



# FAI Sporting Code

*Fédération  
Aéronautique  
Internationale*

---

## **Section 6**

# International Aerobatic Events: Statistical Method for Processing Scores

*Maison du Sport International  
Av. de Rhodanie 54  
CH-1007 Lausanne  
(Switzerland)  
Tél. +41 (0)21 345 10 70  
Fax +41 (0)21 345 10 77  
E-mail: [sec@fai.org](mailto:sec@fai.org)  
Web: [www.fai.org](http://www.fai.org)*

## Version 2018-1

Effective 1st January 2018



**FEDERATION AERONAUTIQUE INTERNATIONALE**  
**MSI - Avenue de Rhodanie 54 – CH-1007 Lausanne – Switzerland**

---

Copyright 2018

All rights reserved. Copyright in this document is owned by the Fédération Aéronautique Internationale (FAI). Any person acting on behalf of the FAI or one of its Members is hereby authorised to copy, print, and distribute this document, subject to the following conditions:

- 1. The document may be used for information only and may not be exploited for commercial purposes.**
- 2. Any copy of this document or portion thereof must include this copyright notice.**
- 3. Regulations applicable to air law, air traffic and control in the respective countries are reserved in any event. They must be observed and, where applicable, take precedence over any sport regulations.**

Note that any product, process or technology described in the document may be the subject of other Intellectual Property rights reserved by the Fédération Aéronautique Internationale or other entities and is not licensed hereunder.



## **TABLE OF CONTENTS**

1.	THE CIVA FAIR PLAY SYSTEM - PURPOSE .....	1
2.	OVERVIEW .....	1
3.	PRE-PROCESSING .....	1
3.1.	Dealing with Hard Zeros and Missed Figures .....	1
3.2.	Identifying Figure Grades for Analysis.....	2
3.3.	Grouping Figure Grades for Analysis .....	2
3.4.	Confirmation of Hard Zero .....	4
3.5.	Treatment of Other “HZ” or of “A” Grades .....	4
3.6.	Treatment of Perception Zero Grades.....	4
4.	DEFINITIONS .....	4
4.1.	The Basic Data Values .....	4
5.	GROUP PROCESSES .....	5
5.1.	Normalisation of a Data Group .....	5
5.2.	Derivation of Fitted Values .....	6
5.3.	Assessment of Anomalous Grades .....	6
5.4.	Treatment of Anomalous Grades .....	6
5.5.	Second Normalisation of the Group .....	7
5.6.	Replacement of Missing and other anomalous Grades .....	7
5.7.	Assembly of Processed Grades by Pilot .....	8
6.	SEQUENCE PROCESSES.....	8
6.1.	Normalisation of Sequence Scores .....	8
6.2.	Derivation of Sequence Fitted Values .....	9
6.3.	Assessment of Sequence Anomalies .....	9
6.4.	Final Sequence Scores .....	10
7.	PROCESS SUMMARY .....	10
7.1.	The process carries out the following analytical steps: .....	10
8.	JUDGING PERFORMANCE ANALYSIS .....	10
8.1.	Ranking Index (RI).....	10
8.2.	Low and High Scoring Anomalies.....	11
8.3.	Discrimination .....	11
8.4.	Hard Zero Anomalies.....	12
8.5.	Sequence Score Anomalies .....	12
	<b>RECORD OF AMENDMENTS .....</b>	<b>13</b>



## **1. The CIVA Fair Play System - Purpose**

- 1.1.1.1. Calculation of grades and scores for an aerobatic competition Programme using a mathematical process to give equal importance to all judges, while replacing anomalous grades with statistically Fitted Values.
- 1.1.1.2. Prior to 2018 the CIVA Fair Play System employed single SD (standard deviation) values for figure (2.24 = 97.5% confidence) and sequence (1.65 = 90.0% confidence) score testing purposes; in each case rejected values were simply replaced by their Fitted Values. From 2018 forwards the processing method is revised such that data identified as anomalous is replaced in proportion to the level of anomaly assessed, using low/high SD pairs 1.96 (95.0%) and 2.43 ((98.5%) for figures and 1.24 (78.5%) and 1.65 (90.0%) for sequence scores. The “60% Rule” at para 5.4.1.4 has been revised to incorporate an improved assessment of the proportionally adjusted figure data. The recalculation system triggered if any sequence scores are below 60% (known) or 50% (unknown) with 30 or more pilots is discontinued. In the calculation of each judges’ Ranking Index (RI) all average (A) grades are resolved to the mean of the scoring judges grades, calculated to the nearest half mark; in this assessment HZ’s and PZ’s are counted, and interpreted with zero value.

