DATS

Capacity Modeling for Controller Workload Evaluation and Optimization of Staff Planning at RTC

Tatiana Polishchuk *Christiane Schmidt* *Billy Josefsson*

Based on joint work with: J. Jakobi, L. Meyer, A. Papenfuss, M. Peukert, L. Sedov

LINKÖPING UNIVERSITY

TEAMWORK with automation

LFV

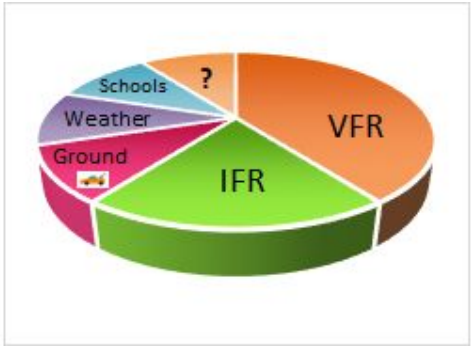
CAPMOD

Project description

- ✓ The project focuses on complete and descriptive ***capacity modelling,*** which will quantify the total controller's workload
- ✓ Continuation of **KODIC**, where we ***designed mathematical models for controllers rostering in a RTC***, using the **number of IFR flight movements** as an indicator of staff *workload*

As IFR traffic accounts for only ***~40% of the workload,*** we need to look at the other ***important aspects***:

- **ground traffic movements**
- **bad weather conditions**
- **VFR and extra traffic movements**



CAPMOD

Capacity **Mod**eling for Controller Workload Evaluation at RTC Arlanda

Motivation

- ✔ ***Mental workload:*** limitation on number of tasks a human can perform during a certain period of time
- ✔ ***Complexity measures influencing workload:*** the number of aircraft in a sector, voice messages, radar screen clicks, ground traffic movements, etc.
- ✔ Several studies for ***en-route traffic***
- ✔ ***New workload factors*** appear in connection with the emerging technologies (CPDLC, RTC).
- ✔ ***A generic single metric for workload measurement is missing***

The *importance of quantitative assessment* of controller mental workload was reported in many of our projects

Research Questions

- ✔ Which ***factors*** contribute to controller's ***workload***?
- ✔ How does the ***workload at RTC*** differ from the workload at ***traditional towers***?
- ✔ How do different ***weather*** conditions influence controller's workload?

Methods

- ✔ Simulation and data analysis
  - DLR simulation data, used Adapted Cooper-Harper Scale
  - Sundsvall validation trials (May-June 2019)
- ✔ Observations and data collection in traditional towers + data analysis
  - Field study at Bromma airport (March 2019) video-recording, questionnaires
- ✔ Objective vs. subjective assessment (workload rating vs. quantitative measures)
- ✔ Mathematical analysis vs. HF

image source:
http://clipart-library.com/clipart/54081.html

Simulation Data Analysis: DLR Dataset

✔ Identification of Complexity Factors Influencing Controllers Workload in Remote Towers (DLR data, used Adapted Cooper-Harper Scale, SID 2018):

Dataset from DLR [C. Möhlenbrink, A. Papenfuss, and J. Jakobi. The role of workload for work organization in a remote tower control center. Air Traffic Control Quarterly, 20(1):5, 2012]

- Six teams of ATCO pairs
- 1 controller + 1 observer (assessing workload)
- 12 ATCOs
- Airports: Erfurt and Braunschweig
- **Multiple remote operation**
- All simulations with "high" traffic volume
- 20 min scenarios
- 222 situations

Adapted Cooper-Harper Scale:

critical
(in terms of safety)

Rating	Evaluation	Question for Evaluation
1	No problems, desirable	Is the situation solvable without major Disturbance?
2	Simple, desirable	
3	Adequate, desirable	
4	Small, but disruptive "delays"	Is the situation solvable by capacity-reducing measures?
5	Medium loss of capacity, which can be improved	
6	Very disruptive, but tolerable difficulties	
7	Problems to predict development of traffic situation	Is the situation solvable if the ATCO works with a reduced situational awareness?
8	Problems in information processing	
9 9	Problems in information reception	
10	Impossible	

Goal: Identify critical complexity factors that drive the workload for a remote tower ATCO

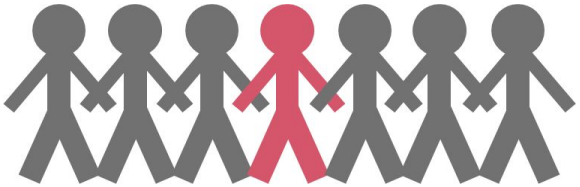
➲ Identify situations at the two controlled airports that induce risk

➲ Aggregate information w.r.t. combination of events: used pairs of events

Simulation Data Analysis: DLR Data

Mean Controller Rating:

- Assume an “average” controller
- Whether situation un-/manageable depends on experience, age, ….
- Targeting a generic measure



