Radiative Transfer Models and their Adjoints

Paul van Delst

Overview

- Use of satellite radiances in Data Assimilation (DA)
- Radiative Transfer Model (RTM) components and definitions
- Testing the RTM components.
- Advantages/disadvantages

Use of satellite radiances in DA

- Adjust the model trajectory with data.
- Iteratively minimise the difference between a model prediction and data using a cost/penalty function, e.g.

$$J(\mathbf{X}) = (\mathbf{X} - \mathbf{X}_b)^T \mathbf{B}^{-1} (\mathbf{X} - \mathbf{X}_b) + (\mathbf{Y}_m - \mathbf{Y}(\mathbf{X}))^T (\mathbf{O} + \mathbf{F})^{-1} (\mathbf{Y}_m - \mathbf{Y}(\mathbf{X})) + J_c$$

  - X, X_b: Input state vector and background estimate
  - Y_m, $Y(X)$: Measurements and forward model
  - B, O, F: Error covariances of X_b, Y_m, and $Y(X)$
- Iteration step direction is determined from $Y(X)$ linearised about X_b,

$$\underbrace{\mathbf{Y}(\mathbf{X})}_{Lx1} = \underbrace{\mathbf{Y}(\mathbf{X}_b)}_{Lx1} + \underbrace{\mathbf{K}(\mathbf{X}_b)}_{LxK} \underbrace{(\mathbf{X} - \mathbf{X}_b)}_{Kx1}$$

- Where the $K(X_b)$ are the Jacobians (K-Matrix) of the forward model for the background state X_b,

$$K^l(X^k) = \left.\frac{\partial Y^l}{\partial X^k}\right|_{X = X_b}$$

RTM components and definitions (1)

- Forward (FWD) model. The FWD operator maps the input state vector, *X*, to the model prediction, *Y*, e.g. for predictor #11:

$$P_{11} = \frac{W}{T^2}$$

- Tangent-linear (TL) model. Linearisation of the forward model about X_b, the TL operator maps changes in the input state vector, δX, to changes in the model prediction, δY,

$$\delta P_{11} = \frac{\partial P_{11}}{\partial W}\delta W + \frac{\partial P_{11}}{\partial T}\delta T$$
$$= \frac{1}{T^2}\delta W - \frac{2W}{T^3}\delta T$$

Or, in matrix form:

$$\begin{bmatrix} \delta P_{11} \\ \delta W \\ \delta T \end{bmatrix}^{n} = \begin{bmatrix} 0 & \frac{1}{T^2} & -\frac{2W}{T^3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta P_{11} \\ \delta W \\ \delta T \end{bmatrix}^{n-1}$$

RTM components and definitions (2)

- Adjoint (AD) model. The AD operator maps in the reverse direction where for a given perturbation in the model prediction, δY, the change in the state vector, δX, can be determined. The AD operator is the transpose of the TL operator. Using the example for predictor #11 in matrix form,

$$\begin{bmatrix} \delta^* P_{11} \\ \delta^* W \\ \delta^* T \end{bmatrix}^{n-1} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{T^2} & 1 & 0 \\ -\frac{2W}{T^3} & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta^* P_{11} \\ \delta^* W \\ \delta^* T \end{bmatrix}^{n}$$

Expanding this into separate equations:

$$\delta^* T^{n-1} = -\frac{2W}{T^3} \delta^* P_{11}^n + \delta^* T^n$$

$$\delta^* W^{n-1} = \frac{1}{T^2} \delta^* P_{11}^n + \delta^* W^n$$

$$\delta^* P_{11}^{n-1} = 0$$

RTM components and definitions (3)

- K-Matrix (K) model. Consider a channel radiance vector, *R*, computed using a single surface temperature, T_{sfc},. For every channel, *l*,

$$R_l = B(\nu_l, T_{sfc}) \qquad \text{FWD}$$

$$\delta R_l = \frac{\partial B(\nu_l, T_{sfc})}{\partial T_{sfc}} \delta T_{sfc} \qquad \text{TL}$$

$$\delta^* T_{sfc} = \frac{\partial B(\nu_l, T_{sfc})}{\partial T_{sfc}} \delta^* R_l + \delta^* T_{sfc} \qquad \text{AD}$$

