Extracting ICAO Carbon Calculator Data

2022-07-17

# 1 Overview

Environmental Sustainability has become a priority topic over the past months. Within PBWG and the bi-lateral work, discussions revolve around establishing an initial “look-up” table for converting operational inefficiencies to associated fuel burn or CO2 estimates (or other emissions).

# 2 Understanding Reading pdf files with R

reading in the pdf file

* pdf\_text() generates a list of all pages of the pdf file

icc\_all <- pdftools::pdf\_text(pdf\_fn)

convert table page to a machine-readable table

page\_17 <- icc\_all[17] # extract page from list  
  
page\_17 <- page\_17 %>%   
 read\_lines() %>% # use line breaks  
 as\_tibble() # convert to tibble  
  
page\_17

## # A tibble: 54 x 1  
## value   
## <chr>   
## 1 " Appendix C: ICAO Fuel Consumption Table"   
## 2 "Equivalent"   
## 3 " ~  
## 4 " Aircraft"   
## 5 " Code 125 250 500 750 1000 1500 2000 2500 ~  
## 6 " 100 1296 2703 3788 5129 6427 8937 11373 13757 ~  
## 7 " 141 1289 2754 3874 5258 6600 9199"   
## 8 " 142 1289 2754 3874 5258 6600 9199 11725"   
## 9 " 143 1324 2874 4105 5621 7100 9986"   
## 10 " 146 1289 2754 3874 5258 6600 9199 11725"   
## # ... with 44 more rows

we can skip the first rows

skip\_rows <- 4  
page\_17 <- page\_17 %>%   
 filter(row\_number() > skip\_rows)  
  
page\_17

## # A tibble: 50 x 1  
## value   
## <chr>   
## 1 " Code 125 250 500 750 1000 1500 2000 2500 ~  
## 2 " 100 1296 2703 3788 5129 6427 8937 11373 13757 ~  
## 3 " 141 1289 2754 3874 5258 6600 9199"   
## 4 " 142 1289 2754 3874 5258 6600 9199 11725"   
## 5 " 143 1324 2874 4105 5621 7100 9986"   
## 6 " 146 1289 2754 3874 5258 6600 9199 11725"   
## 7 " 310 2628 5537 7790 10759 13658 19323 24876 30356 ~  
## 8 " 313 2628 5537 7790 10759 13658 19323 24876 30356 ~  
## 9 " 318 1488 3016 3925 5234 6482 8931 11335 13729 ~  
## 10 " 319 1596 3259 4323 5830 7271 10026 12668 15233 ~  
## # ... with 40 more rows

clean string and split columns

# table has IATA Code and then ranges up-to 8500NM  
col\_names <- c("CODE","125","250","500","750","1000","1500","2000","2500", "3000", "3500","4000","4500","5000","5500","6000","6500","7000","7500","8000","8500")  
  
page\_17 %>%   
 mutate(  
 value = trimws(value)  
 ,value = gsub(pattern = "\\s+", replacement = "\*", x = value)  
 ) %>%   
 separate(  
 col = value  
 , into = col\_names  
 , fill = "right"  
 )

## # A tibble: 50 x 21  
## CODE `125` `250` `500` `750` `1000` `1500` `2000` `2500` `3000` `3500`  
## <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>   
## 1 Code 125 250 500 750 1000 1500 2000 2500 3000 3500   
## 2 100 1296 2703 3788 5129 6427 8937 11373 13757 16104 <NA>   
## 3 141 1289 2754 3874 5258 6600 9199 <NA> <NA> <NA> <NA>   
## 4 142 1289 2754 3874 5258 6600 9199 11725 <NA> <NA> <NA>   
## 5 143 1324 2874 4105 5621 7100 9986 <NA> <NA> <NA> <NA>   
## 6 146 1289 2754 3874 5258 6600 9199 11725 <NA> <NA> <NA>   
## 7 310 2628 5537 7790 10759 13658 19323 24876 30356 35784 41172   
## 8 313 2628 5537 7790 10759 13658 19323 24876 30356 35784 41172   
## 9 318 1488 3016 3925 5234 6482 8931 11335 13729 16130 <NA>   
## 10 319 1596 3259 4323 5830 7271 10026 12668 15233 17741 20203   
## # ... with 40 more rows, and 10 more variables: `4000` <chr>, `4500` <chr>,  
## # `5000` <chr>, `5500` <chr>, `6000` <chr>, `6500` <chr>, `7000` <chr>,  
## # `7500` <chr>, `8000` <chr>, `8500` <chr>

