

**Abschlußbericht  
für das Forschungsvorhaben**

**"Aufbau einer Datenfaktenbank für Reinstoffdaten"**

**BMBF - FKZ 08 G35 27 9**

Prof. Dr. J. Gmehling  
Lehrstuhl für Technische Chemie  
Universität Oldenburg  
Postfach 2503, 26111 Oldenburg  
Telefon: 0441/7983831  
Telefax: 0441/7983330  
Email: gmehling@tech.chem.uni-oldenburg.de

**Oldenburg, 15. 12. 1998**

## Abschlußbericht

### Aufbau einer Datenfaktenbank für Reinstoffdaten

	30.7.98
Gmehling, Jürgen	15.12.98
Prof. Dr. J. Gmehling Lehrstuhl für Technische Chemie Universität Oldenburg Postfach 2503 26111 Oldenburg	08 G35 27 40 10 6 22

Zum Ende des Projekts "Datenfaktenbank Reinstoffdaten" stehen ca. 90 000 Datensätze mit zusammen mehr als 570 000 Datenreihen aus ca. 10 000 Referenzen in computergerechter Form gespeichert zur Verfügung. Die Arbeiten wurden in Zusammenarbeit mit Gruppen in Prag, Berlin, Tallinn und Minsk durchgeführt.

Zusätzlich zur Sammlung der Daten wurden verschiedene Korrelationsmethoden implementiert und zum Test der Daten eingesetzt. Mehr als 8000 Parametersätze für verschiedene Gleichungen und Methoden wurden angepaßt.

Durch die konsequente Erfassung der Molekülstrukturen und die Entwicklung von Programmen zur Eingabe, Speicherung und Zerlegung dieser Strukturen stehen erstmals eine Vielzahl von Gruppenbeitragsmethoden aus der Literatur auch dem ungeübten Benutzer zur Verfügung.

Durch die einfache Möglichkeit der Datenabschätzung ohne langwierige und fehlerbehaftete manuelle Strukturzerlegung wurde ein umfangreicher Test der publizierten Methoden möglich. Dieser Test diente vordringlich der Überprüfung der eingegebenen Daten.

Unter Verwendung der automatischen Strukturzerlegung wurden Programme zur Entwicklung neuer Gruppenbeitragsmethoden entwickelt und erste Abschätzmethoden fertiggestellt.

Die Arbeiten werden mit privatwirtschaftlicher Förderung fortgesetzt.

Datenfaktenbank, Reinstoffdaten, Gruppenbeitragsmethoden, Molekülstrukturen,  
Stoffdatenkorrelation

## Final Report

### Development of a Factual Data Base for Pure Component Properties

Gmehling, Jürgen	08 G35 27
	15.12.98
Prof. Dr. J. Gmehling Lehrstuhl für Technische Chemie Universität Oldenburg Postfach 2503 26111 Oldenburg	40
	10
	6
	22

With the end of the pure component property data base project approx. 90 000 data sets containing more than 570 000 table lines from approx. 10 000 references are available in electronic form. The work was performed in cooperation with groups in Prague, Berlin, Tallinn and Minsk.

In addition to collecting experimental data different correlation methods were implemented and used to test the correctness and consistency of the data. More than 8000 parameter sets were regressed for different correlation equations and physical properties.

By storing the molecular structures of more than 12000 components and developing programs for input, storage and fragmentation of these structures a large number of group contribution methods from literature are readily available also for the untrained user.

With this simple possibility to estimate properties using group contribution methods and the availability of a large database, a broad test of published group contribution methods can be performed. These methods were used for an extensive test of the stored data.

Programs for the development of group contribution methods were implemented and the first predictive methods were developed.

The work will be continued with the help of private funding.

Factual Database, Pure Component Properties, Group Contribution Methods,  
Molecular Structure, Physical Property Correlation

# **INHALTSVERZEICHNIS**

<b>1. FORMALANGABEN</b>	<b>3</b>
1.1 Förderkennzeichen	3
1.2 Thema des Vorhabens	3
1.3 Zuwendungsempfänger	3
1.4 Ausführende Stelle	3
1.5 Projektleiter	3
1.6 Bewilligungszeitraum (einschließlich Verlängerung)	3
1.7 Berichtszeitraum (einschließlich Verlängerung)	3
<b>2. KURZDARSTELLUNG</b>	<b>4</b>
2.1 Aufgabenstellung	4
2.2 Voraussetzungen, unter denen das Vorhaben durchgeführt wurde	5
2.3 Planung und Ablauf des Vorhabens	5
2.4 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde	5
2.5 Zusammenarbeit	6
<b>3. EINGEHENDE DARSTELLUNG</b>	<b>7</b>
3.1 Ergebnisse	7
3.1.1 Detaillierte Darstellung des Inhalts der Datenbank und der Arbeitsleistung der einzelnen Kooperationspartner und Auftragnehmer	8
3.1.1.1 Stand der Arbeiten in Oldenburg	9
3.1.1.2 Stand der Arbeiten in Prag	10
3.1.1.3 Stand der Arbeiten in Berlin	11
3.1.1.4 Stand der Arbeiten in Tallinn	12
3.1.1.5 Stand der Arbeiten in Minsk	13
3.1.1.6 Beitrag der DDBST GmbH	14
3.1.2 Datenbankstruktur	14
3.1.3 Inhalt der Datenbank	14
3.1.4 Test und Korrelation der Daten	14
3.1.4.1 Test und Korrelation der Sättigungsdampfdrücke	15
3.1.4.2 Test und Korrelation der Viskositätsdaten	16
3.1.4.3 Test und Korrelation der kritischen Daten	17
3.1.4.4 Test und Korrelation der Flüssigdichten	19
3.1.4.5 Test und Korrelation der Wärmekapazitäten	21
3.1.4.6 Test und Korrelation der Entropien, Enthalpien, ...	23
3.1.4.7 Test und Korrelation der thermodynamischen Bildungsgrößen	24
3.1.4.8 Test und Korrelation der Wärmeleitfähigkeiten	25
3.1.4.9 Test und Korrelation der Oberflächenspannungen	26
3.1.5 Methoden und Programme zur Strukturverarbeitung	27
3.1.7 Entwicklung eigener Abschätzmethode	30

3.1.7.1	Abschätzung von Normalsiedepunkten nach Cordes und Rarey	30
3.1.7.2	Abschätzung von Flüssigkeitswärmekapazitäten nach Becker und Aufderhaar	32
<b>3.2</b>	<b>Nutzen der Datenbank und der erstellten Programme</b>	<b>35</b>
<b>3.3</b>	<b>Fortschritt bei anderen Stellen</b>	<b>36</b>
<b>3.4</b>	<b>Erfolge und geplante Veröffentlichungen</b>	<b>36</b>
<b>4.</b>	<b>KURZGEFAßTER ERFOLGSKONTROLLBERICHT</b>	<b>37</b>
<b>4.1</b>	<b>Beitrag des Ergebnisses zu den förderpolitischen Zielen des Förderprogramms</b>	<b>37</b>
<b>4.2</b>	<b>Wissenschaftlicher Erfolg des Vorhabens</b>	<b>37</b>
4.2.1	Nebenergebnisse und Erfahrungen	38
<b>4.3</b>	<b>Einhaltung des Finanzierungs- und Zeitplans</b>	<b>38</b>
<b>4.4</b>	<b>Verwertbarkeit der Ergebnisse</b>	<b>38</b>
<b>4.5</b>	<b>Arbeiten, die zu keiner Lösung führten</b>	<b>39</b>
<b>5.</b>	<b>KURZFASSUNG DES SCHLUßBERICHTS</b>	<b>39</b>
<b>6.</b>	<b>ANLAGE</b>	

**Abschlußbericht**  
**für das Forschungsvorhaben**  
**"Aufbau einer Datenfaktenbank für Reinstoffdaten"**

**1. Formalangaben**

**1.1 Förderkennzeichen**

FKZ 08 G35 27 9

**1.2 Thema des Vorhabens**

Aufbau einer Datenfaktenbank Reinstoffdaten

**1.3 Zuwendungsempfänger**

Lehrstuhl für Technische Chemie  
Universität Oldenburg  
Postfach 2503, 26111 Oldenburg  
Telefon: 0441/7983831  
Telefax: 0441/7983330

**1.4 Ausführende Stelle**

Lehrstuhl Technische Chemie

**1.5 Projektleiter**

Prof. Dr. Jürgen Gmehling

**1.6 Bewilligungszeitraum (einschließlich Verlängerung)**

1. 7. 1995 - 31. 7. 1998

**1.7 Berichtszeitraum (einschließlich Verlängerung)**

1. 7. 1995 - 31. 7. 1998

## **2. Kurzdarstellung**

### **2.1 Aufgabenstellung**

Reinstoffdaten sind in allen Bereichen des Chemieingenieurwesens für die Auslegung der verschiedensten Anlagen sowie für die Beurteilung vieler Problemstellungen in Zusammenhang mit den verschiedenen Chemikalien von besonderer Bedeutung. Zuverlässige Reinstoffdaten stellen weiterhin die Grundvoraussetzung bei der Darstellung von Gemischeigenschaften dar.

Im Zuge zunehmender Bedeutung von Spezialchemikalien, der Entwicklung neuer Technologien und der Bearbeitung von Umweltschutzfragestellungen werden immer häufiger Informationen über Komponenten benötigt, für die bisher nur wenig oder keine experimentellen Daten vorliegen. Messungen sind dabei in den meisten Fällen zu zeitaufwendig oder teuer, in vielen Fällen auch nur schwer durchzuführen, da viele Stoffe nicht schnell oder hinreichend rein verfügbar oder besonders giftig bzw. krebserregend sind bzw. sich im Verlauf der Messung zersetzen.

Die heute verfügbaren Sammlungen für Datenkorrelationen (ESDU, PPDS, DIPPR, DETHERM, INFOTHERM, TRC etc.) beschränken sich weitgehend auf die Erfassung von einigen hundert bis tausend häufig verwendeten Chemikalien und beinhalten nicht oder nur in geringem Maße experimentelle Daten. Weiterhin wurde in diesen Datensammlungen die russische, ostasiatische, ... Literatur zum Teil vollends vernachlässigt und bei der Anpassung der Korrelationsparameter nicht sorgfältig genug gearbeitet.

Nachdem im Vorgängerprojekt (08 G35 26 8) eine umfangreiche Datenbank aufgebaut wurde, sollte diese nun in diesem Projekt weiter vervollständigt werden (Literaturbeschaffung, Übersetzung, Auswertung, Dateneingabe). Dabei sollten die eingegebenen Daten mit Hilfe verschiedener Methoden auf Richtigkeit und innere Konsistenz überprüft werden.

Nach dem Aufbau der Datenbank sollen an die gespeicherten Reinstoffdaten mit Hilfe sinnvoller thermodynamischer Modelle und Verwendung des benötigten mathematischen Rüstzeugs Parameter angepaßt werden, so daß zuverlässige Interpolationen und Extrapolationen durchgeführt werden können. Bei der Anpassung soll dabei insbesondere auf eine konsistente Darstellung Wert gelegt werden.

In der Praxis besteht aus diesen Gründen ein großer Bedarf an einer möglichst vollständigen Datensammlung experimenteller Reinstoffdaten sowie an einem Programmpaket, mit Hilfe dessen Berechnungen unter Verwendung der verschiedenen Methoden leicht durchgeführt werden können. Dabei wird die Information über die Molekülstruktur direkt aus einer geeigneten Kodierung übernommen, welche entweder in einer Datenbank vorliegt oder in einfacher Weise durch grafische Eingabe der Strukturformel erhalten werden kann. Es sollten Modelle zur Berechnung von Reinstoffdaten aus Molekülstrukturen mit Hilfe der in der ersten Projekthälfte erstellten Programme bereitgestellt werden.

Die erstellten Datenbanken und Abschätzmethoden sollten unter einer modernen Oberfläche (Windows) unter Berücksichtigung der Qualitätsabschätzung verbunden werden.

Ziel des Projektes war es damit, ein System zu entwickeln, welches die weltweit verfügbaren experimentellen Daten mit von Fachleuten erarbeiteten Korrelationen und verschiedensten Methoden der Stoffdatenabschätzung verbindet und den Anwendern online oder als Inhouse-Version zur Verfügung gestellt werden kann.

## **2.2 Voraussetzungen, unter denen das Vorhaben durchgeführt wurde**

Das Projekt wurde auf der Basis langjähriger Erfahrung auf dem Gebiet der Phasengleichgewichtsthermodynamik und dem Aufbau von Datenbanken in meiner Arbeitsgruppe durchgeführt. Seit 1973 wurden in meiner Arbeitsgruppe verschiedene Datenbanken zum Thema Phasengleichgewichtsverhalten von binären und mehrkomponentigen Stoffsystemen erfolgreich erstellt. Im Rahmen dieser Datenbanken war auch immer die Recherche nach verlässlichen Reinstoffdaten, deren Beurteilung und Speicherung Voraussetzung für den Test und die erfolgreiche Entwicklung von Berechnungsmethoden ( z.B. UNIFAC, mod. UNIFAC, PSRK, ... ).

Die daraus resultierende Erfahrung in Hinblick auf die Programmerstellung, Literaturrecherche und Auswertung der Literaturstellen, sowie die Erstellung eines geeigneten Datenformats waren bei der Bearbeitung des Projekts von großem Vorteil.

Für die programmtechnische Bearbeitung der Programme zur Strukturingabe und -zerlegung wurden Programmierer der DDBST GmbH eingesetzt, die auch schon am ersten Projekt beteiligt waren.

Die bei den Kooperationspartnern und Auftragnehmern bestehenden Voraussetzungen decken sich mit den im ersten Projekt beschriebenen. Durch die inzwischen mehrjährige Mitarbeit seit 1992 besteht ein beträchtliches hinzugewonnenes Know-How bezüglich der zu erledigenden Aufgaben und der speziellen Arbeitsweisen im Projekt.

## **2.3 Planung und Ablauf des Vorhabens**

Das vorliegende Projekt stellt inhaltlich und arbeitstechnisch eine Fortsetzung des Vorgängerprojekts (08 G35 26 8) dar, so daß ein nur geringer Planungsbedarf bestand. Es konnten die im ersten Projekt begonnenen Arbeiten unterbrechungsfrei fortgesetzt werden. Abweichend vom ersten Projekt war es erforderlich, nach umfangreichen Datentests im Jahre 1998 dem FIZ CHEMIE eine vertriebsfähige Version der Datenbank zur Verfügung zu stellen. Dazu wurde im November 1997 der Stand der Datenbank für einige Monate eingefroren und die enthaltenen Daten getestet. Seit April 1998 wurden dann die aufgelaufenen Datenlieferungen zügig integriert. Aufgrund der Datenmenge (ca. 10000 Datensätze nach Einfrieren der Datenbank) erforderte dies einen großen Arbeitsaufwand. Aus diesem Grund wurde das Projekt kostenneutral um einen Monat verlängert. Seit Anfang 1998 hat sich der Antragsteller intensiv bemüht, die Fortführung der Arbeiten nach Projektende sicherzustellen. Es konnten zu diesem Zweck private Förderer (FIZ CHEMIE, DDBST) gefunden werden, so daß eine weitere Pflege, die Eingabe noch fehlender Daten und das regelmäßige Update sichergestellt sind.

## **2.4 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde**

Bei der Bedeutung thermophysikalischer Daten zur Lösung der verschiedensten Fragestellungen ist es nicht verwunderlich, daß bereits vor Beginn des ersten Projekts große Anstrengungen der verschiedensten Institutionen unternommen wurden, die experimentellen Daten bzw. die Parameter von Korrelationsgleichungen in computergerechter Form zu speichern. So wurden insbesondere vom Physical Property Data Service ( PPDS ), ESDU und dem Design Institute for Physical Property Data ( DIPPR ) Datenbanken erstellt, die zu Beginn des Projektes Informationen über ca. 1200 Komponenten ( DIPPR ) enthielten.

Reinstoffdaten wurden auch im DETHERM-Paket ( DECHEMA, Frankfurt ) gespeichert. Zu Beginn des Projekts waren bei DETHERM Reinstoffdaten von ca. 5000 Komponenten (einen

beträchtlichen Teil davon stellen allerdings als Pseudokomponenten behandelte Erdölfraktionen etc. dar) abrufbar. Weiterhin wurden Reinstoffdaten auch in der INFOTHERM-Datenbank ( ZIC, Berlin ), Reinstoffdatenbank Böhlen und COMDOR-Datenbank erfaßt. Bei allen bisher erstellten Reinstoffdatenbanken wurde die schwierig auszuwertende Literatur ( russische, ostasiatische, .. Zeitschriften ) nicht ausreichend berücksichtigt und auch die Parameterkorrelationen nicht immer mit der ausreichenden Sorgfalt durchgeführt.

Für die Anwendung im Chemieingenieurwesen müssen die Stoffdaten aufgrund verbesserter Meßtechniken immer wieder auf den neuesten Stand gebracht werden. Die einmalige Speicherung von Werten wie z.B. bei der SOW-Datenbank ( Sächsische Olefinwerke ) ist deshalb weitgehend ungeeignet. Die Daten müßten immer wieder in Frage gestellt werden

Schon zu Beginn des Folgeprojekts stellte die hier entwickelte Reinstoffdatenbank die größte Datensammlung ihrer Art dar und konnte im hier beschriebenen Projekt deutlich vervollständigt werden

## 2.5 Zusammenarbeit

Aufgrund des großen Umfangs der hier erstellten Datenbank wurden schon in das erste Projekt weitere Arbeitsgruppen integriert. Die Erfassung der Daten sollte wie auch schon in der ersten Projekthälfte vorzugsweise in Ländern mit niedrigem Lohnniveau durchgeführt werden. Dabei durfte das Lohnniveau aber nicht das einzige Kriterium darstellen, da für die Arbeiten teilweise erhebliche organisatorische, fachliche und personelle Ressourcen benötigt werden.

