Ramp Metering Algorithms and Implementations: A Worldwide Overview

Robert L. Bertini

Available at: https://works.bepress.com/rbertini/22/
Ramp Metering
Algorithms and Implementations
-A worldwide overview-

Tom Peter Kristeleit
Benedikt Bracher
Klaus Bogenberger
Robert L. Bertini
Januar 2016
Foreword

Historically, continuously increasing traffic volumes has been a challenge for many countries and departments of transportation, particularly manifested by traffic jams on freeways. These are often caused by overloaded on-ramps in combination with high flows on the mainline and occasionally backups from off-ramps. As it is mostly not possible to dramatically widen the road network, and additionally this does not address the problem adequately, other solutions had to be found. One of the most common and most effective approaches is ramp metering. This report summarizes the history and the current state of existing ramp metering algorithms, as the last comprehensive summary of ramp metering algorithms was completed in 1999 by Bogenberger and May. It presents the importance of this strategy for extending the use of freeways without physical expansions of the existing system. The detailed explanations about the local implemented systems in several countries provide an overview about the diversity in ramp metering algorithms that are currently. It strengthens the assumption that different strategies work better or worse on varying boundary conditions and thus it is necessary to improve the different strategies independently to keep extensive opportunities for the traffic control centers to react to their unique problems in their area of responsibility. In this report ramp metering algorithms and the strategies for implementing them are examined. Therefore, the literature review provides a general overview of the basics of ramp metering. The general method of ramp metering is explained with its implementations and further approaches. To illustrate the importance of ramp metering around the world, an overview of existing algorithms and a detailed summary of the current number of metered ramps in the world is given. This is followed by an overview of implemented ramp metering strategies all over the world, along with an historic explanation of the development of ramp metering algorithms. The purpose of this report was to develop a foundation for future research on ramp metering; summarizing all of the existing history and methodologies.
Table of contents

LIST OF ABBREVIATIONS ...................................................................................................................... 1
1 INTRODUCTION ......................................................................................................................................... 3
  1.1 Motivation .............................................................................................................................................. 3
  1.2 Structure .............................................................................................................................................. 4
2 LITERATURE REVIEW ............................................................................................................................... 6
  2.1 Goals and Benefits of Ramp Metering ................................................................................................. 6
  2.2 Ramp Metering Implementations ......................................................................................................... 7
    2.2.1 Working Modes ............................................................................................................................... 8
    2.2.2 Level of Detail ................................................................................................................................. 10
    2.2.3 Relationships Between Metered Ramps ...................................................................................... 11
    2.2.4 Activation Time ............................................................................................................................. 12
    2.2.5 Field Deployment .......................................................................................................................... 13
    2.2.6 Further Explanations ..................................................................................................................... 15
  2.3 Conclusion of the Literature Review ...................................................................................................... 18
3 ALGORITHMS .......................................................................................................................................... 20
  3.1 Introduction .......................................................................................................................................... 20
  3.2 Local ramp-metering algorithms ........................................................................................................ 21
    3.2.1 Demand-Capacity .......................................................................................................................... 21
    3.2.2 Occupancy Control ...................................................................................................................... 23
    3.2.3 Zone .............................................................................................................................................. 23
    3.2.4 ALINEA ....................................................................................................................................... 24
    3.2.5 Local metering using neural networks ........................................................................................ 27
    3.2.6 MIXCROS .................................................................................................................................. 27
    3.2.7 PRO ............................................................................................................................................ 29
  3.3 Coordinated ramp-metering algorithms .............................................................................................. 33
  3.4 Cooperative .......................................................................................................................................... 33
5.1 Future Research Recommendations .............................................................. 82

WORKS CITED .................................................................................................. 84
LIST OF FIGURES ............................................................................................. 91
LIST OF TABLES ............................................................................................... 92
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADOT</td>
<td>Arizona Department of Transportation</td>
</tr>
<tr>
<td>AHP</td>
<td>Analytic Hierarchy Process</td>
</tr>
<tr>
<td>ALINEA</td>
<td>Asservissement Linéaire d'entrée Autoroutière</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ARMS</td>
<td>Advanced Real-time Metering System</td>
</tr>
<tr>
<td>Caltrans</td>
<td>California Department of Transportation</td>
</tr>
<tr>
<td>CBA</td>
<td>Cost Benefit Analysis</td>
</tr>
<tr>
<td>CCTV</td>
<td>Closed circuit television</td>
</tr>
<tr>
<td>CES</td>
<td>Cost Effectiveness Studies</td>
</tr>
<tr>
<td>CMS</td>
<td>Changeable Message Sign</td>
</tr>
<tr>
<td>CSM</td>
<td>Congestion Management System</td>
</tr>
<tr>
<td>CUA</td>
<td>Cost Utility Analysis</td>
</tr>
<tr>
<td>DoT</td>
<td>Department of Transportation</td>
</tr>
<tr>
<td>FHWA</td>
<td>Federal Highway Administration</td>
</tr>
<tr>
<td>FSP</td>
<td>Freeway Service Patrol</td>
</tr>
<tr>
<td>HAR</td>
<td>Highway Advisory Radio</td>
</tr>
<tr>
<td>HERO</td>
<td>Heuristic Ramp Metering Coordination</td>
</tr>
<tr>
<td>HOV</td>
<td>High Occupancy Vehicle</td>
</tr>
<tr>
<td>LOC</td>
<td>Level of Congestion</td>
</tr>
<tr>
<td>LOS</td>
<td>Level of Service</td>
</tr>
<tr>
<td>Mbottleneck</td>
<td>Modified Bottleneck</td>
</tr>
<tr>
<td>MILOS</td>
<td>Multi, Objective, Integrated, Large-Scale, Optimized, System</td>
</tr>
<tr>
<td>MnDoT</td>
<td>Minnesota Department of Transportation</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>OD</td>
<td>Origin-Destination</td>
</tr>
<tr>
<td>ODOT</td>
<td>Oregon Department of Transportation</td>
</tr>
<tr>
<td>OOCS</td>
<td>Overlapped Occupancy Control Strategy</td>
</tr>
<tr>
<td>OR</td>
<td>Oregon</td>
</tr>
<tr>
<td>PC-RT</td>
<td>Predictive-Cooperative Real-Time Rate Regulation</td>
</tr>
<tr>
<td>PLC</td>
<td>Programmable Logic Controller</td>
</tr>
<tr>
<td>Portal</td>
<td>Portland Oregon Regional Transportation Archive Listing</td>
</tr>
</tbody>
</table>
1 Introduction

This chapter provides a short overview about the motivation and the structure of this report. The objective is to explain why the topic of ramp metering should be studied in more detail and what the general scope of this research is.

1.1 Motivation

Today, traffic is still a growing subject in the world and it has to be handled on existing streets embedded in mostly fixed urban scenarios. It is not possible to widen the space for traffic in such environments. Nevertheless, there is a demand which has to be satisfied for economic growth and welfare. Since the middle of the 20th century there have been aspirations to increase the capacity of existing roads by controlling the flow on the freeway or highway system. The selected method was to regulate the flow of traffic that joins the mainline through the ramps, first implemented on the Eisenhower Expressway in Chicago in 1963 [1]. The success of this system led to further implementations in other states and countries and to automation of this system, called ramp metering, with traffic signals at the on-ramps, splitting up platoons of merging vehicles and controlling the flow on the mainline. The operational system to calculate such a procedure often consists of seven following individual algorithms [2]:

- Release algorithm
- Arbitration algorithm
- Switch on/-off algorithm
- Ramp metering algorithm
- Queue override algorithm
- Queue management algorithm
- Data filtering algorithm

The main focus of this report concerns the original ramp metering algorithm, which calculates cycle times of traffic signals on the on-ramp, correlating with the number of vehicles joining the mainline. In addition, it is not possible to completely exclude other parts of this system like the queue override mechanism or the switch on/-off algorithm. Information concerning these other parts of this structure shall be taken as additional information to fill some gaps, and cannot depict the completeness of these individual algorithms.

The goal is to provide a current overview and summary of all ramp metering algorithms for traffic engineers to decide future deployments in their area of
responsibility. In addition to completing this review of the algorithms, it is also necessary to explain which factors are necessary for evaluating the ramp metering system and what boundaries may occur. The way of fulfilling this goal through this report is explained in the following chapter about the structure of this work.

1.2 Structure

The deployed ramp metering systems have become more sophisticated and varied in their syntax. These algorithms shall be summarized in this report in order to provide a current overview about the situation of ramp metering in the world. The task is to explain how they work and where they are deployed. The structure of this report is a linear path from the general introduction to the more detailed concrete examples. It starts in the first chapter with a general explanation of ramp metering and how it is defined.

In the second chapter, the basic knowledge is given for a unified utilization of terminology and to clarify the several definitions and variations of ramp metering adaptations. It follows the direction to explain the general goals and operation methods of ramp metering in general and concludes with a brief overview of special mathematical methods that are often implemented into ramp metering algorithms.

The third chapter is the overview of all algorithms, which were known and explained to date. The list contains over 25 strategies to control the traffic on freeway entrance ramps to maintain free flow on the mainline. Most of them are deployed or were deployed in the past, but this list also describes some algorithms that were tested by simulations but never deployed in the field. The general structure is divided into sub-chapters about their area of influence, either if it is a local or coordinated algorithm with their particular special cases. The described algorithms in these sub-chapters are sorted by the year of their publication. After that, the brief history of ramp metering is explained to clearly show the development of these strategies. The importance of these algorithms is presented in a world-wide summary of their deployment with a short comparison of the development in the past, to present the growing demand and the trend to use such systems on the freeway. This overview includes the number of ramps metered in each country, and which algorithm is used for this task. A short review about the local characteristics and possible plans to develop the existing system is given in a
more detailed explanation. Since the USA is the leading country in ramp metering and the several states treat this topic in different ways, a more detailed explanation about the use of ramp metering in the USA is given after the world-wide view.

The fifth and following chapters are a summary of the investigations made in this report and an overview of the cited works.
2 Literature Review

This chapter serves the reader as a short overview about the basics of ramp metering and is a reflection of the current state of technology in this field. The actual value of ramp metering and the abilities of this strategy for improving freeway traffic flow will be described. These general explanations are the foundation for the following chapters, which assume a common understanding of this field.

2.1 Goals and Benefits of Ramp Metering

Ramp metering is an often-used system to avoid or respond to congestion on freeways around the world. There are many different strategies to achieve this goal with a variety of equipment, techniques and philosophies to handle the external circumstances. All have in common the objective of freely flowing traffic on the mainline. There are the following subordinate strategies in different intensity in use to guarantee this main goal [3]-(p.1), [4]-(p.4):

- Break up platoons of entering vehicles to facilitate merging onto the freeway
- Avoid spillback from the ramp onto arterial roads
- Find equilibrium between on-ramp delay and mainline flow

These goals lead to benefits according to ramp metering. The impacts are far-ranging, depending on the used algorithm, the traffic demand, local peculiarities and other various properties of the used technique just as well the environment of the metered ramp. Following list concludes studied advantages of ramp metering as average benefits: [1]-(p.10f), [5]-(p.11), [6]-(p.4f)

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainline Speed</td>
<td>20%</td>
</tr>
<tr>
<td>Freeway Capacity</td>
<td>4%</td>
</tr>
<tr>
<td>Travel Time</td>
<td>20%</td>
</tr>
<tr>
<td>Crashes</td>
<td>30%</td>
</tr>
</tbody>
</table>

Table 1 Benefits of Ramp Metering [7]-(p.3f)

Also, there are studies that show decreased ecological damage, because of less fuel consumption and emissions.
Comparing several algorithms, there is no consistent result that could lead to a preferred algorithm, but every study led to improved mainline flow in comparison to a system without ramp metering. These results rarely depend on the evaluated parameters, because like in other sciences, e.g. the law of conservation of energy or the second Maxwell theorem [8], the major delay is often moved from the mainline to the on-ramp. The high diversity of algorithms, which try to handle this situation in an appropriate way, could be explained by a varying performance of these. This depends on many environmental boundaries, e.g. the freeway geometry, general driver's behavior, the quality of the merging area or the available space for queueing.

2.2 Ramp Metering Implementations

Depending on the local situation of traffic demand, environmental boundaries, historical situations or political decisions, ramp metering systems can be modified by adding or removing solely abilities. The chosen features of a ramp metering system depend at first on the ability of the basic algorithms to implement an option, which shall be explained in a later chapter. After that, the features explained here can be evaluated by their weight of fulfilling a pre-defined strategy to satisfy the traffic demand, which contains that some options exclude or improve the value of other. The following explanations are an overview about the state of technology and serve the reader to clarify and unify the used terminology. Figure 1 is an opportunity to understand the whole system and after that, the itemized components are explained in their working method and how they are used in the field.
2.2.1 Working Modes
An algorithm has for the most part three options to act under the influence of the given data. It could be a static mode, which depends on historic data with no consideration of the actual traffic or it could be a real-time mode that reacts directly to its environment by gathering and calculating the latest traffic information. This last method divides into an option to just respond to this collected data, and into a forecast model with computed response to solve bottlenecks that might arise in the future. Especially the pre-timed strategy is going to be superseded by the real-time traffic responsive algorithms, because its invention was based on the former state of technology. At the moment there is no preference on a predictive or a reactive model because both have their advantages and disadvantages concerning the capacity for real time calculations and the boundaries on the quality of measured data. Following, these modes are named and explained in more detail.
2.2.1.1 Pre-Timed

{or fixed-time, time-of-day, preset operation} The basis of any calculations of metering rates is historical data, which is accumulated over time, and merged into a common, recurrent pattern for driver behavior. According to that, the algorithm meters the ramp under the assumption that the current traffic is the average of its data. This technique is only useful for recurrent congestion (e.g. commuter peaks AM and PM on weekdays), but even then, there is no reliability that the metering rate or the times that the system turns on and off fits the actual traffic demand. [5]-(p.30), [9]-(p.1)

2.2.1.2 Adaptive

{or traffic-responsive} Adaptive is a method to react to the actual traffic on the mainline and ramps. The data has to be provided in real-time, which determines constantly the metering rate and start/stop times. Depending on the algorithm there is a need for high resolution data, e.g. speed, flow, density or occupancy to calculate the metering rate. [5]-(30), [9]-(p.6)

2.2.1.3 Predictive

With a predictive algorithm the existing real-time data are analyzed and result in a short-term forecast of the evolution of traffic conditions. With that information the algorithm meters the ramps to keep the density or capacity under its saturation level over the entire period. Although it will not react to real-time data, it will react to predicted data that will be probably provided sometime later. [5]-(p.30) Table 2 provides an overview about the described advantages and disadvantages for better clarity in a direct comparison.
### Table 2 Summary of the Working Modes

<table>
<thead>
<tr>
<th></th>
<th>Pre-Timed</th>
<th>Adaptive</th>
<th>Predictive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>- Easy to implement</td>
<td>- Responds to current situation</td>
<td>- More than responding</td>
</tr>
<tr>
<td></td>
<td>- Good enough for recurrent congestion (commuter peaks)</td>
<td></td>
<td>- Ability to avoid congestion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>- No reliance on current metering rate</td>
<td>- High data input</td>
<td>- Real-time calculation</td>
</tr>
<tr>
<td></td>
<td>- No flexibility under changed circumstances</td>
<td>- Possibility to respond to on-ramp queues</td>
<td>- High data input</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Possibility to work with on-ramp queues</td>
</tr>
</tbody>
</table>

#### 2.2.2 Level of Detail

There are algorithms on a different level and with different levels of detail. It is possible to calculate independent metering rates based on a single ramp, but there are also options to take a broader view and compare computed local metering rates with other ramps in a corridor or networks to find a global equilibrium. This larger system can be added with further information or possibilities to react to the current situation. The implementation of ramp metering systems started obviously with local algorithms and developed to the more sophisticated ones. But there are still a lot of local ramp metering algorithms in use due to easier implementation or because they are a part of area-wide algorithms. It is not given that a more sophisticated strategy works better, because there are a lot of possibilities of mistuning them or maybe there are conditions that do not need a more complex approach. The following levels of detail represent the latest stages that were computed for ramp metering algorithms.

#### 2.2.2.1 Local

(or isolated) The local approach describes an algorithm that calculates the metering rates detached from its environment. It is dependent on the data that
directly belong to its ramp only, and establishes no connection to previous or following ramps. [4]-(p.6), [9]-(p.5), [10]-(p.14), [11]-(p.2)

2.2.2.2 Coordinated

(or system wide) Combines several ramps as a unit to consider what impact a metering rate to one location has on another. This potentially superior system determines the metering rates at local positions to have a wider impact. This could be useful at sections with a bottleneck at the end, because it increases the storage capacity of this section since no longer only the first ramp has to react to this bottleneck with its limited capacity to store vehicles. [4]-(p.6), [9]-(p.5), [10]-(p.14), [11]-(p.2)

Table 3 provides an overview about the described advantages and disadvantages for better clarity in a direct comparison.

