

유기/무기화합물 열역학물성 데이터센터	유기/무기화합물 기본물성 참조데이터 세부 평가기준서	문서번호	MP-C02-0005
		개정번호	3
		개정일자	2013. 10. 29
		쪽	총 38쪽

목 차

1. 목적
2. 적용범위
3. 세부평가기준
 - 3.1 데이터평가기준 및 적용문서
 - 3.2 평가등급 및 평가절차
 - 3.3 세부평가절차
- 부록 1. 고정물성의 평가방법
- 부록 2. 변수의존 물성의 평가방법
- 부록 3. 최대우도를 이용한 매개변수 회귀분석

제.개정 이력부						
관련조항	개정번호	개정일자	제.개정 내용			
전면	0	'09. 05. 30	- ISO 9001의 요건에 따라 제정 - “참조표준 제정 및 보급에 관한 운영요령“ 제9조 (참조데이터 기술평가기준)에 준하여 제정 - 참조표준 표준 품질매뉴얼에 따라 제정			
	1	'11. 06. 15	- 온도 및 압력 의존 물성의 평가 기준 개정			
	2	'11. 10. 09	- 서면평가 반영			
	3	'12. 10. 29	- 표면장력 외 기본물성에 대한 정의 - 변수의존 물성 평가 절차 보완			
			'13. 10. 29	- 무기화합물 포함		
결재	구 분	작 성	작 성	작 성	검 토	승 인
	직 책	연구원			연구원	데이터센터장
	성 명	김선형			신현용	강정원
	서 명					
	일 자	2013.10.29			2013.10.29	2013.10.29

1. 목적

본 세부 평가기준은 유기화합물 물성 데이터 센터의 수집/생산된 유기화합물 및 무기화합물 기본물성 데이터의 세부평가를 목적으로 한다.

2. 적용범위

이 세부 평가기준서는 유기화합물 물성 데이터센터에서 수집/생산된 유기화합물 및 무기화합물 물성 데이터의 기술적인 세부 평가를 통해 참조데이터의 등급을 부여하는 활동에 적용한다.

3. 세부 평가 기준

3.1 데이터 평가기준 및 적용 문서

2006년 7월 31일 산업자원부에서 고시한 “참조표준 제정 및 보급에 관한 운영요령(산업자원부 고시 제2006-86호)”에서는 참조표준의 정의와 분류 및 평가기준을 다음과 같이 명시하고 있다.

3.1.1 참조표준과 참조데이터

요령 제2조에 참조표준과 참조데이터를 다음 정의에 따라 구분하고 있다.

- ◆ 참조표준 : 측정데이터 및 정보의 정확도와 신뢰도를 과학적으로 분석 평가하여 공인함으로써 국가사회의 모든 분야에서 널리 지속적으로 사용되거나 반복사용이 가능하도록 마련된 자료로서 유효참조표준, 검증참조표준, 인증참조표준으로 분류한다.
- ◆ 참조데이터 : 참조표준으로 승인 받지 못한 수치데이터 또는 과학기술적 통계를 말한다.

3.1.1 참조표준의 분류

요령 제10조에서 참조표준의 등급을 유효참조표준, 검증참조표준 및 인증참조표준 등 3가지로 구분하고 등급부여 기준을 다음과 같이 정하고 있다.

- ◆ 유효참조표준 : 데이터 중에서 참조표준의 기본요구조건을 만족시키는 참조표준
- ◆ 검증참조표준 : 유효참조표준 중에서 잘 알려진 연관식과 모델링을 통한 예측 가능성이 검증된 참조표준
- ◆ 인증참조표준 : 검증참조표준 중에서 관련 전문가의 최종종합평가를 거쳐 인증된 최상위 참조표준

3.1.3 데이터의 기술평가기준

요령 제9조에서 데이터를 평가하여 참조표준으로 등급부여하기 위한 기준을 다음과 같이 정하고 있다.

- ◆ 측정대상이 명확하게 명시되어 있는지 여부
- ◆ 측정방법 및 이론계산에 대한 설명 내용 명시 여부
- ◆ 측정방법의 불확도 평가 및 소급성 유지 여부
- ◆ 측정방법의 적절성 및 한계 명기 여부
- ◆ 측정결과에 영향을 줄 수 있는 변수의 적절한 제어 여부
- ◆ 데이터의 오차율 및 정확도 등의 제시 근거
- ◆ 제3자에 의해 측정방법의 재현이 가능하도록 측정방법과 절차가 자세히 명기되었는지의 여부
- ◆ 다른 방법에 의해 얻어진 결과와 측정결과의 일관성 여부
- ◆ 잘 알려진 법칙과 일치하는지 여부
- ◆ 잘 알려진 연관식과 모델링을 통한 데이터의 예측가능성 검증
- ◆ 관련분야 제3의 전문가 종합검토를 받았는지 여부

3.1.4 참조표준의 등급부여 기준

데이터의 기술평가기준 중 참조표준의 기본요건인 제1호 내지 제7호를 만족하는 데이터는 “유효참조표준”으로 등급을 부여하며, 제8호에서 제10호까지의 일관성과 예측가능성 기준을 만족시키는 유효참조표준은 검증참조표준으로 등급을 부여 하며, 전문가의 종합검토를 받아서 모든 기술평가기준을 만족시키는 검증참조표준은 인증참조표준으로 등급을 부여 한다.

3.2 평가 등급 및 평가 절차

모든 참조 데이터는 그림 1에 나타낸 바와 같이 평가 단계를 통하여 다음의 7등급이 평가를 통하여 부여된다.

- (1) 참조데이터 : 평가를 받지 않았거나 “측정대상의 명시” 단계에서 기각된 자료
- (2) 상용데이터 : “측정 방법의 명시” 단계에서 기각된 자료
- (3) 일반데이터 : “소급성 유지” 단계에서 기각된 자료
- (4) 연구데이터 : “재현성” 단계에서 기각된 자료
- (5) 유효참조표준 : “일관성/예측성” 단계에서 기각된 자료
- (6) 검증참조표준 : “전문가검증” 단계에서 기각된 자료
- (7) 인증참조표준 : 모든 평가단계를 통과한 인증된 참조표준

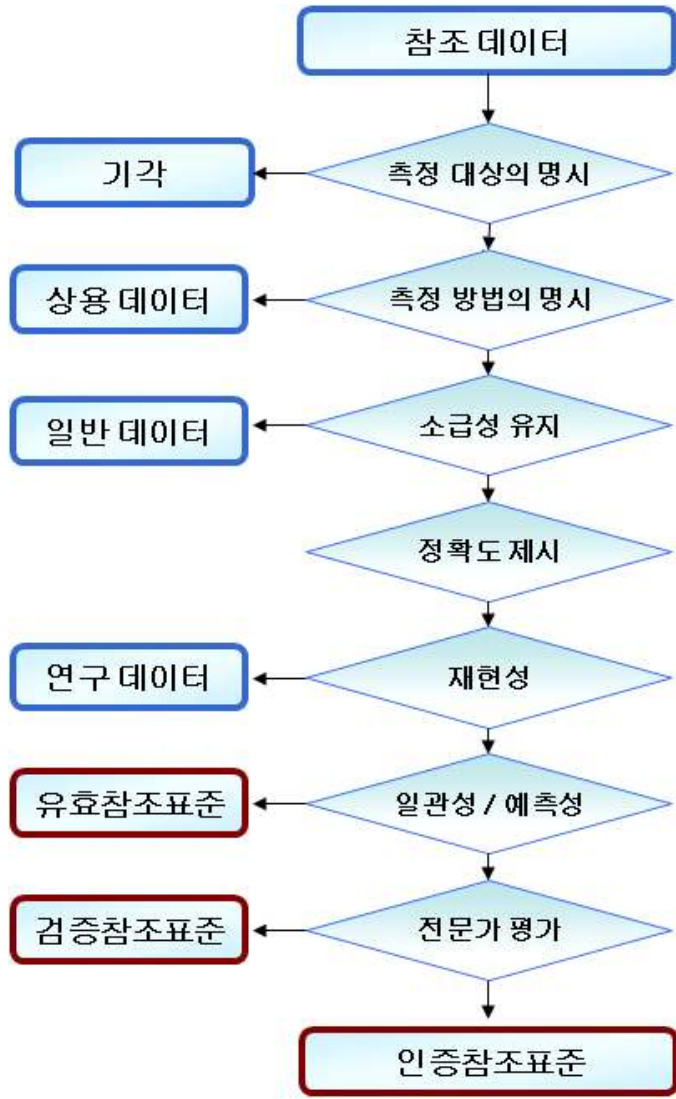


그림 1. 참조데이터의 평가 단계

3.3 세부평가 절차

3.3.1 수집, 평가대상 데이터의 정의

(1) 화합물 물성의 분류

데이터는 그 정의가 명확히 기술되어 있어야 데이터로서의 의미가 있다. 유기/무기 화합물의 기본 물성자료는 대단히 광범위하고 다양하다. 본 절차서 에서 대상으로 하는 유기/무기 화합물의 기본 물성은 다음과 같이 4가지 종류로 분류된다.