Beteiligte Gruppen waren wie auch schon im ersten Projekt:

- Institut für Chemische Technologie ( Institut für Physikalische Chemie, Prag )
- FIZ CHEMIE GmbH, Berlin
- DDBST GmbH, Oldenburg
- die Gruppe von Prof. Kudryavtseva, Tallinn (Estland)
- die Gruppe von Dr. H. Huemer, Graz (Österreich)

Weiterhin wurde ein Gruppe in Weißrußland in geringem Umfang beteiligt. Neben der Exzerption von schwer erhältlichen Quellen war Herr Prof. Shulgin bei mehreren Aufenthalten in Oldenburg auch persönlich bei der Kontrolle von Eingaben aus russischsprachigen Artikeln behilflich:

- Prof. I. Shulgin, Minsk (Weißrußland)

### 3. Eingehende Darstellung

#### 3.1 Ergebnisse

Im Rahmen des Vorhabens wurde in Zusammenarbeit mit

- dem Institut für Chemische Technologie ( Institut für Physikalische Chemie, Prag )
- der FIZ CHEMIE GmbH, Berlin
- der DDBST GmbH, Oldenburg
- der Gruppe von Prof. Kudryavtseva, Tallinn (Estland)
- der Gruppe von Dr. H. Huemer, Graz (Österreich)
- der Gruppe von Prof. I. Shulgin, Minsk (Weißrußland)

eine umfangreiche Datenbank für thermophysikalische Eigenschaften von Reinstoffen ( organische Verbindungen, wichtige anorganische Verbindungen wie Wasser, Ammoniak, .. ) unter Verwendung der weltweit publizierten Informationen aufgebaut.

Die Datenfaktenbank beinhaltet folgende Eigenschaften:

- a) Reinstoffdampfdrücke
- b) Verdampfungsenthalpien
- c) Flüssigkeitsdichten
- d) Flüssigkeitsdichten im Sättigungszustand
- e) Spezifische molare Wärmekapazitäten von Flüssigkeiten
- f) Viskositäten bei nicht zu hohen Drücken
- g) Wärmeleitfähigkeiten
- h) Oberflächenspannungen
- i) Thermodynamische Standardgrößen im idealen Gaszustand
- j) Schmelzpunkte, Schmelzwärmen, Umwandlungsenthalpien ( Kristallformen )
- k) Virialkoeffizienten ( in Kooperation mit Universitätsdozent Dr. Hans Huemer, Institut für Verfahrenstechnik, TU Graz )

Ein gestecktes Ziel bei der Erfassung und der computergerechten Speicherung der experimentellen Daten war dabei, im Gegensatz zu bereits bestehenden Datensammlungen die **gesamte** Literatur, d.h. auch chinesische, russische, ... Fachzeitschriften zu erfassen.

Unter Verwendung der in computergerechter Form gespeicherten Daten wurden Parameter zur Darstellung von Reinstoffgrößen als Funktion der Einflußgrößen ( Temperatur, Druck ) ermittelt.

Die zuverlässige Kenntnis der oben aufgeführten Größen ist insbesondere bei der Auslegung der verschiedensten Prozesse in der Chemischen und Petrochemischen Industrie von großer Bedeutung. Sind keine Daten vorhanden wird zumindest eine Abschätzung der benötigten Reinstoffgrößen erforderlich. Mit Hilfe der im Projekt erstellten Datenbank wurden bereits viele bestehende Abschätzmethoden an einer großen Datenbasis überprüft und erste verbesserte bzw. neue Methoden zur genaueren Vorhersage dieser Größen entwickelt.

Um bestehende oder neu entwickelte Vorausberechnungsmethoden schnell und in benutzerfreundlicher Form an das Programmpaket anschließen bzw. neue Methoden unter Verwendung der gespeicherten Information entwickeln zu können, wurde weiterhin die Struktur von über 12 000 Verbindungen in Form von Konnektivitätsmatrizen gespeichert und ein Programmpaket zur Unterteilung der Struktur in vorgegebene Strukturgruppen integriert. Für die grafische Eingabe der Strukturen und die Ansteuerung der Gruppenzerlegung sowie der verschiedenen integrierten Abschätzmethoden wurde ein Grafikeditor erstellt.

Die im Projekt entwickelte Reinstoffdatenbank wurde in die Datenbank DETHERM (FIZ CHEMIE, DECHEMA) integriert. Die Datenbank wird online und als Inhouse-Version angeboten und wurde über ein Interface an bestehende Prozeßsimulatoren angeschlossen. Diese Schnittstelle steht beispielsweise allen Anwendern des führenden Prozeßsimulators Aspen Plus zur Verfügung. Daten können dabei direkt über das Internet abgerufen werden.

### **3.1.1 Detaillierte Darstellung des Inhalts der Datenbank und der Arbeitsleistung der einzelnen Kooperationspartner und Auftragnehmer**

Die folgende Tabelle liefert eine Übersicht über die Datenlieferungen der beteiligten Gruppen im ersten sowie im zweiten Projekt. Hierbei ist zu beachten, daß es bei diesem Dateninput auch in geringem Umfang zu Doppelterfassungen kommen kann. Dabei handelt es sich nicht zwangsläufig um die doppelte Auswertung der gleichen Literaturstelle. Oft werden gleiche Daten von den Autoren mehrfach in verschiedenen Zeitschriften publiziert. Es ist dabei arbeitsparender, diese doppelten Daten später in der Datenbank zu lokalisieren und zu entfernen als eine doppelte Erfassung zu verhindern. Daneben gibt es weitere Gründe, einige exzerpierten Daten nicht in die endgültige Datenbank aufzunehmen:

- Die Daten sind nicht experimentell (abgeschätzt, geglättet, extrapoliert, zitiert,...)
- Die Daten sind außergewöhnlich fragwürdig.

Ursprünglich sollte das FIZ CHEMIE keine Sättigungsdampfdrücke liefern. Damit fehlten dann aber die kritischen Daten, die in Berlin als Teil der Dampfdruckkurve eingegeben werden. Das FIZ CHEMIE hat inzwischen alle Dampfdruckwerte geliefert, die kritischen Daten werden in Oldenburg herausgefiltert. Die etwa 10000 Tabellenzeilen mit Sättigungsdampfdrücken waren jedoch zuvor schon in Oldenburg exzerpiert worden, wodurch sich der Anteil der als doppelt verworfenen Daten aus Berliner Lieferungen um ca. 12% erhöhte.

In dem im Dezember 1997 bis April 1998 erfolgten intensiven Test von mehr als einer halben Million Tabellenzeilen konnten einige diffizile Probleme nicht geklärt werden, die Daten wurden vorerst nicht in die endgültige Datenbank aufgenommen.

	<b>Gesamt</b>	<b>Erstes Projekt</b>	<b>Zweites Projekt</b>	<b>In der Datenbank</b>	<b>Zurückgelegt oder verworfen</b>	
<b>Oldenburg</b>	218000	94000	124000	218000	0	0%
<b>Berlin</b>	79000	40000	39000	62000	17000	22%
<b>Tallinn</b>	135500	64000	71500	118000	17500	13%
<b>Prague</b>	172000	50000	122000	150000	22000	13%
<b>Minsk</b>	10200		10200	9400	800	8%
<b>Summe</b>	<b>614700</b>	<b>248000</b>	<b>366700</b>	<b>557400</b>	<b>57300</b>	<b>10%</b>

Insgesamt konnten zwischen 87% und 92% der gelieferten Daten in die Datenbank importiert werden. Der Anteil der nicht berücksichtigten Daten aus der Oldenburger Exzerption ist sehr gering, da die Kriterien für DDB-Daten hier schon bei der Eingabe berücksichtigt wurden.

Während im ersten Projekt umfangreiche Arbeiten zur Programmerstellung, Definition der Datenstruktur, Einarbeitung der Mitarbeiter, Aufbau der bibliographischen Datenbanken und Beschaffung der Artikel etc. zu leisten waren, konnte die Datenexzerption im zweiten Projekt auf hohem Niveau weitergeführt werden. Dies führte dazu, daß im zweiten Projekt eine Steigerung des Dateninputs pro Jahr um 50% zu beobachten ist.

Zusätzlich zu den oben genannten Zahlen wurde von der Prager Gruppe noch eine umfangreiche Nachlieferung von voraussichtlich etwa 30000 Tabellenzeilen zugesagt.

Nachdem hier der Zuwachs im zweiten Projekt dargestellt wurde, wird im folgenden bei der Darstellung des Inputs der einzelnen Gruppen jeweils die gesamte Anzahl der eingegebenen Tabellenzeilen aufgeführt, da dies einen besseren Eindruck vom Stand der Arbeiten vermittelt.

### 3.1.1.1 Stand der Arbeiten in Oldenburg

Die hier aufgeführte Tabelle zeigt die Anzahl der Tabellenzeilen, die in der Oldenburger Arbeitsgruppe gesammelt worden sind im Vergleich zum Input aller Gruppen. Mit 35.6% des gesamten Dateninputs nimmt die Oldenburger Gruppe die Spitzenposition ein. Der Schwerpunkt der Arbeit liegt, wie aus der Tabelle hervorgeht, auf den Flüssigkeitsviskositäten, Flüssigkeitswärmekapazitäten und Reinstoffdampfdrücken.

	Anzahl Tabellenzeilen	Summe alle Partner	Anteil Oldenburgs am Dateninput
Viskositäten	63482	109623	<b>57.9%</b>
Dampfdrücke	78030	98085	<b>79.6%</b>
Kritische Größen	162	5126	3.2%
Tripelpunkte	42	126	33.3%
Dichten	9003	95643	9.4%
Schmelzpunkte	1380	10412	13.3%
Wärmekapazitäten ( $c_p$ , $c_v$ )	62904	130921	<b>48.0%</b>
2. Virialkoeffizienten	69	1369	5.0%
Verdampfungswärmen	833	10339	8.1%
Schmelzwärmen	906	4045	22.4%
Oberflächenspannungen	24	17346	0.1%
Wärmeleitfähigkeiten	70	71244	0.1%
Entropien, Enthalpien	18	12646	0.1%
Verbrennungs- und Bildungswärmen	302	39305	0.8%
Umwandlungstemperaturen und -wärmen	323	1780	18.1%
PvT-Daten	726	5796	12.5%
<b>Summe</b>	<b>218274</b>	<b>613806</b>	<b>35.6%</b>

Neben den abgeschlossenen Arbeiten bezüglich der Entwicklung von Programmen zur Verwaltung von Molekülstrukturen (Strukturformeleditor, Datenbank und automatische Strukturanalyse) sowie der Festlegung der Datenbankstruktur war es die Aufgabe der Oldenburger Gruppe, die Arbeit der anderen Gruppen zu koordinieren, deren Daten zu sichten und schließlich in eine DDB-konforme Datenbank aufzunehmen. Weiterhin werden die Daten in Oldenburg abschließend auf Richtigkeit kontrolliert und zum Teil bewertet.

Bei der Literaturrecherche und –beschaffung konnte die Oldenburger Gruppe in gewissem Umfang von den Arbeiten in Prag profitieren. Die mehr als 10000 von Prag als Kopien gelieferten Artikel enthielten auch im größeren Umfang Daten, die dort nicht ausgewertet wurden. Diese Referenzen wurden deshalb in Oldenburg erneut bearbeitet. Dabei konnte auf die Ergebnisse der ersten Durchsicht in Prag zurückgegriffen werden, da die Prager Bibliographie Angaben über die enthaltenen Daten beinhaltet, auch wenn diese dort nicht bearbeitet wurden.

### 3.1.1.2 Stand der Arbeiten in Prag

Die Ergebnisse der Arbeiten in Prag sind als sehr erfreulich zu bewerten. Besonders hervorzuheben war die Kooperationsbereitschaft der Prager Arbeitsgruppe. Durch die bestehende INTERNET-Anbindung ist der Kontakt sehr eng und auf jede Anregung der Oldenburger Arbeitsgruppe wurde umgehend konstruktiv mit dem Willen zur Verbesserung des gemeinsamen Arbeitsergebnisses geantwortet. In Prag erfolgte schwerpunktmäßig die Erfassung von Daten aus Quellen mit geringerer „Datendichte“. Gerade diese Daten erfordern einen großen Arbeitsaufwand und eine systematische Suche nach den Publikationen.

Die Kopien von mehr als 10000 Referenzen wurden von Prag nach Oldenburg gesandt und dienen hier der Kontrolle der Daten.

Die folgende Liste zeigt die detaillierte Auflistung der Größen, die exzerpiert und aufgenommen worden sind.

	<b>Anzahl Tabellenzeilen</b>	<b>Summe alle Partner</b>	<b>Anteil Prags am Dateninput</b>
Viskositäten	9451	109623	8.6%
Dampfdrücke	346	98085	0.4%
Kritische Größen	4680	5126	91.3%
Tripelpunkte	48	126	38.1%
Dichten	20189	95643	21.1%
Schmelzpunkte	6819	10412	65.5%
Wärmekapazitäten ( $c_p$ , $c_v$ )	32637	130921	24.9%
2. Virialkoeffizienten	0	1369	0.0%
Verdampfungswärmen	7915	10339	76.6%
Schmelzwärmen	2332	4045	57.7%
Oberflächenspannungen	14234	17346	82.1%
Wärmeleitfähigkeiten	27493	71244	38.6%
Entropien, Enthalpien	6861	12646	54.3%
Verbrennungs- und Bildungswärmen	37539	39305	95.5%
Umwandlungstemperaturen und -wärmen	1352	1780	76.0%
PvT-Daten	0	5796	0.0%
<b>Summe</b>	<b>171896</b>	<b>613806</b>	<b>28.0%</b>

Ein gesonderter Bericht der Prager Gruppe liegt als Anlage bei.

### 3.1.1.3 Stand der Arbeiten in Berlin

Die Zusammenarbeit mit Berlin gestaltete sich weitgehend reibungslos. Sehr positiv zu vermerken ist die Reaktion auf Anmerkungen der Oldenburger Arbeitsgruppe, die postwendend beantwortet wurden.

Ebenfalls positiv ist die Entwicklung der Datensammlung in Berlin zu bewerten:

Die nachfolgende Tabelle zeigt für die einzelnen Eigenschaften die in Berlin eingegebenen Tabellenzeilen.

	<b>Anzahl Tabellenzeilen</b>	<b>Summe alle Partner</b>	<b>Anteil Berlins am Dateninput</b>
Viskositäten	12717	109623	11.6%
Dampfdrücke	10902	98085	11.1%
Kritische Größen	28	5126	0.5%
Tripelpunkte	16	126	12.7%
Dichten	30154	95643	31.5%

Schmelzpunkte	1731	10412	16.6%
Wärmekapazitäten ( $c_p$ , $c_v$ )	8394	130921	6.4%
2. Virialkoeffizienten	1118	1369	81.7%
Verdampfungswärmen	1377	10339	13.3%
Schmelzwärmen	540	4045	13.3%
Oberflächenspannungen	1005	17346	5.8%
Wärmeleitfähigkeiten	9091	71244	12.8%
Entropien, Enthalpien	880	12646	7.0%
Verbrennungs- und Bildungswärmen	0	39305	0.0%
Umwandlungstemperaturen und -wärmen	0	1780	0.0%
PvT-Daten	0	5796	0.0%
<b>Summe</b>	<b>77953</b>	<b>613806</b>	<b>12.7%</b>

Probleme in der Zusammenarbeit mit Berlin ergaben sich durch das nach wie vor beschränkende Format (ELDAR), das für die Dateneingabe verwendet wird. Dadurch ist es der Arbeitsgruppe bspw. nicht möglich, die Meßmethoden und die Reinheit der verwendeten Substanzen mit zu erfassen.

Kopien der ausgewerteten Artikel wurden der Oldenburger Gruppe in regelmäßigen Abständen zugesandt. Es konnte eine Einigung bzgl einer Systematik zur Zuordnung der Artikel zu den Daten gefunden werden.

#### 3.1.1.4 Stand der Arbeiten in Tallinn

Die Zusammenarbeit mit der Arbeitsgruppe von Frau Prof. Kudryavtseva (seit 1997 Prof. Een Siimer) gestaltet sich sehr erfreulich. Insbesondere zu erwähnen sind die sofortigen und stets befriedigenden Antwort auf alle Anfragen und Wünsche der Oldenburger Arbeitsgruppe. Kommunikation und Datentransfer erfolgen vorwiegend über das Internet.

Sehr positiv an der Zusammenarbeit mit Tallinn ist, daß sie das grundlegendes Problem der großen Anzahl russischsprachiger Artikel löst.

In der nebenstehenden Tabelle sind die Größen aufgelistet, die von der Tallinner Arbeitsgruppe eingegeben worden sind.

Die große Bandbreite an Eigenschaften resultiert daraus, daß keine speziellen Eigenschaften für die Arbeitsgruppe spezifiziert worden sind, sondern sie bestimmte Artikel vollständig auswertet.

	<b>Anzahl Tabellenzeilen</b>	<b>Summe alle Partner</b>	<b>Anteil Tallinn am Dateninput</b>
Viskositäten	22153	109623	20.2%
Dampfdrücke	6821	98085	7.0%
Kritische Größen	202	5126	3.9%
Tripelpunkte	12	126	9.5%
Dichten	35038	95643	36.6%
Schmelzpunkte	252	10412	2.4%

Wärmekapazitäten ( $c_p, c_v$ )	23799	130921	18.2%
2. Virialkoeffizienten	182	1369	13.3%
Verdampfungswärmen	200	10339	1.9%
Schmelzwärmen	221	4045	5.5%
Oberflächenspannungen	2073	17346	12.0%
Wärmeleitfähigkeiten	34289	71244	48.1%
Entropien, Enthalpien	4887	12646	38.6%
Verbrennungs- und Bildungswärmen	1282	39305	3.3%
Umwandlungstemperaturen und -wärmen	64	1780	3.6%
PvT-Daten	3974	5796	68.6%
<b>Summe</b>	<b>135449</b>	<b>613806</b>	<b>22.1%</b>

### 3.1.1.5 Stand der Arbeiten in Minsk

Durch die Mitarbeit der Gruppe von Prof. Shulgin konnten einige ansonsten schwer zugängliche Artikel in die Datenbank integriert werden. Trotz der geringen Anzahl an eingegebenen Daten (1.7% der Datenbank) war diese Zusammenarbeit lohnend. Minsk erhielt nur etwa 1% der Projektmittel.

