A Solution of the Non-linear Problem of Simultaneous Estimation of the Atmospheric Temperature and Underlying Surface Emissivity on the Basis of the Satellite Spectral Measurement of the Microwave Region Spectrum

Youry Plokhenko

Research Center "Planeta" of Hydrometeorological Service of Russia, Moscow, Russia

It is known that MSU data measurements are especially important in the remote sensing of atmospheric parameters under cloudy conditions on the basis of TOVS spectral information of NOAA satellites. Actually cloudiness is considered to be transparent in the MSU microwave region. This assumption simplifies considerably the numerical solution of the inverse problem, because the spectral measurements of the infra-red region of the spectrum on the basis of HIRS/2 information can be totally contaminated by cloudiness .

In the processing of MSU information there is a principal aspect involving the estimation of underlying surface emissivity. There are different numerical schemes of MSU information processing. We will consider them by looking at the example of one algorithm used broadly in practice, which consists in the estimating of emissivity, using the measurement in "clear" channel 1 and than the estimating the vertical temperature distribution, using the measurement in channels 2 through 4. Further we shall discuss the numerical solution of RTE for the microwave region of spectrum, paying especial attention to the question of the mutual dependence of the accuracies of the emissivity and temperature parameters being estimated.

I. Recursive estimation of desired parameters.

Let us consider the model of the measurements, based on RTE and some approximations of its parameters (Weinreb et al. 1981):

$$J = \underbrace{\varsigma T(p_s)\tau(p_s) + \int_{\tau(p_s)}^{1} T(p)d\tau(p)}_{(I)} + \underbrace{(1-\varsigma)\tau^2(p_s)\int_{\tau(p_s)}^{1} T(p)\tau^{-2}(p)d\tau(p)}_{(II)} + \zeta \qquad (1)$$

where: $\mathbf{J}$ is a satellite measurement; ς is an unknown emissivity; $\mathbf{T(p)}$ is a desired temperature profile of pressure $\mathbf{p}$; $\tau(\mathbf{p})$ is a transmittance function of the atmosphere; $(.)_s$ designates value of any parameter on the underlying surface level; ξ - is an error of the measurement. In (1) the terms (I), (II) describe the upwelling and reflected downwelling radiation, respectively. (1) describes a model of "transparent cloudiness". Designating :

$$J_b = T(p_s)\tau(p_s) + \int_{\tau(p_s)}^{1} T(p)d\tau(p), \qquad J_r = -T(p_s)\tau(p_s) + \tau^2(p_s)\int_{\tau(p_s)}^{1} T(p)\tau^{-2}(p)d\tau(p), \qquad \varepsilon = 1-\varsigma,$$

let us transform (1) to the following form:

$$J = J_b + \varepsilon J_r + \xi \qquad (2)$$

(J_b corresponds to RTE with the "black" underlying surface ($\varsigma = 1$)). Let us estimate the order of accuracy of the procedure mentioned above for consecutive estimation of "emissivity than temperature". It follows from (2) that

$$\varepsilon = (J - J_{(1)b}) / J_{(1)r}, \qquad (3)$$

where are denoted: $(.)_1$ parameters of the first MSU channel; $J_{(1)b}(\overline{T}), J_{(1)r}(\overline{T})$ are solutions of the forward problems under the given temperature first guess $\mathbf{T(p)}$. To estimate the accuracy of the approximation (3), substitute (2) into (3) and then rewrite it for the desired temperature variation $\mathbf{T(p)} = \overline{\mathbf{T}}(\mathbf{p}) + \mathbf{x(p)}$:

$$\hat{\varepsilon} = \varepsilon + (J_{(1)b}(x) - \varepsilon J_{(1)r}(x) + \xi) / J_{(1)r}(\overline{T}) \qquad (4)$$