- This is not what you want for DA/retrievals since the sensitivity of each channel is accumlated in the final surface temperature adjoint variable. Simple solution:

$$\delta^* T_{l,sfc} = \frac{\partial B(\nu_l, T_{sfc})}{\partial T_{sfc}} \delta^* R_l \qquad \text{K}$$

- So in the RTM, the K-matrix code simply involves shifting all the channel independent adjoint code inside the channel loop. That's it.

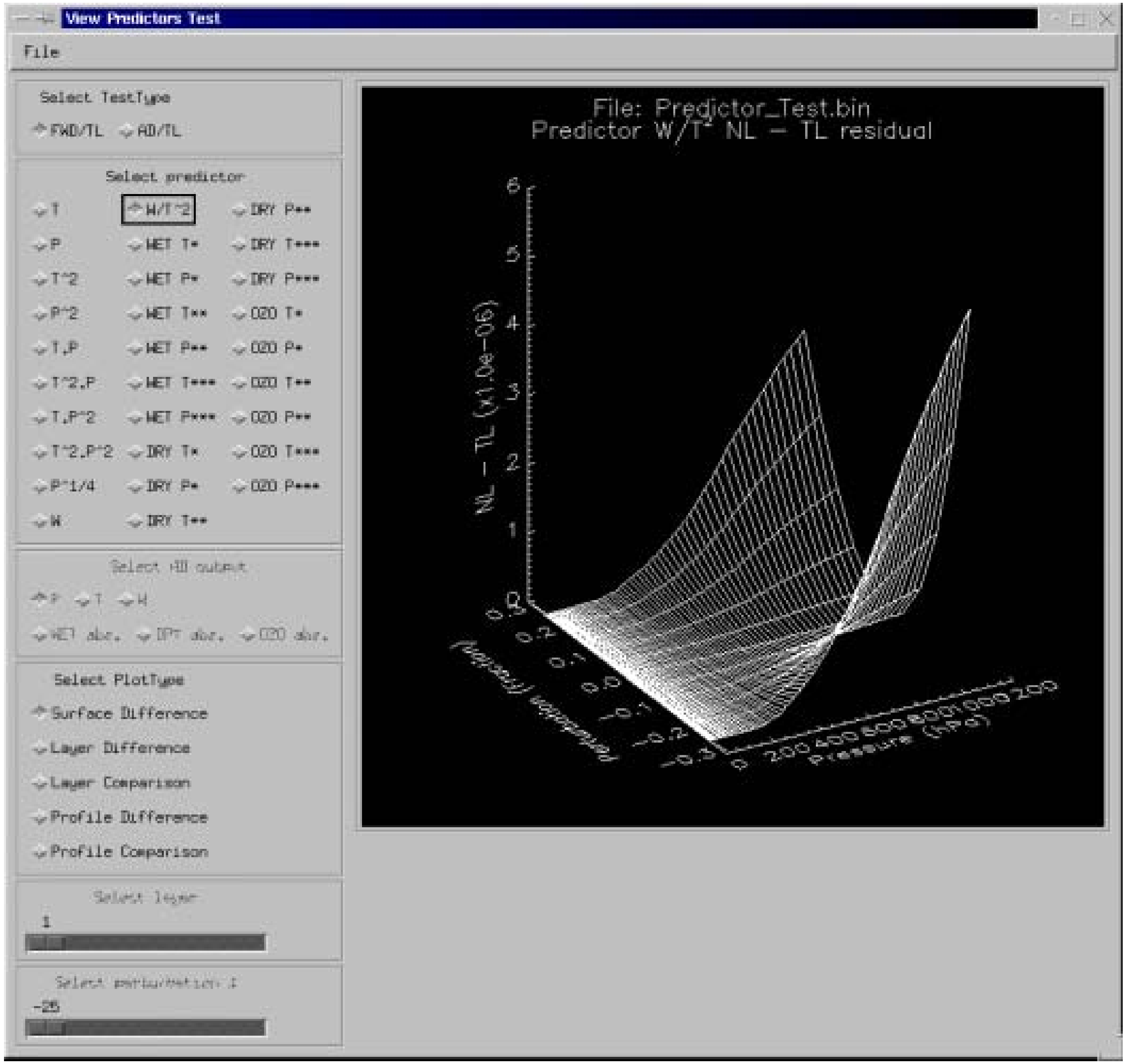
Testing the RTM components – FWD/TL

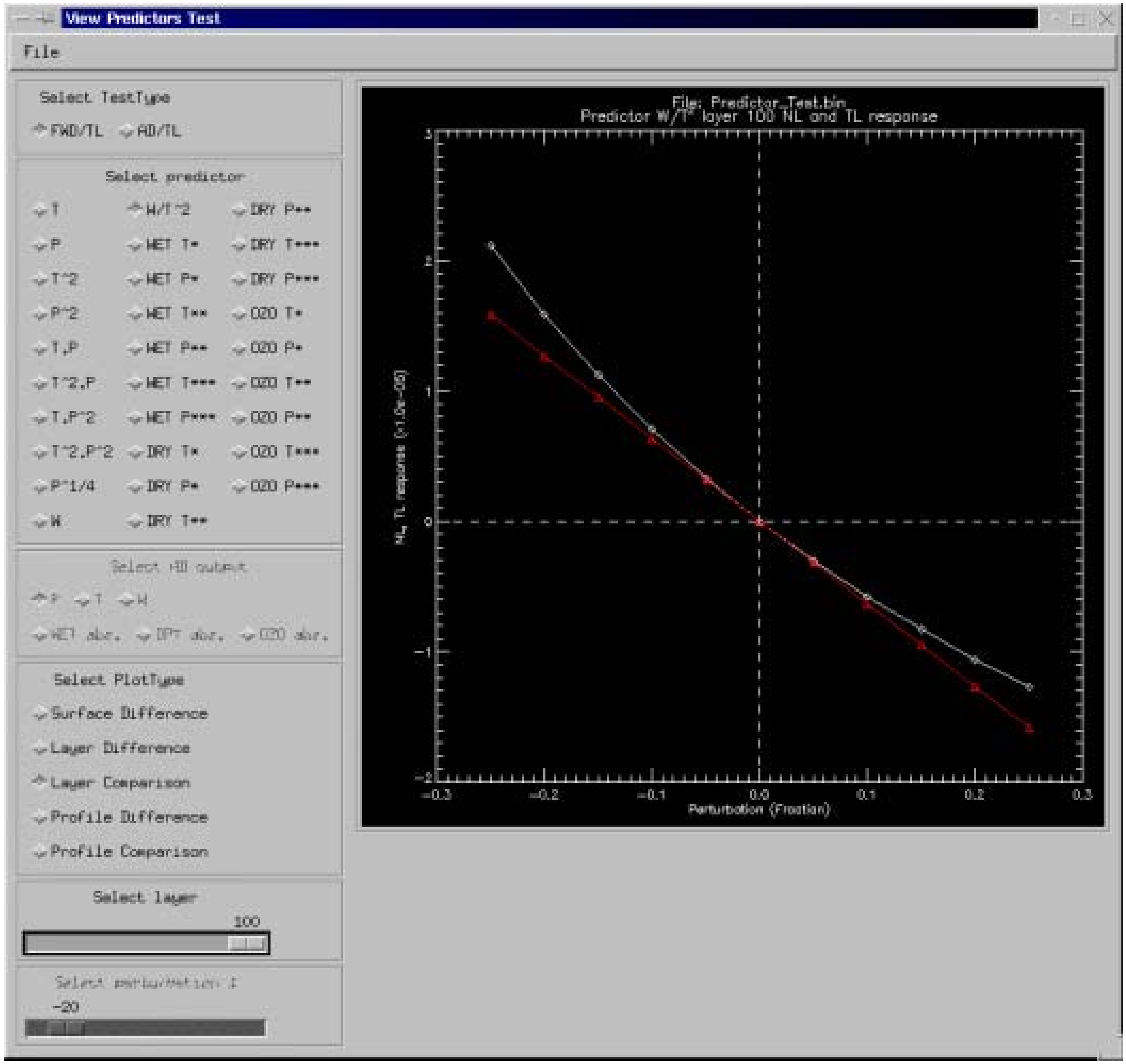
- Start with the assumption that the FWD component is in good shape (e.g. validated with observations, radiosonde matchups, etc).