◆ 고정물성 (Fixed Properties)

- 물질의 고유 상수로서 온도, 압력 등의 외부 조건에 의존하지 않는 고유 특성치
- 고정물성은 불확도가 포함된 실험으로 측정이 가능한 물성으로 제한함. 간접 특성을 통하여 계산된 물성들은 유도물성으로 분류됨.
- 예) 삼중점, 임계온도, 임계압력 등

◆ 유도 물성 (Derived Properties)

- 다른 물성 값이나 분자의 구조 등에 의하여 계산될 수 있는 물질의 특성치
- 예) 이심인자(Acentric Factor), 용해도 매개변수 (Solubility Parameter) 등

◆ 일변수 의존물성 (One Variable Dependent Properties)

- 한 개의 변수 (온도 등)에 의존하는 물질의 고유 특성치
- 예) 증기압, 증발잠열, 이상기체 열용량 등

◆ 이변수 의존 물성 (Two Variable Dependent Properties)

- 두 개의 변수 (일반적으로 온도와 압력임)에 의존하는 물질의 고유 특성치
- 예) 기체의 밀도, 기체의 열용량 등

◆ 혼합물 물성 (Mixture Properties)

- 두 개 이상의 화합물이 혼합된 상태의 물성으로서 온도, 압력 이외에 조성에 의존하는 물질의 고유 특성치
- 예) 혼합물의 밀도, 혼합물의 점도, 상평형 등

유기/무기 화합물의 고유 물성은 많은 종류가 있지만, 본 절차서에는 열역학 물성과 관련된 물성들을 중심으로 범위를 제한하여 기술하고자 한다. 향후 본 센터에서 구축하고자 하는 참조표준 대상의 수가 늘어나는 경우에는 본 절차서의 개정이 필요하다. 다음의 표 1 - 표 5에는 참조표준 구축의 대상이 되는 물성들의 정의를 나타내었다.

표 1. 고정 물성의 종류

물성명 (영문명)	심벌 (기본단위)	물성의 정의
표준 끓는점 Normal Boiling Point	TB (K)	101.325 kPa에서 기체와 액체가 평형을 이루는 온도 The temperature at which the vapor pressure of the liquid is equal to the ambient pressure (= 101.325 kPa) in an open vessel.
표준 녹는점 Normal Freezing Point	TF (K)	101.325 kPa에서 액체와 고체가 평형을 이루는 온도 The temperature at which a solid and a liquid may in equilibrium at 101.325 kPa.
임계온도 Critical Point Temperature	TC (K)	기체상이 더 이상 액화되지 못하는 임계 온도 The temperature of liquid-vapor transition critical point. The temperature above which the substance has no liquid-vapor transition.
임계압력 Critical Point Pressure	PC (kPa)	기체상이 더 이상 액화하지 못하는 임계온도에서의 증기 압력 The corresponding vapor pressure at the critical temperature.
임계밀도 Critical Density	RHOC (kmol/m ³)	임계온도, 임계압력 하에서의 물질의 밀도 The density of a substance at the liquid-vapor critical point. (at critical temperature and critical pressure)
삼중점 온도 Triple Point Temperature	TT (K)	기체-액체-고체, 3개의 상이 동시에 평형을 이루는 온도 The temperature at which three different phases (vapor, liquid and solid) may be in equilibrium.
삼중점 압력 Triple Point Pressure	PT (kPa)	기체-액체-고체, 3개의 상이 동시에 평형을 이루는 압력 The Pressure at which three different phases (vapor, liquid and solid) may be in equilibrium.
쌍극자 모멘트 Dipole Moment	DM (Debye)	분자의 전하 팽창에 의한 1차 모멘트 (보통 기체상을 의미함) The first moment of the electric charge density expansion of a molecule. Values are given for the gaseous state when available, otherwise the state is noted.

표 2. 유도 물성의 종류 - 1

물성명 (영문명)	심볼 (기본단위)	물성의 정의
표준상태의 생성엔탈피 Standard State Enthalpy of Formation	HIDF (kJ/mol)	표준상태 (298.15 K, 1 bar)에서 구성 원소로부터 화합물이 생성되는데 수반되는 엔탈피의 변화량. The change in enthalpy associated with the reaction forming the given chemical in its standard state from the elements in their standard states.
표준상태의 생성 깁스에너지 Standard State Gibbs Free Energy of Formation	GIDF (kJ/mol)	표준상태 (298.15 K, 1 bar)에서 구성 원소로부터 화합물이 생성되는데 수반되는 깁스 자유 에너지의 변화량. The change in Gibbs free energy associated with the reaction forming the given chemical in its standard state from the elements in their standard states.
표준상태의 생성 엔트로피 Standard State Entropy of Formation	SIDF (kJ/mol.K)	표준상태 (298.15 K, 1 bar)에서 구성 원소로부터 화합물이 생성되는데 수반되는 깁스 자유 에너지의 변화량. The change in entropy associated with the reaction forming the given chemical in its standard state from the elements in their standard states.
이심인자 Accentric Factor	ACCF	Pitzer에 의하여 다음의 식으로 제안된 매개변수 $\omega = -1.0 - \log_{10} \left(\frac{P^{sat}}{P_c} \right)_{T_r=0.7}$
임계압축인자 Critical Compressibility	ZC	임계온도, 압력, 밀도에서의 압축인자. 다음의 식으로 정의됨. $Z_c = \frac{P_c}{\rho_c RT_c}$
용해도 매개변수 Solubility Parameter	SOLP (MPa ^{1/2})	Hildebrand Solubility Parameter, cohesive energy density의 제곱근을 의미함. $SOLP = \sqrt{\frac{\Delta H_v - RT}{V_m}}, H_v = \text{Heat of vaporization}$
액체 몰부피 Liquid Molar Volume	VMOL (m ³ /kg-mol)	기준 온도, 압력에서의 액체의 몰부피 (1기압, 298.15 K) Molar volume of the liquid at reference T and P.

표 3. 유도 물성의 종류 - 2

물성명 (영문명)	심벌 (기본단위)	물성의 정의
Van der Waals Volume	VVDW (m ³ /kg-mol)	Hard sphere로 구성된 이상적인 분자모델의 부피. The excluded volume of model molecules treated as hard-sphere beads separated by rigid bonds. Group increments from Bondi are used for calculation.
Van der Waals Area	AVDW (m ² /kg-mol)	Hard sphere로 구성된 이상적인 분자모델의 표면적. The excluded surface area of model molecules treated as hard-sphere beads separated by rigid bonds. Group increments from Bondi are used for calculation.
UNIQUAC 모델의 R_i	RI	UNIQUAC 모델에 사용되는 Ri 매개변수 (Segment Volume) $R_i = VVDW/15.76(cm^3/mol)$
UNIQUAC 모델의 Q_i	QI	UNIQUAC 모델에 사용되는 Qi 매개변수 (Segment Area) $\frac{z}{2} Q_i = AVDW/(2.5 \times 10^9)(cm^2/mol)$
UNIQUAC 모델의 Q_i'	ACCF	UNIQUAC 모델에 사용되는 수정된 Qi 매개변수 (Segment Area) 물, 알코올 등 특수한 경우에만 활용됨.
Dimerization Constant A	DIMA	카르복실 산 등에 대한 Dimerization 반응상수의 매개변수 A $\log_{10}K_D = A + B/T$
Dimerization Constant B	DIMB	카르복실 산 등에 대한 Dimerization 반응상수의 매개변수 B $\log_{10}K_D = A + B/T$

표 4. 일변수 의존 물성의 종류 - 1

물성명 (영문명)	심볼 (기본단위)	물성의 정의
증기압 Vapor Pressure	PB-VL (kPa)	기체-액체 상평형의 경계 압력 Vaporization pressure for liquid-vapor equilibrium
승화압 Sublimation Pressure	PB-VS (kPa)	고체-기체 상평형의 경계 압력 Sublimation pressure for solid-vapor equilibrium.
용융압 Fusion Pressure	PB-LS (kPa)	고체-액체 상평형의 경계 압력 Fusion pressure for solid-vapor equilibrium.
증발잠열 Heat of Vaporization	HT-VL (kJ/kg-mol)	주어진 온도와 증기압에서 기체-액체 상변화에 수반되는 비 엔탈피의 변화량 Difference between the enthalpies of a unit mole of a saturated vapor and saturated liquid of a pure component at any temperature and corresponding vapor pressure
승화열 Heat of Sublimation	HT-VS (kJ/kg-mol)	주어진 온도와 승화압에서 고체-기체 상변화에 수반되는 비 엔탈피의 변화량 Difference between the enthalpies of a unit mole of a saturated vapor and solid of a pure component at any temperature and corresponding sublimation pressure
용융열 Heat of Fusion	HT-LS (kJ/kg-mol)	주어진 온도와 용융압에서 기체-액체 상변화에 수반되는 비 엔탈피의 변화량 Difference between the enthalpies of a unit mole of a liquid and solid a pure component at any temperature and corresponding fusion pressure
이상기체 열용량 Ideal Gas Heat Capacity	CP-IG (kJ/kg-mol.K)	일정한 압력에서 이상기체 1몰을 1도(K) 상승시키는데 필요한 에너지의 양 The amount of energy required to change the temperature of a unit mole of vapor one degree when the vapor is ideal.
포화액체 열용량 Saturated Liquid Heat Capacity	CP-VL (kJ/kg-mol.K)	포화 액체 1몰을 1도 (K) 상승시키는데 필요한 에너지의 양 The amount of energy required to change the temperature of a unit mole of liquid one degree at low pressure.
고체 열용량 Solid Heat Capacity	CP-VS (kJ/kg-mol.K)	낮은 압력에서 고체 1몰을 1도 (K) 상승시키는데 필요한 에너지의 양 The amount of energy required to change the temperature of a unit mole of solid one degree at low pressure.

