ICEAA 2016 Symposium

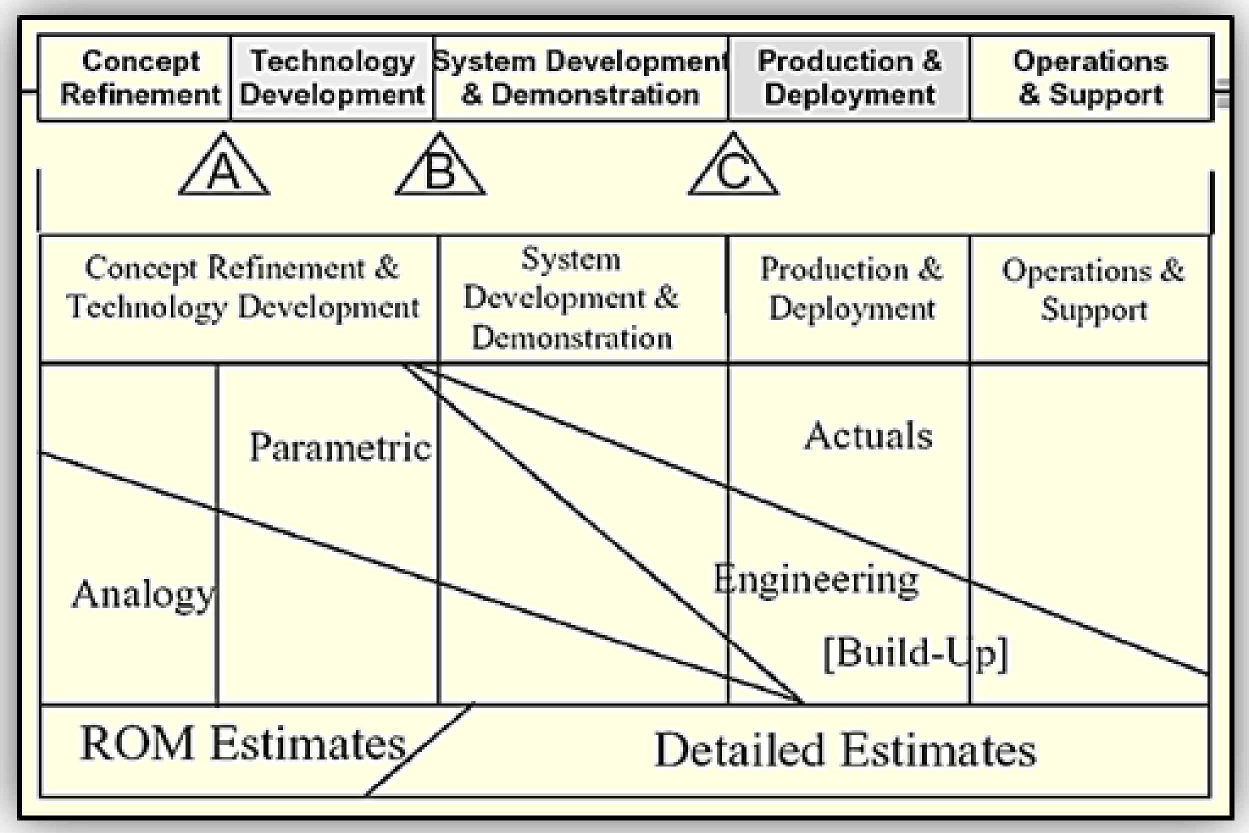
20 Oct 2016

Case Study
A Parametric Model for the Cost per Flight Hour (CPFH)

Maj. Michail Bozoudis
Senior Cost Engineer

General Directorate of
Defence Investments and Armaments,
Ministry of National Defence,
Hellenic Republic

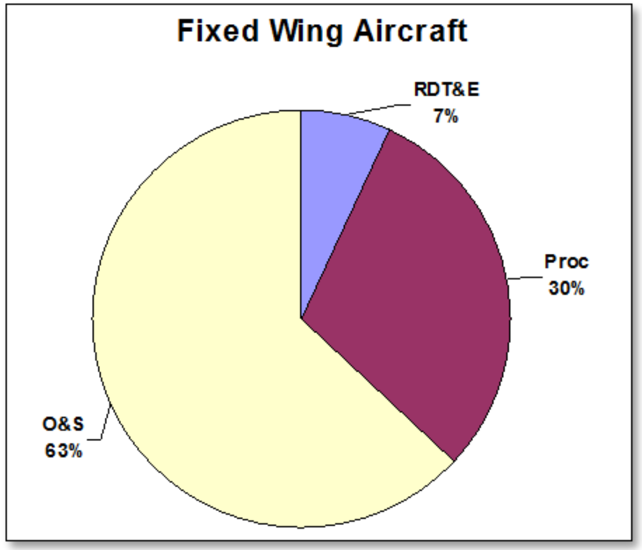
Applicability of Cost Estimating Techniques

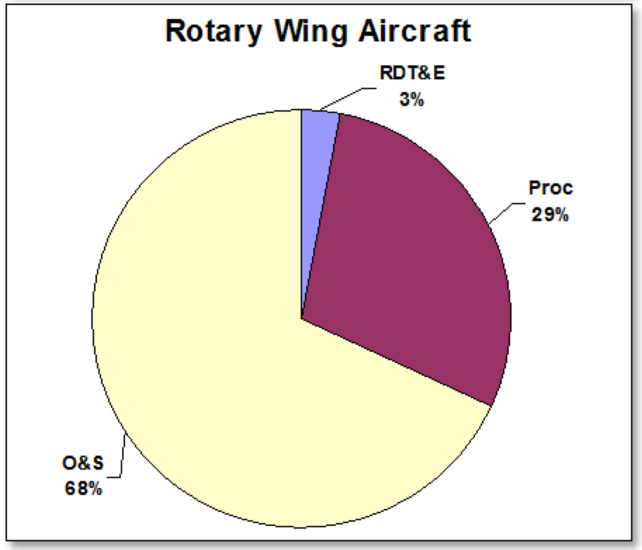


Source: DAU Integrated Defense Acquisition, Technology, and Logistics LCM Framework chart, v5.2 (2008).

