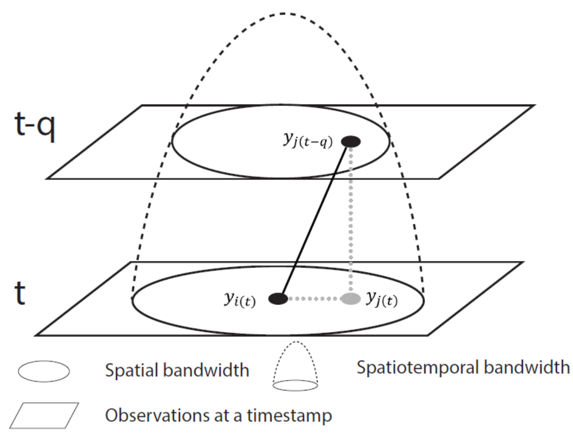
Thank you for your comments. Here are our responses:

1.For the possible major problem:

In our current STWR algorithm, as seen in Equation (4), we use the $y_{i(t)} - y_{j(t-q)}$ (the difference between the regression point i at time t and the observed point j at time $t-q$) rather than the $\Delta y_{j(t-q)}$ (value variation of the observed point j in Δt). The main reason we use $y_{i(t)}$ instead of $y_{j(t)}$ to reflect the rate of change of y_j during the time interval (from $t-q$ to t), is that the y value of the location j at t is often unavailable or may not exist at all, while the y value of the regression point i at t is known (i.e. $y_{i(t)}$).Within the local spatiotemporal bandwidth, the value of $y_{i(t)}$ is close to $y_{j(t)}$ because both values tend to be homogeneous. As shown in the following figure, the dotted line from $y_{j(t-q)}$ to $y_{j(t)}$ can be approximated by the solid line from $y_{j(t-q)}$ to $y_{i(t)}$ within the local spatiotemporal bandwidth. When the observation point j is outside the local spatiotemporal bandwidth, there will be no such approximation. Although the value $y_{i(t)}$ is not actual $y_{j(t)}$, this substitution is also valid. The reason is that both formulations can reflect the consistent temporal effect of the past observation point j on the regression point i at time t. In our STWR algorithm, we need to measure the degree of influence of the observed points at $t-q$ (i.e. $y_{j(t-q)}$) on the regression point i at t (i.e. $y_{i(t)}$). The value of the difference between $y_{i(t)}$ and $y_{j(t-q)}$ divided by $y_{j(t-q)}$, which represents the numerical difference rate, can reflect the degree of temporal influence of the past observation point j ($y_{j(t-q)}$) on the current regression point i ($y_{i(t)}$). Besides, we also have some ideas and suggestions about using $\Delta y_{j(t-q)}$ in Equation (4), which is discussed in Section 6, (page 28 and 29).



2. For the minor problems:

We use three simulation cases and a real-world case for the reasons listing below:

(1) It can verify that this new method can be applied to different situations and is more robust than GTWR. In case 1, two independent variables x_1 and x_2 only changed slightly over time, and the observed time interval is short. In case 2, the x_1 and x_2 changed faster over time, and their observed time interval gets longer. These two cases verify that the performance of GTWR is unstable, which is sometimes better than GWR (case 1), and sometimes worse than GWR (case 2). The model performance of STWR is the best, in both case 1 and case 2, indicating that STWR is more robust than GTWR.

(2) Both case 1 and 2 assumes that three coefficient surfaces keep the same over time, but in case 3, the coefficient surfaces is assumed to vary over time. Results of the case 3 show our new algorithm STWR still outperforms GWR and GTWR models when the coefficient surfaces change over time.

(3) Through the three simulation case studies, we can draw that when the observed data changes faster over time, the outperformance of the STWR model will be more prominent than GWR and GTWR.

(4) Through the real-world case, we verified the effectiveness of our new algorithm STWR, making it more convincing.

We will add the name of the journal in line79, page 31. And we will reduce the decimal number of the AICc of GTWR in Table 2 to keep three decimal places (because some R-squares are close, keeping three digits is more convenient for comparison). Also, we will add more clear explanations and descriptions on the R-square in tables because there are many R-squares for each regression point in GWR, GTWR, and STWR. For the significant difference in the R-Square values, we will add some contents to facilitate the reader's understanding.

Thanks again for your comments.