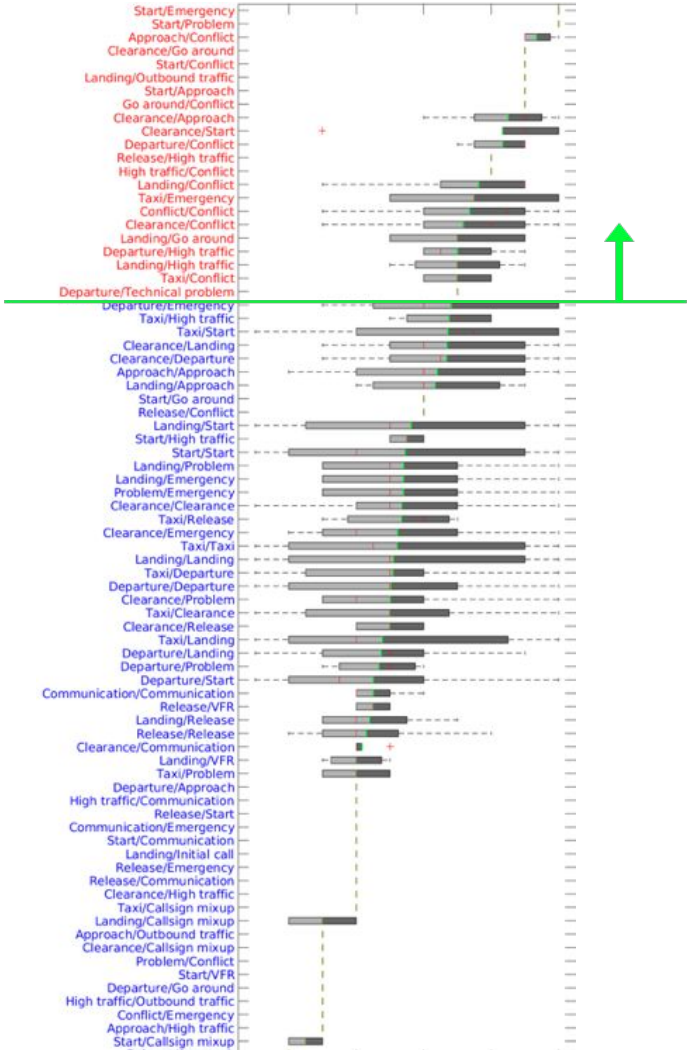
Maximum Controller Rating:

- More conservative
- Possibly only single ATCO rated as critically
- This way we identify all critical factors for the remote tower environment
- Exclude what is unmanageable for any ATCO

B. Josefsson, J. Jakobi, A. Papenfuss, T. Polishchuk, C. Schmidt, L. Sedov Identification of Complexity Factors for Remote Towers. In **SESAR Innovaon Days (SID 2018),** December 3-5, Salzburg.

CAPMOD

Mean Controller Rating



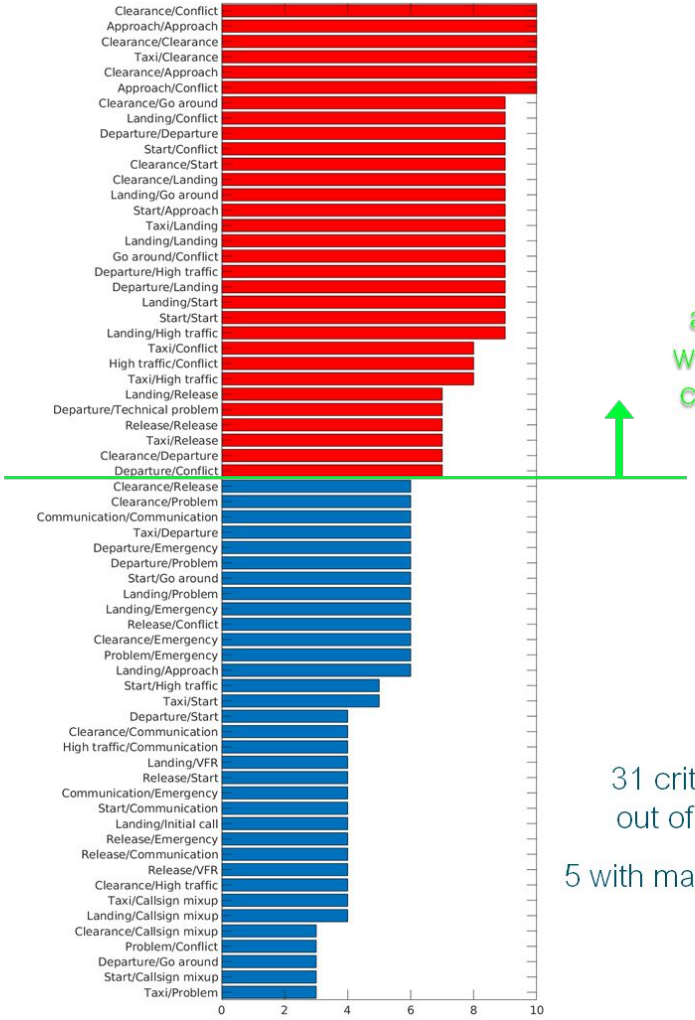
all event pairs with a mean controller rating of at least 7