표 5. 일변수 의존 물성의 종류 - 2

물성명 (영문명)	심볼 (기본단위)	물성의 정의
이상기체 점도 Ideal Gas Viscosity	VS-IG	The shear stress per unit area at any point in a confined Newtonian ideal gas vapor divided by the velocity gradient in the direction perpendicular to the direction of flow
포화액체 점도 Low Pressure Liquid Viscosity	VS-VL	The shear stress per unit area at any point in a confined Newtonian liquid at low P divided by the velocity gradient in the direction perpendicular to the direction of flow
이상기체 열전도도 Ideal Gas Thermal Conductivity	TH-IG	The proportionality constant in Fourier's law of heat conduction which describes the rate at which heat flows through a vapor in ideal gas state.
포화액체 열전도도 Low Pressure Liquid Thermal Conductivity	TH-VL	The proportionality constant in Fourier's law of heat conduction which describes the rate at which heat flows through a liquid at low P.
표면장력 Surface Tension	SRF-VL	The inherent force in the plane of the surface of a gas-liquid interface per unit length of surface which tends to minimize the surface area

(2) 시료 정보

측정 대상에는 유기/무기화합물의 기본물성 측정에 사용된 시료도 포함된다. 시료에 대한 정보가 정확히 기술되어 있지 않다면 자료로서 가치가 의심된다. 따라서 시료에 대한 충분한 정보가 제공되어 있지 않다면 자료로서의 의미가 상실된다. 시료 정보에는 다음과 같은 정보가 포함될 수 있다.

- ◆ 시료의 출처 : 시약회사명, 저자에 의하여 합성, 표준물질, 자연 물질 등
- ◆ 시료의 순도 : 질량 %, 부피 % 등
- ◆ 시료의 순도 결정방법 : 시약회사의 순도 표기, 직접 결정 (측정)
- ◆ 시료의 순도 향상을 위한 전처리 방법
- ◆ 전처리 후 시료의 순도
- ◆ 시료의 밀도
- ◆ 시료의 수분의 함량

최소한의 정보로서 사용된 시료의 출처와 최종순도는 표시되어야 한다.

(3) 표현의 적절성

자료의 표현이 적절히 되어있는지 확인한다. 다음의 항목이 평가된다.

- ◆ 일관성 있게 표현. 즉 상수, 단위의 통일 등
- ◆ 유효숫자의 처리, 데이터 입력방법, 데이터의 내삽 및 외삽 등과 같은 데이터 편집의 적정성 여부

일관성, 단위와 상수의 통일, 유효숫자 처리, 데이터의 내삽 및 외삽 방법 등을 검토하여 합격 여부를 결정한다.

3.3.2 측정 방법

기본 물성의 측정에는 다양한 방법이 활용될 수 있다. 기본물성의 정확한 측정을 위하여 새로운 방법들이 도입되고 있다. 새로운 방법을 도입한 일부 연구결과에는 상세한 측정 방법이 소개되고 있지만, 상당수의 논문에는 실험방법이 기재되어 있지 않고, 참고 문헌 소개로 그치는 경우도 많다. 각 물성별로 잘 알려진 실험방법의 일부 사례는 표 5에 나타낸 바와 같다. 참조표준 데이터의 구축을 위해서 측정방법을 가능한 한 표시하는 것이 바람직 하지만 정확한 표현이 불가능 한 경우에는 실험자료 원자료의 정확한 표시로서 측정 방법을 대신할 수 있다.

표 5. 고정 물성 측정방법 표시의 예

Property Name	Experimental Method
Normal Boiling Point	Manometric method
	Closed cell (static) method
	Isoteniscope
	Distillation
	Ebulliometric method (recirculating still)
	Twin ebulliometer
Normal Freezing Point / Triple Point Temperature	Visual observation
	Heating/cooling curves
	DSC/DTA
	Adiabatic calorimetry
	AC calorimetry
	Resistive pulse heating
	Laser pulse heating
Levitation method	
Critical Point Temperature	Visual observation in an unstirred cell
	Visual observation in a stirred cell
	DSC/DTA
	Derived from PVT data
	Relative permittivity measurements
	Refractive index measurements
Critical Point Pressure	Direct measurement
	Extrapolated vapor pressure
	Derived from PVT data
Critical Volume /Density	Rectilinear Diameter
	Relative permittivity measurements
	Refractive index measurements
	Derived from PVT data
	DSC/DTA

3.3.3 소급성

간접 생산 방식의 경우, 소급성에 대한 정보는 일반적으로 문헌에 표시되지 않는다. 따라서 소급성에 대한 평가는 생략한다.

3.3.4 정확도 제시

모든 유기/무기화합물 물성자료의 경우, 불확도에 대한 정보를 제시하는 것은 매우 중요하다. 자료의 불확도는 표준 합성 불확도로 표시하는 것이 바람직 하지만, 간접생산의 경우 논문의 출판 연도에 따라 자료간의 표준 편차 또는 B형 불확도 만 표시된 경우들이 많으므로, 이를 고려하여야 한다. 유기/무기화합물의 기본 물성 자료의 경우에는 여러 가지 자료의 평균값을 산출하여 계산하는 방식 이므로, 재현성 평가 단계에서 여러 자료들의 값을 통계 처리하여 불확도를 평가하여 이를 전체 자료의 불확도로 제시하도록 한다.

3.3.5 재현성 평가

재현성은 제3자의 실험을 통하여 동일한 결과가 반복될 수 있는지 여부이다. 유기/무기화합물의 기본 물성은 표 1.1에 제시한 바와 같이 순수한 성분에 대한 물성으로서 그 정의가 명확하고 표 1.2에 제시한 바와 같이 일반적인 실험 방법에 제시되어 있다. 유기/무기화합물 순수 물성의 경우, 잘 정의된 물성의 경우, 제 3자의 실험을 통하여 동일한 결과가 반복될 수 있는지 문헌의 내용을 검토하여 평가하도록 한다.

3.3.6 일관성 평가

유기/무기화합물의 기본 물성자료의 일관성 평가과정에서, 고정 물성인 경우와 상변수에 의존하는 물성의 경우 다른 평가 방법이 적용되어야 하며, 그 평가과정은 각 물성 별로 다음과 같이 부록에 표시하였다.

- ◆ 고정물성의 경우 : <부록 1. 고정물성의 평가 방법> 참조
- ◆ 유도물성의 경우 : 계산된 물성이므로 별도의 평가를 거치지 않음
- ◆ 일변수 의존 물성의 경우 : <부록 2. 변수 의존 물성의 평가방법> 참조
- ◆ 이변수 의존 물성의 경우 : <부록 2. 변수 의존 물성의 평가방법> 참조
- ◆ 상평형 물성의 경우 : 향후 추가, 보완 예정

3.3.7 전문가 평가

전문가 평가는 외부전문위원회를 소집하여 최종 인증 이전 단계에서 자료를 평가하는 작업이다. 이 평가 과정에서 전문가 그룹은 다음과 같은 사항을 검토한다.

- ◆자료의 종합 점수 검토
- ◆재현성/일관성 평가의 적절성 검토 (합격/불합격에 포함되지 않은 평가 정보 검토)
- ◆저널의 IF 등 신뢰도 검토
- ◆저널/보고서/리포트 저자의 신뢰도 검토

이 과정에서 기-액 상평형 분야의 전문적인 지식을 통하여 일반적인 평가 절차에서 누락될 수 있는 요소를 발견할 수 있으며, 문제점을 제시할 수도 있다. 전문가 평가를 위한 전문가 그룹은 별도로 선임하여 전문가 평가를 실시하며, 전문가 평가 2개월 전에 선임하도록 한다.

부록 1. 고정 물성의 평가 방법

1. 고정물성 일관성 평가의 개요

고정물성은 1800년대부터 많은 자료들이 문헌상에 보고되어 있으나 불확도 개념이 보급되기 이전의 오래된 자료들은 측정 변수들의 불확도 들이 적절히 보고되어 있지 않은 경우들이 많다. 따라서 본 참조표준의 개발에서는 불확도 정보가 충분한 경우와 그렇지 않은 경우로 나누어 일관성 평가 방법을 달리하고, 평가된 사항에 따라서 등급을 달리 부여하는 방법을 채택하였다. 다음의 그림에 전체적인 일관성 평가 방법을 도시하였다.

