united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

Course on "Inverse Methods in Atmospheric Science"
1 - 12 October 2001

301/1332-11

"Inversion of High Spectral Resolution Radiance from Satellite Infrared Sensors"

C. SERIO
Italian Institute for Condensed Matter (INFM)
Potenza, Italy

Please note: These are preliminary notes intended for internal distribution only.

strada costiera, 11 - 34014 trieste italy - tel. +39 0402240111 fax +39 040224163 - sci_info@ictp.trieste.it - www.ictp.trieste.it

Inversion of High Spectral Resolution Radiance from Satellite Infrared Sensors

Carmine Serio

Italian Institute of Condesed Matter, (INFM)

Gruppo Collegato di Potenza, Unità di Napoli,

C.da Macchia Romana, 85100 Potenza, It

serio@unibas.it

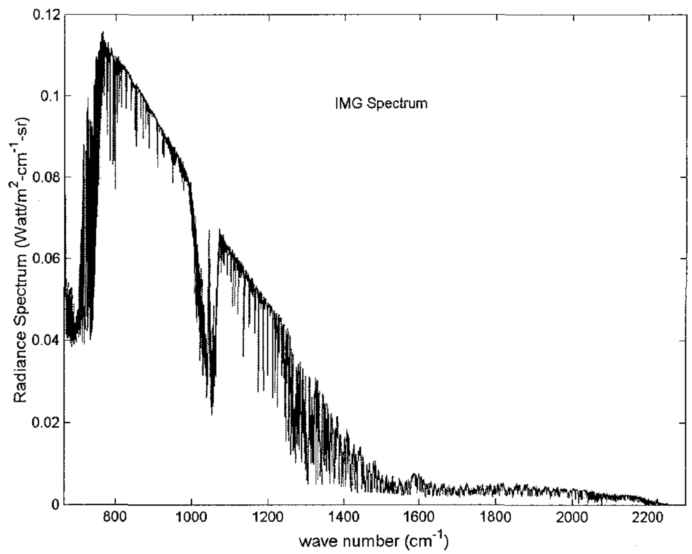
Thanks to U. Amato, I. DeFeis, A.M. Lubrano, G. Masiello and M. Viggiano

Satellite Missions (Planned and Future)

• ADEOS (IMG)	(FTS)
• ENVISAT (MIPAS)	(FTS)
• METOP/1 (IASI)	(FTS)
• EOS Aqua (AIRS)	(Grating Spectrometer)
• EarthCare (IMG follow on?)	(FTS)
• CrIS	(FTS)
• GIFT (geostationary orbit)	(Imaging FTS)
• TES	(FTS)
• REFIR	(FTS)

Key Scientific Issue

Retrieval of Atmospheric Parameters



The need of high spectral resolution

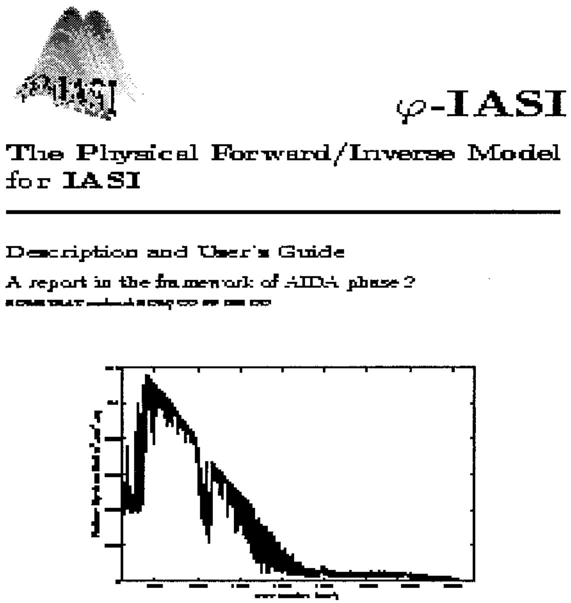
(Kaplan, Chahine, Susskind, and Searl *Appl. Opt*. Vol. 16, No. 2, (1977), 322-325)

- **Numerical Weather Prediction**
  - **Temperature profile with accuracy of 1 K in 1 Km layers in the troposphere**
  - **Water vapour profile with accuracy of 10-20% in 1-2 km layers in the lower troposphere**

- **Climatology**
  - **Total Columnar amount of O_3, CO, CO_2, CH_4, N_2O, CFCs with accuracy within 2-3%**
  - **Chemistry of the upper atmosphere, ozone depletion, and ozone killers**

New Tools to Process High Spectral Resolution Spectra

- Fast Line-by-Line Forward Models
- Physical Inversion Schemes


φ-IASI

The Physical Forward/Inverse Model for IASI

Description and User's Guide

A report in the framework of AIDA phase 2

Outline of the lecture
(First part)

- ***Forward modelling: Basics***
  1) - Monochromatic spectral radiance calculation
  2) - Jacobian calculation
- ***Forward modelling: Applications***
  1) - Checking the linearization of the radiative transfer equation
  2) – Retrieval error analysis

Radiative Transfer along a slant path

$$\frac{dR(\sigma, s)}{ds} = -K(\sigma, s)\,[R(\sigma, s) - B(\sigma, T(s))]\,,$$

$R(\sigma)$ [Watt/m^2-sr-cm^{-1}] σ [cm^{-1}]

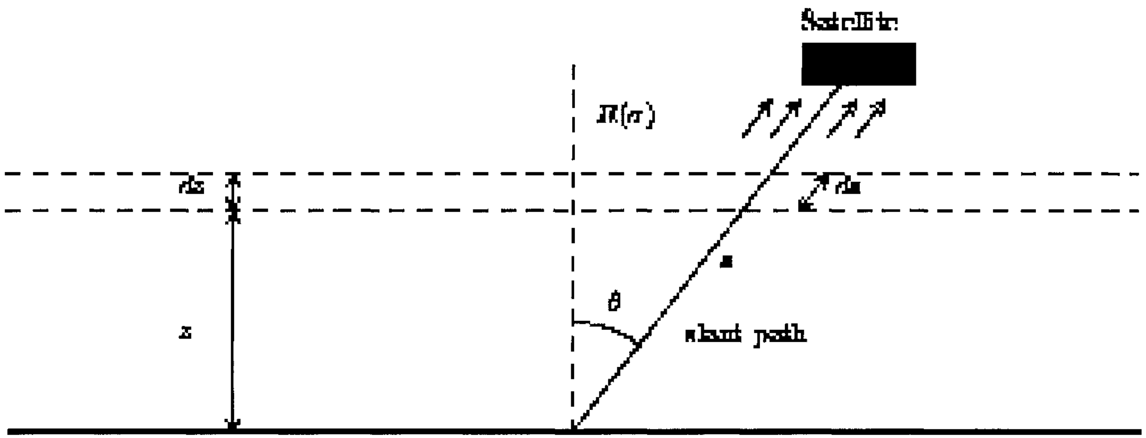
s slant path coordinate

T(s) temperature at *s;*   ***B*** is the Planck Function

$$K(\sigma, s) = k(\sigma, s)\rho(s)$$

$k(\sigma, s)$ [cm^2/gr] $\rho(s)$ [gr/cm^3]

Slant path geometry



$$\mu \frac{dR(\sigma, z)}{dz} = -K(\sigma, z)\left[R(\sigma, z) - B(\sigma, T(z))\right],$$

$$\mu = \cos(\theta)$$

Radiative transfer Equation (upwelling term)

$$R(\sigma) = \varepsilon_g B(T_g)\tau_o + \int_0^{+\infty} B(T)\frac{\partial \tau}{\partial z}dz,$$

τ is the transmittance from the z altitude level to ∞

ε_g is the surface emissivity

T_g the skin temperature

$$\exp\left(\int_{z'}^{z} -K(\sigma, z'')\frac{dz''}{\mu}\right) = \tau(z', z; \sigma, \mu),$$

Radiative transfer Equation (downwelling term)

$$R(\sigma) = \int_{+\infty}^{0} B(T) \frac{\partial \tau^*}{\partial z} dz,$$

$$\tau(z;\sigma,\mu)\tau^*(z;\sigma,\mu) = \tau_o(\sigma,\mu),$$

$$R(\sigma) = -\tau_o \int_0^{+\infty} B(T) \frac{\partial}{\partial z}\left(\frac{1}{\tau}\right) dz.$$

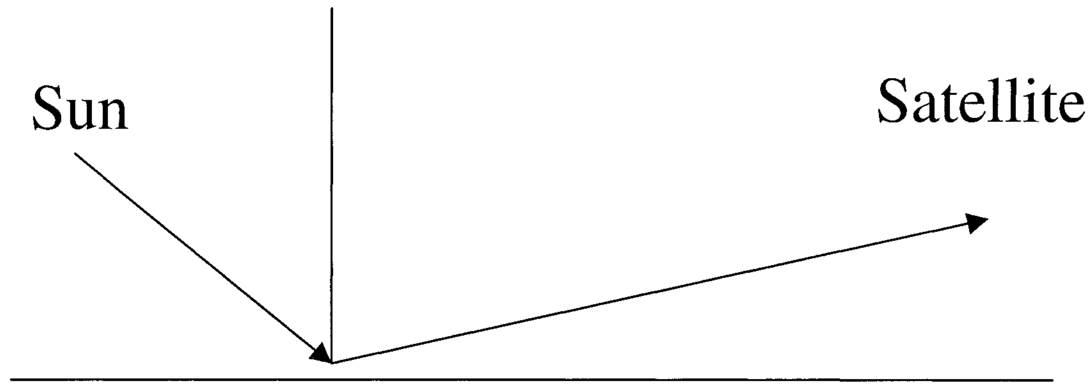
Upwelling Spectral Radiance (Earth Contribution)

$$
\begin{aligned}
R(\sigma) &= \varepsilon_g B(T_g)\tau_o + \int_0^{+\infty} B(T)\frac{\partial \tau}{\partial z} dz \\
&+(\varepsilon_g - 1)\tau_o^2 \int_0^{+\infty} B(T)\frac{\partial}{\partial z}\left(\frac{1}{\tau}\right) dz.
\end{aligned}
$$

This makes the downwelling term of Second Order

Solar Contribution

$$R(\sigma) = \frac{1-\varepsilon_g}{\pi}\tau' \mu_s I_s(\sigma),$$



Geometry for the definition of the two path transmittance

Upwelling Spectral Radiance (Sun+Earth)

$$
\begin{aligned}
R(\sigma) &= \varepsilon_g B(T_g)\tau_o + \int_0^{+\infty} B(T)\frac{\partial \tau}{\partial z} dz \\
&+(\varepsilon_g - 1)\tau_o^2 \int_0^{+\infty} B(T)\frac{\partial}{\partial z}\left(\frac{1}{\tau}\right) dz \\
&+\frac{1-\varepsilon_g}{\pi}\tau' \mu_s I_s(\sigma).
\end{aligned}
$$

Remark 1: The solar term is important above 2000 cm^{-1}

Remark 2: In the form above the radiative transfer equation is suitable only for nadir looking sensors.

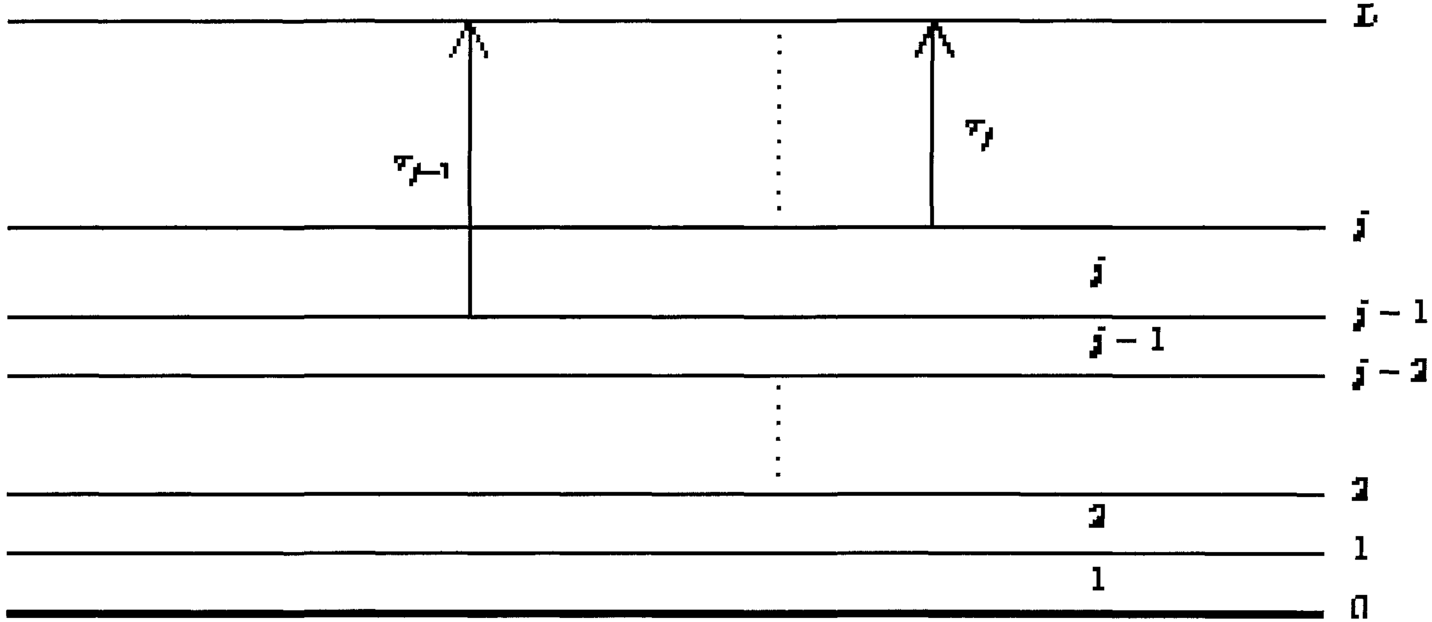
Extension to cloudy sky

$$R(\sigma) = (1-\alpha)\cdot R_o(\sigma) + \alpha \cdot R_c(\sigma);$$

$$R_c(\sigma) = \varepsilon_c B(T_c)\tau_c + \int_{h_c}^{+\infty} B(T)\frac{\partial \tau}{\partial z}dz,$$

α is the cloud fractional amount within the sensor Field of View

Definitions of layers, levels and transmittances (plane parallel atmosphere)



Layer	Pressure (hPa)	Layer	Pressure (hPa)	Layer	Pressure (hPa)
1	1013:25 – 1005:43	16	436:95 – 396:81	31	45:29 – 35:51
2	1005:43 – 985:88	17	396:81 – 358:28	32	35:51 – 27:26
3	985:88 – 957:44	18	358:28 – 321:50	33	27:26 – 20:40
4	957:44 – 922:46	19	321:50 – 286:60	34	20:40 – 14:81
5	922:46 – 882:80	20	286:60 – 253:71	35	14:81 – 10:37
6	882:80 – 839:95	21	253:71 – 222:94	36	10:37 – 6:95
7	839:95 – 795:09	22	222:94 – 194:36	37	6:95 – 4:41
8	795:09 – 749:12	23	194:36 – 167:95	38	4:41 – 2:61
9	749:12 – 702:73	24	167:95 – 143:84	39	2:61 – 1:42
10	702:73 – 656:43	25	143:84 – 122:04	40	1:42 – 0:69
11	656:43 – 610:60	26	122:04 – 102:05	41	0:69 – 0:29
12	610:60 – 565:54	27	102:05 – 85:18	42	0:29 – 0:10
13	565:54 – 521:46	28	85:18 – 69:97	43	0:10 – 0:005
14	521:45 – 478:54	29	69:97 – 56:73		
15	478:54 – 436:95	30	56:73 – 45:29		

Table 1

Radiance Calculation

$$R(\sigma) \approx \varepsilon_g B(T_g)\tau_o + \sum_{j=1}^{L} B(T_j)(\tau_j - \tau_{j-1})$$
$$+(\varepsilon_g - 1)\tau_o^2 \sum_{j=1}^{L} B(T_j)(\tau_j^{-1} - \tau_{j-1}^{-1}) + \frac{(1-\varepsilon_g)}{\pi}\tau' \mu_s I_s(\sigma),$$

Radiance Calculation

$$S^{+} = \sum_{j=1}^{L} B(T_j)(\tau_j - \tau_{j-1}) \quad \text{and} \quad S^{-} = \sum_{j=1}^{L} B(T_j)(\tau_j^{-1} - \tau_{j-1}^{-1}), \quad (14)$$

$$R(\sigma) \approx \varepsilon_g B(T_g)\tau_o + S^{+} + (\varepsilon_g - 1)\tau_o^2 S^{-} + \frac{(1-\varepsilon_g)}{\pi}\tau' \mu_s I_s(\sigma).$$

Definitions of geophysical parameters and radiance vector

- Surface Spectral emissivity, ε
- Skin Temperature, T_g
- Temperature profile, $T(z)$, $(T_1,\ldots,T_N)$
- Gases, $q(z)$, $(q_1,\ldots,q_N)$
- $\mathbf{v}=((T_1,\ldots,T_N,\ q_1,\ldots,q_N,\ T_g,\varepsilon)$, atmosphere state vector
- $R(\sigma)=(R_1,\ldots,R_M)=\mathbf{r}$, radiance vector

$$\mathbf{r} = F(\mathbf{v}); \qquad \frac{\partial F}{\partial(\mathbf{v})} \equiv \text{Jacobian}$$

Derivatives with respect to Emissivity and Skin Temperature

$$\frac{\partial R(\sigma)}{\partial \varepsilon_g} = B(T_g)\tau_o + \tau_o^2 S^- - \frac{\mu_s \tau' I_s}{\pi}.$$

$$\frac{\partial R(\sigma)}{\partial T_g} = \varepsilon_g \tau_o \frac{\partial B}{\partial T_g}, \qquad \frac{\partial B}{\partial T} = \frac{c_1 c_2 \sigma^4 \exp(c_2\sigma/T)}{T^2[\exp(c_2\sigma/T) - 1]^2}.$$

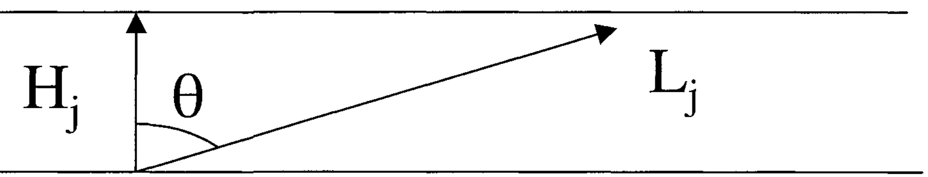
Temperature Jacobian

$$
\begin{aligned}
\frac{\partial R}{\partial T_j} &= D_j^0 + D_j^+ + D_j^-, \\
D_j^0 &= -\left[\varepsilon_g \tau_o B(T_g) + 2(\varepsilon_g - 1)\tau_o^2 S^- + \frac{1-\varepsilon_g}{\pi}\mu_s \tau'(1+\frac{\mu_s}{\mu}) I_s\right] \frac{\partial \nu_j}{\partial T_j} \\
D_j^+ &= \left[\sum_{k=1}^{j} \tau_{k-1}(B(T_k) - B(T_{k-1}))\right] \frac{\partial \nu_j}{\partial T_j} + \frac{\partial B}{\partial T_j}(\tau_j - \tau_{j-1}) \qquad (23) \\
D_j^- &= (\varepsilon_g - 1)\tau_o^2 \left\{ -\left[\sum_{k=1}^{j} \tau_{k-1}^{-1}(B(T_k) - B(T_{k-1}))\right] \frac{\partial \nu_j}{\partial T_j} + \frac{\partial B}{\partial T_j}(\tau_j^{-1} - \tau_{j-1}^{-1}) \right\}
\end{aligned}
$$

$$\nu_j = \sum_h \nu_j^{(h)}$$ Total Optical Depth for the layer j

$$\nu_j^{(h)} = k_j^{(h)} q_j^{(h)} L_j ,$$

$H_j = L_j \cos(\theta)$



$q_j^{(h)}$ is the layer concentration for the species *h*

$k_j^{(h)}$ is the absorption coeffiicient for the species *h*

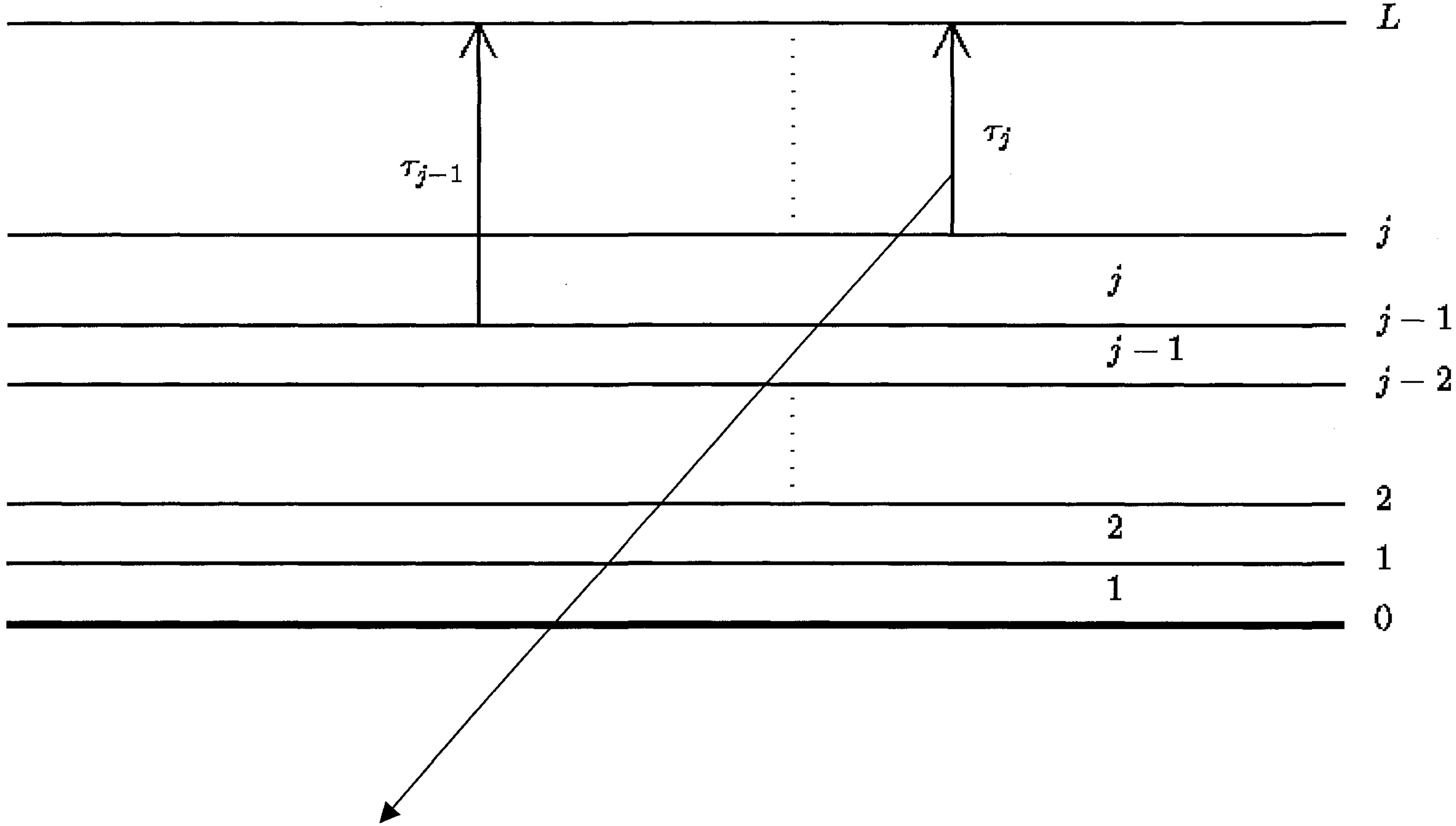
Relation between Optical Depth and transmittance

$$\eta_j^{(h)} = \exp(-\nu_j^{(h)})$$

Layer Transmittance for the Individual species

$$\eta_j = \prod_h \eta_j^{(h)}$$

Total Layer Transmittance for All species



$$\tau_j^{(h)} = \prod_{l=j+1}^{L} \eta_l^{(h)}, \quad \longrightarrow \quad \tau_j = \prod_{h=1}^{N_{\mathrm{GAS}}} \tau_j^{(h)},$$

