



Air Systems

www.thalesgroup.com

Online Learning for Ground Trajectory Prediction

SESAR INNOVATION DAYS 2012

Gaétan Marceau Caron

Marc Schoenauer

Pierre Savéant

Areski Hadjaz



THALES

- **Description of the problem**
- **Our approach**
- **Model Definition**
- **Trajectory Generation and Validation**
- **Optimization Problem**
- **Model Validation Results**
- **Online Learning Methodology**
- **Online Learning Results**
- **Questions**

· How to increase the accuracy of the trajectory prediction ?

◆ Models from Physics (Point-mass, Total-energy, 6DOF, ...)

● Benefits:

- Lot of literature and studies (over 282 documents)
- Well-defined with system dynamics (Mathematics)

● Disadvantages:

- Parameters are required (mass, target speed, aircraft type, ...)
- Can be complicated to manipulate (Optimal Control Law)

◆ Models learned from Data (Neural Network, Regression, ...)

● Benefits:

- The model is automatically learned from the recordings with all observable parameters

● Disadvantages:

- Based on the assumptions that the system is well represented in the dataset and may be time-invariant.

- ◆ **Restrict the scope of the problem to climbing phase**
- ◆ **Use solely the information about the current flight**
- ◆ **Solve an online estimation problem of the parameters**
 - Use the total-energy model and a hybrid system to generate trajectories
 - Tune automatically the parameters of the model by comparing the trajectory and the recordings in a black-box setting.
 - Reiterate when new recordings are available
- ◆ **Tuning is done with a machine learning approach**
- ◆ **The dynamic system is considered as a black-box for the underlying optimization algorithm**
- ◆ **The black-box optimization algorithm is an Evolution Strategy called CMA-ES.**

- **Assumptions:**

1. **Coordinate system as a flat earth model**

- Earth is flat, non-rotating and can be defined as an inertial reference system in the vicinity of the aircraft position
- Gravitational force is constant and perpendicular to the ground

2. **The atmosphere is at rest relative to the earth and atmospheric properties solely depend on the altitude**

3. **The aircraft intention is to reach the flight level with a continuous climb**



• Total-Energy Model and Hybrid System

- ◆ The aircraft state is the altitude and the speed
- ◆ The dynamic system is defined by two equations:

- Evolution of the altitude

$$\dot{h}(V, q) = \frac{T(h) - \Delta T}{T(h)} \left[\frac{(Thr(h) - D(V, h)) \cdot V}{m \cdot g} \right] \cdot f(V, h, q) \quad (1)$$

Atmospheric Effect
Aircraft Dynamics

Energy Share Factor
(Depends on the mode q)

- Evolution of the speed

$$\dot{V}(h, \dot{h}) = \frac{1}{m} (Thr(h) - D(V, h) - m \cdot g \cdot \sin(\gamma)) \quad (2)$$

• Where $\gamma = \frac{\dot{h}}{V}$

- ◆ Mode definition

- Transition Altitude (CAS, MACH)
- Tropopause
- Acceleration



- **Integration with a Runge-Kutta fourth order**
- **The validation of the model consists in determining the expressiveness of the model**
 - ◆ **The expressiveness is the trajectory space obtained by modifying the parameters**
 - **Mass**
 - **Speed Schedule (3 speeds)**
- **One way to evaluate it is to compute the distance between a real trajectory and the closest one in the generated trajectory set**
 - ◆ **Distance: sum of position errors**
- **This can be formulated as an optimization problem.**



$$\theta_{i,j}^* = \operatorname{argmin}_{\theta \in \Theta} \sum_{n=i}^j |\mathcal{H}(\theta; s_0)_n - T_n|$$