In (4) the second term describes the emissivity error estimation due to the uncertainty of temperature. In substituting (4) in (2) for other MSU channels we obtain $J_{(i)}(\delta\varepsilon) = (J_{(1)b}(x) + \varepsilon J_{(1)r}(x) + \xi) J_{(i)r}(\overline{T}+x) / J_{(1)r}(\overline{T})$. The latter means that the approximation error $J_{(i)}(\delta\varepsilon)$ is proportional to the variation $\mathbf{x(p)}$ with the weight $J_{(i)r}(\overline{T}+x)/J_{(1)r}(\overline{T})$, i.e. the disturbance of problem operator in (1) due to the emissivity estimation error $\delta\varepsilon$ generates error approximation correlating with the desired solution of the inverse problem and being proportional to one. It explains why we can degrade the first guess temperature profile located under cloud layer when sounding in heavy cloudy conditions, the only **MSU** measurements are actually used to sound temperature: **if the accuracy of the first guess of used surface temperature $\overline{T}(p_s)$ is significantly poorer than the accuracy of the used atmospheric temperature first guess $\overline{T}(p)$, $|x(p_s)| > |x(p)|$, then the surface temperature error $x(p_s)$ will be propagated by the emissivity estimate $\hat{\varepsilon}$ into other channels.** Such situation is not an exception, but rather the rule. It follows from the above discussion, that we must solve nonlinear problem (1), because its linear approximation has a low efficiency. Let us consider some approach for solving a non-linear problem. The inverse problem solution of (1) is known to be an ill-posed problem, as Fredholms' equation of the first kind. To be solved numerically, (1) must be regularized, i.e., we are approximating the initial problem (1) by some well-posed problem (Tikhonov 1977), whose solution is unique, stable, and corresponds to the solution (1) in some sense. It can been shown that the statistical estimators best correspond to the properties (1). We will consider the minimum variance estimator. Constructing the estimates of desired parameters $(\mathbf{X}',\varepsilon')$, we suggest that in (2) parameters satisfy the following conditions:

- $\mathbf{X}$ is a random value with known covariance matrix $\mathbf{R}$ and zero mean;
- ξ is a random value with known covariance matrix $\mathbf{S}$ and zero mean, not correlated with $\mathbf{x}$;
- ε is free control parameter from some known interval of interest.

Under these conditions we obtain the following estimate for $\mathbf{x(p)}$ (Rao 1965):

$$\hat{x} = R * A(\varepsilon)^T * (A(\varepsilon) * R * A(\varepsilon)^T + S)^{-1} * f(\varepsilon) \qquad (5)$$

where: $A(\varepsilon) = A[J(\varepsilon)]$ is a matrix of the algebraic system approximating (2) and $A(\varepsilon) = A + \varepsilon G$, $A = A[J_b]$, $G = G[J_r]$; $f(\varepsilon)$ is a measurement variation vector; $(.)^T$ and $(.)^{-1}$ are transposition and inversion operators, respectively. To estimate ε we will use the normal solution of (2) under some given mean temperature profile $T(p) = \overline{T}(p)$:

$$\hat{\varepsilon} = q(T) J_r(T)^T * S^{-1} * (J - J_b(T)) \ , \qquad (6)$$

where $q(\overline{T}) = 1/(J_r(\overline{T})^T * S^{-1} * J_r(\overline{T}))$ is a scalar variable. The error estimate (6) will be:

$$\delta\hat{\varepsilon} \approx q(\overline{T}) J_r(\overline{T})^T * S^{-1} * (A(\varepsilon^*) * x + \xi) \ , \qquad (7)$$

where ε^* is an unknown true value of emissivity. (7) is obviously more accurate than (3) due to the smoothing of the influence of (x, ξ). A joint system (5,6) can been solved recursively:

$$\left.\begin{aligned}
&T(n+1)=\bar{T}+x(n)\\
&q(n+1)=1/(J_r(T(n+1))^T*S^{-1}*J_r(T(n+1)))\\
&\varepsilon(n+1)=q(n+1)J_r(T(n+1))^T*S^{-1}*(J-J_b(T(n+1)))\\
&f(n+1)=J-J(\bar{T},e(n+1))\\
&x(n+1)=R*A(\varepsilon(n+1))^T*(A(\varepsilon(n+1))*R*A(\varepsilon(n+1))^T+S)^{-1}*f(n+1)
\end{aligned}\right\} \quad (8)$$

with the initial value $x^{(0)}=0$. Because the sequence of estimates in (8) is monotonic and bound, one will converge to some solutions $(\hat{x},\hat{\varepsilon})$.

