

**METODE *RANDOM SURVIVAL FOREST* DENGAN  
ATURAN PEMISAHAN *LOG-RANK* DAN *LOG-RANK SCORE*  
UNTUK PREDIKSI *DELAY* PENERBANGAN AKIBAT CUACA**

**SKRIPSI**

Diajukan untuk memenuhi sebagian dari syarat untuk memperoleh  
gelar Sarjana Matematika



Oleh:

Jihan Kautsarina

1801457

**PROGRAM STUDI MATEMATIKA  
DEPARTEMEN PENDIDIKAN MATEMATIKA  
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS PENDIDIKAN INDONESIA  
2022**

**METODE *RANDOM SURVIVAL FOREST* DENGAN  
ATURAN PEMISAHAN *LOG-RANK* DAN *LOG-RANK SCORE*  
UNTUK PREDIKSI *DELAY* PENERBANGAN AKIBAT CUACA**

Oleh  
Jihan Kautsarina  
1801457

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana Matematika pada  
Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

© Jihan Kautsarina 2022  
Universitas Pendidikan Indonesia  
Desember 2022

Hak Cipta dilindungi Undang-Undang  
Skripsi ini tidak boleh diperbanyak seluruhnya maupun sebagian dengan cara  
dicetak ulang, fotokopi, atau lainnya tanpa izin dari penulis

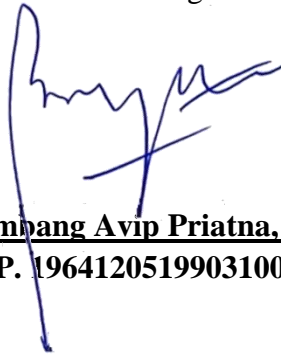
**LEMBAR PENGESAHAN**

JIHAN KAUTSARINA

METODE *RANDOM SURVIVAL FOREST* DENGAN  
ATURAN PEMISAHAN *LOG-RANK* DAN *LOG-RANK SCORE*  
UNTUK PREDIKSI *DELAY* PENERBANGAN AKIBAT CUACA

Disetujui dan disahkan oleh pembimbing:

Pembimbing I



**Dr. Bambang Avip Priatna, M.Si.**

**NIP. 19641205199031001**

Pembimbing II



**Fitriani Agustina, S.Si., M.Si.**

**NIP. 198108142005012001**

Mengetahui,

Ketua Departemen Pendidikan Matematika



**Dr. H. Dadang Juandi, M.Si.**

**NIP. 196401171992021001**

# **METODE *RANDOM SURVIVAL FOREST* DENGAN ATURAN PEMISAHAN *LOG-RANK* DAN *LOG-RANK SCORE* UNTUK PREDIKSI *DELAY* PENERBANGAN AKIBAT CUACA**

## **ABSTRAK**

Metode *Random Survival Forest* (RSF) merupakan metode penggabungan analisis *survival* dengan klasifikasi. Metode ini menjadi alternatif dari model *Cox Proportional Hazard* ketika asumsi *proportional hazard* tidak terpenuhi dalam menganalisis data tersensor. Metode RSF sangat fleksibel untuk mengatasi data berdimensi tinggi dengan banyak kovariat. Dengan pendekatan seperti metode klasifikasi *Random Forest*, metode ini diawali dengan membuat sampel *bootstrap*, kemudian dilakukan *random selection* pada variabel untuk menumbuhkan pohon demi pohon hingga simpul terakhir. Performansi pembuatan RSF sangat bergantung pada aturan pemisahan. Aturan ini akan memecah *node* induk menjadi beberapa *node* anak hingga menghasilkan variabel penting (VIMP) yang paling berkontribusi. Aturan pemisahan yang digunakan dalam penelitian ini adalah *Log-Rank* dengan *Log-Rank Score*, kemudian membandingkan hasil keduanya. Indikator yang digunakan untuk membandingkan performansi kedua aturan pemisahan ini adalah dengan *error rate* atau nilai kesalahan prediksi. Studi kasus dilakukan pada data *delay* penerbangan akibat cuaca di rute bandara Soekarno Hatta menuju bandara I Gusti Ngurah Rai. Metode ini diterapkan untuk menentukan variabel yang berkontribusi pada lamanya *delay* penerbangan apakah lebih dari 30 menit atau kurang dari 30 menit. Hasil penelitian ini menunjukkan bahwa metode RSF dengan aturan pemisahan *Log-Rank* adalah pilihan terbaik untuk memprediksi *delay* penerbangan akibat cuaca dengan nilai *error rate* yang lebih kecil yaitu sebesar 0,3590 atau 35,90%. Ukuran hutan yang dibangun adalah terdiri dari 500 pohon. Urutan variabel penting yang berkontribusi di antaranya *airlines*, kecepatan angin, arah angin, suhu, curah hujan, dan kelembaban. Setelah performa RSF dengan *Log-Rank* dikatakan baik, dibuat *web app* sederhana untuk melakukan pengotomatisan prediksi *delay* penerbangan akibat cuaca.