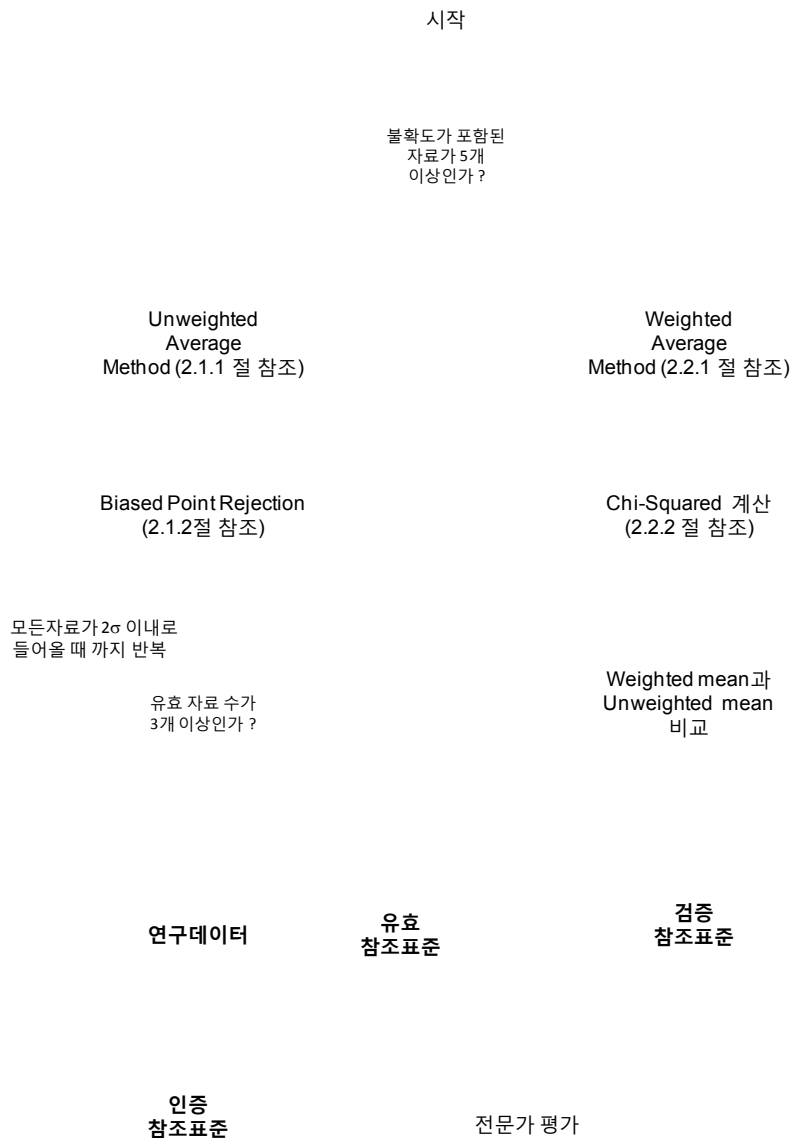


그림 A1.1 고정물성 평가방법의 흐름도

2. 고정물성 평가 방법

2.1 불확도 정보가 불충분한 자료의 평가

불확도 정보가 충분히 주어진 경우, 다음의 사례를 참조하여 유사한 방법을 도입하여 고정물성을 평가한다.

예시) n-Nonane의 Critical Density

표 A1.1 Nonane의 임계밀도 실험자료

값 (kg-mol/m ³)	Uncertainty (kg-mol./m ³)	Reference
1.826	N/A	Daubert, T.E.,Danner, R.P.,Technical Data Book-Petroleum Refining, 5th Edition, American Petroleum Institute, Washington, D.C., (1992)
1.801	0.05	Anselme, M. J.,Gude, M.,Teja, A. S.,Fluid Phase Equilib., 57, 317-26 (1990)
1.838	N/A	Tsonopoulos, C.,Tan, Z.,Fluid Phase Equilib., 83, 127-138 (1993)

(1) 산술 평균 (Unweighted Mean)과 표준 편차 계산

산술평균과 표준 편차는 다음의 수식으로 계산된다.

■ Unweighted Mean : $x_u = \sum_{i=1}^N \frac{x_i}{N}$

■ Standard Deviation : $\sigma_u = \sqrt{\frac{\sum_{i=1}^N (x_i - x_u)^2}{N(N-1)}}$

예제의 경우,

$$x_u = \frac{1.826 + 1.801 + 1.838}{3} = 1.822 \quad \sigma_u = 0.011$$

(2) Biased Point의 Rejection

특히 오래된 문헌상에 오류가 있는 자료들이 많이 있으므로 이를 배제하기 위하여 Biased Point Rejection이 필요하다. Biased Point Rejection을 위해서는 Chauvenet's criterion을 이용하여 다음과 같은 기준으로 Outlier를 제거한다.

- 1) Standard Deviation을 구한다.
- 2) Standard Deviation에서 가장 먼 자료를 택한다.
- 3) 정규 분포에 근거하여 가장 멀리 떨어진 자료를 택할 확률을 구한다.
- 4) (확률)*(자료의 수)가 0.5이하 이면 자료를 버린다.

(3) 최종 값의 보고

최종 값은 다음과 같이 보고한다.

$$x = x_u \pm 2\sigma_u, (k = 2)$$

예제의 경우,

$$\rho_c = 1.822 \pm 0.022 \text{ kg} - \text{mol}/\text{m}^3 \quad (k = 2)$$

제시된 불확도는 확장불확도이며, 불확도의 표현은 2자리로 제한한다. 아울러 보고된 값의 유효 자리 수는 불확도와 일치시키도록 한다.

산술평균으로 처리된 자료는 일관성을 보장할 수 없으므로 유효참조표준으로 등급을 부여한다.

유효한 자료수 (rejection 되지 않은 자료들)가 3개 이하인 경우에는 불확도를 보고하지 않고, 단순히 산술 평균값 만을 보고하며, 연구 데이터로 처리하도록 한다.

2.2 불확도 정보가 충분한 자료의 평가

불확도 정보가 충분히 제공된 자료들은 Weighted Mean을 취하며, Chi-squared 계산과 Unweighted mean과의 비교를 통하여 일관성을 평가한다.

예시) n-Nonane의 Critical Pressure

표 A1.1 Nonane의 임계압력 실험자료

값 (kPa)	Uncertainty (kPa)	Reference
2313.2	N/A	Kay, W. B., Hissong, D. W., Proc. - Am. Pet. Inst., Div. Refin., 47, 653 (1967)
2288.0	10	Ambrose, D., Townsend, R., Trans. Faraday Soc., 64, 2622-2631 (1968)
2293.9	6.9	Pak, S. C., Kay, W. B., Ind. Eng. Chem. Fundam., 11, 255 (1972)
2291.1	0.2	Kay, W. B., Pak, S. C., J. Chem. Thermodyn., 12, 673 (1980)
2270.0	30	Matzik, I., Schneider, G. M., Ber. Bunsen-Ges. Phys. Chem., 89, 551 (1985)
2299.0	5	Brunner, E., J. Chem. Thermodyn., 20, 273 (1988)
2280.0	20	Rosenthal, D. J., Teja, A. S., AIChE J., 35, 1829 (1989)

본 자료의 경우, 첫 번째 실험자료는 불확도 정보가 없으므로 가중 평균을 구하는데 활용할 수 없다. 따라서 가중평균 비교 테스트를 위해서는 이 자료는 포함하지 않으며, 이후 6개의 자료만을 활용한다.

(1) 산술평균과 가중 평균의 계산

먼저 비교를 위한 산술 평균과 표준편차를 구한다.

■ Unweighted Mean : $x_u = \sum_{i=1}^N \frac{x_i}{N}$

■ Standard Deviation : $\sigma_u = \sqrt{\frac{\sum_{i=1}^N (x_i - x_u)^2}{N(N-1)}}$

n-Nonane의 Critical Pressure 의 경우 ,

$$x_u = 2287.00 \sigma_u = 4.27$$

(2) 가중평균과 표준편차를 구한다.