II. The filtering of the impact of error estimation of the underlying surface emissivity into the temperature solution.

Let us examine the estimation $\hat{x}$ as the function of $\hat{\varepsilon}$, $\hat{x}=x(\hat{\varepsilon})$. Because estimation variation $\delta\varepsilon=\hat{\varepsilon}-\varepsilon^*$ is order $\frac{\|x\|}{\|\bar{T}\|}\leq 0.02$ of magnitude, we can linearize $\hat{x}=x(\hat{\varepsilon})$ in the point ε^*:

$$x(\hat{\varepsilon})=x(\varepsilon^*)+\left.\frac{dx(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=\varepsilon^*}\delta\varepsilon \quad (9)$$

Then by substituting (7),(8) for (9) and retaining linear terms of $\hat{x}=x(\hat{\varepsilon})$, we will obtain the following equation:

$$\hat{x}=x(\varepsilon^*)+q(\bar{T})\frac{d}{d\varepsilon}\left\{R*A(\varepsilon)^T*(A(\varepsilon)*R*A(\varepsilon)^T+S)^{-1}*f(\varepsilon)\right\}\Big|_{\varepsilon=\varepsilon^*}*J_r(\bar{T})^T*S^{-1}*A(\varepsilon^*)*x$$

It is worth noting that we can consider the derivatives of $\hat{x}(\varepsilon)$ with respect to ε due to the regularization of problem. Considering $x(\varepsilon^*)\cong\hat{x}(\varepsilon)$, then it follows from the equation above that the value of the linear interpolated estimate $\hat{x}(\varepsilon^*)$ depends on the current value $\hat{x}(\varepsilon)$ as

$$\hat{\hat{x}}(\varepsilon^*)=(E+L(\varepsilon^*))^{-1}*\hat{x}(\varepsilon) \quad , \quad (10)$$

where: E is the identity matrix and

$$L(\varepsilon^*)=q(\bar{T})\frac{d}{d\varepsilon}\left\{R*A(\varepsilon)^T*(A(\varepsilon)*R*A(\varepsilon)^T+S)^{-1}*f(\varepsilon)\right\}\Big|_{\varepsilon=\varepsilon^*}*J_r(\bar{T})^T*S^{-1}*A(\varepsilon^*). \quad (11)$$

To use (10), (11) we must know the value of ε^* , but with the estimation $\hat{\varepsilon}$ being the only known. Let us consider a substitution of the estimate $\hat{\varepsilon}$ into (11) instead of ε^*: look at Appendix containing an feature analysis of $L(\varepsilon)$ and $\delta\varepsilon\frac{dL(\varepsilon)}{d\varepsilon}$ operators , and where the relation $L(\varepsilon^*)\approx L(\hat{\varepsilon})$ is established, i.e.

$$L(\varepsilon^*)\approx -q(\bar{T})\cdot R*A(\varepsilon)^T*D(\varepsilon)^{-1}*J_r(\bar{T})*J_r(\bar{T})^T*S^{-1}*A(\varepsilon)\Big|_{\varepsilon=\hat{\varepsilon}} \quad (12)$$

It is worth noting that the calculation of the solution (10) can be simplified because $\mathrm{rank}\{L(\varepsilon)\}=1$ and by expanding preliminary solution $\{\hat{x}: \hat{x}=\hat{x}_V+\hat{\theta}\eta \,/\, L(\hat{\varepsilon})*\hat{x}_V=0\}$, where η is an arbitrary row of matrix $L(\hat{\varepsilon})$ from (12), $\eta^T*L(\hat{\varepsilon})*\eta=-\omega\ \eta^T*\eta=1,\ 0<\omega\leq 1, \hat{\theta}=\hat{x}^T*\eta$, where $\hat{\theta}$ is a *projection of* $\hat{x}$ *on vector* η : $\hat{\theta}=\hat{x}^T*\eta$.