**Kata Kunci:** *Delay* Penerbangan, Analisis *Survival*, Klasifikasi, *Random Survival Forest*, Aturan Pemisahan.

**RANDOM SURVIVAL FOREST METHOD WITH  
LOG-RANK AND LOG-RANK SCORE SPLITTING RULE  
FOR PREDICTING THE FLIGHT DELAY DUE TO WEATHER**

**ABSTRACT**

*The Random Survival Forest (RSF) method is a method that combining survival analysis with classification. This method is an alternative to the Cox Proportional Hazard model when proportional hazard assumption is not fullfilled in analyzing the censored data. The RSF method is very flexible for handling high-dimensional data with many covariates. Using the approach of classification method such as the Random Forest, this method begins with drawing a bootstrap sample, then selecting variables randomly to grow tree by tree until the terminal node. Generated RSF performance is highly depends on splitting rules. This rule will split the parent node into several daughter nodes to show the most contributing variable importance (VIMP). The splitting rule that used for this research is Log-Rank and Log-Rank Score. We compare these two splitting rules. The indicator that used to compare the performance of the two splitting rules is the error rate or prediction error value. A case study was applied on flight delay data due to weather on the route of Soekarno Hatta airport to I Gusti Ngurah Rai airport. This method is applied to determine the variables that contributing in flight delays time, whether more than 30 minutes or less than 30 minutes. The results of this study indicate that the RSF method with Log-Rank splitting rules is the best choice for predicting flight delays due to weather with a smaller error rate value of 0.3590 or 35.90%. The size of the forest consists 500 trees. The sequence of important contributing variables are airlines, wind speed, wind direction, temperature, precipitation and humidity. After RSF with Log-Rank's performance was qualified as good, a simple web app is created to automate flight delay predictions due to weather.*

**Key Word:** *Flight Delay, Survival Analysis, Classification, Random Survival Forest, Splitting Rule.*

## DAFTAR ISI

LEMBAR PENGESAHAN .....	i
KATA PENGANTAR .....	ii
UCAPAN TERIMA KASIH.....	iii
ABSTRAK .....	iv
ABSTRACT .....	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR .....	ix
DAFTAR LAMPIRAN.....	x
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian .....	5
1.5 Manfaat Penelitian .....	6
BAB II KAJIAN TEORI.....	7
2.1 <i>Delay</i> Penerbangan.....	7
2.2 Faktor Cuaca sebagai Penyebab <i>Delay</i> Penerbangan.....	8
2.3 Analisis <i>Survival</i> .....	10
2.3.1 Waktu <i>Survival</i> .....	11
2.3.2 Penyensoran Data .....	12
2.3.3 Fungsi <i>Survival</i> .....	13
2.3.4 Fungsi <i>Hazard</i> .....	14
2.3.5 Fungsi <i>Hazard</i> Kumulatif.....	15
2.4 Klasifikasi .....	16
2.5 <i>Random Forest</i> (RF) .....	17
2.6 <i>Random Survival Forest</i> (RSF).....	19
2.7 <i>Bootstrap</i> .....	20
2.8 Aturan Pemisahan ( <i>Splitting Rule</i> ).....	21