## **2. Overview**

The rating of a pilot performance for a given flight is an amount of points arising from two separate sources:

- 2.1.1.1. An evaluation of the quality of flown figures and of a flight’s positioning with a grade given by judges observing the flight, on a scale ranging from 0 to 10 in increments of 0.5. These grades are multiplied by difficulty coefficients for each figure and added to derive a score for the programme for each pilot.
- 2.1.1.2. Penalties arising from height or time infringements and/or interruptions of the program sequence and other disciplinary actions.
- 2.1.1.3. The scores from 2.1.1.1 are subject to random and systematic errors due to the inevitable lack of exactness of judging. The purpose of the Fair Play system is to reduce the effect of those errors to a minimum. The penalties from 2.1.1.2 are not subject to the same errors and are simply subtracted from the scores results 2.1.1.1 after they have been calculated as described below.

## **3. Pre-Processing**

### **3.1. Dealing with Hard Zeros and Missed Figures**

- 3.1.1.1. Prior to the scoring data being entered into the computer, the Chief Judge must ascertain the validity of Hard Zero grades. If a figure is determined to have been a Confirmed Hard Zero, this must be designated by the Chief Judge. The grades given to that figure by the grading judges must not be altered prior to being input into the computer.
- 3.1.1.2. For a figure determined not to be a Confirmed Hard Zero, any “HZ” grade given by a grading judge must remain unaltered prior to data entry into the scoring computer.
- 3.1.1.3. Figures that have been missed by a grading judge must be marked “A”. These missing grades will be replaced automatically by the Fair Play system.



### 3.2. Identifying Figure Grades for Analysis

- 3.2.1.1. Prior to the start of the Programme, the data input to the scoring computer will include the number of pilots, the number of figures (including positioning) each pilot will fly, the K-factors of each figure and the number of judges for the programme.
- 3.2.1.2. Each figure will be given a full identifying number in the format **kkkffpp** where:
- kkk* is the K-factor, with leading zeros if necessary, e.g. “037” if K-factor =37
  - ff* is the figure number, with leading zeros if necessary, and
  - pp* is the pilot number, with leading zeros if necessary.
- 3.2.1.3. Note that the number *pp* allocated to a pilot must remain the same throughout a contest and should not be confused with the flight order number any pilot may be allocated for a particular programme.

### 3.3. Grouping Figure Grades for Analysis

- 3.3.1.1. Statistical manipulation must only be carried out on sets of data of reasonable size. Furthermore, such analysis is best conducted on sets of data that share similar source characteristics. To meet these requirements, the grading data from a programme must be combined into appropriate groups.
- 3.3.1.2. For the purpose of the Fair Play analysis, data will be arranged in groups in the following generalised format:

K-factor	Figure #	Pilot #	Judge 1	Judge 2	...	...	Judge j
$Kkk_1$	<i>ff</i>	<i>pp</i>	$Grade_{1,1}$				
$Kkk_2$	<i>ff</i>	<i>pp</i>					
$Kkk_3$	<i>ff</i>	<i>pp</i>					
...	...	...					
...	...	...					
$kkk_n$	<i>ff</i>	<i>pp</i>					$Grade_{n,j}$

- 3.3.1.3. In such a data set, the arrangement of rows will be by ascending value of the full figure Identification Number **kkkffpp**. In compulsory programmes, Known and Unknown, all pilots fly the same figures and the number of rows per data group will normally be the same as the number of pilots. This means that each data group in a compulsory sequence will correspond to a figure of that sequence flown by all pilots, in the form:

K-factor	Figure #	Pilot #	Judge 1	Judge 2	...	...	Judge j
<i>kkk</i>	Figure 1	Pilot 1	$Grade_{1,1}$				
<i>kkk</i>	Figure 1	Pilot 2					
<i>kkk</i>	Figure 1	Pilot 3					
...	...	...					
...	...	...					



kkk	Figure 1	Pilot p								Grade <sub>p,j</sub>
-----	----------	---------	--	--	--	--	--	--	--	----------------------

3.3.1.4. Exceptionally, if the number of pilots is less than 11, the target number of rows for each group (NrmGrp) will be as follows:

Number of Pilots	2	3	4	5	6	7	8	9	10
Group Size	12	12	12	15	12	14	16	18	20

3.3.1.5. Free Known Programmes

- a) In Free Known Programmes, every year CIVA selects five 'Known' or master figures for each category using the code letters "A" to "E" to which each competitor must add five 'Free' or complementary figures each coded "F" to create their sequence.
- b) The Positioning (P and G) figure codes will be set to "1".