Then the solution is defined as

$$\hat{\hat{x}}=\hat{x}_V+\hat{\theta}(1-\omega)^{-1}\eta \quad (13)$$

Equation (13) can be interpreted in the following way:

- in the solution $\hat{\mathbf{x}}$ there is a component η, such that the variation of the solution projection on η is analogous to the variation of emissivity $\delta\varepsilon$ in the sense of the variation of the right-hand side of the equation (1): $\{\delta \mathbf{f}(\mathbf{x} = \theta\eta \neq 0, \delta\varepsilon = 0) = \delta \mathbf{f}(\mathbf{x} = 0, \delta\varepsilon \neq 0)\}$;
- η describes the linear dependence of the two unknown parameters and (13) defines the value of solution projection on the vector l under the condition $\{\delta\varepsilon = 0\}$;
- vector $\hat{\mathbf{x}}_{\mathbf{V}}$ is a unique, stable component of the temperature solution (5) not depending on $\delta\varepsilon$.

Actually, the vector η in solution (5) describes the value of an additional error due to the model error generated by the "ambiguity" of equation (1). It follows from equation (9) that the variance of the perturbing signal due to the model error is approximately equal to $\mathbf{L} * \mathbf{R} * \mathbf{L}^{\mathbf{T}} = \omega^2 \cdot \eta * \mathbf{R} * \eta^{\mathbf{T}} \approx \omega^2 \cdot \sigma_{\mathbf{T_s}}^2$. The influence of the direction η on the accuracy of the solution $\mathbf{x}(\hat{\varepsilon})$ is explained by results of the numerical examination (12) for the measurement content of the MSU radiometer. Fig. 1 contains graphics of vector η, calculated for different kinds of atmospheric conditions and underlying surface types and diagonal matrix $\mathbf{R}$. As seen from the figures, the surface temperature is the principal variable, and there are two maxima in the atmospheric part of the solution component η. The eigenvalue ω as a function of the surface temperature variance $\sigma_{\mathbf{T_s}}^2$ is drawn on fig. 2.

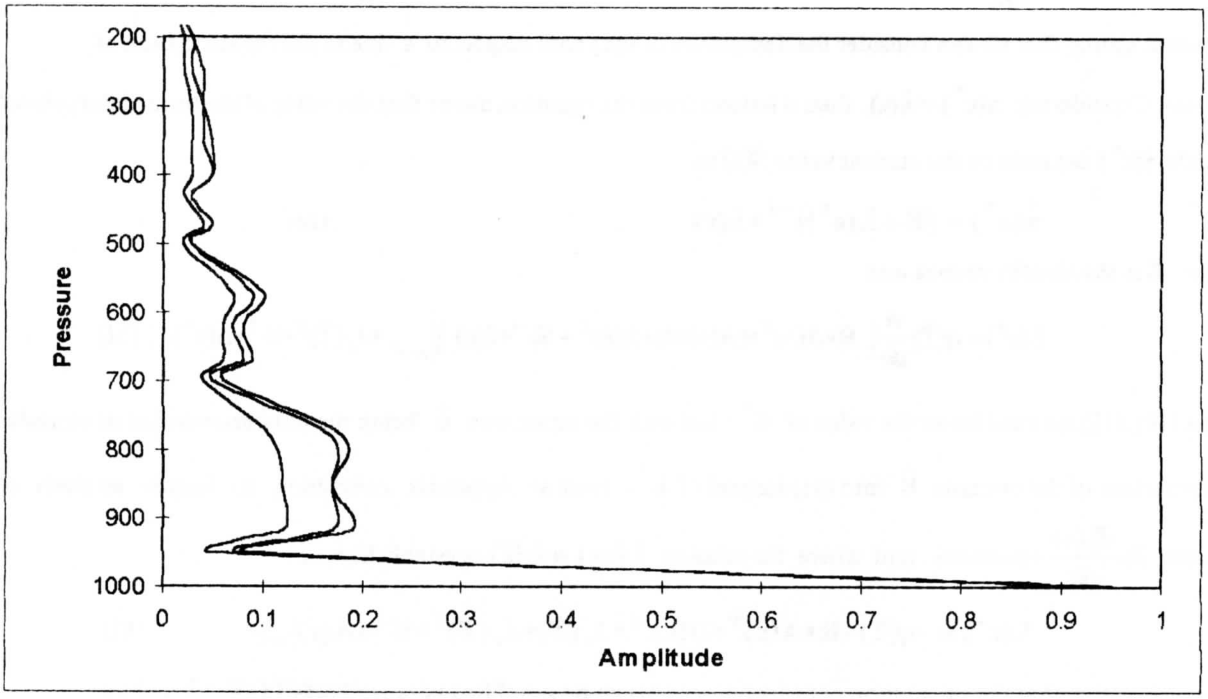

Fig.1 Vector η of vertical distribution of temperature estimation error due to emissivity estimation error for different kinds of atmospheric conditions and surface types