2.8.1	Aturan Pemisahan <i>Log-Rank</i> .....	22
2.8.2	Aturan Pemisahan <i>Log-Rank Score</i> .....	23
2.9	Pemilihan Variabel Penting (VIMP).....	24
2.10	Tingkat Kesalahan Prediksi .....	24
2.10.1	Cara Menentukan <i>C-Index</i> .....	25
2.10.2	Prediksi Kesalahan OOB .....	26
BAB III METODOLOGI PENELITIAN.....		28
3.1	Sumber Data Penelitian.....	28
3.2	Variabel Penelitian .....	29
3.3	Algoritma <i>Random Survival Forest</i> (RSF) .....	31
3.4	Langkah dan Diagram Alur Penelitian.....	32
BAB IV HASIL DAN PEMBAHASAN .....		35
4.1	Analisis Data <i>Delay</i> Penerbangan.....	35
4.2	Pembagian Data <i>In-Bag</i> dan Data <i>Out-of-Bag</i> .....	37
4.3	Analisis <i>Random Survival Forest</i> (RSF).....	38
4.4	Prediksi Data menggunakan RSF <i>Log-Rank</i> .....	39
4.4.1	Nilai Kesalahan Prediksi RSF <i>Log-Rank</i> .....	40
4.4.2	Variabel Penting RSF <i>Log-Rank</i> .....	41
4.5	Prediksi Data Menggunakan RSF <i>Log-Rank Score</i> .....	43
4.5.1	Nilai Kesalahan Prediksi RSF <i>Log-Rank Score</i> .....	44
4.5.2	Variabel Penting RSF <i>Log-Rank Score</i> .....	45
4.6	Perbandingan RSF <i>Log-Rank</i> dan RSF <i>Log-Rank Score</i> .....	46
4.7	Prediksi <i>Delay</i> Penerbangan Akibat Cuaca.....	47
4.8	<i>Web App</i> Sederhana.....	48
BAB V PENUTUP.....		49
5.1	Kesimpulan .....	49
5.2	Saran.....	50
DAFTAR PUSTAKA .....		51
DAFTAR LAMPIRAN.....		57

## DAFTAR TABEL

Tabel 2. 1 Total <i>Delay</i> Penerbangan Januari - Juni 2022 .....	10
Tabel 2. 2 Contoh Penggunaan Nilai <i>C-Index</i> .....	25
Tabel 3. 1 Daftar Variabel Terikat .....	30
Tabel 3. 2 Daftar Variabel Bebas .....	30
Tabel 4. 1 Analisis Deskriptif Variabel Terikat dan Bebas .....	35
Tabel 4. 2 Proporsi Data <i>In-Bag</i> dan <i>Out-of-Bag</i> .....	37
Tabel 4. 3 Hasil Metode RSF <i>Log-Rank</i> .....	39
Tabel 4. 4 Nilai VIMP RSF <i>Log-Rank</i> .....	42
Tabel 4. 5 Hasil Metode RSF <i>Log-Rank Score</i> .....	43
Tabel 4. 6 Nilai VIMP RSF <i>Log-Rank Score</i> .....	46
Tabel 4. 7 <i>Error Rate</i> RSF <i>Log-Rank</i> dan <i>Log-Rank Score</i> .....	46



## DAFTAR GAMBAR

Gambar 2. 1 Skema Algoritma <i>Random Forest</i> .....	18
Gambar 3. 1 Tampilan Laman <i>FlightRadar24</i> .....	28
Gambar 3. 2 Tampilan Laman NASA POWER.....	29
Gambar 3. 3 <i>Flowchart</i> Algoritma RSF.....	32
Gambar 3. 4 Diagram Alur Penelitian.....	34
Gambar 4. 1 <i>Delay</i> Penerbangan berdasarkan Status Tersensor.....	35
Gambar 4. 2 Kurva <i>Survival</i> berdasarkan <i>Airlines</i> .....	36
Gambar 4. 3 Kurva <i>Survival</i> berdasarkan Arah Angin .....	37
Gambar 4. 4 Kurva <i>Error Rate</i> pada RSF <i>Log-Rank</i> .....	40
Gambar 4. 5 Grafik VIMP pada RSF <i>Log-Rank</i> .....	41
Gambar 4. 6 Kurva <i>Error Rate</i> pada RSF <i>Log-Rank Score</i> .....	44
Gambar 4. 7 Grafik VIMP pada RSF <i>Log-Rank Score</i> .....	45
Gambar 4. 8 Tampilan <i>Web App</i> Sederhana .....	48