Analogy

$$Cost = f(Cost')$$





Source: OSD/CAPE Operating and Support Cost-Estimating Guide (2014), Chapter 2, fig. 2-2

$$O\&S\ Cost \approx 2.10 \cdot Procurement$$

$$O\&S\ Cost \approx 2.34 \cdot Procurement$$

$$CPFH = \frac{O\&S\ Cost}{FLHRs}$$

Parametric

$$Cost = f(p_1, p_2, p_3, \ldots, p_n)$$

"Parametric estimating is a technique that develops cost estimates based upon the examination and validation of the relationships which exist between a project's technical, programmatic, and cost characteristics as well as the resources consumed during its development, manufacture, maintenance, and/or modification. Parametric models can be classified as simple or complex. Simple models are cost estimating relationships (CERs) consisting of one cost driver. Complex models, on the other hand, are models consisting of multiple CERs, or algorithms, to derive cost estimates."

Source: ISPA/SCEA Parametric Handbook, 4th Edition (2008)

Engineering (Build-Up)

$$Cost = \sum_{i=1}^{n} Cost_i$$

O&S Cost breakdown

1.0 Unit-level manpower

2.0 Unit operations

3.0 Maintenance

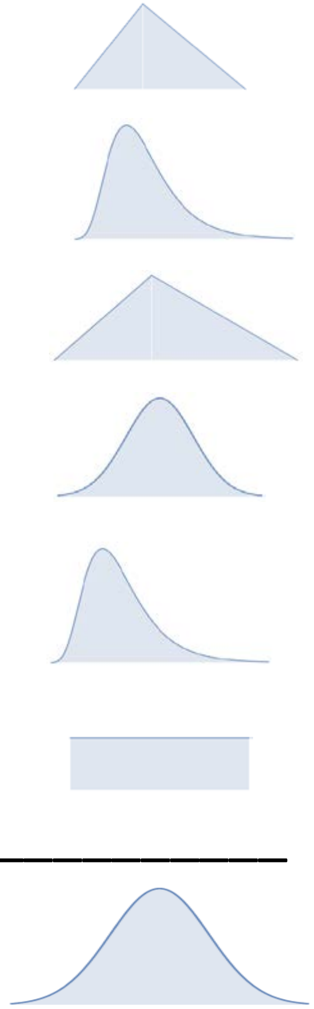
4.0 Sustaining support

5.0 Continuing system improvements

6.0 Indirect cost

____

TOTAL



Source: GAO-09-3SP Cost Estimating and Assessment Guide (2009).

Construction of the parametric model

Objective:
Comparison of alternatives

Pre-Analysis considerations: Constraints & Requirements

- Use the available (small) sample of 22 systems that HAF operates
- Exclude indirect cost
- Search for cost drivers that are easily accessible and quantifiable
- The selected model must:
  - not include more than two cost drivers
  - be significant at the 5% level
  - capture at least 75% of the CPFH variance
  - have valid prediction intervals
  - make sense