<table>
<thead>
<tr>
<th></th>
<th>Local</th>
<th>Coordinated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>- Easy to implement</td>
<td>- Uses more on-ramp storage capacity (section-wide allocation)</td>
</tr>
<tr>
<td></td>
<td>- Often “good enough”</td>
<td>- Solves wider problems (bottlenecks)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Consistency in metering tactic to handle one section</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>- Cannot solve wider problems</td>
<td>- Higher data flow</td>
</tr>
<tr>
<td></td>
<td>- Last ramp before bottleneck gets handicapped</td>
<td>- Sometimes last ramp before bottleneck still handicapped (too reactive)</td>
</tr>
</tbody>
</table>

2.2.3 Relationships Between Metered Ramps

Apart from local ramp metering algorithms, where the ramps do not communicate with each other, there are three possibilities for coordinated algorithms to connect the computed metering cycles for an improved, global solution. Depending on the superior system, each tactic has its advantages and disadvantages and is mostly preset by the used algorithm.
2.2.3.1 Cooperative

A cooperative algorithm is described by calculated metering rates that extend each other with data or responsive possibilities when one of the ramps reaches its limits. [4]-(p.6), [9]-(p.5)

2.2.3.2 Competitive

The competitive strategy calculates at first several metering rates independent of each other at area of response to choose after that the most restrictive calculated metering rate out of all solutions that finally gets deployed. [4]-(p.6), [9]-(p.5)

2.2.3.3 Integrated

The integrated relationship between different levels is characterized by separate computed metering rates that are compared to find the best working consensus of these strategies. [4]-(p.6), [9]-(p.5)

2.2.4 Activation Time

It was proven that ramp metering only has a visible impact on the mainline flow, when there is an increased traffic demand. Nevertheless there are aspirations to keep the ramp metering system ready for incidents and more flexibility. The following list shows the latest common possibilities for using a ramp metering system. In general it is a question of costs over time compared to the use of the gained flexibility. [12]

2.2.4.1 Fixed Time

The ramp metering system gets activated based on a fixed schedule. The switch-on and –off times are depending on historical data about commuting behavior or recurrent, well-known incidents. [13]-(p.101)

2.2.4.2 Dynamic

Dynamic activation describes a strategy where the actual traffic conditions determines that the ramp metering system will be switched-on after the mainline reaches a certain threshold, which is sensitively calibrated to get a chance to avoid congestion and not only to respond to it. [13]-(p.101)
2.2.4.3 Incident

The incident activation is characterized by a certain incident on the mainline. The ramp metering system can be switched-on to support the evacuation of the mainline or to support a temporary changed situation, e.g. during roadwork, crashes or lane closures. [13]-(p.101)

2.2.4.4 Manual

The manual activation strategy is an option to respond to exceptional circumstances. It can be the opportunity to over-ride the algorithm by an operator, e.g. to clear a ramp before an ambulance arrives. [13]-(p.101)

2.2.5 Field Deployment

In general, it is an uncommon possibility to implement a ramp metering system on a newly constructed freeway that is designed for the needs of ramp metering. Thus, the conditions of the profile and traffic demand are given as a task to solve. Options to handle the initial situation are rare and depend mostly on the existing space around the ramp. The following models are opportunities to respond to the circumstances that have an impact on the metered ramp.

2.2.5.1 Single lane, one vehicle per green

The typical option to meter an on-ramp is to let pass one vehicle per cycle to the mainline. This version supports all goals of ramp metering as far as the demand from the arterial road is not too high. [1]-(p.6)

2.2.5.2 Single lane, multiple vehicles per green

{or bulk metering, platoon metering} It describes a strategy to handle a higher demand. There is also the possibility to let pass more than one vehicle per cycle on the mainline. It is not possible to double the flow rate of the version with one vehicle per green because of longer green and red phases to avoid uncertainty of the drivers. [14]-(p.7) This option is in contrast to the goal of ramp metering to split arriving platoons into single units. It is not recommended to let more than 3 vehicle pass the ramp at once. [1]-(p.6)
2.2.5.3 Dual lane metering

If there is enough space, it is a good method to increase the storage capacity of a ramp by adding a second lane to it. It is also useful for mainlines with high occupancy vehicle (HOV) lanes or public transport through the mainline. The green cycle should not be simultaneous to avoid further conflicts. [1]-(p.6)

Table 4 provides an overview about the described advantages and disadvantages for a better clarity in a direct comparison.

<table>
<thead>
<tr>
<th></th>
<th>Single lane, one vehicle</th>
<th>Single lane, multiple vehicles</th>
<th>Dual lane</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>- Supports all goals of ramp metering</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Easy to calculate</td>
<td>- Higher traffic flow to clear the ramp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Easy to understand/use for drivers</td>
<td>- More responsive to high traffic demand and arterial spillback</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Advantage for HOV-lanes and public transport on freeway</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Higher traffic flow</td>
<td>- Higher traffic flow</td>
<td>- Higher ramp storage capacity</td>
</tr>
<tr>
<td></td>
<td>- Higher ramp storage capacity</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>- Sometimes too restricted for traffic demand</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- No double output in comparison to one vehicle</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- “Dilemma zone” how many vehicles are allowed to pass</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Additional traffic sign to explain multiple vehicles</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- In contrast to avoid platoons</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Needs more space</td>
<td>- Increased computing to avoid simultaneous green cycles</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Spacious ramps</td>
<td></td>
</tr>
<tr>
<td><strong>Possible Deployment</strong></td>
<td>- First type to use, for most of the ramps</td>
<td>- Freeway to freeway metering</td>
<td>- High traffic demand and additional space is available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- HOV-lanes or public transport on freeway</td>
</tr>
</tbody>
</table>
2.2.6 Further Explanations

2.2.6.1 Kalman Filter

The Kalman filter is a typical online mathematical tool to filter the noise of empirical measurements and estimates statistically the real system in a more appropriate way, which is described as “the calibration problem: least squared errors and maximum likelihood.” [15]-(p.304) It is based on the assumption that the failures are Gaussian-distributed. [16]

2.2.6.2 Rolling Horizon

Many predictive ramp metering algorithms use a rolling horizon to improve the quality of their prediction. The idea is to smooth the basis for the forecasting model by forming the average of a pre-defined amount of previous measured data and take this as the basis for the following calculations. This method avoids jumps in the prediction that occur regularly by changing traffic situations. [17]

2.2.6.3 Godunov Scheme

This method is a numerical tool to solve partial differential equations with a finite-volume approach and was developed by Godunov in 1959. The conservative variables are piecewise constant over the cells at each time step and the time evolution is determined by solving the Riemann problem. [18]

2.2.6.4 Kinematic Wave Model

The kinematic wave model was first mentioned by Lighthill and Whitham in 1955. They assume that on a long, crowded roadway the vehicles behave like the flow of water. The foundation is the relationship between flow and concentration, which is today better known as density. The assumption is that a change in flow causes a “kinematic wave” upstream with the difference that the height of the wave because of the incompressibility of water is represented by a higher density of the mainline flow. This model only works for a large number of vehicles because it is based on simple continuum models. [19]
2.2.6.5 Feedforward- and Feedback-Control

**Feedforward-Control**

Feedforward-Control, also known as open loop control or non-feedback control, is a simplified system used to regulate a certain behavior in the current state and the predicted result. In cases of ramp metering, an example would be the measured traffic demand on the mainline and on the on-ramp, which results in a pre-defined cycle time for such a situation. At the next measurement cycle, the system only works with the current traffic situation of the next step what is shown as a process in Figure 2. Such a behavior can easily lead to an oscillating condition because the calculation is not smoothened by historic data or an evaluation of what caused its decision. [20]-(p.24f)

![Figure 2 Feedforward Control](image)

**Feedback-Control**

Feedback-Control, also known as closed loop control, works with an evaluation component, which influences the decision of the next step by integrating the results of the previous decision into the calculation. An example for ramp metering would be an upstream measured mainline demand and an on-ramp demand, which results into a cycle time. The influenced situation is measured downstream and this data is included into the next cycle time calculation that is summarized in Figure 3. This behavior probably slows the process down because the algorithm responds more smoothened on the future demand, but in general it is more robust and reliable. [20]-(p.12f)

![Figure 3 Feedback Control](image)

The advantages and drawbacks are summarized for a better overview in Table 5.
Table 5 Comparison Feedforward-Feedback Control

<table>
<thead>
<tr>
<th></th>
<th>Feedforward Control</th>
<th>Feedback Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>- Simple system</td>
<td>- More robust</td>
</tr>
<tr>
<td></td>
<td>- Active control</td>
<td>- Smoothened behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>because of influenced input</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>- Unstable</td>
<td>- Responsive</td>
</tr>
<tr>
<td></td>
<td>- Inconsistent</td>
<td>- Can improve undesirable</td>
</tr>
<tr>
<td></td>
<td>- No response to changed system</td>
<td>properties of the system</td>
</tr>
</tbody>
</table>

**Feedforward-back Controller**

In general, both controllers are often combined to balance the advantages and drawbacks of the single systems, which leads to a fast response because of the feedforward controller and to a more robust result because of the feedback component of the closed loop controller, as shown in Figure 4. [20]-(p.27) In the field of ramp metering, the use of this approach is new and is only implemented in one algorithm, see 3.6.12.

![](image)

Figure 4 Feedforward-Feedback Controller [20]-(p.27)

2.2.6.6 Origin-Destination Approach

To analyze traffic patterns, some ramp metering algorithms base their calculations on assumptions about which specific paths the macroscopic traffic flow follows. This is called Origin-Destination (OD) estimation and can be realized through different approaches that estimate the route choices of single vehicles. Examples to determine this basic data set are traffic surveys which are expensive in time and costs, growth factor models, synthetic techniques or spatial interaction models,
e.g. the connection of gravity centers through routes which are defined by time and cost functions. [21]- (p.14)

The analyzed relationships of the origins and destinations that influence the metered corridor are gathered in an OD matrix and serve the algorithm for traffic pattern assumptions.

2.2.6.7 Evaluation of Individual Algorithms

Distributed over this report, there will be some evaluations and alignments. In the beginning it is to say that ramp metering systems with their various algorithms are not as consistent as desirable because they depend on too many influences in different combinations which distort the unambiguity. The alignments are approximate conclusions of many previous field tests and studies of other researchers and are a summarized component to estimate the value and possible deployments of several ramp metering algorithms and their components. To have a concrete view on a specific algorithm, the attached literature includes tests of nearly all described algorithms, in which the performance, the advantages and disadvantages of the several strategies become obvious.

2.3 Conclusion of the Literature Review

This chapter gives a short overview of the basic possibilities of ramp metering algorithms. It shows the several opportunities to implement features to the sole algorithms, to extend their flexibility and improve their behavior according to the local circumstances. There is the possibility to respond to existing boundaries, e.g. the storage capacity of each ramp or the merging quality of the area in which the on-ramp joins the mainline. The response to the described parameters could be an implementation of a dual lane system or a multiple vehicle per green strategy. Also, there is the possibility to decide how much effort is needed to achieve a certain level of service (LOS). The resulting decision would be a choice of a pre-defined operating schedule with a static activation time, depending on the time of day. Also, it could be an adaptive algorithm which activates automatically when the traffic demand on the mainline starts to increase. The last option would conclude in a continuously monitoring process with a lot of examined data and a constant measuring process. This needs a lot of resources, e.g. manual operator or power, to run appropriately.
This diversity and flexibility of ramp metering strategies is one component of their success, but it is also a hindrance for a best field-implementation, because all these parameters should be recognized in the pre-evaluation process. Possibilities of implementations in reality are explained in the next chapter about the specific algorithms.
3 Algorithms

3.1 Introduction

The number of algorithms grows with the innovations in the field of data collecting and processing. There are several algorithms that reacted on gathered experiences and technical improvements. Some are still in use and some are replaced by sequels or better working algorithms. To get an overview about the latest situation, this chapter summarizes and explains the algorithms that were developed and deployed over time. They are sorted by their characteristic grade of detail either local or, which is also an imprecise order, for their time of creation since the development of these algorithms leads to more complex and system wide strategies. For the beginning and a better overview Figure 5 illustrates chapter 3 in its actual order. The algorithms which are highlighted in dark letters are not mentioned in the compendium of Bogenberger and May.

![Figure 5 Algorithms Overview and Structure](image-url)
3.2 Local ramp-metering algorithms

Ramp metering on the local level is an essential component of all following ramp metering strategies. Former algorithms back to the beginning of ramp metering controlled the ramp solely at the local level. Even though at first sight, more coordinated ramp metering algorithms are used and seem to be more efficient because of their sophisticated structure, there is in every case a local algorithm included solving the local level on-ramp and mainline demand. This is shown in many variations of these local algorithms, e.g. ALINEA, which is transferred in many coordinated systems as a robust and reliable ramp metering algorithm at the local level. [22]-[p.19], [23]

3.2.1 Demand-Capacity

This adaptive algorithm simply computes the gap between measured upstream occupancy and downstream critical capacity that leads to a number of vehicles that are allowed to enter the mainline. The Demand-Capacity (DC) method is a fundamental basis for many other algorithms because of the plausible and transparent idea. This foundation can be flexibly extended, e.g. by setting a minimum metering rate to avoid a permanent red signal at the on-ramp if the critical occupancy is exceeded [10]-[p.16]. The syntax including the described example might be like:

\[
\begin{align*}
\{IF\} \ O_{in} & \leq O_{crit} \ {THEN} \\
 r(t) & = \max(q_{cap} - q_{in}; r_{min}) \\
\{ELSE\} r(t) & = r_{min}
\end{align*}
\]

- \(O_{in}\) = measured mainline occupancy upstream
- \(O_{crit}\) = pre-defined critical mainline capacity downstream
- \(r(t)\) = computed metering rate
- \(q_{cap}\) = pre-defined critical mainline flow
- \(q_{in}\) = measured mainline flow upstream
- \(r_{min}\) = pre-defined minimum metering rate

[24]-[p.7]

3.2.1.1 TANA

The Tangenziale de Napoli (TANA) is a comprehensive road management system, which includes VMS, CCTV and ramp metering and was established in Napoli in
The several parts of the system are not connected with each other automatically, but through a manual operator.

The core algorithm itself seems to be the first installed ramp metering algorithm in the world and is based on the DC theory. This strategy measures the demand and compares it with six pre-defined classes. The lowest class, identified with a zero, induces a permanent green phase, with the peculiarity that the traffic signal is set to red but instantly changes to green when a vehicle arrives in the front of the traffic signal. This setting avoids uncertainties for arriving vehicles, which do not know when this traffic signal could jump to a red phase instantly. This algorithm also contains a queue override detection, which sets the latest cycle class one step lower, when the queue detector is occupied, to slowly clear the ramp. Every part of the TANA strategy can be overridden by the manual operator, who might react more quickly to incidents observed through the CCTV or made experiences.

3.2.1.2 RWS (Rijkswaterstraat)

RWS, named after “the commissioning organization” in the Netherlands is a variation of the DC method and was developed in the Netherlands. The algorithm measures flow and uses speed instead of occupancy. The target is to keep the actual flow below the critical flow at the inflection point as visible in Figure 6.

![Figure 6 Visualization of Traffic-Inflection Point](image)
The metering rate results from a comparison of the pre-defined capacity and the measured flow of the previous time step.

\[ r_t = C - I_{t-1} \]

- \( r_t \) = computed metering rate
- \( C \) = pre-defined capacity
- \( I_{t-1} \) = measured mainline upstream flow of previous time step

3.2.2 Occupancy Control

The occupancy control (OC) strategy is a basic approach for a local ramp metering algorithm and is the standard isolated part of the Bottleneck algorithm 3.5.2. OC is an adaptive system that compares the actual computed solutions, based on measured occupancy data of the mainline by upstream detector loops, with predetermined metering rates corresponding to these occupancy values. The estimated metering rates are created based on the assumption of a relationship between the measured occupancy and the actual flow on the freeway. The calculations are feed-forward based and thereby not as robust as feedback strategies as ALINEA 3.2.4, because it does not include previous calculations in its equations. Because of that, the metering rates can be sometimes inconsistent.

In some simulations the OC method was compared directly with the ALINEA algorithm, e.g. as the local part of the Bottleneck strategy, and it was shown that ALINEA is an improvement every time. This is largely why this algorithm is not used as a standard strategy for ramp metering anymore.

3.2.3 Zone

The Zone algorithm was first deployed in Minneapolis, Minnesota, USA as a pre-timed algorithm, but was updated later to an adaptive system. The region of the Twin Cities is well-known for its aspirations in researching ramp metering including the definitive, comprehensive evaluation of their system as part of an 8-week shut-off experiment in fall 2000.

