I am sorry, for personal reasons, to limit this review to the first part of the rather long document. I hope that my text is still useful for the review process, helping to improve this valuable work, thus making it more understandable.

General comment:
According to the Abstract, the paper should become a comprehensive description. But the present version is far from this goal, especially in the introductory parts.
More fundamental in microwave radiometry than radiative transfer is Kirchhoff's Law for thermal radiation, taking thermodynamic laws into account. This aspect is missing. When a reader comes to Line 71, she or he may not understand the end of the sentence on the capability of the model. The same is true on Line 81 with respect to the unexplained advantages.

Specific comments:
1) On p. 3, Equation (1) is stated to describe isotropic scattering. This is not true. Scattering is ignored because the scattering phase function was put to zero in this approximation. It means that power conservation is violated. The same is true for Kirchhoff's Law because the sum of the coefficients of the three temperatures on the right-hand side differs from 1. A comprehensive treatment cannot omit these facts. Still, a reconciliation with these problems is possible even in the zero-order approximation, see e.g. Chapter 4.2 in *Thermal Microwave Radiation: Applications for Remote Sensing*, IET (2006). The mentioned chapter also contains a comparison of the zero-order model with single scattering and with multiple scattering in the two-stream approach. A reference to this overview should at least be included, and the authors should be aware of this work.

2) Section 2, text to Equation (2): "observed in direction u " is misleading: u is the direction of the antenna axis, i.e. the center of the main beam. Instead of "$ds(u')$ ", I would prefer "$d^2s(u')$ ", to more clearly indicate the integral over two dimensions. The integral is to be taken over all directions (4π) and must be divided by 4π to get the proper normalisation of the antenna gain.

3) Section 2.1, Line 140: Assumption of azimuthal symmetry (better: azimuthal isotropy). The motivation is missing. It is a result of the assumption of horizontal, plane-parallel layers, which was obvious for snowpacks. In case of vegetation, the situation is more delicate. The reason why it is still useful here, is the rather low optical depth of vegetation at L Band. Then the lateral averaging of radiation over a large footprint leads to similar results as the application of the forward model on areal averages of the vegetation parameters. If the vegetation-optical depth gets large, the effective averaging becomes challenging, and azimuthal isotropy may not be true.

4) Further, on Line 153, the use of delta functions means specular reflections. Specular reflection and refraction occur at layer interfaces within snowpacks. Note that volume scattering by snow grains was assumed to be isotropic within each snow layer of Wiesmann and Mätzler (1999), leading to the six-flux equations by separating trapped radiation from radiation that can escape the snowpack. The six fluxes simplified to the two-stream model by the assumption of the plane-parallel snow layers. Specular reflection was assumed at layer interfaces, only. In case of vegetation, delta functions are not plausible. Still, the two-stream approach is more realistic and more physical than the zero-order model.