■ Weight : $w_i = \frac{1}{\sigma_i^2}$

값 (kPa)	Uncertainty (kPa)	Weight	Relative Weight
2313.2	N/A	-	-
2288.0	10	0.01	0.0004
2293.9	6.9	0.021	0.0008
2291.1	0.2	25.0	0.9970
2270.0	30	0.0011	0.00004
2299.0	5	0.04	0.0016
2280.0	20	0.0025	0.0001

- 한 포인트의 weight가 너무 크므로 (0.9970) 이를 0.5로 조정함

값 (kPa)	Uncertainty (kPa)	조정된 Weight	Relative Weight
2313.2	N/A	-	-
2288.0	10	0.01	0.0670
2293.9	6.9	0.021	0.1407
2291.1	0.2	0.0746	0.5000
2270.0	30	0.0011	0.0074
2299.0	5	0.04	0.2680
2280.0	20	0.0025	0.0167

Weighted Mean : $x_w = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i} = 2293.06$

Standard Deviation of Weighted Mean : $\sigma_w = \sqrt{\frac{1}{\sum_{i=1}^N w_i}} = 2.6$

(3) Chi-squared 계산

Chi-squared

$$\chi_i^2 = \frac{(x_i - x_w)^2}{\sigma_i^2}$$

값 (kPa)	Uncertainty (kPa)	Weight
2313.2	N/A	-
2288.0	10	0.256
2293.9	6.9	0.015
2291.1	0.2	96.12
2270.0	30	0.591
2299.0	5	1.411
2280.0	20	0.426
Total Chi-squared		98.21

(4) Reduced Chi-squared와 Birge Ratio 를 구함

$$\text{Reduced Chi-square} : \chi_R^2 = \frac{\chi^2}{N-1} = \frac{98.21}{6-1} = 19.76$$

(5) Unweighted Mean과 Weighted Mean의 comparison

$$|x_u - x_w| \leq \sigma_u + \sigma_w$$

본 예제에서,

$$|2287.00 - 2293.06| = 6.6$$

$$4.3 + 2.6 = 6.9$$

두 개의 평균값은 Overlap 됨 따라서, Weighted mean을 취함.

$$\sigma = \sigma_w * \sqrt{\chi_R^2} = 2.6 \times \sqrt{19.76} = 12$$

(6) 최종 값의 보고

최종 값의 보고는 확장 불확도를 포함하여야 하며, coverage factor를 표시한다.

예제에서,

n-Nonane의 Critical Pressure

$$P_c = 2293 \pm 24 \text{ kPa} (k = 2)$$

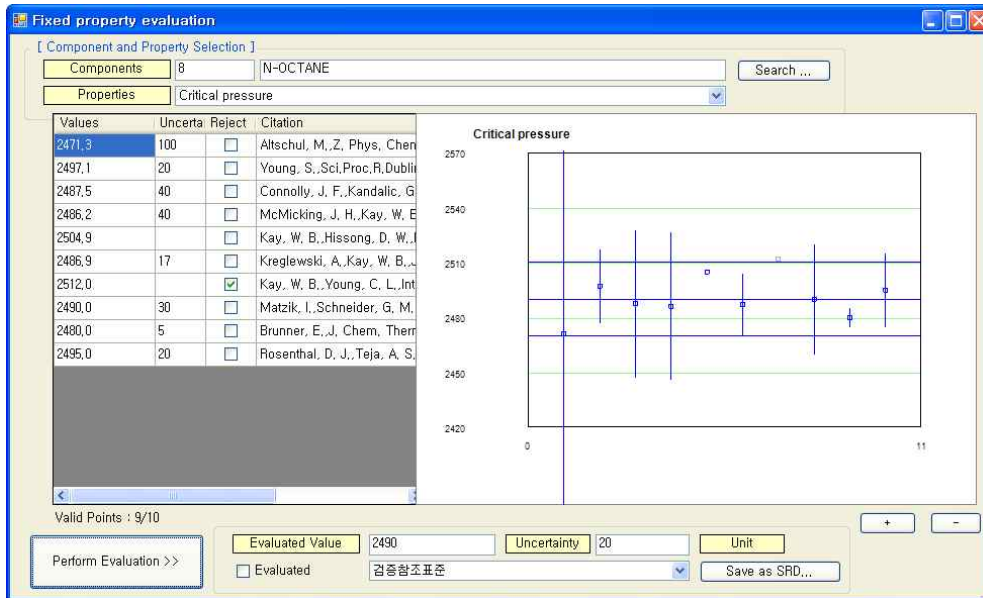


그림 A1.2 여러 개의 값이 있는 경우 참조 표준을 선정하는 프로그램화면

부록 2. 변수의존 물성의 평가 방법

1. 변수의존 물성의 일관성 평가의 개요

온도, 압력 및 조성에 따라 그 값이 변하는 의존물성은 단 한 개의 값을 가지는 물성이 아니므로 참조 표준의 표현 방법이 적절한 수식(표준 모델식)으로 표현되어야 하며, 평가방법도 고정물성과 달라야 한다. 변수 의존 물성들은 선형상관식 (Linear Correlation) 또는 비선형 상관식 (Nonlinear Correlation)으로 표현될 수 있으며, 물성의 특성에 따라 적절한 표준 모델식이 선정되어야 한다.

1.1 선형 상관식 (Linear Corelation)

독립변수 (x : 온도, 압력 또는 조성)와 의존변수 (y : 물성값)의 관계가 선형인 경우를 의미한다.

$$y(x) = \sum_{k=1}^M a_k X_k(x)$$

1.2 비선형 상관식 (Nonlinear Correlation)

독립변수 (x : 온도, 압력 또는 조성)와 의존변수 (y : 물성값)의 관계가 비선형인 경우를 의미한다.

$$y(x) = y(x; \mathbf{a})$$

선형 관계식 또는 비선형 관계식에서 a_k 값이 매개 변수가 되며, 이 값 들을 참조 표준 값으로 보고하여야 한다.

2. 변수 의존 물성의 상관모델

2.1 일변수 의존 물성 (One variable dependent properties)

온도에 의존하는 물성은 다음에 제시된 바와 같이 그동안 널리 활용되어온 모델식들 중하나를 사용하며, 수식번호 1번을 표준 모델식으로 채택한다. 표준 모델식이 적절한 수준으로 물성을 표현하지 못하는 경우 대안식들 중 하나를 선택하여 활용할 수 있다.

(1) 증기압 (Vapor Pressure)

수식 번호	수식명	표현식 (p_{vp} in kPa, T in K, $p_o=1\text{kPa}$)
PB-VP-1	Wagner Eqn. - 1	$\ln p_{vp}/p_c = (T/T_c) [a_1\tau + a_2\tau^{1.4} + a_3\tau^{2.5} + a_4\tau^5], \tau = 1 - T/T_c$
PB-VP-2	Wagner Eqn. - 2	$\ln p_{vp}/p_c = (T/T_c) [a_1\tau + a_2\tau^{1.4} + a_3\tau^3 + a_4\tau^6], \tau = 1 - T/T_c$
PB-VP-3	DIPPR PVP Eqn. - 1	$\ln p_{vp}/p_o = a_1 + a_2/T + a_3 \ln T + a_4 T^2 + a_5/T^2$
PB-VP-4	DIPPR PVP Eqn - 2	$\ln p_{vp}/p_o = a_1 + a_2/T + a_3 \ln T + a_4 T^{a_5}$
PB-VP-5	Antoine Eqn.	$\ln p_{vp}/p_o = a_1 + a_2/(T + a_3)$

(2) 증발 잠열 (Heat of Vaporization)

수식 번호	수식명	표현식 (H_{vp} in kJ/mol, T in K, $H_0=1\text{kJ/mol}$)
HT-VP-1	Watson Eqn	$\ln H_{vp}/H_0 = a_1 + a_2 \ln(1 - T_r) + a_3 T_r \ln(1 - T_r) + a_4 T_r^2 \ln(1 - T_r)$
HT-VP-2	Yaws HVP Eqn.	$H_{vp} = a_1 \left[1 - \frac{T}{T_c} \right]^{a_2}$

(3) 이상기체 열용량 (Ideal Gas Heat Capacity)

수식 번호	수식명	표현식 (C_p^g in kJ/mol.K, T in K)
CP-IG-1	Polynomial Eqn.	$C_p^g = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4$
CP-IG-2	Aly Lee Equation	$C_p^g = a_1 + a_2 [(a_3/T)\sinh(a_3/T)]^2 + a_4 [(a_5/T)\sinh(a_5/T)]^2$
CP-IG-3	Wilhoit Eqn	$C_p^g/R = a_1 + (a_2/T^2)\exp(-a_3/T) + a_4 y^2 + (a_5 - a_6/(T - a_8)^2)y^8$ where $y = (T - a_8)/(T + a_7)$

(4) 저압 액체 열용량 (Liquid Heat Capacity at low P)

수식 번호	수식명	표현식 (C_p^l in kJ/mol.K, T in K)
CP-VL-1	Polynomial Eqn.	$C_p^l = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4$
CP-VL-2	CS Expansion	$C_p^l = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5/\tau$, where $\tau = 1 - T/T_c$

(5) 고체 열용량 (Solid Heat Capacity)

수식 번호	수식명	표현식 (C_p^s in kJ/mol.K, T in K)
CP-VS-1	Polynomial Eqn.	$C_p^s = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4$

(6) 포화 액체 밀도 (Saturated Liquid Density)

수식 번호	수식명	표현식 (ρ_{sat} in kg-mol/m ³ , T in K)
DN-VL-1	VDNS expansion Eqn.	$\rho_{sat} = \rho_c + a_1 \tau^{0.35} + a_2 \tau + a_3 \tau^2 + a_4 \tau^3$, where $\tau = 1 - T/T_c$
DN-VL-2	Rackett Equation	$\rho_{sat} = \rho_c a_1^{-\tau^{a_2}}$, where $\tau = 1 - T/T_c$
DN-VL-3	PPDS Expansion	$\rho_{sat} = \rho_c + a_1 \tau^{0.35} + a_2 \tau^{2/3} + a_3 \tau + a_4 \tau^{4/3}$, where $\tau = 1 - T/T_c$

(7) 이상기체 기체 점도 (Ideal Gas Viscosity)