## DAFTAR LAMPIRAN

Lampiran 1. Data <i>Delay</i> Penerbangan Akibat Cuaca pada Rute Bandara CGK-DPS ....	57
Lampiran 2. <i>Syntax R</i> Metode RSF <i>Log-Rank</i> dan <i>Log-Rank Score</i> .....	58
Lampiran 3. <i>Syntax R</i> Pembuatan <i>Web App</i> Sederhana .....	63
Lampiran 4. Pohon Klasifikasi RSF <i>Log-Rank</i> dengan <i>ntree=2</i> .....	67
Lampiran 5. Pohon Klasifikasi RSF <i>Log-Rank Score</i> dengan <i>ntree=2</i> .....	68
Lampiran 6. Daftar Istilah .....	69

## DAFTAR PUSTAKA

- Badan Kebijakan Transportasi. (2021). *Meteorologi dan Jalur Penerbangan*. [Online]. Diakses dari [baketrans.dephub.go.id](http://baketrans.dephub.go.id).
- Belle, Vanya Van., dkk. (2011). *Support vector methods for survival analysis: a comparison between ranking and regression approaches*. *Artif Intell Med*. doi : [doi.org/10.1016/j.artmed.2011.06.006](https://doi.org/10.1016/j.artmed.2011.06.006)
- Biau, Gerard. (2012). Analysis of a Random Forests Model. *Journal of Machine Learning Research*, 13. France.
- Breiman, Leo. (2001). *Random Forests*. Netherlands : University of California, Statistics Departure. doi : [doi.org/10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
- Breiman, L., & Cutler, A. (2003). *Manual--Setting Up, Using, and Understanding Random Forest V4.0*. [Online]. Accessed from [stat.berkeley.edu/~breiman/Using\\_random\\_forests\\_v4.0.pdf](http://stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf)
- Collett, David. (2003). *Modelling Survival Data in Medical Research Second Edition*. London : Chapman & Hall/CRC.
- Cox, David R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 2. jstor : [jstor.org/stable/2985181](http://jstor.org/stable/2985181)
- Deo, S. V., Deo, V., & Sundaram, V. (2021). Survival analysis-part 2: Cox proportional hazards model. *Journal of Thoracic and Cardiovascular Surgery*. doi : [doi.org/10.1007/s12055-020-01108-7](https://doi.org/10.1007/s12055-020-01108-7).
- Dermadi, Yudi. dkk. (2019). Weather Analysis of Flight Delays at Husein Sastranegara Airport. *Information Technology Engineering Journals Vol. 4 No. 2 Page 89-98*. doi : [doi.org/10.24235/itej.v4i2.31](https://doi.org/10.24235/itej.v4i2.31)
- DiBonaventuro, Frank. (2004). *How does humidity effect the way that an airplane flies?*. *Physics & Astronomy*. [Online]. Accessed from [hcmut.edu.vn](http://hcmut.edu.vn).

- Efron, B., & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. New York : Chapman and Hall.  
doi : doi.org/10.1007/978-1-4899-4541-9
- Ervianto, W.I. (2004). *Teori Aplikasi Manajemen Proyek Konstruksi*. Yogyakarta : Andi.
- Fadholi, Akhmad. (2013). Studi Pengaruh Suhu dan Tekanan Udara terhadap Operasi Penerbangan di Bandara H.A.S Hananjoeddin Buluh Tumbang Belitung Periode 1980-2010. *Jurnal Penelitian Fisika dan Aplikasinya*. Pangkalpinang : Stasiun Meteorologi Deputi Amir.
- Fu, Wei., & Simonoff, Jeffrey S. (2017). *Survival trees for left-truncated and right censored data, with application to time-varying covariate data*. New York : IOMS Department.  
doi : doi.org/10.1093/biostatistics/kxw047
- Galiano, Rodriguez., dkk. (2015). *Machine Learning Predictive Models for Mineral Prospectivity: An Evaluation of NN, Random Forest, Regression Trees, SVM*. Ore Geology.  
doi : doi.org/10.1016/j.oregeorev.2015.01.001
- Gayatri, Dewi. (2005). Mengenal Analisis Ketahanan (Survival Analysis). *Jurnal Keperawatan Indonesia Vol. 9, No. 1*.
- Ghozali, Imam. (2011). *Aplikasi Multivariate dengan Program IBM SPSS 19 Edisi 5*. Semarang : Badan Penerbit Diponegoro.
- Gupta, S. K., Rajeev, B., & Kaur, D. (2015). *Review of Decision Tree Data mining Algorithms: ID3 and C4.5. Proceeding of International Conference on Information Technology and Computer Science*.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Data Mining Concepts And Techniques (Third)*. Morgan Kaufmann Publishers.
- Hermawan, Toto. Nurrohmah, Dwi., & Jannah, I F. (2019). *Estimasi Fungsi Survival dan Fungsi Hazard Kumulatif pada Data Survival Penderita Multiple Myeloma serta Faktor-Faktor yang Mempengaruhi Waktu Survivalnya*. *Jurnal Intersections Vol. 4 No. 2*. Yogyakarta : UCY.
- Hermawati, Fajar Astuti. (2013). *Data Mining*. Yogyakarta : Penerbit Andi.