Gas concentration Jacobian

$$\frac{\partial R(\sigma)}{\partial q_j^{(h)}} = G_j^0 + G_j^+ + G_j^-,$$

$$G_j^0 = -\left[\varepsilon_g \tau_o B(T_g) + 2(\varepsilon_g - 1)\tau_o^2 S^- + \frac{1-\varepsilon_g}{\pi}\mu_s \tau'(1+\frac{\mu_s}{\mu})I_s\right]\frac{\nu_j^{(h)}}{q_j^{(h)}}$$

$$G_j^+ = \left[\sum_{k=1}^{j}\tau_{k-1}(B(T_k) - B(T_{k-1}))\right]\frac{\nu_j^{(h)}}{q_j^{(h)}} \qquad (29)$$

$$G_j^- = -(\varepsilon_g - 1)\tau_o^2\left[\sum_{k=1}^{j}\tau_{k-1}^{-1}(B(T_k) - B(T_{k-1}))\right]\frac{\nu_j^{(h)}}{q_j^{(h)}}.$$

Convolving with the Instrumental Line Shape

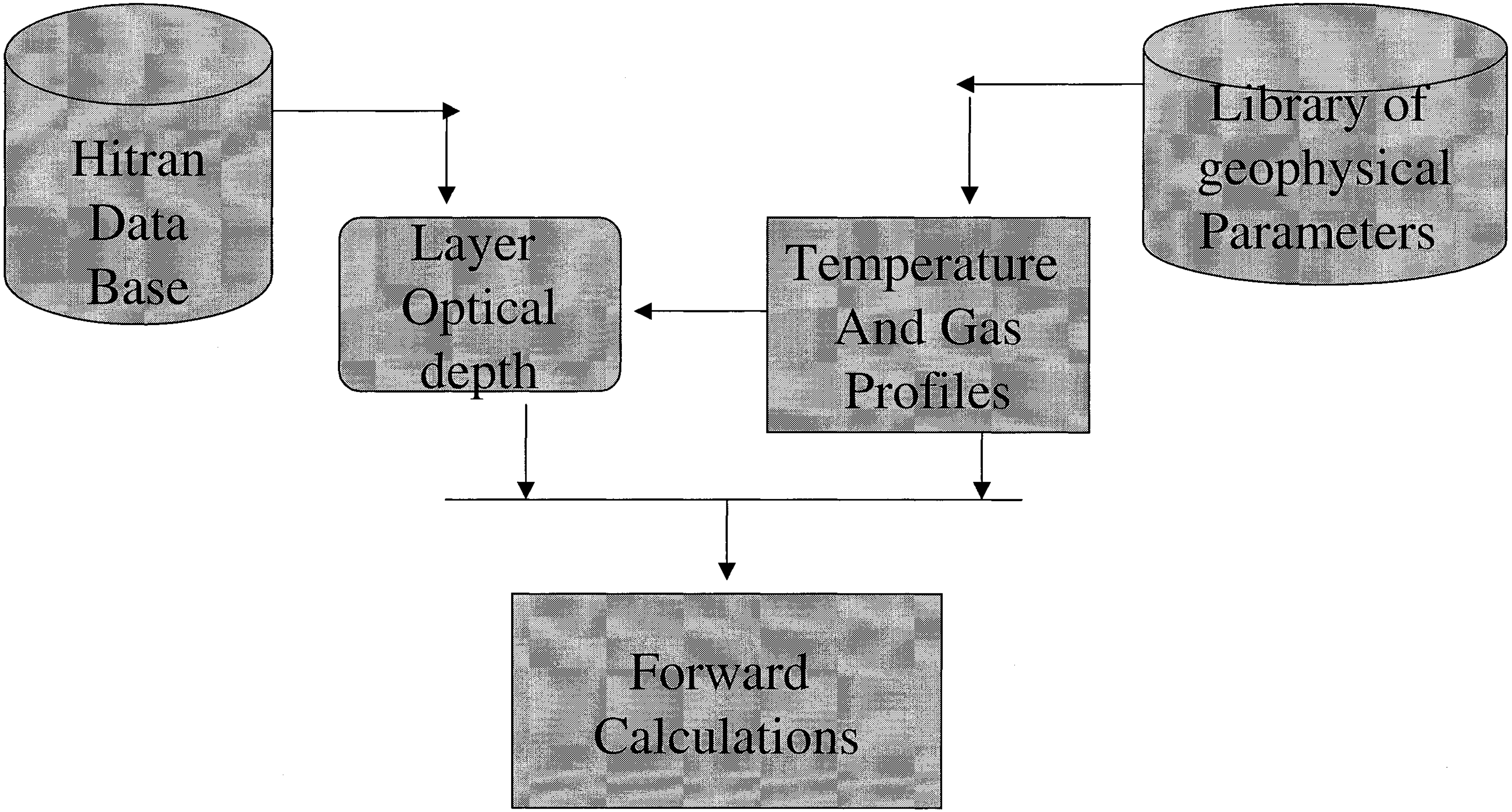
$$\tilde{R}(\sigma) = \int_{-\infty}^{+\infty} f(\sigma, \sigma_o) R(\sigma_o) d\sigma_o$$

Synthetic Radiance

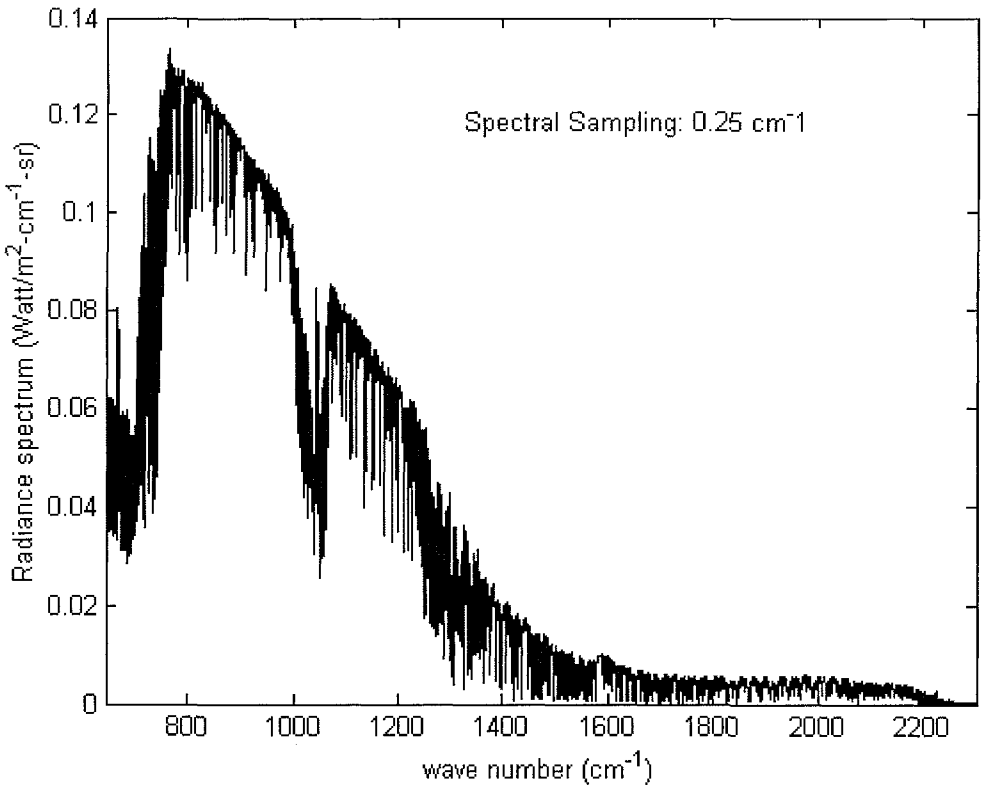
Instrumental Line Shape

Monochromatic Radiance

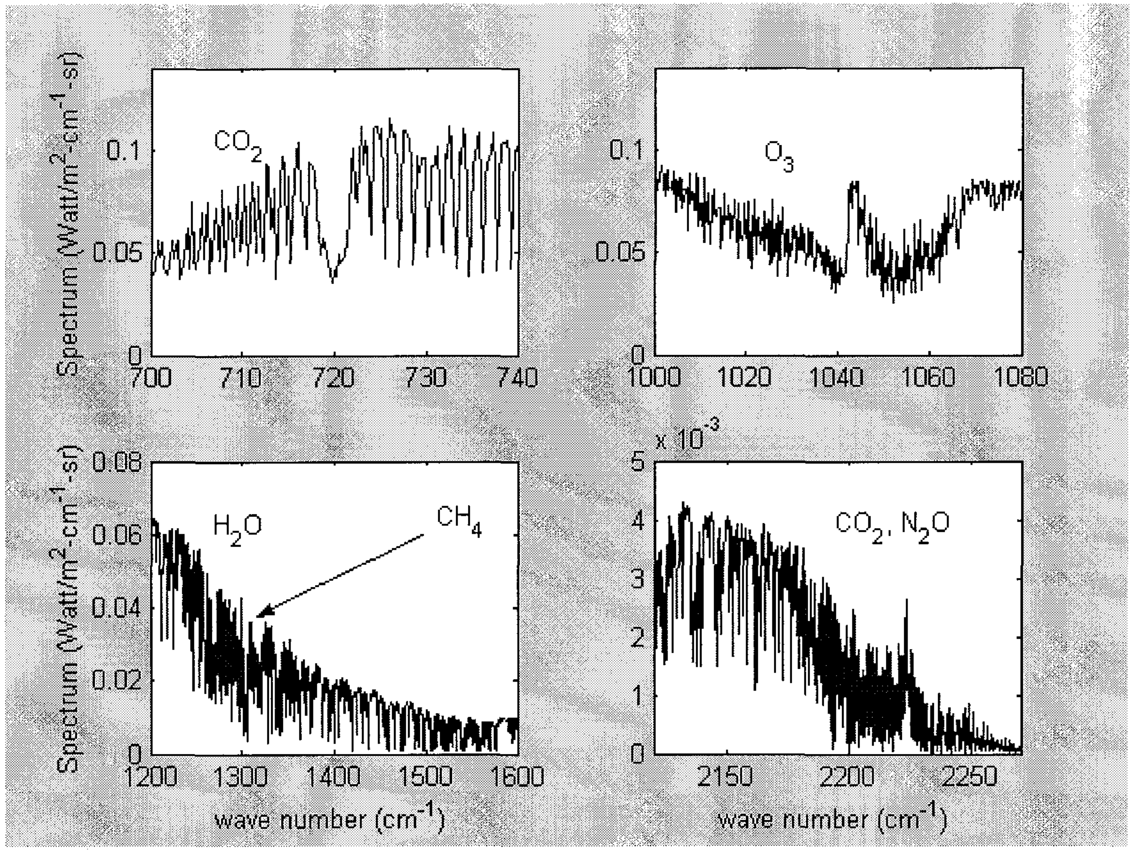
Basic Elements of Forward Calculations



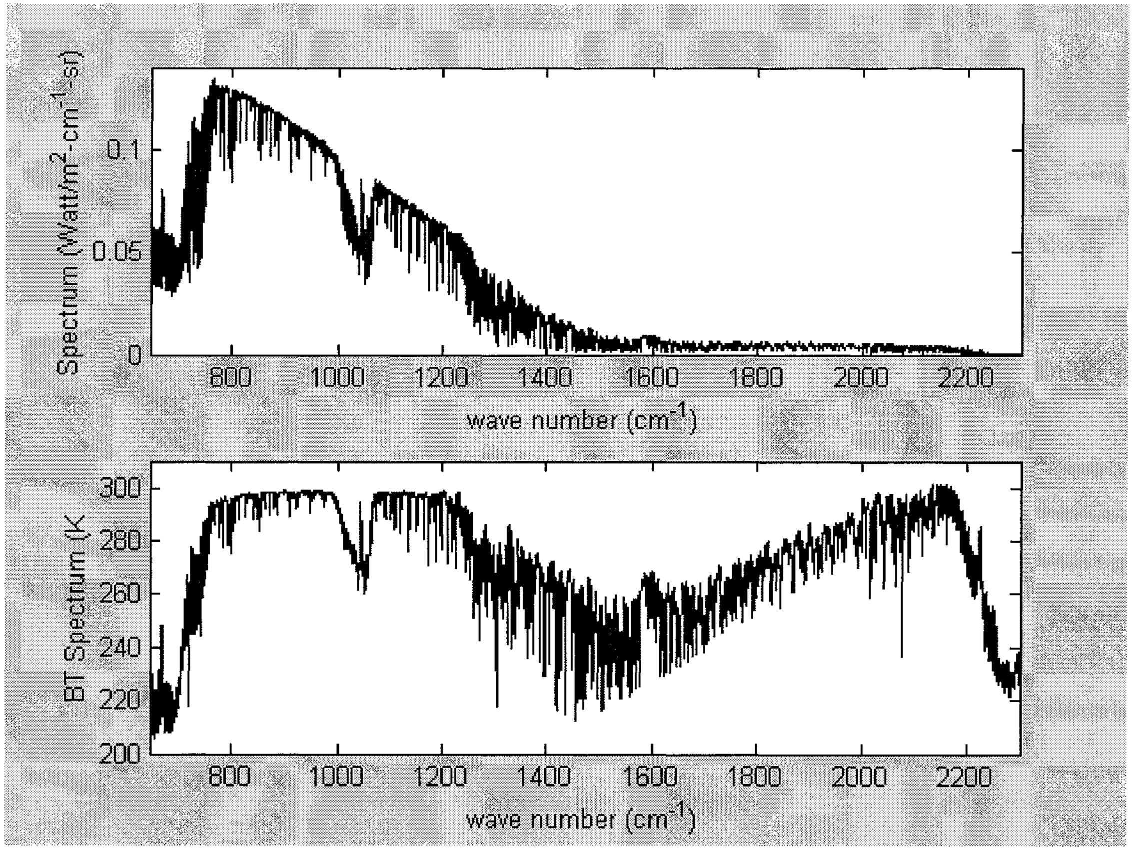
Example of Synthetic Radiance Spectra



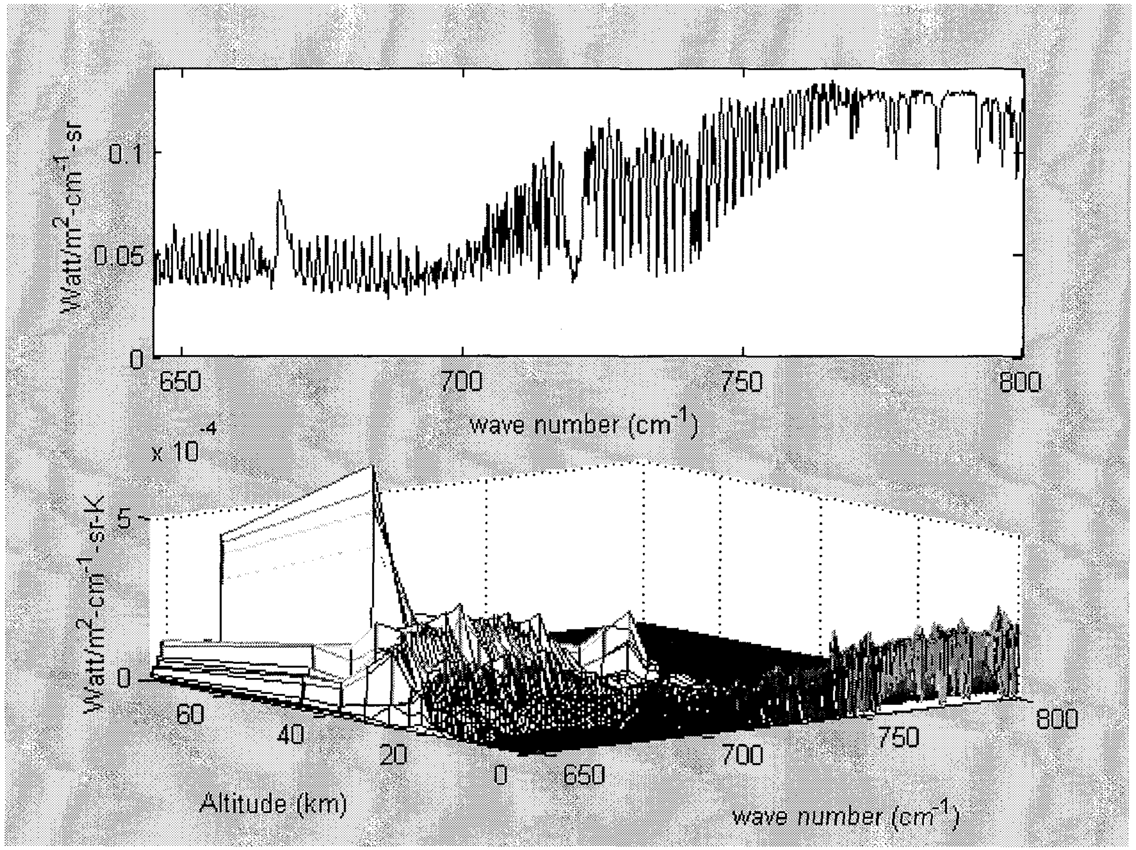
Zooming specific absorption bands



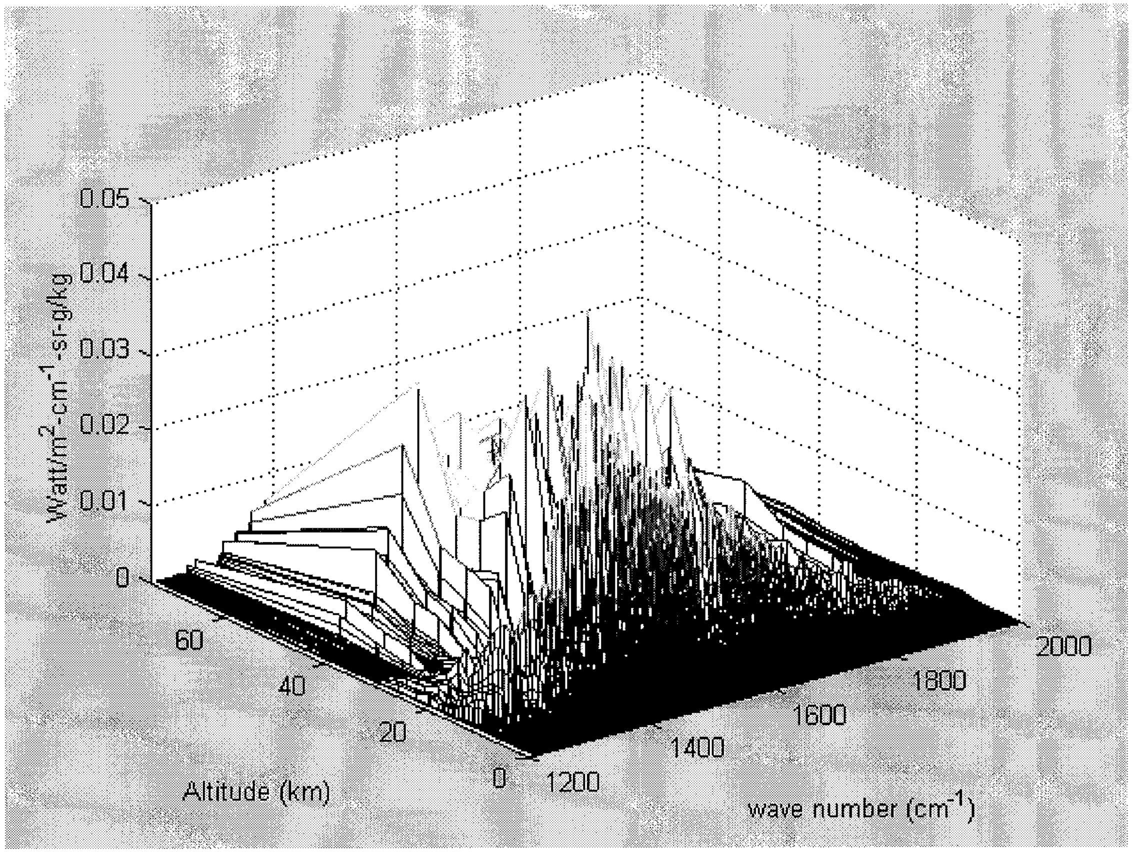
$$T = \frac{c_2 \sigma}{\log\left(\frac{c_1 \sigma^3}{R(\sigma)} + 1\right)}$$



Example of Jacobian Temperature



Example of Jacobian
Water Vapour



Assessing the linearity of RTE

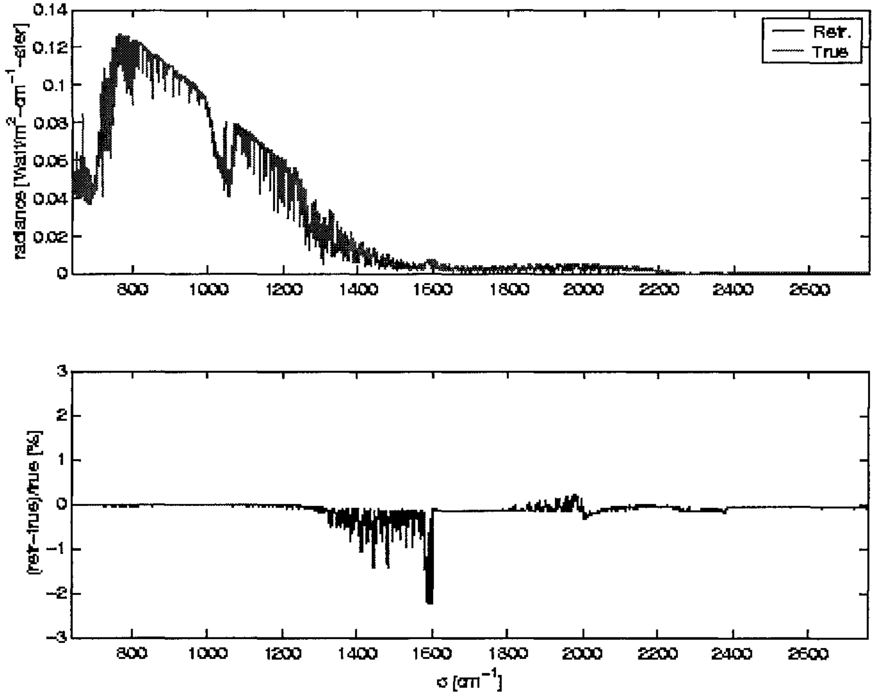
$$\mathbf{r}^{(lin)} = \mathbf{r}_o + \mathbf{K}_T(\mathbf{T} - \mathbf{T}_o) + \mathbf{K}_w(\mathbf{w} - \mathbf{w}_o) =$$

$$\mathbf{r}_o + \mathbf{K}_T\delta\mathbf{T} + \mathbf{K}_w\delta\mathbf{w}; \text{First Order Taylor Expansion}$$

