

Encoder-Decoder Approach to Predict an Airport's Operational Runway Configuration

A case study for Amsterdam Schiphol airport



9th SESAR Innovation Days

Ramon Dalmau & Floris Herrema

Network (NET) and Airports (APT) Research Units

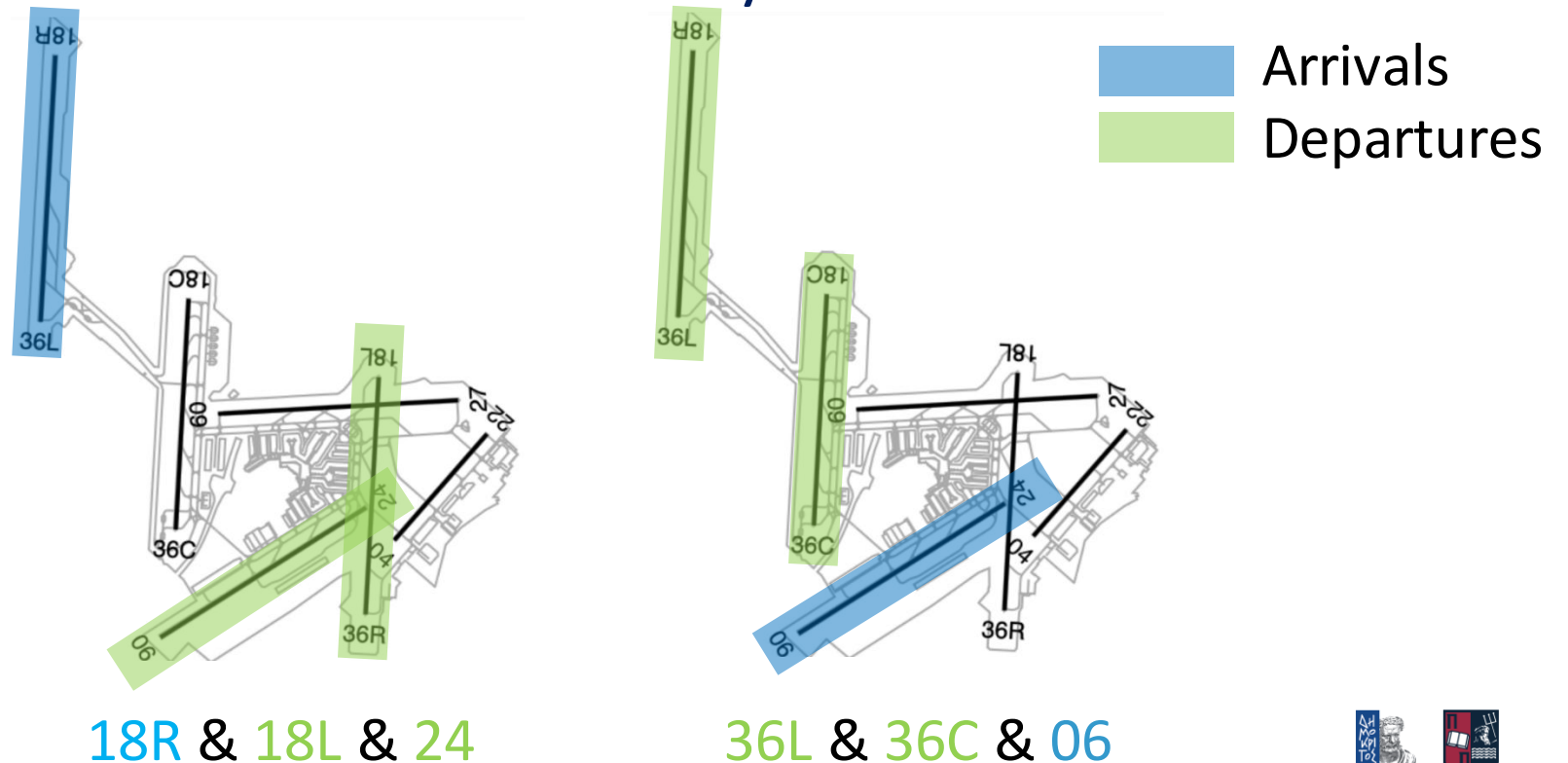
EUROCONTROL Experimental Centre (EEC)

Outline

- Introduction
- Sequence to Sequence (seq2seq) models
- Proposed approach
- Setup of the experiment
- Results
- On-going and future work

