Combined rate and rule-based automatic kinetic model generation

Jessica Aernoudt

Supervisor: Prof. dr. ir. Kevin Van Geem
Counsellor: Ir. Ruben Van de Vijver

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Chemical Engineering

UNIVERSITEIT
GENT

Acknowledgments

While this work is coming to an end, I would like to take the opportunity to acknowledge the persons who have been there for me during the progress of this work. They have encouraged me to exceed limits, to persevere and finally to finish this work with proud and fulfillment.

First of all, I would like to thank Prof. dr. ir. Guy B. Marin for giving me the opportunity to finish this master thesis and my education at the LCT. My deepest gratitude goes out to Prof. dr. ir. Kevin M. Van Geem for his valuable guidance and advice concerning this work. I am very proud and fortunate to have had these persons as my teachers.

I would like to thank my counselor Ruben Van de Vijver for his guidance and support during the entire year. Even from a far distance, help was always just around the corner and when I could not make heads or tails of it, you managed to let me see things clear again, by providing me with many insights and even more answers.

My deepest gratitude goes to Hans-Heinrich Carstensen for providing me with clear insights into the field of kinetic modeling. I admire his enthusiasm and broad knowledge about this subject. Also Ruben De Bruycker and Muralikrishna Khandavilli deserve a word of gratitude for their advice and for providing me with the essentials to finish this work.

Finally I would like to thank everyone at the LCT, especially for the nice coffee breaks. Even though I don't like coffee, the jokes and laughter were the main reason to take a daily trip from 918 to 914. My colleague students who I will miss quite a lot, especially four, Jühl, Charlotte, Jens and Alexander. Together we have adapted to an oxygen deficient atmosphere and to highly elevated temperatures.

Natuurlijk zijn er nog mensen buiten het LCT die ik wil bedanken. Mijn ouders en mijn zussen voor het doorstaan van mijn kuren thuis. Steffi om zaken te relativeren waar nodig, om mij wat discipline en moed in te spreken en om nu en dan mijn geweten te spelen. Voor alle steun en om tot rust te komen tijdens dit jaar was mijn lieve Michel er altijd, nu deze thesis ten einde loopt is het hoogtijd om je de aandacht en liefde te geven die je als de allerliefste verdient. En voornamelijk, bedankt pap om niet alleen voor Laika te zorgen maar ook voor mij.

Jessica Aernoudt

22 mei 2015

Combined Rate and Rule-based Automatic Kinetic Model Generation

Jessica Aernoudt

Master's dissertation submitted in order to obtain the academic degree of

Master of Science in Chemical Engineering

Academic year 2014-2015

Promotor: Prof. dr. ir. Kevin M. Van Geem

Coach: ir. Ruben Van de Vijver

GHENT UNIVERSITY

Faculty of Engineering and Architecture

Department of Chemical Engineering and Technical Chemistry

Laboratory for Chemical Technology

Chairman: Prof. dr. ir. Guy B. Marin

Abstract

A rate-based termination criterion is implemented next to the rule-based criterion in the automatic network generation tool called Genesys. Generating a model in a rule-based fashion relies on constraints which are based on experts knowledge and thus require prior experimental research. This means experimental results are rather validated than "predicted" by a kinetic model. The rate-based algorithm generates the possible species and reactions after which the most significant ones are selected based on their rate of formation. However, this technique suffers from memory limitations and is inapplicable if no accurate rate coefficients and thermodynamic properties are available. The aim of this work is to combine the rate and rule-based method in which the constraints are used to limit the number of generated species and the rate-based criteria is used to select the most significant ones. The performance of the implemented algorithm is studied by constructing a kinetic model for the pyrolysis of n-pentanol. *Keywords*: Termination criteria, Rate-based, Rule-based, Automatic network generation

Combined Rate and Rule-based Automatic Kinetic Model Generation

Jessica Aernoudt

Promotor: Prof. dr. ir. Kevin M. Van Geem
Coach: ir. Ruben Van de Vijver

Abstract: A rate-based termination criterion is implemented next to the rule-based criterion in the automatic network generation tool called Genesys. Generating a model in a rule-based fashion relies on constraints which are based on experts knowledge and thus require prior experimental research. This means experimental results are rather validated than "predicted" by a kinetic model. The rate-based algorithm generates the possible species and reactions after which the most significant ones are selected based on their rate of formation. However, this technique suffers from memory limitations and is inapplicable if no accurate rate coefficients and thermodynamic properties are available. The aim of this work is to combine the rate and rule-based method in which the constraints are used to limit the number of generated species and the rate-based criteria is used to select the most significant species. The performance of the implemented algorithm is studied by constructing a kinetic model for the pyrolysis of n-pentanol.

Keywords: Termination criteria, Rate-based, Rule-based, Automatic network generation

I. INTRODUCTION

Complex processes such as combustion, pyrolysis and oxidation are described by kinetic models involving a large number of reactions and species. Because constructing these model by hands requires too much effort, computational methods have been developed to automatically construct a reaction mechanism and to assign thermochemical data to the species and kinetic parameters to the reactions.

The process of generating new reactions is however not straightforward. First of all, many kinetic model builders rely on the concept of reaction families or reaction rules to generate reactions. This is based on the idea that a functional group inside a molecule governs the reactivity rather than the entire molecule. This does not apply to small molecules, hence, the reactions of these small species are added to the automatically generated reactions to obtain a model that is as complete as possible. Secondly, recombination and addition reactions leading to larger molecules could result in an indefinite kinetic model generation, eventually limited by the computational resources. As a consequence, many insignificant species and reactions are added to the reaction network. These species do not or barely appear in physical reality. Therefore finding an adequate termination criterion is essential for the automatic constructing of kinetic models. Ideally, this criterion should be able to stop the generation process at an appropriate time and exclude insignificant species and reactions from the model.

II. LITERATURE STUDY

A. Termination criteria

Several techniques have been developed to prevent the endless generation of species and reactions by kinetic model builders. The first were developed to arbitrarily neglect certain species. The maximum carbon-count criterion is such an example and suggests that species with a carbon atom number larger than a user-defined number are retained from the model. Broadbelt et al.[1] validated a kinetic model builder with this criterion for the pyrolysis of ethane. The available memory limited the number of species and reactions in the model. To overcome the memory limitations, a second criterion was suggested.[2] The rank-based criterion follows the production rank of each species during the network generation. The rank of species is by definition equal to the order in which this species appears in a reaction mechanism. Species with a rank higher than a predefined value are retrained from the network. This approach is based on the idea that the products formed during a later stage in the reaction network are of less significance. The use of both a rank-based and carbon count criterion results in converging kinetic models for the pyrolysis of ethane. However, still to many insignificant species are included in the model and expected species were not generated.

Later on, two new criteria were developed which are based on more logical reasoning. The rule-based criterion suggests that species are generated with certain constraints. These constraints are defined per reaction family. First, the candidate reacting species is scanned to see if the required reactive moiety is present. If so, the second check controls if the molecule complies with the constraints of the reaction family in question. If not, no new reactions and species are generated. This is an *a priori* network reduction methodology implemented in kinetic model builders such as CASB, COMGEN, RING and Genesys. The constraints are based on knowledge of the chemistry under study and therefore prior experimental research is necessary. These constraints can be imposed on a molecular level or at the level of the atom. Examples of these constraints are the molecules charge, size, or structure.

Susnow et al.[3] were the first to implement the rate-based algorithm in a kinetic model builder. The idea is to use dynamic properties of the species to determine if they are of importance or not. The principle can be explained as follows. During the network generation process, a partial network containing all the important reactions and species exists. All possible products originating from the partial network are not yet added, but are listed as edge species and kept aside

instead. As long as there are species present in the model which have not reacted yet, the network is called "partial", since some important reactions might not have been generated Next, the partial network is combined with an appropriate reactor model and numerically solved to obtain the rates of formation of edge species. The edge species with the highest rate is chosen as most significant product and is added to the partial network. From here on the iterative procedure is repeated. This means that reaction mechanisms are iteratively generated and solved. To prevent endless addition of species with the highest flux, two termination criteria are defined. The network generation stops when the rate of all significant edge species is lower than a minimum rate. This minimum rate is the product of a user-defined precision level ε and the characteristic rate *(1)*, which is defined as the L2-norm of rates of production of the species in the partial network*(2)*.

$$R_{\min} = \varepsilon \times R_{char} \qquad (1)$$

$$R_{char} = \sqrt{\sum_i r_i^2(t)} \qquad (2)$$

For the second criteria, the user has to define a goal conversion for the initial reactant. After each iteration the conversion is calculated and when the goal conversion is reached, the network generation is halted. If multiple reactants are present, the average conversion is calculated.

When the precision level is decreased, the minimum rate will be lower and hence more edge species are added to the model. This however also influences the number of generated edge species. Even though these species are not part of the model, the mechanism generator has to keep track of every edge species and calculate their dynamic properties. The rate-based method has several downsides that should be overcome. Not only could the scarcity of accurate rate coefficients result in the selection of wrong edge species during the network generation, the number of generated edge species can become too high when the precision level is too small. A combination of the rule-based and rate-based method should resolve the memory problems while only significant species are added to the network. Limiting the number of edge species that are generated with the constraints of the reaction families is a promising method.

B. Kinetic parameters

A crucial aspect of the rate-based criterion are the available kinetic parameters. Inaccurate parameters render the rate-based termination method inapplicable to generate an adequate kinetic model. Kinetic parameters are obtained either experimentally, computational or via estimation methods. The origin of kinetic data strongly influences its quality and quantity. Experimental studies have always been the primary source of chemical data and are continuously improving. However, they require financial investment and labor which make them scarce. Often these values are only valid for a limited range of operating conditions. For radical processes specifically, the photochemical impact, the discharge flow and chemical shock tubes techniques are used to produce kinetic data. Most of the measurements are obtained with a standard deviation of 10 %. Off course, the applied detection technique influences the accuracy as well. When acquiring information via experiments becomes too difficult, one could rely on computational methods instead. The computational methods

derive thermodynamics and kinetics from the total energy of the system, after solving the Schrödinger wave equation.

Kinetic properties are obtained via the Transition State Theory, where the potential energy of the wells (reactants and products) and the saddle point on the potential energy surface are calculated. With the increasing size of a molecule, the required computational resources increase steeply, which puts a limitation on the calculations. However, high performance computing capacity allows the use of more accurate methods and a kinetic model builder that is linked with computational chemistry packages to calculate kinetics and species properties on-the-fly will become available in the foreseeable future.

Finally, estimation methods to obtain kinetic data based on semi-empirical correlations are also available. For the estimation of thermochemical properties of species such as enthalpy, entropy and heat capacity the Benson group additivity method is an option.[4] The Benson group addivitiy method has proven to predict properties with chemical accuracy, i.e. within 4 kJ mol^{-1}.[5] The same principle is valid to estimate kinetic parameters. Other correlations suggested from empirical results are either based on the general form of the linear free-energy relationship (LFER):

$$\ln(k_i(T)) = \ln(k^{ref}(T)) + m \cdot (x_i - x^{ref}) \qquad (3)$$

Where $k^{ref}(T)$ is the rate of a known reaction belonging to the same reaction family, m is the characteristic of the reaction family, x_i is a property of reaction i or the species in reaction i and x^{ref} corresponds to the reference reaction that defines the family. Examples of these correlations are the Evans-Polanyi[6], Hammet or Taft correlation[7]. Non-linear relationships exist as well. The Marcus equation[8] or the Blower-Masel[9] correlation are examples of these.

To calculate the kinetic parameters of reversible reactions, the thermodynamic consistency is used, meaning that accurate thermodynamic properties are also necessary to obtain accurate rate coefficients.

III. NUMERICAL SOLVERS

There is a wide variety of software tools available with numerical solvers that are able to solve kinetic models. These solvers not only have to deal with large detailed systems, the stiffness of a system remains a challenge as well. The partial network which is solved during each iteration of the rate-based algorithm contains elementary reactions with their corresponding rate coefficients and thermodynamic data of the included species. To simplify the set of differential equations describing the mechanism, abstraction is made of the mass transport limitations. For this work, the software tool CHEMKIN[10] has been chosen to provide the necessary numerical solution techniques. A solver called TWOPNT is used to solve steady-state CSTR models and the solver DASPK to solve transient problems. By default, an isobaric and isothermal homogeneous batch reactor is chosen to numerically solve the partial network, meaning that the energy equation should not be solved.
CHEMKIN provides the user with a vast choice of reactor models and options, however, a license is required. Other software tools such as Cantera[11] and OpenSMOKE[12] require no license but contain less reactor models. The code of Cantera is open-source meaning that the user can implement its own reactor model. A comparison of the three software tools has shown a perfect superposition of the simulated results for the combustion and pyrolysis of ethanol. [13]

IV. RATE-BASED ALGORITHM

During the network generation process in Genesys, the model is iteratively expanded by repeatedly applying the reaction families. The reactions forming the newly generated edge species are implemented as irreversible reactions. This is because the rate of formation of the edge species is required and not their rate of consumption. After the identification of the edge species, the kinetic parameters of the partial network and the thermodynamic properties of the every species are searched for in the provided databases of Genesys and, in case no entry is available in the databases, the values are estimated.

The structure of a generated model is influenced by the user-defined precision level and goal conversion of the rate-based algorithm. When the precision level is lowered, the minimum rate decreases and more species are added to the model. Hence, more edge species and reactions can be generated. A simple model is generated with Genesys for the steam cracking of ethane to study these influences. The constraints, such as the maximum number of carbon atoms of the reactants, are rather loosely defined which results in the generation of insignificant edge species. The goal is to exclude these species from the partial network based on their rate of production. In Figure 1 the number of edge species and reactions towards edge species that remain after the model is complete are plotted against the precision level on a logarithmic scale.
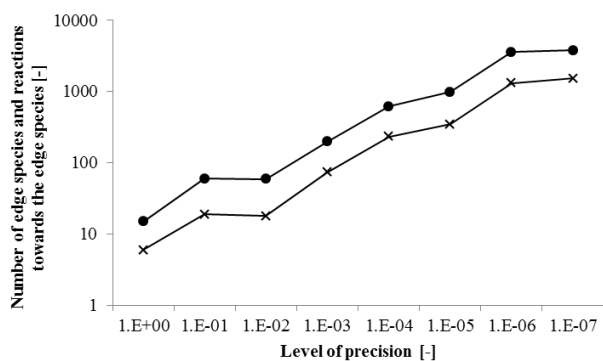

Figure 1 Number of edge species (crosses) and number of reactions towards edge species (dots) as a function of the precision level ε.

This plot shows that the number of edge species depends on the precision level. This number increases exponentially and thus the time required to generate the network increases as well. When there are more edge species present, more thermodynamic properties and rate coefficients have to be calculated and more species have to be checked for a reactive moiety via the graphs isomorphism. Figure 3 illustrates the calculated conversion of ethane during each iteration of the process.
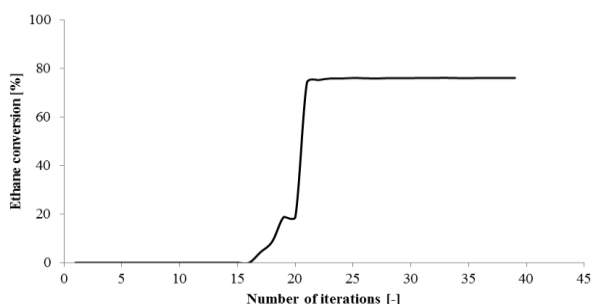

Figure 2 Calculated conversion of ethane per iteration.

The sudden increase after the 20th iteration is caused by the fact that important reactions were added to the model that consume ethane as a reactant. Finally, to emphasize the importance of the reaction family constraints, a model was generated without constraints. The precision level was lowered until Genesys failed to converge. For a precision level of 1E-06 it took approximately 12 hours before memory limitations occurred and if the precision level was further reduced to 1E-07, it took 21hours to exceed the memory capacity. Both simulations had approximately 5500 edge species and 8000 reactions towards the edge species, while the partial network consisted of approximately 80 species and 600 reactions.

V. THERMAL DECOMPOSITION OF N-PENTANOL

To compare the performance of the new implemented rate-based algorithm in Genesys, three kinetic models are constructed for the pyrolysis of n-pentanol. The first model is obtained via the existing rule-based method of Genesys (model I), while the second model is obtained with the combined rate and rule-based method of Genesys (model II). The same reaction families are used as for model I, only the constraints were less tight in order to generate more reactions and species. Eventually, the combination of the rate and rule-based model should result in a detailed kinetic model that can predict the yields of all major products, without running into memory limitations during the network generation process, which is a common issue for the rate-based algorithm. To generate model II the rate-based characteristics as shown in Table 1were used. The last model (model III) is generated with the kinetic model builder RMG[14], which generates models in a purely rate-based fashion.

Table 1 Characteristics of the rate-based criterion.

Precision level ε [-]	0.0001
Conversion n-pentanol [%]	99.99
Reactor type	PFR
Mole fraction n-pentanol	1.0
Pressure [bar]	1.7
Maximum temperature [K]	1046

As mentioned above, the application of reaction families to generate the important reactions is not valid for reactions with small molecules, since the entire molecule instead of a reactive moiety influences the reactivity. To overcome this, reactions and species from the AramcoMech[15] model and initiation reaction, such as important scission reactions, from the work of Heufer et al.[16] were added manually after the network generation. For the generation of model III with RMG, these reactions and species are initially added as seed mechanisms, based on which the further network generation process is continued. In Table 2 the number of reactions and species per model are listed.

Table 2 Number of species and reactions for both models obtained by Genesys and the model obtained by RMG.

	Model I	Model II	Model III
# of generated reactions	2174	2258	4675
# of generated species	320	292	151
# of H-abstraction reactions	1425	1756	1949
# of radical additions	385	285	55
# of radical recombinations	364	217	80
# of extra species added	100	96	129
# of extra reactions added	731	731	731
Total # of reactions	2905	2956	5406

The number of generated radical addition and radical recombination reactions in model II are less than in model I. This could be explained via of the rate-based criterion. If this type of reactions form edge species with a low rate of formation, these edge species are not added to the model and hence these reactions are also not added. In Figure 3 the conversion of pentanol is plotted as function of the temperature in the reactor.
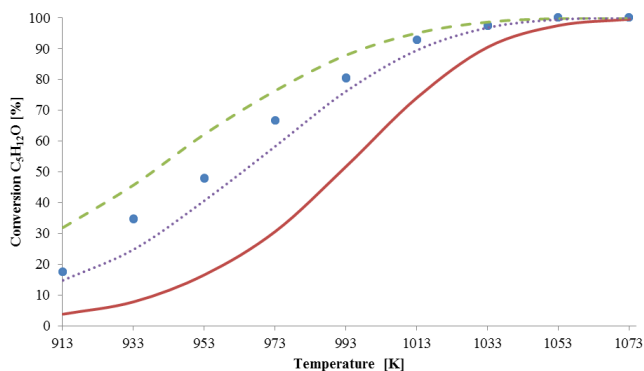

Figure 3 Conversion of pentanol as function of the operating temperature. The red full line corresponds with model I, the purple dotted line with model II and the green dashed line with model III. The dots correspond to experimental data.

Model II simulated the conversion in good agreement with the experimental results. While model I underpredicted the conversion and model III overpredicted the conversion. In Figure 4 the outlet fractions of ethylene and acetaldehyde as function of the conversion of pentanol are depicted. None of the models are able to predict the yield of acetaldehyde, meaning that some important reactions are missing from the models or that the rate coefficients of the important reactions are poorly estimated. It should be mentioned that there is still margin to improve the constraints of model I, based on further experimental research. Looking at the predicted yield of ethylene from model III and the number of reactions and species in model III (Table 2), one might conclude that the generation of the model with RMG was not successful. To obtain more information about the cause of these deviations of the experimental results, a reaction path analysis and a study on the rates of production of important species is necessary.
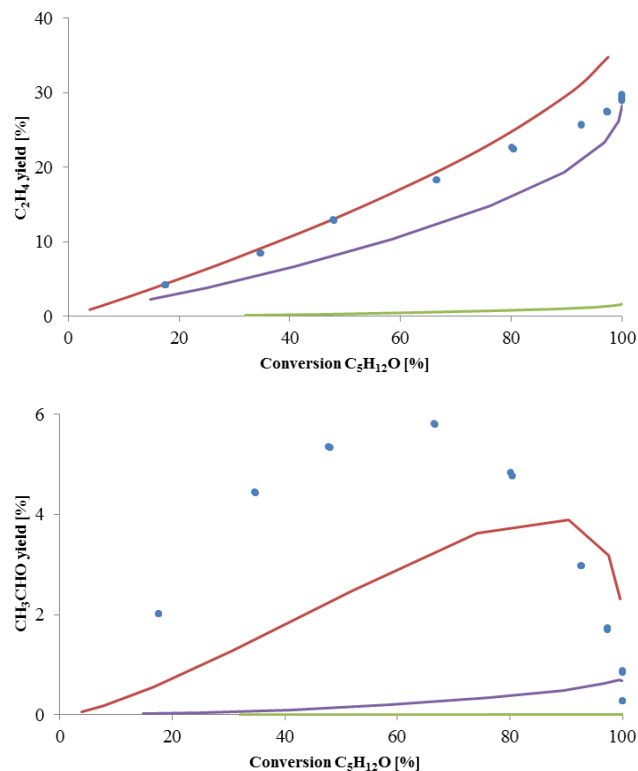

Figure 4 Yield predictions for C_2H_4 (top) and CH_3CHO (bottom) as function of the conversion of pentanol. The red full line corresponds with model I, purple dotted line with model II and the green dashed line with model III. The dots correspond to experimental data.

VI. CONCLUSIONS

A combined rule and rate-based method has been implemented in the code of Genesys in order to stop the network generation process at an appropriate time and to exclude insignificant species from the kinetic model. Typically the rate-based algorithm generates a large number of edge species. The aim was to overcome the memory limitations caused by the size of the edge species list. Constraints imposed on the reacting species of a model prevent the formation of species that are improbable to occur. By applying these constraints with the rate-based method, the number of generated edge species remains within reasonable boundaries. As a result, the user of a kinetic model builder with the combined rate and rule-based algorithm will not be troubled with convergence problems of the software. However, there is still a need for accurate reactions families and constraints even though the latter are now allowed to be less tight. Also, the constraints do not remove the inherent problem of the rate-based algorithm concerning the accuracy of rate coefficients and thermodynamic properties. If the elementary reactions are assigned with incorrect rate coefficients, the rates of formation of the edge species are incorrect as well, and hence species and reactions could be mistakenly added or excluded from the model. The accuracy of availability of kinetic data is none the less improving because of the improved experimental techniques, computational methods and jointly gathering of information by the kinetic modeling community.

REFERENCES

1.   Broadbelt LJ, Stark SM, Klein MT. COMPUTER-GENERATED PYROLYSIS MODELING - ON-THE-FLY GENERATION OF SPECIES, REACTIONS, AND RATES. *Ind. Eng. Chem. Res.* Apr 1994;33(4):790-799.

2.   Broadbelt LJ, Stark SM, Klein MT. TERMINATION OF COMPUTER-GENERATED REACTION-MECHANISMS - SPECIES RANK-BASED CONVERGENCE CRITERION. *Ind. Eng. Chem. Res.* Aug 1995;34(8):2566-2573.

3.   Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. Rate-based construction of kinetic models for complex systems. *J. Phys. Chem. A.* May 1997;101(20):3731-3740.

4.   Benson SW. *Thermochemical kinetics: methods for the estimation of thermochemical data and rate parameters*: Wiley; 1976.

5.   Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J. Phys. Chem. A.* Aug 2005;109(33):7466-7480.

6.   Evans MG, Polanyi M. Inertia and driving force of chemical reactions. *Transactions of the Faraday Society.* 1938;34(0):11-24.

7.   Brown WG. Physical Organic Chemistry (Hammett, L. P.). *Journal of Chemical Education.* 1940/11/01 1940;17(11):551.

8.   Marcus RA. Unimolecular Dissociations and Free Radical Recombination Reactions. *The Journal of Chemical Physics.* 1952;20(3):359-364.

9.   Blowers P, Masel R. Engineering approximations for activation energies in hydrogen transfer reactions. *Aiche J.* 2000;46(10):2041-2052.

10.   Kee RJ, Rupley FM, Miller JA. *Chemkin-II: A Fortran chemical kinetics package for the analysis of gas-phase chemical kinetics*1989. SAND-89-8009; Other: ON: DE90000917 United StatesOther: ON: DE90000917Thu Feb 07 01:43:28 EST 2008NTIS, PC A07/MF A01 - OSTI; GPO Dep.SNL; EDB-89-153716English.

11.   *Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes.* [computer program]. Version 2.1.22014.

12.   Cuoci A, Frassoldati A, Faravelli T, Ranzi E. OpenSMOKE++: An object-oriented framework for the numerical modeling of reactive systems with detailed kinetic mechanisms. *Computer Physics Communications.* (0).

13.   tran L-S. Comparison: Cantera-OpenSMOKE-Chemkin2014.

14.   Song J. *Building robust chemical reaction mechanisms: Next Generation of Automatic model construction software*: Departmen of Chemical Engineering, Massachusetts Institute of Technology; 2004.

15.   Curran HJ. Saudi Aramco Mechanism Release. 2015; http://www.nuigalway.ie/c3/Mechanism_release/frontmatter.html. Accessed 1st May 2015, 2015.

16.   Heufer KA, Sarathy SM, Curran HJ, Davis AC, Westbrook CK, Pitz WJ. Detailed Kinetic Modeling Study of n-Pentanol Oxidation. *Energy & Fuels.* 2012/11/15 2012;26(11):6678-6685.

Table of contents

Nomenclature

Abbreviations

BE	Bond and Electron matrix
B3LYP	Becke, 3-parameter, Lee-Yang-Parr
CASB	Computer Assisted Structure Building
CBS	Chemistry Basis Set
CCM	Computation Cost Minimization
CDK	Chemistry Development Kit
CFD	Computational Fluid Dynamics
COMGEN	Complex Mechanism Generator
CSP	Computational Singular Peturbation method
CSTR	Continuously Stirred Tank Reactor
CVD	Chemical Vapor Deposition
DASPK	Differential Algebraic System Solver Package
DASSL	Differential Algebraic System Solver
DFT	Density Functional Theory
DRG	Directed Relation Graph
FTST	Flexible Transition State Theory
GAV	Group Additivity Value
Genesys	Generation [of reacting] Systems
GRACE	Generalized Reaction Analysis for Creation and Estimation
GUI	Graphical User Interface
HBI	Hydrogen Bond Increment
HF	Hartree-Fock
HMPA	Hexamethyl phosphoric triamide
HPC	High Performance Computers
ILDM	Low Dimensional Manifold method
InChI	International Chemical Identifier

KING	Kinetic Network Generator
KPP	Kinetics PreProcessor
KPPA	Kinetics PreProcessor: Accelerated
MAD	Mean Absolute Deviation
NASA	National Aeronautics and Space Administration
NNI	Non-nearest Neighbor Interaction
NIST	National Institute of Standards and Technology
IUPAC	International Union of Pure and Applied Chemistry
LCT	Laboratory of Chemical Technology
LFER	Linear Free-Energy Relationship
LFP	Laser Flash Photolysis
LIF	Laser Induced Fluoresence
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
PES	Potential Energy Surface
PFR	Plug Flow Reactor
PrIMe	Process Informatics Model
PSR	Perfectly Stirred Reactor
RC-TST	Reaction Class Transition-State Theory
RES	Resonance Correction
RING	Rule Input Network Generator
RMG	Reaction Mechanism Generator
RNG	Reaction Network Generator
RSC	Ring Strain Correction
SMARTS	SMILES Arbitrary Target Specification
SMILES	Simplified Molecular Input Line Entry System
SOL	Structure Oriented Lumping
TS	Transition State
TST	Transition State Theory
TWOPNT	Two Point value boundary solver

VTST	Variational Transition State Theory
QSSA	Quasi-Steady State Approximation
XML	Extended Markup Language

Roman symbols

$\tilde{A}$	Single-event pre-exponential factor	$m^3 mol^{-1} s^{-1}\, or\, s^{-1}$
A	Pre-exponential factor	$m^3 mol^{-1} s^{-1}\, or\, s^{-1}$
C	Vector of molar concentrations	$[mol\, m^{-3}]$
C_i	Polyvalent central atom	$-$
C_p	Constant pressure heat capacity	$J\, K^{-1}$
c_p	Specific constant pressure heat capacity	$J\, mol^{-1} K^{-1}$
E_{system}	Energy of the system	J
E_a	Activation energy	$kJ\, mol^{-1}$
E_0	Evans-Polanyi intrinsic barrier	$kJ\, mol^{-1}$
F_{cent}	Center value of fall-off range	$-$
F_j	Overall sensitivity parameter	$-$
f_n	Rate of production	$mol\, m^{-3} s^{-1}$
G	Iteration matrix of DASSL and DASPK	$-$
H	Enthalpy	$kJ\, mol^{-1}$
h	Planck constant	$6.62\, 10^{-34} J\, s^{-1}$
k	Order in the backward differentiation formula	$-$
k	Rate coefficient	$m^3 mol^{-1} s^{-1}\ or\ s^{-1}$
k_B	Boltzmann constant	$1.38\, 10^{-23} J\, K^{-1}$
K	Equilibrium coefficient	$-$
m	Characteristic of a reaction family	$-$
M_p	Rank of a molecule	$-$
$M_{p,i}$	Rank of a molecule formed by an idiviual reaction	$-$
n_e	Number of single events	$-$
n_{opt}	Number of energetically equivalent optical isomers	$-$
r	Reaction rate	$mol\, m^{-3} s^{-1}$
r	Reaction rate	$kg\, m^{-2} s^{-1}$
R	Net rate of production	$kg\, m^3 s^{-1}$
R	Universal gas constant	$8.314\, J\, mol^{-1} K^{-1}$
R_{char}	Characteristic rate	$mol\, m^{-3} s^{-1}$
R_{min}	Minimum rate	$mol\, m^{-3} s^{-1}$
R_p	Rank of a radical	$-$
$R_{p,i}$	Rank of a radical formed by an individual reaction	$-$

S	Entropy	$J\,K^{-1}mol^{-1}$
T	Temperature	K
t	Time	s
q	Partition function per unit volume	$-$
V	Volume	m^3
$V°$	Molar volume in standard state	$m^3\ mol^{-1}$
X	Conversion	$-$
x_i	Property of the reaction or the species in LFER	$-$
Y	Mass fraction	$-$
z	Distance	m

Greek symbols

α_j	Reaction order of species j	$-$
$\hat{\alpha}$	Constant of DASSL algorithm	$-$
γ	Transfer coefficient for the reaction family	$-$
$\Delta_r G°$	Standard Gibbs energy of a reaction	$kJ\ mol^{-1}$
$\Delta_f H$	Enthalpy of formation	$kJ\ mol^{-1}$
$\Delta_r H$	Reaction enthalpy	$kJ\ mol^{-1}$
Δn	Change in number of moles	$-$
ε	Precision level	$-$
κ	Tunneling correction factor	$-$
λ	Dampening factor	$-$
ρ	Rate of convergence	$-$
ρ_0	Mass density	$kg\ m^{-3}$
σ_{glob}	Total symmetry number	$-$
τ	Residence time	s
υ_j	Stoichiometric coefficient of species j	$-$
ϕ	Solution vector TWOPNT	$-$

Subscripts

av	Average
0	Initial
$-$	Backward
$+$	Forward
ref	Reference

Superscripts

$\ddagger$	Transition state
$'$	Derivative
$°$	Standard

Chapter 1

Introduction

To successfully model a chemical process it is essential to start with the basics. First, a correct kinetic model has to be developed that expresses the dependencies of the rates of chemical transformations on reactor conditions such as pressure and temperature mathematically.[1] A detailed model which has reliable kinetic parameters is the only model that can describe these dependencies accurately. Adequate and accurate kinetic models are indispensable for the design and optimization of chemical applications.[2] A microkinetic model consists of the elementary reactions that occur during the process, reaction rates coefficients and thermodynamic and transport data of the species. Elementary reactions are the basis for understanding how and at what rate the complex chemical reactions occur.[3]

Manually constructing a kinetic model begins with the selection of the most important species, which could be products, reactants or important intermediates that are essential to predict the production rates of major experimentally observed products. The reactions that can occur between these species are then specified with the appropriate thermochemical data.[4] For smaller mechanisms this method is still manageable although for large mechanisms automatic network generation is favored.[5] The automatic generation of reactions networks is a relative new field in kinetic modeling and uses computational techniques to construct kinetic models. The generation of the elementary reactions is based upon chemical and physical principles and translating these principles into useable algorithms is an art in itself. There are many examples of generation codes that have been successfully used to construct reliable kinetic models.[6-15] All of them have common features which will be described briefly in the next section.

The second part of modeling a chemical process is numerically solving the developed kinetic model with a corresponding set of reactor model equations.[16] If possible, the kinetic model should be large enough to capture all the essential chemistry and in the meantime small enough to obtain a relatively fast simulation. Generating a model that contains thousands of species and even more reactions remains the most challenging task in the chemical kinetics. Not only is required information such as kinetic and thermochemical data scarce, even if they would be available, a model of this size will result in a large set of differential equations. Moreover, this set of differential equations could involve a large difference in time scales resulting in a stiff set of differential equations. Hence, numerically solving these systems in combination with a reactor model takes a lot of effort. In Figure 1-1, the size of models for pyrolysis,

oxidation and combustion processes is illustrated. Lu and Law noticed some correlations for these models.[17] The number of species and reactions of a model increases with the size of the molecule, roughly in an exponential trend, while the number of reactions increases linearly with the number of species in the model (Figure 1-1). Pierucci found that the number of generated reactions is correlated to the number of generated species and, as a rule of thumb, that each newly added species to the model generates four new reactions.[18]
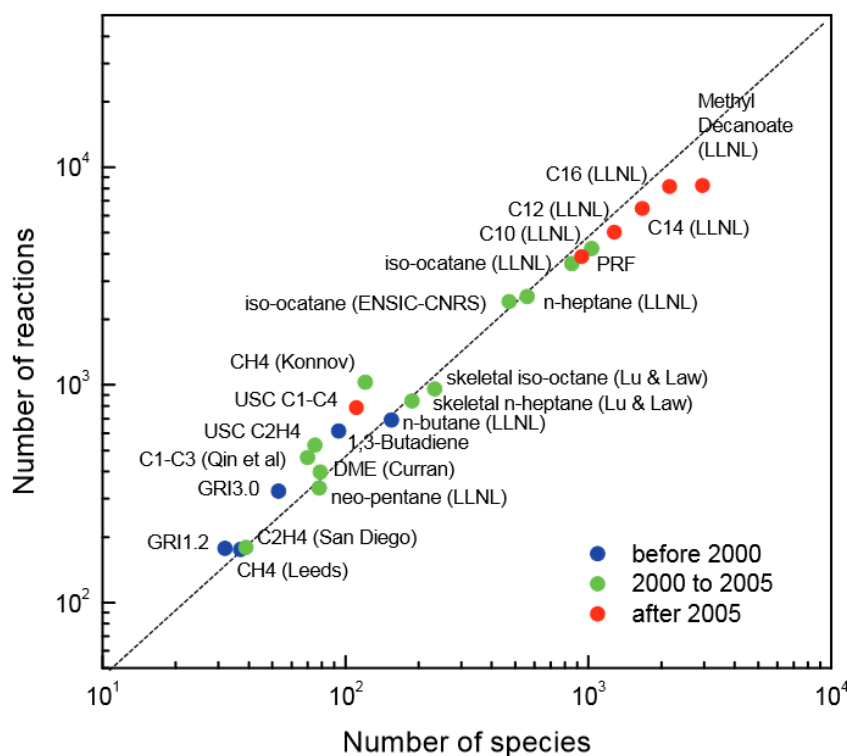

Figure 1-1: Size of various C_1-C_8 hydrocarbon reaction mechanisms superimposed with colors indicating the approximate time when the mechanisms were compiled. [17]

1.1. Common principles of automated network generation

Automatic kinetic model generation is a recursive procedure that is divided into two major sections. First, the construction of the reaction network is performed. Second, the thermodynamic and kinetic properties are assigned. To construct an adequate model, a kinetic model builder should be able to generate the important reactions and species that can occur in a reaction mechanism. An crucial aspect of the automatic network generation is that molecules have to be represented in a unique and unambiguous way. Therefore, an abstract graph representation of molecules and reactions is commonly applied. If the software is unable to recognize identical structures this may lead to unnecessary calculations at best or to infinite loops of calculations.[1] When each molecule has its unique representation, it is straightforward to determine via algorithms if a newly generated species is equal to one of the existing species. The graph isomorphism algorithm compares the graphs of every molecule to each other. The representation of species will be explained later more specifically for the code Genesys in section 1.2.1. For the generation of reactions, the software applies reaction rules or reaction families. The reaction family concept is based on the idea that functional groups inside a chemical species govern the reactivity of a molecule rather than the entire molecule.[2] The effect of functional groups on the reactive center diminishes rapidly with the distance. To search these functional groups in the reactive species, the reaction pattern is translated into a graph representation and the graph isomorphism is used to search for this representation in the graphs of molecules. For reactions involving small molecules, such as those with less than three atoms, the idea that only functional groups influence the reactivity of the entire molecule is not valid anymore. Therefore these reactions should be treated separately. Most kinetic model builders start with a small reaction network as a reaction base containing the small species and the corresponding important reactions. These networks are called seed mechanisms and are eventually enlarged during the generation process.

Generating all possible reactions and species based on reaction rules is not straightforward, since many reactions lead to larger molecules and iteratively continuing the procedure would lead to an indefinite kinetic model generation, which would eventually cause convergence problems. Also, several of these species are insignificant because they are almost not formed in reality. The software should have a proper termination criterion to stop the generation process and to exclude insignificant species and reactions from the reaction mechanism. As mentioned before this remains a challenging task for the automatic network generation process.[3,4] The first termination criteria were developed to arbitrarily neglect certain species based on their production rank or on the amount of carbon atoms of species. These criteria are referred to as the carbon-count and rank criterion. The rule-based and rate-based approach are two methods that rely on more logical reasoning. The rule-based method generates only species that are significant by defining a certain number of constraints. The rate-based approach generates any species

possible starting from an initial set of reactants. Before these new species are added to the reaction mechanism, the rate of formation of these species is calculated. Only the ones with the highest rate are deemed as significant and are added to the model.

