Article

The SSR Brightness Temperature Increment Model Based on a Deep Neural Network

Zhongkai Wen 1,2, Huan Zhang 1,2, Weiping Shu 2, Liqiang Zhang 2, Lei Liu 2, Xiang Lu 2, Yashi Zhou 2, Jingjing Ren 2, Shuang Li 1* and Qingjun Zhang 2,*

1 College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; twenzhongkai@spacechina.com (Z.W.)
2 Institute of Remote Sensing Satellite, China Academy of Space Technology, Beijing 100094, China
* Correspondence: zhangqingjun@spacechina.com

Abstract: The SSS (sea surface salinity) is an important factor affecting global climate changes, sea dynamic environments, global water cycles, marine ecological environments, and ocean carbon cycles. Satellite remote sensing is a practical way to observe SSS from space, and the key to retrieving SSS satellite products is to establish an accurate sea surface brightness temperature forward model. However, the calculation results of different forward models, which are composed of different relative permittivity models and SSR (sea surface roughness) brightness temperature increment models, are different, and the impact of this calculation difference has exceeded the accuracy requirement of the SSS inversion, and the existing SSR brightness temperature increment models, which primarily include empirical models and theoretical models, cannot match all the relative permittivity models. In order to address this problem, this paper proposes a universal DNN (deep neural network) model architecture and corresponding training scheme, and provides different SSR brightness temperature increment models for different relative permittivity models utilizing DNN based on offshore experiment data, and compares them with the existing models. The results show that the DNN models perform significantly better than the existing models, and that their calculation accuracy is close to the detection accuracy of a radiometer. Therefore, this study effectively solves the problem of SSR brightness temperature correction under different relative permittivity models, and provides a theoretical support for high-precision SSS inversion research.

Keywords: sea surface salinity; sea surface brightness temperature forward model; relative permittivity model; SSR brightness temperature increment model; deep neural network

1. Introduction

The SSS is one of the three fundamental components of physical oceanography, and also an important factor affecting global climate changes, sea dynamic environments, global water cycles, marine ecological environments, and ocean carbon cycles. The SMOS satellite [1], Aquarius satellite [2], and SMAP satellite [3] have all demonstrated the feasibility of measuring SSS from space using an L-band microwave radiometer [4,5], and the key to retrieving SSS satellite products is to establish an accurate sea surface brightness temperature forward model [6]. However, the sea surface brightness temperatures calculated with different forward models, which are composed of different relative permittivity models and SSR brightness temperature increment models, are different, and the impact of this calculation difference has exceeded the accuracy requirement of the SSS inversion [7,8]. Therefore, it is necessary to research the SSR brightness temperature increment model under different relative permittivity models.

Empirical models and theoretical models are two types of SSR brightness temperature increment models that are frequently utilized. Empirical models, which are straightforward and convenient to use but typically only applicable to particular payloads and spatiotemporal factors, are fitted using a substantial amount of observational data, such as the
Hollinger model [9], Camps 2003 model [10], WISE model [11], Gabarró model [12], Yueh 2013 model [13], etc. The theoretical model is based on the interaction between rough sea surfaces and electromagnetic waves, and the radiation model of the sea surface can be obtained by solving the theoretical solution of the electromagnetic wave scattering field on the sea surface, such as the Kirchhoff Approximation (KA) [14], Small Perturbation Method (SPM) [15], Two-Scale Model (TSM) [16], Small-Slope Approximation (SSA) [17], Integral Equation Method (IEM) [18], Advanced Integral Equation Method (AIEM) [19], etc. However, both types of models cannot match all the relative permittivity models, and can only obtain good results under the specific relative permittivity model, while the errors under other relative permittivity models are large. Therefore, different SSR brightness temperature increment models should be provided for different relative permittivity models to make the calculation results of different forward models as close to the actual sea surface situation as possible. As a result, the SSR brightness temperature increment models corresponding to different relative permittivity models are provided in this paper based on a deep neural network and offshore experiment data.

The rest of this paper is as follows: In Section 2, an offshore observation experiment is discussed, and the offshore experiment data were measured and processed. In Section 3, different SSR brightness temperature increment models for different relative permittivity models are constructed based on a deep neural network and the offshore experiment data. Finally, the discussion and conclusions are given in Sections 4 and 5.

2. Offshore Observation Experiment and Data Processing

2.1. Offshore Observation Experiment

The 1.4~1.427 GHz frequency band has been accepted internationally for retrieving SSS [20], and the sea surface brightness temperature in this band is highly sensitive to SSS and can realize all-time and all-weather observation. This frequency band is also protected by international treaties and generally does not experience interference from human signals.

An experimental study was conducted on an offshore platform (as shown in Figure 1) from 11 August 2022 to 1 September 2022, with the observation platform situated about 30 km north of Yangma Island in Yantai, China (121.622°E, 37.695°N, as shown in Figure 2).

![Figure 1. Offshore platform.](image-url)
In the experiment, the sea surface brightness temperature was measured at different incident angles through a radiometer; the SST and SSS were synchronously obtained through the CTD (conductivity, temperature, and depth); the atmospheric pressure, atmospheric temperature, atmospheric relative humidity, wind speed, and wind direction were synchronously obtained through the automatic meteorological station. The primary technical indicators of the equipment are shown in Tables 1–3.