- TL test against the forward model. Run the TL model with δX inputs varying from $-\Delta X \rightarrow 0 \rightarrow +\Delta X$ to give δY_{TL}. Run the FWD model with $X+\delta X$ inputs and difference from the zero perturbation case to get the non-linear result δY_{NL}.

- Inspect δY_{TL} and δY_{NL} as a function of δX. TL must be linear (d'oh) for all δX and tangent (d'oh^2) to the NL result at δX=0.

- Linearity of the TL result can be checked by numerical differentiation to give a constant for all δX. Numerical differentiation of NL result should give same value as TL at δX=0. But accuracy of numerical derivative is an issue if the perturbation resolution is low.


View Predictors Test
File
Select TestType
FWD/TL
AD/TL
Select predictor
T
W/T^2
DRY P**
P
WET T*
DRY T***
T^2
WET P*
DRY P***
P^2
WET T**
O2O T*
T,P
WET P**
O2O P*
T^2,P
WET T***
O2O T**
T,P^2
WET P***
O2O P**
T^2,P^2
DRY T*
O2O T***
P^1/4
DRY P*
O2O P***
W
DRY T**
Select PlotType
Surface Difference
Layer Difference
Layer Comparison
Profile Difference
Profile Comparison
1
-25
File: Predictor_Test.bin
Predictor W/T^2 NL − TL residual
NL − TL (X1.0e-06)
Pressure (hPa)


View Predictors Test
File
Select TestType
FWD/TL
AD/TL
Select predictor
T
W/T^2
DRY P**
P
WET T*
DRY T***
T^2
WET P*
DRY P***
P^2
WET T**
O20 T*
T,P
WET P**
O20 P*
T^2,P
WET T***
O20 T**
T,P^2
WET P***
O20 P**
T^2,P^2
DRY T*
O20 T***
P^1/4
DRY P*
O20 P***
W
DRY T**
Select PlotType
Surface Difference
Layer Difference
Layer Comparison
Profile Difference
Profile Comparison
Select layer
100
-20
File: Predictor_Test.bin
Predictor W/T^2 layer 100 NL and TL response
NL, TL response (x1.0e-05)
Perturbation (Fraction)
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
3
2
1
0
-1
-2

Testing the RTM components – TL/AD

- Assume the FWD model input vector, *X*, has *K* elements and the output vector has *L* elements,

$$\mathbf{X} = [X_1, X_2, X_3, \cdots, X_K]$$
$$\mathbf{Y} = [Y_1, Y_2, Y_3, \cdots, Y_L]$$

- Run the TL model $j = 1$ to K times with input,

$$\delta X_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

saving the δY vector output each run to give a LxK matrix, **TL**.