Introduction

- **Runway configuration:** combination of runways that are active at any time

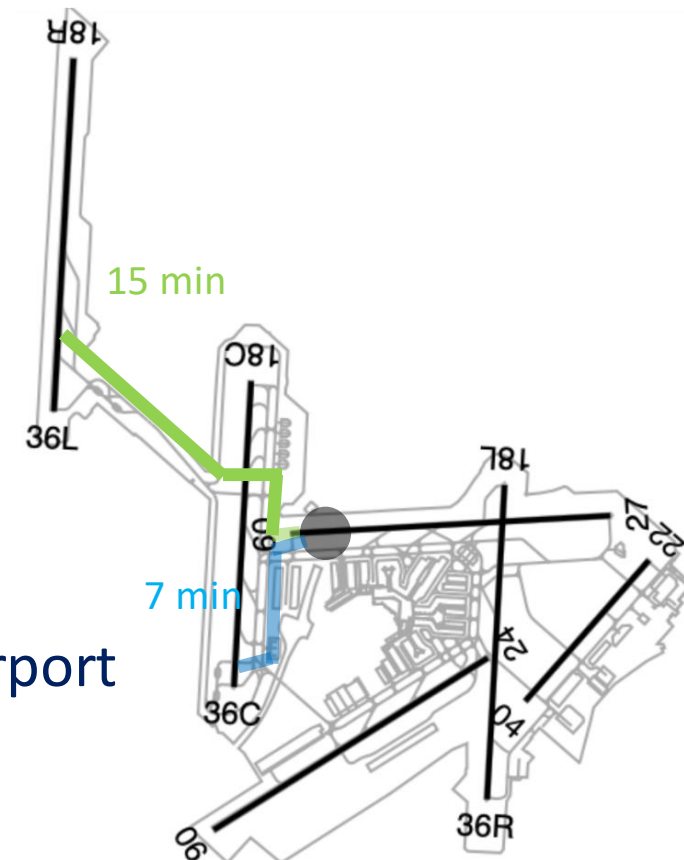


Example of possible runway configurations at Amsterdam Schiphol (**EHAM**).

Introduction

- The runway configuration influences:
 - **Capacity** of the airport
 - **Taxi-in and out times**
 - **Occupation** of parking stands and taxiways
 - Top Of Descent (**TOD**) of arriving aircraft 
 - **Management of traffic** in the airspace surrounding the airport

Knowing the runway configuration well in advance would improve predictability



*Note: Taxi times are illustrative, and **may not correspond to real values***



C. E. V. Gallego, V. F. G. Comendador, M. A. A. Carmona, R. M. A. Valdés, F. J. S. Nieto, M. G. Martínez, *A machine learning approach to air traffic interdependency modelling and its application to trajectory prediction*, Transportation Research Part C: Emerging Technologies (2019) 356–386. doi:10.1016/j.trc.2019.08.015

Introduction

- The selection of runway configuration is based on **human experience**
- Air Traffic Controllers (**ATCOs**) examine several factors:

- Wind direction and speed
- Visibility
- Precipitation / icing
- Temperature

Weather

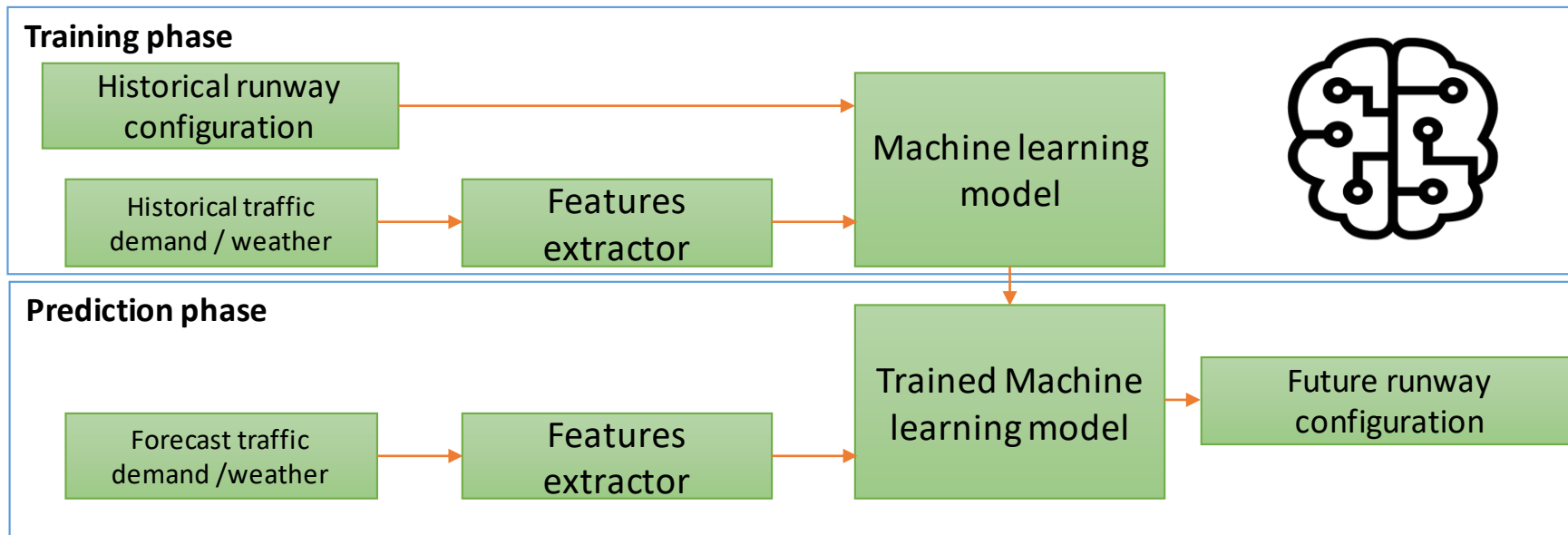
- Planned number of arrivals and departures
- ICAO Wake Turbulence Category of aircraft (Heavy, Medium, Light)
- Environmental rules (i.e., noise nuisance)

Traffic demand

Is it possible to predict ATCOs decisions knowing these factors?