**Table 1.** Primary technical indicators of the radiometer.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Band</td>
<td>L</td>
</tr>
<tr>
<td>Central Frequency</td>
<td>1.415 GHz</td>
</tr>
<tr>
<td>3 dB Bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Polarization</td>
<td>H/V</td>
</tr>
<tr>
<td>Spatial Resolution</td>
<td>≤20°</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>≤0.1 K@1 s</td>
</tr>
<tr>
<td>Integration Time</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Primary technical indicators of the CTD.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Measurement Range</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>0–35 °C</td>
<td>0.002 °C</td>
</tr>
<tr>
<td>Conductivity</td>
<td>0–7 S/m</td>
<td>0.0003 S/m</td>
</tr>
</tbody>
</table>

**Table 3.** Primary technical indicators of the automatic meteorological station.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Measurement Range</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric Pressure (hPa)</td>
<td>850–1050</td>
<td>±1</td>
</tr>
<tr>
<td>Atmospheric Temperature (°C)</td>
<td>−50~50</td>
<td>±0.3</td>
</tr>
<tr>
<td>Atmospheric Relative Humidity (%)</td>
<td>0~100</td>
<td>±3</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0~70</td>
<td>±1</td>
</tr>
<tr>
<td>Wind Direction (°)</td>
<td>0~360</td>
<td>±10</td>
</tr>
</tbody>
</table>

To reduce the impact of the sun and RFI, the entire observing process was conducted with our backs to the sun and not towards the land, and the installation diagram of observation platform equipment is as follows (Figure 3):
Remote Sens. 2023, 15, x FOR PEER REVIEW 4 of 19

Figure 3. Installation diagram of observation platform equipment.

2.2. Sea Salinity Data Processing

In the experiment, the CTD directly measures the conductivity of seawater, and it needs to be converted into seawater salinity using the following equation (2 psu ≤ S ≤ 42 psu, −2 °C ≤ T ≤ 35 °C) [21,22].

\[
S = \frac{5}{i=0} a_i R_{T_i}^{1/2} + \frac{T - 15}{1 + 0.0162(T - 15)} \sum_{i=0}^{5} b_i R_{T_i}^{1/2}
\]  

(1)

where \(T\) (°C) is the temperature of seawater, \(R_T\) is the ratio of the conductivity at a standard atmospheric pressure and a temperature of \(T\) (the superscript represents atmospheric pressure), and the coefficients of Equation (1) are shown in Table 4.

Table 4. Coefficients of Equation (1).

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_0)</td>
<td>0.008</td>
<td>(b_0)</td>
<td>0.0005</td>
</tr>
<tr>
<td>(a_1)</td>
<td>−0.1692</td>
<td>(b_1)</td>
<td>−0.0056</td>
</tr>
<tr>
<td>(a_2)</td>
<td>25.3851</td>
<td>(b_2)</td>
<td>−0.0066</td>
</tr>
<tr>
<td>(a_3)</td>
<td>14.0941</td>
<td>(b_3)</td>
<td>−0.0375</td>
</tr>
<tr>
<td>(a_4)</td>
<td>−7.0261</td>
<td>(b_4)</td>
<td>0.0636</td>
</tr>
<tr>
<td>(a_5)</td>
<td>2.7081</td>
<td>(b_5)</td>
<td>−0.0144</td>
</tr>
<tr>
<td>sum</td>
<td>35</td>
<td>sum</td>
<td>0</td>
</tr>
</tbody>
</table>

2.3. Wind Speed Data Processing

The wind speed data used in SSS inversion is generally the wind speed at a 10 m altitude; however, the installation height of the automatic meteorological station cannot fully meet this requirement. Therefore, the following equation needs to be used to convert the measured wind speed data into the wind speed at a 10 m altitude.

\[
U_z = \frac{U_*}{0.4} \left( \ln \left( \frac{6.84 \times 10^{-5} + 4.28 \times 10^{-3} U_*^2}{U_*} - 4.43 \times 10^{-4} \right) \right)
\]  

(2)

where \(U_z\) (m/s) is the wind speed at \(z\) (m) altitude, and \(U_*\) is the wind friction velocity.

2.4. Removing the Effects of Atmospheric Radiation and Cosmic Radiation

Figure 4 depicts the complete transmission process of sea surface radiation to the offshore platform radiometer.
The sea surface brightness temperature that the offshore platform radiometer received was influenced by SSR, sea surface foam, the atmosphere, and cosmic radiation [23,24]. After removing all RFI-contaminated data and the data with large fluctuations, the brightness temperature data that the radiometer measured can be expressed as follows [25,26]:

\[
T_{BP,MR} = T_{BP,SSR+Foam} + (1 - \epsilon_{p,SSR+Foam}) \cdot (T_{BD} + \Gamma T_{B,COS})
\]  

(3)

where \( T_{BP,MR} \) is the \( p \) polarization brightness temperature received by the radiometer (\( p = H, V \) that stand for the horizontal polarization and vertical polarization of electromagnetic waves, respectively), \( T_{BP,SSR+Foam} \) and \( \epsilon_{p,SSR+Foam} \) are the \( p \) polarization sea surface brightness temperature and the \( p \) polarization sea surface emissivity under the impact of SSR and sea surface foam, respectively, \( T_{BD} \) is the atmospheric downward radiation, \( \Gamma \) is the atmospheric transmittance, and \( T_{B,COS} = 3.7 \text{ K} \) [27] is the cosmic radiation brightness temperature (there are also some papers that take \( T_{B,COS} = 2.73 \text{ K} \) [26,28,29], and this article will explain the reason for not taking this value in subsequent sections). The equations for removing the effects of atmospheric radiation and cosmic radiation are as follows [30]:

\[
T_{BP,SSR+Foam} = \frac{T_{BP,MR} - (T_{BD} + \Gamma T_{B,COS})}{\text{SST} - (T_{BD} + \Gamma T_{B,COS})} \cdot \text{SST}
\]  

(4)

\[
T_{BD} = \sec \theta \int_0^\infty [k_{O_2}(z) + k_{H_2O}(z)] \cdot T_a(z) \cdot \exp(-\tau \sec \theta)dz
\]  

(5)

\[
\Gamma = \exp[-\sec \theta \cdot \tau]
\]  

(6)

\[
\tau = \int_0^h [k_{O_2}(z) + k_{H_2O}(z)]dz
\]  