수식 번호	수식명	표현식 (μ_g in cP, T in K)
VS-IG-1	Polynomial Eqn.	$\mu_g = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4$
VS-IG-2	DIPPR VSG Eqn.	$\mu_g = a_1 T^{a_2} / (1 + a_3 / T + a_4 T^2)$

(8) 포화 액체 점도 (Saturated Liquid Viscosity)

수식 번호	수식명	표현식 (μ_l in cP, T in K, $\mu_g^0 = 1\text{cP}$)
VS-VL-1	Polynomial Eqn.	$\ln(\mu_l / \mu_l^0) = a_1 + a_2 / T + a_3 T + a_4 T^2$
VS-VL-2	Yaws VSL Eqn.	$\ln(\mu_l / \mu_l^0) = a_1 + a_2 / T + a_3 / T^2 + a_4 T^3$
VS-VL-3	PPDS VSL Eqn.	$\ln(\mu_l / \mu_l^0) = a_1 X^{1/3} + a_2 X^{3/4} + \ln(a_5)$ where $X = \frac{a_3 - T}{T - a_4}$

(9) 저압 기체 열전도도 (Low Pressure Gas Thermal Conductivity)

수식 번호	수식명	표현식 (κ_g in W/m.K, T in K)
TH-IG-1	Polynomial Eqn.	$\kappa_g = a_1 + a_2 T + a_3 T^2 + a_4 T^3$

(10) 포화 액체 열전도도 (Saturated Liquid Thermal Conductivity)

수식 번호	수식명	표현식 (κ_l in W/m.K, T in K)
TH-VL-1	Polynomial Eqn.	$\kappa_l = a_1 + a_2 T + a_3 T^2 + a_4 T^3$
TH-VL-2	PPDS THL eqn	$\ln(\kappa_l / \kappa_l^0) = a_1 + a_2 (1 - T/a_3)^{2/7}$

(11) 표면장력 (Surface Tension)

수식 번호	수식명	표현식 (κ_l in W/m.K, T in K)
SF-VL-1	PPDS-14 eqn.	$\sigma/\sigma^0 = a_1 \tau^{a_1} (1 + a_3 \tau), \tau = 1 - T/T_c, \sigma_0 = 1N/m$
SF-VL-2	Yaws eqn.	$\sigma/\sigma^0 = \exp(A) (1 - \frac{T}{T_c})^n, \sigma_0 = 1N/m$
SF-VL-3	Watson eqn.	$\ln(\sigma/\sigma^0) = a_1 + \sum_{i=2}^n a_i T_r^{i-1} \ln(1 - T_r), \sigma_0 = 1N/m$
SF-VL-4	IST Expansion	$\sigma/\sigma_0 = \sum_1^n a_i (1 - T/T_c)^i, \sigma_0 = 1N/m$

2.2 이변수 의존 물성 (Two variable dependent properties)

온도 및 압력에 의존하는 물성은 상태방정식(Equation of State) 등 적절한 모델과 함께 제시되어야 하지만 실험 자료의 상관식이 널리 활용되고 있는 액체의 밀도와 기체상의 소리 전파속도에 한해 다음에 제시된 바와 같은 상관식을 활용한다.

(1) 단일상 액체 밀도 (Single Phase Liquid Density)

수식 번호	수식명	표현식 (ρ in kg/m ³ , T in K, p in kPa)
DN-L-1	Tait Equation (Default)	$\rho = \rho_{sat} / \left[1 - C(T) \ln \left\{ \frac{B(T) + p}{B(T) + p^{sat}} \right\} \right]$ <p>where</p> $B(T) = \sum_{i=1}^{n_{terms}} b_i \tau^{i-1}, C(T) = \sum_{i=1}^{n_{terms}} b_i \tau^{i-1}$ $\tau = 1 - T/T_c$

(2) 기체상의 소리전파 속도 (Speed of Sound in Gas)

수식 번호	수식명	표현식 (v_{sp}^g in m/s, T in K, p in kPa)
SOS-1	Polynomial CPS Eqn.(Default).	$v_{sp}^g = a_1 + a_2 T + a_3 p + a_4 p / T$

2.3 혼합물 물성 (Mixture Properties)

조성에 의존하는 물성은 혼합물의 물성을 직접 나타내거나 또는 과잉물성 (Excess Property)의 형태로 나타내는데, 2성분 혼합물의 경우, 다음과 같은 일반적인 조성에 대한 다항식 또는 Pade approximation의 형태로 표현한다.

수식 번호	수식명	표현식
MIX-1	Mixture properties polynomial	$P^{mix} = P_1x_1 + P_2x_2 + \sum_{i=1}^{nterms} a_i x_1 x_2 (x_1 - x_2)^{i-1}$
MIX-2	Mixture properties Pade eqn.	$P^{mix} = P_1x_1 + P_2x_2 + \frac{\sum_{i=1}^{nterms_1} a_i x_1 x_2 (x_1 - x_2)^{i-1}}{\sum_{j=1}^{nterms_2} b_j x_1 x_2 (x_1 - x_2)^{j-1}}$
EXC-1	Excess property polynomial	$P^E = \sum_{i=1}^{nterms} a_i x_1 x_2 (x_1 - x_2)^{i-1}$
EXC-2	Excess properties Pade eqn.	$P^E = \frac{\sum_{i=1}^{nterms_1} a_i x_1 x_2 (x_1 - x_2)^{i-1}}{\sum_{j=1}^{nterms_2} b_j x_1 x_2 (x_1 - x_2)^{j-1}}$

3. 의존 물성의 회귀분석 방법

온도, 압력 및 조성에 따라 얻어진 실험 자료를 이용하여 매개변수 (a_k)값들을 얻을 때에는 다음과 같은 두 가지 방법을 활용할 수 있다.

- ◆ 최소자승법을 활용한 매개변수 추정 (Method of Least Square)
- ◆ 최대우도를 활용한 매개변수 추정 (Method of Maximum Likelihood)

본 평가기준에서 일변수 의존 물성 및 이변수 의존 물성의 회귀분석에는 최소자승법을 활용하며, 상평형 계산의 경우에는 최대우도를 활용한 매개변수 추정 방법을 활용한다. 최대우도를 활용한 매개변수 추정 방법은 부록.3에 별도로 나타내었다. 최소자승법의 경우, Merit Function 또는 Objective Function을 구성하여 컴퓨터 프로그램을 활용하여 최적의 매개변수 값을 도출한다.

- ◆ 선형 상관식 (Linear Corelation)의 Merit Function

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - \sum_{k=1}^M a_k X_k(x_i)}{\sigma_i} \right]^2$$

위식에서 i 는 실험 자료에 대한 index로서 1부터 N 까지 (N : 실험 자료의 수)합산하며, k 는 선형관계식의 차수에 해당하며, 1부터 M 까지 (M : 최대 차수) 합산한다.

- ◆ 비선형 상관식 (Nonlinear Corelation)의 Merit Function

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - y(x_i; \mathbf{a})}{\sigma_i} \right]^2$$

위식에서 i 는 실험자료에 대한 index로서 1부터 N 까지 (N : 실험자료의 수)합산하며, \mathbf{a} 는 매개변수의 벡터(차원, M)이다.

위와 같은 매개변수들은 선형 또는 비선형의 분류에 따라 적절한 수치해석 방법을 활용하여 매개변수의 값을 도출한다. 상관식의 종류에 따라 적용이 가능한 수치 해석 방법은 다음과 같다.

상관식의 종류	방법명	Reference
선형 상관식	Gauss-Jordan Method	W.H.Press, B.P.Fannery, S.A. Teukolsky, W. T. Vetterling, "Numerical Recipes", Cambridge Press (1986)
	Singlar Decomposition	W.H.Press, B.P.Fannery, S.A. Teukolsky, W. T. Vetterling, "Numerical Recipes", Cambridge Press (1986)
비선형 상관식	Levenberg-Marquardt Method	W.H.Press, B.P.Fannery, S.A. Teukolsky, W. T. Vetterling, "Numerical Recipes", Cambridge Press (1986)
	Maximum-Likelihood Estimation Method	T.F. Anderson, D.S. Abrams and E.A. Grens II, "Evaluation of Parameters for Nonlinear Thermodynamic Models", AIChE J., 24(1), 20-29, 1978

선형 상관식과 비선형 상관식 모두 비선형상관식을 기준으로한 방법이 유효하므로 별도의 방법을 사용하는 것 보다 비선형 상관식 회귀분석 방법인 Levenberg-Marquardt 방법을 동일하게 적용이 가능하므로 본 평가절차에서는 이 방법만을 다루기로 한다.

3.1 Levenberg-Marquardt 알고리즘

다음의 식으로 표시된 함수를 회귀분석 하고자 할 때,

$$y = y(x;a)$$

다음과 같은 merit function을 비선형 회귀 분석을 통하여 최소화(minimization) 한다.

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - y(x_i;a)}{\sigma_i} \right]^2$$

여기에서, σ_i 는 i 번째 측정의 불확도를 나타낸다.

반복계산 과정은 다음과 같다.