3.3.1.6. Grouping of Free Known figures

- a) In the figures coded "A" to "E" the data groups will be formed from within each figure code, unless N<sub>p</sub> is less than 11. The target number of rows for each group (NrmGrp) will be the number of pilots whose flights have been judged, while the minimum group size (MinGrp) will remain 11 rows. When N<sub>p</sub> < 11, then NrmGrp will be as tabulated in 3.3.1.4, and a group may contain figures from more than one figure code.
- b) The data rows with 'Free' or complementary figures coded "F" will be sorted by increasing K value. Where figures have the same K value they will be further sorted by the figure number within the pilots' sequence.
- c) In the figures coded "F" the group size will equal the number of pilots, containing every pilot once with the pilots' lowest remaining K value / figure number. A group may contain more than one K value. The target number of rows for each group (NrmGrp) will be the number of pilots whose flights have been judged, while the minimum group size (MinGrp) will remain 11 rows. When N<sub>p</sub> < 11, then NrmGrp will be as tabulated in 3.3.1.4, and a group may contain every pilot more than once.
- d) In the Positioning and Harmony data sets, the group size will equal the number of pilots, i.e. each will contain the complete Positioning or Harmony data set. If the number of pilots (N<sub>p</sub>) whose flights have been judged is < 11, however, then these data sets will be combined into a group containing them both.

3.3.1.7. Free Unknown Programmes

- a) In Free Unknown Programmes, the International Jury will give the ten common figures code letters "A" to "J" at the time of submission, and these letters must be maintained as the identifier for that figure in each sequence.
- b) The Jury will annotate the linking figures in each sequence, from a single link up to a maximum of four, with the figure codes "L1", "L2", "L3" and "L4" sequentially.
- c) The Positioning (P and G) figure codes will be set to "1".

3.3.1.8. Grouping of Free Unknown figures

- a) In the figures with codes "A" to "J", the data groups will be formed from within each figure code, unless N<sub>p</sub> is less than 11. The target number of rows for each group (NrmGrp) will be the number of pilots whose flights have been judged, while the minimum group size (MinGrp) will remain 11 rows. When N<sub>p</sub> < 11, then NrmGrp will be as tabulated in 3.3.1.4, and a group may contain figures from more than one figure code.



- b) In the linking figures coded “L1” to “L4”, the data groups will be formed from within each figure code, unless  $N_p$  is less than 11. The target number of rows for each group (NrmGrp) will be the number of pilots whose flights have been judged, while the minimum group size (MinGrp) will remain 11 rows. When  $N_p < 11$ , then NrmGrp will be as tabulated in 3.3.1.4, and a group may contain figures from more than one linking figure code.
- c) In the Positioning and Harmony data sets, the group size will equal the number of pilots, i.e. each will contain the complete Positioning or Harmony data set. If the number of pilots ( $N_p$ ) whose flights have been judged is  $< 11$ , however, then these data sets will be combined into a group containing them both.

#### 3.3.1.9. Free programmes

With effect from 2016 the CIVA Known and Free programmes are discontinued. Instructions regarding the coding of data groups for these legacy formats using the Fair Play system can be found in CIVA Section 6 documents from 2015 and before.

### 3.4. Confirmation of Hard Zero

- 3.4.1.1. The first stage of processing is to set to “HZ” all numerical grades given to a figure subsequently deemed to be a Confirmed Hard Zero by the Chief Judge. Any grade thus reduced to “HZ” must result in an increment to the particular judge’s record for determining the HZ anomaly count of the Judging Performance Analysis.
- 3.4.1.2. Once confirmed Hard Zeros have been implemented, each pilot’s score sheet should be printed and made available for inspection along with the judges grading sheets.

### 3.5. Treatment of Other “HZ” or of “A” Grades

- 3.5.1.1. If a figure is not deemed to be a Confirmed Hard Zero, any “HZ” or “A” grades given for that figure must be treated as missing data points. Such grades will therefore be excluded from the calculation of means or standard deviations until such time as they are replaced later in the process. For each “HZ” grade that is not confirmed, an increment will be made to the judge’s HZ anomaly count.