Kinetic model builders do not only have a generator engine, they also have other features in common. These features are required to assign the thermodynamic and kinetic properties during the second stage of process. Among others a large set of databases and algorithms to estimate the properties are also required. It is not the aim of this chapter to discuss every aspect of the network generation process in detail. Instead, the features that are used specifically in the code Genesys will be described in the next section. The advantages of automated kinetic model builders is that the time consuming and error prone task of producing every elementary reaction and species is done in a systematic way. It is still the modelers task to determine the important reaction paths.

1.2. Genesys

Genesys is a kinetic model builder developed at the Laboratory of Chemical Technology (LCT) and is originally intended to generate kinetic models for gas-phase free radical chemistry processes. It is written in the programming language JAVA and is based on several algorithms and functionalities. Genesys can be seen as an ecosystem of many independent modules as illustrated in Figure 1-2. Each of these modules contains an algorithms or a set of algorithms to perform a specific function.
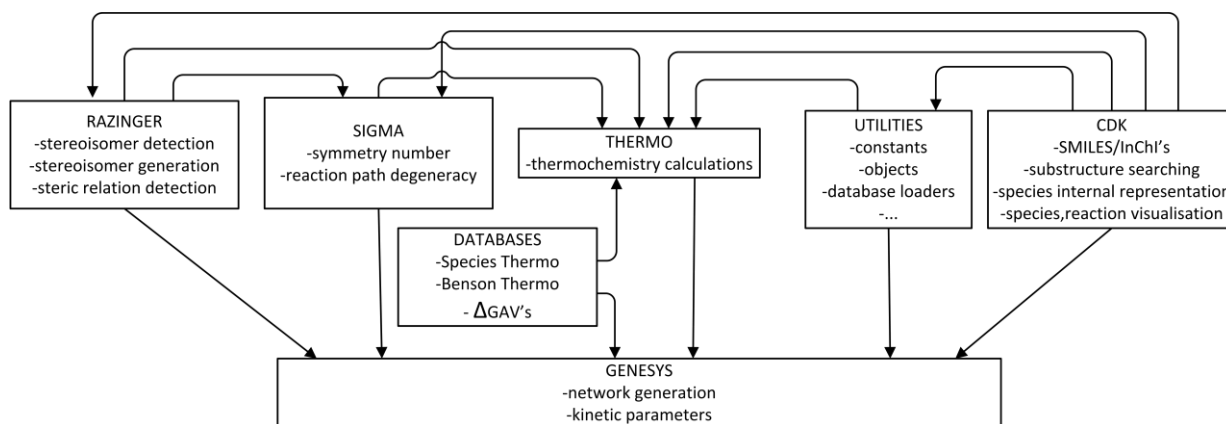
Figure 1-2: Different independent modules present in the kinetic model builder Genesys.

In the THERMO module, the thermochemical properties of species are either looked for in the available databases, and if these are not present, the properties are estimated. Symmetry numbers and number of single events are calculated in the SIGMA module. In RAZINGER, the stereo-center of a molecule is detected and the stereoisomers are generated. And finally, the UTILITIES module is designed to centralize methods and objects that are utilized by the different modules. For example, the values of various constants can be found here. The main advantage of Genesys is that trough the integration of open-source chemo-informatics it is not constrained to specific chemical elements or to specific chemistries.[5] For this purpose the Chemistry Development Kit (CDK)[6] library is used.

1.2.1. Representation of species

When species are generated it is important to determine if these species and the matching reactions are truly new and if they are not already present in the reaction network. Also, for the detection of functional groups in reactive species, it is necessary to have a sub-graph recognition algorithm. In both cases, the graph isomorphism can supply an answer. To successfully use the graph isomorphism algorithms it is fundamental to give each molecule a correct graph representation and a unique identifier.

For this purpose the Simplified Molecular Input Line Entry System (SMILES)[7] and the International Chemical Identifier (InChI)[8] are used in Genesys. With SMILES, graphical molecular information can easily be entered into a computer. This identifier is human readable and can be used to specify the

structure of a molecule. The identifier can contain information about branching, bonds, cyclic structures and atom valance. However, this notation is not-canonical, and thus the numbering of atoms in a molecule is not unique. The more complex InChI language describes chemical structures as hierarchical layers of information, including the atoms and their bonds, isomer information, isotope information, stereochemistry, and electronic charge information. The layers are separated by a slash character. In Figure 1-3 an example of the SMILES and InChI identifiers for two isomers is given. The order of bonds is equal for both molecules but the stereochemistry differs. With an InChI identifier this difference can be noted whereas the SMILES identifier is equal for both molecules.

C/C([H])=C([H])\[C@](O)([H])[CH2]

InChI=1S/C5H9O/c1-3-4-5(2)6/h3-6H,2H2,1H3/b4-3-/t5-/m1/s1

C/C([H])=C([H])\[C@](O)([H])[CH2]

InChI=1S/C5H9O/c1-3-4-5(2)6/h3-6H,2H2,1H3/b4-3-/t5-/m0/s1
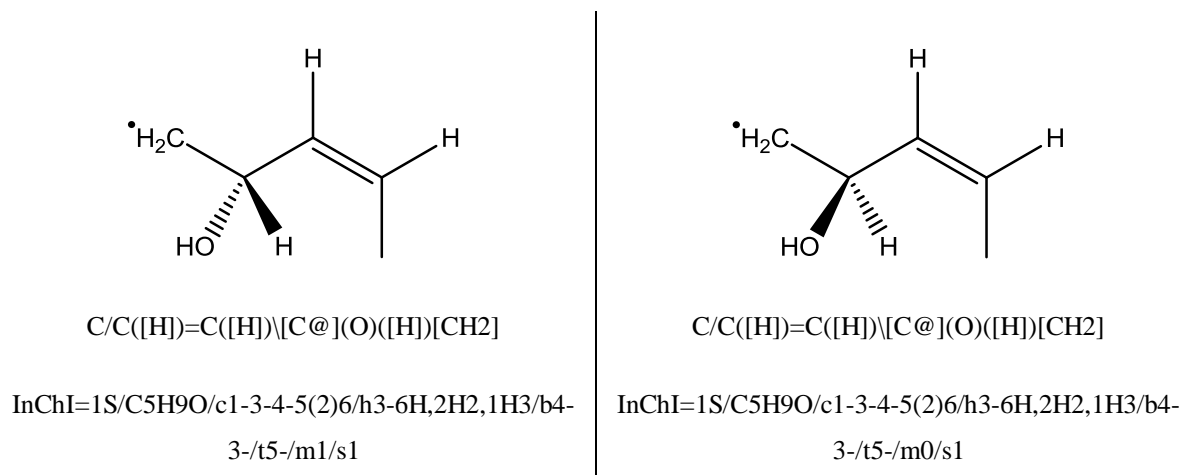
Figure 1-3: SMILES and InChI identifier for two isomers of the C_5H_9O radical.

Algorithms provided by CDK can easily convert InChI's or SMILES into the internal representation and vice versa. Besides this, algorithms for graph isomorphism, automorphism and property identifications are also available.

1.2.2. Reaction families

During the network generation process, new species are formed based on the reaction family description provided by the user. These reactions families are templates written in the Extended Markup Language (XML) language. An example of such a template for the homolytic substitution reaction of a sulfur radical on a sulfur atom is depicted in Figure 1-4. The template should include 1) a description of the reactive moieties required inside potential reactant molecules, 2) a recipe like scheme, where keywords specifying the type of transformations are combined with symbols representing the atoms of the reactive moiety and 3) information on how kinetic parameters of a reaction belonging to this reaction family are calculated. The defined temperature at the beginning of the XML file is used to calculate the tunneling coefficients and the reaction enthalpies used in the kinetic expressions such as the Evans Polanyi relations.

```
<config>
    <inp-temperature>1000</inp-temperature>
    <inp-reaction-family
        name="homolytic substitution reaction of a sulfur centered radical on a sulfur atom">
        <bimolecular>true</bimolecular>
        <inp-recipe>
            <inp-transformation type="FORM_BOND" centers="A,B" />
            <inp-transformation type="BREAK_BOND" centers="B,C" />
            <inp-transformation type="LOSE_RADICAL" centers="A" />
            <inp-transformation type="GAIN_RADICAL" centers="C" />
        </inp-recipe>

        <inp-reactant value="1" smarts="[S;v1]">
            <inp-reactive-center symbol="A" smarts="[S;v1]" />
            <inp-molecule-constraint type="ATOMCOUNT"
                limit="MAX" param="C" value="0" />
            <inp-molecule-constraint type="ATOMCOUNT"
                limit="MAX" param="S" value="1" />
        </inp-reactant>

        <inp-reactant value="2" smarts="[S;X2v2]-[#1]">
            <inp-reactive-center symbol="B" smarts="[$([S;X2v2]-[#1])]" />
            <inp-reactive-center symbol="C" smarts="[$([#1]-[S;X2v2])]" />
            <inp-molecule-constraint type="SINGLEELECTRONCOUNT"
                limit="MAX" value="0" />
            <inp-molecule-constraint type="ATOMCOUNT"
                limit="MAX" param="S" value="1" />
        </inp-reactant>

        <inp-kinetics type="REVERSE">
        </inp-kinetics>
    </inp-reaction-family>
</config>
```

Figure 1-4: Example of a reaction family template, specifically for the homolytic substitution reaction of a sulfur centered radical on a sulfur atom.

The `<inp-recipe>` node contains all the transformations that occur during the specific reaction. These transformations comprise of a set of atom, bond and electronic changes. Inside the `<inp-reactant>` node of the template, a SMARTS language is used to specify the structure of the reactive moiety. By making the SMART string more specific, certain atoms can be excluded from the reactive moiety. SMARTS is a language with rules that are straightforward extensions of SMILES.[9] The reactant node also contains the molecular constraints which are listed under a set of predefined keywords. In the above example, the first reactant contains maximum one sulfur atom and no carbon atoms. The possible molecular constraints and their corresponding keywords are listed in Table 1-1.

Table 1-1: Molecular constraints with their description and corresponding keyword.

Description	Keyword
The number of unpaired electrons of the molecule	SINGLEELECTRONCOUNT
The number of atoms of a given chemical element	ATOMCOUNT
The number of double bonds present in the molecule	DOUBLEBONDCOUNT
The number of elements in the smallest set of smallest rings	SSRINGCOUNT
The number of aromatic atoms	AROMATICATOMSCOUNT

The molecular constraints are verified before the species is scanned for a reactive moiety. If a species does not comply with the constraints, no new reactions of that reaction family are generated. This is the *a priori* network reduction methodology based on chemical principles, that is implemented in Genesys and is also referred to as the rule-based method. The other termination criteria are described in chapter 2 of this master thesis.

Finally, the node `<inp-kinetics>` contains information about the method that is used to obtain the kinetic parameters of elementary reactions belonging to the reaction family. In Table 1-2 a list is given of the possible methods in Genesys.

Table 1-2: Available methods for the assignment of kinetic parameters for elementary reactions with their corresponding keywords.

Method	Keyword
Arrhenius expression	ARRHENIUS
Evans-Polanyi correlation	EVANS_POLANYI
Blowers-Masel correlation	BLOWERS_MASEL
Group additivity	GROUP_ADDITIVITY
"Reverse" kinetics	REVERSE

For the homolytic substitution of a sulfur radical on a sulfur atom the node contains the type "REVERSE" which means that the kinetics are calculated based on the kinetic parameters of the corresponding forward reaction and the thermodynamic consistency. This approach is used to prevent the addition of duplicate reactions in the reaction network, since elementary reactions are considered as reversible and two reaction families can result in the two reactions which are the reverse of each other. The new species generated by a "REVERSE" reaction family are added to the reaction network because they might serve as reactants for other important reactions that are not yet added to the network. Their reactions are added in a later stage when the forward reactions are generated.

1.2.3. Assignment of the thermochemistry

After the network generation process is complete, the THERMO module (Figure 1-2) will assign thermochemical properties to every species present in the reaction network. These properties include the entropy, enthalpy and heat capacity. First it is checked whether these properties are present in the available databases. If not, an estimation is performed via the Benson group additivity scheme for ideal gas phase thermochemical properties.[10-12] This scheme is based on the principle that every group inside a molecule contributes to the thermodynamic properties. These groups are the central, non-terminal, atoms and their nearest neighbors. First, the resonance structures of these species are generated. The thermochemical properties of every resonance structure is obtained after which the structure with the lowest standard enthalpy of formation at the designated temperature is taken as the thermochemistry of

the species. In total there are four types of Benson contributions, the group additive values (GAV) for atom centered groups that take into account the nature of the nearest neighbor ligands, ring strain corrections (RSC) who account for the presence of rings compared to the acyclic counterparts. Non-nearest neighbor interactions (NNI) model interactions between non-bonded sub-molecular fragments and resonance corrections (RES) account for the stabilization effect of electron interactions in a molecule. These contributions can be obtained via experiments or *ab initio* calculations. In addition to the Benson group estimation procedure, the hydrogen atom bond increment (HBI) of Lay et al.[13] is implemented to estimate thermochemical properties of radical species. The contributions are described by either the SMARTS or the Benson notation[12,14] and are also listed in the provided databases. The algorithm used to obtain the contributions from the databases is illustrated in Figure 1-5. The SMARTS description is taken from a database and a matching query is performed on the non-hydrogen atoms of the molecule. If an atom matches the SMARTS description, it will be flagged as "visited" and the GAV will be stored.
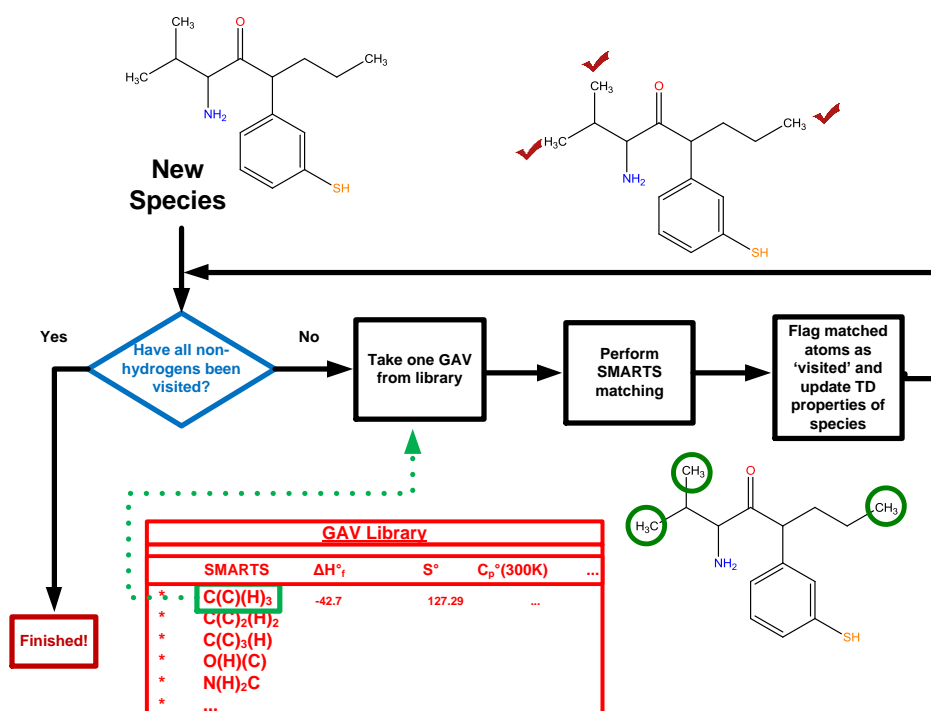
Figure 1-5: Algorithm for the assignment of Benson GAVs to functional groups in a species.[15]

The total entropy of a species is dependent on the symmetry of this species. Because the symmetry is a non-local property of the species, it is not accounted for in the group estimation method. The modules SIGMA and RAZINGER calculate respectively the total symmetry number σ_{glob} and the number of energetically equivalent optical isomers n_{opt} based on the concept of graph automorphisms and stereoisomer enumeration. The assigned thermochemical properties are stored as NASA polynomials in the output of Genesys. These polynomials (Equations 1-1 till 1-3) [16] are easily applied in further simulations.

$$\frac{C_p}{R} = a_1 + a_2 \cdot T + a_3 \cdot T^2 + a_4 \cdot T^3 + a_5 \cdot T^4 \tag{1-1}$$

$$\frac{H}{R \cdot T} = a_1 + \frac{a_2 \cdot T}{2} + \frac{a_3 \cdot T^2}{3} + \frac{a_4 \cdot T^3}{4} + \frac{a_5 \cdot T^4}{5} + \frac{a_6}{T} \tag{1-2}$$

$$\frac{S}{R} = a_1 \cdot \ln(T) + a_2 \cdot T + \frac{a_3 \cdot T^2}{2} + \frac{a_4 \cdot T^3}{3} + \frac{a_5 \cdot T^4}{4} + a_7 \tag{1-3}$$

1.2.4. Assignment of the kinetic parameters

Finally, the kinetic parameters of each elementary reaction present in the reaction network are assigned based on the method defined in the reaction family template (Table 1-2). Ideally, if a library with elementary reactions and their corresponding Arrhenius parameters would be available, the assignment would be straightforward. Unfortunately, for complex reaction networks there is not enough information available to do so and estimation techniques are applied. The Evans-Polanyi [17] and Blowers-Masel [18] are two correlations that are available in Genesys, given that the necessary parameters are provided within the reaction family templates. A group additive method similar to the group additive method for thermodynamic properties is used as well and has proven to give reliable results. [19,20] First a reference reaction with known kinetic parameters is defined. The transition state of the reaction is split up in groups consisting of a single central atom. Every group has a contribution to the different kinetic parameters which is taken into account. There can also be secondary contributions that are present by the influence of groups outside the reactive moiety of the reaction. The kinetic parameters of the entire reaction can be written as a perturbation of the reference reaction, cf. Eq. 1-4 and Eq.1-5.

$$E_a = E_{a,ref} + \sum_i \Delta\Delta GAV_{Ea}^0(C_i) \tag{1-4}$$

$$log\tilde{A} = log\tilde{A}_{ref} + \sum_i \Delta\Delta GAV_{\tilde{A}}^0(C_i) + \log n_e \tag{1-5}$$

Where C_i is the polyvalent central atom of the transition state. The factor n_e is the number of single events which accounts for the number of energetically equivalent reaction pathways. This is a function of the symmetry numbers of the reactants and the transition state and their number of optical isomers. This kinetic group additivity method, together with the Evans-Polanyi and Blower-Masel correlation are described in chapter 3 of this work.

Which method for the calculation of the kinetics is chosen is dependent on the information available for this specific reaction family. For example, for a homolytic scission of a single carbon–carbon bond or for the addition and β-scission of sulfur compounds and for isomerization reactions there are no accurate kinetic group additive values available in literature. There are several ways to assign kinetic parameters in Genesys. The most simple method is presumable to assign one pre-exponential factor A and one activation energy Ea for every elementary reaction of a single reaction family. The number of single events of each reaction is calculated in Genesys and is accounted for in the pre-exponential factor.

1.2.5. Illustration of models generated with Genesys

Next, a few examples are listed which illustrate the performance of Genesys to construct reaction mechanisms and to estimate the kinetic and thermodynamic properties. Nick Vandewiele constructed with Genesys a reaction network for the thermal decomposition of hexamethyl phosphoric triamide (HMPA)[15], a cocking inhibiting additive used in steam cracking applications. The (limited) reaction network was generated with 11 reaction families which were considered relevant for the decomposition of HMPA. The generation of this network illustrated the advantage of the integration of CDK in Genesys, since most of the generated species contained hetero-elements such as phosphor, nitrogen and oxygen.

Ruben Van de Vijver illustrated the use of Genesys for the thermal decomposition of alkylsulfide such as diethyl sulfide and ethyl methyl sulfide.[21] A detailed microkinetic model was generated consisting of 444 reactions and 66 species. The thermochemical properties of molecules and rate coefficients of elementary reactions were estimated through group additivity methods, with parameters obtained from high level *ab initio* calculations. The validation of the simulated model results against experimental data showed overall a promising agreement.

Khandavilli performed a kinetic modeling study of steam cracking of gaseous feeds. Hereby a kinetic model was constructed with different kinetic model builders.[22] The model build with Genesys simulated the yield of major products compared to the experimental results relatively well. In chapter 3 of this work this study is described more into detail.

1.3. Problem formulation

Automatic network generation technology is used to construct reliable and sizable kinetic models for chemical processes involving a large number of species and reactions. The kinetic models provide information on the molecules and reactions involved in a chemical process. To obtain kinetic models with a reasonable size and to end the network generation process there are several termination criteria available. The carbon-count and rank criterion are termination criteria that arbitrarily neglect certain species and therefore result in inadequate kinetic models. The obtained kinetic models were smaller and the generation process converged however most of the important species and reactions were not included. Kinetic model builders that used these criteria have replaced them with either the more accurate rule-based or rate-based criterion.

For the rule-based method the user has to define a certain amount of constraints. These constraints are solely based on knowledge about the chemistry under study. In order to find which species are significant, there is prior experimental research necessary. Thus, the generated kinetic models based on this approach are not able to "predict" but rather to validate experimental results.

The rate-based approach does not rely on prior knowledge of the chemistry but is more computational intensive. The only crucial information that is necessary is highly accurate kinetic data. Another downside is the exponentially growing edge during the network generation. A possibility is to limit the size of this edge with rule-based constraints.

The aim of this master thesis is to implement an algorithm in the code of Genesys that combines the rate-based and (existing) rule-based termination criterion for the automatic network generation. A kinetic model will be generated using the novel approach and validated against experimental data on the thermal decomposition of pentanol. The new algorithm will be compared to existing algorithms found in the literature.

1.4. Outline

The first step in this work consists of a literature study about the different termination criteria used by kinetic model builders. The carbon-count and rank-based criterion as well as the more common rate-based an rule-based termination criteria are studied. Examples of different kinetic model builders using these termination criteria to construct models are given as well. Not only are reactions and species prevented from a model during the network generation process, afterwards reduction methods are occasionally applied to remove unimportant species and reactions, preferably, without losing any essential information.

As mentioned, the rate-based algorithm excludes species from a kinetic model based on their rate of formation. To calculate the rates correctly, accurate kinetic parameters are required. The literature study continues with a description of various sources of kinetic parameters. Either kinetic parameters are obtained experimentally, computational or via estimation methods. Also, the evaluation of kinetic data and an overview of existing databases containing kinetic data from different sources is studied in chapter 3 of this work.

Implementing the rate-based algorithm into the code of Genesys requires the use of a numerical solver. For this there are many options available. In chapter 4 the different software tools and algorithms to solve a set of differential equations or algebraic equations are described. The implementation of the algorithm is further elucidated in the fifth chapter. The characteristics of the rate-based algorithm are discussed and their influence on the generation of a small, simplified model for the pyrolysis of ethane is studied.

To investigate the performance of Genesys with the new algorithm, a case study on the thermal decomposition of n-pentanol is performed. A model is generated and validated against experimental results. The results of this study are provided in chapter 6. The final conclusions and future work are listed in the last chapter of this work.

.

1.5. References

1. Pierucci S, Ranzi E. A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Comput. Chem. Eng.* Apr-May 2008;32(4-5):805-826.

2. Battin-Leclerc F, Blurock E, Bounaceur R, et al. Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models. *Chem. Soc. Rev.* 2011;40(9):4762-4782.

3. Van de Vijver R, Vandewiele NM, Bhoorasingh PL, et al. Automatic mechanism and kinetic model generation: a perspective on best practices, recent advances, and future challenges

4. Green Jr WH. Predictive Kinetics: A New Approach for the 21st Century. In: Guy BM, ed. *Advances in Chemical Engineering.* Vol Volume 32: Academic Press; 2007:1-313.

5. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chem. Eng. J.* Oct 2012;207:526-538.

6. Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen E. The Chemistry Development Kit (CDK): An Open-Source Java Library for Chemo- and Bioinformatics. *ChemInform.* 2003;34(21).

7. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 1988/02/01 1988;28(1):31-36.

8. Stein SE, Heller SR, Tchekhovski D. An Open Standard for Chemical Structure Representation - The IUPAC Chemical Identifier. *Nimes International Chemical Information Conference Proceedings*2003:131-143.

9. Daylight Theory Manual.

10. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J. Phys. Chem. A.* Aug 2005;109(33):7466-7480.

11. Benson SW. *Thermochemical Kinetics, John Wiley & Sons, New York.* 1976.

12. Benson SW. *Thermochemical kinetics: methods for the estimation of thermochemical data and rate parameters*: Wiley; 1976.

13. Lay TH, Bozzelli JW, Dean AM, Ritter ER. Hydrogen Atom Bond Increments for Calculation of Thermodynamic Properties of Hydrocarbon Radical Species. *The Journal of Physical Chemistry.* 1995/09/01 1995;99(39):14514-14527.

14. Benson SW, Cruickshank FR, Golden DM, et al. Additivity rules for the estimation of thermochemical properties. *Chemical Reviews.* 1969/06/01 1969;69(3):279-324.

15. Vandewiele NMU. *Kinetic model construction using chemoinformatics*: Universiteit Gent; 2013.

16. 2013; http://combustion.berkeley.edu/gri_mech/data/nasa_plnm.html. Accessed May, 1st, 2015, 2015.

17. Evans MG, Polanyi M. Inertia and driving force of chemical reactions. *Transactions of the Faraday Society.* 1938;34(0):11-24.

18. Blowers P, Masel R. Engineering approximations for activation energies in hydrogen transfer reactions. *Aiche J.* 2000;46(10):2041-2052.

19. Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Ab initio group contribution method for activation energies for radical additions. *Aiche J.* 2004;50(2):426-444.

20. Saeys M, Reyniers MF, Van Speybroeck V, Waroquier M, Marin GB. Ab initio group contribution method for activation energies of hydrogen abstraction reactions. *ChemPhysChem.* Jan 2006;7(1):188-199.

21. Van de Vijver R, Vandewiele NM, Vandeputte AG, et al. Rule-based ab initio kinetic model for alkyl sulfide pyrolysis. *Chem. Eng. J.* (0).

22. Khandavilli M. An experimental and group additive kinetic modeling study of steam cracking of gaseous feeds2015.

Chapter 2
Termination criteria

2.1. Introduction

When reactions occur that allow the species inside a mechanism to grow, via for example radical addition or recombination reactions, the generation of a network might diverge as endless numbers of new products can be formed, which in turn can react to even bigger species.

To prevent the infinite generation of species by kinetic model builders, several techniques have been developed. In a first attempt to limit the size of a reaction network, there were limitations introduced on the generated species and on the amount of occurring reactions. These limitations were chosen rather arbitrarily and did not take into account the underlying chemistry of the process. Several important pathways and species were neglected, even though the reaction network was of a reasonable size. Other criteria were suggested and implementing these in the existing kinetic model builders showed some promising results.

The main focus of this chapter is to review the several termination criteria that have been implemented in existing automatic kinetic model generation programs. Through understanding their strengths and weaknesses, a comparison can be made. From experience it was found that the first automatically generated kinetic models contained only a small portion of important species and reactions. Trough mechanism reduction methods these models were made smaller.

Application of any convergence criterion is unnecessary for inherently convergent reaction mechanisms. If there is no mechanism for molecular weight growth, the reaction network is convergent. When the model contains radical reactions, the reaction network cannot converge and will grow endlessly. Ideally, all important species are included and all insignificant species are excluded from the reaction network. In certain cases such as oxidation and pyrolysis chemistry, the number of chemically important species can be of the order of $O(10^4)$ because of the non-specificity of the free-radical intermediates[1], making it more desired to leave out the unnecessary species.[2] The determination of only the important species and reactions in a growing mechanism provides a detailed kinetic model.[3]

2.2. Carbon count termination criterion

Broadbelt et al.[1] developed a kinetic model builder inspired on the work of Ugi et al.[2] to implement the graph theory description of molecules in a matrix representation of species and reactions. The kinetic model builder requires three input files: (1) the structure of the reactants, (2) the reaction rules by which the reactants and product species react, and (3) the parameters of a structure/reactivity relation that would provide an estimate of reaction rate coefficients. With this information, the builder provides a reaction network, species properties and rate coefficients. In Figure 2-1 the algorithm that generates the reaction mechanism is illustrated.
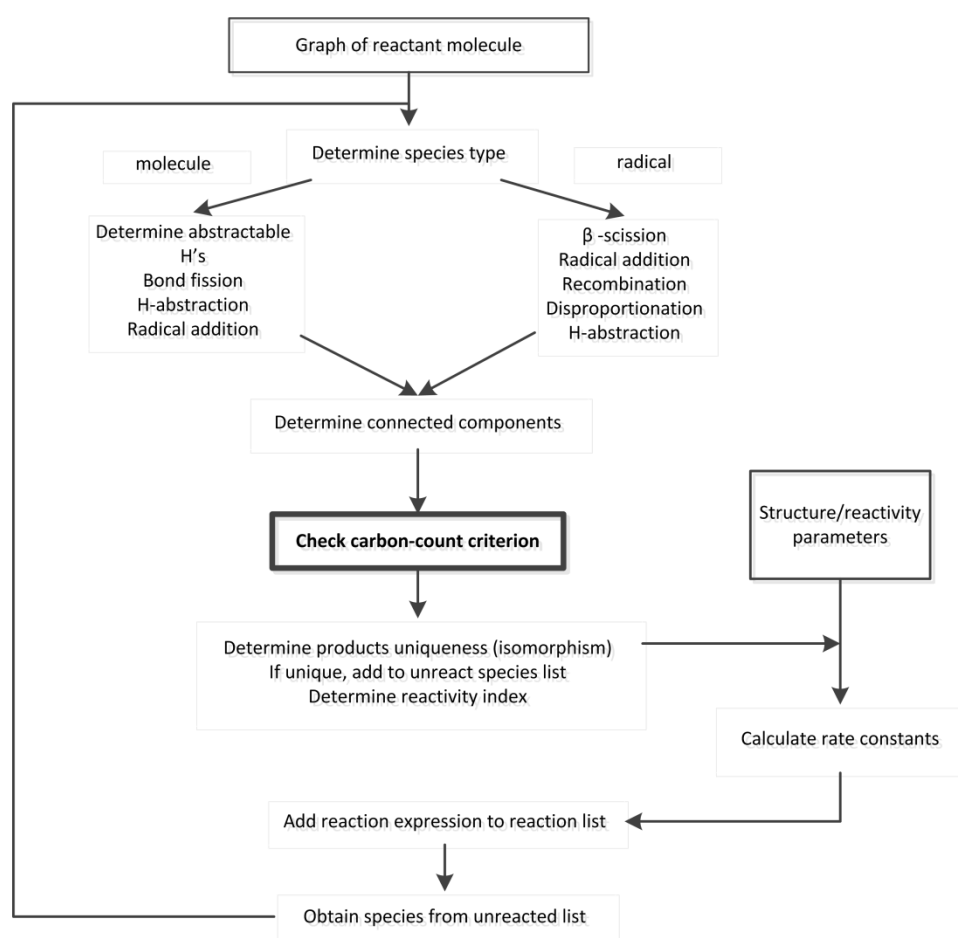
Figure 2-1: Algorithm for the carbon count reaction mechanism generator (Netgen).[1]

The case of ethane pyrolysis was used to validate the new kinetic model builder. Species inside an "unreacted compounds" list are compared to the reaction rules that describe the type of chemical transformations the species can undergo. These reaction rules are user-defined and describe the species' atoms that are involved in a chemical reaction. For instance, every hydrogen in a molecule is allowed to undergo a H-abstraction, or, every carbon-carbon single bond can undergo a bond scission. Products are generated by the addition of a reaction matrix to a reactant bond and electron (BE) matrix. Generated

species are subsequently checked for their uniqueness and added to the unreacted compounds list. The algorithm ends once this list is empty. During pyrolysis chemistry, radical addition reactions play an important role and result in the generation of large molecules. To prevent these generated large molecules from being added to the "unreacted compounds" list, the user has to set a criterion. This criterion is called the maximum carbon count. This condition influences the number of generated species and reactions included in the kinetic model for the pyrolysis of ethane. When the maximum carbon count is equal to two, the model contains 11 species and 55 equations. With a maximum carbon count equal to three, the number of species and reactions increases respectively to 99 and 611. Broadbelt et al. postulated that the number of reactions and species in a model is limited by the memory that is required to store all of the generated information. The used memory is therefore a function of the carbon count number. In megabytes $O(10^{n-1})$ of memory is required where n is equal to the carbon count number.

Essentially, a low carbon count number might assure that the size of the model remains limited, this criterion does not retain or add species and reactions based on their significance. Small kinetic models with only a few important compounds are useless for the prediction of product yields.

2.3. Rank-based termination criterion

To overcome the memory limitations that the kinetic model builder developed by Broadbelt et al.[1] experienced with the simple carbon count criterion, a second termination criterion was implemented. Based upon the carbon atom number and the product rank of a generated species, the species is added to the model. A combination of both criteria results in finite kinetic models with less insignificant compounds and reactions.[3]

The rank-based approach is based on the idea that the products formed during a later stage in the reaction network are of less significance. Therefore, the order in which a product appears in a reaction mechanism is measured. The product rank of a species is by definition equal to this quantitative measure.[4] Traditionally products in a reaction network can be primary, secondary, tertiary, etc. according to their appearance. This terminology is mapped to a numerical value that is termed its rank.[3] This means that primary products have a rank of one and secondary products a rank of two. Some rules are defined regarding the rank-based criterion. Figure 2-2 illustrates two important rules to define a rank to a certain product.
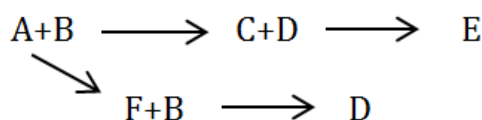
Figure 2-2: Illustration of the species rank in a reaction network.

First of all, the rank of a product is dependent of the rank of the reactants that formed it. This means that product C has rank 1, since it is a primary product formed starting from reactants A and B. Product D is formed starting from F and B, but both reactants have a different rank. By definition the rank of D is two because the first rule states that the rank of a product formed from a reaction of two reactants of different rank has a rank one greater than the maximum of the two reactant's rank. However, product D is also formed as a primary product directly from the bimolecular reaction of A and B. The second rule states that the overall rank of a product formed via several reaction pathways is the minimum of the individual rank.

Radicals and molecules are treated according to analogous yet unique sets of rules. The rank of these species is determined by Equation 2-1 to Equation 2-4.

$$R_{p,i} = max\{R_{r,1}, R_{r,2}, \ldots, R_{r,n}, M_{r,1}, M_{r,2}, \ldots, M_{r,n}\}_i \qquad \text{2-1}$$

$$M_{p,i} = max\{R_{r,1}, R_{r,2}, ..., R_{r,n}, M_{r,1}, M_{r,2}, ..., M_{r,n}\}_i + 1 \qquad \text{2-2}$$

$$R_p = min\{R_{p,1}, R_{p,2}, ..., R_{p,m}\} \qquad \text{2-3}$$

$$M_p = min\{M_{p,1}, M_{p,2}, ..., M_{p,m}\} \qquad \text{2-4}$$

The first two equations calculate the rank of a product formed by an individual reaction. The rank $R_{p,i}$ of a radical species is equal to the maximum rank of the reactants while the rank $M_{p,i}$ of a molecule is always one greater than the maximum rank of its reactants. The overall rank, R_p and M_p, for a species is calculated by the last two equations.

This new convergence criterion is added to the algorithm of the kinetic model builder (Figure 2-1). After the determination of the species uniqueness, the carbon atom number and the product rank of the generated species is checked. If both values are lower than a user-defined stopping criterion, the species will be added to the "unreacted compounds' list.

The results for ethane pyrolysis reveal that a rank-based criterion alone results in kinetic models that do not converge. The coupling of the rank-based with a carbon count criterion provides converging models with large molecular products but the expected products such as butadiene are not formed. When molecules are only constrained by their rank and not by a carbon count number, they can react further to the products that were experimentally found for the pyrolysis of ethane. Coupling of the rank-based criterion with the carbon count criterion imposed on only radical species provides the best capability for predicting reactant conversion and major and minor product yields with the fewest number of excess compounds.

2.4. Rule-based approach

In this paragraph the rule-based approach is discussed. The method requires users' expert knowledge about the chemistry of the process to limit the size of the model and to exclude insignificant species. This knowledge is obtained via experimental and theoretical investigations of the process in the operating conditions at hand.[5] The reaction network is generated based on an algorithm as depicted in Figure 2-3.
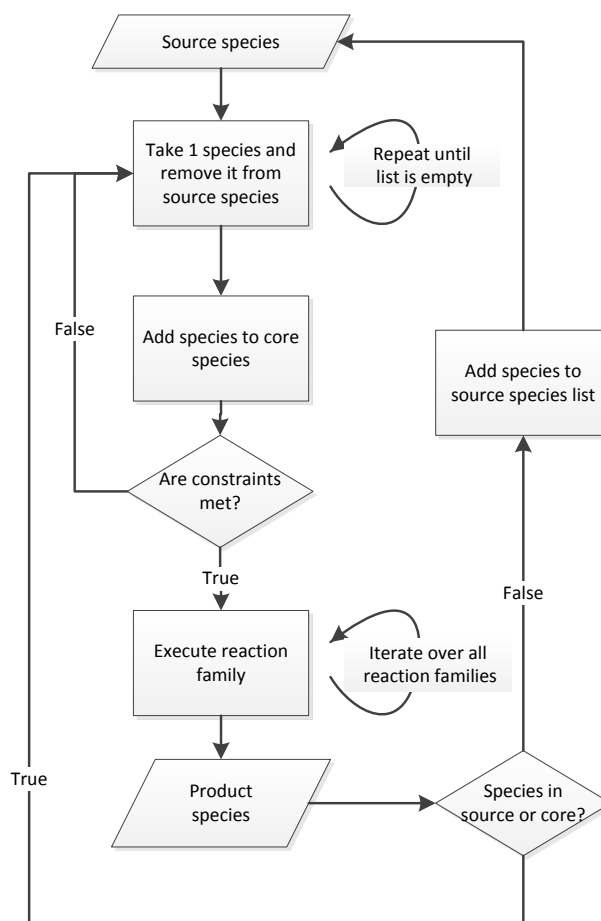
Figure 2-3: Rule-based network generation algorithm.