- Hothorn, Torsten., & Lausen, Berthold. (2003). On the exact distribution of maximally selected rank statistics. Germany : ELSEVIER.  
doi : doi.org/10.1016/S0167-9473(02)00225-6
- ICAO. (2019). *The World of Air Transport in 2019 - Air Transport Statistics*. United Nations. [Online]. Accessed from icao.int.
- Ishwaran, Hemant. (2007). Variable importance in binary regression trees and forests. *Electronic Journal of Statistics*. Cleveland.  
doi : doi.org/10.1214/07-EJS039
- Ishwaran, H., Kogalur, U.B., Blackstone, E .H., & Lauer, M.S. (2008). Random Survival Forests. *The Annals of Applied Statistics*.  
doi : doi.org/10.1214/08-AOAS169
- Ishwaran, H., & Lu, Min. (2019). *Random Survival Forests*. Miami : University of Miami Miller School.  
doi : doi.org/10.1002/9781118445112.stat08188
- Kaho, Norman P.L.B Riwu. (2014). *Panduan Interpretasi dan Respon Informasi Iklim dan Cuaca untuk Petani dan Nelayan*. Kupang : Pikul dan ICCTF.
- Kementrian Perhubungan RI. (2021). *Laporan Tahunan Direktorat Jenderal Perhubungan Udara Tahun 2021*. Jakarta.
- Kim, Myeonghan., & Bae, Jiheon. (2020). Modeling the flight departure delay using survival analysis in South Korea. *Journal of Air Transport Management*.  
doi : doi.org/10.1016/j.jairtraman.2020.101996
- Kleinbaum, D. G., & Klein, M. (2005). *Survival Analysis: A Self-Learning Text Second Edition*. USA : Springer Science+Business Media, Inc. USA.  
doi : doi.org/10.1007/0-387-29150-4
- Kogalur, Udaya B., & Ishwaran, Hemant. (2007). Random Survival Forest for R. *R News Vol. 7/2*. ISSN : 1609-3631.
- Latifah, Esti. (2018). *Random forest untuk mengatasi masalah overfitting pada klasifikasi*. (Skripsi). Universitas Indonesia.
- Lawless, J. F. (2003). *Statistical Model and Methods for Lifetime Data*. New Jersey : John Wiley and Sons.