	Anzahl Tabellenzeilen	Summe alle Partner	Anteil Minsk's am Dateninput
Viskositäten	1820	109623	1.7%
Dampfdrücke	1986	98085	2.0%
Kritische Größen	54	5126	1.1%
Tripelpunkte	8	126	6.3%
Dichten	1259	95643	1.3%
Schmelzpunkte	230	10412	2.2%
Wärmekapazitäten ( $c_p, c_v$ )	3187	130921	2.4%
2. Virialkoeffizienten	0	1369	0.0%
Verdampfungswärmen	14	10339	0.1%
Schmelzwärmen	46	4045	1.1%
Oberflächenspannungen	10	17346	0.1%
Wärmeleitfähigkeiten	301	71244	0.4%
Entropien, Enthalpien	0	12646	0.0%
Verbrennungs- und Bildungswärmen	182	39305	0.5%
Umwandlungstemperaturen und -wärmen	41	1780	2.3%
PvT-Daten	1096	5796	18.9%
<b>Summe</b>	<b>10234</b>	<b>613806</b>	<b>1.7%</b>

Bei mehreren längeren Aufenthalten in Oldenburg war Herr Prof. Shulgin intensiv an Test und Überprüfung der Daten aus russischsprachigen Literaturstellen beteiligt.

### 3.1.1.6 Beitrag der DDBST GmbH

An praktisch allen Arbeiten waren die Programmierer der DDBST GmbH intensiv beteiligt. Dies betrifft sowohl die Entwicklung der Datenstrukturen als auch die Bereitstellung von Programmen zum Test und zur Pflege der Datenbank, Parameterkorrelation, Stoffdatenabschätzung etc.. Praktisch alle Programme sind inzwischen als moderne Windows-Versionen verfügbar, deren Entwicklung von Doktoranden an der Universität nicht geleistet werden kann. Die Durchführung der hier beschriebenen Arbeiten und die Erstellung einer brauchbaren Lieferversion für das FIZ CHEMIE wäre ohne diesen sehr arbeitsintensiven Einsatz nicht denkbar gewesen.

### **3.1.2 Datenbankstruktur**

Die Datenbankstruktur hat sich gegenüber der im Abschlußbericht des ersten Projekts beschriebenen nicht geändert. Es wird deshalb hier auf eine erneute Darstellung verzichtet.

### **3.1.3 Inhalt der Datenbank**

Die Datenbank enthält zum Abschluß des Berichtszeitraums ca. 90 000 Datensätze mit zusammen mehr als 570 000 Datenreihen aus ca. 10 000 Referenzen. Es wurden insgesamt etwa 18 000 Literaturstellen erfaßt. Von etwa 14000 Literaturstellen liegen Kopien vor.

### **3.1.4 Test und Korrelation der Daten**

Zum Test der Daten wurden verschiedene Methoden eingesetzt. Eine sehr leistungsfähige Methode besteht darin, die Daten zuerst einmal für alle Komponenten und Größen, bei denen mehr als ein Datenpunkt vorliegt, in einer Grafik darzustellen. Dabei lassen sich im allgemeinen schnell grobe Fehler oder sogar kleinere Unstimmigkeiten erkennen. Die fraglichen Daten werden dann durch Vergleich der Originalliteratur mit der TFF(Tagged-File-Format)-Datei überprüft. Typische Fehler sind dabei falsche Spezifikation der Einheit oder fehlerhafte Zuordnung der Komponente. Diese Fehler treten nicht zwangsläufig bei der Exzerption der Daten auf, sie finden sich auch in nicht geringer Zahl in der Originalliteratur. Häufig betreffen solche Fehler alle für eine Literaturstelle eingegebene Tabellen, womit auch Fehler bei Komponenten korrigiert werden können, für die keine Vergleichsdaten vorliegen.

Um diese Arbeit zu erleichtern, wurden Programme zur Datenkorrelation und grafischen Darstellung entwickelt, bei dem der Benutzer mit dem Mauszeiger die numerischen Werte und Literaturquellen einzelner Datenpunkte identifizieren kann. Eine weitere Möglichkeit besteht darin, mit Hilfe der verfügbaren Gruppenbeitragsmethoden alle Daten nachzurechnen und die Resultate (experimentelle Werte und berechnete Kurven) in einem Diagramm darzustellen. Beispiele für diese Vorgehensweise finden sich in den folgenden Unterkapiteln. Solche Grafiken geben zusätzlich auch Aufschluß über die zu erwartende Genauigkeit bzw. Verlässlichkeit der verwendeten Abschätzmethoden und ermöglichen es zudem, Fehler bei der Definition der Strukturgruppen und den gespeicherten Gruppenbeiträgen zu finden.

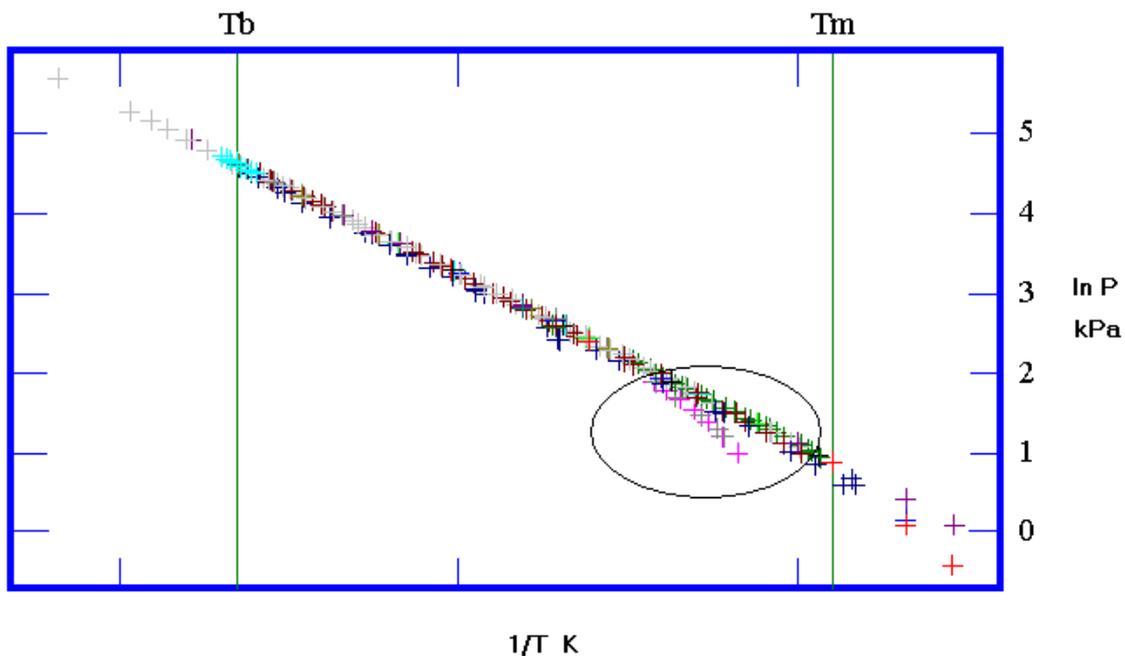
Neben der Überprüfung der Daten wurden auch in größerem Umfang Datenkorrelationen durchgeführt (Oldenburg und Prag). Aufgrund der ständig neu hinzukommenden experimen-

tellen Daten werden diese Korrelationen immer wieder neu überarbeitet. Es sind zur Zeit etwa 8000 Parametersätze für verschiedene Größen und Gleichungen verfügbar.

#### 3.1.4.1 Test und Korrelation der Sättigungsdampfdrücke

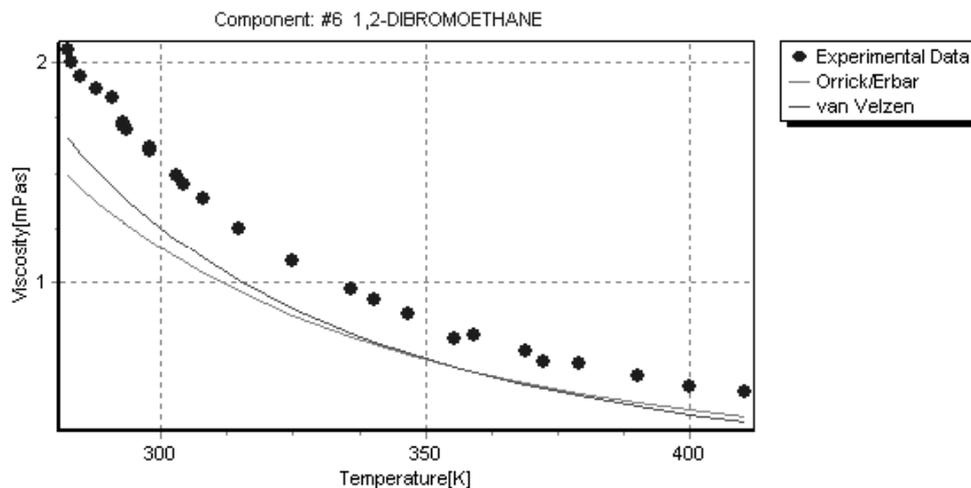
In der Datenbank befinden sich etwa 100000 Datenpunkte mit Sättigungsdampfdrücken in etwa 15000 Datensätzen. Diese sind zu 90% in Oldenburg eingegeben worden. Bei einer solch großen Anzahl von Datenpunkten kommt es gelegentlich zu Fehlern bei der Eingabe oder Verarbeitung der Daten. Ziel muß es somit sein diese Fehler möglichst einfach und effektiv zu entdecken und zu beheben. Jeden Datenpunkt in der Literatur mit den Datenpunkten in der Datenbank zu vergleichen wäre angesichts der großen Datenmenge sicherlich nicht praktikabel. In Oldenburg erfolgt die Datenkontrolle der Sättigungsdampfdrücke mit einem sogenannten „Multiple-Plot“-Programm. Mit diesem Programm können alle Sättigungsdampfdrücke, die sich in der Datenbank befinden, für eine Komponente graphisch dargestellt werden. Aus Geschwindigkeitsgründen werden dabei die Daten für sechs Komponenten in sechs Diagrammen zusammen in einem Programmfenster dargestellt. Zusätzlich zu den Daten können Korrelationen von Antoine, Wagner, DIPPR und Cox mit in die Diagramme eingeblendet werden. Für diese vier Methoden existieren im Oldenburger Programmpaket etwa 6000 Parametersätze, von denen der größte Teil auf Antoine-Parametersätze entfällt. Die Korrelationen bieten weitere Sicherheit bei der Entdeckung von Eingabe- oder Verarbeitungsfehlern. Gleichzeitig werden auch soweit verfügbar der Siedepunkt ( $T_S$ , 1 atm) sowie der kritische Punkt ( $T_{kr}$ ,  $P_{kr}$ ) automatisch eingezeichnet.

Nicht alle fehlerhaften Datenpunkte sind dabei auf falsche Eingabe zurückzuführen, sondern auch auf fehlerhafte Messungen der Autoren. Die Gründe hierfür sind vielfältig (ungenauere Messmethode, verunreinigte Substanzen, mangelndes Know-How) und nicht immer festzustellen. Solche Daten werden nicht aus der Datenbank entfernt, sondern mit einem Qualitätsmerkmal versehen, um dem späteren Nutzer die Auswahl zu erleichtern. Insbesondere bei Sättigungsdampfdrücken, die vor 1900 gemessen wurden, ist Vorsicht geboten. Betont werden muß, daß diese Daten nicht durchweg fehlerhaft sind, aber die Häufigkeit von Abweichungen zu neueren Publikationen doch auffällig ist. Zur Verdeutlichung sind in der folgenden Abbildung sämtliche Sättigungsdampfdrücke (auch Sublimationsdrücke unterhalb der Schmelztemperatur  $T_m$ ) aus der Datenbank für Ameisensäure dargestellt. Es fällt auf, daß die Daten von Kahlbaum und Schröter (Kahlbaum G.W.A., Schröter P., Verhandlungen der Naturforschenden Gesellschaft in Basel, **9**, 573 (1893) , markiert durch den Kreis) aus dem Jahre 1893 stark von den anderen Daten abweichen.

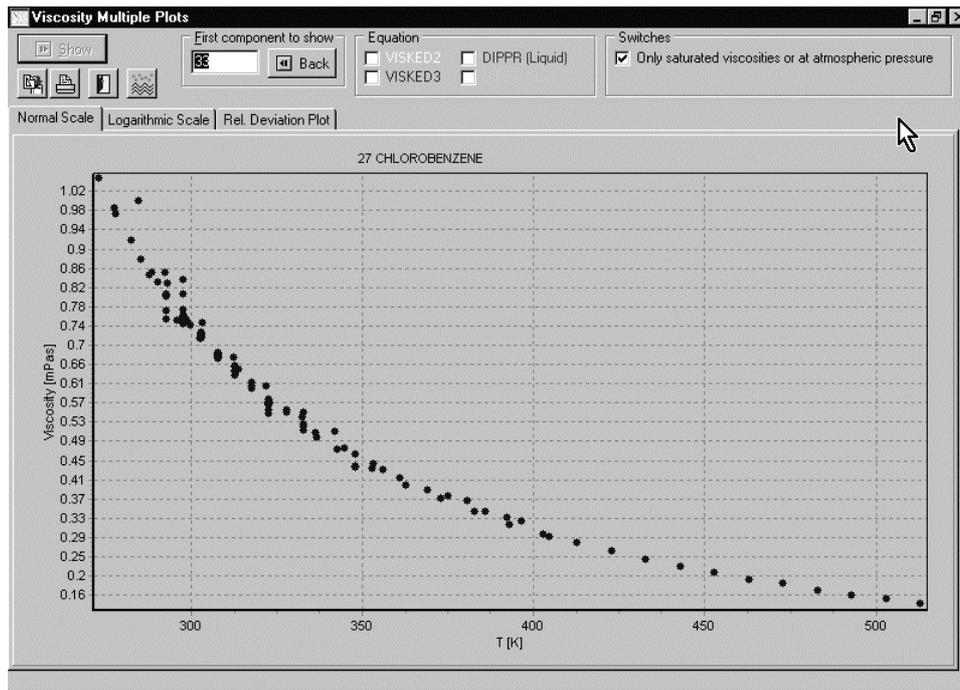


### 3.1.4.2 Test und Korrelation der Viskositätsdaten

Die Datenbasis ist hier sehr groß und besteht zum überwiegenden Teil aus Daten für Flüssigkeitsviskositäten. Für etwa 1600 Komponenten sind drei oder mehr Datenpunkte für die Viskosität der Flüssigkeit verfügbar, so dass für diese Komponenten Parameter angepaßt werden können und Vergleiche mit Vorhersagemethoden mit hoher Aussagekraft möglich sind. In der folgenden Abbildung ist beispielhaft die vorhandene Datenbasis für 1,2-Dibromethan zusammen mit den Vorhersagen nach Orrick/Erbar sowie nach van Velzen/Lopes Carado/Langenkamp dargestellt.



Neben der großen Anzahl an Datenpunkten für die Viskosität von Flüssigkeiten ist für den Test und die Entwicklung von Modellen ebenfalls wichtig, daß die Datenbasis als relativ zuverlässig und konsistent eingestuft werden kann. In der folgenden Abbildung ist als Beispiel hierfür die Datenbasis für die Viskosität von flüssigem Chlorbenzol dargestellt. Die Datenpunkte stammen aus 43 Literaturstellen, die zwischen 1903 und 1998 publiziert wurden. Man erkennt hier, daß Daten über einen großen Temperaturbereich vorliegen, die wenig systematische Abweichungen zueinander zeigen.



### 3.1.4.3 Test und Korrelation der kritischen Daten

Test und Korrelation der kritischen Daten werden in Prag durchgeführt. Für diese Arbeiten wurden von der DDBST GmbH unterstützende Programme entwickelt. So lassen sich beispielsweise kritische Daten für ausgewählte Stoffe oder die gesamte Datenbank abschätzen und in ein Spreadsheet-Programm wie z.B. Excel kopieren. Die folgende Abbildung zeigt das Programmfenster zur Abschätzung des kritischen Drucks:

#DDB	EName	Stoff	PCP	Joback	Ambrose	Lydersen	Gani	Daubert	Somayajul	Klincewicz
1	ACETALDEHYDE	5573		5603	7149	5548	5688	Parameter	6133	5709
2	ACETAMIDE	6600		6009	6657	6604	6600	Group assi	6651	6025
3	ACETONITRILE	4833	4890 4830 4830 4934	4856	4798	4841	Group ass	Group assi	4828	4429
4	ACETONE	4701	4738 4740 4701 4700 4700 47	4803	5045	5008	4880	5034	5040	4569
5	ETHYLENEDIAMINE	6282	6707	5954	6245	6289	6290	5873	6247	5781
6	1,2-DIBROMOETHANE	5336		6181	5857	5914	7344	8891	4442	5091
7	ETHYL BROMIDE	6231	6231	5561	6538	6594	6155	7295	5375	5563
8	1,2-ETHANEDIOL	8200	7698 8200	6653	10071	7528	5490	9522	6824	7256
9	ETHYL IODIDE	5542		4877	5987	5992	5753	Parameter	5332	4879
10	5-ETHYL-2-NONANOL	0		2131	2217	2130	2083	Tc not give	Group ass	2329
11	ETHANOL	6383	6391 6590 6846 6777 6767 61	5757	5399	6400	5557	6486	6054	5669
12	DIETHYL ETHER	3638	3638 3638 3652 3615 3606 36	3815	3766	3788	3778	3638	3763	3752
13	ETHYLENE OXIDE	7194		6339	6552	6511	Group ass	7510	Group ass	6237