### 3.6. Treatment of Perception Zero Grades

- 3.6.1.1. Perception Zero grades are not subject to the same confirmation process as Hard Zeros. They are generally treated as valid numerical grades in the same way as non-zero grades. However, Perception Zero grades should not influence the normalisation of non-zero grades that is described below.

## 4. Definitions

### 4.1. The Basic Data Values

- 4.1.1.1. Define the Raw Grades, for a given sequence, as:

$S(ff, pp, j)$

This is the Grade awarded by Judge  $j$  to Pilot  $pp$  flying Figure  $ff$ .

- 4.1.1.2. These Grades are then divided into semi-homogeneous Groups as defined above, and are now defined as:

$R_g(fp, j)$

This is the Grade awarded by Judge  $j$  to (Pilot  $p$  flying Figure  $f$ ) in Group  $g$ , and is represented physically by a rectangular array of numbers where  $fp$  is the row index and  $j$  is the column index.



4.1.1.3. There should also be a count indicator of values 0 and 1 to indicate 0 for any SZ, HZ or A values. These are designated:  $N_g(fp, j)$

4.1.1.4. Counts

a) Pilot Count = No. Judges who score this pilot/figure combination

$$C_g(fp, *) = \sum_j \{N_g(fp, j)\} \quad (1)$$

b) Judge Count = No. Pilot/figures scored by this judge

$$C_g(*, j) = \sum_{fp} \{N_g(fp, j)\} \quad (2)$$

c) Overall Count = Total number of Scores

$$C_g(*, *) = \sum_{fp,j} \{N_g(fp, j)\} \quad (3)$$

4.1.1.5. Mean Values

a) Pilot Mean

$$mR_g(fp, *) = \sum_j \{R_g(fp, j)\} / C_g(fp, *) \quad (4)$$

b) Judge Mean

$$mR_g(*, j) = \sum_{fp} \{R_g(fp, j)\} / C_g(*, j) \quad (5)$$

c) Overall Mean

$$mR_g(*, *) = \sum_{fp,j} \{R_g(fp, j)\} / C_g(*, *) \quad (6)$$

4.1.1.6. Standard Deviations

a) Judge Standard Deviation

$$sdR_g(*, j) = \sqrt{[\sum_{fp} \{R_g(fp, j)\}^2 - C_g(*, j) * \{mR_g(*, j)\}^2] / [C_g(*, j) - 1]} \quad (7)$$

b) Average Judge Standard Deviation

$$sdR_g(*, ) = \sum_j \{sdR_g(*, j)\} / J \quad (8)$$

## **5. Group Processes**

### **5.1. Normalisation of a Data Group**

a) The first stage of the analysis is to Normalise the non-zero grades in the data group to give each judge's column of grades the same standard deviation. This will give equal importance to each judge's opinion. In the normalisation formula:

b)  $Norm1_g(fp,j)$  is the Normalised grade to replace the Raw grade,  
 $sdR_g(*,j)$  is the standard deviation for a judge's Raw grades in this group,  
 $sdR_g(*, *)$  is the standard deviation for all the Raw grades in this group from all judges,  
 and

$$Norm1_g(fp,j) = mR_g(*,*) + [R_g(fp,j) - mR_g(*,j)] * sdR_g(*, ) / sdR_g(*,j) \quad (9)$$

5.1.1.2. If the result of formula (7) or (8) is zero, then formula (9) cannot be applied and the grades for this judge, or this group, should be set to the overall mean for this group. If the result of formula (9) is less than zero, then it should be set at zero.

5.1.1.3. Perception Zero grades are excluded from this normalisation process because, for each judge, these form part of a second mode of distribution of raw grades. After the non-zero grades are normalised, the Perception Zero grades are set to 0.0 so that they are included in the process of determining Fitted Values and figure anomalies. Hence:





$$\text{If } Rg(fp,j) = 0.0, \text{ Then Norm1}_g(fp,j) = 0.0 \quad (9a)$$

## 5.2. Derivation of Fitted Values

5.2.1.1. Within the data group, a Fitted Value for a figure grade for a pilot is the grade that you would expect a particular judge to give a particular pilot/figure combination, based on an analysis of all the judges' grades for all the pilot/figure combinations in the group, including numerical zeros (SZ) but excluding factual zeros (HZ). In the Fitted Value formula:

- a)  $FV1_g(fp,j)$  is the Fitted Value derived from  $\text{Norm1}_g(fp,j)$ ,  
 $m\text{Norm1}_g(*,j)$  is the mean of the Normalised numerical grades in the group for that judge  
 $m\text{Norm1}_g(fp,*)$  is the mean of the Normalised numerical grades in the group for that pilot/figure,  
 $m\text{Norm1}_g(*,*)$  is the mean of all the Normalised numerical grades for that group for all judges, and

$$FV1_g(fp,j) = m\text{Norm1}_g(*,j) + m\text{Norm1}_g(fp,*) - m\text{Norm1}_g(*,*) \quad (10)$$

## 5.3. Assessment of Anomalous Grades

The normalised grades in each group must be tested for anomalies caused by judging error or partiality.