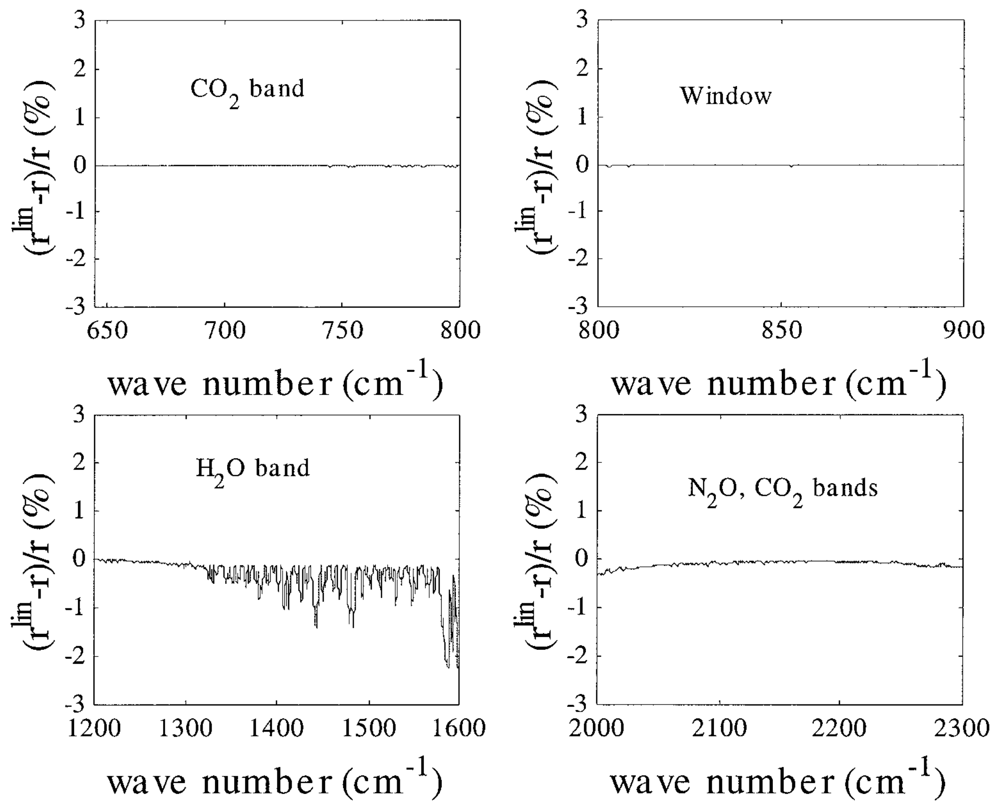
$$\mathbf{r} = F(\mathbf{T}_o + \delta\mathbf{T}, \mathbf{w}_o + \delta\mathbf{w});$$

$$\delta\mathbf{r} = \mathbf{r}^{(lin)} - \mathbf{r} = \begin{bmatrix} 0; & \text{linear} \\ \neq 0; & \text{non linear} \end{bmatrix}$$

Example for a Tropical Atmosphere
$\delta \mathbf{T}$=1 K; $\delta \mathbf{w}$=10%



Pseudo Noise



Expected Retrieval Performance

$\mathbf{C} = (\mathbf{K}^t\mathbf{S}^{-1}\mathbf{K} + \mathbf{B}^{-1})^{-1}$ Covariance Matrix

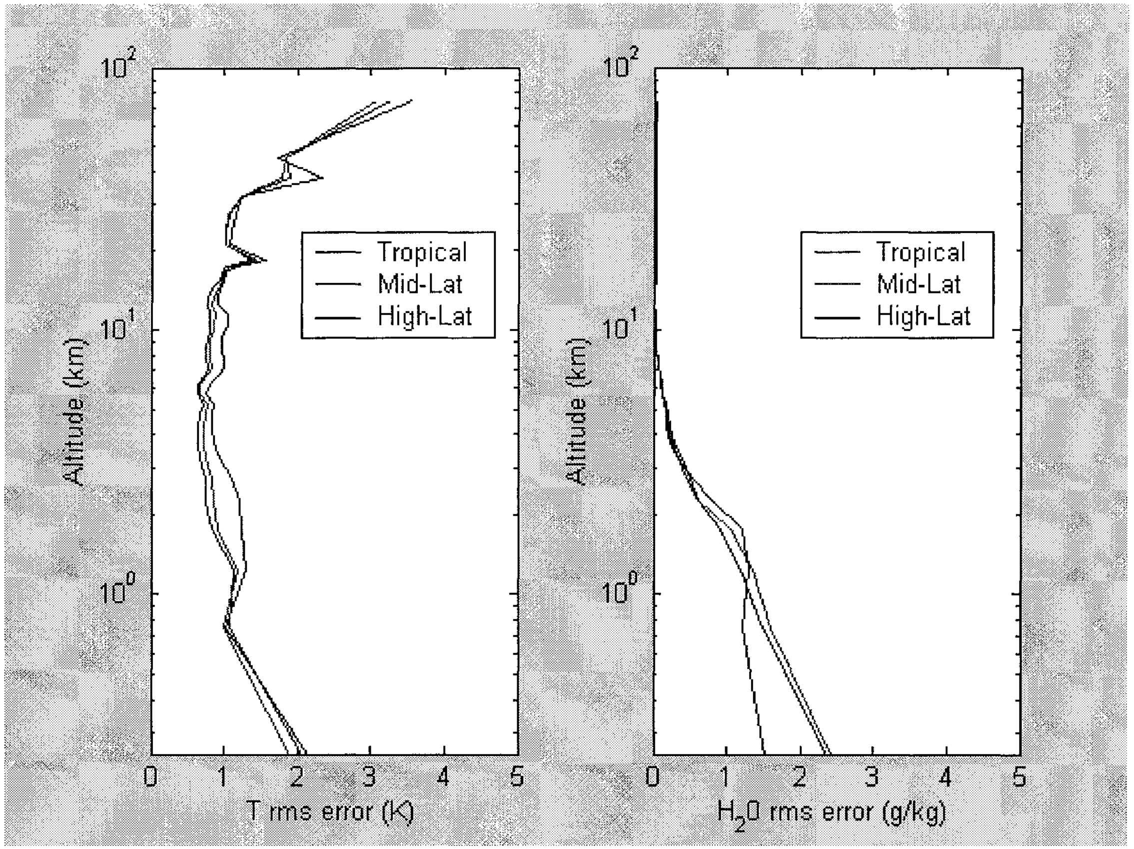
$\mathbf{K}$: is the Jacobian

$\mathbf{S}$: is the Observational Covariance Matrix (depends on Instrument design)

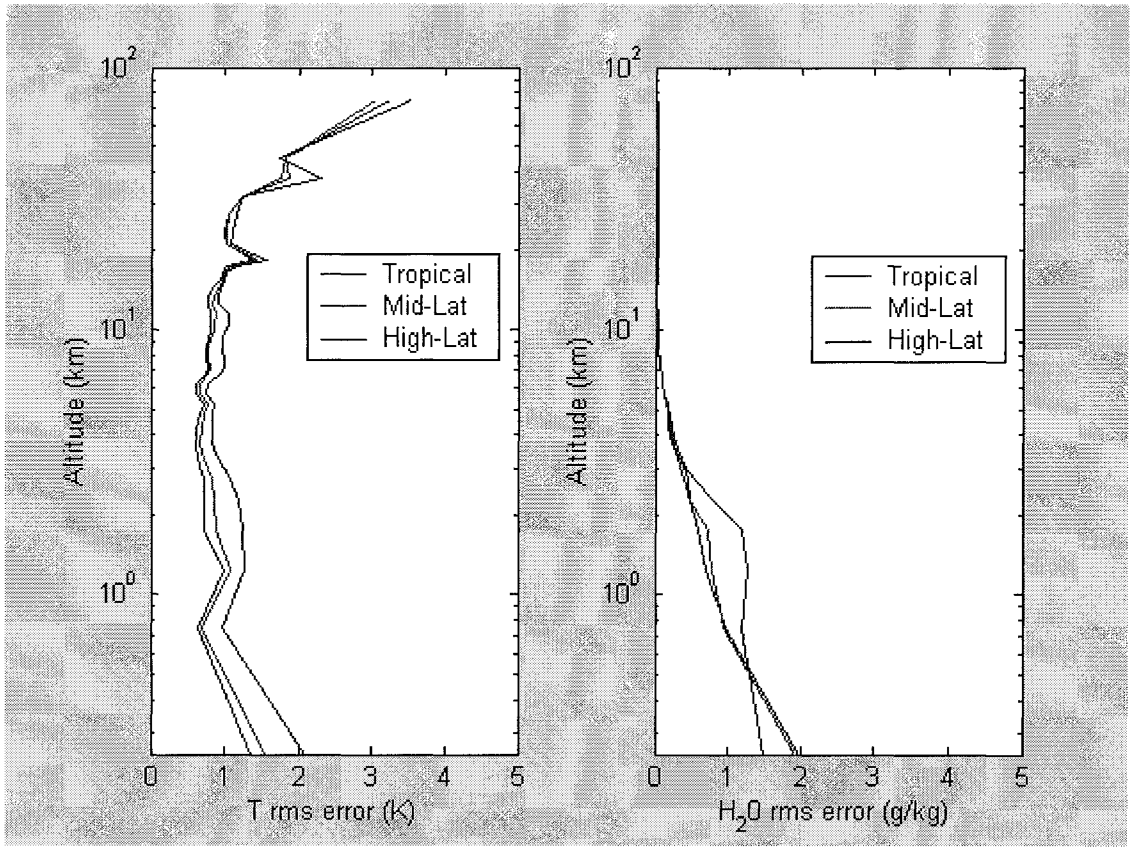
$\mathbf{B}$: is the a-priori Covariance Matrix

Root Mean Square Error: $\sqrt{\mathbf{diag}(\mathbf{C})}$

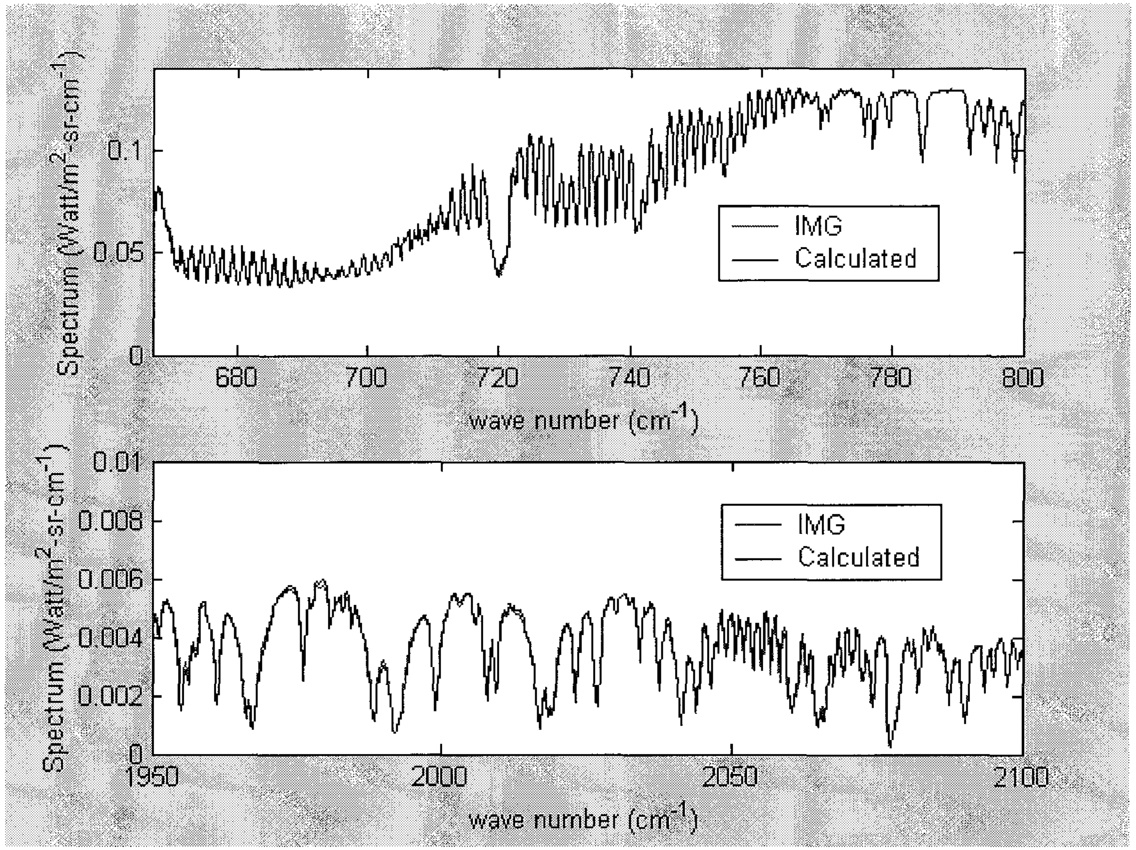
Error Analysis for Simultaneous Retrieval of Temperature and Water Vapour Band 645 to 830 cm^{-1}



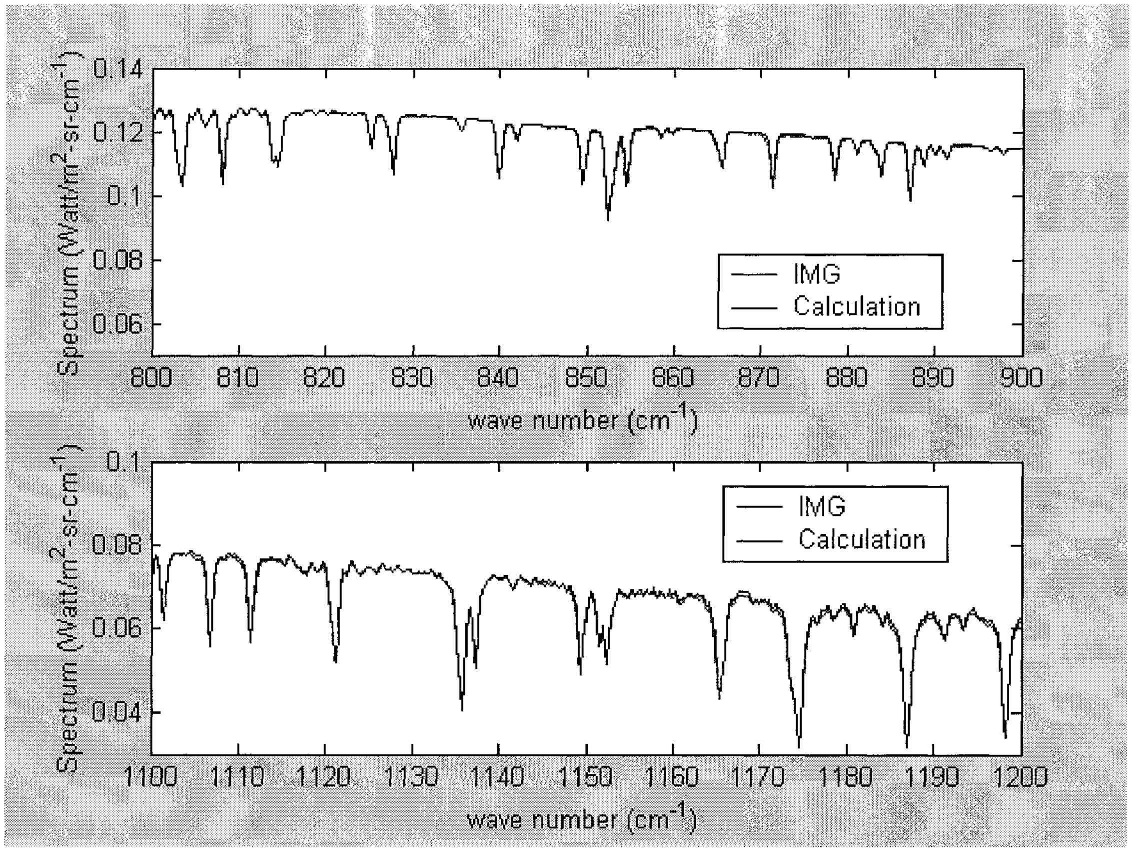
Adding the contribution of IASI band 3
645 to 830 cm^{-1} + 2180 to 2250 cm^{-1}



How consistent we are with reality: Checking calculation against IMG



Atmospheric Window
Checking calculation against IMG



Second Part:
The Inversion Process

- Basics on regularization techniques
- Link to Ridge Regression
- Examples and applications to IMG spectra

Where do we start from?

Expand RTE in Taylor series around a suitable FG atmospheric state

$$
\begin{aligned}
R(\sigma) &= R(\sigma)^{\text{f.g.}} + \sum_{k=1}^{L} \left. \frac{\partial R(\sigma)}{\partial T_k} \right|_{f.g.} \left(T_k - T_k^{\text{f.g.}} \right) \\
&+ \sum_{h=1}^{N_{\text{GAS}}} \sum_{k=1}^{L} \left. \frac{\partial R(\sigma)}{\partial q_k^{(h)}} \right|_{f.g.} \left(q_k^{(h)} - q_k^{(h)\text{f.g.}} \right),
\end{aligned}
$$

Let F and $\mathbf{v}_o$ be the forward model function and the FG state of the atmosphere, respectively

$$\mathbf{R} = \mathbf{R}_0 + \mathbf{K}(\mathbf{v} - \mathbf{v}_0) + \mathrm{H.O.T.},$$

$$\mathbf{R}^t = (R_1, \ldots, R_N)$$

$$\mathbf{v}^t = (v_1, \ldots, v_M),$$

$\mathbf{K}$ is the N by M Jacobian matrix, $\mathbf{K} = \frac{\partial F}{\partial \mathbf{v}}\Big|_{\mathbf{v}=\mathbf{v}_o}$

$$\mathbf{R}_0 = F(\mathbf{v}_0)$$

The straightforward inversion makes no sense: the problem is ill-posed and ill-conditioned

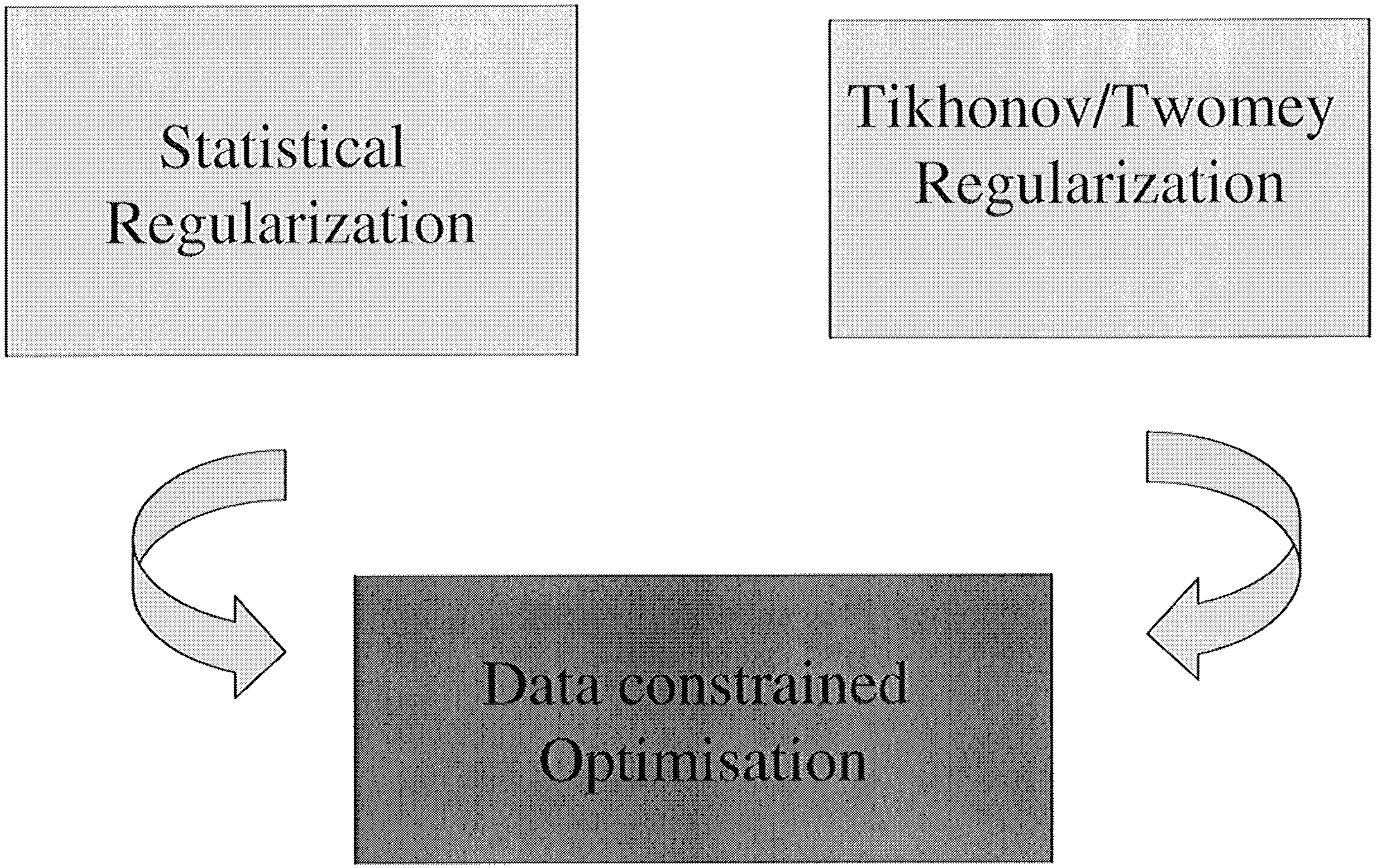
$$\mathbf{x} = \mathbf{v} - \mathbf{v}_0$$

$$\mathbf{y} = \mathbf{R} - \mathbf{R}_0$$

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \text{H.O.T..}$$

$$\downarrow$$

$$\hat{\mathbf{x}} = (\mathbf{K}^t\mathbf{K})^{-1}\mathbf{K}^t\mathbf{y}$$


Statistical Regularization
Tikhonov/Twomey Regularization
Data constrained Optimisation

The subscript g now stands for a suitable background atmospheric state!

$$(\hat{\mathbf{v}} - \mathbf{v}_g)^t \mathbf{L} (\hat{\mathbf{v}} - \mathbf{v}_g) \quad \text{min!}$$

$$(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}})^t \mathbf{S}^{-1} (\mathbf{y} - \mathbf{K}\hat{\mathbf{x}}) = N,$$

$\mathbf{S}^{-}$ Observational Covariance Matrix

$\mathbf{L}$ Smoothing Operator

About **L**

- Twomey's approach

$$\mathbf{L} \equiv \int_0^{+} \left| \frac{d^n \hat{x}}{dh^n} \right|^2 dh;$$

- n=0, 1, 2

- Rodgers' approach

- **L** is intrinsically a covariance operator, **B** in his notation

Let us consider the effect over the L-norm below

$$(\hat{\mathbf{v}} - \mathbf{v}_g)^t \mathbf{L} (\hat{\mathbf{v}} - \mathbf{v}_g)$$

Twomey's L is lacking dimensional consistency! it attempts to Sum unlike quantity, e,g, (K+g/kg)

Rodgers' **L** ensures dimensional consistency

$\mathbf{L}=\mathbf{B}^{-1}$

which makes the norm above dimensionless

Our Choice

- Since we are interested in simultaneous inversion which involves unlike quantities such as Temperature, water vapour concentration and so on we choose
  - $\mathbf{L=B^{-1}}$
- However, the methodology we are going to discuss still hold for any Twomey's **L**

Resorting to Lagrange multiplier minimization

Note: $\gamma=1/\lambda$

$$\hat{\mathbf{v}}-\mathbf{v}_g=(\mathbf{v}_0-\mathbf{v}_g)+\hat{\mathbf{x}}= \quad \mathbf{q}_0+\hat{\mathbf{x}},$$

$$\frac{d}{d\hat{\mathbf{x}}}\left\{(\hat{\mathbf{x}}+\mathbf{q}_0)^t\mathbf{B}^{-1}(\hat{\mathbf{x}}+\mathbf{q}_0)\right.$$
$$\left.-\lambda[(\mathbf{y}-\mathbf{K}\hat{\mathbf{x}})^t\mathbf{S}^{-1}(\mathbf{y}-\mathbf{K}\hat{\mathbf{x}})-N]\right\}=0.$$

$$\hat{\mathbf{x}}=(\gamma\mathbf{B}^{-1}+\mathbf{K}^t\mathbf{S}^{-1}\mathbf{K})^{-1}[\mathbf{K}^t\mathbf{S}^{-1}\mathbf{y}-\gamma\mathbf{B}^{-1}\mathbf{q}_0],$$

Iterization, n=0,1,2,…

$$\hat{\mathbf{x}}_{n+1} = (\gamma_n \mathbf{B}^{-1} + \mathbf{K}_n^t \mathbf{S}^{-1} \mathbf{K}_n)^{-1} [\mathbf{K}_n^t \mathbf{S}^{-1} \mathbf{y}_n - \gamma \mathbf{B}^{-1} \mathbf{q}_n] \tag{34}$$

$$\begin{aligned}
\mathbf{r}_0 &= F(\mathbf{v}_0) \\
\mathbf{y}_0 &= \mathbf{R} - \mathbf{R}_0 \\
\mathbf{K}_0 &= \left.\frac{\partial F(\mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{v}=\mathbf{v}_0} \\
\mathbf{q}_0 &= \mathbf{v}_0 - \mathbf{v}_g .
\end{aligned}$$

Uncovering the elemental constituent of regularization

$$\mathrm{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2}.$$

The square root of **B** is easily computed by SVD

$$(\gamma_n \mathbf{B}_n^{-\frac{t}{2}} \mathbf{B}_n^{-\frac{1}{2}} + \mathbf{J}_n^t \mathbf{J}_n)\hat{\mathbf{x}}_{n+1} = \mathbf{J}_n^t \mathbf{z}_n - \gamma_n \mathbf{B}_n^{-1} \mathbf{q}_n,$$

where $\mathbf{J}_n = \mathbf{S}^{-\frac{1}{2}}\mathbf{K}_n$ and $\mathbf{z}_n = \mathbf{S}^{-\frac{1}{2}}\mathbf{y}_n$.