Hellenic Air Force fleet

Trainers

T-41D


T-6A II


T-2E


Fire fighters

CL-215


CL-415


PZL


Fighters

F-16C/D


F/RF-4E


M2000-5


A-7H


AEW&C

EMB-145H


VIP

EMB-135


G-V


Transporters

C-130H


C-27J


Helicopters

AB-205

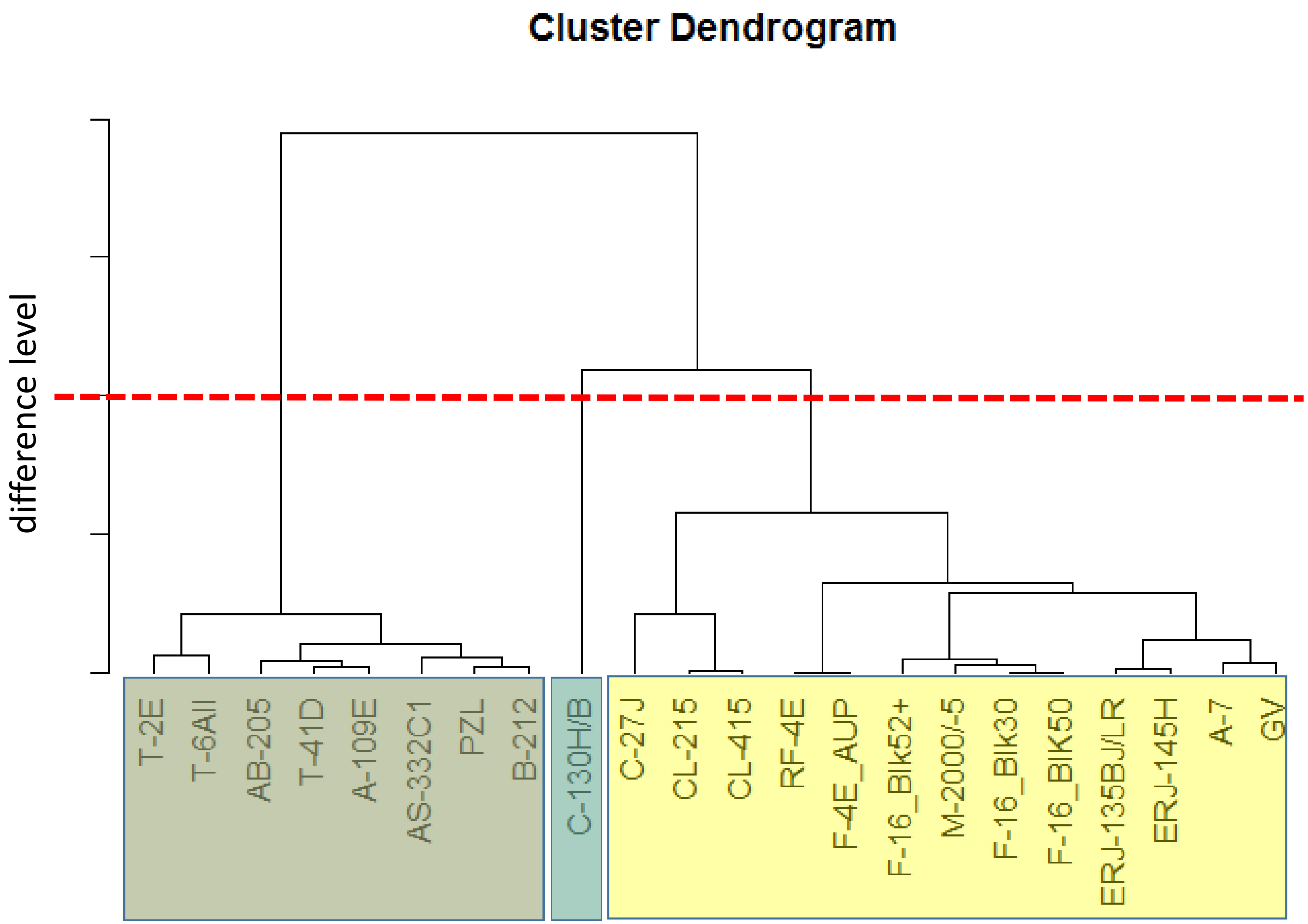

B-212