As the name provides, the Zone algorithm divides the mainline into zones, which were characterized by several metered and some unmetered on- and off-ramps, an area with low incidents upstream and a potential bottleneck downstream.
The idea is to control the functionality of the zone by measuring the incoming and leaving traffic over the whole section with the downstream bottleneck capacity as the regulating factor [11]-[p.3]. This equation is described by:

\[ A + U + M + F = X + B + S \]

- **A** = upstream mainline volume (measured)
- **U** = sum of unmetered entrance ramp volumes (measured)
- **M** = sum of metered ramp volumes (controlled)
- **F** = sum of metered freeway to freeway ramp volumes (controlled)
- **X** = sum of exit ramp volumes (measured)
- **B** = downstream bottleneck capacity
- **S** = space available within the zone (computed volume based on measured variables)

For a more obvious presentation of the possibilities of the algorithm it is necessary to transform the equation which is dependent on the metered ramp volumes. With the assumption that \( S \) is equal to zero this leads to:

\[ M + F = (X + B) - (A + U) \]

[14]-[p.15]

### 3.2.4 ALINEA

The Asservissement Linéaire d’Entrée Autoroutière (ALINEA) algorithm developed by Papageorgiou is a robust and often used local ramp metering algorithm. It was first proposed in 1997 and has been deployed in Paris, Amsterdam, and Munich and is often an improvement on the local area of coordinated ramp metering algorithms, e.g. in test situations for Bottleneck, SWARM or as the standard local algorithm of HERO. [22]-[p.19], [23]

The idea is to keep the traffic under a static, pre-evaluated occupancy and increase thereby the mainline throughput. It is based on the feedback theory and computes the metering rates in a way that the traffic flow remains below the system capacity. The metering rate is calculated with the equation below:

\[ r(t) = r(t-1) + K_R (O_C - O_{out}(t)) \]

- \( r(t) \) = metering rate (volume) in time interval \( t \)
- \( K_R \) = weighting factor >0 to adjust the feedback control
- \( O_C \) = local occupancy at ramp (measured by one mainline detector)
- \( O_{out} \) = pre-defined occupancy value at capacity (varies, depends on local philosophy)
This equation is a closed-loop equation, thus \( r(t) \) is a function of \( r(t-1) \), an advantage in contrast to open-loop algorithms because of a certain consistency in metering rates and a robust behavior in calculations. [3]-(p.39f), [4]-(p.26), [11]-(p.28)

There are many regional variations, improvements and extensions to this algorithm and some of them shall be explained here. It is to summarize that every modification of the original ALINEA led to a worse result in comparison to the original algorithm, except for ALINEA/Q. The deployment of these changed algorithms is explained by the attempt to avoid construction work on the freeway to implement new detector systems by using the existing material. [10]-(30)

3.2.4.1 ALINEA/Q

The additional Q is an additional feature of ALINEA to count the number of vehicles that are stored in a queue or entering the ramp via video monitoring. This extension allows ALINEA to take reference to the queue, which was not possible up to this point and one of the major drawbacks because the mainline flow was guaranteed at the expanse of the arterial network. This extension can be implemented to every ALINEA algorithm including the existing modifications. [10]-(30)

3.2.4.2 MALINEA

MALINEA is an improvement made in 2001, which solves two disadvantages of the original ALINEA. The first the problem is that congestion could arise upstream (and ALINEA just computes metering rates based on the downstream occupancy), the second problem is the localization of the optimal location for the measuring detectors. Both of these subjects are handled in the following revised equation:

\[
Q_r(t+1) = O_u(t + n + 1) - O_u(t) \frac{K}{A} + Q_r(t-n)
\]

\( Q_r(t+1) \) = calculated metering rate for next time step
\( O_u \) = measured upstream occupancy
\( n \) = pre-defined time lag between up- and downstream measurements
\( K \) = regulating factor
\( A \) = derivation of downstream and upstream occupancies relationship

[10]-(p.18)
3.2.4.3 FL-ALINEA

This adjusted ALINEA from 2003 replaced the downstream occupancy input data with the downstream flow measurements to overcome a lack of data. The original ALINEA equation is thus modified in a way that occupancy input is replaced by flow data. A little extension is the connection to a minimum metering rate, which becomes deployed when the freeway capacity is exceeded. [10]-p.19)

3.2.4.4 UP-ALINEA

As the name suggests the input occupancy data is measured upstream of the entrance ramp and is forecasted to a downstream occupancy. This additional step of calculation should only be implemented if there are already upstream detectors, e.g. from former algorithms which were using upstream data like demand-capacity.

The predicted downstream occupancy is calculated by the following equation: 

$$\hat{O}_{out}(k) = O_{in}(k) \ast \left(1 + \frac{q_{ramp}(k)}{q_{in}(k)}\right) \ast \frac{\lambda_{in}}{\lambda_{out}}$$

- $\hat{O}_{out}(k)$ = estimated occupancy downstream at time step k
- $O_{in}(k)$ = measured occupancy upstream at time step k
- $q_{ramp}(k)$ = measured flow from the on-ramp at time step k
- $q_{in}(k)$ = measured flow on the mainline at time step k
- $\lambda_{in}$ = number of mainline lanes upstream
- $\lambda_{out}$ = number of mainline lanes downstream

This calculated downstream occupancy replaces the measured occupancy of the normal ALINEA. The rest of the original equation does not change. [10]-p.19)

3.2.4.5 UF-ALINEA

This modification of ALINEA is an extent to the FL-ALINEA. It uses the same equation as described above but with the difference that the inserted downstream flow is not measured but calculated. It is a simple summation of the measured upstream flow and the on-ramp flow. This step extends the applicability of ALINEA without any modifications on the roadside. [10]-p.20)
3.2.5  Local metering using neural networks

This algorithm is comparable to ALINEA with the difference that it uses Artificial Neural Networks (ANN) to work with input data. This method assigns a weight to the input data before it solves the function to calculate metering rates as output. The traffic is assumed as a kinematic wave that is a good description for the traffic flow. This underlying hydrodynamic model is similar to ALINEA as well as the use of feedback regulation to keep the actual conditions on the freeway below the critical occupancy. The following scheme, Figure 7, shows the sequence of events for the feedback control of the algorithm.

![Non-Linear Feedback Control Process](image)

If there is any congestion, problems can occur because of the lack of a queue override strategy. In this state the computed metering rate would be restrictively overridden, which leads to alternating behavior during congestion. [11]-(p.31)

All in all, this algorithm is as well as ALINEA still ranked as good. [4]-(p.8)

3.2.6 MIXCROS

MIXCROS is originally a local feedback-based algorithm of 2003 that explicitly treats ramp queues in its metering rate calculation. Because of the reduced capability to influence the mainline flow on the local level, there were already expansions to transform MIXCROS into a coordinated algorithm. [28]-(p.73)

The local version of MIXCROS is the basis for the adaptations in their syntax and assumptions. The idea is maintaining the traffic density on the mainline to maximize its throughput, while reducing queues at the on-ramp. The importance of these two influenced areas is weighted by factors for each ramp and is reflected in the following error-equation [28]-(p.73):

\[
e(k) = w_1 \cdot |\rho(k) - \rho_{cr}| + w_2 \cdot queue_{ramp}
\]

\(e(k)\) = error-function to combine the situation on the ramp and the mainline
\(w_1\) = weighting factor to determine importance of the mainline
ρ(k) = measured density of mainline section k
ρ_{cr} = critical density at the section
w_2 = weighting factor to determine importance of the ramp
time_{queue} = actual queue at the on-ramp

The calculated e(k)-factor is finally included in the calculation of how many vehicles shall be allowed to enter the mainline from an on-ramp. This procedure can be described by:

\[ u(k) = \frac{-F(k) - K * e(k)}{G} \]

with

\[ F(k) = \text{sign}(\rho(k) - \rho_{cr}) * w_1 * \left[ \rho(k) - \rho_{cr} + \frac{T}{L_f} * (f_1(k) - q_{out}(k)) \right] \]

\[ G = \text{sign}(\rho(k) - \rho_{cr}) * w_1 * \frac{T}{L_f} - w_2 * \frac{T}{L_r} \]

\[ \text{sign}(\rho(k) - \rho_{cr}) = \begin{cases} 1, & \rho(k) > \rho_{cr} \\ -1, & \rho(k) \leq \rho_{cr} \end{cases} \]

u(k) = computed metering rate in \[ \frac{\text{veh}}{h} \]
k = actual time step
K = defined control gain
T = time step duration
L_f = length of the freeway section
f_1 = measured incoming flow on the mainline
q_{out} = measured flow leaving the section
L_r = length of the entrance ramp

This set of equations is the basis for the coordinated MIXCROS algorithms that exist in a coupled and in a decoupled version [28]-(p.80).
3.2.6.1 D-MIXCROS

The decoupled variant is quite similar to the local MIXCROS algorithm, with the difference of the control gain factor $K$ being constant over a whole section to produce a certain harmony in the calculations of the several ramps. The equation for $n$ ramps in a section would be described by [28]-(p.82):

$$u_n(k) = \frac{-F(k)n - K \cdot e_n(k)}{G_n}$$

As the calculation is identical with the local version of MIXCROS, the advantage of the decoupled variant is that its implementation is easier because of the constant control gain factor.

3.2.6.2 C-MIXCROS

This expansion of the MIXCROS algorithm is the coupled version that combines the weights of several ramps in a section on the foundation of the basic equation of the original algorithm. The calculation is described by the following equation:

$$u_n(k) = \frac{\alpha_n \cdot (-F(k) - K \cdot e(k))}{G_n}$$

The main extension is the weighting factor $\alpha$ which is a certain percent weight for each ramp and must be accumulated equal to 1. With that variable it is possible to respond to different conditions on the several ramps by distributing the burden to other ramps in real-time [28]-(p.80f).

3.2.7 PRO

The Proaktive Rampenoptimierung (PRO) is an algorithm that was first described in 2006 in Germany and has been implemented into the ramp metering system of Baden-Wuerttemberg since 2012. [29]-(p.118) It is a local and predictive strategy which equally tries to handle the mainline flow and the on-ramp demand.

The origin algorithm consists of the unique ramp metering algorithm itself and additional components that calibrate the equation for the use in the field. [30]-(p.81) At first, the core of this strategy shall be explained by the equation and its assumptions:

$$q_{R2S} = q_{Nachfrage \ mak} \cdot \left(\frac{q_{HFB \ mak}}{q_{HFB \ mik}} * f_q\right)$$
The weighting factor $\lambda$ is a possibility to influence the sensitivity of the algorithm. It could be possible to force a higher sensitivity to the current situation on the mainline by setting $\lambda$ to a value greater than 1. This would lead to higher fluctuations of the metering rate. If $\lambda$ is set smaller than 1, the metering rates become more stable and there is less response to the current traffic demand. Tests have shown that $\lambda$ has a different impact on the quality at different traffic demands, e.g., does it have a positive influence on the system when $\lambda$ is small in situation with raising on-ramp demand. [30]-(p.81)

The correcting factor $f_q$ describes the ability to calibrate the algorithm to the local boundaries of the freeway system and is a result of the unique forecasting part of PRO. In contrast to other predictive algorithms, which calculate a future demand by analyzing the data of the near past, PRO predicts the demand by using a second set of detector loops in a certain distance upstream the on-ramp. The variable $f_q$ refers to inaccuracies in the estimated travel time from the detector loops which are further away, caused by lane changing maneuvers and deviations from the pre-set average travel speed. The mentioned input data of the different sets of detector loops are implemented into the equation via $q_{\text{HFB,mik}}$ as the current situation and $q_{\text{HFB,mak}}$ as the forecasted part. [30]-(p.81)

At the origin PRO algorithm time lag between on-ramp entrance and detector loop for the current traffic situation is set to 15 s. The time lag between the on-ramp entrance and the detector loop for the forecasted situation is set to 60 s. The referring distance to the amount of time is calculated by the following equation:

$$D = (T_{\text{HFB,x}} + T_{\text{Schaltung}} + T_{\text{Einfahrt}}) \times v_{\text{prog}}$$

$T_{\text{HFB,x}}$ = time between detector loop and on-ramp entrance

$T_{\text{Schaltung}}$ = estimated time for the data transmission and cycle time calculation

$q_{\text{rs}}$ = calculated metering rate

$q_{\text{Nachfrage,mak}}$ = measured on-ramp demand

$q_{\text{HFB,mak}}$ = measured upstream mainline demand macroscopic

$q_{\text{HFB,mik}}$ = measured upstream mainline demand microscopic

$\lambda$ = weighting factor

$f_q$ = correction factor
\[ T_{Einfahrt} = \text{average time for a car to reach the freeway from the on-ramp meter} \]
\[ v_{prog} = \text{pre-set average travel speed on mainline} \]

The origin PRO algorithm works standardized with \( T_{Schaltung}=5s \), which is an estimated period of time for the transmission of the measured data at the detector loop to the calculating station and includes the time for the algorithm to compute the metering rate. \( T_{Einfahrt} \) is determined to 15s and \( v_{prog} \) is estimated to \( 80 \, \text{km/h} \) and can be adjusted referring to the real situation on the freeway. These assumptions lead to a distance measured from the on-ramp to the first detector loop upstream of 770 m and to the second detector loop of 1760 m. [30]-(p.84) After implementing this ramp metering system on the road, the control scheme of PRO works as illustrated in Figure 8.

---

**Figure 8 PRO Flow Chart** [29]-(p.123)
The description above focuses on the core algorithm of PRO. Additional features are explained in the following paragraph. These features are independent of the algorithm itself and could be used for other strategies, too. In general, none of these improves the results of the algorithm, but partially they are necessary to implement it into a real system.

The minimum cycle time is pre-defined to 4.5 s and leads to a constant green when the computed cycle time is below this threshold. [30]-(p.84) The maximum cycle time is set to 36 s which is equal to 100 vehicles per hour as a minimum flow onto the mainline [30]-(p.85) but tests were conducted with a maximum cycle time of 30 s. PRO also includes strategies to react to incidents on the freeway. To hinder an occurring congestion, PRO sets the cycle time spontaneously to 30 s for a short time, to relieve the impact on the freeway. [30]-(p.86) But this feature turned out to be not as effective as desired. [30]-(p.132) If congestion has already occurred and the average speed is about 40 km/h, PRO turns off automatically because of different driver behavior at this lower speed and the assumption that ramp metering cannot support the system appropriately anymore. [30]-(p.86) The most important feature of PRO is the queue control which avoids queue spillbacks into the arterial network and increases the equity between drivers on the on-ramp and on the mainline. The assumption is that a measured occupancy at the critical point greater than 20% would lead to spillback. To clear the ramp, the metering rate can be increased step by step or it can be set to the minimum cycle time of 4.5 s. [30]-(p.85)

All in all, depending on the tested load on the mainline and on the on-ramp, in micro simulations the algorithm has shown no further improvements to the zero-case study in cases like travel time and emission producing. [30]-(p.111f) This is because the benefits on the freeway always were comparable to the drawbacks on the on-ramps. In this study, ALINEA and a pre-fixed strategy were compared with PRO, too, with the same results. Nevertheless, the last two strategies were worse because of higher probabilities of queue spillback into the arterial network. The advantage of PRO was a statistical reduction of 21.3% of crashes, which was explained by an improved mainline flow through easier merging situations. [30]-(p.113) The new strategy to predict the future traffic demand seems to be very stable if it is matched to the local freeway situation, but the implementation includes the installation of this specific detection structure to fulfill the boundaries...
of the algorithm. An implementation in Baden-Wuerttemberg has shown positive results. This includes an increase of the mainline speed and the mainline capacity as well as a decrease of the travel time. [29](p.134)

3.3 Coordinated ramp-metering algorithms
To improve local ramp metering strategies it was necessary to expand the possibilities of the algorithms. The new idea is to connect the local levels to get an earlier impact on bottlenecks and to enlarge the space to store vehicles on ramps. Coordinated ramp metering algorithms are a common way to avoid congestion and ease the impact on bottlenecks. There are 3 options for how the local ramps can communicate with each other and with the system-wide algorithm, which will be explained based on the associated algorithms of this kind.

3.4 Cooperative
After computing the metering rate at the local level, the solution gets evaluated as independently solvable or as a critical threshold that leads to a communication to other ramps. In general, the previous ramp upstream reacts to the locally unsolvable situation downstream with a reduction of its metering rate to support the following ramp.

3.4.1 Linked-ramp
The original system of the Linked-ramp algorithm was deployed in San Diego, California, USA in 1968 and runs under the name of the San Diego Ramp Metering System (SDRMS). The calibration of this system is based on a pre-evaluation of the section and its influencing sections on previous areas. [10](p.23) The maximum metering rate depends on the capacity of its local area which is affected by the input flow of the mainline and the previous located ramps. This information is selected from historical data. The minimum metering rate depends as well as the maximum metering rate on historical traffic flow data. The (mostly commuter) peaks were analyzed under the consideration of the local possibilities of the ramp to store vehicles. [10](p. 39)
This demand-capacity method is at the SDRMS for example divided into 16 segments of possible metering rates. It is a simplified system of the calculated
period between the maximum and the minimum metering rate divided by the number of segments. The actual metering rate is given by: [4]-(p.10)

\[ \text{Metering flow} = \text{Target flow} - \text{Upstream flow} \]

The highest flow is reached by a free flow where the ramp meter is frozen at green. If the local situation leads to one of the 3 lowest metering rates, the algorithm activates its cooperative part and carries the next previous ramp to lower its metering rate. This response is repeated until all possibilities of the linked-ramps in this section are spent or the traffic demand decreases again. [11]-(p.17)

3.4.2 Helper

The Helper algorithm was first deployed in Denver, Colorado, USA in early 1981 and consists of a local ramp metering algorithm and a superior coordinated algorithm. The freeway is divided into groups consisting of one to seven metered ramps per group which define the section the algorithm can work with.