Continued

$$\mathbf{B}^{-\frac{t}{2}}(\gamma_n \mathbf{I} + \mathbf{B}^{\frac{t}{2}} \mathbf{J}_n^t \mathbf{J}_n \mathbf{B}^{\frac{1}{2}}) \mathbf{B}^{-\frac{1}{2}} \hat{\mathbf{x}}_{n+1} = \mathbf{J}_n^t \mathbf{z}_n - \gamma \mathbf{B}^{-1} \mathbf{q}_n. \tag{37}$$

Defining $\mathbf{G}_n = \mathbf{J}_n \mathbf{B}^{\frac{1}{2}}$ and $\hat{\mathbf{u}}_{n+1} = \mathbf{B}^{-\frac{1}{2}} \hat{\mathbf{x}}_{n+1}$, Eq. (37)

assumes the form of a *ridge* regression

$$(\gamma_n \mathbf{I} + \mathbf{G}_n^t \mathbf{G}_n) \hat{\mathbf{u}}_{n+1} = \mathbf{G}_n^t \mathbf{z}_n - \gamma_n \mathbf{B}^{-\frac{1}{2}} \mathbf{q}_n.$$

Continued

- The same decomposition may obtained for Twomey regularization by putting
    - **L=M^{t}M**
- **M** may be obtained by Cholesky decomposition for any symmetric full rank matrix **L**
- Twomey's **L** is typically singular. Nevertheless the above decomposition may be obtained by resorting to GSVD (Hansen, *SIAM Review*, Vol. 34, pp. 561, (1992))

Conclusion

- The RIDGE regression is the paradigm of any regularizing method,
- The difference between the various methods is the way they normalize the Jacobian and the value they assign to the Lagrange multiplier

Ridge regression is the paradigm of any regularizing method

- **Levenberg-Marquardt**:
- γ is assigned alternatively a small or a large value, Jacobian is not normalized, that is **L=I**
- **Twomey:**
- γ is free-parameter to be determined, Jacobian is normalized through a mathematical operator
- **Rodgers**:
- γ =1, Jacobian is normalized to the a-priori covariance matrix. It is the method which enables dimensional consistency

Having discovered the right paradigm of regularization, can we give any rationale to $\gamma = 1$?

- Let us consider the linear case

$$(\mathbf{G}^t\mathbf{G} + \gamma\mathbf{I})^{-1}\mathbf{G}^t\mathbf{z} = \hat{\mathbf{u}}$$

- Note: with **G** normalized to **B**, the above equation is fully dimensionless

Bias, Covariance Matrix, Root Mean Square Error

- The bias is defined as the difference between retrieval expectation and *true* solution. Under the assumption of Gaussian independent random error, we have

$$\mathbf{b} = E[\hat{\mathbf{u}}] - \mathbf{u} = \mathbf{Iu} - \gamma(\mathbf{G}^t\mathbf{G} + \gamma\mathbf{I})^{-1}\mathbf{u} - \mathbf{Iu} =$$
$$- \gamma(\mathbf{G}^t\mathbf{G} + \gamma\mathbf{I})^{-1}\mathbf{u};$$

$\mathbf{u}$ is the true solution

- Unless **u=0** any regularization method introduces bias

Total Bias, b_t

$$\mathbf{bb}^{\mathrm{t}} = \gamma^2 (\mathbf{G}^{\mathrm{t}}\mathbf{G} + \gamma\mathbf{I})^{-1} \mathbf{uu}^{\mathrm{t}} [(\mathbf{G}^{\mathrm{t}}\mathbf{G} + \gamma\mathbf{I})^{-1}]^{\mathrm{t}}$$

$$b_t = \mathrm{trace}(\mathbf{bb}^{\mathrm{t}}) = \sum_{j=1}^{M} \frac{\gamma^2 P_j}{(\lambda_j + \gamma)^2}$$

with λ_j the eigenvalue s of matrix $\mathbf{G}^{\mathrm{t}}\mathbf{G}$

P_j is the square of the j - th element of the vector $\mathbf{u}$

$$P_j = u_j^2;$$

The bias is zero only for $\gamma = 0$

$$\frac{\partial b_t}{\partial \gamma} = \sum_{j=1}^{M} \frac{2\gamma P_j \lambda_j^2 + 2\gamma^2 P_j \lambda_j}{(\lambda_j + \gamma)^4} = \sum_{j=1}^{M} \frac{2\gamma P_j \lambda_j}{(\lambda_j + \gamma)^3} \geq 0$$

wich is zero only and only if $\gamma = 0$

the Least Square unconstrai ned

solution is unbiased :

$$\hat{\mathbf{u}} = (\mathbf{G}^t \mathbf{G})^{-1} \mathbf{G}^t \mathbf{z}$$

Variance

$$\mathbf{V} = E[(\hat{\mathbf{u}} - E[\hat{\mathbf{u}}])]^2 = (\mathbf{G}^t\mathbf{G} + \gamma\mathbf{I})^{-1}\mathbf{G}^t\mathbf{G}(\mathbf{G}^t\mathbf{G} + \gamma\mathbf{I})^{-1};$$

The total variance is obtained by

$$s^2 = \mathrm{trace}(\mathbf{V}) = \sum_{j=1}^{M} \frac{\lambda_j}{(\lambda_j + \gamma)^2};$$

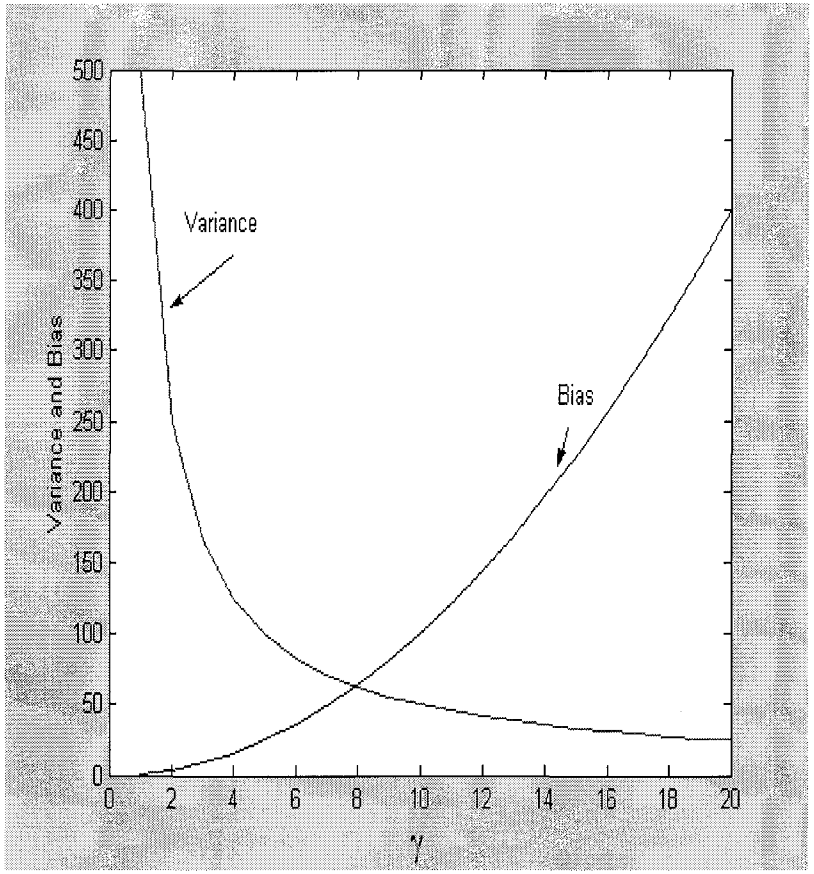
$$\frac{\partial s^2}{\partial \gamma} = \sum_{j=1}^{M} \frac{-2\lambda_j}{(\lambda_j + \gamma)^2} < 0;$$

The Mean Square Error

$$\mathbf{M} = \mathbf{V} + \mathbf{b}\mathbf{b}^{\mathbf{t}}$$

$$D^2 = \text{trace}(\mathbf{M}) = \sum_{j=1}^{M} \frac{\gamma^2 P_j + \lambda_j}{(\lambda_j + \gamma)^2}$$

$$\frac{\partial D^2}{\partial \gamma} = \sum_{j=1}^{M} \frac{\lambda_j (2\gamma P_j - 2)}{(\lambda_j + \gamma)^3}$$



The meaning of $\gamma=1$

$$\frac{\partial D^2}{\partial \gamma} = \sum_{j=1}^{M} \frac{\lambda_j (2\gamma P_j - 2)}{(\lambda_j + \gamma)^3} = 0$$

if and only if $\gamma = 1/P_j \ \forall j$

The meaning of γ=1 (continued)

- We have the true solution vector **u**
- P_j is the squared component u_j^2
- Remember that in our normalized retrieval space the *a-priori* covariance matrix is the Identity matrix and, therefore the *a-priori* variances are all equal to 1
- Since P_j is unknown our best a-priori estimate of it is
    - P_j=*a-priori j-th variance*=1

γ=1 is our best *a-priori* choice if we normalize Jacobian to the *a-priori* covariance matrix

- Nevertheless, γ=1 does not minimize the Mean Square Error, therefore to seek for an alternative choice makes sense even in the context of Statistical Regularization

L-curve method

Hansen, *SIAM Review*, Vol. 34, pp. 561, (1992)

$$\| \mathbf{B}^{-1/2}\hat{\mathbf{x}}_{n+1} \|_2 = \sqrt{\hat{\mathbf{x}}^t_{n+1}\mathbf{B}^{-1}\hat{\mathbf{x}}_{n+1}}$$

$$\| \mathbf{S}^{-1/2}(\mathbf{K}_n\hat{\mathbf{x}}_{n+1} - \mathbf{y}_n) \|_2 = \Big[(\mathbf{K}_n\hat{\mathbf{x}}_{n+1} - \mathbf{y}_n)^t \mathbf{S}^{-1}(\mathbf{K}_n\hat{\mathbf{x}}_{n+1} - \mathbf{y}_n)\Big]^{\frac{1}{2}}$$

L-curve method in our normalized retrieval space

$$
\begin{aligned}
a(\gamma) &= \sqrt{\hat{\mathbf{u}}^t \hat{\mathbf{u}}} \\
b(\gamma) &= \sqrt{(\mathbf{G}\hat{\mathbf{u}} - \mathbf{z})^t (\mathbf{G}\hat{\mathbf{u}} - \mathbf{z})},
\end{aligned}
$$

$$
c(\gamma) = \frac{|\ a'(\gamma) b''(\gamma) - a'(\gamma) b''(\gamma)\ |}{[(a'(\gamma))^2 + (b'(\gamma))^2]^{\frac{3}{2}}},
$$

Let us consider a numerical example

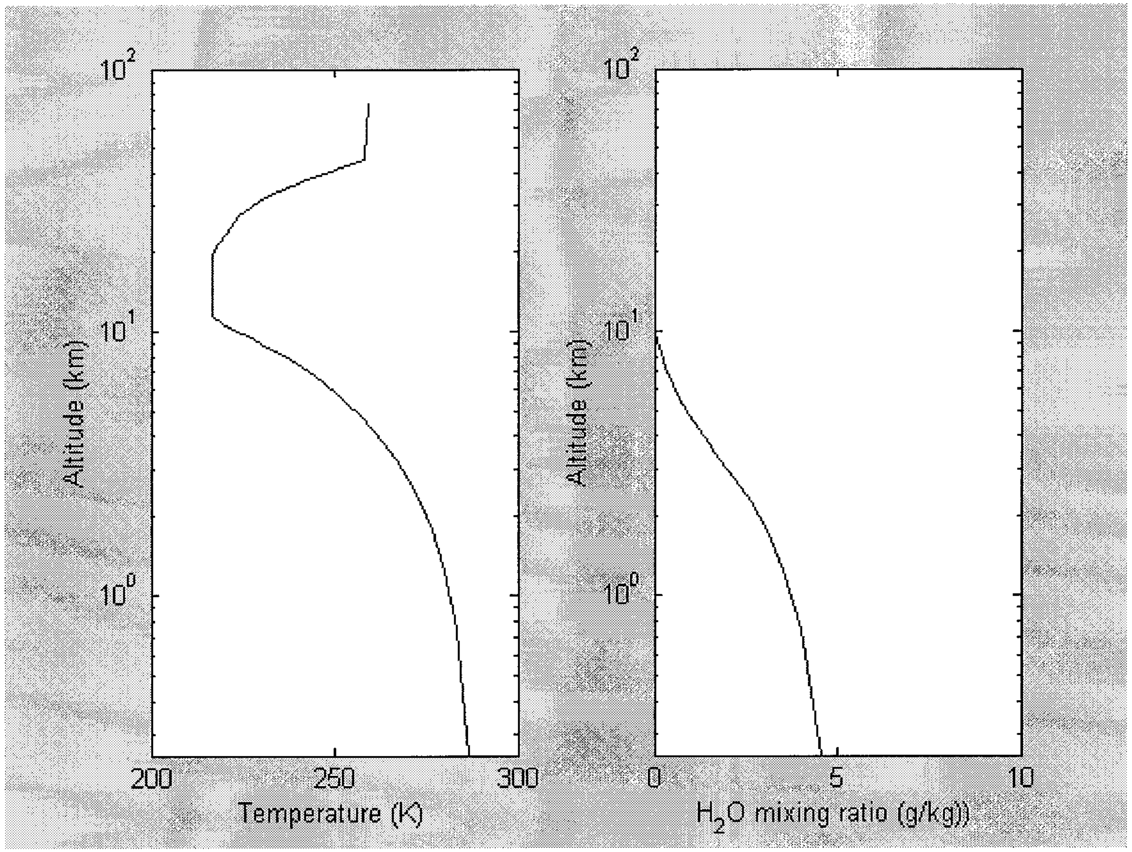
$\mathbf{v} = \mathbf{v}_o; \quad \mathbf{r} = \mathbf{r_o} + \mathbf{n};$ where $\mathbf{n}$ is an additive component random noise

$\mathbf{u} = \mathbf{0}; P_j = 0$

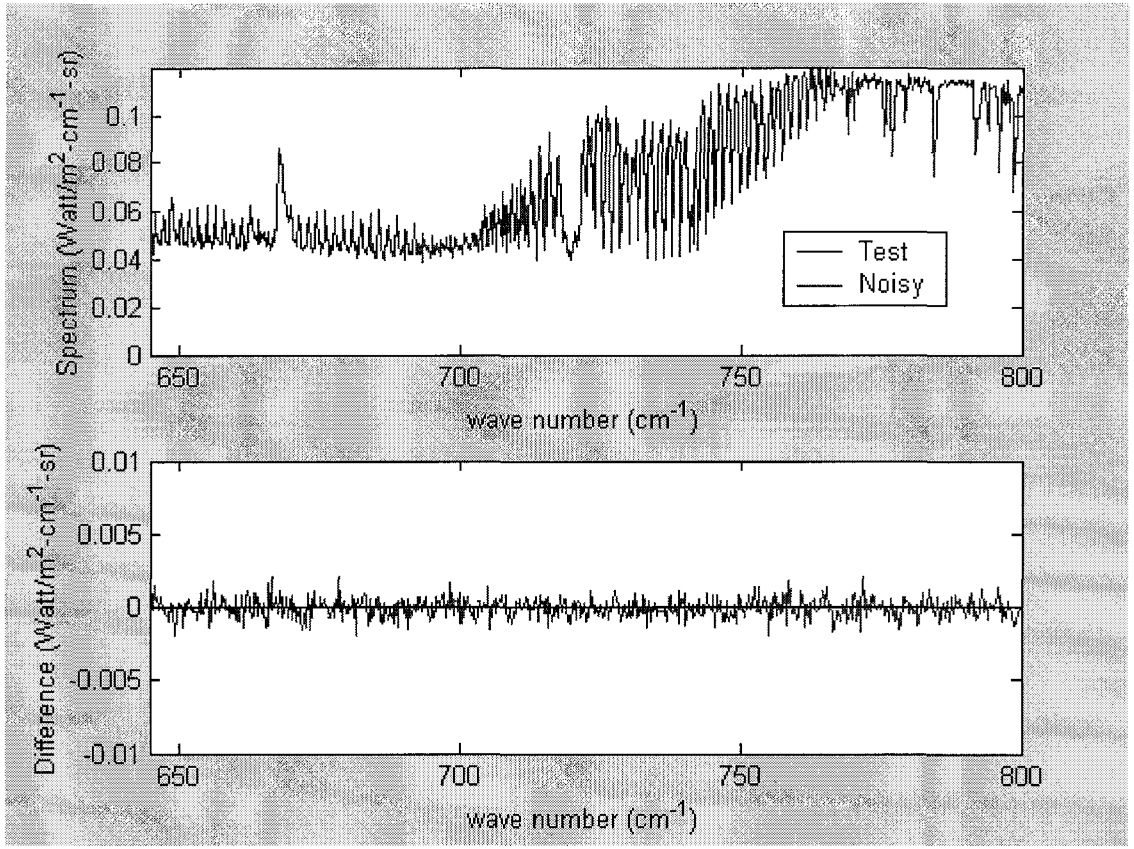
$$\mathbf{M} = \mathbf{V}; \mathrm{trace}(\mathbf{M}) = \mathrm{trace}(\mathbf{V}) = \sum_{j=1}^{M} \frac{\lambda_j}{(\lambda_j + \gamma)^2};$$

The optimal solution correspond s to $\gamma = +\infty$

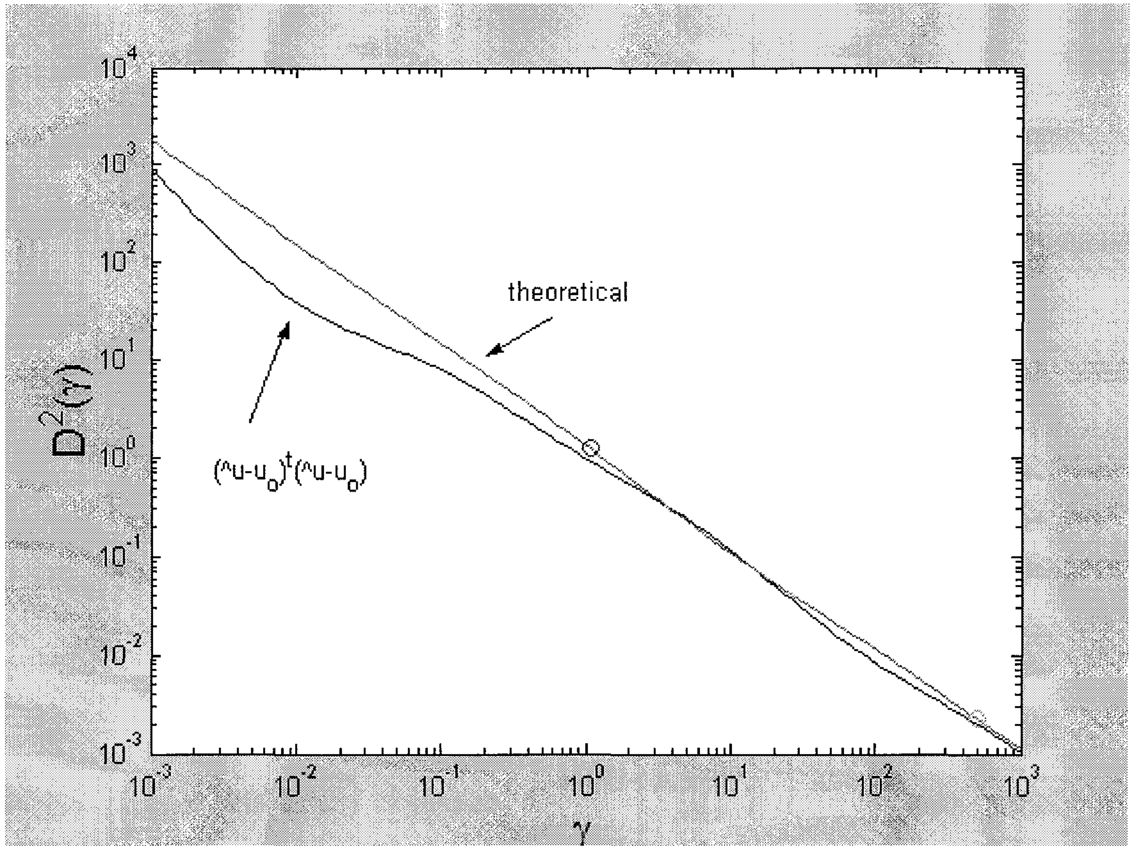
Test case: US standard atmosphere



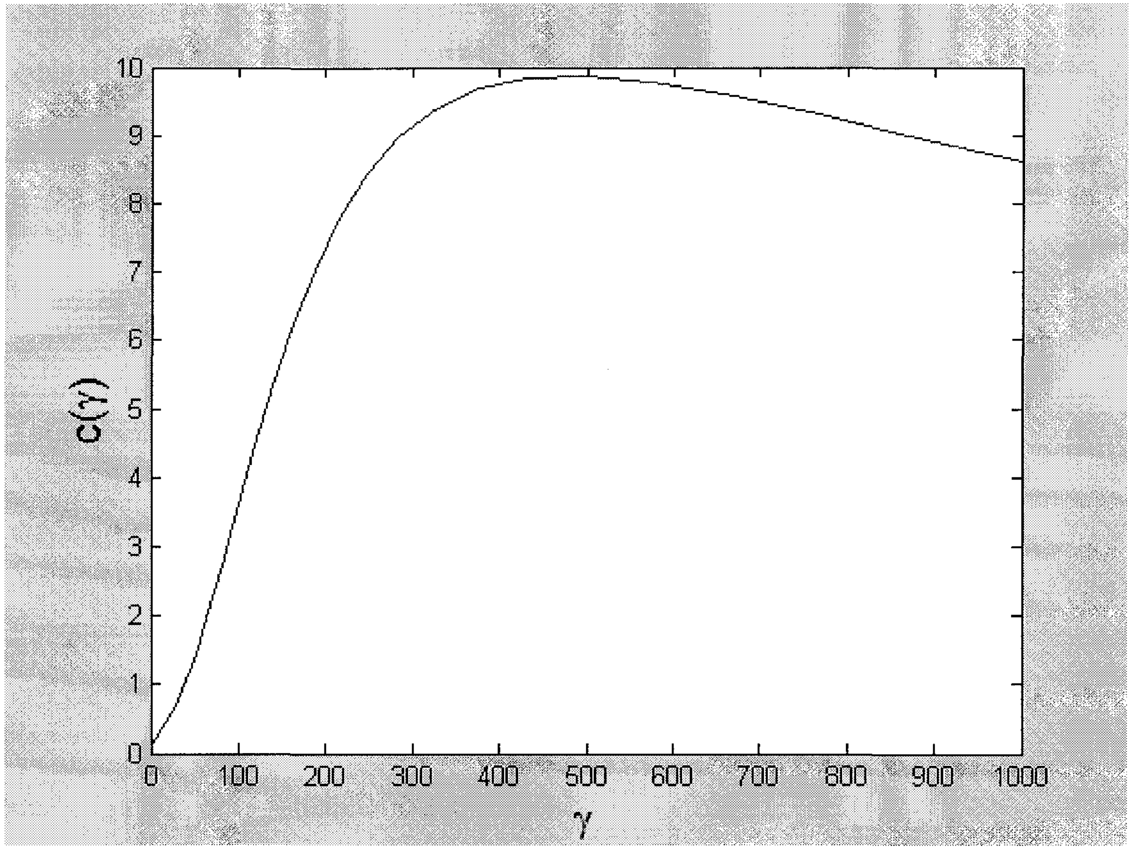
Spectral range 645 to 800 cm^{-1}
Sampling rate=0.25 cm^{-1}