AS-332C1

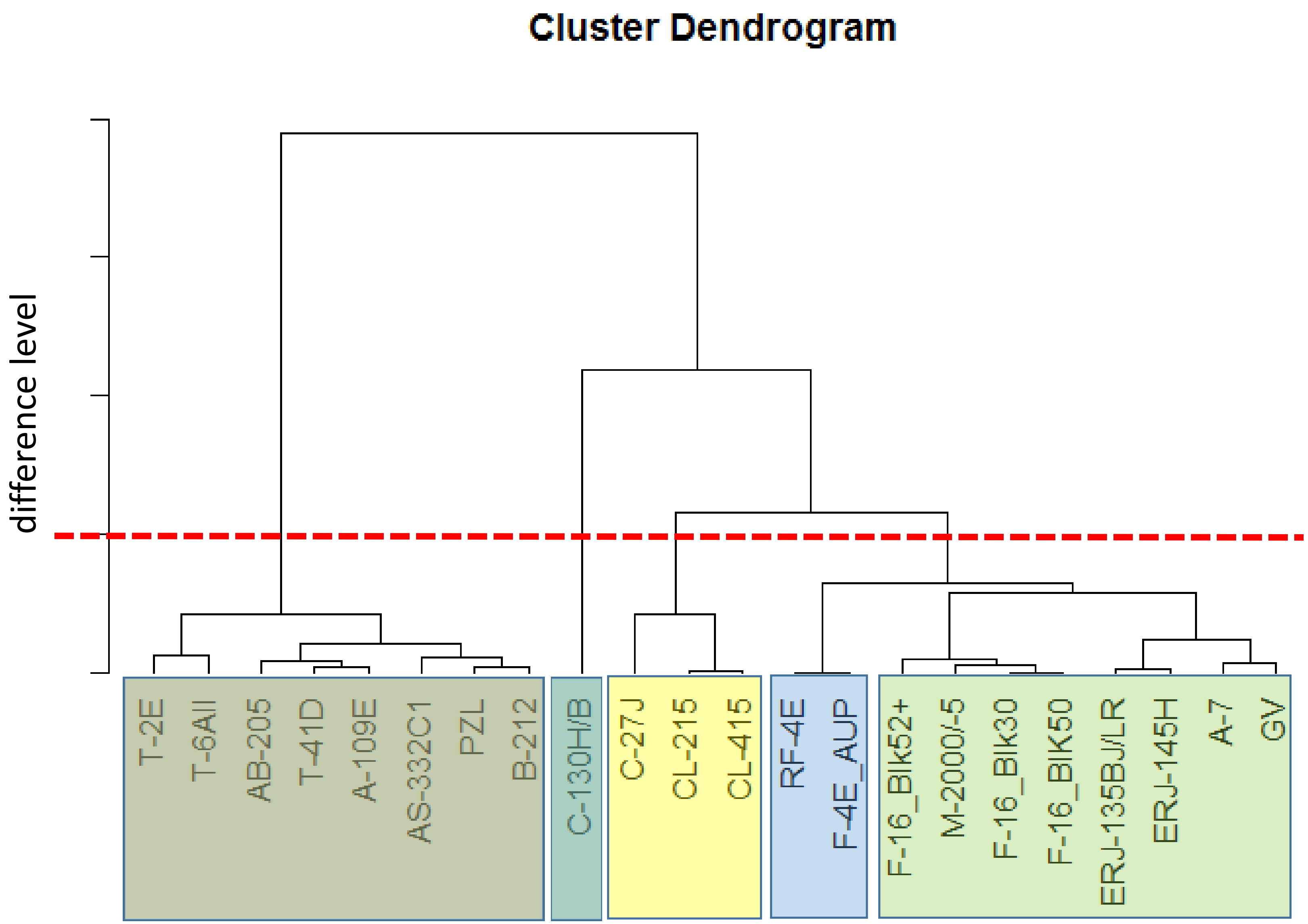

A-109E


Independent variables

- Length
- Empty weight
- MTOW
- SFC (max)
- Speed (max)
- Ceiling

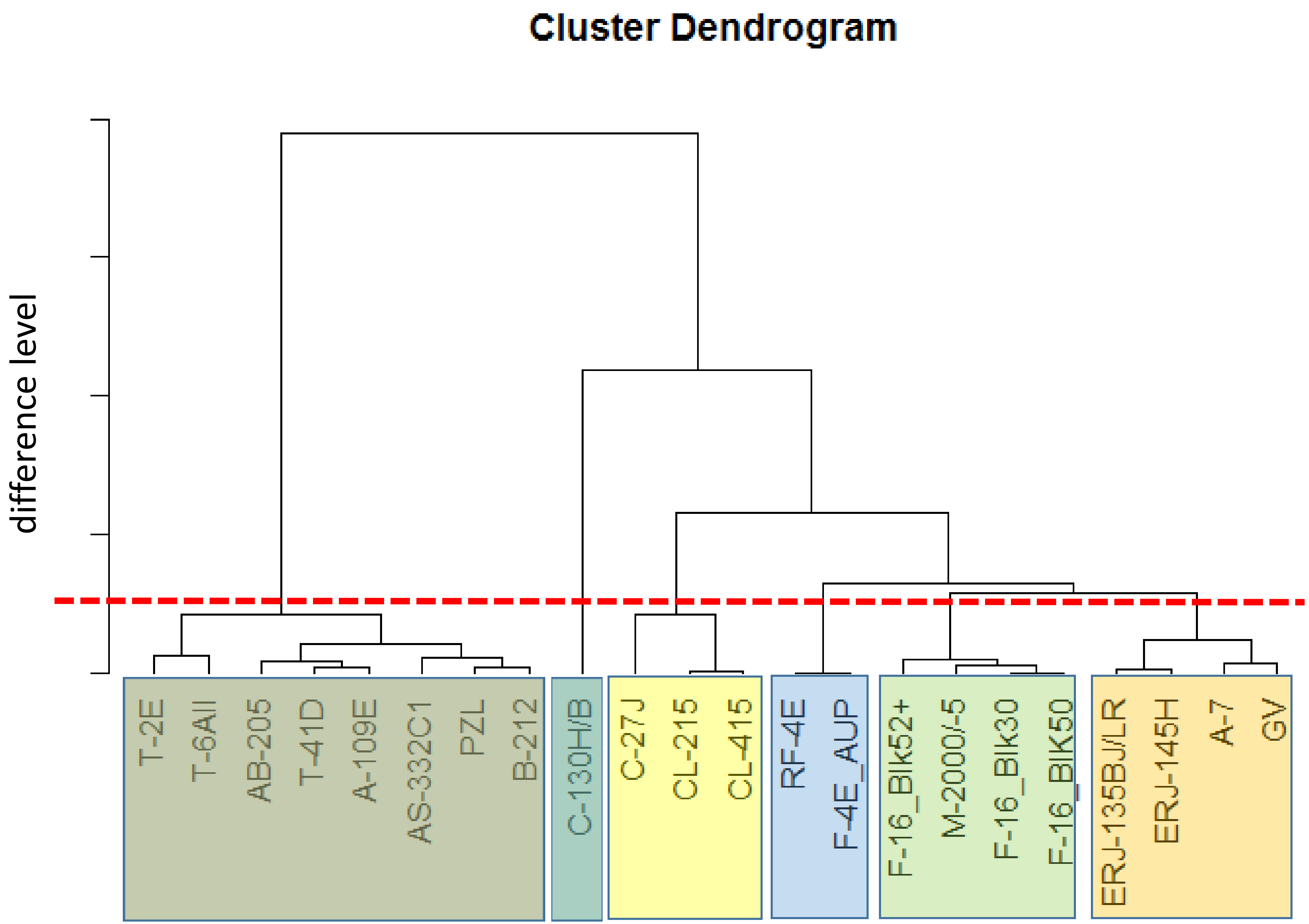


Independent variables
- Length
- Empty weight
- MTOW
- SFC (max)
- Speed (max)
- Ceiling