The kinetic model builder generates new species from initial reactants with pre-defined reaction families. Specific constraints are defined for each reaction family and during the network generation, every species, either provided by the user or generated by the program, that does not comply to the defined constraints will be excluded from undergoing the given reaction family.[6] These constraints are applied to the entire molecule or the reactive moieties present in the molecule. Automatic network generation programs such as Genesys[5,7], COMGEN[8] ,CASB[9] and RING[10], use the rule-based algorithm as a termination criterion. These programs will be discussed to provide more detailed information about the rule-based approach.

2.4.1. CASB (Computer-Assisted Structure building)

Porollo et al.[9] modeled the hemolytic thermolysis of nitromethane with the computer program CASB (Computer-Assisted Structure Building). They suggested that the recombination network approach would solve the problem of theoretical predictions of the possible thermal decomposition pathways of organic substances such as nitromethane. The mechanism of the gas-phase decomposition of nitromethane is well studied but controversial. In this approach, not only the starting materials but also intermediates formed during the generating process are considered as possible participants in various reactions. All the molecules are seen as independent species that can react in any kind of combination, hence the "recombination approach".

The CASB code contains reaction families, or "rule collections", that are collected by the user in separate files. The applicability of these rule collections is limited by nine atom properties and one bond property. Each species is compared to these values by mathematical relations such as "=", "<", "not", "<=" etc. Examples of these properties are the number of neighbor heteroatoms, the number of neighbor hydrogen atoms, charge, the number of unpaired electrons and hybridization. These rule collections can be divided into three groups that differ in the action steps: the transforming, stabilizing and excluding groups. The transforming groups describe the mechanism of the reactions and generate new intermediates. These intermediates are checked by the stabilizing groups. If the generated structures are unstable, such as biradicals, these structures are converted into more stable species. The excluding groups check for undesired species. The entire reaction and the undesired species are removed from the model by these groups. The rules also have a status that can be active of passive, to allow flexible generation of the reaction network.[11]

The kinetic model that Porollo et al. constructed for the thermal decomposition of nitromethane contained every compound that was experimentally observed. Problems occurred when certain new radical intermediates and decomposition pathways were generated. These species and reactions were not observed experimentally and some intermediates are not energetically probable. This means that the rules fail to take into account factors such as the stability of a radical, the energy of bond dissociation, etc.

2.4.2. COMGEN (Complex Mechanism GENerator)

COMGEN (Complex Mechanism GENerator) is an automatic mechanism generator developed by Ratkiewicz to study the combustion of hydrocarbons.[8] This generator is based on a reaction family class approach and a chemical graph theory used by Blurock.[12]

For the generation of a reaction mechanism, a set of pre-defined "reactions patterns" is used. Reaction patterns (or sub-graphs) consist of two to four graphs with vertices and edges that respectively represent

the atoms and bonds in a molecule. Each reaction pattern is an internal representation of a reaction class. An example of a reaction pattern is given in Figure 2-4. The reaction presented by this pattern is the C-H bond dissociation from a secondary carbon. The "-" symbols means that this carbon is connected to two other carbons.
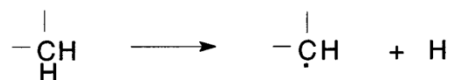

Figure 2-4: Reaction pattern used in COMGEN.

During the generation process, the pattern is compared first to a set of unreacted molecules for the recognition of a substructure. If there is a match, the reaction pattern is applied to the molecules and one or two new structures are generated. This procedure is iteratively repeated until the set of unreacted molecules is empty. The user has several options available to adjust the process. The number of times a given reaction pattern is applied can be chosen. For certain generated species, the user can prevent that any further reaction is generated based on this species as reactant. Also a maximum carbon count for the generated species can be set.

COMGEN was used to generate a kinetic model for the combustion of propane.[8] Alongside 23 reaction patterns, application restrictions were defined in order to limit the kinetic model generation to significant species and reactions. The maximum number of carbon atoms in the recombination products was restricted to three. The number of meta-thesis reactions was also limited by using the concept of β, μ, βμ and Y (resonance-stabilized) free radicals.[13] This means that the radicals are placed into three different groups depending on their structure and based upon their group they can only terminate via specific reaction pathways. This resulted in a mechanism containing 967 reactions and 194 species. Simulations with this mechanism were performed with CHEMKIN and the calculations were compared with the experimental values obtained by Fristrom and Westenberg.[14] The authors of COMGEN suggest some improvements to quantitatively describe combustion systems. To obtain more accurate rate coefficients, on-the fly calculations can be performed with the Reaction Class Transition-State Theory (RC-TST).[15] Another suggestion is the implementation of mechanism reduction techniques such as described in paragraph 2.6 without losing too much information.

As a result, it appears that the COMGEN generator combines rule-based with a carbon count criterion. This is unfortunate because the carbon count criterion is far from a chemical justifiable stopping criterion. This means that COMGEN could possibly miss important reactions of species inside a mechanism.

2.4.3. RING (Rule Input Network Generator)

A third reaction mechanism generator that uses the rule-based criterion is RING (Rule Input Network Generator).[10,16,17] The goal of RING was to build a single platform that could be applied for diverse homogeneous and heterogeneous chemistries. The developers of RING pointed out that the existing kinetic model builders were designed for specific chemistries, thus were very limited in their application. The conversion of biomass to sustainable transportation fuels is a promising alternative for fuels from fossil resources. To understand the chemistry of different biorefinery processes, it would be ideal if a single kinetic model builder could generate models for every different types of chemistry, to compare which process would be best for the conversion of biomass.

For RING, the input requires the initial reactants and the sets of elementary steps as reaction rules. These rules prevent the generation of reaction that are unlikely on the basis of structure and reactivity arguments. The description of a reaction rule is based on a three-step procedure. An example of this procedure is given for the adsorption of a ketone on an acid catalyst (Figure 2-5).
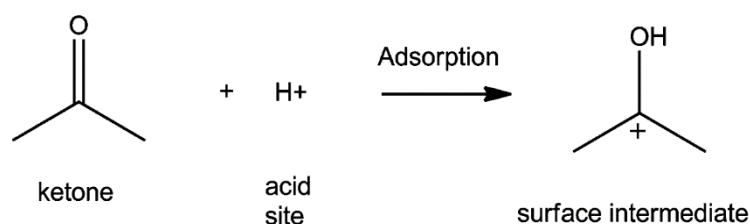
Figure 2-5: Adsorption of a ketone on an acid site to form a carbocation.[16]

The first step is the determination of the reactant pattern, consisting of the set of atoms and bonds participating in a reaction. For the example above, the reaction pattern consists of the keto functional group (C=O) and the acid site of the catalyst. These specific fragments of the molecules will be part of the transformation operation. RING adopts a string representation for reactant patterns based on SMARTS language.[18] During the second step, the transformation operations are described. During the reaction, a bond is formed between the proton and the oxygen. The charge of the proton transfers to the carbon and the double bond weakens until a single bond, C+O, is formed. These transformations are structural changes such as increase or decrease of the bond order and electronic configuration changes such as changes in charge or electron density. The description of constraints is the third and final step. These constraints can be imposed on a molecular level or at the level of the atom. Examples of these constraints are the molecules charge, size or structure. The final reaction rule for the absorption of a ketone is shown in Table 2-1.

Table 2-1: Reaction rule for the absorption of a ketone on an acid site.

Reaction rule
Reactant Pattern
1. C1[!H][!O]=O2
2.H+3
Transformation Operations
1. modify atomtype of 3 to H
2. modify atomtype of 1 to C+
3. decrease bond order of bond (1,2)
4. connect atoms 2 and 3
Constraints
Pattern 1- {s<6}&!{r}&{q0}

The constraints of this reaction rule prevent the size of the species to be more than six atoms and prevents ring structures ({s<6}&!{r}), while the required species has to be a neutral molecule ({q0}). The user can also create a list of global constraints that are to be satisfied by all molecules. Each unreacted molecule is stored in a list and checked against every reaction rule. If the reaction pattern fits and the constraints are satisfied, the transformation rules are applied. As a result, RING reproduces the kinetic models for different chemistries that are described in literature. An important feature of RING is the generation of models for different applications since the reaction rules are described in a generic way.

2.4.4. Genesys

Genesys is the last rule-based automatic kinetic model builder that will be discussed in this work.[5-7,19] This builder was used to illustrate the thermal decomposition of diethyl sulfide and ethyl methyl sulfide (alkyl sulfide pyrolysis).[7]

The input for Genesys is similar to that of RING. An initial set of species and a set of reaction families or rules are required. Reaction family templates are created and contain the reactive moiety that is required to undergo a certain reaction, the transformations that occur during the reaction and information about the assignment of kinetic parameters. Constraints are incorporated in the template to prevent the formation of unimportant species.

In Genesys two types of constraints can be applied for each reaction family: molecular and atomic constraints. Constraints on the atom are, for example, that the atom can be aromatic or not and the atom can have unpaired electrons. It is possible to base these constraints on the SMARTS language restrictions applied on the reactive moiety. Possible molecular constraints are the maximum number of atoms of a

given chemical element, the number of double bonds, the number of smallest set of smallest rings, the number of unpaired electrons and the number of aromatic atoms.

These templates are then repeatedly applied to a pool of initial reactants. When the constraints are met, new reactions and product species are formed. The reactants are added to a list of reacted species and the products are added to a list of unreacted species, if they are not already present in the reacted or unreacted species list. This process continues until the list of unreacted species is empty.

An example of a molecular constraint can be found for the steam cracking of hydrocarbons. Unimolecular reactions are the only kinetically significant reactions for acyclic radicals with more than five carbons atoms. As can be seen in Figure 2-6, this means that two out of five of the generated reactions for the 5-hexenyl radical can be neglected.
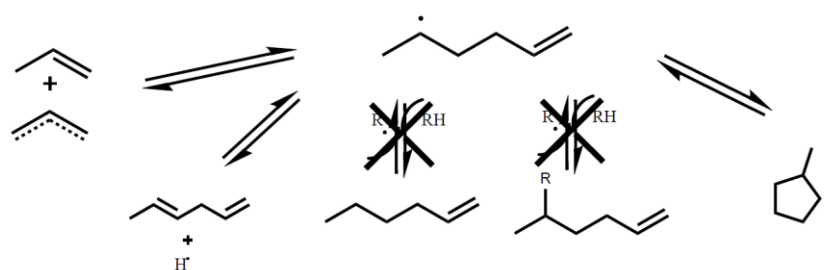
Figure 2-6: Significant and insignificant reactions for the 5-hexenyl radical.

This constraint is based on the βμ rules of Goldfinger-Letort-Niclause.[20] Thus, before an adequate kinetic model can be generated, there is a large amount of knowledge necessary concerning the chemistry that is studied. Otherwise it is impossible to define reaction families with proper constraints.

2.5. Rate-based approach

Susnow et al. were the first to implement the rate-based algorithm to control the size of a mechanism and to terminate the generation process.[21] The program of Broadbelt (described in paragraph 2.2 and 2.3) was modified to use the reactions rates and concentration of species in the model instead of the rank and carbon count criterion. More accurate models were desired and the program had difficulties with adding the important species and leaving insignificant ones out of the model. This rate-based termination criterion is already successfully implemented in several automatic network generation codes.[21-23] A general representation of the algorithm is shown in Figure 2-7.
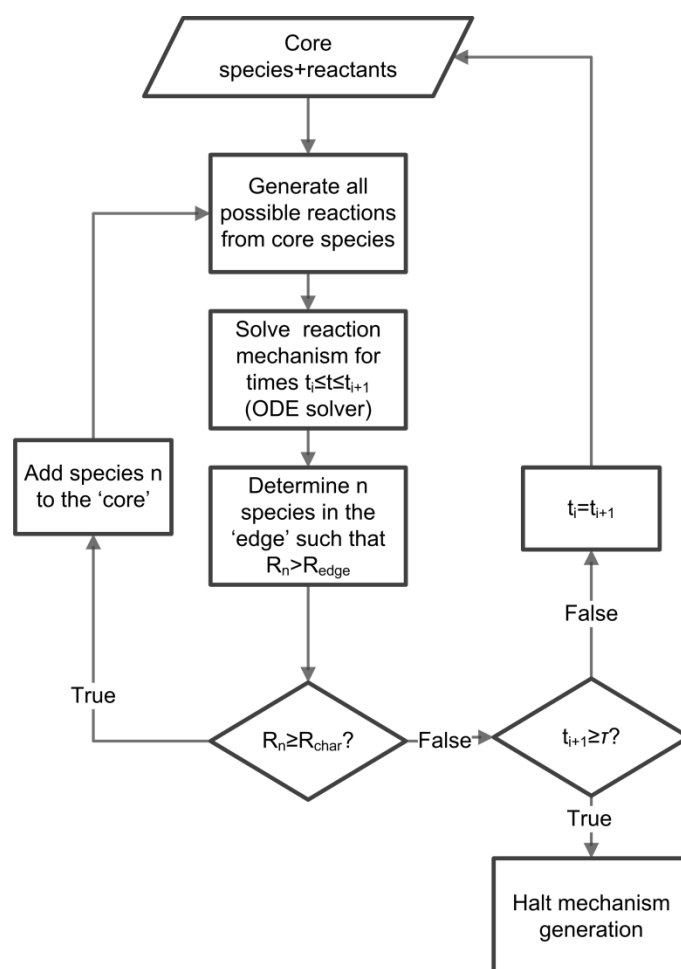
Figure 2-7: Rate-based network generation algorithm.

The principle of this method can be explained as follows. During the generation of a kinetic model, a partial reaction mechanism or "core" contains every important species and reaction. All possible products originating from the core species form the "edge" of this core. Next, one of these edge species

is chosen as the most significant product. This species is added to the core and again all possible reactions are generated, expanding the existing partial reaction mechanism.

The most significant edge species is chosen based on its rate of formation. The rate of formation for every edge species is obtained by solving the differential equations corresponding to the current partial reaction mechanism combined with the appropriate reactor model. No consumption reactions of the edge species are considered thus the formation rates are positive and the differential equations for the edge species are decoupled from the differential equations for the core species. Only the set of equations for the core species has to be solved and yields their concentration profiles. Based on these concentrations, the rate of formation for every edge species is obtained. The reaction mechanism is iteratively generated and solved. To assign rate coefficients to every reaction, thermochemical data is required. In Figure 2-8 this method is represented for the pyrolysis of butanol.
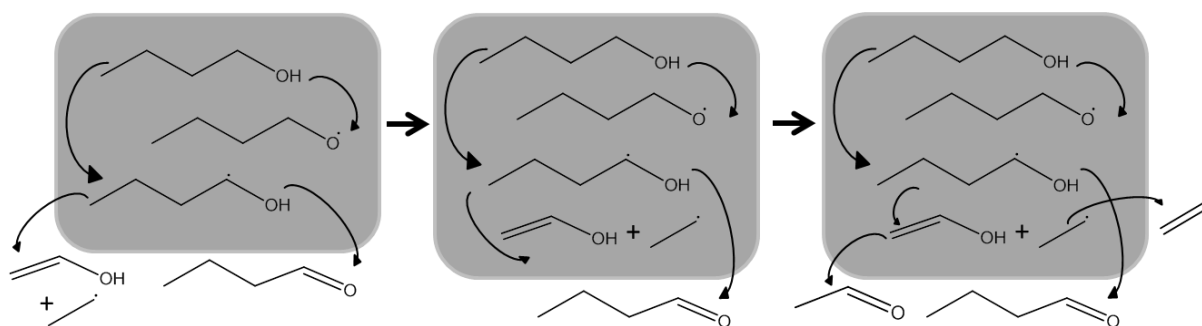
Figure 2-8: The rate-based addition of edge species to the partial reaction mechanism.

The decomposition of butanol starts with the hydrogen abstraction by hydrogen and methylradicals. The first model on the left illustrates that the 1-hydroxybutyl radical present in the network can reactor further into butyraldehyde or into an ethylradical and an ethanol. The latter are the most significant products and are added to the core, after which new edge species are generated. When the rate of formation of the edge species are all less than a minimal rate, R_{min}, the reaction mechanism is considered complete. From here on, no new species will be added to the core and the generation process will terminate.

The minimum rate is defined as the product of a characteristic rate and a precision level ε. This precision level is a user-defined and can be seen as a desired level of precision. When a larger kinetic model is desired, the precision level decreases. This means that the minimal rate decreases as well and more species are included in the reaction mechanism. The definition of the characteristic rate is given by Equation 2-6 and is based on the conversion of a particular reactant A, X_A

$$R_{min} = \varepsilon \times R_{char} \qquad\qquad \text{2-5}$$

$$R_{char} = \frac{[C_{A0} - C_A(\tau)]}{\tau} = \frac{X_A C_{A0}}{\tau} \qquad \text{2-6}$$

C_{A0} and $C_A(\tau)$ are the initial concentration and concentration at time τ of reactant A and τ is the time required to reach a particular conversion X_A.

The influence of the precision level on the generated model for the pyrolysis of ethane is depicted in Figure 2-9. The precision, or value for ε, decreased from 0.5 to 0.0001, while the number of reactions, total species and reacted species increased. The number of reacted species that are included in the model is influenced the least which is beneficial for solving the differential equations of the partial mechanism. However, the number of edge species grows exponentially. This means that generating the unreacted edge species requires a large amount of memory.
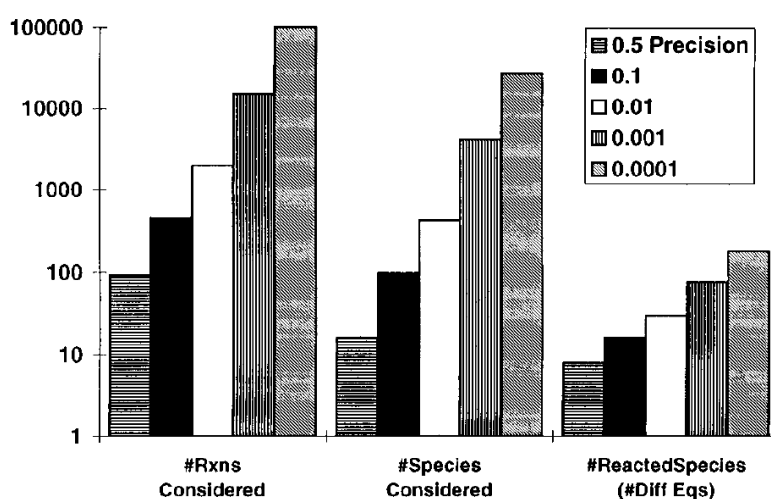
Figure 2-9: Number of reactions, considered species and reacted species for the ethane pyrolysis mechanism as a function of the precision level.[21]

The rate-based algorithm has several disadvantages as well. Susnow et al.[21] discovered that the algorithm would only converge smoothly when the partial reaction scheme was large enough. This means that when the differential equations were solved in the early stages of the scheme, it was possible that kinetically important species and reactions were missing. Therefore, the algorithm starts with the generation of an initial species pool. Again, the users' input is requested to set the size of this pool. The size can be controlled by a carbon count criterion, rank-based criterion or a maximum total amount of species.

De Witt et al.[24] expanded the work of Susnow on the rate-based approach and generated models for long-chain hydrocarbon pyrolysis. Instead of incrementing the reactant conversion X_A after every addition of a new edge species to the core, the reaction time is incremented. The user has to specify a desired total reaction time instead of a total reactant conversion. This adjustment made it easier to apply

the method for systems with multiple reactants, and thus multiple reactant conversions. In this work, the characteristic rate R_{char} is defined as the maximum absolute rate of formation of all core species. R_{char} is not depended anymore on one reactant but on a dynamic set of species. The influence of the precision level on the final kinetic model was brought to attention once again. If the value for ε is too high, important species in the model might be neglected. Reducing the threshold value results in more core species but the number of edge species grows exponentially. An alternative is to keep the threshold value high and to add the kinetically important species into the initial species pool *a priori*.

Another downside of this approach is that the kinetic model is only valid for one specific operating condition.[25] Before the network is generated, the user specifies a temperature and pressure that is required to solve the partial reaction mechanism. These operating conditions preferably are close to the conditions that apply for the system at hand. To compute a model valid over a range of conditions, the program constructs several models for different condition and afterwards forms a grand model based upon the separate models.

It is obvious to see that for this algorithm to work properly, accurate rate estimates are needed. Based on these rate estimates, a decision is be made whether a species is part of the scheme or not and how large the final kinetic model will become. Using slightly different thermochemical parameters of only a few kJ/mol for the same systems results in considerably different models.

2.6. Reduction of automatically generated models

Automatically generated kinetic models can contain unimportant species or reactions resulting in an unnecessary large kinetic model. These models are intrinsically stiff and thus difficult or sometimes impossible to solve numerically. There are various methods to reduce the size of a generated model. These methods can be categorized under either the skeletal reduction methods that eliminate unimportant species and reactions or under the time-scale reduction methods that moderate the stiffness of the model.[26] The goal of mechanism reduction is to remove redundant species and reactions and reduce the size of a kinetic model without losing any essential information.

The skeletal reduction method eliminates species by sensitivity analysis or by lumping of different species into pseudospecies. The lumping approach was successfully adopted by Ranzi et al.[27] to reduce the total number of involved species in partial oxidation, pyrolysis and combustion of hydrocarbon mixtures. Species or reactions of a kinetic model can be lumped together into groups. A collection of elementary reactions are lumped into an "equivalent" apparent single-step reaction. These elementary reactions can be consecutive reactions or related but dissimilar reactions operating as parallel pathways.[28]

Removal of species from a mechanism is performed by identifying important, necessary and redundant species. The important species are defined as species that must be present in a mechanism to predict values in good agreement with experiments or other models. Examples of important species are the initial reactants and the main products of the reactions. Necessary species must be included in the mechanism to produce accurate results for the important species. Finally, species are considered redundant when a change in their concentration does not greatly affect the rate of production f_n of important and necessary species. The effect is given by the sum of squares of normalized Jacobian elements in Equation **2-7**.

$$B_i = \sum_{n=1}^{N} \left(\frac{\partial ln f_n}{\partial \ln c_i} \right)^2 \qquad \text{2-7}$$

If the value of B_i is below a predefined threshold, the direct effect on N important and necessary species can be neglected. Removal of reactions is performed by identifying redundant reactions. The effect of a perturbation in rate parameter k_j on the rate of change of species i is measured via Equation **2-8** and results in an overall sensitivity parameter F_j.

$$F_j = \sum_{n=1}^{N} \left(\frac{\partial f_i}{\partial k_j} \frac{k_j}{f_i} \right)^2 \qquad\qquad 2\text{-}8$$

Reaction j can be removed from the mechanism if F_j is smaller than a predefined threshold value.[29] Species lumping is based on representing information about a group of species using only a single variable. The new lumped variables are related to the original variables by a function called the lumping function.[30] The lumped species are defined as a weighted mixture of similar reactants, intermediates or products (Equation **2-9**).

$$\{A\} = \sum x_j A_j \qquad\qquad 2\text{-}9$$

During pyrolysis of heavy alkanes, the formed isomers are conveniently grouped into a single lumped compound. This is called "horizontal lumping", the lumping of homologous species with a different molecule weight is called "vertical lumping".[31]

There are other techniques that can be applied to further reduce the model, such as principal component analysis[32], detailed reduction[33] and Jacobian analysis.[26] A recently developed reduction method based on graph theory is the directed relation graph (DRG).[34] This method can rapidly and automatically identify and remove redundant species with a high degree of accuracy. An example of this method is given in Figure 2-10.
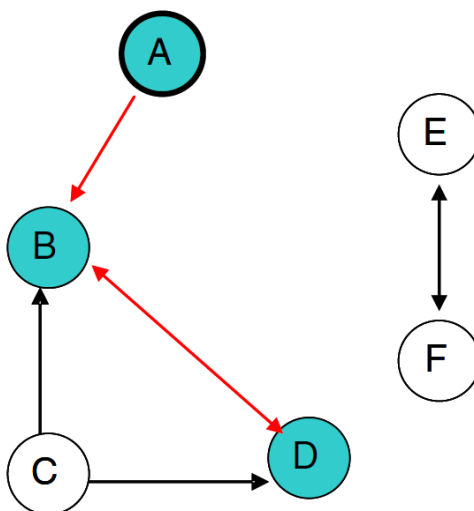
Figure 2-10: Coupling of species in a directed relation graph.

The species presented in this graph are coupled directly or indirectly to each other. Indirect coupled species such as species C and A are coupled with the intermediate species B. The set of species required either directly or indirectly by species A is defined as the dependent set of A. If A was to be removed from the model, the model would incorrectly predict values for species B and C. Therefore, the DRG

method will remove the entire dependent set of A. Species E and F can be eliminated without any effect on A or any species in the dependent set of A. A user specified threshold value determines the accuracy of the reduced kinetic model using the DRG method.[34]

The methods of reduction based on the time-scales include the quasi-steady state approximation (QSSA)[35], the computational singular perturbation (CSP) method[36] and the low dimensional manifold (ILDM) method.[37] There are only a few articles published about the CSP and the ILDM method while the application of the QSSA is a well-established time-scale method. For species that react on a very short time-scale it is assumed that their rate of change can be decoupled from the differential equations. First, these QSS species are identified and their concentration is calculated. Afterwards, the number of differential equations is reduced since some are replaced by algebraic equations.[30] This reduces the overall stiffness of the kinetic model.

Nowadays, reduction methods are still favored for large kinetic models because reactor simulations based on these models remain a difficult task. Tomlin et al.[38] used multiple methods to reduce the chemical mechanism describing the pyrolysis of propane in a cracking tube. The full model was developed by Dente and Ranzi and contains 422 reactions and 48 species.[39] This model showed good agreement with experimental results and results from other models. The objective was to demonstrate that a reduced model can reproduce, within acceptable limits, the yields of major product compounds and the temperature profile predicted by the full model.

Concentration profiles, gas temperature, gas velocity, hydrocarbon partial pressure and total heat input were predicted by conducting reactor simulations. A plug flow reactor model for the cracking of propane at commercial conversions was applied. The operating conditions are described in Table 2-2.

Table 2-2: Operating conditions for the plug flow reactor model.

Inlet temperature (K)	873
Inlet total pressure (bar)	3
Inlet hydrocarbon partial pressure (bar)	1.233
Internal diameter cracking tube (m)	0.05
Heated length (m)	32.9
Reynolds number	>100 000
Firebox temperature range (K)	1400-1500

First, the full kinetic model was simulated to obtain the concentration and temperature profiles. This information is then used in a local rate sensitivity analysis at 12 points along the cracking tube. Redundant, important and necessary species were identified via Equation **2-7**. In total there are 19 necessary and important species. Steam has been identified as a redundant species since the concentration of steam has little effect on the other species. It is of course essential to include steam in

the pressure drop and heat calculations. Next, the redundant reactions are identified. Eventually, 122 reactions were selected as important reactions.

Further reduction of the mechanism was carried out by using overall sensitivity analysis and identifying fast reversible reactions. The latter can be removed in pairs if their overall rate is much slower than the net rate of formation for the species taking part in the reaction. This reduces the number of reactions to 50. The concentration profile and reaction conditions for the full and two reduced mechanisms are compared in Figure 2-11 and Figure 2-12.
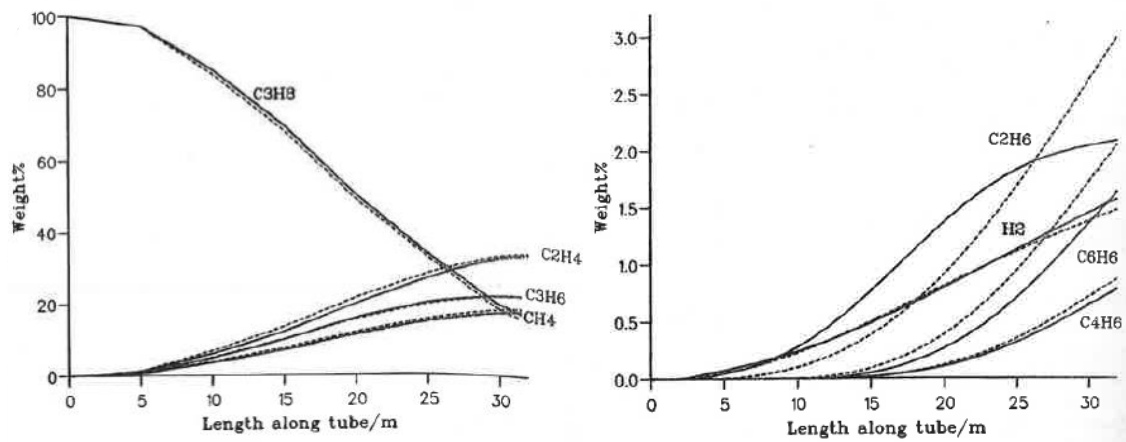
Figure 2-11: Concentration profiles for full model (full line), the reduced model with 122 reactions (long dashed line) and the reduced model with 50 reactions (short dashed line) along the cracking tube with a uniform firebox temperature of 1450 K.
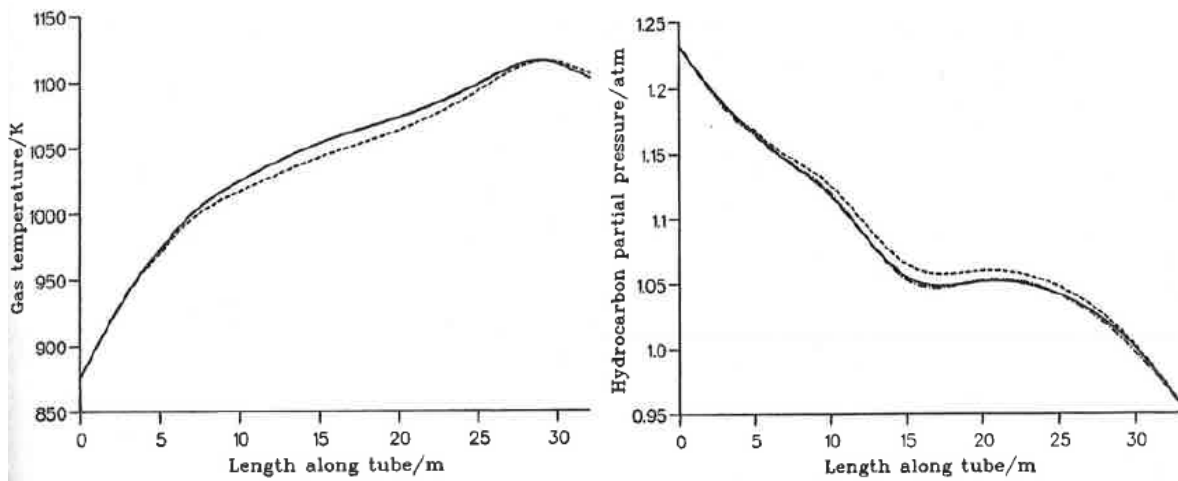
Figure 2-12: Gas temperature (left) and hydrocarbon partial pressure (right) for full model (full line), the reduced model with 122 reactions (long dashed line) and the reduced model with 50 reactions (short dashed line) along the cracking tube with a uniform firebox temperature of 1450 K.

The long dashed line representing the reduced scheme containing 122 reactions is completely covered by the full line. For important products, the weight percent yields remains within 1-5 % for the reduced

scheme. The reduced scheme containing 50 reactions reproduces the concentration of major products such as C_2H_4, C_3H_6 and CH_4 well, however, miner products yields (Figure 2-11, right) are in less agreement with the results of the full scheme. This can be explained by the fact that when less important species are removed completely, the percentages will change of the remaining species while the individual concentrations not differ that much. Finally the QSSA was applied to reduce the number of differential equations. This resulted in 12 QSSA species and minimal errors in the concentrations of the non-steady-state species after the time-scale reduction of the model.

2.7. Conclusion

Automatically generating a kinetic model that can be used to accurately simulate experimental results requires several aspects. First of all, a kinetic model builder should be able to generate the important reactions and species that can occur in a reaction mechanism. Generating all possible reactions and species based on reaction rules is not straightforward, since many of the radical addition reactions lead to large molecules and convergence problems. Also, several of these species are insignificant because they are almost not formed in reality. This results in a large kinetic model that is difficult to use and requires a lot of memory.

Kinetic models builders were developed containing an algorithm that arbitrarily neglects certain species and reaction to keep the model from becoming too big, risking that important reactions will be left out. This results in an inadequate kinetic model as well. The carbon count and rank criterion are two examples of these methods.

The rule-based and rate-based approach are two methods that rely on more logical reasoning. The rule-based method generates only species that are significant by defining a certain amount of constraints. These constraints are solely based on knowledge about the chemistry under study. In order to find which species are significant, there is prior experimental research necessary. Thus, the generated kinetic models based on this approach are not able to 'predict' but rather to validate experimental results.

The rate-based approach does not rely on constraints and generates any species possible starting from an initial set of reactants. Before these new species are added to the reaction mechanism, the rate of formation of these species is calculated. Only the ones with the highest rate are deemed as significant and are added to the model. This approach is more computational intensive but in general, no prior knowledge of the chemistry is necessary. The only crucial information that is necessary is highly accurate thermochemical data. Where this data comes from will be discussed in the next chapter. Another downside is the exponentially growing edge during the network generation. A possibility is to limit the size of this edge with rule-based constraints.

Finally, large kinetic models that are used in combination with reactor models for simulations can be reduced in size. These reduction methods are applied after the network generation and result in easier simulations. Over the years, there are many methods developed that can be used for this purpose. The goal is to lose as little as possible of information during the reduction. Eventually, there will always be a trade-off between a highly accurate model and the size of the model.

2.8. References

1. Broadbelt LJ, Stark SM, Klein MT. COMPUTER-GENERATED PYROLYSIS MODELING - ON-THE-FLY GENERATION OF SPECIES, REACTIONS, AND RATES. *Ind. Eng. Chem. Res.* Apr 1994;33(4):790-799.

2. Ugi I, Marquard.D, Klusacek H, Gokel G, Gillespi.P. CHEMISTRY AND LOGICAL STRUCTURES. *Angew. Chem.-Int. Edit.* 1970;9(9):703-&.

3. Broadbelt LJ, Stark SM, Klein MT. TERMINATION OF COMPUTER-GENERATED REACTION-MECHANISMS - SPECIES RANK-BASED CONVERGENCE CRITERION. *Ind. Eng. Chem. Res.* Aug 1995;34(8):2566-2573.

4. Bhore NA, Klein MT, Bischoff KB. SPECIES RANK IN REACTION PATHWAYS - APPLICATION OF DELPLOT ANALYSIS. *Chem. Eng. Sci.* 1990;45(8):2109-2116.

5. Vandewiele NMU. *Kinetic model construction using chemoinformatics*: Universiteit Gent; 2013.

6. Vandewiele NM. Genesys Manual. UGent: LCT; 2014:65.

7. Van de Vijver R. *Development of a topological symmetry identificat*2012.

8. Ratkiewicz A, Truong TN. Application of chemical graph theory for automated mechanism generation. *J. Chem. Inf. Comput. Sci.* Jan-Feb 2003;43(1):36-44.

9. Porollo AA, Lushnikov DE, Pivina TS, Ivshin VP. Computer representation and generation of possible pathways for thermal decomposition reactions of organic compounds. *Theochem-J. Mol. Struct.* Feb 1997;391(1-2):117-124.

10. Rangarajan S, Bhan A, Daoutidis P. Language-oriented rule-based reaction network generation and analysis: Description of RING. *Comput. Chem. Eng.* Oct 12 2012;45:114-123.

11. Korolev VL, Pivina TS, Porollo AA, Petukhova TV, Sheremetev AB, Ivshin VP. Differentiation of the molecular structures of nitro compounds as the basis for simulation of their thermal destruction processes. *Russ. Chem. Rev.* 2009;78(10):945-969.

12. Blurock ES. Reaction: System for Modeling Chemical Reactions. *J. Chem. Inf. Comput. Sci.* 1995/05/01 1995;35(3):607-616.

13. Warth V, Stef N, Glaude PA, Battin-Leclerc F, Scacchi G, Côme GM. Computer-Aided Derivation of Gas-Phase Oxidation Mechanisms: Application to the Modeling of the Oxidation of n-Butane. *Combust. Flame.* 1998;114(1–2):81-102.

14. Fristrom RM, Westenberg AA. FLAME ZONE STUDIES .4. MICROSTRUCTURE AND MATERIAL TRANSPORT IN A LAMINAR PROPANE AIR FLAME FRONT. *Combust. Flame.* 1957;1(2):217-228.

15. Truong TN. Reaction class transition state theory: Hydrogen abstraction reactions by hydrogen atoms as test cases. *J. Chem. Phys.* Sep 2000;113(12):4957-4964.

16. Rangarajan S, Bhan A, Daoutidis P. Rule-Based Generation of Thermochemical Routes to Biomass Conversion. *Ind. Eng. Chem. Res.* Nov 2010;49(21):10459-10470.

17. Rangarajan S, Kaminski T, Van Wyk E, Bhan A, Daoutidis P. Language-oriented rule-based reaction network generation and analysis: Algorithms of RING. *Comput. Chem. Eng.* 2014;64(0):124-137.

18. Daylight Theory Manual.

19. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chem. Eng. J.* Oct 2012;207:526-538.

20. K.J. L. *Harper & Row, New York.* 1987;3rd ed.

21. Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. Rate-based construction of kinetic models for complex systems. *J. Phys. Chem. A.* May 1997;101(20):3731-3740.

22. Matheu DM, Lada Ii TA, Green WH, Dean AM, Grenda JM. Rate-based screening of pressure-dependent reaction networks. *Computer Physics Communications.* 2001;138(3):237-249.