# 3 Extracting ICAO Carbon Calculator Fuel Burn Estimates

## 3.1 Apendix C

wrap this into a function

col\_names <- c("CODE","125","250","500","750","1000","1500","2000","2500", "3000", "3500","4000","4500","5000","5500","6000","6500","7000","7500","8000","8500")  
  
extract\_icc\_range\_table <- function(.pdf\_page, .skip\_rows = 0, .col\_names = col\_names){  
 range\_table <- .pdf\_page %>%  
 read\_lines() %>% # use line breaks  
 as\_tibble() %>%   
 filter(row\_number() > skip\_rows) %>%   
 mutate(  
 value = trimws(value)  
 ,value = gsub(pattern = "\\s+", replacement = "\*", x = value)  
 ) %>%   
 separate(  
 col = value  
 , into = col\_names  
 , fill = "right"  
 ) %>%   
 # filter all rows where each range column has no value  
 filter(!if\_all(col\_names[2]:col\_names[length(col\_names)], is.na))  
 return(range\_table)  
}

ICC document lists fuel consumption tables on page 17 through 23

page\_nbrs <- 17:23  
fuel\_stages <- page\_nbrs %>%   
 purrr::map\_dfr(.f = ~ icc\_all[.x] %>% extract\_icc\_range\_table) %>%   
 filter(CODE != "Code")

write out fuel\_stages table for future use

write\_csv(fuel\_stages, “./data-analytic/ICC-FuelConsumption-v11-2018.csv”)

## 3.2 Apendix B - Aircraft Mapping

Follow same logic to read in mapping tables. Note that there are multiple Aircraft-Equivalent mappings per page.

extract\_aircraft\_mappings <- function(.pdf\_page){  
 mapping <- .pdf\_page %>%   
 read\_lines() %>%   
 as\_tibble() %>%   
 filter(row\_number() > row\_number()[grepl("^Aircraft", value)]) %>%   
 mutate(  
 value = trimws(value)  
 ,value = gsub(pattern = "\\s+", replacement = "\*", x = value)  
 ) %>% separate(  
 col = value  
 , into = c("Aircraft","Equivalent","Aircraft\_2","Equivalent\_2","Aircraft\_3","Equivalent\_3","Aircraft\_4","Equivalent\_4")  
 , fill = "right"  
 )  
 tmp1 <- mapping[,1:2]  
 tmp2 <- mapping[,3:4]; names(tmp2) = c("Aircraft","Equivalent")  
 tmp3 <- mapping[,5:6]; names(tmp3) = c("Aircraft","Equivalent")  
 tmp4 <- mapping[,7:8]; names(tmp4) = c("Aircraft","Equivalent")  
 mapping <- bind\_rows(tmp1, tmp2, tmp3, tmp4) %>%   
 # remove any incomplete combination / mapping (e.g. NAs, spurious page number)  
 filter(!if\_any(1:2, is.na))   
 return(mapping)  
}

page\_nbrs <- 14:16  
ac\_equi\_map <- page\_nbrs %>%   
 purrr::map\_dfr(.f = ~ icc\_all[.x] %>% extract\_aircraft\_mappings )

write\_csv(fuel\_stages, “./data-analytic/ICC-AircraftMapping-v11-2018.csv”)

## 3.3 ICAO Aircraft Type Designator to IATA Type Code Mapping

The ICC documentation uses - unfortunately - IATA Type Codes. Let’s hack together a mapping table on top of our handwork for the ACERT look-up.

# get something useful from wikipedia  
library(rvest)

## Warning: package 'rvest' was built under R version 4.0.5

##   
## Attaching package: 'rvest'

## The following object is masked from 'package:readr':  
##   
## guess\_encoding

url <- "https://en.wikipedia.org/wiki/List\_of\_aircraft\_type\_designators"  
  
wiki\_html <- read\_html(url)  
wiki\_nodes <- wiki\_html %>%   
 html\_nodes(css = "table")  
  
# only one table in nodes, but to be sure we fetch #1  
wiki\_tbl <- wiki\_nodes %>%   
 nth(1) %>%   
 html\_table()  
  
names(wiki\_tbl) <- c("ICAO", "IATA", "MODEL")

# another source  
url2 <- "https://www.avcodes.co.uk/acrtypes.asp"  
  
avcodes <- read\_html(url2)  
avc\_nodes <- avcodes %>%   
 html\_nodes("table")  
avc\_tbl <- avc\_nodes %>% first() %>% html\_table()  
  
names(avc\_tbl) <- c("IATA", "ICAO\_AVCodes", "Manufacturer-Type-Model\_AVCodes", "WTC")