Search in the parameter space

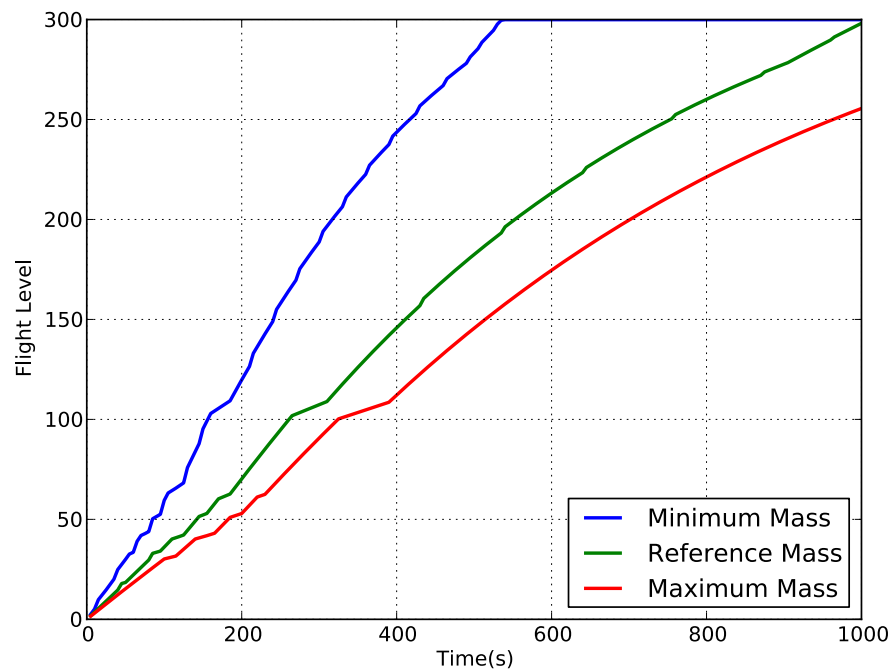


Figure 1. The Effect of the mass parameter

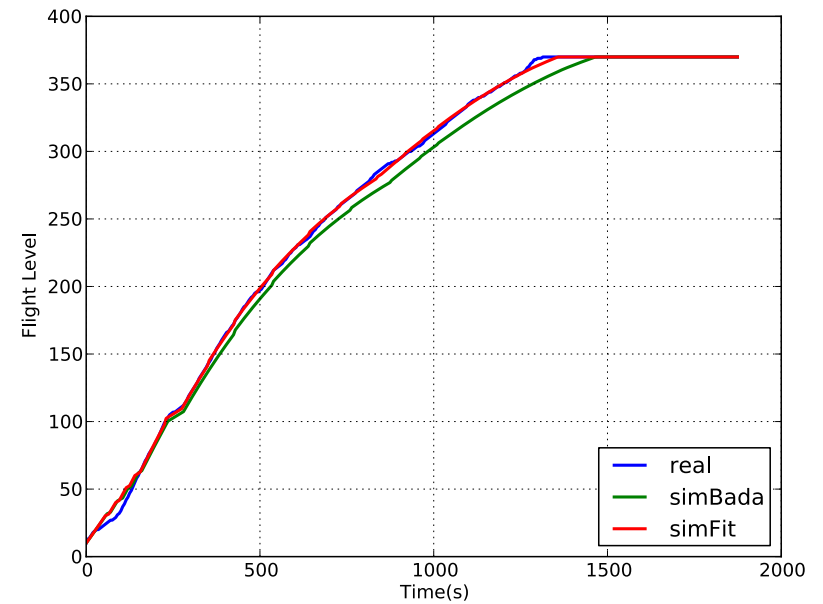
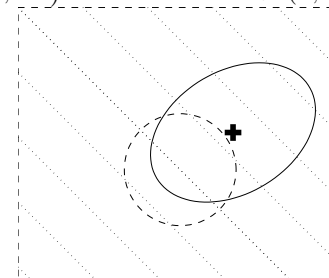
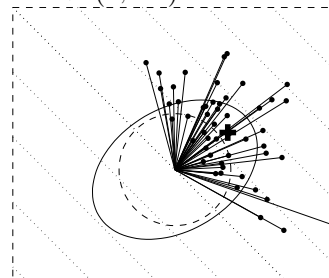
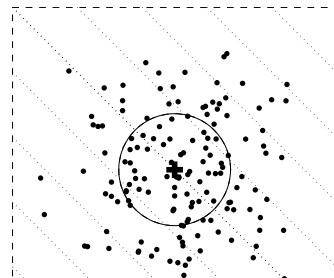
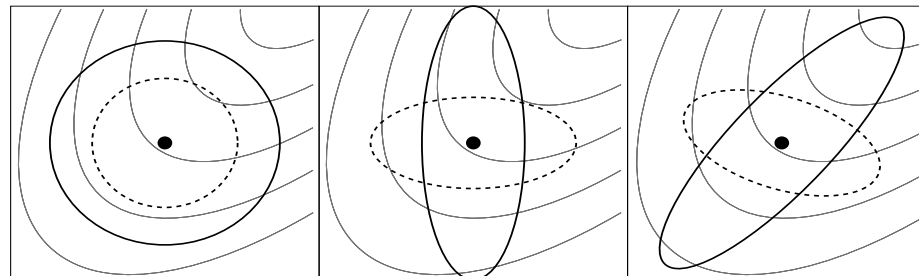