Introduction

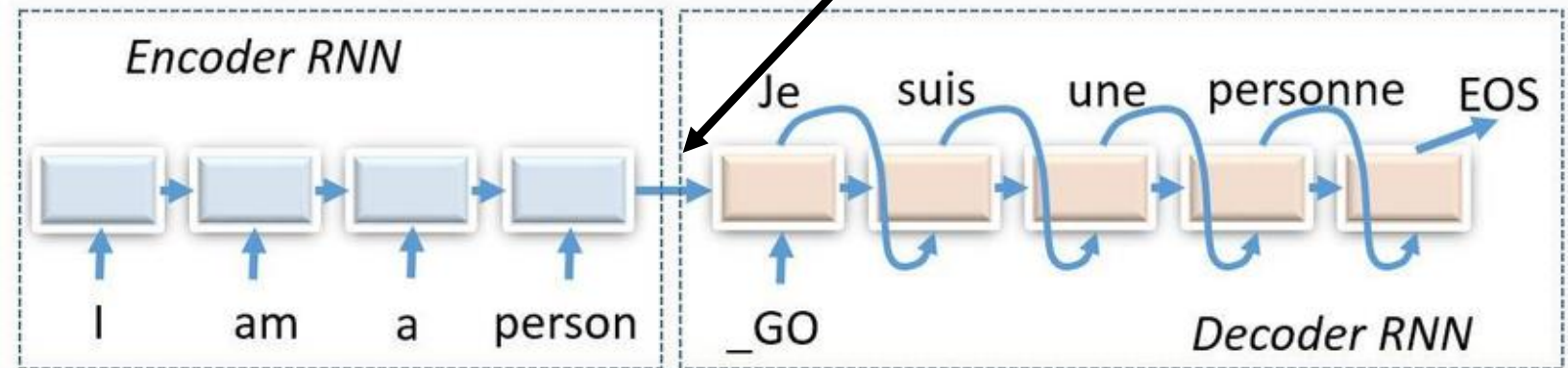
- Coding **rule-based models** for ATCOs decision is difficult and tedious. Each airport may implement different rules. Different models for different airports
- **Machine learning** could be used to train a model based on historical data. Same model with different parameters could be used in different airports



Sequence to sequence (seq2seq) models: Introduction

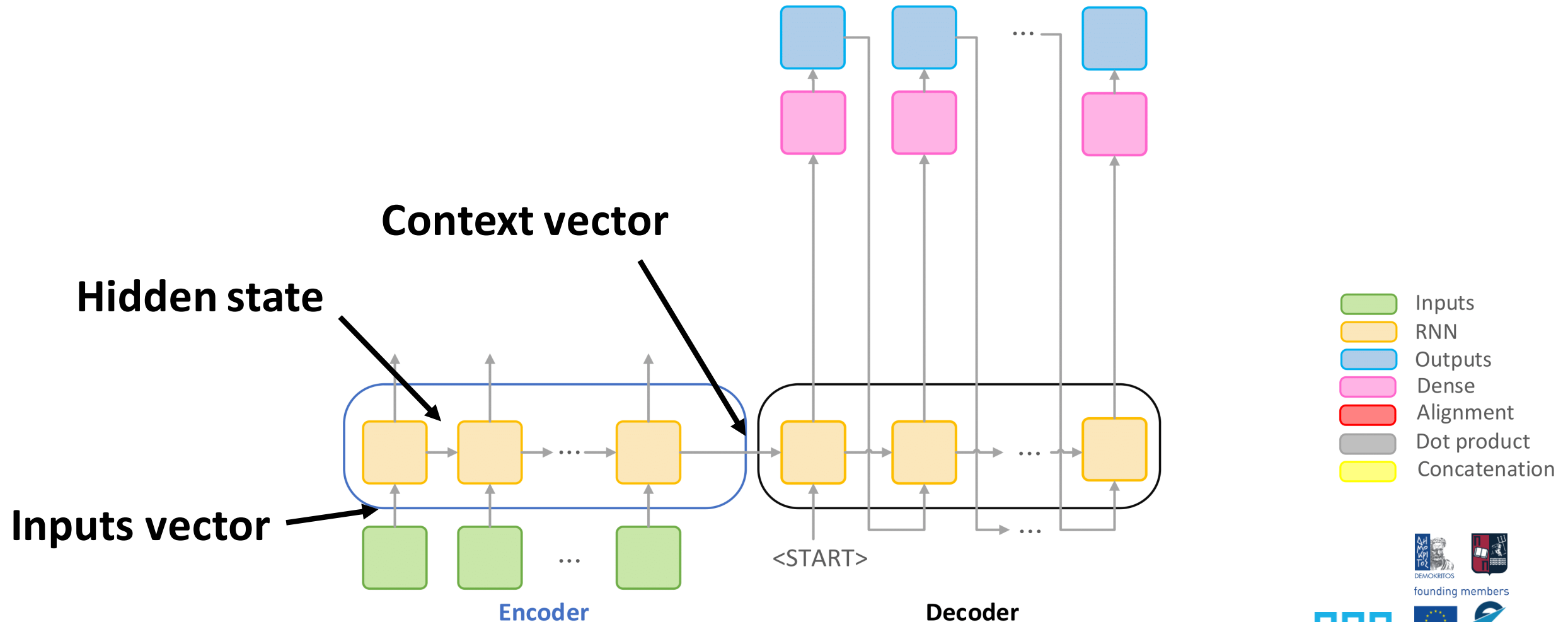
- Examples:
 - Neural machine translation
 - Text summarization
 - Speech recognition
 - Video captioning
 - Online chat-bots