(7)

\[
k_{O_2}(z,f) = 1.1 \times 10^{-2} f^2 \left[ \frac{P_a(z)}{1013} \right] \cdot \left[ \frac{300}{T_a(z)} \right]^\gamma \left[ \frac{1}{(f-f_0)^2 + \gamma^2} + \frac{1}{f^2 + \gamma^2} \right]
\]  

(8)

\[
\gamma = \gamma_0 \left( \frac{P_a}{1013} \right) \left( \frac{300}{T_a} \right)^{0.85}
\]  

(9)

\[
\gamma_0 = \begin{cases} 
0.59, P_a \geq 333 \\
0.59 \cdot [1 + 0.0031 \cdot (333 - P_a)], 25 \leq P_a < 333 \\
1.18, P_a \leq 25
\end{cases}
\]  

(10)

Figure 4. Transmission process of sea surface radiation to the offshore platform radiometer.
where \( \text{SST} \) is the sea surface temperature in Kelvin, \( \theta \) is the incident angle of the radiometer, \( k_0(z) \) is the atmospheric attenuation coefficient (\( a \) can be taken as \( \text{O}_2 \) and \( \text{H}_2\text{O} \), representing oxygen and water vapor, respectively), \( z \) is the altitude in kilometers, \( \tau \) is atmospheric optical depth, \( h \) is the upper boundary of the atmosphere (for most microwave remote sensing of the Earth, and certainly at the L-band, it is only the lower 30 km of the atmosphere that needs to be taken into consideration), \( f \) is the frequency of the radiometer (1.415 GHz), \( f_0 \) is the absorption line frequency (60 GHz), \( \gamma \) (GHz) is the line width parameter, \( \gamma_0 \) is the nonresonant line width, \( k(f, 22) \) is the absorption coefficient for the 22.355 GHz spectral line, \( k_r(f) \) is the residual absorption coefficient, \( \gamma_1 \) (GHz) is the line width parameter, \( T_a(z) \), \( P_a(z) \), and \( \rho_v(z) \) are the temperature profile, the pressure profile, and the water vapor density profile, respectively, \( T_0 \) is the atmospheric temperature in Kelvin at sea level, \( T_a(11) \) is the atmospheric temperature in Kelvin at \( z = 11 \) km, \( P_0 \) is the atmospheric pressure in millibars at sea level, \( \rho_0 \) is the water vapor density in \( \text{g/m}^3 \) at sea level, and \( CH \) is the scale height, which is typically chosen to be between 2 and 2.5 km (in this article, \( CH = 2.25 \) km was taken).

### 2.5. Removing the Effects of Sea Surface Foam and Flat Sea Surface Brightness Temperature

After removing the effects of atmospheric radiation and cosmic radiation from the brightness temperature data received by the radiometer, the processed data can be expressed as follows [31,32]:

\[
T_{BP,SSR+Foam} = T_{BP,Sur} \cdot (1 - F_r) + F_r \cdot T_{BP,Foam} = \left(T_{BP,Flat} + \Delta T_{BP,SSR}\right) \cdot (1 - F_r) + F_r \cdot T_{BP,Foam} \tag{18}
\]

where \( T_{BP,Sur} \) is the \( p \) polarization sea surface brightness temperature under the impact of SSR, \( F_r \) is the whitecap coverage rate, \( T_{BP,Foam} \) is the \( p \) polarization brightness temperatures of the foam-covered sea surface, \( T_{BP,Flat} \) is the \( p \) polarization flat sea surface brightness temperature, and \( \Delta T_{BP,SSR} \) is the \( p \) polarization brightness temperature increment generated by the SSR. The equation for removing the effects of sea surface foam and flat sea surface brightness temperature is as follows:

\[
\Delta T_{BP,SSR} = \frac{T_{BP,SSR+Foam} - F_r \cdot T_{BP,Foam}}{1 - F_r} - T_{BP,Flat} \tag{19}
\]
$F_r$ can be calculated using the following equation [33]:

$$F_r = 1.95 \times 10^{-5} U_{10}^{0.55} \exp(0.0861 \Delta T)$$

where $U_{10}$ is the wind speed at a 10 m altitude, and $\Delta T = SST - T_0$ is the sea air temperature difference.

$T_{BH, Foam}$ can be calculated using the following equations [34]:

$$T_{BH, Foam} = (208 + 1.29f) \cdot \left(1 - 1.748 \times 10^{-3} \theta - 7.336 \times 10^{-5} \theta^2 + 1.044 \times 10^{-7} \theta^3\right)$$

$$T_{BV, Foam} = (208 + 1.29f) \cdot \left(1 - 9.946 \times 10^{-4} \theta + 3.218 \times 10^{-5} \theta^2 - 1.187 \times 10^{-6} \theta^3 + 7 \times 10^{-20} \theta^{10}\right)$$

where $f$ is the frequency of the radiometer, and $\theta$ is the incident angle of the radiometer.

$T_{BP, Flat}$ can be calculated using the following equations [35–37]:

$$T_{BP, Flat} = \left(1 - r_{p, Flat}\right) \cdot SST$$

$$r_{H, Flat} = \left|\frac{\cos \theta - \sqrt{\epsilon_r - \left(\frac{\omega}{c}\right)^2 \sin^2 \theta}}{\cos \theta + \sqrt{\epsilon_r - \left(\frac{\omega}{c}\right)^2 \sin^2 \theta}}\right|^2$$

$$r_{V, Flat} = \left|\frac{\epsilon_r \cos \theta - \sqrt{\epsilon_r - \left(\frac{\omega}{c}\right)^2 \sin^2 \theta}}{\epsilon_r \cos \theta + \sqrt{\epsilon_r - \left(\frac{\omega}{c}\right)^2 \sin^2 \theta}}\right|^2$$