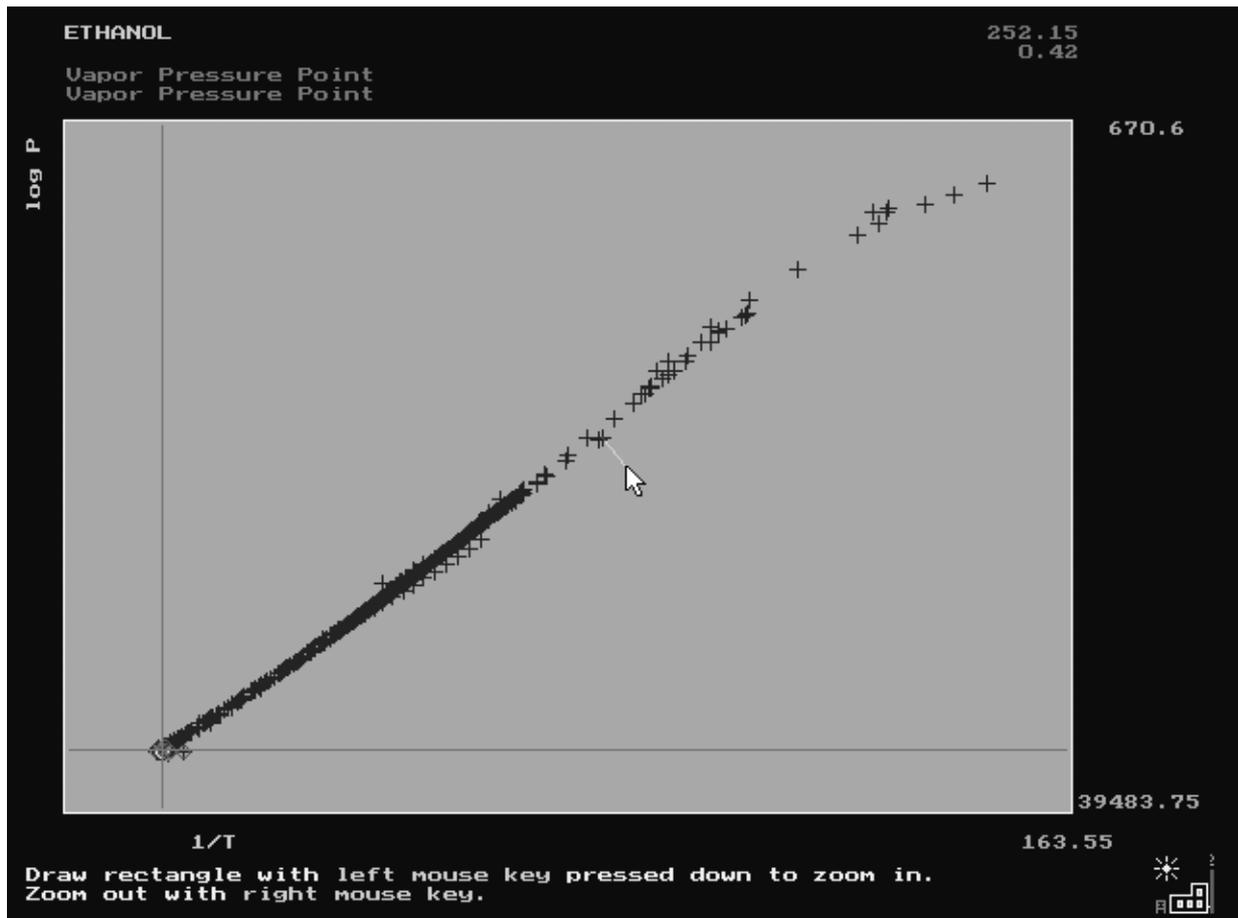
Mit Hilfe des Programms CriSelec lassen sich experimentelle und abgeschätzte Daten gemeinsam bearbeiten und mit Hilfe von Gewichtsfaktoren bewerten:

```

ETHANOL                               Help ? Selection Jan 13 1998
-----
Tc: 490.20 +- 0.00                      Weight: 1
[ES0] Est. <Chein>
Tc: 506.85 +- 0.00                      Weight: 1
[PR9604] (1881) ZHUK' K.,
Tc: 507.75 +- 2.10                      Weight: 1
[PR8607] (1880) HANNAY J.B., HOGARTH J.,
Tc: 508.18 +- 0.00                      Weight: 1          4
[ES0] Est. <Daubert>
Tc: 508.28 +- 0.16                      Weight: 1
[PR7463] (1882) HANNAY J.B.,
Tc: 508.54 +- 0.10                      Weight: 1
[PR7463] (1882) HANNAY J.B.,
Tc: 508.82 +- 0.19                      Weight: 1
[PR7463] (1882) HANNAY J.B.,
-----
Mean Tc: 513.72 Mean Pc: 6298.0 Mean Vc: 166.5 -> zc: 0.2455
Median Tc: 514.05 Median Pc: 6368.6 Median Vc: 167.2 -> zc: 0.2491
a Tc: 0.00 K b Pc: 0.0 kPa c Vc: 0.00 cm3/mol
A ±Tc: 0.00 K B ±Pc: 0.0 kPa C ±Vc: 0.00 cm3/mol
55 items
F10 PCPVIEW Return Accept p Plot against Vapor Pressure # DOS Esc Quit
Delete,- Weight Insert,+ -Weight o List Overview
15:46:53 DDBST GmbH, Industriestr. 1, D-26121 Oldenburg, Germany 15.12.98

```

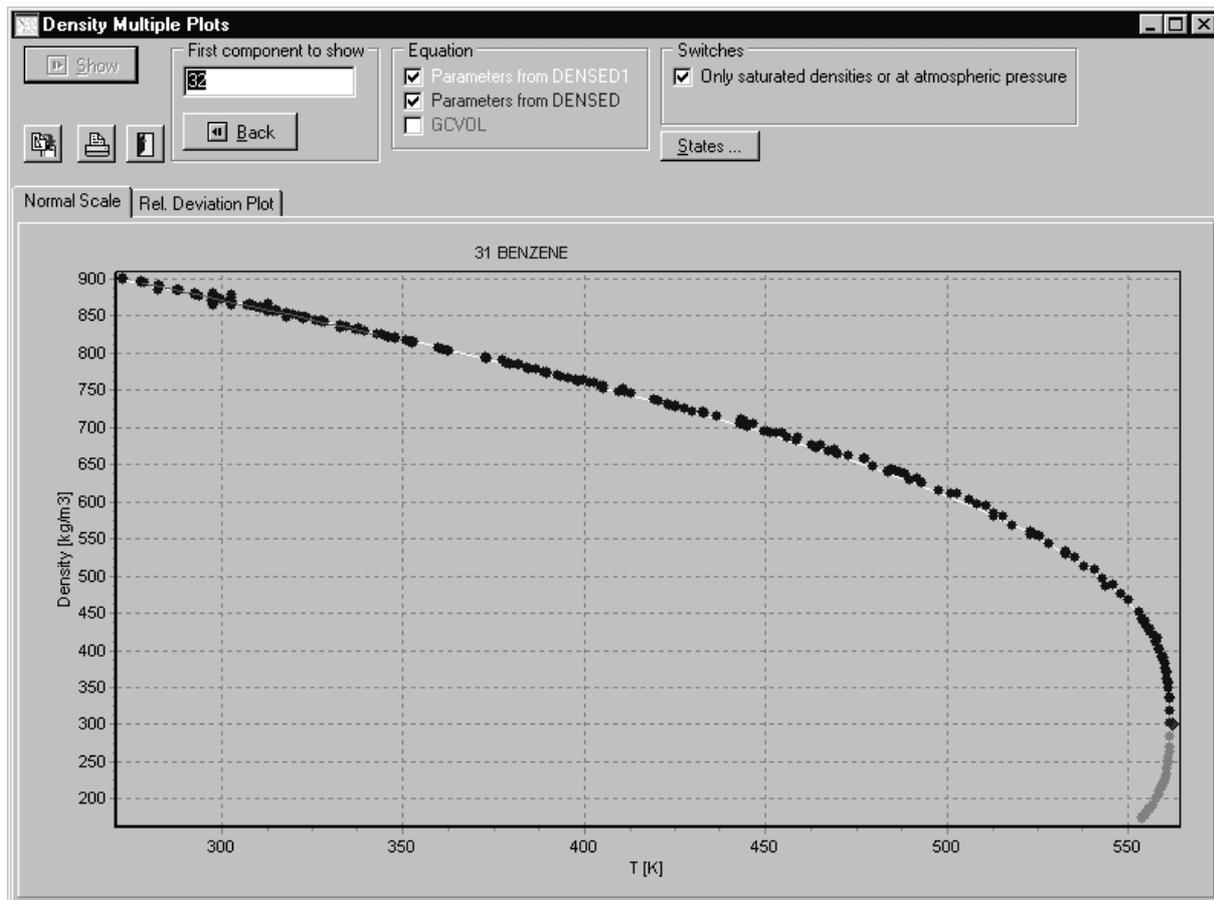
Die ausgewählten Größen  $T_{kr}$  und  $P_{kr}$  lassen sich durch simultane Darstellung mit den gespeicherten Sättigungsdampfdruckdaten weiter überprüfen, da der kritische Punkt den Endpunkt der Dampfdruckkurve darstellt:



#### 3.1.4.4 Test und Korrelation der Flüssigdichten

Für die Stoffgröße Dichte und PvT-Verhalten sind in der Reinstoffdatenbank für über 2650 Komponenten rund 96000 Datenpunkte enthalten. Ca. 15% dieser Daten betreffen Gasdichten im unter- und überkritischen Zustand. Die Dichten wurden überwiegend von den Projektpartnern in Prag, Tallinn und vom FIZ CHEMIE eingegeben.

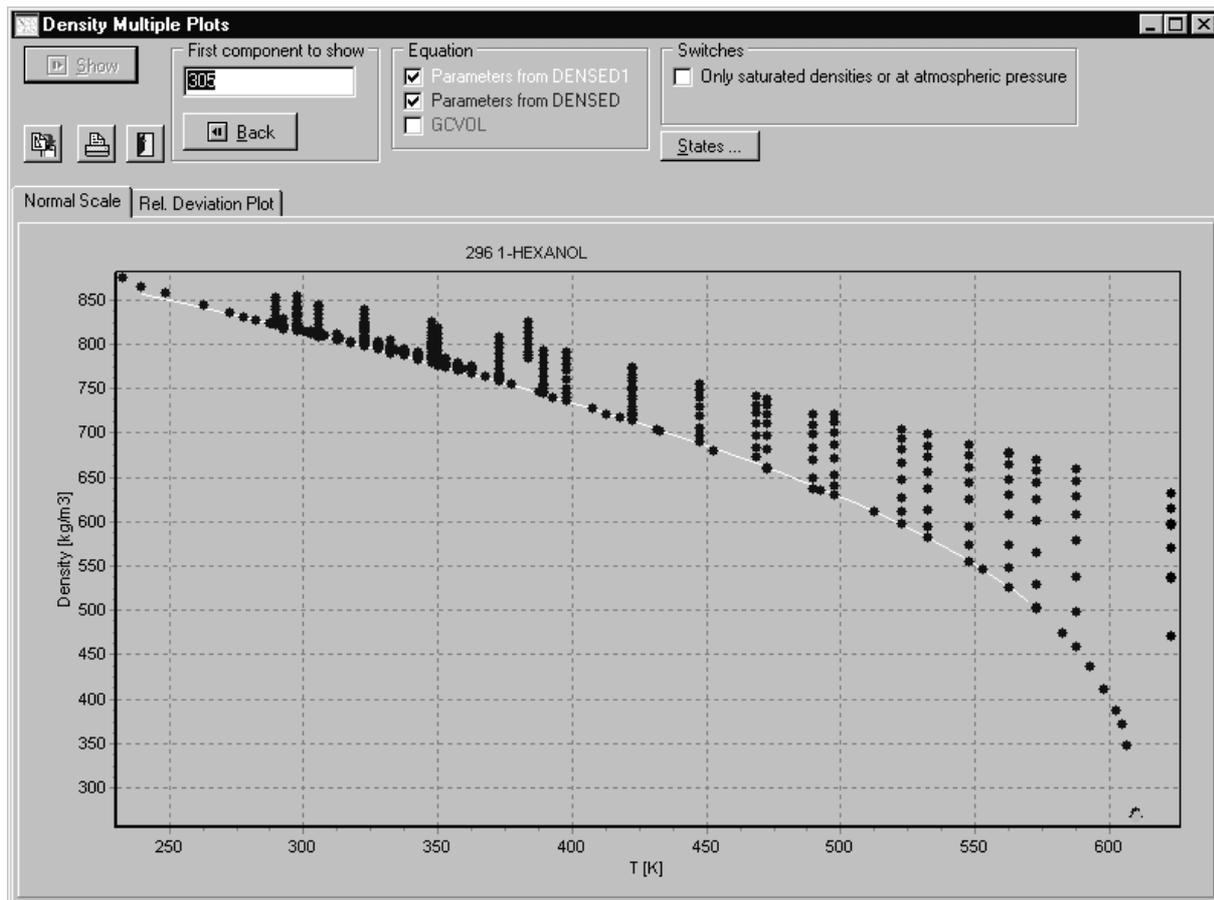
Bei den Flüssigkeitsdichten sind zum größten Teil nur Werte bei Atmosphärendruck oder Sättigungsdampfdruck bzw. leicht erhöhten Drücken enthalten. Dies ist vor allem dadurch begründet, daß von den Prager Projektpartnern laut Vertrag keine Hochdruck-Flüssigkeitsdichten erfaßt werden sollten. Bei den Atmosphärendruck- und Sättigungsdichten ist die Datenbasis schon sehr umfangreich, dies zeigen z.B. die Daten für Benzol in der folgenden Abbildung.



Das hier gezeigte Plotprogramm dient als ein Weg zur Überprüfung der gespeicherten Daten. Neben den experimentellen Werten können hier zum Vergleich auch die Ergebnisse von Korrelationen (DIPPR-Gleichung und Polynom-Gleichung (dieses Projekt)) und die Resultate der in Oldenburg entwickelten Gruppenbeitragsmethode GC-VOL dargestellt werden. Neben simplen Zahlendrehern treten oft auch Fehler bei der Phasenangabe auf, die hier schnell aufgedeckt werden können.

Trotz der schon recht umfangreichen Zahl an Flüssigkeitsdichten weist die Datenbank noch Datenlücken auf. Dazu ist allerdings zu bemerken, daß die von den Prager Projektpartnern beschaffte Literatur zumindest im Bereich der Flüssigkeitsdichten aufgrund des sehr großen Umfangs noch nicht vollständig bearbeitet wurde. Prag wird die Daten aus ca. 2000 Literaturstellen in den nächsten Monaten nachliefern, die weiteren Eingaben werden im privatwirtschaftlich geförderten Folgeprojekt in Oldenburg durchgeführt.

In der folgenden Abbildung sind neben Dichten auf der Sättigungsdampfdruckkurve und bei Normaldruck auch Hochdruck-Flüssigkeitsdichten enthalten. Deren kritische Überprüfung gestaltet sich zur Zeit noch schwieriger. Es wird dazu allerdings in Kürze ein Programm zur dreidimensionalen Darstellung (PpT) von der DDBST GmbH verfügbar sein.



Die in der Datenbank enthaltenen Gasdichten stammen fast ausschließlich aus Tallinn. Vom FIZ CHEMIE wurden vor allem kritische Dichten (mit der Phasenangabe VLE) beigesteuert. Eine größere Zahl an Gasdichten wurde von den Grazer Projektpartnern erfaßt. Diese Daten liegen zwar in Oldenburg vor, sind aber aufgrund der noch ausstehenden endgültigen Absprache nicht in der hier dargestellten Datenbank enthalten.

#### 3.1.4.5 Test und Korrelation der Wärmekapazitäten

Die in der Reinstoffdatenbank enthaltenen Wärmekapazitätsdatenpunkte setzen sich wie folgt zusammen:

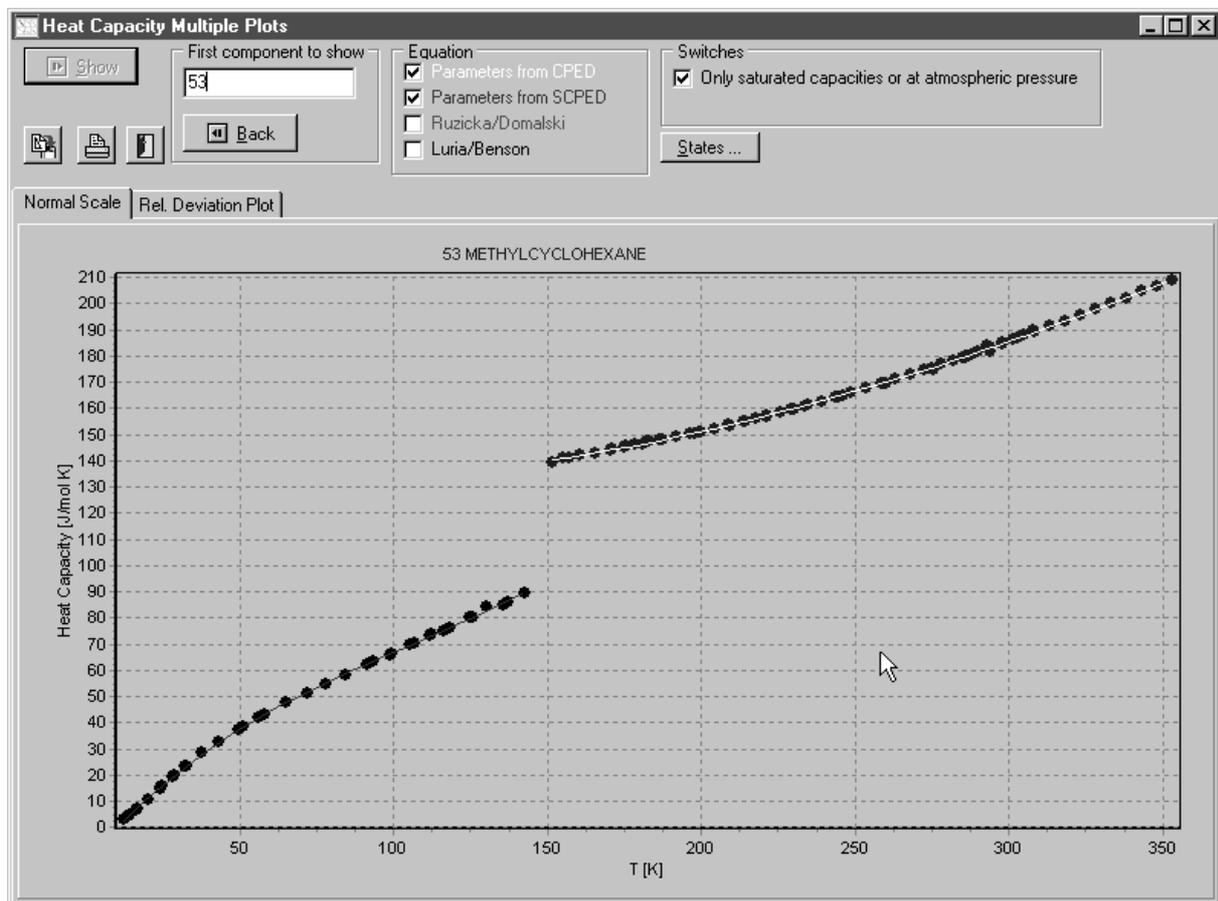
- \* Wärmekapazitäten bei konstantem Druck: 74 %
- \* Idealgaswärmekapazitäten: 20 %
- \* Wärmekapazitäten bei konstantem Volumen: 6 %

Die zahlenmäßig wichtigste Wärmekapazitätsart, die Wärmekapazität bei konstantem Druck, verteilt sich bezüglich ihrer Phase wie folgt:

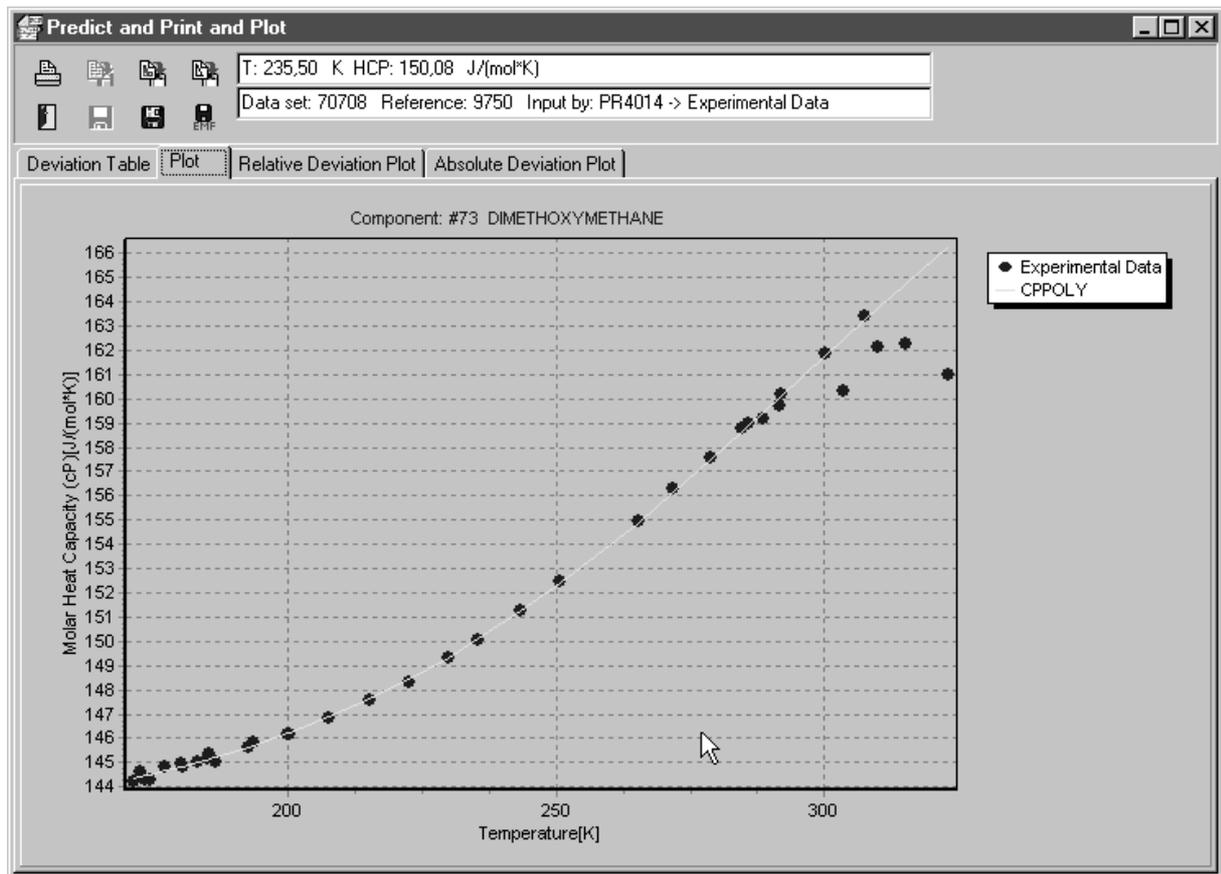
- \* Gasphase: 6 %
- \* Flüssige Phase: 25 %
- \* Feste Phase: 69 %

Bei der festen Phase werden die einzelnen Modifikationen entsprechend der Angaben der Autoren aus der Literatur eingegeben.

Der Test der Daten erfolgt mit Hilfe von graphischen Plot-Programmen. Zusätzlich können mit entsprechenden Programmen Korrelationsparameter an experimentelle Daten angepaßt werden, die dann zusammen mit den experimentellen Daten graphisch angezeigt werden. Für die feste Phase können insgesamt für vier verschiedene Modifikationen Parameter angepaßt werden. In der folgenden Abbildung. sind für Methylcyclohexan die Wärmekapazitätsdaten in der festen und flüssigen Phase zusammen mit den korrelierten Daten angezeigt.



In der nächsten Abbildung sind die Wärmekapazitätsdaten von Dimethoxymethan in der flüssigen Phase dargestellt. Hierbei ist sehr gut zu erkennen, daß mit Hilfe der beschriebenen Vorgehensweise Fehler bzw. experimentelle Daten mit schlechter Qualität erkennbar sind.



### 3.1.4.6 Test und Korrelation der Entropien, Enthalpien, ...

Entropien reiner Stoffe finden sich in der Reinstoffdatenbank als Funktion der Temperatur für den festen, flüssigen oder gasförmigen Zustand. Laut dem 3. Hauptsatz der Thermodynamik ist die Entropie eines idealen Kristalls am absoluten Nullpunkt gleich 0 (Planck). Dadurch erhält man eine absolute Größenskala für die Entropie. Die Entropie bei einer endlichen Temperatur ergibt sich ohne Berücksichtigung von Phasenübergängen über die Gleichung

$$s(T) = \int_{T=0}^T \frac{c_P(T)}{T} dT$$

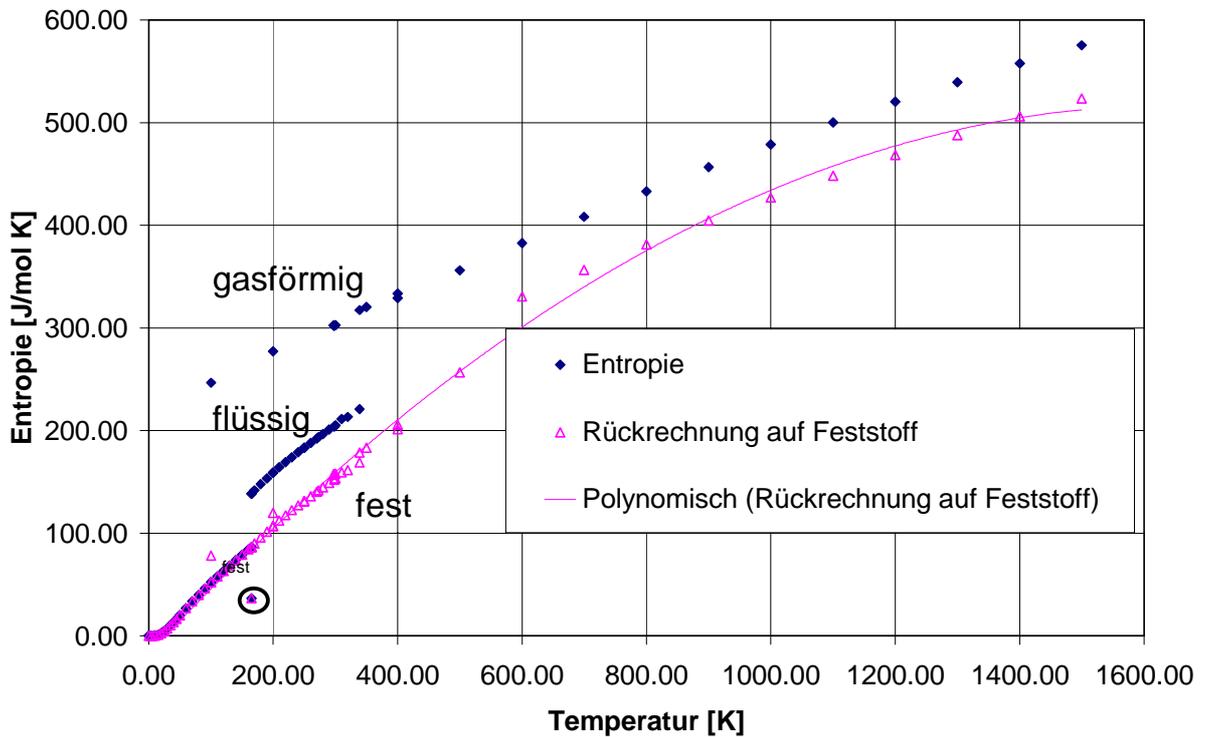
Die Entropie eine Phasenübergangs (U – Umwandlung) ist gleich

$$\Delta s_U = \frac{\Delta h_U}{T_U}$$

Unter Verwendung dieser Zusammenhänge ist es möglich, die Entropien für die verschiedenen Zustände auf die Entropie einer hypothetischen festen Phase rückzurechnen. Bei der Umrechnung vom flüssigen in den festen Zustand benötigt man die Schmelztemperatur und Schmelzenthalpie, die Umrechnung vom Gaszustand in den flüssigen Zustand erfordert die Kenntnis der Verdampfungsenthalpie bei der jeweiligen Temperatur. Diese kann aus der Verdampfungsenthalpie am Normalsiedepunkt mit Hilfe der Watson-Gleichung unter Kenntnis der kritischen Temperatur abgeschätzt werden.

Eine Darstellung der in der Reinstoffdatenbank gespeicherten Entropiedaten für Tetrahydrofuran mit Umrechnung auf den hypothetischen Feststoff zeigt die folgende Abbildung. Mit Hilfe

dieser Vorgehensweise können leicht Fehler lokalisiert werden. Diese treten beispielsweise durch fehlerhafte Eingabe der Einheit, Phase oder des Referenzzustandes auf.



Ein solcher Fehler findet sich in der Abbildung bei einer Temperatur von ca. 170K (eingekreister Datenpunkt). Hier wurde vermutlich die Phase fehlerhaft als Flüssig spezifiziert und es handelt sich um einen Entropiewert für den festen Zustand.

Für Entropien wurden bisher keine Parametersätze mit empfohlenen Parametern angepaßt.

### 3.1.4.7 Test und Korrelation der thermodynamischen Bildungsgrößen

- Standardbildungsenthalpien

Die Überprüfung der Standardbildungsenthalpien für 2413 Komponenten erfolgt in zwei Schritten. In einem ersten Schritt wird überprüft, ob für eine Komponente Daten für die verschiedenen Phasen vorliegen. Wenn dieses der Fall ist, wird die thermodynamische Konsistenz der Daten überprüft, d.h. die Bildungsenthalpie im gasförmigen Zustand muß den größten, im flüssigen Zustand den zweitgrößten Wert und im festen Zustand den kleinsten Wert aufweisen. Zusätzlich kann beim Vorliegen einer Schmelz- bzw. Verdampfungsenthalpie getestet werden, ob die Differenz aus den Standardbildungsenthalpien in den jeweiligen Phasen der vorliegenden Umwandlungsenthalpie entspricht.

In einem zweiten Schritt können die Standardbildungsenthalpiendaten mit aus Gruppenbeitragsmethoden berechneten Daten verglichen werden. Hierfür wurde die Methode von Domalski/Hearing (1993) implementiert, die eine Berechnung der Standardbildungsenthalpien in der gasförmigen, flüssigen und festen Phase für eine sehr große Anzahl von Komponenten erlaubt.

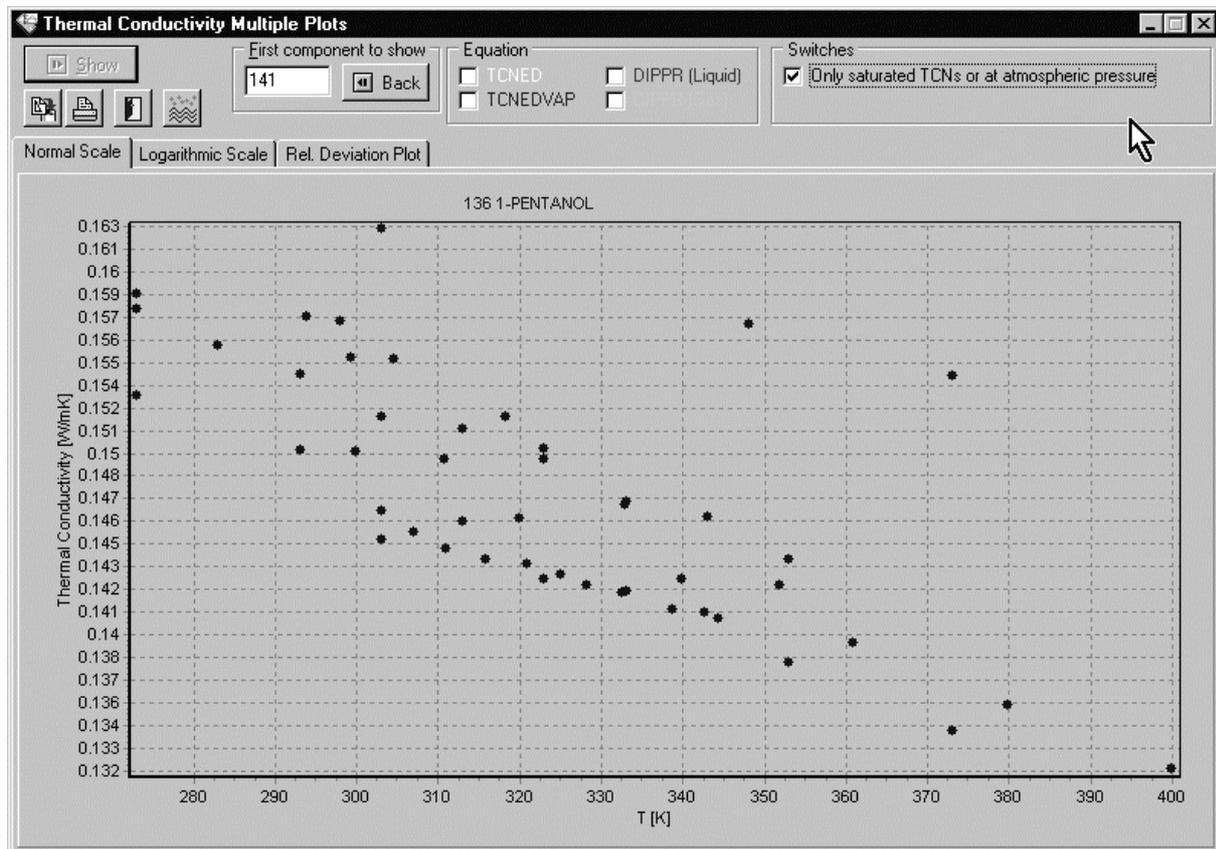
Probleme treten häufig bei schwefelhaltigen Komponenten auf, da hier von den jeweiligen Autoren für die Berechnung der Standardbildungsenthalpie unterschiedliche Schwefelmodifikationen eingesetzt werden ( $S_2(\text{gas})$ ,  $S_8(\text{rhombisch})$ ).

- **Verbrennungsenthalpien**

In der Datenbank befinden sich für knapp 900 Komponenten Verbrennungsenthalpien. Der Test dieser Daten gestaltet sich aufgrund der kleineren Datenbasis etwas schwieriger, da für eine bestimmte Komponente nur selten mehrere Datenpunkte vorliegen. Die Daten können jedoch unter Verwendung des Heßschen Satzes mit den ebenfalls gespeicherten oder abgeschätzten Bildungsenthalpien abgeglichen werden.

### 3.1.4.8 Test und Korrelation der Wärmeleitfähigkeiten

Die Überprüfung und Verifikation der experimentellen Daten für Wärmeleitfähigkeiten gestaltete sich als besonders schwierig. Dies liegt darin begründet, daß die Technik für die experimentelle Bestimmung dieser Größe in den letzten Jahrzehnten erheblich verbessert werden konnte und sich dadurch zeigte, daß viele ältere Daten systematische Fehler von zum Teil erheblicher Größe aufweisen. Daher ist eine sorgfältige Prüfung und Bewertung dieser Daten sehr wichtig. Die folgende Abbildung zeigt einen Screenshot des von der DDBST GmbH entwickelten Programms TCNMultiplePlot, in dem die starke Streuung der Daten für die Flüssigkeitswärmeleitfähigkeit von 1-Pentanol zu sehen ist. Die Daten wurden von verschiedenen Autoren zwischen 1932 und 1993 gemessen und weisen Unterschiede von bis zu 12% auf, was durchaus ein üblicher Wert für die Streuung bei Flüssigkeitswärmeleitfähigkeiten ist.



Bei der Sammlung von Wärmeleitfähigkeitsdaten ist zu erwähnen, daß sich hier die Zusammenarbeit mit den Projektpartnern in Tallinn als besonders nützlich herausgestellt hat, da im russischsprachigen Raum zahlreiche Publikationen mit diesen Daten veröffentlicht wurden.

### 3.1.4.9 Test und Korrelation der Oberflächenspannungen

Oberflächenspannungen von Flüssigkeiten werden schwerpunktmäßig in Prag bearbeitet. Die Datenbank enthält zum Projektende mehr als 17 000 Tabellenzeilen für diese Stoffgröße. Von der DDBST GmbH wurden Programme zur Korrelation, Abschätzung und grafischen Darstellung erstellt. Die folgenden Abbildungen zeigen einige Programmfenster, in denen beispielhaft experimentelle Oberflächenspannungen verschiedener Autoren für Wasser mit den Korrelationen nach DIPPR und der aus dem Reinstoffprojekt (SFTED) sowie der Abschätzung nach Brock, Bird und Miller verglichen werden. Im Falle des Wassers führt die Abschätzung zu ungewöhnlich hohen Abweichungen von bis zu mehr als 50%. Die Darstellung in den Programmfenstern zeigt eine tabellarische Darstellung sowie mehrere grafische Auftragsungen.

