ISSA improves SSAM by reformulating the calculation of PCs (equation 7) to incorporate RCs for missing values (equations 8 to 14). The improvement is small for mostly complete time series and increases as the quantity of missing data increases. I encourage the authors to post ISSA code for others to use.

It appears that ISSA Eigenvectors v are calculated as they are in SSAM from the Toeplitz matrix formed from equation 5. This ISSA step should be added to the manuscript. The eigenvectors are then used to create matrix G. It appears that matrix G must be created and equation 14 solved for each time step i. This is a large increase in computational effort compared to SSAM, which should be stated in the manuscript.

In equation 11, the sums are for all terms in the window with a missing value. The values of the eigenvector do not change with time, so the sum can be replaced with N_m, the number of missing values in the window (e.g. $\Sigma v_{1,j}v_{2,j} = N_m\, v_{1,j}v_{2,j}$). If N_m=0, equation 10 reduces to equation 3.

p. 1953, lines 8-13: Equation 15 is used to compare SSAM and ISSA which is good to include but the approach contains a contradiction that should be explained. To compare their results to SSAM, the authors set nondiagonal elements in equation 11 to zero but also assume $v_{k,i}=L^{-1/2}$, in which case the diagonal elements would equal N_m/L where N_m is the number of missing data points in the window. The authors should explain this contradiction. For the case where $N_m/L << 1$, this contradiction would be minor. Is this contradiction inherently assumed in the formulation of SSAM, and if so, does it explain the relatively improving performance of ISSA as N_m/L (% missing data in table 1) increases? SSAM performance declines when $N_m/L > 0.5$ which is roughly when the diagonal elements of equation 11 become less than the non-diagonal elements – could this be the cause?. Or does the ISSA assumption that missing values can be represented by an RC expression create this contradiction? Missing values are ignored when calculating the eigenvectors in both methods, but ISSA does not ignore missing values when calculating PCs.

Specific comments:

<u>p. 1948</u>:

abstract: Add that the improvement is small for mostly complete time series and increases as the quantity of missing data increases. Because of this, I suggest changing 'much smaller' to 'smaller'.

line 16: define SD

line 17: A difference of 1.2 mg/L (~10%) is within typical measurement error.

lines 25-26: use 'wide' only once in the sentence.

p 1949, line 9: define GNSS

p. 1951, line 14: insert paragraph break where SSAM ends and ISSA starts.

p. 1953, line 8: insert paragraph break where ISSA ends and comparison to SSAM begins (the sentence that introduces eqn. 15).

<u>p. 1954</u>:

lines 2-3: This section is about the synthetic time series, not the real time series, so delete this sentence.

line 20: Delete 'even'.

Equation 18: define T (transpose?)

p. 1955, line 8: delete ‘clear’

p. 1956, line 19: the mean residual in not presented in table 2.

p. 1956, line 22: the difference of r2 of 0.9178 and 0.9046 seems to be minor – is this statistically significant? Autocorrelation would probably have to be considered.

Table 1: Delete ‘As’ in last row, replace with ‘SF’?

p. 1957, lines 7-8: Change ‘With the missing data gets more, the improvements of the relative errors becomes more evident.’ to ‘As the fraction of missing data increases, the improvement of the relative error becomes greater’.

p. 1957, line 12: The SSC improvements are minor and within measurement error.