On the local level the algorithm computes its metering rates based on the upstream mainline occupancy near the ramp and the data of the queue override detector. If the last one detects an approaching spillback onto the arterial roads, the metering rate is increased for one step per cycle time as long the queue override detector is occupied.

In addition to this local response to possible spillback, the coordinated algorithm reacts to this signal. With the Linked-ramp algorithm, the Helper algorithm distributes the load on the previous ramp upstream to relieve the impact on the mainline. This would increase because of the self-solving response of the local algorithm to increase the metering rate to clear the ramp and avoid arterial spillback. The difference with the Linked-ramp algorithm is the separation of this local solution and the coordinated solution.

In general, this separate coordinated part of the Helper algorithm monitors the ramps of its group and categorizes them as critical or not-critical. It considers in a fixed cycle time the two possible events when a ramp would need some help of the other in this section. It checks the metering rate and the queue detector. If one of these two conditions were violated the coordinated algorithm reacts with a distribution of the load on the next upstream ramp. If that leads to a critical condition, the distribution on the next previous ramp gets repeated and so on as illustrated in Figure 9. [11]-(p.5)
This algorithm is considered as sophisticated to calibrate but robust in its syntax. [4]-(p.9)

3.4.3 MILOS
The Multi-Objective, Integrated, Large-Scale, Optimized System (MILOS) is a predictive area-wide algorithm which was developed by the Arizona Department of Transportation (ADOT) during the RHODES-ITMS project in 1999. It is predictive, traffic responsive and interacts between local and coordinated levels in a cooperative way. It uses a macroscopic model to estimate the traffic and solves the optimization with quadratic programming (QP). The algorithm uses measured flow data and computes density and speed for further calculations. [32] The structure of MILOS is based on a classification of the influenced area of the certain modules that have different technologies to handle their area of response. This ranking consists of the strategic, the tactical and the operational strategies, which shall be explained, in more detail. [33]-(p.21)

**Strategic**: the broadest level which controls the entire section and has a horizon of hours, days or seasonal changes. The main task of the strategic view is to update the slowly varying parameters like the overall increase of traffic demand and to transfer the information about sub-networks and other superior data to the lower levels. It contains the modules of anomaly detection and the sub-network identifier.

**Tactical**: the tactical level operates on a horizon of hours and minutes and is described by the coordinated algorithm which reacts on short-term fluctuation, non-recurrent congestion and controls the queues in the section. This level consists of the anomaly detection and the area wide coordination. The computed solutions were fed to the lower level.
**Operational**: this level acts on a horizon of minutes and is responsible for the local level control algorithm which computes an improved strategy for its retarded area on the base of the superior levels. The goal is to reduce ramp queues by responding to short-term fluctuations forecasted by local predictions. As mentioned, these classifications consist of different modules themselves which shall be explained in more detail including their tasks and syntax. The sub-network identifier module is only briefly explained, but at the time of the research there was no functioning algorithm implemented in the complete system of MILOS. The description of the modules begins with the normal structure of the system-wide and local algorithms and is followed by the additional features of statistical process control (SPC), anomaly detection and sub-network identifier. [33]-(p.19)

**Area wide coordination**: this part computes metering rates based on an optimization of queue lengths and mainline throughput maximization. The calculated rates for the system-wide level are given to the next lower level, the local ramp metering algorithm. Under the assumption that the first freeway input cannot be controlled, the following ramps were metered by the quadratic optimization problem below [33]-(p.39f):

\[
\max_{r \in R} \sum_{i=1}^{N} (1 + 2\beta \gamma c_i d_i) * r_i - \beta \gamma c_i r_i^2 - \beta_2 \gamma c_i z_i^2
\]

- \( i \) = number of a certain ramp within the section
- \( N \) = total number of ramps within one section
- \( r \) = computed metering rate
- \( \gamma \) = ratio of the minimum flow entering the mainline from the on-ramps of the section
- \( \beta \) = weighting factor to determine the importance of mainline flow and ramp queues
- \( c_i \) = weighting factor to reflect the congestion conditions at each interchange
- \( d_i \) = on-ramp demand of ramp \( i \)
- \( \beta_2 \) = scaling constant of \( \beta \) to optimize queuing behavior in the section
- \( z_i \) = factor to reflect extra capacity of ramp \( i \) to accommodate the flow at that ramp
The equations which are the basis for this optimization follow below with the itemization of the different variables. The text explanation of a variable is only given once starting at the equation above, to minimize the overview.

\[ \sum_{i=1}^{j} (A_{i,j} \cdot r_i) \leq \text{CAP}_j \quad \forall j \]

\( i \) = defines ramp \( i \)
\( j \) = defines section \( j \)
\( A_{i,j} \) = matrix with flow proportion entering at ramp \( i \) through section \( j \)
\( r \) = metering rate for ramp \( i \)
\( \text{CAP}_j \) = capacity of section \( j \)

\[ (d_i - r_i) \cdot T - z_i \leq Q_i \quad \forall i \]

\( T \) = optimization time horizon
\( Q_i \) = number of vehicles which can be queued on the on-ramp, assuming an average vehicle length

\[ c_i = \frac{\sum_{m=1}^{M} V_{m,i}}{\max_i(c_i)} \quad \forall i \]

\( V_{m,i} \) = offered volume for phase \( m \) at ramp \( i \)
\( C_{m,i} \) = capacity of phase \( m \) at ramp \( i \)

\[ \gamma = \frac{\sum_{i=1}^{N} d_i}{\sum_{i=1}^{N} (c_i \cdot (d_i - r_i, \min)^2)} \]

\( r_{i,\min} \) = minimum metering rate at ramp \( i \)

\[ d_i = \rho_{R,NB} \cdot d_{NB} + \rho_{L,SB} \cdot (1 - \rho_{R,SB}) \cdot d_{SB} + \rho_{T,EB} \cdot d_{EB} + \frac{q_i(0)}{T} \quad \forall i \]

\( \rho_{R,NB} \) = probability to turn right, northbound to enter the on-ramp from arterial road
\( d_{NB} \) = demand of the arterial road, northbound
\( \rho_{L,SB} \) = probability to turn left, southbound to enter the on-ramp from arterial road
\( \rho_{R,SB} \) = probability to turn right, southbound to enter the on-ramp from arterial road

\( d_{SB} \) = demand of the arterial road, southbound

\( \rho_{T,EB} \) = probability to go through, eastbound to enter the on-ramp from arterial road

\( d_{EB} \) = demand of the arterial road, eastbound

\( q_i(0) \) = queue length at the beginning of the optimization iteration

Because this contemplation to involve the turning probabilities to evaluate the real-time ramp flow by minding real-time arterial network demand, Figure 10 shall give a better overview about the explained parameters above.

The whole system maintains itself with the anomaly detector and reacts to infeasible equations, which can occur when an incident on the mainline decreases the capacity, by allowing spillback onto the arterial network.

The capacity of each section can be determined by analyzing historical volume-density curves or calculated with input data about saturation flow rates, number of lanes, merge area restrictions and other information. The goal is to maintain a real-time mainline capacity slightly below the critical capacity for the section to avoid oscillating congestion waves. The actual computed metering rate is between a pre-defined minimum and maximum rate and is automatically set to one of these boundaries for cases of the calculation exceeding the interval. The goal is to maximize the metering rates to minimize the total travel time of the complete freeway system. The computed metering rates were fed to the next lower level,
thus the local algorithm can improve this rate for the more detailed situation at each ramp.

**Local algorithm:** This part of the MILOS system is an optimization tool to improve the suggested metering rates of the upper levels by minding local peculiarities. The main task is to maintain the on-ramp queue and to extend the accuracy by a predictive-cooperative real-time rate regulation (PC-RT) which is only included at the local level to minimize the complexity of the whole system. The prediction refers to the on-ramp demand as well as to the mainline demand and has a horizon of 5 to 7 minutes, but will be re-evaluated after 1 to 2 minutes to improve the metering rates and the forecast as their basis. These predictions do not need to be completely accurate but they must represent the dynamics and the range of possibilities of the system. The computed metering rate only replaces the coordinated rate if it improves the system travel time. It is adjusted to the calculations of the level above over a pre-defined interval $\varepsilon$ in which the local level is allowed to overwrite the coordinated metering rate.

**Anomaly detection:** this SPC-based module is a feature of MILOS to maintain itself and is therefore only activated when the measured data significantly differ from the average of the last assumptions and predictions. This difference is defined by a lower and upper limit with the presumption of an approximately constant demand, which cannot be right at peak periods. For this special occasion, the anomaly detector has implemented jumps to react appropriately on this recurrent traffic behavior. In general, the anomaly detection forces a re-evaluation depending on the detected difference. If a point is exceeded, the PC-RT will re-calculate the metering rate and if a trend is exceeded, the area-wide coordination algorithm will re-evaluate its results.

**Sub-network identification:** normally, the sub-network is pre-defined by a manual operator who regards his decision on local boundaries, e.g. state borders or influence area of traffic control centers. The idea is to include such an automatically identifier into the structure of MILOS, but there is no further explanation available how it could be evolved and how it should work, thus there is no implementation of such a system, yet.

MILOS is a sophisticated algorithm with a lot of possibilities to react to upcoming traffic demand. It is new that the queue treatment is directly included and determining the metering rate and that for this feature the data collection is
expanded to the arterial network. This input data has the advantage in comparison to queue override detector loops that an oscillating metering rate is avoided because there is no certain threshold that determines a change in the computing. It is rather a smooth calculation as known for mainline-based metering rates. The cooperative basis of the algorithm and the two-layer improvement of the metering rate, with a more sophisticated approach on a more detailed level, making MILOS very accurate and flexible in situations of infeasible equations. E.g. if the PC-RT has an infeasible equation, then it will re-activate the area-wide algorithm, which creates an improved metering rate on this new input. This behavior is supported by the different structure of the local and the area-wide level. Additional features, like the response to an incident by decreasing the real-time mainline capacity or the idea of the sub-network identifier make MILOS an interesting project. But all these good ideas lead to the disadvantages that MILOS needs a lot of data even about the arterial network, which can be a significant problem, as tests with SWARM have shown. Furthermore there are a lot of sophisticated real-time calculations with a high data transfer from the upper to the lower level. The algorithm was only tested by ADOT in computer simulations in which the algorithm has shown good results, slightly better than LP.

3.5 Competitive
This mechanism is based on a direct comparison between computed metering rates where the most restrictive rate gets deployed and the other rate has to bow to it. It is often deployed in situations where one important single on-ramp would lead to a collapse of the whole system if it would not be well maintained.

3.5.1 Compass
This algorithm was first deployed in Toronto, Canada in 1975 and is a good example for decisions based on the most restrictive data, i.e. it is a competitive algorithm on local as well as on the system wide level. The ramp metering algorithm is only one part of the whole COMPASS system, which is installed in Toronto, Mississauga, Burlington and Ottawa and contains vehicle detector stations, closed circuit television (CCTV) cameras, changeable message signs and other opportunities for an operator at the traffic control center to have an impact on the traffic on the freeway.
The included algorithm uses the mainline occupancy data near to the ramp, the downstream mainline occupancy and the upstream mainline volume. This information is combined with pre-set thresholds at the local level, downstream occupancies and upstream volume to result in a value that will be compared with a look-up table. Depending on the calculated value of the measurements, the look-up table will lead to the most restrictive metering rate depending on the local and downstream occupancy and the upstream volume and has a resolution of 17 steps. This pre-defined simplified method for data classification is the largest drawback of this method.

Besides the local level, there is an additional off-line optimized metering rate that is influenced by the system wide level. This metering rate is also compared with the metering rate at the local level and again, the most restrictive will be implemented.

In addition, this algorithm includes ramp queue handling that means that the calculated metering rate is set up one level higher than computed. This superior restrictive decision shall lead to a ramp clearance to a value below the on-ramp occupancy threshold. [11]-(p.11)

This algorithm is not as robust as other traffic responsive algorithms because of the pre-defined look-up tables, but it is eventually categorized as a good algorithm. [4]-(p.11) The algorithm was replaced by ALINEA in 2012 in Ontario and thus, there is no location known where the Compass algorithm is still in use. [34]

3.5.2 Bottleneck

This algorithm was first implemented by the Washington State Department of Transportation (WSDOT) on I-5 north of Seattle, Washington, USA in 1981. A well-known representative is the FLOW algorithm of Seattle. It consists of two layers, which are computing metering rates on their certain level, and a comparison of these afterwards. It is a competitive algorithm which means that the most restrictive of these metering rates gets implemented. The following sequence of events in Figure 11 is an overview about the typical calculation steps of a bottleneck algorithm.
At the local level, this algorithm uses the measured real-time upstream occupancy to transfer to a local metering rate. This transition is set up by a pre-defined look-up table made of historical data, which connects the occupancy to a certain volume that can be compared with the local estimated downstream capacity. Finally, the computed metering rate lets a number of vehicles per minute pass to
match the gap between real-time measurement and capacity to create a high LOS. [11]-{p.7}

The system-wide part is activated when two criteria are fulfilled. First, it localizes a violation of a pre-defined occupancy threshold at a bottleneck on the freeway, which represents demand above the capacity, and second, the section upstream of the bottleneck is already storing vehicles, which indicates that more vehicles are joining the freeway than leaving it. The coordinated algorithm determines the gap between demand and supply and reduces the volume of the influencing section by dispersing the burden of storing vehicles to the on-ramps based on a weighting factor. This procedure can be explained with the following equations, at first the reduction of the freeway volume upstream [14]-{p.15}:

$$U_i(t+1) = (q_{in} + q_{on}) - (q_{out} + q_{off})$$

$U(t+1)$ = metering rate reduction for section $i$ for next time step
$q_{in}$ = measured flow on the freeway
$q_{on}$ = measured on-ramp flow
$q_{out}$ = measured flow leaving the bottleneck
$q_{off}$ = measured flow leaving the section on the exit ramps

This reduction must be dispersed onto the upstream on-ramps, weighted because of environmental circumstances:

$$BMR_{ji}(t+1) = U_i(t+1) \cdot \frac{WF_j}{\sum_{j=1}^{n} (WF)_i}$$

$BMR_{ji}(t+1)$ = bottleneck metering rate reduction at section $i$ for ramp $j$ for next time step
$WF_j$ = weighting factor to value ramp $j$

Finally, the computed reduction has to be subtracted from the pre-measured on-ramp flow at each certain ramp:

$$BMR_{ji}(t+1) = q_{ON} - BMR_{ji}(t+1)$$

$BMR_{ji}(t+1)$ = bottleneck metering rate for each section $i$ and ramp $j$ at next time step

The Bottleneck algorithm was ranked as very good even though there are some improvements available like the implementation of the ALINEA algorithm at the local level instead of the original local Bottleneck algorithm. This is tested in an
examination as a small benefit and is called modified Bottleneck (MBottleneck). In the direct comparison the MBottleneck is as good as the ALINEA that could indicate that the effort for the coordinated part of this algorithm has no influence in this certain test. [4]-(p.33)

3.5.3 SWARM
The System Wide Adaptive Ramp Metering (SWARM) algorithm has been developed since 1996 and was tested and implemented in Orange County, California (CA) and Portland, Oregon (OR), USA. This algorithm consists of two separate operating levels, the system-wide algorithm SWARM1 and on the local level any isolated traffic-responsive ramp metering algorithm. The freeway is divided into several sections, which includes a detected bottleneck at the beginning and at the end of it. Between these two bottlenecks there are several entrance and exit ramps. [9]-(p.7), [11]-(p.22)

SWARM1: This part operates on the area-wide level with a predicted traffic demand which is generated through measured data of a few minutes earlier, thus it is pro-active to the actual traffic demand and does not use historical data. The input data which this algorithm needs is the density and a pre-defined critical density, called saturation density, when congestion starts to occur in this section. SWARM1 predicts the density of a pre-defined period of time and compares it to the saturation value of the bottleneck. An exceeding of this threshold leads to a dispersion of the vehicles, correlated to that density, on the ramp storage, thus the saturation density will not be exceeded after the computed amount of time, named $T_{crit}$. Figure 12 and the equations below represent this process in a plausible way. [4]-(p.48f)
\[ d_{\text{tar}} = d_{\text{curr}} - \frac{d_{\text{exc}}}{T_{\text{crit}}} \]

- \( d_{\text{tar}} \) = computed density to stay below saturation value
- \( d_{\text{curr}} \) = measured section mainline density
- \( d_{\text{exc}} \) = computed density above the saturation value
- \( T_{\text{crit}} \) = computed time, when the saturation density is going to be exceeded

\[ V_{\text{red}} = (d_{\text{loc}} - d_{\text{tar}}) \times n \times l \]

- \( V_{\text{red}} \) = computed value, how much the mainline volume has to be reduced
- \( d_{\text{loc}} \) = measured local mainline density
- \( n \) = number of lanes at the local area
- \( l \) = distance to the next ramp

**Local SWARM Algorithm**: as mentioned is every local ramp metering algorithm feasible, e.g. ALINEA (MSWARMI), SWARM 2a and SWARM 2b developed for Orange County, CA or SWARM 2c developed for Portland, OR, where there are no exit ramp detectors, thus it leads to a modification of the algorithms deployed in California. [4]-(p.87), [35]-(p.9f)

The two computed metering rates out of the system-wide and the local perspective are compared and the most restrictive rate gets deployed. SWARM uses an
internal algorithm to exclude inappropriate data from defective detectors, because it needs a reliable basis for an adequate prediction. At the same time, this seems to be the largest drawback of this algorithm because of its high reliability on the demanded data. In the deployed time in Portland, OR, it shows that there are invalid data around 10%, which is a lot in comparison to 1% invalid data at the time of using a fixed-time algorithm. [9]-(p.42) One other possible drawback is the use of density as the metric, since it is not possible to directly measure density using loop detectors.

