	014	Kenny Ting CHIANG
41	2393.56	029	Igor POPOV
42	2370.09	003	Marcio OLIVIERIA
43	2293.08	054	Dietke CLAUSS
44	2246.77	038	Volodymyr PRYSYAZHNYUK
45	2241.88	053	Rainer BERNDT
46	2232.90	023	Cristian IORGOV
47	2232.43	051	Rafa³ PRZYBYLAK
48	2202.21	022	Marek CHOIM
49	2193.82	010	Paul BRICE
50	2191.38	040	Tamara DOVGALENKO
51	2168.81	036	Ralf SINGMANN
52	2005.43	035	Viliam HRNKO
53	1837.37	052	Kathi SUTHAU
54	1708.97	021	Artur STRYCZEK

Total rank changes: 292

FPS2: Figs 1.96 to 2.43, Scores 1.65 to 2.43, Re-Calc @ 11

Rank	Scores	Pno	Name
1	2746.59	032	Vladimir KOTELNIKOV
2	2720.34	048	Craig GIFFORD
3	2683.20	025	Andrei Mihai SERBU
4	2680.57	004	Romain VIENNE
5	2664.97	002	Wim SOENS
6	2661.98	007	Thomas LIBAUD
7	2643.10	045	Cameron JAXHEIMER
8	2636.13	041	Viktor GORBACHENKO
9	2627.61	019	Piotr SZCZEPAŃSKI
10	2620.30	033	Dimitry SAMOKHVALOV
11	2620.11	034	Roman OVCHINNIKOV
12	2608.52	044	Foster BACHSCHMIDT
13	2601.54	042	Valerii NALYVAIKO
14	2594.95	008	Loic LOVICOURT
15	2582.95	050	Mitch WILD
16	2565.32	049	Mark FULLERTON
17	2560.51	016	Eltonas MELECKIS
18	2554.54	028	Natalia POPOVA
19	2538.45	009	Nicolas DURIN
20	2535.98	031	Mikhail BEZDENEZHNYKH
21	2533.60	026	Daniel STEFANESCU
22	2526.86	027	Sergei GONCHAROV
23	2520.20	005	Benoit FAICT
24	2518.56	018	Adrian SZUMAN
25	2511.72	015	Donaldas BLEIFERTAS
26	2507.37	012	David THOMSON
27	2507.20	001	Alexander VAN DEN BROECK
28	2489.28	030	Vasiliy PLOTNIKOV
29	2484.39	020	Robert KOWALIK
30	2464.73	055	Volker FUCHS
31	2458.76	011	Chris BROOK
32	2456.45	017	Gerrit NIJS
33	2451.28	046	John WACKER
34	2430.37	024	George Florin ROTARU
35	2416.36	039	Igor SHCHUTSKYI
36	2415.84	043	Oleksandr DIVYEYEV
37	2415.58	013	Paul TOMLINSON
38	2410.96	014	Kenny Ting CHIANG
39	2402.96	006	Jean-Max VAUTIER
40	2398.59	047	Stanley MOYE
41	2392.82	029	Igor POPOV
42	2371.03	003	Marcio OLIVIERIA
43	2314.81	054	Dietke CLAUSS
44	2271.01	023	Cristian IORGOV
45	2247.65	038	Volodymyr PRYSYAZHNYUK
46	2238.27	053	Rainer BERNDT
47	2223.87	051	Rafa³ PRZYBYLAK
48	2192.80	022	Marek CHOIM
49	2187.16	010	Paul BRICE
50	2183.27	036	Ralf SINGMANN
51	2163.48	040	Tamara DOVGALENKO
52	2004.08	035	Viliam HRNKO
53	1838.04	052	Kathi SUTHAU
54	1712.14	021	Artur STRYCZEK

Total rank changes: 80