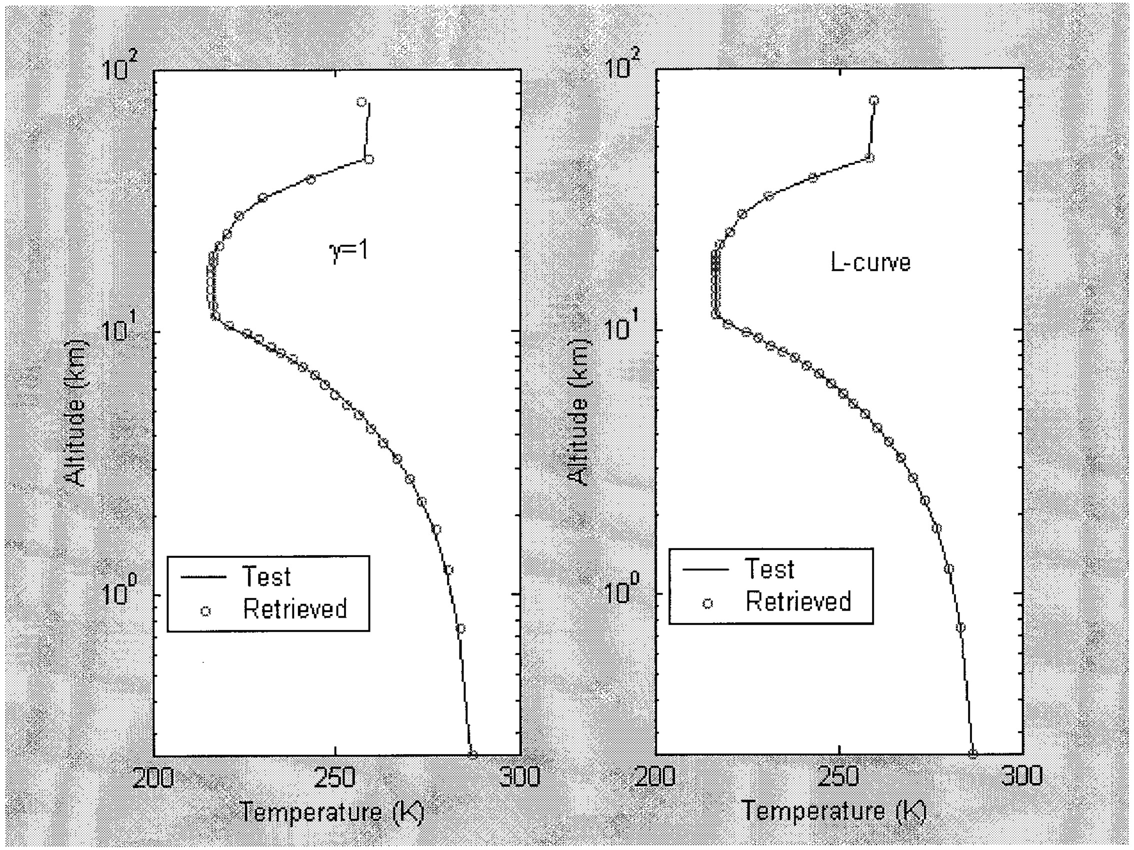
MSE as a function of γ



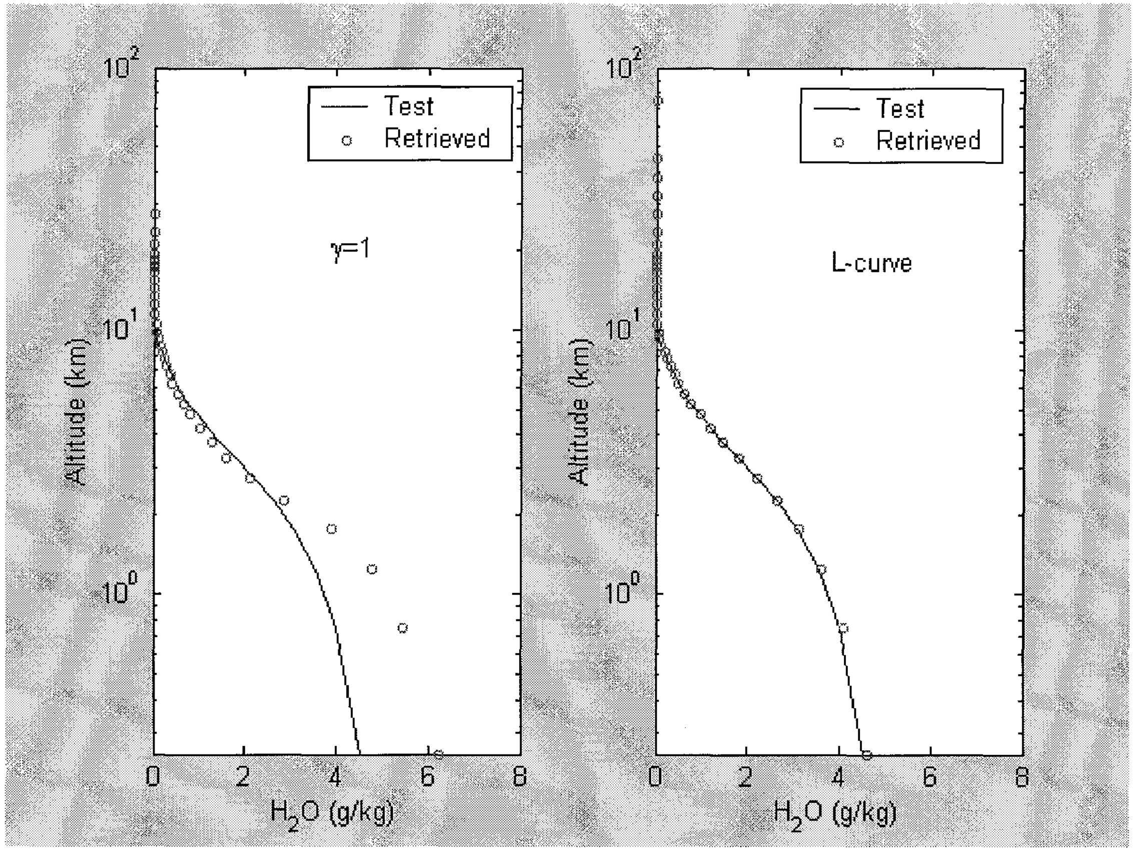
L-curve



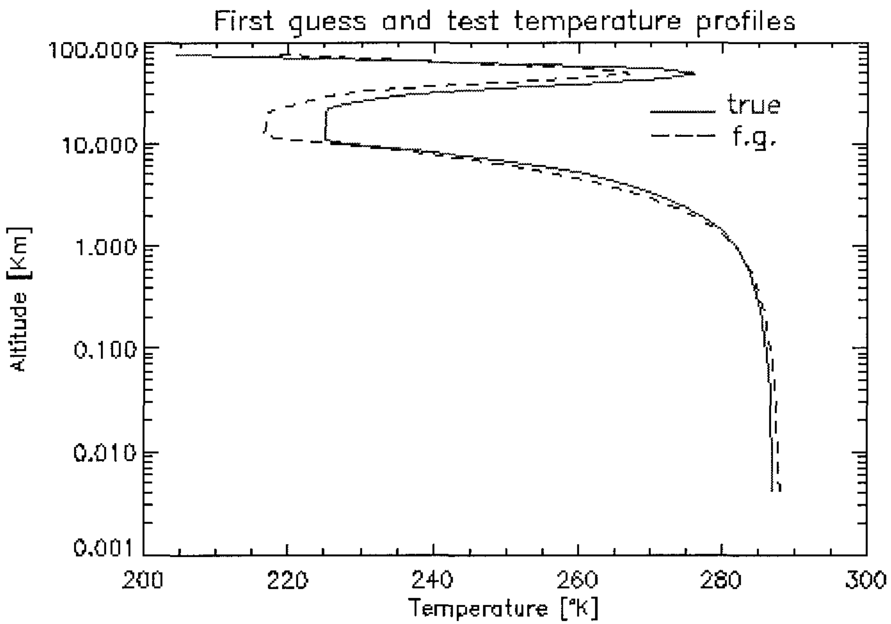
Results for temperature

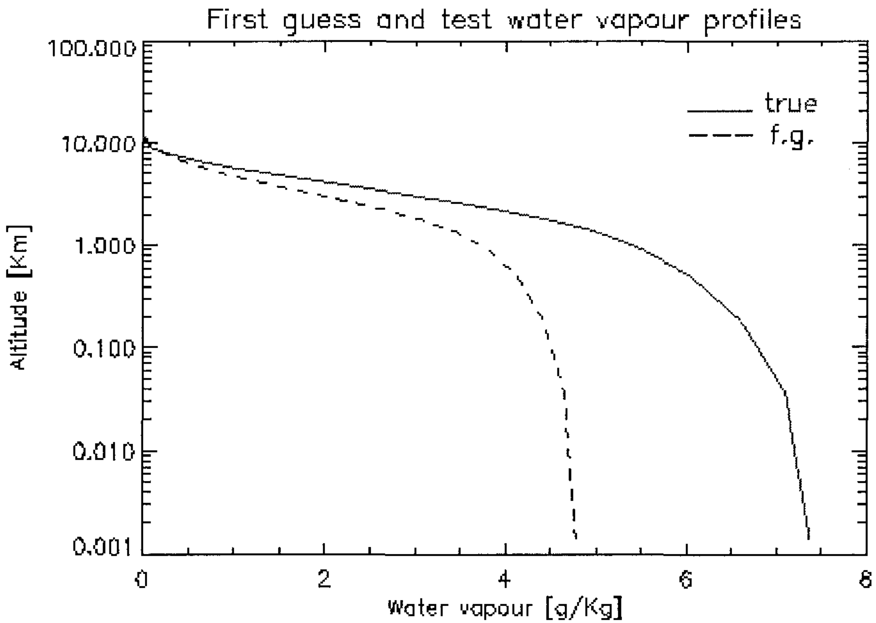


Results for water vapour

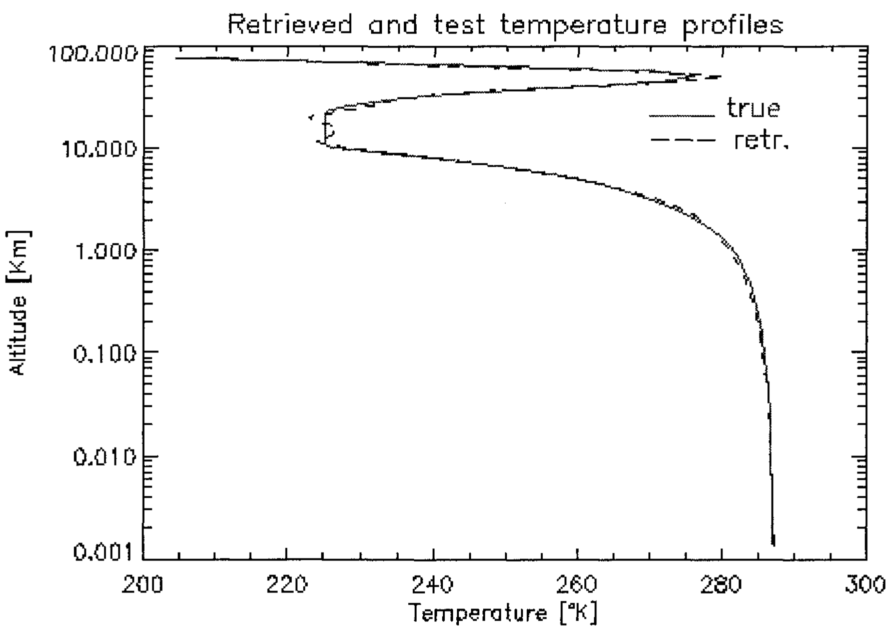


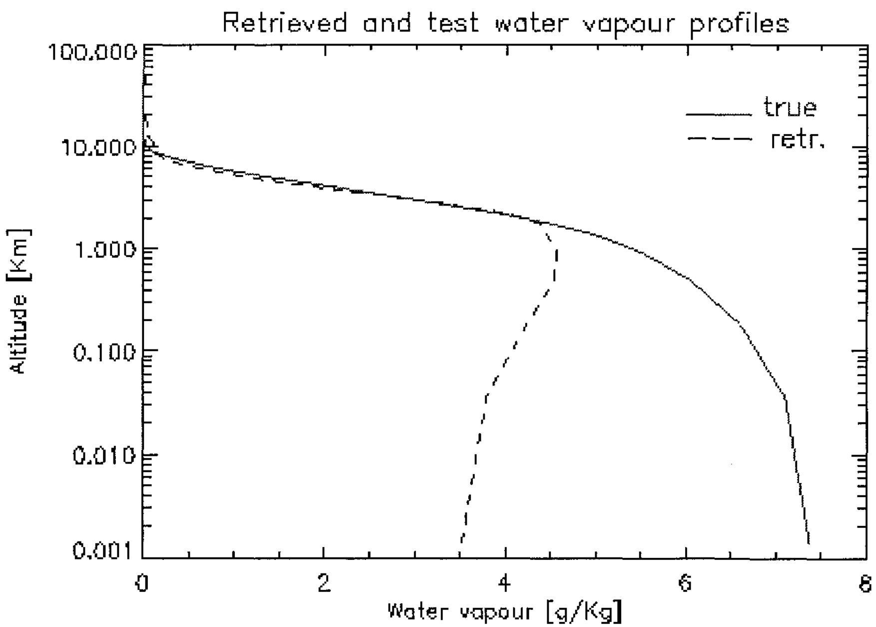
A much more elaborated numerical exercise



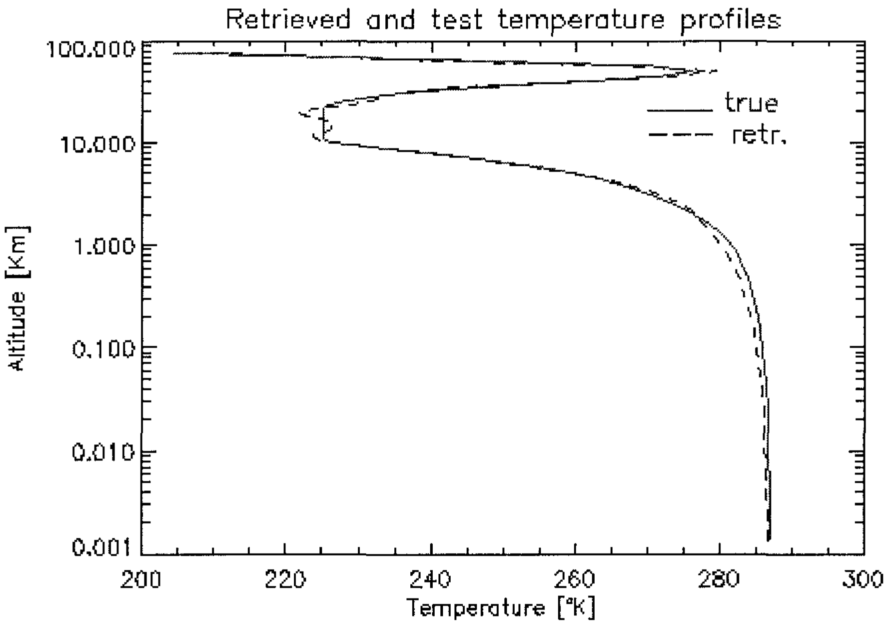


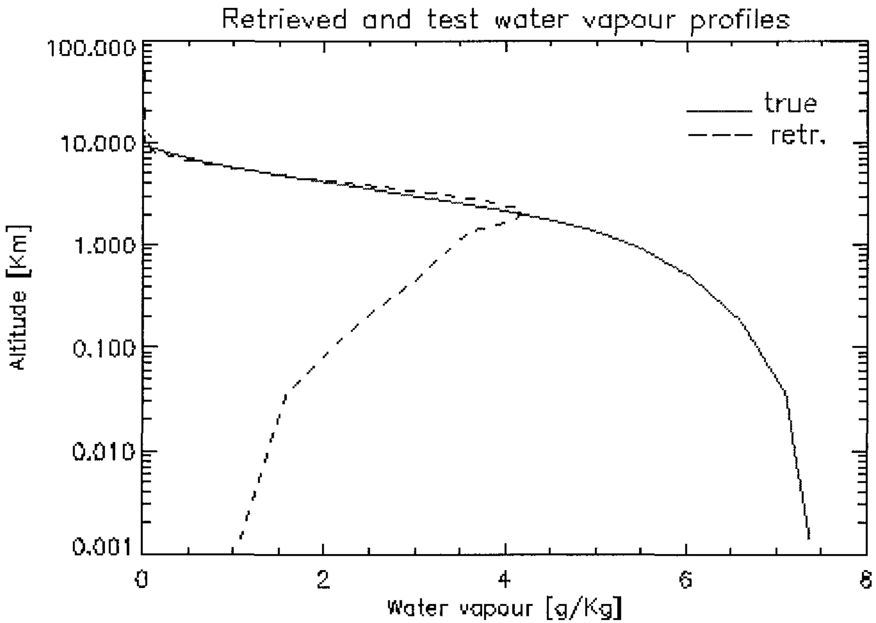
Statistical Regularization
1st Iteration



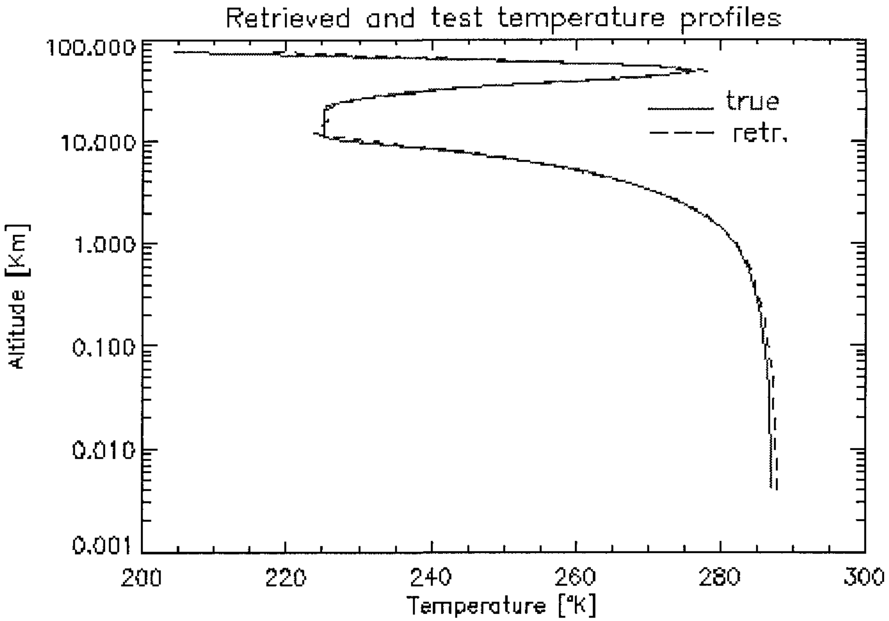


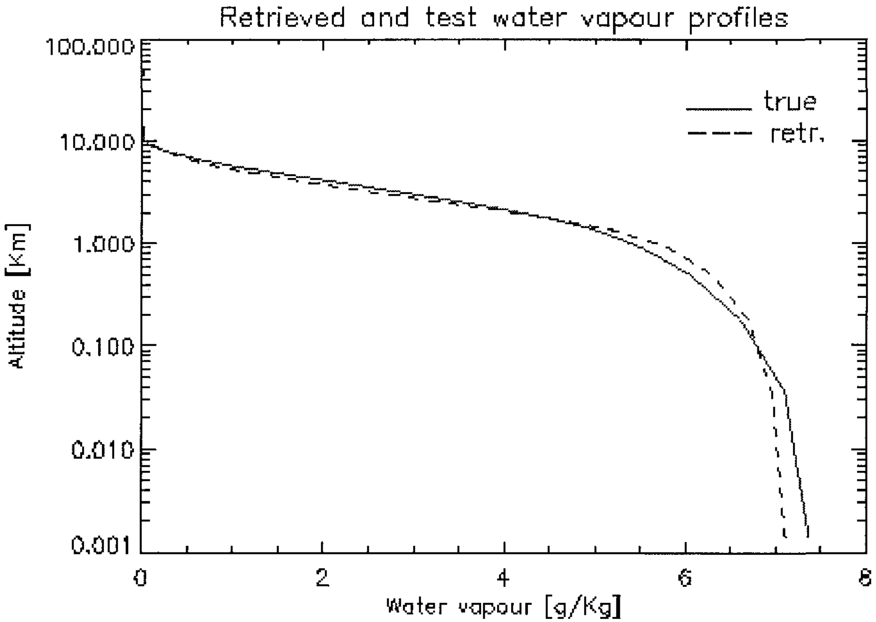
Statistical Regularization
2nd Iteration