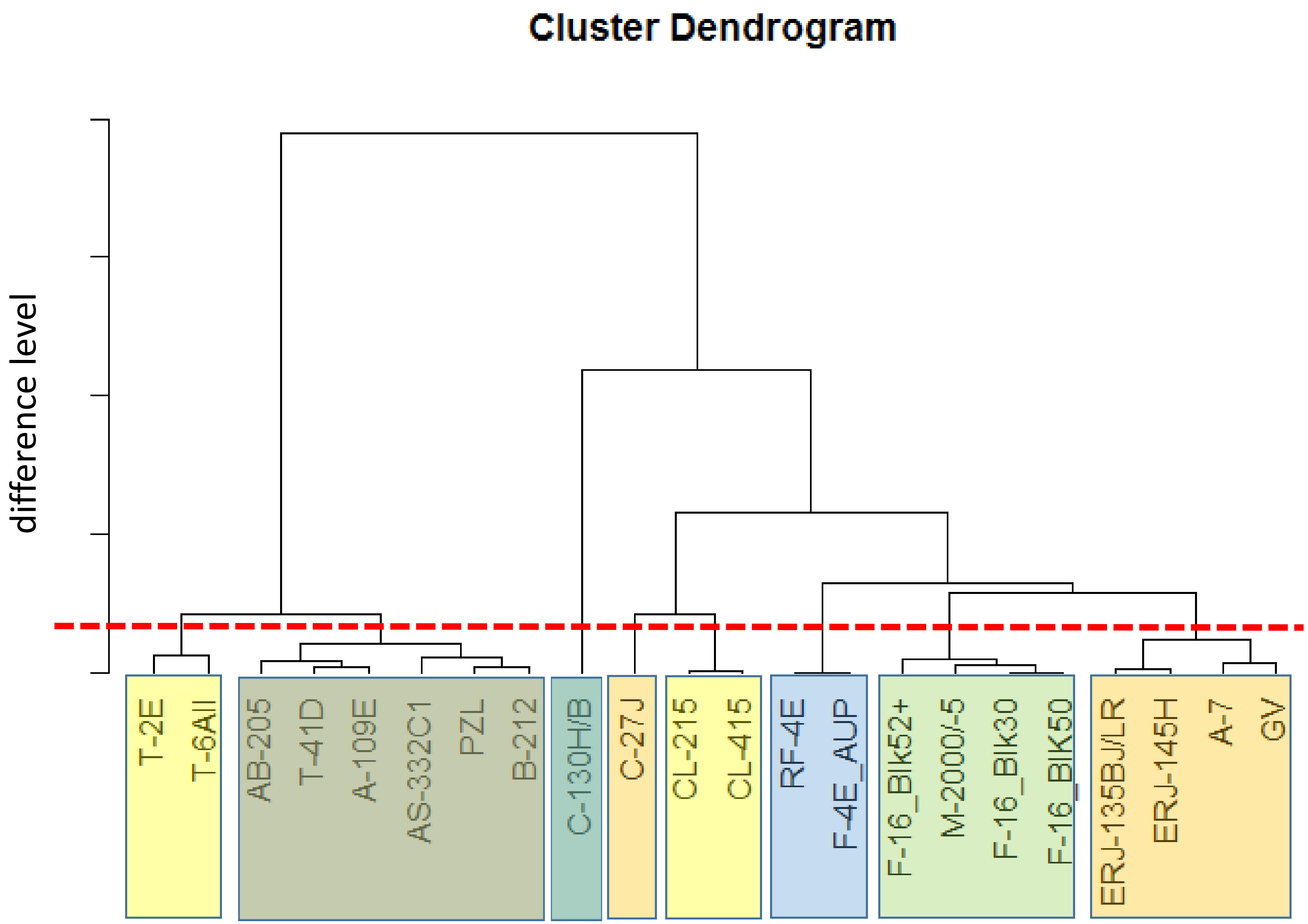


Independent variables

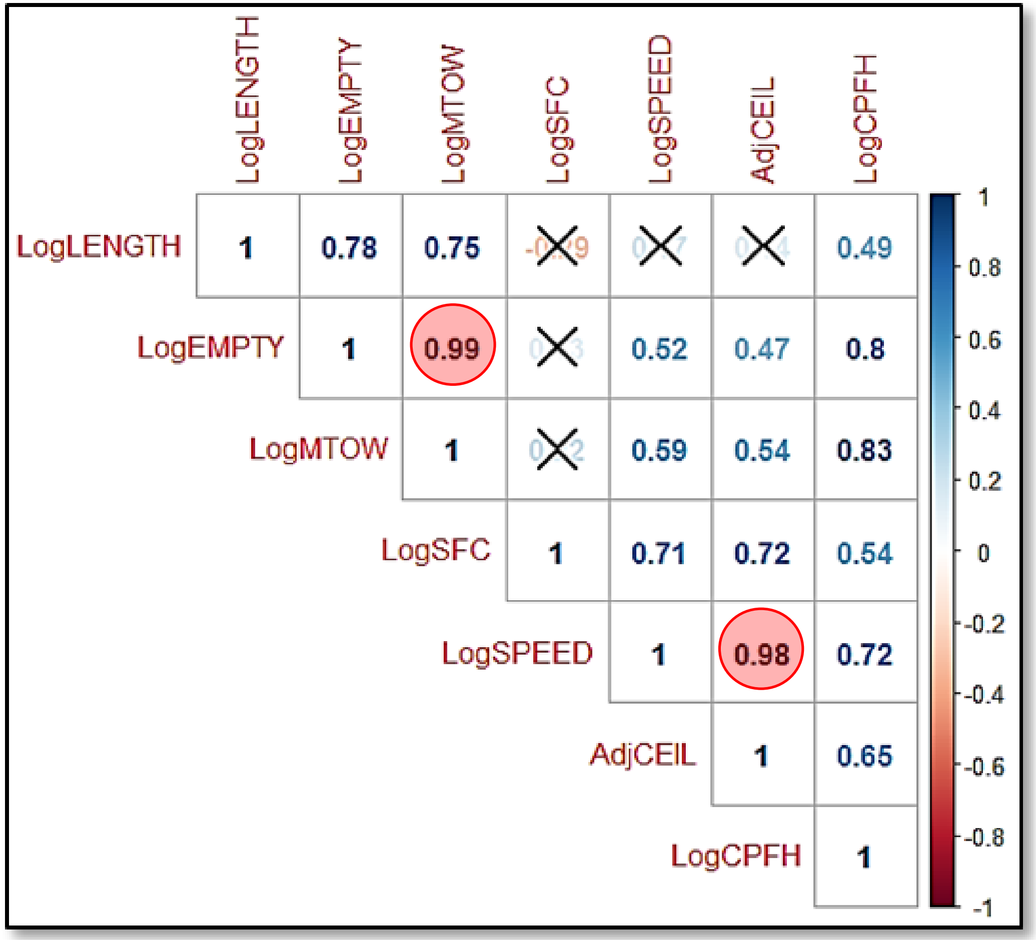
- Length
- Empty weight
- MTOW
- SFC (max)
- Speed (max)
- Ceiling