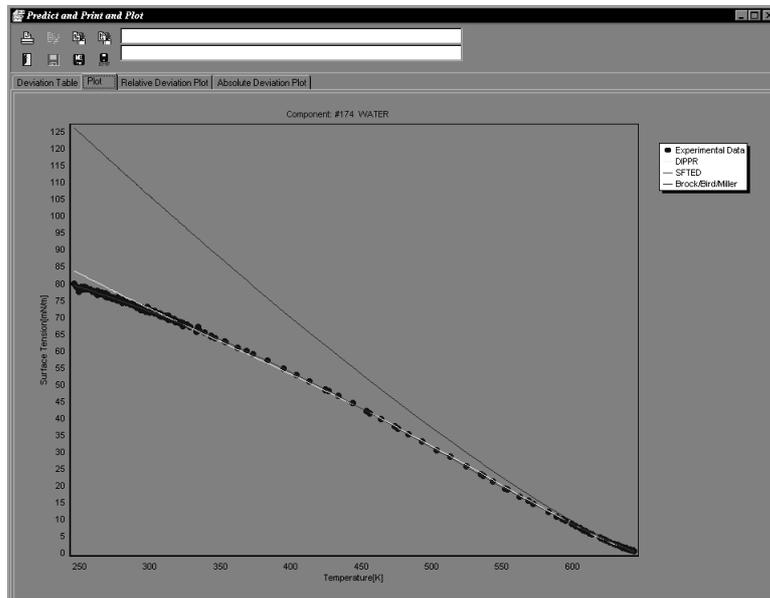
Predict and Print and Plot

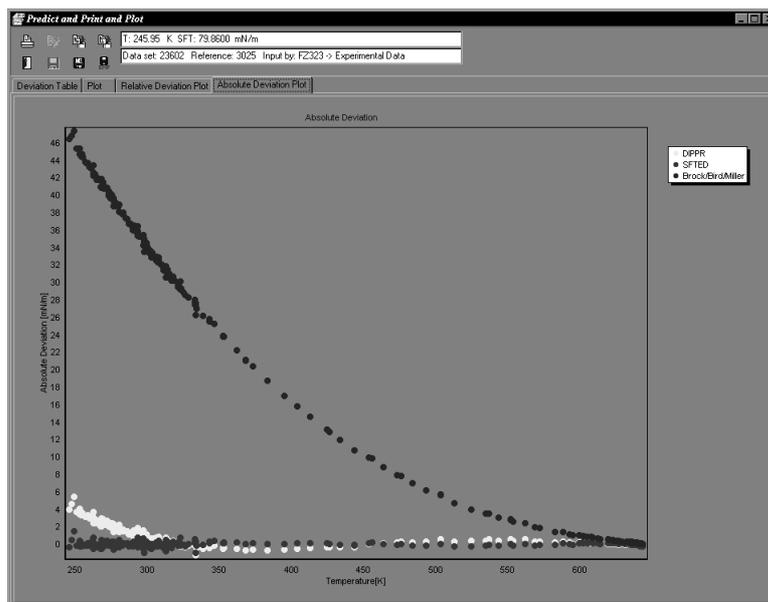
Deviation Table Plot Relative Deviation Plot Absolute Deviation Plot

Estimated property: SURTEN  
Component: WATER  
Calculation/Correlation by DIPPR SFTED Brock/Bird/Miller

VARGATIK H.B., VOLYAK L.D., VOLKOV B.N., TEPLONERGETIKAO(8),80(1973).

T [K]	DIPPR			SFTED		Brock/Bird/Miller	
	Exp: [mN/m]	Calc: [mN/m]	Dev: [%]	Calc: [mN/m]	Dev: [%]	Calc: [mN/m]	Dev: [%]
273.55	75.5900	77.8526	+3.0	75.6593	+0.1	115.7690	+53.2
276.67	75.2400	77.2033	+2.6	75.2053	-0.0	114.5690	+52.3
276.89	75.1400	77.1577	+2.7	75.1731	+0.0	114.5060	+52.4
279.81	74.6600	76.5539	+2.5	74.7445	+0.1	113.4040	+51.9
279.97	74.6100	76.5209	+2.6	74.7209	+0.1	113.3430	+51.9
283.16	74.1800	75.8654	+2.3	74.2486	+0.1	112.1420	+51.2
283.19	74.1800	75.8592	+2.3	74.2482	+0.1	112.1300	+51.2
286.08	73.7600	75.2687	+2.0	73.8127	+0.1	111.0430	+50.5
286.16	73.7300	75.2523	+2.1	73.8007	+0.1	111.0130	+50.6
289.27	73.3300	74.6202	+1.8	73.3324	+0.0	109.8460	+49.8
290.05	73.0600	74.4622	+1.9	73.2143	+0.2	109.5540	+50.0
290.33	73.1710	74.4055	+1.7	73.1718	+0.0	109.4490	+49.6
290.60	73.0600	74.3509	+1.8	73.1309	+0.1	109.3480	+49.7
292.76	72.7800	73.9146	+1.6	72.8020	+0.0	108.5390	+49.1
292.78	72.7500	73.9106	+1.6	72.7989	+0.1	108.5310	+49.2
303.59	71.0700	71.7484	+1.0	71.1221	+0.1	104.5000	+47.0
303.72	71.0600	71.7225	+0.9	71.1016	+0.1	104.4520	+47.0
313.62	69.6800	69.7682	+0.1	69.5185	-0.2	100.7850	+44.6
313.62	69.6400	69.7682	+0.2	69.5185	-0.2	100.7850	+44.7
323.48	67.9000	67.8405	-0.1	67.8955	-0.0	97.1575	+43.1
323.49	67.8900	67.8385	-0.1	67.8938	+0.0	97.1539	+43.1
333.43	66.1800	65.9105	-0.4	66.2110	+0.0	93.5250	+41.3
333.43	66.1700	65.9085	-0.4	66.2093	+0.1	93.5213	+41.3
343.26	64.4300	64.0082	-0.7	64.4946	+0.1	89.9540	+39.6
343.27	64.4000	64.0062	-0.6	64.4929	+0.1	89.9504	+39.7
352.84	62.7100	62.1594	-0.9	62.7764	+0.1	86.5020	+37.9
352.84	62.6500	62.1594	-0.8	62.7764	+0.2	86.5020	+38.1
353.41	62.5400	62.0494	-0.8	62.6727	+0.2	86.2974	+38.0
353.43	62.5200	62.0455	-0.8	62.6691	+0.2	86.2902	+38.0
362.29	60.8700	60.3346	-0.9	61.0356	+0.3	83.1212	+36.6
362.30	60.8600	60.3327	-0.9	61.0337	+0.3	83.1177	+36.6



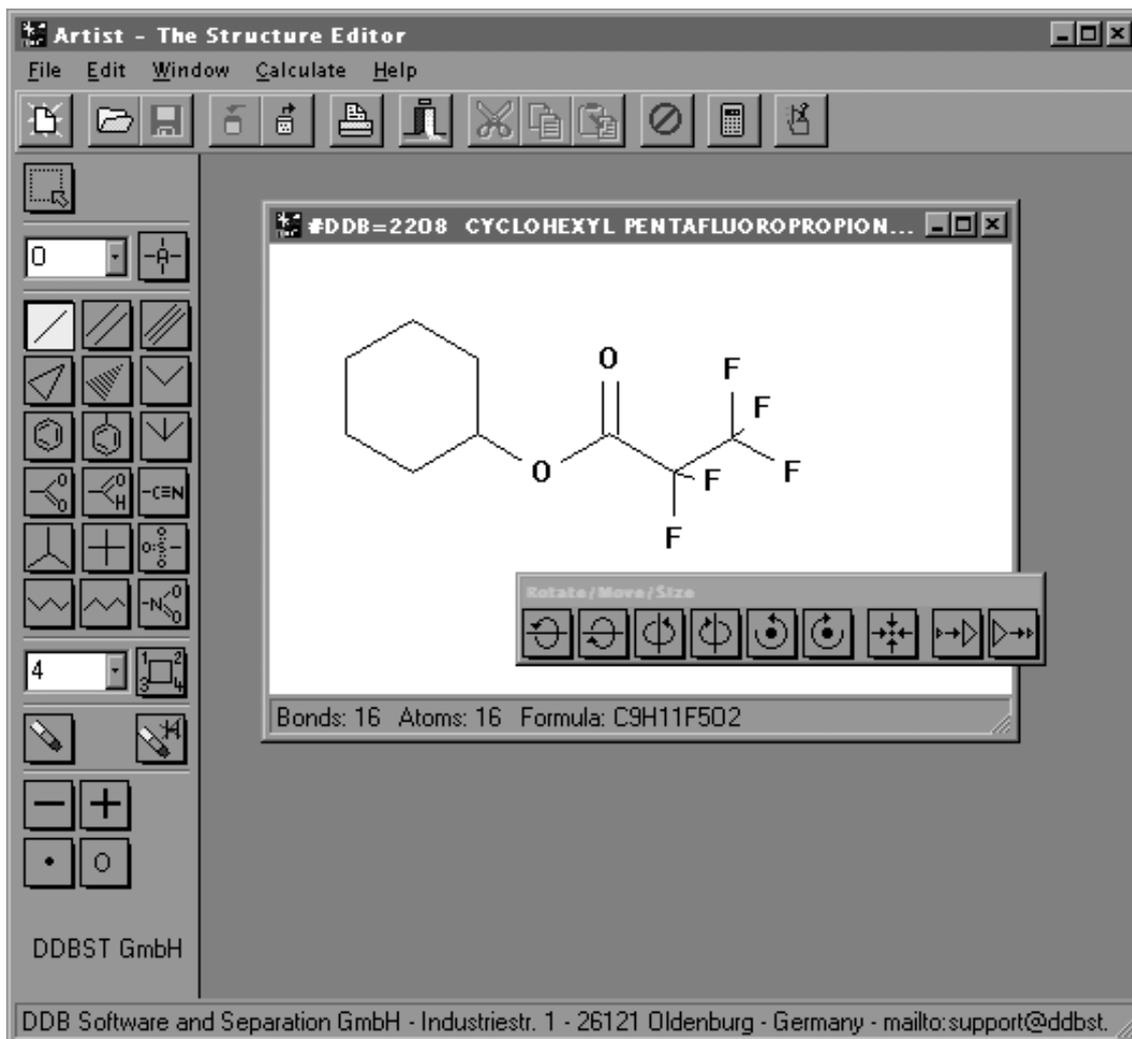


Die hohe Qualität der in diesem Projekt durchgeführten Regression ist sicherlich auf die besonders breite Datenbasis zurückzuführen. Bemerkenswert ist ebenfalls die gute Übereinstimmung der Daten verschiedenster Autoren.

### 3.1.5 Methoden und Programme zur Strukturverarbeitung

Zur Eingabe, Speicherung und Zerlegung von Molekülstrukturen wurden die Programme Artist, ChemDB und AutoInkr entwickelt. Alle Programme wurden zuerst in ANSI-C möglichst ohne die Verwendung spezieller Programmbibliotheken verfaßt und sollten dadurch ohne zu großen Aufwand auf andere Hardwareplattformen portierbar sein (Artist nutzte das Borland Graphic Interface). Aufgrund der weiten Verbreitung der Betriebssysteme Windows-9x und Windows-NT wurden die Programme in geeigneter Weise an diese Plattformen angepaßt. Dabei dient der Grafikeditor gleichzeitig als Benutzerschnittstelle. Komponentenauswahl, Strukturdatenbank und Berechnungsmethoden incl. Strukturzerlegung wurden als COM-Module (COM – Common Object Module) implementiert.

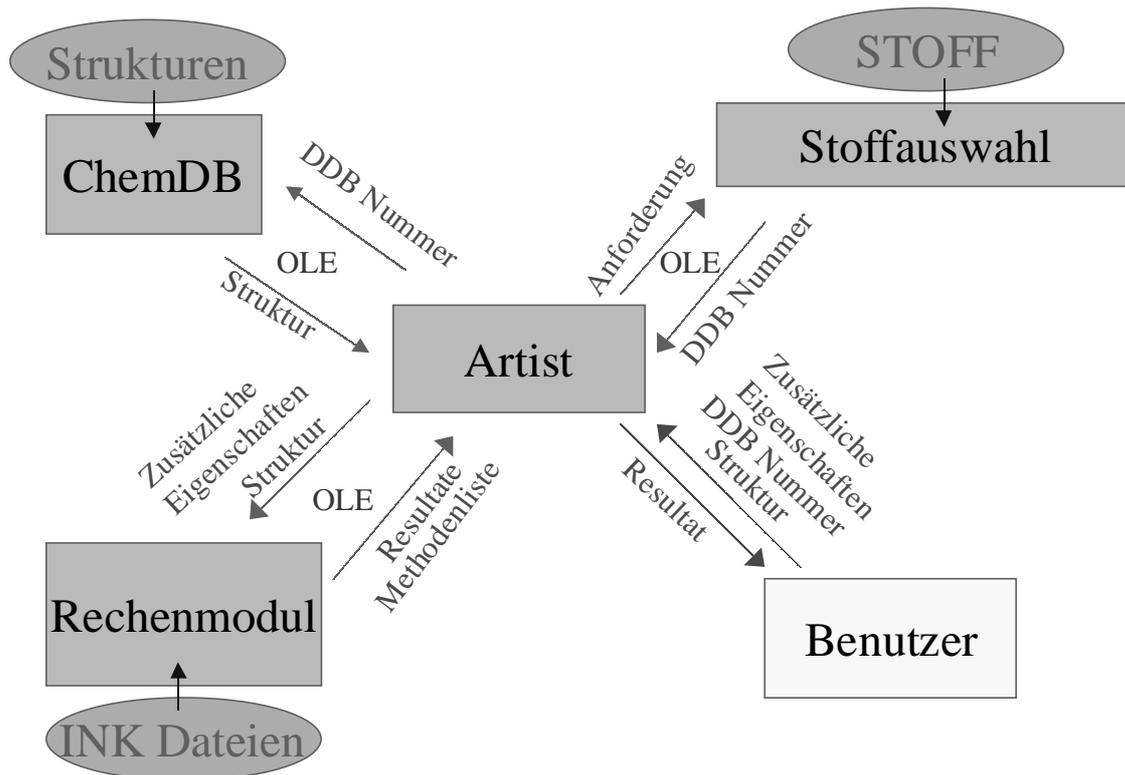
Einen Bildschirmauszug des Grafikeditors Artist zeigt die folgende Abbildung:



Von der grafischen Oberfläche aus können

- Moleküle gezeichnet und als Datei oder in der Datenbank abgelegt werden.
- Moleküle nach allen verfügbaren Regelsystemen in Gruppen zerlegt werden (Es werden die entsprechenden Gruppensystemdateien benötigt).
- Berechnungen mit Hilfe der verschiedenen Gruppenbeitragsmethoden durchgeführt werden.

Die folgende Abbildung erläutert die Struktur des Programmpakets ARTIST.



### 3.1.6 Abschätzung von Reinstoffdaten und Test der Gruppenbeitragsmethoden

Zur Stoffdatenabschätzung wurden eine größere Anzahl von Gruppenbeitragsmethoden implementiert. Eine in etwa vollständige Liste der Methoden ist der folgenden Abbildung zu entnehmen. Diese Liste enthält auch Methoden, deren Implementation noch nicht endgültig getestet ist.

- [-] Critical Temperature
  - ... Chein-Hsiun Tu
  - ... Gani/Constantinou
  - ... Ambrose
  - ... Daubert
  - ... Lydersen
  - ... Joback
  - ... Klinciewicz/Reid
  - ... Somayajulu
- [-] Critical Pressure
  - ... Gani/Constantinou
  - ... Ambrose
  - ... Daubert
  - ... Lydersen
  - ... Joback
  - ... Klinciewicz/Reid
  - ... Somayajulu
- [-] Critical Volume
  - ... Gani/Constantinou
  - ... Ambrose
  - ... Lydersen
  - ... Joback
  - ... Klinciewicz/Reid
  - ... Somayajulu
- [-] Normal Boiling Point
  - ... Cordes/Rarey
  - ... Gani/Constantinou
  - ... Lydersen
  - ... Stein/Brown
  - ... Devotta/Rao
  - ... Joback
- [-] Freezing Point
  - ... Gani/Constantinou
  - ... Joback
- [-] Heat Capacity of Ideal Gases
  - ... Seaton
  - ... Joback
  - ... Benson
- [-] Heat Capacity of Liquids
  - ... Ruzicka/Domalski
  - ... Missenard
- [-] Heat Capacity of Liquids at 25 Deg C
  - ... Luria/Benson
  - ... Shaw
  - ... Chueh/Swanson
- [-] Energy of Vaporization
  - ... Horvath
- [-] Entropy of Vaporization
  - ... Hoshino
  - ... Ma/Zhao
- [-] Gibbs Energy of Formation
  - ... Gani/Constantinou
  - ... Joback
- [-] Heat of Formation
  - ... Gani/Constantinou
  - ... Benson
  - ... Joback
- [-] Heat of Vaporization
  - ... Basarova/Svoboda
  - ... Tu/Liu
- [-] Standard Heat of Vaporization
  - ... Gani/Constantinou
  - ... Basarova/Svoboda
  - ... Tu/Liu
  - ... Horvath
- [-] Liquid Viscosity
  - ... Orrick/Erbar
  - ... van Velzen
  - ... Thomas
  - ... Souders
  - ... Sastri/Rao
- [-] Skubla
  - ... Joback
- [-] Gas Viscosity
  - ... Reichenberg
- [-] Vapor Pressure
  - ... Gomez-Nieto/Thodos
  - ... Riedel
  - ... Miller
- [-] Acentric Factor
  - ... Han/Peng
  - ... Hoshino/Nagahama/Hirata
- [-] Liquid Density
  - ... Tsibanogiannis/Kalospiros/T
  - ... Campbell/Thodos
  - ... GCVDL
  - ... Kolbasov
- [-] Liquid Volume at NBP
  - ... Schroeder
  - ... Tyn/Calus
  - ... Sastri/Mohanty/Rao
  - ... Le Bas
- [-] Van der Waals Volume
  - ... V.d.W. -Covolume by Horvath
  - ... Bondi
- [-] Motoc/Balaban
  - ... Thermal Conductivity
  - ... Nagvekar/Daubert
  - ... Sastri/Rao
  - ... Lakshmi
- [-] Upper Flammability Limit
  - ... High
- [-] Surface Tension
  - ... Brock
- [-] Water Solubility
  - ... Kuehne
- [-] Second Virial Coefficient
  - ... McCann/Danner
- [-] Domalski/Hearing
  - ... Heat of Formation (Gas)
  - ... Heat of Formation (Liquid)
  - ... Heat of Formation (Solid)
  - ... S0 (Gas)
  - ... S0 (Liquid)
  - ... S0 (Solid)
  - ... Cp0 (Gas)
  - ... Cp0 (Liquid)
  - ... Cp0 (Solid)

Die Methoden wurden intensiv zum Test der Datenbank auf grobe Fehler eingesetzt. Leider schwankt auch die Qualität der implementierten Modelle außerordentlich stark.

Bei der Implementation der Modelle ergeben sich häufiger Fragen nach der genauen Spezifikation der Strukturgruppen. Die Autoren der Originalpublikationen verzichteten fast grundsätzlich auf eine exakte Festlegung, in welchen Fällen eine Strukturgruppe anzuwenden ist, und überlassen dies dem chemischen Einfühlungsvermögen des mit der Methode häufig nur wenig vertrauten Anwenders.