# another source  
url3 <- "https://stringfixer.com/nl/List\_of\_aircraft\_type\_designators"  
  
stringfix <- read\_html(url3)  
sfix\_nodes <- stringfix %>% html\_nodes("table")  
sfix\_tbl <- sfix\_nodes %>% first() %>% html\_table()  
  
names(sfix\_tbl) <- c("ICAO\_sfix", "IATA\_sfix", "MODEL\_sfix")

combine all - for the time being

rq <- wiki\_tbl %>% left\_join(avc\_tbl %>% rename(ICAO = ICAO\_AVCodes)) %>% left\_join(sfix\_tbl %>% rename(ICAO = ICAO\_sfix))

## Joining, by = c("ICAO", "IATA")  
## Joining, by = "ICAO"

write\_csv(rq, “./data-analytic/Aircraft-ICAO-IATA-Types.csv”)

## 3.4 Augment ICC Fuel Consumption

Appendix B recognises “equivalent” emitting aircraft types. Thus, we have to augment the look up to account for these additional aircraft.

fuel\_stages2 <- fuel\_stages %>%   
 mutate(VERSION = "ICC-v11-2018 (direct)")  
  
which\_equiv <- ac\_equi\_map %>%   
 mutate(CHECK = Aircraft != Equivalent) %>%   
 filter(CHECK == TRUE)  
  
which\_equiv <- which\_equiv %>% select(-CHECK) %>%   
 inner\_join(fuel\_stages %>% rename(Equivalent = CODE)) %>%   
 mutate(VERSION = paste0("ICC-v11-2018 (equiv: ", Equivalent, ")")) %>%   
 rename(CODE = Aircraft)

## Joining, by = "Equivalent"

which\_equiv <- which\_equiv %>%   
 filter(CODE != "787") # double accounting of 787  
  
fuel\_stages2 <- bind\_rows(fuel\_stages2, which\_equiv) %>%   
 rename(IATA = CODE, EQUIVALENT = Equivalent)

fuel\_stages2 <- fuel\_stages2 %>%   
 left\_join(avc\_tbl %>% rename(ICAO = ICAO\_AVCodes))

## Joining, by = "IATA"

fuel\_stages2 <- fuel\_stages2 %>% select(ICAO, IATA, everything())

missin\_icao <- fuel\_stages2 %>%   
 filter(ICAO == "" | is.na(ICAO))

fix by hand - check that we only have unique ICAO !

missin\_icao %>% left\_join(sfix\_tbl, by = c(“IATA”=“IATA\_sfix”)) %>% select(ICAO\_sfix, ICAO, IATA, everything()) %>% filter(!is.na(ICAO\_sfix)) # A tibble: 5 x 28 ICAO\_sfix ICAO IATA 125 250 500 750 1000 1500 2000 2500 3000 3500 4000 4500 5000 5500 1 A388 NA 388 5851 12016 17623 24940 32211 46695 61160 75638 90143 104681 119255 133865 148512 163196 2 A321 NA 32B 1909 3925 5270 7157 8970 12456 15818 19094 22308 NA NA NA NA NA  
3 A158 NA A58 1543 3087 4064 5306 6478 NA NA NA NA NA NA NA NA NA  
4 E75L NA E7W 1113 2240 2989 3953 4890 6725 NA NA NA NA NA NA NA NA  
5 E50P NA EP1 290 580 764 997 1218 NA NA NA NA NA NA NA NA NA

write\_csv(fuel\_stages2, “./data-analytic/ICC-FuelConsumption-v11-2018-augmented.csv”)

There exists already a A321, E75L. Fixed by editing the csv (outside R).

# 4 Conclusion

This document summarises the data preparatory steps for generating a look-up table for PBWG and the bi-regional comparison work. It is based on the extraction of fuel burn estimates from the ICAO Carbon Calculator Methodology document, v11 2018.

The data has been extracted from the respective Appendix C. With Appendix B, the lookup has been expanded for a set of aircraft types that show - in accordance with the methodology document - similar fuel burn characteristics.

The comparison work is based on ICAO Aircraft Type Designators (*Aircraft Type Designators, Doc 8643\_2022* 2022).

== initial lookup table!

*Aircraft Type Designators, Doc 8643\_2022*. 2022. *Aircraft Type Designators, Doc 8643*. 50th ed. ICAO.