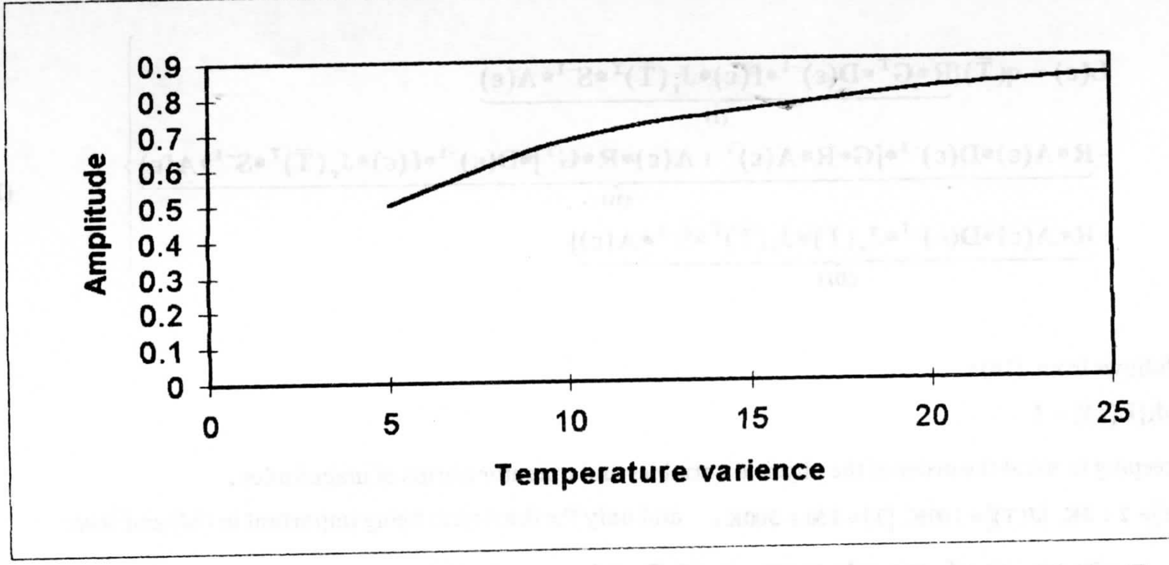

Fig.2 Dependance the weight of emissivity estimation error ω on the surface temperature variance $\sigma_{T_s}^2$

As seen in fig. 2, the uncertainties of surface temperature or the corresponding error of the emissivity estimate will propagate into the solution $\hat{\mathbf{x}}(\varepsilon)$ along the direction η, being proportional to the value $\sigma_{T_s}^2$. This means that in the atmosphere the value of the temperature error due to the disturbance of operator will be about **0.5-1.0 K** on the respective levels of the vector η under the variation of surface temperature of about **5. K**.

Finishing the consideration let us formulate an algorithm of simultaneous non-linear estimation of atmospheric temperature profile and surface emissivity, based upon the results of the analysis being carried out above:

1. Calculate the estimation (6) $\hat{\varepsilon} = \mathbf{q(T)J_r(T)^T * S^{-1} * (J - J_b(T))}$.

2. Calculate the estimation (5) $\hat{\mathbf{x}}(\varepsilon) = \mathbf{R} * \mathbf{A}(\varepsilon)^T * (\mathbf{A}(\varepsilon) * \mathbf{R} * \mathbf{A}(\varepsilon)^T + \mathbf{S})^{-1} * \mathbf{f}(\varepsilon)$ in two points of $\varepsilon = \hat{\varepsilon}$ and $\varepsilon = \hat{\varepsilon} + \delta\varepsilon$, where $|\delta\varepsilon| \leq .01$ is an arbitrary disturbance of the estimation.

