In the article, authors have described a technique to estimate the wind velocity from thermal images. The estimated wind velocity is comparable with the wind speed measurement qualitatively. But quantitatively, e.g. $R^2$ analysis, p-value) the agreement between them is poor. Also many details of the analysis are missing. Thus, I won’t recommend the paper for publication in current form. Following are the itemized suggestions:

1) How authors calibrate each pixel of the thermal image to physical distance, m? The pixel sizes vary based on the camera viewing angle. Pixels closer to the camera have smaller physical sizes compared to the further away. How authors have corrected this camera viewing distortion of the thermal image? What are the image sizes in physical distance?

2) How did authors perform the spatial average of the 15m x 15m patch? From figure 5, the velocities are estimated when temperature gradient is present. Does the patch size ensure no missing wind velocity estimate for the pixels considered? Why specific size of patch is chosen?

3) For comparison with the sonic anemometer measurement, how the location of 15m x 15m patch is chosen? Does it lie inside the footprint of sonic anemometer? If not, why authors have not considered some flux footprint model, as described in Garai et al. 2013, Boundary Layer Meteorology?

4) What are the weights for averaging wind velocities from different perturbation? How the weights are chosen?
5) Figure 7, how the difference is calculated? Not all the pixels result velocity estimation.

6) Lower value of p-estimation means that the assumed null hypothesis does not hold. For present study, what is the null hypothesis, is it TIV corresponds to wind velocity? If so, then the reported small p-value means there is no correlation.

7) Figure 9a. Too many wiggly lines make it difficult to read. Authors should consider separate out the temperature and velocity curves in two separate figures.

8) Figure 9 shows that the TIV and sonic anemometer have some temporal lag. It also looks like the lag and wind direction are not constant for the time period considered. How the authors account for these effects when calculating $R^2$, p-value and histograms for quantitative comparison? Also how these variable effects the comparison.

9) Figure 10, What are the markers in the figure?

10) Page 17, line 261: What do you mean by positive wind speed? Wind speed is always positive.

11) For Turf-T2 why authors have not considered to have two thermocouple arrays on the two surfaces, instead of putting one thermocouple array in the mixed surface. As the surface properties are changing, a new boundary layer will start to develop. How authors account for that in the analysis.

12) Garai and Kleissl 2013, Journal of Turbulence, reported that the different temporal filtering result thermal structures corresponding different scale. How authors account for that when comparing averaged TIV from different temporal perturbation with sonic anemometer? The small flow structure giving the TIV 5s perturbation may not be registerd at the sonic anemometer.

13) English in the article is poor.