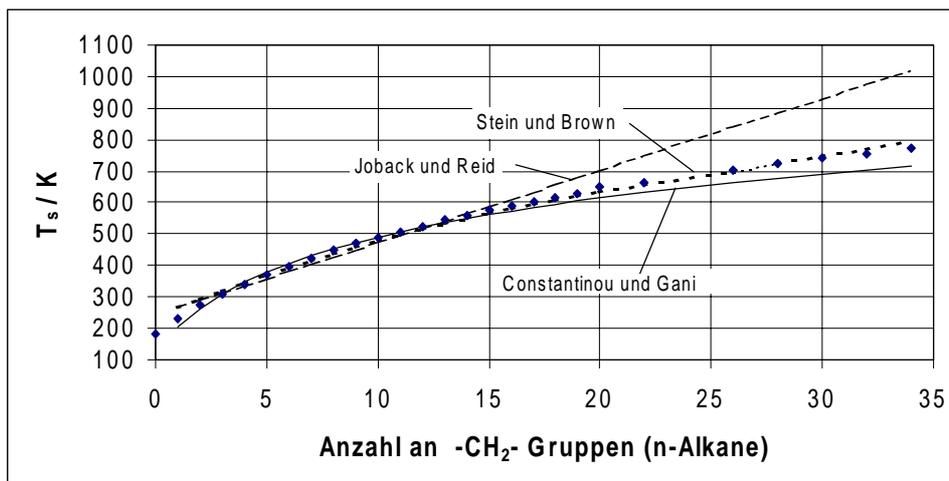
### 3.1.7 Entwicklung eigener Abschätzmethoden

Die Entwicklung eigener Gruppenbeitragsmethoden war nicht primäres Ziel des Projekts. Wie im Antrag beschrieben, sollten die im Projekt erarbeiteten Ergebnisse jedoch geeignet sein, solche Entwicklungen zukünftig deutlich zu erleichtern. Aus diesem Grunde sollen Ergebnisse von Arbeiten in dieser Richtung hier an 2 Beispielen dargestellt werden.

#### 3.1.7.1 Abschätzung von Normalsiedepunkten nach Cordes und Rarey

Sättigungsdampfdrücke gehören zu den technisch wichtigsten Eigenschaften von Reinstoffen. Grundsätzlich sind diese bei Kenntnis des kritischen Punktes mit Hilfe des 2-Parameter Korrespondenzprinzips berechenbar. Allerdings führt diese Methode bei den meisten Stoffen zu großen Abweichungen von den experimentellen Ergebnissen. Aus diesem Grunde wurde das Korrespondenzprinzip durch Einführung des azentrischen Faktors  $\omega$  (Pitzer) zum 3-Parameter Korrespondenzprinzip erweitert. Zur Berechnung des azentrischen Faktors benötigt man den Sättigungsdampfdruck bei  $0.7 T_{kr}$ . Dieser liegt in der Nähe des Normalsiedepunktes, so daß ersatzweise auch die Kenntnis des Normalsiedepunktes ausreicht.

In der Literatur werden verschiedene Methoden zur Abschätzung des Normalsiedepunktes beschrieben. Die Vorausberechnung der Normalsiedepunkte mit Hilfe der drei vermutlich wichtigsten Methoden (Joback und Reid, Stein und Brown sowie Constantinou und Gani) sind in der folgenden Abbildungen den experimentellen Daten gegenübergestellt.



Es fällt auf, daß der von Joback und Reid postulierte lineare Zusammenhang zwischen der Summe der Inkremente und der Normalsiedetemperatur

$$T_s = A + \sum N_i \cdot C_i$$

nur über einen kleinen Bereich der Molekülgröße annähernd gültig ist. Stein und Brown nutzten ebenfalls diesen linearen Zusammenhang, korrigierten später aber die erhaltenen Abweichungen durch ein empirisches Polynom:

$$T_s = f_{\text{corr}}(T_s)$$

Durch die nachträgliche Korrektur ergeben sich allerdings numerische Probleme bei der Anpassung der Gruppenbeiträge, so dass diese Methode nicht das erreichbare Optimum darstellt.

Gani und Constantinou verwendeten einen logarithmischen Ansatz sowie Korrekturen zweiter Ordnung:

$$T_s = 204.359 \ln (\sum N_i \cdot C_i + \sum M_i \cdot D_i)$$

- mit  $N_i$  - Anzahl der Gruppen 1. Ordnung Typ i  
 $C_i$  - Gruppenbeitrag der Gruppe 1. Ordnung i  
 $M_i$  - Anzahl der Gruppen 2. Ordnung Typ i  
 $D_i$  - Gruppenbeitrag der Gruppe 2. Ordnung i

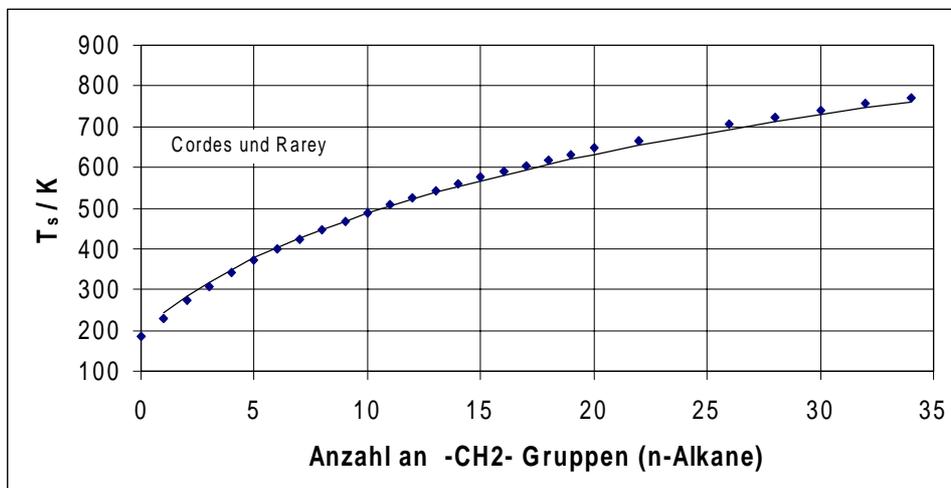
Diese Gleichung ist deutlich besser geeignet als der lineare Ansatz von Joback und Reid, allerdings standen hier zur Anpassung der Gruppenbeiträge nur Daten für 392 Komponenten zur Verfügung (Joback und Reid: 438 Komponenten, Stein und Brown: 11010 Komponenten aus dem Aldrich-Katalog und dem CRC-Handbook).

Auf der Basis der im Projekt aufgebauten Reinstoffdatenbank entwickelten Cordes und Rarey eine Abschätzmethode, die auf einem anderen funktionellen Zusammenhang basiert:

$$T_s = (a + \sum N_i \cdot C_i) / (n^b + c) + d$$

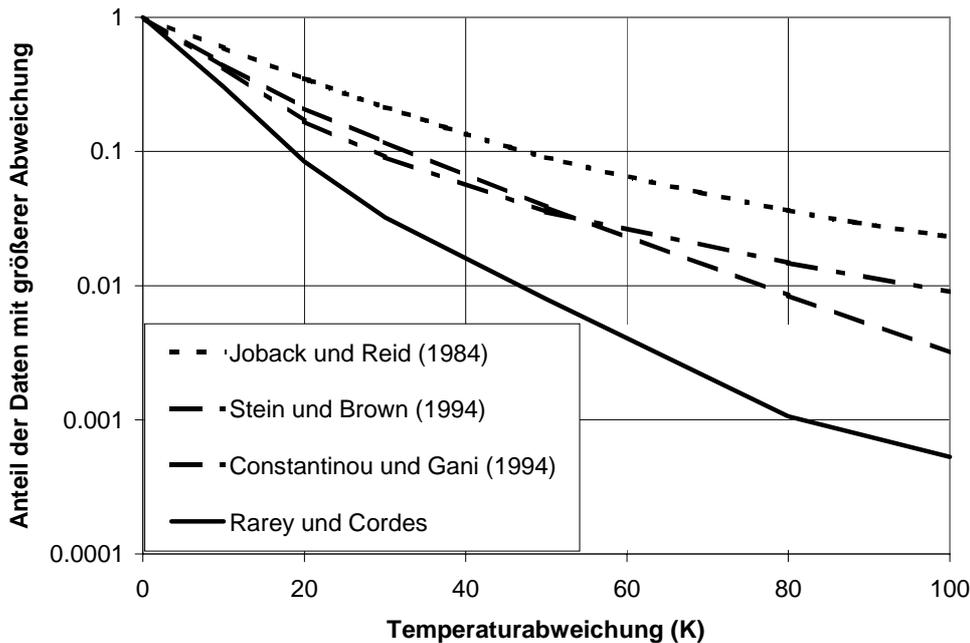
- mit a,b,c,d - anpaßbare Parameter  
 $N_i$  - Anzahl der Gruppen 1. Ordnung Typ i  
 $C_i$  - Gruppenbeitrag der Gruppe 1. Ordnung i  
n - Anzahl der Atome im Molekül (außer H)

Die Resultate dieser Methode für n-Akane zeigt die folgende Abbildung. Es ist erkennbar, daß die gewählte funktionelle Form gut zur Beschreibung des Verlaufs der Siedepunkte in der homologen Reihe geeignet ist.



Ermöglicht wurde die Anpassung der Gruppenbeiträge durch ein besonderes Verfahren, bei dem lineare und nichtlineare Regression in einem Programm in geeigneter Weise kombiniert werden.

Die Abschätzung von Normalsiedepunkten ist für die üblichen Substanzen nur selten erforderlich, da hier Daten im großen Umfang vorliegen. Technisch wichtig wird die Abschätzung bei eher exotischen Stoffen wie schwer zu isolierenden Neben- oder Zwischenprodukten. Hier ist man nicht auf eine hochgenaue Abschätzung angewiesen, die Qualität einer Abschätzmethode muß hier eher durch die Sicherheit vor großen Fehlern beurteilt werden. Dafür ist die folgende grafische Darstellung sehr aussagekräftig:

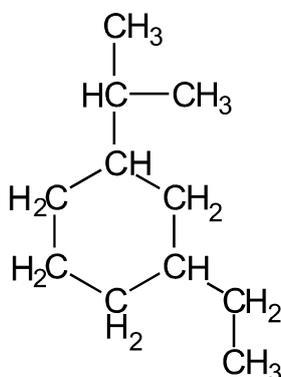


Es findet sich darin der Anteil der Komponenten gegen die Temperaturabweichung aufgetragen, für den eine größere Abweichung erhalten wird. Es läßt sich daraus ablesen, daß bei der Methode von Cordes und Rarey bei 0.1% der Komponenten mit einer Abweichung von mehr als 80K zu rechnen ist. Bei den Methoden von Gani und Constantinou bzw. Stein und Brown liegt dieser Anteil bei etwa 1%, bei Joback bei etwa 5%. Der Grafik liegen Daten und Abschätzungen für etwa 2500 Komponenten zugrunde.

### 3.1.7.2 Abschätzung von Flüssigkeitswärmekapazitäten nach Becker und Aufderhaar

Es existieren verschiedene Gruppenbeitragsmethoden für die Vorhersage von Flüssigkeitswärmekapazitäten [1, 2, 3, 4, 5, 7, 8, 10]. Diese Methoden erlauben die Vorhersage von Wärmekapazitäten nur bei einer einzigen Temperatur oder bei wenigen diskreten Temperaturen. Neben Luria und Benson [6] haben Ruzicka und Domalski [9] eine Gruppenbeitragsmethode entwickelt, mit deren Hilfe Wärmekapazitäten bei beliebigen Temperaturen berechnet werden können. Bei diesen Methoden handelt es sich um Gruppenbeitragsmethoden 2. Ordnung, da Inkremente 2. Ordnung verwendet werden, allerdings werden Korrekturen für Abweichungen der Wärmekapazität, die sich aus Ringspannungen z.B. bei Cyclohexan ergeben, mit Hilfe von Inkrementen 3. Ordnung durchgeführt. Wird eine bestimmte Anzahl von Substanzen betrachtet, so ist die mögliche Anzahl der Inkremente 3. Ordnung, da diese gesamte Molekülstrukturen beschreiben, gleich der Anzahl der Moleküle der Substanzen. Das Ziel einer Gruppenbeitragsmethode ist jedoch, mit einer möglichst geringen Anzahl von Inkrementen eine möglichst große Anzahl von Molekülen zu beschreiben. Dies kann erreicht werden, wenn auf Korrekturen für

Ringspannungen mit Inkrementen 3. Ordnung verzichtet wird und statt dessen Inkremente 2. Ordnung eingeführt werden, die speziell für Ringe gelten. Zur Veranschaulichung ist in der folgenden Abbildung Isopropyl-3-Ethyl-Cyclohexan zusammen mit den Inkrementen, die sich bei einer Vorhersage mit Inkrementen 3. Ordnung bzw. Inkrementen ausschließlich 2. Ordnung ergeben, dargestellt.



Vorhersage 3. Ordnung

Vorhersage 2. Ordnung

3 \* C-(H)<sub>3</sub>(C)

3 \* C<sub>k</sub>-(H)<sub>3</sub>(C<sub>k</sub>)

5 \* C-(H)<sub>2</sub>(C)<sub>2</sub>

1 \* C<sub>k</sub>-(H)<sub>2</sub>(C<sub>k</sub>)<sub>2</sub>

3 \* C-(H)(C)<sub>3</sub>

1 \* C<sub>k</sub>-(H)<sub>k</sub>(C<sub>r</sub>)(C<sub>k</sub>)<sub>2</sub>

Korrekturinkrement:

4 \* C<sub>r</sub>-(H)<sub>2</sub>(C<sub>r</sub>)<sub>2</sub>

Cyclohexan korr.

2 \* C<sub>r</sub>-(H)(C<sub>k</sub>)(C<sub>r</sub>)<sub>2</sub>

Die Indizes k bzw. r bei der Vorhersage 2. Ordnung bedeuten, daß sich das Inkrement in einer Kette bzw. in einem Ring befindet. Um Mißverständnissen vorzubeugen muß betont werden, daß nur das Korrekturinkrement bei der Vorsage 3. Ordnung ein Inkrement 3. Ordnung ist, die anderen Inkremente sind 2. Ordnung. Um die beiden Methoden unterscheiden zu können wird aber im folgendem die Bezeichnung Vorhersage 3. Ordnung bzw. Vorhersage 2. Ordnung beibehalten.

Es fällt auf, daß bei der Vorhersage 2. Ordnung ein Inkrement mehr zur Beschreibung des Moleküls benötigt wird, als bei der Vorhersage 3. Ordnung. Allerdings können die mit dem Index r versehenen Inkremente bei der Vorhersage 2. Ordnung auch für die Beschreibung von anderen Ringen, wie z.B. Cyclopropan, Cyclopentan und Cycloheptan verwendet werden. Bei der Vorhersage 3. Ordnung müßte jedesmal ein anderes Korrekturinkrement verwendet werden.

Es wurde untersucht, wie sich die Vorhersagegenauigkeit einer Gruppenbeitragsmethode ändert, wenn statt einer Vorhersage 3. Ordnung eine Vorhersage 2. Ordnung benutzt wird. Zu diesem Zweck wurde im Rahmen einer Diplomarbeit eine Gruppenbeitragsmethode entwickelt, mit der sowohl eine Vorsage 3. Ordnung, als auch 2. Ordnung möglich ist. Diese Gruppenbeitragsmethode wurde in Anlehnung an die Methode von Luria und Benson [6] entwickelt. So wurden die Inkremente von Luria und Benson mit leichten Modifikationen übernommen. Das

Ziel der entwickelten Gruppenbeitragsmethode war allerdings nicht die Ergebnisse von Luria und Benson zu reproduzieren oder gar zu verbessern, sondern zu zeigen, wie sich die Vorhersagegenauigkeit einer Gruppenbeitragsmethode ändert, wenn auf Inkremente 3. Ordnung verzichtet wird.

Grundlage für die Überprüfung und Entwicklung der Gruppenbeitragsmethode waren in der Reinstoffdatenbank gespeicherte Flüssigkeitswärmekapazitätsdaten von 97 Kohlenwasserstoffen. Die Daten stammen aus der Reinstoffdatenbank der DDB. Zur Vereinfachung der Programme, die für die Gruppenbeitragsmethode benötigt werden, wurde nicht direkt mit den experimentellen Daten gearbeitet, sondern an die experimentellen Flüssigkeitswärmekapazitätsdaten wurde für jede Substanz ein Polynom maximal vierten Grades angepaßt:

$$c_p = a_0 + a_1T + a_2T^2 + a_3T^3 + a_4T^4$$

Mit Hilfe dieser Polynome wurden die für die Entwicklung der Gruppenbeitragsmethode benötigten Flüssigkeitswärmekapazitäten bei verschiedenen Temperaturen für die jeweilige Substanz berechnet.

Die Anpassung für die Vorsage 3. Ordnung ergab 29 Parametersätze (ein Parametersatz besteht aus drei Parametern) für 29 Inkremente und 97 Substanzen. Für die Anpassung 2. Ordnung wurden 25 Parametersätze (ein Parametersatz besteht aus drei Parametern) für 25 Inkremente und 97 Substanzen erhalten. Bei der Vorhersage 2. Ordnung beträgt die durchschnittliche prozentuale Abweichung über alle Datenpunkte 1.72%. Bei der Vorhersage 3. Ordnung ist dieser Wert mit 2.24% etwas höher.