Figure 2. Trajectory Fitting

• Covariance Matrix Adaptation Evolution Strategy (CMA-ES)

- ◆ Stochastic, derivative-free method for numerical optimization in black-box context.
- ◆ Parameters are sampled with a multivariate Gaussian Distribution
- ◆ Pairwise dependencies between variables are represented by a covariance matrix



$$C_{\mu}^{(g+1)}$$

- ◆ References: <http://www.lri.fr/~hansen/cmaesintro.html>

- **In the model validation, all the points of the trajectory are known. (Complete information)**
- **Dataset : 262 trajectories of A320**
 - ◆ Cruise level at least of 300 FL
 - ◆ One position / 5 seconds
 - ◆ Continuous climbing

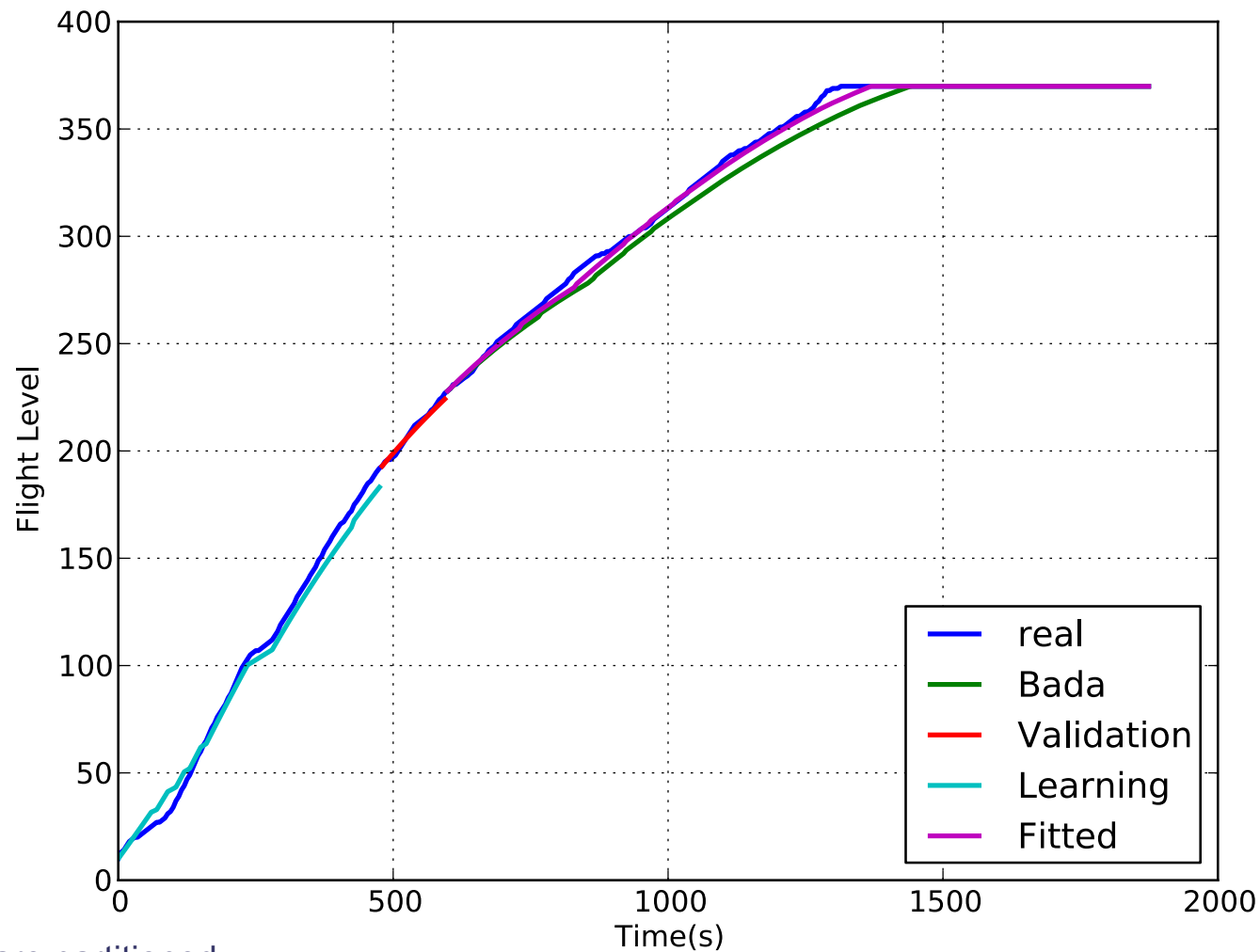


TABLE I
MODELIZATION ERRORS - MEAN AND STANDARD DEVIATION

Time after takeoff	Nominal (FL)	Tuned (FL)
2min.	4.9195 (3.1422)	3.0929 (2.3133)
5min.	7.1416 (4.8556)	2.5496 (2.5282)
10min.	9.6714 (6.6146)	1.4057 (1.7441)
15min.	10.9441 (9.0016)	2.1957 (2.2600)
20min.	11.8008 (8.8068)	2.0546 (2.1367)

- ◆ **Naive approach: find the trajectory in the trajectory space that is the closest to the observed one so far**
 - Leads to trajectories that are unlikely to occur (Overfitting)
 - Reminder: at 2 minutes, the model mean error is the largest (initial climb)
- ◆ **A solution to overfitting is to use regulation**
 - The hyper-parameter λ is learned through a cross-validation process.

$$\theta_{0,t}^* = \underset{\theta \in \Theta}{\operatorname{argmin}} \left[\underbrace{\sum_{i=0}^t \alpha_i |\mathcal{H}(\theta; s_0)_i - O_i|}_{\text{Weighted-sum of the errors}} + \lambda \underbrace{\sum_{i=0}^{|\theta|} |\theta_i - \theta_i^d|}_{\text{Penalty for deviating from the BADA nominal}} \right] \quad (4)$$