3.5.4 SZM
The stratified zone metering (SZM) algorithm was first implemented in Minnesota, USA in 2002, as the replacement for the local algorithm ZONE, which is the basis for the new algorithm. The progression was demanded by the public after discomfort about excessively long waiting periods on the entrance ramps. This led first to the shut-down experiment in 2000, which showed that ramp metering is still an improvement in comparison to an unmetered ramp situation, and secondly it led to a desire to decrease the waiting times on the on-ramps. The SZM includes a calculation method comparable to the original ZONE algorithm and furthermore it takes the on-ramp waiting times into account. [27]-(p.1f)
In general, the freeway section is divided into zones which may overlap and include different numbers of metered entrance ramps. The several detector stations shall have a distance of a half a mile with maximum 7 stations in one zone, which leads to a highest zone length of 3 miles. Comparable zones are gathered in layers, e.g. every zone with 2 metered on-ramps belongs to layer 1 and layer 2 includes zones which contain 3 metered on-ramps and so on, illustrated in Figure 13.
The computed metering rates are based on the idea to keep the current capacity under a certain threshold. For the coordinated level, this is determined through:

\[ M \leq B + X + S - A - U \]

- **M** = total allowed volume entering the freeway through entrance ramps
- **B** = pre-defined downstream mainline capacity (not an especially bottleneck)
- **X** = total measured off-ramp volume
- **S** = calculated spare capacity
- **A** = measured upstream mainline volume
- **U** = total measured volume entering the freeway through non-metered ramps

The spare capacity minds extra absorption ability during a time of free flowing on the mainline with low traffic. It is calculated through:

\[ S = (d_{full} - d_{cur}) \times n_l \]

- **d_{full}** = pre-estimated full zone density, approximately 32 vehicles per mile, per lane
- **d_{cur}** = measured current density
- **n_l** = number of lanes within the zone

At the local level, every ramp computes a minimum metering rate to mention a maximum waiting time on the on-ramp. This upper limit is determined through public surveys to estimate an accepted period of time. For Minnesota, this is a maximum waiting time on ordinary entrance ramps of 4 minutes and on freeway to freeway ramps the limit is set to 2 minutes. These thresholds continuously determine a minimum metering rate through the following equation:

\[ R_{min} = \frac{N}{T_{max}} \]
This minimum metering rate allows the last vehicle on the ramp to enter the freeway after the pre-determined maximum waiting time. The weight of this boundary is very high which is represented through the additional behavior of the algorithm, that it increases the metering rate if the queue detector is occupied for 25% of the time, which indicates an upcoming or still existing queue length which cannot guarantee the maximum waiting time. The metering rate is set one level higher every time step, which means that 150 more vehicles per hour can depart the ramp to establish the promised time again. \[D_t = D_{t-1} + 150\]

The new metering rate is calculated every 30 seconds which leads to the maximum release rate of 1716 vehicles per hour being reached within 5 minutes in case calculation has to start form the lowest limit of 240 vehicles per hour. The upper limit is pre-determined through the fixed minimum cycle times and the circumstance that Minnesota has a two-lane on ramp system.

In the end, every computed metering rate for one ramp will be compared and the most restrictive one gets deployed. This is a substantial redundancy because every ramp belongs to multiple zones. On top of that, the algorithm excludes single zones if the detectors provide invalid data or when the density between two measurements drops spontaneously more than 50 vehicles per hour, which implies an incident. \[27\]-(p.3f)

All in all, this algorithm outperformed the ZONE algorithm. Although it led to increased mainline delay, the total travel time was reduced because of the decreased ramp delay. It is not evaluated, yet, whether the increased stops on the mainline would lead to more crashes. But in general, this algorithm represents the transition of DOTs to put more effort in a fair distribution of waiting time by restricting the mainline flow. \[27\]-(p.8)
3.6 Integral

The most sophisticated approach to link the systems to each other would be to compute metering rates separately at the local level and then compare the results on an open perspective. The goal of the comparison would be to find a consensus that fits best in the local and system wide situation. The advantage of this method is that no computed solution gets lost by overly restrictive boundaries. [4]-(p.12f)

3.6.1 Linear programming

Linear programming is in general a mathematical technique to solve a problem of several variables fixed in linear inequalities to define the maximum or minimum of this system.

As a ramp metering algorithm it is first deployed in 1970 on the Hanshin Expressway in Kobe, Japan. This algorithm needs the occupancy and demand data of the mainline accurately and at a high resolution to solve the constraint equations simultaneously and in an adequate way. The several equations represent the ramps and the function to minimize or maximize these, leading to an optimal metering rate, according to the mainline flow. In addition it is possible to tune the boundaries, e.g. with weighted ramps to vary their importance in the section, or it would be feasible to calculate the real-time capacity for the sections to react even if the congestion already occurred. [11]-(p.15f)

The algorithm follows the pattern below: [11]-(p.16f)

1. The roadway is divided into segments (h) between ramps (i)
2. Speed-detection (Yh) for each section \( \rightarrow \) calculate real-time capacity due to congestion results in real-time capacity for each section (Ch)
3. Detection of queue length (Ni) and determine ramp demand (Di) from historical data or queue detection
4. Pre-define storage capacity of a ramp \( \rightarrow \) queue length maximum (Li)
5. Pre-define weight-factor (Qin) with historical occupancy and demand data \( \rightarrow \) scales the traffic flow from i to the mainline downstream flow on h
6. Pre-define weighting factor (A) to estimate the importance of the ramps \( \rightarrow \) guide drivers behavior by penalize unfavorably ramps
7. Maximize function for ramp flow at each ramp (Ui)

\[
Z = \sum_{i=1}^{n} (A_n \cdot U_n)
\]
The constraints for this equation are listed below:

\[ Q_{\text{main}} \cdot U_i + \sum_{n=1}^{i} (Q_{nh} \cdot U_n) \leq C_h \quad \forall h \]

\[ \land \]

\[ (0 \leq U_i \leq N_i + D_i) \quad \lor (\text{ramp demand } + \text{ramp queue} \leq \text{ramp flow rate}) \]

\[ \land \]

\[ (N_i + D_i - U_i \leq L_i) \quad \lor (\text{ramp queue } + \text{ramp demand} \leq \text{ramp flow rate} \leq L_i) \]

\[ \land \]

\[ (U_{\text{min}} \leq U_i \leq U_{\text{max}}) \quad \lor (\text{minimum rate} \leq \text{metering rate} \leq \text{maximum rate}) \]

The disadvantages of linear programming are, as always for coordinated algorithms that, the accuracy of the occupancy and demand data and the equations are pre-defined and static. It works with the assumption that the travel time is a constant. Nevertheless, the algorithm is ranked as good. [4]-(p.14)

3.6.2 Sperry

The Sperry algorithm, 1999 deployed in Arlington, Virginia (VA), USA works on the system wide level in a so called restrictive mode where it solves a demand-capacity equation to ensure that the critical capacity is not exceeded by the demand. The data for that equation is provided by detectors at each ramp (entrance and exit) to maintain an overview about the total number of vehicles in the section. The algorithm distributes the vehicles on the ramps in an average way to keep the demand below the critical mainline capacity which is pre-defined for certain weather conditions. If there is a violated queue length threshold on a ramp, the algorithm switches into the non-restrictive mode to avoid a spillback on to the arterial roads by increasing the metering rate for the affected ramp. [4]-(p.13)

Special features of Sperry are to react to manual operation by adjusting previous ramps downstream in an adequate way or that the algorithm can interact with other sources like variable message signs (VMS). In general, it is trained to favor the mainline by keeping the metering rate a little bit under its possibility to maintain
a higher level of service (LOS) on the freeway. Furthermore the algorithm has the ability to work with predictive traffic flow information and beyond it adjusts its metering rates in an appropriate way, thus the more restrictive rate is set in a time it is really needed and not prematurely. [11]-p.10

The disadvantages of Sperry are obvious the sole use of volumes and pre-defined capacities. Which is harder to handle and depending on the variable capacities, which often do not match the reality.

3.6.3 METALINE

This algorithm is an expansion of the ALINEA and is developed in Paris, France in 1991, shortly after publishing the local traffic responsive ramp metering algorithm ALINEA. The additional element is the system-wide algorithm that causes a transition from scalar values to vectors and matrices. The rest is similarly built to ALINEA, a closed loop algorithm that compares measured occupancies with defined capacities and uses the ramp storage to keep the actual situation below the critical capacity. The modified equation to determine the metering rate is given below: [11]-p.19f), [36]-p.28

\[ \tilde{r}(k) = \tilde{r}(k - 1) - K_1(\tilde{o}(k) - \tilde{o}(k - 1)) - K_2(\tilde{\sigma}(k) - \tilde{\sigma}_c) \]

\( \tilde{r} \) = vector of metering rates for several ramps at time interval \( k \)

\( k \) = time interval

\( \tilde{o} \) = vector of measured occupancies in the section

\( \tilde{\sigma} \) = vector of measured occupancies downstream of each ramp (\( O \subseteq o \))

\( \tilde{\sigma}_c \) = vector of corresponding (to occupancy) capacity values

\( K \) = matrix of weighting factors for the impact of a station on a certain ramp

The area-wide level is established through the matrix \( K_1 \) and \( K_2 \) of weighting factors. These factors determine how much impact one detector should have on the metering rate of a certain ramp. In fact, the \( K_1 \) matrix contains the importance of the measured occupancy of a detector on a certain ramp and the \( K_2 \) matrix determines the impact of the critical detectors to the metered rates at each ramp. If one measurement shall not have any impact on a certain ramp, this value is set to zero. This weight of the several detectors and their influence on the system in certain areas determines the efficacy of the algorithm. [10]-p.22, [11]-p.20
3.6.4 ARMS
The Advanced Real-time Metering System (ARMS) is developed by the Texas Transportation Institute, Texas (TX) USA in 1993 and is defined in contrast to other algorithms to actively risk congestion. This algorithm is divided in two calculating (or even three, if the predictive algorithm is counted as an own) parts. The sequence of events in Figure 14 presents how these individual algorithms work together within the ARMS strategy.

![Figure 14 ARMS Operational Flow Chart][1]

The first part computes a mainline free flow over the whole section and is pictured on the right site of the operational chart. The calculated boundaries are allocated to the several ramps by analyzing the traffic-occupancy and demand data. This process is supported by a predictive algorithm that has a dynamic foundation to improve its own forecast based on patterns and adjusting it with real time measurements. The second part handles congestion which has already occurred as seen on the left side of the chart. The algorithm estimates the congestion clearance time and computes the storage of vehicles in that section. After that, the
burden of storing vehicles is dispersed over the several ramps regarding the occupancy and demand information as in the first part. This additional level is probably due to the implemented congestion risk of the first part. [4]-(p.16), [11]-(p.28f)

ARMS is not implemented in any real system, yet.

3.6.5 RAMBO

Ramp Adaptive Metering Bottleneck Optimization (RAMBO) is an algorithm divided into two parts and developed in Texas in 1991. These two parts are called RAMBO I for the local level and RAMBO II for the system wide calculation of ramp metering rates. Originally both are DOS stand-alone applications that are also implementable in Windows. The unique characteristic is that this tool does not have a direct impact on the ramp metering rate. It is a planning tool which provides ramp metering plans, which can be implemented into the system manually by an operator or to pre-evaluate the functionality of a ramp metering system project. [14]-(p.15), [37]-(p.1)

RAMBO I

This local algorithm computes metering plans for single ramps or groups of isolated ramps. An operator has to type in the demanded information about traffic flow, the target capacity downstream and general boundaries on which the algorithm will calculate 4 different metering plans. [37]-(p.1) It is the decision of the operator to change the input data or to choose one of those plans or a variety which he wants to implement on the real system. [37]-(p.3)

RAMBO II

The coordinated part of the RAMBO system can control up to 12 metered on-ramps and other 12 exit ramps in one section. It is a predictive algorithm that uses linear programming. This program, like RAMBO I, needs certain information about the geometry of the mainline and the ramps, e.g. number of lanes, distance between ramps, definition if it is an exit or an on-ramp. Also it needs the traffic data and other factors like merge quality, capacity analysis or the maximum queue length. [14]-(p.25f), [37]-(p.31f)

All in all, this algorithm is not in use anymore because of its dependence on a manual operator, thus it stays being a mix of a pre-timed algorithm and a slow adaptive ramp metering system. The modeling of the freeway with its conditions also has some aspects of a simulation tool. This diversity and the lack of detailed
information about its structure made it hard to categorize this algorithm in an adequate way and the goal of listing it here is to have a mostly complete list of every developed ramp metering algorithm.

3.6.6 Metering model for non-recurrent congestion
This algorithm is a standard scheme for ramp control, first explained in 1994 and includes most of the features that can possibly be implemented. It computes the traffic flow as a kinematic wave and has predictive parts in it to estimate forecasted densities, which are the basis for maintaining the mainline flow at its highest possible throughput. The algorithm considers also the ramp queue length and creates metering rates based on all named conditions. The data is smoothened by Kalman filtering which is explained in more detail in 2.2.6.1. The equations were solved with linear programming. The sequence of events in Figure 15 is a good overview about the computing steps of this algorithm. [11]-(p.32) It is mentioned that there shall be some improvements for a faster and more robust work, like the exchange of the Kalman filter with the Godunov scheme which would reduce the calculation time. In addition to that, the occupancy-demand flow is not detected directly. It is more determined by the time-varying occupancy-demand data on exit ramps. Even though this algorithm is ranked as very good, there is no further information available about any implementations in a real system. [4]-(p.16)
3.6.7 Coordinated metering using ANN
This ramp metering algorithm from at least 1996 is a V/C based algorithm with the ability to evolve its metering rate results by using Artificial Neural Networks (ANN). At first this algorithm computes a basic metering plan which is going to improve over the time depending on the reaction of the environment. Therefore volume-capacity ratios have to be measured on the mainline upstream and downstream of the ramp and the ramp queue length has to be determined. Figure 16 shows the individual steps of this algorithm. [11]-(p.30f)
This information is matched together at the local level to determine a metering rate based on the pre-defined pattern. This metering rate is deployed by the local ramp and is a further input data via the hidden layer for the other ramps. Thus, the local ramps are connected on a system wide level in an integral way. The next step would be to receive the new measured data and combine them with the previous solution to evolve the metering rate to keep the V/C ratio beneath one. For a better understanding, Figure 17 presents this process in a simple graphical way.

Coordinated metering using ANN is not implemented in any real system, yet.
3.6.8 Fuzzy
This adaptive algorithm initialized and first deployed in the Netherlands in 1989 was also used in 1995 and 1999 in Seattle. Fuzzy is a sophisticated algorithm which depends on human abilities to respond in an appropriate way to the upcoming traffic.

The algorithm can be divided into 3 components which edit the input data step by step. The first component, named fuzzyfication, takes the input data and transfers them into a pre-defined textual description of the actual condition to a certain degree, e.g. 30% small, 70% medium and 0% high for a description of the measured occupancy on the mainline. This fuzzy data is forwarded to the next component, called the inference, where plausible {IF}-{THEN}-{ELSE} rules processing the data. These rules should be kept as few as possible, to keep the syntax simple and the calculation time low. But the manual operator has a lot of opportunities to influence the depth of complexity in this system. The last step is the defuzzyfication, which transfers the result of the used rules into a real metering rate. This component-system is presented in Figure 18 to clarify the single steps.

![Figure 18 Fuzzy Process](image)

The Fuzzy logic algorithm has a great potential because of its plausible structure and the realistic evaluation of input data, i.e. that a certain value is not every time a hard fact, but can be interpreted as multiple states. The transitions are more smoothed and the behavior of the algorithm is more logical and more robust. But the problem of this algorithm is especially these declared rules, which demand a high insight of the operator at the time he programs them, because the robustness
of the algorithm depends on the logical structure which must be implemented manually. In addition, the rules need to be adjusted to the real conditions on the freeway at the time of implementation and also later, if conditions are changing. [4]-(p.13) To defuse this drawback of this algorithm, there were efforts to create a real adaptive algorithm based on the Fuzzy Logic which shall be explained in more detail.