where $r_{p, Flat}$ is the $p$ polarization flat sea surface reflectivity, $\theta$ is the incident angle of the radiometer, $n = n' - jn''$ is the complex refractive index of seawater, and $\epsilon_r$ is the relative permittivity at an infinite frequency, and $\epsilon_\infty$ is the relative permittivity at an infinite frequency, $\epsilon_s$ is the static relative permittivity, $\omega$ is the radian frequency with $f$ in hertz, $\tau$ is the relaxation time in seconds, $\alpha$ is an empirical parameter that represents the distribution of relaxation times, $\sigma$ is the ionic conductivity in mhos/meter, $\epsilon_0 \approx 8.85419 \times 10^{-12}$ is the permittivity of free space in farads/meter, $\epsilon_1$ is the relative permittivity at an intermediate frequency, $\tau_1$ and $\tau_2$ are the first and second debye relaxation times in seconds, and $f_1$ and $f_2$ are the first and second debye relaxation frequencies in hertz.

The parameters of the Cole–Cole equation and the double debye equation can be expressed as a function of $SST(T)$ and $SSS(S)$ with the L-band relative permittivity model, such as KS (1977) [40], ModKS (2000) [41], BA (2003) [42], BA (2004) [43], GW (2017) [44], and GW (2021) [45] can all be used to depict the parameters in the Cole–Cole equation, and MW (2004) [46], FASTEM (2011) [47], and MW (2012) [48] can all be used to depict the parameters in the double debye equation. By way of illustration, the MW (2004) depicts the parameters in the double debye equation as the following functions of $T$ and $S$:

$$\epsilon_s(T, S) = \epsilon_s(T, 0) \cdot \exp\left[\epsilon_0 S + b_1 S^2 + b_2 TS\right]$$
\[ \varepsilon_s(T,0) = \frac{3.70886 \times 10^4 - 8.2168 \times 10^1 T}{4.21854 \times 10^2 + T} \]  
(29)

\[ f_1(T,S) = f_1(T,0) \cdot \left[ 1 + S \cdot \left( b_3 + b_4 T + b_5 T^2 \right) \right] \]  
(30)

\[ f_1(T,0) = \frac{45 + T}{a_3 + a_4 T + a_5 T^2} \]  
(31)

\[ \varepsilon_1(T,S) = \varepsilon_1(T,0) \cdot \exp \left[ b_6 S + b_7 S^2 + b_8 T S \right] \]  
(32)

\[ \varepsilon_1(T,0) = a_0 + a_1 T + a_2 T^2 \]  
(33)

\[ f_2(T,S) = f_2(T,0) \cdot \left[ 1 + S \cdot (b_9 + b_{10} T) \right] \]  
(34)

\[ f_2(T,0) = \frac{45 + T}{a_8 + a_9 T + a_{10} T^2} \]  
(35)

\[ \varepsilon_\infty(T,S) = \varepsilon_\infty(T,0) \cdot \left[ 1 + S \cdot (b_{11} + b_{12} T) \right] \]  
(36)

\[ \varepsilon_\infty(T,0) = a_6 + a_7 T \]  
(37)

\[ \sigma(T,S) = \sigma(T,35) \cdot \frac{R_T(S)}{R_{15}(S)} \]  
(38)

\[ \sigma(T,35) = 2.903602 + 8.607 \times 10^{-2} T + 4.738817 \times 10^{-4} T^2 - 2.991 \times 10^{-6} T^3 + 4.3047 \times 10^{-9} T^4 \]  
(39)

\[ R_{15}(S) = S \cdot \frac{37.5109 + 5.45216 S + 1.4409 \times 10^{-2} S^2}{1004.75 + 182.283 S + S^2} \]  
(40)

\[ \frac{R_T(S)}{R_{15}(S)} = 1 + \frac{a_0 (T - 15)}{(a_1 + T)} \]  
(41)

\[ a_0 = \frac{6.9431 + 3.28415 - 9.9486 \times 10^{-2} S^2}{84.850 + 69.024 S + S^2} \]  
(42)

\[ a_1 = 49.843 - 0.2276 S + 0.198 \times 10^{-2} S^2 \]  
(43)

\[ \varepsilon_0 = \frac{1}{1.7975 \cdot 2\pi} \times 10^{-10} \]  
(44)

The coefficients of Equations (28)–(44) are shown in Table 5. The equations of other relative permittivity models will not be listed in this article. Finally, 18 sets of SSR brightness temperature increment data (including 9 sets of H-polarization data and 9 sets of V-polarization data) were generated by using the aforementioned nine relative permittivity models to process the data.
Table 5. Coefficients of Equations (28)–(44).

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>5.7230</td>
<td>$b_0$</td>
<td>$-3.56417 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_1$</td>
<td>$2.2379 \times 10^{-2}$</td>
<td>$b_1$</td>
<td>$4.74868 \times 10^{-6}$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$-7.1237 \times 10^{-4}$</td>
<td>$b_2$</td>
<td>$1.15574 \times 10^{-5}$</td>
</tr>
<tr>
<td>$a_3$</td>
<td>5.0478</td>
<td>$b_3$</td>
<td>$2.39357 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_4$</td>
<td>$-7.0315 \times 10^{-2}$</td>
<td>$b_4$</td>
<td>$-3.13530 \times 10^{-5}$</td>
</tr>
<tr>
<td>$a_5$</td>
<td>$6.0059 \times 10^{-4}$</td>
<td>$b_5$</td>
<td>$2.52477 \times 10^{-7}$</td>
</tr>
<tr>
<td>$a_6$</td>
<td>3.6143</td>
<td>$b_6$</td>
<td>$-6.28908 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_7$</td>
<td>$2.8841 \times 10^{-2}$</td>
<td>$b_7$</td>
<td>$1.76032 \times 10^{-4}$</td>
</tr>
<tr>
<td>$a_8$</td>
<td>$1.3652 \times 10^{-1}$</td>
<td>$b_8$</td>
<td>$-9.22144 \times 10^{-5}$</td>
</tr>
<tr>
<td>$a_9$</td>
<td>$1.4825 \times 10^{-3}$</td>
<td>$b_9$</td>
<td>$-1.99723 \times 10^{-2}$</td>
</tr>
<tr>
<td>$a_{10}$</td>
<td>$2.4166 \times 10^{-4}$</td>
<td>$b_{10}$</td>
<td>$1.81176 \times 10^{-4}$</td>
</tr>
<tr>
<td>$a_{11}$</td>
<td>$1.57883 \times 10^{-4}$</td>
<td>$b_{11}$</td>
<td>$-2.04265 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>$1.57883 \times 10^{-4}$</td>
<td>$b_{12}$</td>
<td>$1.57883 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