Die Vorhersage 2. Ordnung ist zwar über alle Datenpunkte betrachtet besser, jedoch gibt es für die jeweilige Substanz z.T. große Unterschiede in der Vorhersagegenauigkeit beider Methoden. Die Qualität der Vorhersage ist abhängig von der betrachteten Substanz. Für die acyclischen Alkane und Aromaten sollte es keine Unterschiede zwischen den Vorhersagegenauigkeiten der Vorhersage 3. Ordnung und 2. Ordnung geben, da bei diesen Methoden, auch bei der Vorhersage 3. Ordnung, keine Korrektur durchgeführt wird. In der Tat gibt es keine wesentlichen Unterschiede in der Vorhersagegenauigkeit beider Methoden, mit Ausnahme von Propan und 1,2'-Binaphthylmethan. Die Unterschiede sind in der größeren Anzahl der Inkremente zu suchen, an die bei der Vorhersage 2. Ordnung angepaßt wird. Bei der Vorhersage 2. Ordnung werden Parameter für 25 Inkremente, bei der Vorhersage 3. Ordnung hingegen nur für 17 Inkremente. Die Parameter für die einzelnen Inkremente sind nicht isoliert, sondern sie beeinflussen sich gegenseitig, aus diesem Grund kommt es zu den Unterschieden zwischen den Vorhersagegenauigkeiten der Vorhersage 3. Ordnung und 2. Ordnung bei den Alkanen und Aromaten. Für die cyclischen Verbindungen wird bei der Vorhersage 3. Ordnung eine Korrektur mit Hilfe von Korrekturinkrementen berücksichtigt. Bei diesen Substanzen sind die Differenzen der Vorhersagegenauigkeiten zwischen der Vorhersage 3. Ordnung und 2. Ordnung größer als bei den Alkanen und Aromaten, wobei keine Methode grundsätzlich besser ist. Es fällt auf, daß bei kleinen unsubstituierten Ringen, z.B. Cyclopropan, Cyclobutan und Cyclopentan, die Abweichungen der Vorhersage 3. Ordnung sehr viel kleiner sind als bei der Vorhersage 2. Ordnung. Auf der anderen Seite sind die Abweichungen der Wärmekapazitäten bei 1-Ethyldecalin und 1-Propyldecalin mit der Vorhersage 3. Ordnung größer, als bei der Vorhersage 2. Ordnung. Besonders extrem ist dies bei Spiropentan, obwohl es ein spezielles Korrekturinkrement für diese Komponente gibt.

Die im Rahmen dieser Arbeit entwickelte Vorhersage 2. Ordnung, bei der auf Korrekturinkremente verzichtet wird, stellt eine gute Alternative zur bisher verwendeten Vorhersage 3. Ordnung dar. Wie hier gezeigt wurde, kommt eine Methode 2. Ordnung letztendlich mit weniger Inkrementen aus, als eine Methode 3. Ordnung, dies erhöht die Übersichtlichkeit bei der Ent-

wicklung und Anwendung einer Gruppenbeitragsmethode. Es muß bei der Vorhersage 2. Ordnung allerdings berücksichtigt werden, daß die Vorhersagequalität bei kleinen unsubstituierten Ringen nicht sehr gut ist. Dieser Umstand ist aber nicht entscheidend, denn eine Gruppenbeitragsmethode ist in erster Linie für die Vorhersage der Wärmekapazitäten von Substanzen, für die keine experimentellen Werte vorhanden sind, konzipiert. Für kleine unsubstituierte Ringe, wie z.B. Cyclopropan, Cyclobutan und Cyclopentan, sind jedoch experimentelle Wärmekapazitätsdaten vorhanden, eine Abschätzung mit Hilfe von Gruppenbeitragsmethoden ist für diese Komponenten nicht erforderlich.

- [1] Chu J.Y., Nguyen T.T., King K.D., J. Phys. Chem., **86**, 443 (1982)
- [2] Chueh C.F., Swanson A.C., Can. J. Chem. Eng., **51**, 596 (1973)
- [3] Chueh C.F., Swanson A.C., Chem. Eng. Progress, **69**(7), 83 (1973)
- [4] Domalski E.S., Hearing E.D., J. Phys. Chem. Ref. Data, **17**, 1637 (1988)
- [5] Johnson A.I., Huang C.-J., Can. J., Technol., **33**, 421 (1955)
- [6] Luria M., Benson S.W., J. Chem. Eng. Data, **22**, 90 (1977)
- [7] Missenard F.A., C. R. Acad. Sc. Paris, **260** (24 mai 1965), groupe 7, 5521 (1965)
- [8] Ogiwara K., Arai Y., Saito S., J. Chem. Eng. Japan, **14**, 156 (1981)
- [9] Ruzicka V., Domalski E.S., J. Phys. Chem. Ref. Data, **22**(3), 607 (1993)
- [10] Shaw R., J. Chem. Eng. Data, **14**, 461 (1969)

### 3.2 Nutzen der Datenbank und der erstellten Programme

Mit der Datenbank steht ein Handwerkszeug zur Verfügung, mit dem sich viele Fragestellungen von praktischer Bedeutung bearbeiten lassen. In allen Bereichen des Chemieingenieurwesens ist für die Simulation der einzelnen Grundoperationen oder ganzer Anlagenteile neben genauen Informationen über das Gemischverhalten auch die Kenntnis der Eigenschaften reiner Stoffe erforderlich.

Für die Berechnung von Dampf-Flüssig Gleichgewichten sind die Sättigungsdampfdrücke der beteiligten Komponenten von vorrangiger Bedeutung. Verdampfungsenthalpien geben Auskunft über die Temperaturabhängigkeit der Sättigungsdampfdrücke und sind von großer Bedeutung bei der Berechnung von Rektifikationskolonnen, Wärmetauschern, Siedekühlung von Reaktoren, ...

Flüssigdichten und Viskositäten sind entscheidend für den Energiebedarf beim Transport von Flüssigkeiten durch Rohre, Mischen und Rühren, ... Viskositäten und Oberflächenspannungen entscheiden oft über die Wirkungsgrade von Packungskolonnen, thermodynamische Standardgrößen, speziell Bildungsgrößen ermöglichen die Berechnung der Gleichgewichtslage chemischer Reaktionen.

Durch die Verfügbarkeit von Daten für eine Vielzahl von Komponenten lassen sich weitaus verbesserte Abschätzmethoden entwickeln, mit deren Hilfe dann auch Daten von Verbindungen berechnet werden können, für die keine experimentellen Daten vorliegen.

Auf der Basis solcher Abschätzmethoden ist es dann beispielsweise möglich, Molekülstrukturen mit optimalen Eigenschaften, z.B. als Selektivlösungsmittel für die Lösung von Trennproblemen, zu entwerfen (molecular design).

Durch die Entwicklung der automatischen Strukturzerlegung zusammen mit einer umfangreichen Strukturdatenbank und einer grafischen Eingabeoberfläche sind diese Abschätzmethoden auch für den Nicht-Spezialisten leicht zugänglich.

### 3.3 Fortschritt bei anderen Stellen

Es ist nicht bekannt, daß an anderen als den in Kapitel 2.4 genannten und schon zu Projektbeginn bekannten Stellen am Aufbau einer vergleichbaren Datenbank gearbeitet wird. Der Fortschritt des DIPPR-Projekts läßt sich unserer Kenntnis nach bei weitem nicht mit der Qualität und Quantität des Dateninputs und der Programmentwicklung im hier durchgeführten Projekt vergleichen.

Arbeiten zur Erfassung von Stoffdaten werden auch vom Beilstein-Institut seit langem kontinuierlich betrieben. Versuchsweise wurde im Rahmen dieses Projekts auch Literaturverweisen aus dieser Datenbank nachgegangen. Es hat sich dabei die schon zu Beginn des Projekts gefaßte Meinung bestätigt, daß die dort gespeicherten Daten und Literaturhinweise für chemisch-technische Anwendungen nur von begrenztem Wert sind. So werden dort oft zitierte oder abgeschätzte Werte erfaßt, Fehler in der Literatur oder bei der Eingabe der Daten werden nicht durch Überprüfung z.B. auf thermodynamische Konsistenz korrigiert. Aufgrund der mehr organisch-chemischen Ausrichtung treten naturgemäß auch Verwechslungen auf. So findet man beispielsweise unter kritischen Temperaturen auch untere oder obere kritische Entmischungstemperaturen entmischender Systeme.

### 3.4 Erfolgte und geplante Veröffentlichungen

Tagungen, Messen

- AIChE Annual Meeting (Miami Beach, Nov. 1996) (Poster)
- Sitzung des GVC Fachausschusses „Thermische Zerlegung von Gas- und Flüssigkeitsgemischen“ in Luzern (1996) (Vortrag)
- CHISA 1996 in Prag (Poster)
- ACHEMA 1997 in Frankfurt (Poster, Vorträge)
- Sitzung des GVC Fachausschusses „Thermodynamik“ in Frankfurt (1997) (Vortrag)

Zeitschriften, Tagungsbände

- Rarey, J., Fiege, C., „Reinstoffdaten für die Chemische Technik“, Fraunhofer UMSICHT, Tagungsband UMSICHT-Tage 1998, 14.1-14.12

## **4. Kurzgefaßter Erfolgskontrollbericht**

### **4.1 Beitrag des Ergebnisses zu den förderpolitischen Zielen des Förderprogramms**

Das durchgeführte Projekt zielte auf den Aufbau einer Datenfaktenbank ab. Der Aufbau von Datenfaktenbanken stellte einen wichtigen Teil des inzwischen ausgelaufenen Fachinformationsprogramms des BMBF dar. Datenfaktenbanken sind insbesondere bei der Synthese, Auslegung und Optimierung ( Reduktion der Umweltbelastung und Energiekosten, Vergleich von Verfahrensvarianten, .. ) der unterschiedlichsten Prozesse von großer Bedeutung. Im Gegensatz zu den Literaturdatenbanken, die lediglich Hinweise auf die entsprechenden Literaturstellen liefern, die evtl. nicht oder nur sehr schwer ( z.B. Fernleihe ) beschafft werden können und dann noch ausgewertet werden müssen, wobei mehr als 20 Sprachen berücksichtigt werden müssen ( zumindest im Falle der Dortmunder Datenbank richtig ), erlaubt die erstellte Datenfaktenbank ein sofortiges Abrufen der gewünschten Information, eine Abschätzung mit Hilfe einer Vielzahl von Korrespondenz- oder Gruppenbeitragsmethoden oder auch die Berechnung der Größe für vorgegebene Bedingungen mit Hilfe der gespeicherten Parameter ( recommended values, recommended parameters).

Nach der Förderung anderer chemisch-technischer Datenbanken (Azeotrope Daten, Fest-Flüssig Gleichgewichte, ...) und der Integration der vom Projektleiter in Eigenleistung im Rahmen der Dortmunder Datenbank (DDB) erstellten Datenbanken ( Dampf-Flüssig Gleichgewichte, Mischungsenthalpien, Aktivitätskoeffizienten bei unendlicher Verdünnung, ...) stellt die Datenfaktenbank Reinstoffdaten einen wichtigen Bestandteil des Fachinformationsprogramms des BMBF dar.

### **4.2 Wissenschaftlicher Erfolg des Vorhabens**

Die Datenbank enthält zum Abschluß des Berichtszeitraums mehr als 90 000 Datensätze mit zusammen mehr als 570 000 Datenreihen aus annähernd 10 000 Referenzen. Es wurden insgesamt etwa 18 000 Literaturstellen erfaßt Von etwa 14 000 Literaturstellen liegen Kopien vor. Die verbleibenden 4000 Literaturstellen konnten noch nicht als Kopie erhalten werden. Dabei handelt es sich teilweise um interne Berichte oder in der Literatur fehlerhaft zitierte Quellen. Zusätzlich zur Sammlung der Daten wurden verschiedene Korrelationsmethoden implementiert und zum Test der Daten eingesetzt.

Durch die konsequente Erfassung der Molekülstrukturen und die Entwicklung von Programmen zur Eingabe, Speicherung und Zerlegung dieser Strukturen stehen erstmals eine Vielzahl von Gruppenbeitragsmethoden aus der Literatur auch dem ungeübten Benutzer zur Verfügung.

Durch die einfache Möglichkeit der Datenabschätzung ohne langwierige und fehlerbehaftete manuelle Strukturzerlegung wurde ein umfangreicher Test der publizierten Methoden möglich.

Unter Verwendung der automatischen Strukturzerlegung wurden Programme zur Entwicklung neuer Gruppenbeitragsmethoden entwickelt. Erste Methoden stehen zur Verfügung

### **4.2.1 Nebenergebnisse und Erfahrungen**

Beim Aufbau der Datenbank und dem Umgang mit den Daten und Berechnungsmethoden zeigte sich eine Vielzahl bisher im Arbeitskreis nicht behandelter Phänomene und Datenarten. So ergab sich aus der Diskussion mit Fachleuten, daß oft für die genaue Spezifikation eines Schmelzpunktes die Angabe des Lösungsmittels erforderlich sein kann, aus dem dieser umkristallisiert wurde.

Beim Test der Daten mit Hilfe verschiedener Gruppenbeitragsmethoden zeigte sich deren Anwendungsbreiten und Grenzen zum Teil sehr deutlich. Diese Erfahrungen haben sich bei der Entwicklung eigener Abschätzmethode als sehr nützlich erwiesen.

Bei der Entwicklung der automatischen Strukturzerlegung zeigte sich, wie relativ ungenau die Spezifikation der Strukturgruppen in den Originalpublikationen gehalten ist. Präzise Definitionen von Strukturgruppen für Gruppenbeitragsmethoden sollten eindeutige Regeln für die erlaubte und verbotene chemische Umgebung einer Gruppe im Molekül enthalten.

### **4.3 Einhaltung des Finanzierungs- und Zeitplans**

Der Finanzierungsplan wurde eingehalten. Der Zeitplan wurde um 1 Monat verlängert.

### **4.4 Verwertbarkeit der Ergebnisse**

Mit der Datenbank steht ein Handwerkszeug zur Verfügung, mit dem sich viele Fragestellungen von praktischer Bedeutung bearbeiten lassen. In allen Bereichen des Chemieingenieurwesens ist für die Simulation der einzelnen Grundoperationen oder ganzer Anlagenteile neben genauen Informationen über das Gemischverhalten auch die Kenntnis der Eigenschaften reiner Stoffe erforderlich.

Für die Berechnung von Dampf-Flüssig Gleichgewichten sind die Sättigungsdampfdrücke der beteiligten Komponenten von vorrangiger Bedeutung. Verdampfungsenthalpien geben Auskunft über die Temperaturabhängigkeit der Sättigungsdampfdrücke, den Energiebedarf von Rektifikationskolonnen und sind von großer Bedeutung bei der Berechnung von Destillations- und Rektifikationsanlagen, Wärmetauschern, Siedekühlung von Reaktoren, ... Die Bedeutung zeigt sich z.B. in der Tatsache, daß in etwa 40000 Rektifikationskolonnen ca. 3% des gesamten Primärenergiebedarfs der USA verbraucht wird.

Flüssigdichten und Viskositäten sind entscheidend für den Energiebedarf beim Transport von Flüssigkeiten durch Rohre, Mischen und Rühren, ... Viskositäten und Oberflächenspannungen entscheiden oft über die Wirkungsgrade von Packungskolonnen, Thermodynamische Standardgrößen, speziell Bildungsgrößen ermöglichen die Berechnung der Gleichgewichtslage chemischer Reaktionen. Ebenso sind thermodynamische Größen als Funktion von Temperatur und Druck von vorrangiger Bedeutung bei der Suche nach alternativen Arbeitsmedien (CFC-Ersatz) z.B. für Wärmepumpen.

Durch die Verfügbarkeit von Daten für eine Vielzahl von Komponenten wurden erste verbesserte Abschätzmethode entwickelt, mit deren Hilfe dann auch Daten von Verbindungen berechnet werden können, für die keine experimentellen Daten vorliegen. Weitere Methoden sind im Aufbau.

Auf der Basis solcher Abschätzmethoden ist es dann beispielsweise möglich, Molekülstrukturen mit optimalen Eigenschaften, z.B. als Selektivlösungsmittel für die Lösung von Trennproblemen, zu entwerfen (molecular design).

Durch die Entwicklung der automatischen Strukturzerlegung zusammen mit einer umfangreichen Strukturdatenbank und einer grafischen Eingabeoberfläche sind diese Abschätzmethoden auch für den Nicht-Spezialisten leicht zugänglich.

Die im Projekt entwickelte Datenfaktenbank wird über die DECHEMA, die FIZ CHEMIE GmbH, die DDBST GmbH und Aspen Tech Inc. weltweit als Inhouse-Version sowie online über das Internet vertrieben. Ein direkter Zugriff aus dem Prozeßsimulator Aspen Plus auf die Daten (via Internet) wurde ebenfalls realisiert. Damit stehen die Projektergebnisse den potentiellen Anwendern auf schnellstem Wege zur Verfügung, die damit verbundenen Einnahmen garantieren die Fortführung der Arbeiten zur regelmäßigen Aktualisierung der Datenbanken.

#### **4.5 Arbeiten, die zu keiner Lösung führten**

Im Verlauf des Berichtszeitraums wurden alle darin geplanten Arbeiten durchgeführt. Der Umfang der Datenexzerption ist unter Berücksichtigung der verfügbaren Ressourcen sehr zufriedenstellend und konnte im Vergleich zum ersten Projekt um 50% gesteigert werden. Damit ist eine vollständige Erfassung der Literatur noch nicht erreicht. Es wurde aber inzwischen die Weiterführung der Arbeiten mit privatwirtschaftlicher Förderung sichergestellt.

### **5. Kurzfassung des Schlußberichts**

Zum Ende des Projekts "Datenfaktenbank Reinstoffdaten" stehen ca. 90 000 Datensätze mit zusammen mehr als 570 000 Datenreihen aus ca. 10 000 Referenzen in computergerechter Form gespeichert zur Verfügung. Die Arbeiten wurden in Zusammenarbeit mit Gruppen in Prag, Berlin, Tallinn und Minsk durchgeführt.

Zusätzlich zur Sammlung der Daten wurden verschiedene Korrelationsmethoden implementiert und zum Test der Daten eingesetzt. Mehr als 8000 Parametersätze für verschiedene Gleichungen und Methoden wurden angepaßt.

Durch die konsequente Erfassung der Molekülstrukturen und die Entwicklung von Programmen zur Eingabe, Speicherung und Zerlegung dieser Strukturen stehen erstmals eine Vielzahl von Gruppenbeitragsmethoden aus der Literatur auch dem ungeübten Benutzer zur Verfügung.

Durch die einfache Möglichkeit der Datenabschätzung ohne langwierige und fehlerbehaftete manuelle Strukturzerlegung wurde ein umfangreicher Test der publizierten Methoden möglich. Dieser Test diente vordringlich der Überprüfung der eingegebenen Daten.

Unter Verwendung der automatischen Strukturzerlegung wurden Programme zur Entwicklung neuer Gruppenbeitragsmethoden entwickelt. Die Publikation erster Abschätzmethoden wird vorbereitet.

Die Arbeiten werden mit privatwirtschaftlicher Förderung fortgesetzt.

Oldenburg, 15. 1. 1999

.....

Prof. Dr. J. Gmehling