23. Song J. *Building robust chemical reaction mechanisms: Next Generation of Automatic model construction software*: Departmen of Chemical Engineering, Massachusetts Institute of Technology; 2004.

24. De Witt MJ, Dooling DJ, Broadbelt LJ. Computer generation of reaction mechanisms using quantitative rate information: Application to long-chain hydrocarbon pyrolysis. *Ind. Eng. Chem. Res.* Jul 2000;39(7):2228-2237.

25. Green WH, Barton PI, Bhattacharjee B, et al. Computer Construction of Detailed Chemical Kinetic Models for Gas-Phase Reactors. *Ind. Eng. Chem. Res.* 2001/11/01 2001;40(23):5362-5370.

26. Law CK. Combustion at a crossroads: Status and prospects. *Proc. Combust. Inst.* 2007;31(1):1-29.

27. Ranzi E, Dente M, Goldaniga A, Bozzano G, Faravelli T. Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures. *Prog. Energy Combust. Sci.* 2001;27(1):99-139.

28. Pierucci S, Ranzi E. A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Comput. Chem. Eng.* Apr-May 2008;32(4-5):805-826.

29. Whitehouse L. Using automated techniques to generate reduced mechanisms: University of Leeds Department of Chemistry.

30. Tomlin AS, Turányi T, Pilling MJ. Chapter 4 Mathematical tools for the construction, investigation and reduction of combustion mechanisms. In: Pilling MJ, ed. *Comprehensive Chemical Kinetics.* Vol Volume 35: Elsevier; 1997:293-437.

31. Marin GB. *Advances in Chemical Engineering*: Elsevier Science; 2011.

32. Vajda S, Valko P, Turányi T. Principal component analysis of kinetic models. *International Journal of Chemical Kinetics.* 1985;17(1):55-81.

33. Wang H, Frenklach M. Detailed reduction of reaction mechanisms for flame modeling. *Combust. Flame.* 1991;87(3–4):365-370.

34. Lu T, Law CK. Linear time reduction of large kinetic mechanisms with directed relation graph: n-Heptane and iso-octane. *Combust. Flame.* 2006;144(1–2):24-36.

35. Brad RB, Tomlin AS, Fairweather M, Griffiths JF. The application of chemical reduction methods to a combustion system exhibiting complex dynamics. *Proc. Combust. Inst.* 2007;31(1):455-463.

36. Lam SH, Goussis DA. The CSP method for simplifying kinetics. *International Journal of Chemical Kinetics.* 1994;26(4):461-486.

37. Battin-Leclerc F. Detailed chemical kinetic models for the low-temperature combustion of hydrocarbons with application to gasoline and diesel fuel surrogates. *Prog. Energy Combust. Sci.* 2008;34(4):440-498.

38. Tomlin AS, Pilling MJ, Merkin JH, Brindley J, Burgess N, Gough A. Reduced Mechanisms for Propane Pyrolysis. *Ind. Eng. Chem. Res.* 1995/11/01 1995;34(11):3749-3760.

39. Albright LF, Crynes BL, Corcoran WH. *Pyrolysis, theory and industrial practice*: Academic Press; 1983.

Chapter 3
Kinetic data

Over the years, databases containing chemical information have grown tremendously. They have become more diverse and reliable. However, gathering the necessary data via experiments or computational chemistry calculations remains time consuming. The quality of the obtained values strongly depends on their origin, yielding a wide variety in availability and accuracy of data. During the automatic generation of a kinetic model, databases are required to determine rate coefficients, thermodynamic properties and transport properties. This chapter focuses on kinetic parameters which are crucial for rate-based network generation, as discussed in the previous chapter. These values are also essential for reactor simulations that describe the concentration profiles at certain operating conditions. An important question is how the error variance on these values would influence the generated model and secondly, how this information can be stored efficient in open-source databases.

Commonly, gathering information is done in-house and not distributed to the outside. In many cases, these databases have been established to address specific interests of their developer in biological or chemical systems with relatively little attention paid to the integration of different types of biological, chemical and literature data.[1] If the chemical data was to be shared inside the scientific community, this would save a lot of effort and progress in the field would accelerate. Another important issue is de compatibility of data originating from different sources. The difference in accuracy can have a certain influence on a generated kinetic model. To what extent this influence reaches is not yet extensively described in literature, but the user must be aware of the possible deviations from reality when validating the model.

3.1. Sources of kinetic data

Formerly, modelers relied on thermodynamic and kinetic data from carefully designed measurements. As the need for more detailed simulations is growing, the kinetic models became more complex and contained more species and reactions. Especially for combustion, oxidation and pyrolysis where radical chemistry plays a major role, the generated reaction network can contain up to thousands of elementary reactions.[1] For these networks, accurate thermodynamic and kinetic parameters need to be at hand.[2] Experiments require financial investment and labor which makes them scarce. More often computational chemistry provides a complementary source to the chemical information due to the increase in computational power and the reduced cost of simulations. Figure 3-1 illustrates how the performance of processors for standard computers has evolved over the years. Developments in processor technology such as the use of multicore processors improve the performance of computers. These developments are beneficial for computational chemistry.
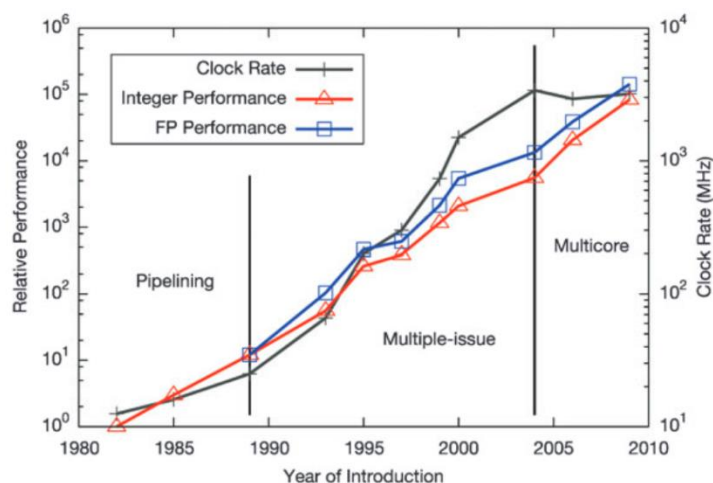
Figure 3-1: Development of clock rate and performance of Intel chips from 1982 until 2010.[3]

The quality and quantity of kinetic data depend strongly on its source. Experimental studies have always been the primary source of chemical data and are constantly improving. Computational methods can be used when acquiring kinetic parameters experimentally becomes too difficult. Although, large computational time for large species puts a limitation on this method as well and the accuracy of the results increases with computational resources. Finally, estimation methods based on semi-empirical correlations that were derived over the past decades can be used. A short overview of these three methods and their accuracy will follow.

3.1.1. Experimental

When a kinetic modeler needs chemical data, experimentally measured values are his primary relief. Since these values can be derived from different experimental setups and different research groups, the

information is mostly widely scattered and the quality can be variable.[4] Also, these values are often only valid for a limited range of operating conditions[5] and regression would introduce additional uncertainties.

The recent experimental techniques available to produce kinetic data for radical processes can be divided in two groups: techniques based on photochemical or electron impact to produce radicals and chemical shock tube techniques.[6] Their range of operating conditions, accuracy and limitations will be briefly discussed in this paragraph.

3.1.1.1 Chemical shock tubes

Chemical shock tubes are used to measure rate coefficients at a temperature above 1000 K and over a pressure of 1 to 1000 bar.[6] A simplified scheme is given in Figure 3-2. When a gas at high pressure is suddenly expanded in a gas at low pressure, a shock wave is created. This shock wave will move through a long closed tube while it produces a rapid decrease of pressure and temperature of the reactive mixture. Inside this tube a radical precursor, a diluent and a reactant is present. This well-designed precursor is the source of the radicals that will take part in the elementary reaction under study. The radicals are produced by pyrolysis of the precursor on a timescale that is short compared with the reaction time of the occurring reactions.
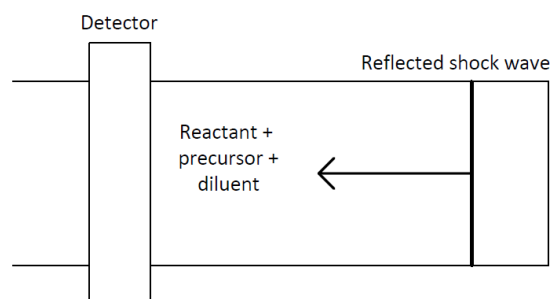
Figure 3-2: Simplified principle scheme of a shock tube.

The detection method used at the end of the tube requires high sensitivity because of the high temperatures that are reached. When working at high temperatures, there are usually other complex chemistry reactions that influence the rate of removal or production for the elementary reaction.[4] These secondary reactions are a potential source of error. If the involved chemistry is well known, it is possible to take this effect into account and improve the accuracy. Thanks to the application of advanced measuring techniques, the accuracy of shock tube results has been improved.

3.1.1.2 Photochemical impact techniques

Pulsed photolysis or photochemical impact techniques produce radicals by laser or a plasma. The photolytic nature makes the radical source independent of the temperature. This results in a error of maximum 10% in most instances. These methods are restricted to lower temperatures (below 1000 K) and are mainly developed to determine kinetic parameters for reactions of interest in atmospheric chemistry.[7]

During laser flash photolysis (LFP), UV radiation from a pulsed laser is used to dissociate a radical precursor. The laser power is designed to keep the concentration of the formed radical low enough which minimizes radical recombination reactions. Several detection techniques can be used such as laser induced fluorescence (LIF), UV and IR absorption spectroscopy and mass spectroscopy.

3.1.1.3 Discharge flow method

With the discharge flow method (or the fast flow tube method) a continuous production of radicals is obtained by microwaves discharges. A simplified scheme is shown in Figure 3-3.
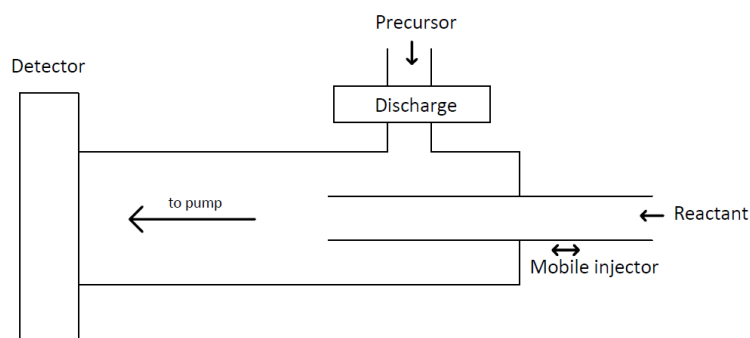
Figure 3-3: Simplified principle scheme of a fast flow tube

The discharge flow tube consists of the quartz or glass tube in which a carrier gas with atoms or radicals flows and a reactant can be injected through a movable injector. There is a continuous generation of radicals when the precursor passes the microwave discharge. At the end of the tube there is an optical pulsed or continuous detector to measure the radical concentration. The reactant can be added inside the tube at various distances. If the mixing region of the reactants and radicals is changed, the reaction time is changed as well and various kinetic information is obtained. The formed radicals can enter via the mobile injector instead of the reactants as well. Another radical behavior will be measured with this setup because the radicals can interact more with the wall of the tube, resulting in a great radical decay. The precision of the measurements lie in the same region as for the photochemical methods, which is a standard deviation of about 10 %.

Other techniques such as flame experiments, flow and stirred reactors can be used to determine rate coefficients but have a less general applicability. The measurements can be validated with sensitivity analysis to show which elementary reactions were the major contributors to the production or formation of the studied radicals.[8] When rate coefficients are published, occasionally the error limits are quoted which are based on the precision of the measurements by reproducing the same experiment several times and by measuring the deviations.

3.1.2. Computational chemistry methods

Computational methods can be used as an alternative to experiments to calculate the required chemical data for a kinetic model. These methods can vary from *ab initio* calculations, based on first principles, to empirical methodologies.[9] Whenever computational chemistry is applied for the calculation of thermodynamic and kinetic data, an appropriate theoretical method must be applied. Depending on the available computational resources, the problem at hand and the required accuracy for the data, a best suited method is chosen.[10]

Essentially, chemical properties such as thermodynamics and kinetics are derived from the total energy of a system, which is found after solving the fundamental Schrödinger wave equation. Exact solutions for the wave equation only exist for small systems. To solve the equation for larger systems, assumptions are made such as the Born Oppenheimer. This assumption allows to factorize the wave equation into a set of basis functions and results in a potential energy surface (PES). Once the PES is known, thermochemical properties can be derived. Kinetic properties require extra theories such as the Transition State Theory (TST), which are discussed further in this paragraph. There are two common methods that are used to solve the wave equation, both based on different assumptions. The Hartree-Fock (HF) theory is the simplest method and relies on the mean-field approximation which neglects the electron configuration energy. The Density Functional Theory (DFT) calculates the electron density and not the wave function itself.[11] A commonly used DFT method to calculate the partition function and optimize the geometry is the B3LYP (Becke, 3-parameter, Lee-Yang-Parr) method.[12] In Figure 3-4 a classification of the different computational methods is given.
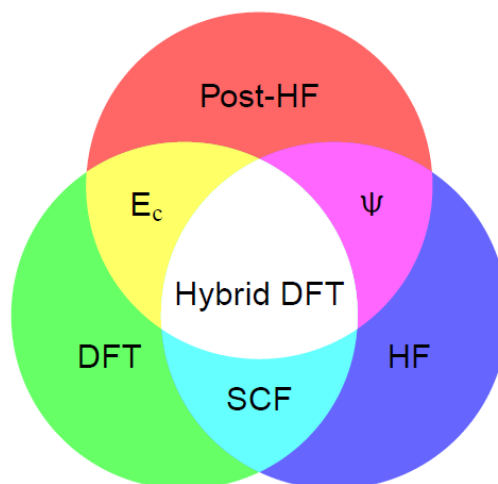
Figure 3-4: Classification of computational methods[13]

In some cases a large error can be made when neglecting the correlation energy. There are several methods to compute this energy such as the Møller-Plesset theory, Configuration Interaction (CI) and Coupled-Cluster (CC) methods. These methods are summarized under the name Post-HF methods.[14] Another type of method can be referred to as the composite method and relies on a combination of high-level single point calculations followed by a correction scheme. The most popular composite methods are the Gaussian methods[15], complete basis set (CBS) methods of Petersson[16] and Weizmann methods (Wn) of Martin.[17]

For the calculation of the kinetic properties, methods based on the conventional transition state theory in the high pressure limit for unimolecular or bimolecular reactions, Equation 3-1 and Equation 3-2 are used.

$$k_{unimolecular}(T) = n_e \kappa(T) \frac{k_B T}{h} \frac{q_{\ddagger}}{q_A} e^{-\frac{\Delta^{\ddagger}E}{RT}} \qquad \text{3-1}$$

$$k_{bimolecular}(T) = n_e \kappa(T) \frac{k_B T}{h} \frac{q_{\ddagger}}{q_A q_B} e^{-\frac{\Delta^{\ddagger}E}{RT}} \qquad \text{3-2}$$

Where T is the absolute temperature, h the Planck constant, k_B the Boltzmann constant, $\kappa(T)$ the tunneling correction factor and n_e the number of single events. TST is a statistical theory because it calculates the potential energy of the saddle point on the PES. This saddle point is stationary and represents a transition state structure of the elementary reaction. However, it is not guaranteed that this energy is really the highest free energy of any point along the reaction coordinate. The TST will therefore always overestimate the rate coefficient. Variational Transition State Theory (VTST) and Flexible Transition State Theory (FTST) account for the dynamics of this rate coefficient.[18,19] The VTST

variationally moves the reference position along the reaction coordinate that is employed for the computation of the activated complex free energy.[20] The disadvantage of computational chemistry is the steep increase of required resources with the size of the molecules. Group contribution methods developed to estimate the thermodynamic properties of species can be used to estimate kinetic properties as well. These methods will be described in the next paragraph.

There are a number of computational methods that can be used to calculate the necessary properties. A trade-off between the computational cost and accuracy will always be present. The better the correlation energy is accounted for and the larger the size of the basis set, the higher the level of theory the method has. Normally, the geometry optimization of the molecule is performed at a lower level of theory. After the optimization, the calculation of the total energy of the complex can be performed with a higher level of theory. Vandeputte et al. evaluated the performance of various computational methods for reactions involving organosulfur compounds by comparison with experimental data.[2] The extended level of theory study assessed the accuracy of 82 *ab initio* methods. The study demonstrated that two composite methods, the G3B3 and CBS-QB3, and one DFT method, the BMK/cc-pVTZ, proved to be powerful tools to calculate both thermodynamic and kinetic data with a high accuracy.

Genesys uses kinetic group additive values that are determined using the CBS-QB3 method of Montgomery et al.[16]. Eventually, the use of *ab initio* methods to obtain kinetic coefficients remains an intensive task, therefore the use of estimation methods is an alternative. However, high performance computing capacity allows to increase the use of more accurate methods and a reaction network generator that is linked with computational chemistry packages to calculate kinetics and species properties on-the-fly is being developed.

3.1.3. Semi-empirical methods to estimate rate coefficients

When the necessary data is lacking, the kinetic model builder will depend upon estimation methods to obtain thermodynamic and kinetic data. These methods give the possibility to rapidly obtain data for a large set of reactions.[6] The Benson group additivity method has proven its ability to estimate thermodynamic properties for stable molecules and radicals.[21-23] Given a set of accurate values for smaller molecules and reactions, obtained via experiments or *ab initio* calculations, values for larger molecules and reactions are calculated. If the thermochemistry of the transition state is known, the macroscopic formulation of the rate coefficient, Equation 3-3, can be used.

$$k(T) = n_e \kappa(T) \frac{k_B T V°}{h} \exp(\frac{\Delta S^{\ddagger}}{k_B}) exp(-\frac{\Delta H^{\ddagger}}{k_B T}) \qquad \text{3-3}$$

Where V° is the molar volume in the standard state. The standard enthalpy of formation of a transition state is determined accurately with Benson's group additivity method.

Another group additivity method is shown in Figure 3-5. Here the transition state of a radical addition reaction is divided into primary groups. The C_1, C_2 and C_3 carbon atoms change during the reaction and are the central atoms of the primary groups. The dotted line represents the reactive moiety.
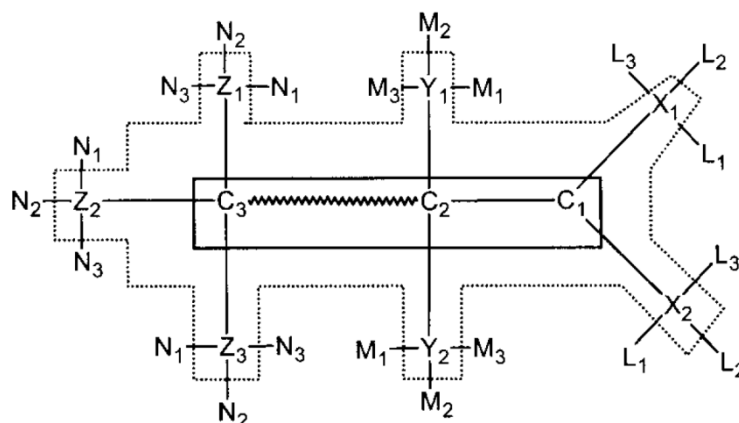
Figure 3-5: Transition state for a β-scission and a radical addition reaction[24]

The three primary groups are equal to C_1^{TS}-$(C_2^{TS})(X_1)(X_2)$, C_2^{TS}-$(C_3^{TS})(C_1^{TS})(Y_1)(Y_2)$ and C_3^{TS}-$(C_2^{TS})(Z_1)(Z_2)(Z_3)$. The corresponding transition-state specific GAVs are then determined for these groups. Saeys et al. validated the group concept and additivity approximation for 67 hydrogen radical addition and β-scission reactions with *ab initio* data.[24]

There are plenty of other correlations suggested from empirical results that parameterize the rate or the activation energy between members of the same reactions family. The general form of these relationships is given by Equation 3-4.[25] These correlations are categorized under the linear free-energy relationships (LFERs).

$$\ln(k_i(T)) = \ln(k^{ref}(T)) + m \cdot (x_i - x^{ref})$$
3-4

Where $k^{ref}(T)$ is the rate of a known reaction belonging to the same reaction family, m is the characteristic of the reaction family, x_i is a property of reaction i or the species in reaction i and x^{ref} corresponds to the reference reaction that defines the family. The most popular correlations are the Evans-Polanyi (Equation **3-5**), Hammet or the Taft correlation.

$$E_a^i = E_a^0 + \gamma \Delta_r H_i^\circ$$
3-5

Evans and Polanyi considered the pre-exponential factor constant for all the reactions of the same reaction family and the activation energy linear dependent on the reaction enthalpy. Here, $\Delta_r H_i^\circ$ is the reaction enthalpy of reaction i, E_a^{ref} is the *intrinsic barrier*, which is the activation energy for a thermoneutral reaction, hence noted with the '0' superscript and γ the transfer coefficient for the reaction

family. The advantage of this method is that the activation energy can be calculated from the heat of reaction which is easily estimated.[26] This relationship is linear as a function of the reaction enthalpy while for most reactions there is a non-linear dependence of the activation enthalpy on the reaction enthalpy, thus the accuracy is mediocre. This is illustrated in Figure 3-6 where experimental activation energies for 302 reactions are plotted as a function of the reaction enthalpy. There appears to be a nonlinear relationship showing that the Polanyi relationship can only be used for a small range of data.
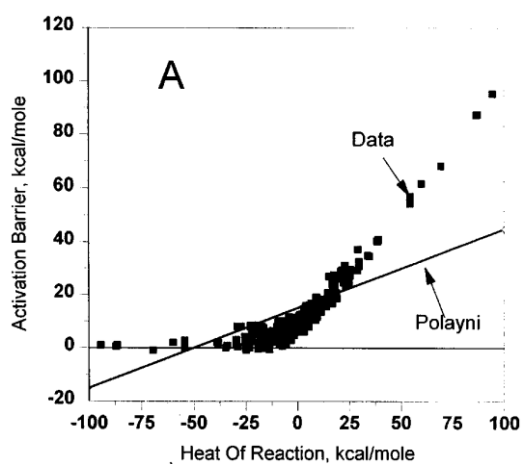
Figure 3-6: Activation barriers as a function of reaction enthalpy[26]

Alternative estimation methods try to describe this nonlinear relationship more accurate. A very popular nonlinear relationship is the Marcus equation (Equation **3-6**).

$$E_a^i = E_a^{ref} \cdot \left(1 + \frac{\Delta_r H_i^\circ}{8 \cdot E_a^0}\right)^2 \qquad \text{3-6}$$

The pre-exponential factor is again assumed constant for all the reactions of a same reaction family. The Marcus equation however fails to describe activation barriers for very endothermic and very exothermic reactions. Blowers and Masel[26,27] extended the approach of Evans and Polanyi by accounting for the nonlinear behavior. They described the activation barriers of 151 hydrogen transfer reactions and compared these with the results of the Evans-Polanyi relation and Seminov's extension of the former relationship. They found that 142 reactions are within the range of 4 kJ/mol of the model.[26] There are other correlations (Hammet (1937), Taft (1952), Shustorovich (1990, 1998)) whose description fall outside the scope of this introductory paragraph.

3.2. Evaluation of kinetic data

After all the species and reactions are generated, the automatic kinetic model builder will assign thermodynamic and/or transport properties to the molecules and rate coefficients to the reactions. Ideally, all the required chemical data is available with a "chemical accuracy". This chemical accuracy is defined in literature as to within 4 kJ/mol of the experimental data for thermodynamic properties. In computational chemistry a mean absolute deviation (MAD) of about 2 kJ/mol is required to achieve the same chemical accuracy as for the experimental methodology.[28] For kinetic data, comparable accuracy cannot be reached yet, it is not exceptional that this data deviates with a factor of 10 of the experimental data.

The rate coefficients that are derived experimentally or computationally need to be evaluated on their reliability before they can be used in a kinetic model. A quantification of the degree of confidence for the values over a given temperature range can be useful to validate the uncertainty of the generated model.[29]

There are two large groups, IUPAC (International Union of Pure and Applied Chemistry) and NASA (National Aeronautics and Space Administration), who use their expertise in a panel and meet regularly to evaluate data on photochemical and chemical kinetics for atmospheric modeling. Smaller groups and individuals still producing evaluations periodically are Tsang (NIST), Baulch (Leeds), Cohen (Aerospace Corporation) and Atkinson (Riverside, CA).[4] The original purpose of these groups was to obtain new insights into the atmospheric chemistry and in the chemistry of exhaust gasses. The process of evaluating kinetic data can be divided into several steps (Figure 3-7). First, the data is collected from published articles. There are several requirements for this data. Primarily, sufficient detail about the experimental technique or the computational method used should be given to assess the quality of the work. Results should be reported in sufficient detail to be reanalyzed and reinterpreted by others. And there should be estimates of the precision and, if possible, the accuracy of the results.[30]
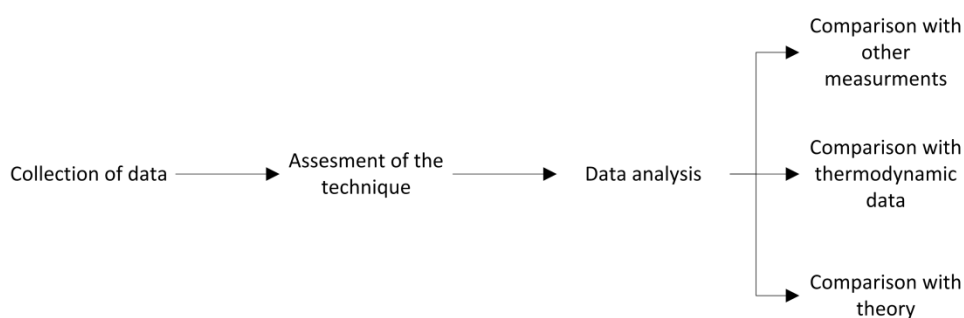
Figure 3-7: Procedure for the evaluation of kinetic data.

In the next step, an assessment of the techniques used to acquire the data is performed, followed by a data analysis. There will always be a systematic error or scatter of results with experimental techniques. When computational methods are used the assumptions that are made result in certain deviations. Also, transport limitations should not influence reaction kinetics during the experiments.[31] Throughout the evaluation, an effort should be made to identify the source and the magnitude of these errors.[32]

Finally, a comparison is made with existing data, thermodynamic and theoretical data. Sometimes, the evaluator can have a number of different measurements obtained with the same technique but from different sources. If there is no good agreement, the evaluator will either recommend one measurement in preference to the others with good arguments or he will choose a measurement with wider error limits. Thermodynamic data such as the equilibrium coefficient K_c is often used to evaluate if the rate coefficients are consistent via Equation **3-7**.

$$K_c = \frac{k_+}{k_-}$$
<div align="right">3-7</div>

Finally, experimental values can be compared to kinetic information obtained via the transition state theory. Eventually, when kinetic data is recommended after evaluation, an expression is derived in which the rate coefficient is expressed as a function of temperature, and if necessary, as function of the pressure. The critical evaluation of kinetic data is accepted as an essential part of the modeling process.[4]

3.3. Database and format

Currently many of the kinetic parameters required for a kinetic model can be found in online databases such as those of NIST and IUPAC. Most of this information has been critically evaluated as described in the previous paragraph. If databases contain reliable kinetic information, the effort a modeler has to put into collecting information himself is greatly reduced, and the accuracy of the generated kinetic model increases.

The goal is to save experimental data and details on kinetic models in open formats that will be easily and permanently accessible in the future.[33] Kinetic data that are not published are usually lost within a few years.[34] Different methods (high throughput screening[35], computational methods with high performance computers (HPC) and various experimental setups) have led to an exponential growth in the amount of chemical data. Thus databases stored as hardcopies are more difficult to update and could become rapidly outdated. The compatibility of different electronic formats that are used to store the information is important as well as the recent databases that are stored in PDF format online. This makes the information accessible to anyone but requires some transformation of data when implemented in kinetic model builders. The Extensible Mark Language (XML) is probably favored the most as electronic format.

3.3.1. IUPAC

The IUPAC[36] datasheets consist of two types, the sheets for individual thermal reactions and those for individual photochemical reactions. Not every reported value for a reaction is present in this database but only the preferred values that are carefully evaluated.

Compared to other databases, IUPAC dedicates more effort into providing the user with an estimation of the reliability. Uncertainties for the rate parameters for the thermal reactions are a subjective assessment of the evaluators. These are based on the knowledge of the techniques, the difficulties of the experimental measurements, the potential for systematic errors and the number of studies conducted.

3.3.2. NIST

Often, group additive values obtained from *ab initio* calculations are compared with experimental values taken from the NIST Chemical Kinetics web site. [37,38] In this database over 38 000 separate reactions for over 11 700 distinct reactant pairs are recorded. Kinetics for thermals gas-phase reactions from over 12 000 papers are collected. The objective was initially on reactions important in combustion chemistry.

Records for rate coefficients contain the rate parameters A, Ea and n and, if reported in literature, the uncertainty in A. The temperature range for which the parameters are valid is reported as well. Besides

this kinetic database, NIST has also a WebBook page that provides thermodynamic, spectral and other data.

The database does not recommend for a specific reaction the most accurate value but instead, gives the user every reported value that can be found in literature. Also, error limits are rarely reported. This means that the user of a kinetic model cannot conclude if a generated kinetic model is incomplete or incorrect due to large deviations for the kinetic data.

3.3.3. PrIMe

The Process Informatics structure, referred to as Process Informatics Model (PrIMe), has as primary goals the gathering, validation and transformation into usable form of information concerning predictive models for chemical reaction systems. PrIMe consists of components such as a data depository where data provided by the community is stored, a data library where evaluated data is stored and computer-based tools to process and assemble data into predictive models.[15] The data depository and data library are part of the PrIMe Data Warehouse (Figure 3-8). This warehouse can be seen as a "living" collection of data for combustion chemistry. One objective is to expand this database for atmospheric chemistry and astrophysics. The database contains hierarchical "layers" of information like molecular species, elementary reactions and reaction models.
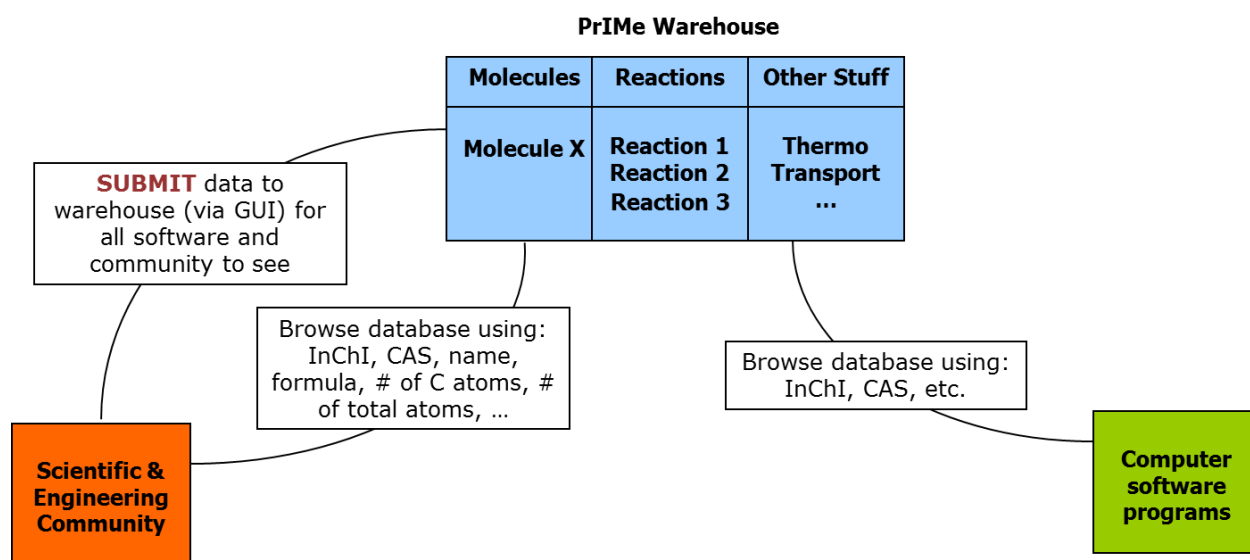
Figure 3-8: Structure of PrIMe.[39]

When an experimenter wants to submit data to PrIMe, he will make a request to deposit his observations via automated web submission forms. Afterwards, PrIMe will evaluate the data as described in the previous section. The entire collection is open and everyone can examine, criticize or make recommendations on the collected information.

3.3.4. New initiatives

Nowadays, PrIME has not yet been accepted by the scientific community as the main database for storing of thermochemical information. The Combustion Institute cyberinfrastructure task force has presented a similar initiative trough a SMARTCAT COST action.[40] The project itself is in an early stage since no implementation strategies of any kind have been set up and the discussion is still ongoing. The main object is to decentralize the data and to perform an extensive survey in the initial phase of the project. By decentralizing data in some form of "social network" the user can choose to keep his data local or to share data within the user's research network. This decentralization of data guarantees that maintenance and cost of the database systems are spread across the users. With a preliminary survey the needs of the users will be investigated before the development of the project. This survey helps to shape the interface of the project to the user's requirements, this way the interface will be adapted instead of the user having to adapt to the interface.

3.3.5. Comparison of kinetic data for steam cracking of ethane

To illustrate the influence of different sources of thermochemical information on the automatic network generation, a small model for the steam cracking of ethane is generated. The rate coefficients for this model will originate from two different sources. The first kinetic model has rate parameters assigned by Genesys during the final stage of the network generation while the second model contains the same elementary reactions but has rate parameters obtained from the NIST kinetics database[37].

The generated model is relatively small and contains 41 species and 57 elementary reactions. Leaving out insignificant reactions based upon their reaction rate reduces the model to 23 species and 33 reactions. Notice that this model is not able to describe a realistic ethane cracking process since to many reactions forming large products were left out. The goal of this model is not to describe experimental results but to compare results originating from different kinetic data. The elementary reactions of this model are listed in Appendix A. The "full model" and the reduced model were both simulated with an ideal plug flow reactor model to check if no important reactions for the simulation of some molar fractions were left out. As a result, the molar fraction of ethane, hydrogen, ethylene and methane simulated by the reduced model were within a range of 0.01 % of the molar fractions simulated with the full model. To model the plug flow reactor, the operating conditions as described in Table 3-1 were used.

Table 3-1: Operating conditions for the plug flow reactor simulation.

Coil inlet temperature (K)	498
Absolute coil outlet pressure (10^5 Pa)	1.90
Ethane mass flow rate (kg s^{-1})	8.3 10^{-4}
Steam mass flow rat (kg s^{-1})	5.8 10^{-4}
Steam dilution (kg$_{steam}$ kg$_{ethane}^{-1}$)	0.7
Internal diameter cracking tube (mm)	10
Heated length (m)	23.135

A temperature profile in the axial direction was taken from Pyl. et al[41] (Figure 3-9). The pressure was set constant to the value reported in Table 3-1. The energy and momentum equation were not solved because the temperature and pressure were provided along the axial coordinate of the reactor.
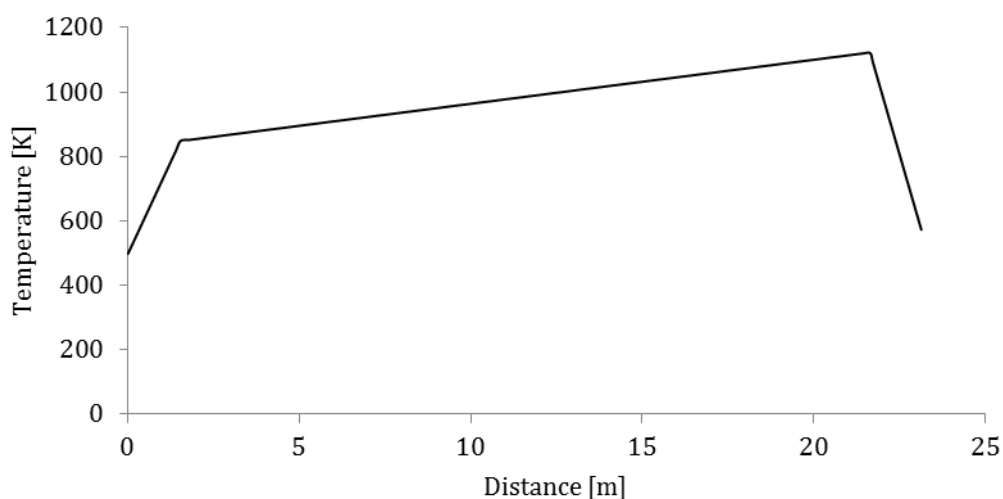
Figure 3-9: Temperature profile along the axial coordinate of the plug flow reactor.

The thermodynamic properties of each species are described by NASA polynomials and are equal for both models. These properties were calculated based on databases containing thermochemical properties of species using the best available *ab initio* estimates. If the properties of a species were not found in this database, the enthalpy of formation, entropy and heat capacity were calculated using the Benson's group additivity method with GAV's originating from literature.[21,22,42]

The kinetics for the first model were estimated using a group additivity method for the Arrhenius parameters as described in section 3.1.3. These estimates are also reported in Appendix A. Next to these values, the rate parameters from the NIST database are listed. After comparing the pre-exponential factor of each elementary reaction it is concluded that for only nine of the 33 elementary reactions the order of magnitude is equal to each other. For the other reactions, the order of magnitude is in close proximity. The activation energies of each elementary reaction differ less compared to the pre-exponential factors. The NIST kinetics database contains a collection of reported experimental and computational values

from literature. There are no recommended values for an elementary reaction and therefore the Arrhenius parameters were chosen based upon the temperature range in which they were valid. The earliest reported value that had a temperature range containing the temperature profile as described in Figure 3-9 was taken from the database.