3. Deep Neural Network Model

The universal approximation theorem [49,50] shows that a feedforward neural network with a linear output layer and at least one hidden layer with any kind of “squeeze” property activation function can approximate any Borel measurable function from one finite dimensional space to another with any precision as long as enough neurons are given to its hidden layer. Therefore, the feedforward neural network was chosen in this paper.

The research ranges of $\theta$, $U_{10}$, and WD (wind direction) in 18 sets of SSR brightness temperature increment data were set to $[0^\circ, 60^\circ]$, $[0, 12 \text{ m/s}]$, and $[0, 360^\circ]$, respectively, and 18 datasets, which correspond to 9 H-polarization DNN models and 9 V-polarization DNN models, were uniformly and randomly selected from the research ranges, and each dataset contains 35,000 sets of data in the form of $[\theta, U_{10}, WD, \Delta T_{Bp, SSR}]$; then, each dataset was randomly divided into a training set and a testing set in an 8:2 ratio.

Before DNN training, training sets and test sets should be normalized. The data normalization process is to convert all data into data between 0 and 1, which is to eliminate the order of magnitude difference of data of various dimensions and avoid large network prediction errors caused by a large order of magnitude difference in input data. The data normalization method in this paper is the max–min method [51].

$$x_k = \frac{(x_k - x_{\min})}{(x_{\max} - x_{\min})}$$  \hspace{1cm} (45)

where $x_k$ is a set of input data, and $k$ is the serial number of a datum. $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of corresponding data.

The output of the DNN model needs to be renormalization to convert to the original range, which is the inverse process of normalization. The equation is as follows:

$$y_k = y_k \cdot (y_{\max} - y_{\min}) + y_{\min}$$  \hspace{1cm} (46)

where $y_k$ is a set of input data, and $k$ is the serial number of a datum. $y_{\max}$ and $y_{\min}$ are the maximum and minimum values of corresponding data of the training sets.

The training model in this paper is a six-layer DNN model, with the input layer having three neurons for $\theta$, $U_{10}$, and WD, the output layer having one neuron for $\Delta T_{Bp, SSR}$, and the hidden layer having 100 neurons; the random initial values were used as the DNN parameters, and the fundamental architecture of the DNN model is as follows (Figure 5):
The PReLU [52] was used as the activation function, as shown below (Figure 6):

\[
PReLU(x) = \begin{cases} 
  x, & \text{if } x > 0 \\
  \alpha x, & \text{if } x \leq 0
\end{cases}
\]  

(47)

The learnable parameter \( \alpha \), which is updated continually throughout the training phase, is the slope of the negative semi-axis of PReLu. Additionally, PReLU permits a group of neurons to share an \( \alpha \) as well as allowing various neurons to have a distinct \( \alpha \). When \( \alpha \) is a very small constant, PReLU can be thought of as LeakyReLU [53], and when \( \alpha = 0 \), PReLU becomes ReLU [54].

The MSEloss was used as the loss function of the output layer, as shown below [55]:

\[
\text{MSEloss} = (\Delta T_{Bp,\text{SSR},\text{out}} - \Delta T_{Bp,\text{SSR},\text{target}})^2
\]  

(48)
where $\Delta T_{Bp, SSR, \text{out}}$ denotes the output value of the output layer neuron and $\Delta T_{Bp, SSR, \text{target}}$ denotes its target value.

The Adam [56] and Backpropagation algorithms [57,58] were used as the training algorithms, as shown below:

$$u_k = u_{k-1} - \frac{\eta w_k}{\sqrt{v_k} + \gamma}$$  (49)

$$\begin{cases} W_k(h + 1) = W_k(h) - \eta \delta_k \cdot A^{k-1} \\ B_k(h + 1) = B_k(h) - \eta \delta_k \end{cases}$$  (50)

where $h$ is the number of iterations, $u$ is the parameter to be optimized, $\eta$ is the learning rate, $m_k$ is the first-order momentum, $v_k$ is the second-order momentum, $\gamma$ is a small positive number, $W_k$ is the weight matrix between layers $k - 1$ and $k$, $\delta_k$ is the error matrix of layer $k$, $A^k$ is the output matrix of layer $k$, and $B^k$ is the bias matrix of layer $k$.

The training process and forecast accuracy of the DNN models are shown in Table 6.