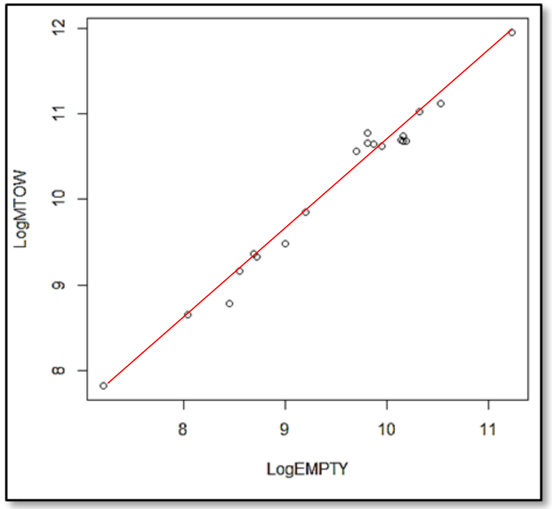


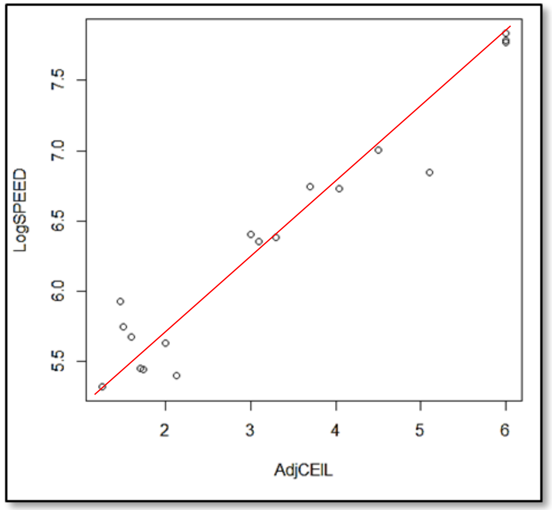
Independent variables
- Length
- Empty weight
- MTOW
- SFC (max)
- Speed (max)
- Ceiling



Multicollinearity issues







Different variables contain the same information!!!
(They are highly correlated and one can be linearly predicted from the other(s))

Model selection

Simple CER

	LogLENGTH	LogEMPTY	LogMTOW	LogSFC	LogSPEED	AdjCEIL	LogCPFH
LogLENGTH	1	0.78	0.75	✕	✕	✕	0.49
LogEMPTY		1	0.99	✕	0.52	0.47	0.8
LogMTOW			1	✕	0.59	0.54	0.83
LogSFC				1	0.71	0.72	0.54
LogSPEED					1	0.98	0.72
AdjCEIL						1	0.65
LogCPFH							1

$$\text{Log}(CPFH) = a_0 + a_1 \cdot \text{Log}(MTOW)$$

$$R^2 = 0.69$$

Complex CER

	LogLENGTH	LogEMPTY	LogMTOW	LogSFC	LogSPEED	AdjCEIL	LogCPFH
LogLENGTH	1	0.78	0.75	✕	✕	✕	0.49
LogEMPTY		1	0.99	✕	0.52	0.47	0.8
LogMTOW			1	✕	0.59	0.54	0.83
LogSFC				1	0.71	0.72	0.54
LogSPEED					1	0.98	0.72
AdjCEIL						1	0.65
LogCPFH							1

Akaike Criterion

$$\text{Log}(CPFH) = \beta_0 + \beta_1 \cdot \text{Log}(Empty\ weight) + \beta_2 \cdot \text{Log}(SFC)$$