17 critical event pairs out of 65

green: mean
red: median

CAPMOD

Maximum Controller Rating



all event pairs with a maximum controller rating of at least 7

31 critical event pairs out of 65 event pairs

5 with maximum rating of 10

Consequences of Events and Their Causing Factors

Situations / Consequences	Taxi	Clearance	Departure	Landing	Release	Start	Approach	Go around	Problem	Initial call	Technical problem	Callsign mixup	High traffic	Conflict	Commu-nication	Outbound traffic	VFR	Emergency
Monitoring problem	11.1%	0.0%	14.3%	13.6%	0.0%	0.0%	20.0%	0.0%	0.0%	0.0%	0.0%	0.0%	28.6%	0.0%	0.0%	0%	0.0%	0.0%
Small delay	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	20.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0%	0.0%	0.0%
Mix-up of airports	3.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%	0.0%	0.0%	0.0%	0%	0.0%	0.0%
Switching airports	3.7%	0.0%	0.0%	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0%	0.0%	0.0%
Communication problem	3.7%	40.9%	4.8%	6.8%	25.0%	4.5%	20.0%	0.0%	0.0%	0.0%	0.0%	10.0%	14.3%	12.5%	40.0%	0%	100.0%	0.0%

Problematic consequence can be indicator of risky situation

- ✦ Monitoring problem
- ✦ Small delay
- ✦ Mix-up of airports
- ✦ Switching airports
- ✦ Communication problem

OBSERVATIONS

- 40% of communication led to communication problem
- 100% of VFR traffic (when mentioned!!) led to communication problem
- 100% of mentions of VFR traffic coincided with communication problem
- Several situations never caused a problematic consequence (e.g., go-arounds)

B. Josefsson, J. Jakobi, A. Papenfuss, T. Polishchuk, C. Schmidt, L. Sedov Identification of Complexity Factors for Remote Towers. In **SESAR Innovaon Days (SID 2018),** December 3-5, Salzburg.

CAPMOD

Goal: proof of concept for the **validation of quantitative indicators on their workload predictability** in a conventional tower and in a Remote Tower

- ✔ Sundsvall validation trials (May-June 2019)
- ✔ Field study at Bromma airport (March 2019) video-recording, questionnaires

Methods

Derive quantitative measures from recorded video and communication data collected during two studies, candidate measures are, for example, the number of ATCO tasks and the response time to Situation Present Assessment Method (SPAM) queries.

ATCO tasks: arrival, clearance, communication, abnormal situation, departure, secondary task, taxi

Workload Rating

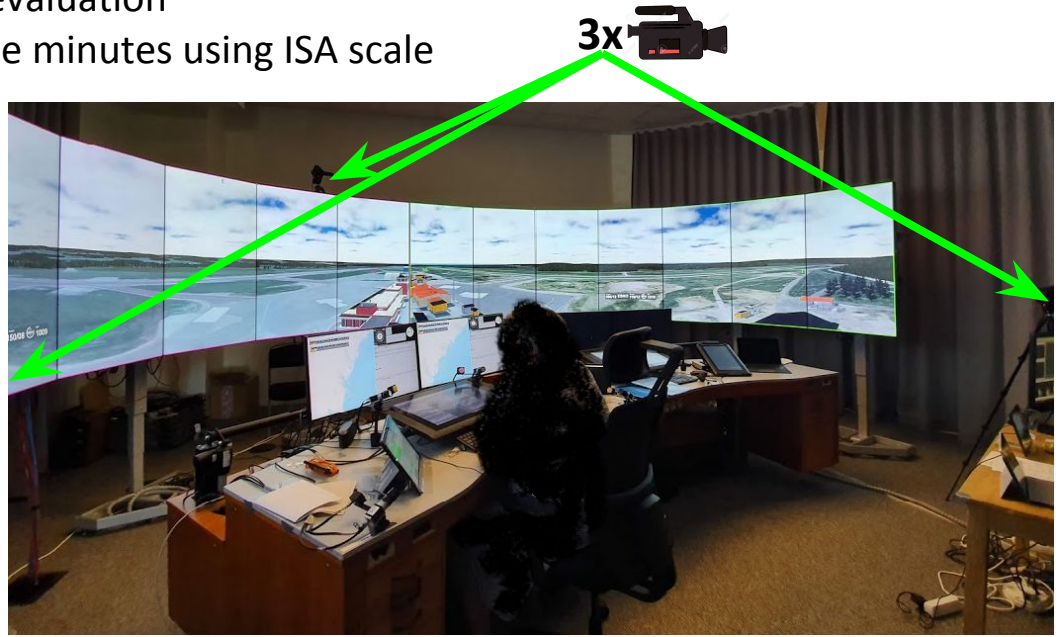
Rating	Evaluation	Question for Evaluation
1	No problems, desirable	Is the AT solvable without major Disturbance?
2	Simple, desirable	
3	Adequate, desirable	
4	Small, but disruptive "delays"	Is the AT solvable by capacity-reducing measures?
5	Medium loss of capacity, which can be improved	
6	Very disruptive, but tolerable difficulties	
7	Problems to predict development of traffic AT	Is the AT solvable if the ATCO works with a reduced ATal awareness?
8	Problems in information processing	
9	Problems in information reception	
10	Impossible	

Rating	Workload	Spare Capacity	Description	Possible Interpretation of CHS values
1	Underutilized	Very much	Little or nothing to do. Rather boring.	1
2	Relaxed	Ample	More time than necessary to complete the tasks. Time passes slowly.	2,3
3	Comfortable	Some	The controller has enough work to keep him/her stimulated. All tasks under control.	4,5,6
4	High	Very little	Certain nonessential tasks are postponed. Could not work at this level very long. Controller is working at the limit. Time passes quickly.	7,8,9
5	Excessive	None	Some tasks and not completed. The controller is overloaded and does not feel in control.	10

- ✔ Different scales for workload rating
- ✔ Adapted Cooper-Harper Scale (CHS)--as used for the DLR data
- ✔ Instantaneous Self Assessment (ISA) scale of workload

Simulation Data Analysis: Sundsvall Validation Trials 2019

- Remote Tower Center simulator
- Subjective vs. Objective workload evaluation
- Worload rating by ATCO every three minutes using ISA scale
- Three ATCOs
- Video data analysis:
  - # ATCO tasks
  - Communication time
  - Reaction time SPAM queries



subjective | objective

Workload vs. #ATCO Tasks/Weighted #ATCO Tasks

Weights?