Context vector: representation of the input sequence

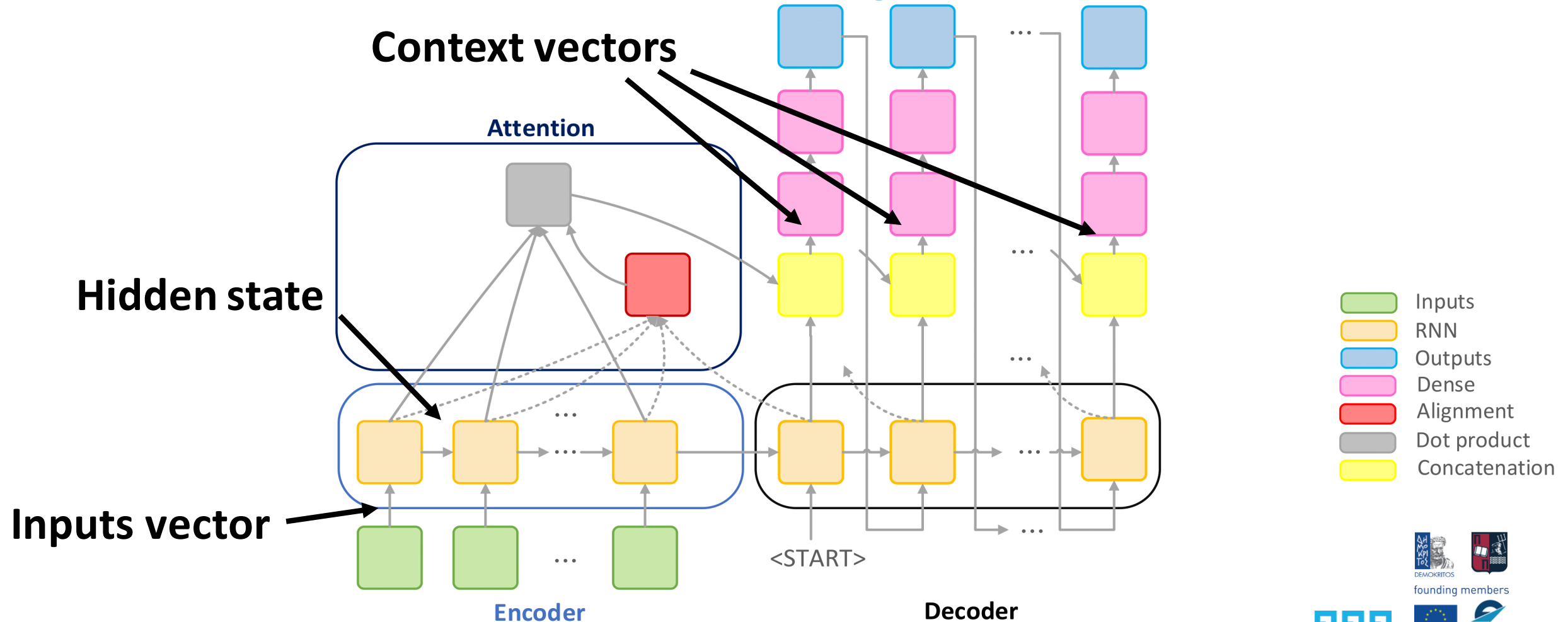


Example: A sequence-to-sequence RNN English to French translator with the encoder and decoder unrolled to show the flow of messages. Source: Fun with Recurrent Neural Nets - One More Dive into CNTK and TensorFlow, by Dennis Gannon

Sequence to sequence (seq2seq) models: basic encoder-decoder



Sequence to sequence (seq2seq) models: encoder-decoder with Luong's attention



Sequence to sequence (seq2seq) models

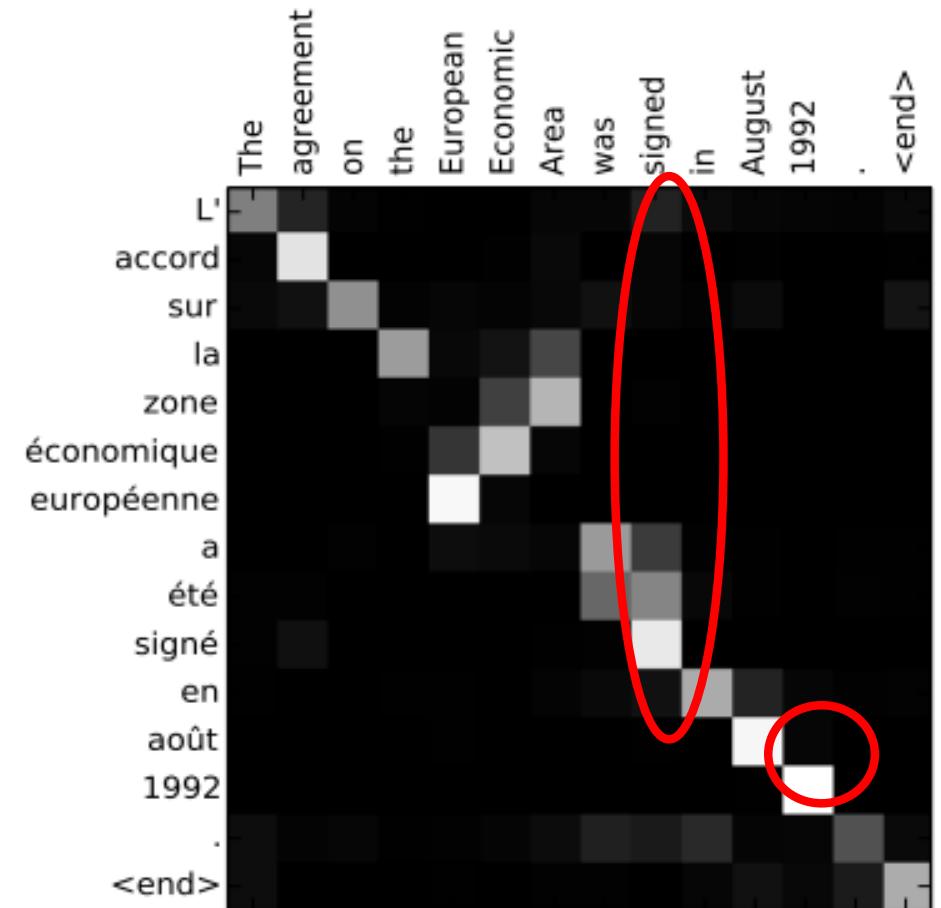
Attention weights example

Attention weights vector at decoding time step t :

$$\vec{\alpha}_t = \text{Alignment}(\vec{s}_t, h) = \frac{\exp(\text{score}(\vec{s}_t, h))}{\sum_{k=1}^{T_x} \exp(\text{score}(\vec{s}_t, \vec{h}_k))}$$