- Run the AD model $j = 1$ to L times with input,

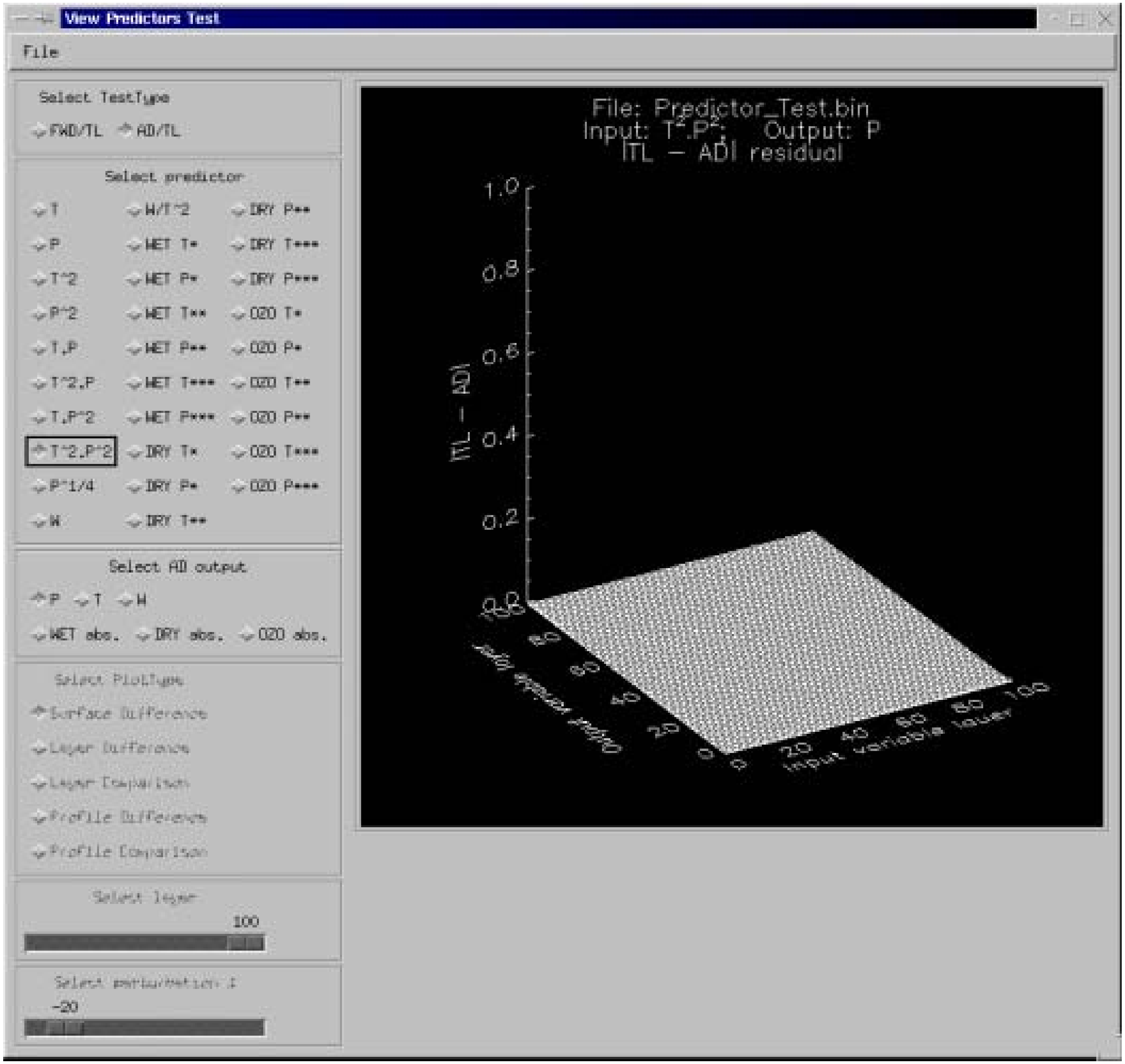
$$\delta^* Y_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad \boldsymbol{\delta}^* \mathbf{X} = \mathbf{0}$$

saving the $\delta^* X$ vector output each run to give a KxL matrix, **AD**. Then, to within numerical precision,