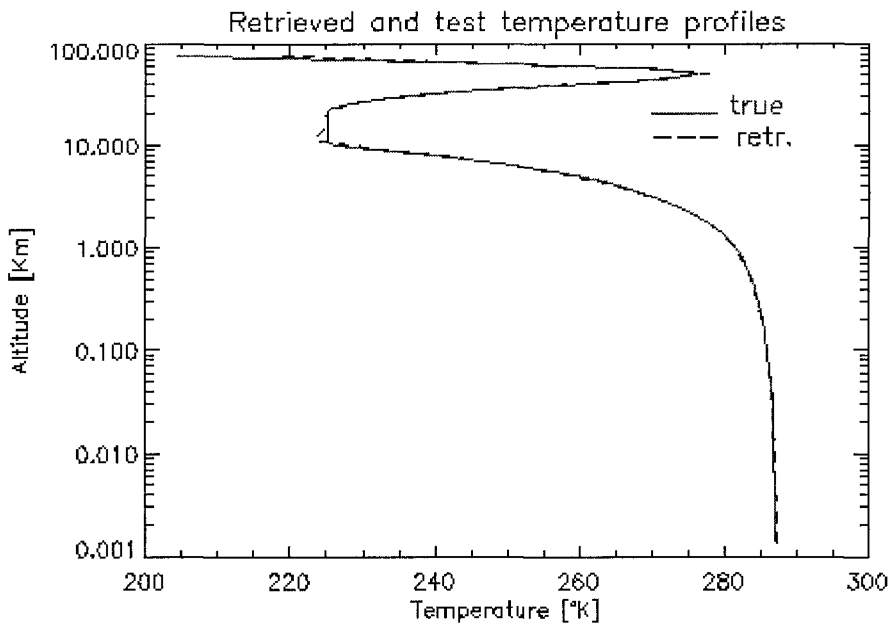


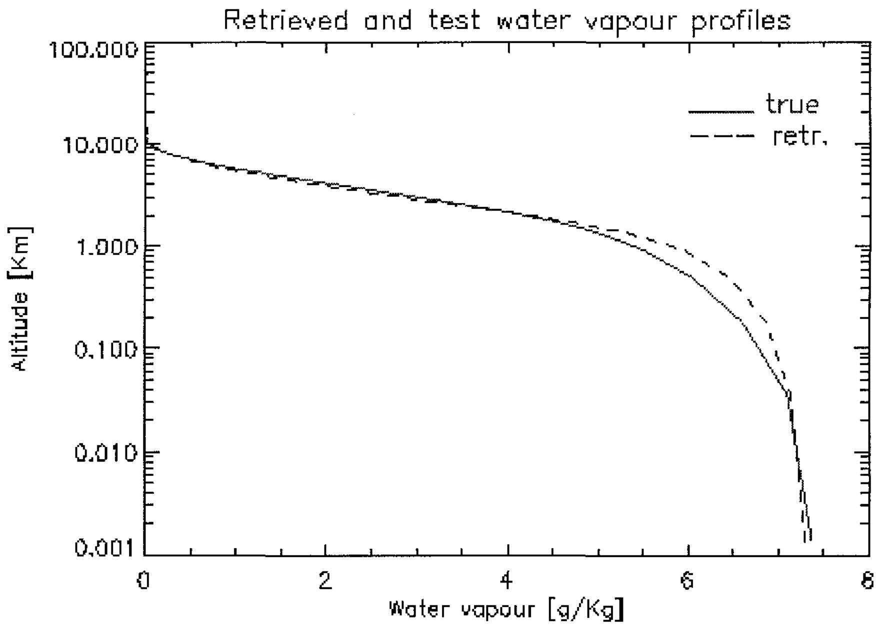
L-Curve, 1st Iteration





L-Curve, 2nd Iteration

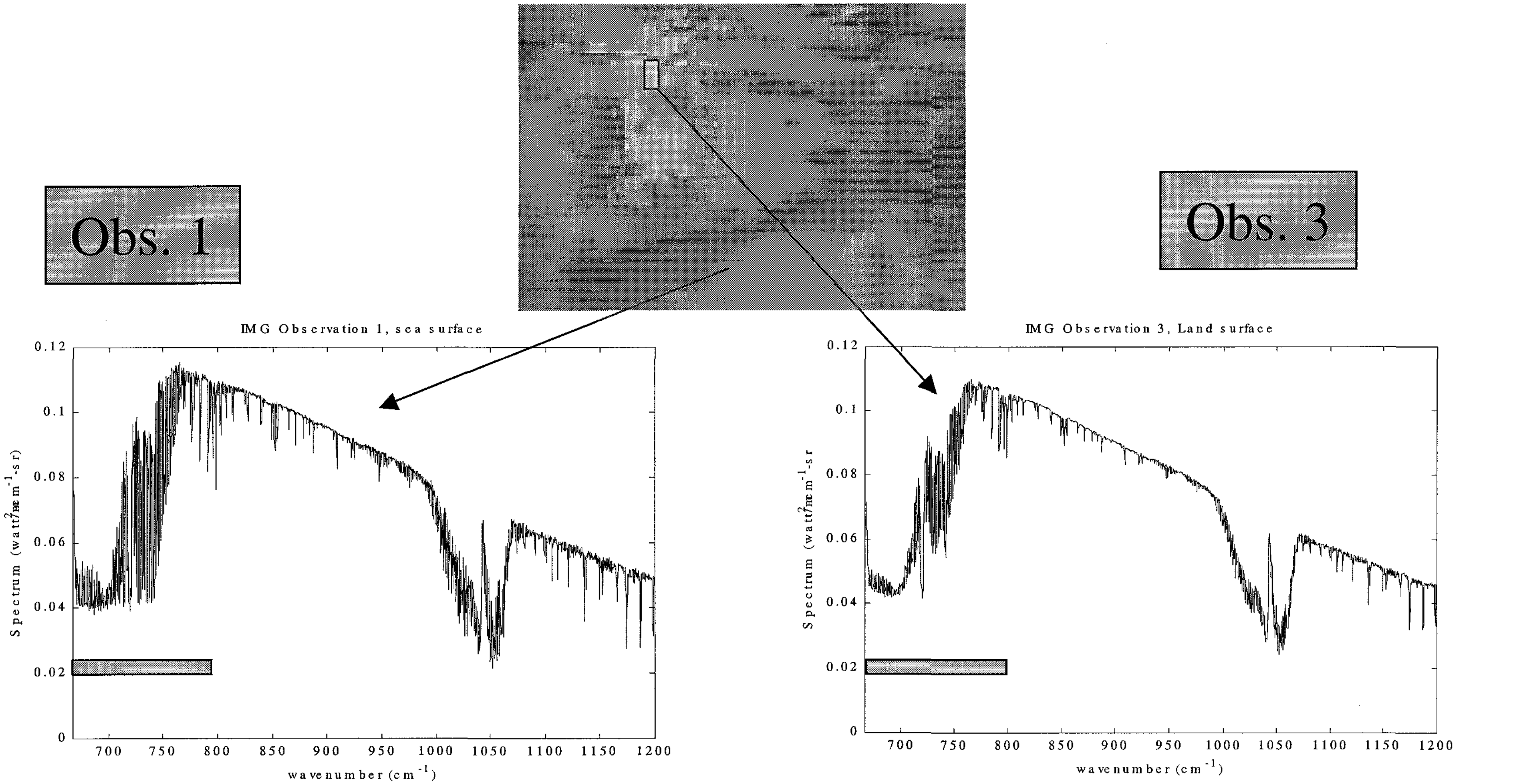




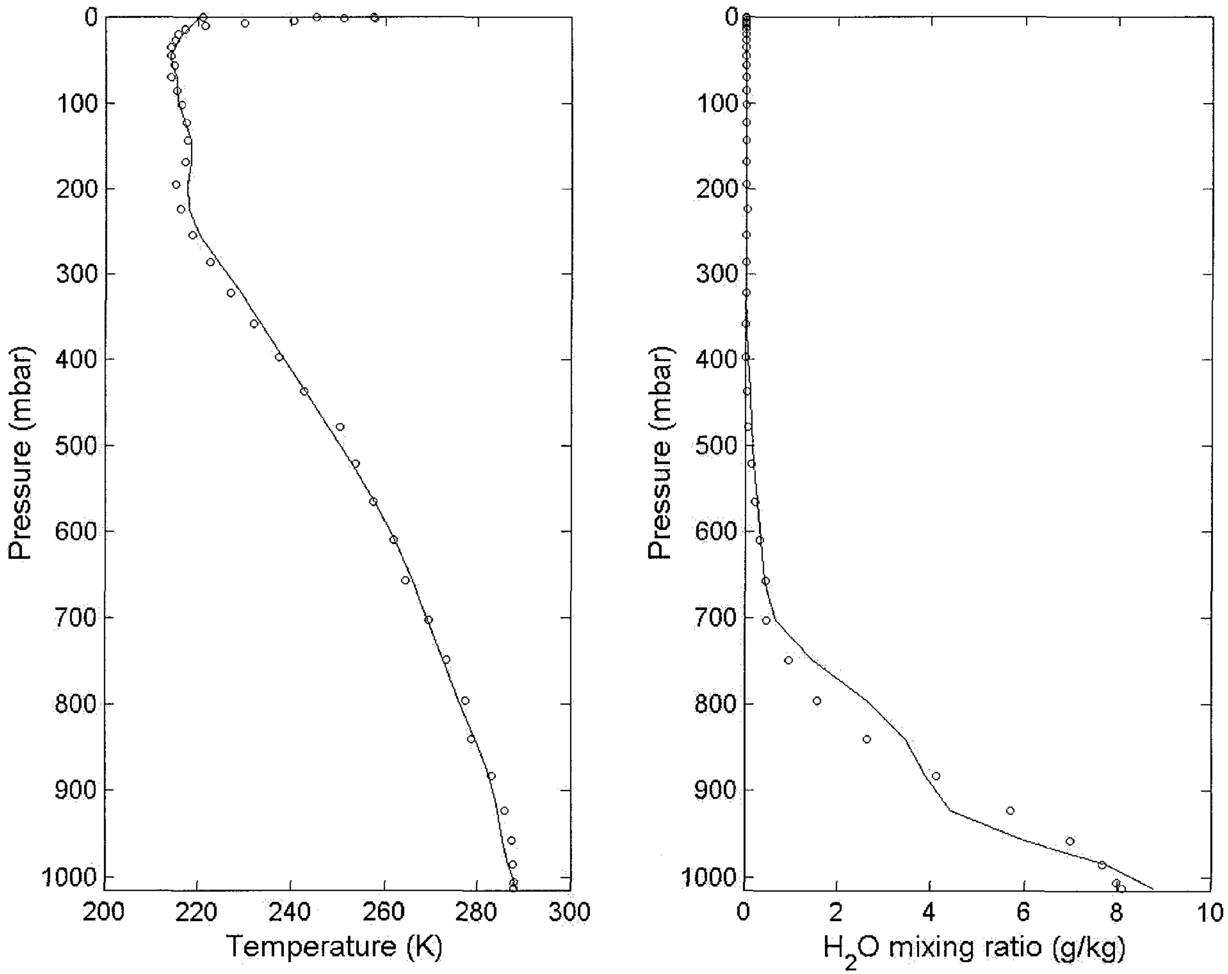
Inversion from IMG spectra Skin temperature: sea surface

IMG ID	*hs* index	Retrieved Skin Temperature (K)	ECMWF Skin Temperature (K)
#167229, Obs1	0.06	301.54	301.37
#310807, Obs. 1	0.27	288.03	288.28
#327906, Obs. 1	0.26	287.83	287.94
#327906, Obs. 5	0.27	285.98	285.72
#327906, Obs. 6	0.27	286.21	286.07
#374125, Obs. 6	0.07	302.06	302.26
#374126, Obs. 2	0.10	302.39	302.33

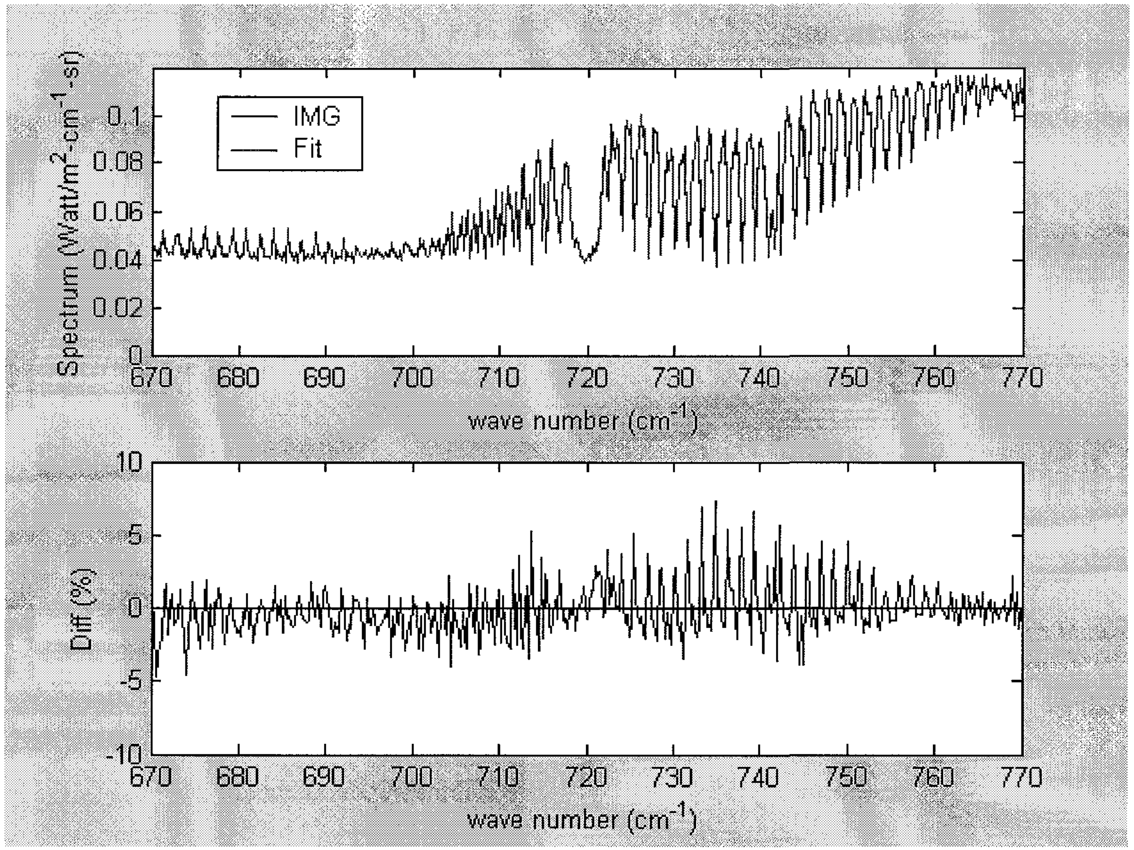
Mediterranean sea



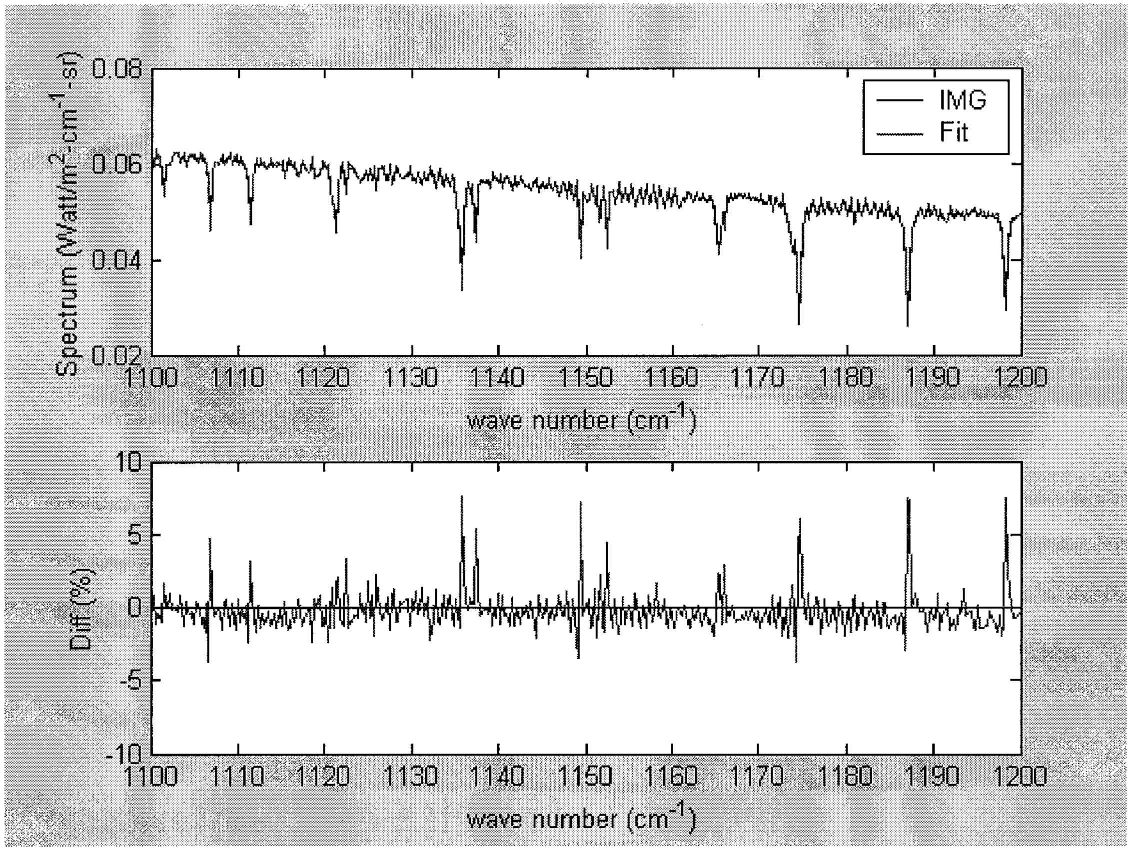
Mediterranean sea, Obs 1



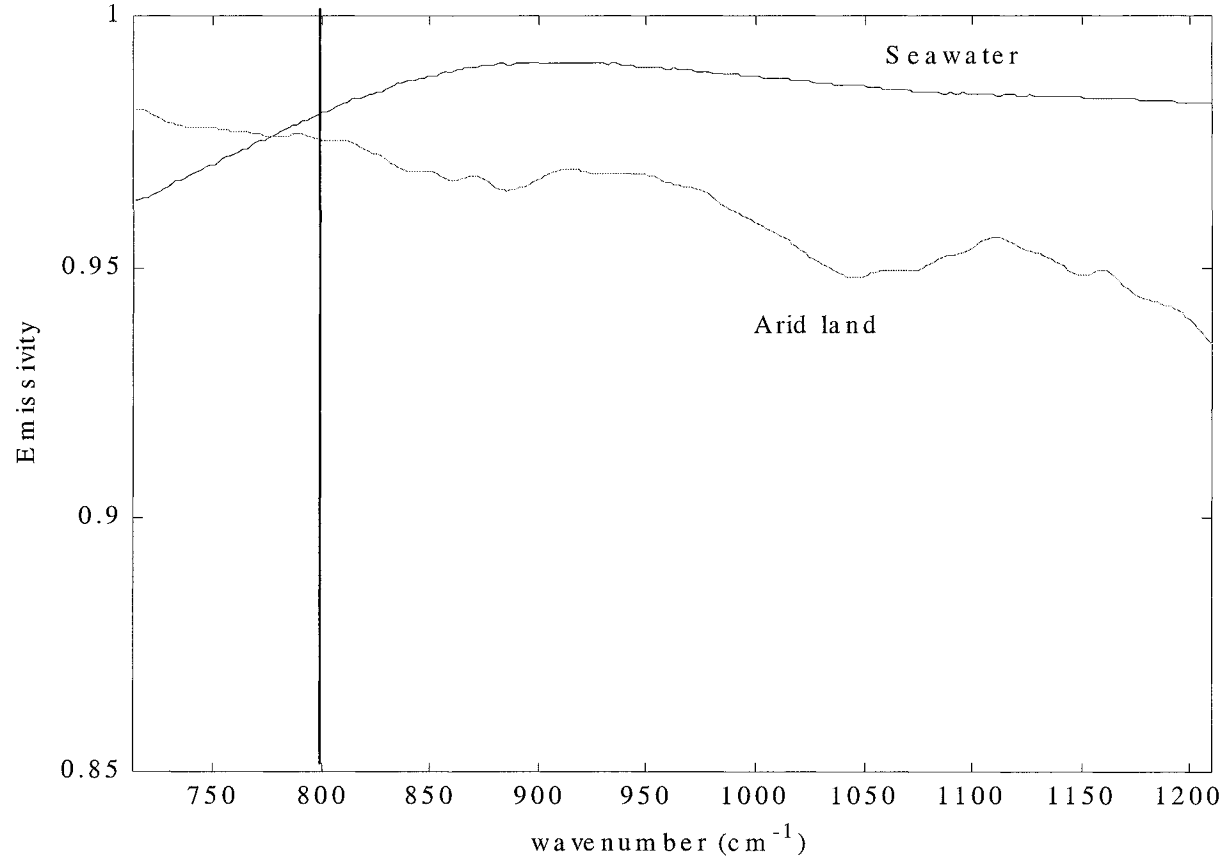
Mediterranean sea, Obs 1



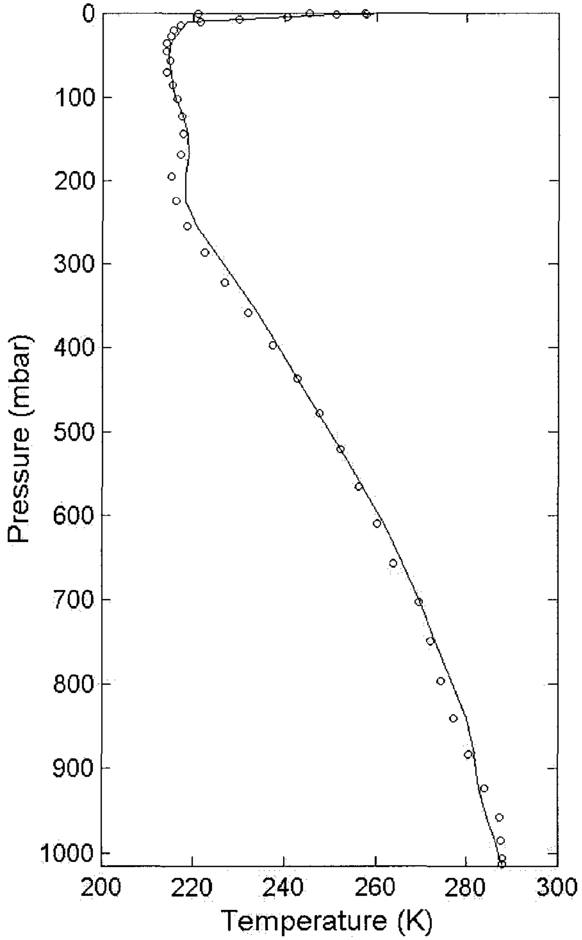
Mediterranean sea, Obs 1
Diagnostic check, Band 1100-1200 cm^{-1}

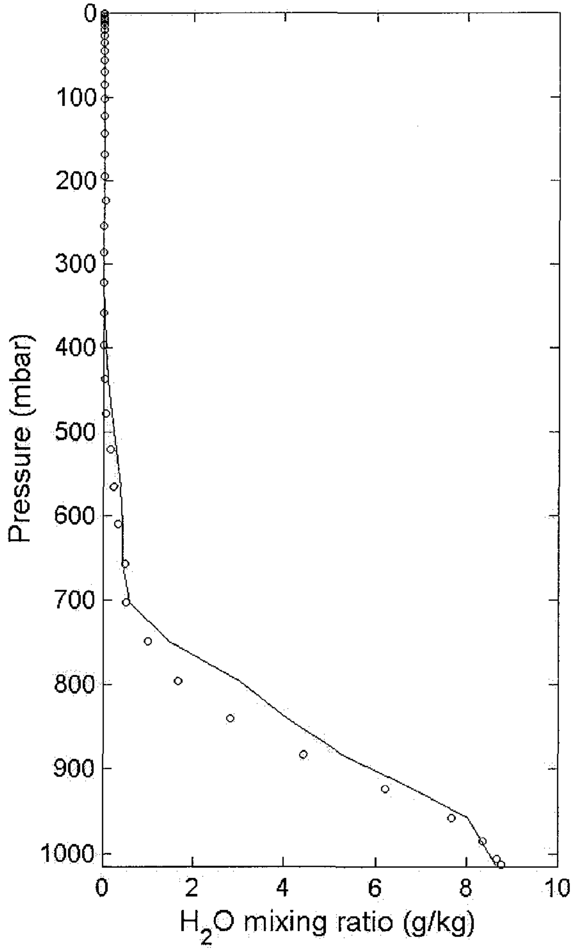


Mediterranean area, IMG Observation 3, Land Surface Seawater and Arid Land Spectral Emissivity