5.3.1.1. The Uncertainty of Any Individual Data Point

- a) A data point (grade) will be considered anomalous if its uncertainty exceeds a given threshold value. This uncertainty is derived by a two-way analysis of variance and starts with the calculation of the Residual for each data point. In the Residual formula:
- b)  $\text{Res1}_g(fp,j)$  is the Residual value for each data point in the group after the first normalisation, and,

$$\text{Res1}_g(fp,j) = \text{Norm1}_g(fp,j) - FV1_g(fp,j) \quad (11)$$

- c)  $\text{RSS1}_g$  is the Residual Sum of Squares for the data group after normalisation and,

$$\text{RSS1}_g = \sum_{fp,j} \{\text{Res1}_g(fp,j)\}^2 \quad (12)$$

5.3.1.2. The Degrees of Freedom of the data group is determined by:

- a)  $D_g$  is the value of the Degrees of Freedom of the data group  
 $FP_g$  is the number of pilot/figure rows in the group  
 $J_g$  is the number of judges in the programme (columns in the data group)  
 $Nm_g$  is the number of missing values (HZ or A) in the group, and

$$D_g = \{[FP_g - 1] * [J_g - 1]\} - Nm_g \quad (13)$$

5.3.1.3. The Residual Standard Deviation of the data group,  $\text{RSD1}_g$ , is determined by:

$$\text{RSD1}_g = \text{sqrt}\{\text{RSS1}_g / D_g\} \quad (14)$$

5.3.1.4. Finally, the uncertainty of each individual data point,  $U1_g(fp,j)$  is calculated:

$$U1_g(fp,j) = \text{ABS}\{\text{Res1}_g(fp,j)\} / \text{RSD1}_g \quad (15)$$

## 5.4. Treatment of Anomalous Grades

5.4.1.1. If the uncertainty of an individual grade,  $U1_g(fp,j)$ , exceeds 2.43 it has an uncertainty of approximately 98.5%. This degree of anomaly, or more, is to be expected in the case of a small number of Perception Zeros for a figure which generally attracts a majority of high grades. Similarly, such an anomaly might occur if a single judge missed a large pilot error that led all other judges to award a very low grade. Anomalies such as this should be treated



as though they were missing values. This treatment will give the benefit of the doubt to the pilot in situations where it is possible that a very significant judging error has been made.

5.4.1.2. The raw grade for any data point showing such an anomaly should be set to “Missing” in the original Raw Data  $R_g(fp, j)$  – call it  $R2_g(fp, j)$ . The judge concerned should have an increment made to Low or High Score anomaly count in the Judging Performance Analysis, as appropriate, for each grade replaced.

5.4.1.3. When making judgements based on the perception of the quality of flick rolls or spins, the panel of judges might produce a series of grades in which the distribution is bi-modal rather than Gaussian. For example, a set of grades might possibly include a number of Perception Zeros and a number of high grades. In extremely rare cases, this difference of opinion may be so great that the majority of raw grades might be considered anomalous by this analysis. In this situation it is not fair to assume that the remaining grades are truly representative of the pilot’s performance of the figure concerned.

5.4.1.4. To address this issue the uncertainty of each normalised figure  $U1_g(fp,j)$  for each pilot is accrued to provide a ‘Level of Anomaly’ (LOA) value as follows:

a) Where the  $U1_g(fp,j)$  is equal to or greater than 1.96 (95.0% confidence) and less than 2.43 (98.5%) the LOA is increased from its initial zero value by an amount given by:

$$[ABS(U1_g(fp,j) - 1.96) / (2.43 - 1.96)] \quad (16)$$

b) Where the  $U1_g(fp,j)$  is equal to or greater than 2.43 the LOA is increased by one (1.0).

If the final LOA for a figure exceeds 60% of the number of scoring judges then all judges’ grades for this figure/pilot are set to ‘Missing’ prior to the second normalisation of the group.