- (1) 매개변수의 추정치(initial estimate) 를 제공한다. 비선형 정도가 큰 모델인 경우에는 수렴여부가 추정치의 선택에 많은 영향을 받을 수 있다. 초기 추정치는 선형 분석을 결과를 활용하거나 측정값으로부터 얻어진 그래프의 모양으로 추정할 수 있다.
- (2) 초기 추정치 값으로부터 merit function의 값을 계산한다.
- (3) Levenberg-Marquardt 방법을 활용하여 모델이 자료를 잘 표현할 수 있도록 매개변수의 값을 향상시킨다. 모델들은 매개변수에 대한 analytical derivative 값이 필요하다. 만일, analytical derivative를 얻기 어려운 모델인 경우에는 수치 해석적 방법으로 미분을 제공하여야 한다.
- (4) 개선된 매개변수로 새로운 merit function을 계산하여 이전 반복계산의 결과와 비교한다.
- (5) 3,4 과정을 merit function의 변화가 거의 없을 때 까지 반복 계산한다.
- (6) Goodness of Fit Statistics를 계산한다.

3.2 회귀분석의 Goodness-of-fit statistics

비선형 회귀분석의 Goodness of Fit는 다음의 양들을 계산하여 분석할 수 있다.

■ Fitted Parameter

회귀 분석을 통하여 얻어진 매개변수를 $\hat{\beta}_j$ 로 나타낸다.

■ Predicted Value

회귀 분석을 통하여 얻어진 매개변수 값에서 종속 변수 y 의 계산된 값을 다음과 같이 표시할 수 있다.

$$\hat{y} = f(x, \hat{\beta}_j)$$

i 번째 계산된 값은 다음과 같이 표시한다. \hat{y}_i

■ Residual

계산된 값과 예측된 값의 차를 나타낸다.

- i-번째 residual = $(y_i - \hat{y}_i)$

■ Residual Sum of Squares

계산된 값과 예측된 값의 차의 제곱을 나타낸다.

$$\text{SSE} = \text{Residual or Error Sum of Squares (absolute)} = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

자료의 불확도가 고려된 경우에는 각 자료의 Weight를 다음의 식과 같이 정규화 (Normalization)하여 상대적인 SSE값을 구한다.

$$W_i = \frac{1.0}{\sigma_i^2}, \quad \sum_{i=1}^N W_i = N$$

$$\text{SSE} = \text{Residual or Sum of Squares (relative)} = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \times W_i$$

■ Variance Analysis

\bar{y}_i 를 측정의 평균값이라고 할 때, Variance Analysis를 위한 각 항들은 다음과 같이 계산된다.

$$SSR = \text{Regression Sum of Squares} = \sum_{i=1}^N (\hat{y}_i - \bar{y}_i)^2 \times W_i$$

$$SSE = \text{Residual or Sum of Squares (relative)} = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \times W_i$$

$$SST = \text{Total Sum of Squares} = \sum_{i=1}^N (y_i - \bar{y}_i)^2 \times W_i$$

$$SST = SSE + SSR$$

- SST는 측정된 값의 전체 variation을 나타낸다.
- SSR은 “설명이 가능한” variation을 나타낸다.
- SSE는 “설명이 불가능한” variation을 나타낸다.

Coefficient of Multiple Determination (R^2) 값은 다음과 같이 나타낼 수 있다.

$$R^2 = \frac{SSR}{SST} = 1.0 - \frac{SSE}{SST}, \quad (0 \leq R^2 \leq 1)$$

만일, R^2 값이 0.95이면 95%의 variation을 모델을 이용하여 설명할 수 있다는 의미가 될 것이다. R^2 값이 1에 가까울수록 모든 point가 회귀 분석된 모델식으로 표현 가능하다는 의미가 되며, 0에 가까울수록 모델은 전혀 의미가 없으며 단순 평균으로 표현하는 것이 나올 수 있다는 의미가 된다.

Variance Analysis Table은 다음과 같이 표현된다.

Source	DOF	Sum of Squares	Mean Squares	F-ratio
Regression	$(p - 1)$	SSR	$SSR / (n - 1)$	$\frac{SSR / (p - 1)}{SSE / (n - p)}$
Error	$(n - p)$	SSE	$SSE / (n - p)$	
Total	$(n - 1)$	SST		

위 테이블은 null hypothesis $H_0 : \beta_1 = \beta_1 = \dots = \beta_p = 0$ (모든 매개변수가 0인 가설) 와 적어도 한 개의 매개변수는 0이 아니라는 가설 $H_1 : at least one \beta_j \neq 0$ (최소한 한 개의 매개변수는 0이 아니라는 가설)중 하나를 받아들이는 근거로 활용할 수 있다. 만일 계산된 F-ratio 값이 특정 F 함수값 ($F_{\alpha, p-1, n-p-1}$) 보다 크면 H_0 를 기각하고 H_1 을 받아들인다.

3.3 Chi-squared 회귀 분석의 신뢰도 구간 계산 방법

- (1) 실험 자료의 개수는 N, 매개변수의 수는 M으로 표시한다.
- (2) 관심의 매개변수의 개수를 ν 로 표시하며, $\nu \leq M$ 이다.
- (3) 원하는 신뢰도 구간을 p로 표시한다. (예를 들어 p=0.68 또는 p=0.95 등)
- (4) 주어진 자유도에서 chi-squared 변수가 $\Delta\chi^2$ 보다 작을 확률이 p가 되는 $\Delta\chi^2$ 값을 찾는다. 이 값은 다음의 테이블 또는 컴퓨터 프로그램을 찾을 수 있다.

p	ν					
	1	2	3	4	5	6
68.3 %	1.00	2.30	3.53	4.72	5.89	7.04
90 %	2.71	4.61	6.25	7.78	9.24	10.6
95.4 %	4.00	6.17	8.02	9.70	11.3	12.8
99 %	6.63	9.21	11.3	13.3	15.1	16.8
99.73 %	9.00	11.8	14.2	16.3	18.2	20.1
99.99%	15.1	18.4	21.1	23.5	25.7	27.8

5) Chi-square fit의 covariance matrix를 $[C] = [\alpha]^{-1}$ 을 취한다. diagonal element를 취한 matrix를 $[C_{proj}]$ 라고 부른다.

6) 원하는 confidence limit의 ν 차원의 subspace는 다음의 식과 같다.

$$\Delta\chi^2 = \delta a' [C_{proj}] \delta a'$$

여기에서 $\delta a'$ 는 ν 차원의 벡터이다.

따라서 매개변수의 신뢰도 구간은 다음과 같이 표시될 수 있다.

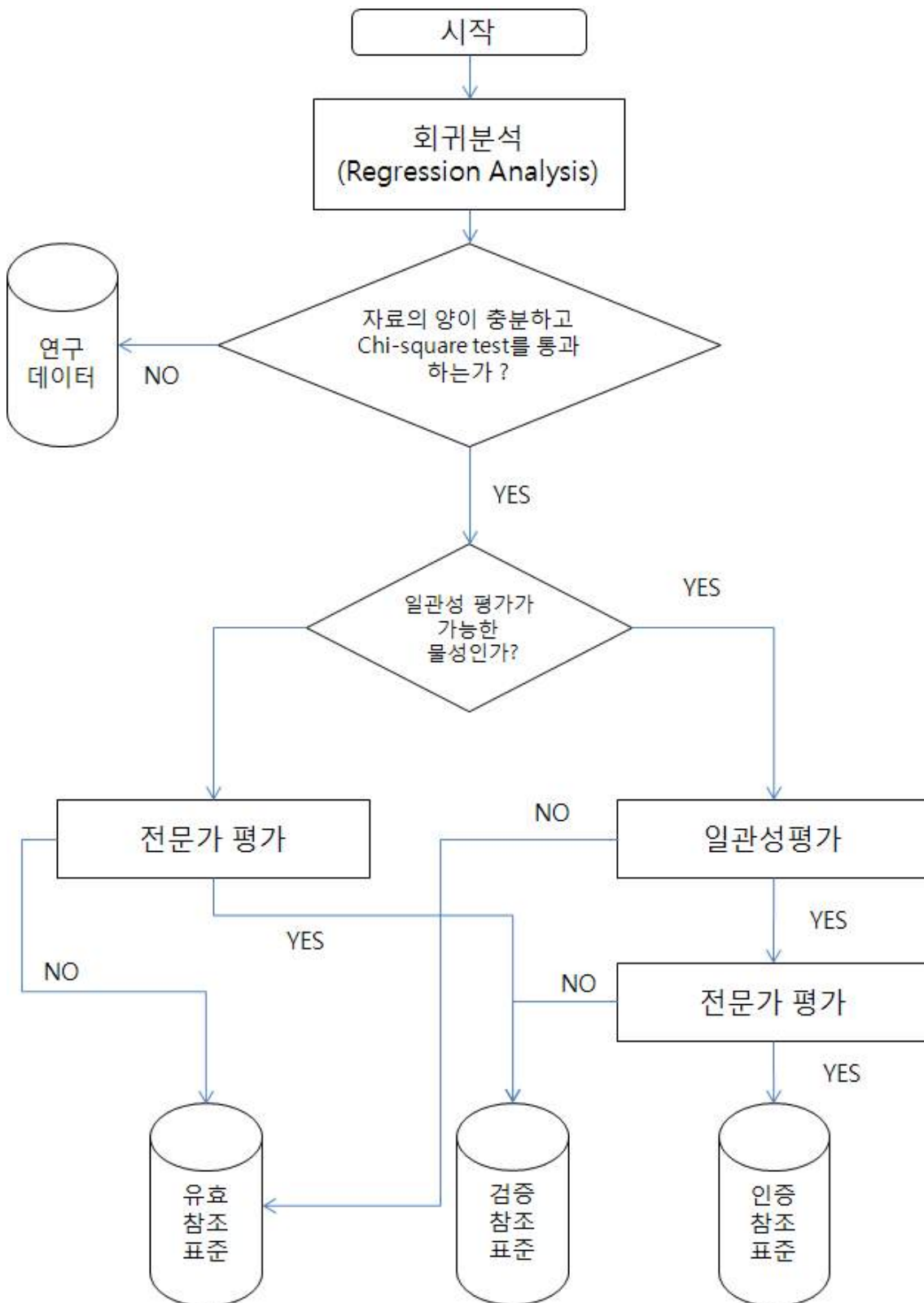
$$\delta\beta_j = \pm \sqrt{\Delta\chi^2_\nu} \sqrt{C_{jj}}$$