Figure 3-10 shows molar fractions as a function of the residence time for ethane and the most important products such as hydrogen, ethylene and methane. For none of the compounds both models are in good agreement. The model based on the kinetics from the NIST database overestimates the molar fractions of the products hydrogen, ethylene and methane while it underestimates the molar fraction of the reactant ethane. The sudden change in molar fraction between a residence time of 0.5 and 0.6 seconds can be explained by the sudden drop in temperature as shown in Figure 3-9. This comparison shows that a modeler should take into account the influence of the kinetic data that is used to construct the kinetic model.
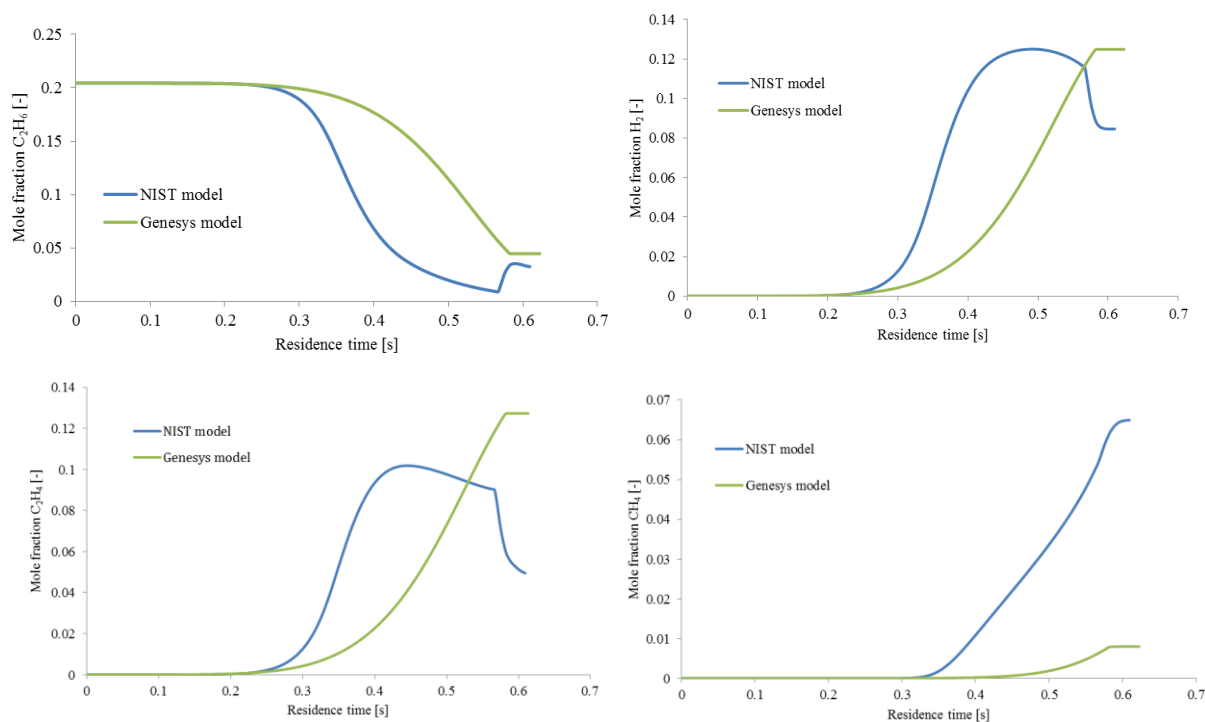
Figure 3-10: Mole fraction of ethane (top, left), hydrogen (top, right), ethylene (bottom, left) and methane (bottom, right) as a function of the residence time.

3.4. Comparison of kinetic models for steam cracking of gaseous feeds

Simulating the quantitative results from experiments with kinetic models remains a challenging task.[43] Most of the uncertainties of a model are caused by the uncertainties of rate parameters and sometimes also by the uncertainty of thermodynamic data.[44,45] To investigate the influence of kinetics on model predictions, two different kinetic models are constructed for the steam cracking of various C_1-C_4 hydrocarbon feedstocks. The first model is generated with the rule-based kinetic model builder Genesys[46]. The rate parameters are based on the group additivity estimation method. This method relies on *ab initio* CBS-QB3 rate coefficients for reference reactions and temperature independent group additive values. For some elementary reactions such as the recombination, Diels-Alder and cyclization reactions there are no GAV's available in the database. For these reactions calculated CBS-QB3 rate coefficients were used instead. The generated model consists solely of elementary reactions. There are in total 116 species and 556 reactions of which not all are important for the production of the major C_1-C_4 cracking products. Via principle component analysis the model was reduced to 77 species and 340 irreversible reactions. The simulated results of the reduced model matched those of the "full" model. The model was further reduced by assuming that all radicals of carbon number three and higher behave as μ-radicals. This means that these radicals solely undergo monomolecular decomposition reactions, reducing the total number of irreversible reactions from 340 to 277.

The second kinetic model was generated with RMG, the kinetic model builder developed by Green et al[45,47]. This model builder uses the rate-based approach to build a network of important species and reactions. The kinetics are derived from a combination of highly accurate *ab initio* calculations and estimation methods. The predictions of both models are compared to the results of 113 experiments with a feed of C_1-C_4 paraffin/olefin mixture at steam cracking conditions. The coil has an internal diameter of 1 cm and a length of 23.137 m. The steam dilution was in the range of 0-50 wt% while the hydrocarbon feed flow was in the range of 2 to 5 kg/hr. The feed composition, the temperature and pressure profile were varied for each experiment. This resulted in product flow compositions with hydrogen, methane, ethane, ethylene, propylene, propane, n-butane, isobutane, isobutene, n-butene, acetylene, propadiene, 1,3-cyclopentadiene, 1,3-butadiene and benzene as main products. In Figure 3-11 the parity plots for the mass fraction of ethylene, propylene, butadiene and hydrogen are given for both models. The kinetic model constructed by Genesys predicts the experimental results better than the RMG model. This difference in predicted values can be caused by the method used to construct the kinetic model (rule-based or rate-based) or by the kinetic data. Another explanation is the influence of the thermodynamic data. For the reversible reactions the reverse rate coefficient is calculated via the thermodynamic consistency. Using the thermodynamic data calculated by Genesys for the model generated with RMG could result in additional deviations. Either way, these results show that the impact

of kinetic and thermodynamic data on model predictions cannot be neglected. More results for the comparison of the Genesys model with other models and experimental results can be found in the work of Murali et al.[48]
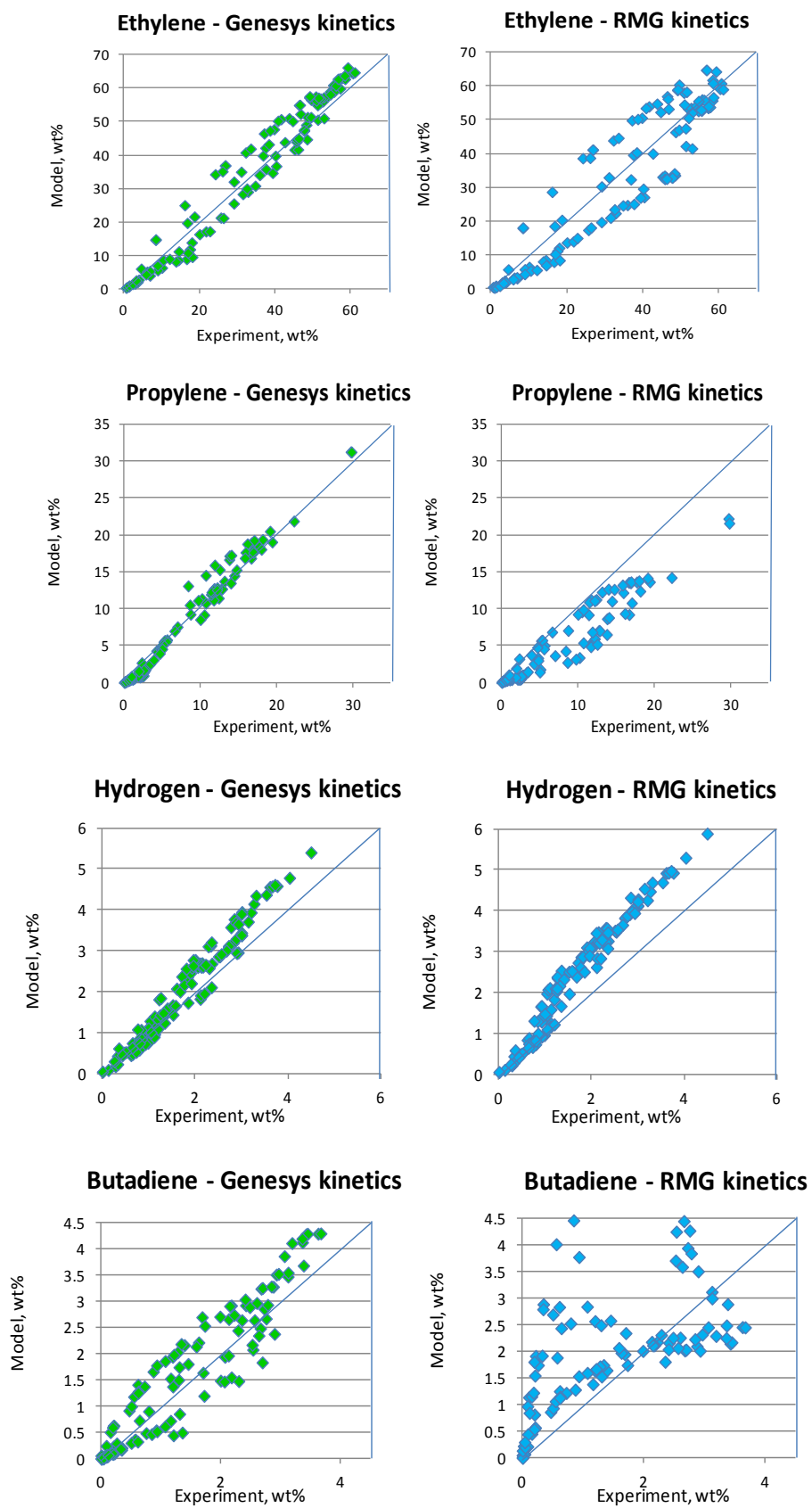
Figure 3-11: Parity plots for some major products (ethylene, propylene, hydrogen and butadiene) of the steam cracking of C1-C4 feed for the kinetic model generated with Genesys (left) and RMG (right).[48]

3.5. Conclusion

When a rate-based algorithm is applied, it is necessary to provide accurate kinetic parameters. These parameters show whether a reaction should be added to the model or not. If a certain elementary step is left out because the rate of formation was to low due to inaccurate kinetic parameters, important reaction pathways can be forgotten and this results in an incomplete kinetic model.

If the model is incomplete, commonly one tries data fitting and adjust/force the kinetic parameters to compensate for the elementary steps that were left out (this causes the physicochemical basis to be diminished). Off course, the result of these adjustments depend on the available experimental datasets, and extrapolations outside the covered conditions such as temperature, pressure, reactor geometries, etc. become impossible. Even if a kinetic model is complete, accurate thermochemical data is required to solve the model in combination with a reactor model. Otherwise the simulated concentration profiles will deviate from experimentally obtained results.

For successfully modeling a chemical reaction, it is important to look at the compatibility of different data. Mostly, many of these sources have a different error variance which could influence the generated kinetic model. It is therefore necessary for the modeler to keep this in mind while evaluating the results. Little is known about the extent of this influence.

About de degree of accuracy for rate coefficients there is no straightforward answer. Mainly this all depends on how accurate the user wants its model to be. A kinetic model can't be more accurate than the rate coefficients it employs. There will always be a trade-off between the accuracy of the information and the time spent to obtain these values. Nevertheless, when the scientific community moves towards open-source information, the amount of available information can expand rapidly and the accuracy will increase as well.

3.6. References

1. Sabbe MK, Van Geem KM, Reyniers MF, Marin GB. First Principle-Based Simulation of Ethane Steam Cracking. *Aiche J.* Feb 2011;57(2):482-496.

2. Vandeputte AG, Reyniers M-F, Marin GB. A theoretical study of the thermodynamics and kinetics of small organosulfur compounds. *Theoretical Chemistry Accounts.* Aug 2009;123(5-6):391-412.

3. Law CK. Combustion at a crossroads: Status and prospects. *Proc. Combust. Inst.* 2007;31(1):1-29.

4. Baulch DL. Chapter 3 Kinetics databases. In: Pilling MJ, ed. *Comprehensive Chemical Kinetics.* Vol Volume 35: Elsevier; 1997:235-292.

5. Carstensen H-H, Dean AM. Rate Constant Rules for the Automated Generation of Gas-Phase Reaction Mechanisms†. *The Journal of Physical Chemistry A.* 2009/01/15 2008;113(2):367-380.

6. Battin-Leclerc F, Blurock E, Bounaceur R, et al. Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models. *Chem. Soc. Rev.* 2011;40(9):4762-4782.

7. Battin-Leclerc F, Simmie JM, Blurock E. *Cleaner Combustion: Developing Detailed Chemical Kinetic Models*: Springer London; 2013.

8. Pilling MJ. From elementary reactions to evaluated chemical mechanisms for combustion models. *Proc. Combust. Inst.* 2009;32(1):27-44.

9. Wagner AF. The challenges of combustion for chemical theory. *Proc. Combust. Inst.* 2002;29(1):1173-1200.

10. Gotwals R. Determination of the Choice of Theoretical Method for a Pyridine/Protonated Pyridine System. *Comp Chem Moodle.* 2007. Accessed 28 octobre 2014.

11. Lewars E. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*: Springer; 2010.

12. Becke AD. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of Chemical Physics.* 1993;98(2):1372-1377.

13. Andrew G. Introduction to Computational Quantum Chemistry: Theory2007.

14. Cramer CJ. *Essentials of Computational Chemistry: Theories and Models*: Wiley; 2013.

15. Baboul AG, Curtiss LA, Redfern PC, Raghavachari K. Gaussian-3 theory using density functional geometries and zero-point energies. *The Journal of Chemical Physics.* 1999;110(16):7650-7657.

16. Montgomery JA, Frisch MJ, Ochterski JW, Petersson GA. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *The Journal of Chemical Physics.* 1999;110(6):2822-2827.

17. Vandeputte AG. *The thermochemistry and decomposition mechanism of organosulfur and organophosphorus compounds*, Ghent University; 2012.

18.   Truhlar DG, Garrett BC, Klippenstein SJ. Current Status of Transition-State Theory. *The Journal of Physical Chemistry.* 1996/01/01 1996;100(31):12771-12800.

19.   Garrett BC, Truhlar DG. Chapter 5 - Variational transition state theory. In: Scuseria CEDFSKE, ed. *Theory and Applications of Computational Chemistry*. Amsterdam: Elsevier; 2005:67-87.

20.   Van Belleghem J. *Kinetic modeling of scission and recombination reactions of hydrocarbons and of O- and S- containing compounds*. Ghent, Ghent University; 2012.

21.   Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J. Phys. Chem. A.* Aug 2005;109(33):7466-7480.

22.   Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *J. Phys. Chem. A.* Nov 2008;112(47):12235-12251.

23.   Ackermann T. Thermochemical kinetics. Methods for the estimation of thermochemical data and rate parameters. Von S. W. Benson. John Wiley & Sons, Inc., New York 1976. 2. Aufl., XI, 320 S., 15 Abb., geb. $ 28.50. *Angewandte Chemie.* 1977;89(12):921-921.

24.   Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Ab initio group contribution method for activation energies for radical additions. *Aiche J.* 2004;50(2):426-444.

25.   Sumathi R, Green WH. A priori rate constants for kinetic modeling. *Theoretical Chemistry Accounts.* Oct 2002;108(4):187-213.

26.   Blowers P, Masel R. Engineering Approximations for Activation Energies in Hydrogen Transfer Reactions. *AIChE Journal.* 2000;46(10):2041-2052.

27.   Blowers P, Masel RI. Extensions of the Marcus equation for the prediction of approximate transition state geometries in hydrogen transfer and methyl transfer reactions. *Theor Chem Acc.* 2000/11/01 2000;105(1):46-54.

28.   Peterson K, Feller D, Dixon D. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges. *Theoretical Chemistry Accounts.* 2012/01/11 2012;131(1):1-20.

29.   Turányi T, Tomlin AS. *Analysis of Kinetic Reaction Mechanisms*: Springer Berlin Heidelberg; 2015.

30.   Tomlin AS, Turányi T, Pilling MJ. Chapter 4 Mathematical tools for the construction, investigation and reduction of combustion mechanisms. In: Pilling MJ, ed. *Comprehensive Chemical Kinetics.* Vol Volume 35: Elsevier; 1997:293-437.

31.   Van de Vijver R. *Development of a topological symmetry identificat*2012.

32.   Hampson RF, Garvin D. Evaluation and compilation of reaction rate data. *The Journal of Physical Chemistry.* 1977/12/01 1977;81(25):2317-2319.

33.   West RH, Magoon GR. Letter to the editor: Electronic data formats for open storage of published chemical kinetic data. *International Journal of Chemical Kinetics.* 2010;42(3):200-201.

34.   Green PWH. Editorial. *International Journal of Chemical Kinetics.* 2009;41(5):295-295.

35. Hann M, Green R. Chemoinformatics - a new name for an old problem? *Curr. Opin. Chem. Biol.* Aug 1999;3(4):379-383.

36. R. Atkinson DLB, R. A. Cox, J. N. Crowley, R. F. Hampson, R. G. Hynes, M. E. Jenkin, M. J. Rossi, and J. Troe. Task Group on Atmospheric Chemical Kinetic Data Evaluation. 2004; http://iupac.pole-ether.fr/, 2015.

37. J. A. Manion, R. E. Huie, R. D. Levin, et al. NIST Chemical Kinetics Database, NIST Standard Reference Database 17. Version 7.0 (Web Version), Release 1.6.8, Data version 2013.03:http://kinetics.nist.gov/

38. Sabbe MK, Vandeputte AG, Reyniers M-F, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics.* 2010;12(6):1278-1298.

39. Michael Harper SS, William Green, Kevin Van Geem. Modeling butanol combustion: Automated mechanism generation utilizing the Process Informatics Model (PrIMe) data depository. Paper presented at: 20th International Symposium on Gas Kinetics2008; Manchester.

40. European Coorperation in Science and Technology. 2015; http://www.cost.eu/COST_Actions. Accessed 23/02/2015, 2015.

41. Pyl SP. *Duurzame productie van lichte alkenen: van fossiele naar hernieuwbare grondstoffen* [Dissertation], Ghent University; 2012.

42. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates. *Chemistry – A European Journal.* 2013;19(48):16431-16452.

43. Sarathy SM, Oßwald P, Hansen N, Kohse-Höinghaus K. Alcohol combustion chemistry. *Prog. Energy Combust. Sci.* 2014;44(0):40-102.

44. Zádor J, Zsély IG, Turányi T, Ratto M, Tarantola S, Saltelli A. Local and Global Uncertainty Analyses of a Methane Flame Model. *The Journal of Physical Chemistry A.* 2005/11/01 2005;109(43):9795-9807.

45. Green WH, Barton PI, Bhattacharjee B, et al. Computer Construction of Detailed Chemical Kinetic Models for Gas-Phase Reactors. *Ind. Eng. Chem. Res.* 2001/11/01 2001;40(23):5362-5370.

46. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chem. Eng. J.* Oct 2012;207:526-538.

47. Song J. *Building robust chemical reaction mechanisms: Next Generation of Automatic model construction software*: Departmen of Chemical Engineering, Massachusetts Institute of Technology; 2004.

48. Khandavilli M. An experimental and group additive kinetic modeling study of steam cracking of gaseous feeds2015.

Chapter 4

Numerical solvers

This chapter introduces the available tools to solve the reaction network during the rate-based network generation process. Important and common tools are CHEMKIN, Cantera and OpenSMOKE. Some of these are open-source while others are commercial and require the purchase of a license.

During the rate-based network generation process, the reaction network is solved each iteration step. Therefore, it is important that the numerical solver can give results in a fast and efficient way. Generating very complex and large mechanisms requires several hours to complete and eventually, the numerical solver will have to deal with large kinetic models that are very time-consuming to simulate. A solver should be chosen that has the right advanced numerical techniques to reduce the computational cost without sacrificing the accuracy and the robustness of the calculations.[1]

4.1. Introduction

A reactor simulation is based on a combined description of the physical and chemical phenomena that occur in a chemical reactor. A kinetic model describes the chemical reactions and a reactor model solves the conservation laws and incorporates the physical transport phenomena.[2] In general, the numerical solver requires a kinetic model containing chemical reactions and rate coefficients, and in some cases thermodynamic data of the included species and transport properties. To simplify the current discussion, abstraction is made of the mass transport limitations. This means that the system is homogeneous and that temperature and composition fields are limited by the kinetics of the system and not by effects of transport phenomena such as diffusion, convection and conduction. Examples where the transport phenomena of species and energy are rate-limiting are diffusion flames and chemical vapor deposition systems.[3,4]

Moreover, if transport limitations are taken into account, the mechanism is described by a set of partially differential equations (PDE). Computational fluid dynamics (CFD) simulations would have to be performed and solving PDE's is much more complex and expensive than solving ordinarily differential equations (ODE). For now, Genesys is used to construct kinetic models for processes such as steam cracking, pyrolysis and combustion. The presence of a homogeneous system is assumed, which valid if

these processes occur in gas-phase reactors with a behavior that closely resembles an ideal reactor. For example, tubular reactors with high Reynolds numbers resemble plug flow reactors.

While neglecting transport phenomena, the reaction mechanism dynamics will be described by a specific set of ODE's. The formation rate of any edge species can be described by the continuity equation given by Equation 4-1

$$r(t) = \frac{dC}{dt} = f(C, t, k(T)) \qquad C(t_0) = C_0$$

4-1

Where C is a vector of concentrations, $k(T)$ is a vector of kinetic parameters, t is the time variable and C_0 is the vector of initial concentrations. This set of ODE's can be solved if the initial conditions and the kinetic parameters are available.

Thermal phenomena in chemical reactors are accounted for based on the law of conservation of energy. This law leads to the following energy balance:

$$\frac{dE_{system}}{dt} = E_{in} - E_{out} \qquad [J.s^{-1}]$$

4-2

The total energy of the system, E_{system}, consists of the internal energy of the gas. Heat supply or removal through the walls of the reactor is neglected. Overall the energy balance remains simple, for example, kinetic and potential energy of the streams as well as surface energy are neglected. Depending on the type of reactor, this balance can take many forms. For instance, the energy equation for a plug flow reactor is reduced to Equation 4-3.

$$c_p \rho_0 \frac{dT}{dz} = c_p \rho_0 \frac{(T_0 - T)}{\tau_0} + \sum_i (-\Delta H_{r,i}) r_i$$

4-3

Where c_p is the specific heat capacity of the mixture, $\Delta H_{r,i}$ the reaction enthalpy of reaction i and ρ_0 the mass density. Finally, the reactor is described as isobaric, meaning that the conservation law for momentum should not be included in the model equations. The software package CHEMKIN has been chosen for this master thesis to provide the necessary numerical solution techniques. This package contains a large set of available reactor models and is, after a steep initial learning curve, user friendly. To implement the rate-based algorithm in the code of Genesys, the ecosystem is extended with two extra modules, namely, the ChemkinParallelizer and the ModelAnalyzer, as can be seen in Figure 4-1.
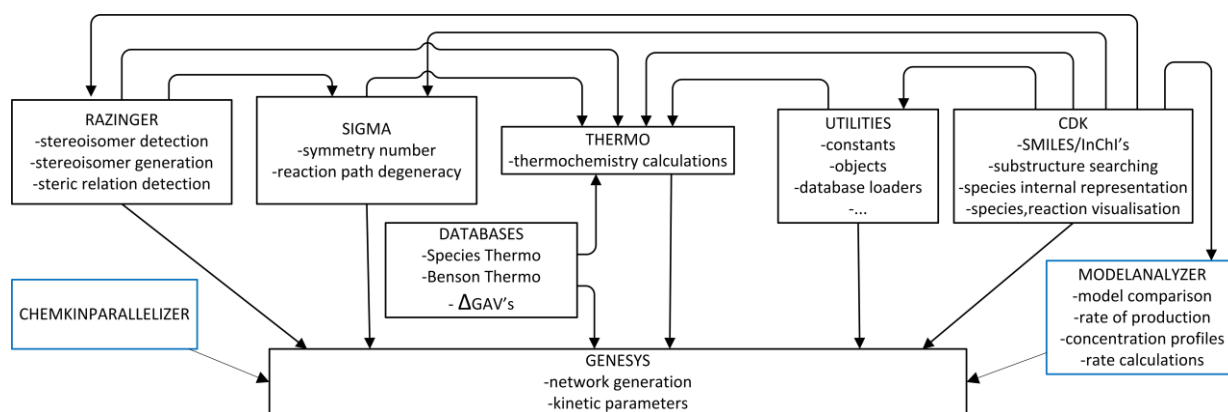
Figure 4-1: Different independent modules present in the kinetic model builder Genesys, extended with the modules ChemkinParallelizer and ModelAnalyzer .

The ChemkinParallelizer is a module written in Java and its purpose is to bypass the CHEMKIN GUI and to run multiple reactor simulations in parallel. The user needs to supply a kinetic model input file, a reactor configurations file, an XML file with the directory of the previously mentioned files and of the CHEMKIN software. Once the ChemkinParallelizer has obtained all the necessary files, it will call the required sub-processes of CHEMKIN itself, the same way the CHEMKIN GUI performs its steps. The benefit if this module is that it is much more prone to automatization, which is ideally for the rate-based algorithm. The option to run parallel simulations will not be used in this project. Since this module is written in Java, the integration with Genesys is straightforward. Instead of using the XML file with configuration input, the code in Genesys is altered to call the main class of the ChemkinParallizer when the network has to be solved. Simultaneously, the required information such as the location of the reactor configuration input and the CHEMKIN directory are provided via program arguments. Once the numerical solver is finished with the calculations, the ChemkinParallelizer writes the solution data to an Excel workbook.

4.2. Stiff numerical solvers

Not only the size of detailed kinetic models is a challenge for the numerical solvers but also the stiffness of non-linear chemical equations. The stiffness is related to the existence of a wide range of characteristic chemical times related to the large number of radical species in quasi steady-state and fast reversible reactions in partial equilibrium.[5] When a set of ordinary differential equations is stiff, special care needs to be taken with the choice of numerical integration scheme. There are many algorithms and software developed to solve ordinary differential equations initial value problems. In this chapter, DASSL and DASPK[6], the two most applied codes used to solve these stiff systems are described. To solve transient systems such as chemical processes occurring in plug flow reactors or CSTR's, the CHEMKIN package relies on the DASPK solver.

4.2.1. DASSL

The code Differential Algebraic System Solver (DASSL)[7,8] is designed to solve implicit systems of differential/algebraic equations (DEA) (Eq. 4-4).

$$F(t, y, y') = 0$$

$$y(t_0) = y_0 \qquad \text{4-4}$$

$$y'(t_0) = y_0'$$

Where F, y and y' are N dimensional vectors and the initial values y_0 and y_0' are given. Standard ODE solvers require the above equations to be rewritten in an explicit form to solve the system for y'. DASSL will solve these equations in their original form. The system is solved in each time step by replacing the derivative in Equation 4-4 via the Backward Differentiation Formulas (BDF)[9] and then solve the resulting non-linear equations for the solution at the current time t_n using Newton's method.[10] Replacing the derivative by its backward difference results in the Equation 4-5.

$$F\left(t_n, y_n, \frac{y_n - y_{n-1}}{\Delta t_n}\right) = 0 \qquad \text{4-5}$$

With Newton's method the following equation is obtained:

$$y_n^{m+1} = y_n^m - \left(\frac{\partial F}{\partial y'} + \frac{1}{\Delta t_n}\frac{\partial F}{\partial y}\right)^{-1} F\left(t_n, y_n, \frac{y_n - y_{n-1}}{\Delta t_n}\right) \qquad \text{4-6}$$

Where m is the iteration index. In the DASSL algorithm, the derivative is approximated using the kth order backward differentiation formula, where k ranges from 1 to five. The order of k and the time stepsize Δt_n are determined during every step based on the behavior of the solution. DASSL will begin

every time step by making an initial guess of the solution and its derivative at t_{n+1}.[11] The initial guess y_n^0 is obtained by evaluating the polynomial which interpolates the computed solution at the last $k+1$ times $t_{n-1}, t_{n-2}, \dots, t_{n-(k+1)}$ at the current time t_n. Equation 4-5 is rewritten as Equation 4-7 to discretize the implicit formula and to solve the non-linear equation efficiently.

$$F(t, y, \hat{\alpha}y + \beta) = 0 \qquad \text{4-7}$$

$\hat{\alpha}$ is a constant, which changes when the stepsize or order of k changes and β is a vector which depends on the solution at past time steps. $t, y, \hat{\alpha}$ and β are evaluated at t_n. The modified version of the Newton method is given by

$$y^{m+1} = y^m - c\left(\frac{\partial F}{\partial y'} + \alpha\frac{\partial F}{\partial y}\right)^{-1} F(t, y^m, \hat{\alpha}y^m + \beta) \qquad \text{4-8}$$

This linear system is solved via a direct linear solver (dense or banded) solver.[12] The iteration matrix G is nothing more than:

$$G = \frac{\partial F}{\partial y'} + \alpha\frac{\partial F}{\partial y} \qquad \text{4-9}$$

This matrix is computed, factorized and is then used for as many time steps as possible. The rate of convergence of Equation 4-5, ρ, is estimated by:

$$\rho = \left(\frac{||y^{m+1} - y^m||}{||y^1 - y^0||}\right)^{\frac{1}{m}} \qquad \text{4-10}$$

The norms are scaled depending on the error tolerance specified by the user. Finally, the iteration has converged when

$$\frac{\rho}{1 - \rho} < ||y^{m+1} - y^m|| < 0.3 \qquad \text{4-11}$$

When $\rho > 0.9$ or $m > 4$ and the iteration has not yet converged, the stepsize is reduced and the step is attempted again.

4.2.2. DASPK

The solver Differential Algebraic System Solver Package (DASPK) has a much more improved initial condition calculation algorithm compared to DASSL. DASPK also includes iterative Krylov methods[13] for the linear systems that arise, in addition to the direct (dense/banded) methods in DASSL.[5,14]

The DASPK solver is based on the integration methods in the solver DASSL. To find a numerical solution for the large system of differential algebraic equations described by Equation 4-4, the dense or banded direct linear system solvers of DASSL are replaced with a preconditioned iterative method. For this method, there is always a preconditioner needed for the DAE's. The integration methods and strategies for time-stepping are identical to those described in section 4.2.1. The main difference lies within the calculation of the solution for matrix G (Equation 4-9). With the DASPK solver, a preconditioner matrix P is used which is an approximation of matrix G.[8] This matrix is also used as many times as possible. The main advantage is that the matrix G does not have to be stored and calculated each iteration step, only the value of G multiplied with a vector product has to be stored. This type of methods used to solve large-scale linear and sparse system is referred to as the polynomial acceleration method or the Krylov method.

Both DASSL and DASPK are applied by the CHEMKIN package to solve large sets of ODE's. It is not the goal of this chapter to go more into detail about the specific algorithms used in these solvers. For a more detailed explanation about these solvers, the reader is referred to the work of Petzold, Brenan and Brown.[7,8,14]

4.3. CHEMKIN

CHEMKIN is a program package suitable for the simulation of gas kinetic systems.[15] It was first developed for internal research by the group from the Sandia National Laboratory in 1975. In 1995 the rights of CHEMKIN were transferred to ReactionDesign which further developed CHEMKIN as a commercial code. A Fortran subroutine library is at the center of the program. This library facilitates the incorporation of complex chemical kinetics into simulations of reacting flows under various modeling scenarios.[16] A brief description of the steady state solver TWOPNT that is developed by ReactionDesign follows. This solver is important when a CSTR in steady state is chosen for the chemical system.

4.3.1. TWOPNT

The numerical solver TWOPNT is used to determine the solution of a set of algebraic equations for steady-state problems concerning homogeneous systems. This type of problem is solved by first applying a damped modified Newton algorithm to the set of non-linear algebraic equations. These algebraic equations are the global mass balance equation, the species equations and optionally the gas energy equation. The modified damped Newton's method performs a set of iterations to find the solution vector ϕ, such that

$$F(\phi) = 0 \qquad\qquad 4\text{-}12$$

when it is substituted in the governing equations. The solution vector contains the gas temperature T and the gas species mass fractions Y. The approximate solution is obtained during each iteration via Equation 4-13.[10]

$$\phi^{(n+1)} = \phi^{(n)} - \left(\frac{\partial F}{\partial \phi}\right)^{-1}_{\phi^{(n)}} F(\phi^{(n)}) \qquad\qquad 4\text{-}13$$

The advancement of the approximate solution $\phi^{(n)}$ to the next approximate solution $\phi^{(n+1)}$ is damped by a factor $\lambda^{(n)}$.[3] The modified algorithm becomes then

$$\phi^{(n+1)} = \phi^{(n)} - \lambda^{(n)} \left(\frac{\partial F}{\partial \phi}\right)^{-1}_{\phi^{(n)}} F(\phi^{(n)}) \qquad\qquad 4\text{-}14$$

where $0 < \lambda^{(n)} < 1$. The Jacobian matrix $J^{(n)}$ is equal to $\frac{\partial F}{\partial \phi}$ and is time consuming to solve. Therefore, the Jacobian matrix is not generated every iteration step but instead, it is kept constant until the solution does not converge anymore. Another option is to reduce the dampening factor by half if convergence fails. With a user-defined absolute and relative tolerance, a standard convergence criterion is defined.

When the entire Newton algorithm fails to converge, a time stepping procedure is used. This procedure will integrate the time-dependent version of these equations over a fixed number of time steps in order to find an estimate for the solution. With this new starting estimate, which is closer to the steady-state solution, the Newton algorithm is applied again. This increases the likelihood of convergence for the Newton method. If however the algorithm fails again, additional time steps on the transient solution are taken to further improve the initial estimate. In the end the Newton iteration method converges on the steady state solution.

If no initial guess of the solution is provided by the user, the solver will calculate the equilibrium composition by minimizing the Gibb's free energy of the system. These equilibrium gas-phase mole fractions are used as the starting initial guess. When the user has not specified a fixed temperature of the systems and wants the energy equation to be solved, the solver will use a two step approach. First, the species composition at a fixed temperature will be calculated. Next, the obtained solution will be used as initial guess to solve simultaneously the energy equation and the species composition. This two-step method is more robust for thermal systems.

4.3.2. Reactor Type

In CHEMKIN the reactor types listed in Table 4-1 are available.

Table 4-1: Available reactor models in CHEMKIN.

Closed 0-D reactors	Internal combustion engine
	Homogeneous batch reactor
	Partially stirred reactor
	Plasma reactor
Open 0-D reactors	Perfectly stirred reactor (CSTR)
	Plasma perfectly stirred reactor
	Partially perfectly stirred reactor
Flow reactors	Plug flow reactor
	Plasma plug flow reactor
	Planar shear flow
	Cylindrical shear flow
	Honeycomb monolith reactor
Flame simulators	Premixed laminar burner-stabilized flame simulator
	Flame speed simulator
	Opposed-flow Flame simulator
Chemical vapor deposition reactors	Stagnation flow CVD reactor
	Rotating disk CVD reactor
Shock tube reactors	Normal incident shock
	Normal reflected shock

For now, only the plug flow reactor, homogeneous batch reactor and perfectly stirred reactor are supported by the ChemkinParallelizer. Also the ignition delay and the flame speed can be calculated but

are of no importance to the rate-based algorithm. There are several issues worth mentioning concerning the characteristics of each ideal reactor type. First of all, by default, an isobaric and isothermal homogeneous batch reactor is chosen to numerical solve the network during the rate-based network generation process, meaning that the energy equation should not be solved. For radical chemistry the corresponding set of differential equations is inherently stiff which can lead to a situation where the solver might not converge towards a solution. Simplifying the reactor model as much as possible makes the calculations less complex and has a positive effect on the convergence of the system. Secondly, the mathematical equations for a batch reactor and a steady state plug flow reactor are analogous besides the batch time t and the residence time τ_0 meaning that the concentration profiles are equal. If two models were to be generated with a batch and a steady-state PFR, the reactions and species included in the models should be equal to each other. For a CSTR, the content is perfectly mixed and the bulk concentration for every species is assumed. These bulk concentrations influence the rates of the occurring reactions. How this reactor model influences the generated kinetic model is studied in section 5.4.3.

4.4. Other numerical solver tools

The CHEMKIN software tool provides the user with a vast choice of reactor models and options, however a license is required. Preferably one would like to be independent of licenses and therefore open-source tools could provide the answer. In the next section two solver tools are briefly introduced which are possible alternatives for CHEMKIN.

4.4.1. OpenSMOKE

OpenSMOKE is a collection of open-source C++ libraries that has specifically been developed to manage large, detailed kinetic schemes. The software does not have a graphical interface. A graphical postprocessor is present that can rapidly obtain temperature or species profiles. The output is provided in TXT and XML files. These XML files can however only be read by the graphical post processor of OpenSMOKE which is a downside if one would like to use the results directly in another software tool. There are a large number of ODE solvers available in OpenSMOKE, including the DAPSK solver. The OpenSMOKE tools rely on the Computation Cost Minimization (CCM) techniques.[1] CCM refers to a set of strategies that optimize the efficiency of simulations with little to no loss of accuracy.[17] A list of the reactor models available in OpenSMOKE is given in Table 4-2.