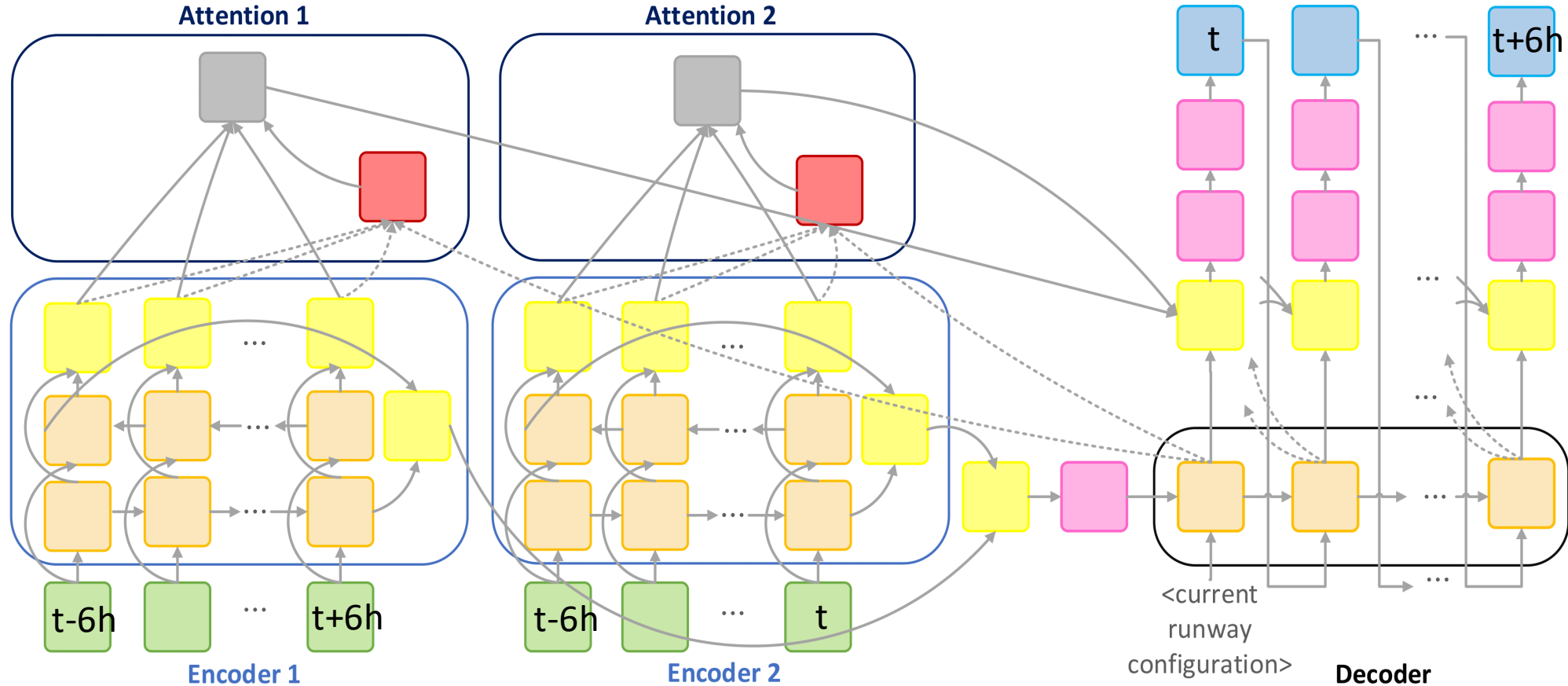
Context vector at decoding time step t :

$$\vec{c}_t = \vec{\alpha}_t \cdot h = \sum_{k=1}^{T_x} \alpha_{tk} \vec{h}_k$$



Source: Neural Machine Translation by Jointly Learning to Translate and Align by
Dzmitry **Bahdanau**, KyungHuyn Cho, and Yoshua Bengio. ICLR 2015.

Proposed approach: Architecture



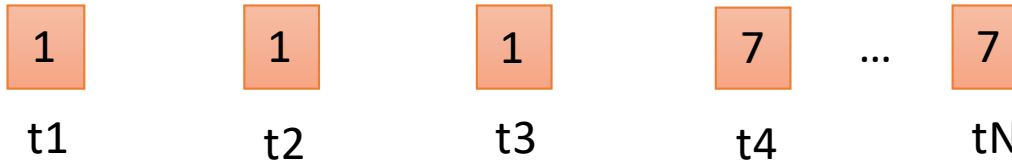
Encodes **observed** and **forecast weather** and **traffic demand** sequence