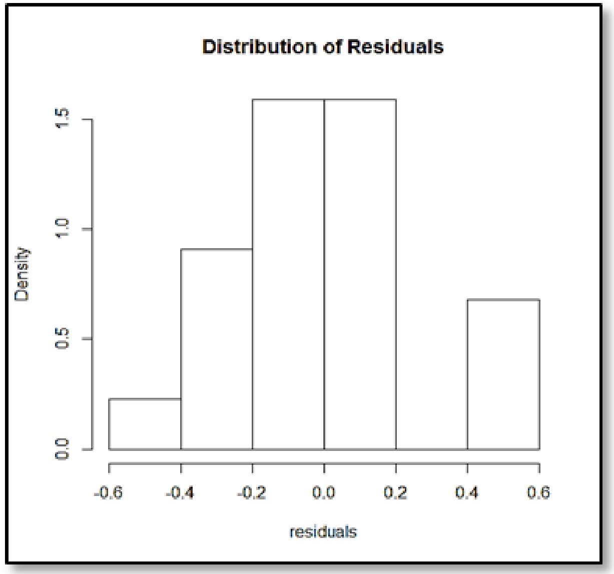
$$R^2_{adj} = 0.82$$

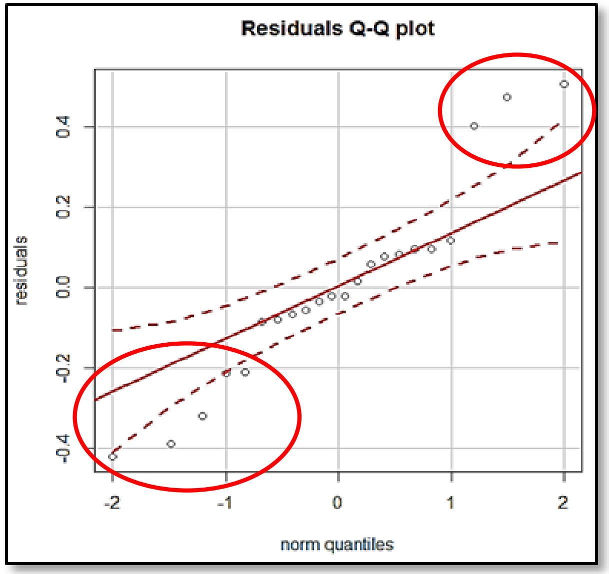
ANOVA table

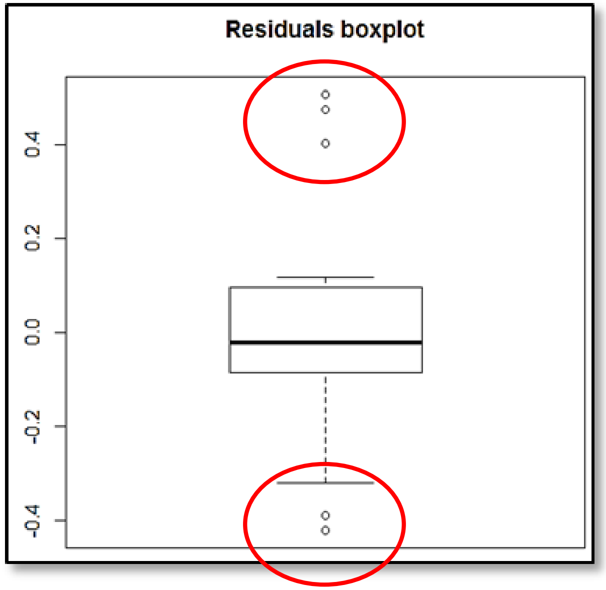
```
Call:
lm(formula = LogCPFH ~ LogEMPTY + LogSFC)

Residuals:
     Min       1Q   Median       3Q      Max
-0.42125 -0.08515 -0.02154  0.09199  0.50650

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  ........  .....     6.570 2.74e-06 ***
LogEMPTY     ........  .....     7.984 1.73e-07 ***
LogSFC       ........  .....     4.827 0.000117 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2553 on 19 degrees of freedom
Multiple R-squared:  0.8385,            Adjusted R-squared:  0.8215
F-statistic: 49.31 on 2 and 19 DF,        p-value: 3.009e-08

Correlation of Coefficients:
         (Intercept) LogEMPTY
LogEMPTY -0.99
LogSFC    0.17       -0.13
```

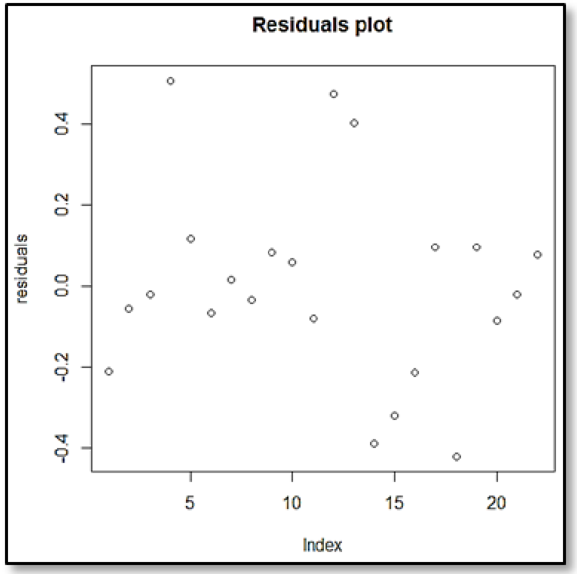
Residuals diagnostics

Test	Null hypothesis	*p*-value	Reject the null hypothesis at the 5% sig. level?
Shapiro-Wilk normality test	normality	0.161	NO
Breusch-Pagan test for heteroscedasticity	constant variance	0.332	NO
Durbin-Watson test for autocorrelation	no autocorrelations	0.342	NO









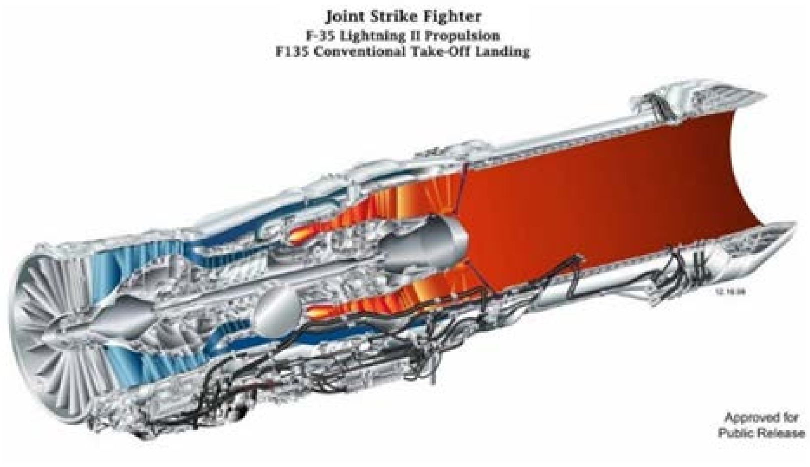
Review of the selected model

Constraints & requirements	Results
Use the sample of 22 aircraft operated by the Hellenic Air Force.	**OK**.
Use the appropriate cost information.	**OK**. Current CPFH data used, excluding the *indirect support* cost category.
Use cost drivers (independent variables) that are easily accessible and quantifiable.	**OK**. The cost drivers are physical and performance characteristics.
The model must be as less complex as possible and include no more than two cost drivers.	**OK**. The selected model includes two independent variables.
The model should be statistically significant at the 5% level.	**OK**. p-value = $3 \cdot 10^{-8}$
The model should capture at least 75% of the CPFH variance.	**OK**. $R^2_{adj} = 0.82$
The model's prediction intervals must be valid.	**OK**. The residuals pass all tests. There are many outliers though.
The model's mathematical expression should make sense.	**OK**. The model suggests that the aircraft weight and the engine specific fuel consumption correlate positively with the CPFH.

Post-Analysis considerations

- Small sample → high uncertainty
- Diverse systems → poor precision, robust CERs
- Many outliers → unreliable prediction intervals
- Outliers → why are they far away from the "mainstream"?
- Tailored model → no generalizations
  - ❖ Why was the model built?
  - ❖ Which question does the model actually answer?
  - ❖ How does the model perform on the training sample?
  - ❖ How can the model be useful?