Table 4-2: Available reactor models in OpenSMOKE

Closed 0-D reactors	Homogeneous batch reactor
Open 0-D reactors	Perfectly stirred reactor
Flow reactors	Plug flow reactor
Flames	Premixed laminar burned-stabilized flame
	Opposed diffusion flame

A comparison between the computational results of OpenSMOKE and CHEMKIN for the simulation of ethanol pyrolysis has been performed in literature.[18] In Figure 4-2 some results can be seen showing a perfect superposition for the results of both software tools.
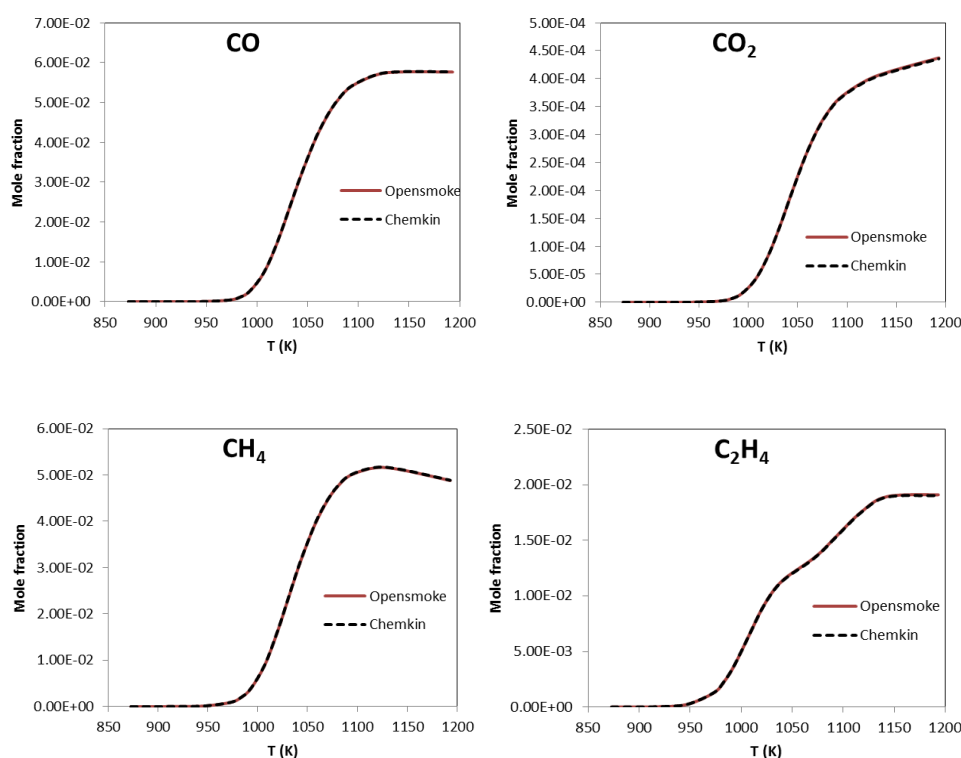
Figure 4-2: Pyrolysis of ethanol, simulated using CHEMKIN and OpenSMOKE. The kinetic model is obtained from the work of Tran et al.[19]

Recently the OpenSMOKE framework was incorporated into the laminarSMOKE solver for the simulation of multi-dimensional laminar flames.[20] The calculation of thermodynamic and transport properties, as well as the integration of independent stiff ODE systems representing the chemical reactions is managed by OpenSMOKE.

4.4.2. Cantera

The last software discussed in this chapter is the open-source Cantera[21]. It is currently used for applications regarding combustion, electrochemical energy conversion, various types of plasma and chemical vapor deposition. It is a very useful and free alternative to the CHEMKIN software package of which the community can participate in the further development of Cantera. Cantera can be seen as a collection of object-oriented software tools.[22] Cantera is linkable with most common programming languages making it possible for the user to build a new reactor model. This is not possible with software tools such as CHEMKIN and OpenSMOKE. A comparison of the simulation results of CHEMKIN and Cantera against experimentally obtained results for the ethanol combustion in a flame burner show again a perfect superposition (Figure 4-3).[18]
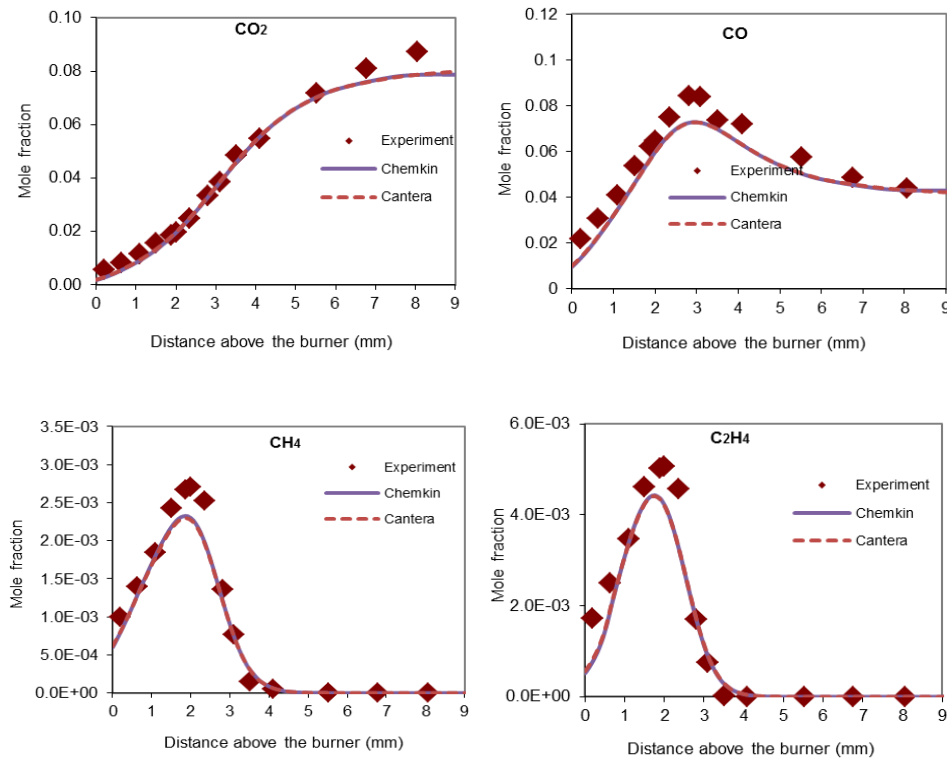
Figure 4-3: Combustion of ethanol in a premixed flame, simulated using CHEMKIN and Cantera. Both simulated results are compared against experimental data from Tran et al.[19]

The available reactor types in Cantera are listed in Table 4-3. As mentioned before, own reactor models can be programmed and used with Cantera. A downside however is the less user-friendly interface compared to CHEMKIN and the lack of code documentation and technical support.[18]

Table 4-3: Available reactor models in Cantera

Closed 0-D reactors	Homogeneous batch reactor
Open 0-D reactors	Perfectly stirred reactor
Flow reactors	Plug flow reactor
	Monolith reactor
Flames	Premixed laminar burned-stabilized flame
	Premixed laminar burned-stabilized stagnation flame
	Opposed diffusion flame
	Flame-speed

4.5. KPPA

During this master thesis, the program KPPA (Kinetics PreProcessor: Accelerated) was studied. The process of solving a chemical system can be split into two steps. First, the chemical reactions are translated into differential equations. Secondly, these differential equations are numerically solved. KPPA is a preprocessor that uses a totally integrated method to do both tasks. The required input of KPPA is a kinetic model description and a description of the target architecture (meaning which numerical integrator scheme is preferred to solve the system). The chemical mechanism is described in a specific KPPA language, thus it is first required to transform the kinetic model into the correct format. KPPA is designed to parse these equations and generate an appropriate code with the chemical kinetic solvers in FORTRAN, Matlab or C. The Jacobian matrix of the ODE system in either spare or full format is generated as well. KPPA is written in the C++ programming language, it is not an open-source code and it is profoundly based on the program KPP (Kinetic PreProcessor). The goal is to facilitate the numerical solution of chemical reaction network problems.[23] In Figure 4-4 the program flow of KPPA is illustrated.
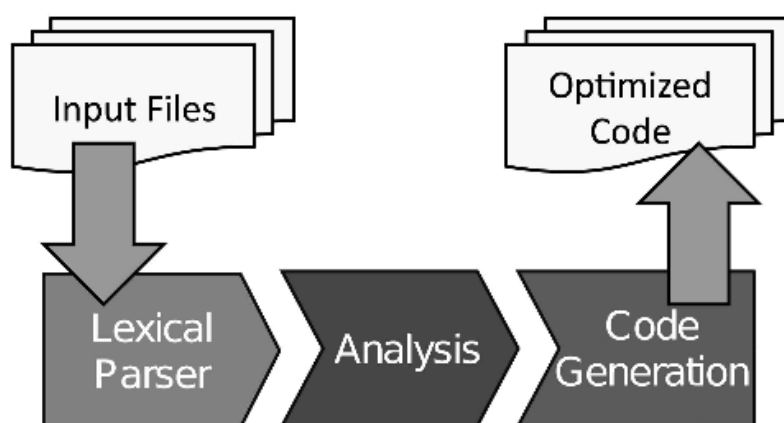
Figure 4-4: Program flow of KPPA and the principle components.

To understand the functionalities of KPPA, the next paragraph will go more into detail concerning the functionalities of KPP, since KPPA can be seen as the next generation of KPP.

4.5.1. KPP

The computer code KPP is a simulation program specific for chemical kinetics problems and has been primarily used for atmospheric chemistry.[24] Based on a reaction mechanism, the code generates a kinetic system of differential equations by assuming mass action kinetics.[15] The latter is the chemical kinetic law of mass action first formulated by Waage and Guldberg[25] to calculate rates of elementary reactions. According to this law a rate can be written as Equation 4-15.

$$r = k(T) \prod_{j=1}^{N_s} Y_j^{v_j} \qquad \text{4-15}$$

Where N_s is the number of species, Y_j is the molar concentration of species j and v_j its stoichiometric coefficient. This equation looks similar to Equation 4-16,

$$r = k(T) \prod_{j=1}^{N_s} Y_j^{\alpha_j} \qquad \text{4-16}$$

Where α_j is a positive real number or zero and equals the reaction order with respect to species j. For elementary reactions the reaction order α_j and the stoichiometric coefficient v_j are commonly the same. For an overall reaction equation (such as $2H_2 + O_2 \leftrightharpoons 2H_2O$) this is not the case, since the overall reaction consists of several intermediate steps.

The ODE and the Jacobian of the system are afterwards transported in formats such as Fortran 77, Fortran 90, Matlab or C. To solve stiff ODE systems, KPP provides methods implementing the spare matrix routines.[26]

This program is mainly used to generate chemical systems code for atmospheric chemistry. In the atmosphere there is a coexistence of very stable (CH_4) and very reactive species. Implicit integration methods are suitable to solve stiff chemical systems. To do so, they require to evaluation of the Jacobian of the derivate function (Equation 4-17)

$$J = \frac{\partial f(t,y)}{\partial y} = \begin{bmatrix} \dfrac{\partial f_1}{\partial y_1} & \dfrac{\partial f_1}{\partial y_2} & \cdots & \dfrac{\partial f_1}{\partial y_n} \\ \dfrac{\partial f_2}{\partial y_1} & \dfrac{\partial f_2}{\partial y_2} & \cdots & \dfrac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial y_1} & \dfrac{\partial f_n}{\partial y_2} & \cdots & \dfrac{\partial f_n}{\partial y_n} \end{bmatrix} \qquad \text{4-17}$$

It is important to mention that KPP is used to simulate spatially homogeneous reaction kinetic systems with user-defined functions utilized for the rate parameters. For systems characterized by rate-coefficients that may change substantially as function of the temperature, pressure and possibly also the gas composition, the temperature and pressure during the course of the reaction should be calculated as well. This is however not possible with KPP.[15]

4.6. Conclusion

To implement the rate-based algorithm in the code of Genesys, numerical techniques are required. These numerical techniques are indispensable to solve the reaction network during each iteration step. For this reason, CHEMKIN has been implemented via the ChemkinParallelizer module. By default the reactor model describes an isobaric and isothermal homogeneous batch reactor. However, the user can choose the reactor model just as he or she wishes and the solution data is obtained in such a way that it is relatively easy to interpret. In the next chapter the influence of the chosen reactor model on the generated kinetic model is studied.

The CHEMKIN software package contains complicated algorithms such as the TWOPNT solver for stead-state problems and the DASSL and DASPK solver for transient problems. Especially the latter are known to handle stiff ordinary differential equations initial value problems in an orderly fashion. CHEMKIN provides the user with a vast choice of reactor models and options, there is however a license required. Other software tools such as Cantera and OpenSMOKE contain less reactor models but require no license. The code of Cantera is open-source meaning that the user can implement its own reactor model. A comparison of the three software tools has shown a perfect superposition of the simulated results. Meaning that the other software tools besides CHEMKIN could be used for the rate-based algorithm without losing any accuracy. Finally the program KPPA was studied for this master thesis. The program contains a broad variety of different solvers which are usefull to solve large chemical systems. However, for processes regarding combustion, pyrolysis and oxidation, the program is less applicable

4.7. References

1. Cuoci A, Frassoldati A, Faravelli T, Ranzi E. OpenSMOKE++: An object-oriented framework for the numerical modeling of reactive systems with detailed kinetic mechanisms. *Computer Physics Communications.* (0).

2. Sabbe MK, Van Geem KM, Reyniers MF, Marin GB. First Principle-Based Simulation of Ethane Steam Cracking. *Aiche J.* Feb 2011;57(2):482-496.

3. CHEMKIN Theory manual: Reaction Design: San Diego; 2010.

4. CHEMKIN 10101 Input Manual: Reaction Design: San Diego; 2010.

5. Lam SH. Using CSP to Understand Complex Chemical Kinetics. *Combustion Science and Technology.* 1992;89:375-404.

6. Li S, Petzold L. Software and algorithms for sensitivity analysis of large-scale differential algebraic systems. *Journal of Computational and Applied Mathematics.* 2000;125(1–2):131-145.

7. Petzold LR. *Description of DASSL: a differential/algebraic system solver*: Sandia National Labs., Livermore, CA (USA);1982.

8. Brenan KE, Campbell SL, Petzold LR. *Numerical Solution of Initial-value Problems in Differential-algebraic Equations*: Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104); 1996.

9. Shampine LF. *Numerical Solution of Ordinary Differential Equations*: Taylor & Francis; 1994.

10. Gilat A, Subramaniam V. *Numerical methods for engineers and scientists : an introduction with applications using MATLAB*. Hoboken, NJ: Wiley; 2011.

11. Liu L, Felgner F, Frey G. Comparison of 4 numerical solvers for stiff and hybrid systems simulation. Paper presented at: Emerging Technologies and Factory Automation (ETFA), 2010 IEEE Conference on; 13-16 Sept. 2010, 2010.

12. Soetaert K, Cash J, Mazzia F. *Solving Differential Equations in R*: Springer; 2012.

13. Gutknecht M. A Brief Introduction to Krylov Space Methods for Solving Linear Systems. In: Kaneda Y, Kawamura H, Sasai M, eds. *Frontiers of Computational Science*: Springer Berlin Heidelberg; 2007:53-62.

14. Brown PN, Hindmarsh AC, Petzold LR. Using Krylov methods in the solution of large-scale differential-algebraic systems. *SIAM J. Sci. Comput.* 1994;15(6):1467-1488.

15. Turányi T, Tomlin AS. *Analysis of Kinetic Reaction Mechanisms*: Springer Berlin Heidelberg; 2015.

16. Kee RJ, Rupley FM, Miller JA. *Chemkin-II: A Fortran chemical kinetics package for the analysis of gas-phase chemical kinetics*1989. SAND-89-8009; Other: ON: DE90000917 United StatesOther: ON: DE90000917Thu Feb 07 01:43:28 EST 2008NTIS, PC A07/MF A01 - OSTI; GPO Dep.SNL; EDB-89-153716English.

17. Lu TF, Law CK. Toward accommodating realistic fuel chemistry in large-scale computations. *Prog. Energy Combust. Sci.* Apr 2009;35(2):192-215.

18. tran L-S. Comparison: Cantera-OpenSMOKE-Chemkin2014.

19. Tran LS, Glaude PA, Fournet R, Battin-Leclerc F. Experimental and Modeling Study of Premixed Laminar Flames of Ethanol and Methane. *Energy & Fuels.* Apr 2013;27(4):2226-2245.

20. Cuoci A, Frassoldati A, Faravelli T, Ranzi E. A computational tool for the detailed kinetic modeling of laminar flames: Application to C2H4/CH4 coflow flames. *Combust. Flame.* 2013;160(5):870-886.

21. *Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes.* [computer program]. Version 2.1.22014.

22. developers C. Cantera Introduction. 2012; http://www.cantera.org/docs/sphinx/html/cti/intro.html. Accessed 1st May, 2015 2015.

23. Zhang H, Linford JC, Sandu A, Sander R. Chemical Mechanism Solvers in Air Quality Models. *Atmosphere.* 2011;2(3):510-532.

24. Sandu A, Sander R. Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1. *Atmos. Chem. Phys.* 2006;6(1):187-195.

25. Waage P, Gulberg CM. Studies concerning affinity. *Journal of Chemical Education.* 1986/12/01 1986;63(12):1044.

26. Linford JC, Michalakes J, Vachharajani M, Sandu A. Automatic Generation of Multicore Chemical Kernels. *Parallel and Distributed Systems, IEEE Transactions on.* 2011;22(1):119-131.

Chapter 5
Rate-based algorithm

5.1. Introduction

In the previous chapters was concluded that the rule-based and rate-based algorithm are two appropriate methods to terminate the automatic network generation and to include only significant species in the kinetic model. The rule-based method applies molecular and atomic constraints for each reaction family based upon expert knowledge concerning specific chemical mechanisms. The downside of this method is that these rules have to be supported by experimental evidence. In this chapter a new algorithm will be considered that extends the existing rule-based network generation process. This rate-based algorithm is already discussed briefly in chapter 2 and hence, the goal of this chapter is not to explain the method again but to discuss how its characteristics are used in Genesys. After the algorithm is implemented, influences of parameters on the performance of the network generation will be evaluated. Whether a combination of both the rate- and rule-based method or solely one method will predict values that are closest in range with experiments is discussed as well.

The basic idea is to use dynamic properties of species to determine if they are of importance for the generated kinetic model. For example, only the newly generated species with a flux high enough at a certain set of reaction conditions are selected, and this logical reasoning guarantees that no insignificant species are part of the model, which is in reality also the case. The rate-based algorithm is illustrated in Figure 2-8 for the pyrolysis of n-butanol. Butanol decomposes into a 1-hydroxyd-butyl radical which can react further into butyraldehyde or into an ethylradical and ethanol. According to Harper[1], the reaction rate of the latter is the highest which means that these species are added to the network. The second model is now obtained and after reacting the two new species against the reaction families, ethylene and acetaldehyde are added to the edge species. As can be seen from the third model, the rates of the three reactions that form the edge species ethanol, acetaldehyde and butyraldehyde are compared to each other.

When the chemistry of a mechanism is not completely understood, it is important that only the correct species are included in the model based upon their flux or concentration. To prevent endless addition of species with the highest flux, two termination criteria are added to the new algorithm. First, the flux of the selected species should be higher than a minimum rate and secondly, the conversion of the initial

reaction should not exceed a predefined value. The minimum rate can be adjusted by the user via a so called precision level, ε. Both termination criteria are discussed later in this chapter. Harper et al. reported that during the network generation for the pyrolysis of butanol, the rate-based kinetic model builder RMG runs into memory limitations if the precision level is lower than 0.01.[1] When the precision level is lowered, the minimum rate decreases and more of the species are added to the model. Hence, more edge species and reactions can be generated. Even though these species are not part of the model, the mechanism generator has to keep track of every edge species and calculate their dynamic properties. This requires a lot of memory capacity. De Witt et al.[2] investigated the influence of the precision level on the number of generated species and the total number of significant species. For the pyrolysis of tetradecane the level had to be reduced until 5E-05 before all of the experimentally important species were included in the model. Figure 5-1 illustrates that if the level is further reduced under 1.0E-05, the number of reactive species and edge species increases significantly.[2] The lowest curve indicates the number of edge species that are generated. When the precision level is equal to 1 there are already more than 5000 edge species generated. This number increases with decreasing precision level until the level is equal to 7E-06. With this precision level almost 20 000 edge species are generated. The dotted curve shows the number of species that are present in the generated kinetic model. When looking at the right axis it appears that this number increases as well with decreasing precision level, but remains two orders of magnitude smaller than the number of edge species. Moreover, with an increasing amount of species the computational time increases as well. Eventually, 98% of the edge species and 80% of the reactions are not included in the reaction network.
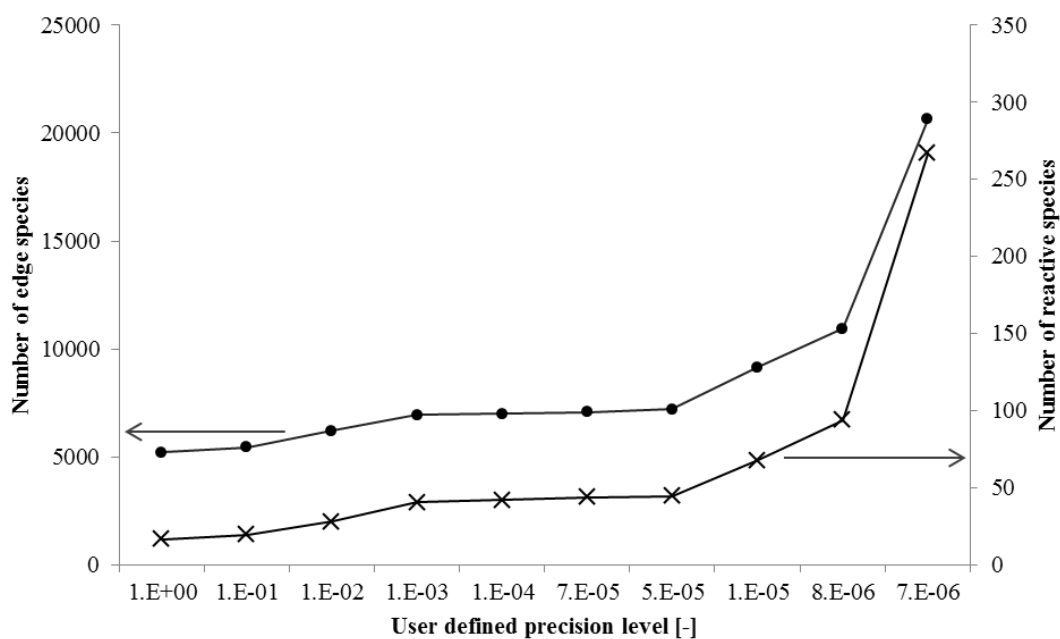
Figure 5-1: Number of edge species (left axis) and reactive species(right axis) as function of the user-defined precision level

The rate-based method has several downsides that should be overcome. Not only could the scarcity of accurate rate coefficients result in the selection of wrong edge species during the network generation, as was mentioned in chapter 2. The amount of generated edge species can become too high when the precision level is too small. A combination of the rule-based and rate-based method should resolve the memory problems while only significant species are added to the network. Limiting the amount of edge species that are generated with the constraints of the reaction families is a promising method.

5.2. Characteristics of the rate-based algorithm

In Genesys the generation scheme that is used to iteratively expand the model is already present. During the network generation, all possible reactions are determined by repeatedly applying the available reaction families provided by the user. First, these families are applied to the starting pool of molecules and afterwards to new molecules that have not reacted yet. With the rate-based algorithm a crucial change is made at this stage of the process. Instead of adding the newly derived species to the unreacted species list, they are added to the edge species list. The reactions that form these edge species as products are added to the partial network but are marked as forward reactions instead of reversible reactions. As long as there are species present in the network that have not reacted yet, some important reactions might not have been generated yet and therefore the network is called "partial". Afterwards, the kinetic parameters of the partial network and the thermodynamic properties of all the species are calculated. Either these values are obtained from databases or estimated via procedures described in chapter 3. This is a second major change in the network generation process. Where the assignment of thermochemical and kinetic parameters was executed after generation of the entire reaction network, it is now necessary to perform this step earlier since these properties are essential to numerically solve the reaction mechanism. By default the system is assumed homogeneous and the kinetic model will be solved for an isothermal batch reactor with user specified operating conditions. As described in the previous chapter, CHEMKIN has a large choice of reactor models and hence not only the default reactor system but also other ideal reactors (such as PFR and CSTR) can be used to obtain the dynamic properties of the partial network. This means that the generated kinetic model is tailored to the specific process and thus predictions of the output yields will be more accurate. However, if one would like to use such a specific model for a different range of operating conditions, the model might become inadequate. How the ordinary differential equations that describe the system are solved is thoroughly described in the chapter 4.

5.2.1. Rate of production

Edge species will be selected using their rate of production. This dynamic property is easy to calculate once the molar fraction of every species is known. Firs the reaction rate of every reaction is calculated based on the rate coefficients in the kinetic model and the concentrations of the reactants and products. If the reaction is reversible, the net reaction rate is equal to the forward rate minus the reverse rate, which is obtained via the equilibrium coefficient K_p. If the reaction is implemented as irreversible, the net reaction rate is equal to the forward rate. For a CSTR the steady state concentrations are available while for a batch reactor or PFR the concentrations at each integration step, time or distance, are used. For every species i present in the system the total rate of production, R_i, is a linear combination of the rates of reactions wherein this species takes part.

$$R_i = \sum_k v_{ik} r_k \qquad\qquad \text{5-1}$$

When a species is consumed or formed during a reaction; the reaction rate is added to respectively the total disappearance rate or the total production rate of the species. In practice, there are no reactions that consume the edge species present in the partial network and therefore the edge species total disappearance rate is equal to zero. As mentioned before, because of the steady state assumption in a CSTR, there will only be one molar fraction for every species in the network. This means that the concentrations and reaction rates are easily calculated. For other ideal reactors, the molar fractions are numerically obtained across the entire length or time interval of the system. To get the average rates, this set of discrete points is numerically integrated and divided by the number of subintervals. The closed trapezoidal method is used to approximate the integrand with a linear function that is easily integrated.[3] The trapezoidal rule is based on Newton's form of interpolating polynomials and is given by the general Equation 5-2. The integral over the interval [*a,b*] is divided into subintervals.[3]

$$I(f) = \int_a^b f(x)dx \approx \frac{1}{2}\sum_{i=1}^{N}\left[\frac{f(x_i) + f(x_{i+1})}{2}\right](x_{i+1} - x_i) \qquad\qquad \text{5-2}$$

When the entire interval is divided into subintervals of the same width, that is, if

$$(x_2 - x_1) = (x_3 - x_2) = \cdots = (x_{i+1} - x_i) = \cdots = (x_N - x_{N-1}) = h \qquad\qquad \text{5-3}$$

then Equation 5-2 can be simplified to Equation 5-4.

$$I(f) \approx \frac{h}{2}\sum_{i=1}^{N}[f(x_i) + f(x_{i+1})] \qquad\qquad \text{5-4}$$

Because the size of each subinterval is small compared to the size of the entire interval, the approximation of the integrand with a linear function is valid, which means that the introduced error is negligible. Dividing the above integral with the number of subintervals results in the average production rate. Once the rate of production for every edge species is known, the edge species with the largest production rate is selected. This species and the reactions forming this species are not added to the partial network yet. A termination criterion based on the characteristic rate R_{char} is used to determine if this species is significant enough to be added.

5.2.2. Characteristic rate

The characteristic rate depends on the system and can be calculated in several ways. Some of the suggested definitions for R_{char} found in literature are described in this section. Susnow proposed a first definition for the characteristic reaction rate. In the original rate-based algorithm, the characteristic rate is estimated as the L1-norm of rates of production of the species in the partial network (Equation 5-5).

$$R_{char} = \sum_i |r_i(t)| \qquad\qquad 5\text{-}5$$

Where $r_i(t) = \frac{dC_i}{dt}(t)$ is the rate of change of the concentration of the ith reacted species and $0 < t < \tau$. The total reaction time, τ, is also a user specified parameter that indicates the end time of the system.

Song proposed a slightly different definition and estimates the characteristic rate as the L2-norm of rates of production of the species in the partial network.[4,5]

$$R_{char} = \sqrt{\sum_i r_i^2(t)} \qquad\qquad 5\text{-}6$$

Where $r_i(t) = V\frac{dC_i}{dt}(t)$, which is the molar rate of change of the ith species in the partial network.

Another definition proposed by Susnow is valid when the focus lies on the uniform conversion of a single primary reactant of the system.[6] In this case Equation 5-7 is valid.

$$R_{char} = \frac{n_{A,conv}}{t_{A,conv}} \qquad\qquad 5\text{-}7$$

Where $n_{A,conv}$ is the amount of the reactant A that is converted and $t_{A,conv}$ is the time it takes for this conversion. To use this equation, the user must specify a time scale of interest. When the partial network is solved, the time is incremented in regular intervals until a reactant conversion X_A is reached at time τ. The user must also specify a desired conversion of the initial reactant, $X_{A,max}$. At the beginning of the network generation process, this conversion is set to a smaller value and gradually incremented to $X_{A,max}$ while the network increases in size. This is done because smaller networks are expected to be only accurate for smaller conversions. Equation 5-7 can then be rewritten as the following,

$$R_{char} = \frac{[C_{A0} - C_A(\tau)]}{\tau} = X_A C_{A0}/\tau \qquad\qquad 5\text{-}8$$

where C_{A0} is the initial concentration of the reactant, C_A is the concentration of the same reactant at conversion X_A and τ is the time required to reach conversion X_A. Susnow also remarks that there are other, sometimes more appropriate definitions for R_{char} when the kinetics of the system change with time

or when there is more than one initial reactant. Worth mentioning is that for each edge species i, the maximum rate of production $R_{i,max}$ for the interval $0 < t < \tau$ is determined. This is used rather than a time-averaged value for the rate of production to properly account for the kinetic significance of species that may only be important for a short period. Finally, the rate-based algorithm with the above characteristics was tested for the pyrolysis of ethane. The number of generated reactions, edge species and selected species for this model as function of the defined precision level are depicted in Figure 2-9 on a logarithmic scale. The precision level was raised from 0.5 to 0.0001 and the complexity of the kinetic model increased rapidly. To determine which precision level is required to obtain a model of reasonable size including all important reactions and species, the concentration profiles for the major species at various precision levels are compared. For the pyrolysis of ethane, a conversion of 70 % and a precision level of 0.1 resulted in a relative complete kinetic model.

De Witt et al. improved the rate-based algorithm of Susnow and used it to construct a model for the pyrolysis of tetradecane.[2] A slight modified approach to determine the important species in the model was proposed. The characteristic rate is equal to the largest absolute value of r_i. This way, R_{char} is based on the most dynamic species in the partial network instead of on the kinetics of the initial reactant.

$$R_{char} = \max(|r_i|) \qquad \text{5-9}$$

Again, r_i is the net rate of production of the species in the partial network over the interval of integration. Finally, De Witt assessed the influence of the controlling parameter, ε, on the generation of a model for tetradecane pyrolysis. The results of the study were already briefly discussed in the introduction of this chapter. The models were deemed complete when the 2-propenyl radical was included in the model. Only when the precision level was equal to 7.0E-6, this radical was present in the kinetic model. Eventually, a total of 19 052 species and 479 206 reactions were generated to build a model that included 289 species and 102 257 reactions. It should be noted that this number of reactions is much higher than the number of unique reactions. In this case, the multiplicity of reactions and both the reverse and forward reactions are not yet excluded from the model. Even though this model is adequately complete, it took 209 hours for the kinetic model builder to complete the generation process.[2]

In this master thesis, the characteristic rate will be based on Equation 5-6. This definition is favored above Equation 5-5 because the conversion stop criterion instead of the time stop criterion will be used. This choice will be explained in section 5.4. Because the calculated rates are typically smaller than 1, the L2-norm is preferred above the L1-norm.

5.2.3. Precision level

Once the characteristic rate is calculated, the minimum rate R_{min} is obtained via Equation 5-10 :

$$R_{min} = \varepsilon * R_{char}$$

5-10

The precision level ε has a value between 0 and 1 and is specified by the user before the network generation process starts. The structure of the final kinetic model is influenced by the value of ε. If the users requires a detailed model that includes also less significant species, the precision level is lowered. This results in a smaller minimum rate and thus more species that are added to the network. As mentioned before, the time required to generate this model will increase as well. Vice versa, a smaller model is obtained when the value of ε is increased. The question remains which precision level the user should choose to obtain an optimal tradeoff between the accuracy of the model and the computational time. For the combustion of ethane, Susnow found that for different models generated with a precision level lower than 0.1 the major species concentrations were almost identical. In general, values between 0.5 and 0.0001 are typical for rate-based kinetic model builders.[7]

5.2.4. Conversion stop

Besides the first termination criterion that is introduced in the new rate-based algorithm, a second criterion has to be fulfilled during the network generation. Namely, the conversion of the initial reactants should not exceed a user specified value. This value should be comparable to the conversion that is achieved during experiments at similar operating conditions. On a molar base, the fractional conversion of a component k is defined as

$$X_k = \frac{F_k^0 - F_k}{F_k^0} \qquad\qquad 5\text{-}11$$

Where F_k^0 and F_k are the initial flow rate and flow rate of component k. If the reaction volume is constant, Equation 5-12 can be used.

$$X_k = \frac{C_k^0 - C_k}{C_k^0} \qquad\qquad 5\text{-}12$$

Because a kinetic model can be generated for processes with multiple initial reactants, an average conversion is introduced. This average conversion is the sum of the molar conversion of every initial reactant divided by the number of initial reactants, N. This equation is written as:

$$X_{av} = \frac{\sum_{i=0}^{N} X_i}{N} \qquad\qquad 5\text{-}13$$

During the generation process, the molar conversion is obtained from the CHEMKIN results after the partial network is solved. This output variable is derived by the post-processor of CHEMKIN and is defined as the percentage conversion of gas-phase species based on either the initial and final mole fraction in case of a close reactor or the inlet and outlet mass flux in case of an open reactor.[8]

The reaction time can also be used as a controlling variable in the rate-based algorithm. De Witte choses the reaction time while Susnow favors the conversion for the termination criterion. In RMG the user has to option to choose either one of them.[7] In Genesys, it was chosen to use the conversion as controlling variable of the algorithm. For radical chemistry there are many different time scales important. If the user would have to set a final reaction time, it might be possible that there are some significant species not added to the network yet. Therefore setting the conversion is a preferred option.

5.3. Selection of the edge species

When Genesys was developed, it was opted to construct the kinetic model considering all elementary steps as reversible.[9] The calculations of the reverse rate coefficient is based on the rate coefficient of the forward reaction and the equilibrium coefficient, via the thermodynamic consistency:

$$\frac{k_+(T)}{k_-(T)} = K_{eq}(T) = \exp\left(\frac{-\Delta_r G^\circ}{RT}\right)\left(\frac{RT}{p}\right)^{-\Delta n} \qquad \text{5-14}$$

When a new reaction is generated via a reaction family of which the kinetics are defined as "REVERSE", the reaction is not yet added to the partial network. Not until the corresponding forward reaction, of which the kinetics are calculated via Arrhenius parameters or group additivity, is added to the network will the reverse reaction be added. This should be taken into account in the rate-based algorithm. For example, during the pyrolysis of ethane the carbon-carbon scission reaction forms the radicals CH_3, H and C_2H_5. The kinetics of this type of reaction family is defined as "REVERSE" and hence, these reactions are not added to the partial network as long as the forward reaction, the radical recombination reaction, is not present in the network. During the network generation in Genesys the unreacted species, in this case ethane, are iteratively reacted against all the reaction families. This results in a set of new reactions and species. Before solving the partial network with these new reactions and species, it is checked if these species are truly products of the new reactions. If not, this means that the species are the product of a reverse reaction. This is the case for the radicals CH_3, H and C_2H_5 and hence they are added to the unreacted species list instead of the edge species list. One could call this "seeding" of the mechanism or adding important species to the network at the initial stage of the generation process. Next, it is checked if the newly generated species appear in the unreacted species list or in the partial network itself, to prevent that any species is listed as edge species by mistake.

5.4. Influence of parameters

The structure of the generated kinetic model is influenced by the user-defined parameters of the rate-based algorithm. The influence of the precision level on the number of edge species has already been discussed briefly. This is the one of the major concerns of the algorithm. If the precision level is decreased, the amount of edge species that is generated increases exponentially, as well as the computational time. Too much effort goes into estimating kinetic parameters of reactions with edge species as products, while eventually only a small fraction of these species and reactions will be included in the final model. By combining this with the rule-based method, the number of generated edge species is constrained and the computational time is reduced while significant species are still added to the network.

As one can suspect, the second termination criterion that utilizes the conversion of the initial reactant or the reaction time will have an influence as well. If the conversion is increased, it will take longer to fulfill the criterion during the network generation process and the size of the model increases. In this section, a simple model is generated with Genesys for the steam cracking of ethane. The reaction families and their constraints are listed in Appendix B. These constraints are rather loosely defined which results in the generation of insignificant edge species. The goal is to exclude these species from the partial network based on their rate of production. To solve the partial network, the following conditions for the isothermal batch reactor are chosen:

Table 5-1: Operating conditions for a homogeneous and isothermal batch reactor.

Temperature [K]	1100
Initial pressure [atm]	1
Volume [cm³]	90
Molar fraction C_2H_6 [-]	1.0

5.4.1. Precision level

First, the precision level is decreased from 1 to 1E-06. In Figure 5-2 the number of edge species and reactions towards the edge species that remain after the model is complete are plotted against the precision level on a logarithmic scale. The time required to generate the kinetic model and the calculated conversion of ethane are listed in Table 5-2. These values increase slowly and the model is generated within a reasonable time while the precision level is increased. However, once the precision level is smaller than 1E-06, too many edge species are added to the network, which means even more edge species can be created. This leads to an exponential increase in the amount of species and the computational time.
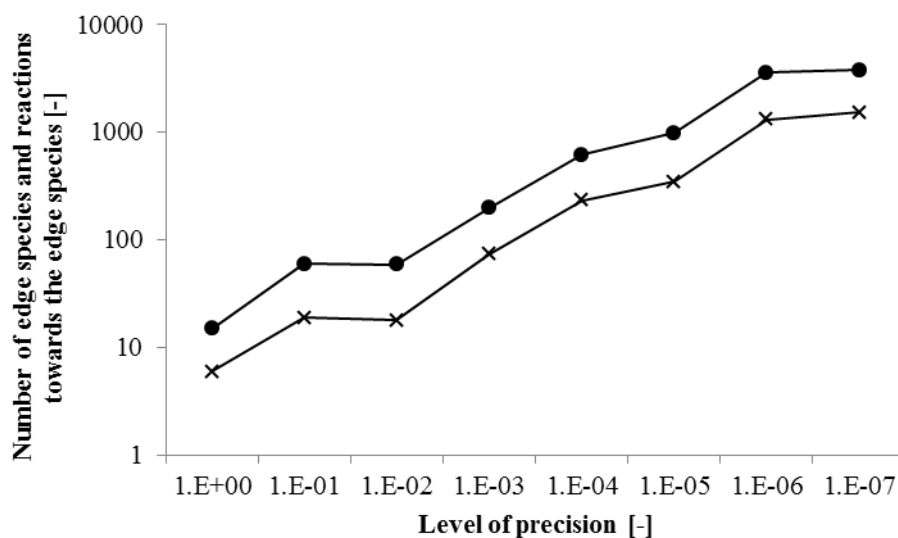
Figure 5-2: Number of edge species (crosses) and number of reactions towards edge species (dots) as a function of the precision level ε.