3. Design vector $\eta = \hat{\mathbf{x}}(\hat{\varepsilon}) - \hat{\mathbf{x}}(\hat{\varepsilon} + \delta\varepsilon)$ and filter out the projection of the solution $\hat{\mathbf{x}}(\varepsilon)$ on it: $\hat{\hat{\mathbf{x}}}(\hat{\varepsilon}) = \hat{\mathbf{x}}(\hat{\varepsilon}) - \theta\eta$

Appendix

When considering the substitution $\hat{\varepsilon}$ from (7) into (11) and analyzing features of the operator $\mathbf{L}(\varepsilon)$, we use following designations and relations::

$$\frac{d\mathbf{A}}{d\varepsilon} = \mathbf{G}, \ \mathbf{D}(\varepsilon) = \mathbf{A}(\varepsilon) * \mathbf{R} * \mathbf{A}(\varepsilon)^T + \mathbf{S}, \quad \frac{d\mathbf{D}}{d\varepsilon} = \mathbf{G} * \mathbf{R} * \mathbf{A}(\varepsilon)^T + \mathbf{A}(\varepsilon) * \mathbf{R} * \mathbf{G}^T,$$

$$\frac{d\mathbf{D}(\varepsilon)^{-1}}{d\varepsilon} = -\mathbf{D}(\varepsilon)^{-1} * \frac{d\mathbf{D}(\varepsilon)}{d\varepsilon} * \mathbf{D}(\varepsilon)^{-1}, \quad \frac{d\mathbf{f}(\varepsilon)}{d\varepsilon} = -\mathbf{J}_r(\overline{\mathbf{T}}).$$

Then we obtain:

$$\left.\begin{array}{l} L(\varepsilon) = q(\bar{T})\{\underbrace{R*G^T*D(\varepsilon)^{-1}*f(\varepsilon)*J_r(\bar{T})^T*S^{-1}*A(\varepsilon)}_{(I)} \\ \underbrace{-R*A(\varepsilon)*D(\varepsilon)^{-1}*[G*R*A(\varepsilon)^T + A(\varepsilon)*R*G^T]*D(\varepsilon)^{-1}*f(\varepsilon)*J_r(\bar{T})^T*S^{-1}*A(\varepsilon)}_{(II)} \\ \underbrace{-R*A(\varepsilon)*D(\varepsilon)^{-1}*J_r(\bar{T})*J_r(\bar{T})^T*S^{-1}*A(\varepsilon)\}}_{(III)} \end{array}\right\} \quad (14)$$

As it follows from (14):

- $\mathbf{rank}\{L(\varepsilon)\} = 1$
- - keeping in mind the order of the physical variables and operator norms of magnitudes,

$\|r\| \approx \|x\| \approx 2 \div 5K$, $\|J(T)\| \approx 100K$, $\|T\| \approx 250 \div 300K$, and only the third term being important in (14) and then:

$$L(\varepsilon) \approx -q(\bar{T}) \cdot R*A(\varepsilon)^T*D(\varepsilon)^{-1}*J_r(\bar{T})*J_r(\bar{T})^T*S^{-1}*A(\varepsilon),\ \|L(\varepsilon)\| \le 1 .$$

Obviously, repeating the above transformation concerning the operator $\delta\varepsilon \dfrac{dL(\varepsilon)}{d\varepsilon}$, we obtain that

$\left\|\dfrac{dL(\varepsilon)}{d\varepsilon}\right\| \le 1$, $\left\|\delta\varepsilon \dfrac{dL(\varepsilon)}{d\varepsilon}\right\| \le \dfrac{\|x\|}{\|\bar{T}\|} \le .02$, proving the use of current estimation $\hat{\varepsilon}$ to calculate the operator

$L(\varepsilon\bullet) \approx L(\hat{\varepsilon})$.

III. References

Rao, C.R., 1965: Linear statistical inference and its applications.--New York, Viley.

Tikhonov, A.N., and V.Ya. Arsenin, 1977: Solutions of ill-posed problems.--Washington : Winston.

Weinreb, M.P.; H.E. Fleming; L.M. McMillin; and A.C. Neuendorffer, 1981: Transmittances for the TIROS Operational Vertical Sounder, NOAA Tech. Rep. NESS 85.

TECHNICAL PROCEEDINGS OF

THE NINTH INTERNATIONAL TOVS STUDY CONFERENCE

Igls, Austria

20-26 February 1997

Edited by

J R Eyre

Meteorological Office, Bracknell, U.K.

Published by

European Centre for Medium-range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, U.K.

May 1997