## 5.5. Second Normalisation of the Group

5.5.1.1. If anomalies have been removed from the raw grades, the data set will have more missing values. It would therefore be necessary to normalise the data group for a second time. Again, Perception Zero grades must be excluded from the Normalisation and these grades must remain 0.0. Using only the remaining non-zero grades new values must be determined for  $mNorm_g(*,j)$ ,  $mN_g(fp, *)$ ,  $mN_g(*, *)$  and thus  $FV_g(fp,j)$ .

a) Hence,

$$Norm2_g(fp,j) = mR2_g(*,*) + [R2_g(fp,j) - mR2_g(*,j)] * sdR2_g(*, ) / sdR2_g(*,j) \quad (17)$$

b) and,

$$FV2_g(fp,j) = mNorm2_g(*,j) + mNorm2_g(fp,*) - mNorm2_g(*, *) \quad (18)$$

5.5.1.2. These new Fitted Values will have been determined free from the influence of any anomalous grades and are thus robust and give the benefit of any doubt to the pilot in the case of minority Perception Zeros for an otherwise highly-graded figure.

## 5.6. Replacement of Missing and other anomalous Grades

5.6.1.1. These  $FV2_g(fp,j)$  values are then used to replace anomalous and other selected normalised grades carried forward from the preceding analysis in the following manner:

a) HZ, A and grades that were set ‘Missing’ in the first normalisation group are replaced by their Fitted Value  $FV2_g(fp,j)$ .



- b) Normalised figure data in the second group whose uncertainty  $U1g(fp,j)$  in the first group was equal to or greater than 1.96 (95.0%) are replaced by new data  $Nfig(p,j)$  proportionately between their second normalised value and their Fitted Value, given by:

$$\text{Norm}2_g(fp,j) + [(U2_g(fp,j) - 1.96) / (2.43 - 1.96)] \times (FV2_g(fp,j) - \text{Norm}2_g(fp,j)) \quad (19)$$

5.6.1.2. The judge concerned should have an increment made to his Low Score or High Score anomaly count in the Judging Performance Analysis, as appropriate, for each anomalous grade whose  $U1g(fp,j)$  was equal to or greater than 2.43, as well as to the HZ anomaly count for any HZ replaced.

5.6.1.3. After these replacements, the second normalised grades will be the final processed grades for each data group.

### **5.7. Assembly of Processed Grades by Pilot**

5.7.1.1. After processing in the separate data groups, the final processed grades must be combined into a single matrix and this table sorted by ascending value of the Pilot identification number and then the figure number. These grades are then multiplied by the respective K-factor for each figure and totalled to give:

- a)  $SR(p,f,j)$  an overall score for each pilot on each figure from each judge

5.7.1.2. These can then give

- a)  $SR(p,f,*)$  an overall score for each pilot for each figure over all judges, where:

$$SR(p,f,*) = \sum_j SR(p,f,j) \quad (20)$$

- b)  $SR(p,*,j)$  an overall score for each pilot for each judge over all figures, where:

$$SR(p,*,j) = \sum_f SR(p,f,j) \quad (21)$$

- c)  $SR(p,*,*)$  an overall score for each pilot, where:

$$SR(p,*,*) = \sum_{f,j} SR(p,f,j) \quad (22)$$

5.7.1.3. These data should be printed and passed to each pilot at the earliest possible stage, so that the changes made during the processing stage can be understood.

## **6. Sequence Processes**

### **6.1. Normalisation of Sequence Scores**

6.1.1.1. It is now necessary to repeat the normalisation process at the sequence stage, once again to ensure that the opinion of each judge is given the same importance.

6.1.1.2. The sequence score data,  $SR(p,j)$ , can be set out in a matrix form as shown here.

Pilot #	Judge 1	Judge 2	Judge 3	...	...	...	Judge j
Pilot 1	$SR(1,1)$	$SR(1,2)$	...	...	...	...	...
Pilot 2	$SR(2,1)$	...	...	...	...	...	...
Pilot 3	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
Pilot p	...	...	...	...	...	...	$SR(p,j)$



6.1.1.3. From this table:

- a)  $mSR(*,j)$  is the mean of all the scores given by Judge j.  
 $sdSR(*,j)$  is the standard deviation of all the scores given by Judge j.  
 $sdSR(*,*)$  is the average standard deviation of all the scores given to all the pilots by all the judges, and

$$\text{NormS}(p,j) = mSR(*,j) + [SR(p,j) - mSR(*,j)] * sdSR(*,*) / sdSR(*,j) \quad (23)$$