Encodes the **most recent runway configuration** sequence

Proposed approach: Outputs of the decoder

Single-output multi-class

Configuration

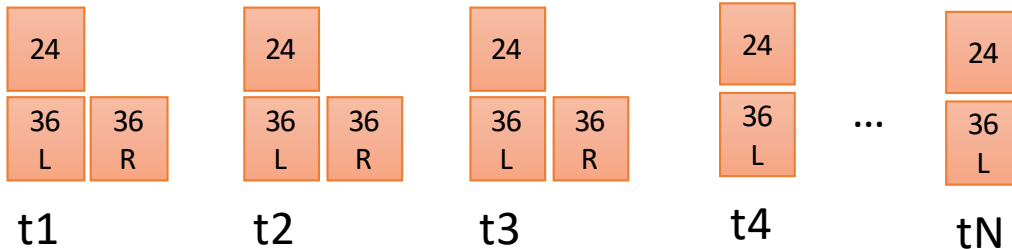


Active runways	ID
36L & 36R & 24	1
36L & 24	2
...	
18R & 24	M

Multi-output multi-label

Landing RWYs

Take-off RWYs



One sample every **15 minutes**
Forecast horizon of **6 hours**

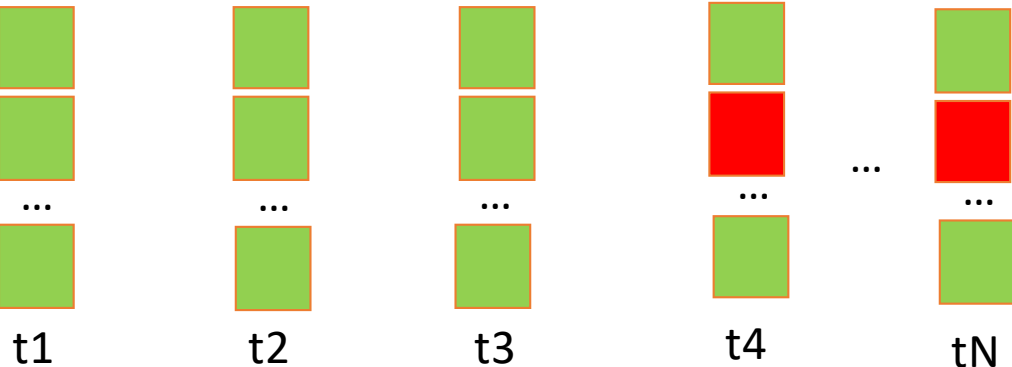
Multi-output multi-class

36L RWY state

36R RWY state

...

24 RWY state



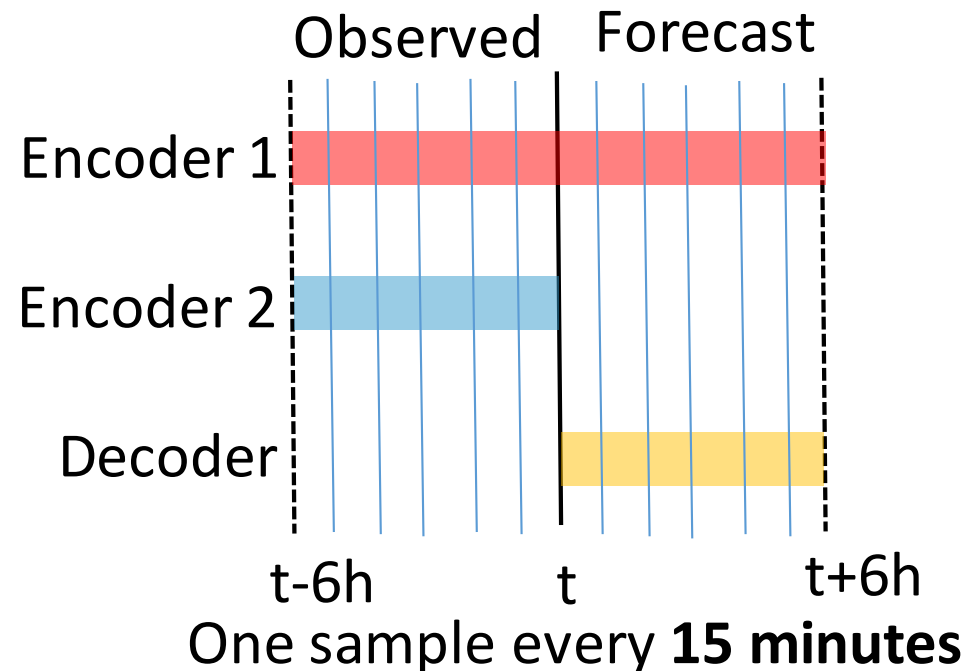
Proposed approach: Inputs of the encoders

Encoder 1

Group	Type	Feature
Weather	Continuous	wind direction wind speed wind gust temperature dew point ceiling visibility cloud height
	Categorical	cloud cover cloud type precipitation obscuration
Demand	Continuous	# of H departures in a 15/30/60 min interval # of M departures in a 15/30/60 min interval # of L departures in a 15/30/60 min interval # of H arrivals in a 15/30/60 min interval # of M arrivals in a 15/30/60 min interval # of L arrivals in a 15/30/60 min interval
Calendar	Categorical	Hour of the day Day of the week Week of the month Month of the year