<table>
<thead>
<tr>
<th>Model Forecast Accuracy (K)</th>
<th>Learning Rate 0.001 (Iterate 8000 Times)</th>
<th>Learning Rate 0.003 (Iterate 30,000 Times)</th>
<th>Learning Rate 0.001 (Iterate 50,000 Times)</th>
<th>Learning Rate 0.0003 (Iterate 80,000 Times)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KS-H</td>
<td>-0.38~0.47</td>
<td>-0.50~0.24</td>
<td>-0.34~0.25</td>
<td>-0.14~0.29</td>
</tr>
<tr>
<td>KS-V</td>
<td>-0.29~0.35</td>
<td>-0.52~0.14</td>
<td>-0.39~0.15</td>
<td>-0.24~0.15</td>
</tr>
<tr>
<td>ModKS-H</td>
<td>-0.16~0.69</td>
<td>-0.66~0.12</td>
<td>-0.30~0.21</td>
<td>-0.23~0.12</td>
</tr>
<tr>
<td>ModKS-V</td>
<td>-0.26~0.35</td>
<td>-0.44~0.21</td>
<td>-0.23~0.15</td>
<td>-0.17~0.22</td>
</tr>
<tr>
<td>BA-H(03)</td>
<td>-0.39~0.47</td>
<td>-0.28~0.38</td>
<td>-0.26~0.12</td>
<td>-0.22~0.06</td>
</tr>
<tr>
<td>BA-V(03)</td>
<td>-0.32~0.68</td>
<td>-0.32~0.39</td>
<td>-0.31~0.26</td>
<td>-0.14~0.22</td>
</tr>
<tr>
<td>BA-H(04)</td>
<td>-0.38~0.43</td>
<td>-0.63~0.22</td>
<td>-0.37~0.30</td>
<td>-0.29~0.28</td>
</tr>
<tr>
<td>BA-V(04)</td>
<td>-0.23~0.76</td>
<td>-0.44~0.40</td>
<td>-0.14~0.38</td>
<td>-0.19~0.23</td>
</tr>
<tr>
<td>MW-H(04)</td>
<td>-0.35~0.40</td>
<td>-0.06~0.77</td>
<td>-0.63~0.03</td>
<td>-0.31~0.36</td>
</tr>
<tr>
<td>MW-V(04)</td>
<td>-0.45~0.75</td>
<td>-0.56~0.59</td>
<td>-0.19~0.31</td>
<td>-0.14~0.25</td>
</tr>
<tr>
<td>FASTEM-H</td>
<td>-0.42~0.48</td>
<td>-0.40~0.17</td>
<td>-0.31~0.11</td>
<td>-0.28~0.18</td>
</tr>
<tr>
<td>FASTEM-V</td>
<td>-0.20~0.73</td>
<td>-0.79~0.40</td>
<td>-0.51~0.35</td>
<td>-0.27~0.28</td>
</tr>
<tr>
<td>MW-H(12)</td>
<td>-0.83~0.45</td>
<td>-0.63~0.55</td>
<td>-0.41~0.62</td>
<td>-0.23~0.36</td>
</tr>
<tr>
<td>MW-V(12)</td>
<td>-0.41~0.48</td>
<td>-0.36~0.27</td>
<td>-0.27~0.21</td>
<td>-0.17~0.27</td>
</tr>
<tr>
<td>GW-H(17)</td>
<td>-0.49~0.77</td>
<td>-0.24~0.52</td>
<td>-0.27~0.37</td>
<td>-0.12~0.37</td>
</tr>
<tr>
<td>GW-V(17)</td>
<td>-0.55~0.52</td>
<td>-0.72~0.29</td>
<td>-0.18~0.23</td>
<td>-0.13~0.22</td>
</tr>
<tr>
<td>GW-H(21)</td>
<td>-0.49~0.46</td>
<td>-0.78~0.32</td>
<td>-0.64~0.31</td>
<td>-0.13~0.22</td>
</tr>
<tr>
<td>GW-V(21)</td>
<td>-0.23~0.41</td>
<td>-0.59~0.22</td>
<td>-0.22~0.14</td>
<td>-0.24~0.13</td>
</tr>
</tbody>
</table>

The forecast accuracy of the DNN models can be further improved by further lowering the learning rate and continuing training, and the final performance of the DNN models is shown in Table 7.

Where RMSE and MAE are the Root Mean Square Error and the Mean Absolute Error [59,60], and their calculation equations are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$$   (51)

$$MAE = \frac{\sum_{i=1}^{n} |x_i|}{n} = \frac{|x_1| + |x_2| + \cdots + |x_n|}{n}$$   (52)

This article chooses SSR brightness temperature increment models like Hollinger [9], WISE (2003) [10], WISE (2004) (containing two models calculated with the incident angle and wind speed, which are called WISE1 and WISE2, respectively, according to the order in reference [11]), TSM [16], and SSA [17] for research based on the data characteristics of offshore platform experiments, and 108 forward models (including 54 H-polarization
models and 54 V-polarization models) can be obtained by combining each SSR brightness temperature increment model with the aforementioned 9 relative permittivity models. The test sets were substituted into the 108 forward models, and the accuracy is shown in Table 8.

Table 7. Final performance of the DNN models.