3.6.8.1 ACCEZZ

The adaptive and coordinated control of entrance ramps with Fuzzy Logic (ACCEZZ) was published in 2002 and extends the capabilities of the original Fuzzy controller in a way that the pre-set rules adjust themselves to the conditions on the mainline. It uses the mainline upstream occupancy and the flow-capacity ratio to compute its metering rates. The main improvement is the self-adjusting feature of the algorithm, which allows it to define the components of fuzzification and defuzzification dynamically. For this learning behavior is an approach of neural networks in use which defuses the main drawbacks of the Fuzzy logic that an operator has to tune the algorithm after implementing it and the manual adjusting when traffic patterns are changed. [39]-(p.1)

The algorithm is constructed in two layers. The first is the original Fuzzy Logic at the local ramp as described above. The second layer is the system-wide controller that includes the genetic part to evolve the Fuzzy algorithm. This coordinated portion uses a macroscopic model to compute a minimum travel time within the system. For that it uses the local solutions and develops a best matching area wide result. This forwarded optimization repeats for a pre-fixed number of steps or until the evolved parameters approximate a certain value. In the end, the algorithm chooses the best solution out of all iterations. [39]-(p.2)
This extension of the original Fuzzy logic already has five modifications that are characterized by the type of algorithm they use to evolve the solutions and by the availability of its service, i.e. if it is activated once a day for estimation or is it a real-time forecasting model, which responds every 15 minutes to the traffic flow. In general, the algorithm to evolve the system is either an artificial neural network or based on a genetic theory. There is also an idea to implement a reinforcement learning technique that is dissolved from the supervised training of a neural network theory. [40]-(p.2f) Tests with that technique show that this algorithm has very good results in unknown demand situations because of its real adaptive behavior. [39]-(p.3)

3.6.9 Dynamic metering control
This algorithm developed in 1997 consists of the four following layers: [11]-(p.32)

1) State estimation
2) O-D prediction
3) Local metering algorithm
4) System-wide metering algorithm
Below is a flow chart demonstrating the general structure of this algorithm to illustrate the implementation of the listed layers. Following Figure 20, the strategy of this algorithm shall be briefly explained.

![Flow Chart](image)

Figure 20 Dynamic Metering Control Structure [11]- (p. 32)

The system-wide metering algorithm computes metering rates on predictive assumptions about demand. Its goal is to minimize the travel time of the whole area. This includes the mainline travel time as well as the delay on the ramps. Thus, the local ramp metering algorithm maintains the estimated metering rates of the system-wide algorithm and improves this rate for the actual situation at the local level. [11]- (p. 32)

The state estimation and the O-D prediction layer were used to generate the estimated forecast model, on which the system-wide algorithm calculates the metering rates. The combination of the local and the system-wide level is established by following equation:

\[ r_t = \bar{r} - K(o_t - \bar{o}_k) \]

- \( r_t \) = local ramp metering rate at time \( t \)
- \( \bar{r} \) = system-wide ramp metering rate
- \( o_t \) = local occupancy at time \( t \)
- \( \bar{o}_k \) = system-wide occupancy

This is a sophisticated algorithm that might produce good metering rates but it depends substantially on the provided data. Even though this algorithm is ranked as very good, there is no information about an implementation of this algorithm in reality. [4]- (p. 17)

3.6.10 RMS 2000

This algorithm is a result of a project of the Federal Highway Administration (FHWA) that began in 1997 and was completed in 1999. The Ramp Metering System 2000 (RMS 2000) is better known through former summaries as Ball
Aerospace/FHWA [11]-(p.26) [4]-(p.15) and is a mixed pre-timed and adaptive, coordinated strategy. The coordination is guaranteed through an offline simulation that considers the entire corridor and creates metering plans on this foundation. [11]-(p.27) The exact contents shall be explained in the following subchapter. The general structure of RMS 2000 is a chain of three consecutive steps, which convert the simplified algorithm to a more sophisticated one, if needed. It starts first with a **Time-of-Day Plan Generator**, which uses archived data about the environmental conditions, e.g. storage capacity of the several ramps, and historic traffic data to create metering rates. There is a certain number of connected traffic demand with an associated metering rate deposited in the archive of RMS 2000. This also includes information about the expected traffic behavior after implementing the associated cycle time. [21]-(p.6f) These expected changes are the foundation for the second part, the **Local or Segment Regulator**, which gets activated when the measured traffic behavior differs from the pre-defined assumptions. It replaces the old fixed rate with another that might better fit the current situation, and it adjusts this new rate based on pre-defined parameters. [21]-(p.7) The third step is the **Ramp Queue Management**, which controls the implemented metering rates with the on-ramp demand and calculates the probability of a spillback on to the arterial network based on this information. [21]-(p.7) To summarize this general structure, Figure 21 illustrates these three major steps.

![Figure 21 General Structure of RMS 2000](21)-(p.6)
These explained major steps are broken down into five modules which realize the described features and procedures of RMS 2000. These functions are:

1. Roadway Modeling
2. Traffic Modeling
3. Generation of Model-Based Metering Plans
4. Evaluation of Current Traffic Conditions
5. Generation of Real-Time Metering Plans

The first three modules are the foundation for future metering rates which are determined through the best matching plan to a measured pattern. The last two modules guarantee an adaptive behavior which responds to incidents. To clarify the task of every single module, the explanation of the general structure above is itemized and combined to the executive module.

Roadway Modeling: it contains the general boundaries, e.g. the number of lanes or the situation of the detection devices. [21]-(p.8f)

Traffic Modeling: this module generates traffic demand estimations based on historical data. This step can be done from fully manually operated to fully automated. [21]-(p.9)

Generation of Model-Based Metering Plans: creates time of day metering plans for every on-ramp meter with expected traffic conditions for each plan. Also there are adjustments possible to improve the plan-based cycle times through a comparison of the expected traffic conditions with the occurred conditions. [21]-(p.9f)

Evaluation of Current Traffic Conditions: this module uses the surveillance devices to compare the current traffic situation with the pre-defined patterns and it searches for inconsistencies. [21]-(p.10f)

Generation of Real-Time Metering Rates: the final step computes the cycle time based on the measured traffic conditions. It chooses the best matching pattern from the model based plan and looks for any incidents on the roadway. If there are deviations, the pre-chosen plan gets replaced by another better matching plan. The taken plan is adjusted to the real conditions via pre-defined parameters and the resulting metering rates are forwarded to the host traffic control center. [21]-(p.11f)
The interactions of the described modules are presented in Figure 22. Alternative data collecting strategies. However, these approaches reflect that the quality of the data is an important factor. [5]-(p.27)

The explained traffic pattern estimations are based on OD-estimation and use DYMIN, a dynamic network equilibrium traffic model of macroscopic flow, to create these patterns. [21]-(p.19f) DYMIN is also able to work with HOV lanes and bus bypasses because of multiple vehicle type recognition. The step, which detects the best model based metering plan, uses the pattern of DYMIN and non-linear programming to minimize travel time. This combination is called OPDYMIN. [21]-(p.49) The referring cycle times are computed by an improved ALINEA algorithm with a shortened iteration time and with the explained queue management. [21]-(p.80) In the case of detector loops failures, the input data is set to the previous information or historic data is used to keep the metering rate consistent. [21]-(p.86) In general, RMS 2000 has the capability to implement the arterial road network with surveillances of the choice of alternative routes. [21]-(p.7f) But in the end, RMS 2000 is not implemented in any real system, yet. [41]

3.6.11 HERO
The HEuristic Ramp metering coOrdination algorithm is a new method to calculate metering rates and was first tested in 2006 in Paris, France (in a restricted version
with the name Coordinated Heuristic Control (CORDIN) [36]-[p.23]) and Amsterdam, Netherlands (with the RWS algorithm on the local level [24]-[p.19]), and is now implemented in Australia. Concerning to previous coordinated ramp metering algorithms and their disadvantages, this algorithm was simplified and optimized to use real-time measurements, but without doing real-time calculations. This way the used capacities decrease substantially.

This algorithm consists of a local algorithm, which is originally ALINEA but it would be possible to implement any other local algorithm as happened at the Rijkswaterstraat test, and the coordinated algorithm, which is called HERO. [24]-[p.2] This supplement responds if a critical queue length threshold is reached at one ramp by setting this critical ramp as a “master-ramp” and recruiting other ramps upstream as “slaves”. This procedure is similar to other coordinated algorithms to use the ramp storage capacity of the whole section to relieve the critical ramp. HERO uses gradually up to six upstream ramps to disperse the burden, which is measured and maintained by the control of the queue length with following equation: [23]-[p.6f]

\[
q_{\text{min}}(k) = \frac{q_{\text{max}}(k) \cdot \sum_{i=1}^{n} q_{i}}{\sum_{i=1}^{n} q_{i,\text{max}}}
\]

\( q_{\text{min}}(k) \) = the minimum queue length, which has to be reached at slave ramp \( k \)
\( q_{\text{max}}(k) \) = the maximum queue length at the slave ramp \( k \)
\( q_{i} \) = the actual queue length at each ramp (master + slave) in the coordinated control
\( q_{i,\text{max}} \) = the maximum queue length at each ramp (m + s) in the coordinated control

This information about the desired minimum queue length is transferred to the local level. This local algorithm will decrease the cycle time at slave-ramps as long as the desired queue length is not reached. The local algorithm at the master-ramp will act normally as so as there is no coordinated control. One example for an implementation of the mentioned recalculation of the cycle time is the test at Rijkswaterstraat with the following procedures:

\[
\text{Cycletime}_{\text{rest, capacity}} = \frac{n_{\text{lane}} \cdot n_{\text{veh}} \cdot 3600}{\text{Capacity}_{\text{mainline}} - \text{Flow}_{\text{mainline}}}
\]
Cycletime$_{\text{rest, capacity}}$ = generated cycle time to use the rest capacity on the mainline

$n_{\text{lane}}$ = number of lane on the on-ramp

$n_{\text{veh}}$ = number of vehicle per green per lane

Capacity$_{\text{mainline}}$ = pre-defined mainline capacity

Flow$_{\text{mainline}}$ = measured mainline flow upstream the on-ramp

If the computed cycletime$_{\text{rest, capacity}}$ is larger than the pre-used cycle time at the local level, then the algorithm proves the following boundaries:

\[
\text{Cycletime}_{\text{min, queue}} = \frac{n_{\text{lane}} \cdot n_{\text{veh}} \cdot 3600}{(\text{queue}_{\text{curr}} - \text{queue}_{\text{min, desired}}) \cdot \frac{3600}{\text{Control\_Interval}}} \]

{\text{IF}} \text{queue}_{\text{curr}} < \text{queue}_{\text{min, desired}}

{\text{THEN}}

\[
\text{Cycletime} = \max\{\text{Cycletime}_{\text{rest, capacity}} ; \text{Cycletime}_{\text{min, queue}}\}
\]

Cycletime$_{\text{min, queue}}$ = generated cycle time to reach the minimal queue

queue$_{\text{curr}}$ = the current queue on the on-ramp

queue$_{\text{min, desired}}$ = the minimum queue which is desired by the coordinated algorithm

Control\_Interval = the time between the checks of the current situation

Finally, there is the possibility that the computed cycle time of the maximum queue length (Cycletime$_{\text{max, queue}}$) is less than Cycletime$_{\text{rest, capacity}}$. The following equations will explain how the Cycletime$_{\text{max, queue}}$ will be calculated and solve the last boundary. [24]-p.18f

\[
\text{Cycletime}_{\text{max, queue}} = \frac{n_{\text{lane}} \cdot n_{\text{veh}} \cdot 3600}{(\text{queue}_{\text{curr}} - \text{queue}_{\text{max, admissible}}) \cdot \frac{3600}{\text{Control\_Interval}} + \text{Flow}_{\text{ramp}}}\
\]

{\text{IF}} \text{queue}_{\text{curr}} > \text{queue}_{\text{max, admissible}}

{\text{THEN}}

\[
\text{Cycletime} = \min\{\text{Cycletime}_{\text{rest, capacity}} ; \text{Cycletime}_{\text{max, queue}}\}
\]

Cycletime$_{\text{max, queue}}$ = generated cycle time to stay below the maximum queue on the on-ramp
Queue_{\text{max,admissible}} = \text{maximum admissible queue length on the on-ramp}

Flow_{\text{ramp}} = \text{incoming flow on the on-ramp}

The coordinated control of HERO disengages the slave ramps when the queue length at the master ramp falls below a deactivation threshold. This threshold should be lower than the activation threshold to produce unambiguous conditions.

3.6.12 SRMS
The Sydney Coordinated Adaptive Traffic System (SCATS) Ramp Metering System (SRMS) from Australia is a strategy to compute cycle times at intersections in urban areas, in general. It was developed in 2005 to develop an adaptive system for traffic signals, which responds automatically to the actual traffic demand. The SRMS algorithm is a part of the SCATS System and shall be explained in more detail below. [31]-(p.2)

The strategy of SRMS consists of several parts which are listed and explained below. These features are the foundation of the algorithm to compute an optimized and integrated solution. [31]-(p.8f)

1. Dynamic bottleneck location identification: this features allows algorithms to localize the critical bottlenecks automatically and is thereby less dependent on the manual operator and fixed traffic pattern assumptions.
2. Simultaneous coordinated response: the ability to respond to a bottleneck with several coordinated ramps at the same time.
3. Data-fusion of multiple mainline fundamental traffic measures: the equation of the algorithm calculates solutions based on speed and occupancy to guarantee reliable mainline capacity estimations.
4. Real-time integration with arterial traffic signals: this is an approach to connect the calculations for the on-ramp metering with data from the mainline and the arterial network.

The first two features are provided several ramp metering strategies, e.g. SZM or Bottleneck. But the last two features are an unique approach of the SRMS even though MILOS tried to connect the arterial network with the mainline, but not with real time measurements as SRMS is doing.
The following two equations describe the foundation of SRMS to compute cycle times: [31]-(p.9)

$$\varepsilon_j(k) = o_{j}^{\text{crit}} - o_{j}^{\text{meas}}(k)$$

$$\alpha_j(k) = \alpha_j(k - 1) + \Delta \alpha$$

- \(\varepsilon_j(k)\) = calculated error at detector station \(j\) at time step \(k\)
- \(o_{j}^{\text{crit}}\) = defined critical occupancy at detector station \(j\)
- \(o_{j}^{\text{meas}}\) = measured current occupancy at detector station \(j\)
- \(\alpha\) = accumulated conversion of \(\varepsilon\) through a pre-defined dependency
- \(\Delta \alpha\) = direct dependent summand from conversion of \(\varepsilon\)

The dependency of \(\alpha\) and \(\varepsilon\) is given by the alpha increment function, which is pre-defined by the operator and visualized below.

![Alpha increment function](image)

This system is called the overlapped occupancy control strategy (OOCS) and is realized by controlling the situation at detector loops with several on-ramps. The regulation after an exceeded threshold at any detector station responds by limiting the associated ramps. This is more responsive than comparable algorithms because of the simultaneous use of several ramps and not a slowly increased regulation by adding ramps step by step, if the goal is not reachable with less regulated ramps. The local control at the on-ramps is a modified ALINEA, depending on the trend of the alpha increment function, where a straight line through the origin would reflect the original ALINEA. [31]-(p.10)

The OOCS is a feedback controller with the general advantages and disadvantages described in 2.2.6.5, thus it is expanded by a Feedforward Disturbance Compensation (FDC). This unique implementation first introduces a feedback-feedforward scheme to a ramp metering algorithm and is a new
approach of the SRMS. Including the FDC, the cycle time is computed by the following equation: [31]- (p.12)

\[ q(k) = \frac{n}{z} \cdot \left[ 1 - \max_j (\alpha_j(k)) \right] \]

- \( q(k) \) = computed cycle time at time step \( k \)
- \( n \) = describes how many vehicles are allowed to join the zone (see M+F at 3.2.3)
- \( z \) = number of ramps in the zone

Finally, there is an idea to expand the possibilities of SRMS by adding the opportunity for the manual operator to define IF-THEN-ELSE rules. This allows the user to respond to queue spillback and other arterial network related problems and increases the complexity of the algorithm. This approach may help to control the global system with a ramp metering algorithm and can be a new classification of integrated algorithms. [31]- (p.13)

This algorithm is still in enhancement, e.g. is one goal to implement data from arterial network detectors to estimate the on-ramp demand and possible queue spillbacks more precisely. Nevertheless, SRMS was tested in Auckland, New Zealand on several freeways in comparison to former data on these roads and has shown good results.