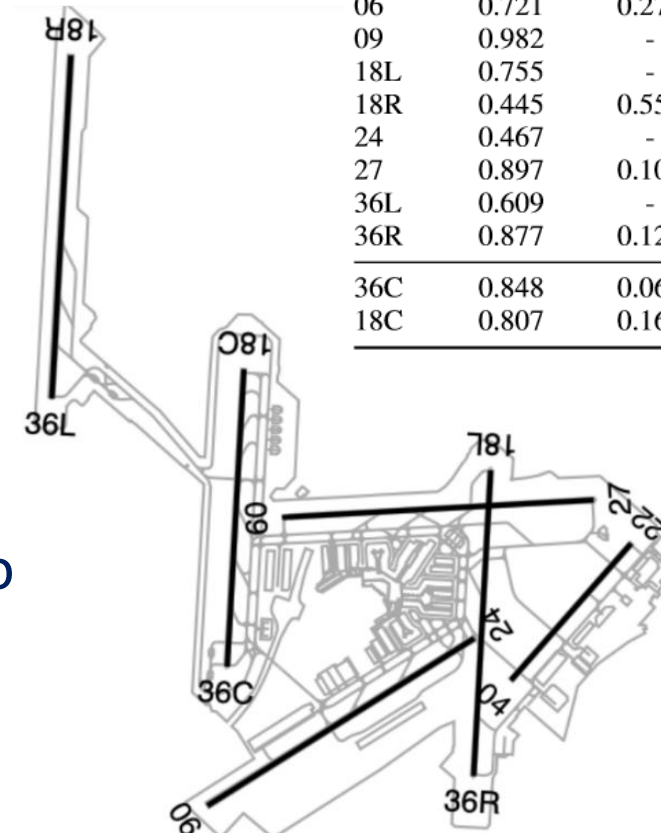
Encoder 2

Group	Type	Feature
Runway use	Categorical (binary)	runway r state $\forall r \in \mathcal{R}_b$
	Categorical (multi-class)	runway r state $\forall r \in \mathcal{R}_m$



Set up of the experiment: Data Set

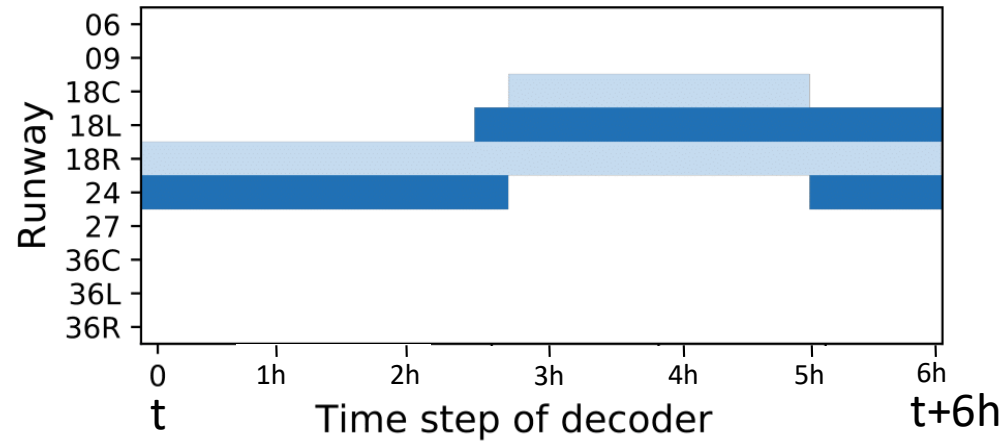
- 1st of January 2016 to 1st of May 2019
- Number of arrivals and departures from EFD
- Aircraft category from ICAO DOC 8643
- Weather data from METAR
- Runway configuration history from LVNL
- Re-sampled in intervals of 15 minutes
- Only 20 of the 142 observed runway configurations where used more than 1% of the time. The accumulated frequency of these 20 runway configurations is 87%



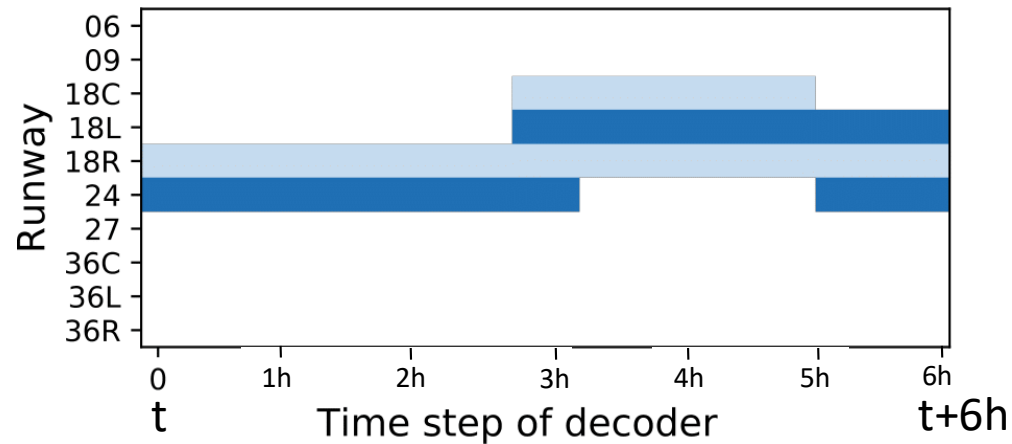
RWY	Frequency			type
	Inactive	Landing	Takeoff	
06	0.721	0.279	-	binary
09	0.982	-	0.018	
18L	0.755	-	0.245	
18R	0.445	0.555	-	
24	0.467	-	0.533	
27	0.897	0.103	-	
36L	0.609	-	0.391	
36R	0.877	0.123	-	multi-class
36C	0.848	0.066	0.086	
18C	0.807	0.164	0.029	