<table>
<thead>
<tr>
<th>DNN Model</th>
<th>Forecast Accuracy (K)</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>KS-H</td>
<td>-0.13~0.14</td>
<td>0.0157</td>
<td>0.0103</td>
</tr>
<tr>
<td>KS-V</td>
<td>-0.11~0.13</td>
<td>0.0156</td>
<td>0.0090</td>
</tr>
<tr>
<td>ModKS-H</td>
<td>-0.13~0.13</td>
<td>0.0161</td>
<td>0.0101</td>
</tr>
<tr>
<td>ModKS-V</td>
<td>-0.13~0.16</td>
<td>0.0185</td>
<td>0.0114</td>
</tr>
<tr>
<td>BA-H(03)</td>
<td>-0.14~0.16</td>
<td>0.0177</td>
<td>0.0108</td>
</tr>
<tr>
<td>BA-V(03)</td>
<td>-0.15~0.16</td>
<td>0.0191</td>
<td>0.0118</td>
</tr>
<tr>
<td>BA-H(04)</td>
<td>-0.17~0.13</td>
<td>0.0183</td>
<td>0.0121</td>
</tr>
<tr>
<td>BA-V(04)</td>
<td>-0.15~0.15</td>
<td>0.0173</td>
<td>0.0109</td>
</tr>
<tr>
<td>MW-H(04)</td>
<td>-0.15~0.17</td>
<td>0.0182</td>
<td>0.0124</td>
</tr>
<tr>
<td>MW-V(04)</td>
<td>-0.15~0.13</td>
<td>0.0168</td>
<td>0.0106</td>
</tr>
<tr>
<td>FASTEM-H</td>
<td>-0.14~0.15</td>
<td>0.0169</td>
<td>0.0106</td>
</tr>
<tr>
<td>FASTEM-V</td>
<td>-0.14~0.17</td>
<td>0.0176</td>
<td>0.0102</td>
</tr>
<tr>
<td>MW-H(12)</td>
<td>-0.17~0.16</td>
<td>0.0189</td>
<td>0.0113</td>
</tr>
<tr>
<td>MW-V(12)</td>
<td>-0.17~0.16</td>
<td>0.0190</td>
<td>0.0110</td>
</tr>
<tr>
<td>GW-H(17)</td>
<td>-0.15~0.15</td>
<td>0.0177</td>
<td>0.0107</td>
</tr>
<tr>
<td>GW-V(17)</td>
<td>-0.14~0.17</td>
<td>0.0166</td>
<td>0.0095</td>
</tr>
<tr>
<td>GW-H(21)</td>
<td>-0.13~0.17</td>
<td>0.0170</td>
<td>0.0100</td>
</tr>
<tr>
<td>GW-V(21)</td>
<td>-0.17~0.13</td>
<td>0.0184</td>
<td>0.0105</td>
</tr>
</tbody>
</table>

Table 8. Performance of the forward models on the test sets.

<table>
<thead>
<tr>
<th>Model</th>
<th>Forecast Accuracy (K)</th>
<th>Hollinger</th>
<th>WISE</th>
<th>WISE1</th>
<th>WISE2</th>
<th>TSM</th>
<th>SSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>KS-H</td>
<td>-1.57~1.49</td>
<td>-1.55~1.71</td>
<td>-1.51~1.69</td>
<td>-1.65~1.54</td>
<td>-0.47~0.61</td>
<td>-0.60~0.81</td>
<td></td>
</tr>
<tr>
<td>KS-V</td>
<td>-0.79~1.29</td>
<td>-0.79~1.11</td>
<td>-0.79~1.18</td>
<td>-0.81~0.98</td>
<td>-0.35~0.55</td>
<td>-0.48~0.69</td>
<td></td>
</tr>
<tr>
<td>ModKS-H</td>
<td>-2.38~0.59</td>
<td>-2.35~0.80</td>
<td>-2.32~0.79</td>
<td>-2.46~0.63</td>
<td>-1.28~0.26</td>
<td>-1.41~0.05</td>
<td></td>
</tr>
<tr>
<td>ModKS-V</td>
<td>-1.80~0.11</td>
<td>-1.80~0.07</td>
<td>-1.80~0.01</td>
<td>-1.82~0.20</td>
<td>-1.33~0.58</td>
<td>-1.47~0.49</td>
<td></td>
</tr>
<tr>
<td>BA(03)-H</td>
<td>-2.04~0.59</td>
<td>-2.02~0.80</td>
<td>-1.98~0.79</td>
<td>-2.12~0.63</td>
<td>-0.94~0.50</td>
<td>-1.07~0.32</td>
<td></td>
</tr>
<tr>
<td>BA(03)-V</td>
<td>-4.03~1.14</td>
<td>-4.20~1.33</td>
<td>-4.12~1.26</td>
<td>-4.36~1.45</td>
<td>-3.89~1.94</td>
<td>-3.51~1.75</td>
<td></td>
</tr>
<tr>
<td>BA(04)-H</td>
<td>-0.31~2.55</td>
<td>-0.29~2.77</td>
<td>-0.25~2.76</td>
<td>-0.39~2.60</td>
<td>0.79~1.46</td>
<td>0.66~1.65</td>
<td></td>
</tr>
<tr>
<td>BA(04)-V</td>
<td>-1.37~1.41</td>
<td>-1.53~1.23</td>
<td>-1.47~1.29</td>
<td>-1.69~1.10</td>
<td>-1.23~0.61</td>
<td>-0.85~0.81</td>
<td></td>
</tr>
<tr>
<td>MW(04)-H</td>
<td>-1.25~1.48</td>
<td>-1.23~1.70</td>
<td>-1.19~1.69</td>
<td>-1.33~1.53</td>
<td>-0.15~0.41</td>
<td>-0.28~0.59</td>
<td></td>
</tr>
<tr>
<td>MW(04)-V</td>
<td>-2.82~0.02</td>
<td>-2.98~0.20</td>
<td>-2.91~0.13</td>
<td>-3.14~0.33</td>
<td>-2.67~0.81</td>
<td>-2.30~0.62</td>
<td></td>
</tr>
<tr>
<td>FASTEM-H</td>
<td>-1.34~1.38</td>
<td>-1.32~1.60</td>
<td>-1.28~1.59</td>
<td>-1.42~1.43</td>
<td>-0.24~0.30</td>
<td>-0.37~0.48</td>
<td></td>
</tr>
<tr>
<td>FASTEM-V</td>
<td>-2.98~0.17</td>
<td>-3.15~0.35</td>
<td>-3.07~0.28</td>
<td>-3.30~0.48</td>
<td>-2.84~0.97</td>
<td>-2.46~0.77</td>
<td></td>
</tr>
<tr>
<td>MW(12)-H</td>
<td>-1.25~1.48</td>
<td>-1.23~1.70</td>
<td>-1.19~1.69</td>
<td>-1.33~1.53</td>
<td>-0.15~0.41</td>
<td>-0.28~0.59</td>
<td></td>
</tr>
<tr>
<td>MW(12)-V</td>
<td>-2.82~0.02</td>
<td>-2.98~0.20</td>
<td>-2.91~0.13</td>
<td>-3.14~0.33</td>
<td>-2.67~0.81</td>
<td>-2.30~0.62</td>
<td></td>
</tr>
<tr>
<td>GW(17)-H</td>
<td>-1.55~1.50</td>
<td>-1.52~1.72</td>
<td>-1.49~1.70</td>
<td>-1.63~1.55</td>
<td>-0.45~0.62</td>
<td>-0.58~0.81</td>
<td></td>
</tr>
<tr>
<td>GW(17)-V</td>
<td>-0.82~1.59</td>
<td>-0.82~1.41</td>
<td>-0.82~1.48</td>
<td>-0.84~1.28</td>
<td>-0.34~0.79</td>
<td>-0.51~0.99</td>
<td></td>
</tr>
<tr>
<td>GW(21)-H</td>
<td>-1.58~1.48</td>
<td>-1.56~1.69</td>
<td>-1.52~1.68</td>
<td>-1.66~1.52</td>
<td>-0.48~0.60</td>
<td>-0.61~0.80</td>
<td></td>
</tr>
<tr>
<td>GW(21)-V</td>
<td>-0.80~1.28</td>
<td>-0.80~1.09</td>
<td>-0.80~1.16</td>
<td>-0.82~0.97</td>
<td>-0.36~0.53</td>
<td>-0.49~0.67</td>
<td></td>
</tr>
</tbody>
</table>