4. 변수의존 물성의 일관성 평가

선형 또는 비선형상관식으로 상변수에 의존하는 물성자료의 일관성 평가과정은 다음과 같다.

- (1) 최소한 2개 이상의 다른 연구자에 의한 실험 세트가 존재하여야 참조표준으로 등록이 가능함. 각 세트는 변수 (온도, 압력 및 조성) 구간에서 최소 10개 이상의 자료가 존재하여야 함. 이 조건이 만족되지 않을 경우, 연구 데이터 등급을 부여함.
- (2) 최소한 2개 이상의 실험자료 세트를 적절한 표준 모델식 또는 대안식으로 회귀 분석한 후 chi-square test를 통과한 경우 유효참조표준 등급을 부여함. (신뢰도 구간 99%).
- (3) 물성별로 일치성을 확인할 방법이 있는 경우 (열역학적 건전성 검증) 일치성 확인 절차를 통과한 경우 검증 참조 표준으로 선정함. 검증 참조표준은 전문가 평가를 통하여 인증 참조표준으로 등록될 수 있음.
- (4) 물성별로 일치성을 확인할 방법이 없는 경우, 전문가 평가를 통하여 검증 참조표준의 등급을 최대 부여함. 일치성 확인이 불가능한 물성의 최대 부여 등급은 검증 참조표준임.
- (5) 추정된 매개변수의 불확도와 계산값의 불확도는 Chi-square test의 결과를 활용한다.

전체적인 과정은 다음의 그림에 표시하였다.



부록 3. 최대우도를 이용한 매개변수 추정 방법

최대우도를 이용하여 매개변수를 추정하는 방법은 다음의 문헌에 소개된 방법을 활용한다.
(T.F. Anderson, D.S. Abrams and E.A. Grens II, "Evaluation of Parameters for Nonlinear Thermodynamic Models", AIChE J., 24(1), 20-29, 1978)

열역학 모델에 이 방법을 적용한 사례는 다음의 문헌에 제시되어 있다.

(J. M. Prausnitz, T.F. Anderson, E. A. Grens, C. A. Eckert, J. P. O'Connell, "Computer Calculation for Multi-component Vapor-Liquid Equilibria", Prentice-Hall, 1980)

상기 논문들에 소개된 방법은 다변수 열역학 모델의 매개변수 예측에 가장 적절한 방법으로 알려져 있으며, 측정된 모든 변수들이 Random Error를 가진다고 가정하고 Maximum Likelihood Principle에 근거하여 매개변수와 Covariance Matrix 및 매개변수의 불확도와 상관식의 계산값 불확도 레벨을 구하는 방법이다.

열역학 모델의 매개변수를 구하기 위하여 총 M 개의 실험을 수행하였다고 할 경우, 각 i 번째 실험에서 M 개의 측정된 변수를 가진다고 할 때, \mathbf{x}_i 는 측정된 변수의 벡터를 나타내며 θ 는 매개변수의 벡터를 나타낸다. 컴퓨터 프로그램이 수행하는 계산은 다음의 Constraint Equation를 만족시켜야 한다.

$$F(\mathbf{x}_i, \theta) = 0 \quad \text{constraint equation 1}$$

$$G(\mathbf{x}_i, \theta) = 0 \quad \text{constraint equation 2}$$

Constraint Equation은 단일 물성인 경우 선형 상관식 또는 비선형 상관식으로 표시되는 Merit Function과 동일할 수도 있지만 기-액 상평형의 Bubble Point Pressure 계산의 경우에는 계산된 압력과 계산된 기체 조성이 일치하여야 된다는 2개의 Constraint를 가질 수도 있다.

Maximum likelihood method에 의하여 매개변수를 구하는 방법은 최초로 Fisher (1922)에 의하여 도입되었으며, Bard과 Lepidus(1968)에 의하여 상세히 기술되었으나 이 방법들은 의존변수 (dependent variable)이 불확도를 가질 수 있다는 가정 하에 개발되어 왔다. 의존변수 뿐 만이 아니라 독립 변수들도 불확도를 가질 수 있다는 가정을 도입하면, 변수들의 확률밀도함수가 필요하다. 측정된 변수들의 확률밀도함수가 정규분포를 따른다는 가정을 도입하면, Morrison (1967)이 제시한 Multi-variate joint probability density function을 이용

할 수 있다.

$$\Phi_i(\mathbf{x}_i^m, \mathbf{x}_i) = \frac{1}{(2\pi)^{1/2M} |\boldsymbol{\beta}_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_i^m - \mathbf{x}_i)^T \boldsymbol{\beta}_i^{-1}(\mathbf{x}_i^m - \mathbf{x}_i)\right\}$$

여기에서 ϕ_i 는 i 번째 실험에서 실제값이 \mathbf{x}_i 이고, 측정값으로 \mathbf{x}_i^m 을 얻을 확률이다. β 는 MxM variance-covariance matrix이다.

문제를 정리하면 다음의 식을 최적화하는 것과 동일해 진다.

$$S = \sum_{i=1}^N (\mathbf{x}_i^m - \mathbf{x}_i)^T \boldsymbol{\beta}_i^{-1} (\mathbf{x}_i^m - \mathbf{x}_i)$$

Constraint equation 1,2를 푼 방정식을 다음과 같이 표현하면,

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}, \boldsymbol{\theta})$$

$$\mathbf{Z} = \mathbf{g}(\mathbf{X}, \boldsymbol{\theta})$$

상기 Equation을 정리하면,

$$S = (\mathbf{X} - \mathbf{X}^m)^T \boldsymbol{\lambda} (\mathbf{X} - \mathbf{X}^m) + (\mathbf{Y} - \mathbf{Y}^m)^T \boldsymbol{\gamma} (\mathbf{Y} - \mathbf{Y}^m) + (\mathbf{Z} - \mathbf{Z}^m)^T \boldsymbol{\delta} (\mathbf{Z} - \mathbf{Z}^m)$$

S를 최적화하기 위한 조건은 다음과 같다.

$$\frac{\partial S}{\partial \mathbf{X}} = 0 \quad \frac{\partial S}{\partial \boldsymbol{\theta}} = 0$$

각 반복계산에서 매개변수와 변수들의 값을 변경시키는 수식은 다음과 같다.

$$\Delta \boldsymbol{\theta} = (\boldsymbol{\theta}^{r+1} - \boldsymbol{\theta}^r) = -[\mathbf{T} - \mathbf{R}^T \mathbf{D}^{-1} \mathbf{R}]^{-1} [\mathbf{U} - \mathbf{R}^T \mathbf{D}^{-1} \mathbf{Q}]$$

$$\Delta \mathbf{X} = (\mathbf{X}^{r+1} - \mathbf{X}^r) = -\mathbf{D}^{-1} [\mathbf{Q} + \mathbf{R} \Delta \boldsymbol{\theta}]$$

각 행렬들은 다음과 같이 정의된다.

$$\mathbf{D} = \boldsymbol{\lambda} + \mathbf{f}_x^T \boldsymbol{\gamma} \mathbf{f}_x + \mathbf{g}_x^T \boldsymbol{\delta} \mathbf{g}_x$$

$$\mathbf{R} = \mathbf{f}_x^T \boldsymbol{\gamma} \mathbf{f}_0 + \mathbf{g}_x^T \boldsymbol{\delta} \mathbf{g}_0$$

$$\mathbf{T} = \mathbf{f}_0^T \boldsymbol{\gamma} \mathbf{f}_0 + \mathbf{g}_0^T \boldsymbol{\delta} \mathbf{g}_0$$

$$\mathbf{U} = \mathbf{f}_0^T \boldsymbol{\gamma} \Delta \mathbf{Y}^m + \mathbf{g}_0^T \boldsymbol{\delta} \Delta \mathbf{Z}^m$$