3.6.13 TIME

The Traffic Introduction Metering (TIME) is a rarely-described adaptive algorithm which was developed by MIZAR Automazione and is implemented on three ramps in Stockholm, Sweden and on the Mestre Ring Road near Venice, Italy. [7]-(p.7)

The general idea is to keep the mainline flow stable, but it tries to avoid queue spillbacks at the on-ramp on to the arterial network. [42] This ramp metering strategy is able to interact with other sources, e.g. external traffic prediction, VMS or CCTV. [7]-(p.7)

This algorithm is divided into a central part, called Omnivue, which can be overviewed by manual operators, and the algorithm itself, which computes cycle times based on all traffic input data at all ramps. This optimized calculation is sent to the local controllers, called SPOT, which try to implement the optimized solution based on the local condition. [7]-(p.4) The local controllers do not influence each other but they exchange information, e.g. the downstream flow information from one corridor is sent to the following and is used there as upstream flow input data.
The Omnivue component is more restrictive in every case. It is able to change single output information as well as it can switch the system off or let the system run based on a pre-timed data set. Unfortunately, there is no further information about the working method of the algorithm itself available.

### 3.7 Conclusion of the Algorithm Overview

The listed algorithms are a summary of all known strategies to compute metering rates which were available at the time of research. Because of the diversity and the local peculiarities, this list cannot fulfill a complete overview of all algorithms and their varieties, thus it is possible that there are some algorithms or small local modifications left out. Also, this overview only contains strategies which are implemented or ready for an implementation, including these, which are not even used anymore, for a complete overview of the state of technology. Efforts to invent a new algorithm are only known for the area of Switzerland. Other researchers mostly improve existing strategies and create adjustments for an improved local implementation, e.g. for the HERO algorithm in different countries.

All in all, this summary shows that ramp metering algorithms are not bound to a strict direction of developments. Next improvements and possible developments are only estimated in a rough overview, e.g. that the ramp metering algorithms will be connected with other controlling sources like VLS or VMS or a combined global work between freeway and arterial network. But the syntax to fulfill this goal depends on the creativity of certain researchers on this field. Some new ideas could lead the development in another direction, e.g. it was a goal in the middle of the 1990s to complicate and sophisticate the algorithms to increase their options to respond to the mainline demand, but some new approaches try to simplify the calculations for a more stable system. It seems that there is no general optimized method, reflected by the variety of implemented algorithms all over the world with different ramp metering strategies, ranging from local, pre-timed to adaptive, dynamic implementations. To clarify the world-wide situation, the next chapter presents an overview about the various approaches of different countries and shows how the use of ramp metering algorithms has grown.
4 Use of Ramp Metering World-Wide

In countries around the world there has been a growing demand to improve the effectiveness of existing freeways without consuming more space. Ramp metering is a well-known method used in a wide range of countries with varying congestion problems and driver behaviors. To illustrate the increasing use of ramp metering, this chapter provides an overview about the formation of ramp metering and its dispersion around the world. The most accurate account of the latest state of ramp metering strategies is compared to former publications. Following, an additional chapter highlights the outstanding situation of the USA in development, deployment and research efforts in ramp metering. The chapter also summarizes the current use of ramp metering in the USA and demonstrates the diversity of algorithms used throughout the country.

4.1 Brief Overview of the History of Ramp Metering

Ramp metering is an American invention developed as a reaction to the accumulation of traffic incidents following the deployment of highways and freeways in the mid-1950s. As the use of transportation increased, so did the occurrences of congestion and crashes. Given right-of-way constraints, it is usually not possible to respond to the demand by increasing the size of the roadway. The first attempt to mitigate this growing congestion by local ramp control, metered by police officers, was implemented along the Eisenhower Expressway in Chicago, 1963. It should be noted that earlier experiments in New York City tested manual traffic metering in tunnels. The improved situations led to a dissemination of ramp metering strategies to other states and to improvements and tests on the system itself. For example, a total ramp closure instated in Los Angeles, USA in 1967 and in the implementation of prioritized bypass lanes in Minneapolis, USA in 1972. These are states that are still leading in the research on ramp metering systems in the USA. [1]-{p.5f)

The development and deployment of ramp metering systems in Europe started in England in 1986 and continued in the Netherlands in 1989. The total amount of metered ramps does not reach the number of installed systems in the USA. Nevertheless, there were many invented algorithms; such as the ALINEA developed in France in 1990, which is even today a commonly used local ramp metering algorithm. The world-wide uses of such systems lead to the deployment
of many algorithms that consider the local conditions throughout the continents. [2]-(p.4f)

The time line shown in Figure 24 is an illustration of invented ramp metering algorithms and their deployments in various cities. Due to the sharp increase in the exchange of knowledge in the beginning of the 1990’s, the segmented time line ends and an era of algorithm information and knowledge exchange overlaps. It is not possible to determine an exact date of an invention of an algorithm anymore because there are pre-tests, deployments of unfinished algorithms that were improved over the time and numerous publications where it is not clear if it is ready to deploy or if the research is still in progress. Some dates are given based on a publication date, a first simulation or a first field implementation. In addition to the improvement of the algorithms themselves, there has been progress in the use of collected data: from static time schedules for the metering rates to traffic responsive up to predictive systems, where the trend of traffic demand is analyzed and results in a short term forecast. This demonstrates that ramp metering itself is an adaptive system that is improvable over the time as the hardware and software of detection, computing and communication advances. At the moment, sophisticated algorithms are being deployed which are coordinating several ramps under system wide circumstances using short term traffic forecasts. The limit of these systems is the provision of accurate real time data. [1]-(p.6), [9]-(p.79)

There are two main approaches available to manage the data. The first is to implement procedures into the algorithm to evaluate the measured input information and whether they are feasible or not. This process excludes the infeasible data from subsequent calculations, such as invalid overdriven or broken detector loops. The disadvantage of this method is that there might be a sizable loss of data. [9]-(p.42) Another approach is to change the strategy of collecting these data and to find a method that removes the least amount of invalid data as possible. One alternative way might be wireless data transmission. Another advantage of wireless transmission is that it avoids the problem of wire theft. [6]-(p.4) An additional technique is the measurement of traffic with radar systems or probe vehicles. It is not certain if there are notable improvements with these alternative data collecting strategies. However, these approaches reflect that the quality of the data is an important factor. [5]-(p.27)
Figure 24 Overview of Invented Algorithms
In addition to the research on data restriction algorithms, research on ramp metering algorithms used to test the connectivity of on-ramp metering with other sources on the mainline, e.g. variable message signs (VMS) or variable speed limits (VSL), using one calculation being developed to increase the potential to influence traffic. [43] This more sophisticated approach is in contrast to another strategy which tries to reduce the complexity of the algorithm and the required data. A typical example of the first method is the MILOS algorithm, and for the reduced strategy HER0 is a good example.

4.2 Explanation World Map
When ramp metering was first implemented in Chicago, USA (as described in section 4.1), it spread internationally due to its success to expand the capacity and increases the mainline flow; decreasing congestion of freeway systems without additional lanes. Today, there are various approaches to implement the algorithms and to adapt these to the local situation to improve their results. The map in Figure 25 illustrates the latest numbers of metered ramps in these regions. It was not possible to get information from the last two years from every traffic control center from every country and there is wide diversity within each region. Thus, the following explanation describes merely the number at known sites with additional information about its date and a precise spatial alignment.

Figure 25 Amount of Metered On-Ramps in the World
Argentina (2013): It was assumed that Argentina has ramp metering systems but the most recent source denied this and added that there is no attempt to implement such a system in the near future. [44]-[p.6]

Australia (2014): There are 77 metered ramps in the region of Melbourne under the organization of VicRoads, where 63 of them are situated on the M1 and 14 are installed on the M80. In the area of Queensland are 8 metered on-ramps. All are using the HERO algorithm. There is a project in development in Western Australia to install a ramp metering system which will use HERO. However, no funding is currently available for the project. [45], [46], [47]

Canada (2014): Eleven metered ramps are situated in the area of Mississauga, Ontario and are an additional feature of the operating COMPASS system. In 2012, the original ramp metering algorithm, also named Compass, was replaced by ALINEA as a more robust solution. [34] There are no ramp metering systems in Toronto, Burlington or Ottawa. [48] There was no further information about other provinces available.

France (2014): France has made some minor efforts to implement ramp metering. Two ramp meters are located on a trial in Bordeaux. There is also unconfirmed evidence about such a trial in Marseille. Ramps have been installed in Paris: Six installations on the A6 and there is also a planned program for Grenoble, which may be implemented at the end of 2015. [49]

Germany (2014): Germany has 104 ramp meters unevenly dispersed over the country. The following itemization will give a better overview of their locations. Both states listed with a zero are in this summary, Lower Saxony and Hesse, is a way to distinguish between which assumed ramp meters are actually implemented and which not.

- Lower Saxony: 0
- Hesse: 0
- Baden-Wuerttemberg: 3
- Bavaria: 4
- North Rhine-Westphalia: 97

Bavaria and North Rhine-Westphalia are using an ALINEA algorithm to meter their ramps. Baden-Wuerttemberg is using PRO. North Rhine-Westphalia will implement an additional 15 metered ramps in the next few years. They will also replace the PRO-algorithm in their area with ALINEA. There are upcoming efforts
to implement three more ramp metering systems in the Bavaria region. [50], [51], [52], [53], [54]

**Great Britain**: There is some evidence that there are 100 ramp meters in the UK. [47] However, there is no source available to confirm this assumption and there is not itemized where these metered ramps are situated.

**Italy (2014)**: An unknown number of metered ramps are installed on the Mestre Ring Road north of Venice. The system is called MARCO and the implemented algorithm is TIME. It appears to be the only place in Italy where ramp metering is used but there is no further information about the details of the system. [55]

**Japan**: Formerly, metered ramps were known to exist on the Hanshin Expressway. [11]-p.1 There is no source available to confirm this information and no known number of metered ramps.

**Netherlands (2014)**: The Netherlands conducts a lot of research in the field of ramp metering algorithms, e.g. tests with new algorithms like Fuzzy, V-ALINEA or HERO. The number of deployed ramp meter has grown from 20 installed systems in 2000 to 120 metered on-ramps by 2014. At the moment, there are tests with the ALINEA algorithm in combination with type of HERO algorithm on the coordinated level. In the field, most of the metered ramps were calculated by the Rijkswaterstraat algorithm, which is a feed forward control algorithm based on flow and capacity. [56]

**New Zealand (2014)**: An installed ramp metering system is only used in the area of Auckland and is computed by the SCATS (SRMS) algorithm. [57]

**Russia**: There is some evidence that there are metered ramps situated in Moscow, based on some clues dispersed over out-of print publications. Unfortunately there is no reliable source available which confirms this assumption.

**South Africa**: (2014): Some ramp metering trials were conducted in Durban as part of a comprehensive upgrade of the freeway system of Pretoria/Johannesburg in 2005. Today, there are ramp metering capabilities situated on this corridor. However the ramp meters were never turned on because of the existing toll system on the corridor and the uncertainty of these two strategies running simultaneously. [58]

**South Korea**: The extant of information on the number of metered ramps situated in South Korea is that they are located in three sections in the area of Seoul. There is no further explanation available about how many ramps one section
includes. South Korea is probably using a fuzzy self-adaptive PID controller algorithm. [59]

**Spain:** It is unlikely that ramp metering in Spain will be implemented due to the fact that the major routes are privatized and are imposed with a toll. It would be counterproductive for the owner of the highway to hinder the traffic demand onto their road. [60] Trials were conducted in Barcelona and Murcia and a former system in Madrid, which was shut down because of political will. There may be a system installed in Zaragoza, but neither the scope of the project nor implementation has been confirmed. [61]

**Sweden (2014):** There are three metered ramps located in the area of Stockholm. They are computed by the TIME algorithm with a switch-on algorithm based on ALINEA. There are additional ramp metering systems in Gothenburg, which will be reactivated soon, probably with ALINEA as the computing algorithm. [62]

**Switzerland (2014):** There is research exploring the possibility of implementing ramp metering system in Switzerland. At the moment are there simulations to develop an algorithm which can handle the local situation of short, urban-near ramps by connecting it with other road control mechanisms (integrated solution with VMS and VSL). The simulation will last until 2014 and could be expanded to a field testing. A paper about the solutions is in progress. [43], [63]

**USA:** the number of total metered ramps in the USA is itemized and explained in chapter 4.3.

### 4.3 Explanation USA Map

The USA is the leading country in the deployment of ramp metering on their highways and freeways. This is reasoned historically as the ramp metering system was first invented and deployed in Chicago in 1963. It may also have to do with the geometry of the freeways. Freeways in the US lack the same type of acceleration ramps found in other countries; introducing vehicle platoons that have a greater impact on the mainline flow than in countries that normally have these lanes. But even within the USA there is a diversity of implemented metered ramps which represents the local aspirations on the field of ramp metering in different
states. This plurality can be itemized and explained in more detail to get a true overview about the situation of ramp metering in the USA.

The number of ramps is extracted out of a general survey of the Federal Highway Administration published in 2014. The survey itemized several state DOTs and assigned metered ramps to them, which is presented in Table 6. The number of ramps in each state is summed and assigned to the appropriate state. This final conclusion is presented in Figure 26. Each of the summed number of ramp meters for each state is located at the center of gravity of the state. To compare the development of ramp metering, the data of published studies from 1999 and 2010 are also included as maps. The disadvantage of the surveys is that they ask for
actual metered ramps at the time of the survey, which explains the alternating numbers of metered ramps in some states.
Table 6 contains, as mentioned before, the number of activated metered ramps at the individual locations. The table also displays data from previous years in order to obtain a better overview of where ramp metering capability exists or which states decided against this strategy for managing their traffic demand. The bold numbers represent the sum of each state ramp meters, which is presented in Figure 26. The others are, as far as it is known, the itemized number of ramp meters from each state DOT. The first column of activated ramp meters from 2014 is considered to be the clearest picture of the current situation. The next column contains the counts from 2010, which were provided through a survey of the Research and Innovative Technology Administration (RITA) in 2010. [64] Note that the far right column compares the approximate data from 1999. Unfortunately, the number of metered ramps from 1999 is not itemized to the DOTs, except for Texas. [11]-(p.1)
Table 6: Itemized Overview: Number of Metered Ramps in the USA [11](p.1), [64]

<table>
<thead>
<tr>
<th>State, Referring City</th>
<th>Assigned DoT</th>
<th>2014</th>
<th>2010</th>
<th>1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZ, Phoenix</td>
<td>Arizona DOT Statewide TOC</td>
<td>202</td>
<td>182</td>
<td>65</td>
</tr>
<tr>
<td>CA, Bakersfield</td>
<td>Caltrans District 6</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CA, Los Angeles</td>
<td>Caltrans District 7 - Los Angeles TMC</td>
<td>1016</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td>CA, Fresno</td>
<td>Caltrans District 6</td>
<td>64</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>CA, San Bernardino</td>
<td>Caltrans District 8</td>
<td>226</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>CA, San Diego</td>
<td>Caltrans District 11 TMC</td>
<td>0</td>
<td>288</td>
<td></td>
</tr>
<tr>
<td>CA, San Francisco</td>
<td>Caltrans District 4</td>
<td>425</td>
<td>298</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO, Denver-Aurora</td>
<td>Colorado Department of Transportation</td>
<td>70</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td>DC-VA-MD-WV, Washington</td>
<td>Virginia DOT - NRO Traffic Operations Center</td>
<td>24</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>FL, Miami-Fort Lauderdale</td>
<td>Florida DOT-District 6 – Sun Guide TMC</td>
<td>22</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>GA, Atlanta Sandy Springs</td>
<td>Georgia Department of Transportation</td>
<td>170</td>
<td>154</td>
<td>5</td>
</tr>
<tr>
<td>IL-IN-WI, Chicago</td>
<td>Illinois Department of Transportation</td>
<td>0</td>
<td>113</td>
<td>110</td>
</tr>
<tr>
<td>LA, Baton Rouge</td>
<td>Louisiana Department of Transportation</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td></td>
<td>0</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td>0</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>MN-WI, Minneapolis</td>
<td>Minnesota Department of Transportation</td>
<td>425</td>
<td>425</td>
<td>370</td>
</tr>
<tr>
<td>MO-IL, St. Louis</td>
<td>Illinois Department of Transportation</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MO-KS, Kansas City</td>
<td>Missouri Department of Transportation</td>
<td>28</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NV, Las Vegas</td>
<td>Nevada Department of Transportation</td>
<td>70</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>NY-NJ-PA</td>
<td>New York State DOT - Long Island Region 10</td>
<td>0</td>
<td>82</td>
<td>0</td>
</tr>
<tr>
<td>OH, Columbus</td>
<td>Ohio Department of Transportation</td>
<td>24</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>OH-KY-IN, Cincinnati</td>
<td>TRW/ARTIMIS OCC for Ohio DoT</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR, Portland</td>
<td>Oregon Department of Transportation</td>
<td>142</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>OR, Eugene</td>
<td>Oregon Department of Transportation</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>146</td>
<td>140</td>
<td>60</td>
</tr>
<tr>
<td>PA-NJ-DE-MD, Philadelphia</td>
<td>Pennsylvania DoT District 6-0</td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>PA-NJ</td>
<td>Pennsylvania DoT Allentown</td>
<td>0</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>TX, Dallas</td>
<td></td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>TX, Houston</td>
<td></td>
<td>0</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>
4.4 Conclusion of the World-Wide Overview

This chapter is a description of the implementations of the previous explained ramp metering algorithms. It presents the importance of this strategy for extending the use of freeways without greater physical expansion of the existing system. Countries, which implemented a ramp metering algorithm once, increase their efforts in this field over time, because of visible effects which were discussed in chapter 3. The detailed explanations about the local implemented systems in the various countries provide an overview about the still predominant diversity in algorithms used in practice. It strengthens the assumption that different strategies work better or worse on varying boundaries and thus it is necessary to improve the different strategies independently to keep extensive opportunities for the traffic control centers to respond to their unique problems in their area of responsibility. This overview, about the development and the deployment of ramp metering algorithms, concludes the general summary of the latest situation in the world and can serve as a basic starting point for further research. It is basic knowledge to understand development strategies in different regions and their varying approaches in their evaluation processes.