Results: Illustrative example

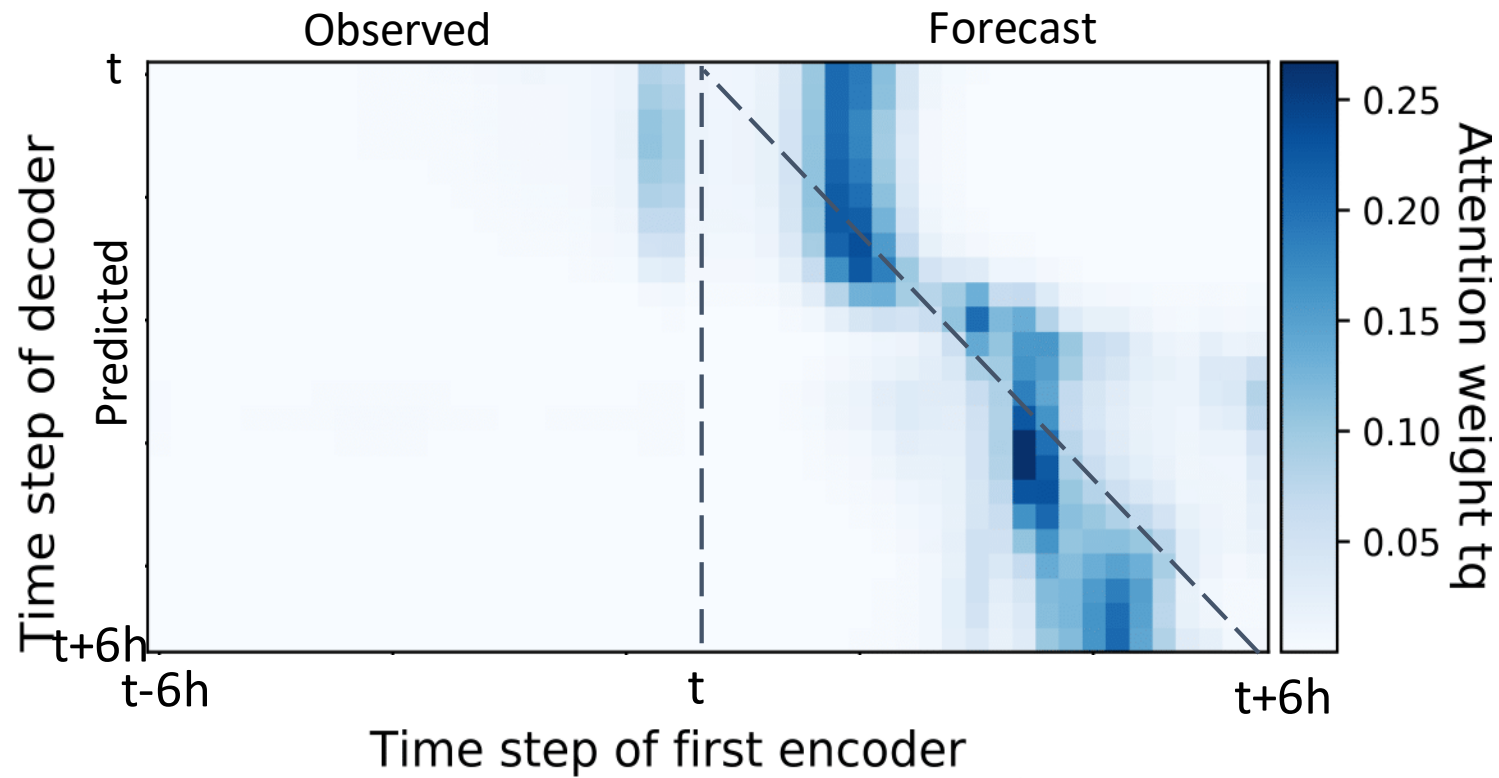
Predicted



Ground truth



Results: Illustrative example



Results: Optimal hyperparameters

hyperparameter	value
Attention score function	general (see Ref. )
Type of RNN encoder / decoder	LSTM / LSTM
Units of the encoder's RNN (n_h)	16
Units of the decoder's RNN (n_s)	32
Attention vector length (n_a)	32
batch size	64
learning rate	0.001
γ	2
Training epochs / early stopping patience	50 / 5



T. **Luong**, H. Pham, and C. D. Manning, “Effective Approaches to Attention-based Neural Machine Translation,” in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal: Association for Computational Linguistics, September 2015, pp. 1412–1421

Results: Aggregated performance metrics

RWY	precision			recall			f1-score		
	(0,2]	(2,4]	(4,6]	(0,2]	(2,4]	(4,6]	(0,2]	(2,4]	(4,6]
06	0.98	0.96	0.96	0.98	0.97	0.97	0.98	0.97	0.96
09	0.99	0.99	0.99	1.0	1.0	1.0	0.99	1.0	0.99
18C	0.9	0.86	0.83	0.9	0.86	0.84	0.9	0.86	0.84
18L	0.86	0.82	0.82	0.91	0.9	0.88	0.89	0.86	0.85
18R	0.92	0.87	0.84	0.91	0.85	0.8	0.91	0.86	0.82
24	0.9	0.88	0.88	0.83	0.73	0.71	0.86	0.8	0.79
27	0.98	0.97	0.96	0.98	0.97	0.97	0.98	0.97	0.96
36C	0.95	0.94	0.94	0.96	0.94	0.94	0.96	0.94	0.94
36L	0.97	0.95	0.95	0.97	0.96	0.96	0.97	0.96	0.95
36R	0.99	0.98	0.98	0.98	0.97	0.97	0.98	0.98	0.97

On-going and future work

- Feature importance analysis
- METAR vs TAF
- Actual vs planned demand
- Model's complexity vs accuracy trade-off
- Use output to improve taxi time prediction
- SID and STAR prediction

Thank you!