- ✔ Average call duration for each ATCO task (AT) type in single and multiple mode (for each ATCO and as average over all ATCOs)
- ✔ Normalized for weights

	ATCO 3 single	ATCO 3 multiple	ATCO 2 single	ATCO 2 multiple	ATCO 1 single	ATCO 1 multiple	average single	average multiple
Arrival	24	9.2	28.5	13.67	10.83	11.5	21.11	11.46
Clearance	12.71	25.8	13.17	13.5	13	22.17	12.96	20.49
Comm	9.11	12.47	10.62	11.5	8.63	13.69	9.45	12.55
Taxi	20	18.2	8.75	5.33	12.6	8.5	13.78	12.04

Communication shows significantly higher values in multiple than in single mode (one-sided U-test, p-value 1.65%)

Increase in average communication times related to arrivals from multiple to single was nearly significant (one-sided Utest, p-value 7.57%)

Communication for clearances shows nearly significantly higher values in multiple than in single mode (one-sided U-test, p-value 6.7%)

probably caused by risk compensation behavior by the operator to avoid risk at the expense of time

CAPMOD

subjective

objective

ATCO WL: ISA scale

Workload vs. #ATCO Tasks/Weighted #ATCO Tasks Single Mode

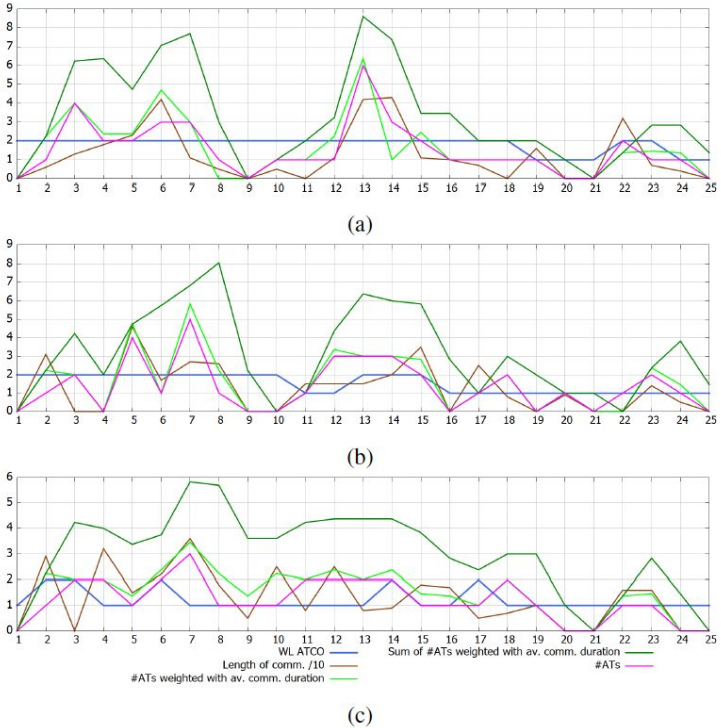
ATCO WL

#ATs

length comm.

#ATs weighted with av. comm. duration

sum of #ATs weighted with av. comm. duration



- **Necessary condition** for an increase in workload?
- The number of ATs is not a necessary condition for an increase in workload.
- ATCO 1: increase in workload rating is accompanied by an increase in all measures that take the communication time into account.
- ATCO 2: increase in the workload rating is accompanied by an increase in the sum of the number of ATs weighted with the average communication duration for two consecutive time periods.
- ATCO 3: *all but one* increase in workload rating is accompanied by an increase in the sum of average-communication-duration weighted ATs.

Increase in WL is accompanied by an increase in at least one of the metrics

CAPMOD

subjective

objective

ATCO WL: ISA scale

Workload vs. #ATCO Tasks/Weighted #ATCO Tasks Multiple Mode

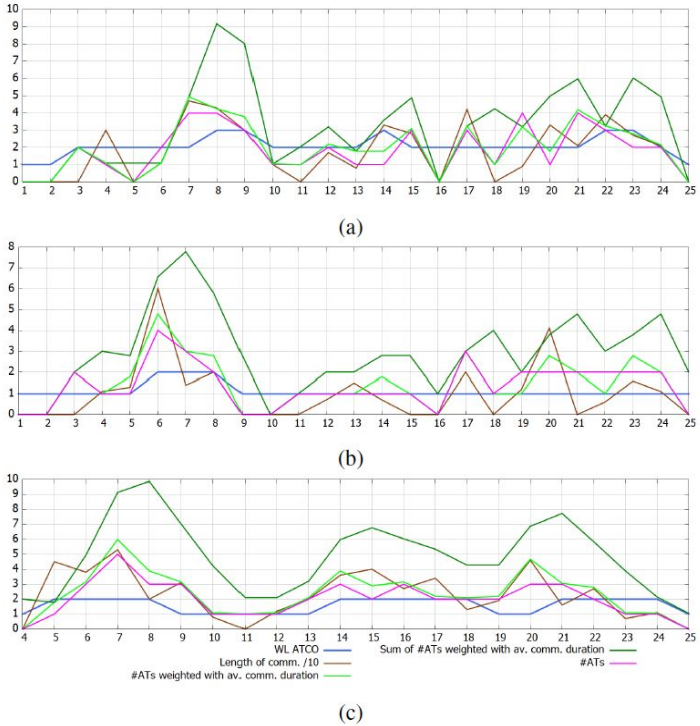
ATCO WL

#ATs

length comm.