## 6.2. Derivation of Sequence Fitted Values

6.2.1.1. Next sequence Fitted Values are derived from the normalised scores to enable calculation of standardised residuals at the sequence level. In this derivation:

- a)  $mNormS(*,j)$  is the mean of all the normalised scores given by Judge j.  
 $mNormS(p,*)$  is the mean of all the normalised scores given to Pilot p.  
 $mNormS(*,*)$  is the mean of all normalised scores given by all judges to all pilots, and

$$FVS(p,j) = mNormS(*,j) + mNormS(p,*) - mNormS(*,*) \quad (24)$$

## 6.3. Assessment of Sequence Anomalies

6.3.1.1. Despite the replacement of anomalous figures at the earlier stage of the process, it might be possible for slight, consistent favouritism or subconscious bias to influence unduly a Judge's overall score for a pilot. Such a score might be high or low and should be replaced if its degree of uncertainty reaches approximately 90%.

6.3.1.2. Therefore the analysis must next derive the residuals for the sequence scores:

$$ResS(p,j) = NormS(p,j) - FVS(p,j), \text{ and} \quad (25)$$

$$RSS_s = \sum_{p,j} \{ResS(p,j)\}^2 \quad (26)$$

6.3.1.3. The number of degrees of freedom for the sequence data set is calculated where:

- a)  $D_s$  is the value of the Degrees of Freedom of the sequence data  
 $P_s$  is the number of Pilots in the sequence  
 $J_s$  is the number of judges in the sequence  
 $Nm_s$  is the number of missing values (confirmed HZ for all figures by a pilot), and

$$D_s = \{[P_s - 1] * [J_s - 1]\} - Nm_s \quad (27)$$

6.3.1.4. The Residual Standard Deviation for the sequence is given by:

$$RSD_s = \text{sqrt}\{RSS_s / D_s\} \quad (28)$$

6.3.1.5. The uncertainty of each sequence score is given by:

$$US_s(p,j) = ABS[ResS(p,j)] / RSD_s \quad (29)$$

6.3.1.6. If this uncertainty exceeds 1.24 (78.5%) it must be replaced by a new score  $Nscore(p,j)$  calculated in the following manner:

- a) Normalised score data whose uncertainty  $US_s(p,j)$  are equal to or greater than 1.65 (90.0%) are replaced by their Fitted Value  $FVS(p,j)$ .
- b) Normalised score data whose uncertainty  $US_s(p,j)$  are equal to or greater than 1.24 and less than 1.65 are replaced by new data  $Nscore(p,j)$  proportionately between their normalised value and their Fitted Value, given by:

$$\text{NormS}(p,j) + [(US_s(p,j) - 1.24) / (1.65 - 1.24)] * (FVS(p,j) - \text{NormS}(p,j)) \quad (30)$$



## 6.4. Final Sequence Scores

6.4.1.1. The processed sequence score for each pilot will be the sum of the normalised sequence scores over judges, after replacement of anomalous and other selected values of NormS(p,j):

$$PS(p) = \sum_j \{NormS(p,j) \text{ or } Nscore(p,j)\} / N_j \quad (31)$$

6.4.1.2. Penalties awarded for whatever reason are subtracted from this processed score to give each pilot's final overall score for the sequence.

$$FS(p) = PS(p) - Pen(p) \quad (32)$$

6.4.1.3. The final ranked order of all pilots for the sequence can now be published.

## 7. Process Summary

### 7.1. The process carries out the following analytical steps:

7.1.1.1. Sets confirmed Hard Zeros to HZ for all judges

7.1.1.2. Treats unconfirmed HZ and A grades as "Missing" at this stage.

7.1.1.3. Arranges figure grades into data groups for further analysis.

7.1.1.4. Within each data group:

- a) Normalises the grades to give equal importance to each judge.
- b) Derives Fitted Values for each judge for each figure.
- c) Determines if any normalised grades are more than 98.5% uncertain and disregards them by setting them to "Missing".
- d) Derives revised normalised grades and Fitted Values taking account of the new missing data.
- e) Replaces missing and selected normalised grades with revised data in proportion to the degree of anomaly detected.

7.1.1.5. At the sequence level:

- a) Normalises the scores to give equal importance to each judge.
- b) Derives Fitted Values for each judge for each pilot.
- c) Determines if any scores are more than 90% uncertain, and if so replaces them and selected normalised scores with new data in proportion to the degree of anomaly detected.