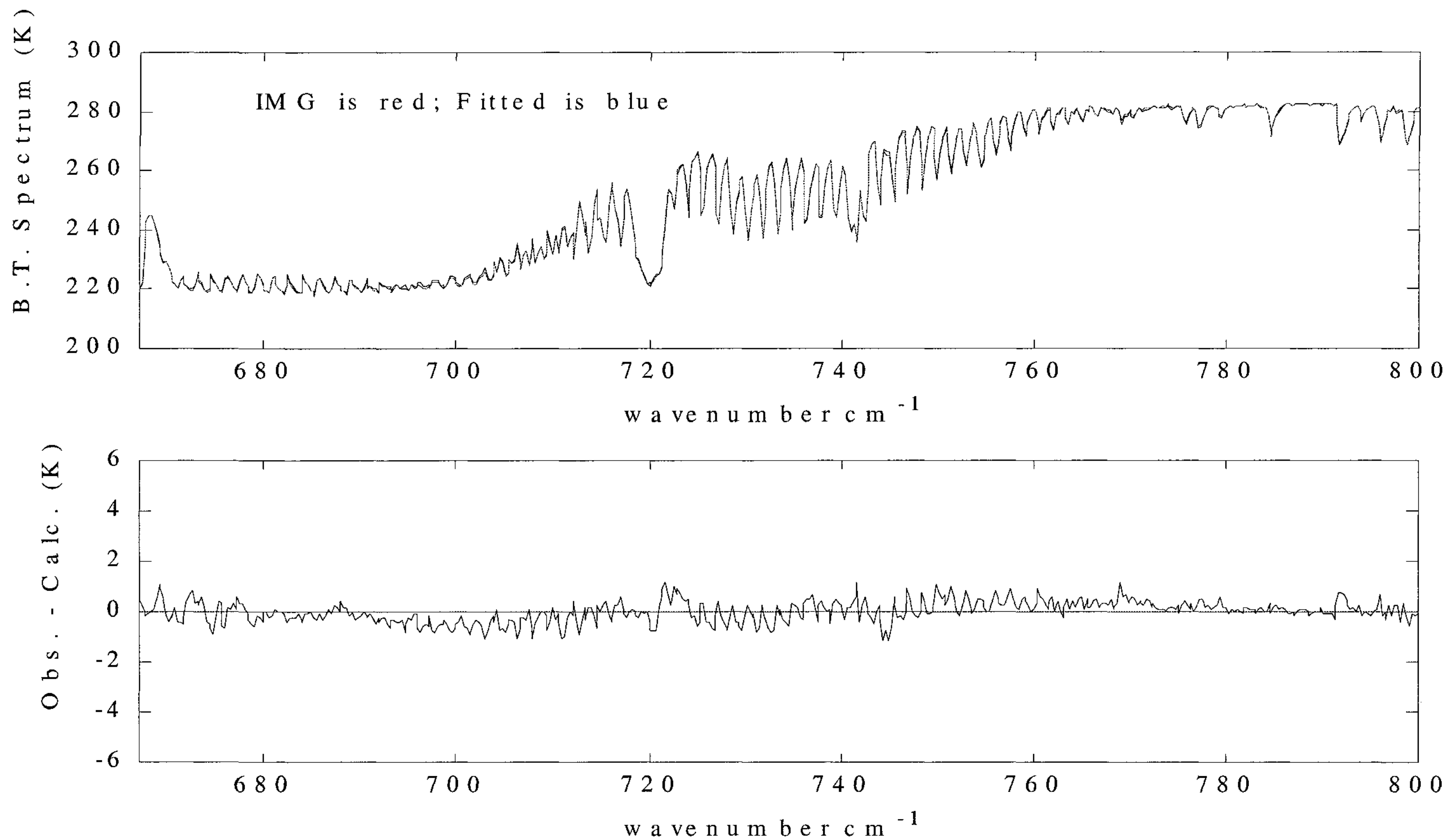


Mediterranean area, Obs 3: Retrieval

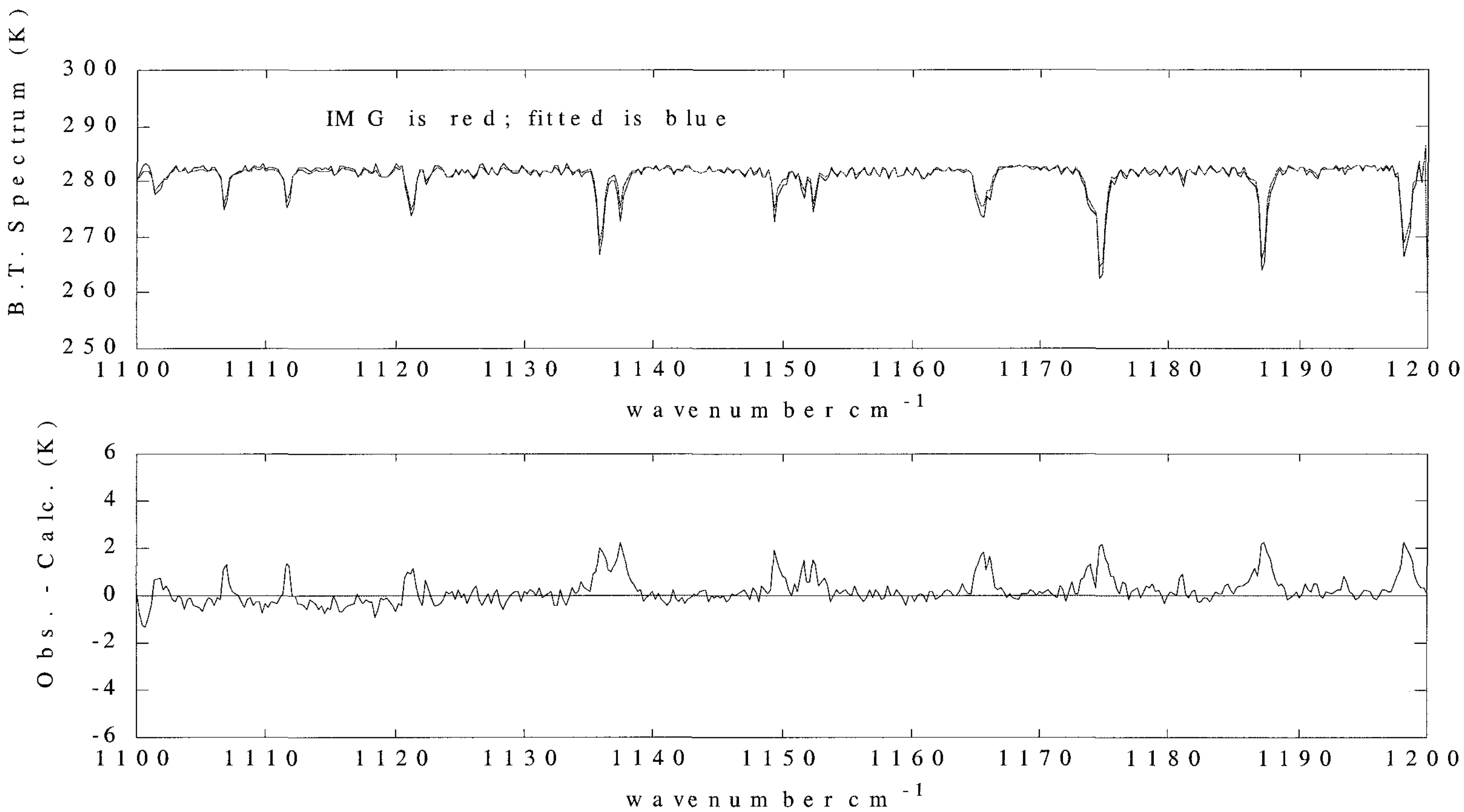




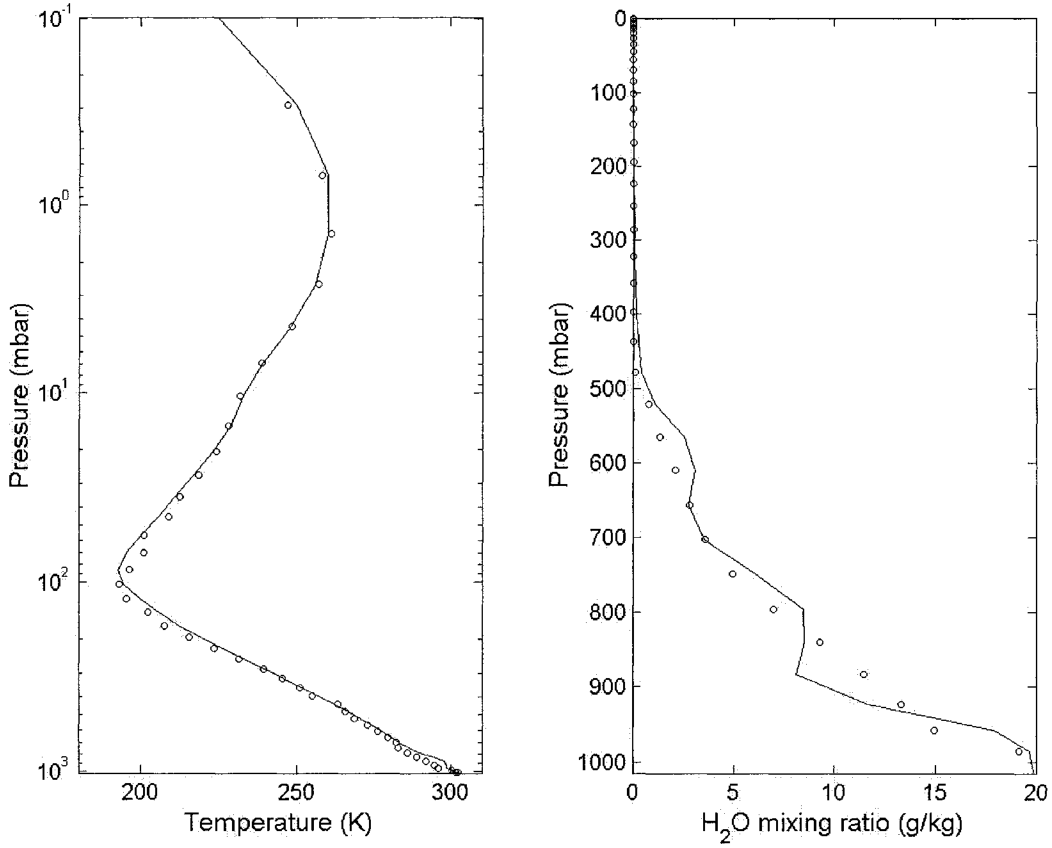
Mediterranean area, Obs 3
Consistency check, Band 667-800 cm^{-1}



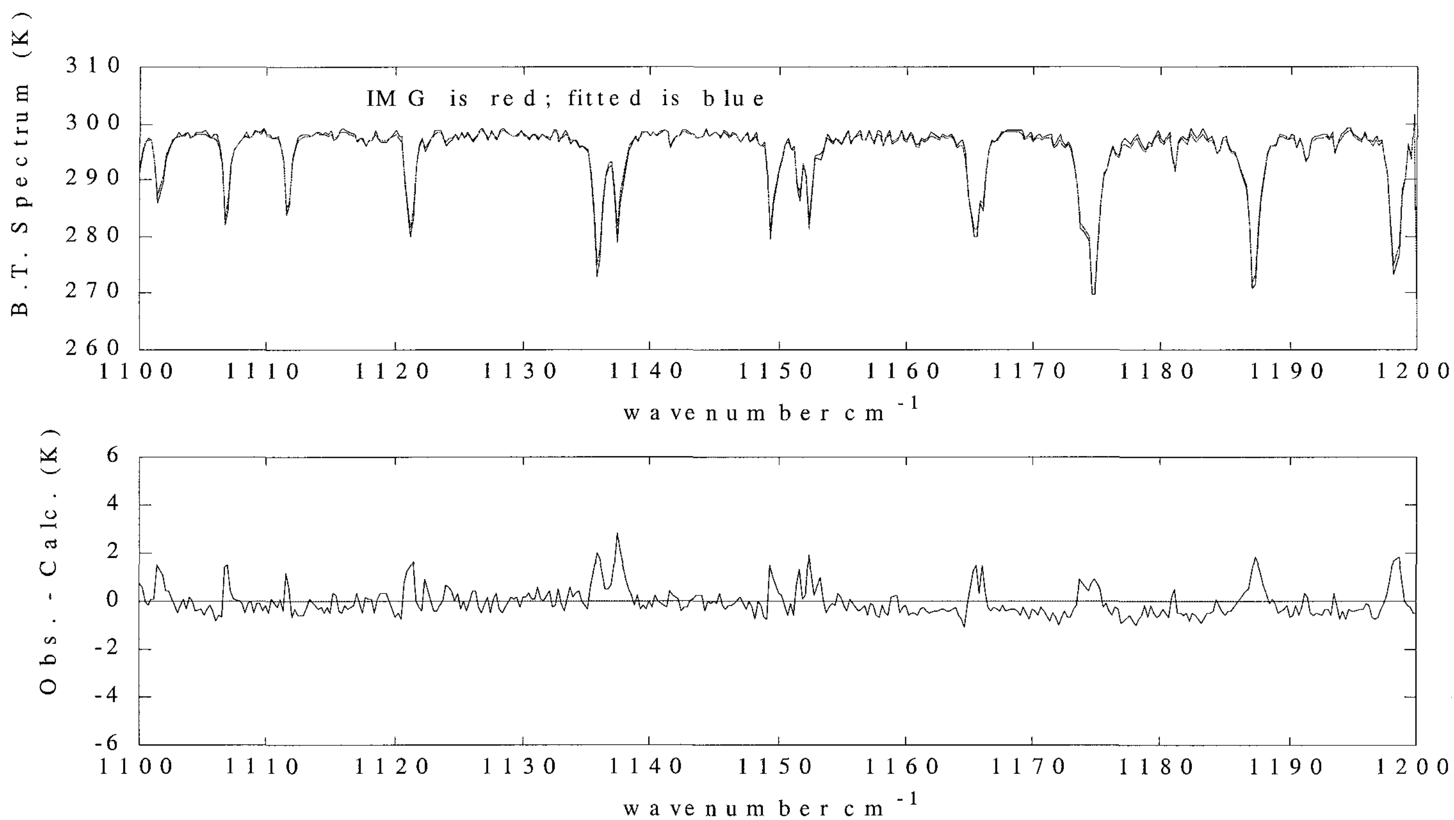
Mediterranean area, Obs 3
Diagnostic check, Band 1100-1200 cm^{-1}



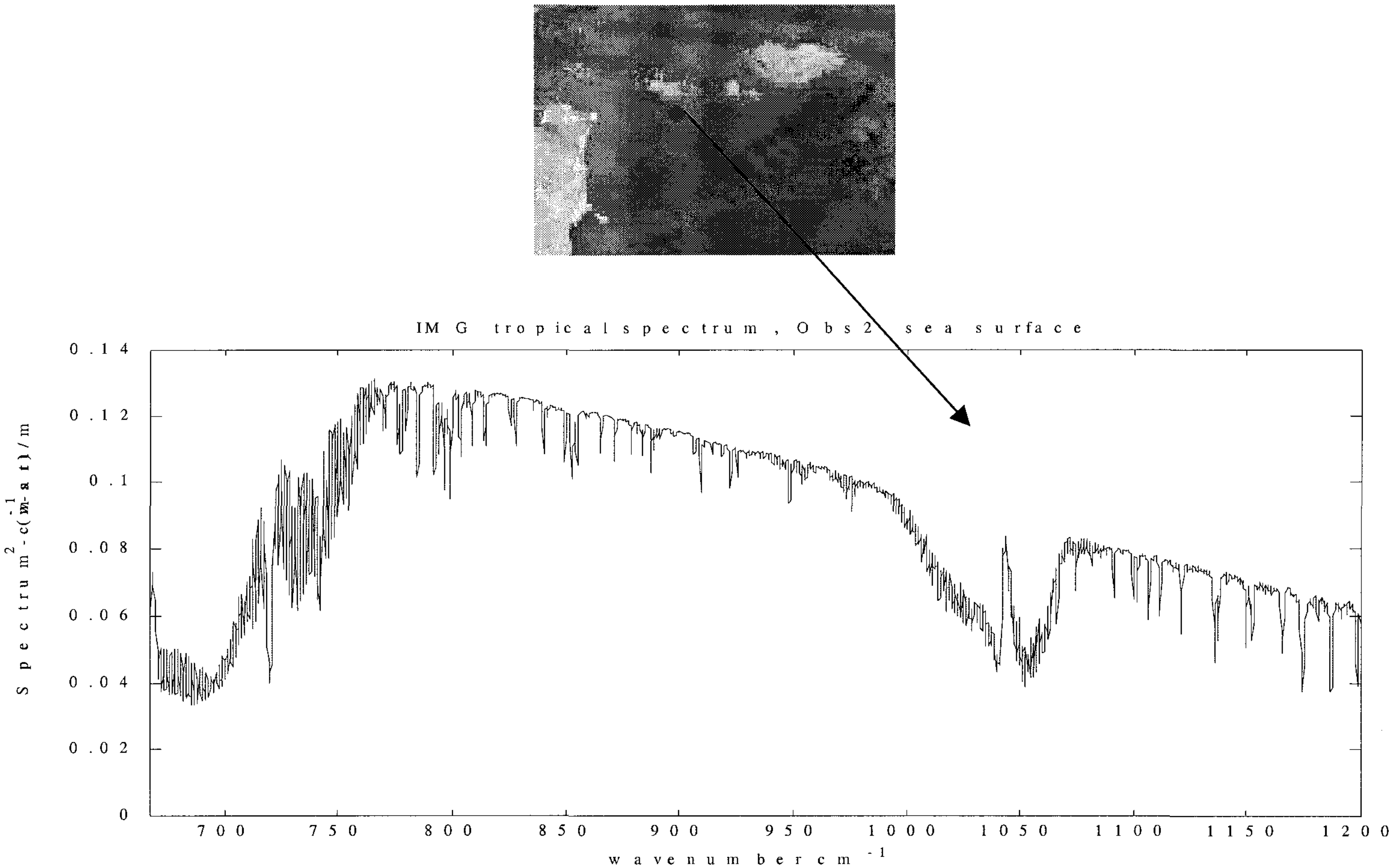
Indian Ocean sea, Obs 2: Retrieval



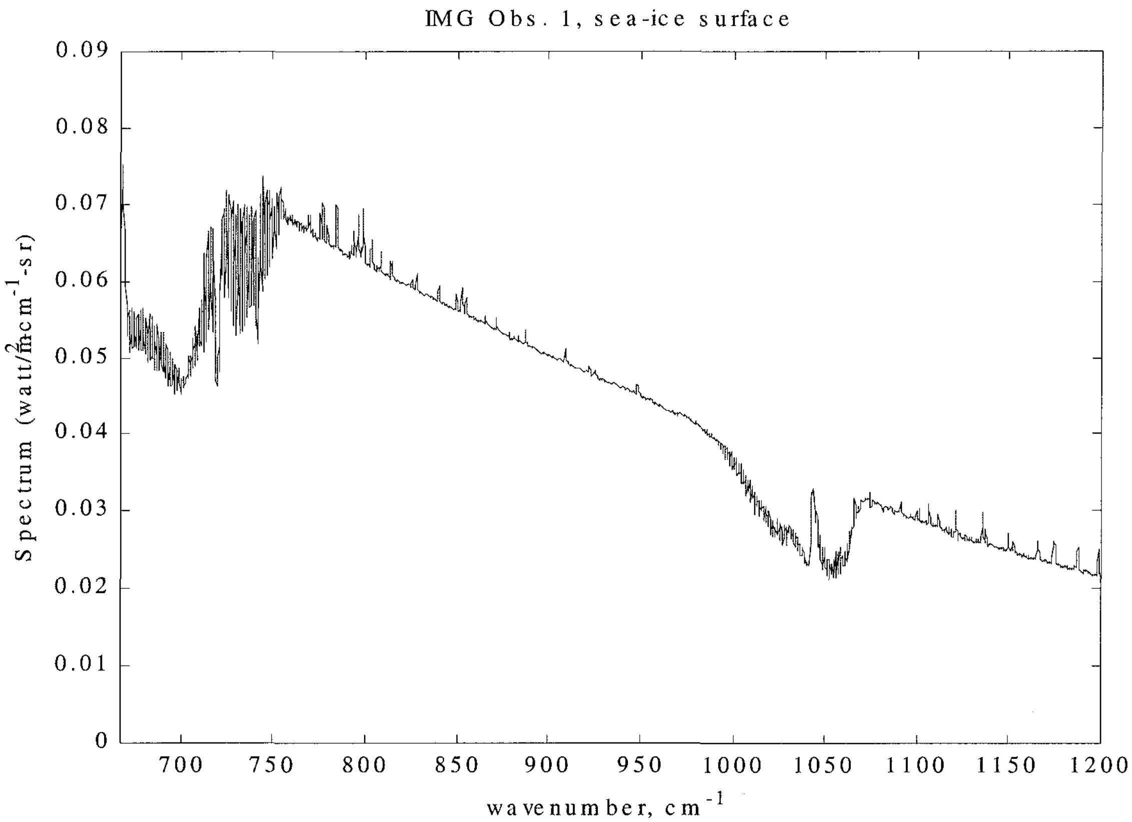
Indian Ocean, Obs 2
Diagnostic check, Band 1100-1200 cm^{-1}