#ATs weighted with av. comm. duration

sum of #ATs weighted with av. comm. duration



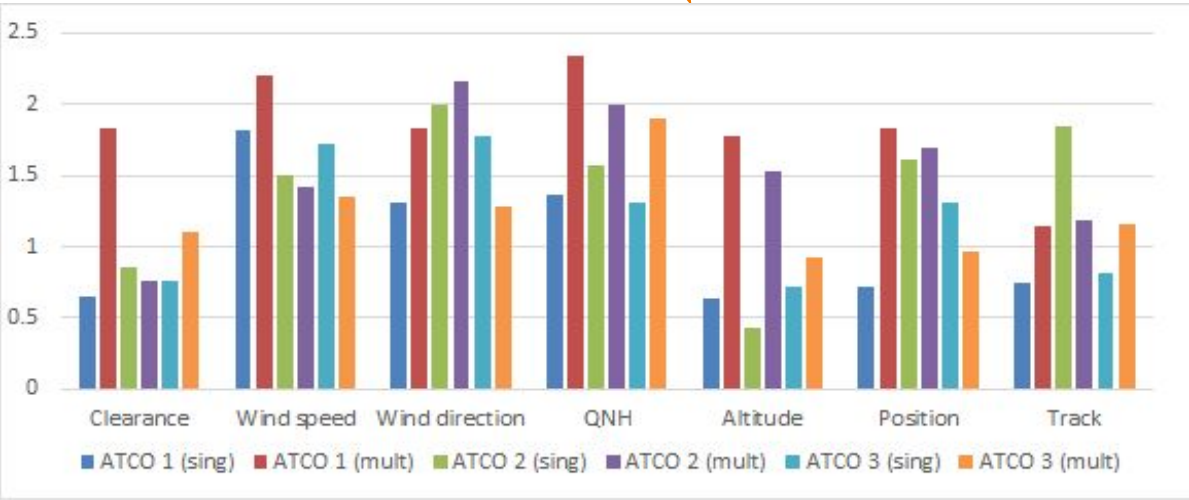
ATCO 1 endorsement only for Sundsvall,
➲ ATCO 1 was confronted with an unknown working environment

- **Necessary condition** for an increase in workload?
- Each increase in the workload rating (for all ATCOs) is accompanied by an increase of the duration of communication at that time interval or by an increase in the sum of average communication-duration weighted ATs for two consecutive time periods
- **Regression analysis**: results quite good (small data set, human subjects)
- Number of ATs weighted with the average communication duration for two consecutive time periods can be a good predictor for ATCO workload

Increase in WL is accompanied by an increase in at least one of the two metrics

Reaction Times: Single vs. Multiple Mode

Average reaction time for the three ATCOs for each SPAM query



- Most queries: reaction time in multiple mode increases vs. single mode.
- More tasks ➲ risk compensation behaviour ➲ can be indicator for increased stress
- Trend not true for all queries (ATCO 1 new environment in multiple, the others not)
- New working environment as stressor
- Emphasizes importance of **training**

Reaction times good indicator for stress, which might be caused by increased WL

Bromma Airport: Field Studies

- ✔ Real operations
- ✔ Again, video-recording, questionnaires
- ✔ Objective vs. subjective assessment
- ✔ # pre-defined ATCO tasks
- ✔ Communication duration
- ✔ Weather (snow sweeping)

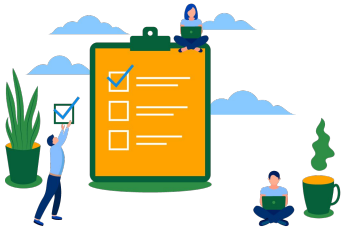
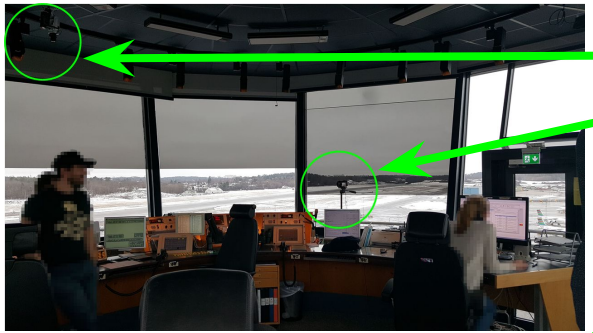
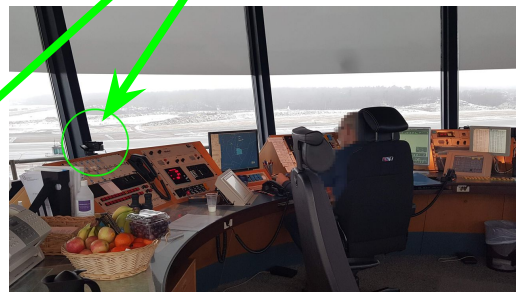