## 8. Judging Performance Analysis

The Fair Play System generates judging analysis data from the raw and FPS-processed scores. A Ranking Index is derived and judging errors totalled in a number of different categories. This data is gathered by the FPS and made available to the International Jury after each sequence and may, at the Jury's discretion, be passed on the Chief Delegates of each national team at the completion of each Programme, or at the end of the contest.

### 8.1. Ranking Index (RI)

8.1.1.1. The Ranking Index measures how closely an individual judge's pilot ranking for a programme conforms to the overall ranking based on all judges' assessments. It is derived by comparing the ranking each judge would have provided for each pilot if no other judges



were present (*Jrank*) with the overall ranking calculated from the combined judging panel after all anomalies have been resolved and prior to any penalties being included (*Prank*). In the event that *Jrank* and *Prank* differ, this difference is weighted by an amount proportional to the differences in the underlying sequence scores that lead to this difference. A ranking error is penalised more heavily when it involves a larger difference in scores than when only minor differences in scores are the cause.

8.1.1.2. To derive the “raw” Judge rankings (*Jrank*), all HZ and PZ grades are set to numerical zero (0.0) and any “A” grade is replaced by the mean of the scoring judges grades for the figure/pilot in question, calculated to the nearest half mark; if the ‘remainder’ is less than 0.25 the integer value is used; at or above 0.25 and less than 0.75 the integer value plus 0.5 is used; at or above 0.75 the next higher integer is used. This array of data is then normalised, and a pilot total score is calculated by multiplying each normalised figure mark by the relevant K Factor. If a PZ mark is determined to be an anomaly, it is set to "Missing" and replaced with a FV after resolution of any “A” marks but before final calculation of the Judge rankings.

8.1.1.3. After the full FPS process is complete, a full panel score (*Pscore*) and ranking (*Prank*) is determined for each pilot prior to the application of any penalties.

8.1.1.4. To derive the “Raw” Judge scores (*Jscore*), a final normalisation is conducted to rescale each judge’s all-pilot average “raw” scores and all-pilot average standard deviation of “raw” scores (as calculated in 8.1.1.2) to be the same as the all-pilot average and standard deviation of the full panel score (calculated in 8.1.1.3).

8.1.1.5. An index is now derived from these data for each judge. This takes account of differences in ranks and differences in total scores. Thus:

$$RI = \left\{ \sum_1^{Np} \frac{abs(Jrank - Prank) * abs(Jscore - Pscore)}{Pscore} \right\} * \frac{20}{(0.0057 * Np^2) + (0.1041 * Np)}$$

8.1.1.6. Typically, a judge’s Ranking Index will be in the range 1 to 50, lower numbers indicating that a judge’s individual rankings are closer to the overall panel rankings. When there has been a second FPS iteration in accordance with rule 6.5, the RI is calculated using only data from the second iteration, i.e. pilots excluded from the second iteration are also excluded from the RI calculation.

## **8.2. Low and High Scoring Anomalies**

8.2.1.1. A Low or High Scoring Anomaly is determined to have occurred each time a judge grades a figure significantly lower or higher than the consensus view of the judges.

8.2.1.2. For each figure, examine the normalised scores. If a judge’s score for the figure has been determined ‘Low’ or ‘High’ at the approved confidence level, then add one to that judge’s aggregate of errors under the appropriate heading.

## **8.3. Discrimination**

8.3.1.1. Judges show differing degrees of “Discrimination” in that they score over wider or narrower ranges of raw grades. The FPS will keep track of these raw grades for subsequent publication and analysis in the form of histograms and banded totals.



**8.4. Hard Zero Anomalies**

- 8.4.1.1. The occurrence of Hard Zeros is determined by majority voting or by video conference. The scoring system determines the validity of each HZ from the “CHZ” box on the score sheets.
- 8.4.1.2. In the event that an individual judge fails to identify a confirmed Hard Zero, then add one to that judge’s aggregate of errors under this heading. Similarly, if a judge gives a grade of HZ when no such error occurred, add one to the aggregate of errors under this heading.

**8.5. Sequence Score Anomalies**

- 8.5.1.1. A Sequence Anomaly occurs whenever a judge grades a whole sequence significantly higher or lower than the consensus view of the judges.
- 8.5.1.2. For each pilot, examine the normalised sequence scores. If a judge’s score for the sequence has been determined ‘High’ or ‘Low’ at the approved confidence level, then add one to that judge’s aggregate of errors under this heading.