Estimating the thermochemical properties of each of these species and the kinetic parameters of every reaction towards the edge species becomes the most time consuming part of the entire network generation process. For a precision level of 1E-07 the increase is smaller. When looking at the network generation process with this precision level, it appears that the last edge species that was added to the network did not generate any new reactions, perhaps due to the constraints of the reactions families. The process was not terminated because the rate of the edge species was to small but because the user-defined rules were not satisfied. The calculated conversion of the initial reactant remains approximately 76 % once the precision level is lower than 1E-03. This indicates that there are no more reactions added to the network that consume ethane. This does not mean that the kinetic model is adequate. To verify the adequacy, one should also investigate the yields of the major products.

Table 5-2: Time required to generate the kinetic model and the final calculated conversion of ethane.

Precision level ε [-]	Computational time [s]	Calculated conversion [%]
1E00	15	48
1E-01	27	56.6
1E-02	32	57.1
1E-03	73	76.2
1E-04	201	76.2
1E-05	387	76.1
1E-06	2218	76.2
1E-07	2287	75.3

Also the size of the final kinetic model is influenced by the precision level. This can be seen in Figure 5-3, where the number of reactions and species of the models are illustrated.
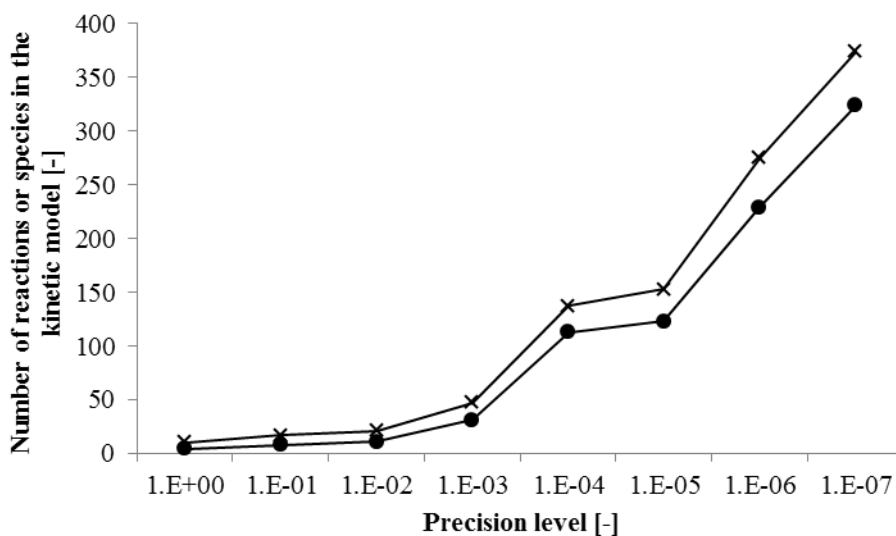
Figure 5-3: Number of species (crosses) and number of reactions (dots) in the final kinetic model as function of the precision level ε.

Overall, the increase in computational time and edge species remains reasonable compared to the rate-based algorithm without any constraints for the reaction families.

5.4.2. Conversion stop

Figure 5-4 shows the number of edge species and the computational time as function of the conversion. The influence of the conversion is less pronounced compared to the influence of the precision level. For a conversion higher than 80%, there is no influence at all on both variables. This is due to the fact that the precision level limits the formation of edge species instead of the user-defined conversion, since this high conversion is never reached. Since the precision level for every simulation is set to 0.001, the last three simulations result in the same kinetic model.
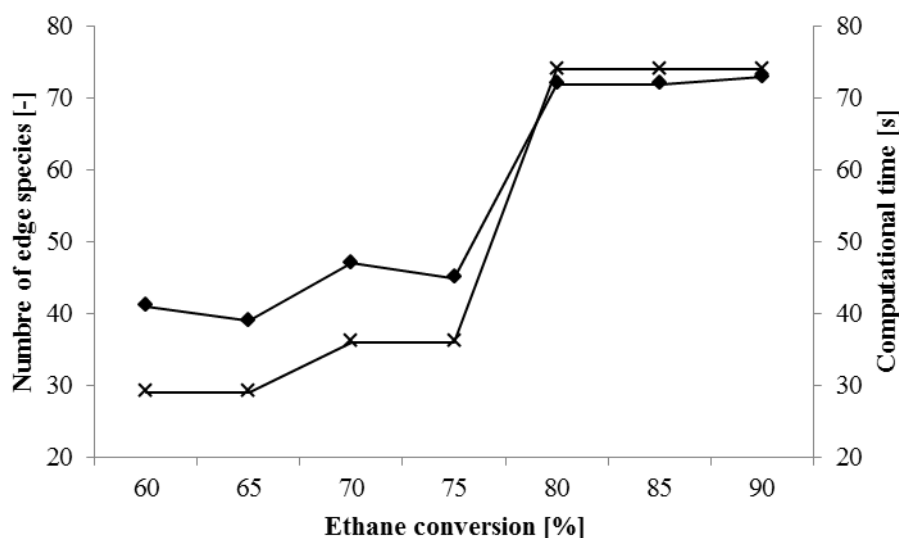
Figure 5-4: Number of edge species (dotted line) and time required to generate the kinetic model (full line), as a function of the conversion of ethane.

The size of the model, as can be seen in Table 5-3, is also less influenced by the conversion. When the conversion increases, it takes longer to fulfill the second termination criterion. Hence, more reactions are generated and added to the network. However, this increasing conversion has no influence on whether an edge species is selected and added to the network or not. Even if the conversion criterion is not met yet, if there are no more edge species with a rate of production that is high enough and the kinetic model will not be enlarged. For a conversion higher than 80 %, the minimum rate criterion is fulfilled instead of the conversion criterion. This also explains the stagnant trend for higher conversions in Figure 5-4.

Table 5-3: Size of the generated kinetic model, number of edge species and of the reactions towards the edge species for models generated with different precision levels.

Defined Conversion	# species	# reactions	# edge species	# reactions towards edge species	Calculated conversion
60	13	23	30	57	68.3
65	13	23	30	57	68.3
70	14	29	37	70	75.8
75	14	29	37	70	75.8
80	16	31	74	125	**76.2**
85	16	31	74	125	**76.2**
90	16	31	74	125	**76.2**

One could say that for the last three models the level of precision should be lowered in order to get a higher calculated conversion. But the results of the simulations with varying precision level (Table 5-2)

show that the calculated conversion remains stagnant around 76.2 %. This can be explained by looking at average conversions for ethane during experiments. For the same set of operating conditions, the conversion of ethane will be equal to this calculated conversion.

Finally, the characteristic rate R_{char} and the calculated conversion during each iteration in the network generation process are illustrated in Figure 5-5 and Figure 5-6. The sudden increase in the characteristic rate after the 20th iteration is caused by the fact that important reactions are now added to the kinetic model that consume ethane as a reactant and form important products. For example, the carbon-carbon scission reaction forming the radicals H, CH_3 and C_2H_5 were already generated but not added to the network in the beginning of the process. These reactions are considered as reverse reactions and only when the forward reactions are generated, they are added to the network.
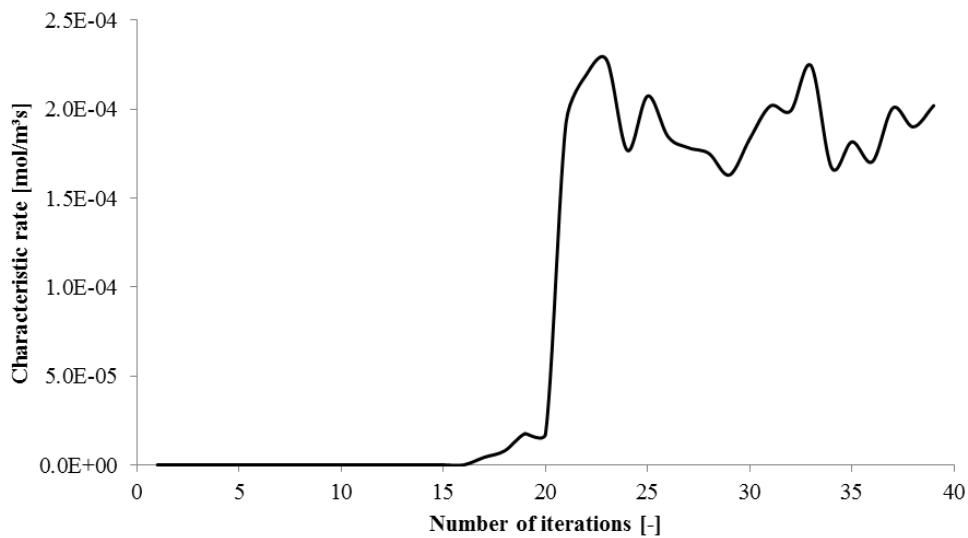
Figure 5-5: Characteristic rate per iteration.

This process also explains the sudden increase in the fractional conversion of ethane. After 20 iterations, all the important reactions that consume ethane are present in the network and the conversion remains stagnant.
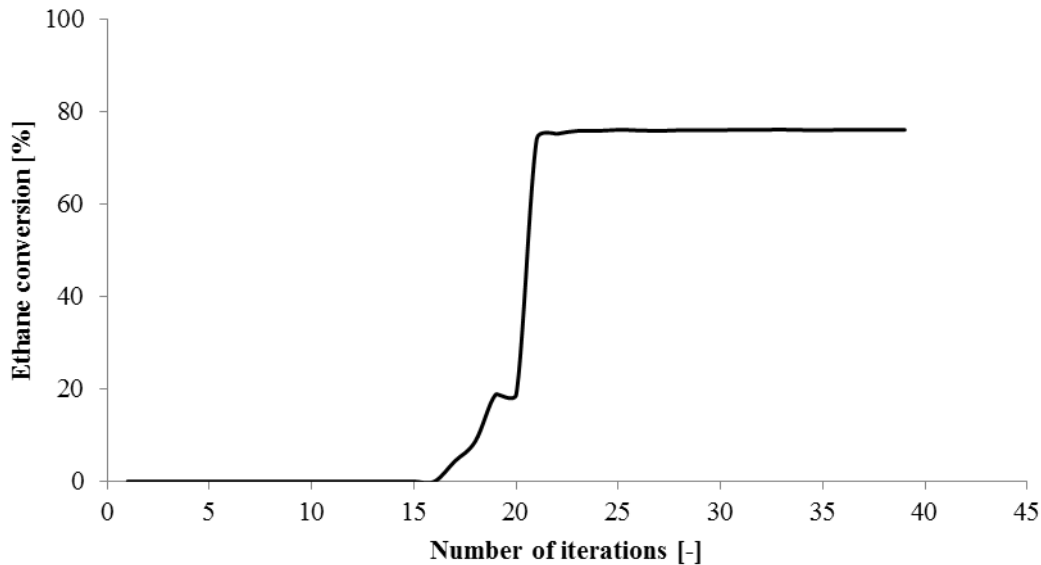
Figure 5-6: Calculated conversion of ethane per iteration.

Depending on the complexity of the process, the user of a rate-based kinetic model builder should put more effort into selecting an appropriate precision level instead of the conversion. Most of the time, he or she already has an indication of the conversion based on experimental results, while the precision level can range from 1 to 0.0001. To verify this assumption, this algorithm should be further studied for other, more complex processes.

5.4.3. Reactor type

By default the reactor type that is used to solve the partial network during each iteration is a perfectly stirred reactor or an ideal batch reactor. If the user is able to choose between other types of ideal reactors, it would become possible to tailor the generated kinetic model towards the process it will eventually be used for. As described in chapter 4, for a CSTR a steady state solver is used to obtain the dynamic properties of the process. This might influence the rates of the species and therefore also the selection of the edge species. To evaluate this assumption, the model for the steam cracking of ethane is used again. A conversion of 90 % and a precision level of 0.001 is chosen.

For the batch reactor with a volume of 90 cm³, the temperature is 1100 K and the initial pressure 1 atm. The molar fraction of the initial reactant ethane was set to 1.0. The total integration time is equal to 0.1 seconds. For the CSTR, the volume is 90 cm³ and the average residence time, τ, 0.1 seconds. The molar fraction, temperature and pressure were equal to the values for the batch reactor. Finally, the PFR has a length of 23.137 m and a diameter of 1 cm. The temperature and pressure profile were assumed constant

across the entire length and were set equal to 1100 K and 1 atm respectively. The mass flow rate is 6.04 g/s, which results in a residence time of 0.1 seconds.

Table 5-4 presents the differences for each kinetic model generated with a different reactor. The largest model was generated with the CSTR, which means that the concentrations and hence the rates of the edge species are larger when a steady state solver is used. For the same precision level and defined conversion, the model will be more accurate, that is, more species are deemed significant.

Table 5-4: Results for the network generation using different types of ideal reactors while solving the partial network.

Reactor type	# reactions	# species	# edge species	Computational time [s]	Conversion [%]
Batch	31	16	74	119	76.20
PFR	32	17	87	149	56.97
CSTR	51	22	131	372	46.02

5.5. Rate-based network generation without constraints

To emphasize the importance of the constraints of each reaction family, the new rate-based algorithm is tested without constraints. A model for steam cracking of ethane is generated with the same reaction families as described in Appendix B. The last column containing the constraints is ignored, except for the constraint on the number of unpaired electrons. The species can either have zero or one unpaired electron to prevent the formation of biradicals. These type of radicals are not important for the steam cracking of ethane and most of their thermochemical properties are not listed in the database. Therefore it is safe to prevent them from being generated during the process. If they would be generated, this would result in to species of which no accurate thermochemical information can be estimated. The second reason why these species are excluded lies within the code itself. Each reaction that is generated is checked on its uniqueness by comparing the transition state of the reaction itself. When Genesys encounters multiple reactions with the same products and reactants, the transition state graphs stored for every elementary reaction will be compared to each other. If these graphs differ, the reactions are considered to be different and are added to the reaction network. This results in a network containing duplicate elementary reactions. An example is given in Figure 5-7. The formation of C_4H_{10} and the biradical C_4H_8 formed from two radicals C_4H_9 can occur via two different transition states.
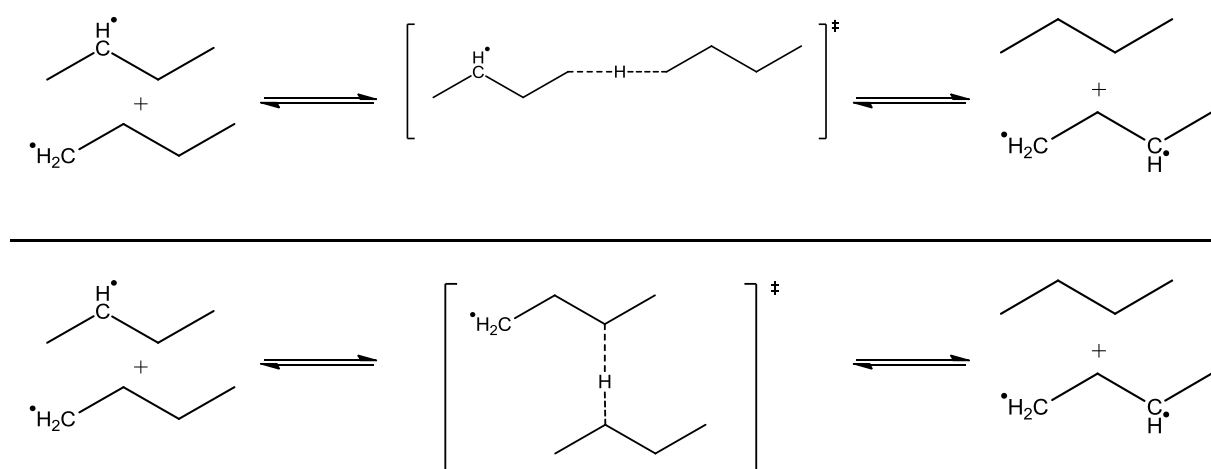
Figure 5-7: H abstraction by a C_4H_9 radical from a C_4H_9 radical forming C_4H_{10} and C_4H_8 via different transition states.

To investigate the results of excluding the constraints, the model is generated using an isothermal batch reactor with a volume of 90 cm³. The temperature is set to 1100 K and the initial pressure to 1 atm. The goal conversion of ethane is set to 80 %. The precision level is reduced for each simulation until memory limitations occur. When the precision level is equal or lower than 1E-06 Genesys fails to complete the network generation process. For the simulation performed with a precision level equal to 1E-06 it took a little over 12 hours before memory limitations occurred. If the precision level was further reduced to

1E-07 it took approximately 21 hours to exceed the memory capacity. Overall the precision level should be reduced a fair amount compared to the other pure rate-based kinetic model builders. Generating a kinetic model with such a small precision level will results eventually in a kinetic model that is too large and contains too many insignificant species. It is fair to say for this mechanism that one should not have to worry about memory limitations. However, there are only 11 reactions families defined to construct this model. In reality, more reactions families are necessary to obtain an accurate kinetic model for the steam-cracking of ethane. There are usually more reaction families necessary resulting in more possible reactions and new species that are generated. The list of edge species in these cases would be bigger and memory limitations would become more of an issue.

5.6. Conclusion

A new rate-based algorithm was implemented in Genesys. This algorithm is used next to the existing rule-based algorithm to include significant species during the network generation process and to terminate the process at an appropriate time. The rate-based method itself has some shortcomings that are described in literature. The method requires a large amount of computational effort when a highly accurate kinetic model is required. To many insignificant species have to be generated and their properties have to be estimated, which is a computationally demanding task. When rules or constraints are used to create these species, their total amount is significantly reduced. Another shortcoming of the rate-based method is the dependency of the model on the predefined operating conditions such as the temperature, pressure and reactor type. The range of operating conditions for which this model can be applied will be limited, but hopefully the model will predict better yields at conditions it was built for. In literature, a solution for this problem is to include a seed mechanism that contains species and reactions which must be included in the model, independently of the operating conditions.

The precision level, conversion and reactor type are parameters that the user must define before Genesys can generate a kinetic model. These parameters influence the structure of this model. When the precision level decreases, the number of reactions and species in the model increases and the computational time rises. However, this increase is limited because the amount of edge species that is generated is limited by the reaction families constraints. When the goal conversion increases, the model increases as well. This conversion has less influence on the structure of the kinetic model compared to the precision level.

5.7. References

1. Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combust. Flame.* 2011;158(1):16-41.

2. De Witt MJ, Dooling DJ, Broadbelt LJ. Computer generation of reaction mechanisms using quantitative rate information: Application to long-chain hydrocarbon pyrolysis. *Ind. Eng. Chem. Res.* Jul 2000;39(7):2228-2237.

3. Gilat A, Subramaniam V. *Numerical methods for engineers and scientists : an introduction with applications using MATLAB.* Hoboken, NJ: Wiley; 2011.

4. Song J. *Building robust chemical reaction mechanisms: Next Generation of Automatic model construction software*: Departmen of Chemical Engineering, Massachusetts Institute of Technology; 2004.

5. Song J, Stephanopoulos G, Green WH. Valid parameter range analyses for chemical reaction kinetic models. *Chem. Eng. Sci.* 2002;57(21):4475-4491.

6. Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. Rate-based construction of kinetic models for complex systems. *J. Phys. Chem. A.* May 1997;101(20):3731-3740.

7. Green PWH. RMG Documentation. *Release 4.0.1*2013. Accessed 6 may 2013.

8. CHEMKIN Visualization Manual: Reaction Design: San Diego; 2010.

9. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chem. Eng. J.* Oct 2012;207:526-538.

Chapter 6

Case study: Pentanol

pyrolysis

6.1. Introduction

The purpose of this chapter is to generated two kinetic models with Genesys. The first model is obtained via the rule-based method meaning that the species and reactions are included in the final reaction network based on user-defined constraints. For the generation of the second model, the flux of the generated species is used as well. The reaction family constraints for the latter model will be less tight in order to generate more species and reactions. This means that more insignificant species are generated as well. However, based on their flux, only the important species will be included in the final model. If the constraints of the rule-based method are ill-conditioned or if the precision level ε of the rate-based criteria is set too high, the consequence could be that some important species and reactions are missed in the final model. Eventually, the combination of the rate and rule-based method should result in a detailed kinetic model that can predict the major product yields accurately without computational issues such as memory limitations, which are common for the rate-based algorithm. A third model will be generated via the kinetic model builder RMG. This code is purely based on the rate-based criterion. Finally, the three models are compared to experimental results. For this purpose, the models will be constructed for the pyrolysis of n-pentanol. Studying the thermal decomposition pathways of higher molecular weight alcohols is currently an important field of study. The number of kinetic modeling studies on alternative fuels have been increasing and biofuels appear to be a promising alternative source of energy. They have a lower pollutant emission compared to fossil fuels and are produced from renewable sources.[1] The amount of biofuels in fuels used for transport in Europe has increased from 0.1 % in 1996 up to 5 % in 2012 (Figure 6-1).
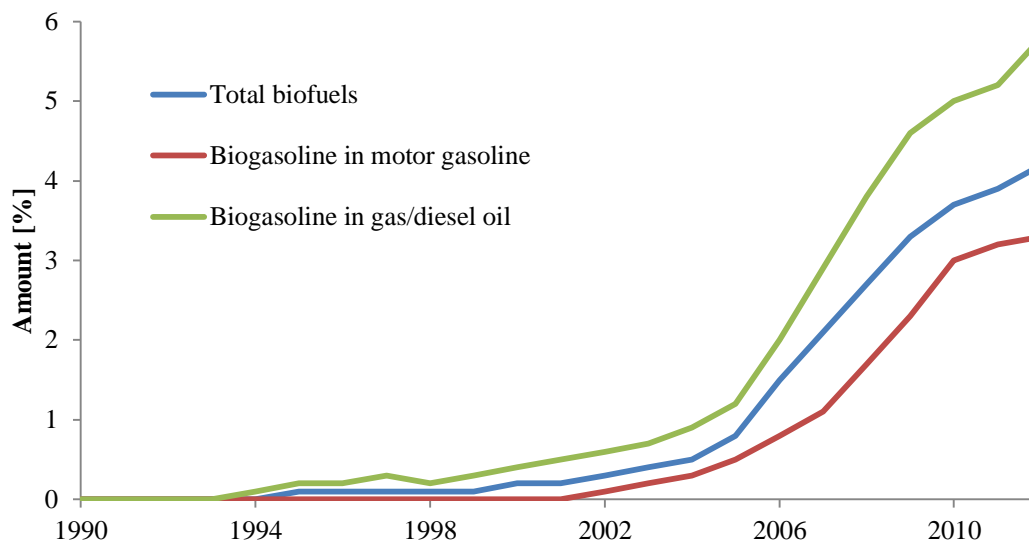
Figure 6-1: Biofuels share in transport fuels from 1990 to 2012.[2]

The advantage of n-pentanol is its higher energy density, lower vapor pressure and better fuel capability.[3] There are already detailed reaction mechanisms available based upon the chemistry of n-butanol. However, pentanol could be much better blended with fossil fuels.[4] Although it is not a good octane improver, it may be applicable in diesel engines with auto ignition.[5] Pyrolysis reactions are an essential part of any combustion mechanism. The goal of this chapter is to compare the performance of different generated models for the same chemical process. This case study is not intended to perform a detailed kinetic study.

6.2. Methodology

6.2.1. Experimental data

The experimental data for the pyrolysis of n-pentanol is obtained from a bench scale reactor. The feed of pure pentanol was diluted with nitrogen at a molar ratio of nitrogen to pentanol equal to four. During the experiments several conversions of n-pentanol are obtained at different temperature profiles along the tubular reactor. The set of conditions correspond to a range of n-pentanol conversion from 17 % to nearly complete conversion. The experimental data was measured in a tubular reactor of 147.5 cm long with an internal diameter of 6 mm. The experimental conditions are described in Table 6-1.

Table 6-1: Experimental conditions for the pyrolysis for n-pentanol on a bench scale reactor.

Temperature range [K]	913-1013
Pressure [bar]	1.7
Feed flow rate [g/h]	48
Nitrogen flow rate [g/h]	61

6.2.2. Model generation with Genesys

Models for pyrolysis are less complex than oxidation models but they have a higher number of species and reactions.[6] The three important reaction families for a free radical mechanism are 1) the carbon-carbon and carbon-hydrogen bond scissions and the reverse radical-radical recombination, 2) the intra- and intermolecular hydrogen abstraction reactions and 3) the radical addition to olefins and the reverse β-scission of radicals. These common reaction families, which were also used in chapter 5 for the pyrolysis of alkanes, are expanded with extra reaction families. In Appendix C an overview of the reaction families can be found. The reactions families concern also the scission, abstraction and addition reactions as described above, except that the element oxygen is allowed to appear in the reacting species. The reaction family templates for the hydrogen abstraction of aldehydes and the ketene formation are also listed in Appendix C but no reactions of this type will be generated because the reactions families forming a carbon-oxygen double bond are not defined. For the estimation of kinetic parameters and thermodynamic properties, the group additive values of Sabbe[7,8] and Paraskevas[9-11] incorporated in the Genesys libraries were used.

As mentioned in chapter 1, the application of reaction families to generate the important reactions is not valid for reactions with small molecules, since the entire molecule instead of a reactive moiety influences the reactivity. Therefore, a detailed kinetic model is added to include the pyrolysis of C_1-C_4 based hydrocarbons and oxygenated hydrocarbon species.[12] This model is called the AramcoMech 1.3 and is developed by the Combustion Chemistry Centre in NUI Galway.[13] It is valid for a wide range of conditions and has been developed from "the bottom up". First, a C_1 sub-mechanism was used and

expended with larger carbon species such as ethane, ethylene, acetylene, etc. This model has been validated against a large number of experimental results.

Secondly, for carbon-carbon scission and carbon-hydrogen scission reactions it is difficult to estimate accurate kinetic parameters based on group additive values. Because these reactions are important initiation reactions for the decomposition of n-pentanol, they are added manually after the network generation process. The included reactions are based on the work of Heufer.[5] Heufer et al. investigated the oxidation of n-pentanol based on modeling rules used for n-butanol. The obtained model was validated against ignition delay time data, speciation data from a jet-stirred reactor and laminar flame velocity measurements.[5] Table 6-2 lists the reactions which were selected from this work and added to the three models described in this chapter.

Table 6-2: Carbon-carbon scission reactions and carbon-hydrogen reactions that were manually added to the model. [5]

Elementary Reaction	A [cm³ mol^{-1} s^{-1}]	n	Ea [cal mol^{-1}]
C5H12O⇋CH3+C4H9O	3.790E24	-2.230	8.807E04
LOW	1.7820E60	-1.2280E01	8.398E04
	a	T$_3$	T$_2$	T$_1$
TROE	2.3520E-01	7.2400E+02	5.0000E+09	5.0000E+09
C5H12O⇋C2H5+C3H7O	3.080E24	-2.269	8.844E04
LOW	8.0250E113	-27.650	9.645E04
	a	T$_3$	T$_2$	T$_1$
TROE	9.7028E05	2.6840E02	2.6840E02	4.7401E03
C5H12O⇋nC3H7+C2H4OH	5.530E24	-2.230	8.901E04
LOW	6.632E59	-1.2130E01	8.472E04
	a	T$_3$	T$_2$	T$_1$
TROE	2.4380E-01	7.4406E02	5.0000E09	5.0000E09
C5H12O⇋C4H9+CH2OH	3.020E23	-1.880	8.571E04
LOW	1.416E59	-1.1930E01	8.398E04
	a	T$_3$	T$_2$	T$_1$
TROE	7.6460E-01	8.3440E09	7.2480E02	8.2140E09
C5H12O⇋OH+C5H11	6.330E20	-1.370	9.493E04
LOW	6.902E51	-1.0210E01	8.920E04
	a	T$_3$	T$_2$	T$_1$
TROE	7.0300E-01	9.8820E09	6.3468E02	1.7860E09
C5H12O⇋C5H10+H2O	3.520E13	0.00	6.723E04
LOW	1.690E75	-1.7040E01	6.475E04
	a	T$_3$	T$_2$	T$_1$
TROE	8.0000E-02	1.0000E00	9.9240E09	9.9240E09
C5H12O⇋H+C5H11O*	4.00E13	0.0	0.0
C5H12O⇋H+C5H11O_1**	4.00E13	0.0	0.0
C5H12O⇋H+C5H11O_2***	4.00E13	0.0	0.0
C5H12O⇋H+C5H11O_3****	4.00E13	0.0	0.0
C5H12O⇋H+C5H11O_4*****	4.00E13	0.0	0.0

*C5H11O corresponds to [CH2]CCCCO1 (SMILES notation)
**C5H11O_1 corresponds to C[CH]CCCO (SMILES notation)
*** C5H11O_2 corresponds to CC[CH]CCO (SMILES notation)
**** C5H11O_3 corresponds to CCC[CH]CO (SMILES notation)
***** C5H11O_4 corresponds to CCCC[CH]O (SMILES notation)

The five carbon-hydrogen bond scission reactions listed in the lower part of Table 6-2 were generated by Genesys but the kinetic parameters did not correspond to the ones from Heufer. Therefore the estimated kinetic parameters were replaced with the kinetic parameters listed in Table 6-2. The carbon-carbon bond scission reactions in Table 6-2 are pressure-dependent reactions. To describe the pressure dependence the F-center formalism of Troe[14,15] is used. Hereby, ten parameters are used to determine the rate coefficients of a reaction at a specified temperature and pressure. The normal Arrhenius parameters are the high-pressure modified Arrhenius parameters, while the second line contains the low-pressure modified Arrhenius parameters, and the third line contains four parameters a, T_3, T_2 and T_1 which are used to determine the F-center value which describes the center of the fall-off range.[16] Using these parameters in the following equations the falloff function $F(T, P_r)$ is obtained.

$$\log_{10} F(T, P_r) = \frac{\log_{10} F_{cent}(T)}{1 + f_1^2} \qquad \text{6-1}$$

$$F_{cent}(T) = (1 - a) \exp\left(-\frac{T}{T_3}\right) + a\, exp\left(-\frac{T}{T_1}\right) + \exp\left(-\frac{T_2}{T}\right) \qquad \text{6-2}$$

$$f_1 = (\log_{10} P_r + C)/(N - 0.14(\log_{10} P_r + C)) \qquad \text{6-3}$$

$$C = -0.4 - 0.67 \log_{10} F_{cent} \qquad \text{6-4}$$

$$N = 0.75 - 1.27 \log_{10} F_{cent} \qquad \text{6-5}$$

It took approximately 728 seconds to generate the rule-based model with Genesys. For the network generation via the combined rate and rule-based method in Genesy, a model was obtained after 4523 second (or 75 minutes). In Table 6-3 an overview is given of the number of species and the reactions consisting in the three models. From now on, the model generated with the rule-based method will be referred to as model I , the model generated with a combination of both rate and rule-based methods as model II and the model generated with RMG as model III. The rows with the extra number of species added and extra number of reactions added in Table 6-3 consist of reactions and species from the AramcoMech model and the model from Heufer that were added manually afterwards or, in the case of RMG, that were used as seed mechanisms. This also explains why there are less species added to the model of Genesys compared to the model of RMG. Species that are already present in the model Genesys which are also present in the reactions of AramcoMech or Heufer are not added again. In RMG al 126 species from AramcoMech and Heufer are used since the network generation process starts with these species as initial reactants.

Table 6-3: Number of species and reactions for both models obtained by Genesys and the model obtained by RMG.

	Genesys rule-based	Genesys rule and rate-based	RMG
# of species	320	292	151
# of reactions	2174	2258	4675
# of hydrogen abstraction reactions	1425	1756	1949
# of radical additions	385	285	55
# of radical recombinations	364	217	80
# of extra species added	100	96	129
# of extra reactions added	731	731	731
Total # of reactions	2905	2956	5406

Notice that the total number of reactions for model II is not equal to the sum of the number of generated reactions and the number of extra reactions that are added. This is because some of the reactions from the AramcoMech or from Heufer were already present in model II. These duplicate reactions and their rate coefficients were removed from the model of Genesys and replaced.

Comparing model I to model II shows some interesting results. There are more species but fewer reactions present in model II. Looking at the number of radical addition and radical recombination reactions in model I and model II, there appear to be less of these reactions present in the second model. This could be explained via of the rate-based criterion. If this type of reactions forms edge species with a low rate of formation, these edge species are not added to the model and hence these reactions are also not added. To generate model II the rate-based characteristics described in Table 6-4 were chosen.

Table 6-4: Rate-based criterion characteristics.

Precision level ε [-]	0.0001
Conversion $C_5H_{12}O$ [%]	99.99
Reactor type	PFR
Mole fraction $C_5H_{12}O$ [-]	1.0
Pressure [bar]	1.7

The goal conversion of pentanol was set to a high value because experiments show that at high temperatures almost all of the pentanol is consumed. To solve the partial network, an isobaric plug flow reactor was chosen with a temperature profile in axial direction as can be seen in Figure 6-2. This temperature profile is obtained from the experimental results where 48 g/h pentanol and 61 g/h N_2 were used as feed.
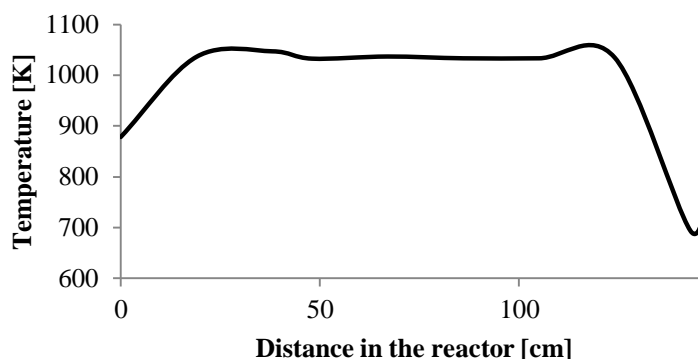
Figure 6-2: Temperature profile along the bench scale reactor for the pyrolysis of n-pentanol with as feed 48 g/h pentanol and 61 g/h N_2.

6.2.3. Model generation with RMG

To generate the model with RMG the following characteristics were used:

Table 6-5: Characteristics of the rate-based model generation for the n-pentanol pyrolysis with RMG.

Temperature [K]	1053.15
Pressure [bar]	1.7
Initial concentration $nC_5H_{12}O$ [mol cm^{-3}]	1.0
Initial concentration N_2 [mol cm^{-3}]	1.0
Precision level ε [-]	0.1
Goal conversion $nC_5H_{12}O$	0.6

A model consisting of 5406 reactions and 151 species was generated in a little over 34 minutes. Two seed mechanisms were used, the model from AramcoMech and a model with the scission reactions from Heufer. The tolerance and conditions were tuned by trial and error to obtain a kinetic model without facing memory limitations. I.e., by adjusting the precision level and the goal conversion the edge remained small enough to allow conversion of the model generation. The model has an edge with 110 863 reactions and 21 039 species. However, due to the low number of species in the model, there is reason to believe that this model will not be adequate to predict the product yields.

Once the three models were generated, the CHEMKIN software was used predict the product composition. For the reactor simulations a plug flow reactor was assumed. Radial temperature gradients were neglected, as was also assumed by Harper[17] and Froment[18,19]. The conditions and reactions configuration as described section 6.2.1 and in Table 6-1 were used for the reactor simulations.

6.3. Comparison with experimental data
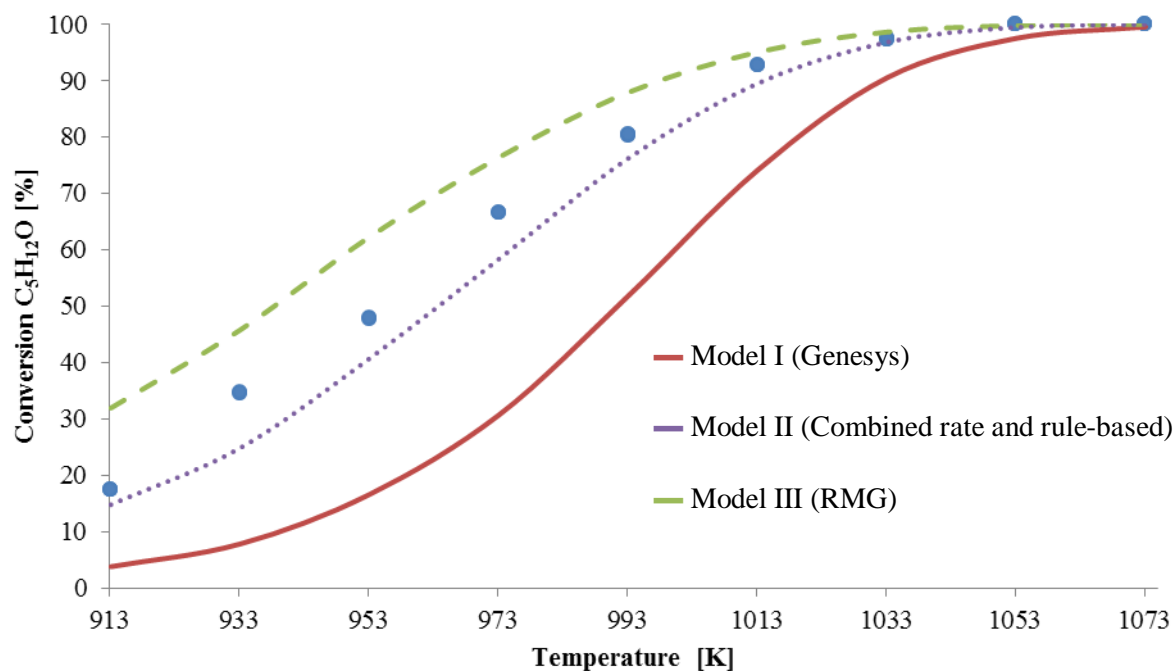
Figure 6-3: Conversion of pentanol as function of the operating temperature. The green dashed line corresponds to the RMG model, the red full line with the Genesys model and the purple dotted line with the combined rate and rule-based model. The dots correspond to experimental data.