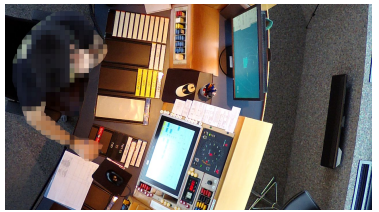
Image sources: www.typeform.com, Clipart Library, DZApk, Wikipedia

Bromma Airport: Experimental Setup

- ✔ 5 mounted video cameras
  - 3 facing ATCOs
  - 2 facing runway ends
- ✔ 4 hours of recording
- ✔ 4-27 movs (increasing intensity)
- ✔ 3 ATCOs at work
- ✔ 2x2 observers
- ✔ Subjective ratings using C+H scale
- ✔ Audio recordings (communication)

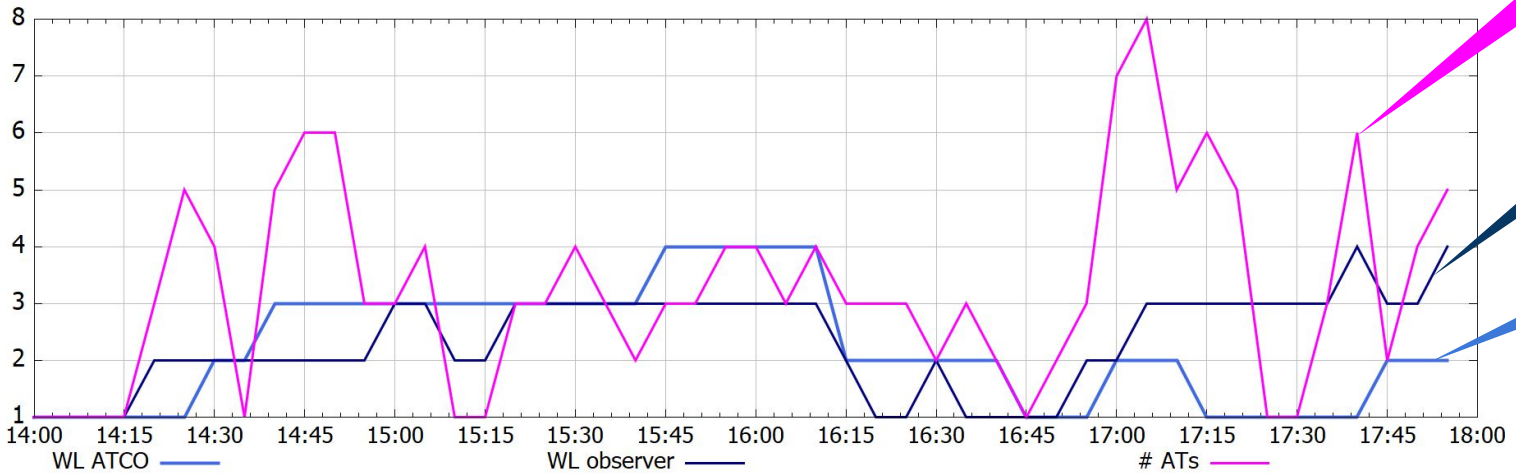






Bromma Airport: Field Studies

Subjective vs. Objective workload evaluation



ATCO tasks

Observer WL

ATCO WL

Increase in WL is accompanied by an increase in # ATCO tasks in the current or prev. time period

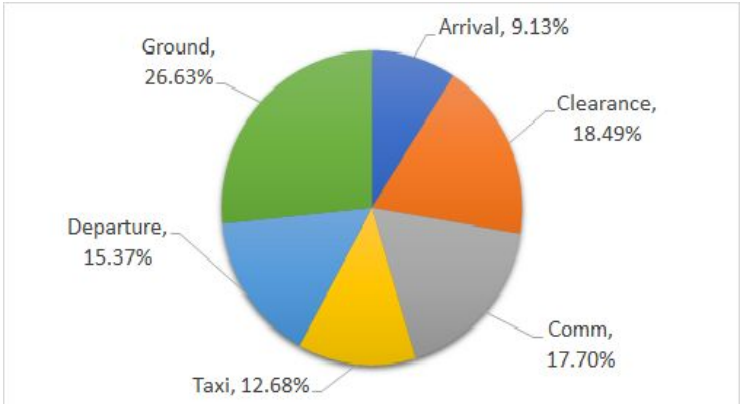
 # ATCO tasks alone does not explain variations in ATCO workload ratings

Bromma Airport: Field Studies

Radio Communication Duration

	Arrival	Clearance	Comm	Taxi
Average (in s)	10.04348	20.34783	11.2	10.7
Sum (in s)	231	468	448	321
Percentage	9.13%	18.49%	17.70%	12.68%
Range	6-16	6-57	4-72	5-28

	Departure	Ground	Total	
Average (in s)	11.44118	13.48	∅	
Sum (in s)	389	674	2531	
Percentage	15.37%	26.63%	100%	
Range	5-27	3-37		