Observation are partitioned:

- Learning Set
- Validation Set

- ◆ **Wilcoxon signed-rank test is used to determine if the Online Predictor is **statistically better** than the nominal BADA.**
 - H0: Mean difference between the pairs is zero
 - H1: Mean difference is not zero
 - If $p\text{-value} < 0.05$, choose H1

TABLE II
COMPARISON OF ONLINE MODELS AT $P = 400s$

Time after takeoff	Nominal (FL)	Tuned (FL)	p-value
2min.	3.3029 (2.6698)	3.1699 (2.6740)	0.3401
5min.	6.7553 (5.6084)	6.5518 (5.6578)	0.6726
10min.	8.7851 (7.0757)	9.1846 (7.5687)	0.4541

TABLE III
COMPARISON OF ONLINE MODELS AT $P = 500s$

Time after takeoff	Nominal (FL)	Tuned (FL)	p-value
2min.	4.0406 (3.2758)	3.2834 (2.7237)	5.612e-4
5min.	8.1290 (6.0885)	7.0567 (4.8281)	0.02049
10min.	9.0872 (7.0085)	9.4205 (6.7658)	0.7939

TABLE IV
COMPARISON OF ONLINE MODELS AT $P = 600s$

Time after takeoff	Nominal (FL)	Tuned (FL)	p-value
2min.	4.5110 (3.4354)	3.5912 (2.4845)	1.289e-05
5min.	6.7936 (4.9209)	5.7231 (4.0456)	1.289e-03
10min.	8.5131 (6.6410)	9.4992 (7.8805)	0.09098



- **Define a hybrid system for aircraft simulation with a Runge-Kutta fourth order Integration Scheme**
- **Use an Evolving Strategy (CMA-ES) to do parameter estimation of a Hybrid System**
- **Cast the problem into a Machine Learning Problem to avoid the overfitting problem**
- **Demonstrate that we can gain accuracy by doing parameter estimation only with observed positions**

Thank you for your attention

Questions or comments ?