CPFH point estimate for an “unknown” system



F-35A empty weight = 29,098 lb

F135-PW-100 specific fuel consumption ≈ 1.95 lb/lbf·h

$$\text{Log}(CPFH) = \beta_0 + \beta_1 \cdot \text{Log}(29{,}098) + \beta_2 \cdot \text{Log}(1.95)$$

CPFH prediction interval

Theoretical approach

$$\text{Prob}\left[\hat{Y}_0 - s(\hat{Y}_0) \cdot t_{n-p,\frac{a}{2}} \leq Y \leq \hat{Y}_0 + s(\hat{Y}_0) \cdot t_{n-p,\frac{a}{2}}\right] = 1 - a$$

where: $s^2(\hat{Y}_0) = STE^2 \cdot [\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]$, and $Y = \text{Log}(CPFH)$

CPFH prediction interval

Simulation approach

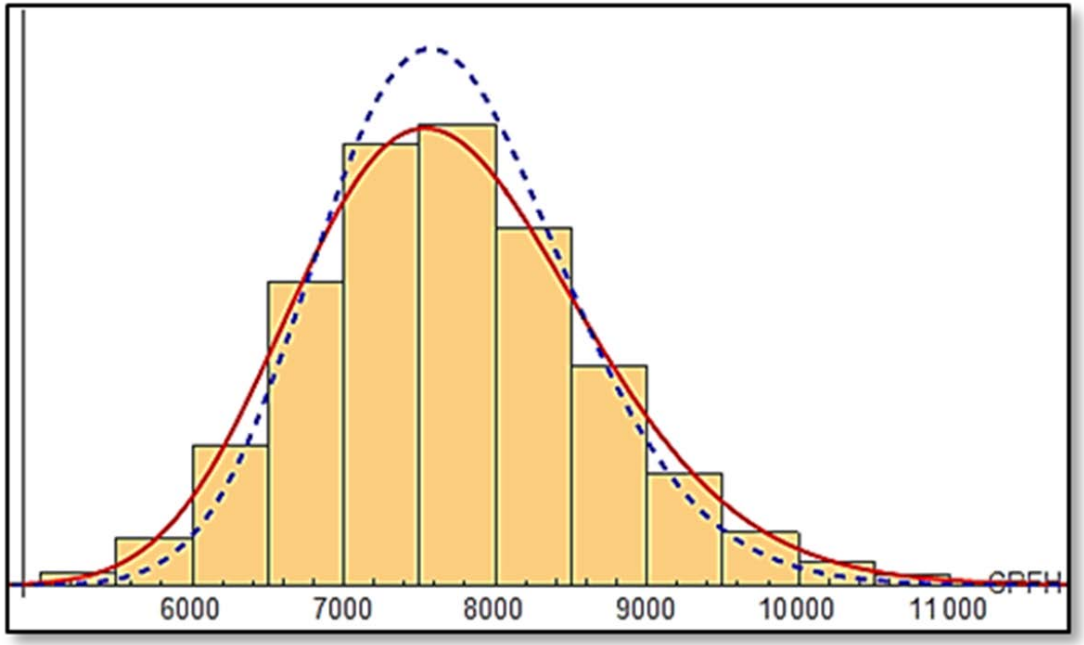
STEP 1: Generate one set of pseudorandom values for the coefficients β_0, β_1, β_2, using a multivariate Student-t distribution, according to the coefficients correlation matrix and their standard errors.

STEP 2: Use the model's equation to get a value for Y (denoted y_1).

STEP 3: Repeat the steps above to get a set of y_i , $i = 1,2, \ldots , n$.

STEP 4: $\mathrm{E}[Y] = \frac{1}{n}\sum_{i=1}^{n} y_i$ $\quad$ $s^2[Y] = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - E[Y])^2$

CPFH estimate for the F-35A



Theoretical output (dashed line)

Simulation output (solid line)

Property	Theoretical output	Simulation output
Log(CPFH) mean	8.9434	8.9435
Log(CPFH) standard deviation	0.1066	0.1254
CPFH mean	7,701 €	7,719 €
CPFH median	7,658 €	7,658 €
CPFH mode	7,571 €	7,539 €
CPFH standard deviation	823 €	973 €
CPFH 80th percentile	8,376 €	8,481 €
CPFH 95% prediction interval	6,214 to 9,436 €	5,975 to 9,822 €
Prob(CPFH > 10,000 €)	0.61%	1.83%
Cost risk (80th percentile - mode)	805 €	942 €

REFERENCES

[1] AAP-20 NATO Program Management Framework (2015)

[2] AAP-48 NATO System Life Cycle Stages and Processes (2013)

[3] AFMC Air Force Analyst's Handbook, by C. Feuchter (2000)

[4] ALCCP-1 NATO Guidance on Life Cycle Costs (2008)

[5] DoD 5000.4-M Cost Analysis Guidance and Procedures (1992)

[6] FAA Guide to Contacting Business Case Cost Evaluations (2015) Accessible at http://www.ipa.faa.gov/Tasks.cfm?PageName=Parametric%20Cost%20Estimating

[7] GAO-09-3SP Cost Estimating and Assessment Guide (2009) Accessible at http://www.gao.gov/new.items/d093sp.pdf

[8] Hellenic Air Force official site https://www.haf.gr

[9] ISPA/SCEA Parametric Handbook, 4th Edition (2008) Accessible at http://www.galorath.com/images/uploads/ISPA_PEH_4th_ed_Final.pdf

[10] NATO Continuous Acquisition and Lifecycle Support (CALS) Handbook, v.2 (2000)

[11] NASA Cost Estimating Handbook v.4 (2015) Accessible at http://www.nasa.gov/pdf/263676main_2008-NASA-Cost-Handbook-FINAL_v6.pdf

[12] OSD/CAPE Operating and Support Cost-Estimating Guide (2014) Accessible at http://www.cape.osd.mil/files/OS_Guide_v9_March_2014.pdf

[13] TO 00-20-2 Maintenance Data Documentation, Change 2 (2007) Accessible at http://everyspec.com/USAF/USAF-Tech-Manuals/download.php?spec=TO_00-20-2.004007.pdf

[14] USAF Cost Risk and Uncertainty Analysis (CRUA) Handbook (2007) Accessible at http://www.amsaa.army.mil/Documents/Air%20Force%20Analyst's%20Handbook.pdf