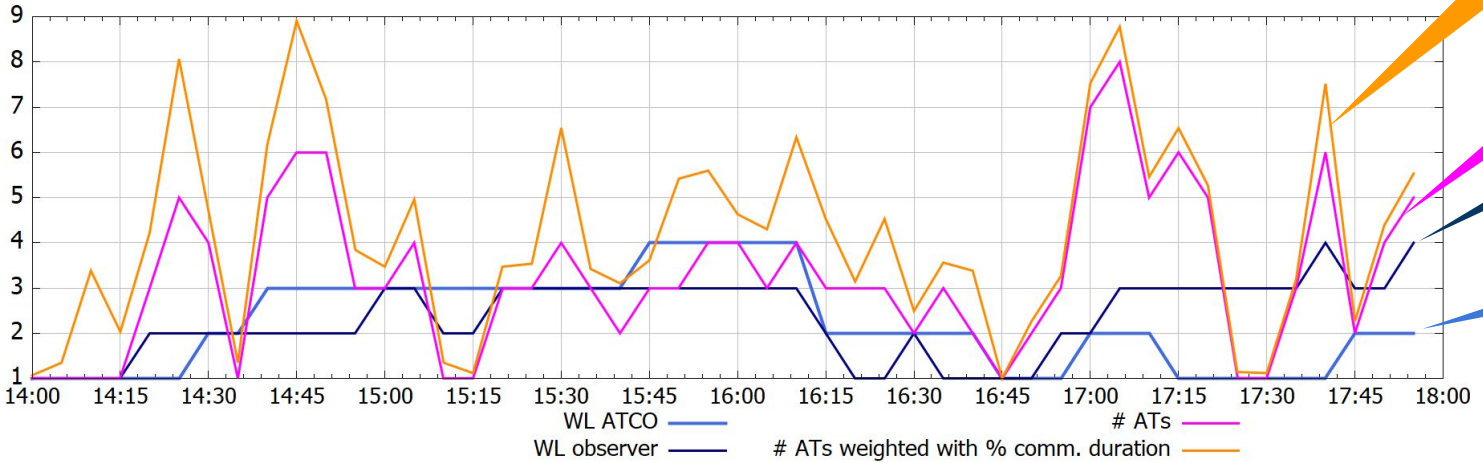


Clearances have the highest average communication duration

Ground communication takes the largest share in total communication duration

Bromma Airport: Field Studies

Subjective vs. Objective workload evaluation



ATCO tasks weighted with % comm. duration

ATCO tasks

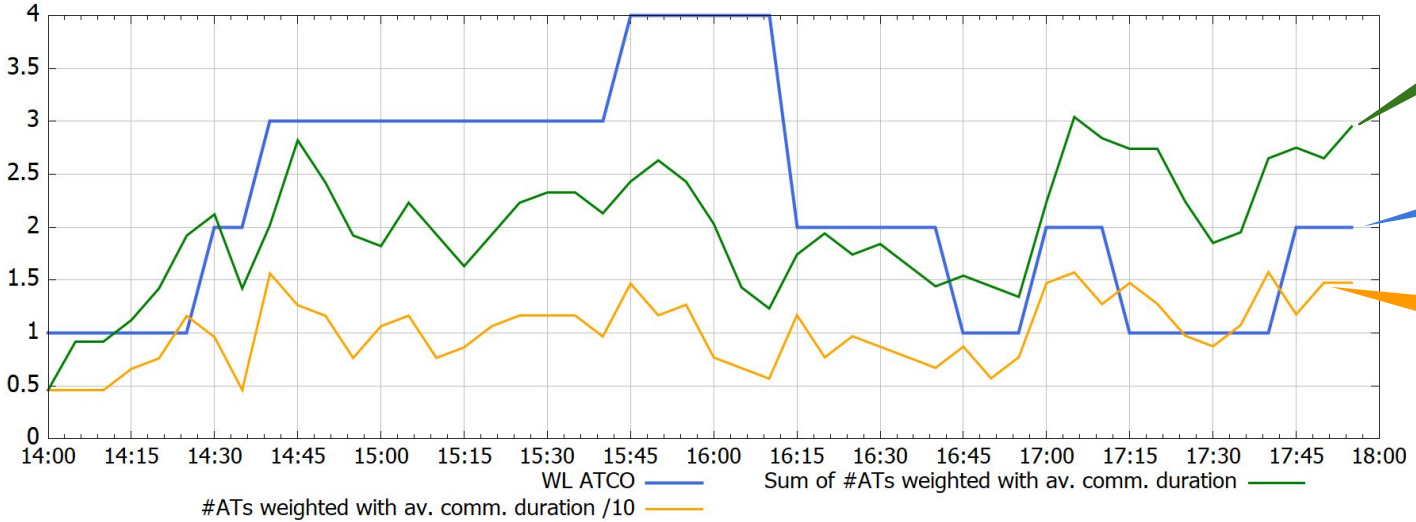
Observer WL

ATCO WL

Increase in WL is accompanied by an increase in at least one of the metrics in the current or previous time period