<table>
<thead>
<tr>
<th>Location</th>
<th>Implementation</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT, Provo-Orem</td>
<td>Utah DoT Region 3</td>
<td>24</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>UT, Salt Lake City</td>
<td>Utah DoT Region 1</td>
<td>0</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>UT, Salt Lake City</td>
<td>Utah DoT Region 2</td>
<td>19</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td></td>
<td>43</td>
<td>51</td>
<td>8</td>
</tr>
<tr>
<td>WA, Seattle</td>
<td>Washington State DoT Northwest Region</td>
<td>222</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>WA, Seattle</td>
<td>Washington State DoT - Olympic Region TMC</td>
<td>21</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td></td>
<td>243</td>
<td>219</td>
<td>54</td>
</tr>
<tr>
<td>WI, Janesville</td>
<td>Wisconsin DoT District 1</td>
<td>141</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>WI, Milwaukee</td>
<td>Wisconsin DoT District 1</td>
<td>141</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td></td>
<td>282</td>
<td>195</td>
<td>43</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td></td>
<td>3500</td>
<td>2996</td>
<td>2738</td>
</tr>
</tbody>
</table>
5 Conclusion
In this report ramp metering algorithms and the strategy for implementing them are examined. Therefore, the literature review gives a general overview of the basics of ramp metering. It explains how using a ramp metering approach to handle freeway free flow in comparison with arterial demand works, and which goals it seeks to achieve. In preparation for the following chapters, the different approaches and ideas as to how ramp metering algorithms work are explained. Additionally, some major background information is given which is used in later chapters, but does not necessarily solely concern ramp metering, such as the Kalman filter, which is used in many other fields too.

After this general overview follows a summary of all known ramp metering algorithms that existed at the time this study took place. The different algorithms are sorted by their operating mode, which is either on a local level, or on the coordinated level. The algorithms of the latter mode are divided again into their relationship at the local level, which is either cooperative, competitive or integrated. All algorithms are explained as to how they work, or where they are or have been implemented. If available, a short description of their performance is provided. Then, an overview of implemented ramp metering strategies all over the world is provided, along with an historic explanation of the development of ramp metering algorithms. After this, a map of the world and of the USA is shown, which contains the number of implemented ramp metering systems in the different states and countries.

5.1 Future Research Recommendations
This report is a summary of the basic knowledge of ramp metering. It contains general applications of the algorithms and briefly describes mathematical procedures that are often implemented. To extend the general overview about ramp metering, all known algorithms were gathered and their syntax briefly described, including how they used to work.

Currently, there is only one hint about an algorithm in development, but nevertheless, this summary should be maintained from time to time, since its scope doubled in comparison to the last great compendiums from Bogenberger and May in 1999, or from Zhang in 2001. The research revealed that many recent implementations work with either general well-known algorithms, or regional
project specific algorithms. This may be a sign of an information gap about all possibilities of ramp metering algorithms, and can be dissolved through a current and updated summary of the different strategies. A global review, including states, countries, or DOTs would be useful. These experienced departments can be the basis for other possible research. It seems that newly implemented strategies are based on personal recommendations rather than on a standardized manual. It would be necessary to explain in a universal way how such a ramp metering system should be installed to achieve the highest possible efficiency. Especially in places where the old strategy is replaced by a new one, the DOT tries to handle the situation with the existing infrastructure. It would be necessary to explain the possible different locations of detector stations and comparison of detector devices, which vary a lot with many different advantages and possible implementations, and give standardized recommendations for a completely new ramp metering solution. In fact, there are guidelines for implementing ramp metering systems, but they are not a universal tool that helps the organization to handle even the necessary details because they are mostly general explanations of how ramp metering works.

In addition to the latest overview of ramp metering which reveals the increasing use of ramp metering in the world, the efficiency of ramp metering performing under different environmental boundaries should be researched. Most of the evaluations were made in the US, with an acceleration lane that generally differs of the ones used in other countries. The validity of the reported 20% reduction in travel time resulting from ramp metering is questionable because the drivers’ behavior also tends to change with a longer acceleration lane on test sites, which in itself causes improvements in the merging area.
Works Cited


[34] D. Tsui, Interview about ramp metering in Canada, David.Tsui@ontario.ca, 2014.


V. Punzo, *Interview about Ramp Metering in Italy*, vinpunzo@unina.it, 2014.

H. Dr. Taale, *Interview about ramp metering in the Netherlands*, H.Taale@tudelft.nl, 2014.


C. Venter, *Interview about ramp metering in South Africa*, christo.venter@up.ac.za, 2014.


C. J. Lopez, *Interview about Ramp Metering in Spain*, c.lopez@s3transportation.com, 2014.

B. Holmgren, *Interview about ramp metering in Sweden and about TIME*, bjarne.holmgren@trafikverket.se, 2014.

List of figures

Figure 1 Overview of Ramp Metering Implementations ........................................... 8
Figure 2 Feedforward Control ............................................................................. 16
Figure 3 Feedback Control ................................................................................. 16
Figure 4 Feedforward-Feedback Controller ....................................................... 17
Figure 5 Algorithms Overview and Structure .................................................... 20
Figure 6 Visualization of Traffic-Inflexion Point ................................................ 22
Figure 7 Non-Linear Feedback Control Process ................................................. 27
Figure 8 PRO Flow Chart .................................................................................. 31
Figure 9 Working Mode of the Helper Algorithm ............................................. 35
Figure 10 Visualization of Turning Propabilities ............................................... 38
Figure 11 Bottleneck Operational Flow Chart ................................................... 42
Figure 12 Visualization of SWARM Forecast Operation ................................... 45
Figure 13 Layer Classification of SZM ............................................................... 47
Figure 14 ARMS Operational Flow Chart ......................................................... 52
Figure 15 Maryland Algorithm Operational Flow Chart .................................... 55
Figure 16 Taiwan Algorithm Operational Flow Chart ....................................... 56
Figure 17 ANN Conceptual Operation ............................................................... 56
Figure 18 Fuzzy Process .................................................................................... 57
Figure 19 ACCEZZ Process .............................................................................. 59
Figure 20 Dynamic Metering Control Structure .............................................. 60
Figure 21 General Structure of RMS 2000 ......................................................... 61
Figure 22 Overview of the Interaction of the Modules of RMS 2000 ................. 63
Figure 23 Alpha increment function .................................................................. 67
Figure 24 Overview of Invented Algorithms ..................................................... 72
Figure 25 Amount of Metered On-Ramps in the World .................................... 73
Figure 26 Detailed Summary of Ramp Metering in the USA ............................ 79
List of tables

Table 1 Benefits of Ramp Metering ................................................................. 6
Table 2 Summary of the Working Modes ........................................................... 10
Table 3 Summary of the Grade of Detail ........................................................... 11
Table 4 Summary of the Field Deployment ....................................................... 14
Table 5 Comparison Feedforward-Feedback Control ....................................... 17
Table 6 Itemized Overview: Number of Metered Ramps in the USA ................. 80
Schriftenreihe des Instituts für Verkehrswesen und Raumplanung
ISSN 0172-8326

Heft 1 (1978)
A. Schmuck: Verkehrsplanung und Straßenwesen an der Hochschule der Bundeswehr München (vergriffen)

Heft 2 (1979)
C. Pingel: Die Verbesserung der Wohnqualität städtischer Quartiere durch verkehrliche Maßnahmen und ihre Bewertung - Materialien für die Entwicklung eines Zielsystems als methodischer Ansatz für ein Planungs- und Bewertungskonzept (vergriffen)

Heft 3 (1979)
A. Schmuck, G. Oefner: Investitionsrechnung als Entscheidungshilfe bei Straßenbautechnischen Fragestellungen (vergriffen)

Heft 4 (1980)
A. Schmuck, G. Oefner: Grundlagen zur Entwicklung von Verhaltens- und Managementmodellen für den Straßenoberbau (vergriffen)

Heft 5 (1981)

Heft 6 (1981)
C. Pingel: Zur Methodik der systematischen Vorbereitung von Entscheidungen im Bereich der Straßenerhaltung (vergriffen)

Heft 7 (1981)
M. Löffler: - Weißenburg in Bayern - Eine Studie zur verkehrlichen Situation in der Altstadt und Gedanken zur Verbesserung der Stadtqualität durch verkehrliche Maßnahmen (vergriffen)

Heft 8 (1981)
H. Becker, G. Oefner, H. Rohling, W. Scherer: Beiträge zu Wirtschaftlichkeitsuntersuchungen im Straßenwesen (vergriffen)

H. Becker: Ein gesamtwirtschaftlicher Berechnungsansatz zur Optimierung Straßenbautechnischer Entscheidungen

Heft 10 (1982)
H. Becker, J. Gutschker, G. Oefner, A. Schmuck, G. Schönberger: Wirtschaftlichkeitsuntersuchungen in der Straßenbautechnik (vergriffen)

Heft 11 (1983)

B. Breiter: Energieverbrauch im Straßenbau - Exemplarischer Vergleich der Auswirkungen bautechnischer und entwurfstechnischer Parameter

Heft 12 (1983)

G. Oefner: Ein Beitrag zur modellgestützten Vorbereitung von straßenbautechnischen Entscheidungen (vergriffen)

Heft 13 (1983)

A. Böker, A. Schmuck: Verkehrsstärkenspezifische aktualisierte Baulastträgerkosten - Daten zur Wirtschaftlichkeit von Straßennetzelementen (vergriffen)

Heft 14 (1983)

H. Becker, W. Emde, H. Hamester, A. Schmuck: Verkehrsablauf und Unfallgeschehen an Autobahnbaustellen

Heft 15 (1983)

H. Koch: Füllernachbildung an Sanden durch thermische und mechanische Beanspruchungen bei der Herstellung von bituminösem Mischgut

Heft 16 / 17 / 18 (1984)


Heft 19 (1984)

C. Pingel: Entwurf eines Leitfadens für die visuelle Zustandsaufnahme und Zustandsbewertung auf zweistreifigen Außerortsstraßen mit bituminöser Befestigung

Heft 20 (1984)

G. Schönberger: Entwicklung von Empfehlungen für eine rationale Straßenerhaltung (vergriffen)

Heft 21 (1984)

B. Breiter: Untersuchungen zur Bestimmung optimaler Bauabschnittslängen für Erhaltungsmaßnahmen auf Autobahnen

Heft 22 (1984)

A. Schmuck u. a.: 10 Jahre Verkehrsplanung und Straßenwesen an der Hochschule
der Bundeswehr München (1974 - 1984) - Berichte, wissenschaftliche Beiträge, Dokumentation -

Heft 23 (1984)
M. Löffler: Thermisch bedingtes Längenänderungsverhalten von Stoffen unter besonderer Berücksichtigung von straßenbautechnischen Aspekten

Heft 24 (1986)
B. Breiter: Ein Beitrag zur Optimierung straßenbautechnischer Entscheidungen unter energiewirtschaftlichen Aspekten

Heft 25 (1986)
W. Toepel: Die institutionelle Situation auf europäischen Flughäfen. The Institutional Scene in European Airports

Heft 26 (1987)
F. Hofmann: Ein Beitrag zur Beurteilung der Beanspruchung von Betonfahrbahnplatten unter besonderer Berücksichtigung des Fugenbereiches

Heft 27 (1988)
G. Oefner: Handbuch für die wirtschaftliche Vergleichsrechnung im Management der Straßenerhaltung – Baulastträger- und Nutzerkosten –

Heft 28 (1988)
A. Schmuck: Straßenbaubedarf und Finanzierung des Straßenbaues (vergriffen)

Heft 29 (1988)
Festschrift zum 60. Geburtstag von Univ.-Prof. Dr.-Ing. Alfred Schmuck (vergriffen)

Heft 30 (1989)

Heft 31 (1989)
W. Ressel: Diagramme zur Ermittlung von Straßennutzerkosten

Heft 32 (1991)
B. Steinauer: Stochastische und periodische Unebenheiten auf Fahrbahnen

Heft 33 (1991)
M. Sršen: Das HDM-III-Modell (The Highway Design and Maintenance Standards Model) – Untersuchung zur Übertragbarkeit auf mitteleuropäische Verhältnisse (vergriffen)

Heft 34 (1992)
Vorträge, Seminar- und Arbeitsberichte I

Schriftenreihe IVR, Heft 59
Heft 35 (1993)
A. Schmuck: Kommunaler Straßenbau - Beiträge zum Bau- und Erhaltungsmanagement - (vergriffen)

Heft 36 (1993)
F. Hehenberger: Grundlagen zur Entwicklung von Expertensystemen für das Management der Straßenerhaltung

Heft 37 (1994)
W. Ressel: Untersuchungen zum Verkehrsablauf im Bereich der Leistungsfähigkeit an Baustellen auf Autobahnen (vergriffen)

Heft 38 (1994)
B. Steinauer: Harmonisierung zwischen Befahrbarkeitszustand und Ausbaustandard von Straßen

Heft 39 (1994)
A. Schmuck, G. Krause: Langzeitverhalten von ungebundenen Standardbauweisen der ländlichen Neuordnung in Bayern (vergriffen)

Heft 40 (1994)
Aus Wissenschaft und Praxis - Zum Ausscheiden von Universitätsprofessor Alfred Schmuck

Heft 41 (1996)
E. Kienlein: Funktionale Relativierung von Zustandsgrößen und Zustandsreihung in der Straßenerhaltung (vergriffen)

ISSN 1436-9982

Heft 42 (1999)
Ch. Ferchland: Interaktive Fahrplanerstellung für den intermodalen ÖPV mit Methoden der Netzplantechnik

Heft 43 (2000)
K. Klingberg: Das Straßenzustands- und Wetterinformations-System (SWIS): Einsatzbeispiel Autobahndirektion Nordbayern

Heft 44 (2001)
E. Kienlein, C. Nußrainer, R. Sobotta: Dringlichkeitsreihung von streckenbezogenen Erhaltungsmaßnahmen
Heft 45 (2001)

Heft 46 (2003)
M. Pfeifle: Verkehrsplanung in Stadt und Region Stuttgart - Fazit aus 150 Jahren Planungsgeschichte

Heft 47 (2004)
R. Biel: Verkehrswirtschaftliche Standortuntersuchung

Heft 48 (2006)
W. Wirth: Planfreie Knotenpunkte und Stadtautobahnen: Bausteine für die RAA

Heft 49 (2006)
R. Sobotta: Überprüfung von Entwurfsparametern für Kreisverkehre mit empirischen Schleppkurven

Heft 50 (2007)
A. Opel: Laboranalytische Überprüfung der Rezepturtreue bei der Asphaltherstellung im praktischen Mischanlagenbetrieb

Heft 51 (2007)
U. Häp: Bewertungsverfahren für Planungsvarianten von Start- und Landebahnen bei einem Flughafenausbau

Heft 52 (2007)
S. Stroh: Entwicklung einer Dokumentationssystematik für planfreie Knotenpunkte mit digitaler Lageplandarstellung und Elementdatenbank

Heft 53 (2008)
E. Kienlein, A. Opel: Bauzeitenermittlung für Erhaltungsmaßnahmen zur Minimierung der baustellenbedingten Nutzerkosten für das Erhaltungsmanagement/PMS

Heft 54 (2009)
A. Opel: Herstellung von Niedrigtemperaturasphalt (Walzasphalt) unter Verwendung von Schaumbitumen

Heft 55 (2010)
S. Klein: Bemessung und Bewertung von Passagierabfertigungseinrichtungen in Flughafenterminals
Heft 56 (2011)
V. Spahn: Standardisierte Wirksamkeitsanalyse von sicherheitsverbessernden Maßnahmen an Unfallhäufungen auf Außerortsstraßen

Heft 57 (2013)
S. Çalışkan: Die straßenkategoriebezogenen Geschwindigkeitsvorgaben in den Richtlinien für integrierte Netzgestaltung - Praktische Überprüfung und wirtschaftliche Bewertung

Heft 58 (2015)
R. Seign: Model-Based Design of Free-Floating Carsharing Systems

Heft 59 (2016)
T. Kristeleit, B. Bracher, K. Bogenberger, R. Bertini: Ramp Metering Algorithms and Implementations