$$\mathbf{TL} - \mathbf{AD}^T = \mathbf{0}$$

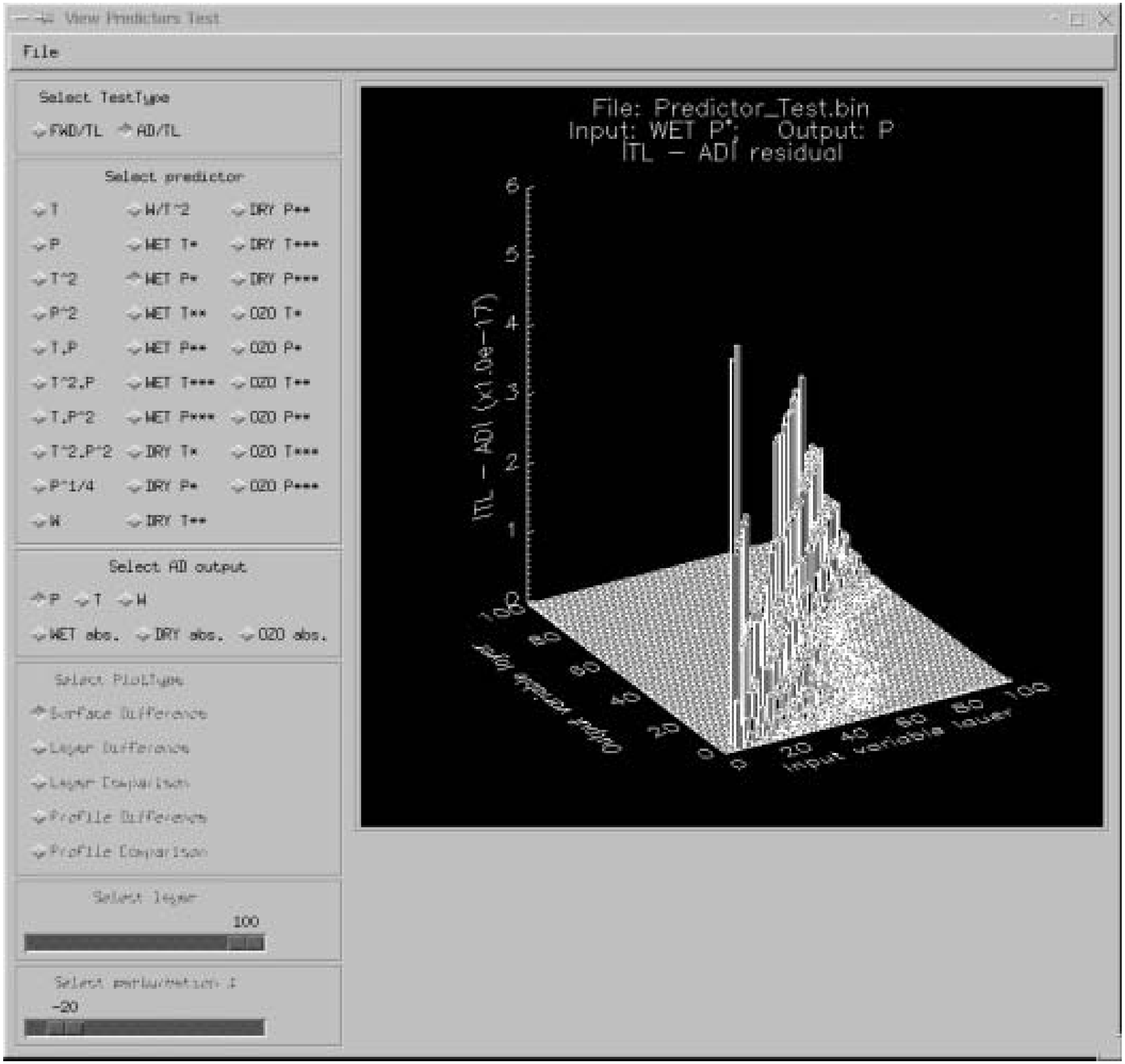
Input: T^2P^2

Output: P



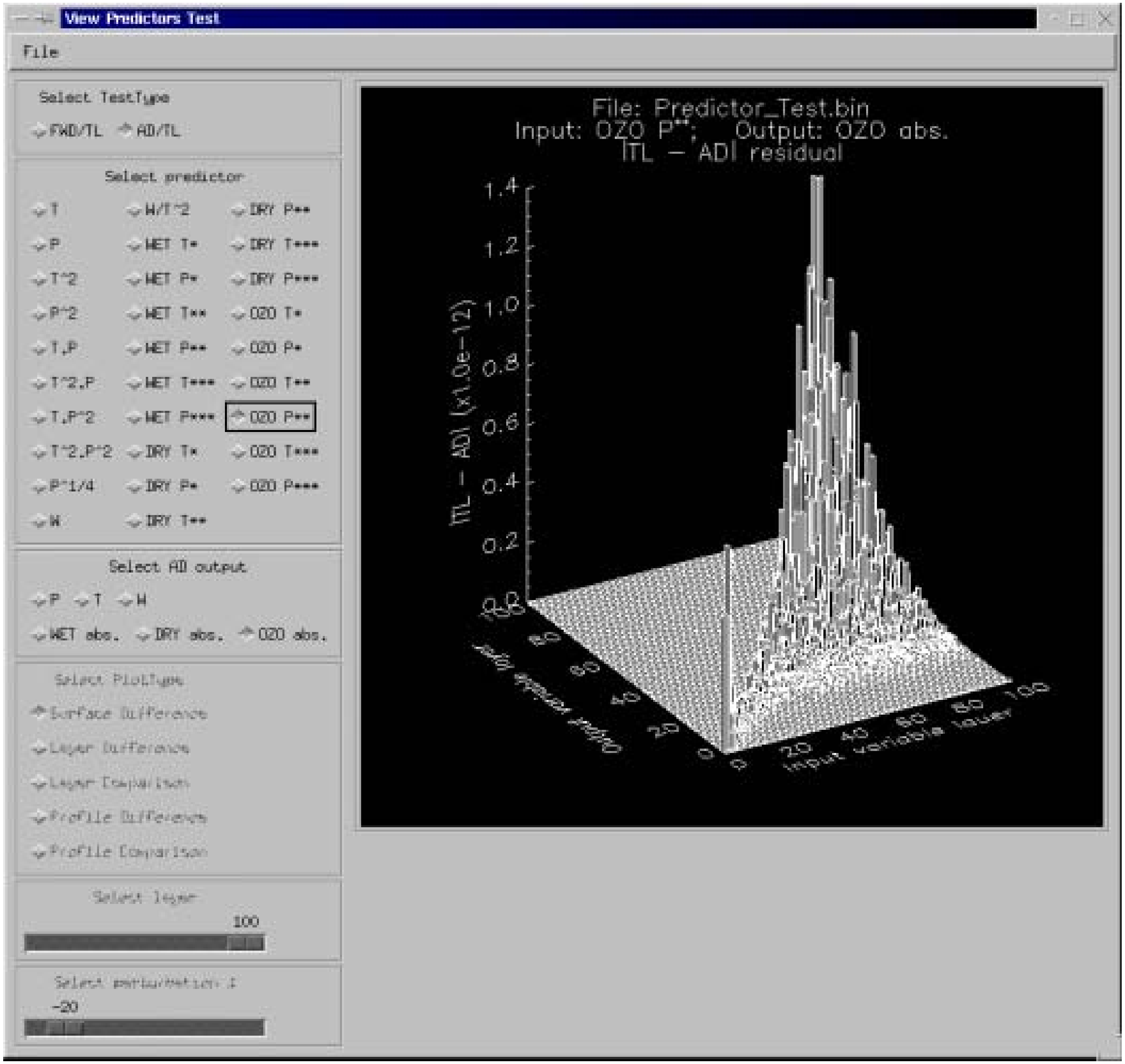
Input: Wet P^*

Output: P



Input: Ozo P^{**}

Output: Ozo A



Advantages/Disadvantages

- Advantages
  - Adjoint method produces Jacobians fully consistent with the forward model.
  - Well defined set of rules for applying method to code.
  - Code tests are straightforward and definitive – particularly for the TL/AD test.
  - Easy to incorporate model changes, improvements, additions, etc.
  - Good for sensitivity analyses. TL used to investigate impact of small disturbances, AD can be used to investigate origin of the anomaly.
- Disadvantages
  - Complexity. Compared to finite differences (if one can live with using them), adjoint coding can be a bit of a brain teaser.
  - *Very* easy to produce code slower than a snail in a straitjacket. Up front code design is an important step.
  - Have to be careful when vectorising and optimising code.