Bromma Airport: Field Studies

Subjective vs. Objective workload evaluation



Sum # ATs weighted with av. comm. duration

ATCO WL

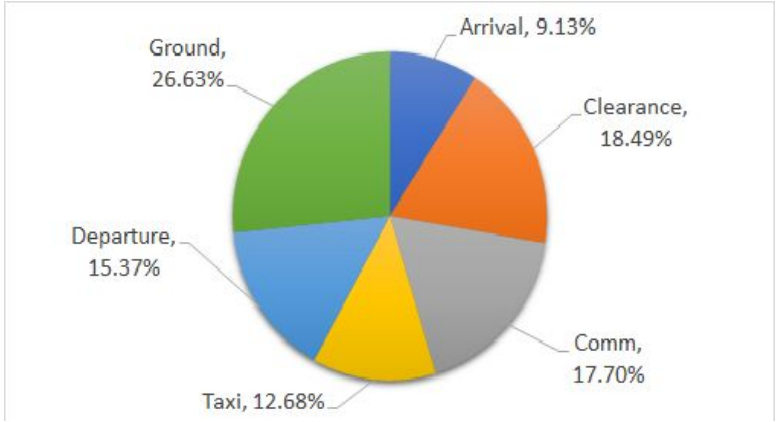
#ATs weighted with av. comm. duration

Sum of av. comm.-weighted #ATs over 2 periods generally replicates the ATCO WL

Necessary, but not sufficient condition colds - can not be used as a WL predictor!

Bromma Airport: Field Studies

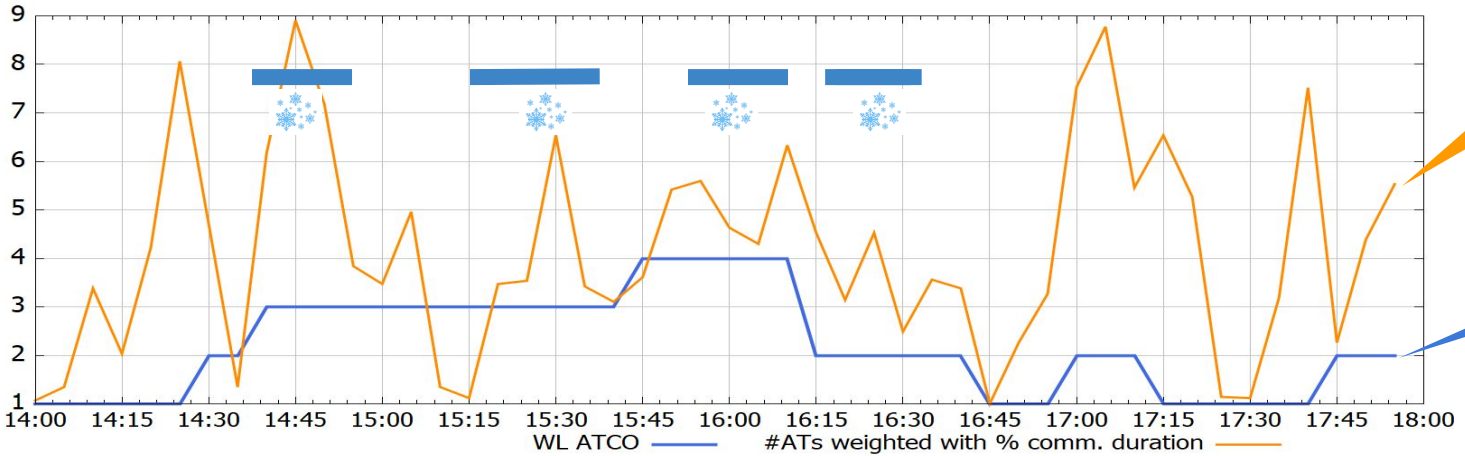
Weather Impact



Snow sweeping coordination is a major part in ground communication

Bromma Airport: Field Studies

Impact of Weather on ATCO WL?



ATCO tasks weighted with %comm. duration

ATCO WL

ATCO WL ratings are higher during the snowy period despite the traffic volume

Conclusions and Outlook

- ✔ Goals: deliver universal flexible automation tools for robust staff rostering
- ✔ Main results: outlined challenges in RTC staff planning
- ✔ Challenges: avoid potential conflicts in schedules and controllers overload and fatigue
- ✔ Future work: better indicators of WL

  identify the WL drivers

  more data in different working conditions

  investigate staff solutions in different weather scenarios

Open for discussions and collaboration

References

T. Andersson, P. Axelsson, J. Petersson, T. Polishchuk, V. Polishchuk, C. Schmidt. Configuration and Planning of the Remote Tower Modules in a Remote Tower Center. In **ICRAT'16**, Philadelphia, USA.

B. Josefsson, T. Polishchuk, V. Polishchuk, C. Schmidt. A Step Towards Remote Tower Center Deployment: Optimizing Staff Scheduling. In **ATM Seminar 2017**, Seattle, USA.

B. Josefsson, T. Polishchuk, V. Polishchuk, C. Schmidt. Scheduling Air Traffic Controllers at the Remote Tower Center. In **DASC 2017**, St. Peterburg, USA.

B. Josefsson, T. Polishchuk, V. Polishchuk, C. Schmidt. Scheduling Air Traffic Controllers at the Remote Tower Center. In Swedish Transportation Research Conference, Stockholm.

J. Dahlberg, T. Polishchuk, V. Polishchuk, C. Schmidt. Stakeholder Cooperaon for Improved Predictability and Lower Cost Remote Services. In **SESAR Innovaon Days (SID 2017)**, Belgrad.

B. Josefsson, J. Jakobi, A. Papenfuss, T. Polishchuk, C. Schmidt, L. Sedov Identification of Complexity Factors for Remote Towers. In **SESAR Innovaon Days (SID 2018)**, December 3-5, Salzburg.

THANK YOU!