The RMSE and MAE of the 108 forward models and the 18 DNN models on the test sets are shown in Figures 7–10.

It can be seen from Tables 7 and 8, and Figures 7–10, that the DNN models perform significantly better than the existing models, and that their calculation accuracy is close to the detection accuracy of the radiometer (0.1 K).

Furthermore, this study processed offshore experiment data using $T_{R,COS} = 2.73$ K and employed the same methodology for its research and analysis, and the findings demonstrate that although the 18 DNN models can also be trained to the level of Table 7, the errors of the 108 forward models in Table 8 will increase. As a result, the experimental data in this
article were processed with $T_{B,\text{COS}} = 3.7$ K (the processed results of $T_{B,\text{COS}} = 2.73$ K will not be listed in this article).

Figure 7. The RMSE of the H-polarization models.

Figure 8. The RMSE of the V-polarization models.
Figure 9. The MAE of the H-polarization models.

Figure 10. The MAE of the V-polarization models.

4. Discussion

The forecast accuracy of the DNN models should be set above 0.1 K because the detection accuracy of the radiometer utilized in the offshore observation experiment is 0.1 K. Excessive training (such as setting the forecast accuracy within 0.1 K) will make the DNN models learn too much error from the radiometer itself, which will lower the reliability of the prediction results of the DNN models. As a result, the threshold for ending
training regarding the DNN models in this paper is set to 0.1~0.2 K, and Table 7 shows that all DNN models met the standards.

The reason for the low forecast accuracy of Hollinger, WISE, WISE1, and WISE2 is that they do not consider the influence of wind direction. Although the forecast accuracy of TSM and SSA is relatively high, they are unable to match the detection accuracy of the radiometer (0.1 K), which may be due to the error of the ocean wave spectrum of the offshore area used by the two algorithms.

The relative permittivity model also has an impact on the forward model. It can be seen from Table 8 that KS (1977), MW (2004), FASTEM (2011), MW (2012), GW (2017), and GW (2021) have good performances for the H-polarization forward models; KS (1977), GW (2017), and GW (2021) have good performances for the V-polarization forward models; and KS (1977) and GW (2021) perform similarly and well on the two polarization forward models.

5. Conclusions

The calculation results of different forward models, which are composed of different relative permittivity models and SSR brightness temperature increment models, are different, and the impact of this calculation difference exceeds the accuracy requirement of the SSS inversion. The existing SSR brightness temperature increment models, which primarily include empirical models and theoretical models, cannot match all the relative permittivity models, and can only obtain good results under the specific relative permittivity model, while the errors under other relative permittivity models are large. Therefore, different SSR brightness temperature increment models should be provided for different relative permittivity models to make the calculation results of the forward model as close to the actual sea surface situation as possible. In order to address this problem, this paper proposes a universal DNN model architecture and corresponding training scheme, and provides 18 different SSR brightness temperature increment models (including 9 H-polarization models and 9 V-polarization models) for 9 widely used L-band relative permittivity models utilizing DNN based on the offshore experiment data, and compares them with the existing models. The results show that the DNN models perform significantly better than the existing models, and that their forecast accuracy is close to the detection accuracy of the radiometer (0.1 K). Therefore, this study effectively solves the problem of SSR brightness temperature correction under different relative permittivity models, and provides a theoretical support for high-precision SSS inversion research.

Author Contributions: Conceptualization, Z.W., S.L. and Q.Z.; methodology, Z.W.; software, Z.W. and L.L.; validation, Z.W., H.Z., W.S., L.Z. and J.R.; formal analysis, Z.W., X.L. and Y.Z.; investigation, Z.W.; resources, Z.W. and H.Z.; data curation, Z.W.; writing—original draft preparation, Z.W.; writing—review and editing, Z.W.; visualization, Z.W.; supervision, Z.W.; project administration, Z.W.; funding acquisition, Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number: 62031005.

Data Availability Statement: The data are unavailable due to privacy.

Acknowledgments: The offshore observation platform of this article was provided by the National Satellite Ocean Application Service of China, and the L-band microwave radiometer was provided by the China Academy of Space Technology, Xi’an.

Conflicts of Interest: The authors declare no conflict of interest.

References
45. Zhou, Y.; Lang, R.H.; Dinnat, E.P.; Le Vine, D.M. Seawater Debye Model Function at L-Band and Its Impact on Salinity Retrieval


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.