- Lee, Elisa T., & Wang, John Wenyu. (2003). *Statistical Methods for Survival Data Analysis Third Edition*. New Jersey : John Wiley and Sons.
- Menteri Perhubungan Republik Indonesia. (2015). *Peraturan Menteri Perhubungan RI Nomor PM 89 Tahun 2015 Tentang Penanganan Keterlambatan Penerbangan (Delay Management) pada Badan Usaha Angkutan Udara Niaga Berjadwal di Indonesia*. Jakarta.
- Neiburger, Morris., dkk. (1982). *Understanding Our Atmospheric Environment*. New York : W.H Freeman and Company.
- Nasejje, J. B., Mwambi, H., Dheda, K., & Lesosky, M. (2017). A comparison of the conditional inference survival forest model to random survival forests based on simulation study as well as on two applications with time-to-event data. *BMC Medical Research Methodology*.  
doi : doi.org/10.1186/s12874-017-0383-8
- OAG (*Official Airline Guide*). (2020). *On-Time Performance*. [Online] Accessed from oag.com.
- Octapiani, S. (2021). *Random Survival Forest versus Cox Proportional Hazard dalam Kasus Menyeimbangkan Kembali pada Sistem Berbagi Sepeda*. (Skripsi). UIN Sunan Gunung Djati Bandung.
- Purwakinanti, Rengganis., dkk. (2014). Aplikasi Metode Momen Probabilitas Terboboti untuk Estimasi Parameter Distribusi Pareto Terampat pada Data Curah Hujan. *Jurnal Gaussian Vol. 3 No. 4*.  
ejournal-s1.undip.ac.id/index.php/gaussian.
- Pakan, Welly. (2012). On Time Performance Levels of Sultan Hasanudin Airport in 2011. *Jurnal Penelitian Perhubungan Udara Vol. 38 No. 2*. Jakarta : Warta Ardhia.
- Prasetyo, Pipi Dwi S. (2019). *Penentuan Delay Penerbangan Akibat Cuaca Menggunakan Decision Tree*. (Skripsi). Kediri : Universitas Nusantara PGRI Kediri.
- Rizki, Yuke Sri., & Samsudin, Rosidin. (2014). *On Time Performance Airline Scheduled at The Airport I Gusti Ngurah Rai-Bali*. Jakarta : Badan Litbang Perhubungan.

- Saadati, M., & Bagheri, A. (2019). Comparison of Survival Forests in Analyzing First Birth Interval. Iran : *Jorjani Biomedicine Journal*.  
doi : doi.org/10.29252/jorjanibiomedj.7.3.11
- Sumathi, K., Kannan, S., & Nagarajan, K. (2016). Data Mining: Analysis of Student Database Using Classification Techniques. *International Journal of Computer Applications*.  
doi : doi.org/10.5120/ijca2016909703
- Sungkono, Joko. (2015). Bootstrap Resampling Observasi pada Estimasi Parameter Regresi Menggunakan Software R. *Jurnal Magistra No. 92 Th. XXVII*. Klaten : FKIP Universitas Widya Dharma.
- Sutton, Clifton D. (2005). *Classification and Regression Trees, Bagging, and Boosting*. Handbook of Statistics Vol. 24.  
doi : doi.org/10.1016/S0169-7161(04)24011-1
- Tjaysono, Bayong. (2004). *Penuntun Praktikum Agroklimatologi Program Studi Agronomi*. Medan : Fakultas Pertanian Universitas Sumatera Utara.
- Turban, E., Aronson, Jay E., & Liang, Ting-Peng. (2007). *Decision Support System and Intelligent System Seventh Edition*. New Delhi : Prentice-Hal.
- Tusmar, Tito., & Mora, Minda. (2015). *Perkembangan Perekonomian Wilayah dan Kargo Udara : Korelasi atau Kausalitas*. Jakarta : Jurnal Perhubungan Udara.
- Wanyonyi, Kennedy Sifuna. (2019). *Comparison of Random Survival Forests Split Rules in Selecting the Determinants of Under-Five Mortality in Kenya using 2014 DHS Data*. (Master Thesis). University of Nairobi.
- Wienke, Andreas. (2007). *Frailty Models in Survival Analysis*. Jerman : ULB Sachsen-Anhalt.  
doi : dx.doi.org/10.25673/2651
- Wright, Marvin N., Dankowski, Theresa., & Ziegler, Andreas. (2017). *Unbiased split variable selection for random survival forests using maximally selected rank statistics*.  
doi : doi.org/10.1002/sim.7212

Wu, Cheng-Lung., & Truong, Tiffany. (2014). Improving the IATA Delay Data Coding System for Enhanced Data Analytics. *Journal of Air Transport Management*.

doi : [dx.doi.org/10.1016/j.jairtraman.2014.06.001](https://dx.doi.org/10.1016/j.jairtraman.2014.06.001)

Yang, Bo Suk., Di, Xiao., & Han Tian. (2008). Random forests classifier for machine fault diagnosis. *Journal of Mechanical Science and Technology* 22.

doi : [doi.org/10.1007/s12206-008-0603-6](https://doi.org/10.1007/s12206-008-0603-6)