Figure 6-3 shows the conversion of n-pentanol as function of the temperature in the reactor. At higher temperatures, the three models all predict the conversion correctly, specifically, when almost the entire amount of pentanol is decomposed. If the temperature decreases, the error in the estimation increases for every model. Only model II, the model generated with the combined rate and rule-based method, predicts the conversion relatively well. RMG overestimates the conversion while Genesys underestimates the conversion. There are two possible explanations for these results. First, it is possible RMG deemed more reactions that consume pentanol as significant reactions, this could explain the higher amount of pentanol that is converted. On the other hand, Genesys could have missed some reactions that are essential to the decomposition pathways of pentanol. A second possibility is that the rates of the reactions in both models are either too slow or too fast, due to wrong group additive values for thermodynamic properties or rate coefficients. In this case, the rule-based network generation process is not faulty but the available information inside the databases is. For the rate-based process, incorrect information in the database influences the selection of species and reactions. Looking at the prediction of the conversion obtained with model II, it could be that the faulty prediction of model I is caused by missing some important reactions. To support this assumption, a rate of production analysis for the most important species is required. If inaccurate rate coefficients were at the base of the errors, model II would give a worse estimation as well. In Figure 6-4 the predictions of the major product yields

are compared for the three models against the experimental results. Model I predicts the yields of C_2H_4, H_2 and H_2O quite well but underestimates the yields of CO, CH_4 and acetaldehyde. Especially for acetaldehyde there is a large difference visible. However, the trend is the same: there is a maximum in the acetaldehyde fraction as a function of the conversion. One of the reasons why the acetaldehyde yield is predicted the worse can be found in the reaction families template. As mentioned before, the reaction families that form carbon oxygen double bonds were not added to the model. The yield that is seen now is caused by reactions from the AramcoMech involving acetaldehyde.
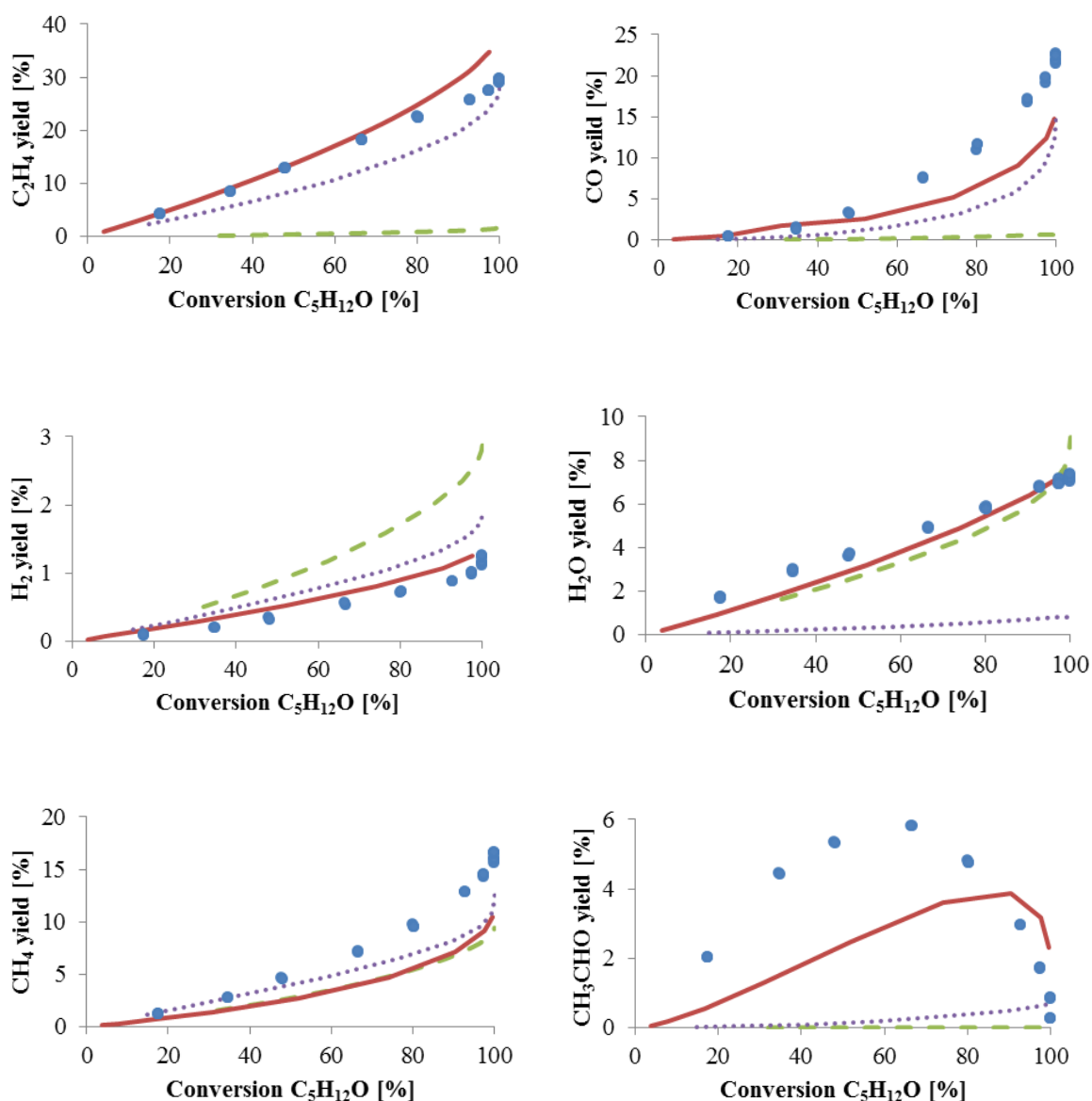
Figure 6-4: Yield predictions for C_2H_4, CO, H_2, H_2O, CH_4 and acetaldehyde as function of the conversion of pentanol. The green dashed line corresponds to the RMG model, the red full line with the Genesys model and the purple dotted line with the combined rate and rule-based model. The dots are experimental values.

As can be seen for model III, which was generated with RMG, the predictions are worse for most of the products. The predictions for CH_4 and H_2O are compared to the predictions of Genesys almost equal to each other. For the prediction of H_2 there is an overestimation. Finally the yields of acetaldehyde, C_2H_4 and CO are underpredicted. This could be caused by important formation reactions of these products that are not added to the model. It can already be concluded that the generation of the model via RMG was done incorrectly. A possible way to solve the memory limitations with RMG is by setting certain constraints on the generated species. RMG offers the possibility to limit the maximum amount of carbon atoms, oxygen atoms, radical sites and ring structures present in species. By default these values are set to 25 but the user can choose the lower the values and hence limit the size of the generated edge species. To obtain model III these values were already lowered, however, there is still margin to reduce them further. It should be investigated if a better model would be obtained even if this means that the concept of "rate-based" is undermined by setting these constraints. Looking at some kinetic parameters for reactions which occur in every model, there is a large difference noticeable, especially for the activation energy (Table 6-6). In Figure 6-5 the corresponding rate coefficients for every reaction is plotted as function of the reciprocal temperature.

Table 6-6: Elementary reactions with the estimated kinetic parameters. The kinetic parameters of model I and model II were equal, since they were based on group additivity values of the same database.

	Model I & Model II			Model III		
	A [m³ mol⁻¹s⁻¹]	n	Ea [kJ/mol]	A [cm³ mol⁻¹s⁻¹]	n	Ea [kJ/mol]
C3H5+C5H12O⇋C3H6+C5H11O	2.680E07	0.0	113.9	2.806E09	-0.062	95.54
C4H8O+H⇋C4H9O	6.460E07	0.0	22.52	1.180E07	0.0	159.1
C2H5O+C5H12O⇋C2H6O+C5H11O	1.750E07	0.0	77.25	3.716E-14	5.994	48.23
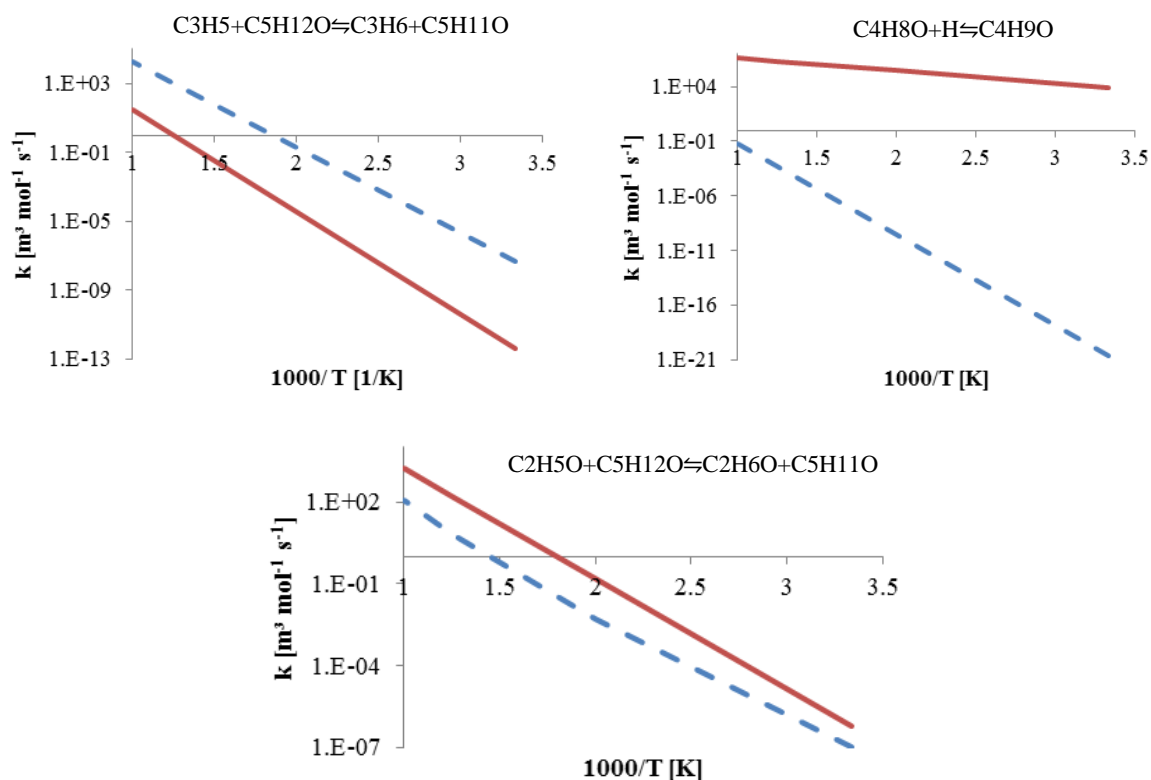
Figure 6-5: Rate coefficients as function of the reciprocal temperature. The red full line correspond to the model obtained with Genesys(model I and model II), the blue dashed lines correspond to the model obtained with RMG (model III).

The predictions of model II lie close to the predictions of model I, except for the yield of H_2O and acetaldehyde. These yields are significantly underpredicted compared to the experimental values. This could be the result of the lower number radical addition reactions and their reverse reactions that were included in the model. Perhaps the rate-based algorithm has mistakenly excluded them from the model based on incorrect kinetic parameters. However, this prediction of the yield is in contrast with the better prediction of the conversion of pentanol. In this case, model II predicted the conversion better. Apparently model II has included most of the important decomposition reactions of pentanol but has failed to include the important formation reactions of the major products. Again, to confirm this, a reaction path analysis and a study of the rates of production are required. One could say the worse estimation is due to the missing reaction families in the reaction families template of Genesys. However, if this was the case, model I has to predict the yield of H_2O faulty as well.

6.4. Conclusion

Three kinetic models were generated for the pyrolysis of n-pentanol. The first two models were generated with the kinetic model builder Genesys. Model I was generated with the rule-based method while model II was generated with a combined version of the rate and rule-based method. The last model, model III, was generated with the kinetic model builder RMG, a code which is purely based on the rate-based method. For every model, reactions concerning small hydrocarbons and oxygenates were added from the AramcoMech model and reactions concerning carbon-hydrogen and carbon-carbon scission reactions were taken from the work of Heufer et al. These reactions were added manually after the network generation with Genesys, while for the network generation with RMG these reactions were used as seed mechanisms.

After the network generation was complete, the models were compared to each other and to experimental results obtained from a bench scale reactor. The size of model I and II were almost equal to each other. However, model III had a very low number of species compared to the number of reactions. This was an indication that the predictions of the final model had to be looked at with great care.

To obtain the yield predictions of major products, the three models were used in reactor simulations with a plug flow reactor model. The best prediction of the conversion of n-pentanol as a function of temperature was obtained by model II, model I underestimated the conversion and model III overestimated the conversion. For the yields of products such as C_2H_4, CO, CH_4, H_2O, H_2 and acetaldehyde the results varied. All models predicted the yield of acetaldehyde poorly. This could result due to various reasons. Not only could this be the result of important reaction families that are missing reaction families template in Genesys, it could also be that case that the kinetic or thermodynamic parameters are poorly estimated due to wrong group additive values. Comparison of some kinetic parameters and rate coefficients for reactions that appear in every model show that there is a large difference between these values, especially for the activation energy.

Based solely on the conversion of pentanol it could be concluded that model II is the most adequate model to predict the pyrolysis of pentanol. However, the product yields dispute this conclusion. It can be concluded that rate of production analysis is necessary to see whether there are some important reactions missing in the models.

6.5. References

1. Battin-Leclerc F, Simmie JM, Blurock E. *Cleaner Combustion: Developing Detailed Chemical Kinetic Models*: Springer London; 2013.

2. European Union. European Commission. Directorate-Genaral for E. *EU Energy in Figures 2014*: Publications office of the European Union; 2014.

3. Qianqian L, Shamel SM, Green Jr WH, Huang J. A kinetic modeling study of high temperature chemistry of n-pentanol. In: Institute WSSotC, ed. *8th U. S. National Combustion Meeting*. Utah2013.

4. Sciubba E, Manfrida G, Desideri U. *ECOS 2012 The 25th International Conference on Efficiency, Cost, Optimization and Simulation of Energy Conversion Systems and Processes (Perugia, June 26th-June 29th, 2012)*: Firenze University Press; 2012.

5. Heufer KA, Sarathy SM, Curran HJ, Davis AC, Westbrook CK, Pitz WJ. Detailed Kinetic Modeling Study of n-Pentanol Oxidation. *Energy & Fuels.* 2012/11/15 2012;26(11):6678-6685.

6. Pierucci S, Ranzi E. A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Comput. Chem. Eng.* Apr-May 2008;32(4-5):805-826.

7. Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *J. Phys. Chem. A.* Nov 2008;112(47):12235-12251.

8. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J. Phys. Chem. A.* Aug 2005;109(33):7466-7480.

9. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates. *Chemistry – A European Journal.* 2013;19(48):16431-16452.

10. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Kinetic Modeling of α-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *ChemPhysChem.* 2014;15(9):1849-1866.

11. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos NG, Marin GB. Kinetic Modeling of α-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *The Journal of Physical Chemistry A.* 2014/10/09 2014;118(40):9296-9309.

12. Metcalfe WK, Burke SM, Ahmed SS, Curran HJ. A Hierarchical and Comparative Kinetic Modeling Study of C1 − C2 Hydrocarbon and Oxygenated Fuels. *International Journal of Chemical Kinetics.* 2013;45(10):638-675.

13. Curran HJ. Saudi Aramco Mechanism Release. 2015; http://www.nuigalway.ie/c3/Mechanism_release/frontmatter.html. Accessed 1st May 2015, 2015.

14. Troe J. Theory of Thermal Unimolecular Reactions in the Fall-off Range. I. Strong Collision Rate Constants. *Berichte der Bunsengesellschaft für physikalische Chemie.* 1983;87(2):161-169.

15. Gilbert RG, Luther K, Troe J. Theory of Thermal Unimolecular Reactions in the Fall-off Range. II. Weak Collision Rate Constants. *Berichte der Bunsengesellschaft für physikalische Chemie.* 1983;87(2):169-177.

16. Warnatz J, Maas U, Dibble RW. *Combustion: Physical and Chemical Fundamentals, Modeling and Simulation, Experiments, Pollutant Formation*: Springer; 2006.

17. Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combust. Flame.* 2011;158(1):16-41.

18. Chen Q, Froment GF. THERMAL-CRACKING OF SUBSTITUTED AROMATIC-HYDROCARBONS .1. KINETIC-STUDY OF THE THERMAL-CRACKING OF I-PROPYLBENZENE. *J. Anal. Appl. Pyrolysis.* Sep 1991;21(1-2):27-50.

19. Froment GF, Bischoff KB, De Wilde J. *Chemical Reactor Analysis and Design*: Wiley; 2010.

Chapter 7

Conclusions and future work

7.1. Conclusions

In literature one can find many examples of detailed kinetic models successfully generated by automatic network generation programs. Next to these accomplishments the challenges that constructors of the software still have to overcome is brought to our attention as well. Generating all possible reactions and species based on reaction rules is not straightforward, since many reactions lead to large molecules and convergence problems. Also, several of these species are insignificant because they are improbable to be formed in reality. This results in large kinetic models that are difficult to use and require a lot of computer memory. Ideally the model would contain only those reactions and species which are important under the operating conditions at hand. The algorithms that currently exist to terminate the network generation and to exclude insignificant species from the model still require some optimization. The carbon count and rank-based criterion are two examples of algorithms that arbitrarily neglect species and reactions from the model risking that important reactions will be left out, hence they are not used anymore in the existing kinetic model builders. The rule-based and rate-based approach are two methods that rely on more logical reasoning. The rule-based method generates only species that are significant by defining a certain amount of constraints. These constraints are solely based on knowledge about the chemistry under study. The rate-based approach does not rely on constraints and generates any species possible starting from an initial set of reactants. Only the ones with a high rate of formation are deemed as significant and are added to the model. This approach is more computational intensive but, in general, no prior knowledge of the chemistry is necessary. The rate-based algorithm has proven its capability to generated accurate kinetic models. A downside is the exponential growth of the generated edge species list resulting in memory limitations and hence more often does the rate-based generation method run into convergence problems. Another crucial factor of the rate-based algorithm is the necessity of highly accurate thermochemical data. Kinetic parameters decide whether a reaction should be added to the model or not. If a certain elementary step is left out because the rate of formation was to low due to

inaccurate kinetic parameters, important reaction pathways can be forgotten and this results in an incomplete kinetic model. There is good reason to expect that the accuracy and availability of thermochemical data will continue to increase rapidly. Experimental techniques and computational methods are improving while the scientific community moves towards open-source information meaning the amount of available information can expand rapidly.

The rate-based algorithm is implemented next to the existing rule-based algorithm in the code of Genesys. The software tool CHEMKIN has been chosen to provide the necessary numerical techniques. These techniques are required to solve the reaction network during each iteration step. When rules or constraints are used to create the edge species, their total amount is significantly reduced. The precision level, conversion and reactor type are rate-based parameters that the user must define before network generation process can start. These parameters influence the structure of the model. When the precision level decreases, the number of reactions and species in the model increases and the computational time rises. However, this increase is limited because the number of edge species that is generated is limited by the reaction families constraints.

The newly developed algorithm has been tested and validated using n-pentanol pyrolysis as test case. A model was generated with the kinetic model builder Genesys using the combined version of the rate and rule-based method. A second model was generated with Genesys as well, but the rule-based method was used instead. Finally, a third model was generated with the kinetic model builder RMG, a code which is purely based on the rate-based method. After the network generation was complete, the models were compared to each other and to experimental results obtained from a bench scale reactor. The best prediction of the conversion of n-pentanol as a function of temperature was obtained by the model generated with the combined rate and rule-based method. The rule-based model underestimated the conversion and the model generated with RMG overestimated the conversion. For the yields of products such as C_2H_4, CO, CH_4, H_2O, H_2 and acetaldehyde the results varied. All models predicted the yield of acetaldehyde poorly. This could result due to various reasons. Not only could this be the result of important reaction families that are missing reaction families template in Genesys, it could also be that case that the kinetic or thermodynamic parameters are poorly estimated due to wrong group additive values. Comparison of some kinetic parameters and rate coefficients for reactions that appear in every model show that there is a large difference between these values. Based solely on the conversion of pentanol it could be concluded that model generated with the combined rate and rule-based method is the most adequate model to predict the pyrolysis of pentanol. However, the product yields dispute this conclusion. It can be concluded that rate of production analysis is necessary to see whether there are some important reactions missing in the models and if some kinetic parameters of certain reactions should be more accurate.

A combined rate and rule-based termination criterion succeeds to circumvent the memory issues which are common for the rate-based algorithm. However, this combination does not resolve the issue of inaccurate thermodynamic and kinetic data. If no accurate data is available the rate-based algorithm might still include or exclude species based on an incorrectly calculated rate of formation. The only solution therefore, is to provide highly accurate information in the databases of Genesys.

7.2. Future work

During the estimation of thermochemical properties of species generated with the combined rate and rule-based method, there is a possibility that for some species no properties could be estimated. This is due to some group additive values which are missing from the databases. If a species is generated without any thermodynamic properties, it is impossible to solve the partial network. To circumvent this issue during this work, species of which no properties could be estimated, are assumed to be insignificant for the model and are removed from the partial network. The reactions in which they occur as products are removed from the network as well. Instead of removing the species it is better to provide an estimation for this species anyway. This estimation could be based on thermochemical properties of species with a similar structure. The user should be notified when such an estimation is used for a species, so he or she can account for this during further simulations with the kinetic model.

The software tool used to numerically solve the partial reaction network can be changed from CHEMKIN to tools such as Cantera or OpenSMOKE, since these require no license. Also, if the numerical solver fails to converge it is currently not implemented in the code of Genesys what steps should be taken to resolve the issue. First of all, the user should be made aware of the convergence problem. The problems often originate from a kinetic model with inaccurate kinetic parameters. One possibility is to let Genesys calculate the rate coefficients of every elementary reaction present in the kinetic model at certain temperatures and search for the minimum and maximum coefficients. Usually, if there is a relatively large difference in order of magnitude, this might be the cause of the convergence problem. This means that a reaction is occurring much faster or slower compared to reality. Incorrect thermodynamic properties could be the cause of this.

For now, the code is altered to treat products of reactions having "REVERSE" as kinetic source listed in the reaction family template in a separate way. These species are not considered as edge species and are added to the partial reaction network instead. This means that these reactions are not added to the network yet until the forward reactions have been generated. If the products would be considered as edge species, without any reactions forming these edge species, this would give errors during the numerical solving of the partial network. For most of the reaction families, the carbon-carbon bond scission reactions are considered as reverse reactions, meaning that the kinetics are estimated based on the recombination reactions. As a results, these initiation reactions are only added to the model once the

formed radical are present in the model and the recombination reactions are generated. It would be ideal if the code was adjusted in such a way that if a reverse reaction was found, the forward reaction is searched for in a following step.

Concerning the generated models for the pyrolysis of n-pentanol, there is still room to improve the constraints of the considered reaction families. The model generated with RMG should be further improved as well. Overall, there is a rate of production analysis required to identify the missing reactions or the incorrect rate coefficients.

A
Kinetic models for steam cracking of ethane

Elementary reactions	Genesys			NIST		
	A	n	E_a	A	n	E_a
C2H5r+Hr=C2H6	1.21E+14	0	1.3	1.00E+14	0	0
C2H5r+C3H6_Olef=C2H6+C3H5r_Olef	1.42E+12	0	89.8	1.02E+09	3.5	27.77
Hr+C2H4_Olef=H2+C2H3r_Olef	7.23E+14	0	72.1	5.42E+14		62.36
C3H6_Olef+C2H5r=C5H11r_2	9.42E+10	0	47.3	1.20E+10	2.48	25.65
CH3r+C2H5r=C3H8	1.61E+13	0	-4.9	3.37E+13	0	0
Hr+C3H8=H2+C3H7r	5.74E+14	0	52.3	2.55E+12	2.54	28.27
Hr+C4H10=H2+C4H9r_1	5.74E+14	0	52.3	5.22E+11	2.4	10.81
C2H5r+C2H5r=C4H10	1.20E+13	0	-1.4	1.15E+13	0	0
Hr+C4H8_Olef=H2+C4H7r_Olef_1	2.87E+14	0	52.3	1.72E+14	0	33.42
Hr+C3H6_Olef=H2+C3H5r_Olef_1	7.28E+13	0	38.4	2.61E+11	2.5	10.39
C2H4_Olef+C2H5r=C4H9r_1	7.76E+11	0	41	1.81E+10	2.48	25.65
CH3r+C2H6=CH4+C2H5r	4.56E+13	0	74	4.33E+09	4	34.67
Hr+C4H10=H2+C4H9r	3.98E+14	0	41.2	3.48E+12	2.54	28.27
C3H6_Olef+CH3r=C4H9r	1.83E+12	0	41.7	2.41E+10	2.48	25.65
C3H6_Olef+Hr=C3H7r_1	5.56E+13	0	16.1	7.77E+12	0.51	5.15
Hr+C3H6_Olef=H2+C3H5r_Olef_2	1.79E+14	0	60	1.20E+12	2.5	51.38
C3H6_Olef+Hr=C3H7r	6.46E+13	0	22.5	4.57E+12	0.51	10.96
C2H4_Olef+C2H3r_Olef=C4H7r_Olef_1	5.56E+12	0	27.8	2.00E+11	0	8.4
CH3r+CH3r=C2H6	2.18E+13	0	-8.3	6.20E+13	-1.1	1.33
CH3r+C2H4_Olef=CH4+C2H3r_Olef	3.50E+13	0	84.4	4.16E+12	0	46.56
C2H3r_Olef+C2H5r=C4H8_Olef	1.04E+13	0	-0.4	3.91E+13	0	0

Elementary reactions	Genesys			NIST		
	A	n	E_a	A	n	E_a
Hr+C2H6=H2+C2H5r	7.03E+14	0	53.5	7.41E+12	1.5	31.01
Hr+C3H8=H2+C3H7r_1	2.44E+14	0	42.4	1.13E+12	2.4	18.71
C2H3r_Olef+C2H6=C2H4_Olef+C2H5r	4.43E+13	0	53.9	8.79E+10	3.3	43.9
C2H5r+C3H6_Olef=C2H6+C3H5r_Olef_1	9.17E+11	0	63.3	6.93E+10	0	21.7
CH3r+Hr=CH4	2.84E+14	0	-2.3	2.11E+14	0	0
C2H4_Olef+CH3r=C3H7r	7.43E+12	0	43.1	2.41E+10	2.48	25.65
C2H3r_Olef+C3H6_Olef=C2H4_Olef+C3H5r_Olef_1	5.49E+12	0	37.8	1.01E+09	3.5	19.62
Hr+CH4=H2+CH3r	1.28E+15	0	66.4	2.63E+11	3.16	36.63
C2H4_Olef+Hr=C2H5r	2.13E+14	0	18.5	7.53E+12	1.07	6.07
CH3r+C3H8=CH4+C3H7r_1	1.63E+13	0	64.4	9.70E+08	3.65	29.93
CH3r+C3H6_Olef=CH4+C3H5r_Olef_1	5.65E+12	0	57.9	1.01E+09	3.5	23.78
CH3r+C3H8=CH4+C3H7r	4.56E+13	0	74	5.48E+08	3.46	22.95

B
Reaction families for the steam cracking of ethane

Reaction family	Bimolecular	Reactants	Molecule constraints	Kinetics
Hydrogen abstraction to a carbon centered group by a carbon centered radical	True	Reactant 1	Maximum 7 carbon atoms	Group additivity
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2: carbon radical	Maximum 7 carbon atoms	
			Maximum 1 single electron	
Carbon centered radical addition (non-cyclic components)	True	Reactant 1:	Maximum 7 carbon atoms	Group additivity
			No rings	
			No single electrons	
		Reactant 2:carbon radical	Maximum 7 carbon atoms	
			Maximum 1 single electron	
Hydrogen radical addition to hydrocarbons (non -cyclic components)	True	Reactant 1:	Atom not aromatic?	Group additivity
			No single electrons	
			Maximum 7 carbon atoms	
		Reactant 2: hydrogen radical	Maximum 1 single electron	
Carbon-centered beta-scission onto a carbon-centered radical	False	Reactant 1: carbon radical	Maximum 7 carbon atoms	Reverse
			No rings	
			Maximum 1 double bond	

		Maximum 1 single electron		
Hydrogen-centered beta-scission onto a carbon-centered radical	False	Reactant 1: carbon radical	Maximum 4 carbon atoms	Reverse
			Maximum 2 double bonds	
			Maximum 1 single electron	
H abstractions by a hydrogen radical (non-cyclic components)	True	Reactant 1: hydrogen radical	Maximum 1 single electron	Group additivity
		Reactant 2	No single electrons	
			Maximum 7 carbon atoms	
H abstractions by a carbon radical onto H_2	True	Reactant 1: Carbon radical	Maximum 7 carbon atoms	Reverse
			Maximum 1 single electron	
		Reactant 2: H_2	No single electrons	
Hydrogen recombination	True	Reactant 1: carbon radical	Maximum 7 carbon atoms	Group additivity
			Maximum 1 single electron	
		Reactant 2: hydrogen radical	Maximum 1 single electron	
Carbon-carbon scission	False	Reactant	Maximum 7 carbon atoms	Reverse
			Minimum 2 carbon atoms	
			No single electrons	
Carbon hydrogen scission	False	Reactant	Maximum 7 carbon atoms	Reverse
			Minimum 1 carbon atoms	
			No single electrons	
Bimolecular radical recombination	True	Reactant 1: carbon radical	Maximum 1 single electron	Group additivity
			Maximum 7 carbon atoms	
		Reactant 2: carbon radical	Maximum 1 single electron	
			Maximum 7 carbon atoms	

C
Reaction families for the pyrolysis of n-pentanol

Reaction Family	Bimolecular	Reactants	Molecular constraints	Kinetics
H abstractions to a carbon centered group by a carbon centered radical	True	Reactant 1: Carbon radical	Maximum 2 carbon atoms	Group additivity
			Maximum 1 oxygen atom	
			1 single electron	
		Reactant 2	Maximum 5 carbon atoms	
			Maximum 1 oxygen atom	
			No single electrons	
			Maximum 1 double bon	
H abstractions of aldehydes to a carbon centered group by a carbon centered radical	True	Reactant 1: carbon radical	1 single electron	Arrhenius
			Maximum 3 carbon atoms	
			Maximum 1 oxygen atom	
		Reactant 2: aldehyde	No single electrons	
			Maximum 5 carbon atoms	
			Maximum 1 oxygen atom	
			Maximum 1 double bond	
Carbon-centered radical addition1_4	True	Reactant 1:	Maximum 4 carbon atoms	Group additivity
			Maximum 1 oxygen atom	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2: Carbon radical	Maximum 1 carbon atom	
			No oxygen atoms	
			1 single electron	

Reaction Family	Bimolecular	Reactants	Molecular constraints	Kinetics
Carbon-centered radical addition 1_4bis	True	Reactant 1:	Maximum 4 carbon atoms	Group additivity
			Maximum 1 oxygen atom	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2: Carbon radical	Maximum 1 carbon atom	
			Maximum 1 oxygen atom	
			1 single electron	
Carbon centered radical addition 2_3	True	Reactant 1:	Maximum 3 carbon atoms	Group additivity
			Maximum 1 oxygen atom	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2: carbon radical	Maximum 2 carbon atoms	
			No oxygen atoms	
			1 single electron	
Carbon centered radical addition 2_3bis	True	Reactant 1:	Maximum 3 carbon atoms	Group additivity
			No oxygen atoms	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2: carbon radical	Maximum 2 carbon atoms	
			Maximum 1 oxygen atom	
			1 single electron	
Carbon centered radical addition 3_2	True	Reactant 1	Maximum 2 carbon atoms	Group additivity
			Maximum 1 oxygen atom	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2:Carbon radical	Maximum 3 carbon atoms	
			No oxygen atoms	
			Maximum 1 double bond	
			1 single electrons	
Carbon centered radical addition 3_2bis	True	Reactant 1	Maximum 2 carbon atoms	Group additivity
			No oxygen atoms	
			Maximum 2 double bonds	
			No single electrons	
		Reactant 2:Carbon radical	Maximum 3 carbon atoms	
			Maximum 1 oxygen atom	
			Maximum 1 double bond	
			1 single electrons	

Reaction Family	Bimolecular	Reactants	Molecular constraints	Kinetics
Hydrogen radical addition (non-cyclic components)	True	Reactant 1:	No single electrons	Group additivity
			Maximum 5 carbon atoms	
			Maximum 1 oxygen atom	
			Maximum 2 double bond counts	
		Reactant 2: hydrogen radical	1 single electron	
Carbon-centered beta-scission onto a carbon-centered radical	False	Reactant: carbon radical	1 single electron	Reverse
			No rings	
			Maximum double bond count	
			Maximum 5 carbon atoms	
Hydrogen-centered beta-scission onto a carbon-centered radical	False	Reactant: carbon radical	1 single electron	Reverse
			Maximum 5 carbon atoms	
			Maximum 1 double bond	
H abstraction by a hydrogen radical	True	Reactant 1: hydrogen radical	1 single electron	Group additivity
		Reactant 2	No single electrons	
			Maximum 1 oxygen atom	
			Maximum 5 oxygen atoms	
			Maximum 1 double bond	
H abstractions of aldehydes by a hydrogen radical	True	Reactant 1: hydrogen radical	1 single electron	Arrhenius
		Reactant 2: aldehyde	No single electrons	
			Maximum 5 carbon atoms	
			Maximum 1 oxygen atom	
			Maximum 1 double bond	
H abstractions by a carbon radical onto H_2	True	Reactant 1 : carbon radical	1 single electron	Reverse
			Maximum 5 carbon atoms	
			Maximum 1 double bond	
		Reactant 2: H_2	No single electrons	

Carbon recombination 4_1	True	Reactant 1: carbon radical	1 single electron	Group additivity
			Maximum 4 carbon atoms	
			Maximum 1 oxygen atom	
		Reactant 2: carbon radical	1 single electron	
			Maximum 1 carbon atom	
			No oxygen atoms	
Carbon recombination 3-2	True	Reactant 1: carbon radical	1 single electron	Group additivity
			Maximum 3 carbon atoms	
			Maximum 1 oxygen atom	
		Reactant 2: carbon radical	1 single electron	
			Maximum 2 carbon atoms	
			No oxygen atoms	
Carbon recombination 4-1bis	True	Reactant 1: carbon radical	1 single electron	Group additivity
			Maximum 4 carbon atoms	
			No oxygen atoms	
		Reactant 2: carbon radical	1 single electron	
			Maximum 1 carbon atom	
			Maximum 1 oxygen atom	
Carbon recombination 3-2bis	True	Reactant 1: carbon radical	1 single electron	Group additivity
			Maximum 3 carbon atoms	
			No oxygen atoms	
		Reactant 2: carbon radical	1 single electron	
			Maximum 2 carbon atoms	
			Maximum 1 oxygen atom	
Hydrogen recombination	True	Reactant 1: carbon radical	1 single electron	Group additivity
			Maximum 5 carbon atoms	
			Maximum 1 oxygen atom	
		Reactant 2: hydrogen radical	1 single electron	
Carbon-carbon scission	False	Reactant	No single electrons	Reverse
			Minimum 2 carbon atoms	
			Maximum 5 carbon atoms	
Carbon-hydrogen scission	False	Reactant	No single electrons	Reverse
			Minimum 1 carbon atom	
			Maximum 5 carbon atoms	
			Maximum 1 double bond	
CO formation	False	Reactant	1 single electron	Arrhenius
			Minimum 2 carbon atoms	
			Maximum 5 carbon atoms	
Ketene formation	False	Reactant	1 single electron	Arrhenius
			Maximum 5 carbon atoms	

D
References to lab-journal

Model generation

Topic	Page lab journal	Directory
Pentanol model generation files RMG	4	\Case Study Pentanol\Model generation\RMG
Pentanol model generation files rule-based	4	\Case Study Pentanol\Model generation\Genesys rule based
Pentanol model generation files ombined rate and rule-based	4	\Case Study Pentanol\Model generation\Genesys combined rule and rate based
Ethane model generation rule-based	5-6	\Model Generation Ethane Cracking\rule_based
Ethane model generation combined rate and rule-based with different precision level and conversion	5-6	\Model Generation Ethane Cracking\influence_parameters
Ethane model generation combined rate and rule-based with no constraints	5-6	\ \Model Generation Ethane Cracking\no_constraints
Ethane model generation with energy equation solved	5-6	\Model Generation Ethane Cracking\energyequationbatch

CHEMKIN simulations

Topic	Page lab journal	Directory
RMG simulations pentanol 4 to 1 dilution (640-800°C)	7-8	\Case Study Pentanol\Chemkin simulations\Genesys combined rule and rate based
Rule-based model simulations pentanol 4 to 1 dilution (640-800°C)	7-8	\Case Study Pentanol\Chemkin simulations\Genesys rule based
Combined rate and rule-based model simulations 4 to 1 dilution (640-800°C)	7-8	\Case Study Pentanol\Chemkin simulations\Genesys combined rule and rate based
Chemkin Simulation of ethane steam cracking model with NIST kinetic parameters	9	\Comparison NIST Genesys\Ethane steam cracking reduced

Algorithms

Topic	Page lab journal	Directory
Launch ChemkinParallelizer	2	\Genesys\src\parameterestimator\LaunchCheminParallelizer
Launch ModelAnalyzer	3	\Genesys\src\parameterestimator\LaunchModelAnalyzer