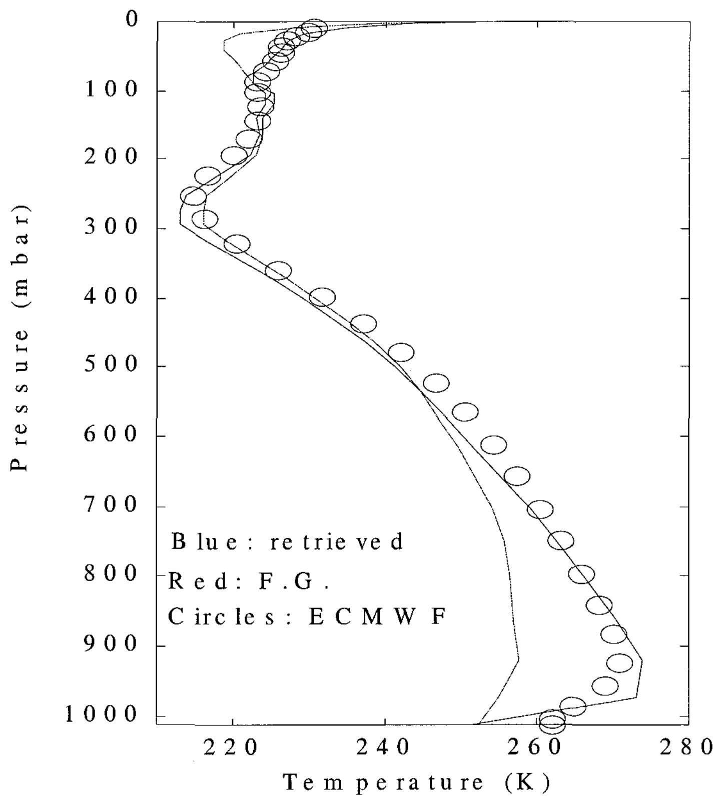
Tropical area, Indian Ocean

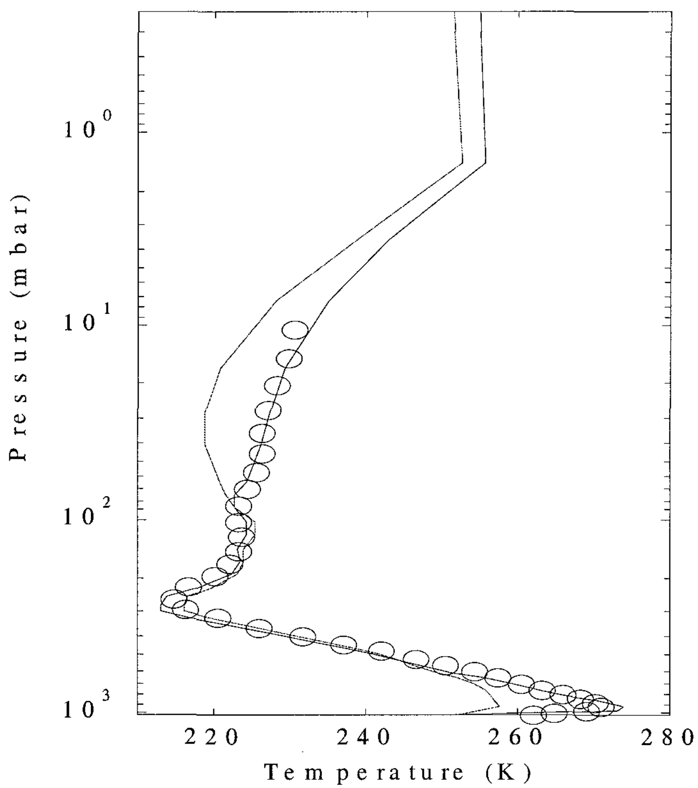


Arctic Region, Point Barrow, Alaska

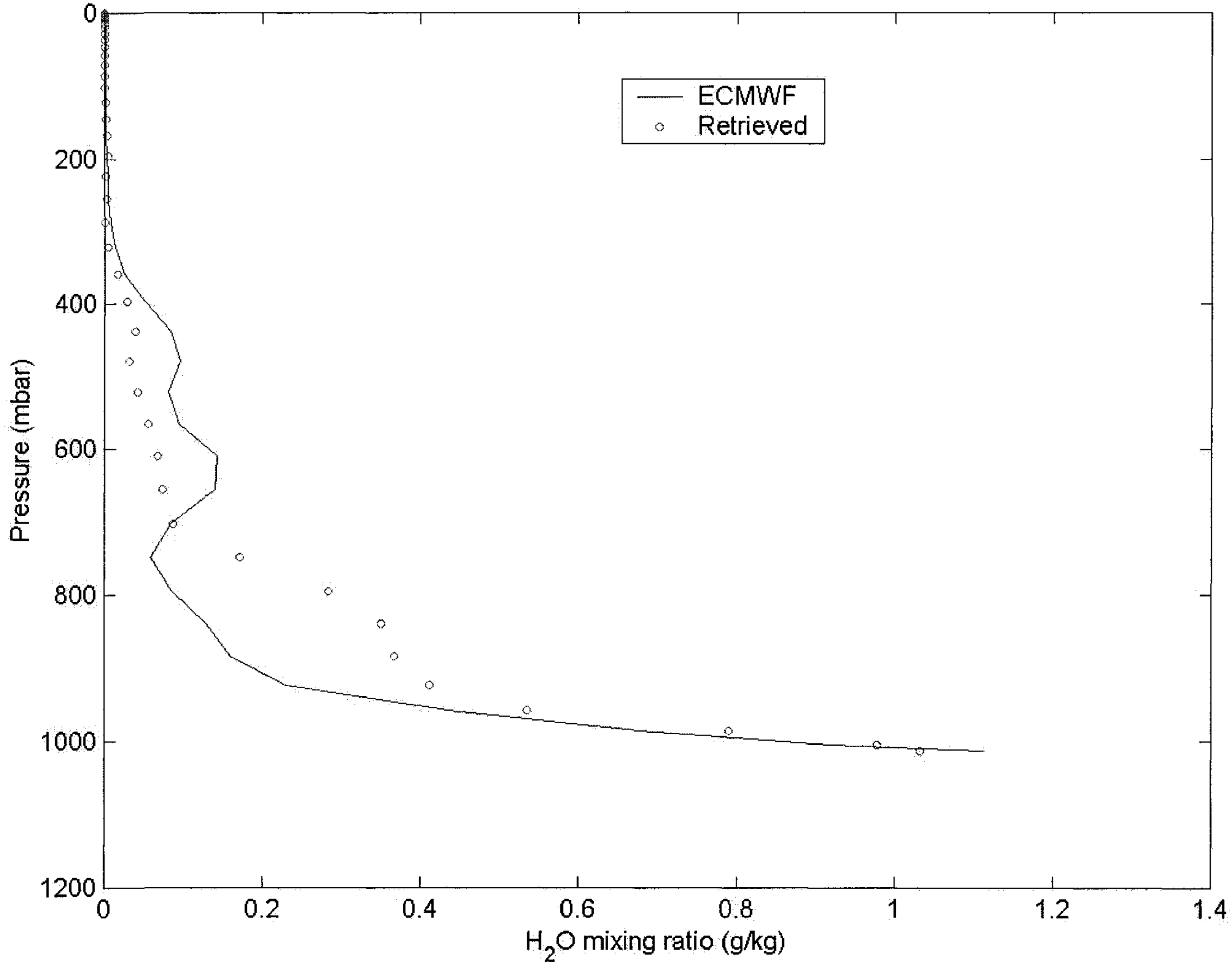


Arctic area sea-ice, Obs 1
Temperature Retrieval

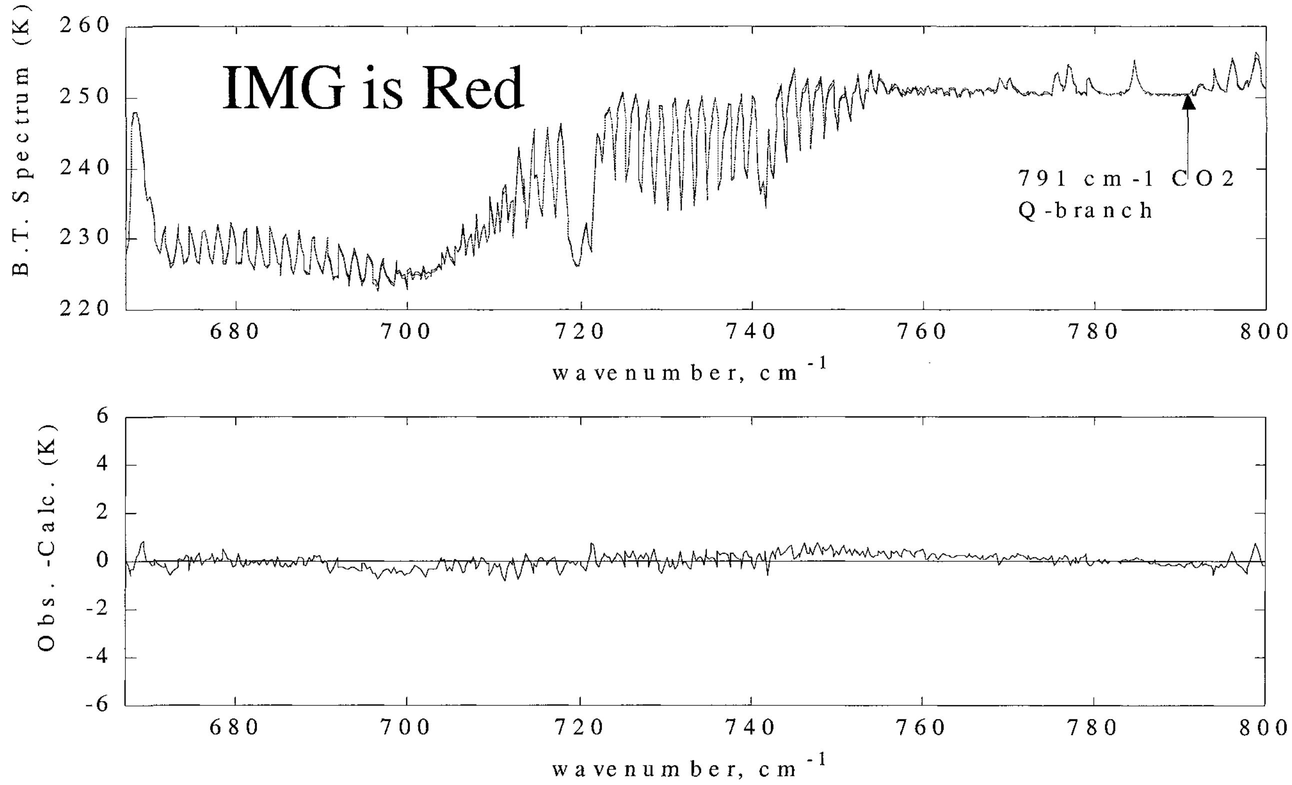




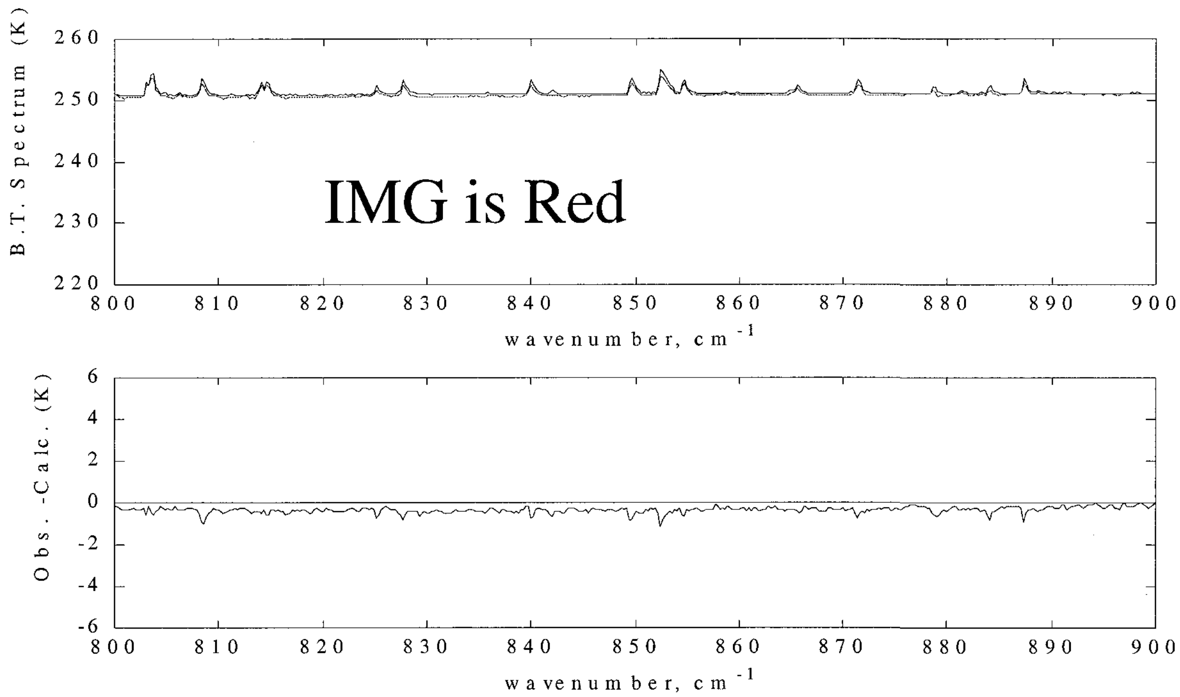
Arctic area sea-ice, Obs 1



Arctic Region, Obs 1
Consistency check, Band 667-800 cm^{-1}

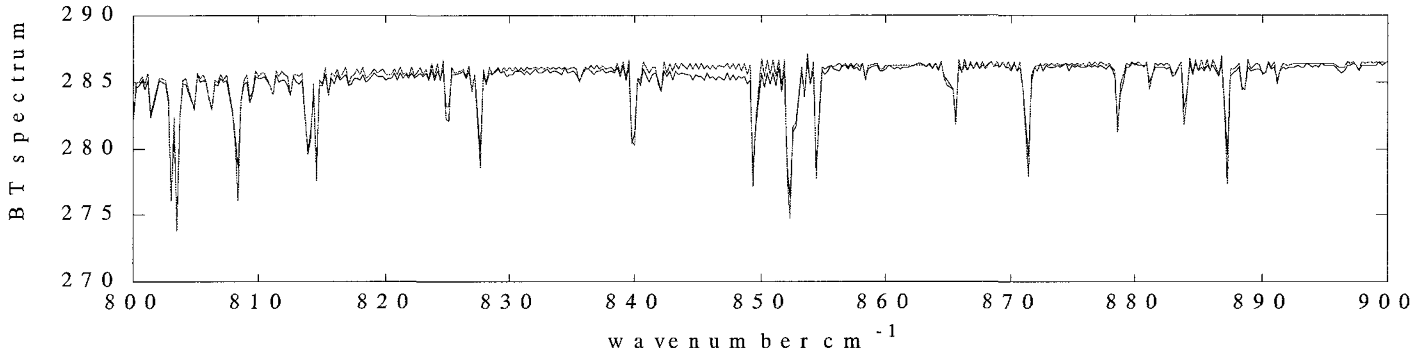


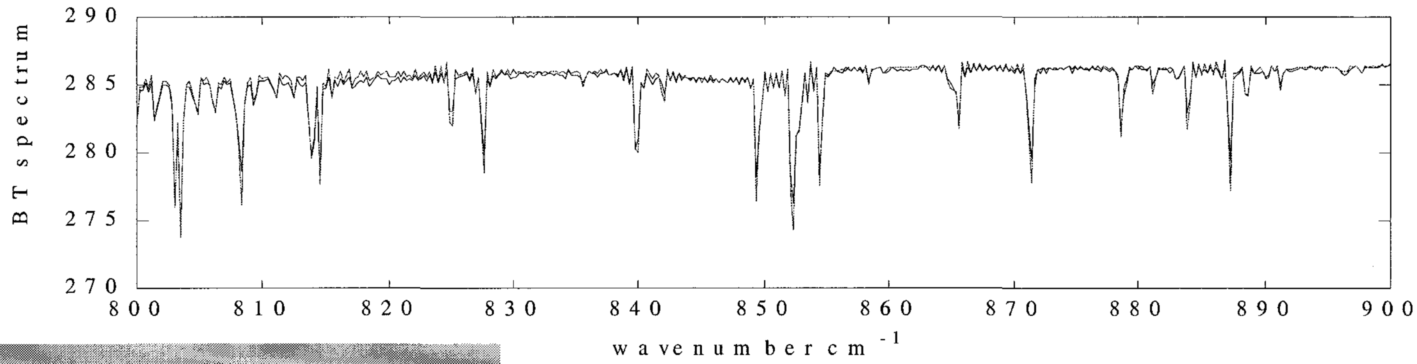
Arctic Region, Obs 1
Diagnostic check, Band 800-900 cm^{-1}



Effect of CFC in the window region 800-900 cm^{-1}
Sardinia Obs. 1

IMG is blue

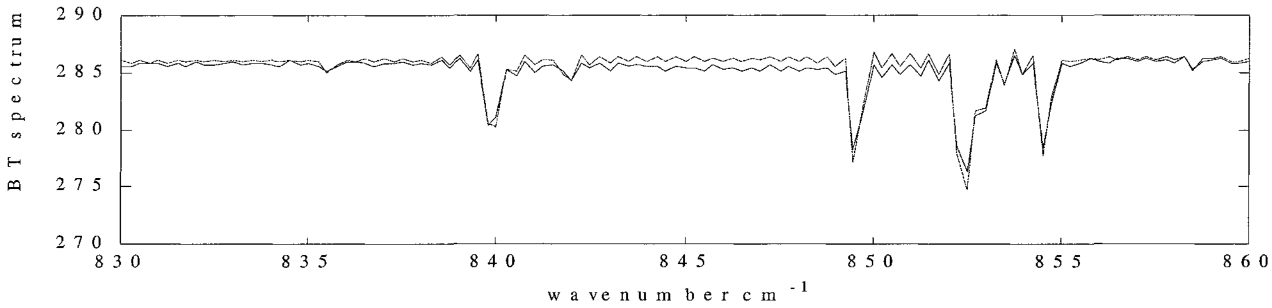


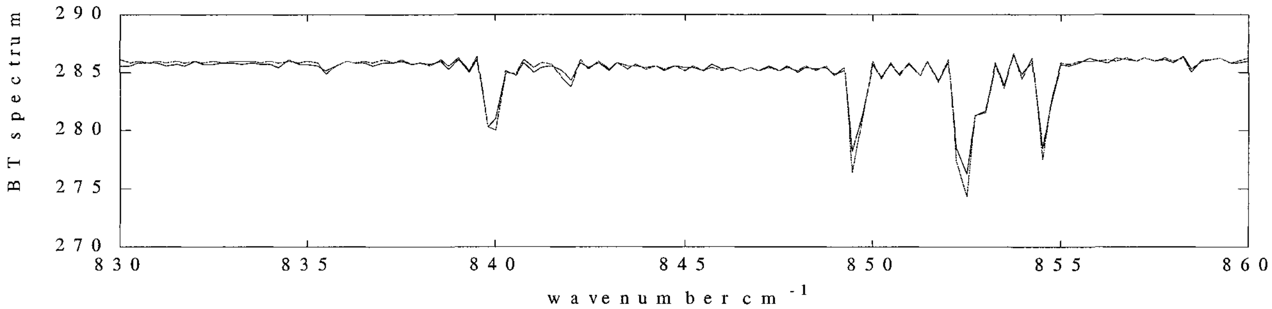


CFC m.r.= 275 pptv

Effect of CFC in the window region 800-900 cm^{-1}
Sardinia Obs. 1

IMG is blue

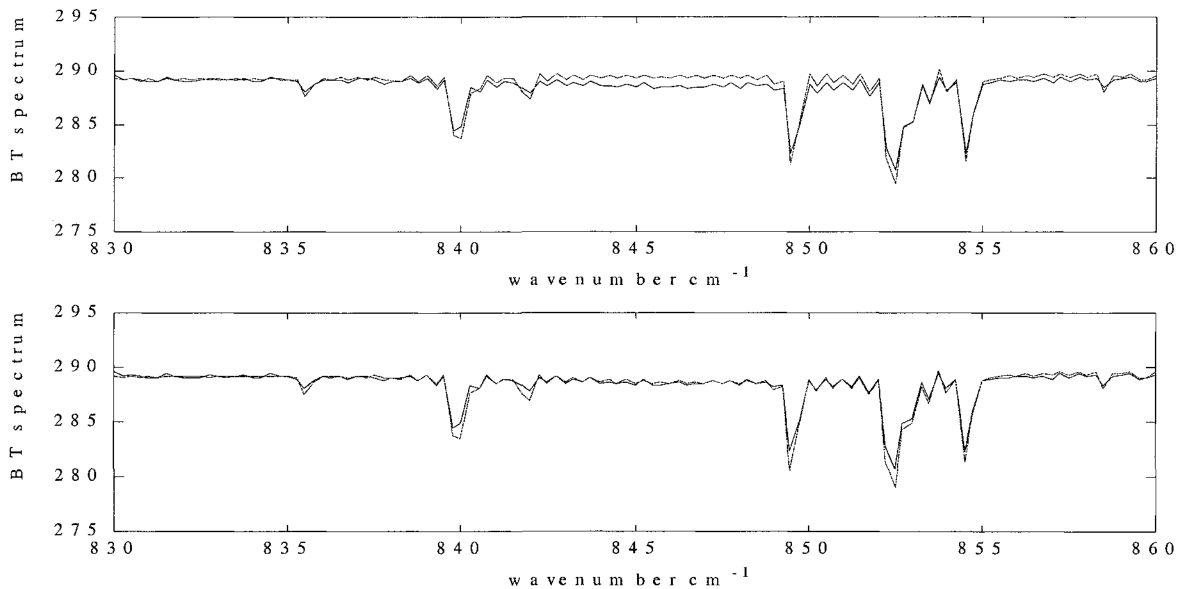




CFC m.r.= 275 pptv

Effect of CFC in the window region 800-900 cm^{-1}
Atlantic Ocean Obs. 1

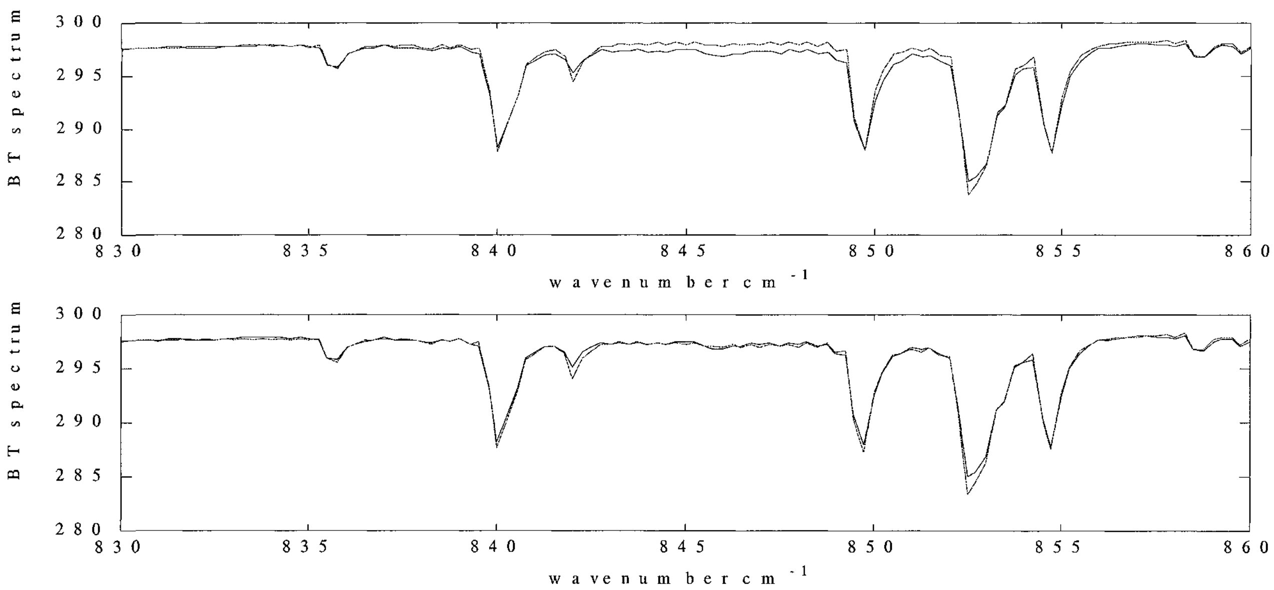
IMG is blue



CFC m.r.= 275 pptv

Effect of CFC in the window region 800-900 cm^{-1}
Indian Ocean Obs. 2

IMG is blue



CFC m.r.= 275 pptv

Summary and Conclusions

- The physics and mathematics of radiative transfer suitable for high spectral resolution infrared sensor has been reviewed
- The basic aspects of mathematical inversion for geophysical parameters have been discussed
- The *ridge regression nucleus* of regularization has been highlighted
- Statistical regression is the only scheme which fully achieves physical and dimensional consistency
- The choice $\gamma=1$ has a precise meaning in the context of statistical regularization, although $\gamma=1$ does not necessarily minimize the root mean square error
- L-curve seems to provide a better scheme to achieve rms error minimization,
- IMG spectra have been considered for inversion of geophysical parameters which have evidenced:

A) The high consistency of state-of-art line-by-line forward models

B) The fine accuracy of retrieval which seems to be in line with what it is expected.